THE ROLE OF TRIPLEX DNA IN THE CELL

A Thesis Submitted to the College of
Graduate Studies and Research in
Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Biochemistry
University of Saskatchewan
Saskatoon

by

Carolyn Ashley

Fall, 1999

UNIVERSITY OF SASKATCHEWAN

College of Graduate Studies and Research

SUMMARY OF DISSERTATION

Submitted in partial fulfillment

of the requirements for the

DEGREE OF DOCTOR OF PHILOSOPHY

By

Carolyn Ashley

Department of Biochemistry
University of Saskatchewan

Fall 1999

Examining Committee:

Dr. B. Juurlink	Dean/Associate Dean/Dean's Designate, Chair, College of Graduate Studies and Research
Dr. L.T.J. Delbaere	Chair, Department of Biochemistry
Dr. R.L. Khandelwal	Graduate Chair, Department of Biochemistry
Dr. J.S. Lee	Supervisor, Department of Biochemistry
Dr. K. Bonham	Department of Biochemistry
Dr. G.H. Rank	Department of Biology
Dr. W.J. Roesler	Department of Biochemistry
Dr. R.C. Warrington	Department of Biochemistry

External Examiner:

Dr. A.R. Morgan
Department of Biochemistry
University of Alberta
Edmonton, Alberta
T6G 2H7

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any materials in my thesis.

Requests for permission to copy or make other use of material in this thesis in whole or in part should be addressed to:

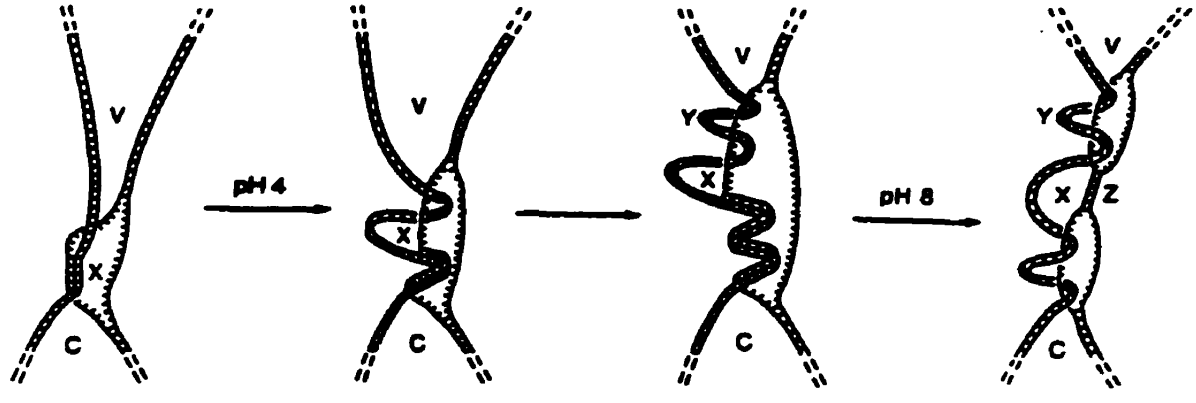
Head of the Department of Biochemistry
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

ABSTRACT

Polypurine•polypyrimidine (pur•pyr) tracts are a run of all purines on one strand and all pyrimidines on the complementary DNA strand. Statistical overrepresentation of the tracts in eukaryotes suggests a cellular role or roles. The tracts form triplex DNA *in vitro* and there is evidence for triplex DNA *in vivo*.

Several cellular roles are possible for triplex DNA. The presence of the tracts in gene 5' flanking regions suggests a regulatory role. This work investigates the role of triplex DNA in the cell, particularly in the regulation of transcription.

Proteins mediate DNA looping in the regulation of transcription and in its condensation in chromosomes. Such looping may also be mediated by transmolecular triplexes, formed between separated pur•pyr tracts. Formation of pyr•pur•pyr transmolecular triplexes was investigated using linear and circular plasmid models containing separated pur•pyr tracts able to form a triplex with each other, but not within a tract. Transmolecular triplex loops (T-loops) formed in circular DNA, suggesting a possible regulatory or structural role *in vivo*. The following model shows a T-loop formed at pH 4. At pH 8, a duplex partially reforms and single-stranded region(s) trap the structure.



T-loops were used as a model to test the idea that a single-strand extruded by triplex formation in the 5' flanking region of a gene could

promote transcription. Transcription was inhibited in T-loops, suggesting such structures could block transcriptional elongation if formed *in vivo*.

The ability of polyamine analogues to promote triplex formation was also tested using T-loops. Pentamines promoted T-loop formation at lower concentrations than tetramines. Spatial distribution of charge was also important.

A triplex role in transcriptional regulation was investigated using two examples of human genes with 5′ flanking pur•pyr tracts. The effect of triplex-specific antibodies on expression of c-*myc* was investigated using agarose-encapsulated nuclei. Triplex formation between c-*src* promoter pur•pyr tracts was visualized as gel band shifts due to dimerization between linear plasmid fragments containing individual tracts. A transmolecular triplex was proposed as one way in which the c-*src* tracts could form a triplex *in vivo* which might be involved in the regulation of transcription.

DEDICATION

To my parents, for their strong encouragement of my academic endeavors over the years, and especially to my father, who had already established a family tradition of obtaining graduate degrees later than usual in life!

ACKNOWLEDGMENTS

First, and most importantly, I would like to thank my supervisor, Dr. Jeremy Lee, for his patience, knowledge, and guidance during my sojorn as a Ph.D student. I am very grateful for the opportunity to learn so much in his laboratory. Financial support throughout my studies was provided by the College of Medicine Graduate Scholarship Fund and a Medical Research College of Canada grant to Dr. Lee.

I would like to thank the members of my committee for their guidance and very thorough reading of my thesis. Additionally, I appreciate the many informative discussions I had with Dr. Bonham on the promoter region of human c-*Src* regarding its potential for triplex formation and factors acting in *trans*. I would also like to thank him for instruction on run-off transcription procedures using ^{32}P. I am grateful to Dr. Bill Roesler for discussions on aspects of transcription.

Dr. Vikram Misra of the Department of Veterinary Microbiology, Western College of Veterinary Medicine generously supplied blocks of time out of his busy day to advise on procedures, loan supplies and equipment, and provide moral support. I also appreciate the help received from Drs. Rai Lu and Ping Yang of his laboratory. I thank the College of Veterinary Medicine for use of its gel documentation system.

Thanks are also extended to Dr. Selvaraj of the Plant Biotechnology Institute, Saskatoon, for the use of the PhosphorImager in his laboratory, and to the Plant Biotechnology Institute for use of equipment on its premises.

I would like to thank the staff in the Biochemistry Department Student Laboratory for their tolerance of my frequent incursions into their territory. Indeed, I am appreciative of their considerable help and friendship. In particular, I would like to thank Bill Wilcox for never failing to get any faltering piece of equipment up and running again!

I thank all the other staff of the department, and of course, my fellow graduate students, for their friendship and help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A,G,C,T,U	adenine, guanine, cytosine, thymine, uracil
A_{260}	absorbance at a wavelength of 260 nm
agarose-encapsulated nuclei	agarose-encapsulated and permeabilized nuclei
ATP	adenosine 5'-triphosphate
bp	base pair(s)
BSA	bovine serum albumin
CAT	chloramphenicol acetyl transferase
Ci	Curie(s)
CMF-PBS	calcium- and magnesium-free phosphate-buffered saline
CPM	counts per minute
CTP	cytidine 5'-triphosphate
d	deoxyribo prefix
ddH_20	distilled, deionized water
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
E. coli SSB protein	*Escherichia coli* single-strand binding protein
EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
EtBr	ethidium bromide
γ	gamma
GTP	guanosine 5'-triphosphate
h	hour(s)
Inr	initiator element
IPTG	isopropyl β-D-thiogalactoside
kb	kilobase(s)
kD	kilodaltons
LB	Luria-Bertani
Lk	linking number
MAR	matrix-attached region
MGMT	O^6-methylguanine-DNA methyltransferase

min	minute(s)
mRNA	messenger RNA
NSE	nuclease sensitive element
pK	negative logarithm of dissociation constant of an acid
Pol II	RNA polymerase II
pur•pyr	polypurine•polypyrimidine
r	ribo prefix
R.T.	room temperature (23 °C)
RNA	ribonucleic acid
RNase	ribonuclease
RPM	revolutions per minute
rRNA	ribosomal RNA
σ	sigma
SAR	scaffold-associated element
SDS	sodium dodecyl sulphate
sec	second(s)
SPy	*src* pyrimidine-binding element
T-loops	transmolecular triplex loops
T/E	Tris/EDTA
TAF	TBP-associated factor
TBP	TATA-binding protein
TCA	trichloroacetic acid
TFO	triplex-forming oligonucleotide
tRNA	transfer RNA
TTP	thymidine 5'-triphosphate
UTP	uridine 5'-triphosphate
X-Gal	5-bromo-4-chloro-3-indolyl β-D-galactoside
Ω	omega

1.0 INTRODUCTION

1.1 Nucleic Acid Structure

Triplex DNA is made up of a DNA duplex plus a single DNA strand. Therefore, a good way to introduce the topic of the role of triplex DNA in the cell is to review the basic structure of nucleic acids. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are nucleotide polymers in which the 3'-hydroxyl of one nucleotide is joined via a phosphodiester linkage to the 5' hydroxyl of the next. By convention, nucleic acid sequences are written in a 5' to 3' direction. A nucleotide consists of a nitrogenous base, a pentose sugar, and a phosphate group. The bases in nucleic acids are derivatives of either purine or pyrimidine. Common purine bases are adenine (A) and guanine (G), and pyrimidine bases are cytosine (C) and thymine (T) in DNA or uracil (U) in RNA (Figure 1.1). The structure of uracil is the same as thymine except that in uracil the C5 carbon is unmethylated. The bases are attached at N1 (pyrimidines) or N9 (purines) via an N-glycosidic bond to C'1 of the pentose sugar. There are two positions of the base relative to the sugar (Figure 1.2). The *anti* conformation is favored energetically in most polynucleotides. The pentose sugar is not a planar molecule and there are 4 major nonplanar conformations of the 2' and 3' carbons relative to the other atoms in the ring (Figure 1.3).

DNA forms a double helix consisting of the isomorphous base pairs (bp), A•T and G•C (Figure 1.1). The term "isomorphous" means that both bp have the same C1'-C1' distance and N9-C1'-C1' and N1-C1'-C1' angles. The duplex is stabilized by stacking interactions between successive bp and also by hydrogen bonding between complementary bp, two for A•T and three for G•C. The paired DNA strands are antiparallel with the 5' to 3' phosphodiester linkages running in opposite directions. As the pK of the phosphates in the phosphodiester linkages is close to 1, the phosphate backbone of DNA is highly negatively charged at physiological pH. This results in a repulsion between polynucleotide molecules. This is the model proposed by Watson and Crick in 1953 to account for the observed X-ray diffraction pattern and the A=T and G=C

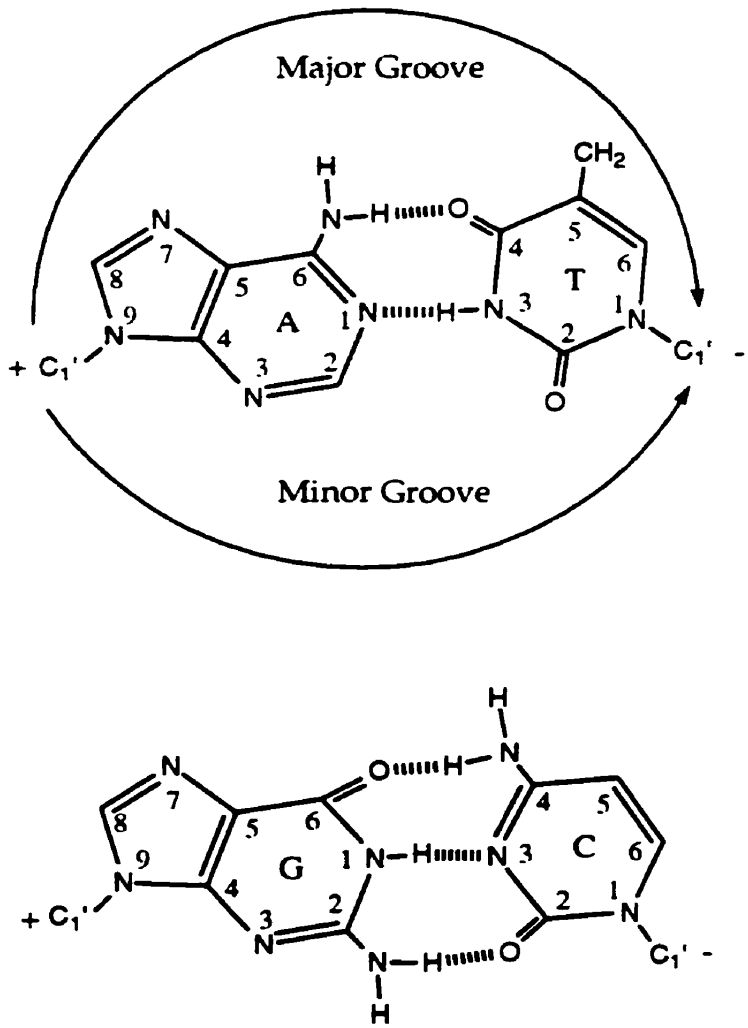


Figure 1.1. The structures of the complementary base pairs of duplex DNA. The atoms are numbered as shown. Positions of hydrogen bonds and major and minor grooves are shown. The bases are in *anti* conformation relative to the sugars, and are uncharged at physiological pH.

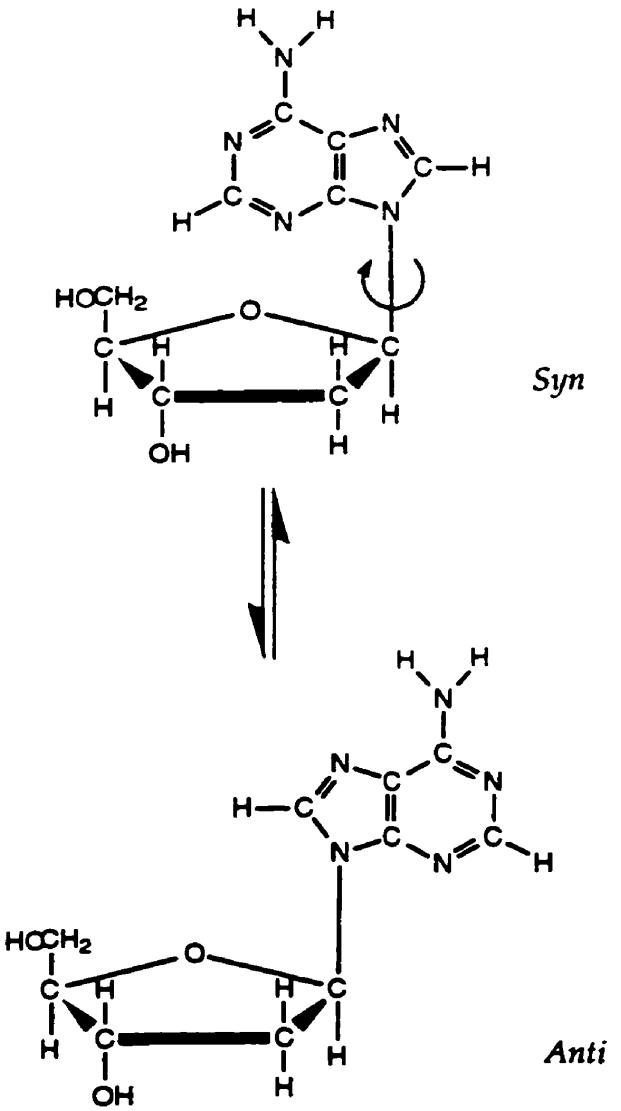
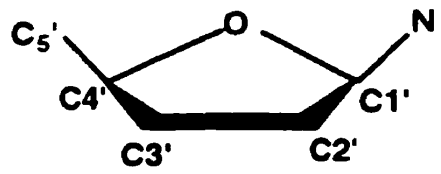
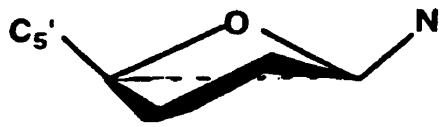


Figure 1.2. *Syn* and *anti* conformations of the bases relative to the sugar ring in deoxyadenosine.

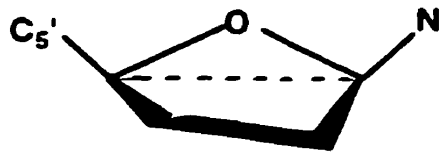


Planar - never seen

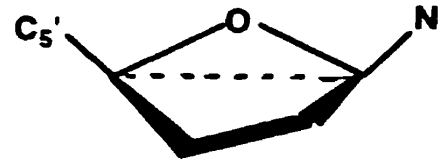


C2'-endo

C3'-endo



C2'-exo



C3'-exo

Figure 1.3. Non-planar conformations of ribose and deoxyribose.

base equivalencies of DNA. It is now known as B form DNA and is thought to be the most common form in the cell.

Because the topic of this thesis is triplex DNA, the potential of certain base nitrogens to gain or lose protons as the pH is lowered or raised is relevant. Protons determine the ability of nitrogen atoms to form hydrogen bonds with atoms in the bases of other DNA strands, thus stabilizing multiple-stranded DNA structures. Of particular importance is the potential of cytosine to become protonated as the pH is lowered. The pK of position N3 of cytosine as a free base is 4.3 (Inman, 1964a,b; Figure 1.4). However, in polycytidylic acid the pK of cytosine increases to 7.3 to 7.5 (Gray *et al.*, 1987). The increase in the pK when the base is part of a polynucleotide is important in a consideration of whether a particular type of triplex would be able to form under physiogical conditions. This issue will be discussed in detail in this thesis.

The parameters of a number of different DNA forms have been determined by X-ray analysis of DNA fibers (reviewed in Soyfer and Potaman, 1996). In B-DNA the strands wind around each other in a right-handed helix, there are 10 bp per helical turn, there is a rise of 3.4 Å per bp, the orientation of bases to sugar is *anti*, and the sugar conformation is primarily C2'-endo (Figure 1.3). The close association of other duplex molecules in DNA fibers apparently influenced these parameters slightly; using a technique involving enzymatic cutting from one side of the molecule, the number of bp per helical turn of B-DNA in solution was later adjusted to 10.5 (Rhodes and Klug, 1980).

A form DNA is another type of helix, favored in relatively dehydrated solutions. A-DNA also has a right-handed helix. However, the helix is underwound (see below) compared to B-DNA. There are 11 bp per turn, a rise of 2.9 Å per bp and although the base-sugar orientation is *anti*, the sugar conformation is C3'-endo. When a DNA molecule adopts the A form it is therefore shorter and wider than B-DNA.

Z-DNA can form in alternating purine-pyrimidine sequences such poly[d(GC)] and poly[d(TG)]•poly[d(CA)]. In contrast to the other forms of DNA, the helix is left-handed. Glycosidic bonds alternate in orientation, being *anti* for pyrimidines and *syn* for purines, and the sugar conformations also alternate between C3'-endo and C2'-endo. The sugar-phosphate backbone has a zig-zag shape.

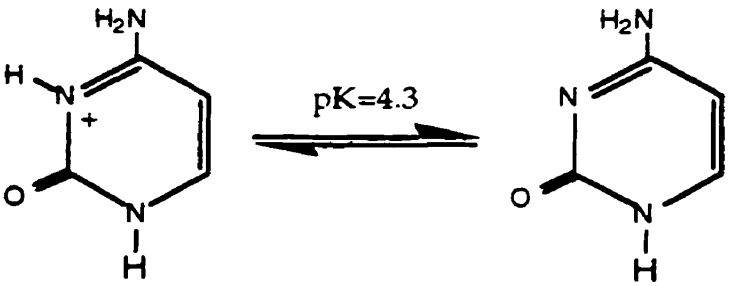


Figure 1.4. The protonation of cytosine. The equilibrium pK and the atom where the proton is gained or lost is shown.

Most DNA molecules *in vivo* are present as closed loops (Vinograd *et al.*, 1965). A good example is a small, circular plasmid. If both strands are intact, the circular DNA is in a covalently closed circular form. If phosphodiester bonds in one or both strands are broken, the circle is nicked or open circular. The DNA in a covalently closed circular form is termed relaxed if there are 10.5 bp per helical turn. However, most DNA *in vivo* is underwound, meaning there are more than 10.5 bp per helical turn. Underwinding introduces thermodynamic strain which can be accomodated either by strand separation sufficient to accomodate the "missing" number of bp, or supercoiling. If circular duplex DNA is described as being in helical turns about an axis, then supercoiling is twisting of that axis about itself. Linking number (Lk) refers to the number of helical turns in a closed circular DNA molecule when it is constrained to lie in a plane, and is a topological property. That is, it does not change with supercoiling - the only way to change Lk is to nick one of the strands, rotate one of the ends around the helix, and rejoin the ends. By convention, in a right-handed helix, Lk is positive. In underwound DNA the change in linking number from the relaxed state can be defined as follows:

$$\Delta Lk = Lk - Lk_0$$

where Lk_0 is the number of helical turns in relaxed DNA. The specific linking difference (sigma, σ) or superhelical density is defined as:

$$\sigma = \frac{\Delta Lk}{Lk_0}$$

As DNA is usually underwound in naturally occurring molecules, σ is negative and the DNA is said to be negatively supercoiled. For cellular DNA, σ falls into the range of -0.05 to -0.07. That is, from 5 to 7% of helical turns have been removed.

In vivo, supercoiling can be unrestrained, as in a circular plasmid or toroidally-coiled eukaryotic DNA free of protein. Restrained supercoiling occurs where DNA is bound by protein. Underwound DNA facilitates strand separation promoting cellular processes such as transcription. Strand separation and the energy provided by supercoiling is also thought to drive

the formation of sequence-specific DNA structures like cruciforms and triplexes (reviewed in Palecek, 1991; Yagil, 1991).

1.2 Early Studies on Triple Helical Nucleic Acids

Felsenfeld *et al.* (1957) reported the first formation of a triple helical nucleic acid using synthetic ribonucleotide homopolymers. The authors mixed poly(rU) with poly(rA), keeping constant the total concentration of bases, while varying the mole ratio of one polymer to the other. The observations relied on the hypochromic effect, or decrease in UV absorption, that occurs when polynucleotide strands form complexes in solution. The maximum decrease in absorption occurs with the greatest production of complex. From a plot of UV absorption versus mole ratio of polymers, the authors determined that in the absence of $MgCl_2$, the complex contained poly(rU) and poly(rA) in a 1:1 ratio. However, the presence of $MgCl_2$ drove the formation of a three-stranded complex containing poly(rU) to poly(rA) in a 2:1 ratio.

The complex had a sedimentation coefficient approximately 50% larger than the complex with a 1:1 ratio of polymers. Felsenfeld *et al.* (1957) proposed that the poly(rU) strand was hydrogen bonded to either adenine or uracil in a helical groove of the Watson and Crick poly(rA)+poly(rU) duplex to form the three-polymer complex. They noted, however, that two strong hydrogen bonds could only be formed with adenine. These bonds would later be referred to as Hoogsteen hydrogen bonds (see below). Felsenfeld *et al.* also proposed that the cation Mg^{2+} promoted the formation of the triple helical nucleic acid complex by neutralizing the negative charges on the phosphate backbones, thus reducing interstrand repulsion. The effect may have been through a non-specific shielding of the negative charges, but it seemed more likely that Mg^{2+} bound specifically to the nucleotides, as the extent of complex formation depended on $MgCl_2$ concentration. Complex formation was also promoted by other divalent cations such as Mn^{2+}, Ca^{2+}, and Zn^{2+}, while monovalent cations such as Na^{+} were less effective.

These findings were supported in other reports which came immediately after. Felsenfeld and Rich (1957) found that a specific reaction occurred between a duplex and incoming third strand bases; that is, poly(rA)+poly(rU) would react with poly(rU), but not with poly(rC) or

poly(rA). Furthermore, triplex formation was not limited to the polyribonucleotides poly(rA) and poly(rU). By UV mixing curve analysis and determination of sedimentation coefficients, Rich (1958) showed the formation of the all purine polyribonucleotide complex poly(I)+poly(rA)+poly(I). During this time, it was thought that synthetic triple stranded structures might be models of some (as yet undiscovered) biologically important structure consisting of an RNA strand wound about a duplex DNA.

Hoogsteen (1959, 1963) crystallized 9-methyladenine:1-methylthymine. Hydrogen bonding between the bases differed from that in the Watson and Crick scheme (Figure 1.1). In the latter, thymine N3 bonded to adenine N1. In the Hoogsteen bonding scheme, N3 of thymine was bonded to N7 of adenine. This bonding scheme confirmed the position of the third strand base relative to the purine of the duplex in the triple-stranded complex of Felsenfeld *et al.* (1957).

The study of complexes containing poly(G) was initially hampered by the tendancy of guanine homopolymers to form self-structures (Ralph *et al.*, 1962). Additionally, poly(G) did not react with poly(C)•poly(G) to form a triplex due apparently to the rigidity of the duplex. Lipsett (1963, 1964) avoided this problem by using short guanine oligomers (2-4 nucleotides) in combination with poly(rC). This allowed strand conformations to adjust to each other. Lipsett observed formation of complexes containing 2G:1C at neutral pH and 1G:2C as the pH was lowered below 6. This latter complex depended on the protonation of one-half of the cytosines and the author postulated that although the structure was unknown, the bases of the second poly(rC) strand were hydrogen bonded to guanines via Hoogsteen bonding. Howard *et al.* (1964) in studying the same complex reported the existence of one neutral and one protonated cytosine base for each guanine base. Using infrared spectroscopy, they determined the groups involved in base pairing. As cytosine N3 was protonated, they proposed either Hoogsteen or reverse Hoogsteen bonds (see Section 1.4.1) between cytosine in the third strand and guanine of the duplex and either parallel or antiparallel poly(C) strands.

Later work in the 1960's and early 1970's showed that triplexes could be formed using both RNA and DNA homopolymers. Riley *et al.* (1966) and Murray and Morgan (1973) studied the hybrid three-stranded complex poly(dT)•poly(dA)•poly(rU). (Note: in this thesis the convention in writing

a base triad will be that the base on the far right side of the triad is the one in the third strand.) It was found that in general, if two polymers were to interact to form a triplex, one needed to contain all purines and the other all pyrimidines (Howard *et al.*, 1964; Inman, 1964b; Lipsett, 1964; Riley *et al.*, 1966; Morgan and Wells, 1968; reviewed in Felsenfeld and Miles, 1967). Usually, the triplexes formed were of the double pyrimidine type (i.e. contained two pyrimidine strands). However, an all purine triplex poly(I)•poly(rA)•poly(I) had been formed (Rich, 1958) and also a double purine triplex dIn•dIn•dCn (Inman, 1964b).

On the basis of the deduced structures of homopolymer triple helical nucleic acids such as [poly(dT)•poly(dA)•poly(rU)] (Riley *et al.*, 1966) and [poly(C)•poly(G)•poly(CH^+)] (Howard *et al.*, 1964; Lipsett, 1964) and on the basis of model building, it was felt that in triple helical nucleic acids the ribonucleotide polymer bound via Hoogsteen bonding in the major groove of the duplex. If the sugars of the ribonucleotide polymer were in an *anti* orientation regarding the phosphate backbone as normally found in duplex DNA, then the RNA pyrimidine polymer strand would be antiparallel to the DNA pyrimidine of the duplex.

The next phase in work on triple helical nucleic acids was carried out by Morgan and Wells (1968) who investigated the specificity of base interactions in triplex formation using homopurine and homopyrimidine polymers (heteropolymers containing di-, tri-, and tetranucleotide repeats) consisting of a DNA duplex and a ribonucleotide polymer. It was found that poly(rUC) interacted with poly(dTC)•poly(dGA) under conditions of low pH to form the triplex poly[d(TC)]•poly[d(GA)]•poly[r(C^+U)]. However, no conditions were found for triplex formation between a polyribonucleotide and poly[d(TG)]•poly[d(CA)] or poly[d(AT)]•poly[d(AT)]. Additionally, the random copolymer $r(U,C)_n$ did not interact to form a three-stranded complex. However, it was apparent that many different bases in both RNA and DNA could interact to form triplexes.

The fact that triplex formation was a general phenomenon yet required specific interactions between bases suggested the possibility of a biological role. However, in considering a biological role for three-stranded structures containing a DNA duplex and an RNA third strand, Morgan and Wells (1968) stated,

stretches of a naturally occurring DNA which contain purines in one strand and the complemenary pyrimidines in the other strand, are statistically rare. For example, the probability of getting 20 pyrimidines in a row is one in 2^{20} or about one in 10^6.

However, Morgan and Wells (1968) also showed that the triplex poly[d(TC)]•poly[d(GA)]•poly[r(C^+U)] inhibited transcription. They therefore proposed that where they did occur, such asymmetric stretches of DNA might have an *in vivo* function. At this time, there was evidence that stretches of pyrimidine-rich bacteriophage DNA could facilitate the entry of RNA polymerase (Szybalski *et al.*, 1966; Summers and Szybalski, 1968). Szybalski *et al.* (1966) had also proposed that promoter regions may contain polypurine•polypyrimidine (pur•pyr) regions where triplex DNA could be formed and hence transcription regulated. Pur•pyr tracts are regions of DNA with all purines on one strand and the complementary pyrimidines on the other. A few years later Murray and Morgan (1973) also suggested, based on the results of detailed enzymatic and physical studies of the triplex $dT_n \bullet dA_n \bullet rU_n$, that triplex formation could play a role in the regulation of transcription and possibly even replication.

In other relevant work, Suwalski *et al.* (1969) reported X-ray crystallographic studies which concluded that the naturally-occurring, tetravalent cation spermine bound in the minor groove of duplex DNA. Murray and Morgan (1973) found that spermine bound equally well to the duplex poly(dT)•poly(dA), and to the triplex poly(dT)•poly(dA)•poly(rU). From modelling results, they also proposed that spermine lay across the minor groove, forming ionic and hydrogen bonds with the phosphate groups and hence binding the two duplex strands together.

At this time, a role for triplexes containing an RNA strand bound to duplex DNA was also proposed in chromosome condensation. It had been noted that *Escherichia coli (E. coli)* chromosomes relaxed after treatment with ribonuclease (Pettijohn and Hecht, 1973; Worcel *et al.*, 1973). Arnott and Bond (1973) suggested a pyrimidine RNA strand could form a triplex involving two widely separated pur•pyr tracts, with the duplex DNA in the middle looping out. Soon after, pur•pyr tracts of around 750 bp in length were found in *Drosophila* (Birnboim, 1975; Sederoff *et al.*, 1975). Therefore, it appeared that the pur•pyr tracts necessary for the formation of such looped structures existed.

Stereochemical models of the structures of the three strands in a triplex were based on X-ray fiber diffraction studies. X-ray data on both the ribonucleotide triplex poly(rU)•poly(rA)•poly(rU) (Arnott and Bond, 1973) and the deoxyribonucleotide triplex poly(dT)•poly(dA)•poly(dT) (Arnott and Selsing, 1974; Arnott *et al.*, 1976) was consistent with the three strands all being A form DNA. In the DNA triplex, the sugar conformation was C3'-endo and there were 12 base triads per helical turn. A model where the triplex strands were in A conformation was appealing as the major groove in A form is deeper than B form, and would accomodate a third strand more easily.

Johnson and Morgan (1978), following the work of Morgan and Wells (1968) and Murray and Morgan (1973), attempted to form the all DNA triplex $d(TC)_n \bullet d(GA)_n \bullet d(C^+T)_n$ from the duplex $d(TC)_n \bullet d(GA)_n$ at pH 5.5. However, the final structure appeared to be a tetraplex. In 1979, Lee *et al.* reported that based on a number of enzymatic and physical methods such as nuclease digestion (S1, DNase I and II), buoyant density determinations, temperature-absorbance transitions, and circular dichroism spectra, the structure formed was a double pyrimidine triplex. Earlier confusion as to the structure of the complex had arisen from a physically-linked but not hydrogen bonded purine strand $d(GA)_n$.

At the time this paper was published, there were already several publications recognizing that long pur•pyr stretches were common in both prokaryotes and eukaryotes (Sederoff *et al.*, 1975; Denniston-Thomson *et al.*, 1977; Subramanian *et al.*, 1977; Birnboim, 1978; Sures *et al.*, 1978).

1.3 S1 Nuclease Sensitive Polypurine•Polypyrimidine Tracts

At the time of the initial studies on chemical and physical properties of triplexes there was no evidence that triplexes might exist *in vivo*. Speculation that such structures could influence transcription or replication was confined to a few authors, and triplexes remained a laboratory phenomenon of synthetic polymers and oligonucleotides. More widespread interest in the possible biological significance of triplexes arose out of the discovery that pur•pyr tracts were statistically overrepresented in a wide range of organisms.

Researchers were aware early on of pyrimidine clustering (isostychs) in genomic DNA (Tamm *et al.*, 1953; Spencer and Chargaff, 1963). These tracts were detected by the process of acid depurination particularly in analyses of human spleen and calf thymus DNA, and to a lesser extent they were also found in the bacterial genome (Shapiro and Chargaff, 1957). Later, pyrimidine clusters were commonly found in the repetitive DNA (satellite and intermediate C_0t regions) of murine L cells (Straus and Birnboim, 1974; Birnboim *et al.*, 1976). Such tracts were also found in the repetitive DNA of the sea urchin (Case and Baker, 1975) and in many other species (Birnboim *et al.*, 1979). At this point it was unknown whether pyrimidine tracts also occurred in non-repetitive regions of DNA. Interest in these tracts was awakened with the discovery of an S1 nuclease sensitive pur•pyr tract in the chicken β globin gene promoter (Larsen and Weintraub, 1982) and in the sea urchin histone gene (Hentschel, 1982).

Eukaryotic DNA is constrained in chromatin by winding around histone proteins. In order for regulatory proteins to bind, the DNA must unwind from the histones. At this point it becomes DNase I hypersensitive (reviewed in Gross and Garrard, 1988; Freeman and Garrard, 1992). In the late 1970's and throughout the 1980's it was discovered that many DNase I hypersensitive areas in chromatin were also hypersensitive to single-strand specific nucleases. The main single-strand nuclease used was S1, which has maximum activity at pH 4.2 (Vogt, 1973). Analyses of S1 sensitivity were carried out both by digesting nuclei in low pH buffer, and *in vitro* on sequences cloned into supercoiled plasmids. Research carried out by many authors led to the discovery that many S1 nuclease sensitive regions, which often mapped to the promoter regions of genes, contained pur•pyr tracts (Elgin, 1982; Hentschel, 1982; Larsen and Weintraub, 1982; Goding and Russel, 1983; Nickol and Felsenfeld, 1983; Schon *et al.*, 1983; Shen, 1983; Weintraub, 1983; Evans *et al.*, 1984; McKeon *et al.*, 1984; Ruiz-Carillo, 1984; Christophe *et al.*, 1985; Margot and Hardison, 1985; Financsek *et al.*, 1986; Fowler and Skinner, 1986; Hoffman-Liebermann *et al.*, 1986; Siegfried *et al.*, 1986; Yavachev *et al.*, 1986; Boles and Hogan, 1987; Finer *et al.*, 1987; reviewed in Wells *et al.*, 1988; Htun and Dahlberg, 1989; Hoffman *et al.*, 1990; reviewed in Bucher and Yagil, 1991; Tripathi and Brahmachari, 1991).

There is a very significant statistical overrepresentation of pur•pyr tracts in eukaryotic genomes. It has been estimated that the tracts represent

up to 1% of total DNA in eukaryotes (Birnboim *et al.*, 1979; Gillies *et al.*, 1984; Hoffman-Liebermann *et al.*, 1986; Manor *et al.*, 1988; Wong *et al.*, 1990, Bucher and Yagil, 1991; Tripathi and Brahmachari, 1991; Yagil, 1993). As an example, Hoffman-Libermann *et al.* (1986) estimated that pur•pyr tracts 200-300 bp long occur in the human genome more than 100,000 times. Behe (1987) observed that the human β globin sequence is strongly biased in favor of long pur•pyr stretches: an analysis of 67 kb of the gene showed that pur•pyr tracts 10 or more nucleotides long occurred 3.5 to 4 times more than one would expect if they occurred randomly. Beasty and Behe (1988) found a similar overrepresentation of pur•pyr tracts in a few eukaryotic viruses, but with one exception, not in*E. coli.* or bacteriophage as a group.

Bucher and Yagil (1991) designed a program to analyze the length and frequency of occurrence of pur•pyr tracts in 163 transcribed genes in eukaryotes. This data bank was evenly representative of different species as well as class of gene. In addition, the complete genomes of SV40, the chloroplast of *N. tobacum*, yeast 2 micron plasmid, phage lambda, pBR322, and the *lac* operon of *E. coli* were analyzed. Pur•pyr tracts were highly significantly overrepresented in all the transcribed genes as well as in the chloroplast. In prokaryotic genes, only tracts $\geq$ 12 bases long were overrepresented and these were found near regulatory regions.

Recently, Raghavan *et al.* (1997) analyzed the distribution of pur•pyr tracts in the 16 completely sequenced chromosomes of the yeast *Saccharomyces cerevisiae*. Pur•pyr tracts $\geq$ 15 nucleotides in length were present on average 15 times more than expected ($p < 0.001$).

In studies too numerous to list, from the mid-1980's to the present, pur•pyr tracts have been found distributed throughout the eukaryotic genome: the 5' flanking regions of genes, 3' ends of genes, gene coding regions, repetitive DNA, recombination sites, and intron sequences. It has been said that the tracts are mainly located in gene regulatory regions (Birnboim *et al.*, 1979; Hoffman-Liebermann *et al.*, 1986; Manor *et al.*, 1988; Wong *et al.*, 1990). While many tracts have been found in the 5' flanking regions of genes, recent researchers might disagree with the statement that they are mainly located in regulatory regions. For example, Raghavan *et al.* (1997) in the study mentioned above, found that pur•pyr tracts $\geq$ 15 nucleotides in length were approximately equally represented in coding and noncoding regions. However, as the ratio of the noncoding to coding

regions is about 1.5, the occurrence of the tracts in the noncoding regions was statistically more significant.

1.4 Intramolecular Triplex DNA *in Vitro*

1.4.1 Molecular Structure

The S1 nuclease sensitivity of pur•pyr tracts both *in vivo* and when cloned into supercoiled plasmids led to speculation that they formed a non-B-DNA structure with a single-stranded region (reviewed in Wells, 1988). Classical biophysical techniques used in the early studies of triplexes were unable to detect conformation changes confined to a few bp in plasmids. Enzymes such as the single-strand specific nucleases P1 and S1, restriction enzymes, and chemicals such as bromo- and chloroacetaldehyde, osmium tetroxide, potassium permanganate, and diethylpyrocarbonate, which react with functional groups of unpaired bases, all proved extremely useful in elucidating localized conformations down to the individual base level in supercoiled plasmids (reviewed in Soyfer and Potaman, 1996).

The S1 sensitivity of pur•pyr tracts was puzzling as the tracts were unable to adopt cruciform or Z-DNA conformations. A number of non-triplex structures were initially proposed by way of explanation. One of the earliest proposals was a slipped structure (Hentschel, 1982; Glikin *et al.*, 1983; McKeon *et al.*, 1984). A slipped structure was shown not to be the correct model for most tracts by chemical analysis and DNase I footprinting at the bp level. Another proposal based on energetics was the formation of a left-handed non-Z DNA structure (Cantor and Efstratiadis, 1984; Margot and Hardison, 1985). Later, 2-D gel electrophoresis results showed that the S1 nuclease sensitive structure was right- rather than left-handed (Collier *et al.*, 1988; Lyamichev *et al.*, 1985; Pulleyblank *et al.*, 1985; Hanvey *et al.*, 1988a). Lyamichev *et al.* (1985) proposed an early model of H-form DNA, requiring H^+ ions for stabilization. This model was based on S1 nuclease sensitivity data and 2-D gel electrophoresis results showing the relaxation of supercoiled plasmid as the pH is lowered. A structure was proposed which contained a single purine strand and a pyrimidine columnar structure consisting of C•C^+ bp with thymines extruded. This was the first model to

account for the observed increased stability of the S1 nuclease sensitive structure at low pH.

Pulleyblank *et al.* (1985) proposed that the sequence $(A \bullet G)_n$ adopted a structure where $G \bullet C^+$ Hoogsteen bp alternated with A•T Watson-Crick bp. This was the first model to propose that the reason for the stability of the structure at low pH was cytosine protonation. It also proposed Hoogsteen base pairing where protonated cytosine paired with guanine. S1 nuclease sensitivity would occur at sites of conformational adjustment required to take up strain from conformational differences in the two strands.

Lee *et al.* (1984) proposed a structure which consisted of a duplex loop held together by a triplex formed between two pur•pyr tracts (Figure 1.5). This was quite similar to the hairpin triplex structure of Christophe *et al.* (1985). Finally, Evans and Efstratiadis (1986) suggested a heteronomous structure for the sequence $(AG)_n$, where the stacking differences of a two-stranded structure would produce a different backbone conformation for each strand.

Today, the accepted model for the structure adopted by most S1 nuclease-sensitive pur•pyr tracts is the intramolecular triplex, also commonly referred to as H-DNA (Figure 1.6; Lyamichev *et al.*, 1986; Mirkin *et al.*, 1987; Htun and Dahlberg, 1988, 1989). This is very similar to the structure initially proposed by Lee *et al.* (1984). In the formation of H-DNA, the strands in an insert sequence between two halves of a mirror repeat denature. A nucleation event follows where a single strand from one-half of the tract rotates and folds back to form the first base triad with the duplex bp in the other half of the tract. The single strand continues to wind back along the other half of the tract, forming Hoogsteen bp with the purine strand. The other unwound one-half strand remains unpaired.

The triplex third strand lies in the major groove of the duplex and after binding, the diameter of the resulting triplex is only a few angstroms wider than the original 20 angstroms of the duplex (Laughton and Neidle, 1992). The minor groove is too narrow to accomodate the third strand stably so that it can form Hoogsteen hydrogen bonds with the purine bases of the duplex (Cheng and Pettitt, 1992).

Duplex purines have two faces for forming hydrogen bonds (Sinden, 1994). Each face is capable of forming two hydrogen bonds with a base in another strand. This is why a pur•pyr tract is necessary for triplex

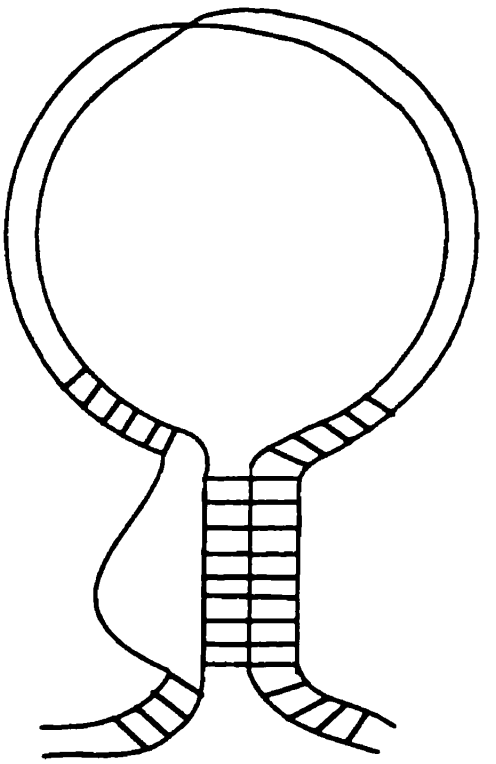


Figure 1.5. Model proposed by Lee *et al.* (1984). A duplex loop of DNA is constrained at its base by triplex formation.

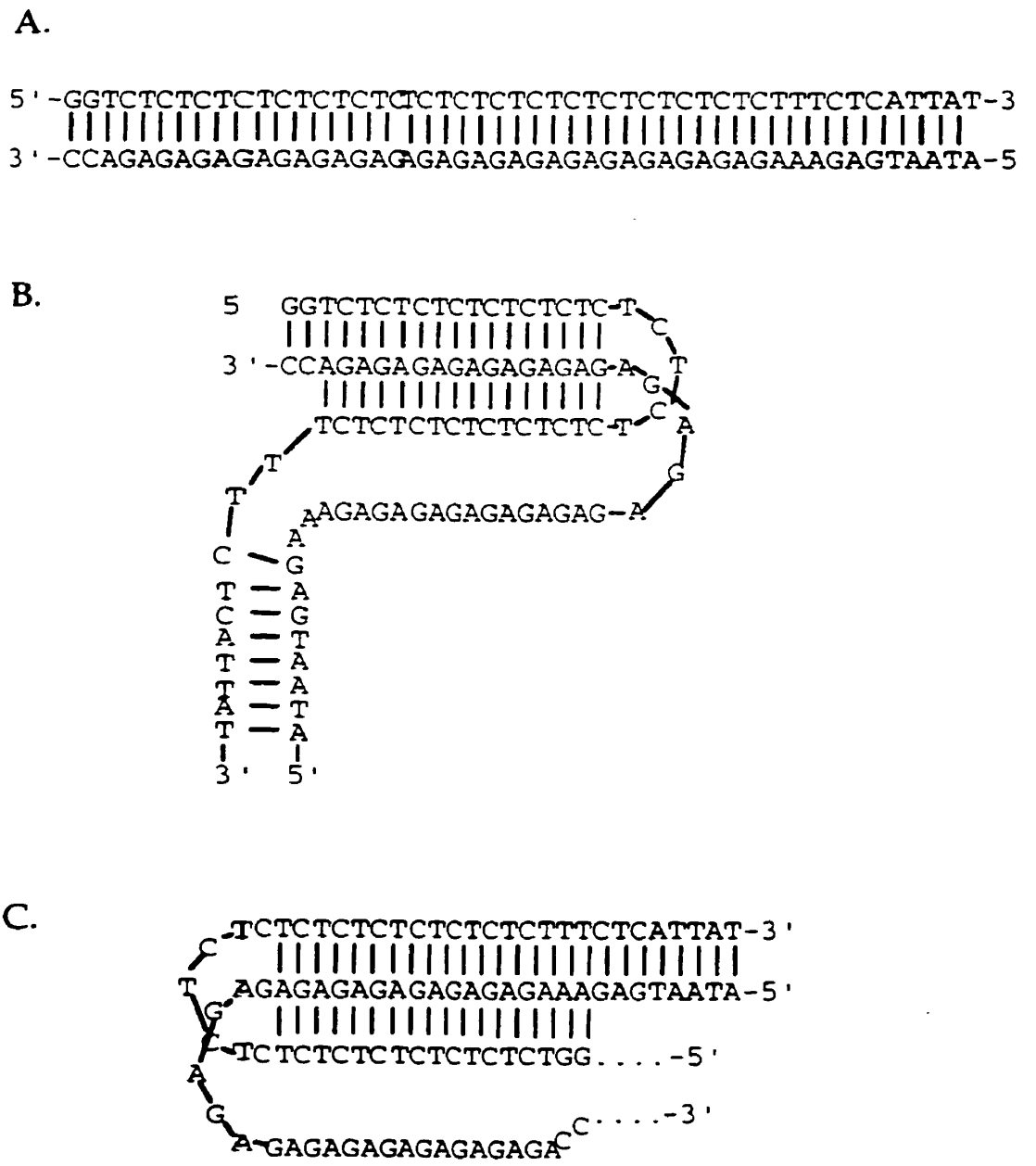


Figure 1.6. The two possible isomers of H-DNA. A. Mirror repeat pur•pyr tract. B. H-y3. C. H-y5.

formation. A triplex where the purine bases alternated randomly between the two strands of the duplex would be energetically unfavorable as it would result in distortion of the sugar-phosphate backbone of the third strand and a loss of stacking interactions (Soyfer and Potaman, 1996).

The third strand in an intramolecular triplex can be either purine or pyrimidine: both types of bases can form two hydrogen bonds with the duplex purine base (Felsenfeld *et al.*, 1957; Miles, 1964; Broitman *et al.*, 1987). The double pyrimidine intramolecular triplex, pyr•pur•pyr class of triplex, or H-DNA (Figure 1.6) was discovered first (Lyamichev *et al.*, 1985, 1986). The term H-DNA was originally used to refer to the requirement for H^+ ions in the medium (i.e an acidic pH). This was to fulfill the requirement for protonation of cytosine at the N3 position (Figure 1.4), so that two Hoogsteen bonds could be formed with the purine strand of the duplex. H-DNA also refers to the hinged structure of the intramolecular triplex. Common base triads in this type of triplex are the isomorphous base triads, $C•G•C^+$ and T•A•T (Figure 1.7).

The third strand bases in the model of H-DNA proposed by Lyamichev *et al.* (1986) are in *anti* conformation with respect to their sugar rings. Hence, the third strand is parallel to the purine strand of the duplex and anti-parallel to its identical strand in the duplex (Cheng and Pettitt, 1992a). This strand orientation agrees with the X-ray data of Arnott and Selsing (1974), based on the poly(dT)•poly(dA)•poly(dT) triplex. The structure of the pyr•pur•pyr triplex has also since been confirmed by several high resolution proton NMR studies showing hydrogen bonding schemes, strand positions, and backbone structures (de los Santos *et al.*, 1989; Rajagopal and Feigon, 1989a,b; Mooren *et al.*, 1990; Pilch *et al.*, 1990a).

H^*-DNA refers to an intramolecular triplex containing two purine strands (pyr•pur•pur class of triplex; Bernues *et al.*, 1989). This class of triplex generally does not require an acidic pH for formation but does require multivalent cations for stabilization. The common base triads in H^*-DNA, C•G•G and T•A•A, are not quite isomorphous (Figure 1.8). Therefore, there is some distortion of the sugar-phosphate backbone in the third strand. A number of different techniques have shown that a small number of T•A•T base triads can also be accomodated in this type of triplex (Radhakrishnan *et al.*, 1991). In H^*-DNA, the bases are again in an *anti* conformation with respect to the sugars. Reverse-Hoogsteen bonds are

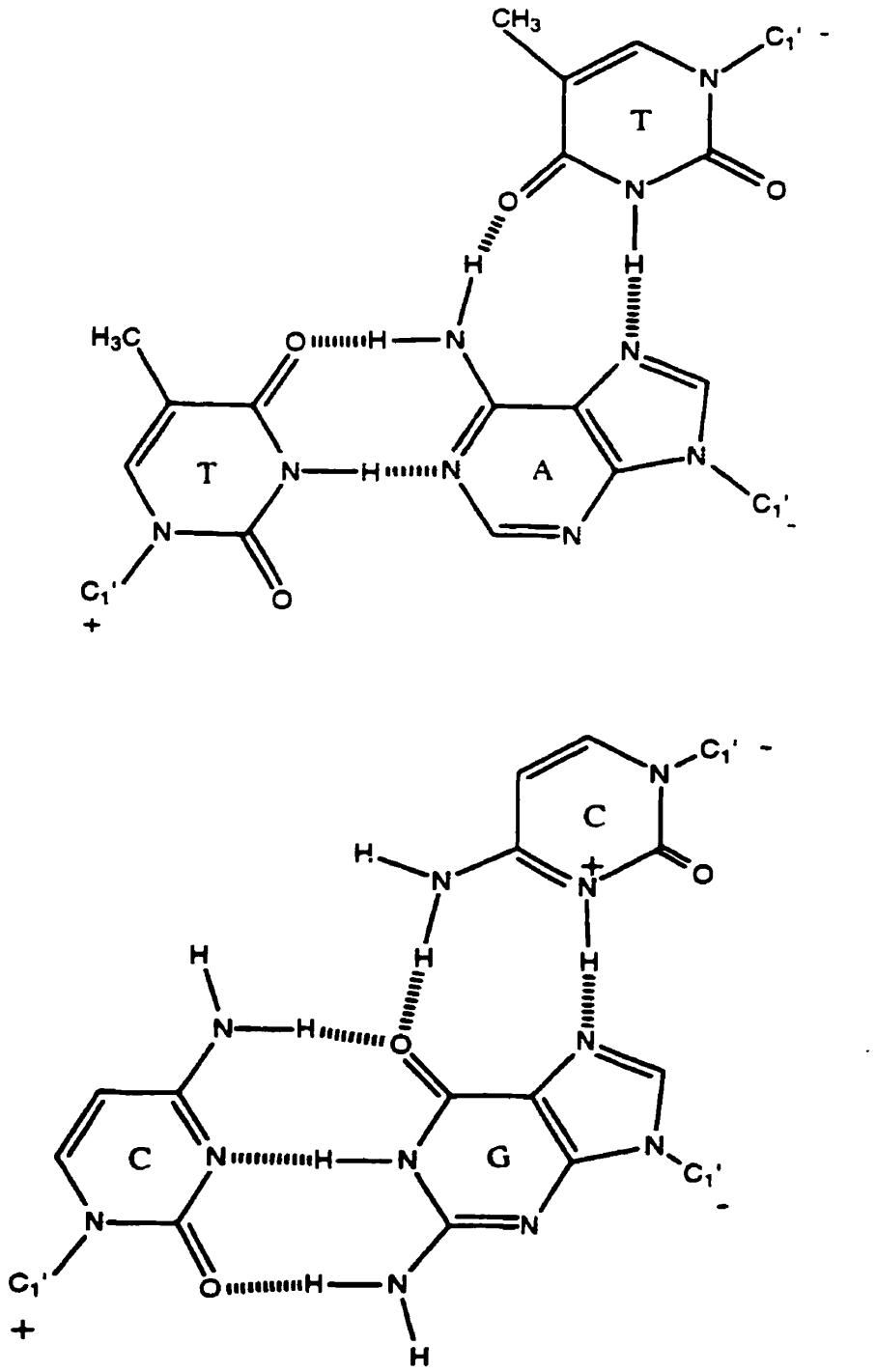


Figure 1.7. The isomorphous base triads in a pyr•pur•pyr triplex.

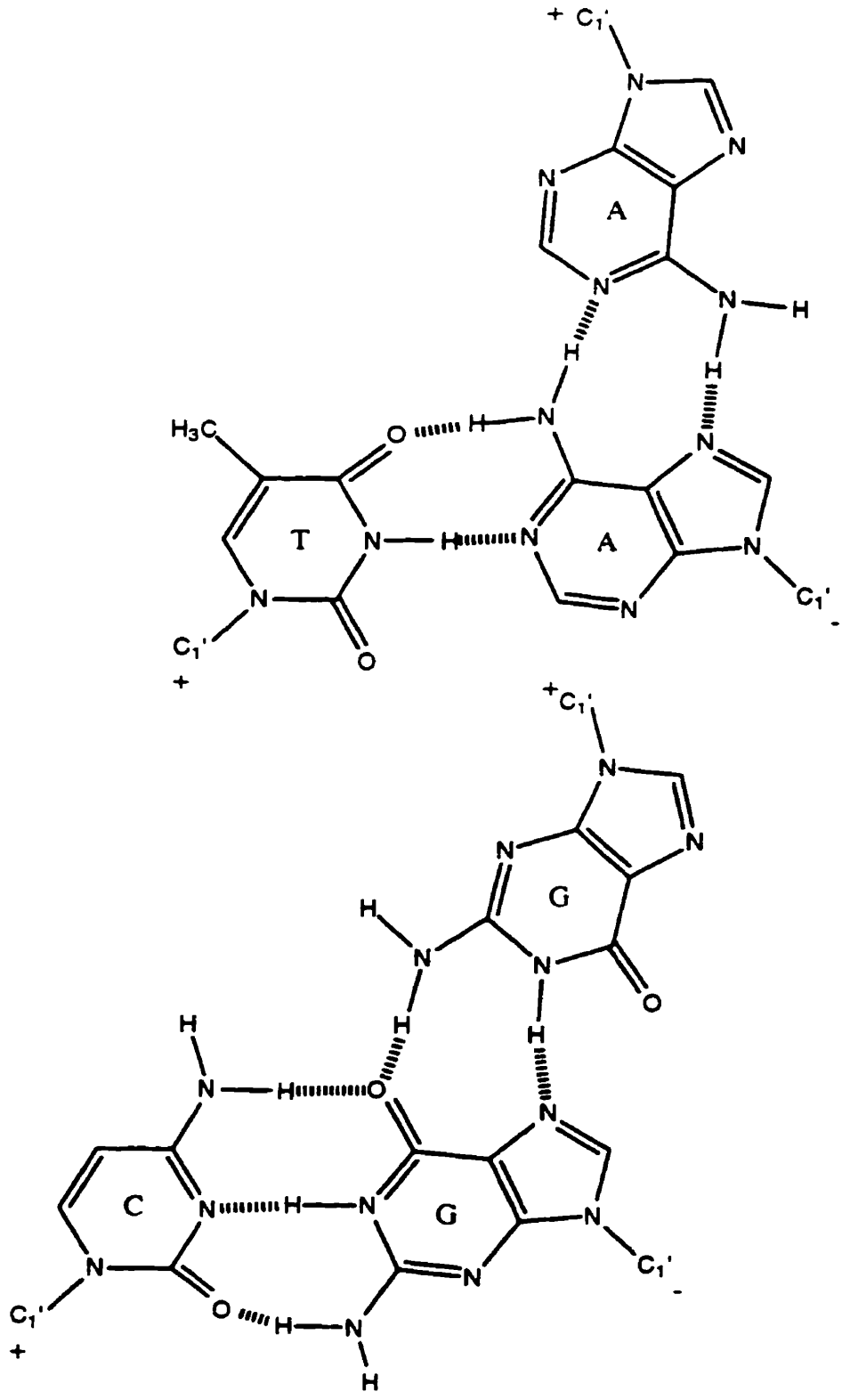


Figure 1.8. Base triads in a pyr•pur•pur triplex.

formed between the bases of the third strand and the purines of the duplex; the third (purine) strand is therefore antiparallel to the like (purine) strand of the duplex (Kohwi and Kohwi-Shigematsu, 1988). The *anti* orientation of the bases with respect to the sugars and the antiparallel orientation of like strands in the pyr•pur•pur triplex was confirmed by NMR (Radhakrishnan *et al.*, 1991; 1993).

A number of authors have also proven the models of H-DNA and H*-DNA by chemical probing to the level of individual bases (Vojtiskova and Palecek, 1987; Htun and Dahlberg, 1988; Johnston, 1988; Kohwi and Kohwi-Shigematsu, 1988; Vojtiskova *et al.*, 1988; Voloshin *et al.*, 1988).

Two isomers of both H-DNA and of H*-DNA are possible due to the mirror repeat symmetry of the pur•pyr tracts from which they are formed (Mirkin *et al.*, 1987; Htun and Dahlberg, 1989). The H-DNA conformer in which the 3' half of the pyrimidine tract forms the third strand of the triplex is designated H-y3, and the conformer in which the 5' half of the pyrimidine tract folds back to become the third strand of the triplex is designated H-y5 (Figure 1.6). Similar terminology is applied to the H*-DNA isomers, H-r3 and H-r5. For structural reasons, formation of the H-y3 isomer relaxes more supercoils than formation of the H-y5 isomer. Formation of the H-y5 isomer is therefore energetically unfavorable at high levels of negative supercoiling, but becomes possible at low levels (Htun and Dahlberg, 1989).

Initially, the strands in a triplex were considered to conform to A type DNA based on X-ray fibre studies. Data gathered using a number of different methods over the last few years has shown the strands in a triplex assuming an A or B form, or a combination of characteristics of each, depending on the sequences involved and the sugar-phosphate backbone type (reviewed in Soyfer and Potaman, 1996).

1.4.2 Factors Stabilizing and Promoting Formation of H-DNA

Many factors are important in promoting triplex formation and stability (reviewed in Cheng and Pettitt, 1992; Soyfer and Potaman, 1996). As the research in this thesis involves formation of pyr•pur•pyr or double pyrimidine triplexes, the discussion below focuses on the major factors shown *in vitro* to stabilize this class of intramolecular triplex.

There is an interdependence of pH, superhelical density, and pur•pyr tract length in the formation of H-DNA (Soyfer and Potaman, 1996). In the absence of other factors, the upper pH limit for H-DNA formation is 5.0 to 5.5. Apart from the requirement for protonating cytosines so that two Hoogsteen bonds can be formed with duplex guanine bases, protonated cytosines also stabilize the triplex by interacting with the negatively-charged phosphate backbone of the duplex (Soyfer and Potaman, 1996). However, although an acidic pH may be required for formation of a pyr•pur•pyr triplex, once formed, it may not be required for maintaining a stable structure. Hampel *et al.* (1991) observed that after pyr•pur•pyr triplexes were formed under acidic conditions, they were stable at neutral pH in the presence of polyvalent cations.

H-DNA formation is also promoted by the energy of supercoiling (Htun and Dahlberg, 1988, 1989). There is an interdependence between the pH at which H-DNA can be extruded and plasmid negative superhelical density (Lyamichev *et al.*, 1985). For example, at neutral pH, a very high negative superhelical density ($-\sigma=\sim 0.1$) was required to form H-DNA; at pH 4, no supercoiling was required.

One supercoil is relaxed for every 10 or 11 bp of pur•pyr tract incorporated into an intramolecular triplex (Htun and Dahlberg, 1989; Collier and Wells, 1990). Formation of an intramolecular triplex requires a tract length of about 20 bp (Soyfer and Potaman, 1996). For such a length, it is estimated that triplex formation would require a superhelical density of about -0.05 and a pH of about 5 (Lyamichev *et al.*, 1985; Htun and Dahlberg, 1989). The longer the tract, the more readily H-DNA is formed. With longer tracts, less superhelical density is required and/or H-DNA can be formed at higher pHs (Htun and Dahlberg, 1989; Lyamichev *et al.*, 1989; Lyamichev *et al.*, 1991; Collier and Wells, 1990; Soyfer *et al.*, 1992). For example, Collier and Wells (1990) noted triplex formation at neutral pH and a moderate level of supercoiling in the tract $d(GA)_{37}$•$d(TC)_{37}$.

Monovalent cations (especially Na^+) (Felsenfeld *et al.*, 1957; Felsenfeld and Rich, 1957), divalent cations (especially Mg^{2+} and Zn^{2+}) (Campos and Sabirana, 1991), and polyamines such as spermine and spermidine (Hampel *et al.*, 1991) play a very important role in triplex stability. Cations shield the negative charges on the sugar phosphate backbones and permit the three DNA strands to approach each other. Pyr•pur•pyr triplexes become more

stable as the concentrations of monovalent cations such as Na^+ or K^+ are increased beyond 100 mM (Krakauer and Sturtevant, 1968; Durand *et al.*, 1992). Wilson *et al.* (1994a), however, noted that the shielding is not as effective in triplexes containing protonated cytosines. Triplexes containing C•G•C^+ triads are stabilized with up to 0.4 M Na^+ and in fact are more stable than T•A•T-containing triplexes. At higher Na^+ concentrations, the trend reverses and the Na^+ has a destabilizing effect (Volker and Klump, 1994). Additionally, increasing concentrations of monovalent cations have a destabilizing effect on triplexes already stabilized with polyvalent cations (Lipsett, 1964; Lee *et al.*, 1984; Latimer *et al.*, 1989; Maher *et al.*, 1990; Lyamichev *et al.*, 1991; Singleton and Dervan, 1993). Na^+ and K^+ can also destabilize triplexes by promoting tetraplexes in G rich sequences (Olivas and Maher, 1995).

Much lower concentrations (1 to 10 mM) of divalent cations such as Mg^{2+}, Mn^{2+}, Zn^{2+}, and Ca^{2+} are as least as effective as the monovalents in stabilizing triplexes (Felsenfeld and Rich, 1957; Felsenfeld *et al.*, 1957). This is because the ions bind to the phosphates and have a higher charge density in lower ionic radii (Soyfer and Potaman, 1996; CRC Handbook of Chemistry and Physics, 1994). Multivalent cations, either divalent metal ions or polyamines, are also necessary for the formation of the pyr•pur•pur class of triplex (reviewed in Soyfer and Potaman, 1996).

Micromolar concentrations of the polyamine spermine stabilize both pyr•pur•pyr and pyr•pur•pur triplexes (Lyamichev *et al.*, 1991). Hampel *et al.* (1991) found that spermine promoted the formation of a pyr•pur•pyr triplex containing both cytosines and thymines at a faster rate at acidic pH. The naturally-occurring polyamines spermidine and spermine bind tightly to nucleic acids and therefore play important roles in cell proliferation and differentiation (Tabor and Tabor, 1984). The binding site of spermine to DNA has been difficult to determine. Recently, in a crystal structure, spermine was observed bound across the minor groove of a B-DNA fragment rather than within it (Tari and Secco, 1995).

Polyamines stabilize triplexes in a structure dependent manner, particularly with regards to charge. Of the naturally-occurring polyamines, spermine (+4) is more effective than spermidine (+3), which is more effective than putrescine (+2) (Tabor, 1962; Glaser and Gabbay, 1968; Hampel *et al.*, 1991; Thomas and Thomas, 1993). Considering synthetic polyamines,

divalent cations which have spacing between the positive charges are more stabilizing than Mg^{2+}, which has a point charge (Thomas and Thomas, 1993). Spermine and spermidine containing hydrophobic substituents have also been found to stabilize triplexes. Charge separation is also important in substituted polyamines for promoting triplex stability (Thomas and Thomas, 1993). However, increasing the size of hydrophobic moieties has a destabilizing effect (Glaser and Gabbay, 1968).

Tract sequence is also relevant to triplex stability (Roberts and Crothers, 1992). Any variation of mirror repeat symmetry makes intramolecular triplex formation much less favorable, and formation of the structure requires a higher level of supercoiling (Mirkin *et al.*, 1987). A high G•C content in the tract was also found to be stabilizing and required a lower level of supercoiling for extrusion of the H-form (Hanvey *et al.*, 1988a). In long tracts, more than one intramolecular triplex may be formed. A short sequence is required between the two halves of the mirror repeat, preferably one that is AT-rich (Mirkin *et al.*, 1987; Shimizu *et al.*, 1994a); the denaturation bubble necessary to initiate formation of H-DNA occurs more readily in the easily melting AT-rich sequences. Energetically, the optimum size of the spacer sequence is four to five bp (Harvey *et al.*, 1988). The absolute minimum size for the extrusion of the H-form is two bases (Veselkov *et al.*, 1993). Longer loop lengths increase the energetic cost of triplex formation. For example, increasing the size of the spacer sequence from four to 12 bases required a higher level of supercoiling for formation of H-DNA (Shimizu *et al.*, 1989).

Effect of intercalating agent, ethidium bromide (EtBr)

The cationic intercalating agent EtBr is used as a stain in the detection of double-stranded RNA and DNA. Upon intercalation, it fluoresces and is therefore useful for staining nucleic acids in gels and in solution as an indicator of nucleic acid concentration. EtBr also stains triplex DNA, although the interaction is not well understood. EtBr intercalates between T•A•T base triads, producing more fluorescence upon intercalation into a poly(dT)•poly(dA)•poly(dT) triplex than the parent duplex, poly(dT)•poly(dA) (Morgan *et al.*, 1979; Scaria and Shafer, 1991). In fact, EtBr was shown to stabilize the formation of a poly(dT)•poly(rA)•poly(dT) triplex (Pilch and Breslauer, 1994). In contrast, binding of EtBr to C•G•C^+-

containing triplexes is low. This is likely due to electrostatic repulsion between protonated cytosines and the positively-charged EtBr (Lee *et al.*, 1979).

1.5 Proposed Cellular Roles for Polypurine•Polypyrimidine Tracts

The overabundance and ubiquitous occurrence of pur•pyr tracts strongly suggests that they have a biological role or roles within the cell. A number of possibilities have been considered by researchers and these are discussed below. It is entirely possible that the tracts could have more than one role, depending on their composition and location within the genome.

1.5.1 No Role

The first possibility which should be addressed is that pur•pyr tracts may not have a role. Bucher and Yagil (1991) considered whether they might be a result of variation in gene structure or 'evolutionary remnants'. According to these authors, if the first possibility were the case, it is hard to understand why these tracts would not have been deleted over time. As for the second possibility, the tracts could only be very early evolutionary remnants, as they are present to the same degree in invertebrates and plants. Additionally, Bucher and Yagil noted that when pur•pyr tracts are present in the same gene in different species, the position and composition of the tracts varies. This suggested a transposition event rather than evolutionary conservation.

In contrast, Mirkin and Frank-Kamenetskii (1994) have expressed the view that pur•pyr tracts in the GC rich promoters of housekeeping genes such as c-*src* may merely represent a random juxtaposition of similar nucleotides, rather than having a functional significance.

1.5.2 Condensation of DNA in Chromosomes

The basic higher order structure of chromosomal DNA in eukaryotes throughout the cell cycle consists of loops about 70 kilobases (kb) in size which are tethered to form a scaffold network (Laemmli *et al.*, 1977). DNA regions known as scaffold-associated elements (SARs), also called MARs

(matrix-attached regions), may form the base of the loops. SARs have been characterized as extremely A-T rich sequences several hundred bp long in a number of species including yeast and humans (Gasser and Laemmli, 1987; Laemmli *et al.*, 1992). The tracts are a strong binding site for topoisomerase II (Sander and Hsieh, 1985) which is the major scaffold protein (Gasser *et al.*, 1986). The condensation and decondensation of chromosomes requires topoisomerase II (Adachi *et al.*, 1991; Kas *et al.*, 1993).

Other sequences potentially able to form triplexes are present in SARs (Boulikas, 1993). For example, Opstelten *et al.* (1989) found a number of long pur•pyr tracts including some with the sequence poly[d(GGAAA)].

Originally it was proposed that RNA was involved in the organization of looping in chromosomes (Pettijohn and Hecht, 1973; Sinden and Pettijohn, 1981). Further, Pettijohn and Hecht (1973) and Lee and Morgan (1982) proposed that RNA transcripts could form triplexes with pur•pyr tracts thus stabilizing such chromosomal loops of duplex DNA.

The distribution of pur•pyr tracts throughout the eukaryotic genome has recently led to speculation that triplex formation between separated pur•pyr tracts may mediate the looping of DNA in the condensation of chromosomes (Burkholder *et al.*, 1988; Hampel *et al.*, 1993). In support of such a model, linear plasmids with a pyr•pur tract at each end were able to form looped structures as a result of triplex formation between the tracts (Hampel *et al.*, 1993). Indirect evidence for triplex-mediated chromosomal condensation was provided by the low pH-dependent mobility change of chromosomes in the second dimension of pulsed field gel electrophoresis (Hampel and Lee, 1993).

Bucher and Yagil (1991) suggested that the large number of pur•pyr tracts found in their analysis of chloroplast intergenic regions and for example in high C_0t DNA by Birnboim *et al.*, (1976) supports a structural role for the tracts. On the other hand, Bucher and Yagil (1991) found the tracts were just as plentiful in mammalian viruses, yet the structural organization of viral DNA is minimal compared to the level required in eukaryotic chromosomes. Therefore, Bucher and Yagil proposed that the tracts would likely not fulfill a structural role in viruses.

1.5.3 Regulation of Transcription

The most frequently proposed cellular role for triplex DNA is in the regulation of transcription. This was first proposed by Lee *et al.* (1984) based on the model shown in Figure 1.5. Many pur•pyr tracts located in the 5' flanking regions of genes have been shown to form intramolecular triplexes when cloned into supercoiled plasmids (Lyamichev *et al.*, 1985, 1991; Kinniburgh, 1989; Glaser *et al.*, 1990; Pestov *et al.*, 1991; Horwitz *et al.*, 1994; Raghu *et al.*, 1994). The S1 nuclease sensitivity of many promoter pur•pyr tracts *in vivo* suggests the possibility of intramolecular triplex formation. It seems very likely that formation of a triplex in the regulatory region of a gene would have an effect on gene expression.

1.5.3.1 Intramolecular Triplex Formation in Promoter Polypurine•Polypyrimidine Tracts

A number of pur•pyr tracts in the 5' flanking regions of genes have been shown to be important for promoter function. A decrease in transcription resulted when pur•pyr tracts were deleted from the promoter regions of a number of genes. Some of these are: the human epidermal growth factor receptor (Johnson *et al.*, 1988), c-*myc* (Davis *et al.*, 1989; Postel *et al.*, 1989), *Drosophila hsp*26 (Gilmour *et al.*, 1989; Glaser *et al.*, 1990), *ets*-2 (Mavrothalassitis *et al.*, 1990), mouse c-Ki-*ras* (Hoffman *et al.*, 1990), actin (Chung and Keller, 1990), TGF-β3 (Lafaytis *et al.*, 1991), and decorin (Santra *et al.*, 1994).

In considering the effect that intramolecular triplex formation in the promoter region of a gene might have, the first possibility is that it could upregulate transcription. One way this could happen is for an H-DNA structure itself to have a positive *cis*-acting effect. Kato and Shimizu (1992) found the presence of an H-forming pur•pyr tract in the promoter of the β-lactamase gene in a plasmid led to a higher level of transcription than when the pur•pyr tract was not present. In another example, introduction of H-DNA-destabilizing mutations (*in vitro*) into the promoter pur•pyr tract of c-*myc* led to decreased transcriptional efficiency *in vivo* (Firulli *et al.*, 1994).

Another way in which intramolecular triplex formation may upregulate transcription is that the single DNA strand extruded on

formation may provide a binding site for RNA polymerase. Upon binding, RNA polymerase locally unwinds DNA (Saucier and Wang, 1972), which is a rate limiting step in transcription (Mangel and Chamberlin, 1974). However, RNA polymerase has been shown to bind to supercoiled plasmids containing single-stranded regions (Beard *et al.*, 1973; Germond *et al.*, 1974). Lee *et al.* (1984) proposed that the single strand extruded from an H-DNA structure could facilitate transcription by promoting the binding of RNA polymerase. Typically, such an H-DNA structure could arise from chromosomal reorganization prior to initiation of transcription (Soyfer and Potaman, 1996), or it could result from transcription of a downstream gene (Wang and Lynch, 1993).

Results from three model systems support the idea that the single-strand of H-DNA might promote the binding of RNA polymerase, thus upregulating transcription. Daube and von Hippel (1992) created a model of a functional transcriptional complex using an RNA-DNA bubble hybrid duplex containing a noncomplementary sequence 12 nucleotides long. An RNA primer was hybridized to the noncomplementary sequence. *E. coli* RNA polymerase and T7 RNA polymerase carried out efficient transcription from this construct. In addition, the *E. coli* RNA polymerase core enzyme [without the sigma (σ) factor responsible for the binding of the polymerase] also formed a functional transcription complex, suggesting that a locally unwound structure might facilitate the binding of RNA polymerase.

Mollegaard *et al.* (1994) created a D-loop structure in which a strand of DNA was displaced by a homopyrimidine peptide nucleic acid binding to duplex DNA. *E. coli* RNA polymerase holoenzyme initiated transcription from this template, creating a transcript the length of the DNA. As the peptide nucleic acid does not have a 3' hydroxyl group, the polymerase must have initiated transcription without a primer. The level of transcription initiated from this D loop, where σ factor played no role, was equivalent to that initiated from the *E. coli lac* UV5 promoter.

Finally, Aiyar *et al.* (1994) showed transcription initiation in both directions from the edges of a synthetic 12-bp mismatched bubble construct lacking promoter sequences, using both *E. coli* core RNA polymerase and holoenzyme. The results from these three model systems suggest that RNA polymerase entry is facilitated by an unwound DNA structure.

On the other hand, formation of an intramolecular triplex in the promoter region of a gene could downregulate transcription. Several ways can be envisaged (Sinden, 1994). First, the triplex DNA conformation could be unfavorable for the entry or passage of RNA polymerase. An interesting case is the failure of the transcriptional switch from expression of the fetal γ-globin gene to the β-globin gene which normally occurs at birth in humans, resulting in a disease called the hereditary persistence of fetal hemoglobin. Point mutations associated with the condition, located in the 5' flanking region of the γ-globin gene, have a destabilizing influence on an H-DNA structure *in vitro* (Ulrich *et al.*, 1992; Bacolla *et al.*, 1995). If the H-DNA structure is important for regulating gene expression, then the destabilizing point mutations associated with the disease may result in a failure to repress transcription.

Another way intramolecular triplex formation could down regulate transcription is that single strand specific proteins could bind to the extruded strand, preventing binding by RNA polymerase. A number of genes have been identified which have promoter pur•pyr tracts able to form H-DNA and either the pyrimidine or purine strand of the tracts binds proteins. One such gene is the *Drosophila* heat shock gene encoding hsp26 (Gilmour *et al.*, 1989). Chemical and nuclease probing has shown that the sequence $d(CT)_{10}$•$d(AG)_{10}$, located at -100, is able to form H-DNA *in vitro*. A *Drosophila* protein was identified which binds the pyrimidine strand of the sequence. In another example, in a situation reminiscent of the switch from expression of fetal γ-globin to β-globin, a nuclear factor has been identified which binds to a pur•pyr tract upstream of the human δ globin gene (O'Neill *et al.*, 1991). The pur•pyr tract is located between fetal and adult β-globin-like genes. The factor binds at much higher levels in erythroid cells expressing adult globin than in cells expressing either embryonic or fetal globin. Also, the factor is only expressed in hemopoetic cells lines. The fact that the factor binds to a sequence which has potential to form a triplex, and that the factor only binds in adult erythroid cells, suggests that both the factor and the sequence play a role in the developmental switch from fetal to adult hemoglobin. Many other proteins have been identified which bind to one or the other strand of a promoter pur•pyr tract (Biggin and Tjan, 1988; Brunel *et al.*, 1991; Yee *et al.*, 1991; Kennedy and Rattner, 1992; Muraiso *et al.*, 1992; Kolluri *et al.*, 1992; Aharoni *et al.*, 1993; Goller *et al.*, 1994;

Hollingsworth *et al.*, 1994; Horwitz *et al.*, 1994; Ito *et al.*, 1994; Kovacs *et al.* 1996). Apart from the possibility of downregulating transcription by binding to the single strand extruded by triplex formation and preventing entry of RNA polymerase, these single strand binding proteins may be playing other roles. Such possible roles are discussed elsewhere in this thesis as applicable.

Finally, triplex formation could prevent binding of B-DNA specific transcription factors which would otherwise activate transcription. A number of proteins have been identified which specifically bind to promoter pur•pyr tracts in the duplex form. For example, NSEP binds the c-*myc* promoter (Davis *et al.*, 1989), BPG1 binds to the (G)n•(C)n tract of the chicken β-globin gene (Clark *et al.*, 1990), NFκB binds to the interleukin-2 receptor α gene (Grigoriev *et al.*, 1992), and the transcription factor Sp1 binds to the pur•pyr tracts of c-*myc* (Kinniburgh, 1989) and the dihydrofolate reductase gene (Gee *et al.*, 1992). In the latter case, triplex formation in the pur•pyr tract was shown to prevent binding of Sp1 (Gee *et al.*, 1992). An S1 hypersensitive promoter pur•pyr tract important in promoter function is located about 100 bp upstream of multiple initiation sites in the mouse c-ki-*ras* oncogene (Hoffman *et al.*, 1990; Pestov *et al.*, 1991). Using band shift assays, Hoffman *et al.* (1990) found a nuclear protein which bound to the tract.

An inhibitory effect of intramolecular triplex formation on transcription has been shown in at least two experimental systems. Kohwi and Kohwi-Shigematsu (1991) used poly(dC)•poly(dG) able to form an intramolecular pur•pur•pyr triplex to investigate the influence of triplex DNA on gene expression. The effect on transcription in mouse LKT$^-$ cells using a reporter gene construct correlated directly with tract length. A tract of 27-30 nucleotides stimulated transcription at a level equivalent to the polyoma enhancer. A tract of length greater than or equal to 35 nucleotides lacked the stimulatory effect. The authors showed that a supercoiled plasmid containing a tract 30 nucleotides long competed for a binding factor whilst a tract 35 nucleotides in length or longer did not compete for the factor. The longer tract formed an intramolecular triplex when cloned in a supercoiled plasmid. When similarly cloned the shorter tract (30 nucleotides or less) remained as B-DNA. These results suggested the longer tract was driven by supercoiling *in vivo* to form an intramolecular triplex.

The intramolecular triplex prevented a duplex DNA-specific transcription factor from binding, thus downregulating gene expression.

In another series of experiments, Brahmachari *et al.* (1997) found that incorporation of a pur•pyr tract upstream of lacZ led to a several fold reduction in gene expression in mammalian cells. A similar reduction in transcription was observed when a cassette containing the tract upstream of lacZ was integrated into a yeast chromosome upstream of a promoter.

1.5.3.2 B-DNA/H-DNA Molecular Switch

There may be an interplay between the formation of H-DNA and factors acting in *trans* in the regulation of expression of a gene having a promoter pur•pyr tract. *Trans*-acting factors would be purine or pyrimidine strand-specific proteins or triplex DNA-specific proteins, and proteins binding to the duplex conformation of the tract. More than one author has proposed that interplay between such factors could influence the equilibrium between different conformational states of the same promoter (Mirkin and Frank-Kamenetskii, 1994; Kinniburgh *et al.*, 1994). The effect on transcription would depend on whether it was the triplex or duplex conformation of the pur•pyr tract that promoted transcription. The interaction between *trans*-acting factors and promoter conformation in the regulation of transcription is considered further in Section 1.8, using the specific example of the c-*myc* gene.

1.5.3.3 Summary

Although there is evidence that pur•pyr tracts in 5' flanking regions of some genes are important in the regulation of transcription, it is not clear whether it is the H-DNA structure itself which is involved in the regulation. At present, it is equally possible that regulation of transcription is via proteins which bind to the duplex DNA form of the pur•pyr tract. A problem with proposing a regulatory role for pur•pyr tracts is that only rarely is the homology in the location of the tracts in the same gene from different species (Bucher and Yagil, 1991; Soyfer and Potaman, 1996). For example, Behe (1987) noted that the single strand nuclease sensitive dG tract of the chicken β globin promoter is not found in rabbit, human or mouse

promoters. Bucher and Yagil (1991) investigated the positions of pur•pyr tracts in the histone H4 genes from *X. laevis*, chicken, mouse, and sea urchin. Eighty percent of the bases between base 36 and 72 were purines. However, only the promoter of the mouse H4 gene had a virtually continuous stretch of purines from -148 to -115. Thus, for this gene at least, if the pur•pyr tract plays a regulatory role, it may differ between species.

1.5.3.4 Supercoiling Density

In general, an increased level of supercoiling facilitates transcription (Wang and Lynch, 1993). The free energy of supercoiling may assist RNA polymerase in locally unwinding and denaturing duplex DNA in the initiation of transcription (Parvin and Sharp, 1993). Another possibility to consider, which may be related to a role in the regulation of transcription, is that triplex formation in the 5' flanking region of a gene may regulate superhelical density. Formation of non-B DNA structures has been shown to relax superhelical tension in circular DNA (Wells *et al.*, 1988; Palecek, 1991). In non-transcribed genes, the DNA is restrained in negative supercoils by nucleosomes. Prior to transcription, the DNA is removed from the nucleosomes. H-DNA, along with other non-B DNA structures such as Z DNA, may be a way of relaxing the resulting unrestrained negative supercoiling (Weintraub, 1983; van Holde and Zlatanova, 1994; Sinden, 1994).

While increasing the level of supercoiling in general facilitates transcription (Wang and Lynch, 1993), supercoiling at levels above that found naturally has been shown to hinder transcription in both bacterial and eukaryotic genes (Brahms *et al.*, 1985). Kato and Shimizu (1992) have therefore suggested that formation of H-DNA may be a way of maintaining optimal superhelical density by removing excessive tension from supercoiling.

1.5.3.5 Other Possibilities for Triplex Involvement in Transcriptional Regulation

Another way that triplex formation could down regulate transcription is by blocking transcription elongation rather than initiation. Pausing or

blocking has been noted during transcription of HIV-1, human and mouse c-*myc*, adenosine deaminase (ADA), and the *Drosophila hsp*70 genes (reviewed in Spencer and Groudine, 1990). Sarkar and Brahmachari (1992) using degenerate codons, engineered a 38 nucleotide pur•pyr tract into the middle of the coding region of the β-galactosidase gene in pBlueScriptIISK$^+$. Following transformation of *E. coli*, the cells showed an 80% reduction in transcription of the β-galactosidase gene compared to the same strain transformed with a construct containing the same amino acids but using less preferred codons and lacking the pur•pyr tract. An explanation for these results was that H-DNA formed within the coding region and blocked transcription.

In another study, Duval-Valentin *et al.* (1992) designed an oligonucleotide (13-mer) to target a pur•pyr sequence in the β-galactosidase gene of *tn3* located just downstream from the RNA polymerase binding site. The effect of binding of the oligonucleotide on transcription of the β-galactosidase gene was tested *in vitro*. Transcription by RNA polymerase was blocked at the start site in a temperature-dependent manner. The stability of the triplex formed was directly correlated with temperature. Again, the presence of an H-DNA structure was presumed to block transcription elongation in the coding region of the gene.

Looping via interaction between distant pur•pyr tracts may also be important in the regulation of transcription (Soyfer and Potaman, 1996). Such interactions may substitute for the well-known situation where a distant sequence such as an enhancer or upstream activating sequence is brought close to the promoter of a gene in the initiation of transcription (Guarente, 1988; Collado-Vides *et al.*, 1991), with the looping mediated by duplex DNA-binding proteins (reviewed in Schleif, 1992). A number of genes have two pur•pyr tracts upstream of the start site of transcription (Boles and Hogan, 1987; Gilmour *et al.*, 1989; Kohwi, 1989; Firulli *et al.*, 1992; Lu *et al.*, 1993). In the case of the *Drosophila hsp*26 gene, both tracts were found to act together in promoter transcriptional efficiency (Lu *et al.*, 1993).

1.5.3.6 Polypurine•Polypyrimidine Tracts as B-DNA Protein Binding Sites

Sometimes there is no evidence that a pur•pyr tract is involved in the regulation of transcription as H-DNA. For example, some tracts are too

short to form H-DNA (Chung and Keller, 1990; Lafaytis *et al.*, 1991). Additionally, the ability of a promoter pur•pyr tract to form H-DNA and its transcriptional efficiency are not always related. There are several cases in the literature where an H-DNA structure was not important for the transcriptional efficiency of a promoter. In one example, an H-DNA destabilizing point mutation introduced into the pur•pyr tract of the *Drosophila hsp*26 gene had no effect on transcription levels (Glaser *et al.*, 1990). Additionally, the level of transcription did not change when the triplex-forming (CT)n•(GA)n tract was substituted with the tract (C)n•(G)n. In another study, point mutations were introduced into the c-Ki-*ras* pur•pyr tract (Raghu *et al.*, 1994). Some of these destabilized H-DNA *in vitro* while others did not. However, a comparable decrease in promoter transcriptional efficiency occurred with both types of mutation.

These and similar results have led to the proposal that duplex pur•pyr tracts are binding sites for specific proteins (Glaser *et al.*, 1990; Raghu *et al.*, 1994). Formation of H-DNA is not required - it is the proteins which regulate transcription. Mutations in the tracts lead to decreased transcriptional efficiency not because the H-DNA structure is destabilized but because the protein binds to the mutated site with decreased affinity.

Mutations which destabilize an H-DNA conformation *in vitro* may in some cases change an imperfect binding site to one more specific for certain transcription factors *in vivo* (Soyfer and Potaman, 1996). An example to revisit is the failure of the transcriptional switch at birth from expression of the fetal γ-globin gene to the β-globin gene, resulting in the hereditary persistence of fetal haemoglobin. As discussed earlier, if an H-DNA structure is involved in regulating transcription from the globin gene locus, then the point mutations associated with the condition may result in a failure of transcriptional repression. Another possibility is that the duplex DNA pur•pyr tract is an imperfect binding site for one or more of the following transcription factors: SP1, the octamer-binding protein OTF-1, or the erythroid regulatory factor NFE-1 (Fischer and Nowock, 1990; Sykes and Kaufman, 1990; Jane *et al.*, 1993). Mutations in the sites may permit binding by one or more of the factors leading to inappropriate expression of the gene. At this time, not enough information is available to chose between these possibilities (Soyfer and Potaman, 1996).

1.5.4 Other Roles

1.5.4.1 Replication

Pur•pyr tracts may form triplexes and terminate replication in a similar way to which they may interrupt elongation of transcription. Murray and Morgan (1973) observed that binding of a third strand to a polynucleotide duplex inhibited replication. A $(GA)_{27}$•$(TC)_{27}$ sequence located 2 kb from the integration site of polyoma virus strongly terminated replication in rat cells. The tract paused replication enzymes *in vitro* under conditions where a triplex could form (Baran *et al.*, 1983, 1991; Lapidot *et al.*, 1989). Using two-dimensional gel electrophoresis to detect *in vivo* replication intermediates, Rao (1994) showed that replication fork movement slowed in SV40 variants containing the cloned sequence $d(GA)_n$•$d(TC)_n$.

If triplex formation is important in regulating replication, various types of triplex structures could be involved. Dayn *et al.* (1992) studied the effect of preformed H-r3 and H-r5 isomers of H*-DNA consisting of C•G•G and T•A•T base triads. The site of replication blockage in each case was mapped using chemical probes of the triplex structures. Progress of the DNA polymerase 3' to 5' on one strand of DNA was interrupted either at the start of the H*-DNA structure or in the middle of it, depending on which isomer was present.

Another possibility is that a triplex structure blocking transcription could be created by the act of replication. DNA polymerase carrying out replication on a single strand of DNA has been observed to stall in the middle of a homopyrimidine or homopurine tract (Lapidot *et al.*, 1989; Baran *et al.*, 1991; Samadashwily *et al.*, 1993). When the new DNA strand is synthesized into the homopurine or homopyrimidine tract, one half of the tract may fold back forming a triplex trapping the DNA polymerase (Baran *et al.*, 1991). In replication involving duplex DNA, another possibility is that the nontemplate strand may fold back on itself at the site of the tract downstream of the replication fork and create a block to replication (Samadashwily *et al.*, 1993).

1.5.4.2 Recombination

A number of different types of three-stranded structures have been proposed to mediate RecA protein-dependent homologous recombination (Radding, 1991; Camerini-Otero and Hsieh, 1993; Zhurkin *et al.*, 1994). These triple-stranded structures (R form DNA) differ however, from the triplexes which are the topic of this thesis. Pur•pyr tracts are not required, like strands are aligned in a parallel fashion, and the pattern of hydrogen bonding differs from that in triplexes.

More recently it has been proposed that H-DNA may be involved in recombination. DNA supercoiling appears to have an effect on homologous and site-specific recombination. Transcription activates recombination and also results in negative supercoiling in upstream regions (Thomas and Rothstein, 1989). The negative supercoiling drives the formation of non-B DNA conformations such as H-DNA upstream and such a structure may provide single-stranded entry sites for proteins involved in recombination. The structure may also be directly involved in recombination (Soyfer and Potaman, 1996).

A number of pur•pyr tracts have been identified close to recombination points (Collier at al., 1988). The H-DNA forming sequence $(AGGAG)_{28}$ is located in the murine IgA heavy chain switch region. Formation of H-DNA in this sequence could provide a single strand which could initiate recombination by pairing with an homologous region on the same or different chromosome. Another interesting example is that of unequal sister chromatid exchange involving $(TC)_n$ repeats which results in part of one heavy chain being duplicated on one chromosome and deleted from the other (Weinreb *et al.*, 1990).

One of the models of H-DNA involvement in genetic recombination is a displacement mechanism in which the extruded single-strand invades a duplex region, forming a D-loop. Another involves the interaction of H and H^* isoforms from the same pur•pyr tract (Sinden, 1994).

1.5.4.3 Mutation

Triplex formation may play a role in mutation. Pur•pyr tracts able to form H-DNA are susceptible to deletion. Jaworkski *et al.* (1989) found

through deletion analysis that a long $(GAA)_n \bullet (TTC)_n$ tract potentially able to form an intramolecular triplex was unstable when cloned into the tetracycline resistance gene of *E. coli*.

Several reasons have been proposed for the instability of pur•pyr tracts. First, as described in Section 1.5.4.1, the formation of H-DNA has been shown to block replication. A high mutation frequency is associated with a paused replication fork (Bebenek *et al.*, 1989). Sinden *et al.* (1991) proposed that a stalled DNA polymerase could reiterate pur•pyr tract repeat sequences thus increasing the length of the tract. If the expanded tract were able to generate a hairpin structure, it could be deleted. Another possiblilty is the increased likelihood of slipped alignment of simple repeats such as $(G)_n \bullet (C)_n$ and $(GA)_n \bullet (TC)_n$ during replication, resulting in expansion or deletion of the pur•pyr tract (reviewed in Sinden and Wells, 1992).

1.5.4.4 Telomeres

The ends of eukaryotic chromosomes are stabilized by structures called telomeres which consist of the general motif $(T/A)_m G_n$ reiterated many times (Blackburn and Szostak, 1984; Zakian, 1989). The telomere secondary structure consisting of a duplex and a single-stranded overhang has been shown to form a quadruplex structure called a G quartet in the presence of Na^+ and K^+ (reviewed in Williamson, 1994). However, Veselkov *et al.* (1993) showed that the single-stranded overhang $(T_2G_4)_2$ of *Tetrahymena* could fold back to form a pyr•pur•pur triplex at physiological pH in the presence of Mg^{2+} ions. As triplexes are not usually a substrate for proteins, the formation of such a structure may explain the resistance of telomeres to degradation and recombination.

1.5.4.5 Exon Coding

The presence of pur•pyr tracts in the coding region of a gene may be related to structural requirements of the proteins encoded (Bucher and Yagil, 1991). For example, proteins interacting with DNA or RNA need to be rich in positive charges and hence contain positively-charged amino acids. The positively-charged amino acid lysine is encoded by the all purine codons AAA and AAG, while arginine is encoded by the two purine codons AGA

and AGG as well as by 4 mixed codons. The majority of the SV40 genome encodes the proteins VP1-3, t, and T antigen, all of which interact with DNA. In most of the many pur•pyr tracts present in the SV40 genome, the purine tract is on the coding strand, meaning the proteins encoded are rich in positively-charged amino acids.

N. tobacum chloroplast genes are another example discussed by Bucher and Yagil. Ribosomal proteins are expected to contain positive charges as they interact with RNA. The chloroplast gene containing the longest pur•pyr tract (17 bases), with the purine tract on the coding strand, encodes the basic protein ribosomal S14. In contrast, proteins important in energy metabolism (mainly membrane associated) should contain an excess of hydrophobic amino acids such as leucine (encoded by CTY and YTR, where Y and R stand for any pyrimidine or any purine, respectively) and phenylalanine (encoded by TTT and TTC). In the photosystem, genes PS I and PS II which encode membrane-associated proteins, 14/19 and 12/19, respectively, of pur•pyr tracts $\geq$ 8 bases long are all pyrimidine in the coding strand.

1.5.4.6 RNA Processing

A consensus oligopyrimidine tract near the 3' end of mammalian introns appears to be important in mRNA splicing, as substituting a pyrimidine in the tract can interfere with this process (Ruskin and Green, 1985). The tract, however, has not been found in yeast or higher plants (Goodall and Filipowicz, 1989). Bucher and Yagil (1991) examined 163 genes and found that in many cases the frequency of pur•pyr tracts in introns did not exceed the expected number. Where a long pur•pyr tract was present in an intron it was usually more than 100 bases away from the 3' end of the intron. Bucher and Yagil therefore proposed that the tracts when present in introns, may have roles in addition to splicing.

1.5.4.7 Nucleosomal Spacing

Formation of a non-B DNA structure such as a triplex and involvement in the nucleosome appear to be mutually exclusive possibilities for a pur•pyr tract (reviewed in van Holde and Zlatanova, 1994).

Beasty and Behe (1988) suggested that some pur•pyr tracts may have a role in excluding nucleosomes, as long poly(dA)•poly(dT) tracts do not easily form nucleosomes. Additionally, pur•pyr tracts theoretically bend less readily than mixed sequences and therefore should not be as easily incorporated into nucleosomes (Zhurkin, 1985). However, Birnboim *et al.* (1976) in an analysis of micrococcal nuclease resistant sequences in L cell DNA found the tracts occur with the same frequency in nucleosomal regions as in nucleosome-free regions. More recently, some pur•pyr tract sequences have been shown to have affinity for nucleosomes (Kiyama and Kiyama, 1994; Puhl and Behe, 1995). Brahmachari *et al.* (1995) showed that mammalian $(TC)_n$ sequences form compact nucleosomes.

Raghavan *et al.* (1997) analyzed all 16 chromosomes of the completely sequenced genome of *Saccharomyces cerevisiae* for the distribution of pur•pyr tracts. In noncoding regions, there was a positional preference for tracts to occur within one unit nucleosomal length, i.e. within 200 bp both upstream and downstream, of gene open reading frames. The majority of these tracts occurred upstream of translation start sites. The authors suggested that due to their nucleosomal affinity, these tracts might have a role in modulating the position of nucleosomes. This would have implications in transcription initiation as it could prevent entry of transcription factors.

1.6 Evidence for Triplex DNA *in Vivo*

The overrepresentation of pur•pyr tracts in eukaryotes and their nonrandom distribution suggests a biological role. Several experiments have provided evidence that these tracts can form triplex structures *in vivo*.

1.6.1 Triplex-Specific Antibodies

Two triplex-specific monoclonal antibodies have been produced. Binding studies carried out on eukaryotic chromosomes using these antibodies have provided evidence that triplexes exist *in vivo* in chromosomes. As the two antibodies preferentially bind different sequences, the studies have also shown that triplexes are not randomly distributed in chromosomes. The first antibody, Jel 318, was made by immunizing mice

with the triplex poly[d(Tm^{5}C)]•poly[d(GA)]•poly[d(m^{5}CT)] which was stable at neutral pH (Lee *et al.*, 1987). The triplex was formed from duplex poly[d(Tm^{5}C)]•poly[d(GA)] prepared enzymatically in a reaction containing salts, primer DNA, *E. coli* DNA polymerase I, and deoxyribonucleotides (Lee *et al.*, 1984). Jel 318 shows preference for the sequence it was immunized with, poly[d(Tm^{5}C)]•poly[d(GA)]•poly[d(m^{5}CT)]. It also binds poly[d(T)]•poly[d(A)]•poly[d(T)]. By many criteria the antibody was shown to be specific for T•A•T-rich pyr•pur•pyr triplexes and did not bind calf thymus DNA or the duplex poly[d(TG)]•poly[d(CA)] which cannot form a triplex. Nor did it bind the pyr•pur duplex poly[d(TC)]•poly[d(Gm^{6}A)] which cannot form a triplex because the methyl group at position 6 prevents formation of Hoogsteen bonds.

Jel 318 binds to specific areas of fixed, eukaryotic chromosomes (Lee *et al.*, 1987; Burkholder *et al.*, 1988, 1991). Immunofluorescent microscopy was used to assess binding to chromosomes in mouse myeloma cells that had been fixed in methanol/acetic acid. Although only weakly immunogenic compared to a duplex-specific control monoclonal antibody, the antibody bound to chromosomes in metaphase and to interphase nuclei. The binding was removed by addition of triplex, but not by *E. coli* DNA. It also binds to the chromosomes and nuclei of unfixed, permeabilized cells (Burkholder *et al.*, 1991). As low pH favors formation of triplex and acid fixation has been shown to change DNA structure (Hill and Stollar, 1983), cells were also fixed in cold acetone. A weak immunofluorescent staining of the nuclei was still observed (Lee *et al.*, 1987). Finally, mouse chromosomes that had been gently decondensed at low ionic strength and neutral pH still reacted with the antibody, therefore fixation was not necessary for immunofluorescence (Burkholder *et al.*, 1988). One problem, however, is that it is still unknown whether decondensation makes the triplexes in chromatin more accessible to the antibodies, or provides conditions favorable for triplex formation (Stoller, 1992).

Burkholder *et al.* (1991) stained fixed *Chironomus tentans* and *Drosophila* polytene chromosomes with Jel 318. Intense immunofluorescent staining of chromosome areas consisting of condensed chromatin was observed. Interband areas stained less densely, suggesting more triplex DNA was present in condensed chromatin. In *Chironomus*, areas that were decondensed and therefore likely transcriptionally active, i.e.

nuclear organizer regions and Balbiani rings, were not immunofluorescent. These results suggested triplex DNA may be important in organizing areas of condensation in chromosomes.

Recently, the X-ray crystallographic structure of the Fab (antigen-binding fragment) of Jel 318 was solved (Mol *et al.*, 1994) and a model was proposed for how the monoclonal antibody interacts with the triplex DNA. It is suggested that the CDR-H2 residues of the antibody contact bases in the minor groove of the triplex. Jel 318 is isotype IgG2b with κ light chains.

The second triplex-specific antibody, Jel 466, was also prepared by immunizing mice with poly[d(Tm^{5}C)]•poly[d(GA)] (Agazie *et al.*, 1994). This antibody has a different specificity from Jel 318, preferring the triplex form of the duplex, poly[d(TC)]•poly[d(GA)]. It also weakly binds the triplex derived from poly[d(G)]•poly[d(C)], but unlike Jel 318, does not bind poly[d(T)]•poly[d(A)]•poly[d(T)]. No binding was observed to single-stranded DNAs or to duplexes not capable of forming triplexes. Jel 466 has the IgG2a isotype with κ light chains. There is very little sequence homology between the two monoclonal antibodies, although both have a large number of positively-charged amino acids.

Agazie *et al.* (1994) carried out immunofluorescent studies of human and mouse chromosomes using the two antibodies. The results showed that the staining patterns of the antibodies matched their sequence preferences and that triplexes are not randomly distributed in chromosomes. Hoechst 33258 intensely stains mouse C-band chromosome regions (Hilwig and Gropp, 1972). Both Jel 318 and Hoechst 33258 stained C- and G-bands, but not R-bands. C and G bands are A-T-rich whereas R bands are G-C rich. Conversely, Jel 466 stained R bands rather than C and G bands.

Finally, crude cell extracts have been successfully probed for the presence of triplexes using Jel 318 (Lee *et al.*, 1989). As a control, triplexes were not detected in *E. coli* cell extracts. All the above findings strongly suggest that triplex DNA is a naturally-occurring component of eukaryotic chromosomes.

In 1996, Agazie *et al.* used a different procedure to show that Jel 318 and Jel 466 bound to nuclei. Binding studies were carried out using agarose-encapsulated and permeabilized (agarose-encapsulated) MOPC nuclei. Varying concentrations of the duplex-specific monoclonal antibody, Jel 275, and the two triplex specific antibodies were added to agarose-encapsulated

nuclei. Binding was detected using ^{125}I-labeled sheep antimouse IgG as a secondary antibody. Both Jel 318 and Jel 466 bound at significant levels to the nuclei, providing further evidence that triplexes might exist *in vivo*. Competition experiments were performed using the triplex-forming duplex, poly[d(Tm^{5}C)]•poly[d(GA)]. Poly[d(Tm^{5}C)]•poly[d(GA)] inhibited the binding of both triplex-specific antibodies by 80-85%.

1.6.2 Bacteria

Chemical probing has shown that pyr•pur tracts may be present as triplexes within bacterial cells under very specific (nonphysiological) environmental conditions and high levels of DNA supercoiling. Karlovsky *et al.* (1990) incubated *E. coli* cells containing the plasmid pEJ4, which has pur•pyr tracts, in pH 4.5 medium. Probing with osmium tetroxide and bipyridine showed a modification pattern consistent with the formation of H-DNA at pH 4.5 to 5.0. When the superhelical density was increased by incubating the cells in a higher ionic strength medium containing 0.35 M NaCl, the modification pattern became stronger A criticism of these experiments was that the cells were exposed to low pH and the chemicals osmium tetroxide and bipyridine for 30 min, which could affect viability. *E. coli* cells maintain a normal intracellular pH of close to 7.6 for short intervals during incubation at pH 5.5. However, they cannot maintain such homeostasis during longer incubation intervals (Slonczewski *et al.*, 1981). Therefore in the above experiments the intracellular pH may have approached 4.5 with the possibility of cells remaining viable.

Kohwi *et al.* (1992) used chloroacetaldehyde to probe triplex formation in a plasmid containing a pyr•pur tract within cells grown at neutral pH. Mg^{2+} ions and chloramphenicol, which promotes supercoiling, were used to create conditions favorable for the formation of H*-DNA. A pattern of modification was observed *in situ* which indicated the formation of H*-DNA. A similar pattern indicating H*-DNA was observed in the sequence *in vitro*.

Ussery and Sinden (1993) used trimethylpsoralen in a photochemical assay to detect formation of H-DNA (H-y3 conformer) within cells grown at pH 5 but not at pH 8. The trimethylpsoralen bound at a lower rate to A•T nucleotides between a pair of pur•pyr tracts than to A•T outside the triplex

forming region. This was interpreted to mean the DNA was single-stranded at the A•T nucleotides between the two pur•pyr tracts, suggesting H-DNA formation. The use of mutant *E. coli* cells deficient in topoisomerase I resulted in a higher level of supercoiling which also promoted triplex formation within the cells.

The *E. coli* enzyme *Dam* methylase methylates the sequence GATC in double-stranded B DNA but not in other DNA conformations. This sequence was undermethylated compared to other *Dam* sites in the same plasmid when located in the middle of or close to a pur•pyr mirror repeat in a plasmid grown in *E. coli* JM 101 (Parniewski *et al.* 1990; Klysik, 1996). Undermethylation at the same site was also detected in supercoiled plasmid molecules *in vitro*. However, the sequence was not undermethylated in non-JM strains of *E. coli*. Mutational analysis and chemical footprinting showed that the undermethylation was related to the ability of the tract to form H-DNA (Klysik, 1996).

1.6.3 Factors Promoting/Stabilizing Formation

If triplexes form in the cell and play a significant biological role, stabilizing factors must be present. A number of factors were discussed in Section 1.4.2 which have been shown on their own to induce or stabilize triplexes, particularly H-DNA, *in vitro*. The following discussion focuses on how these factors might promote triplex formation *in vivo*. It may be that the factors act in concert to produce a local environment favoring triplex extrusion.

The first factor to consider is supercoiling. As a result of various cellular processes, nuclear DNA exists in a state of superhelical tension (Cook and Brazell, 1975; Giaever and Wang, 1988). During transcription, waves of negative supercoiling ($Lk<Lk_0$) develop behind RNA polymerase and positive waves ($Lk>Lk_0$) develop ahead (Liu and Wang, 1987; Wu *et al.*, 1988; Frank-Kamenetskii, 1989; Tsao *et al.*, 1989; Bowater *et al.*, 1994). As mentioned in Section 1.4.2, triplex formation *in vivo* might take up superhelical tension and serve to regulate superhelical density (Wang and Lynch, 1993). In the case of Z DNA, it has been shown that superhelical tension generated by transcription is taken up by the formation of Z DNA and that the level of Z DNA is regulated by the degree of superhelical

tension (Wittig *et al.*, 1989, 1991, 1992). Other cellular processes which might create local superhelical tension and thus favor the formation of triplexes are the binding of proteins such as transcription factors, bringing together two distant strands of DNA via protein binding, topoisomerase gyrase activity, and the acetylation of histones (reviewed in van Holde and Zlatanova, 1994).

Another factor is the screening of the negatively-charged phosphate backbones in order to bring the three strands of DNA together to form a triplex. Cellular concentrations of monovalent cations such as Na^+ and K^+ are inadequate for triplex stabilization as these ions are only present at about 0.1 to 0.2 M (Darnell *et al.*, 1986). When triplex formation involves neutral bases, concentrations of up to 1 M Na^+, Li^+, or K^+ are necessary as counterions to shield the repulsion of the phosphate groups (Felsenfeld and Rich, 1957; Felsenfeld *et al.*, 1957; Felsenfeld and Miles, 1967; Michelson *et al.*, 1967). When triplex formation involves protonated cytosines, increasing concentrations of monovalent cations beyond a certain level actually have a decreased stabilizing effect (Wilson *et al.*, 1994a,b). In fact, increasing levels of monovalent cations may destabilize triplexes formed in the presence of multivalent cations (Lee *et al.*, 1984; Latimer *et al.*, 1989; Singleton and Dervan, 1993; Hampel *et al.*, 1993).

Only millimolar concentrations of divalent cations such as Mg^{2+}, Ca^{2+}, Zn^{2+}, or Mn^{2+} are required for triplex formation (Felsenfeld and Rich, 1957; Felsenfeld *et al.*, 1957). However, the concentration of intracellular Mg^{2+} is only about 1 mM (Darnell *et al.*, 1986). Such a low concentration may only partially stabilize pyr•pur•pyr triplexes and C•G•G base triad pyr•pur•pur triplexes (Felsenfeld and Rich, 1957; Kohwi and Kohwi-Shigematsu, 1988; Lyamichev *et al.*, 1991; Soyfer *et al.*, 1992; Singleton and Dervan, 1993). It is not sufficient to stabilize a mixed C•G•G and T•A•A pyr•pur•pur triplex (Malkov *et al.*, 1993). The high concentration of Zn^{2+} required to stabilize T•A•A triad type triplexes is present in the cell. However, it is inaccessible as it is bound within proteins at the active centre of metalloproteins (Soyfer and Potaman, 1996).

As mentioned previously (Section 1.4.2), the stabilizing effect of polyamines on triplexes depends both on structure and charge (Tabor, 1962; Glaser and Gabbay, 1968). The concentrations of the naturally-occurring polyamines putrescine, spermidine, and spermine in the nucleus may be as

high as 5 mM (Sarhan and Seiler, 1989). However, they are largely bound in macromolecules (Davis *et al.*, 1992). Spermine and spermidine are present free in concentrations up to one millimolar (Mach *et al.*, 1982; Sarhan and Seiler, 1989; Tabor and Tabor, 1984). Spermine (4^+) is the best promoter of both pyr•pur•pyr triplexes (Glaser and Gabbay, 1968; Hampel *et al.*, 1991; Lyamichev *et al.*, 1991; Soyfer *et al.*, 1992; Singleton and Dervan, 1993; Thomas and Thomas, 1993) and pyr•pur•pur triplexes (Beal and Dervan, 1991; Soyfer *et al.*, 1992) in the range 0.1 mM to 5 mM. Hampel *et al.* (1991) found that pyr•pur•pyr triplexes could be formed at neutral pH in the presence of micromolar (physiological) concentrations of polyamines. It seems likely that the concentration of free polyamines in the eukaryotic nucleus may be sufficient to provide significant triplex stabilization.

Perfect mirror repeat pur•pyr tracts are relatively rare, and some nonstandard base triads are considered possible in the formation of H- and H*-DNA (reviewed in Soyfer and Potaman, 1996). A•T•G and C•G•T triads containing single Hoogsteen bonds can be stably incorporated into pyr•pur•pyr triplexes. Pyr•pur•pur triplexes have been shown to accomodate both G•C•T and and the protonated $C•G•A^+$ base triads, which requires an acidic pH for formation.

RNA has been proposed as another stabilizing factor. Newly synthesized RNA may bind to single-stranded or duplex DNA near an active RNA polymerase (Reaban and Griffin, 1990; Reaban *et al.*, 1994) and stabilize a non-B DNA structure formed as a result of superhelical tension. The mRNA may bind the single extruded strand of an H-DNA structure and prevent it reverting to the duplex form. In support of this, Belotserkovskii *et al.* (1992) were able to stabilize H-DNA at physiological pH with an oligonucleotide.

Lee *et al.* (1984) suggested that a triplex structure formed between two distant pur•pyr tracts could be stabilized by proteins binding to the single extruded strand. A number of proteins which bind either to the purine single strand or the pyrimidine strand of a pur•pyr tract have been identified (Section 1.5.3.1). One role that has been proposed for these proteins is in stabilizing the triplex structure. These proteins may be able to bind a single strand of DNA as the strands "breathe" or torsional stress causes DNA denaturation. When one strand is stabilized, the other strand may spontaneously bind to the target pur•pyr tract and/or the triplex DNA

structure may be further stabilized or induced by other factors (Kolluri *et al.*, 1992; Aharoni *et al.*, 1993).

There is some evidence, however, that in pairing with single DNA strands, single-stranded DNA binding proteins may in fact destabilize triplexes. *E. coli* single strand binding (SSB) protein facilitates recombination by removing secondary structures including folded single-stranded regions from DNA (Muniyappa *et al.*, 1984). It also facilitates transcription (Baran *et al.*, 1991). It stabilizes locally unwound regions in superhelical DNA and increases the size of the denaturation bubble (Glikin *et al.*, 1983; Langowski *et al.*, 1985). However, Klysik and Shimizu (1993) determined that interaction of *E. coli* SSB protein with the single-stranded loop of H-DNA in supercoiled plasmid DNA resulted in unpairing in the triple-standed region, thus destabilizing the triplex.

It is also unknown what length of single-stranded DNA is adequate for the binding of single-stranded DNA binding proteins. Furthermore, the influence of duplex DNA in proximity to the single-stranded region is unknown (Soyfer and Potaman, 1996). In general, binding studies on the above proteins binding either purine or pyrimidine strands of pur•pyr tracts were carried out using oligonucleotides of at least 30 nucleotides in length. In the study of Aharoni *et al.* (1993), reducing the oligonucleotide length from 30 nucleotides to six resulted in a four-fold decrease in binding. The *E. coli* SSB protein functional complex, consisting of four, 20 kilodalton (kD) subunits, covers approximately 70 nucleotides of single-stranded DNA (Krauss *et al.*, 1981). As mentioned above, when this protein bound to single-stranded DNA in a supercoiled plasmid where H-DNA was present, more duplex DNA was unwound and the triplex disappeared (Klysik and Shimizu, 1993). Many naturally occuring pur•pyr tracts thought capable of forming H-DNA would do so producing single strands only 10 to 20 nucleotides long (Soyfer and Potaman, 1996). Thus, if the binding proteins required longer sites, H-DNA would be depleted rather than stabilized.

There are now three reports of triplex-binding proteins in the literature. All the proteins have been isolated from HeLa cells. Kiyama and Camerini-Otero (1991) used conventional and triplex affinity column chromatography to isolate a 55 kD protein which bound the triplex $(dT)_{34}•(dA)_{34}•(dT)_{34}$ with much higher affinity than duplex or single-stranded DNA. More recently, another 55 kD triplex-specific protein was

isolated which also recognized double pyrimidine triplexes (Guieysse *et al.*, 1997). More research is needed to determine if these are the same proteins. Most recently, triplex DNA-binding proteins were identified which recognized double purine triplex motifs (Musso *et al.*, 1998). Although a role can be proposed for these proteins in stabilizing the triplex conformation, there is no evidence as yet that such is the case.

1.7 Transcription

A major objective of the research presented in this thesis was to investigate a possible role of triplex DNA in the regulation of transcription. This question was investigated from several different angles. Accordingly, the remainder of the Introduction reviews individual topics to put the experiments carried out into context.

1.7.1 Prokaryotes

Transcription is the process of synthesizing a single-stranded RNA molecule from a locus of double-stranded DNA. The two strands of DNA being transcribed have different names. The template strand is the template upon which RNA is polymerized, with nucleotides inserted complementary to the template sequence. The coding strand is the one cited 5' to 3' in published sequences. The nucleotide sequence of the RNA transcript is essentially identical to the coding strand, with substitution of the base uracil for thymine in DNA.

Transcription in prokaryotes is reviewed in Lehninger *et al.* (1993). Proteins bind to DNA sites near the promoter and regulate transcription by either activating transcription (facilitating the binding of RNA polymerase) or repressing transcription by interfering with the activity of RNA polymerase. The single *E. coli* RNA polymerase is a large enzyme (390,000) consisting of 5 core subunits. The enzyme also contains a σ factor (70,000), also called σ^{70}. Its function is to direct the enzyme to promoter sequences. Together, the core and the σ factor constitute the holoenzyme. Two important consensus sequences are located at approximately position -35 and -10 in the bacterial promoter. The -10 sequence is also called the Pribnow Box. Footprinting studies have shown *E. coli RNA* polymerase covers

between 60 and 90 bp of the promoter, including the -10 and -35 regions (Ozoline and Tsyganov, 1995).

There are three stages in the initiation of transcription: closed complex, open complex, and initial transcribing complex in which short RNA strands are released prior to the polymerase complex leaving the promoter region. After binding to the DNA, RNA polymerase migrates to the promoter and binds to the -35 sequence, forming what is known as the closed complex. The polymerase then moves to the -10 region and the DNA duplex unwinds about 17 bp to form a transcription bubble, or open complex and expose the initiation site. The RNA polymerase binds more tightly in the open complex. The initiation of transcription does not require a primer. However, as mentioned previously, binding of RNA polymerase is facilitated by supercoiling. The DNA template nucleotide corresponding to the first nucleotide of the RNA transcript is designated +1 by convention. During transcription, the polymerase continues to keep about 17 bp unwound. As the DNA is unwound ahead of the polymerase complex, and rewound behind, rotation of the duplex is required, generating the postive and negative supercoiling previously discussed.

The σ factor is required only for the recognition of the promoter, and falls off as the polymerase core continues transcription beyond the promoter. Transcription continues until a termination signal, or terminator, is encountered. In *E. coli*, there are at least two classes of terminators, rho-dependent and rho-independent. Apart from the normal σ^{70}, there are other σ factors which recognize the promoters of genes important under certain physiological conditions. An example is σ^{32}, so-called because it has a molecular weight (MW) of 32,000. σ^{32} directs the RNA polymerase to heat shock promoters when the cell is under environmental stress.

1.7.2 Eukaryotes

There are three RNA polymerases in eukaryotes (Lehninger *et al.* (1993). Each is responsible for transcribing a different class of genes. RNA polymerase I transcribes only one class of RNA, 45S RNA, which is the precursor to 28S, 18S, 5.8S ribosomal RNA (rRNA). RNA polymerase II (Pol II) is a 12-subunit complex that is responsible for transcribing nuclear genes encoding messenger RNAs (mRNAs) and a few small nuclear RNAs

(Sentenac, 1985). RNA polymerase III transcribes transfer RNAs (tRNAs), 5S ribosomal RNA (rRNA), and the class of molecules known as small nuclear RNA.

Chromatin proteins such as histones and the HMG proteins package genes into nucleosomal arrays in a transcriptionally repressed, inactive state (reviewed in Shilatifard, 1998). Promoter and enhancer sequences are inaccessible to Pol II and basal initiation factors. Prior to the initiation of transcription, genes are activated by such proteins as histone acetyltransferases and chromatin remodelling complexes. This relieves nucleosomal binding in the vicinity of the core promoter so that Pol II, transactivators, and basal initiation factors can bind.

Pol II promoter sequences contain combinations of core (basal) promoter elements, proximal elements, and distal enhancer elements (reviewed in Nikolov and Burley, 1997). The minimal TATA-containing promoter which Pol II recognizes consists of a TATA box of consensus sequence TATAAA located about -30, and sometimes a pyrimidine-rich initiator (Inr) element at the site of transcription initiation (Fassler and Gussin, 1996). A CAAT box (about -75) is found in some but not all promoters. The multisubunit complex TFIID mediates recognition of the minimal promoter. TFIID consists of TATA-binding protein (TBP) and several TBP-associated factors (TAFs). Association of TBP with the promoter is regulated by transcriptional activators. The binding of TFIID is followed by the other general initiation factor TFIIB, then TFIIF and Pol II, followed by TFIIE and TFIIH to form a large preinitiation complex. The preinitiation complex is functionally equivalent to the *E. coli* holoenzyme. Promoter melting at the transcription start site is facilitated by the energy of negative supercoiling and by the ATP-dependent DNA helicase subunit of TFIIH.

Once the preinitiation complex is assembled, in the presence of nucleoside triphosphates, the strands separate to form an open complex at the site of transcription initiation. A long polypeptide tail extending from the C-terminal domain of the large subunit of Pol II is phosphorylated, presumably by serine/threonine kinase activity of TFIIH. Pol II then initiates transcription and is released from the promoter. During elongation *in vitro*, TFIID remains associated with the core promoter, supporting reinitiation of transcription by Pol II.

In TATA-less Pol II promoters, an Inr element is functionally analogous to the TATA box and mediates assembly of Pol II transcription complexes (reviewed in Smale, 1997). The sequences of Inr elements are very variable. In the consensus Inr sequence an A at +1, a T or A at +3, and a pyrimidine at -1 are the critical requirements in determining the strength of the Inr element. At least some of the positions surrounding the core nucleotides need to be filled by pyrimidines; in fact, activity is enhanced when pyrimidines are present in all the positions. Transcription complexes formed at Inr elements contain the same subunits as in complexes formed at TATA-containing promoters, including the TBP. The specific recognition of Inr elements by proteins is poorly understood; some candidate Inr binding proteins which include some TAFs have been identified, and Pol II is thought to bind weakly to the sequence in some cases. Some TATA-less promoters initiate transcription from a single nucleotide whilst others initiate transcription from multiple start sites ranging either from a few located in a cluster to dozens ranging over hundreds of bp. As in TATA-containing promoters, the Inr element directs initiation of transcription by Pol II when stimulated by an upstream activator such as Sp1.

Transcriptional activators bind to proximal promoter elements located from -50 to -200 and regulate transcription by interacting with the basal transcription complex. The location and orientation of enhancer sequences vary from one gene to the next (Fassler and Gussin, 1996). Such sequences can be located within a gene or can be hundreds or even thousands of bp away from the start site of transcription. The mechanisms by which the effects are exerted are obscure. Activation of promoters appears to be mediated by protein-protein interactions between activators, perhaps intermediary adaptors (cofactors), and components of the transcription machinery.

Approximately 10^{-5} µg of RNA is present in the typical mammalian cell (Farrell, 1993). Eighty to eighty-five percent of the RNA is found in the ribosomes, mostly as 28S and 18S rRNA, and is mainly transcribed by RNA polymerases I and III. Pol II transcription products amount to 20-40% of the RNA (depending on cell type and state). However, mRNA accounts for only 1-4% of the total RNA in the cell as about 95% of all RNA transcribed by Pol II is degraded in the nucleus.

1.7.3 Regulation by Triplex-Forming Oligonucleotides

Two basic types of triplex DNA structures exist *in vitro*. Intramolecular triplexes have already been described in Section 1.4. The other type of triplex is the intermolecular triplex which forms between a duplex pur•pyr tract and an added oligonucleotide (triplex-forming oligonucleotides, TFOs) with a sequence complementary to the purine strand of the tract. The incoming DNA strand forms the third strand of the triplex and can be either mostly purine or mostly pyrimidine. As in intramolecular triplexes, the third strand binds to the duplex purine strand via Hoogsteen or reverse Hoogsteen bonds. This section discusses intermolecular triplexes and their potential use in gene therapy (reviewed in Soyfer and Potaman, 1996).

As in intramolecular triplexes, there are two classes of intermolecular triplex, pyr•pur•pyr and pyr•pur•pur, which are composed of the same base triads as in intramolecular triplexes. The same factors are important in the formation and stabilization of intermolecular triplexes as in intramolecular triplexes, with the exception of supercoiling and the requirement for mirror repeat symmetry in the pur•pyr tract.

In a landmark study, Morgan and Wells (1968) found that a homopyrimidine RNA polymer formed a triplex with a duplex DNA template and prevented transcription. Triplex forming oligonucleotides, which bind to sequences upstream of gene coding regions, are envisioned in having potential in the regulation of gene expression (Helene, 1991). The wealth of information currently available on triple helixes suggests a strong potential for therapeutic use to suppress transcription of disease-related genes. Therapeutic uses are feasible as it has been shown that oligonucleotides may enter the cell via liposomes, receptor-mediated endocytosis, or in some cases are taken up directly and survive in the cell for several hours (Postel *et al.*, 1991)

Several mechanisms have been proposed for the "antigene" action of triplex-forming oligonucleotides. Triplex formation at the promoter has been shown to block the binding of transcription factors such as Sp1 (Maher *et al.*, 1992; Mayfield *et al.*, 1994) and the NFκB (Grigoriev *et al.*, 1992), thus inhibiting transcription. Another possible mechanism is the prevention of transcription elongation (Young *et al.*, 1991).

There are practical considerations in the use of triplex-forming oligonucleotides in antigene therapy. Problems to be overcome include the slow rate of oligonucleotide uptake, the low stability of triplex DNA under physiological conditions, and the sensitivity of oligonucleotides to nucleases (Wagner, 1995). Additionally, in targeting the promoter of a single gene, TFOs can be designed that are sequence specific. To be useful, the sequence specificity of the TFO must be combined with a high binding constant. In this regard, the ability of polyamines to stabilize triplex DNA is of interest. In the search for agents which will better stabilize triplexes, naturally-occurring polyamines have been found amenable to modification (Musso *et al.*, 1997). In addition, it is felt that as the negative charge density of triplexes is higher than duplexes, polyamines will shift the equilibrium to favor triplex formation with TFOs (Thomas and Thomas, 1993). T.J. Thomas and co-workers are investigating the use of polyamine analogues as ligands to stabilize triplex formation using triplex-forming oligonucleotides (Thomas and Thomas, 1993; Thomas *et al.*, 1995).

1.7.4 Agarose-Encapsulated and Permeabilized Nuclei

1.7.4.1 Advantages

The chromatin in agarose-encapsulated nuclei most closely resembles that in the cell as it is isolated under physiological salt concentrations (Jackson and Cook, 1985a; Jackson *et al.*, 1988). When conventional procedures are used to isolate nuclei under physiological salt concentrations (isotonic conditions), the nuclei aggregate and jellify (Ohlenbusch *et al.*, 1967). Nuclei are therefore generally isolated under hypo- or hypertonic conditions in the presence of stabilizing divalent cations. Typically, cells are immersed in hypotonic buffer, then the swollen cells are homogenized. This procedure removes many proteins and the divalent cations activate nucleases which destroy the chromatin by cleaving DNA and removing supercoiling. The non-physiological salt concentrations may also introduce artefacts, as nuclear scaffolds are seen in nuclei prepared under hypertonic conditions, but not in nuclei prepared in hypotonic buffer (Jackson and Cook, 1985b).

A further advantage of agarose-encapsulated nuclei is that the chromatin is protected from shearing and degradation; therefore, the nuclei can be washed and transferred to different buffers (Jackson and Cook, 1985a,b). Agarose-encapsulated nuclei are made as follows. Microbeads form when an aqueous suspension of cells in molten 0.5% agarose is mixed with liquid paraffin. On cooling, the immiscible agarose gels into droplets to form microbeads containing cells. On average, the microbeads are 0.05 mm in diameter. The microbeads are pipetted and transferred into a buffer containing a mild, non-ionic detergent, Triton-X-100, where cell lysis occurs within seconds. The cationic composition of the buffer is the approximate equivalent of that of the cytoplasm (Jackson *et al.*, 1988). Greater than 70% of total cell protein, including most soluble cytoplasmic proteins and RNA, diffuse out through the pores within 10 min. Inner and outer nuclear and cell membranes are mostly removed, leaving chromatin surrounded by insoluble cytoskeletal elements. Some heterochromatin decondensation occurs upon lysis and removal of the nuclear membrane; therefore, the chromatin is unaggregated and less dense than in intact cells. However, it is accessible to probes such as antibodies, as the microbeads contain pores large enough for diffusion of all but the largest macromolecular complexes consisting of chromosomal DNA (Jackson *et al.*, 1988). Proteins up to 1.5 X 10^8 daltons, 1,000 x the size of an IgG antibody molecule (1.5 X 10^5 daltons) (Roitt *et al.*, 1998), can enter and exit. Therefore, antibodies diffuse freely throughout the microbeads in seconds.

The DNA template, RNA polymerase II, and nascent RNA remain associated with the nucleoskeleton in the agarose-encapsulated nuclei. Cell and nuclear lysis interrupt the transcript elongation activity of Pol II on the DNA template, but nascent RNA is not removed and the polymerases retain their ability to transcribe (Jackson and Cook, 1985b). Although the absolute rate of transcription *in vivo* is unknown, transcription in agarose-encapsulated nuclei is thought to occur at rates equal to or greater than *in vivo* (Jackson *et al.*, 1988). Run-off transcription in encapsulated nuclei is more efficient than in nuclei conventionally prepared in hypotonic buffer by homogenization, as that procedure removes half the nascent RNA (Jackson and Cook, 1985a).

1.7.4.2 Triplex-Specific Antibodies and Transcription

Agazie *et al.* (1996) incubated agarose-encapsulated MOPC nuclei with the triplex-specific monoclonal antibodies, Jel 318 and Jel 466. Both antibodies inhibited overall transcription by about 20%. This was measured by trichloroacetic acid (TCA)-precipitable counts from run-off transcription. By comparison, when the nuclei were incubated with the control antibody Jel 42 specific for the bacterial protein, HPr, there was no significant effect on transcription, which was the same as when the nuclei were incubated with bovine serum albumin (BSA).

Run-off transcription is a procedure for localizing and measuring the density of RNA polymerases on transcribed genes (Weber *et al.*, 1977). Isolated nuclei are incubated with radiolabelled nucleotides so that polymerases which paused on the template at the time of cell lysis continue transcription. The resulting labelled RNA is hybridized to immobilized DNA probes to determine genes or regions of genes involved in transcription. The intensity of the hybridization signal is proportional to the density of the polymerases on the genes.

1.8 C-*Myc*

1.8.1 Function of gene

The human proto-oncogene c-*myc* was initially identified due to its homology with the transforming viral gene, v-*myc* (Vennstrom *et al.*, 1982). Although c-*myc* has been intensively studied, its role in the cell is perplexing. It has conflicting functions sometimes being a transcriptional activator and sometimes a repressor. It is central in cell cycle progression, apoptosis, and cellular immortalization. There are two gene products: c-Myc 1 which suppresses cellular growth and c-Myc 2 which stimulates growth (Figure 1.9) (reviewed in Ryan and Birnie, 1996).

1.8.2 Regulation of Transcription

C-*myc* expression is under the control of four promoters. Normally, the major promoters are P1 (75-90% of cytoplasmic c-*myc* mRNAs) and P2

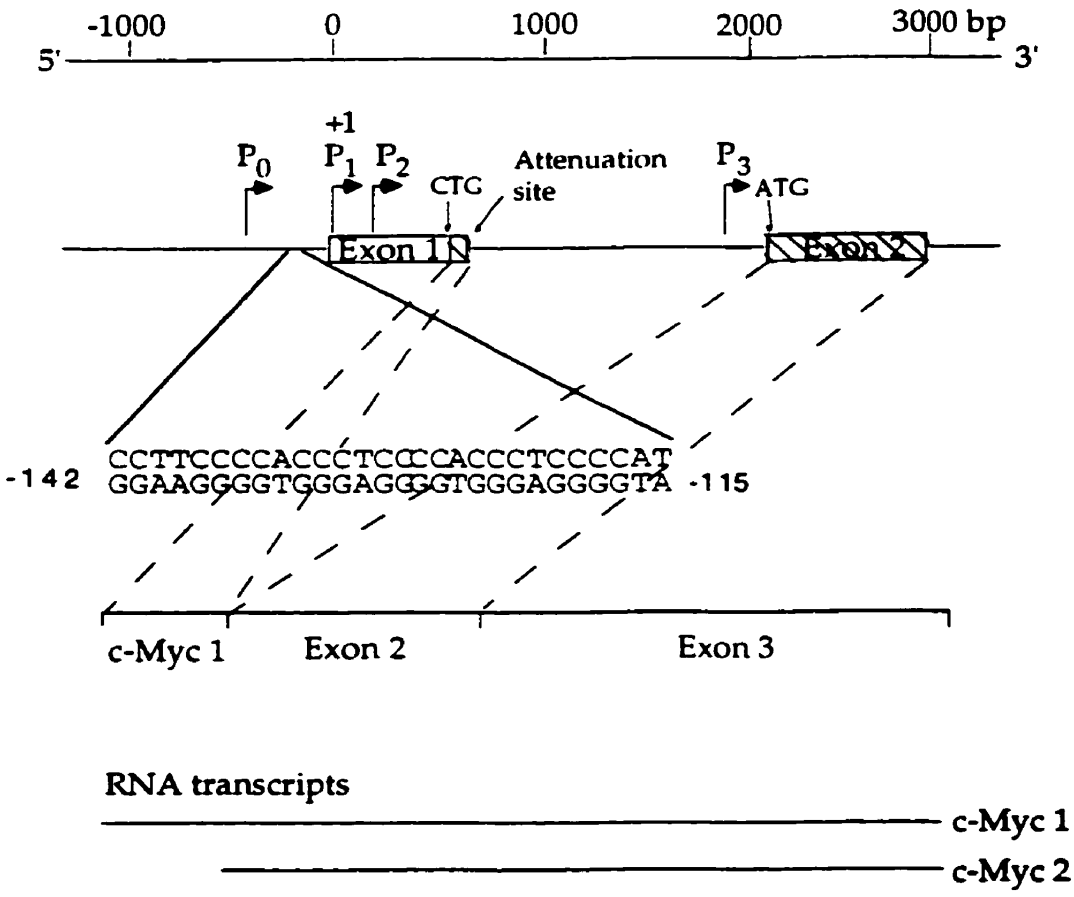


Figure 1.9. The promoters, nuclease sensitive element, and RNA transcripts of the human proto-oncogene *c-myc*.

(10-25% of cytoplasmic c-*myc* mRNAs) (Figure 1.9) (Stewart *et al.*, 1984). Promoter P3 lies approx 1500 bp downstream close to the translation start site and only contributes 5% of c-*myc* mRNAs (Ray and Robert-Lezenes, 1989). The fourth promoter is located about 600 bp upstream from P1 and P2 and is only found in human c-*myc*. It also has only about 5% activity (Bentley and Groudine, 1986b). The usage of the four promoters is variable in different cell lines and in deregulated gene expression (Siebenlist *et al.*, 1988). Bentley and Groudine (1986a) found the initial control of c-*myc* RNA levels was attenuation of transcription. The site of attenuation was a 95 bp section at the 3' end of exon 1, downstream from promoter P2 (Figure 1.9) (Krumm *et al.*, 1992).

1.8.3 Expression in HT-29 Cells

The colon cancer cell line, HT-29, expresses c-*myc* at high levels (Trainer *et al.*, 1988). Expression appears to be deregulated as c-*myc* is expressed under various growth conditions (presence or absence of serum), in various growth phases (exponential and stationary), and in non-differentiated as well as differentiated states due to glucose deprivation. Southern blot analysis of genomic DNA did not reveal any rearrangements in the gene or in upstream or downstream flanking sequences. Therefore, deregulated c-*myc* expression was thought to be due to altered transcription or mRNA processing rather than gene rearrangement.

1.8.4 Nuclease Sensitive Element, H-DNA Model, and Transcription

The human c-*myc* proto-oncogene has pur•pyr region located -142 to -115 in the 5' region (Figure 1.9). This region was initially found as an S1 nuclease sensitive element (NSE) located 125 bp upstream of the start of the P1 promoter (Boles and Hogan, 1987; Kinniburgh, 1989; Postel *et al.*, 1989). Using S1 nuclease mapping to locate regions with altered secondary structure in c-*myc* cloned in supercoiled plasmids, Boles and Hogan (1987) located a conformational isomer 270 bp upstream from the major transcriptional origin, P2. The S1 mapped site consisted of five tandem repeats of the consensus sequence 5'-CCCTCCCC-3'. A second region was located at 1400 bp (53 bp upsteam of P3). This region, designated CT-I_2,

contains a single copy of the repeat sequence in reverse orientation (Desjardins and Hay, 1993).

The c-*myc* NSE was the first promoter element known to be able to form H-DNA (Davis *et al.*, 1989; Firulli *et al.*, 1992). It is a positive regulatory element, accounting for 75-85% of transcriptional activity (determined by point mutations and deletion studies) (Davis *et al.*, 1989; Firulli *et al.*, 1994; Kinniburgh *et al.*, 1994). The NSE has been proposed to form an unusual tandem H-DNA structure containing both H-y3 and H-y5 isoforms (Figure 1.10). Studies of chemical reactivity revealed a large single-stranded region that was only present at pH 4. The requirement for low pH suggested an H-DNA structure rather than a C•G•G triplex.

A number of factors have been identified which bind to one strand or the other or to the duplex form of the c-*myc* NSE. A ribonucleoprotein and protein factor(s) which bound the sequence were found in nuclear extracts from HL-60 cells (Davis *et al.*, 1989). Postel *et al.* (1989) reported a protein called PuF which contacted the GGGTGGG motif. Kolluri *et al.* (1992) cloned and sequenced a gene encoding a protein, NSEP-1, which preferentially bound the pyrimidine-rich strand. The protein also bound the duplex. Interestingly, the protein also binds pur•pyr tracts from other promoters such as c-ki-*ras* and the epidermal growth factor receptor (EGFR). This led the authors to suggest that a structural feature such as the formation of H-DNA was being recognized, rather than primary sequence. Heterogeneous nuclear ribonucleoprotein K (hnRNP K) preferentially bound the pyrimidine-rich strand (Takimoto *et al.*, 1993; Tomonaga and Levens, 1995), while cellular nucleic acid binding protein (CNBP) was found to bind the purine strand (Michelotti *et al.*, 1995). It was suggested that as hnRNP K and CNBP bind opposite strands of the NSE that they co-regulate the element. In 1996, Sakatsume *et al.* reported a zinc finger protein called THZif-1 which bound to the pyrimidine-rich strand. The duplex DNA-binding transcription factor Sp1 and the zinc finger protein ZF87 bind both the tandem sequence elements upstream of P1 and the CT-I_2 element (Desjardins and Hay, 1993).

C-*myc* is a good example of a gene illustrating the possible equilibrium between an H-DNA conformer and the B-DNA conformer of a pur•pyr tract and how factors acting in *trans* might affect the equilibrium. The ability of the NSE to form H-DNA under acidic conditions *in vitro*

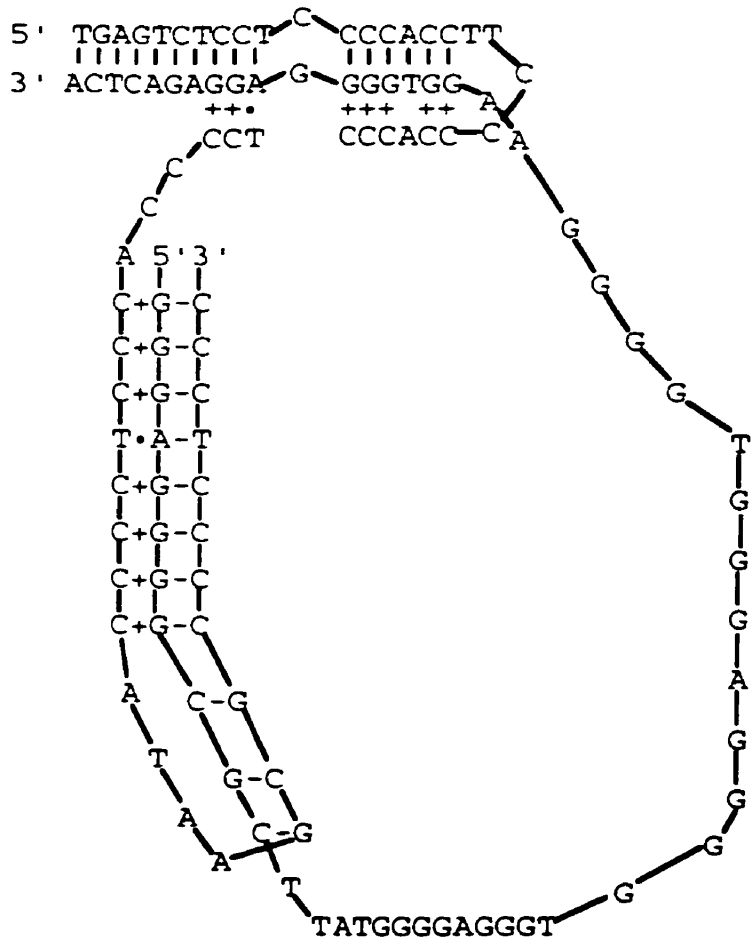


Figure 1.10. **The tandem H-DNA structure proposed for the NSE of c-*myc*** (Firulli *et al.*, 1992; Kinniburgh *et al.*, 1994).

correlated positively with its ability to drive transcription of a reporter gene *in vivo* (Kinniburgh *et al.*, 1994). The authors hypothesized that c-*myc* is turned on when the sequence is in H-DNA conformation and turned off when the sequence returns to B-DNA. Because of the requirement for either low pH or high levels of supercoiling for extrusion of the H-form, the H-DNA structure will not form spontaneously *in vivo*. Kinniburgh *et al.* feel this makes the NSE an ideal structure for an on/off switch for transcription. The authors suggest that some of the factors identified as binding to the region may alter the topology. In binding, such factors could thus reduce the activation energy for formation of H-DNA by either inducing or stabilizing the triplex structure.

1.9 C-*Src*

1.9.1 Function of Gene

The human proto-oncogene c-*src* is a cellular homologue of the oncogene v-*src* of the chicken Rous sarcoma virus (reviewed in Schwartzberg, 1998). C-*src* encodes a non-receptor membrane-bound tyrosine kinase, pp60$^{c\text{-}src}$, which has been extensively studied as a model tyrosine kinase. pp60$^{c\text{-}src}$ is a member of a group of evolutionarily conserved, related enzymes which are widely expressed in a number of tissues. Accordingly, the enzymes are thought to play important roles in cell regulation. pp60$^{c\text{-}src}$ is involved in signal transduction pathways leading to mitogenesis and cellular transformation. The activation or overexpression of this protein is implicated in many cancers, especially colon and breast cancer.

1.9.2 Promoter Region and Polypurine•Polypyrimidine Tracts

The 5' region of the gene was isolated and important regulatory sequences were described by Bonham and Fujita (1993) (Figure 1.11). The promoter region was initially identified through the presence of a CpG island. CpG islands are often found around promoter regions and transcriptional start sites of vertebrate genes, especially housekeeping genes (Lindsay and Bird, 1987). The promoter region is GC rich and contains no

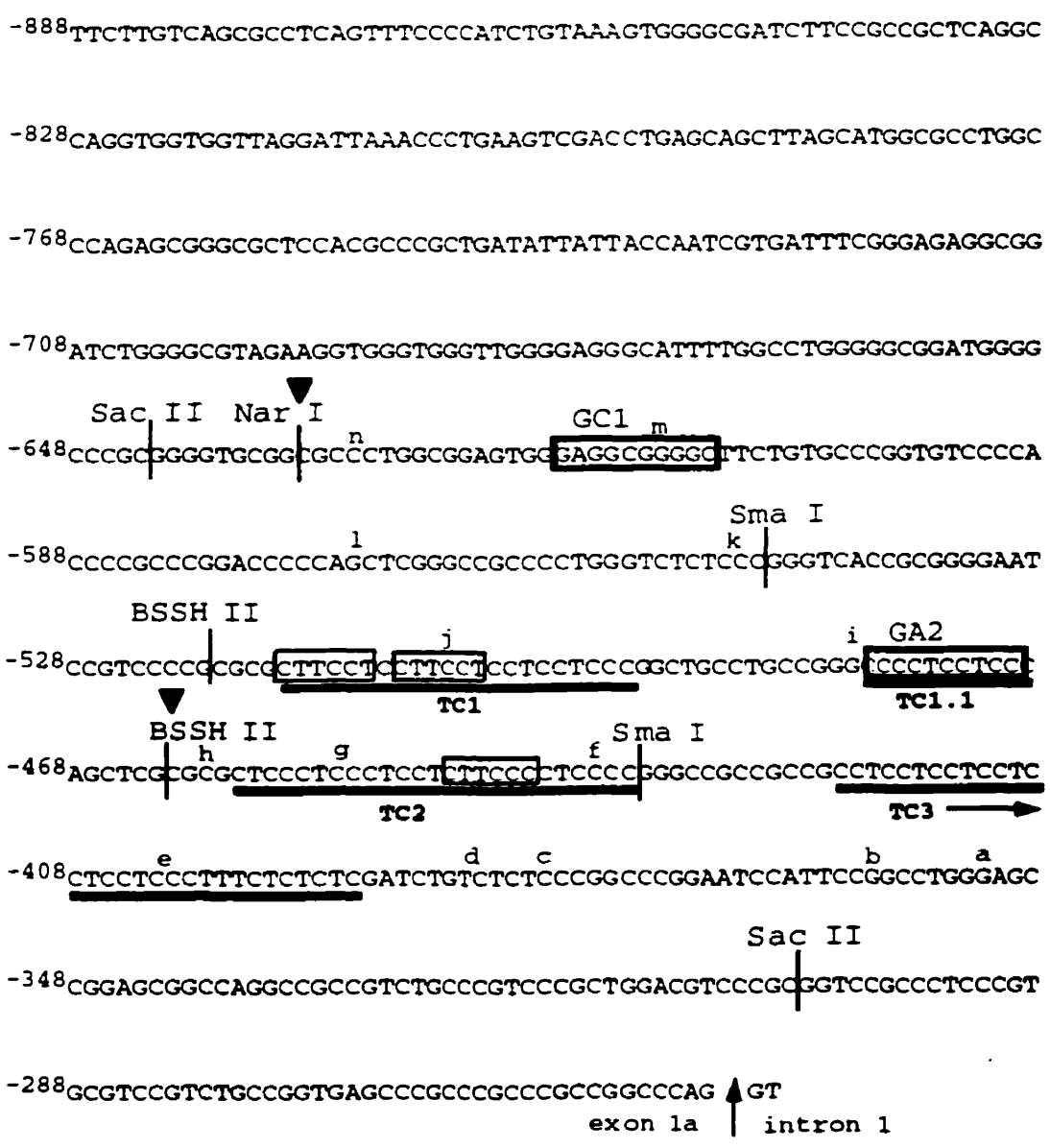


Figure 1.11. Sequence of the 5' flanking region of the human proto-oncogene, c-*src*. The pur•pyr tracts are underlined. The black arrowheads indicate the region necessary for maximal promoter activity; letters above the sequence indicate initiation sites for transcription. Solid boxes - Sp1 binding sites, boxes with dashed lines - SPy binding sites.

TATA or CAAT boxes. C-*src* is therefore a typical housekeeping gene. Conversely, lack of TATA or CAAT boxes is associated with GC rich promoters and multiple initiation sites (Garvin *et al.*, 1988; Gutkind *et al.*, 1991). Other proto-oncogenes such as human *ets*-2 (Mavrothalassitis *et al.*, 1990), human EGFR (Johnson *et al.*, 1988), human *fgr* (Patel *et al.*, 1990), and murine c-Ki-*ras* (Hoffman *et al.*, 1990) also lack TATA and CAAT boxes.

Bonham and Fujita (1993) observed several pur•pyr tracts in the cloned 5' region of the gene. The tract called TC1 was not originally identified due to sequencing misinterpretation, and TC1.1 was recently recognized (K. Bonham, personal communication). The authors mapped an S1 nuclease sensitive site to TC3 which often extended into TC2 in supercoiled plasmids. They also mapped an *in vivo* DNase I hypersensitive site to this region. In recent studies using purine oligonucleotides antiparallel to the TC tract purine strands, Bonham and coworkers showed the formation of intermolecular triplexes in the promoter region of c-*src* (Ritchie and Bonham, 1998).

1.9.3 Regulation of Transcription

Until recent studies by Bonham and coworkers (Bonham and Fujita, 1993; unpublished observations), little was known about the regulation of the gene. The c-*src* promoter was functional in transfection assays with promoter-chloramphenicol acetyl transferase (CAT) gene constructs. CAT assays using deletion constructs in transient transfections of mouse fibroblasts and Human HeLa cells defined a sequence of approximately 180 bp upstream of exon 1a (-634 to -462) with maximal promoter activity (Figure 1.11; Bonham and Fujita, 1993).

S1 nuclease protection assays using c-*src* mRNA from human cell lines revealed the presence of 14 different transcriptional start sites which mapped to a 250 bp region upstream of noncoding exon 1a (Figure 1.11). In all cell lines studied to date, with the exception of testes, the major initiation sites are c, d, e, and f, with f being the most commonly used (K. Bonham, personal communication). Initiation sites in colon cancer cell lines have not yet been studied. Bonham and Fujita (1993) found only a few minor human c-*src* mRNAs not originating from the promoter they described.

C-*src* is activated in many breast and colon cancers (Ritchie and Bonham, 1998). This results in part from the presence of more protein. In recent unpublished observations, Bonham and coworkers found that c-*src* is constitutively overexpressed in some colon cancer cell lines. The high levels of mRNA and protein observed in one colon cancer cell line (HT29), however, were not due to message stability as the mRNA had a relatively short half-life. Instead, overexpression was due to the transcriptional upregulation of *c-src* .

The pur•pyr tracts appear to be a critical *cis* factor regulating transcription. Subcloning and sequence analysis originally showed the 5' flanking region had several consensus Sp1 binding sites (Bonham and Fujita, 1993). Sp1 is a strong transactivator of transcription in *Drosophila* SL2 cells cotransfected with *src* plasmid constructs (K. Bonham-personal communication). A strong Sp1 binding site (GA2) is located in TC1.1 and an additional binding site (GC1) is located upstream of TC1 (Figure 1.11). In transient transfection experiments carried out in mouse, human, and *Drosophila* cells, mutations at both of these sites were necessary to prevent transcription. Therefore, Sp1 appears to be responsible for most of the transcriptional regulation of *c-src*.

Additionally, a cell extract factor called SPy (*src* pyrimidine binding element) has been shown by gel shift analysis to bind very strongly to the pyrimidine strand of TC1 (K. Bonham - personal commun.). Competition with an oligonucleotide corresponding to the pyrimidine strand but not the purine strand of TC1 removed this factor from the TC1 tract. The factor could also be removed using a duplex corresponding to the TC1 tract. However, the competition was very much weaker using the duplex. TC1 contains two binding sites for this factor and TC2 contains one (Figure 1.11). SPy was originally thought to belong to the Ets family of initiator binding factors because the site where it binds in TC1 is Ets-like. However, Ets factors bind exclusively to double-stranded DNA. Furthermore, SPy was not bound with anti-Ets antisera.

1.10 Research Objectives

The preceding literature review has shown that if pur•pyr tracts form triplex DNA within the cell, there are a number of roles this non-B DNA

structure could play. The research presented in this thesis is primarily an investigation into a possible role of triplexes in the regulation of transcription.

The project was mainly approached through the formation of a new type of triplex, the transmolecular triplex. A transmolecular triplex forms between two complementary but widely separated pur•pyr tracts on the same tract of duplex DNA. DNA looping is important both in condensing DNA within chromosomes and in the regulation of transcription. Both events are known to be mediated by proteins. The hypothesis on which the research presented in this thesis is based is that interactions between widely-separated pur•pyr tracts in a sea of duplex DNA, resulting in transmolecular triplex formation, can also mediate such looping. Transmolecular triplexes could therefore have a structural role in the condensation of DNA within chromosomes, as well as a regulatory role in gene expression.

Plasmid models were incubated *in vitro* at low pH in the presence of spermine which promoted the formation of protonated pyr•pur•pyr, or C•G•C$^+$-containing triplexes. Transmolecular triplex formation was initially modelled using a linearized plasmid form containing a pair of separated but complementary pur•pyr tracts. The tracts were designed to be able to form a triplex with each other but not within an individual tract. Because DNA in chromosomes is constrained in loops, plectonemic problems may interfere with the formation of triplexes between separated pur•pyr tracts *in vivo*. Accordingly, a circular form of the same plasmid was used in further experiments to provide a topologically-equivalent model.

Triplex formation in the 5' region of a gene could have the effect of either promoting transcription or inhibiting it. This question was investigated using as a model system the transmolecular triplex loops (T-loops) formed by incubating the circular plasmid at low pH. Synthetic polyamine analogues have been proposed as triplex-stabilizing ligands for use with TFOs in antigene therapy. Therefore, in other experiments, the efficacy of synthetic polyamine analogues in promoting triplex formation was determined using their ability to promote the formation of T-loops.

From plasmid modelling *in vitro*, the research turned to triplex formation *in vivo*. Incubation with triplex-specific antibodies had previously been shown to bring about an overall reduction in transcription in agarose-encapsulated nuclei. The analysis was extended to determine if

incubation with triplex-specific antibodies caused a decrease in expression of a particular gene. The human proto-oncogene c-*myc* has a pur•pyr tract in the 5' region which may form a triplex important in the regulation of transcription. C-*myc* is expressed at high levels in HT-29 cells. Therefore, agarose-encapsulated HT-29 nuclei were incubated with triplex-specific antibodies. RNA was isolated following run-off transcription to determine whether there was a change in expression of c-*myc*.

Finally, the 5' region of the proto-oncogene c-*src* contains four consecutive pur•pyr tracts which might come together *in vivo* to form a triplex involved in the regulation of transcription. In an *in vitro* analysis, each of three tracts was obtained alone in a linear plasmid fragment of unique length. Linear fragments were incubated together in pairs. Triplex formation was monitored as dimerization between two different fragments resulting in a band shift on gel electrophoresis. Based on the results, a model was proposed for how a transmolecular triplex might be formed in the c-*src* promoter *in vivo*. Such a triplex might be involved in the regulation of transcription.

2.0 MATERIALS AND METHODS

2.1 Chemicals, Biological Reagents, Radioactive Isotopes, Equipment, and Supplies

Table 2.1 contains a list of chemical and biological reagents, radioisotopes, equipment, and supplies used in this research. A list of companies and their addresses is found in Table 2.2.

2.2 Cloning in *Escherichia coli*

2.2.1 Bacterial Growth Media

Bacterial growth media was made up according to formulations in Sambrook *et al.* (1989).

2.2.2 Transformation

Two millilitres of Luria-Bertani (LB) broth was inoculated with a single colony of an *E. coli* strain from an LB plate. LB broth contained 10 g bacto-tryptone, 5 g bacto-yeast extract, and 10 g NaCl in a litre of distilled, deionized water (ddH_2O) with the pH adjusted to 7 with NaOH. Plates were made with 15 g bacto-agar/litre. Broth culture was grown up for at least 6 h (or overnight) in a shaker incubator at 300 revolutions per min (RPM) (37 ^oC). Five hundred microlitres of culture were added to 20 ml of LB broth in a 200 ml Erlenmeyer flask. This was again incubated with shaking until early log growth phase (about 0.3 A_{600}). Ten millilitres of the resulting culture was centrifuged (3,000 x g) for 5 min at 4 ^oC. The supernatant was discarded and the cell pellet was gently resuspended in 5 ml of ice cold 100 mM $CaCl_2$ and recentrifuged. The cell pellet was again resuspended in 5 ml of ice cold $CaCl_2$ and left on ice for 0.5 to 2 h. Cells were centrifuged again and resuspended in 0.5 ml of 100 mM $CaCl_2$. One hundred microlitres of the cell suspension was added to plasmid

Table 2.1 Chemical and biological reagents, radioisotopes, equipment, and supplies.

Chemical and biological reagents	
2-Mercaptoethanol	Sigma
2 [N-Morpholino] ethanesulfonic acid (MES)	Sigma
2'-deoxyadenosine 5'-triphosphate (dATP)	Pharmacia
2'-deoxycytidine 5'-triphosphate (dCTP)	Pharmacia
2'-deoxyguanosine 5'-triphosphate (dGTP)	Pharmacia
2'-deoxythymidine 5'-triphosphate (dTTP)	Pharmacia
3-(N-morpholino)propanesulfonic acid (MOPS)	GIBCO/BRL
5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal)	Promega
Acetic acid	BDH
Actinomycin D	Sigma
Adenosine 5'-triphosphate (ATP)	Pharmacia
Agarose (low melting point)	BIO-RAD
Agarose (Ultra Pure)	GIBCO/BRL
Agarose Type VII: low gelling temperature	Sigma
Ammonium sulphate	BDH
Ampicillin	Sigma
Bacto-agar	Fisher
Bacto-tryptone	Fisher
Bacto-yeast extract	Fisher
Bovine serum albumin (BSA)	GIBCO/BRL, Sigma
Bovine serum albumin (BSA), Fraction V	GIBCO/BRL, Sigma
Bromophenol blue	BIO-RAD
Calcium chloride	BDH
Cesium chloride	Boehringer Mannheim
Chloroform	BDH
Cytidine 5'-triphosphate (CTP)	Pharmacia
Diethyl pyrocarbonate (DEPC)	BDH
Dimethylsulfoxide (DMSO)	Sigma

Dipotassium hydrogen orthophosphate	BDH
Disodium hydrogen orthophosphate	BDH
Dithiothreitol (DTT)	Sigma
DNA (from calf thymus)	Sigma
DNA ladder [100 base pairs (bp)]	Pharmacia
DNA molecular weight markers	Boehringer Mannheim, Promega, Gibco/BRL
DNA polymerase I	Pharmacia
DNase I	Pharmacia
Dulbecco's Modified Eagle Medium (D-MEM), high glucose with L-glutamine	GIBCO/BRL
E. coli RNA polymerase core enzyme	Epicentre Technologies
E. coli RNA polymerase holoenzyme	Epicentre Technologies
Ethidium bromide (EtBr)	Sigma
Ethylenediaminetetraacetic acid (EDTA)	BDH, GIBCO/BRL
Fetal bovine serum	Sigma
Formaldehyde solution	BDH
Formamide	Sigma
Gentamicin sulfate	Novopharm
Glycerol	BDH
Guanidine thiocyanate	BDH
Guanosine 5'-triphosphate (GTP)	Pharmacia
Isoamyl alcohol	Fisher
Isopropanol	BDH
Isopropyl β-D-thiogalactopyranoside (IPTG)	Promega
Klenow fragment of DNA polymerase I, FPLCpure	Pharmacia
Klenow fragment of DNA polymerase I, labelling grade	Boehringer Mannheim
Magnesium chloride	BDH
MEM Non-essential Amino Acids Solution 10 mM 100X	GIBCO/BRL
MEM Sodium Pyruvate Solution 100 mM 100X	GIBCO/BRL
MEM Vitamin solution 100X	GIBCO/BRL

NP-40	Sigma
Oligonucleotides	University Core DNA Services
P1 nuclease	Pharmacia
Paraffin liquid, colourless light	BDH
Phenol (Ultra Pure) or buffer-saturated	GIBCO/BRL
Potassium chloride	BDH
Potassium dihydrogen orthophosphate	BDH
Potassium hydroxide	BDH
Prehybridization/Hybridization solution	GIBCO/BRL
Restriction enzymes	GIBCO/BRL, NEB, Pharmacia, and Promega
RNAguard™	Pharmacia
RNase A	Pharmacia
S-adenosyl methionine	Sigma
Sarcosyl™	Sigma
Sodium acetate	BDH
Sodium chloride	BDH
Sodium dodecyl sulfate (SDS)	Sigma, Pharmacia
Sodium hydroxide	BDH
Sodium pyrophosphate	Sigma
Spermidine-3HCl	Sigma
Spermine-4HCl	Sigma
Sucrose	BDH
T4 DNA ligase	Pharmacia
Topoisomerase I	Gibco/BRL
Tri-sodium citrate	BDH
Trichloroacetic acid (TCA)	Sigma
Tris-[hydroxymethyl] aminomethane (Tris)	Sigma, BDH
Triton X-100	BIO-RAD
Trizol™ reagent	GIBCO/BRL
Trypan blue dye	Sigma
Trypsin-EDTA solution	GIBCO/BRL
Uridine 5'-triphosphate (UTP)	Pharmacia
Xylene cyanol FF	BIO-RAD

Yeast transfer RNA (tRNA)	GIBCO/BRL
Radioisotopes	
Deoxycytidine 5'-[α-^{32}P] triphosphate, triethylammonium salt-[α-^{32}P] dCTP	Amersham
Uridine 5'- [α-^{32}P] triphosphate, tetra(triethylammonium) salt-[α-^{32}P] UTP	NEN
Uridine 5'-(α-^{35}S] thiotriphosphate, triethylammonium salt-[α-^{35}S] UTP	NEN, Amersham
Equipment and Supplies	
Bio-Dot SF Microfiltration Apparatus	BIO-RAD
Centricon-100 concentrators	Amicon
Chroma Spin™-100 columns	Clontech
Dialysis tubing	Fisher
Double-stranded nested deletion kit	Pharmacia
GF/C filter discs	Whatman
Hybridization bags	GIBCO/BRL
NucTrap® probe purification column	Stratagene
^{T7}DNA polymerase sequencing kit	Pharmacia
Tissue culture dishes (disposable, sterile, 100 x 200 mm)	Corning
Wizard Plus Minipreps DNA Purification System	Promega
X-Omat-AR film	Kodak
Zeta Probe GT Genomic blotting Membranes	BIO-RAD

Table 2.2 Companies and addresses.

Company	Address
Amersham	Amersham Canada Ltd., Oakville, ON, Canada
Amicon	Fisher Scientific, Winnipeg, MB, Canada
BDH	British Drug House, Saskatoon, SK, Canada
BIO-RAD	BIO-RAD Laboratories, Mississauga, ON, Canada
Boehringer Mannheim	Boehringer Mannheim Canada, Laval, PQ, Canada
Clontech	CLONTECH Laboratories, Inc., Palo Alto, CA, USA
Epicentre Technologies	InterScience, Markham, ON, Canada
Fisher	Fisher Scientific, Winnipeg, MB, Canada
GIBCO/BRL	Bethesda Research Laboratories, Burlington, ON, Canada
Kodak	Amersham Canada Ltd., Oakville, ON, Canada
NEB	New England Biolabs Inc., Mississauga, ON, Canada
NEN	New England Nuclear, Mandel Scientific Company Ltd., Guelph, ON, Canada
Novopharm	Schaumburg, IL, USA
Pharmacia	Pharmacia Biotech Inc., Baie d'Urfé, PQ, Canada
Promega	Fisher Scientific, Winnipeg, MB, Canada
Sigma	Sigma Chemical Company, St. Louis, MO, USA
Stratagene	Stratagene, La Jolla, CA, USA

University Core DNA Services	University of Calgary, Calgary, AB, Canada
Whatman International Ltd.	Chromatographic Specialties Inc., Brockville, ON, Canada

DNA in microfuge tubes and gently mixed. Cells were left on ice for 0.5 to 2 h, heat-shocked at 42 °C for 2 min, then plated as described in the relevant sections below.

2.2.3 Plasmid Mini-Preparations

Unless otherwise stated, plasmid mini-preparations were by the alkali lysis method of Sambrook *et al.* (1989).

2.2.4 Storage of Transformants

Transformed *E. coli* cells from liquid cultures or plates were stored in small, screw-capped tubes in LB broth plus glycerol at -70 °C as described in Sambrook *et al.* (1989).

2.2.5 Large Scale Plasmid Preparations

Three different methods of preparing large quantities of plasmid were used (described in the relevant sections below). The quality of plasmid preparations (i.e. the presence of contaminating chromosomal DNA and RNA) was assessed by appearance on agarose gel electrophoresis and by determination of the A_{260}/A_{280} ratio of the DNA solution.

2.2.6 Digestion with Restriction Enzymes

Standard procedures were used for restriction digestion of plasmid DNA as described in Sambrook *et al.* (1989) and Ausubel *et al.* (1995).

2.3 Agarose Gel Electrophoresis and Photography

Agarose gel electrophoresis was carried out in gel electrophoresis buffer at pH 8 [40 mM Tris-Acetate (pH 8.0), 20 mM sodium acetate, 0.1 mM ethylenediaminetetraacetic acid (EDTA)] or pH 4 (40 mM sodium acetate, 0.1 mM EDTA) in Pharmacia GNA-100 gel boxes. Occasionally, buffers of other pHs were used and are described where relevant. The percentage of agarose used in the gels is also given as required. Agarose gels and tracking dye

were made up in electrophoresis buffer of the appropriate pH and unless otherwise mentioned, tracking dye contained 0.25% bromophenol blue and 30% glycerol.

Much of the work in this thesis involved electrophoresis of plasmid structures based on triplex formation. Therefore, because these forms were heat sensitive, electrophoresis was carried out at 3.6 V/cm for 4.5 to 10 h in ice. Electrophoresis buffer and gel surface temperatures were monitored with a thermometer. Electrophoresis of linearized plasmid DNA following restriction enzyme digestion was generally carried out at 5.7 V/cm for one to two h. Following electrophoresis, gels were stained with EtBr (2 µg/ml) and photographed under shortwave ultraviolet light (254 or 302 nm) unless otherwise stated. Photographs were labelled using Canvas™ 3.5.3.

2.4 Dimers and Omega Loops

These experiments were carried out in collaboration with Dr. K. J. Hampel.

2.4.1 Plasmid DNA

E. coli strain JM 109 was transformed (Section 2.2.2) with the pUC19-derived plasmid, pKHa3PYL/PUL. Construction of this plasmid is described in Hampel (1994). The plasmid contains a pair of asymmetric pur•pyr tracts, $d(CTTCCT)_{11}$ and $d(TCCTTC)_{12}$, designed to form a triplex with each other, but not within a tract itself. White colonies were selected from LB plates containing ampicillin (50 µg/ml) spread with 40 µl of a 20 mg/ml solution of isopropyl β-D-thiogalactoside (IPTG) and 20 µl of a 50 mg/ml solution of 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal). Large quantities of plasmid were prepared by alkali lysis followed by polyethylene glycol precipitation as described in Sambrook *et al.* (1989). In the final step, DNA was ethanol precipitated with 3M sodium acetate (pH 7.0). The plasmids were linearized at the unique Cla I restriction site. Pure linear DNA was obtained by excision of the appropriate band from 1.5% pH 8 preparative agarose gels. DNA was then electroeluted into 1 inch dialysis tubing in 1/4 X pH 8 gel electrophoresis buffer at 4.2 V/cm for 2 h.

The current was reversed for 1 min prior to transferring dialysate into a 15 ml tube. The DNA (at 37 °C) was then reduced under a stream of air to a volume of 200-300 µl. It was extracted with phenol/chloroform (1:1), and chloroform, then ethanol precipitated with 3M sodium acetate (pH 7.0). The pellet was resuspended at 0.5 µg/µl in Tris/EDTA buffer (T/E) [10 mM Tris-HCl (pH 8.0) and 0.1 mM EDTA (pH 8.0)].

2.4.2 Formation

Dimers and omega (Ω) loops were formed by incubating linear plasmid (1.5 µg) in pH 4 triplex-forming buffer [25 mM sodium acetate (pH 4), 1 mM EDTA (pH 8.0)] and spermine in a total volume of 10 µl for 18 h at room temperature (R.T., 23 °C). Two µl of pH 4 tracking dye was added and the whole sample was run in the lane of a 1.8% pH 4 or pH 8 agarose gel.

2.4.3 pH, Temperature, and Ionic Stability

Dimers and loops were formed as in Section 2.4.2 and the pH of the buffer was then changed to 5, 6, or 8 using Chroma Spin™-100 columns. Formulations of the buffers were: pH 8.0, 50 mM Tris-HCl plus 0.1 mM EDTA, pH 6.0, 56 mM MES-NaOH plus 0.1 mM EDTA, and pH 5.0, 25 mM sodium acetate plus 0.1 mM EDTA. Dimers and loops (1.6 µg) were incubated in the new buffers for 19 h at R.T. with/without 50 µM spermine. In a different experiment, following buffer exchange into T/E, dimers and loops were incubated overnight at 50 °C or on ice. Dimers and loops were also incubated overnight at R.T. following buffer exchange into 62.5 mM sodium acetate (pH 4) plus 0.1 mM EDTA, with/without added 200 mM NaCl. In all cases, 2 µl of pH 8 tracking dye was added to each sample which was then added to the lane of a pH 8, 1.8% agarose gel.

2.4.4 Incubation with Topoisomerase I

Dimers and Ω loops were formed in a larger volume of 100 µl by incubating 16 µg of pKHa3PYL/PUL linearized at the *Cla* I site in pH 4 triplex-forming buffer (Section 2.4.2) and 100 µM spermine-4HCl for 18 h. The buffer was exchanged for STE buffer [100 mM NaCl, 10 mM Tris-HCl

(pH 8) and 1 mM EDTA (pH 8.0)] on a Chroma Spin™-100 column according to the manufacturers instructions. Plasmid (1.6 µg) was then incubated with 1.7 to 7 units of topoisomerase I at 37 °C for 3.75 h in a total volume of 12 µl. One unit of topoisomerase I relaxes 0.5 µg of supercoiled ϕX174 RF in 30 min at 37 °C. Two microlitres of pH 8 tracking dye was added to each sample which was then added to the lane of a 1.8% pH 8 agarose gel. As a control, supercoiled pKHa3PYL/PUL was incubated with 7 units of topoisomerase I.

2.4.5 P1 Nuclease Sensitivity

Dimers and Ω loops were formed as described in Section 2.4.2 and the buffer was exchanged into STE buffer without EDTA using a Chroma Spin™-100 column. Plasmid (1.6 µg) was incubated with 0.005 to 0.5 unit of P1 nuclease diluted in a buffer of 50 mM Tris-HCl (pH 7.5) plus 100 mg of BSA/ml, added in a 1 µl volume. Reactions were started by the addition of 10 mM $MgCl_2$ in a 1 µl volume, and were carried out in a total volume of 12 µl. After a 5 min incubation at R.T., reactions were stopped by the addition of 3 µl of pH 8 tracking dye containing 200 mM EDTA and tubes were placed in ice water. Samples were immediately added to lanes in a 1.8%, pH 8 agarose gel.

2.5 T-Loops

2.5.1 Plasmid DNA

In large scale preparations of the plasmid pKHa3PYL/PUL, the basic protocol for preparation of crude lysate by alkaline hydrolysis was followed (Ausubel *et al.*, 1995). However, the pellet from a 500 ml overnight culture grown in Terrific Broth was divided into 2 tubes and each tube was treated with stated solution volumes. Terrific broth was made by adding 12 g bacto-tryptone, 24 g bacto-yeast extract, and 4 ml of glycerol to 900 ml of ddH_2O. The solution was then sterilized and cooled and 100 ml of a sterile salt solution containing 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4 was added. Final plasmid purification was according to the Polyethylene Glycol Precipitation alternate protocol, with the contents of the 2 tubes being pooled for

extraction with buffered phenol and chloroform/isoamyl alcohol. The procedures in Ausubel *et al.* (1995) were used as they consistently yielded plasmid preparations with no contaminating chromosomal DNA and a high A_{260}/A_{280} ratio.

2.5.2 Formation

Open circular plasmid forms were created by gamma- (γ-) irradiating native pKHa3PYL/PUL (Figure 2.1) and pKHa3PY/PU plasmid preparations for 7 min (21,500 Rads). Plasmid pKHa3PY/PU is the same as pKHa3PYL/PUL except that in the former the number of repeats in the two pur•pyr tracts is six instead of 12. "Native" plasmid is plasmid directly purified from *E. coli*, usually containing some open circular as well as supercoiled forms. T-loops were made by incubating the γ-irradiated plasmid (1.25 µg) for 19 - 22 h at R.T. with 37.5 - 50 µM spermine in pH 4 triplex-forming buffer (Section 2.4.2) in a volume of 10 µl. Two microlitres of tracking dye was added, the whole volume was added to a gel lane, and electrophoresis was carried out in pH 4 or pH 8, 0.85% agarose gels. Larger volumes of T-loops were made by increasing incubation constituents proportionately.

2.5.3 Relaxed Plasmid

Relaxed plasmid was made by incubating 9 µg of native pKHa3PYL/PUL with 100 units of Topoisomerase I in 0.15 M NaCl plus 0.02 M EDTA in a total volume of 50 µl for 2 h at 37 °C. This was followed by exchange into T/E buffer on a Chroma Spin™-TE 100 column. Two microlitres of tracking dye was added to 10 µl aliquots which were added to the lanes of either pH 4 or pH 8 0.85% agarose gels.

2.5.4 Formation at pH 4.5 or 5.0

Plasmids pKHa3PY/PU and pKHa3PYL/PUL (1.25 µg), γ-irradiated for 7 min, were incubated for 19 - 22 h at R.T. with 37.5 - 50 µM spermine in pH 4.5 (25 mM sodium acetate, pH 4.5 and 1 mM EDTA) or 5.0 buffer (25 mM sodium acetate, pH 5.0 and 1 mM EDTA) in total volumes of 10 µl. Two

microlitres of tracking dye was added and electrophoresis was carried out at pH 4.5 or pH 5.0 in 0.85% agarose gels.

2.5.5 Incubation with P1 Nuclease

T-loops were made as described in Section 2.5.2. Incubation was followed by buffer exchange on a Chroma Spin™-STE 100 column. P1 nuclease, diluted in 50 mM Tris HCl (pH 7.5) plus 100 µg BSA/ml, was added to reaction tubes in 1 µl volumes. Digestion reactions were carried out at R.T. and were started by adding 1 µl of 120 mM $MgCl_2$ to T-loop preparations (0.6 µg DNA) in a total volume of 7 µl. The reactions were stopped after 5 min by adding 3 µl of pH 8 tracking dye containing 200 mM EDTA. The total volume of each tube was added to the lane of a pH 8, 0.85% agarose gel.

2.5.6 Temperature and Ionic Strength Stability

T-loops were made as described in Section 2.5.2. The buffer was exchanged to either T/E or T/E with 0.2 M NaCl on a Chroma Spin™-100 column. Ten microlitre aliquots were then incubated for 24 h at -20, 4, 22, 37, and 50 °C. The total volume of each tube was added to the lane of a pH 8, 0.85% agarose gel.

2.5.7 Electron Microscopy

T-loops were made (Section 2.5.2) and added to the well of an 0.85% pH 4 preparative gel. Following electrophoresis at 3.4 V/cm on ice and staining in EtBr, bands containing T-loop structures were excised with a scalpel blade. DNA was extracted from the agarose by centrifugation through a small quantity of glass wool in a microfuge tube. Glass wool and microfuge tubes had been presoaked in pH 4 buffer and the buffer removed by a brief centrifugation. Centrifuge speeds for DNA extraction were 4,000 x g for 8 min followed by 16,000 x g for 10 min.

Electron microscopy was carried out by Roger Bradley in the laboratory of Dr. Douglas G. Scraba, Department of Biochemistry at the University of Alberta, Edmonton, Alta. T-loops in pH 4 gel electrophoresis

buffer were adsorbed to freshly cleaved mica. This was followed by 8° rotary shadowing with 20 Å platinum and replication with 150 Å carbon in a vacuum evaporator (Balzers BA511M) equipped with electron gun evaporator sources and quartz crystal film thickness monitor (Lurz *et al.*, 1986). Carbon replicas were floated on water, mounted on 300-mesh copper grids and examined in an electron microscope (Philips EM420) operated at 100 kV. Photographs were obtained using 35 mm film and the STEM accessory in bright field/dark field mode and a high resolution photomonitor. Measurements of molecules were made using a Hewlett Packard 9874A digitizer linked to a Tektronix 4051 computer.

2.5.8 Incubation with *Escherichia coli* SSB Protein

E. coli SSB protein (a gift of Dr G. Chaconas, Department of Biochemistry, University of Western Ontario, London, Ontario) was received at a concentration of 3.8 µg/µl. Plasmid pKHa3PYL/PUL (0.1 µg/µl) which had been γ-irradiated (Section 2.5.2) was incubated at R.T. with 60 µM spermine-HCl, pH 4 triplex-forming buffer, and 0.38 or 0.76 µg SSB protein (diluted in ddH_2O), in a total volume of 10 µl. Control preparations contained no SSB protein or contained SSB protein dilution buffer [20 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 0.1 mM dithiothreitol (DTT), 0.2 M NaCl, 50% glycerol] diluted in the same manner as SSB protein. Two microlitres of a tracking dye containing 500 mM Tris-HCl (pH 8), 50 mM EDTA, 0.1% sodium dodecyl sulphate (SDS), 0.25% bromophenol blue, 0.25% xylene cyanol FF, and 40% sucrose made up in water was added to each sample which was then stored at -20 °C until gel electrophoresis at pH 8.

In a different experiment, T-loops were made in 10 µl volumes as described in Section 2.5.2. After the incubation period, 0.38 µg SSB protein/µl was added. Control preparations contained no SSB protein or contained SSB protein dilution buffer as previously described. A further incubation was carried out for 22 h at R.T.. Two microlitres of the tracking dye containing 500 mM Tris-HCl was added to each sample which was then stored at -20 °C until gel electrophoresis at pH 8.

2.5.9 Formation using Polyamine Analogues

The following polyamine analogues were synthesized according to procedures described in He *et al.* (1994) and Igarashi *et al.* (1995) and were received from Dr. T.J. Thomas, R.W. Johnson Medical Complex, New Brunswick, N.J., U.S.A: 1,11-diamino-4,8-diazaundecane (norspermine, 3-3-3), N^1,N^{11}-bis(ethyl)norspermine (BE-3-3-3), N^1,N^{12}-bis(ethyl)spermine (BE-3-4-3), 1,14-diamino-5,10-diazatetradecane (homospermine, 4-4-4), N^1N^{14}-bis(ethyl)homospermine (BE-4-4-4), 1,15-diamino-4,8,12-triazapentadecane (3-3-3-3), 1,15-bis(ethylamino)-4,8,12-triazapentadecane (BE-3-3-3-3), 1,19-diamino-5,10,15-triazanonadecane (4-4-4-4), and 1,19-bis(ethylamino)-5,10,15-triazanonadecane (BE -4-4-4-4). T-loops were formed by incubating pKHa3PYL/PUL (1.25 µg; γ-irradiated, Section 2.5.2) for 24 h at R.T. in pH 4 triplex-foming buffer (Section 2.4.2) with individual polyamine analogues in a total volume of 10 µl. After the addition of 2 µl of tracking dye, the whole sample was added to the lane of a pH 8, 0.85% agarose gel.

2.6 Transcription in T-loops

2.6.1 Gel Electrophoresis of T-Loop Preparations

Comparative transcription studies employing *E. coli* RNA polymerase holo- and core enzymes were carried out using isolated T-loop, supercoiled, open circular, and linear plasmid forms of pKHa3PYL/PUL. Therefore, preparative agarose gels were run to isolate each plasmid form in sufficient quantity for use in the experiments. Large quantities of T-loops were typically made by incubating 125 µg of γ-irradiated pKHa3PYL/PUL (Section 2.5.2) in pH 4 triplex-forming buffer and spermine in a total volume of one ml. After this, all solutions used in the isolation of the plasmid forms were either treated with diethyl pyrocarbonate (DEPC) or used DEPC-treated ddH_2O and ribonuclease- (RNase)-free conditions were maintained. Following addition of 200 µl of pH 8 tracking dye, the whole amount of the T-loop incubation was added to four, 0.85% pH 8 preparative agarose gels and electrophoresis was carried out as described in Section 2.3. Gels were then stained in EtBr for about 10 min. T-loop bands were visualized under UV light (302 mn) and quickly excised with a scalpel blade.

Bands were not clearly visible under long wave UV light (365 nm) which would have had less potential to nick the DNA. Bands containing supercoiled and open circular plasmid forms were also excised from the same gels. As the amount of supercoiled plasmid form in these gels was small, additional supercoiled bands were excised from preparative gels of native plasmid DNA. The quantity of linear plasmid form on the T-loop preparative gels was also small, therefore linear plasmid forms were excised from preparative gels of pKHa3PYL/PUL linearized with Cla I. Gel slices containing the plasmid forms were stored at 4 °C for a maximum of 2 days before plasmid DNA was isolated.

2.6.2 Isolation of Plasmid Forms from Gel Bands

Individual gel slices containing plasmid forms were placed in dialysis tubing in 100 to 200 µl of pH 8 gel electrophoresis buffer. The dialysis tubing had been autoclaved in DEPC-treated ddH_2O. The DNA was electroeluted in pH 8 gel electrophoresis buffer in ice in the dark for 2 to 2.5 h at 3.4 V/cm. Completeness of electroelution was determined by quickly viewing gel slices under shortwave UV light. Shortwave UV light was used as any remaining fluorescence in the agarose slice was not visible under longwave UV light. The pH 8 buffer containing the DNA was collected from around the gel slice using a Pasteur pipette, and pooled for each form. The volume of liquid was reduced from about 3 ml to about 50 µl using Centricon-100 concentrators according to the manufacturer's instructions. These devices were used as they enabled a considerable reduction in volume without a comcomitant increase in ionic strength which could have been detrimental to the preservation of the T-loop forms. Prior to use, concentrators were prerinsed with 0.1 N NaOH as suggested by the supplier to remove preservative agents, then rinsed with pH 8 gel electrophoresis buffer. Whenever feasible during the extraction procedures, plasmid forms were kept on ice in the dark. This was to reduce as far as possible the tendancy of isolated T-loop forms to revert to open circular forms, and isolated supercoiled forms to be nicked to become open circular forms.

2.6.3 Determination of Plasmid Form Concentrations

Concentrations of the isolated plasmid forms were estimated using a Hitachi F-2000 Fluorescence Spectrophotometer at an incident wavelength of 525 nm and an emission wavelength of 600 nm. Plasmid forms were added to 2 ml of pH 8 EtBr fluorescence assay buffer containing 5 mM Tris-HCl (pH 8.0), 0.5 mM EDTA, and 0.5 µg EtBr/ml. Plasmid form concentrations were determined by comparison with known concentrations of linear, open circular, or native plasmid controls. Prior to use in transcription assays, the concentration of each isolated plasmid form was adjusted to 0.17 µg/µl using pH 8 gel electrophoresis buffer.

2.6.4 Transcription Assays

Transcription assays were carried out using both *E. coli* RNA polymerase core and holoenzyme. Both core and holoenzyme catalyze the incorporation of one nanomole of radiolabelled ribonucleoside triphosphates into RNA in 10 min at 37 °C. Assay solutions were prepared using DEPC-treated ddH_20 or were treated with DEPC. RNase-free conditions were maintained throughout the transcription assay. Each transcription reaction contained in a volume of 40 µl: 1x assay buffer, 125 µM each of adenosine 5'-triphosphate (ATP), cytidine 5'-triphosphate (CTP), and guanosine 5'-triphosphate (GTP), 0.425 µM uridine 5'-triphosphate (UTP), 0.17 µg of DNA, 0.0425 µM [α-^{35}S] UTP (1.7 µCuries (µCi) at >1000 Ci/mmol), and 5 or 10 units of core or holoenzyme. Once plasmid DNA was added, tubes were kept in the dark as much as possible. Assay buffer was made according to the supplier's product information sheet as a 5x solution (except that Tris-HCl was used at pH 7.6 instead of pH 7.9) and contained 200 mM Tris-HCl, 750 mM KCl, 50 mM $MgCl_2$, 0.5 mM EDTA, and 0.5 mM DTT. The 1x assay buffer for reactions using holoenzyme also contained 2.5 µg BSA/µl. Epicentre storage buffer contained 50 mM Tris-HCl (pH 7.5), 250 mM NaCl, 0.1 mM EDTA, 1 mM DTT, and 50% glycerol. As both core and holoenzymes were supplied in storage buffer, the volume of storage buffer in reactions was adjusted so that all tubes contained an equal volume. Reactions were started by adding enzyme, briefly centrifuging tube contents to the bottom of tubes in a microfuge, and gently

tapping the tubes to mix the contents. Transcription reactions were carried out at 37 °C for 30 min in the dark in a waterbath covered with aluminum foil. The waterbath was used to ensure all tubes were subjected to the same temperature conditions. Reactions were stopped by a 10 min incubation at 75 °C. The tubes were cooled in ice and then briefly spun in a microfuge.

2.6.5 Trichloroacetic Acid Precipitable Counts

A Whatmann GF/C filter disc placed on a holder connected to a vacuum pump was wetted with approximately 2 ml of a wash solution consisting of 5% TCA and 20 mM sodium pyrophosphate using a Pasteur pipette. Vacuum was briefly applied and turned off. One-half of a transcription assay reaction volume was added to the filter, the vacuum was turned on again, and the filter was washed with three Pasteur pipettefuls of the wash solution (about 18 ml) followed by 2 ml of 95% ethanol. The filter was placed in a scintillation vial, dried under a heat lamp, and 10 ml of scintillation fluid was added. Radioactive counts were determined using the ^{35}S channel of a Beckman LS 6500 Scintillation Counter.

Calculations of the concentrations of RNA in the T-loop and open circular assays following transcription were based on the counts obtained and the specific activity of the ^{35}S-UTP radioisotope.

2.7 Triplex-Specific Antibodies and Gene Expression in Agarose-Encapsulated and Permeabilized HT-29 Nuclei

2.7.1 Run-off Transcription Using Hypotonically-Prepared HT-29 Nuclei

The triplex-specific antibodies Jel 318 and Jel 466, and the control antibody, Jel 42, stored in S-200 column buffer [50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.5 M NaCl] at -20 °C, were dialyzed into transcription buffer [5 mM Tris-HCl (pH 7.5), 2.5 mM $MgCl_2$, 150 mM KCl]. Nuclei, prepared using a hypotonic lysis buffer containing NP-40 (Ausubel *et al.*, 1995), were supplied by Dr. K. Bonham, Saskatoon Cancer Centre. Prior to use, nuclei ($\geq 10^7$) were thawed on ice for 15 to 20 min, washed twice in transcription buffer, then resuspended in 200 µl of transcription buffer. Nuclei were

incubated with antibodies (30 µg; 1.53 µg/µl) in microfuge tubes for 3.75 h at R.T. with occasional mixing. Nuclear run-off buffer [2x; 200 µl; 10 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 0.3 M KCl, 5 mM DTT, 1 mM each of CTP, GTP, and ATP] together with RNAguard™ (32 Units) and 0.044 µM [α-^{32}P] UTP (50 µCi at 3,000 Ci/mmole) was added to 15 ml polypropylene tubes. Nuclei and antibodies were added to the tubes and run-off transcription was carried out for 30 min at R.T. DNase I (10 units) was added and the tubes were incubated for 5 min, then yeast tRNA (50 µg) was added as a carrier. RNA was extracted according to the method of Chomczynski and Sacchi (1987). Solution D [600 µl; 4 M guanidine thiocyanate, 25 mM sodium citrate (pH 7.0), 0.5% Sarcosyl™, 0.1 M 2-mercaptoethanol], 2 M sodium acetate (pH 4, 100 µl, 10% vol), water-saturated phenol (1 ml), and chloroform isoamyl alcohol (0.2 ml) were added individually with vortexing after each step. Tubes were placed on ice for 15 min, then centrifuged at 3,000 revolutions per min (RPM) for 5 min in a bench top Sorvall. Supernatants were transferred to new 15 ml tubes, 95% ethanol (2.5 x vol) was added, and tube contents were mixed then left 20 min to overnight at -20 °C.

Tubes were centrifuged at 3,000 RPM for 20 min at 4 °C. The supernatants were poured off and tube insides were wiped with laboratory tissue. The pellets were air-dried and resuspended in 300 µl ddH_2O. They were then reprecipitated for 1 h at -20 °C with 2 M sodium acetate (pH 4; 1/10 volume) and 95% ethanol (2.5 volumes), then warmed briefly to allow free nucleotides to go back into solution. The tubes were spun for 5 min at 3,000 RPM in the Sorvall bench top centrifuge and pellets were washed with 75% ethanol (-20 °C) then resuspended in 300 µl ddH_2O. The pellet precipitation and centrifugation steps were repeated. Pellets were dried after the second wash in 75% cold ethanol, and resuspended in 100 µl of ddH_2O. RNA (1 µl) from each resuspended pellet was counted on a ^{32}P channel in a scintillation counter.

A Zeta Probe membrane was soaked in ddH_2O and placed in a Bio-Dot SF Microfiltration apparatus according to the manufacturer's instructions. Plasmids (kindly provided by Dr. K. Bonham, Saskatoon Cancer Centre, Saskatoon) contained the following cloned genes which were used as probe DNA: c-*src*, ß-actin, c-*myc*, glucose-6-dehydrogenase, and 18s rRNA. A control plasmid, pBS, lacked cloned genomic DNA. Plasmid DNA (estimated excess amount of 35 µg) was denatured by adding 700 µl of

0.4 M NaOH in a microfuge tube and boiling for 5 to 10 min. DNA (5.8 µg; 100 µl) was added to each slot of the apparatus and bound to the membrane by washing through with 500 µl of 0.4 M NaOH/slot under vacuum. The membrane was removed and neutralized by washing once in 3 x SET [0.45 M NaCl, 0.09 M Tris-HCl (pH 8.0), 6 mM EDTA]. It was then dried and cut into individual strips containing DNA from all the plasmids. The strips were placed in hybridization bags and prehybridization was carried out in 2 or 3 ml of Prehybridization/Hybridization solution at 68 °C in a slow shaker bath for 2.5 h. Bags were cut open at one end, adjusted amounts of RNA (containing equal counts) were added to each bag, and the bags were resealed. Hybridization was carried out for 48 h with gentle agitation at 68 °C.

Following hybridization, strips were removed from the hybridization bags and washed in 1 x SET containing 1% SDS, 0.1 x SET containing 1% SDS, and rinsed in ddH_2O. The washes and rinse were carried out at 68 °C with gentle agitation for 20 min in 200 ml solutions prewarmed to 68 °C. The strips were lined up without drying on filter paper, wrapped in plastic wrap, and placed under X-Omat-AR film with an enhancer screen in a film cassette for 24 h at -70 °C.

2.7.2 Transcription Using Agarose-Encapsulated and Permeabilized HT-29 Nuclei

2.7.2.1 Tissue Culture

The human colon carcinoma cell line HT-29 was a gift of Dr. S. Laferté, Department of Biochemistry, University of Saskatchewan. Cells were grown in Dulbecco's Modified Eagle Medium (high glucose with L-glutamine) supplemented with 10% fetal bovine serum, MEM sodium pyruvate solution (1 mM), MEM Non-essential Amino Acids solution (0.1 mM), MEM Vitamin solution, and 48 µg gentamicin sulfate/ml at 37 °C in an atmosphere of 5% CO_2. Cells grew as a monolayer in 100 mm tissue culture dishes and were split 1:7 to 1:9 twice a week while in logarithmic growth phase (subconfluent). The day before splitting, the culture medium was replaced with fresh medium. During splitting, the medium was removed and each plate was washed with about 5 ml of a solution of 1 mM

EDTA in calcium- and magnesium-free phosphate buffered saline (CMF-PBS; 1.5 mM KH_2PO_4, 8.1 mM Na_2HPO_4, 2.7 mM KCl, 137 mM NaCl) at R.T. Trypsin-EDTA solution (1.5 ml; 0.003% trypsin, 0.027 mM EDTA•4Na in CMF-PBS) warmed to 37 °C was then added to the plate which was then replaced in the incubator for about 6 min. A 5 or 10 ml pipette was used to remove the trypsinized cells which were then either split or used in a transcription experiment.

2.7.2.2 Preparation of Agarose-Encapsulated and Permeabilized Nuclei

Agarose-encapsulated HT-29 nuclei were prepared according to standard procedures (Jackson and Cook, 1985a; Jackson *et al.*, 1988; Agazie *et al.*, 1996). Transcription experiments were begun on the day cells were split. Following removal of most trypsinized cells with a pipette, plates were washed with an additional 2 to 5 ml of R.T. CMF-PBS and cells and washings were collected in 50 ml disposable, polypropylene tubes. The tubes were centrifuged at 200 x g for 5 min at R.T., and resulting cell pellets were washed twice in 50 ml of R.T. CMF-PBS. Pellets in all tubes were pooled and gently resuspended in R.T. CMF-PBS. Viable cells were counted using trypan blue dye (0.4% in 0.85% NaCl) and a haemocytometer, resuspended at 2.5 x 10^7/ml in CMF-PBS, and placed in a 39 °C water bath.

Except when the nuclei were to be used in transcription experiments followed by TCA-precipitable counts, the following procedures were carried out under RNase-free conditions and all solutions were either treated with DEPC or used DEPC-treated ddH_2O. Chemicals were purchased new and kept RNase-free. Low gelling Type VII agarose (1/4 cell suspension volume; 2.5% in CMF-PBS), liquified and cooled to 39 °C, and light liquid paraffin (2.5 x cell suspension volume), also warmed to 39 °C, were added to HT-29 cells in a 250 ml Erlenmeyer flask. Flask contents were emulsified at 400 RPM for 30 seconds (sec) at R.T. in a rotary shaker, then equilibrated to 0 °C in an ice water bath for 5 min. Ice-cold CMF-PBS (5 x cell suspension volume) was added and flask contents were immediately mixed manually by shaking the flask, then were quickly distributed into 30 ml glass centrifuge tubes and centrifuged at 3,500 x g at 0 °C for 5 min. The supernatant was removed by aspiration and pellets (10 to 12 ml of microbeads) were washed 3 x with 20 ml of ice-cold CMF-PBS and once in 3

volumes of ice-cold isotonic buffer, with gentle resuspending of the microbeads during each wash. Isotonic buffer (Jackson *et al.*, 1988) contained 130 mM KCl, 10 mM Na_2HPO_4, 1 mM $MgCl_2$, 1 mM ATP, and 1 mM DTT, with the pH adjusted to 7.4 with 100 mM KH_2PO_4 or 5 M KOH. Pellets were then resuspended in 3 volumes of ice-cold isotonic buffer containing 0.5 % Triton X-100 and kept in an ice-water bath for 20 min with gentle stirring. Permeabilized nuclei were collected by centrifugation at 4,800 x g for 5 min and washed 3 x with 5 volumes of ice-cold isotonic buffer, the final wash containing isotonic buffer supplemented with 0.2 % BSA, Fraction V.

2.7.2.3 Incubation with Triplex-Specific Antibodies, Run-Off Transcription, and Trichloroacetic Acid-Precipitable Counts

Antibodies, stored in S-200 column buffer were dialyzed into transcription buffer (above) prior to use. Procedures were as for MOPC nuclei (Agazie *et al.*, 1996). Agarose-encapsulated HT-29 nuclei were washed once with 5 volumes of ice-cold transcription buffer [10 mM Tris-HCl (pH 7.6), 50 mM KCl, 2 mM EDTA, 5 mM spermidine, 100 mM $(NH_4)_2SO_4$, 1 mM DTT, 10% glycerol] and pelleted at 3,500 x g at 0 ^oC for 5 min. An equal volume of transcription buffer was added and 500 µl of suspended nuclei were dispensed by weight (500 µg) into preweighed microfuge tubes. Jel 318, Jel 466, Jel 42, or BSA (each at 120 µg), Actinomycin D (7.5 µg), or transcription buffer were added to the nuclei and tube contents were vortexed gently, then incubated for 2 h at R.T.. ATP, CTP, and GTP (each at 500 µM), UTP (100 µM), 10 mM $MgCl_2$, 4 x 10^{-6} µM [α-^{35}S] UTP (4 µCi), and transcription buffer were added to the nuclei to bring the final volume of each reaction to 1 ml. Following run-off transcription for 1.5 h at 37 ^oC, 100 µl aliquots were removed and 250 µl of 2% SDS was added. Each aliquot plus SDS (350 µl) was spotted onto a Whatman GF/C disc and washed under vacuum with 3 Pasteur pipettefuls of 5% TCA plus 20 mM sodium pyrophosphate and 1 pipetteful of 95% ethanol. Filters were then placed in scintillation vials, dried for 10 min under a heat lamp, 10 ml of scintillation fluid was added, and scintillation counts were obtained.

2.7.2.4 Incubation with Triplex-Specific Antibodies, Run-Off Transcription, RNA Isolation and Hybridization, and PhosphorImage Analysis

Large scale plasmid preparations were according to the basic protocol for preparation of crude cell lysate by alkaline hydrolysis in Current Protocols in Molecular Biology (Ausubel *et al.*, 1995), except that the pellet from a 500 ml overnight culture was divided into two tubes and each tube was treated with required solution volumes. Final plasmid purification was according to the polyethylene glycol precipitation alternate protocol, with the contents of the two tubes being pooled for extraction with buffered phenol and chloroform/isoamyl alcohol. The procedures in Current Protocols were used as they consistently yielded plasmid preparations with no contaminating chromosomal DNA and a high A_{260}/A_{280} ratio.

The following procedures were carried out under RNase-free conditions and all solutions were either treated with DEPC or used DEPC-treated ddH_2O. Antibodies in S-200 column buffer were dialyzed into transcription buffer (below) without spermidine. Agarose-encapsulated HT-29 nuclei were washed once with 5 volumes of ice-cold transcription buffer with spermidine [100 mM Tris-HCl (pH 7.6), 50 mM KCl, 2 mM EDTA, 10 mM spermidine, 200 mM $(NH_4)_2SO_4$, 1 mM DTT, 10% glycerol; Wittig *et al.*, 1991], and pelleted at 3,500 x g at 0 ^oC for 5 min. Nuclei (500µl = 500 mg) were weighed into preweighed, 15 ml polypropylene tubes on ice. The following solutions were added ice-cold: antibodies (Jel 318, Jel 466, or Jel 42; 120 - 750 µg), actinomycin D (30 µg), or BSA (120 - 750 µg), together with RNAguard™ (1,000 units), and transcription buffer with spermidine so that the final volume in each tube during transcription was 1 ml. All tubes contained an equal volume of transcription buffer without spermidine. Higher concentrations of antibodies than in Section 2.8.2.3 were used as 500 µl of pelleted microbeads were used versus 250 µl in the former section. Additionally, higher levels of transcription have been observed with higher concentrations of antibody (Wittig *et al.*, 1991). Tubes were incubated for 2 h at R.T. then cooled in ice for 10 min. Then the following were added ice-cold to each tube: 500 µM each of ATP, CTP, and GTP, 10 mM $MgCl_2$, 1,000 U RNAguard™, 50 µM s-adenosyl methionine, and 0.027 µM [α-^{32}P] UTP (8 µl; 80 µCi). Run-off transcription was carried out for 1.5 h at 37 ^oC.

The tubes were placed on ice and RNA was isolated from the agarose beads using Trizol™ reagent according to the supplier's protocol except for two modifications suggested by the supplier noted below. Briefly, Trizol™ reagent (6 ml) was added to each tube which was then extensively vortexed (about 1 min) until the solution completely cleared. Tubes were left 5 min at R.T. to allow dissociation of nucleoprotein complexes, then stored overnight at -70 °C. In the morning, tube contents were thawed. Then, as a small insoluble pellet (agarose; high carbohydrate) which could contain RNA remained after Trizol™ extraction, tubes were heated at 42 °C for 5 min to release RNA from agarose prior to chloroform extraction (first modification). Chloroform (1.2 ml) was immediately added and each tube was shaken vigorously by hand for 15 sec. The tubes were then incubated at R.T. for 2 to 3 min then centrifuged at 680 x g in a benchtop International Clinical Centrifuge at 4 °C for 45 min. Upper aqueous layers were removed with Pasteur pipettes and transfered to clean tubes. Isopropanol (1.5 ml) and 0.8 M Na citrate plus 1.2 M NaCl (1.5 ml; the second modification to the procedure) were added to each tube, the tubes were vortexed, incubated at R.T. for 10 min, then centrifuged for 50 min in the benchtop clinical centrifuge.

The supernatants were poured off the RNA pellets and the tubes were drained. Ethanol (75%; 7 ml; -20 °C) was added, tube contents were mixed by vortexing, then centrifuged for 15 min as above. Supernatants were again poured off, the tubes were drained, then tube insides were dried with laboratory tissues. Pellets were partially dried (5 to 10 min) on the bench. Each pellet was then solubilized in 300 µl of DEPC-treated ddH_2O plus 0.1 mM EDTA and kept on ice. An aliquot (5 µl) was removed for scintillation counting so that equal counts could be added to each hybridization bag.

Membrane strips were prepared and RNA hybridization and washing steps were carried out basically as described in Section 2.8.1 with the following changes. These plasmids were added to slots (5.8 µg in 200 µl/slot): pMI (complete *c-myc* genomic sequence) and pMGMT, a plasmid containing the gene O^6-methylguanine-DNA methyltransferase (a gift of Dr. W. Xiao, Department of Microbiology, University of Saskatchewan, Saskatoon). The plasmid pUC 19 was used as a control lacking genomic sequences. Several plasmids containing regions of *c-src* (gifts of Dr.

Bonham, Saskatoon Cancer Centre), e.g. 114C, were also used to detect level of expression of c-*src*. The membrane was neutralized in 3 x SET (Section 2.7.1), cut into strips with each plasmid present twice (in the same order on each half of the strip), then air dried. Strips were stored in plastic wrap at -20 °C until use. Two millilitres of Prehybridization/Hybridization solution was added to bags and stringency was optimized so that hybridization and washing were carried out at 60 °C. Washes following hybridization consisted of 3 x SET plus 1% SDS, 0.3 x SET plus 1% SDS, and 0.3 x SET. There was no final rinse in ddH_2O.

Radioactivity on the strips was detected and quantified using a PhosphorImager SI scanning instrument connected to a Microsoft® Windows NT™-based computer containing ImageQuaNT™ and Scanner Control software in the laboratory of Dr. Selvaraj at the Plant Biotechnology Institute, Saskatoon.

2.7.2.5 RNA Agarose Gel Electrophoresis

One percent RNA gels were made according to Sambrook *et al.* (1989). Agarose was melted in DEPC-treated ddH_2O (3.5 parts), cooled to 60 °C, then mixed with one part 12.3 M formaldehyde solution and 1.1 parts of 5 x formaldehyde gel running buffer [0.1 M MOPS (pH 7.0), 40 mM sodium acetate, 5 mM EDTA (pH 8.0)]. The gel was cast in a chemical hood. RNA (1-5 µg in 3-5 µl) was mixed with 1.5 µl formaldehyde and 4.5 µl formamide, heated to 60 °C for 8-10 min, cooled on ice and centrifuged for 5 sec. Following addition of 1 µl of loading buffer (50% glycerol, 1 mM EDTA, 0.4% bromophenol blue), samples were loaded into the lanes of a gel which had been pre-run for 5 min at 75 V. Electrophoresis was carried out at 75-100 V in 1 x formaldehyde gel running buffer. The gel was stained for 5-10 min in 0.5 µg EtBr/ml, destained in 1x formaldehyde gel running buffer overnight, then visualized under ultraviolet light.

2.7.2.6 Detection of RNase Activity in Transcription Assay Components

The pH 8 EtBr fluorescence assay described in Section 2.6.3 was used. Cuvettes were rinsed with tap water, 3 x with DEPC-treated ddH_2O, and 2 ml of pH 8 EtBr fluorescence assay buffer. Then 2 ml of pH 8 EtBr assay

buffer was dispensed into the tubes and RNA (5.2 µg; 1µl) was added to each tube. Fifty-five microlitres of the following were added to the tubes: transcription buffer without spermidine, BSA (200 µg in transcription buffer without spermidine), antibodies (Jel 42, Jel 466; 200 µg in transcription buffer without spermidine), or RNase A (2.5 µg in transcription buffer without spermidine). Tubes were incubated at 20 °C for 1.5 h and readings were taken every 15 min starting at 0 min. Before each reading, the fluorimeter was blanked with a tube containing 2,056 µl of pH 8 EtBr assay buffer.

2.8 Triplex Formation in the C-*Src* Promoter Region

Plasmids containing the promoter region of human c-*src* were kindly donated by Dr. Keith Bonham of the Saskatoon Cancer Centre, Saskatoon, Saskatchewan. Construction of these plasmids is described in Bonham and Fujita (1993).

2.8.1 Large Scale Plasmid Preparations

Two methods described in Sambrook *et al.* (1989) were used to prepare large quantities of plasmid. One was alkali lysis followed by polyethylene glycol precipitation which produced plasmid with a lot of contaminating chromosomal DNA. Plasmid DNA was then purified by cutting open circular and supercoiled bands from low melting agarose preparative gels without EtBr staining, heating the agarose to 70 °C to melt it, phenol/chloroform extracting the DNA, and reprecipitating it with 3M sodium acetate (pH 7.0). The second method was cesium chloride gradient centrifugation which enabled one to obtain pure plasmid DNA without contaminating chromosomal DNA.

2.8.2 Initial Studies on pBS-Src and p0.2SrcCAT

pBS-Src contains all the c-*src* promoter pur•pyr tracts, TC1, TC1.1, TC2, and TC3. Open circular forms of the plasmid were created by γ-irradiation for 7 min (Section 2.5.2). Plasmid (1.25 µg) was then incubated 19 h in 50 µM spermine and pH 4 triplex-forming buffer (Section 2.4.2) in a

total volume of 10 µl. The whole amount was added to the lane of a pH 4 or pH 8 0.85% agarose gel. The presence of any bands with the mobility of T-loop-like structures (i.e moving faster than open circular/linear forms but slower than supercoiled) was assessed following electrophoresis.

p0.2srcCAT contains the *c-src* promoter pur•pyr tracts TC2 and TC3 only. This plasmid was cut at the unique *Sma* I site producing linearized plasmid with TC2 at one end and TC3 at the other. Linearized p0.2srcCAT alone, or with pBS-Src linearized at the unique *Sac* I site upstream from the pur•pyr tracts, was incubated with spermine in pH 4 triplex-forming buffer in a total volume of 10 µl for 19 h. The whole amount was added to the lane of a pH 4 or pH 8, 1% agarose gel.

2.8.3 Linear Plasmid Fragments Containing Either TC2 or TC3

The pur•pyr tracts TC2 and TC3 were each obtained alone in linear plasmid fragments of unique length called TC2-linear and TC3-linear. TC2-linear and TC3-linear were obtained as follows: p0.2 SrcCAT was completely digested by *Sma* I as confirmed by agarose gel electrophoresis. The restriction digestion preparation was extracted with phenol/chloroform, precipitated with ethanol, and the resulting pellet was resuspended in T/E. The DNA was then digested to completion with *Sca* I and run on a pH 8, 1% agarose preparative gel. Bands containing TC2-linear and TC3-linear were identified by their mobility (second fastest and fastest of 3 bands; estimated sizes 1.25 and 0.91 kb, respectively). The bands were excised using a scalpel blade. The DNA in each band was then electroeluted at 5.7 V/cm for 2 h into pH 8 gel electrophoresis buffer. The volume of liquid containing the DNA was reduced at 37 °C under a stream of air to about 400 µl. Plasmid DNA was then extracted with phenol/chloroform and precipitated with ethanol using 3 M sodium acetate (pH 5.3).

2.8.4 End-Labelling TC1-Containing *BSS* HII Cassette and Autoradiography

To determine whether the pur•pyr tract TC1 formed a triplex with TC2 or TC3, the *BSS* HII cassette from pBS-Src containing TC1 was end-labelled with ^{32}P. End-labelling was carried out as follows: pBS-Src (500 µg) was digested to completion with *BSS* HII. The extent of digestion was

confirmed by an agarose gel and the DNA was extracted from the digest with phenol/chloroform. The aqueous layer was added to a 3% pH 8 preparative agarose gel and following electrophoresis, the *BSS* HII fragment containing TC1 was cut out. The DNA was electroeluted from the agarose band, reduced in volume under a stream of air, extracted with phenol/chloroform, ethanol precipitated, and the resulting pellet was resuspended in T/E.

A 20 µl end-labelling reaction contained the following: 120 ng plasmid *BSS* HII cassette, 1X Gibco/BRL React 2 buffer [50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM NaCl], 0.05 mM each of dATP, dGTP, and dTTP, 30 µCi [α-^{32}P] dCTP (3,000 Ci/mmol), and 2 units of the Klenow fragment of DNA polymerase I, labelling grade (added last). Everything was kept on ice until all ingredients were added. End-labelling was then carried out for 30 min at R.T. The reaction mix was brought up to 70 µl by the addition of 50 µl of STE buffer and the end-labelled fragment was purified on a NucTrap® probe purification column according to the manufacturer's instructions.

End-labelled fragment (1 µl) was incubated at R.T. for 21.5 h with TC2-linear or TC3-linear (0.83 µg) using pH 4 triplex-forming buffer and increasing concentrations of spermine in a total volume of 10 µl. As a control, linear fragments were incubated in 11 mM Tris-HCl (pH 8.0), 10 mM NaCl, 0.2 mM EDTA, and no spermine. In a separate experiment to determine if non-specific binding were occurring between the end-labelled fragment and TC3-linear, TC3-linear was digested with *Bam* HI to remove the TC3 pur•pyr tract. TC3-linearΔTC3 was identified as a 180 bp band on a 3% pH 8 preparative gel. The band was excised and the DNA was electroeluted, phenol chloroform extracted, precipitated, and resuspended in T/E. TC3ΔTC3 was incubated with the end-labelled fragment as described above. The whole contents of an incubation mix were added to the lane of a 2%, pH 4 agarose gel. After gel photography, gels were wrapped in plastic wrap and placed under X-Omat-AR film with an enhancer screen in a film cassette for 6 to 14 h at -70 °C.

2.8.5 Exonuclease III Digestion to Obtain TC1 Alone

A linear plasmid fragment of unique length containing only TC1 was obtained by exonuclease III digestion. pBS-Src (58 µg) was cut in the

multiple cloning site with the restriction enzyme *Bam* HI. The digestion reaction was heat-inactivated, then cooled, and the *Bam* HI site was blunted with Klenow as follows (Ausubel *et al.*, 1995). The Klenow fragment of DNA Polymerase I, FPLCpure (75 units) and 100 mM each of dCTP and dGTP were added to the heat-inactivated digestion reaction in a total volume of 525 µl. The Klenow reaction was carried out for 15 min at 30 °C, then heat-inactivated for 10 min at 75 °C and cooled to R.T.. Linearized pBS-Src was then cut with *Sac* I. This produced a fragment containing TC1, TC2, and TC3 which was purified as a band of 372 bp from a pH 8, 2% agarose gel. The DNA in the band was electroeluted, blown down, phenol/chloroform extracted, precipitated with ethanol, and resuspended in T/E.

The multiple cloning site of pUC19 was digested with *Sac* I and *Sma* I. The DNA in the digest was extracted with phenol/chloroform and the large DNA fragment was gel purified on a 1% pH 8 agarose gel. The band was cut out, the DNA was electroeluted, the volume of electroeluant was reduced under a stream of air, and the DNA was precipitated with ethanol. The resulting pellet was resuspended in T/E. The *Sac* I-*Bam* HI insert of pBS-Src was then cloned into pUC 19. The ligation reaction was carried out at R.T. for 24 h in a buffer containing 50 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 20 mM DTT, 1 mM ATP, and 50 µg BSA/ml in a volume of 10 µl. Reaction conditions consisted of 0.04 µg of the vector, 0.05 µg of the insert, and 60 cohesive end units of T4 DNA ligase. *E. coli* strain JM109 was transformed with the products of the ligation reaction and plated on LB agar plates containing 50 µg ampicillin/ml and spread with 40 µl of a 20 mg/ml solution of IPTG and 20 µl of a 50 mg/ml solution of X-gal. Plasmid DNA from white colonies was prepared using the Wizard Plus Minipreps DNA Purification System. A clone containing a vector with the entire TC1 tract was identified by digesting plasmid DNA with *BSS* HII and observing a linear band on agarose gel electrophoresis.

After this, an exonuclease III digestion was carried out to remove the TC2 and TC3 tracts from the new plasmid, leaving only TC1. The Double-stranded Nested Deletion kit was used according to the instructions supplied. Briefly, restriction digestion of the *Sph* I site of pUC 19 gave a 3' overhang and a portion of the digest was checked on an agarose gel to ensure linearization was complete. Overhanging 3' ends of 3 or more

nucleotides in length are resistant to exonuclease III. The plasmid was then cut with *Bam* HI which gave a 5' overhang. Exonuclease III removes nucleotides progressively from the 3' end of a DNA strand where there is a 5' overhang.

The single-stranded region was then removed with S1 nuclease. A deletion of only 184 bp was required to remove TC3 and TC2, leaving TC1. The exonuclease III digestion was therefore carried out at R.T.. An initial sample was taken at time 0 and further samples were removed at 2 min intervals up to 16 min. A final sample was removed at 30 min as a control to ensure the exonuclease digestion had worked. *E. coli* HB101 competent cells were prepared and transformed (Section 2.2.2.) with recircularized (ligated) plasmid DNA from each time point. The transformed cells were plated out on LB plus ampicillin (50 μg/ml). After overnight growth, minipreps were done on 2 randomly selected colonies from each time point using the Wizard Plus Minipreps DNA Purification System.

Miniprep DNA was screened by fragment size following digestion with *Eco* RI and *Hin* dIII. The *Eco* R I-*Hin* dIII fragment was 441 bp without exonuclease digestion. In clones where TC2 and TC3 had been removed, the fragment would be only 200 bp. Following restriction enzyme digestion with *Eco* RI and *Hin* dIII, plasmid DNA was analyzed by agarose gel electrophoresis along with a 100 bp DNA ladder for estimation of size of fragments. This was difficult due to a diffuse spreading of fragment bands on gels, thought to be due to nuclease activity from the cultures. However, several clones from different time points where the *Eco* RI-*Hin* dIII digestion fragment appeared to run in approximately the right position were selected for sequencing. The clones were sequenced using the ^{T7}DNA Polymerase Sequencing Kit and the Universal (-20) primer according to the procedure supplied. A clone from the 10 min time point was identified which contained the c-*src* promoter sequence to the end of TC1 plus 1 nucleotide. A large scale plasmid preparation of this clone was made according to Ausubel *et al.* (1995). Plasmid DNA from this clone was linearized at the *Hin* dIII site to create a 2,806 bp fragment (TC1-linear) for use in dimerization experiments. The completeness of linearization was checked by agarose gel electrophoresis, then the DNA in the digestion reaction was extracted with phenol/chloroform. The aqueous layer was added to a pH 8, 1% preparative agarose gel. The linear band was cut out,

the DNA was electroeluted, followed by concentration under a stream of air. The concentrated DNA was extracted again with phenol/chloroform, ethanol precipitated with 3 M sodium acetate, pH 7 at -70 °C, then finally resuspended in T/E.

2.8.6 Dimer Formation and Gel Electrophoresis

Equimolar amounts of the linear plasmid fragments TC1-linear, TC2-linear, and TC3-linear were incubated together in pairs (1 µg of DNA in total) in pH 4 triplex-forming buffer and spermine for at least 19 h at R.T. in a volume of 10 µl. Two µl of pH 4 loading buffer was added to the incubation mix and the whole volume was added to a lane of a pH 4, 0.85% agarose gel. The composition of linear plasmid fragments in gel bands following electrophoresis was determined by estimating the size of the bands in bp. This was done by comparing band mobilities with a standard curve of $\log_{10}$ of the MW marker band sizes in bp versus distances travelled on the gel.

To determine whether the third strand of the triplex in TC1/TC3 heterodimers was contributed by TC1 or TC3, an oligonucleotide corresponding to the pyrimidine tract of TC1 (30 nucleotides; CTTCCTCCTTCCTCCTCCTCCC) or TC3 (31 nucleotides; CCTCCTCCTCCTCCTCCTCCCTTTCTCTCTC) was added to the incubation of TC1-linear and TC3-linear.

2.8.7 Gel Imaging

Following staining in EtBr, gels were either photographed (Section 2.3) or imaged using AlphaImager 3.24i software and an AlphaImager™ 2,000 Documentation and Analysis System. Images were cropped and arranged using Adobe Photoshop™ 4.0 and labelled using Canvas™ 3.5.3.

3.0 RESULTS

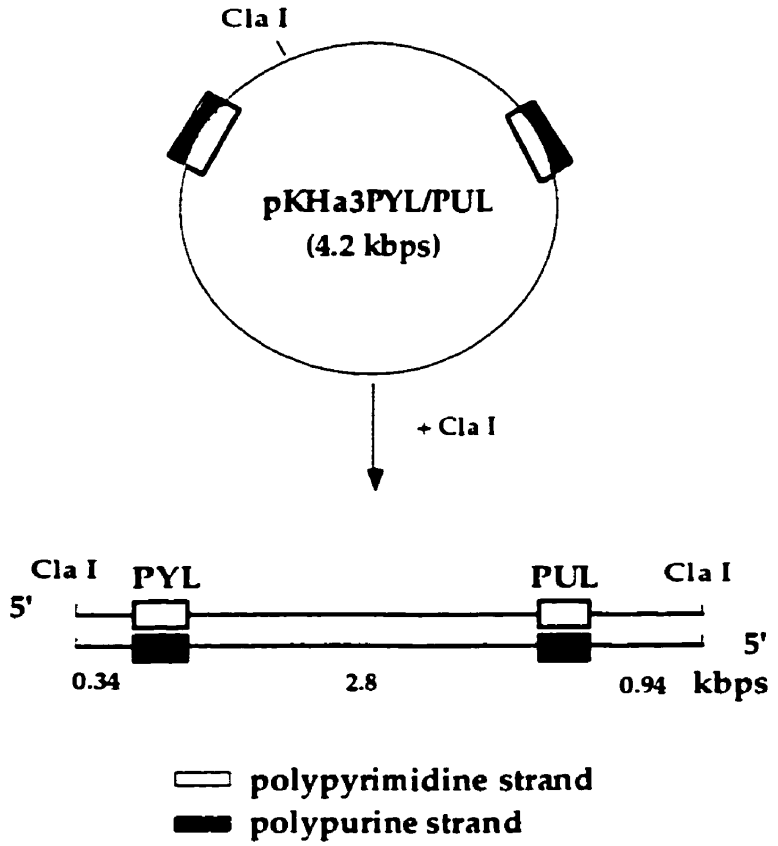
3.1 pKHa3PYL/PUL Linear X-Dimers and Omega Loops

Looping in which distant tracts of DNA are brought together is known to be important in the condensation of chromosome length DNA (Gasser and Laemmli, 1987; Uemura *et al.*, 1987; Adachi *et al.*, 1991). Looping, in which regulatory sequences are brought together, is also important in gene regulation. Both of these looping processes are known to be mediated by proteins. It has been postulated that triplex formation may be involved in the condensation of DNA in chromosomes (Hampel and Lee, 1993). This first set of experiments was carried out in collaboration with Dr. K.J. Hampel to test whether DNA looping could occur as a result of triplex formation between distant but complementary pur•pyr tracts on the same tract of duplex DNA. A triplex formed between such separated pur•pyr tracts is referred to as a transmolecular triplex.

The pUC19-derived plasmid pKHa3PYL/PUL contains a pair of asymmetric 72 bp pur•pyr tracts, $d(CTTCCT)_{11}$ and $d(TCCTTC)_{12}$ (Figure 3.1). These tracts are complementary in that they are designed to form a triplex with each other, but not within a tract itself because of the requirement for the pyrimidine strands to be antiparallel (Morgan and Wells, 1968; Arnott and Selsing, 1974). Cutting pKHa3PYL/PUL at the unique Cla I site produces a linear plasmid monomer of 4.2 kb. Linearized pKHa3PYL/PUL was used as a model to study looping in duplex DNA mediated by transmolecular triplex formation.

3.1.1 Formation of Dimers and Loops at Low pH

Incubation at pH 4 in the presence of spermine promotes the formation of a pyr•pur•pyr or double pyrimidine triplex between pur•pyr tracts. This is because the low pH protonates cytosines in the incoming third (pyrimidine



Plasmid Inserts

PYL
5'-CCC(CTTCCT)$_{11}$CTTCCGGGG-3'
3'-GGG(GAAGGA)$_{11}$GAAGGCCCC-5'

PUL
5'-CCC(TCCTTC)$_{12}$GGG-3'
3'-GGG(AGGAAG)$_{12}$CCC-5'

Figure 3.1. pKHa3PYL/PUL showing unique *Cla* I site and sequence of the pur•pyr tracts, PYL and PUL. kbps - kilobase pairs.

strand) of the triplex, enabling the formation of hydrogen bonds with the purine strand of the duplex (Lipsett, 1964). Figure 3.2 shows linear X-dimer and Ω loop structures which could result from the formation of a pur•pyr triplex. The term "X-dimer" denotes the fact that if two linear plasmid monomers dimerize via triplex formation, the join will be at pur•pyr tracts located some distance from the ends of either plasmid. Note that in this case the X-dimers and Ω loops can be formed in two ways, depending on which pur•pyr tract, PUL or PYL, contributes the third strand of the triplex.

The formation of these structures was tested by the appearance on an agarose gel of new, acid pH-dependent bands with a different mobility from the linear plasmid monomer. When the linearized plasmid was incubated at pH 4, two new bands with a slower mobility than the linear monomer were seen on a pH 4 gel (lane 2, Figure 3.3). By trial and error it was found that a 1.8% agarose gel gave the best separation of the new bands from those containing linear and open circular plasmid forms. Formation of the new bands was dependent on incubation with spermine: the amount of linear plasmid converted to the new forms increased as the concentration of spermine in the incubation increased from 25 μM (lane 1) to 100 μM (lane 4, Figure 3.3). Pyr•pur•pur or double purine triplexes are a different class of triplex which forms at neutral pH but requires a high concentration of divalent cations and supercoiling (Kohwi and Kohwi-Shigematsu, 1988; Kohwi *et al.*, 1992). No bands other than the linear monomer were seen following incubation of the linearized plasmid at pH 8 at low ionic strength, even in the presence of 10 mM $MgCl_2$ (Hampel, 1994). Therefore, the structures in the new bands were based on the formation of a double pyrimidine triplex.

Hampel *et al.* (1993) identified the structures in the slowest moving of the new bands as linear X-dimers (lanes 1 and 2, Figure 3.3). The faster moving band (lane 2, Figure 3.3) could not have been a linear dimer as its mobility was faster than the open circular form (lane 5, Figure 3.3). This new band with faster mobility than the open circular form only appeared following incubation of plasmids containing both the asymmetrical pur•pyr tracts, PUL and PYL (Hampel, 1994). The band therefore contained the Ω loop structures of Figure 3.2. The expected loop size and length of the long and short arms has since been confirmed by measurements of electron micrographs of molecules isolated from the band (not shown). The slower moving closely-packed bands on the gel following incubation at higher spermine concentrations are multimers (lanes 3

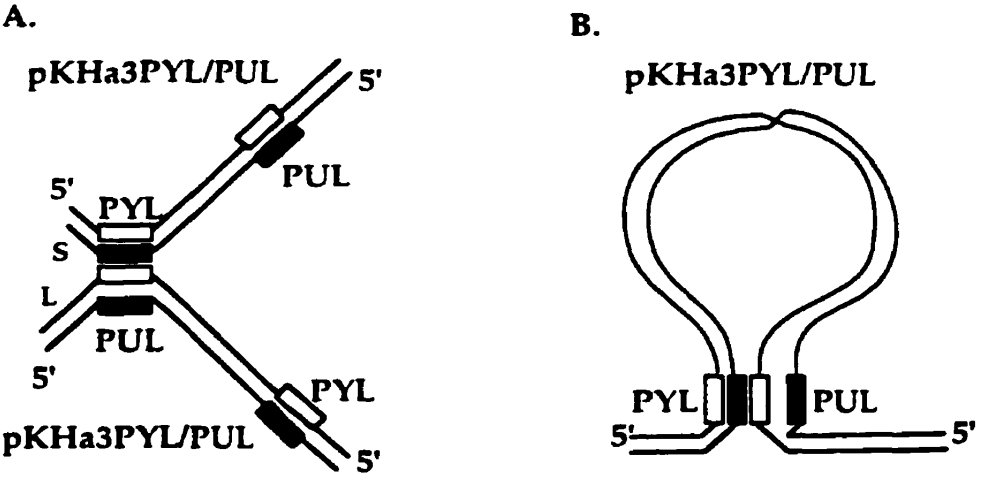


Figure 3.2. Possible structures of linear X-dimers and Ω loops. The structures were formed from pKHa3PYL/PUL linearized with Cla I and incubated at low pH in the presence of spermine. PUR and PYR: pur•pyr tracts as in Figure 3.1. Open and closed boxes represent pyrimidine and purine strands, respectively. A. Linear X-dimer. Arm lengths: S - 0.34 kb, L - 0.94 kb. B. Ω loop.

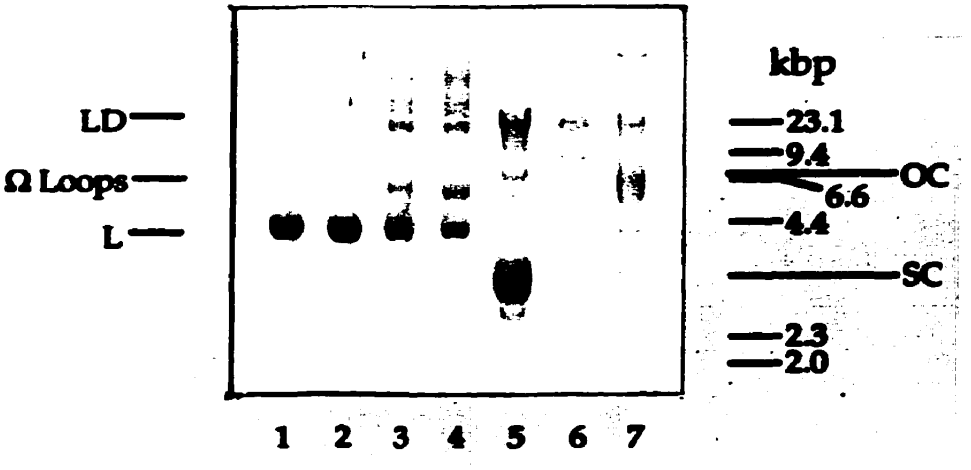


Figure 3.3. Formation of pKHa3PYL/PUL linear X-dimers and Ω loops. pKHa3PYL/PUL linearized with *Cla* I was incubated in 25 mM sodium acetate (pH 4), 1 mM EDTA (pH 8.0), and 25 (lane 1), 50 (lane 2), 75 (lane 3), and 100 (lane 4) μM spermine. Lane 5 - native plasmid, lane 6 - λ plus *Hin* dIII MW markers, lane 7 - native plasmid γ-irradiated for 5 min. Positions of plasmid forms are indicated: OC - open circular, L - linear monomer, SC - supercoiled, LD - linear X dimer, Ω - omega loop. Electrophoresis was carried out for 4.5 h in a pH 4, 1.8% agarose gel.

and 4, Figure 3.3). These would have arisen due to additional triplex interactions promoted by the higher spermine concentrations. The position of the open circular band in a control lane was shifted in the direction of faster moving DNA due to the large amount of DNA in the band (lane 7, Figure 3.3). Native plasmid is that grown in *E. coli* before further manipulations (lanes 5 and 7, Figure 3.3).

3.1.2 Effect of pH, Temperature, and Ionic Strength on Dimer and Loop Stability

Once formed at low pH, pyr•pur•pyr triplexes tend to show pH hysteresis in that they are stable at higher pHs (Lee *et al.*, 1984; Hampel *et al.*, 1991). The linear X-dimers and Ω loops were formed at a low, non-physiological pH. Therefore, their stability at higher pHs was investigated to determine their possible biological relevance. Dimers and loops were formed as described in the Materials and Methods, then transferred using spin columns into pH 5, 6, and 8 buffers of the same ionic strength. The dimers and loops were then re-incubated (Figure 3.4). Some smearing of bands is visible in the dimer and loop control (lane 4, Figure 3.4). This may be due to precipitation of the DNA as a result of storage at pH 4 in the concentration of 100 μM spermine used to form the dimers and loops prior to running the gel. Dimer and loops were stable following re-incubation at pH 5 and 6 without and with spermine (lanes 5-8, Figure 3.4). Less DNA was present in all bands following re-incubation at pH 5 and 6 in the presence of 50 μM spermine (lanes 6 and 8, Figure 3.4). In these experiments, this concentration of spermine was too low to cause DNA precipitation. As the size of all bands in lanes 6 and 8 decreased equally, including those containing plasmid monomers, it is likely that some DNA was lost from the wells on loading the gel. Strong dimer and loop bands were present following re-incubation at pH 8 (lanes 9 and 10, Figure 3.4). Therefore, the structures were stable at physiological pH. Although spermine was required for the formation of the loops and dimers, once formed its presence was not necessary for their stability (lane 9, Figure 3.4).

The effect of temperature and ionic strength on the stability of the dimers and loops is investigated in Figure 3.5. Dimer and loop bands were smaller and multimer bands were absent in a control compared to the rest of the gel (lanes 1 and 2, Figure 3.5). This was because a lower concentration of spermine was

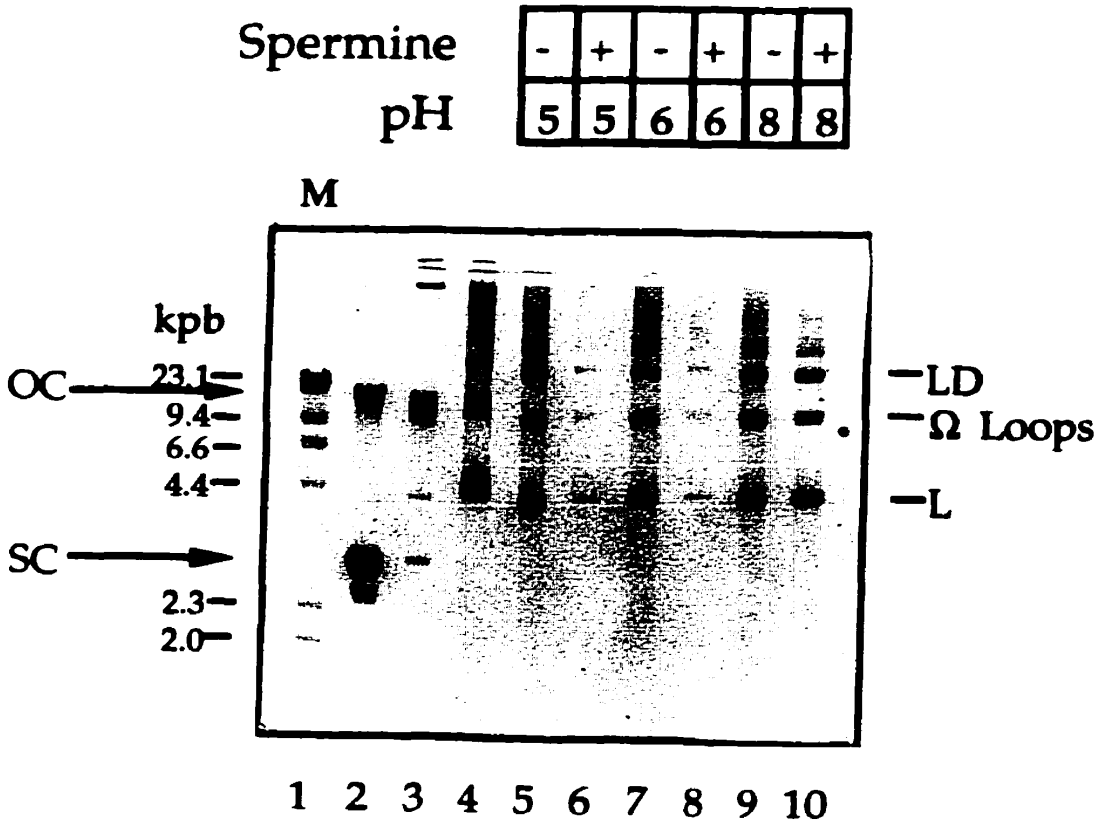


Figure 3.4. Stability of pKHa3PYL/PUL dimers and Ω loops after re-incubation at various pHs without and with spermine. Lane 1 - λ plus *Hin* dIII MW markers (M), lane 2 - native plasmid DNA, lane 3 - native plasmid γ-irradiated for 5 min, lane 4 - dimers and loops with no re-incubation. Dimers and loops were formed at pH 4 in the presence of spermine then incubated for 19 h at R.T. at pH: 5.0 - lanes 5 and 6, 6.0 - lanes 7 and 8, 8.0 - lanes 9 and 10. Incubations in lanes 6, 8, and 10 contained 50 μM spermine. Abbreviations as for Figure 3.3. pH 8, 1.8% agarose gel.

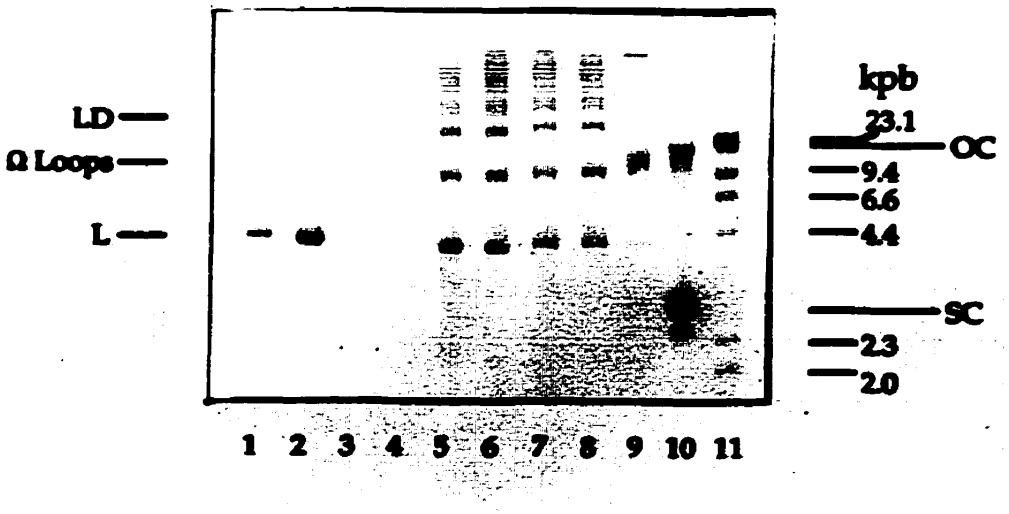


Figure 3.5. Temperature and ionic strength stability of pKHa3PYL/PUL dimers and Ω loops. Lanes 1 and 2 - control (dimers and loops formed in 25 mM sodium acetate and 50 μM spermine), lanes 3 and 4 - blank. Dimers and loops were formed at pH 4 in 100 μM spermine then re-incubated for 17 h under the following conditions: lane 5 - 50 °C in T/E, lane 6 - 0 °C in T/E, lane 7 - 62.5 mM sodium acetate (pH 4), lane 8 - 62.5 mM sodium acetate (pH 4) plus 200 mM NaCl. Lane 9 - native plasmid γ-irradiated for 5 min, lane 10 - native plasmid, lane 11 - λ plus *Hin* dIII MW markers. Abbreviations as for Figure 3.3. pH 8, 1.8% agarose gel.

used in their formation than on the rest of the gel. Strong dimer, loop, and multimer bands remained following incubation at pH 8 at low ionic strength (lanes 5 and 6, Figure 3.5). Additionally, the structures were extremely heat-stable, as the bands were the same size whether incubated at 50 °C (lane 5, Figure 3.5) or on ice (lane 6, Figure 3.5). Dimers and loops also remained at pH 4 when the stabilizing polyvalent cation spermine necessary for their formation was substituted with the monovalent cation Na^+ (lanes 7 and 8, Figure 3.5). There was no difference in the size of the bands following re-incubation in 62.5 mM sodium acetate (pH 4) (lane 7, Figure 3.5) or at a much higher ionic strength when the concentration of the monovalent cation Na^+ in the buffer was increased by the addition of 200 mM NaCl (lane 8, Figure 3.5). The structures therefore showed considerable hysteresis at the pH of formation when the stabilizing polyvalent cation was removed. Na^+ is not able to neutralize phosphate charges sufficiently to bring about the condensation of DNA strands and hence cannot stabilize a triplex (Wilson and Bloomfield, 1979).

3.1.3 Effect of Topoisomerase I and P1 Nuclease on Dimers and Loops

Two models were proposed for the X-junction at the root of the linear X-dimer and Ω loops. The stability of the dimers and loops at pH 8 and high temperatures suggested that they were not merely formed by a twisting of the two duplex DNA strands at the pur•pyr tracts when the triplex is formed (Figure 3.6A). If this were the case, the structures would be expected to unwind at least partially when incubated at 50 °C. In previous studies Y dimers were formed from linear plasmids containing the same pur•pyr tracts (Hampel *et al.*, 1993). The difference was that in the previous experiments the pur•pyr tract was located at the end of the linear plasmid monomer which joined the other linear monomer to form a dimer. In this previous work, very few Y dimers remained following electrophoresis at pH 7 and none were seen at pH 8. The Y dimers were therefore unwound at higher pHs without any topological trapping. The stability of the linear X-dimers and Ω loops in the current study suggested that the structures were somehow topologically constrained upon transfer to a higher pH or on heating. It was unlikely that a protonated triplex would be stable under the experimental conditions at pH 8. This led to an investigation of whether the join at the pur•pyr tracts contained a single-stranded region. The nature of the join was analyzed at pH 8 using

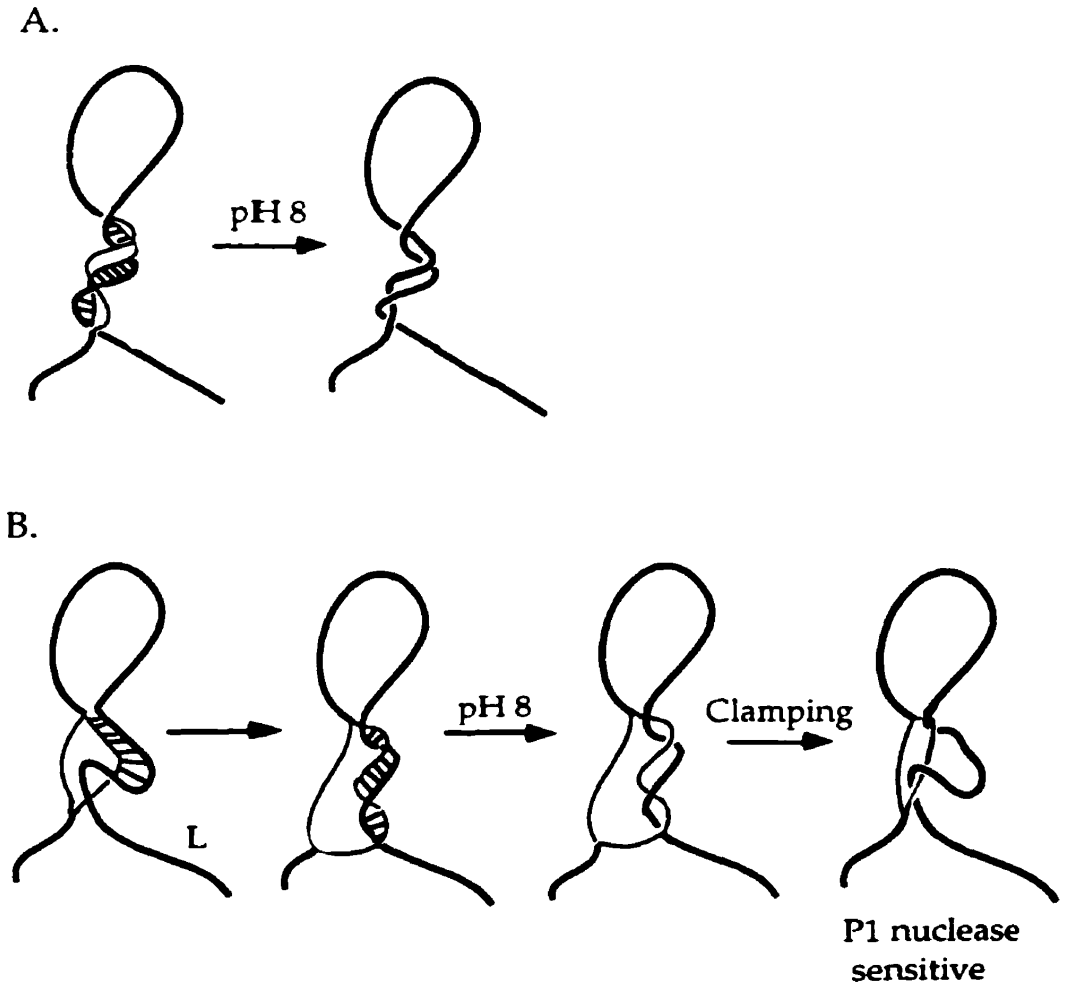


Figure 3.6. Models for the junction in the linear X-dimers and Ω loops. Dashed lines indicate the triplex formed between one of the pur•pyr tracts and the pyrimidine strand of the other pur•pyr tract. A. twisted structure. B. hydrogen bonded or braided knot. L - duplex end which passes once through the single-stranded bubble for each turn of triplex.

topoisomerase I and P1 nuclease. The X-dimers and Ω loops were not unwound with topoisomerase I under conditions relaxing supercoiled plasmid (data not shown). Topoisomerase I catalyzes removal of supercoils in covalently closed circular DNA by transient breakage and resealing of phosphodiester bonds. However, both linear dimers and loops reverted to linear monomers following incubation with 0.5 unit of P1 nuclease, a phosphodiesterase preferentially cleaving single-stranded DNA (lane 4, Figure 3.7). Therefore, the X-junction contains a single-stranded region at pH 8. Figure 3.6B is a model accounting for the stability of the linear X-dimers and Ω loops and the presence of the single-stranded region. The X-junction consists of a braided or hydrogen-bonded knot. A single-stranded bubble accompanies triplex formation at pH 4. When the pH is changed to 8, the duplex partially reforms in the region of the bubble, clamping the duplex loop of the end marked L in position. The partially reformed duplex explains the P1 nuclease sensitivity of the structures.

Finally, as a control, the major supercoiled form in a native plasmid was not nicked with P1 nuclease as there was no change in band size (lanes 5-8, Figure 3.7). However, a smaller and faster moving band, which likely contained more tightly supercoiled plasmid (lanes 5 and 6, Figure 3.7), was converted to linear plasmid monomers with 0.5 unit of P1 nuclease (lane 7, Figure 3.7). Again, this was evidence for single-stranded regions and supported the premise that this faster moving form was a more highly supercoiled form of plasmid. Constrained, highly supercoiled duplex DNA is known to contain regions that are single-stranded (Beard *et al.*, 1973; Germond *et al.*, 1974).

3.2 T-Loops

3.2.1 Transmolecular Triplex Formation in a Circular Plasmid

As the DNA in chromosomes is topologically constrained in loops plectonemic problems could arise in the formation of transmolecular triplexes *in vivo*. Therefore, a topologically equivalent model for studying transmolecular triplex formation should be provided by a circular plasmid with separated complementary pur•pyr tracts. Accordingly, the open circular form of pKHa3PYL/PUL (Figure 3.1) was used in the next set of experiments.

Following γ-irradiation for 7 min (21,500 Rads), approximately 95% of pKHa3PYL/PUL was in the open circular form. When the open circular forms

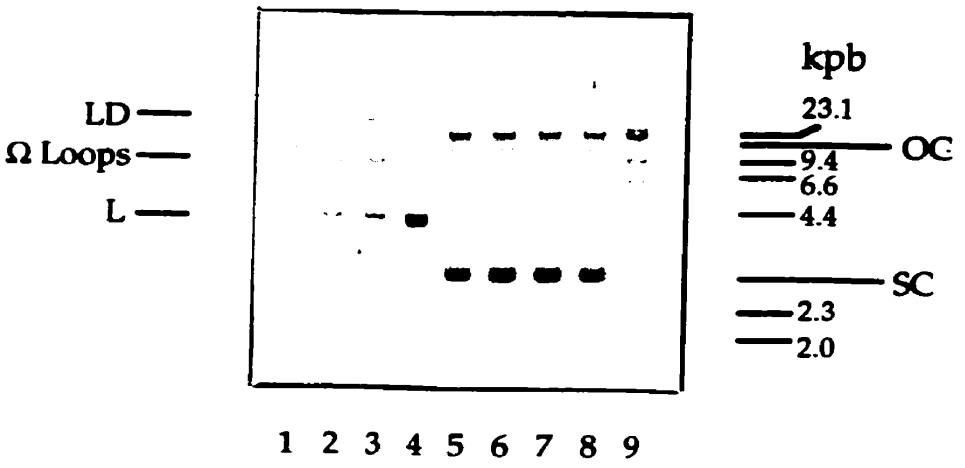


Figure 3.7. Effect of P1 nuclease on pKHa3PYL/PUL linear X-dimers and Ω loops. Dimers and loops were incubated with the following units of P1 nuclease: lane 1 - control (no P1 nuclease), lane 2 - 0.005, lane 3 - 0.05, lane 4 - 0.5. Lanes 5-8 - native pKHa3PYL/PUL control incubated with same concentrations of P1 nuclease as lanes 2-4. Lane 9 - λ plus *Hin* dIII MW markers. Abbreviations as in Figure 3.3. pH 8, 1.8% agarose gel.

were incubated at pH 4 in the presence of spermine, a new band with a faster mobility than linear plasmid but slower than supercoiled form appeared on both pH 4 and pH 8 agarose gels (lane 6, Figure 3.8A,B). Linear pKHa3PYL/PUL ran slightly faster than open circular plasmid on both pH 4 and pH 8 gels, however this was most clearly visible on pH 8 gels (lane 5, Figure 3.8B). The supercoiled form had the highest mobility (lanes 1, 2, 5, and 6, Figure 3.8A,B). By trial and error, an 0.85% agarose gel was found to give the greatest separation between open circular and supercoiled plasmid forms. The forms in the new band with a faster mobility than linear plasmid were referred to as T-loops, for "transmolecular triplex loops". A control experiment showed that in comparison with relaxed plasmid, the mobility of the T-loops was equivalent to a molecule containing one or two supercoils. The T-loop band was not present when spermine was absent from the incubation mix (lane 5, Figure 3.8A,B). T-loops were also formed when pKHa3PY/PU, a plasmid with pur•pyr tracts having the same complementary sequences as pKHa3PYL/PUL but one-half the length (32 bp), was incubated at pH 4 in the presence of spermine (not shown). Therefore, T-loops also formed from circular plasmids having much shorter pur•pyr tracts. On the other hand, the acid pH requirement for T-loop formation was very severe - no band formed when γ-irradiated pKHa3PYL/PUL was incubated with spermine at pH 4.5 or above (not shown). Finally, both pur•pyr tracts were required as T-loops did not form when an open circular plasmid having only one tract was incubated at pH 4 in the presence of spermine (not shown). Therefore, T-loop formation was extremely acid pH-dependent, required the presence of spermine, and only occurred when both asymmetric pur•pyr tracts were present in the same plasmid. This was strong evidence that a pyr•pur•pyr triplex was at the root of the T-loop structure.

Relaxed forms of the plasmid were made by incubating native plasmid with topoisomerase I. Following incubation of relaxed plasmid at pH 4 in the presence or absence of spermine, a new form with increased mobility also appeared on pH 4 gels (lane 3 and 4, Figure 3.8A). This form was not present upon electrophoresis at pH 8 (lane 3, Figure 3.8B) except for a small percentage of the forms which had been incubated with spermine (lane 4, Figure 3.8B). Linear dimers formed in the presence of spermine and were seen on both pH 4 and pH 8 gels (lane 8, Figure 3.8A,B). These were the linear X-dimers of Section 3.1. Dimers formed from open circular and relaxed plasmid forms and had a

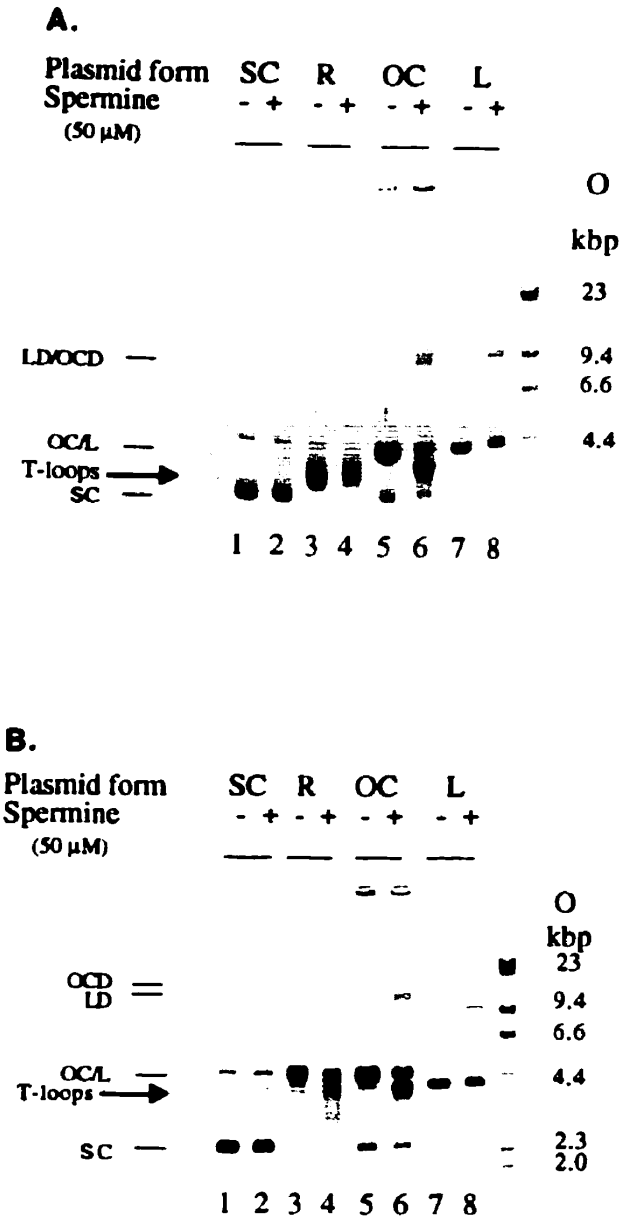


Figure 3.8. Formation of T-loops in open circular pKHa3PYL/PUL. Plasmid forms indicated above the lanes: SC - supercoiled, R - relaxed, OC - open circular, L - linear. Lanes contained 0.9 μg of DNA and electrophoresis was at 3.6 V/cm for about 5 h at 10 °C. The positions of linear/open circular dimers (LD/OCD), open circular plasmid (OC), linear plasmid (L), T-loops, and supercoiled plasmid (SC) are indicated. Molecular weight markers are λ DNA cut with *Hin* dIII. A. pH 4, 0.85% agarose gel. B. pH 8, 0.85% agarose gel.

similar mobility to linear dimers (lanes 4 and 6, Figure 3.8A,B). A small T-loop band also appeared on a pH 8 gel following incubation of supercoiled DNA with spermine (lane 2, Figure 3.8B). This must have arisen from the small numbers of open circular forms present in the supercoiled DNA. Even if triplexes formed between the separated pur•pyr tracts in the supercoiled molecules, there would not have been a mobility change as the molecules were already tightly supercoiled.

Following formation at pH 4 in the presence of spermine, T-loops were very stable at physiological pH without the presence of the stabilizing cation spermine. For example, they were stable for at least two weeks at 4 °C in T/E (data not shown). They were also very stable when heated in T/E: T-loops incubated at 37 °C for 24 h did not revert to the open circular form - a temperature of 50 °C was required (lanes 4 and 5, Figure 3.9). Surprisingly, when the ionic strength of the incubation buffer was increased by addition of 200 mM NaCl, the T-loops were less stable at high temperatures and reverted to open circular forms at 37 °C (lane 10, Figure 3.9). This may have been due to the increased flexibility of DNA at higher ionic strength (Thomas and Bloomfield, 1983a). The stability of T-loops at pH 8 was reminiscent of the topological trapping seen in linear X-dimers and Ω loops made from the same plasmid (Section 3.1). Again, as with the dimers and loops, it seemed unlikely that a pyr•pur•pyr triplex remained under the extreme conditions at pH 8. As the topological trapping at pH 8 in the dimers and loops was found to be due to the presence of a single-stranded region, the effect of P1 nuclease on the stability of T-loops was investigated. Following incubation with 0.5 unit of P1 nuclease, T-loops reverted to open circular forms (lane 6, Figure 3.10). As before, this was evidence for the existence of single-stranded region topologically trapping the structures when the pH was increased to 8.

The band containing the T-loops was usually present as a doublet (e.g. see lane 6, Figure 3.8A,B). The two bands were excised separately from a pH 4 agarose gel and plasmid DNA was extracted by centrifugation through glass wool. The structures of the T-loops in the bands were examined using electron microscopy. Approximately 50% of the well-defined plasmid molecules extracted from the slower moving T-loop band were dumbell or figure 8 in shape (Figure 3.11A). Usually, the two loops in the dumbell were of unequal size (Figure 3.11 A,B, and D). Analysis of small and large loop sizes by histogram showed that size distribution was not random (Figure 3.12). Thirty-

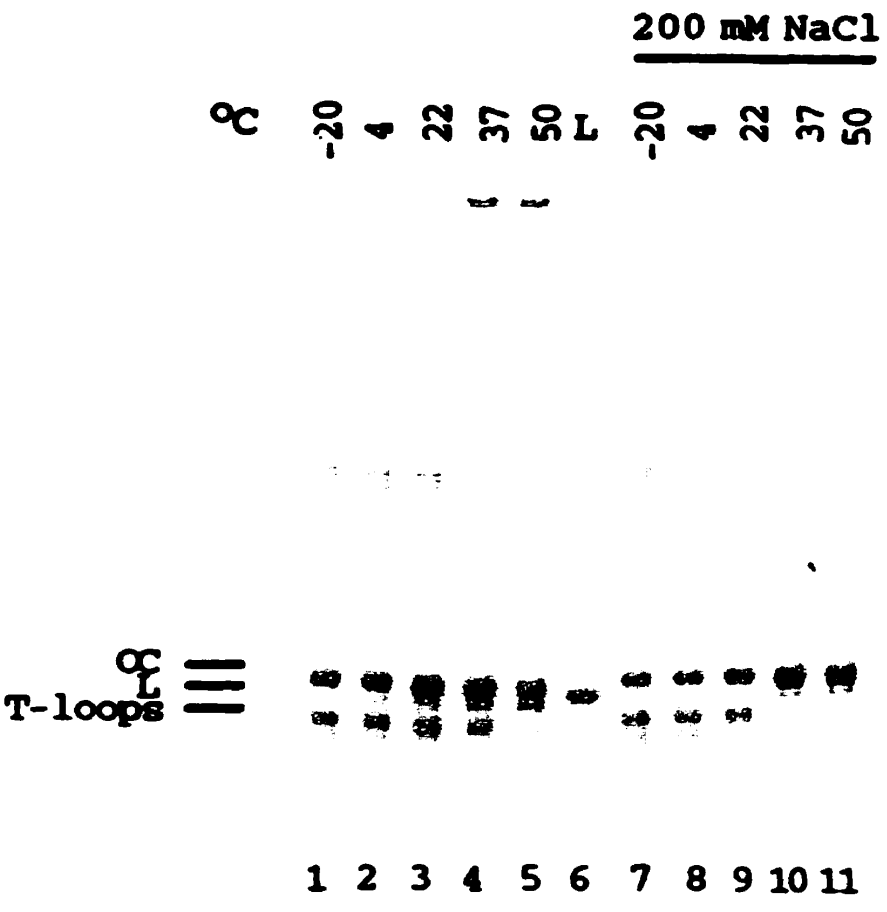


Figure 3.9. Temperature and ionic strength stability of T-loops. T-loops were incubated for 24 h at the indicated temperatures in T/E - lanes 1 to 5 and in T/E plus 200 mM NaCl - lanes 7 to 11. Plasmid form abbreviations as for Figure 3.8.

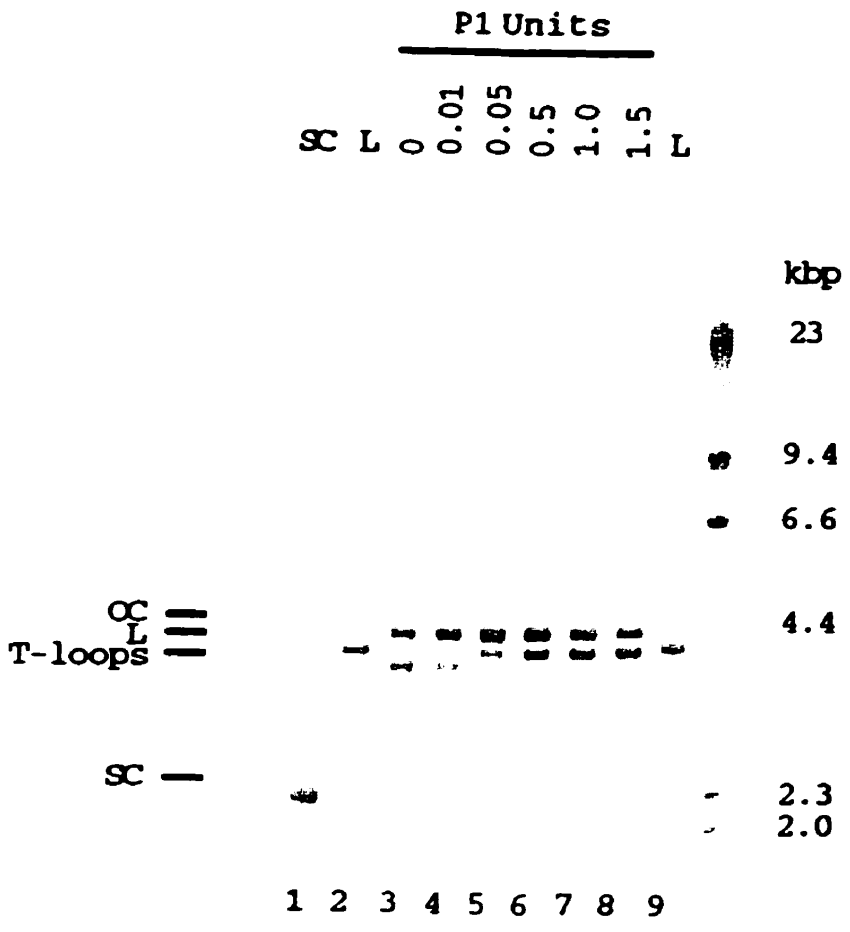


Figure 3.10. P1 nuclease sensitivity of T-loops. Plasmid form abbreviations as for Figure 3.8. Positions of λ plus *Hin* dIII MW markers are indicated. Units of P1 nuclease in the incubation are given above the lanes. pH 8, 0.85% agarose gel.

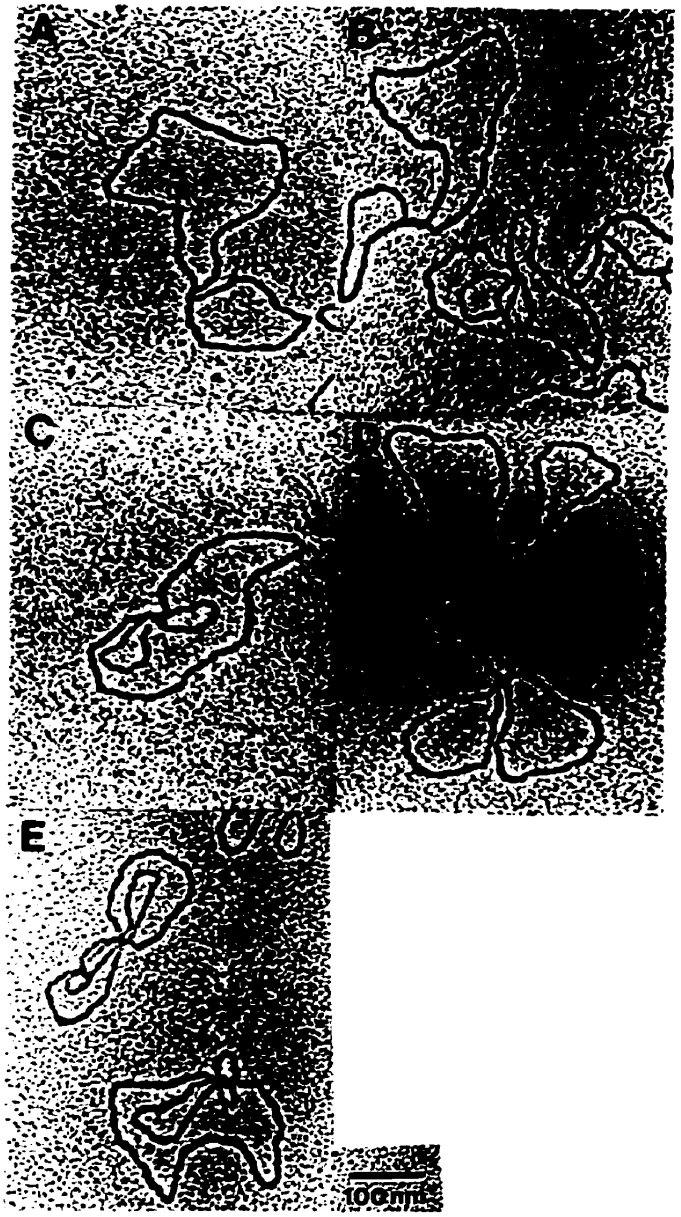


Figure 3.11. Electron microscopy of T-loops. A. dumbell, B. dumbell with linker, C. trefoil, D. dumbell with linker and trefoil, E. two tetrafoils.

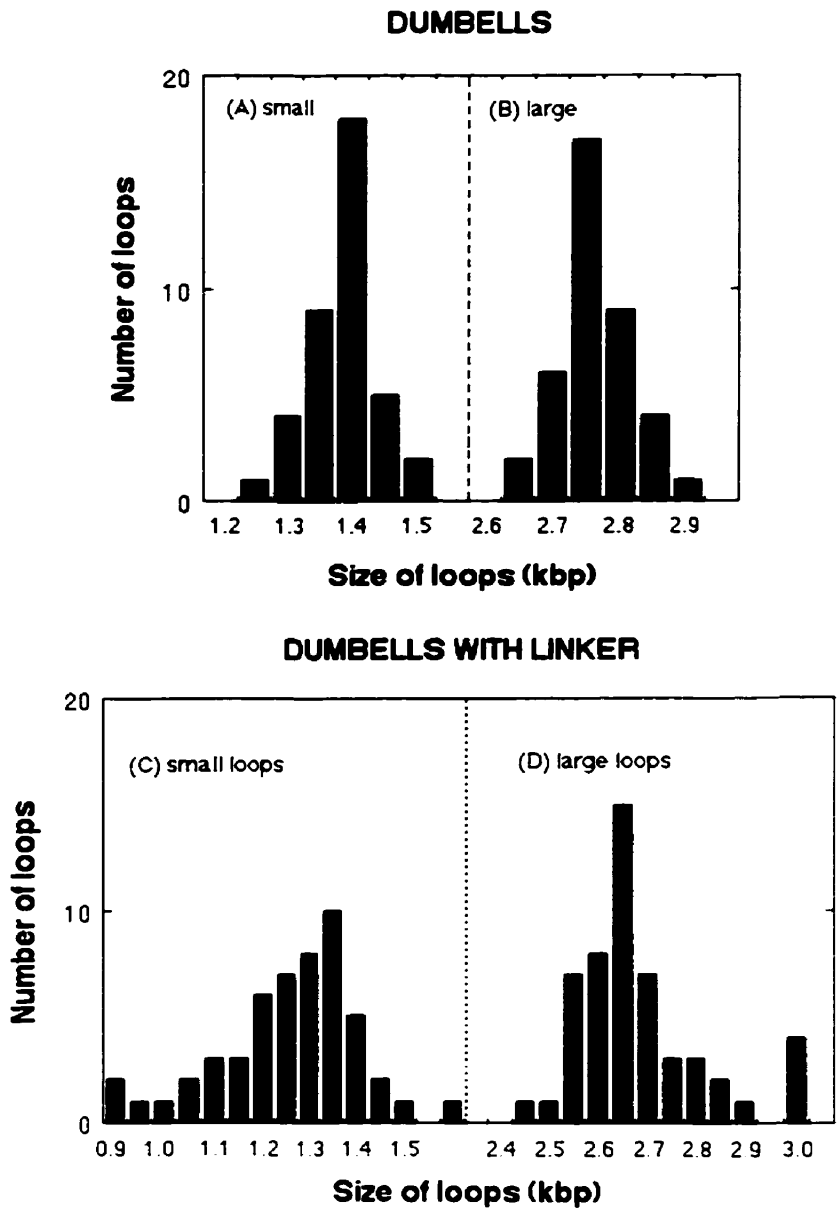


Figure 3.12. Histograms showing distribution of small and large loop sizes. kbp - kilobase pairs.

nine dumbells without a linker were measured. The size of the small loop was 1,411 ± 55 bp (average ± standard deviation) and the large loop size was 2,787 ± 56 bp [top, Figure 3.12]. By comparison, if the join were to occur at the pur•pyr tracts, the expected loop sizes would be 1,430 and 2,770 bp. The results of the histogram are therefore in agreement with the join occurring between the pur•pyr tracts. In most dumbells, a linker of variable length was present between the two loops. Fifty-nine dumbell molecules with a linker were measured and the linker averaged 112 ± 47 bp. The small loop was 1,304 ± 191 bp and the large loop was 2,679 ± 186 bp [bottom, Figure 3.12].

About 25% of the structures in the faster moving T-loop band were trefoils (Figure 3.11 D). Trefoil loop size was more variable. However, 25 were measured and 12 trefoils had a small loop size of 1,394 ± 41 bp and 13 had a large loop size of 2,712 ± 121 bp. Tetrafoils were also occasionally seen (Figure 3.11 E). As with the dumbells, the loops in the trefoils and tetrafoils originated from a single point.

The nonrandom distribution of the loop sizes, the presence of the dumbell linker, and the fact that the loops in the trefoils or tetrafoils joined at a single point is not consistent with T-loops being supercoiled molecules. The model in Figure 3.13 shows how the loops seen in the electron microscope could arise at pH 4, and how the structures become trapped at pH 8. The C loop arises when a transmolecular triplex forms at pH 4 between one of the duplex pur•pyr tracts and the single pyrimidine strand of the complementary pur•pyr tract. The duplex loop X, which becomes the third loop in a trefoil if large enough, is formed by the duplex pur•pyr tract winding about the single pyrimidine strand from the complementary tract. X turns once through the single-stranded bubble for every turn of triplex. If X remains small, it is visible in the electron microscope as the dumbell linker. If the triplex contains more than two helical turns (i.e. the duplex winds through the bubble more than twice), a fourth loop designated Y may become visible. Necessarily, the size of V varies with the size of X, and Y if it is present. When the pH is increased to 8, the triplex falls apart, but the loops remain as they are trapped in position by a partial reforming of the duplex, marked Z. This structure at pH 8, referred to as a hydrogen-bonded knot, retains single-stranded regions, explaining the sensitivity of the T-loops to P1 nuclease.

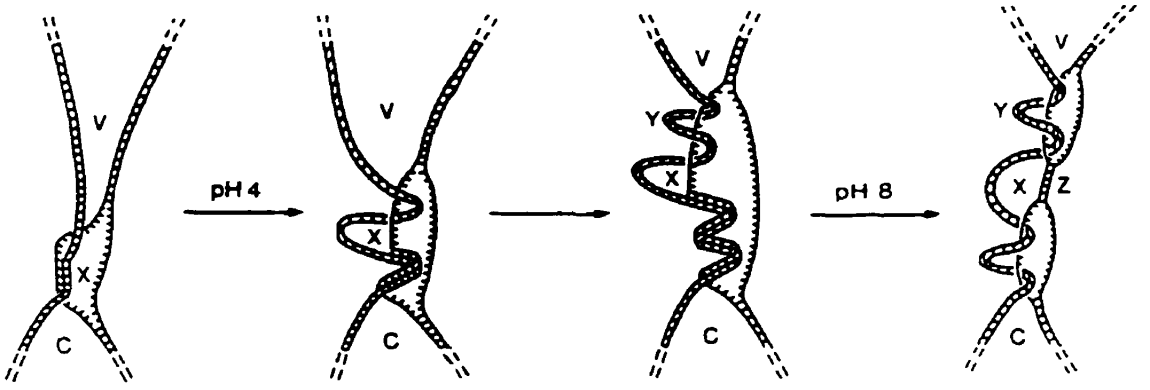


Figure 3.13. Model of T-loop formation at pH 4 and the hydrogen-bonded knot at pH 8. Loops: C-constant, V-variable, X-additional trefoil loop, Y-additional tetrafoil loop. Z-partial duplex reformed at pH 8.

3.2.2 *Escherichia coli* SSB Protein and T-Loop Stability

It has been speculated that single-stranded DNA-binding proteins may stabilize a triplex by binding to the extruded single strand (Lee *et al.*, 1984). On the other hand, *E. coli* single-stranded DNA binding protein, SSB, has been shown not only to bind to single-stranded regions in unwound DNA, but to increase the size of the denaturation bubble (Glikin *et al.*, 1983; Langowski *et al.*, 1985). It has also been shown to destabilize the triplex in H-DNA by increasing the size of the denatured region upon binding to the extruded strand (Klysik and Shimizu, 1993). As described above, T-loops are a model for how transmolecular triplexes may form *in vivo* between separated pur•pyr tracts in chromosomal loops of DNA. Because T-loops were shown to contain a single-stranded region, the effect of SSB on T-loops was investigated next. The results would reflect the effect that single-stranded DNA-binding proteins might have on the stability of transmolecular triplexes should they form *in vivo*.

If *E. coli* SSB protein had a visible effect in stabilizing or destabilizing T-loops, it would result in an increase or decrease in the size of the T-loop bands on a gel, compared to controls. No such effect of SSB protein on the size of T-loop bands was seen in these experiments. T-loop band sizes were not significantly different when SSB protein was included in the incubation mix forming T-loops, compared to controls where T-loops were made as usual, or when the incubation mix contained SSB dilution buffer (not shown). When preformed T-loops were reincubated with SSB protein, again there was no difference in the size of the T-loop bands compared to controls (not shown). Therefore, the presence of SSB protein had no observable effect on the stability of the triplex at the root of the T-loop structure at pH 4, nor did it have an obvious effect on the stability of the T-loop hydrogen-bonded knot at pH 8.

3.2.3 Transcription in T-Loops

The melting of the duplex DNA strands by RNA polymerase as it enters the promoter region has been described as a rate limiting step in the initiation of transcription (Mangel and Chamberlin, 1974). It has therefore been proposed that a pre-existing single-stranded region in the promoter, such as that which might arise from triplex formation between promoter pur•pyr tracts, could facilitate transcription (Lee *et al.*, 1984). T-loops were shown to contain a single-

stranded region at pH 8 (Figure 3.10). According to the T-loop model, this is located at the site of the hydrogen-bonded knot (Figure 3.13). In the next set of experiments, T-loops were used as a model system to investigate whether the single-stranded region could promote transcription. If transcription were favored in T-loop plasmid forms as opposed to other plasmid forms such as open circular or supercoiled, and if T-loops are a valid model for the formation of transmolecular triplexes *in vivo*, then this could be one way in which transcription is turned on in the cell.

This idea was investigated using T-loop, open circular, linear, and supercoiled plasmid forms isolated from pH 8 agarose preparative gels as described in the Materials and Methods. The presence of the hydrogen-bonded knot at pH 8 constrained the T-loops, preventing large scale reversion to the open circular form during isolation. This meant that a higher percentage of T-loops were isolated intact from the T-loop band in a pH 8 gel than from a pH 4 gel. Even so, in spite of keeping the forms on ice and in the dark as much as possible during isolation, it was not possible to isolate the T-loops without a small percentage reverting to the open circular form. About 10% of the DNA in isolated T-loops was present as open circular plasmid when re-analyzed on an agarose gel (lane 1, Figure 3.14). About the same percentage of linear forms were also present. These must have arisen from nicking of the open circular forms across from the first nick during the isolation. However, the majority of the plasmid DNA was in the form of T-loops. The T-loop band was not very large in the T-loop control (lane 2, Figure 3.14). However, it is sufficient for identifying the T-loop band in lane 1. The isolated supercoiled forms must have also undergone some nicking: about 10% of the forms were present as open circular plasmid DNA when rerun on a gel (lane 7, Figure 3.14).

In vitro transcription experiments were carried out using the isolated plasmid forms and both *E. coli* RNA polymerase core enzyme (Figures 3.15) and holoenzyme (Figure 3.16). The core enzyme lacks σ^{70} which specifically directs the enzyme to initiate transciption at bacterial promoter sequences. Transcription in these experiments was measured as TCA-precipitable counts based on the incorporation of ^{35}S-UTP. With one exception, less transcription (measured as % counts incorporated) occurred in all plasmid forms when the core enzyme was used (Figure 3.15) than with the holoenzyme (Figure 3.16). The exception was the linear plasmid form where slightly more than 5% of counts were incorporated using 10 units of core enzyme and slightly less than

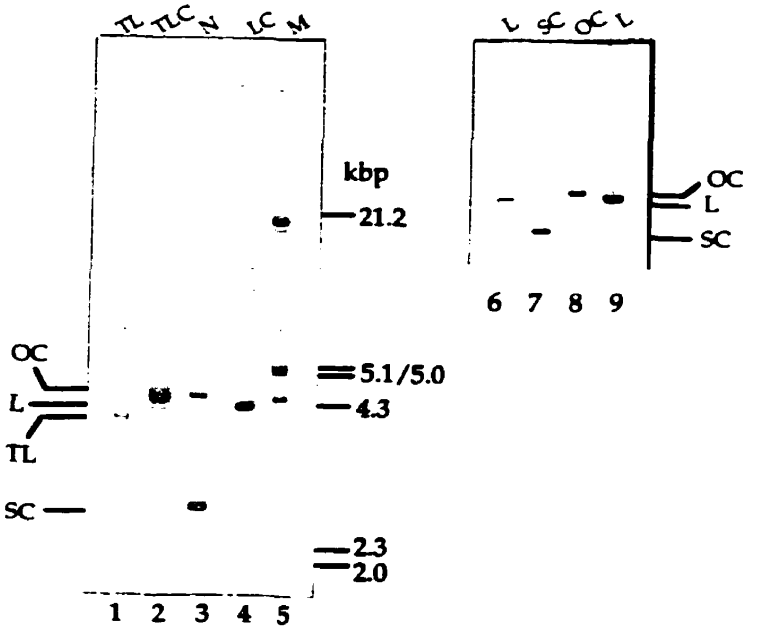


Figure 3.14. Isolated plasmid forms used in the transcription experiments. Plasmid forms added to each lane are indicated at the top of the gel. Plasmid forms present in each lane are indicated beside the gel where they occur. Abbreviations: T-loop (TL), open circular (OC), linear (L) supercoiled (SC), λ plus *Hin* dIII and *Eco* RI MW marker (M), native plasmid cut with *Cla* I as a linear control (LC), T-loop preparation as control (TLC), native plasmid preparation (N). Quantities of DNA added: lanes 1-3 (1.25 μg), lanes 6-9 (0.43 μg). pH 8, 0.85% agarose gel.

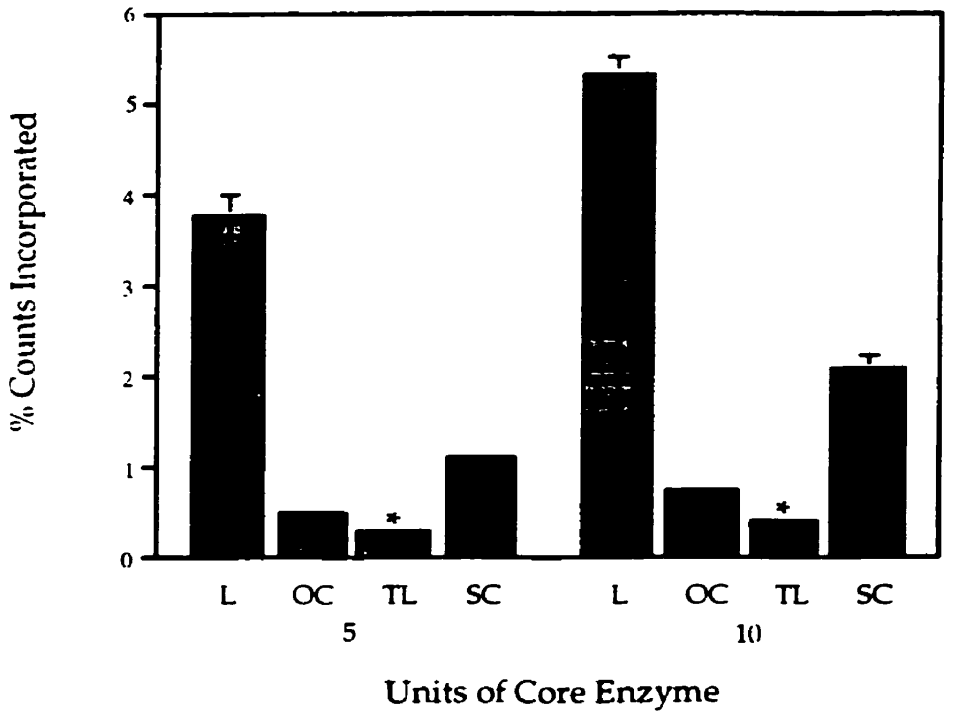


Figure 3.15. Transcription in plasmid forms using *E. coli* RNA polymerase core enzyme. Plasmid forms : linear (L), open circular (OC), T-loop (TL), supercoiled (SC). Bars are mean % counts incorporated ± SEM. * - $p<0.001$ by Student's t-test, compared with mean for OC form. Number of assays on which the means are based: L - 6; OC, TL, and SC - 7. Counts were adjusted to reflect background counts (0.01% of total counts).

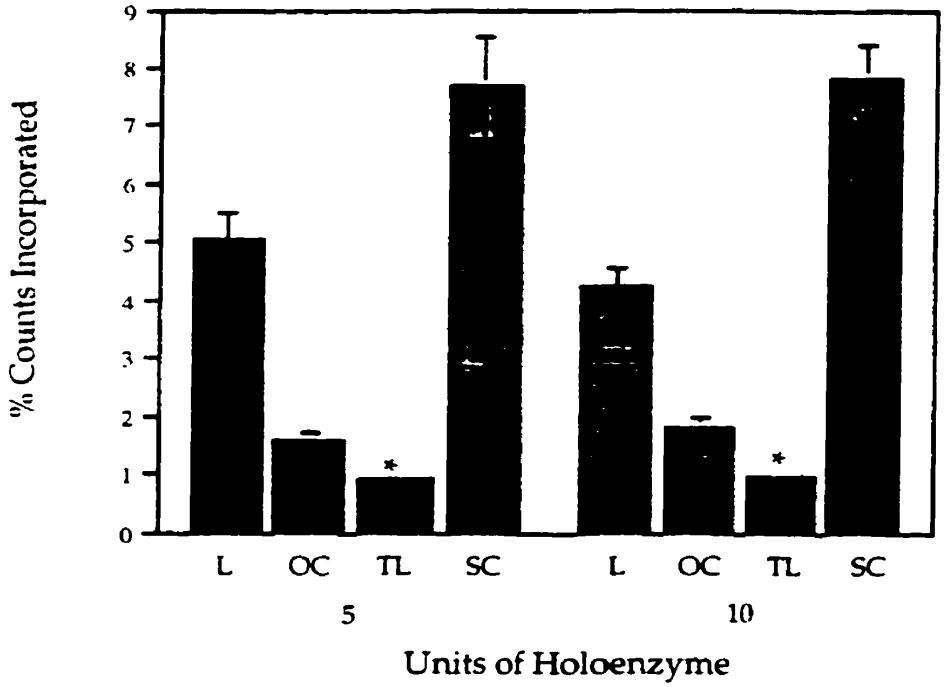


Figure 3.16. Transcription in plasmid forms with *E. coli* RNA polymerase holoenzyme. Abbreviations and explanations are the same as in Figure 3.14. Number of assays on which the counts are based: using 5 units of holoenzyme - 3 (L, SC), 7 (OC, TL); using 10 units of holoenzyme - 6 (OC, TL), 4 (L), 3 (SC). Counts were adjusted for background (0.02% of total counts).

5% of counts were incorporated using the same amount of holoenzyme. This apparent reversal of the trend of higher transcription using the holoenzyme was likely statistical variation. It is expected that higher counts would be obtained using the holoenzyme due to the presence of the σ factor. Of great interest was the observation that significantly less transcription occurred from T-loop plasmid forms than from open circular, linear, or supercoiled plasmid forms. Less than 1% of counts were incorporated using 5 and 10 units of both *E. coli* RNA polymerase core (Figure 3.15) and holoenzyme (Figure 3.16). Table 3.1 compares the concentrations of RNA produced in the assays using the T-loop and open circular plasmid forms. As the T-loops differed from the open circular forms in the presence of the hydrogen-bonded knot, this suggests that the knot in some way blocked transcription.

The greatest amount of transcription occurred in the supercoiled plasmid form using the holoenzyme (7.5% of counts incorporated; Figure 3.16). This was very much more than using the open circular forms and higher than in the linear. The difference between the open circular and the supercoiled forms would have been the single-stranded nick and the lack of supercoiling in the open circular form. This shows the importance of supercoiling in facilitating the opening of promoter regions by the holoenzyme in initiating transcription. The free energy of supercoiling is thought to assist RNA polymerase in locally unwinding and denaturing duplex DNA in the initiation of transcription (Parvin and Sharp, 1993; Wang and Lynch, 1993). A significantly higher level of transcription also occurred in the supercoiled forms than the open circular forms when core enzyme was used (Figure 3.15). Although overall transcription levels were much lower with the core enzyme, this suggests that there was a difference between the supercoiled and open circular plasmid forms which promoted transcription by the core enzyme. The initiation of transcription must have been non-specific as the core enzyme lacked the sigma factor. As it was likely due to the presence of single-stranded regions in the supercoiled forms, it provides evidence that single-stranded regions can non-specifically promote entry of RNA polymerase.

Again using the core enzyme, the highest level of transcription occurred with the linear form (Figure 3.15). This would have been due to non-specific initiation of transcription from both ends of the duplex DNA strands (Vogt, 1969). With the holoenzyme, levels of transcription were lower from the linear

Table 3.1. RNA produced per nM of DNA template in transcription assays. [a]

	Enzyme (units)	
	5	10
Core Enzyme		
T-loops [b]	1.0 nM [c]	1.4 nM
Open circular	1.6 nM	2.6 nM
Holoenzyme		
T-loops	2.6 nM	2.9 nM
Open circular	5.2 nM	5.8 nM

[a] Each 40 µl assay contained 1.55 nM of DNA.
[b] plasmid form
[c] Concentrations were calculated from TCA-precipitable counts and specific activity of the incorporated ^{35}S-UTP.

than from the supercoiled form (Figure 3.16). This again shows the importance of supercoiling in facilitating the entry of RNA polymerase holoenzyme to promoter sites. However, some non-specific initiation of transcription by the holoenzyme must have occurred at the linear duplex ends of the plasmid, otherwise the rate of transcription in the linear plasmid forms would have been as low as in the open circular forms. Linear forms of the plasmid pKHa3PYL/PUL were created by cleavage at the *Cla* I site. pKHa3PYL/PUL was created by the insertion of a fragment of M13mp19 containing the *Cla* I site into a pre-existing plasmid. Although the M13mp19 gene III is present in this insert, its promoter is located approximately 1 kb away from the *Cla* I site. Therefore the proximity of a gene promoter to one end of the linear plasmid forms is not a likely reason for transcription being higher in the linear than in the open circular forms, in transcription using both the core and the holoenzymes.

In conclusion based on these results, if the T-loop model is an accurate representation of transmolecular triplex formation *in vivo*, the chromosomal structures formed would be expected to downregulate transcription rather than facilitate expression from a particular promoter.

3.2.4 T-Loop Formation Using Polyamine Analogues

The naturally-occurring polyamines spermine and spermidine are present in the cell in millimolar concentrations, with the majority being bound to nucleic acids (Sarhan and Seiler, 1989). A major role of polyamines is in the condensation of DNA through charge neutralization (Gosule and Schellman, 1976; Thomas and Bloomfield, 1983b). Polyamines are thought to promote triplex formation by shielding the negative charges on the three DNA strands permitting the approach and binding of the third strand in the major groove of the duplex. TFOs which bind to pur•pyr tracts in the regulatory regions of genes are being investigated as a way of regulating gene expression in antigene therapy. It is desirable for such oligonucleotides to have a very high binding constant while retaining specificity for their target sequences. The use of polyamines as secondary ligands has been proposed as a way of increasing the triplex-forming potential of oligonucleotides (Thomas and Thomas, 1993; Thomas *et al.*, 1995).

In this final set of experiments using T-loops, a number of polyamine analogues were compared with spermine and spermidine in their ability to promote triplex formation. The polyamine analogues differed both in their charge ($^{+}4$ or $^{+}5$) and in the charge separation (Table 3.2). The ability of the polyamines to promote triplex formation was tested by substituting them for spermine in incubation reactions to form T-loops. The appearance of the T-loop bands on agarose gels were then analyzed. Figure 3.17 is a representative set of gels comparing the ability of two polyamine analogues, 4-4-4-4 and 3-3-3-3, to form T-loops with that of spermine and spermidine. Spermine strongly promoted the formation of T-loops: at 30 μM approximately 50% of the open circular forms were converted to the faster mobility form (lane 5, Figure 3.17; Table 3.3). More than 10 times the amount of spermidine was required for an equivalent formation (lane 5, Figure 3.17; Table 3.3). The polyamine analogue 4-4-4-4 promoted the greatest conversion to T-loops and at the lowest concentration of polyamine: some T-loops were visible at only 10 μM (lane 4, Figure 3.17). At the optimum concentration of 16 μM, the majority of the open circular form was converted to T-loops (lane 6, Figure 3.17; Table 3.3). Conversion appeared at an even lower concentration of the other polyamine analogue, 3-3-3-3 (4 μM; lane 4, Figure 3.17). However, formation of T-loops with 3-3-3-3 was very poor, appearing as a faster running smear ahead of the linear band. At 8 μM and above the DNA precipitated and did not enter the gel (lane 6, Figure 3.17). Based on such gels, Table 3.3 lists for each polyamine the concentration range over which T-loops formed, the optimum concentration for formation, and the percent conversion of open circular forms to T-loops. Spermine, with a charge of $^{+}4$ has previously been reported as being superior in promoting triplex formation to spermidine with a lower charge of $^{+}3$ (Hampel et al, 1991). The results obtained here support this trend: total charge was the most important factor in the efficacy of the polyamines in promoting triplex formation. Optimum concentrations at which pentamines (e.g. 4-4-4-4) with a charge of $^{+}5$ formed T-loops was lower than the tetramines (e.g. 4-4-4) with a total charge of $^{+}4$ (Table 3.3). Spacing between the charges was also important: polyamines with trimethylene bridging (e.g. 3-3-3) were able to promote T-loop formation at a lower concentration than molecules with tetramethylene bridging (e.g. 4-4-4). Finally, bisethylation generally resulted in slight increase in the minimum concentration at which T-loops were formed. Therefore, the presence of the ethyl groups seemed to produce some steric hindrance.

Table 3.2. Polyamines used to test triplex formation.

Chemical Structure	Abbreviation
$H_3N^+(CH_2)_3N^+H_2(CH_2)_4N^+H_3$	spermidine
$H_3N^+(CH_2)_3N^+H_2(CH_2)_4N^+H_2(CH_2)_3N^+H_3$	spermine (3-4-3)
$EtN^+H_2(CH_2)_3N^+H_2(CH_2)_4N^+H_2(CH_2)_3N^+H_2Et$	BEspermine (BE-3-4-3)
$H_3N^+(CH_2)_3N^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_3$	3-3-3
$EtN^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_2Et$	BE-3-3-3
$H_3N^+(CH_2)_3N^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_3$	3-3-3-3
$EtN^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_2(CH_2)_3N^+H_2Et$	BE-3-3-3-3
$H_3N^+(CH_2)_4N^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_3$	4-4-4
$EtN^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_2Et$	BE-4-4-4
$H_3N^+(CH_2)_4N^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_3$	4-4-4-4
$EtN^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_2(CH_2)_4N^+H_2Et$	BE-4-4-4-4

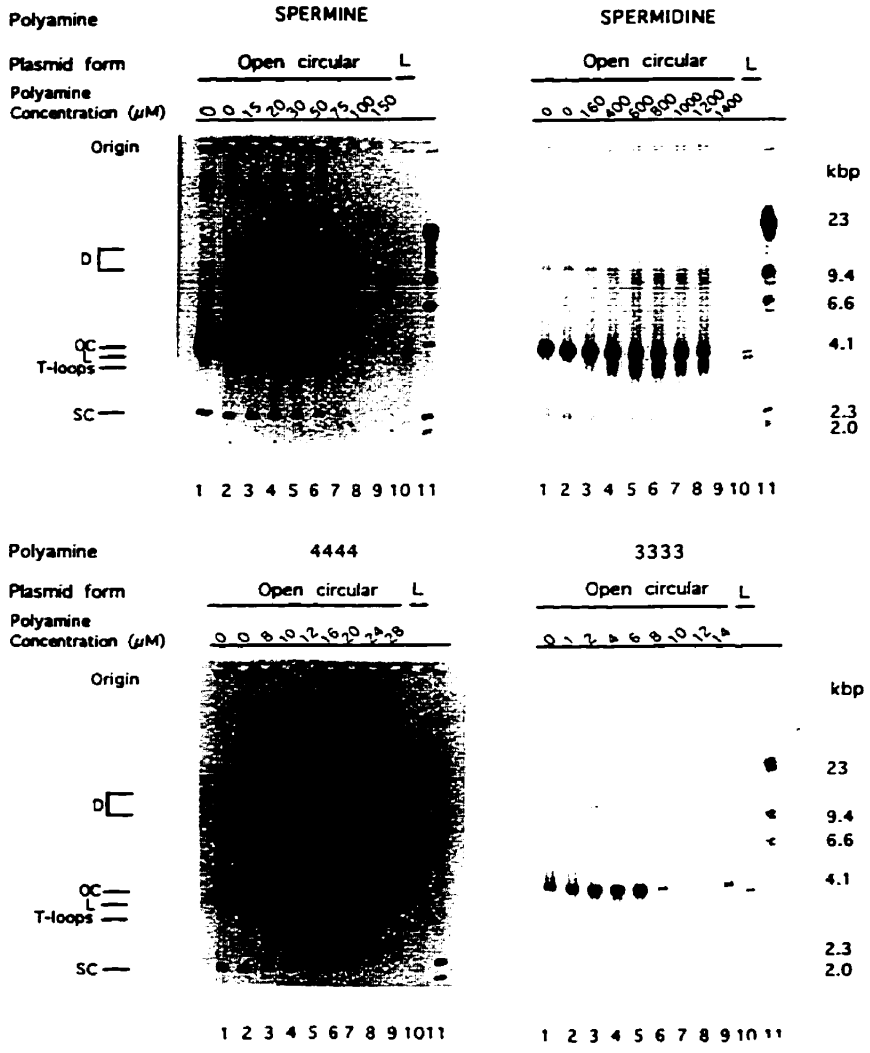


Figure 3.17. T-loop formation with spermine, spermidine, and the polyamine analogues 4-4-4-4 and 3-3-3-3. pKHa3PYL/PUL (γ-irradiated) was incubated at pH 4 with the polyamine concentrations at the top of the gel lanes. Positions of plasmid forms in lanes are indicated: D - dimers, OC - open circular, L - linear, T-loops, SC - supercoiled.

Table 3.3. Polyamine concentrations for T-loop formation.

	Concentration (μM)			Percent
Polyamine	Minimum	Maximum	Optimum	Conversion to T-Loops [a]
Spermidine	400	1,200	600	50
Spermine	30	50	30	50
BE-spermine	20	60	35	40
3-3-3	20	50	40	10
BE-3-3-3	20	70	60	30
3-3-3-3	4	4	4	<10
BE-3-3-3-3	10	10	10	~5
4-4-4	25	50	50	30
BE-4-4-4	30	80	30	20
4-4-4-4	10	28	20	90
BE-4-4-4-4	12	28	20	50

[a] Estimated by comparing size of open circular band when polyamine not included in incubation mix with T-loop band size following incubation with optimum concentration of polyamine.

3.3 Triplex-Specific Antibodies and Gene Expression in Agarose-Encapsulated and Permeabilized HT-29 Nuclei

In previous work in this laboratory, Agazie *et al.* (1996) incubated agarose-encapsulated MOPC nuclei with the triplex-specific antibodies, Jel 318 and Jel 466. Both antibodies were found to inhibit overall transcription by about 20%, measured by TCA-precipitable counts from run-off transcription. By comparison, when the nuclei were incubated with the control antibody, Jel 42, specific for a bacterial protein, there was no reduction in transcription over incubation with BSA. These results suggested that triplexes were involved in gene expression.

Agazie *et al.* (1996) used agarose-encapsulated nuclei as the chromatin most closely resembles that in the cell (Jackson and Cook, 1985a; Jackson *et al.*, 1988). Additionally, the encapsulated nuclei can be transferred to different buffers and contain pores large enough to permit the ready diffusion of antibodies (Jackson *et al.*, 1988). The encapsulated nuclei are permeabilized during preparation, removing all macromolecular complexes except chromatin surrounded by cytoskeletal elements. The DNA template, Pol II, and nascent RNA remain associated with the nucleoskeleton. The transcript elongation activity of Pol II is interrupted, but the polymerases retain their ability to transcribe (Jackson and Cook, 1985b).

No measurable initiation of transcription by Pol II has been detected in unencapsulated nuclei prepared conventionally or in agarose-encapsulated nuclei (Marzluff, 1978; Jackson and Cook, 1985a; Jackson *et al.*, 1988). However, whether or not some reinitiation occurs during run-off transcription is a point of controversy (Ausubel *et al.*, 1995). Another reason why agarose-encapsulated nuclei were used by Agazie *et al.* (1996) is that it was thought there may be some reinitiation of transcription in these nuclei. This was because the protocol followed by Agazie *et al.* (1996) was based on procedures where transcription was carried out for 2 h (Jackson *et al.*, 1988). In comparison, conventional run-off transcription is only carried out for 30 min (Ausubel *et al.*, 1995).

Formation of triplexes in pur•pyr tracts in the 5' regions of genes might either inhibit or stimulate transcription. Therefore, because of the reduction in overall transcription following incubation with triplex-specific antibodies observed by Agazie *et al.* (1996), the next set of experiments was undertaken to determine if incubation with triplex-specific antibodies had an effect on

transcription from specific genes containing pur•pyr tracts in their promoter regions. C-*myc* was chosen as it is expressed at high levels in the human colon carcinoma cell line HT-29 (K. Bonham - personal communication). It also has a well-characterized 5' pur•pyr tract located close to sequences important in transcriptional regulation. Of the greatest interest is that triplex-destabilizing mutations in the pur•pyr tract cloned in plasmids led to decreased transcription from a reporter gene (Firulli *et al.*, 1994). If a triplex structure forms in the c-*myc* pur•pyr tract *in vivo*, it may be involved in transcriptional regulation of the gene. Therefore, experiments on the effect of the triplex-specific antibodies on c-*myc* transcription were carried out using agarose-encapsulated HT-29 nuclei.

3.3.1 Transcription in Hypotonically-Prepared HT-29 Nuclei

Run-off transcription experiments incorporating ^{32}P were initially done using HT-29 nuclei isolated under hypotonic conditions. This was for the purpose of comparison with later results using agarose-encapsulated nuclei, and to become familiar with run-off transcription procedures labelling RNA with ^{32}P-UTP. Total counts per min (CPM) in resuspended RNA pellets following run-off transcription in hypotonically-prepared nuclei ranged from 1.2×10^5 to 4.2×10^5 when ^{32}P-UTP was used within two weeks of the reference date (data not shown). In a densitometry scan of a blot of RNA from run-off transcription, the ratio $\frac{\text{c-}myc\text{ mRNA}}{\text{18S rRNA}}$ was reduced by about 50% following incubation with Jel 466 compared to incubation with Jel 42 (Table 3.4). The gene for 18S rRNA was used as a control as it is expressed at extremely high levels in the cell and does not contain a pur•pyr tract in its 5' region, although it is transcribed by RNA polymerase I rather than Pol II. In contrast, the ratio $\frac{\text{c-}myc\text{ mRNA}}{\text{18S rRNA}}$ increased about 20% following incubation with Jel 318 compared with Jel 42. Following incubation with Jel 466, the $\frac{\text{ß-actin mRNA}}{\text{18S rRNA}}$ ratio was also reduced compared with that following incubation with Jel 42 (last column, Table 3.4). This was interesting as ß-actin contains a pur•pyr sequence in its 5' region which may be able to form a triplex. This latter ratio did not increase following incubation with Jel 318 compared with Jel 42. A plasmid containing the c-*src* gene which has 5' pur•pyr tracts, was also use as a probe in analyzing RNA from run-off transcription. However, no band was visible on the

Table 3.4. Gene expression in hypotonically-prepared HT-29 nuclei following incubation with triplex-specific antibodies.[a]

	gene transcribed [b]			ratios	
Antibody	c-*myc*	ß-actin	18S rRNA	c-*myc* mRNA / 18S rRNA	ß-actin mRNA / 18S rRNA
Jel 318	6.81×10^7	1.56×10^6	4.10×10^5	166	3.80
Jel 466	4.40×10^7	1.55×10^6	7.84×10^5	56.1	1.98
Jel 42	6.80×10^7	3.14×10^6	5.12×10^5	133	6.13

[a] Nuclei (>10^7) were incubated with 30 µg of antibody in a total volume of 220 µl for 3.75 h at R.T. Following this, run-off transcription was carried out in a total volume of 420 µl for 30 min at R.T. Values shown are from a single experiment.

[b] Values are arbitrary units representing areas under peaks from densitometry scanning of an autoradiogram. The autoradiogram was produced by hybridizing ^{32}P-UTP-labelled RNA from run-off transcription to gene probes bound to a membrane.

autoradiogram. Therefore, c-*src* was expressed at too low a level in these cells to be of any use. A 50% decrease in transcription of c-*myc* following incubation with Jel 466, compared with Jel 318 and Jel 42, was found in an additional experiment. However, mistakenly, nuclei were incubated in a higher concentration of Jel 466 (34.3 µg) than for the other antibodies (Jel 318 - 22.9 µg; Jel 42 - 15.3 µg) (data not shown). Hypotonically-prepared nuclei were not used in further experiments as it was thought that better results would be obtained using encapsulated nuclei as the pores in the agarose would allow the antibodies to diffuse freely to the DNA template in the nucleus.

3.3.2 Trichloroacetic Acid-Precipitable Counts

TCA-precipitable counts were carried out following run-off transcription in agarose-encapsulated HT-29 nuclei as done by Agazie *et al.* (1996) using MOPC nuclei. This was to determine if the decrease in TCA-precipitable counts observed with MOPC nuclei was reproducible in the HT-29 nuclei. Compared with counts following incubation with BSA, a 75% decrease in TCA-precipitable ^{35}S counts occurred following incubation with the negative control, actinomycin D (Table 3.5). The antibiotic actinomycin D inhibits elongation of RNA chains by intercalating into the DNA helix, preventing the movement of RNA polymerase along the template. A reduction in TCA-precipitable counts of 63% and 66%, respectively, occurred following incubation of agarose-encapsulated HT-29 nuclei with the triplex-specific antibodies Jel 318 and Jel 466. However, an 18% reduction in transcription also occurred with the control antibody, Jel 42. The decreases following incubation with the triplex-specific antibodies were considerably higher than the 20% observed by Agazie *et al.* (1996). Additionally, these authors did not observe a decrease in transcription following incubation with the control antibody, Jel 42.

3.3.3 Gene Expression

The next steps were to extract RNA from the encapsulated HT-29 nuclei following run-off transcription, then determine the level of expression of individual genes. It was not possible to isolate any RNA from agarose-encapsulated HT-29 nuclei following transcription without using Trizol reagent and strict precautions against RNase contamination. RNA isolated under these

Table 3.5. TCA-precipitable ^{35}S counts following run-off transcription in agarose-encapsulated HT-29 nuclei.

Incubation[a]	CPM	Percent of maximum CPM[b]
Actinomycin D	889	25
BSA	3,526	100
Jel 42	2,884	82
Jel 318	1,314	37
Jel 466	1,204	34

[a] Agarose-encapsulated nuclei (500 µl packed volume) were incubated with 120 µg of BSA and the three antibodies, or 7.5 µg of actinomycin D, for 2 h at R.T in a total volume of 500 µl. Run-off transcription was then carried out for 2 h at 37 ^{o}C.

[b] Following transcription, 100 µl of suspended agarose-encapsulated nuclei were mixed with 250 µl of 2% SDS and spotted onto a filter prior to TCA precipitation under vacuum. Counts on the filter were determined in a scintillation counter. CPM were expressed as a percentage of the counts following incubation with BSA. Values shown are from a single experiment.

conditions from encapsulated nuclei after run-off transcription showed 18S and 28S rRNA bands which were somewhat degraded on an RNA agarose gel (not shown). Some degradation of the RNA was acceptable as it would improve hybridization with DNA on the membrane strips (Ausubel *et al.*, 1995).

The run-off transcription experiments of Agazie *et al.* (1996) used 2.5 x 10^6 agarose-encapsulated nuclei (250 µl packed volume) in a total volume of 1 ml per assay. As Ausubel *et al.*, (1995) suggest a minimum of 5 x 10^6 nuclei for a successful run-off transcription assay, the number of agarose-encapsulated nuclei in these experiments was increased to 5 x 10^6. This gave a packed volume of 500 µl, still using a total final volume of 1 ml.

Following run-off transcription incorporating ^{32}P-UTP and extraction of the RNA from the agarose microbeads, the RNA pellet was resuspended in a volume of 300 µl. Total CPM in resolubilized RNA pellets were more variable than those from hypotonically-prepared nuclei (above) and ranged from 6.8 x 10^4 to 3.4 x 10^6 (Table 3.6). The average total count, however, was 1.1 x 10^4 which was more than ten-fold lower than the average counts from hypotonically-prepared nuclei (Section 3.3.1). Use of carrier yeast tRNA did not improve counts. Overall, Table 3.6 shows a trend towards decreased counts following incubation with increasing concentrations of both the triplex-specific antibodies and with Jel 42. A similar decrease in counts was not seen with increasing concentrations of BSA: incubating agarose-encapsulated nuclei with 750 µg of BSA caused no significant decrease in counts over no pre-incubation with a protein (row 1, Table 3.6; other results not shown).

Following the run-off transcription incorporating ^{32}P-UTP, RNA was isolated from the nuclei and used to probe the genes for c-*myc* and 18S rRNA, and the DNA repair gene, MGMT. This latter gene was a good control as it lacked a pur•pyr tract in its promoter region and was transcribed by Pol II at very high levels in the HT-29 cells. Because RNA quantities added to hybridization bags were adjusted so that equal counts were added to each bag, enough RNA was often available for hybridization to more than one slot blot strip, particularly following incubation with BSA. This provided data for limited statistical analysis. A PhosphorImager was used to quantitate the radioactivity on the membrane strips following hybridization with the ^{32}P-labelled RNA. One useful property of this technique is that exposure time is about one-tenth that of X-ray film autoradiography. Other important advantages are that the linear dynamic range is 40 times that of film and the

Table 3.6. ^{32}P CPM in resolubilized RNA pellets following incubation with BSA or antibodies and run-off transcription.

Incubation		Experiment Number [a]				
Protein	$\frac{\mu g}{ml}$	16	17	18	19	20
None						$1.4x10^5(102.9)$
	120					
BSA	300	$3.4x10^6(100)$				
	500		$3.0x10^5(100)$	$7.7x10^5(100)$		
	750				$2.4x10^5(100)$	$1.4x10^5(100)$
	120	$2.8x10^6(83.6)$				
Jel	300	$1.4x10^6(42.1)$	$2.3x10^5(74.5)$		$1.4x10^5(56.1)$	$9.4x10^4(67.1)$
42	500	$2.1x10^6(62.1)$	$1.9x10^5(61.0)$	$5.8x10^5(74.5)$		
	750				$7.3x10^4(30.0)$	$7.6x10^4(54.2)$
	120	$1.4x10^6(42.9)$				
Jel	300	$1.2x10^6(35.7)$		$7.1x10^5(92.3)$		
318	500	$9.6x10^5(28.6)$	$1.4x10^5(44.4)$	$7.3x10^5(95.0)$		
	750				$6.8x10^4(28.0)$	
Jel	300		$1.8x10^5(60.3)$	$4.6x10^5(58.9)$		$1.2x10^5(86.8)$
466	500		$1.3x10^5(43.7)$	$4.2x10^5(54.9)$		
	750					$1.7x10^5(118.2)$

[a] Numbers given are total ^{32}P CPM after solubilizing the RNA pellet in 300 µl DEPC-treated ddH_2O plus 0.1 mM EDTA. Numbers in brackets are percentage of counts compared to counts obtained following incubation with BSA in the same experiment. CPM obtained following incubation with BSA or antibodies in each experiment are single values.

sensitivity is 10 to 250 times. Furthermore, analysis is simplified as results can be analyzed quantitatively by image analysis software. Therefore, in spite of the low counts in the isolated RNA, clear PhosphorImages were obtained following a two-day exposure of the Phosphor plate to the hybridized membranes. When PhosphorImage results were averaged from five experiments, a general trend was found towards decreasing levels of transcription in all three genes with increasing concentrations of both the triplex-specific antibodies and Jel 42 (Table 3.7). The largest decrease occurred following incubation with 750 μg/ml. On the other hand, no reduction in expression of any of the genes was seen following incubation with BSA, even at 750 μg/ml.

3.3.4 Transcription Assay Components and RNase Activity

To summarize, following incubation with increasing concentrations of both the triplex-specific antibodies and the control antibody Jel 42, both the total counts in the solubilized RNA pellets and the levels of expression of c-*myc* and the non-pur•pyr tract control genes decreased. As the same decrease did not occur with increasing concentrations of BSA, it was thought that the antibodies might be contaminated with RNases. For the experiments, BSA had been prepared RNase-free and the antibodies were dialyzed into RNase-free transcription buffers. However, RNases were likely present in the original antibodies due to the omnipresence of RNases in the environment. Accordingly, RNA was incubated with Jel 42, Jel 466, BSA, and transcription buffer at R.T. for 1.5 h in pH 8 fluorescence assay buffer. No decreases occurred in the fluorimeter readings over that time. In contrast, readings dropped dramatically in a control tube containing RNase A (data not shown). Therefore, there was no evidence that RNase contamination in the antibodies was responsible for the decreasing levels of RNA observed in these experiments with increasing concentrations of antibodies.

3.4 Transmolecular Triplex Formation in the C-*Src* Promoter

The 5' region of human c-*src* contains four pur•pyr tracts in series (TC1, TC1.1, TC2, and TC3; Figure 1.11). Studies using plasmid CAT constructs in transient transfection experiments showed that the tracts are located in a region of the promoter essential for maximum levels of expression (Bonham and Fujita,

Table 3.7. Effect of monoclonal antibodies on gene expression in agarose-encapsulated HT-29 nuclei.

Protein	$\frac{\mu g}{ml}$	Gene [a] ± S.D.		
		c-myc	MGMT	18S rRNA
BSA	various [b]	116.3 ± 69.8 (16)	103.9 ± 41.7 (13)	100 ± 25.7 (15)
Jel 42	120	66.4 ± 13.1 (4)	76.5 ± 9.8 (2)	66.4 ± 7.7 (4)
	300	171.0 ± 172.7 (8)	88.5 ± 34.7 (7)	108.4 ± 34.6 (7)
	500	177.1 ± 117.2 (8)	102.7 ± 28.6 (6)	92.9 ± 37.0 (6)
	750	43.1 ± 23.4 (4)	42.4 ± 28.2 (4)	52.6 ± 26.1 (4)
Jel 318	120	113.9 ± 43.1 (2)	161.1 (1)	113 ± 17.3 (2)
	300	114.5 ± 20.7 (4)	123.3 ± 47.6 (3)	90.4 ± 52.1 (4)
	500	146.5 ± 19.5 (5)	79.7 ± 63.8 (4)	90.6 ± 37.3 (4)
	750	75.7 ± 0.5 (2)	62.8 ± 22.9 (2)	71.4 ± 30.3 (2)
Jel 466	300	105.6 ± 102.7 (6)	47.8 ± 25.6 (6)	60.9 ± 40.0 (5)
	500	129.9 ± 13.0 (4)	77.4 ± 47.9 (4)	56.1 ± 3.7 (3)
	750	36.7 ± 13.2 (4)	56.5 ± 18.0 (4)	65.7 ± 10.8 (4)

[a] The level of expression of each gene is given as a percentage of the level of expression of the 18S rRNA gene in that experiment following pre-incubation with BSA. Percentages are averaged from 5 experiments and numbers in brackets are the total number of hybridizations upon which the average is based.

[b] A single concentration of BSA used in an experiment. This was the same as for the antibodies (120 to 750 μg/ml). In some cases, more than one concentration of antibody was used. However, control experiments showed that BSA concentration did not affect levels of gene expression.

1993). Sequence similarities amongst the tracts suggests that one or more pairs of tracts might be complementary and able to form a triplex *in vitro*. A similar triplex formation could form *in vivo* and be involved in the regulation of expression of the gene.

The next set of experiments used plasmids containing cloned regions of the human c-*src* promoter to investigate *in vitro* triplex formation between the three major tracts, TC1, TC2, and TC3. The approach was to identify triplex formation between pairs of tracts through dimerization upon agarose gel electrophoresis. In most experiments, this was visualized as shift in band mobility. The type of triplex investigated again was a pyr•pur•pyr triplex, as the incubation conditions were low pH in the presence of spermine.

3.4.1 pBS-Src and p0.2SrcCAT Linear Dimers

The plasmid pBS-Src contained all four c-*src* pur•pyr tracts together with the immediately surrounding sequences from the c-*src* promoter region (Figure 3.18). Any triplex formed between a pair of the c-*src* promoter pur•pyr tracts would be a transmolecular triplex rather than an intramolecular triplex (H-DNA), as the tracts were separated from each other by more than a few nucleotides. An initial investigation was therefore carried out to determine if a transmolecular triplex could form between two of the pur•pyr tracks in an open circular form of the plasmid, in the same way as it did between the two separated pur•pyr tracts in pKHa3PYL/PUL (Section 3.2.1). pBS-Src γ-irradiated to create open circular forms was incubated under the same conditions (pH 4 triplex-forming buffer and 50 to 125 μM spermine) used to form T-loops in γ-irradiated pKHa3PYL/PUL. No extra bands with a mobility intermediate between open circular/linear and supercoiled forms suggesting the presence of T-loop-like structures were seen on pH 4 or pH 8 agarose gels (not shown). However, the pur•pyr tracts were much closer together in pBS-Src: it was only 125 bp from the 5' end of TC1 to the 3' end of TC3 in a 3.3 kb plasmid (Figure 1.11). In pKHa3PYL/PUL, the complementary pur•pyr tracts were separated by 1.28 kb in a 4.2 kb plasmid (Figure 3.1). Because the tracts were so close together in pBS-Src, if a transmolecular triplex had formed between a pair of tracts, it may not have resulted in a big enough mobility change to be visible on a gel.

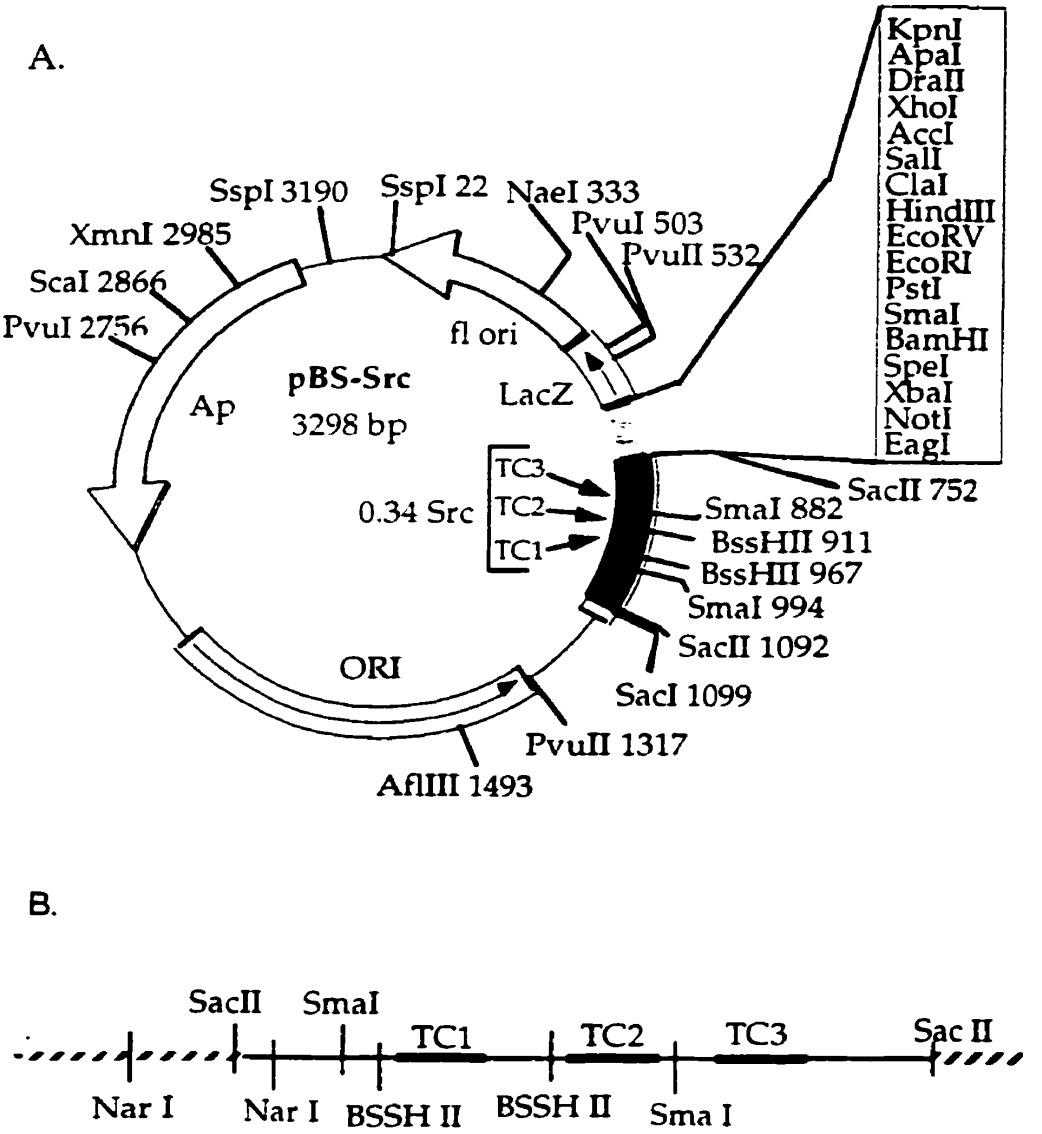


Figure 3.18. pBS-Src. A. Location of genes, restriction sites, and c-*src* promoter pur•pyr tracks. B. Details of region containing cloned c-*src* promoter fragment. Diagonally-striped flanking regions are vector sequences.

Instead, following the incubation of γ-irradiated pBS-Src at pH 4 in the presence of spermine, two new bands of slower mobility than open circular forms were seen. A band with the same mobility as the faster moving of the two new bands also appeared when pBS-Src linearized at *Sac* I upstream from the TC1 tract (Figure 3.18) was incubated with spermine at pH 4 (lane 7, Figure 3.19A). This must have been a linear dimer arising due to triplex formation between the pur•pyr tracts in pairs of linearized plasmid. The slower moving band on the gel following incubation of γ-irradiated pBS-Src at pH 4 with spermine must have contained open circular dimers.

The plasmid p0.2SrcCAT contained the c-*src* pur•pyr tracts TC2 and TC3 only (Figure 3.20). When p0.2SrcCAT cut at the unique *Sma* I site was incubated in the presence of spermine at pH 4, an additional band appeared on a pH 4 gel (lane 9, Figure 3.19A). Again, this band was thought to contain linear dimers arising from triplex formation. Another band with mobility intermediate between that of the two preceeding linear dimers appeared when linearized p0.2SrcCAT and linearized pBS-Src were incubated together with spermine at pH 4 (lane 5, Figure 3.19A). The intermediate mobility of this band suggested it was a dimer formed between the two linearized plasmids. All three bands were also visible on pH 8 agarose gels, showing that the three dimers were stable at pH 8 (not shown). The dimer formed between p0.2SrcCAT linearized at the *Sma* I site and pBS-Src linearized at *Sac* I was likely a Y dimer. The Y junction could have contained either TC2 or TC3 of p0.2SrcCAT. This would have been similar to Y dimers formed between mixtures of p3PU and p3PY, plasmids related to pKHa3PYL/PUL but containing only a single pur•pyr tract in each plasmid (Hampel *et al.*, 1993). In those experiments, Y dimers formed between one plasmid linearized with the pur•pyr tract two nucleotides from an end, and the other plasmid linearized so the tract was in the middle. These Y dimers disappeared upon electrophoresis at pH 7. The Y dimers in the present study were stable at pH 8. It was thought that the Y dimers in the study of Hampel *et al.* (1993) lacked stability at pH 7 because the pur•pyr tract was located right at the junction. Hence, there was no topological trapping to prevent unwinding of the duplex end from the triplex as the pH was raised. The *Sma* I site in p0.2SrcCAT is located immediately at the end of TC2 and only 12 nucleotides from the end of TC3. However, the 12 nucleotide arm beyond TC3 must have been long enough for topological trapping to occur when the pH was increased to 8. Therefore, TC3 may have been the pur•pyr tract in the triplex.

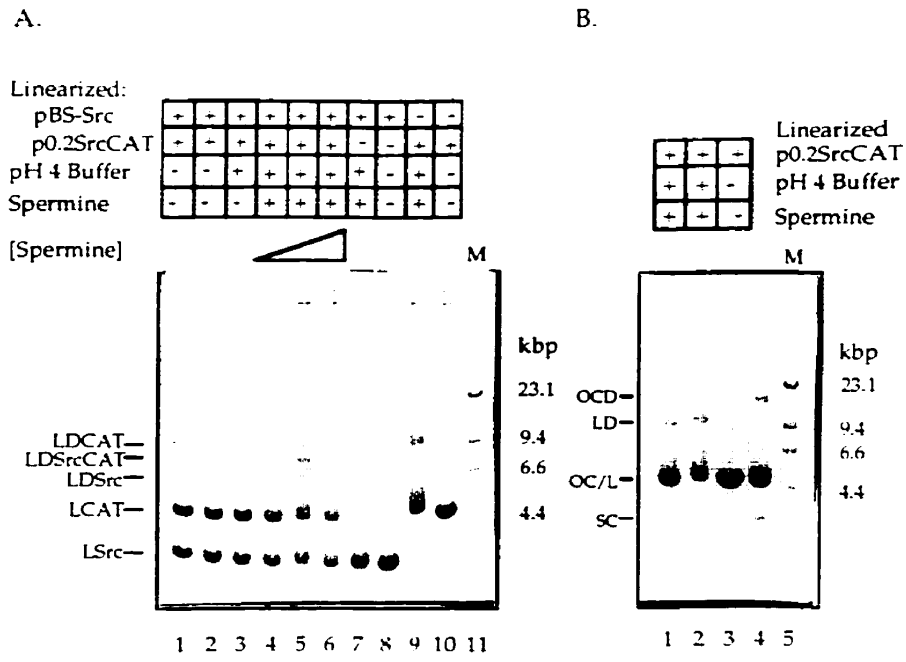
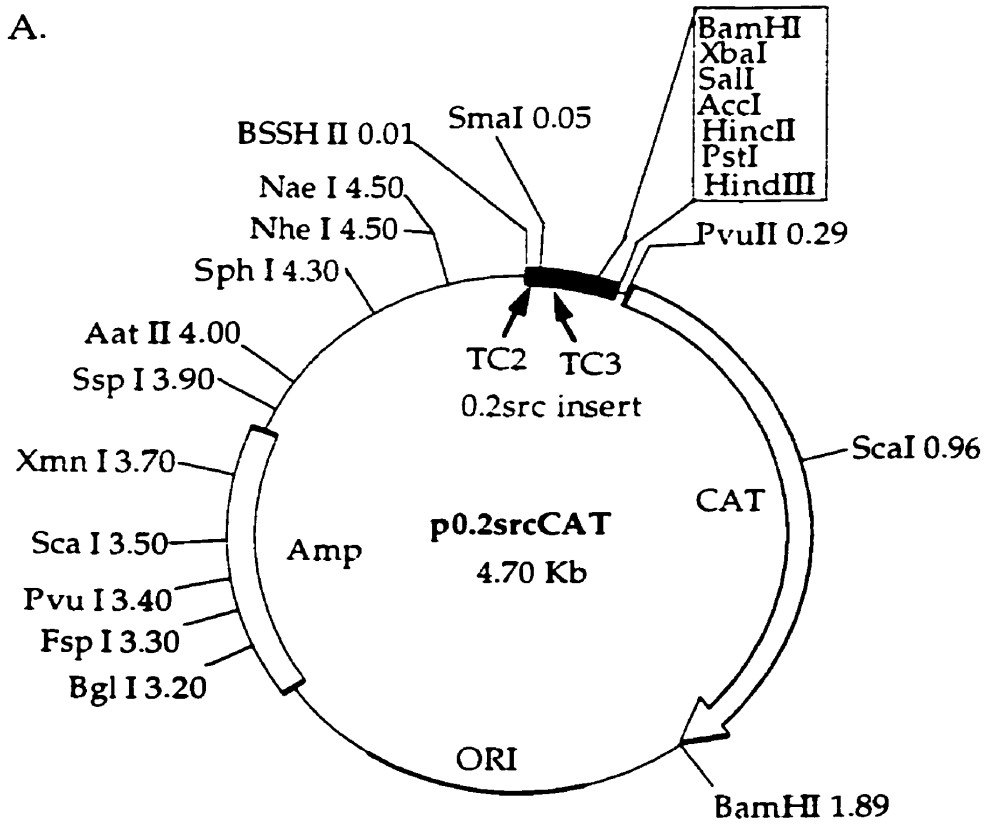


Figure 3.19. Dimer formation with linearized pBS-Src and linearized p0.2Src-CAT. λ plus *Hin* dIII markers (M). pH 4, 1.8% agarose gels. Incubation components are given above the lanes. A. Dimers between linearized pBS-Src and linearized p0.2Src-CAT. Positions of linear monomers and dimers are indicated: linear pBS-Src (LSrc), linear p0.2SrcCAT (LCAT), linear pBS-Src dimers (LDSrc), linear p0.2SrcCAT dimers (LDCAT), dimers between linear pBS-Src and linear p0.2SrcCAT (LDSrcCAT). Lanes contained 2.6 μg of DNA. B. Dimerization in linearized p0.2SrcCAT and position of the open circular form. Plasmid forms: supercoiled (SC), plasmid γ-irradiated for 7 min (OC), open circular dimer (OCD), linear dimer (LD), linear (L). DNA added to lanes: lanes 1-3 (1.6 μg), lane 4 (0.6 μg).



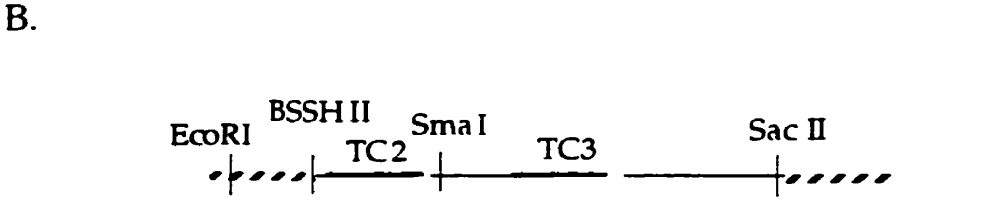


Figure 3.20. p0.2SrcCAT. A. Major restriction sites, genes, and the two pur•pyr tracks, TC2 and TC3. B. Expanded diagram of cloned region of c-*src* promoter. The flanking diagonally-striped areas are vector sequences.

Accordingly, the junction in this present study was an X-junction formed in a similar way to that seen in pKHa3PYL/PUL linear X-dimers and Ω loops (Figure 3.6).

The question arose as to whether the extra band following incubation of p0.2SrcCAT linearized with *Sma* I at pH 4 with spermine (lane 9, Figure 3.19A) was in fact due to linear dimers (either between like tracts or different tracts) or to dimerization between TC2 and TC3 located at either end of the linearized plasmid so that a loop structure formed. The band must have been due to the formation of linear dimers, as it had a slower mobility (lanes 1-3, Figure 3.19B) than the open circular form of the plasmid (lane 4, Figure 3.19B). Some loop structures may have also formed. However, their mobility would not have been distinguishable from the open circular forms on this 1.8% gel, as the linear monomers and open circular forms ran in the same position.

3.4.2 Autoradiography of Linear Dimers with Radiolabelled TC1

The results above show that the pur•pyr tracts in the c-*src* promoter can dimerize as a result of triplex formation. To define further which pairs of c-*src* pur•pyr tracts can form a triplex, TC2 and TC3 were each isolated alone in linear plasmid fragments of unique length called TC2-linear and TC3-linear, respectively. TC2-linear and TC3-linear were made by cutting p0.2SrcCAT with *Sma* I, then by digesting with *Sca* I (Figure 3.21). It was not possible to isolate TC1 alone in a linear fragment of convenient size, using simple restriction digestion of pBS-Src. This was because of the limited restriction sites and the proximity of the other pur•pyr tracts (Figure 3.18). For example, a pair of *Sma* I sites flanked TC1 and TC2 contained close together in a cassette in pBS-Src. TC1 itself was contained in a 56 bp cassette flanked by a pair of unique *BSS* HII sites. However, digestion with *BSS* HII yielded a fragment too small to be isolated practically from preparative gels. In any case, it would have been difficult to detect any change in mobility of TC2-linear or TC3-linear following dimerization with the *BSS* HII cassette containing TC1 because of its small size. Attempts to obtain TC1 at the end of a linear plasmid fragment of useful size, following partial digestion with *BSS* HII, were unsuccessful.

Instead, the following approach was taken. The *BSS* HII cassette containing TC1 was purified from pBS-Src by digesting the plasmid with *BSS* HII. The cassette was then end-labelled with ^{32}P. Radiolabelled TC1 (TC1*)

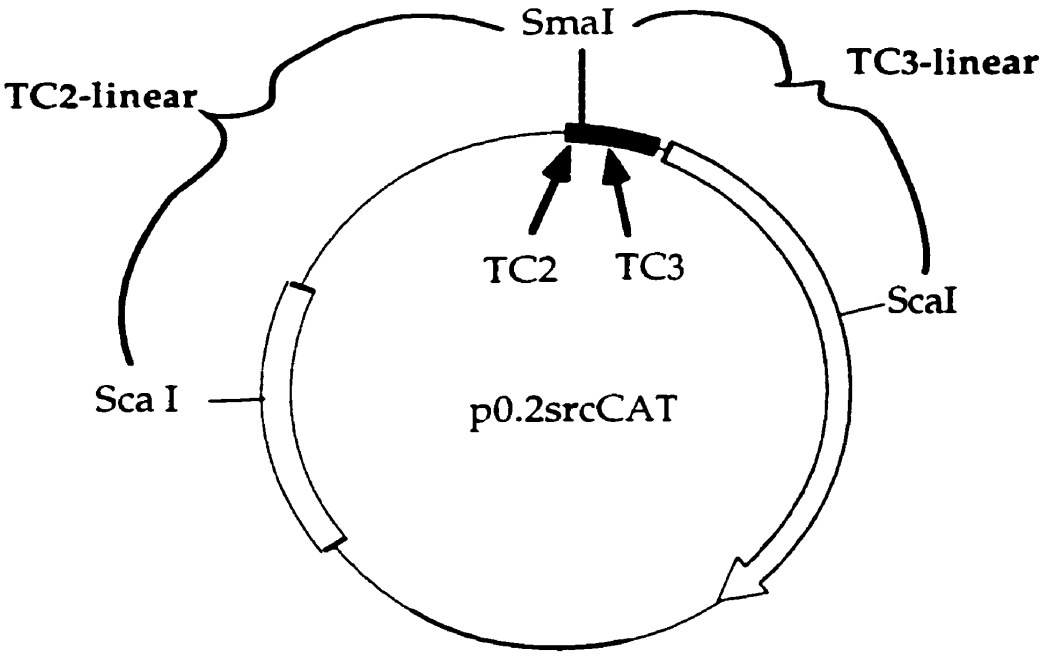


Figure 3.21. p0.2 srcCAT showing the location of the plasmid fragments TC2-linear and TC3-linear. These were produced by cleavage at the *Sma* I and *Sca* I restriction sites. The locations of the c-*src* promoter pur•pyr tracts TC2 and TC3 are indicated by arrows.

was then incubated with TC2-linear or TC3-linear at pH 4 in the presence of spermine in the same experiment (Figure 3.22). The strength of dimerization between TC1* and TC2-linear and/or TC1* and TC3-linear could then be compared by band formation on the same autoradiograph. When TC1* was incubated with TC2-linear at pH 8 without spermine, a weak band appeared suggesting non-specific binding between TC1* and TC2-linear (lanes 1 and 2, Figure 3.22). However, following incubation with increasing concentrations of spermine at pH 4, the band became stronger indicating the presence of TC1*/TC2-linear heterodimers based on triplex formation between the tracts TC1 and TC2 (lanes 3-5, Figure 3.22). Again, when TC1* was incubated with TC3-linear, non-specific binding occurred at pH 8 without the presence of spermine (lane 6, Figure 3.22). However, strong bands appeared when incubation was carried out at pH 4 (lane 7, Figure 3.22) and in the presence of increasing concentrations of spermine (lanes 8-10, Figure 3.22). This was a result of heterodimer formation between TC1* and TC3-linear, based on triplex formation between TC1 and TC3. Faint heterodimers formed by the dimerization of TC1* with TC2-linear homodimers are visible at higher spermine concentrations (lanes 4 and 5, Figure 3.22). Much stronger heterodimers are visible between TC1* and TC3-linear homodimers (lanes 8-10, Figure 3.22). In spite of nonspecific binding occurring, this autoradiograph suggests that a stronger triplex is formed between the TC1 and TC3 pur•pyr tracts than between TC1 and TC2.

Non-specific binding between TC1* and TC3-linear was confirmed by incubating TC1* with TC3-linear with the pur•pyr tract TC3 removed (Figure 3.23). Weak dimerization is evident following incubation of TC1* with TC3-linear without TC3 (lanes 3-9, Figure 3.23).

3.4.3 Dimers with TC1-Linear, TC2-Linear, and TC3-Linear

In order to determine whether a triplex could form between TC2 and TC3, and that a stronger triplex formed between TC1 and TC3 than between TC1 and TC2, TC1 was obtained alone in a linear plasmid fragment of unique length. In keeping with the nomenclature for the other two pur•pyr tract-containing fragments, TC2-linear and TC3-linear, this fragment was called TC1-linear. TC1-linear was made by cloning the portion of pBS-SRC containing all four consecutive pur•pyr tracts into pUC 19 (Figure 3.24). An exonuclease III

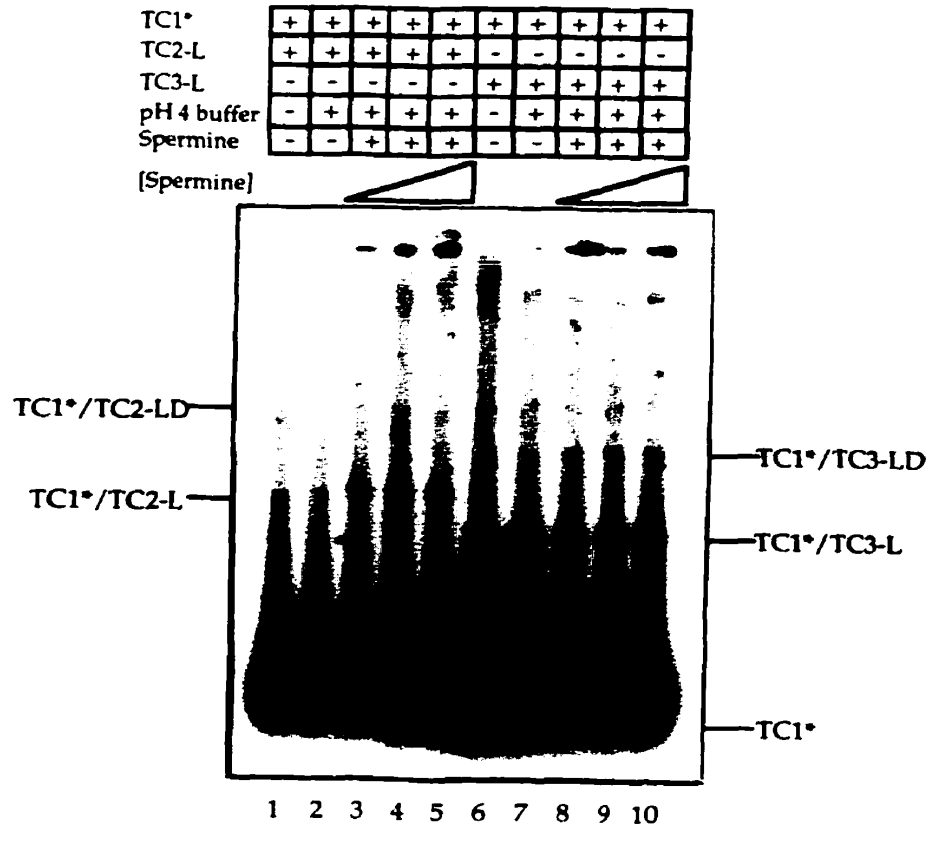


Figure 3.22. Autoradiograph showing heterodimerization of radiolabelled TC1 with TC2-linear and TC3-linear. Positions of TC1 (TC1*), TC1/TC2-linear heterodimer (TC1*/TC2-L), TC1/TC3-linear heterodimer (TC1*/TC3-L), TC1*/TC2-linear homodimer heterodimers (TC1*/TC2-LD), and TC1/TC3-linear homodimer heterodimers (TC1*/TC3-LD) are indicated. Except for the TC1 band, the positions of all the other bands were determined by comparison with gel bands prior to autoradiography. Incubation components in each lane are given above the gel. Concentrations of spermine increased from 50 μM (lanes 3 and 8) to 100 μM (lanes 5 and 10).

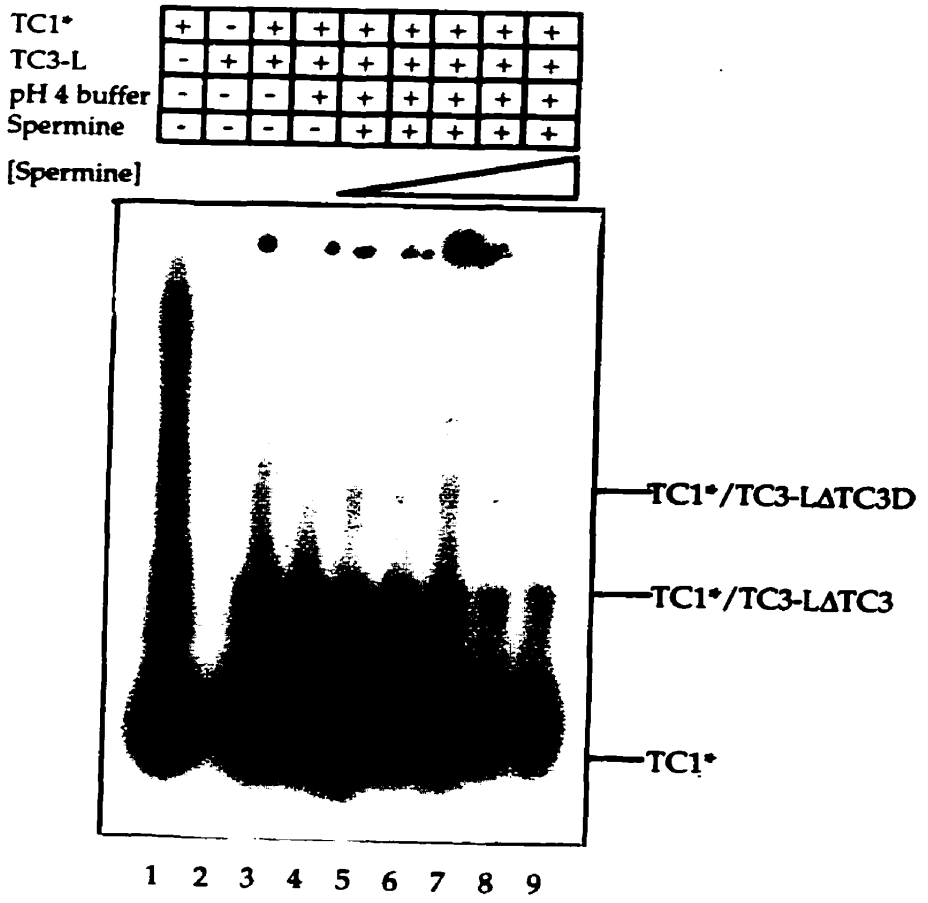


Figure 3.23. Autoradiograph showing weak non-specific dimerization between radiolabelled TC1 and TC3-linear without TC3. Positions of TC1 (TC1*), TC1/TC3-linear without TC3 (TC1*/TC3-LΔTC3), and TC1/TC3-LΔTC3 homodimer heterodimer (TC1*/TC3-LΔTC3D) are indicated. Except for TC1, the positions of the bands were determined by comparison with gel bands prior to autoradiography. Incubation components in the lanes are listed above the gel.

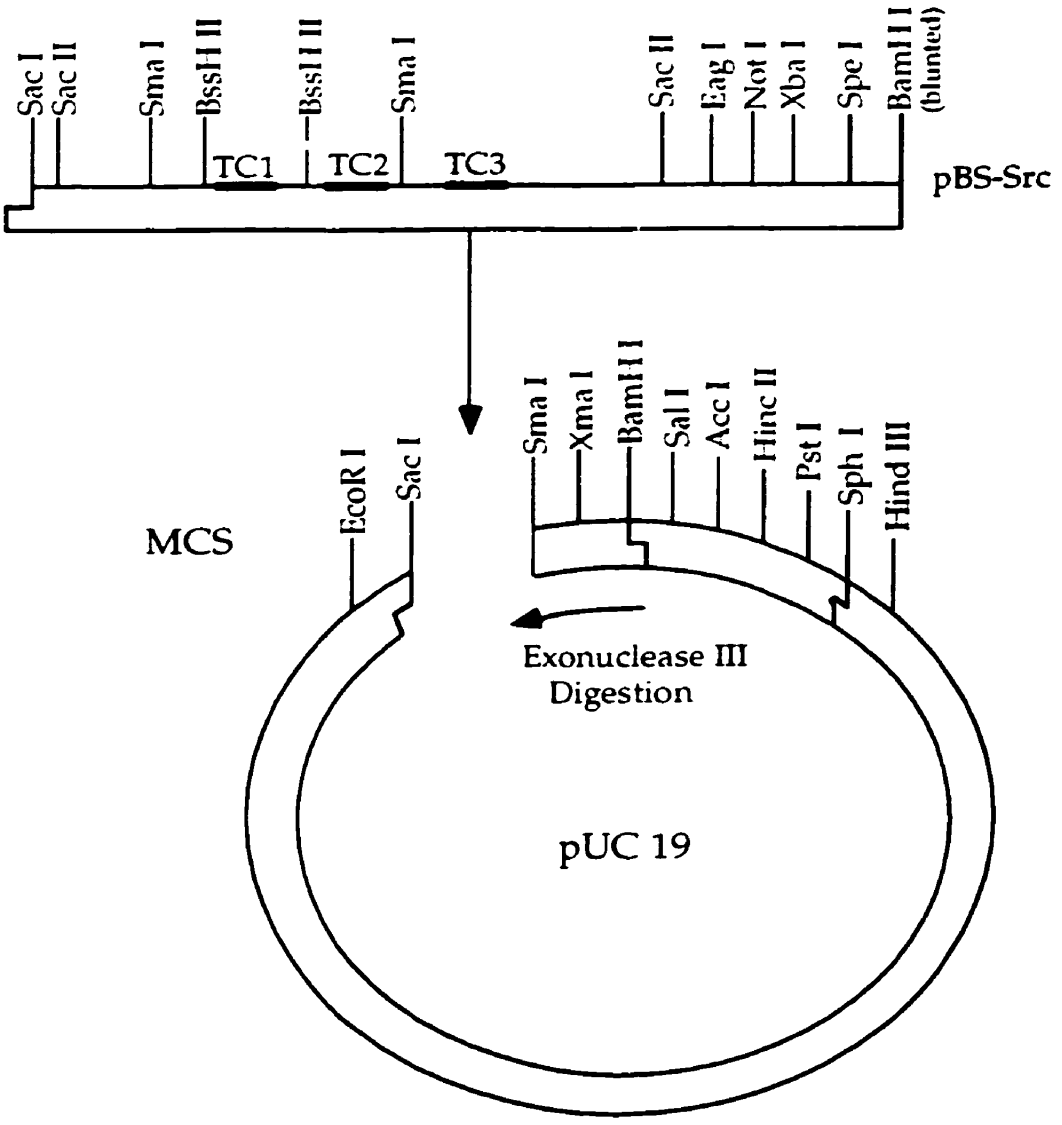


Figure 3.24. pUC 19 and position of cloned pur•pyr tracts from pBS-Src. The direction of exonuclease III digestion to create a plasmid containing only TC1 is shown. MCS - multiple cloning site.

digestion was then carried out to remove the *c-src* promoter sequences containing the other pur•pyr tracts to the 3' end of TC1 plus one nucleotide. The extent of the exonuclease III digestion was determined by sequencing (not shown). Following religation, this new plasmid was linearized at *Hin* dIII in the multiple cloning site of pUC 19 to create the linear plasmid fragment, TC1-linear.

TC1-linear, TC2-linear, and TC3-linear were then incubated together in pairs at pH 4 in the presence of spermine. Any triplex formation between two tracts would be visible as a dimer with an altered mobility upon gel electrophoresis. Dimerization between linear fragments containing the same tract (homodimers versus heterodimers) had already been observed with TC-2 linear and TC-3 linear. However, the sizes of TC1-linear, TC2-linear, and TC3-linear were designed so that the identity of the linear fragments comprising a dimer could be determined from gel mobility. Table 3.8 lists the predicted sizes of homo- and heterodimers when TC1-linear, TC2-linear, and TC3-linear are incubated together.

Figure 3.25 shows the results of incubating TC1-linear with TC2-linear. TC1-linear and TC2-linear were each incubated alone, first under non-triplex-forming conditions, then at pH 4 in the presence of spermine (lanes 2-5, Figure 3.25). The band containing TC1-linear was estimated to be 2,700 bp, by comparison with the molecular weight markers. A TC1-linear homodimer formed and was estimated to be 5,300 bp in size (lane 3, Figure 3.25). The fastest moving bands on the gel, estimated to be about 1,100 bp in size, contained TC2-linear (lanes 3 and 4, Figure 3.25). A strong band estimated to be 2,330 bp containing TC2-linear homodimer formed following incubation with 100 μM of spermine (lane 5, Figure 3.25). These estimated band sizes were close to the predicted sizes (Table 3.8). TC1-linear and TC2-linear were incubated together in lanes 6 to 11 (Figure 3.25). The predicted size of a TC1-linear/TC2-linear heterodimer was 4,056 bp (Table 3.8). Therefore a band estimated to be a little less than 4,000 bp in size must have contained TC1-linear/TC2-linear heterodimers (lanes 7-11, Figure 3.25). The strongest TC1-linear/TC2-linear heterodimers followed incubation with 100 μM spermine (lane 10, Figure 3.25). No TC2-linear homodimers were visible at lower spermine concentrations when TC1-linear and TC2-linear were incubated together (lanes 6-8, Figure 3.25). However, faint bands corresponding to TC2-linear homodimers were visible following incubation with 100 μM spermine (lane 10, Figure 3.25). Therefore,

Table 3.8. Predicted sizes of gel bands with homo- and heterodimers of TC1-linear, TC2-linear, and TC3-linear.[a]

			Heterodimer with:	
Linear Fragment	Size	Homodimer	TC2-linear	TC3-linear
TC1-linear	2,806	5,612	4,056	3,716
TC2-linear	1,250	2,500		2,160
TC3-linear	910	1,820		

[a] Sizes are given in bp and are calculated by adding the sizes of the constituent linear plasmid fragments.

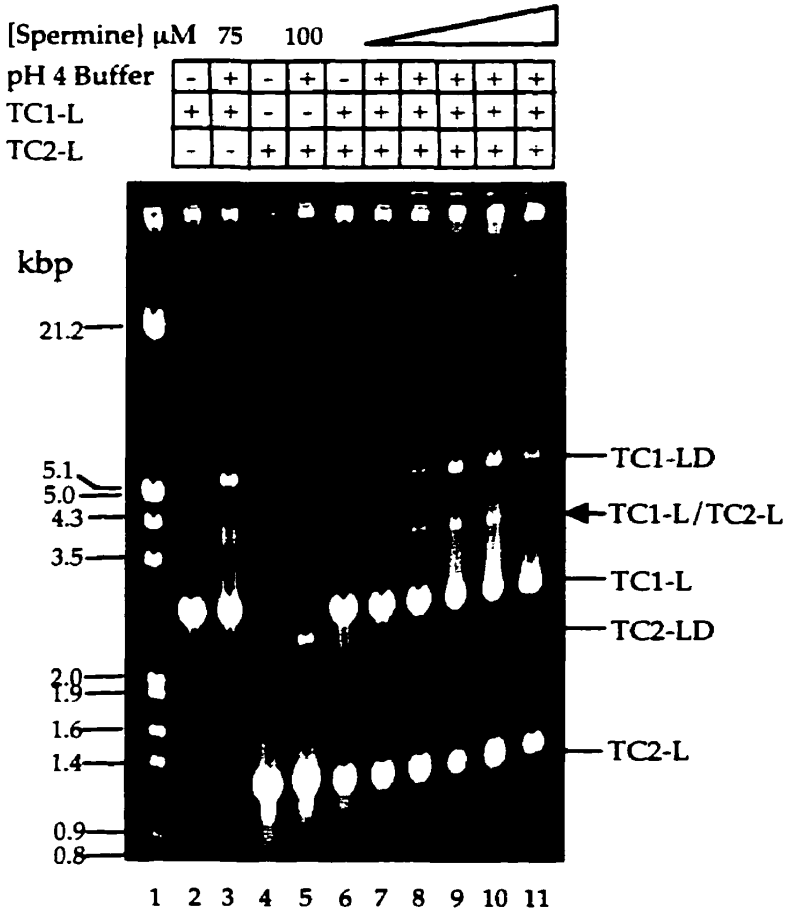


Figure 3.25. Incubation of TC1-linear with TC2-linear. Lane 1, phage λ DNA cut with *Eco* RI plus *Hin* dIII. The sizes in bp of the MW marker bands are indicated on the left. The positions of linear plasmid monomers and dimers are shown beside the gel where they occur. Abbreviations: TC1-linear (TC1-L), TC1-linear homodimer (TC1-LD), TC2-linear (TC2-L), TC2-linear homodimer (TC2-LD), TC1-linear/TC2-linear heterodimers (TC1-L/TC2-L). The presence of each linear plasmid fragment, spermine concentration, and pH 4 triplex buffer in the incubation mix is given above the lanes. The concentration of spermine in the incubation mix increased from 25 μM in lane 7 to 125 μM in lane 11. Lanes contained 1 μg of DNA and linear fragments were present in equimolar amounts. pH 4, 0.85% agarose gel.

the formation of TC1-linear/TC2-linear heterodimers was favored over TC2-linear homodimers (lanes 7-11, Figure 3.25). An apparent band of about 4,000 bp appeared in lane 3 (Figure 3.25). However, this had a diffuse appearance and must have been an artefact (reflection) of the digital imaging program as it was not visible on a photograph of the gel.

The results of incubating TC2-linear with TC3-linear are shown in Figure 3.26. TC2-linear and TC3-linear were each incubated alone in lanes 2-5 (Figure 3.26), first at pH 8 without spermine, then at pH 4 in the presence of spermine. The fastest moving bands in the lanes corresponded to the TC3-linear fragment, of estimated size 900 bp. The next was the TC2-linear fragment, of estimated size 1,200 bp. The positions of bands containing TC3-linear homodimers and TC2-linear homodimers are shown. A band of about 2,000 bp in size moving between TC3-linear and TC2-linear homodimers contained heterodimers formed between TC2-linear and TC3-linear (lanes 7-11, Figure 3.26). This agrees well with the predicted size of this heterodimer of 2,160 bp (Table 3.8). The other bands which ran slower in the lanes than the TC2-linear homodimers must have been multimers.

The strongest heterodimers in these experiments, based on band density, formed when TC1-linear and TC3-linear were incubated together (lane 3, Figure 3.27). The TC1-linear/TC3-linear heterodimer band was identified as such on control gels (not shown). A competition experiment was therefore carried out by adding oligonucleotides corresponding to the TC1 and TC3 pyrimidine tracts to the incubation mix (lanes 4-11, Figure 3.27). This was to determine whether the pyrimidine strand of TC1 or TC3 formed the third strand of the pyr•pur•pyr triplex at the root of the TC1-linear/TC3-linear heterodimer. The heterodimer was removed by addition of the TC3 pyrimidine oligonucleotide at an oligonucleotide to plasmid molar ratio of 1:4 (lane 9, Figure 3.27). Therefore, the third strand of the triplex at the root of the heterodimer was the pyrimidine strand of the pur•pyr tract TC3 in TC3-linear. These results agree with those obtained using autoradiography where stronger dimers appeared to form between radiolabelled TC1 and TC3-linear than between radiolabelled TC1 and TC2-linear (Figure 3.22). These results also support but do not prove the proposal that it was the TC3 tract of linearized p0.2SrcCAT rather than the TC2 tract which was located at the junction in the Y dimers formed with linearized pBS-Src (Figue 3.19).

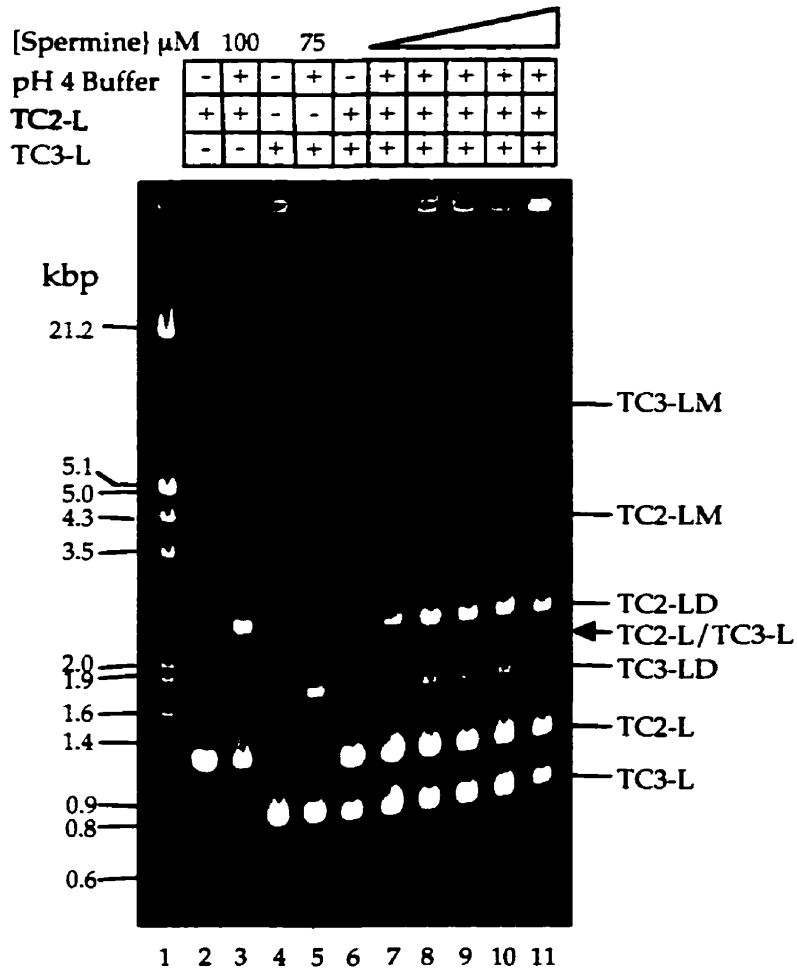


Figure 3.26. Incubation of TC2-linear with TC3-linear. Lane 1, phage λ DNA cut with *Eco* RI plus *Hin* dIII. The sizes of the MW marker bands are indicated on the left. The positions of homo- and heterodimers are indicated: TC2-linear (TC2-L), TC3-linear (TC3-L), TC2-linear homodimers (TC2-LD), TC3-linear homodimers (TC3-LD), TC2-linear/TC3-linear heterodimers (TC2-L/TC3-L), TC2-linear multimer (TC2-LM), TC3-linear multimer (TC3-LM). The presence of each tract, spermine concentration, and pH 4 triplex buffer in the incubation mix is given above the lanes. The concentration of spermine in the incubation mix increased from 25 μM in lane 7 to 125 μM in lane 11. Lanes contained 1 μg of DNA and plasmid tracts were present in equimolar amounts. pH 4, 0.85% agarose gel.

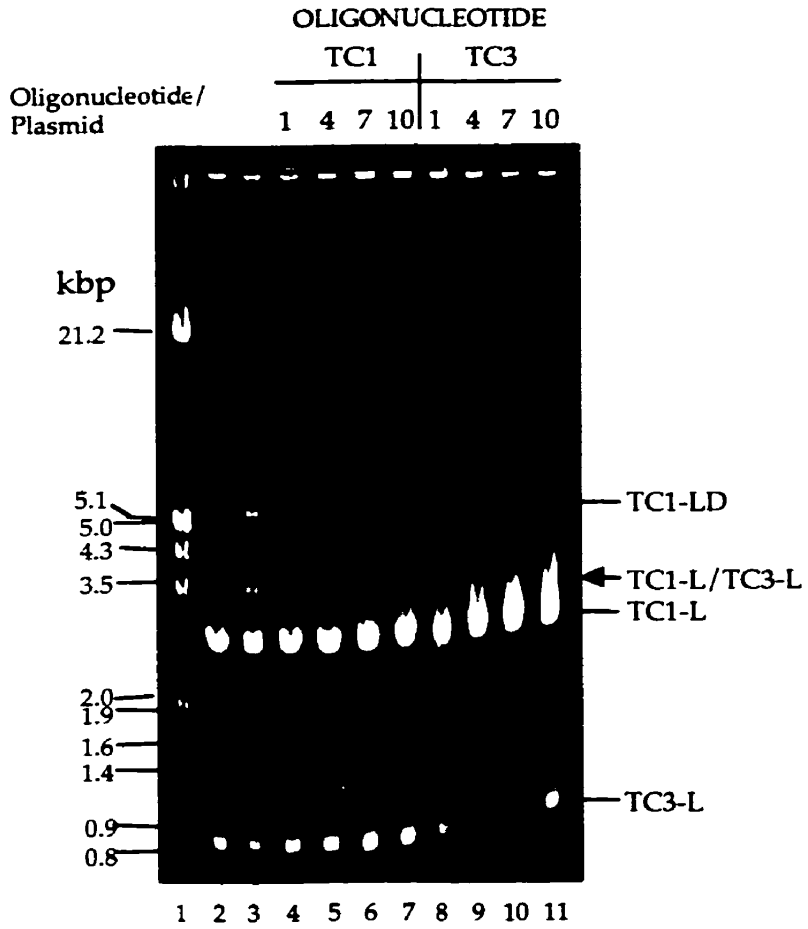


Figure 3.27. Incubation of TC1-linear with TC3-linear and competition using TC1 and TC3 oligonucleotides. Lane 1, λ phage cut with *Eco* RI and *Hin* dIII. Lane 2, TC1-linear and TC3-linear incubated at pH 8, no spermine. Lanes 3-11: incubations were carried out in pH 4 triplex-forming buffer and contained 75 μM spermine. Lane 3, TC1-linear and TC3-linear incubated together. Lanes 4-7, TC1-linear and TC3-linear incubated with TC1 pyrimidine oligonucleotide. Lanes 8-11, TC1-linear and TC3-linear incubated with TC3 pyrimidine oligonucleotide. Positions of homo- and heterodimers indicated on the right: TC1-linear (TC1-L), TC3-linear (TC3-L), TC1-linear homodimers (TC1-LD), TC3-linear homodimers (TC3-LD), TC1-linear/TC3-linear heterodimers (TC1-L/TC3-L). The molar oligonucleotide to plasmid ratio is given at the top of the lanes. Lanes contained 1 μg of DNA and plasmid tracts were present in equimolar amounts. pH 4, 0.85% agarose gel.

The TC3 oligonucleotide also removed the TC1-linear homodimers (lane 9, Figure 3.27). This would have been due to triplex formation between the TC3 oligonucleotide and the TC1 tracts in the TC1-linear homodimers. The TC1 pyrimidine oligonucleotide brought about partial removal of both the heterodimers and the TC1-linear homodimers (lanes 4-7, Figure 3.27). This would have been due to sequence similarities between TC1 and TC3.

4.0 DISCUSSSION

Researchers have proposed a number of roles that triplex DNA might play in the cell. The laboratory of Dr. J.S. Lee has focused on investigating two of these - a role in the condensation of DNA in chromosomes and a regulatory role in transcription. Two antibodies with different specificities for pyr•pur•pyr triplexes had previously been produced. These antibodies detected triplexes in both fixed and unfixed mammalian chromosomes. They also showed that the distribution of triplexes within chromosomes was not random. In later work, preincubation of agarose-encapsulated nuclei with these antibodies brought about a reduction in transcription. The work with the antibodies confirmed the idea that triplexes might have both a structural and regulatory role.

In other research in the laboratory, a low-pH-dependent chromosome condensation was detected in the second dimension of pulsed-field gel electrophoresis. This suggested a structural role for pyr•pur•pyr triplexes. As a result of the pulsed-field work, and more recent work involving looping between separated pur•pyr tracts in plasmids, it was proposed that the looping of DNA in the condensation of chromosomes could be mediated by triplexes. These triplexes could form the base of the large loops known to be present in at least one level of condensation of DNA within the chromosomes (Hampel and Lee, 1993; Hampel *et al.*, 1993). Additionally, it was postulated that triplexes could form elsewhere on the loops as a result of pur•pyr tract interaction with such tracts on neighbouring loops. This latter type of triplex could stabilize the loops in an orderly array around the MAR. Finally, it was suggested that triplex-mediated looping could also be involved in the regulation of transcription. This type of triplex, which forms across a single molecule between separated but complementary pur•pyr tracts in a sea of duplex DNA, is being referred to as a transmolecular triplex.

The work in this thesis continues the plasmid modeling of transmolecular triplex formation, and extends it to investigate in detail a possible role of this type of triplex in the regulation of transcription. In

addition, the investigation of the role of triplexes in the regulation of transcription was applied to two *in vivo* situations.

4.1 X-Junctions Between Polypurine•Polypyrimidine Tracts in Duplex DNA

The important role that duplex DNA looping plays in chromosome condensation and the regulation of transcription has been well documented. Duplex DNA condensation within chromosomes is known to be mediated via large loops (Paulson and Laemmli, 1977; Sinden and Pettijohn, 1981). Additionally, in the regulation of transcription, distant regions are thought to be brought together via looping. In both cases, proteins have been shown to mediate the looping (Schleif, 1992; Hofmann *et al.*, 1989). Recently, it has been proposed that such looping may also be mediated by triplex formation between separated pur•pyr tracts (Hampel and Lee, 1993; Hampel *et al.*, 1993). This type of looping was initially investigated using intermolecular triplex formation between single but complementary pur•pyr tracts positioned in the middle or at one end of two different plasmids, or between two separated tracts positioned at either end of the same plasmid (Hampel *et al.*, 1993). In these experiments, there was no barrier to unwinding during triplex formation when at least one tract was positioned at the end of a linear plasmid. These experiments showed that separated pur•pyr tracts embedded in duplex DNA could interact and form a triplex leading to a tertiary structure.

A more relevant model of DNA looping in the condensation of DNA in chromosomes or in the regulation of transcription via triplex formation should be provided by separated pur•pyr tracts present on the same linear plasmid but without free ends. This was the first plasmid model of transmolecular triplex formation investigated in this thesis. Plasmid pKHa3PYL/PUL was designed with two separated pur•pyr tracts able to form a triplex with each other but not within a tract itself (Figure 3.1). The plasmid was linearized so that the two pur•pyr tracts were located at some distance from the ends. Acidic incubation conditions in the presence of spermine promoted the formation of linear X-dimers and Ω loops (Figure 3.2). The structures proved to be extraordinarily stable at pH 8, yet were resolved to linear monomers with P1 nuclease. A model was presented explaining the stability of the structures at pH 8. The presence of single-

stranded DNA constrained the structures, forming an X-junction containing a hydrogen-bonded knot (Figure 3.6B).

For several reasons, the results were consistent with the formation of a double pyrimidine triplex at pH 4. First is the requirement of low pH which would facilitate the formation of C•G•C$^+$ base triads. Secondly, in results reported elsewhere, formation of the dimers and loops was blocked by adding an homologous single-stranded oligopyrimidine but not an oligopurine (Hampel, 1994). Additionally in earlier work, control plasmids showed that both the PYL and PUL tracks of pKHa3PYL/PUL (Figure 3.1) must be present in the same plasmid for circles to form (Hampel *et al.*, 1993).

The formation of the X-junction is modelled in Figure 3.6B. In the initiation of triplex formation between the two separated pur•pyr tracts, a single-stranded bubble is created when one of the pur•pyr tracts denatures at the top. The acceptor duplex (labelled "L" in the diagram) initiates triplex formation by passing once through the single-stranded bubble for every turn of triplex. Although the loop is topologically closed, no supercoiling occurs. The process forming the linear X-dimers would be similar. An equilibrium state exists between the dimers or loops and the linear monomers at pH 4. When the pH or the temperature is increased, the triplex falls apart and the donor duplex partially reforms, "clamping" the acceptor duplex and constraining the Ω loop or linear X-dimer in position by means of a braided or hydrogen-bonded knot. At this point there would be no equilibrium with the monomer as the structure would be constrained. Nicking the X-junction with P1 nuclease allows the acceptor duplex to pass back through and the structure reverts to a linear duplex monomer. An alternative twisted structure would not be resolved by P1 nuclease (Figure 3.6A). This is the first instance of "topological trapping" by a hydrogen-bonded knot presented in this thesis. The structure at the join is called a hydrogen-bonded knot rather than a true knot, as it can be undone without breaking any covalent bonds.

Topoisomerase I, which removes positive or negative supercoils by cutting one strand of duplex DNA and rotating it about the other strand failed to resolve the structures. Hampel (1994) also found that incubation with topoisomerase II was unsuccessful in resolving the structures. It is unknown why topoisomerase I and II failed to remove the hydrogen-bonded knot from the linear X-dimers and the Ω loops. A possibility is that enzyme concentrations in the local environment of the braided knot were too low.

In a supercoiled molecule, a single or double-stranded nick introduced anywhere in the structure would be capable of relaxing it. In the case of the braided knot, the activity would need to be focused on the area of the knot.

The X-junctions showed that triplex formation could occur between two separated pur•pyr tracts in duplex DNA without the presence of a free end. Ω loop formation may be a more realistic model for how triplexes could be involved in the cellular processes of DNA condensation in chromosomes and transcription. The condensation of DNA via loops may reflect the existence of a "chromosome folding code" based on repetitive DNA sequences (Vogt, 1990,1992). Pur•pyr tracts capable of forming triplex DNA are present in MARs (Opstelten *et al.*, 1989; Boulikas, 1993). Additionally, it is interesting to note that many MAR pur•pyr tracts are S1 nuclease sensitive (Boulikas, 1993; Gromova *et al.*, 1995). Topoisomerase II is the major protein component of the MAR and is required for chromosome condensation and decondensation (Adachi *et al.*, 1991; Uemura *et al.*, 1987). Topoisomerase II may provide the nicking and sealing activity necessary for the formation of loops. The structure could be removed if one of the duplex single strands were nicked in such a way as to release its "clamping" action. The requirement for topoisomerase II in chromosome condensation and decondensation may reflect the presence of triplexes at some level in the process. These results, however, do not eliminate the possibility that the alternative twisted structure of Figure 3.6A could also form at the base of the loops *in vivo*.

Ω loops could also play a regulatory role in the initiation of transcription. Pur•pyr tracts have frequently been found in the regulatory regions of genes. Where these contain the appropriate mirror-repeat sequence it is thought that the tracts might form H-DNA (Htun and Dahlberg, 1989). It is hypothesized that H-DNA might be involved in the regulation of transcription. For example, the single-strand of DNA extruded by the formation of the structure might facilitate the entry of RNA polymerase and thus promote transcription. In a similar way, if an Ω loop structure resulted from triplex formation between a promoter pur•pyr tract and a distant tract, even one several hundred kb distant, the single-stranded region present could be envisaged as facilitating the entry of RNA polymerase.

4.2 Transmolecular Triplex Formation in Circular DNA

Although it is easy to see how two separated duplex pur•pyr tracts in linear DNA could come together to form a transmolecular triplex, it is not immediately obvious how two separated pur•pyr tracts in circular DNA might come together. Yet this question may be very relevant as chromosomal DNA is constrained in loops. The results of experiments in this thesis showed that there is no topological impediment to the formation of such transmolecular triplexes.

Transmolecular triplex loops (T-loops) were formed at low pH in the presence of spermine from the open circular form of pKHa3PYL/PUL, the same plasmid used in the previous experiments (Figure 3.1). As with the linear X-dimers and Ω loops the incubation conditions promoted the formation of a double pyrimidine triplex. Furthermore, as with the previous structures based on the formation of an X-junction, the T-loops were found to be stable at pH 8 and sensitive to P1 nuclease. Again, this suggested topological trapping due the presence of a single DNA strand. A model for the formation of T-loops was proposed which again involved a hydrogen-bonded knot (Figure 3.13).

It is unknown which of the two pur•pyr tracts, PUL or PYL, of pKHa3PYL/PUL contributed the third strand pyrimidine in the formation of the T-loops. In fact, for several reasons it seems possible that the third strand was equally contributed by both tracts. In the Htun and Dahlberg model of H-DNA, the H-y3 isoform is favored energetically at high levels of supercoiling while the H-y5 isoform is favored at low or no supercoiling (Htun and Dahlberg, 1988, 1989). The experiments discussed here, however, do not involve H-DNA. Instead they involve transmolecular triplex formation between two tracts separated by 2.8 kb on the bottom half of the plasmid and 1.4 kb on the top half (Figure 3.1). This would mean that the DNA in the loop regions would be expected to remain as a duplex. The open circular forms from which the T-loops were made were estimated to be 95% nicked (open circular). This would mean they averaged 3 nicks/molecule (Pettijohn and Pfenninger, 1980). The nicking would have allowed free rotation of the duplex DNA strands in the formation of the T-loops. In the formation of transmolecular triplexes *in vivo*, it is hypothesized that nicking activity might be provided by topoisomerase I. Because of the

nicking leading to free rotation of the strands, there would have been no supercoiling present. Therefore, there should have been no energetic barrier to the formation of either "isoform". The orientation of the strands in the pur•pyr tracts in Figure 4.1 shows that if the third pyrimidine strand in the formation of the T-loop is contributed by PUL, an "H-y5" isoform results, and if it is contributed by PYL, an "H-y3" isoform results. However, as mentioned above, the resulting structure is not H-DNA, and the terms are not used in that context. Instead, they are simply used to describe the two ways in which the triplex at the root of the T-loop structure may be formed. The unequal distance between the two pur•pyr tracts at the top and bottom of the plasmid in Figure 4.1 shows the reason for the large and small loops in the dumbells in the electron micrographs (Figure 3.11A). However, the correspondence between the "C" and "V" loops in the model (Figure 3.13) and the large and small dumbell loops is unknown. It is also unknown from which end of the pyrimidine strands triplex formation is initiated.

H-DNA formation is driven by negative supercoiling (Htun and Dahlberg, 1988, 1989). It is thought that the waves of negative supercoiling generated behind a transcribing RNA polymerase may promote or stabilize intramolecular triplexes in the promoter region of a gene. Depending on how the structures form, supercoiling is not necessarily envisaged to be a factor influencing the formation of transmolecular triplexes *in vivo*. This is partly because the tracts could be several hundred kb apart. On the other hand, restrained negative superhelical tension in the *in vivo* equivalent of loop "C" (T-loop model, Figure 3.13), perhaps generated by protein binding in the loop, is one situation where supercoiling might drive formation of the triplex. On the other hand, supercoiling in the "V" loop is not seen as having an effect on the formation of the loops "Y" and "Z" as these loops are independent of the triplex structure.

Because chromosomal DNA is constrained in loops, compared with the Ω loop model, T-loops may more realistically portray the structural role that transmolecular triplexes could play in the condensation of DNA in chromosomes. For example, T-loops may form the base of chromosome-sized loops involved in condensation. Additionally, a T-loop structure could be envisaged as mediating the interaction between two complementary pur•pyr tracts located on different duplex DNA chromosome loops. This would stabilize the loops in an array around the

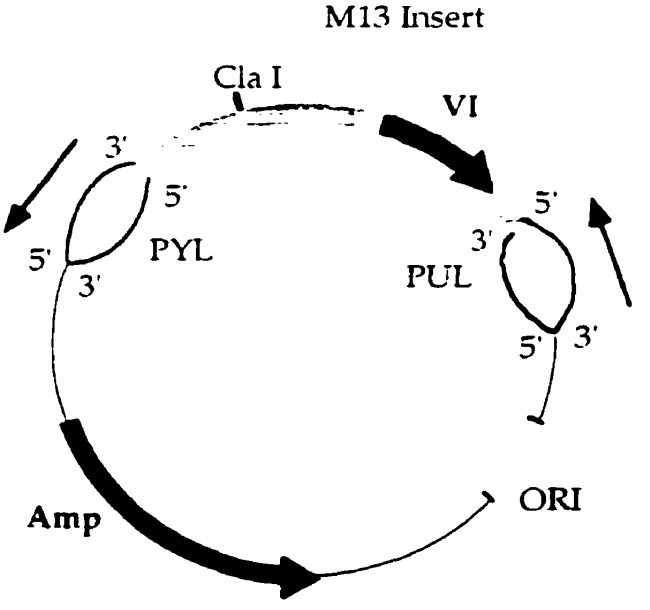


Figure 4.1. Transcription in pKHa3PYL/PUL. The positions of genes in the M13 insert (**VI**) and the pUC 19-derived portions of the plasmid (**Amp**) are shown and the directions of transcription indicated. The PYL and PUL pur•pyr tracts of Figure 3.1 are shown as single-stranded bubbles and strand orientation is indicated. Arrows beside the plasmid indicate directions of putative non-specifically-initiated transcription from the single-stranded polypurine tracts of T-loops.

nuclear matrix. The T-loop structure also has the potential to play a role in the regulation of transcription. This role is considered in experiments discussed in Section 4.4 below.

4.3 SSB Protein and T-Loops

Due to the presence of single-stranded region(s) in the hydrogen-bonded knot, T-loops provided a model system for studying potential interactions of single-strand-specific DNA-binding proteins with the single strand extruded by intra- or transmolecular triplex formation. It has been proposed that proteins specific for either the purine or pyrimidine strand of pur•pyr tracts may bind and either promote the formation of, or stabilize, an intramolecular triplex. Such a triplex formed in the promoter region of a gene might possibly have a role in the regulation of transcription (Lee *et al.*, 1984).

E. coli SSB protein was used as an example of a protein which specifically binds single-stranded DNA. In these experiments, 0.38 μg SSB protein had no significant effect either on the formation of T-loops or on their stability at pH 8, once formed. The amount of SSB protein used was based on a recommendation of 7 to 8.5 μg/μg of single-stranded DNA (Pharmacia catalogue, 1995). This corresponded to 0.2 μg if the estimated size of the single-stranded region in T-loops was 100 nucleotides. An increased amount of SSB (0.57 μg) also had no effect on T-loop formation. According to the supplier, this amount should been adequate for use on a single-stranded region of more than 200 nucleotides. Therefore, the concentration of SSB used in these experiments was sufficient.

There are two possible expanations for the lack of effect. At least in the formation of the T-loops, SSB protein may have lacked activity due to the low pH. However, Christiansen and Baldwin (1977) reported a significant level of SSB protein activity at pH 4. A more likely possibility is that the single-stranded region(s) in the T-loops did not form a suitable substrate for SSB protein. For example, the single-stranded region(s) may have been too short for the protein to bind, both at pH 4 and pH 8. It is unknown what length of single-stranded DNA is adequate for the binding of single-stranded binding proteins. However, SSB protein binds to single-stranded DNA as a tetramer, covering a region of about 70 nucleotides

(Krauss *et al.*, 1981). In general, binding studies on proteins specific for either purine or pyrimidine strands of pur•pyr tracts have been carried out using oligonucleotides of at least 30 nucleotides in length. Reducing an oligonucleotide length from 30 nucleotides to six brought about a four-fold decrease in binding (Aharoni *et al.*, 1993). The pur•pyr tract in pKHa3PYL/PUL is 72 bp long. During the formation of the T-loop triplex at pH 4, the size of the single-stranded bubble would depend on the size of the triplex formed and could possibly be as long as 72 bp. However, when the duplex reforms at pH 8, the P1 nuclease sensitive single-stranded region(s) in the hydrogen-bonded knot would likely be very much shorter.

4.4 Transcription in T-Loops

Separated duplex DNA sequences important in gene expression are thought to be brought into proximity via looping. Although this has traditionally been thought of as occurring via protein-protein interactions, this could also be accomplished via triplex formation between two separated pur•pyr tracts. The melting of promoter duplex strands by RNA polymerase is a rate limiting step in the initiation of transcription. Therefore, it has been proposed that a pre-existing single-stranded bubble in the region could promote transcription (Lee *et al.*, 1984). Formation of an intramolecular triplex in a promoter pur•pyr tract is one way in which a single strand may arise. S1 nuclease sensitive regions in the 5' flanking regions of genes have traditionally been thought to be due to the presence of a single-stranded region resulting from the formation of an intramolecular triplex. Another way for such S1 nuclease sensitive regions to arise is from T-loop formation between a promoter pur•pyr tract and another complementary pur•pyr tract in the promoter region or a tract at a more distant site. Because T-loops contain a single-stranded region or regions, it was hypothesized that they might non-specifically promote the entry of RNA polymerase, thus stimulating the initiation of transcription. Additionally, an interesting aspect of the transmolecular triplexes discussed in this thesis is the inherent remarkable stability or long-term hysteresis of T-loops, once the factors precipitating their formation have been removed. For example, the formation of T-loops *in vivo* could be promoted by the binding of a protein specific for the purine strand and spermine. Once the T-loops are formed

and the stabilizing factors are removed, the clamping action of the single strand would lock the structure in position. If such a structure were formed in the promoter region of a gene, for example, it may provide a mechanism for stably turning a gene on. Finally, regulation of transcription can also occur at the elongation or termination stages (Spencer and Groudine, 1990). Accordingly, T-loops could also be involved in the regulation of gene expression at these stages.

The experimental results in this thesis showed that when tested in parallel with three pKHa3PYL/PUL plasmid control forms (linear, open circular, and supercoiled) the lowest level of transcription occurred in the T-loop forms (Figures 3.15, 3.16). At both concentrations of *E. coli* core and holoenzyme, transcription in the T-loop forms was much lower than in the open circular forms (Table 3.1). In fact, with the holoenzyme, transcription in the T-loops was 50% of that in the open circular forms. Transcription was therefore blocked or inhibited in the T-loop forms. As the T-loops were made by incubating the open circular forms at pH 4, the only difference between the two forms would have been the presence of the hydrogen-bonded knot. Therefore, the hydrogen bonded knot must have brought about the reduction in transcription.

About 80% of the DNA in the isolated T-loop forms in the transcription experiments was present as T-loops (lane 1, Figure 3.14). The rest was open circular (10%) and linear (10%). Figure 4.1 shows the two possible directions for non-specific initiation of transcription, if it occurred, from the "Hy-3" and "Hy-5" isoforms. It is not possible to say for certain whether or not there was any non-specific initiation of transcription from the single-stranded region(s) in the area of the hydrogen-bonded knot in the T-loops. However, the experimental results suggest that it did not occur. Considering transcription with the core enzyme, the following analysis can be applied. As the T-loop forms were made from the open circular forms, the average number of single-stranded nicks in each of the forms would be the same. It has been reported that some non-specific initiation of transcription by RNA polymerase can occur from nicks in open circular DNA (Vogt, 1969; Hinkle *et al.*, 1972). This most likely accounted for most if not all of the initiation by the core enzyme in the open circular forms. The most likely explanation for the reduction in the T-loops with the core enzyme is that as with the open circular forms, transcription initiated non-

specifically at the nicks. However, it was terminated when it encountered the hydrogen-bonded knot at the end of the loop it started in, or an attenuator at the end of one of the two genes present. It is possible that the "Hy-3" and "Hy-5" T-loop isoforms were formed in equal numbers. Although energetically this is possible (Section 4.2), it also cannot be known for certain. Overall, this would have the effect of cutting the amount of RNA in half that was transcribed in the T-loops. Transcription in the T-loops was slightly better than half of that in the open circular forms. The extra amount could have been provided by the presence of a low percentage of linear and open circular forms in the isolated loops. This, however, is a strong argument for there being no significant stimulatory effect of the single-stranded regions in the T-loops on transcription.

This result contradicts other reports in the literature. *E. coli* RNA core and holoenzymes initiated transcription and elongation non-specifically from a 12-bp mismatch DNA bubble (Daube and von Hippel, 1992). In another example, non-specific initiation of transcription was reported from a D-loop created by a peptide nucleic acid (Mollegaard *et al.*, 1994). One possible explanation for the lack of transcription initiation from single-stranded regions in the T-loops is that the gaps in the hydrogen-bonded knot may have been too short to permit entry of RNA polymerase. Alternatively, the presence of the duplex strand winding through the single-stranded area (Figure 3.13) may have blocked elongation by the RNA polymerase.

The fact that transcription using both the core enzyme and the holoenzyme is about 50% less in the T-loop forms than in the open circular forms suggests two hypotheses. One is that RNA transcription, once initiated, continued all the way around the plasmid, unless it encountered a block such as the multistranded component of the T-loop or an attenuator. Second, the results with the holoenzyme again suggest that the two T-loop isoforms, "Hy-3" and "Hy-5" are present in equal numbers. Figure 4.1 shows that there are two genes present in pKHa3PYL/PUL. The holoenzyme may initiate transcription from these equally. If an attenuator is present at the end of gene VI in the M13 insert, holoenzyme transcription would stop before it encountered the multistranded hydrogen-bonded knot in the "Hy-5" isoform. Therefore, this T-loop isoform would have no significant effect on transcription of the gene. However, in the "Hy-3" isoform, depending on how the original triplex was formed, loops corresponding to "X" and "Y" of

the model (Figure 3.13) could interfere with transcription of gene VI. On the other hand, if an attenuator is present at the end of the Ampicillin gene, formation of the "Hy-3" isoform would have no effect on transcription of the Ampicillin gene by the holoenzyme. However, again depending on how the original triplex is formed, loops "X" and "Y" of an "Hy-5" isoform could conceivably interfere with Ampicillin gene transcription.

In conclusion, the T-loops are unable to provide evidence for non-specific initiation of transcription from single-stranded regions of DNA. Instead, they show that triple-stranded structures can block transcription. The T-loops and their effect on transcription have been used as an *in vitro* model of what might occur *in vivo* in the promoter of a eukaryotic gene. Unlike the situation with *E. coli* RNA polymerase, there have been no studies done to investigate the effect of a pre-existing single-stranded region in a promoter on stimulating assembly and initiation of transcription by the eukaryotic pre-initiation complex containing Pol II. Although the experiments conducted here have not shown it, the possibility remains that this could occur. If the tract which contributed the pyrimidine strand to a T-loop were located in the promoter region of a gene, the single-stranded region(s) in the hydrogen-bonded knot, if of sufficient length, could promote the binding of the Pol II complex and stimulate transcription, if the orientation of the "extruded" strand were correct. The fact that the gene promoter regions in the T-loops did not coincide with the single-stranded region(s) arising from triplex formation may have been a problem in using the T-loops to investigate this hypothesis. However, from the above results it certainly appears likely that the presence of the triple stranded part of the T-loop structure in the promoter region, the coding region, or the 3' region of a gene would have the effect of blocking or terminating transcription. These results also suggest that formation of T-loops could provide a block to replication. There are several examples in the literature where a multistranded structure has inhibited the progress of DNA polymerase (Section 1.5.4.1).

4.5 T-Loop Formation with Polyamine Analogues

Research into the use of TFOs in the control of gene expression in antigene therapy is a very important application of the study of triplexes

(reviewed in Soyfer and Potaman, 1996; Maher, 1996). Intermolecular triplexes do not form as readily as intramolecular triplexes: the stimulatory effect of supercoiling is absent and the TFO is able to dissociate more readily (Soyfer and Potaman, 1996). A TFO of about 15 nucleotides is suggested for specificity when used as an antigene agent in the promoter of a gene (Maher, 1996). However, the ideal length for a TFO is about 12 nucleotides as both longer and shorter nucleotides bind more weakly to their target duplex sequence (Cheng and Van Dyke, 1994). The reason for the weaker binding of a longer oligonucleotide appears to be due to structural misalignments between the duplex and the added TFO (Broitman *et al.*, 1987).

Thomas and Thomas (1993) had previously proposed that polyamines could be used as secondary ligands to promote triplex formation in antigene therapy. The polyamines could be free, attached to oligonucleotides, or attached to pyrimidine bases (Tung *et al.*, 1993; Barawkar *et al.*, 1996). Possible benefits of using polyamines as secondary ligands with the TFO are increased nuclease resistance, increased transport into the cell due to reduction of negative charge, and increased triplex formation.

A number of polyamine analogues were tested and compared with the naturally-occuring polyamines spermidine and spermine in their ability to promote triplex formation. T-loops had already been characterized by their altered mobility in an agarose gel. In many ways, formation of the transmolecular triplex at the root of the T-loops resembles the interaction of a TFO with its target duplex in intermolecular triplex formation. In the T-loops the third strand is donated by a duplex located a considerable distance from the target duplex tract. Furthermore, supercoiling is not a stabilizing factor in formation. Accordingly, a relevant way to assess the ability of the polyamine analogues to promote the formation of stable intermolecular triplexes was by observing the strength of formation of T-loops on agarose gels.

Total charge was the most important factor in the efficacy of the polyamines in promoting triplex formation. The higher the charge, the lower the concentration at which T-loops were formed (Table 3.3). Of the polyamines studied, the most efficient overall for promoting T-loop formation was 4-4-4-4. This agrees with previous reports which show that the ability of naturally-occurring polyamines to stabilize triplexes depends on their charge: spermine (+4) is more effective than spermidine (+3), which

is more effective than putrescine (+2) (Tabor, 1962; Glasser and Gabbay, 1968; Hampel *et al.*, 1991; Thomas and Thomas, 1993). In synthetic polyamines, pentamines have been found to be more efficient than tetramines for stabilizing triplex DNA (Musso *et al.*, 1997)

Spacing between the charges was also important as polyamines with trimethylene bridging promoted T-loop formation at lower concentrations than polyamines with tetramethylene bridging. For example, the optimum concentration at which T-loops formed with the synthetic polyamine 3-3-3-3 was 4 μM while with 4-4-4-4 it was 16 μM (Table 3.3). However, as noted, there was poor T-loop formation with 3-3-3-3. This was likely because of the high level of condensation promoted. Charge separation has previously been reported as important in substituted polyamines for promoting triplex stability (Thomas and Thomas, 1993). For example, divalent cations which have spacing between the positive charges are more stabilizing than Mg^{2+}, which has a point charge (Thomas and Thomas, 1993). In the crystal structure of spermine bound to a distorted B DNA hexamer, spermine lay across the minor groove rather than within it (Tari and Secco, 1995). This may provide an explanation of why the spacing of the charges in polyamines affects how well triplex formation is promoted - the spacing of the phosphate backbones of the DNA strands may be the critical factor. However, it is not known whether polyamines bind at the major or minor groove of a triplex. The third strand of the triplex binds in the major groove and hence the resulting high density of negative charges would provide a strong binding site for the polyamines. On the other hand, the minor groove is narrower in a triplex, and this might provide a site of the appropriate width for polyamine binding (Schmid and Behr, 1991).

Finally, bisethylated polyamines were studied as ligands for use with TFOs because they are resistant to catabolism within the cell (Porter and Sufrin, 1986). However, Table 3.3 shows that bisethylation generally resulted in a slight increase in the minimum concentration at which T-loops were formed. Therefore, the ethyl groups appeared to produce some steric hindrance. While spermine and spermidine containing hydrophobic substituents have been found to stabilize triplexes, increasing the size of hydrophobic moieties has been shown to have a destabilizing effect (Glasser and Gabbay, 1968; Musso *et al.*, 1997).

4.6 Gene Expression in Agarose-Encapsulated Nuclei

Agazie *et al.* (1996) observed a 20% reduction in TCA-precipitable counts in transcription following the incubation of agarose-encapsulated MOPC nuclei with triplex-specific antibodies. Explanations for the reduction may depend on whether transcription is reinitiated in such nuclei. It is still debated whether or not reinitiation of transcription by Pol II occurs in conventionally-prepared nuclei (Marzluff, 1978; Ausubel *et al.*, 1995). However, a number of researchers have reported reinitiation (Tata and Baker, 1985; Zhang-Keck and Stallcup, 1988; Thorburn and Knowland, 1988; Mennes *et al.*, 1992). In two of these, initiation of transcription of dormant vitellogenin genes was observed following incubation of nuclei with a cell extract (Tata and Baker, 1985) and cloned estrogen receptor (Thorburn and Knowland, 1988). However, in nuclei not exposed to an extraneous stimulus, the percentage of total RNA synthesis due to reinitiation is low - only about 3% (Zhang-Keck and Stallcup, 1988; Mennes *et al.*, 1992).

Researchers using agarose-encapsulated nuclei state that reinitiation does not occur in these nuclei (Jackson and Cook, 1985a,b; Jackson *et al.*, 1988; Thorburn and Knowland, 1988). On the other hand, the results of Agazie *et al.* are easier to explain if there is some reinitiation. This may occur as *in vitro* TFIID remains at the promoter initiation site where it could possibly facilitate further rounds of transcription (Section 1.7.2). If this is true, there might not be any requirement for renewed binding of transcription factors for the reinitiation of transcription.

Agazie *et al.* (1996) used the triplex-specific antibodies Jel 318 and Jel 466 at 60 µg/ml. This concentration would have been expected to stabilize pre-existing triplexes. It was likely too low to induce the formation of triplexes, as it falls in the so-called binding plateau in a graph of binding (CPM) versus antibody concentration (Wittig *et al.*, 1989; Agazie *et al.*, 1996). At the binding plateau, antibody concentrations can be increased approximately 100 fold without increasing binding. At much higher antibody concentrations, the binding increases exponentially, suggesting the induction of Z DNA (Wittig *et al.*, 1989) or triplexes (Agazie *et al.*, 1996).

A promoter triplex might be required to reform in order to take up negative supercoiling generated behind reinitiated transcription. In this case, preincubation with triplex-specific antibodies could fix the triplex

formed as a result of the first round of transcription. This would have the effect of hindering elongation when transcription was reinitiated. Topoisomerase I, which is known to be present in agarose-encapsulated nuclei (Wittig *et al.*, 1989), releases torsional strain generated by transcription (Zhang *et al.*, 1988). Consequently, it can be supposed that without the presence of the antibodies, the release of the torsional strain might have led to the reversion of the triplex structure to B-DNA. However, in spite of the torsional strain being released, if the triplex structure is bound by antibody, it would not be able to revert to B-DNA. The action of topoisomerase I has been described as continuous (Zhang *et al.*, 1988). If topoisomerase I and the H-DNA structure act in concert to release torsional strain generated by transcription, the antibody binding may have an effect on transcript elongation even without reinitiation of transcription occurring.

A much greater decrease in TCA-precipitable counts followed incubation with triplex-specific antibodies and run-off transcripition using agarose-encapsulated HT-29 nuclei in the present experiments compared with those of Agazie *et al.* (1996) (Table 3.5). This could be explained by a different cell line being used. HT-29 is an adherent human colon carcinoma cell line whereas MOPC cells are a suspended mouse plasma tumor cell line (ATCC Catalogue, 1992). Additionally, the packed cell volume of HT-29 cells used was doubled from 250 µl per ml total volume to 500 µl per ml, and the amount of antibody in the incubation was increased proportionally to 120 µg. Doubling the packed cell volume and concomitantly doubling the antibody concentration, while keeping the same final volume, could have affected the results. Additionally, procedures were modified to incorporate initial trypsinization steps to remove the adherent HT-29 cells from plate surfaces. Finally, the same packed volume of agarose beads could have held a significantly different number of HT-29 nuclei than MOPC nuclei. However, none of these possibilities explains why, using the HT-29 nuclei, there was also a 20% reduction in transcription following incubation with the control antibody, Jel 42.

Isolation of RNA following run-off transcription in agarose-encapsulated nuclei has not been previously reported in the literature. There were two technical problems to overcome in developing a procedure to isolate the RNA. The first was the extreme lability of RNA and the two-day period for preparing the agarose-encapsulated nuclei, carrying out run-

off transcription, isolating RNA, and performing the hybridizations. Degradation of RNA was not a problem in the overall transcription experiments of Agazie *et al.* (1996). However, TCA-precipitation following run-off transcription does not involve the isolation of RNA from agarose beads. The second technical problem was the nuclei being embedded in agarose. This complicated RNA extraction as standard procedures for isolating RNA yielded an insoluble agarose pellet. Both technical problems were solved by using a commercially-available solution designed for extracting and preserving RNA, and by using a procedural modification recommended by the supplier for use with high carbohydrate-containing tissues.

The experiments were carried out to determine if preincubation with triplex-specific antibodies had an effect on the expression of c-*myc*, a gene with a promoter pur•pyr tract. No specific change in expression of c-*myc* was seen in response to incubation with the triplex-specific antibodies. Overall CPM in the RNA from run-off transcription were variable, but generally decreased with increasing concentrations of both the triplex-specific antibodies and with Jel 42, but not with BSA (Table 3.6). Hybridization results of probing specific genes with the RNA were very variable both within an experiment and from one experiment to the next. This can be seen in the large S.D. values in Table 3.7. Furthermore, expression of c-*myc*, and MGMT and the gene for 18S rRNA which lack promoter pur•pyr tracts, all decreased with increasing concentrations of both the triplex-specific antibodies and with Jel 42, but again not with BSA (Table 3.7).

There were a number of sources of the experimental variability seen. Preparation of the agarose-encapsulated nuclei and isolation of RNA was more complex than parallel procedures with conventionally-prepared nuclei. Some variability therefore likely arose from differences in how samples were handled. Again, because of the technical complexity, it was not feasible to process more than six samples per experiment. As a result, opportunities for statistical analysis were limited. Second, the degree to which RNA was extracted from the agarose beads following run-off transcription varied. Finally, the greatest variability arose from the hybridization and washing steps. Gene probes were duplicated on each half of a membrane strip. Frequently, the hybridization band was darker on one

half of the strip than on the other. Additionally, the density of hybridization to the same gene probe on different membrane strips varied. Every effort was made to correct for the variation in background between different halves of the strips and between different strips: individual background controls were used for each half strip. However, hybridizations were carried out in bags made in the laboratory. These differed in size, although the same amount of buffer was added to each. Leakage of buffer and RNA occurred when it was added to bags. Finally, exposure of each probe to the RNA solution may have been inconsistent in the "static" system where the bags were floated in a dish on the surface of a water bath.

An average total CPM of 1.1×10^4 was obtained in the RNA following its isolation after run-off transcription (Table 3.6). This was approximately ten-fold lower than in the hypotonically-prepared nuclei (Section 3.3.1). However, it can be explained by the fact that the number of agarose-encapsulated nuclei used were about ten-fold less than the hypotonically-prepared nuclei. Conventional run-off transcription procedures allow for the incorporation of $1\text{-}10 \times 10^6$ CPM, although good results have been obtained with 1×10^6 CPM (Ausubel *et al.*, 1995). The CPM obtained with the agarose-encapsulated HT-29 nuclei was therefore likely too low for adequate sensitivity. This was a direct result of the fact that only 5×10^6 nuclei were used in these experiments whereas incorporation of labelled UTP into RNA has been shown to increase significantly between 5×10^6 and 5×10^7 nuclei, with 5×10^7 being recommended (Ausubel *et al.*, 1995).

From the results, it was concluded that, as used, the agarose-encapsulated nuclei were not a good system for studying the expression of individual genes following incubation with triplex-specific antibodies. The assay is inherently insensitive due to the low numbers of nuclei used. However, the procedures could be improved in several ways. The variation in results resulting from the hybridization and washing steps could be removed by using commericially-available hybridization tubes and hybridization oven.

Second, a binding curve could be constructed to show the concentration of the triplex-specific antibodies falling into the plateau range of binding to the nuclei. The initial antibody concentration of 120 µg/ml was used as being twice that used by Agazie *et al.* (1996), as double the number of nuclei were used in the present experiments. Although the same

antibody binding plateau may be present in the HT-29 nuclei as in the MOPC nuclei (Agazie *et al.*, 1996), it was not investigated. It is possible that for the agarose-encapsulated HT-29 nuclei, 120 µg/ml of antibody falls into the range where it begins to induce triplex formation.

The highest concentrations of antibodies used likely caused non-specific condensation. Evidence for this is that CPM in the RNA pellet decreased with increasing concentrations of antibodies, including Jel 42. Additionally, expression of all three genes decreased with increasing concentrations of all three antibodies. Neither of these trends was seen with high concentrations of BSA. High antibody concentrations in the range of 100 to 1,000 µg/ml, when 250 µl of packed nuclei are present, have been hypothesized to drive the formation of Z-DNA (Wittig *et al.*, 1989). Because they are anti-DNA antibodies, both Jel 318 and Jel 466 have positively charged amino acids in the combining region (Mol *et al.*, 1994; Agazie *et al.*, 1994). This could have promoted non-specific binding to DNA at higher concentrations. Finally, non-specific binding may have occurred with Jel 42 used at a concentration of 120 µg/ml. This is suggested by the 20% inhibition of transcription seen in the TCA-precipitable counts.

4.7 Triplex Formation in the Human C-*Src* Promoter

The promoter region of human c-*src* contains four pur•pyr tracts (Figure 1.13). *In vitro* experiments were carried out to determine if the tracts could form a triplex. This was done by investigating which pair of the three major tracts formed the strongest dimers on an agarose gel, following incubating under conditions favoring the formation of a pyr•pur•pyr triplex. TC1, TC2, and TC3 were each isolated alone, together with their nearby promoter sequences, in linear plasmid fragments of unique length. TC1 and TC3 formed the strongest triplex when incubated together, with the third strand being contributed by the pyrimidine strand of TC3 (Figure 3.27).

These results can be used to propose how TC1 and TC3 might form a triplex *in vivo* which is involved in the regulation of transcription (Figure 4.2). In this model, the pyrimidine strand of TC3 is opposite in orientation to its like strand in the duplex, as required (Arnott and Selsing, 1974). TC3 contains a site of frequent transcription initiation, labelled "e" (Dr. K. Bonham, personal communication). Therefore, one possibility is that the

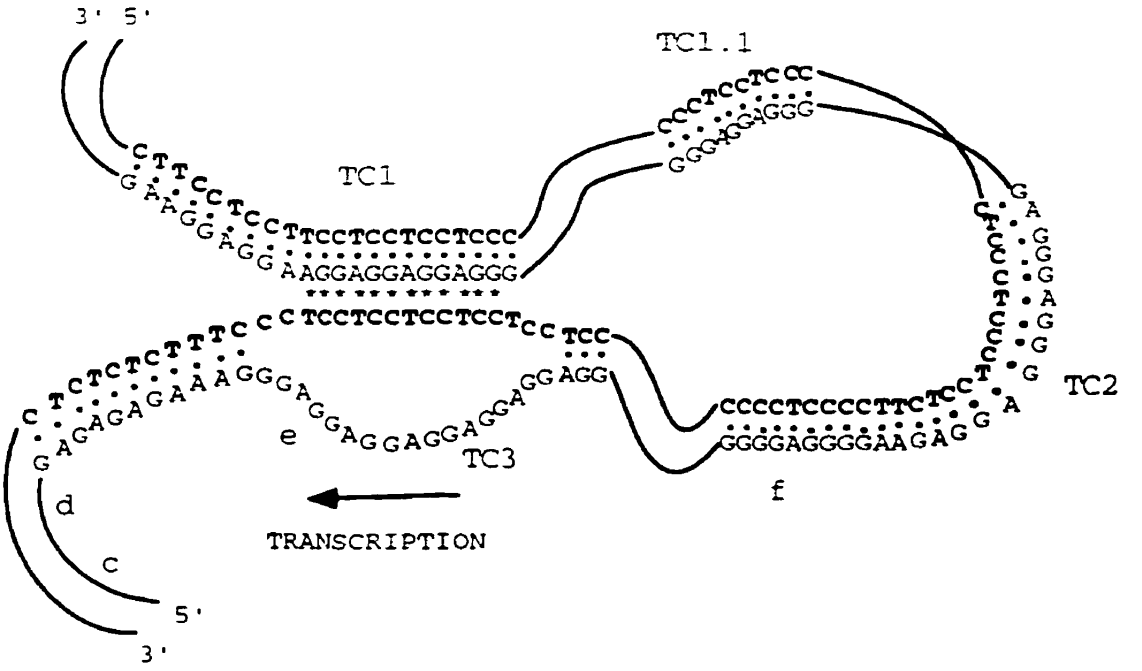


Figure 4.2. Model of possible transmolecular triplex formed *in vivo* between TC1 and TC3 in the c-*src* promoter. c, d, e, f - common sites for initiation of transcription. * - Hoogsteen hydrogen bond between third strand of triplex and purine strand of duplex, • - hydrogen bond between duplex strands. The direction of transcription from the purine strand of TC3 is shown.

single purine template strand extruded by triplex formation could promote entry of RNA polymerase to initiate transcription at this site. There are 72 bases between the 3' end of TC1 and the 5' end of TC3 (Figure 1.13). Therefore the tracts are widely separated and such a triplex would not be classified as H-DNA. Rather, a transmolecular triplex is formed. One can imagine that triplex formation could proceed in a similar way to that for the formation of the T-loops (Figure 3.13). The TC1 duplex would wrap around the pyrimidine strand of TC3, passing once through the bubble for each turn of triplex. It is conceivable that some small loops might form as with T-loops. When the factors promoting the formation of the triplex are removed, the structure could be maintained by a hydrogen-bonded knot as in the T-loops. The triplex drawn is 12 bases long. In fact, a longer triplex is possible if non-standard base triads are formed (see discussion below for Figure 4.3).

As TC1 and TC2, and TC2 and TC3 also formed weak dimers, there are two other possibilities for triplex formation (Figure 4.3A,B). Besides the standard C•G•C$^+$ and T•A•T base triads, some license has been taken in the depiction of the alternative possibilities. Two unusual triads have been used in modelling how the triplexes may form. The nonstandard C•G•T is relatively easily accomodated into a pyr•pur•pyr triple helix (Soyfer and Potaman, 1996). However, T•A•C would be less stable, distorting the triple helix more. A triplex 22 bases long is proposed between TC1 and TC2 (Figure 4.3A). However, eight of the base triads are nonstandard. Two sets of conventional base triads five and six bases in length, respectively, flank the nonconventional triads in the model. This structure could favor the initiation of transcription at site "f". In fact, "f" is the major transcription initiation site for all cell lines examined, except testes, although no colon cancer cell lines have been examined (K. Bonham - personal communication).

A triplex 19 bases long containing five nonstandard base triads is proposed between TC2 and TC3 (Figure 4.3B). This proposed triplex contains a straight series of nine conventional base triads. In this alternative model, access by RNA polymerase to site "e" would be blocked, thus repressing transcription. Nothing can be said about initiation from two other commonly used sites, c and d, from the models proposed.

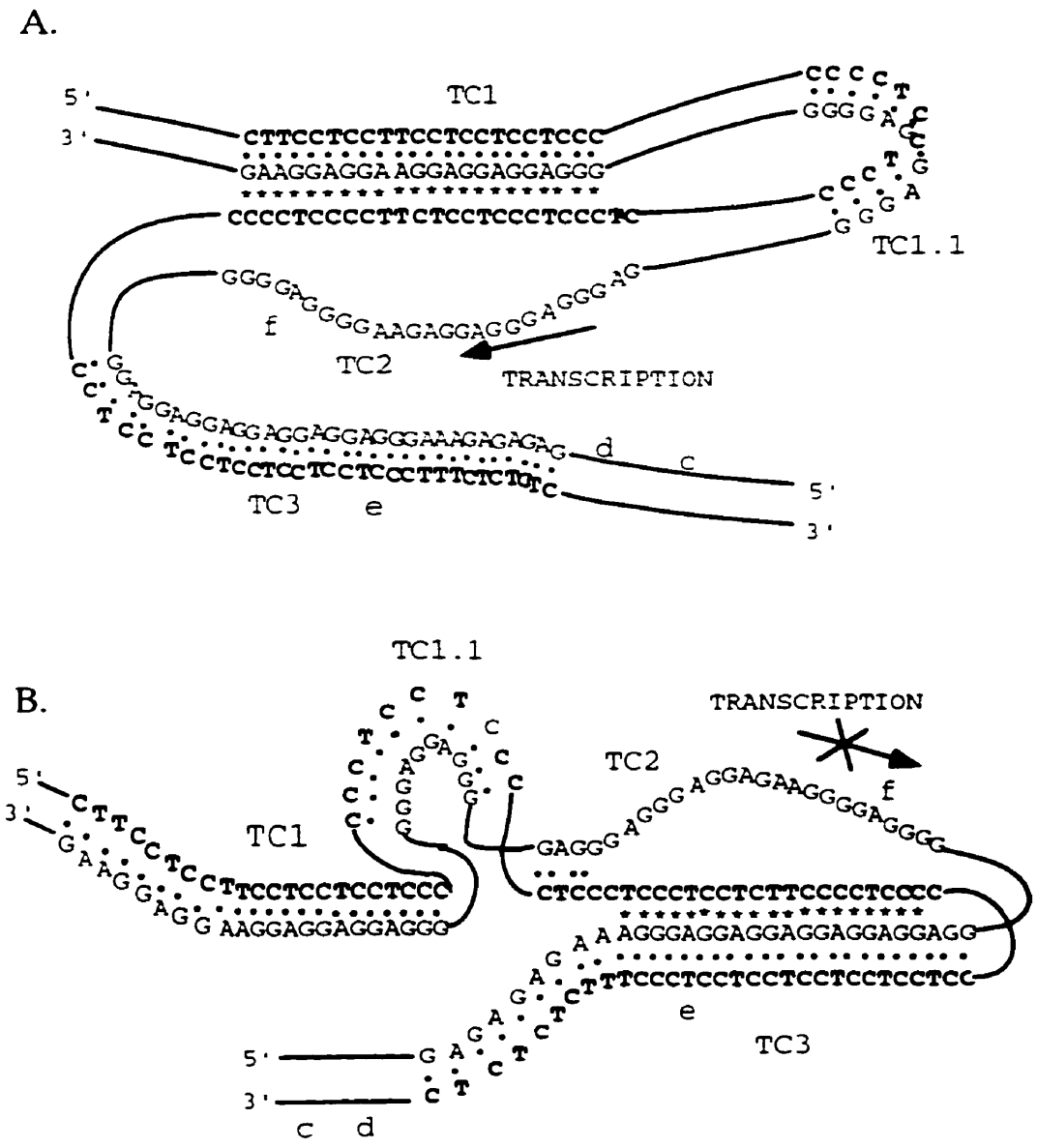


Figure 4.3. Alternative models for transmolecular triplex formation in the c-*src* promoter. Explanations as in Figure 4.2. A. Triplex formation between TC1 and TC2. B. Triplex formation between TC2 and TC3.

The experiments did not investigate whether TC1.1 could form a triplex with the other pur•pyr tracts. However, an examination of the base triads possible shows that a triplex eight bases in length could be formed between the TC1.1 duplex and the pyrimidine strand of TC2, using conventional C•G•C$^+$ and T•A•T triads. A triplex of nine conventional base triads could be formed between the TC1.1 duplex and the pyrimidine strand of TC3. Thus on the basis of sequence complementarity, TC1.1 in theory appears capable of forming triplexes with TC2 or TC3 which are at least as stable as the proposed structures in Figure 4.3A and B.. Thus, even just considering the double pyrimidine class of triplex, a number of possibilites for *in vitro* triplex formation exist for these pur•pyr tracts.

However, triplex formation between pairs of pur•pyr tracts *in vitro* does not mean that the same triplexes could be formed *in vivo*. Triplex formation *in vivo* would depend on possible supercoiling providing favorable energetics, and the presence of stabilizing factors such as single-strand binding proteins and cations. It is difficult to assess how favorable triplex formation between TC1 and TC3 would be compared to triplex formation between other pairs of tracts. Longer tracts were proposed for the two alternative possibilities in Figure 4.3. However, the effect of incorporating several nontraditional triads on the stability of the proposed triplex structures is unknown. The nontraditional triad possibilities may in fact be destabilizing as stronger dimers were formed between TC1 and TC3 *in vitro*. The model in Figure 4.2 shows that a longer tract of conventional base triads can be formed between TC1 and TC3 than between the pairs of tracts in either of the two alternative models in Figure 4.3, or as discussed above with TC1.1. Triplex formation between TC1 and TC3 should therefore be more favorable *in vivo* on the basis of greater sequence complementarity between the two tracts.

Formation of a transmolecular triplex between TC1 and TC3 would require a duplex loop of 72 bp. It is uncertain how energetically favorable the reasonably tight bend to form the triplex would be. However, the length of duplex DNA which winds 1.8 times around a nucleosome is only 146 bp. Additionally, at least some pur•pyr tracts are capable of bending around a nucleosome (Birnboim *et al.*, 1976; Puhl and Behe, 1995; Brahmachari *et al.*, 1995). The presence of a number of dispersed A•T bp in the bend region

might aid in the compression. Thus the bend in the duplex DNA to form the transmolecular triplex of Figure 4.2 appears possible.

Triplex formation between the other pairs of tracts could be regarded as forming intramolecular triplexes because of the shorter distances between the pairs of tracts. Theoretically, all of the triplex structures discussed are of sufficient length for stability as an H-DNA structure. A lower limit of 15 bp has been postulated for H-DNA formation (Lyamichev *et al.*, 1989). The longer the tract involved, the lower the density of supercoiling required to induce it (Htun and Dahlberg, 1989). However, H-DNA formation requires a short insert between the two halves of a mirror repeat pur•pyr tract, with the most energetically favorable insert length being four to five bases (Harvey *et al.*, 1988). As the length of the insert increases, for example from four to 12 bases, formation of H-DNA is less energetically favorable, and requires a higher supercoiling density (Parniewski *et al.*, 1989; Shimizu *et al.*, 1989). Thus, triplex formation *in vivo* between TC1 and TC2 would be highly unfavorable because of the insert length of 35 bases (Figure 1.13). This would require either a highly energetically unfavorable length of single strand foldback or very tightly bent duplex DNA. Intramolecular triplex formation between TC1 and TC1.1 or TC1.1 and TC2, would also be relatively unfavorable as the lengths of the insert sequences in the single strand folding back would be 14 and 10 bp, respectively. Likewise, triplex formation between TC2 and TC3 would also be unfavorable as the length of the insert sequence would be 12 bases. Therefore, this could be another reason why a triplex between TC1 and TC3 would be more likely to occur *in vivo*.

However, if high levels of supercoiling were generated from regions downstream as the gene is transcribed or even from a gene 3' to c-*src*, sufficient torsional strain might be present to promote the formation of one of these alternative intramolecular triplexes. Proteins binding along this region of the promoter could also be envisaged as having a role in constraining the strands and either providing the torsional strain to drive triplex formation or preventing formation.

The c-*src* promoter has been shown to adopt a non-B conformation in supercoiled plasmids (K. Bonham - personal commun.). Both strands in the region of TC3 and often part of TC2 were S1 nuclease sensitive. This is evidence that the region of TC3 is single-stranded and supports the model in Figure 4.2. It is possible that if a transmolecular triplex formed between TC1

and TC3 involving the pyrimidine strand of TC3, that some of TC2 might also be single-stranded.

Expression of *c-src* has recently been found to be under control of the B-DNA binding Sp1 family of transcription factors (K. Bonham - personal communication). The promoter region -634 to -462, containing the pur•pyr tracts, TC1 and TC1.1, is required for full activity (Figure 1.13). This region contains two perfect Sp1 consensus binding sequences, GC1 located upstream of TC1, and GA2 which overlaps TC1.1. These two binding sites are critical for optimal transcriptional activity: mutations or deletions in one or both of TC1 and TC1.1 leads to a decrease in transcription. In particular, mutation of the TC1.1/GA2 Sp1 binding site results in about a 50% reduction in transcription. This means that TC1.1 is an essential regulatory site for the gene. TC1, TC2, and TC3 each contain weaker binding sites for Sp1. Sp3 was found to be a negative regulator of expression from the gene, binding to the same sites. A simple model for the regulation of transcription from this promoter could be that transmolecular triplex formation between TC1 and TC3 would prevent the binding of Sp1, by physically making the site inappropriate for Sp1 binding. If Sp3 binds to the same tract as a repressor, then triplex formation could turn the gene on.

A factor called SPy has been found to bind very strongly to the pyrimidine strand of TC1, which contains two binding sites for the factor (Figure 1.13; K. Bonham - personal communication). SPy also binds to the duplex form of the tract, but appears to bind preferentially to the pyrimidine strand. Another binding site is present in TC2. It is not yet known if this factor is specific for these tracts, or if the interaction is non-specific.

A number of TATA-less promoters in mammalian genes contain pur•pyr tracts. Among these is the gene for the androgen receptor (Chen *et al.*, 1997). The *c-src* promoter resembles that of the androgen receptor gene. The androgen receptor gene promoter region, like that of *c-src*, lacks a TATA or CAAT box, and contains a pur•pyr tract. The tract is located upstream of a strong Sp1-binding GC box. Examination of this region of the promoter using S1 nuclease showed that both strands were sensitive, indicating the tract could adopt either an H-DNA or H*-DNA conformation. The pur•pyr tract weakly bound Sp1 and its pyrimidine strand bound a single-strand specific protein called ssPyrBF. The binding of these two proteins might be expected to be mutually exclusive. The function of ssPyrBF is unknown, but

Sp1 is a strong promoter of transcription. Therefore, binding of ssPyrBF could play a negative regulatory role. The authors postulated that multiple weak Sp1 binding sites in the pur•pyr tract served as a reservoir for Sp1 when the tract was in the duplex conformation. Sp1 binding in the TATA-less promoter functions to recruit and stabilize TFIID which is essential for assembly of the Pol II initiation complex (Pugh and Tjian, 1991). These weak binding sites could supply Sp1 to the adjacent GC box, facilitating the assembly of TFIID and the rest of the initiation complex at the downstream initiation site. Intramolecular triplex formation would prevent the interaction of the pur•pyr tract with Sp1 and would hence downregulate expression from the gene. Both H-DNA and H*-DNA could form, but interaction with ssPyrBF would preferentially stabilize the H*-DNA conformation. Gee *et al.* (1992) had previously observed that triplex formation prevented Sp1 binding to the promoter of the dihydrofolate reductase gene. A negative regulatory role of a single strand pyrimidine tract binding protein has also been proposed for the pur•pyr tract of the promoter of the malic enzyme gene, which has also been shown to form an intramolecular triplex (Xu and Goodridge, 1996).

Several TATA-less promoters such as are found in the dihydrofolate reductase gene contain more distant Sp1 binding sites that have been shown to interact with Sp1 bound to the promoter GC box via duplex DNA looping and protein-protein interactions (Su *et al.*, 1991). This may be the case in the c-*src* promoter as the other strong binding site, GC1, is located about 120 bp upstream. As is the situation in the androgen receptor gene promoter pur•pyr tract, the TC1 tract in the c-*src* promoter binds both a pyrimidine strand-specific protein, SPy, and provides a weak binding site for Sp1. One difference is that the androgen receptor promoter pur•pyr tract provides multiple weak binding sites for Sp1 (Chen *et al.*, 1997). However, the weak Sp1 binding site in TC1 may be a reservoir to provide Sp1 to the immediately adjacent strong Sp1 binding site in TC1.1/GA2. This could conceivably be supplemented by Sp1 from the additional weak binding site just downstream in TC2. An attractive hypothesis is that in binding to TC1.1, Sp1 could recruit the binding of TFIID which would facilitate the assembly of the initiation complex, promoting transcription from the commonly used downstream sites c, d, e, and f.

Again, the function of SPy binding is unknown, but as was the case for ssPyrBF, its binding could play a negative regulatory role. Regulation of transcription from this promoter in c-*src* could depend on the ratio of Sp1 to SPy. In the presence of a dominating amount of Sp1 some could bind weakly to the site in TC1 and perhaps in TC2, providing a reservoir for the strong site in TC1.1 leading to transcription initiation. If the ratio of SPy is higher, it could bind to the pyrimidine strand of TC1. Formation of both a double pyrimidine and double purine triplex might be possible. However, as yet no single-stranded binding protein specific for the purine strand of TC3 has been identified. Such a protein would favor the formation of a double pyrimidine triplex. Accordingly, SPy would drive the equilibrium to forming a double purine triplex. This would downregulate transcription as Sp1 would be unable to bind.

One hypothesis held in the laboratory of Dr. K. Bonham is that a pyr•pur•pur class of triplex is formed in the c-*src* promoter, resulting in the extrusion of the pyrimidine strand of TC1. In footprinting studies of the promoter using purine oligonucleotides designed to form triplexes with each of the four pur•pyr tracts, it was discovered that the TFO designed to form a triplex with TC1 also formed a triplex with TC3 (Ritchie and Bonham, 1998). However, the opposite was not true. That is, the purine oligonucleotide designed to form a triplex with TC3 did not form a triplex with TC1. These results suggested that a pyr•pur•pur triplex could be formed *in vivo* with the third strand being the purine strand of TC1. The finding of *in vitro* strong pyr•pur•pyr triplex formation between TC1 and TC3, together with the TFO cross-reactivity between TC1 with TC3, confirms that if a triplex forms in these sequences *in vivo*, it is most likely to form between TC1 and TC3. Whether or not a double purine or double pyrimidine triplex forms may depend on the presence of the appropriate stabilizing factors. The fact that S1 nuclease digestion showed sensitivity of both tracts in the region of TC3, extending into TC2, suggests that a triplex formed between TC1 and TC3 could exist in a conformational equilibrium between the two classes of triplex, as suggested for the androgen receptor gene. It is also possible that a mixed structure, containing both a double pyrimidine triplex and a double purine triplex, could occur here. In the case of intramolecular triplexes, multiple nucleation events have been observed in pur•pyr tracts $\geq$ 35 bp resulting in a mix of conformers (Htun and

Dahlberg, 1989; Panyutin and Wells, 1992; Kohwi and Kohwi-Shigematsu, 1993).

The discussion above shows that very little is known as yet about the regulation of transcription in c-*src*. An analogy has been drawn with the promoter of another gene which is also TATA-less and contains a pur•pyr tract. The results show that formation of a transmolecular double pyrimidine triplex between the TC1 duplex and the pyrimidine strand of TC3 is possible *in vivo*. This conclusion is based on it being the most favorable triplex energetically out of several options from a theoretical point of view, and on its *in vitro* stability. However, other possibilities exist for triplex formation. A double purine triplex between TC1 and TC3 has been proposed based on footprinting studies. In addition, a protein has been found which binds the pyrimidine strand of TC1. The conformation or conformations adopted by this promoter *in vivo* will be due to the combined interaction of stabilizing factors such as supercoiling, polyamines, and DNA binding factors. The different possibilities may represent alternative conformational states of the promoter which are involved in the regulation of transcription. The fact that transcription is initiated from a number of sites in the promoter supports this possibility. Finally, the results discussed here emphasize the mutually exclusive interplay that could exist between transcription factors and triplex DNA in the regulation of transcription.

4.8 Conclusions and Future Directions

4.8.1 Biological Relevance of Polypyrimidine•Polypurine•Polypyrimidine Triplexes

Most of the work presented in this thesis has involved the formation of a pyr•pur•pyr class of triplex. Even the triplex-specific antibodies used in the experiments involving agarose-encapsulated nuclei were directed against T•A•T and C•G•C^+ triads. *In vitro*, this class of triplex requires an acidic, non-physiological pH for formation. As the results presented in this thesis are *in vitro* studies aimed at investigating the possible role that triplexes might have in the cell, it is appropriate at this point to review why the double pyrimidine class of triplex has the potential to be relevant biologically.

It has frequently been proposed that pyr•pur•pur triplexes have more biological relevance as they do not require low pH for formation *in vitro*. Indeed, the requirement of protonation of cytosine for Hoogsteen base pairing in pyr•pur•pyr triplexes was initially a cause for concern in considering a model for such triplex formation *in vivo*. The pK of N3 of cytosine as a free base is 4.3 (Inman, 1964a,b). On the other hand, nuclear pH is 7.1 to 7.5 (Cody *et al.*, 1993; Masuda *et al.*, 1998). However, transient shifts in intracellular pH have been documented (Hardin, 1993). Furthermore, the pK of polycytidylic acid is 7.3 to 7.5 (Gray *et al.*, 1987). In mixed sequence DNA stretches within the chromosome, it seems likely that some cytosines will be protonated at physiological pH. This has been confirmed using NMR (Sklenar and Feigon, 1990). A triplex containing both T•A•T and C•G•C^+ triplets was formed at pH 5.5. When the pH was raised to 8, about 50% of the cytosines remained protonated. Recently, the pK of the protonation of third strand cytosines when embedded within a triplex was estimated taking into account the strong electrostatic field in a triple helix. Calculations were based on three different assumptions about the interior dielectric constant of the triplex helix. In the resulting three estimates of the pK of cytosine in the helix were 4.04, 8.70, and 10.00 (Pack *et al.*, 1998). Therefore, the requirement for the protonation of cytosine *in vitro* does not appear to rule out the formation of pyr•pur•pyr triplexes *in vivo*, especially considering the existence of other stabilizing factors.

As discussed in Section 1.6.3, a number of other factors are present in the nucleus which may promote and stabilize the formation of a double pyrimidine class of triplex. Single-stranded DNA-binding proteins specific for the pyrimidine strand of pur•pyr tracts have been discovered, as well as one or more proteins which preferentially bind double pyrimidine triplexes. Local supercoiling density might also favor the formation of H-DNA and possibly some transmolecular triplexes. Other promoting factors would be length of tract. For example, H-DNA was observed in the tract $d(GA)_{37}$•$d(TC)_{37}$ at neutral pH in the presence of a moderate amount of supercoiling (Collier and Wells, 1990). Other factors would be the local mix of cations, particularly polyvalent cations. Hampel *et al.* (1991) observed triplex formation at neutral pH in the presence of micromolar concentrations of polyamines. In fact, because the polyamines spermine and spermidine are present in the nucleus in up to millimolar concentrations,

these authors proposed that a pyr•pur•pyr triplex conformation may be preferred over a B-DNA conformation for pur•pyr tracts.

Formation of a pyr•pur•pyr triplex may in fact be more relevant physiologically as it could occur more readily. First, which may be of importance for the looping in the formation of a transmolecular triplex, the partial protonation of the third strand may aid its recognition of the complementary duplex. Secondly, formation of a pyr•pur•pur triplex requires high $Mg2^+$ concentrations (Pilch *et al.*, 1991). Such high ionic strength may hinder the breathing of DNA which is a prerequisite to initiation of triplex formation by reducing DNA flexibility (Thomas and Bloomfield, 1983a). The low pH and low ionic strength conditions required for the formation of the pyr•pur•pyr triplex may encourage breathing of the duplex. Finally, a study using 14 combinations of base triads showed that T•A•T and $C•G•C^+$ triads formed more readily and were more specific (Fosella *et al.*, 1993).

4.8.2 Transmolecular Triplex Formation *In Vivo*

The work presented in this thesis has been a detailed enquiry into the role that triplex DNA might play in the regulation of transcription. A number of different aspects of this topic were investigated. A special emphasis was placed on the possible role of transmolecular triplex formation. *In vitro* studies showed there is no topological impediment to the formation of transmolecular triplexes in circular and hence chromosomal loops of DNA. Such triplexes could play both a structural and a regulatory role. In transmolecular triplex formation there is no requirement for supercoiling or mirror repeat symmetry in individual pur•pyr tracts as is the case for intramolecular triplex formation. It has been observed that not all pur•pyr tracts, whether occurring in the 5' region of a gene or elsewhere, have mirror repeat symmetry or are sufficiently long enough to form intramolecular triplexes without additional stabilizing factors such as very high levels of supercoiling (Bucher and Yagil, 1991). The possibility of such tracts being involved in transmolecular triplex formation increases the number of pur•pyr tracts which may form triplexes *in vivo*.

4.8.3 Future Directions

Several aspects of the topics investigated in this thesis present themselves for further investigation. Single-strand DNA-binding proteins specific for one or the other strand of a pur•pyr tract have been identified and it has been proposed that they may play a role in promoting and/or stabilizing intramolecular (and in this thesis) transmolecular triplexes. There is some evidence, however, that in pairing with single DNA strands, single-stranded DNA binding proteins may in fact destabilize triplexes. Prokaryotic SSB has been used in the literature as a representative single-strand binding protein and triplex destabilization effects have been reported. The effect of single-strand binding proteins on triplex stability could be explored further using specific examples of promoter pur•pyr tracts able to form H-DNA and their purine strand binding proteins. There may be difficulty in providing the appropriate *in vitro* conditions for such experiments as a low pH is required, which might reduce the activity of the proteins.

Another possible line of investigation involves an application of agarose-embedded nuclei and a new use for the triplex-specific antibodies developed in the laboratory of Dr. J.S. Lee. If successful, this work could significantly add to the direct evidence already produced with the antibodies that triplexes exist in the cell. The triplex-specific antibodies could be used in an attempt to "pull out" the pur•pyr tract in the 5' flanking region of the human γ-globin gene in triplex conformation. There is good evidence that this tract exists as a triplex at some stage in development, as it is involved in the switch from expression of fetal to adult hemoglobin. This topic was discussed in some detail in Section 1.5.3. Erythroid cell lines at different developmental stages could be used for the investigation. The protocol for identifying promoter sequences in Z-DNA conformation using anti-Z-DNA antibodies has been well worked out and has been used successfully to identify sequences present as Z-DNA in the promoters of several genes (Wittig *et al.*, 1992; Müller *et al.*, 1996).

Finally, *in vitro* studies reported in this thesis and by Ritchie and Bonham (1998) using linear fragments have shown that the c-*src* TC1 and TC3 tracts can come together to form either a double purine or a double pyrimidine triplex. A number of triplex structures are possible when one

considers all four pur•pyr tracts, even if some appear less energetically feasible. Because of the complexity of the four pur•pyr tracts in close proximity, the next analysis to be carried out on this region would be to see how the tracts come together in a supercoiled plasmid *in vitro*. Chemical modification studies could be carried out to analyze exactly which tracts are single-stranded. One could incubate the plasmid under conditions both promoting the formation of a double pyrimidine triplex and a double purine triplex. This would enable one to see how closely any triplex in a supercoiled plasmid matched the models in Figures 4.2 and 4.3. The analysis could then be extended to the *in situ* situation to analyze how the tracts come together when cloned in a plasmid in *E. coli*.

5.0 REFERENCES

Adachi, Y.A., Luke, M., and Laemmli, U.K. (1991). Chromosome assembly *in vitro*: Topoisomerase II is required for chromosome condensation. Cell 64: 137-148.

Agazie, Y.M., Burkholder, G.D., and Lee, J.S. (1996). Triplex DNA in the nucleus: direct binding of triplex specific antibodies and their effect on transcription, replication and cell growth. Biochem. J. 316: 461-466.

Agazie, Y.M., Lee, J.S., and Burkholder, G.D. (1994). Characterization of a new monoclonal antibody to triplex DNA and immunofluroescent staining of mammalian chromosomes. J. Biol. Chem. 269: 7019-7023.

Aharoni, A., Baran, N., and Manor, H. (1993). Characterization of a multisubunit human protein which selectively binds single-stranded $d(GA)_n$ and $d(GT)_n$ sequence repeats in DNA. Nucleic Acids Res. 21: 5221-5228.

Aiyar, S.E., Helmann, J.D., and deHaseth, P.L. (1994). A mismatch bubble in double-stranded DNA suffices to direct precise transcription initiation by *Escherichia coli* RNA polymerase. J. Biol. Chem. 269: 13179-13184.

Arnott, S., and Bond, P.J. (1973). Structures for poly (U)•poly(A)•poly(U) triple-stranded polynucleotides. Nature New Biol. 44: 99-101.

Arnott, S., and Selsing, E. (1974). Structures of poly d(A)•poly d(T) and poly d(T)•poly d(A)•poly d(T). J. Mol. Biol. 88: 509-521.

Arnott, S., Bond, P.J., Selsing, E., and Smith, P.J.C. (1976). Models of triple-stranded polynucleotides with optimized stereochemistry. Nucleic Acids Res. 3: 2459-2470.

ATCC Catalogue of cell lines and hybridomas (1992). 7th ed., American *Type Culture* Collection, Maryland, U.S.A.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds) (1995). Current Protocols in Molecular Biology. 4th ed. John Wiley and Sons, Toronto, Canada.

Bacolla A., Ulrich, M.J., Larson, J.E., Ley, T.J., and Wells, R.D. (1995). An intramolecular triplex in the human γ-globin 5'-flanking region is altered by point mutations associated with hereditary persistence of fetal hemoglobin. J. Biol. Chem. 270: 24556-24563.

Baran, N., Lapidot, A,., and Manor, H. (1991). Formation of DNA triplexes accounts for arrests of DNA synthesis at $d(TC)_n$ tracts. Proc. Natl. Acad. Sci. U.S.A. 88: 507-511.

Baran, N., Neer, A., and Manor, H. (1983). "Onion skin" replication of integrated polyoma virus DNA and flanking sequences in polyoma-transformed rat cells: termination within a specific cellular DNA segment. Proc. Natl. Acad. Sci. U.S.A. 870: 105-109.

Barawkar, D.A., Rajeev, K.G., Kumar, V.A., and Ganesh, K.N. (1996). Triplex formation at physiological pH by 5-Me-dC-N^4-(spermine)[X] oligonucleotides: non protonation of N3 in X of X*G:C triad and effect of base mismatch/ionic strength on triplex stabilities. Nucleic Acids Res. 24: 1229-1237.

Beal, P.A., and Dervan, P.B. (1991). Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. Science 251: 1360-1363.

Beard, P., Morrow, J.F. and Berg, P. (1973). Cleavage of circular, superhelical simian virus 40 DNA to a linear duplex by S1 nuclease. J. Virol. 12: 1303-1313.

Beasty, A.M. and Behe, M.J. (1988). An oligopurine sequence bias occurs in eukaryotic viruses. Nucleic Acids Res. 16: 1517-1528.

Bebenek, K., Abbots, J., Roberts, J.D., Wilson, S.H., and Kunkel, T.A. (1989). Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J. Biol. Chem. 264: 16948-16956.

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.

Belotserkovskii, B.P., Krasilnikova, M.M., Veselkov, A.G., and Frank-Kamenetski, M.D. (1992) Kinetic trapping of H-DNA by oligonucleotide binding. Nucleic Acids Res. 20: 1903-1908.

Bentley, D.L. and Groudine, M. (1986a). A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. Nature 321: 702-706.

Bentley, D.L. and Groudine, M. (1986b). Novel promoter upstream of the human c-*myc* gene and regulation of c-*myc* expression in B-cell lymphomas. Mol. Cell. Biol. 6: 3481-3489.

Bernues, J., Beltran, R., Casasnovas, J.M., and 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 J. 8: 2087-2094.

Biggin, M.D., and Tjian, R. (1988). Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. Cell 53: 699-711.

Birnbiom, H.C., Sederoff, R.R., and Paterson, M.C. (1979). Distribution of polypurine•polypyrimidine segments in DNA from diverse organisms. Eur. J. Biochem. 98: 301-307.

Birnboim, H.C. (1975). Polypyrimidine segments in *Drosophila melanogaster* DNA: I. Detection of a cryptic satellite containing polypyrimidine/polypurine DNA. Cell 5: 173-181.

Birnboim, H.C. (1978). Spacing of polypyrimidine regions in mouse DNA as determined by poly(adenylate, guanylate) binding. J. Mol. Biol. 121: 541-559.

Birnboim, H.C., Holford, R.M., and Seligy, V.L. (1976). Random phasing of polypyrimidine/polypurine segments and nucleosome monomers in chromatin from mouse L cells. Cold Spring Harbor Symp. Quant. Biol. 39: 1161-1165.

Blackburn, E.H., and Szostak, J.W. (1984). The molecular structure of centromeres and telomeres. Ann. Rev. Bioch. 53: 163-194.

Boles, T.C., and Hogan, M. E. (1987). DNA structure equilibria in the human *c-myc* gene. Biochemistry 26: 367-376.

Bonham, K. and Fujita, D.J. (1993). Organization and analysis of the promoter region and 5' non-coding exons of the human c-*src* proto-oncogene. Oncogene 8: 1973-1981.

Boulikas, T. (1993). Nature of DNA sequences at the attachment regions of genes to the nuclear matrix. J. Cell. Biochem. 52: 14-22.

Bowater, R., Chen, D., and Lilley, D.M.J. (1994). Elevated unconstrained supercoiling of plasmid DNA generated by transcription and translation of the tetacycline resistence gene in eubacteria. Biochemistry 33: 9266-9275.

Brahmachari, S.K, Meera, G., Sarkar, P.S., Balagurumoorthy, P., Tripathi, J., Raghavan, S., Shaligram, U., Pataskar, S. (1995). Simple repetitive DNA sequences in the genome: structure and functional significance. Electrophoresis 16: 1705-1714.

Brahmachari, S.K., Sarkar, P.S., Raghavan, S., Narayan, M., and Maiti, A.K. (1997). Polypurine/polypyrimidine sequences as cis-acting transcriptional regulators. Gene 190 (S1): 17-26.

Brahms, J.G., Dargouge, O., Brahms, S., Ohara, Y., and Vagner, V. (1985). Activation and inhibition of trancription by supercoiling. J. Mol. Biol. 181: 455-465.

Broitman, S.L., Im, D.D., and Fresco, J.R. (1987). Formation of triple stranded polynucleotide helix, poly(A•A•U). Proc. Natl. Acad. Sci. U.S.A. 84: 5120- 5124.

Brunel, F., Alzari, P.M., Ferrara, P., and Zakin, M.M. (1991). Cloning and sequencing of PYBP, a pyrimidine-rich specific single strand DNA-binding protein. Nucleic Acids Res. 19: 5237-5245.

Bucher, P. and Yagil, G. (1991). Occurrence of oligopurine-oligopyrimidine tracts in eukaryotic and prokaryotic genes. DNA Sequence 1: 157-172.

Burgi A.W., Robinton J., Carlson C.L. (1974). Studies on the folded chromosome of *Eschcrichia coli.* Cold Spring Harbor Symp. Quant. Biol. 38: 43-51

Burkholder, G.D., Latimer, L.J.P., and Lee, J.S. (1988). Immunofluorescent staining of mammalian nuclei and chromosomes with a monoclonal antibody to triplex DNA. Chromosoma 97: 185-192.

Burkholder, G.D., Latimer, L.J.P., and Lee, J.S. (1991). Immunofluorescent localization of triplex DNA in polytene chromosomes of *Chironomus* and *Drosophila*. Chromosoma 101: 11-18.

Camerini-Otero, R.D. and Hsieh, P. (1993). Parallel DNA triplexes, homologous recombination, and other homology-dependent DNA interactions. Cell 73: 217-223.

Campos, J.L. and Subirana, J.A. (1991). The influence of Mg2+ and Zn2+ on polypurine/polypyrimidine multistranded helixes. J. Biomol. Struct. Dyn. 8: 793-800.

Cantor, C.R., and Efstratiadis, A. (1984). Possible structures of homopurine-homopyrimidine S1-hypersensitive sites. Nucleic Acids Res. 12: 8059-8072.

Case, S.T. and Baker, R.F. (1975). Detection of eukaryote-specific pyrimidine runs in repetitive DNA sequences and their relation to single-stranded regions in DNA isolated from sea urchin embryos. J. Mol. Biol. 98: 69-92.

Chen, S., Supakar, P.C., Vellanoweth, R.L., Song, C.S., Chatterjee, B., and Roy, A.K. (1997). Functional role of a conformationally flexible homopurine/homopyrimidine domain of the androgen receptor gene promoter interacting with Sp1 and a pyrimidine single strand DNA-binding protein. Mol. Endocrinol. 11: 3-15.

Cheng, A.-J. and Van Dyke, M.W. (1994). Oligodeoxyribonucleotide length and sequence effects on intermolecular purine-purine-pyrimidine triple-helix formation. Nucleic Acids Res. 22: 4742-4747.

Cheng, Y.K. and Pettitt, B.M. (1992). Stabilities of double and triple-strand helical nucleic acids. Prog. Biophys. Mol. Biol. 58: 225-257.

Cheng, Y.K. and Pettitt, B.M. (1992a). Hoogsteen versus reversed-Hoogsteen base pairing: DNA triple helices. J. Am. Chem. Soc. 114: 4465-4474.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochem. 162: 156-159.

Christiansen, C. and Baldwin, R.L. (1977). Catalysis of DNA reassociation by the *Escherichia coli* DNA binding protein: a polyamine-dependent reaction. J. Mol. Biol. 115: 441-454.

Christophe, D., Carber, B., Bacolla, A., Targovnik, H., Pohl, V., and Vassat, G. (1985). An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene. Nucleic Acids Res. 13: 5127-5144.

Chung, Y.T. and Keller, E.B. (1990). Regulatory elements mediating transcription from the *Drosophila melanogaster* actin 5C proximal promoter. Mol. Cell. Biol. 10: 206-216.

Clark, S.P., Lewis, C.D., and Felsenfeld, G. (1990). Properties of BPG1, a poly(dG)-binding protein from chicken erythrocytes. Nucleic Acids Res 18: 5119-5126.

Cody S.H., Dubbin, P.N., Beischer, A.D., Duncan, N.D., Hill, J.S., Kaye, A.H., and Williams, D.A. (1993). Intracellular pH mapping with SNARF-1 and confocal microscopy: I. A quantitative technique for living tissues and isolated cells. Micron 24: 573-580.

Collado-Vides, J., Magasanik, B., and Gralla, J.D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. Microbiol. Rev. 55: 371-394.

Collier, D.A. and Wells, R.D. (1990). Effect of length, supercoiling and pH on intramolecular triplex formation. J. Biol. Chem. 265: 10652-10658.

Collier, D.A., Griffin, J.A., and Wells, R.D. (1988). Non-B right-handed DNA conformations of homopurine•homopyrimidine sequences in the murine immunoglobin C alpha switch region. J. Biol. Chem. 263: 7397-7405.

Cook, P.R., and Brazell, I.A. (1975). Supercoils in human DNA. J. Cell Sci. 19: 261-279.

CRC Handbook of Chemistry and Physics (1974). D.R. Lide (Ed. in Chief), 74th ed., 1993-1994, CRC Press, Boca Raton.

Darnell, J., Lodish, H., and Baltimore, D. (1986). Molecular biology of the cell. Scientific American Books, New York, p. 618.

Daube, S.S. and von Hippel, P.H. (1992). Functional transcription elongation complexes from synthetic RNA-DNA duplexes. Science 258: 1320-1324.

Davis, R.H., Morris, D.R., and Coffino, P. (1992). Sequestered end products and enzyme regulation: the case of ornithine decarboxylase. Microbiol. Rev. 56: 280-290.

Davis, T.L., Firulli, A.B., and Kinniburgh, A.J. (1989). Ribonuclear protein and protein factors bind to an H-DNA-forming *c-myc* DNA element:

Possible regulators of the *c-myc* gene. Proc. Natl. Acad. Sci. U.S.A. 86: 9682-9686.

Dayn, A., Samadashwily, G.M., and Mirkin, S.M. (1992). Intramolecular DNA triplexes: Unusual sequence requirements and influence on DNA polymerization. Proc. Natl. Acad. Sci. U.S.A. 89: 11406-11410.

de los Santos, C., Rosen, M., and 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-triplet formation. Biochemistry 28: 7282-7289.

Denniston-Thomson, K., Moore, D.D., Kruger, K.E., Furth, M.E., and Blattner, F.R. (1977). Physical structure of the replication origin of bacteriophage lambda. Science 198: 1051-1056.

DesJardins, E. and Hay, N. (1993). Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-*myc* promoters. Mol. Cell. Biol. 13: 5710-5724.

Durand, M., Peloille, S., Thuong, N. T., and Maurizot, J.C. (1992). Triple-$(dT)_{12}$ sequences bridged by two hexaethylene glycol bridges. Biochemistry 31: 9197-9204.

Duval-Valentin, G., Thuong, N.T., and Helene, C. (1992). Specific inhibition of transcription by triplex helix-forming oligonucleotides. Proc. Natl. Acad. Sci. U.S.A. 89: 504-508.

Elgin, S.C.(1982). Chromatin structure, DNA structure. Nature 300: 402-403.

Evans, T. and Efstratiadis, A. (1986) Sequence-dependent S1 nuclease hypersensitivity of a heteronomous DNA duplex. J. Biol. Chem. 261: 14771-14780.

Evans, T., Schon, E., Gora-Maslak, G., Patterson, J. and Efstratiadis, A. (1984). S1-hypersensitive sites in eukaryotic promoter regions. Nucleic Acids Res. 12: 8043-8058.

Farrell, R.E. (1993). RNA methodologies: a laboratory guide for isolation and characterization. Academic Press, San Diego.

Fassler, J.S. and Gussin, G.N (1996). Promoters and basal transcription machinery in eubacteria and eukaryotes: concepts, definitions, and analogies. Methods Enzymol. 273: 3-29.

Felsenfeld, G. and Rich, A. (1957). Studies on the formation of two- and three-stranded polyribonucleotides. Biochim. Biophys. Acta. 26: 457-468.

Felsenfeld, G., and Miles, H.T. (1967). The physical and chemical properties of nucleic acids. Ann. Rev. Biochem. 36: 407-448.

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

Financsek, I. Tora, L., Kelemen, G. and Hedvegi, E.J. (1986). Supercoil induced S1 hypersensitive sites in the rat and human ribosomal RNA genes. Nucleic Acids Res. 14: 3263-3277.

Finer, M.H., Aho, S., Gerstenfeld, L.C., Boedtker, H., and Doty, P. (1987). Unusual DNA sequences located within the promoter region and the first intron of the chicken pro-α1(I) collagen gene. J. Biol. Chem. 262: 13323-13332.

Firulli, A.B., Maibenco, D.C., and Kinniburgh, A.J. (1992). The identification of a tandem H-DNA structure in the c-*myc* nuclease sensitive promoter element. Biochem. Biophys. Res. Commun. 185: 264-270.

Firulli, A.B., Maibenco, D.C., and Kinniburgh, A.J. (1994). Triplex-forming ability of a c-*myc*-promoter element predicts promoter strength. Arch. Biochem. Biophys. 310: 236-242.

Fischer, K.D., and Nowock, J. (1990). The T->C substitution at -198 of the A globin-gene associated with the British form of HPFH generates overlapping recognition sites for two DNA-binding proteins. Nucleic Acids Res. 18: 5685-5693.

Fossella, J.A., Kim, Y.J., Shih, H., Richards, E.G., and Fresco, J.R. (1993). Relative specificities in binding of Watson-Crick base pairs by third strand residues in a DNA pyrimidine triplex motif. Nucleic Acids Res. 21: 4511-4515.

Fowler, R.F. and Skinner, D.M. (1986). Eukaryotic DNA diverges at a long and complex pyrimidine•purine tract that can adopt altered conformations. J. Biol. Chem. 261: 8994-9001.

Frank-Kamenetskii, M. D. (1989). Waves of supercoiling. Nature 337: 206.

Freeman, L.A. and Garrard, W.T. (1992). DNA supercoiling in chromatin structure and gene expression. Crit. Rev. Euk. Gene Exp. 2: 165-209.

Garvin, A.M., Pawar, S., Marth, J.D., and Perlmutter, R.M. (1988). Structure of the murine *lck* gene and its rearrangement in a murine lymphoma cell line. Mol. Cell. Biol. 8: 3058-3064.

Gasser, S. M., and Laemmli, U.K. (1987). A glimpse at chromosomal order. Trends Genet. 3: 16-22.

Gasser, S.M., Laroche, T., Falquet, J., Boy de la Tour, E., and Laemmli, U.K. (1986). Metaphase Chromosome Structure: Involvement of Topoisomerase II. J. Mol. Biol. 188: 613-629.

Gee, J.E., Blume, S., Snyder, R.C., Ray, R., and Miller D.M. (1992). Triplex formation prevents Sp1 binding to the dihydrofolate reductase promoter. J. Biol. Chem. 267: 11163-11167.

Germond, J.E., Vogt, V.M. and Hirt, B. (1974). Characterization of the single-strand-specific nuclease S1 activity on double-strand supercoiled polyoma DNA. Eur. J. Biochem. 43: 591-600.

Giaever, G.N. and Wang, J.C. (1988). Supercoiling of intracellular DNA can occur in eukaryotic cells. Cell 55: 849-856.

Gillies, S.D., Folsom, V., and Tonegawa, S. (1984). Cell type-specific enhancer element associated with a mouse MHC gene, EB. Nature 310: 594-597.

Gilmour, D.S., Thomas, G.H., and Elgin, S.C.R. (1989). *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. Science 245: 1487-1490.

Glaser, R.L., Thomas, G.H., Siegfried, E.S., Elgin, S.C.R., and Lis, J.T. (1990). Optimal heat-induced expression of the *Drosophila hsp26* gene requires a promoter sequence containing $(CT)_n \bullet (GA)_n$ repeats. J. Mol. Biol. 211: 751-761.

Glasser, R., and Gabbay, E.J. (1968). Topography of nucleic acid helices in solutions. III. Interactions of spermine and spermidine derivatives with polyadenlyic-polyuridylic and polyinosinic-polycytidylic acid helices. Biopolymers 6: 243-254.

Glikin, G.C., Gargiulo, G., Rena-Descalzi, L, and Worcel, A. (1983). *Escherichia coli* single-strand binding protein stabilizes specific denatured sites in superhelical DNA. Nature 303: 770-774.

Goding, C.R. and Russell, W.C. (1983). S1 sensitive sites in adenovirus DNA. Nucleic Acids Res. 11: 21-36.

Goller, M., Funke, B., Hehe-Becker, C., Kroger, B., Lottspeich, F., and Horak, I. (1994). Murine protein which binds preferentially to oligo C-rich single-stranded nucleic acids. Nucleic Acids Res. 22: 1885-1889.

Goodall, G.J. and Filipowicz, W. (1989). The AU rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell 58: 473-483.

Gosule, L.C. and Schellman, J.A. (1976). Compact form of DNA induced by spermidine. Nature 259: 333-335.

Gray, D.M., Ratliff, R.L., Antao, V.P., and Gray, C.W. (1987). CD spectroscopy of acid induced structures of polydeoxyribonucleotides: importance of $C \bullet C^+$ base pairs. In (Sarma, R.H., and Sarma, M.H. Eds.) Structure and Expression, Vol 2: DNA and its Drug Complexes. Adenine Press, New York, pp. 147-166.

Grigoriev, M., Praseuth, D., Robin, P., Hemar, A., Saison-Behmoaras, T., Dautry-Varsat, A., Thuong, N.T., Hélène, C., and Harel-Bellan (1992). A triple helix-forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor via inhibition of NFkB binding to interleukin-2 receptor alpha-regulatory sequence. J. Biol. Chem. 267: 3389-3395.

Gromova, I.I., Nielsen, O.F., and Razin, S.V. (1995). Long-range fragmentaion of the eukaryotic genome by exogenous and endogenous nucleases proceeds in a specific fashion via preferential DNA cleavage at matrix attachment sites. J. Biol. Chem. 270: 18685-18690.

Gross, D.S. and Garrard, W.T. (1988). Nuclease hypersensitive sites in chromatin. Annu. Rev. Biochem. 57: 159-197.

Guarente, L. (1988). UACs and enhancers: common mechanism of transcriptional activation in yeast and mammals. Cell 52: 303-305.

Guieysse, A.L., Praseuth, D., and Helene, C. (1997). Identification of a triplex DNA-binding protein from human cells. Annu. Rev. Biochem. 267: 289-298.

Gutkind, J.S., Link, D.C., Katamine, S., Lacal, P., Miki, T., Ley, T.J., and Robbins, K.C. (1991). A novel c-*fgr* exon utilized in Epstein-Barr virus-infected B lymphocytes but not in normal monocytes. Mol. Cell. Biol. 3: 1500-1507.

Hampel, K.J. (1994). Triplex DNA in plasmids and chromosomes. Ph.D. thesis. Department of Biochemistry, University of Saskatchewan, Saskatoon, Sask., Canada.

Hampel, K.J. and Lee, J.S. (1993). Two-dimensional pulsed-field electrophoresis of yeast chromosomes: evidence for triplex-mediated DNA condensation. Biochem. Cell. Biol. 71: 190-196.

Hampel, K.J., Burkholder, G.D., and Lee, J.S. (1993). Plasmid dimerization mediated by triplex formation between polypyrimidine-polypurine repeats. Biochemistry 32: 1072-1077.

Hampel, K.J., Crosson, P. and Lee, J.S. (1991). Polyamines favor DNA triplex formation at neutral pH. Biochemistry 30: 4455-4459.

Hanvey, J.C., Klysik, J., and Wells, R.D. (1988a). Influence of DNA sequence on the formation of non-B right-handed helices in oligopurine•oligopyrimidine inserts in plasmids. J. Biol. Chem. 263: 7386-7396.

Hardin, C.C., Corregan, M., Brown II, B.A., and Frederick, L.N. (1993). Cytosine-cytosine$^+$ base pairing stabilizes DNA quadruplexes and cytosine methylation greatly enhances the effect. Biochemistry 32: 5870-5880.

Harvey, S.C., Luo, J., and Lavery, R. (1988). DNA stem-loop structures in oligopurine-oligopyrimidine triplexes. Nucleic Acids Res. 16: 11795-11809.

He, Y., Suzuki, T., Kashiwagi, K., Kusama-Eguchi, K., Shirahata, A., and Igarashi, K. (1994). Correlation between the inhibition of cell growth by bis(ethyl)polyamine analogues and the decrease in the function of the mitochondria. Eur. J. Biochem. 221: 391-398.

Helene, C. (1991). The anti-gene strategy: control of gene expression by triplex-forming oligonucleotide. Anticancer Drug Des. 6: 569-584.

Hentschel, C.C. (1982). Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. Nature 295: 714-716.

Hill, R.J., and Stollar, D.B. (1983). Dependence of Z-DNA antibody binding to polytene chromosomes on acid fixation and DNA torsional strain. Nature 305: 338-340.

Hilwig, I and Gropp, A (1972). Staining of constitutive heterochromatin in mammalian chromosomes with a new fluorochrome. Exp. Cell Res. 75: 122-126.

Hinkle, D.C., Ring, J., and Chamberlin, M.J. (1972). Studies of the binding of *Escherichia coli* RNA polymerase to DNA III. Tight binding of RNA polymerase holoenzyme to single-stranded breaks in T7 DNA. J. Mol. Biol. 70: 197-207.

Hoffman, E.K., Trusko, S.P., Murphy, M., and George, D.L. (1990). An S1 nuclease-sensitive homopurine/homopyrimidine domain in the c-Ki-*ras* promoter interacts with a nuclear factor. Proc. Natl. Acad. Sci. U.S.A. 87: 2705-2709.

Hoffman-Liebermann, B., Liebermann, D., Troutt, A., Kedes, L.H., and Cohen, S.N. (1986). Human homologues of TU transposon sequences: polypurine/polypyrimidine sequence elements that can alter DNA conformation *in vitro* and *in vivo*. Mol. Cell. Biol. 6: 3632-3642.

Hofmann, J.F., Laroche, T., Brand, A.H., and Gasser, S.M. (1989). RAP-1 factor is necessary for DNA loop formation *in vitro* at the silent mating type locus *HML*. Cell 57: 725-737.

Hollingsworth, M.A., Closken, C., Harris, A., McDonald, C.D., Pahwa, G.S., and Maher, L.J., III. (1994). A nuclear factor that binds purine-rich, single-stranded oligonucleotides derived from S1-sensitive elements upstream of the CFTR gene and the MUC1 gene. Nucleic Acids Res. 22: 1138-1146.

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

Hoogsteen, K. (1963). The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. Acta Crystallogr. 16: 907-916.

Horwitz, E.M., Maloney, K.A., and Ley, T.J. (1994). A human protein containing a "cold shock" domain binds specifically to H-DNA upstream from the human γ-globin genes. J. Biol. Chem. 269: 14130-14139.

Howard, F.B., Frazier, J., Lipsett, M.N., and Miles, H.T. (1964). Infrared demonstration of two- and three-strand helix formation between poly

C and guanosine mononucleotides and oligonucleotides. Biochem. Biophys. Res. Commun. 17: 93-102.

Htun, H., and Dalhberg, J. E. (1988). Single strands, triple strands, and kinks in H-DNA. Science 241: 1791-1796.

Htun, H., and Dalhberg, J.E. (1989). Topology and formation of triple-stranded H-DNA. Science 243: 1571-1576.

Igarashi, K., Koga, K., He, Y., Shimogori, T., Ekimoto, H., Kashiwagi, K., and Shirahata, A. (1995). Inhibition of the growth of various human and mouse tumour cells by 1,15-bis(ethylamine)-4,8,12-triazapentadecane. Cancer Res. 55: 2615-2619.

Inman, R. B. (1964a). Transitions of DNA homopolymers. J. Mol. Biol. 9: 624-637.

Inman, R. B. (1964b). Multistranded DNA homopolymer interactions. J. Mol. Biol. 10: 137-146.

Ito, K., Sato, K., and Endo, H. (1994). Cloning and characterization of a single-stranded DNA binding protein that specifically recognizes deoxycytidine stretch. Nucleic Acids Res. 22: 53-58.

Jackson, D.A. and Cook, P.R. (1985a). A general method for preparing chromatin containing intact DNA. EMBO J. 4: 913-918.

Jackson, D.A. and Cook, P.R. (1985b). Transcription occurs at a nucleoskeleton. EMBO J. 4: 919-925.

Jackson, D.A., Yuan, J, and Cook, P.R. (1988). A gentle method for preparing cyto- and nucleo-skeletons and associated chromatin. J. Cell Sci. 90: 365-378.

Jane, S.M., Gumucio, D.L., Ney, P.A., Cunningham, J.M., and Nienhuis, A.W. (1993). Methylation-enhanced binding of Sp1 to the stage selector element of the human γ-globin gene promoter may regulate developmental specifity of expression. Mol. Cell. Biol. 13: 3272-3281.

Jaworski, A., Blaho, J.A., Larson, J.E., Shimizu, M., and Wells, R.D. (1989). Tetracycline promoter mutations decrease non-B DNA structural transitions, negative linking differences and deletions in recombinant plasmids in *Eschericlia coli*. J. Mol. Biol. 207: 513-526.

Johnson, C.A., Jinno, Y., and Merlino, G.T. (1988). Modulation of epidermal growth factor receptor proto-oncogene transcription by a promoter site sensitive to S1 nuclease. Mol. Cell. Biol. 8: 4174-4184.

Johnson, D., and Morgan, A.R. (1978). Unique structures formed by pyrimidine•purine DNAs which may be four-stranded. Proc. Natl. Acad. Sci. U.S.A. 75: 1637-1641.

Johnston, B.H. (1988). The S1-sensitive form of $d(C\text{-}T)_n \bullet d(A\text{-}G)_n$: Chemical evidence for a three-stranded structure in plasmids. Science 241: 1800-1804.

Karlovsky, P., Pecinka, P., Vojtiskova, M., Makaturova, E., and Palecek, E. (1990). Protonated triplex DNA in *E. coli* cells as detected by chemical probing. FEBS Lett. 274: 39-42.

Kas, E., Poljak, L., Adachi, Y., and Laemmli, U.K. (1993). A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. EMBO J. 12: 115-126.

Kato, M. and Shimizu, N. (1992). Effect of potential triplex DNA region on the *in vitro* expression of bacterial β-lactamase gene in superhelical recombinant plasmids. J. Biochem. 112: 492-494.

Kennedy, G.C. and Rutter, W.J. (1992). Pur-1, a zinc-finger protein that binds to purine-rich sequences, transactivates an insulin promoter in heterologous cells. Proc. Natl. Acad. Sci. U.S.A. 89: 11498-11502.

Kinniburgh, A.J. (1989). A *cis*-acting transcription element of the c-*myc* gene can assume an H-DNA conformation. Nucleic Acids Res. 17: 7771-7778.

Kinniburgh, A.J., Firulli, A.B., and Kolluri, R. (1994). DNA triplexes and regulation of the c-*myc* gene. Gene 149: 93-100.

Kiyama, R. and Camerini-Otero, R.D. (1991). A triplex DNA-binding protein from human cells: purification, and characterization. Proc. Natl. Acad. Sci. U.S.A. 88: 10450-10454.

Klysik, J. (1996). Both an altered DNA structure and cellular proteins are involved in protecting a triplex forming an oligopurine-rich sequence from Dam methylation in *E. coli*. Biochem. Genetics 34: 165-178.

Klysik, J. and Shimizu, M. (1993). *Escherichia coli* single-stranded DNA-binding protein alters the structure of intramolecular triplexes in plasmids. FEBS Lett. 333: 261-267.

Kohwi, Y. (1989). Cationic metal-specific structures adopted by the poly (dG) region and the direct repeats in the chicken adult βA globin gene promoter. Nucleic Acids Res. 17: 4493-4502.

Kohwi, Y. and Kohwi-Shigematsu, T. (1991). Altered gene expression correlates with DNA structure. Genes Dev. 5: 2547-2554.

Kohwi, Y. and Kohwi-Shigematsu, T. (1993). Structural polymorphism of homopurine-homopyrimidine sequences at neutral pH. J. Mol. Biol. 231: 1090-1101.

Kohwi, Y., and Kohwi-Shigematsu, T. (1988). Magnesium ion-dependent triple-helix structure formed by homopurine-homopyrimidine sequences in supercoiled plasmid DNA. Proc. Natl. Acad. Sci. U.S.A. 85: 3781-3785.

Kohwi, Y., Malkhosyan, S.R., and Kohwi-Shigematsu, T. (1992). Intramolecular dG•dG•dC triplex detected in *Escherichia coli* cells. J. Mol. Biol. 223: 817-822.

Kolluri, R., Torrey, T.A., and Kinniburgh, A.J. (1992). A CT promoter element binding protein: definition of a double-strand and a novel single strand binding motif. Nucleic Acids Res. 20: 111-116.

Kovacs, A., Kandala, J.C., Weber, K.T., and Guntaka, R.V. (1996). Triple helix-forming oligonucleotide corresponding to the polypyrimidine sequence in the rat α1(I) collagen promoter specifically inhibits factor binding and transcription. J. Biol. Chem. 271: 1805-1812.

Krakauer, H., and Sturtevant, J.M. (1968). Heats of the helix-coil transitions of the poly A-poly U complexes. Biopolymers 6: 491-512.

Krauss, G., Sinderman, H., Schomburg, U., and Maass, G. (1981). *Escherichia coli* single-strand deoxyribonucleic acid binding protein: stability, specificity, and kinetics of complexes with oligonucleotides, and deoxyribonucleic acid. Biochemistry 20: 5346-5352.

Krumm, A., Meulia, T., Brunvand, M., and Groudine, M. (1992). The block to transcriptional elongation within the human c-*myc* gene is

determined in the promoter-proximal region. Genes Dev. 6: 2201-2213.

Laemmli, U.K., Cheng, S.M., Adolph, K.W., Paulson, J.R., Brown, J.A., and Baumbach, W.R. (1977). Metaphase chromosome structure: the role of nonhistone proteins. Cold Spring Harbor Symp. Quant. Biol. 42: 351-360.

Laemmli, U.K., Käs, E., Poljak, L., and Adachi, Y. (1992). Scaffold associated regions: cis-acting determinants of chromatin structural loops and functional domains. Curr. Opin. Genet. Dev. 2: 275-285.

Lafyatis, R., Denhez, F., Williams, T., Sporn, M., and Roberts, A. (1991). Sequence specific protein binding to and activation of the TGF-beta 3 promoter through a repeated TCCC motif. Nucleic Acids Res. 19: 6419-6425.

Langowski, J., Benight, A.S., Fujimoto, B.S., Schurr, J.M., and Schomburg, U. (1985). Change of conformation, and internal dynamics of supercoiled DNA upon binding of *Esclierichia coli* single-strand binding protein. Biochemistry 24: 4022-4028.

Lapidot, A., Baran, N., and Manor, H. (1989). $(dT\text{-}dC)_n$ and $(dG\text{-}dA)_n$ tracts arrest single stranded DNA replication *in vitro*. Nucleic Acids Res. 17: 883-900.

Larsen, A., and Weintraub, H. (1982). An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chicken globin gene chromatin. Cell 29: 609-622.

Latimer, L.J.P., Hampel, K., and Lee, J.S. (1989). Synthetic repeating sequence DNA containing phosphorothioates: nuclease sensitivity and triplex formation. Nucleic Acids Res. 17: 1549-1561.

Laughton, C.A., and Neidle, S. (1992). Molecular dynamics simulation of the DNA triplex $d(TC)_5 \bullet d(GA)_5 \bullet d(C^+T)_5$. J. Mol. Biol. 223: 519-529.

Lee, J.S. and Morgan, A.R. (1982). Novel aspects of the structure of the *Esclierichia coli* nucleoid investigated by a rapid sedimentation assay. Can J. Biochem. 60: 952-961.

Lee, J.S., Burkholder, G.D., Latimer, L.J.P., Haug, B.L., and Braun, R.P. (1987). A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes. Nucleic Acids Res. 15: 1047-1061.

Lee, J.S., Johnson, D.A., and Morgan, R.A. (1979). Complexes formed by (pyrimidine)$_n$•(purine)$_n$ DNAs on lowering the pH are three-stranded. Nucleic Acids Res. 6: 3073-3091.

Lee, J.S., Latimer, L.J.P., and Sibley, J.T. (1989). Detection of unusual nucleic acids with monoclonal antibodies specific for triplexes, poly (ADP-ribose), DNA-RNA hybrids, and Z-DNA. Anal. Biochem. 178: 373-377.

Lee, J.S., Woodsworth, M.L., Latimer, L.J.P., and Morgan, A.R. (1984). Poly(pyrimidine)•poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. Nucleic Acids Res. 12: 6603-6614.

Lehninger, A.L., Nelson, D.L., and Cox, M.M. (1993). Principles of Biochemistry, Second Ed., Worth Publishers, New York.

Lindsay, S. and Bird, A.P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. Nature 327: 336-338.

Lipsett, M.N. (1963). The interactions of poly C, and guanine trinucleotide. Biochem. Biophys. Res. Commun. 11: 224-231.

Lipsett, M.N. (1964). Complex formation between polycytidylic acid and guanine oligonucleotides. J. Biol. Chem. 239: 1256-1260.

Liu, L.F., and Wang, J.C. (1987). Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. U.S.A. 84: 7024-7027.

Lu, Q., Wallrath, L.L., Granok, H., and Elgin, S.C.R. (1993). (CT)$_n$•(GA)$_n$ repeats, and heat shock elements have distinct roles in chromatin structure, and transcriptional activation of the *Drosophila hsp*26 gene. Mol. Cell. Biol. 13: 2802-2814.

Lurz, R., Grote, M., Dijk, J., Reinhardt, R., and Dobrinski, B. (1986). Electron microscopic study of DNA complexes with proteins from the Archaebacterium *Sulfolobus acidocaldarius*. EMBO J. 5: 3715-3721.

Lyamichev, V.I., Mirkin, S.M., and Frank-Kamenetskii, M.D. (1985). A pH dependent structural transition in the homopurine-homopyrimidine tract in superhelical DNA. J. Biomol. Struct. Dyn. 3: 327-338.

Lyamichev, V.I., Mirkin, S.M., and Frank-Kamenetskii, M.D. (1986). Structures of homopurine-homopyrimidine tract in superhelical DNA. J. Biomol. Struct. Dyn. 4: 667-669.

Lyamichev, V.I., Mirkin, S.M., Kumarev, V.P., Baranova, L.V., Vologodskii, A.V., and Frank-Kamenetskii, M.D. (1989). Energetics of B-H transition in supercoiled DNA carrying $d(CT)_x \bullet d(AG)_x$ and $d(C)_n \bullet d(G)_n$ inserts. Nucleic Acids Res. 17: 9417-9423.

Lyamichev, V.I., Voloshin, O.N., Frank-Kamenetskii, M.D., and Soyfer, V.N. (1991). Photofootprinting of DNA triplexes. Nucleic Acids Res. 19: 1633-1638.

Mach, M., Ebert, P., Popp, R., and Ogilvie, A. (1982). Compartmentalization of polyamines in mammalian cells. Biochem. Biophys. Res. Comm. 104: 1327-1334.

Maher, L.J. (1996). Prospects for the therapeutic use of antigene oligonucleotides. Cancer Invest. 14: 66-82.

Maher, L.J., Dervan, P.B., and Wold, B.J. (1992). Analysis of promoter-specific repression by triple-helical DNA complexes in a eukaryotic cell-free transcription system. Biochemistry 31: 70-81.

Maher, L.J., III, Dervan, P.B., and Wold, B. (1990). Kinetic analysis of oligodeoxyribonucleotide-directed triplex-helix formation on DNA. Biochemistry 29: 8820-8826.

Malkov, V.A., Voloshin, O.N., Veselkov, A.G., Rostapshov, V.M., Jansen, I., Soyfer, V.N., and Frank-Kamenetskii, M.D. (1993). Protonated pyrimidine-purine-purine triplex. Nucleic Acids Res. 21: 105-111.

Mangel, W.F. and Chamberlin, M.J. (1974). Studies of ribonucleic acid chain initiation by *Escherichia coli* ribonucleic acid polymerase bound to T7 deoxyribonucleic acid. Assay for the rate, and extent of ribonucleic acid chain initiation. J. Biol. Chem. 249: 3007-3013.

Manor, R., Rao, B.S., and Martin, R G. (1988). Abundance and degree of dispersion of genomic $d(GA)_n \bullet d(TC)_n$ sequences. J. Mol. Evol. 27: 96-101.

Margot, J.B. and Hardison, R.C. (1985). DNase I and nuclease S1 sensitivity of the rabbit B1 globin gene in nuclei and in supercoiled plasmids. J. Mol. Biol. 184: 195-210.

Martens, P.A. and Clayton, D.A. (1977). Strand breakage in solutions of DNA and ethidium bromide exposed to light. Nucleic Acids Res. 4: 1393-1407.

Marzluff, W.F. (1978). Transcription of RNA in isolated nuclei. Methods Cell Biol. 19: 317-331.

Masuda, A., Oyamada, M., Nagaoka, T., Tateishi, N., and Takamatsu, T. (1998). Regulation of cytosol-nucleus pH gradients by K+/H+ exchange mechanism in the nuclear envelope of neonatal rat astrocytes. Brain Res. 807: 70-77.

Mavrothalassitis, G.J., Watson, D.K., and Papas, T.S. (1990). Molecular, and functional characterization of promoter of ETS2, the human c-*ets*-2 gene. Proc. Natl. Acad. Sci. U.S.A. 87: 1047-1051.

Mayfield, C., Ebbinghaus, S., Gee, J., Jones, D., Rodu, B., Squibb, M., and Miller D. (1994). Triplex formation by the human Ha-*ras* promoter inhibits Sp1 binding and *in vitro* transcription. J. Biol. Chem. 269: 18232-18238.

McKeon, C., Schmidt, A., and de Crombrugghe, B. (1984). A sequence conserved in both the chicken and mouse alpha 2(I) collagen promoter contains sites sensitive to S1 nuclease. J. Biol. Chem. 259: 6636-6640.

Mennes, A.M., Quint, A., Gribnau, J.H., Boot, C.J.M., van der Zaal, E.J., Maan, A.C., and Libbenga, K.R. (1992). Specific transcription and reinitiation of 2,4-D-induced genes in tobacco nuclei. Plant Mol. Biol. 18: 109-117.

Michelotti, E.F., Tomonaga, T., Krutzsch, H., and Levens, D. (1995). Cellular nucleic acid binding protein regulates the CT element of the human c-*myc* protooncogene. J. Biol. Chem. 270: 9494-9499.

Michelson, A.M., Massoulie, J., and Guschlbauer, W. (1967). Synthetic polynucleotides. Progr. Nucleic Acids Res. Mol. Biol. 6: 83-141.

Miles, H.T. (1964). The structure of the three-stranded helix, poly (A+2U). Proc. Natl. Acad. Sci. U.S.A. 51: 1104-1109.

Mirkin, S.M. and Frank-Kamenetskii, M.D. (1994). H-DNA and related structures. Annu. Rev. Biophys. Biomol. Struct. 23: 541-576.

Mirkin, S.M., Lyamichev, V.I., Drushlyak, K.N., Dobrynin, V.N., Filippov, S.A., and Frank-Kamenetskii, M.D. (1987). DNA H form requires a homopurine-homopyrimidine mirror repeat. Nature 330: 495-497.

Mol, C.D., Muir, A.K.S., Cygler, M., Lee, J.S., and Anderson, W.F. (1994). Structure of an immunoglobulin Fab fragment specific for triple-stranded DNA. J. Biol. Chem. 269: 3615-3622.

Mollegaard, N.E., Buchardt, O., Egholm, M., and Nielsen, P.E. (1994). Peptide nucleic acid•DNA strand displacement loops as artificial transcription promoters. Proc. Natl. Acad. Sci. U.S.A. 91: 3892-3895.

Mooren, M.M.W., Pulleyblank, D.E., Wijmenga, S.S., Blommers, M.J.J., and Hilbers, C.W. (1990). Polypurine/polypyrimidine hairpins form a triple helix structure at low pH. Nucleic Acids Res. 18: 6523-6529.

Morgan, A.R., and 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.

Morgan, A.R., Lee, J.S., Pulleyblank, D.E., Murray, N.L., and Evans, D.H. (1979). Ethidium fluorescence assays. Part I. Physicochemical studies. Nucleic Acids Res. 7: 547-569.

Müller, V., Takeya, M., Brendel, S., Wittig, B., and Rich, A. (1996). Z-DNA-forming sites within the human β-globin gene cluster. Proc. Natl. Acad. Sci. U.S.A. 93: 780-784.

Muniyappa, K., Shaner, S.L., Tsang, S.S., and Radding, C.M. (1984). Mechanism of the concerted action of recA protein, and helix-destabilizing proteins in homologous recombination. Proc. Natl. Acad. Sci. U.S.A. 81: 2757-2761.

Muraiso, T., Nomoto, S., Yamazaki, H., Mishima, Y., and Kominami, R. (1992). A single-stranded DNA binding protein from mouse tumor cells specifically recognizes the C-rich strand of the (AGG:CCT)n repeats that can alter DNA conformation. Nucleic Acids Res. 20: 6631-6635.

Murray, N.L. and Morgan, A.R. (1973). Enzymic and physical studies on the triplex $dT_n \bullet dA_n \bullet Ru_n$ [(deoxyribosylthymine)$_n$•(deoxyadenosine)$_n$•(ribo uridine)$_n$]. Can. J. Biochem. 51: 436-449.

Musso, M., Nelson, L.D., and Van Dyke, M.W. (1998). Characterization of purine-motif triplex DNA-binding proteins in HeLa extracts. Biochemistry 37: 3086-3095

Musso, M., Thomas, T., Shirahata, A., Sigal, L.H., Van Dyke, M.W., and Thomas, T.J. (1997). Effects of chain length modification and bis(ethyl) substitution of spermine analogs on purine-purine-pyrimidine triplex

DNA stabilization, aggregation, and conformational transitions. Biochemistry 36: 1441-1449.

Nickol, J.M. and Felsenfeld, G. (1983). DNA conformation at the 5' end of the chicken adult beta-globin gene. Cell 35: 467-477.

Nikolov, D.B. and Burley, S.K. (1997). RNA polymerase II transcription initiation: A structural view. Proc. Natl. Acad. Sci. U.S.A. 94: 15-22.

O'Neill, D., Bornschlegel, K., Flamm, M., Castle, M., and Bank, A. (1991). A DNA-binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human delta-globin gene. Proc. Natl. Acad. Sci. U.S.A. 88: 8953-8957.

Ohlenbusch, H.H., Olivera, B.M., Tuan, D. and Davidson, N. (1967). Selective dissociation of histones from calf thymus nucleoprotein. J. Mol. Biol. 25: 299-315.

Olivas, W.M. and Maher, L.J., III (1995). Competitive triplex/quadruplex equilibria involving guanine-rich oligonucleotides. Biochemistry 34: 278-284.

Opstelten, R.J.G., Clement, J.M.E., and Wanka, F. (1989). Direct repeats at nuclear matrix-associated DNA regions and their putative control function in the replicating eukaryotic genome. Chromosoma 98: 422-427.

Ozoline, O.N. and Tsyganov, M.A. (1995). Structure of open promoter complexes with *Escherichia coli* RNA polymerase as revealed by the DNase I footprinting technique: compilation analysis. Nucleic Acids Res. 23: 4533-4541.

Pack, G.R., Wong, Ll., and Lamm, G. (1998). pKa of cytosine on the third strand of triplex DNA: preliminary Poisson-Boltzmann calculations. Internatl. J. Quantum. Chem. 70: 1177-1184.

Palecek, E. (1991). Local supercoil-stabilized DNA structures. Crit. Rev. Biochem. Mol. Biol. 26: 151-226.

Panyutin, I.G. and Wells, R.D. (1992). Nodule DNA in the $(GA)_{37}\bullet(CT)_{37}$ insert in superhelical plasmids. J. Biol. Chem. 267: 5495-5501.

Parniewski, P., Galazka, G., Wilk, A., and Klysik, J. (1989). Complex structural behavior of oligopurine-oligopyrimidine sequence cloned within the supercoiled plasmid. Nucleic Acids Res. 17: 617-629.

Parniewski, P., Kwinkowski, M., Wilk, A., and Klysik, J. (1990). Dam methylase sites located within the loop region of the oligopurine-oligpyrimidine sequences capable of forming H-DNA are undermethylated *in vivo*. Nucleic Acids Res. 18: 605-611.

Parvin, J.D. and Sharp, P.A. (1993). DNA topology, and a minimal set of basal factors for transcription by RNA polymerase II. Cell 73: 533-540.

Paulson, J.R., and Laemmli, U.K. (1977). The structure of histone-depleted metaphase chromosomes. Cell 12: 817-828.

Pestov, D.G., Dayn, A., Siyanova, E.Y., George, D.L., and Mirkin, S.M. (1991). H-DNA, and Z-DNA in the mouse c-Ki-*ras* promoter. Nucleic Acids Res. 19: 6527-6532.

Pettijohn, D.E. and Hecht, R. (1973). RNA molecules bound to the folded bacterial genome stabilize DNA folds, and segregate domains of supercoiling. Cold Spring Harbor Symp. Quant. Biol. 38: 31-41.

Pettijohn, D.E. and Pfenninger, O. (1980). Supercoils in prokaryotic DNA restrained *in vivo*. Proc. Natl. Acad. Sci. U.S.A. 77: 1331-1335.

Pharmacia catalogue (1995). Pharmacia Biotech Inc., Baie d'Urfé, PQ, Canada

Pilch, D.S. and Breslauer, K.J. (1994). Ligand-induced formation of nucleic acid triple helices. Proc. Natl. Acad. Sci. U.S.A. 91: 9332-9336.

Pilch, D.S., Levenson, C., and Shafer, R.H. (1990). Structural analysis of the $(dA)_{10}$•$2(dT)_{10}$ triple helix. Proc. Natl. Acad. Sci. U.S.A. 87: 1942-1946.

Pilch, D.S., Levenson, C., and Shafer, R.H. (1991). Structure, stability, and thermodynamics of a short intermolecular purine-purine-pyrimidine triple helix. Biochemistry 30: 6081-6087.

Porter, C.W. and Sufrin, J.R. (1986). Interference with polyamine biosynthesis and (or) function by analogs of polyamines and methionine as a potential anticancer chemotherapeutic strategy. Anticancer Res. 6: 525-542.

Postel, E.H., Flint, S.J., Kessler, D.J., and Hogan, M.E. (1991). Evidence that a triplex-forming oligodeoxyribonucleotide binds to the c-*myc* promoter in HeLa cells, therby reducing c-*myc* mRNA levels. Proc. Natl. Acad. Sci. U.S.A. 88: 8227-8231.

Postel, E.H., Mango, S.E., and Flint, S.J. (1989). A nuclease-hypersensitive element of the human c-*myc* promoter interacts with a transcription initiation factor. Mol. Cell. Biol. 9: 5123-5133.

Pugh, B.F. and Tjian, R. (1991). Transcription from TATA-less promoter requires a multisubunit TFIID complex. Genes Dev. 5: 1935-1945.

Puhl, H.L. and Behe, M.J. (1995). Poly (dA)•poly (dT) forms very stable nucleosomes at higher temperatures. J. Mol. Biol. 245: 559-567.

Pulleyblank, D.E., Haniford, D.B., and Morgan, A.R. (1985). A structural basis for S1 nuclease sensitivity of double-stranded DNA. Cell 42: 271-280.

Radding, C.M. (1991). Helical interactions in homologous pairing, and strand exchange driven by RecA protein. J. Biol. Chem. 266: 5355-5358.

Radhakrishnan, I., de los Santos, C. and Patel, D.J. (1991). Nuclear magnetic resonance structural studies of intramolecular purine•purine•pyrimidine DNA triplexes in solution - base triple pairing alignments, and strand direction. J. Mol. Biol. 221: 1403-1418.

Radhakrishnan, I., de los Santos, C., and Patel, D.J. (1993). Nuclear magnetic resonance structural studies of A•AT base triple alignments in intramolecular purine•purine•pyrimidine DNA triplexes in solution. J. Mol. Biol. 234: 188-197.

Raghavan, S., Burma, P.K., and Brahmachari, S.K. (1997). Positional preferences of polypurine/polypyrimidine tracts in *Saccharomyces cerevisiae* genome: implications for *cis* regulation of gene expression. J. Mol. Evol. 45: 485-498.

Raghu, G., Tevosian, S., Anant, S., Subramanian, K.N., George, D.L., and Mirkin, S.M. (1994). Transcriptional activity of the homopurine-homopyrimidine repeat of the c-Ki-*ras* promoter is independent of its H-forming potential. Nucleic Acids Res. 22: 3271-3279.

Rajagopal, P. and Feigon, J. (1989a). NMR studies of triple-strand formation from the homopurine-homopyrimidine deoxyribonucleotides $d(GA)_4$ and $d(TC)_4$. Biochemistry 28: 7859-7870.

Rajagopal, P., and Feigon, J. (1989b). Triple-stranded formation in the homopurine•homopyrimidine DNA oligonucleotides $d(G-A)_4$ and $d(C-T)_4$. Nature 339: 637-640.

Ralph R.K., Conners, W.J., and Khorana, H.G. (1962). Secondary structure and aggregation in deoxyguanosine oligonucleotides. J. Am. Chem. Soc. 84: 2265-2266.

Rao, B.S. (1994). Pausing of simian virus 40 DNA replication fork movement *in vivo* by $(dG\text{-}DA)_n \bullet (dT\text{-}dC)_n$ tracts. Gene 140: 233-237.

Ray, D. and Robert-Lezenes, J. (1989). Coexistence of a c-*myc* mRNA initiated in intron 1 with the normal c-*myc* mRNA and similar regulation of both transcripts in mammalian cells. Oncogene Res. 5: 73-78.

Reaban, M.E. and Griffin, J.A. (1990). Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. Nature 348: 342-344. (Erratum Nature 351: 447-448).

Reaban, M.E., Lebowitz, J., and Griffin J.A. (1994). Transcription induces the formation of a stable RNA•DNA hybrid in the immunoglobulin α switch region. J. Biol. Chem. 269: 21850-21857.

Rhodes, D. and Klug, A. (1980). Helical periodicity of DNA determined by enzyme digestion. Nature 286: 573-578.

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., and 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.

Ritchie, S. and Bonham, K. (1998). The human c-*src* proto-oncogene promoter contains multiple agents for triplex-forming oligonucleotides. Antisense Nucleic Acid Drug Dev. 8: 391-400.

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

Roitt, I., Brostoff, J., and Male, D. (1998). Immunology. 5th ed., Mosby International Ltd., London.

Ruiz-Carillo, A. (1984). The histone H5 gene is flanked by S1 hypersensitive structures. Nucleic Acids Res. 12: 6473-6492.

Ruskin, B. and Green, M.R. (1985). The role of the 3' splice site concensus sequence in mammalian pre-mRNA splicing. Nature 317: 732-734.

Ryan, K.M. and Birnie, G.D. (1996). *Myc* oncogenes: the enigmatic family. Biochem. J. 314 (Pt 3): 713-721.

Sakatsume O., Tsutsui, H., Wang, Y., Gao, H., Tang, X., Yamauchi, T., Murata, T., Itakura, K., and Yokoyama K.K. (1996). Binding of THZif-1, a MAZ-like zinc finger protein to the nuclease-hypersensitive element in the promoter region of the *c-myc* protooncogene. J. Biol. Chem. 271: 31322-31333.

Samadashwily, G.M., Dayn, A., and Mirkin, S.M. (1993). Suicidal nucleotide sequences for DNA polymerization. EMBO J. 12: 4975-4983.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory, New York.

Sander, M., and Hsieh, T.-S. (1985). *Drosophila* topoisomerase II double-stranded DNA cleavage: analysis of DNA sequence homology at the cleavage site. Nucleic Acids Res. 13: 1057-1072.

Santra, M., Danielson, K.G., and Iozzo, R.V. (1994). Structural, and functional characterization of the human decorin gene promoter. A homopurine-homopyrimidine S1 nuclease-sensitive region is involved in transcriptional control. J. Biol. Chem. 269: 579-587.

Sarhan, S. and Seiler, N. (1989). On the subcellular localization of polyamines. Biol. Chem. Hoppe-Seyler 370: 1279-1284.

Sarkar, P.S. and Brahmachari, S.K. (1992). Intramolecular triplex potential sequence within a gene down regulates its expression *in vivo*. Nucleic Acids Res. 20: 5713-5718.

Saucier, J.M. and Wang, J.C. (1972). Angular alteration of the DNA helix by *Escherichia coli* RNA polymerase. Nature 239: 167-170.

Scaria, P.V. and Shafer, R.H. (1991). Binding of ethidium bromide to a DNA triplex helix: evidence for intercalation. J. Biol. Chem. 266: 5417-5423.

Schleif, R. (1992). DNA looping. Ann. Rev. Biochem. 61: 199-223.

Schmid, N., and Behr, J.-P. (1991). Location of spermine and other polyamines on DNA as revealed by photoaffinity cleavage with polyaminobenzenediazonium salts. Biochemistry 30: 4357-4361.

Schon, E., Evans, T., Welsh, J., and Efstratiadis, A. (1983). Conformation of promoter DNA: fine mapping of S1-hypersensitive sites. Cell 35: 837-848.

Schwartzberg, P.L. (1998). The many faces of *src*: multiple functions of a prototypical tyrosine kinase. Oncogene 17 (11 Reviews): 1463-1468.

Sederoff, R. and Lowenstein, L. (1975). Polypyrimidine segments in *Drosophila melanogaster* DNA: II. Chromosome location and nucleotide sequence. Cell 5: 183-194.

Sentenac, A. (1985). Eukaryotic RNA polymerases. CRC Crit. Rev. Biochem. 18: 31-90.

Shapiro, H.S. and Chargaff, E. (1957). Characterization of nucleotide arrangement in DNA through stepwise acid degradation. Biochem. Biophys. Acta 23: 451-452.

Shen, C.-K.J. (1983). Superhelicity induces hypersensitivity of a human polypyrimidine•polypurine DNA sequence in the human α2-α1 globin intergenic region to S1 nuclease digestion - high resolution mapping of the clustered cleavage sites. Nucleic Acids Res. 11: 7899-7910.

Shilatifard, A. (1998). Factors regulating the transcriptional elongation activity of RNA polymerase II. FASEB J. 12: 1437-1446.

Shimizu, M., Hanvey, J.C., and Wells, R.D. (1989). Intramolecular DNA triplexes in supercoiled plasmids. I. Effect of loop size on the formation and stability. J. Biol. Chem. 264: 5944-5949.

Shimizu, M., Kubo, K., Matsomoto, U., and Shindo, H. (1994). The loop sequence plays crucial roles for isomerization of intramolecular triplexes in supercoiled plasmids. J. Mol. Biol. 235: 185-197.

Siebenlist, U., Bressler, P., and Kelly, K. (1988). Two distinct mechanisms of transcriptional control operate on c-*myc* during differentiation of HL60 cells. Mol. Cell. Biol. 8: 867-874.

Siegfried, E., Thomas, G.H., Bond, U.M., and Elgin, S.C.R. (1986). Characterization of a supercoil-dependent S1 sensitive site 5' to the *Drosophila melanogaster hsp*26 gene. Nucleic Acids Res. 14: 9425-9444.

Sinden, R.R. (1994). DNA structure and function. Chapter 6. Academic Press, Toronto. pp. 217-256.

Sinden, R.R. and Wells, R.D. (1992). DNA structure, mutations, and human genetic disease. Curr. Opin. Biotechnol. 3: 612-622.

Sinden, R.R., and Pettijohn, D.E. (1981). Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling. Proc. Natl. Acad. Sci. U.S.A. 78: 224-228.

Sinden, R.R., Zheng, G., Brankamp, R.G., and Allen, K.N. (1991). On the deletion of inverted repeated DNA in *E. coli*: effects of length, thermal stability, and cruciform formation *in vivo*. Genetics 129: 991-1005.

Singleton, S.F. and Dervan, P.B. (1993). Equilibrium association constants for oligonucleotide-directed triple helix formation at single DNA sites: linkage to cation valence, and concentration. Biochemistry 32: 13171-13179.

Sklenar, V., and Feigon, J. (1990). Formation of a stable triplex from a single DNA strand. Nature 345: 836-838.

Slonczewski, J.L., Rosen, B.P., Alger, J.R., and Macnab, R.M. (1981). pH homeostasis in *Escherichia coli*: measurement by ^{31}P nuclear magnetic resonance of methylphosphonate and phosphate. Proc. Natl. Acad. Sci. U.S.A. 78: 6271-6275.

Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta. 1351: 73-88.

Soyfer, V.N. and Potaman, V.N. (1996). Triple-helical nucleic acids. Springer-Verlag, New York.

Soyfer, V.N., Voloshin, O.N., Malkov, V.A., and Frank-Kamenetskii, M.D. (1992). Photofootprinting of inter, and intramolecular DNA triplexes. In (Sarma, R.H. and Sarma, M.H., Eds.). Structure and Function, Vol 1: Nucleic Acids. Adenine Press, New York, pp. 29-41.

Spencer, C.A. and Groudine, M. (1990). Transcription elongation and eukaryotic gene regulation. Oncogene 5: 777-785.

Spencer, J.H. and Chargaff, E. (1963). Studies on the nucleotide arrangement in deoxyribonucleic acids. VI. Pyrimidine nucleotide clusters: frequency and distribution in several species of the AT-type. Biochim. Biophys. Acta 68: 18-27.

Stewart, T.A., Bellve, A.R., and Leder, P. (1984). Transcription and promoter usage of the *myc* gene in normal somatic and spermatogenic cells. Science 226: 707-710.

Stollar, B.D. (1992). Immunochemical analyses of nucleic acids. In (Cohn, W.E. and Moldave, K., Eds). Progress in Nucleic Acids Research and Molecular Biology, Vol 42, Academic Press, San Diego, California, pp. 40-77.

Straus, N.A. and Birnboim, H.C. (1974). Long pyrimidine tracts in L cell DNA: Localization to repeated DNA. Proc. Natl. Acad. Sci. U.S.A. 71: 2992-2995.

Su, W., Jackson, S., Tjian, R., and Echols, H. (1991). DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. Genes Dev. 5: 820-826

Subramanian, K.N., Dhar, R., and Weissman, S.M. (1977). Nucleotide sequence of a fragment of SV40 DNA that contains the origin of DNA replication and specifies the 5' ends of "early" and "late" viral RNA. III. Construction of the total sequence of EcoRII-G fragment of SV40 DNA. J. Biol. Chem. 252: 355-367.

Summers, W.C. and Szybalski, W. (1968). Totally asymmetric transcription of coliphage T7 *in vivo*: correlation with poly G binding sites. Virology 34: 9-16.

Sures, I., Lowry, L, and Kedes, L.H. (1978). The DNA sequence of sea urchin (*S. purpuratus*) H2A, H2B and H3 histone coding and spacer regions. Cell 15: 1033-1044.

Sykes, K. and Kaufman, R. (1990). A naturally occurring gamma globin gene mutation enhances Sp1 binding activity. Mol. Cell. Biol. 10: 95-102.

Szybalski, W., Kubinski, H., and Sheldrick, P. (1966). Pyrimidine clusters on the transcribing strand of DNA, and their possible role in the initiation of RNA synthesis. Cold Spring Habor Symp. Quant. Biol. 31: 123-127.

Tabor, F. and Tabor, C. (1984). Polyamines. Annu. Rev. Biochem. 254: 749-783.

Tabor, H. (1962). The protective effect of spermine, and other polyamines against heat denaturation of deoxyribonucleic acid (DNA). Biochemistry 1: 496-501.

Takimoto, M., Tomonaga, T., Matunis, M., Avigan, M., Krutzsch, H., Dreyfuss, G., and Levens, D. (1993). Specific binding of heterogeneous ribonucleoprotein particle protein K to the human c-*myc* promoter, *in vitro*. J. Biol. Chem. 268: 18249-18255.

Tamm, C., Shapiro, H.S., Lipshitz, R., and Chargaff, E. (1953). Distribution density of nucleotides within a DNA chain. J. Biol. Chem. 203: 673-688; 689-698.

Tari, L.W. and Secco, A.S. (1995). Base-pair opening and spermine binding-B-DNA features displayed in the crystal structure of a *gal* operon fragment: implications for protein-DNA recognition. Nucleic Acids Res. 23: 2065-2073.

Tata, J.R. and Baker, B.S. (1985). Specific switching on of silent egg protein genes *in vitro* by an S-100 fraction in isolated nuclei from male *Xenopus*. EMBO J. 4: 3253-3258.

Thomas, B.J. and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. Cell 56: 619-630.

Thomas, T. and Thomas, T.J. (1993). Selectivity of polyamines in triplex DNA stabilization. Biochemistry 32: 14068-14074.

Thomas, T.J., and Bloomfield, V.A. (1983a). Chain flexibility and hydrodynamics of the B and Z forms of poly (dG-dC)•poly (dG-dC). Nucleic Acids Res. 11: 1919-1930.

Thomas, T.J., and Bloomfield, V.A. (1983b). Collapse of DNA caused by trivalent cations: pH and ionic specificity effects. Biopolymers 22: 1097-1106.

Thomas, T.J., Faaland, C.A., Gallo, M.A., and Thomas, T. (1995). Suppression of c-*myc* oncogene expresion by a polyamine-complexed triplex forming oligonucleotide in MCF-7 breast cancer cells. Nucleic Acids Res. 23: 3594-3599.

Thorburn, A. and Knowland, J. (1988). A novel nuclear transcription system which responds correctly to cloned estrogen receptor. Nucleic Acids Res. 16: 10469-10476.

Tomonaga, T. and Levens, D. (1995). Heterogeneous nuclear ribonucleoprotein K is a DNA-binding transactivator. J. Biol. Chem. 270: 4875-4881.

Trainer, D.L., Kline, T., McCabe, F.L., Faucette, L.F., Feild, J., Chaikin, M., Anzano, M., Rieman, D., Hoffstein, S., Li, D.J., et al. (1988). Biological characterization and oncogene expression in human colorectal carcinoma cell lines. Int. J. Cancer 41: 287-296.

Tripathi, J. and Brahmachari, S.K. (1991). Distribution of simple repetitive (TG/CA)n and (CT/AG)n sequences in human and rodent genomes. J. Biomol. Struct. Dyn. 9: 387-397.

Tsao, Y.P., Wu, H.Y., and Liu, L.F. (1989). Transcription-driven supercoiling of DNA: direct biochemical evidence from *in vitro* studies. Cell 56: 111-118.

Tung, C.-H., Breslauer, K.J., and Stein, S. (1993). Polyamine-linked oligonucleotides for DNA triple helix formation. Nucleic Acids Res. 21: 5489-5494.

Uemura, T., Ohkura, H., Adachi, Y., Marino, K., Shiozaki, K., and Yanagida, M. (1987). DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in *S. pombe*. Cell 50: 917-925.

Ulrich, M.J., Gray, W.J., and Ley, T.J. (1992). An intramolecular DNA triplex is disrupted by point mutations associated with hereditary persistence of fetal hemoglobin. J. Biol. Chem. 267: 18649-18658.

Ussery, D.W., and Sinden, R.R. (1993). Environmental influences on the *in vivo* level of intramolecular triplex DNA in *Escherichia coli*. Biochemistry 32: 6206-6213.

van Holde, K. and Zlatanova, J. (1994). Unusual DNA structures, chromatin and transcription. BioEssays 16: 59-68.

Vennstrom, B., Sheiness, D., Zabielski, J., and Bishop, J.M. (1982). Isolation and characterization of c-*myc*, a cellular homolog of the oncogene (v-*myc*) of avian myelocytomatosis virus strain 29. J. Virol. 42: 773-779.

Veselkov, A.G., Malkov, V.A., Frank-Kamenetskii, M.D., and Dobrynin, V.N. (1993). Triplex model of chromosome ends. Nature 364: 496.

Vinograd, J., Lebowitz, J., Radloff, R., Watson, R., and Laipis, P. (1965). The twisted circular form of polyoma viral DNA. Proc. Natl. Acad. Sci. U.S.A. 53: 1104-1111.

Vogt, P. (1990). Potential genetic functions of tandem repeated DNA sequence blocks in the human genome are based on a highly conserved "chromatin folding code". Hum. Genet. 84: 301-336.

Vogt, P. (1992). Code domains in tandem repetitive DNA sequence structures. Chromosoma 101: 585-589.

Vogt, V. (1969). Breaks in DNA stimulate transcription by core RNA polymerase. Nature 223: 854-855.

Vogt, V.M. (1973). Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae*. Eur. J. Biochem. 33: 192-200.

Vojtiskova, M. and Palecek, E. (1987). Unusual protonated structure in the homopurine•homopyrimidine tract of supercoiled and linearized plasmids recognized by chemical probes. J. Biomol. Struct. Dyn. 5: 283-296.

Vojtiskova, M., Mirkin, S., Lyamichev, V., Voloshin, O., Frank-Kamenetskii, M., and Palecek, E. (1988). Chemical probing of the homopurine•homopyrimidine tracts in supercoiled DNA at single-nucleotide resolution. FEBS Lett. 234: 295-299.

Volker, J. and Klump, H.H. (1994). Electrostatic effects in DNA triple helices. Biochemistry 33: 13502-13508.

Voloshin, O.N., Mirkin, S.M., Lyamichev, V.I., Belotserkovskii, B.P., and Frank-Kamenetskii, M.D. (1988). Chemical probing of homopurine-homopyrimidine miror repeats in supercoiled DNA. Nature 333: 475-476.

Wagner, R.W. (1995). The state of the art in antisense research. Nature Med. 1: 1116-1118.

Wang, J.C. and Lynch, A.S. (1993). Transcription and DNA supercoiling. Curr. Opin. Genet. Dev. 3: 764-768.

Weber, J., Jelinek, W., and Darnell, J.E. (1977). The definition of a large viral trancription unit late in Ad2 infection of HeLa cells: mapping of nascent RNA molecules labeled in isolated nuclei. Cell 10: 611-616.

Weinreb, A., Collier, D.A., Birshtein, B.K., and Wells, R.D. (1990). Left-handed Z-DNA and intramolecular triplex formation at the site of an unequal sister chromatid exchange. J. Biol. Chem. 265: 1352-1359.

Weintraub, H. (1983). A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. Cell 32: 1191-1203.

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

Wells, R.D., Collier, D.A., Hanvey, J.C., Shimizu, M., and Wohlrab, F. (1988). The chemistry and biology of unusual DNA structures adopted by oligopurine•oligopyrimidine sequences. FASEB J. 2: 2939-2949.

Williamson, J.R. (1994). G-quartet structures in telomeric DNA. Ann. Rev. Biophys. Biomol. Struct. 23: 703-730.

Wilson, R.W., and Bloomfield, V.A. (1979). Counterion-induced condensation of deoxyribonucleic acid. A light scattering study. Biochemistry 18: 2192-2196.

Wilson, W.D., Hopkins, H.P., Mizan, S., Hamilton, D.D., and Zon, G. (1994a). Thermodynamics of DNA triplex formation in oligomers with and without cytosine bases: influence of buffer species, pH, and sequence. J. Am. Chem. Soc. 116: 3607-3608.

Wilson, W.D., Mizan, S., Tanious, F.A., Yao, S., and Zon, G. (1994b). The interaction of intercalators and groove-binding agents with DNA triple-helical structures: the influence of ligand structure, DNA backbone modifications and sequence. J. Mol. Recogn. 7: 89-98.

Wittig B, Dorbic, T, and Rich, A (1989). The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. J. Cell Biol. 108: 755-764.

Wittig B, Dorbic, T, and Rich, A (1991). Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. Proc. Natl. Acad. Sci. U.S.A. 88: 2259-2263.

Wittig, B., Wolfl, S., Dorbic, T., Vahrson, W., and Rich, A. (1992). Transcription of human c-*myc* in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. EMBO J. 11: 4653-4663.

Wong, A.K., Yee, H.A., van de Sande, J.H., and Rattner, J.B. (1990). Distribution of CT-rich tracts is conserved in vertebrate chromosomes. Chromosoma 99: 344-351.

Wu, H.-Y., Shyy, S., Wang, J. C., and Liu, F. L. (1988). Transcription generates positive and negative supercoiled domains in the template. Cell 53: 433-440.

Xu, G. and Goodridge, A.G. (1996). Characterization of a polypyrimidine/polypurine tract in the promoter of the gene for chicken malic enzyme. J. Biol. Chem. 271: 16008-16019.

Yagil, G. (1991). Paranemic structures of DNA and their role in DNA unwinding. Crit. Rev. Biochem. Mol. Biol. 26: 475-559.

Yagil, G. (1993). The frequency of two-base tracts in eukaryotic genomes. J. Mol. Evol. 37: 123-130.

Yavachev, L.P., Georgiev, O.I., Braga, E.A., Avdonina, T.A., Bogomolova, A.E., Zhurkin, V.B., Nosikov, V.V. and Hadjiolov, A.A. (1986) Nucleotide sequence analysis of the spacer regions flanking the rat tRNA transcription unit and identification of repetitive elements. Nucleic Acids Res. 14: 2799-2810.

Yee, H.A., Wong, A.K., van de Sande, J.H., and Rattner, J.B. (1991). Identification of novel single-stranded d(TC)n binding proteins in several mammalian species. Nucleic Acids Res. 19: 949-953.

Young, S.L., Krawczyk, S.H., Matteucci, M.D., and Toole, J.J. (1991). Triple helix formation inhibits transcription elongation *in vitro*. Proc. Natl. Acad. Sci. U.S.A. 88: 10023-10026.

Zakian, V.A. (1989). Structure and function of telomeres. Annu. Rev. Genet. 23: 579-604.

Zhang, H., Wang, J.C., and Liu, L.F. (1988). Involvement of DNA topoisomerase I in transcription of human ribosomal RNA genes. Proc. Natl. Acad. Sci. U.S.A. 85: 1060-1064.

Zhang-Keck, Z-Y. and Stallcup, M.R. (1988). Optimized reaction conditions and specific inhibitors for initiation of transcription by RNA polymerase II in nuclei from cultured mammalian cells. J. Biol. Chem. 263: 3513-3520.

Zhurkin, V.B. (1985). Sequence-dependent bending of DNA and phasing of nucleosomes. J. Biomol. Struct. Dyn. 2: 785-804.

Zhurkin, V.B., Raghunathan, G., Ulyanov, N.B., Camerini-Otero, R.D., and Jernigan, R.L. (1994). A parallel DNA triplex as a model for the intermediate in homologous recombination. J. Mol. Biol. 239: 181-200.

PUBLICATIONS

1. T.J. Thomas, **C. Ashley**, T. Thomas, A. Shirahata, L.H. Sigal, and J.S. Lee, 1997. Pyrimidine•purine•pyrimidine triplex DNA stabilization in the presence of tetramine and pentamine analogues of spermine. *Biochem. Cell Biol.* **75**, 207-215.

2. **C. Ashley**, M. Morhart, R. Rennie, and B. Ziola, 1997. Release of *Candida albicans* yeast antigens upon interaction with human neutrophils *in vitro*. *J. Med. Microbiol.* **46**, 747-755.

3. J.S. Lee, **C. Ashley**, K.J. Hampel, R. Bradley and D.G. Scraba, 1995. Triplex mediated knotting in circular DNA: T-loops. *Biol. Struct. Dynam.* **2**, 121-128.

4. J.S. Lee, **C. Ashley**, K.J. Hampel, R. Bradley and D.G. Scraba, 1995. A stable interaction between separated pyrimidine•purine tracts in circular DNA. *J. Mol. Biol.* **252**, 283-288.

5. K.J. Hampel, **C. Ashley**, and J S. Lee, 1994. Kilobase-range communication between polypurine•polypyrimidine tracts in linear plasmids mediated by triplex formation: a braided knot between two linear duplexes. *Biochemistry* **33**, 5674-5681.

6. **C. Ashley**, 1993. *Candida albicans* antigens released upon interaction with human neutrophils *in vitro*. M.Sc. Thesis, Department of Microbiology, University of Saskatchewan, Saskatoon.