

**SYNTHESIS PHYSICOCHEMICAL AND BIOLOGICAL STUDIES
ON OLIGONUCLEOTIDES CONTAINING *D*-ARABINOSE**

by

Anne Marietta Noronha

Department of Chemistry
McGill University
Montreal, Canada.

A Thesis submitted to the Faculty of
Graduate Studies and Research
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

© Copyright by Anne Marietta Noronha

July 1999

ABSTRACT

Arabinonucleic acid (ANA), the 2'-stereoisomer of RNA, was synthesized *via* solid phase synthesis in order to test for its ability to associate with single and double stranded nucleic acids (DNA and RNA), as well as to examine priming, RNase H activation, and *in vitro* inhibition of HIV-1 reverse transcription.

The monomer 9-[2'-O-acetyl- β -D-arabinofuranosyl]-[5'-O-monomethoxytrityl] (N²-isobutyryl) O⁶-[2-(4-nitrophenyl)ethyl]guanine was prepared using a combination of synthetic strategies developed in the Damha and Pfeleiderer labs with an overall yield of 19- 22% starting from guanosine. Oligoarabinonucleotides of mixed base composition were found to hybridize to single stranded RNA, but only weakly (if at all) to single stranded DNA. A pyrimidine (Py) ANA strand was shown to form triple-helical complexes only with duplex DNA, and hybrid DNA (purine):RNA (pyrimidine) with an affinity that was slightly lower relative to the corresponding pyrimidine DNA strand. Neither the ANA or DNA pyrimidine strands were able to bind to duplex RNA, or hybrid RNA (Pu): DNA (Py). Such understanding can be applied to the design of sequence selective oligonucleotides which interact with double-stranded nucleic acids, and emphasizes the role of the 2'-OH group as a general recognition and binding determinant of RNA.

A V-shaped oligonucleotide containing a central arabinoadenosine unit and two parallel decahydymidilic acid (dT₁₀) strands joined *via* 2'-5' and 3'-5'-phosphodiester linkages was capable of forming triple-helical complexes in the presence of complementary deoxyadenylate strands with minimal structural effects. These complexes represent the lesser known "anti-parallel" motif, or T*A:T reverse Hoogsteen/Watson-Crick triplex. Hybridization, stoichiometric measurements and CD spectroscopy confirmed the formation of this triplex. Magnesium ions and DNA binding ligands such as Benzo[e]pyridole (BePI), Ethidium Bromide and Hoescht 33258 can also increase the stability of these triplexes. Further developments with this chemistry and the idea of an

arabinose branch point are anticipated to contribute much to our understanding of the native role of branched RNA.

The insertion of arabinonucleotide units within a DNA strand yielded chimeric molecules that exhibited interesting physicochemical and biological characteristics. Although the structural changes of DNA strands caused by arabino inserts were minimal, the biological outcomes were manifold. For example, DNA with araC as the 3'-terminus (DNA-araC) was found to prime RT-catalyzed DNA synthesis, but significant pausing was noted after the addition of 4 nucleotides ($n + 4$ product); no significant amounts of shorter ($n + 1 \dots 3$) products were observed. The appearance of the $n + 4$ product seems to be sequence independent, since it is noted with several 3'-terminal arabinonucleotide oligomers targeted to different sequences on the RNA template. This is a fascinating observation, with potential therapeutic significance.

The same DNA-araC chimera was found to be a very effective inhibitor of several important stages of HIV replication, including RT-catalyzed synthesis of full-length (-) strong-stop DNA as well as RT-mediated strand transfer. Maximal inhibition was noted at a 1:1 antisense oligomer: target RNA ratio, significantly lower than that noted in other studies.

Finally ANA strands when hybridized to their respective target RNA, were found to activate RNase H, an enzyme that has been implicated in the mechanism of action of antisense drugs. These observations are likely to have a very significant impact in clinical chemotherapy and biotechnology.

RÉSUMÉ

L'acide arabinonucléique (AAN), l'épimère en 2' de l'ARN, a été synthétisé sur phase solide afin de vérifier son association avec l'ADN ou l'ARN à simple ou double brins et pour tester l'amorçage, l'activation de l'ARNase H et l'inhibition in vitro de la transcription inverse du VIH-1. Le monomère 9-[2'-O-acétyl- β -D-arabinofuranosyl]-[5'-O-monométhoxytrityl](N²-isobutyryl) O⁶-[2-(4-nitrophényl)éthyl]guanine a été préparé à partir de la guanosine avec un rendement de 22% en utilisant une combinaison de stratégies développées dans les groupes de Damha et Pfeleiderer. L'appariement d'oligoarabinonucléotides de séquence mixte a été observé avec un simple brin d'ARN et avec un simple brin d'ADN, mais très faiblement dans le dernier cas. Un brin AAN contenant des pyrimidines (Py) a formé une triple hélice avec une double hélice d'ADN et un hybride ADN (purine):ARN (Py) avec une plus faible affinité que le brin ADN (Py) correspondant. Par opposition, les mêmes brins pyrimidines AAN et ADN n'ont pas montré d'association avec une double hélice ARN ou l'hybride ARN (Pu):ADN (Py). Ceci peut être appliqué à l'élaboration d'oligonucléotides qui interagissent avec les acides nucléiques double brin et montre bien le rôle prédominant de l'hydroxyle 2' dans l'ARN.

Un oligonucléotide en V contenant une arabinoadénosine centrale et deux brins d'acide désoxythymidylique (dT₁₀) parallèles unis par des liens phosphodiester 2'-5' et 3'-5' a formé une triple hélice en présence d'un brin désoxyadénylate complémentaire. Cette triple hélice représente le motif «antiparallèle» ou triple hélice T*A:T Hoogsteen inversé/W-C et a été caractérisée par différent appariement, mesure stoechiométrique et dichroïsme circulaire. Des développements futurs dans l'utilisation d'un arabinose comme agent de branchement va permettre l'approfondissement de nos connaissances sur l'ARN branché.

L'insertion d'arabinonucléotides dans un brin d'ADN a très peu affecté la structure du brin d'ADN, mais a eu des conséquences importantes sur l'activité biologique. La substitution de dC par araC à la position 3' n'a pas modifié l'amorçage de la synthèse

d'ADN par la transcriptase inverse, mais un arrêt prononcé a été observé après l'addition de 4 nucléotides (produit n+4); aucun autre produit plus court a été observé. L'obtention de ce produit n+4 n'a pas semblé dépendre de la séquence puisque le même effet a été observé avec plusieurs oligomères d'ADN modifiés en 3' par un arabinonucléotide. Ceci est une observation fascinante, avec possibilité d'application thérapeutique.

Le même hybride ADN-araC a démontré un effet inhibiteur sur quelques étapes importantes de la réplication du VIH, incluant la synthèse par la transcriptase inverse de l'ADN «(-) strong-stop» et le transfert de brin facilité par la transcriptase inverse. L'inhibition maximale a été observée à un ratio oligomère antisens: ARN-cible 1:1, ratio plus petit que les autres études.

Un brin d'ADN apparié à un ARN complémentaire a activé la ARNase H, un enzyme impliqué dans le mode d'action des drogues antisens. Cette observation pourrait avoir un impact significatif en biotechnologie et chimiothérapie clinique.

ACKNOWLEDGMENTS

First and foremost, I would like to express my thanks to Professor Masad Damha for allowing me the opportunity to explore the field of nucleic acid chemistry over the course of my studies. His enthusiasm for this project was a constant source of encouragement. For the development and progress of the projects, his patience, inspiration and wisdom - I feel a deep sense of gratitude. “Muchas Gracias!”

Some of this work is a synergistic result of a team effort. In particular I would like to thank Dr. Michael Parniak and his research associates Dr. Gadi Borkow and Dr. Dominique Arion of the McGill AIDS Center for an extremely fruitful collaboration and for giving “life” to my molecules “ANA” in the context of the cell.

For training on various instruments and spectra acquisition I would like to thank Dr. F. Sauriol (NMR), Dr. J. Turnbull (Concordia’s CD Facilities), Dr. N Saade (FAB) and Antisar Hlil (MALDI-MS).

A special thank you to Dr. Suzanne Black for giving me the opportunity to teach and to learn with the freshman students in the General Chemistry Laboratory. Her faith in my capacity to work and learn not only allowed me to realize my dreams, but also opened new avenues.

Thanks for all their administrative assistance: Renee Charron, Carol Brown and especially Sandra Aerssen (for “where’s that thesis Anne?”); things were made a lot easier because of their help. Thanks to the Chemistry Department (McGill University) for support in the form of teaching assistantships, and for giving me the opportunity to study here.

Thanks to

-all my colleagues in room 207, both past and present, for their support and feedback; special thanks to Kazim Ally Agha, Maria Mangos and Anita Liscio for their expert proofreading of this thesis and insightful editorial suggestions.

-Dr. Kristina Kutterer, Dr. Robert Hambalek, Dr. Irene Iziak, Christine Allen and Margherita Scartozzi for their friendship and encouragement over the years.

-Irene Iziak, for helping me through a difficult decision.

-Dr. Louis Cuccia for the many tutorial sessions on the XL-500 as well as for making the Otto Maass Chemistry Building a more pleasant place.

“Merci à Isabelle Paquin pour son amitié, et pour avoir partagé ses connaissances sur Window Excel et PowerPoint. J’apprécie également son excellent travail qu’elle a effectué sous ma supervision.”

-To my dear colleagues but more importantly, friends, Christopher Wilds, Dr. Sébastien Robidoux and Dr. Marita Wasner.

-Marita, thanks for teaching me the art of perseverance and for the insightful discussions on the synthesis of the dreaded (black sheep) ara-G.

-Sébastien, sincere thanks for translating the abstract, your many excellent suggestions throughout the course of my study here, for proof-reading this work, and for helping me to better understand the difference between theory and practice in the lab. Your integrity both personal and professional is admirable.

-Many thanks to Christopher for the editing of the first manuscript of this work (a year ago!), for refining and testing the material in several contexts, for his reliability and for actually making some of the future directions of this work a reality; (and I am still learning!). Thanks again.

Finally I would like to thank my parents whose emphasis on education throughout my life inspired me to go on as far as I did. Thanks:

-to the happy memory of my *late* mother.

-to my father for his devotion to education and for his constant demonstrations of love.

-to all my family for their love, interests, insights and support, especially Steve, J.K, Mike and Ed.

To everyone else whom I forgot to mention, thank you.

Anne Noronha

TABLE OF CONTENTS

Dedication	ii
Abstract	iii
Résumé	v
Acknowledgments	vii
Table of Contents	ix
Abbreviations	xiii
List of Figures	xvii
List of Tables	xxii
I INTRODUCTION	
1.1 IMPORTANCE OF NUCLEIC ACIDS	1
1.2 REGULATION OF GENE EXPRESSION BY OLIGONUCLEOTIDE ANALOGUES	3
1.2.1 Antisense Strategy	4
1.2.2 Triple Helix Formation and the Antigene Approach	13
1.3 HYBRIDIZATION OF NUCLEIC ACIDS	
1.3.1 UV Spectroscopy	19
1.3.2 Circular Dichroic Spectroscopy	21
1.4 PROJECT BACKGROUND AND OBJECTIVES	24
1.4.1 Biochemical and Physiological Studies of Arabinonucleic Acids	25
1.4.2 Structural Studies	26
II SYNTHESIS AND CHARACTERIZATION OF OLIGOARABINO-NUCLEOTIDES OF MIXED BASE COMPOSITION	
2.1 BACKGROUND ON ARABINONUCLEOSIDES AND ARABINONUCLEIC ACIDS	29
2.1.1 Arabinonucleosides	29
2.1.2 Oligoarabinonucleotides	31
2.1.3 Chemical Synthesis of Arabinoguanosine	34

2.2	CHARACTERIZATION OF ARABINOGUANOSINE NUCLEOSIDE DERIVATIVES	40
2.3	SYNTHESIS OF OLIGOMERS OF ARABINONUCLEIC ACIDS (ANA)	44
2.4	CHARACTERIZATION OF OLIGOMERS	45
2.5	CONCLUSION	48
III	PHYSICOCHEMICAL STUDIES OF OLIGONUCLEOTIDES BASED ON D-ARABINOSE	
3.1	ROLE OF 2'-STEREOCHEMISTRY ON DUPLEX FORMATION	
3.1.1	Introduction	50
3.1.2	Specific Objectives	55
3.1.3	Results	58
3.1.4	Discussion	68
3.1.5	Conclusions	71
3.2	SYNTHESIS AND BIOPHYSICAL PROPERTIES OF BRANCH-NUCLEIC ACIDS CONTAINING ARABINOSE AT THE BRANCH POINT	
3.2.1	Introduction	72
3.2.2	Results and Discussion	77
3.2.3	Stabilization of T*A:T (antiparallel) triplexes <i>via</i> DNA binding ligands	94
3.2.4	Conclusions	99
3.3	TRIPLE HELICES CONTAINING ARABINONUCLEOTIDES IN THE THIRD (HOOGSTEEN) STRAND: EFFECT OF INVERTED STEREOCHEMISTRY AT THE 2'-POSITION OF THE SUGAR MOIETY	
3.3.1	Introduction	101
3.3.2	Experimental Design	104
3.3.3	Results	106
3.3.4	Discussion	119
3.3.5	Conclusion	121

3.4	POSSIBLE FORMATION OF AN ARABINONUCLEIC ACID DOUBLE HELIX	
3.4.1	Introduction	123
3.4.2	Results	125
3.4.3	Discussion	127
IV	BIOLOGICAL APPLICATIONS OF ARABINONUCLEIC ACIDS AND THEIR ANALOGUES	
4.1	INHIBITION OF HIV-1 RT SYNTHESIS	
4.1.1	Introduction	130
4.1.2	Results	135
4.1.3	Discussion	139
4.2	FUNCTIONAL CONSEQUENCES OF ARABINOSYL NUCLEOTIDE INSERTS IN DNA ON PRIMING BY HIV-1 RT	
4.2.1	Introduction	141
4.2.2	Results and Discussion	142
4.2.3	Conclusions and Future Directions	151
4.3	MODULATION OF RIBONUCLEASE H (RNase H) ACTIVITY BY HIV-1 RT USING ARABINONUCLEIC ACIDS	
4.3.1	Introduction	153
4.3.2	Results	156
4.3.3	Discussion	158
4.3.4	Conclusions and Biological Implications	163
V	CONTRIBUTIONS TO KNOWLEDGE	164
VI	EXPERIMENTAL	
6.1	GENERAL METHODS	171
6.1.1	General Reagents	171
6.1.2	Chromatography	171
6.1.3	Instruments	172
6.2	OLIGONUCLEOTIDE SYNTHESIS	
6.2.1	Reagents for Derivatization of Nucleosides	173
6.2.2	Derivatization of Solid Support	174
6.2.3	Automated Oligonucleotide Synthesis	174

6.3	PURIFICATION OF OLIGONUCLEOTIDES	
6.3.1	Polyacrylamide Gel Electrophoresis (PAGE)	176
6.3.2	Gel Mobility Shift Assays (Native gels)	177
6.3.3	Visualization of Oligonucleotides	177
6.3.4	Desalting of Oligonucleotides	178
6.3.5	Summary of Sequences Prepared and Purified	179
6.4	CHARACTERIZATION OF OLIGONUCLEOTIDES	
6.4.1	Bio-Physical Characterization: Hybridization Properties	180
6.4.2	Matrix-Assisted Laser Desorption/Ionization Time of Flight Spectrometry	183
6.5	BIOLOGICAL STUDIES	184
6.5.1.	Inhibition of (-) Strong Stop DNA Synthesis	184
6.5.2.	RNase H induction and Priming Assays	185
6.6.	MONOMER PREPARATION	186
	BIBLIOGRAPHY	191

ABBREVIATIONS AND SYMBOLS

A	adenosine
Å	angstrom
A ₂₆₀	UV absorbance measured at 260 nm
Ac	acetyl
Ade	adenine
AIDS	acquired immunodeficiency syndrome
AMV	avian myelocytomatosis virus
ANA	arabinonucleic acid
AP	alkaline phosphatase
APS	ammonium persulphate
araC	arabinocytosine
araCMP	1-β-D-arabinofuranosylcytosine-5'-monophosphate
araCTP	1-β-D-arabinofuranosylcytosine-5'-triphosphate
araNTP	arabinofuranosyl-triphosphate
Asn	asparagine
atm	atmospheres
AZT	azidothymidine
bp	base pairs
B	base
BePI	benzo[e]pyridole
BIS	N,N'-methylene-bis(acrylamide)
bNA	branched nucleic acid
BPB	bromophenol blue
bRNA	branched RNA
Bz	benzoyl
C	Celsius
ca.	approximately
calc	calculated
CEO-	2-cyanoethoxy (β-cyanoethoxy)
COSY	correlated spectroscopy
CPG	controlled pore glass
Cys	cysteine
Cyt	cytosine
d	doublet (NMR)
DBU	1,8-diazabicyclo[5,4,0] undec-1-ene
DCE	1,2-dichloroethane
DCM	dichloromethane
dCTP	2'-deoxycytidine-5'-triphosphate
ddCTP	2',3'-dideoxycytidine-5'-triphosphate
dd	doublet of doublets (NMR)
DEC	1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride
DIPEA	N,N'-diisopropylethylamine
DMAP	4'-(dimethylamino)pyridine

DMF	N,N-dimethylformamide
DMSO	dimethylsulphoxide
DMT	dimethoxytrityl
dN	2'-deoxynucleotides
DNA	2'-deoxyribonucleic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	disodium ethylenediaminetetraacetate.2H ₂ O
<i>e.g.</i>	for example
EtBr	Ethidium Bromide
EtOH	ethanol
EtOAc	ethyl acetate
eq	equivalent
FAB MS	fast atom bombardment mass spectrometry
G	guanosine
ΔG°	Standard Gibbs free energy change
g	gram
Gln	glutamine
Gua	guanine
ΔH°	Standard enthalpy change
h	hour
%H	percent hyperchromicity
HIV-1	human immunodeficiency virus type 1
HMPA	hexamethylphosphoramide
HMQC	heteronuclear multiple quantum correlation
HOAc	glacial acetic acid
HPLC	high performance liquid chromatography
Hz	Hertz
<i>i.e.</i>	that is
<i>i-, n-, t-,</i>	iso, normal, tertiary
<i>i</i> -Bu	isobutyryl
<i>i</i> -Pr	isopropyl
J	coupling constant
λ	wavelength
LCAA-CPG	long-chain alkylamine controlled pore glass
M	molar
m	multiplet
MALDI TOF	matrix assisted laser desorption ionization time of flight
max	maximum
m/c	mass to charge ratio
MeOH	methanol
min	minute
mL	milliliter
mg	milligram
mM	millimolar
μ M	micromolar

MMTr	monomethoxytrityl
mol	mole
mRNA	messenger RNA
MS	mass spectrometry
MW	molecular weight
NBA	nitrobenzyl alcohol
nm	nanometre
NMR	nuclear magnetic resonance
NP1	nuclease P1
npe	<i>para</i> -nitrophenylethyl
nt	nucleotides
OD	optical density
ONT	oligonucleotides
OPC	oligonucleotide purification cartridge
PAGE	polyacrylamide gel electrophoresis
PBS	primer binding site
ppm	parts per million
pre	precursor
Pu	purine
Py	pyrimidine
py	pyridine
q	quartet (NMR)
R	repeat region of the HIV genome
®	register trademark
RAN	random sequence
R _r	retardation factor (TLC mobility)
Rf	radio-frequency
rN	ribonucleotide
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
RT	reverse transcriptase
rt	room temperature
ΔS°	Standard entropy change
s	singlet (NMR)
SEC	size exclusion chromatography
sec	second
SVPDE	Snake Venom Phosphodiesterase
ss	single stranded
T	thymidine
t	triplet (NMR)
T•AT	parallel (or Hoogsteen) TAT base triplet
T*AT	antiparallel (or reverse Hoogsteen) TAT base triplet
TBDMS	<i>t</i> -butyldimethylsilyl
TBAF	tetra- <i>n</i> -butylammonium fluoride
TBE	TRIS/boric acid/EDTA buffer

3TC	lamuvidine or Epivir™
TCA	trichloroacetic acid
TEA	triethylamine
TEMED	N,N,N',N'-tetramethylethylenediamine
TFO	triplex forming oligonucleotides
Thy	thymine
THF	tetrahydrofuran
TIPS	1,1',3,3'tetraisopropylidisiloxane
TLC	thin layer chromatography
T _m	thermal melt transition temperature(melting temperature)
TM	trade mark
TMS	trimethylsilyl
T/P	template/ primer complex
TREATHF	triethylamine-tris(hydrogen fluoride)
TRIS	2-amino-2-(hydroxymethyl)-1,3-propanediol
tRNA	transfer RNA
U	Uridine
Ura	Uracil
UV	ultraviolet
UV-VIS	ultraviolet-visible
vs.	versus
v/v	volume by volume
WC	Watson-Crick
w/v	weight by volume
yDBE	yeast debranching enzyme
XC	xylene cyanol

LIST OF FIGURES

Figure 1.1	Structural representation for 2'-deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).	2
Figure 1.2	Eukaryotic Gene Expression: The flow of biological information in living cells.	6
Figure 1.3	Possible chemical modifications of (a) phosphate groups, (b) bases and (c) sugar groups of natural oligonucleotides.	9
Figure 1.4	Solid phase synthesis cycles using nucleoside-phosphoramidite synthons <i>via</i> the convergent methodology.	11
Figure 1.5	Triple helical formation (pyrimidine motif).	14
Figure 1.6	Triple helix formation: configurations for T•A:T and C ⁺ •G:C triplex motifs.	16
Figure 1.7	Hypothetical melting curve and melting temperature T_m (°C).	20
Figure 1.8	CD spectra illustrating (A) differences between A- and B- form duplexes; (B) differences in triplex and duplex signatures for T•A:T systems.	23
Figure 1.9	Structures of DNA, RNA and arabinonucleic acids (ANA).	24
Figure 1.10	Complexes formed by branched V shaped molecules with deoxyadenylate third strands.	27
Figure 2.1	Arabinonucleosides used as anticancer and antiviral drugs.	29
Figure 2.2a	First chemical synthesis of oligoarabinonucleotides.	33
Figure 2.2b	Formation of O ² -2'-anhydronucleoside synthons limited only to pyrimidine arabinonucleosides.	34
Figure 2.3a	Synthesis of arabinoguanosine <i>via</i> the condensation of 2,6-dichloropurine with xylofuranose tetraacetate.	35
Figure 2.3b	Synthesis of arabinoguanosine <i>via</i> the 8,2'-anhydroguanosine intermediate.	35
Figure 2.3.c	Oxidation-reduction conversion of riboG to araG.	36

Figure 2.4	Synthesis of the completely derivatized <i>arabinoguanosine</i> monomer for solid phase synthesis.	37
Figure 2.5 B	^1H -NMR (500 MHz) of O^6 -[2-(4-Nitrophenyl)ethyl]-9-{3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)-2'-O-[(trifluoromethyl)sulfonyl]- β -D-ribofuranosyl} guanine (2.3).	41
Figure 2.5 B	^1H -NMR (500 MHz) of O^6 -[2-(4-Nitrophenyl)ethyl]-9-{2'-O-acetyl-3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)- β -D-arabinofuranosyl} guanine (2.4).	42
Figure 2.6	Denaturing polyacrylamide gel illustrating migration of single stranded ANA.	46
Figure 2.7	MALDI-TOF Mass Spectrum of sequence 3.31 ara (CCU CUC CUC CCU).	49
Figure 3.1.1	A form and B form duplexes adopted by nucleic acids along with their average parameters.	51
Figure 3.1.2	Sugar puckering equilibria of RNA, DNA and ANA.	53
Figure 3.1.3	The R region sequence of the HIV-1 genomic RNA.	56
Figure 3.1.4	Variation of T_m as a function of RNA content in the antisense strand.	60
Figure 3.1.5	Circular dichroic (CD) spectra of duplexes formed between DNA containing arabino inserts and complementary RNA.	61
Figure 3.1.6 A	CD spectra of DNA:RNA, (DNA-RNA):RNA and RNA:RNA duplexes.	62
Figure 3.1.6 B	CD spectra of DNA:RNA, (DNA-RNA):RNA and RNA:RNA duplexes.	63
Figure 3.1.7	Hybridization of DNA, RNA and ANA to (A) complementary DNA and (B) complementary RNA.	64
Figure 3.1.8 A	CD spectra of single strands (DNA, RNA and ANA)	66
Figure 3.1.8 B	CD spectra of duplexes.	67
Figure 3.2.1	Types of intra and intermolecular triplexes.	73

Figure 3.2.2	Schematic representation of the branched molecules and looped controls.	74
Figure 3.2.3	Primary structure of branched oligomer 3.19 showing the conformation for the branch-point arabinoadenosine unit.	76
Figure 3.2.4 A	Melting curves of complex formed by 3.19 with 1 and 2 equivalents dA ₁₀ in magnesium buffer, at 260 nm.	77
Figure 3.2.4 B	Melting curves of complex formed by 3.19 with 1 equivalent dA ₁₀ in magnesium buffer, at 260 and 284 nm.	78
Figure 3.2.5	Determination of the stoichiometric interaction for 3.19 and 3.20a with dA ₁₀ by the method of continuous variation.	81
Figure 3.2.6	CD spectra as a function of temperature for the complex 3.19 :dA ₁₀ in magnesium buffer.	83
Figure 3.2.7	Comparative CD spectra of the complexes formed by molecules 3.19 and 3.22a with complementary dA ₁₀ (1 equivalent) in magnesium and physiological media.	84
Figure 3.2.8	Concentration dependence of melting temperature (T_m) for the complex 3.19 :dA ₁₀ in magnesium buffer.	87
Figure 3.2.9	The effects of varying the adenylyate chain length (n) on the melting temperature of complex of 3.19 and 3.22a with 1 equivalent of the various dA _n .	89
Figure 3.2.10	Effect of the target chain length and sequence on the T_m of branched complexes.	92
Figure 3.2.11	Complex formation between 3.19 and adenylic acid base on the sugar-phosphate backbone.	93
Figure 3.2.12	The structures of the DNA-binding ligands.	95
Figure 3.2.13	Effect of ligand induced stability on complex formation (Hoescht 33258, EtBr).	97
Figure 3.2.14	Effect of buffer on BePI induced complex formation.	98
Figure 3.3.1	Hairpin duplexes (DD, DR, RD and RR), single strands (D, R and A) and complements (D' and R').	105

Figure 3.3.2	CD Spectra of single strands in acetate buffer.	106
Figure 3.3.3	Gel shift mobility assay of single strands under (A) denaturing conditions and (B) native conditions.	107
Figure 3.3.4	Gel shift mobility assay of mixtures of the thirds strands and their complements.	109
Figure 3.3.5	CD of duplexes, in 100 mM sodium acetate buffer.	111
Figure 3.3.6	UV melting curves of complexes in acetate buffer. (A) to DD target and (B) to DR targets.	113
Figure 3.3.7	Gel shift mobility assay of potential triplexes under non-denaturing conditions.	115
Figure 3.3.8	(A) CD spectra of hairpin duplexes. (B) CD spectra of mixtures of hairpins and single strands. Target is the DD hairpin.	117
Figure 3.3.8	CD spectra of mixtures of hairpins and single strands (C) Target is DR target hairpin and (D) RD target hairpin.	118
Figure 3.3.8 E	CD spectra of mixtures of hairpins and single strands. Target is RR hairpin.	119
Figure 3.4.1	Possible association forms of the Dickerson-Drew dodecamer.	124
Figure 3.4.2	CD spectra of single-stranded oligomers. (Dickerson-Drew study).	126
Figure 3.4.3	Absorbance spectra of single-stranded oligomers. (Dickerson-Drew study).	126
Figure 4.1	The mechanism of reverse transcription in viral gene expression HIV-1 reverse transcriptase.	133
Figure 4.2	Schematic illustration of the anticipated DNA polymerization products in the <i>in vitro</i> DNA polymerization reactions testing the effect of the oligonucleotide on (-) strong stop DNA synthesis.	136
Figure 4.3	DNA polymerization product profiles in the presence of the various oligonucleotides.	137

Figure 4.4	Graphical representation of the quantitation of antisense oligonucleotide RT-catalyzed full-length products.	138
Figure 4.5	<i>In vitro</i> priming assay.	143
Figure 4.6	Autoradiograph of <i>in vitro</i> priming assay.	144
Figure 4.7	(A) Association, (B) pausing, (C) dissociation and (D) extension of reverse transcriptase with the araC terminal nucleotide to the primer:template duplex.	146
Figure 4.8	Pausing or dissociation of the HIV-1 RT after extension of the 4 nucleotides in the primer strand as a result of the arabinocytosine residue at the 3' terminus of the primer.	148
Figure 4.9	Induction of RNase H activity.	153
Figure 4.10	Ribonuclease H degradation of various 18-bp oligonucleotide hybrid duplexes.	157
Figure 4.11	Position of the 2'OH groups in ANA vs. RNA in duplexes.	161

LIST OF TABLES

Table 2.1	^1H and ^{13}C Assignments O^6 -[2-(4-Nitrophenyl)ethyl]-9-{3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)-2'-O-[(trifluoromethyl)sulfonyl]- β -D-ribofuranosyl} guanine (2.3) <i>via</i> HMQC Experiments	43
Table 2.2	^1H and ^{13}C Assignments of O^6 -[2-(4-Nitrophenyl)ethyl]-9-{2'-O-acetyl-3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)- β -D-arabinofuranosyl} guanine (2.4) <i>via</i> HMQC Experiments	43
Table 2.3	Sequence Description of Arabinonucleic Acids (ANA) and Controls Characterized	45
Table 2.4	Observed and Experimentally Determined Molecular Weights of the Oligoarabinonucleotides	48
Table 3.1.1	DNA Sequences Containing Various Ara Inserts and Stretches of Ribonucleotides	57
Table 3.1.2	Thermal Stability of Antisense DNA Strands Containing Arabinonucleotide <i>Inserts</i> with RNA and DNA Sense Strands	58
Table 3.1.3	Thermal Stability of Antisense DNA-RNA Chimera with RNA and DNA Sense Strands	59
Table 3.1.4	Thermal Denaturation Studies Involving Uniformly Substituted Modified Oligomers	65
Table 3.2.1 A	Melting Temperatures (T_m) and Hyperchromicity (%H) for Complexes in Magnesium and Manganese Based Buffers	79
Table 3.2.1 B	Melting Temperatures (T_m) and Hyperchromicity (%H) for Complexes in a Buffer which Approximates the Intracellular Cationic Environment	80
Table 3.2.2	Calculated Thermodynamic Parameters for the Hoogsteen and Reverse Hoogsteen T/A:T Base Triplets	86
Table 3.2.3	Thermal Melt Data for the $\text{aA}^{\text{T10}}_{\text{T10}}$ (3.19): dA_n Complexes	88

Table 3.2.4	ΔT_m and %H Values from Melting Curves (260 nm) of Mismatched Relative to Matched Complexes	91
Table 3.2.5	Melting Temperatures (T_m) and Hyperchromicity (%H) for Complexes Formed Between 3.19 and Ribodecaadenylic Acid, rA ₁₀	93
Table 3.3.1	Thermal Denaturation of Duplexes	108
Table 3.3.2	Triplex Hybridization Properties of Single Strands with Hairpins	114
Table 3.4.1	Dickerson-Drew Oligomers Selected for Study	125
Table 3.4.2	Thermal UV Studies of the Dodecamers in Various Buffers	127
Table 4.1	Priming of Antisense Oligomers by HIV-1 RT	149
Table 6.1	Summary of Sequences Prepared and Purified	179
Table 6.2	Designation and Molar Extinction Coefficients of Various Sequences	181

“The biological scientist is steeped in the satisfaction of studying a well defined area; He is studying “the central point between nothing and all,” the mechanism of life and, therefore of man, and he knows of no more rewarding a way to spend a life”.

Ernest Borek

CHAPTER I INTRODUCTION

1.1 IMPORTANCE OF NUCLEIC ACIDS^{1,2,3,4,5}

Nucleic acids are the most fundamental and important constituents of all living cells. They are the “databases,” of the cell which contain “all the information” required by a cell to live and develop. They also govern a number of the functions of the cell including metabolism and reproduction. Although they play a major role in the processes of life, they are by themselves inert outside the context of a cell.

There are two types of nucleic acids in a cell: DNA (deoxyribonucleic acids) and RNA (ribonucleic acids), both of which are polymers of alternating sugar and phosphate residues (**Figure 1.1**). Attached to the C1' of the sugar atom is a purine or pyrimidine base. The configuration of this linkage is β , as the base lies on the same side as the C5' hydroxymethyl group. The DNA chain is said to have a “directionality” - one end of the chain has a free 5'-hydroxyl (OH) group and the other end a free 3'-OH.

During the majority of a cell's life cycle, DNA is double stranded (ds). The two strands are held together in an antiparallel fashion by Watson-Crick (W-C) hydrogen-bonds between composite purine and pyrimidine bases. Three of the bases are common to both DNA and RNA: adenine, cytosine, guanine. The fourth DNA base, thymine, is replaced by uracil in RNA. The principles of life and evolution (the genetic information) lie in the sequence of these bases and are characteristic for each organism. This complementarity between the bases constitutes one of the major stabilizing interactions (hydrogen bonds) and is also responsible for the intrinsic double helical nature of DNA. The second major interaction is hydrophobic and results in the vertical stacking of π (π) electron systems of the coplanar heterocyclic bases along the vertical axis of the double helix. These two base-derived stabilizing interactions, along with steric and hydrophobic interactions, impart on the DNA molecule the stability necessary to preserve life. The four nucleobases ensure fidelity of genetic information transfer while the sugar-phosphate backbone imparts structure, flexibility and hydrophilicity.

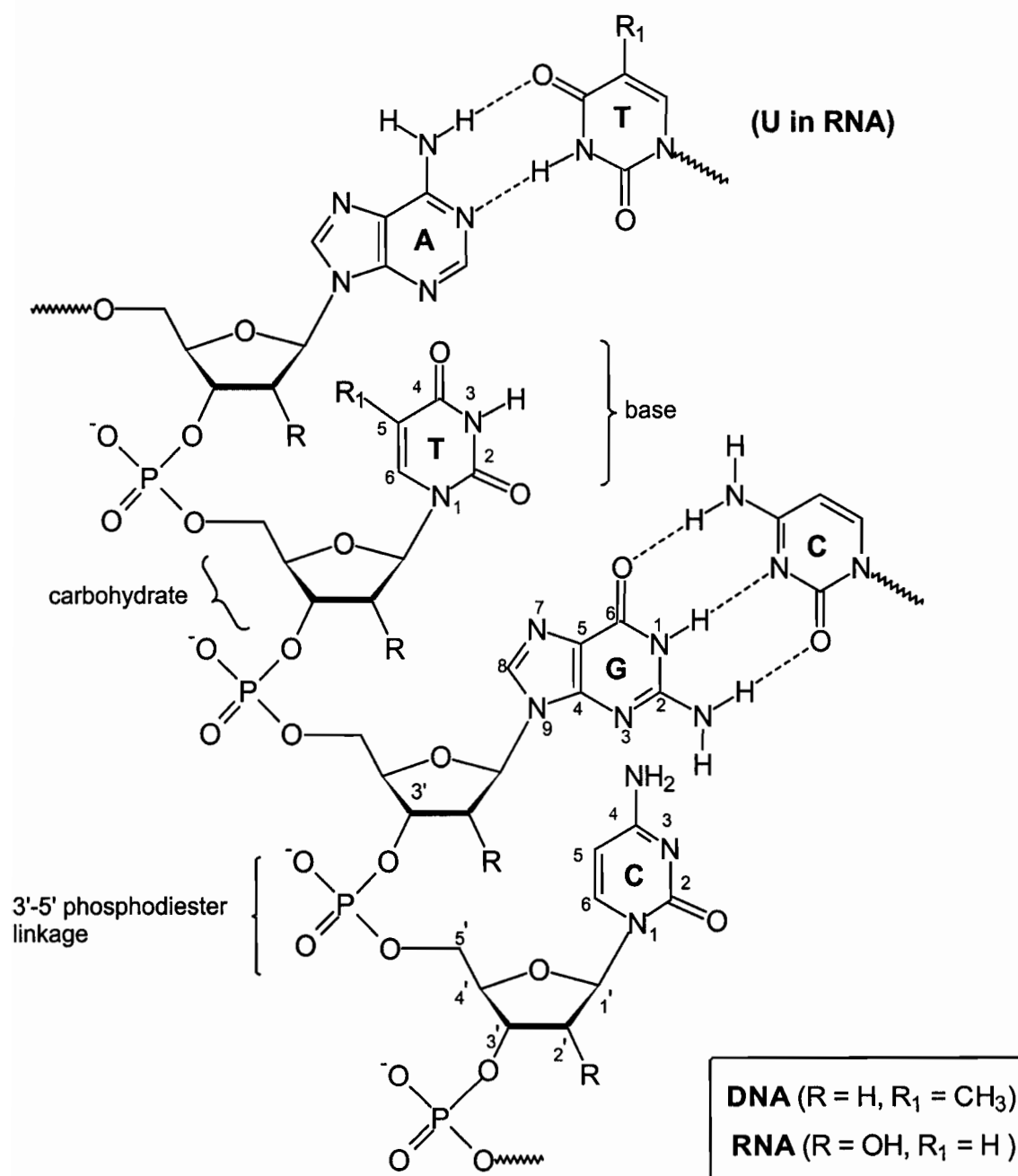


Figure 1.1: Structural representation for 2'-deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Numbering for pyrimidines (C,T,U), purines (A,G) and carbohydrates is given. DNA contains the carbohydrate β -D-2-deoxyribofuranose while RNA contains β -D-2-ribofuranose. The A:T and G:C Watson Crick base pairs are also illustrated.

RNA molecules are usually single stranded (ss) and composed of ribonucleotide units. They are generally shorter than DNA, as they are copied from a limited region of DNA (chromosomes), and have a wide array of tertiary structures. Three major types of RNA exist: messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA). The various roles of RNA in gene expression clearly indicate the significance of these molecules in living cells, for without the intermediary “RNA” the genetic message stored in DNA would have no means of expression. During the course of evolution, DNA, which is chemically more stable, emerged better suited for preserving genetic information.⁶ RNA’s instability lies in the presence of an additional hydroxyl (OH) function at the 2’ position of the sugar ring, making RNA sequences more labile or sensitive than DNA, due to hydroxyl participation in degradation (both enzymatically and chemically).⁷

1.2 REGULATION OF GENE EXPRESSION BY OLIGONUCLEOTIDE ANALOGUES

The path from DNA to protein synthesis begins in a cell’s nucleus, when the double-stranded DNA helix is unwound from one point to another, typically the length of 12 base pairs which is enough for transcription to occur (**Figure 1.2**). RNA polymerase binds one of the strands of the DNA duplex at the promoter region (3’ end), unwinding *ca.* 1 turn of the DNA helix, which now serves as a template for RNA synthesis. RNA is then synthesized, one nucleotide residue at a time, from the template strand of the unwound DNA. This nascent RNA strand grows in the 5’ to 3’ direction, transiently forming a DNA:RNA hybrid of *ca.* 12 nucleotides at the growing 3’ end. The DNA:RNA hybrid then dissociates and the melted region of the DNA rewinds to form the original DNA duplex. This is followed by the RNA polymerase dissociating from the DNA template and then from the enzyme itself. Known as the primary transcript or precursor RNA (pre-mRNA), the nascent RNA is a complementary copy of the original DNA strand. The primary RNA transcript thus formed now requires the excision of its introns (noncoding regions) which occurs in the nucleus in order to form mature mRNA before transport to the cytoplasm. Entry into the cytoplasm signals the ribosomes to recognize the leader region at the 5’-end of the mRNA which then form a ribosomal-mRNA complex.

Everything is now in the right place to begin the translation of mRNA to the polypeptide (protein), of which the first amino acid is always *N*-formylmethionine. This transfer of biological information from a stored form (typically double helical DNA) to functional polypeptides *via* mRNA intermediates comprises the “central dogma of molecular biology”.⁸ Artificial inhibitors of biological functions have traditionally been targeted towards the final protein products of this pathway (“Sense” approach). However in the last few decades, approaches have been suggested for the inhibition of intermediate information transfer reactions prior to “protein” synthesis. Such strategies include “antisense” and “antigene” interactions of oligonucleotides with messenger RNA and the genome, respectively. These interactions may be applicable to the inhibition of many DNA characterized diseases, thereby constituting a potential route to rational drug design.⁹

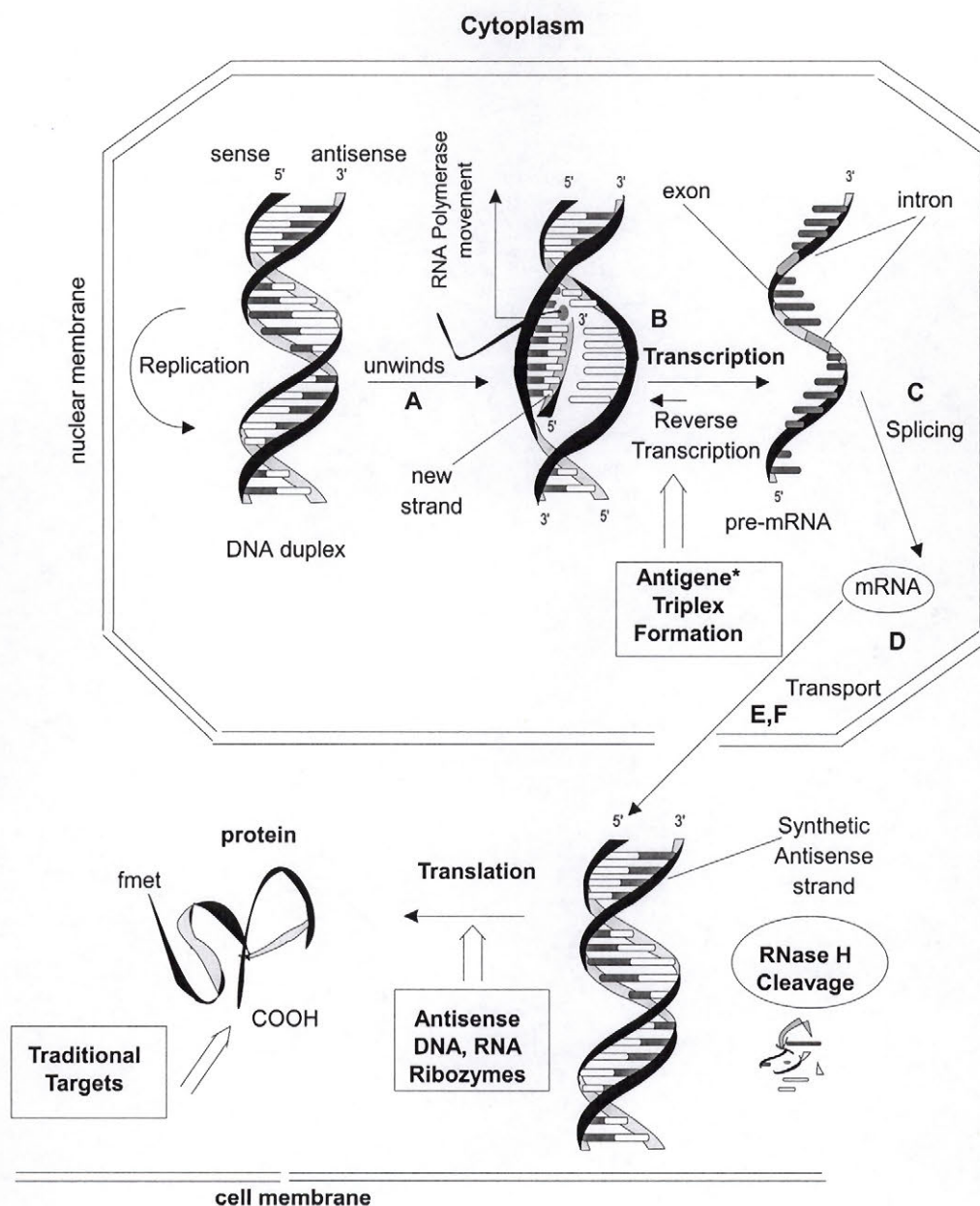
1.2.1 Antisense Strategy^{10,11}

Antisense technology, one of the most exciting fields in genetic research, holds the promise of novel therapies for hard-to-treat diseases. Part of the power of antisense rests on the fact that the genetic code of every known organism is made up of the same handful of nucleotides. As described above, the four-letter chemical alphabet, when combined into a sequence called a gene, encodes the synthesis of a specific protein. Ordinarily, mRNA leaves the nucleus and enters the cytoplasm, where it becomes the physical template for protein synthesis in the cell. If it is prevented from forming, or its genetic signals are blocked, the protein it encodes cannot be synthesized. Several diseases occur when a mutated gene is translated into a “dysfunctional” protein, when a gene over-expresses or underexpresses a given protein, or when the genetic code of a micro-organism or virus codes for proteins that cause disease. One obvious way to prevent a “harmful” protein from forming is to block the mRNA strand that codes for it. Thus the principle of the antisense approach involves the synthesis of antisense molecules which are usually short, chemically modified oligonucleotide sequences that hybridize specifically to a complementary sense sequence to form an “antisense:sense” duplex resulting in the suppression of information flow to the next stage, namely translation or virus production. Although control of gene expression of natural antisense oligoribonucleotides by RNA is

known to occur in prokaryotes,¹² usually DNA analogues are preferred over RNA for synthetic and therapeutic purposes due to their greater stability and sometimes less complicated synthesis.

Traditionally, most drugs target proteins rather than mRNA. However, once protein synthesis begins, millions of copies of a given protein can be produced, the amount depending on the gene coding for it - all of which have to be deactivated. But compared with the protein it produces, the amount of mRNA is trivial. If a drug blocks mRNA rather than a protein, it has a much easier task ahead of it in terms of amount of target to be deactivated. Moreover site-specificity constraints considerably limit the development of drugs that can antagonize the effects of these proteins. Thus antisense technology could provide therapies capable of pinpoint accuracy - effective therapies that could diffuse through the body, striking only “trouble spots” but leaving their surroundings untouched. In other words antisense drugs have in principle, the added advantage of specificity; by blocking the synthesis of only one protein, they are less likely to cause side effects (toxicity).

Modes of Antisense Inhibition.^{13,14a,b} The informational route from RNA to proteins can be prevented by two possible mechanisms. As described above, the first one results as a consequence of a stable duplex formed between the antisense oligonucleotide and either a strand of unwound DNA or the mRNA thus inhibiting transcription (DNA to mRNA) or translation of the nucleotidic sequence to the peptidic sequence respectively. In principle, this “translation arrest” can occur at different stages during the life of the mRNA (**Figure 1.2** Steps B through F). Hybridization of the antisense molecule to the target mRNA can occur within the nucleus, before transcription is completed, during splicing of pre-mRNA to mRNA, or during passage of the mature transcripts from the nucleus to the cytoplasm. Thus, hybridization of the antisense strand to the local unwound loop (DNA) inhibits polymerase movement along the DNA and stops “transcription” by physical arrest (step B). Hybridization of antisense oligomers to the pre-mRNA obstructs formation of mature mRNA by either disrupting pre-mRNA tertiary structure formation or preventing spliceosome assembly.



Strategy:	Antigen*	Antisense	Sense
Interaction :	Hoogsteen	Watson/Crick	Protein/DNA
Inhibition of:	Transcription	Translation	Protein Function

Figure 1.2. Eukaryotic Gene Expression: The flow of biological information in living cells. Transfer of information between macromolecules occurs at replication, transcription, reverse transcription and translation. Proteins have generally been traditional targets for therapeutic intervention. Antisense and antigen strategies seek to interrupt translation and transcription, respectively, by sequence-specific binding to informational macromolecules.

Antisense hybridization at the intron-exon junction inhibits the splicing (cleavage) of the intron and/or the ligation of the exons (step C). Complementary binding of antisense molecules to nuclear mRNA can also interfere with transport of mRNA to the cytoplasm (step D). Finally translation of the mRNA to polypeptides can be blocked by hybridization of the antisense molecule to cytosolic mRNA *via* antisense binding to initiation factors and/or assembly of the ribosomal subunits at the start codon or by preventing the ribosome from scanning the mRNA message in the target sequence (steps E and F). This strategy essentially targets mRNA *upstream of the AUG* promoter with “passive” oligonucleotides (which do not modify the target). Thus the mode of action is a steric block or physical arrest mechanism.

The second, more attractive and accepted route, is the activation of RNase H, an intracellular enzyme that degrades the RNA strand of a hybrid antisense-DNA:RNA duplex.^{14a} Inhibition will not generally occur when targeting the coding region of mRNA unless the resulting antisense-DNA:RNA complex is cleavable by RNase H. For example, because certain antisense oligonucleotide analogues do not elicit RNase H cleavage when targeted downstream of the AUG initiation codon, they must consequently be targeted to a region from the 5' cap site to a few nucleotides upstream of the AUG initiation codon in order to elicit a “physical arrest mechanism”. The contribution of these two mechanisms is primarily dependent on the location of the target along the mRNA, but the same ultimate result is effected in either case (*i.e.* protein synthesis is inhibited). However the location of action of antisense oligonucleotides within the cell is still uncertain.

As described below, antisense oligonucleotides can also be linked to active groups such as metal complexes or photosensitizers which irreversibly damage the target mRNA by cleavage or crosslinkage, respectively.¹⁵ In this case, the location of the target is less important as the irreversible damage ultimately prevents protein synthesis.

Development of Antisense Oligonucleotide Analogues: Antisense drugs are essentially designed to display the following properties: (i) formation of more stable duplexes with RNA under conservation of the base pair specificity (ii) efficient penetration through the cell membranes in order to reach their targets in the cytoplasm or the nucleus (iii) high

stability to intra- and extra-cellular nucleases. (iv) a size that is longer than 16 units, in order to statistically inhibit the expression of a single gene among the entire human genome¹⁶ and (v) high specificity for their target relative to protein and unintended targets.

The chemical nature of the DNA molecule permits modifications at a number of different sites (**Figure 1.3**). Ideally, chemical modifications should result in structural analogues with hybridization properties similar to or better than the parent molecule under physiological conditions. The rapid advances made in the development of a complete set of reliable protecting groups and coupling methodologies permitted oligonucleotide synthesis *via* automated solid phase procedures. With the advent of such automated synthesis¹⁷ more than 200 different analogues have been reported over the past several years.¹⁸ The most widely used chemical route for solid-phase oligonucleotide synthesis is the phosphite triester method introduced by Letsinger¹⁹ and latter refined by Beaucage and Caruthers (**Figure 1.4**).²⁰ It relies on nucleoside phosphoramidite synthons and long-chain alkylamine derivatized controlled pore glass (LCAA-CPG) as a solid-phase matrix supporting the synthesis. It was estimated that greater than two million oligonucleotide sequences are assembled annually employing CPG-based supports.²¹ This approach, as illustrated in **Figure 1.4**, is suitable for not only DNA and RNA synthesis,²² but also for nucleic acids bearing various modifications (**Figure 1.3**).

The modifications can be categorized into three main classes. Oligonucleotides bearing modifications at (a) the phosphate-backbone (internucleoside linkage), (b) the base, and (c) the pentofuranosyl sugar. So far, most of the approaches have focused on the modification of the phosphodiester groups, as in this situation the base pairing necessary for specificity and the orientation of the backbone are still maintained.^{23,24} This category itself includes two types of modifications.

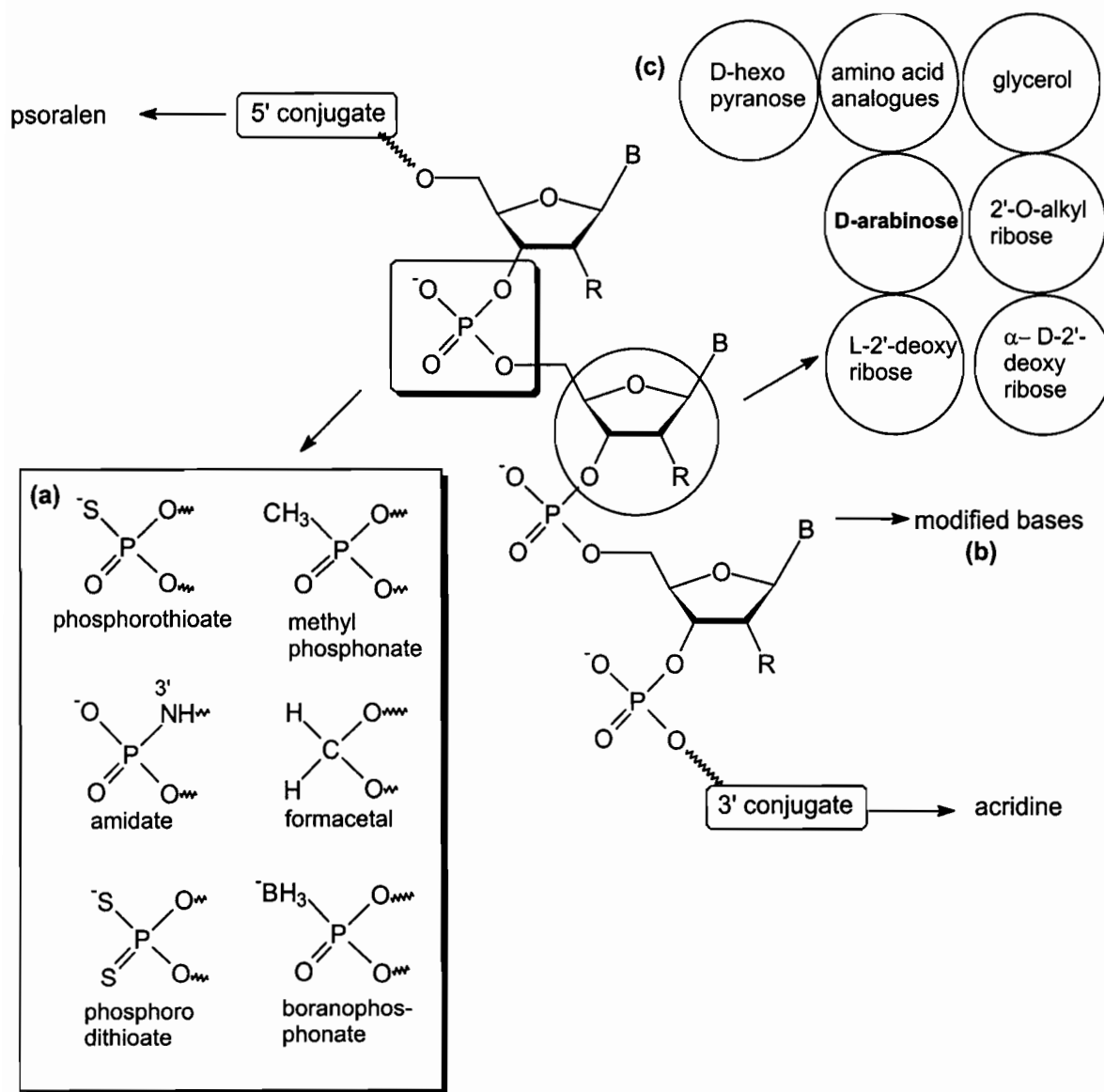


Figure 1.3: Possible chemical modifications of (a) phosphate groups, (b) bases and (c) sugar groups of natural oligonucleotides.

The first requires replacement of the internucleotide phosphate residue by other functional groups (“Dephosphono” oligonucleotides). Examples of this include oligonucleotides containing one or several carbonate,^{25a,b} carbamate,^{21c,d} dialkyl or diaryl silyl,^{24c} methylene^{24f} (methylamino) and amide^{24g,h} groups, among many such internucleotidic linkages. The second class encompasses modifications of only the atoms around the phosphorus atom, or “phosphate” modified oligonucleotides. Some examples include phosphorothioates,²⁶ phosphorodithioates,²⁷ methylphosphonates,²⁸ boranophosphonates²⁹ and phosphoramidates³⁰ (**Figure 1.3**).

Phosphorothioates (S-oligos), wherein one nonbridging phosphate oxygen atom is replaced by sulfur, were introduced by Belikova³¹ and used in stereochemical and mechanistic studies by Eckstein and coworkers.³² These were later shown to possess several desirable properties of antisense agents: they are resistant to nuclease degradation,³³ have sufficient binding capacity to the target and also induce RNase H cleavage of the RNA target.³⁴ However like their phosphodiester counterparts, they do not sufficiently penetrate cells and are required in relatively high concentrations to achieve the desired effects. In addition, these analogues are believed to be implicated in non-specific antisense inhibition with a potential risk of toxicity.³⁵

Methylphosphonates, wherein a non-bonding oxygen is replaced by a methyl group are not only nuclease resistant but, due to the non-ionic nature of the internucleotide bridge increase cellular uptake considerably. However, the binding affinity of these molecules is drastically decreased,³⁶ and they do not elicit RNase H cleavage when bound to the target RNA.

Since the specificity of the antisense strategy is based on Watson-Crick base pairing rules, any modifications to the bases must preserve this base pairing and, as a consequence, this approach has a rather finite scope.²³ Among the base modifications studied, some examples include 5-methylcytosine, 5-bromocytosine,³⁷ 5-ethynyluracil,³⁸ 5-(1-propynyl)-2'-deoxyuracil/deoxycytosine,³⁹ pseudouracil⁴⁰ and 2,6-diaminopurine.⁴¹ Oligonucleotides incorporating these bases have been shown to hybridize to complementary RNA but further modifications in the sugar-phosphate backbone are necessary in order to provide appropriate resistance towards degradative enzymes.

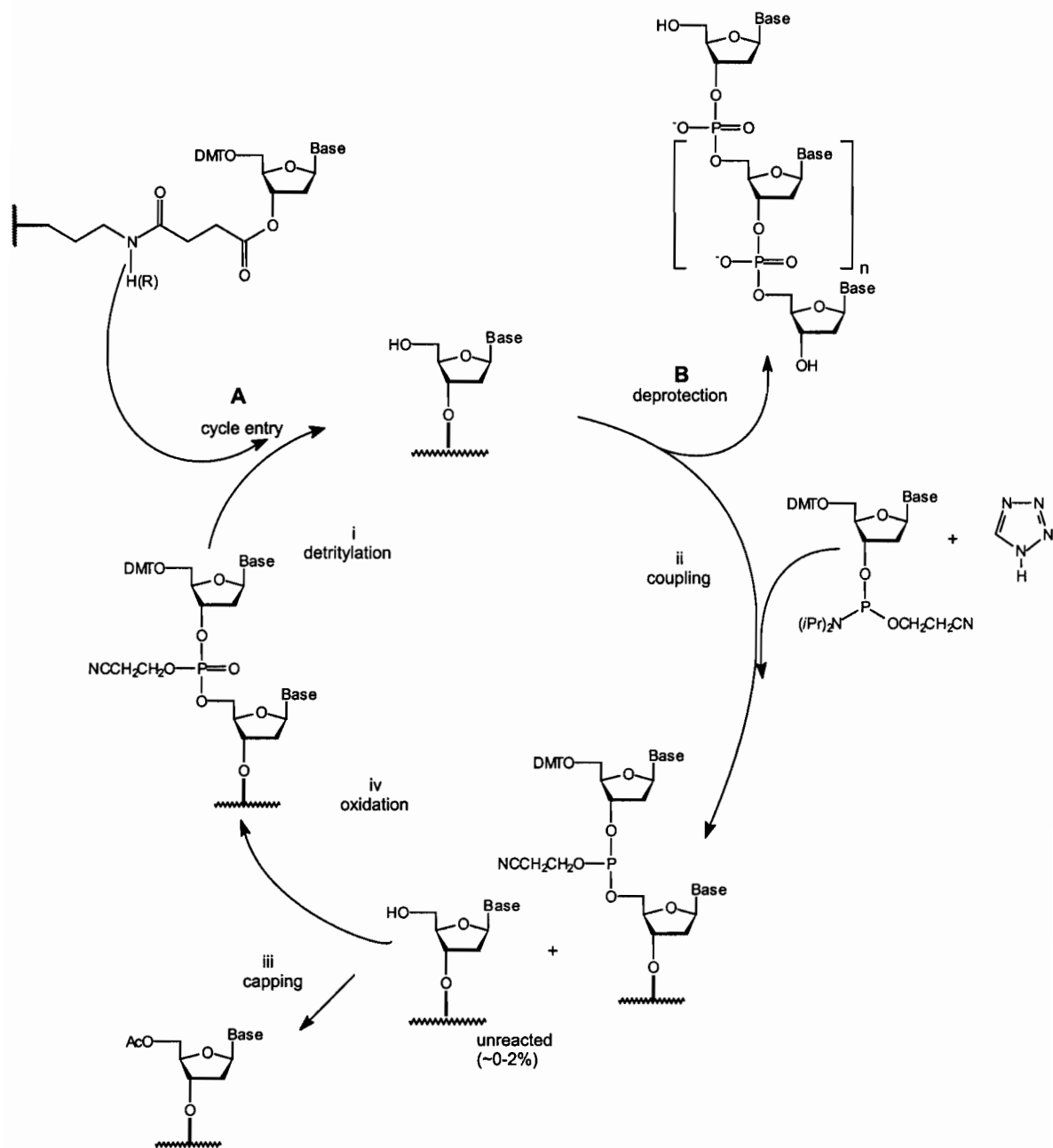


Figure 1.4: Solid phase synthesis cycles using nucleoside-phosphoramidite synthons.¹⁹⁻²¹
(A) Cycle entry with support bound protected nucleoside. (i) Detritylation: trichloroacetic acid in dichloroethane. (ii) Coupling: nucleoside-3'-phosphoramidite and tetrazole. (iii) Capping: acetic anhydride, N-methylimidazole and collidine in THF. (iv) Oxidation: iodine, pyridine and water in THF. **(B)** Cycle exit: Cleavage/Deprotection: the support-bound protected oligomer is concomitantly cleaved and deprotected by treatment with concentrated aqueous ammonia after $n + 1$ cycles.

Recently there has been a considerable interest in the synthesis and properties of oligonucleotide analogues with “sugar” components different from those of natural DNA and RNA.⁴² Structural changes include direct modifications in the pentofuranosyl moiety, for example altering the configuration of the anomeric position (β to α glycosidic bond), or introducing pendant moieties at different positions of the sugar ring.¹⁸ Oligodeoxynucleotides based on L- and D- deoxynucleotides,⁴³ D-glyceronucleosides,⁴⁴ peptido-nucleosides (peptide nucleic acid or PNA)⁴⁵ and morpholinonucleosides⁴⁶ are some notable examples of antisense constructs.

Sugar modifications at the 2' carbon of the pentose ring are gaining popularity. The beneficial effects of a C2' substitution include cellular uptake *in vitro* and *in vivo*, nuclease resistance or binding affinity.^{16,47} These include RNA analogues with 2'-fluoro,⁴⁸ 2'-O- methyl,⁴⁹ and 2'-O-allyl substituents instead of the natural 2'-OH group.⁴² The synthesis of oligonucleotides containing hexopyranoses instead of pentofuranose sugars has also been reported.⁵⁰ A few of these pyranose analogues have increased enzymatic stability but generally suffer from a reduced duplex forming ability with the target sequence. A notable exception is the 6'→4' linked oligomers constructed from 1,5-anhydrohexitol units which, due to their highly pre-organized sugar structure, form very stable complexes with RNA.⁵¹ In fact oligomers composed of arabinopyranose sugars specifically hybridize to complements containing arabinopyranose oligomers but not to RNA or ssDNA.⁵¹ However, none of these hexopyranose oligonucleotide analogues have been shown to elicit RNase H activity.⁵⁰

Lastly, compounds linked at 3' or 5' ends of the phosphate linkage have also been investigated and some of these are illustrated in **Figure 1.3**.^{23,52} Covalently linking lipophilic moieties such as cholesterol to an oligonucleotide effectively improves both its transport and hybridization properties.⁵³ In addition, attachment of photosensitive groups (*e.g.* psoralen) to the termini of oligomers have been shown to damage RNA or ssDNA targets after hybridization.¹⁵

Some interesting applications of the antisense technology include the *in vitro* synergistic inhibition of HIV-1 reverse transcriptase by antisense oligonucleotides and nucleoside analogues,⁵⁴ and inhibition of human melanoma metastasis, by targeting mRNA coding for protein kinase c- α ,⁵⁵ both cases occurring *in vitro*. Other applications

that have advanced to *in vivo* studies are: selective inhibition of α (1B)-adrenergic receptor expression and function,⁵⁶ inhibition of influenza virus RNA polymerase and nucleoprotein expression by phosphorothioate oligonucleotides,⁵⁷ and inhibition growth of human hepatoma cells overproducing Insulin Growth like Factor (IGF-II).⁵⁸

In mid 1998, the U.S. Food and Drug Administration approved fomivirsen (Vitravene)TM, an antisense drug (thioate-DNA) made by ISIS Pharmaceuticals in California, that is injected into the eye to inhibit cytomegalovirus (CMV) retinitis. It specifically targets the viral RNA, thus inhibiting production of a protein that the virus needs to replicate. CMV retinitis is a common infection seen in AIDS patients that can lead rapidly to blindness. Before Vitravene[®] the only effective treatment available for CMV retinitis was the nucleoside analogue ganciclovir developed by former McGill professor Kelvin K. Ogilvie. Patients with low tolerance of ganciclovir[®] are administered Vitravene[®]. Alternatively, a combination of Vitravene[®] and ganciclovir[®] are indicated.⁴⁴

1.2.2 Triple Helix Formation and the Antigene Approach

The selective recognition of DNA base sequences plays a central role in molecular biology. Gene expression is controlled by sequence specific regulatory proteins, many of which have been shown to recognize the major groove of DNA.⁵⁹ Antigene inhibition is designed to target a specific gene in order to modulate transcription of the targeted gene (**Figure 1.2** step A). Because it targets duplex DNA rather than single stranded RNA, it is different from antisense inhibition and intrinsically more difficult. Sequence specific transcriptional arrest demands that oligonucleotides interact with chromatin in a fashion that inhibits transcription. To do this, oligonucleotide analogues must either invade double stranded DNA structures (strand displacement) or form triple stranded structures. Either task is both a chemical and a biological challenge. Given the difficulty of this approach, it is not surprising that progress is at a reduced pace relative to that of the antisense strategy.⁶⁰ This disappointingly slow progress is in spite of the several unique advantages that triplex gene targeting offers, at least in principle.

One of the advantages of triplex gene targeting is the high probability, on a statistically random basis, that a ~16 base pair target sequence will occur uniquely in the

human genome. This means that only low levels of a third strand oligonucleotide of appropriate sequence will be needed for triplex formation. Furthermore, the resulting down-regulation of transcription can be long-lasting, in contrast to the antisense targeting of mRNA, where sustained levels of oligonucleotide may well be required in order to down-regulate multiple mRNA molecules as they are synthesized.

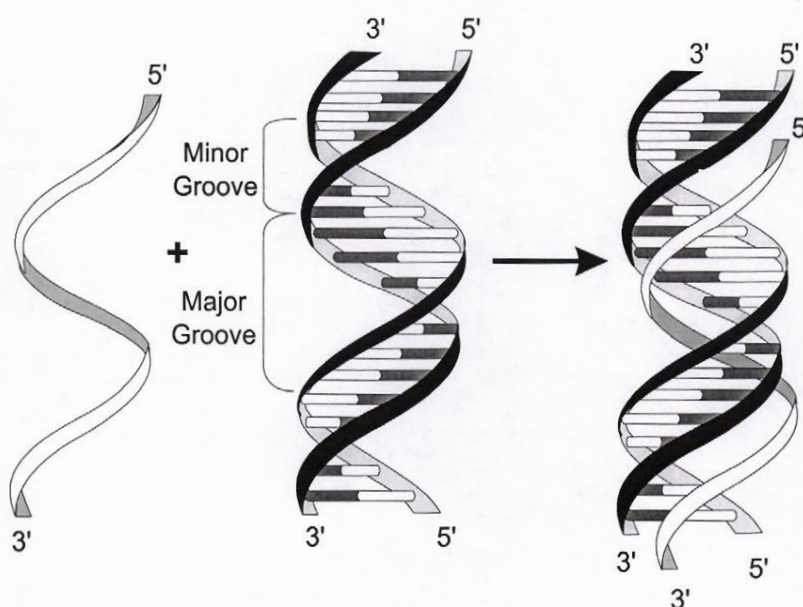


Figure 1.5: Triple helical formation (pyrimidine motif). Ribbon model of the local triple-helical structure formed by the binding of an oligonucleotide to a target DNA sequence. The Watson-Crick duplex strands are depicted as black ribbons, while the third oligonucleotide is depicted as a white ribbon.

Triple-helical RNA complexes have been recognized since 1957.⁶¹ Subsequently, it was established that ribo- as well as deoxyribohomopyrimidine strands ('third strands') could associate with complementary double helical homopurine-homopyrimidine complexes to form triple helices (**Figure 1.5**).^{62,63} At least two structural classes of DNA triple helices exist that differ in hydrogen bonding, sequence compositions of the third strand, relative orientations and relative positions of the sugar-phosphate backbones.⁶⁴ Both classes contain oligonucleotides binding in the major groove of duplex DNA, by

specific interactions between the third strand⁶⁵ and the purine strand of the target duplex. Since polypurine sequences are widespread in eukaryotic genomes and often occur in positions that flank transcribed genes,⁶⁶ the antigene approach to gene regulation can in principle be put into practice.

In the first more commonly described ‘pyrimidine’ motif, the third strand binds *via* Hoogsteen base pairing, with thymines and/or uracils recognizing only A:T^{61,67} base pairs, and protonated cytosines recognizing only G:C base pairs (**Figure 1.6**).^{63,64,68,69} Protonation of the N³ of the pyrimidine ring of C or 5-Me-C is necessary for this binding motif and generally limits its use to acidic conditions (pH 5.5-6.5) or targets which are relatively poor in G:C base pairs.⁶⁴ Despite this limitation, pyrimidine rich oligonucleotides have been used to inhibit the binding of transcription factor *Sp1*⁷⁰ and RNA polymerase II transcription *in vitro*.⁷¹ The pH dependence of binding can be overcome using analogues of C which have the correct hydrogen bond donors at pH 7; examples include 2'-O-methyl-pseudocytidine,⁷² 6-methyl-8-oxo-2'-deoxyadenosine⁷³ and 3-methyl-5-amino-1-pyrazol-[4,3-d] pyrimidin-7-one.⁷⁴ These bases have all been identified as capable of triplex interaction with G:C base pairing at physiological pH, allowing for sequence-specific inhibition under physiological conditions. Oligonucleotide-directed triple helix formation in the ‘pyrimidine’ motif is stabilized by increasing oligonucleotide length (11-15 nt) as expected by analogy with double helical-DNA.⁹ The terms “parallel” and “antiparallel” are sometimes used in the literature to refer to pyrimidine Hoogsteen (•) and purine reverse-Hoogsteen (*) triplexes respectively.

In the second class termed the ‘purine’ motif, a purine third strand binds in an antiparallel fashion to the duplex purine strand through reverse-Hoogsteen base pairing.^{64,75,76} Base pairing in this motif includes guanosine binding to guanosine (G*G:C) and either adenosine or thymidine binding to adenosine (A*A:T or T*A:T).

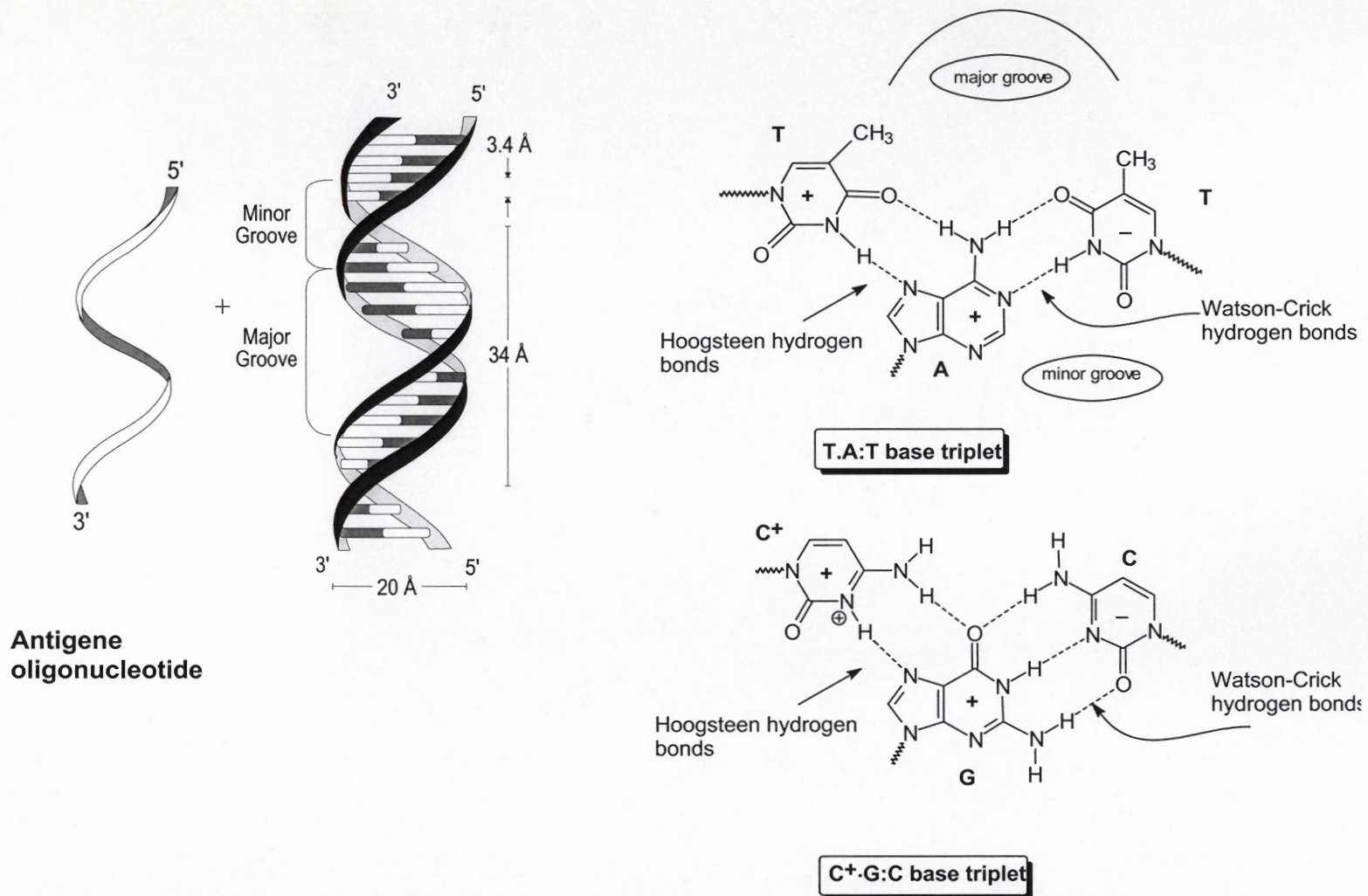


Figure 1.6: Triple helix formation: Configurations for T•A:T and C⁺•G:C base triplets conforming to the more common “pyrimidine” motif. The (+) and (-) signs within the bases represent the polarities of the strand. The third strand lies in the major groove parallel to the purine strand of the Watson-Crick duplex and binds *via* Hoogsteen H-bonding. Adapted from references 75 and 208.

Unlike the pyrimidine motif triplexes, these purine motif (Pu*Pu:Py) triplexes have been shown to exist at physiological pH, and are stabilized by polyvalent cations such as Mg^{2+} , but not by monovalent cations such as K^+ , Rb^+ , NH_4^+ and Na^+ .^{77,78} Oligonucleotides which bind in the purine motif have been shown to form specific complexes with a number of biologically significant targets including the promoters for human epidermal growth factor receptor gene, “proto-oncogene” *HER2* and *c-myc* gene,⁷⁹ and inhibit transcription from the latter *in vitro*⁷⁵ and possibly *in vivo*.⁸⁰

Applications and Limitations of the Antigene strategy: The concept of the antigene strategy has been clearly verified *in vitro* for both purine and pyrimidine motifs. Oligonucleotides bound to duplex DNA were able to inhibit *in vitro* transcription⁸¹ by altering DNA-protein interactions (RNA polymerase or transcription factors) or by blocking transcription, elongation and unwinding of the duplex DNA.⁸² Other biomedical applications involve the mapping of megabase DNA through the selective cleavage of specific DNA duplex domains.^{64,83,84} The effective development of such applications requires a knowledge of the affinities and specificities of third strand probes for target DNA:DNA, RNA:RNA and RNA:DNA duplex domains. In recognition of this need, several studies have focused on elucidating the influence of ‘strand composition’ (RNA versus DNA) on the stabilities and energetics of triplex structures.^{85,86,87}

Antigene strategies suffer from the following problems: (a) third strand association with duplex is invariably weak, so a high local concentration of the third strand would be required in cells; (b) the $C^+ \bullet G:C$ triplet in pyrimidine triplexes is only stable at $pH < 6.5$, so a C:G-rich parallel triplex is generally unstable at physiological pH unless there exists an accompanying “helper effect” such as protein binding, or psoralen cross-linking as described above;⁸⁸ (c) an adequate level of third strand has to reach the nucleus, and thus needs to transcend two membrane barriers - there is evidence that this does indeed occur, at least to some extent;⁸⁹ (d) the target site may be inaccessible due to its masking by chromosomal and other nuclear proteins, although this is less likely to be a problem when a gene is being actively transcribed; and (e) despite much effort, the triplex recognition code is not general for all sequences. Triplex formation has been considered to be most effective when directed against perfect tracts of oligopurine:oligopyrimidine DNA sequences, although an increasing number of studies

have suggested that ‘purine motif’ triplexes with a significant number of sequence imperfections are stable under cell culture conditions.⁹⁰

Although triple helices were known since 1957,⁶¹ Dervan,⁶⁴ and Hélène⁹¹ only recently demonstrated the feasibility of the antigene concept. Triplex stability in both cell-free systems and *in vitro* (human P450 aromatase gene)^{88a} were enhanced by intercalating groups such as acridine linked to the 5’end of the third strand. A more direct demonstration of triplex formation has been made by means of a primer extension assay, in which *Taq* polymerase is stopped at the site of a psoralen cross-linked lesion at the promoter of the interleukin-2 receptor α subunit. Inhibition of transcription has been reported to persist in HeLa cells up to 72 h using a 5’-psoralen linked 19-mer third strand triplex.^{92a} A similar oligomer targeting the *supF* gene in the mouse genome was also shown to have varying effects^{92b} in repair-deficient cells that cause xeroderma pigmentosum.^{92c-e} Several other cellular studies have highlighted the large number of possible targets for triplex formation within a gene.

Triplex formation competes with binding of transcription factors, although the possibility of third-strand interaction with these proteins rather than DNA cannot be ruled out. Non-toxic third strand oligonucleotide concentrations of $\sim 2\mu\text{M}$ have successfully demonstrated inhibitory effects of gene expression *via* triplex formation mechanism *e.g.*, in NF- κB ^{93a} promoter sequences and human tumor factor^{93b} (TNF gene). Within the viral genomes, the potential for triplex specific targeting has been shown for HIV-1 provirus in chronically infected cells.^{93c}

Very little is known about the mechanisms that govern triplex-stabilizing binding events. There is a need to define solution conditions and structural modifications which could predictably alter triplex-stabilizing properties. Some of these requirements will be investigated in Chapter 3, Section 2. Specifically the pyrimidine triplex motif, which is less known and more difficult to form at physiological conditions, will be examined.

1.3 HYBRIDIZATION OF NUCLEIC ACIDS

1.3.1 UV Spectroscopy⁹⁴

A fundamental biophysical intrinsic property of single stranded DNA and RNA is their ability to associate or *hybridize* to other single stranded nucleic acids. This interaction is sequence specific and may be intermolecular or intramolecular in nature. Intramolecular interactions arise when self complementary regions exist in the nucleic acid itself. Unimolecular, bimolecular, as well as higher order complexes such as triplexes,^{63,95} tetraplexes⁹⁶ and three- and four-stranded junctions are also known to exist.⁹⁷

Monitoring Duplex and Triplex Formation:

The nucleobases are chromophores absorbing UV light at wavelength (λ) *ca.* 260 nm. Hence UV spectrophotometrical changes that accompany the dissociation and association of nucleic acid complexes can be used to monitor and consequently determine if the desired complexes (duplexes or triplexes) are indeed being formed.⁹⁸ This is because nucleic acid polymers are known to absorb UV light to a “lesser” extent than the sum of their monomeric components. Moreover, aqueous solutions of nucleic acids absorb less UV light at lower than at higher temperatures. This occurs as a result of greater electronic interactions between the chromophores in the ordered complex compared to the single strands (random coil). The ordered structures are predominant at lower temperatures while the single strand population increases as temperatures increase. This decrease in UV absorbance for the ordered complexes is known as “*Hypochromicity*”. Hypochromicity is usually reported as a percentage, according to the equation $\Delta H\% = [(A_f - A_i) / A_f] \times 100$, where A_i = solution absorbance at initial temperature and A_f = solution absorbance at final temperature. $\Delta H\%$ is thus an indication of the chromophoric base stacking interactions between the strands. (For a detailed explanation of the currently accepted hypochromicity theory refer to the literature review).⁹⁸ The “Melting Temperature” (T_m) is the temperature at which only one half of the complex-population exists in the hybridized form and the remaining half is fully dissociated. T_m measurements involve mixing the strands of interest together, followed by incubation in the appropriate buffer at low temperatures in order to promote annealing. The UV absorption behaviour is then monitored at 260 nm over a specific temperature range (265nm for C⁺:C rich systems, 284 nm T•A:T system, 300nm C⁺•G:C system). The resulting profile

of absorbance *versus* temperature is known as the "Melt Curve" or "Melt Profile" (Figure 1.7).

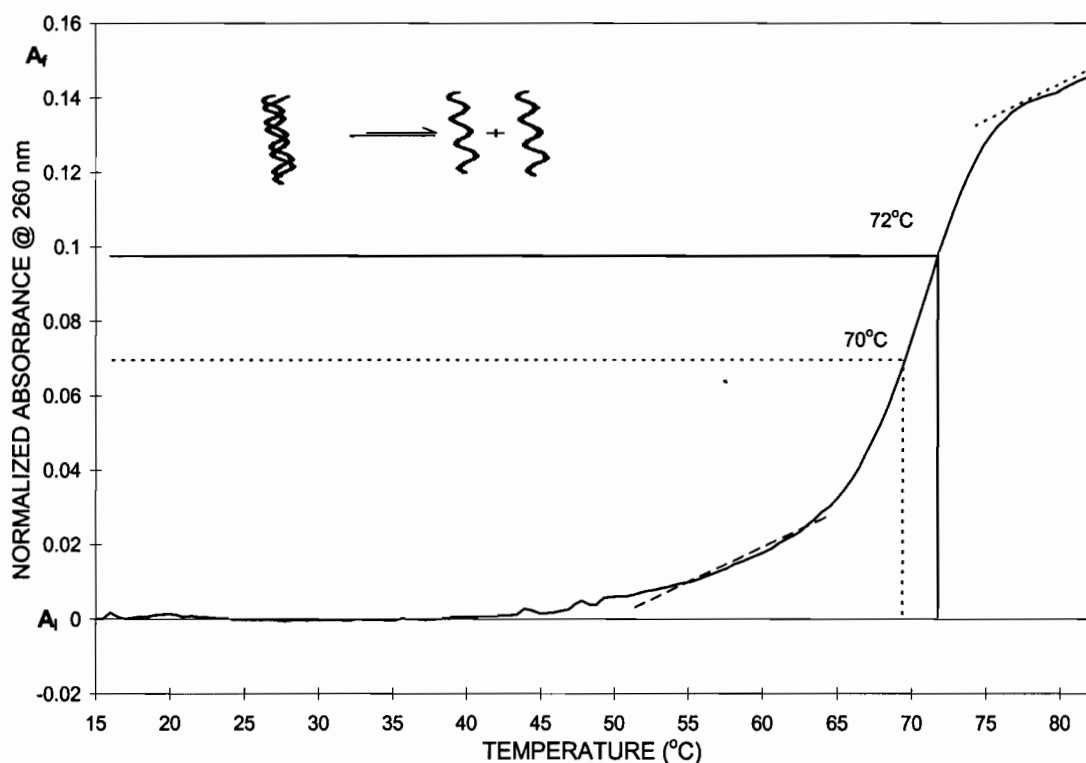


Figure 1.7: Hypothetical melting curve from which the melting temperature or T_m is determined. The tangents on either sides of the steepest rise are determined. A_f represents the absorbance at high temperature and A_i absorbance at lower temperatures. The base-line method for T_m determination involves determining the average ($\frac{A_f + A_i}{2}$) (absorbance axis) which is then extrapolated onto the temperature axis to give the T_m (ca. 70°C). The T_m can also be determined by measuring T at $dA/dT = \text{maximum}$ (first derivative method); in this case $T_m = 72^\circ\text{C}$. The base-line method was used in the studies described in this thesis.

Melting of the third strand from the underlying duplexes, or of the duplexes themselves can be observed as monophasic melt profiles with the appropriate choice of wavelength. Loss of the third strand from a C:G rich system coincides with a hypochromic UV absorbance change at 300 nm and no change at 284, while the

denaturation of the duplex coincides with a hyperchromic UV absorbance at 284 and no change at 300 nm (Chapter 3, Section 3). Dissociation of T/A:T triplexes (both Hoogsteen and reverse-Hoogsteen) into duplex and single strands (dT)_n is accompanied by a hyperchromic absorbance change at 284 nm, while the denaturation of the duplex shows no change at this wavelength⁹⁹ (Chapter 3, Section 2). However the denaturation of both the Watson-Crick and Hoogsteen paired strands is accompanied by a hyperchromic change at 260 nm.

1.3.2 Circular Dichroic Spectroscopy¹⁰⁰

Circular Dichroism (CD) is a powerful technique for studying nucleic acid structure, *i.e.* absolute configuration, complexation of small molecules, and three-dimensional structure (conformation, helical arrangement). This is achieved by alternately measuring the absorbance of left and right handed circularly polarized light as a function of wavelength (λ),¹⁰¹ in the UV range. The CD spectrum usually detects a region of rapid change with respect to λ termed the Cotton effect. Both the sign and absolute intensity of the Cotton effects are sensitive to molecular stereochemistry. The CD spectra of many types of nucleic acid complexes have been determined and certain patterns and signatures established. Hence CD analysis proves useful in structural elucidation of new complex formation. Although an ‘absolute’ method for determining stereochemistry of small molecules, CD is used more as a comparative than an absolute method for the studies involving larger biomolecules such as the ones described in this thesis.

The nucleobases themselves have a plane of symmetry, and thus are not intrinsically optically active. When the bases are attached to the C1' of the asymmetric pentose, the resulting nucleoside is now also asymmetric and this induces CD in the absorption bands of the chromophoric bases. As a result of this secondary effect the intensity of the CD peaks is unfortunately low. This situation is however different for the nucleosides in oligomers and nucleic acid complexes. As mentioned above the hydrophobic planes, the hydrophilic edges, as well as charge-charge interactions, cause the bases to stack and the polymer as a whole to assume a helical structure (**Figure 1.1**). Thus the electronic transitions of the chromophoric bases in a “duplex” are now in close proximity, and can interact to give CD spectra of high intensity. The nucleobase is the

biggest contributor to the CD spectrum and consequent analyses. Thus a large number of π - π^* -type transitions are involved and begin at about 300 nm, with high intensity. Other sources of electronic transitions include (a) phosphate groups (170 nm, high energy), the sugar moiety (190 nm, low energy) and nonbonding electrons, the latter giving rise to n - π^* transitions of low intensity that are buried under the π - π^* transitions. Since CD is dependent on base-base interactions, the technique is sensitive to nucleic acid secondary and primary structure.

In the case when polymorphism is dependent on solvent conditions, CD becomes the method of choice for following changes in secondary structure as a function of solvent conditions. For example the B to A transition observed when ethanol is added to a duplex solution can usually be monitored by CD. Circular dichroism will be used in this thesis to determine the relative conformation of nucleic acid complexes (*i.e.* A, B form etc.), and the nature of the complex formed (*i.e.* triplex or duplex). Some characteristic features of these complexes are exemplified in **Figure 1.8**.

B form versus A form Helices: B-form helices display a positive band centered near 275 nm, a negative band near 240 nm and a crossover around 260 nm (the wavelength maximum for the normal absorption). The CD band in the longer wavelength region does not appreciably change upon melting of the duplex, whereas the high intensity CD bands in the short-wavelength regions are sensitive to the G:C content of the particular DNA and show large changes in intensity as a result of melting. A-form duplexes exhibit a strong wavelength band centered around 260 nm, and a fairly intense negative band around 210 nm. (**Figure 1.8A**).

Nature of Complex: Triplex signatures for T•A:T and C⁺•G:C motifs: The spectra for sequences of both C⁺•G:C (data not shown) and T•A:T systems exhibit changes in the 240-300 nm region with increasing triplex:duplex ratios (**Figure 1.8 B**).

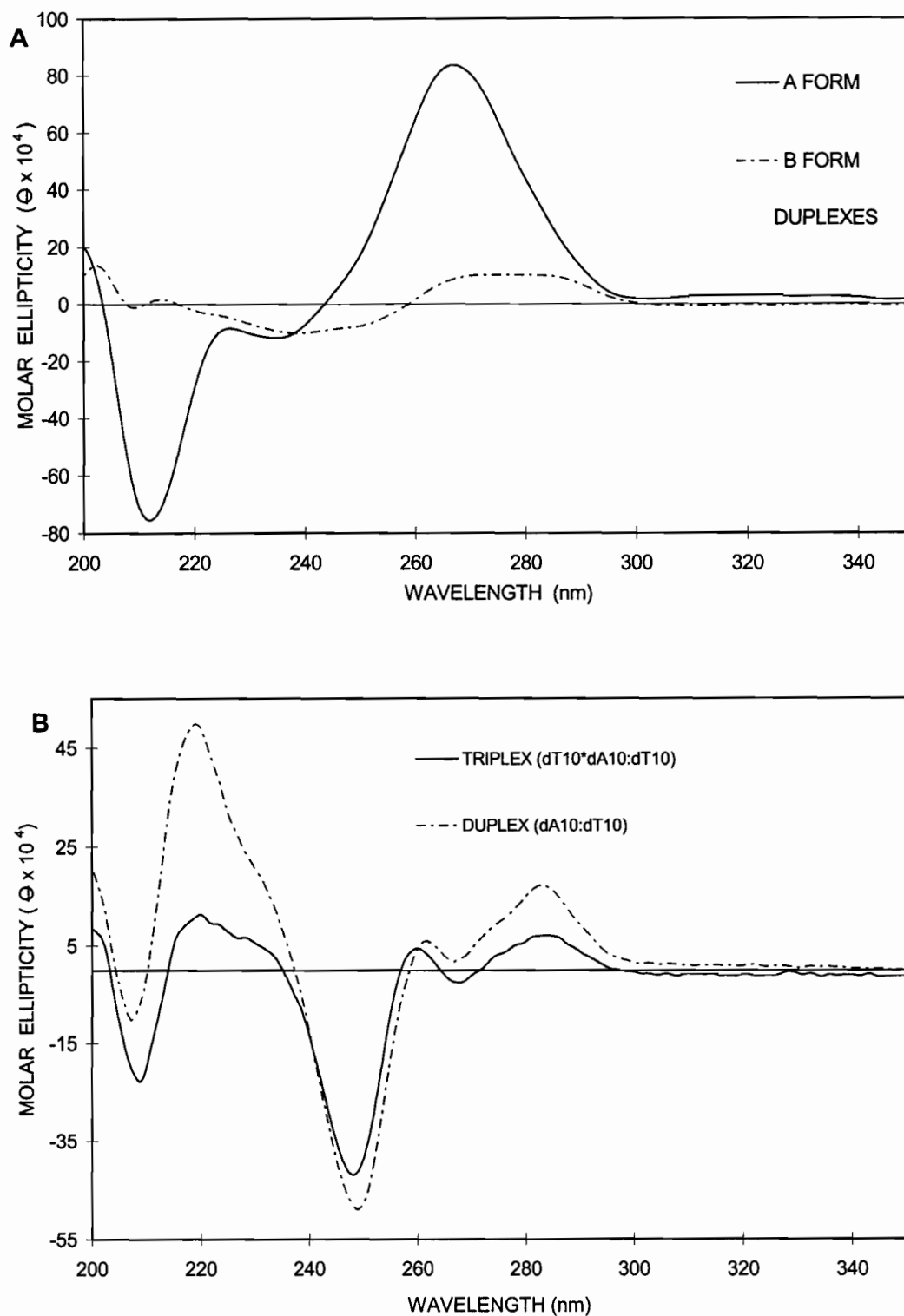


Figure 1.8 CD spectra illustrating (A) differences between A- and B- form duplexes; (B) differences in triplex and duplex signatures for T*A:T systems.

However, the nature of these changes appears to be unique to the specific sequence. Unlike the spectral changes in the 240-300 nm region, those in the 210-240 nm region are more independent of sequence. With increasing triplex:duplex ratios, the spectra of both sequences show a marked amplitude decrease, broadening and red-shift of the negative band at 210 nm. The similarity in spectral variations that both sequences display in the 210- 240 nm region upon alteration of the triplex:duplex ratio suggests that this spectral region presents a useful circular dichroic signature of triplex formation irrespective of base sequence¹⁰² (See Chapter 3, Section 2 - 3).

1.4 PROJECT BACKGROUND AND OBJECTIVES

This thesis work focuses on the synthesis and use of oligonucleotides based on D-*arabinose* instead of the natural D-ribose (**Figure 1.9**). Three independent driving forces exist for studying oligoarabinonucleotides, namely, (I) to obtain information on nucleic acid structure, (II) to study protein:nucleic acid interactions (*e.g.*, interactions of HIV-1 RT and RNase H with nucleic acids), and (III) to assess their use in antisense and antigene strategies.

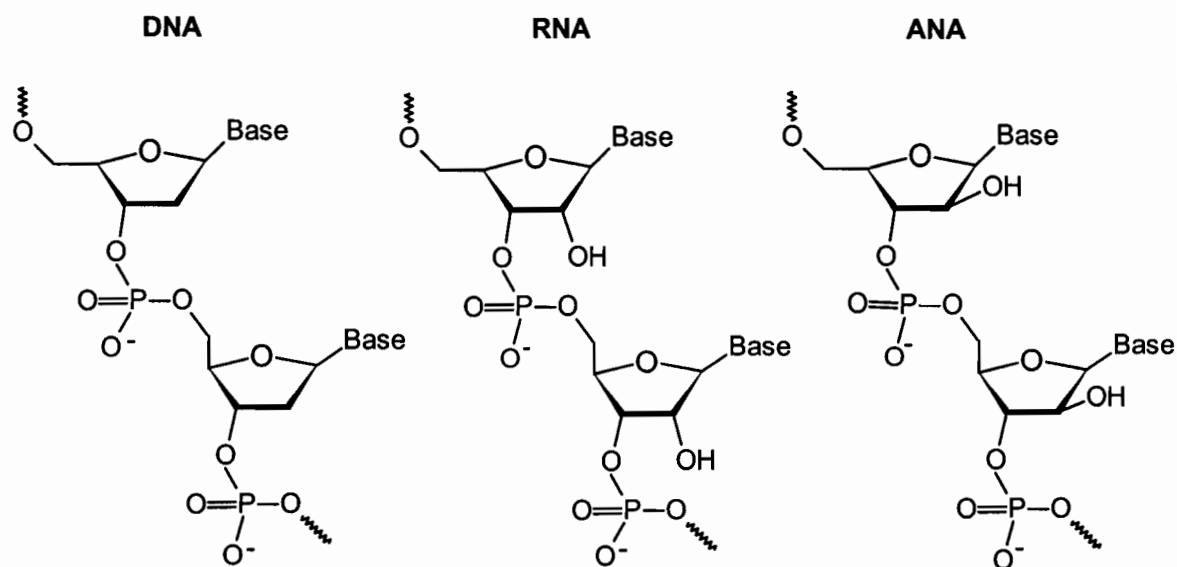


Figure 1.9: Structures of DNA, RNA and arabinonucleic acids (ANA). ANA like RNA, has a hydroxyl group (OH) at the C2' position, but unlike RNA this is *trans* to the phosphodiester linkages.

1.4.1 Biochemical and Physicochemical Studies of Arabinonucleic Acids

Antisense oligonucleotides hold great promise as highly specific therapeutic agents against a wide range of human diseases including viral infections and cancer. Antisense oligomers that modulate gene expression by more than one mechanism of action are highly desirable as this increases the potential efficacy of these compounds *in vivo*.¹⁰³ In addition to the antisense oligomer binding tightly to its complementary target RNA, another essential requirement in the antisense approach is the ability of the resulting antisense oligomer:RNA hybrid to serve as a substrate of RNase H. The added effect is likely to have therapeutic value by enhancing the antisense effect relative to oligomers that are unable to activate this enzyme. Apart from phosphorothioate (PS-DNA) and phosphorodithioate (PS₂-DNA), boranophosphonate-linked DNA and oligomers containing an internal PS-DNA segment, there are no other examples of fully modified oligonucleotides that elicit RNase H activity.¹⁰⁴ Although oligonucleotide phosphorothioates are performing well in clinical trials and are moving rapidly towards New Drug Applications (NDA), sequence-nonspecific biological effects are often encountered in the evaluation of these compounds.¹⁰⁵ Moreover, as indicated above, the therapeutic applications of the phosphorothioate molecules are somewhat limited since they form less stable duplexes with the target nucleic acid than do normal phosphodiester oligomers and are also less efficient than the corresponding phosphodiester-linked oligodeoxynucleotides with respect to RNase H activity.¹⁰⁶ For all of these reasons and because of the problems encountered with PS-oligomers (*e.g.* non-specific antisense effects and potential risk of toxicity), alternative oligonucleotide analogues that selectively block gene expression through the activation of RNase H activity are constantly being sought. From this point of view, oligoarabinonucleotides may have advantageous and interesting properties, since the arabinonucleosides are well known for their clinical potential as antiviral and anti-neoplastic agents.^{107,108} Moreover arabinonucleotides appear to be more stable to enzyme degradation,¹⁰⁹ depurination¹¹⁰ and may be promising alternatives to normal deoxynucleotides for use in antisense and antigene strategies.^{111,112,113} It therefore appears important, from a practical and theoretical point of view, to study the thermal stability of ANA:RNA duplexes as well as

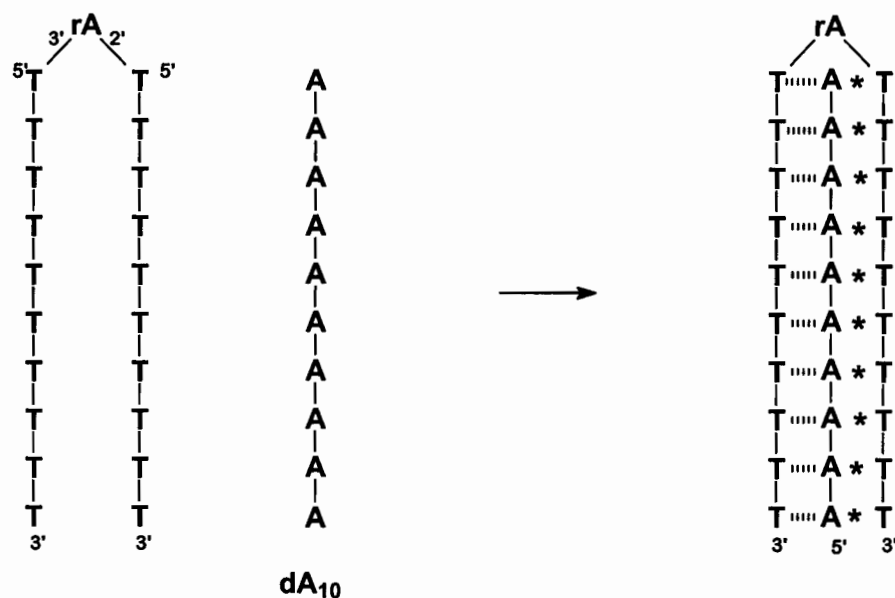
to assess their susceptibility to RNase H cleavage. This would be of interest, not only for the antisense area, but also in the study of the mechanism of action of different enzymes. Thus as a further extension of the studies conducted earlier in this research group (Damha's lab),¹¹⁴ the synthesis of arabinonucleic acids (ANA, **Figure 1.9**) of mixed base (A, U, C, G) composition will be conducted for the first time (Chapter 2). The properties of the complexes formed between ANA and RNA, and ANA and ssDNA will be investigated by physicochemical methods (Chapter 3, Section 1). In addition, a self complementary ANA dodecamer adapted from the pioneering work of Dickerson and Drew¹¹⁵ will be synthesized and studied for the potential formation of pure ANA:ANA duplexes (Chapter 3, Section 4). Since the Dickerson-Drew DNA duplex was one of the first to be studied at high resolution (x-ray diffraction), this is an excellent model to adopt and compare.

Two other studies will involve the synthesis of an ANA-DNA chimeric series of oligomers (18 nucleotides in length) complementary to the R region of the HIV-1 genome, in order to (a) investigate their ability to inhibit viral production (Chapter 4, Sections 1 and 2) and (b) to ascertain the relative roles of ANA in mediating RNase H cleavage (Chapter 4, Section 3). In the first study, oligodeoxynucleotides containing a single arabinonucleoside at or near the 3'-end will be prepared (*e.g.*, 5'DNA-araC3') and targeted to genomic HIV-1 RNA. The idea is to investigate whether the chimera will inhibit polymerization by HIV reverse transcriptase (RT). *i.e.* by a steric blockage mechanism, and then to explore the extent of the sensitivity of RT to subtle changes in the primer strand of the DNA-araC:RNA duplex on such polymerization.

1.4.2 Structural Studies

Triplexes: As part of a long term interest in structure-function correlation of branched oligonucleotides, considerable effort will focus on reverse-Hoogsteen T*A:T triplexes formed by oligonucleotides containing arabinoadenosine as a branch-point (Chapter 3, Section 2). A vast repertoire of synthetic branched nucleic acids (both in terms of composition and size) have already been synthesized and thoroughly investigated in the Damha laboratory, but these were based on ribose rather than arabinose branch-points (**Figure 1.10**).¹¹⁶

(A) Previous work (Hudson, 1995) ¹¹⁶



(B) This work:

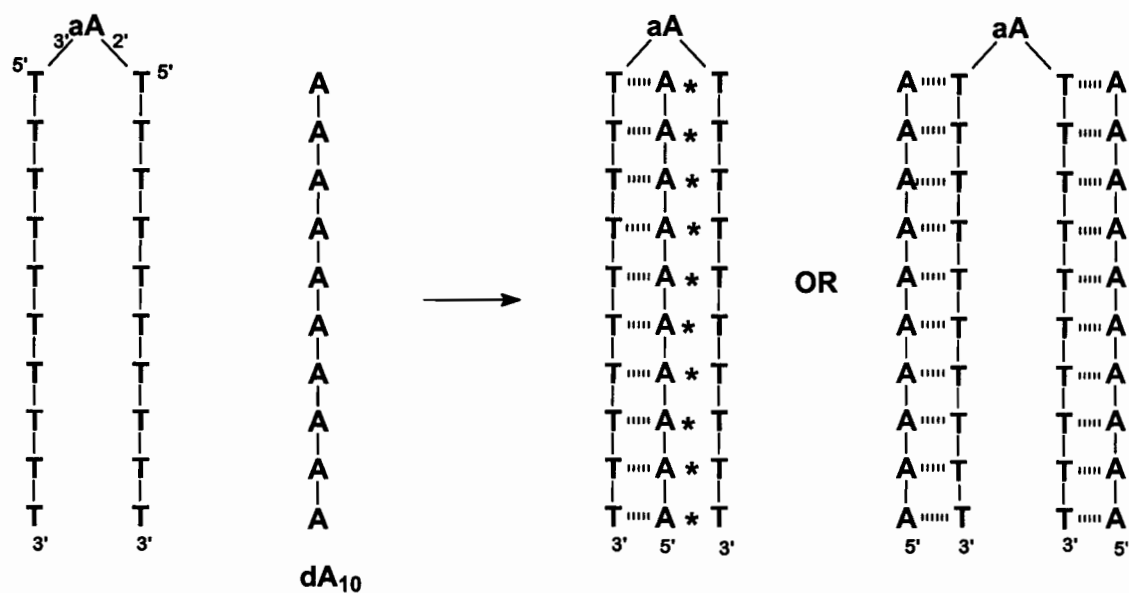


Figure 1.10: Complexes formed by branched V shaped molecules with deoxyadenylate third strands. (""") Watson-Crick bonds and (*) reverse-Hoogsteen bonds. [The following symbols will be used throughout this thesis for base pairing: Watson-Crick (:), Hoogsteen (•) and reverse-Hoogsteen (*)]. (A was studied by Hudson *et al.*) ¹¹⁶

Specifically, a “V”-shaped 21-mer, containing a single arabinoadenosine branch-point esterified at the 2' and 3' positions with decathymidilic acid will be synthesized, and its association with complementary DNA and RNA investigated. An examination of melting temperatures (T_m) for complex dissociation and nucleobase mismatch tolerance on hybridization will be measured in order to reveal fundamental physicochemical properties of branched oligonucleotides. This “branched strategy” for recognition of nucleic acids may be important in antisense applications that target viral ssDNA or RNA.

Finally the formation of triplexes by a polypyrimidine ANA third strand with duplex DNA, duplex RNA and DNA:RNA hybrids will be investigated for the first time (Chapter 3, section 3). “Will the chemically more stable ANA (the 2'-epimer of RNA), bind to duplex DNA?” These studies are important since RNA third strands are known to bind with high affinity to duplexes, but their short half-life *in vivo* precludes their use as antigene agents.

CHAPTER II SYNTHESIS AND CHARACTERIZATION OF OLIGOARABINO-NUCLEOTIDES OF MIXED BASE COMPOSITION

2.1 BACKGROUND ON ARABINONUCLEOSIDES AND ARABINONUCLEIC ACIDS

2.1.1 Arabinonucleosides

Arabinonucleosides are stereoisomers of the naturally occurring ribonucleosides, differing only in the configuration at the 2'-position of the sugar ring. This simple alteration of the sugar ring leads to nucleoside analogues with diverse biological activities, including anticancer^{107,117} and antiviral activities¹⁰⁸ (**Figure 2.1**).

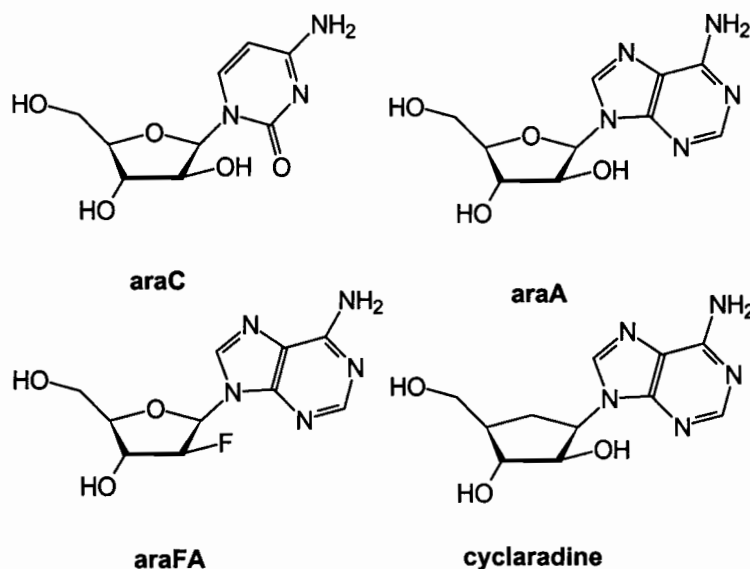


Figure 2.1: Arabinonucleosides used as anticancer and antiviral drugs.

To date, 1-β-D arabinofuranosylcytosine (araC) is the most successful nucleoside anticancer agent¹¹⁸ and is widely used either in combination therapy, or at high doses as a single agent to treat patients with leukemias (acute nonlymphocytic and human myeloblastic).¹¹⁹ Amphidicolin and araC elicit synergistic inhibition of DNA repair

synthesis in human fibroblasts, possibly due to the inhibition of DNA repair synthesis of pre-existing DNA damage caused by amphotericin.¹²⁰ AraC has only weak antiviral activity¹²¹ and its usefulness is limited by its *in vivo* deamination to inactive arabinofuranosyluracil (araU).

9- β -D-arabinofuranosyladenine (araA) and its 5'-monophosphate derivative¹²² have substantial antiviral activity towards Herpes Simplex Virus (HSV) and hematological malignancies.¹²³ Large doses of araA are needed however because of its conversion by adenosine deaminase to the less active 9- β -D-arabinofuranosylhypoxanthine (araH). AraA is therefore more effective against HSV infection when externally applied as an ointment. Adenosine analogues including 9- β -D-arabinofuranosyl-2-fluoroadenine¹²⁴ and cyclaridine¹²⁵ have been developed to circumvent the deamination of the adenine ring (**Figure 2.1**).

In order to achieve a therapeutic effect, it is necessary for these drugs to be phosphorylated by nucleoside kinases once inside the cell.¹²⁶ The difference in substrate specificity between cellular and viral induced nucleoside kinases, serves as a basis for the selectivity of several clinically used antiviral nucleoside analogues. Interestingly, carbocyclic-dG, araG, araH and araA are all substrates for the enzyme 2'-deoxyguanosine kinase, whereas 3'-modified dG analogues such as 3'-fluoro-2',3'-dideoxyinosine and 3'-azido-2',3'-dideoxy-2,6-diaminopurine riboside are not.¹²⁶ AraC is converted by a series of kinases to the active metabolite araCTP.¹²⁷ AraCTP and araATP are known to be incorporated into elongating DNA strands¹²⁸ and to work as inhibitors of DNA polymerase α , thus inhibiting DNA synthesis. Arabinofuranosyl nucleosides could also induce apoptosis (programmed cell death) in different cell lines such as human myeloid leukemia cell lines¹²⁹ without being incorporated into DNA. Thus at the cellular level, the major biochemical consequence of arabinonucleosides in treatment is suppression of replication¹³⁰ and of DNA repair.¹³¹

Recently 2'-O-allyl arabinofuranosides have been investigated with respect to their biological importance in the inhibition of ribonuclease reductase.¹³² The 2'-O-allyl araC derivative is approximately three orders of magnitude less cytotoxic than the parent compound araC, possibly due to a reduced phosphorylation efficiency by nucleoside kinases and/or a decreased affinity of the metabolite for its target enzyme.

While some arabinonucleosides have significant antitumor activity, their use is limited by the fact that they are equally toxic to rapidly proliferating normal cells. In this regard, it is worth noting that araG is selectively cytotoxic to T-lymphoblasts *in vitro* compared to B-lymphoblasts and null cells.¹³³ As such, araG is an attractive candidate as a selectively acting chemotherapeutic agent.^{134,135,136} The decreased catabolism of araG in T cells relative to B cells results in selective accumulation of the nucleoside¹³⁷ and consequently greater exposure of such cells to the activated araGTP, ultimately inhibiting DNA synthesis.^{134,136}

AraG also serves as a substrate for a number of other enzymes. For example it is phosphorylated by deoxycytidine kinase¹³⁸ in both T and B cells.^{135,139} It is also a substrate of mitochondrial deoxyguanosine kinase¹⁴⁰ and an inhibitor of both mitochondrial purine nucleoside phosphorylase¹⁴¹ and *Giardia lamblia* DNA synthesis.¹⁴² It has proven to be an effective *ex vivo* agent for purging residual malignant T cells from human bone marrow.¹³³

Although araG has been studied for over a decade as a selective T-cell cytotoxic agent, no clinical trials of araG have yet been initiated. The two greatest clinical deterrents are its poor water solubility and its elaborate chemical synthesis. The former concern was addressed by the development of 2-amino-6-methoxypurine arabinoside,¹⁴³ a water-soluble “pro-drug” that is converted to araG by adenosine deaminase whereas the issue of chemical synthesis was somewhat eased by the development of chemo-enzymatic synthetic methods.¹⁴⁴

2.1.2 Oligoarabinonucleotides

The above data clearly demonstrate that arabinonucleosides are valuable and unique agents that merit further investigation. In this regard, oligonucleotides constructed from arabinonucleotides have been under investigation from various different aspects.¹⁴⁵ Incorporation of these arabinonucleosides into a nucleotide chain¹⁴⁶ have been investigated in an attempt to improve the solubility of arabinonucleoside therapeutics. Specifically incorporation of araC into oligonucleotide strands has also been the focus of research to understand the mechanism of action of this anticancer drug.^{147,148} For example, DNA strands containing arabinocytosine have been a subject of

a number of structural studies. In the crystal, DNA duplexes containing araC adopt a normal B-type double helix with only small conformational perturbations at the araC-dG base pair.^{149,150} The association properties of uniformly modified oligoarabinonucleotides have been investigated in this research group¹⁰⁹ and independently by Pfeleiderer and coworkers.¹⁵¹ Pfeleiderer *et al.* have described the chemical synthesis of a 'transfer Arabino Nucleic Acid (ANA)', and several other groups have reported on the hybridization properties of DNA:DNA and DNA:RNA duplexes incorporating at most a single β -D-arabinonucleotide residue.^{148,152}

Oligomers constructed solely from α -arabinofuranosylthymine (α -ara-T) exhibited a large decrease in melting temperature toward complement DNA relative to control DNA (β -dT).¹⁵³ Duplexes formed by the association of either α -araT₁₅ or dT₁₅ with complementary RNA (poly-rA) were found to be of similar stability. More recently, Wengel and co-workers reported the synthesis and association properties of DNA oligomers containing one and two β -2'-OMe-araT inserts.¹¹³ These oligomers showed moderately lower thermal stabilities towards both ssDNA and RNA, compared to unmodified DNA controls. The same authors reported that oligomers constructed from α -2'-OMe araT units exhibited increased affinity towards a riboadenylate target compared to normal DNA controls. However the α -2'-OMe araT strand did not display any additional advantage relative to the known α -dT oligomers.

Aside from all the above mentioned studies, oligoarabinonucleotides containing the four natural nucleobases have not been extensively studied in biological and antisense fields, probably due to the problems of differential chemical protection of arabinose's hydroxyl groups.^{22b,154} The selective protection of different OH functions is one of the crucial problems in the synthesis involving nucleosides and other polyfunctional compounds, such as DNA and particularly RNA.

There exist a few reported strategies for the chemical synthesis of arabinonucleotides which were then incorporated into oligomers. The first and original work of Wechter¹⁴⁶ involved coupling of arabinonucleotides without any 2' hydroxyl protection, thus resulting in the formation of mixed 2'-5' and 3'-5' linked nucleotides (**Figure 2.2a**).

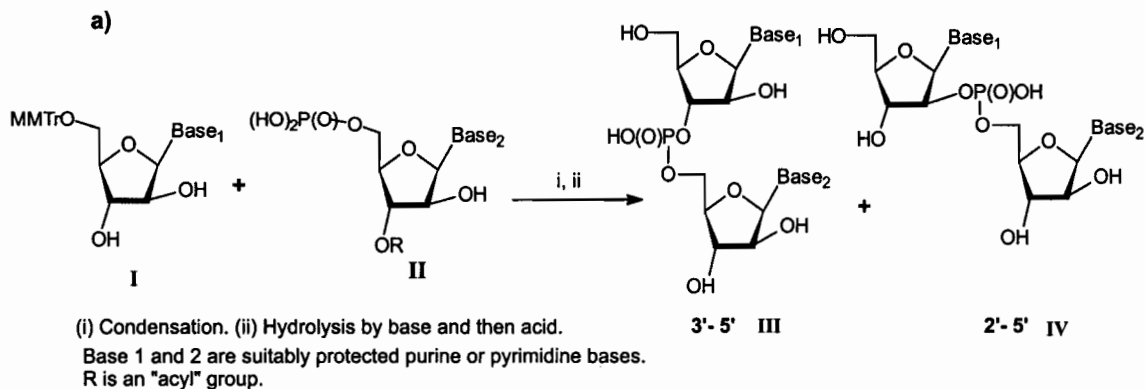


Figure 2.2a: First chemical synthesis of oligoarabinonucleotides.¹⁴⁶

Another strategy makes use of 5',2'-O-protected arabinonucleosides as precursors, which upon phosphorylation of the remaining 3'-OH function afford suitable building blocks for the oligoarabinonucleotides.^{155,156} The problem with this approach is the selective protection of the 2'-OH function which is on the sterically more hindered β face of the sugar. In practice this involves the use of a Markiewicz's bifunctional silylating reagent (1,3-dichloro-1,1',3,3'-tetraisopropylidisiloxane)¹⁵⁷ to temporarily protect the 3' and 5' positions and followed by protection of the 2'-OH group. Removal of the disiloxane function, followed by 5'-tritylation and 3' phosphitylation generates the completely functional arabinonucleoside building block for automated synthesis.¹⁵⁶ It was possible to introduce both the 2' and the N⁴ protecting groups in a single step leading to an overall simplified synthesis of the araC.¹⁵⁵ A third synthetic method, involving the formation of O²-2'-anhydronucleoside synthons, is limited only to pyrimidine arabinonucleosides (**Figure 2.2.b**).¹⁵⁸

A fairly recent strategy introduced by the Damha research group involved the sequential and regioselective tritylation (at 5') after the usual base protection, followed by phosphitylation (at 3'), and acetylation (at 2') of arabinonucleosides.¹⁰⁹ Thus, by taking advantage of the different reactivities of the hydroxyl groups (5'OH > 3'OH > 2'OH) this method significantly minimized protection steps of arabinonucleoside units and afforded only the 3'-5' linked oligonucleotide in good yields.¹⁵⁹ It is worth noting that acyl protecting groups (acetyl and benzoyl) could be employed for the arabinose 2'-OH group because its removal at the end of the chain assembly does not cause internucleotide cleavage or 3'-5' to 2'-5'-isomerization of the phosphodiester moiety. These problems are observed when 2'-O-

acylated RNA chains are deblocked under the same conditions, since the 2'-OH now has the right orientation (stereochemistry) to effect the internucleotide cleavage.¹⁵⁶

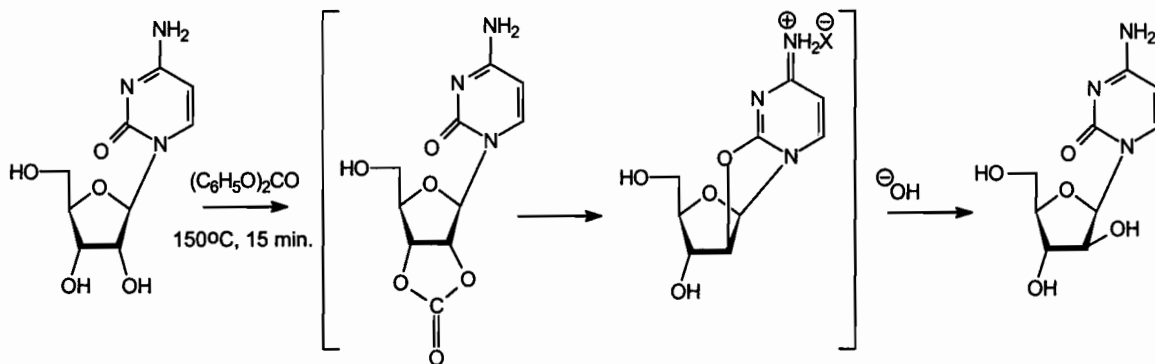


Figure 2.2.b: Formation of O²-2'-anhydronucleoside synthons limited only to pyrimidine arabinonucleosides.¹⁵⁸

When this project was initiated, oligoarabinonucleotides containing all four nucleobases had not been thoroughly investigated due in part to the difficulties in the synthesis of the 9-β-D-arabinofuranosylguanine (araG). This chapter will first describe a new method for the synthesis of araG. Oligoarabinonucleotides of mixed base composition will then be prepared and characterized.

2.1.3 Chemical Synthesis of Arabinoguanosine

Arabinoguanosine was first synthesized by Reist and Goodman in 1964, by the initial condensation of 2,6-dichloropurine with xylofuranose tetraacetate followed by its transformation into araG (**Figure 2.3a**).¹⁶⁰ Since then there have been several attempts to develop practical syntheses of araG by glycosylation of a purine base with the appropriately protected ara-sugar moiety.¹⁶¹ Unfortunately this approach often leads to unseparable mixtures of α/β-anomers.

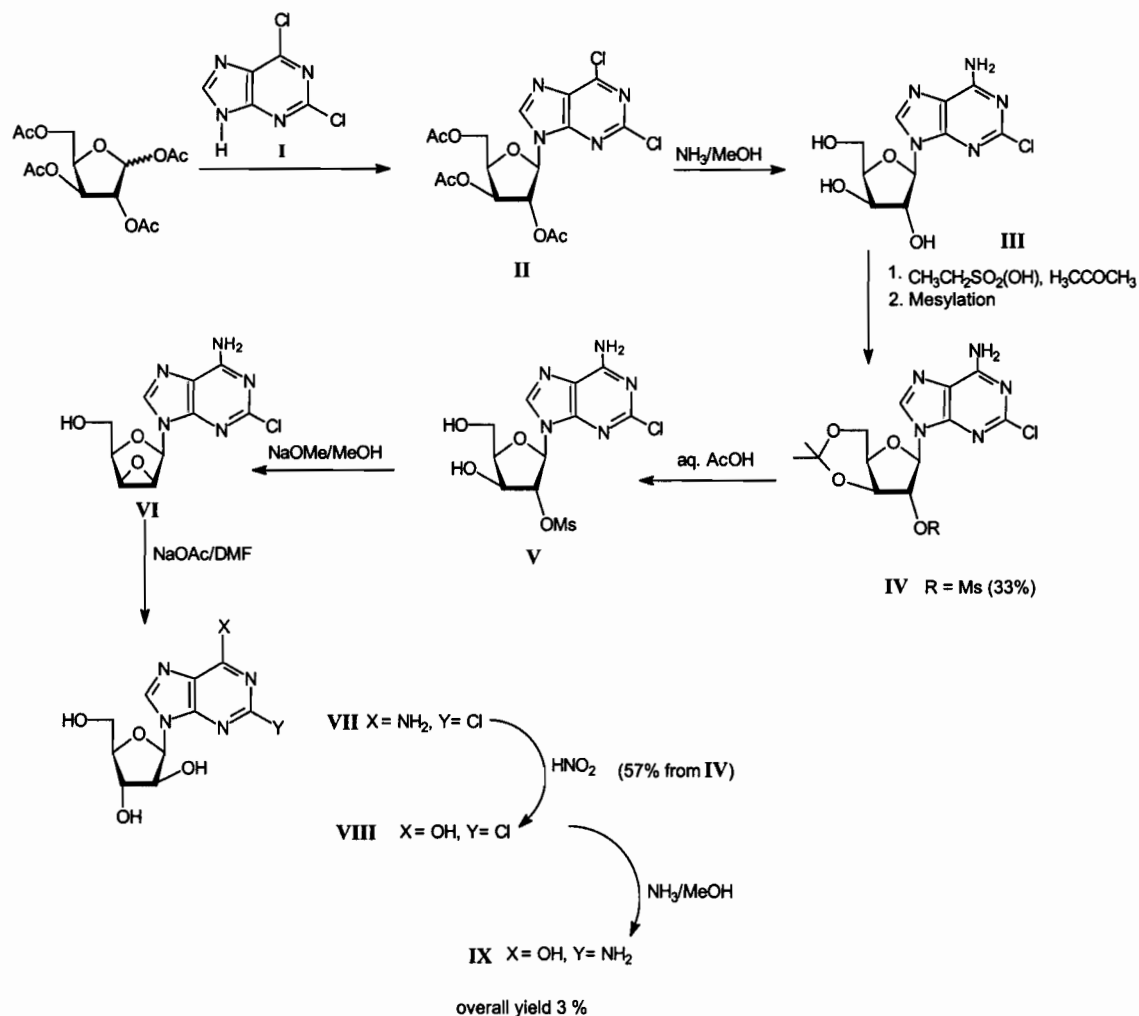


Figure 2.3a: Synthesis of arabinoguanosine *via* the condensation of 2,6-dichloropurine with xylofuranose tetraacetate.¹⁶⁰

In another approach described by Chattopadhyaya and Reese,¹⁶² guanosine is converted into araG *via* the 8,2'-anhydroguanosine intermediate (**Figure 2.3b**).¹⁶³

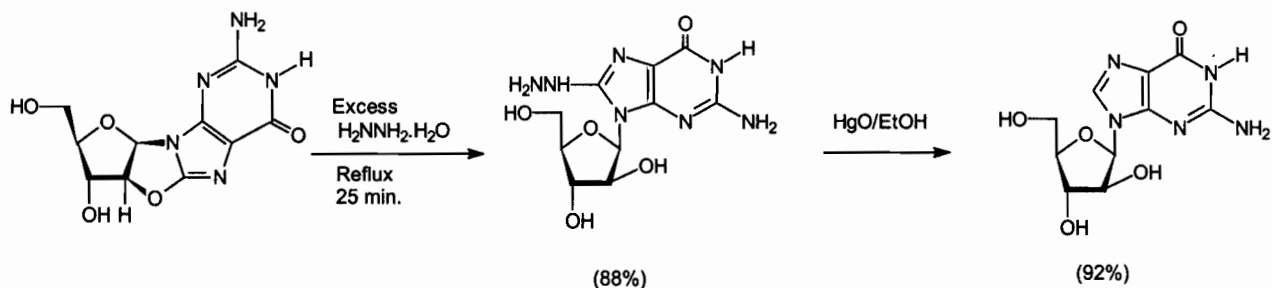


Figure 2.3b: Synthesis of arabinoguanosine *via* the 8,2'-anhydroguanosine intermediate.¹⁶²

Although this offers an attractive and reproducible route to araG, the overall yield from riboG is usually quite low. Hansske *et al.* have reported the conversion of guanosine to araG *via* a 2'-ketonucleoside;^{161c,164} however this method leads to unseparable mixtures of ara/ribonucleosides (**Figure 2.3c**).

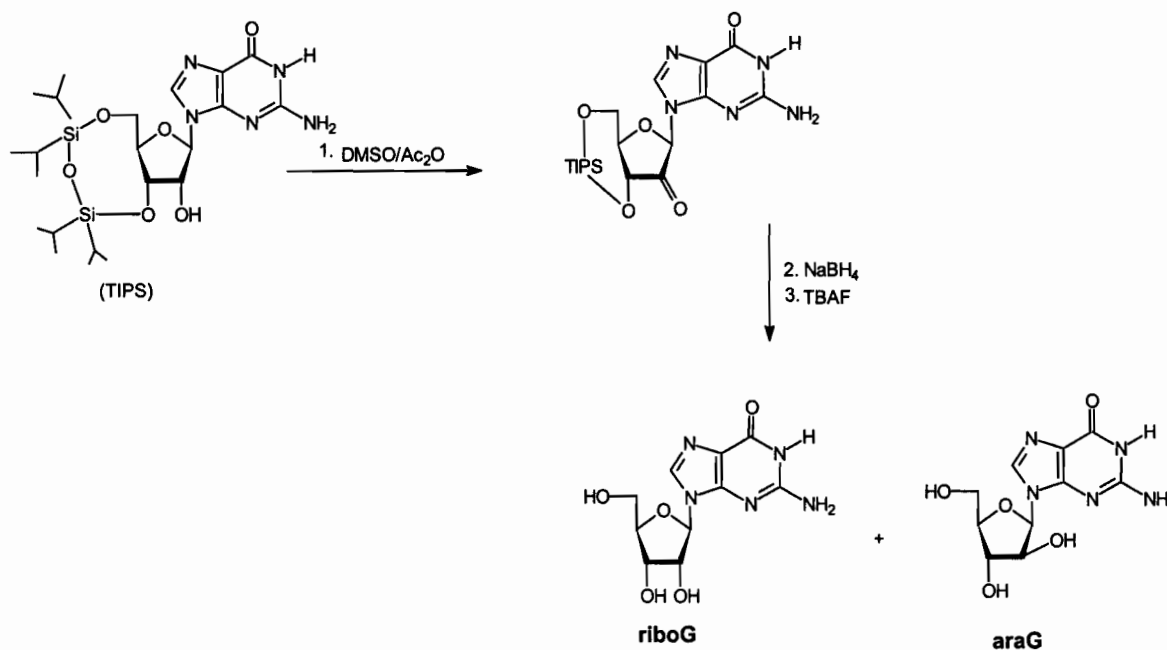
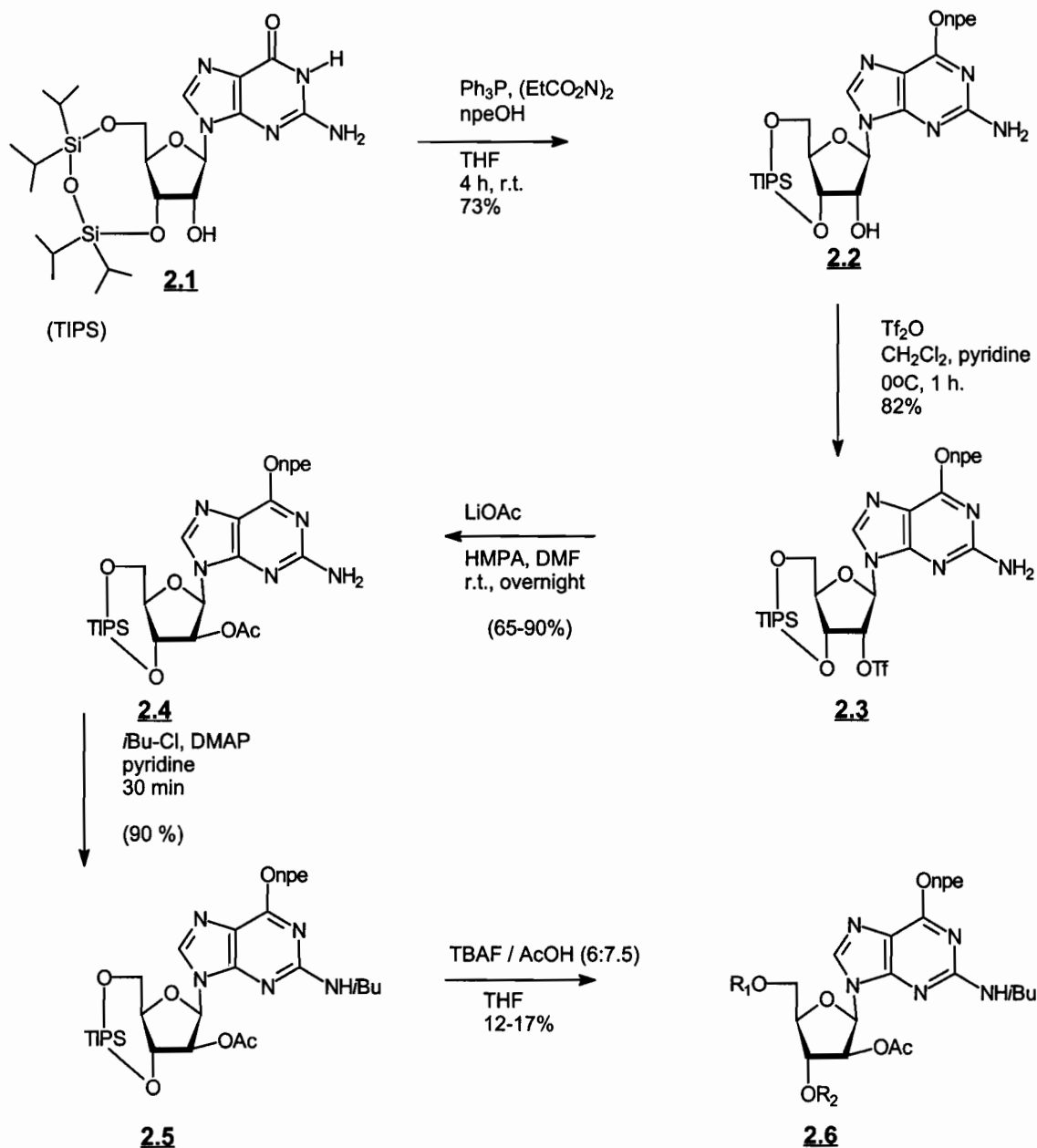


Figure 2.3c: Oxidation-reduction conversion of riboG to araG.¹⁶⁴

Enzymatically catalyzed transglycosylations¹⁶⁵ have been used to convert guanosine to the 2,6-diaminoarabinoside, followed by enzymatic deamination with adenosine deaminase to give the desired araG nucleoside.¹⁶⁶ There remains some definite practical limitations in all of the above mentioned approaches because of low-yielding, laborious and time-consuming steps.

Recently Resmini and Pfeleiderer reported a successful chemical conversion of guanosine to arabinoguanosine by triflating the 2'-OH and subsequent nucleophilic displacement with LiOAc,¹⁶⁷ as previously reported for the synthesis of araA.¹⁶⁸ This synthesis appeared attractive and thus was adopted in the current work with a few modifications. For example, it was desirable to synthesize a fully protected araG monomer that was compatible with the DNA and RNA synthesis strategy developed by Damha *et al.*¹⁶⁹ (**Figure 2.4**).



		R ₁	R ₂	yield
MMTrCl (1.7eq) pyridine, r.t. 24h	2.6	H	H	
CIP(OCE)NiPr ₂ DIPEA, THF, 0°C	2.7	MMTr	H	69%
	2.8	MMTr	P(OCE)NiPr ₂	64%

Figure 2.4: Synthesis of the *arabinoguanosine* phosphoramidite monomer for solid phase synthesis.

Thus the first step involved the simultaneous protection of the 3' and 5' hydroxyl groups with the Markiewicz's reagent (1,3-dichloro-1,1',3,3'-tetraisopropylidisiloxane) with some modifications to the original protocol.¹⁵¹ This step regiospecifically blocks the 5' and 3' OH groups and also affords a compound (**2.1**) that is less polar than the starting materials used by Resmini. As a result, silica gel chromatography is less tedious.

A long standing problem in nucleoside and oligonucleotide chemistry is dealing with side-reactions occurring at the O⁶-position of guanosine residues.^{170,171} Therefore, protection of the exocyclic amine and amide O⁶ function is an obvious necessity if these side reactions are to be eliminated. Thus protection of the O⁶ atom with the *para*-nitrophenylethyl group (npe) was accomplished by the Mitsunobu reaction^{172,173} with nitrophenylethyl alcohol. Both its introduction and removal are fairly straightforward. An important consideration in this step was the order in which the reagents were added. Triphenylphosphine and diethyl-azodicarboxylate were first mixed and the nucleoside **2.1** was added last to reduce potential reaction at the free 2'-OH group. After workup, triphenylphosphine oxide was removed by precipitation of the crude **2.2** in ether followed by silica gel column chromatography. The desired O⁶-npe protected nucleoside **2.2** was isolated in 73% yield. Visualization of the desired product by TLC using long wavelength UV (365 nm), and NMR, confirmed the presence of the npe group.

Conversion of **2.2** to the more reactive 2'-O-triflate derivative **2.3**, was accomplished with trifluoromethanesulphonic anhydride in the presence of pyridine and 4-(dimethylamino)pyridine (4-DMAP) in CH₂Cl₂. The crude product **2.3** was obtained in moderate yields (75-80%) and was generally used without purification in the next reaction.

Conversion of ribonucleoside **2.3** to the desired arabinonucleoside was accomplished by subsequent nucleophilic displacement of **2.3** with lithium acetate, in hexamethylphosphoramide (HMPA)/DMF solvent. Following work-up, **2.4** was obtained as an amorphous powder which inevitably turned to a gum on exposure to the atmosphere. This called for some modifications to the original protocol. The gummy residue was thus dissolved in ethyl acetate, with drying over MgSO₄, filtering and evaporation to give **2.4** as a sticky 'foam'. Yields varied between 65 and 90%.

Protection of the guanine exocyclic amino group with isobutyryl chloride (*i*BuCl) was postponed to a later stage, in order to avoid depurination during the triflation step. This side reaction appears to be favored for nucleosides with the npe-O⁶ and NH-*i*Bu protected base.¹⁶⁷ When isobutyrylation was attempted with isobutyric anhydride [(*i*Bu)₂O] instead, no reaction occurred even after 3 days, 6 additional equivalents of (*i*Bu)₂O and 4-DMAP as catalyst. It is possible that the steric bulk and the electron withdrawing ability of the O⁶-npe group causes the NH₂ group to be less reactive towards the anhydride, which itself is a larger and less reactive acylating agent than *i*BuCl. Interestingly in riboG synthesis, the isobutyryl function can be introduced *via* (*i*Bu)₂O prior to the Mitsunobu reaction step, without any complications.¹⁷⁴

Cleavage of the disiloxane (TIPS) group in **2.5** was achieved with 1.0 M tetrabutylammonium fluoride/THF/acetic acid (TBAF(THF)/AcOH/nucleoside) in the ratio 6:7.5:1 (molar equivalence). This afforded **2.6** in only 12-30% yields, which was significantly less than those reported for analogous reactions (*ca.* 70%).¹⁶⁷ The dramatic loss of material in the current approach could be explained by the high polarity of **2.6** which makes it “stick” to silica gel during purification by column chromatography. It should be noted that the npe group was not affected by excess of F⁻, under the conditions used (as ascertained by mass spectrometry and NMR). The next step entailed selective tritylation of the 5'-OH^{169,175} with MMTrCl in pyridine at r.t. to give **2.7** in moderate yields (69-73%).

The completely protected monomer was obtained by reacting **2.7** in the presence of 1.1 equivalents of N,N-diisopropyl-β-cyanoethylphosphoramidic chloride and N,N-diisopropylethylamine (DIPEA) in THF as described for ribomonomers.¹⁷⁶ The overall yield of the completely protected araG monomer was 19-22% (from guanosine and without carrying out the purification step of compound **2.6**).

2.2 CHARACTERIZATION OF ARABINOGUANOSINE NUCLEOSIDE DERIVATIVES

AraG derivatives were characterized by Fast Atom Bombardment Mass Spectrometry (FAB MS) and Nuclear Magnetic Resonance Spectroscopy (NMR). The NMR techniques used were one-dimensional proton (^1H), two-dimensional Correlated Spectroscopy (^1H -COSY) and Heteronuclear Multiple Quantum Correlated Spectroscopy (HMQC).

Usually ^1H -NMR allows for the specific identification of ribo- and arabinonucleosides because of certain characteristic signals that result from different structural features of the protecting groups used. **Figure 2.5** shows the proton NMR spectra of ribo and arabino compounds **2.3** and **2.4**, respectively. The insets show the anomeric proton resonances with their respective coupling constants $J_{1,2'}$. The chemical shifts are summarized in **Tables 2.1** and **2.2** respectively. It is clear that the introduction of the (trifluoromethyl)sulphonyl group (**2.3**) causes a typical downfield shift (5.56 ppm) of H-C(2') relative to that caused by acetyl group in **2.4** (5.52 ppm) or the free 2'OH (4.78 ppm). Furthermore, a typical signal down-field shift of the anomeric proton is seen upon conversion of **2.3** (6.08 ppm) to **2.4** (6.32 ppm). Coupling constants ($J_{1,2'}$) for the anomeric protons in the disiloxane protected nucleoside, as shown in each spectrum are typical for ribo ($J_{1,2'} = 0\text{--}1.5\text{ Hz}$) and ara ($J_{1,2'} = 6\text{--}6.5\text{ Hz}$) nucleosides. These results are in accordance with earlier observations made by Robins¹⁷⁷ *et al.* and Resmini *et al.*¹⁶⁷ and suggest that the furanose ring adopts primarily the C3'-endo conformation, as shown by the structures given in **Figures 2.5 A and B**.

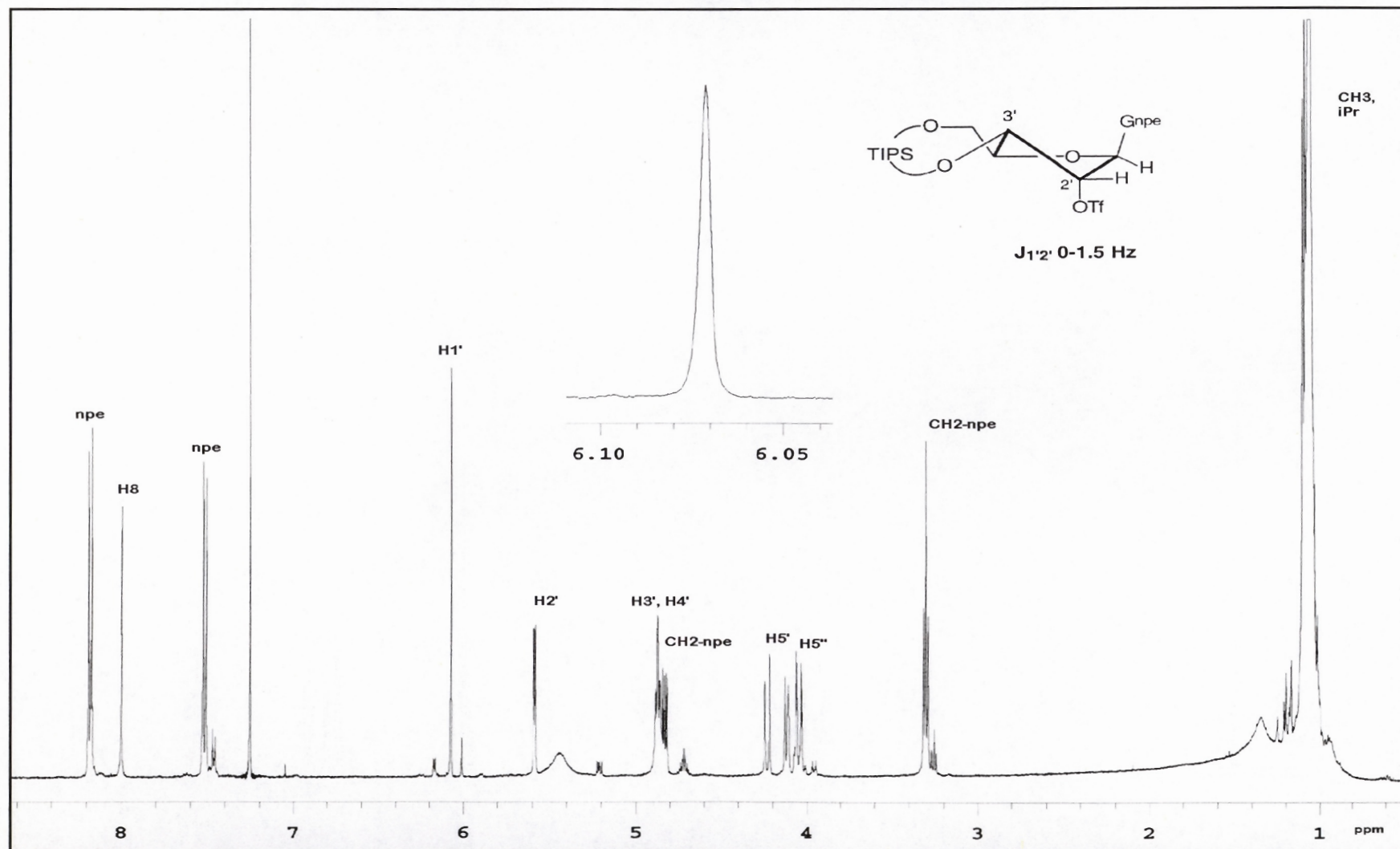


Figure 2.5 A: ^1H -NMR (500 MHz) of O^6 -[2'-(4-Nitrophenyl)ethyl]-9-{3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)-2'-O-[(trifluoromethyl)sulfonyl]- β -D-ribofuranosyl} guanine (**2.3**) in CDCl_3 .

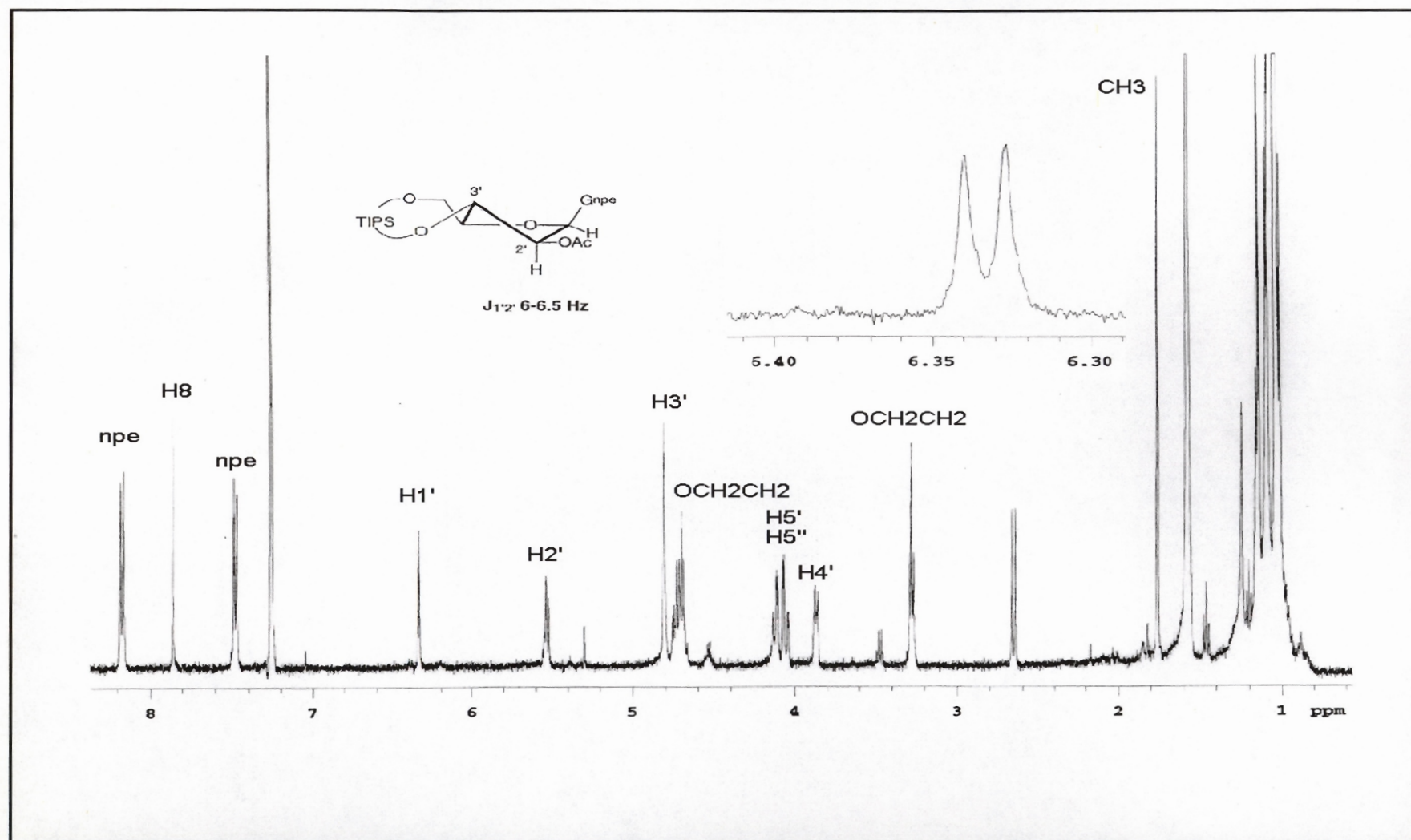


Figure 2.5 B: ^1H -NMR (500 MHz) of O^6 -[2-(4-Nitrophenyl)ethyl]-9-{2'-O-acetyl-3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)} β -D-arabinofuranosyl}- guanine (**2.4**) in CDCl_3 .

Table 2.1: ^1H and ^{13}C -NMR Assignments of O^6 -[2-(4-Nitrophenyl)Ethyl]-9-{3',5'-O-(1,1',3,3'-Tetraisopropyldisiloxane-1,3-Diyl)-2'-O-[(Trifluoromethyl)Sulfonyl]- β -D-Ribofuranosyl}Guanine (2.3) via HMQC Experiments

^1H (δ ppm)	split	^1H assignment	^{13}C (δ ppm)
8.17	d	2H <i>o</i> to NO_2	138.3
8.09	s	H-C(8)	123.9
7.53	d	2H <i>m</i> to NO_2	130.2
6.10	s	H1' J= 0Hz	86.7
5.56	d	H2'	87.0
4.93	m	H3'	69.8
4.80	t	OCH_2CH_2	81.7
4.26	d	H5'	59.1
4.16	m	H4'	67.3
4.02	dd	H5''	59.1
3.31	t	OCH_2CH_2	34.7
1.02	m	Me from <i>i</i> -Pr	12.5, 16.5

Table 2.2: ^1H and ^{13}C -NMR Assignments of 9-[2'-O-Acetyl-3',5'-O-(1,1',3,3'-Tetraisopropyldisiloxane-1,3-Diyl)- β -D-Arabinofuranosyl] O^6 -[2-(4-Nitrophenyl)Ethyl]Guanine(2.4) via HMQC Experiments.

^1H (δ ppm)	split	^1H assignment	^{13}C (δ ppm)
8.16	d	2H <i>o</i> to NO_2	137.8
7.87	s	H-C(8)	123.2
7.43	d	2H <i>m</i> to NO_2	129.8
6.32	d	H1' J= 6.0 Hz	80.0
5.52	m	H2'	76.0
4.70	m	H3', OCH_2CH_2	65.5, 71.0
4.05-4.18	m	H5, H5''	60.0
3.87	m	H4'	80.0
3.27	dd	OCH_2CH_2	34.5
1.76	s	OCH_3	20
1.02	m	Me from <i>i</i> -Pr	17

The other three required arabinonucleoside derivatives, namely, 5'-monomethoxytrityl-2'-OAc-3'-O-(β -cyanoethyl-N,N-diisopropylphosphoramidite) of ara-A (N⁶-Bz), araC (N⁴-Bz) and araU were prepared by variations of published procedures.^{109,169}

2.3 SYNTHESIS OF OLIGOMERS OF ARABINONUCLEIC ACIDS (ANA)

Oligoarabinonucleotides were readily prepared by the conventional solid-phase methodology employing the phosphoramidite chemistry described in Chapter 1. A number of oligoarabinonucleotides were synthesized in order to investigate the influence of the arabinose sugar-backbone in different studies that will be described in detail in Chapters 3 and 4. The oligomers were prepared on a 1 μ mole scale, using an automated DNA synthesizer,¹⁷⁸ and controlled-pore glass (CPG) as the solid support. CPG consists of a uniformly prepared glass matrix with pores of defined size and is the preferred solid support in oligonucleotide synthesis. The preparation begins with the attachment of the 3'-hydroxyl group of the first nucleoside to the solid support as previously published.¹⁷⁹ Typical arabinonucleoside loadings on CPG were 20-40 μ mole nucleoside/g. The amidites were introduced as 0.11M ara (C, A, U) or 0.15 M araG concentrations in acetonitrile. Prior to chain assembly, the support was treated with the capping reagents, namely acetic anhydride/N-methylimidazole/4-N,N-dimethylaminopyrimidine. The coupling efficiencies, based on the yield of the monomethoxytrityl cation released after each coupling were consistent with that observed for ribonucleotide synthesis and varied between 79 and 109%. After deprotection with aqueous ammonia/ethanol (3:1, r.t. 24-48 h.), the oligomer was purified by preparative gel electrophoresis (PAGE) on denaturing gels (14% acrylamide/7M urea). Desalting by size exclusion chromatography (Sephadex G-25)TM afforded the purified oligomers in good yields. For oligomers containing araG units it was necessary to subject the partially protected oligomer to an additional step in order to cleave the *p*-nitrophenylethyl guanosine O⁶ protecting group. Thus following the ammonia treatment and evaporation step, the oligomer was treated with a solution of 1M TBAF/ THF (50 μ L, r.t. 16 h) prior to PAGE analysis. (For a detailed list of the yields obtained, refer to the Experimental Section, Chapter 6)

2.4 CHARACTERIZATION OF OLIGOMERS

Oligoarabinonucleotides (ANA) were analyzed by polyacrylamide gel electrophoresis (PAGE), UV spectroscopy and MALDI-TOF mass spectrometry. Although a large number of sequences were synthesized for this thesis work, only a few ANA oligomers (Table 2.3) will be discussed here to illustrate the characterization. A complete list of all sequences synthesized can be found in Chapter 6, Tables 6.1 and 6.2.

Table 2.3: Examples of arabinonucleic acids (ANA) and ssDNA and RNA controls synthesized

Sequence Number	Sequence Description	Associated Chapter
<u>3.1</u>	d(AGC TCC CAG GGT CAG ATC)	3.1
<u>3.14</u>	r(AGC UCC CAG GGU CAG AUC)	3.1
<u>3.15</u>	a(AGC UCC CAG GGU CAG AUC)	3.1
<u>3.27</u>	d(CCT CTC CTC CCT)	3.3
<u>3.29</u>	r(CCU CUC CUC CCU)	3.3
<u>3.31</u>	a(CCU CUC CUC CCU)	3.3
<u>3.32</u>	d(CGC GAA TTC GCG)	3.4
<u>3.33</u>	r(CGC GAA UUC GCG)	3.4
<u>3.34</u>	a(CGC GAA UUC GCG)	3.4

Polyacrylamide Gel Electrophoresis (PAGE): A first tool of analysis involved PAGE on denaturing gels. **Figure 2.6** shows a representative gel illustrating the relative mobilities of ANA and corresponding DNA and RNA single strands of same chain length. Generally the ANA oligomers were present as a single species, and the migration on the gels was comparable to that of unmodified sequences.

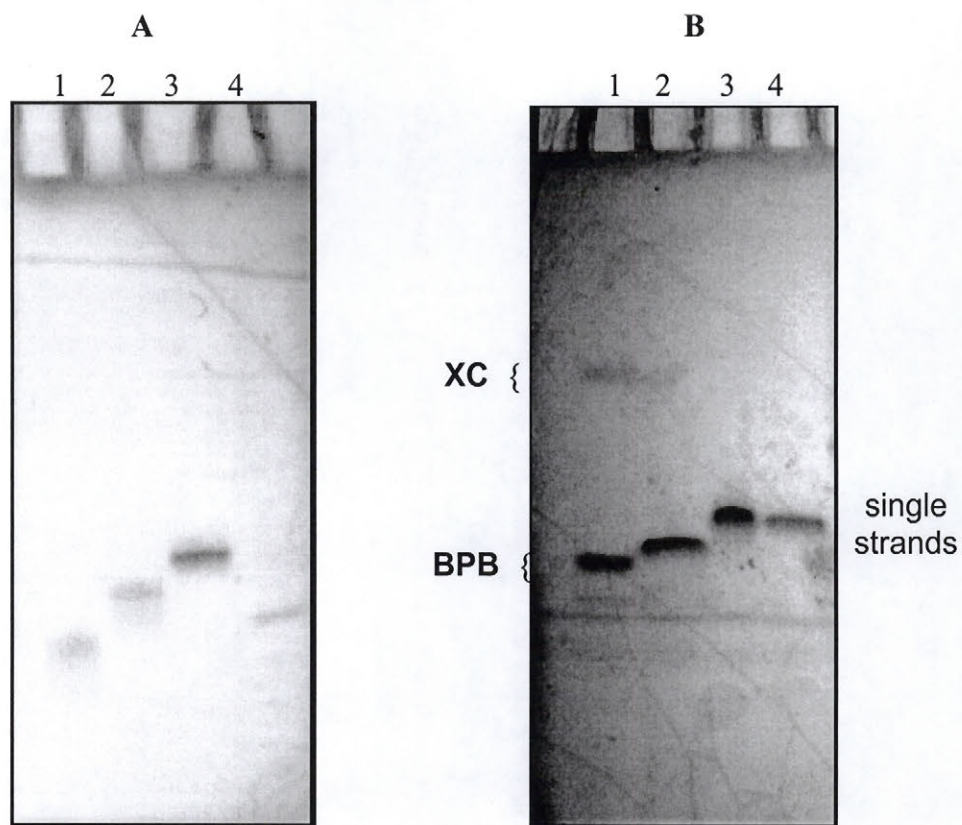


Figure 2.6: Polyacrylamide gel electrophoresis of oligonucleotides. Conditions, 20% polyacrylamide, 7 M Urea, pH 8.0 Bands were visualized by UV shadowing. **(A)** Lanes: (1) DNA, (2) RNA, (3) ANA, (4) marker dyes: xylene cyanol (XC) and bromophenol blue (BPB). Sequence: CGCGAAUUCGCG^{*}; **(B)** Lanes: (1) xylene cyanol (XC) and bromophenol blue (BPB), (2) DNA, (3) RNA & (4) ANA. Sequence: CCU CUC CUC CCU. ^{*} In DNA U is replaced by T.

MALDI-TOF mass spectrometry (MS): A third method of analysis involved molecular weight determination of the oligomers by MALDI-TOF MS. This is a recently introduced, mild ionization technique that allows desorption and ionization of very large molecules, and has been successfully used for mass determination of polypeptide and nucleic acids.¹⁸⁰

For MALDI-TOF-MS analysis, a small amount of the sample (analyte, *ca.* 0.5mM) is mixed uniformly with a crystalline matrix, typically a small organic species with a high extinction coefficient, such as 2,4,6-trihydroxyacetophenone or 6-aza-2-thiothymine in 20 mM ammonium citrate:acetonitrile. In an environment in which a large excess of the matrix is present, the sample molecules are surrounded by the matrix molecules in crystalline or amorphous states. Direct laser irradiation can supply the resonant excitation energy that is needed for ionization. The use of a matrix circumvents the problematic photodissociation of bonds, sometimes associated with the irradiations. The matrix, which has resonance absorption at the laser wavelength used (337 nm), absorbs the laser energy and causes rapid heating, resulting in expulsion and mild ionization of the sample molecules without fragmentation. In other words, the role of the matrix is to absorb and transfer laser energy to the analyte in a controllable efficient way.

The sample must be dilute to prevent complexation of the analyte. A laser is focused on a sample crystal, which is then ionized by short pulses of laser energy. After ionization, the sample is accelerated to a fixed kinetic energy by a potential difference. The ions then pass through a field-free region and a detector which records the time of flight. The time of flight of the ions is proportional to their velocity, which in turn is related to the mass of the ions, thus the name Matrix-Assisted Laser Desorption/Ionization Time of Flight MS.

The molecular weight of ANA sequences were thus confirmed by MALDI-TOF MS. These are reported in **Table 2.4** and were in agreement with the calculated values. A representative MALDI-TOF mass spectrum of oligomer **3.31** is shown in **Figure 2.7**.

Table 2.4: Observed and Experimentally Determined Molecular Weights of Oligoarabinonucleotides.

Oligomer	Sequence	Calculated (Expected) MW (g mol^{-1})	Observed MW (g mol^{-1}) $\pm 0.1\%$	Interpretation
3.15	a(AGC UCC CAG GGU CAG AUC)	5648	5653	M^+
3.19	aA ^{dT10} _{dT10}	6332	6357	$\text{M} + \text{Na}^+$
3.31	a(CCU CUC CUC CCU)	3602	3602	M^+

Samples contained *ca.* 0.2 mmol of oligomer in a matrix of 6-aza-2-thio thymine dissolved in 20 mM ammonium citrate:acetonitrile.

2.5 CONCLUSION

AraG was prepared using a combination of synthetic strategies developed in Pfeleiderer's and Damha's laboratories. The overall yield starting from guanosine to arabinoguanosine phosphoramidite was 19-22%. Completely modified oligoarabinonucleic acids were synthesized *via* solid phase synthesis according to known procedures. The molecular weights as measured by MALDI-TOF MS correspond with the expected values.

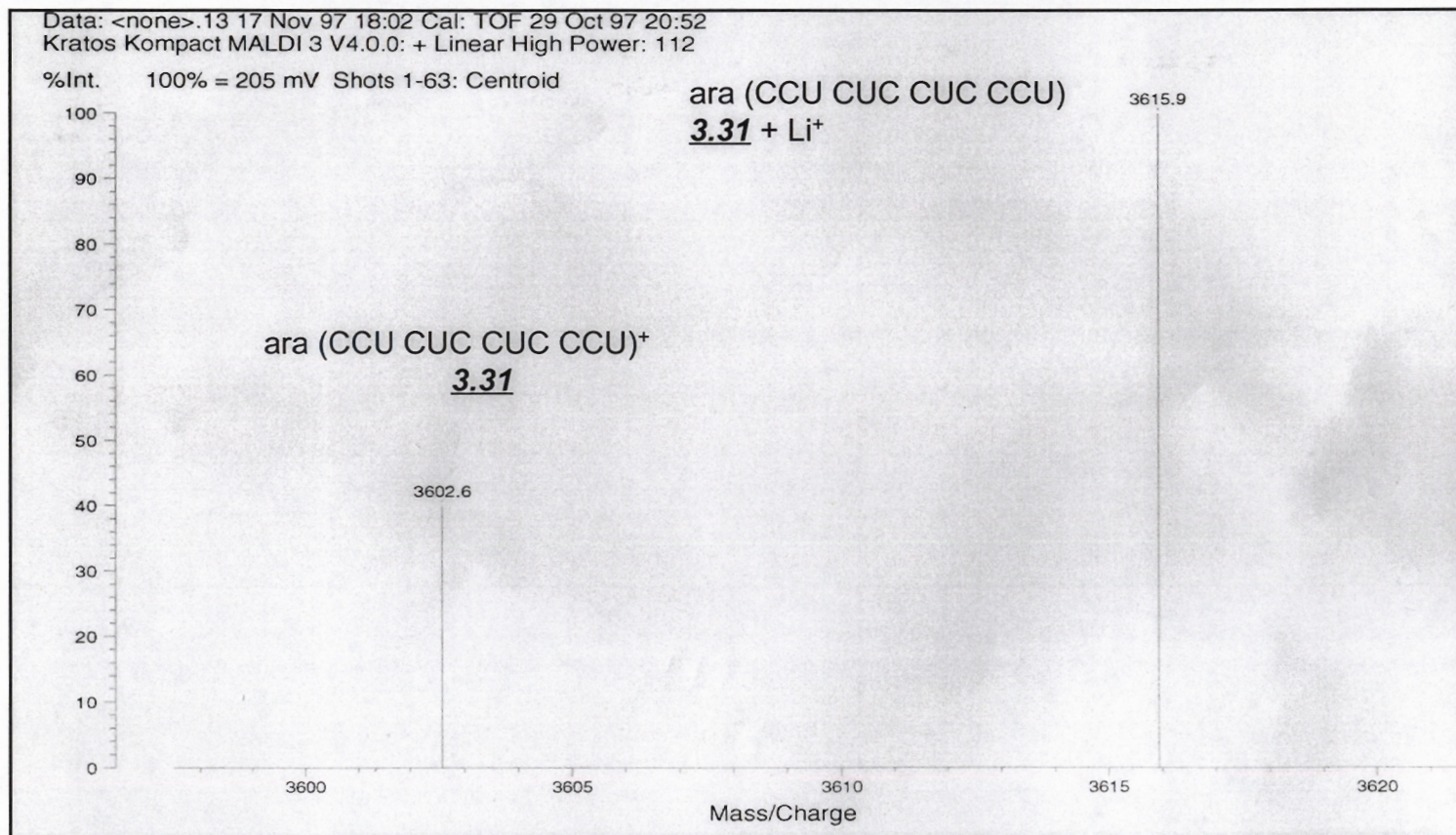


Figure 2.7: MALDI-TOF spectrum of ara(CCUCUC CUC CCU) 3.31 . Matrix :6-aza-2-thio thymine in 20 mM ammonium citrate/acetonitrile buffer.

CHAPTER III PHYSICOCHEMICAL STUDIES OF OLIGONUCLEOTIDES BASED ON D-ARABINOSE

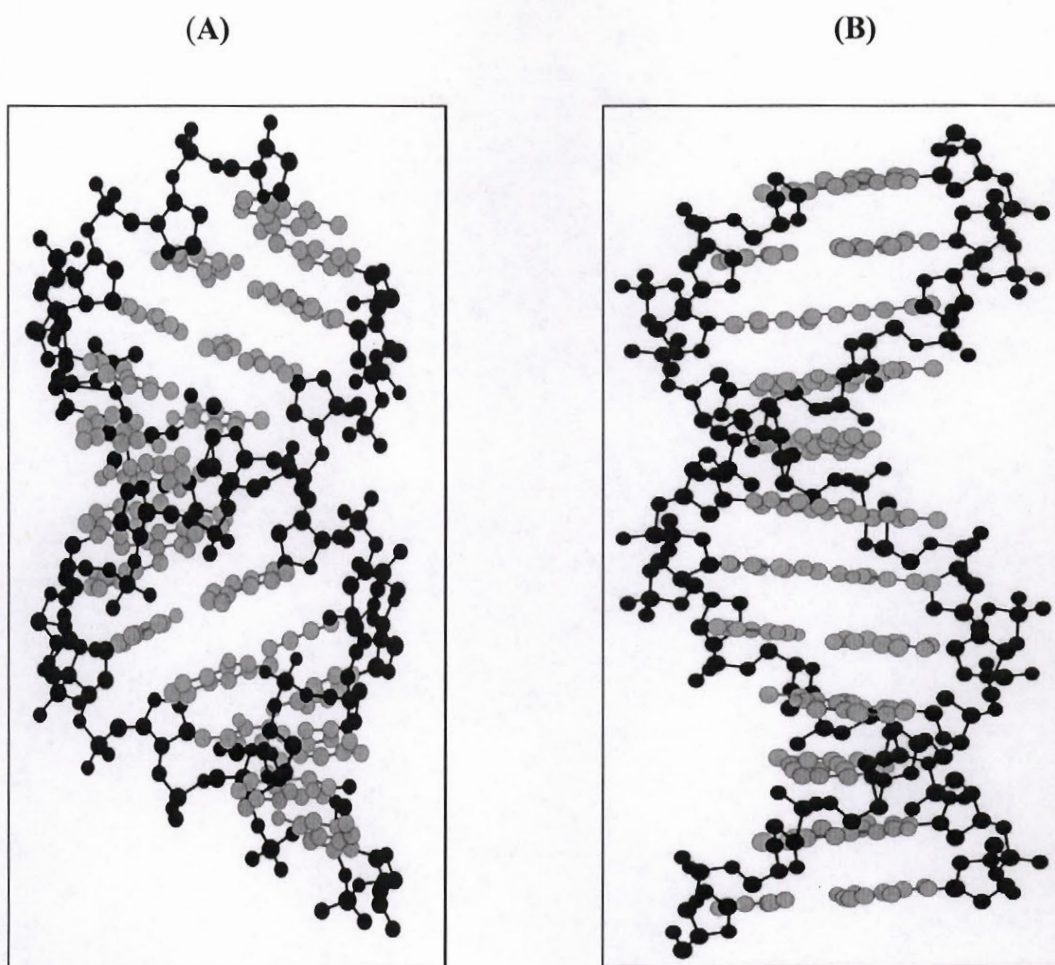
3.1 ROLE OF 2'-STEREOCHEMISTRY ON DUPLEX FORMATION

3.1.1 Introduction¹

DNA and RNA play different biological roles which are closely associated with the preferred conformations that they adopt. In general the biologically relevant conformation of DNA is the B genus right-handed double helix, while the RNA double helix adopts the A conformation.¹

In B-DNA, the bases stack predominantly above their neighbors in the same strand and are perpendicular to the helical axis (**Figure 3.1.1**). The helix has major and minor grooves of different depth. Although the major groove displays more distinct structural features, the minor groove is believed to be the locus of action of a larger number of groove binding drugs. *e.g.* Hoechst 33258 and Netropsin. Enzymes are known to bind in both grooves for example, RNA polymerases bind to the major groove whereas DNase I binds in the minor groove. The broad major groove is coated by a unimolecular layer of water molecules which interacts with the exposed C=O, N, and NH functions of the bases and extensively solvates the phosphate backbone. More significantly, the narrow minor grooves contain well-ordered, zigzag chains of two water molecules per base-pair. Hence high humidity favors the B-form over the A-form and as such B-form DNA is prevalent under physiological conditions.¹

A-DNA, like B-DNA, is a right-handed double helix made up of antiparallel strands held together by Watson-Crick base pairing. As shown in **Figure 3.1.1** the A-helix is wider and shorter than the B-helix. The adjacent phosphates in the same chain are a little further apart in B-DNA (P-P = 6.7 Å) than in A-DNA (P-P = 5.4 Å). Moreover the minor groove of A-DNA is so shallow that it nearly vanishes. Thus compared to a B-form duplex, the minor groove of an A-helix receives little stabilization through hydration or magnesium ions, and as a consequence, dehydration favors the A-form.

**Characteristics:**

Helical Sense	Right	Right
Residues /turn	11	10.4
Twist /bp	32.7°	36°
Displacement bp	4.5 Å	-0.2 to - 1.8 Å
Rise/bp	2.56 Å	3.3-3.4 Å
Base Tilt	20°	-6°
Sugar Pucker	C3'-endo	C2'-endo
Groove (Minor)	broad and shallow	narrow and deep
(Major)	narrow and deep	wide and deep
Helix Diameter	25.5 Å	23.7 Å

Figure 3.1.1: (A) A-form and (B) B-form duplexes adopted by nucleic acids along with their average parameters. Adapted from Blackburn, G. and Gait, M. 1996.²

Many of the structural differences between the two types of helices arise from different puckerings of the sugar units (**Figure 3.1.2**). The furanose rings are twisted out of plane (“puckered”) in order to minimize non-bonded eclipsing interactions (steric and torsional strain) between their substituents. This “puckering” is described by identifying the major displacement of 2' and 3' carbons from the median plane of C1'-O4'-C4'. Thus, if the *endo* displacement of C2' is greater than the *exo* displacement of C3', the conformation is called C2'-endo and *vice versa* (the term “endo” refers to the endo face of the furanose being on the same side as C5' and the base, while the “exo” face is on the opposite face to the base). These sugar puckers are located in the south (S) and north (N) domains of the pseudorotation cycle of the furanose ring¹⁸¹ and as a result spectroscopists frequently use “S” and “N” designations which, incidentally, reflect the relative shapes of the C1'-C2'-C3'-C4' bond paths in the C2'- and C3'-endo forms respectively (**Figure 3.1.2**). In solution, the C2' and C3' conformations are in rapid equilibrium and are separated by an energy barrier of less than 20 kJ mol⁻¹. A-DNA has predominantly C3'-endo sugar units, whereas B-DNA contains sugars in the C2'-endo form. Changes in the conformation of the furanose are driven by the relative strengths of stabilizing stereoelectronic effects, namely, gauche and in particular anomeric effects.¹⁸² Thus, in 2'-deoxyribonucleosides, the 5'-OH and 3'-OH groups prefer a gauche orientation with respect to the O4' (as is the case in the C2'-endo form or S state), whereas in ribonucleosides the conformational equilibrium is affected by three more gauche effects, namely those between the 2'-OH and 3'-OH, 2'-OH and O4', and finally 2'-OH and the nucleobase nitrogen atoms. The 3'-OH gauche effect with O4' in DNA favors the C2'-endo pucker¹⁸² particularly in aqueous solution.

The anomeric effect¹⁸³ also exists in nucleosides between the C1' substituent (base) and the ring oxygen (O4'). It describes the preferred antiperiplanar orientation which exists between the α lone pair of the O4' and the C1'-N bond and that furnishes an optimal overlap between the non-bonding (O4') and the antibonding σ^* (C1'-N) orbitals.¹⁸² As a result of this effect and the O2' \rightarrow O4' gauche effects the sugars in RNA adopt primarily the C3'-endo N pucker, regardless of whether the RNA is found in single-stranded, double or triple-helical forms.

An interesting and related issue pertains to the structural consequence of inverting 2'-OH to the *arabino* configuration, and to the incorporation of such arabinonucleotides into a DNA molecule. Based on the analysis above, arabinonucleotides are expected to mimic deoxyribonucleotides since a combination of $O2' \rightarrow O4'$, and $O3' \rightarrow O4'$ *gauche* effects would stabilize the C2'-endo geometry (Figure 3.1.2).

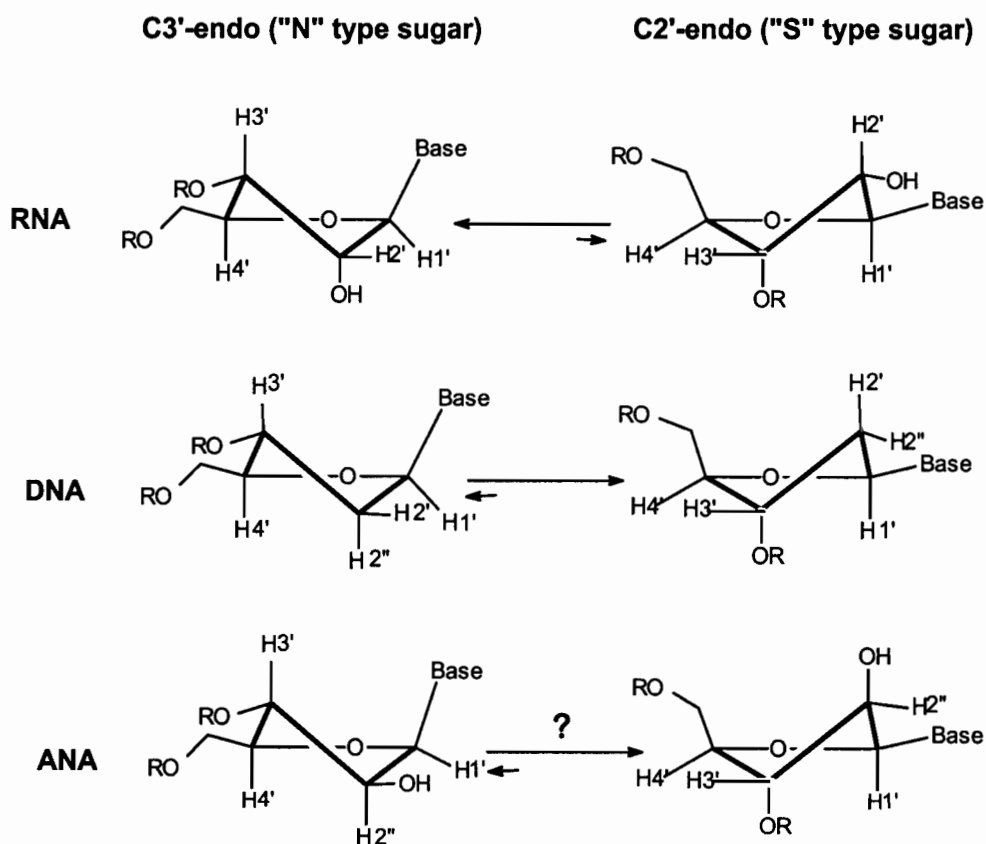


Figure 3.1.2: Sugar pucker equilibria of RNA, DNA and ANA.

The actual pucker of the arabinofuranose is still being debated, and the conformation of ANA strands alone or in a complex has not yet been determined. Structural models of an arabinonucleotide in B-form DNA show that the 2'- β OH group of the arabinose sugar is directed toward the major groove, while in the case of duplex RNA the 2'- α OH groups are projected into the minor groove. Although there seems to be free space for new modifications in this β position, it is in this same position that the β -2' OH groups of the arabinofuranosyl units could potentially interfere with the binding and catalytic functions of those enzymes that recognize the DNA:RNA by contacting the major groove. For example T7 RNA Polymerase is sensitive to the arabinose sugar within the DNA strands.^{117,148} On the other hand ribose 2'- α OH groups affect enzymes that bind to the minor groove of hybrid DNA:RNA duplexes such as RNase H.¹¹¹

Another molecular modelling study indicated that strong helix specific perturbations would arise if an araC-dG base pair were to exist in an A-form helix environment (all sugars C3'-endo).¹⁴⁹ These perturbations appeared to be much greater than those introduced by an araC residue in B-form DNA. Since it has been observed that araC is not enzymatically incorporated into RNA to any appreciable degree,¹⁸⁴ the authors of the above modelling study¹⁴⁹ hypothesized at that time, that steric effects may partially or wholly account for this exclusion, *i.e.*, araCTP would not be introduced into RNA because it could not be made to fit into the A-form DNA:RNA hybrid which is believed to exist at the site of catalysis within the RNA polymerase ternary complex. Furthermore Altona and co-workers discovered by NMR analysis that araC inserted into DNA *i.e.* duplex d(CGaraCTAGCG)₂ adopts exclusively, a hairpin structure.¹⁸⁵ Lastly araC was also found to facilitate B form \rightarrow Z form conversion when it was incorporated into an alternating dC-dG sequence [d(CCGaraCGCG)]₂.¹⁸⁶ The Z form is usually observed for dC-dG rich unmodified DNA duplexes.

Little is known, however, about the properties of oligoarabinonucleotides (ANA) and the duplexes formed by them (*e.g.* ANA:DNA, ANA:RNA and ANA:ANA). A relevant question relating to oligoarabinonucleotides is "*How is the structure of the resulting duplex, if formed, affected by substitution of arabinofuranosyl residues within it?*" Previous experimental approaches to this question have focused on the use of

homopolymers of arabinonucleotides such as ara(A_p)₇A, by previous members of Damha's group¹⁰⁹ who found that inversion of stereochemistry at the C2' position of ribonucleotides in general had no effect on hybridization properties. In fact, the significantly higher T_m observed for complexes araA₈:polydT and araA₈:polyrU relative to complexes rA₈:polydT and rA₈:polyrU indicated that the 2'-OH of ANA exerts, a stabilizing effect, while studies involving ara(U)_n led to the opposite conclusion. Other independent studies have shown that the araC-dG base pair has comparable stability to the dC-dG base pair in duplex DNA,¹⁵⁵ and that 2'-*fluoro*-arabino modifications within DNA are generally stabilizing although the reason for this effect is not well understood.^{187,188}

Arabinonucleotides are expected to possess an inherent stability comparable to that of deoxyribonucleic acids. For example, with respect to enzyme modifications and cleavage of the phosphodiester linkages, the unfavorable *trans* relationship between the 2'-OH and the 3'-5' phosphodiester linkage would prevent participation of the 2'-OH in intranucleotide cleavage.^{146,156,169} In fact, Giannaris and Damha¹⁰⁹ have reported that oligoarabinonucleotides are completely stable to treatment with 0.1N KOH for 24 h, conditions that fully hydrolyze oligoribonucleotides; partially stable to snake venom phosphodiesterase (SVPDE); and just as stable as RNA to other endonucleases such as Nuclease P1, Ribonuclease S1 and the 5'-exonuclease calf-spleen phosphodiesterase. These results have now stimulated further experimental work on ANA oligomers. Specifically the thermal stability of complexes (duplexes and triplexes) formed between ANA oligomers of mixed base composition and DNA and RNA will be studied for the first time (Chapter 3).

3.1.2 Specific Objectives

In an attempt to address the question of the structural consequence of the C2'-OH inversion of ANA, a series of oligonucleotides, 18nt in length, and complementary to the R region near the 5'-LTR of the HIV-1 genomic RNA, were synthesized (**Figure 3.1.3**).

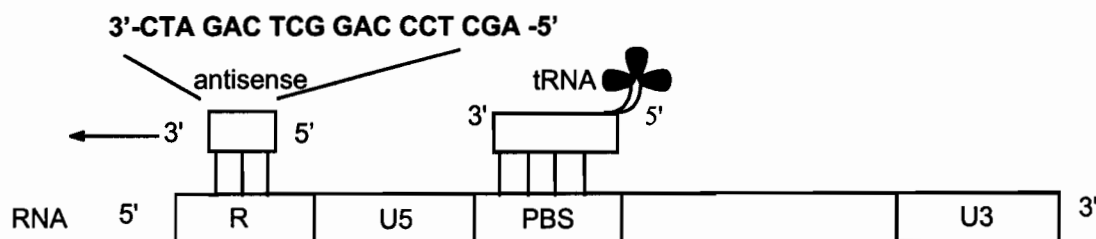


Figure 3.1.3: The R region sequence of the HIV-1 genomic RNA. The sequence of the antisense molecule is 5'-AGC TCC CAG GCT CAG ATC-3'. tRNA (shown) is the natural primer of RT mediated DNA synthesis. Watson-Crick hydrogen bonds are represented by vertical lines.

The structures of the complexes formed with RNA and DNA targets were also then investigated by CD and UV spectroscopy. This investigation itself contains three sub-sections based on the composition of the antisense oligomers. These are (a) DNA oligomers containing arabinonucleotide inserts, (b) DNA-RNA chimeras, and (c) pure ANA oligomers.

Table 3.1.1 A shows prepared DNA sequences containing a single ara insert, with the exception of **3.7**, which contains two inserts. Note the sequential movement of the insert away from the 3' end. Entry **B** illustrates chimeric DNA-RNA sequences containing ribonucleotide runs of differing length and placement within an 18 nt DNA control. The association of ribonucleotide strands with deoxynucleotide strands to form a hybrid RNA:DNA double helix is one of the principle reactions of the biological information transfer system and these are found at initiation DNA replication forks (Okazaki fragments).¹⁸⁹ As such, a closely related project involved DNA-RNA chimeras associating to target DNA and RNA with the same 18 nt sequence shown in **Figure 3.1.3** (above). Entry **C** shows a uniformly modified (pure) oligoarabinonucleotide. The physicochemical properties of this sequence along with those of the corresponding DNA and RNA controls are described below. Also, the biological studies of some of the above sequences will be summarized in Chapter 4.1 - 4.3.

Table 3.1.1: DNA Sequences Containing Various Ara Inserts and Stretches of Ribonucleotides

Designation	Sequence
A. Antisense DNA containing arabinose inserts	
aC ₁	5'-AGC TCC CAG GCT CAG AT _a C -3'
aU ₂	5'-AGC TCC CAG GCT CAG A _a UC -3'
aA ₃	5'-AGC TCC CAG GCT CAG _a ATC -3'
aG ₄	5'-AGC TCC CAG GCT CA _a G ATC -3'
aA ₅	5'-AGC TCC CAG GCT C _a AG ATC -3'
aC ₁₆	5'-AG _a C TCC CAG GCT CAG ATC -3'
aC(1&8)	5'-AGC TCC CAG G _a CT CAG AT _a C -3'
B. Antisense oligonucleotides: DNA, RNA and DNA-RNA chimeras	
D17R ₁	5'-AGC TCC CAG GCT CAG AT _r C -3'
D13R ₅	5'-AGC TCC CAG GCT Cr(AG AUC) -3'
D9R ₉	5'-AGC TCC CAG r(GCU CAG AUC) -3'
R9D ₉	5'-r(AGC UCC CAG) GCT CAG ATC -3'
R17D ₁	5'-r(AGC UCC CAG GCU CAG AU)dC -3'
C. Completely modified oligoarabinonucleotide (3,15) and control DNA and RNA sequences	
D18	5'-d(AGC TCC CAG GCT CAG ATC) -3'
R18	5'-r(AGC UCC CAG GCU CAG AUC)-3'
A18	5'-a(AGC UCC CAG GCU CAG AUC) -3'

3.1.3 Results

Thermal Denaturation Studies

Duplex formation between the 18 nt oligomers (referred as “antisense” strands) and the respective targets was monitored by both UV and CD spectroscopy. All duplexes gave single cooperative, melting transitions from which accurate T_m values were extracted. This cooperative nature indicates two-state melting behavior for duplexes. The overall shapes of the curves were similar but as expected the T_m values varied (Table 3.1.2).

Table 3.1.2: Thermal Stability of Duplexes Formed by DNA Strands Containing Arabinonucleotide *Inserts* and Target RNA and DNA Single Strands

Antisense oligomer	Antisense sequence with ara inserts	Target T_m (°C)	
		RNA	DNA
D18	5'-AGC TCC CAG GCT CAG ATC -3'	72.3	68.0
aC1	5'-AGC TCC CAG GCT CAG ATaC -3'	71.6	68.0
aU2	5'-AGC TCC CAG GCT CAG AaUC -3'	71.7	65.5
aA3	5'-AGC TCC CAG GCT CAG aATC -3'	71.7	65.5
aG4	5'-AGC TCC CAG GCT CAaG ATC -3'	69.9	64.0
aA5	5'-AGC TCC CAG GCT CaAG ATC -3'	71.7	66.2
aC at 3' & 8	5'-AGC TCC CAG GaCT CAG ATaC -3'	69.1	63.7
aC16	5'-AGaC TCC CAG GCT CAG ATC -3'	71.5	65.0

All mixtures contained 1:1 antisense:target strands in 140 mM KCl, 1mM MgCl₂, 5 mM Na₂HPO₄, pH 7.2, and adjusted with HCl.

The antisense strands containing one or two arabino inserts formed complexes with both RNA and DNA targets, and had T_m values in the ranges of 68-72°C and 63-68°C respectively. The data suggest that the terminal araC-G base pair has comparable stability to the dC-dG and dC-rG base pairs in both DNA:DNA and DNA:RNA duplexes (Table 3.1.2). A decrease in T_m was observed as the ara insert was moved away from the

3' terminus, with the largest decrease detected for araG at position 4. The range of T_m change was about 5°C for this study.

Table 3.1.3 summarizes the effect of deoxyribose/ribose substitutions in the antisense oligonucleotide. The presence of either one *deoxy* or one *ribose* unit at the 3' terminus had negligible effect on binding to the targets relative to the R18 and D18 sequences as expected. In the case of the other chimeras namely D13R5, D9R9 and R9D9 the T_m s were lower, and appeared to average out. When the target was single stranded DNA, the T_m s also varied in a range of about 4°C as seen in the study above. Associations of R17D1 and R18 to RNA substantially increased the T_m to about 10-12°C over the DNA richer sequences.

Table 3.1.3 Thermal Stability of Duplexes Formed by DNA-RNA Chimeras and ssRNA and ssDNA Sense Strands

Antisense oligomer	DNA-RNA chimera sequence	Target T_m (°C)	
		RNA	DNA
D18	5'-AGC TCC CAG GCT CAG ATC -3'	72.3	68.0
D17R1	5'-AGC TCC CAG GCT CAG ATrC -3'	71.1	68.0
D13R5	5'-AGC TCC CAG GCT Cr(AG AUC) -3'	68.6	65.8
D9R9	5'-AGC TCC CAG r(GCU CAG AUC) -3'	67.6	64.7
R9D9	5'-r(AGC UCC CAG) GCT CAG ATC -3'	66.6	65.2
R17D1	5'-r(AGC UCC CAG GCU CAG AU)C -3'	82.0	68.0
R18	5'-r(AGC UCC CAG GCU CAG AUC) -3'	84.6	66.2

All mixtures contained 1:1 antisense:target strands in 140 mM KCl, 1mM MgCl₂, 5 mM Na₂HPO₄, pH 7.2, and adjusted with HCl.

In summary binding of DNA-RNA chimera to target DNA had smaller variations in T_m than binding to target RNA. **Figure 3.1.4** illustrates these variation of T_m as a function of RNA residues in the antisense strand. The minima for the two curves occur when the deoxyresidues equals the number of ribonucleotide residues. The minima appears near

68°C (RNA target) and 66°C (DNA target). However for this particular sequence, targeting RNA is more sensitive to the R → D substitutions relative to targeting ssDNA. In fact targeting DNA yields duplexes of similar stability irrespective of the degree of substitution of the antisense strand.

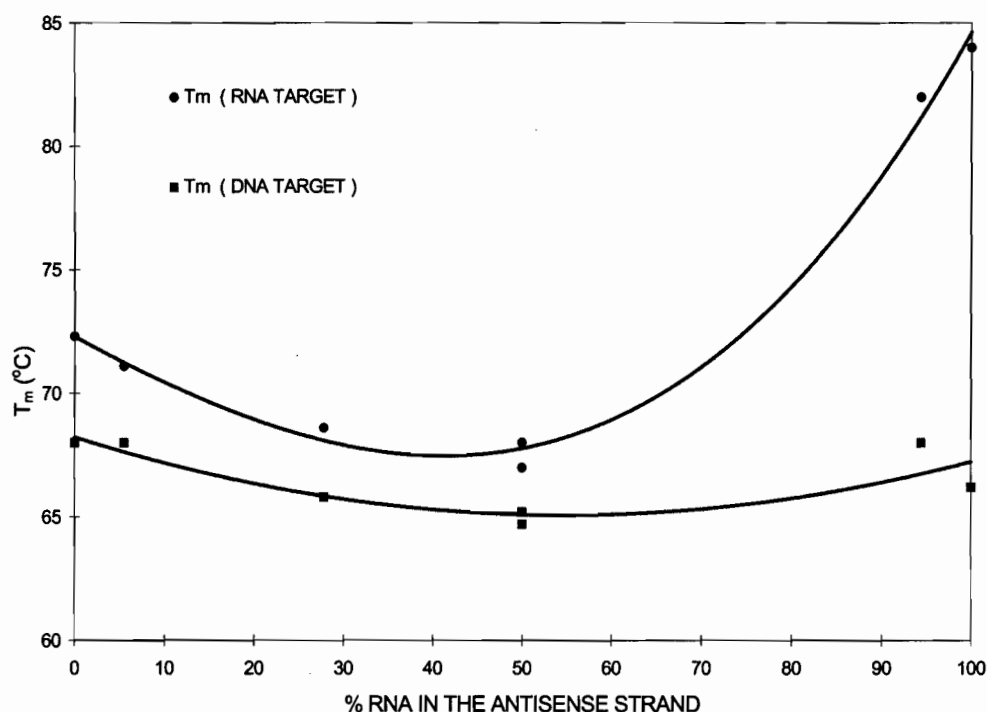


Figure 3.1.4: Variation of T_m as a function of RNA content in the antisense strand

Circular Dichroism of Duplexes

Effect of arabinose substitution on DNA:RNA hybrids

As pointed out in Chapter 1 (pp. 21-24), A- and B- form duplexes have very different CD spectral signatures. The A-form spectrum is characterized by a strong negative band at 217 nm, while the B-form spectrum of DNA:DNA duplexes has relatively equal positive and negative bands of moderate intensity. Consistent with these earlier findings one observes an obvious trend in the amplitudes and wavelengths of the CD maxima as the araX residue is shifted from the 3' terminus towards the interior, suggesting the existence of a gradient in conformation of the duplexes (**Figure 3.1.5A and B**). The overall conformation is “A like” and, interestingly, this A-helical nature

appears to diminish when the ara insert is at the fourth position from the 3' end. Loss of A character is evident by the gradual amplitude decrease of the key negative peak at 210 nm as the araX shifts into the interior of the strand.

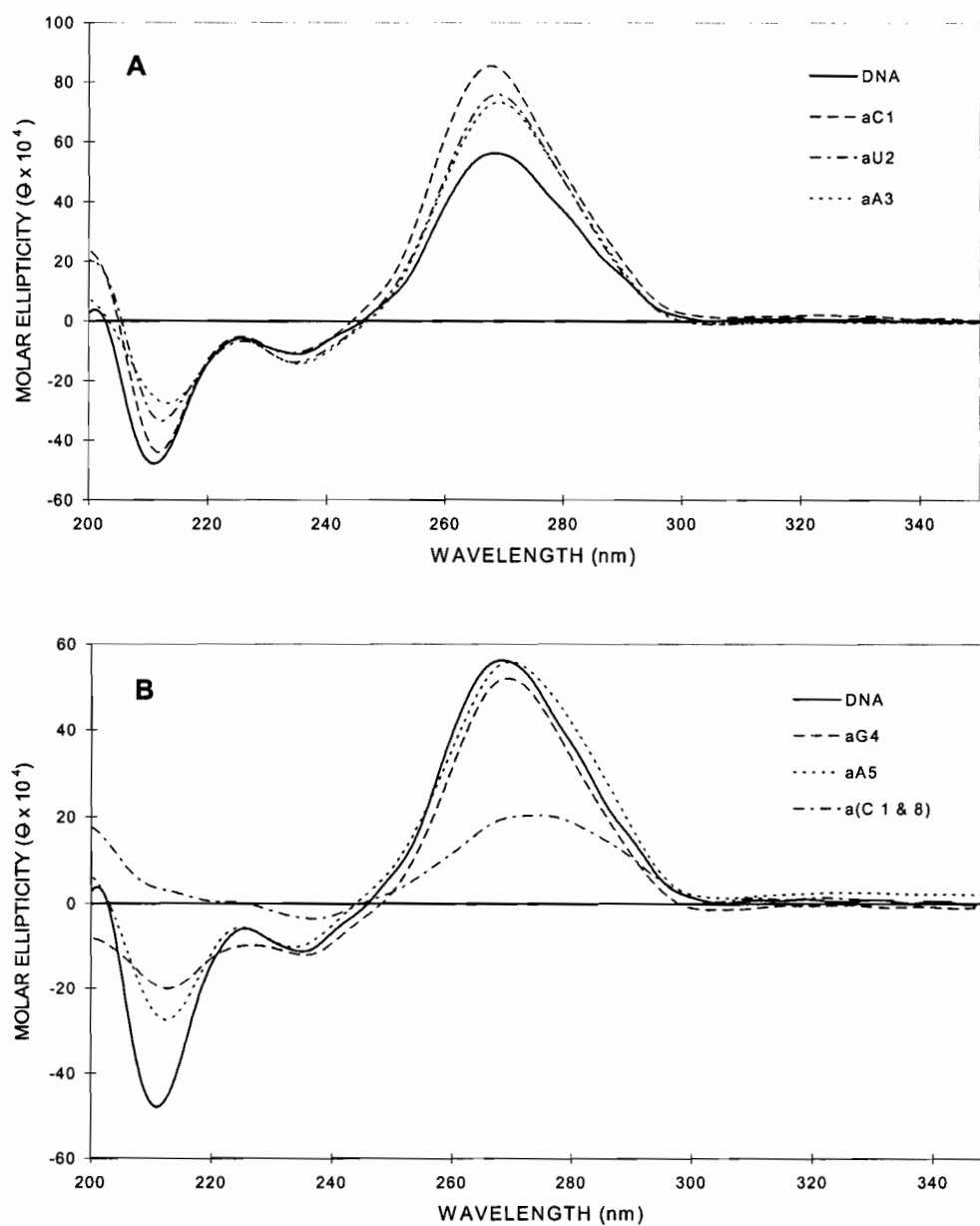


Figure 3.1.5 A and B: CD spectra of duplexes formed between DNA containing arabinoside inserts and complementary RNA. Base Sequence is shown in **Table 3.1.1 A**

Effect of ribose substitution on the DNA strand of DNA:RNA hybrids

RNA (purine-rich):DNA hybrids have more A-type character than RNA (pyrimidine-rich):DNA hybrids.¹⁹⁰ Close inspection of the spectra shown in **Figure 3.1.6** reveals that the pure RNA duplex, R18:RNA, and the corresponding hybrid duplexes, R17D1:RNA and R9D9:RNA fall into the A conformation family, whereas those derived from D17R1, D13R5, and D9R9 are significantly less A-like. For instance, in the spectra of the latter series the “A” signature peak at 212 nm is reduced gradually as the DNA content of the antisense strand increases while the spectrum of the R18:RNA duplex is the most A like (**Figure 3.1.6A & B**).

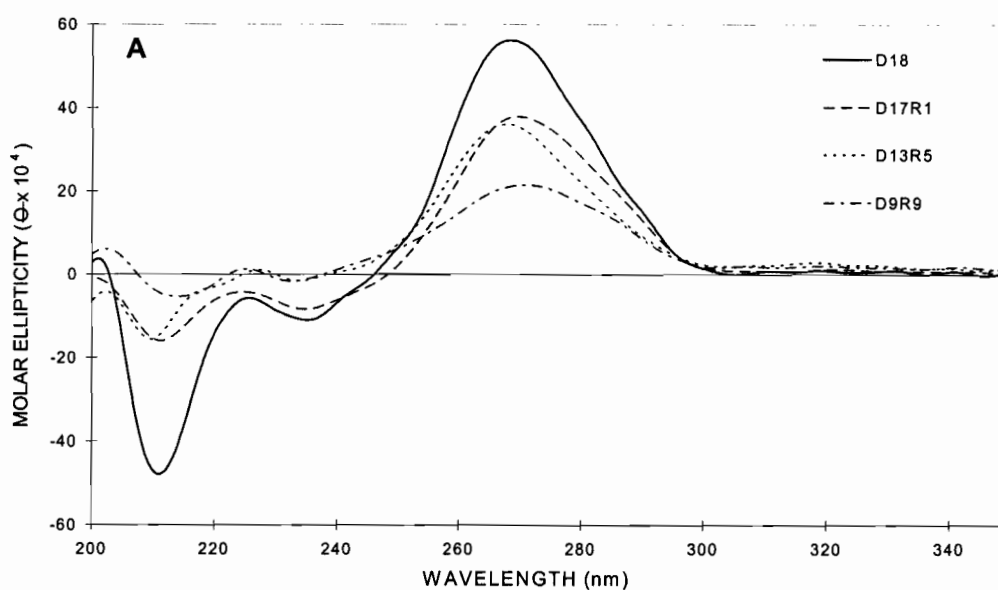


Figure 3.1.6 A: CD spectra of DNA:RNA, (DNA-RNA):RNA and RNA:RNA duplexes. Base Sequence as shown in **Table 3.1.1 B**

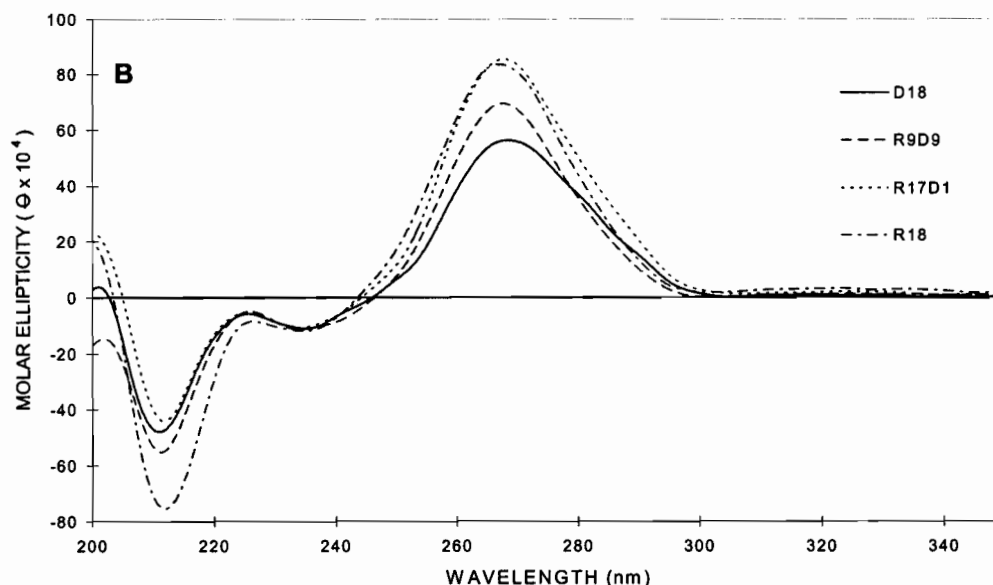


Figure 3.1.6 B: CD spectra of DNA:RNA, (DNA-RNA):RNA and RNA:RNA duplexes. Base Sequence as shown in **Table 3.1.1 B**

UV Melting Studies of Duplexes Formed by Uniformly Sugar Modified Oligonucleotides (ANA, RNA and DNA)

The affinity of 3'-5' linked oligoarabinonucleotides (ANA) for target DNA and RNA strands was likewise determined by thermal denaturation studies (**Table 3.1.4** and **Figure 3.1.7**). Oligoarabinonucleotide (A18) (**3.15**) formed a stable duplex with RNA but not with ssDNA. For instance, the UV-melting curve of ANA + DNA (1:1) at 260 nm showed a weak cooperative transition centered at $\sim 26^{\circ}\text{C}$, indicating weak or no base pairing in this case. Resmini *et al.*¹⁹¹ have also studied the association of ANA strands containing araT instead of araU and have shown decreased stabilities of -1.0°C to $1.6^{\circ}\text{C}/\text{base}$ when hybridizing to a target DNA. The smaller decrease observed in their study is probably a consequence of composition, primary sequence and/or substitution of the araT for araU in the oligomer. ANA (**3.15**) bound to target RNA with a decrease of $-1.5^{\circ}\text{C}/\text{base}$ relative to DNA **3.1**. Thus the order of stability for this series is R18:RNA (85°C) > D18:RNA (73°C) > A18:RNA (44°C) (**Figure 3.1.7 B**).

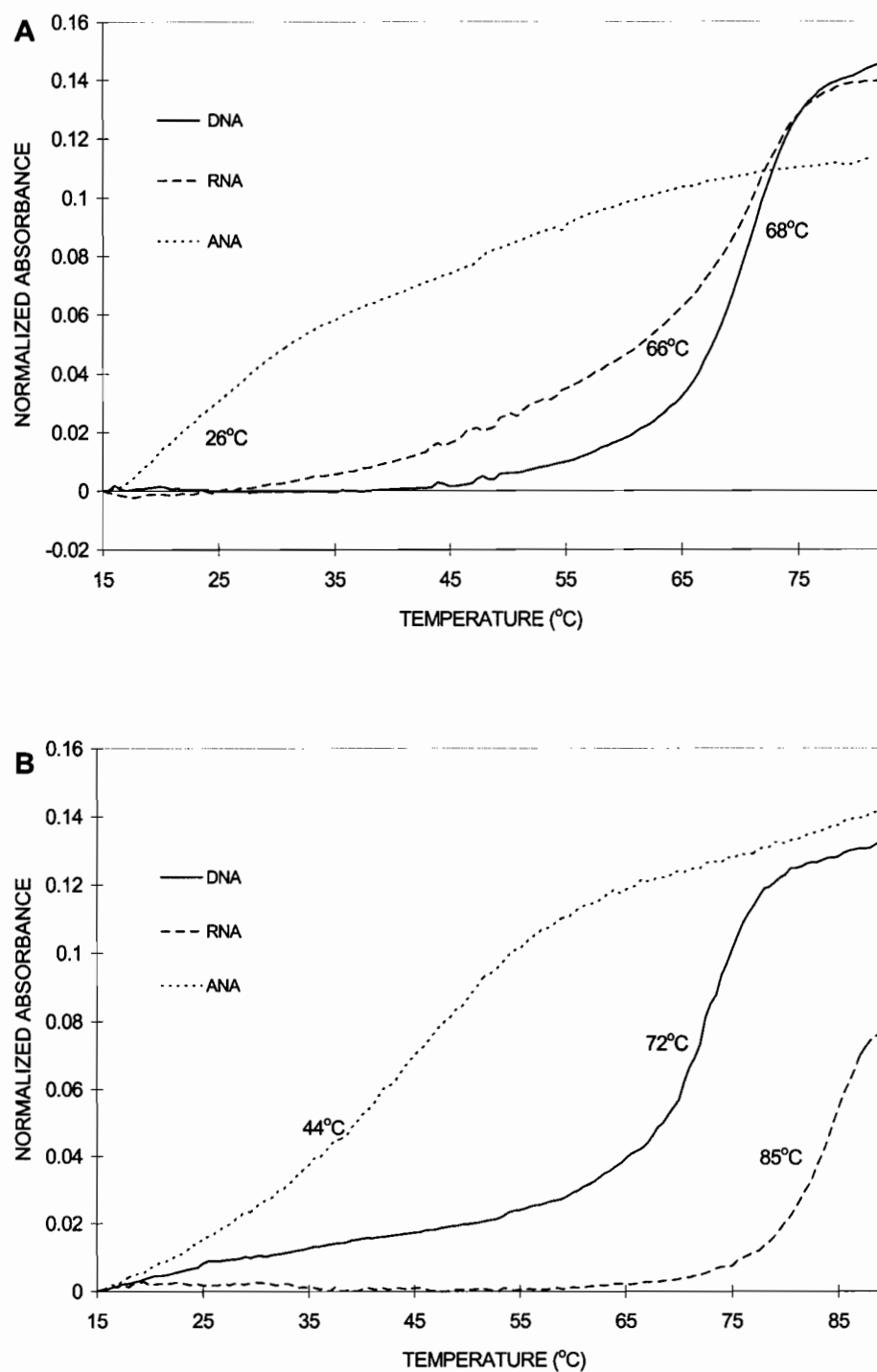


Figure 3.1.7: Hybridization of DNA, RNA and ANA to (A) complementary DNA and (B) complementary RNA. Solutions contain 140 mM KCl, 5 mM Na_2HPO_4 , 1 mM MgCl_2 , pH 7.3. Base sequences of the strands are given in **Table 3.1.4**.

Table 3.1.4: Thermal Denaturation Studies of Uniformly Substituted Modified Oligomers

Antisense oligomer		Target T_m (°C)	
		RNA	DNA
DNA	d(AGC TCC CAG GCT CAG ATC)	72.3	68.0
RNA	r(AGC UCC CAG GCU CAG AUC)	84.6	66.2
ANA	<i>a</i> (AGC UCC CAG GCU CAG AUC)	44.0	26.0 ^a
Thio-ANA	<i>a_s</i> (AGC UCC CAG GCU CAG AUC)	38	n.m. ^b

The mixtures contained antisense and target strands in (1:1 ratio) 140 mM KCl, 1mM MgCl₂, 5 mM Na₂HPO₄, pH 7.2, adjusted with HCl. ^a Very broad transition. ^b n.m.: not measured.

Phosphorothioate substitution in the backbone of ANA (**3.17**) decreased the affinity against RNA even further, relative to the non-thioated duplexes. The T_m decrease for thioate-DNA (**3.18**):RNA was -0.5°/base while that of the thioate - ANA:RNA was (-) 0.3°/base. This is consistent with the decreased stability associated with the PO to PS (thioate) substitution of ssDNA.³⁵

From all of the above studies (inserts, chimeras and completely modified oligomers) one can conclude that modification of an oligodeoxynucleotide (DNA) strand with arabinonucleotides led to considerable reduced binding affinity ($\Delta T_m \sim 2.3^\circ\text{C}/\text{modification}$) with complementary target DNA, and moderate reduced binding affinity ($\Delta T_m \sim 1.2^\circ\text{C}$) with target RNA. Also, a pure ANA oligomer of mixed base composition was shown to form a less stable duplex with RNA under physiological conditions, compared to DNA:RNA and RNA:RNA. The ANA sequence also exhibited unusual selectivity for complementary RNA over ssDNA. This study indicates that the inversion of configuration at the C2' position of ribonucleotides has a significant effect on the hybridization.

Circular Dichroism of ANA:RNA duplexes

The CD spectrum of single stranded ANA in a physiological medium displayed positive and negative peaks of equal and moderate intensity, although slightly blue shifted relative to the analogous ssDNA sequence (**Figure 3.1.8 A**).

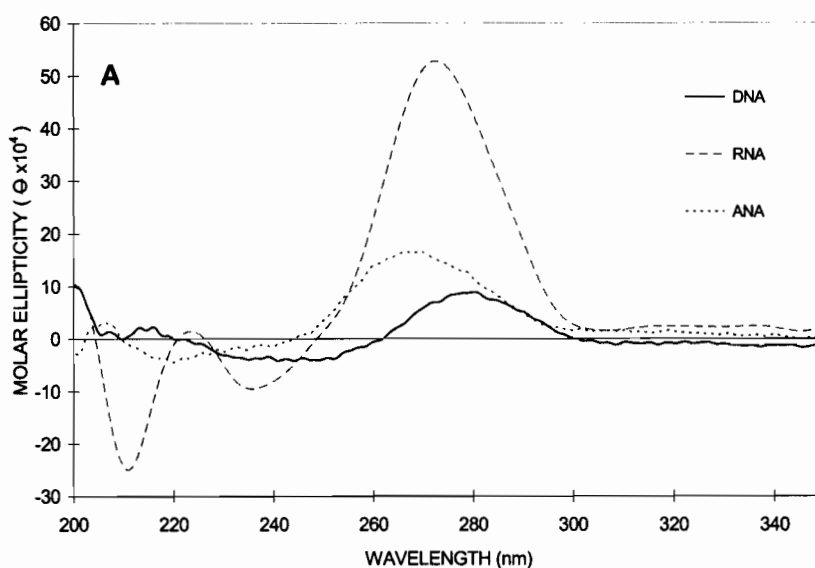


Figure 3.1.8 A: CD spectra of single strands. Spectra were recorded in 140 mM KCl, 5 mM Na₂HPO₄, 1 mM MgCl₂, at 5°C (pH 7.3).

The CD spectra of the duplexes are illustrated in **Figure 3.1.8 B**. The appearance of a large positive CD band at ~267 nm and a large negative CD band near 210 nm in the spectrum of the pure RNA duplex is characteristic of the A-conformation as previously discussed. In sharp contrast, the CD spectra of the corresponding DNA duplex has positive and negative CD bands of smaller magnitudes at wavelengths above 220 nm and a crossover point at ~260-265 nm, characteristic of the B conformation. The large negative peak at 210 nm that characterizes A-form RNA duplexes, is nearly absent in the spectra of the pure DNA duplex.

Each hybrid shows both differences from and similarities to the spectrum of the pure RNA duplex (R:R). For example the ANA:RNA (A:R) spectrum is similar to the DNA:RNA (D:R) spectrum, and displayed an intermediate form between pure R:R and pure D:D helices.

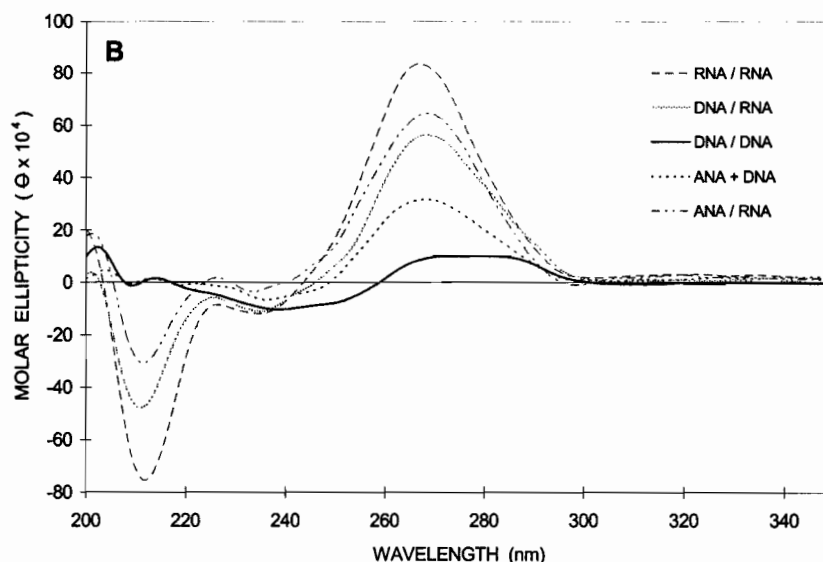


Figure 3.1.8 B: CD spectra of duplexes. Spectra were recorded in 140 mM KCl, 5 mM Na_2HPO_4 , 1 mM MgCl_2 , at 5°C (pH 7.3).

This is expected since the relatively high purine content of the target RNA sequence (55%) suggests an energetic propensity for the heteroduplex to base stack in the A form conformation as seen with DNA:RNA hybrids.¹⁹⁰ Closer inspection of the CD spectra in **Figure 3.1.8 B** show that the strong negative band below 210 nm, also present in the spectrum of R:R, was much reduced in the A:R and D:R spectra. Also the positive band at 260 nm was more reduced for D:R and A:R relative to R:R. The spectral similarities between the A:R and D:R hybrids may reflect the similar strand orientation of the base chromophores in the heteronomous conformation.¹⁹² The CD spectrum of “ANA:DNA” is not as clear cut. From the UV melting profiles, the rather low T_m of the ANA+DNA mixture illustrates that ANA binding to DNA is rather weak and what one observes in the CD spectra could be predominantly a “sum” of the ssANA and ssDNA spectra.

The above results suggest that the global conformations of ANA:RNA and DNA:RNA hybrid duplexes are similar which are themselves intermediate between those of the pure DNA (B form) and RNA (A form) duplexes.

3.1.4 Discussion

Very little has been reported in the literature regarding the properties of arabinonucleic acids. Among the few reports available the main emphasis is on the effects of incorporating a single arabinonucleotide residue into DNA duplexes.^{193,194} For example it has been shown that when araC was inserted into a self complementary DNA duplex the T_m was lowered by 2°C/araC insert relative to the unmodified duplex.¹⁴⁷ Likewise the single arabino inserts studied here did not greatly affect thermal stability relative to the D18 control (all DNA), with the exception of araG at an internal position (position 4) and the case of two araC inserts within the 18nt DNA strand (position 1 and 8). The apparent destabilization cause by arabino units is significantly smaller than that created by mismatches at similar positions (not shown), suggesting that the arabino units in these duplexes retain classical base pairing interactions. Thus one can conclude that duplex stability decreases slightly by substitution of one ara→deoxy residue within DNA strands.

For the chimeric (R-D:R) hybrids, the T_m s observed were only slightly lower than those of the pure (D:D and R:R) duplexes. The stability of the duplexes averages out when the number of deoxy- and ribo- nucleotides are more or less equal in the R-D (antisense) strand irrespective of whether the stretch of ribonucleotides were at the beginning (R9D9) or at the end (D9R9) of the antisense strand. In fact the absolute value of the T_m s of these sequences are rather close as seen in **Figure 3.1.4**.

In the third study which deals with pure arabinonucleic acid strands the physicochemical characteristics observed were not as clear-cut. Usually mixed-sequence heteroduplexes are reported to have reduced thermodynamic stability compared with pure homoduplex DNA (B-form) or RNA (A-form).¹⁹⁵ The T_m s of the controls used in this study were D:D (68°C), R:D (64°C), D:R (72°C) and R:R (85°C). Differences in T_m values of these complexes may arise from a combination of factors, which include differences in the conformational freedom (entropy) of the single strands, altered water structure around the backbone of the duplexes and altered stacking interactions. Moreover determination of the thermodynamic properties of the heteroduplexes relative to their homoduplexes is complicated by the greater sequence-dependent variation in

thermal stability encountered for different RNA:DNA pairs. The mixed ANA sequence appeared to have a binding selectivity to RNA over ssDNA, as seen in **Table 3.1.4**. To investigate the global helical conformation of complexes formed by ANA with target ssDNA or RNA, CD spectra of the various duplexes were recorded. As pointed out earlier in this chapter the structure of A-form duplexes is characterized by a wide and shallow minor groove, while in the structure of B-form duplexes it is a narrow and deep minor groove.¹ The conformation of DNA:RNA hybrids are not as clear cut and they appear to adopt a more complex, sequence-dependent structure that is generally intermediate between the A-RNA form and B-DNA form structures.^{85,190,196} Structurally this heteroduplex topology is closer to the A-form of pure RNA duplexes and thus is commonly described to as an “A-like” helical structure. In fact ANA:RNA and DNA:RNA hybrids appear to share this “A-like” helical form.

Striking differences in thermal stability of ANA:RNA complexes of different base sequences were observed. For example a mixed 12 nt arabino pyrimidylate *ara*(CCUCUCCUCCU) complexed with RNA has a larger T_m (48°C, see Chapter 3, Section 3) than the 18 nt oligomer *ara*(AGC UCC CAG GCU CAG AUC) (44°C) studied above, despite the shorter length. This can be reasoned out on the basis of a higher G:C content (67%) present in the shorter duplex which also possesses a higher purine content in the RNA target strand (100%), relative to the 18 nt duplex (C:G = 61%; purine RNA target content = 55%). These results further imply that RNA with long purine repeats should be good targets for mixed DNA or ANA antisense oligomers from the standpoint of duplex stability.¹⁹⁰

Another factor affecting duplex stability may be the actual base sequence *i.e.* homopolymeric versus mixed base sequences. Previously Giannaris and Damha have shown that complexes formed between *homopolymers* of oligoarabinonucleotides and complementary ssDNA and RNA, are rather stable exhibiting comparable melting temperatures and, in some cases, greater than the corresponding unmodified oligomers.¹⁰⁹ In new work currently in progress in the Damha group a *homopolymer* of arabinoadenylate residues showed varying results under different cationic conditions.¹⁹⁷ Furthermore there appears to be no selectivity of the *araA*₁₈ strand to the riboU₁₈ (RNA) over the ssDNA (dT₁₈) target,¹⁰⁹ and hence for this system the selectivity rule appears to break down.

Thus it is clear that there are dramatic differences among the stabilities of hybrid duplexes formed by ANA and DNA, or ANA and RNA that are magnified in some sequences.

The selective hybridization properties of ANA strands of mixed base composition towards RNA is intriguing. DNA is known to be more “flexible” than RNA, which manifests itself in the fact that duplex DNA is polymorphic whereas double helices of the latter usually form only A-type helices.¹ Thus from this perspective, one would expect that any modified oligomer that binds to RNA should also complex to DNA. However this was not the case for ANA (3.15) as observed by UV and CD spectroscopy (Figure 3.1.7 and 3.1.8) and gel shift mobility assays (data not shown). This finding together with the extent to which ANA binding occurs for different sequences and target type (DNA *vs.* RNA), suggests that molecular recognition is primarily based on general conformational features (secondary) and to some extent by specific sequences effects (primary). For example, the selective recognition of RNA by ANA may be primarily determined by shape complementarity modulated by sequence-specific effects. While a DNA target has methyl group at C5 of thymine, the corresponding RNA target instead has a hydrogen atom (uracil). This together with the fact that the 5-methyl dT groups and ara-2'-OH may sterically interact in the major groove may explain, at least in part, the relatively weak binding of ANA (3.15) to DNA. Clearly, high field NMR analysis and crystallographic work on both ANA:DNA and ANA:RNA duplexes are needed in order to gain an understanding of this ‘RNA selectivity’.

Because of the many diseases caused by RNA viruses, including AIDS, compounds capable of selective binding to RNA should be considered in developing effective chemotherapeutic agents. Bacterial disease sources are also susceptible to this approach of drug treatment. Thus, the opportunity for developing therapeutic agents targeted to distorted RNA structures are manifold. For this approach to succeed, and to avoid unwanted side effects it is apparent that such agents should possess selectivity for the RNA target in comparison to ssDNA. Based on recent literature the lower relative affinity of modified oligomers for DNA appears to be a rather frequent phenomenon. Among these oligonucleotide analogues are the 2'-5' linked systems (2'-5' DNA and 2'-5' RNA) and 2'-O-Methyl (3'-5') RNA.^{198,199} Apart from these, ANA appears to be a potential candidate

as an effective antisense agent based on this discrimination and its other interesting properties examined in Chapter 4.^{200,201}

3.1.5 Conclusions:

The binding of DNA oligomers containing one to two arabinonucleotide inserts was not greatly affected relative to the binding of an unmodified DNA control. Chimeras containing equal amounts of RNA and DNA residues in the antisense strand have similar stabilities but the conformation of the overall complexes formed are rather different. When the number of RNA residues in the antisense increases the thermal stability of the R-D:R duplexes also increases and the overall conformation resemble the A-type from RNA helices. ANA strands have been shown to bind selectively to RNA, which could derive from shape complementarity. The resulting conformation of the ANA:RNA hybrid is A-like as suggested by CD spectroscopy. Poor DNA binding exhibited by the ANA oligomer may contribute to their therapeutic efficacy by minimizing potential toxicity due to ssDNA binding.

New questions have now arisen from these studies that need to be addressed, for example “*what is the conformations of arabinonucleotides in ANA-RNA chimera complexed to ssRNA?*” Some of these issues are already being addressed in the Damha research group. Furthermore functionalization of the 2'-position of ANA with new groups (2'-OMe, NH₂, etc.) and moieties such as intercalators and groove binders may yield molecules with many interesting biological properties.

3.2 SYNTHESIS AND BIOPHYSICAL PROPERTIES OF BRANCH NUCLEIC ACIDS CONTAINING ARABINOSE AT THE BRANCH-POINT

3.2.1 Introduction

The formation of triple helices has become an area of great interest to chemists and biologists for their possible role in natural and artificial regulation of gene expression at the level of transcription,^{202,203} and for use in analytical, diagnostic or synthetic methods.^{90,204} It also provides a versatile structural motif for the design of molecules capable of sequence-specific recognition of double-helical DNA, a novel function that is 10^6 more specific than restriction enzymes.^{64,75,205} However there is still a need to extend the triplex recognition code which is restricted to homopyrimidine sequences and thus a general solution to this problem is eagerly awaited.²⁰⁶ The use of combinatorial approaches may well be useful, and one such study using solely natural bases²⁰⁷ has found that novel G•A:T triplets appear to be consistently stable in purine (anti-parallel) triple helices.

Although much work has begun to focus on the physical and chemical requirements for triplex formation, the precise conditions required have not yet been fully elucidated. Dissecting the relative contributions of all factors controlling triple helix formation will be pivotal when considering the use of modified oligonucleotides for *in vivo* applications where temperature, pH and ionic conditions are strictly controlled. The absolute requirements for triplex formation (especially reverse-Hoogsteen motifs) have not yet been established and thus there seems to be a growing focus on the physical and chemical conditions required to promote their formation.^{75,77,85,86,208,209,210}

Studies involving natural and unnatural nucleic oligomers^{23,24} with chemically modified backbones (connectivities and sugar moieties), circular,²¹¹ branched^{114,116} and hairpin-like oligomers containing non-nucleosidic loops have appeared in the literature.^{212,213,214} Some of the branched molecules include “comb”-like nucleic acids, three stranded or “Y”-shaped branched DNA,²¹⁵ wherein the branchpoint is riboadenosine, β -D-3'-deoxypsicothymidine,²¹⁶ or bases involving a linker²¹⁷ or even wherein branching is *via* base-to-base linkers.²¹⁸ There is also a specific interest in the

design and study of nucleic acids containing two pyrimidine domains and that have the ability to form triple helical complexes when they encounter single stranded nucleic acids (“clamp effect”)²¹² (**Figure 3.2.1**). One domain serves as a duplex forming region hybridizing with purine rich target sequences through Watson-Crick hydrogen bonds while the other forms Hoogsteen hydrogen bonds with the common purine strand.^{219,220,221,222} Thus triplex forming nucleic acids can be classified as circular, hairpins, fold back and V-type molecules as schematically represented in **Figure 3.2.1**.

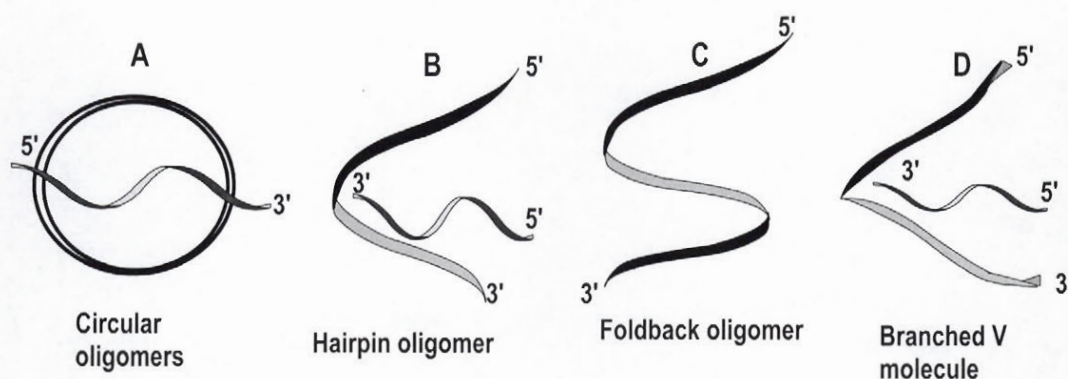


Figure 3.2.1: Types of intra and intermolecular triplexes; filled ribbon: Watson-Crick purine strand; shaded ribbon: Watson-Crick pyrimidine strand.

Recent investigations pertaining to branched and dendritic oligonucleotides have been carried out in Damha's group. Such branched RNA (bRNA) containing vicinal 2'-5' and 3'-5' phosphodiester linkages on a single sugar moiety, were first detected in nuclear polyadenylated RNA from HeLa cells by Wallace and Edmonds.²²³ Branched “V” oligonucleotides (with two binding domains connected) are choice candidates because they possess less freedom of internal bond rotations than their linear precursors, and as a result are expected to complex to a target with a lower entropic cost.^{220,221,224}

Indeed Damha's group has shown that joining the Watson-Crick and Hoogsteen components of a potential triple helix to an adenosine branch-point was found to promote the formation of parallel T•A:T triple helices²²⁵ and the less common antiparallel T*A:T (**Figure 3.2.2**).¹¹⁶ The following notation, henceforth will be used in this thesis. (“:” Watson-Crick bonding; “•” Hoogsteen; and “*” reverse-Hoogsteen). In two of these

studies the "V"-shaped compound A^{T10}_{T10} (**3.20a** and **3.20b**), consisted of a riboadenosine branch-point nucleoside with 2' and 3' "dT₁₀ tails". In the V-compound **3.20a**, the 2' tail runs parallel to, or has the same polarity as the 3' tail ($A^{5'T103'}_{5'T103'}$), while in **3.20b**, the polarities of the two tails are opposite each other.

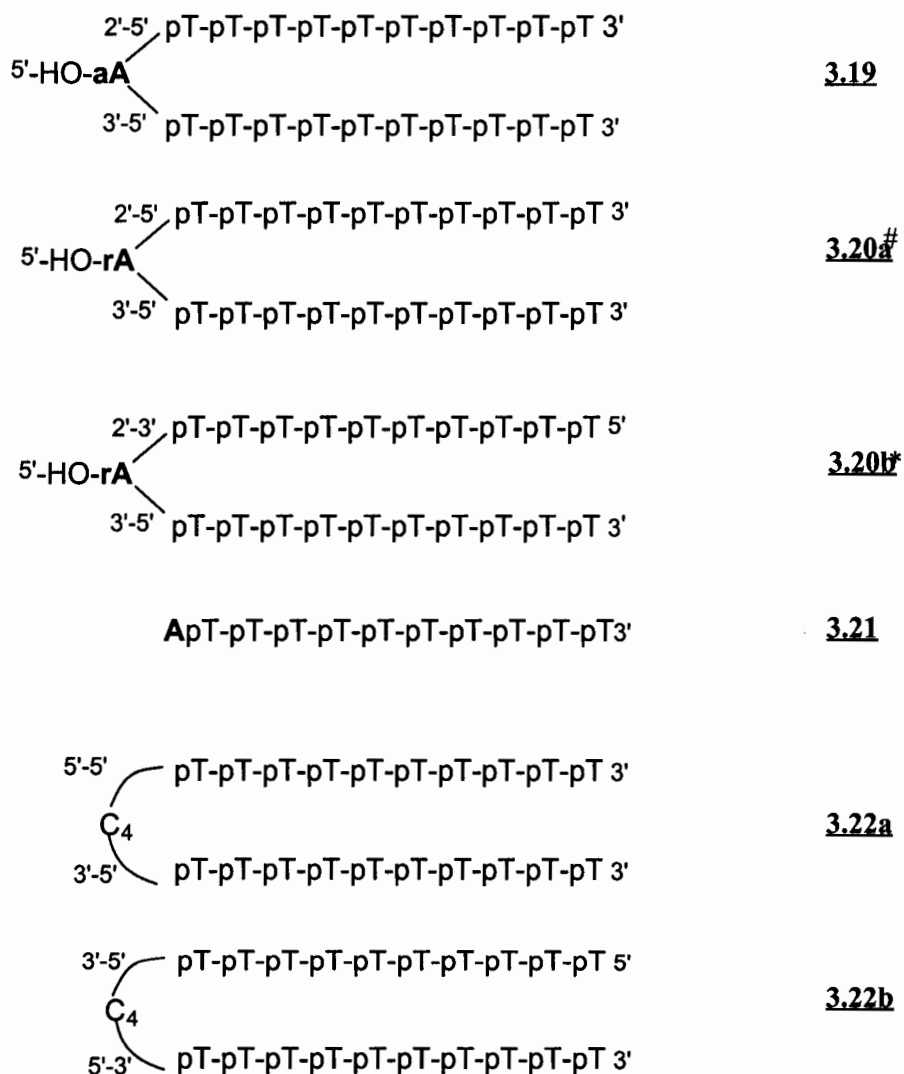


Figure 3.2.2: Schematic representation of the branched molecules and looped controls used in this study. aA and rA represent arabinoadenosine and riboadenosine, # refers to reference 116 and * to reference 225.

The thermal stabilities of triplex helical complexes formed with the V molecules were much higher than that of the linear trimolecular $T_{10} \bullet A_{10} : T_{10}$ complex as expected.¹¹⁶ The ability of these two “V” molecules in forming stable triple-stranded helical complexes with dA_{10} via T^*AT and $T \bullet A : T$ triplets respectively, prompted the investigation of molecule **3.19** wherein the branchpoint nucleoside is arabinoadenosine (aA) instead of riboadenosine (rA).

Structural Implications

The Hoogsteen $T \bullet A : T$ triplex is the most studied and understood of these triplex systems.⁹⁹ It is characterized by the regular Watson-Crick duplex and a third pyrimidine strand that is hydrogen bonded (Hoogsteen) to its major groove, *parallel* to the duplex purine strand (Chapter 1, Section 2).²²⁶ Divalent cations such as Mg^{2+} and Mn^{2+} , and higher concentrations of the univalent cations, such as Na^+ (1-2M) promote such triplex formation.²²⁷ CD spectroscopy reveals very distinct signatures for the $T_{10} \bullet dA_{10} : T_{10}$ triplex system and the underlying duplex $dA_{10} : dT_{10}$. Key features exhibiting intermolecular triplex formation are: (a) A characteristic depression of the positive amplitudes at *ca.* 225, 259 and 284 nm; (b) an amplitude increase of the negative Cotton effect *ca.* 248 nm; and (c) the presence of a unique negative CD band at 208 (pp 21-24).²²⁸

Figure 3.2.3 illustrates the branched molecule **3.19** with its arabinoadenosine branch-point. Recent studies on small branched RNA fragments *e.g.* the trinucleotide diphosphates $rA^{[2'-5'U]}_{[3'-5'U]}$ and $aA^{[2'-5'rU]}_{[3'-5'rU]}$, indicate that the sugar-phosphate framework of the adenosine residue is rigid, a result of strong base stacking between the adenine at the branch-point and adjacent 2'-uridine.²²⁹ The 2'-5' and 3'-5' phosphodiester linkages of $rA^{[2'-5'U]}_{[3'-5'U]}$, are nearly parallel to one another (*i.e.* display significant preference for the ϵ^- and ϵ'^- conformation about the C3'-O3' and C2'-O2' bonds as shown in **Figure 3.2.3**). The furanose ring of the arabinoadenosine residue shows a high preference for the C2'-endo pucker conformation a common feature of purine sugars linked to a pyrimidine *via* a 2'-5' linkage.

The branched molecule **3.19** when complexed with dA₁₀, would involve antiparallel T*A:T triplets, and much less is known about this system. Like **3.20a** it contains dT₁₀ tails at each of the 2' and 3' positions of the branch-point nucleoside. The polarities of the dT₁₀ tails are identical being 5' to 3' from the core extending outwards, a consequence of the synthesis methodology employed. Compound **3.19** differs from compound **3.20a** only at the branch-point nucleoside, specifically its 2' dT₁₀-tail has the “ β ” configuration as opposed to α in **3.20a**.

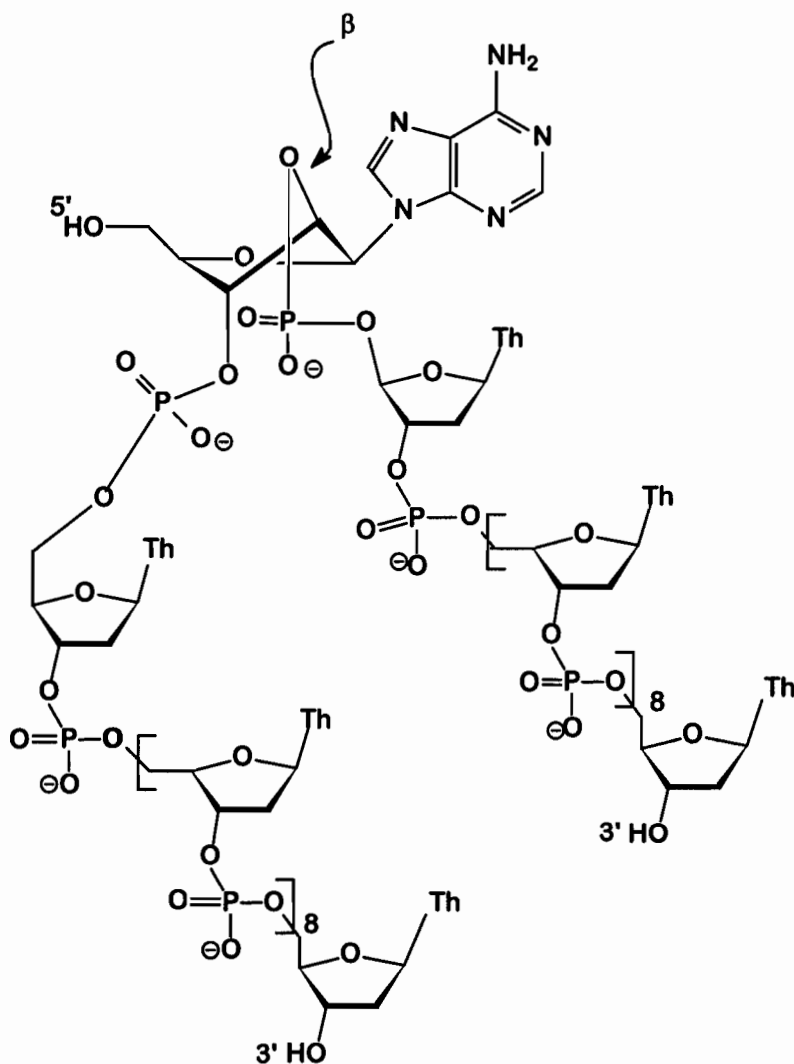


Figure 3.2.3: Primary structure of branched oligomer **3.19** showing the conformation for the branch-point arabinoadenosine unit.

The topology of **3.19** is yet to be determined but, like that of **3.20a**, can be considered as a linear molecule with a 'kink' (Figure 3.2.1). Thus this section deals with the association properties and biophysical characterization of molecule **3.19**. Soon after this work was completed a similar report had appeared involving the arabinonucleoside 2'-methyluracil as branch-point.²³⁰

3.2.2 Results And Discussion

Association of **3.19** with deoxyadenylic acid

The UV melting profiles were recorded in a buffer (10 mM TRIS-HCl, 50 mM MgCl₂, pH 7.3) which favors triple helix formation.^{64,75,208} The melting curve at 260 nm, for the complex formed between compound **3.19** and dA₁₀ showed a monophasic, cooperative transition at 32.9°C and a hyperchromicity of 15% (Figure 3.2.4). The stability of this complex is comparable to that of the duplex formed between dT₁₀ and dA₁₀ ($T_m = 32.2^\circ\text{C}$) as well as the putative triplex formed between **3.20** and dA₁₀ (32.7°C).^{99,231}

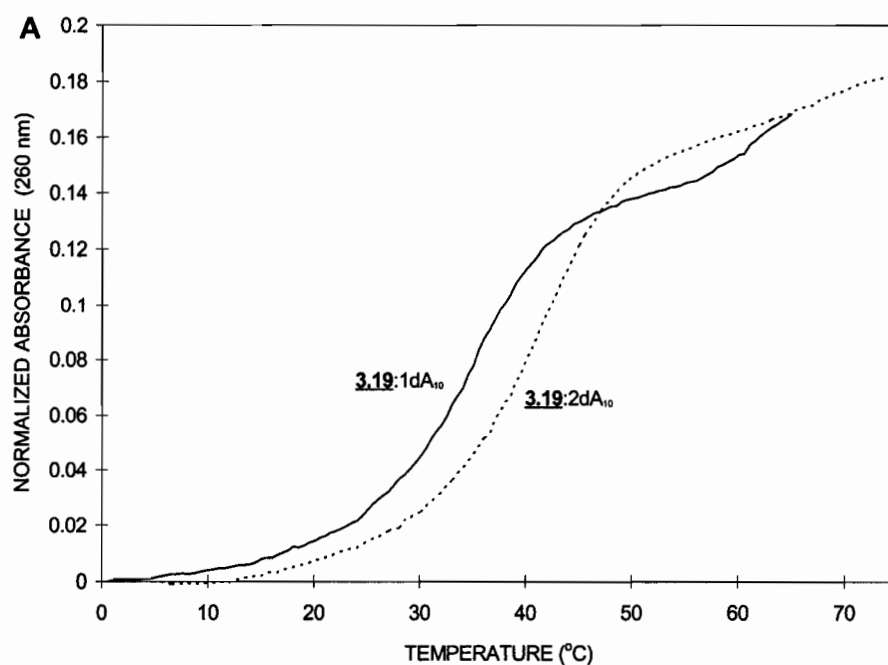


Figure 3.2.4 A: Melting curves of complex **3.19**:dA₁₀ in 50 mM MgCl₂, 10 mM TRIS-HCl pH 7.3, with 1 and 2 equivalents dA₁₀ at 260 nm.

Supporting a triple helical structure for **3.19** + dA₁₀ are the observations of a detectable melting transition of the complexes at 284 nm, a wavelength at which pyr•pur:pyr triplexes composed entirely of T/A:T base triads (both Hoogsteen and reverse-Hoogsteen) display significant changes in absorbance, but at which duplex A:T pairs do not (**Figure 3.2.4 B**). Addition of a second dA₁₀ equivalent to the **3.19** + dA₁₀ (1:1) mixture increased the T_m by *ca.* 4°C, similar to what was observed for **3.20a**:dA₁₀ (1:2). This implies that there exists some obstruction in a full length base pairing of the **3.19** with dA₁₀ which could only be complete in the presence of the second dA₁₀ strand.

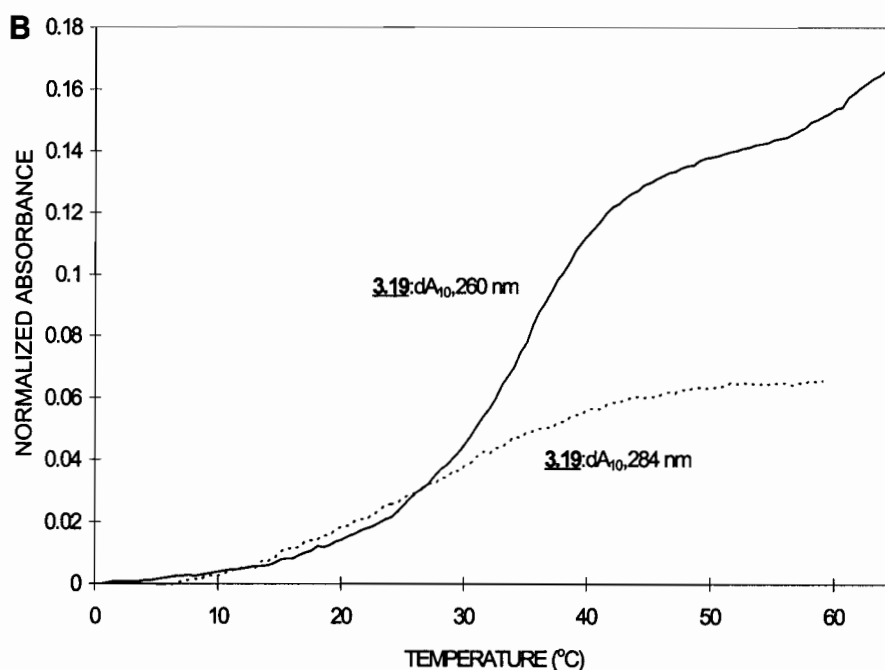


Figure 3.2.4 B: Melting curves of complex **3.19**:dA₁₀ in 50 mM MgCl₂, 10 mM TRIS-HCl pH 7.3 with 1 equivalent of dA₁₀ at 260 and 284 nm.

Effect of Cations on Triplex Formation

(i) *Magnesium:* The promotion of triplex formation in Mg²⁺ buffers is often interpreted as a manifestation of the ability of Mg²⁺ to efficiently neutralize the phosphate↔phosphate backbone repulsions in DNA triple helices.⁷⁵ Mg²⁺ ions have a smaller activity coefficient (0.07) than Na⁺ (0.37) and thus can bind to the duplex DNA

more tightly relative to Na^+ .²³² As a result, the experiments were conducted with Mg^{2+} buffer unless otherwise specified.

(ii) *Manganese*: Fox and coworkers have reported that within the context of G*G:C triplets the order of antiparallel triplex stability was $\text{T}^*\text{A}:\text{T} (\text{Mn}^{2+}) > \text{T}^*\text{A}:\text{T} (\text{Mg}^{2+})$. They also showed that $\text{T}^*\text{A}:\text{T}$ triplets are not strong enough to form in the absence of either a Hoogsteen helper section (*i.e.* $\text{T}\bullet\text{A}:\text{T}$) or reverse-Hoogsteen section (G*G:C). From **Table 3.2.1 A** the T_m s of the complexes formed with dA_{10} all increased on substituting Mg^{2+} for Mn^{2+} (compare buffer B versus C). These results suggest that the reverse-Hoogsteen $\text{T}^*\text{A}:\text{T}$ triplex appear to form in manganese without the helper effects mentioned above. To confirm the formation of such triplexes in Mn^{2+} , thermal studies were conducted at 284 nm, and transitions were indeed detected for only **3.19a**, **3.20a** and **3.22b** with the dA_{10} complement but not for **3.21** and **3.22a**. Based the above observations it is possible that the rigid branch-point of the V molecules is sufficient to allow triplex formation in Mn^{2+} . Triple helix formation is also supported by the experiments described below.

Table 3.2.1 A: Melting Temperatures (T_m) and Hyperchromicity (%H) for Complexes in Magnesium and Manganese Based Buffers^{a-c}

Complex	Buffer ^a		Buffer ^b		Buffer ^c	
	T_m (°C)	%H	T_m (°C)	%H	T_m (°C)	%H
Oligomer: dA_{10}						
$\text{aA}^{\text{T}_{10}}_{\text{T}_{10}}$ (3.19)	32.9	15	28.4	19	31.7	22
$\text{rA}^{\text{T}_{10}}_{\text{T}_{10}}$ (3.20a)	32.2	17	26.4	16	33.2	28
dAT_{10} (3.21)	32.2	19	27.6	9	30.0	19
$3'\text{T}_{10}\text{C}_4\text{T}_{10}3'$ (3.22a)	36.6	20	28.0	14	34.8	20
$5'\text{T}_{10}\text{C}_4\text{T}_{10}3'$ (3.22b)	46.5	23	40.0	21	40.8	18

Conditions: ^a 50mM MgCl_2 , 10 mM TRIS-HCl, pH 7.3; ^b 5 mM MgCl_2 , 10 mM NaCl, 10mM TRIS-HCl, pH 7.5. ^c 5 mM MnCl_2 , 10 mM NaCl, 10mM TRIS-HCl, pH 7.5. Concentration of each strand was 1 μM . Error limits for individual measurements are estimated at $\pm 0.5^\circ\text{C}$.

(iii) *Potassium*: The formation of complexes under physiological type conditions were then investigated. The **3.19** and **3.20a** molecules did not appear to form triplexes with

dA₁₀ in the presence of potassium (K⁺) as exemplified by the absence of the 284 profiles and a drop of *ca.* 10°C in T_m values relative to those in the magnesium buffer (**Table 3.2.1.B**). This is consistent with previous reports which suggest that K⁺ ions are strongly inhibitory to reverse-Hoogsteen triplex formation when present in a solution containing both K⁺ and Mg²⁺ ions at or above millimolar concentrations.^{77,78,233} Not surprising the second dA₁₀ equivalent gave a negligible increase in T_m and H% for both V molecules.

Table 3.2.1 B: Melting Temperatures (T_m) and Hyperchromicity (%H) for Complexes in a Buffer which Approximates the Intracellular Cationic Environment

Complex Oligomer: dA ₁₀	Buffer	
	T _m (°C)	%H
aA ^{T10} _{T10} (3.19)	21.0	11
rA ^{T10} _{T10} (3.20a)	21.9	11
dAT ₁₀ (3.21)	22.8	14
3'T ₁₀ C ₄ T ₁₀ 3' (3.22a)	19.3	11
5'T ₁₀ C ₄ T ₁₀ 3' (3.22b)	28.4	18

Conditions: 140 mM KCl, 5 mM Na₂HPO₄, 1mM MgCl₂, pH 7.2. Concentration of each strand was 1 μM. Error limits for individual measurements are estimated at ± 0.5°C.

Stoichiometry of Interactions²³⁴

To confirm that **3.19** can in fact form a triplex structure, titrations of **3.19** with dA₁₀ were carried out in both, the Mg²⁺ and K⁺ buffers (**Figure 3.2.5**). The control **3.20a** which is known to form a 1:1 complex with dA₁₀ was also titrated alongside for comparison.¹¹⁶ As shown in **Figure 3.2.5A & B** a break in the absorbance plot occurred at the point where dA₁₀ and the **3.19** and **3.20** concentrations were equal (Mg²⁺ buffer). No further break was found as the concentration of the adenylate oligomer was increased. This result strongly suggests that the two compounds **3.19** and **3.20** form a 1:1 complex with dA₁₀, wherein both dT₁₀ “tails” of **3.19** associate simultaneously with the dA₁₀ target to form a triple helix.

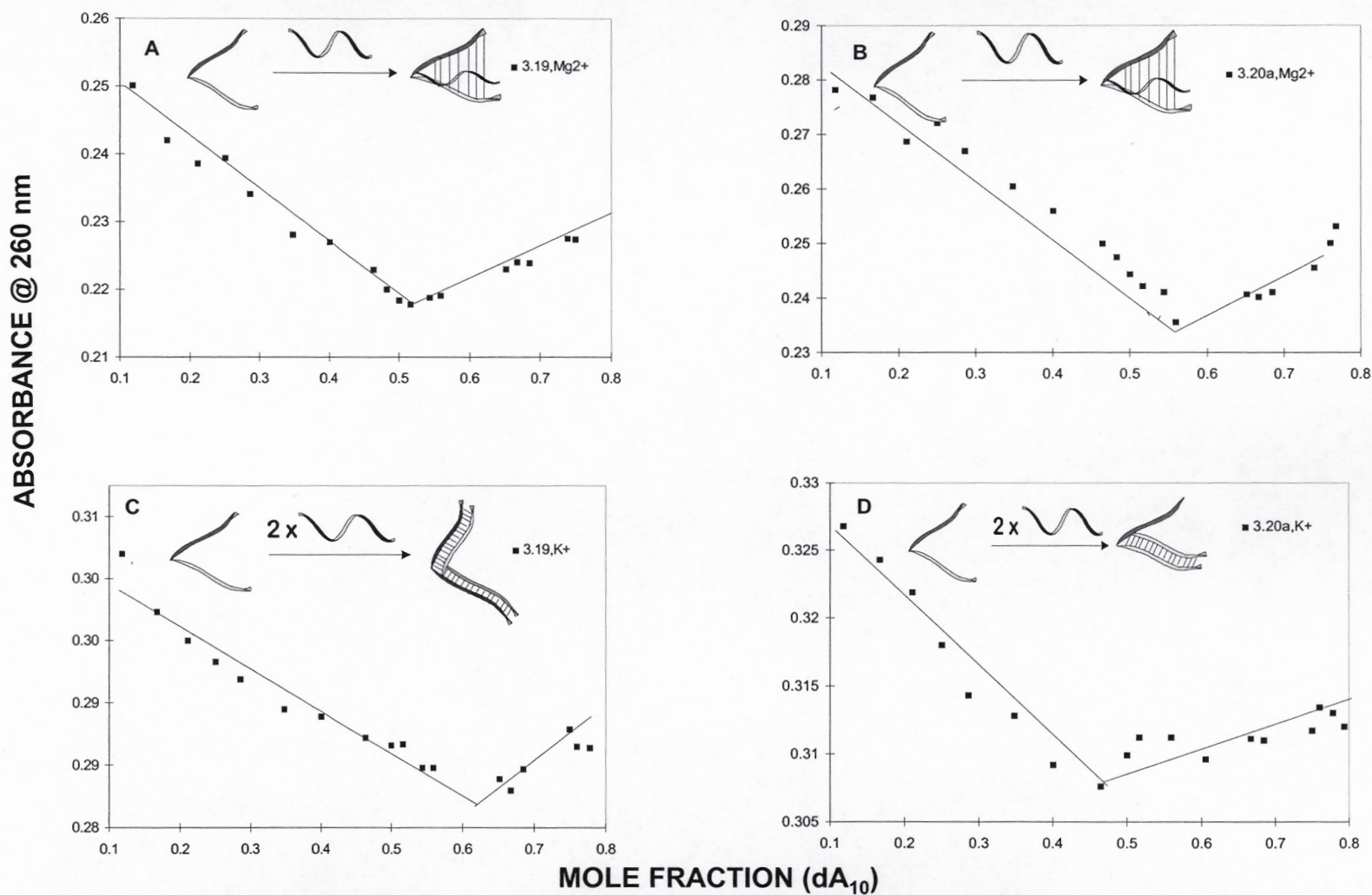


Figure 3.2.5: Determination of the stoichiometric interaction for **3.19** and **3.20a** with dA_{10} by the method of continuous variation. **A** and **B** are in 50 mM MgCl_2 10 mM TRIS HCl, pH 7.2, whereas **C** and **D** are in conditions that represent intracellular media, namely 140 mM KCl, 5 mM Na_2HPO_4 , 1 mM MgCl_2 , pH 7.2; refer to the experimental section for details.²³⁴

As opposed to a third strand binding to a preformed duplex, the “V” molecules behave like a clamp sequestering dA₁₀ in a single all-or none process. This is supported by the monophasic curves, wherein triplex formation requires the presence of a dA₁₀ complementary strand, to first form the Watson-Crick duplex with one of the arms, and Hoogsteen hydrogen bonds with the other arm (**Figure 3.2.4**).

Figure 3.2.5 C, reveals that in the presence of K⁺, a single inflection point is observed when the molar ratio of **3.19**:dA₁₀ is 1:2 consistent with the formation of a duplex in this case. This is in contrast to what is observed for **3.20a** which under the K⁺ conditions binds only one equivalent of dA₁₀ (break at 50%, **Figure 3.2.5 D**). Two scenarios can be therefore envisioned: (i) **3.20a** cannot take up the second dA₁₀ because it is too crowded around the branch-point, a consequence of the stereochemistry of the ribose 2'carbon. Thus the second arm of the **3.20a** remains unhybridized (single stranded), an observation that agrees with the 284 profiles and Job plots; (ii) in **3.19** one of the dT₁₀ tails (2') can bend upwards to accommodate an additional dA₁₀ molecule.

Circular Dichroic Spectroscopy

Complex formation is readily monitored by changes in both the amplitude (magnitude) and position of the CD bands. Compound **3.19** with dA₁₀ at 5°C showed a CD spectrum qualitatively similar to that of **3.20a**:dA₁₀ in Mg²⁺, but only the results of the arabino system will be discussed here. The CD spectrum of the dA:dT duplex system exhibits two maxima at 218 nm (large) and 283 nm (moderate), two minima at 208 nm and 249 nm, a crossover at 258 nm which occurs near the λ_{max} observed in the UV spectrum. The negative 208 nm band has been attributed to the dA_n component of the A:T complex.²³⁵ These features are also present in the triplex. However, as the triplex:duplex ratio increases (for example, as dA₁₀ + duplex \rightleftharpoons triplex equilibrium is shifted to the right) notable changes are obvious. These are the relative increase in amplitudes of negative 208 and 249 nm bands and the corresponding more obvious decreases of the positive 283nm and 219 nm bands. The binding of **3.19** and dA₁₀ is also accompanied by the appearance of a distinct shoulder at 260 nm, a characteristic for dT/dA:dT triplexes.^{99,102}

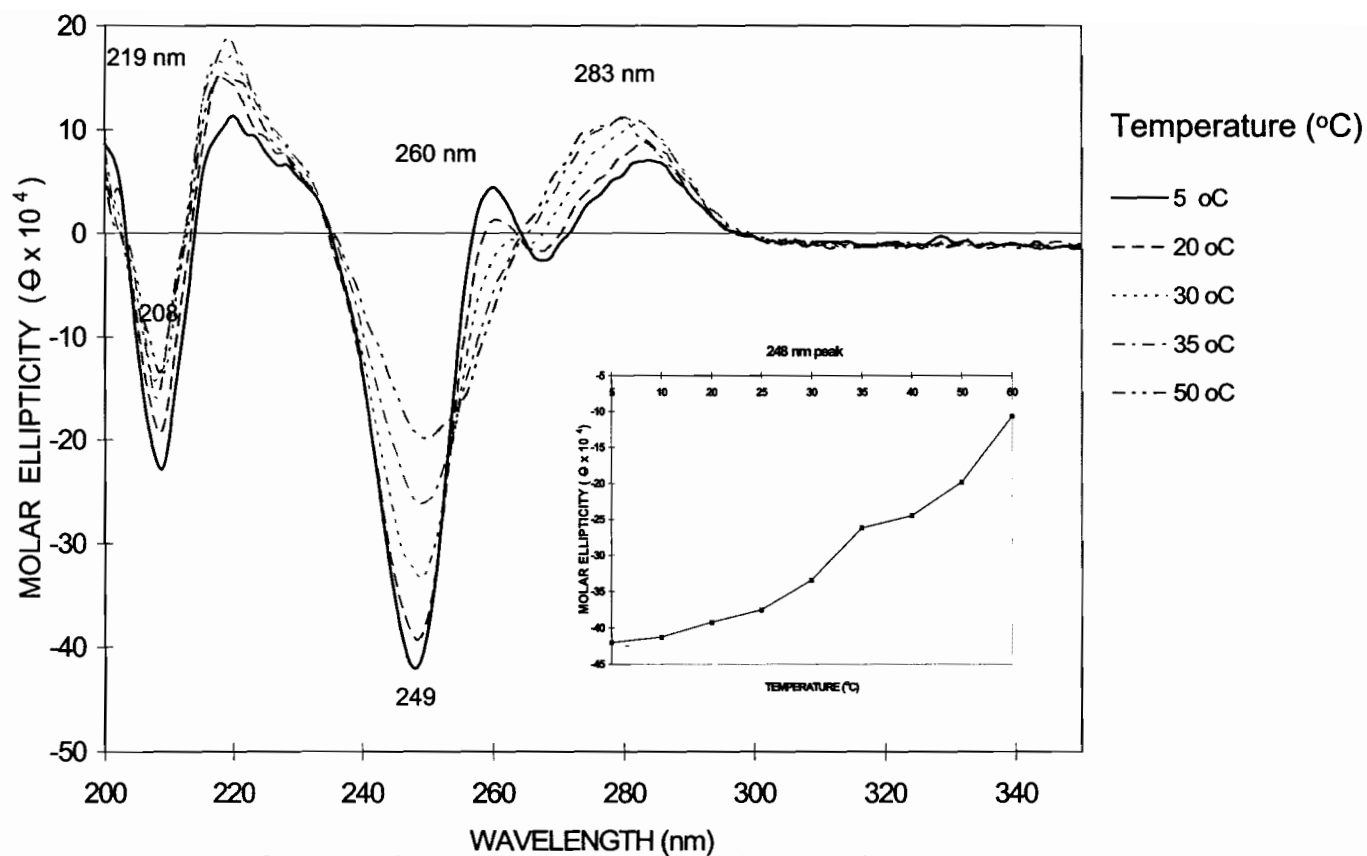


Figure 3.2.6: CD spectra as a function of temperature for the complex **3.19**:dA₁₀ in 50 mM MgCl₂, 10 mM TRIS-HCl, pH 7.3. The inset shows a change in molar ellipticity $\Delta \theta$ (10^{-4}) at 248 nm as a function of temperature.

It has been documented that the major influence on CD spectra of poly [d(T):d(A)] by addition of the third poly d(T) strand is the apparent loss of the positive 217 nm CD band,²³⁵ a feature which is believed to be indicative of triplex formation for this particular sequence.⁹⁹ For the given study the 217 peak corresponds to one observed at 219 nm (Figure 3.2.6 and 3.2.7).

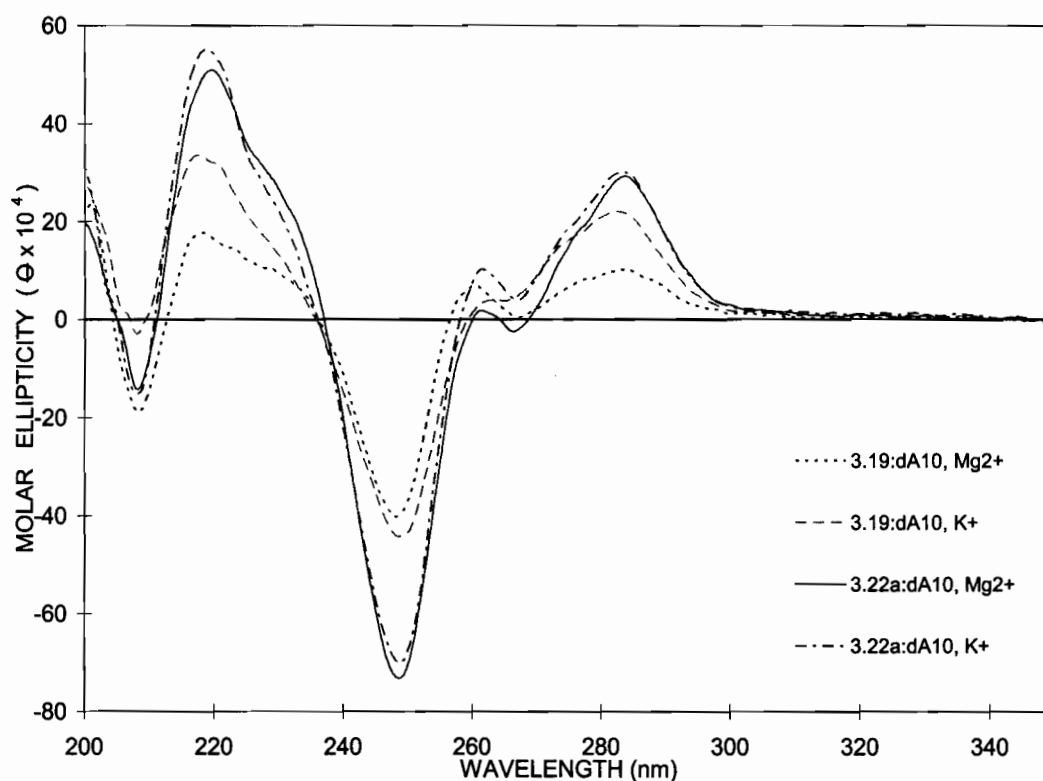


Figure 3.2.7: Comparative CD spectra of the complexes formed by molecules **3.19** and **3.22a** with complementary dA₁₀ (1 equivalent) in magnesium (50 mM MgCl₂, 10 mM TRIS-HCl, pH 7.2) and physiological media (140 mM KCl, 5 mM Na₂HPO₄, 1 mM MgCl₂, pH 7.2).

Figure 3.2.6 shows the effect of increasing temperature on the CD spectrum of complex **3.19**:dA₁₀. As the temperature was increased from 5°C to 60°C, marked changes in both amplitude and wavelength occurred, *i.e.*, there was an apparent loss of the negative bands at 208 nm and 248 nm, and a concurrent amplitude increase of the positive bands at 283 nm and 219 nm. The spectral changes were more pronounced in the 30°- 35°C temperature range, which correspond to the T_m of the **3.19**: dA₁₀ complex measured by UV spectroscopy. The other distinct feature is the marked disappearance of

the shoulder at 259 nm and blue shift of the band at 283 nm. The 1:1 mixture of **3.19**:dA₁₀ in Mg⁺² (60° C > T_m) displays a spectrum that is similar to a superimposition of the spectra of the free, single stranded **3.19** and dA₁₀ (data not shown). The inset displays the change in molar ellipticity as a function of temperature at 248 nm. The greatest change occurred around 35°C, corroborating the fact that the complex melt around this temperature.²³⁶

From the comparisons between the complexation of **3.19** and **3.22a** to dA₁₀ in two buffers the following conclusions can be made: The CD spectrum of **3.19**:dA₁₀ and **3.22a**:dA₁₀ in K⁺ buffer is similar (**Figure 3.2.7**) to that of the duplex formed by dT₁₀ + dA₁₀ (not shown) indicating that a double-stranded helical structure is formed predominantly in this buffer. This further confirms the results obtained from thermal denaturation profiles, and UV mixing curves. *i.e.* K⁺ does not promote triplex formation for **3.19**:dA₁₀. Furthermore the spectrum of **3.19**:dA₁₀ (1:2, K⁺) was almost identical to that of **3.19**:dA₁₀ (1:1) except for a slightly broader shoulder at 260 nm in the former case, and a further drop in amplitude of the positive band at 283 nm (data not shown). Finally the branched architecture of **3.19** is required for triplex formation (under Mg²⁺) since the corresponding linear **3.22a** does not appear to form a triplex under the same conditions (compare CD spectra of **3.19**:dA₁₀ to **3.22a**:dA₁₀ in Mg²⁺, **Figure 3.2.7**).

Thermodynamic Data of Complex Formation

Thermodynamic data of triplex formation were obtained from melting curves by the concentration T_m variation method.²³⁷ As expected for an intermolecular process, the association of **3.19** and dA₁₀ displayed a concentration dependent melting denaturing profile. **Figure 3.2.8** shows a plot of reciprocal melting temperature $1/(T_m)$ versus the natural logarithm of the total strand concentration (2.16 μ M to 0.21 mM range). The van't Hoff plot can be fit to a straight line with a negative slope, supporting the view that the monophasic transition manifested by melting curves of **3.19** is a bimolecular, all-or none, process. **Table 3.2.2** summarizes the thermodynamic parameters derived from the above analysis. The thermodynamic parameters for denaturation of complexes **3.20b**:dA₁₀ are also included for comparison purposes. The latter system was studied by R. Braich of the Damha group.²²⁵

Table 3.2.2: Calculated Thermodynamic Parameters for the Hoogsteen and Reverse-Hoogsteen T/A:T Base Triplets

Thermodynamic Data	3.19	3.20b
T_m (°C)	33	45
ΔH° kcal/mol	-8.9	-8.3
ΔS° cal°/mol.	-26.1	-23.5
ΔG°_{25} kcal/mol	-1.1	-1.3
Binding of third strand	reverse-Hoogsteen	Hoogsteen

Conditions: 50mM MgCl₂, 10 mM TRIS-HCl, pH 7.3. Melting temperatures were determined using a total strand concentration that varied between (2.16 μ M to 0.21mM)

Although a direct comparison of the thermodynamic data of the two branched complexes is compromised by the different branch-point structure (ribo vs. ara) some conclusions can be drawn from **Table 3.2.2**. The enthalpy change of association for **3.19**:dA₁₀ is very similar to that of **3.20b**:dA₁₀. This value of 8.3 kcal/mol of base triplets for **3.20b**:dA₁₀²²⁵ is in excellent agreement with those for the dT₁₀•dA₁₀:dT₁₀ intermolecular triplex, under the same conditions.¹⁰² The similarity between all three complexes (**3.19**:dA₁₀, **3.20b**:dA₁₀ and dT₁₀•dA₁₀:dT₁₀) could be taken as an indicator of

the presence of 10 hydrogen bonded base triplets in the branched helical complexes. The enthalpy and entropy changes for triplex formation involving **3.19** were more negative than those of **3.20b**. This implies that whereas the **3.19**:dA₁₀ complex has the most favorable pairing enthalpy change (suggesting better stacking, hydration, bonding etc.), it also has a more unfavorable pairing entropy change ΔS° (suggesting a better stacked geometry of the complex relative to the single stands). This unfavorable standard entropy change is compensated by the favored enthalpy change and the resulting free energy change (ΔG°_{25}) is relatively close to that of **3.20b**. The calculated standard free energy of **3.20b**:dA₁₀ formation corresponds to the observed melting behaviors in which **3.20b**:dA₁₀ shows the highest melting temperature. The greater stability observed for the **3.20b**:dA₁₀ triplex (Hoogsteen, W-C) relative **3.19a**:dA₁₀ (reverse-Hoogsteen, W-C) is consistent with the notion that reverse-Hoogsteen interactions are of known lower stability relative to Hoogsteen interactions.^{116,225}

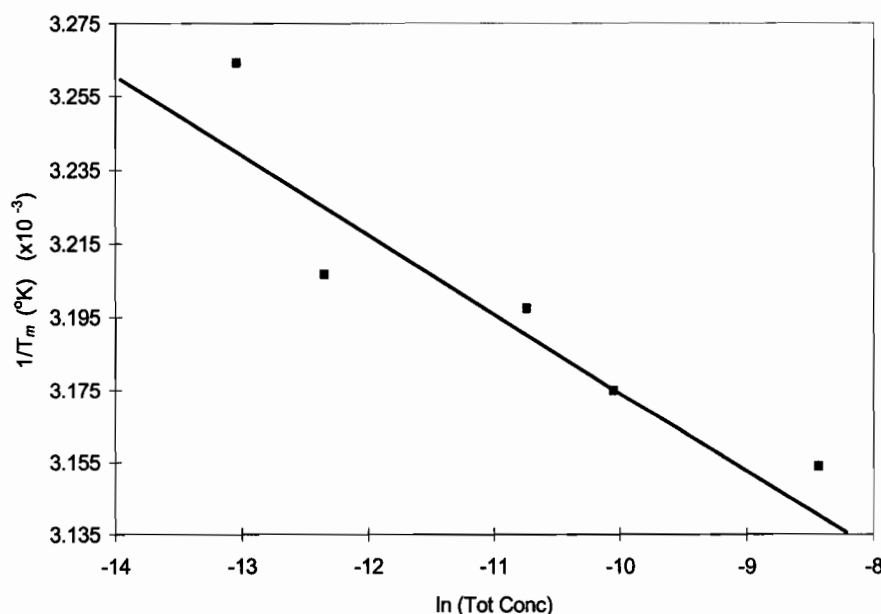


Figure 3.2.8: Concentration dependence of melting temperature (T_m) for the complex **3.19**:dA₁₀ in 50 mM MgCl₂, 10 mM TRIS-HCl pH 7.2. Melting temperatures were determined using a total strand concentration that varied between (2.16 μ M to 0.21 mM).

Effect of Target Chain Length on Binding and Base Pairing

The study described in the preceding sections involved investigating differences in T_m as a function of increasing amounts of the complement dA_n . This section deals with T_m measurements as a function of increasing length of the dA_n target when both the branched molecule **3.19** and the dA_n target were in a 1:1 ratio (strand:strand and A base:T base respectively).

Table 3.2.3: Thermal Melt Data for the aA^{T10}_{T10} (3.19**): dA_n Complexes ($n = 3-10$)**

Target strand (dA_n)	Equivalents of dA_n	T_m (°C)	H%
dA_3	1.0	-	-
dA_3	3.3	16.1	4.2
dA_4	1.0	broad	7.9
dA_4	2.5	23.1	8.9
dA_5	1.0	22.6	15.4
dA_5	2.0	24.5	17.3
dA_7	1.0	24.1	14.6
dA_7	1.4	27.3	14.6
dA_{10}	1.0	32.0	13.9
dA_{10}	2.0	35.2	16.3

Conditions: 50mM $MgCl_2$, 10 mM TRIS-HCl, pH 7.3. Error limits for individual measurements are estimated at $\pm 0.5^\circ C$. Concentration of each strand was 1 μM .

Examination of the data obtained (Table 3.2.3) leads to the following conclusions: 1 strand equivalent of dA_3 and dA_4 bound weakly or not at all to **3.19** as illustrated by ambiguous broad transitions obtained. Larger sequences namely dA_5 , dA_7 and dA_{10} displayed enhanced binding upon addition of more equivalents. For example a slight increase in T_m was observed when the second equivalent of dA_7 and dA_{10} was added. Two things can account for this, (i) cooperativity which is the enhancement in the specific binding of a ligand to a site on DNA, due to the presence of similar ligands, bound at neighboring sites²³⁸ and (ii) an increase in the number of Watson-Crick

hydrogen bonds formed, as well as base stacking of the terminal thymine bases in the third strand of the triple helix.²³⁹

A question that arises in this study is whether all T residues in the branched compound **3.19** are involved in base-pairing interactions. This question was addressed by measuring complex stability with dA_n oligomers of different chain lengths (*i.e.* $n = 3, 5, 7, 10, 15, 18, 20$ and 24) with **3.19** and the control **3.22a**. Figure 3.2.9 illustrates the relationship between the melting temperatures of these complexes (**3.19**: dA_n and **3.22a**: dA_n) as a function of adenylate chain length “ n ” where both target and dA_n oligomer are in a strand:strand ratio of 1:1.

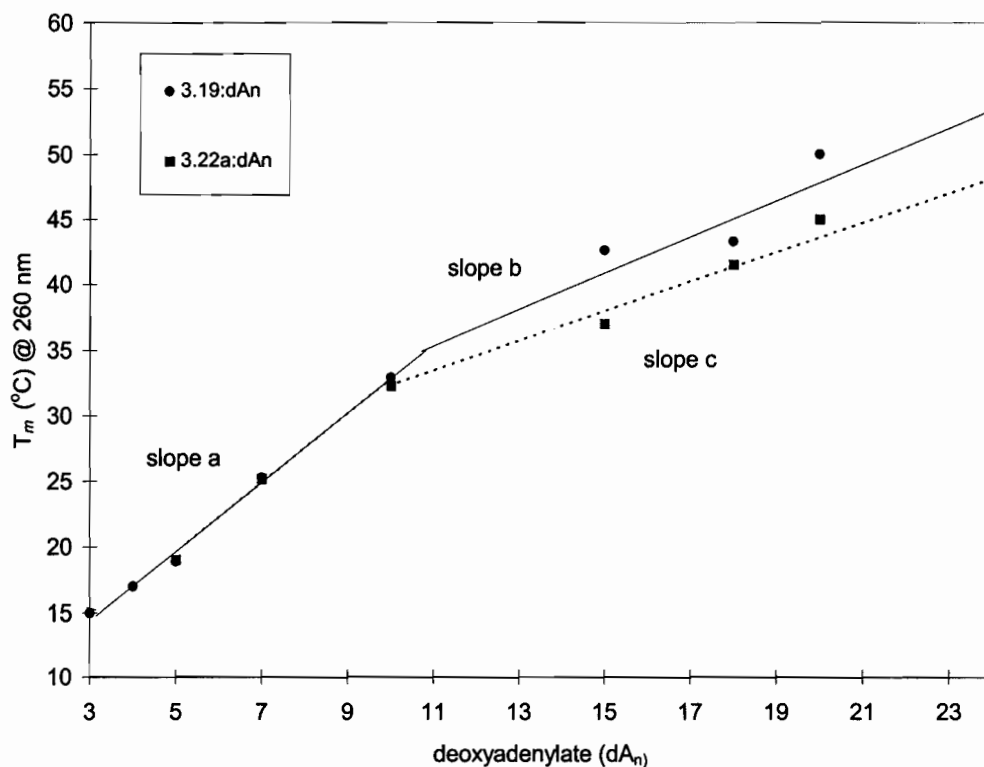


Figure: 3.2.9: The effects of varying the adenylate chain length (n) on the melting temperature of complex of **3.19** and **3.22a** with 1 equivalent of dA_n ($n = 3$ to 24). The calculated values of slopes a, b and c are 2.63, 1.42 and 1.14 °C/nt respectively.

For complexes involving the branched oligomers **3.19** and **3.22a** with dA_n , the T_m s are linearly dependent on target chain length (n) over a limited range. There are two obvious ranges ($5 \leq n \leq 10$) and ($n > 10$). In the first range the behavior of the two are very similar. Both lines then change slopes at n close to 10 which approximately corresponds to the number of thymines present in each arm of the branched. Beyond this value the T_m s of the complexes still remain linearly dependent on n , but the slope for **3.19** is more steep. This most likely reflects differences in the stability of the complexes and supports the idea that, in the case of **3.19**: dA_n the adenylate strand is held to one dT_{10} by Watson-Crick bonding and to the other strand by reverse-Hoogsteen bonding. The increasingly higher T_m values beyond $n = 10$ are most likely the effect of stability as a result of reduced fraying at the ends of the complexes.²⁴⁰ Significant hyperchromicity at 260 nm and to a lesser extent at 284 nm was likewise observed for both **3.19** and **3.22a**, characteristic of $dT \bullet dA : dT$ triplexes.²⁴¹ However the 284 transitions for **3.22a** were very weak relative to **3.19**.

Pairing Selectivity: Effect of Mismatch on the Target dA_n Sequence

In order to study the effects of base mismatches on triplex formation, UV melting curves were recorded for **3.19**:target and **3.21**:target complexes containing a mismatch in the centre of the target purine strand or at the residue that base pairs with the branch-point adenosine (Table 3.2.4 and Figure 3.2.10).

(i) Mismatch within the target purine strand

The nucleotide at position 5 in the oligoadenylate (dA_{10}) was changed to a dT and dG and the corresponding thermal stabilities measured. Both mismatches led to a decrease in T_m of 9-11°C, reinforcing the stringency of proper base pairing requirements in the antiparallel T*AT triplex, as reported previously (Table 3.2.4).¹¹⁶ A mismatch could presumably disrupt base stacking either by preventing a base from aligning favorably with respect to the base across the junction, or it could disrupt favorable electrostatic interactions by changing the electronic distribution of the mismatched base itself.²⁴² The T*G:T mismatch appears to be less stable than the T*T:T mismatch, in the **3.19**: dA_{10} complex. Moreover triplexes are usually equally or more sensitive to base triplet mismatches than are Watson-Crick duplexes to corresponding base pair

mismatches, because of lower stabilities than those of duplexes.^{209,210} In the case presented here the T_m decrease for the mismatch “duplexes” (dAT₁₀:T₁₀) was found to be more significant (*ca.* 17°C) than those observed for the triplexes. This probably reflects the intramolecular nature of the reverse-Hoogsteen strand which serves not only to increase the local concentration of the third strand and but also to compensate for the destabilization created by the mismatches present in the dA_n strand.

Table 3.2.4: ΔT_m and %H Values from Melting Curves (260 nm) of Mismatched Relative to Matched Complexes

Target Oligomer	3.19		3.21	
	ΔT_m (°C)	$\Delta H\%$	ΔT_m (°C)	$\Delta H\%$
5'-AAAAAAAAAAA 3'	0	0	0	0
5'-AAAATAAAAA 3'	-8.1	-2.0	-17.0	-0.5
5'-AAAAGAAAAA 3'	-11.0	-0.9	-16.9	-2.0

Conditions: 50 mM MgCl₂, 10 TRIS-HCl, pH 7.2. Error limits for individual measurements are estimated at $\pm 0.5^\circ\text{C}$. Concentration of each strand was 1 μM .

(ii) Mismatch at the branch-point

To determine whether the arabinoadenosine branch-point is involved in any base pairing interactions, **3.19** was allowed to hybridize with target DNA sequences 5'-d(A₁₀C₃)3' and 5'-d(A₁₀TC₂)3' in the presence of Mg²⁺ or Na⁺ (see **Figure 3.2.10**). The complex **3.19**:d(A₁₀C₃) contains a rA/dC mismatch and thus would be expected to have a lower T_m relative to the fully complementary complex **3.19**:d(A₁₀TC₂). Indeed, the lower melting temperature observed for **3.19**:d(A₁₀C₃) is consistent with the notion that the branch-point adenosine can base pair. The contributions of the branch-point aA:dT pair to the thermal stability of the complex is similar to the one observed by Hudson *et al.* for **3:20a**:d(A₁₀C₃).¹¹⁶

Mismatch effects were more pronounced in Na⁺ buffer than in Mg²⁺ which may be explained on the basis of magnesium's ability to better shield the ionic repulsion of the sugar phosphate backbone. It appears that the role of cation on stability of the triplex

is more important than the C-A(branch-point) mismatched triplex probably because of a fraying or end effect of such a mismatch. A central mismatch (see previous section) has a greater effect on T_m than one at the end as in $d(A_{10}C_3)$ consistent with previous findings.²⁴³




	Complexes	T_m (°C)
I		32.1°C (Mg^{2+}) 29.7°C (Na^+)
II		35.7°C (Mg^{2+}) 35.3°C (Na^+)
III		33.6°C (Mg^{2+}) 27.5°C (Na^+)

Figure 3.2.10: Effect of the target chain length and sequence on the T_m of branched complexes. Experimental conditions: Mg^{2+} (50 mM TRIS-HCl, pH 7.2) and Na^+ (1M NaCl, 100 mM Na_2HPO_4 , pH 7.2).

Effect of Sugar Composition of Adenylate Strand on Triplex Forming Ability

In order to study the effect of strand composition on triplex stability the deoxyadenylate (dA_{10}) target was replaced by riboadenylate (rA_{10}) and hybridized to **3.19** under different buffer conditions. As shown in **Table 3.2.5**, the complex formed between **3.19** and rA_{10} was considerably weaker than that formed between **3.19** and dA_{10} (ΔT_m ca. -10°C). Addition of a second molar equivalent of rA_{10} resulted in a negligible T_m change. The absence of an observable transition at 284 nm, as well as a lack of key triplex signatures in the CD spectrum of **3.19**: rA_{10} (data not shown) all point to the formation of a duplex in this case. This is consistent with the “Purine Exclusion Principle” which states that a ribopurine (rPu) third strand cannot form either a rPu •duplex or a rPu*duplex triple helical complex.²⁴⁴ Moreover it is generally believed that sodium enhances duplex formation by increasing the rate of triplex dissociation, possibly

because additional energy has to be expended to distort a stable duplex in the process, further reflecting duplex formation.⁸⁵

Table 3.2.5: Melting Temperatures (T_m) and Hyperchromicity (%H) for Complexes Formed Between 3.19 and Decaadenylic acid, rA_{10} .

Oligomer	Buffer ^a	Target rA_{10}					
		1 equivalent			2 equivalents		
		T_m (°C)	(ΔT_m) (°C) ^b	%H	T_m (°C)	(ΔT_m) (°C)	%H
<u>3.19</u>	S	27.3	-2.4	12	29.1	-3.7	14
	M	23.3	-8.8	14	25	-11	15

^aBuffer S: 1M NaCl, 100mM Na_2HPO_4 , pH 7.1. Buffer M: 50 mM $MgCl_2$, 10 mM TRIS-HCl, pH 7.2. ^b $(\Delta T_m) = (T_m \text{ of } \underline{3.19}:rA_{10}) - (T_m \underline{3.19}:dA_{10})$. Concentration of each strand was 1 μ M. Error limits for individual measurements are estimated at $\pm 0.5^\circ C$.

This study is presently being extended to the use of arabinonucleic acids (and analogues, *e.g.*, 2'-fluoro-ANA) in the complementary (third) strands.

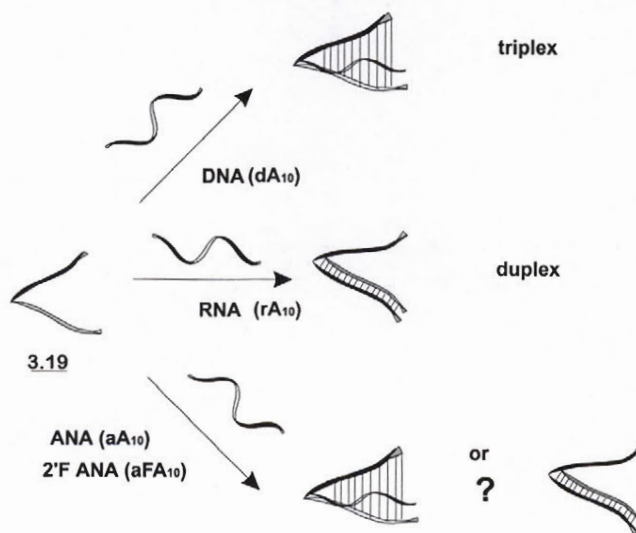


Figure 3.2.11: Complex formation between 3.19 and adenylic acids dA_{10} , rA_{10} , aA_{10} and aFA_{10} .

3.2.3 STABILIZATION OF T*AT (ANTIPARALLEL) TRIPLEXES VIA DNA BINDING LIGANDS

Introduction

The use of triple helix-forming oligonucleotides is often limited by the low stability of triple helical complexes relative to double helices, especially under *physiological* conditions. Triplexes have a bigger negative potential and hence their requirement for positive cations such as Mg^{2+} , Mn^{2+} and spermine is imperative.²⁴⁵ The recognition of the potential roles and biomedical applications of triplex structures has now focused considerable attention on increasing triplex stability, as well as on their interactions with nucleic acid-binding ligands. A triplex specific ligand is a small molecule designed to promote the duplex-to-triplex transition, displacing the equilibrium in favor of the triplex.

The ligand 'Hoechst 33258' developed by Luck and coworkers in 1984,²⁴⁶ is a synthetic compound that contains two consecutive benzimidazole rings with a phenolic and a N-methylpiperazine group at either end of the U-shaped elongated molecule (**Figure 3.2.12**). Hoechst, like other small molecules (netropsin, berenil, pentamidine and distamycin) bind in the minor groove of duplex DNA, with a marked preference for A:T-rich regions.²⁴⁷ In the case of Hoechst 33258, the minimum size required for binding is four consecutive A:T base pairs. Ligands often contain a series of linked aromatic rings structures which make them inherently flat and crescent in shape. The flexible nature of Hoechst's ring system and its cationic nature permits the molecule to adopt a conformation that follows the contours of the minor groove and thus optimizes binding to duplex DNA.²⁴⁸ A basic asymmetry also arises from the fact that Hoechst 33258 provides the donor hydrogen bonding groups, whereas the oligonucleotide molecule participates with its acceptor groups located at the bottom and walls of the DNA minor groove.

Benzo[e]pyridoles (BePIs) were the first reported molecules to bind triplexes more tightly than to duplexes, and indeed provide strong stabilization to such triplexes.^{249,250} BePI's are also crescent shaped, tetracyclic, aromatic compounds that optimize stacking interactions with base triplets. The stronger stacking interactions

present in BePI in BePI/ triplex relative to BePI/duplex is the key factor enhancing the stability triple helices relative to double helices. BePI interacts more strongly with T•A:T than C⁺•G:C triplets,²⁵¹ probably because electrostatic repulsions between BePI's positively charged side chain and protonated cytosines located in helix groove.

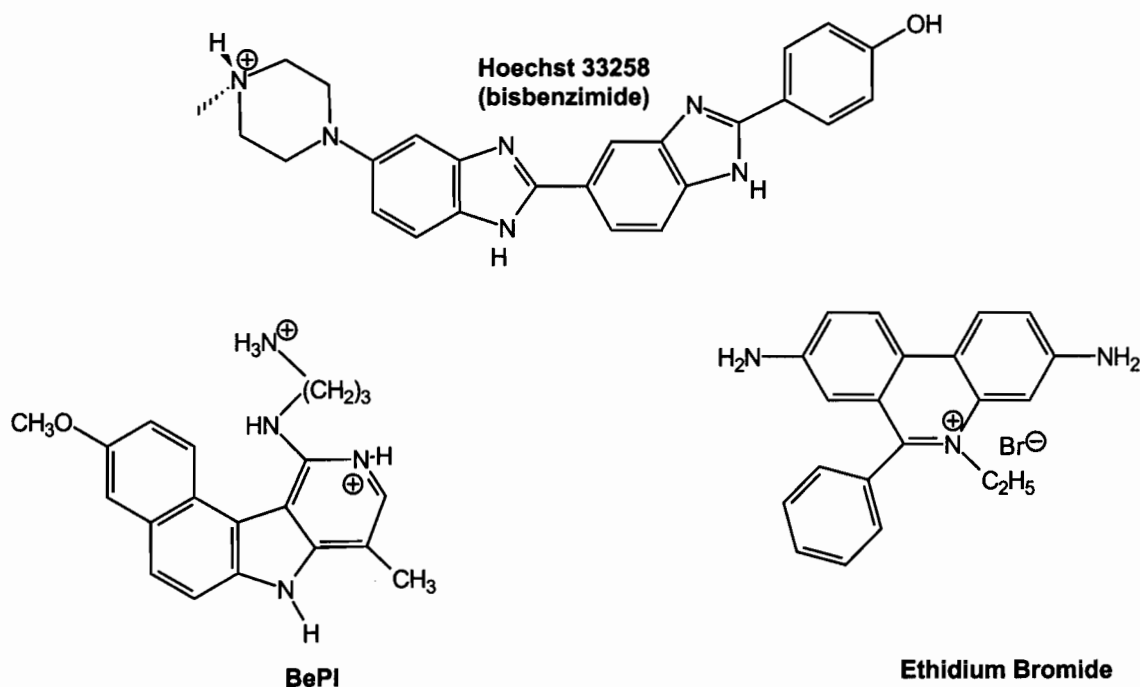


Figure 3.2.12: The structures of the DNA-binding ligands: Hoechst 33258, BePI and Ethidium Bromide

Ethidium Bromide (EtBr) is a cationic planar chromophore. It not only intercalates better but it also binds with a stronger affinity to the triple helix poly(dT)•poly(dA):poly(dT) than to the A-T duplex. The elegant and thorough work of Scaria and co-workers demonstrate that the nature of interaction of the EtBr to duplex gives rise to CD bands that are of opposite sign to those resulting from the more favorable EtBr/triplex interactions.²⁵²

This concept of ligand induced triplex stabilization, has been extended to several other related polycyclic aromatic systems such as the alkaloid Coralyne,²⁵³ unfused aromatic cations,^{254,255} and 2,6-disubstituted amidoanthraquinones.²⁵⁶ Conjugation of these ligands to the third strands at either end or internally, results in dramatic

improvements in triplex stability,²⁵⁷ with stable parallel (Pu•Pu:Py)triplexes being formed at conditions close to physiological conditions. Polyamines such as 1,2-diaminopropane and other endogenous amines (spermine), have been extensively characterized as triplex stabilizers,²⁵⁸ and have the additional advantage that they can be selectively transported into many tumor cells.

This study utilized BePI, EtBr and Hoechst 33258 to explore the triplex and duplexes formed by the **3.19**:dA₁₀ system in the hope of further understanding the nature of such complexes. The interaction of these molecules with the antiparallel T*A:T motif had not been investigated except for the preliminary study of **3.20a**:dA₁₀ by A. Uddin in Damha's group.²⁵⁹

Results and Discussion

A five-fold excess of each ligand namely, Hoechst 33258, Ethidium and BePI was added to the preformed complexes (triplexes and duplexes), and the mixture incubated for 1 day at 4°C. UV denaturing studies were run on the mixtures before and after the addition of the ligands. The differences in the two sets of T_m , indicates the extent of stability contributed by the ligand to the complex. The following molecules were also studied, which served as controls: dAT₁₀ (linear control), **3.22a** and **3.22b** (looped controls) (**Figure 3.2.2**). The latter two molecules have both Watson-Crick and Hoogsteen domains joined together *via* a loop of 4 deoxycytosine residues. These serve as linear and looped controls that lack the rigid adenosine branch core. **3.20a** was previously studied in the Damha's group and is also included for comparison purposes. The results are presented in **Figures 3.2.13** and **14** and can be summarized as follows: Hoechst addition stabilizes the **3.21**:dA₁₀ duplex, and to a lesser extent, the complexes formed by **3.19** and **3.20a**. It has a negligible effect on the triplex formed by **3.22b**:dA₁₀. All of these solutions exist as a dynamic equilibrium of both the duplex and triplex, the ratio of which is different in each situation.

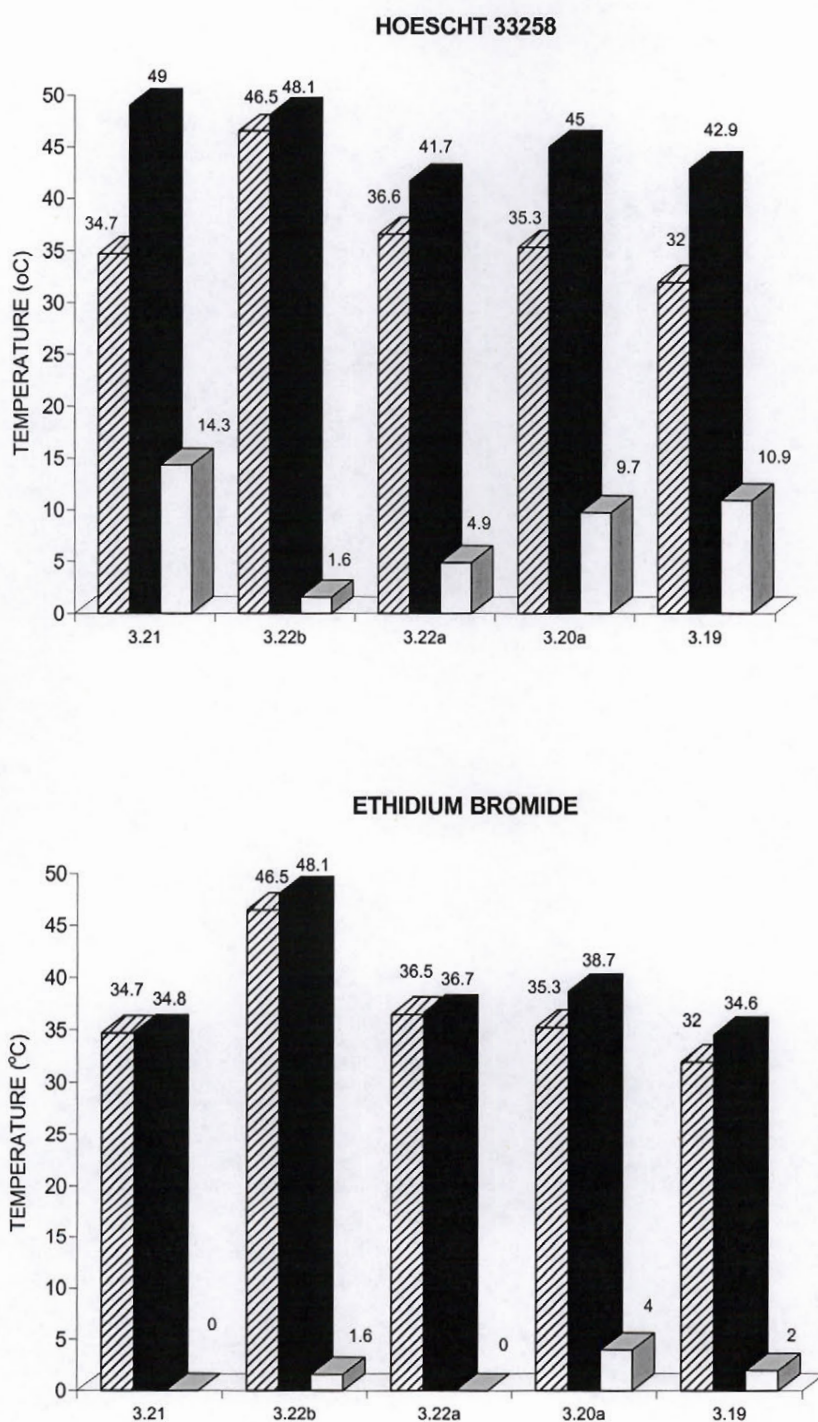


Figure 3.2.13: Effect of ligand induced stability on complex formation. (▨) T_m before addition of ligand, (■) T_m after addition of ligand and (□) ΔT_m i.e. difference in $T_m = (T_m \text{ of complex + ligand}) - (T_m \text{ of complexes in the absence of ligand})$. Ligands as given above. Conditions in magnesium buffer (50 mM $MgCl_2$, 10 mM TRIS-HCl, pH 7.2)

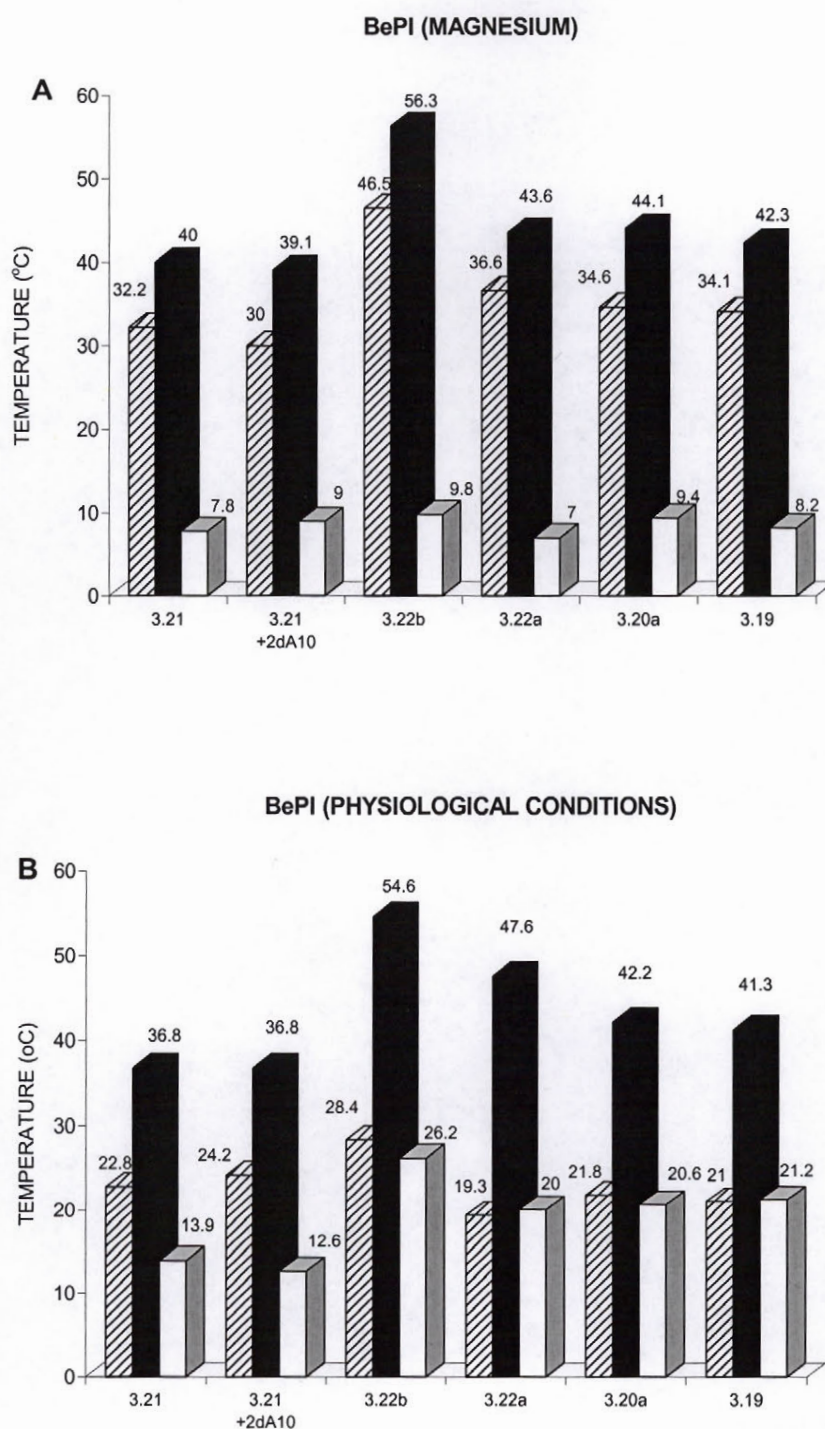


Figure 3.2.14: Effect of buffer on BePI induced complex formation. (▨) T_m before addition of BePI, (■) T_m after addition of BePI and (□) ΔT_m i.e. difference in $T_m = (T_m \text{ of complex} + \text{BePI}) - (T_m \text{ of complexes in the absence of BePI})$. Conditions (A) magnesium buffer (50 mM MgCl_2 , 10 mM TRIS-HCl, pH 7.2) and (B) physiological buffer (140 mM KCl, 5 mM Na_2HPO_4 , 1 mM MgCl_2 , pH 7.2).

It is believed that Hoechst stabilization of the duplex is not because it prevents triple helix formation but more so because triplex stability is reduced in preference of duplex formation.²⁶⁰

There is a slight increase in T_m when EtBr is added to the complexes of **3.19**, **3.20a** and **3.22b** with dA₁₀, but virtually no change is observed for **3.21**:dA₁₀ and **3.22a** dA₁₀. This differential stabilization almost categorizes the molecules by their triplex forming abilities: **3.19** and **3.20a** appear to form triplexes while **3.22a** like **3.21** appears to form predominantly a duplex.

The most interesting observations were made in the case of BePI stabilization. The data shown in **Figure 3.2.13** suggests that BePI stabilization (Mg^{2+}) was most significant for **3.22b**:dA₁₀ followed by **3.22a**:dA₁₀ and then **3.19**:dA₁₀. Substituting Mg^{2+} with K^+ in the buffer, which was shown earlier to destabilize the triplex, increased the extent of stabilization of the **3.19**, **3.20a**, **3.22a** and **3.22b** to a value double the original (in the absence of the ligand). This suggests that these molecules which do not form triple helical complexes under the K^+ conditions (with the exception of **3.22b**), could be induced to do so in the presence of BePI under physiological conditions.

3.2.4 Conclusions

Hybridization, stoichiometric measurements and CD spectroscopy indicated that oligopyrimidines connected *via* vicinal phosphodiester linkages to an arabinoadenosine branch-point were capable of forming triple-helical complexes in the presence of complementary deoxyadenylate strands. These complexes and those reported earlier by Hudson and coworkers¹¹⁶ (riboadenosine branch-point), represent the lesser known “anti-parallel”, or T*A:T reverse-Hoogsteen/W-C triplex.

The length of the complementary deoxyadenylate is required to be equal to or greater than 5 nucleotide for a stable triplex to be observed. As expected, triplex stability increased with increasing chain length of the target adenylate.

A single mismatch in the complementary dA₁₀ strand decreased binding substantially, while mismatches in complements hybridizing to the branch-point nucleotides were less destabilizing.

Binding of ribopurine complements (rA_{10}) did not result in the formation of a triplex, but rather a double helix.

The presence of araA at the branch-point of **3.19** endows it with a peculiar characteristic not seen for branched molecules containing riboA. Due to the β orientation of the 2'-5' linked dT_{10} branch, **3.19** was found to accommodate a second molecule of dA_{10} under physiological (K^+) conditions. This was the only notable difference between **3.19** and its ribo counterpart **3.20a**.

BePI appears to induce formation of $\text{T}^*\text{A}:\text{T}$ in a medium that is representative of cationic physiological conditions (140 mM KCl, 1 mM MgCl_2 , 5 mM Na_2HPO_4 , pH 7.2)

In an applied approach, the capture of single-stranded target molecules *via* triple helix formation with *branched* nucleic acid analogues could represent a new model for the 'antisense' and 'antigene' strategy. The fixed structural orientation in the branched nucleic acids may give the opportunity to study the properties of uncommon triple helices, such as the one presented here. Also, further developments with this chemistry and the idea of an *arabinose* branch-point are anticipated to contribute much to our understanding of both the native role of branched RNA and the use of synthetic nucleic acids for the artificial manipulation of gene expression.

3.3 TRIPLE HELICES CONTAINING ARABINONUCLEOTIDES IN THE THIRD (HOOGSTEEN STRAND): EFFECT OF INVERTED STEREOCHEMISTRY AT THE 2'-POSITION OF THE SUGAR MOIETY

3.3.1 Introduction

Triple helix formation depends on Hoogsteen hydrogen bonds between thymine with an A:T base pair (T•A:T) and protonated cytosines with a G:C base pair (C⁺•G:C). However these interactions are optimally stable at non-physiological pH 5.6-6.0, which greatly limits oligonucleotides as candidates for antigene therapeutics. Several strategies employing modified cytosine bases have been applied to overcome the protonation barrier posed by C⁺•G:C triplet formation.²⁶¹ 5-Methyl cytosine is a good example but has only a limited ability in this regard. Examples of successful substitutions for cytosine which result in two hydrogen bonds to the guanine second strand at physiological pH include the 6-keto derivative of 2'-deoxy-5-methylcytosine²⁶² and N⁷-deoxyguanosine.²⁶³ The stabilizing properties of the latter were found to be sequence dependent, probably on account of its larger size compared with a pyrimidine base. The synthetically accessible cytosine mimetic 2-aminopyridine is considered more basic than cytosine, and thus when incorporated into a third strand, forms effective triplexes at physiological pH. Its unnatural α -anomer can be accommodated at certain positions within the third strand without loss of stability, and is stable in serum-containing media.²⁶⁴ Cytosine mimetics have not yet been used in biological studies, with the exception of 5-methyl cytosine itself, probably because of their synthetic inaccessibility. It is hoped that the more recently developed 2-aminopyridine will be used universally.

Another issue of concern deals with the contributions of the sugar-phosphate backbone to such triplex stability. A number of studies have focused on the stability of triple helices containing both RNA and DNA strand combinations.^{84f,85,86 87,190} Initially it was suggested that duplex DNA undergoes a change from a B-DNA to an A-DNA conformation upon triple helix formation,⁶⁸ but more recent evidence points to the formation of an altered B-DNA conformation.^{62e,265}

To be able to fully analyze the stability of triplex formation resulting from mixed combinations of DNA or RNA strands, one must first understand the basis of the stability of the DNA:RNA (D:R) “hybrid” duplexes. *i.e.* the exact accommodation of two the types of backbones in a hybrid molecule. The existence of DNA:RNA hybrid duplexes in important biological processes is well known and such structures cannot be underestimated. Known as Okazaki²⁶⁶ fragments RNA:DNA duplexes occur frequently in DNA replication forks and also during reverse transcription.²⁶⁷ A model DNA:RNA duplex in the A-form DNA:RNA hybrid has been built into the cleft of the RNase H and RNA polymerase active sites of reverse transcriptase.²⁶⁸ Despite the fact that very little is known about the structure and the stability of “hybrid” duplexes, a developing mode of disease treatment based on RNA targeting (“the antisense therapy”) has been a major focus in the past years (**Section 1.2 and 4.1**).

There are a number of different views regarding the actual conformation of DNA:RNA (D:R) hybrid duplexes. The most popular belief is that D:R duplexes belong the A conformational family,¹ while some researchers believe that they are intermediate (heteronomous) between pure A and B-type duplexes.²⁶⁹ Indeed, this intermediate, heteronomous structural feature of the homopolymer hybrid poly[r(A):d(T)] was confirmed by solid state ³¹P-NMR and Raman spectroscopy studies.²⁷⁰ The detailed study of Roberts and Crothers on hairpin duplexes wherein the complementary stem strands of the hairpin are pure RNA, pure DNA, or chimeras of all purine and pyrimidine stretches has been well documented and referenced (**Figure 3.3.1**). The order of stability of the four hairpin duplexes [DD, D(pu)R(py), R(pu)D(py), and RR], was found to be RR > RD > DD > DR. From these results and the evaluation of several other published hybrid data,^{85,86,87,190} it can be predicted that a hybrid duplex with a purine-rich RNA strand will have a CD spectrum, and probably conformation, similar to that of A-form duplexes. It is also predicted that such hybrids will be more stable than a corresponding hybrid duplex with fewer purine bases in the RNA strand. A number of intramolecular interactions of the ribose-2'-hydroxyl group are believed to contribute to this stabilization. Thus, one can safely conclude that DNA:RNA hybrids are polymorphous and their conformation may depend on various parameters, such as solvent conditions and base composition.

Roberts and Crothers⁸⁵ first described the sensitivity of triple helix stability to the backbone composition (DNA *versus* RNA). An RNA third strand was found to have strong effects on the stability of Py•Pu:Py or ‘pyrimidine motif’ triplexes. Importantly, an RNA pyrimidine third strand was found to bind to all four possible duplex combinations (*i.e.*, DD, DR, RD, RR), whereas a DNA pyrimidine third strand bound only when the polypurine strand of the duplex is DNA (*i.e.*, DD and DR). More recently Damha and Noronha have shown that 2’-5’ RNA (*R**) third strands were also found to mimic the effects of DNA third strands.²⁷¹ The structural basis for these effects is unknown. Two major hypotheses can be used to explain these observations: (a) DNA and 2’-5’ RNA favor the C2’-endo conformation and differ from RNA which favors the C3’-endo conformation.¹ These conformations are maintained when these third strands fit into a duplex major groove.^{272,273} For instance, Taillandier and co-workers have shown the existence of C2’-endo sugars in all three strands of the dT•dAdT (D•DD) triplex,²⁷² whereas C3’-endo sugars are observed in the third strand of R•DD triplexes.²⁷³ Replacing the duplex DNA purine strand to a RNA purine strand (*e.g.*, DD → RD) changes the sugar pucker from a C2’-endo to a C3’-endo form; this conformation would be compatible with the binding of an RNA third strand (C3’-endo), but not with the binding of DNA and 2’-5’ RNA strands (C2’-endo).⁸⁷ This would account for the observed selectivity of *D* and 2’-5’RNA strands for DD and DR duplexes, over RD and RR duplexes; (b) The sugars of an RNA third strand favor the formation of short contacts between the α-2’-OH groups and phosphate groups of the duplex purine strand hence the stabilization of R•DD triplexes compared to D•DD.⁸⁷ This proposed intermolecular contact may account for the observation that RNA (but not DNA and 2’-5’ RNA) recognizes RR and RD duplexes. Such a mechanism cannot occur with DNA, and may not be possible for 2’-5’ RNA (α-3’OH) strands since their sugar hydroxyl groups are oriented differently.

In order to gain a better understanding of the effect of sugar composition on triplex stability, arabinonucleic acids (ANA, or A) were employed to recognize pure duplex DNA, pure duplex RNA, and RNA:DNA hybrids. This section specifically deals with the questions “Can an oligoarabinonucleotide fit into a duplex major groove, and if

so, how is triplex stability influenced by the stereochemistry at the 2' position of the third strand? In terms of duplex recognition, does ANA 'mimic' a DNA third strand, or the regioisomeric RNA third strand?"

3.3.2 Experimental Design

To study the interaction of arabinonucleic acids with duplexes, the experimental design of Roberts and Crothers, which has been successful in analyzing the effects of DNA and RNA composition of 'pyrimidine motif' triple helices was adopted. The target duplexes are Pu:Py hairpins and contain the four possible combinations of DNA and RNA strands (designated DD, DR, RD and RR, where the first letter describes the 5'-homopurine stem strand, and the second letter the 3'-homopyrimidine sequence) (**Figure 3.3.1**). The oligoarabinopyrimidine strand (A) was synthesized to explore triple helix formation with the hairpin duplexes. For the purpose of comparisons, the known oligoribopyrimidine (R) and oligodeoxyribopyrimidine (D) sequences were also examined. To study the duplex formation of the third strands, the complementary DNA and RNA designated respectively by prime letters D' and R' were synthesized and hybridized appropriately. The ability of oligomers (D, R and A) to associate with either the hairpins (antigene strategy), or with the complementary D' and R' strands (antisense strategy) was determined from UV melting experiments, native gel electrophoresis and CD spectroscopy, in a solution containing 100 mM sodium acetate and 1 mM ethylenediamine tetraacetate (EDTA), pH 5.5 (acetate buffer).

Although CD spectra and gel electrophoresis mobilities do not give detailed structural information, they can provide information about the global conformations of the hybrid duplexes, especially on a comparative basis.

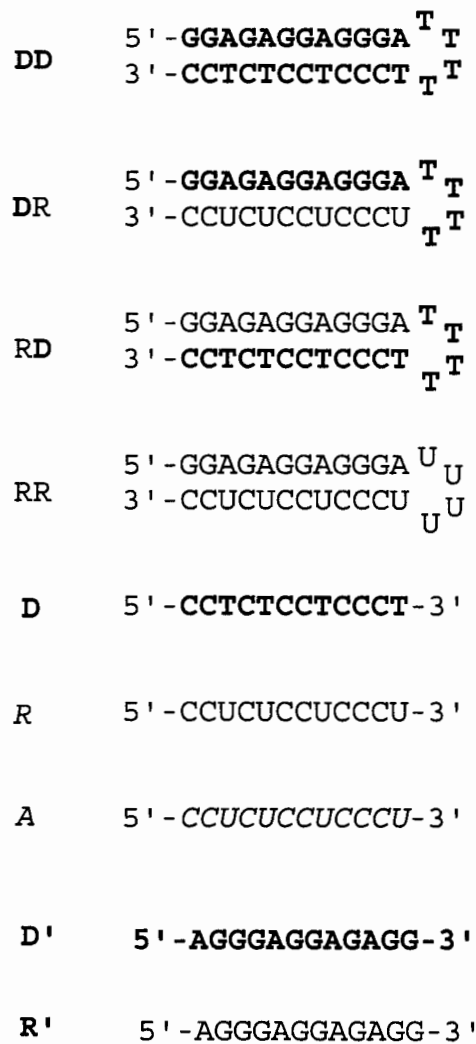


Figure 3.3.1: Hairpin duplexes (DD, DR, RD and RR), single strands (D, R and A) and complements (D' and R'); DNA sequences shown in bold and ANA italicized.

3.3.3 Results

(i) Characteristics of single strands D, R and A

UV melting curves were run to gain preliminary knowledge on the properties of the single stranded oligomers. In comparison to the D (DNA) Py strand which forms a fairly stable pH- dependent self-structure²⁷⁴ (T_m 23°C, CD spectrum, **Figure 3.3.2**) the self-association of A (ANA) was weak, while non-existent for the R (RNA) strand.

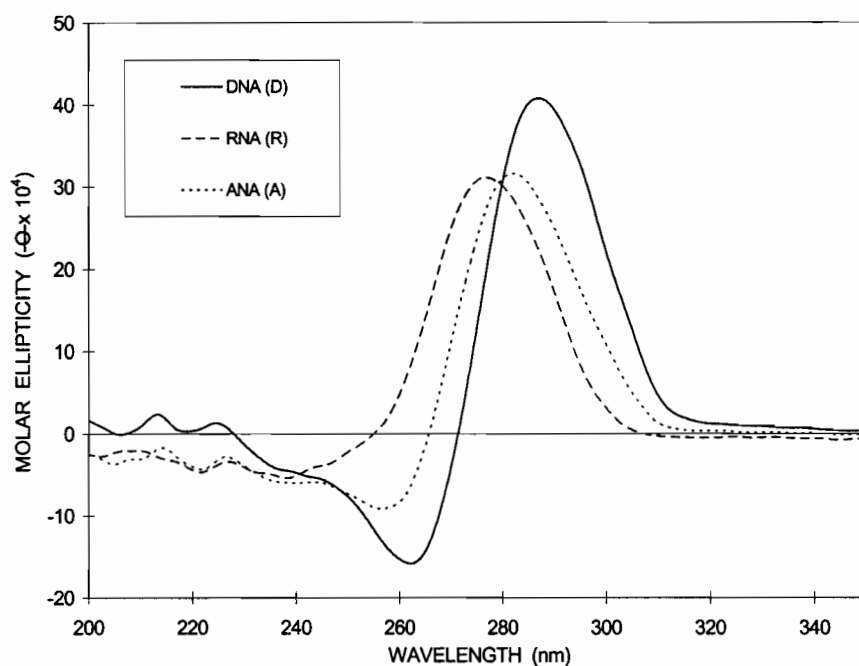


Figure 3.3.2: CD Spectra of single strands in 100 mM sodium acetate, 1mM EDTA, pH=5.5. Base sequences of D, R and A strands are given in **Figure 3.3.1**.

When tested alone on a *denaturing* gel, D, A and R appeared as a single, well-defined band with the expected electrophoretic mobility (**Figure 3.3.3 A**). The situation is different under *non-denaturing* conditions, where D and A (but not R) form self-structures detectable by the presence of numerous bands of low electrophoretic mobility (**Figure 3.3.3 B**).

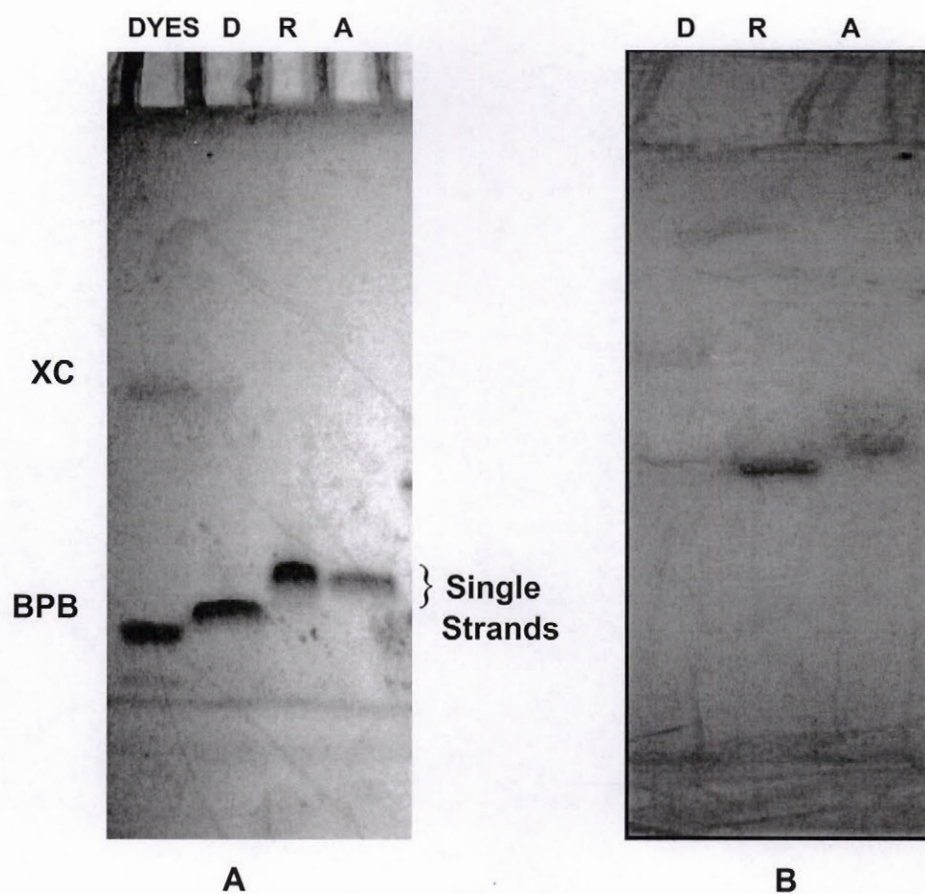


Figure 3.3.3: Gel shift mobility assay of single strands under (A) denaturing conditions, 7M urea, 14% polyacrylamide, pH 8; and (B) non-denaturing conditions, 7 M urea, 20% polyacrylamide, pH 5.0. Lanes are D (DNA), R (RNA), A (ANA) and the dyes xylene cyanol (XC) and bromophenol blue (BPB).

The self-associating structures of D and other deoxycytidine-rich sequences have been described in detail,^{275,276,277} and are believed to be very compact containing hemi protonated base pairs ($C^+ \bullet C$). Consistent with this notion, the CD spectra of A and D (**Figure 3.3.2**) exhibited a positive red- shifted band at *ca.* 282 nm, and a negative band centered at *ca.* 260 nm, which is strongly indicative of such $C^+ \bullet C$ base pairs.^{278,279,280} The D strand (T_m 23 °C, H 12%) appears to be more structured than the A strand (T_m 22°C, H 4%), as assessed by the amplitudes of the Cotton effects, gel mobility and thermal denaturation at both 260 and 300 nm.

(ii) Behavior of duplexes

The interaction of D, R and A with their corresponding complements (D' and R') gave varied results. The duplexes formed by the purine-rich R' strands always melted at higher temperatures than the corresponding duplexes formed by D' (purine). Only R:R', D:R', A:R', D:D' and R:D' (but not A:D') were detectable by UV profiles, and each displayed a single melting transition with similar curve shapes but with rather different T_m values. This is consistent with a two-state melting behavior. The order of thermal duplex stability under these conditions were $R:R' > D:R' > A:R' = D:D' > R:D'$ (**Table 3.3.1**). The stability of the hybrid duplex, A(py):R'(pu) was similar to that of the pure D(py):D'(pu) duplex, and nearly non existent for the A(py) + D'(pu), under the same conditions. Thus, ANA appears to bind to target ssRNA but not to target ssDNA.

Table 3.3.1: Thermal Denaturation of Duplexes

Pyrimidine Strand	Purine Target			
	D'		R'	
	T_m (°C)	H%	T_m (°C)	H%
D	48.6	20	65.7	15
R	45.1	11	71.3	10
A	22.1 ^a	6 ^a	48.7	9

All mixtures contained 2μM of each strand in 100 mM NaOAc, 1mM EDTA, pH 5.5 and adjusted with acetic acid; ^a weak transition.

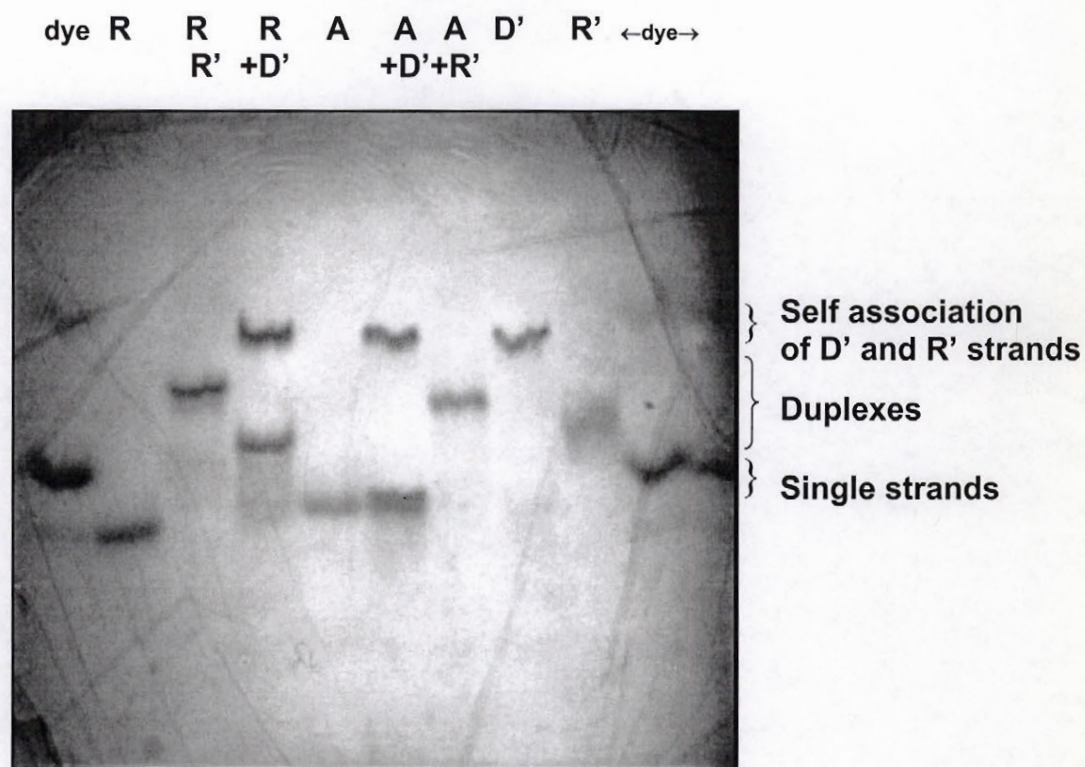


Figure 3.3.4: Gel shift mobility assay of mixtures of the thirds strands (D, R and A) and their complements (D' and R') as described above under non-denaturing conditions, 7M urea, 20% polyacrylamide, pH 7.5.

In an attempt to prevent the self-association of D or A strands when monitoring duplex formation with their complements, the pH was increased to 7.4 (100mM NaOAc, 1mM EDTA). Under these conditions the only apparent duplex formed with the D' target was R:D' (**Figure 3.3.4**) and D:D' (not shown). In the case of the R' target, all the three duplexes with D, R and A strands were observed to form (**Figure 3.3.4**, D:R' not shown). The migration of the A:R' duplex was similar to the R:R' duplex. However even under these conditions another curious observation was made: both the DNA (D') and RNA (R') purine strands self associate as evident from a slower migrating species in each lane, with the D' self structure being more prominent. This has previously been reported for d(AG)₃₀ strands below 20°C.²⁸¹ This feature further complicates the interpretation, *i.e.*, it is possible that A:D' and D:D' duplexes were not observed because the D' complex is thermodynamically more stable than those expected to form by the association of D' with A or D strands.

Figure 3.3.5 shows the CD spectra of the relevant duplexes. The D:D' duplex displayed a B-type CD spectrum with similar intensity positive (272 nm) and negative (237 nm) bands, while the R:R' duplex had a larger positive band at 270 nm and a smaller negative band at 222 nm, which is characteristic of an A-type duplex structure. The hybrid duplexes (D:R' and R:D') have CD features intermediate between pure A- and B- conformations. The positive peak of A:D' was red shifted to 281 nm. Based on the T_m and native gels and the fact that the positive peak appears at 281 nm, this spectrum could simply be derived in part from the self association of D' and/or the A single strands, without any complex formation. Other features of the A:R' spectrum included a prominent shoulder *ca.* 254 nm, a cross over at 250 nm, and a negative bands at 237 similar to the D(py):R'(pu) hybrid and R:R' duplexes. The positive peak of A:R' was shifted to 275 nm and the shoulder was almost non existent relative to those observed for D:R' and R:R'. It can be concluded from the CD studies, as well as Section 3.1 that ANA:RNA, like DNA:RNA hybrids, belong to a distinct structural class, that appears to be intermediate between the pure A and B type helices.

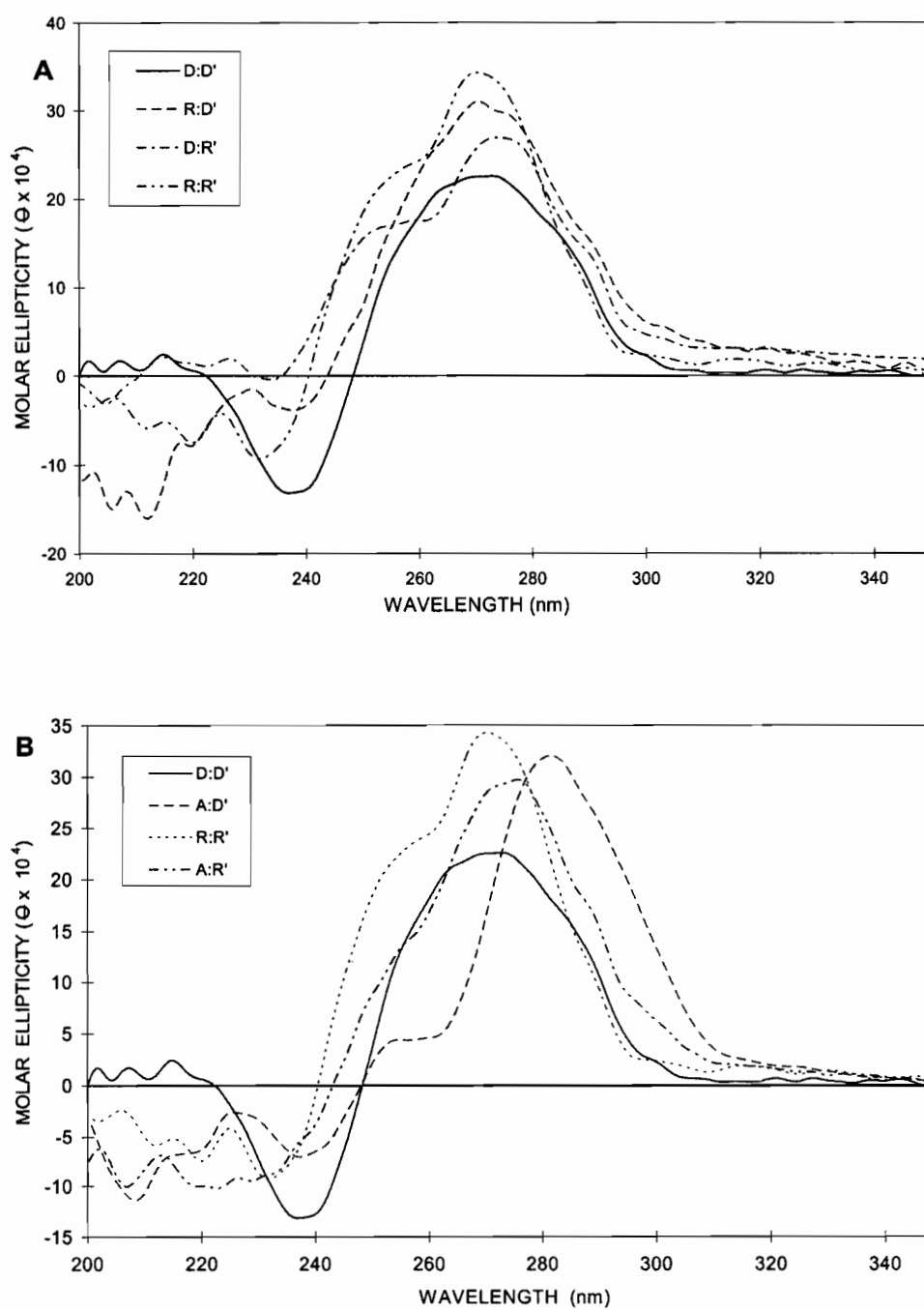


Figure 3.3.5 : Circular dichroism (CD) of duplexes at 5°C. Concentration is 2 μM of each strand, and the buffer is 100 mM sodium acetate, 1 mM EDTA, pH= 5.5.

(iii) *Triplexes*

The results of the melting experiments are shown in **Table 3.3.2** and **Figure 3.3.6**. Of the four possible triplexes containing A as the third strand, only A•DD and A•DR were observed to form. This was indicated by the presence of two transitions in the absorbance *versus* temperature profile when solutions containing equal concentrations of each A and duplex DD or DR were heated at a rate of 0.5°/min (**Figure 3.3.6**). The low-temperature transition corresponds to the dissociation of the arabino strand (A) from the target DD and DR duplexes. The high-temperature transition is assigned to the melting of the hairpin duplexes, since it was also observed when a solution of duplex alone was heated under identical conditions. This assignment is based upon the observation that the low-temperature transition disappears at neutral pH, whereas the high-temperature transition was essentially independent of pH over the range studied, as well as the presence of detectable absorbance changes at 300 nm (pH 5-7; data not shown). Melting of the arabino (A) strand would be expected to be sensitive to pH because its association involves C⁺•G:C triads in which the hydrogen bonded *arabinocytidine* residues are protonated. On addition of BePI to these mixtures at pH 5.5, only the second transition was observed. It is known that triplex formation that requires protonation of the cytosines in the third strand, disfavors intercalation of the positively charged BePI at sites involving the C⁺•G:C.²⁵⁰ As can be seen from the melting curves shown in **Figure 3.3.6**, the A strand has a higher affinity for the DR duplex than the DD duplex (T_m 43° vs. 34 °C) and was *ca.* 2-6 °C less than those for the corresponding triplexes formed by D.

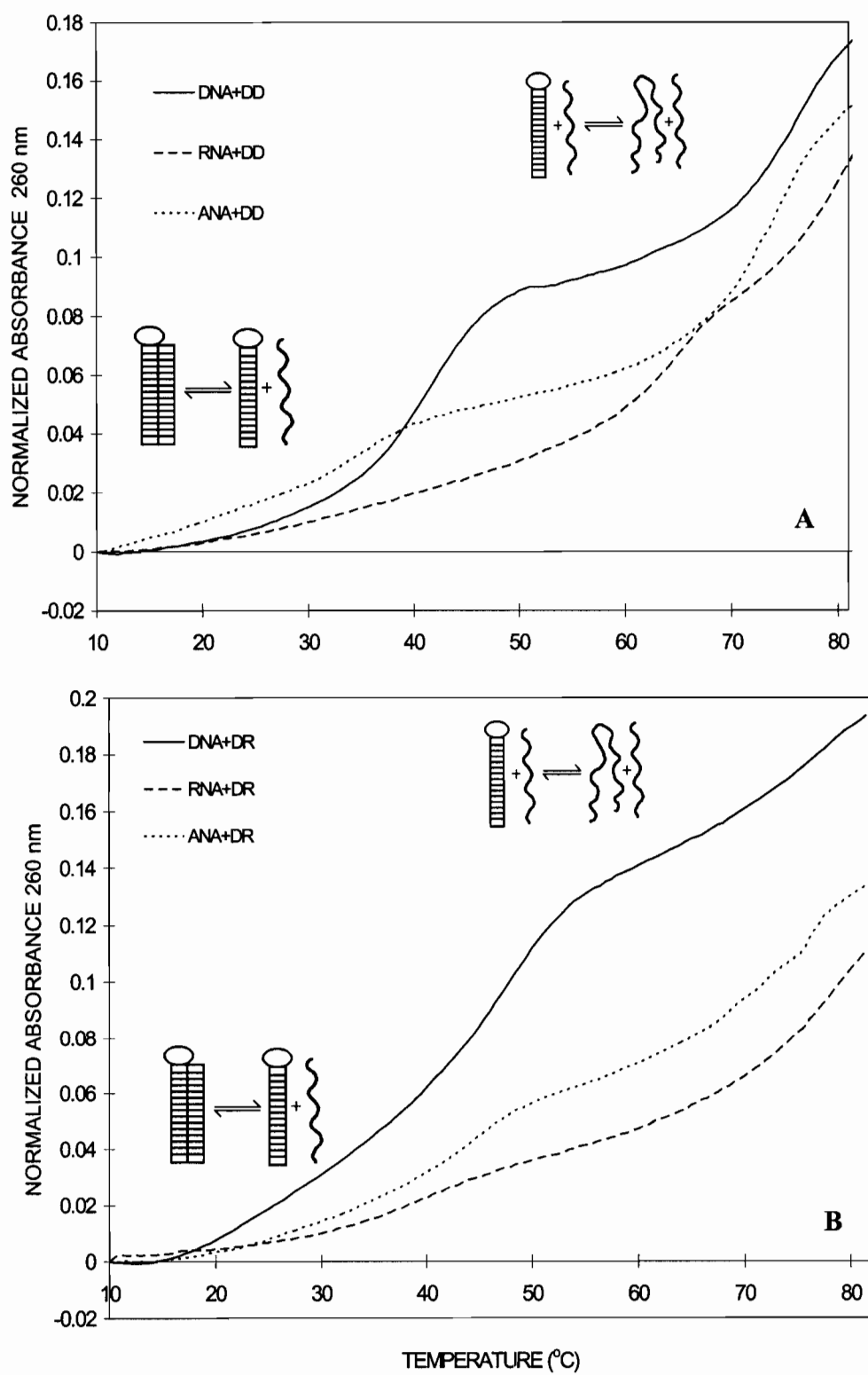


Figure 3.3.6: UV melting curves of complexes (2 μ M) in 100 mM sodium acetate, 1 mM EDTA, pH=5.5. (A) DD target and (B) DR targets.

Biphasic melting behavior was also observed for the control triplexes D•DD and D•DR, in agreement with the results of Roberts and Crothers.⁸⁵ For mixtures A+RD and A+RR only the transition corresponding to duplex melting was observed, which exactly parallels what was observed for D+RD and D+RR mixtures (Table 3.3.2).^{85,86,87,282} This is in clear contrast to the oligoribonucleotide (R) which formed stable triplexes with all DD, DR, RD and RR duplexes. In summary, these results show that *A and D associate only with DD and DR, whereas R associate with all four duplexes.*

Table 3.3.2: T_m Hybridization Data of Single Strands With Hairpins

Duplex	Third Strand	T_m (°C)	
		1 st	2 nd
DD	D	40	75
	R	62	77
	A	34	75
DR	D	45	75
	R		69 ^a
	A	43	73
RD	D	-	84
	R	43	86
	A	-	84
RR	D	-	84
	R	45	85
	A	-	84

All mixtures contained 2 μ M in each strand in 100 mM NaOAc, 1mM EDTA pH 5.5 and adjusted with acetic acid. ^aSingle transition for triplex \rightarrow R + DR (hairpin) \rightarrow DR (coil) processes

Gel Mobility Shift Detection of Triplex Formation

Triplex formation in the Py•Pu:Py motif is more difficult to analyze by spectroscopic experiments because C⁺-containing third strands often self-associate and the hyperchromic effects associated with the triplex-to-duplex transition are smaller. To confirm the above observations, triple helix formation was monitored by non-denaturing gel electrophoresis (Figure 3.3.7 A & B).

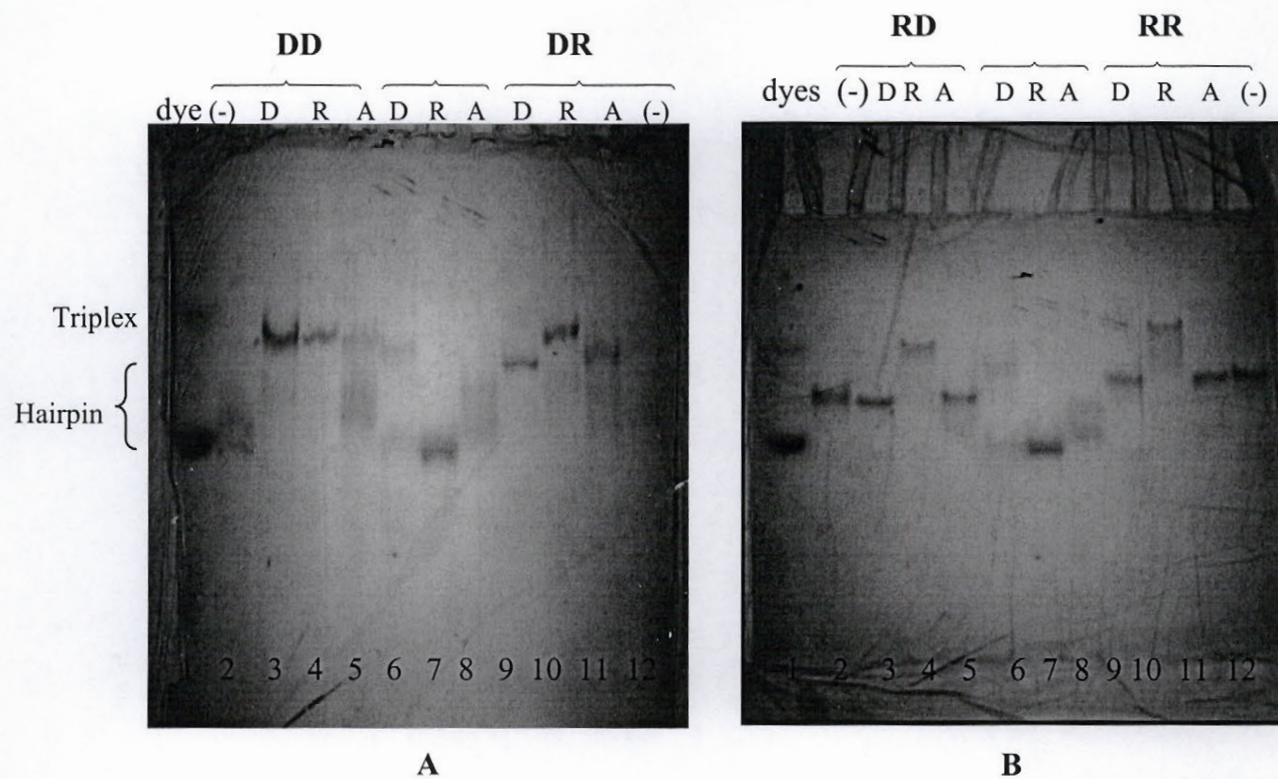


Figure 3.3.7 A & B: Gel shift mobility assay under non-denaturing conditions, 20% polyacrylamide, at 4°C and pH 5.0. Lane (1) marker dyes xylene cyanol (XC) and bromophenol blue (BPB). Hairpins either alone (-) or in the presence of (+) DNA, RNA and ANA as indicated. The DR hairpin is not clearly seen on lane 12 (A), however, it moves faster than the complexes it forms, very close the BPB dye. Lanes 6-8 represent the single stranded third strands.

Strand A interacted with DD and DR, as evidenced by the appearance of a new band of reduced mobility (**Figure 3.3.7 A**). In polyacrylamide gels, the mobility of the triple helix was retarded relative to the corresponding duplex, as shown in previous studies.^{76,283} Not all of A was shifted to triplexes under these conditions (A:duplex, 1:1 stoichiometry) as seen in **Figure 3.3.7 B**. This is in contrast to the incubation of D with DD and DR which, under the same conditions, produced D•DD and D•DR in quantitative yields. The mobility of RD and RR was unaffected by incubation with either A or D, confirming the UV thermal melting results that triplexes A (or D) •RD and A (or D)•RR do not form under these conditions. Finally, incubation of R with any duplex gave rise to the expected triplexes with decreased mobilities (**Figures 3.3.7 A & B**).

The triplexes constitute a more homogenous electrophoretic mobility group than the duplexes. Data from a 20% native gel indicated three likely groups: (I) A•DD (34°C); (II) D•DD, D•DR, A•DR, R•RR and R•RD (40-45°C); (3) R•DD, and R•DR (62°C and 69°C). These results suggest that the global conformations of hybrid triplexes formed by ANA and DNA third strands are similar, but different from those formed by RNA. The mobilities follow a pattern similar to that observed in the T_m and CD studies. Though the T_m s are close, the targets for group II include all four types of hairpins, *i.e.*, DD, DR RD and RR.

Circular Dichroism (CD)

As noted by Roberts and Crothers,⁸⁵ the hairpin duplexes exhibited considerable differences in CD spectra (**Figure 3.3.8A**). The CD spectra of DR hybrid is closer to that of the pure DD duplex, while the CD spectrum of RD resembles that of the pure RR duplex. The situation is different in the case of the triplexes, which exhibited appreciable spectral similarities (**Figure 3.3.8B-E**). For example, the spectra of A•DR is strikingly similar to that of R•DR, being only slightly different to the D•DR spectra (**Figure 3.3.8C**). These similarities are most likely the result of the conformation of the underlying duplex, *e.g.*, DR, which dominates the CD-spectra, rather than the three-dimensional arrangement (*e.g.*, sugar puckering) of the constituent strands (see Discussion Section). The differences in CD spectra are mainly located in the region around 280 nm, where the Y•DR (and Y•DD) spectra show 'red-shifted' Cotton effects,

Y being the pyrimidylate third strand. The CD spectra of a 1:1 mixture of A and RD, or A and RR, show reduced band intensities and did not differ significantly from the calculated average of the spectra of A + RD, or A + RR, consistent with the lack of association of these strands.

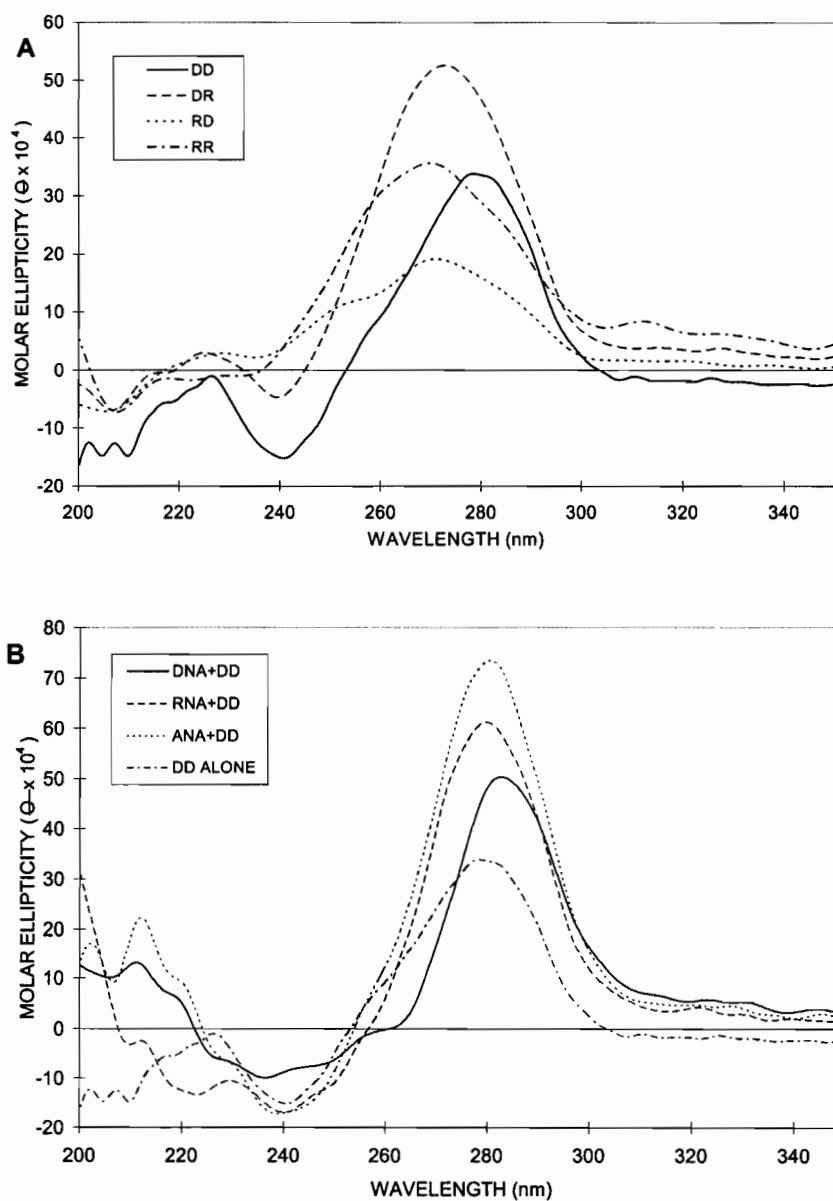


Figure 3.3.8 (A): CD spectra of hairpin duplexes. **(B)** CD spectra of mixtures of hairpins and single strands. Target is the DD hairpin. Concentration is 2 μM in each strand and the buffer is 100 mM sodium acetate, 1mM EDTA pH 5.5 for both **A** and **B**.

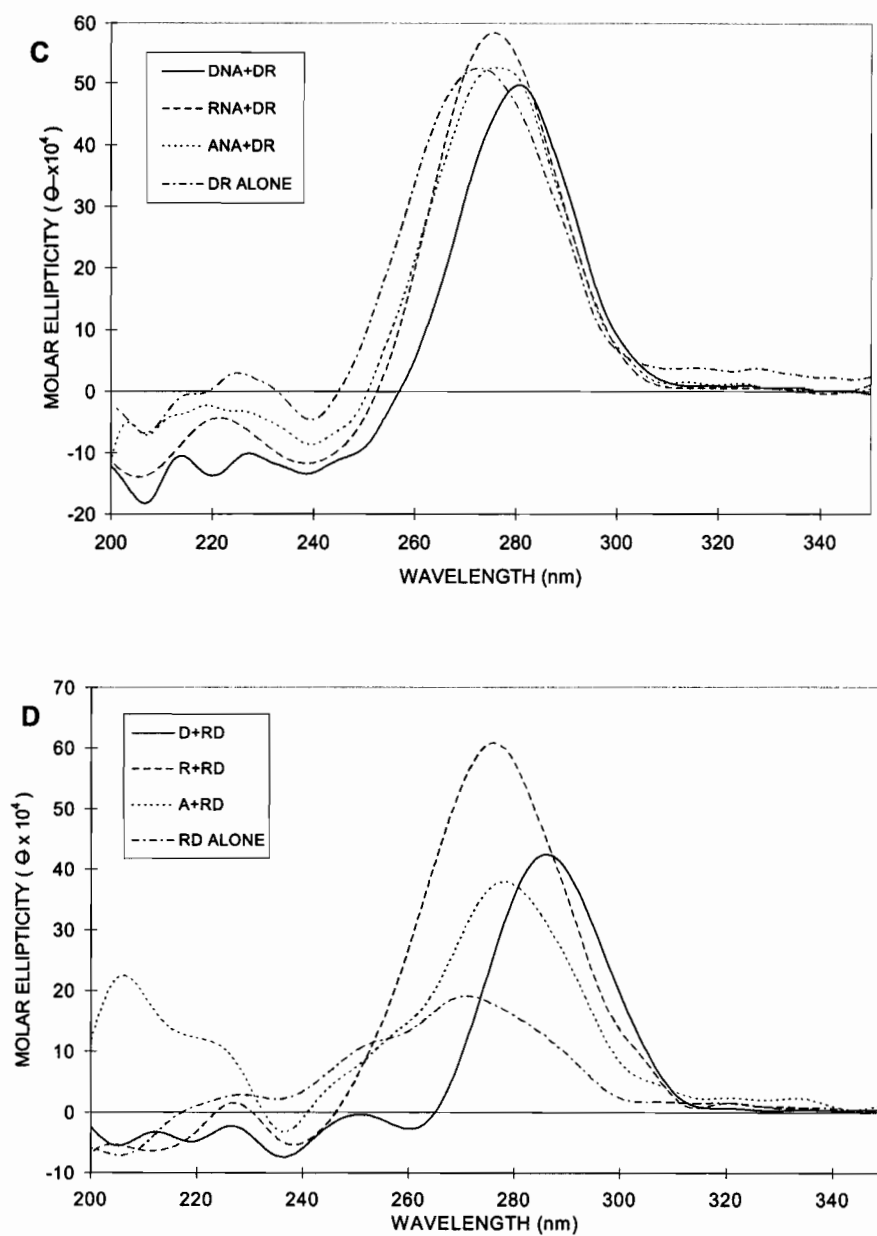


Figure 3.3.8: CD spectra of mixtures of hairpins and single strands. Target is the DR hairpin, (Panel C) and RD target hairpin (Panel D). Conditions as reported above in Panel A.

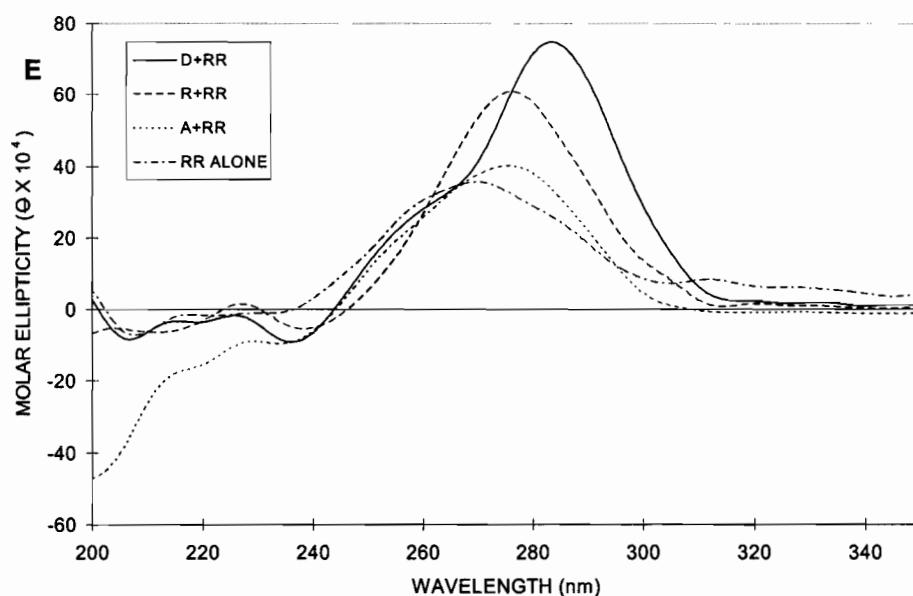


Figure 3.3.8 E: CD spectra of mixtures of hairpins and single strands. Target is RR target hairpin. Conditions as reported above in Panel A.

3.3.4 Discussion

A goal of this study was to investigate the associative properties of an oligoarabinopyrimidylate, and its ability to form a complex with complementary ssDNA, ssRNA and duplexes. The relative stabilities were determined by T_m experiments, while CD and gel electrophoresis were used to investigate conformation and structure.

A pure RNA duplex (R:R') had the highest stability. The hybrid duplex (A:R') obtained by replacing the R (py) strand with an A (py) strand is of lower stability than the RNA duplex, but comparable in stability to the pure DNA duplex (D:D'). The A+D' complex was not detected under the conditions of this study, and suggests that ANA binds selectively to ssRNA over ssDNA complements. In the case of the hairpin duplexes there seems to be two groups whose classifications are based on the T_m , CD and gels profiles. *i.e.* DD and DR (T_m 77°C and 73°C), and RR and RD (T_m 82°C and 85°C). RNA substitutions in the pyrimidine strand of D(py):D'(pu) duplexes do not appear to have dramatic effects on hybridization patterns and CD signatures. These results are in agreement with previous studies that suggest that oligomers with a higher RNA purine content ultimately dictate the conformation and thermal stability of the duplexes.

The influence of the sugars in the duplex strands on triplex stability is more difficult to interpret. The results presented above demonstrate that pyrimidine oligoarabinonucleotides do not form triplexes with duplex RNA (RR) or hybrid RNA (purine): DNA (pyrimidine)(RD). However, arabinopyrimidines can act as the third strand in recognition of duplex DNA (DD) and hybrid DNA (purine): RNA (pyrimidine) (DR). This selectivity parallels exactly what has been previously observed for DNA third strands.^{85,86,87,282} In contrast, RNA (the 2'-epimer of ANA) shows a different behavior, forming stable triplexes with all four DD, DR, RD and RR duplexes. Comparison of the T_m data of the various triplexes revealed that those with ANA third strand were thermally less stable than those with RNA strands, but similar to those with DNA third strand. Interestingly, the ANA had a slightly higher affinity for a DR than DD. Indeed the T_m s of the ANA and DNA binding to double helical DNA•RNA are comparable (43-45°C). Hairpin DD, being composed only of DNA should have a pure B type conformation, with a minor and major groove that is typical of B helices, while DR is expected to have a hybrid form intermediate between pure B and A form duplexes. The favorable binding of A to D:R could be useful for targeting such duplexes *in vivo*.

Various reasons may be invoked to explain (1) the contrasting hybridization behavior of ANA and RNA third strands, and (2) the similar binding characteristics of ANA and DNA strands. A possible interpretation for (1) is that in the case of RNA, the stereochemistry of the sugar favors formation of short contacts between the 2'-OH groups of the third strand with the purine strand phosphates,²⁷² as well as with the 5' oxygen as predicted by computational models,^{87,284} and x-ray crystallography²⁸⁵ respectively. The ribose (α) 2'OH not only induces a preferred N-type pucker but may also be involved in direct or water-mediated intrastrand and interstrand interactions.¹ In the case of the 2'-epimeric ANA strands such a mechanism may not be possible probably because its β C2'-OH points in a different direction. This could then be involved in interaction with the neighboring sugar on the 5'-side of the purine which is now sterically unfavorable in such triplex structures.

Another possible explanation, which reconciles both (1) and (2) above, is that *arabinonucleotides mimic deoxyribonucleotide rather than ribonucleotide conformations*. The sugars in RNA adopt primarily the C3'-endo pucker, regardless of

whether the RNA is found in single-stranded, double or triple-helical forms. Such conformation is governed by anomeric effects of the C1' and C4' substituents with the ring oxygen, and/or stereochemical requirements of complex formation.^{1,182} This situation differs for DNA, where an O3' gauche effect to O4' favors the C2'-endo pucker, particularly in aqueous solution (**Figure 3.1.2**, pp. 53). For example, Dagneaux *et al.* showed that the sugars of DNA triplexes (T•A:T) assume the C2'-endo conformation, while triplexes containing a ribo third strand have mainly C3'-endo-type, e.g., R (C3'-endo)•D(C2'-endo):D(C2'-endo), and R(C3'-endo)•R(C3'-endo):R(C3'-endo).²⁷³ Arabinonucleotides are expected to mimic deoxyribonucleotides since a combination of O2'→O4', and O3'→O4' *gauche effects* would stabilize the C2'-endo geometry (**Figure 3.1.2**).

On the basis of the available data, it cannot be decided which of the above two effects is more important, but it is suspected that more than one of these is operating in an important way. It is tempting to speculate that in order to recognize all four possible dispositions of DNA and/or RNA strands in Watson-Crick duplex, oligonucleotide analogues must be RNA-like, *i.e.*, have both the C3' endo-like sugars and a 2' α-OH group, whereas those adopting the C2'-endo pucker without the appropriately oriented 2'-OH group will recognize only DD and DR duplexes. Two recent findings support this view wherein 2'-5'-linked RNA (C2'-endo, and with a 3' α-OH)²⁷¹ and a 2' fluoro-RNA²⁸⁶ (C3'-endo, but lacking a 2'-OH) binds only to DD and DR, but not RD and RR. The common sugar pucker in 2'-5' RNA, DNA and ANA may explain, at least in part, the similar thermal stabilities of (R*,D,A)•DD and (R*,D,A)•DR triplexes.

3.3.5 Conclusion

Arabinonucleic acids are able to recognize double helical complexes, demonstrating that the stereochemistry of the 2'-OH groups of a triplex forming RNA strand can be inverted, but not without affecting the hybridization properties of such strands, and the stability of the complexes formed. Such an understanding can be applied to the design of sequence selective oligonucleotides which interact with double-stranded nucleic acids.

With respect to the hybridization behavior of ANA, *versus* the regioisomeric RNA, the following principles apply: If the target nucleic acid is single-stranded DNA, or a double helix with a RNA in the purine strand, only RNA will bind. If the target is single-stranded RNA, or a double helix containing a DNA in the purine strand, RNA, ANA, DNA or 2'-5' RNA will bind. In all cases, the duplexes or triplexes formed by ANA are thermally less stable than those formed by RNA. The results presented here show that, similar to effects encountered with the double helix, triplex stability is governed not only by base sequence and the chemical nature of their strands (*e.g.*, third strand ribose *versus* arabinose) and sequence but more precisely by their backbone conformation. In analogy to the 2'-5' RNA and 2'-deoxyribose third strands, the possible C2'-endo pucker of arabinoses together with the lack of an α -2'-OH group are believed to be responsible for the selective binding of ANA to DD and DR duplexes, over RR and RD duplexes. Recent work in Damha's group has attempted to uncover the relative importance of the proposed 2'-OH/phosphate contact mechanism²⁸⁷ and the sugar conformation of the third strands, utilizing 2'-deoxy third strands with 'locked' C2'-endo and C3'-endo puckers associating to DD and RR duplexes. This report supports the notion that the major force governing duplex selectivity is the contact formed by the 2'-OH of the third strand the phosphate groups of the duplex purine strand. This report along with the results presented here emphasize the role of the ribose 2'-OH group as a general recognition and binding determinant of RNA.

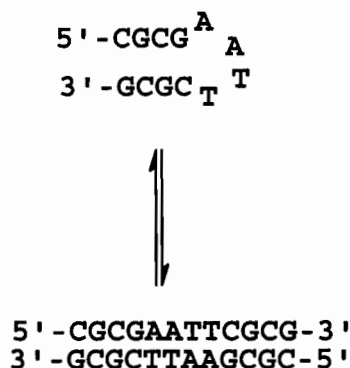
3.4 POSSIBLE FORMATION OF AN ARABINONUCLEIC ACID DOUBLE HELIX

3.4.1 Introduction

Our knowledge of DNA structure and function has undergone rapid enrichment in recent years with discoveries of new DNA structures. In addition to the most common form of DNA, namely the duplex, several other three-dimensional forms such as hairpins, cruciforms, triple-stranded structures and four stranded structures (Holliday junctions, G-quadruplex, i-motifs)^{277,288} have been discovered and investigated in detail. Hairpin, dumbbell and cruciform DNAs occur as a consequence of inverted repeats in a single strand of the DNA and are known to play crucial roles in genetic recombination and regulation. Sequences capable of hairpin formation are often seen near regulatory and promoter sites in DNA, and are stabilized by interaction of the loops with specific enzymes.²⁸⁹ The process of genetic expression requires unwinding and opening of the duplex to single stranded DNA, and then depending on the base sequence and environmental conditions, these ssDNA could form secondary structures that are specifically recognized by proteins during the course of biological functions. A molecular level understanding of all these phenomena requires a detailed knowledge of the three-dimensional structure of the individual DNA forms. The structural transitions between the various forms of DNA could have important consequences *in vivo*.

The hybridization properties of ANA strands were examined in Chapter 3. These dealt with duplex and triplex formation, and it was concluded that the incorporation of arabinonucleotides within a DNA strand has small structural perturbations but the biological outcomes are manifold (Chapter 4, Sections 1-2). ANA was selective to RNA over ssDNA in two different studies and conditions (Chapter 3, Section 1 and 3). When the branch-point of a V-shaped oligomer is changed from ribose to arabinose structural effects are rather minimal (Chapter 3, Section 2). When ANA is combined with a duplex it forms triplexes, but only if the purine strand of the Watson-Crick duplex is DNA (Chapter 3, Section 3). With all of these findings a natural follow-up question is "Is it possible for two ANA strands to associate and form a duplex?"

To answer this question the ANA analogue of a well known DNA sequence (CGCGAATTCGCG) was investigated. This oligonucleotide is popularly known as the “Dickerson-Drew Dodecamer” and self-associates to form slightly more than one complete turn of B form duplex DNA. Under normal conditions of neutral pH and ionic strength, the molecule can also form a hairpin as shown below.



Low temperatures and high DNA concentrations

Figure 3.4.1: Possible association forms of the Dickerson-Drew dodecamer

The hairpin form has a 4 nt loop in the central portion, and a 4 base pair GC rich stem. The shift from hairpin monomers to helical duplexes is favored by lower temperatures and by higher DNA and salt concentrations, the same conditions which also favor crystal growth. An additional significance of the molecule is the *EcoRI* restriction site, G-A-A-T-T-C, in the sequence. Moreover, because it contains C-G-C-G chains at the ends, it also offers a test for the tendency of mixed-sequence DNA to adopt another high-concentration, high-salt alternative to the B-duplex: the so called ‘left-handed or Z DNA double helix’. At lower pH the molecule may be anticipated to further associate and form triplex or higher molecular structures. Whether a hairpin or a duplex is achieved, the resulting complex that is formed should still homogeneously be composed of ANA strands. The weak tetrameric forms of ANA observed earlier (Chapter 3, Section 2) involved $\text{C}^+ \bullet \text{C}$ ‘base pairing’ and not Watson-Crick base pairing which is the objective of the present study.

Another fundamental question which is still uncertain is the conformation of ssANA. Studies in earlier sections of this chapter point to greater similarities of ANA to

DNA, rather than to RNA. Thus ANA, like the DNA strands may have some flexibility as a single strand and its observable conformation in a duplex may be dictated by the identity of the target sequence. Therefore, an ANA strand with the same sequence as the DNA Dickerson Dodecamer was synthesized and subjected to hybridization and conformational analysis by CD and UV. The DNA and RNA analogues were also prepared and likewise studied for comparative purposes.

Table 3.4.1: Dickerson-Drew Oligomers Selected For Study

OLIGOMER	DESIGNATION	SEQUENCE #
d(CGCGAATTCGCG)	DNA	<u>3.32</u>
r(CGCGAAUUCGCG)	RNA	<u>3.33</u>
a(CGCGAAUUCGCG)	ANA	<u>3.34</u>

3.4.2 Results

Circular Dichroic and UV Spectroscopy

CD and UV spectra of the single stranded Dickerson-Drew oligomers are shown in **Figures 3.4.2** and **3.4.3** respectively. In contrast to the UV absorbance profiles, the measured CD spectra displayed significant differences between the three strands. CD spectra for DNA exhibited characteristic positive (282 nm) and negative peaks (250 nm) of equal intensities, while the RNA spectra had a positive band at *ca.* 263 nm and a very significant negative peak at 208 nm characteristic of A-form RNA duplexes. The ssANA had a very unusual CD signature with a positive band of intermediate amplitude centred around 268 nm. This spectrum can be seen to have features of both DNA and RNA.

The other interesting observation was the markedly different crossovers for the D, R and A, (Dickerson-Drew strands). DNA had two crossovers at 270nm and 239 nm, the RNA had several such cross overs while the ANA had only one at 250 nm coinciding with one of the RNA.

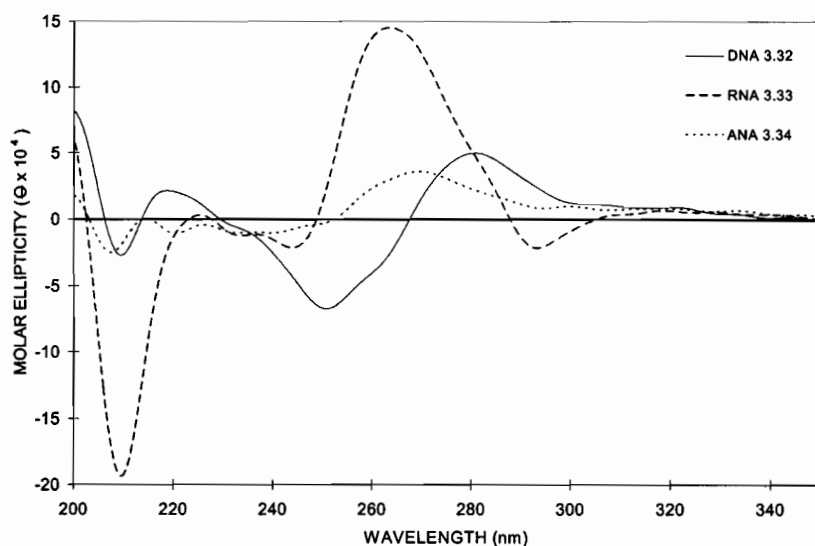


Figure 3.4.2: CD spectra of single-stranded oligomers **d(CGCGAATTCGCG)** (—) ; **r(CGCGAAUUCGCG)** (- -); **a(CGCGAAUUCGCG)** (.....) in water, 20°, pH 6.9.

Above 230 nm the UV absorption spectrum of the ANA strand was very similar to the spectra of the DNA and RNA sequences. However the maximum absorption of ANA (268 nm) is red shifted relative to both DNA (258 nm) and RNA (260 nm).

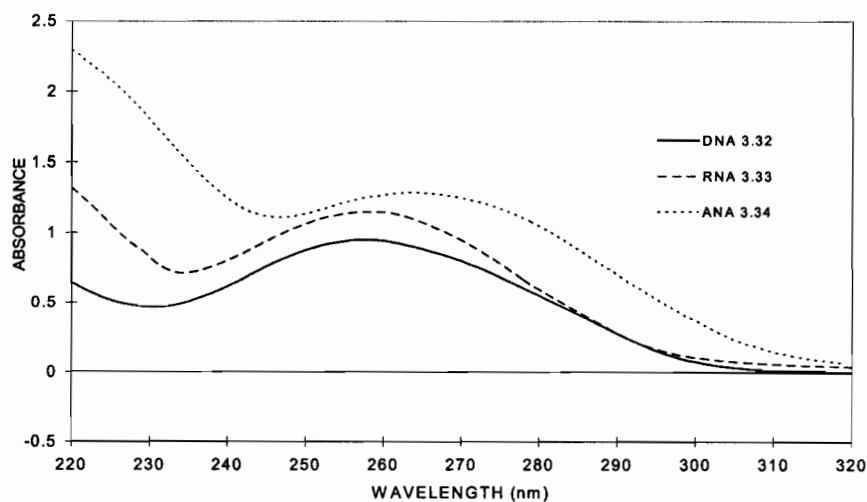


Figure 3.4.3: Absorbance spectra of single-stranded oligomers **d(CGCGAATTCGCG)** (—) ; **r(CGCGAAUUCGCG)** (- -); **a(CGCGAAUUCGCG)** (.....) in water, 20°, pH 6.9.

Melting Studies

The strands were then subjected to hybridization studies in different buffer systems. The oligonucleotides, each at a strand concentration of 4.4 μM , were incubated on their own at 90°C for 20 min., allowed to cool to room temperature and then stored at 4°C (overnight). The melting curves (A_{260} nm) were then determined with a heating rate of 0.5°C /min over 5-90°C. From the results obtained it is evident that a solution of ANA does not give a detectable T_m hyperchromic transition in any of the buffers used (**Table 3.4.2**). This is in contrast to R and D which showed sharp transitions in all buffer systems. Thus for this base sequence the stability order is $A < D < R$, and one is then tempted to conclude that ANA is unable to form double helical structures.

Table 3.4.2: Thermal UV Studies Of The Dodecamers In Various Buffers

Oligomeric Strands	1M NaCl ^a		140 mM KCl ^b		pH 5.5 ^c	
(4.4 μM)	T_m (°C)	%H	T_m (°C)	%H	T_m (°C)	%H
D	60.9	9.0	54.4	9	54.1	12
R	64.0	8	62.6	6	54.1	14
A	-		-		-	

^a 1M NaCl, 100 mM Na₂HPO₄, pH 7.2; ^b 140 mM KCl, 5 mM Na₂HPO₄, 1 mM MgCl₂, pH 7.2; ^c 1 mM EDTA, 100 mM. NaOAc pH 5.50.

3.4.3 Discussion

The Dickerson Drew dodecamer was one of the first sequences to have been solved at high resolution and as such has been a choice candidate for several studies involving modified oligonucleotide analogues. Previous studies that have dealt with the same DNA sequence usually containing a single arabino insert.^{147,155,188} In general the dodecamers containing arabino inserts show a decrease in T_m by *ca.* 2°C/insert relative to the DNA control. Other studies also containing 1-2 ara inserts in the dodecamer but fluorinated at the C2' carbon gave mixed results. The T_m increased for araFT inserts

while it decreased for ara-FC inserts. Moreover Kois and coworkers¹⁸⁸ have shown that modifications of the dodecamer with ara-FT or ara-FC dramatically increased the catalytic efficiency of *EcoRI* endonuclease relative to the unmodified sequence by several orders of magnitude. It has been suggested that the 2'- β -fluoro modification influences the local structure of the molecule and/or the electrostatic environment including water concentration in the activated complexes.

The spectral studies reported here were conducted in acidic and neutral pH. No apparent melting was observed for sequences involving the all ANA Dickerson-Drew dodecamer. The CD spectra of ANA (Dickerson) differed considerably from CD spectra of DNA and RNA analogues. The spectrum appeared to be intermediate between that exhibited by the corresponding RNA or DNA strands. The reasons for the absence of a detectable ANA:ANA complex is unknown at the present time.

However it is a well established fact that the conformational and solution properties of nucleic acids are strongly dependent on base-composition, sequence and chemical structure.²⁹⁰ For example, duplexes containing rA.rU base pairs in the middle of the helix CGCGAAAUUCGCG are of lower stability relative to those containing dA:dT base pairs at the same positions. Similarly, while a short d(A_nT_n) tract stabilizes the DNA helix, an analogous r(A_nU_n) tract substantially destabilizes the RNA helix.¹⁹⁵ Clearly, the influence of thymine methyl groups is substantial. The same effect may operate for ara(A_nT_n) and in retrospect, this study should have also included ara(CGCGAAATTTCGCG).

As this study was being completed, we became aware of the Ph.D., thesis work of Resmini¹⁹¹ which reports that aU₁₀ associates with the aA₁₀ to form a complex having a T_m of 36 °C (1M Na⁺, pH 7.0). This has also been confirmed by C. Wilds of our laboratory for the complex aU₁₈:aA₁₈ (T_m, 50°C). However, the ability of two ANA strands to associate appears to be limited to homopolymeric sequences since the mixed base complementary strands ara(AICUCCCAIICUCAIAUC) and ara(LAUCUIAICCUIIIAICU) do not associate.²⁹¹

The fact that ANA (mixed base sequence) is unable to form ANA:ANA duplexes may be of interest to researchers trying to uncover the criteria that Mother Nature followed in selecting the RNA structure over the course of evolution. Since ANA (mixed

base) does not have the base-pairing capability exhibited by DNA and RNA, it can be suggested that ANA would be incapable of self-replication under potentially natural conditions. In fact Eschenmoser has argued that "alternative" nucleic acid analogues can be dropped from the list of potentially relevant evolutionary competitors of RNA if, in fact, that analogue is not a base-pairing system.²⁹² The pioneering study by Eschenmoser has also dealt experimentally with the question of "why pentose and not hexose nucleic acids?",²⁹² and more recently in a report that appeared as this thesis was being submitted - "why ribose and not another pentose?"²⁹³ The outcome of these studies led to the conclusion that six-membered pyranose-nucleic acids not only exhibited much stronger Watson-Crick base pairing than DNA, but that it could not have acted as viable nucleic acids in early evolution because of steric bulk of the hexopyranoses ("too many atoms"). They also find that while pentopyranose systems (*i.e.* six-membered pentose rings) display stronger base pairing properties than RNA, pentofuranoses (specifically the araU:araA system) have lower base-pairing potential than RNA. Based on the above, Eschenmoser concluded "whatever the chemical determinants by which nature selected RNA as a genetic system, maximization of base-pairing strengths within the domain of pentose-derived (furanose, but particularly pyranose) was not the critical selection criterion".

CHAPTER IV BIOLOGICAL APPLICATIONS OF ARABINONUCLEIC ACIDS AND THEIR ANALOGUES²⁹⁴

4.1 INHIBITION OF HIV-1 RT SYNTHESIS

4.1.1 Introduction²⁹⁵

The notion that oligonucleotide (ONT) analogues might be designed to bind to m-RNA and disrupt the production of various gene products²⁹⁶ revealed a great paradox in its apparent theoretical simplicity, but extraordinary technical and practical complexity. In fact, the concept raised the possibility of creating an entirely new area of pharmacology. A potential new drug: oligonucleotides; a new target: m-RNA; a new target binding motif: Watson-Crick hybridization and a new class of post-target binding events, *e.g.* RNase H-mediated degradation of the target RNA. Considered in this context, it was obvious that an enormous investment would be required and that only as examples of drugs based on this technology were tested, would one have the opportunity to understand the technology, its potential, its problems and its limitations.

The success of the surviving companies using this approach is a sign that the biotechnology industry is maturing. It seems to hold special promise, as in a period of only ten years or so, about a dozen antisense-based drug candidates have reached phase I clinical trials. ISIS's VitraveneTM (fomivirsen) represents the first compound based on antisense technology to be approved by the US Food and Drug Administration (August 26, 1998). It is used in the treatment of cytomegalovirus (CMV)-induced retinitis that occurs in AIDS patients and is uniformly progressive to blindness. In 1993 Hybridon's anti-HIV compound GEM 91 became the first antiviral antisense drug to be given directly to patients. ISIS Pharmaceuticals based in Carlsbad, California has several compounds in clinical trials. To name a few: ISIS 2302 (inhibitor of CAM-1) is used for treatment of Crohn's disease, ISIS 351/CGP 6412A, ISIS 5132/CGP 69846 and ISIS 2503 are being tested against cancer.

Antisense technology builds on some of the fundamental tenets of modern molecular biology. It is both simple and wide-reaching: it uses synthetic copies of the basic components of life - the nucleotides that make up DNA - to block disease processes at their origin. This is achieved by synthesizing a compound that would bind to the target RNA

strand by the Watson-Crick base pairing as detailed earlier in Chapter 1, Section 2. Some of the therapeutic challenges in greatest need of breakthroughs are viral infections, one of which is the Human Immunodeficiency Virus (HIV). With the development of synthetic oligonucleotides (ONTs), the basic strategy of antisense therapy has finally become feasible, and Zamecnik, in collaboration with Robert C. Gallo, reported for the first time that antisense DNA directed against HIV-1 in cell culture significantly reduced viral replication.²⁹⁷ Although many therapies exist for HIV-1, there is still a tremendous need for new drugs, particularly because the virus develops resistance to current drugs. For example, the 3'-azido thymidine (AZT) and other 2',3'-dideoxynucleosides (3TC, ddI, and d4T), inhibit reverse transcription of viral RNA but, have no effect on the chromosome integrated proviruses. Because studies show that a high viral burden exists in clinically late AIDS patients, this chemical approach may have little benefit beyond transient prevention of virus spread. In addition, the clinical use of nucleoside analogues is associated with severe toxic side effects. As a result, the cumulative benefits of AZT therapy may be questionable. Other approaches such as recombinant vaccines and gene-therapy remain uncharacterized for their efficacies in patients. Thus, faced with the emergence of the above mentioned hurdles, novel strategies to treat AIDS are needed. Not surprisingly and entirely appropriate, many programs have focused on discovering and developing antisense-based antiviral drugs. Although a long way from declaring victory, progress has been gratifying.

Clearly, these are exciting times in antisense, and one is poised to ask crucial questions both prior to and in the clinic. It would be unreasonable to expect every drug based on antisense to work. However it is hopeful and one is buoyed by the activities that have been observed for antisense ONTs developed in the Damha research group particularly those containing arabinose sugars.²⁹⁸ As shown below these molecules deserve thorough evaluation and the technology merits continued investigation.

In this chapter antisense drugs that interfere with the synthesis of several gene products, thought to play a role in the life cycle of the HIV-1 RT are considered. **Section 4.1** deals with the blockage of polymerization and strand switching reactions all catalyzed by HIV-1 reverse transcriptase (RT). **Section 4.2** illustrates the pausing of DNA elongation, which in turn produces defective proviral products, and **Section 4.3** looks at digestion and destruction of the RNA template by inducing RNase H cleavage activity.

Background of HIV-1 Reverse Transcriptase

The discovery that the HIV is the cause of the acquired immunodeficiency syndrome (AIDS) has fostered a plethora of research into the underlying mechanisms of the viral infectious cycle and pathogenesis. Between the point of infection and the onset of AIDS lies an unexplained disease process. Does HIV directly destroy the immune system, or does it provoke the immune system into self-destruction? Or do both processes occur? Can HIV alone cause AIDS, or does it need help from other factors?²⁹⁹

Studies on these mechanisms have provided researchers with an ever-increasing number of molecular targets for the development of antivirals. An ideal antiviral drug should show strong potency in inhibition of viral replication reduced host toxicity. Nucleoside derivatives have been, and continue to be, the predominant molecular prodrugs candidates for the clinical therapy of AIDS. Antivirals against HIV-1 can be categorized by their mode of action. They include inhibitors of reverse transcriptase (RT), competitors for viral entry into cells, vaccines, protease inhibitors, and an emerging group referred to as the “genetic antivirals.”³⁰⁰ Antisense ONTs belong to this latter class and they differ from other gene therapy vaccines which attempt to stimulate cytotoxic T-lymphocytic response against the HIV-1 envelope protein. Generally anti-HIV ONTs have targeted a single stage of viral replication. But the popularity of genetic antivirals lies in their ability to attack HIV simultaneously at multiple loci/sites in the HIV-1 genome, thereby minimizing the emergence of resistant viruses. *i.e.* antisense ONTs may inhibit multiple stages of viral replication when targeted to an appropriate sequence on HIV-1 genomic RNA.³⁰¹

RT is a multifunctional enzyme that possesses RNA-dependent DNA polymerase (RDDP), DNA-dependent DNA polymerase (DDDP), strand transfer, strand displacement and RNase H activities (see **Figure 4.1**).

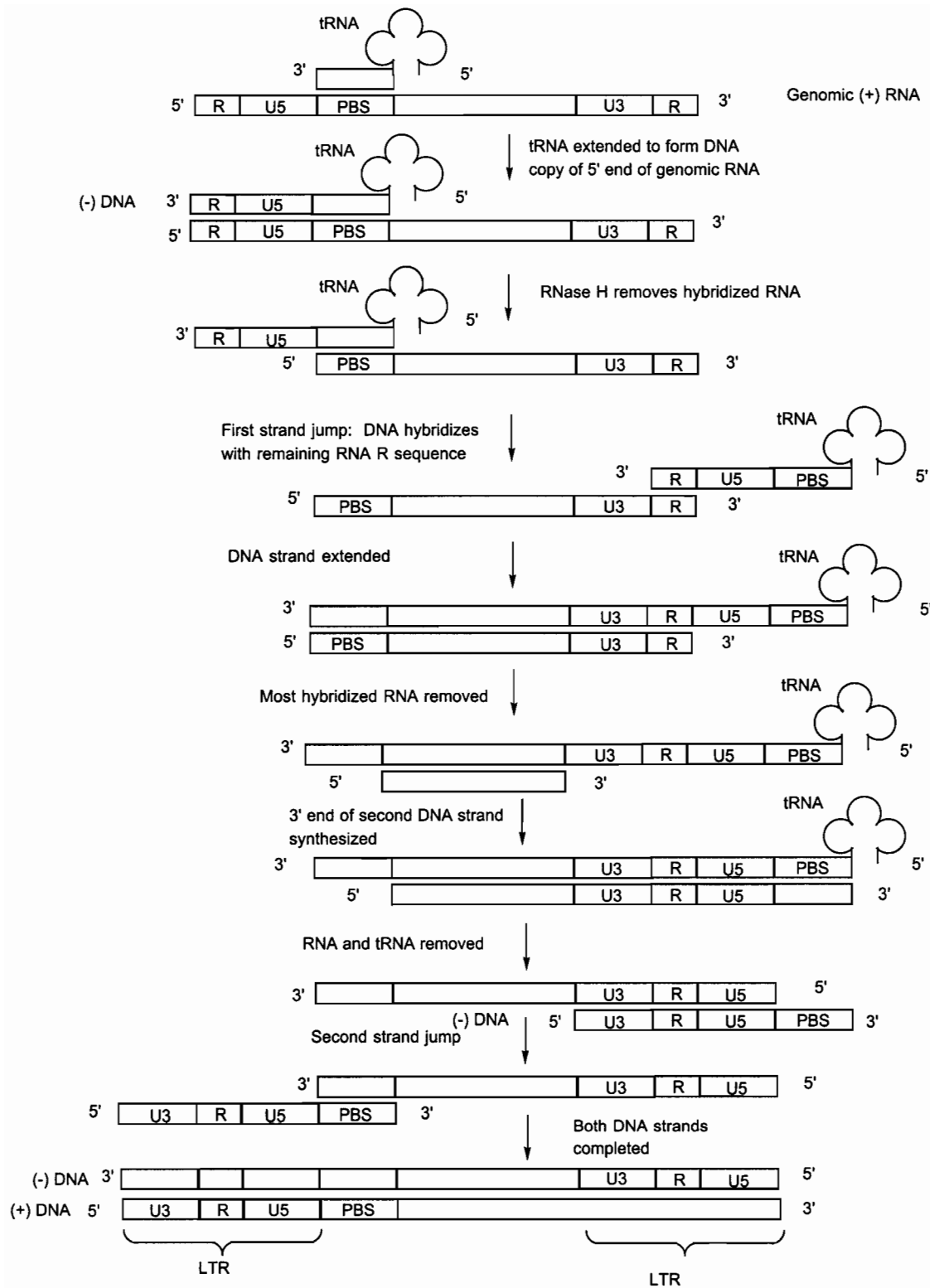


Figure 4.1: The mechanism of reverse transcription in viral gene expression. Reverse transcriptase (RT) performs all of the above functions. When it acts as a RNA dependent-DNA polymerase RT catalyzes formation of a complementary DNA to the viral genomic RNA. Its RNase H function results in degradation of RNA of the RNA:DNA hybrid. RT also functions as a DNA-dependent DNA polymerase, where it forms double stranded viral DNA ready for integrating into a host genome. Once viral DNA duplex has been incorporated into a host eukaryote DNA duplex the replication follows the same procedure outlined for eukaryotes. Adapted from L. Stryer.⁵

It is known to act within a tight cytoplasmic complex that includes viral genomic RNA, tRNA, viral nucleocapsid proteins and deoxynucleoside triphosphates. The RDDP activity is responsible for the synthesis of the minus strand (-) of DNA and also creates the primer for plus strand (+) synthesis, which is catalyzed by the DDDP activity of the RT. Additionally, RT catalyzes at least two strand transfer events and strand displacement synthesis to complete the process of reverse transcription.³⁰² Thus the central role of RT in the viral replication cycle needs no further emphasis and hence the wisdom of continuing to target the RT itself can be justified.

RNase H versus physical blockage by antisense oligomers

As mentioned in Chapter 1, to be successful, antisense therapeutics have to fulfill several criteria, including sufficient nuclease resistance, biodistribution and ease of synthesis.¹⁰ These issues have spurred a quest for antisense oligonucleotide analogues with favorable specificity, affinity and stability (Section 1.2).²³ Since many types of nucleotide modifications which provide nuclease resistance also reduce the stability of the ONT:RNA duplex, increasing affinity of the ONT towards the target RNA is of paramount importance.^{16,303,304,305} An open question is whether the induction of RNase H activity by an antisense oligonucleotide is necessary for efficient inhibition of gene expression or whether binding of the oligonucleotide to its target sequence with high affinity is sufficient to achieve this goal. Several recent studies support the first assumption that the more potent antisense effects are obtained when RNase H is activated.³⁰⁶

Wainberg and co-workers³⁰⁷ of the McGill AIDS center have developed an *in vitro* reverse transcription assay to study the properties of priming, RNA-dependent DNA polymerization and template switching by HIV-1 RT, *i.e.* the same reactions that occur in infected cells.

In principle, the antisense approach can be effected in this particular study *via* several modes of action: (I) antisense oligomer hybridization to viral RNA prevents transcription, (II) competition with tRNA priming for PBS region, (III) direct interaction with the RT, (IV) extension of the antisense oligonucleotides *via* priming (defective proviral DNA) and (V) termination of cDNA polymerization *via* steric blockage of DNA

synthesis. Thus ONTs (18 nt in length) and complementary to the R region near the 5' end of viral HIV-1 genomic RNA were synthesized in order to evaluate their antisense potential in the inhibition of HIV-1 replication. The R region sequences are conserved in the HIV-1 genome and are present at both ends of the genome. These characteristics make the R sequences very attractive targets in genetic antivirals therapeutics *via* mode I, III, IV and V.

4.1.2 Results

Inhibition of (-) strong stop DNA synthesis by antisense oligonucleotides

This section briefly describes experiments demonstrating the inhibitory effects of the synthetic ONTs bearing modifications at the 3' terminus on the continued synthesis of (-) strand proviral DNA.³⁰¹ These sequences are 5'AGC TCC CAG GCT CAG ATC 3' (Control DNA **3.1**) and 5' TAC GCA CGT CAC GTA CCG 3' (Random Control **3.18**). The modified ONT contained a single araC residue at the 3' terminus *i.e.* 5'AGC TCC CAG GCT CAG ATaC 3' (**3.6**). The hybridization properties of **3.1** and **3.6** are similar and were described earlier in Section 3.1.

The DNA polymerization products expected to arise from *in vitro* reverse transcription in reactions employing the pHIV-PBS donor RNA template, in the absence and in the presence of antisense ONTs, are illustrated schematically in **Figure 4.2**. In the absence of antisense ONT inhibitors, the expected full-length (-) strong stop DNA obtained from PBS oligonucleotide-primed reverse transcription on the pHIV-PBS RNA template is 192 nt. Inhibition by the specific antisense ONTs used here should result in a final DNA product 162 nt in length.

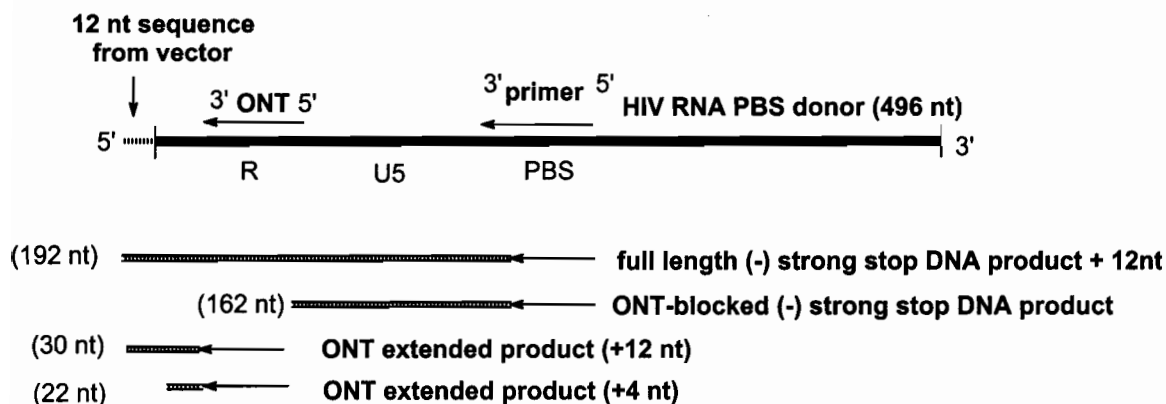


Figure 4.2: Schematic illustration of the anticipated DNA polymerization products in the *in vitro* DNA polymerization reactions testing the effect of the oligonucleotide on (-) strong-stop DNA synthesis, using pHIV-PBS RNA containing an additional 12 nt of vector-derived sequences.

Polymerization from the PBS ONT annealed to the pHIV-PBS RNA template in the absence of antisense ONTs, resulted in the expected (-) strong stop DNA (192 nt) product (95% of the total polymerization product). When the antisense ONTs were added RT-catalyzed synthesis of full-length (-) strong stop DNA was inhibited, as evidenced by the appearance of significant amounts of the predicted shorter 162 nt polymerization product (90 % of the total polymerization products) (**Figure 4.3** and **Figure 4.4**). This 162 nt, polymerase product is not observed in the absence of antisense ONT inhibitors. Concomitantly, significant decreases in the amount of full-length (-) strong stop DNA were noted, with full-length products comprising only 5-10% of the total DNA in the presence of each of the specific-sequence antisense ONTs. Interestingly, near optimal antisense ONT inhibition was noted at a ratio of 1:1 (ONT:donor RNA template) as illustrated in **Figure 4.3**.

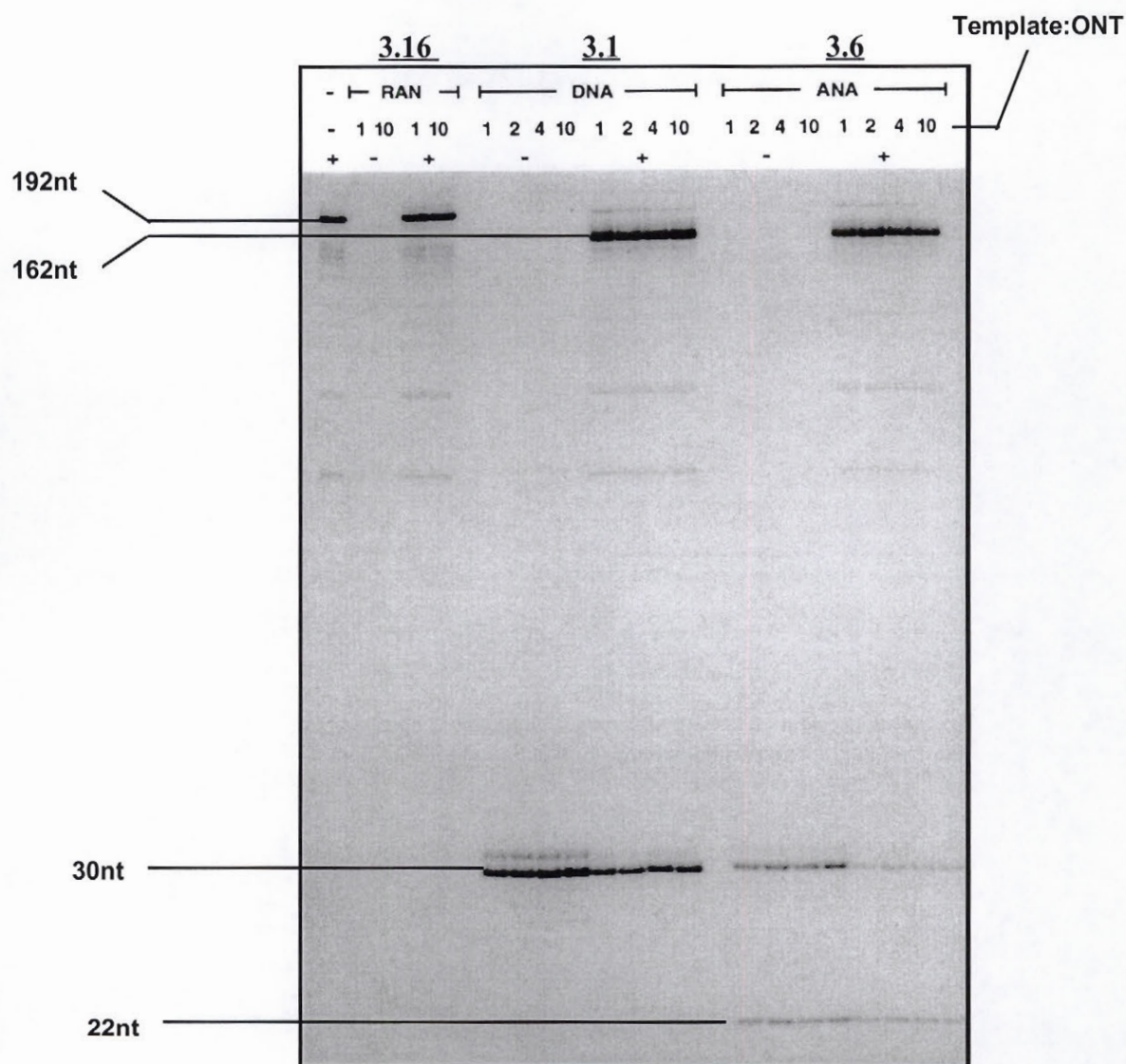


Figure 4.3: DNA polymerization product profiles in the presence of the various oligonucleotides. Reactions were carried out either with (+) a preannealed PBS primer:RNA template (2:1 molar ratio) or in the presence of the RNA template alone (-). Molar ratios of 1:1, 1:2, 1:4 or 1:10 (RNA template:ONT) were used, as indicated in the figure.

The polymerization product distribution in the presence of the random sequence antisense ONT RAN sequence **3.16** was similar to that noted in the absence of antisense oligonucleotides. Thus, the inhibition of (-) strong stop DNA synthesis by the antisense ONTs **3.1** and **3.6** are most likely due to specific interactions with the RNA template.

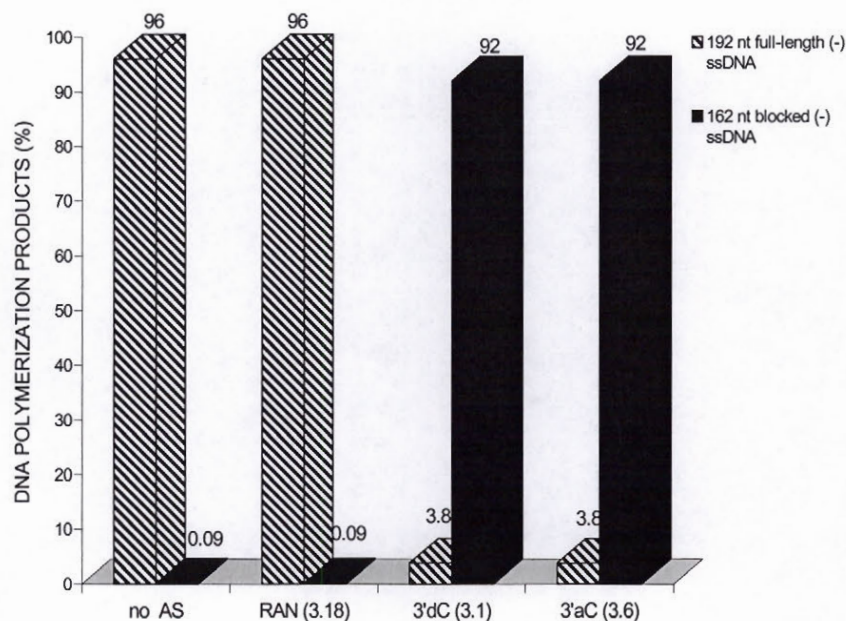


Figure 4.4: Graphical representation of the effect of antisense oligonucleotide RT-catalyzed full-length products illustrated in **Figure 4.3** and quantitated by densitometry. Open bars: full-length (-) strong stop DNA polymerization products (192 nt); solid bars oligonucleotide-blocked DNA polymerization product (162 nt).

Another 30 nt DNA product was noted in the reaction containing **3.1**, both in the absence and in the presence of the PBS primer (**Figure 4.3**). This product results from the RT-catalyzed 12 nt extension from the 3'-end of the antisense ONT, and is seen in the reactions containing chimera **3.6**, although the amount of this 30 nt product was significantly less than that seen in **3.1** DNA reactions. Interestingly, a smaller 22 nt product, corresponding to a 4 nt extension of the 18 nt ONT, was also noted in the reactions with the ara chimera seq **3.6**. This suggests that RT can incorporate 4 nt effectively and that the enzyme pauses or idles at this position on the T/P. This

phenomenon will be discussed in detail in Section 4.2. Neither the 22 nt nor the 30 nt products were noted in the reaction carried out with the RAN, indicating that this sequence could not serve as a polymerization primer for RT.

4.1.3 Discussion

The development of HIV-1 resistance in response to chemotherapy remains a major obstacle in the search for an effective treatment for AIDS. Antisense ONT therapeutics might be refractory to the development of resistance since single base changes in targeted sequences of viral genomic RNA are unlikely to significantly diminish the strength of ONT hybridization to target RNA. However, unmodified ssDNA and ssRNA are rapidly degraded by serum exonucleases, particularly 3'-exonucleases. In the present section, an oligonucleotide with a structural modification in the sugar component of the 3'- terminal nucleotide, namely β -D arabinocytosine, was examined. Arabinonucleotides have been shown to be more resistant to such exonucleases.¹⁰⁹ The nuclease-free *in vitro* system described allows rapid testing of potential efficacy of modified oligomers prior to initiation of more costly *in vivo* analyses.³⁰¹

Several sequences within the HIV-1 viral RNA have been explored for their suitability for serving as targets for antisense inhibitors.³⁰⁸ Previously, ONTs were primarily designed to block a single process in HIV-1 replication. The present work utilizes highly conserved targets in the genomic RNA. (**Figure 4.1**) These repeat (R) sequences are located at both the 5' and 3' ends of the viral RNA and they are essential to enable both the first and the second strand-transfer reactions required for proviral DNA synthesis. Moreover the RNA target is devoid of secondary structure and the potential mutability of the viral RNA sequence selected is very low.

Interaction of the R-region specific ONTs with the HIV-1 genomic RNA could inhibit HIV-1 reverse transcription at several levels as detailed in Section 4.1.1. Antisense inhibition produces a truncated (-) strong stop DNA product, which could be less effective in the first strand transfer process owing to the decreased complementarity with the acceptor RNA template as well as act as a physical block of this transfer. Indeed

this has been demonstrated recently in collaboration with Dr. M.A. Parniak of the McGill AIDS center.³⁰¹

HIV-1 RT binds single-stranded RNA and DNA with high affinity in a non-sequence specific manner.³⁰⁹ With appropriate cell delivery, the intracellular concentration of ONT therapeutics is likely to be substantially higher than that of HIV genomic RNA. As a result these ONTs may act as competitive inhibitors of RT-viral genomic RNA interaction, thereby slowing the rate of viral reverse transcription. In fact *in vitro* reverse transcription was decreased in the presence of excess antisense oligomer (**Figure 4.3**). Interestingly, near maximal inhibition (*in vitro*) was noted at an antisense ONT:target RNA template ratio of 1:1 even at the physiological temperature of 37°C. This interaction between target RNA and ONT is an important feature in such inhibition, because even at this temperature a balance between association and dissociation of the antisense ONT and the viral RNA occurs. As a result of this equilibrium, 100% blocking of reverse transcription in the *in vitro* assays is not possible, as seen by trace amounts of 192 nt product in **Figure 4.3**. Maximization of inhibition by the antisense ONTs at 37°C should occur by increasing the ratio between the antisense ONTs and the target RNA. However, when the ratio of ONTs:template was increased an overall decrease in the RT polymerization products was observed, even when the random ONTs (RAN) was used (data not shown). This is probably because of antisense ONT competition with the normal template/primer (T/P) for interaction with the RT template/primer cleft.

In addition, deoxynucleotide antisense ONTs of appropriate size may induce RNase H cleavage of the complementary RNA strand. This RNase H degradation of the RNA template would also result in a truncated form of the (-) ss DNA produced by elongation of the PBS primer and would be similar in size to that produced by antisense blockage of RT polymerization. The contribution of the antisense oligonucleotide directed RNase H degradation of the RNA template to the total amount of the 162 nt truncated (-) ss DNA was not assessable under these conditions, but nevertheless is suspected to be a minor contribution. Other investigators have shown that the R-region of HIV-1 genomic RNA is relatively refractory to RNase H degradation.³¹⁰

4.2 FUNCTIONAL CONSEQUENCES OF ARABINOSYL NUCLEOTIDE INSERTS IN DNA ON PRIMING BY HIV-1 RT

4.2.1 Introduction

Beardsley and Mikita¹⁴⁷ have undertaken a number of studies to investigate the effects of araC incorporation on biological phenomena. For example araCMP at the primer terminus dramatically reduces the rate of the next nucleotide addition by polymerases (*Escherichia coli* polymerase I, T4 polymerase, HeLa, cell polymerase α_2 and AMV reverse transcriptase). Polymerases with associated 3'-5' exonuclease activity preferentially excise araCMP from the primer terminus prior to chain elongation, and furthermore araCMP-terminated fragments are ligated more slowly than control fragments by T4 DNA ligases. *In vitro* araCTP inhibits various DNA polymerases, notably DNA polymerase α ,³¹¹ by competing with dCTP, and then acting as a chain terminator. In their most recent work¹⁹⁴ the authors have shown that araC incorporated at different positions in a series of DNA duplex substrates containing a T7 RNA polymerase promoter has different and dramatic outcomes. Transcription was very sensitive to incorporation of araC at positions before nucleotide 10 in the coding (sense) strand. This indicates that a narrow window of vulnerability exists, where transcription output is severely reduced (~100-fold) by a subtle DNA lesion caused by such araC incorporations. Recently, Thompson and Kuchta¹¹⁷ found that primers were elongated by up to 35 nt *via* efficient polymerization of corresponding araNTPs (alternate substrates) in place of missing dNTPs during elongation of *primase*-synthesized primers. Such araNTPs ended up in internucleotide linkages and did not result in chain termination. During elongation of *exogeneously* added template/primers however, araNTPs were not readily polymerized, and if they were, they resulted in strong chain termination. Together these results suggest that when araC is introduced into DNA oligomers, the small anomaly in the structure of the sugar (presence of the 2'OH) although not affecting base coding properties has a profound effect on the utilization of these templates or primers as substrates for DNA polymerases.³¹²

Based on previous and ongoing studies, one is led to believe that the resultant DNA duplex containing arabinonucleotide in one of the DNA strands is structurally perturbed. It is hoped that such studies involving perturbations on the processes of replication may also provide basic insights into the structural requirements of “reverse” transcription by HIV-1 RT. Reverse transcriptase being a DNA polymerase, requires a primer to initiate DNA synthesis. In this particular work the manner in which araX monomer substitutions create sites of DNA dysfunction will be reported by first studying the sensitivity of RT priming to such exogenous primers containing arabinose (in the absence of the PBS primer as opposed to the inhibition assay described in Section 4.1). Second, the effect of duplex conformation (DNA-RNA chimera):RNA on priming will also be investigated. Finally, an attempt to correlate the three - dimensional changes of the substrate duplexes with the products of HIV-1 RT priming *in vitro* will be addressed.

It should be noted that while Section 4.1 deals with the direct blockage of reverse transcription (DNA polymerization) by antisense oligomer, this section also deals with the synthesis of truncated (-) strong stop transcripts which could impair protein expression of RT; *i.e.* antisense ONTs that bind to the R-region at both-ends of the viral genomic RNA and “prime” DNA synthesis would result in the production of severely truncated proviral DNA. The resulting extended antisense oligomer would lack both the U region and the tRNA primer “tail” essential for the incorporation of the primer binding site into the (+) strand proviral DNA.

4.2.2 Results and Discussion

Analysis of the Priming Products from Sequencing Gels:

Sequences **3.1** to **3.15** (Chapter 3.1) were examined in fixed time assays as described in the Experimental Section (6.6.1).³⁰¹ Briefly, the template was mixed with the 18mer antisense ONTs in the absence of the PBS primer for the appropriate times (**Figure 4.5**). The reaction mixtures contained 50 pmol pre-formed T/P, and each of the four dNTPs, with the tracer dCTP [α - ³²P]. After fixed incubation times the reactions were stopped and then resolved on a sequencing gel (**Figure 4.6**). The total dNMP incorporations were determined by two methods namely densitometry and liquid scintillation.

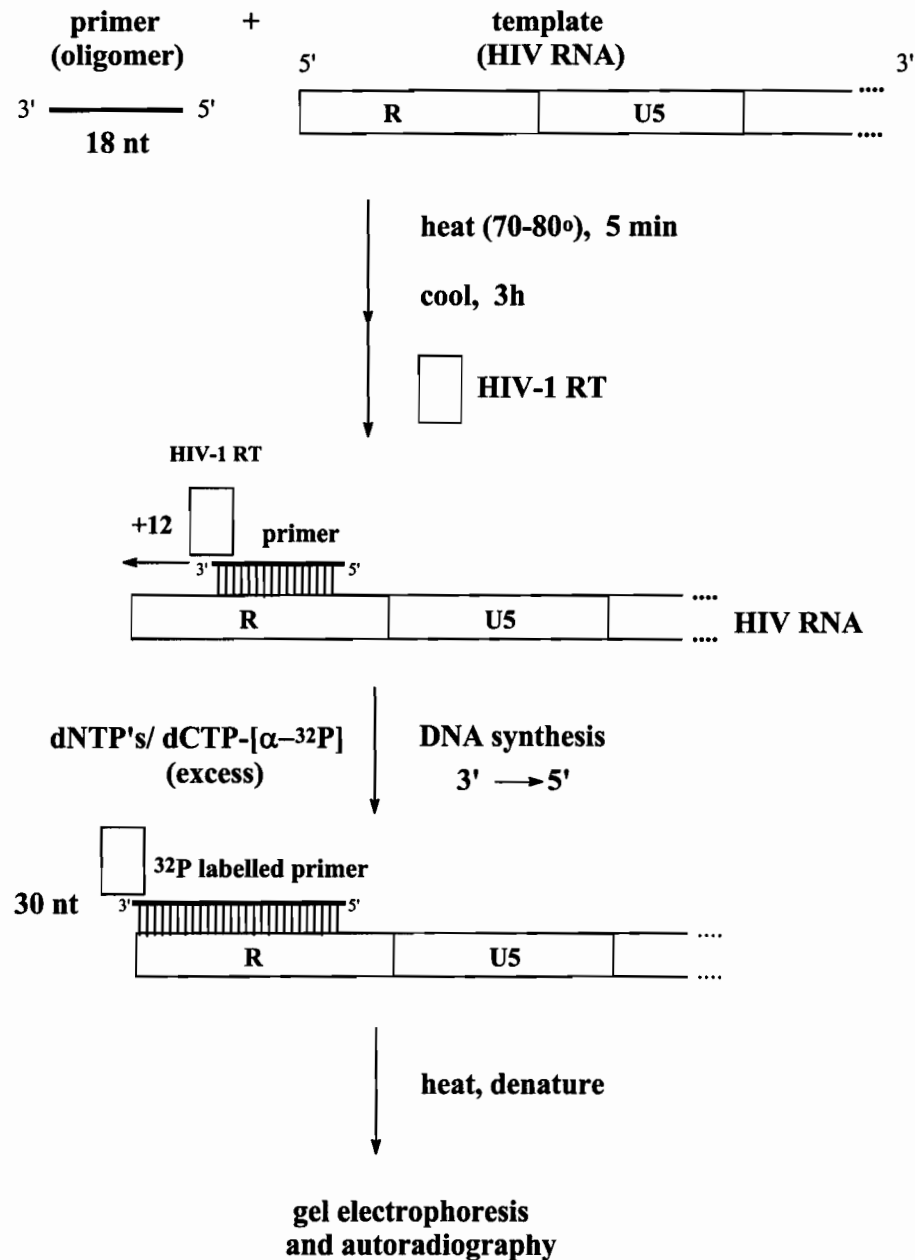


Figure 4.5: *In vitro* priming assay. For details refer to the experimental section 6.5.3.

Effect of incorporation of arabino inserts into priming strands

From the gel (**Figure 4.6**) one observes that an araX nt at the 3' terminus of a primer (sequence **3.6**) slows down the priming.

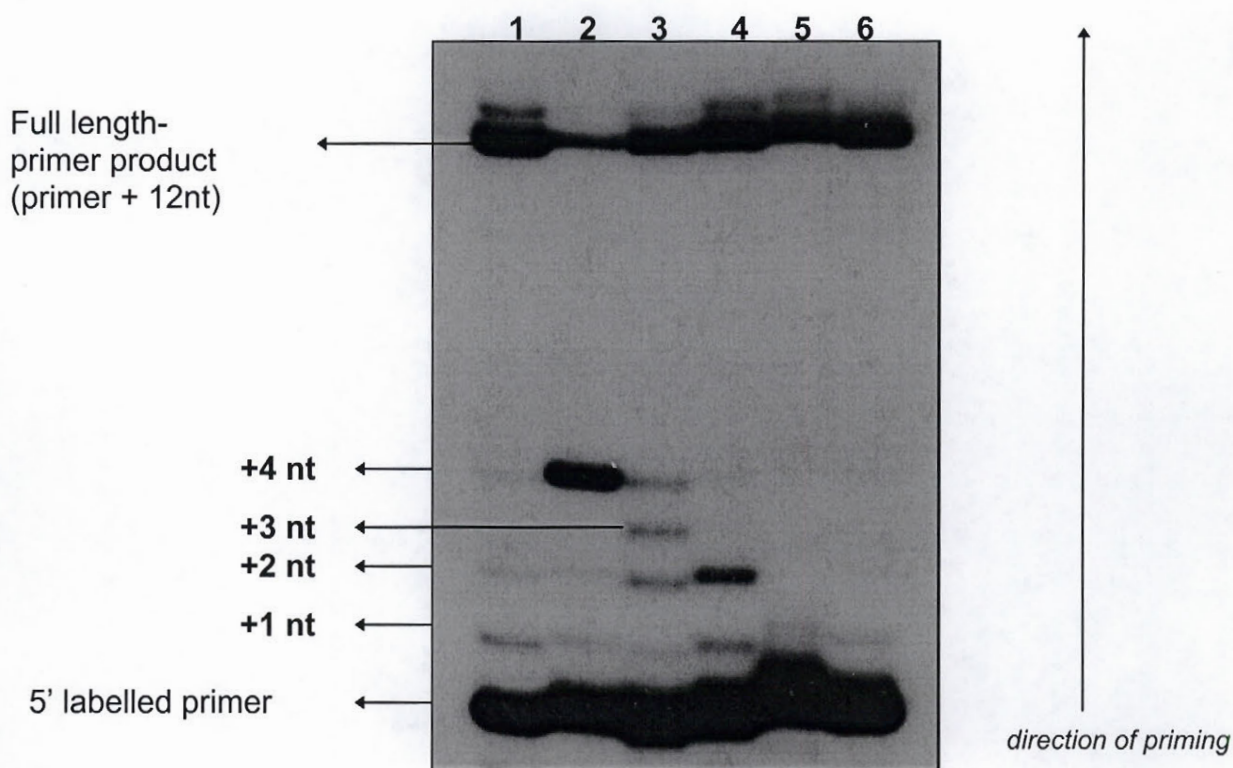


Figure 4.6: Autoradiogram of *in vitro* priming assay: extension of chimeric DNA-ANA primers each annealed separately with RNA templates. The ^{32}P -labelled chimeras were used as primers for polymerase reactions catalyzed by HIV-1 RT. Lanes 1 through 6 show addition products resulting from the extension of the primers catalyzed by RT. Lanes: (1) DNA (3.1), (2) aC₁ 3.6, (3) aU₂ 3.2, (4) aA₃ 3.3 (5) aG₄ 3.4 and (6) aA₅ 3.5.

Designation	Sequence
aC ₁	5'-AGC TCC CAG GCT CAG AT <u>a</u> C -3'
aU ₂	5'-AGC TCC CAG GCT CAG A <u>a</u> UC -3'
aA ₃	5'-AGC TCC CAG GCT CAG <u>a</u> ATC -3'
aG ₄	5'-AGC TCC CAG GCT CA <u>a</u> G ATC -3'
aA ₅	5'-AGC TCC CAG GCT C <u>a</u> AG ATC -3'

The total amount of completely primed product decreases to 30 % and instead a predominant +4 product is observed. When the araX nt is moved more into the interior of the primer strand, priming increases but with the corresponding changes in the other side products to +3, +2 and +1 nt. This pattern appears to correlate with the sequential movement of the ara nt away from the priming end. There appears to be no change in priming relative to the DNA control at and past the fifth position (compare lanes 1 with 6 in **Figure 4.6**). The insertion of the arabinose sugar in the primer strand though not affecting thermodynamic stability dramatically (Chapter 3, Section 1), exhibits a striking feature of a partial priming block that occurs within a defined window of 5 nt from the 3' end. **Figures 4.3** and **4.6** show that HIV-1 RT readily extends the antisense ONT DNA to form a 30 nt product. This primer extension was significantly reduced with the oligonucleotide containing the arabinose residue at the 3' end. Interestingly, an ONT containing riboC at the 3'-end (D17R1) was fully extended to the 30 nt product without accumulation of the +4 band (22 nt product).

Thus there seem to be two aspects to the inhibition process: (a) the presence and the stereochemistry of the extra 2' hydroxyl on the 3' terminal nucleotide; (b) the presence of the additional 2' hydroxyl on an arabino residue "within" the antisense strand. One hypothesis is that there are important polymerase-DNA contacts along the sugar-phosphate groups of the DNA which are required for normal binding and activity of the enzyme but which are somehow perturbed by the presence of the arabinose sugar. When araX is located at the 3' terminus of the primer, the aberrant sugar moiety of the nucleotide is directly involved in the success or failure of chain elongation. It is possible that the additional hydroxyl group, while not affecting the stability of the primer terminus, interferes directly with the mechanism by which the polymerase catalyzes phosphodiester bond formation between the 3'-hydroxyl of the terminal sugar moiety and the α phosphorus of the incoming nucleotide. However when araX occupies an internucleotide position in the primer, the sugar moiety serves only as a structural component of the T/P sugar-phosphate backbone and is not directly involved in phosphodiester bond formation. Moreover these sugar-phosphate contacts might also be critical for the processive functioning of the enzyme, *i.e.* some tertiary structure of the enzyme may contact the last 4 nt pairs of the duplex and this causes stalling or pausing,

with subsequent dissociation of the RT from the T/P complex. At the fifth position there is no apparent pausing suggesting a lack of steric contacts with the enzyme. At positions 4, 3, 2 and at the 3' terminus, +1, +2, +3 and +4 extended products are formed respectively, with possibly the simultaneous dissociation of the RT from the substrate duplexes (**Figure 4.7** and **Figure 4.8**).

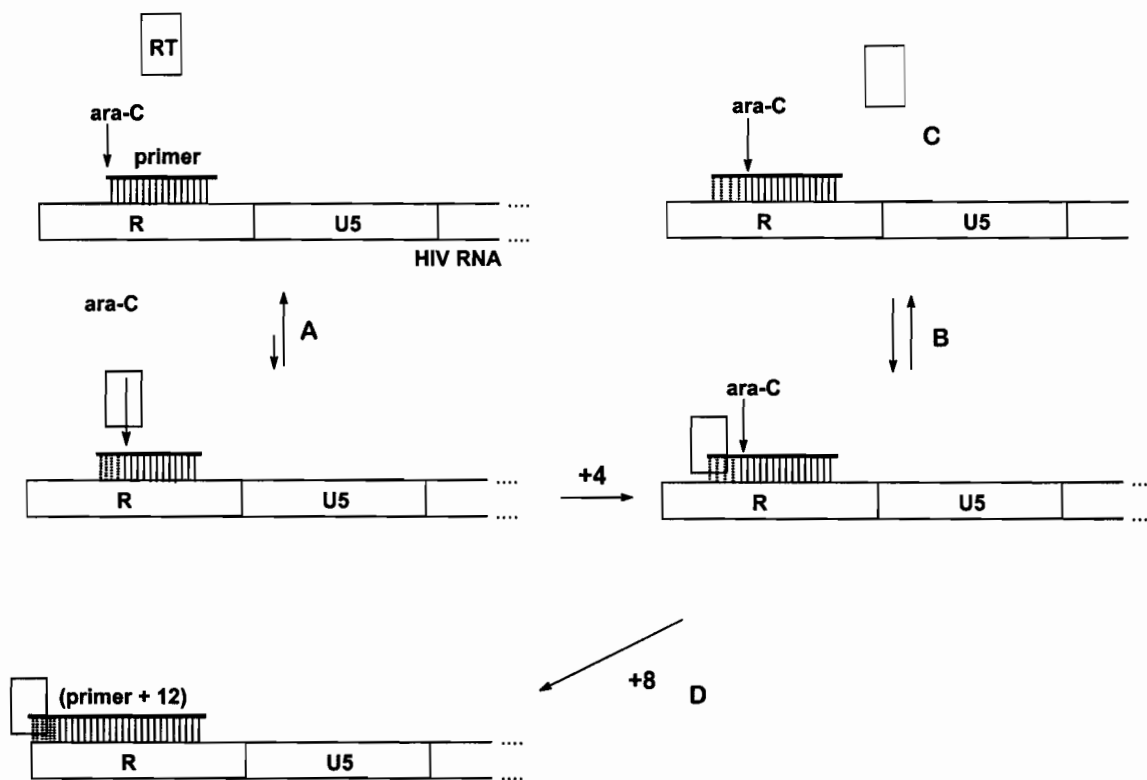


Figure 4.7: (A) Association, (B) pausing, (C) dissociation and (D) extension of reverse transcriptase with the araC terminal nucleotide to the primer:template duplex.

All these cases indicate that the enzyme encounters a probable steric interference with a window equivalent to about half a turn of the duplex. The pausing activity by the araNTPs which differ from the native DNA sequence by the inclusion of a single stereocenter (C2'), is surprising and suggests that the araNTPs ($n = 2$ to 4), even though internal, are still intimately involved in the polymerase mechanism. Consistent with this pausing one observes an obvious trend in the amplitudes and wavelengths of the maxima in the CD spectra, discussed in Chapter 3, Section 1 (**Figure 3.1.5 A and B** pp. 61). The overall conformation is more A like and this interestingly A-helical nature appears to

diminish when the ara insert is at the fourth position from the 3' end. *i.e.* Loss of A character is evident in the amplitude decrease of the key negative peak at 210 nm as the araX shifts into the interior of the strand. Moreover it is clear that the identity of the ara sugar and not the specific nucleobase moiety is responsible for the pausing of RT, *i.e.* both araA residues at position 3 and 5 behave differently in the polymerase active site of the enzyme.

The araX nt near the 5' end of the primer (position 16) does not appear too critical for efficient priming. This could be interpreted in two ways: (i) the availability of more nts beyond the dysfunctional site may increase the affinity of the RT for its substrate. This differential binding affinity could manifest itself as either a reduced rate of dissociation from the longer template when a block to synthesis is encountered or a higher rate of reassociation with the partially extended substrate. In other words, the ara is locked in, and may regenerate the formation of normal base-pairs resulting in increased bypass frequency. This is probably true for nts from the 5' terminus to the position 5 (dA) on the oligomer; (ii) alternatively the 5' end of the primer may lie at the RNase H active site while the 3' terminus is at the polymerase active site. *i.e.* the araX at the 5' end should not severely affect priming.

The results presented above are consistent with previous reports of Beardsley¹⁴⁷ who suggest that the RTs are less sensitive to the presence of araCMP in the *template* relative to other DNA polymerases. This permissivity may be related to the much higher spontaneous error rate this enzyme exhibits *in vitro* compared to that of the other polymerases because it lacks a 3'-5' exonuclease activity.³¹³ On the other hand Gotfredson and co-workers have recently illustrated that when 2'-OCH₃-ara-T is positioned in an oligonucleotide it favours resistance to the 3'-exonuclease SVPDE (a dual DNA/RNA nuclease).^{112,113} The increased stability may be caused by steric interference from the 2'-OCH₃ group in the arabino configuration so that the shape of the oligomer does not fit into the active site of the enzyme. Thus, since the hydrophilic 2'-OH is tucked in a hydrophobic cavity (major groove), it is possible that HIV-1 RT (DNA-binding protein) may have amino acid side chains that interact unfavorably with this hydrophobic area. As a result such enzyme binding to the araX substituted DNA may not be nearly as well as the normal DNA primer.

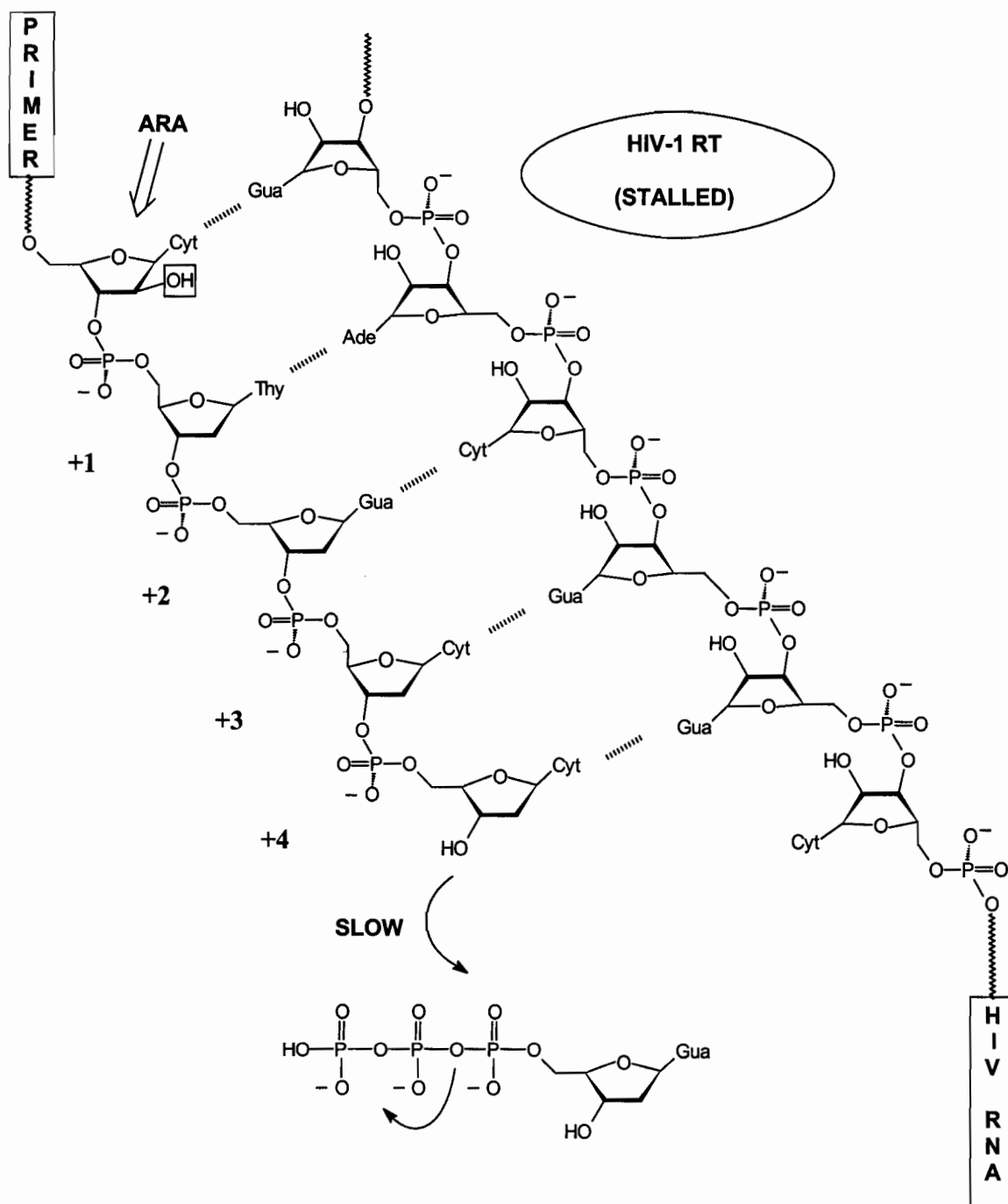


Figure 4.8: Pausing or dissociation of the HIV-1 RT after extension of the 4 nucleotides in the primer strand as a result of the arabinocytosine residue at the 3' terminus of the primer.

At the molecular level these steric effects can be interpreted as follows. The O2' oxygen of araX which lies at the edges of the major groove of the helix appears to be in close contact with the residue on the 3' side. Such a close contact (a little shorter than the sum of the Van der Waals radii) could push the 3' base slightly away from it causing a small destacking between the two bases.¹⁴⁹ This adjustment of the strand conformation in response to the araX structure could further involve the movement of the entire motif of the nucleotide components (*i.e.* either as a buckle or kink effect) to a distorted conformation of the nt in the complementary strand opposite the araX.

Factors Affecting Priming With Respect To DNA/RNA Placement And Length In The Primer-Chimeras.

This study deals with the DNA-RNA chimera:hybrids. A comparison of the product of seven primers (D18, D13R5, D9R9, R9D9, R17D1, D17R1 and R18) indicate that only the D18, D17R1 and R9D9 act as primers for HIV-1 RT. Note that the ribose substitutions are at contiguous positions and at either end of the primer molecule. The choice of these sequences allows one to study the effect of "absolute identity" of sugar as well as "location" within the priming strands on polymerization. A summary of the results of the priming from the DNA-RNA chimera study are shown in **Table 4.1**.

Table 4.1: Priming Of Antisense Oligomers By HIV-1 RT

Designation	Sequence	Priming
D18	5'-AGC TCC CAG GCT CAG ATC -3'	++++
D17R1	5'-AGC TCC CAG GCT CAG ATrC -3'	+++
D13R5	5'-AGC TCC CAG GCT Cr(AG AUC) -3'	-
D9R9	5'-AGC TCC CAG r(GCUCAG AUC) -3'	-
R9D9	5'-r(AGC UCC CAG) GCT CAG ATC -3'	+++
R17D1	5'-r(AGC UCC CAG GCU CAG AU)C -3'	-
R18	5'-r(AGC UCC CAG GCU CAG AUC) -3'	-

Conditions are the same as reported in Figures 4.5 and 4.6; extent measured by densitometry and liquid scintillation.

In addition to absolute identity of the sugars there seems to be two distinct overall conformations at termini which appear to govern the priming process. When ribose residues are incorporated at the 3' end of the primers such as D9R9 or D13R5 the priming appears to be significantly retarded or abolished.¹⁴⁷ On the other hand chimera R9D9 annealed to the RNA template was efficiently used as a primer to catalyze the nt addition reaction by HIV-1 RT. The results from **Table 4.1** strongly suggest that a necessary requirement for polymerization by the RT, is that at least five nts in the duplex at the 3' end of the primer be in the “hybrid form” and not in the pure A RNA form. However at least one ribo-residue at the 3' terminus can be accommodated by the polymerase active site of the RT.

Although it is known that HIV-1 RT may have a relatively higher affinity for RNA than for DNA oligonucleotides bound to a template, the RT prefers to extend RNA primers only when sequences are very similar to that of the polypurine tract (PPT) and are unable to extend RNA primers of other sequences.³¹⁴ This ensures that the site of initiation is very specific. In the case of the R region it is possible that the nature of the RNA primer is similarly recognized but not polymerized.

The above related observations may be associated with a number of factors, including the structural distortion caused by ribose, its interference with polymerase action, the induced change in DNA properties (such as conformational rigidity, helical diameter) or a combination of these factors. This is apparent from the absence of priming observed with the R17D1:RNA substrate. From all of these results it appears that it is not only the phosphate backbone of T/P that is solely essential for recognition and binding by the polymerase class of enzymes, but also the diameter of the duplex region which has to be accommodated in the template binding groove. Thus it is possible that R9D9:RNA and D9R9:RNA have different diameters of duplex region at the T/P binding groove and this is consistent with the CD profiles of the two complexes where D9R9:RNA is more B like, while R9D9:RNA exhibits an A type helical conformation **Figure 3.1.6 (A and B).**³¹⁵

4.2.3 Conclusions and Future Directions

DNA (with ara inserts) and some of the DNA-RNA chimera targeted to the R region of the viral RNA genome are able to pause the reverse transcription of HIV-1 RT suggesting that these molecules when bound to RNA template, may be oriented in the polymerase mode of the enzyme. However it is clear from the patterns in both the priming experiments and the CD spectra that there is a correlation between the structures and function which arise from different sources within these two studies. Internalizing the ara inserts causes a direct steric interaction with the enzyme's polymerase active site that lies closer to the 3' terminus of the primer (+4 and smaller products are due the pausing of the RT as it encounters the additional 2'-OH relative to deoxyC). This is accompanied by a decrease in the A-type duplex character as shown in **Figure 3.1.5A** and **B**, by CD spectroscopy (Chapter 3. Section 1). D17R1 which had a riboC instead of araC at the terminus did prime which indicates that RT polymerization is also sensitive to the 2'-stereochemistry at the 3'-terminus. However changing the chimeric combinations of RNA and DNA within the priming strand may have less apparent potential for such steric perturbation but directly affects global helical conformation which is itself also important for the priming process (**Figure 3.1.6**). In this context the completely modified ANA oligomer (**3.15**), which is not a substrate for the polymerase of the RT, perhaps has an overall duplex conformation that is not the typical hybrid for the polymerase activity of RT, but is something different and nonetheless useful for inducing other properties namely the activation of RNase H (discussed in the next section).

It is still unclear how the small structural modification in the DNA backbone could affect the catalytic function of the enzyme and/or binding affinity for the T/P. The impact of introducing the araX into "RNA strands" to be used as a primer or as a template probe will be yet another important issue. Although incorporation of an araX residues into the core sequence displaces the associated 2'-OH relative to the corresponding ribose derivative by approximately 109° (as a result of the inversion of the C2' stereocenter), the associated changes in sugar puckering could result in more extensive changes in location of the OH function (180°) when present in an all-RNA sequence. Thus these perturbations are expected to be more severe for the RNA double helix. Presumably the RNA polymerase would not be able to insert an araC into the

elongating RNA chain (transcription) due to steric hindrances; this could explain why araC is not incorporated into RNA molecules. A logical continuation of this study is to test the inhibitive capacity of antisense DNA-araX3' chimeras in combination with antiviral agents currently in use (*e.g.* AZT, 3TC) on the replication of retroviruses such as HIV. If these oligonucleotide analogues could direct enhanced incorporation of chain-terminating dideoxynucleoside drugs to specific positions during DNA synthesis by viral and cellular DNA polymerases, then this strategy could provide therapeutic benefits in the treatment of diseases such as HIV-1 infection and rapidly proliferating cancers.

4.3 MODULATION OF RIBONUCLEASE H (RNase H) ACTIVITY BY HIV-1 RT USING ARABINONUCLEIC ACIDS

4.3.1 Introduction

RNase H activity has been the subject of extensive studies. RNase H forms a family of enzymes with both endonuclease and 3'-5' exonuclease activities. It is unique in that it specifically cleaves the RNA strand of DNA:RNA duplexes in the presence of divalent ions (Mg^{2+} or Mn^{2+}),³¹⁶ to yield a 5'-phosphate and a 3'-hydroxyl at the hydrolysis sites.^{317,318} The RNase H activity of RT plays an important role in the life cycle of HIV-1 by directing DNA replication, and possibly in the destruction of mRNA during antisense gene regulation.^{23,319} (Figure 4.1 and 4.9 below).

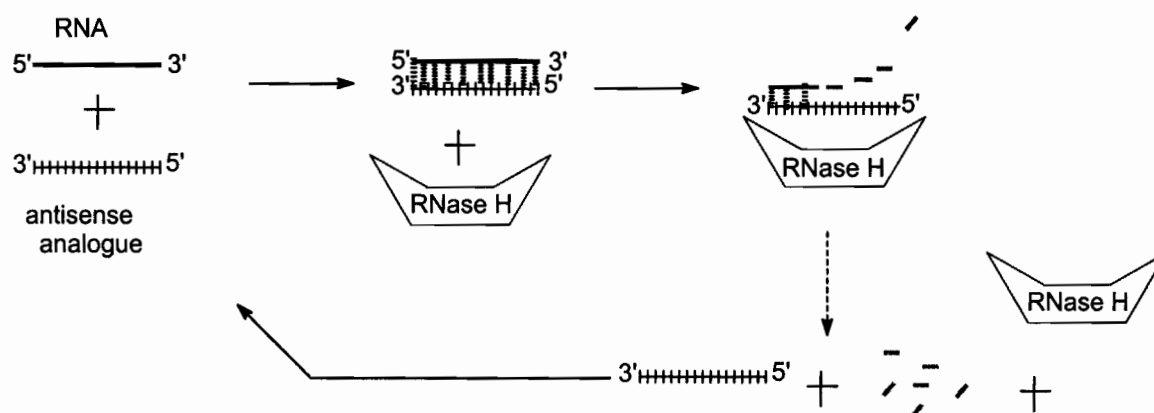


Figure 4.9: Induction of RNase H Activity. Proposed mechanism of action of degradation of the RNA template of DNA:RNA hybrids by RNase H. RNase H cleaves only the RNA target strand but this action is dependent on the structure of antisense analogues, *e.g.* antisense DNA and thioate-DNA, but not RNA, induce cleavage.

The *E. coli* enzyme consists of a single polypeptide chain with 155 amino acid residues and exhibits sequence similarity to the RNase H domains of retroviral RTs, including HIV-1.^{317,320} The crystal structure of *E. coli* RNase H^{321,322} demonstrates that the enzyme has a handle region with a high percentage of cationic amino acids (Cys¹³, Asn^{14,44} and Gln⁷²) which present a highly positive charge within a confined space.³²³ This handle is probably responsible for the initial non-specific binding of the enzyme to the heteroduplex structure *via* electrostatic interactions. On the other hand HIV-1 RT³²⁴ lacks this handle and therefore cannot bind to a duplex substrate when separated from the rest of the RT. The Mg²⁺ binding site of *E. coli* RNase H is located in the β sheet and consists of three amino acids, Asp¹⁰, Glu⁴⁸ and Asp.^{70,323} The three catalytic acidic residues, the Mg²⁺ ion and the hybrid substrate bind close to each other within this active site.³¹⁷

If the substrate is of high molecular weight (> 28 nt) then both the endonuclease and processive 3'-5' exonuclease activity for *E. coli* RNase H act on it, while only the endonuclease activity is observed to act on the low molecular weight substrate.³²⁵ The enzyme is thought to bind in the minor groove of a DNA:RNA double helix by recognizing the *sugar conformation* of the DNA nucleotide complementary to the site of RNA cleavage.^{323,326} Moreover pendant groups located in the major groove have no significant effect on the enzyme-substrate binding in the minor groove; this agrees with the effects of minor-groove modifications reported recently by Crooke.³²⁷ Thus it has been proposed that hydrolysis is initiated throughout the heteroduplex region except in the inhibited areas that result from substituents in the minor groove.³²⁸

For HIV-1 RT, the physical distance spanning the RNase H and Polymerase active sites is about 18-20 nt.³²⁹ These two domains are interdependent during the entire process. Thus the RNase H activity itself has been classified as polymerase-dependent and polymerase-independent RNase H.³³⁰ Polymerase-dependent RNase H cleavage advances upon primer extension yielding fragments of 18 base pairs, and remaining at a fixed distance from the 3'-OH terminus of the elongating DNA. This activity is not sufficient to eliminate all of the template RNA,^{331,332} and is not an absolute pre-requisite for (+) strand DNA synthesis. Thus RNA fragments that remain annealed to the (-) strand DNA are generally degraded before or during synthesis of the (+) strand DNA by

the polymerase-independent activity of RNase H, which yields fragments as small as seven bases. Fuentes and coworkers³⁰² have shown that if the RNA:DNA hybrid is too small to be recognized by the RNase H domain alone, it may not be cleaved at all.

During reverse transcription a variety of nucleic acid structures are created: RNA:RNA, RNA:DNA and DNA:DNA hybrids are present simultaneously.³⁰² In order for genomic replication *via* reverse transcription to be successful, RT must perform its multifunctional activities in an appropriate order. One of the ways in which the RT could achieve this is to have an order of preference for binding and reacting with these structures. The relative displacement of the two active sites along the T/P duplex implies that during the first steps of DNA synthesis, the RNase H domain of HIV-RT is positioned over the RNA:RNA homoduplex, and only after 18-20 steps will it encounter the junction between primer tRNA and the nascent DNA strand.³³² A residual RNA of about 18 base pairs remains undigested during the DNA synthesis when RT reaches the 5' end of the template. Presumably the conformation of the tRNA-DNA "junction" provides a particularly good target for RNase H activity irrespective of whether chimeric RNA-DNA structure is complexed with (+) strand RNA (initiation of the reverse transcription) or (-) strand DNA (primer removal).³³² Taken together these results indicate that during initiation of (-) strong stop DNA synthesis by HIV-1 RT, the first RNase H mediated endonucleolytic cut of the genomic RNA is dictated mainly by the length of the nascent DNA and not by sequence except that of the polypurine tract.

The principle aim of this study was to explore the impact of chemically modified oligonucleotides on RNase H cleavage ability of HIV-1 RT and *E. coli*. These oligonucleotides were designed to be completely modified. ANA was shown to form a stable duplex with the RNA target. In a related report, metal identity and concentration in addition to these synthetic modifications were shown to further effect the extent of hydrolysis by different RTs.³³³ UV and CD spectroscopy, the main tools used to study the synthetic duplex model, have been discussed earlier in Section 3.1. ANA (3.15) and the corresponding DNA (3.1) and RNA (3.14) analogues were studied in order to probe the importance of the 2'-OH stereochemistry in RNase H mediated cleavage of the duplex structures.

4.3.2 Results

In order to test whether ANA induces RNase H cleavage of a RNA target when bound to it, RNase H assays were performed. They were conducted by Dr. G. Borkow and D. Arion at the McGill AIDS center. The thermal dissociation (T_m) data for the duplexes formed between ANA (3.15) and complementary RNA, and those of the controls, are summarized in Table 3.3 of Chapter 3.1. Figure 4.10 shows the products from a sequencing gel, of the HIV-1 RT and *E. coli* RNase H degradation as a function of time. Panels (A) and (C) represent digestion with HIV-1 RT RNase H with 10 mM $MgCl_2$ and 0.1 mM $MnCl_2$ respectively; Panels (B) and (D) represent *E. coli* RNase H with 10 mM $MgCl_2$ and 0.1 mM $MnCl_2$. As seen in autoradiograph, *E. coli* and HIV-1 RT associated RNase H was able to cleave the RNA of a DNA:RNA hybrid efficiently as observed by the disappearance of the full length ^{32}P -RNA target under these conditions. However the patterns of hydrolysis differed between the two enzymes as well as between the two substrates, namely DNA:RNA and ANA:RNA.

In the case of ANA, degradation of the RNA was also observed, with 100-250 μ mol of $MnCl_2$ being the optimal amount. It is clear that RNase H cleavage activity of RNA in the ANA:RNA duplex is more pronounced in $MnCl_2$ than in $MgCl_2$. ANA:RNA also appears to be a better substrate of the *E. coli* enzyme than HIV-1 RT as indicated by the nearly complete disappearance of the target. In fact according to the scanning results the cleavage of RNA template when bound to ANA for *E. coli* was $\sim 90\%$ (Mn^{2+}) and $\sim 60\%$ (Mg^{2+}).

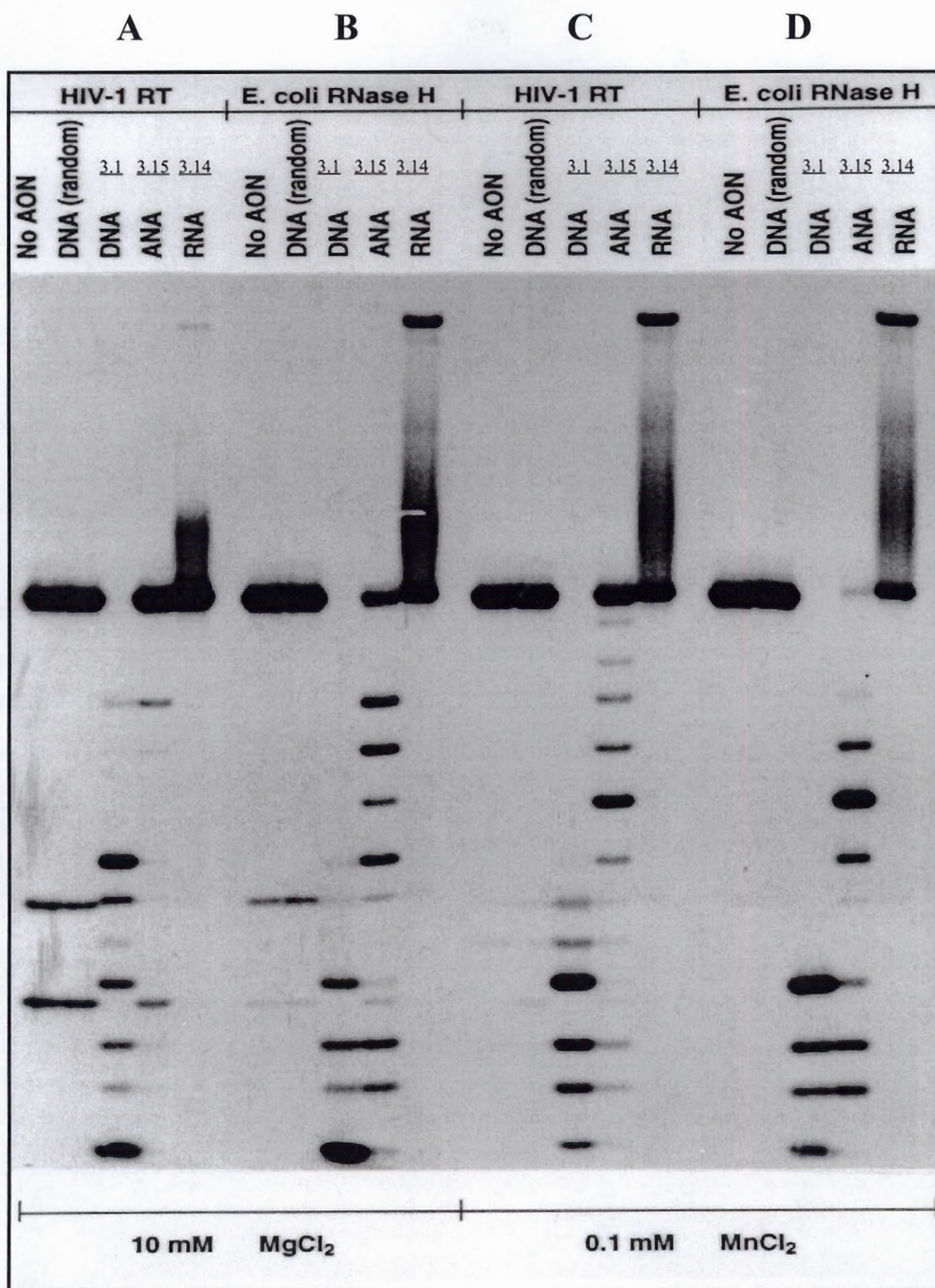


Figure 4.10: Ribonuclease H degradation of various 18-bp oligonucleotide hybrid duplexes (37°C). An 18-nt 5'-³²P-labeled 3'-5' RNA was preannealed with the complementary DNA (3.1), RNA (3.14), ANA (3.15) and RAN (3.16, is the random DNA) oligonucleotides and then added to reaction assays containing HIV-1 RT or *E. coli* RNase H at the given concentrations of MgCl₂ (10 mM) and MnCl₂ (0.1mM). Panels A and B: HIV-1 RT RNase H; Panels C and D: *E. coli*. RNase H. No AON means no antisense oligonucleotide. See Materials and Methods for experimental conditions.

Note also that the RNA (3.14) did not induce RNase H cleavage under the conditions used, as expected *i.e.* the RNA:RNA duplex is not a substrate for RNase H activity. In fact the RNA:RNA ($T_m \sim 86^\circ\text{C}$) was observed on the sequencing gel running as a high molecular weight species, despite the denaturing conditions used to run the gel (and additional boiling to 100°C prior to loading).

4.3.3 Discussion

Structural properties of chemically modified heteroduplexes.

The detailed work of Jacobo-Molina and coworkers³³⁴ infer that a duplex DNA (18 base pair region) bound to the polymerase and RNase H sites adopt both A and B form geometry. The A form corresponds to the first 5-7 bp (-1 to -5/-7) at the DNA polymerase domain, followed by a bend (40° - 45°) and the remainder of the nucleic acid is primarily B (-10 to -18).^{335,336} After RNase H binds to the duplex, the enzyme needs a way to identify true substrates (*i.e.* DNA:RNA duplexes *vs.* non substrates such as RNA:RNA duplexes). A recognition handle region (positive charges) lies near the active site and apparently makes the key sugar contacts that identify a suitable substrate.^{323,326} Although HIV-1 RT RNase H can differentiate between (DNA:RNA) hybrid and the (RNA:RNA) homoduplex, both appear to traverse the substrate binding groove between the active sites similarly. When RT is arrested, the RNA template is cleaved at a fixed distance of 18 bp from the primer terminus, independently of whether the duplex between the two active sites is entirely dsRNA, partly RNA:DNA and partly RNA:RNA or entirely RNA:DNA. This suggests that the duplex is bound similarly to the enzyme regardless of its precise composition and the discrimination between RNA:DNA and dsRNA is accomplished *locally* near the RNase H active site, rather than *globally* throughout the general interaction of the duplex with RT. Moreover the entire binding interaction appears to comprise a single helical turn of the substrate duplex.

Although there seems to be a lot of controversy as to whether the hybrids are of the A or B forms, Salazar *et al.*³³⁷ have shown that the DNA strand in the hybrid is neither in the B nor the A form, but in an intermediate form in solution (DNA sugars *O4'-endo*, RNA sugars *C3'-endo*). Despite the similarities in helical parameters of both

duplexes, the width of the minor groove is narrower for the hybrid³²⁶ than for the RNA homoduplex, leading the authors to propose that *E. coli* RNase H is unable to cleave dsRNA because critical interactions cannot form with a duplex in a wide minor groove A-type conformation. In the DNA:RNA duplex both nucleic acid strands are contacted by RNase H and this triggers cleavage. Therefore, although *E. coli* RNase H does not display discernible activity against dsRNA, the relative substrate discrimination mechanism of RT-RNase H and *E. coli* RNase H may not in fact be substantially different, both being mediated by local structural differences between dsRNA and RNA:DNA as predicted for the *E. coli* enzyme. For the RT-associated RNase H activity, such local differences in structures may then be necessary, but not sufficient for substrate discrimination in the context of reverse transcription. RT must accept dsRNA as a substrate in order to initiate reverse transcription. Consequently polymerization to rapidly remove the RNase H domain from the dsRNA target can be seen as an additional requirement for retrovirus-associated RNase H substrate discrimination.³³² Based on solution NMR studies, Reid reported that more than four successive DNA residues were required to generate this active conformation in the RNA:DNA duplex, which is consistent with the inhibition when the length of successive DNA residues is less than four.³²⁵

The ability of retroviral RT to degrade heteropolymeric duplexes has been demonstrated in this section. In this study a *uniformly* sugar-modified analogue, *i.e.* ANA, has been shown to induce RNase H activity. To our knowledge, ANA represents the first such example, despite the fact that a very large number of studies have focused on the design of antisense molecules that induce RNase H activity. Popular design motifs include “chimeric” or “splint” oligonucleotides that are capped at either one or both ends, with a contiguous string of 2'OMe ribonucleotides. These methylated “wings” enhance the metabolic stability, binding affinity and protection from exonuclease degradation. The remaining internal sequence or ‘gap’ is then a phosphorothioate oligodeoxynucleotide which serves as the substrate for RNase H.³²⁸ However it appears that the same properties that enhance the binding affinity 2'OMe RNA or 2'F-RNA for RNA, result in a loss of RNase H activity.³²⁸

Nature of enzyme substrate interaction

There are two binding steps in this process: (a) Binding between the individual strands from which the T'_m is a reliable indication and (b) binding of enzymes to the heteroduplex, which depend on duplex dimensions such as groove widths and helical diameter. Crooke and coworkers claim that the second binding interaction between the enzyme and substrate although predominantly electrostatic, is more specific than a mere attraction between a positively charged protein surface and polyanions.³²⁷ Moreover the enzyme binds preferentially to phosphate groups which are spatially oriented within a helical structure as opposed to a random coil as evidenced by the higher binding to the RNA homoduplex over single stranded RNA. In addition it is not only a helix, but also the appropriate helical conformation that appears to be important, with greater binding affinity exhibited for A-type helices (duplex RNA) than for B-form helices (duplex DNA).

From **Figure 4.10** it is clear that the RNA:RNA complexes did not elicit any RNase H activity, in spite of the complex adopting a typical A- type conformation (**Figure 3.1.8.B**). Thus, although RNA homoduplexes adopt a preferred conformation with respect to RNase H binding, their 2'OH's are positioned within the minor groove where they presumably interfere with enzyme recognition and the ensuing cleavage (hydrolysis step). This is not the case for ANA which displays its 2'OH groups in the major groove of the hybrid (**Figure 4.11**).

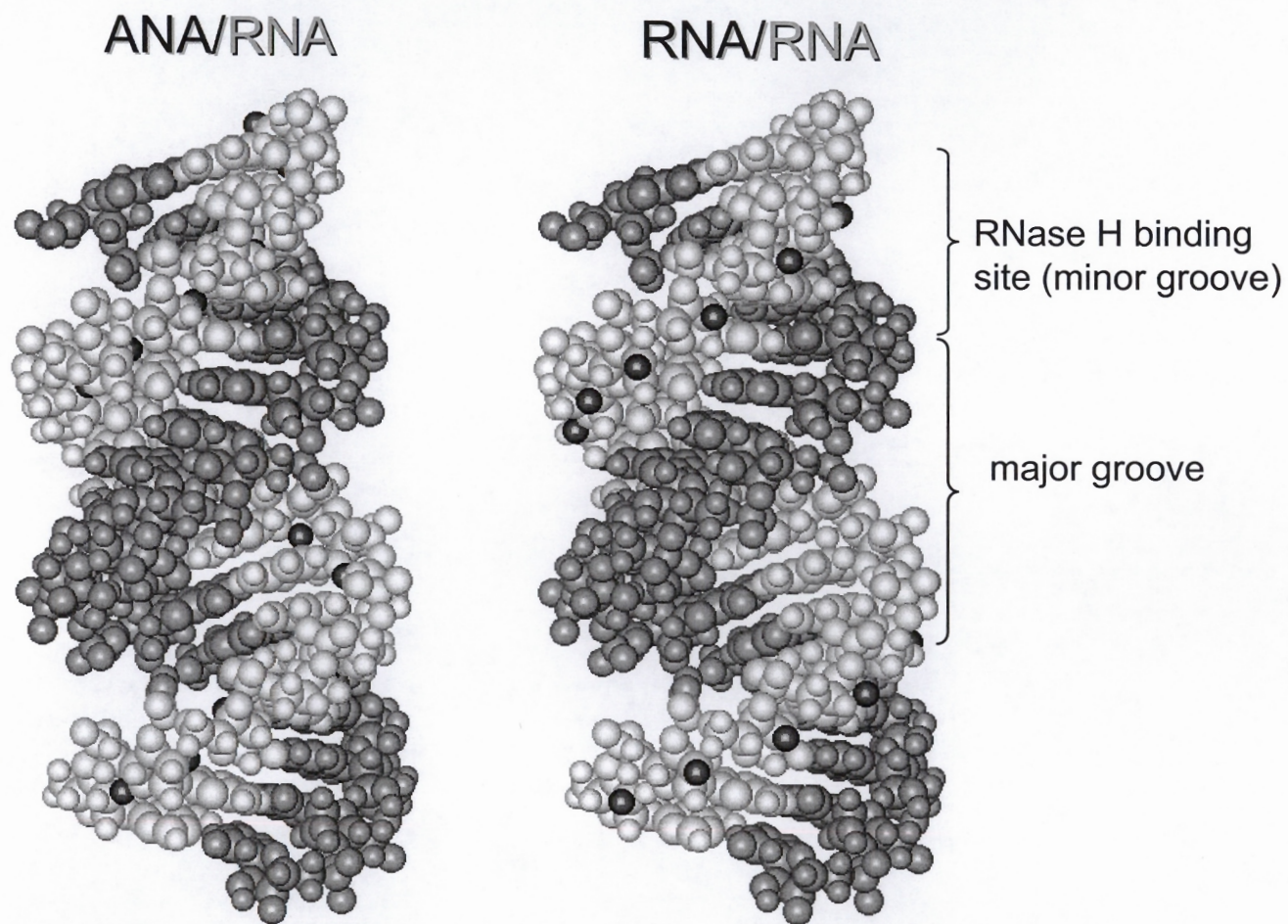


Figure 4.11: Position of the 2'-OH groups in ANA (○) *versus* RNA (○) in hybrid duplexes. For clarity hydrogens are not shown. The green strands (●) represent the target RNA. Note the disposition of the 2'oxygens (●) in ANA:RNA (major groove) and RNA:RNA (minor groove).

This is a probable mechanism by which RNase H differentiates between the ANA:RNA and RNA:RNA duplexes. It is also possible that the ANA strand in the ANA:RNA hybrid remains flexible after enzyme binding, and that this “flexibility condition” is important for RNA cleavage.

At this point it is not entirely clear whether the reduced cleavage of RNA in the ANA:RNA versus DNA:RNA hybrids is a result of weaker binding of ANA to the RNA (T_m 44°C) relative to DNA (T_m 72°C) or, due to imperfect binding of the enzyme to ANA:RNA hybrid relative to the natural DNA:RNA heteroduplex. The fact that significantly more cleavage of the target RNA is observed upon decreasing the temperature of the RNase H assay from 37°C to 25°C, suggests that ANA:RNA hybrids bind to the enzyme rather well.³³⁸ Crooke and coworkers clearly show that the conformation of a (DNA:RNA) hybrid duplex is altered by enzyme binding. Following binding of the enzyme by CD spectroscopy, the maximum is blue-shifted from 263 nm to 258 nm and a new minimum at 288 nm appears.³²⁷ This involves bending of the duplex, an alteration in the base stacking, and an overall global helical conformation change of the hybrid from the pure A-form. Although the ANA hybrid appears to be more like the DNA:RNA heteroduplex (A-like, **Figure 3.1.8 B**) the alteration after binding to the enzyme has not yet been explored relative to the natural substrate.

For ANA to be used as an antisense agent, its binding to RNA needs to be improved. On a microscopic level this translates to tuning the pucker of the ara sugar to a more DNA like C2'-endo conformation, and/or perhaps reducing the size of the 2' group. The 2'O → 2'F substitution may satisfy both conditions.²⁹⁸ This work is in progress in the Damha lab, where the 2'β moiety is being replaced by a fluorine atom in an effort to increase the percentage of S over N conformers. Indeed, C. Wilds of Damha's group has recently demonstrated that 2'F-ANA:RNA hybrids are of higher thermal stability than ANA:RNA and natural DNA:RNA hybrids, and are also substrates for *E. coli*. RNase H.³⁰¹

4.3.4 Conclusions And Biological Implications

Overall the RNA cleavage efficiency effected by RNase H and an antisense ONT probe decreases as follows : unmodified DNA > Thio-DNA > ANA> Thio-ANA (data not shown) >> RNA. Differences between the binding affinity of the ONT to RNA cleavage efficiency and cleavage pattern observed suggest that the binding interaction as well as the catalytic process are both sensitive to these types of structural modifications. The weak but definite binding affinity of ANA towards target RNA may under certain circumstances be advantageous, *e.g.* increase in the turnover rate of the enzyme may be observed as ANA dissociate from cleaved target and reassociate to a new target (**Figure 4.9**). This is particularly desirable if the enzyme concentration is limiting with respect to the substrate concentration and for improving the rate at which the enzyme scans A-form duplexes for the intended heteroduplex substrate. If a biological termination event is suspected to be the result of an RNase H mechanism, then assuming it is rate-limiting, then some compromise in binding affinity of the antisense ONT is inevitable, *i.e.* “more” (high affinity) may not be necessarily better.

CHAPTER V: CONTRIBUTIONS TO KNOWLEDGE

SYNTHESIS OF ARABINOGUANOSINE

A fully protected arabinoguanosine (araG) monomer for solid-phase synthesis was prepared from guanosine (rG) in moderate yields. Inversion of configuration at C(2') was achieved by introduction of the (trifluoromethyl)sulphonyl group and subsequent displacement by the acetate nucleophile. The guanine moiety was protected at the amide (O⁶) function by the 2-(4-nitro-phenyl)ethyl (npe) group, and at the amino group by the isobutyryl protecting group.

SYNTHESIS AND PROPERTIES OF OLIGONUCLEOTIDES BASED ON D-ARABINOSE

Physicochemical Generalities

Oligodeoxynucleotides containing one or two arabinonucleosides as well as completely modified oligomers (ANA) both with phosphodiester and phosphorothioate backbones have been synthesized. These sequences were 18 nucleotides in length, and were designed to bind to the 5'-long terminal region (LTR) of HIV-1 genomic RNA (strain 3B). The DNA-ANA chimera displayed similar binding properties towards both single stranded DNA and RNA, relative to the all-DNA oligomer. However ANA (*i.e.*, containing only D-arabinose sugars) exhibited reduced affinities towards both ssDNA and RNA as compared to an unmodified DNA oligomer. Phosphorothioate (PS) ANA displayed lower binding affinity relative the unmodified phosphodiester (PO)-ANA oligomer. Furthermore, it has been demonstrated that ANA oligomers of mixed base composition (U, C, A, G) have preferential binding to complementary RNA rather than ssDNA, a phenomenon generally not seen for unmodified DNA and RNA oligomers. Although this property may not be critical for antisense applications (where the target is RNA), it may offer new insights into the factors that influences nucleic acid duplex stability.

Triplex Helix Formation

Structural implications of the antigene technology: The ability of an oligoarabinopyrimidylate (ANA third strand) to associate and form a triplex with duplex DNA and hybrid DNA:RNA was assessed for the very first time. ANA, like DNA & 2',5'-RNA, was found to bind to duplex DNA and hybrid DNA(Pu): RNA(Py), but not to duplex RNA or hybrid DNA (Pu): RNA (Py). This is in contrast to the 2'-epimeric RNA, which binds to all four types of helices (DD & DR, and RR & RD). These results support the notion that the presence of ribose 2'-OH groups in the third strand contributes significantly to the stability of triple helical complexes (particularly R•R:R and R•RD). Such understanding can also be applied to the design of sequence selective oligonucleotides which interact with double-stranded nucleic acids, and emphasizes the role of the 2'-OH group as a general recognition and binding determinant of RNA.

Nucleic acids containing an araA branch-point: Hybridization, stoichiometric measurements and CD spectroscopy suggested that branched oligopyrimidines with an arabinoadenosine branch-point are capable of forming triple-helical complexes in the presence of complementary deoxyadenylate strands. Such complexes represent the lesser known "anti-parallel", or T*A:T Reverse-Hoogsteen/W-C motif.

In a buffer that mimics intracellular cationic conditions, araA(dT)₁₀dT₁₀ binds to dA₁₀ to form a 1:2 W-C complex, whereas its epimeric riboA(dT)₁₀dT₁₀ forms a 1:1 W-C complex under the same conditions. This interesting observation may put into perspective some recent reports that have particularly emphasized the use of nucleic acids as building elements of 'nanostructures'. Another application of oligonucleotides containing D-arabinose would be to synthesize branched substrates that would be useful for investigation of fundamentally important biological processes, such as pre-mRNA splicing. These research endeavours are in progress in the Damha research laboratory.

Priming and Inhibition of HIV-1 RT Translocation

In a collaborative effort with the McGill AIDS Centre, DNA-ANA chimeras (*e.g.*, DNA containing araC at the 3' terminus) were found to disrupt reverse transcription by a

physical block mechanism. The stereochemical flexibility of the critical hydroxyls present in the core sequence (aU₂, aA₃, aG₄), as well as others such as aA₅, aC₈, aC₁₆ have been examined by simply altering the absolute configuration of the corresponding C2'- carbons within the priming-substrate complex. Alteration of the configuration at these sites was achieved by simple replacement of the desired deoxy- residues with ara residues. Seven DNA:RNA complexes were prepared in which dC₁, dT₂, dA₃, dG₄, dA₅, dC₈, and dC₁₆ were replaced by the ara residues in the DNA priming strand, and the activity of these complexes has been compared with that of the corresponding native complex. Two parameters that affect the ability of HIV RT polymerases to synthesize past ara monomers *in vitro*, have been identified. These are the identity of the sugar in the backbone (ara vs. ribo or deoxy), and base identity (araC, araA, araG, araU). All of the results show clearly that the araX substitution in the primer constitutes a significant block to DNA synthesis for HIV-1 RT polymerase. They demonstrate that an anomaly in the sugar moiety of the T/P backbone, which does not disrupt the thermodynamic stability of base pairing, or appear to introduce a bulky side group into the DNA, can still induce partial synthetic arrest. These results suggest a complex mechanism for the priming of the antisense strand by the HIV-1 RT polymerase. While the 3'-hydroxyls are necessary for efficient priming by the RT complex, little is known about the optimal positioning of these functional groups within the active catalytic complex.

Induction of RNase H Activity by Arabinonucleic Acids

One of the most important contributions of this thesis work, was the finding that ANA:RNA hybrids serve as substrates for RNase H (*E. coli* and HIV-RT). To the best of our knowledge, there are no previous examples of RNase H induction resulting from hybridization of RNA with a *uniformly sugar-modified* oligonucleotide. The ability of RNase H to degrade RNA in ANA:RNA hybrids (2'-F or 2'-OH) may result from (a) the similarity of structure of these hybrids to that of the normal DNA:RNA substrate (A-like form), and (b) the fact that the 2'-OH groups of arabinose projects into the major groove of the helix, at a site where it should not interfere with RNase H's binding and catalytic processes. Work continued by C.J. Wilds of our laboratory, and by members of the

McGill AIDS Center, have shown that ANA and 2'F-ANA exhibit enhanced nuclease-resistance to serum and cellular nucleases compared to DNA strands, although less than to phosphorothioate PS-DNA. However, unlike PS-DNA, arabinonucleic acids show little nonspecific binding to cellular proteins a property that may result in a significantly improved interaction with cellular RNA *in vivo*. These properties combined establish that arabinonucleic acids may serve as excellent models of antisense agents, and as valuable tools for studying and controlling gene expression in cells and organisms.

PUBLICATIONS/INVENTION DISCLOSURES

Noronha, A. and Damha, M.J. “Hybridization Properties of Arabinonucleic Acids (ANA) Influence of Stereochemistry at 2’ on the Stability of Double and Triple Helices”. *Journal of Biomolecular Structure and Dynamics*, **1997**, 14, 805-806 (abstract format).

Borkow, G.; Arion, D.; Noronha, A.; Scartozzi, M.; Damha, M.J. and Parniak, M.A. “Probing Two Important Steps in the Replication of HIV-1 with Antisense Oligonucleotides: Priming and Template-Switching Reactions”. *The International Journal of Biochemistry and Cell Biology*, **1998**, 29, 1285-1295.

Wasner, M.; Arion, D.; Borkow, G.; Noronha, A.; Uddin, A.H.; Parniak, M.A. and Damha, M.J.; “Physicochemical and Biochemical Properties of 2’,5’-Linked RNA and 2’,5’-RNA:3’,5’-RNA “Hybrid” Duplexes”. *Biochemistry*, **1998**, 37, 7478-7486.

Noronha A. and Damha, M.J. “Triple Helices containing arabinonucleotides in the third (Hoogsteen) Strand: Effects of inverted stereochemistry at the 2’ position of the sugar moiety”. *Nucleic Acids Res.*, **1998**, 26, 2665-2671.

Damha, M.J. and Noronha, A. “Recognition of nucleic acids double helices by homopyrimidine 2’,5’-linked RNA”. *Nucleic Acids Res.*, **1998**, 26, 5152-5156.

Damha, M.J. ; Wilds, C.J.; Noronha, A.; Brukner, I.; Borkow, G.; Arion, D. and Parniak, M.A. “Hybrids of RNA and Arabinonucleic Acids (ANA and 2’F-ANA) are Substrates of Ribonuclease H.” *J. Am. Chem. Soc.*, **1998**, 120, 12976-12977.

"Antisense oligonucleotide constructs based on β -arabinofuranose and its analogues", Damha, M.J.; Parniak, M.A.; Noronha, A.; Wilds, C.J.; Borkow, G. and Arion, D. Canadian Provisional Patent #2,241,361 filed June 19, 1998; Patent Cooperation Treaty (PCT) filed June 16, 1999.

CONFERENCE PRESENTATIONS

Wilds, C.J.; Lok, C.-N.; Noronha, A.; Viazovkina, E.V.; Parniak, M.A.; Gehring, K.; Pon, R.T. and Damha, M.J. Synthesis and biochemical properties of arabinonucleic acids (ANA). 218th American Chemical Society National Meeting New Orleans, Louisiana, Aug. 22nd- 26th, **1999**.

Noronha, A.; Wilds, C.J.; Damha, M.J.; Arion, A. and Parniak M.A. "Oligonucleotide Analogues which Inhibit HIV-1 Reverse Transcriptase", 81st Canadian Society for Chemistry Conference and Exhibition, Whistler, British Columbia, May 31- June 4, **1998**.

Wilds, C.J.; Noronha, A. and Damha, M.J.; "Formation and Stability of Complexes with Arabino-Fluoro Nucleic Acids: Triplexes and Tetraplexes", 81st Canadian Society for Chemistry Conference and Exhibition, Whistler, British Columbia, May 31- June 4, **1998**.

Liscio, A.; Braich, R. S.; Chen, J.; Noronha, A.; Hudson, R.H.E. and Damha, M.J. "Yeast RNA Lariat Debranching Enzyme: Base Sequence and Substrate Specificity." 81st Canadian Society for Chemistry Conference and Exhibition, Whistler, British Columbia, May 31- June 4, **1998**.

Noronha, A and Damha, M.J. "Role of 2' Stereochemistry in Directing Nucleic Acid Complex Stability and Formation: Duplexes and Triplexes ", 4th Workshop of the Montreal Joint Center for Structural Biology (MJCSB), Bio-Research Institute (NRC), Montreal, Quebec, May 22, **1998**.

Wilds, C.J.; Noronha, A. and Damha, M.J. "Structural Studies of Oligoarabinonucleotides", 1st Concordia University Chemistry and Biochemistry Graduate Research Conference, Montreal, Quebec, February 6, **1998**.

Noronha, A.; Wilds, C.J. and Damha, M.J. "The Synthesis and Biophysical Studies of 2'-Modified Oligoarabinonucleotides". 8th Annual Quebec/Ontario Minisymposium in Bio-organic and Synthetic Organic Chemistry, Quebec. Nov 7-9, **1997**.

Wasner M.; Borkow, G.; Uddin, A.; Noronha, A. Parniak, M.A. and Damha, M.J. "Selective Association and Antisense Properties of 2'5' Linked Oligoribonucleotides." 4th Cambridge Symposium, Oligonucleotide Chemistry and Biology, Cambridge, UK. August 30-September 4, **1997**.

Noronha, A.; Borkow, G.; Arion, D.; Scartozzi, M.; Damha, M.J. and Parniak, M.A. "Probing Important Steps in the Replication of HIV-1 with Antisense Oligonucleotides. 7th Annual Quebec/Ontario Minisymposium in Bio-organic and Synthetic Organic Chemistry, Waterloo, Ontario. October 26-27, **1996**.

Noronha, A.; Arion, D.; Borkow, G.; Uddin, A.; Scartozzi, M.; Parniak, M.A. and Damha, M.J. "Probing Two Important Steps in the Replication of HIV-1 with Antisense Oligonucleotides: Priming and Template-Switching Reactions"; 79th Canadian Society for Chemistry, Conference and Exhibition, Saint Johns, Newfoundland, June 2-5, **1996**.

Damha, M.J.; Noronha, A.; Scartozzi, M.; Uddin, A; Parniak, M.A.; Arion, D.; and Borkow, G. "*In Vitro* Inhibition of HIV-1 Reverse Transcription by Antisense Oligonucleotides" 78th Canadian Society for Chemistry, Conference and Exhibition, Guelph. Ontario, June 4-7, **1995**.

CHAPTER VI EXPERIMENTAL

6.1 GENERAL METHODS

6.1.1 General Reagents

In general, solvents were dried and fractionally distilled, under reduced pressure for high boiling point solvents. Dichloromethane (DCM) was dried over CaCl_2 and then refluxed over CaH_2 prior to use. A simple distillation was carried out for hexanes (BDH, British Drug Houses, Toronto, ON) used in column chromatography. Pyridine (Caledon Laboratories Ltd., Georgetown, ON), collidine and N,N-dimethylformamide (DMF) (BDH) were dried by refluxing over calcium hydride (CaH_2) and stored over 4Å molecular sieves. N,N-diisopropylethylamine (Aldrich Chemical Company, Milwaukee, WI) was stirred over CaH_2 (BDH) and distilled before use. Chloroform, methanol (BDH), hydrochloric acid, diethyl ether (Caledon) and triethylamine tris(hydrofluoride) or TREAT HF (Aldrich) were used as obtained.

The following reagents were all used as received: 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (Aldrich). Analytical reagent grade glacial acetic acid, ammonium acetate, anhydrous sodium sulphate, magnesium chloride, sodium chloride, 95% ethanol, acetone, triethylamine, analytical grade ethylenediamine tetra acetate (EDTA), magnesium chloride were all obtained from BDH, disodium hydrogen phosphate (Fisher Scientific Fairlawn, NJ), manganese chloride, sodium acetate (Aldrich), TRIS (Bio-Rad).

Hoechst 33258 and Ethidium Bromide (2,7-diamino-9-phenylphenanthridinium-10-ethyl bromide) were purchased from Aldrich Chemical Company. Benzo[e]pyridoles (BePI) was obtained as a gift from Dr. Claude Hélène (Paris, France) and is now available from Aldrich.

6.1.2 Chromatography

Column chromatography was performed with silica gel [40-63 micron Silica gel 60 (EM science Gibbstown, NJ)].

Thin-layer chromatography (TLC) was achieved using Merck Kieselgel 60 F-254 plastic-back analytical silica gel sheets (0.2 mm thickness, EM Science, Gibbstown, NJ). Compounds were visualized on TLC plate by illumination with a UV light source (Mineralite, emission wavelength *ca.* 254 nm), and further stained by exposure to HCl or trifluoroacetic acid vapours for trityl-containing compounds. Both UV and non UV active compounds could also be visualized with Mohr's solution (2.5 g of ammonium molybdate and 1 g of ceric sulphate in 10% sulphuric acid (w/v). This was achieved by first dipping the plate in the solution followed by heating to obtain purple coloured spots/regions.

6.1.3 Instruments

NMR Spectroscopy All spectra were obtained at ambient temperature, on a Varian XL-500 spectrophotometer, and the chemical shifts are reported in ppm downfield from tetramethylsilane (TMS). For the new compounds synthesized, all ^{13}C and ^1H assignments were made using 2-D NMR experiments, including homonuclear correlated spectra (COSY), and ^1H -detected heteronuclear multiple quantum coherence spectra (HMQC). Deuterated solvents: acetone- D_6 (CDN Isotopes Quebec, Canada), dimethylsulfoxide- D_6 (Cambridge Isotope Laboratories Andover, MA), chloroform- CDCl_3 (Isotec Inc. Miamisburg, OH).

UV Spectroscopy. UV-VIS spectra were recorded using a Varian Cary 1 UV-VIS spectrophotometer (Varian:Mulgrave, Victoria, Australia). Thermal denaturation of oligonucleotides was followed in the ultraviolet spectrum by Varian CARY I spectrophotometer equipped with a thermostated cell holder with data being collected using manufacturers supplied software (version:CARY 1.3e) and a personal computer. (see section 6.4 for details)

CD Spectroscopy. CD spectra were collected on a Jasco J-710 spectropolarimeter equipped with a thermoelectrically controlled external constant temperature NESLAB RTE-111 circulating bath. The data were processed on a PC computer using Windows based software supplied by the manufacturer (JASCO, Inc.).

FAB-Mass Spectrometry. Fast atom bombardment mass spectra were collected using a Kratos MS25RFA high resolution mass spectrometer. Nitrobenzyl alcohol (NBA) matrix was used.

MALDI TOF Mass Spectrometry. Matrix-assisted laser desorption/ionization time of flight mass spectra (MALDI-TOF) were recorded on a Kratos Kompact-III TOF instrument with a minimum laser output of 6 mW at a wavelength of 337 nm light, 3 ns pulse width, 100 mm diameter spot. The MALDI instrument was operated in a positive (reflectron and linear) mode.

6.2 OLIGONUCLEOTIDE SYNTHESIS

6.2.1 Reagents for Derivatization of Nucleosides

All nucleosides (ribo and deoxy) were obtained from Dalton Chemical Laboratories (DCL, Toronto, ON) while arabinonucleosides of adenine, uracil and cytosine were from Sigma (St. Louis). N,N-diisopropyl-2-cyanoethylphosphonamidic chloride was purchased from Dalton Chemical Laboratories; chlorotrimethylsilane (TMS-Cl), p-anisylchlorodiphenylmethane (MMTrCl), 4-dimethylaminopyridine (4-DMAP), triphenyl phosphine, diethyl azodicarboxylate, 2-(4-nitrophenyl)-ethanol, trifluoromethanesulphonic anhydride, hexamethylphosphoramide, isobutyryl chloride, 1M tetrabutylammonium fluoride (TBAF), and benzoyl chloride were obtained from Aldrich. Lithium acetate (LiOAc) (ACROS Organics, Nepean, ON), 1,3-dichloro-1,1',3,3'-tetraisopropylidisiloxane from United Chemical Technologies, Inc. (Bristol PA).

Reagent grade acetic anhydride (Ac_2O), trichloroacetic acid (TCA), aqueous ammonia (BDH), 1,2-dichloroethane (DCE) (Caledon), N-methylimidazole (N-MeIm), and iodine (Aldrich) and DNA synthesis grade tetrazole (DCL) were used as received.

Double distilled water (Millipore, Mississauga ON), was treated with diethyl pyrocarbonate (Aldrich) to form a 0.1% solution and autoclaved (1 h., 121°C, 1.3 atm).

6.2.2 Derivatization of Solid Support

Long-chain alkylamine controlled-pore glass (LCAA-CPG, 500°A pore size, density: *ca* 0.4 g/mL; DCL) was derivatized according to literature procedure^{339,340} with the additional prolonged acid activation step. Activation of commercial LCAA-CPG, by shaking over 5% TCA in DCE (w/v) for 24-72 h. was found to be necessary to achieve the desired nucleoside loading in the subsequent derivatization steps.³⁴¹ Nucleoside loadings were determined by spectrophotometric mono- and di- methoxytrityl cation assay. The support was dried *in vacuo* 24 h. before use, loaded into an empty column with replaceable filters (ABI), crimped closed with aluminum seals (Pierce) and installed on the instrument.

6.2.3 Automated Oligonucleotide Synthesis

Monomers for Automated Synthesis: 5'-O-dimethoxytrityl-3'-O-(2-cyanoethyl)N,N-diisopropylphosphoramidite derivatives of various 2'-deoxy and ribo nucleosides were purchased from DCL. N⁶-benzoyl-5'-O-monomethoxytritylarabinoadenosine-2',3'-bis((cyanoethyl)N,N-diisopropylphosphoramidite)) was prepared according to the method of Damha *et al.*^{175,342} These phosphoramidite reagents were stored at -20°C and desiccated under vacuum (over P₂O₅) for 24 hr. prior to use. The 5' amidites were synthesized according to the previously published procedures.³⁴¹

a. Chain Assembly

Reagents for the solid phase synthesizer of oligonucleotides were purchased from Applied Biosystems Inc. (ABI, Foster City, CA), DCL or prepared as described below.

Syntheses were carried out on an Applied Biosystems DNA/RNA 381A synthesizer. Anhydrous acetonitrile (Caledon) was pre-dried by distillation over P₂O₅ (BDH) and then refluxed over calcium hydride under dry nitrogen prior to use. Tetrahydrofuran (BDH) was refluxed over CaH₂, filtered and then distilled over sodium (Aldrich)/benzophenone (Aldrich). Oligomers were prepared on a 1 µmol scale using LCAA- controlled pore glass (500 Å) derivatized with 3'-nucleoside described above. Assembly of sequences was carried out using the following reagents and synthesis steps:

(1) detritylation: 3% trichloroacetic acid in dichloroethane delivered in 100 s (+ 40 s 'burst') steps. To determine condensation yields, the eluate from this step was collected and the absorbance measured by UV spectroscopic quantitation of trityl for DMT⁺ ($\lambda = 504$ nm, $\epsilon = 76000 \text{ L.mol}^{-1}.\text{cm}^{-1}$, deoxy and ribo) and MMT⁺ ($\lambda = 478$ nm, $\epsilon = 56000 \text{ L.mol}^{-1}.\text{cm}^{-1}$, ara); (2) phosphoramidite coupling time of 90s (DNA) and 7.5 min. (RNA and ANA) and 30 min. for the branch-point nucleosides; (3) Capping: acetic anhydride/2,4,6-collidine/THF 1:1:8 (volume ratio, solution A) and 1-methyl-1*H*-imidazole/THF 16:84 (volume ratio, solution B) delivered in 15 s + 600 s "wait" steps; (4) Oxidation: 0.2M iodine in THF/ pyridine/water 25:20:2, delivered in 20 s + 20 s "wait" steps.

Prior to oligonucleotide assembly, the derivatized support was capped to block undesired reactive sites^{179,340} with the capping reagents described above using the capping cycle provided by ABI. Phosphoramidite reagents were dissolved in dry, freshly distilled acetonitrile retrieved from the collection bulb of a continuous reflux apparatus *via* a syringe fitted with a stainless steel needle that was introduced through a septa-sealed stopcock. The final concentrations of the amidites were 0.1 M (deoxy monomers) and 0.15 M ribo monomers. The concentrations of the arabino amidites were as follows: 0.11M for araG and 0.15M for araA, araC & araU; while that of the branch-point N⁶-Benzoyl-5'-O-monomethoxytrityl-arabinoadenosine-2',3'-O-bis(2-cyanoethyl)N,N-diisopropyl phosphoramidite was at approximately 0.015M. All custom made amidite solutions were filtered through a 0.45 μm pore Teflon[®] filter by use of a 'swinny' filter apparatus (Millipore, Mississauga, ON) before placement on the automated synthesizer. The range of couplings efficiencies as monitored by the trityl assay was found to be 65- 109%.

Thio-ANA was synthesized on a 1 μmole scale using Beaucage's reagent^{20,343} (0.05M in acetonitrile) as the sulfurizing reagent, with a delivery time of 12 s and wait step of 120 s. The coupling times were kept the same as for the ANA oligomers, while the couplings efficiencies were much lower for the Thio-ANA (55-113%). The thio DNA oligomers were obtained from the University of Calgary DNA Synthesis Laboratory (Calgary, Alberta).

b. Deprotection of Oligomers

Following chain assembly, the CPG linked oligomer was taken out of the column and divided into two portions (0.5 μmol each). Each portion was treated with aqueous ammonia/ethanol (3:1 v/v, 1.2 mL) for 2-3 days at rt to cleave the oligomers from the solid support as well as to remove the exocyclic amino and cyanoethyl blocking groups. Under these conditions; dT₁₀ was deprotected for 8 h; oligomers containing rG^{iBu} and dG^{iBu} units for 48 h, and oligomers containing mixed bases but lacking guanine for 24 h. After centrifugation, the supernatant was collected and the support washed with EtOH. The supernatant and EtOH washings were evaporated. In the case of RNA oligomers the resulting pellet was treated with neat TEA:3HF at rt for 48 h [(triethylamine tris(hydrogenfluoride) or TREATHF)]³⁴⁴ to remove the silyl protecting groups.³⁴⁵ The amount of TREATHF was 100 μL per 1 μmol of oligomer. In the case of the oligomers containing araG^{iBu,npe} units, the crude oligomer was treated with TBAF (1 M in THF; 15 μL per npe group). Caution: TREATHF does not effect the removal of the npe groups. The oligomeric solutions were lyophilized (Savant Industries Speed-Vac[®]) and purified by PAGE as described below.

6.3 PURIFICATION OF OLIGONUCLEOTIDES

6.3.1 Polyacrylamide Gel Electrophoresis (PAGE)

Crude oligomers were purified by vertical slab polyacrylamide gel electrophoresis (PAGE) using a Bio-Rad Proteam II or Hoefer Scientific units (San Francisco, CA). Electrophoresis grade acrylamide, N,N'-methylene-bis(acrylamide) (BIS), ammonium persulphate (APS), N,N,N',N'-tetramethylethylenediamine (TEMED), bromophenol blue (BPB) and xylene cyanol (XC) were from Bio Rad. Other electrophoresis reagents were obtained as follows: boric acid, formamide and disodium ethylenediaminetetraacetate dihydrate (EDTA, BDH), Trishydroxy-methylaminomethane (Tris), sucrose (Aldrich) and urea (Caledon).

The thickness of the gels were 0.75 mm and 1.5 mm for analytical and preparative gels respectively. Most commonly gels were composed of 14-24% (w/v) acrylamide and

employed TBE buffer (89 mM Tris/boric acid, 2.5 mM EDTA, pH = 8.3)³⁴⁶. Deionized formamide was prepared for use as loading buffer by stirring over a mixed bed ion-exchange resin (Bio Rad AG 501-X8), 8:2 formamide/10 x TBE was used as a denaturing loading buffer. All oligonucleotides were purified by denaturing PAGE. The amount of material was quantitated by UV absorbance at $\lambda = 260$ nm and then lyophilized to a pellet. About 30 ODs were typically loaded onto preparative gels in the loading buffer. The desired bands were excised and shaken overnight in 5 mL water. Subsequently, the oligomers were water-extracted from crushed gel pieces and desalted as discussed below.

6.3.2 Gel Mobility Shift Assays (Native gels)

As the name suggests these gel systems contain no detergent or denaturing agent and are designed to allow the nucleic acids under study to retain their complex activity. It is well known that the helical periodicities for double-stranded RNA (11.3 ± 0.1 bp/turn) and for RNA:DNA hybrid in solution (10.9 ± 0.1 bp/turn) can be detected with gel mobility shift assays and are reported to be the same as estimated from the crystal. The value estimated for RNA:DNA hybrid obtained *via* this shift assay technique is intermediate between dsRNA and the known periodicity of dsDNA and was found to be 10.6 base pairs/turn. (Chapter 3, Section 3).³⁴⁷ For the native gels, the solutions of oligonucleotides were lyophilized, incubated in 10 μ L of 30% sucrose in 10 x NAE (100 mM Sodium Acetate, 1mM EDTA, pH 5.5) at 80°C for 15 min., cooled to room temperature, and finally incubated at 4°C for 1-7 days. The samples were loaded in 30% sucrose onto the gel. The running buffer contained 1x NAE (pH=5.00). These were run at 100V for 6-8 h at 4°C.

6.3.3 Visualization of Oligonucleotides

After removal of the PAGE glass plates, the gels were covered with Saran Wrap[®] and photographed over a fluorescent TLC plate illuminated by a hand-held UV lamp using Polaroid PolaPan[®] (4x5" Instant Sheet Film, #52, medium contrast, ISO 400/21°C; f 4.5, 16 second) and a Kodak Wratten gelatin filter (#58).

Gels were also better visualized by use of All-stain (Bio-Rad)[®],³⁴⁸ a solution containing 1-ethyl-2-(3-[1-ethylnaphtol(1,2-d)-thiazolin-2-ylidene]-2-methylpropenyl)naptho(1,2-d)-thiazolium bromide. The filter was removed and the same film (f 11, 0.30s) used, by placing the gel on a white background.

6.3.4 Desalting of Oligonucleotides

The desired oligonucleotides were separated from low molecular weight impurities and salts by size-exclusion chromatography (SEC) using Sephadex G-25 fine (Pharmacia, Baie d'Urfe, QC). SEC matrix was generously washed with sterile water to remove fines, autoclaved (1 h, 120° C, 1.3 atm) using an All American Electric Steam Sterilizer - Model No. 25X (Wisconsin Aluminium Co., Inc. Manitowoc, WI) and then allowed to hydrate overnight. SEC was run in appropriately-sized, sterile, disposable 10 mL syringe barrels (Becton Dickinson & Co., Franklin Lakes, NJ) plugged with silanized glass wool (Chromatographic Specialities Inc., Brockville, ON). Doubly deionized distilled, autoclaved water was used to hydrate SEC media as well as to pack the column and elute the oligonucleotides. The size of the column was determined by the amount of nucleotide material loaded. In general the volume of the column was ten times the volume of the sample. Fractions (0.5 to 1.5 mL) were collected and the amount of oligonucleotide in each was quantitated by UV absorbance spectrophotometry (260 nm). The oligonucleotide containing tubes were then pooled to give a stock which was stored at -20° C. SEC media was also used for linear and branched oligonucleotides that were ten bases or longer. The recovery of material from the column was generally very good (70 - 90 %).

The overall resulting yields in A₂₆₀ unit is reported in **Table 6.1**. The crude yield is given as total number of optical density units (or "OD units"). An OD unit represents the number of absorbance units in 1 mL of water as quantitated at 260 nm. Analysis was achieved by PAGE, UV spectroscopy and MALDI. Commercial DNA oligomers were used as received.

6.3.5 Summary of sequences prepared and purified (Table 6.1)

Compound	Description	crude yield (OD units ^f)	Pure (OD)	yield	yield OD (%)
Chapter 3.1					
3.1	D18	DCL	-	-	-
3.2	aU ₂	51	40	-	78
3.3	aA ₃	50	43	-	87
3.4	aG ₄	40	4.3	-	10.8
3.5	aA ₅	76	35	-	46
3.6	aC ₁	68	23	-	34
3.7	aC(1&8)	109	8.4	-	7.7
3.8	aC16	219	36	-	17
3.9	D17R1	DCL	-	-	-
3.10	D13R5	DCL	-	-	-
3.11	D9R9	DCL	-	-	-
3.12	R9D9	DCL	-	-	-
3.13	R17D1	43	3.1	-	15
3.14	R18	75	52	-	69
3.15	A18	42	4.8	-	12
3.16	RAN	DCL	-	-	-
3.17	Thio-ANA	22	1.7	-	8
3.18	Thio-DNA	RTP	-	-	-
Chapter 3.2					
3.19	aA ^{T10} _{T10}	57	13	-	23
3.20	rA ^{T10} _{T10}	93	14	-	15
3.21	linear AT ₁₀	DCL	-	-	-
3.22a	3'T ₁₀ C ₄ T ₁₀ 3'	106	18	-	17
3.22b	5'T ₁₀ C ₄ T ₁₀ 3'	DCL	-	-	-
Chapter 3.3					
3.23	DD	RTP	-	-	-
3.24	DR	83	75	-	90
3.25	RD	25	15	-	60
3.26	RR	38	3.4	-	9
3.27	D	RTP	-	-	-
3.28	D'	RTP	-	-	-
3.29	R	147	16	-	10.9
3.30	R'	27	17	-	62
3.31	A	64	35	-	54
Chapter 3.4					
3.32	DNA	RTP	-	-	-
3.33	RNA	167	11	-	7
3.34	ANA	40 (1/2)	-	-	17

Scale of solid-supported synthesis was 1 μ mol in all cases. ^f The crude yield is total number of optical density units. RTP: Oligomers obtained from R.T. Pon, University of Calgary; DCL Oligomers obtained from Dalton Chemical Laboratories (Toronto). These were used as received.

6.4 CHARACTERIZATION OF OLIGONUCLEOTIDES

6.4.1 Bio-Physical Characterization: Hybridization Properties

a. UV-Thermal Melt Studies

The hybridization characteristics of oligonucleotides were investigated by following the change in UV- absorbance with temperature, by generating UV denaturing profiles (“melting curves”). These studies were conducted using a Varian Cary I UV-VIS spectrophotometer equipped with a Peltier temperature controller. Spectra were collected while the samples were held in Hellman QS-1.000-104 cells. In general, the temperature was ramped at 0.5°C intervals at a rate of 0.5°C/min., while the absorbance was recorded at 260nm. The spectrophotometer was set on dual optical mode to reduce optical drift. The absorbance vs. temperature data was converted from a binary file format to an ASCII matrix using the reports “table function” (software x-y pairs, format Cary1/3). This form of file could be imported into Microsoft Excel and further manipulated for presentations. In order to compare relative overall changes in absorbance, normalized ΔA plots were constructed according to the method of Zon and Wilson,²²⁷ by use of the formula: $(A_t - A_i)/A_f$ where A_i is the initial absorbance, A_f is the final absorbance. This facilitates comparison on a more uniform basis, for parallel runs, especially at identical concentrations and in the same buffer.⁹⁴ Hyperchromicity values (H%) are reported as the percent increase in absorbance at the wavelength of interest with respect to the final absorbance and the T_m s were calculated using the base-line method in accordance with that of Puglisi and Tinoco.⁹⁴ T_m values were determined in duplicate, and have generally an uncertainty of $\pm 0.5^\circ\text{C}$.

Molar extinction coefficients for DNA and RNA strands were estimated and calculated from those of the mononucleotides and dinucleotides according to nearest-neighbouring approximations.⁹⁴ Molar extinction coefficients for ANA were assumed to be the same as those of normal RNA strands, and reported in Table 6.2 (in $10^4 \text{ M}^{-1} \text{ cm}^{-1}$ units). The extinction coefficient of the DR and RD hybrid molecules were assumed to be similar to the sum of their D plus R components. Extinction coefficients of ANA-DNA chimeric sequences were assumed to be the same as those of normal DNA sequences.

Table 6.2: Designation and Molar Extinction Coefficients of Various Sequences

Compound	Designation	ϵ ($10^4 \text{ M}^{-1} \text{ cm}^{-1}$ units)	Compound	Designation	ϵ ($10^4 \text{ M}^{-1} \text{ cm}^{-1}$ units)
Chapter 3.1^ζ			Chapter 3.2^τ		
3.1	D18	17	3.19	aA ^{T10} _{T10}	18.4
3.2	aU2	17	3.20	rA ^{T10} _{T10}	18.4
3.3	aA3	17	3.21	linear AT ₁₀	9.9
3.4	aG4	17	3.22a	3'T ₁₀ C ₄ T ₁₀ 3'	19.9
3.5	aA5	17	3.22b	5'T ₁₀ C ₄ T ₁₀ 3'	19.9
3.6	aC1	17	Chapter 3.3^ψ		
3.7	aC(1&8)	17	3.23	DD	26.5
3.8	aC16	17	3.24	DR	27.1
3.9	D17R1	17	3.25	RD	26.7
3.10	D13R5	17	3.26	RR	27.7
3.11	D9R9	17.5	3.27	D	9.1
3.12	R9D9	17.5	3.28	D'	14
3.13	R17D1	17.5	3.29	R	9.6
3.14	R18	17.5	3.30	R'	14
3.15	A18	17.5	3.31	A	9.6
3.16	RAN	17	Chapter 3.4^χ		
3.17	Thio-ANA	17.5	3.32	DNA	9.1
3.18	Thio-DNA	17	3.33	RNA	9.6
			3.34	ANA	9.6

Concentrations were: ^ζ: 2.7 μM in each strand; ^τ: 1.0-1.6 μM in each strand; ^ψ: 1.8 μM for the hairpins and 3.5-5.0 μM for the single strands; ^χ: 5.0 μM in each strand; All buffers were adjusted with HCl or acetic acid. No correction was made to account for thermal expansion of buffer.

Samples for thermal denaturation analysis were prepared by lyophilizing an equimolar mixture of complementary strands to dryness and then re-dissolving in the appropriate buffer: (a) 140 mM KCl, 5 mM NaH₂PO₄ & 1 mM MgCl₂, pH 7.2; (b) 5 mM MgCl₂, 10 mM NaCl, 10 mM Tris-HCl, pH = 7.5; 5mM MnCl₂, 10 mM NaCl, 10 mM Tris-HCl, pH = 7.5. Samples were heated to 80-90°C for 15 min., then cooled slowly to room temperature, and stored at 4°C overnight before measurements. Prior to the thermal run, samples were degassed by placing them in a Savant lyophilizer (2 min.).

b. Circular Dichroism Spectroscopy (CD)

CD spectra were collected on a Jasco J-710 spectropolarimeter equipped with a constant temperature NESLAB RTE-111 circulating bath. Fused Quartz cells (Hellma 165-QS) with 1-cm optical path lengths were used. Oligonucleotide solutions for CD measurements were prepared with the appropriate buffers in a manner similar to that used for UV melting. Before data acquisition, samples were allowed to equilibrate for 5-10 min. at the appropriate temperatures. Each spectrum was an average of 5 scans and was collected at a rate of 100 nm/min. with a band width of 1 nm and sampling wavelength of 0.2 nm. The CD spectra were recorded from 350 to 200 nm at 5°C and normalized by subtraction of the background scan with buffer. The molar ellipticity was calculated from the equation $[\theta] = \theta/Cl$, where θ is the relative ellipticity (mdeg), C is the molar concentration of oligonucleotides (moles/L), and l is the path length of the cell (cm). To determine changes in CD profiles of triplex as a function of temperature in the absence and presence of the third strand, spectra of oligonucleotides were collected from 5° to 60°C, at 5 or 10 degree increments, with 5 min. equilibration times between measurements. The data were processed on a PC computer using Windows™ based software supplied by the manufacturer (JASCO, Inc.).

c. Stoichiometric Studies

The proportion in which (complementary) oligonucleotide strands associate can be determined by titrating a solution of one strand with an equimolar (*ca.* 1 μ M per strand) solution of the second strand as described by Job.²³⁴ Titrations were performed in quartz micro cuvettes (Hellma®, Concord, ON, QS-104-283, 10 mm path length). Absorbance measurements were collected on the spectrophotometer equipped with a Peltier block. Titrations with dA₁₀ (*ca.* 1 μ M) were performed at 10°C, and the initial reading was made after at least 20-30 min. to allow for equilibrium. Subsequent readings were made 15 min. after each addition of the complementary strand. Three wavelengths were used to monitor the complex formation (260, 275 and 284 nm). The hybridization buffers were: (a) 50 mM MgCl₂, 10 mM Tris-HCl, pH 7.3 adjusted with HCl. (b) 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄, pH 7.2.

d. Calculation of Thermodynamic Data

Thermodynamic parameters of complexation were extracted from the melt curves, *via* van't Hoff plots assuming a two state (all -or-none) model, *i.e.*, that the monophasic helix-coil transition of the **3.19**:dA₁₀ and **3.20b**:dA₁₀ triplexes corresponds to the equilibrium reaction: [branch oligomer: dA₁₀] \rightleftharpoons branch oligomer + dA₁₀. Thermodynamic data presented in **Figure 3.7** were calculated from plots of reciprocal melting temperature, $1/T_m$, versus the natural logarithm of total strand concentration, $\ln C_T$ (2-200 μ M concentration range), according to :

$$1/T_m = (R \ln C_T / \Delta H^\circ) + (\Delta S^\circ - R \ln 4) / \Delta H^\circ \quad (\text{eq. 1})$$

The slope of this plot can be used to calculate ΔH° and the intercept permits calculation of ΔS° . These two thermodynamic parameters can then be used to calculate ΔG° by application of the Gibb's free energy equation, $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$. R is gas constant and C_T : the total strand concentration [*i.e.* concentration of **3.19** and dA₁₀] present in the UV cell cuvette. The enthalpy is related to the slope of the curve ($-R/\Delta H^\circ$) and the entropy to the y-intercept ($\Delta S^\circ - R \ln 4 / \Delta H^\circ$). The free energy of association, or ΔG° , was calculated at 25°C. All plots were analyzed by linear regression.

6.4.2 Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectra (MALDI-TOF)

The gel purified oligomers **3.15**, **3.19**, **3.31** and **3.34** were analyzed by MALDI-TOF mass spectrometry using as matrix: 20 mM ammonium citrate (Fluka) in acetonitrile:water (1:1, v/v) buffer containing 6-aza-2-thiothymine (10 mg/mL, Aldrich).³⁴⁹ The positive ion MALDI-TOF mass spectra obtained in the reflectron (**3.19**, **3.34**) and linear mode (**3.15**, **3.32**), gave correct molecular mass for the desired oligomers, with excellent signal to noise ratio.

Samples were prepared by dissolving the oligomers in water at a concentration of 0.8 - 1 mM. A 5- μ L aliquot was pipetted into an eppendorf tube, to which was added 5 μ L

of the matrix. The final solution was shaken briefly, and 1-3 μL was applied to a stainless steel sample slide and air-dried prior to analysis.

6.5 BIOLOGICAL STUDIES

Biological studies described in this section were performed by Drs. D. Arion and G. Borkow of the McGill AIDS Centre. Interpretation of results and design of the experiments were carried out by the author (A. Noronha), Dr. M.J. Damha and members of the McGill AIDS Centre.

Recombinant heteropolymeric p51/p66HIV-1 RT was purified from lysates of *E.coli* JM-105, transformed with expression plasmids pRT66 and pRT51 using a rapid single-step purification method recently described.³⁰¹ Polynucleotide kinase T4 was purchased from Pharmacia Biotech. The “donor” RNA template used in reverse transcription reactions was prepared by *in vitro* transcription from plasmids pHIV-PBS, using T7 polymerase. Details and maps of this plasmid has been described previously.³⁵⁰ Ultrapure dNTPs, T7 RNA polymerase and *E.coli* RNase H were obtained from Pharmacia. [γ - ^{32}P]-ATP was purchased from Amersham. All 18 mer oligonucleotides (modified and unmodified) (5'-AGC TCC CAG GCT CAG ATC-3') were synthesized as described in Section 6.2.3, unless where indicated. All other materials were of the highest purity grade.

The 18mer RNA template corresponding to the R-region near the 5'-LTR of HIV-1 genome (5'-GAU CUG AGC CUG GGA GCU 3') was ^{32}P , 5'-end labelled using T4 polynucleotide kinase and then purified by 16% polyacrylamide gel electrophoresis (PAGE).

6.5.1 Inhibition of (-) Strong Stop DNA Synthesis

The effect of the various oligonucleotides on RT RNA Dependent DNA Polymerase (RDDP) activity of HIV-1 RT was examined in fixed-time assays by using the 496 nucleotide pHIV-PBS donor RNA template primed with a synthetic 18 nt DNA oligomer complementary to the HIV-1 primer binding sequence (PBS) in the donor RNA template. Template/primer was prepared by mixing pHIV-PBS RNA (250 pmol) with 18 nt PBS primer (500 pmol) in a final volume of 100 μL of 50 mM Tris-HCl (pH 7.8, 37°C).

Mixtures were heated at 85°C for 15 min., cooled to 50°C for 30 min. to allow specific annealing of the PBS primer to the template, and then cooled further to 37°C for 30 min. and 4°C for 30 min. to allow for restoration of template RNA secondary structure. Reaction mixtures for the (-) strong stop DNA synthesis (50 μ L total volume) contained 50 pmol preformed T/P, 250-2500 pmol oligonucleotide, 100 μ M of each of the four dNTPs and tracer [α -³²]-dCTP in 50 mM Tris-HCl (pH 7.8, 37°C) containing 10 mM dithiothreitol, 60 mM KCl and 10 mM MgCl₂. Reactions were initiated by the addition of 25-50 ng p51/p66 RT, followed by incubation at 37°C. After an appropriate incubation period, two aliquots were removed. One aliquot was quenched with 500 μ L of ice cold 10 % trichloroacetic acid containing 20 mM sodium pyrophosphate, kept on ice for 30 min., then filtered on Whatman 934-AH glass fibre filters and counted in a liquid scintillation spectrometer in order to determine total dNMP incorporation. The other aliquot was immediately placed on ice, and polymerization products were extracted with phenol/chloroform, followed by precipitation with sodium acetate/ethanol, and centrifugation at 12000 g for 30 min. The resulting pellet was washed with 70% ethanol then dissolved in 25 μ L loading buffer (22.5 mM Tris-borate, pH 8, containing 98% deionized formamide, 10 mM EDTA, 1mg/mL Bromophenol Blue and 1mg/mL xylene cyanol) and heated at 100°C for 5 min. Polymerization reaction products were resolved on a 10% sequencing gel containing 7M urea in 22.5 mM Tris-borate, pH 8, 0.5mM EDTA, and visualized by autoradiography (Kodak X-OMAT film). Polymerization products were quantified by densitometry and/or by excision of appropriate bands from the gel followed by analysis by liquid scintillation spectrometry. The amount of each product was normalized based on the number of possible insertion sites for the radiolabelled substrate [α -³²P]-dCTP in the sequence. Similar results were obtained with both methods of quantitation.

6.5.2. RNase H Induction and Priming Assays

A typical RNase H induction assay was carried out at rt, (25°C), 30°C or 37°C in 10 μ L aliquots which included 60mM Tris-HCl (pH 7.8, 37°C), 2mM dithiothreitol, 60 mM KCl, 7.5 pmols of antisense oligonucleotides, 2.5 pmol of ³²P 5'-end labelled template RNA, and 2.5 mM Mg²⁺ or 0.1 mM Mn²⁺ in concentrations specified in the text. The reaction was

started by adding 25 ng (0.2 pmol) of HIV-1 RT or 3.75 unit of *E.coli* RNase H (GIBCO/BRL). The reaction was quenched after 2 h. (overnight incubation) by adding 2x loading buffer (TBE buffer containing 98% deionized formamide, 10mM EDTA, 1 mg/mL bromophenol blue and 1 mg/mL xylene cyanol) and heated at 100°C for 5 min. In order to be denatured, the samples containing 3'-5' RNA antisense were boiled for 20 min. due to the very high affinity of the RNA template and the ONT. The reaction products were resolved on a 16% sequencing gel containing 7 M urea in TBE buffer and visualized by autoradiography (Kodak X-OMAT film). Polymerization products were quantified by densitometry and /or cutting the desired band from the gel and measuring the radioactivity using a scintillation counter.

The priming assay employed identical conditions with the exception that only Mg^{2+} was used, and the Mg^{2+} concentration was 10 mM instead of 2.5 mM. The concentration of the RT varied between 5-10 nM.

6.6. MONOMER PREPARATION

6.7.1 The 2'-O-acetyl-5'-O-monomethoxytrityl-3'-O-N,N-diisopropyl- β -cyanoethyl phosphoramidites of N⁴-Benzoyl-5'-O-arabincytosine, N⁶-Benzoyl- arabinoadenosine, and arabinouridine were prepared according to published procedures.^{109,169}

6.7.2 N⁶-Benzoyl-5'-O-monomethoxytritylarabinoadenosine-2',3'-bis-O-(N,N-diisopropyl- β -cyanoethylphosphoramidite) was synthesized according to procedures of Damha and Ogilvie.¹⁷⁵

6.7.3 N²-Isobutyryl-O⁶-[2-(4-nitro-phenyl)ethyl]-9-[2'-O-acetyl-3'-O-(N,N-diisopropyl- β -cyanoethylphosphoramidite)-5'-O-monomethoxytrityl- β -D-arabinofuranosyl] guanine (**2.8**).

a. 9-[3',5'-O-(1,1',3,3'-Tetraisopropylidisiloxane-1,3-diyl)- β -D-ribofuranosyl]guanine (2.1**).**¹⁵⁷

Anhydrous guanosine (2.04g, 2.83 mmol) was co-evaporated in absolute pyridine (3 times) and then further dissolved in a solution of 30 mL pyridine and 6 mL dry DMF.

To the stirred mixture was added 1,3-dichloro-1,1',3,3'-tetraisopropyldisiloxane (2.6 mL, 2.83 mmol), *via* a syringe and the mixture stirred overnight at rt. The reaction was complete within 20 h. and was thus stopped by the addition of water (35 mL). The white precipitate was suction filtered, washed with water and recrystallized from methanol (20 mL) to which 20 mL of diethyl ether was added to allow maximum formation of a white precipitate. This was filtered and dried to yield 2.29 g (84%). TLC: EtOAc/Tol/MeOH (4.5:4.5:1) R_f 0.25. FAB MS (Nitrobenzyl alcohol (NBA)) calc. M:525, found: MH^+ 526.

1H -NMR: (DMSO- D_6): 10.6 (s, NH amide); 7.73 (s, 1H, H-C(8)); 6.5 (s, NH_2); 5.65 (d, H1', $J_{1,2'} = 1.5$ Hz); 4.35 (t, H2'); 4.25 (m, H3'); 4.1 (d, H4'); 3.97 (dd, H5' and H5''); 1.00 (m, Si-CH(CH $_3$) $_2$ and CH $_3$).

b. O⁶-[2-(4-Nitrophenyl)ethyl]-9-[3',5'-O-(1,1',3,3'-tetraisopropyldisiloxane-1,3-diyl)- β -D-ribofuranosyl]guanine (2.2**).¹⁷³**

To a solution of triphenyl phosphine (0.6 g, 2.8 mmol) and 2-(4-nitrophenyl)-ethanol (507 mg, 3.0 mmol), in dry THF (38.0 mL) was added first diethylazodicarboxylate (0.33 mL 2.8 mmol) *via* a syringe and finally the nucleoside **2.1** (1.0 g 1.9 mmol). The reaction was stopped after 2 h. by evaporating to dryness (2.8g). The orange gummy resin obtained, was dissolved in CHCl $_3$, worked up with NaHCO $_3$, brine and MgSO $_4$. A column attempted in toluene/EtOAc (2:1) did not separate the components. Treatment with diethyl ether resulted in a white spongy solid (mp of 113-120°C, actual 150°C), believed to be Ph $_3$ PO. A second column on the resulting mixture using a gradient 3:1 to 1:1 toluene/EtOAc yielded a yellow foam of 0.9g. (73%). TLC:(toluene/EtOAc) 1:1, R_f 0.19. (CH $_2$ Cl $_2$ /MeOH) 9:1, R_f :0.48. FAB MS (NBA) calc. M:674, found: MH^+ 675.

1H -NMR: (DMSO- D_6):8.16, (d, 2H *o* to NO $_2$); 7.89, (s, H-(C-8)); 7.61 (d, 2H *m* to NO $_2$); 6.46 (s, NH_2); 5.74 (d, H1', $J_{1,2'} = 1.5$ Hz); 5.61 (d, OH-C2'); 4.65 (t, OCH $_2$ CH $_2$); 4.36 (m, H2' and H3') 4.04 (m, H4', H5', H5''); 3.22 (t, OCH $_2$ CH $_2$); 1.12-0.98 (m, Si-CH(CH $_3$) $_2$ and CH $_3$);

c. O⁶-[2-(4-Nitrophenyl)ethyl]-9-{3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)-2'-O-[(trifluoromethyl)sulfonyl]- β -D-ribofuranosyl}guanine(2.3**).**

A solution of **2.2** (0.52g 0.77 mmol), 4-(dimethylamino)pyridine (0.37g, 3.08 mmol), and absolute pyridine (0.6 mL, 6.93 mmol) in absolute CH₂Cl₂ (10.0 mL) was stirred in an ice-bath for 1 h. After addition of trifluoromethanesulfonic anhydride (0.194mL, 1.15 mmol) over 5 min. at 0°C, the mixture was stirred for another 3h at rt. The reaction was stopped by dilution with CH₂Cl₂ (10 mL), followed by successive washing with H₂O (10.0 mL), 0.4M HCl (2 x 12.5 mL), sat. NaHCO₃ (2 x 20mL), NaCl (2 x 20 mL), and finally dried over MgSO₄. Flash column chromatography (FC) (toluene/EtOAc 6:1) yielded 0.16 g of **2.3** as an off-white foam and 0.3g a second yellow crop (82%). TLC (toluene/EtOAc 4:1) R_f: 0.3. Purification by FC was found to be unnecessary as it reduced yields. FAB MS (NBA) calc. M:806, found: MH⁺ 807.

¹H-NMR: (CDCl₃):8.17 (d, 2H, *o* to NO₂); 8.09(s, H-C(8)); 7.53 (d, 2H *m* to NO₂); 6.10 (s, H1', J_{1,2} = 0 Hz); 5.56 (d, H2'); 4.93 (m, H3'); 4.80 (t, OCH₂CH₂); 4.16 (m, H4'); 4.26 and 4.0 (H5', H5''); 3.31 (t, OCH₂CH₂); 1.02 (m, Si-CH(CH₃)₂ and CH₃ from 4 *i*-Pr).

d. O⁶-[2-(4-Nitrophenyl)ethyl]-9-{2'-O-acetyl-3',5'-O-(1,1',3,3'-tetraisopropylidisiloxane-1,3-diyl)- β -D-arabinofuranosyl}guanine (2.4**)¹⁶⁷**

To a solution of **2.3** (100 mg, 0.12 mmol) in DMF (1.2 mL) and hexamethylphosphoramide (HMPA, 0.5 mL) was added LiOAc (50 mg, 0.72 mmol) and the mixture stirred overnight at rt. Under stirring, the mixture was dropped into ice-water (10.0 mL) and the resulting precipitate filtered (suction), washed with H₂O, and dried (high vacuum), resulting in a yellow powder (58 mg, 65%). TLC (tol/EtOAc 1:1) R_f: 0.5.

Attempts on larger scales produced compound **2.4** an orange gum. This was dissolved in EtOAc and dried over MgSO₄ resulting in a sticky orange foam. FAB MS (NBA) calc. M:716, found: MH⁺ 717 and 1157, 2(NBA + Na⁺).

¹H-NMR: (CDCl₃):8.16 (d, 2H, *o* to NO₂); 7.87 (s, H-C(8)); 7.48 (d, 2H, *m* to NO₂); 6.32 (d, H1', J_{1,2} = 6.0 Hz); 5.52 (d, H2'); 4.70 (m, H3', OCH₂CH₂); 4.08 and 4.8 (m, H5', H5''); 3.87 (m, H4'); 3.27 (t, OCH₂CH₂); 1.95 (s, CH₃ from acetyl); 1.02 (m, Si-CH(CH₃)₂ and CH₃ 4 *i*-Pr).

e. N²-Isobutyryl-O⁶-[2-(4-nitrophenyl)ethyl]-9-[2'-O-acetyl-3',5'-O-(1,1,3,3-tetra isopropylidisiloxane-1,3-diyl)-β-D-arabinofuranosyl]guanine(2.5**)**

To a solution of **2.4** (170 mg, 0.23 mmol) in absolute pyridine (5 mL) was added 4-DMAP (5 mg) followed by isobutyryl chloride (25 μL, 0.25 mmol) at rt. After stirring for 30 min. the reaction mixture was evaporated to dryness. The resulting orange foam was dissolved in EtOAc, washed with sat. NaHCO₃ (2x), brine (2x) and finally dried with anhydrous MgSO₄. The organic layer on coevaporation with toluene, CH₂Cl₂, and Et₂O yielded 160 mg. of **2.5** as a light orange foam 160 mg. (90%). TLC (DCM:EtOAc 1:1) R_f 0.75. FAB MS (NBA) calc. M: 787, found MH⁺ 787.

¹H-NMR: (CDCl₃): 8.14 (d, 2H, *o* to NO₂); 8.03 (s, NH); 7.78(s, H-C(8)); 7.49 (d, 2H, *m* to NO₂); 6.64 (d, *J* = 6.5 Hz, H1'); 5.52 (d, H2'); 4.8 (m, H3'); 4.68 (t, OCH₂CH₂); 4.08, (m, 2H-C(5')); 3.87 (d, H4'); 3.09 (t, OCH₂CH₂); 1.69 (s, CH₃ from acetyl); 1.20 (d, CH(CH₃)₂); 1.02 (m, Si-CH(CH₃)₂ and CH₃ from 4 *i*-Pr).

f. N²-Isobutyryl-O⁶-[2-(4-nitrophenyl)ethyl]-9-[2'-O-acetyl-β-D-arabinofuranosyl]guanine (2.6**).**

To a solution of the nucleoside **2.5** (2.63g, 1.52 mmol) in anhydrous THF (35.0 mL), was added glacial acetic acid (0.68 mL 11.4 mmol) followed by 1.0 M TBAF in THF (2.6 mL, 9.1 mmol). The mixture was stirred at rt for 24 h. EtOAc (120 mL) was added to the mixture and the organic layer washed 2x with sat. NaHCO₃, followed by brine, and dried over MgSO₄. On evaporation of the solvent, a light yellow foam was obtained (80%). Purification *via* FC (CH₂Cl₂/MeOH 0-5%) yielded a white foam but yields were poor (*ca.* 12%) probably due to decomposition on the silica gel column. TLC (CHCl₃:MeOH 19:1), R_f: 0.1. Thus upon scaling up of this product FC purification was avoided. MS: (FAB-NBA)) calc. M:544, found: MH⁺ 545.

¹H-NMR: (CDCl₃): 8.15 (d, 2H, *o* to NO₂); 8.07 (s, N-H); 7.9 (s, H-C(8)); 7.5 (d, 2H, *m* to NO₂); 6.5 (s, *J* = 6.5 Hz, H1'); 5.36 (d, H2'); 4.77 (m, H3', OCH₂CH₂); 4.0 (m, H4'); 3.9 (H5' and H5''), 3.3 (t, OCH₂CH₂, *J* = 10Hz); 1.73 (s, CH₃ from acetyl); 1.2 (m, Si-CH(CH₃)₂ and CH₃ from 4 *i*-Pr).

g. N²-Isobutyryl-O⁶-[2-(4-nitrophenyl)ethyl]-9-[2'-O-acetyl-5'-O-monomethoxy trityl-β-D-arabinofuranosyl]guanine(2.7)^{109,176}

Nucleoside **2.6** (2.2 g, 4.6 mmol) was coevaporated with pyridine (50 mL x 3) and dissolved in 30 mL pyridine. A solution of MMTrCl (2.12g, 6.9 mmol) in pyridine (10 mL) was added and the mixture allowed to stir at rt overnight. An additional 0.25 equivalent of MMTrCl was added (0.5 g) and the reaction was allowed to proceed for another 16 h. The reaction was stopped by addition of EtOAc (50 mL). The reaction mixture was evaporated to remove as much of the pyridine and then worked up with sat. NaHCO₃, brine and finally dried over MgSO₄. FC purification 0-5% MeOH/ CH₂Cl₂. Yield (2.59 g, 69%) of a light yellow foam. TLC: R_f: 0.52 (1:10 MeOH/CH₂Cl₂). FAB MS (NBA)) calc. M:817, found: MH⁺ 817.

¹H-NMR: (CDCl₃): 8.17 (d, 2H *o* to NO₂); 7.95 (s, N-H); 7.86 (s, H-C(8)); 7.51 (d, 2H *m* to NO₂); 6.81 -7.46 aromatic; 6.66 (s, *J*_{1,2'} = 4.5 Hz, H1'); 5.29 (d, H2'); 4.79 (t, OCH₂CH₂); 4.59 (m, H3'); 4.08, (m, H5'); 4.17(d, H4'); 3.78 (s, OMe, trityl); 3.47 (m, H5''); 3.32 (m, OCH₂CH₂); 1.78, (Me from acetyl); 1.23 (m, Si-CH(CH₃)₂) and 1.02 (m, 4 *i*-Pr).

h. N²-Isobutyryl-O⁶-[2-(4-nitro-phenyl)ethyl]-9-[2'-O-Acetyl-3'-O-(N,N-diisopropyl-β-cyanoethylphosphoramidite)-5'-O-monomethoxytrityl-β-D-arabinofuranosyl] guanine (2.8).

To a solution of nucleoside **2.7** (0.12g, 0.15 mmol) in dry THF (2 mL) was added N,N-diisopropylethylamine (356 μL, 2.0 mmol) followed by N,N-diisopropyl-β-cyanoethylphosphoramidic chloride (35μL, 1.05 eq.). The reaction turned cloudy after 15 min., and was complete in 1 hr. as shown by TLC: EtOAc/TEA (99:0.05) R_f 0.84 and 0.76. Pre-washed EtOAc (20 mL, 5% in NaHCO₃) was added to stop the reaction, which was worked up as in the previous step. The crude gum (148mg) was purified by column chromatography (Hexanes/ CH₂Cl₂/TEA 50:45:5) to yield 96 mg of a white foam (64%); FAB MS (NBA) calc. M:1017, found: MH⁺ 1017. ³¹P-NMR: (acetone-D₆):150.91, 150.69.

BIBLIOGRAPHY

- ¹ Saenger, W. *In Principle of Nucleic Acid Structure*. Cantor, C.R. (Ed.) Springer-Verlag Inc.: New York, NY, USA, **1984**.
- ² Blackburn, G.M. and Gait, M.J. *In Nucleic Acids in Chemistry and Biology*. Blackburn, G.M. and Gait, M.J. (Eds.) Oxford University Press: New York, **1990**.
- ³ Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K. and Watson, J. *In Molecular Biology of the Cell*. Garland Publishing Inc: New York, NY, USA, **1983**.
- ⁴ Darnell, J.; Lodish, H. and Baltimore, D. *In Molecular Cell Biology*, W.H. Freeman and Company, New York, NY, USA, **1990**.
- ⁵ Stryer, L. *In Biochemistry*, W.H. Freeman and Company, New York, NY, USA, **1988**.
- ⁶ Baltimore, D. *Nature*, **1970**, 226, 1209.
- ⁷ (a) Westheimer, F. H. *Acc. Chem. Res.*, **1968**, 1, 70. (b) Westheimer, F.H. *Science*, **1987**, 235, 1173.
- ⁸ Crick, F. *Nature*, **1970**, 227, 561.
- ⁹ Maher III, J. L.; Dervan, P. D. and Wold, B. *In Prospects for Antisense Nucleic Acid Therapy of Cancer and AIDS*. E. Wickstrom, (Ed.) Wiley-LISS, Inc.; New York, NY. **1991**, 227.
- ¹⁰ Crooke, S. T. and Lebleu, B. (Eds). *In Antisense Research and Applications*. CRC Press: Boca Raton, Florida, **1993**.
- ¹¹ Crooke, S. T. *Anti-Cancer Drug Design*, **1997**, 12, 311.
- ¹² Simons, R. W. and Kleckner, N. *Cell*, **1983**, 34, 683.
- ¹³ (a) Cazenave, C.; Chevrier, M.; Thuong, N.T. and Hélène, C. *Nucleic Acids Res.*, **1987**, 10507. (b) Minshall, J and Hunt T. *Nucleic Acids Res.*, **1986**, 114, 6433.
- ¹⁴ (a) Walder, R. Y. and Walder, J. A. *Proc. Natl. Acad. Sci. USA*, **1988**, 85, 5011. (b) Stavrianopoulos, J. G.; Gambino-Giuffrida, A. and Chargaff, E. *Proc. Natl. Acad. Sci. USA*, **1976**, 73, 1087.
- ¹⁵ Boutorine, A. S.; Brault, D.; Takasugi, M.; Delgado, O. and Hélène C. *J. Am Chem. Soc.*, **1996**, 118, 9469, and references therein.
- ¹⁶ Milligan J. F.; Matteucci, M. D. and Martin, J. C. *J. Med. Chem.*, **1993**, 36, 1923.

-
- ¹⁷ Caruthers, M. H. *Acc. Chem. Res.*, **1991**, 24, 278.
- ¹⁸ Crooke, S. T., Oligonucleotides Therapeutics. In *Burger's Medicinal Chemistry and Drug Discovery*. Wolff, M.E. Ed. Wiley: NY. **1995**, 1, 863 and references therein.
- ¹⁹ Letsinger R. I.; Finnan, J. L.; Heavner, G. A. and Lunsford, N. B. *J. Am. Chem. Soc.*, **1975**, 97, 3278.
- ²⁰ Beaucage, S. L. and Caruthers M.H. *Tet. Lett.*, **1981**, 22, 1859.
- ²¹ Pon, R.T. In *Methods in Molecular Biology, Protocols for Oligonucleotides and Analogues*, Agrawal, S., Ed. Humana Press Inc.: Totowa, NJ. USA, 20, **1993**, 465.
- ²² (a) Ogilvie, K. K.; Damha, M.J. and Usman, N. and Pon, R. T. *Pure & Appl. Chem.* **1987**, 59, 325. (b) Usman, N.; Ogilvie, K. K.; Jiang, M.-Y. and Cedergren, R. J. *J. Am. Soc.*, **1987**, 109, 7845.
- ²³ Uhlmann, E. and Peyman, A. *Chem. Rev.* **1990**, 90, 543.
- ²⁴ (a) Englisch, U and Gauss, D. H. *Angew. Chem. Int. Ed. Eng.*, **1991**, 30, 613. (b) Goodchild, J. *Bioconjugate Chem.*, **1990**, 1, 165.
- ²⁵ (a) Mertes, M. P and Coates, E.A. *J. Med. Chem.*, **1969**, 12, 154. (b) Tillenson, J. R. *J. Chem Soc. C.*, **1971**, 5656. (c) Mungall, W.S. and Kaiser, J.K. *J. Org. Chem.* **1977**, 42, 703 (d) Coull, J. M; Carlson, D. V. and Weith, H. L. *Tet. Lett.*, **1987**, 28, 745. (e) Cormier, J. F. and Ogilvie, K. K. *Nucleic Acids Res.*, **1988**, 16, 4583. (f) Vasseur, J. J.; Debart, F.; Sanghvi Y. S. and Cook, P. D. *J. Am. Chem. Soc.*, **1992**, 114, 4006. (g) Idziak, I.; Just, G.; Damha, M. J. and Giannaris, P.A. *Tet Lett.*, **1993**, 34, 5417. (h) Lebreton, J.; De Mesmaeker, A.; Waldner, A.; Frutsch, V.; Wolf, R. M. and Freier, S. M. *Tet. Lett.*, **1993**, 34, 6383.
- ²⁶ Matsukura, K.; Shinozuka, K.; Zon, G.; Mitsuya, H.; Reitz, M.; Cohen, J. S. and Broder, S. *Proc. Natl. Acad. Sci. USA*, **1987**, 84, 7706.
- ²⁷ Seeberger, P.H.; Yen, E. and Caruthers, M.H. *J. Am. Chem. Soc.*, **1995**, 117, 472.
- ²⁸ Miller, P. S.; McParland, K.B.; Jayaram, K. and Ts'o, P.O.P. *Biochemistry*, **1981**, 20, 1874. (b) Miller, P. S. and Ts'o P.O.P. *Anti Cancer Drug Design*, **1987**, 2, 117. (c) Miller, P. S.; Yano, E.; Carroll, C.; Jayaram, K. and Ts'o, P. O. P. *Biochemistry*, **1979**, 18, 5134.

-
- ²⁹ Li, H.; Porter, K.; Huange, F. and Shaw, B.R.; *Nucleic Acids Res.*, **1995**, 23, 4495 (b) Higson, A. P.; Sierzchala, A.; Brummel, H.; Zhao, Z.; Caruthers, M. H. *Tet Lett*, **1998**, 39, 3899.
- ³⁰ (a) Letsinger, R. L.; Singman, C. L.; Histan, G. and Salunkhe, M. *J.Am. Chem. Soc.*, **1988**, 110, 4470.
- ³¹ Belikova, A. M.; Zarytova, V. F. and Grineva, N. I., *Tet. Lett.*, **1967**, 37, 3557.
- ³² Eckstein, F. *Ann. Rev. Biochem.*, **1985**, 54, 367.
- ³³ Matsukura, M. *In Antisense RNA and DNA*, Wiley-Liss, NY, **1992**, 285.
- ³⁴ Furdon, P. J.; Dominski, Z. and Kole, R. *Nucleic Acids Res.*, **1989**, 17, 9193.
- ³⁵ Altmann, K.-H.; Fabbor, D.; Dean, N. N.; Geiger, T.; Monia, B.P.; Muller, M and Necklin, P. *Biochem. Soc. Trans.*, **1996**, 24, 630.
- ³⁶ Heillila, R.; Schuab, G.; Wickstrom, E.; Loke, S. L.; Pluznik, D. H.; Watt, R. and Neckers, L. M. *Nature*, **1987**, 328, 445.
- ³⁷ Sanghvi, Y.; Hoke, G. D.; Freier, S. M.; Zounes, M.C.; Gonzalez, C.; Cummins, L. Sasmor, H. and Cooke, P. D. *Nucleic Acids Res.*, **1993**, 21 3197.
- ³⁸ Biala, E.; Jones, A.S. and Walker, R. T. *Tetrahedron*, **1980**, 36, 155.
- ³⁹ Froehler, B. C.; Wadwani, S.; Terhorst, T. J. and Gerrad, S.R.; *Tet. Lett.*, **1992**, 33, 5307.
- ⁴⁰ Hall, K. B. and McLaughlin, L. W. *Biochemistry*, **1991**, 30, 1795.
- ⁴¹ Cheong, C.; Tinoco, I. and Chollet, A. *Nucleic Acids Res.*, **1988**, 16, 5115.
- ⁴² De Mesmaeker, A.; Haner, R.; Martin, P. and Moser, H. E.; *Acc. Chem. Res.* **1995**, 28, 366 and references cited therein.
- ⁴³ Damha, M. J.; Giannaris, P. A. and Marfey, P. *Biochemistry*, **1993**, 33, 7877.
- ⁴⁴ (a) Ogilvie, K. K.; Nguyen, N.; Gillen, M. F.; Radatus, B.K.; Cheriyan, V. O.; Hanna, H.R.; Smith, S. M. and Galloway, K. S. *Can. J. Chem.*, **1984**, 62, 241. (b) Schneider, K.C. and Benner, S. A. *J. Am. Chem. Soc.*, **1990**, 112, 453. (c) Smith, K.O.; Galloway, K.S.; Hodges, S.L.; Ogilvie, K.K.; Radatus, B.K.; Kalter, S.S. and Heberling, R.L. *Am. J. Vet. Res.*, **1983**, 44, 1032.

-
- ⁴⁵ Nielson, P. E.; Eghlom, M.; Berg, R. H. and Buchardt, R. H. *Science*, **1991**, 254, 453.
(b) Egholm, M.; Buchardt, O.; Nielson, P. E. and Berg, R. H. *J. Am. Chem. Soc.*, **1992**, 114, 1895.
- ⁴⁶ Stirehak, E. P.; Summerton, J. E. and Weller, D. D. *Nucleic Acids Res.*, **1989**, 17, 6129.
- ⁴⁷ (a) Nielson, P. E.; Egholm, M.; Berg, R. M. and Buchardt, O. *Anti-Cancer Drug Design*, **1993**, 8, 53. (b) Cohen, J. S. *In Antisense Research and Applications*, Crooke, S. T. and Lebleu, B., Eds.: CRC Press: Boca Raton, Florida, **1993**.
- ⁴⁸ Inoue, H.; Hayase, Y.; Imura, A.; Miura, K. and Ohtsuka, E. *Nucleic Acids Res.*, **1987**, 15, 6131.
- ⁴⁹ (a) Furukawa, Y.; Kobayashi, Y. K.; Kanai, Y. and Honjo, M. *Chem. Phar. Bull.*, **1965**, 12, 1273. (b) Iribarren, A. M.; Sproat, B. S.; Never, P.; Sulston, U.; Ryder, U. and Lamond, A. I. *Proc. Natl. Acad. Sci. USA*, **1990**, 87, 7741.
- ⁵⁰ Herdewijn, P.; De Winter, H.; Doboszewski, B.; Verheggen, I.; Augustyns, K.; Hendrix, C.; Saison-Behmoaras, T.; De Ranter, C. and Van Aershot, A. *In 'Carbohydrate Modifications in Antisense Research' ACS Symposium Series 580*; Sanghvi, Y.S. and Cook, P. D. Eds. American Chemical Society. Washington D.C. **1994**, 80-99.
- ⁵¹ De Bouvere, B.; Kerremans, L.; Hendrix, C.; De Winter, H.; Schepers, G.; Van Aerschot, A. and Herdewijn, P. *Nucleosides and Nucleotides*, **1997**, 16, 973.
- ⁵² Sanghvi, Y.S. and Cook, P. D. *In Nucleosides and Nucleotides as Antitumor and Antiviral Agents*; Chu, C.K., Baker, D. C., Eds. Plenum Press. New York, **1993**, 331. (and reference therein).
- ⁵³ (a) Leonetti J. P.; Rayner, B.; Lemaitre, M.; Gaynor, C.; Milhared, P. G.; Imbach, J. L. and Lebleu, B. *Gene*, **1988**, 72, 323. (b) Letsinger, R. L.; Zhang, G.; Sun, D. K.; Ikeuchi, T. and Sarin, P. S. *Proc. Natl. Acad. Sci. U.S.A.*, **1989**, 86, 6553. (c) Asseline, U.; Delaire, M.; Lancelot, G.; Toulme J. J.; Thuong, N. T.; Monteray-Garestier, T. and Hélène, C. *Proc. Natl. Acad. Sci. U.S.A.*, **1984**, 81, 3297. (d) Gryaznov, S. and Lloyd, D.H. *Nucleic Acids Res.*, **1993**, 21, 5909.
- ⁵⁴ Veal, G. J.; Agrawal, S. and Byrn, R. A. *Antiviral Research*, **1998**, 38, 63.

-
- ⁵⁵ Dennis, J. U.; Dean, N. M.; Bennett, C.F.; Griffith, J. W.; Lang, C. M. and Welch, D. R. *Cancer Letters*, **1998**, 128, 65.
- ⁵⁶ Gonzlezcabrera, P. J.; Iversen, P.L.; Liu, M. E. and Jeffries, W. B. *Molecular Pharmacology*, **1998**, 53, 1034.
- ⁵⁷ Abe, T.; Suzuki, S.; Hatta, T.; Takai, K.; Yokota, T. and Takaku, H. *Antiviral Chemistry and Chemotherapy*, **1998**, 9, 253.
- ⁵⁸ Lin, S. B.; Hsieh, S. H.; Hsu, H. L.; Lai, M.Y.; Kan, L. S. and Au, L. C. *Journal of Biochemistry*, **1997**, 122, 717.
- ⁵⁹ Hélène, C. and Lancelot, G. *Prog. Biophysics Molecular Bio.*, **1982**, 39, 1.
- ⁶⁰ Crooke, S. T. *Cancer Investigations*, **1996**, 14, 89.
- ⁶¹ Felsenfeld, G.; Davies, D. R. and Rich. A. *J. Am Chem. Soc.*, **1957**, 79, 2023.
- ⁶² Riley, M.; Maling, B. and Chamberlin, M. J. *J. Mol. Biol.*, **1966**, 20, 359. (b) Lipsett, M. N. *J. Biol. Chem.*, **1964**, 239, 1256. (c) Thiele, D. and Guschbauer, W., *FEBS Lett.* I, **1968** 173. (d) Morgan, A. R. and Wells, R. D. **1968**, 37, 63. (e) Howard, F. B. and Miles, H.T. *Biochemistry*, **1977**, 16, 4647. (f) Felsenfeld G. and Miles, H. T. *Annu. Rev. Biochem.*, **1967**, 36, 407. (g) Michelson, A. M.; Massoulie, J. and Guschbauer, W. *Prog. Nucleic Acids Res. Mol. Biol.*, **1967**, 6, 83.
- ⁶³ (a) Arnott, S. and Bond, P.J. *Nature New Biol.* **1973**, 244, 99. (b) Arnott, S. and Selsing, E. J. *Mol. Biol.* **1974**, 88, 509.
- ⁶⁴ Moser, H. E. and Dervan, P. D. *Science*, **1987**, 238, 645.
- ⁶⁵ Hoogsteen, K. *Acta Crystallogr.*, **1959**, 12, 822.
- ⁶⁶ Wells, R. D.; Collier, D. A.; Hanvey, J.C.; Shimizu, M. and Wohlrab, F. FASEB, **1988**, 2, 2939. (b) Kilpatrick, M.; Torru, A.; Kang, D.; Engler, J. and Wells R. *J. Biol. Chem.*, **1986**, 261, 11350.
- ⁶⁷ (a) Rajagopal, P. and Feigon, J. *Nature*, **1989**, 339, 637. (b) Rajagopal, P. and Feignon, J. *Biochemistry*, **1989**, 28, 7859. (c) De los Santos, C.; Rosen, M. and Patel, D. *Biochemistry*, **1990**, 29, 8820.
- ⁶⁸ (a) Letai, A. G.; Pallandino, M.A.; Fromm, E.; Rizzo, V. and Fresco, J. R. *Biochemistry*, **1988**, 27, 9108. (b) Arnott, S.; Bond, P.J.; Selsing. E. and Smith, P. J. C. *Nucleic Acids Res.*, **1976**, 3, 2459.

- ⁶⁹ Hattori, M.; Frazier, J and Miles, T.D. *Biopolymers*, **1976**, 15, 523. (b) Howard, F. B.; Frazier, J.; Lipsett, M. N. and Miles, H.T. *Biochem, Biophys. Res. Commun.*, **1964**, 17, 93. (c) Live, D. H.; Radhakrishnan, I.; Misra, V. and Patel, D. J. *J. Am. Chem. Soc.*, **1991**, 113, 4687.
- ⁷⁰ Maher III, L. J.; Wold, B. and Dervan, P. B. *Science*, **1989**, 215, 725.
- ⁷¹ Maher, L. J.; Dervan, P. B. and Wold B. *Biochemistry*, **1992**, 31, 70.
- ⁷² (a) Ono, A.; Ts'o P.O.P and Kan, L. *J. Am Chem Soc.*, **1991**, 113, 4032.
- ⁷³ (a) Young, S. L.; Krawczyk, S. H.; Matteucci, M. D. and Toole, J. J. *Proc. Natl. Acad. Sci., USA*, **1991**, 88, 10023. (b) Krawczyk, S. H.; Milligan .F.; Wadwani, S.; Moulds, C.; Froehler, B. and Matteucci, M. J. *Proc. Natl. Acad. Sci. USA*, **1992**, 9, 3761.
- ⁷⁴ Koh, J. S. and Dervan, P.B. *J. Am. Chem Soc.*, **1992** , 114, 1470.
- ⁷⁵ Cooney, M.; Czernuszewicz, G.; Postel E. H.; Flint, S. J. and Hogan, M. E., *Science*, **1988**, 241, 456.
- ⁷⁶ Durland, R. H.; Kessler, D. J; Gunnell, S.; Duvic, M.; Pettitt, B. M. and Hogan, M. E. *Biochemistry*, **1991**, 30, 9246.
- ⁷⁷ Cheng, A.-J.; Van Dyke, M. W. *Nucleic Acids Res.*, **1993**, 21, 5630.
- ⁷⁸ Milligan, J. F.; Krawczyk, S. H.; Wadwani, S. and Matteccu, M. D. *Nucleic Acids Res.*, **1993**, 21, 327.
- ⁷⁹ Pilch, D. S.; Levenson, C. and Shafer, R. H.; *Biochemistry*, **1991** 30, 6081.
- ⁸⁰ Postel, E. H.; Flint, S. J.; Kessler, D. J and Hogan, M. E. *Proc. Natl. Acad. Sci. USA.*, **1991**, 88, 8227.
- ⁸¹ Grigoriev, M.; Praseuth, D.; Guieysse, A. L.; Robin, P.; Thuong, N. T.; Hélène, C. and Harel-Bellan, A. *Proc. Natl. Acad. Sci., USA*, **1993**, 90, 3501.
- ⁸² Rando, R. F.; De Paolis, L.; Durland, R. H.; Jayaram, K.; Kessler, D. J. and Hogan, M. E. *Nucleic Acids Res.*, **1994**, 22, 678.
- ⁸³ Shimizu, M.; Inoue, H. and Ohtsuka, E. *Biochemistry*, **1994**, 33, 606.
- ⁸⁴ (a) François, J.-C.; Saison-Behmoaras, T.; Barbier, C.; Chassignol, M.; Thuong, N. T. and Hélène C. *Proc. Natl. Acad. Sci. USA.*, **1989**, 86, 9702. (b) François, J.-C.; Saison-Behmoaras, T.; Chassignol, M.; Thuong, N.T. and Hélène C. *J. Biol. Chem.* **1989**, 264, 5891. (c) Hélène C.; Thuong, N.T.; Saison-Behmoaras, T. and François, J.-C. *Trends*

Biotech. **1989**, 7, 310. (d) Strobel, S.A.; Moser, H.E. and Dervan, P.B. *J.Am. Chem. Soc.*, **1988**, 110, 7927. (e) Strobel, S.A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D.E. and Dervan, P.B. *Science*, **1991**, 254, 1639. (f) Perrouault, L.; Asseline, U.; Rivalle, C.; Thuong, N.T.; Bisagni, E.; Giovannangeli, C.; Le Doan, T. and Hélène C. *Nature (London)* **1990**, 344, 358.

⁸⁵ Roberts, R. W and Crothers D. M. *Science*, **1992**, 258 1463.

⁸⁶ Han, H. and Dervan, P. B. *Proc. Natl. Acad. Sci. USA*, **1993**, 90, 3806.

⁸⁷ Escudé, C.; François, J-C.; Sun, J-S.; Ott, G.; Sprinzl, M. Garestier, T. and Hélène, C. *Nucleic Acids Res.*, **1993**, 21, 5547.

⁸⁸ (a) Macaulay, V. M.; Bates, P. J.; McLean, M. J.; Rowlands, M. G.; Jenkins, T.C.; Ashworth, A. and Neidle, S., *FEBS Letters*, **1995**, 372, 222. (b) Guieysse, A.-L.; Praseuth, D.; Griogoriev, M.; Harel-Bellan, A. and Hélène C., *Nucleic Acids Res.*, **1996**, 24, 4210.

⁸⁹ Clarenc, J. P.; Lebleu, B. and Leonetti, J.P. *J. Biol. Chem.*, **1993**, 268, 5600.

⁹⁰ Neidle, S. *Anti-Cancer Drug Design*, **1997**, 12, 433.

⁹¹ (a) Hélène, C. *Anti Cancer Drug Design*, **1991**, 6, 569. (b) Thuong, N. T. and Hélène, C. *Angew Chem. Intl. Ed.*, **1993**, 32, 666.

⁹² (a) Musso, J.; Wang, C. J. and van, Dyke, M. W. *Nucleic Acids Res.*, **1996**, 24, 4924. (b) Gunter, E. J; Havre, P. A.; Gasparro, F. P. and Glazer, P. M. *Photochemistry and Photobiology*, **1996**, 63, 207. (c) Raha, M.; Wang, G.; Seidman, M. M. and Glazer, P. M. *Proc. Natl. Acad. Sci., USA*, **1996**, 93, 2941. (d) Wang, G.; Seidman, M. M. and Glazer, P.M. *Science*, **1996**, 271, 802. (e) Wang, G. and Glazer, P. M. *J. Biol. Chem.* **1995**, 270, 22595.

⁹³ (a) Kochetkova, M. and Shannom, M. F. *J. Biol. Chem.*, **1996**, 271, 14438. (b) Aggarwal, B.B. Schwarz, L.; Hogan, M. E. and Rando, R. F. *Cancer Res.*, **1996**, 56, 5156. (c) Giovannageli, C.; Diviacco, S.; Labrousse, V.; Gryaznov, S.; Charneau, P. and Hélène, C. *Proc. Natl. Acad. Sci., USA*, **1997**, 94, 79.

⁹⁴ Puglisi, J. D.; Tinoco, I. Jr. *In Methods in Enzymology*, Dahlberg, J. E. and Abelson, J. N., (Eds.) Academic Press, Inc.; New York, N. Y., USA, **1989**, 180, 304.

⁹⁵ (a) Stevens, C. L. and Felsenfeld, G. *Biopolymers*, **1964**, 2, 293.

-
- ⁹⁶ Arnott, S.; Chandrasekaran, R. and Martilla, C. M. *Biochem. J.*, **1974**, 41, 537.
- ⁹⁷ (a) Leontis, N.B.; Hills, M.T.; Piatto, M.; Malhotra A.; Nassbaum, J. and Goresnstein, D.G.J. *Biomolec. Struct. Dynam.*, **1993**, 11 215. (b) Marky, L.A.; Kallenbach, N.R.; McDonough, K.A.; Seeman, N.C. and Breslauer, K. J. *Biopolymers*, **1987**, 26, 1621.
- ⁹⁸ a) Rich, A. and Tinoco, I. Jr. *J. Am. Chem. Soc.*, **1960**, 82, 6409. (b) Weissbluth, M. *Quarterly Rev. Biophys.*, **1971**, 1,1.
- ⁹⁹ Pilch, D. S.; Levenson, C. and Shafer, R. H. *Proc. Natl. Acad. Sci. USA*, 87, **1990**, 1942.
- ¹⁰⁰ Harada, N. and Nakanishi, K. *In Circular Dichroic Spectroscopy, Exciton Coupling in Organic Stereochemistry*. Columbia University, New York, USA University Science Books, **1983**. 1-31. (b) Johnson, W. C. *In Circular Dichroic Principles and Applications*, Nakanishih, K.; Berova, N. and Woody, R.W. (Ed.) Library of Congress, **1994**.
- ¹⁰¹ Atkins, P. W. *In Physical Chemistry* 4th ed.; W. H. Freeman and Company: New York, NY **1990**, 654.
- ¹⁰² Pilch, D.S., Brousseau, R. and Shafer, R.H. *Nucleic Acids Res.*, **1990**, 18, 5743.
- ¹⁰³ Crooke, S.T. *In Therapeutic Applications of Oligonucleotides, Chem. and Ind.* **1996**, 90.
- ¹⁰⁴ Stein, C.A.; Subasinghe, C.; Shinozuka, K. and Cohen, J.S. *Nucleic Acids Res.*, **1998**, 16, 3206. (b) Cummins, L.; Graff, D.; Beaton, G.; Marshall, W. S. and Caruthers M.H. *Biochemistry*, **1996**, 4272.
- ¹⁰⁵ Brach, A.D. *TIBS*, **1998**, 23, 45.
- ¹⁰⁶ Agrawal, S.; Mayrand, S. H.; Zamenick, P.; Pederson, T.; *Proc. Natl. Acad. Sci. USA*, **1990**, 87, 1401.
- ¹⁰⁷ Young R.S.K. and Fisher, G.A. *Biochem. Biophys. Res. Commun.*, **1968**, 32, 23.
- ¹⁰⁸ (a) Reist, E.J.; Sturm, P.A.; Pong, R.Y.; Tanga, M.J. and Sidwell, R.W. *In Nucleotide Analogues As Antiviral Agents*. J. C. Martin Ed. American Chemical Society. Washington, D.C. **1989**. 17-34. (b) Rideout, J .L. and Beacham, III (Eds). *In Nucleosides, Nucleotides and Their Biological Activity*. Academic Press, New York, **1983**, and

references cited therein. (b) Wright, G. E. and Brown, N.C. *Pharmacol. Ther.*, **1990**, 47, 447.

¹⁰⁹ Giannaris, P. A. and Damha, M. J. *Can. J. Chem.* **1994**, 72, 909.

¹¹⁰ Verhoef, V. and Fridland, A. *Cancer Res.*, **1985**, 45, 3646.

¹¹¹ Sproat, B. S.; Lamond, A. I.; Beijer, B.; Nener, P. and Ryder, U. *Nucleic Acids Res.* **1989**, 17, 3373.

¹¹² Gotfredsen, C. H.; Jacobsen, J. P. and Wengel, J. *Tet. Lett.*, **1994**, 35, 6941.

¹¹³ Gotfredsen, C. H.; Jacobsen, J. P. and Wengel, J. *Bioorg. Med. Chem.*, **1996**, 4, 1421.

¹¹⁴ Hudson, R. H. E. and Damha, M. J. *J. Am. Chem. Soc.*, **1993**, 115, 2119.

¹¹⁵ Wing, R. M.; Drew, H.R.; Takano, T.; Broka, C.; Tanaka S.; Itakura, K. and Dickerson, R. E. *Nature*, **1987**, 287, 755.

¹¹⁶ Hudson, R. H. E., Uddin, A. H. and Damha, M. J. *J. Am. Chem. Soc.*, **1995**, 117, 12470.

¹¹⁷ Kuchta, R. D. and Thompson, H. C. *Biochemistry*, **1995**, 34, 11198.

¹¹⁸ (a) Hart, J.S.; Shirakawa, S.; Trujillon, J. and Frei, E. III. *Cancer Res.*, **1969**, 29, 1325. (b) Bodey, G. P.; Freiereich, E.J.; Monto, R.W. and Hewlett, J.S. *Cancer Chemother. Rep.*, **1969**, 53, 59.

¹¹⁹ (a) Carey, R.W.; Ribas-Mundo, M. and Ellison, R.R. *Cancer*, **1976**, 36, 1516. (b) Clarkson, B.; Dowling, M.D.; Gee, T.S.; Cunningham, J.H. and Burchenal, J.H. *Cancer*, **1976**, 36, 775. (b) Kufe, D.W. and Spriggs, D.R. *Semin. Oncol.*, **1985**, 12, 34.

¹²⁰ Zittoun J.; Marquet, J.; Maniey, D and Zittoun, R. *Cancer Chemother. Pharmacol.*, **1989**, 24, 251. (b) Mirzayans, R.; Anrais, B and Paterson, M.C. *Int. J. Radiat. Biol.*, **1992**, 62, 417. (c) Gunji, H.; Kharbanda, S. and Kufe, D.W. *Cancer Res.*, **1991**, 51, 741. (d) Kufe, D. W.; Munroe, D.; Herrick, D.; Egan, E. and Spriggs, D. *Mol. Pharmacol.*, **1984**, 26, 124.

¹²¹ (a) Chien, L. T.; Schabel, F.M. Jr. and Alford, C.A. Jr. In *Selective Inhibitors of Viral Functions*, Carter, W.A. (Ed.) CRC Press, Cleveland Ohio, **1975**, 227-256. (b) Pavan-Langston, D.; Buchanan, R.A.; Alford, C.A. Jr., (Ed.). In *Adenine Arabinoside: An Antiviral Agent*. Raven Press, New York, N.Y., **1975**. (c) Pursoff, W.H. and Ward, D.C. *Biochem. Pharmacol.*, **1976**, 25, 1223.

- ¹²² Sidwell, R.W.; Allen, L.B.; Huffman, J.H.; Khwaja, T.A.; Tolman, R.L. and Robins, R.K. *Chemotherapy*, **1973**, 19, 325. (b) Kaufman, H.E. and Varnell, E.D. *Antimicrob. Agents Chemother.*, **1976**, 10, 885.
- ¹²³ Grever, M.R.; Collman, C.A.; Files, J.C.; Greenberg, B.K.; Hutton, J.J.; Talley, R.L.; Von Hoff, D.D. and Balcerzak, S.P. *Blood*, **1986**, 68, 771.
- ¹²⁴ Brokman, R.W.; Schabel, F.M. and Montgomery, J.A. *Biochem. Pharmacol.*, **1977**, 26, 2193.
- ¹²⁵ (a) Vince, R. and Daluge, S.J. *Med Chem.* **1977**, 20, 612. (b) Vince, R.; Daluge, S.; Lee, H.; Shannon, W. M.; Arnette, G.; Schafer, T.W.; Nagabusham, T.L.; Reichert, P. and Tsai, H. *Science*, **1983**, 221, 1405. (b) Nobuya, K.; Matsushashi, Y.; Kokufuda, H.; Takebayashi, M. and Kaneko, C. *Tet. Lett.*, **1997**, 38, 1961.
- ¹²⁶ Wang, L.; Karlsson, A.; Arner, E.S. and Eriksson, S. *J. Biol. Chem.*, **1993**, 268, 2847.
- ¹²⁷ Coleman, C.N.; Stroller, R.G.; Drake, J.C. and Chabner, B.S. *Blood*, **1975**, 46, 791. (b) Hande, K. and Chabner, B. *Cancer Res.*, **1978**, 38, 579.
- ¹²⁸ Bell D.E and Fridland, A. *Biochim Biophys Acta*, **1980**, 606, 57. (b) Kufe, D W.; Major, P. P.; Munrow, D.; Egan, M. and Herrick, D. *Cancer, Res.*, **1983**, 43, 2000.
- ¹²⁹ Kuwakado, K.; Kubota, M.; Hirota, H.; Adachi, S.; Matsubaru, K.; Kasai, Y.; Akiyama, Y. and Mikawa, H. *Biochemical Pharmacology*, **1993**, 46, 1909.
- ¹³⁰ Graham, F and Whitmore, C. *Cancer Res.*, **1970**, 30, 2636. (b) Heinz, N. and Hamlin, J. *Biochemistry*, **1983**, 22, 3557.
- ¹³¹ Collins, A. *Biochim Biophys. Acta*, **1977**, 478, 461.
- ¹³² Manfredini, S.; Baraldi, P.G.; Bazzanini, R.; Simoni, D.; Balzarinin, J. and De Clercq, (Eds.) *Bioorg. Med. Chem. Lett.*, **1997**, 1, 473.
- ¹³³ Cohen, A. and Gel, L.J. *Blood*, **1983**, 61, 660.
- ¹³⁴ Fridland, A and Verhoef, V. *Proc. Soc. Exp. Biol. Med.*, **1985**, 179, 456.
- ¹³⁵ Ullman, B and Martin, D.W. Jr. *J. Clin. Invest.*, **1984**, 74, 951.
- ¹³⁶ Verhoef, V. and Fridland, A. *Cancer Res.*, **1985**, 45, 3646.
- ¹³⁷ Fairbanks, L. D.; Taddeo, A.; Duley, H. A. and Simmonds, H. A. *J. Immunol.*, **1990**, 144, 585.

- ¹³⁸ Krenitsky, T.A.; Tuttle, J.V.; Koszalka, G.W.; Chen, I.S.; Beacham L.M.; Rideout, J.L. and Elion, G.B. *J Biol Chem.*, **1976**, 251, 4055.
- ¹³⁹ Shewach, D. S. and Mitchell, B. S. *Cancer Res.*, **1989**, 49, 6498. (b) Shewach, D.S.; Daddona, P.E.; Ashcraft, E. and Mitchell, B.S. *Cancer Res.*, **1985**, 45, 1008.
- ¹⁴⁰ Lewis, R.A. and Link, L. *Biochem. Pharmacol.*, **1989**, 38, 2001.
- ¹⁴¹ Lewis, R.A. and Haag, R.K. *In Purine and Pyrimidine Metabolism in Man VII Part B*, Harkness, R.A., Elion, G.B. and Zollner, N. (Eds.) Plenum Publishing Corp., New York **1991**, 181-184.
- ¹⁴² Miller, R.L.; Nelson, D.J.; LaFon, S.W.; Miller, W.H. and Krenitsky, T.A. *Biochemical Pharmacology*, **1987**, 30, 2519.
- ¹⁴³ Lambe, C. U; Averett, D.R.; Paff, M.T.; Reardon, J.E.; Wilson, J.G. and Krenitsky, T. *Cancer Res.*, **1995**, 55, 3352.
- ¹⁴⁴ Krenitsky, T.A.; Koszalka, G.W.; Tuttle, J.V.; Rideout, J.L. and Elion, G.B. *Antimicrob. Agents Chemother.*, **1991**, 35, 851.
- ¹⁴⁵ Vince, R.; Daluge, S.; Kee, H.; Shannon, W.M.; Arnette, G.; Schafer, T.W.; Nagabhusan, T.L.; Reichert, P. and Tsai, H. *Science*, **1983**, 221, 1405.
- ¹⁴⁶ Wechter, W. *J. Med. Chem.*, **1967**, 10, 762.
- ¹⁴⁷ Mikita, T. and Beardsley, G.P. *Biochemistry*, **1988**, 27, 4698.
- ¹⁴⁸ Mikita, T. and Beardsley, G.P. *Biochemistry*, **1995**, 94, 33.
- ¹⁴⁹ Gao, Y.-G.; van der Marel, G.A.; van Boom, J.H. and Wang, A, H.-J. *Biochemistry*, **1991**, 30, 9922.
- ¹⁵⁰ (a) Chwang, A.K. and Sundarlingam, M. *Nature*, **1973**, 243, 78. (b) Teng, M.-K.; Liaw, Y.-C.; van, der Marel, G.A.; van Boom, J.H. and Wang, A. H.-J. *Biochemistry*, **1989**, 28, 4923.
- ¹⁵¹ (a) Resmini, M. and Pfeleiderer, W. *Helv. Chim. Acta*, **1993**, 76, 158. (b) Pfeleiderer, W. 10th Int. Roundtable: *Nucleosides and Nucleotides*, Park City, Utah, Sept 16 **1992**.
- ¹⁵² (a) Pieters, J.L.M.; de Vroom, E.; van der Marel, G.A.; van Boom, J.H.; Koning, T.M.G.; Kaptein, R. and Altona, C. *Biochemistry*, **1990**, 29, 788.
- ¹⁵³ Adams, A. D; Petrie, C.R. and Meyer, R.B. Jr. *Nucleic Acids Res.*, **1991**, 19, 3647.

- ¹⁵⁴ (a) Ogilvie, K.K.; Schefman, A.L. and Penney, C.L. *Can. J. Chem.*, **1979**, 57, 2230. (b) Ogilvie, K.K.; Beaucage, S.L.; Schiffman, A.L.; Theriault, N.Y. and Sadana, K.L. *Can. J. Chem.*, **1978**, 56, 2768. (c) Pon, R.T. and Ogilvie, K.K. *Nucleosides and Nucleotides*, **1984**, 3, 485.
- ¹⁵⁵ Beardsley, G. P.; Mikita, T.; Klaus, M and Nussbaum, A.L. *Nucleic Acids Res.*, **1988**, 16, 9165.
- ¹⁵⁶ (a) Gioeli, C.; Chattopadhyaya, J.B.; Drake, A.F. and Oberg, B. *Chem. Scr.* **1982**, 19, 13. (b) Ogilvie, K.K.; McGee, D.P.C.; Boisvert, S.M.; Hakimelahi, G.H. and Proba, Z.A. *Can. J. Chem.*, **1983**, 61, 1204. (c) McGee, D.P.C. M.Sc. thesis, McGill University, Montreal, Canada **1981**.
- ¹⁵⁷ Markiewicz, W.T. *J. Chem. Res.*, **1979**, 24, 181.
- ¹⁵⁸ (a) Provenzale G. and Nagyvary, J. *Biochemistry*, **1970**, 9, 1744. (b) Nagyvary J. and Provenzale, G. *Biochemistry*, **1969**, 8, 4769. (c) Ogilvie, K. and Iwacha, D. *Can. J. Chem.*, **1970**, 48, 862. (e) Schramman, G. *Angew. Chem. Int. Ed. Engl.*, **1967**, 6, 460.
- ¹⁵⁹ Damha, M.J. *Tet. Lett.*, **1987**, 28, 1633.
- ¹⁶⁰ Reist, E.J. and Goodman, L. *Biochemistry*, **1964**, 3, 15.
- ¹⁶¹ Cheriyan, U.O. and Ogilvie, K.K. *Nucleosides and Nucleotides*, **1982**, 1, 233. (b) Hanna, N.B.; Ramaswamy, K.; Robins, R.K. and Revankar, G.R. *J. Heterocycl. Chem.*, **1988**, 25, 1899. (c) Sakairi, N.; Hirao, I.; Zama, Y. and Ishido, Y. *Nucleosides and Nucleotides*, **1983**, 2, 221.
- ¹⁶² Chattopadhyaya, J.B. and Reese, C.B. *Synthesis*, **1978**, 908.
- ¹⁶³ Ikehara, M. and Maruyama, T. *Tetrahedron*, **1975**, 31, 1369.
- ¹⁶⁴ Hansske, F. Madey, D. and Robins, M.J., *Tetrahedron*, **1984**, 40, 125.
- ¹⁶⁵ Morisawa, H.; Utagawa, T.; Miyoshi, T.; Yoshinaga, F.; Yamazaki, A. and Mitsugi, K. *Tet. Lett.*, **1980**, 21, 479.
- ¹⁶⁶ Robbins, M.J.; Zou R.; Hansske, F. and Wnuk, S.F. *Can. J. Chem.*, **1997**, 75, 762.
- ¹⁶⁷ Resmini, M. and Pfeleiderer, W. *Helv. Chim. Acta.*, **1994**, 77, 429.
- ¹⁶⁸ Ranganathan, R. and Larwood, D. *Tet Lett.*, **1978**, 16, 4341. (b) Jiang, C.; Suhadolnik, R.J. and Baker, D.C. *Nucleosides and Nucleotides*, **1988**, 7, 271.
- ¹⁶⁹ Damha, M.J.; Usman, N. and Ogilvie, K.K. *Can. J. Chem.*, **1989**, 67, 831.

- ¹⁷⁰ Reese, C.B. *Tetrahedron*, **1978**, 34, 3143. (b) Broka, C.; Hozumi, T.; Arentzen, R. and Itakura, K. *Nucleic Acids, Res.*, **1980**, 8, 5461. (c) Chattopadhyaya, J.B. and Reese, C.B. *Nucleic Acids. Res.*, **1980**, 8, 2039.
- ¹⁷¹ Mateucci, M.D. and Caruthers, M.H. *J. Am. Chem. Soc.*, **1981**, 103, 3185,
- ¹⁷² Himmelbach, F.; Schulz, B.S.; Trichtinger, T.; Charubala, R. and Pfleiderer, W. *Tetrahedron*, **1984**, 40, 1.
- ¹⁷³ Mitsunobu, O. *Synthesis*, **1981**, 1.
- ¹⁷⁴ Pon, R.T.; Usman, N.; Damha, M.J. and Ogilvie, K.K. *Nucleic Acids Res.*, **1986**, 14, 6453.
- ¹⁷⁵ Damha, M.J. and Ogilvie, K.K. *J. Org. Chem.*, **1988**, 53, 3710.
- ¹⁷⁶ Damha, M.J. and Ogilvie, K.K. Agrawal, S. (Ed.) *In Methods in Molecular Biology - Protocols for Oligonucleotides and Analogues*. Humana Press, Totowa, NJ, **1993**, 20, 81.
- ¹⁷⁷ Robins, M.J.; Wilson, J.S.; Sawyer, L. and James, M.N.G. *Can. J. Chem.*, **1982**, 81, 1911.
- ¹⁷⁸ Alvarado-Urbina G.; Sathe G.M.; Liu W.C.; Gillen M.F.; Duck P.D.; Bender R. and Ogilvie, K.K. *Science*, **1981**, 214, 270.
- ¹⁷⁹ Damha, M.J.; Giannaris, P.A. and Zabarylo, S.V. *Nucleic Acids Res.*, **1990**, 18, 3813.
- ¹⁸⁰ Karas, M. and Hillenkamp, F. *Anal. Chem.*, **1988**, 60, 2299. (b) Hillenkamp, F.; Kara, M.; Beavis, R.C. and Chait, B.T. *Anal. Chem.*, **1992**, 64, 2866. (c) Piele, U.; Zücher, W.; Schär, M. and Moser, H.E. *Nucleic Acids Res.*, **1993**, 21, 391. (d) Beavis, R.C. and Chait, B.T. *Rapid Commun. Mass Specrom.*, **1989**, 3, 233. (e) Bahr, U.; Karas, M. and Hillenkamp, F. *Anal Chem.*, **1992**, 64, 2866.
- ¹⁸¹ Altona, C. and Sundaralingam, M. *J. Am. Chem. Soc.*, **1972**, 94, 8205-8212.
- ¹⁸² (a) Plavec, J.; Thibadeau, C. and Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, 116, 6558. (b) Thibadeau, C.; Plavec, J. and Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, 116, 8033.
- ¹⁸³ (a) Lemieux, R. U. and Chü, N. J. *Abstr. Papers Amer. Chem. Soc. Meetings*, **1958**, 133, 31 N. (b) Edward, J.T. *Chem. Ind.*, **1955**, 1102.
- ¹⁸⁴ (a) Kufe, D.W.; Major, P.P.; Egan, E.M. and Beardsley G.P. *J. Biol. Chem.*, **1980**, 255, 8997. (b) Kufe, D.W. and Spriggs, D.R. *Semin, Oncol.* **1985**, 12, 34.

-
- ¹⁸⁵ (a) Pieters, J.M.; de Vroom, E; van der Marel, G.A.; van Boom, J.H. and Altona, C. *Eur. J.Biochem.* **1989**, 184, 415. (b) Pieters, J.M.; de Vroom, E; van der Marel, G.A.; van Boom, J.H.; Konig, T. M., Kapstein, R. and Altona, C. *Biochemistry*, **1990**, 29, 788.
- ¹⁸⁶ Teng, M.-K., Liaw, Y.C.; van der Marel, G.A.; van Boom, J.H. and Wang, A.H.-J.; *Biochemistry*, **1989**, 30, 9922.
- ¹⁸⁷ Berger, I.; Tereshko. V.; Ikeda, H.; Marquez, V.E.; *Nucleic Acids Res.*, **1998**, 26, 2473.
- ¹⁸⁸ (a) Kois, P. and Watanabe, K.A. *Nucleic Acids Symposium Series.*, **1993**, 29, 215. (b) Kois, P.; Tocik, Z.; Spassova, M.; Ren, W.-Y.; Rosenberg, I.; Soler, J.F. and Watanabe, K.A. *Nucleosides and Nucleotides* **1993**, 12, 1093.
- ¹⁸⁹ Okazaki, R.; Sugino, A.; Hirose, S.; Okazaki, T.; Imae, Y.; Kainuma-Kuroda, R.; Ogawa, T.; Arisawa, M.; and Kurosawa, Y. In *DNA Synthesis in Vitro*, Wells, R.D. and Inman, R.B. (Eds.) University Park Press, Baltimore MD, **1973**, 83-106.
- ¹⁹⁰ Ratmeyer, L.; Vinayak, R.; Zhong, Y.Y.; Zon, G. and Wilson, W.D. *Biochemistry*, **1994**, 33, 5298.
- ¹⁹¹ Resmini, M. Ph.D. Thesis Konstanz University, Konstanz, Germany, Hartung-GorreVerlag, **1993**.
- ¹⁹² González, C.; Stec, W.; Kobylanska, A.; Hogrege, R.I.; Reynolds, M. and James, T.L. *Biochemistry*, **1994**, 33, 1106 and references therein.
- ¹⁹³ Wang, A.C.; Kim, S.G.; Flynn, P.F.; Chou, S.-H.; Orban, J. and Reid, B.R. *Biochemistry*, **1992**, 31, 3940.
- ¹⁹⁴ Mikita, T and Beardsley, G.P. *Biochemistry*, **1994**, 33, 9195.
- ¹⁹⁵ Conte M.R.; Conn, G. L.; Brown, T. and Lane, A.N. *Nucleic Acids Res.*, **1997**, 25, 2627.
- ¹⁹⁶ Hung, S.-H.; Yu, Q.; Gray, D.M. and Ratliff, R.L. *Nucleic Acids Res.*, **1994**, 25, 4098.
- ¹⁹⁷ Wilds, C.J and Damha M. J. unpublished results.
- ¹⁹⁸ Damha, M.J.; Meng, B.; Wang, D.; Yannopoulos, C.G. and Just, G. *Nucleic Acids Res.*, **1995**, 19, 3967.
- ¹⁹⁹ Fujimoro, S.; Shudo, K. and Hashimoto, Y. *J. Am. Chem Soc.*, **1990**, 112, 7436.

- ²⁰⁰ Matteucci, M. *Nucleosides and Nucleotides*, **1991**, 12, 231. (b) Ashley, G.W. *J. Am. Chem. Soc.*, **1992**, 114, 9731.
- ²⁰¹ Kawai, S.H.; Wang, D.; Giannaris, P.A.; Damha, M.J and Just, G. *Nucleic Acids Res.*, **1993**, 21, 1473. (b) Meng, B.; Kawai, S.H.; Wang, D.; Just, G.; Giannaris, P.A. and Damha, M.J. *Angew. Chem., In. Ed.; Engl*, **1993**, 32, 729.
- ²⁰² Hélène, C. *Anti- Cancer Drug Design*, **1991**, 6, 569.
- ²⁰³ Maher, L.J. III. *Bioessays*, **1992**, 14, 807.
- ²⁰⁴ Mirkin, S. M. *Annu. Rev. Biochem.*, **1995**, 64, 65.
- ²⁰⁵ Le Doan, T.L.; Perrounault, L; Praseuth, N.; Decoult, J. L.; Thuong, N.T.; L'homme, J. and Hélène, C. *Nucleic Acids Res.*, **1987**, 15, 7749.
- ²⁰⁶ Zimmerman, S.C. and Schmitt, P. *J. Am. Chem. Soc.*, **1995**, 117, 10769.
- ²⁰⁷ Hardenbol, P. and Van Dyke, M.W. *Proc. Natl. Acad. Sci, USA*, **1996**, 93, 2811.
- ²⁰⁸ (a)Skoog, J.U. and Maher, L.J. III. *Nucleic Acids Res*, **1993**, 21, 2131. (b) Huamin, J. and Smith, L.M. *Anal. Chem.* **1993**, 65, 1323. (b) Luebke, K.J. and Dervan P.B. *J Am. Chem Soc.*, **1989**, 111, 8333.
- ²⁰⁹ Frossella, J.A.; Kin, Y.J.; Shih, H.; Richards, E.G. and Fresco, J.R. *Nucleic Acids Res.*, **1993**, 21, 4511.
- ²¹⁰ Wilson, W.D.; Hopkins, H.P.; Mizan, S.; Hamilton, D.D. and Zon, G. *J. Am. Chem. Soc.* **1994**, 116, 3607.
- ²¹¹ Booher, M.A.; Wong, S. and Kool, E.T. *Biochemistry* **1994**, 33, 4645 and references therein.
- ²¹² (a) Giovannangeli, C.; Garestier, T.; Rougee, M.; Chassignol, M.; Thuong, N.T. and Hélène, C. *J. Am. Chem. Soc. B*, **1991**, 113, 7775. (b) Azhayeva, E.; Azhayev, A.; Guzaev, A.; Hovinen, J. and Lonnberg, H. *Nucleic Acids Res.*, **1995**, 23, 1170. (c) Durland, M.; Peloille, N.T.; Thuong, N.T. and Maurizot, J.C.; *Biochemistry*, **1992**, 31, 9197. (d) Rumney, S. and Kool, E.T. *J. Am. Chem. Soc.* **1995**, 117, 5635.
- ²¹³ (a) Chen, G. and Seeman, N.C.; *Nature*, **1991**, 350, 631 (b) Hoffman, R. *American Scientist*, **1994**, 82, 308.
- ²¹⁴ (a) Ma, M.Y.-X.; McCallum, K.; Climie, S.C.; Kuperman, R.; Lin, W.C.; Sumner-Smith, M. and Barnett, R.W.; *Nucleic Acids Res.*, **1993**, 21, 2568. (b) Ma, M.Y.-X.; Reid,

- L.S.; Climie, S.C.; Lin, W.C.; Kuperman, R.; Sumner-Smith, M. and Barnett R.W. *Biochemistry*, **1993**, 423, 1751. (c) Gao, H.; Chidambaram N.; Chen, B.C.; Pelham, D.E.; Patel, R.; Yang, M.; Zhou, L.; Cook, A. and Cohen, J.S. *Bioconjugate Chemistry*, **1994**, 5, 445. (d) Gao, H.; Yang, M. and Cook, A. F. *Nucleic Acids Res.*, **1995**, 23, 285.
- ²¹⁵ (a) Seeman, N.C. *DNA and Cell Biol.*, **1991**, 10, 475.
- ²¹⁶ (a) Azhayez, A.; Gouzaev, A.; Hovinen, J.; Azhayeva, E. and Lönnberg, H. *Tet. Lett.*, **1993**, 34, 1993. (b) Azhayev, A.; Gouzaev, A.; Hovinen, J.; Azhayeva, E. and Lönnberg, H. Eleventh International Roundtable, Nucleosides, Nucleotides and Their Biological Applications, Katholieke Universiteit Leuven, Belgium, Sept. 7-11, **1994**, 119.
- ²¹⁷ Uddin, A.H.; Roman, M.A.; Anderson, J.R. and Damha, M.J. *J. Chem. Commun.* **1996**, 171.
- ²¹⁸ Zhou, B.W.; Marchand, C.; Asseline, U.; Thuong, N.T.; Sun, J.S.; Garestier, T. and Hélène, C. *J. Am. Chem Soc.*, **1995**, 117, 10425.
- ²¹⁹ Xodo, L.E.; Manzini, G. and Quadrifoglio. *Nucleic Acids Res.*, **1990**, 18, 3557.
- ²²⁰ Kool, E.T. *J. Am. Chem. Soc.*, **1991**, 113, 6265.
- ²²¹ Prakash, G. and Kool, E. T. *J. Chem. Soc. Commun.*, **1991**, 1161.
- ²²² Kandimalla, E.R. and Agarwal, S.J. *Gene*, **1994**, 149, 115.
- ²²³ Wallace, J.C. and Edmonds, M. *Proc. Natl. Acad. Sci. USA*, **1983**, 80, 850.
- ²²⁴ Prakash, G. and Kool, E.T. *J. Am. Chem Soc.*, **1992**, 114, 3523.
- ²²⁵ Braich, R. and Damha, M.J. *Bioconjugate Chem.*, **1997**, 370.
- ²²⁶ Hoogsteen, K. *Acta Crystallogr.*, **1963**, 16, 907.
- ²²⁷ Kibler-Herzog, L.; Zon, G.; Whittier, G.; Shaikh, M. and Wilson, W.D. *Anti-Cancer Drug Design*, **1993**, 8, 65.
- ²²⁸ Chalikian, T.V.; Plum, G.E.; Sarvazyan, A.P. and Breslauer, K.J. *Biochemistry*, **1994**, 33, 8629.
- ²²⁹ Damha, M.J. and Ogilvie, K.K. *Biochemistry*, **1988**, 27, 6403.
- ²³⁰ Wengel J. *Biorg. Med. Chem. Lett.*, **1995**, 8, 791 (b) Wengel J. *Tetrahedron*, **1995**, 51, 8491.
- ²³¹ Blake R.D. Massoulie, J. and Fresco, J.R. *J. Mol. Biol.*, **1967** 30, 291.
- ²³² Johnson, K.H.; Gray, D.M and Sutherland J.C. *Nucleic Acids Res.*, **1991**, 19, 2275.

-
- ²³³ (a) Musso, M. and Van Dyke, M. W. *Nucleic Acids Res.*, **1995**, 23, 2320.
- ²³⁴ Job, P. *Ann. Chim. (Paris)* **1928**, 9, 113
- ²³⁵ Gray, D.M., Ratcliff, R.L. and Vaughan, M.R. *Methods in Enzymology*, **1992**, 211, 389.
- ²³⁶ Gray, D.M.; Hung, S.-H.; and John, K.H.; *Methods in Enzymology*, **1995** 246, 19-34 and references therein.
- ²³⁷ Marky, L.A. and Breslauer, K.J. *Biopolymers*, **1987**, 9, 1601.
- ²³⁸ Strobel, S.A. and Dervan, P.B. *J. Am Chem. Soc.*, **1989**, 111, 2656.
- ²³⁹ Colocci, N.; Distefano, M.D. and Dervan, P.B. *J. Am. Chem. Soc.*, **1995**, 117, 4781.
- ²⁴⁰ Leijon, M. and Graslund, A. *Nucleic Acids Res.*, **1992**, 20, 5539.
- ²⁴¹ Riley, M.J.; Maliing, B. and Chamberlin, M. J. *J. Mol. Biol.* **1966**, 20 359.
- ²⁴² Colocci, N.; Distefano, M.D. and Dervan, P.B. *J. Am. Chem Soc.*, **1993**, 115, 4468.
- (b) Colocci, N. and Dervan, P.B. *J. Am. Chem Soc.*, **1995**, 117, 4781.
- ²⁴³ Mergny, J.-L.; Sun, J.-S.; Rougee, M.; Garestier, T.; Barcelo, F.; Chomilier, J. and Hélène, C. *Biochemistry*, **1991**, 30, 9791.
- ²⁴⁴ Semerad, C.L. and Maher III L.J. *Nucleic Acids Res.*, **1994**, 22, 5321.
- ²⁴⁵ Xu, Z.; Pilch, D.S.; Srinivasan, A.R.; Olson, W.K.; Geacintov, N.E. and Breslauer, K.J. *Bioorg. Med. Chem*, **1997**, 5, 1137.
- ²⁴⁶ Luck, G.; Zimmer, C. and Baguley, B.C. *Biochim, Biophys Acta*, **1984**, 782, 41.
- ²⁴⁷ Abu-Day, A.; Brown, P.M. and Fox, K.R. *Nucleic Acids Res.*, **1995**, 23, 3385.
- ²⁴⁸ Goodsell, D. and Dickerson, R.E. *J Med. Chem.*, **1986**, 29, 727.
- ²⁴⁹ Mergny, J-L.; Duval-Valentin, G.; Nguyen, C.H.; Perrouault, L.; Faucon, B.; Rougée, M.; Garestier, T.; Bisagani, E. and Hélène, C. *Science*, **1992**, 256, 1681.
- ²⁵⁰ Escudé, C.; Nguyen, C.H.; Mergny, J.-L.; Sun, J-S.; Bisagni, E.; Garestier, T and Hélène, C. *J. Am. Chem. Soc.*, **1995**, 117, 10212.
- ²⁵¹ (a) Pilch, D.S.; Waring, M.J.; Sun, S.J.; Rougée, M.; Nguyen, C. H.; Bisagani, E.; Garestier, T. and Hélène, C. *J. Mol. Biol* **1993**, 232, 926. (b) Radhakrishnan, I. and Patel, D.J. *J. Mol Biol*. **1994**, 241, 600.
- ²⁵² Scaria, P.V. and Shafer, R.H. *J Biol. Chem.*, **1991**, 266, 5417.
- ²⁵³ Lee, J.S., Latimer, L.J.P. and Hampel, K.J. *Biochemistry*, **1993** 32, 5591.

-
- ²⁵⁴ Strekowski, L. Gulevich, Y.; Baranowski, T.C.; Parker, A.N. Kiselyov, A.S.; Lin, S.-Y.; Tanious, F.A. and Wilson, W.D. *J. Med. Chem.* **1996** 39, 3980.
- ²⁵⁵ Fox, K.R.; Polucci, P.; Jenkins, T.C. and Neidle, S. *Proc. Natl. Acad. Sci. USA*, **1995**, 94, 79.
- ²⁵⁶ Fox, K.R.; Polucci, P.; Jenkins, T.C. and Neidle, S. *Proc. Natl. Acad. Sci., USA*. **1995**, 92, 7887.
- ²⁵⁷ Silver, G.C.; Sun, J.-S.; Nguyen, C.H. Boutrine, A.S. Bisagni, E. and Hélène, C. *J. Am. Chem. Soc.*, **1997**, 119, 263.
- ²⁵⁸ Thomas, T.J.; Farland, C.A.; Gallo, M.A. and Thomas, T. *Nucleic Acids Res.*, **1995**, 23, 3594.
- ²⁵⁹ Uddin, A. H. Ph.D. Thesis, McGill University. Montreal, Canada, **1996**.
- ²⁶⁰ Durland, M.; Thuong, N.T. and Maurizot, J.C. *Biochimie*, **1994**, 76, 181.
- ²⁶¹ Bédu, E.; Benhida, R.; Devys, M. and Fourrey, J.-L. *Tet. Lett.* **1999**, 40, 835.
- ²⁶² Xiang, G.; Bogacki, R. and McLaughlin, L.W. *Nucleic Acids Res.*, **1996**, 24, 1963.
- ²⁶³ Brunar, H. and Dervan P.B. *Nucleic Acids Res.*, **1996**, 24, 4176
- ²⁶⁴ Bates, P.J.; Laughton, C.A.; Jenkins, T.C.; Capaldi, D.C.; Roselt, P.D.; Reese, C.B. and Neidle, S. *Nucleic Acids Res.*, **1996**, 24, 4176. (c) Hildbrand, S. and Leumann, C. *Angewante Chime Intl. Ed.*, **1996**, 35, 1968.
- ²⁶⁵ (a) Macaya, R.; Schultz, P. and Feignon, J. *J. Am. Chem. Soc.*, **1992**, 114, 781. (b) Macaya, R.; Wang, E.; Schultz, P.; Sklenar, V. and Feigon, J. *J. Mol. Biol.* **1992**, 225, 755. (c) Park, Y.-W. and Breslauer, K.J *Proc. Natl. Acad. Sci. USA.*, **1992**, 89, 6653. (d) Ranghunathan, G.; Miles, H.T. and Sasisekharan, V. *Biochemistry*, **1993**, 32, 455.
- ²⁶⁶ Okazaki, R.; Sugino, A.; Hirose, S.; Okazaki, T.; Imae, Y.; Kainuma-Kuroda, R.; Ogawa, T.; Arisawa, M. and Kurosawa, Y. *In DNA Synthesis in Vitro*. Wells, R.D. and Inman, R.B., Eds. University Park Press, Baltimore MD. **1973**, 83-106.
- ²⁶⁷ Kurosawa, Y.; Okazaki, R; *J. Mol. Biol.* **1975**, 94, 229.
- ²⁶⁸ Kohlstaedt, L.A.; Wang, J.; Friedman, J.M.; Rice, P.A. and Steitz, T.A. *Science*, **1992**, 256, 1783.

- ²⁶⁹ (a) Gray, D.M. and Ratliff, R.L. *Biopolymers*, **1975**, 14, 487. (b) Zimmerman, S.B. and Pfeiffer, B.H. *Proc. Natl. Acad. Sci. USA*, **1981**, 78, 78. (c) Arnott, S.; Chandrasekaran, R.; Millane, R. and Park, H.-S. *J. Mol. Bio.* **1986**, 188, 631.
- ²⁷⁰ (a) Shindo, H. and Matsumoto, U. *J. Biol. Chem.*, **1984**, 259, 8682. (b) Benevides, J.M. and Thomas, G.J. Jr., *Biochemistry*, **1988**, 27, 3868.
- ²⁷¹ Damha, M.J. and Noronha, A. *Nucleic Acids Res.*, **1998**, 26, 5152.
- ²⁷² Liquier, J.; Coffinier, P.; Firon, M. and Taillandier, E. *J. Biomol. Struct. Dyn.*, **1991**, 9, 437.
- ²⁷³ Dagneaux, C.; Liquier, J. and Taillandier, E. *Biochemistry*, **1995**, 34, 16618.
- ²⁷⁴ Gray, D. M. and Vaughan, M. *Nucleic Acids, Res.*, **1980**, 8, 3695.
- ²⁷⁵ Gehring, K.; Leroy, J.L. and Gueron, M. *Nature*, **1993** 363, 561.
- ²⁷⁶ Chen, L.; Cai, L.; Zhang, X. and Rich, L. *Biochemistry*, **1994**, 33, 13540.
- ²⁷⁷ Singh, S., Patel, P.K. and Hosur, R.V. *Biochemistry*, **1997**, 36, 13214.
- ²⁷⁸ Volker, J. and Klump, H.H. *Biochemistry*, **1994**, 33, 13502.
- ²⁷⁹ Kanehara, H.; Mizuguchi, M.; Tajima, K. and Makino, K. *Biochemistry*, **1997**, 36, 1790.
- ²⁸⁰ Manzini, G., Yathindra, N. and Xodo, L.E. *Nucleic Acids Res.*, **1994**, 22, 4634.
- ²⁸¹ Antao, V.P. *Nucleic Acids Res*, **1988**, 16, 719. (b) Lee, J.S. *Nucleic Acids Res*, **1990**, 18, 6057. (c) Lee, J.S. *Nucleic Acids Res*, **1980**, 8, 4305.
- ²⁸² Wang, S. and Kool, E.T. *Nucleic Acids Res.*, **1994**, 22, 2326.
- ²⁸³ Young, S.L.; Krawczyk, S.H.; Matteucci, M.D. and Toole, J.J., *Proc. Natl. Acad. Sci.* **1991**, 88, 8277.
- ²⁸⁴ (a) Shimizu, M.; Konishi, A.; Shimada, Y.; Inoue, H. and Ohtsuka, E. *FEBS Letters*, **1992**, 302, 155. (b) Escudé, C.; Sun, J.S.; Rougee, M.; Garestier, T. and Hélène, C. *C.R. Acad. Sci. Paris, Serie III*, **1992**, 315, 521.
- ²⁸⁵ Egli, M.; Usman, N. and Rich, A. *Biochemistry*, **1993**, 32, 3221.
- ²⁸⁶ Wilds, C.J. and Damha, M.J. *Bioconjugate Chemistry*, **1999**, 10, 299.
- ²⁸⁷ Lalitha, V. and Yathindra, N. *Current Science*, **1995**, 68, 68.

-
- ²⁸⁸ Wyatt, J.R.; Vikers, T.A.; Robertson, J.L.; Buckeheit, R.W.J.; de Baets, T.K.E.; Davies, P.W.; Rayner, B.; Imbach, J.L. and Ecker, D.J. *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 1356.
- ²⁸⁹ Cheng, K.Y. and Tinoco, I. I. *Proc. Natl. Acad. Sci. U.S.A.*, **1994**, 91, 8705.
- ²⁹⁰ (a) Nelson, H.C.M.; Finch, J.T.; Luisi, B.F. and Klug, A.; *Nature*, **1987**, 330, 221. (b) Prive, G.G.; Heinemann, U.; Chandrasegaran, S.; Kan, M.L. and Dickerson, R.E. *Science*, **1987**, 238, 498.
- ²⁹¹ Viazovkina, E. and Damha, M.J unpublished results.
- ²⁹² Eschenmoser, A. *Science*, **1999**, 284, 2118.
- ²⁹³ Beier, M.; Reck, F.; Wagner, T.; Krishnamurthy, R. and Eschenmoser, A. *Science* **1999**, 283, 699.
- ²⁹⁴ These studies have been conducted in collaboration with M.A.Parniak, D.Arion, G. Borkow of the Lady Davis Institute for Medical Research, and McGill AIDs Centre, SMBD-JGH, Montreal Qc., Canada.
- ²⁹⁵ The Sciences May/June S. Agarwal, **1997**, 24.
- ²⁹⁶ Zamecnik, P.C. and Stephenson, M.L. *Proc. Natl. Acad. Sci. USA*, **1978**, 230, 281.
- ²⁹⁷ Zamecnik, P.C. and Agrawal S. *Nucleic Acids Symposium Series*, **1991**, 24, 127.
- ²⁹⁸ Damha, M. J.; Wilds, C. J.; Noronha, A.; Brukner, I.; Borkow, G.; Arion, D. and Parniak, M. A. *J. Am. Chem. Soc.*, **1998**, 49,12976.
- ²⁹⁹ Brown, P. *New Scientist*, **1992**, 31.
- ³⁰⁰ Dropulic, B. and Jeang, K-T. *Human Gene Therapy*, **1994**, 5, 927 and references therein.
- ³⁰¹ Borkow, G.; Arion, D.; Noronha, A.; Scartozzi, M.; Damha, M.J. and Parniak, M.A. *Int. J. Biochem. Cell Biol.*, **1997**, 29, 1285.
- ³⁰² Fuentes, G.M.; Fay, P.J. and Bambara, R.A. *Nucleic Acids Res.*, **1996**, 24, 1719.
- ³⁰³ Chiang, M.Y.; Chan, H.; Zounes, M.A.; Freier, S.M.; Lima W.F. and Bennett, C.F. *J. Biol. Chem.*, **1991**, 266, 18162.
- ³⁰⁴ Monia, B.P.; Lesnik. E.A.; Gonzalez, C.; Lima, W.F.; McGee, D.; Guinosso,C.J.; Kawazaki, A.M.; Cook, P.D. and Frier, S. M. *J. Biol. Chem.*, **1993**, 268, 14514.

-
- ³⁰⁵ Dean, N.M.; McKay, R.; Condon, T.P. and Bennett, C.F. *J. Biol. Chem.*, **1994**, 269, 16416.
- ³⁰⁶ Heidenreich, O.; Gryaznov, S.; Nerenburg, M. *Nucleic Acid Res.*, **1997**, 25, 776.
- ³⁰⁷ Arts, E.J.; Li, X.; Gu, Z.; Lkeunab, L.; Parniak, M.A. and Wainberg, M.A. *J. Biol. Chem.*, **1994**, 269, 13472.
- ³⁰⁸ (a) Sczakiel G.; Pawlita, M. and Kelnheinz, A. *Biochem. Biophys. Res. Commun.*, **1990**, 169, 643. (b) Bordier, B.; Perala-Heape, M.; Degols, G.; Lebleu, B.; Litcak, S.; Sarih-Cottin, L. and Hélène, C. *Proc. Natl. Acad. Sci. USA*, **1995**, 92, 9383. (c) Volkmann, S.; Dannull, J. and Moeling, K. *Biochemie*, **1993**, 75, 71. (d) Demirhan, I.; Hasselmayer, O.; Hofmann, D.; Chandra, A.; Svinarchuk, F.P.; Vlassov, V.V.; Engels, J. and Chandra, P. *Virus Genes*, **1994**, 9, 113. (e) Anazodo, M.I.; Salomon, H.; Friesen, A.D.; Wainberg, M.A. and Wright, J.A. *Gene*, **1995**, 166, 227. (f) Biasoslo, M.A.; Radaelli, A.; Del Pup, L.; Franchin, E.; Giuli-Morghen, C. and Palu, G. *J Virol.*, **1996**, 70, 2154.
- ³⁰⁹ Arion, D; Harada, R.; Li, X.; Wainberg, M.A. and Parniak, M.A. *Biochem Biophys. Res. Commun.*, **1996**, 225, 839.
- ³¹⁰ Ben-Artzi, H.; Zeelon, E.; Amit, B.; Wortzel, A.; Gorecki, M. and Panet, J. *Biol. Chem.*, **1993**, 268, 16465.
- ³¹¹ Momparler, R. *Mol. Pharmacol.*, **1972**, 8, 362-370. (b) Dicioccio, R. and Srivasdtava, B. I. S. *Eur. J. Biochem.*, **1977**, 79, 411.
- ³¹² (a) Moore, P. and Strauss, B. *Nature*, **1979**, 278, 664 (b) Sagher, D. and Strauss, B. *Biochemistry*, **1983**, 22, 4518.
- ³¹³ Loeb, L.A. and Kunkel, T. *Annu. Rev. Biochem.*, **1982**, 51, 429.
- ³¹⁴ Fuentes. G.M.; Rodriguez-Rodriguez, L.; Fay, P.J and Bambara, R.A. *J. Biol. Chem.* **1995**, 270, 28169.
- ³¹⁵ Misra, H.S.; Pandey, P.K.; Modak, M.J.; Vinayak, R. and Pandey, V.N. *Biochemistry*, **1998**, 37, 1917.
- ³¹⁶ Crouch, R.J. and Dirksen, M.-L. *In Nuclease*; Linn, S.M. and Roberts, R.J. (Eds.) Cold Harbour Spring Laboratory Press, Plainview New York. **1982**, 211-241.

- ³¹⁷ Uchiyama, Y.; Inoue H., Ohtsuka, E.; Nakai, C.; Kanaya, S.; Ueno, Y. and Ikehara, M. *Bioconjugate Chem.*, **1994**, 5, 327.
- ³¹⁸ Oda, Y.; Iwai, S.; Ohtsuka, E.; Ishikawa, M.; Ikehara, M. and Nakamura, H.S. *Nucleic Acids Res.*, **1993**, 21, 4690.
- ³¹⁹ Toulme, J.J.; Boizau, C.; Larrouy, B.; Frank, P.; Albert, S. and Ahmadi, R. *In DNA and RNA Cleavers and Chemotherapy of Cancer and Viral Diseases*. B. Meunier, (Ed.) Kluwer: The Netherlands, **1996**, 479, 271.
- ³²⁰ Doolittle, R.F.; Feng, D.F.; Johnson, M.S. and McClure, M.A. *Quart. Rev Biol.* **1989**, 64, 1. (b) Johnson, M.S.; McClure, M.A.; Feng, D.F.; Gray, J. and Doolittle, R.F. *Proc. Natl. Acad. Sci.*, **1986**, 7648.
- ³²¹ Katayanagi, K.; Iyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Ikehara, M.; Matsuzaki, T. and Morikawa, K. *Nature (London)*, **1990**, 347, 306.
- ³²² Yang, W.; Hendrickson, W.A.; Crouch, R.J. and Satow, Y. *Science*, **1990**, 249, 1398.
- ³²³ Nakamura, H.; Oda, Y.; Iwai, S.; Inoue, H.; Ohtsuka, E.; Kanaya, S.; Kimura, S.; Katsuda, C.; Katayanagi, K.; Morikawa, K.; Miyashiro, H. and Ikehara, M. *Proc. Nat. Acad. Sci., U.S.A.* **1991**, 88, 11535.
- ³²⁴ Davies, J.; Hostomka, Z. Hostomsky, Z.; Jordan, S. and Mathews, D. *Science*, **1991**, 252, 88.
- ³²⁵ Daniher, A.T.; Xie, J.; Mathur, S and Bashkin, J.K., *Bioorg. Med. Chem.*, **1997**, 5, 1037.
- ³²⁶ Fedoroff, O. Y.; Salazar, M. and Reid, B.R. *J. Mol. Biol.*, **1993**, 233, 509.
- ³²⁷ Crooke, S.T.; Lemonidis, K.M.; Neilson, L.; Griffey, R.; Lesnik, E.A. and Monia, B.P. *Biochem. J.*, **1995**, 312, 599.
- ³²⁸ Lima, W.F. and Crooke, S.T. *Biochemistry*, **1997**, 36, 390.
- ³²⁹ Ben-Artzi, H.; Shemesh, J.; Zeelon, E.; Amit. B.; Kleinman, L; Gorecki, M. and Panet. A. *Archives of Biochem and Biophysics*, **1996**, 325, 209.
- ³³⁰ Furfine, E.S. and Reardon, J.E. *J.Biol. Chem.*, **1991**, 266, 406. (b) Peliska, J.A. and Benkovic, S.J. *Science*, **1992**, 258, 1112.
- ³³¹ De Stefano, J.J.; Buiser, R.G.; Mallaber, L.M.; Myers, T.M.; Bambara, R.A. and Fay, P.J. *J. Biol. Chem.*, **1991**, 266, 7423. (b) Dudding, L.R.; Nkabinde, N.C. and Mizrahi, V.

-
- Biochemistry*, **1991**, 30, 10498. (c) Kati, W.M.; Johnson K.A.; Jerva, L.F. and Anderson, K.S. *J. Biol. Chem.* **1992**, 267, 25988.
- ³³² Gotte, M.; Fackler, S.; Hermann, T.; Perola, E.; Le Grice, S.F.J. and Heumann, H. *The Embol Journal*, **1995**, 14, 833.
- ³³³ Borkow, G. unpublished results .
- ³³⁴ Jacobo-Molina, A.; Ding, J.; Nanni, R.G.; Clark, A.D. Jr.; Lu, X.; Tantillo, C.; Williams, R.L.; Kamer, G.; Ferris, A.L.; Clark, P.; Hizi, A.; Hughes, S.H. and Arnold, E. *Proc. Natl. Acad. Sci., USA*, **1993**, 90 6320.
- ³³⁵ Metzger, W. Hermann, T. Schatz, O.; Le Grice, S.F.J. and Hermann, H. *Proc. Natl. Acad. Sci. USA.*, **1993**, 90, 5909.
- ³³⁶ Wohrl, B.M.; Tantillo, C.; Arnold, E. and Le Grice, F.J. *Biochemistry*, **1995**, 34, 5343.
- ³³⁷ Salazar, M.; Fedoroff, O.Y.; Miller, J.M.; Ribeiro, S. N. and Reid, B.R. *Biochemistry*, **1993**, 32, 4207.
- ³³⁸ Lok, C.-N, McGill AIDS Centre, personal communication.
- ³³⁹ Mergny, J.L.; Collier, D.; Rougee, M.; Garestier, T. and Hélène, C. *Nucleic Acids Res.*, **1991**, 19, 1521.
- ³⁴⁰ Pon, R.T.; Usman, N. and Ogilvie, K.K. *Biotechniques*, **1988**, 6, 768.
- ³⁴¹ Hudson, R.H.E. Ph.D. Thesis, University of Toronto, Toronto, Canada, **1995**.
- ³⁴² Damha, M.J.; Pon, R.T. and Ogilvie, K.K. *Tet. Lett.*, **1985**, 26, 4839.
- ³⁴³ Iyer, R.P.; Egan, W.; Regan, J.B. and Beaucage, S.L. *J. Am. Chem. Soc.*, **1990**, 112, 1253.
- ³⁴⁴ McClinton, M.A. *Aldrichimica Acta*, **1995**, 28, 31.
- ³⁴⁵ Gasparutto, D.; Livache, T.; Bazin, H.; Duplaa, A.-M.; Guy, A.; Khorin, A.; Molko, D.; Roget, A. and Teoule, R. *Nucleic Acids, Res.*, **1992**, 20, 5159.
- ³⁴⁶ Grierson, D. *In Electrophoresis of Nucleic Acids, A Practical Approach*. D. Rickwood and B.D. Hanes, (Ed.) IRL press Ltd.: Oxford, England, **1982**, 15.
- ³⁴⁷ Bhattacharyya, A.; Murchie, A.I.H. and Lilley, D.M. *Nature*, **1990**. 343, 484.
- ³⁴⁸ Dahlberg, A.E. *J. Mol. Biol.* **1969**, 41, 139.
- ³⁴⁹ Lecchi, P.; Le, H.M.T. and Pannell, L.K. *Nucleic Acids Res.*, **1995**, 23,1276.

³⁵⁰ Arts, E.J.; Li, X.; Gu.; Keliman, L.; Parniak, M.A. and Wainberg, M.A. *J. Biol. Chem.*, **1994**, 268, 16465.