Expanding the Size and Shape of Nucleic Acids: studies on Branched and Heptose based Nucleic Acids

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

By

David Sabatino

May 2007

Department of Chemistry
McGill University
Montreal, Quebec, Canada

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Dedicated to my parents Mirka and Claudio,
for their support, guidance and love.

ABSTRACT

The generation of synthetic oligonucleotides is dependent on an efficient solid-phase synthesis methodology. This thesis evaluates the 2'-*O*-levulinyl (Lv) and 2'-*O*-monomethoxytrityl (MMT) ribonucleosides, as possible synthons for RNA and branched RNA synthesis. A key feature of this RNA and bRNA synthesis procedure is their removal while still attached to the solid support and under conditions that prevent isomerization or cleavage of the nascent strands. For the first time, the stability of 3'-5'-internucleotide phosphate triesters (and diesters) adjacent to a ribose 2'-hydroxyl group was determined on a solid support. These studies are not only relevant to the proper assembly of branched and linear RNA species, but also to the stability of an unusual branched RNA species ("RNA X") proposed to form during the pre-mRNA splicing reactions in vitro. These studies are also important to the development of large quantities of native and chemically modified short interfering RNA (siRNA) for animal and human studies.

The 2'-*O*-Lv and 2'-*O*-MMT ribonucleoside monomers served as building blocks for the assembly of a series of branched nucleic acid species (bRNA, bDNA, msDNA and hyperbranched or "dendritic" DNA/RNA) with discrete length, base composition and structure. These structures were synthesized via an iterative divergent-growth strategy, which facilitates the regioselective branching, deblocking and chain lengthening steps from a branchpoint core. These structures served as useful materials (bio-probes) as demonstrated by the biological studies performed with *E. coli* RNaseH and the yeast lariat RNA debranching enzyme (yDBr1). These studies not only led to the identification of novel branched nucleic acid inhibitors of yDBR1 and RNase H, but also provided new insights about the substrate specificity of these important enzymes.

This thesis also describes the synthesis of a new nucleic acid form, the so-called "oxepane nucleic acids" (ONAs), in which the pentofuranose ring of DNA and RNA was replaced with a 7-membered heptose sugar ring. ONA were found to be much more resistant towards nuclease degradation than natural DNA, an important feature if these analogues are to be used in biological media. Furthermore, ONAs exhibited cross-pairing with complementary RNA and were found to elicit *E. coli* RNaseH mediated degradation of the RNA strand. These finding are significant because oligonucleotide-directed RNase

H degradation of the target RNA is a key determinant for the gene-specific inhibitory potency of antisense oligonucleotides. When comparing the rates of RNase H-mediated degradation induced by 5 (furanose), 6 (2'-ene-pyranose) and 7 (oxepane) membered ring oligonucleotides, the following trend was observed: DNA > 2'-ene-pyranose NA > ONA. The implications of these results are discussed in the context of our current understanding of the catalytic mechanism of the enzyme, particularly with regard to the required flexibility of the oligonucleotide strands that bind to the RNA target. Hence, ONAs are useful tools for biological studies and provide new insights into the structure/function of natural and alternative genetic systems.

RÉSUMÉ

L'efficacité dans la préparation d'oligonucléotides de synthèse dépend fortement de la mise au point de méthodologies performantes de synthèse en phase solide. Au cours de cette thèse les ribonucléosides 2'-*O*-levulinyl (Lv) et 2'-*O*-monomethoxytrityl (MMT) ont été évalués en tant que synthons potentiels pour la synthèse d'ARN et d'ARN ramifiés (bARN).

L'étape clé de cette stratégie de synthèse d'ARN et de bARN est leur décrochage du support solide dans des conditions empêchant toute isomérisation ou dénaturation des ARN. Pour la première fois, la stabilité des liens internucléotidiques (phosphotriesters et diesters) ont été étudiés sur support solide. Cette méthodologie est appropriée à la synthèse d'ARN linéaires et ramifiés, et peut être étendue à des structures ramifiées plus rares d'ARN ("ARN X") formées au cours de l'épissage des pre-mRNA. Ces travaux sont également d'intérêt pour la production de quantités importantes de siARN pour applications biomédicales.

Les ribonucléosides 2'-*O*-Lv et 2'-*O*-MMT ont servi comme synthons pour la synthèse d'une série d'acides nucléiques ramifiés (bARN, bADN, msADN et hyper ramifiés ou ADN/ARN "dendritique") avec longueur, séquence et structure bien définies. Ces structures ont été synthétisées par une stratégie itérative de croissance divergente facilitant un embranchement régiosélectif à partir d'un point de branchement. Ces nouvelles structures ont servi comme bio-sondes pour des études biologiques réalisées avec la RNaseH de l'*E. coli* et l'enzyme de débranchement d'ARN de lasso de levure (yDBr1). Ces études ont non seulement mené à l'identification de nouveaux inhibiteurs des acides nucléiques de la yDBR1 et de la RNase H, mais ont également fourni de nouvelles perspectives sur la spécificité de substrat de ces enzymes.

Cette thèse décrit également la synthèse d'une nouvelle forme d'acide nucléique: 'les oxepanes' (AONs), dans lesquels la partie osidique (pentofuranose) de l'ADN et de l'ARN a été remplacée par un heptose. Les AONs se sont avérés beaucoup plus résistants à la dégradation enzymatique (nucléase) que l'ADN non modifié. Cette stabilité est d'importance pour l'utilisation de ces analogues en milieu physiologique. En outre, une hybridation des AONs avec leur ARN complémentaire a été mise en évidence. Ces résultats sont déterminants car la dégradation de l'ARN dirigée par la RNase H est une

des causes majeures des oligonucléotides antisens. En comparant les taux de dégradation de la RNase H induits par des oligonucléotides comprenant des parties osidiques de diverses tailles, 5 (furanose), 6 (2'-ene-pyranose) et 7 (oxepane) la tendance suivante a été observée: ADN > 2'-ene-pyranose > AON's. Ces résultats sont d'importance dans notre étude du mécanisme catalytique de l'enzyme, en particulier en ce qui concerne la flexibilité nécessaire des brins d'oligonucléotides qui se lient à la cible d'ARN. Par conséquent, les AONs sont des outils très utiles pour l'étude des relations structure-activité biologique des systèmes génétiques.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

ABBREVIATIONS AND SYMBOLS

A	adenosine
Å	angstrom
A_{260} or Abs	UV absorbance measured at 260 nm
AB	alkaline buffer
Ac	acetyl
Ace	acetone
Ace-d_6	deuterated acetone
Ac_2O	acetic anhydride
Ade	adenine
AE HPLC	Anion Exchange HPLC
AIDS	Acquired Immunodeficiency Syndrome
ANA	arabinonucleic acid
AON(s)	antisense oligonucleotide(s)
APS	ammonium persulfate
ara	arabino
ATT	6-aza-2-thiothymine
A.U.	Absorbance Units
AZT	azidothymidine
B	Base
bc	bicyclo
BIS	*N,N'*-methylene-bisacrylamide
bDNA	branched deoxyribonucleic acid
bNA	branched nucleic acid
bRNA	branched ribonucleic acid
bp	base pair
BPB	bromophenol blue
Bz	benzoyl
C	cytidine
ca.	*circa* (approximately)

CAB	carboxylic acid buffer
calc	calculated
CD	circular dichroism
CEt or CNEt	2-cyanoethyl
CeNA	cyclohexene nucleic acid
CI-MS	chemical ionization mass spectrometry
Ci	Curie
CMV	cytomegalovirus
COSY	homonuclear correlation spectroscopy
Cyt	cytosine
d4T	didehydro-deoxythymidine
d	doublet
dd	doublet of doublets
DCI	4,5-dicyanoimidazole
ddH$_2$O	doubly-distilled and deionized water
D$_2$O	deuterated water
DCC	1,3-dicyclohexylcarbodiimide
DCM	dichloromethane
DEPC	diethyl pyrocarbonate
DIPEA	*N,N*-diisopropylethylamine
DMAP	4-(dimethylamino)-pyridine
DMF	*N,N*-dimethyl formamide
DMSO	dimethyl sulfoxide
DMSO-*d$_6$*	deuterated dimethyl sulfoxide
DMT	dimethoxytrityl
DMT-Cl	dimethoxytrityl chloride
dN (N = A,G,C,T)	2'-deoxynucleoside
DNA	2'-deoxyribonucleic acid
DSC	differential scanning calorimetry
ds	double stranded
dT	thymidine

dt	doublet of triplets
DTT	dithioreitol
E	eastern
E. coli	bacterium *Escherichia coli*
EDTA	ethylene-diamine tetraacetate dihydrate
e.g.	for example
ESI-MS	electrospray ionization mass spectrometry
Et	ethyl
EtOAc	ethyl acetate
EtOH	ethanol
ETT	ethyl thiotetrazole
2'-F-ANA	2'-fluoro-2'-deoxyarabinonucleic acid
2'-F-RNA	2'-fluoro-2'-deoxyribonucleic acid
fN	2'-fluoro-2'-deoxyarabino nucleotide
FBS	fetal bovine serum
G	guanosine
g	gram
GNA	glycol nucleic acid
Gua	guanine
h	hours
% H	percent hyperchromicity
HATU	*O*-(7-azabenzotrizol-1-yl)-1,1,3,3-tetramethyl-uronium hexafluorophosphate
h-bNA	hyperbranched nucleic acid
HBTU	*O*-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyl-uronium hexafluorophosphate
HCV	hepatitis C virus
HEPES	2-[4-(2-hydroxyethyl)-1-piperazine]ethanesulfonic acid
Hex	hexanes
HIV-1	human immunodeficiency virus type I
HMQC	heteronuclear multiple quantum correlation spectroscopy

HNA	hexitol nucleic acid
HOAc	acetic acid
homo DNA	$(4'{\rightarrow}6')$linked oligo(2',3'-dideoxy-β-D-glucopyranosyl)nucleotides
HPLC	high performance liquid chromatography
Hz	Hertz
i-Bu	iso-butyryl
i-Pr	iso-propyl
IC$_{50}$	concentration of inhibitor which causes 50% inhibition
i.e.	that is
IP RP HPLC	ion-pairing reverse phase HPLC
J	coupling constant
λ	wavelength
LCAA-CPG	long-chain alkylamine controlled pore glass
LNA	locked nucleic acid
Lv	levulinyl
M	molar
m/z	mass to charge ratio
MALDI-TOF-MS	matrix assisted laser desorption ionization time-of-flight mass spectrometry
max	maximum
MeCN	acetonitrile
MeOH	methanol
MeOH-*d$_4$*	deuterated methanol
μL	microliter
μM	micromolar
mM	milimolar
MMT	monomethoxytrityl
min	minimum
min	minutes
MOE	2'-methoxyethyl

mol	mole
MON(s)	modified oligonucleotide(s)
mRNA	messenger (mature) RNA
MS	mass spectrometry
msDNA	multicopy single-stranded DNA
MW	molecular weight
N	northern
NA	nucleic acid
NEt$_3$	triethylamine
NH$_4$OH	ammonium hydroxide
nm	nanometers
nM	nanomolar
NMI	*N*-methylimidazole
NMP	*N*-methylpyrrolidinone
NMR	nuclear magnetic resosonace
NNRTI	non-nucleoside reverse transcriptase inhibitors
NOE	nuclear overhauser enhancement
NOESY	nucler overhauser and exchange spectroscopy
NRTI	nucleoside reverse transcriptase inhibitors
nt(s)	nucleotide(s)
O.D.U	optical density unit
oA	oxepane adenine
ON	oligonucleotide
oN	oxepane nucleotide
ONA	oxepane or oxepine nucleic acid
oT	oxepane thymine
oT*	oxepine thymine
PAGE	polyacrylamide gel electrophoresis
pNA	ene-pyranose nucleic acid
PNA	peptide nucleic acid
PNK	polynucleotide kinase

PPB	physiological phosphate buffer
ppm	parts per million
pre-mRNA	precursor mRNA
pT	2'-ene-pyranose thymine
Pu	purine
Py	pyrimidine
Q-linker	hydroquinone-O,O'-diacetyl linker
®	registered trademark
r.t.	room temperature
R_f	retention factor (*i.e.* TLC analysis)
RISC	RNAi induced silencing complex
rN (N = A,G,C,U)	ribonucleoside
RNA	ribonucleic acid
RNAi	RNA interference
RNase H	ribonuclease H
RP-IP HPLC	reverse phase ion-pairing HPLC
rU	ribouridine
RT	reverse transcriptase
S	southern
s	singlet
SEC	Size exclusion chromatography (Sephadex® G-25)
siRNA	short interfering RNA
SNA	seco nucleic acid
snoRNA	small nucleolar RNA
ss	single stranded
succ linked LCAA CPG	succinyl linked LCAA CPG
3TC	3-thiocytosine
T	thymidine
t	triplet
t-	tertiary
TAG	tri-O-acetyl D-glucal

T*AT	parallel (or Hoogsteen) TAT triplex
T·AT	antiparallel (or reverse Hoogsteen) TAT triplex
TBAF	tetra-*n*-butylammonium fluoride
TBDMS	*tert*-butyl dimethylsilyl
TBE	Tris/boric acid/EDTA buffer
TCA	trichloroacetic acid
TEA	triethylamine
TEAA	triethylammonium acetate
TEMED	*N,N,N',N'*-tetramethylethylenediamine
TFO	triplex forming oligonucleotides
THF	tetrahydrofuran
Thy	thymine
TLC	thin layer chromatography
TMS	trimethylsilyl
TMSOTf	trimethylsilyltrifluoromethanesulfonate
T_m	thermal melting temperature
TM	trademark
TNA	α-L-threofuranosyl nucleic acid
Tr	trityl
TREAT HF	triethylamine trihydrofluoride
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
U	units (of enzyme)
U	uridine
Ura	uracil
UV	ultraviolet
UV-VIS	ultraviolet-visible
v/v	voulme per volume
VIS	visable
wt/v	weight per volume
XC	xylene cyanol
yDBr1	yeast debranching enzyme

CHAPTER 1

CHAPTER 5

LIST OF TABLES

CHAPTER 1: GENERAL INTRODUCTION

1.1 FUNCTION AND STRUCTURE OF NUCLEIC ACIDS WITHIN THE CELL

Nucleic acids are vital molecules required for the construction, storage and transmission of the genetic material found in living cells.[1] Nucleic acids are assembled by ubiquitous polymerase enzymes within the cells to yield two different types of genetic material: Deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). In most eukaryotic cells, DNA is a hybrid duplex that is capable of self-replication and transcription (transmission of the genetic material) into single-stranded RNA. The synthesis, splicing and maturation of RNA into messenger RNA, (mRNA), functionalizes this genetic material for the translation and synthesis of polypeptides which are assembled into mature protein form to execute the genetic code. Other synthesized RNA molecules such as transfer RNAs, (tRNAs), and ribosomal RNAs, (rRNAs), are responsible for the translation of mRNA into polypeptide sequences. The orderly relationship in the evolution of nucleic acids structure and function from DNA to RNA and protein synthesis is the most common method of gene expression in living organisms and is referred to as the *Central Dogma of Molecular Biology* (**Figure 1.1**).[2]
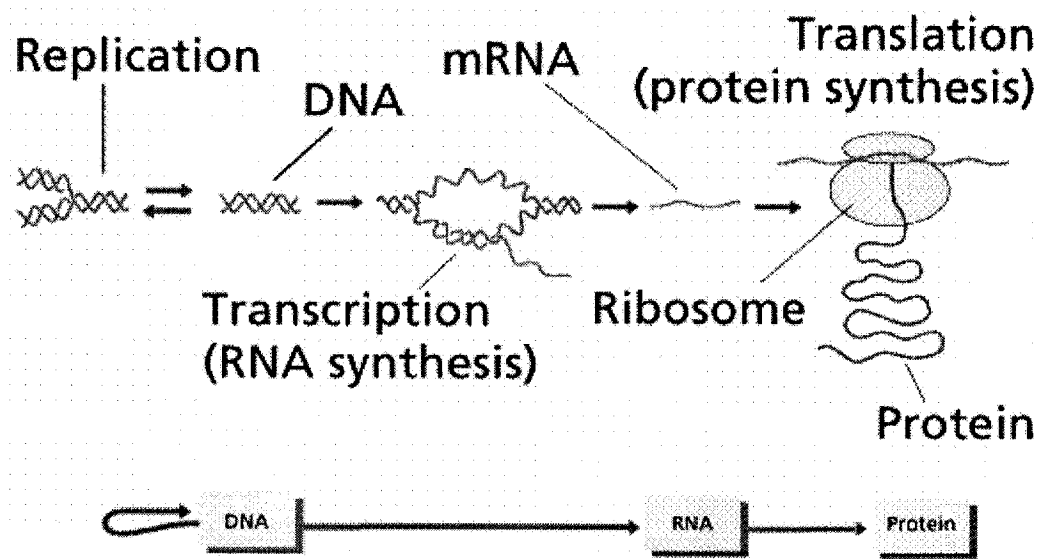
Figure 1.1: The "*Central Dogma of Molecular Biology*" describing the flow of the genetic material from DNA→ RNA→ protein in the nucleus of eukaryotic cells. The figure was taken from http://www.mcb.mcgill.ca/~hallett/GEP/Lecture2/Image12.gif.

1

DNA and RNA are constructed from different nucleotide monomers comprised of a D-ribofuranose sugar phosphate backbone and a nitrogenous purine or pyrimidine base which often favors DNA and RNA to assemble into higher ordered macromolecular structures. These two polymers (DNA and RNA) are distinguished chemically and structurally by the β-D-ribofuranose sugar moiety, which contains a 2'-hydroxyl functional group in RNA (*i.e.* ribose sugar) and a 2'-methylene group in DNA (*i.e.* 2-deoxyribose sugar) (**Figure 1.2, i.**). They are also distinguishable by the composition of the pyrimidine nitrogenous bases. The four bases found in DNA are the pyrimidines thymine and cytosine, and the purines adenine and guanine. In RNA, the purine bases remain the same with the single alteration being the thymine base substituted for a uracil base (no C5-methyl substituent). The relative stereochemistry of the *N*-glycosidic bond at the anomeric carbon of the sugar is fixed in the β-orientation and in the sterically more favored *anti* conformation, such that the heterocyclic functional groups of the base are pointing away from the sugar ring (**Figure 1.2, i).**

1.2 NUCLEIC ACID SECONDARY STRUCTURE

The changes in phenotypes in living organisms are often associated with the characteristic base sequence distribution within the coding region of the genetic material (DNA and RNA). This is contingent with the evolution of molecular biology which can be traced to James Watson and Francis Crick and their discovery of the X-Ray crystal structure and model of DNA.[3] The structure of DNA consists of a double helix in which two complementary oligonucleotide strands are associated together in an antiparallel orientation (in which the 5'$\rightarrow$ 3' directions are opposing in either strands) and both strands are referred to as the sense and antisense strands. This association is most heavily attributed to the hydrogen-bonding network, (*i.e.* H-bonding), which is generally referred to as base-pair fidelity and with the steric and hydrophobic interactions of the nitrogenous nucleotide bases, which themselves are planar and stack on top of each other along the helical axis. More specifically, these interactions stabilize the hybrid duplex in a specific Watson-Crick base pairing mode; in which the pyrimidines Cyt and Thy(Ura) hydrogen bond specifically with the purines Gua and Ade, respectively. The stronger base pairing interaction occurs with the Cyt:Gua base pairs which contain three H-bonding

interactions as opposed to Ade:Thy which only contains two H-bonding interactions **(Figure 1.2, ii.)**.

i.

A : adenine

G : guanine

C : cytosine

T : thymine X: CH₃
U : uracil X: H

ii.

Ade : Thy

Gua : Cyt

DNA : R = H (β-D-2'-deoxyribose)
RNA : R = OH (β-D-ribose)
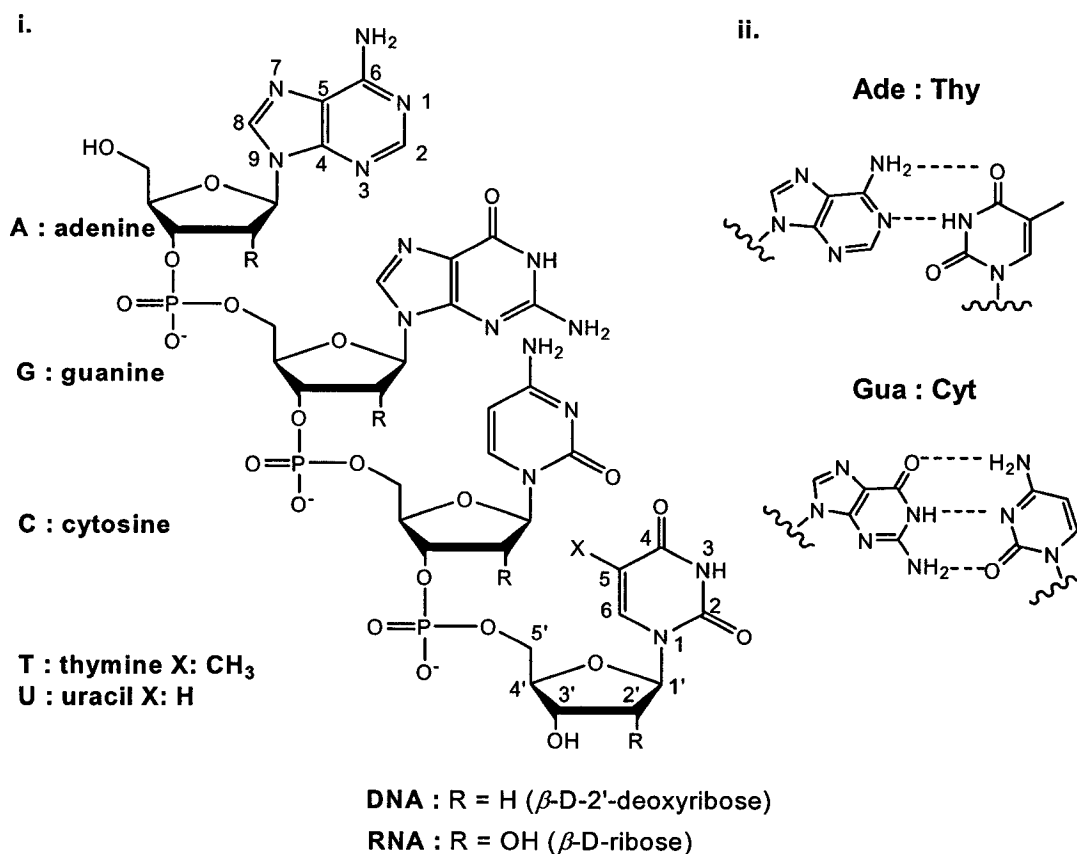
Figure 1.2: i. Primary structure of DNA and RNA containing the four nitrogenous bases (adenine, guanine, thymine and cytosine) for DNA and the four bases (adenine, guanine, uracil and cytosine) for RNA. The conventional numbering system for the β-D-deoxyribofuranose (DNA) and 2-β-D-ribofuranose (RNA) sugar moiety with the purine (Ade and Gua) and pyrimidine (Thy/Ura and Cyt) nitrogenous bases. ii. Hydrogen bonding interactions for Watson-Crick base pairs (Ade:Thy; 2 H-bonding interactions and Gua:Cyt; 3 H-bonding interactions).

The global conformation of the duplex is mostly restricted by the sugar pucker conformation of the individual nucleotide residues within the oligonucleotide strands.[4] This pucker is dependent on the displacement of the C2' and C3'-carbons from the median plane of the C1'-O4'-C4' carbons within the sugar ring **(Figure 1.3)**. In DNA, the duplex structure adopts a B-like conformation, in which the nucleotide residues form a *Southern* or C2'-*endo* sugar conformation on the pseudorotational cycle[5] (*i.e.* with C2' above the median plane) in high humidity or low salt concentrations. This nucleotide

conformation provides a duplex conformation (B-DNA) with characteristic distances between base-pairs along the helix axis of 3.4 Å with approximately 10 base pairs per helical turn. The bases are stacked above one another on the helix axis producing *major* and *minor grooves* of similar depths and wider and narrower *groove* widths, respectively **(Figure 1.3)**.

Alternatively, the highly dynamic conformation of DNA can be converted to a more compact A-like geometry in conditions of low humidity and high salt concentrations. This anti-parallel, complementary, right-handed, helical duplex structure is pre-organized by a nucleotide sugar pucker conformation which adopts a *Northern* or C3'-*endo* geometry. The pucker restricts the C3' and C5' carbons above the median plane of C1'-O4'-C4' carbons such that the C5'-C3'-phosphate diester distance is reduced to 5.9 Å (as opposed to 7.0 Å for the more extended C2'-*endo* conformation in the B-DNA duplex). As a result, the global structure of A-DNA contains a more compact conformation with 11 base pairs per helical turn and distances between base-pairs along the helix axis of 2.6 Å. Furthermore, the bases are displaced by 4.5 Å from the parallel helix axis and thereby creating a hollow core within the center of the duplex structure, resulting in a deep but narrow *major groove* and shallow but broad *minor groove* depth and width **(Figure 1.3)**.

Another less common DNA helical structure is Z-DNA[6]; a left-handed helix that is narrower and longer then A or B-DNA with a characteristically narrow *minor groove*. This conformation is usually favored with high salt concentrations and in the presence of metal counterions (*i.e.* 4 M NaCl and 100 mM NiCl$_2$) and with specific strand sequences which contain alternating purine-pyrimidine bases. The base pairs are oriented perpendicular to the helix axis generating a *zig-zag* phosphodiester backbone in the DNA duplex. These characteristic properties can be attributed to a unique nucleotide pairing conformation in which G most readily adopts a *syn*, C2'-*endo* conformation and C adopting an *anti*, C3'-*endo* conformation **(Figure 1.3)**.

More recently, DNA has also been shown to self-assemble into higher ordered structures by virtue of unique pairing and stacking interactions within the individual strands. More specifically, Helene[7] and Gehring[8], have shown that C-rich sequences can form compact, intramolecular structures involving C · C$^+$ (protonated cytosine) base pairs

4

in acid or even physiological pH by intercalation of the C · C$^+$ pairs of two parallel stranded duplexes in a global conformation referred to as i-motif DNA.

A-form RNA   B-form DNA   Z-form DNA

5.9 Å   C3'-endo

7.0 Å   C2'-endo

	A-RNA	B-DNA	Z-DNA
Helical Sense:	Right	Right	Left
Sugar Conformation:	C3'-endo	C2'-endo	C2'/C3'-endo
Helix Diameter:	25.5 Å	23.7Å	18.4 Å
Residues/Turn:	11	10.4	12
Base-pair distance from axis center:	4.5 Å	-0.2-0.8 Å	1.2 Å
Major Groove: width	2.7 Å	11.7Å	8.8 Å
depth	13.5 Å	8.8 Å	3.7 Å
Minor Groove: width	11 Å	5.7 Å	2.0 Å
depth	2.8 Å	7.5 Å	13.8 Å
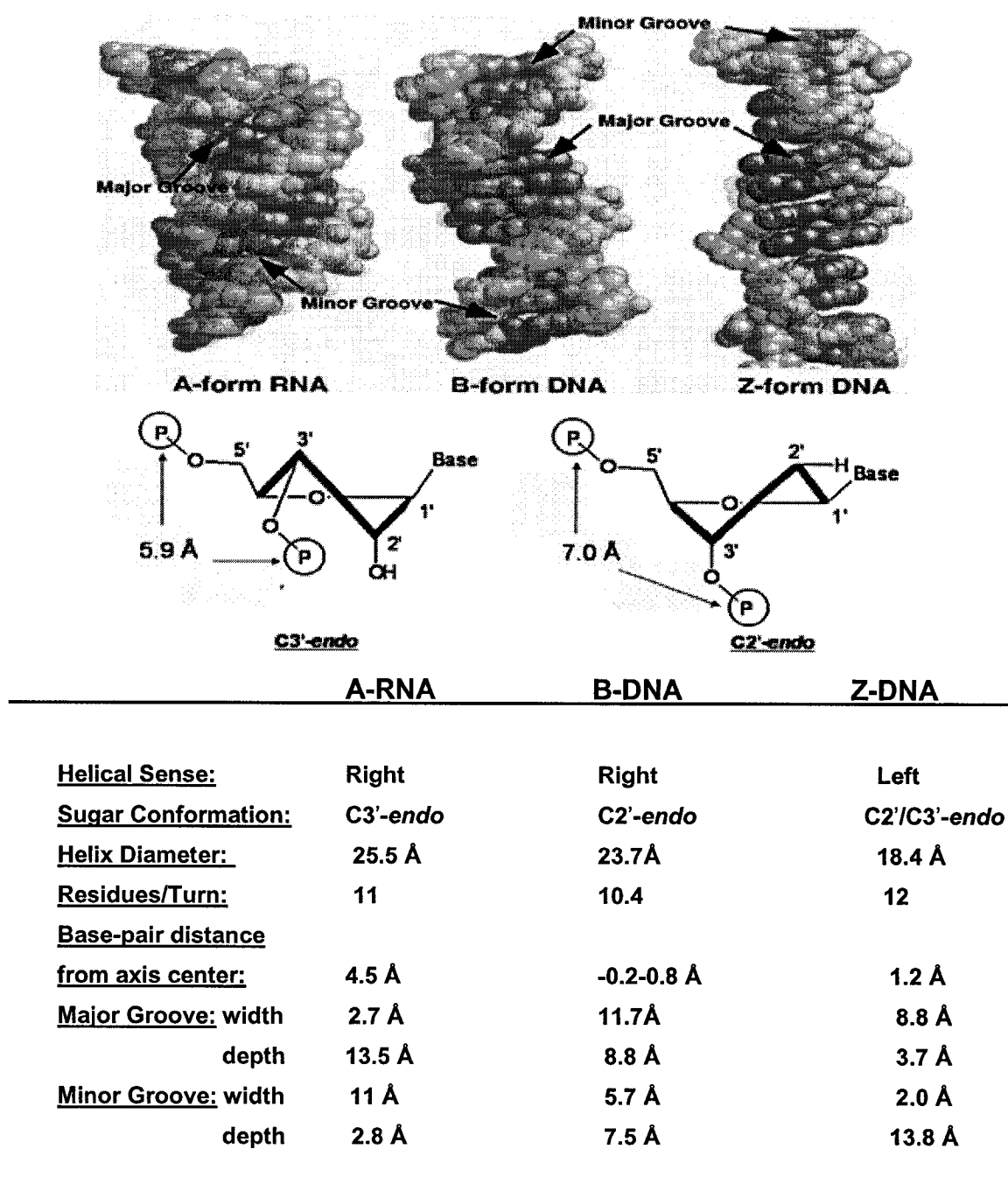
Figure 1.3: The global helical conformations of A-RNA, B-DNA and Z-DNA, their preferred sugar conformations and duplex parameters. Structures and data adapted from Blackburn; G.M. and Gait; M.J. In *Nucleic Acids in Chemistry and Biology*; Oxford University Press: New York, **1996.** and Dickerson; R.E., Drew; B.N., Conner; R.M., Wing; R.M., Fratini; A.V. and Kopka; M.L. *Science* **1982,** *216*, 475-485.

5

Another interesting and biologically more relevant tetraplex (relative to the i-motif DNA) is the so-called G-tetrad DNA, in which guanine-rich sequences can also self-assemble to generate planar stretches of four H-bonded guanines stacked in layers of three or four looped units containing strand polarity that can be either parallel, anti-parallel or both; and in which the *N*-glycosidic bond of the guanosine yields both *syn* and *anti* conformations in the global geometry.[9]

The ribonucleic acids (RNA) are functional in their single-stranded form, but double stranded RNA has also been shown to exist and perform various functional and catalytic activities within the cell.[10] The RNA nucleotide monomers are restricted to a *Northern* or C3'-*endo* conformation which generates an A-like helix geometry with the RNA duplex. This is partially due to the steric effect of the 2'-hydroxyl functional group and the nitrogenous base at the C1' position forcing the equilibrium from the C2'-*endo* to the more favored C3'-*endo* sugar pucker. This conformation is also stabilized by the O4'-O2' *anomeric effect* and the *gauche effects* of the O2'-C2'-C1'-O4' torsion angles.[11] These characteristics generate the geometric parameters found in A-RNA and A-DNA type duplexes (**Figure 1.3**).[12,13]

1.3 CHEMICAL SYNTHESIS OF OLIGONUCLEOTIDES: GENERATION OF THE SOLID-PHASE PHOSPHORAMIDITE APPROACH

During the period of the structure elucidation of the DNA duplex, pioneering work by Khorana and Todd had already been initiated for the in-solution chemical synthesis of simple dimers and short oligonucleotides by virtue of a carbodiimide-mediated coupling of the nucleotide monomers via a *phosphodiester* based approach.[14] Some of the limitations of this approach included a modest product yield and purity (*i.e.* RNA syntheses generally yield 60% product), specifically in the synthesis of trimers and longer oligonucelotide sequences. This procedure also produced a significant quantity of pyrophosphate byproducts in addition to cleavage side reactions which originated from the reaction of the internucleotide phosphodiester linkages and the reactive starting materials.

To overcome some of these limitations, particularly in the synthesis of longer, biologically relevant oligonucleotide sequences, Letsinger's group have developed an

alternate strategy for the in-solution and solid support mediated synthesis of oligonucleotides via the *phosphotriester* approach.[15] The emergence of an insoluble polymer support enabled the synthesis of longer oligonucleotide sequences with higher yields and purity due to the ability to drive the coupling reactions to completion with excess soluble monomers and coupling reagents. The solid support mediated synthesis also provided a considerable advantage relative to the in-solution approach, in the purification of the oligonucleotides. The extensive 'work-up' and chromatography (or recrystallization) procedures were substituted with expedient "washing steps" for the removal of side products and unreacted reagents or monomers from the oligonucleotide bound support. However, limitations of this procedure were still apparent, particularly with the high reactivity of the coupling reagents which often produced irreversible and competing side reactions with the oligonucleotide phosphodiester backbone.

The most popular route today for the chemical synthesis of oligonucleotides was initiated by Letsinger[16] and later optimized by Beaucage and Caruthers[17]. The development of the *phosphite triester* approach was initiated with the use of a highly reactive nucleoside 3'-chlorophosphite derivative for coupling with the 5'-hydroxyl functional group of a nucleoside (or oligonucleotide) and subsequent *in-situ* oxidation to convert the highly reactive phosphite triester to the less reactive phosphate triester backbone. However, the practical utility of the reactive nucleoside 3'-chlorophosphites limited their applicability for the automated oligonucleotide synthesis on solid support. The development of the nucleoside 3'-*N,N*-dialkylphosphoramidite in the presence of a weak acid, such as tetrazole (pKa ~ 4.9) provided the more stable coupling reagents and conditions for the *in-situ* activation and quantitative coupling of the monomers to the nascent oligonucleotide. The coupling reactions were driven to completion on the order of minutes and without competing side-reactions. This fundamental *phosphate approach* was further exploited by Ogilvie and co-workers[18] at McGill University during the early 1980s for the early developments of an automated solid phase synthesis of DNA and RNA by virtue of the Bio Logicals® *gene machine*. Today, the automated phosphoramidite approach remains the method of choice for the small (microgram to milligram) and large (gram to kilogram) scale syntheses of DNA and RNA.
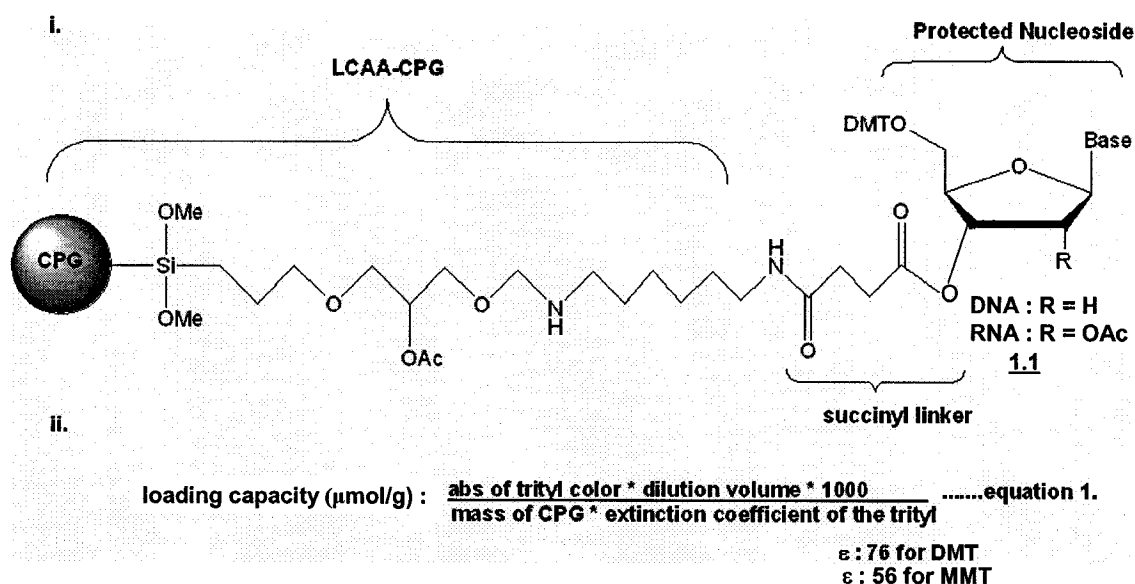
Figure 1.4: i. The structure of the succinyl linked long chain alkyl amino controlled pore glass derivatized with the 5'-MMT or DMT nucleoside. **ii.** Equation used to calculate the loading capacity (μmol/g) of the nucleoside attached to the solid support.

The automated solid phase oligonucleotide synthesis is performed on commercially available 5'-dimethoxytritylated nucleosides connected to the solid support through a base labile succinate ester. The solid support of choice is long chain alkylamine controlled pore glass (LCAA CPG), *i.e.* silica beads with pore size of 500 - 1000 Å[19] with a loading capacity in the order of 10-200 μmol (nucleoside) per gram of support. The loading is determined by treating a small quantity of **1.1** with an acidic (3% trichloroacetic acid:dichloromethane) solution to remove the 5'-dimethoxytrityl protecting group, and to relate the nucleoside loading on the solid support to the absorbance of the trityl cation by colorimetric analysis (**Figure 1.4, ii.**).

The phosphoramidite building blocks for the automated solid phase synthesis of DNA and RNA contain orthogonal protecting groups that are essential for the efficient chemo and regio-selective coupling, deprotection and chain lengthening reactions (**Figure 1.5**). These include: 1) the 5'-dimethoxytrityl, DMT, or the more stable 5'-monomethoxytrityl, MMT, groups, which are cleaved by acid to liberate the 5'-end hydroxyl group for coupling with additional DNA or RNA 3'-phosphoramidites[20], 2) the 2-cyanoethyl protecting groups on the non-bridging oxygen of the phosphate internucleotide linkage. These protecting groups are cleaved by base and prevent side reactions from occurring at

the non-bridging phosphorous oxygen during DNA or RNA chain assembly.[21] 3) The 2'-O-tert-butyldimethylsilyl, TBDMS, fluoride labile protecting groups in RNA phosphoramidites prevent chemical modifications from occurring at the 2'-hydroxyl groups in addition to chain isomerization and cleavage reactions with the acidic, neutral and basic conditions found in the oligoribonucleotide synthesis and deprotection steps.[22] They are only removed during the final step of the post synthesis 'work-up' procedure by treatment with fluoride reagents (i.e. 1M TBAF in THF or TREAT-HF).[24,257] 4) The exocyclic purine and pyrimidine base labile protecting groups are removed during the alkaline cleavage and deprotection steps after synthesis, and used to prevent side reactions at the exocyclic amino positions of the bases.[23] These include the isobutyryl, i-Bu, group on guanine and the benzoyl, Bz, group on adenine and cytosine and with the pyrimidines, thymine (DNA) and uracil (RNA) do not require base protection.

5'-O-Trityl Protection

$R_1 = R_2 = H$, -OTr
$R_1 = H$ and $R_2 = OCH_3$,   MMT
$R_1 = OCH_3$ and $R_2 = OCH_3$.   )DMT

Base Protecting Groups (Base P.G.)

2'- Protection
R : H, **DNA**
R : O-Si(Me)$_2$t-Bu or TBDMS, **RNA**

2 -cyanoethyl or -OCEt

N,N-diisopropylamine or · N(i-Pr)$_2$

Phosphoramidite Diastereomer
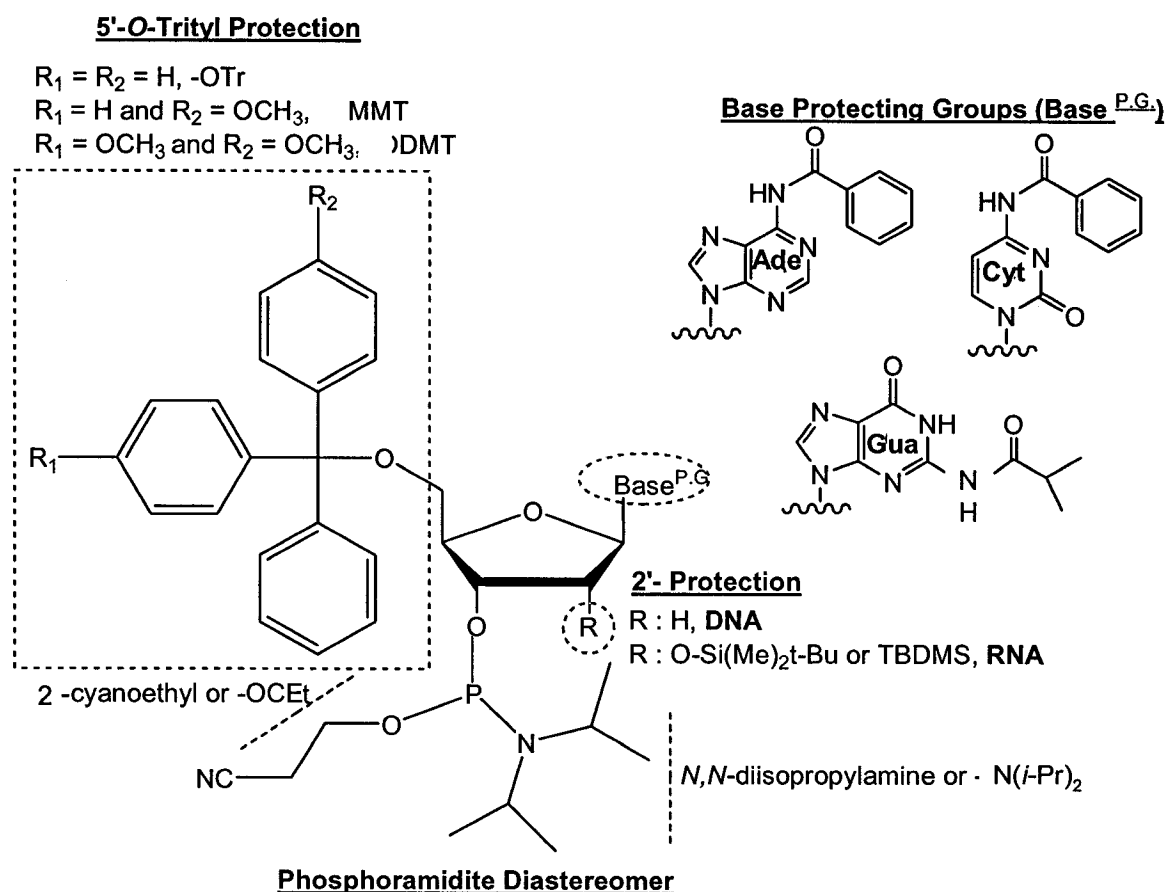
Figure 1.5: Structures of phosphoramidite diastereomers and their protecting groups for RNA and DNA automated solid phase synthesis.

The conventional automated solid phase synthesis cycle for the small scale synthesis (*e.g.* 1 μmol synthesis on a 500 Å CPG support) of DNA and RNA is illustrated in **Figure 1.6**.[24] Prior to the very first 5'-detritylation step, the CPG is treated with Cap A (1: 1:8 v/v/v acetic anhydride/ pyridine/ tetrahydrofuran) and Cap B (10% N-methyl imidazole in tetrahydrofuran) reagents to dry the solid support and to acetylate (or 'cap') unmasked functional groups on the CPG surface. In the case of RNA synthesis, it also serves to acetylate the 2'/3'-OH group of the nucleoside attached to the support (**Figure 1.4**). The detritylation step is then performed with a 120 – 200 second delivery of 3% trichloroacetic acid in dichloromethane (3% TCA:DCM), for the removal of the 5'–DMT (or MMT) protecting groups followed by a washing step with DCM to remove any excess acid prior to the coupling reaction. The DNA phosphoramidite monomers are dissolved as 0.1 M solutions in anhydrous acetonitrile, MeCN. These, in addition to an activator solution (*e.g.* 0.5 M tetrazole in MeCN) are injected into the column containing the nucleoside tethered solid support and coupled for 2 minutes. The RNA amidites are less reactive due to their bulky 2'-TBDMS protecting groups which hinders the coupling reaction between the 5'-hydroxyl group and the activated nucleoside 3'-phosphoramidite. Therefore, RNA amidites are added as a more concentrated (0.15 M) solution with the more reactive activators (*i.e.* 0.5 M DCI or 0.25 M ETT in MeCN) and coupled for an extended period of time (*i.e.* 10 minutes and G requires 15 minutes coupling due to the additional steric effects of the base N2-*i*-Bu protecting group). Since the stepwise coupling efficiencies for DNA and RNA synthesis proceed with >99% and 97-98%, respectively, the accumulation of "failure" sequences during assembly is minimal. Nevertheless, it must be ensured that these failure sequences do not continue to grow in subsequent coupling reactions, and are therefore 5'-acetylated by capping immediately after the coupling step. The support is then subjected to an oxidation procedure for 20 - 40 seconds with a 0.1 M oxidant solution (I_2 in 75/20/5 v/v/v tetrahydrofuran/ pyridine/ water) to convert the reactive phosphite triester to the more stable pentavalent phosphate triester backbone. The support is extensively washed and dried with MeCN and argon prior to detritylation of the terminal 5'-end trityl protecting group and the cycle repeated again for continued chain growth. The synthesized strand is released from the CPG support with a solution of 3:1 v/v ethanolic ammonium hydroxide (3:1 NH_4OH:EtOH).

This step also cleaves the phosphate and base protecting groups (12 - 16 hours at 55°C). After evaporation of this solution, the resulting crude oligomers are purified either by anion exchange (AE) or reverse phase (RP) HPLC, or polyacrylamide gel electrophoresis (PAGE). In the case of RNA oligmers, deprotection requires an additional treatment with fluoride to cleave the 2'-TBDMS protecting groups. After evaporation of triethylamine hydrofluoride (the fluoride source), the RNA oligomers are precipitated into cold butanol and then purified as describe above for DNA oligonucleotides.
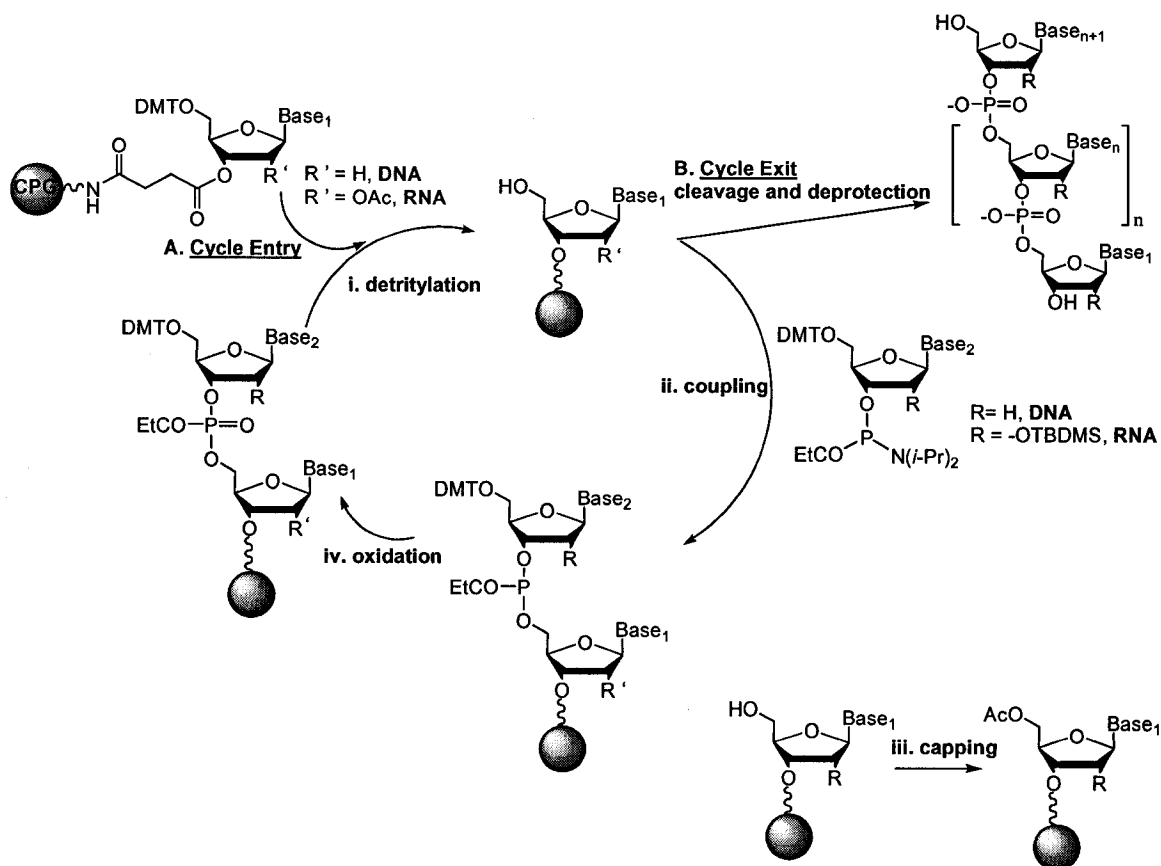
Figure 1.6: The automated solid phase synthesis cycle of oligonucleotides by the use of the phosphoramidite method.

1.4 SOLID-PHASE SYNTHESIS OF HIGHER ORDERED DNA AND RNA PRIMARY STRUCTURES: BRANCHED AND HYPERBRANCHED NUCLEIC ACIDS

The automated synthesis of DNA and RNA has provided a method for the iterative and stepwise assembly of nucleic acid building blocks into discrete higher ordered

macromolecular architectures with distinct length, size, shape and base sequence composition (**Figure 1.7**). This approach has facilitated the need for the synthesis of more complex biologically relevant DNA and RNA structures such as the branched[25] and intronic lariat RNA[26] motifs for their study as intermediates involved in the splicing mechanism of mRNA formation. More specifically, branched nucleic acids can also be used for the study of the biochemical processes associated with triple helix formation in branched oligonucleotides[27], as fiber-optic biosensors[28] and as well as substrates and inhibitors for the structure and mechanistic elucidation of the branchpoint sugar hydrolysis by the yeast and human debranching enzymes.[29] Alternatively, these automated solid phase synthesis methods have also generated discrete higher ordered macromolecular structures such as "hyperbranched" or "dendritic" nucleic acids[41,42] for their utility in biomaterials such as the generation of polylabeled biosensor probe materials[30] for their applications in the cellular detection of biomolecules.[31]
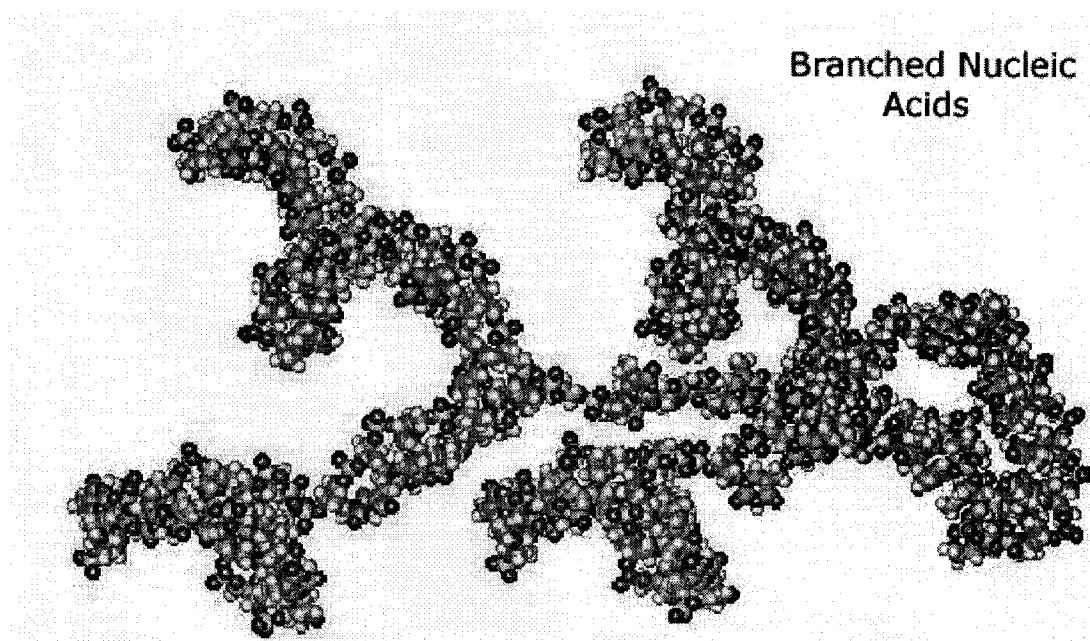
Figure 1.7: The schematic representation of hyperbranched nucleic acid as depicted by Hudson and Damha *J. Am. Chem. Soc.* **1993**, *115*, 2119.

The earliest reports on the synthesis of branched nucleic acids, bNAs, were performed in-solution and were of considerable challenge due to the requirement of orthogonal protecting groups for the regiospecific branchpoint coupling and deprotection reactions. These synthetic bNAs yielded a unique branchpoint 2',5' and 3',5' phosphodiester bond

within the structure.[32] The chain lengthening steps from the branchpoint also required mild conditions in order to prevent isomerization and cleavage of the phosphate by the vicinal hydroxyl groups with the neutral, acid or basic conditions found in the synthesis steps.[33] Vicinal phosphate triesters are labile under these conditions and were converted to the more resistant phosphate diester internucleotide linkages prior to the removal of the vicinal hydroxyl protecting groups.[34] Despite modification and optimization of the in-solution bNA synthesis protocol, these methods were found to be of considerable challenge, particularly in their limited applicability in synthesizing larger biologically relevant branched RNA structures.

Therefore, alternate methods such as the template directed chemical ligation synthesis of bNAs[26,35], enzymatic ligation with deoxyribozymes[36] and the solid phase synthesis of bNAs on solid support have been developed to address many of the limitations involved in the solution phase approach. The more conventional method for their synthesis, the solid phase methodology, can be performed by an extension of the in-solution *phosphite triester* methodology proposed by Damha and Ogilvie.[32a,b] The *convergent approach* utilizes a branchpoint 2',3'-bisphosphoramidite synthon capable of coupling with two adjacent and identical oligonucleotide strands at their terminal 5'-hydroxyl ends. This generates a V-like structure with vicinal 2',5' and 3',5'-phosphodiester linkages (**Figure 1.8 and Compound 1.2**). The branchpoint synthon is detritylated at the 5'-end and reacted with DNA or RNA phosphoramidites to yield the Y-shaped bNA bound to the CPG support.[25,37] A limitation of this *convergent approach* for the synthesis of bNAs prevents the formation of chemically diverse bNAs with asymmetric base sequence compositions. To fulfill this requirement, the *divergent approach* for the automated support bound synthesis of bNAs permits the regiospecific synthesis of mixed sequence bNAs[40], and more specifically the synthesis of prokaryotic and/or eukaryotic multicopy single-stranded DNA (msDNA).[38]

This unique bNA, msDNA, consists of a DNA and RNA chimera covalently linked at a branchpoint position, and has been shown to be implicated in the mechanism of RNA transcription and processing in bacteria.[39] The key feature in their synthesis is related to the use of a commercially available RNA branchpoint synthon phosphoramidite with chemically non-equivalent protecting groups at the 2' and 5'-branchpoint position. This

allows for separate and regiospecific chain extension, deprotection and branching reactions with DNA and RNA phosphosphoramidites (**Figure 1.9** and Compound **1.3**).[40]
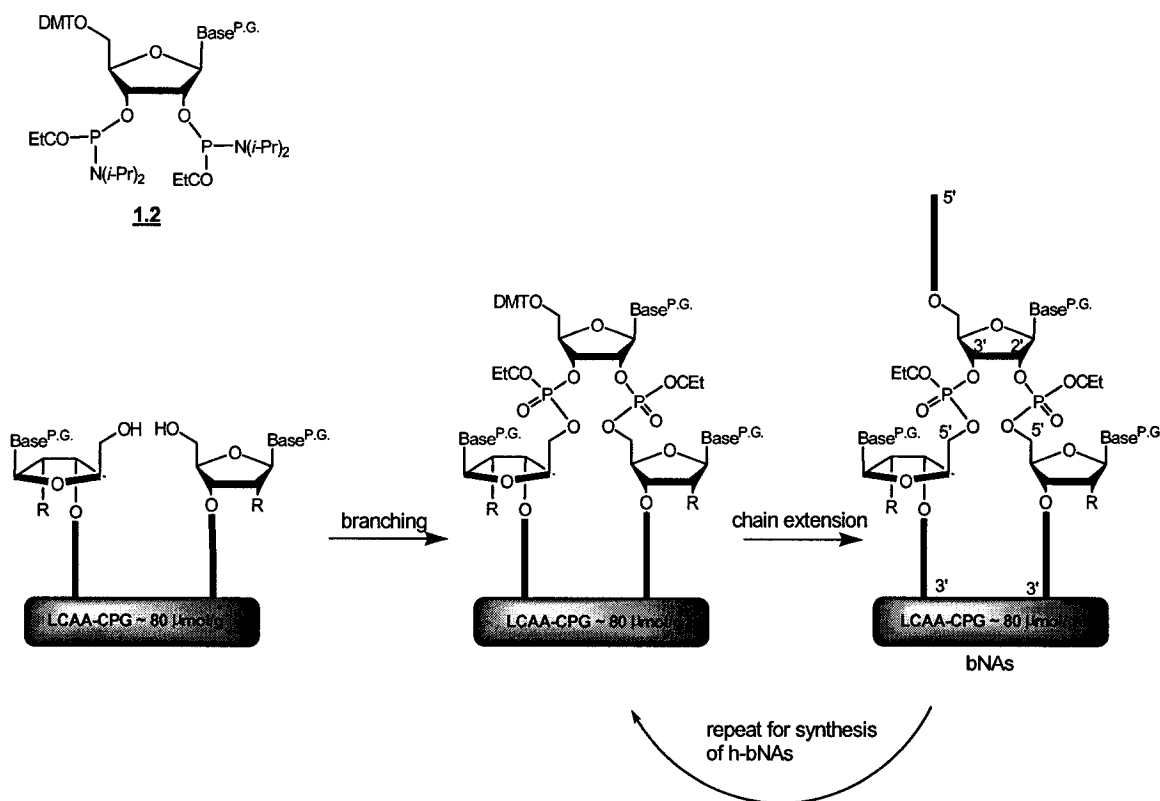
Figure 1.8: The convergent synthesis for symmetric branched and hyperbranched nucleic acids; where R = H, DNA or R = OTBDMS, RNA and Base$^{P.G.}$= Ura/Thy, Ade^{N-Bz}, Cyt^{N-Bz}, Gua^{iBu}. Compound **1.2** is a nucleoside bisphosphoramidite that couples two adjacent and identical oligonucleotides on the surface of LCAA CPG.

An extension to the synthesis of bNAs was also introduced by Damha and co-workers for the *convergent*[41] and *divergent*[42] solid phase synthesis of higher ordered bNAs, the hyperbranched, h-bNAs, or *dendritic* nucleic acids. Their synthesis was performed in a similar fashion to the previously described synthesis of bNAs, with the use of branchpoint synthons containing orthogonal protecting groups which allowed for the selective removal and iterative branching and chain lengthening steps involved in the multiply branched h-bNA structure.

14
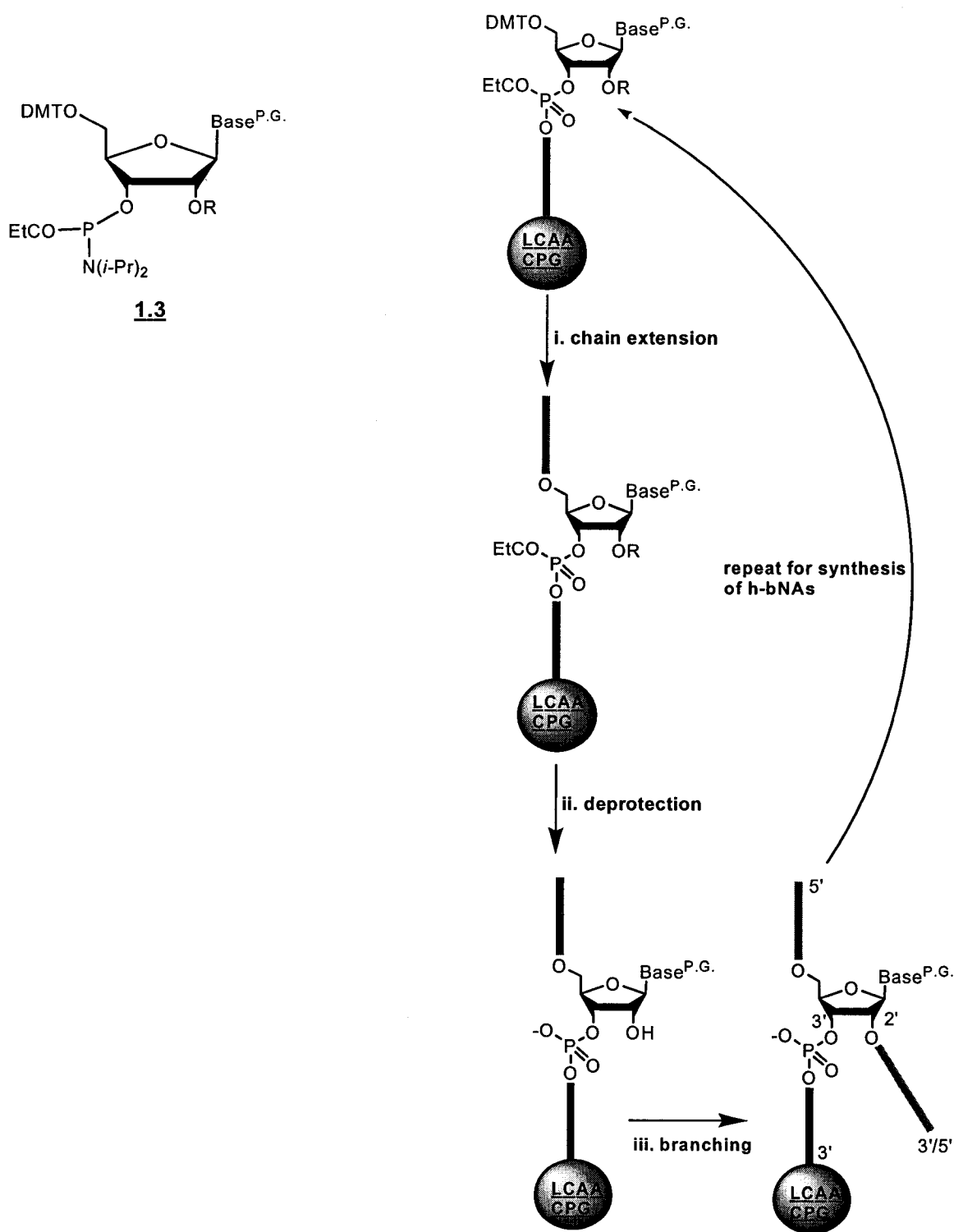
Figure 1.9: The divergent synthesis for asymmetric branched and hyperbranched nucleic acids; where R = TBDMS or 2'-*O*-[1-(2-fluorophenyl)-4-methoxypiperidin-4yl], 2'-Fpmp, RNA and Base$^{P.G.}$ = Ura/Thy, Ade^{N-Bz}, Cyt^{N-Bz}, Gua^{iBu}. Compound **1.3** illustrates the branchpoint synthon phosphoramidite diastereomer utilized for the divergent synthesis of bNAs and h-bNAs.

1.5 POTENTIAL FOR USING BRANCHED AND HYPERBRANCHED NUCLEIC ACIDS IN THE STUDY OF NUCLEAR pre-mRNA SPLICING: THE ENZYMATIC DEBRANCHING ACTIVITY

After the discovery and structure elucidation of the first branched nucleic acids, from the extract of nuclear polyadenylylated RNA in HeLa cells[43], much effort has been devoted towards understanding its role as a transient intermediate in the processing of mRNA. Their occurrence has been linked to a splicing process that excises intervening intron sequences at their 5'- and 3'-end splice sites followed by ligation of the two adjacent exons.[44] This ligation process produces an intronic lariat structure that contains unusual vicinal 2',5'- and 3',5'-phosphodiester linkages at the branchpoint, one of which (2',5') is hydrolyzed by the lariat debranching enzyme resulting in intron linearization (**Figure 1.10**).[45,46]

This unique 2',5'-phosphodiesterase activity was initially discovered in HeLa cells for a splicing derived lariat species. Interestingly, debranching did not occur when the lariat was isolated directly from the splicesome[47] indicating that a protein factor was protecting the 2',5'-phosphodiester bond against hydrolysis by the debranching enzyme. Subsequent isolation and purification of this cytosolic fraction in HeLa cells, demonstrated that the catalytic debranching activity was also associated and favored with divalent metal cations.[48] Furthermore, simple 'V-like' bNAs generated from the digestion of RNA lariats with nuclease P1 [*i.e.* $pA^{(2'p5X)}{}_{3'p5'}Y$] were also found to be substrates to the debranching activity. The discovery of this unique debranching activity prompted the gene encoding for the yeast (*Saccharomyces Cerevisiae*) lariat debranching enzyme which demonstrated 40% sequence homology to the human debranching enzyme.[49] The enzyme was found to be capable of digesting a variety of bNAs associated with intron lariats, multi-copy single-stranded msDNA, and the synthetic branched oligoribonucleotides.[50,51] Thus, this debranching activity has been shown to be responsible for the formation of intronic RNA elements which may have specific coding capabilities for other cellular proteins.[52] The bNAs and h-bNAs also provide tools for investigating the substrate requirements and inhibition properties of the debranching activity. Furthermore, the development of inhibitors against the yDBr1 enzyme also provides the potential for co-crystallization of the substrate and enzyme in order to better

understand and evaluate the debranching mechanism and activity associated with the DBr1 enzyme in the pre-mRNA splicing process.
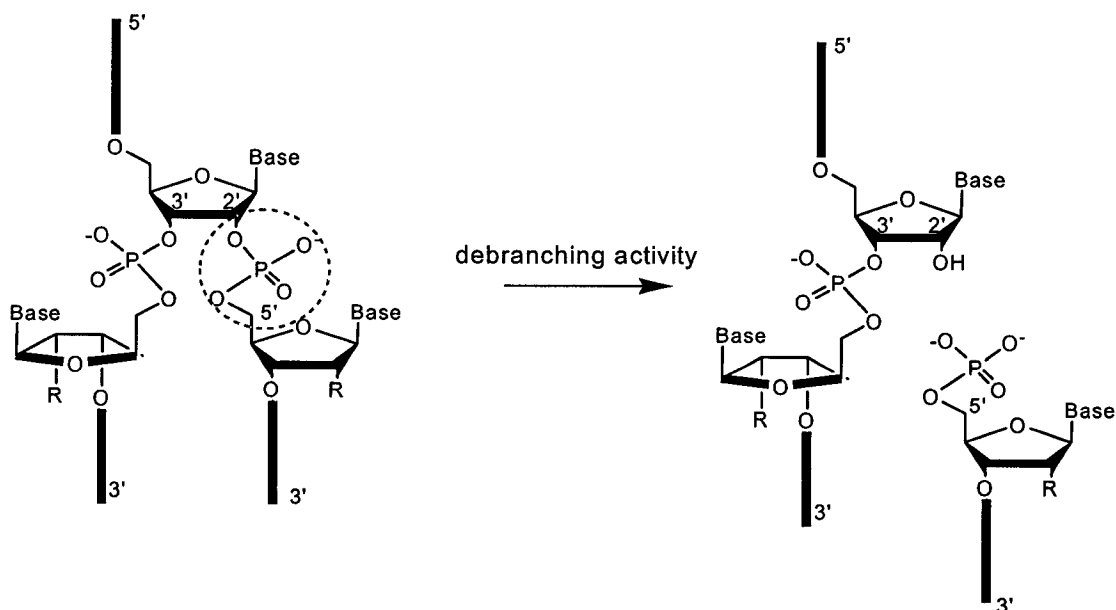
Figure 1.10: The debranching activity of the debranching enzyme is associated with the regiospecific 2',5'-phsophodiesterase activity at the branchpoint. The structure of the RNA branchpoint contains the Bases = Ura, Ade, Cyt, Gua with either DNA or RNA branching arms (R = H; DNA or R = OH; RNA).

1.6 REGULATION OF GENE EXPRESSION BY MODIFIED OLIGONUCLEOTIDE ANALOGUES: THE ANTISENSE, ANTIGENE AND RNA INTERFERENCE STRATEGIES

The expression of genetic information is carried out through a series of events linking transcription and translation of RNA to a functional protein. A key step during transcription is the formation of a DNA/RNA hybrid in which the single stranded DNA region (transcription DNA bubble) serves as a template for RNA synthesis. This is followed by post synthesis RNA modification procedures and the translation of mRNA into polypeptides and proteins. Moreover, the study of the DNA/RNA hybrid involved in the mRNA transcription process is key in developing antisense based therapeutics.[53,54]

When a mixed sequence DNA is hybridized to its RNA complement, the heteroduplex forms an intermediate right-handed duplex structure conformation between the DNA pure A or B-like helix conformation.[53] Alternatively, the RNA homopurine and DNA homopyrimidine hybrid resembles the A-like RNA duplex whereas a DNA homopurine

17

and RNA homopyrimidine hybrid resembles a B-DNA structure, an indication of the conformational flexibility of the nucleotide conformers within the duplex structure. This structural diversity is ultimately linked to the nucleotide sugar pucker conformation, which in mixed sequences adopts a C3'-endo geometry in the RNA sense strand and an intermediate O4'-endo pucker in the complementary DNA antisense strand.[54] The contraction in the DNA nucleotide structure (P-P nucleotide distance of *ca.* 6.5 Å at the monomer level) reduces the *minor groove* width in the DNA/RNA hybrid to 9 Å. This lies in between the canonical A and B-like hybrids and as a result, the structure has been proposed to adopt an AB-like or a Hybrid (H)-form geometry (**Figure 1.11**).[53,54] More unexpectedly though, the P-P distance in the DNA strand was found to be the same as that of B-DNA while the RNA strand consists of P-P distances that are typical of the A-type helix structure. These values suggest that this parameter is a more conservative property of the entire DNA strand and not the sugar pucker conformation.[55]
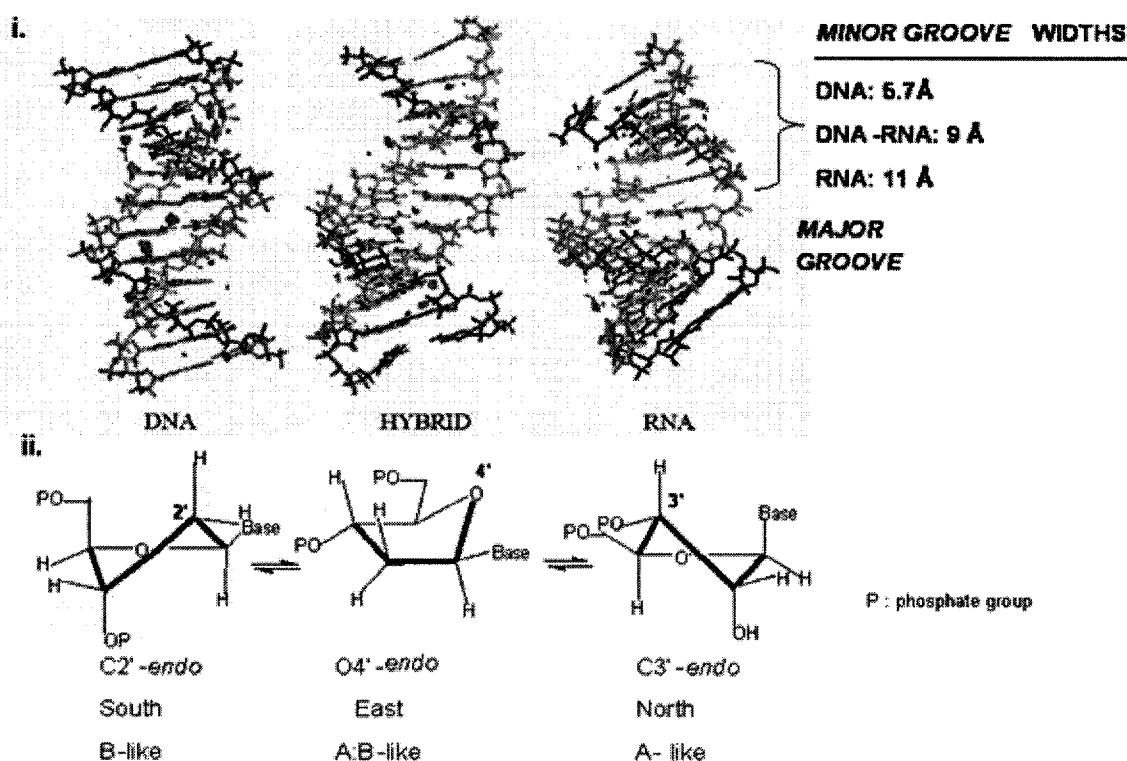
Figure 1.11: i. The energy minimized DNA, DNA/RNA and RNA hybrids. The purple indicates water solvation, which is most abundant in RNA due to the presence of the 2'-OH group.[55] **ii.** The energy favored nucleotide conformations in the antisense oligonucleotide.

18

1.6.1 Antisense Strategy

The antisense strategy was first put to practice by Zamecnik and Stephenson during their seminal 1978 study demonstrating that a synthetic unmodified DNA oligonucleotide complementary to a viral mRNA target inhibited the *Rous Sarcoma* virus *in-vitro*.[56] This mRNA-*targeting* approach is currently being considered for the treatment of genetic diseases that are commonly associated with: i) mutated and irreparable genes formed during synthesis and that translate into malfunctioning proteins (*i.e.* inherited diseases such as cystic fibrosis)[57], ii) genes that are over-expressed and result in deleterious cell growth (*i.e.* tumor cell-line growth)[60a], iii) genes that are under-expressed and hinder synthesis of functional proteins that are vital to the cell (*i.e.* insulin level in patients suffering from diabetes)[58], and in iv) genes that are expressed by viruses and bacteria that are harmful to the host[57,59]. Therefore, rather than targeting the function of harmful protein products, the antisense approach targets its precursor, *e.g.* mRNA, bypassing traditional medicinal chemistry efforts (*e.g.*, high-throughput screening of compound libraries) that search for small molecule inhibitors against the target protein.

The antisense strategy[60] is based on a chemically modified antisense oligonucleotide (usually of 18-21 nucleotide base length) capable of penetrating the cellular membrane and to form a duplex with the specific sense mRNA strand in the cytosol of the cell. This assumes that the chemically modified antisense oligonucleotide is capable of penetrating the cellular membrane and form a duplex with the specific sense RNA strand within the cell. The resulting "antisense:sense" duplex sterically hinders the ribosomal complex assembly and translation of mRNA. This mechanism is referred to as 'translation arrest'. An alternative and more effective mechanism of action is the recruitment of a ubiquitous RNaseH enzyme that can recognize the canonical DNA/RNA hybrid and selectively degrade the RNA segment within the duplex (**Figure 1.12**). The antisense strand is thus released and further recycled for continued mRNA binding and RNaseH mediated RNA degradation. However, the following criteria must be satisfied before antisense can be routinely used, particularly for *in vivo* applications, i) a method for efficient cellular delivery across the plasma membrane, ii) an increase in nuclease resistance within the cell, iii) the regiospecific and stable RNA sense strand targeting and binding, iv) efficient

recruitment of RNaseH activity and v) tolerable pharmacokinetics and metabolism of the antisense oligonucleotide (*i.e.* no cytotoxicity).
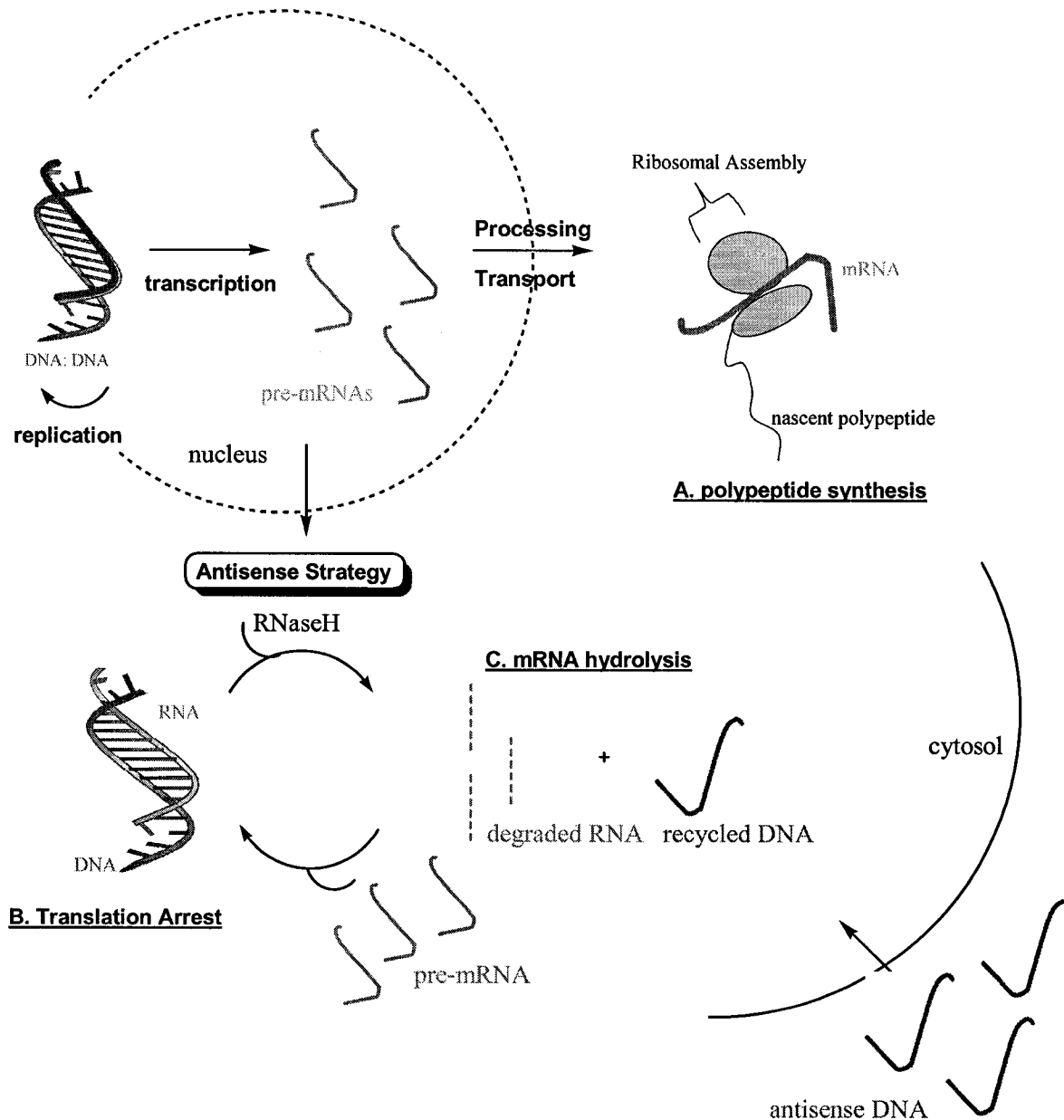
Figure 1.12: The consequences of RNA maturation and processing in eukaryotic cells. **A.** Normal undisrupted polypeptide synthesis and protein translation from the ribosomal subunits. **B.** Antisense Strategy involving translation arrest or steric inhibition of polypeptide synthesis and **C.** Antisense Strategy involving translation arrest followed by RNaseH mediated degradation of the mRNA sequence.

Chemical modifications of the DNA nucleotide core generally consists of alterations to the i) deoxyribose pentofuranosyl sugar unit, ii) the heterocyclic nitrogenous purine or pyrimidine base, iii) modifications to the phosphodiester core and iv) 5' or 3'-end modifications of the oligonucleotide strand termini.

A combination of the in-solution and solid phase synthesis methods is necessary to convert the modified oligonucleotide, MON, analogues into potent antisense drugs.

Chemical modifications to the pentofuranosyl carbohydrate skeleton have generated some of the more promising antisense analogues. The challenge in re-constructing the DNA carbohydrate skeleton is linked to generating a nucleotide conformation that is similar to the DNA sugar conformation found in the DNA/RNA hybrid, such that the duplex is stable and capable of eliciting RNaseH activity.

Arabinonucleic acids. Arabinonucleic acids (ANA), the 2'-epimer of RNA, were first exploited as antisense agents by our reseach group. The inversion of configuration at the C2'-position switches the sugar pucker from the North C3'-*endo* RNA geometry to the South/East (C2'-*endo* / O4'-*endo*) conformation while providing both chemical and nuclease resistance within the oligonucleotide structure (**Figure 1.13**).[61] The inversion in stereochemistry in oligoarabinonucleotides also led to a decrease in binding affinity or cross-paring to the complementary native DNA or RNA sequences (ΔT_m = -0.5– 2.1°C/modification relative to DNA) as the β-hydroxyl group of the arabinofuranose sugar was found to project into the major groove of the helix and thereby disrupting base-pairing.[62] Interestingly, the CD spectral signatures denoted similar hybrid conformations for the DNA/RNA and ANA/RNA duplexes, such that the ANA/RNA hybrid was capable of supporting RNaseH mediated degradation of the sense RNA strand.[62,63] Despite some of the remarkable features of ANA as antisense constructs, the replacement of the β-hydroxyl group with a smaller, electronegative 2'-fluorine atom provided the 2'-fluoroarabino nucleic acids, 2'-F-ANA, which had excellent affinity towards both DNA and RNA complements (ΔT_m = + 0.5-1°C/modification relative to DNA).[64] The sugars of 2'-F-ANA adopt the Eastern O4'-*endo* conformation, also prevalent in the DNA antisense strand of the native DNA/RNA hybrids. As a result, 2'-F-ANA oligonucleotides, like DNA, was found to elicit RNaseH activity when bound to complementary RNA.[62,64]

2'-O-alkyl modifications: 2'-O-Me RNA. Most C2'-modifications of the ribofuranosyl sugar which project an electronegative α-C2'-substituent within the minor groove of the MON/RNA duplex function to stabilize the A-RNA type conformation. Some of the more common modifications include 2'-*O*-methyl (2'-OMe)[65], and the 2'-ribofluoro, (2'-riboF)[66], modifications which generally function to stabilize the antisense oligonucleotide from chemical or nuclease degradation. In fact, 2'-F-RNA/RNA and 2'-OMe RNA/RNA hybrids have greater thermodynamic stability than RNA/RNA duplexes. Although these modifications do not elicit RNaseH activity, they can compete with the natural substrate for binding at the active site and slow the rate of hydrolysis.[62,65,66] A widely used antisense construct, referred to as "gapmer", consists of two 2'-OMe RNA segments (2-6 nt) flanking an internal DNA sequence (10-15 nt). The 2'-OMe RNA segments provide nuclease resistance and good RNA binding affinity, whereas the internal DNA strand serves to promote RNaseH hydrolysis when it hybridizes to the target RNA strand.[62]

Ribo- configuration
X = OH; RNA
X = F; 2'-F-RNA
X = OMe; 2'-OMe RNA

Arabino- configuration
X = OH; ANA
X = F; 2'-F-ANA
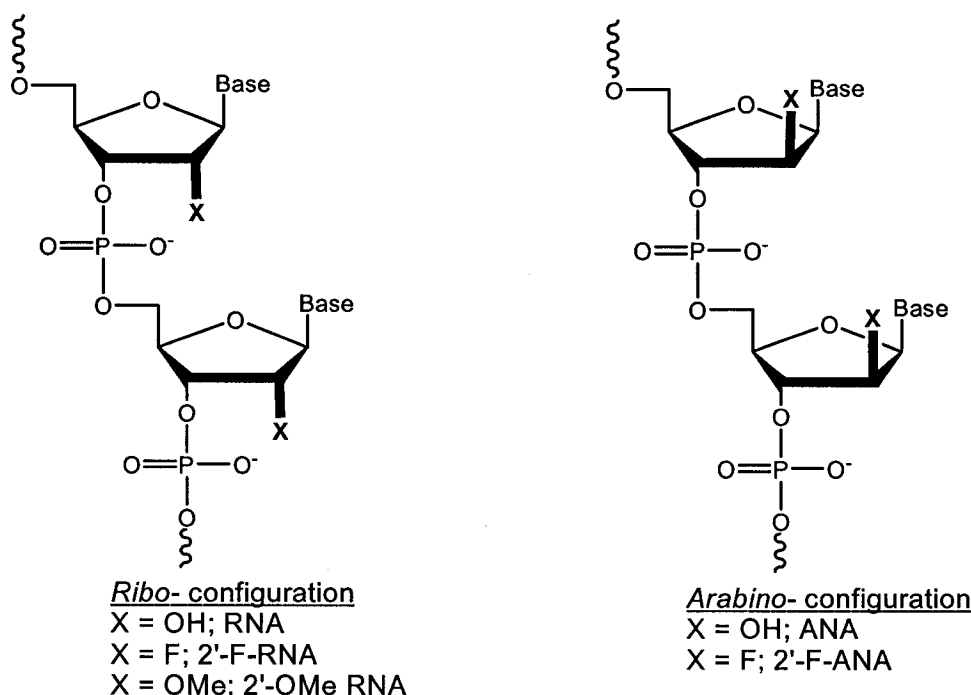
Figure 1.13: The comparison of the C2'-epimers of the ribofuranosyl sugar unit in the ribose configuration (RNA, 2'-F-RNA and 2'-OMe RNA) with the arabino configuration (ANA and 2'-F-ANA).

Locked Nucleic Acids. The 'locked' nucleic acids, LNAs, have been designed to entropically favor binding to their target strand. The strategic bridging connectivity

within the nucleoside sugar conformation has been shown to restrict or 'freeze' the sugar pucker into an RNA (C3'-*endo*)[67] conformation, resulting in enhanced binding to target RNA. Interestingly, it has also been possible to lock the sugar in the O4'-*endo* geometry, to afford oligomers (*i.e.,* arabinose LNA) that in principle should elicit RNase H (**Figure 1.14- A** and **B**). Surprisingly, bicyclo ANAs were unable to promote RNase H mediated hydrolysis of RNA, indicating that these compounds exhibit a high energy barrier for enzymatic hydrolysis by sterically preventing binding or manipulation of the rigid bicycle/RNA substrate once bound to RNase H.[67a]

A. East-Locked *Arabino* LNA

B.South-Locked *Carbocyclic* LNA   **C.North-Locked *β-D-ribose* LNA**
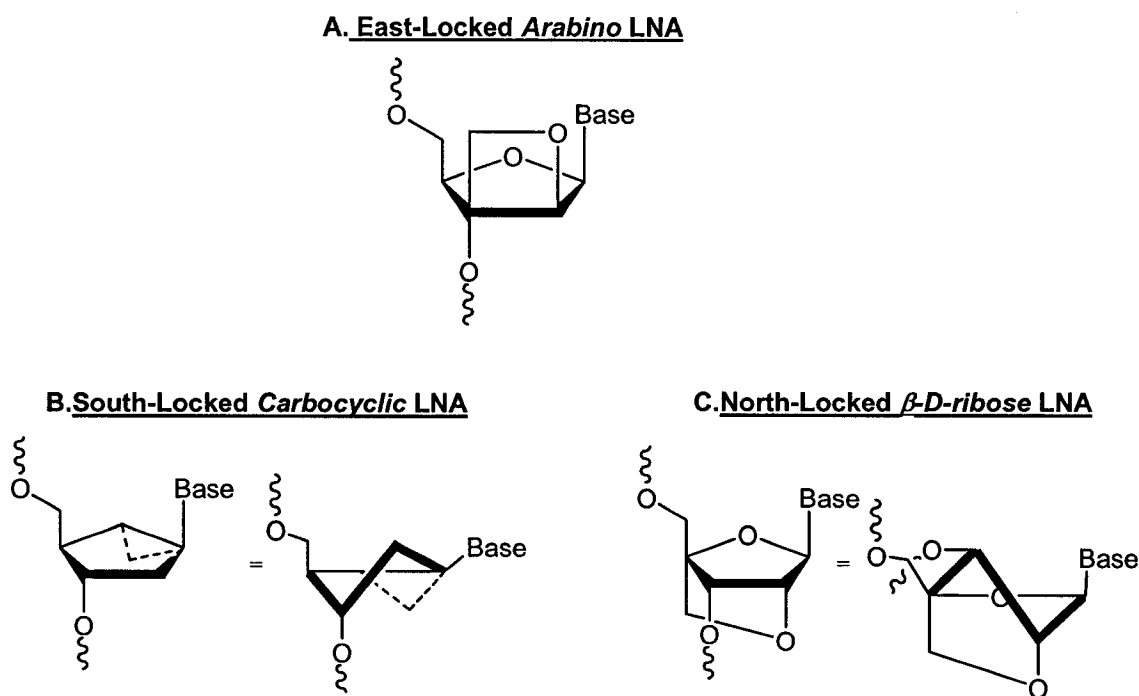
Figure 1.14: The structures of the conformationally restricted LNAs that are capable of adopting the nucleoside sugar pucker found in the naturally occurring DNA and RNA.

Acyclic Nucleic Acids. The construction of acyclic nucleic acid analogues such as the *seco* nucleic acids, SNAs[68], peptide nucleic acids, PNAs[69], glycerol nucleic acids[70], or glycol nucleic acids, GNAs[71] have also provided comparable binding and conformational properties relative to the native DNA and RNA complements. In addition to their antisense properties, these modifications have served as a link into the evolution of nucleic acids structure and function.[72] Also, a requirement for RNaseH recognition and

23

activation is based on a flexible sugar phosphate backbone, capable of adjustments within the active site of the enzyme for optimal substrate binding.[67a,68] Antisense DNA and 2'-F-ANA have been modified with internal acyclic seconucleotide residues resulting in chimeras with interspersed glycosyl or alkyl 'flexible-linker' moieties which provided enhanced RNaseH activity.[68]
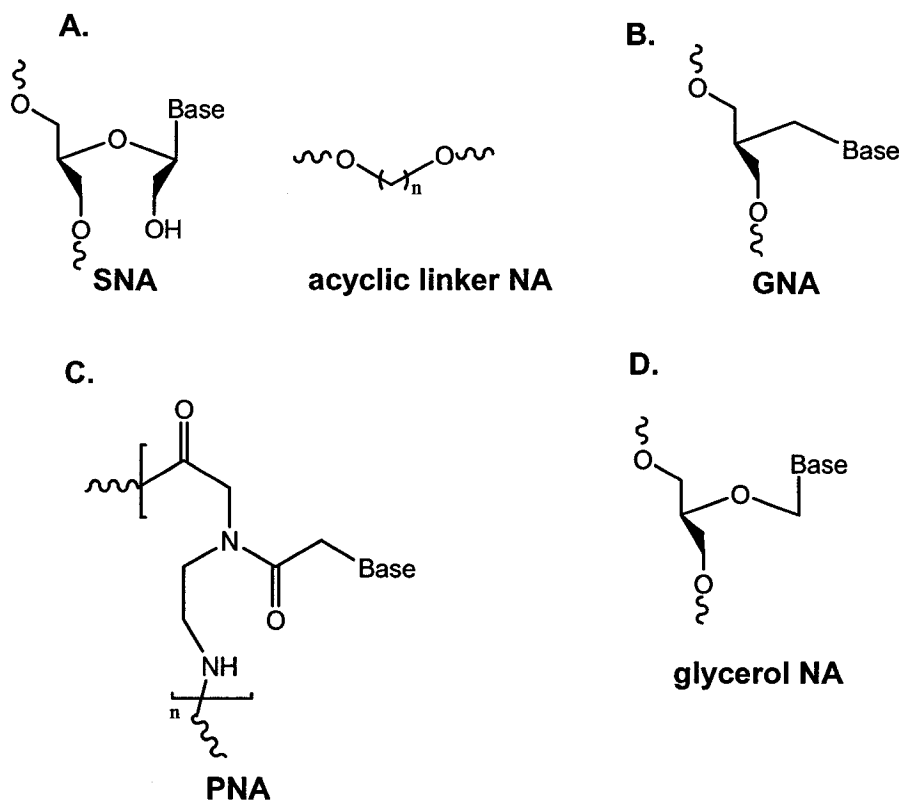
Figure 1.15: The structures of the acyclic nucleic acid analogues: **A.** seco-nucleic acids, SNAs and acyclic linkers bridging adjacent nucleic acid residues, **B.** glycol nucleic acids, GNAs, **C.** peptide nucleic acids, PNA and **D.** glycerol nucleic acids.

Threose and Pyranose Nucleic Acids. Eschenmoser and Herdewijn have systematically studied the influence of the carbohydrate ring size on nucleic acid hybridization.[73,75,76] To date, the largest expansion of the sugar moiety in DNA is that reported with the six-membered pyranose carbohydrate (**Figure 1.16- A.**). For example, the expansion of the furanose to the six-membered pyranose carbohydrate provided an oligonucleotide structure based on a β-(6',4')-hexopyranosyl oligonucleotide *i.e.* homo-DNA.[75] This conformationally restricted pyranose carbohydrate adopts a rigid chair-like conformation[74] with a modified oligonucleotide capable of only pairing to its pyranose

complement and not cross-pairing with the single stranded DNA or RNA complement.[75,76] The hexitol nucleic acids, HNAs,[77] adopt an A-like helix conformation when bound to complementary RNA without eliciting RNaseH activity. Since then, conformationally less restricted unsaturated 6-membered ring nucleic acids, e.g. the cyclohexene nucleic acids, CeNA, have been developed to improve both binding to DNA and RNA while eliciting RNaseH activity (**Figure 1.16- B.**).[78,79] Alternatively, reducing the carbohydrate to a four carbon skeleton, i.e. an α-L-threofuranosyl (TNA) sugar, maintains a furanose half-chair sugar pucker conformation similar to the native nucleic acids. Consequently, TNA is capable of pairing to its TNA complement in addition to cross-pairing with single stranded DNA and RNA.[80,81] TNA can also serve as a template for DNA synthesis and conversely, TNA synthesis can occur on a DNA template.[82] As a result, TNA has been considered a link between RNA and DNA in the evolution of nucleic acids.[72]

Additional modifications within antisense oligonucleotides have been implemented at the phosphodiester backbone.[83] The DNA phosphorothioates (PS-DNA) contain a substituted sulfur atom at one of the non-bridging oxygen atoms of the 3′,5′-phosphodiester linkages. These antisense oligonucleotides provide nuclease resistance and cell-permeability without compromising RNaseH activity.[84] The drawbacks of the PS-DNA include a reduced binding affinity towards mRNA and non-specific binding to serum and cellular proteins.[85] Despite these disadvantages, the first phosphorothioate antisense drug, Vitravene,[TM] was approved by the Food and Drug Administration in 1998 for the treatment of human cytomegalovirus induced retinitis in patients suffering from AIDS.[86]

Modifications to the heterocyclic bases are less popular due to their importance in retaining base-pair fidelity within the hybrid duplex structure. They also require additional sugar phsophate backbone modifications for increased nuclease resistance. However, Kool and co-workers have shown that the addition of ring expanded bases can also hybridize and retain the helical conformation when hybridized to DNA and RNA. These have potential applications in biotechnology due to the fluorescence properties of the size expanded bases.[87]

A.

TNA   DNA   homo- DNA

B.

<u>1st generation pyranose DNA</u>

- no binding to RNA/DNA
- no RNaseH activity

HNA   CeNA

<u>2nd generation pyranose DNA</u>   <u>3rd generation pyranose DNA</u>

- binding to RNA/DNA   - binding to RNA/DNA
- no RNaseH activity   - RNaseH activity
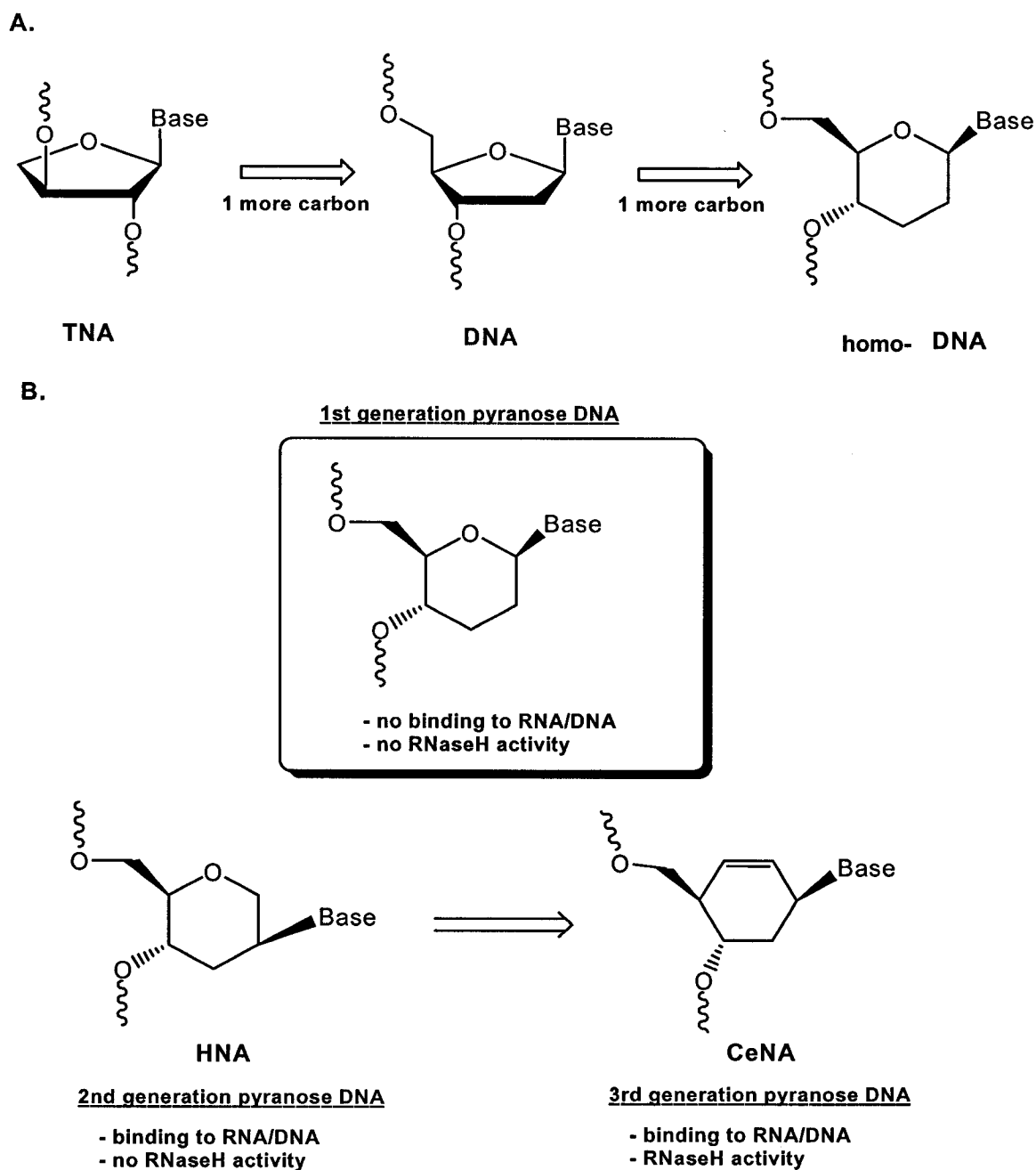
Figure 1.16: The carbohydrate ring modified nucleic acids. **A.** The conversion of the nucleic acid structure from TNA to DNA to 6-membered pyranose nucleic acids. **B.** Structures and properties of 6-membered ring nucleic acids.

Certain modifications to the 5' or 3'-end of the antisense oligonucleotides enhance cellular permeability. These modifications include lipophilic compounds that can compensate the hydrophilic polyanionic phosphate diester backbone of the

oligonucleotide, favoring their transport across the amphiphilic cellular membrane (Figure 1.17).[88]
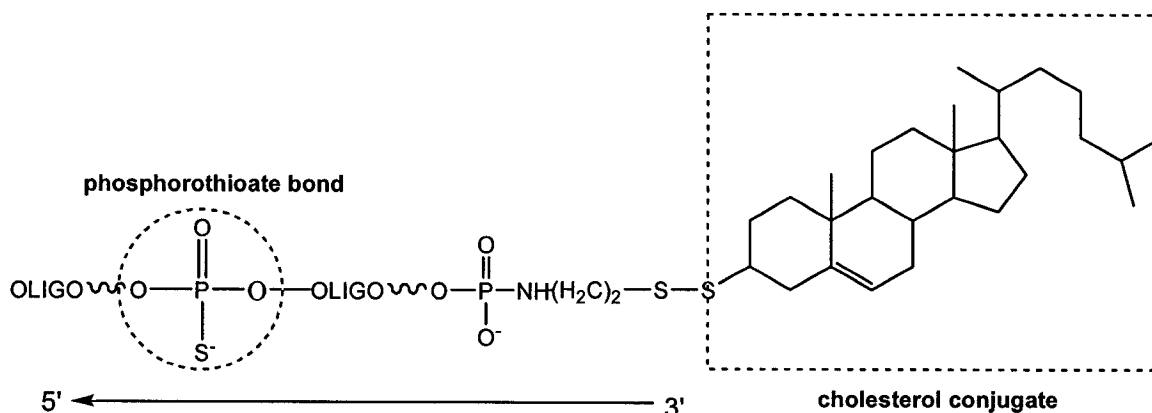
Figure 1.17: The structure of a phosphorothioate antisense oligonucleotide tethered to a 3'-cholesterol conjugate.

1.6.2 Triple Helix formation and the Antigene Strategy

Gene expression can also be controlled by directly targeting the DNA helix. In this approach an "antigene" oligonucleotide binds with a high specificity of recognition to the major groove of the DNA helix by forming Hoogsteen type bonds with purine bases of the Watson-Crick base pairs, resulting in triple helix (*i.e.* triplex) formation. Although the potential target sequences were originally restricted to polypurine-polypyrimidine sequences, considerable efforts have been devoted towards developing base modified oligonucleotides that recognize duplexes of mixed base composition (Figure 1.18-A).[90,91] Gene suppression via triplex formation is a much more difficult method of inhibition (relative to the antisense strategy) due to their difficulty in formation and with the reduced stability of the triplex structures.[89]

Triple helical RNA structures were initially observed in 1957 by Rich and co-workers, in which homopyrimidines associated with homopurines in a 2:1 stoichiometry to yield triple-stranded structures that were stabilized in the presence of divalent metal counterions.[90,91] The first class of the triplex binding patterns identified is referred to as the "pyrimidine" motif, which consists of a third polypyrimidine (*i.e.* thymine/uracil or protonated cytosine) complementary strand binding via Hoogsteen interactions with either an A:T[92] or a G:C[93] base pairs, respectively.
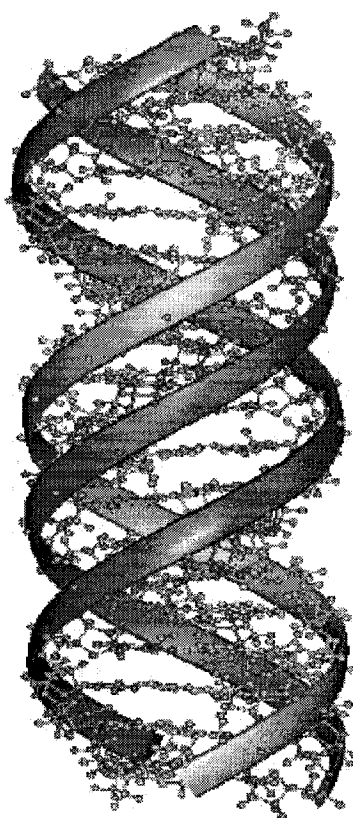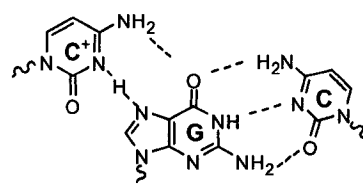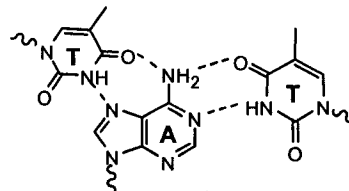
27
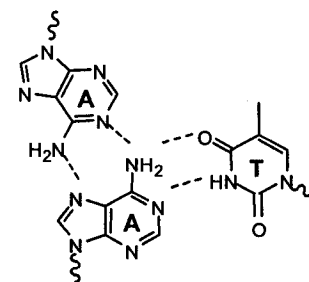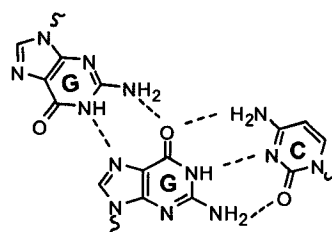
Figure 1.18: The antigene approach: **A.** Triplex DNA Structure **B.** Purine and Pyrimidine motifs in triplex structure.

These pyrimidine rich strands hybridize via Hoogsteen H-bonding, (·), parallel to the duplex homopurine strand to yield triplex structures such as T·A:T and C⁺·C:G **(Figure 1.18- B)**. The pyrimidine rich strands which contain cytosines must be protonated, C⁺, at the N3-position of the base with acidic conditions (pH: 5.5 - 6.5) in order for this binding motif to occur, thereby making their applications at physiological conditions limited. The

second class of triplex structures is referred to as the "purine" motif, in which a third polypurine strand pairs via reverse Hoogsteen interactions to either an A:T or G:C base pair (**Figure 1.18- B**).[94] These purine rich strands have the tendency to hybridize with reverse Hoogsteen H-bonding, (*), antiparallel to the duplex homopurine strand to yield triple helices in the form of A*A:T or G*G:C (**Figure 1.18- B**). Unlike the pyrimidine triple helix motif, the purine motif is stable at physiological pH but with limited applications due to their requirement of high concentrations of metal counterions such as K^+. Therefore, current strategies involving modified triplex forming oligonucleotides (TFOs) have provided stable triplex structures for antigene applications at physiological conditions.[95]

1.6.3 RNA interference Strategy

RNA interference (RNAi) is one of several post-transcriptional defense mechanisms that silences the expression of mRNA.[96] Within the cells, double-stranded RNA, (dsRNA) triggers a gene silencing enzyme, *dicer,* which initially cleaves dsRNA (*ca.* > 150 base pairs) into 21 - 23 nt duplexes called short interfering RNA (siRNA), by virtue of an RNase III-like endonuclease activity. These fragments associate with a large protein assembly or RNA-induced silencing complex, RISC, which functions to unwind the siRNA for the target specific recruitment and binding with mRNA. The complex harbors a specific catalytic activity, *slicer,* which selectively hydrolyzes the mRNA strand without affecting the guide siRNA strand within RISC (**Figure 1.19**).[97] Therefore, gene suppression via RNAi differs from the antisense strategy by virtue of the enzymatic cleavage activity (*i.e.* RISC factor in RNAi and the RNaseH enzyme in the antisense strategy) in addition to the target substrate (*i.e.* dsRNA in RNAi and DNA/RNA in antisense). Moreover, the RNAi methods require important functional criteria for the siRNAs, and these include: i) 5'-phosphorylation at the end of the siRNA fragments for recognition by RISC, ii) the presence of a 5'-A/U rich region, iii) the additional 2 or 3 nucleotide overhangs at either the 5' or 3'-ends of the siRNAs and iv) antisense and sense strand oligonucleotides that are capable of forming siRNAs with an entirely RNA, A-like helix conformation. The initial applications of siRNAs were limited because the unmodified RNAs had modest cellular delivery, permeability and higher (relative to

29

antisense oligonucleotides) susceptibility towards 3'-endonuclease degradation. Therefore, the development of nuclease resistant and cell-permeable MONs capable of assuming the hybrid RNA C3'-*endo* conformations are better suited for applications of the siRNAs within the RNAi pathway.[98] However, there still remain only a few exceptions of completely modified siRNAs (*i.e.* 2'-F-RNA, LNA, 2'-*O*-alkyl RNA and phosphorothioate RNA can impair activity[99]) and current efforts are currently being aimed at conquering some of these limitations.[100]
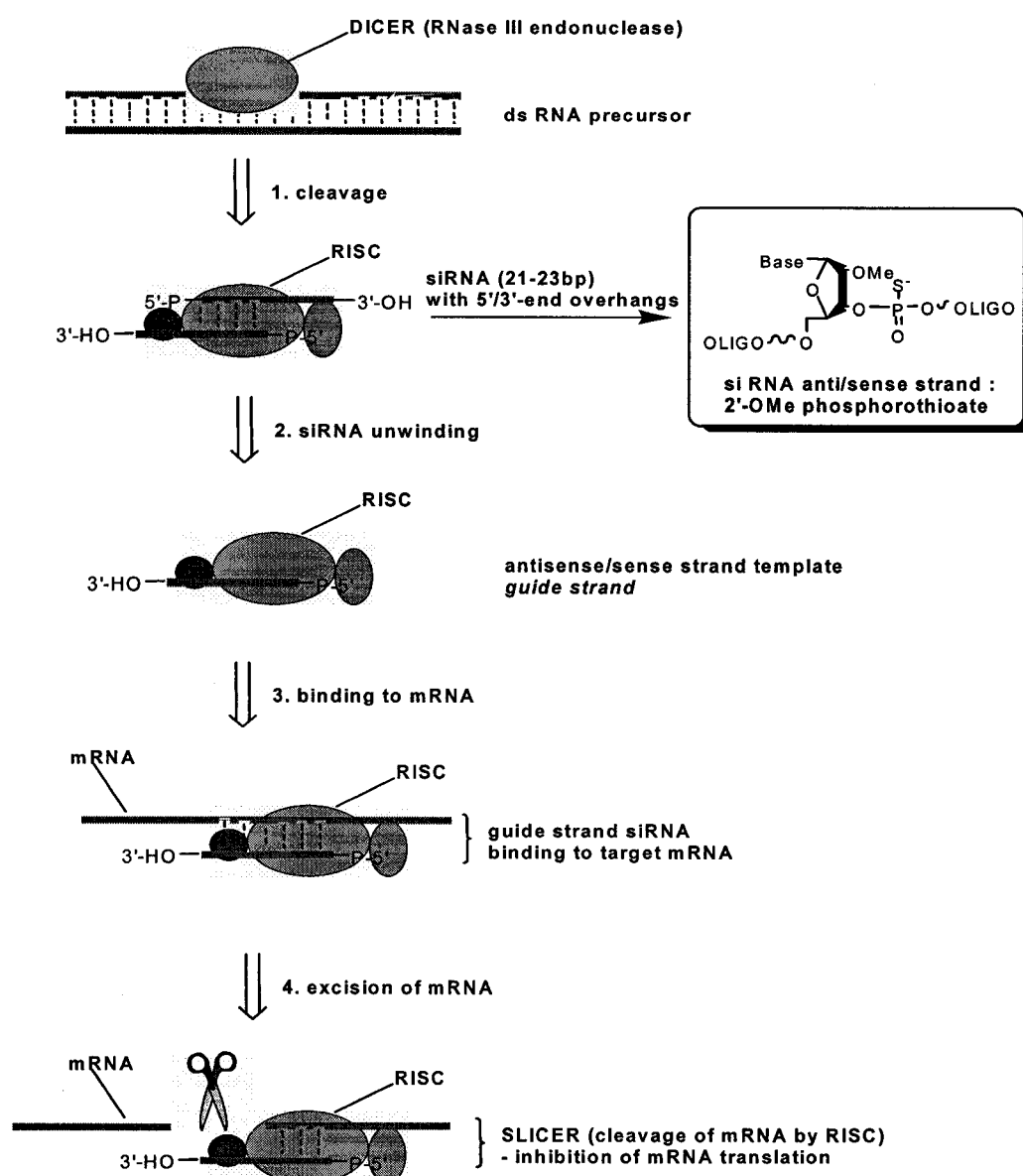
Figure 1.19: The RNAi pathway. A 2'-OMe phosphorothioate modified siRNA is recognized and loaded into the RISC factor, before recruitment and hydrolysis of mRNA.

1.7 THESIS OBJECTIVES

The creation of synthetic oligonucleotides is dependent on an efficient solid-phase synthesis methodology. The advent of the RNA interference (RNAi) and its application to therapeutics has created an urgent and growing need for the synthesis of large quantities of native and chemically modified short interfering RNA (siRNA) for animal and human studies.[101] The conventional monomers and others recently introduced for the automated synthesis of RNA have provided a facile and efficient route for their synthesis, however, they are limited to small and medium size scale applications due to their cost, coupling efficiencies and the requirement of excess quantities of reagent.[16,17,18,24] Chapter 2 of this thesis evaluates 2'-O-levulinyl (Lv) and 2'-O-monomethoxytrityl (MMT) ribonucleosides, as possible synthons for the synthesis of RNA and branched RNA (bRNA). The latter are key intermediates implicated in the splicing of pre-mRNA in eukaryotic cells. A key feature of the RNA and bRNA synthesis is the removal of 2'-protecting groups while the oligomer is still attached to the solid support. Conditions are developed in Chapter 3 which prevent isomerization or cleavage of the internucleotide linkages of nascent RNA and bRNA strands. Chapter 3 also examines, for the first time, the chemical stability of 3'-5'-internucleotide phosphate triesters (and diesters) adjacent to a 2'-hydroxyl group of a ribose oligomer attached to a solid support. These studies are not only relevant to the proper assembly of linear and branched RNA species, but also in the stability and properties of phosphate diesters and triesters as in the case of an unusual branched RNA termed "RNA-X".[140] Chapter 4 examines the use of 2'-O-Lv and 2'-O-MMT ribonucleoside monomers as building blocks for the assembly of a series of branched nucleic acid species (bRNA, bDNA, msDNA and hyperbranched or "dendritic" DNA/RNA) with discrete length, base composition and structure. These structures are synthesized via an iterative *divergent-growth* strategy, which facilitates the regioselective branching, deblocking and chain lengthening steps from a branchpoint core. These structures served as useful materials (bio-probes) as will be demonstrated by the biological studies performed with *E. coli* RNase H and the yeast lariat RNA debranching enzyme (yDBr1). These studies not only identified novel branched nucleic acid inhibitors of yDBR1 and RNase H, but also provided new insights about the substrate specificity of

these important enzymes. Chapter 5 will describe the synthesis of a new nucleic acid form, the so-called "oxepane nucleic acids" (ONAs), in which the pentofuranose ring of DNA and RNA is replaced with a 7-membered heptose sugar ring. The oxepane adenine and thymine monomers were prepared from the ring expansion reaction of a cyclopropanated glucose derivative and their conversion into phosphoramidite derivatives allowed efficient assembly of pentadecanucleotide ONAs on a solid support. ONA are much more resistant towards nuclease degradation than natural DNA, an important feature if these analogues are to be used in biological media. As it will be shown, ONAs exhibit cross-pairing with complementary RNA and elicit *E. coli* RNase H mediated degradation of the RNA strand. These finding are significant because oligonucleotide-directed RNase H degradation of the target RNA is a key determinant for the gene-specific inhibitory potency of antisense oligonucleotides. When comparing the rates of RNase H-mediated degradation induced by DNA, 2'-ene-pyranose and oxepane nucleic acids, the following trend is observed: DNA > 2'-ene-pyranose NA > ONA. The implications of these results will be discussed in the context of our current understanding of the catalytic mechanism of the enzyme, particularly with regard to the required flexibility of the oligonucleotide strands that binds to the RNA target. We will demonstrate that ONAs are useful tools for biological studies and their properties provide new insights into the structure/function of natural and alternative genetic systems.

CHAPTER 2: DEVELOPMENT OF NOVEL RIBONUCLEOSIDE SYNTHONS FOR THE SOLID-PHASE SYNTHESIS OF RNA

2.1 INTRODUCTION

The conventional solid phase synthesis cycle for oligoribonucleotides by the phosphoramidite approach has been outlined in Chapter 1 (**Figure 1.6**).[24] Although the method is ideal for the small to medium scale synthesis of DNA and RNA, some limitations for the large scale synthesis of RNA still exist. More specifically in the advent of synthesizing gram to kilogram scale quantities of RNA and DNA for therapeutic applications aimed towards antisense[57,58,59,60] and RNAi[98,99,100] based strategies, a more efficient and cost effective procedure is highly desirable for their bulk production.

Although some of the alternative synthetic strategies based on the in-solution synthesis[16,17] and enzymatic[102] ligation and polymerization reactions have been used for the synthesis of DNA and RNA, these strategies are of limited commercial applicability.[103] A more amenable method for the large scale commercial synthesis is associated with the use of orthogonal protecting groups for the regiospecific chemical synthesis of oligoribonucleotides.[104] This procedure is based on a block condensation approach that couples oligoribonucleotide segments with a 3'-phosphate and a 5'-hydroxyl functional group.[105] This method is also limited by the multiple chain lengthening steps, potentially undesirable side-reactions and tedious, iterative purification procedures.

The automated solid phase synthesis procedure still remains the current method of choice for the commercial scale synthesis of DNA and RNA in milligram, gram and kilogram scale quantities.[24,106] The reaction can be tuned to larger μmol or mmol scale syntheses[107] on a high density functionalized CPG support[108] and with the use of an optimized automated synthesis cycle which complements high phosphoramidite coupling efficiencies with lower required amounts (close to stoichiometric quantities) of reagents.[24] The solid phase synthesis strategy is particularly advantageous for the automated synthesis of medium sized length oligomers (*i.e.* 18 to 23 base sequence length is optimal for antisense and RNAi applications) in which the stepwise failure sequence accumulation is minimal (<5%) allowing for efficient and practical purification

procedures. However, there remain certain disadvantages to the solid phase synthesis approach, specifically in the large scale production of RNA.[109] These drawbacks include: 1) expense of the solid support, 2) expense of the phosphoramidite building block synthons (particularly for chemically modified monomers), 3) careful optimization of the synthesis cycle to yield a 'DNA-like' synthesis efficiency, *i.e.* stepwise coupling yields > 99%, and lastly 4) the requirement for faster and more efficient deprotection and purification steps for the recovery of the purified oligoribonucleotides.

2.2 PROJECT OBJECTIVES

The search for readily available protecting groups that favor efficient amidite coupling yields, facile orthogonal deprotection and purification is key in developing synthetic RNA. A suitable 2'-*O*-protecting group must be resistant during the assembly of oligoribonucleotides and throughout the post-synthesis work-up procedures. Their removal with acid or alkaline conditions can lead to isomerization and/or cleavage of the oligoribonucleotide.[22]

The 5',2'-di-*O*-monomethoxytrityluridine (5',2'-diMMT rU) 3'-*O*-phosphoramidite, **2.5**, and 2'-*O*-levulinyl-5'-*O*-monomethoxytrityluridine 3'-*O*-phosphoramidite **2.10**, (**Figure 2.1**) are examples of alternative monomers that we considered for the automated solid-phase synthesis of RNA and branched RNA. These building block synthons (**2.5** and **2.10**) will be employed to compare the compatibility of parallel (diMMT) and orthogonal (MMT/Lv) protecting groups with the automated solid phase synthesis of oligouridylic acid. A key requirement for the protecting groups in the RNA synthesis procedure is related to their regiospecific, chemoselective and quick deprotection conditions. This becomes very important for the diMMT derivative **2.5** where the 5'-MMT group needs to be selectively removed in the presence of the 2'-MMT group for the synthesis of linear 3'-5' RNA. If this were not possible, we envisaged that monomers such as **2.5** could be used for the *divergent-growth* synthesis of branched RNA sequences (see Chapter 4). Likewise, the MMT/Lv monomers such as **2.10** would also be tested as potential synthons for the synthesis of RNA and branched oligonucleotides. Preliminary studies directed towards the attainment of these goals are presented below.
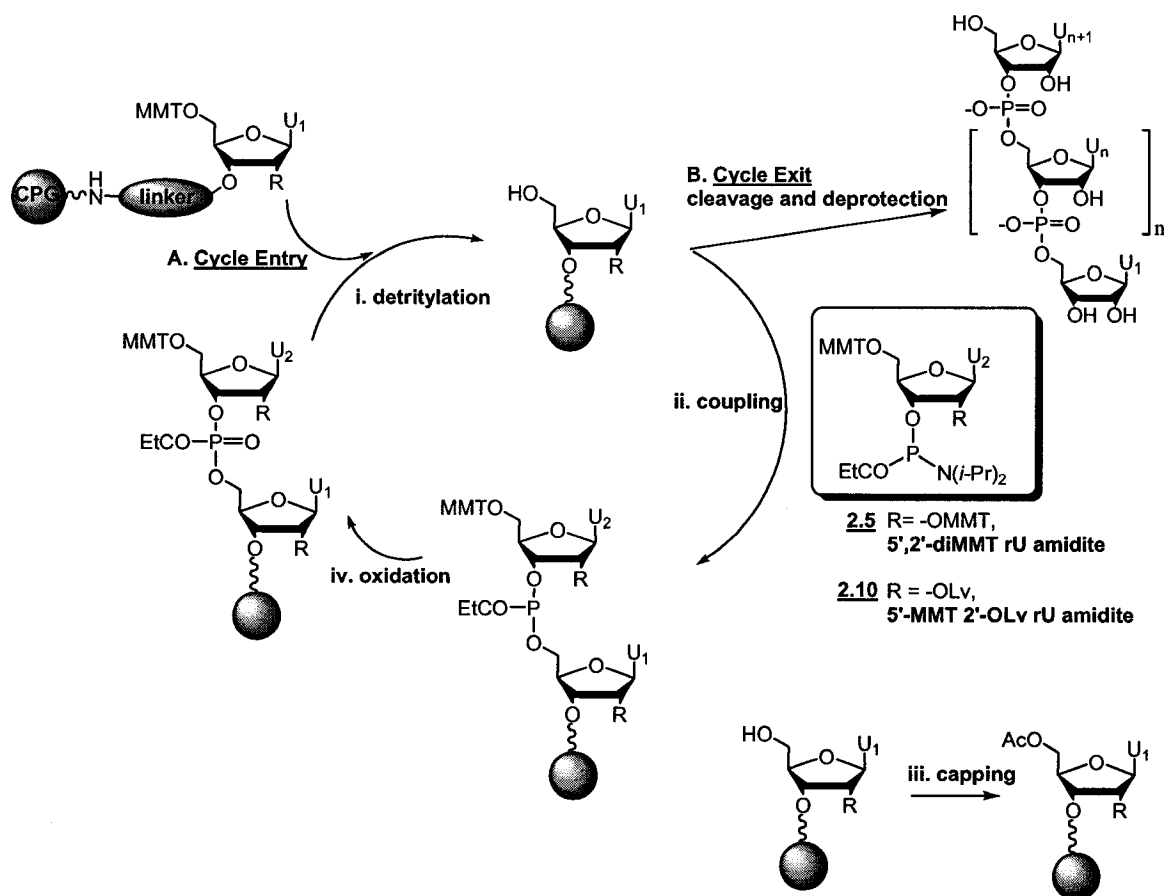
Figure 2.1: Solid phase synthesis cycle for the assembly of oligouridylate sequences through the use of 5',2'-diMMT rU amidite **2.5** and 5'-MMT 2'-OLv rU amidite **2.10.**

2.3 SYNTHESIS OF 5',2' AND 5',3'-diMMT rU PHOSPHORAMIDITES

Ditritylation of uridine (**2.1**) was performed by a slight modification of the MMT-Cl/AgNO₃ procedure described by Ogilvie and co-workers.[110] The reaction progress was firstly monitored by TLC with 5% MeOH in CH₂Cl₂ as eluent system and produced separable regioisomer products with R_f: 0.44 and 0.33 for the 2',5' and 3',5'-diMMT rU nucleosides, respectively. The less polar product, **2.2**, was easily separated from regioisomer **2.3**, and 5'-MMT rU **2.4** (R_f: 0.18) by column chromatography with an eluent gradient of 1 to 3% MeOH in CH₂Cl₂. Under optimal conditions the reaction afforded the following yields: 5',2'-diMMT rU **2.2** (50%), 5',3'-diMMT rU **2.3** (35%) and 5'-MMT rU **2.4** (10%). The diMMT rU regioisomers were characterized by ¹H NMR as described below.[110] The regioisomers were subsequently phosphitylated to yield the corresponding 2'-*O* and 3'-*O*-phosphoramidite derivatives in 66% isolated yields (**Figure**

35

2.2). In either case the formation of the desired products proceeded with the concomitant formation of $Cl^{-+}HNEt(i\text{-}Pr)_2$ which precipitated from the reaction mixture. The phosphoramidites were purified by flash chromatography and the purity was confirmed by ^1H-decoupled ^{31}P NMR. The 2'-O-phosphoramidite regioisomer **2.6** consisted of a mixture of a pair of diastereoisomers (phosphorus epimers; R_f: 0.65, 0.55), and was shown to be less polar than the corresponding 3'-O-phosphoramidite regioisomer **2.5** (phosphorus epimers; R_f: 0.59, 0.51) (TLC 1:1 EtOAc:hexanes). The diMMT rU phosphoramidite derivatives were characterized by NMR and ESI-MS as described below (section 2.3.1).
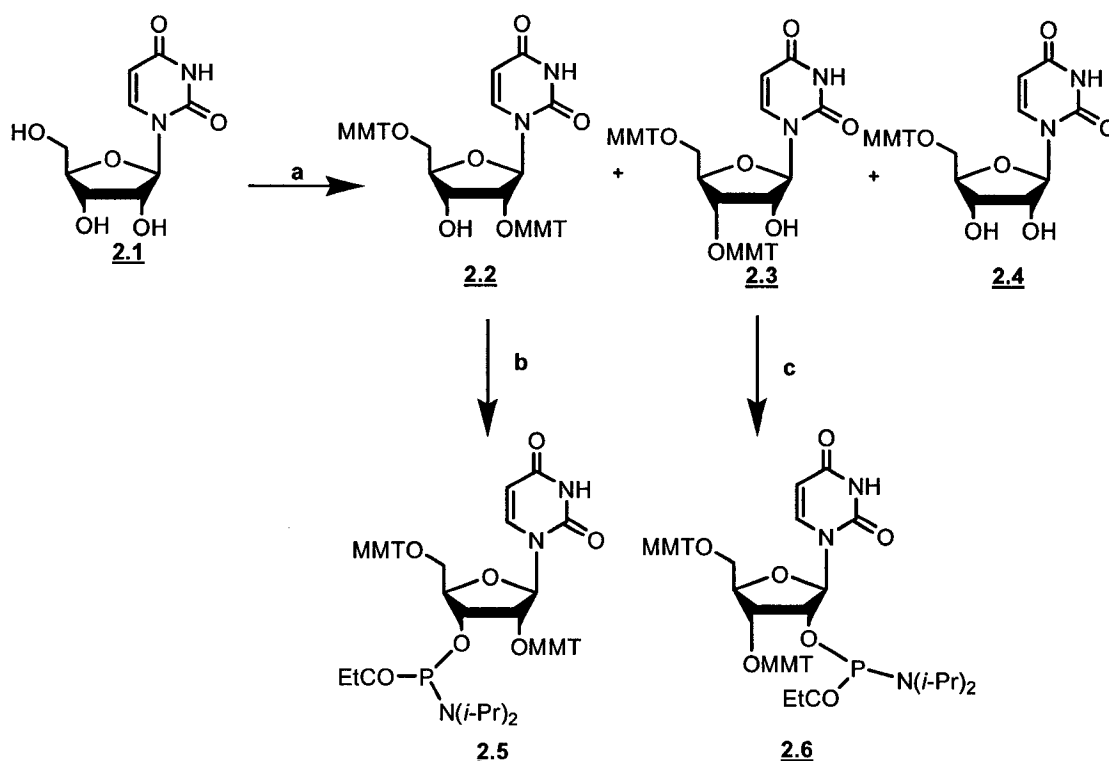
Figure 2.2: Reagents and Conditions: a. 2eq. MMT-Cl, 2eq. AgNO3, 2:1 pyr:THF, r.t., 24 h (**2.2**: 50%, **2.3**: 35%, **2.4**:10% yield); b. Cl-P(OCEt)N(i-Pr)2, EtN(i-Pr)2, THF, r.t. 3h, 66%; c. Cl-P(OCEt)N(i-Pr)2, EtN(i-Pr)2, THF, r.t., 3h, 66%.

2.3.1 Characterization of diMMT rU nucleosides and phosphoramidites

Compounds **2.2-2.6** were characterized by 1-D, 2-D NMR and MS techniques. ^1H-^1H COSY experiments in combination with deuterium mixing experiments with D2O-DMSO-d_6 as solvent were performed to unambiguously assign the chemical shifts of the sugar and hydroxyl protons in **2.2** and **2.3** (Figure 2.3). The presence of a free 3'-OH

(assigned by its coupling to H3') confirmed the positions of the MMT groups in nucleoside **2.2** at 5' and 2' positions. The identification of the 5',3'-diMMT rU isomer was carried out in the same manner (**Figure 2.4- A** and **B**). The proton NMR data of **2.2** and **2.3** were in full agreement with the literature data.[110] The ^{31}P-NMR spectra of **2.5** and **2.6** each consisted of a pair of signals at *ca.* 150 ppm (pairs of P-epimers), with little if any H-phosphonate by-products.[24,111,112] ESI-MS provided confirmation of the molecular weights of the sodium adducts of **2.5** and **2.6** ($C_{49}H_{44}N_2O_8Na$ calc. 812 found 812).
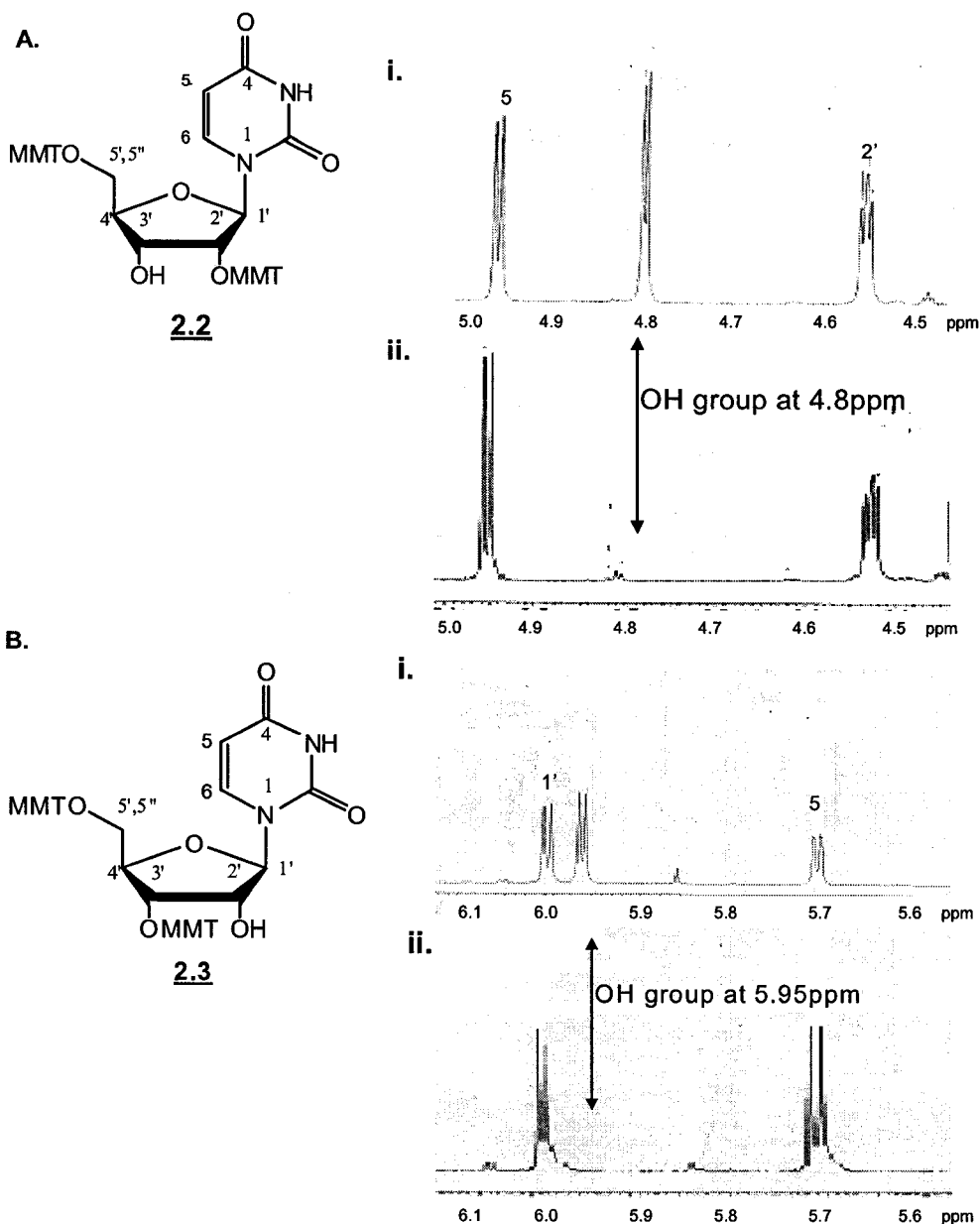
Figure 2.3: The location of the sugar OH group in **A. 2.2** and **B. 2.3** was determined unequivocally by ^{1}H NMR experiments in **i.** DMSO-d_6 and **ii.** with D_2O mixing.
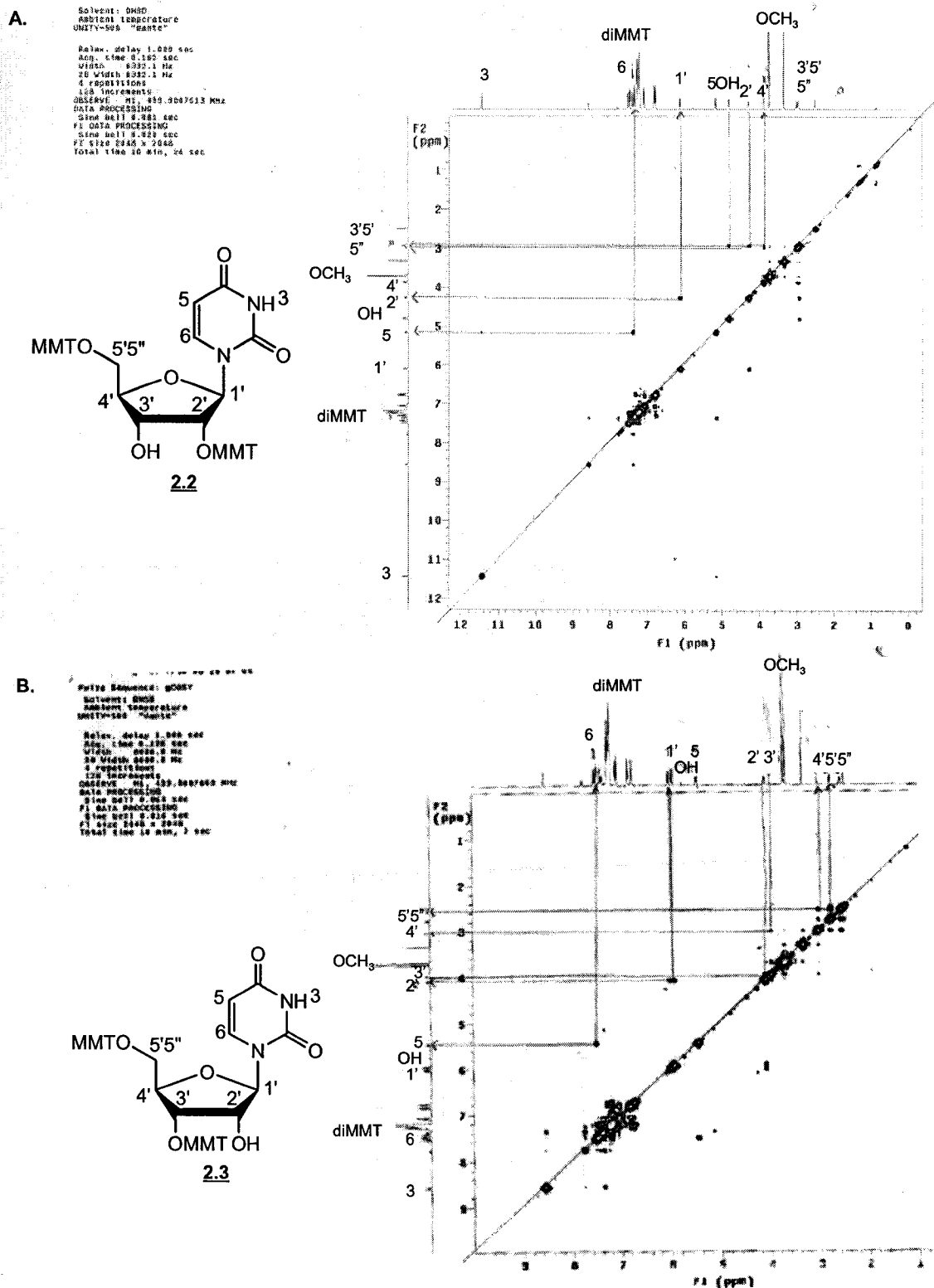
Figure 2.4: 2D ^{1}H-^{1}H COSY NMR spectrum for **A. 2.2** and **B. 2.3** in DMSO-d_6 at 500 MHz. The coupling between the OH and the sugar protons unambiguously identify their location on the furanose ring.

38

2.4 ATTEMPTED REGIOSELECTIVE 5'-DETRITYLATION OF diMMT rU NUCLEOSIDES 2.2 and 2.3

Discerning detritylation conditions are necessary for the removal of the 5'-MMT group during the solid phase RNA synthesis cycle using monomer **2.5** (**Figure 2.1**). Acid detritylation conditions have been developed to selectively remove the 5'-trityl protecting groups in the presence of other acid sensitive functional groups, including 3'-trityl groups.[113]

Chelating reagents such as $ZnBr_2$ and $TiCl_4$, have also been employed for the regioselective 5'-detritylation reaction of 5',3'-O-bis-DMT 2'-deoxynucleosides.[114] Selectivity is achieved through bidentate chelation of the metal to O5' and O4' leading to selective cleavage at the C5'-O5' bond and formation of the trityl cation. The rate of acid catalyzed depurination or depyrimidation under these conditions is also significantly slower. However, $ZnBr_2$ is mostly insoluble in organic solvents (*i.e.* CH_2Cl_2, MeOH, CH_3NO_2) limiting its utility in automated solid phase RNA synthesis. Kierzek and co-workers have remedied some of these limitations by dissolving $ZnBr_2$ with an anhydrous polar protic solvent system comprised of 85:15 (v/v) CH_2Cl_2:*i*-PrOH.[114b] This alternative solvent system retained the selective 5'-detritylation properties of $ZnBr_2$ while making it suitable for solid-phase DNA synthesis.

A small sample of each nucleoside, **2.1**, **2.2**, **2.3** and **2.4** was first treated with 3% TCA in DCM, and their detritylation monitored by TLC (5% MeOH in CH_2Cl_2). The analysis of the reaction progress by TLC indicated a non-selective detritylation reaction that proceeded via the mono-MMT intermediates to afford unprotected uridine (**2.1**). These results were also confirmed by colorimetric analysis of the MMT$^+$ in CH_2Cl_2 after each detritylation reaction. The absorbance measurements at 478 nm indicated a doubling in the absorbance values for **2.2** and **2.3** (removal of 2x MMT groups) relative to the detritylation reaction with the 5'-MMT ribouridine nucleoside, **2.4**.

Next, the regioselective detritylation of **2.2** and **2.3** with 1 M $ZnBr_2$ in CH_2Cl_2:*i*-PrOH (85:15; v/v) was attempted. Each detritylation reaction was carried out for 1-4 minutes and monitored by TLC with the eluent system 5% MeOH in CH_2Cl_2. Under these conditions, selective detritylation of **2.2** and **2.3** occurred within 2 minutes to afford a monotritylated product consistent with compound **2.4** (R_f: 0.18). As above, the

monodetritylation was confirmed by colorimetric analysis of the MMT$^+$ released during the reaction. Prolonged exposure of compounds **2.2** and **2.3** to ZnBr$_2$ led to complete detritylation, likely resulting from moisture (and release of HBr) during the reaction.[114a]

2.5 SOLID-PHASE SYNTHESIS OF OLIGOURIDYLATE SEQUENCES THROUGH THE USE OF 5',3'-diMMT rU PHOSPHORAMIDITE, 2.5

Short oligouridylate sequences (*i.e.* rU$_4$-4mers) were synthesized on the automated ABI-381A gene machine synthesizer according the conventional 1-µmol RNA phosphoramidite synthesis cycle.[24] The phosphoramidite, **2.5**, was dissolved as a 0.15 M solution (typical for RNA phosphoramidites) in anhydrous CH$_2$Cl$_2$ for enhanced solubility (**2.5** is mostly insoluble in MeCN, the standard solvent). The synthesis was performed on a uridine tethered to the support (**1.1**), with a loading capacity of 45 µmol/g, prepared according to the literature procedure.[19] The uridine phosphoramidite, **2.5**, was coupled with the activator solution (0.25 M ETT in MeCN) during an extended coupling time of 30 minutes. The use of shorter coupling times prevented complete coupling of the monomers and this was confirmed by colorimetric analysis of the MMT$^+$ collected after each stepwise detritylation step (data not shown). This poor coupling reactivity is likely due to the steric influence of the bulky 2'-MMT protecting group vicinal to the activated 3'-phosphoramidite moiety. Coupling was followed by the capping and oxidation steps and the UpU dimer bound to the support was detritylated (2 minutes) with an anhydrous solution of 1 M ZnBr$_2$ in 85:15 (v/v) CH$_2$Cl$_2$:*i*PrOH. The oligouridylate sequence (rU$_4$) was grown to completion and subjected to an 'on-column' removal of the phosphate triester 2-cyanoethyl group for 90 minutes with 4:6 (v/v) NEt$_3$:MeCN at room temperature conditions (22°C) followed by extensive washing with THF and MeCN.[40b,c] This step functions to selectively generate a more stable phosphodiester backbone in the RNA strand. This is important as RNA phosphate triesters, and to a lesser extent diesters are prone to acid catalyzed isomerization and hydrolysis reactions.[33,34] The rU$_4$ bound support was divided into 2 portions and each of these re-loaded into separate synthesizer columns. The first column was re-attached to the synthesizer and the 2'-detritylation reaction was performed with 3% TCA:DCM for 5 minutes to completely detritylate and deprotect the sequence. The complete detritylation

was monitored by the MMT$^+$ color collected during the reaction. This deprotected sequence was cleaved from the solid support with mildly basic conditions (anhydrous 2 M NH_3 in EtOH) for 30 minutes at 22°C.[24d,e] The reaction was quenched on dry ice for 20 minutes and evaporated to a pellet on a Speed Vac® concentrator and prior to the extraction of the crude rU$_4$ in sterilized water (recoveries of 2.56 ODs, 5% total yield) for analysis by anion exchange HPLC, AE HPLC. This analysis indicated a degradation pattern consistent with the cleavage of the rU$_4$ sequence under the basic conditions required to cleave the oligomer from the support (**Figure 2.5- A**).[22,33,34,115]

In order to overcome this limitation, a second portion of the solid support was first treated with 3:1 NH_4OH:EtOH (4 h, 55°C)[24] to cleave the 2'-MMT protected RNA sequence from the solid support. The reaction was quenched on dry ice for 20 minutes and evaporated to a pellet on a Speed Vac® concentrator prior to extraction of the rU$_4$ sequence with sterile 50% water in acetone. This was followed with the 2'-detritylation reaction with a solution of 3% trifluoroacetic acid in water (50 µL; 3% TFA:H_2O, pH = 2.5).[116] The reaction mixture was vortexed for 1 hour on a wrist action shaker at room temperature (22°C) for complete reaction. The crude reaction mixture was evaporated to a solid pellet on a Speed Vac® concentrator, re-dissolved in sterile water (recoveries of 5.66 ODs; 16% total yield) and analyzed for purity by AE HPLC. The analysis of the crude produced a cleavage pattern with little, if any, detection of the desired rU$_4$ sequence (**Figure 2.5-B**).

This is an indication that the detritylation conditions (*i.e.* 3% TCA:DCM or 1 M $ZnBr_2$, 85:15 CH_2Cl_2:*i*PrOH) are not selective for the regiospecific deprotection of the MMT protecting groups while on solid support. The degradation of rU$_4$ can result from the complete detritylation and exposure of the 2'-OH group to the vicinal phosphate triester/diester during the basic work-up treatment after synthesis which will quickly degrade the tetraribonucleotide sequence.[22,33,34] Therefore, we turned our attention to alternative RNA phosphoramidite reagents and strategies.
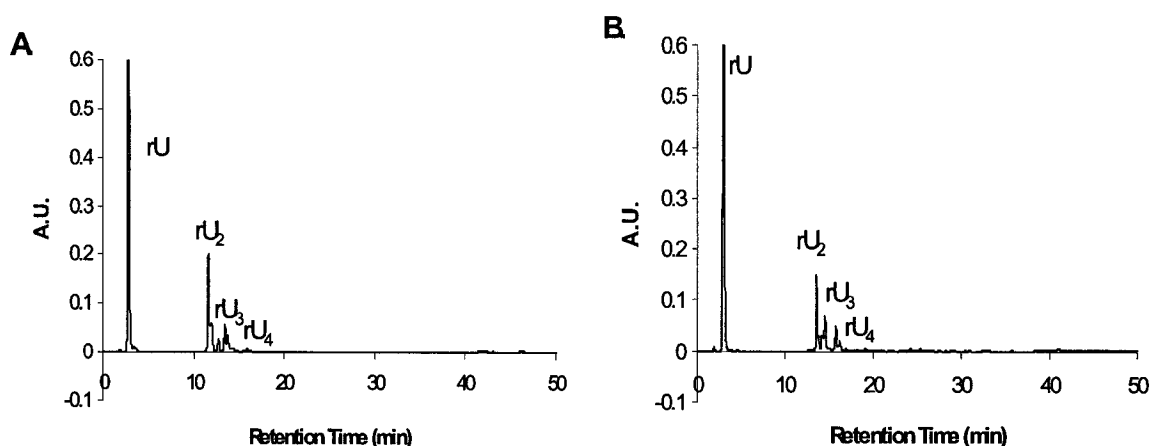
Figure 2.5: The analysis of crude rU_4 sequences by AE HPLC with a gradient up to 30% 1M $LiClO_4$ in water. The tetramer was synthesized by **A.** complete on-column deprotection followed by ammonia treatment, and **B.** cleaving the 2'-MMT tetramer from the solid support followed by acid treatment. Neither conditions yielded significant amounts of full length product; rather, the HPLC trace indicates chain degradation.

2.6 THE LEVULINYL GROUP AS 2'-PROTECTING GROUP FOR RNA SYNTHESIS.

The 2'-O-levulinyl ribonucleosides were next considered as possible synthons for RNA synthesis.[117] At the onset, it was recognized that 2'-O-acyl protecting groups (e.g., benzoyl) in RNA synthesis had been largely unsuccessful due to ease of 2'-3' migration,[22] and lack of specificity and compatibility with other groups. In fact, for these reasons, 2'-O-acyl protecting groups have only very rarely been used in oligoribonucleotide synthesis[117] and are more commonly used in the synthesis of branched RNA.[118] The levulinyl (Lv) group has been previously described by van Boom and others, as a 'transient' 5'-protecting group for ribonucleosides.[117a] Ogilvie was the first to utilize the Lv group for transient 3'-protection in conjunction with the procedures their group developed for block coupling of 3',5'-oligoribonucleotides.[117b] This present study builds on Ogilvie's work as 2'-O-Lv ribonucleoside phosphoramidites are examined as synthons for RNA synthesis. As documented in an earlier study with 2'/3'-O-(2-chlorobenzoyl) ribonucleosides,[119] we reasoned that O-phosphitylation of a mixture of interconverting 2'-O- and 3'-O-Lv ribonucleoside isomers would provide separable, regioisomerically stable, phosphoramidite monomers that would be suitable for RNA synthesis. To demonstrate the applicability of this method, the uridine derivatives and an oligo-rU sequence were prepared first.

2.6.1 Chemical Synthesis of 5'-MMT 2'-Lv and 5'-MMT 3'-Lv rU Phosphoramidtes.

The regioselective acylation of 5'-MMT uridine, **2.4**, was attempted under kinetically controlled conditions.[120] The 5'-MMT uridine, **2.4**, was allowed to react with one equivalent of the acylating agent levulinic acid/2-chloro-1-methylpyridinium iodide/DABCO at 0°C for 15 min to afford predominantly a mixture of isomers **2.7** and **2.8** (60 %) along with **2.9** (10%). The crude products were purified by silica gel column chromatography, yielding **2.7** and **2.8** as an inseparable mixture of regioisomers (3:1 ratio by NMR, in favor of **2.7**). It is well known that 2' or 3'-O-acyl ribonucleosides isomerizes under, neutral, acid or alkaline conditions.[22,121] This inseparable mixture was phosphitylated[122] to generate the 3' and 2'-ribouridine phosphoramidite diastereomers, **2.10** and **2.11**, where an unexpected reversal in isomeric ratio was observed, i.e., the 2'-O-phosphoramidite derivative **2.11** was now the major component in the mixture (**2.11:2.10**, ~2:1 ratio, determined by ^{1}H NMR). This suggests that the minor component, the 3'-O-Lv isomer (**2.8**), reacts faster than the 2'-O-Lv isomer (**2.7**). The eluent conditions were optimized for the purification of the 4 phosphoramidite diastereomer regioisomers, (**2.10** and **2.11**, R_f values 0.18 to 0.25) with an eluent system of 1:1 to 4:1 (v/v) EtOAc:Hex. The 3' and 2'-phosphoramidites were separated and recovered in 20% and 60% yields, respectively (**Figure 2.6**).

4 separable phosphoramidite diastereomers
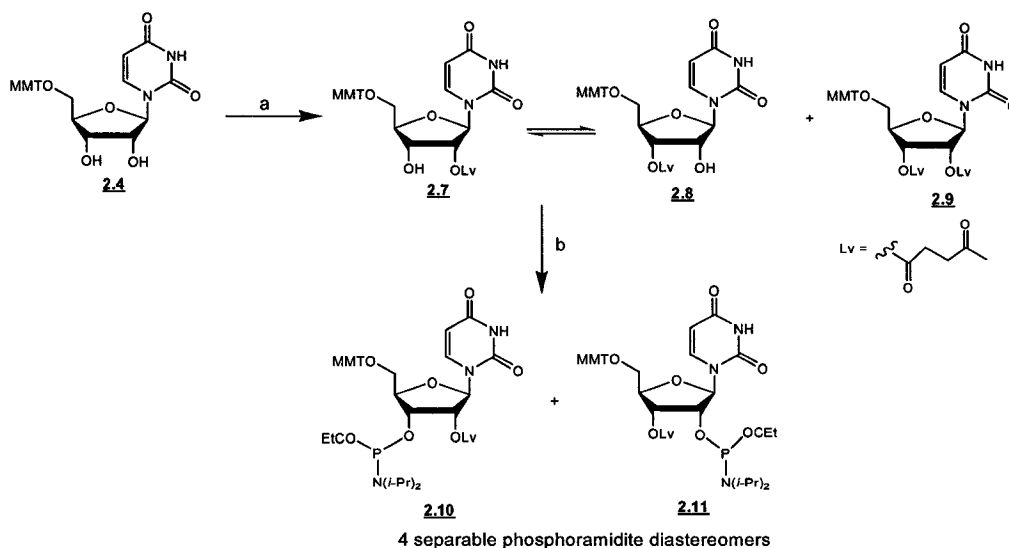
Figure 2.6: Reagents and Conditions: a. Levulinic acid, 2-chloro-1-methylpyridinium iodide, 1,4-diazabicyclo [2.2.2] octane (DABCO), 1,4-dioxane, MeCN, 15 min, 0°C, **2.7** and **2.8**; 60%; **2.9**, 10%. b. P(OCEt)[N(i-Pr)₂]₂, 1-H-tetrazole/Et-N(i-Pr)₂, MeCN, 5 h, 22 °C **2.10**, 20%; **2.11**; 60%.

2.6.2 Characterization of MMT Lv rU nucleosides and phosphoramidites

As described above, the selective 2' and 3'-O-levulination reaction of **2.4** proceeded quickly and efficiently affording 60% of **2.7** and **2.8** as inseparable mixtures. The more non-polar diastereomers, **2.11** (R_f : 0.36 and 0.30), were collected as the major products of the reaction in purified yields of 60%. The more polar diastereomers, **2.10**, R_f : (0.24 and 0.20), were collected as the minor products of the reaction in collective yields of 20% after purification. The mixture was characterized by NMR and MS. The integration of the NMR peaks indicated the relative abundance of 3:1 for **2.7** and **2.8**, respectively. ^{1}H-^{1}H COSY experiments in combination with deuterium exchange experiments were performed to unambiguously assign the chemical shifts of the sugar and hydroxyl protons (**Figure 2.7 A** and **B**; **Figure 2.8**). The ^{31}P-NMR spectra of **2.10** and **2.11** each consisted of a pair of signals at $ca.$ 150 ppm (pairs of P-epimers). ESI-MS provided confirmation of the molecular weights ($C_{43}H_{51}N_4O_{10}PNa$ calc. 838, found 838). Compounds **2.10** and **2.11** were further characterized by 2D-hetereonuclear correlation ^{1}H-^{31}P CIGAR NMR (**Figure 2.9**).
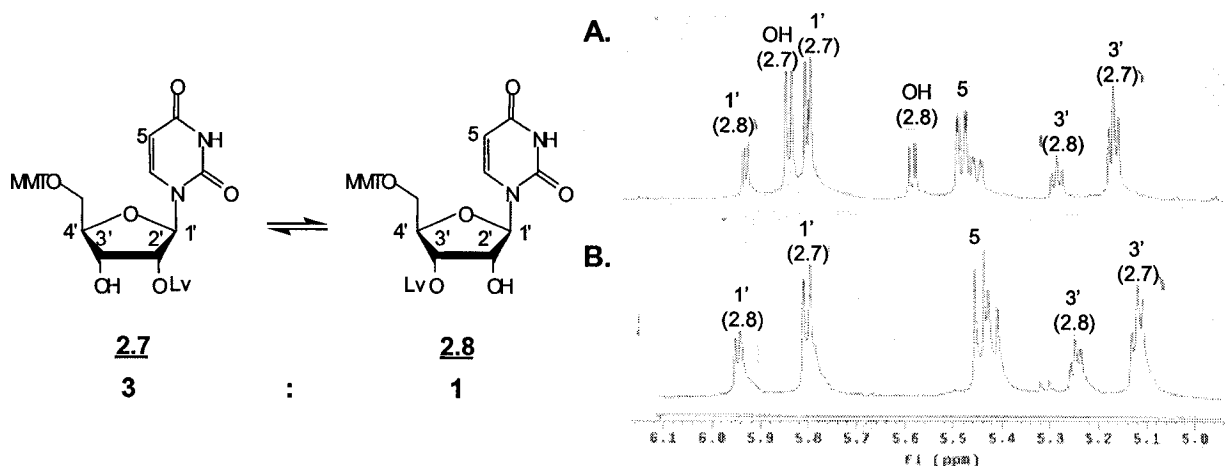
Figure 2.7: The structures and assignments of ^{1}H-NMR signals of **2.7** and **2.8** (5-6 ppm region). **A.** ^{1}H NMR spectrum in **A.** DMSO-d_6 solvent and **B.** D$_2$O and DMSO-d_6.
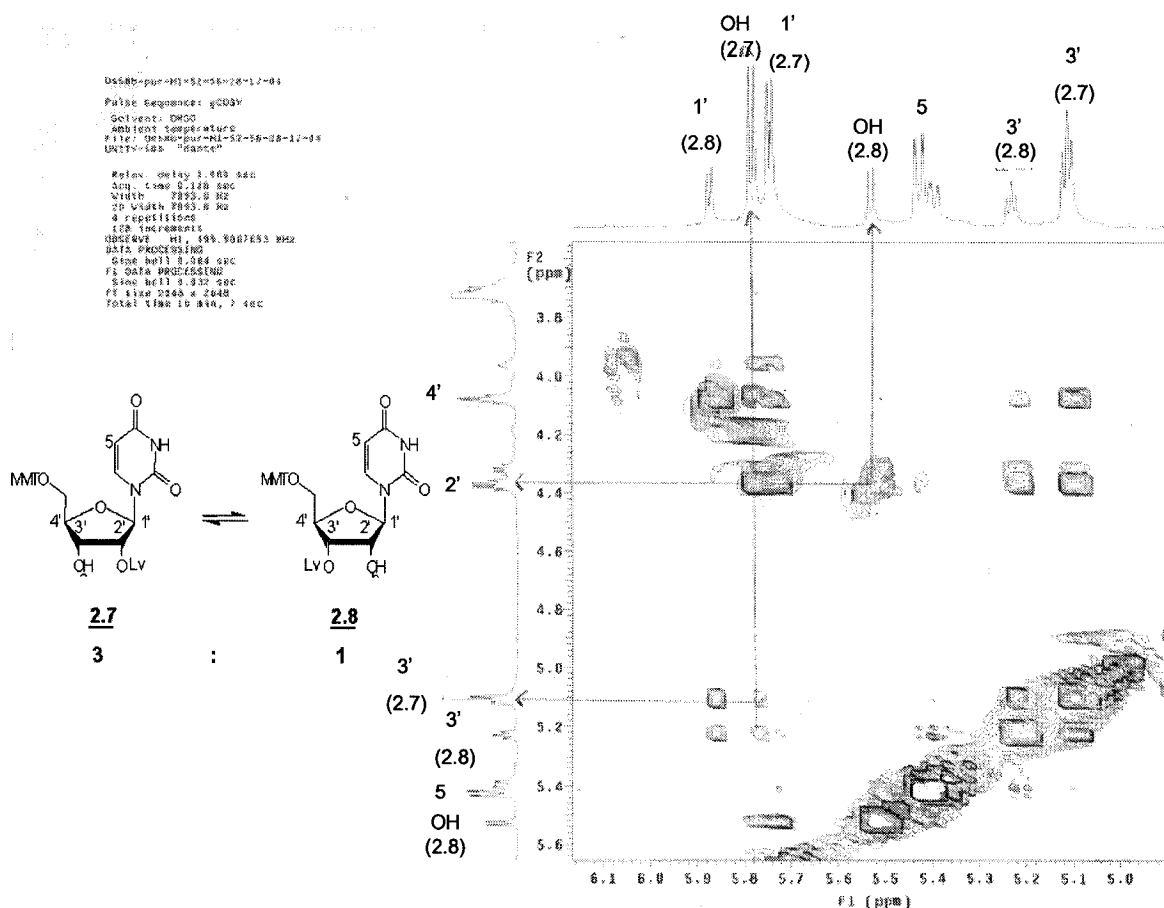
Figure 2.8: ^{1}H-^{1}H homonuclear correlation COSY spectrum revealing the coupling of H3' to 3'-OH, and H2' to 2'-OH, for **2.7** and **2.8**, respectively. The spectrum also revealed the relative abundance of **2.7** and **2.8** (3:1 ratio, respectively).

The individual regioisomer diastereomers were characterized by virtue of a coupling crosspeak of the phosphorous to the 2'- or 3'-protons of the nucleosides **2.11** and **2.10**, respectively (**Figure 2.8**). The purification and characterization of high purity phosphoramidite regioisomer diastereomers, **2.10** and **2.11**, is a requirement for the efficient solid phase synthesis of oligonucleotides containing 3',5' and 2',5'-internucleotide phosphodiester linkages.
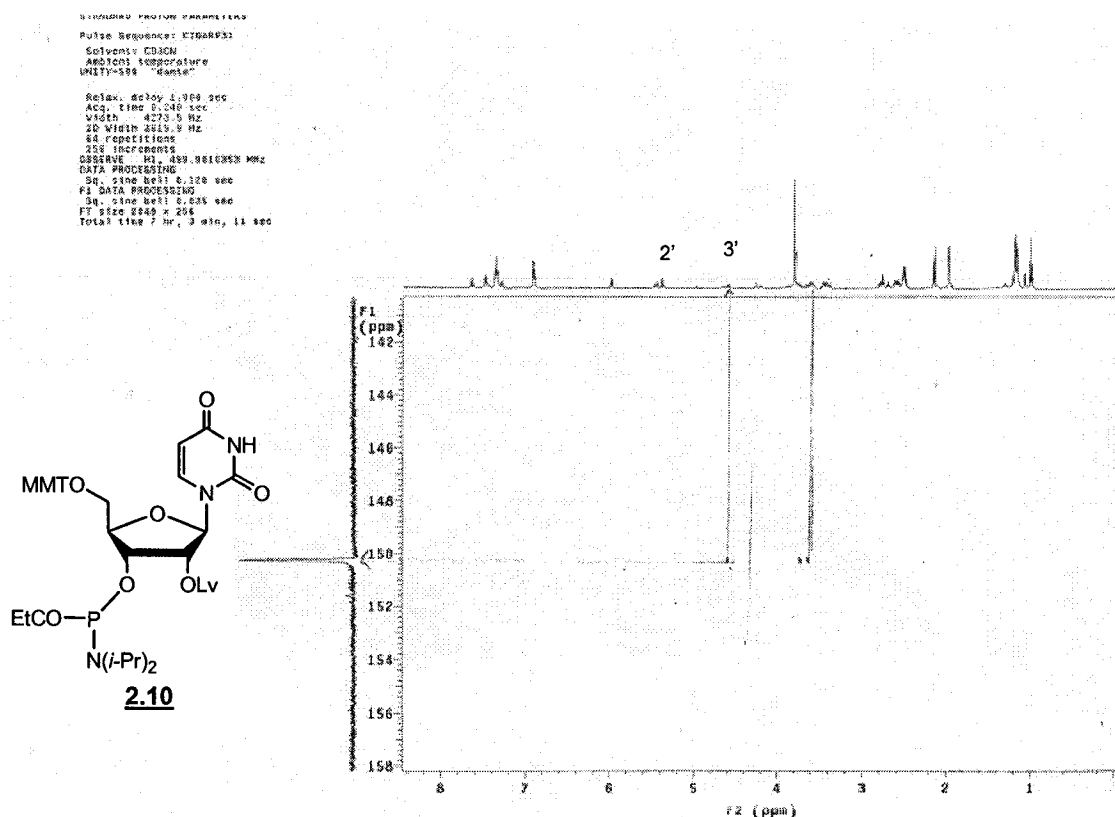
Figure 2.9: The 2D heteronuclear ^{1}H-^{31}P CIGAR NMR of **2.10** (single isolated phosphorus epimer). The spectrum illustrates coupling of the H3' with the phosphorus signal, establishing the position of the phosphoramidite moiety at C3'.

2.7 SOLID-PHASE SYNTHESIS OF OLIGOURIDYLATE SEQUENCES THROUGH THE USE OF 5'-MMT 2'-OLv rU 3'-*O*-PHOSPHORAMIDITE, 2.10

With phosphoramidite synthon **2.10** in hand, the oligopyrimidine test sequence 5'-dTTTTTT-rU-dTT-3' was first synthesized. The synthesis was performed on a 1 μmol scale starting from the thymidine-tethered support, **1.1**, with a loading capacity of 45 μmol/g.[19] The ribouridine phosphoramidite **2.10** (0.15 M in MeCN) was coupled after assembly of dTT-3', using 0.25 M ETT in MeCN as activator and a 10 minute coupling cycle. Chain extension was continued with six more couplings using dT phosphoramidites to yield the target sequence. The monomer **2.10** coupled with 95% efficiency (trityl assay). The support bound oligomer was decyanoethylated to convert the labile phosphate triester to the more stable phosphate diester backbone.[14,22,33,34] This was followed by a mildly basic treatment with 0.5 M NH$_2$NH$_2$-H$_2$O in 3:2 (v/v)

pyr:HOAc buffer, to selectively remove the 2'-levulinyl protecting group.[120] The completely deprotected oligomer was cleaved from the succinyl linked CPG support with anhydrous 2M NH_3 in EtOH, to minimize the 2'-OH assisted internucleotide cleavage.[24d,e] This solution was evaporated after 20 minute reaction at 22°C, and the crude oligomer was extracted in sterile water (25% yield) prior to AE HPLC analysis and purification. Not surprisingly, the HPLC trace indicated that some site-specific cleavage at the uridine insert had occurred (**Figure 2.10-A**). Nevertheless, the desired product was also formed in 20% yield.

In order to avoid internucleotide cleavage under the basic conditions required to cleave the oligomer from the solid support, an alternate support will be necessary. Since the completion of this preliminary work, Jeremy Lackey a Ph.D. student has continued this work[124], and has shown that Pon's hydroquinone-O,O-diacetic acid ("Q")-linker CPG solid support[123] permits the release of the oligonucleotide strand using fluoride ions, under conditions that do not lead to cleavage of the unprotected RNA strand. A brief summary on the preliminary results acquired by Jeremy Lackey for the optimized synthesis of a homopolymeric 9-nt oligouridylate sequence is presented. The conditions developed here were optimized for the synthesis of a completely modified 9-mer oligouridylate sequence such that the required quantity of uridine phosphoramidite **2.10** was minimized to solution concentrations of 0.1 M and the coupling efficiency increased to stepwise yields of 99% with short coupling times (1-2 min) that are similar to those achieved during DNA synthesis. The optimized deprotection conditions included a 90 minutes decyanoethylation step with a solution of 4:6 (v/v) NEt_3:MeCN followed by a 2'-delevulination step for 30 minutes with a solution of 0.5 M NH_2NH_2-H_2O in 3:2 (v/v) pyr:HOAc. The oligouridylate sequence was cleaved from the 'Q-linker' support (15 minutes, 1 M TBAF in THF) in good yields and with purity of 85% (AE HPLC analysis, **Figure 2.10-B**). The 9-nt product was further characterized by analytical (HPLC and PAGE) comparison with control 9-nt oligouridylate sequences synthesized with the conventional RNA 2'-TBDMS phosphoramidite chemistry (data not shown). The structure of the target sequence was confirmed by molecular weight analysis with MALDI-TOF MS (Calcd 4506 g/mol; found: 4502). This is an indication that oligoribonucleotides can be deprotected and/or cleaved from the support with non-basic

conditions (*i.e.* F⁻) for the recovery of full length sequences without concomitant isomerization or cleavage of the nascent strand. Therefore, orthogonal Lv and MMT protecting groups in RNA phosphoramidites provide the potential for developing an alternate method for the solid phase synthesis of RNA oligonucleotides.
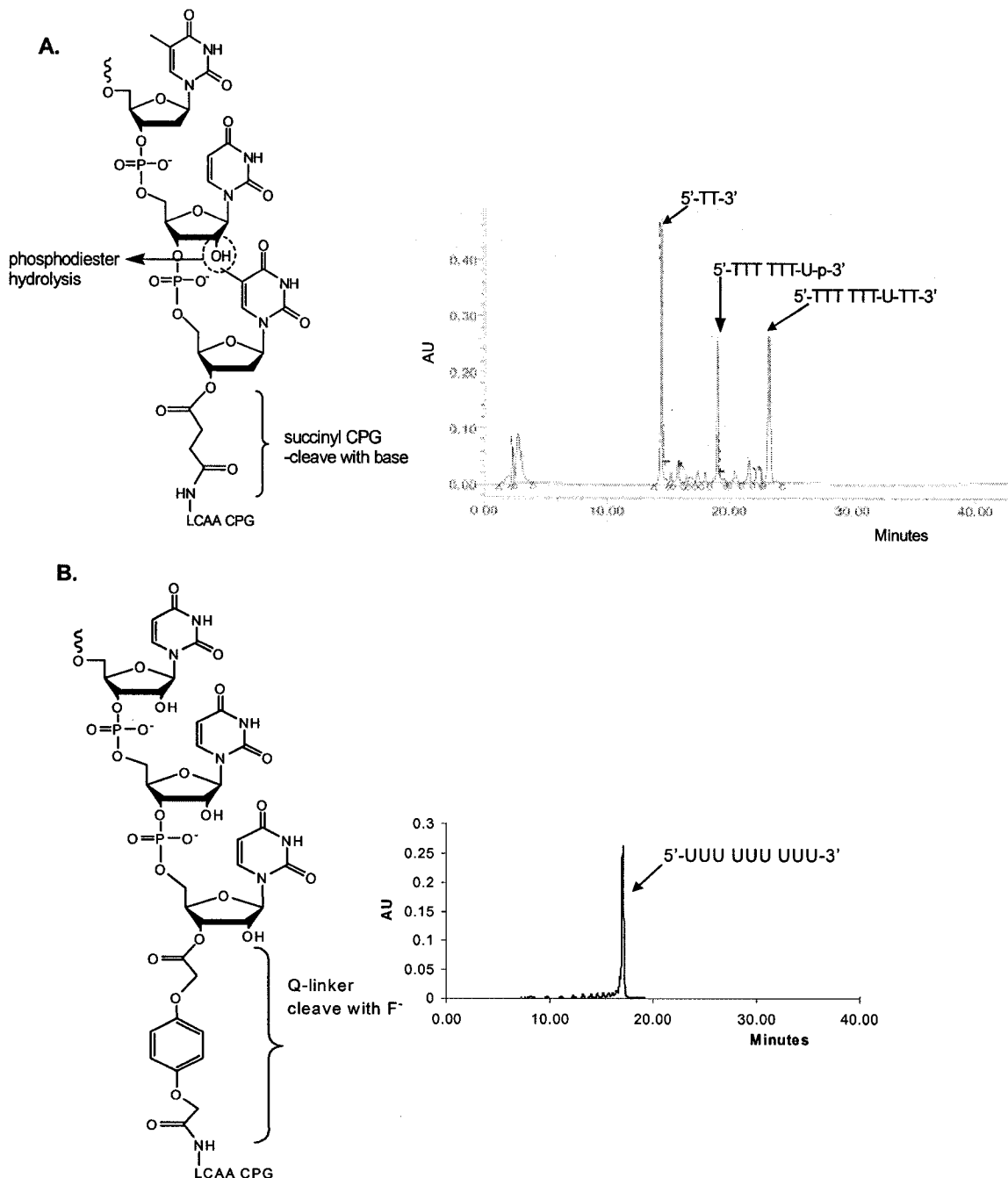
Figure 2.10: AE HPLC analysis of RNA oligonucleotides; 30% 1 M LiClO₄ in water. The sequences were synthesized on A. succinyl linked and B. Q-linked support.

2.8 CONCLUSIONS

The advent of oligoribonucleotide-based therapeutics such as siRNAs has increased the demand for bulk quantities of RNA production with sequence specificity, high purity and at a cost efficient rate. Although RNA 2'-TBDMS phosphoramidites remain the monomers of choice for routine RNA synthesis, and synthesis on large scale using these monomers has also been initiated, there is still a need for improvement on cost, coupling efficiencies while minimizing monomer waste, and time of deprotection and purification steps.

We have shown that the Lv group is a suitable protecting group for the 2'-hydroxy functions of uridine building blocks. The major advantage of the Lv group over other 2'-protecting groups is in the on-column unblocking step at the end of the synthesis which greatly simplifies and speeds up post-synthesis processing. While this provides a clear cost-advantage over some current protection schemes, the arduous separation of derivatives such as **2.10** and **2.11** must be taken into account.

Jeremy Lackey, a Ph.D. student in our lab has continued this research making the remaining RNA phosphoramidites.[124] His investigations have led to the discovery of an alternative and complementary method for the solid phase synthesis of RNA oligonucleotides containing biologically relevant mixed base sequence compositions (siRNAs). Our lab continues to develop novel improved methods for the automated *large scale* synthesis of DNA and RNA using ionic supports.[125]

CHAPTER 3: CHEMICAL STABILITY OF PHOSPHATE TRIESTERS AND DIESTERS DURING SOLID PHASE RNA SYNTHESIS

3.1 INTRODUCTION

The solid phase approach has emerged as a leading strategy for the efficient production of biomolecules in excellent yield and purity.[24,126] An example is related to the phosphoramidite method, which is the leading solid phase synthesis strategy for DNA and RNA synthesis.[24] While conditions are optimized for each of the 4 steps of the DNA/RNA synthesis cycle, side reactions such as acid mediated depurination during the detritylation step[127], incomplete coupling of the nucleoside phosphoramidite yielding oligonucleotide failure sequences (N-1, N-2 *etc.*)[109,128] formation of higher molecular weight species (N+1, N+2 sequences) and incomplete detritylation, all affect the integrity and purity of the target oligomer.[24] Since these reactions occur *in-situ* during the automated cycle, it is often difficult to predict the success of a synthesis until after the oligonucleotide chain has been assembled, cleaved from the support and completely deprotected for analysis.

One such example is the ambiguous chemistry occurring at the phosphorus center of the phosphite/phosphate triester backbone during the solid phase synthesis. The phosphate bridge is generally protected as a triester to prevent side reactions from occurring at the non-bridging phosphate oxygen atom.[21] However, phosphate or phosphite *O*-alkyl protecting groups may be modified during solid phase oligonucleotide synthesis making their function and utility somewhat doubtful. For example, Caruthers and co-workers have demonstrated that, *in solution*, the 2-cyanoethyl (CNEt) type phosphate protecting group is prone to cleavage and transesterification reactions during the iodine oxidation step that converts the phosphite to a phosphate triester.[129] The loss of alkyl phosphite protecting groups during oxidation has also been observed by Letsinger and co-workers on the solid support. Specifically, they observed that methyl phosphite triesters can undergo selective oxidation without loss of the alkyl group when the oxidation was carried out with mild oxidants such as *tert*-butylhydroperoxide (*t*-BuOOH); however, under oxidizing conditions that are routinely used in the solid phase synthesis protocol, *i.e.* 0.1 M iodine in tetrahydrofuran/pyridine/water [20/9/1 (v/v/v), 2

minutes], an appreciable amount of the phosphate diester was also recovered.[130] [31]P cross polarization (CP) magic angle spinning (MAS) nuclear magnetic resonance (NMR) experiments[131] have revealed the loss of CNEt protecting groups during the oxidation of di and tri-nucleotides attached to a CPG support. Interestingly, decyanoethylation also took place with the use of milder *t*-BuOOH oxidant, leading to the proposal that CNEt removal was related to the interactions of the phosphate linkages with the CPG surface.[131] Therefore, it is difficult to un-ambiguously monitor the *in-situ* chemistry of the oligonucleotide phosphate backbone as phosphate triesters are labile in solution to the bases and nucleophiles [pyridine, (pyr), water, (H_2O), hydroxide, (OH^-), and iodide and fluoride, (I^- or F^-) ion] commonly used as reagents during the course of the solid phase synthesis cycle.[129,130]

Another complication that arises during RNA (and branched RNA) synthesis, is the removal of the 2'-protecting group under conditions that may lead to internucleotide cleavage and/or isomerization (**Figure 3.1** and Chapter 2).[22] This has been extensively studied, and linked to the stability of oligoribonucleotide phosphate diesters and triesters and to their mechanisms and kinetics of hydrolysis by ribozymes and ribonucleases.[132,133,134] Acid or base catalyzed isomerization/cleavage reactions in solution are extremely rapid, occurring within seconds for RNA triesters,[135,136] which are typically more reactive than the corresponding diester by a significant factor of 10^5-10^6,[137] resulting in RNA hydrolysis typically on the order of minutes to hours.[138] The neutral internucleotide linkage and the vicinal 2'-OH both contribute to the rate acceleration, since no base-catalyzed hydrolysis of DNA occurs after 1 hour reaction at 100°C.[138,139] These studies would also lend evidence or challenge recent reports about "RNA X", a peculiar branched RNA species identified during *in-vitro* pre-mRNA splicing assays.[140] RNA X is structurally unique because it contains a phosphotriester linkage vicinal to a free 2'-OH group, a motif that has been shown by Todd[135] to be highly unstable over a wide pH range. While extensive work has been described by Breslow[141] and Lonnberg[142] on the mechanistic and kinetic aspects of phosphate diester and triester hydrolysis in RNA in solution, very little is known about similar reactions on the solid support. This thesis Chapter examines the stability of ribosyl 3'-phosphates on a solid support, and specifically, we report on the use of a branchpoint synthon, diMMT rU

amidite **2.5**, for studying the 2'-OH assisted cleavage of phosphodiester and phosphotriester linkages under neutral, basic and acidic conditions (**Figure 3.2**). These studies are relevant not only for the development of strategies for the synthesis of RNA[124] and branched RNA[40], but also for furthering our understanding of the stability of phosphate diesters and triesters.[135,136,137,138]
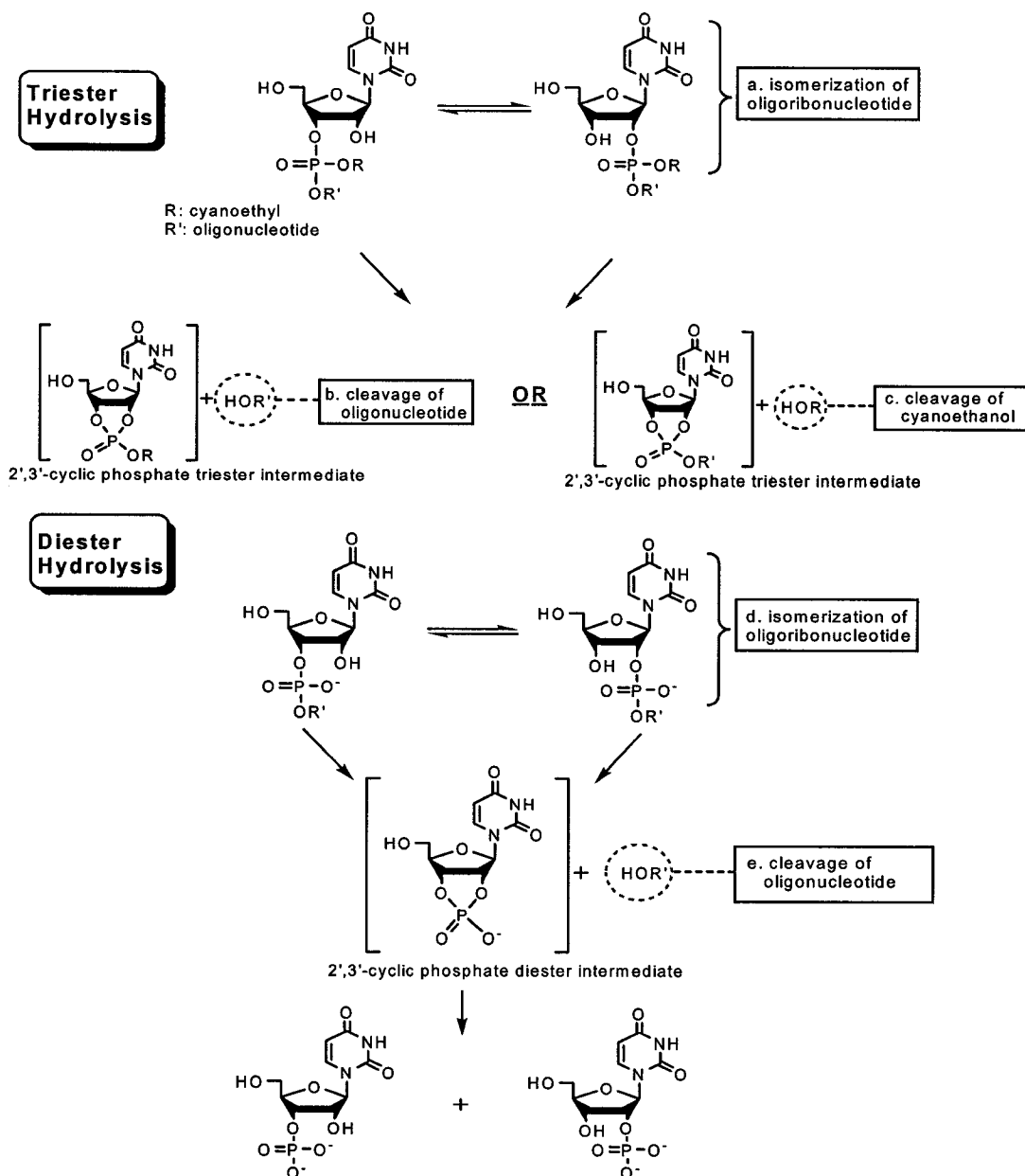
Figure 3.1: The cleavage and isomerization reactions of phosphate diesters and triesters in RNA. The vicinal 2'-OH group can facilitate the isomerization (**a.** or **d.**) and the cleavage (**b.**, **c** and **e**) reactions by the release of a 2',3'-cyclic phosphate triester/diester intermediate which in turn hydrolyzes to the more stable nucleoside 2'/3'-phosphate monoester regioisomers.

3.2 EXPERIMENTAL DESIGN FOR MONITORING THE STABILITY OF RNA PHOSPHATE DIESTERS AND TRIESTERS ON A SOLID SUPPORT

The experiment designed to monitor the stability of RNA phosphate triesters and diesters is shown in **Figure 3.2**. The strategy consists of growing dT_5 on the solid support, followed by coupling of monomer **2.5**. Treatment of the resulting hexanucleotide under acidic conditions would unmask the hydroxyl groups at 2' and 5' positions, from which chain extensions (i.e., growth of dT_6) can ensue to afford a Y-shaped or branched oligonucleotide structure that is 18-nt in length (**Figure 3.2**). To monitor the stability of the phosphate triester or diester, the sequence was either decyanoethylated[40] (phosphate diester backbone) or kept un-modified (phosphate triester backbone) prior to the acid detritylation conditions which favored the release of the 2',5'-diMMT protecting groups (*i.e.* 200 second automated detritylation step). The deprotected ribouridine was subjected to acidic, neutral and alkaline conditions to determine the extent of the isomerization and cleavage reaction with the liberated 2'-hydroxyl group vicinal to either the phosphate diester or triester. The reaction was conducted during fixed time points, followed by extensive washing and drying of the support bound oligomer prior to the automated branching with dT_6 "arms" from the liberated 2' and 5'-hydroxyl groups of the branchpoint synthon to yield the desired branched product, **3.1**, in addition to the potential side-products originating from the isomerization, *i.e.*, **3.2** and cleavage reactions, *i.e.* **3.3**. Following chain assembly, the solid support was treated with ethanolic ammonium hydroxide,[24] and the oligonucleotides released and analyzed by Anion Exchange[143] (AE), Ion-Pairing Reverse Phase[144] (IP-RP) HPLC and by denaturing PAGE[24]. The analyses were designed to monitor the relative amounts of the branched 18-mer regioisomers (pathway c, **3.1** and pathway d, **3.2**, **Figure 3.2**) and the linear 11-mer oligothymidylate that would result if dT_5 coupled to dT_6 following release (cleavage) of rU (pathway a, **Figure 3.2**). Alternatively, the possible release of 2-cyanoethanol during acid treatment (pathway b, **Figure 3.2**) which upon chain extension would yield the branched 18-mer regioisomers (**3.1** and **3.2**) and a linear 12-mer sequence with an internal ribouridine unit that would be labile to the basic (ethanolic ammonium hydroxide) deprotection conditions used for the cleavage of the oligomer from the support. Internal cleavage of this sequence would result in the formation of dT_5 (**3.4**) and

dT$_6$rUp* (**3.5** and **3.6**, where p* is a 2' or 3' phosphate monoester). Therefore, in order to characterize all possible reaction products, oligomers **3.1**-**3.6** were prepared as standards according to established methods.[24]
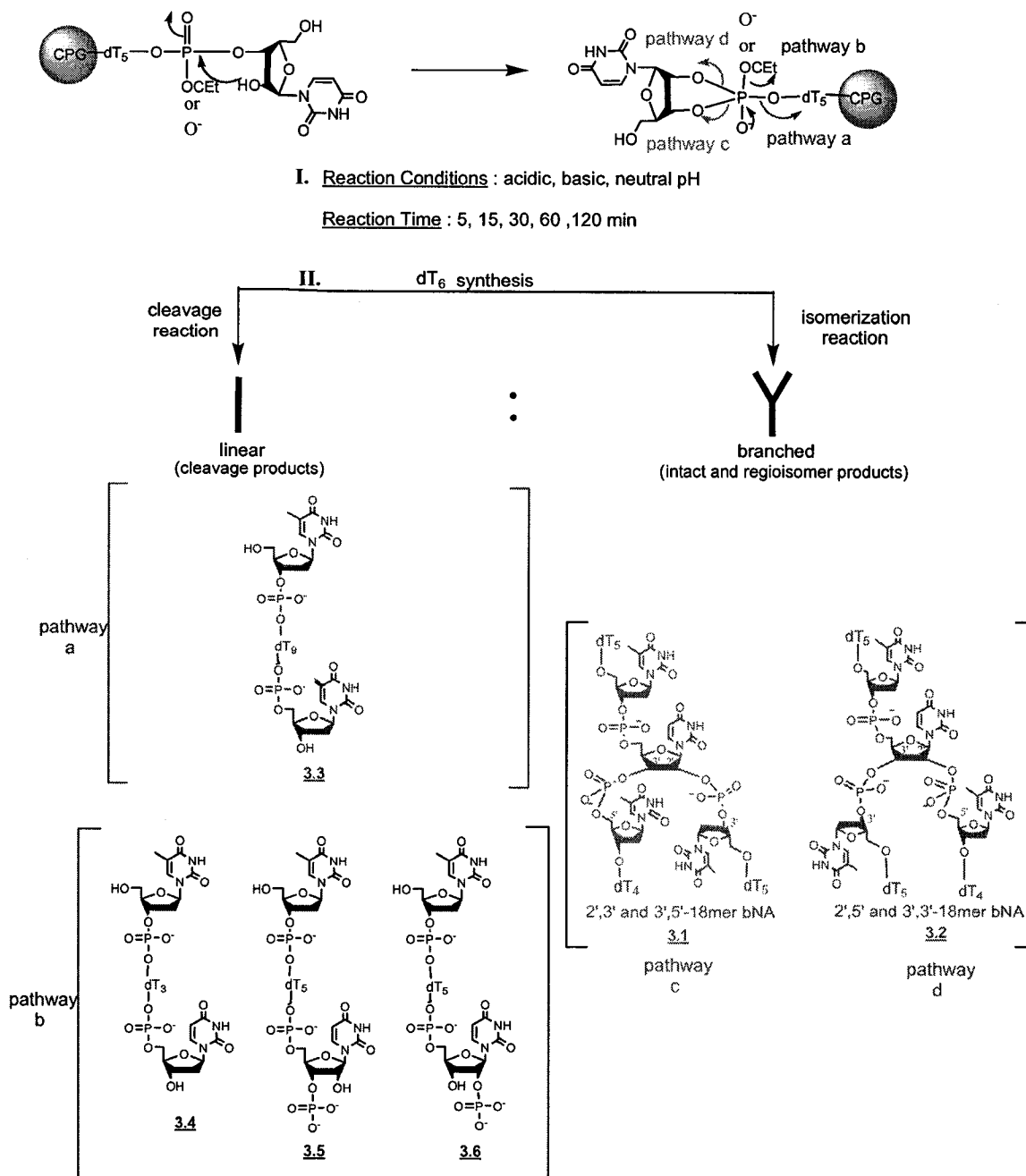
Figure 3.2: The experimental design for monitoring the stability of RNA phosphate triesters and diesters on solid support. The analysis of the reaction produced the concomitant formation of the branched 18-mer regioisomers (**3.1** and **3.2**) in addition to either 5, 7-mer or 11-mer linear cleavage products (**3.3** – **3.6**). The complete analysis and characterization of the reaction products was conducted with AE HPLC, RP-IP HPLC and denaturing PAGE techniques.

3.3 STABILITY OF PHOSPHATE CNEt PROTECTING GROUPS DURING SOLID PHASE OLIGORIBONUCLEOTIDE SYNTHESIS

In oligonucleotide synthesis, the phosphate backbone of the growing oligonucleotide exists as phosphate triesters due to the use of 2-cyanoethyl (CNEt) protecting groups, which consist of activated alkyl moieties capable of C-O bond cleavage when exposed to alkaline conditions or to certain nucleophiles, yielding phosphate diesters and acrylonitrile[21,129,130] Phosphate diester and triester isomerization and cleavage reactions in the presence of a vicinal 2'-hydroxyl group were examined in order to resolve the ambiguity surrounding the presence (or absence) of the phosphorus CNEt protecting groups during the course of the automated solid phase synthesis cycle (see section 3.1). The CPG-dT$_5$-rU-5' sequence was synthesized using the standard solid phase synthesis protocol[24] with anhydrous 3% TCA:DCM detritylation step (200 sec), which liberated both the 5'- and 2'-hydroxyl groups from the branchpoint synthon for subsequent reaction. To study the extent to which the branchpoint synthon undergoes acid catalyzed isomerization and cleavage reactions under the detritylation conditions, the 3% TCA:DCM treatment was continued for an additional 5, 15, 30, 60, and 120 minutes prior to chain extension from the branchpoint uridine with dT$_6$, from the liberated 5'- and 2'-hydroxyl groups (**Figure 3.2**). The oligonucleotide was cleaved from the support and deprotected using ethanolic ammonium hydroxide at 55°C, for 3 hours, followed by evaporation and extraction of the oligomers from the support and analysis by either HPLC or PAGE. The data collected indicated the extent of the branchpoint isomerization by comparing the desired intact 18-mer, **3.1**, (*i.e.* with branchpoint connectivity of 3',5' and 2',3') and the branched oligonucleotide regioisomer, **3.2**, (*i.e.* with branchpoint connectivity of 3',3' and 2',5') and also the degree of the cleavage reaction, which produced a linear 11-mer oligothymidylate byproduct (dT$_{11}$ **3.3**). No additional major side products were observed during the course of the stability experiments. The reaction process, analyzed using HPLC and denaturing 24% analytical PAGE, produced 80% of the cleavage byproduct, **3.3**, after the 120 minutes exposure to the detritylation conditions but the cleavage reaction was found to be the slower of the two processes (**Figure 3.3**). This was determined by further analysis of the crude reaction mixture using RP-IP HPLC, which indicated a fast isomerization, resulting in the formation of 40% of the

regioisomer after only 5 minute exposure to the detritylation conditions. The branched regioisomers were separated, analyzed and their identities confirmed using RP-IP HPLC and MALDI-TOF MS, to yield similar molecular weights (*i.e.* calc. 5417; found **3.1**: 5416, and **3.2**: 5418).
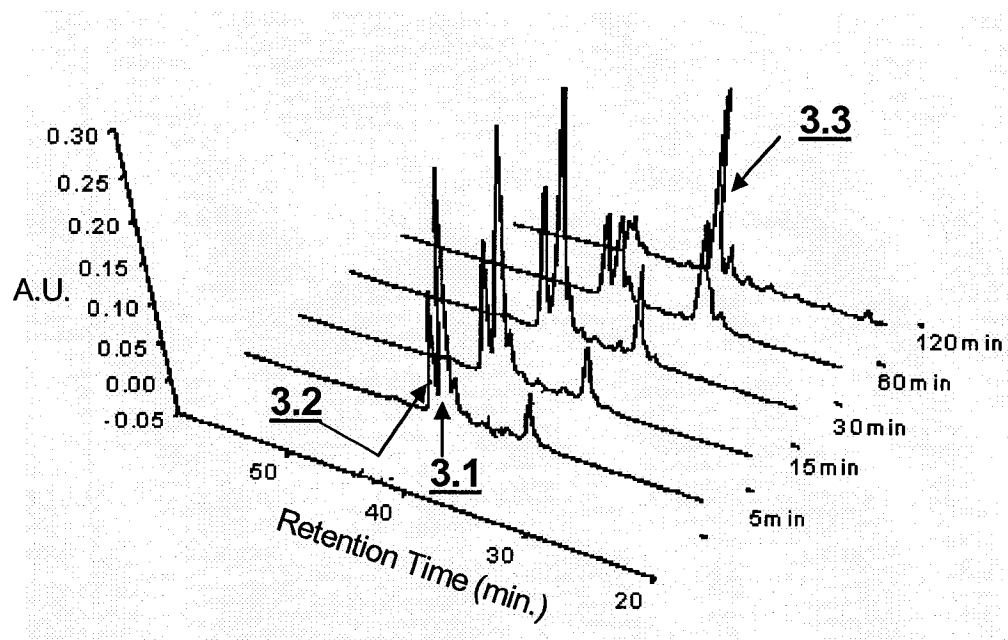
Figure 3.3: The analysis after the acid detritylation reaction of the uridine branchpoint phosphate triester by RP-IP HPLC with an elution gradient up to 13% MeCN in 0.1 M TEAA for a total run time of 70 min. The analysis indicates the separate regioisomers **3.1** and **3.2** and the cleavage linear 11-mer byproduct, **3.3** without the formation of additional major side-products up to 2 h reaction.

A parallel experiment was performed where the synthesis cycle on the CPG-dT$_5$-rU-5' sequence was interrupted once the branchpoint synthon, **2.5**, was coupled, prior to detritylation so that the 2'- and 5'-MMT groups were still attached. The synthesis column was subsequently removed from the instrument and exposed to a 10 mL volume of a 4:6 (v/v) triethylamine:acetonitrile (TEA:MeCN) solution over a 90 minute period at room temperature (22°C), followed by a washing of the support with 30 mL of MeCN and 30 mL of THF. These conditions have been shown to effectively decyanoethylate the support bound oligonucleotide, yielding a phosphate diester backbone, without cleaving the oligonucleotide from the support or the occurrence of any other observable side reactions.[40] The column was subsequently re-attached to the synthesizer and the

detritylation procedure was conducted for the standard 200 seconds. The determination of the chemical stability of the phosphate diester with the vicinal 2'-hydroxyl group was performed in the same manner as for the triester case, using a 3% TCA:DCM treatment for 5, 15, 30, 60 and 120 minutes. After the allotted exposure, the synthesis was continued with dT_6, from the liberated 5'- and 2'-hydroxyl groups and the products were cleaved from the support as before, with the alkaline treatment followed by evaporation and extraction in sterile water prior to the analysis with HPLC and PAGE, showing only a 50% cleavage to yield the linear 11-mer oligothymidylate sequence, **3.3**, after 120 min exposure to 3% TCA:DCM (**Figure 3.4**). The RP-IP HPLC analysis of the crude reaction mixture identified the presence of a single branched 18-mer oligonucleotide product, **3.1**, (*i.e.* MALDI TOF MS calcd. 5417, found 5418) with little if any regioisomer, **3.2** observed, even after the 2 hour exposure (**Figure 3.4**).
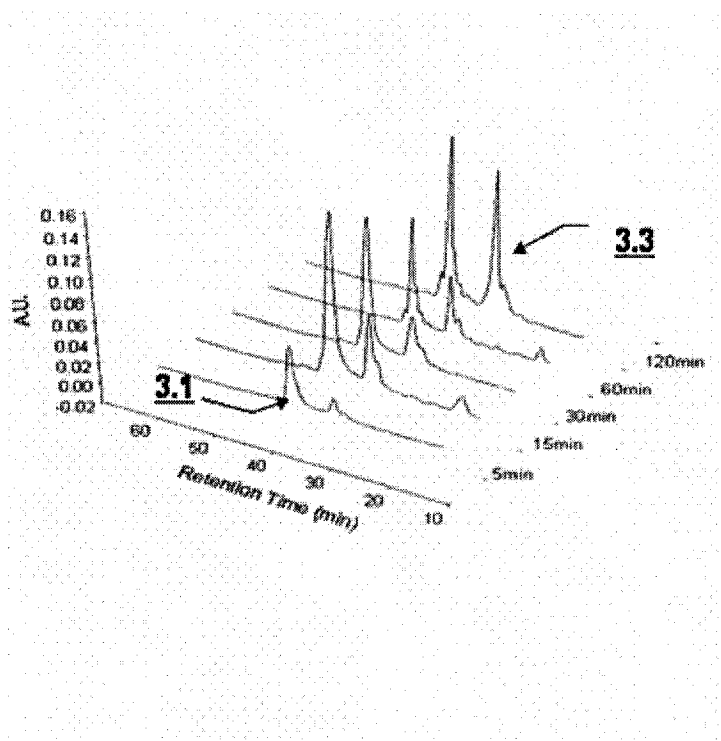
Figure 3.4: The analysis after the acid detritylation reaction of the ribouridine branchpoint phosphate diester by RP-IP HPLC with an elution gradient up to 13% MeCN in 0.1 M TEAA for a total run time of 70min. The analysis indicates the intact product, **3.1** the cleavage linear 11-mer byproduct, **3.3** without the formation of additional major side-products during the 2 h reaction.

The change in reactivity of the cleavage reaction and the absence of the branched regioisomer side products in the experiment involving manual 2-cyanoethyl deprotection, indicates that the CNEt protecting groups (*i.e.* phosphate triesters) are stable and present throughout the solid phase synthesis protocol (**Figure 3.5**). Moreover, in the presence of a vicinal 2'-hydroxyl group, the phosphate triester backbone is prone to rapid isomerization (seconds) but an unexpectedly slow cleavage reaction ($t_{1/2} \sim$ 1 hour) under the anhydrous acid catalyzed detritylation conditions. This observation is significant relative to the in-solution data for the chemical stability of CNEt protected phosphate triesters.[145] This data indicated that a CNEt protected nucleoside 3'-phosphate triester, exposed to the same detritylation conditions exhibited cleavage rates with $t_{1/2} \sim$ 12 minutes by ^{31}P NMR.[145] Therefore, the decrease in the rate of the cleavage reaction of phosphate triesters generated on solid support (*i.e.* CNEt present) *vs.* the solution phase model can, in part, be associated with the possibility of the release of 2-cyanoethanol (**Figure 3.1- c**), which would lead to the formation of identical branched oligonucleotide regioisomers upon cleavage of the 2',3' cyclic phosphate triester intermediate. The careful analysis of many more protecting groups[146] with varied leaving group capabilities relative to CNEt is necessary, in order to obtain further kinetic data for the relative rates of the cleavage and isomerization reactions on solid support. Additionally, it can also be suggested that these types of reactions in the solid phase are less tolerant, resulting in a decrease in the reactivity.[148]
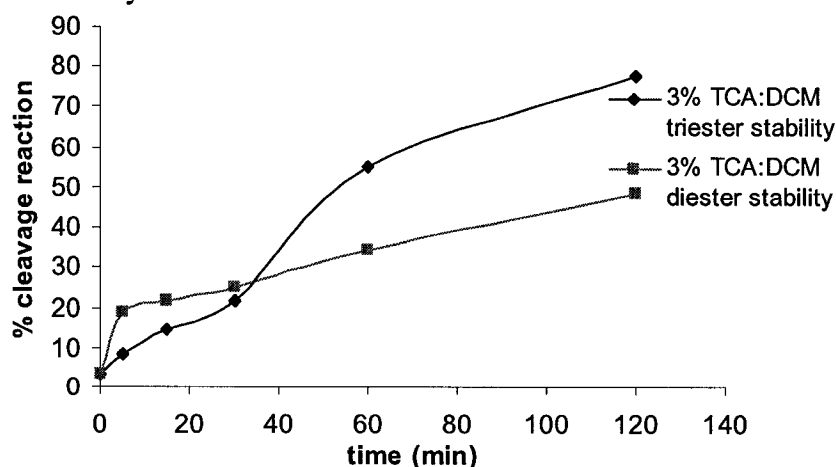
Figure 3.5: The relative stability of the ribouridine phosphate triesters and diesters during the acid detritylation conditions (*i.e.* 3% TCA:DCM). The plot indicates the extent of the cleavage reaction for a 5'-terminal ribouridine tethered to a pentameric polythymidylate sequence while on the solid support

58

3.4 THE CHEMICAL STABILITY OF PHOSPHATE TRIESTERS AND DIESTERS DURING SOLID PHASE OLIGORIBONUCLEOTIDE SYNTHESIS

The reactions described above all occurred under anhydrous trichloroacetic acid solutions, and thus are not representative of RNA cleavage reactions that occur under physiological conditions (in vivo). Thus, the chemical stability of RNA phosphate diesters and triesters on solid support was further studied in aqueous buffer conditions.[138,139,140] Therefore, aqueous buffers, based on a trichloroacetic acid : sodium acetate system, were prepared for acidic (pH ~ 2) and alkaline (pH ~ 10) conditions in conjunction with a neutral, physiologically relevant phosphate buffer (pH ~ 7), in order to monitor the hydrolysis reaction on solid support for the CPG-dT$_5$-rU-5' sequence containing either a phosphate diester or a triester vicinal to the 2'-hydroxyl group.

3.4.1 The Chemical Stability of Phosphate Diesters and Triesters in Neutral Physiological Phosphate Buffer Conditions

In order to duplicate the RNA phosphate diester and triester hydrolysis conditions previously used in the solution phase experiments,[138,139,140] the phosphate buffer (PPB), consisting of 140 mM KCl , 1 mM MgCl$_2$, and 5 mM Na$_2$HPO$_4$, adjusted to pH: 7.2, was injected into the synthesizer column containing the CPG-dT$_5$-rU-5' sequence with the 5'-terminal ribouridine phosphate *triester* after the 3-min detritylation step for 5, 15, 30, 60 and 120 minute exposure times. The column was subsequently washed with 15 mL of H$_2$O, 15 mL of MeOH and 60 mL of MeCN followed by drying with argon and the synthesis (dT$_6$ extension) was continued as previously described. The resulting products were cleaved from the support under aqueous ammonia conditions for 3 hours at 55°C followed by evaporation of the reagent and extraction of the crude product in water. HPLC and analytical PAGE analyses showed the formation of the linear cleavage product (**3.3**) and the branched products (**3.1** and **3.2**) from the hydrolysis reaction (**Figure 3.6**). The analysis did not indicate the formation of any additional oligonucleotide products. Furthermore, the cleavage reaction (monitored by measuring the appearance of dT$_{11}$, **3.3**) was similarly found to be relatively slow, with only 25% of dT$_{11}$ formed after the 2 hour treatment. Alternatively, RP-IP HPLC analysis of the reaction time points indicated that the isomerization reaction occurred quickly, resulting in 30% conversion to the

regioisomer branched oligonucleotide byproduct (**3.2**) after only 5 minutes exposure to buffer (**Figure 3.6**).
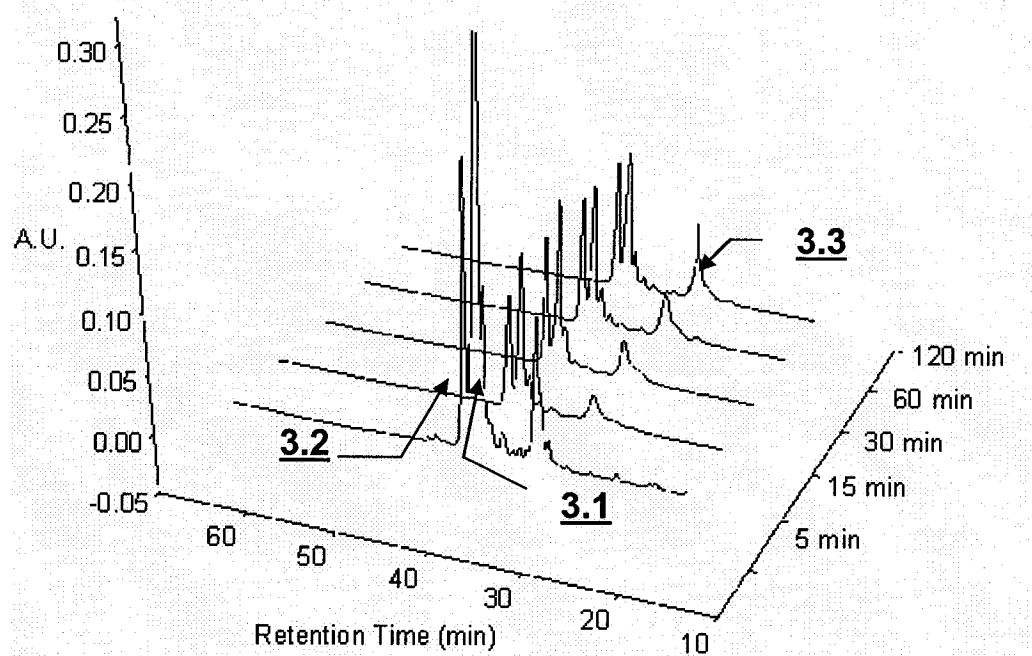
Figure 3.6: The analysis after the hydrolysis reaction with neutral buffer of the ribouridine branchpoint phosphate triester by RP-IP HPLC with an elution gradient up to 13% MeCN in 0.1 M TEAA for a total run time of 70 min. The analysis indicates the separate regioisomers **3.1** and **3.2** and the cleavage linear 11-mer byproduct, **3.3** without the formation of additional side-products during the 2 h reaction.

The same experiment was performed on the phosphate *diester* species by treating the CPG-dT$_5$-rU-5' sequence with the decyanoethylation mixture[40], followed by the detritylation (3 min) in order to release the vicinal 2'-hydroxyl group. The physiological phosphate buffer was again injected into the column. At specific time intervals of 5, 15, 30, 60 and 120 minutes, a sample of the oligonucleotide bound support was washed and dried extensively prior to continuation of the solid phase synthesis with dT$_6$, from the liberated 5'- and 2'-hydroxyl groups. The samples were cleaved from the support, evaporated to yield crude pellets and extracted in sterile water prior to the analysis. The analyses with HPLC and analytical PAGE showed a much slower rate of cleavage than for the triester case (**Figure 3.7**). In fact, since we observed similar recoveries (*ca.* 20%)

of the linear cleavage side-product at each time interval, we assume that the more labile 2',3'-phosphodiester linkage in **3.1** is slowly being cleaved during our analyses (section 4.4.1). This would result in the release of dT_{11} with an internal rU unit. This was also seen in the absence of buffer and during the short exposure time of the phosphate diester to the acid detritylation conditions (**Figure 3.4**, 5 min). The analysis with RP-IP HPLC, showed that the phosphate diester backbone is resistant to the isomerization reaction under neutral conditions, without noticeable isomerization occurring even after 60 minutes of treatment (**Figure 3.7**). These results are consistent with the significant hydrolytic stability of phosphodiester bonds in RNA ($t_{1/2}$ ~100 years).[137]
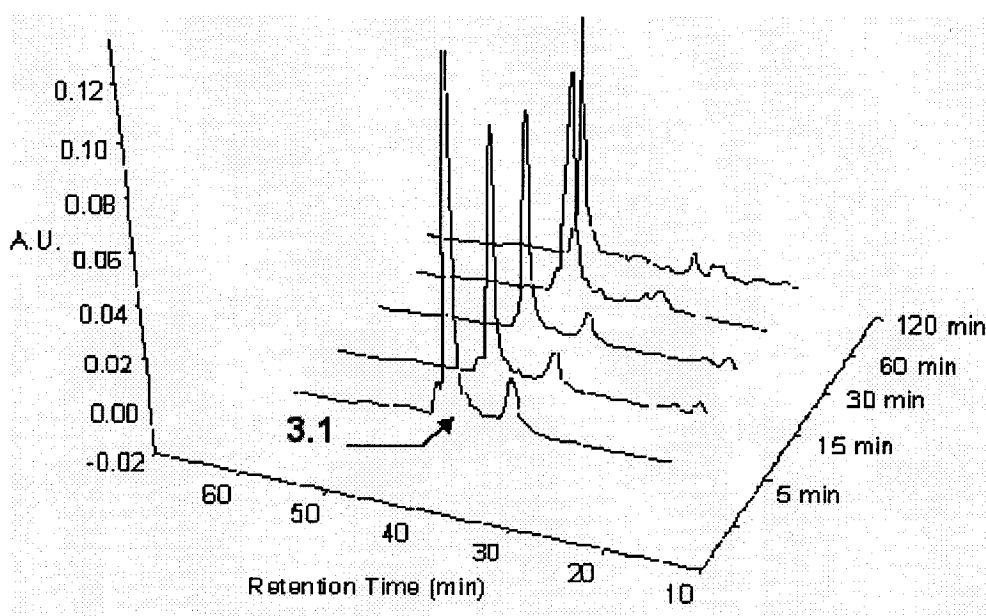
Figure 3.7: The analysis after the hydrolysis reaction with neutral buffer of the ribouridine branchpoint phosphate diester by RP-IP HPLC with an elution gradient up to 13% MeCN in 0.1 M TEAA for a total run time of 70 min. The analysis indicates the intact product, **3.1**, without the increasing formation of the linear cleavage side-products during the 2 h reaction.

Furthermore, the change in reactivity and the absence of branched regioisomers in the diester case (when compared to the triester experiments) again shows that CNEt phosphate protecting groups are not removed during solid-phase synthesis (**Figure 3.8**). In neutral hydrolytic conditions, the phosphate triester backbone is thus equally prone towards rapid isomerization (seconds) as under acidic conditions, but the oligonucleotide

cleavage reaction (25% cleavage after 2 hours) is significantly slower. This is in part due to the requirement of an acid/base catalysis for the hydrolysis reaction of phosphate triesters and diesters.[136,138]
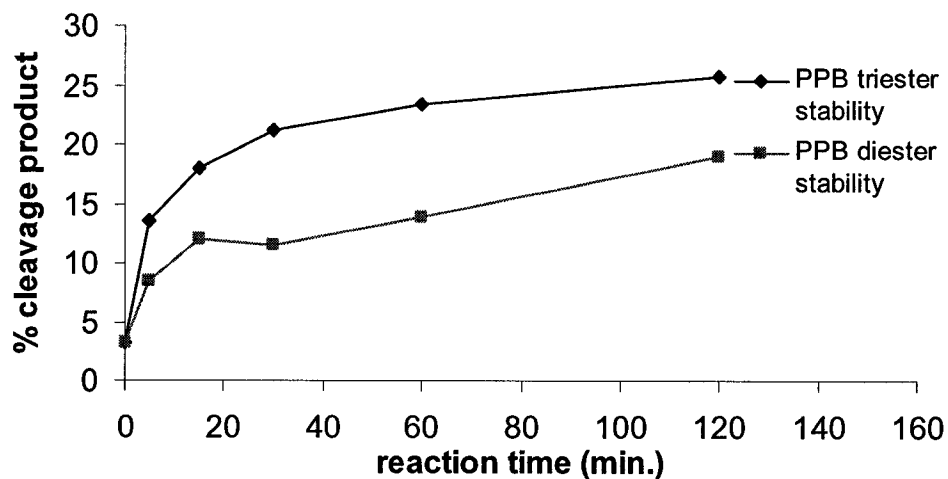
Figure 3.8: The relative stability of the ribouridine phosphate triesters and diesters during the neutral physiological phosphate buffer conditions (*i.e.* 140 mM KCl , 1 mM MgCl$_2$, and 5 mM Na$_2$HPO$_4$, adjusted to pH: 7.2). The plot indicates the extent of the cleavage reaction of the branchpoint 2',3'-phosphodiester linkage.

3.4.2 The Chemical Stability of Phosphate Diesters and Triesters with Aqueous Carboxylic Acidic Buffer Conditions

The hydrolysis reaction of RNA phosphate diesters and triesters *in-vivo* has been shown to be largely accelerated by the acid/base catalytic activity of ubiquitous enzymes such as the ribonucleases and ribozymes.[134,147] An aqueous buffer system of trichloroacetic acid : sodium trichloroacetate was prepared in order to monitor the influence of the acid and base catalysis with the hydrolysis reaction of RNA phosphate triesters and dieters on solid support.

A carboxylic acid buffer, CAB, consisting of a 1 mM solution of 3:1 (v/v) trichloroacetic acid : sodium trichloroacetate pH: 1.4, was prepared in order to determine the stability of the phosphate triester vicinal to a 2'-hydroxyl group. The CPG-dT$_5$-rU-5' sequence was subjected to the CAB conditions for reaction time points of 5, 15, 30, 60 and 120 minutes. Each reaction was quenched and flushed sequentially with 15 mL of CH$_2$Cl$_2$, 15 mL of MeOH and 60 mL of MeCN followed by drying with argon prior the

solid phase synthesis with dT$_6$, from the liberated 5'- and 2'-hydroxyl groups. The products were cleaved from the succinyl linked solid support with the 3 hours treatment with aqueous ammonia at 55°C, and this was followed by evaporation of the reagent and extraction of the products in water. The analysis of the crude reaction mixture at each of the time points was achieved by HPLC (**Figure 3.9**) and 24% analytical PAGE (data not shown) which indicated the presence of the branched products, (**3.1** and **3.2**) and the linear cleavage side-product, (**3.3**) without additional oligomer side-products. The cleavage reaction of the phosphate triester linkage occurs to 50% completion after a 2 hour reaction time with the carboxylic acid buffer (**Figure 3.9**). Furthermore, the analysis of the crude reaction time points with RP-IP HPLC confirmed a fast isomerization reaction of the branchpoint synthon phosphate triester in acid catalyzed buffer conditions with the formation of up to 40% of the regioisomer (**3.2**: 2',5' and 3',3' branchpoint connectivity) after only 5 minutes exposure (**Figure 3.9**). Thus, the phosphate triester cleavage reaction was found to be acid catalyzed in carboxylic acid buffer (50% cleavage in 2 hours) relative to the neutral phosphate buffer (25% cleavage after 2 hours) conditions, resulting in a *ca.* 2-fold enhancement in the reactivity during the aqueous acid catalyzed conditions.
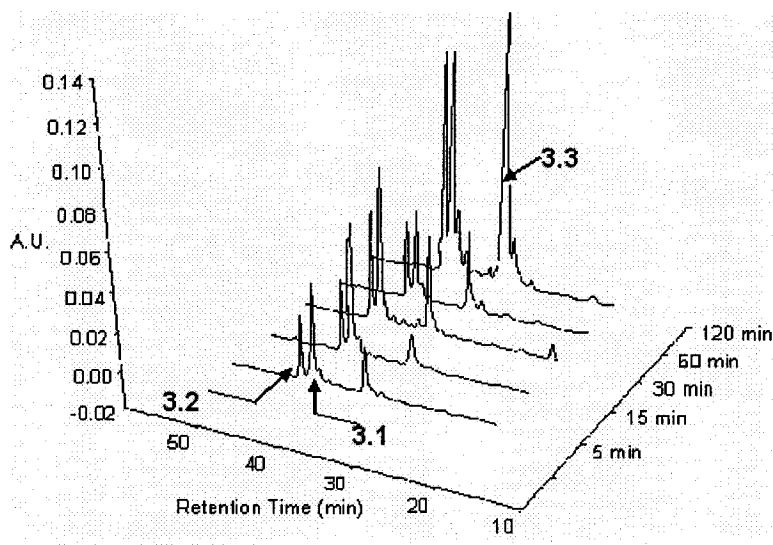
Figure 3.9: The analysis after the hydrolysis reaction with the carboxylic acid buffer of the ribouridine branchpoint phosphate triester by RP-IP HPLC with an elution gradient up to 13% MeCN in 0.1 M TEAA for a total run time of 70 min. The analysis indicates the separate regioisomers **3.1** and **3.2** and the cleavage linear 11-mer byproduct, **3.3** without the formation of additional side-products during the 2 h reaction.

63

Similarly, the putative sequence, (CPG-dT$_5$-rU-5') was synthesized with the terminal 2',5'-diMMT protecting groups for decyanoethylation prior to acid detritylation, and treatment with CAB for time point reactions during 5, 15, 30, 60 and 120 minutes followed by washing and drying of the oligomer bound support for the automated assembly with the dT$_6$ sequences, from the liberated 5'- and 2'-hydroxyl groups. These were cleaved from the support, evaporated and extracted in sterile water prior to the analysis. The HPLC and PAGE analyses for the phosphate diester stability experiments indicated 31% cleavage product after 120 minutes for the cleavage reaction (**Figure 3.10**). The phosphate diester was also found to be resistant to the isomerization reaction, without noticeable isomerization (**Figure 3.10**).
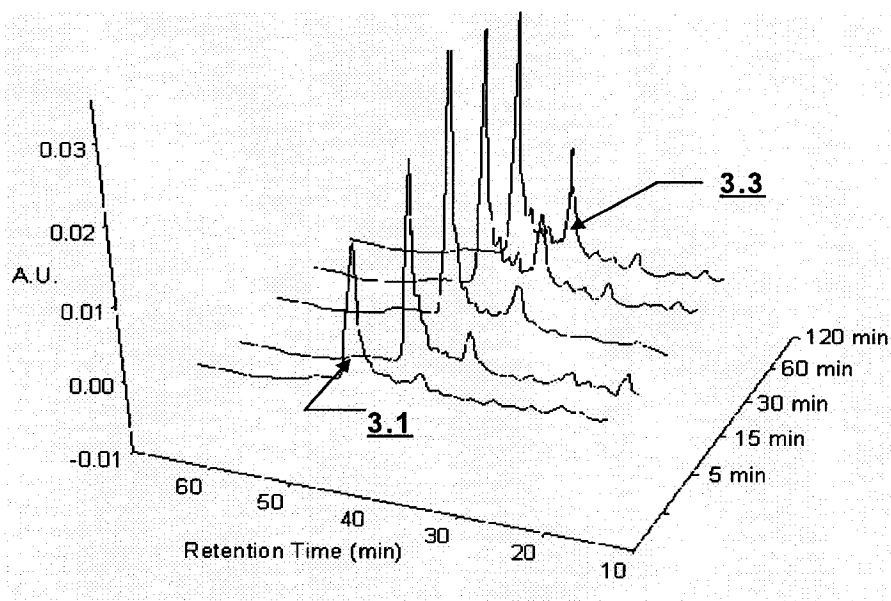
Figure 3.10: The analysis after the hydrolysis reaction with the carboxylic acid buffer of the ribouridine branchpoint phosphate diester by RP-IP HPLC with an elution gradient up to 13% MeCN in 0.1 M TEAA for a total run time of 70 min. The analysis indicates the intact product, **3.1** the cleavage linear 11-mer byproduct, **3.3** without the formation of additional side-products during the 2 h reaction.

The reactivity of the phosphate triester in the CAB buffer was found to be impeded in aqueous conditions resulting in a 1.5-fold decrease (50% cleavage after 2 hours) relative to the anhydrous acid catalyzed detritylation conditions (80% cleavage after 2 hours) (**Figure 3.11- A**). It can be rationalized that the aqueous conditions could generate a negative kinetic solvent effect in which the addition of water impedes the

64

transesterification reaction or more specifically the initial attack of the vicinal 2'-hydroxyl.[132b] Additionally, it can also be suggested that reactions in the solid phase are less tolerant with the hydrolytic conditions resulting in a decrease in the reactivity.[148] Comparison of the relative rates for the cleavage reaction also provided a faster cleavage reaction for the phosphate diester in acid catalyzed anhydrous detritylation conditions (t ½ ~ 2 hours) relative to the aqueous carboxylic acid buffer conditions (30% cleavage after 2 hours) (**Figure 3.11- B**).
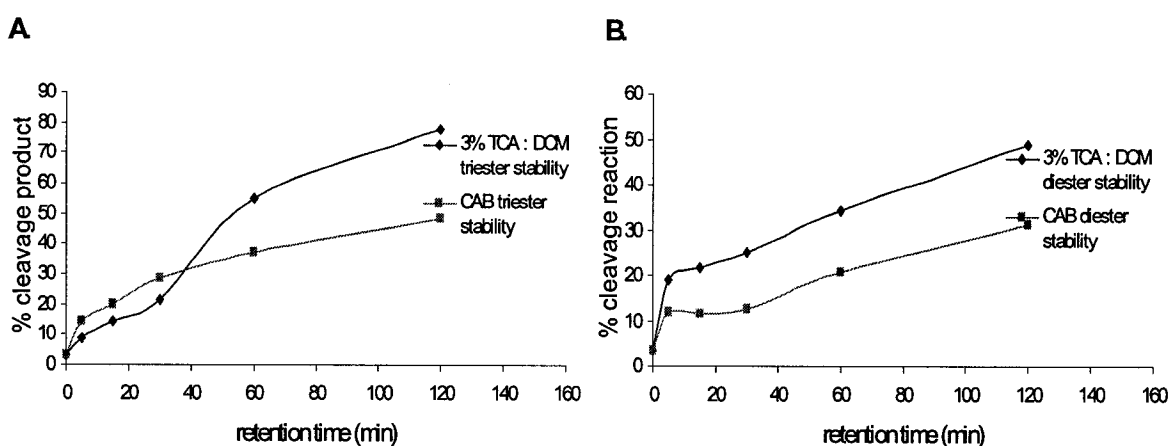
Figure 3.11: The extent of the cleavage reaction with aqueous and anhydrous acid catalyzed conditions of the ribouridine phosphate **A.** diester and **B.** triester linked to a pentameric polythymidylate sequence while on the solid support.

Again, the change in the reactivity (*i.e.* triester *vs.* diester) and the absence of the branched regioisomer side-product in the diester case indicates that the CNEt phosphate protecting groups were not removed by the treatment with CAB (**Figure 3.12**). In addition, under CAB conditions, the phosphate triester backbone is equally prone towards rapid isomerization (seconds) but the oligonucleotide cleavage reaction is significantly slower ($t_{1/2}$ ~ 2 hours) relative to the anhydrous conditions. Furthermore, these observations are significantly different (*i.e.* slower reactivity) than the in-solution data for the chemical stability of CNEt protected phosphate triesters.[145] This decrease in reactivity with the cleavage reaction of phosphate triesters, can in part be associated with the concomitant release of 2-cyanoethanol (**Figure 3.1- c**) which leads to the slow and minimal formation of the 11-mer oligothymidylate sequence, (3.3) (**Figure 3.9 and 3.10**).
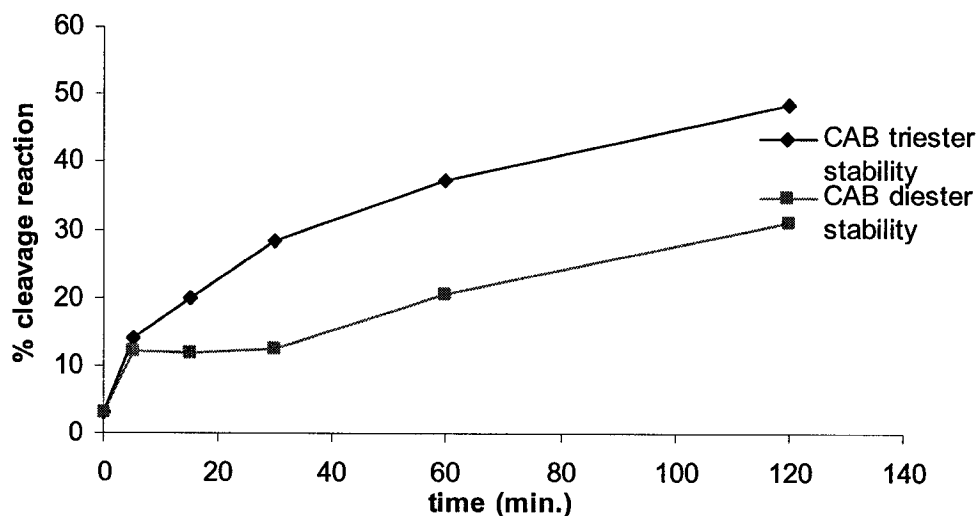
Figure 3.12: Relative stability of ribouridine phosphate triesters and diesters during the CAB conditions [1 mM solution of 3:1 (v/v) trichloroacetic acid : sodium trichloroacetate pH: 1.4]. The plot indicates the extent of the cleavage reaction for a 5'-terminal ribouridine linked to a dT₅ sequence while on the solid support.

3.4.3 The Chemical Stability of Phosphate Diesters and Triesters with the Alkaline Buffer Conditions

An aqueous alkaline buffer, AB, was prepared from a 1 mM solution of sodium trichloroacetate, pH: 12.1 for the treatment of the test sequence, (*i.e.* CPG-dT₅-rU-5') while on a solid support during 5, 15, 30, 60 and 120 minute reaction times. The solid support for each trial experiment was subsequently washed successively with 15 mL of MeOH, 15 mL CH_2Cl_2 and 60 mL of MeCN to ensure complete removal of the buffer. The wet CPG was subsequently dried with argon and the solid phase synthesis was continued with the dT₆ sequences from the liberated 5'- and 2'-hydroxyl groups. The crude reaction time points were analyzed with HPLC and 24% denaturing analytical PAGE which indicated 52% cleavage to yield the linear side product after a 2 hour exposure to the alkaline buffer conditions (**Figure 3.15**). The analysis also confirmed the formation of a lower molecular weight oligomer species which may result from the cleavage of the CNEt protecting group (**Figure 3.1-c**) and the formation of linear oligomer byproducts (**3.4** to **3.6**).
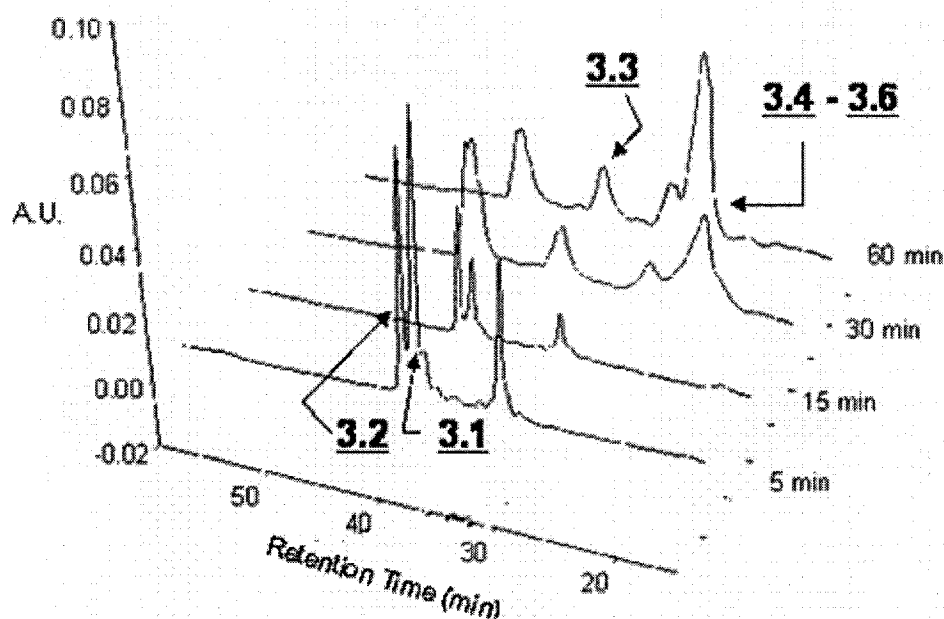
Figure 3.13: The analysis of the AB hydrolysis reaction of the branchpoint phosphate triester by RP-IP HPLC with the eluent gradient up to 13% MeCN in 0.1 M TEAA and a total run time of 70 min. The analysis indicates the branched regioisomers **3.1** and **3.2**, the linear cleavage product, **3.3** and the formation of a lower MW oligomer species (with **3.4** to **3.6** as possible structures) with increasing exposure time with the AB conditions. The exact structure of this byproduct has not been confirmed.

Similarly, with a slightly modified experimental procedure, (*i.e.* decyanoethylation of the 2',5'-diMMT ribouridine phosphate triester followed by detritylation) the isolated phosphate diester with the vicinal 2'-hydroxyl group was exposed to the AB conditions for fixed time point reactions in order to determine the extent of the hydrolysis reaction in the base-catalyzed conditions. The sample reactions were washed, dried and the automated syntheses were continued with the dT_6 sequences from the liberated 2' and 5'-hydroxyl groups of the branchpoint unit. The trial experiments were cleaved from the support with alkaline treatment and evaporated prior to extraction of the crude oligomer samples with sterile water. The analyses of the crude samples by PAGE and HPLC indicated that the cleavage reaction occurred to 55% completion after 2 hours (**Figure 3.15**). The analysis by RP-IP HPLC confirmed a single branched product, **3.1**, that was resistant to the isomerization reaction after 60 minute treatment with the AB conditions (**Figure 3.14**).

67
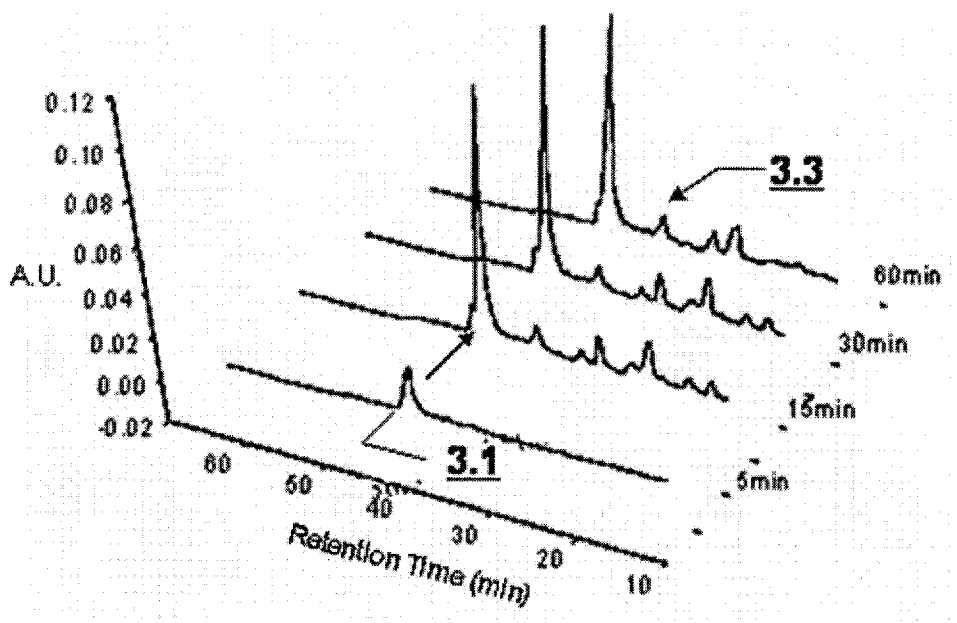
Figure 3.14: The analysis of the AB hydrolysis reaction of the branchpoint phosphate diester by RP-IP HPLC with the eluent gradient up to 13% MeCN in 0.1 M TEAA and a total run time of 70 min. The analysis indicates the single regioisomer **3.1** by RP-IP HPLC with little accumulation of the cleavage linear 11-mer byproduct, **3.3.**

The analyses confirmed that in the AB phosphate triester and diester hydrolysis reactions, the extent of the cleavage reaction was found to be relatively similar *i.e.* $t_{1/2} \sim 2$ hours **(Figure 3.15)**. This observation provided additional evidence that the decyanoethylation process is selective with the alkaline conditions required during the post-synthesis work-up procedure.[21]
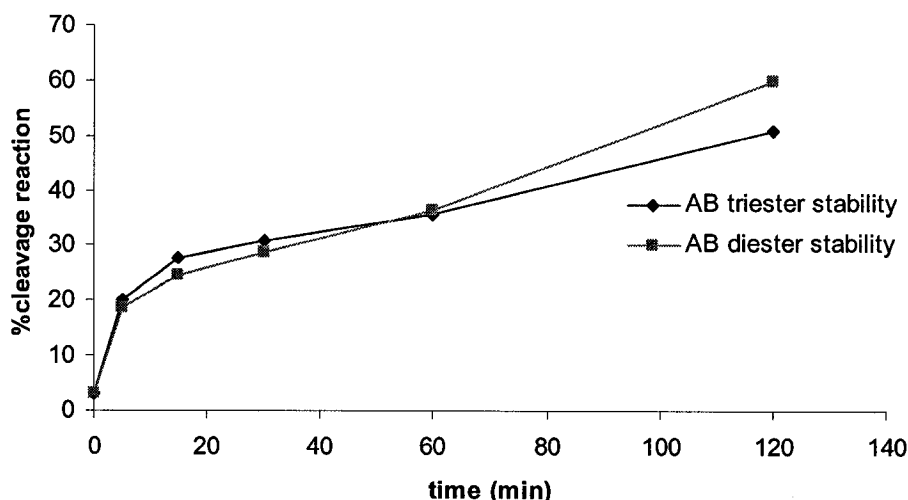
Figure 3.15: The relative stability of ribouridine phosphate triesters and diesters during the AB conditions (*i.e.* 1 mM solution of sodium trichloroacetate buffer adjusted to a basic pH: 12.1). The plot indicates the extent of the cleavage reaction for a 5'-terminal ribouridine linked to a pentameric polythymidylate sequence while on a solid support.

3.5 CHARACTERIZATION OF THE BRANCHPOINT REGIOISOMERS

3.5.1 Characterization of the branchpoint regioisomers by RP IP HPLC

To characterize the branchpoint regioisomers (**3.1** and **3.2**) by RP IP HPLC, the branched regioisomers were prepared as a mixture and in their pure form. This was performed by the previously described procedures, in which the mixture of regioisomers were generated during the detritylation reaction (3% TCA:DCM, 200 seconds) to remove the MMT protecting groups from the ribouridine branchpoint synthon, and the pure regioisomers were synthesized separately, with branchpoint synthons **2.5** and **2.6** and with the key decyanoethylation step prior to removal of the MMT protecting groups.

The sequences were analyzed by RP-IP HPLC, and the pure products confirmed the structures of **3.1** and **3.2** within the mixture of regioisomers (**Figure 3.16**). All sequences were purified by RP-IP HPLC and their identities confirmed with MALDI-TOF MS, (*i.e.* calc. 5417 and found 5416, **3.1**, and 5418, **3.2**). Also, **3.1** and **3.2** were not substrates of the yeast debranching enzyme, a phosphodiesterase that is highly specific for branched oligonucleotides containing vicinal 2'-5' and 3'-5' phosphodiester linkages (absent in **3.1** and **3.2**).[149]
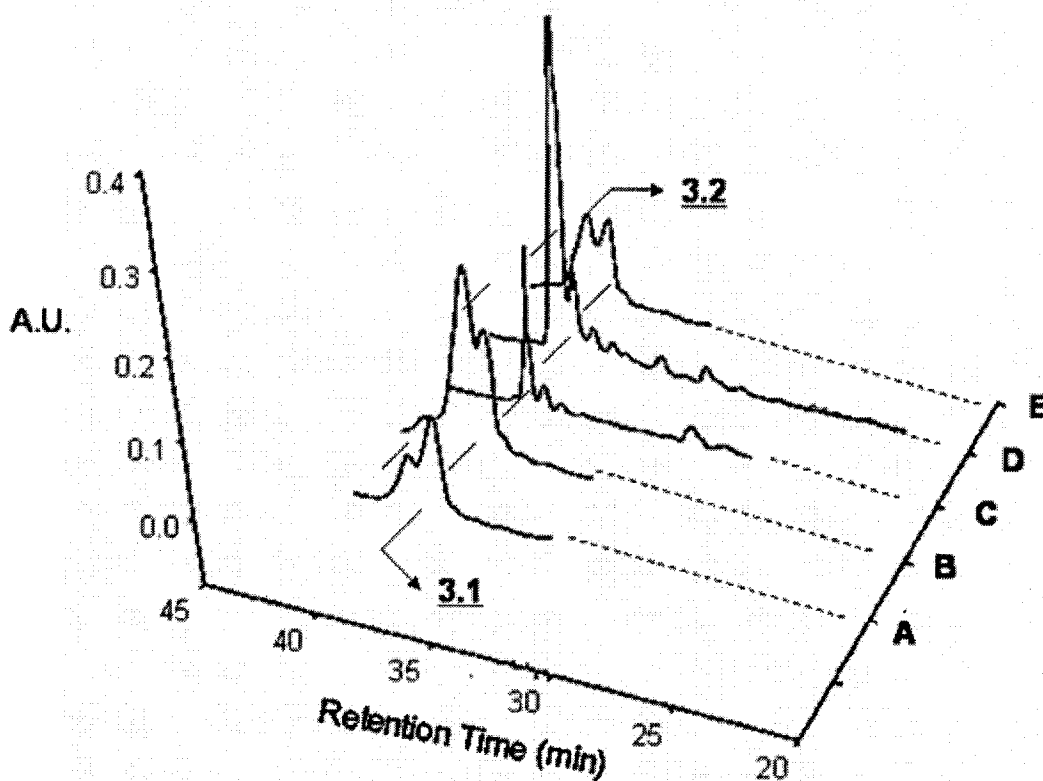
Figure 3.16: The RP IP HPLC analysis with eluent gradient up to 13% MeCN in 0.1 M TEAA for the characterization of **3.1** and **3.2**. The analysis consists of **A.** the synthesis of the mixture of regioisomers **3.1** and **3.2** generated with **2.5** as the branchpoint synthon, **B.** the synthesis of the impure regioisomers **3.1** and **3.2** generated with **2.6** as the branchpoint synthon, **C.** the synthesis of pure sequence **3.1** with **2.5** as the branchpoint synthon, **D.** the synthesis of the pure sequence **3.2** with **2.6** as the branchpoint synthon and **E.** a standard solution generated by mixing **3.1** and **3.2** in a 1:1 ratio.

3.6 CONCLUSIONS

The stability of a ribonucleoside 3'-phosphate triester and diester on a solid support has been monitored by HPLC and PAGE over a range of conditions. The experiment was performed on a rU(2'OH)pdT$_5$ sequence linked to a CPG support, and the extent of the cleavage/transesterification reaction following branch DNA synthesis was assessed by HPLC and PAGE analyses. The structure of the branchpoint isomerization products were confirmed by MS, RP IP HPLC. The cleavage and isomerization reactions of phosphate

70

triesters were faster than those observed with phosphate diester backbones. This provided evidence that the phosphate CNEt protecting group is indeed resistant to the solid phase synthesis protocol, and labile to the alkaline conditions required during deprotection of the oligonucleotide chain. Moreover, the reactivity of phosphate triesters observed on the solid supports was slower than those of analogous reactions in solution-phase.

In the context of branched oligonucleotide synthesis, these results confirm that unmasking the phosphate triester (*i.e.*, decyanoethylation) must precede the removal of the adjacent 2'-protecting group in order to maintain the integrity (*i.e.* prevent isomerization of) the branchpoint. This is also relevant in the synthesis of biologically useful oligoribonucleotide sequences such as the siRNAs[124] and with the synthesis of branched RNA and DNA for the splicing[43] and mechanistic[160,161] studies involved in the maturation process of mRNA. These studies are also important towards understanding the stability of the unsual "RNA-X" species that has been detected during pre-mRNA splicing in vitro, in which a phosphate triester is present and vicinal to a 2'-OH group. Robert Donga, a current Ph.D. student in our lab and others[147] have synthesized this unstable phosphate triester in order to study its stability in solution.

CHAPTER 4: SYNTHESIS AND PROPERTIES OF BRANCHED OLIGONUCLEOTIDES WITH VICINAL 2',5' AND 3',3', OR, 2',3' AND 3',5'-PHOSPHODIESTER LINKAGES

4.1 INTRODUCTION

The discovery and structure elucidation of nucleic acid lariats, from the extract of nuclear polyadenylylated RNA in HeLa cells, led to several studies directed towards understanding their role during the processing of pre-mRNA.[43,44] Y-shaped RNAs have also been detected in mRNA trans-splicing reactions of parasitic nematodes,[43,46] whereas Y-shaped DNA/RNA chimeras, better known as multi-copy single-stranded DNA, occur in *Myxoccocus xanthus*, a species of soil-dwelling bacteria, and within certain strains of the bacterium *Escherichia coli*.[47] Evidence for the catalytic potential of spliceosomal small nuclear RNAs was recently provided by Manley and co-workers, which furnished ample evidence for the formation of an unprecedented branched RNA structure, termed RNA-X.[140,150] The natural occurrence of these has prompted many groups to investigate facile synthetic methodologies for creating such compounds, or compounds reminiscent of the lariat RNA itself (*i.e.* Y-shaped oligonucleotides) in order to study their properties and potential applications. For example, synthetic branched and hyperbranched nucleic acids, or DNA dendrimers have been explored as delivery vehicles for (antisense) drugs,[151] as tools for investigating nucleic acid structure and function,[152] as diagnostic reagents for radioactive or fluorescent signal amplification via oligonucleotide-based arrays,[153] and as probes to examine the substrate specificity of the lariat RNA debranching enzyme.[49]

The chemical synthesis of biologically relevant branched nucleic acids (bNAs) has been accomplished via the solid-phase[25], template-directed chemical and enzymatic ligation[35], or through the use of deoxyribozymes[36]. The more conventional solid-phase process can be performed by the *convergent approach* introduced by Zabarylo and Damha in 1989 (**Figure 1.8**).[25] This method utilizes a branchpoint 2',3'-bis-phosphoramidite synthon, **1.2**, capable of coupling with two adjacent (and identical) oligonucleotide strands through their terminal 5'-hydroxyls, thus generating a V-like structure with vicinal 2',5'and 3',5'-phosphodiester bonds. Chain extension from the apex

of this V-like molecule generates a Y-shaped NA.[25,37] Alternatively, such bNA species can be prepared via a stepwise *divergent approach* (**Figure 1.9**) allowing for separate and regiospecific synthesis of asymmetric bNAs.[38,39,46] The important feature of this method is the use of a branchpoint synthon with chemically non-equivalent protecting groups at the branching positions, **1.3** which allows for separate and regiospecific chain extension and branching reactions to yield the asymmetric bNA.[40]

The groups of Tomalia, Newkome and Frechet have also designed a methodology for the convergent and divergent synthesis of hyperbranched or dendritic-like structures.[154] Similarly, hyperbranched nucleic acids, h-bNAs, or "dendritic" DNA oligonucleotides of finite length and base composition have been assembled on a solid support through the use of these *convergent*[41] (**Figure 1.8**) or *divergent*[42] (**Figure 1.9**) approaches by repeating the branching and chain elongation steps several times. However, applications of this method towards the synthesis of bNAs and h-bNAs of mixed backbone composition (DNA/RNA chimeras) and unnatural branchpoints (*e.g.* vicinal 2',3' and 3',3' linkages) have been limited. Such synthetic branched analogues could be particularly useful for probing the structural requirements of branch point recognition during splicing and developing inhibitors of the lariat debranching enzymes.[27,28,40,161]

4.2 PROJECT OBJECTIVES AND METHODOLOGY

In this study, the automated *divergen-growtht* synthesis of homopolymeric branched and hyperbranched oligouridylate sequences is described through the use of 5',2' and 5',3'-diMMT rU phosphoramidite synthons **2.5** and **2.6**. Novel features of the method described are the simultaneous removal of two protecting groups (MMT) at the branchpoint, simultaneous divergent growth of two 'arms' to rapidly assemble Y-shaped NA structures, and the introduction of 'un-natural' linkages at the branchpoint, *i.e.* 2',5' and 3',3' or 2',3' and 3',5' phosphodiester bonds (**Figure 4.1- B** and **C**). The regiochemistry at the branchpoint can be controlled by the conversion of the phosphate triester backbone to the more acid-resistant phosphate diester to prevent the isomerization/cleavage of the backbone during bNA assembly (see Chapter 3).[33,34] The divergent synthesis of these Y-shaped bNAs will be compared to the more traditional *convergent approach* previously established in our laboratory.[25,37] It was rationalized that

yields and purity of bNA and h-bNA species prepared by a divergent approach could be vastly improved over those previously achieved by convergent methods, where branching is highly influenced by the proximity of neighboring NA chains.[41] This divergent approach provides bNAs with unique branchpoint connectivites (*i.e.* 2',5' and 3',3'or 2',3' and 3',5') that may have different stabilities towards chemical and enzymatic hydrolysis relative to the native bNAs (*i.e.* 2',5' and 3',5' linkages) (**Figure 4.1-** A and C). In fact, these studies led to the identification of some inhibitors of the yeast debranching enzyme (yDBR), making these compounds potential good choices for future enzyme co-crystallization and X-ray analysis of the yDBR active site structure.
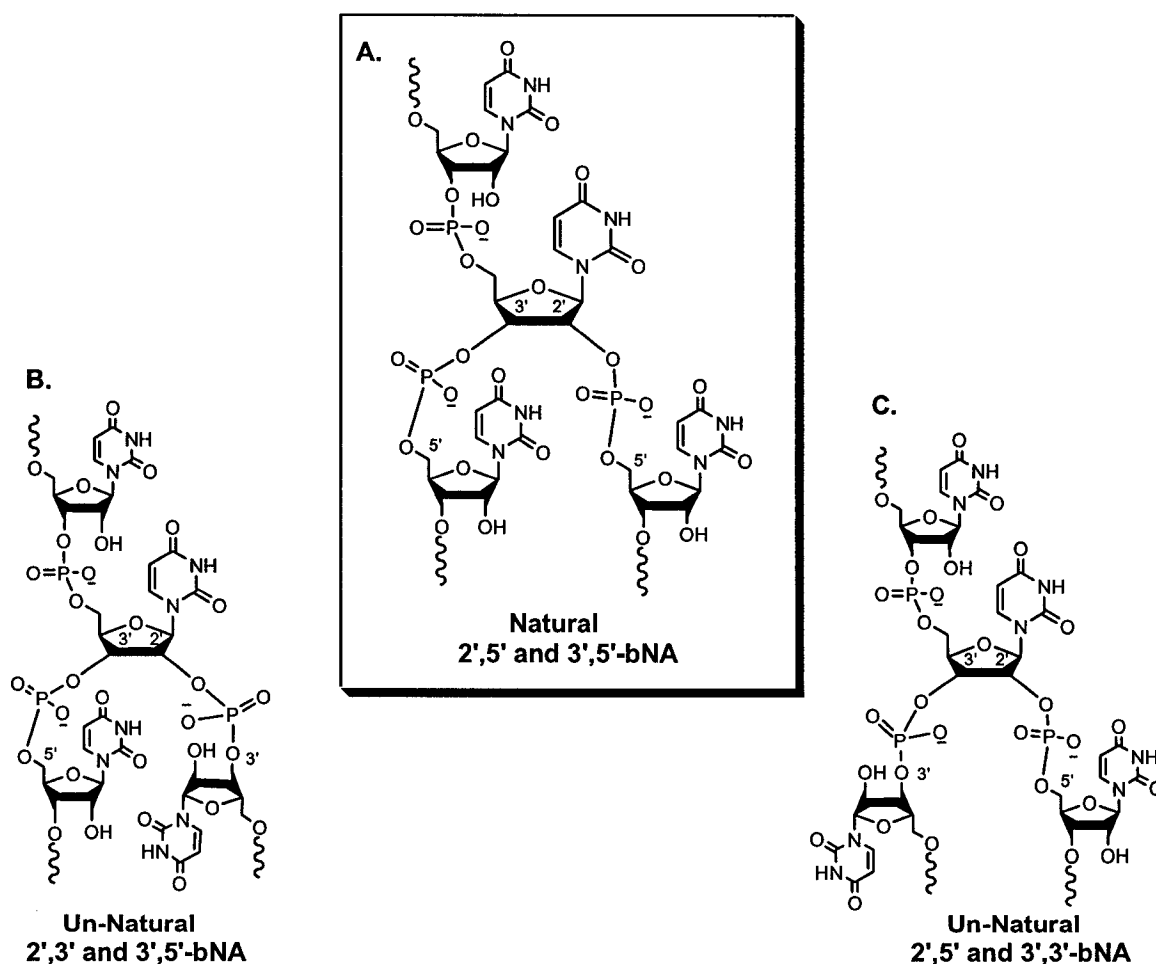
Figure 4.1: Structure and branchpoint diversity of bNAs synthesized for this study.

74

4.3 SOLID PHASE SYNTHESIS OF BRANCHED AND HYPERBRANCHED OLIGOURIDYLIC ACID

4.3.1 Solid Phase Synthesis and Characterization of Branched Oligouridylate Sequences containing 'natural' (2',5' and 3',5') and 'un-natural' (2',5' 3',3' or 2',3' 3',5') Branchpoint Linkages

Phosphoramidite derivatives **2.5** and **2.6** were used as branchpoint synthons for the divergent solid phase synthesis of branched oligouridylate sequences (**Figure 4.2- A**). Synthesis was performed on a 1 μmol reaction scale using CPG solid support (500 Å pores; succinyl linker) functionalized with 5'-DMT rU.[19] A linear RNA strand (rU$_9$) was first assembled using commercially available 5'-DMT-2'-TBDMS-rU-3'-phosphoramidite. The diMMT rU branchpoint synthons, **2.5** (or **2.6**) were coupled next. The phosphate's CNEt protecting groups were cleaved (4:6 v/v NEt$_3$:MeCN, 90 min, 22°C) to produce an RNA chain containing a more stable phosphate diester backbone.[40] The synthesis column was re-attached to the synthesizer and an anhydrous detritylation step (200 seconds, 3% TCA:DCM) was performed to remove both MMT protecting groups at the branchpoint. Simultaneous chain extension (2 × rU$_8$) from the 5' and 2' (or 3') ends produced a 26-nt Y-shaped oligouridylate sequence with vicinal 2',5'/3',3' (**4.1**), or 2',3'/3',5' (**4.2**) phosphodiester linkages at the branchpoint (**Figure 4.2- A**). The sequences were cleaved from the solid support by treatment with 3:1 (v/v) NH$_4$OH:EtOH (55°C, 4 h). After evaporation, the pellet was re-suspended in TREAT-HF (48 h, 22°C) to cleave the 2'-TBDMS protecting groups. The crude bNAs were precipitated from the reaction mixture with n-BuOH to recover *ca.* 120 optical density units (OD at A$_{260}$) or 50% of the bNAs. The crude bNA sequences were analyzed and purified[155] by AE HPLC or denaturing PAGE which indicated 65% and 70% purity for **4.1** and **4.2**, respectively (**Figure 4.5**; PAGE analysis of all sequences synthesized in this study). The major failure sequences (30-35% of the crude material) were identified as the N-1, N-2, N-3 *etc.* sequences (5-10% yield), an indication of incomplete coupling reactions during linear RNA chain growth. The more predominant byproduct (*ca.* 15% yield), was a linear rU$_{18}$ sequence generated as a result of unsuccessful coupling at the hindered 3' (or 2')-hydroxyl position (**Figure 4.3- A and B**). Any unreacted 3' (or 2'OH) groups would be

acylated (3' or 2'-OAc) in the ensuing capping step. An additional lower molecular weight oligomer species (*i.e.* 9-mer; *ca.* 10% yield) was also observed by AE HPLC (**Figure 4.3- A and B**) and is likely derived from the fragmentation of rU_{18} assisted by the internal 3' (or 2')-OH released during the ammonia treatment. Alternatively, it is also conceivable that a small amount of bNAs **4.1** and **4.2** are "debranched" at the more energetically strained linkages (*i.e.* 3',3' or 2',3') during the basic ammonia treatment resulting in the release of an additional rU_8.

Next, a 26-nt Y-shaped RNA (**Figure 4.2- B, 4.3**) having 'natural' vicinal 2',5' and 3',5' linkages was synthesized by the conventional *convergent approach* under optimized reaction conditions.[25,37,155] That is, a solid support with high density loading (*i.e.* 50-80 μmol/g) was used along with a low concentration of branching bisphosphoramidite synthon **1.2** (0.02 - 0.03 M) to maximize the probability of coupling **1.2** to neighboring CPG-bound oligonucleotide chains.[25,37,155] The support bound bNA was cleaved, deprotected and precipitated prior to HPLC/PAGE analyses.[155] The crude product was recovered in 21% yield (50 OD units), and constituted 75% of the crude mixture. The major failure sequence (23%) was confirmed as the 2'/3'-phosphorylated linear 18-nt regioisomers (**Figure 4.3- C**). All bNAs were purified by AE HPLC and/or PAGE purification protocols prior to desalting with Nap® 10 columns (*i.e.* Sephadex® G-25 size exclusion columns) for characterization by MALDI-TOF MS (*i.e.* Calcd. 8273, found **4.1**: 8296; **4.2**: 8273; and **4.3**: 8273).

A.

CPG: oligouridylate derivatized
500Å LCAA CPG

couple with
diMMT rU amidite
(0.15 M, 30 min)

2.6 **2.5**

or

decyanoethylate
4 : 6 v/v NEt₃ : MeCN
90 min., r.t.

or

i. di-detritylation
(200 sec. 3% TCA : DCM)

ii. chain extension

: oligouridylate
(rU amidites)

or

cleavage and deprotection

or

(2'5')   UUU UUU UUU-3'-OH
HO-5'-UUU UUU UU—(3'5')—rU
(3'3')   UUU UUU UU-5'-OH

4.1

(2'3')   UUU UUU UU-5'-OH
HO-5'-UUU UUU UU—(3'5')—rU
(3'5')   UUU UUU UUU-3'-OH

4.2

77

Figure 4.2: Solid phase synthesis of 26-nt oligouridylate bNAs by the **A.** novel *divergent-growth approach* with diMMT rU amidite branchpoint synthons **2.5** and **2.6** and **B.** traditional *convergent-growth approach* with bisphosphoramidite branchpoint synthon **1.2**. The methods produced bNAs with unnatural (**4.1** and **4.2**) and native (**4.3**) branchpoint linkages.

78

A.

0.7
0.6
0.5
0.4
A.U. 0.3
0.2
0.1
0
-0.1

HO-5'-UUU UUU UU$^{(3'5')}$rU$^{(2'5')}$UUU UUU UUU-3'-OH

$^{(3'3')}$UUU UUU UU-5'-OH

U UU UUU UUU-3'-OH

HO-5'-UUU UUU UU$^{(3'5')}$rU$^{(2'5')}$

3'-OH

0   10   20   30   40   50   60   70

P/HO-5'-UUU UUU UUU-3'-OH   Retention time (min)
or
P-O-3'-UU UUU UUU-5'-OH

B.

1.4
1.2
1
0.8
A.U. 0.6
0.4
0.2
0
-0.2

HO-5'- UUU UUU UU$^{(3'5')}$rU$^{(2'3')}$UUU UUU UU-5'-OH

$^{(3'5')}$UUU UUU UUU-3'-OH

2'-OH

HO-5'-UUU UUU UU$^{(3'5')}$rU

$^{(3'5')}$UUU UUU UUU-3'-OH

0   10   20   30   40   50   60   70

P/HO-5'-UUU UUU UUU-3'-OH   Retention Time (min)
or
P-O-3'-UU UUU UUU-5'-OH

C.

0.4
0.35
0.3
0.25
A.U 0.2
0.15
0.1
0.05
0
-0.05

2'-O-P

HO-5'-UUU UUU UUU$^{(3'5')}$rU

$^{(3'5')}$UUU UUU UU-3'-OH

or

HO-5'-UUU UUU UUU$^{(3'5')}$rU$^{(2'5')}$UUU UUU UU-3'-OH

HO-5'-UUU UUU UUU$^{(3'5')}$rU$^{(3'5')}$UUU UUU UU-3'-OH

$^{(3'5')}$UUU UUU UU-3'-OH

$^{(2'5')}$UUU UUU UU-3'-OH

3'-O-P

0   10   20   30   40   50   60   70

Retention Time (min)

Figure 4.3: AE HPLC analysis of crude bNA and byproduct sequences (0-30% gradient of 1 M LiClO$_4$ in H$_2$O; total run time: 70 min). A. 4.1 (65% purity); B. 4.2 (70%); C. 4.3 (75%).

4.3.2 Solid Phase Synthesis and Characterization of the Hyperbranched Oligouridylate Sequences containing 'natural' (2',5'/3',5') and 'un-natural' (2',5'/3',3' or 2',3'/3',5') Branchpoint Linkages

This newly developed *divergent-growth* method was next applied to the synthesis of hyperbranched nucleic acids. Although yields of 4.1 and 4.2 (prepared via a *divergent-growth* method) were similar to those obtained for 4.3 (*convergent-growth*), the divergent synthesis of h-bNAs is in principle, expected to produce significantly increased yields of full length h-bNAs, by removing the steric and geometric constraints of the convergent approach. Also, the ability to use a stoichiometric excess of reagents to "force" branching reactions to near completion is in contrast to the convergent strategy where, as indicated above, special control of reactions conditions are required to ensure maximum branching

(**Figure 4.4**).[25,37,40,41,42] The *divergent-growth* synthesis of 39-mer h-bNAs, **4.4** and **4.5** (**Figure 4.4- A**) was performed on a 1 μmol synthesis scale with a 500 Å succinyl linked CPG support highly loaded with 5'-DMT rU (80 μmol/g).[19] While compounds **4.1-4.3** represent "first generation" dendrons, each with a single branchpoint and having Y-like shape, compounds **4.4** and **4.5** are "second generation" dendrons containing three branchpoints.[156] Thus, branchpoint synthons, **2.5** and **2.6** were added as 0.15 M solutions in anhydrous CH_2Cl_2 and MeCN, respectively, and coupled to the nascent oligouridylate strand for 20-30 minutes. This coupling reaction was followed by decyanoethylation, removal of both MMT groups, chain extension and the entire procedure repeated until the desired h-bNAs, **4.4** and **4.5**, were obtained (**Figure 4.4- A**). These sequences were cleaved from the solid support, desilylated and precipitated over n-BuOH.[155] The crude h-bNAs **4.4** and **4.5** were recovered in yields of 65 OD (20% yield) and 100 OD (35% yield) units, respectively. Analysis of the crude synthesis by AE HPLC and 24% denaturing PAGE indicated 15% and 22% purity for the 2nd generation dendritic h-bNAs, **4.4** and **4.5**, respectively. Further analysis of the reaction byproducts indicated incomplete branchpoint coupling and significant amounts of linear and branched failure sequences (**Figure 4.5**). The synthesis of the h-bNAs was repeated using a larger excess of branching synthons relative to the nucleoside bound CPG. This was achieved by performing the solid phase synthesis on a smaller scale (0.25 μmol), a CPG solid support of lower loading (500 Å; 40-50 μmol/g), higher concentration of the branching (0.15 M or higher) and normal amidite synthons (0.3 M), and extended coupling times (30-40 min). This resulted in the recovery of higher purity h-bNAs, more specifically, 45% and 60% yield recovery for the intact **4.4** and **4.5**, respectively (**Figure 4.5**). The structures of the h-bNAs, **4.4** and **4.5** were confirmed by MALDI TOF MS (Calcd. 12290, found **4.4**: 12292 and **4.5**: 12302).

A.

i. couple with diMMT rU amidite (0.15 M, 40 min)

ii. decyanoethylation

2.6   **2.5**

i. di-detritylation (200 sec. 3% TCA : DCM)

ii. chain extension

⌇⌇⌇ : oligouridylate

cleavage and deprotection

HO-5'-UUU UU$_{(3'3')}$ (2'5')
⌐rU—UUU UU
HO-5'-UUU UU$^{(3'5')}$ $_{(3'3')}$ (2'5')
rU—UUU UUU-3'
HO-5'-UUU UU$_{(3'3')}$ (2'5') $^{(3'5')}$
rU—UUU UU
HO-5'-UUU UU$^{(3'5')}$

4.4

or

HO-5'-UUU UU$_{(3'2')}$ (3'5')
⌐rU—UUU UU
HO-5'-UUU UU$^{(3'5')}$ $_{(3'2')}$(3'5')
rU—UUU UUU-3'-OH
HO-5'-UUU UU$_{(3'2')}$ (3'5') $^{(3'5')}$
rU—UUU UU
HO-5'-UUU UU$^{(3'5')}$

4.5

81

B.

HO'5

HO'5

EtCO–P=O EtCO–P=O

CPG

i. couple with
5'-DMT 2',3'-bis rU amidite
(**1.2** : 0.03 M, 30 min)

ii. detritylate

iii.chain elongate ∿∿∿ : oligouridylate

DMTO

EtCO–P–N(*i*-Pr)₂
   N(*i*-Pr)₂   EtCO

1.2

5'OH

EtCO–P=O   EtCO–P=O

CPG

cleavage and deprotection

HO-5'-UUU UUU —(3'5')— rU

(3'5')— UU UUU —(3'5')— rU —(3'5')— UU UUU-3'-OH
                              (2'5')— UU UUU-3'-OH

(2'5')— UU UUU —(3'5')— rU —(3'5')— UU UUU-3'-OH
                              (2'5')— UU UUU-3'-OH

4.6

Figure 4.4: The solid phase synthesis of 39-nt poly-rU h-bNAs by **A.** *divergent-growth synthesis* using diMMT rU branchpoint amidite synthons **2.5** and **2.6** and **B.** the traditional *convergent-growth approach* with bisphosphoramidite branchpoint synthon **1.2**. The methods produce h-bNAs with 3',3'/2',5' (**4.4**), 2',3'/3',5' (**4.5**) and native 2',5'/3',5' (**4.6**) branchpoint linkages.

82

Compound **4.6**, with the 'natural' (2',5'/3',5') branchpoint connectivity was next prepared using the previously reported *convergent-growth approach* with the optimized conditions described by Hudson and Damha (CPG, 50-80 µmol/g; 0.03 M bisamidite synthon in MeCN).[41] The crude 39-mer h-bNA product, **4.6**, was then compared to the *divergent-growth* products **4.4** and **4.5** of the same length (**Figure 4.4- B**). The crude h-bNA **4.6** was recovered in 15% yield (45 OD units). PAGE and AE HPLC analyses of the crude product revealed a purity of only 25% (**Figure 4.5**). The structure of the h-bNA, **4.6**, was also confirmed by MALDI TOF MS (Calcd. 12290, found: 12292).

linear controls		bNAs Convergent Divergent optimized			h-bNAs Convergent Divergent optimized non-optimized optimized				
rU$_9$	rU$_{18}$	**4.3**	**4.2**	**4.1**	**4.6**	**4.5**	**4.4**	**4.5**	**4.4**
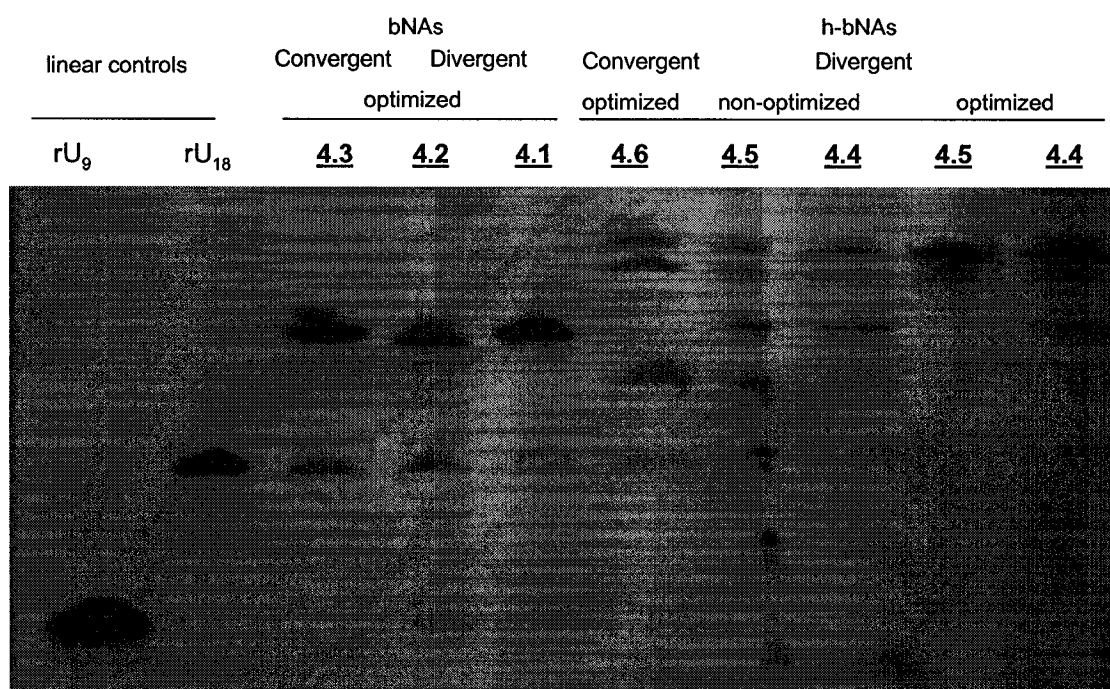
Figure 4.5: 24% denaturing PAGE analysis of the crude bNAs and h-bNAs synthesized via the *convergent-* and *divergent-growth* syntheses. The crude bNAs and h-bNAs samples were ran against linear oligouridylate controls (rU$_{18}$ and rU$_9$). The gel indicates higher yield of second generation h-bNAs with the *divergent* method.

A comparison of the synthesis, analysis and purification data illustrates that the *divergent-growth* approach, when optimized, affords higher yields of h-bNAs relative to the *convergent-growth* method (**Table 4.3.2**).

Table 4.3.2: Synthesis parameters and yields of bNAs and h-bNAs.

Sequence	scale (μmol)	coupling yield[1] —rU<	crude yield[2] ODs (%)	purity[3] (%)	purification[4] (avg. % recovery)
4.3	1	0.65	50 (21)	75	60
4.2	1	1.50	126 (48)	65	68
4.1	1	1.46	124 (47)	70	65
4.6	1	0.72	45 (15)	25	15
4.4non-opt	1	1.10	65 (20)	22	10
4.5non-opt	1	1.00	100 (35)	15	12
4.4opt	0.25	1.55	32 (33)	45	35
4.5opt	0.25	1.50	31 (33)	60	50

[1] colorimetric ratio of the absorbance values of the trityl colors in CH_2Cl_2 for the 5'-DMT 2'-TBDMS rU amidites (DMT, ε: 76) before and after branchpoint coupling; [2] determined from the absorbance values at A_{260}; [3] determined by AE HPLC (1 M LiClO₄ in H_2O; 0-30% gradient; total run time of 70 min; [4] after HPLC purification and desalting through Sephadex Nap® 10 columns.

4.4 CHEMICAL AND ENZYMATIC DEBRANCHING PROPERTIES OF bNAs.

Previous studies have indicated that isolated 2',3' or 3',3' phosphodiester bonds in oligonucleotides are important for increasing nuclease resistance of DNA/RNA-based therapeutics[157], altering strand polarity in branched, hairpin and triplex structures[158] and in assisting in mechanistic and kinetic studies on the hydrolytic stability of RNA-X.[147]

4.4.1 Stability of bNAs with 'natural' (2',5' and 3',5') and 'un-natural' (2',5'and 3',3' or 2',3' and 3',5') branchpoint linkages under basic and acidic conditions.

The secondary 3',3' and 2',3'-linkages are inherently less stable than the naturally occurring primary (2',5' and 3',5') phosphodiester bonds found in RNA due to the steric environment of the neighboring functional groups and the good leaving group feasibility of the 2' or 3'-phosphate.[159] Indeed, when ^{32}P-5'-end labeled bNAs **4.1** and **4.2** were treated with a pH range of 1.5 −10.7 (10 min., 95°C) both molecules underwent cleavage to produce linear ^{32}P-5'-RNAs, 18-nt in length. That is, the 3',3'-linkage of **4.1** and the 2',3' linkage of **4.2** were selectively hydrolyzed to afford ^{32}p-rU₉-2'-p-5'-rU₉, and ^{32}p-rU₁₈ respectively, along with rU₈ **(Figure 4.6)**. In sharp constrast, bNA, **4.3**, was more tolerant

to the neutral, acid and alkaline hydrolysis conditions (10 min, 95°C), indicating only a modest "RNA ladder" degradation profile (N-1, N-2, N-3 *etc.*) with no selective debranching occurring at the vicinal 2',5'/3',5' linkages (**Figure 4.6**).
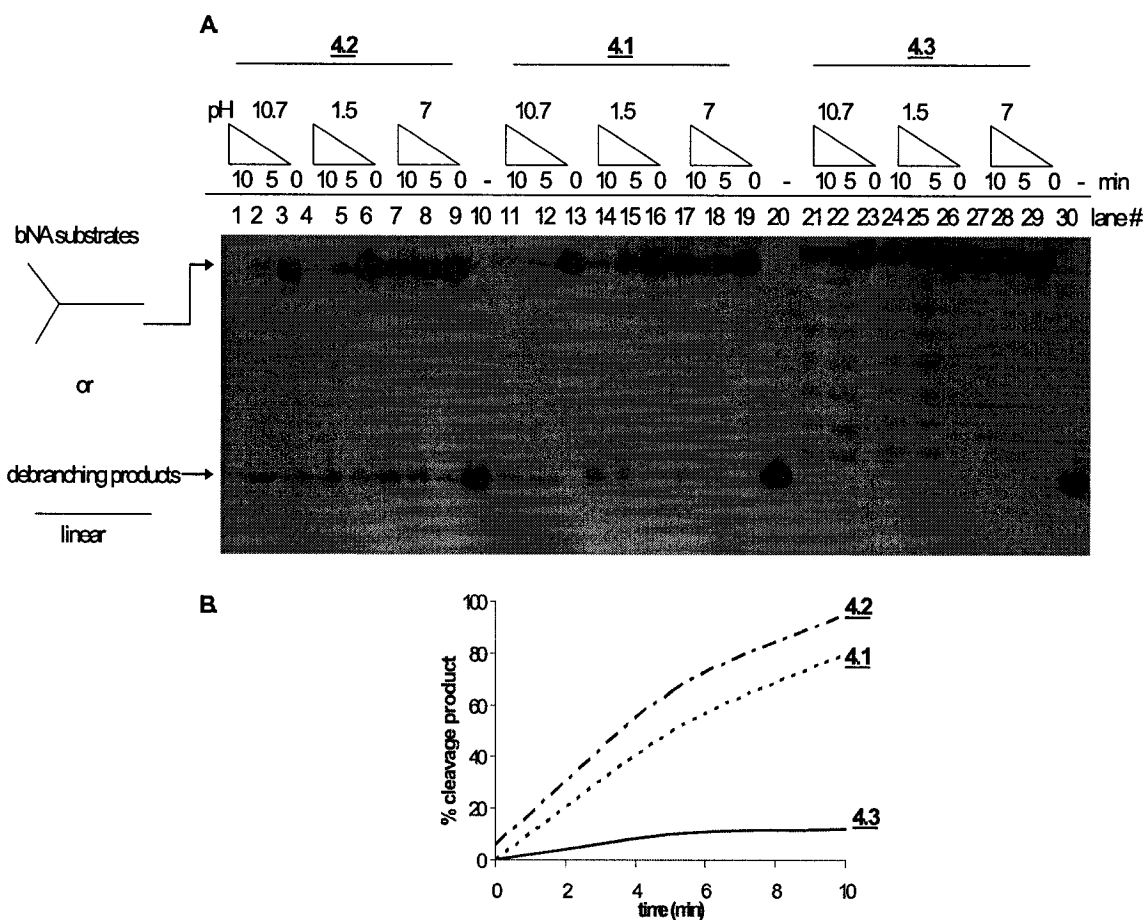
Figure 4.6: **A.** 16% PAGE analysis of the bNAs and linear controls. The bNA substrates and linear controls (200 pmol) were 5'- radiolabeled with γ-^{32}P ATP and PNK. Each bNA (*ca.* 0.5 pmol) was incubated in 10 µL of alkaline buffer (100 mM Na_2CO_3 pH: 10.7) or 10 µL of acidic buffer (100 mM HCl pH: 1.2) or 10 µL neutral buffer (50 mM Tris HCl pH: 7.4, 5 mM DTT, 1 mM EDTA, 10% glycerol, 500 mM NaCl) at 95°C for 10 min. Lanes 10, 30 contain the labeled linear ^{32}P-rU$_{18}$ sequence and lane 20 the ^{32}P-rU$_9$-2'-p-5'-U$_9$, both loaded with 50% formamide in H_2O. **B.** Chart demonstrating the extent of debranching of **4.1** (3',3'), **4.2** (2',3') and **4.3** (3',5' or 2',5') over time under alkaline conditions (pH: 10.7).

4.4.2 The Enzymatic Stability of bNAs with 'natural' (2',5' and 3',5') and 'unnatural' (2',5' and 3',3' or 2',3' and 3',5') Branchpoint Linkages

A debranching assay with the specific 2'-phosphodiesterase activity of the yeast debranching enzyme (yDB1) was performed next. yDBr1 converts RNA lariats and branched oligonucleotides into linear molecules by exclusive hydrolysis of the 2',5'-phosphodiester linkage, resulting in an oligonucleotide containing a 2'-hydroxyl group and a 5'-terminal phosphate.[51,160] Unlike other phosphodiesterases (e.g. snake venom phosphodiesterase) which cleave both 2',5' and 3',5' phosphodiester bonds, the RNA debranching enzyme (yDBr1) hydrolyzes only 2',5'-branchpoint linkages that are vicinal to 3',5'-bonds, thereby leaving all other phosphodiester bonds intact. Since purified yDBr1 was made available to us by Dr. Beate Schwer (Cornell University), we used this enzyme to further our understanding of the enzyme specificity towards bNAs containing unnatural linkages at the branchpoint (e.g., **4.1** and **4.2**).[161] Of much interest to us was to determine whether yDBr1 was capable of hydrolyzing the 2',5' phosphodiester linkage of bNA **4.1**, in which the 2',5'-linkage is adjacent to an atypical 3',3'-linkage The three radiolabeled bNAs were incubated with yDBr1 and the digestion products resolved by denaturing PAGE (**Figure 4.7**). Hydrolysis of ^{32}P-labeled bNA **4.3** resulted in the quantitative formation of rU_{18} within 40 min at 37°C (**Figure 4.7**). By contrast, the bNA lacking a 2',5'-phosphodiester linkage altogether (i.e., **4.2**) appeared to be completely resistant to the 2'-phosphodiesterase activity, however, some hydrolysis of the labile 2',3'-phosphodiester bond **4.2** occurred spontaneously in the absence of yDBr1 to produce the linear rU_{18} sequence (**Figure 4.7**). Hydrolysis of **4.1** resulted in the formation of a 17-nt cleavage product, $5'-rU_9-3'-p-3'-U_8-5'$, whose migratory behavior coincided with the control rU_{17} linear sequence (compare lane 2 with 11-14). Although debranching did not proceed quantitatively (30% debranching of **4.1** relative to 90% debranching of **4.3**), these results indicate, for the first time, that yDBr1 can cleave 2',5' linkages adjacent to unnatural 3',3' linkages. This debranching assay fully corroborates the assigned branchpoint structures for the bNAs **4.1-4.3**.
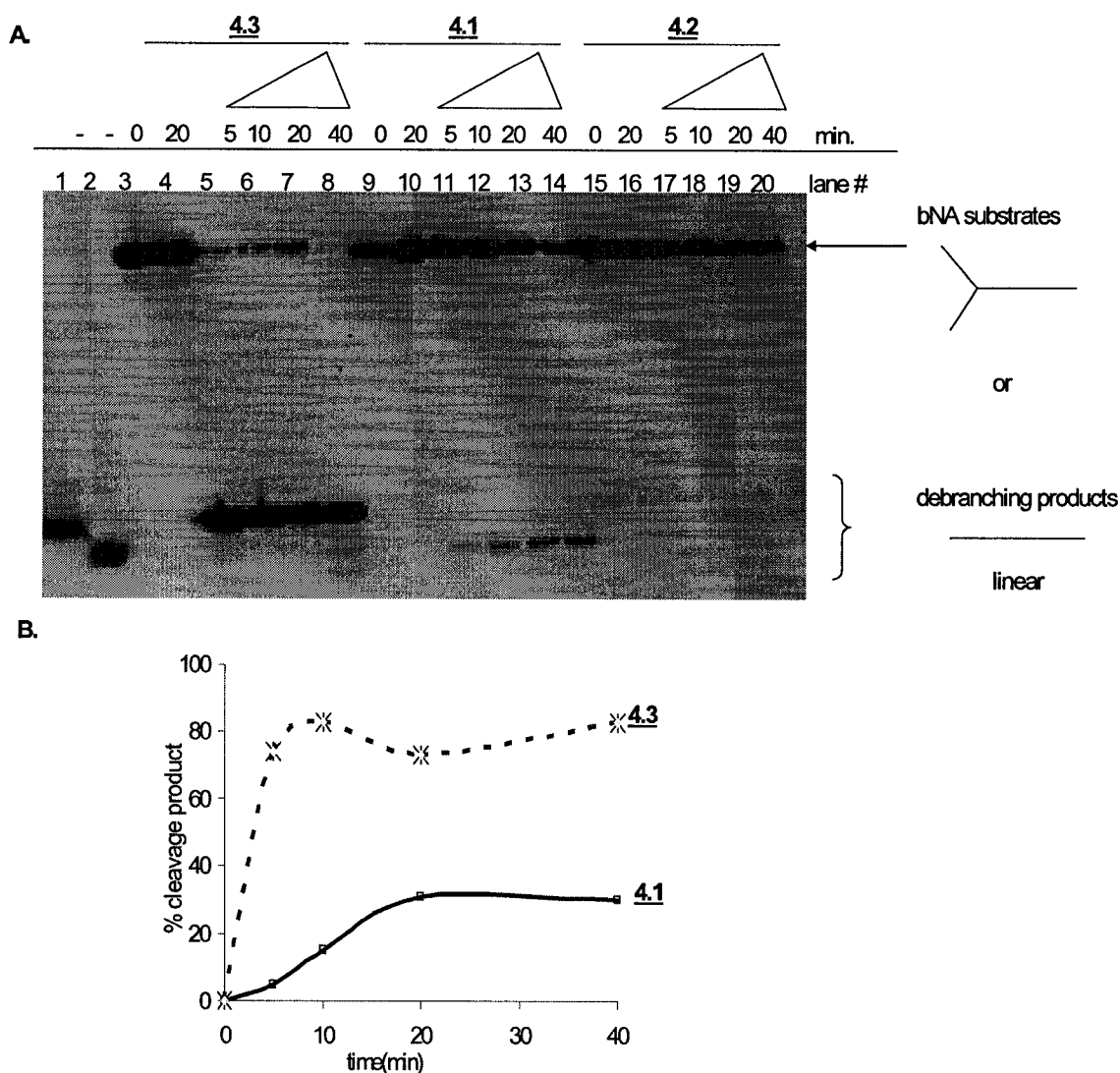
Figure 4.7: **A.** 16% PAGE analysis of bNAs. Approximately 1-2 pmol of the 5'-end ^{32}P radiolabeled bNAs were incubated with 10 μL of a solution containing yDBr1 (0.28 mg/mL, 1 μL), buffer (1 μL of 500 mM Tris-HCl pH: 7, 20 mM DTT, 5 mM MnCl$_2$) and water (8 μL) at 37°C for 40 min. They were also incubated with buffer alone at 37°C, for 20 min (lanes 3, 4, 9, 10, 15, 16). The control linear oligomers rU$_{18}$ (lane 1) and rU$_{17}$ (lane 2) were both loaded with 50% formamide in H$_2$O. **B.** Chart demonstrating the greater extent of debranching of **4.3** over **4.1** by yDBr1 over a 40-min period.

4.5 SUMMARY AND CONCLUSIONS FROM THIS STUDY

In summary, a novel *divergent-growth* method for the synthesis of high purity bNAs and h-bNAs oligouridylate sequences has been developed as a complementary method to

87

the *convergent-growth* synthesis of symmetric branched and hyperbranched oligonucleotides. The bNAs **4.1**, and **4.2** underwent selective debranching under acid, neutral or basic conditions at the 3',3' and 2',3' phosphodiester linkages, respectively (95 °C). The release of linear RNA products from Y-shaped RNA at physiological conditions (*e.g.* bNA **4.2**), and the fact that nucleic acid-based dendrimers have enhanced cellular uptake[151], make bNAs such as **4.2** (vicinal 2',3' and 3',5' phosphodiester linkages) potential "pro-drug" candidates for therapeutic applications (*i.e.* encapsulated bNA → slow release of 2 RNA strands → siRNA duplex).[124,162] Interestingly, bNA (**4.3**) was resistant to hydrolysis under similar conditions, underscoring nature's choice for the more stable 2',5'/3',5' branchpoint configuration. Furthermore, **4.1**, relative to the sequence with 'naturally' occurring branchpoint, **4.3**, exhibited a 3-fold decrease in the enzymatic 2',5'-*phosphodiesterase* debranching activity with yDBr1, making them potential yDBr1 inhibitors for their co-crystallization with the enzyme and mechanistic investigations related to the splicing mechanism of mRNA[160,161] This generates a specific requirement for developing an efficient synthesis strategy for the creation of asymmetric bNAs and h-bNAs with sequence compositions relevant to biological structures (*i.e.* msDNA).

4.6 REGIOSPECIFIC SYNTHESIS AND PROPERTIES OF BRANCHED AND HYPERBRANCHED NUCLEIC ACIDS OF MIXED BASE COMPOSITION

The strategy developed above has permitted the facile synthesis of branched oligo-rU sequences containing atypical branchpoint linkages; however, the method is not regioselective. Since following coupling of the branchpoint nucleotide, the acid step simultaneously unmasked both 5' and 2' (or 3')-hydroxyl groups at the apex of the molecule and subsequent chain growth proceeded with the simultaneous assembly of two identical 'arms', which limits the synthesis of naturally occurring msDNAs and their analogues (**Figure 4.8**). Given this limitation, we directed our attention toward alternative strategies for the regiospecific synthesis of bNAs and h-bNAs containing natural (2',5'/3',5') and the unnatural branchpoint linkages described above (3',3'/2',5' and 3',5'/2',3'). As a starting point, we chose to regiospecifically synthesize msDNA analogues consisting of two RNA arms and one DNA arm joined to a uridine branchpoint (**Table 4.9.1**). Our choice for this branchpoint nucleotide was based on a personal

communication with Dr. S. Virgin (Washington University), informing us that the norovirus specific branched oligonucleotides found in the RAW 264.7 macrophage cell line contained uracil at their branchpoint.[163,164]
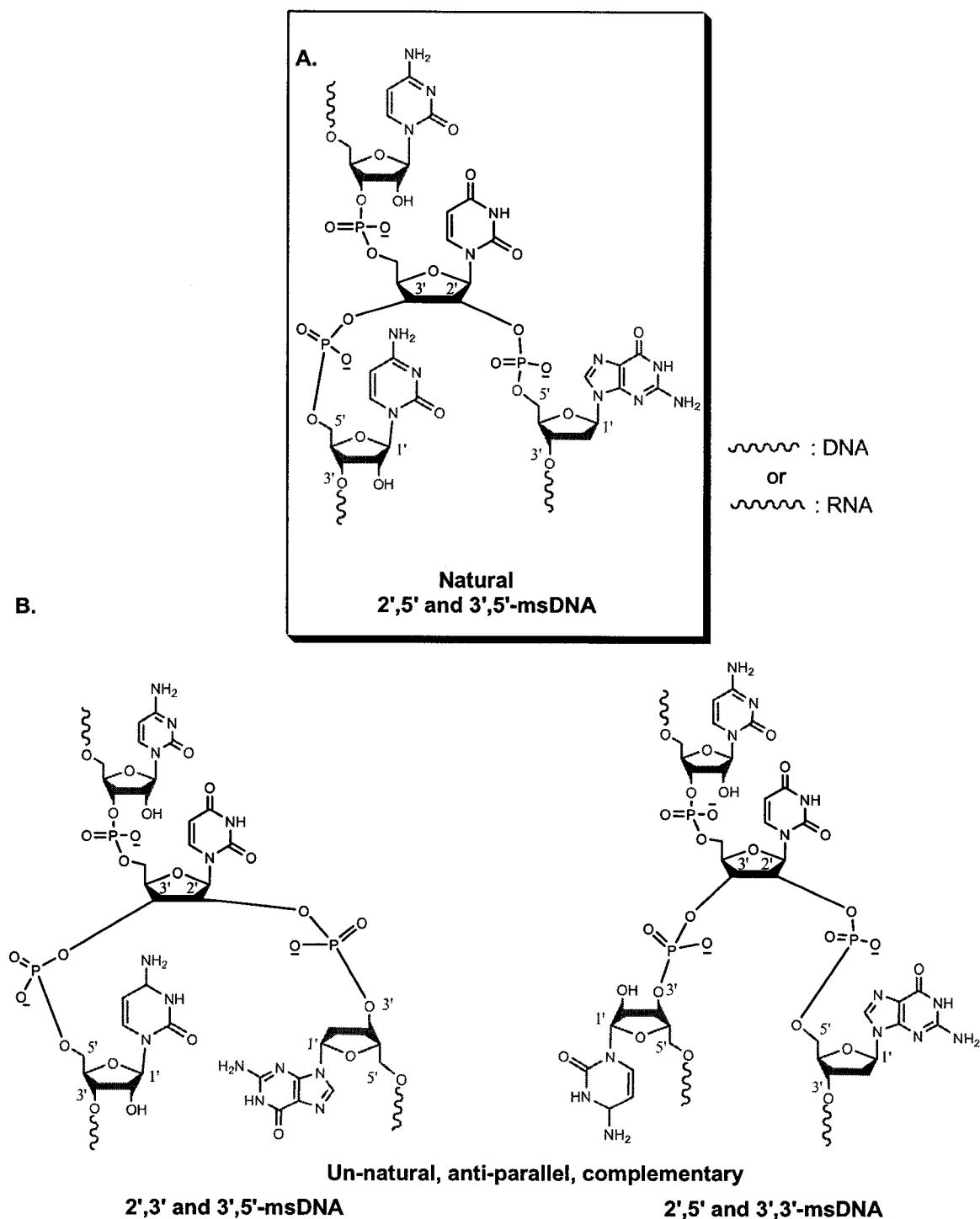
Figure 4.8: Structure of the branchpoint linkages in the branched msDNAs

4.7 PROJECT OBJECTIVES AND METHODOLOGY

The regiospecific synthesis of one msDNA oligonucleotide had previously been reported by our group, however, it required the use of 2'-Fpmp (acetal) protected RNA phosphoramidites which limited the ease and speed of deprotection of the sequences.[40c] Furthermore, the required removal of an orthogonal 2'-TBDMS group at the branchpoint sometimes lead to the release of oligonucleotide chains from the solid support, resulting from the fluoride-mediated 'etching' of the glass beads.[40b] While this could be largely prevented by a brief treatment with fluoride (<15 min), the reaction depended on the batch and quality of the TBAF reagent. Sometimes, desilylation did not proceed quantitatively reducing the yield of bNA synthesis.

Here we describe a facile method for the regiospecific synthesis of msDNA and hyperbranched msDNA. The strategy introduces novel branchpoint synthons containing the orthogonal levulinyl (Lv) and MMT protection (*i.e.*, **4.10** and **4.11**, and the previously described regioisomeric pairs **2.10** and **2.11**). We (Chapter 2) and others have demonstrated that Lv groups can be removed rapidly and efficiently under nearly neutral conditions without releasing nascent RNA or DNA chains from the solid support.[124,168]

The synthesis of branched oligonucleotides of mixed base composition in which deoxyribose is substituted for ribose sugars in the normal (2',5' or 3',5') or atypical (3',3' and 2',3') configuration would be useful for studying the substrate specificity of debranching enzymes, for evaluating the requirements of the branchpoint phosphodiester linkages in yDBr1-bNA interactions, and developing inhibitors of both yDBr1 (for potential co-crystallization studies) and noroviruses. The method is also amenable to the synthesis of msDNAs with complementary arms **(Figure 4.8)** thus capable, in principle, of base pairing interactions. These bNAs will serve as models for hybridization (*i.e.* T_m), structural (*i.e.* CD) and biological (*i.e.* *E. coli* RNaseH) investigations. The strategy for nucleoside synthesis, the automated assembly of msDNAs and their repeated synthesis, provided discrete 2nd generation dendrimer[165,170] and 3rd generation cascade dendrimer[156] or hyperbranched msDNA structures **(Figure 4.9)**.

R_x : DMT and R_y : P(OCEt)N(i-Pr)$_2$; **reverse DNA synthon**

or

R_x : P(OCEt)N(i-Pr)$_2$ and R_y : DMT; **DNA synthon**

RNA synthon

linear RNA   **branch msDNA (1st generation)**   **dendrimer msDNA (2nd generation)**   **cascade dendritic msDNA (3rd generation)**

Oligonucleotide Structures

R : Lv and R' : MMT; **branchpoint synthon (4.10, 4.11)**

or

R : MMT and R' : Lv; **branchpoint synthon (2.10, 2.11)**

Building Block Monomers
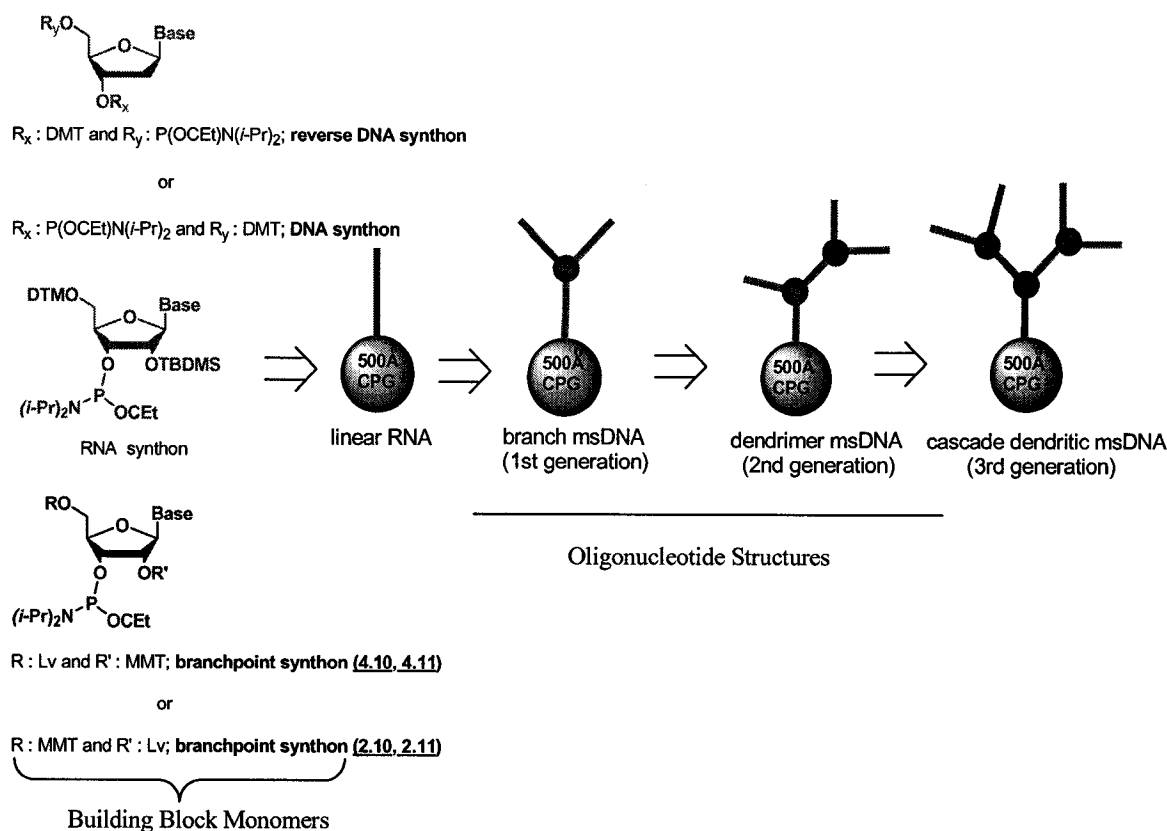
Figure 4.9: The solid phase synthesis assembly of linear, branched and hyperbranched (dendrimer and cascade dendritic) msDNA structures through the use of commercially available DNA, reverse DNA, RNA, and the branchpoint phosphoramidite synthons synthesized in this study, namely **2.10**, **2.11**, **4.10** and **4.11.**

4.8 SYNTHESIS OF 5'-Lv, 2' AND 3'-MMT URIDINE PHOSPHORAMIDITE BRANCHPOINT SYNTHONS

The synthesis of 5'-MMT, 2'(or 3')-Lv rU synthons **2.10** and **2.11** have been previously described along with their applications in RNA synthesis (Chapter 2). The chemoenzymatic synthesis of **4.10** and **4.11** was initiated by regioselective 5'-levulination of uridine (**2.1**) with the immobilized form of lipase B from *Candida Antarctica* (Novozyme® 435) in the presence of levulinic anhydride[166] as the acyl donor. The reaction was monitored to completion by TLC and yielded cleanly the desired product, **4.7,** after overnight reaction at room temperature. The slow reaction is partly due to the nature of the substrate, as the chemoenzymatic levulination reaction works best with selective nucleosides and ribonucleosides.[167,168] The enzyme was filtered off, washed

91

with MeOH and H_2O to remove organic material and re-used. In fact, we found that the enzyme can be recycled up to four times yielding quantitative conversion of **2.1** into **4.7** (recovered 80%).[168] Without purification, **4.7** reacted with monomethoxytrityl chloride (MMT-Cl) in the presence of silver chloride catalyst[110] to yield a mixture of 5'-Lv, 2'-MMT rU (**4.8**) and 5'-Lv, 3'-MMT rU (**4.9**) in *ca.* 2:1 ratio. After separation by silica gel column chromatography (R_f, 5% MeOH in $CHCl_3$: 0.25 and 0.15, with **4.8** eluting faster), the regioisomers were characterized by ^1H NMR and ^1H-^1H correlation COSY NMR and distinguished by identifying the locations of the secondary hydroxyl groups (**Figure 4.11**). The protected ribonucleosides **4.8** and **4.9** were converted to their corresponding phosphoramidite derivatives, **4.10** and **4.11**, in 60-65% yield ($C_{43}H_{51}N_4O_{10}PNa$ Calcd: 838, found in both cases: 837).
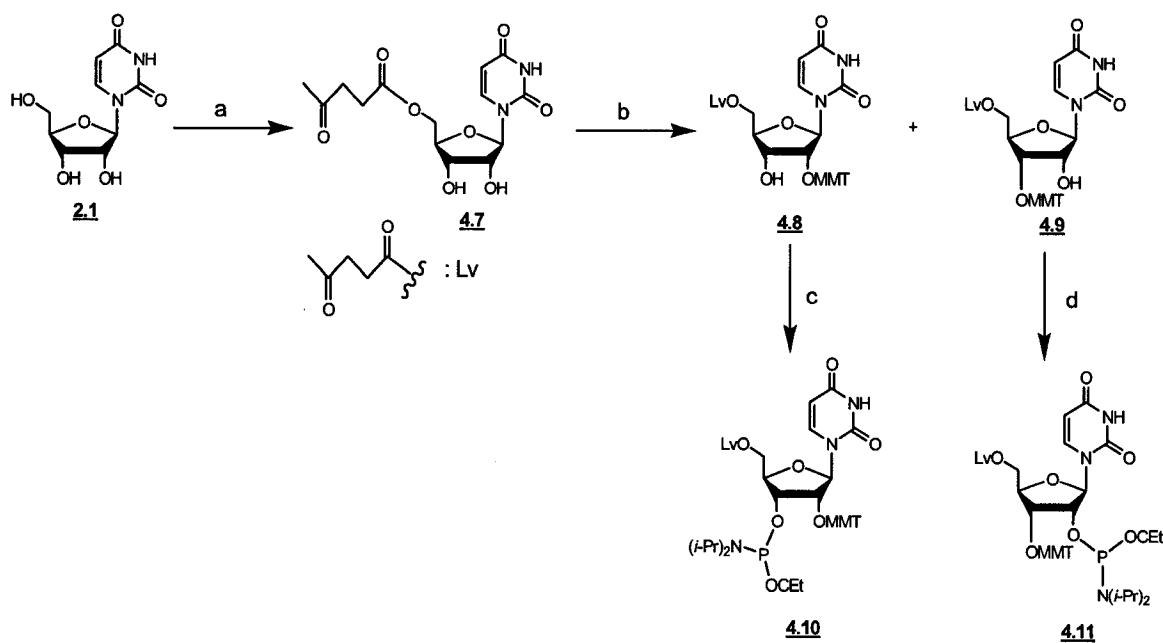
Figure 4.10: Reagents and Conditions: a. Lv_2O, Novozyme® 435, 1,4-dioxane, 22°C, overnight, 80%. b. MMT-Cl, $AgNO_3$, THF, pyr. 22°C, overnight, **4.8**: 40% and **4.9**: 20%, c. and d. Cl-P(OCEt)N(i-Pr)$_2$, Et-N(i-Pr)$_2$, THF, 22°C, 3 hours, **4.10**: 60% and **4.11**: 65%.

Figure 4.11: The ^{1}H-^{1}H COSY NMR for **A. 4.8** and **B. 4.9** in DMSO-d_6 at 400 MHz showing correlation of the OH protons to H3' and H2', respectively.
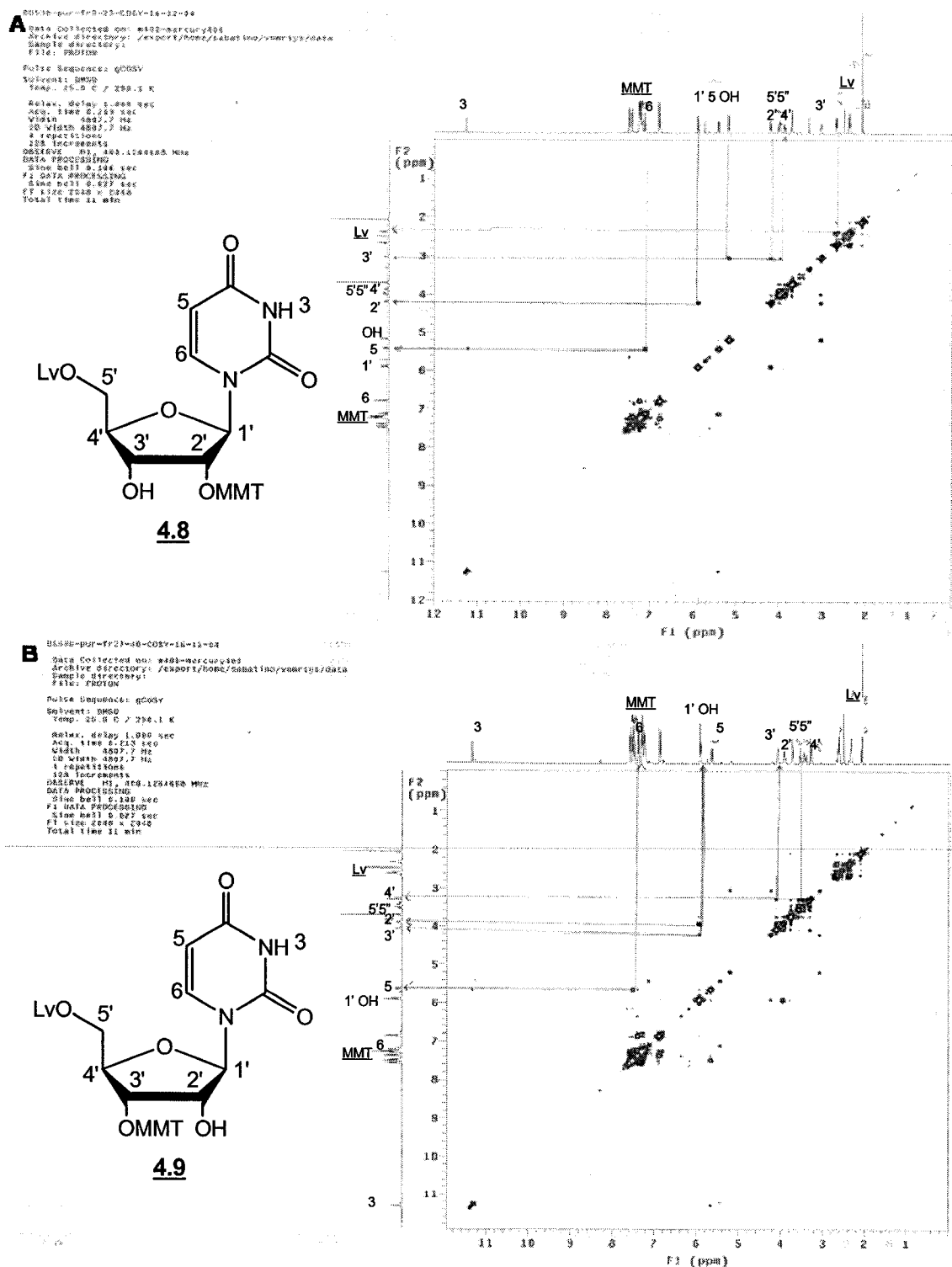
4.9 SOLID PHASE SYNTHESIS OF BRANCHED AND HYPERBRANCHED msDNA SEQUENCES

4.9.1 Solid Phase Synthesis and Characterization of Branched msDNA of Mixed Base Composition

With synthons **4.10**, **4.11** and **2.10**, **2.11** in hand, solid phase synthesis of msDNAs were initiated by synthesizing a linear RNA or DNA strand on CPG solid supports derivatized with either 5'-DMT dT[19] inverted 3'-DMT dT[157] or with the universal Unylinker® support[169] (**Figure 4.13**). Following coupling of the branchpoint synthons **4.10** and **4.11** the solid support was treated with triethylamine/acetonitrile solution (4:6 v/v, 90 min) to affect the removal of all cyanoethyl phosphate protecting groups, thus providing an oligomer with a intrinsically more stable phosphodiester backbone (Chapter 3). This is necessary because deblocking of a 2'-OH (or 3'-OH) group vicinal to a 3',5'- (or 2',5'-)phosphotriester linkage under neutral, acidic, or basic conditions has been shown to lead to phosphoryl migration and/or chain cleavage (Chapter 3).[22] The 2' (or 3')-MMT group was removed from the branching rU residue by treatment of the support 3% TCA/DCM (2.5 min), followed by washing with acetonitrile. At this point, the second 'arm' was synthesized in the 5'-to-3' direction using commercially available DNA 5'-phosphoramidites ("reverse amidites"), or in the normal 3'-to-5' direction using conventional RNA or DNA 3'-phosphoramidite derivatives. This was followed by a manual delevulination procedure by passing through a solution of 0.5 M NH_2NH_2-H_2O in pyr:HOAc (3:2) through the solid support (20 min).[120,124,168] The column containing the solid support was then washed with MeCN (5 min), reattached to the gene machine, and RNA synthesis under standard conditions[24] was then continued until the desired 18-nt msDNAs were assembled (**Table 4.9.1**).

The msDNA sequences were cleaved from the support and deprotected by treatment, with concentrated aqueous ammonia (3:1 NH_4OH:EtOH, 55°C, 12-16 h) followed by evaporation. The solid residue was then resuspended in TREAT-HF, to effect deblocking of the TBDMS groups (48 h, 22°C), and the resulting mixture precipitated in n-butanol. Crude samples were recovered (40-60%) in sterile water and purified by either AE HPLC or PAGE. The analyses indicated that 50-70% of the material were the desired target

msDNA sequences (**Figure 4.12** and **4.16**). The yield and purity was comparable to the *divergent-growth approach* previously reported by our group.[40c] The sequences were desalted by size exclusion chromatography (Sephadex® G-25) and their molecular weights confirmed by MALDI-TOF MS (**4.12a**: Calcd. 5808, found 5809; **4.13a**: Calcd. 5496, found 5497; **4.14**: Calcd. 10251, found 10250; and **4.15**: Calcd. 10250, found 10249).

The synthesis of msDNAs was also performed with branchpoint synthons **2.10** and **2.11** (**Figure 4.14**). Switching the locations of Lv and MMT groups on the ribose moiety (*e.g.* **4.10** *vs* **2.10**) led to a slightly different msDNA assembly strategy (**Figure 4.14**). For example, msDNA sequences **4.12b** and **4.13b** were assembled by first synthesizing a normal 3',5'-linked RNA (13-nt) oligomer before branching from the *internal* rU 2' (or 3') Lv residue. In the case of msDNA **4.12a** and **4.13a**, which required the 5'Lv, 2'/3'-MMT monomers (**4.10** and **4.11**), branching occurs earlier, *i.e.*, from the *terminus* of the chain, immediately after the branchpoint synthons are incorporated (compare **Figures 4.13** and **4.14**). As a result, we found that the internal 2' (3')-Lv groups required longer exposure to hydrazine to effect their removal, compared to the terminal 5'-Lv groups (*e.g.*, 30 *vs* 20 min). Also, based on the PAGE profiles shown in **Figure 4.12**, it is evident that branching from an internal hydroxyl position (msDNA **4.12b** and **4.13b**) proceeds less efficiently relative to when branching occurs earlier (msDNA **4.12a** and **4.13a**) from the less crowded terminal position (refer to "ladder" of bands immediately below the products, **Figure 4.12**). Nevertheless, overall yields (*ca.* 60%) and purity of the recovered msDNA sequences **4.12b** (65%) and **4.13b** (70%) were quite good. MS characterization [MALDI-TOF MS; **4.12b**: Calcd. 5808, found 5810; **4.13b**: Calcd. 5496, found 5497] and the fact that a sample of **4.12a** (produced from synthon **4.10**) was the same as a sample of **4.12b** (produced from synthon **2.10**), unequivocally establishes the assigned structures.

Sequence	Branched msDNA	Branchpoint Synthons
1. **4.12a**	2'3'-r(CGG GU)-5' 5'-r(CCG AUC)-3',5'-rU 3'5'-d(GTG AAT)-3'	**4.10**
2. **4.13a**	2'5'-r(CGG U)dT-3' 5'-r(CCG AUC)-3'5'-rU 3'3'-d(GTG AAT)-5'	**4.11**
3. **4.14**	2'5'-d(GTG AAA TGA G)-3' 5'-r(CUA CCG AUC)-3'5'-rU 3'5'-r(CGG GUG GAC G)-3'-p-3'-dT-5'	**4.10**
4. **4.15**	2'5'-r(CGG GUG GAC G)-3'-p-3'-dT-5' **4.11** 5'-r(CUA CCG AUC)-3'5'-rU 3'5'-d(GTG AAA TGA G)-3'	
5. **4.12b**	2'3'-r(CGG GU)-5' 5'-r(CCG AUC)-3'5'-rU 3'5'-d(GTG AAT)-3'	**2.10**
6. **4.13b**	2'5'-r(CGG U)dT-3' 5'-r(CCG AUC)-3'5'-rU 3'3'-d(GTG AAT)-5'	**2.11**
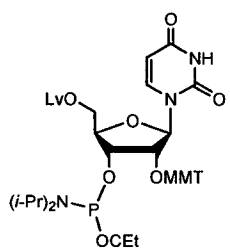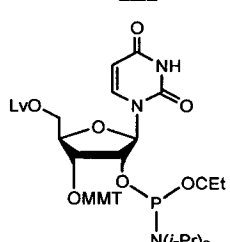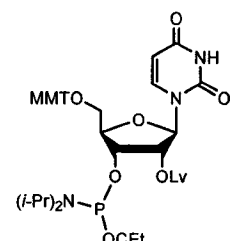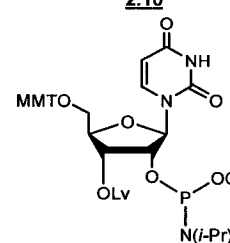
Table 4.9.1: The branched msDNA sequences and their corresponding branchpoint synthons synthesized in this study. The lowercase, d, denotes the DNA sequences and the lowercase, r, denotes the RNA sequences.

controls   msDNA   controls   msDNA

linear   4.12a   4.12b   linear   4.13a   4.13b
13-mer   18-mer   6-mer   12-mer   18-mer
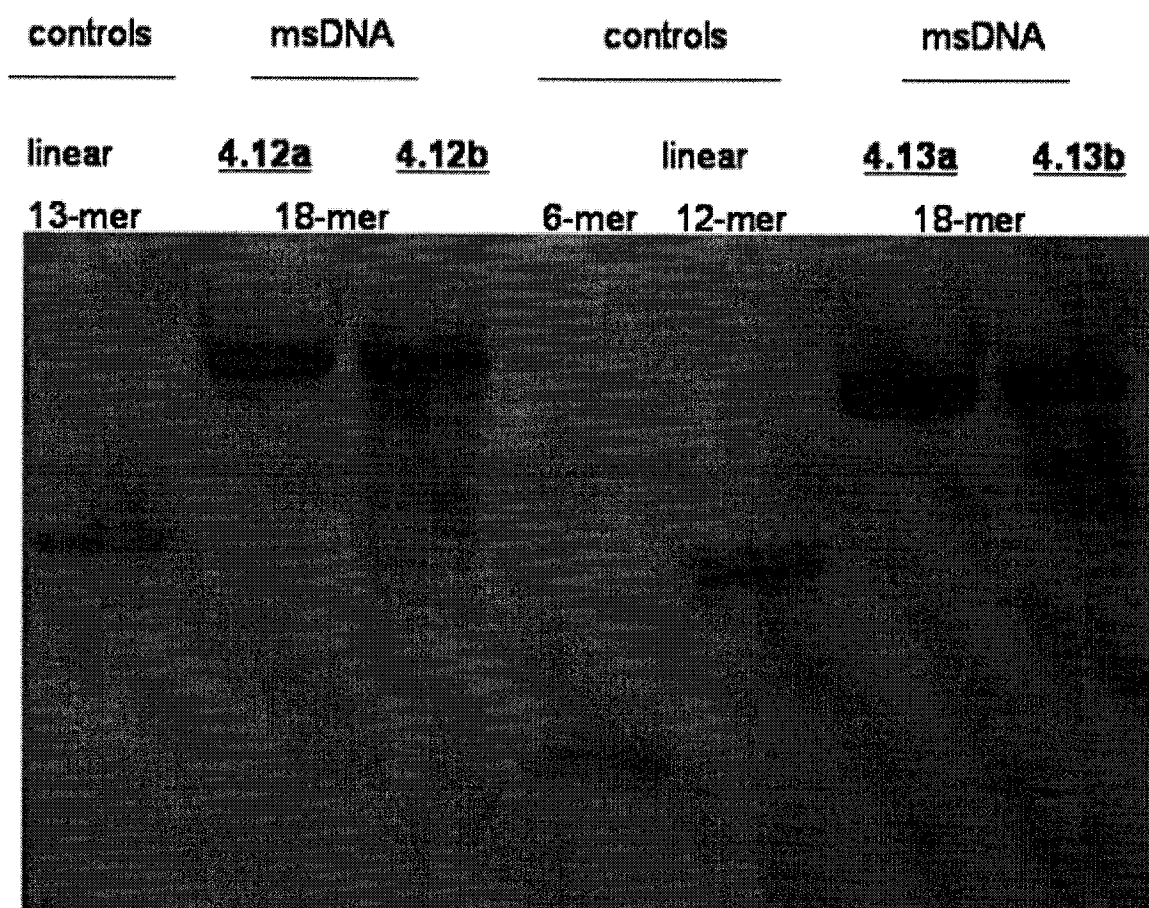
Figure 4.12: 24% denaturing PAGE comparison of the msDNA sequences synthesized with the exact base sequence compositions (*i.e.* 4.13a and 4.13b with 4.11 and 2.11 and 4.12a and 4.12b with 4.10 and 2.10). The sequences were compared to a linear 13-mer [5'-r(CCG AUC)-3'5'-rU-2'5'-3'5'-d(GTG AAT)-3'], 6-mer [5'-d(TAA GTG)-3'], and 12-mer [5'-r(CCG AUC)-3'5'-rU-2'5'-r(CGG U)dT-3'] controls. The lowercase, d, denotes the DNA sequences while the lowercase, r, denotes the RNA sequences.

97
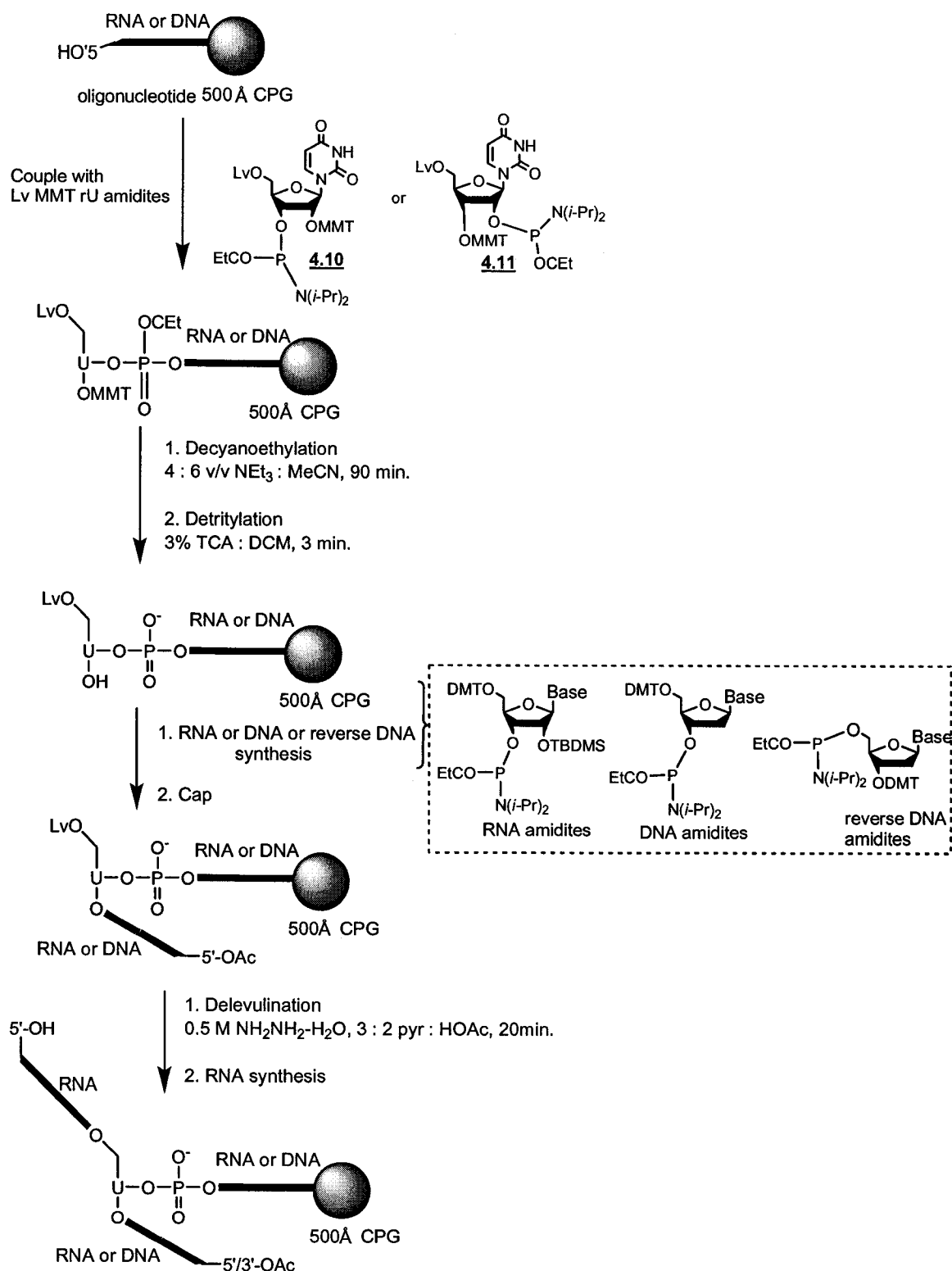
Figure 4.13: Regiospecific *divergent-growth* synthesis of branched msDNAs through the use of branchpoint synthons **4.10** and **4.11**.
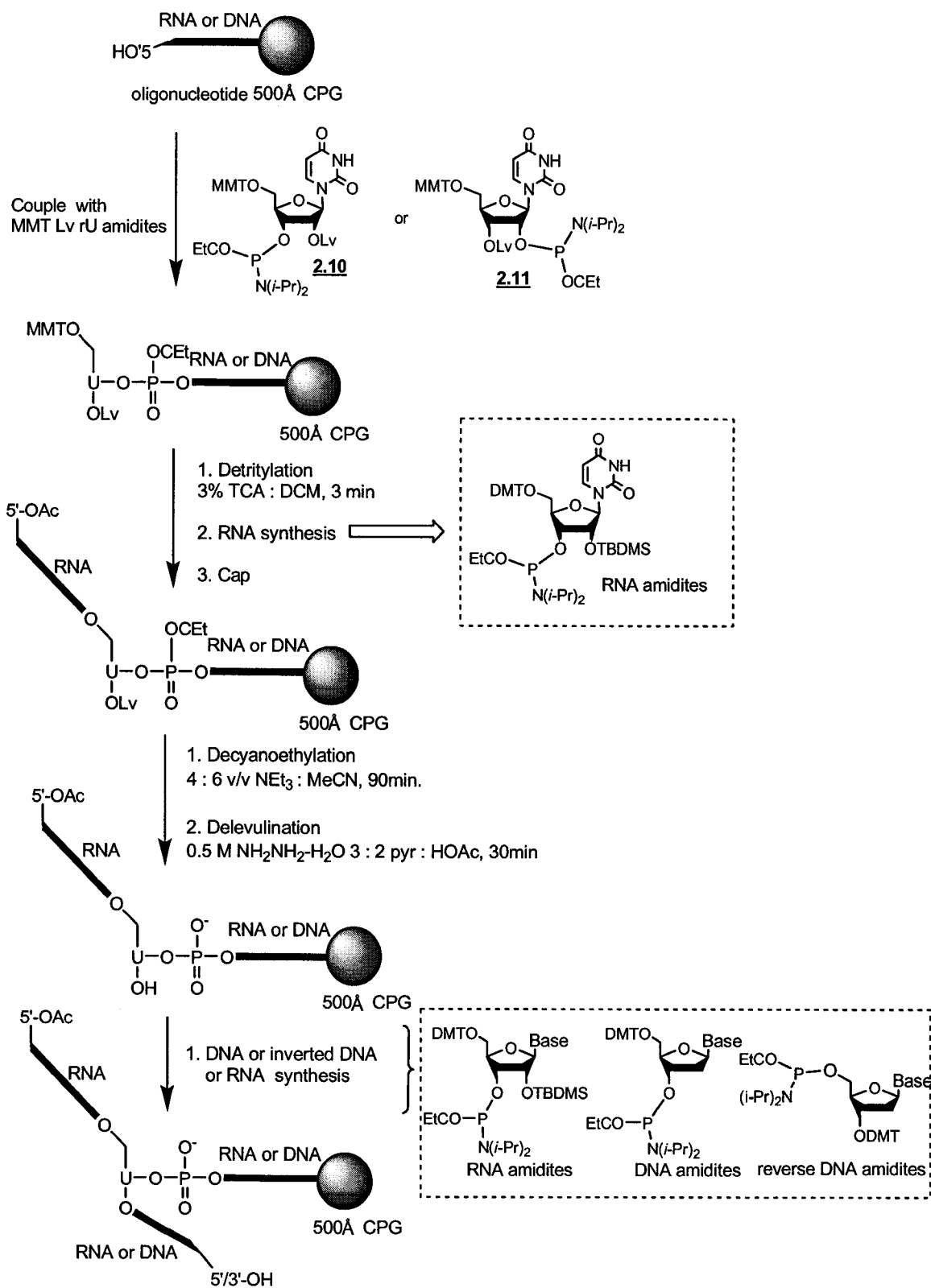
Figure 4.14: Regiospecific *divergent-growth* synthesis of branched msDNAs through the use of branchpoint synthons **2.10** and **2.11**.

4.9.2 Solid Phase Synthesis and Characterization of Hyperbranched msDNA Sequences of Mixed Base Compositions

The automated synthesis of discrete 2nd generation[165,170] and 3rd generation[156] msDNA dendrimers (**Table 4.9.2**: **4.16** and **4.17**) with base sequence composition found in the norovirus specific oligonucleotide RAW 264.7 macrophage cell line[163,164] were constructed next (**Figure 4.15**). Divergent-growth synthesis of these hyperbranched msDNAs was performed on a 0.5 μmol scale using a 3'-DMT dT derivatized CPG as solid support (500 Å pore size; succinyl linker; loading of 40-50 μmol dT/g of CPG).[157] These conditions were thought to minimize any steric effects of the nascent hyperbranched structure within the CPG pore, and provide a larger excess of monomer to push the reaction to completion. The branchpoint synthon **4.10** was added as a 0.15 M solution in MeCN, and coupled for extended 20 minute reaction times using 5-ethylthiotetrazole as activator (0.25 M ETT in MeCN). Furthermore, to ensure efficient chain growth, the first DNA 5'-phosphoramidite or RNA 3'-phosphoramidite addition at each branchpoint was carried out with a more concentrated solution (0.15 M for DNA and 0.3 M for RNA) and longer coupling time (5 min for DNA and 20-30 min for RNA) (**Figure 4.16**). Removal of the 2'-MMT group followed by DNA synthesis and acetylation at the 3' terminus was proceeded with the cleavage of the 5'-Lv group and RNA chain extension provided the first generation msDNA. At this point the whole process was repeated (branchpoint **4.10** coupling, detritylation, DNA synthesis, capping, removal of the Lv and RNA synthesis) to yield a 2nd generation, 31-nt msDNA dendrimer (**Figure 4.15- A**). The CPG bound dendrimer was deprotected under standard conditions (concentrated aqueous ammonia in ethanol, followed by removal of silyl groups as described above for msDNA synthesis; section **4.9.1**) to yield **4.16** as crude product (50% yield; 35% purity as assessed by HPLC and PAGE). This similar iterative divergent growth procedure was followed for the synthesis of a 46-nt, 3rd generation, msDNA dendrimer (**4.17**) (**Figure 4.15- B**). The product was recovered in crude yields of 40% and analyzed by AE HPLC and PAGE which indicated a purity of 37% (**Figure 4.16**).

1. 4.16

           2'5'-d(GTG AAT)-3'

5'-r(CCG AUC)-3'5'-rU   2'5'-d(GTG AA)-3'

       3'5'-r(CCG AUC)-3'5'-rU

           3'5'-r(CGG GU)-3'-p-3'-dT-5'

2. 4.17

          2'5'-d(GTG AAT)-3'

5'-r(CCG AUC)-3'5'-rU   2'5'-d(GTG AAT)-3'

      3'5'-r(CCG AUC)-3'5'-rU   2'5'-d(GTG AAT)-3'

         3'5'-r(CUA GCC)-3'5'-rU

           3'5'-r(CUA GCC)-3'-p-3'-dT-5'
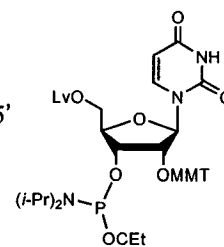
4.10

Table 4.9.2: The hyperbranched msDNA sequences and their corresponding branchpoint synthons synthesized in this study. The lowercase, d, denotes the DNA sequences and the lowercase, r, denotes the RNA sequences.

A.

B.
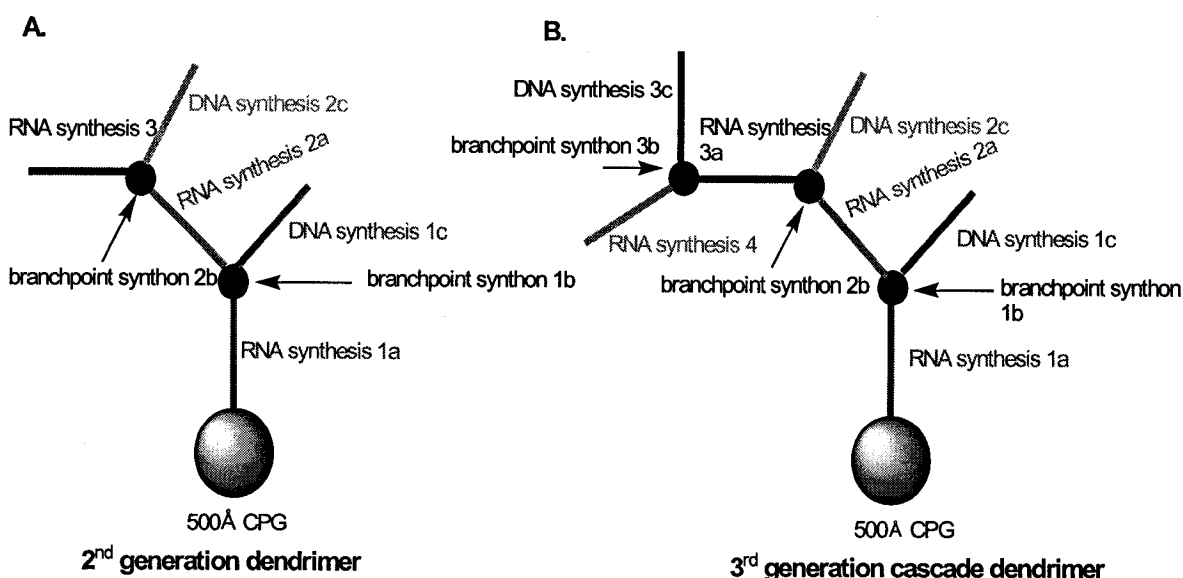
Figure 4.15: The regiospecific *divergent-growth* synthesis of hyperbranched msDNA A. **4.16** and B. **4.17.**

The recoveries and purity were found to be comparable to the previously established methods for the synthesis of hyperbranched DNA and RNA.[41,42] Moreover, these novel methods generated alternative routes for the synthesis of asymmetric branched and

hyperbranched sequences having biologically relevant compositions. The hyperbranched sequences were desalted by size exclusion chromatography (Sephadex® G-25) and their structures confirmed by MALDI-TOF MS (<u>4.16</u>: Calcd. 13498, found 13482 and <u>4.17</u>: Calcd. 16614, found 16693).

Divergent								Convergent		control
h-bNAs			bNAs					h-bNA	bNA	linear
<u>4.17</u>	<u>4.16</u>	<u>4.15</u>	<u>4.14</u>	<u>4.13a</u>	<u>4.13b</u>	<u>4.12a</u>	<u>4.12b</u>	<u>4.6</u>	<u>4.3</u>	rU$_{18}$
46-mer	31-mer	31-mer			18-mer			39-mer	26-mer	18-mer
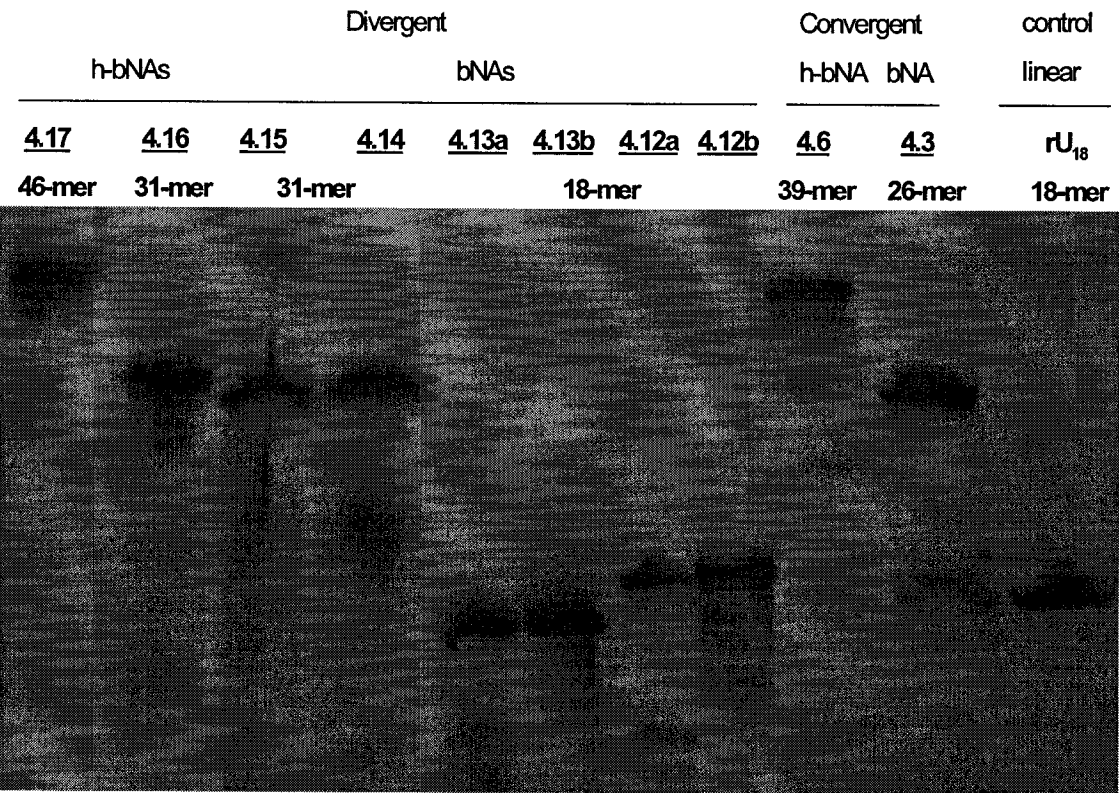
Figure 4.16: 24% denaturing PAGE comparison of the msDNA branched and hyperbranched sequences (bNAs and h-bNAs) synthesized in this study. The sequences were compared to linear 18-mer (rU$_{18}$), branched 26-mer (**4.3**), and hyperbranched 39-mer (**4.6**) control sequences generated by the convergent approach as previously described (section **4.3.1** and **4.3.2**).

4.10 SYNTHESIS, HYBRIDIZATION AND STRUCTURAL PROPERTIES OF msDNA WITH SELF BASE PAIRING ABILITY

4.10.1 Synthesis of Complementary msDNA Sequences

Branched oligonucleotides with two connected binding domains are choice candidates for hybridization and complexation studies because they possess the propensity for rigid binding at a low entropic energy cost.[27,28,40,171] Self-complementary 43-nt msDNAs sequences (**Table 4.10.1**) were synthesized on a Unylinker[169] CPG support by the newly developed *divergent-growth* synthesis protocol using **4.10** (sequence **4.18**) or **4.11** (sequence **4.21**) as branchpoint synthons (refer to section **4.9.1**, **Figure 4.13**). Their strand orientation favored the formation of an antiparallel duplex region as shown in **Figure 4.8**. Thus, the strand polarities and base sequences of the two longer arms of msDNAs **4.18** and **4.21** favor the formation of DNA/RNA hybrids (**Table 4.10.1**). Likewise, **4.20** would self-associate to produce an RNA/RNA duplex. In order to evaluate the steric contribution of the branchpoint position on duplex structure and hybridization, additional complementary 53-nt msDNA sequences were synthesized containing non-complementary 5 nt residues from the branchpoint position (sequences **4.19** and **4.22**). Linear control heteroduplexes (hybrids) were also synthesized by the conventional methods.[24] All of these sequences were isolated by standard methods[155] then purified and desalted by denaturing PAGE and by Sephadex® G-25 prior to characterization by MALDI-TOF MS. The thermal stability and global helical structure of the resulting duplexes were determined spectroscopically (UV T_m melting experiments and circular dichroism). Finally, the propensity of these duplexes to act as substrate for RNase H, an enzymatic activity that hydrolyzes the RNA strand of DNA/RNA hybrids, was also evaluated. As described below, some interesting differences in RNase H mediated cleavage were observed among the various msDNA structures.
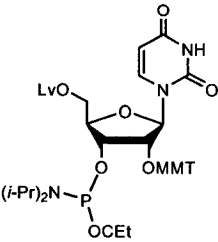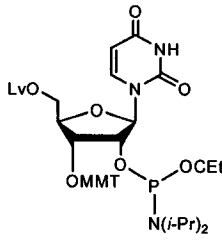
Sequence	Branched Complementary	Branchpoint Synthon

1. 4.18

2'3'-d(GCC TAC CTG CGA ACT GGA)-5'
5'-r(CCG AUC)-3'5'-rU
3'5'-r(CGG AUG GAC GCU UGA CCU)-3'

LvO—[ribose ring]—N(uracil base, O, NH, O)
O—OMMT
(*i*-Pr)₂N—P
OCEt

4.10

2. 4.19

3'5'-r(CCG AU CGG AUG GAC GCU UGA CCU)-3'   **4.10**
5'-r(CCG AUC)-3'5'-rU
2'3'-d(CCG AT GCC TAC CTG CGA ACT GGA)-5'

3. 4.20

3'5'-r(CGG AUG GAC GCU UGA CCU)-3'   **4.10**
5'-r(CCG AUC)-3'5'-rU
2'3'-r(GCC UAC CUG CGA ACU GGA)-5'

4. 4.21

3'3'-r(CGG AUG GAC GCU UGA CCU)-5'
5'-r(CCG AUC)-3'5'-rU
2'5'-d(GCC TAC CTG CGA ACT GGA)-3'

LvO—[ribose ring]—N(uracil base, O, NH, O)
OMMT—O—P—OCEt
N(*i*-Pr)₂

4.11

5. 4.22

3'3'-r(CCG AU CGG AUG GAC GCU UGA CCU)-5'   **4.11**
5'-(CCG AUC)r-3'5'-rU
2'5'-d(CCG AT GCC TAC CTG CGA ACT GGA)-3'

Table 4.10.1: Self-complementary msDNA and branched RNA sequences and the corresponding branchpoint synthons from which they are derived. The lowercase, d, denotes the DNA sequences and the lowercase, r, denotes the RNA sequences. The highlighted region indicates the complementary sequences.

4.10.2 Hybridization of Self-Complementary msDNA Sequences– T_m and CD Studies.

The msDNA samples (3.04 µM) were denatured at 95°C for 10 minutes, and then allowed to hybridize overnight at 4°C in a phosphate buffer that mimics physiological conditions (140 mM KCl, 5 mM NaH₂PO₄, 1 mM MgCl₂ pH: 7.2). Thermal stability data for the msDNA (melting temperature, T_m and % hyperchromicity, %H) were similar to

those of the linear DNA/RNA and RNA/RNA control duplexes (*e.g.* compare **4.18** to sequence 7, and **4.20** to sequence 6, **Table 4.10.2**). However, this does not imply that all the residues of msDNA duplex **4.18** are capable of base pairing because unimolecular duplexes are generally more stable than bimolecular duplexes of the same base composition.[172] In fact, comparison of **4.18** (T_m: 64 °C) and **4.19** (T_m: 67 °C) suggest that the branchpoint sterically constrains base pairing[40], although based on the ΔT_m value observed (-3 °C), the effect appears to be small. The same trend applied to the RNA duplexes **4.20** *vs* sequence 8 (ΔT_m = -4 °C) (**Table 4.10.2**). Comparison of msDNA **4.18** (T_m: 64 °C) and **4.21** (T_m: 58 °C) suggests that the 3',5'/2',3' branched framework leads to more stable duplexes relative to the 3',3'/2',5' branchpoint configuration.

Sequence[#]	T_m (°C)	% H
1. **4.18**	64	18.6
2. **4.19**	67	18.5
3. **4.20**	79	18.0
4. **4.21**	58	12.7
5. **4.22**	64	16.4
6. 5'-r(CGG AUG GAC GCU UGA CCU)-3' 3'-r(GCC UAC CUG CGA ACU GGA)-5'	80	21.0
7. 5'-r(CGG AUG GAC GCU UGA CCU)-3' 3'-d(GCC TAC CTG CGA ACT GGA)-5'	65	18.8
8. 5'-r(CCG AU CGG AUG GAC GCU UGA CCU)-3' 3'-r(CCG AU GCC UAC CUG CGA ACU GGA)-5'	83	21.3
9. 5'-r(CCG AU CGG AUG GAC GCU UGA CCU)-3' 3'-d(CCG AT GCC TAC CTG CGA ACT GGA)-5'	63	20.7

Table 4.10.2: Thermal melting temperature, T_m, and the % Hyperchromicity, % H, of branched and linear oligonucleotide duplexes. [#]duplex concentration: 3.04 μM in a buffer consisting of 140 mM KCl, 5 mM NaH_2PO_4, 1 mM $MgCl_2$ pH 7.2. The lowercases r and d denote RNA and DNA sequences, respectively. The highlighted region indicates the complementary sequences.

The msDNA and linear duplexes were analyzed by circular dichroism (CD) spectroscopy (310 – 200 nm range) to assess their global helical conformation. The msDNA displayed CD spectral profiles characteristic of A-form helices[173] (**Figure 4.17**) and were in complete agreement with those obtained for the linear control duplexes. This served as further evidence that the branchpoint junction does not considerably influence the structure of the base complementary 'branching' segments.
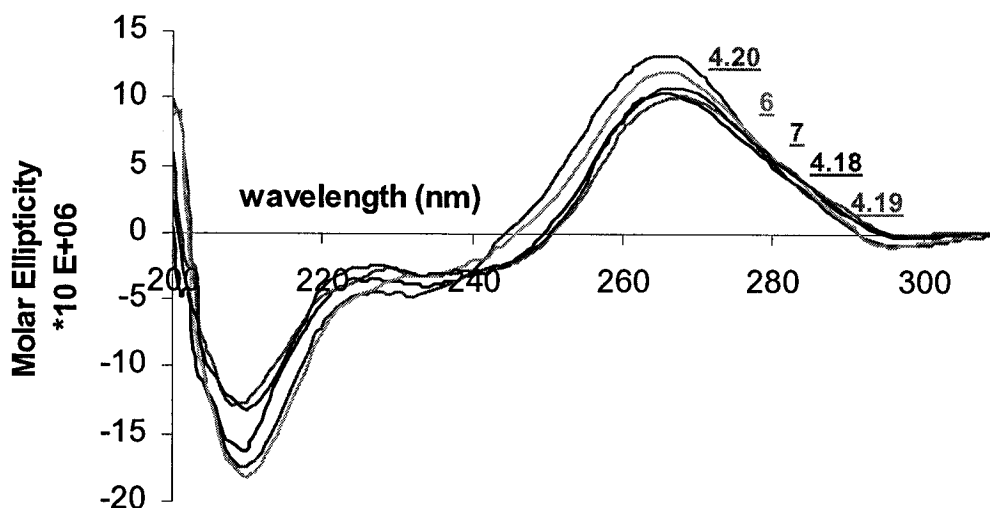
Figure 4.17: CD spectra of complementary branched msDNA, RNA and linear controls. Duplex concentration: 3.04 μM in a buffer consisting of 140 mM KCl, 5 mM NaH$_2$PO$_4$, 1 mM MgCl$_2$, pH: 7.2.

4.11 BIOLOGICAL PROPERTIES OF BRANCHED AND HYPERBRANCHED msDNA SEQUENCES

4.11.1 Enzymatic Debranching of mixed base msDNAs

As stated above (section **4.4.2**), the yeast debranching enzyme (yDBr1) selectively hydrolyzes 2',5'-phosphodiester linkages in bNAs[51,160] and this was applied to confirm the branchpoint connectivities of the msDNAs synthesized (*i.e.* 2',5'/3',5' for **4.14** and **4.15**; 2',5'/3',3' for **4.13a**; and 2',3'/3',5' for **4.12a**). Previous studies on the yDBr1 hydrolytic efficacy have mainly examined as test substrates Y-shaped and RNA lariats with vicinal 2',5' and 3',5'-phosphodiester bonds.[51,160,161,174] The studies described above

(section **4.4.2**) with oligouridylate sequences revealed that yDBr1 is capable of cleaving 2',5'-linkages adjacent to either a native 3',5' or a 3',3'-phosphodiester bonds. Furthermore these studies revealed that branchpoint 2',3' linkages (vicinal to 3',5' linkages) are not cleaved by yDBr1. The relative rates of cleavage observed were: 2',5'/3',5' (native) > 2',5'/3',3' >> 2',3'/3',5' (not cleaved). The same trend is observed with msDNAs of mixed base composition prepared in this study, *i.e.* rate of cleavage of msDNA **4.14, 4.15** (2',5'/3',5') > msDNA **4.13a** (2',5'/3',3') >> msDNA **4.12a** (not cleaved). The cleavage of **4.13a** by yDBr1 confirms that while the 3',5'-linkage is necessary for efficient hydrolysis of the vicinal 2',5'-phosphodiester bond, its presence is not absolutely essential for cleavage. A recent model for the recognition of bNAs by yDBr1 at its active site has recently been proposed,[160,161] however, its focus is related to the hydrolytic mechanism of the scissile 2',5'-linkage by yDBr1, without any indication of the role of the vicinal 3'-nucleotide. Though there is no structural data available on yDBr1:bNA complexes, the results presented here are consistent with those obtained by a previous member of our research group (A. Liscio, M.Sc. 2000, McGill University) who showed that the minimum substrate of yDBr1 is the trimer rA3'p5'rA$^{(3'p*)}$2'p5'rA (p* = phosphate monoester). The 2',5'-linkage of this compound was cleaved by yDBr1, although not as efficiently (5% over 16h) as the branched tetranucleotide rA3'p5'rA$^{(3'p5''G)}$2'p5'rG (*ca.* 50% under the same conditions). The presence of the 3'-phosphate monoester at the internal residue of rA3'p5'rA$^{(3'p*)}$2'p5'rA was essential, since the linear control rA3'p5'rA2'p5'rA (lacking the internal 3'-phosphate group) remained intact upon exposure to the yDBr1 enzyme.
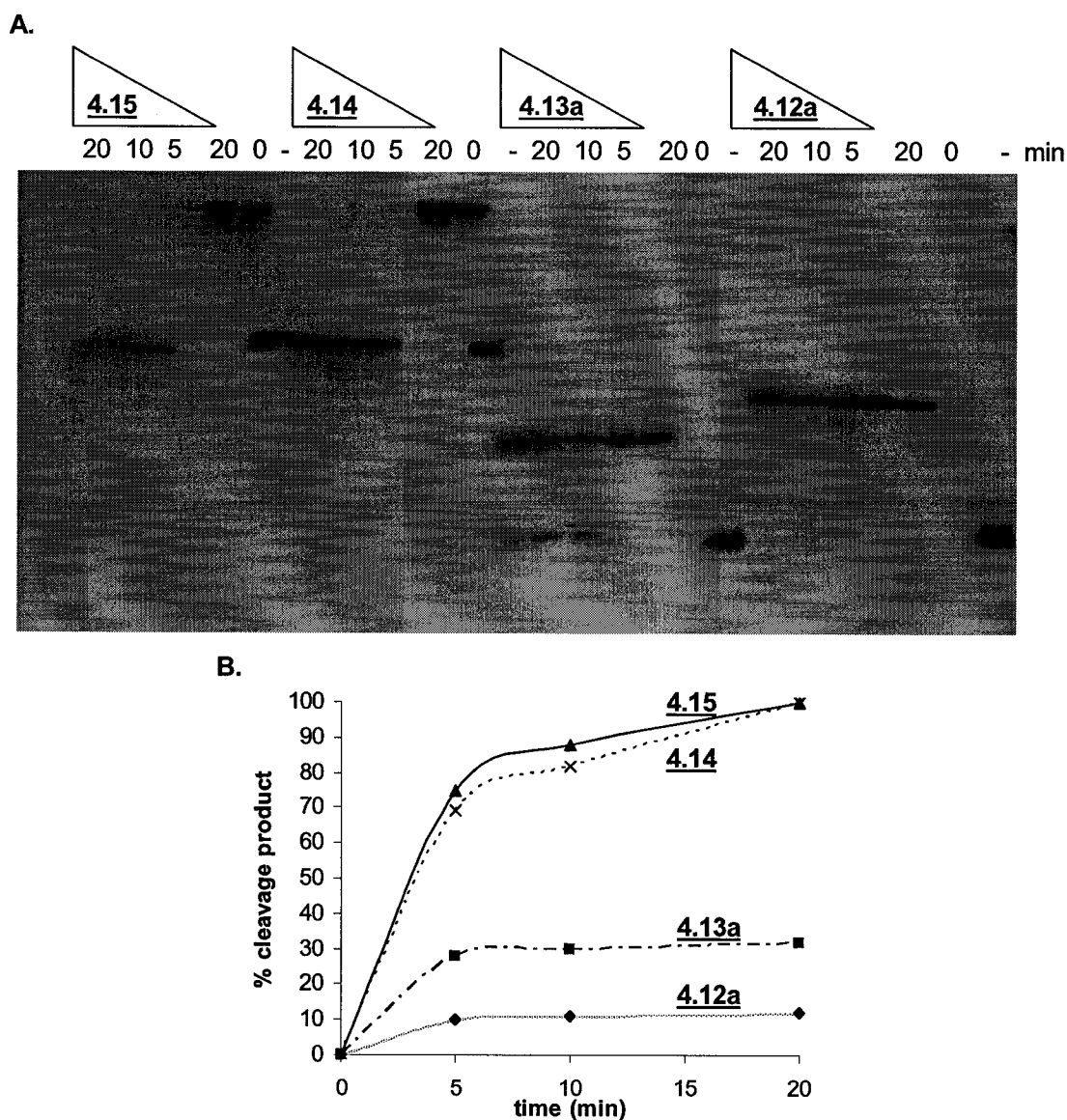
Figure 4.18: A. Enzymatic debranching (yDBr1) of msDNA structures. Lanes marked as "-" refer to the linear controls released upon debranching of **4.12a**: [5'-r(CCG AUC)-3'5'-rU-3'5'-d(GTG AAT)-3'], **4.13a**: [5'-r(CCG AUC)-3'5'-rU-3'5'-d(GTG AAT)-3'], **4.14**: [5'-r(CUA CCG AUC)-3'5'-rU-3'5'-r(CGG GUG GAC G)-3'-p-3'-dT-5'], **4.15**: [5'-r(CUA CCG AUC)-3'5'-rU-3'5'-d(GTG AAA TGA G)-3'], debranched products, **4.13a** : [5'-r(CCG AUC)-3'5'-rU-3'3'-d(GTG AAT)-5'], **4.14**: [5'-r(CUA CCG AUC)-3'5'-rU-3'5'-r(CGG GUG GAC G)-3'-p-3'-dT-5'] **4.15**: [5'-r(CUA CCG AUC)-3'5'-rU-3'5'-d(GTG AAA TGA G)-3']. Approximately 0.5-2 pmol of the 5'-end ^{32}P radiolabeled bNAs were incubated with 10 µL of a solution containing yDBr1 (0.28 mg/mL, 1 µL), buffer (1 µL of 500 mM Tris-HCl pH: 7, 20 mM DTT, 5 mM MnCl$_2$) and water (8 µL) at 20°C for 20 min. **B.** Chart demonstrating the trend in debranching of **4.15** > **4.14** > **4.13a** >> **4.12a** (no debranching) by yDBr1 over a 20-min period.

108

Inhibition studies on yDBr1. The debranching assays described in the previous section show that msDNAs with natural ribose branch-point linkages are substrates of the debranching enzyme (*e.g.* msDNAs **4.14** and **4.15**) whereas other modified msDNAs were not recognized (*e.g.* **4.12a**), or moderately recognized (*e.g.* msDNA **4.13a**), making these compounds potential good choices for future enzyme co-crystallization and X-ray analysis of the yDBr1's active site structure. To this end, we next evaluated whether branchpoint modified msDNA **4.12a**, which was found to be inactive towards yDBr1 hydrolysis, could instead serve to inhibit yDBr1's debranching activity toward wild-type like substrates, *e.g.* **4.14**. Thus, increasing concentrations (5 nM-50 μM) of **4.12a**, **4.13a** and **4.16** were added in a debranching reaction with $5'-^{32}P$ radiolabeled msDNA **4.14** and yDBr1 (**Figure 4.19**). As it can be seen, msDNA **4.13a** did not show any concentration-dependent inhibition pattern. By contrast, msDNA **4.12a** and hyperbranched msDNA **4.16** exhibited good debranching inhibition with an IC_{50} in the 1-2 μM range. The inhibition of debranching by increasing amouns of msDNA **4.16** (itself a potential substrate of yDBr1) may be due to inhibition of formation of the yDBr1:**4.14** complex (*i.e.* yDBr1's activity is 'overwhelmed' hydrolyzing the excess of **4.16** present). In the case of **4.12a** (not a substrate of yDBr1), further studies will be needed to determine whether the inhibition is or isn't competitive. Previous studies indicate the debranching inhibition by ara-A branched RNA, and by $2'-CH_2O-$ modified branched RNA to have an IC_{50} in the 0.5-0.1 μM range.[161] Thus, msDNA **4.12a** exhibits comparable potency in debranching inhibition with those recently reported by our group. Any of these inhibitors would signify a potential choice for enzyme co-crystallization and X-ray analysis of yDBR:msDNA complexes.
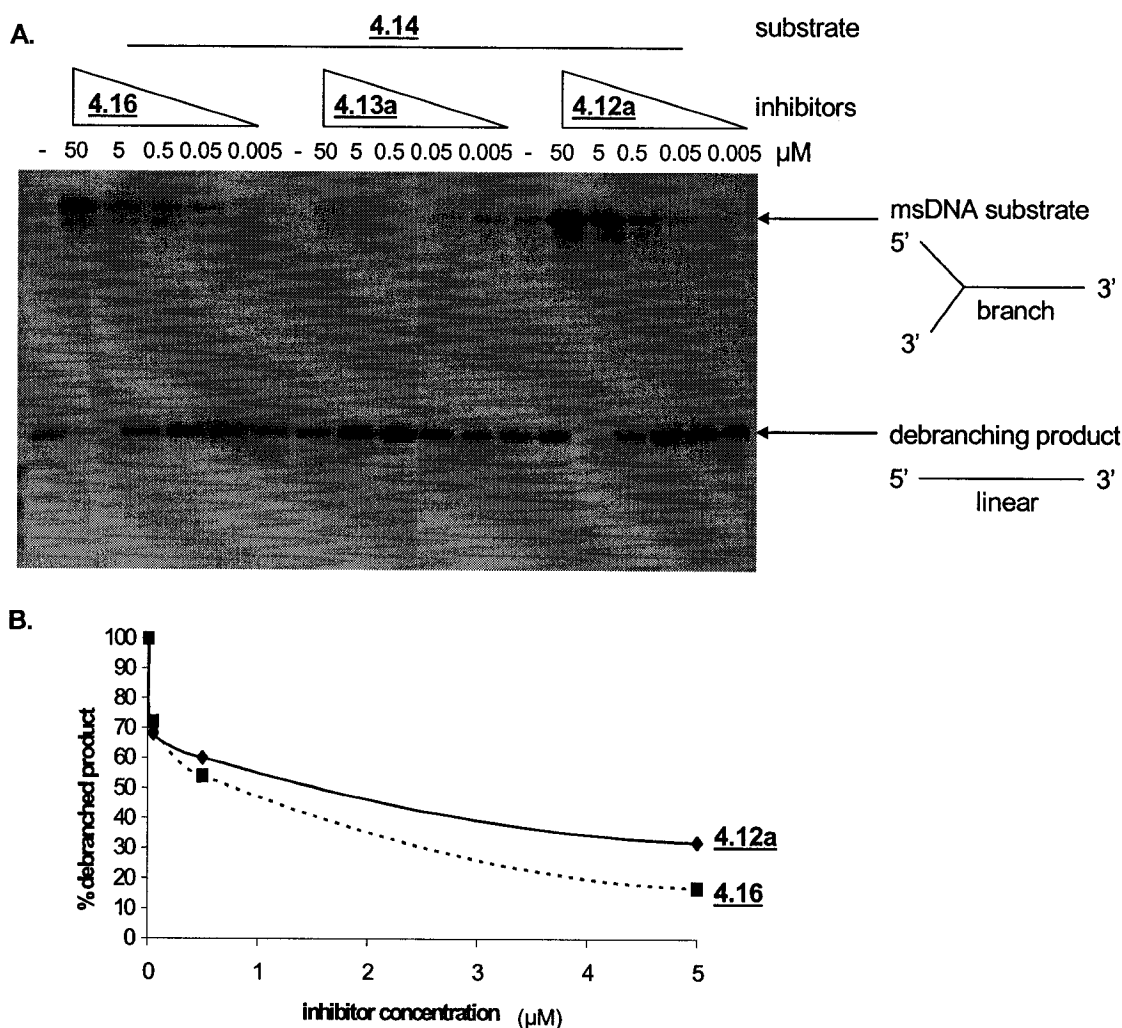
Figure 4.19: The inhibition of yDBr1 with branched and hyperbranched msDNA. **A.** The inhibition of yDBr1 (0.28 mg/mL, 0.5 μL) with inhibitors, **4.12a**, **4.13a** and **4.16** added in increasing concentrations (0.005-50 μM) with 0.5 pmol of the 5'-^{32}P radiolabeled msDNA substrate (**4.14**) at 20°C for 5 minute reaction in yDBr1 buffer (10x buffer: 500 mM Tris-HCl pH: 7, 20 mM DTT, 5 mM MnCl$_2$). **B.** Quantity of enzymatically debranched oligonucleotide product with increasing concentration of msDNA branched and hyperbranched inhibitors. The plot indicates an IC$_{50}$ of 1 and 2 μM for the inhibitors **4.16** and **4.12a**, respectively.

4.11.2 *E. coli* RNaseH Degradation of the Complementary msDNAs

The enzyme RNase H is a ribonuclease that cleaves the RNA strand in a DNA/RNA hybrid to produce 3'-hydroxyl and 5'-phosphate terminated RNA fragments.[175] However, the RNA/RNA duplexes are not generally cleaved by this enzyme.[66] Members of the

RNase H family can be found in nearly all organisms, from archaea to eukaryota. Retroviral RNase H, a part of the viral reverse transcriptase enzyme (RT), is an important pharmaceutical target, as it is absolutely necessary for the proliferation of retroviruses, such as HIV and HBV. Inhibitors of RNase H could therefore provide new drugs against AIDS and Hepatitis B.[176] Activating and inhibiting RNaseH activity with naturally occurring and modified DNA or RNA is an attractive area of research in the development of antisense-based therapeutics[177,178,179,180,181] and studies aimed at understanding the substrate specificity and mechanism of this enzyme.[54,182]

Structural studies have suggested that RNase H exhibits a selective binding directionality with respect to the RNA of the DNA/RNA hybrid substrate such that the binding region ("B" – see **Figure 4.20**) of the enzyme is positioned a few residues from the 5'-end of to the catalytic region ("C" – see **Figure 4.20**). Binding occurs in the minor groove of the hybrid near the 3'-end with respect to the DNA and induces endonucleotidic cuts in the RNA. Then the enzyme exonucleolytically degrades the RNA in the 3'-to-5' direction with respect to the RNA strand.[183]

An objective of this present study was to examine whether msDNAs of opposite strand polarities and backbone composition (*e.g.* **4.18** *vs* **4.21**), upon binding to *E. coli* RNase H, are degraded by the enzyme. This would not only serve to characterize the molecules themselves, but also probe the binding directionality of the RNase H enzyme. Hence, the susceptibility of ^{32}P-labeled msDNAs **4.18**-**4.22** to degradation by *E. coli* RNase H was first examined in 100 mM Tris-HCl, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT (pH: 7.5). This was done in conjuction with both the positive DNA/RNA control and negative RNA:RNA controls (**Figure 4.20**). Once the samples were exposed to the enzyme, they were quenched with the stop solution (deionized formamide with 50 mM EDTA) and analyzed by PAGE (**Figure 4.20**). The degradation of the native DNA/RNA control confirms the viability of the assay. Furthermore, as expected, neither the control RNA:RNA duplex (negative control) or bNA **4.20** served as substrates. By contrast, the enzyme readily hydrolyzed msDNA sequence **4.18** to generate a ladder of bands that is much narrower than that observed for the control DNA/RNA duplex (**Figure 4.20**). This indicates that RNase H binds at the extremity of duplex **4.18** opposite to the branchpoint (**Figure 4.20**). Processive cleavage of the RNA strand in the required 3' to 5' direction

111

would then proceed until RNase H's binding domain sterically clashes with the branchpoint region. This scenario is fully consistent with the degradation pattern observed in the degradation of **4.18** (sequence shown in **Table 4.10.1; Figure 4.20**). Here the 5 nt extension between the duplex region and the branchpoint of the molecule (*i.e.* msDNA **4.18** → **4.19**) led to widening of the ladder-like pattern by approximately the same extent (**Figure 4.20**).

The above model is also fully consistent with the degradation pattern observed with msDNAs **4.21** (and **4.22**). These compounds are regioisomers of **4.18** (and **4.19**), differing in the connectivity (and polarity) of the DNA and RNA chains (see **Table 4.10.1**). Again, due to steric reasons, and the preferential binding of RNase H on the 5'-end of the RNA, interaction of the enzyme with these msDNAs is expected to occur away from their branchpoint. In these cases, endonucleotidic cleavage of the labeled RNA strand is expected to generate a short ^{32}P-labeled RNA fragment, with fast electrophoretic mobility. Indeed, this was observed to be the case, although the extent of cleavage was only (*ca.* 10%) (**Figure 4.20**). Since exonucleolytic cleavage can only occur in the 3'-to-5' direction, the enzyme is now unable to cleave these msDNAs any further. Alternatively, one may argue that **4.21** and **4.22** interact weakly with the enzyme, however, the competitive binding assay described below shows this is not the case. At this point, it was decided to conduct a series of competitive binding studies to determine whether RNase H was actually binding to the poor substrates identified. The reduced (or absent) RNaseH activity for the complementary msDNAs, **4.21**, **4.22** and the bRNA duplex **4.20** also provided potential inhibitors for the *E. coli* RNaseH enzyme. The inhibitors were prepared by serial dilutions of a 500 μM stock solution in buffer (5x RNaseH buffer: 100 mM Tris-HCl, pH: 7.5, 100 mM KCl, 10 mM MgCl$_2$, 1 mM DTT) and added in increasing concentrations (5 nM – 50 μM) to the reaction containing the dT$_{15}$/^{32}p-rA$_{15}$ hybrid as the substrate[184] and the *E. coli* RNaseH enzyme at 20°C for a 5 minute reaction. After digestion by the enzyme, the samples were quenched with the stop solution and the reaction samples were analyzed by PAGE (**Figure 4.21**). The results confirmed that the best inhibitors were, in addition of the bRNA duplex **4.20**, the bNAs with the 3'-end of the RNA strand linked the branchpoint core (*e.g.* **4.22**). These

112

sequences exhibited potent inhibition of the enzyme at low nM concentrations; IC_{50} values of **4.22** and **4.20**, were measured as 25 and <5 nM, respectively (**Figure 4.21**).
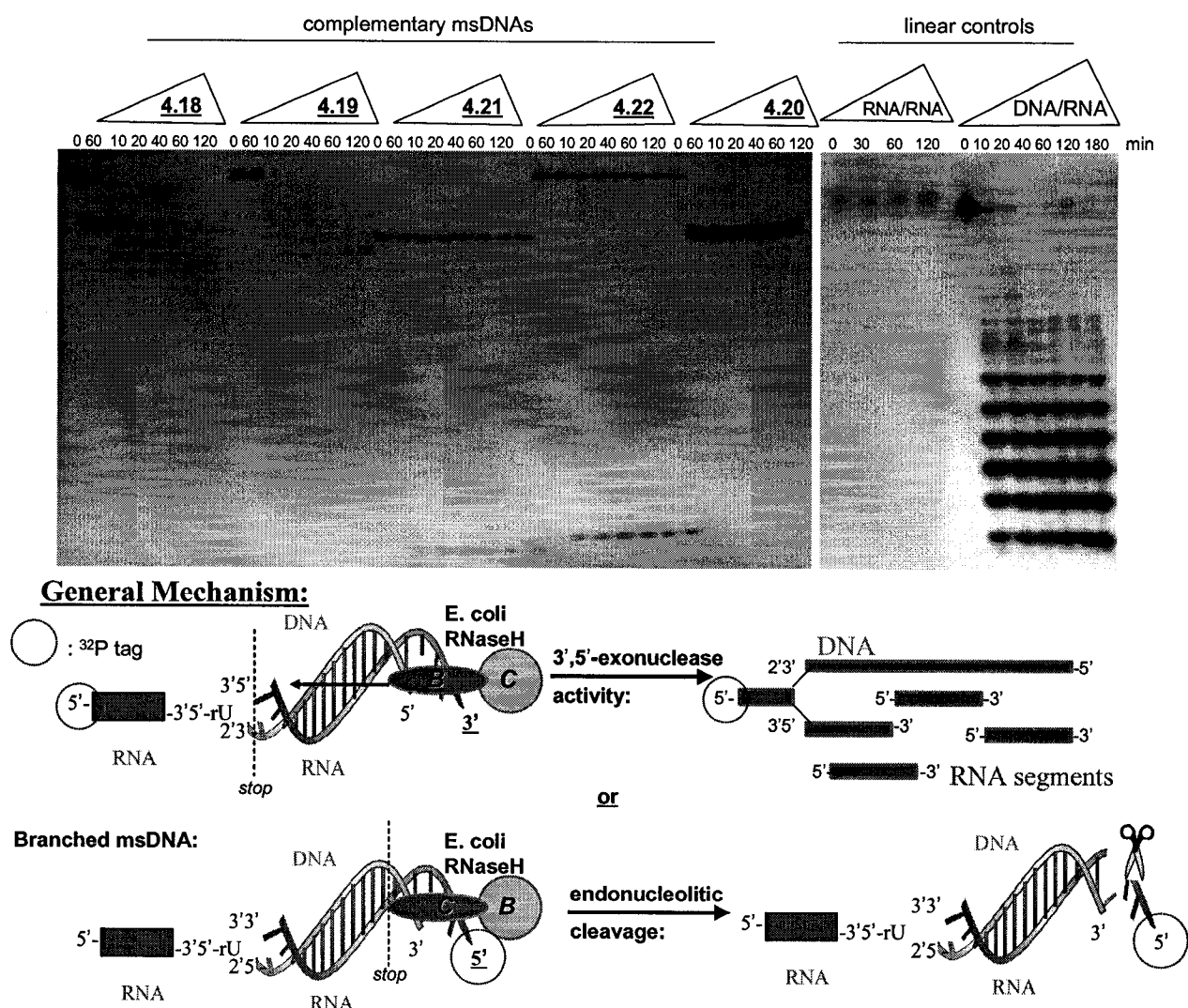
Figure 4.20: The complementary branch and linear msDNA hybrids (0.5 pmol substrate) were completely resistant to hydrolysis in buffer alone (100 mM Tris-HCl, pH: 7.5, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT) at 20°C for 60 min. *E. coli* RNaseH hydrolysis of complementary branched msDNA sequences **4.18**, **4.19**, **4.21** and **4.22** in addition to branched RNA hybrid, **4.20**. The linear RNA/RNA (template RNA sequence: 5'-CGG AUG GAC GCU UGA CCU-3') and the linear DNA/RNA (template RNA sequence: 5'-CGG AUG GAC GCU UGA CCU-3') were used as controls and indicated the expected enzymatic response with the *E. coli* RNaseH. The branched and linear msDNA substrates were added in *ca.* 0.5 pmol quantities in 10 µL reaction volumes at 20°C up to 2 hours reaction with 0.5 µL of *E. coli* RNaseH (5 units/µL) and 2 µL of buffer.

113
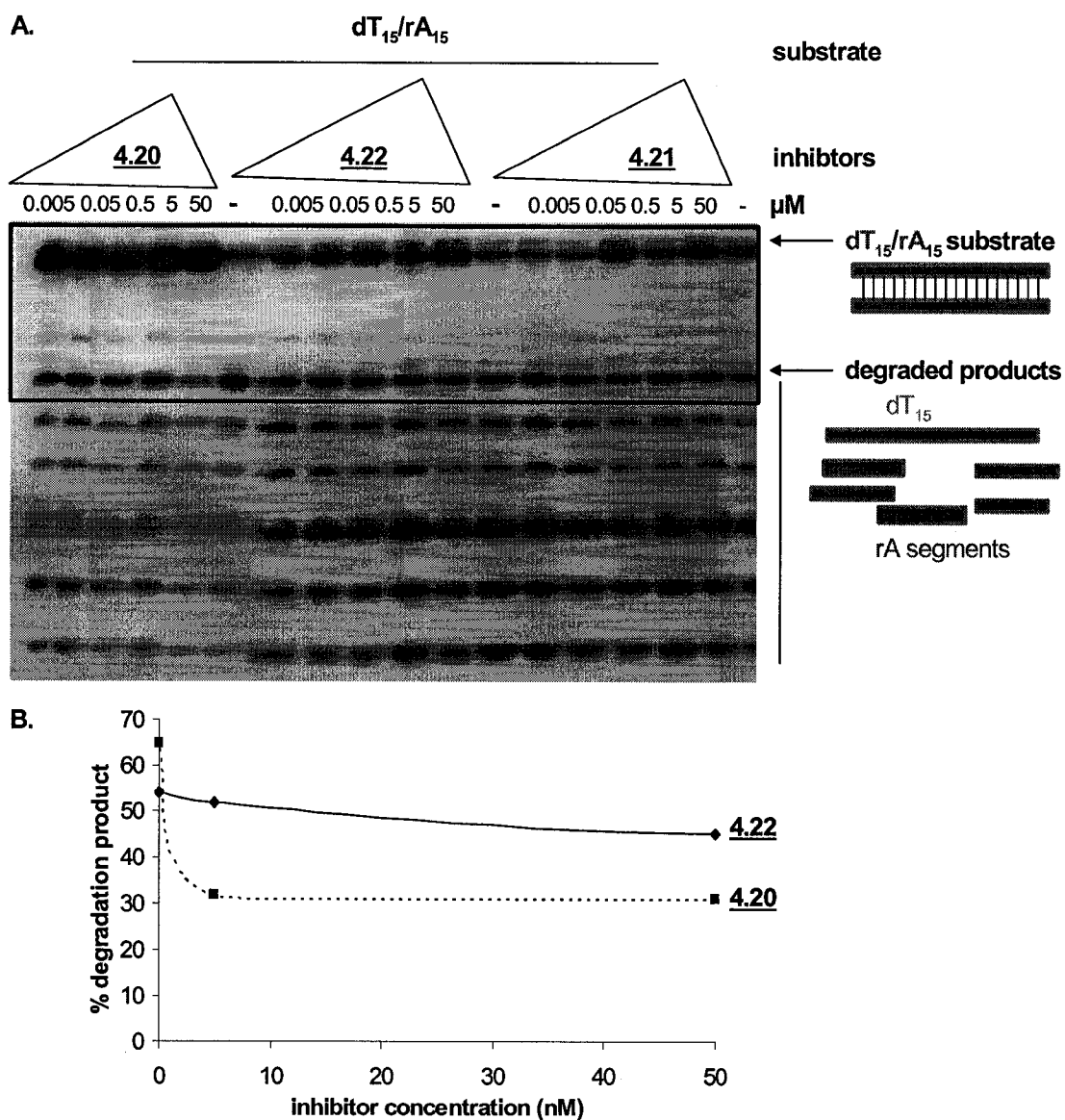
Figure 4.21: Inhibition of *E. coli* RNaseH by msDNA and bRNA sequences. A. PAGE analysis for the inhibition of *E. coli* RNaseH (5 units/μL, 0.5 μL) by inhibitors, **4.21**, **4.22** and **4.20** added in increasing concentrations (0.005 -50 μM) in the presence of 0.5 pmol of ^{32}P radiolabeled substrate (dT_{15}/rA_{15}) and 2 μL of buffer (100 mM Tris-HCl, pH: 7.5, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT) at 20°C for 5 min. B. Chart demonstrating the densiometric ratio of intact full length product to the enzymatically degraded oligonucleotide product (within the reactangular box) with increasing concentration of the msDNA branched inhibitors. The plot indicates an IC_{50} of <5 and 25 nM for inhibitors **4.20** and **4.22**, respectively.

114

4.12 SUMMARY AND CONCLUSIONS FROM THIS STUDY

A novel and efficient divergent method for the synthesis of discrete, and chemically diverse branched and hyperbranched msDNA sequences was developed. The synthesis of msDNAs containing atypical branchpoint linkages (2',5'/3',3' or 2',3'/3',5') were synthesized, and fully characterized by mass spectrometry, and chemical and enzymatic assays. These modified Y-shaped oligonucleotides serve to probe the substrate specificity of the yeast lariat RNA debranching enzyme (yDBr1). For example, these studies revealed that yDBr1 is capable of debranching the 'un-natural' 2',5'/3',3' branchpoints. While debranching activity was considerably less than that observed with msDNAs with the natural (2',5' and 3',5'), the results show that the presence of a 3',5'-linkage is not absolutely necessary for debranching activity. These studies also led to the identification of a potent inhibitors of yDBr1, making these compounds potential good choices for future enzyme co-crystallization and X-ray analysis of the yDBR-msDNA complexes. Additionally, branched msDNA duplex structures provided mechanistic information related to the hydrolysis reaction catalyzed by *E. coli* RNaseH. These sequences generated substrates and inhibitors to the enzyme for potential applications in developing better antisense (refer to Chapter 5) based therapeutics[185] and inhibitors of the RNaseH enzyme.[180]

115

CHAPTER 5: SYNTHESIS AND PROPERTIES OF OLIGONUCLEOTIDES BEARING 6 AND 7-MEMBERED RING CARBOHYDRATES

5.1 INTRODUCTION

Oligonucleotide (ON) analogs are important for purposes of structural studies as well as potential therapeutic agents for gene silencing applications (*i.e.* antisense[54,185], antigene[186], RNAi[99] and aptamer[187] based modalities). In the optimal case for antisense ONs, gene silencing occurs catalytically in concert with a constitutive ribonuclease H (RNase H) enzyme, capable of recognizing and cleaving the RNA strand of ON/mRNA base-paired complexes. Thus, understanding the mechanisms of substrate recognition by ubiquitous nucleases present in most viral, bacterial and eukaryotic organisms provides guidance to the design of future ONs for therapeutic use.[54,188]

Previous studies have indicated that enhanced flexibility of the ON backbone enables better alignment of the ON/RNA hybrid into the active site of RNase H, thereby conferring accelerated enzyme-catalyzed degradation of the mRNA.[60a,238] Accordingly, our group recently synthesized chimeric ONs comprised of 2'-deoxy-2'-fluoro-β-D-arabinose or 2'-β-D-deoxyribose (2'-F-ANA) sugar units with interspersed acyclic (*e.g.* 2',3'-seconucleotide or butyl) residues.[68] These compounds combined both pre-organization (fluoroarabinose moiety) and flexibility (acyclic linker moiety) within the host duplex, leading to efficient RNase H-mediated cleavage of complementary RNA. To further test this hypothesis, we turned our attention to ON backbones that lack the furanose ring altogether, but that would be flexible enough to adjust to the required duplex trajectory for RNase H recognition and cleavage.

To date, the largest expansion of the sugar moiety in DNA is that reported for six-membered ring hexopyranosyl nucleic acids and their assembly into (4'→6') linked oligo(2',3'-dideoxy-β-D-glucopyranosyl)nucleotides (*i.e.* "homo-DNA").[73] This architecture exhibits much stronger Watson-Crick base pairing than DNA, but it does not cross-pair with native DNA or RNA strands. The reason for the tighter base pairing was postulated to be a consequence of the higher rigidity of a pyranose compared to a furanose ring resulting in a preorganization of the single strand's backbone conformation favoring a compact quasi-linear duplex.[75,76] Since then, more-flexible unsaturated 6-

116

membered ring nucleic acids (*e.g.* cyclohexene NA) have been developed to promote binding to the natural systems in hybrid conformations that mimic the native duplexes.[79,225] Alternatively, contracting the carbohydrate moiety to a four carbon skeleton, as in the α-L-threofuranosyl sugar, maintains a furanose half-chair sugar pucker conformation similar to native systems, allowing α-L-threofuranosyl nucleic acids, TNAs, to pair with its complement as well as to single stranded, ssDNA and ssRNA.[80,81]

To this effect, we propose that the nucleosides and oligonucleotides comprising a 7-membered heptose ring structure, adopt a conformation that supplies the necessary balance between structural pre-organization and flexibility for efficient pairing interactions (H-bonding and base stacking) for RNase H induction when binding to complementary RNA.[193,222] A survey of the conformational requirements for binding and RNaseH activity is achieved comparing the 5 (furanose, dT_{15}), 6 (2'-ene-pyranose, pT_{18}) and 7-membered (oxepane, oT_{15}) ring oligonucleotides, in order to develop trends with the inherent flexibility of the constituent sugar rings, *i.e.* furanose[189] (5) > oxepane[222c] (7) > 2'-ene-pyranose[190] (6).

5.2 CONFORMATIONAL ANALYSIS OF oT, pT, dT AND rU

Molecular modeling of nucleoside conformation is a commonly used method for studies of the carbohydrate geometry in oligonucleotide duplex structures.[54] Unsaturated 6-membred rings have been shown to be most stable in their twisted half chair conformations with barriers to interconversion of 8-12 kcal/mol, oscillating predominantly via the boat (bent) conformation.[190,191] Previous computational studies have predicted that the unsaturated pyranose nucleic acids (pNA or pT), favors an $^{o}H_{5'}$ half-chair conformation such that the dihedral angle, δ: 85° for C4'-C5' is in the range of ribose, δ: 77 – 97° for C3'-C4', in A-RNA type helices; a requirement for binding to complementary RNA (**Figure 5.1- C**).[192] The latter property is highly desirable for antisense ONs that target cellular mRNA for recognition and activation of the ubiquitous RNase H enzyme.[238]

The pseudorotational equilibrium for oxepane carbohydrates is more complex and involves many low energy transitions between the boat, (B), and twist boat, (TB), chair, (C), and the favored twist-chair, (TC), forms.[193,222] The most common transitions from C

to TC geometry proceeds with relatively low interconversion energy barriers of 0.5-2 kcal/mol whereas some of the less populated TC to B, or B to TB transitions have interconversion energy barriers of 5-10 kcal/mol.[222] An *ab-initio* geometry optimization was performed on oT nucleoside, **5.28**, in-vacuo, with the 3-21 G basis set using HyperChem version 7.0. The Polak-Ribiere conjugate gradient algorithm was used, with a root-mean-square gradient termination cut-off of 0.05 kcal/Å/mol. The calculations provided compact energy minimized structures where the heptose carbohydrate was found to be either in a chair or slightly twisted chair conformation ($^{O'6'}TC_{4'5'}$), with C4' and C5' lying below and the ring oxygen and C6' above the plane of the molecule defined by C1'-C2'-C3' (**Figure 5.1- D**). This conformation restricted the susbstituents to a sterically equivalent isoclinal trajectory that was further confirmed by the ^{1}H NMR spectrum of the titled compound, **5.28**, and the overlapping of the 2', 3' and 4' methylene proton signals (*ca.* 2 ppm).

This preliminary molecular modeling was also performed on the dT and rU nucleoside structures for conformational analysis using the previously defined parameters (**Figure 5.1- A and B**). Comparison of the oT conformation with the C2'-*endo*[194] sugar pucker of dT, (**Figure 5.1- A and D**) resulted in an expanded oxepane carbohydrate structure in which the in-ring diameter of the molecule was found to be extended relative to the pentofuranose ring by almost 1 Å (*i.e.* oT: C1'-C5': 3.42 Å and dT: C1'-C4': 2.38 Å), while the O-O distance in oT was found to be reduced by almost 0.8 Å (*i.e.* oT: C7'OH-C5'OH: 3.35 Å and dT: C5'OH-C3'OH: 4.19 Å). Therefore, the oxepane carbohydrate conformation in oT could provide a more compact oligonucleotide structure that can, in principle, hybridize to DNA; and even more favorably, to the more compact C3'-*endo*[194] geometry found in RNA (*i.e.* oT: C7'OH-C5'OH: 3.35Å and rU: C5'OH-C3'OH: 3.87 Å) (**Figure 5.1- B and D**).

Nucleoside **Conformation**
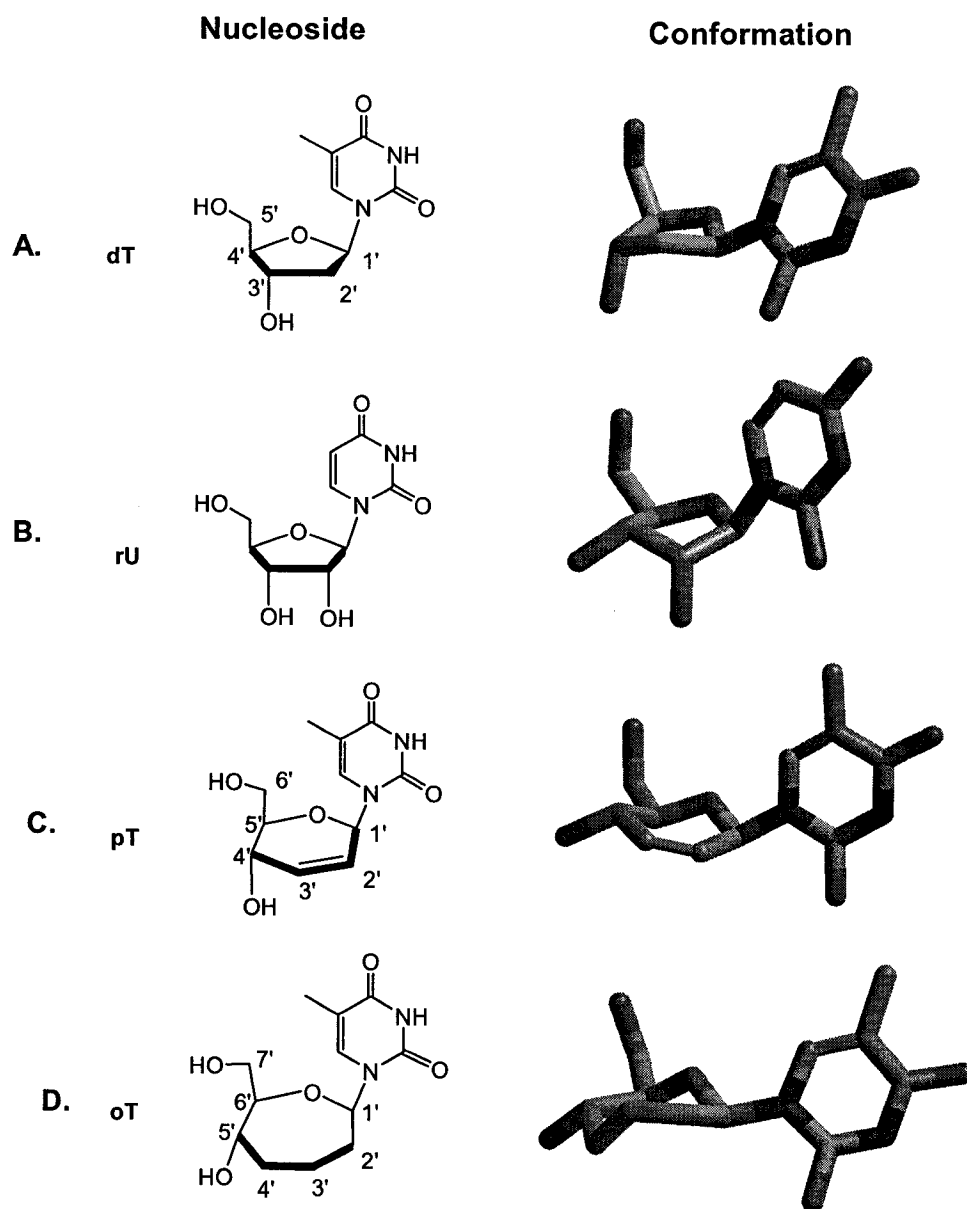
Figure 5.1: Nomenclature, structure and conformation of **A.** dT **B.** rU **C.** pT **D.** oT nucleic acids

5.3 UNSATURATED 6 AND 7-MEMBERED RING NUCLEOSIDES AND OLIGONUCLEOTIDES

5.3.1 Objectives.

This study describes the conversion of a commercially available glycal, 3,4,6-tri-*O*-acetyl D-glucal, TAG, into unsaturated ene-pyranose, (*i.e.* **5.7** and **5.12**) and oxepine, (*i.e.*

119

5.20) nucleosides. It was rationalized that the existing glycosylation methods (**Figure 5.3**) could facilitate their syntheses with persilylated nucleobase thymine (2,4- bis-*O*-trimethylsilyl thymine) in the presence of trimethylsilyl trifluoromethanesulfonate (TMSOTf) as Lewis acid catalyst. These modified nucleosides can function as inhibitors of DNA synthesis[195] and their conversion into DNA phosphoramidite analogues provided a means for incorporating them into oligonucleotides for antisense[196] and siRNA[97] applications.

A synthetic requirement for modified nucleoside and oligonucleotides is linked with developing a compatible solid phase synthesis procedure.[197] The generation of modified oligonucleotides, MONs, requires high purity phosphoramidites capable of efficient coupling on a solid support without competing side reactions that can modify the oligonucleotide. Moreover, MONs must also be resistant to the post synthesis 'work-up' and purification procedures prior to characterization and evaluation of their properties.

This work discusses some of the preliminary results based on the synthesis of modified 6 (pT and pT*) and 7 (oT*) -membered ring nucleoside analogues and their solid phase synthesis into MONs (**Figure 5.2**) for antisense applications. These include the generation of nuclease resistant oligonucleotides that can mimic the natural function of DNA (*i.e.* hybridize to complementary RNA) and elicit RNaseH activity.[54,60]

ene-pyranose nucleic acids
(pNA)

oxepine nucleic acids
(ONA)
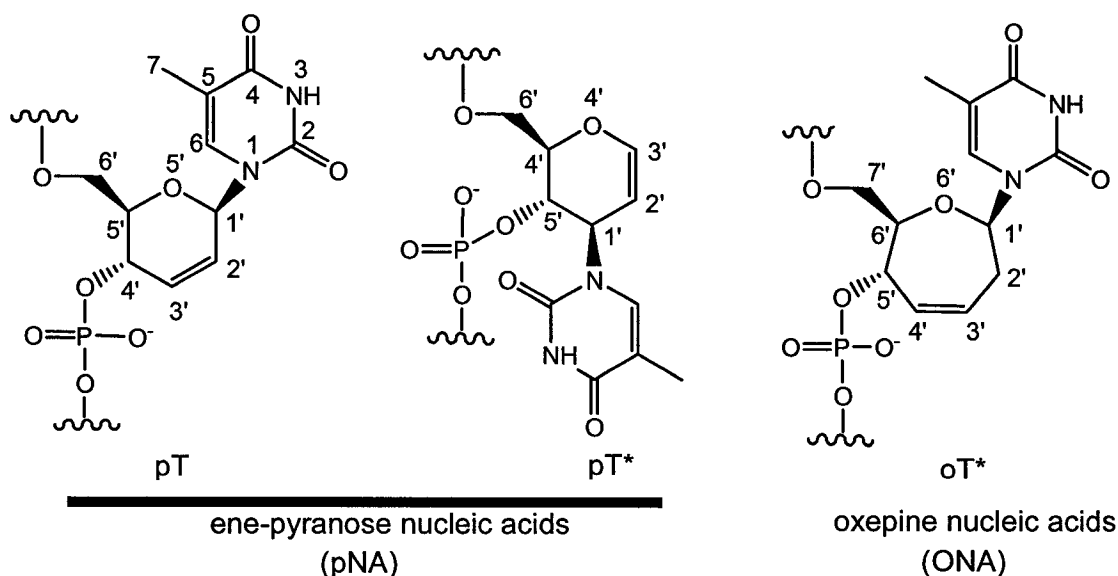
Figure 5.2: Structures of ene-pyranose and oxepine nucleotides and oligonucleotides

5.3.2 Synthesis and Characterization of Nucleosides and Phosphoramidite Derivatives for Solid Phase Oligonucleotide Synthesis.

1,2-Unsaturated *O*-glycosides have been used as chiral precursors for the total synthesis of important naturally-occurring and biologically active oligosaccharides, glycopeptides, modified carbohydrates and nucleosides.[198,199,200,201] Their ease in preparation from commercially available and economical pyranose starting materials as well as their reactivity in regio- and stereo-selective addition[202] and elimination[203] reactions makes the glycal a valuable synthon for the *de-novo* design and synthesis of chiral carbohydrate derivatives. The synthesis of vinylic 6-membered ring oxiranes *i.e.* unsaturated pyranose carbohydrates, can be accomplished by Ferrier-type[204] glycosylation reactions in the presence of a Lewis acid catalyst. These reactions commonly proceed with a reaction mechanism that involves nucleophilic addition to either the C1 or the C3 position of the carbohydrate ring in regio- and diastereo-selectivity that depend on the nucleophile, Lewis acid catalyst, the glycal starting material and reaction conditions (**Figure 5.3- A**).[205]

A specific example is related to the synthesis and applications of expanded nucleic acids (sugar ring and heterocyclic base), with promising biological and functional properties as nucleoside and oligonucleotide analogues.[75,79,206] The commercially available glycal, TAG[207], has been converted to the 2',3'-unsaturated pyranose nucleoside in the presence of Lewis acid catalysts such as TMSOTf[208], trityl perchlorate[209] or $SnCl_4$[210] for the potential application in developing anti-virals such as the AZT-like compounds[211] and antisense oligonucleotides[212]. However, this modification has not yet been evaluated as an antisense construct (*i.e.* 18-21 base length oligonucleotide), bound to complementary RNA for gene-silencing applications and RNaseH induction for the degradation of mRNA.

The ring expanded sugars such as the 7-membered heptoses are important in nature and posess interesting biological properties. These include the diterpenoid oxepane natural products (*i.e.* Zoapatanol[213]), and the oxepane (or oxepine) fused polycyclic marine toxins[214]. However, there exist only few examples of synthetic 7-membered carbohydrate rings and their derivatives (*i.e.* 7-membered ring nucleic acids) due to the difficulty in

their synthesis. Hoberg and co-workers have reported an efficient procedure for their synthesis (*i.e.* 3,4-unsaturated oxepins- **Figure 5.3- B**).[215]
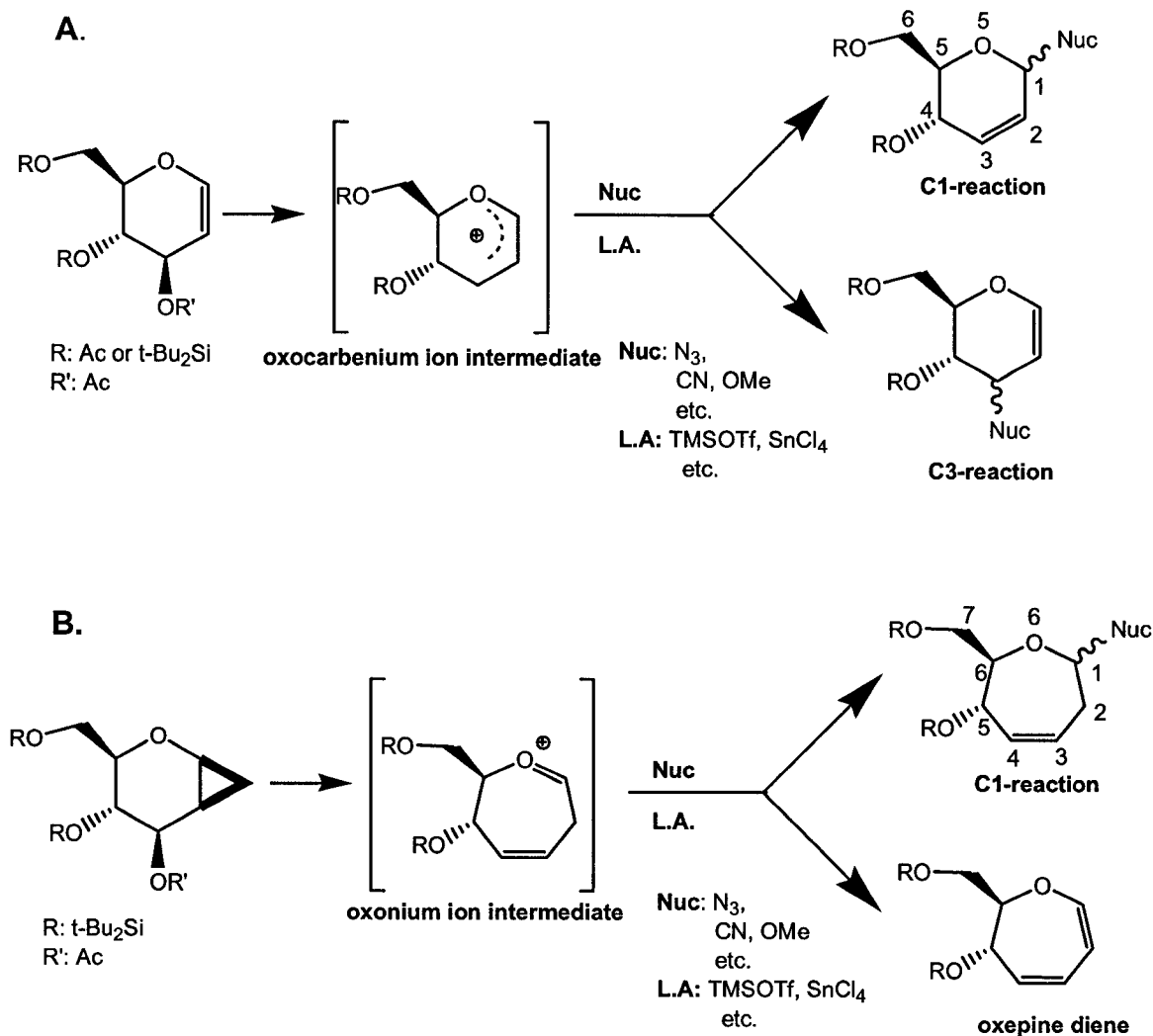
Figure 5.3: Glycosylation reactions of 1,2-unsaturated glycosides and their derivatives. **A.** Synthesis of 1,2 and 2,3-unsaturated pyranose derivatives from the C3 and C1 addition reactions, respectively and **B.** the synthesis of 3,4-unsaturated oxepine derivatives from the C1 addition and the carbohydrate ring expansion reaction and the oxepine diene formed from the concomitant elimination reaction of the oxonium ion intermediate.

C3-addition with persilylated thymine (Figure 5.4). The starting material TAG, **5.1**, was initially deacetylated in basic conditions (0.1 M NaOMe in MeOH) until complete reaction was indicated by TLC. This was then silylated regioselectively at the C4 and C6 positions of the sugar unit with di-*t*-butylsilyl bis(trifluoromethanesulfonate), *t*-

Bu$_2$Si(OTf)$_2$, under kinetically controlled conditions.[216] The silylated glycal, **5.3**, was acetylated prior to the Vorbrüggen-like[217] coupling reaction of the persilylated nucleobase and the protected glycal intermediate **5.4**. The coupling reaction required silylating thymine with hexamethyldisilazane, HMDS, during reflux in MeCN. The silylation reaction was monitored by the dissolution of the product, 2,4-bis-*O*-TMS thymine in MeCN.[218] This was added to the glycal, **5.4**, and reacted with the Lewis acid catalyst, TMSOTf (30 mol %), in MeCN with reflux during 4 hours. This yielded the unsaturated pyranose nucleoside diastereomers, **5.5** and **5.6**, in an 89% total yield with a 1:2 ratio in $\beta:\alpha$ selectivity at the pseudo-anomeric C3 position of the glycal. The β-anomer, **5.5**, was isolated by chromatography and desilylated with 1 M TBAF in THF to yield the deprotected nucleoside **5.7**. This was converted to the C6'-MMT nucleoside, **5.8**, which was phosphitylated at the C5' hydroxyl group to yield the phosphoramidite, **5.9**.
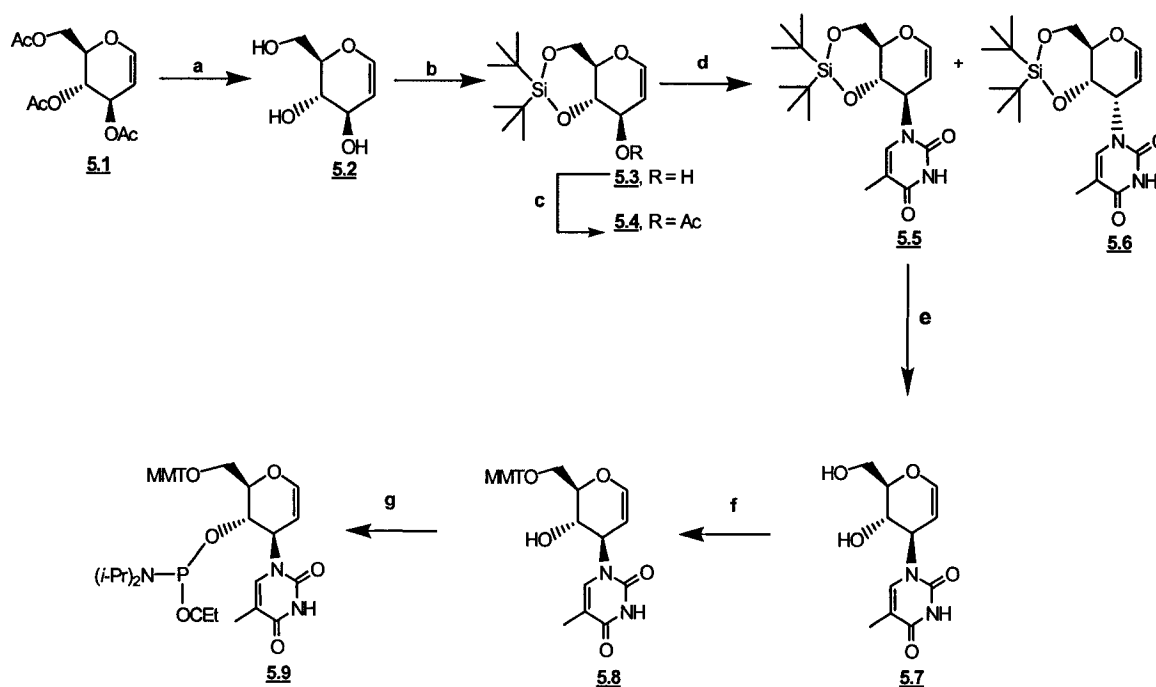
Figure 5.4: Reaction and Conditions: a. 0.1 M NaOMe in MeOH, 22°C, 3 h (99%) b. *t*-Bu$_2$Si(OTf)$_2$, DMF, pyr, -40°C, 45 min (88%) c. Ac$_2$O, DMAP, pyr, 22°C 1 h (88%) d. i. HMDS, thymine, MeCN, 90°C, 3 h ii. TMSOTf, MeCN, 6 h , 90°C (89%) e. 1 M TBAF in THF, 0°C, 1 h (93%) f. MMT-Cl, pyr, 22°C, 5 h (60%) g. Cl-P(OCEt)N(*i*-Pr)$_2$, EtN(*i*-Pr)$_2$, THF, 22°C, 1 h (83%).

The nucleoside diastereoisomers, **5.5** and **5.6**, were separated by chromatography and their structures assigned by ^{1}H-^{1}H NOESY correlation experiments in which the more non-polar β-isomer, (R$_f$: 2:1 Hex:EtOAc, 0.25), exhibited a [thymine H6]$\leftrightarrow$[C6' methylene protons] NOE cross-peak whereas the α-isomer, (R$_f$: 2:1 Hex:EtOAc, 0.19), displayed an NOE cross-peak for H6$\leftrightarrow$H4' (**Figure 5.6- A**). The modest diastereoselectivity can be rationalized by the absence of the anomeric effect or anchimeric assistance during the coupling and rearrangement reaction of the 1,2-unsaturated glycal with the persilylated thymine nucleophile.[204b]

The regioselectivity of the reaction was further ascertained by ^{1}H-^{13}C HMQC NMR (**Figure 5.6- B**) which indicated that the characteristic anomeric carbon, (*i.e.* δ: 50-80) was transposed to the pseudo-anomeric C1' position of the glycal moiety. This was also supported by the chemical shift values of the vinylic carbons, (δ: 140 and 100), which were found at the C2' and C3' positions of the nucleoside. The regioselectivity of the coupling reaction (**Figure 5.3- A**) can be partially rationalized by the influence of the rigid bicylic conformation of the oxocarbenium ion of the glycal intermediate, **5.4**, which favors the equilibrium conversion to the thermodynamically more favored product, **5.5** and **5.6** in an allylic rearrangement reaction.[219]

The β-diastereomer, **5.5** was isolated by silica gel column chromatography and desilylated to the deprotected nucleoside **5.7**. The conformational analysis for **5.7** by the ^{1}H and ^{13}C NMR data generated a nucleoside conformation, pT*, which favored the $^{5'}H_{4'}$ half-chair geometry with C5' and C4' lying above and below the plane of the molecule defined by C1'–C2'–C3'–O4' (**Figure 5.5**). This was in agreement with the literature data providing coupling constants of $J_{1'2'}$: 3 Hz and $J_{4'5'}$: 5.6 Hz and the chemical shift of the pseudo-anomeric proton which was found to be relatively shielded (by comparison with **5.12**: see C1 addition with thymine) by ^{1}H NMR δ: 5.04 and by ^{13}C NMR δ: 50.[219]
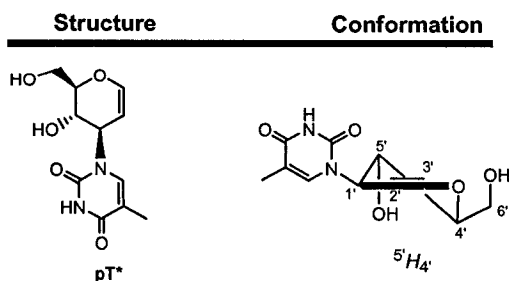
Figure 5.5: Structure and conformation of (**5.7**) pT*.

The nucleoside, **5.7**, was converted to the monomethoxytritylated nucleoside phosphoramidite, **5.9**, and analyzed for purity by ^{31}P (^{1}H decoupled) NMR (**5.9**: δ 150.2 and 149.0). Their structures were confirmed by ^{1}H NMR and ESI mass spectrometry (*i.e.* $C_{40}H_{47}N_{4}O_{7}PNa$ calc. 749.8, found 749.3).
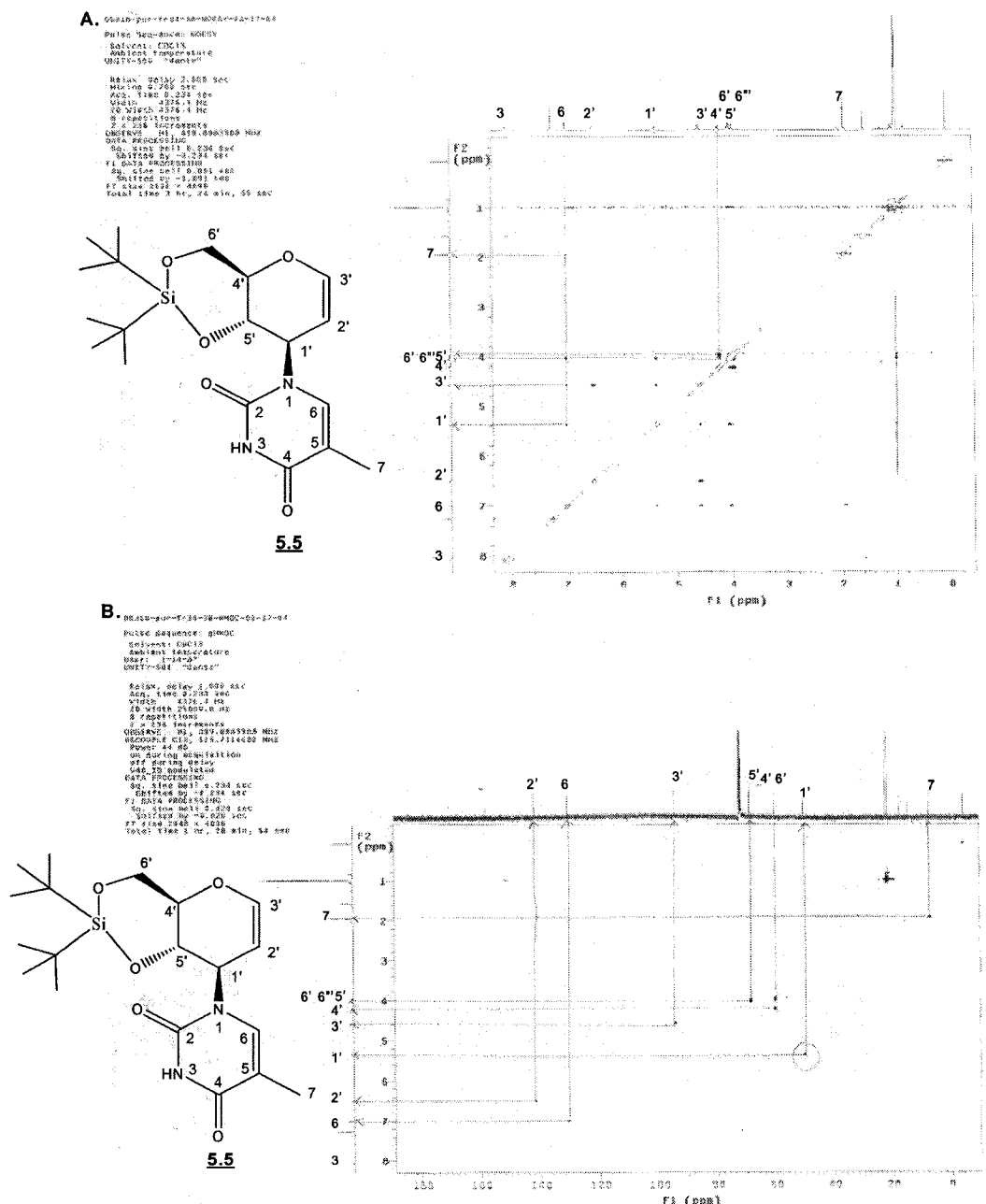
Figure 5.6: Characterization of **5.5** by: **A.** ^{1}H-^{1}H NOESY NMR indicating the β-diastereomer with the NOE cross-peak with C6 proton of thymidine base with the C6' protons of the glycal sugar and **B.** ^{1}H-^{13}C HMQC NMR indicating the regiochemistry of the thymine base by the pseudo-anomeric proton at C1'.

125

C1 addition with thymine (Figure 5.7). In order to compare the reactivity and selectivity of the glycosylation reactions, **(Figure 5.4** and **5.7)** with persilylated thymine as nucleophile, the synthesis of nucleoside **5.12** was accomplished by the literature procedure.[220] This coupling reaction was performed with TAG, which contains the C4 and C6 acetyl groups that favor a more conformationally labile glycal conformation relative to the di-*t*-butylsilyl protecting group in **5.4**.[216,219] The reaction of persilylated thymine with TAG and TMSOTf (30 mol %) as Lewis acid catalyst, was refluxed to completion in MeCN for 3 hours. Purification of the crude product afforded **5.10** and **5.11** in a 1:1.1 β:α ratio (75% overall yield). The structures of **5.10** and **5.11** were confirmed by the ^{1}H and ^{13}C NMR chemical shift data, which were in complete agreement with the literature reported data.[209] These were purified by silica gel column chromatography and the β-anomer, **5.10**, was deacetylated with 0.1 M NaOMe in MeOH for 1 hour generating **5.12** in 93% yield. This was monomethoxytritylated at the primary C6' (78%), and phosphitylated (80%) at the secondary C4' hydroxyl functional groups to yield the phosphoramidite, **5.14**, as a mixture of P-diastereomers suitable for solid phase oligonucleotide synthesis **(Figure 5.7)**.
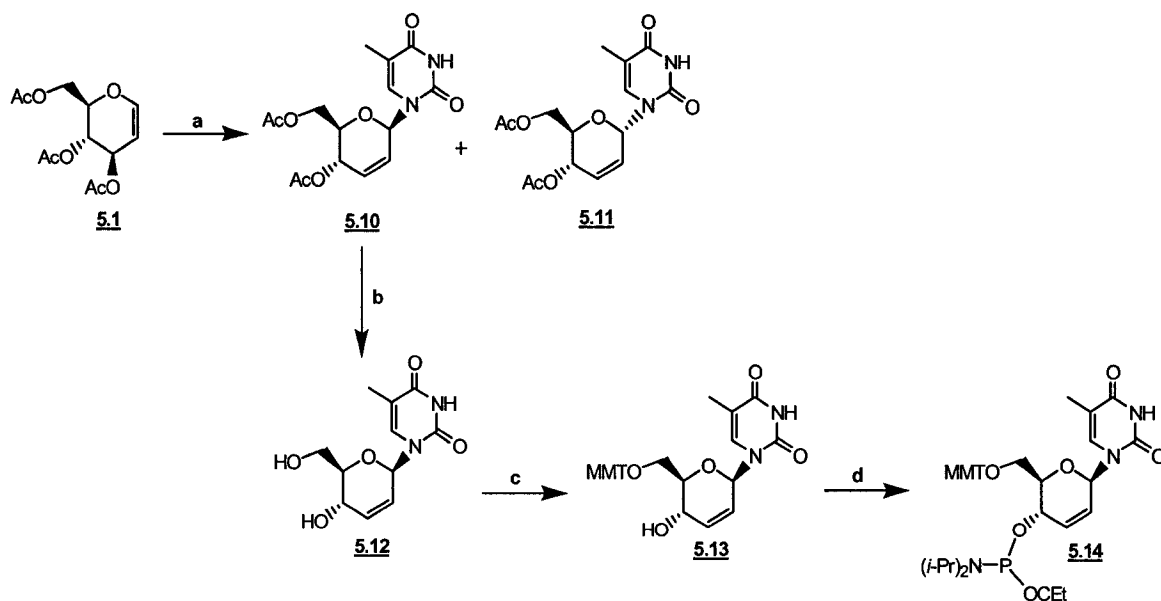
Figure 5.7: Reaction and Conditions. a. i. HMDS, thymine, MeCN, reflux 3 h ii. TMSOTf, persilylated thymine, MeCN, reflux 3 h (75%), b. 0.1 M NaOMe in MeOH 1 h (93%), c. MMT-Cl, pyr, 22°C, 3 h (78%), d. Cl-P(OCEt)N(*i*-Pr)₂, EtN(*i*-Pr)₂, THF, 22°C, 1 h (80%).

Figure 5.8: Characterization of **5.10** by: **A.** NOESY NMR indicating the β-diastereomer containing the NOE cross-peak of C1' with the C5' protons in **5.10** and **B.** ^{1}H-^{13}C HMQC NMR indicating the position of the thymine base by the anomeric proton at C1'.

127

The coupling reaction of persilylated thymine with TAG, was monitored to completion by TLC and produced a 75% total yield for the more non-polar β-anomer, **5.10** (R_f: 2:1 Hex:EtOAc, 0.29) and the α-anomer, **5.11** (R_f: 2:1 Hex:EtOAc, 0.20) in a 1:1.1 ratio, respectively. Their structures were confirmed by the ^{1}H and ^{13}C NMR chemical shift values which were in agreement with the literature data.[209] The diastereoselectivity of the reaction was confirmed by NOESY NMR which indicated a NOE cross-peak for the anomeric C1' and C5' protons for the β-anomer, **5.10** (**Figure 5.8- A**). Alternatively, the α-anomer, **5.11**, exhibited an NOE cross-peak for the anomeric C1' and C4' protons (data not shown). The regioselectivity of the reaction was also confirmed by the ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMQC NMR (**Figure 5.8- B**) experiments. The anomeric proton and the carbon chemical shift values were deshielded (δ: 6.52 for H1' and δ: 77.19 for C1') relative to **5.7** (δ: 5.04 for H1' and δ: 50 for C1').

The nucleoside, **5.12**, was finally converted to the methoxytritylated nucleoside phosphoramidite diastereomers, **5.14**, and analyzed for purity by ^{31}P (^{1}H decoupled) NMR (**5.14:** δ 151.8 and 150.1) and the structure confirmed by ^{1}H NMR and by molecular weight with ESI mass spectrometry (*i.e.* $C_{40}H_{47}N_4O_7PNa$ calc. 749.8, found 749.2). The pure phosphoramidite derivative, **5.14**, was employed for their automated incorporation into oligonucleotides.

C1 addition and ring expansion with thymine (Figure 5.9). The chemical synthesis of oxepine nucleosides was inspired by the elegant work of Hoberg.[215] The synthesis strategy was initiated by a deacetylation of TAG followed by the regioselective silylation at the C6 and C4 hydroxyl groups with the reagent t-Bu$_2$Si(OTf)$_2$. The use of the allylic alcohol finctional group in **5.3** provided excellent diastereoselectivity towards the Simmons-Smith cyclopropanation reaction according to the modified Furukawa procedure to yield **5.15** in 92% isolated yield.[221] After acetylation of **5.15**, the completely functionalized glycal, **5.16**, was subjected to the ring expansion reaction of the sugar with persilylated thymine as nucleophile to yield the desired unsaturated 7-membered oxepine nucleoside, **5.17**. The coupling reaction was performed in a two-step procedure, in which thymine, was initially silylated with HMDS, and refluxed to yield the persilylated base. The glycal starting material **5.16**, was transferred to the persilylated base and with

TMSOTf as Lewis acid catalyst (30 mol %) the reaction was refluxed in MeCN and produced the oxepine nucleoside diastereomers, **5.17** and **5.18**, and the oxepine diene side product, **5.19**. The chemoselectivity of the reaction (coupling of base versus elimination reaction) was found to be dependent on the potency of the nucleophile (refer to section 5.5.1).[215] The reaction with silylated thymine as nucleophile was slow, resulting in the complete reaction of **5.16** after reluxing in MeCN for 24 hours. The product distribution produced a modest chemoselectivity of 40 %:30 % in favor of the nucleosides, **5.17** and **5.18**, relative to **5.19**. Furthermore, the slow reaction with thymine produced a diastereoselectivity of 10:1 at the anomeric position in favor of the desired β-anomer, **5.17**. The reaction products, **5.17**, **5.18**, and **5.19** were purified and isolated by chromatography prior to characterization by NMR and ESI MS. The desired oxepine nucleoside **5.17** was desilylated with a 1 M TBAF in THF solution during a 1 hour reaction. The reaction progress was monitored by TLC and by the poor solubility of the deprotected nucleoside, **5.20**, in THF. The deprotected oxepine nucleoside, **5.20** was monomethoxytritylated at C7' and phosphitylated at the C5' hydroxyl group forming phosphoramidite **5.22**, suitable for solid phase oligonucleotide synthesis (**Figure 5.9**).
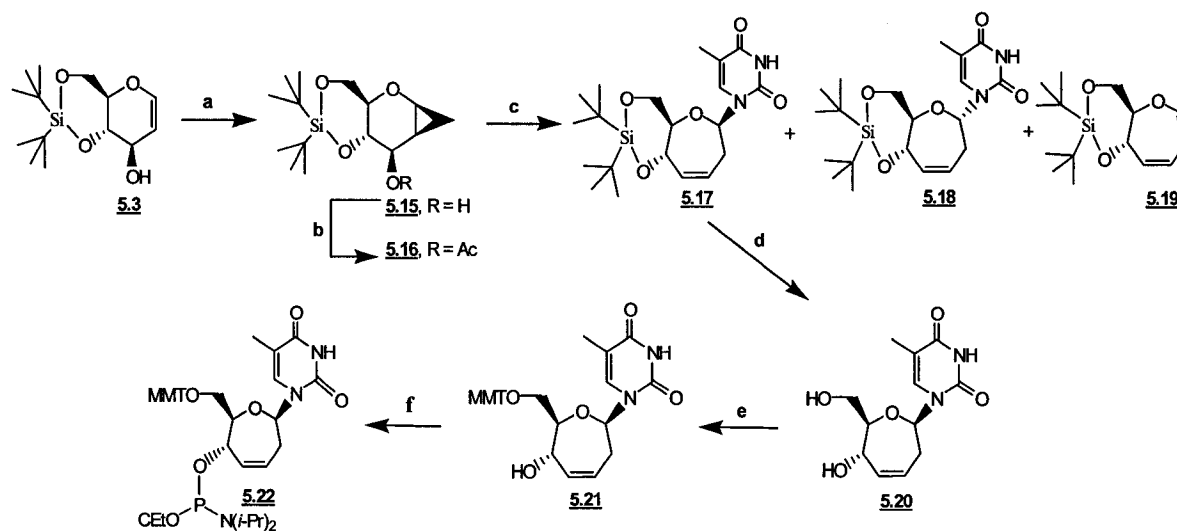
Figure 5.9: Reaction and Conditions. a. 1 M ZnEt2:hexanes, CH2I2, Et2O, 0°C, 4 h (92%) b. Ac2O, DMAP, pyr 1 h 22°C (92%) c. i. HMDS, thymine, MeCN, reflux, 3 h ii. persilylated thymine, TMSOTf, MeCN, reflux overnight (67%) d. 1 M TBAF in THF 0°C, 1 h (90%) e. MMT-Cl, pyr 22°C, 3 h (65%) f. Cl-P(OCEt)N(*i*-Pr)2, EtN(*i*-Pr)2, THF, 22°C, 2 h (80%).

The stereochemistry at the anomeric position for the oxepine nucleoside diastereomers, **5.17** and **5.18** was ascertained by NOESY NMR, which indicated a NOE cross-peak for the C1' and C6' protons in the designated β-anomer, **5.17**, and the C1' and C5' protons for the α-anomer, **5.18** (**Figure 5.10**). The characteristic ^{1}H signals for the more polar α-anomer (R_f: 2:1 Hex:EtOAc, 0.20) were found to be deshielded relative to those of the β-anomer (R_f: 2:1 Hex:EtOAc, 0.22); e.g. the H1' at δ: 5.75 and δ: 5.99 for the β and α anomers, respectively.
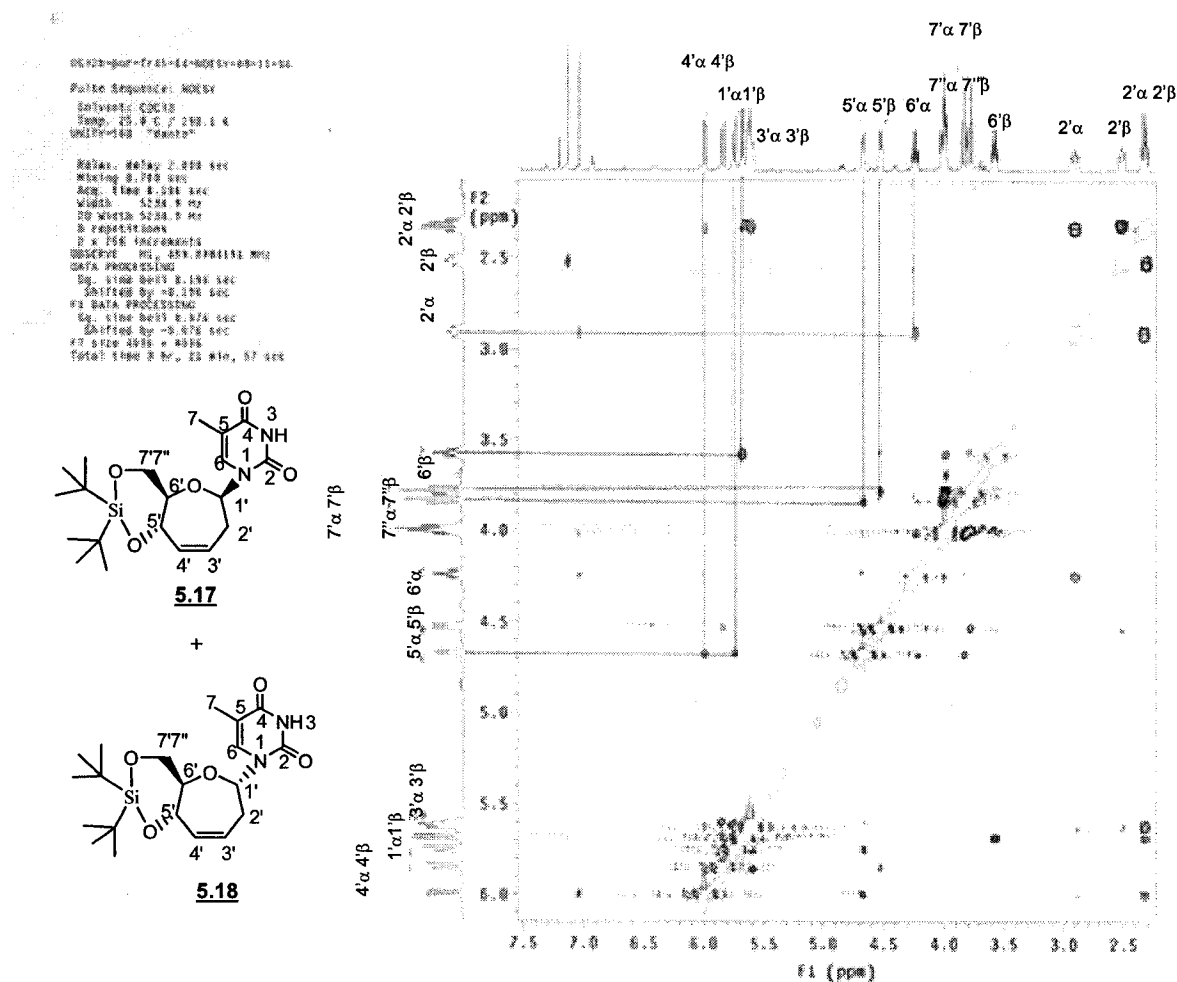
Figure 5.10: Characterization of **5.17** and **5.18** by: NOESY NMR indicating the β-diastereomer with the NOE cross-peak for C1' with the C6' protons and the α-diastereomer with the NOE cross-peak for C1' with the C5' protons of the oxepine ring.

The pseudorotational equilibrium for 7-membered carbohydrates involves many low energy transitions between the boat (B), and twist boat (TB), chair (C), and the favored

twist-chair (TC), forms.[222] The most common transitions from C to TC geometry proceed with relatively low interconversion energy barriers of 0.5-2 kcal/mol. Preliminary molecular modeling for the conformational analysis on **5.20** was performed with a starting nucleoside conformation having torsion angles C7'-OH and C5' as 30.9° and O6', C1' and C6 of the pyrimidine base as 84.6°. An *ab-initio* geometry optimization was performed *in-vacuo*, with the 3-21 G basis set using HyperChem version 7.0. The calculation was performed with the Polak-Ribiere conjugate gradient algorithm, with a root-mean-square gradient termination cut-off of 0.05 kcal/Å/mol to find the energy minimum conformation. The calculated torsion angles were measured between C7'-OH and C5' as 17.5° and between O6', C1'and C6 of the pyrimidine base as 62.8° in the anti region (*i.e.* for C2 of the pyrimidine oriented on the bottom face of the carbohydrate) for the geometry optimized oxepine nucleoside, oT*. The calculated C7' to C5' O-O distance for the hydroxyl groups was found to be 3.36 Å. The calculations provided a compact energy minimized structure where the carbohydrate was found to be in a twist-chair ($^{5',6'}TC_{O,1'}$) heptose conformation, with C1'/O lying below and C5'/C6' above the plane of the molecule defined by C2' – C3' and C4' (**Figure 5.11**).[222] The ^{1}H NMR data ($J_{1'2'}$: 10 Hz and $J_{5'6'}$: 7.5 Hz) was in full agreement with this conclusion.[223]

The nucleoside, **5.20**, was finally converted to the tritylated nucleoside phosphoramidite **5.22**, and analyzed for purity by ^{31}P (^{1}H decoupled) NMR (**5.22**: δ 146.7 and 144.9 ppm). The structure was confirmed by ^{1}H NMR and ESI mass spectrometry (*i.e.* $C_{41}H_{49}N_4O_7PNa$ calc. 763.3, found 763.2).

Structure	Conformation
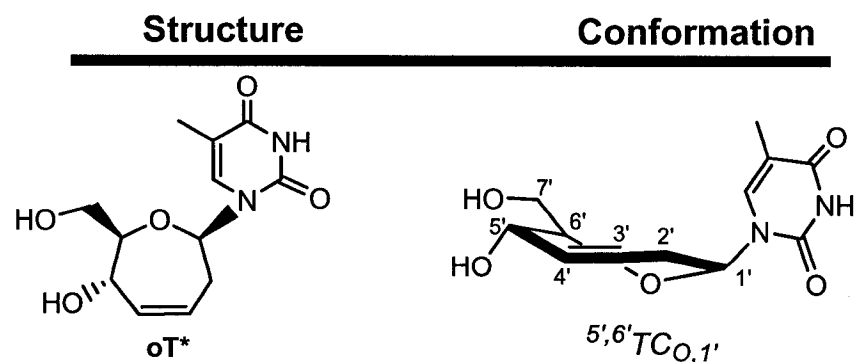
Figure 5.11: Structure and conformation for (**5.20**) oT*.

5.3.3 Solid Phase Synthesis of oxepine and ene-pyranose oligonucleotides

The unsaturated nucleoside phosphoramidites synthesized in this study, namely **5.9**, **5.14** and **5.22** were prepared as 0.15 M solutions in anhydrous MeCN for solid phase oligonucleotide syntheses. The octadecanucleotide sequences prepared are shown in **Table 5.3.1**. Nucleoside phosphoramidite **5.9** was recovered in sufficient quantities to only carry out 1 and 3 inserts within an otherwise unmodified oligothymidylate sequence. Slight modifications of the standard solid phase oligonucleotide synthesis protocol was used for the construction of the oligonucleotides on a 1 μmol scale.[24,197] The coupling (0.25 M ETT in MeCN as activator) and detritylation (3% TCA in DCM) steps were extended to 30 min and 2.5 min, respectively to ensure maximum yields (normally these steps are run for 2 min and 1.5 min, respectively). The stepwise coupling efficiency was determined by trityl color absorbance measurement (colorimetric analysis) after release of the MMT cation at the end of each coupling cycle. The average stepwise coupling efficiency was 98-99%. The oligomers were cleaved from the succinyl linked support and deprotected by treatment with 3:1 (v/v) concentrated NH_4OH:EtOH for 3-4 hours at 55°C. After evaporation, the oligothymidylate sequences were dissolved in water and quantitated by UV absorption measurements (260 nm). Yields obtained for the fully modified sequences were inconsistent with the average coupling yields observed: oT^*_{18} (7.4 ODs; 5.5%); pT_{18} (27 ODs; 20%). By contrast, the amount of material obtained in the synthesis of sequence 8, [5'-d(TTT TTT T)-p(T*T*T*)-d(TT TTT TTT)-3', **Table 5.3.1**] was 62 ODs or 43%. Furthermore, the AE HPLC analysis of the crude reaction mixtures indicated that the full length products oT^*_{18} and pT_{18} had been formed in 2.3% and 37% yield, respectively, whereas the oligothymidylate sequence with the three pT* inserts appeared to be reasonably pure (90% purity) (**Figure 5.13**). Based on this data, we concluded that the oT* modification was unstable to the ammonia treatment used to deprotect and release the oligomer from the solid support.

To investigate this further, a small amount of oligothymidylate-bound CPG solid support was treated with NH_4OH/EtOH and the amount of oligomer released was monitored as a function of time by AE HPLC (**Figure 5.12**). The analysis indicated a quick degradation for oT^*_{18} under these conditions ($t_{1/2}$ < 20 min, 20 °C). By contrast, recovery of pT_{18} decreased gradually from 33% to 12% over a 50 hour period at 55°C.

The native oligothymidylate sequence, dT_{18}, was completely resistant to the ammonia treatment as expected (50 h; 55°C). Likewise, the sequence 8 (**Table 5.3.1**) was also resistant to degradation during the extended ammonia treatment (16 h; 55°C) (**Figure 5.12**).
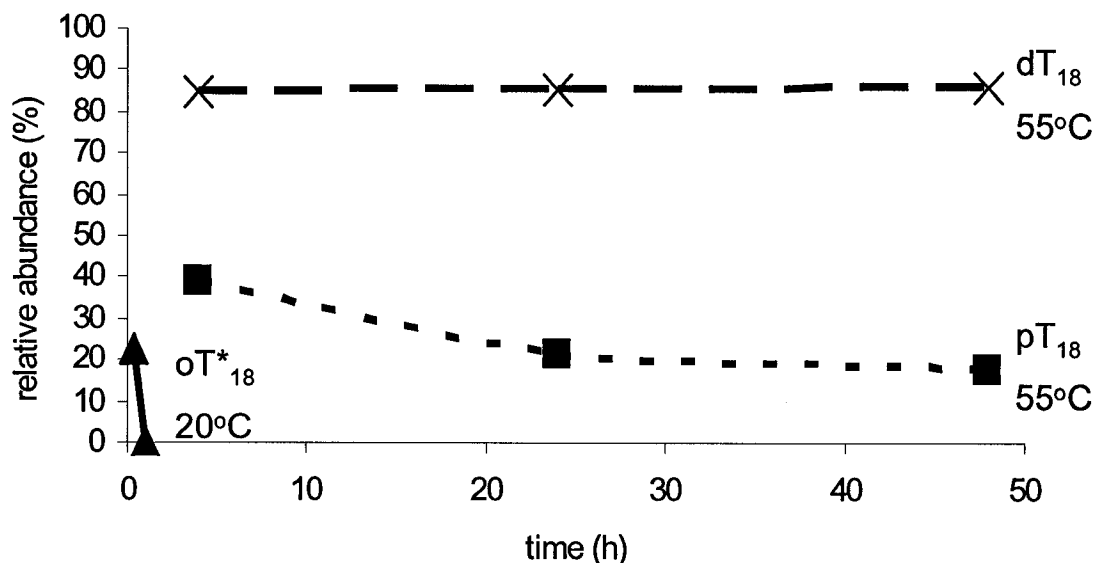
Figure 5.12: AE HPLC analyses of the stability oligonucleotides dT_{18}, pT_{18} and oT^*_{18} oligothymidylate sequences under the deprotection conditions 3:1 NH$_4$OH:EtOH at 20°C or 55°C. The oligothymidylate sequences pT_{18} and oT^*_{18} underwent degradation as indicated by the decrease in full length oligomer over time (conditions: gradient up to 30 % 1 M LiClO$_4$ in H$_2$O over 70 minute).

Since the degradation observed was persistent with the modified nucleotide inserts that contain the acidic allylic hydrogen (pT and oT*), it was rationalized that abstraction of the acidic allylic hydrogen led to an elimination reaction resulting in the concomitant release of the 5'-phosphorylated oligonucleotide.[224] This reaction would result in degradation of the oligothymidylate sequence and the probable release of the unsaturated carbohydrate diene nucleoside monomers (**Figure 5.13- A: 5.23** and **B: 5.24**). The recovered oligonucleotides products were isolated after purification by AE HPLC and desalted by size exclusion chromatography (Sephadex® G-25) and their structures confirmed by MALDI TOF MS (**Table 5.3.1**).[230]
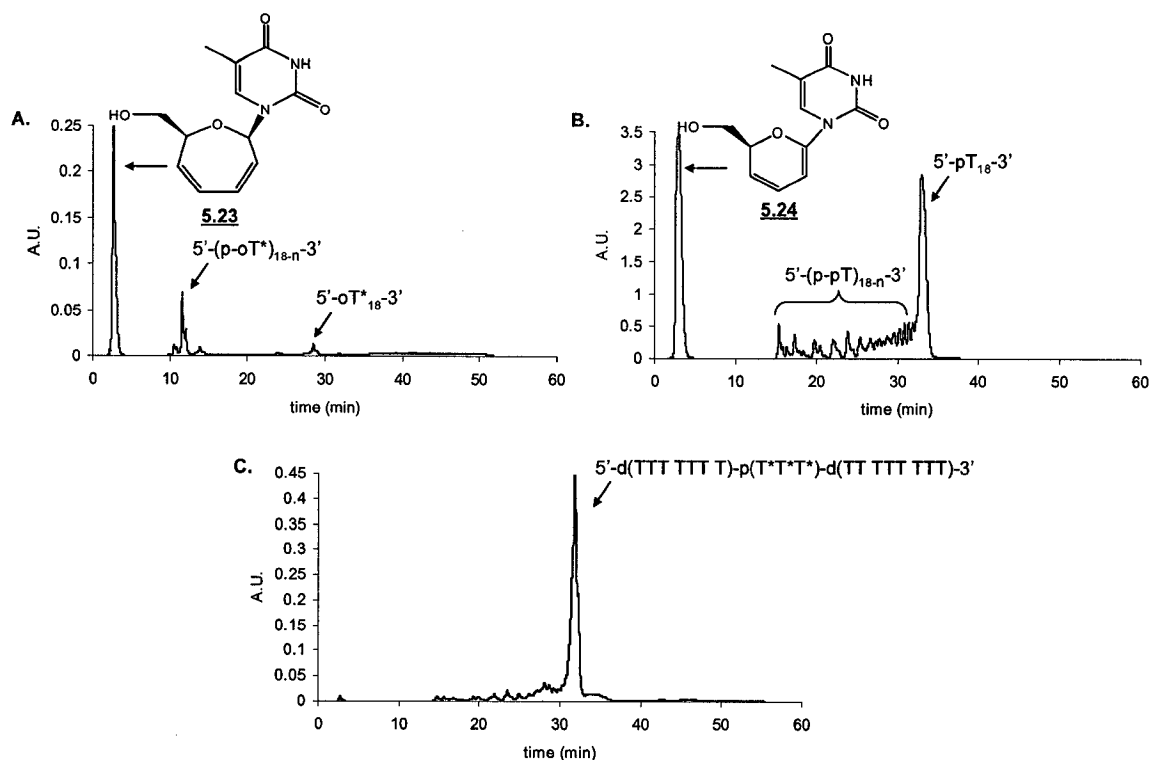
Figure 5.13: The AE HPLC analysis with an eluent gradient up to 30 % 1 M LiClO$_4$ in H$_2$O during a 70 minute run time for the crude oligothymidylate sequences: **A.** oT*$_{18}$ after 20 min treatment at 20°C with 3:1 NH$_4$OH:EtOH, **B.** pT$_{18}$ after 4 h treatment at 55°C with 3:1 NH$_4$OH:EtOH and **C.** 5'-d(TTT TTT T)-p(T*T*T*)-d(TT TTT TTT)-3'after 16 h treatment at 55°C with 3:1 NH$_4$OH:EtOH.

Sequences	Theoretical MW (g/mol)	Experimental MW (g/mol)
1. 5'-d(TTT TTT TTT)-oT*-d(TT TTT TTT)-3'	5440	5441
2. 5'-d(TTT TTT T)-o(T*T*T*)-d(TT TTT TTT)-3'	5492	n/a
3. 5'-oT*$_{18}$-3'	5882	n/a
4. 5'-d(TTT TTT TTT)-pT-d(TT TTT TTT)-3'	5426	5423
5. 5'-d(TTT TTT T)-p(TTT)-d(TTT TTT TT)-3'	5450	5486
6. 5'-pT$_{18}$-3'	5630	5635
7. 5'-d(TTT TTT TTT)-pT*-d(TT TTT TTT)-3'	5426	5432
8. 5'-d(TTT TTT T)-p(T*T*T*)-d(TT TTT TTT)-3'	5450	5450

Table 5.3.1: Oligonucleotide sequences prepared in this study, along with MALDI TOF characterization data.

5.3.4 Hybridization Properties of the Partially and Completely Modified Oligothymidylate Sequences

To determine the impact of the chemically modified thymidine units on helix stability, the MONs (**Table 5.3.1**) were hybridized to their complementary native DNA (dA_{18}) and RNA (rA_{18}) strands in a phosphate buffer (140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4, pH: 7.2). (**Table 5.3.2**). The native hybrids dT_{18}/dA_{18} and dT_{18}/rA_{18} served as controls. Duplex formation was monitored by UV spectroscopy, and the thermal data (T_m and %H) were extracted from the melting transitions observed. The data shown in **Table 5.3.2** indicated that incorporation of a single oT* unit [sequence 1: 5'-d(TTT TTT TTT)-oT*-d(TT TTT TTT)-3'] has a significant destabilizing effect on binding to both DNA and RNA targets (ΔT_m -8°C and -6°C, respectively, relative to dT_{18}). The same trend was observed for the pT and pT* insert. This decrease in duplex stability may be ascribed by the steric bulk effect[73] and carbohydrate conformations[225] of the modified 6 and 7-membered ring inserts within the dT_{18}/rA_{18}[226] or dT_{18}/dA_{18}[227] native structures.

Similarly, the completely modified pT_{18} sequence was shown to form a significantly less stable complex compared to the native dT_{18} sequence. Its hybridization with complementary dA_{18} and rA_{18} (**Figure 5.14**) indicated broad transitions with small overall hyperchromicity (% H), pT_{18}/dA_{18} (34°C and 11.3 %) and pT_{18}/rA_{18} (25°C and 15.3 %) suggesting the formation of a weak complex [compare to dT_{18}/dA_{18} (T_m : 42°C and 19.4 %) and dT_{18}/rA_{18} (T_m : 37°C and 19.1 %)] (**Figure 5.14; Table 5.3.2**).

Sequence	dA_{18}	rA_{18}	ΔT_m
1. 5'-d(TTT TTT TTT)-oT*-d(TT TTT TTT)-3'	34 (23.1)	31 (24.5)	-8 and -6
2. 5'-d(TTT TTT TTT)-pT-d(TT TTT TTT)-3'	34 (25.1)	30 (23.0)	-8 and -7
3. 5'-d(TTT TTT T)-p(TTT)-d(TTT TTT TT)-3'	26 (23.0)	22 (24.9)	
4. 5'-pT$_{18}$-3'	34[#] (11.3)	25[#] (15.3)	
5. 5'-d(TTT TTT TTT)-pT*-d(TT TTT TTT)-3'	32 (26.5)	32 (28.2)	-10 and -5
6. 5'-d(TTT TTT T)-p(T*T*T*)-d(TT TTT TTT)-3'	22 (22.4)	25 (25.3)	
7. 5'-dT$_{18}$-3'	42 (19.4)	37 (19.1)	

Table 5.3.2: Thermal denaturation data of duplexes in 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4, pH: 7.2. [#] represents broad transition in the T_m curve.
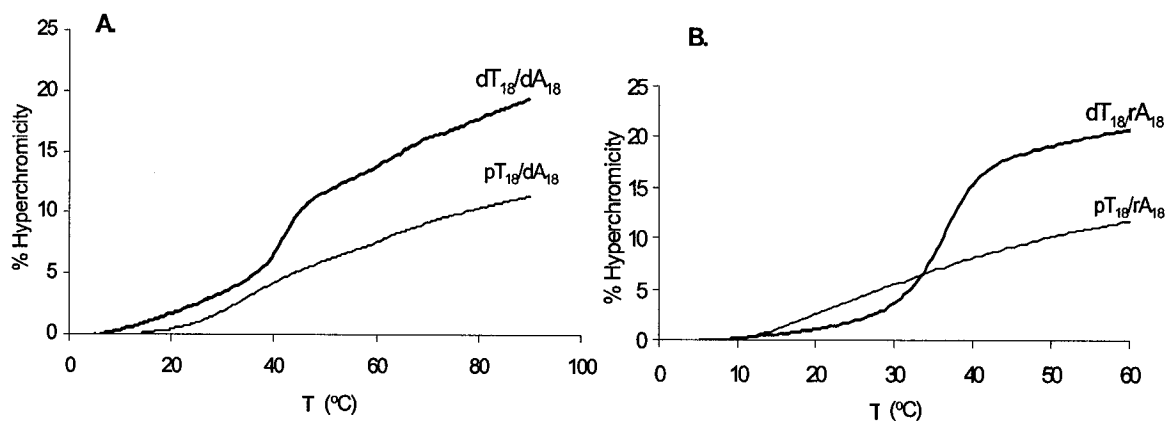
Figure 5.14: Thermal denaturation transition curves of pT_{18} (and dT_{18}) mixed with dA_{18} and rA_{18} in 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4, pH: 7.2. Hybrid concentration: 3.04 μM.

Circular Dichroism (CD) is particularly well suited for studying duplex helical structure and conformation.[173] The hybrid conformation for dT_{18}/dA_{18} is that of a B form helix which displays a CD spectrum with a positive band near 217 nm, a negative band at 247 nm and two positive peaks of low intensity at 285 and 260 nm.[228] The dT_{18}/rA_{18} hybrid conforms to an intermediate A-B form helix, with a large positive peak at 255 – 290 nm, and a negative peak at 247 nm.[226]

The CD spectra of the modified duplexes were similar to the native duplex suggesting that the global hybrid geometry is largely unaffected by the minimal number of modified units (data not shown). In contrast, hybridization of the completely modified pT_{18} with dA_{18} resulted in a CD "signature" that was shifted to longer wavelengths relative to the native dT_{18}/dA_{18} duplex (**Figure 5.15- A**). However, the spectra of pT_{18}/rA_{18} indicated a duplex structure conformation that was similar to the native dT_{18}/rA_{18}, both displaying a strong positive band at 265 nm, a strong negative band near 248 nm, and a crossover point at around 254 nm (**Figure 5.15- B**). These similarities can be ascribed to the conformation of the underlying rA_{18} strand which greatly influences the CD spectrum (**Figure 5.15- C**). In support of this notion, the CD trace of pT_{18} alone (**Figure 5.15- C**) is very different from that of the duplex, whereas the CD profile of pT_{18}/rA_{18} and rA_{18} alone are indeed very similar. Thus, it appears that pT_{18} can readily adjust to the dominant conformation of the purine strand (rA_{18} or dA_{18}) to favor hybridization. This suggests that enzymes that generally recognize DNA/RNA hybrids (such as RNase H)

136

may also recognize the pT_{18}/rA_{18} hybrid system. This hypothesis is tested in section 5.6.3.
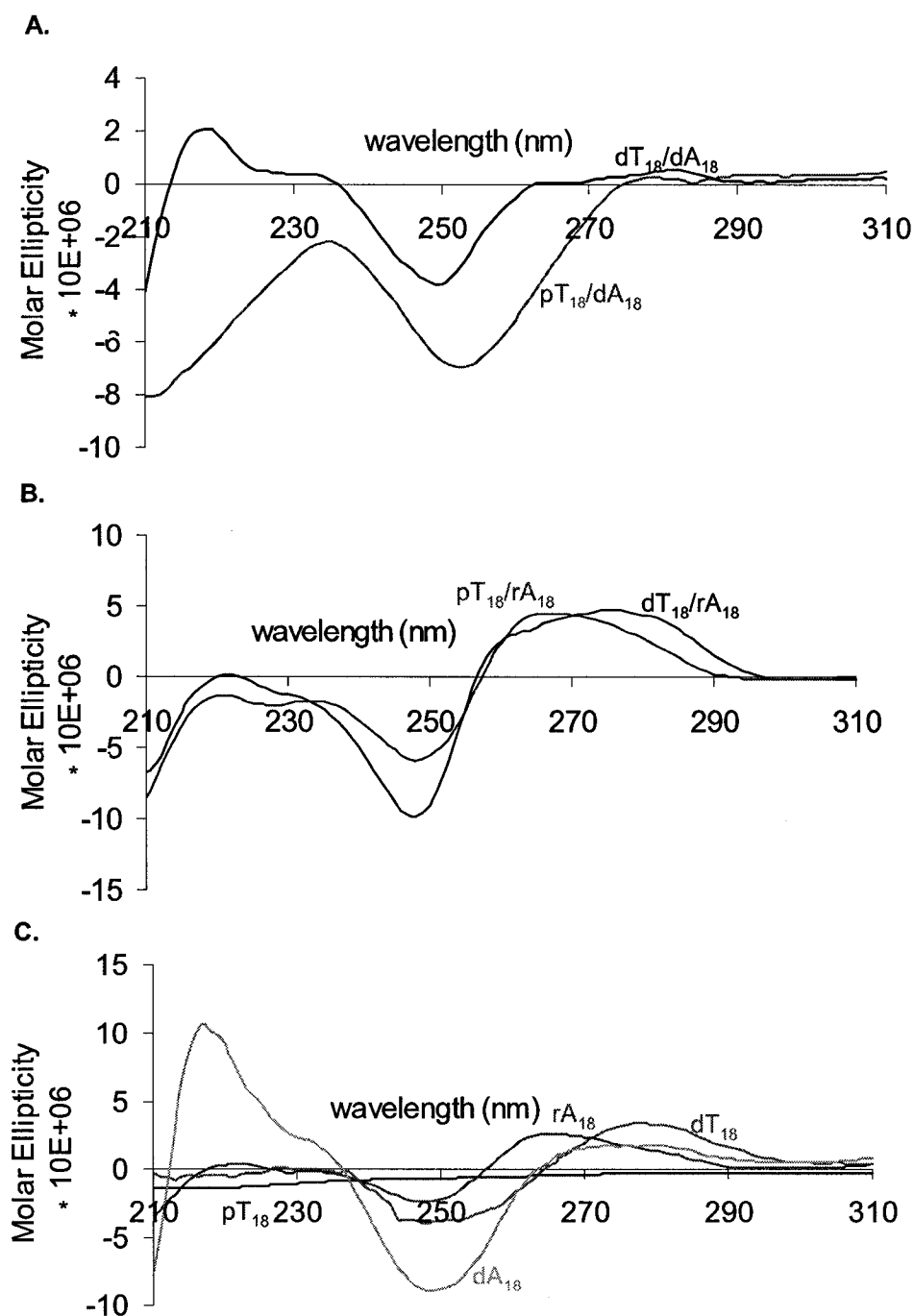
Figure 5.15: CD spectra of **A.** pT_{18}/dA_{18} and dT_{18}/dA_{18}, **B.** pT_{18}/rA_{18} and dT_{18}/rA_{18} and **C.** the single strands dT_{18}, dA_{18}, rA_{18} and pT_{18}. Hybrid concentration: 3.04 μM; buffer: 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4 adjusted to pH: 7.2.

5.4 SUMMARY AND CONCLUSIONS FROM THIS STUDY

Commercially available unsaturated glycosides such as TAG are appropriate starting reagents for the synthesis of modified 6 and 7-membered unsaturated nucleoside analogues. Incorporation of these units into oligonucleotides via the phosphoramidite methodology is possible, however, degradation of the chain was observed during the basic ammonia step used to deprotect and cleave the oligomers from the solid support. This degradation prevented the recovery of a fully modified oligonucleotide bearing a 7-membered (oxepine) carbohydrate ring. All modifications (pT, pT* and oT*), when incorporated into an otherwise unmodified oligo-dT chain were detrimental to the stability of the duplexes formed with target DNA and RNA strands. This, together with the inherent instability of the unsaturated (*e.g.*, oxepine) oligomers, prompted us to examine the saturated oxepane systems (**Figure 5.16**).

A.

oT
oxepane nucleic acids
(ONA)

B.

oA
oxepane nucleic acids
(ONA)
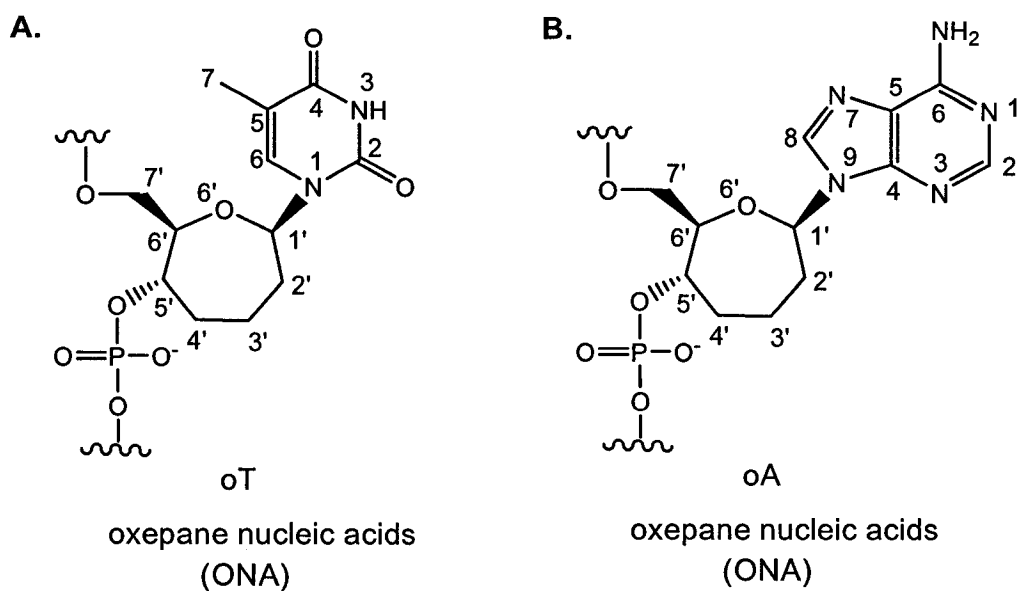
Figure 5.16: Structures of oxepane nucleotides and oligonucleotides

5.5 CHEMICAL SYNTHESIS OF NUCLEOSIDES AND OLIGONUCLEOTIDES BEARING A 7-MEMBERED (OXEPANE) SUGAR RING

5.5.1 Synthesis and Characterization of Oxepane Nucleosides and Phosphoramidites

The oxepane nucleosides, **5.28** (oT) and **5.29** (oA), were prepared by a procedure similar to the ring expansion reaction of cyclopropanated glycal sugars reported by Hoberg[215] and illustrated earlier in the synthesis of 3',4'-unsaturated oxepine nucleosides, oT* (section 5.3.2). This reaction was performed by introducing persilylated thymine and N^6-benzoyladenine as nucleophiles. These Vorbrüggen-type[229] glycosylation reactions with sugar **5.16** afforded the unsaturated oxepine nucleosides **5.25**, along with the diene byproduct, **5.19**. The chemo and diastereo-selectivity for the reaction was found to be dependent on the potency of the nucleophile. Coupling with N^6-benzoyladenine proceeded more rapidly and more efficiently than with thymine (0.5 days *vs* 1 day reflux with a chemoselectivity of 40%:30% for **5.25a** and **5.19**, and 45%:15% for **5.25b** and **5.19**) at the expense of a poorer diastereoselectivity (2:1 β:α for N^6-benzoyl adenine and 10:1 β:α for thymine), (**Table 5.5.1**). The isolated β-anomers of the oxepine nucleosides **5.25a** and **5.25b** were desilylated with 1 M TBAF in THF until TLC predicted complete reaction. The oxepine nucleosides were hydrogenated with palladium/charcoal, (Pd/C) and the oxepane nucleosides, **5.28**, (oT) and **5.29**, (oA) were tritylated and phosphitylated to the 7'-MMT 5'-phosphoroamidite diastereomers, **5.32** and **5.33** (**Figure 5.17**).

The structural assignment for the β-anomer of **5.25a** contained a ^{1}H NMR spectra with characteristic chemical shift values for the anomeric proton, 1' (δ: 5.74) in addition to the vinylic protons, 3' (δ: 5.68) and 4' (δ: 5.91) which contained a strong vinylic through bond coupling $^3J_{3',4'}$: 12 Hz (**Figure 5.18- A**). Similarly, the structural assignment for the β-anomer of **5.25b** contained a ^{1}H NMR spectra with the anomeric proton 1' at δ: 6.06 and the vinylic protons 3' (δ: 5.98) and 4' (δ: 5.90) with a coupling constant, $^3J_{3',4'}$: 12.5 Hz (**Figure 5.18- B**). The absolute stereochemistry at the anomeric position was confirmed by NOESY experiments which indicated a strong through space coupling cross-peak for the anomeric 1' and 6' protons for the β-anomers of the oxepine nucleosides **5.25a** and **5.25b**. The reduction of the oxepine double bond was monitored by ^{1}H NMR, which indicated complete conversion to the saturated oxepane nucleosides, **5.28**, (oT) and **5.29**, (oA) by the disappearance of the oxepine vinylic hydrogen protons, 3' and 4' at approximately δ: 6.0. These were finally converted to the tritylated nucleoside phosphoramidite diastereomers, **5.32** and **5.33**, and analyzed for purity by ^{31}P (^{1}H decoupled) NMR (**5.32**: δ 149.5 and 149.0 and **5.33**: δ 148.4 and 147.9) and their

structures confirmed by ¹H NMR and by molecular weight with ESI mass spectrometry (*i.e.* **5.32**: $C_{41}H_{51}N_4O_7PNa$: 765.9, found 765.2 and **5.33**: $C_{48}H_{54}N_7O_6PNa$ 878.9, found 878.3). The pure phosphoramidite diastereomers, **5.32** and **5.33**, were employed for their automated incorporation into oligonucleotides.
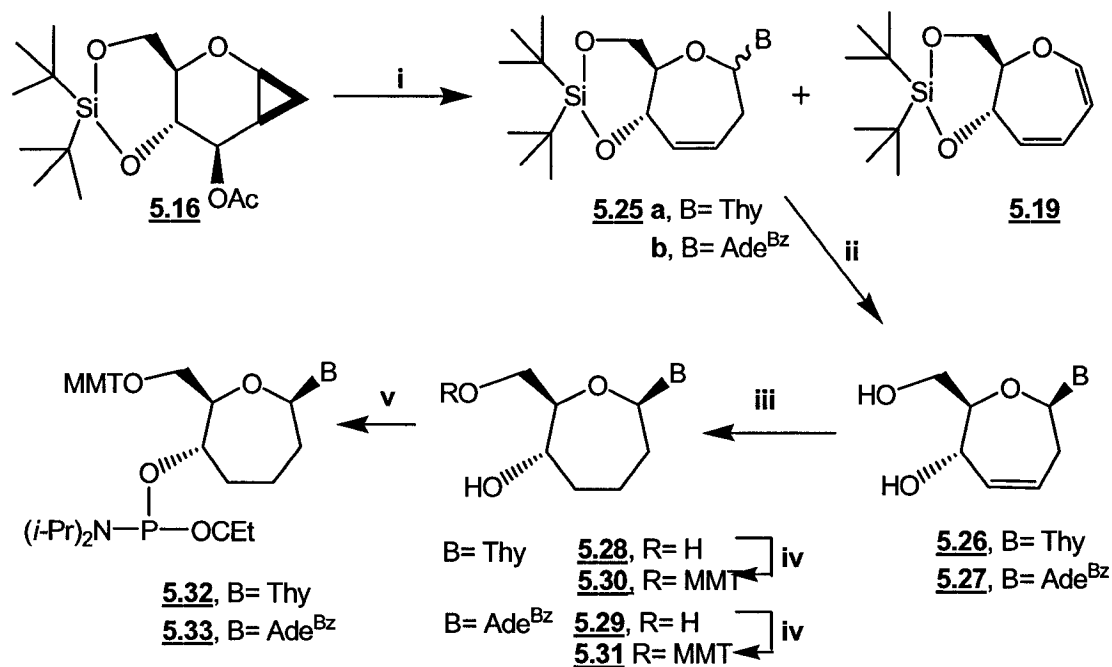
Figure 5.17: Conditions and Reagents: i. persilylated base, TMSOTf, MeCN, reflux, 12-24 h, Thy : 40% and Ade^{NBz} : 45%, ii. 1 M TBAF/THF, 0°C, 1 h, Thy : 90% and Ade^{NBz} : 61%, iii. 1 atm H₂, Pd/C, MeOH, 22°C, 4 h, Thy : 70% and Ade^{NBz} : 99% iv. MMT-Cl, pyridine, 22°C, 3 h, Thy : 66% and Ade^{NBz} : 50%, v. Cl-P(OCEt)N(*i*-Pr)₂, Et-N(*i*-Pr)₂, THF, 22°C, 2 h, Thy : 80% and Ade^{NBz} : 90%.
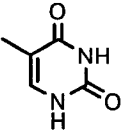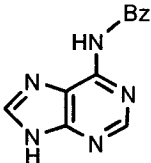
Base	Diastereoselectivity ($\beta : \alpha$)	Chemoselectivity (**5.25** : **5.19**)
T	10 : 1	1.3 : 1 40% : 30%
A	2 : 1	3 : 1 45% : 15%

Table 5.5.1: Selectivity of the glycosylation reaction of **5.16**
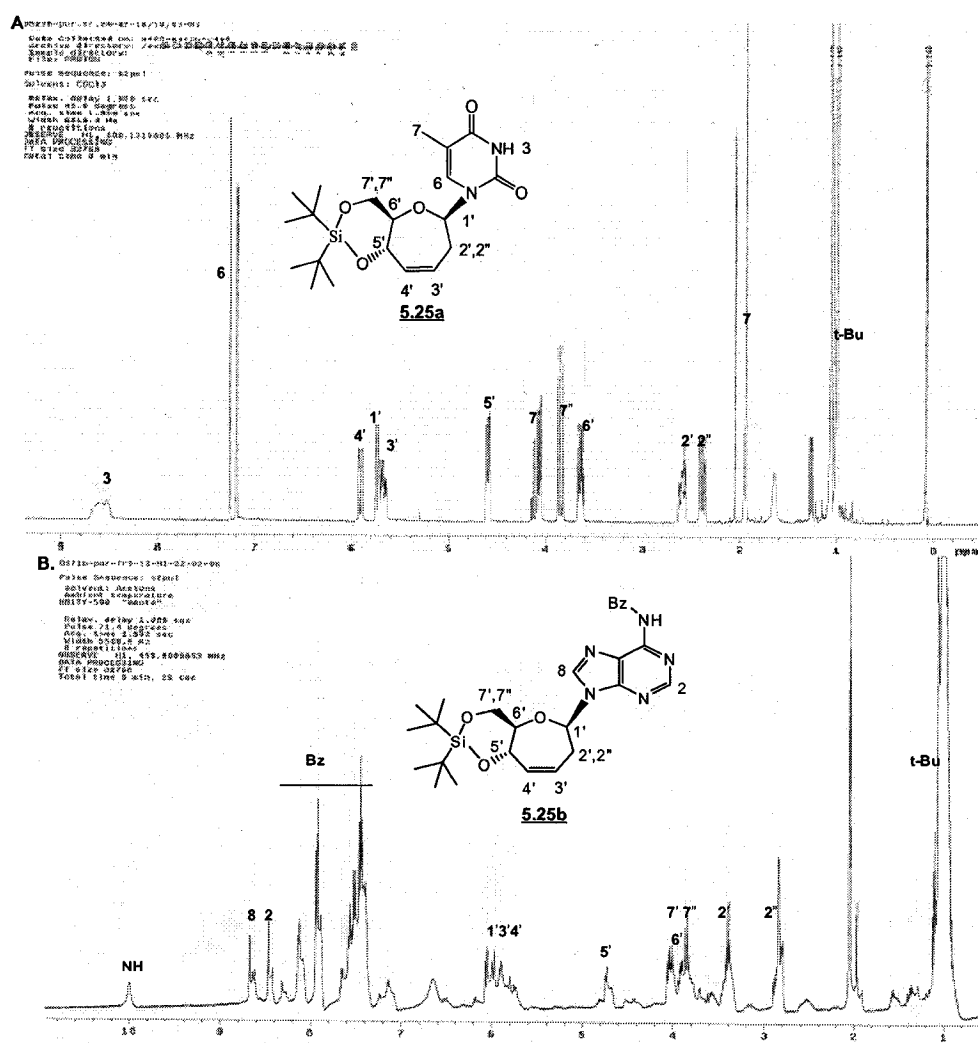
Figure 5.18: The 500 MHz ^{1}H NMR and characterization of the β-anomers for **A. 5.25a** in CDCl$_3$ and **B. 5.25b** in acetone-d_6

5.5.2 Solid Phase Synthesis and Characterization of Oxepane Oligonucleotides

The oxepane phosphoramidite diastereomers, **5.32** and **5.33** were assembled into 15-mer homopolymers on a commercially available 500 Å Unylinker[169] CPG support by the conventional automated solid phase synthesis cycle.[24,197] With a limited supply of oT and oA monomers in hand, phosphoramidites **5.32** and **5.33** were used as dilute 0.05 M solutions, in anhydrous CH_2Cl_2 for the more non-polar **5.32** (R_f: 2:1 Hex:EtOAc, 0.52) and in anhydrous MeCN for the more polar **5.33** (R_f: 2:1 Hex:EtOAc, 0.33). Oligonucleotide syntheses were conducted on a 0.5 μmol scale using extended amidite coupling times (30 min) and 5-ethylthiotetrazole, ETT, as the activator (0.25 M in MeCN) in order to ensure efficient coupling of the oxepane amidites. The detritylation step was extended to 2.5 minutes for the complete removal of the MMT groups. All other sequences were synthesized using conventional procedures.[24] Recoveries of oT_{15} and oA_{15} from the solid support were *ca.* 25 OD units, (25-40% yield). Furthermore, the desired oligonucleotides constituted 65-70% of the crude material isolated after alkaline deprotection, indicating that the monomers coupled with 98-99% stepwise efficiency and were chemically resistant to the deprotection conditions (**Figure 5.19**). Following purification by AE HPLC and/or denaturing PAGE conditions, the oligomers were desalted by size exclusion chromatography (Sephadex® G-25). The structure of the oligonucelotides was confirmed by MALDI-TOF mass spectrometry[230] (**Table 5.5.2**).

Sample	Theoretical M.W. (g/mol)	Experimental M.W. (g/mol)
1. rU_{15}	4532	4547
2. rA_{15}	4878	4931
3. dT_{15}	4503	4530
4. dA_{15}	4638	4640
5. oT_{15}	4924	4927
6. oA_{15}	5056	5080

Table 5.5.2: MALDI-TOF MS data for oligonucleotides synthesized in this study. The samples were dissolved in 1 μL H_2O, 1 μL of 6-aza-2-thiothymine/spermine and 1 μL of 50 mM aqueous L-fucose mixture prior to MS analysis.

A.

B.
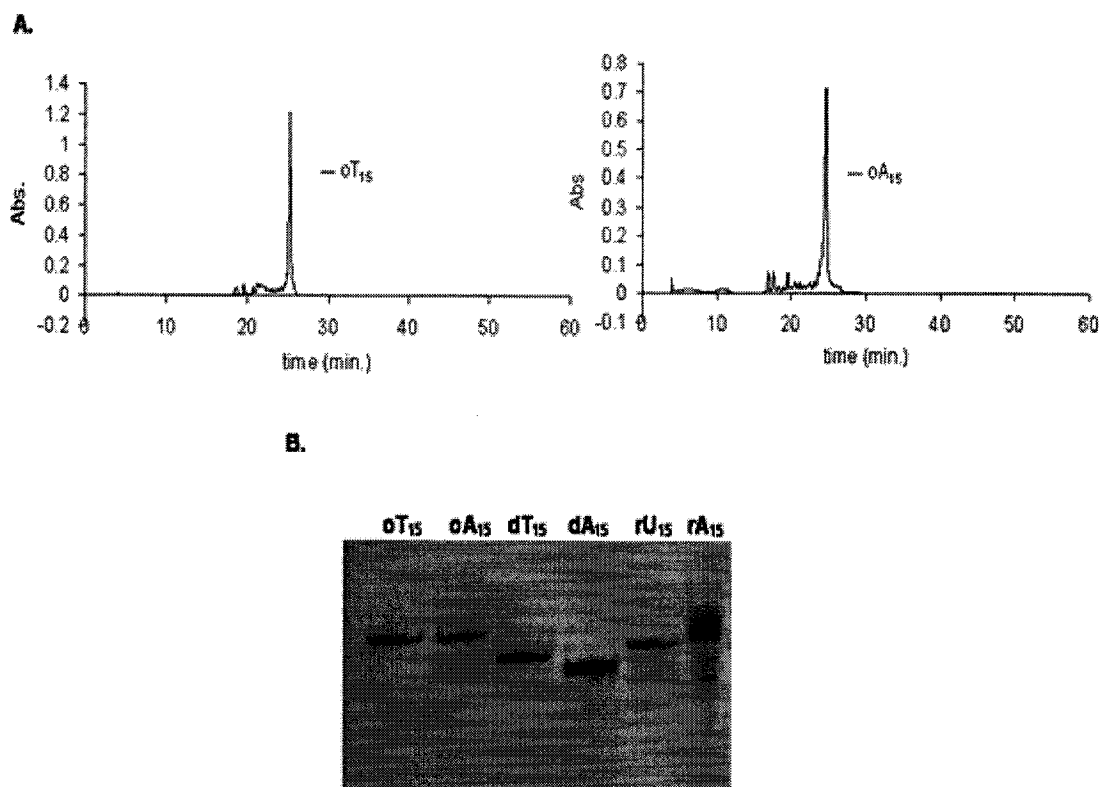
Figure 5.19: Characterization of the crude sequences by A. AE HPLC with eluent gradient up to 30 % 1 M LiClO$_4$ in H$_2$O analysis of oT$_{15}$ after 4 h treatment at 55°C with 3:1 NH$_4$OH:EtOH and oA$_{15}$ after 16 h treatment at 55°C with 3:1 NH$_4$OH:EtOH, B. Denaturing 24% PAGE analysis of homopolymeric ONA (oT$_{15}$ and oA$_{15}$), DNA (dT$_{15}$ and dA$_{15}$) and RNA (rU$_{15}$ and rA$_{15}$).

5.6 PROPERTIES OF OLIGONUCLEOTIDES BEARING A 7-MEMBERED RING OXEPANE CARBOHYDRATE

5.6.1 Hybridization and Structural Properties of Oxepane Oligonucleotides

oT$_{15}$/oA$_{15}$ Duplex. Hybridization of oligonucleotide strands was first assessed by UV-monitored thermal denaturation experiments with physiological phosphate buffers (140 mM KCl, 1 mM MgCl$_2$, 5 mM Na$_2$HPO$_4$, pH: 7.2). As anticipated from previous studies,[231] the control duplexes dT$_{15}$/dA$_{15}$, rU$_{15}$/rA$_{15}$, dT$_{15}$/rA$_{15}$, rU$_{15}$/dA$_{15}$ all exhibited a single-phase transition with 20-30% hyperchromicity (Table 5.6.1). By contrast, binding experiments on oT$_{15}$/oA$_{15}$ showed a weak but clearly detectable early transition curve (T_m : 12 °C) (Figure 5.20- B). In addition, changes in hyperchromicity, (% H), were quite

143

small, and suggesting the formation of a weaker complex relative to the native duplex dA_{15}/dT_{15} (T_m: 37 °C and % H: 25 %) (**Figure 5.20- A; Table 5.6.1**). Experiments with the individual strands alone showed that in the case of dA_{15} and oA_{15} a slight increase in absorbance with temperature (**Figure 5.20- A** and **B**), consistent with a good stacking propensity observed for the purine strands.[231] This was not the case for oT_{15}, which displayed a decrease in absorbance with temperature (**Figure 5.20- B**). Since the hyperchromicity change in the UV melting curve of oT_{15}/oA_{15} was small, more experiments were needed to evaluate whether the putative oT_{15}/oA_{15} duplex was formed. Data from UV mixing (Job Plot)[232] studies were more conclusive, strongly suggesting that oT_{15} and oA_{15} base pair into a complex with 1:1 stoichiometry at 5°C (**Figure 5.21- A**). The native dT_{15}/dA_{15} duplex also showed, as expected, UV mixing data consistent with a 1:1 stochiometry (**Figure 5.21- B**).

Sequences	T_m (°C)	% H
1. $oT_{15} + oA_{15}$	12*	4.5
2. $oT_{15} + dA_{15}$	<5	8.7
3. $oT_{15} + rA_{15}$	13*	11.3
4. $dT_{15} + oA_{15}$	<5	3.3
5. $rU_{15} + oA_{15}$	12*	3.4
6. $dT_{15} + dA_{15}$	37	24.8
7. $dT_{15} + rA_{15}$	32	25.4
8. $rU_{15} + rA_{15}$	25	28.7
9. $rU_{15} + dA_{15}$	16	20.7

Table 5.6.1: Comparison of the UV thermal melt, T_m and hyperchromicity, % H values for pairing and cross-pairing of oxepane, oT_{15} and oA_{15} oligonucleotides and the control DNA (dT_{15} and dA_{15}) and RNA (rU_{15} and rA_{15}) sequences. * Represents rough T_m due to early and/or broad transition curve.

144

A.

%Hyperchromicity

30
25
20
15
10
5
0
-5

0   20   40   60   80   100
T (°C)

dT₁₅/dA₁₅
dA₁₅
dT₁₅

B.

%Hyperchromicity

5
4
3
2
1
0
-1
-2
-3
-4

0   20   40   60   80   100
T (°C)
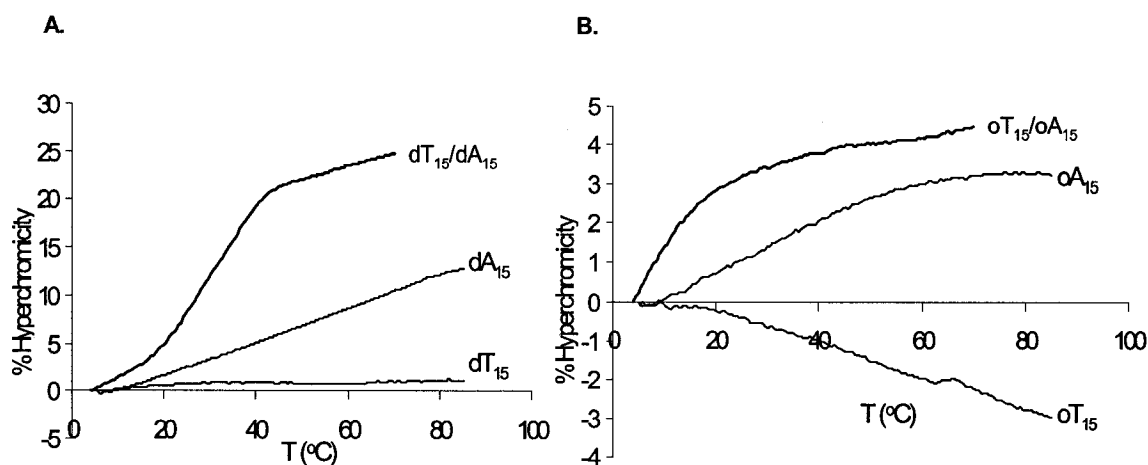
oT₁₅/oA₁₅
oA₁₅
oT₁₅

Figure 5.20: Comparison of the T_m plots for **A.** dT_{15}/dA_{15} and **B.** oT_{15}/oA_{15} in addition to their single-stranded oligonucleotides. The experiments were performed with a duplex concentration of 3.04 µM in sodium phosphate buffer: 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4 adjusted to pH: 7.2.

A.

Absorbance at 260nm

1.1
1.05
1
0.95
0.9
0.85
0.8

0   0.2   0.4   0.6   0.8   1
mol fraction oT

B.

Absorbance at 260nm

1.2
1.1
1
0.9
0.8
0.7

0   0.2   0.4   0.6   0.8   1
mol fraction T
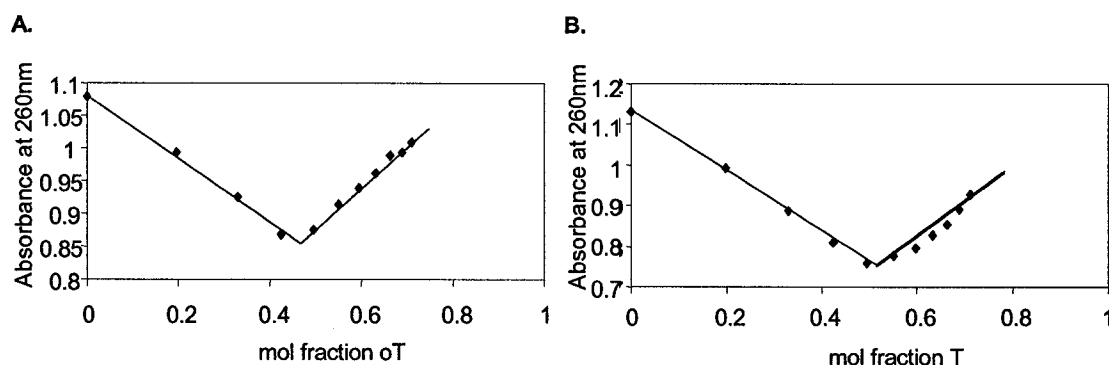
Figure 5.21: UV-mixing curves at 5°C for **A.** oT_{15}/oA_{15} and **B.** dT_{15}/dA_{15}. The experiments were performed with a concentration of each oligonucleotide stock solution 5 µM, in duplex binding buffer 10 mM Na_2HPO_4, 100 mM NaCl, pH: 7.2.

As another test for oT_{15}/oA_{15} complexation, we performed CD-monitored melting experiments (**Figure 5.22- A**). CD is particularly well suited for monitoring the melting of duplexes,[233] as single-phase transition from helix to coil state is often associated with the largest change in amplitude in the Cotton effect (molar ellipticity) at a given wavelength.[173] Samples were prepared as described above for the T_m experiments and the CD spectra normalized against a blank solution containing the phosphate buffer. A parallel experiment was conducted with the native dT_{15}/dA_{15} duplex (**Figure 5.22- B**), and the individual component strands (**Figure 5.23- B**).
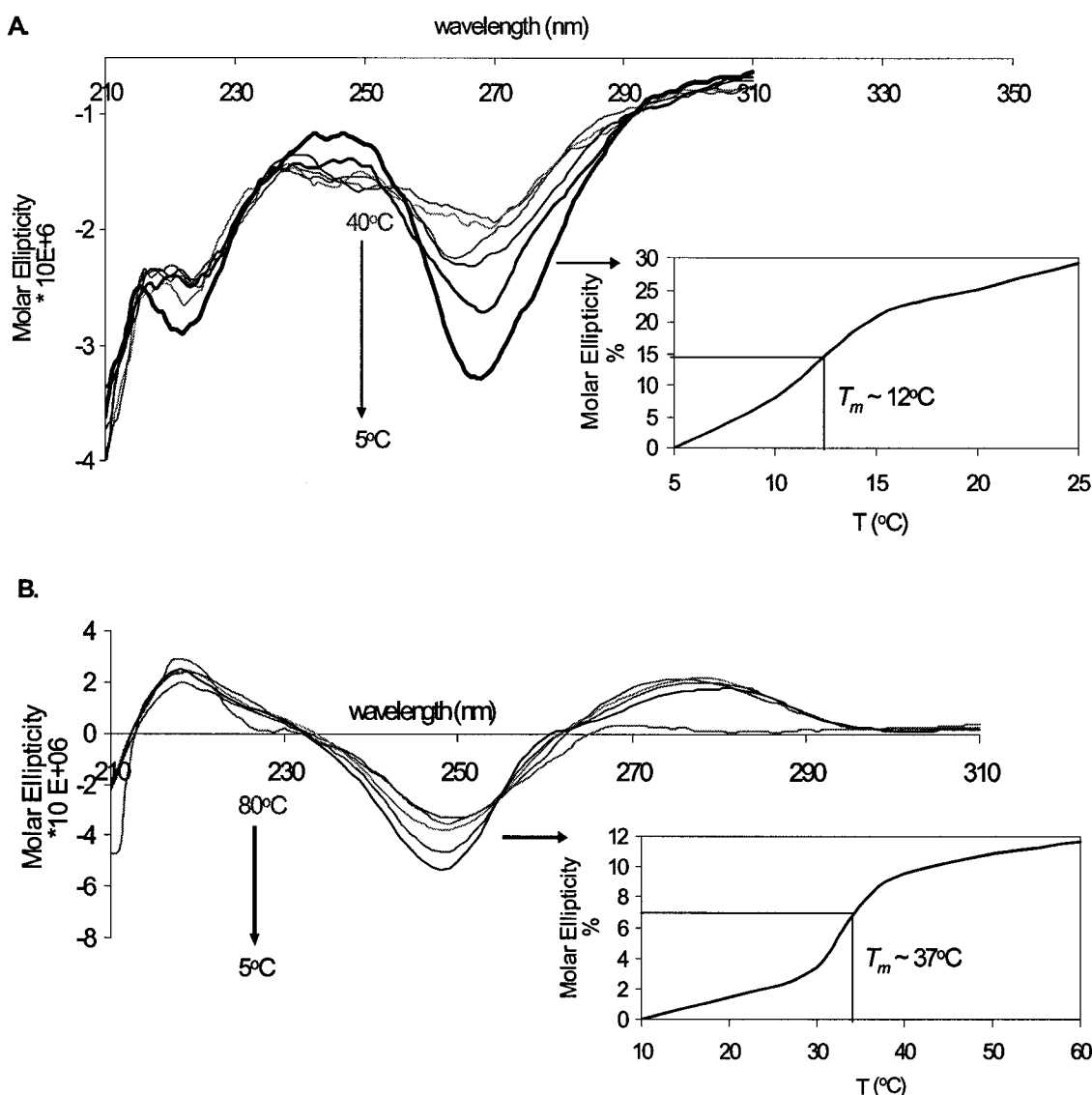
Figure 5.22: The temperature dependent CD curves and representative plots for the change in molar ellipticity with temperature for **A.** oT_{15}/oA_{15} at λ_{max} 248 nm or λ_{min} 265 nm and **B.** dT_{15}/dA_{15} at λ_{max} 278 nm or λ_{min} 248 nm. The inset displays the % change in molar ellipticity as a function of temperature at 265 or 248 nm and provides a sigmoidal thermal denaturation curve. The experiments were performed with a duplex concentration of: 3.04 μM in sodium phosphate buffer: 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4 pH: 7.2.

The data obtained provided additional evidence for the interaction between oT_{15} and oA_{15}. Specifically, the CD spectra of the mixed oligomers oT_{15} and oA_{15} showed clear differences from those of the component single strands (**Figure 5.23-** **A** and **B**), and a

plot of the change in molar ellipticity as a function of temperature at 265 nm for oT_{15}/oA_{15} produced a sigmoidal hyperchromic transition from which a T_m value of $ca.$ 12 °C was determined (**Figure 5.22- A**). This was in complete agreement with the T_m value obtained from UV melting measurements. Of note, oT_{15}/oA_{15} exhibited a strikingly different CD profile from the dT_{15}/dA_{15} or the rU_{15}/rA_{15} hybrids (**Figure 5.23- A**). The dT_{15}/dA_{15} hybrid exhibited the expected CD profile for a B-form helix, including a negative peak at 248 nm and a positive one at 278 nm.[228]

Figure 5.23 (chart A and B showing CD spectral data)
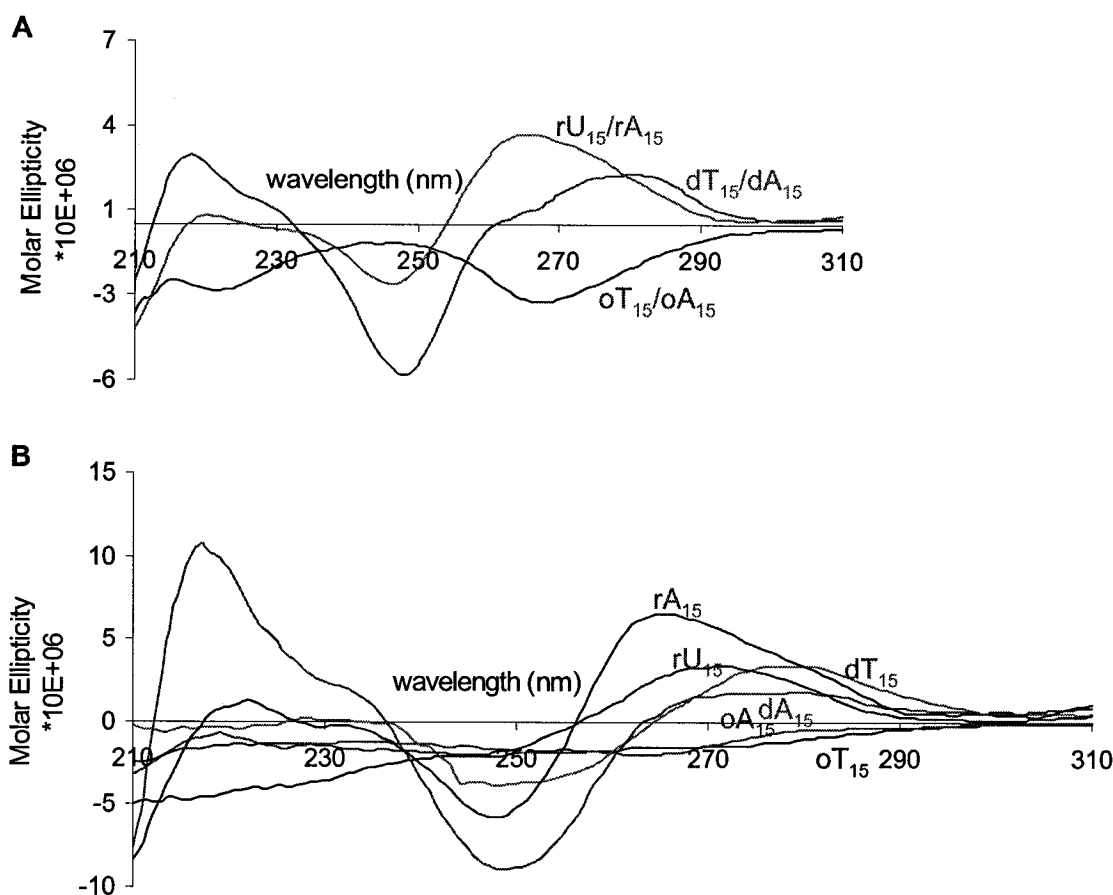
Figure 5.23: The CD spectral signatures at 5°C for single stranded and duplex oligonucleotides in **A.** oT_{15}/oA_{15}, dT_{15}/dA_{15} and rU_{15}/rA_{15} duplexes. **B.** oT_{15}, oA_{15}, dT_{15}, dA_{15}, rU_{15} and rA_{15} single strands. The experiment was performed with duplex concentration: 3.04 μM and single strand concentration: 1.52 μM in buffer, 140 mM KCl, 1 mM MgCl$_2$, 5 mM Na$_2$HPO$_4$ pH: 7.2.

Similarly, the rU_{15}/rA_{15} hybrid provided the typical CD profile for an A-form helix, including a negative peak at 248 nm and a positive broad one between 255 - 290 nm.[228]

By contrast, the oT$_{15}$/oA$_{15}$ system displayed a nearly opposite signature that was characterized by negative minima at 265 and 220 nm, and a negative maximum at 248 nm, suggesting the presence of a different helical form (**Figure 5.23- A**). While the trace observed may be suggestive of a left-handed helix,[6,234] NMR and/or X-ray crystallographic studies will be required to conclusively establish the oT$_{15}$/oA$_{15}$ duplex structure.[235]

Oxepane nucleic acid cross pairs with RNA but not single stranded DNA. The single strands, oT$_{15}$ and oA$_{15}$ were found to associate with their respective complementary RNA strands in 1:1 stoichiometric ratios (**Figure 5.24- A and B**) to form oT$_{15}$/rA$_{15}$ (T_m: 13 °C) and oA$_{15}$/rU$_{15}$ (T_m: 12 °C) hybrids, respectively (**Table 5.6.1**); by the UV absorption experiments. These values were in agreement with the T_m values estimated from the temperature dependent CD experiments (**Figure 5.25- A and B**). The corresponding native hybrids dT$_{15}$/rA$_{15}$ and dA$_{15}$/rU$_{15}$ also indicated a 1:1 stoichiometry with duplex formation (**Figure 5.24- C and D**) and had melting temperatures of 32 °C, and 16 °C, respectively (**Table 5.6.1**); by the UV absorption experiments.[231]
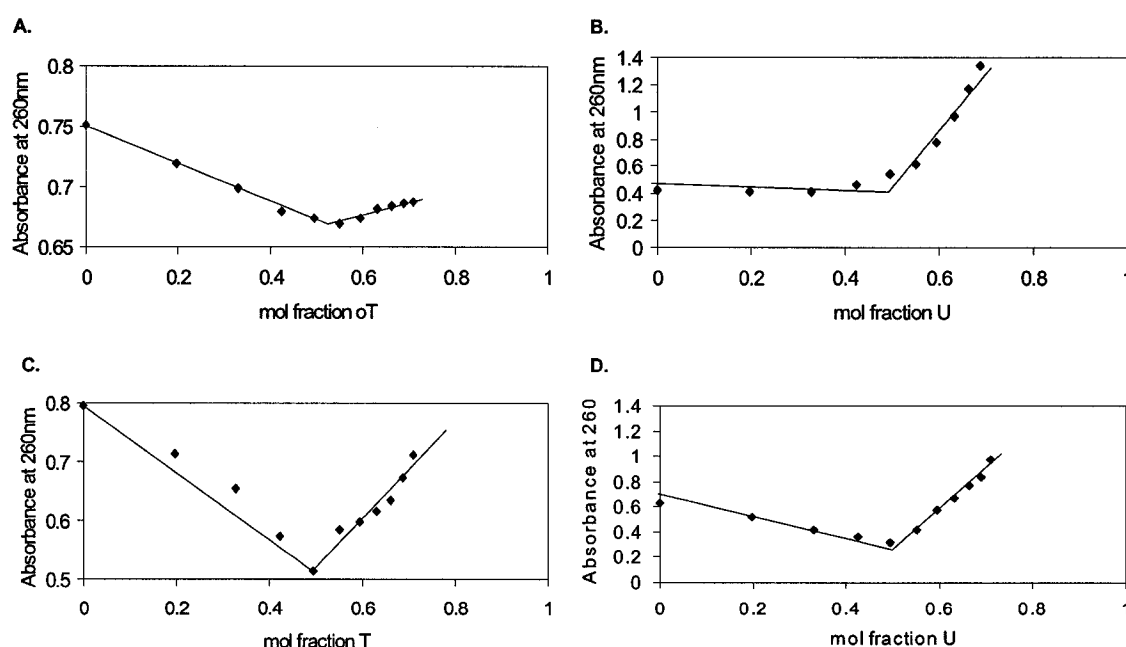
Figure 5.24: UV-mixing curves at 5°C for **A.** oT$_{15}$/rA$_{15}$ and **B.** rU$_{15}$/oA$_{15}$ **C.** dT$_{15}$/rA$_{15}$ **D.** rU$_{15}$/dA$_{15}$. The experiments were performed with a concentration of each oligonucleotide stock solution 5 µM, in duplex binding buffer 10 mM Na$_2$HPO$_4$, 100 mM NaCl, pH: 7.2.
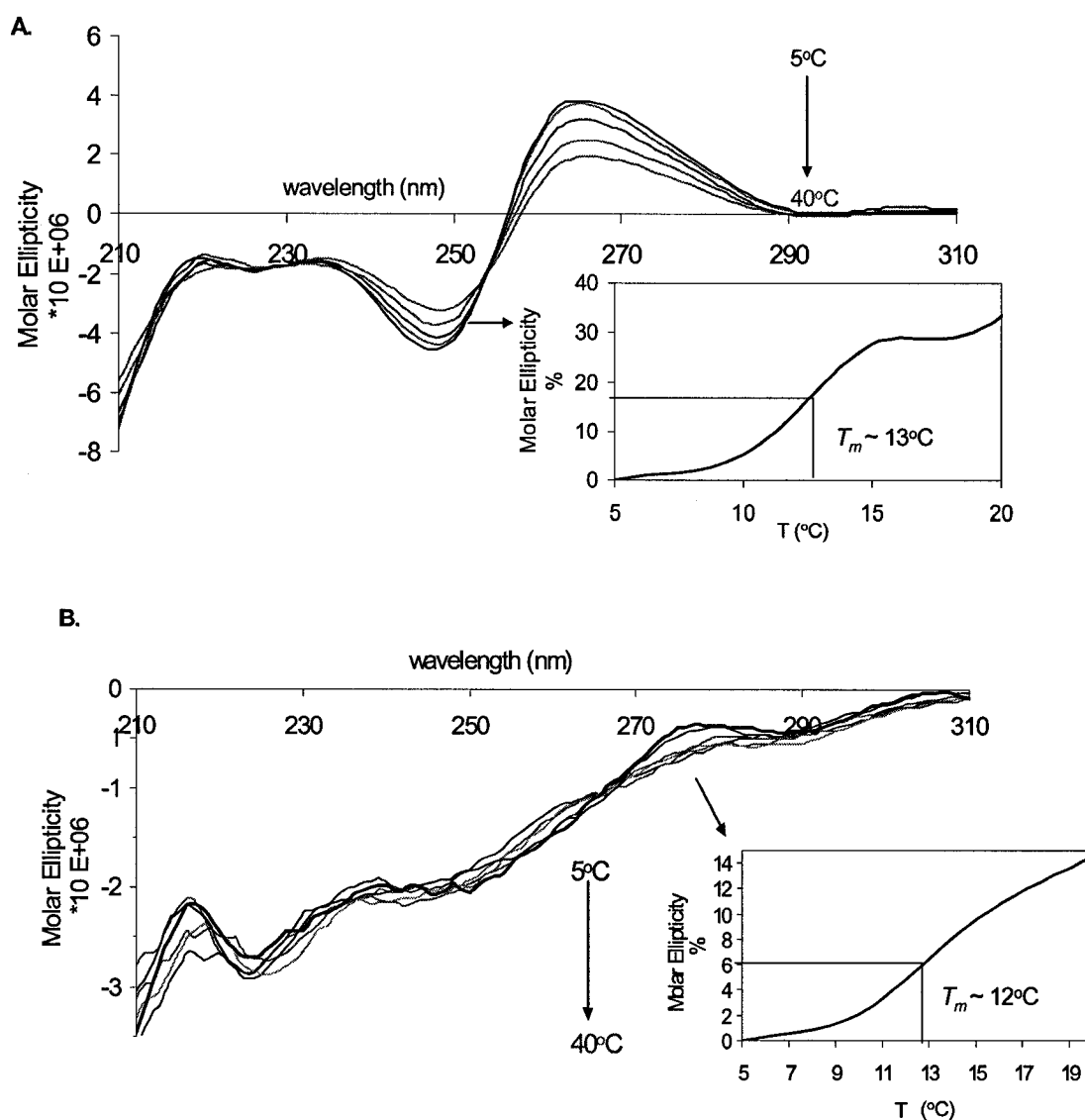
Figure 5.25: The temperature dependent CD curves and representative plots for the change in molar ellipticity with temperature for **A.** oT_{15}/rA_{15} at λ_{min} 248 nm or λ_{max} 265 nm and for **B.** rU_{15}/oA_{15} at λ_{min} 248 nm or λ_{max} 278 nm. The inset displays the % change in molar ellipticity as a function of temperature at λ_{max} 265 or 278 nm and provides a sigmoidal thermal denaturation curve. The experiments were performed with a duplex concentration: 3.04 μM in sodium phosphate buffer: 140 mM KCl, 1mM $MgCl_2$, 5mM Na_2HPO_4 pH: 7.2

No association $(T_m < 5^{\circ}C)$ was detected between the oxepane oligomers and their complementary ssDNA strands by UV (**Table 5.6.1**) and the CD-thermal denaturation experiments (**Figure 5.26-** **A** and **B**). Consistent with this notion, the CD spectrum of

oT_{15}/dA_{15} and dT_{15}/oA_{15} exhibited only a modest change in amplitude of the molar ellipticity with respect to the single strands, and the CD-monitored melting curve of the $oT_{15}+dA_{15}$ and the $dT_{15}+oA_{15}$ mixtures produced only a broad early transition of less than 5°C (**Figure 5.26-** **A** and **B**). Similarly, the UV-monitored mixing experiments (Job plots) indicated no association between the ONA and DNA single strands at 5°C (data not shown).
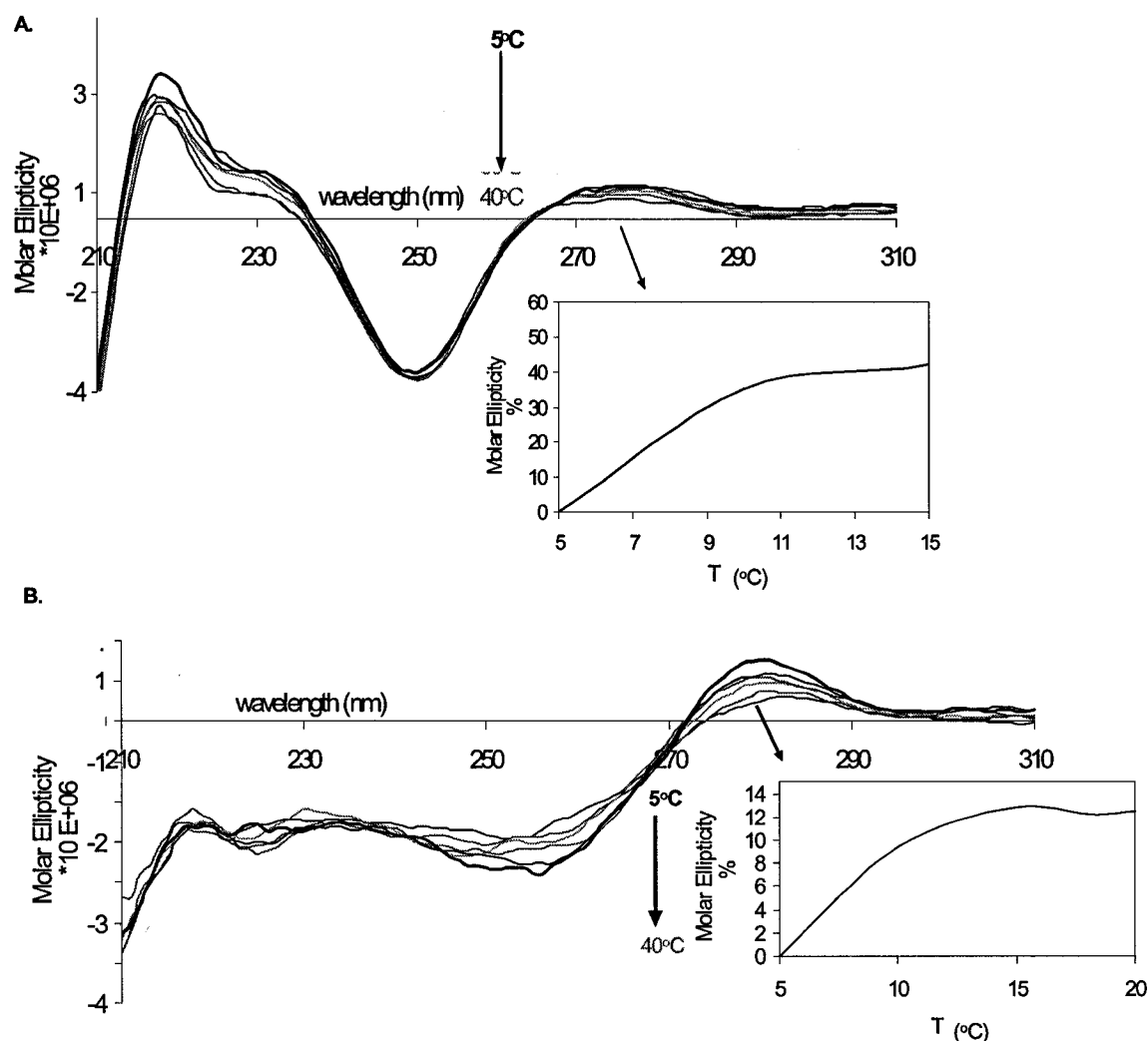
Figure 5.26: The temperature dependent CD curves and representative plots for the change in molar ellipticity with temperature for **A.** oT_{15}/dA_{15} at λ_{min} 248 nm and λ_{max} 278 nm and **B.** dT_{15}/oA_{15} at λ_{min} 255 nm and λ_{max} 278 nm. The inset displays the % change in molar ellipticity as a function of temperature at λ_{max} 278 nm and provides broad thermal denaturation curve. The experiments were performed with a duplex concentration: 3.04 μM in sodium phosphate buffer: 140 mM KCl, 1mM $MgCl_2$, 5mM Na_2HPO_4 pH: 7.2.

oT₁₅/rA₁₅ and dT₁₅/rA₁₅ form an A-like helix. The CD profiles of oT_{15}/rA_{15} and the native dT_{15}/rA_{15} systems are characteristic of A-like helices,[173,228] both displaying a strong positive band at 265 nm, a strong negative band near 248 nm, and a cross-over point at around 254 nm (**Figure 5.27- A**). These similarities can be ascribed to the conformation of the underlying rA_{15} strand which greatly influences the CD spectrum (**Figure 5.23- B**). Given that the CD trace of the oT_{15} single strand alone (**Figure 5.23-B**) is so different (a unique spectral signature with small negative bands and no cross-over points) from that of rA_{15} and the oT_{15}/rA_{15} duplex, we conclude that oT_{15} readily adjusts to the dominant conformation of the purine RNA strand (rA_{15}) during hybridization. Pairing between oT_{15} and rA_{15} provided a 1:1 complex at 5°C (**Figure 5.24- A**) whose helical conformation very closely resembled that of the dT_{15}/rA_{15} hybrid (**Figure 5.27- A**).

A.
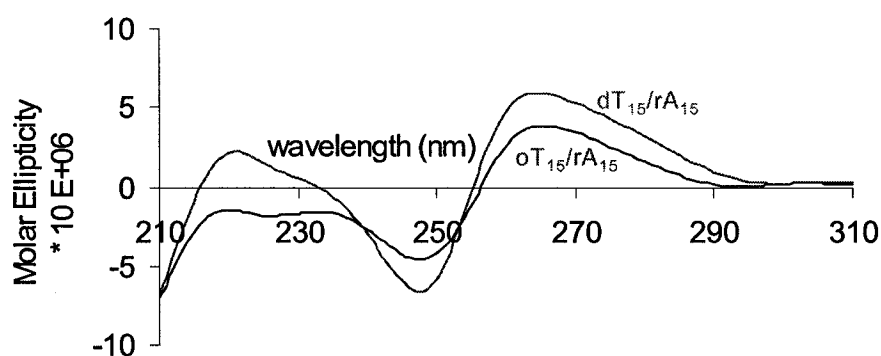
B.
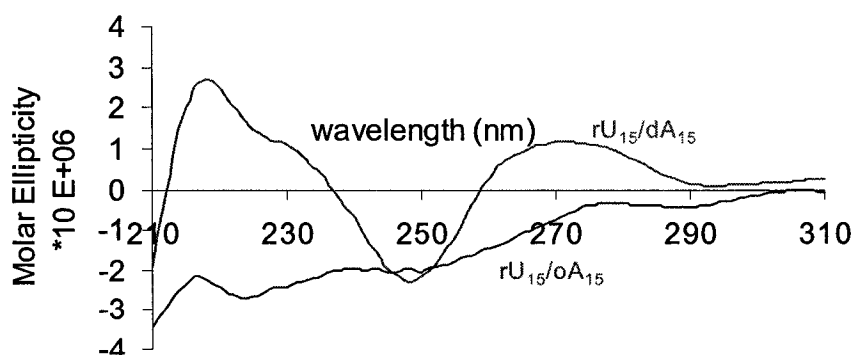
Figure 5.27: The CD spectral signatures at 5°C for duplex oligonucleotides in **A.** oT_{15}/rA_{15} and dT_{15}/rA_{15} duplexes and **B.** rU_{15}/oA_{15} and rU_{15}/dA_{15}. The experiment was performed with duplex concentration: 3.04 μM in 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4 pH: 7.2.

The CD signature for the rU_{15}/oA_{15} hybrid was less characteristic of the A-like geometry for the putative rU_{15}/dA_{15} duplex, an indication of the inherently weaker binding affinity of the two complementary sequences (**Figure 5.27- B**).[173,228] This weaker heteroduplex association has been partly attributed to the decrease in base stacking efficiency of the dA single strand (relative to rA) conferring an inherently weaker, rU_{15}/dA_{15} duplex.[231b] This has partly compromised some of the biological properties (RNaseH activity) of the rU_{15}/dA_{15} duplex.[184]

Oxepane nucleic acids do not form triplex structures. Previous studies have indicated that DNA triplexes exist *in-vivo* and can be used as chemotherapeutic agents for gene-silencing applications (see **Chapter 1.6.2** for more details).[92,94,236] However, gene suppression via triplex formation (*antigene strategy*) is a much more difficult method of inhibition (*i.e.* relative to the antisense strategy) due to their difficulty in formation and in the reduced stability of triplex structures.[91]

Hybridization of the oligonucleotide strands was assessed by UV and the CD-monitored thermal denaturation experiments with binding conditions favoring hybrid duplex formation (3.04 μM duplex containing 1.52 μM of each single strand T and A in 140 mM KCl, 1 mM $MgCl_2$, 5 mM Na_2HPO_4 pH: 7.2) and triplex formation (3.6 μM triplex containing 2.4 μM of T and 1.2 μM of A in 10 mM Na_2HPO_4, 50 mM $MgCl_2$ pH 7.3). As anticipated from previous studies[232b], the control hybrid structures dT_{15}/dA_{15} (**Figure 5.20- A** and **5.22- B**) exhibited a single-phase transition consistent with the melting of the hybrid duplex structure at $T_m \sim 37°C$ and $2(dT)_{15}/(dA)_{15}$ (**Figure 5.28- B**) indicated a two phase transition curve, characteristic of the initial disassociation of the weakly binding third complementary dT_{15} strand ($T_m \sim 12°C$) followed by the melting of the dT_{15}/dA_{15} duplex ($T_m \sim 37°C$). In contrast, binding experiments with oT_{15} and oA_{15} with the previously described binding conditions indicated, in either case (*i.e.* duplex and triplex favoring situations), a single-phase transition consistent with the melting of the hybrid duplex structure, oT_{15}/oA_{15} at $T_m \sim 12°C$ (**Figure 5.20- B** and **5.22, 5.28- A**).

As another test for oT_{15}/oA_{15} complexation, the data from the UV mixing (Job Plot)[232] studies were suggestive that the oT_{15} and oA_{15} base pair into a complex with 1:1 stoichiometry at 5°C (**Figure 5.21- A**), even with conditions favoring triplex formation

(**Figure 5.29- A**). The native dT_{15}/dA_{15} duplex also showed, as expected, UV mixing data consistent with a 1:1 stoichiometry (**Figure 5.21- B**) and $2(dT)_{15}/(dA)_{15}$ indicated the expected 2:1 stoichiometry for the putative triplex structure at 5°C (**Figure 5.29- B**). Furthermore, oxepane nucleic acids (oT_{15}) do not associate with complementary DNA and only form a 1:1 hybrid duplex structure with RNA, even with conditions favoring triplex formation (data not shown).
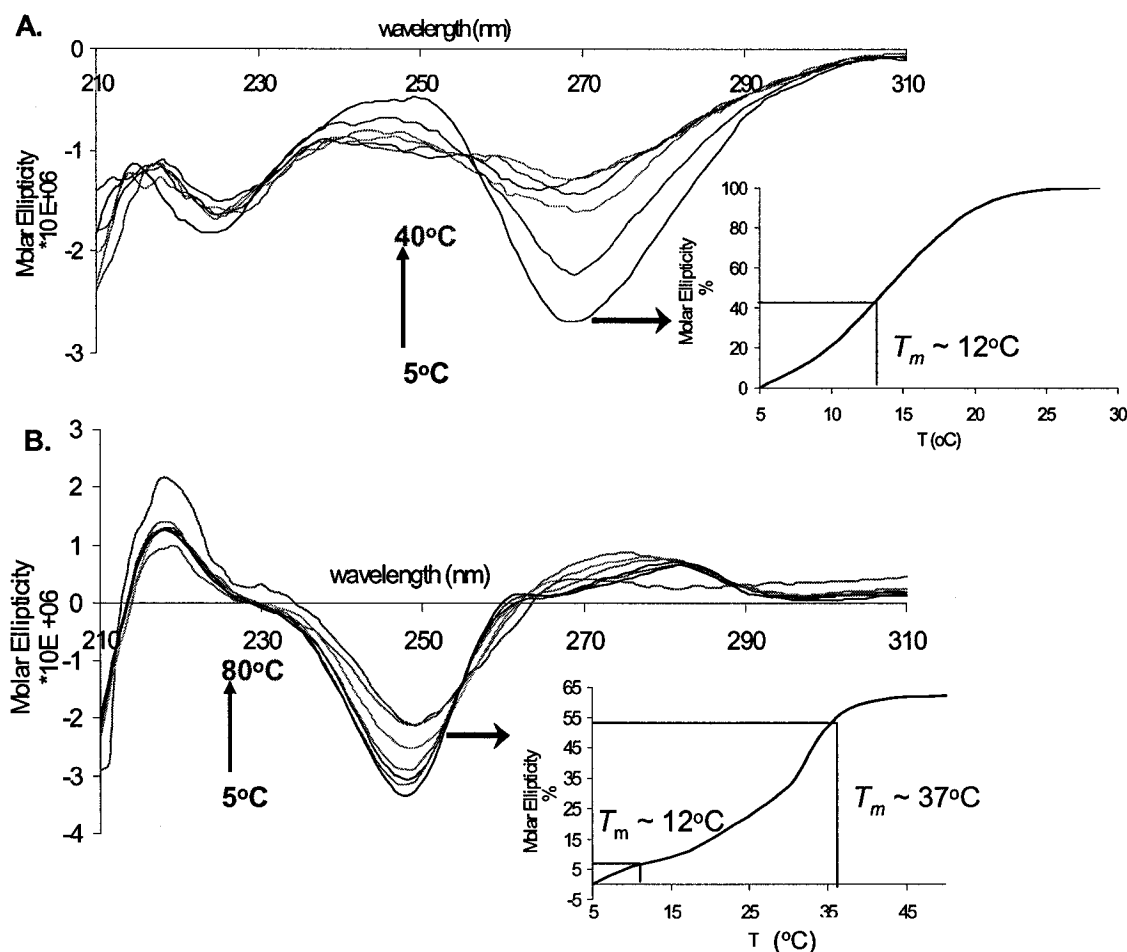
Figure 5.28: The temperature dependent CD curves and representative plots for the change in molar ellipticity with temperature for **A.** $2(oT)_{15}/(oA)_{15}$ at λ_{min} 265 nm and **B.** $2(dT)_{15}/(dA)_{15}$ at λ_{min} 248 nm. The inset displays the % change in molar ellipticity as a function of temperature for **A.** at λ_{min} 265 nm which provides a sigmoidal hybrid thermal denaturation curve for $2(oT)_{15}/(oA)_{15}$, indicative of a duplex melting transition and **B.** at λ_{min} 248 nm which provides a bi-phasic thermal denaturation curve for $2(dT)_{15}/(dA)_{15}$, indicative of a triplex melting transition. The experiments were performed with a triplex concentration of 3.6 µM containing 2.4 µM of T and 1.2 µM of A in 10 mM Na_2HPO_4, 50 mM $MgCl_2$ pH: 7.3.

A.   B.
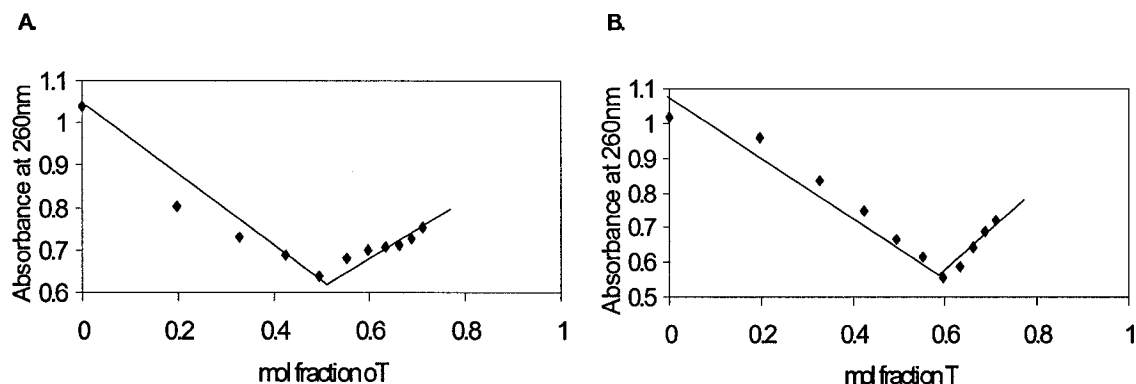
Figure 5.29: UV-mixing curves at 5°C for **A.** 2(oT)$_{15}$/(oA)$_{15}$ and **B.** 2(dT)$_{15}$/(dA)$_{15}$. The experiments were performed with a triplex concentration of 5.4 μM containing 3.6 μM of T and 1.8 μM of A in 10 mM Na$_2$HPO$_4$, 50 mM MgCl$_2$ pH 7.3

Of note, the 2(dT)$_{15}$/(dA)$_{15}$ triplex exhibited a different CD profile than the duplex B-form helix, including a negative peak at 248 nm and positive bands at 259 nm and 284 nm **(Figure 5.30- B)**.[228,231] By contrast, the 2(oT)$_{15}$/(oA)$_{15}$ system displayed a CD signature that was characteristic of the duplex form **(Figure 5.30- A)**, with a negative minima at 265 and 220 nm, and a negative maximum at 248 nm, which are suggestive of a different helical form than for A or B-type helices **(Figure 5.23- A)**.

A.   B.
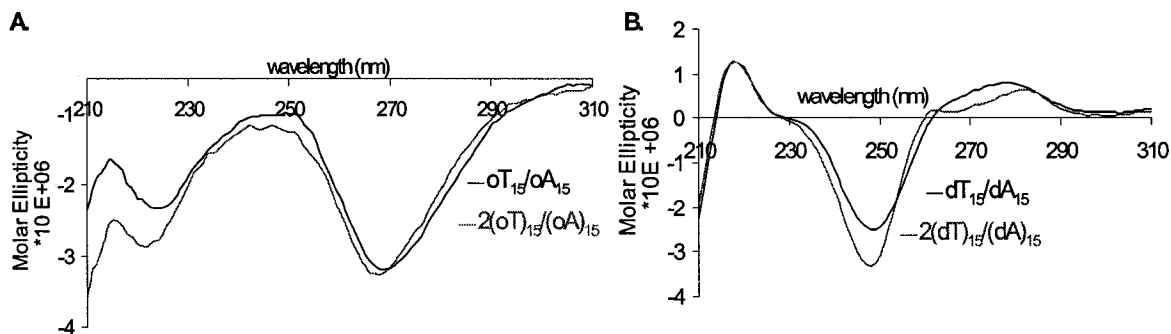
Figure 5.30: The CD spectral signatures at 5°C for hybrid oligonucleotides in **A.** oT$_{15}$/oA$_{15}$ and 2(oT)$_{15}$/(oA)$_{15}$ duplexes and **B.** dT$_{15}$/dA$_{15}$ and 2(dT)$_{15}$/(dA)$_{15}$. The experiments were performed with a triplex concentration of 3.6 μM containing 2.4 μM of T and 1.2 μM of A in 10 mM Na$_2$HPO$_4$, 50 mM MgCl$_2$ pH: 7.3.

5.6.2 Hybridization and Structural Properties of Mixed-Base DNA/DNA, DNA/RNA and RNA/RNA Duplexes Incorporating oT and oA units

To determine whether oxepane nucleotide substitutions (oT and oA) are well tolerated within DNA/DNA, DNA/RNA and RNA/RNA duplexes, these modifications were incorporated at various sites within these duplexes. Firstly, certain DNA (dT and dA) residues within an antisense strand complementary to the coding region (+1056 to +1073) of the luciferase mRNA were replaced with oT and oA.[237] Similarly, oT and oA were incorporated within the sense and antisense strand of an siRNA duplex directed at the coding regions of the firefly luciferase mRNA gene of a recombinant HeLa X1/5 cell line.[100] All oligonucleotide sequences synthesized for this study were purifed (PAGE) and their identities confirmed by MALDI-TOF MS (data not shown). The effect of oxepane (oN) modifications on the thermal stability of duplexes was compared to the known 2'-deoxy-2'-fluoro-β-D-arabinonucleotide (fN) modification extensively study in our lab.[68] Duplexes containing mismatch bases (dN) as the same positions served as controls. The data are summarized on **Tables 5.6.2** and **5.6.3**

Antisense DNA Sequence	**Target**	
	ssDNA	**RNA**
	T_m °C (ΔT_m)	T_m °C (ΔT_m)
1. 5'-ATA TCC TTG TCG TAT CCC-3'	63	67
2. 5'-ATA TCC TTG CCG TAT CCC-3'	52 (-11)	60 (-7)
3. 5'-ATA TCC AAG CCG TAT CCC-3'	43 (-6.5)	47 (-6)
4. 5'-ATG TCC AAG CCG TCT CCC-3'	35 (-5.5)	38 (-5.5)
5. 5'-ATA TCC TTG oTCG TAT CCC-3'	49 (-14)	52 (-15)
6. 5'-ATA TCC oToTG oTCG TAT CCC-3'	41 (-7)	41 (-9)
7. 5'-AToA TCC TTG TCG TAT CCC-3'	56 (-7)	60 (-7)
8. 5'-AToA TCC TTG TCG ToAT CCC-3'	47 (-8)	51 (-8)

Table 5.6.2: Comparison of the T_m and change in T_m per each modified oxepane insert (oT and oA), base mismatch (A, G, C) when incorporated into a DNA sequence. Targets were: DNA, 5'-GGG ATA CGA CAA GGA TAT-3' and RNA, 5'-GGG AUA CGA CAA GGA UAU-3'. Duplex concentration: 3.04 μM; sodium phosphate buffer: 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄, pH: 7.

155

RNA/RNA Duplexes	T_m °C (ΔT_m)
9. 5'-GCU UGA AGU CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	62
10. 5'-GCU UGA AG_A_ CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	54 (-8)
11. 5'-GCU UG**C** AG**C** CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	44 (-9)
12. 5'-GCU UGA AG**fT** CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	60 (-2)
13. 5'-GCU UGA AGU CUU UAA **fTfT**A Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	62 (0)
14. 5'-GC**fT fT**GA AGU CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	60 (-1)
15. 5'-GCU UGA AGU CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU **fT**-5'	64 (+2)
16. 5'-GCU UGA AGU CUU UAA UUA Att-3' 3'-ggCG**fA** ACU UCA GAA AU**fT** AAU U-5'	61 (-0.5)
17. 5'-GCU UGA AG**oT** CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	51 (-11)
18. 5'-GCU UGA AGU CUU UAA **oToT**A Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	60 (-1)
19. 5'-GC**oT oT**GA AGU CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU U-5'	52 (-5)
20. 5'-GCU UGA AGU CUU UAA UUA Att-3' 3'-ggCGA ACU UCA GAA AUU AAU **oT**-5'	62 (0)
21. 5'-GCU UG**oA** AG**oT** CUU UAA U**oT**A Att-3' 3'-ggCG**oA** ACU UCA GAA AU**oT** AAU U-5'	39 (-5)
22. 5'-GCU UGA AGU CUU UAA UUA Att-3' 3'-ggCG**oA** ACU UCA GAA AU**oT** AAU U-5'	52 (-5)

156

Table 5.6.3: Comparison of the T_m and (change in T_m) per each modified oxepane insert (**oT** and **oA**), modified 2'-F-ANA inserts (**fT** and **fA**) or base mismatch (**A, C**) within a an RNA duplex. Duplex concentration of 3.04 μM in sodium phosphate buffer, 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄, pH 7. The sense RNA strand is the top strand. The antisense strand that is complementary to luciferase mRNA is the bottom strand.

In all cases, the oxepane modifications destabilized DNA/DNA and DNA/RNA helix formation (**Table 5.6.2**). This de-stabilization effect was at times more significant to that observed by the incorporation of base mismatches at the same positions (compare entries 2 and 5; 3 and 6; **Table 5.6.2**), however, this effect was not additive. For example, substitution of **oT** for dT at position 10 of sequence 5 (**Table 5.6.2**) led to a significant drop in T_m for both DNA/DNA and DNA/RNA duplexes (ΔT_m -14 and -15 °C, respectively). The destabilization effect is compensated by the incorporation of two additional **oT** units upstream of the DNA strand (sequence 6, **Table 5.6.2**), providing a less destabilizing effect per oxepane substitution (ΔT_m -7 and -9 °C, respectively). However, this property is also observed for the corresponding mismatched DNA/DNA duplex (entry 3, **Table 5.6.2**), so the compensatory effect appears to be related to this particular sequence. The data also indicated that oxepane modifications are better tolerated in RNA/RNA duplexes (ΔT_m -1 to -5 °C), provided that they are not introduced in the center of the helix, where destabilization in this particular duplex was significant (ΔT_m -11 °C) (**Table 5.6.3**). The poorer binding affinity of DNA and RNA strands containing oxepane units may be due, at least in part, to local steric disruptions and/or missalignment of the sugar-phosphate backbone at the site of modification (section 5.2) within a highly organized duplex structure.

157
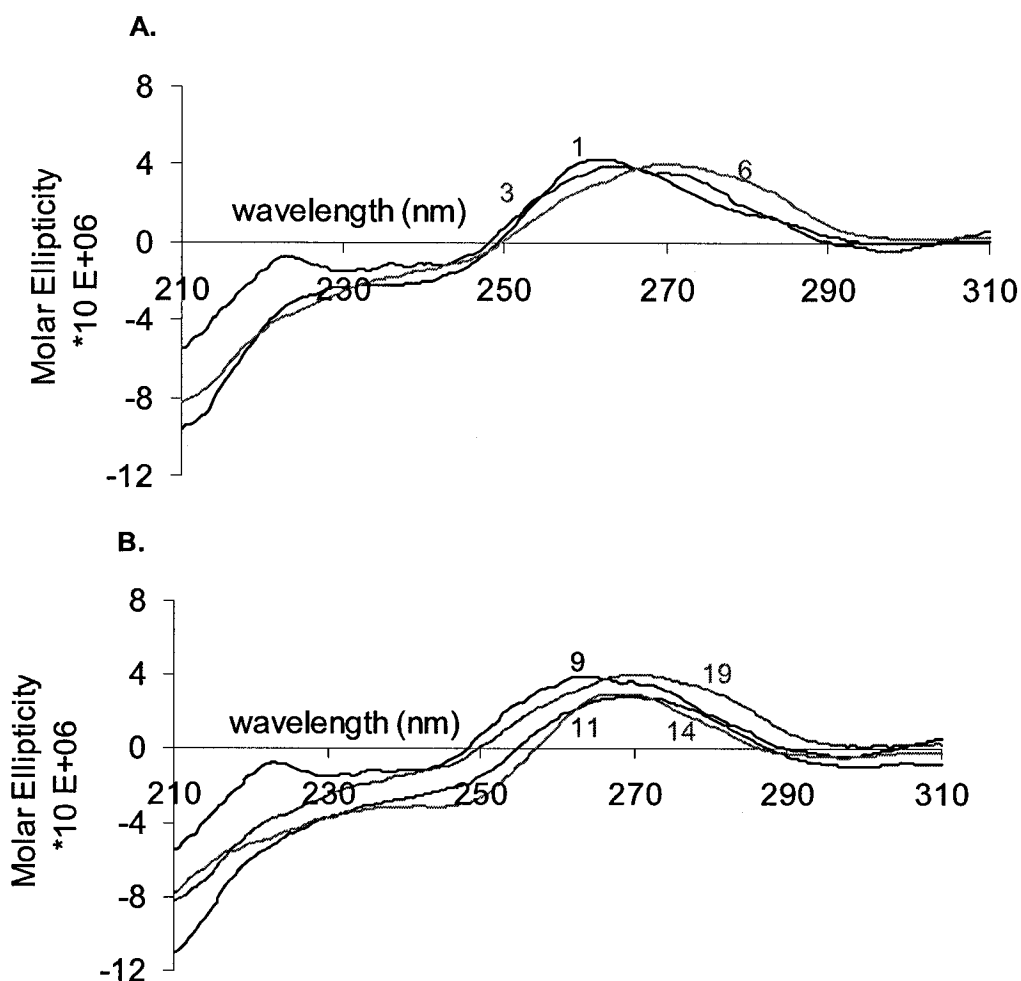
Figure 5.31: The CD spectral signatures at 10°C for duplex oligonucleotides in **A.** Modified DNA sequences 1, 3, 6 hybridized to the DNA target (**Table 5.6.2**), and **B.** RNA duplexes 9, 11, 14 and 19 (**Table 5.6.3**). The experiments were performed with a duplex concentration of 3.04 μM in sodium phosphate buffer: 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄, adjusted to pH: 7.

The de-stabilization effect created by oN was also apparent from the CD structural studies (**Figure 5.31**). While the signatures of the oN, fN and mismatched modified duplexes all conform to the A-like helical conformation, introduction of the oN units led to a red shift and broadening of the positive peak centered at around 270 nm. Given that the number of oN units in these 21-bp duplexes are small, such change is in fact significant.

5.6.3 Biological Properties of Oligonucleotides

Recognition and Cleavage of dT_{18}/rA_{18} and pT_{18}/rA_{18} by E.coli RNaseH. The insertion of a double bond in cyclohexane nucleic acids (*i.e.* cyclohexene NAs) is necessary for cross-pairing with RNA and activation of the RNaseH enzyme.[77] To test whether the same effect operates in a pyranose system, we prepared a variant of the homo-DNA structure, the 2'-enopyranosyl modification (pT_{18}) previously reported by Felder *et al.*[192] The pT modification was found to be of particular interest as an antisense construct because previous studies had indicated that the pT oligomer was nuclease resistant[211,212] and adopted a carbohydrate conformation that would, in principle, favor pairing with RNA.[192] The latter property was confirmed by the hybridization (T_m) and structural (CD) studies described in section 5.3.4 (**Figure 5.14- B and 5.15- B**) which indicated that pT_{18} hybridized to its target RNA (rA_{18}) to afford a duplex with a DNA/RNA hybrid-like conformation. This prompted us to evaluate the susceptibility of the former hybrid to cleavage by RNase H.

The assay was carried out with pT_{18}/rA_{18} and the control heteroduplexes, dT_{18}/rA_{18} and $(rU_{2'-OMe})_{18}/rA_{18}$ which activate and eliminate RNaseH activity, respectively. The catalytic efficiency of E. *coli* RNase H was determined at 37°C and lower temperatures (10 and 20 °C), since the T_m of pT_{18}/rA_{18} was about 25°C (**Figure 5.14**). This ensures that a sufficient population of the modified duplex exists for probing enzymatic cleavage. In all cases, the rate of cleavage of dT_{18}/rA_{18} was the greatest, *i.e.*, $dT_{18}/rA_{18} > pT_{18}/rA_{18} >> (rU_{2'-OMe})_{18}/rA_{18}$ (no cleavage) (**Figure 5.32**). Maximum cleavage of pT_{18}/rA_{18} occurred at 20°C (25 % cleavage after 2 hours). At 10°C and 37°C the extent of cleavage was 10 % and 15 %, respectively (2 hours).

These results support the notion that effective antisense MONs contains comparable hybrid stability and conformational properties relative to the native DNA/RNA hybrid for efficient RNase H catalysis.[238]
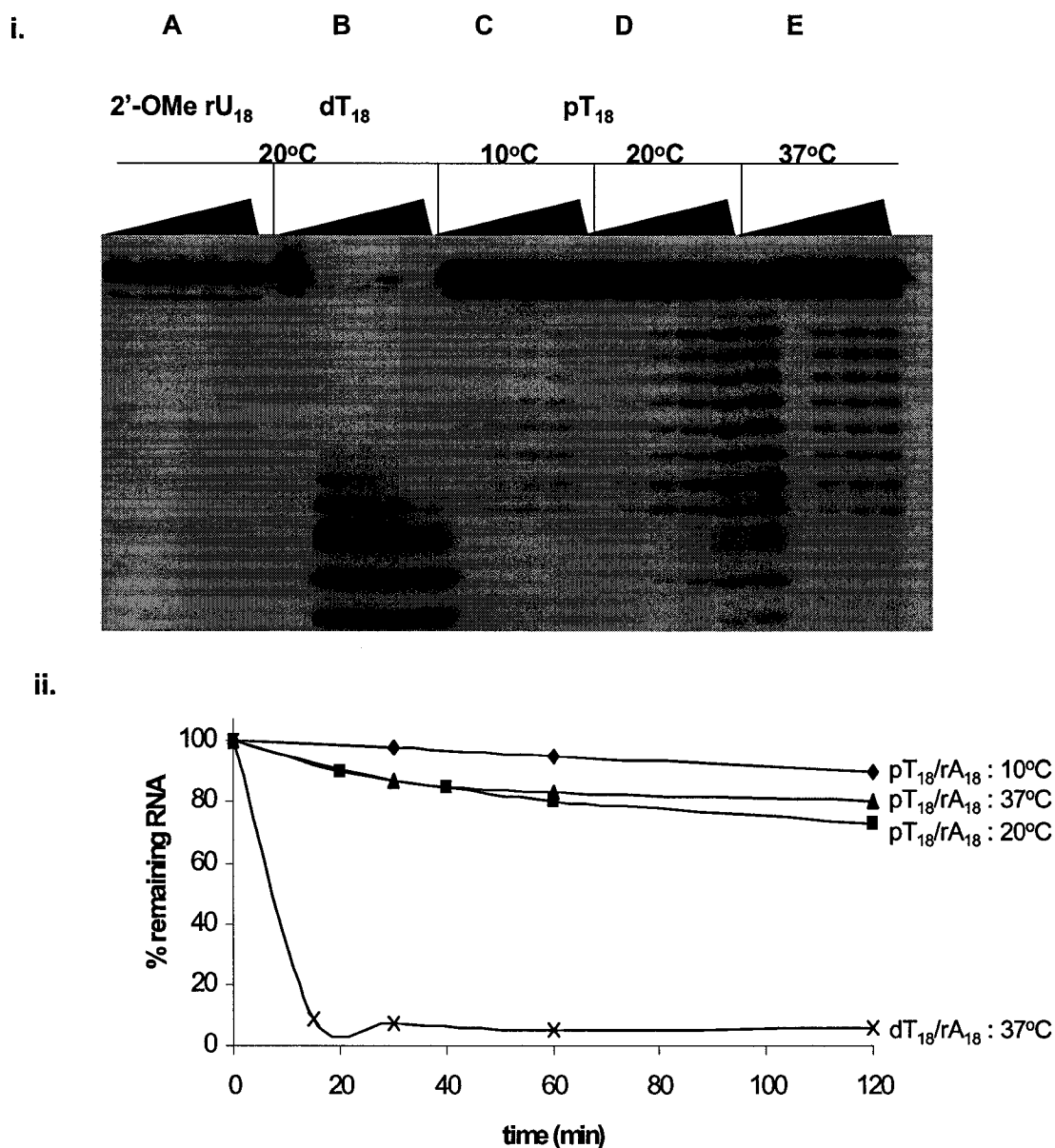
Figure 5.32: Susceptibility of 2'-OMe rU_{18}/rA_{18}, dT_{18}/rA_{18}, pT_{18}/rA_{18}, and pT_{18}/rA_{18} to *E. coli* RNase H hydrolysis.

E. coli **RNase H-mediated degradation of oT_{15}/rA_{15}.** The RNase H family comprises a class of enzymes that recognize and cleave the RNA strand of RNA/DNA heteroduplexes having topologies that are intermediate between pure A- or B-form helices adopted by dsRNA and dsDNA, respectively.[53,54,239] Among the expanding list of sugar-modified ONs that elicit RNase H-mediated cleavage of RNA are arabino-

(ANA)[61,62] and 2'-deoxy-2'-fluoro-β-D-arabinonucleic acids (2'-F-ANA),[64,237] cyclohexene nucleic acids (CeNA)[79] and α-L-*ribo*-configured locked nucleic acids (α-L-LNA).[240] Chemical changes of the furanose or alterations in the orientation of the furanose to the base can dramatically diminish RNase H activation.[241] The discovery that 2'-*arabino*-configured locked bicyclonucleotide, [3.3.0]bc-ANA, satisfies the conformational requirements for RNase H recognition and cleavage without activating this enzyme suggests that the flexibility of the ON strand is critical as well.[242] A rigid ON seems to be unable to distort to allow its RNA complement to assume the conformation necessary for hydrolysis.[68,243]

The study of other carbohydrate systems, including non-furanose sugars, allows us to extend our understanding of the impact of backbone flexibility on RNase H activity.[60b,238] The oxepane backbone provides a 7-membered ring structure that is rendered more flexible than homo-DNA[73,75] and [3.3.0]bc-ANA,[242] and conceivably more likely to engage in interactions (*i.e.* H-bonding and base-stacking) with the opposing RNA strand of the formed ON/RNA heteroduplex. This is demonstrated in **Figure 5.33**, which shows that oT$_{15}$/rA$_{15}$ supports detectable cleavage by *E. coli* RNase H. As previously observed for the pT$_{18}$/rA$_{18}$ hybrid (**Figure 5.32**), less hydrolysis occurs when compared with the native substrate which can be partly rationalized by the lower affinity of oT$_{15}$ for the RNA target ($T_m \sim 13°C$). This property acts to present a lower effective concentration of substrate duplex to the enzyme, thereby diminishing the overall rate of catalysis. Consistent with this notion, RNase H activity for the oT$_{15}$/rA$_{15}$ duplex was essentially lost at the optimal enzyme temperature of 37°C (**Figure 5.33**).
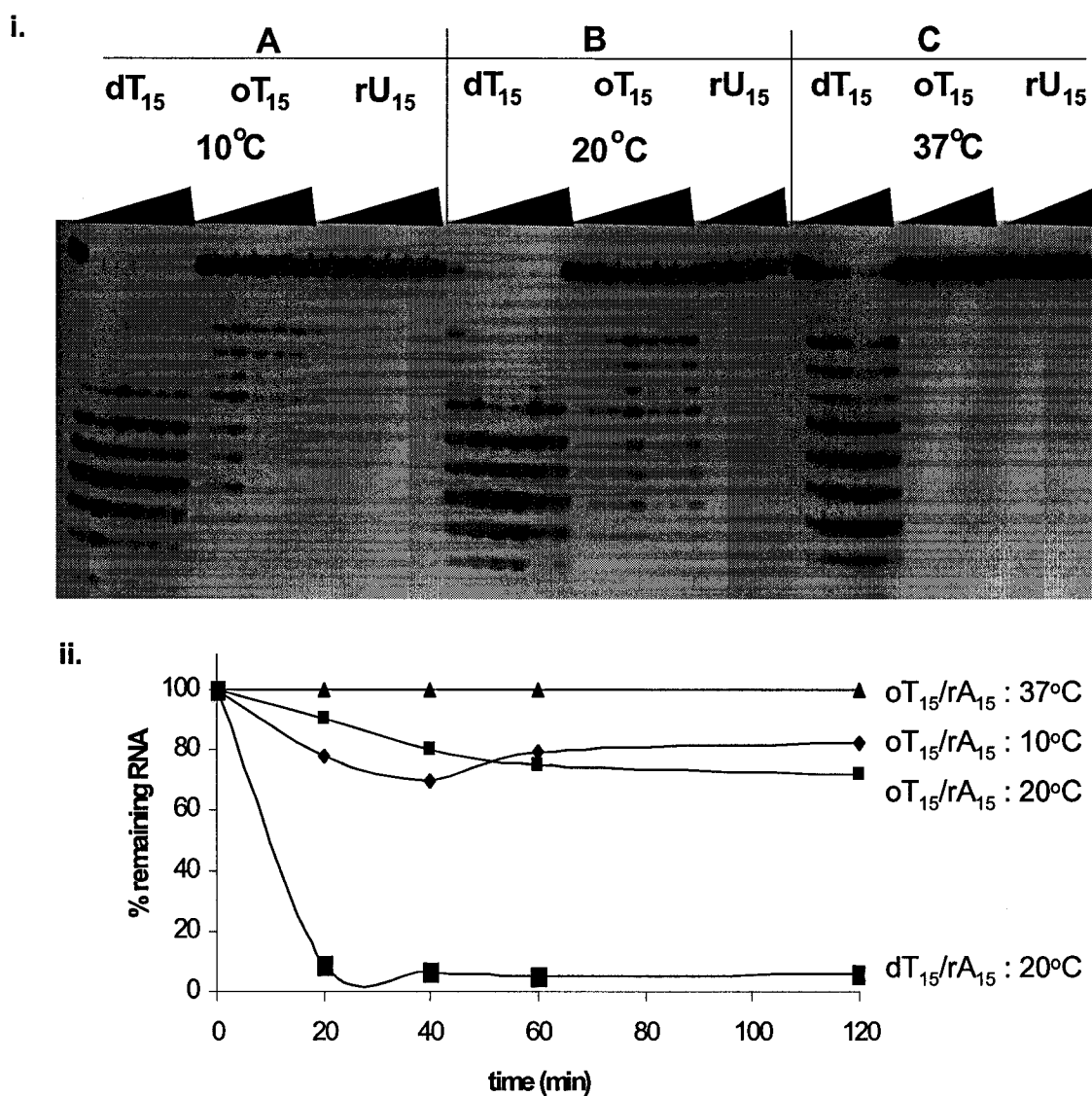
Figure 5.33: i. *E. coli* RNase H-mediated degradation of the RNA strand of dT_{15}/rA_{15}, oT_{15}/rA_{15}, and rU_{15}/rA_{15} duplexes at various temperatures. ii. Rate of the hydrolysis of oT_{15}/rA_{15} at various temperatures. Comparison to the rate of cleavage at dT_{15}/rA_{15} at 20°C is provided.
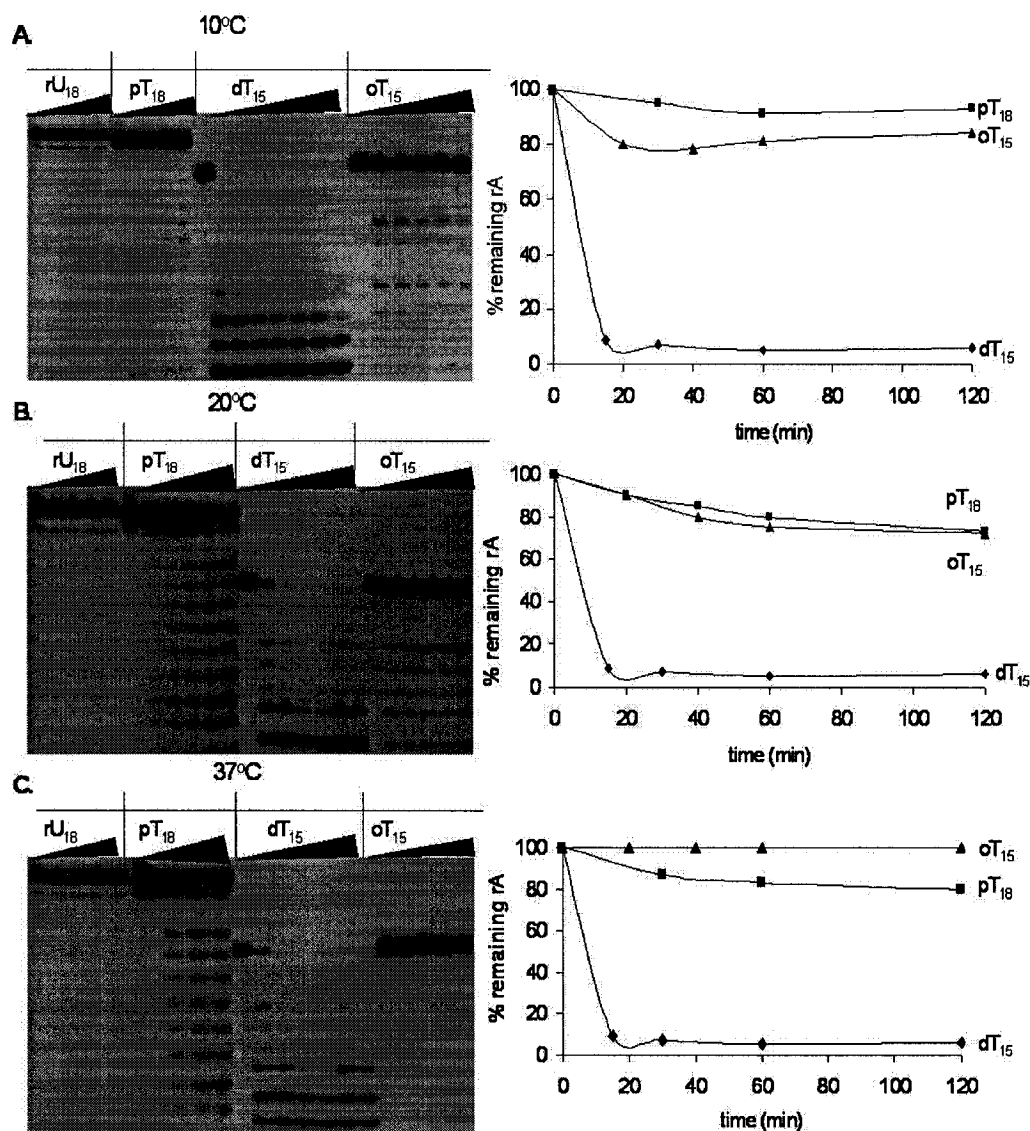
Figure 5.34: Rate of *E. coli* RNase H-mediated degradation of the RNA strand of 2'-OMe rU_{18}/rA_{18}, pT_{18}/rA_{18}, dT_{15}/rA_{15}, and oT_{15}/rA_{15}, duplexes at various temperatures.

E.Coli **RNase H-mediated degradation of dT_{15}/rA_{15}, oT_{15}/rA_{15} and pT_{15}/rA_{15}.** When comparing the rates of RNase H-mediated degradation (**Figure 5.34**) induced by 5 (furanose, dT_{15}), 6 (2'-ene-pyranose, pT_{18}) and 7-membered (oxepane, oT_{15}) ring oligonucleotides, at a temperature that ensures the highest hybrid population without significantly affecting RNase H activity, *i.e.* $10°C$ (**Figure 5.34- A**), the following trends are observed: $dT_{15} \gg oT_{15} > pT_{18}$ which parallels the decreasing conformational

163

flexibility of the constituent sugar rings, *i.e.* (5) furanose[189] > (7) oxepane[222c] > (6) 2'-ene-pyranose[190]. These results, and the fact that the pT and oT oligomers elicit RNase H activity at all, strongly support the notion that the plasticity of the DNA/RNA hybrid is essential for efficient RNase H catalysis, in which an enzyme-induced altered trajectory of the bound substrate facilitates optimal interaction with RNase H's catalytic site. Thus, the pT and oT oligomers studied here represents new classes of chemically-modified ONs capable of activating RNase H when bound to RNA, the others being, ANA,[61,62] 2'F-ANA,[64,237] CeNA,[79] and α-L-LNA.[240] These results also further demonstrate that a furanose-based antisense oligonucleotide strand is not a pre-requisite for enzyme activation. As my research supervisor once said, *an "overweight" nucleic acid (e.g., ONA) should be able to trigger RNase H activation provided that it can "flex"!*

RNAi activity of oxepane containing oligonucleotides. As mentioned in the Introduction (Chapter 1) RNA interference (RNAi) is a mechanism for the regulation of gene expression in which duplex RNA inhibits the expression of genes by targetting the mRNA encoded by such genes.[97] To test whether oxepane-modified siRNAs entered the RNAi pathway, the modified siRNAs 17-22 (**Table 5.6.3**) and the control siRNA duplex 9, were tested in the Hela cell line that over-expresses the firefly luciferase protein. Assays were carried out by Dr. Pelltier's group (McGill University), following procedures recently reported by our group.[100] Samples were transfected at 5 different concentrations for 24 h and the cells harvested for determination of the protein firefly luciferase counts normalized against a scrambled control. As this control does not bind to the firefly mRNA, any effects on the firefly production due to the transfecting agent (lipofectamine 2000 from InVitrogen Inc.) will be taken care of. The data (**Figure 5.35**) shows that oxepane-modified siRNAs were found to be less active relative to the control sequence 9, however, certain sequences particularly those containing the oN substitution in the sense RNA strand (*e.g.* duplexes 17, 18 and 19) maintained RNAi activity up to the low nM concentrations (1-10 nM). Modifications introduced in the critical antisense RNA strand (*e.g.*, 20, 21 and 22) and at the 5'-end of the siRNA duplex (20) were also found to be less tolerated. However, this effect was found to be compensated to a certain extent by re-introducing oN modifications in the sense strand (*e.g.* 21 *vs.* 22). The poor

activity of the siRNA bearing a 5'-end oT unit is likely related to the inability of such siRNA to be 5'-phosphorylated by the human Clp1 kinase[244] an essential step of the RNAi pathway.[97,100]
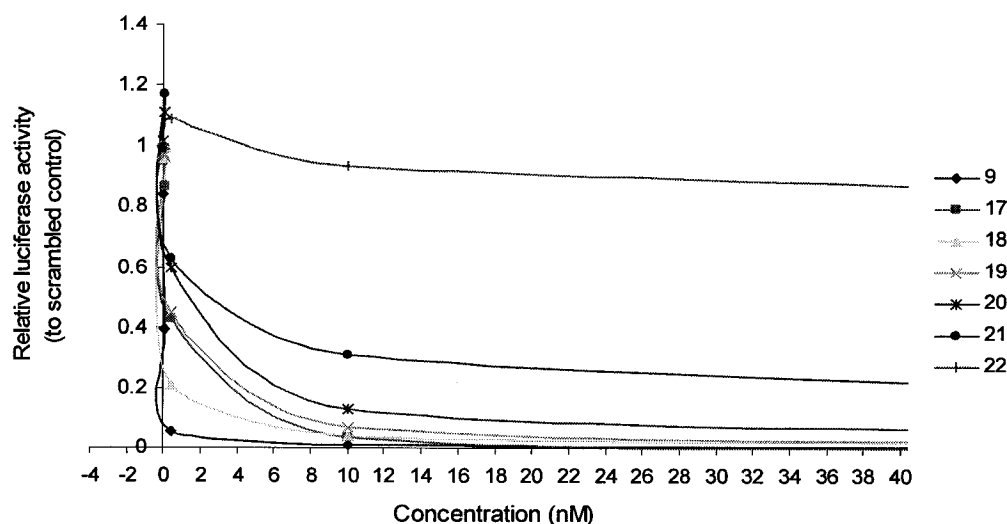
Figure 5.35: Dose-response curves of modified siRNAs targeting the luciferase firefly mRNA (duplex sequences are shown in **Table 5.6.3**).

Serum stability of oT$_{15}$ and oA$_{15}$. Resistance to *exo-* and *endo-*nucleases is imperative for *in vivo* applications of oligonucleotide-based therapeutics such as AONs, siRNAs and aptamers.[245] To determine the extent of nuclease resistance, the thymidylate and adenylate 15-mer sequences were subjected to 10% fetal bovine serum (FBS) and incubated at 37°C during a 24 hour reaction period.
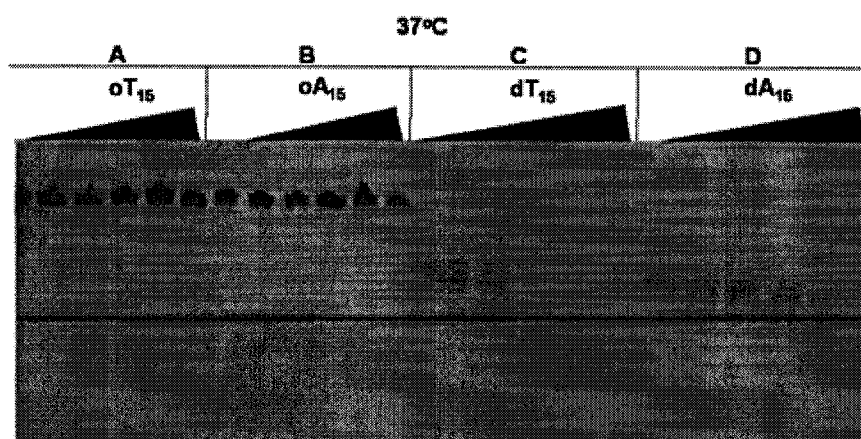
Figure 5.36: Nuclease resistance of oT$_{15}$, oA$_{15}$, dT$_{15}$, and dA$_{15}$. Oligonucleotides were incubated with fetal bovine serum at 37°C and quenched and analyzed after various incubation times, *i.e.*, 0, 0.5, 1, 2, 8, and 24 h.

A sample of the reaction mixture was quenched at successive time intervals before analysis by gel electrophoresis. Under these conditions, dT_{15} and dA_{15} were degraded within 2 and 8 hours, respectively (**Figure 5.36**). In marked contrast, the oxepane-modified oligomers (oT_{15} and oA_{15}) were completely resistant to cleavage under the same conditions with no noticeable degradation observed after 1 day. The 2'-ene-pyranose (pT) oligonucleotides were also found to be nuclease resistant according to the literature.[211,212]

5.7 CONCLUSIONS

Replacement of the furanose carbohydrate (dT_{15}) with oxepane (oT_{15}) confers exceptional resistance to nucleases while retaining the ability to direct RNase H-mediated cleavage of a target RNA. The oT_{15} and pT_{18} oligomers of the present study represent two new cases of chemically-modified ONs capable of activating RNase H when bound to RNA (the others being, in chronological order: ANA[61,62], 2'-F-ANA[64,237], CeNA[79], and α-L-LNA[240]). Although the extent to which cleavage occurs is lower than that observed for the wild type hybrid (dT_{15}/rA_{15}), it is significant that oT_{15} is able to affect cleavage within the RNA component at all, in light of the dramatically different structure of its sugar moiety (**Figure 5.1**) and oligonucleotide conformation (**Figure 5.23- A**). This highlights the importance of sugar conformer flexibility along the ON strand which likely acts in concert with the global helical architecture of the duplex to govern interactions between RNase H and its substrate.

Preliminary results on the biological activity of ONA-modified siRNA were gathered. While it is clear that the oxepane modification was detrimental to siRNA activity, it is encouraging that the ONA provides a highly nuclease resistant scaffold that may be exploited to increase the in vivo half-life of antisense and siRNA sequences. We have begun to explore alternative ONA structures (see Contributions to Knowledge, Chapter 6). For example, synthesis of more functionalized ONA nucleosides from oxepine nucleoside, **5.17**, have led to number of interesting nucleoside structures (D. Sabatino and M.J. Damha, unpublished results). These and other structures will continue to be tested in the Damha lab as potential inhibitors of DNA synthesis, or as ONAs that may bind cellular RNA (siRNA and antisense) and protein targets (aptamers).

CHAPTER 6: CONTRIBUTIONS TO KNOWLEDGE AND FUTURE WORK

6.1 FUTURE WORK

6.1.1 Synthesis of Novel Oxepane Nucleoside Derivatives

Background. The discovery of nucleosides with antiviral and anticancer activity generally relies on the rational approach by which they are designed to act through (a) initial conversion to their 5'-triphosphate derivatives and inhibition of nucleotide polymerase through chain termination of the growing viral DNA or RNA chain[246], (b) inhibition of polymerase through mechanisms other than chain termination[247], (c) incorporation into the viral genome, thereby disrupting expression of genetic information or (d) inhibition of a metabolic pathway necessary for viral replication[248]. Despite significant advances in antiviral therapies, several viruses [*e.g.,* West Nile virus, hepatitis C virus (HCV), influenza virus] have become a serious threat in North America, with only a few potent and selective inhibitors reported to date. Nearly 170 million individuals are infected with HCV, and currently, there is no effective treatment for this infection. Acquired immunodeficiency syndrome (AIDS), caused by the human immunodeficiency virus (HIV), has become one of the most lethal chronic diseases for which no cure has yet been identified. Of the numerous lead compounds studied, only those that specifically target HIV-1 reverse transcriptase or HIV-1 protease enzyme, and, more recently, the cell entry process, have been approved for HIV therapy. The common nucleoside reverse transcriptase inhibitors (NRTIs) e.g., the chain terminators AZT, 3TC, and d4T, and the non-nucleoside reverse transcriptase inhibitors (NNRTIs) *e.g.,* nevirapine, efavirenz, and delavirdine effectively block viral DNA replication and slow the onset and progression of AIDS.[249]

Despite the tremendous success associated with antiretroviral combination therapy for HIV, which may also include inhibitors of the viral protease, the development of resistance cannot be prevented and accounts for a major cause of treatment failure. Moreover, studies have suggested that a significant number of newly infected individuals in Europe and North America harbor resistant variants of the virus. The prevalence and

transmission of drug resistant variants is expected to increase, due to the extensive use of NRTIs, NNRTIs and protease inhibitors, which is an important factor that contributes to the limitations of treatment options for millions of infected individuals. HIV-2 is nearly resistant to all known NNRTIs, which is an additional factor that can severely compromise and restrict current treatment strategies. Vaccination against HIV-1 is still a long-range goal. Thus, the development of novel antiretroviral agents (such as novel nucleoside analogs) with potency against resistant HIV variants is of highest priority.

As emphasized in the Introduction Chapter 1, and Chapter 5, there is also a continued need for improved oligonucleotide antisense and siRNA chemistries. Further advances in these areas will come through improved chemistries that exhibit better efficacy and higher safety profiles, and are suitable to treat a wider variety of diseases. The unsaturated heptose-based (oxepine) nucleoside synthons provides the opportunity to synthesize a large number of nucleoside derivatives with potential antiviral/anticancer activity, and is an area that should (and will) be pursued in the Damha laboratory. Preliminary experiments and results carried out by the author (D. Sabatino) are briefly described below.

Synthesis of oxepane nucleoside derivatives. The functionalization of the oxepine double bond in oT* (**5.17** or **5.20**) provided a variety of modified ONA derivatives (**Figure 6.1**). These modifications were inspired by previous work on the syntheses of biologically important nucleoside mimics containing bicylcic and hydroxylated carbohydrate moieties.[250,251]

For example, epoxidation of **5.17** with *m*-chloroperbenzoic acid (*m*CPBA) was successively achieved to yield target compound **6.1**, in 50% yield. It was reasoned that the cyclic siloxane protecting group in **5.17** would assist in the stereoselectivity of the epoxidation reaction by facilitating the delivery of *m*CPBA to the α -bottom face of the oxepine carbohydrate.[252] However, this was not the case, as only a modest diastereoselectivity was observed, favoring the α-epoxy nucleoside to yield an inseparable mixture of 2α:1β after 24 hours reaction at 40°C. The structure of the α-cis

oxirane nucleoside, **6.1** was confirmed by the NOESY cross-peak assignments and the strong *syn* coupling between H3' and H4' ($^3J_{3'4'}$: 10.6 Hz).

The dihydroxylation reaction of **5.17** with catalytic (7 mol %) osmium tetroxide (OsO_4) in the presence of *N*-morpholine *N*-oxide, (NMO) as re-oxidant was found to generate the diol, **6.2**, in a 50% yield. This reaction was slow due to the steric effect of the bulky siloxane protecting group and the poor nucleophilicity of the oxepine double bond.[253] The reaction was completed after 5 hours generating the diastereomers, **6.2** as a 1:1 inseparable mixture of α:β cis hydroxylated isomers.

The cyclopropanation reaction with **5.17** following Furukawa's methods[221] yielded the cyclopropanated oxepane nucleoside, **6.3**, in modest yields of 30%. The reaction progress was followed by TLC for 24 hours, which resulted in *ca.* 50% conversion of the starting material. It was rationalized that the slow reaction was due to the poor reactivity of the double bond in **5.17**, as the reaction works best with sterically unhindered allylic alcohols as substrates.[215] It was also reasoned that the silyl ether protecting group could assist in the diastereoselectivity of the reaction.[254] The reaction yielded selectively, the α-cis cycloproponated oxepane nucleoside, **6.3**. The structure was confirmed by assignment of the COSY and NOESY crosspeaks and the strong *syn* proton coupling constant $^3J_{3'4'}$: 12 Hz.

The regioselective mono-hydroxylation reaction with **5.17** was also attempted for the synthesis of target compound **6.4**. Thus, hydroboration with borane-THF (BH_3-THF) proceeded slowly, requiring overnight reaction for completion. This was followed by the base hydrolysis (H_2O, NaOH) and oxidation (H_2O_2) reactions for 2 hours which generated the product diastereomers (2.5:1 ratio; 67% total yield) in favor of the β-cis adduct, **6.4**. The β-cis stereochemistry of **6.4** was established by NOESY cross-peaks and was rationalized by the steric influence of the neighboring 5'-silyl ether protecting group which prevented attack from the bottom face of the oxepine ring.[255] This protecting group also favored a 10:1 regioselectivity by facilitating the delivery of BH_3 to the C4' (relative to the C3') position of the reagent, **5.17**.
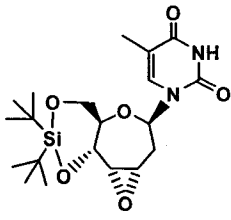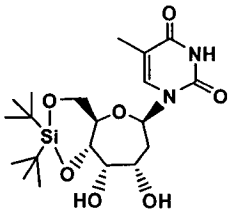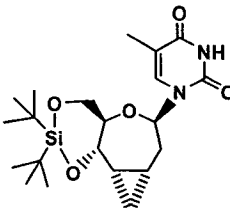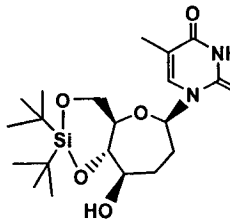
	6.1	**6.2**	**6.3**	**6.4**
Yield (total)	50%	50%	30%	67%
Diastereo-selectivity	2 : 1	1 : 1	100%	2.5 : 1

Figure 6.1: Yield and selectivity of synthesis of oxepane nucleoside derivatives prepared from **5.17.** Yields are isolated after chromatographic purification.

Conclusions. Oxepane nucleoside derivatives (**6.1** to **6.4**) are novel nucleosides containing a 7-membered heptose carbohydrate moiety. Each of these derivatives (in their completely deprotected form) represents potential nucleoside antivirals, or after elaboration into phosphoramidite derivatives, building blocks for ONA solid-phase synthesis. However, this is also contingent on developing high yielding, stereoselective methods for their synthesis. These 2nd generation oxepane-based nucleosides and oligonucleotides will allow us to further develop our investigations on nucleoside derivatives with potential antiviral/anticancer activity and on the chemical etiology of nucleic acids, inspired by the work of Herdewijn and Eschenmoser.[73]

6.2 GENERAL CONCLUSIONS AND CONTRIBUTIONS TO KNOWLEDGE

Novel Solid Phase Strategies toward RNA and bRNA Synthesis. The advent of oligoribonucleotide based therapeutics (siRNAs for RNAi strategies) has increased the demand for bulk quantities of biologically relevant oligoribonucleotides with high purity and at low cost. The selection of 2'-protecting groups for the large-scale, facile synthesis of RNA remains an area of intense research. Most of the previously described 2'-protecting groups all share the same requirements for a manual solution-phase deprotection. The need for manual 2'-deprotection is time and labor intensive, particularly for large-scale synthesis, and a potential source for material losses and ribonuclease contamination. In this thesis, 2'-*O*-levulinated rU was evaluated as possible

synthon for RNA synthesis. This monomer coupled with 1-2 min and 98-99% stepwise efficiency at 0.1 M concentration, thus reducing coupling time and consumption of reagent. More importantly, a ribouridylic acid strand synthesized with this monomer was deprotected rapidly under virtually neutral conditions (15 minutes, 1 M TBAF in THF). Despite these advantages, the 2'-O-Lv protecting group still requires further investigation and improvements, e.g. improving yield and separation of phosphoramidite 2'/3' isomers. Jeremy Lackey, a Ph.D. student in our lab has continued this research and addressed some of these limitations.[124]

The Chemical Stability of RNA Phosphate Diesters and Triesters on a Solid Support. The stability of the phosphorus center during the solid phase RNA synthesis is ambiguous since the phosphate or phosphite O-alkyl protecting groups can be modified with reagents and/or conditions that are commonly used in the synthesis cycle.[129,130] Since these reactions occur *in-situ*, it is often difficult to monitor and predict the success of a synthesis before the oligonucleotide has been synthesized and cleaved from the support. The chemical stability of a phosphate triester and diester bond at a 5'-terminal rU moiety of support bound rUdT$_5$ sequence has been monitored by AE, RP IP HPLC and PAGE over a range of chemical and pH conditions. The cleavage and isomerization reactions of the phosphate triester on solid support are faster than that of the same oligonucleotide sequence containing a rU 3'-phosphate diester linkage. This provided evidence that the phosphate CNEt protecting group is stable to the solid phase synthesis protocol and removed only during the ammonia conditions required during deprotection.

In the context of branched oligonucleotide synthesis, these results confirm that unmasking the phosphate triester (*i.e.*, decyanoethylation) must proceed removal of the adjacent 2'-protecting group in order to maintain the integrity (*i.e.* prevent isomerization of) the branchpoint point. This is also relevant in the synthesis of biologically useful oligoribonucleotide sequences such as the siRNAs[124] and with the synthesis of branched RNA and DNA for the splicing[45] and mechanistic[160,161] studies involved in the maturation process of mRNA. This includes the stability the unsual "RNA-X" species that has been detected during pre-mRNA splicing in vitro, in which a phosphate triester is present and vicinal to a 2'-OH group.[150] Robert Donga, a current Ph.D. student in our lab and

171

others[147] have synthesized this unstable phosphate triester in order to study its stability in solution.

The Solid Phase Synthesis of Branched and Hyperbranched oligonucleotides. A novel *divergent-growth* method for the synthesis of high purity bNAs and h-bNAs oligouridylate sequences has been developed as a complementary method to the *convergent-growth* synthesis of symmetric branched and hyperbranched oligonucleotides. The bNAs **4.1**, and **4.2** underwent selective debranching under acid, neutral or basic conditions at the 3',3' and 2',3' phosphodiester linkages, respectively (95 °C). The release of linear RNA products from Y-shaped RNA at physiological conditions (*e.g.* bNA **4.2**), and the fact that nucleic acid-based dendrimers have enhanced cellular uptake[151], make bNAs such as **4.2** (vicinal 2',3' and 3',5' phosphodiester linkages) potential "pro-drug" candidates for therapeutic applications (*i.e.* encapsulated bNA → slow release of 2 RNA strands → siRNA duplex).[124,256] Interestingly, bNA (**4.3**) was resistant to hydrolysis under similar conditions, underscoring nature's choice for the more stable 2',5'/3',5' branchpoint configuration. Furthermore, **4.1**, relative to the sequence with 'naturally' occurring branchpoint, **4.3**, exhibited a 3-fold decrease in the enzymatic 2',5'-*phosphodiesterase* debranching activity with yDBr1, making them potential yDBr1 inhibitors for their co-crystallization with the enzyme and mechanistic investigations related to the splicing mechanism of mRNA[160,161] This generates a specific requirement for developing an efficient synthesis strategy for the creation of asymmetric bNAs and h-bNAs with sequence compositions relevant to biological structures (*i.e.* msDNA).

A novel and efficient divergent method for the synthesis of discrete, and chemically diverse branched and hyperbranched msDNA sequences was developed. The synthesis of msDNAs containing atypical branchpoint linkages (2',5'/3',3' or 2',3'/3',5') were synthesized, and fully characterized by MS, and chemical and enzymatic assays. These modified Y-shaped oligonucleotides serve to probe the substrate specificity of the yeast lariat RNA debranching enzyme (yDBr1). For example, these studies revealed that yDBr1 is capable of debranching the 'un-natural' 2',5'/3',3' branchpoints. While debranching activity was considerably less than that observed with msDNAs with the natural (2',5' and 3',5'), the results show that the presence of a 3',5'-linkage is not

absolutely necessary for debranching activity. These studies also led to the identification of potent inhibitors of yDBr1, making these compounds potential good choices for future enzyme co-crystallization and X-ray analysis of the yDBR-msDNA complexes. Additionally, branched msDNA duplex structures provided mechanistic information related to the hydrolysis reaction catalyzed by *E. coli* RNaseH. These sequences generated substrates and inhibitors to the enzyme for potential applications in developing better antisense (refer to Chapter 5) based therapeutics[185], inhibitors of the RNaseH enzyme[180] and potentially good choices for future enzyme co-crystallization and X-ray analysis of the RNaseH-msDNA complexes.

Synthesis, Characterization and Properties of Nucleosides and Oligonucleotides Bearing 6 and 7-Membered Ring Carbohydrates. Replacement of the furanose carbohydrate (dT_{15}) with oxepane (oT_{15}) confers exceptional resistance to nucleases while retaining the ability to direct RNase H-mediated cleavage of a target RNA. The oT_{15} and pT_{18} oligomers of the present study represent two new cases of chemically-modified ONs capable of activating RNase H when bound to RNA (the others being, in chronological order: $ANA^{61,62}$, $2'\text{-}F\text{-}ANA^{64,237}$, $CeNA^{79}$, and $\alpha\text{-}L\text{-}LNA^{240}$). Although the extent to which cleavage occurs is lower than that observed for the wild type hybrid (dT_{15}/rA_{15}), it is significant that oT_{15} is able to affect cleavage within the RNA component at all, in light of the dramatically different structure of its sugar moiety (**Figure 5.1**) and oligonucleotide conformation (**Figure 5.23- A**). This highlights the importance of sugar conformer flexibility along the ON strand which likely acts in concert with the global helical architecture of the duplex to govern interactions between RNase H and its substrate.

Preliminary results on the biological activity of ONA-modified siRNA were gathered. While it is clear that the oxepane modification was detrimental to siRNA activity, it is encouraging that the ONA provides a highly nuclease resistant scaffold that may be exploited to increase the in vivo half-life of antisense and siRNA sequences. We have begun to explore alternative ONA structures (see Contributions to Knowledge, Chapter 6). For example, synthesis of more functionalized ONA nucleosides from oxepine nucleoside, **5.17**, have led to number of interesting nucleoside structures (D. Sabatino and

173

M.J. Damha, unpublished results). These and other structures will continue to be tested in the Damha lab as potential inhibitors of DNA synthesis, or as ONAs that may bind cellular RNA (siRNA and antisense) and protein targets (aptamers).

6.3 PUBLICATIONS, INVENTION DISCLOSURES AND CONFERENCE PRESENTATIONS

Manuscripts submitted and/or accepted.

1. David Sabatino and Masad J. Damha **Synthesis and Properties of oligonucleotides containing a 7-membered (Oxepane) Sugar Ring** *Nucleosides, Nucleotides and Nucleic Acids*, **2007**, in press (refereed conference proceeding Manuscript number: XVII-06-311 (Oct 9/2006).

2. Jeremy G. Lackey, David Sabatino and Masad J. Damha **Solid-Phase Synthesis and On-Column Deprotection of RNA from 2'- (and 3')-*O*-levulinated (Lv) Ribonucleoside Monomers.** *Organic Letters* **2007**, *9*(5), 789-792.

3. David Sabatino and Masad J. Damha **Oxepane Nucleic Acids (ONA): Synthesis, Characterization and Properties of Oligonucleotides Bearing a 7-Memebered Carbohydrate Ring** *Journal of the American Chemical Society* **2007**, *129*(26), 8259-8270.

Manuscripts Completed and In Preparation (targeted journals given).

1. David Sabatino and Masad J. Damha **Synthesis of Nucleoside Analogues based on a 7-Membered Carbohydrate Ring** (*Journal of Organic Chemistry*).

2. David Sabatino and Masad J. Damha **Synthesis and Properties of Unsaturated 6 and 7-Membered Carbohydrate Nucleic Acids and Oligonucleotides** (*Nucleic Acids Research*).

3. David Sabatino and Masad J. Damha **Chemical Stability of Phosphate Triesters and Diesters during Solid Phase Oligoribonucleotide Synthesis** (*Journal of Organic Chemistry*).

4. David Sabatino, Maxie Roessler and Masad J. Damha **Divergent Synthesis, Characterization and Properties of branched and hyperbranched RNA that contain vicinal 2',5'/3',3' and 2',3'/3',5' branchpoint linkages** (*Bioconjugate Chemistry*).

5. David Sabatino, Jeremy G. Lackey and Masad J. Damha **Novel Divergent Approach for the Solid Phase Synthesis of Branched and Hyperbranched msDNA,** (*Angewandte Chimie International Edition in English*).

Patent Disclosures.

1. David Sabatino and Masad J. Damha **Oxepane Nucleosides and Oligonucleotides** *US Provisional Patent Application*, filed August 31, **2006**, 24 pp.

Conference Presentations.

1. David Sabatino and Masad J. Damha **Synthesis and Physical Properties of Oxepine Nucleosides (poster presentation).** The 6th annual Chemistry and Biochemistry Graduate Research Conference, Concordia University, Montreal Quebec, November 2003.

2. David Sabatino and Masad J. Damha **Synthesis and Physical Properties of Oxepine Nucleosides (poster presentation).** The 14th annual Quebec-Ontario minisymposium in Synthetic and Bioorganic Chemistry, Universite de Montreal a Quebec, Montreal Quebec, December 2003.

3. David Sabatino and Masad J. Damha **Synthesis and Properties of Hexenepyranoside and Oxepine Oligonucleotides (oral presentation).** The 87th annual Canadian Chemistry Conference and Exhibition, Western University, London Ontario, May 2004.

❖ **Awarded the R.U. Lemieux award for outstanding oral presentation**

4. David Sabatino and Masad J. Damha **Novel Approaches towards the Synthesis of Branched and Dendritic Nucleic Acids (oral presentation).** The 2004/2005 RiboClub meeting seminar, University of Sherbrooke, Sherbrooke Quebec, February 2005.

5. David Sabatino and Masad J. Damha **Synthesis, Characterization and Biological Properties of Linear, Branched and Hyperbranched Nucleic Acids (oral presentation).** The 2005/2006 McGill University Organic Seminar, Montreal Quebec, February 2006.

6. David Sabatino and Masad J. Damha **Synthesis, Methodology and Characterization of Branched and Hyperbranched Oligonucleotides (oral presentation).** The 2005/2006 American Chemical Society Meeting & Exposition, Atlanta, GA USA, March 2006.

7. David Sabatino and Masad J. Damha **Synthesis and Characterization of Oxepane Oligonucleotides: further insight into the chemical etiology of nucleic acids (poster presentation).** The 2005/2006 Bern IRT International Round Table on Nucleosides, Nucleotides and Nucleic Acids, Bern, Switzerland, September 2006.

CHAPTER 7: EXPERIMENTAL SECTION

7.1 GENERAL METHODS

7.1.1 Solvents and Reagents

The solvents for anhydrous reactions were either dried and distilled prior to use or purchased from Aldrich Inc. The general solvents used, dichloromethane (DCM), pyridine (pyr), acetonitrile (MeCN) and 1,4-dioxane were dried over CaH_2, refluxed and distilled prior to use. Similarly, tetrahydrofuran (THF) and diethyl ether (Et_2O) were dried, refluxed with Na/benzophenone and distilled. Anhydrous methanol (MeOH), N,N-dimethyl formamide (DMF) and formamide were purchased from Aldrich and used as received. Most solvents were stored over activated 4 Å molecular sieves and septa sealed under dry N_2 until further use. Conventional ACS reagent-grade chemical solvents used for chromatography and/or work-up procedures [$i.e.$ ethyl acetate (EtOAc), hexanes (Hex), triethylamine (TEA), DCM, MeOH, cholorform ($CHCl_3$), ethanol (EtOH), and acetone (ace)] were not dried prior to use. Spectrophotometric grade solvents (MeOH, ace, CH_2Cl_2, THF) were used for instrumental analyses ($i.e.$ HPLC, ESI or CI MS, gene machine synthesizer $etc.$) and employed directly as received. Millipore water filtered with a 0.45 μm nylon membrane was also used for instrumental analyses. The Millipore filtered water (1 L) was autoclaved with 50 μL of diethylpyrocarbonate (DEPC) for 1.5 h at a pressure of 10 - 15 psi for oligonucleotide applications. Deuterated solvents for NMR analyses ($CDCl_3$, D_2O, DMSO-d_6, MeOH-d_4, acetone-d_6) were also used as received.

The following reagents (unless otherwise stated) were purchased from Aldrich Inc. and used as received: tri-O-acetyl-D-glucal (TAG), sodium methoxide (NaOMe) powder, di-t-butylsilyl di-(trimethylsilyl)trifluoromethane sulfonate, t-$Bu_2Si(OTf)_2$, 1 M diethyl zinc ($ZnEt_2$) in hexanes, diiodomethane (CH_2I_2), acetic anhydride (Ac_2O), thymine (T) or N^6-benzoyl adenine (N^6-Bz Ade), 1 M tetrabutylammonium fluoride (TBAF) in THF, 10% palladium charcoal (Pd/C) catalyst, monomethoxytrityl chloride (MMT-Cl), diisopropylethylamine [$EtN(i$-$Pr)_2$], N,N-diisopropylamino-β-cyanoethylphosphonamidic chloride- {Chemgenes Inc., [Cl-$P(OCEt)N(i$-$Pr)_2$]}, hexamethyldisilazane (HMDS), m-chloroperbenzoic acid (mCPBA), osmium tetroxide (OsO_4), 4-methylmorpholine N-oxide (NMO), borane tetrahydrofuran complex (BH_3-THF), Levulinic acid (LvOH), 1,3-dicyclohexylcarbodiimide solution (1M DCC in CH_2Cl_2), Novozyme® 435, silver nitrate

176

(AgNO₃), 2-chloro-*N*-methylpyridinium iodide (2-chloro-NMP iodide), 1,4-diazabicyclo[2.2.2]octane (DABCO), 1-H tetrazole (1-H tet), 2-cyanoethoxy bis(diisopropylamino)phosphane [P(OCEt)(N(*i*-Pr)₂)₂]. Analytical reagent grade glacial acetic acid (HOAc), hydrogen chloride (HCl), sodium hydroxide (NaOH), trichloroacetic acid (TCA), ammonium acetate (NH₄OAc), anhydrous sodium or magnesium sulfate (Na₂SO₄ or MgSO₄), sodium hydrogen sulfate (NaHSO₄), hydrogen peroxide (H₂O₂), ammonium chloride (NH₄Cl), sodium chloride (NaCl), sodium bicarbonate (NaHCO₃), sodium carbonate (Na₂CO₃), magnesium or manganese chloride (MgCl₂ or MnCl₂), analytical grade dissodium ethylenediaminetetracetate dihydrate (EDTA), boric acid, disodium hydrogen phosphate (Na₂HPO₄), sodium acetate (NaOAc), 2-amino-2-hydroxymethyl-1,3-propanediol (Tris) were obtained from BDH Inc. or Bio-Rad Inc. and used as received.

7.1.2 Chromatography

Column chromatography was performed with silica gel [40-63 micron Silica gel 60 (EM science Gibbstown NJ)]. Thin layer chromatography was performed with Merck Kieselgel 60 F-254 aluminum back analytical silica gel sheets (0.2 mm thickness, EM Science, Gibbstown, NJ.) and the plates were visualized with a solution of 10% sulfuric acid (H₂SO₄) in MeOH and/or UV light.

7.1.3 Instrumentation

UV Spectroscopy. UV spectra for oligonucleotide quantitation (absorbance measurements) were measured at 260 nm on a Varian Cary I or 300 UV-VIS dual beam spectrophotometer. These were converted into stoichiometric values by correlation with the molar extinction coefficient (ε_{260}) of the oligonucleotide sequence. The molar extinction coefficients were based on those of the corresponding mono- and di-nucleotides using the nearest neighbor approximation method of Puglisi and Tinoco.[260]

Thermal denaturation of oligonucleotide complexes was followed in the ultraviolet spectrum (260 or 284 nm) using a Varian Cary I or 300 UV-VIS dual beam spectrophotometer equipped with a thermoelectrically controlled cell holder with data collection by software supplied by the manufacture (Cary Win™ UV versions 1.3e or

2.00). All readouts were acquired on personal computer with Microsoft™ Office Software (Excel™ XP).

NMR Spectroscopy. Compound characterization and analyses were performed at ambient temperatures on a Varian 400-500 MHz spectrophotometer with chemical shift values reported in ppm (δ) downfield from tetramethylsilane (TMS) internal standard. The characterization was performed by ^{1}H, ^{13}C and ^{31}P NMR (^{31}P NMR spectras were collected at 80 MHz as a ^{1}H decoupled experiment with a Varian 200 or 500 MHz spectrophotometer). For the new compounds synthesized, the ^{1}H and ^{13}C and ^{31}P chemical shift assignments were confirmed by the 2-D NMR experiments which included homonuclear (^{1}H-^{1}H COSY) and heteronuclear (^{1}H-^{13}C HMQC, ^{1}H-^{31}P CIGAR) correlated NMR spectroscopy. Furthermore, the absolute stereochemistry of certain compounds was ascertained by Nuclear Overhauser and Exchange spectroscopy (1-D NOE and 2-D NOSY experiments). All samples (3-10 mg) were prepared in deuterated solvents (700 μL) prior to the NMR analyses.

Mass Spectrometry. The compounds were further characterized by molecular weight analysis using electrospray or chemical ionization mass spectrometry, ESI-MS or CI-MS, with a Finnigan LCQ DUO mass spectrometer in the negative mode for the sample (1-3 mg) dissolved in spectrophotometric grade methanol or acetone (1-3 mL).

Additionally, oligonucleotide samples were analyzed by molecular weight with matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI TOF MS) on a Kratos Kompact-III instrument with a minimum laser output of 6 mW at a wavelength of 337 nm light, 3 ns pulse width and 100 mm diameter spot. The spectra were collected in the negative linear mode with samples (*ca.* 200 pmol) suspended in a matrix consisting of 6-aza-2-thiothymine/spermine (80 mg/mL ATT with 1:1 H$_2$O: MeCN) and a 50 mM aqueous L-fucose solution.

Circular Dichroism Spectroscopy. The circular dichroism spectra were analyzed using a Jasco Model J-710 or J-810 spectropolarimeter. This was performed with a temperature controlled sample cell using a circulating bath (NESLAB® RTE-111). The

178

CD spectra were generated from the wavelength range of 310 – 200 nm and as the sum of 3 replicate scans which were finally baseline corrected from the physiological phosphate buffer. The samples were generally prepared in a duplex concentration of *ca.* 3.04 μM in 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄ pH: 7.2.

7.2 AUTOMATED SOLID PHASE OLIGONUCLEOTIDE SYNTHESIS

7.2.1 Reagents for Derivitization of Nucleosides and their assembly into oligonucleotides

The reactions for the preparation of the nucleoside monomers and their automated assembly into oligonucleotides are sensitive to moisture and require anhydrous glassware, reagents and handling. The solvents and reagents for phosphoramidite preparation, nucleoside derivitization to the solid support and oligonucleotide syntheses were of the highest quality and kept completely anhydrous during use.

7.2.2 Derivitization of the Solid Support

The following procedure was used to couple the tritylated nucleoside monomers with the CPG support. Into a dried 5 mL glass vial was added the tritylated nucleoside monomer (0.7 mmol) with coupling reagents *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HATU), or *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU) (0.1 mmol), and *N,N*-dimethylaminopyridine (DMAP) (0.1 mmol) and the solid support [succinyl linked 500 Å long chain alkyl amino controlled pore glass *i.e.* succinyl linked LCAA CPG (250 mg)] prepared according to literature procedure.[19] The reaction was completed at room temperature (22°C) with N₂ atmosphere in anhydrous MeCN (1 mL). The extent of the coupling reaction is contingent with the reaction time (20-40 min produces low nucleoside loading 20-40 μmol/g while longer reaction times can yield higher loadings of *ca.* 80 μmol/g). The nucleoside derivatized solid support was filtered, washed successively with 25 mL of DCM, 50 mL of MeOH and additional 25 mL of DCM prior to determining the nucleoside loading. The nucleoside loadings were determined by spectrophotometric mono- and dimethoxytrityl cation colorimetric assay (Chapter 1,

Figure 1.4). The support was dried *in-vacuo* for 24 h before use, loaded into an empty synthesizer column with replaceable filters (ABI), crimped closed with aluminum seals (ABI), and installed on the instrument.

Alternatively, Manoharan and co-workers have recently published a facile procedure for the synthesis and applications of a 'universal linker' (Unylinker™ support) LCAA CPG for the automated synthesis of DNA, RNA and modified oligonucleotides.[169] This is extremely desirable in the case of chemically modified oligonucleotides because the tritylated nucleoside phosphoramidites can be coupled directly to the Unylinker™ support. This Unylinker™ support LCAA CPG is now commercially available and was purchased from Chemgenes Inc. for the synthesis of certain sequences.

7.2.3 Automated Solid Phase Synthesis of Oligonucleotides

Monomers and Reagents for Automated Synthesis. The solid phase syntheses of oligonucleotides were conducted on either an ABI 3400 or 381A gene machine synthesizer. The reagents for the solid phase synthesis procedure included: 1) the *detritylation reagent* (3% solution of TCA in DCM), 2) the *coupling reagent* (0.25 M ethylthiotetrazole in acetonitrile or 0.25 M ETT in MeCN), 3) the *capping reagents* (Cap A : 1:1:8 v/v/v Ac$_2$O:pyr:THF, and Cap B: 10% *N*-methyl imidazole in THF or *N*-Me Im in THF), 4) the *oxidation reagents* (0.1 M iodine in 75:20:5 v/v/v THF:pyr:H$_2$O) and 5) *acetonitrile wash* (Biotech grade purchased from EMI with low water content and 99.999% purity). These reagents were purchased and used as anhydrous reagents and solvents from Chemgenes Inc.

The oligonucleotide syntheses were performed on 0.3 − 1 μmol scales with 500 Å succinyl linked LCAA CPG derivitized with the tritylated nucleoside monomers[19] or directly on the universal linker[169] LCAA.

The modified nucleoside phosphoramidites were prepared in 0.05-0.15 M solutions with anhydrous CH$_2$Cl$_2$, or MeCN. The coupling times for these were extended to 30 min with 0.25 M ETT in MeCN as activator and the detritylation time was also extended to 2.5 min. For the conventional DNA and RNA syntheses[24], DNA phosphoramidites were prepared as 0.1 M solutions in anhydrous MeCN and coupled to the solid support for 2 min. The RNA amidites were prepared as 0.15 M solutions in anhydrous MeCN and the

180

coupling times were extended to 10 min reactions (rG amidites required 15 min coupling times).

The assembly of oligonucleotides was performed with the following sequential synthesis steps: 1) *detrityaltion*: a DCM wash step for 40 s followed by 3% TCA in DCM delivery for 120 s which cleaves the trityl protecting group that can be used for determining coupling yields by UV spectroscopic quantitation (DMT$^+$: λ: 504 nm, ε: 76 000 L/mol cm^{-1} and MMT$^+$: λ: 478 nm, ε: 56 000 L/mol cm^{-1}), 2) *coupling*: a delivery of phosphoramidite dissolved in CH$_2$Cl$_2$ or MeCN (0.05 − 0.15 M solutions) with the activator (0.25 M ETT in MeCN) for a coupling time of 90 s (DNA), 10 or 15 min (RNA) and 30 min (modified nucleoside amidites), 3) *capping*: the delivery of Cap A and Cap B for 15 s and a wait time for 45 s followed by, 4) *oxidation*: the delivery of the oxidant (0.1 M iodine in 75 : 20 : 5 v/v/v THF : pyr : water) for 20 s and an additional wait step for 20 s. Prior to the oligonucleotide assembly, the derivitzed support was capped (*capping cycle* provided by manufacture) to block the undesired reactive sites.

7.2.4 Complete Deprotection of Synthetic Oligonucleotides

The CPG bound oligonucleotide sequences were dried with argon in the synthesizer column (10 min) and transferred into an autoclaved, 1.5 mL screw cap microtube. The oligomer bound support, was treated with a 1 mL solution of 3:1 v/v ammonium hydroxide (NH$_4$OH) in absolute EtOH. The cleavage reaction of the oligomer from the support and the protecting group deprotection reactions were performed at 55°C for 4 - 6 h, and for mixed base sequences for 16 − 24 h to ensure the complete removal of the oligonucleotide base protecting groups.[24] After cleavage and deprotection, this solution was evaporated to dryness and the oligomer was re-suspended in autoclaved water for determining the yield by UV absorbance measurements.

Also, a selective decyanoethylation procedure for the removal of the phosphate 2-cyanoethyl protecting groups converts the phosphate triester to the more stable diester backbone.[40] This was performed during the automated synthesis of bNAs and h-bNAs in order to prevent un-desired branchpoint isomerization or cleavage reactions (Chapter 3).[22,33,34] The 'on-column' decyanoethylation procedure was performed for 90 min with a solution of 4:6 v/v NEt$_3$:MeCN, followed by washing of the oligomer bound support with

30 mL or THF and MeCN.

Alternatively, RNA oligoribonucleotides require a 2'-desilylation reaction.[40] This was performed with 300 – 500 μL of anhydrous triethylamine trihydrofluride, TREAT HF[257], (Aldrich) on an over-head shaker for 48 h at ambient temperature (22°C).[24] Alternatively, a faster desilylation procedure[24,258] was followed by adding 0.3 mL of a solution of 0.75 mL NMP, 1.0 mL TEA and 1.5 mL TREAT HF at 65 °C for 90 min. After complete reaction, the oligoribonucleotide was precipitated directly with 25 μL of a 3 M NaOAc solution and 1 mL n-BuOH. The precipitation process was optimized with dry ice for 2 h, centrifuged and the supernatant removed prior to dissolving the crude oligomer in autoclaved water. The recovery of crude oligoribonucleotide sequence from precipitation was determined by UV absorbance and this was followed by purification of the crude.

7.3 PURIFICATION OF OLIGONUCLEOTIDES

7.3.1 General Reagents

These reagents were prepared prior to purification of the oligonucleotide sequences. These included the preparation of a 10x TBE buffer solution [500 mL: TRIS, (54.5 g), boric acid (27.83 g) and EDTA (1.86 g) in millipore water (filtered on a 0.45 μm filter)]. This buffer was also diluted (0.5 - 1x TBE) with millipore H_2O for use as a running buffer solution for PAGE purification procedures. These buffers can be stored at 4°C and re-cycled after PAGE purification of the sequences. The preparation of gel loading buffer was prepared with 10 mL of formamide (Aldrich Inc.) which was de-ionized with a mixed bed resin (lot # AG 501-X8 purchased from BioRad). This mixture was stirred for 30 min for complete de-ionization and filtered prior to use. This solution can be stored for an indefinite period of time at -20°C. Dye solutions (dye indicators while running PAGE) were prepared in two separate Falcon® Tubes with de-ionized formamide (2 x 10 mL) and minimal bromophenol blue (BPB) and xylene cyanol (XC) indicators. These solutions were vortexed to homogeneity and can be stored indefinitely at -20°C. The denaturing polyacrylamide solutions for PAGE analysis and purification of oligonucleotides was prepared as a 24% gel solution with 7 M urea (42.04 g), 40% acrylamide solution (contains acrylamide and *N,N*-methylene bisacrylamide, 60 mL), 10x

TBE buffer (10 mL), or as a 16% gel solution with 7 M urea (42.04 g), 40% acrylamide solution (contains acrylamide and *N,N*-methylene bisacrylamide, 40mL), 10x TBE buffer (10 mL) and each was diluted to 100 mL with millipore water. Aqueous solutions of lithium perchlorate, $LiCLO_4$ (Aldrich) for anion-exchange HPLC, (AE HPLC) and triethylammonium acetate (0.1 M TEAA in MeCN pH: 7) for reverse phase ion-pairing HPLC, (RP IP HPLC) analyses were filtered through a 0.45 μm pore nylon membrane filter and degassed under vacuum prior to use.

7.3.2 Polyacrylamide Gel Electrophoresis (PAGE)[24]

The crude oligomers were purified by vertical slab polyacrylamide gel electrophoresis (PAGE) using a Hoefer® Scientific Unit. The thickness of the gels was 0.75 mm and 1.5 mm for analytical and preparative gels, respectively. The gels were polymerized with of a 10% (wt/v) ammonium persulfate (APS) (300 μL) and TEMED (20 μL) (Amersham BioSciences Inc.) at ambient temperature (22°C) for 1 – 2 h. The crude olignucleotide samples (*ca.* 0.5 A_{260} units for analytical and 20 – 30 A_{260} units for preparative PAGE) were evaporated to dryness and re-suspended in loading buffer (formamide) prior to analysis or purification. These samples were run in addition to the dye solutions, (BPB and XC) which were used as indicators during purification. The gels were run at 500 – 800 V until the faster BPB dye had migrated to two-thirds down the length of the gel surface. For analytical PAGE, the gel plates were removed and the gel was carefully placed on an acetate transparency. The gel was photographed over a fluorescent plate while illuminated by a hand-held UV lamp. The oligomers can also be visualized with Stains-All® (Sigma) solution (25 mg Stains-All®, 50 mL *i*-PrOH, 25 mL formamide, 125 mL water, stirred to homogeneity) for 2 - 4 h at ambient temperatures (22°C). Similarly, for PAGE purification procedures, the desired oligomer products were excised from the gel, transferred into a sterile culture tube and extracted in autoclaved water at ambient temperatures (22°C) for 12 – 16 h. The purified oligomers were desalted (Sephadex® G-25) prior to further use.

7.3.3 Anion Exchange (AE) HPLC[143]

The crude oligomers were also analyzed and purified by Anion Exchange (AE) HPLC

using a Waters® 1525 HPLC instrument equipped with 2 solvent delivery pumps, a dual wavelength absorbance detector (Waters® model 2487), in-line degasser and internal column heater driven by the Windows-based software (*i.e.* Breeze® v. 3.20). The crude samples were resolved on a Protein Pak™ DEAE-5PW anion exchanger column, consisting of DEAE-bonded, hydrophilic, porous (1000 Å pore size) polymer particles 10 μm in size and with overall column dimensions of 75 mm x 7.5 mm (internal diameter). The crude oligomer samples for analytical (0.5 – 1 A_{260} units) and preparative (15 – 20 A_{260} units) injections were dissolved in autoclaved water and separated by AE HPLC using a linear elution gradient of 0 – 30% 1 M $LiClO_4$ (with millipore water) and a flow rate of 1 mL/min at a column temperature of 60°C during 70 min. The detector was tuned to monitor the absorbance at 260 nm for analytical scales and preparative separations were simultaneously monitored at dual wavelengths of 260 and 290 nm in order to avoid saturating the detector. The desired oligomer products were collected in 1 mL micro-tubes and the samples were desalted (Sephadex® G-25) prior to further use.

7.3.4 Reverse Phase Ion-Pairing (RP IP) HPLC[144]

In certain instances, the crude oligomers were also analyzed and purified by reverse phase ion pairing HPLC (RP IP HPLC). This was performed on a Waters® 1525 HPLC instrument equipped with 2 solvent delivery pumps, a dual wavelength absorbance detector (Waters® model 2487), in-line degasser and internal column heater driven by the Windows-based software (*i.e.* Breeze® v. 3.20). The crude samples were dissolved in autoclaved water (*ca.* 0.5 A_{260} units for analytical runs and 10-15 A_{260} units for preparative runs) and injected into a Waters® Symmetry C-18 reverse phase column (4.6 x 150 mm, 5 μM particle size) using a linear elution gradient of 95 – 83% (0.1 M TEAA in MeCN, pH: 7.1 and spectrophotometric grade MeCN) with a flow rate of 1 mL/min and a column temperature of 60°C during 70 min. The desired products were collected in 1.5 mL micro-tubes and evaporated to dryness in a Speed® Vac concentrator prior to further use.

7.3.5 Desalting of Oligonucleotides

The purified samples (from PAGE and AE HPLC) were desalted from water soluble

counterions and lower molecular weight impurities by gel filtration with Nap® 10 or 25 size exclusion chromatography columns containing Sephadex® G-25 Superfine medium (Amersham Inc.), prepared by cross-linking dextran with epichlorohydrin.[259] This formed a gel in autoclaved water and was used to elute the purified oligonucleotide samples (*ca.* 0.5 – 1 A_{260} units in 1 mL autoclaved water with Nap® 10 columns and *ca.* 15 – 20 A_{260} units in 2.5 mL autoclaved water with Nap® 25 columns) in 1 mL fractions collected in sterile 1.5 mL microtubes.

7.4 BIOPHYSICAL CHARACTERIZATION OF OLIGONUCLEOTIDES

7.4.1 UV Thermal Denaturation Studies[231]

Oligonucleotide hybridization experiments were performed on a Varian Cary I or 300 UV-VIS spectrophotometer equipped with a Pelltier temperature controller interfaced to a PC running Windows® based software (Win 3.1 or Win 2000 Professional). The thermal melt experiments for hybridized complementary oligonucleotide strands were monitored by hyperchromic changes in the UV absorbance at 260 nm (and 284 nm for triplex studies) with increasing temperatures (5 - 80°C). The melting curves were collected with a data interval and temperature gradient of 0.5°C/min under constant N_2 flow to prevent condensation at lower temperature. Molar extinction coefficients (ε_{260}) for the single strands were calculated based on those of the mono- and di-nucleotides using the nearest neighbor approximation method of Puglisi and Tinoco.[260] The melting temperature, (T_m) for the complexes was calculated from the first derivative plots of the melting curve, which produces a maximum value corresponding to the inflection point of the melting transition and represents the temperature at which 50% of the complex has disassociated. The spectra were acquired in duplicate scans to ensure reproducibility. The absorbance versus temperature data were converted to ASCII binary format and imported into a spreadsheet with the Microsoft Excel™ XP Software for further manipulation. Hyperchromicity values (% H) were calculated as relative changes in the absorbance (260 nm) at a given temperature. This was calculated with the equation: % H = $[(A_T - A_o)/A_f]$ x 100, where H is the hyperchromicity, A_T is the absorbance at a given temperature (T),

A_o is the initial absorbance at the start temperature and A_f is the absorbance at the final temperature.

Samples for T_m studies were generally prepared by evaporating an equimolar mixture of complementary strands to dryness with a Speed® Vac concentrator and then re-dissolving them in 1 mL of the appropriate buffer for a duplex concentration of 3 – 5 μM. The buffer typically consisted of a physiologically relevant phosphate buffer (*i.e.* 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄ adjusted to pH: 7.2). The component singles strands (1.5 – 2.5 μM) were also analyzed by T_m experiments. The solutions were heated at 90°C for 10 – 15 min to denature the complex, and then slowly cooled to room temperature for 1.5 – 2 h and annealed overnight (12 – 16 h) at 4°C prior analysis. The hybridized samples were quickly transferred into pre-chilled (on ice) Hellma® QS- 1.000 quartz cells and sealed with stopper and conserved with parafilm to prevent solvent evaporation during the thermal analysis. The solutions were degassed by sonication for 5 – 10 sec and further equilibrated at 5°C for 5 min in the cell chamber with N₂ flow prior to the analysis. The N₂ was continuously flushed through the chamber to prevent condensation at low temperatures (5 – 25°C).

7.4.2 CD Hybridization and Structural Studies[173,228]

Samples for the CD experiments were prepared similarly to the T_m experiments with typical phosphate buffers (*i.e.* 3 – 5 μM complex in 140 mM KCl, 1 mM MgCl₂, 5 mM Na₂HPO₄ adjusted to pH: 7.2). The component singles strands (1.5 – 2.5 μM) were also analyzed by the CD experiments. Since CD is also useful in monitoring the melting of hybrid complexes, the relative change in ellipticity (mdeg) at a given wavelength (310 – 200 nm) with temperature (5 – 80°C) produced a thermal transition curve from which T_m values were determined.

Briefly, equimolar mixtures of the single strands were annealed by pre-heating the solution (90°C) for *ca.* 10 min and cooling to room temperature (22°C) for approximately 2 h and at 4°C for 12 – 16 h. The samples were transferred to Hellma QS-1.000 (Cat # 114) fused quartz cells and maintained within the sample cell holder at 5°C with N₂ for 10 – 15 min prior to the spectral acquisition. The CD spectra were collected on a Jasco J- 710 spectropolarimeter equipped with a thermoelectrically controlled external constant

186

temperature (NESLAB® RTE-111 circulating bath). Samples were dispensed in cells (Hellma® QS- 1.000 quartz cells) with a 1 cm path length. Before the CD spectra were determined, the samples were equilibrated at a given temperature for approximately 5 min with N_2. Each spectrum was collected as an average of 3 scans at a rate of 100 nm/min and band width of a 1 nm interval. The sampling wavelength was adjusted to 0.2 nm and the spectra were analyzed between 350 and 200 nm. The raw data was processed using J-700 Windows® software (version 1.00) as supplied by the manufacture, and was normalized by subtraction of the buffer, noise reduction (*i.e.* line smoothing), and concentration such that the molar ellipticity was calculated from the equation $[\theta] = \theta /cl$, where θ is the relative ellipticity (mdeg), c is the molar concentration of the oligonucleotides (M) and l is the path length of the cell (cm). The data was imported and processed with Microsoft Excel™ XP spreadsheet after manipulation with J-700 Windows® software (version 1.00).

7.4.3 UV Stoichiometric Studies (mixing curves or Job plots)[232]

The proportion in which (complementary) strands associate can be determined by monitoring the relative change in absorbance values at a given wavelength (260 nm) with titration of a solution containing one strand to an equimolar solution of the second complementary strand. This study was conducted on a UV-VIS Cary 300 dual beam spectrophotometer. Equimolar stock solutions of each oligonucleotide strand (2.5 nmol) were prepared in 0.5 mL of buffer (10 mM Na_2HPO_4, 100 mM NaCl, pH: 7.2 or 10 mM Na_2HPO_4, 50 mM $MgCl_2$ pH: 7.3). At low temperature (5°C), 100 μL aliquots (0.5 nmol) of T/U solution was titrated into the stock solution containing A, and allowed to hybridize for 5-10 min prior to measuring the UV absorbance at 260 nm under constant flow of N_2. Absorbance values were measured at 260 nm for each mol fraction of T titrated into a fixed concentration of a complementary A stock solution to determine the stoichiometry of the hybridization interaction.

7.5 GENERAL MOLECULAR BIOLOGY TECHNIQUES AND STUDIES: OLIGONUCLEOTIDE LABELING, CHARACTERIZATION AND ENZYME PROPERTIES

7.5.1 5'-End [^{32}P]-Labeling of Synthetic Oligoribonucleotides

Synthetic oligoribonucleotides were 5'-end radiolabeled with a radioactive phosphate probe and the enzyme T4 polynucleotide kinase (T4 PNK) according the manufacture's directions (MBI Fermentas® Life Sciences, Burlington, ON).

This was performed with a reaction mixture consisting of a purified and 'dried-down' RNA substrate (200 pmol), suspended in 2 μL of 10x reaction buffer (500 mM Tris-HCl, pH: 7.6, 100 mM MgCl$_2$, 50 mM DTT, 1 mM spermidine and 1 mM EDTA), 1 μL of T4 PNK enzyme (20 U in 20 mM Tris-HCl, pH: 7.5, 25 mM KCl, 0.1 mM EDTA, 2 mM DTT and 50% glycerol), 10 μL of [γ-^{32}P]-ATP (6000 Ci/mmol, 10 mCi/mL, Amersham Biosciences Inc.) and diluted to 20 μL with autoclaved water. The mixture was incubated at 37°C for 45 - 60 min to complete the labeling reaction, heated at 95°C for 5 min to denature the enzyme and the reaction was evaporated to a solid pellet.

The crude reaction mixture was dissolved in 10 μL of loading dye buffer (deionized formamide with XC and BPB dyes) and purified by PAGE with a 16 % sequencing gel that was run for 2.5 h at 2000 V, 30 mA and 55 W.

The gel plates were disassembled; the gel was placed on an autoradiogram film and covered with Saran® Wrap. This was carefully placed in a Kodak® X-Omatic cassette and the labeled oligoribonucleotides were visualized by autoradiography. The desired products were excised from the gel, placed in an autoclaved Eppendorf® and the labeled oligoribonucleotides were extracted in 1 mL of autoclaved water for 12 – 16 h at 37°C. The purified oligoribonucleotides were desalted with autoclaved water (1.5 mL) using size exclusion chromatography (Nap® 10 columns with Sephadex® G-25). The samples were evaporated to dryness with a Speed Vac® concentrator and the amount of ^{32}P-5'-RNA label was determined (CPM) with a Bioscan® Quick Count QC-2000 high energy benchtop beta counter (Bioscan Inc.). The incorporation of the label was usually greater than 90%, with isolated yields of ^{32}P-5'-RNA following gel extraction of 50%.

7.5.2 RNaseH Induction Assays

An aliquot (7 pmol) of the radiolabeled $5'-^{32}P$-RNA strand (200 pmol) was added with a complementary (1.8 fold excess) DNA single strands and annealed with 5x reaction buffer (20 pmol in 10 µL buffer; 100 mM Tris HCl pH: 7.5, 100 mM KCl, 50 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT) and autoclaved water (100 µL, total volume). The complementary sequences were denatured for 5 – 10 min at 95°C, slowly cooled to room temperature for 1.5 – 2 h and annealed overnight for 12 – 16 h at 4°C prior to the enzymatic reaction. Similarly, complementary msDNA sequences were 5'-end radiolabeled, purified, desalted and annealed directly by the previously described methods.

E. coli RNaseH assays were performed at different temperature conditions (10°C, 20°C and 37°C) in 10 µL reactions containing 2 pmol of duplex substrate with 2 µL of reaction buffer (5x RNaseH buffer, Amersham Biosciences Inc.) and 0.5 µL of *E. coli* RNaseH enzyme (Amersham Biosciences Inc., concentration of 5 units/µL in storage buffer 20 mM Tris HCl pH: 7.9, 100 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM DTT and 50% glycerol). Each reaction was quenched at various time points by heating to 90°C and with the addition of "stop solution" (10 µL, 50 mM EDTA in formamide with BPB and XC dyes) prior to analysis by 16% PAGE that was run for 2.5 h at 2000 V, 30 mA and 55 W.

The reactions were analyzed and visualized by autoradiography. The extent of the cleavage reaction for the radiolabeled RNA portion of the RNA/DNA hybrid was determined quantitatively by densiometric analysis (UN-SCAN-IT™ software) with the disappearance of the full length RNA and/or the appearance of the smaller RNA degradation products.

7.5.3 RNaseH Inhibition Assays

A control RNA (rA_{15}) sequence was 5'-end radiolabeled, purified, desalted and annealed to the complementary DNA (dT_{15}) sequence as previously described (section 7.5.1 and 7.5.2). This was initially treated with the *E. coli* RNaseH enzyme to determine the optimized reaction conditions for the enzymatic hydrolysis of the substrate (section 7.5.2).

The inhibition of the *E. coli* RNAseH enzyme with msDNAs was performed by preparing 10 μL stock solutions of 'cold' (non-labeled) msDNA inhibitors (500 - 0.05 μM) in 5x RNaseH buffer. These were denatured for 5 – 10 min at 95°C, slowly cooled to room temperature for 1.5 – 2 h and annealed overnight for 12 – 16 h at 4°C prior to the enzymatic reaction. The inhibitors (1 μL) were added in increasing concentrations (50 - 0.005 μM) to the substrate (0.5 – 2 pmol) in the presence of 0.5 μl of the *E. coli* RNaseH enzyme, with 2 μL of 5x RNaseH buffer and diluted to a 10 μL reaction volume with autoclaved water. The reactions were performed for 5 min at 20°C and the reaction was stopped with 10 μL of "stop solution" prior to analysis.

The reactions were analyzed by autoradiography and determined quantitatively by densiometric analysis (UN-SCAN-IT™ software).

7.5.4 Debranching with the yDBr1 enzyme

The branched oligonucleotides (200 pmol) were initially radiolabeled at their 5'-end(s) with γ-^{32}P ATP and the enzyme T4 PNK as previously described (section 7.5.1). These were purified, desalted and concentrated to dryness.

The enzymatic debranching reactions were performed with the pure radiolabeled bNAs (0.5 - 2 pmol) as substrate with 1 μL of yDBr1 enzyme (0.28 mg/mL in 50 mM Tris HCl, pH: 7.4, 10% glycerol, 500 mM NaCl, 5 mM DTT, 1 mM EDTA), 1 μL of debranching 10x DBr1 buffer (500 mM Tris-HCl pH: 7, 20 mM DTT, 5 mM MnCl₂), and diluted to 10 μL volume with autoclaved H₂O. The yDBr1 enzyme was generously donated by Dr. Schwer (Cornell University) stored at -78°C, and thawed to 0°C prior to use. The reactions were performed for 20 min at 20°C and each reaction was quenched at various time points with the addition of 10 μL of the stop solution (prior to analysis by PAGE.

The reactions were analyzed by autoradiography and the cleavage reaction was determined quantitatively by densiometric analysis (UN-SCAN-IT™ software).

7.5.5 Inhibition of the yDBr1 enzyme

A control bNA sequence was 5'-end radiolabeled, purified, desalted and initially treated with the yDBr1 enzyme to determine the optimized reaction conditions for the enzymatic debranching of the substrate (section 7.5.4).

The inhibition of the yDBr1 enzyme with bNAs was performed by preparing 10 μL stock solutions of 'cold' (non-labeled) bNAs inhibitors (500 - 0.05 μM) in 10x DBr1 buffer. The inhibitors (1 μL) were added in increasing concentrations (50 - 0.005 μM) to the substrate (0.5 – 2 pmol) in the presence of the 1 μL of yDBr1 enzyme (0.28 mg/mL), with 1 μL of 10x yDBr1 buffer and diluted to a 10 μL reaction volume with autoclaved water. The reactions were performed for 5 min at 20°C and the reaction was stopped with 10 μL of "stop solution" prior to analysis by PAGE.

The reactions were analyzed by autoradiography and the cleavage reaction was determined quantitatively by densiometric analysis (UN-SCAN-IT™ software).

7.5.6 Serum Stability of Oligonucleotides

Pure oligothymidylate and oligoadenylate (0.7 ODs) were evaporated to dryness. To each dried oligonucleotide sample was added 300 μL of 10% FBS (0.2 μM, in 10% fetal bovine serum diluted with eagle's medium) and incubated at 37°C up to 24 h incubation. For each time point, a 50 μL aliquot of the reaction mixture was removed and frozen on dry ice for 10 min to stop the reaction, followed by evaporation in a Speed-Vac® concentrator. Aliquots were re-dissolved in deionized formamide and analyzed by denaturing 24% PAGE, and visualized by Stains-All® dye.

7.6 MONOMER PREPARATION

7.6.1 Specific Reaction Procedures and Product Characterization

2',5' and 3',5'-di-(monomethoxytrityl) ribouridine (2.2 and 2.3). The starting reagents, ribouridine (2 g, 8.19 mmol), MMT-Cl (5.06 g, 16.4 mmol), and AgNO$_3$ (2.78 g, 16.4 mmol) were dried under high vacuum overnight prior to the start of the reaction. The reaction flask was flushed with N$_2$ and stirred to a slurry mixture with anhydrous THF, (250 mL) and pyr (7 mL). The reaction mixture was stirred to complete reaction after 24

h with N_2 at room temperature, (22°C) and the reaction progress was monitored to completion by TLC with R_f for **2.2**: (5% MeOH in $CHCl_3$) 0.44 and R_f for **2.3**: (5% MeOH in $CHCl_3$) 0.33. The crude reaction mixture was filtered to remove the AgCl precipitate, and the filtrate was collected and diluted with 140 mL of CH_2Cl_2. The organic solution was quenched and washed successively with 140 mL of 5% $NaHCO_3$, water and dried over $MgSO_4$ prior to solvent evaporation. The crude product was purified by silica gel column chromatography with an eluent system based on 1 to 3% MeOH in CH_2Cl_2. The ditritylated regioisomers were isolated as white foams in an overall 85% yield, with **2.2**, (3.14 g, 50%) and **2.3**, (2.24 g, 35%).

^{1}H NMR (**2.2**, 500 MHz, DMSO-d_6) δ : 11.4 (1H, s, NH), 7.33-6.43 (28H, m, MMT), 7.32 (1H, d, J = 8.5 Hz, H6), 6.10 (1H, d, J = 7.5 Hz, H1'), 5.13 (1H, d, J = 8.4 Hz, H5), 4.80 (d, J = 5.5 Hz, OH), 4.28 (1H, t, J = 5 Hz, H2'), 3.88 (1H, s, H4'), 3.77 (3H, s, OCH$_3$), 3.72 (3H, s OCH$_3$), 2.99 (1H, dd, J = 7, 2.5 Hz., H5''), 2.93 (1H, dd, J = 4, 2.5 Hz, H3'), 2.93 (1H, dd, J = 7, 2.5 H5'). ESI-MS Calcd for $C_{49}H_{44}N_2O_8Na$: 812, found: 812 .

^{1}H NMR (**2.3**, 500 MHz, DMSO-d_6) δ : 11.4 (1H, d, J = 4 Hz, NH), 7.47 (1H, d, J = 7 Hz, H6), 7.33-6.43 (28H, m, MMT), 6.01 (1H, d, J = 7 Hz, H1'), 5.94 (d, J = 6.5 Hz, OH), 5.46 (1H, J = 1.5, 7.5 Hz, dd, H5), 4.09 (1H, dd, J = 2.5, 6.5 Hz, H2'), 3.97 (1H, d, J = 3.5, H3'), 3.73 (3H, s, OCH$_3$), 3.68 (3H, s, OCH$_3$), 3.02 (1H, s, H4'), 2.77 (1H, dd, J = 4.5, 9.5 Hz., H5'), 2.56 (1H, dd, J = 4.5, 10 Hz., H5''), ESI-MS Calcd for $C_{49}H_{44}N_2O_8Na$: 812, found: 812.

5'-(monomethoxytrityl) 2' and 3'-O-levulinyl ribouridine (2.7 and 2.8). In separate pre-dried round bottom flasks, the starting material, **2.4** (0.5 g, 0.968 mmol), 2-chloro-NMP iodide (0.5 g, 2mmol) (labeled solution A) and separately, LvOH (0.169 mL, 1.45 mmol), DABCO (0.52 g, 4.64 mmol) (labeled solution B) were dried overnight under high vacuum prior to initiating the reaction. Solution A was prepared by adding anhydrous dioxane (98.4 mL), MeCN (23.6 mL) and this mixture was stirred to a white slurry. Solution B was prepared with dry MeCN (50 mL) and this was transferred to solution A and the reaction continued to completion under N_2 for 15 min at 0°C. The reaction was monitored by TLC which generated the more non-polar side-product, **2.9**,

192

with R_f (5% MeOH in $CHCl_3$) 0.38 and the desired regioisomers, **2.7**, with R_f (5% MeOH in $CHCl_3$) 0.23 and **2.8**, with R_f(5% MeOH in $CHCl_3$) 0.21. The crude reaction mixture was evaporated to a viscous oil and dried with high vacuum prior to purification by silica gel column chromatography. The crude material was purified with an eluent gradient of 1 - 2% MeOH in $CHCl_3$ and **2.7**, **2.8** were recovered as an inseparable mixture of isomers (400 mg, 60%) isolated from the di-levulinated product, **2.9** (90 mg, 10%).

^1H NMR (**2.7**, 500 MHz, DMSO-d_6) δ : 11.41 (1H, d, J = 13.5 Hz, NH), 7.39-6.86 (14H, m, MMT), 7.24 (1H, d, J = 7.6 Hz, H6), 5.78 (1H, d, J = 6 Hz, 3'OH), 5.74 (1H, t, J = 6 Hz, H1'), 5.43 (1H, dd, J = 2, 7.5 Hz, H5), 5.11 (1H, t, J = 5 Hz, H3'), 4.37 (1H, dd J = 6, 10 Hz, H2'), 4.08 (1H, dd, J = 4, 7.5 Hz, H4'), 3.73 (3H, s, OCH_3), 3.35 (2H, m H5' and H5''), 2.71 (2H, t, J = 7 Hz, Lv CH_2), 2.53 (2H, t, J = 13 Hz, Lv CH_2), 2.09 (3H, s, Lv CH_3). ESI-MS: Calcd. for $C_{34}H_{34}N_2O_9Na$: 637.8, found : 637.1.

^1H NMR (**2.8**, 500 MHz, DMSO-d_6) δ : 11.39 (1H, d, J = 13.5 Hz, NH), 7.39-6.86 (14H, m, MMT), 7.24 (1H, d, J = 7.6 Hz, H6), 5.87 (1H, t, J = 4.5 Hz, H1'), 5.53 (1H, d, J = 6.5 Hz, 2'OH), 5.39 (1H, dd, J = 2, 7.5 Hz, H5), 5.22 (1H, t, J = 5 Hz, H3'), 4.32 (1H, dd, J = 6, 10 Hz, H2'), 3.96 (1H, d, J = 4 Hz, H4'), 3.73 (3H, s, OCH_3), 3.35 (2H, m H5' and H5'') 2.71 (2H, t, J = 7 Hz, Lv CH_2) 2.53 (2H, t, J = 13 Hz, Lv CH_2) 2.09 (3H, s, Lv CH_3). ESI-MS: Calcd. for $C_{34}H_{34}N_2O_9Na$: 637.8, found : 637.1.

5'-(monomethoxytrityl)-2'and 3'-O-levulinyl ribouridine phosphoramidite (2.10 and 2.11). The starting material **2.7** and **2.8**, (80 mg, 0.122 mmol) was dried under high vacuum overnight prior to the start of reaction. The phosphitylation reagent was prepared *in-situ* by reaction of 1-H tet (2.5 mg, 0.032 mmol) with P(OCEt)[(N(*i*-Pr)$_2$]$_2$ (40 µL, 0.127 mmol) and this was added to the starting material in anhydrous MeCN (2 mL). The reaction was completed after 5 h at room temperature (22°C) with N_2 as was confirmed by TLC with R_f (2:1 EtOAc:Hex with 0.5% TEA) 0.24 and by R_f (2:1 EtOAc:Hex with 0.5% TEA) 0.36, for the corresponding 2' and 3'-phosphoramidite diastereomers, respectively. The crude reaction mixture was diluted with 6 mL of EtOAc, quenched and washed with 2 mL of saturated $NaHCO_3$. The upper organic layer was filtered over $MgSO_4$ and the solvent evaporated to dryness. The crude was purified by silica gel

column chromatography with the eluent gradient of 2 to 4:1 EtOAc:Hex with 5% TEA, and the purified phosphoramidite regioisomers diastereomers, **2.11**, was collected as a white crystalline solid in yields of 58 mg (58%) while **2.10**, was also collected as a white solid in yields of 22 mg (22%).

^{1}H NMR (**2.10**, CDCl$_3$, 500 MHz) δ : 9.01 (1H, bs, NH), 7.76 (1H, d, J= 8 Hz, H6), 6.08 (1H, d, J= 5 Hz, H1'), 5.28 (1H, m H2'), 5.28 (1H, m, H5), 4.53 (1H, ddd, J = 4.5, 5, 5.5 Hz, H3'), 4.15 (1H, t, J=2.5 Hz, H4'), 3.69 (2H, m, CH$_2$), 3.54 (2H, m, CH$_2$), 3.42 (2H, m, H5' and H5"), 2.70 (2H, m, Lv CH$_2$), 2.53 (2H, m, Lv CH$_2$), 2.11 (3H, s, CH$_3$), 1.11 (12H, m, i-Pr CH$_3$); ^{31}P NMR (CDCl$_3$, 80 MHz ^{1}H decoupled 200 MHz) δ : 151.5 and 151.4; ESI MS Calcd for C$_{43}$H$_{51}$N$_4$O$_{10}$PNa: 837.3, found: 837.2.

^{1}H NMR (**2.11**, 500 MHz, CDCl$_3$) δ : 8.40 (1H, s, NH), 7.76 (1H, d, J= 8 Hz, H6), 6.05 (1H, d, J= 4.5 Hz, H1'), 5.33 (1H, t, J= 5 Hz, H3'), 5.23 (1H, d, J= 8 Hz, H5), 4.55 (1H, m, H2') 4.15 (1H, m, H4') 3.80 (2H, m, CH$_2$), 3.53 (2H, s, CH$_2$), 3.43 (2H, m, H5'H5"), 2.72 (2H, m, Lv CH$_2$) 2.56 (2H, m, Lv CH$_2$) 2.10 (3H, s, CH$_3$) 1.10 and 1.07 (12H, m, i-Pr CH$_3$); ^{31}P NMR (CDCl$_3$, 80 MHz, ^{1}H decoupled 200 MHz) δ : 151.5 and 150.7; ESI MS: calcd for C$_{43}$H$_{51}$N$_4$O$_{10}$PNa: 837.3, found: 837.1.

5'-O-levulinyl ribouridine (4.7). The reagents, LvOH (8.2 mL, 49 mmol) and 1 M DCC in CH$_2$Cl$_2$ (6.6 mL, 40 mmol) were stirred in anhydrous Et$_2$O (260 mL) to convert LvOH to the anhydride. The reaction was continued to completion for 5 h at room temperature conditions (22°C) and under N$_2$ atmosphere. The DCU byproduct was filtered from the solution, and the filtrate was evaporated to a viscous oil. In a separate flask, the starting material ribouridine (2 g, 8.19 mmol) was vacuum dried overnight and flushed with N$_2$. This was suspended in anhydrous dioxane (44 mL) with the lipase enzyme, Novozyme® 435 (3 g) and a solution of the crude Lv$_2$O in dioxane (146 mL). The chemoenzymatic reaction was completed overnight (12 – 16 h) at ambient temperatures (22°C) with N$_2$ and the product, **4.7** was confirmed by TLC with R$_f$: (15% MeOH in CHCl$_3$) 0.33. The enzyme was filtered (recycled and washed with MeOH for additional use) and the filtrate was evaporated to dryness prior to purification by silica gel column chromatography. The product, **4.7** was purified with 10% MeOH in CHCl$_3$ and recovered as a white crystalline solid in yields of 2.33 g (80%).

194

^{1}H NMR (**4.7**, 500 MHz, DMSO-d_6) δ : 11.36 (1H bs, NH), 7.64 (1H, d, J = 8 Hz, H6), 5.66 (1H, d, J = 8 Hz, H5), 5.74 (d, J = 5 Hz, H1'), 5.48 (1H, d, J = 4.5 Hz, 2'OH), 5.29 (1H, d, J = 5 Hz, 3'OH), 4.24 (1H, d, J = 11.5 Hz, H5'), 4.14 (1H, d, J = 5.5Hz, H5"), 4.12 (d, J = 5.5 Hz, H2'), 3.97 (1H, s H3'), 3.33 (2H, s, Lv CH$_2$), 3.25 (1H, d, J = 5 Hz, H4'), 2.72 (2H, s, Lv CH$_2$), 2.10 (1H, s, CH$_3$); ESI MS Calcd for $C_{14}H_{18}N_2O_8Na$ 365.3, found: 365.1.

4,6-*O*-(di-tert-butylsilanediyl)-D-glucal (5.3). The starting material, **5.1**, (8 g, 29.38 mmol) was vacuum dried overnight prior to initiating the reaction the following day. At ambient temperatures (22°C) and with N$_2$, MeOH (40 mL) was added to dissolve the starting material and the reaction was initiated with the addition of a freshly prepared solution of 2 M NaOMe in MeOH (0.5 mL, 0.13 M). The reaction was completed after 1.5 h, as was indicated by TLC, with R$_f$: (9: 1 CH$_2$Cl$_2$: MeOH) 0.25. The crude reaction mixture was concentrated *in-vacuo* to a crude oil and purified by silica gel column chromatography. The crude was eluted with 9:1 v/v CH$_2$Cl$_2$:MeOH and the purified product **5.2**, was collected in yields of 4.30 g (99%) as a white crystalline solid after drying with high vacuum.

D-glucal, **5.2**, (4.30 g, 29.42 mmol) was flushed with a constant flow of N$_2$ and at -40°C was added anhydrous DMF (145 mL, 2.87 mol) and stirred to complete solution. The reaction was initiated with the slow (dropwise for 5 min) addition of *t*-Bu$_2$Si(OTf)$_2$, (15 mL, 40 mmol) and the reaction was completed after 45 min by confirmation with TLC, R$_f$: (5 : 1 Hex : EtOAc) 0.33. The crude reaction mixture was quenched with anhydrous pyr (3.6 mL, 44 mmol) and stirred for an additional 15 min. The crude product was extracted in Et$_2$O (600 mL), and quenched, washed with saturated NaHCO$_3$ (150 mL) and H$_2$O (150 mL). The organic solution was dried with MgSO$_4$ and the solvent evaporated prior to silica gel column chromatography. The purified product was collected as a white crystalline solid in yields of 7.4 g (88%).

^{1}H NMR (**5.3**, 400 MHz, CDCl$_3$) δ : 6.25 (1H, d, J = 6 Hz, H1), 4.75 (1H, d, J = 6.4 Hz, H2), 4.29 (1H, d, J = 6.8 Hz, H3), 4.17 (1H, dd, J = 4.8, 10.4 Hz, H6), 3.95 (1H, t, J = 10.4 Hz 1H, H6'), 3.90 (1H, d, J = 10.4 Hz, H4), 3.84 (1H, dd, J = 4.8, 10.4 Hz, H5), 2.36 (1H, s, OH), 1.07 (9H, s, t-Bu Me), 0.99 (9H, s, t-Bu Me).

195

^{13}C NMR (**5.3**, 100 MHz, ^{1}H decoupled 400 MHz, CDCl$_3$) δ : 144 (C1), 103(C2), 77.4 (C4), 72.6 (C5), 70.1(C3), 66.1 (C6), 27.9 (t-Bu Me), 27.4 (t-Bu Me), 23.25 (t-Bu), 20.27 (t-Bu); ESI-MS Calcd. for C$_{14}$H$_{26}$O$_4$Si: 286.4, found: 286.1.

3-acetyl-4,6-*O*-(di-tert-butyl-silanediyl)-D-glucal (5.4). The starting material **5.3** (1.86 g, 6.51 mmol) and DMAP catalyst (41 mg, 0.33 mmol) were dried overnight under high vacuum prior to reaction. The reaction was initiated at room temperature (22°C) with N$_2$ by the addition of anhydrous pyr (5 mL) and Ac$_2$O (1.25 mL, 13.2 mmol) until complete reaction, (1 h) was indicated by TLC with R$_f$: (5:1 v/v Hex:EtOAc) 0.65. The crude product was extracted in Et$_2$O (100 mL) and washed with H$_2$O (35 mL). The organic solution was dried with MgSO$_4$ and the solvent was evaporated prior to purification by silica gel column chromatography. The product was purified with eluent 9 to 5:1 Hex: EtOAc and collected as a viscous oil in yields of 1.84 g (88%).

^{1}H NMR (**5.4**, 400 MHz, CDCl$_3$) δ : 6.31 (1H, dd, *J*= 1.6, 6.4 Hz, H1), 5.37 (1H, dt, *J*= 2, 7.6 Hz, H3), 4.72 (1H, dd, *J* = 2, 6.2 Hz H2), 4.19 (1H, dd, J = 4.8, 10 Hz, H6), 4.14 (1H, m, H4), 3.98 (1H, t, J = 10.4Hz, H6'), 3.92 (1H, dd, *J* = 4, 7.8 Hz, H5), 2.11 (3H, s, OAc), 1.06 (9H, s, t-Bu Me), 0.99 (9H, s, t-Bu Me).

^{13}C NMR (**5.4**, 100 MHz, ^{1}H decoupled 400 MHz, CDCl$_3$) δ : 171.1 (OAc), 145 (C1), 100.8 (C2), 73.93 (C4), 72.57 (C3), 73.22 (C5), 66.07 (C6), 27.77 (t-Bu Me), 27.25 (t-Bu Me) 23.11 (t-Bu), 21.62 (t-Bu), 20.26 (OAc); ESI-MS Calcd. for C$_{12}$H$_{19}$O$_5$Si: 271.3, found : 271.1.

1-(2,3-Dideoxy-*β*-D-erythro-hex-2-enopyranosyl)thymine (5.12). The starting material, **5.10**, (0.7 g, 2.07 mmol) was vacuum dried overnight prior to reaction. Under N$_2$ and at 22°C, the starting material was dissolved in 10 mL of anhydrous MeOH followed by the dropwise addition of 0.1 M NaOMe:MeOH (7 mL). The reaction was proceeded for 1 h, and the product was confirmed by TLC with R$_f$: (9:1 v/v CH$_2$Cl$_2$:MeOH) 0.2. The crude reaction mixture was concentrated to a white slurry and purified by silica gel column chromatography with eluent system 9:1 v/v CH$_2$Cl$_2$:MeOH. The purified product was collected as a white foam in yields of 0.485 g (93%).

196

^{1}H NMR (**5.12**, 400 MHz, DMSO-d_6) δ : 11.34 (s, NH), 7.16 (1H, s, H6), 6.23 (1H, s, H1'), 6.08 (1H, d, J = 11Hz, H3'), 5.64 (1H, d, J = 10.4 Hz, H2'), 5.27 (1H, s, OH), 4.71 (1H, s, OH), 4.03 (1H, d, J = 6.8 Hz, H4'), 3.67 (1H, d, J = 11 Hz, H6'), 3.46 (1H, m, H6''), 3.30 (1H, s, H5'), 1.76 (1H, s, H7).

^{13}C NMR (**5.12**, 125.7 MHz, ^{1}H decoupled at 500 MHz, DMSO-d_6) δ : 163.8 (CO), 150.5 (CO), 137.3 (C3'), 136.5 (C6), 124.9 (C2'), 109.9 (C5), 81.06 (C5'), 77.99 (C1'), 61.24 (C4'), 60 (C6'), 11.99 (C7); ESI-MS Calcd for $C_{11}H_{14}N_2O_5$: 254.3, found: 254.2.

3-acetyl-1,5-anhydro-2-deoxy-1,2-C-methylene 4,6-O-(di-t-butyl-silanediyl)-D-glucal (5.16). The starting material, **5.3**, (4.60 g, 16.06 mmol) was dried overnight under vacuum prior to reaction. The starting material was dissolved and stirred to solution with anhydrous Et_2O (56 mL) under a flow of N_2 while on ice (0°C). The reaction was initiated with dropwise addition of 1 M $ZnEt_2$ in Hex (11 mL, 96.5 mmol) and CH_2I_2 (4 mL, 49 mmol). The reaction was completed after 4 h and this was confirmed by TLC with R_f (5:1 v/v Hex:EtOAc) 0.27. The crude reaction mixture was quenched with saturated NH_4Cl (130 mL) and the product extracted with Et_2O, (2 x 200 mL). The organic solution was washed with saturated $NaHCO_3$ and brine (2 x 130 mL) and dried over $MgSO_4$ prior to evaporation and purification of the crude by silica gel column chromatography. The crude product was purified by silica gel flash chromatography with eluent 5:1 Hex:EtOAc and collected as a white solid in yields of 4.42 g (92%).

This product, **5.15**, (4.42 g, 14.71 mmol) with DMAP (91 mg, 0.74 mmol) were dried overnight under vacuum prior to reaction. At room temperature conditions (22°C), and with N_2 atmosphere, the reagents were dissolved in pyr (11 mL) and reacted with Ac_2O (3 mL, 30 mmol) for 1 h, until TLC indicated complete conversion to the product with R_f (5:1 Hex:EtOAc) 0.6. The crude reaction mixture was diluted with Et_2O (220 mL) and quenched, washed with H_2O (2 x 70 mL). The organic solution was dried with $MgSO_4$ and evaporated to a viscous oil prior to silica gel column chromatography. The crude product was purified with eluent gradient of 9 to 5:1 Hex:EtOAc, and collected as a viscous oil in yields of 4.6 g (92%).

^{1}H NMR (**5.16**, 400 MHz, CDCl₃) δ : 5.18 (1H, t, J = 4.8 Hz, H3), 4.09 (1H, dd, J = 6.4, 10 Hz, H6), 4.09 (1H, dd, J = 6.4, 10 Hz, H2), 3.69 (1H, t, J = 6.8 Hz, H4), 3.65 (1H, t, J

= 6.4 Hz, H6'), 3.42 (1H, m, H5), 2.13 (3H, s, OAc), 1.58 (1H, m, H1), 1.01 (9H, s, t-Bu Me), 0.98 (9H, s, t-Bu Me), 0.71 (2H, m, CH$_2$).

^{13}C NMR (**5.16**, 100MHz, ^{1}H decoupled 400 MHz, CDCl$_3$) δ : 75.93 (C4), 75.11 (C3), 73.5 (C5), 66.2 (C6), 55.29 (C2), 27.95 (t-Bu Me), 27.51 (t-Bu Me), 23.23 (t-Bu), 21.87 (t-Bu), 16.71 (C1), 13.19 (CH$_2$); ESI-MS Calcd for C$_{16}$H$_{28}$O$_5$SiNa : 342.5, found: 343.3.

1-[(3S,4S)-3,4-epoxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-β-oxepanyl] thymine (6.1).
The starting materials, **5.17** (200 mg, 0.49 mmol) and *m*CPBA (88 mg, 0.511mmol) were dried overnight under high vacuum prior to reaction. The reagents were flushed with N$_2$, and at room temperature (22°C), anhydrous MeCN (3.5 mL), CH$_2$Cl$_2$ (2.0 mL) were added successively, and the mixture was stirred to solution. The reaction was completed after 24 h and at 40°C. The complete conversion of **5.17** to the product was confirmed by TLC with R$_f$: (2 : 1 Hex : EtOAc) 0.19. The reaction mixture was diluted with CH$_2$Cl$_2$ (10 mL) and this was quenched, washed with saturated NaHCO$_3$ (2 x 10 mL) and with H$_2$O (10 mL).[261] The organic solution was dried with MgSO$_4$ and the solvent was evaporated prior to silica gel column chromatography. The crude product, **6.1** was purified with 2:1 to 1:2 Hex:EtOAc which yielded a white crystalline solid, 100 mg (50%).

^{1}H NMR (**6.1**, 400 MHz, CDCl$_3$) δ : 9.50 (1H, s, 1NH), 7.03 (1H, s, H6), 5.76 (1H, dd, J = 1.6, 11 Hz, H1'), 4.22 (1H, d, J= 9 Hz, H5'), 4.02 (1H, dd, J= 6.4, 11.2 Hz, H6'), 3.70 (2H, m, H7',H7"), 3.37 (1H, d, J = 4.8 Hz, H4'), 3.25 (1H, dd, J = 6.4, 10.6 Hz, H3'), 2.46 (1H, ddd, J= 1.6, 4.8, 15 Hz, H2'), 2.18 (1H, dd, J= 10.8, 15 Hz, H2"), 1.85 (3H, s, H7), 0.99 (9H, s, t-Bu Me), 0.95 (9H, s, t-Bu Me).

^{13}C NMR (**6.1**, 125 MHz, ^{1}H decoupled 500 MHz, CDCl$_3$) δ : 171.7(C2), 163.1 (C4), 134.3 (C6), 110.34 (C5), 80.85 (C1'), 74.6 (C5'), 71.9(C7'), 65.1 (C6'), 59.6 (C4'), 50.66 (C3'), 33.38 (C2'), 26.34 (t-Bu Me), 25.99 (t-Bu Me), 22.74 (t-Bu),11.52 (C7), 19.13 (t-Bu) ; ESI-MS Calcd for C$_{20}$H$_{32}$N$_2$O$_6$Si: 424.6, found: 425.3.

1-[(3S,4S)-3,4-dihydroxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-β-oxepanyl]thymine (6.2). The starting materials, **5.17** (125 mg, 0.3 mmol), OsO$_4$ (4 mg, 0.02 mmol) and NMO (150 mg, 1.25mmol) were dried overnight under high vacuum prior to reaction. With N$_2$ and at room temperature (22°C), the reagents were dissolved with 5:1 v/v ace:

H₂O (5 mL), and stirred to complete reaction for 5 h. The complete conversion of **5.17** to the product was confirmed by TLC, with R_f: (2:1 v/v EtOAc:Hex), 0.25. The reaction was quenched with saturated $NaHSO_3$ (3.8 mL) for an additional 30 min at 22°C. The crude reaction mixture was diluted with $CHCl_3$ (40 mL) and the organic solution was treated with saturated NaCl (2 x 30mL) and extracted again with $CHCl_3$ (3 x 40 mL). The organic solution was dried with $MgSO_4$ and evaporated prior to silica gel column chromatography. The crude product was purified with 2:1 EtOAc:Hex and was collected as a white crystalline solid in yields of 65 mg (50%).

^{1}H NMR (**6.2**, 400 MHz, $CDCl_3$) δ : 9.34 (1H, s, 1NH), 7.11 (1H, s, H6), 6.03 (1H, t, J = 6.4 Hz, H1'), 4.20 (1H, s, H3'), 4.06 (1H, m, H5'), 3.96(2H, m, H7'H7"), 3.78 (1H, m, H6'), 3.71 (1H, s, 3'OH), 3.50 (1H, m, H4'), 2.90 (1H, s, 4'OH), 2.44 (1H, ddd, J = 6.4, 8.8, 15.3 Hz, H2'), 2.12 (1H, m, H2"), 1.85 (3H s, H7), 0.99 (9H, s, t-Bu Me), 0.98 (9H, s, t-Bu Me).

^{13}C NMR (**6.2**, 125.7MHz, ^{1}H decoupled 500 MHz, $CDCl_3$) δ : 162.8 (C2), 149.0 (C4), 134.93 (C6), 110.3 (C5), 81.26 (C1'), 74.29 (C5'), 66.93 (C3'), 65.48 (C6'), 62.90 (C7'), 53.37 (C4'), 36.22 (C2'), 26.05 (t-Bu Me), 21.62 (t-Bu Me), 19.04 (t-Bu), 18.92 (t-Bu), 11.55 (C7); ESI-MS Calcd for $C_{20}H_{34}N_2O_7SiNa$: 465.6, found: 465.2.

1-[(3S,4S)-3,4-*C*-methylene-(5S,6R)-5,7-di-tert-butylsilanediyl)-*β*-oxepanyl] thymine (6.3). The starting material **5.17**, (120 mg, 0.294 mmol) was dried overnight under high vacuum prior to the start of the reaction. Under N_2 and at low temperature (0°C), **5.17**, was diluted with anhydrous Et_2O (1.2 mL) and the reaction was initiated with the slow addition (dropwise, 5 min) of 1 M $ZnEt_2$ in Hex (0.3 mL, 2 mmol) and CH_2I_2 (80 μL, 0.9 mmol). The reaction progress after 24 h and with a gradual temperature increase (0 - 30°C) was monitored by TLC which indicated 40% conversion of the starting material to a product with R_f: (5:1 v/v Hex : EtOAc) 0.14. The crude product was extracted in Et_2O (5 mL), treated with saturated NH_4Cl (2.5 mL) and extracted, washed again with Et_2O (4 mL) and H_2O (2.5 mL), saturated NaCl (2.5 mL). The organic solution was dried with $MgSO_4$ and the solvent was evaporated prior to silica gel column chromatography. The crude product was purified with 5:1 to 2:1 Hex:EtOAc, as a white crystalline solid in yields of 35 mg (30%).

^{1}H NMR (**6.3**, 400 MHz, CDCl$_3$) δ : 7.12 (1H, s, H6), 5.85 (1H, d, J = 12 Hz, H4'), 5.73 (1H, dd, J = 1.6, 10 Hz, H1'), 5.62 (1H, m, H3'), 4.53 (1H, dd, J = 2.4, 9.4 Hz, H5'), 3.99 (1H, dd, J = 4.8, 10.4 Hz, H7'), 3.79 (1H, dd, J = 10.4,14.8 Hz, H7"), 3.57 (1H ,dd, J = 4.8, 9.4 Hz, H6'), 3.28 (2H, s, CH$_2$), 2.52 (1H, ddd, J = 2.8, 13.4, 14.6 Hz, H2'), 2.31 (1H, ddd, J = 8.8, 14.6 Hz, H2"), 1.88 (1H, d, H7), 0.98 (1H, s, t-Bu Me), 0.92 (1H, s, t-Bu Me).

^{13}C NMR (**6.3**, 125.7MHz, 500MHz ^{1}H decoupled, CDCl$_3$) δ : 162.4 (C2), 149.6 (C4), 138.8 (C4'), 132.0 (C6), 120.97 (C3'), 109.2 (C5), 83.31 (C1'), 77.14 (C6'), 76.26(C5'), 65.5 (C7'), 35 (C2'), 26.99 (>C), 26.38 (t-Bu Me), 25.97 (t-Bu Me), 21.52 (t-Bu), 18.96 (t-Bu),12.31 (C7); EI-MS Calcd for C$_{21}$H$_{34}$N$_2$O$_5$Si: 422.6, found: 422.

1-[2,3-dideoxy-(4R)-4-hydroxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-β-oxepanyl]

thymine (6.4). The starting material, **5.17**, (25 mg, 0.06 mmol) was dried overnight under high vacuum prior to initiating the reaction the following day. With N$_2$ and at low temperature (0°C), the starting material was dissolved with anhydrous THF (0.3 mL) and the reaction was initiated with the dropwise addition of BH$_3$-THF (20 µL, 0.21 mmol) and stirred to completion for 20 h with a gradual temperature increase (0 - 22°C). The extent of the reaction was monitored by TLC with eluent system 2:1 v/v EtOAc:Hex which indicated a product with R$_f$: 0.30. The reaction was placed in an ice bath and treated with H$_2$O (0.5 mL), 3 M NaOH (0.15 mL) and 30% H$_2$O$_2$ (0.15 mL). This mixture was stirred for and additional 2 h, or until TLC indicated complete reaction, R$_f$: (2:1 EtOAc:Hex) 0.36. The reaction mixture was diluted with Et$_2$O (3 mL), treated with saturated NaCl (2 x 0.15 mL) and the crude product was extracted in Et$_2$O (3 x 3 mL). The organic solution was dried with MgSO$_4$ and evaporated prior to silica gel column chromatography. The crude product was purified by chromatography with eluent system, 2:1 EtOAc:Hex, and collected as a white crystalline solid in yields of 17 mg (67%).

^{1}H NMR (**6.4**, 500 MHz, CDCl$_3$) δ : 8.16 (1H, s, 1H), 7.09 (1H, s, H6), 5.80 (1H, t, J = 6.5 Hz, H1'), 4.05 (1H, dd, J = 5.5, 10.8 Hz, H7'), 3.77 (1H, t, J = 10.5 Hz, H7"), 3.61 (1H ,m, J = 5 Hz, H5'), 3.56 (1H, dt, J = 1.5, 10 Hz, H4'), 3.46 (1H, dt, J = 6.5, 8.8 Hz, H6'), 2.95 (1H, s, 4'OH), 2.10 (1H, m, H2') 1.94 (1H, m, H2"), 1.72 (2H, m, H3'H3"), 1.87 (1H, s, H7) 1 (9H, s, t-Bu Me), 0.93 (9H, s, t-Bu Me).

^{13}C NMR (**6.4**, 125.7MHz, 500MHz ^{1}H decoupled, CDCl$_3$) δ : 153.13 (C4), 134.80 (C6), 110.2 (C5), 82.19 (C1'), 81.26 (C4'), 74.20 (C5'), 65.45 (C6'), 65.45(C7'), 28.43 (C2'), 26.42 (t-Bu Me), 25.95 (t-Bu Me), 24.93 (C3'), 21.63 (t-Bu), 18.93 (t-Bu), 11.59 (C7); EI-MS Calcd for C$_{20}$H$_{34}$N$_2$O$_6$Si : 426.6, found : 427.3.

7.6.2 General Reaction Procedures and Product Characterization

General procedure for the glycosylation reaction. The starting material (4.7 mmol) was dried overnight under high vacuum. Similarly, in a separate flask, the base (thymine or N^6-benzoyl adenine) (24 mmol) and drying reagent (NH$_4$)$_2$SO$_4$ (2.5 mmol) were also dried under vacuum. Under a N$_2$ atmosphere and at ambient temperatures (22°C), 85 mL of dry MeCN was added to the flask containing the base and (NH$_4$)$_2$SO$_4$. Hexamethyldisilazane, HMDS (38 mmols) was added dropwise to the resulting suspension, and the reaction was refluxed for 3 - 4 h until the MeCN-soluble silylated base was formed. The solvent was evaporated and a solution of the starting material in 20 mL of dry MeCN was added and stirred at 22°C under N$_2$. The reaction was initiated with TMSOTf (1.6 mmol) and completed with reflux (90°C) until TLC indicated complete conversion to the produt. The crude reaction mixture was 'worked up' by diluting with EtOAc (150 mL) followed by quenching and washing the upper organic layer with 100 mL each of saturated NaHCO$_3$ and H$_2$O. The upper organic layer was dried over MgSO$_4$ and evaporated to dryness, and the residue was purified on a column of silica gel (eluent 4:1 to 2:1 v/v Hex : EtOAc).

1-(2,3-Dideoxy-5,6-di-O-tert-butylsilanediyl-$β$-D-erythro-hex-2-enopyranosyl) thymine (5.5). The product **5.5** was collected as its pure $β$-anomer in yields of 135 mg (35%) and as a white foam R$_f$: (2:1 Hex:EtOAc) 0.25.

^{1}H NMR (**5.5**, 500 MHz, CDCl$_3$) δ : 7.99 (1H, s, NH), 7.36 (1H, s, H6), 6.49 (1H, dd, J = 2.5, 6 Hz, H2'), 5.36 (1H, d, J = 2.5 Hz, H1'), 4.56 (1H, dd, J = 2, 6.5 Hz, H3'), 4.21 (1H, dd, J = 6.5, 10.5 Hz, H6"), 4.04 (1H, dd, J = 4.5, 10 Hz, H6'), 4.02 (1H, dd, J = 4.5, 6 Hz H5'), 3.96 (1H, dd, J = 4, 10 Hz, H4'), 1.93 (3H, s, H7), 0.98 (9H, s, t-Bu Me), 0.97 (9H, s, t-Bu Me).

^{13}C NMR (**5.5**, 125.7 MHz, 500 MHz ^{1}H decoupled, CDCl$_3$) δ : 55.70 (C1'), 147.49

(C2'), 99.88 (C3'), 74.02 (C4'), 74.15 (C5'), 65.71 (C6'), 163.7 (C2), 151.2 (C4), 111.4 (C5), 12.79 (C7), 136.15 (C6), 27.52 (t-Bu Me), 26.96 (t-Bu Me), 22.84 (t-Bu), 20.13 (t-Bu); ESI-MS Calcd. for $C_{19}H_{30}O_5N_2Si$: 394.5, found: 395.0.

1-(2,3-Dideoxy-5,6-di-*O*-tert-butylsilanediyl-*α*-D-erythro-hex-2-enopyranosyl) thymine (5.6). The product **5.6** was collected as its pure *α*-anomer in yields of 215 mg (55%) and as a white foam R_f: (2:1 Hex:EtOAc) 0.19.

^{1}H NMR (**5.6**, 500 MHz, CDCl$_3$) δ : 8.19 (1H, bs, NH), 7.36 (1H, s, H6), 6.72 (1H, dd, J = 1, 6 Hz H2'), 5.45 (1H, t, J = 5 Hz, H1'), 4.71 (1H, t, J = 6 Hz, H3'), 4.32 (1H, dd, J = 6, 10.5 Hz, H5'), 4.25 (1H, dd, J = 5, 10.5 Hz, H6'), 3.90 (1H, t, J = 10 Hz, H6"), 3.70 (1H, ddd, J = 5, 10.5,12.75 Hz, H4'), 1.89 (3H, s, H7), 1.03 (9H, s, t-Bu Me), 0.81 (9H, s, t-Bu Me); ESI-MS Calcd. for $C_{19}H_{30}O_5N_2Si$: 394.5, found: 395.0.

1-(2,3-dideoxy-4,6-di-*O*-acetyl-*β*-D-*erythro*-hex-2-enopyranosyl)thymine (5.10). The product **5.10** was collected as its pure *β*-anomer in yields of 765 mg (35%) and as a white foam R_f: (2:1 Hex:EtOAc) 0.29.

^{1}H NMR (**5.10**, 400 MHz, CDCl$_3$) δ : 8.16 (s, NH), 6.96 (1H, s, H5), 6.52 (1H, s, H1'), 6.16 (1H, d, H3'), 5.75 (1H, d, J = 10 Hz, H2'), 5.38 (1H, d, J = 6.5 Hz, H4'), 4.20 (2H, d, J = 3.5 Hz, H6'6"), 4.00 (1H, t, J = 4.5 Hz, H5'), 2.11 (3H, s, OAc), 2.08 (3H, s, OAc), 1.91 (3H, s).

^{13}C NMR (**5.10**, 125.7 MHz, 1H-decoupled 500 MHz, CDCl$_3$) δ : 169.3 (CO), 168.8 (CO), 161.7 (CO), 148.8 (CO), 131.2 (C3'), 134.1 (C6), 126.3 (C2'), 110.77 (C5), 77.19 (C1'), 73.77 (C5'), 62.89 (C4'), 61.53 (C6'), 19.83 (OAc), 19.73 (OAc), 11.42 (C7); ESI-MS Calcd. for $C_{15}H_{18}N_2O_7$: 338.9, found: 338.9.

1-(2,3-dideoxy-4,6-di-*O*-acetyl-*α*-D-*erythro*-hex-2-enopyranosyl)thymine (5.11). The product **5.11** was collected as its pure *α*-anomer in yields of 935 mg (40%) and as a white foam R_f : (2:1 Hex:EtOAc) 0.20.

^{1}H NMR (**5.11**, 400 MHz, CDCl$_3$) δ : 8.18 (s, NH), 7.24 (1H, s, H5), 6.40 (1H, s, H1'), 6.30 (1H, d, J = 10 Hz, H3'), 5.86 (1H, d, J = 10.5 Hz , H2'), 5.24 (1H, d, J = 3 Hz, H4'),

4.27 (1H, dd, $J = 12$ Hz, H6'), 4.16 (1H, dd, $J = 3$, 12.25 Hz, H6''), 4.01 (1H, m, $J = 4.5$ Hz, H5'), 2.12 (s, OAc), 2.09 (s, OAc), 1.93 (3H, s, 7); ESI-MS Calcd for $C_{15}H_{18}N_2O_7Na$: 361.9 found: 362.1.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-β-oxepinyl]thymine

(5.17). The product **5.17** was collected as its pure β-anomer and as a white foam (650 mg, 35%). R_f: (2:1 Hex:EtOAc) 0.22.

^{1}H NMR (**5.17**, 400 MHz, CDCl$_3$,) δ : 8.56 (bs, NH), 7.19 (1H, d, $J = 1$ Hz, H6), 5.91 (1H, ddd, $J = 2.4$, 2.4, 12 Hz, H4'), 5.74 (1H, dd, $J = 6$, 10 Hz, H1'), 5.61 (1H, m, H3'), 4.59 (1H, dd, $J = 2.4$, 8.8 Hz, H5'), 4.06 (1H, dd, $J = 8.4$, 10.5Hz H7'), 3.85 (1H, dd, $J = 8.4$, 10.5Hz H7''), 3.64 (1H, m, H6'), 2.59 (1H, m, H2'), 2.39 (1H, m, H2''), 1.94 (3H, d, $J = 1$Hz, H7), 1.05 (9H s, t-Bu Me), 0.99 (9H s, t-Bu Me).

^{13}C NMR (**5.17**, 100 MHz, ^{1}H decoupled 400 MHz CDCl$_3$) δ : 163.4 (CO), 149.8 (CO), 140 (C4'), 135.3, 122 (C3'), 111.36, 83.81 (C1'), 78.5 (C6'), 77 (C5'), 66.8 (C7'), 36.44 (C2'), 27.82 (t-Bu Me), 27.40 (t-Bu Me), 23.02 (t-Bu), 20.42 (t-Bu), 13.04; ESI-MS: Calcd for $C_{20}H_{32}O_5N_2Si$ 408.6; found 408.8.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-α-oxepinyl]thymine

(5.18). The product **5.18** was collected as a mixture of α and β-anomer and as a white foam (100 mg, 15%). R_f: (2:1 Hex:EtOAc) 0.20.

^{1}H NMR (**5.18**, 400 MHz, CDCl$_3$,) δ : 8.56 (bs, NH), 7.13 (1H, d, $J = 1$ Hz, H6), 5.99 (1H, ddd, $J = 2.4$, 2.4, 12 Hz, H4'), 5.75 (1H, dd, $J = 2$, 10 Hz, H1'), 5.68 (1H, m, H3'), 4.67 (1H, dd, $J = 2.4$, 8.5 Hz, H5'), 4.24 (1H, dd, $J = 8.4$, 10 Hz H6'), 4.03 (1H, dd, $J = 8.4$, 10 Hz H7'), 3.84 (1H, m, H7''), 2.91 (1H, m, H2'), 2.33 (1H, m, H2''), 1.88 (3H, d, $J = 1$Hz, H7), 0.98 (9H s, t-Bu Me), 0.91 (9H s, t-Bu Me). ESI-MS: Calcd for $C_{20}H_{32}O_5N_2Si$ 408.6; found 408.8.

1(R)-1-[(2,3,4-trideoxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-β-oxepinyl]-N^6-benzoyladenine (5.25b- β-anomer). The product **5.25b** was collected as its pure β-anomer in yields of 710 mg (30%) and as a white foam R_f : (2:1 EtOAc:Hex) 0.24.

^{1}H NMR (**5.25b**- β-anomer, 500 MHz, acetone-d_6) δ : 10.0 (1NH, s, H6), 8.68 (1H, s, H8), 8.47 (1H, s, H2), 8-7.2 (5H, m, ar), 6.02 (1H, d, J= 9 Hz H1'), 5.90 (1H, dd, J= 8, 12 Hz, H3'), 5.76 (1H, dd, J = 8.5, 11 Hz, H4'), 4.74 (1H, d, J = 9 Hz, H5'), 4.04 (1H, ddd, J= 14, 10, 6 Hz, H7'), 3.91 (1H, dd, J = 9.5, 4.5 Hz, H6'), 3.86 (1H, d, J = 18.5 Hz, H7''), 3.40 (1H, d, J = 14 Hz, H2'), 2.84 (1H, m, H2''), 1.07 (9H, s, t-Bu Me), 1.04 (9H, s, t-Bu Me).

^{13}C NMR (**5.25b**- β-anomer, acetone-d_6, 125.7 MHz, ^{1}H decoupled 500 MHz): δ 152.16, 141.65, 139.3 (C3'), 132.6, 131.4, 128.75, 128.5, 128.4, 127.8, 127.6, 123.2 (C4'), 84.29 (C1'), 77.84 (C6'), 77.31 (C5'), 66.58 (C7'), 35.30 (C2'), 27.11 (t-Bu-Me), 26.78 (t-Bu-Me), 22.46 (t-Bu), 19.86 (t-Bu); ESI-MS: Calcd. for $C_{27}H_{35}N_5O_4SiNa$: 544.7, found: 544.1.

1(R)-1-[(2,3,4-trideoxy-(5S,6R)-5,7-di-tert-butylsilanediyl)-α-oxepinyl]-N^6-benzoyladenine (5.25b- α-anomer). The product **5.25b** was collected as its pure α-anomer in yields of 375 mg (15%) and as a white foam R$_f$: (2:1 EtOAc:Hex) 0.18.

^{1}H NMR (**5.25b**- α-anomer, 500 MHz, acetone-d_6) δ : 9.95 (1NH, s, H6), 8.61 (1H, s, H8), 8.51 (1H, s, H2), 8.4-7.4 (5H, m, ar), 6.31 (1H, d, J= 10.75 Hz H1'), 5.89 (1H, dd, J = 8, 12 Hz, H3'), 5.89 (1H, dd, J = 8.5, 11 Hz, H4'), 5.14 (1H, d, J = 10.5 Hz, H6'), 4.66 (1H, dd, J = 12, 6 Hz, H5'), 4.04 (1H, d, J = 12 Hz, H2'), 3.91 (1H, dd, J = 14, 4.5 Hz, H7'), 3.78 (1H, dd, J = 14.5, 10.5 Hz, H7''), 2.77 (1H, m, H2''), 1.06 (9H, s, t-Bu Me), 1.04 (9H, s, t-Bu Me); ESI-MS: Calcd. for $C_{27}H_{35}N_5O_4SiNa$ 544.7; found 544.1.

General procedure for the desilylation reaction.

The starting material, (3.3 mmol) was dried overnight with vacuum. Under N_2 and at 0°C, the starting material was dissolved in 7 mL of THF. A solution of 1 M TBAF in dry THF (7 mL) was added with stirring over 5 min. The reaction mixture turned slightly cloudy as the deprotected nucleoside slowly began to precipitate from the solvent. The reaction was complete by TLC after 1 h, so the solvent was removed and the resulting viscous oil was purified by column chromatography (9:1 CH_2Cl_2:MeOH) to give a white foam.

1-(2,3-Dideoxy-β-D-erythro-hex-2-enopyranosyl)thymine (5.7). The purified product was collected as a white foam in yields of 0.210 g (93%) with R_f: (9:1 CH$_2$Cl$_2$:MeOH) 0.27.

^{1}H NMR (**5.7**, 400 MHz, DMSO-d_6) δ : 11.14 (1H, s, NH), 7.25 (1H, s, H6), 6.56 (1H, dd, J = 3, 7 Hz, H2'), 5.45 (1H, J = 6 Hz, d, 2'OH), 5.06 (1H, d, J = 8.4 Hz, H1'), 4.68 (1H, t, J = 5.6 Hz, 3'OH), 4.51 (1H, dd, J = 2, 6 Hz, H3'), 3.72 (1H, m, H5'), 3.76 (2H, m, H6' and H6"), 3.62 (1H, m, H4'), 1.78 (3H, s, H7).

^{13}C NMR (**5.7**, 125.7 MHz, 500 MHz ^{1}H decoupled, DMSO-d_6) δ : 141.3 (C2'), 94.17 (C3'), 74.63 (C5'), 59.63 (C4'), 54.64 (C6'), 49.81 (C1'); ESI-MS Calcd. for C$_{11}$H$_{14}$O$_5$N$_2$: 254.2, found: 243.4.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-hydroxymethyl)-β-oxepinyl]thymine (5.20). The purified product was collected as a white foam in yields of 0.854 g (90%) and with R_f: (9:1 CH$_2$Cl$_2$:MeOH) 0.37.

^{1}H NMR (**5.20**, 500 MHz, DMSO-d_6,) δ : 11.30 (1NH, s, H3), 7.57 (1H, s, H6), 5.74 (1H, t, J = 9Hz, H4'), 5.60 (1H, d, J = 10Hz, H1'), 5.55 (1H, d, J = 9 Hz, H3'), 4.05 (1H, d, J = 7.5Hz, H5'), 3.60 (1H, d, J = 11Hz H7'), 3.46 (1H, d, J = 11Hz, H7"), 3.36 (1H, t, J = 6.5 Hz, H6'), 3.14 (1H, s, 2'OH), 2.67 (1H, t, J = 12Hz, H2'), 2.32 (1H, dd, J = 6, 7.5 Hz, H2"), 1.76 (3H, s, H7), 1.55(1H, s, 3'OH).

^{13}C NMR (**5.20**, 125 MHz, ^{1}H decoupled 500 MHz, DMSO-d_6) δ : 164.22 (CO), 150.33 (CO), 138.03 (C3') , 136.31 (C6), 83.45 (C1'), 34.89 (C2'), 122.27 (C4'), 68.91 (C5'), 84.80 (C6'), 61.95 (C7'), 23.06 (C7), 108.93 (C5); ESI-MS: Calcd. for C$_{12}$H$_{16}$N$_2$O$_5$: 268.7; found: 269.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-hydroxymethyl)-β-oxepinyl]-N^6-benzoyladenine (5.27). The purified product was collected as a white foam in yields of 0.184 g (61%) with R_f: (9:1 CH$_2$Cl$_2$: MeOH) 0.33.

^{1}H NMR (**5.27**, 500 MHz, MeOH-d_4) δ : 8.71 (1H, s, H8), 8.58 (1H, s, H2), 8.08 (2H, d, ar.), 7.65 (1H, m, ar.), 7.56 (2H, m, ar), 6.09 (1H, dd, J = 9.5, 2.5 Hz, H1'), 5.92 (ddd, 1H, J = 2.5, 13 Hz, H3'), 5.76 (1H, m, H4'), 4.28 (1H, dd, J = 2, 9 Hz, H5'), 3.89 (1H, dd, J = 4.5, 9.5 Hz, H7'), 3.79 (1H, ddd, J = 2.5, 5, 9 Hz, H6'), 3.68 (1H, dd, J = 6, 11.5 Hz,

H7″), 3.17 (1H, m, H2′), 3.05 (1H, s, 2'OH), 2.90 (1H, ddd, J = 2, 7, 16.5 Hz, H2″), 1.65 (1H, s, 3'OH).

^{13}C NMR (<u>5.27</u>, 125.7 MHz, 500 MHz ^{1}H decoupled, MeOH-d_4) δ : 175.23 (CO), 152.0 (C8), 142.2 (C2), 137.6 (C3′), 132.7 (ar.), 128.6 (ar), 128.2 (ar), 122.3 (C4′), 84.90 (C1′), 84.29 (C6′), 69.92 (C5′), 62.83 (C7′), 35.23 (C2′); ESI-MS Calcd for $C_{19}H_{19}N_5O_4$: 402.6; found 404.

General procedure for the hydrogenation reaction.

The product from the desilylation reaction (0.565 mmol) and 152 mg of 10% Pd/C catalyst were dried overnight under vacuum. Dry MeOH (11.5 mL) was added to the evacuated flask and the dark suspension was stirred at room temperature (22°C). A balloon filled with H_2 was attached to the flask by piercing the septum with a needle. Small aliquots were periodically withdrawn, evaporated to dryness and the extent of the reaction was verified by ^{1}H NMR. After 4 h, the remaining reaction mixture was filtered and evaporated to dryness. The crude product was purified by flash silica gel column chromatography in 9:1 CH_2Cl_2:MeOH.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-hydroxymethyl)-β-oxepanyl]thymine (5.28). The extent of reaction could not be accurately monitored by TLC as the R_f : 0.21, values for the starting material and product were found to be identical in eluent system 9: 1 CH_2Cl_2:MeOH. The purified product was collected after chromatography in yields of 104 mg (70%) as a white foam.

^{1}H NMR (<u>5.28</u>, 500 MHz, MeOH-d_4) δ : 7.56 (1H, s, H6), 5.78 (1H, d, J = 9.5Hz, H1′), 3.77 (1H, s, H5′), 3.70 (1H, d, J = 11.5 Hz, H7′), 3.55 (1H, m, H6′), 3.55 (1H, m, H7″), 1.97 (1H, m, H2′), 1.90 (2H, m, H3'H3″), 1.90 (3H, s, H7), 1.90 (2H, m, 2'OH, 3'OH), 1.88 (2H, m, H4'H4″), 1.66 (1H, d, J = 6 Hz, H2″).

^{13}C NMR (<u>5.28</u>, 125 MHz, 500 MHz ^{1}H decoupled, MeOH-d_4) δ : 150.9 (CO), 137 (C6), 110 (CO), 95 (C5), 86.64 (C6′), 86.5 (C1′), 70.8 (C5′), 63.37 (C7′), 34.87 (C4′), 33.29 (C3′), 17.9 (C2′) 11.18 (C7); ESI-MS Calcd for $C_{12}H_{18}N_2O_5Na$: 279.3; found 293.1.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-hydroxymethyl)-β-oxepanyl]-N^6-benzoyladenine (5.29). The extent of reaction could not be accurately monitored by TLC as the R_f : 0.22, values for the starting material and product were found to be identical in eluent system 9:1 CH_2Cl_2:MeOH. The purified product was collected in yields of 149 mg (99%) as a white foam.

^{1}H NMR (**5.29**, 500 MHz, MeOH-d_4) δ : 9.81 (s, 1NH), 8.71 (1H, s, H8), 8.58 (1H, s, H2), 8.08 (2H, ar), 7.65 (1H, ar), 7.56 (2H, ar), 6.06 (1H, dd, J = 5, 5.5 Hz, H1'), 3.80 (1H, dd, J = 4, 8,25 Hz, H6'), 3.76 (1H, dd, J = 3, 6 Hz, H4'), 3.58 (1H, dd, J = 6.35, 12 Hz, H4''), 3.12 (1H, t, J = 8 Hz, H5'), 2.39 (2H, m, H2'H2''), 2.04 (2H, m, H3'H3''), 1.93(1H, m, H7'), 1.81 (1H, m, H7''), 1.67 (1H, s, 2'OH), 1.42 (1H, s, 3'OH).

^{13}C NMR (**5.29**, 125.7 MHz, ^{1}H decoupled 500 MHz, MeOH-d_4) δ : 151.9 (C8), 142 (C2), 132.7 (ar), 128.6 (ar), 128.3 (ar), 87.25 (C1') 70.68 (C5'), 63.54 (C4'), 52.90 (C6'), 35.08 (C2'), 33.76 (C3'), 18.0 (C7'); ESI-MS Calcd for $C_{19}H_{21}N_5O_4Na$: 406.4; found: 406.4.

General Procedure for the tritylation reaction.

The deprotected nucleoside (0.377 mmol) and MMT-Cl (0.44 mmol) were dried overnight under vacuum. Pyridine (1.5 mL) was added at 22°C under nitrogen. The reaction was stirred for 4 h until TLC (9:1 CH_2Cl_2:MeOH) indicated completion. The reaction was diluted with EtOAc (60 mL) and washed with saturated aqueous $NaHCO_3$ (2 × 60 mL). The organic layer was then dried over $MgSO_4$, concentrated and purified by silica gel chromatography with eluent system 20 to 9:1 CH_2Cl_2:MeOH.

5'-O-levulinyl 2' -monomethoxytrityl ribouridine (4.8). The product was purified and dried as a white foam in yields of 1.75 g (43%) and confirmed by TLC R_f: (5% MeOH in $CHCl_3$ with 0.5% TEA) 0.25.

^{1}H NMR (**4.8**, 400 MHz, DMSO-d_6) δ : 11.27 (1H, d, J = 3 Hz, NH), 7.49-6.77 (14H, m, MMT), 7.12 (1H, d, J = 8 Hz, H6), 5.90 (1H, d, J = 8.5 Hz, H1'), 5.42 (1H, dd, J = 3, 8 Hz, H5), 5.18 (1H, d, J = 7.5 Hz, 3'OH), 4.20 (1H, t, J = 7 Hz, H2'), 4 (1H, m, H5' and H5''), 3.87 (1H, dd, J = 5, 11.5 Hz, H4'), 3.71 (3H, s, OMe), 3.03 (1H, ddd, J = 1, 3.5, 4

Hz, H3'), 2.68 (2H, t, J = 7.5 Hz, Lv CH$_2$) 2.38 (2H, dd, J = 7.5, 8 Hz, Lv CH$_2$) 2.10 (3H, s, CH$_3$); ESI-MS, Calcd for C$_{34}$H$_{34}$N$_2$O$_9$Na: 637.6, found: 637.1.

5'-O-levulinyl 3'-monomethoxytrityl ribouridine (4.9). The product was purified and dried as a white foam in yields of 885 mg (21%) and confirmed by TLC R$_f$: (5% MeOH in CHCl$_3$ with 0.5% TEA) 0.15.

^{1}H NMR (**4.9**, 400 MHz, DMSO-d_6) δ : 11.34 (1H, s NH), 7.59-6.86 (14H, m, MMT), 7.24 (1H, d, J = 6.8Hz, H6), 5.91 (1H, t, J = 6.8 Hz, H1'), 5.91 (1H, t, J = 6.8 Hz, 2'OH), 5.62 (1H, d, J = 6.8 Hz, H5), 4.07 (1H, s, H3'), 3.92 (1H, dd, J = 6.4, 12 Hz, H2'), 3.73 (3H, s, OMe), 3.54 (1H, d, J = 12 Hz, H5"), 3.43 (1H, dd J = 5, 12 Hz, H5'), 3.26 (1H, s, H4') 2.62 (2H, t, J = 6.8 Hz, Lv CH$_2$) 2.33 (2H, t, J = 6 Hz, Lv CH$_2$) 2.05 (3H, s, CH$_3$); ESI-MS, Calcd for C$_{34}$H$_{34}$N$_2$O$_9$Na: 637.6, found: 637.0.

1–{5-hydroxy-6-[4-(methoxyphenyl)diphenyl]-2,3-dideoxy-β-D-erythro-hex-2-enopyranosyl}thymine (5.8). The product was purified and dried as a white foam in yields of 230 mg (60%) and confirmed by TLC R$_f$: (9:1 CH$_2$Cl$_2$:MeOH) 0.67.

^{1}H NMR (**5.8**, 500 MHz, DMSO-d_6) δ : 9.14 (1H, bs, NH), 7.06 (1H, s, H6), 7.46 – 6.83 (14H, m, MMT), 6.65 (1H, d, J = 4.5 Hz, H2'), 5.33 (1H, d, J = 8.5 Hz, H1'), 4.52 (1H, d, J = 6 Hz, H3'), 4.02 (1H, dd, J = 5, 6 Hz, H4'), 3.92 (1H, t, J = 9 Hz H5'), 3.78 (3H, s, -OCH$_3$), 3.51 (1H, dd, J = 3, 11 Hz, H6'), 3.42 (1H, dd, J = 4.5, 10 Hz, H6"), 1.90 (1H s, H7).

^{13}C NMR (**5.8**, 125.7 MHz, 500 MHz ^{1}H decoupled, CDCl$_3$) δ : 163.8 (C2), 158.9 (C4), 149.69 (C2'), 148.7, 144.3, 136.6, 135.3 (C6), 130.6, 129.5, 128.5, 128.2, 128.1, 127.4, 127.3, 124.2, 113.4, 111.8, 109.9 (C5), 98.43 (C3'), 78.40 (C4'), 69.33 (C5'), 63.06 (C6'), 56.08 (C1'), 55.45, 12.77 (C7); ESI-MS Calcd. for C$_{31}$H$_{30}$O$_6$N$_2$Na: 549.6, found: 550.1.

1–{4-hydroxy-6-[4-(methoxyphenyl)diphenyl]-2,3-dideoxy-β-D-erythro-hex-2-enopyranosyl}thymine (5.13). The product was purified and dried as a white foam in yields of 780 mg (78%) and confirmed by TLC R$_f$: (9:1 CH$_2$Cl$_2$:MeOH) 0.68.

^{1}H NMR (**5.13**, 400 MHz, CDCl$_3$) δ : 8.61 (NH), 7.40, 7.27, 6.84, (MMT), 6.94 (1H, d, J = 2 Hz, H6), 6.42 (1H, d, J = 2Hz, H1'), 6.20 (1H, ddd, J = 2, 10 Hz H3'), 5.64 (1H,

ddd, $J = 1.8, 3.6, 10$ Hz, H2'), 4.30 (1H, dd, $J = 2.4, 8.4$ Hz, H4'), 3.82 (OCH$_3$), 3.80 (1H, m, H5'), 3.58 (1H, dd, $J = 4.8, 9.2$ Hz, H6'), 3.31 (1H, dd, $J = 6.4, 9.2$ Hz, H6''), 1.90 (1H, d, 5).

^{13}C NMR (**5.13**, 100 MHz, ^{1}H decoupled at 400 MHz, CDCl$_3$) δ : 163.4 (CO), 158.8 (CO), 150.4, 149.1, 143.6, 136.3 (C2'), 135.89 (C6), 134.7, 130.4, 128.3, 128.2, 127.9, 127.4 (C3'), 125.4, 123.96 (C6'), 111.7 (C5), 87.8, 78.64 (C1'), 77.64 (C5'), 66.3 (C4'), 65.5 (C6'), 55.6 (OMe), 13.01 (C7); ESI MS Calcd. for C$_{31}$H$_{30}$N$_2$O$_6$: 549.4, found: 549.1.

(1R)-1-{(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-[4-(methoxyphenyl)diphenyl])-oxepinyl} thymine (5.21). The product was purified and dried as a white foam in yields of 700 mg (65%) and confirmed by TLC R$_f$: (9:1 CH$_2$Cl$_2$:MeOH) 0.73.

^{1}H NMR (**5.21**, 500 MHz, CDCl$_3$) δ : 8.62 (1H, s, NH), 7.70 – 6.83 (14H, m, MMT), 7.35 (1H, s, H6), 5.85 (1H, d, $J = 12.5$ Hz, H4'), 5.75 (1H, d, $J = 9$ Hz, H1'), 5.68 (1H, t, $J = 9.75$ Hz, H3'), 4.55 (1H, d, $J = 8$ Hz, H5'), 3.79 (3H, s, OCH$_3$), 3.42 (1H, dd, $J = 4, 10$ Hz, H7'), 3.65 (1H, t, $J = 4$ Hz, H6'), 3.35 (1H, dd, $J = 4, 10$ Hz, H7''), 2.60 (1H, t, H2', $J = 11.75$ Hz), 2.46 (1H, dd, $J = 8, 12.5$ Hz, H2''), 1.97 (3H, s, H7).

^{13}C NMR (**5.21**, 125.7 MHz, 500 MHz ^{1}H decoupled, CDCl$_3$): δ 163.78 (C2), 158.9 (C4), 149.9, 144.12 137.3, 135.5, 135.4 (C3'), 135.1 (C6), 130.4, 128.4, 128.2, 127.4, 125, 122.8, 113.4 (C4'), 84.11 (C1'), 82.06 (C6'), 72.29 (C5'), 65.57 (C7'), 55.56, 36.35 (C2'), 13.08 (C7); ESI-MS Calcd for C$_{31}$H$_{32}$N$_2$O$_7$Na 563.6, found 563.1.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-[4-(methoxyphenyl)diphenyl])-β-oxepanyl]thymine (5.30). The product was purified and dried as a white foam in yields of 135 mg (66%) and confirmed by TLC R$_f$: (9:1 CH$_2$Cl$_2$:MeOH) 0.68.

^{1}H NMR (**5.30**, 500 MHz CDCl$_3$) δ : 8.85 (1H, s, NH), 7.34 (4H, ar), 7.2 (1H, s, H6), 7.19 (8H, ar), 6.76 (2H, ar), 5.74 (1H, dd, $J = 3.6, 9.8$ Hz H1'), 3.83 (H1, ddd, $J = 3.6, 4.8, 7.6$ Hz, H5'), 3.72 (s, OMe), 3.6 (1H, dd, $J = 6, 12.6$ Hz, H6'), 3.31 (1H, dd, $J = 5.6, 9.6$ Hz, H7), 3.11 (1H, dd, $J = 5.6, 9.6$ Hz, H7''), 1.98 (1H, dd, $J = 3.6, 9$ Hz, H3'), 1.78 (1H, m, H3''), 1.80 (2H, m, H4'H4''), 1.86 (3H, d, $J = 1$ Hz, H7), 1.62 (2H, dd, $J = 8.4, 14.4$ Hz, H2'H2'');

¹³C NMR (**5.30**, 125.7MHz, 1H decoupled 500MHz, CDCl₃) δ : 163.79 (CO), 158.93 (CO), 149, 144.3,144,1, 135.9 (C6), 135.2, 130.6, 128.5, 128.2, 127.3, 113.5, 110.9, 86.53 (C1'), 87.16 (C5), 83.58 (C6'), 73.41 (C5'), 65.90 (C7'), 55.47 (OMe), 35.81 (C3'), 33.48 (C4'), 18.31 (C2'), 12.86 (C7); ESI-MS Calcd for C₃₂H₃₄N₂O₆Na: 565.6, found: 565.1.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-hydroxy-7-[4-(methoxyphenyl)diphenyl])-β-oxepanyl]-N⁶-benzoyladenine (5.31). The product was collected as a white foam in yields of 120 mg (50%) and confirmed by TLC; R_f: (9:1 CH₂Cl₂:MeOH) 0.65.

¹H NMR (**5.31**, 500 MHz, CDCl₃) δ : 9.03 (1H, s, H4), 8.14 (1H, s, H2), 8.72 (1H, s, H8), 7.95 (2H, ar), 7.52 (1H, ar), 7.44 (2H, ar), 7.29 (3H, ar), 7.17 (8H ar), 6.71 (2H, ar.), 5.99 (1H, dd, J = 2.8, 10 Hz, H1'), 3.90 (1H, t, J = 8.4 Hz, H6'), 3.79 (1H, dd, J = 5.6, 12.6 Hz, H5'), 3.71 (OMe), 3.30 (1H, dd, J = 5.6, 9.6 Hz, H4'), 3.14 (1H, dd, J = 6, 9.6 Hz, H4''), 2.27 (1H, ddd, J = 3, 6.5, 15.6 Hz, H2'), 2.16 (1H, ddd, J = 4.8, 9, 15.6 Hz, H2''), 1.99 (1H, dd, J = 4.8, 17.2 Hz, H3'), 1.92 (1H, dd, J = 2.8, 17.2 Hz, H3''), 1.82 (1H, ddd, J = 4, 8.4, 16Hz, H7'), 1.73 (1H, t, J = 8.4 Hz, H7'');

¹³C NMR (**5.31**, 125.7 MHz, 500 MHz ¹H decoupled, CDCl₃) δ : 158.9 (CO), 152.8 (C8), 144.3, 140.6 (C2), 144.1, 135.2, 132.98, 130.5, 129.5, 129.1, 128.48, 128.13, 128.1, 128.08, 127.4, 127.3, 113.5, 95, 87.2, 86.44 (C1'), 83.71 (C5'), 73.28 (C6'), 65.88 (C4'), 55.44 (OMe), 36.08 (C2'), 33.61 (C7'), 18.37 (C3'); ESI-MS Calcd for C₃₉H₃₇N₅O₅: 655.9 found: 655.7.

General procedure for the phosphitylation reaction.

The tritylated nucleoside (0.224 mmol) was dried under vacuum overnight prior to reaction. Dry THF (1.2 mL) was added under N₂. To the resulting solution was added dropwise over a span of 10 min, EtN(i-Pr)₂ (0.89 mmol) and Cl-P(OCEt)N(i-Pr)₂ (0.246 mmol). The reaction mixture was stirred for 2 h at 22°C and the reaction progress monitored by TLC (2:1 Hex:EtOAc). The progression of the reaction is also observable by the formation of a white precipitate, Cl⁻⁺NH(Et)(iPr)₂. After the reaction reached completion, EtOAc (15 mL) was added and the mixture was washed twice with saturated

aqueous NaHCO₃, dried over MgSO₄ and concentrated to a yellowish foam, which was purified by silica gel chromatography (Hex:EtOAc v/v 2:1 to 1:2 with 3% TEA).

2', 5'-di-(monomethoxytrityl) 3'-phosphoramidous ribouridine (2.5). The purified phosphoramidite diastereomers were collected in yields of 2.50 g (66%) as a white foam and confirmed by TLC with R_f: (1:1 Hex:EtOAc with 5% TEA) 0.59 and 0.51.

^{31}P NMR (**2.5**, CDCl₃, 80 MHz, ^{1}H decoupled at 200 MHz) δ : 153.15 and 148.96. ESI-MS Calcd. for $C_{58}H_{61}N_4O_9PNa$: 1012, found: 1012.

3', 5'-di-(monomethoxytrityl) 2'-phosphoramidous ribouridine (2.6). The purified phosphoramidite diastereomers were collected in yields of 1.71 g (66%) as a white foam and confirmed by TLC with R_f: (1:1 Hex:EtOAc with 5% TEA) 0.63 and 0.54.

^{31}P NMR (**2.6**, CDCl₃, 80 MHz, ^{1}H decoupled at 200 MHz) δ : 152.83 and 150.38. ESI-MS: Calcd. for $C_{58}H_{61}N_4O_9PNa$: 1012, found : 1012.

5'-*O*-levulinyl 2'-(monomethoxytrityl) 3'-phosphoramidous ribouridine (4.10). The purified phosphoramidite diastereomers were collected in yields of 1.13 g (60%) as a white foam and confirmed by TLC with R_f: (2 : 1 EtOAc : Hex) 0.32.

^{31}P NMR (**4.10**, 80 MHz, ^{1}H decoupled at 200 MHz, CDCl₃) δ : 156.7 and 153.59; ESI-MS Calcd. for $C_{43}H_{51}N_4O_{10}PNa$: 837.9, found: 837.1.

5'-*O*-levulinyl 3'-(monomethoxytrityl) 2'-phosphoramidous ribouridine (4.11). The purified phosphoramidite diastereomers were collected in yields of 604 mg (65%) as a white foam and confirmed by TLC with R_f: (2:1 EtOAc:Hex) 0.28.

^{31}P NMR (**4.11**, 80 MHz, ^{1}H decoupled at 200 MHz, CDCl₃) δ : 156.5 and 154.72; ESI-MS Calcd. for $C_{43}H_{51}N_4O_{10}PNa$: 837.9, found : 838.1.

1-{2,3-dideoxy-5-phosphoramidous-6-[4-(methoxyphenyl)diphenyl]-*β*-D-erythro-hex-2'-enopyranosyl}thymine (5.9). The purified phosphoramidite diastereomers were collected in yields of 210 mg (83%) as a white foam and confirmed by TLC with R_f: (2:1 EtOAc:Hex) 0.60.

211

^{31}P NMR (**5.9**, 80.99 MHz, ^{1}H decoupled 200 MHz, CDCl$_3$) δ : 150.2 and 149.0; ESI-MS Calcd. for C$_{40}$H$_{47}$O$_7$N$_4$PNa: 749.8, found: 750.2.

1-[(2,3-dideoxy-4-phosphoramidous-6-[4-(methoxyphenyl)diphenyl]-β-D-erythro-hex-2'-enopyranosyl)thymine (5.14). The purified phosphoramidite diastereomers were collected in yields of 825 mg (80%) as a white foam and confirmed by TLC with R$_f$: (2:1 EtOAc:Hex) 0.59.

^{31}P NMR (**5.14**, 80 MHz, ^{1}H decoupled at 200 MHz, CDCl$_3$) δ : 151.8 and 150.1; ESI-MS Calcd for C$_{40}$H$_{47}$N$_4$O$_7$PNa: 749.8, found: 749.2.

(1R)-1-{(2,3,4-trideoxy-(5S,6R)-5-phosphoramidous-7-[4(methoxyphenyl)diphenyl])-oxepinyl}thymine (5.22). The purified phosphoramidite diastereomers were collected in yields of 800 mg (80%) as a white foam and confirmed by TLC with R$_f$: (2:1 EtOAc: Hex) 0.58.

^{31}P NMR (**5.22**, 80.99 MHz, ^{1}H decoupled 200 MHz, CDCl$_3$) δ : 146.7 and 144.9; Expt. ESI-MS Calcd for C$_{41}$H$_{49}$N$_4$O$_7$PNa: 763.6, found: 763.2.

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-phosphoramidous-7-[4(methoxyphenyl)diphenyl])-β-oxepanyl]thymine (5.32). The purified phosphoramidite diastereomers were collected in yields of 131 mg (80%) as a white foam and confirmed by TLC with R$_f$: (2:1 Hex: EtOAc) 0.52.

^{31}P NMR (**5.32**, 80.99 MHz, ^{1}H decoupled 200 MHz, CDCl$_3$) δ : 149.5 and 149.0; ESI-MS Calcd for C$_{41}$H$_{51}$N$_4$O$_7$PNa: 765.86, found: 765.2 .

(1R)-1-[(2,3,4-trideoxy-(5S,6R)-5-phosphoramidous-7-[4(methoxyphenyl)diphenyl])-oxepanyl]-N^6-benzoyladenine (5.33). The purified phosphoramidite diastereomers were collected in yields of 135 mg (92%) as a white foam and confirmed by TLC with R$_f$: (2:1 Hex:EtOAc) 0.33.

^{31}P NMR (**5.33**, 80.99 MHz, ^{1}H decoupled 200 MHz, CDCl$_3$) δ : 148.4 and 147.9; ESI-MS Calcd for C$_{48}$H$_{54}$N$_7$O$_6$PNa 878.9, found 878.3.

References :

[1] Avery, T.O.; Macleod, C. and McCarthy, M. *J. Exp. Med.,* **1944**, *79*, 137-158.

[2] Crick, F. *Nature,* **1970**, *227*, 561-563.

[3] Watson, J.D. and Crick, F.H.C. *Nature,* **1953**, *171*, 964-967.

[4] Blackburn, G.M. and Gait, M.J. *Nucleic Acids in Chemistry and Biology;* Oxford University Press: New York, USA, **1996**.

[5] Sun, G.; Voigt, J.H.; Filippov, I.V.; Marquez, V.E. and Nicklaus, M.C. *J. Chem. Inf. Comput. Sci.,* **2004**, *44*, 1752-1762.

[6] Wang, A.H-J.; Quigley, C.J.; Kolpak, F.J.; Crawford, J.L.; van Boom, J.H.; van der Marel, G. and Rich, A. *Nature,* **1979**, *282*, 680-686.

[7] Leroy, J-L.; Gueron, M.; Mergny, J-L. and Helene, C. *Nucleic Acids Res.,* **1994**, *22*(9), 1600-1606.

[8] Leroy, J.L.; Gehring, K.; Kettani, A. and Gueron, M. *Biochemistry,* **1993**, *32*, 6019-6031.

[9] a. Guschlbauer, W.; Chantot, J-F. and Thiele, D. *J. Biomol. Struct. Dyn.,* **1990**, *8*, 491-511. b. Smith, F.W.; Francis, L.W. and Feigon, J. *Proc. Natl. Acad. Sci. USA.,* **1994**, *91*, 10546-10550.

[10] a. Cech, T.R.; Zuang, A.J. and Grabowski; P.J. *Cell,* **1981**, *27*, 487-496. b. Guerrier-Takada, C. and Altman; S. *Science,* **1984**, *223*, 285-286.

[11] a. Thibaudeau, C.; Plavec, J.; Watanabe, K.A. and Chattopadhyaya, J. *J. Chem. Soc. Chem. Comm.,* **1994**, 537-540. b. a. Thibaudeau, C.; Plavec, J.; Garg, N.; Papchikhin, A. and Chattopadhyaya, J. *J. Am. Chem. Soc.,* **1994**, 116, 4038-4043. c. Acharya, P.; Nawrot, B.; Sprinzl, M.; Thibaudeau, C. and Chattopadhyaya, J. *J. Chem. Soc. Perkin Trans. 2,* **1999**, 1531-1536.

[12] Saenger, W. In *Principles of Nucleic Acid Structure.* Cantor, C.R. Ed. Springer-Verlag Inc.; New-York, NY, USA, **1984**.

[13] Dickerson, R.E.; Drew, B.N.; Conner, R.M.; Wing, R.M.; Fratini, A.V. and Kopka, M.L. *Science,* **1982**, *216*, 475-485.

[14] a. Khorana, H.G. and Todd, A.R. *J. Chem. Soc. Abs.,* **1953**, 2257-2260. b. Gilham, P.T. and Khorana, H.G. *J. Am. Chem. Soc.,* **1958**, *80*, 6212-6222.

[15] a. Letsinger, R.L. and Mahadevan, V. *J. Am. Chem. Soc.,* **1965**, *87*, 3526-3527. b. Letsinger, R.L. and Mahadevan, V. *J. Am. Chem. Soc.,* **1966**, *88*, 5319-5324. c. Letsinger, R.L. and Ogilvie, K.K. *J. Am. Chem. Soc.,* **1967**, *89*, 4801-4803. d. Letsinger, R.L. and Ogilvie, K.K. *J. Am. Chem. Soc.,* **1969**, *91*, 3350-3355.

[16] a. Letsinger, R.L. and Lunsford, W.B. *J. Am. Chem. Soc.,* **1976**, *98*, 3655-3661. b. Matteucci, M.D. and Caruthers, M.H. *Tetrahedron Lett.,* **1980**, *21*, 719-722. c. Ogilvie, K.K. and Nemer, M.J. *Tetrahedron Lett.,* **1980**, *21*, 4159-4148.

[17] a. Beaucage, S.L. and Caruthers, M.H. *Tetrhedron Lett.,* **1981**, *22*, 1859-1862. b. McBride, L.J. and Caruthers, M.H. *Tetrahedron Lett.,* **1983**, *24*, 245-248.

[18] Adlvaredo-Urbina, G.; Sathe, G.M.; Liu, W.C.; Gillen, M.F.; Duck, P.D.; Bender, R. and Ogilvie, K.K. *Science,* **1981**, *214*, 270-274.

[19] a. Pon, R.T.; Yu, S. and Sanghvi, Y.S. *Bioconj. Chem.,* **1999**, *10*, 1051-1057. b. Damha, M.J.; Ganeshan, K.; Hudson, R.H.E. and Zabarylo, S.V. *Nucleic Acids Res.,* **1992**, *20*(24), 6565-6573.

[20] a. Schaller, H.; Weimann, G.; Lerch, B. and Khorana, H.G. *J. Am. Chem. Soc.,* **1963**, *85*, 3821-3827. b. Hakimelahi, G.H.; Proba, Z.A. and Ogilvie, K.K. *Can. J. Chem.,* **1982**, *60*, 1106-1113.

[21] Tener, G.M. *J. Am. Chem. Soc.,* **1961**, *83*, 159-168.

[22] Reese, C.B. In *Current Protocols in Nucleic Acid Chemistry,* **2000**, 2.2.1-2.2.24.

[23] Ti, G.S.; Gaffney, B.L. and Jones, R.A. *J. Am. Chem. Soc.,* **1982**, *104*, 1316-1319.

[24] a. Caruthers, M.H.; Barone, A.D.; Beaucage, S.L.; Dodds, D.R.; Fisher, E.F.; McBride, L.J.; Mateucci, M.; Stabinsky, Z. and Tang, J-Y. *Methods in Enzymology* **1987**, *154*, 287-313. b. Damha, M.J. and Ogilvie, K.K. In *Methods in Molecular Biology: Protocols for Oligonucleotides and Analogs* 1993, *20*, Humana Press Inc. Totowa, N.J. c. Schulhof, J.C.; Molko, D. and Teoule, R. *Nucleic Acids Res.* **1987**, *15*(2), 397-416. d. Wu, T.; Ogilvie, K.K. and Pon, R.T. *Nucleic Acids Res.* **1989**, *17*(9), 3501-3517. e. Wincott, F.; DiRenzo, A.; Shaffer, C.; Grimm, S.; Tracz, D.; Workman, C.; Sweedler, D.; Gonzalez, C.; Scaringe, S. and Usman, N. *Nucleic Acids Res.,* **1995**, *23*(14), 2677-2684.

[25] Damha, M.J. and Zabarylo, S.V. *Tetrahedron Lett.,* **1989**, *30*(46), 6295-6298.

[26] Carriero, S. and Damha, M.J. *J. Org. Chem.,* **2003**, *68*, 8328-8338.

[27] Hudson, R.H.E.; Uddin, A.H. and Damha, M.J. *J. Am. Chem. Soc.*, **1995**, *117*, 12470-12477.

[28] Uddin, A.H.; Piunno, P.A.E.; Hudson, R.H.E., Damha, M.J. and Krull, U.J. *Nucleic Acids Res.*, **1997**, *25*, 4139-4146.

[29] a. Nam, K.; Hudson, R.H.E.; Chapman, K. B.; Ganeshan, K.; Damha, M.J. and Boeke, J.D. *J. Bio. Chem.*, **1994**, *269*(32), 20613-20621. b. Carriero, S. and Damha, M.J. *Nucleic Acids Res.*, **2003**, *31*(21), 6157-6167. c. Khalid, M.F.; Damha, M.J.; Shuman, S. and Schwer, B. *Nucleic Acids Res.*, **2005**, *33*(19), 6349-6360.

[30] Shchepinov, M.S.; Udalova, I.A.; Bridgman, A.J. and Southern, E.M. *Nucleic Acids Res.*, **1997**, *25*(22), 4447-4454.

[31] a. Capaldi, S.; Getts, R.C. and Sumedha, D.J. *Nucleic Acids Res.*, **2000**, *28*(7), e21. b. Mora, J.R. and Getts, R.C. *BioTechniques* **2006**, *41*(4), 420-424.

[32] a. Damha, M.J.; Pon, R.T. and Ogilvie, K.K. *Tetrahedron Lett.*, **1985**, *26*, 4839-4842. b. Damha, M.J. and Ogilvie, K.K. *J. Org. Chem.*, **1988**, *53*, 3710-3722. c. Zhou, X.; Remaud, G. and Chattopadhyaya, J. *Tetrahedron Lett.*, **1988**, *44*, 6471-6489. Foldesi, A.; Balgobin, N. and Chattopadhyaya, J. *Tetrahedron Lett.*, **1989**, *30*, 881-884.

[33] a. Reese, C.B. and Skone, P.A. *Nucleic Acids Res.*, **1985**, *13*, 5215-5231., b. Huss, S.; Gosselin, G.; Stawinski, J.; Stromberg, R. and Imbach, J-L. *Nucleosides, Nucleotides and Nucleic Acids* **1988**, *7*, 321-337.

[34] a. Kierzek, R.; Kopp, D.W.; Edmonds, M. and Caruthers, M.H. *Nucleic Acids Res.*, **1986**, *14*, 4751-4764. b. Huss, S.; Gosselin, G.; Imbach, J-L. *Tetrahedron Lett.*, **1987**, *28*, 415-418.

[35] a. Fujimoto, K.; Ogawa, N.; Hayashi, M.; Matsuda, S. and Saito, S. *Tetrahedron Lett.*, **2000**, *41*, 9437-9440. b. Mitra, D. *Ph.D. Thesis,* **2007**, McGill University.

[36] Coppins, R.L. and Silverman, S.K. *Biochemistry,* **2005**, *44*, 13439-13446.

[37] Damha, M.J.; Ganeshan, K.; Hudson, R.H.E. and Zabarylo, S.V. *Nucleic Acids Res.*, **1992**, *20*, 6565-6573.

[38] a. Furuichi, T.; Inouye, S. and Innouye, M. *Cell,* **1987**, *48*, 55-62. b. Dhundale, A.; Lampson, B.; Furuichi, T.; Innouye, S. and Innouye, M. *Cell,* **1987**, *51*, 1105-1112.

[39] a. Reilly, J.D.; Wallace, J.C.; Melhem, R.F.; Kopp, D.W.; Edmonds, M. *Methods in Enzymology,* **1989**, *180*, 177-191. b. Jeong, M-A. and Lim, D. *J. of Microbiol.,* **2004**, *42(3)*, 200-204.

[40] a. Sproat, B.S.; Beijer, B.; Grotli, M.; Ryder, U.; Morand, K.L. and Lamond, A.I. *J. Chem. Soc. Perkin Trans. 1,* **1994**, 419-431. b. Braich, R.S. and Damha, M.J. *Bioconj. Chem.,* **1997**, *8*, 370-377. c. Damha, M.J. and Braich, R.S. *Tetrahedron Lett.,* **1998**, 39, 3907-3910. d. Sorensen, M.D.; Meldgaard, M.; Rajwanshi, V. K. and Wengel, J. *Bioorg. Med. Chem. Lett.,* **2000**, *10*, 1853-1856.

[41] Hudson, R.H.E. and Damha, M.J. *J. Am. Chem. Soc.,* **1993**, *115*, 2119-2124.

[42] Hudson, R.H.E.; Robidoux, S. and Damha, M.J. *Tetrahedron Lett.,* **1998**, *39*, 1299-1302.

[43] Wallace, J.C.; Edmonds, M. *Proc. Natl. Appl. Sci. USA,* **1983**, *80*, 950-954.

[44] a. Grabowski, P.J.; Padgett, R.A. and Sharp, P.A. *Cell,* **1984**, *37*, 415-427. b. Berget, S.M.; Moore, C.; Sharp, P.A. *Proc. Natl. Appl. Sci. USA,* **1977**, *74*, 3171-3175.

[45] Ooi, S.L; Samarsky, D.A; Fournier, M.J. and Boeke, J.D. *RNA,* **1998**, *4*, 1096-1110.

[46] Furuichi, T.; Inouye, S. and Innouye, M. *Cell,* **1987**, *48*, 55-62. b. Dhundale, A.; Lampson, B.; Furuichi, T.; Innouye, S. and Innouye, M. *Cell,* **1987**, *51*, 1105-1112.

[47] Ruskin, B. and Green, M.R. *Science,* **1985**, *229*, 135-140.

[48] Arenas, J. and Hurwitz, J. *J. Biol. Chem.,* **1987**, *262*, 4274-4279.

[49] Chapman, K.B. and Boeke, J.D. *Cell,* **1991**, *65*, 483-492.

[50] Furuichi, T.; Dhundale, A.; Inouye, M. and Innouye, S. *Cell,* **1987**, *48*, 47-53.

[51] Nam, K.; Hudson, R.H.E.; Chapman, K.B.; Ganeshan, K.; Damha, M.J. and Boeke, J.D. *Methods Enzymol.,* **1994**, *269*, 20613-20621.

[52] Petfalski, E.; Dandekar, T.; Henry, Y. and Tollervey, D. *Mol. Cell Biol.,* **1998**, *18*, 1181-1189.

[53] Ratmeyer, L.; Vinayak, R.; Zhong, Y.Y.; Zon, G. and Wilson, D.W. *Biochemistry,* **1994**, *33*, 5298-5304.

[54] Noy, A.; Perez, A.; Marquez, M.; Luque, J. and Orozco, M. *J. Am. Chem. Soc.,* **2005**, *127*(13), 4910-4920.

[55] Fedoroff, O.Y.; Salazar, M. and Reid, B.R. *J. Mol. Biol.,* **1993**, *233*, 509-523.

216

[56] Zamecnik, P.C. and Stephenson, M.L. *Proc. Natl. Acad. Sci. USA*, **1978**, *75*, 280-284.

[57] Friedman, K.J.; Kole, J.; Cohn, J.A.; Knowles, M.R.; Silverman, L.M. and Kole, R. *J. Biol. Chem.*, **1999**, *274*(51), 36193-36199.

[58] Zinker, B.A. *et. al. Proc. Natl. Acad. Sci. USA,* **2002**, *99*(17), 11357-11362.

[59] Good, L. and Neilsen, P.E. *Nature Biotechnology,* **1998**, *16*, 355-358.

[60] a. Mangos, M.M. and Damha, M.J. *Curr. Topics in Med. Chem.* **2002**, *2*, 1147-1171. b. Dean, N.M. and Bennett, F.C. *Oncogene,* **2003**, *22*, 9087-9096.

[61] Noronha, A.M. and Damha, M.J. *Nucleic Acids Res.,* **1998**, *26*, 2665-2671.

[62] Noronha, A.M.; Wilds, C.J.; Lok, C-N.; Viazovkina, K.; Arion, D.; Parniak, M.A. and Damha, M.J. *Biochemistry,* **2000**, *39*(24), 7050-7062.

[63] Minasov, G.; Teplova, M.; Neilsen, P.; Wengel, J. and Egli, M., *Biochemistry,* **2000**, *39*(13), 3525-3532.

[64] a. Damha, M.J.; Wilds, C.J.; Noronha, A.M.; Brukner, I.; Borkow, G.; Arion, D. and Parniak, M.A. *J. Am. Chem. Soc.,* **1998**, *120*, 12976-12977. b. Wilds, C.J. and Damha, M.J. *Nucleic Acids Res.,* **2000**, *28*, 3625-3635.

[65] Venkateswarlu, D.; Lind, K.E.; Mohan, V.; Manoharan, M. and Ferguson, D.M. *Nucleic Acids Res.,* **1999**, *27*, 2189-2195.

[66] Manoharan, M., *Biochim. Biophys. Acta.,* **1999**, *1489,* 117-130.

[67] a. Christensen, N.K.; Peterssen, M.; Neilsen, P.; Jacobsen, J.P.; Olsen, C.E. and Wengel, J. *J. Am. Chem. Soc.,* **1998**, *120*, 5458-5463. b. Altman, K-H.; Imwinkelried, R.; Kesselring, R. and Rihs, G. *Tetrahedron Lett.,* **1994**, *35*, 7625-7628. c. Wengel, J.; Koshkin, A.; Singh, S.K.; Neilsen, P.; Meldgaard, M.; Rajwanshi, V.K.; Kumar, R.; Skouv, J.; Nielsen, C.B.; Jacobsen, J.P.; Jacobsen, N. and Olsen, C.E. *Nucleosides and Nucleotides,* **1999**, *18*, 1365-1370.

[68] Mangos, M.M.; Min, K-L.; Viazovkina, E.; Galarneau, A.; Elzagheid, M.I.; Parniak, M.A. and Damha, M.J. *J. Am. Chem. Soc.,* **2003**, *125*, 654-661.

[69] a. Neislen, P.A.; Egholm, M.; Berg, R.H. and Buchardt, O. *Science,* **1991**, *254*, 1497-1500. b. Egholm, M.; Buchardt, O.; Neilsen, P.E. and Berg, R.H. *J. Am. Chem. Soc.,* **1992**, *114*, 1895-1897. c. Nelson, K.E.; Levy, M. and Miller, S.L. *Proc. Natl. Acad. Sci. USA,* **2000**, *97*, 3868-3871.

[70] a. Joyce, J.F.; Schwartz, A.W.; Miller, S.L. and Orgel, L.E. *Proc. Natl. Acad. Sci. USA,* **1987,** *84,* 4398-4402. b. Schneider, K.C. and Benner, S.A. *J. Am. Chem. Soc.,* **1990,** *112,* 453-455.

[71] Zhang, L.; Peritz, A. and Meggers, E. *J. Am. Chem. Soc.* **2005,** *127,* 4174-4175.

[72] Joyce, G. *Nature* **2002,** *418,* 214-221.

[73] Eschenmoser, A. *Science* **1999,** *284,* 2118-2124.

[74] Stoddart, J.F. *Stereochemistry of Carbohydrates* New York, Wiley InterScience, **1971.**

[75] Hunziker, J.; Roth, H. J.; Boehringer, M.; Giger, A.; Diederichsen, U.; Goebel, M.; Krishnan, R.; Jaun, B.; Leumann, C. and Eschenmoser, A. *Helv. Chim. Acta,* **1993,** *76*(1), 259-352.

[76] a. Augustyns, K.; Vandendriessche, F.; Van Aerschot, A.; Busson, R.; Urbanke, C. and Herdewijn, P. *Nucl. Acids Res.,* **1992,** *20*(18), 4711-4716. b. Augustyns, K.; Van Aerschot, A.; Urbanke, C. and Herdewijn, P. *Bull. Soc. Chim. Belg.* **1992,** *101,* 119-130. c. Lescrinier, E.; Froeyen, M. and Herdewijn, P. *Nucleic Acids Res.,* **2003,** *31*(12), 2975-2989.

[77] a. Hendrix, C.; Rosemeyer, H.; De Bouvere, B.; Van Aerschot, A.; Seela, F. and Herdewijn, P. *Chem. Eur. J.,* **1997,** 3, 1513-1520. b. Maurinsh, Y.; Rosemeyer, H.; Esnouf, R.; Medvedovici, A.; Wang, J.; Ceulemans, G.; Lescrinier, E.; Hexndrix, C.; Busson, R.; Sandra, P.; Seela, F.; Van Aerscot, A. and Herdewijn, P. *Chem. Eur. J.* **1999,** *5*(7), 2139-2150.

[78] Lescrinier, E.; Esnouf, R.; Schraml, J.; Busson, T.; Heus, H.A.; Hibers, C.W. and Herdewijn, P. *Chemistry & Biology,* **2000,** *7*(9), 719-731.

[79] Wang, J.; Verbeure, B.; Luyten, I.; Lescrinire, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aershot, A. and Herdewijn, P. *J. Am. Chem. Soc.,* **2000,** *122,* 8595-8602.

[80] Schoning, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science,* **2000,** *290*(5495), 1347-1351.

[81] Wilds, C.J.; Wawrzak, Z.; Krishnamurthy, R.; Eschenmoser, A.; Egli, M.. *J. Am. Chem. Soc.,* **2002**, *124*, 13716-13721.

[82] a. Chaput, J.C.; Ichida, J.K. and Szostak, J.W. *J. Am. Chem. Soc.,* **2003**, *125*(4), 856-857. b. Chaput, J.C. and Szostak, J.W. *J. Am. Chem. Soc.,* **2003**, *124*(31), 9274-9275.

[83] a. Eckstein, F. *Ann. Rev. Biochem.,* **1985**, *54*, 367-402. b. Marshall, W.S. and Caruthers, M.H. *Science,* **1993**, *259*, 1564-1570. c. Higson, A.P.; Sierzchala, A.; Brummel, H.; Zhao, Z. and Caruthers, M.H. *Tetrahedron Lett.,* **1998**, *39*, 3899-3902.

[84] Agrawal, S. and Zhao, Q. *Curr. Opin. Chem. Biol.,* **1998**, *2*, 519-528.

[85] Crooke, S.T. *Antisense & Nucleic Acid Drug Development,* **1998**, *8*, 115-122.

[86] De Smet, M.D.; Meenken, C. and van den Horn, G. *Ocular Immunology and Inflamation,* **1999**, *7*(3-4), 189-198.

[87] a. Liu, H.; Gao, J.; Lynch, S.R.; Saito, Y.D.; Maynard, L. and Kool, E.T. *Science,* **2003**, *302*, 868-871. b. Liu, H.; Gao, J.; Maynard, L.; Saito, Y.D. and Kool, E.T. *J. Am. Chem. Soc.,* **2004**, *126*, 6900-6905. c. Liu, H.; Lynch, S.R. and Kool, E.T. *J. Am. Chem. Soc.,* **2004**, *126*, 11826-11831. d. Lee, A.H.F. and Kool, E.T. *J. Am. Chem. Soc.,* **2005**, *127*, 3332-3338.

[88] a. Uhlmann, E. and Peyman, A. *Chem Rev.,* **1990**, *90*, 543-585. b. Godard, G.; Boutorine, A.S.; Behmoaras-Saison, E. and Helene, C. *Eur. J. Chem.,* **1995**, *232*, 404-410.

[89] Jen, K-Y. and Gerwitz, A.M. *Stem Cells,* **2000**, *18*(5), 307-319.

[90] Felsenfeld, G.; Davies, D.R. and Rich, A. *J. Am. Chem. Soc.,* **1957**, *79*, 2023-2024.

[91] a. Hoogsteen, K. *Acta Crystallogr.,* **1957**, *12*, 822-823. b. Moser, H.E. and Dervan, P.D. *Science,* **1987**, *238*, 645-650.

[92] a. Rajagopal, P. and Feigon, J. *Nature,* **1989**, *339*, 637-640. b. Rajagopal, P. and Feigon, J. *Biochemistry,* **1989**, *28*, 7859-7870. c. De los Santos, C.; Rosen, M. and Patel, D. *Biochemistry,* **1990**, *29*, 8820-8826.

[93] a. Arnott, S.; Bond, P.J.; Selsing, E. and Smith, P.J.C. *Nucleic Acids Res.,* **1976**, *3*, 2459-2470. b. Letai, A.G.; Pallandino, M.A.; Fromm, E.; Rizzo, V. and Fresco, J.R. *Biochemistry,* **1988**, *27*, 9108-9112.

[94] a. Cooney, M.; Czernuszewicz, G.; Postel, E.H.; Flint, S.J. and Hogan, M.E. *Science*, **1988**, *241*, 456-459. b. Durland, R.H.; Kessler, D.J.; Gunnell, S., Duvic, M.; Pettit, B.M. and Hogan, M.E. *Biochemistry*, **1991**, *30*, 9246-9255.

[95] a. Ferdous, A.; Watanabe, H.; Akaike, T.; Marayuma, A. *Nucleic Acids Res.*, **1998**, *26*(17), 3949-3954. b. Torigoe, H.; Hari, Y.; Sekiguchi, M.; Obika, S. and Imanishi, T. *J. Biol. Chem.*, **2001**, *276*(4), 2354-2360.

[96] Hannon, G.J. Ed. *RNAi. A Guide to Gene Silencing*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, **2003**.

[97] Sontheimer, E.J. and Carthrew, R.W. *Science*, **2004**, *305*, 1409-1410.

[98] a. Hornung, V.; Guenthner-Biller, M.; Bourquin, C.; Ablasser, A.; Schlee, M.; Uematsu, S.; Noronha, A., Manoharan, M.; Akira, S.; de Fougerolles, A.; Endres, S. and Hartmann, G. *Nature Med.*, **2005**, *11*, 263-270. b. Kraynak, B.A. and Baker, B.F. *RNA*, **2006**, *12*, 163-176.

[99] a. Braasch, D.A.; Jensen, S.; Liu, Y.; Arar, K.; White, M.A. and Corey, D.R. *Biochemistry*, **2003**, *42*(26), 7967-7975. b. Amarzguioui, M.; Holen, T.; Babaie, E. and Prydz, H. *Nucleic Acids Res.*, **2003**, *31*, 589-595.

[100] a. Dowler, T.; Bergeron, D.; Tedeschi, A-L.; Paquet, L.; Ferrari, N. and Damha, M.J. *Nucleic Acids Res.*, **2006**, *34*(6), 1669-1675. b. Watts, J.K.; Choubdar, N.; Sadalapure, K.; Robert, F.; Wahba, A.S.; Pelltier, J.; Pinto, M.B. and Damha, M.J. *Nucleic Acids Res.*, **2007**, 1-11.

[101] a. Morrissey, D.V.; Blanchard, K.; Shaw, L.; Jensen, K.; Lockridge, J.A.; Dickinson, B.; McSwiggen, J.A.;Vargeese, C.; Bowman, K.; Shaffer, C.S.; Polisky, B.R. and Zinnen, S. *Hepatology*, **2005**, *41*(6), 1349-1356. b. Bartlett, D.W. and Davis, M.E. *Nucl. Acids Res.*, **2006**, *34*(1), 322-333.

[102] a. Nilsson, M.; Antson, D-O.; Barbany, G. and Landegren, U. *Nucleic Acids Res.*, **2001**, *29*(2), 578-581. b. McGinness, K.E. and Joyce, G.F. *Chemistry & Biology*, **2002**, *9*, 297-307.

[103] a.Wyatt, J.R.; Chastain, M.; and Puglisi, J.D. *BioTechniques*, **1991**, *11*, 764-769. b. Yeung, M.C.; Lau, A.S; *BioTechniques*, **1993**, *15*(3), 381-382. c. Lukavsky, P.J. and Puglisi, J.D. *RNA*, **2004**, *10*, 889-893.

[104] Muller, S.; Wolf, J. and Ivanov, S.A. *Curr. Org. Syn.*, **2004**, *1*(3), 293-307.

[105] a. Ogilvie, K.K. and Nemer, M.J. *Can. J. Chem.*, **1980**, *58*, 1389-1397. b. Ohstuka, E.; Yamane, A. and Ikehera, M. *Nucl. Acids Res.*, **1983**, *11*(5), 1325-1335.

[106] Day, R.T.; Williams, D.; Soriano, P. and Sanghvi, Y.S. *Nucleosides, Nucleotides and Nucleic Acids*, **2005**, *24*(5-7), 897-899.

[107] Song, Q.; Wang, Z. and Sanghvi, Y.S. *Nucleosides, Nucleotides and Nucleic Acids*, **2003**, *22*(5-8), 629-633.

[108] Patnaik, A.K.; Rao, N.S.; Kumar, P.; Sharma, A.K.; Garg, B.S. and Gupta, K.C. *Helv. Chim. Acta*, **2000**, *83*(2), 322-327.

[109] Usman, N.; Egli, M. and Rich, A. *Nucleic Acids Res.*, **1992**, *20(24)*, 6695-6699.

[110] Hakimelahi, G.H.; Zbignew, A.D. and Ogilvie, K.K. *Can. J. Chem.*, **1982**, *60*, 1106-1113.

[111] Cramer, H. and Pfleiderer, W. *Helv. Chim. Acta*, **1999**, *82*, 614-632.

[112] Lyttle, M.H.; Wright, P.B.; Sinha, N.D.; Bain, J.D. and Chamberlain, R.A. *J. Org. Chem.*, **1991**, *56*, 4608-4615.

[113] a. Bessodes, M.; Komoitis, D. and Antonakis, K. *Tetrahedron Lett.*, **1986**, *27*(5), 579-580. b. Yang, S.G.; Lee, D.H. and Kim, Y.H. *Heteroatom Chem.*, **1997**, *8*(5), 435-438.

[114] a. Mateucci, M.D. and Caruthers, M.H. *Tetrahedron Lett.*, **1980**, *21*, 3243-3246. b. Kierzeck, R.; Ito, H.; Bhatt, R. and Itakura, K. *Tetrahedron Lett.*, **1981**, *22*(38), 3761-3764.

[115] a. Breslow, R.; Anslyn, E. and Huang, D-L. *Tetrahedron*, **1991**, *47*, 2356-2376. b. Kosonen, M.; Hakala, K.; Lonnberg, H. *J. Chem. Soc. Perkin Trans. 2*, **1998**, 663-670.

[116] Gilar, M. and Bouvier, E.S.P *J. Chrom.*, **2000**, *890*(1), 167-177.

[117] a. van Boom, J.H. and Burgers, P.M.J. *Tetrahedron Lett.*, **1976**, *17*, 4875-4878. b. Ogilvie, K.K. and Nemer, M.J. *Can. J. Chem.*, **1980**, *58*, 1389-1397. c. Iwai, S. and Ohtsuka, E. *Nucleic Acids Res.* **1988**, *16*(20), 9443-9456. d. Ganeshan, K.; Tadey, T.; Nam, K.; Purdy, W.C.; Boeke, J.D. and Damha, M.J. *Nucleosides and Nucleotides*, **1995**, *14*(3-5), 1009-1013.

[118] a. Sund, C.; Agback, P. and Chattopadhyaya, J. *Tetrahedron,* **1991**, *47*(46), 9659-9674. b. Meldgaard, M.; Nielsen, N.K.; Bremner, M.; Pedersen, O.S.; Olsen, C.E. and Wengel, J. *J. Chem. Soc. Perkin Trans. 1,* **1997**, 1951-1955.

[119] Rozners, E.; Renhofa, R.; Petrova, R.; Popelis, J.; Kumpins, V; Bizdena, E *Nucleosides Nucleotides,* **1992**, *11,* 1579-1593.

[120] Ohtsuka, E.; Sasaki, T. and Shigenari, I. *Tetrahedron,* **1990**, *46*(19), 6673-6680.

[121] Reese, C.B. and Trentham, D.R. *Tetrahedron Lett.,* **1965**, *29*, 2467-2472.

[122] a. Lyttle, M.H.; Carter, T.G.; Dick, D.J. and Cook, R.M. *J. Org. Chem.,* **2000**, *65*(26), 9033-9038. b. Lyttle, M.H.; Walton, T.A.; Dick, D.J.; Carter, T.G.; Beckman, J.H. and Cook, R.M. *Bioconj. Chem.,* **2002**, *13*(5), 1146-1154.

[123] Pon, R.T. and Yu, S. *Nucleic Acids Res.,* **1997**, *25*(18), 3629-3635.

[124] Lackey, J.G.; Sabatino, D. and Damha, M.J. *Org. Lett.,* **2007**, *9*(5), 789-792.

[125] a. Donga, R.A. ; Khaliq, S.M. ; Chan, T-H. And Damha , M.J. *J. Org. Chem.,* **2006**, *71*(20), 7907-7910. b. Donga, R.A.; Chan; T-H. and Damha, M.J. *Can. J. Chem.,* **2007**, manuscript in press.

[126] Merrifield, B. *Biosciences Rep.,* **1985**, *5*(5), 353-376.

[127] Ravikumar, V. T.; Kumar, R. K.; Capaldi, D.C. and Cole, D. L. *Nucleosides, Nucleotides & Nucleic Acids,* **2003**, *22*(5-8), 1421-1425.

[128] Gilar, M. *Anal. Biochem.* **2001**, *298*, 196-206.

[129] Neislen, J. and Caruthers, M.H. *J. Am. Chem. Soc.,* **1988**, *110*(18), 6275-6276.

[130] Alul, R.H.; Singman, C.N.; Zhang, G. and Letsinger, R.L. *Nucleic Acids Res.,* **1991**, *19*(7), 1527-1532.

[131] Macdonald, P.M.; Damha, M.J.; Ganeshan, K.; Braich, R. and Zabarylo, S.V. *Nucleic Acids Res.,* **1996**, *24*(15), 2868-2876.

[132] a. Hershlag, D.; Piccirilli, J.A. and Cech, T.R. *Biochemistry,* **1991**, *30*, 4844-4854. b. Breslow, R. and Xu, R. *J. Am. Chem. Soc.,* **1993**, *115*, 10705-10713.

[133] Herschlag, D. *J. Am. Chem. Soc.,* **1994**, 116(26), 11631-11635.

[134] Lonnberg, H. and Korhonen, J. *J. Am. Chem. Soc.,* **2005**, *127*, 7752-7758.

[135] a. Cohn, W.E. and Volkin, E. *Arch. Biochem. Biophys.,* **1952,** *35,* 465. b. Cohn, W.E. and Volkin, E. *Nature,* **1951,** *167,* 483. c. Brown, D.M.; MaGrath, D.I. and Todd, A.R. *J. Chem. Soc.,* **1954,** 1442.

[136] Kosonen, M.; Seppannen, R.; Wichman, O. and Lonnberg, H. *J. Chem. Soc. Perkin Trans. 2,* **1999,** 2433-2439.

[137] Chandler, A.J.; Hollfelder, F.; Kirby, A.J. and O'Carroll, F. and Stromberg, R. *J. Chem. Soc. Perkin Trans. 2,* **1994,** 327-333.

[138] a. Brown, D.M.; Magrath, D.I. and Todd, A.R. *J. Chem. Soc.,* **1955,** 4396-4401. b. Oivanen, M.; Kuusela, S. and Lonnberg, H. *Chem. Rev.,* **1998,** *98,* 961-990.

[139] Bannon, P. and Verly, W. *Eur. J. Biochem.,* **1972,** *31(1),* 103-111.

[140] Valadkhan, S and Manley, J.L. *Nature,* **2001,** *418,* 701-707.

[141] Breslow, R.; Anslyn, E.; Huang, D.-L. *Tetrahedron,* **1991,** *47,* 2365-2376.

[142] Kosonen, M.; Hakala, K. and Lonnberg, H. *J. Chem. Soc. Perkin Trans. 2,* **1998,** 663-670.

[143] a. Van Boom, J.H. and De Rooy, J.F.M. *J. Chrom. A.,* **1977,** *131,* 169-177. b. Ausserer, W.A. and Biros, M.L. *BioTechniques,* **1995,** *19*(1), 136-139.

[144] Gilar, M.; Fountain, K.J.; Budman, Y.; Neue, U.D.; Yardley, K.R.; Rainville, P.D.; Russell II, R.J. and Gebler, J.C. *J. Chrom. A.,* **2002,** *958,* 167-182.

[145] Damha, M.J. *Ph.D. Thesis,* **1988,** McGill University.

[146] a. Wada, S. and Sekine, M. *Tetrahedron Lett.,* **1994,** *35*(5), 757-760. b. Guzaev, A. and Manoharan, M. *J. Am. Chem. Soc.,* **2001,** *123*(5), 783-793. c. Cheruvallath, Z.S.; Eleuteri, A.; Turney, B. and Ravikumar, V.T. *Org. Process Res. Dev.,* **2006,** *10*(2), 251-256.

[147] Lonnberg, T.; Kiiski, J. and Mikkola, S. *Org. Biomol. Chem.,* **2005,** *3,* 1089-1096.

[148] a. Adlercreutz, P. *Eur. J. Biochem.,* **1991,** *199,* 609-614. b. Clough, J.M.; Jones, R.V.H.; McCann, H.; Morris, D.J. and Wills, M. *Org. Biomol. Chem.,* **2003,** *1,* 1486-1497.

[149] a. Ruskin, B. and Green, M.R. *Science,* **1985,** *229,* 135-140. b. Reed, R. and Maniatas, T. *Genes Dev.,* **1988,** *2,* 1268-1276. c. Zhuang, Y.; Goldstein, A.M. and Weiner, A.M. *Proc. Natl. Acad. Sci. USA,* **1989,** *86,* 2752-2576.

[150] Valadkhan, S. and Manley, J.L. *RNA,* **2003,** *9,* 892-904.

[151] a. Akhtar, S.; Hughes, M.D.; Khan, A.; Bibby, M.; Hussain, M.; Nawaz, Q.; Double, J. and Sayeed, P. *Adv. Drug Del. Rev.,* **2000,** *44,* 3-21. b. Snoussi, K.; Bulte, J.W.M.; Gueron, M. and van Zijl, P.C.M. *Magnetic Resonance Med.,* **2003,** *49*(6), 998-1005. c. Chaltin, P.; Margineanu, A.; Marchand, D.; Van Aerschot, A.; Rozenski, J.; De Schryver, F.; Herrmann, A.; Mullen, K.; Juliano, R.; Fisher, M.H.; Kang, H.; De Feyter, S. and Herdewijn, P. *Bioconjugate Chem.* **2005,** *16,* 827-836.

[152] Nilsen, T.W.; Grayzel, J. and Prensky, W. *J. Theor. Biol.,* **1997,** *187*(2), 273-284.

[153] a. Urdea, M.S.; Running, J.A.; Horn, T.; Clyne, J.; Ku, L. and Warner, B.D. *Gene,* **1987,** *61,* 253-264. b. Schepinov, M.S.; Udalova, I.A., Bridgman, A.J. and Southern, E.M. *Nucleic Acids Res.,* **1997,** *25(22),* 4447-4454.

[154] a. Newkome, G.R.; Moorefield, C.N. and Baker, G.R. *Aldrichim Acta.* **1992,** *25,* 31-37. b. Tomalia, D.A. *Aldrichim Acta.* **1993,** *26,* 91-101. c. Frechet, J.M. *Science* **1994,** *263,* 1710-1715.

[155] Carriero, S. and Damha, M.J. in *Current Protocols in Nucleic Acid Chemistry,* **2002,** 4.14.1-4.14.32.

[156] Buhleier, E.; Wehner, W. and Voegtle, F. *Synthesis,* **1978,** *2,* 155-158. b. Tomalia, D.A. *Aldrichimica Acta.* **2004,** *37(2),* 39-57.

[157] Chaix, C.; Radhakrishnan, P.I. and Agrawal, S. *Bioorg. Med. Chem. Lett.,* **1996,** *6(7),* 827-832.

[158] Kandimalla, E.R. and Agrawal, S. *Biochemistry,* **1996,** *35,* 15332-15339.

[159] a. Brown, D.M. and Todd, A.R. *J. Chem. Soc.,* **1952,** 52. b. Rayner, B.; Reese, C.B. and Ubagawa, J. *J. Chem. Soc.,* **1980,** 972. c. Kiviniemi, A.; Lonnberg, T. and Ora, M. *J. Am. Chem. Soc.,* **2004,** *126,* 11040-11045.

[160] Ooi, S.L.; Dann, C. III; Nam, K.; Leahy, D.; Damha, M.J. and Boeke, J.D. *Methods in Enzymology,* **2001,** *342,* 233-250.

[161] a. Carriero, S.; Mangos, M.M.; Agha, K.A.; Noronha, A.M. and Damha, M.J. *Nucleosides, Nucleotides and Nucleic Acids,* **2003,** *35(9),* 1599-1602. b. Mourani, R. and Damha, M.J. *Nuclesoides, Nucleotides and Nucleic Acids,* **2006,** *25(2),* 203-229. c. Peng, C.G. *Ph. D. Thesis,* **2007,** McGill University.

[162] a. Guerlavais-Dagland, T.; Meyer, A.; Imbach, J.L. and Morvan, F. *Eur. J. Org. Chem.*, **2003**, *12*, 2327-2335. c. Garjkowski, A.; Pedras-Vasconcelos, J.; Wang, V.; Ausin, C.; Hess, S.; Verthelyi, S. and Beaucage, S.L. *Nucleic Acids Res.*, **2005**, *33(11)*, 3550-3560.

[163] Sosnovtsev, S.V.; Belliot, G.; Chang, K.-O.K.; Prikhodko, V.G.; Thackray, L.B.; Wobus, C.E.; Karst, S.M.; Virgin, H.W. and Green, K.Y. *J. Virol.*, **2006**, *80*(16), 7816-7831.

[164] a. Sidiropoulos, P.; Liu, H.; Mungre, S.; Anderson, L.; Thimmapaya, B. and Pope, R.M. *Gene Therapy*, **2001**, *8*, 223-231. b. Stiriba, S-E.; Frey, H. and Haag, R. *Angew. Chem. Int. Ed,.* **2002**, *41*(8), 1329-1334. c. Virgin, H.W.; Wobus, C. and Karst, S. *U.S. Pat. Appl. Publ.* **2006**, 64 pp.Stiriba, S-E.; Frey, H. and Haag, R. *Angew. Chem. Int. Ed,.* **2002**, *41*(8), 1329-1334.

[165] Murer, P.K.; Lapierre, J.M.; Greiveldinger, G. and Seebach; D. *Helv. Chim. Acta*, **1997**, *80*(5), 1648-1681.

[166] Hassner, A.; Strand, G; Rubinstein, M. and Patchornik, A. *J. Am. Chem. Soc.*, **1975**, *97(6)*, 1614-1615.

[167] Garcia, J.; Fernandez, S.; Ferrero, M.; Sanghvi, Y.S. and Gotor, V. *J. Org. Chem.*, **2002**, *67*(13), 4513-4519.

[168] a. Moris, F. and Gotor, V. *J. Org. Chem.*, **1993**, *58*, 653-660. b. Meldgaard, M.; Neilsen, N.K.; Bremner, M.; Pedersen, O.S.; Olsen, C.E. and Wengel, J. *J. Chem. Soc. Perkin Trans. 1*, **1997**, 1951-1955. c. Garcia, J.; Fernandez, S.; Ferrero, M.; Sanghvi, Y. and Gotor, V. *Tet. Asymm.* **2003**, *14*, 3533-3540.

[169] Guzaev, A.P. and Manoharan, M. *J. Am. Chem. Soc.*, **2003**, *125*(9), 12380-12381.

[170] Murer, P. and Seebach, D. *Angew. Chem. Int. Ed. Engl.*, **1995**, *34*(19), 2116-2119.

[171] a. Noronha, A.M. *Ph.D. Thesis* **1999,** McGill University. b. Avino, A.; Grimau, M.G.; Freiden, M. and Eritja, R. *Helv. Chim. Acta*, **2004**, *87*(2), 303-316.

[172] a. Antao, V.P.; Lai, S.Y. and Tinocco; I.J. *Nucl. Acids Res.*, **1991**, *19*(21), 5901-5905. b. Hannoush, R.N. and Damha, M.J. *Nucleosides and Nucleotides*, **2005**, *25*, 1519-1530.

[173] Gray, D.M.; Hung, S.-H. and John, K.H. *Methods in Enzymology*, **1995**, *246*, 19-34.

[174] Gendron, D.; Carriero, S.; Garneau, D.; Villemaire, J.; Klinck, R.; Abou Elela, S.; Damha, M.J. and Chabot, B. *BMC Biotechnology,* **2006,** *6*(5), 12pp.

[175] Stavrianopoulos, G-G.A. and Chargaff, E., *Proc. Natl. Acad. Sci. USA,* **1976,** *73*, 1087-1091.

[176] Arts, E.J. and Wainberg, M.A. *Adv. Virus Res.,* **1998,** *46*, 97-163.

[177] Song, J-J.; Smith, S.K.; Hannon, G.J. and Joshua-Tor, L. *Science,* **2004,** *305*, 1434-1437.

[178] a. Okada, T.; Amanuma, H.; Okada, Y.; Obata, M.; Hayashi, Y.; Yamaguchi, K. and Yamashita, J. *Biochem. and Biophys. Res. Comm.,* **1997,** *240*, 203-207. b. Opalinska, J.B. and Gewirtz, A.M. *Nature Rev. Drug Disc.,* **2002,** *1*, 503-514. c. Janowski, B.A.; Huffman, K.E.; Schwartz, J.C.; Ram, R.; Hardy, D.; Shames, D.S.; Minna, J.D. and Corey, D. R. *Nature: Chem. Biol.,* **2005,** *4*(1), 216-22.

[179] Hammond, S.M.; Boettcher, S.; Caudy, A.A.; Kobayashi, R. and Hannon, G.J. *Science,* **2001,** *293*, 1146-1150.

[180] a. Li, T-K.; Barbieri, C.M.; Lin, H.-C.; Rabson, A.B.; Yang, G.; Fan, Y.; Gaffney, B.L.; Jones, R.A. and Pilch, D.S. *Biochemistry,* **2004,** *43*, 9732-9742. b. Hanoush, R. N.; Carriero, S.; Min, K-L. and Damha, M.J. *ChemBioChem,* **2004,** *5*, 527-533.

[181] Walder, R.Y. and Walder, J.A. *Proc. Natl. Acad. Sci. USA,* **1988,** *85*(14), 5011-5015.

[182] Song; J-J., Smith; S.K., Hannon; G.J. and Joshua-Tor; L. *Science,* **2004,** *305*, 1434-1437.

[183] Crooke, S.T.; Lemonidis, K.M.; Neilson, L.; Griffey, R.; Lesnik, E.A. and Monia, B.P. *Biochemistry Journal,* **1995,** *312*, 599-606.

[184] Stavrianopoulos; G-G.A. and Chargaff; E., *Proc. Natl. Acad. Sci. USA,* **1976,** *73*, 1087-1091.

[185] Minshull, J. and Hunt, T. *Nucleic Acids Res.,* **1986,** *14*, 6433-6451.

[186] Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.L.; Thoung, N.T.; Lhomme, J. and Helene, C. *Nucleic Acids Res.,* **1987,** *15*, 7749-7760.

[187] Shoji, A.; Kuwahara, M.; Ozaki, H. and Sawai, H. *J. Am. Chem. Soc.,* **2007,** *129*(5), 1456-1464.

[188] Itaya, M. and Kondo, K. *Nucleic Acids Res.,* **1991,** *19*, 4443-4449.

[189] Gabb, H.A. and Harvey, S.C. *J. Am. Chem. Soc.,* **1993,** *115,* 4218-4227.

[190] Choo, J.; Lee, S-N. and Lee, K-H. *Bull. Kor. Chem. Soc.,* **1996,** *17*(1), 7-11.

[191] a. Ferrier, R.J. and Sankey, G.H. *J. Chem. Soc. C,* **1966,** 2345-2349. b. Laane, J. and Choo, J. *J. Am. Chem. Soc.,* **1994,** *116,* 3889-3891.

[192] Felder, E.; Gattlen, R.; Ossola, F. and Baschang, G. *Nucleosides Nucleotides,* **1992** *11*(9), 1667-1671.

[193] DeMatteo, M.P.; Snyder, N.L.; Morton, M.; Baldisseri, D. M.; Hadad, C. M.; Peczuh, M.W.; *J. Org. Chem.,* **2005,** *70,* 24-38.

[194] a. Pauling, L. and Corey, R.B. *Proc. Natl. Acad. Sci. USA,* **1953,** *39*(2), 84-97. b. Allen, F.H.; Kennard, O.; Watson, D.G.; Brammer, L.; Orpen, A.G. and Taylor, R. *J. Chem. Soc. Perkin Trans. 2,* **1987,** S1-S19. c. Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S-H.; Olson, W.K. and Berman, H.M. *J. Am. Chem. Soc.,* **1996,** *118,* 519-529.

[195] Ichikawa, E. and Kato, K. *Curr. Med. Chem.,* **2001,** *8,* 385-423.

[196] Crooke, S.T. *Annu. Rev. Med.,* **2004,** *55,* 61-95.

[197] Goodchild, J. *Bioconjugate Chem.,* **1990,** *1*(3), 165-187.

[198] Danishefsky, S.J. and Bilodeau, M.T. *Angew. Chem., Int. Ed. Eng.,* **1996,** *35*(13/14), 1380-1419.

[199] Dorgan, B.J. and Jackson, R.F.W. *Synlett,* **1996,** 859-861.

[200] Di Bussolo, V.; Caselli, M.; Pineschi, M. and Crotti, P. *Org. Lett.,* **2002,** *4*(21), 3695-3698.

[201] a. Leutzinger, E.E.; Meguro, T. and Townsend, L.B. *J. Org. Chem.,* **1972,** *37*(23), 3695-3703. b. El-Laghdach, A.; Matheu, M.I. and Castillon, S. *Tetrahedron,* **1994,** *50*(42), 12219-12234.

[202] Gonzalez, R.; Lesage, S. and Perlin, A. *Carbohydrate Res.,* **1975,** *42,* 267-274.

[203] Tatsuta, K.; Fujimoto, K. and Kinoshita, M. *Carbohydrate Res.,* **1977,** *54,* 85-104.

[204] a. Ferrier, R.J. and Prasad, N. *J. Chem. Soc.,* **1969,** 570-575. b. Ferrier, R.J. and Zubkov, O.A. *Org. React.,* **2003,** *62,* 569-736.

[205] a. Di Bussolo, V.; Caselli, M.; Pineschi, M. and Crotti, P. *Org. Lett.,* **2003,** *5*(12), 2173-2176. b. Di Bussolo, V.; Caselli, M.; Romano, R.; Pineschi, M. and Crotti, P. *J. Org. Chem.,* **2004,** *69,* 8702-8708. and references therein.

227

[206] Zhang, P.; Zhang, N.; Korba, B.E. and Hosmane, R.S. *Bioorg. Med. Chem. Lett.*, **2005**, *15*, 5397-5401.

[207] Sushil, R. and Bhanu, P.B.A. *Arkivoc*, **2003**, *3*, 16-24.

[208] Ueda, T. and Watanabe, S. *Chem. & Pharm. Bull.*, **1985**, *33*(9), 3689-3695.

[209] Herscovici, J.; Montserret, R. and Antonakis, K. *Carbohydrate Res.*, **1988**, *176*, 219-229.

[210] Kondo, T.; Nakai, H. and Goto, T. *Agric. Biol. Chem.*, **1971**, *35*(12), 1990-1991.

[211] Tari, L.W. and Secco, A.S. *Can. J. Chem.*, **1992**, *70*, 894-899.

[212] Tari, L.W.; Sadana, K.L. and Secco, A.S. *Nucleosides & Nucleotides*, **1995**, *14*(1&2), 175-183.

[213] Kocienski, P.; Love, C.; Whitby, R. and Roberts, D.A. *Tet. Lett.*, **1988**, *29*, 2867-2870.

[214] Yasumoto, T. and Murata, M. *Chem. Rev.*, **1993**, *93*, 1897-1909.

[215] Hoberg, J.O. *J. Org. Chem.*, **1997**, *62*, 6615-6618.

[216] Hoberg, J.O. *Carbohydrate Res.*, **1997**, *300*, 365-367.

[217] Skalmowski-Bennua, B.; Krolikiewicz, K. and Vorbrüggen, H. *Tetrahedron Lett.*, **1995**, 36(43), 7845-7848.

[218] Elzagheid, M.I.; Viazovkina, E. and Damha, M.J. *Current Protocols in Nucleic Acid Chem.*, **2002**, 1.7.1-1.7.19.

[219] Ferrier, R.J. and Ponpipom, M.M. *J. Chem. Soc. C*, **1971**, *3*, 553-559.

[220] Pedersen, H.; Pedersen, E.B. and Neilsen, C.M. *Heterocylces*, **1992**, *34*(2), 265-272.

[221] a. Furukawa, J.; Kawabata, N. and Nishimura, J. *Tetrahedron*, **1968**, *24*, 53-58. b. Hoveyda, A.H. Evans, D.A. and Fu, G.C. *Chem. Rev.*, **1993**, *93*, 1307-1370.

[222] a. Bocain, D.F. and Straus, H.L. *J. Am. Chem. Soc.*, **1977**, *99*, 2876-2882. b. Jennings, B.W.; Rutherford, M.; Agarwal, S.K.; Boyd, D.R.; Malone, J.F. and Kennedy, D.A. *J. Chem. Soc. Chem. Comm.*, **1986**, 970-972. c. Espinosa, A.; Gallo, M.A.; Entrena, A. and Gomez, J.A. *J. Mol. Struct.*, **1994**, *323*, 247-256.

[223] a. Jakobs, J.J. and Sundaralingam, M. *J. Chem. Soc. Chem. Comm. D*, **1970**, *3*, 157-158. b. Stevens, J.D.; Beale, J.P. and Stephenson, N.C. *J. Chem. Soc. Chem. Comm. D*, **1971**, *10*, 484-486. c. Chou, W.N.; White, J.B. and Smith, W.B. *J. Am. Chem. Soc.*,

1992, *114*(12), 4658-4667. d. Peczuh, M.W.; Snyder, N.L. and Fyvie, W.S. *Carbohydrate Res.,* **2004**, *339*, 1163-1171.

[224] Balasubramanian, S.; Abell, C. and Coggins, J.R. *J. Am. Chem. Soc.,* **1990**, *112*, 8581-8583.

[225] Kerremans, L.; Schepers, G.; Rozenski, J.; Busson, R.; Van Aerschot, A. and Herdewijn, P. *Org. Lett.,* **2001**, *3*(26), 4129-4132.

[226] Benevides, J.M. and Thomas Jr., G.J. *Biochemistry,* **1988**, *27*, 3868-3873.

[227] Aymaan, J.; Coll, M.; Frederick, C.A.; Wang, A.H-J. and Rich, A. *Nucleic Acids Res.,* **1989**, *17*, 3229-3245.

[228] Steely, H.T.; Gray, D.M. and Ratliff, R.L. *Nucleic Acids Res.,* **1986**, *14*, 10071-10090.

[229] Skalmowski, B. B.; Krolikiewicz, K.; Vorbrüggen, H. *Tetrahedron Lett.,* **1995**, *36*(43), 7845-7848.

[230] Distler, A.M. and Allison, J. *Anal. Chem.,* **2001**, *73*, 5000-5003.

[231] a. Kim, K. and Jhon, M.S. *Biochim. Phys. Acta,* **1979**, *565*, 131. b. Martin, F.H. and Tinoco, I.J. *Nucleic Acids Res.,* **1980**, *8*, 2295-2299. c. Aymaan, J.; Coll, M.; Frederick, C.A.; Wang, A.H-J. and Rich, A. *Nucleic Acids Res.,* **1989**, *17*, 3229-3245. d. Hall, K.B. and McLaughin, L.W. *Biochemistry,* **1991**, *30*, 10606-10613. e. Ratmeyer, L.; Vinayak, R.; Zhong, Y.Y.; Zon, G. and Wilson, W.D. *Biochemistry,* **1994**, *33*, 5298-5304.

[232] a. Job, P. *Ann. Chim. (Paris),* **1928**, *9*, 113-203. b. Pilch, D.S.; Levenson, C. and Shafer, R.H. *Proc. Natl. Acad. Sci. USA,* **1990**, *87*, 1942-1946.

[233] Chan, S.S.; Breslauer, K.K.; Hogan, M.E.; Kessler, D.J.; Austin, R.H.; Ojemann, J.; Passner, J.M. and Wiles, N.C. *Biochemistry,* **1990**, *29*, 6161-6171.

[234] a. Nejedly, K.; Klysik, J. and Palecek, E. *FEBS Lett.,* **1989**, *243*(2), 313-317. b. Kagawa, T. F.; Howell, M. L.; Tseng, K. and Ho, P.S. *Nucleic Acids Res.,* **1993**, *21*(25), 5978-5986.

[235] Tomasz, M.; Barton, J.K.; Magliozzo, C.C.; Tucker, D.; Lafer, E.M. and Stollar, B.D. *Proc. Natl. Acad. Sci. USA,* **1983**, *80*, 2874-2878.

[236] Praseuth, D.; Perrouault, L.; LeDoan, T.; Chassignol, M.; Thuong, N. and Helene, C. *Proc. Natl. Acad. Sci. USA,* **1988**, *85*, 1349-1353.

[237] Lok, C-N; Viazovkina, E.; Min, K-L.; Nagy, E.; Wilds, C.J.; Damha, M.J. and Parniak, M.A. *Biochemistry,* **2002**, *41*, 3457-3467.

[238] a. Verbeure, B.; Lescrinier, E.; Wang, J. and Herdewijn, P. *Nucleic Acids Res.,* **2001** *29*(24), 4941-4947. b. Yazbeck, D.R.; Min, K-L. and Damha, M.J. *Nucleic Acids Res.,* **2002**, *30*(14), 3015-3025.

[239] a. Uchiyama, Y.; Miura, Y.; Inoue, H.; Ohtsuka, E.; Ueno, Y.; Ikehara, M. and Iwai, S. *J. Mol. Biol.,* **1994**, *243*, 782-791. b. Nishizaki, T.; Iwai, S.; Ohtsuka, E. and Nakamura, H. *Biochemistry,* **1997**, *36*, 2577-2585.

[240] Sorensen, M. D.; Kvaerno, L.; Bryld, T.; Hakansson, A. E.; Verbeure, B.; Gaubert, G.; Herdewijn, P. and Wengel, J. *J. Am. Chem. Soc.,* **2002**, *124*, 2164-2176.

[241] Baker, B. F. and Monia, B. P. *Biochim. Biophys. Acta,* **1999**, *1489*, 3-18.

[242] Nielsen, P.; Pfundheller, H. M. and Wengel, J. *Chem. Comm.,* **1997**, 825-826.

[243] Li, J. and Wartell, R. M. *Biochemistry,* **1998**, *37*, 5154-5161.

[244] a. Trotta, C.R. *Nature,* **2007**, *447*, 156-157. b. Weitzer, S. and Martinez, J. *Nature* **2007**, *447*, 222-227.

[245] Kandimalla, E.R.; Manning, A.; Zhao, Q.; Shaw, D.R.; Byrn, R.A., Sasisekharan, V. and Agrawal, S. *Nucleic Acids Res.,* **1997**, *25*(2), 370-378.

[246] Marquez, V.E.; Tsipi, B-K.; Barchi, J.J.; Green, K.M.; Nicklaus, M.C. and Agbaria, R. *J. Am. Chem. Soc.,* **2004**, *126, 543-549.*

[247] Gunaga, P.; Moon, H.R.; Choi, W.J.; Shin, D.J.; Park, J.G. and Jeong, L.S. *Curr. Med. Chem.,* **2004**, *11*, 2585-2637.

[248] Schinazi, R.F. and Liotta, D.C. *Frontiers in Nucleosides and Nucleic Acids*, IHL Press, **2004**.

[249] a. Mitsuya, H.; Yarchoan, R. and Broder, S. *Science,* **1990**, *249*, 1533-1544; b. Merluzzi, V.J., *Science,* **1990**, *250*, 1411-1413; c. Kohlstaedt, L.A.; Wang, J.; Friedman, J.M.; Rice, P.A. and Steitz, T.A. *Science*, **1992**, *256*, 1783-1790.

[250] Joshi, B.V.; Moon, H-R.; Fettinger, J.C.; Marquez, V.E. and Jacobson, K.E. *J. Org. Chem.,* **2005**, *70*, 439-447.

[251] Vanheudsen, V.; Busson, R.; Herdewijn, P. and Carlenbergh, S.V. *J. Org. Chem.*, **2004**, *69*(13), 4446-4453.

[252] Achmatowicz, O. and Burzynska, M.H. *Carbohydrate Res.*, **1985**, *141*, 67-76.

[253] Wong, J.C.Y.; Lacombe, J. and Sturino, C. *Tetrahedron Lett.*, **1999**, *40*, 8751-8754.

[254] Corsaro, A.; Chiacchio, U.; Adamo, R.; Pistara, V.; Rescifina, A.; Romeo, R.; Catelani, G.; D'Andrea, F.; Mariani, M. and Attolino, E. *Tetrahedron*, **2004**, *60*, 3787-3795.

[255] Murali, R. and Nagarjan, M. *Carbohydrate Res.*, **1996**, *280*, 351-355.

[256] a. Guerlavais-Dagland, T.; Meyer, A.; Imbach, J.L. and Morvan, F. *Eur. J. Org. Chem.*, **2003**, *12*, 2327-2335. c. Garjkowski, A.; Pedras-Vasconcelos, J.; Wang, V.; Ausin, C.; Hess, S.; Verthelyi, S. and Beaucage, S.L. *Nucleic Acids Res.*, **2005**, *33(11)*, 3550-3560.

[257] Gasparutto, D.; Livache, T.; Bazin, H.; Duplaa, A.M.; Guy, A.; Khorlin, A.; Molko, D.; Roget, A.; Teoule, R. *Nucleic Acids Res.*, **1992**, *20*, 5159-5166.

[258] Sproat, B.; Colonna, F.; Mullah, B.; Tsou, D.; Andrus, A.; Hampel, A. and Vinayak, R. *Nucleosides & Nucleotides*, **1995**, *14*(1 & 2), 255-273.

[259] a. Porath, J.; Flodin, P. *Nature*, **1959**, *183*, 1657-1659. b. Porath, J. *Clinica Chim. J.* **1959**, *4*, 776-778.

[260] Puglisi, J.D. and Tinoco, I. Jr. *Methods Enzymol.*, **1989**, *180*, 304-325.

[261] NMR analysis indicated that $NaHCO_3$ is not sufficient in completely neutralizing the excess acid into the aqueous layer. The work-up required more basic conditions to completely eliminate the residual acid and this can also result in the ring opening reactions of the epoxide.