SYNTHESIS, CHARACTERIZATIONS AND APPLICATIONS OF C2'-MODIFIED OLIGONUCLEOTIDE ANALOGUES

Chang Geng Peng

Department of Chemistry
McGill University
Montreal, Canada

February 2007

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

Your file Votre référence
ISBN: 978-0-494-32317-5
Our file Notre référence
ISBN: 978-0-494-32317-5

This thesis is dedicated to my wife Mu Qing

Abstract

During the past two decades, oligonucleotide analogues have drawn considerable attention as potential therapeutic and diagnostic agents. Gene silencing through "RNA interference" (siRNA) or the more mature "antisense" technology (AONs) have proven to be powerful tools for studying gene functions. Chemical modifications of these compounds are generally required to improve their "drug-like" properties such as potency, selectivity and delivery, particularly in the development of oligonucleotide-based therapeutics. Aptamers are another emerging class of oligonucleotide therapeutics and diagnostics.

This thesis focuses on oligonucleotides containing 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (2'-α-*hm*-dT, abbreviated as "**H**") and 2'-deoxy-2'-fluoroarabinonucleotides (2'F-araN), and their applications. A major component of this work focused on the synthesis of 2'-α-*hm*-dT (**H**) and the first investigation of oligoribonucleotides containing this nucleoside analogue. Specifically, 2'-CH_2O-phosphoramidite and 3'-*O*-phosphoramidite derivatives of **H** were synthesized and incorporated into both 2',5'-RNA and RNA chains. Incorporation of 3',5'-linked **H** units into a DNA, 2',5'-RNA or RNA strand led to significant destabilization of duplexes formed with unmodified RNA targets. 2',5'-Linked **H** units into 2',5'-RNA or RNA caused significantly less destabilization, and in fact, they were shown to stabilize the loop structure of some RNA hairpins. These results were rationalized in terms of the "compact" and "extended" conformations of nucleotides.

A series of branched RNAs (Y-shaped) related to yeast pre-mRNA splicing intermediates were synthesized incorporating both natural (i.e., ribose) and nonnatural (i.e., **H**, and acyclic nucleoside) branch points in order to examine the effect of sugar conformation and phosphodiester configuration on yeast debranching enzyme (yDBR) hydrolytic efficiency. The results indicate that 2'-phosphodiester scission with yDBR occurs only with a ribose-phosphate backbone at the branch point, whereas some of the **H**-containing branched RNAs were found to competitively inhibit yDBR hydrolytic activity.

This thesis also examines the stabilization of DNA guanine-quadruplexes (G-quadruplexes) by replacing the deoxyribose sugar by a 2-deoxy-2-fluoroarabinose. The effect of this substitution was assessed in the well-known thrombin-binding DNA aptamer $d(G_2T_2G_2TGTG_2T_2G_2)$, the telomeric DNA $d(G_4T_4G_4)$ sequence and a phosphorothioate octanucleotide PS-$d(T_2G_4T_2)$, all of which are known to fold into G-quadruplex structures. Stabilization of the G-quadruplexes was possible provided that the arabinose sugar was introduced at guanosine residues adopting an *anti* *N*-glycosidic bond conformation. Some of the arabinose modified thrombin-binding aptamers not only exhibited superior thermal stability and nuclease resistance, but also maintained high thrombin binding affinity.

Finally, this thesis examines the ability of DNA polymerases to recognize and utilize 2'-deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs) as building blocks for the synthesis of 2'-deoxy-2'-fluoro-β-D-arabinonucleic acids (2'F-ANA). The results obtained indicate that a few DNA polymerases can synthesize 2'F-ANA and 2'F-ANA-DNA chimeras on a DNA template. Conversely, certain enzymes were shown to catalyze 2'F-ANA template-directed DNA synthesis. While it was not possible to synthesize 2'F-ANA strands on a 2'F-ANA template, it is possible for some DNA polymerases to catalyze the formation of multiple 2'F-ANA:2'F-ANA base pairs within a DNA-FANA chimeric duplex. These results suggest that it should be possible to evolve FANA-modified aptamers *via* SELEX.

Résumé

Au cours des deux dernières décennies, une attention particulière a été portée à l'utilisation des oligonucléotides tant sur le plan diagnostique que thérapeutique. Par exemple, l'inactivation épigénétique post-transcriptionnelle comme l'ARN interférant (ARNi) et la technologie plus maîtrisée des oligonucléotides antisens (OA) se sont révélés de bons outils pour l'étude de la fonction des gènes. Des modifications chimiques sont généralement nécessaires pour améliorer leurs propriétés thérapeutiques, notamment l'efficacité, la sélectivité et la libération, en particulier dans la mise au point des thérapies basées sur des oligonucléotides. Les aptamères représentent une autre façon émergente d'utiliser les oligonucléotides sur les plans diagnostique et thérapeutique.

Dans cette thèse, nous nous concentrons sur deux modifications chimiques importantes, la 1-(2-désoxy-2-α-hydroxyméthyl-β-D-ribofuranosyl)thymine (ou 2'-α-*hm*-dT, abrégé **H**) et l'acide 2'-désoxy-2'-fluoro arabinonucléique (2'-AFAN), et leurs applications. Une des parties importantes de la thèse concerne la synthèse de **H** et la première investigation des oligoribonucléotides contenant cette modification chimique. Spécifiquement, les dérivés 2'-CH_2O-phosphoramidite et 3'-*O*-phosphoramidite de **H** ont tous deux été synthétisés et incorporés dans des chaînes de 2',5'-ARN et d'ARN. L'incorporation de la modification **H** lié en 3',5' dans les brins d'ADN, ARN ou 2',5'-ARN déstabilisait de façon significative les duplexes formés avec un brin d'ARN complémentaire. Par contre, le **H** lié en 2',5' a un effet déstabilisant moindre sur les duplexes 2',5'-ARN/ARN et ARN/ARN, et peut même stabiliser les boucles de quelques structures d'ARN en épingle à cheveux. Ces résultats ont été rationalisés en termes de conformations compactes ou allongées des nucléotides.

On a synthétisé une série de ARNs branchés (en forme de Y) qui sont similaires aux intermèdes de l'épissage pre-ARNm en levure, qui incorporent les points de branchement naturels (i.e. ARN) et modifiés (i.e. **H**, et nucléoside acyclique), pour examiner l'effet de la conformation du sucre et la configuration du phosphodiester sur l'efficacité de l'hydrolyse de l'enzyme du débranchement en levure (yDBR). Les résultats indiquent que le clivage du 2'-phosphodiester est possible seulement avec un

squelette naturel (ribose-phosphate) aux points de branchement. Certains des ARNs branchées qui contenaient **H** sont des inhibiteurs compétitifs de yDBR.

Une autre partie de la présente thèse traite de la stabilisation de la structure en quadruplexe de guanine (G-quadruplexe) par le remplacement du sucre désoxyribose par 2'-désoxy-2'-fluoroarabinose. Cet effet a été souligné dans un aptamère ADN de liaison avec la thrombine $d(G_2T_2G_2TGTG_2T_2G_2)$, dans un télomère ADN $d(G_4T_4G_4)$ ainsi que dans un octanucléotide phophorothioate PS-$d(T_2G_4T_2)$, qui possèdent tous des structures G-quadruplexe. On observe une stabilisation et un maintien de la conformation G-quadruplexe tant que les unités 2'-AFAN remplacent les guanosines *anti* (*anti*-Gs). Il a aussi été démontré que les aptamères de liaison avec la thrombine modifiés par des unités 2'-AFAN offrent non seulement une stabilité thermale et une résistance aux nucléases supérieures, mais aussi un maintien de l'affinité de liaison avec la thrombine.

Finalement, cette thèse examine la capacité des polymérases de reconnaître et utiliser les arabinonucléosides 2'-désoxy-2'-fluoro-5'-triphosphates (2'F-araNTP) comme précurseurs pour la synthèse de l'acide 2'-désoxy-2'-fluoroarabinonucléique (2'-AFAN). On a trouvé que certaines polymérases peuvent synthétiser la 2'-AFAN ou les chimères ADN-2'-AFAN sur un précurseur de brin en ADN. Inversement, certaines enzymes ont pu catalyser la synthèse de l'ADN sur un précurseur de brin en 2'-AFAN. Tandis qu'il n'est pas possible de synthétiser la 2'-AFAN sur un précurseur de brin en 2'-AFAN, certaines polymérases ont la capacité de catalyser la formation de plusieurs paires de bases 2'-AFAN: 2'-AFAN au sein d'un duplexe chimérique ADN: 2'-AFAN. Ces résultats suggèrent qu'il serait possible de créer des aptamères contenant de la 2'-AFAN par l'évolution *in vitro* (SELEX).

Acknowledgements

Confucius said that *you could find a teacher among any two people you walk with*, because one person cannot know everything, but everyone must know something. It is very true that I cannot do things alone. I feel very grateful to have met many nice people who all taught me something on different occasions. My first deep appreciation goes to my supervisor Dr. Masad Damha, a great nucleic acid chemist. I am very lucky that he accepted me as a student and passed on to me not only his wealthy knowledge in science but also his valuable life experiences. His inspiration, guidance, patience and help were always companying me in every step I made. His sense of humour and deep understanding mean a lot to me. His great teaching from class presentations will always be a model for me to follow.

Special thanks to Dr. George Just for his kindness to me and always support for my applications. His encouragement during my study at McGill is greatly appreciated!

I also want to thank Dr. Patrick Farrell for his laughs and encouragement in my first year, Dr. C.J. Li for sharing with me his experience on how to become a professor, Dr. Karine Auclair for her encouragement in my presentation and also for letting me use her CD instrument.

I owe many thanks to Dr. Paul Xia for teaching me how to use NMR, Nadim Saade for training me how to use MS, Antisar Hlal for running MALDI-TOF MS experiments for me, Dr. Jerry Pelletier in Biochemistry Department for lending me the nitrocellulose filter binding apparatus, Dr. B. Schwer at Cornell University for sending us a debranching enzyme as a gift, Dr. Helene D'Anjou at Topigen for showing me how to operate the real-time PCR. I thank Dr. Yingfu Li at McMaster University for accommodating me in his lab last July, and Dr. Srinivas Kandadai in Li's lab for teaching me how to run SELEX.

I want to thank Chantal Marotte, Sandra Aerssen and Fay Nurse for making my graduate life convenient.

I would like to show my deep thanks to both the former and current members of Dr. Damha's group. I thank Drs. Rami Hannoush, Kyung-Lyum Min, Ekaterina Viazovkina, Mohammad El-Zagheid and Kazim Agha for their help in my first year. I

thank Drs. Maria Mangos and Sandra Carriero for sharing with me the experience of being a graduate student and helping me proofread my writings. I also thank Benedicte Patureau and Jonathan Watts for translating my abstract from English to French. I thank Anna-Lisa Tedeschi for helping me synthesize some oligonucleotides at Topigen. I especially thank the current members Robert Donga, Siara Isaac, Jeremy Lackey, Debbie Mitra, David Sabatino, Alexander Wahba and Jonathan Watts for their friendship and all kinds of help, especially in my times of difficulty; they made my graduate life very enjoyable. "Chang Geng, I give you ten points for your performance" was such a wonderful encouragement. I am so grateful to have worked with these great colleagues, scientists and friends. In my thesis writing period, Jonathan, Robert, Alex, Siara, Jeremy and Maria helped me proofread my manuscripts in their busy time. Your help means a lot to me! My wonderful experience in Damha's lab will be rooted in my head as beautiful memories in my life.

I would like to thank Jonathan Watts and David Sabatino again for their friendship. We have shared many important moments in our graduate studies. Their ever-readiness to help, understanding and encouragement touches me.

I also thank Dr. Auclair's group and Dr. Sleeman's group for their help. My best wishes for you all.

I appreciate McGill Clifford Wang Major Fellowship and Le Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) for their financial support for my graduate studies, and McGill and Department of Chemistry for traveling awards.

I thank the Chinese Christian Gospel Church for being a big family to me, and for the many wonderful moments we have spent together in the past five years.

I thank my family members in China for their unconditional love and patience.

And certainly not least of all – thank you, Mu Qing!

Contributions of Authors

This thesis represents a collection of five manuscripts. Chapter II has been published in *Nucleic Acids Research*; Chapter V has been submitted to *the Journal of the American Chemical Society*; the work described in the remaining Chapters (Chapters III, IV, and VI) will soon be submitted for publication. All the work was independently conducted by me under the supervision of my research supervisor, Prof. Masad J. Damha. Dr. Maria M. Mangos, a former graduate student in Prof. Damha's research group, provided the *seco*-U nucleoside used for debranching inhibition studies described in Chapter III.

Table of Contents

Abbreviations and Symbols

λ	wavelength
®	register trademark
Å	angstrom
A	adenosine
A_{260}	UV absorbance measured at 260 nm
Ac	acetyl
$AgNO_3$	silver nitrate
AMD	age-related macular degeneration
ANA	arabinonucleic acid
AON	antisense oligonucleotide(s)
ap	antiperiplanar
ara	arabino
araF	2'-deoxy-2'-fluoro-arabino
AS-ODNs	antisense oligodeoxynucleotides
B	base
bFGF	basic fibroblast growth factor
Bn	benzyl
bp	base pairs
BPB	bromophenol blue
br	broad
bRNA	branched RNA
Bst	*Bst* DNA polymerase large fragment
Bz	benzoyl
Bz-Cl	benzoyl chloride
°C	Celsius degree
CaH_2	calcium hydride
CD	circular dichroism
CEO	2-cyanoethoxy (β-cyanoethoxyl)
CH_2Cl_2	dichloromethane
CH_3CN	acetonitrile
COSY	correlated spectroscopy
CPG	controlled pore glass
d	doublet
DCE	1,2-dichloroethane
DCI	4,5-dicyanoimidazole
DCM	dichloromethane
dd	doublet of doublets
ddd	doublet of doublet of doublets
DIPEA	*N*-ethyl-*N,N*-diisopropylamine
DMAP	*N,N*-dimethyl-4-aminopyridine
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulphoxide
DMT	dimethoxytrityl

DNA	2'-deoxyribonucleic acid
dNTPs	deoxyribonucleoside 5'-triphosphates
dsRNA	double-stranded RNA
DTT	dithiothreitol
DV	Deep Vent (3'→5' exo-) DNA polymerase
E. coli	*Escherichia coli*
e.g.	for example
EDTA	disodium ethylenediaminetetraacetate dihydrate
eq.	equivalent(s)
EtOAc	ethyl acetate
EtOH	ethanol
2'F-ANA	2'-deoxy-2'-fluoro-β-D-arabinonucleic acid
2'F-araN	2'-deoxy-2'-fluoro-β-D-arabinonucleoside
2'F-araNTP	2'-deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphate
2'F-rN	2'-deoxy-2'-fluoro-β-D-ribonucleoside
2'F-RNA	2'-deoxy-2'-fluoro-β-D-ribonucleoside
2'F-RNA	2'-deoxy-2'-fluoro-β-D-ribonucleic acid
2'F-rNTP	2'-deoxy-2'-fluoro-β-D-ribonucleoside 5'-triphosphate
g	gram(s)
G	guanosine
GMP	guanosine monophosphate
G-quartet	guanine quartet
anti-Gs	*anti* glycosidic guanosines
syn-Gs	*syn* glycosidic guanosines
h	hours
HIV-1	human immunodeficiency virus type 1
hKGF	human keratinocyte growth factor
HMDS	1,1,1,3,3,3-hexamethyldisilazane
HOAc	glacial acetic acid
HPLC	high performance liquid chromatography
Hz	Hertz
i.e.	that is
i-Bu	isobutyryl
IFN-γ	interferon γ
IDT	Integrated DNA Technology
J	coupling constant
KF	Klenow fragment DNA polymerase (3'→ 5' exo-)
LCAA-CPG	Long chain alkylamine controlled pore glass
M	molar
m	multiplet
MALDI TOF	matrix assisted laser desorption ionization time of flight
max	maximum
Me	methyl
MeOH	methanol
mg	minigram
min	minute(s)

mL	miniliter
mM	minimolar
MMLV	Moloney Murine Leukemia Virus reverse transcriptase
MMT	monomethoxytrityl
mRNA	message RNA
MS	mass spectrometry
MW	molecular weight
9N	9° N_m™ DNA polymerase
N	north
NEB	New England Biolabs
NF-IL6	nuclear factor for human interleukin 6
nm	nanometer
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enhancement
NOESY	nuclear Overhauser and exchange spectroscopy
NTPs	nucleoside 5'-triphosphates
O.D.U.	optical density unit
PAGE	polyacrylamide gel electrophoresis
PCR	polymerization chain reaction
Pd/C	palladium on carbon powder
Ph	Phusion™ High-Fidelity DNA polymerase
ppm	part per million
pre	precursor
Pu	purine
py	pyridine
q	quartet (NMR)
R_f	(TLC mobility) retardation factor
RNase	ribonuclease
rNTPs	ribonucleoside 5'-triphosphate
rRNA	ribosomal RNA
RNAi	RNA interference
rt	room temperature
RT	reverse transcriptase
s	singlet (NMR)
S	south
sc	synclinal
sec	second
SELEX	systematic evolution of ligands by exponential enrichments
siRNA	small interfering RNA
$SnCl_4$	tin (IV) tetrachloride
SVPDE	snake venom phosphodiesterase
t	tertiary, triplet (NMR)
T	thymidine
Taq	*Taq* DNA polymerase
TBDMSCl	*tert*-butyldimethylsilyl chloride
TBDMSOTf	*tert*-butyldimethylsilyl triflate

Th	Therminator™ DNA polymerase, a mutated 9N
THF	tetrahydrofuran
TNA	α-L-threofuranosyl nucleic acids
T-NPP	nitrophenylpropyl
TREAT HF	triethylamine tris(hydrofluoride)
μM	micromolar
UV	ultraviolet
VEGF	vascular endothelial growth factor
VPF	vascular permeability factor
yDBR	yeast debranching enzyme

CHAPTER I

CHAPTER I: General Introduction

1.1.Importance of Nucleic Acids[1,2]

The elucidation of the double helical structure of deoxyribonucleic acid (DNA) by Watson and Crick[3] signified the beginning of our understanding of life at a molecular level. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) play significant roles as the information carrying molecules in cells. They are known to be involved in a pathway of information flow within cells that has been termed the *Central Dogma of Molecular Biology*.[4] Briefly, DNA is the storage medium of all genetic information and passes that information along through RNA, via the transcription process, to protein, via the translation of RNA. Transcription is the process by which an RNA molecule is synthesized on a DNA template by an enzyme RNA polymerase. Translation is the process by which an amino acid sequence (polypeptide) is assembled from an RNA molecule. DNA can also make copies of itself for cell propagation via the replication process catalyzed by DNA polymerase and faithfully transfer the genetic information from one generation to another. One notable exception to the central dogma is found in retroviruses, such as human immunodeficiency virus (HIV), where viral genetic information is stored as RNA and is first reverse transcribed into DNA for eventual integration into the host DNA.

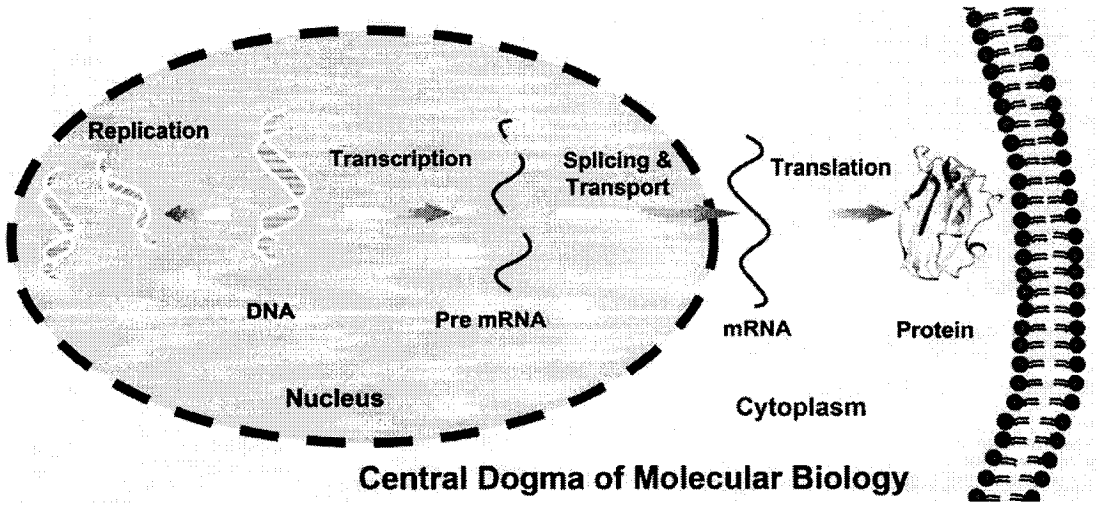


Figure 1.1: *Central Dogma of Molecular Biology* for the information flow.

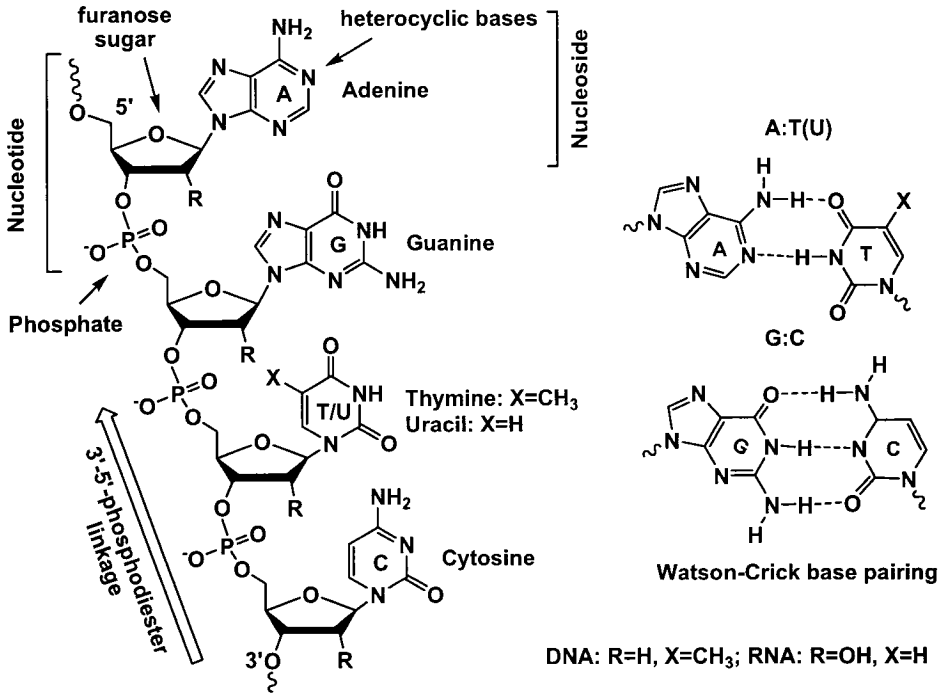


Figure 1.2: Primary structure of 2'-deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) and the Watson-Crick base paring of complementary nucleoside bases (A:T(U); G:C)

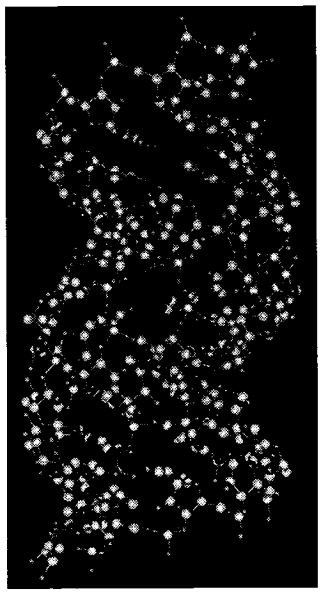
From a chemical point of view, the primary structure of both DNA and RNA is a linear polymer composed of repeating monomeric nucleotide units (Figure 1.2). A nucleotide consists of three molecular sub-units, namely a sugar, a phosphate linkage and a heterocyclic base. The term "nucleoside" refers to a structure comprised of one of four heterocyclic bases attached through a β-glycosidic linkage to a cyclic furanoside ribose or deoxyribose sugar. A nucleotide, therefore, is a nucleoside phosphorylated at the 3' and/or 5'-hydroxyl (-OH) group of sugar. The heterocyclic bases are divided into purines including adenine (A) and guanine (G), and pyrimidines including cytosine (C) and thymine (T) in DNA, while RNA substitutes uracil (U) for T. The whole nucleic acid strand is connected through a phosphodiester linkage in a defined direction, that is, 3'->5' connectivity (Figure 1.2).

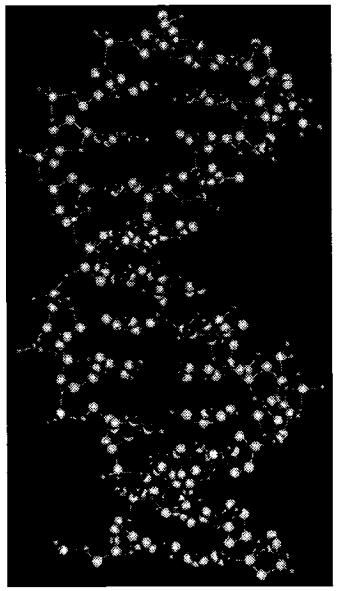
The biological importance of nucleotides in DNA is not in their local structure but

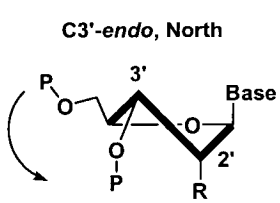
rather in their length and sequence, in other words as the genetic codes. Even though there are only four common heterocyclic bases, the number and sequence in which their corresponding nucleotides may be organized in DNA can vary in extraordinary ways. Another recent breakthrough in biological sciences has been the actually deciphering of these codes in the human genome. Since the veil was lifted[5] for the first time on the complete picture of human genome, we have had to humble ourselves in imagining how such a simple system is able to control the activities necessary to life in such a sophisticated way. There is still a long way to go in actually understanding exactly how these codes are related to certain biological activities, therefore, furthering the understanding of nucleic acids and their applications in biology and medicine certainly will benefit human well-being.

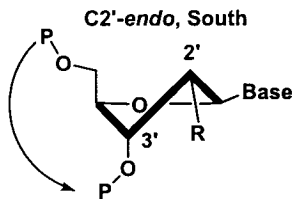
1.2 Secondary Structure of Nucleic Acids

DNA has been shown to adopt different structural polymorphs. The two helical types most commonly found in nature are called the A-form, present in low humidity conditions and high salt concentration, and the B-form, existing in high humidity conditions with low salt concentration (Figure 1.3). These two conformations of helices both display right-handed rotations with two antiparallel strands. The strands recognize each other through Watson-Crick base-pairing, where A pairs with T (or U in RNA) through two hydrogen bonds while G pairs with C through three hydrogen bonds (Figure 1.2). The duplex is further stabilized by hydrophobic interactions and base stacking (π-π-interaction) among the flanking heterocyclic bases at the interior of the helix.[1] Both A- and B-form helices possess characteristic structure parameters as shown in Figure 1.3. Overall, A-DNA is wider, more compact, with more bases per turn, and more tilted from the helical axis as compared with B-DNA. The double helix may also be defined by the two types of grooves found in both A and B forms, namely the minor and major grooves. In A-form, the major groove is narrow and deep and the minor groove broad and









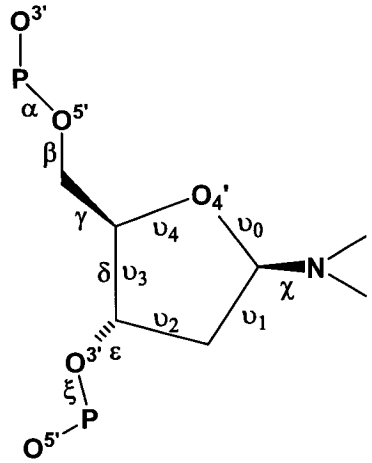
	A-form	B-form
Helicity:	Right-handed	Right-handed
Helix diameter (Å):	25.5	23.7
Residues/turn:	11	10.4
Rise/bp (Å):	2.56	3.3 to 3.4
Displacement/bp (Å):	4.5	-0.2--0.8
Sugar pucker:	North, C3'-*endo*	South, C2'-*endo*
Major groove:	Narrow and deep	Wide and deep
Minor groove:	Broad and shallow	Narrow and deep

Figure 1.3: Global helical conformation and average structure parameters for nucleic acids of A and B form (adapted from Blackburn and Gait, 1996[2])

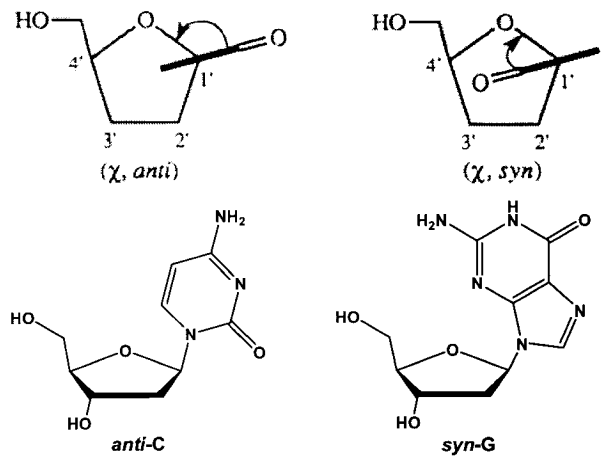
shallow; while in B-form, the major groove is wide and deep and the minor groove narrow and deep (Figure 1.3). The DNA structure originally proposed by Watson and Crick[3] is a B form helix. On the other hand, RNA, which exists predominantly in a single stranded form in cells, prefers an A-form conformation with a complementary RNA strand. The 2'-OH groups in RNA would have steric conflict with the phosphate linkage and the heterocyclic base if RNA were to adopt a B-form helix. Finally, RNA is more susceptible to hydrolysis of the phosphate linkages due to the possibility of neighboring group participation of the 2'-OH, which probably explains why DNA was chosen to store the genetic information.[6]

1.3 Shapes of Nucleosides and Nucleotides [1]

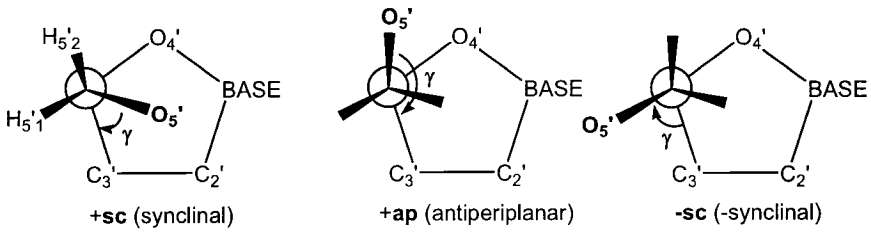
The overall shape of a nucleoside or nucleotide can be measured by bond lengths, bond angles and rotations of groups of atoms. These rotations, expressed as a torsion angles, are important since they define the overall conformational detail of nucleoside/nucleotide units (Figure 1.4 A). According to IUPAC nomenclature,[2] torsion angles are expressed as Greek letters, including α, β, γ, δ, ε, ξ in the phosphate backbone, υ_0 - υ_4 in the furanose ring, and χ for the glycosidic linkage (Figure 1.4 A). Since many of these torsion angles are interdependent, a complete description of the shape of nucleosides involves three principal parameters: (1) the glycosidic torsion angle χ (Figure 1.4 A) that determines the *syn* or *anti* disposition of the base relative to the sugar moiety (*anti* when the H6 of pyrimidines or H8 atom of purines lies above the sugar ring, such as *anti*-cytidine shown in Figure 1.4 B; *syn* when these atoms are oriented in the opposite direction such as *syn*-guanosine in Figure 1.4 B);[7,8] (2) the torsion angle γ, which determines the orientation of C4'-C5' represented by three main rotamers, namely, +*sc*, *ap*, -*sc* (Figure 1.4 C); and (3), more commonly discussed, the sugar pucker and its deviation from planarity (Figure 1.4 D). Sugar pucker is described by the maximum out-of-plane pucker υ_{max} and the phase angle of pseudorotation P (0 - 360°) which are correlated with the equation $\upsilon_{max} = \upsilon_0 / \cos P$).[9,10] The value of P depends on the five endocyclic sugar



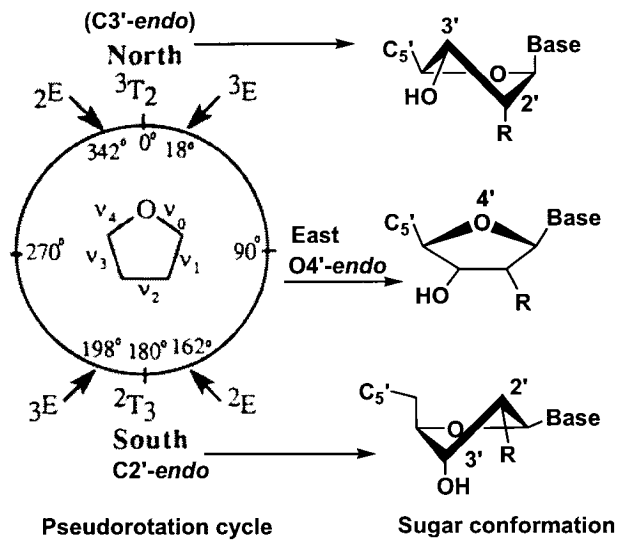
A. Torsion angle notation



B. Glycosidic conformations



C. C4'-C5' bond orientation



D. Sugar conformation

Figure 1.4: Shapes of nucleoside and nucleotide

torsion angles υ_0 - υ_4. The conformation of the furanose ring around the pseudorotational cycle alternates every 18° between envelope (*E*) and twist (*T*) conformations. By convention, the sugar pucker is called north, or C3'-*endo,* at $0^\circ \leq P \leq 36^\circ$ and south, or C2'-*endo*, at $144^\circ \leq P \leq 190^\circ$. An absolute north conformation possessing a symmetrical twist form ${}^3T_2$ (P = 0°), its south antipode, ${}^2T_3$ (P = 180°), and an east conformation ${}^O\text{E}$ (P = 90°) are shown in Figure 1.4 D. The superscripts and subscripts represent, respectively, atoms displaced above or below the plane relative to other atoms in the five-membered ring. Atoms displaced on the same side as C5′ are referred to as *endo* and those on the opposite side are termed *exo*.[11,12] Sugar conformation is influenced by the substitution on the sugar ring, the orientation of the base (*anti* and *syn*) and other factors.

Ideal B-DNA is C2'-*endo* (South) and ideal A-RNA is C3'-*endo* (North) (Figure 1.2). In this thesis, chemical modifications are often discussed in terms of their influence on the shapes of nucleosides/nucleotides, which help to explain some of the interesting properties that these modifications have elicited (discussed in detail in **Section 1.8.1**).

1.4. RNA Splicing and Intron Branched Structure

RNA splicing is a post-transcriptional event that excises non-coding introns and links coding exons.[13,14] The intron excisions result in intermediate, branched ribonucleic acids (bRNAs) structures containing vicinal 2'-5' and 3'-5' phosphodiester linkages at the branch point (Figure 1.5). Branched RNA was first detected in nuclear polyadenylated RNA from HeLa cells by Wallace and Edmons[15] and these species were subsequently shown by others to be mRNA splicing intermediates.[16] Since then, two unique branched RNA configurations have been detected: loops with a tail containing a branch point (i.e., a "lariat" configuration) in monomolecular *cis*-splicing reactions,[16] or branches between two linear RNAs (i.e., a "Y"-like structure) in the case of bimolecular *trans*-splicing reactions.[17] Both in yeast and in mammalian cells, the branch point nucleotide is an adenosine,[18,19] which can have stronger base stacking ability within the branch structure

than other RNA branch points. This could be a possible reason why all lariat RNA introns conserve adenosine.[20]

Branched RNAs are not stable in cells and are hydrolyzed by a specific enzyme that was first identified in HeLa cell extracts and named the RNA lariat debranching enzyme.[21] The enzyme is involved in the intron turnover pathway found in mammalian cells[21] and in yeast[22,23]. The yeast lariat debranching enzyme (yDBR) is a unique manganese-dependent phosphodiesterase, which selectively cleaves the 2'-5'-phosphodiester linkage at the branch site of either lariat or Y-shaped introns, thus converting them into linear molecules (Figure 1.5).[24] yDBR does not digest 3'-5' phosphodiester bonds differentiating it from snake venom phosphodiesterase, which cleaves both 2'-5' and 3'-5' phosphodiester bonds. Debranching efficiency of yDBR is highly substrate-specific. Previous studies have shown that the branch point adenosine is more favored than other natural counterparts[25,26] and chemically modified branch points with different sugar conformations and phosphodiester configurations result in inactivation of debranching activity.[27] A vicinal 3'-5' phosphodiester bond (i.e. one nucleotide connecting to the 3' position) at the branchpoint is required since yDBR cannot not hydrolyze a 2'-5' connection in a linear fashion.[21] The nucleotide directly connected to the branchpoint is also important to the debranching activity and a 2'-purine

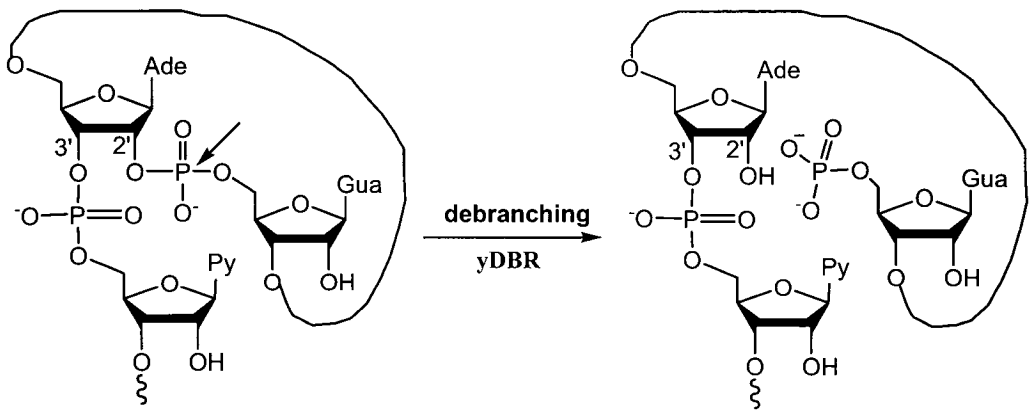


Figure 1.5: A lariat intron RNA undergoes debranching (the arrow shows the cleavage site)

(A) is debranched more efficiently than either of the 2'-pyrimidines (C/T).[26,28] Synthetic Y-shaped RNAs have been shown to be valuable tools for understanding the molecular requirements of RNA splicing[29] and debranching[24,26,27] with the help of the development of branched oligonucleotide synthesis methodologies (see **Section 1.5.1** Y-shaped RNA solid-phase synthesis).[30-32]

1.5 Solid Phase Oligonucleotide Synthesis

Solid phase synthesis was first applied to polypeptide preparation by Merrifield[33] and Letsinger[34] and later adapted to oligonucleotides by Letsinger.[35] The advantages of solid phase synthesis include the fact that reactions can be forced to high yield end points by using a large excess of soluble monomers and coupling reagents; the covalent attachment of the growing chain to a solid support facilitates separation and purification; and the assembly of oligomers can be automated. The phosphite triester method has been the most popular chemical route for DNA synthesis. The method was initially developed by Letsinger and later refined by Beaucage and Caruthers.[36-39] Controlled pore glass (CPG) is widely applied as a solid support due to its uniform pore size and non-swelling feature during synthesis. It can be easily functionalized, first with a long chain alkylamine moiety (LCAA), followed by a succinyl linker[40] and the first 5'-protected nucleoside (Figure 1.6A). Recently a "universal linker" has been developed and been found to be extremely convenient for synthesizing modified oligonucleotides (Figure 1.6B).[41] Other polystyrene-based supports are available and allow for very high loading of nucleoside.

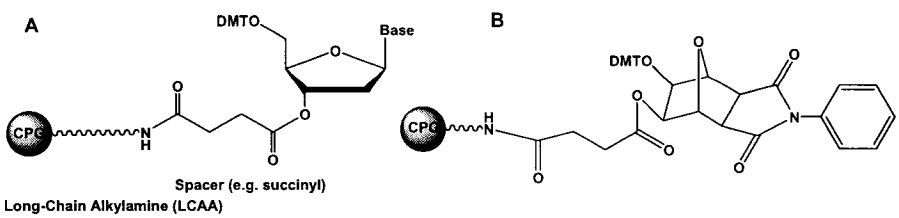


Figure 1.6: (**A**) Solid support for oligonucleotide synthesis and (**B**) a "unilinker" (ChemGene)

The monomers used for the automated solid phase synthesis use orthogonal protecting groups for heterocyclic bases and the 5'-OH (and 2'-OH in RNA) (Figure 1.7). The protecting groups for bases are introduced by acylation of the exocyclic amino groups of adenine and cytosine with benzoyl (Bz) groups and guanine with the *iso*-butyryl (*i*-Bu) group. The 5'-OH is usually protected by a trityl group, either 4'-monomethoxytrityl (MMT) or 4',4'-dimethoxytrityl (DMT),[42,43] which can be easily removed under mild acid condition during each synthesis cycle. Furthermore, the coupling efficiency can be quantitated by measuring the coloured trityl cations via spectrophotometry in the visible region during the detritylation step (Figure 1.8). The most common protecting group for 2'-OH in RNA is the *tert*-butyl dimethylsilyl (TBDMS) groups originally employed in nucleosides by Ogilvie and coworkers.[44,45] TBDMS is stable to the mild acid and base conditions employed in solid-phase cycles and can be removed in the presence of fluoride ions.[46] The phosphoramidite moiety is protected by a β-cyanoethyl protecting group, which is stable in synthesis cycles and only removed during the oligonucleotide deprotection step.

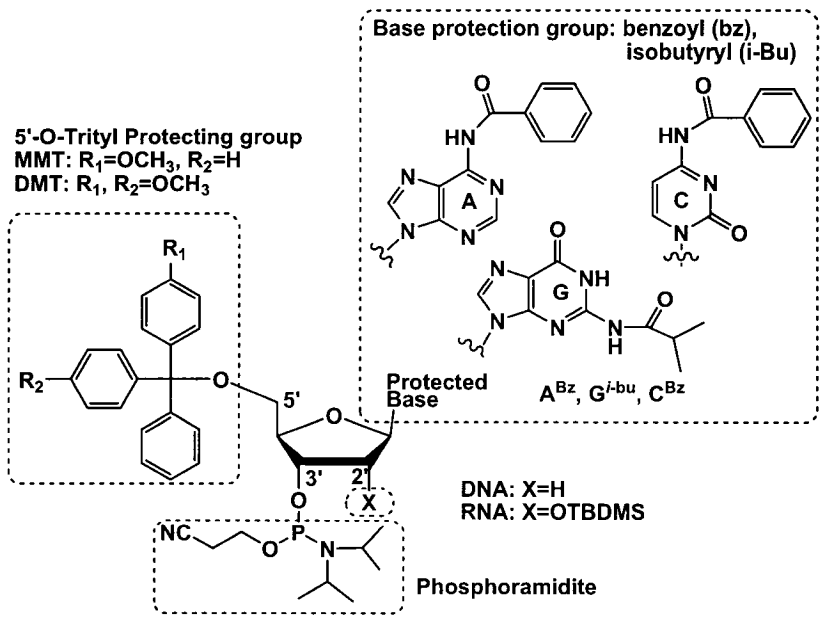


Figure 1.7: Building block (phosphoramidite) in DNA solid-phase synthesis

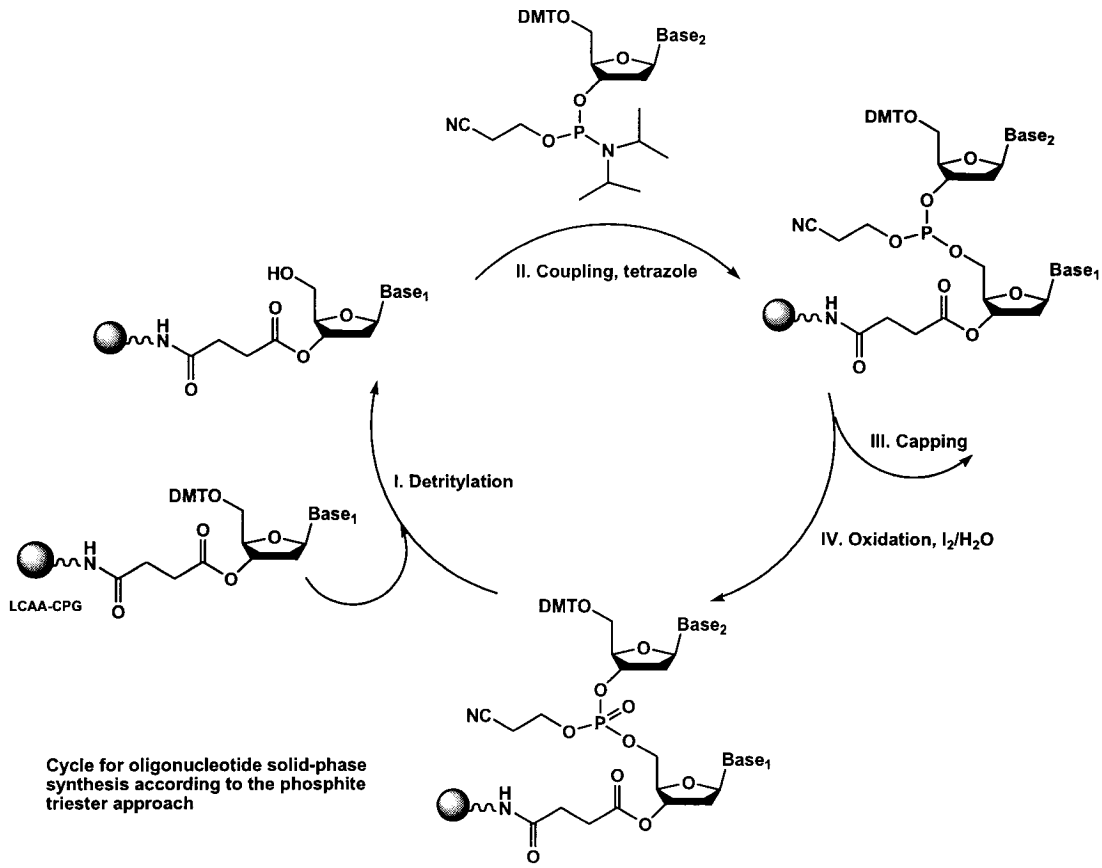


Figure 1.8: Automated DNA solid phase synthesis cycle

The automated solid-phase DNA synthesis strategy via the phosphite triester approach is shown in Figure 1.8. Briefly, the DNA synthesis cycle starts with the cleavage of the 5'-protecting group by acid treatment to afford the free 5'-hydroxyl group, followed by the coupling with next monomeric 3'-phosphoramidite in the presence of an activator such as tetrazole. Since the coupling reaction may not be quantitative, any unreacted 5'-OH is "capped" by acetylation. Oxidation of the trivalent phosphite triester to a more stable pentavalent phosphate triester is achieved using aqueous iodine.[38,39] Repeating the synthesis cycle N-times results in oligonucleotide products with N+1 nucleotides (Figure 1.8). The DNA oligonucleotides are then cleaved under basic condition (aqueous ammonia) and purified by HPLC or polyacryamide gel

electrophoresis (PAGE). An extra step is required for RNA deprotection to remove the silyl protecting group on 2'-OH with a fluoride reagent such as *tert*-butylammonium fluoride (TBAF)[44,45] or triethylamine trihydrofluoride (TREAT-HF).[47]

1.5.1 Y-shaped Branched RNA Synthesis

Solid phase synthesis to make linear oligonucleotides can be easily applied to making Y-shaped branched RNAs (bRNAs) using well-established convergent methodology developed by Damha and coworkers.[31,48] Briefly, the introduction of the branch point nucleoside is achieved by coupling of two adjacent CPG-bound oligonucleotide chains with a 2',3'-bis-phosphoramidite. The coupling time of a bis-phosphoramidite is extended to 30 min and it is introduced at a lower concentration (0.01-0.03 M) than typical amidite solutions during solid phase synthesis.[32] A low bis-phosphoramidite concentration reduces the formation of mostly "extended" linear isomeric failure sequences when the branch point is connected with only one of two support-bound oligonucleotide chains.

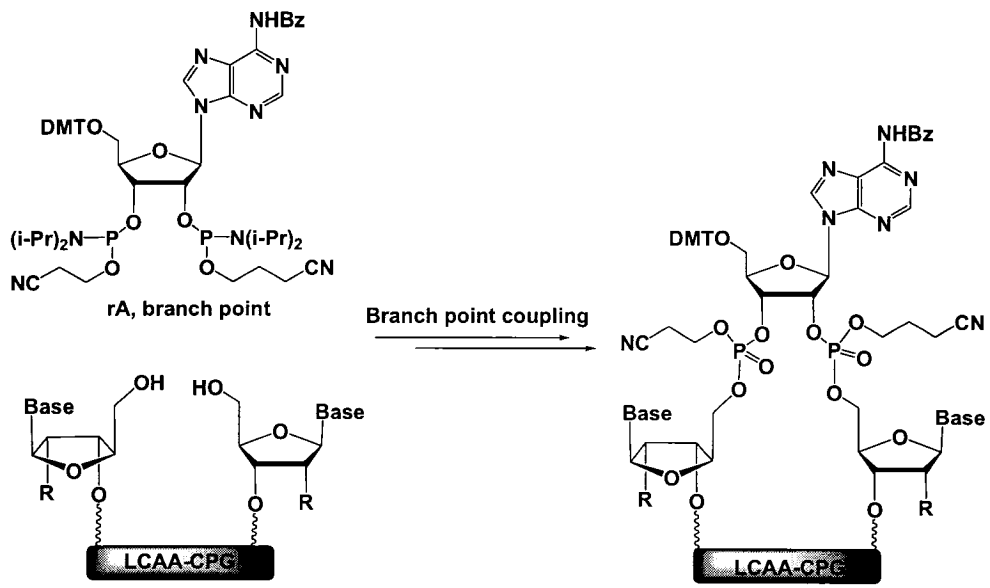


Figure 1.9 Convergent branched oligonucleotide synthesis developed by Damha and coworkers [31] (R=H in DNA; R=OTBDMS in RNA)

High CPG loadings (> 30 μmol/g) are required to ensure two vicinal oligonucleotide chains are close enough to react with the branch point bis-phosphoramidite. The Y-shaped constructs are purified by denaturing PAGE based on a very characteristic gel mobility (slower) as compared with linear failure sequences.[32]

1.6 Techniques Applied for the Study of Nucleic Acids

Nucleic acids are often studied by various techniques to understand their structure, dynamics, binding characteristics and biological properties. A few common techniques employed in this work are described below.

1.6.1 Ultraviolet (UV) Spectroscopy[49]

UV spectroscopy is the most commonly used technique to study oligonucleotides since they exhibit strong UV absorption around 260 nm as a result of absorption by the chromophoric purine and pyrimidine bases. UV absorption can be employed to quantitate the amount of oligonucleotides by measure the optical density units (O.D.U.). The measurement is very sensitive since the molar extinction coefficients of oligonucleotides are at the magnitude of 10^4 M^{-1} cm^{-1} per nucleotide. UV absorption can also be employed to measure the relative affinity of an oligonucleotide to a complementary oligonucleotide target. UV absorption profile obtained with the increase of temperature is called thermal melting curve from which the melting temperature (T_m) of a nucleic acid duplex is defined by the midpoint of thermal melting curve, indicating 50% of nucleic acid is in a duplex state and 50% in a single stranded state (dissociated). The fact that UV absorption decrease with the decrease of temperature is known as hypochromicity which arises from the dipole-dipole interactions of the bases when they stack.[50] The T_m experiment can offer insight to the kinetics and thermodynamic parameters of that system. For example, hysteresis observed in the heating and cooling process often suggests a slow kinetic process of duplex formation. Unimolecular complex (e.g. hairpin structure) formation

does not depend on concentrations, while intermolecular complex (e.g. intermolecular duplex, guanine-quadruplex) should be concentration-dependent in T_m measurement.[49,51,52]

1.6.2 Circular Dichroism (CD) Spectroscopy[53,54]

CD spectroscopy has been used as a highly sensitive but qualitative technique by nucleic acid chemists to study the conformation of nucleic acids in solution. CD spectra record the differential absorption of left- and right-handed circularly polarized light as a function of wavelength as a result of the chirality of oligonucleotides, providing a CD "pattern" that is related to the overall structure of oligonucleotides. For example, two canonical A and B-form duplexes show distinct CD characteristics: A-form shows positive peak at ~ 265 nm, negative peaks at ~ 230 and 210 nm and a cross line at 240 nm; while B-form shows a positive peak at 280 nm, a negative peak at ~ 240 nm and a cross line at 260 nm. CD spectra are a sensitive probe to the conformation of other unusual nucleic acid structures. For example, CD spectra have been used to distinguish between parallel and antiparallel guanine-quadruplexes (See more information in **Section 1.7.1**).

1.6.3 Electrophoresis[55,56]

Electrophoresis serves as an important purification and characterization technique for nucleic acid structures. Oligonucleotides, due to their polyanionic nature, can move through a porous gel under the influence of an electric field based on their charge, size and molecular shape. Usually small oligonucleotides move faster relative to larger molecules, and branched molecules move slower than linear molecules. The cross-linked polymer used in gel electrophoresis in this work is polyacrylamide and the density of cross-linkers depends on the amount of polyacrylamide used in the polymerization. These gels can be run under denaturing (with urea added and elevated temperature) or native conditions (no urea and low temperature). Denaturing gels are typically employed for

purification purposes, whereas native gels are used to assess the formation of a nucleic acids complex such as nucleic acid and protein complex structure. The oligonucleotides can be visualized by UV shadowing or staining the gel bands with a dye such as StainsAll®(Bio-Rad).

1.7 Other Secondary Structures of Nucleic Acids

Since the DNA "alphabet" is made up of only four basic building blocks, far less than that of the proteome, DNA must have the ability to adopt multiple structural polymorphs and is required to be intrinsically more flexible in order to meet its biological roles in cells.[57,58] As mentioned in **Section 1.2**, aside from the two canonical A- and B-forms, DNA can adopt many other secondary structures. For example, Z-DNA[59,60] is a left-handed helix, contrasted to the right-handed rotation of the A- and B-form. In Z-DNA, the orientation of the glycosidic bonds among the bases alternate between the *syn* and *anti* conformations (Figure 1.4). Z-DNA is preferentially formed by a sequence of alternating purines and pyrimidines. Compared with *anti*-parallel strands found in A- and B-form, double helices can also be organized in a parallel orientation.[61-64] In addition to the typical double stranded structure, DNA can also form triplexes (or triads)[52,65], quadruplexes (or tetrads) with an *i*-motif[66] or a guanine (G)-quartet structure[67-69] (Figure 1.10), even higher ordered structures such as pentads, hexads and heptads (see review in this topic[70,71]). These structures involve non-canonical hydrogen-bonding among base pairs,[72] stronger base stacking, base-backbone interactions, backbone-backbone interactions, interactions of the complex with solvent and the complex with metal ions. A good example, shown in Figure 1.10, is that of a DNA sequence with human telomeric repeats (a telomere is a specific nucleoprotein structure at the end of eukaryotic chromosomes, which is essential for chromosome stability[73,74]). The pair of oligonucleotides consisting of a telomeric DNA with the sequence 5'-d(AGGG(TTAGGG)$_3$)-3' and its complementary strand is able to form a

typical duplex by Watson-Crick base pairing (A:T, C:G) together; a unimolecular *i*-motif quadruplex by C:C$^+$ base pairing (C$^+$ refers to a protonated cytosine) involving the complementary strand alone, and a G-quadruplex by G:G base pairing involving the telomeric DNA strand alone.[75] The G-quadruplex structure will be focused on in the next section.

1.7.1 G-quadruplex Structure

G-quadruplex, by its name, is a four stranded nucleic acid structure containing a unique planar structural motif: G-quartet with four Hoogsteen G:G base pairs (Figure 1.11). The structure of G-quartets was first observed more than 40 years ago by Gellert et al.[67] when they found that guanosine 3'-monophosphate (3'-GMP) and 5'-monophosphate (5'-GMP) form layers of hydrogen-bonded tetramers.

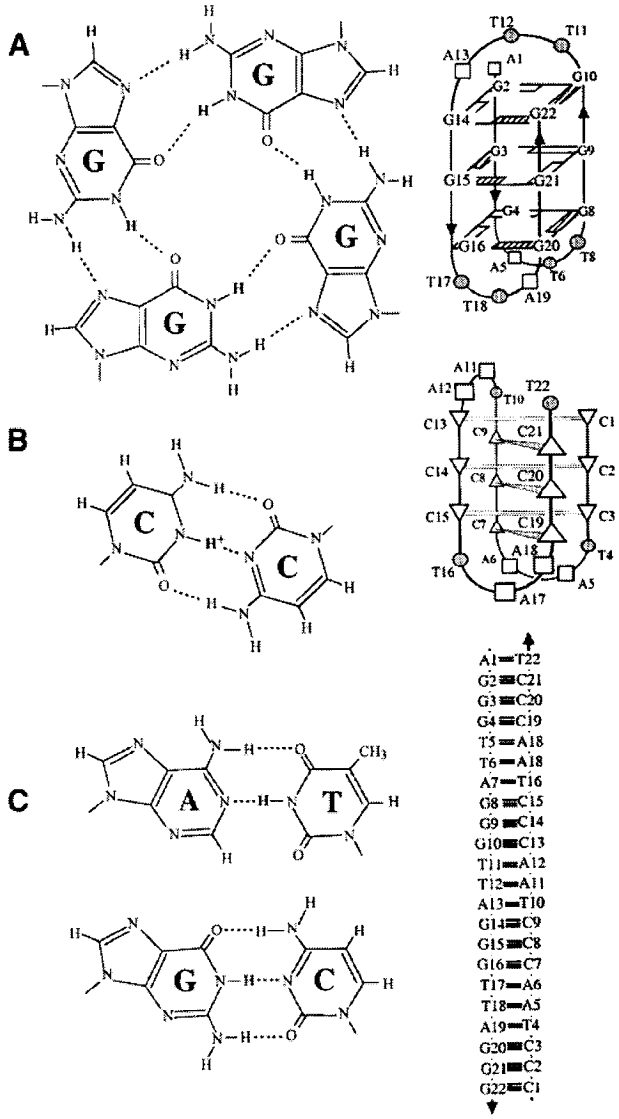


Figure 1.10: Human telomeric DNA sequence 5'-d(AGGG(TTAGGG)$_3$-3' and its complementary strand can form different secondary structure: (**A**) G-quartet (left) and schematic structure of the intramolecular Na$^+$ G-quadruplex (right). (**B**) C:C+ base pair (left) and schematic structure of the intramolecular *i*-motif (right). (**C**) Watson-Crick A:T and G:C base pairs (left) and schematic structure of the Watson:Crick duplex (right). (adapted from Phan and Mergny 2004 [75])

They elucidated a planar structure where four guanines are arranged in a way that the N1-H and N2-H in one guanine forms hydrogen-bonds with the N7 and O6, respectively, of another guanine (Figure 1.11). This structure is stabilized through the coordination of alkali-metal cations, particularly K^+ and Na^+, to the carbonyl oxygens (C=O) of the bases, otherwise the electrostatics of this cyclic arrangement would be unfavorable.[76]

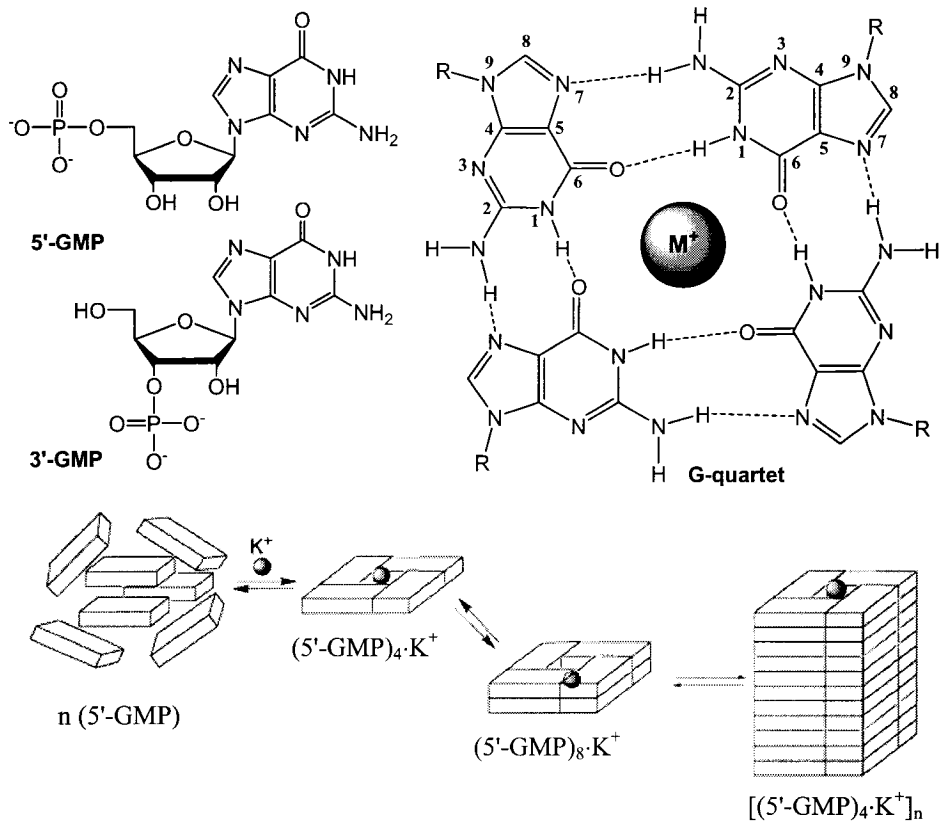


Figure 1.11. Schematic presentation of the formation of G-quartet and G-quadruplex. Poorly defined aggregates are formed in the absence of cations. Cations can stabilize the formation of the planar G-quartets and helical G-quadruplex stacks (adapted from Davis, 2004 [77])

Sen and Gilert[78] first demonstrated guanine-rich DNA oligonucleotides formed parallel 4-stranded complexes (i.e. G-quadruplex) with G-quartet structures. Subsequently, G-quartets have been shown to be the basic structural motif of G-quadruplexes formed by a diverse number of guanosine-rich oligonucleotides *in vitro*.

These G-rich sequences are originated from biologically significant regions of the genome, such as in telomeres (see below),[79,80] immunoglobulin switch regions,[78] gene promoter regions,[81,82] in sequences associated with human disease[83] and in sequences identified as potential therapeutics through screening techniques such as SELEX (see **Section 1.8.3.1**). Although these sequences have been demonstrated to form G-quadruplex structures *in vitro,* indirect evidence suggests that G-quadruplexes exist *in vivo* as well and are involved in important biological functions such as the stabilization of chromosome structure by telomeres.[73,74] The extreme 3'-end of telomeres consists of single-stranded G-rich DNA overhangs.[84,85] G-quadruplexes are thought to form in these regions and interact with telomerase,[86] an enzyme that adds a specific DNA repeat sequence, such as "d(TTAGGG)" for all vertebrates, to the ends of telomeres. Several small molecules have been shown to interact with G-quadruplexes and have been found to be credible selective inhibitors of telomerase,[87,88] suggesting G-quadruplexes are involved in telomerase binding and function. Further evidence for the existence of G-quadruplexes *in vivo* is found in the fact that the formation of G-quadruplexes can be promoted by certain proteins.[89-93] Lastly, the existence of G-quadruplexes *in vivo* is strongly supported by the discovery of helicases capable of unwinding G-quadruplex DNA.[94,95]

Nucleic acid sequences containing G-quadruplex structures have drawn great attention not only for their biological functions in cells but also for their potential as therapeutic agents and as targets for anticancer therapeutics.[96] Nucleic acid quadruplexes have been demonstrated to inhibit human thrombin activity,[97,98] HIV infection,[98] HIV-1 integrase activity,[99,100] and as candidates to inhibit telomerase activity in cancer cells.[101] Excellent reviews can be found to cover different aspects of G-quadruplex activity in biology,[74,93,96] its biophysical properties,[102,103] structure[68,69] and uses in materials science.[77]

1.7.2 Polymorphism of G-quadruplexes

G-quadruplexes display an extraordinary array of structural polymorphisms,[68,92,104] elucidated by a number of modern spectroscopy techniques including NMR spectroscopy, X-ray crystallography, thermal denaturation (see **Section 1.6.1**) and circular dichroism (see **Section 1.6.2**). Various G-quadruplex structures may be discussed through the following variations.

1.7.2.1 Strand Stoichiometry

A single sequence can form different types of structures by associating in various stoichiometries, for example four strands associating may form tetrameric G-quadruplexes (Figure 1.12A), while two strands may form dimeric G-quadruplexes (Figure 1.12B,C,D,E,F), or a single strand may self-associate to form a unimolecuar G-quadruplex (Figure 1.12 G,H,I). Though complexes composed of three stranded structures can be imagined, none has been characterized to date.[104]

1.7.2.2 Strand Alignments: Parallel and Antiparallel Configurations

G-quadruplex strands can be assembled in both parallel or antiparallel alignments. As shown in Figure 1.12, d(TG_4T) (Figure 1.12A) forms linear intermolecular parallel structures; while dimeric edge-loop hairpins (Figure 1.12B, E, F) and unimolecular structures (Figure 1.12G) form in antiparallel orientations. Diagonal-loop hairpins (Figure 1.12C, D) and unimolecular G-quadruplexes (Figure 1.12H, I) may contain a mix of parallel and antiparallel strands.[68] Previous studies have suggested that the parallel stranded structure is much more stable,[105] therefore short linear strands usually form parallel G-quadruplexes.[106] Loop connectivity, however, will change the strand alignments. edge-loops and diagonal-loop hairpins with the same sequences have different strand alignments, as shown in Figure 1.12. A recently reported unimolecular G-quadruplex d($AAAG_3(T_2AG_3)_3AA$)[107] has an unusual loop structure leading to the

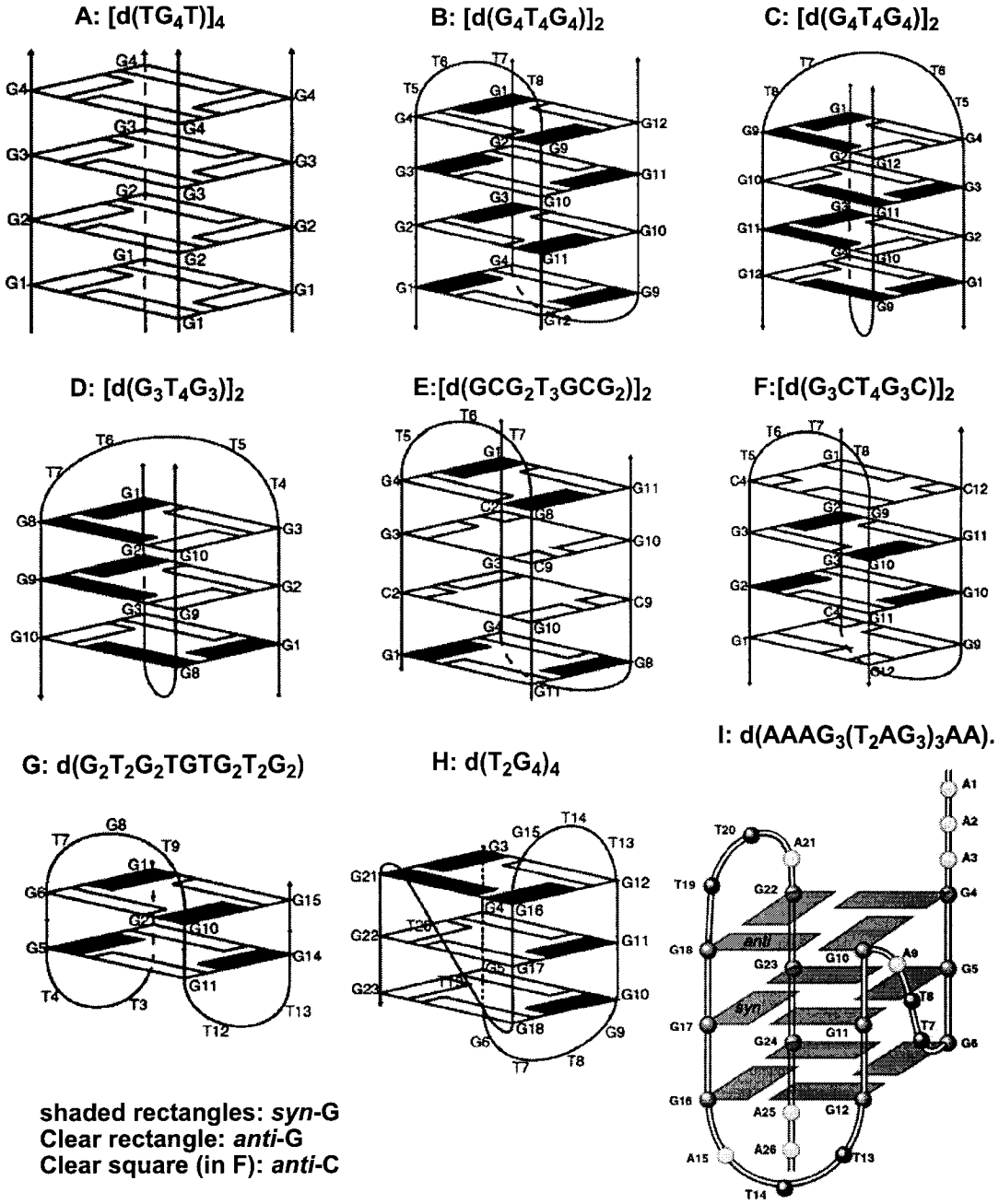


Figure 1.12 Schematic diagrams of a four-stranded linear quadruplex (A), some two-stranded dimeric hairpin quadruplexes and unimolecular quardruplexes, which are based on x-ray crystal structures of (**A**) $[d(TG_4T)]_4$[108] and (**B**) $[d(G_4T_4G_4)]_2$[109] and the NMR solution structures of (**C**) $[d(G_4T_4G_4)]_2$;[110,111] (**D**) $[d(G_3T_4G_3)]_2$;[112-114] (**E**) $[d(GCG_2T_3GCG_2)]_2$;[115] (**F**) $[d(G_3CT_4G_3C)]_2$;[116,117] (**G**) $d(G_2T_2G_2TGTG_2T_2G_2)$;[118,119] (**H**). $d(TG_4)_4$;[120] (**I**) $d(AAAG_3(T_2AG_3)_3AA)$.[107] A solid line represents the sugar phosphate backbone. A solid circle indicates the 5' end of the strand and an arrow indicates the 3' end of the strand (adapted from Keniry, 2000 [104] except for (**I**))

co-existence of parallel and antiparallel configurations, which again shows the amazing polymorphic versatility of G-quadruplexes.

1.7.2.3 Glycosidic Conformations[68]

The studies of Mohanty and Bansal[121] have suggested that G-quadruplexes prefer specific sugar conformations. The guanosines that compose a G-quartet can adopt either entirely *syn* glycosidic conformations (*syn*-Gs), entirely *anti* conformations (*anti*-Gs), or a mix of both (the definition of *syn* and *anti* glycosidic conformation is shown in Figure 1.4). It has been established that the strand alignments in G-quadruplexes have a strong correlation to guanosine glycosidic conformations.[68] For example, molecular modeling has indicated that a parallel-stranded structure with exclusively *anti*-Gs is more stable than a structure with alternating *syn* and *anti*-Gs and all syn-Gs; while antiparallel fold-back structures (e.g. hairpin dimeric G-quadruplex in Figure 1.12A-F) prefer the alternating *syn–anti* conformations. The theoretical calculations support numerous G-quadruplex structures observed in experiments.[103,104]

1.7.2.3 Ion-binding Geometry[103]

Various cations (generally monovalent) interact with the G-quartet structures with differing binding affinities. The general trend of monovalent cation-induced stabilization is found to be $K^+ > Rb^+ > Na^+ > Cs^+ > Li^+$. The fact that K^+ is better than Na^+ in stabilization of quartets is likely due to the difference in dehydration energy ($K^+ < Na^+$) between the two rather than a better size match at the binding site of K^+ over Na^+.[103] A change in the type of cation bound to the G-quartet can have an effect on the cation to ligand stoichiometry. Furthermore, potassium-induced G-quartet structures can differ from their sodium-induced counterparts in terms of both stability and conformation. For example, a telomeric DNA sequence $d(G_4T_4G_4)$ (Figure 1.12 B & C) can form two different kinds of dimeric quadruplex, with a diagonal loop observed when Na^+ is the

cation used (Figure 1.12 B)[110,122] and an edge-loop when K^+ is used (Figure 1.12 C).[109]

1.7.3 Three Examples of G-quadruplexes

Three G-quadruplexes of relevance to this work are a thrombin-binding aptamer (Figure 1.12G), a telomeric DNA sequence (Figure 1.12C), and an HIV infection inhibitor containing phosphorothioates linkages (Figure 1.12A).

The thrombin-binding aptamer, generated via *in vitro* selection, is a 15 nucleotide long oligomer $d(G_2T_2G_2TGTG_2T_2G_2)$[97] that has been intensively studied for its ability to bind and inactivate thrombin, a key enzyme in the blood clotting cascade. NMR spectroscopy has revealed that it adopts a unique folded structure with two stacked G-quartets connected through edge-loops (one comprised of TGT and two of TT) with an antiparallel alignment of adjacent strands (Figure 1.12G).[118,119,123,124] A seemingly contradictory crystal structure of this aptamer-thrombin complex [125,126] was later shown to be in agreement with the previous NMR structure.[126,127] Potassium ions (K^+) stabilize the structure of G-quadruplex by coordination to the G residues as previously discussed.[123] The G-quartet displays an alternating *syn* and *anti* conformation of guanosines (*syn*- and *anti*-Gs) within the same plane and 5'-*syn*-G/3'-*anti*-G confirmation along each strand of the quadruplex (i.e. 5'-*syn*-*anti*-3' connectivity where *syn*-Gs include G1, G5, G10 & G14 and *anti*-Gs include G2, G6, G11 & G15 in example G in Figure 1.12) while all the thymidines are in the *anti* conformation.[123]

NMR and X-ray crystallographic analysis of a telomeric DNA sequence $d(G_4T_4G_4)$ has shown that it can form two kinds of symmetrical dimeric quadruplexes with four G-quartets via a diagonal loop in the presence of Na^+ (Figure 1.12C)[110,122] and an edge-loop formed in the presence of K^+ (Figure 1.12B).[109] The dimeric diagonal hairpin complex (Figure 1.12C) consists of *5'-syn*-G/3'-*anti*-G along each G-strand, *syn-syn-anti-anti* conformations in a G-quartet and all deoxynucleotides in the *south* conformation.[122,128]

The final example is a therapeutically interesting G-quadruplex consisting of a phosphorothioate linked PS-d($T_2G_4T_2$), which was first combinatorially selected for its inhibition ability of HIV envelope-mediated cell fusion during *in vitro* experiments (HIV infection).[98,129] PS-d($T_2G_4T_2$) forms a parallel-stranded tetramer stabilized by four G-quartets with all *anti*-Gs, the same structure as that of PO-d(TG_4T) shown in Figure 1.12A.[78,98]

1.8 Overview of Oligonucleotide-based Therapeutics and Diagnostics

Chemically modified oligonucleotides have drawn great interest in the scientific community and in public eye for their important roles in gene silencing and regulations, and their tremendous potential for use in the development of innovative therapeutics and diagnostics. The different mechanisms of action and the various biological pathways that synthetic oligonucleotides are involved in as well as their applications will be briefly reviewed in this section.

1.8.1. Antisense Oligodeoxynucleotides[130,131]

Antisense molecules generally consist of a single strand of DNA, referred to as an antisense oligodeoxynucleotide (AS-ODN). The first demonstration of an AS-ODS to act as a gene silencing agent was reported by Zamecnick and Stephenson.[132] Theoretically, chemically modified AS-ODNs can be introduced into cells and bind with very high specificity to targeted intracellular messenger RNA strands (mRNA) through Watson-Crick base pairing (Figure 1.13). The binding of the AS-ODN to the mRNA will arrest the translation machinery of the cell by physically blocking its progression along the mRNA and then its presence may activate an RNA nuclease, RNase H,[133] an enzyme that recognizes DNA/RNA duplexes and selectively cleaves the RNA strand, releasing the antisense strand (AS-ODN) untouched.[131,134] Though simple physical arrest of the translation machinery is effective in inhibiting gene expression, a therapy based on this

technique would require higher doses than one where RNase-H mediated mRNA cleavage occurs. Eliciting RNase-H activity is thus desirable as it is much more efficient since it allows the AS-ODN to act catalytically, with a single antisense strand repeatedly involved in the mRNA cleavage processes, allowing for a much smaller dose of therapeutic to be used. The length of AS-ODNs is typically 15-20 nucleotides in length in order to ensure optimal identification and binding of a unique targeted sequence, avoiding unintended silencing of other genes.[135,136] The ability of an AS-ODN to form a stable hybrid with mRNA depends on its binding affinity and sequence-specificity, evaluated by determining the melting temperature (T_m), that is the temperature at which one half of a population of duplexes dissociate into single strands. The recognition of AS-ODN/RNA duplexes by RNase-H is, unfortunately, restricted to a small number of chemically modified, conformationally DNA-like nucleotides.[137]

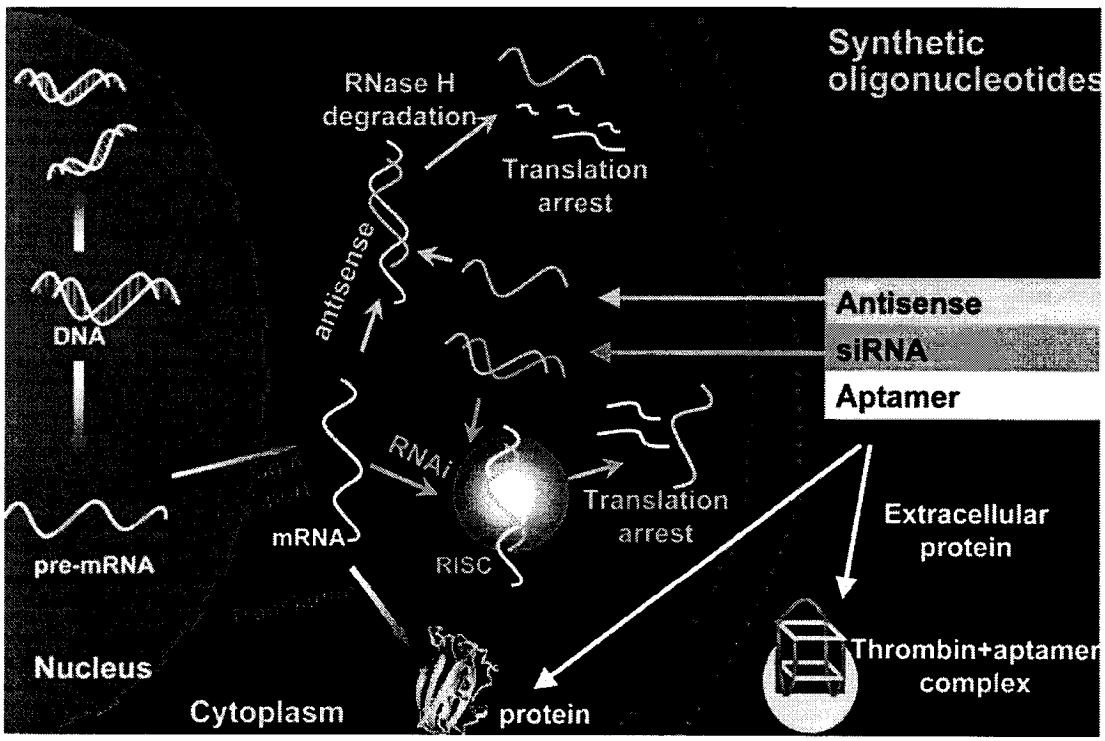


Figure 1.13: Different biological functions of synthetic oligonucleotides

1.8.2 Small Interfering RNA (siRNA)[138]

Andrew Fire and Craig Mello were awarded the 2006 Nobel Prize for Physiology or Medicine for their discovery of RNA interference (RNAi) in 1998. RNAi has been proven to be a powerful gene silencing mechanism triggered by long double-stranded RNA (dsRNA, 299nt long).[139] RNAi most likely evolved as a natural defense mechanism against viral or bacterial infection. It was first observed in worms[139] and plants[140] and was then found in mammalian cells.[141] Silencing is achieved by the action of small interfering RNAs (siRNAs), 21 to 25 nucleotides long,[141-143] which is either generated in cells from longer dsRNA, through ATP-dependent cleavage by an endogenous ribonuclease, a dsRNA specific RNase III enzyme called Dicer,[144] or introduced into cells as exogenous synthetic siRNAs.[141] These siRNAs are 5'-monophosphorylated and have free hydroxyl groups at the 3'-termini. In addition, there are overhangs at the 3'-ends of each strand, consisting of two nucleotides each. siRNAs containing these overhangs are found to be much more effective in gene silencing than duplexes with blunt ends.[143] Regardless of the source, the siRNAs are then taken up by an inactive ribonucleoprotein complex, in which ATP-dependent unwinding of the siRNAs by a helicase allows the antisense strand (guide strand)[145] to bind to the targeted mRNA and yield an active RNA-induced silencing complex (RISC). The RNA target is then selectively cleaved by a sequence-dependent endonuclease present in RISC (Figure 1.13).[146] This endonuclease contains the Argonaute proteins, whose crystal structure indicates that the Argonaute PIWI domain, responsible for the "Slicer" activity, is very similar to known RNase H structures.[147] RNAi has an antisense type of mode of action since it ultimately employs a single stranded RNA to bind to the target mRNA through Watson-Crick base pairing, recruiting a ribonuclease to degrade the target (Fig. 1.13). Synthetic siRNAs have been shown to be more effective in gene silencing than AS-ODNs [141,148-150] and have great potential for use in the development of siRNA based therapeutics.[151,152]

1.8.3. Aptamers[153-157]

Oligonucleotide aptamers, usually derived from a *Systematic Evolution of Ligands by EXponential enrichment* (SELEX) process (see **Section 1.8.3.1**),[158-160] have drawn great attention for their potential as therapeutics and diagnostics. The word "aptamer" is derived from the Latin word "*aptus*" meaning "to fit" and the Greek "*meros*" meaning "part or region").[159] Aptamers can form unique three dimensional structures that can selectively bind to almost any kind of molecular target including proteins,[161] nucleic acids[162] and small molecules.[163] Double-stranded DNA that can bind to transcription factors (i.e. proteins) and inhibit the transcription process, so as to regulate gene expression, is often termed a "decoy" oligodeoxynucleotide (ODN).[164,165] Aptamers are regarded as a type of oligonucleotide "antibody" due to their specificity and high binding affinity, usually measured at the nanomolar scale.[153] Furthermore, aptamers offer advantages over antibodies, such as ease of screening, uncomplicated and cost-effective large-scale manufacturing, and free of cell-culture-derived contaminants, by contrast, antibodies require complicated manufacturing processes using cell-based (eukaryotic or prokaryotic) expression systems. Binding affinity may be further enhanced to the picomolar scale by using SELEX with photo-cross-linkable aptamers, which allow the covalent attachment of aptamers to their bound proteins, while maintaining high selectivity. These aptamers have been shown great potential for the diagnostic applications.[155] Functional groups such as fluorescence dyes or chemically reactive groups can be readily attached to the nucleotides during aptamer synthesis. These functional groups allow aptamers to serve as probes of enzymes, beads or radiotracers in diagnostic assays.[153,166,167] Shown in Figure 1.13 is an example of a DNA aptamer that can bind to the extracellular protein thrombin, a key enzyme involved in the blood clotting process.[97] Although they are only recent additions to the large number of nucleic acid-derived molecules being investigated as potential therapeutics, aptamers have gained recognition with the recent FDA approval of Macugen®, an RNA aptamer for the

treatment of neovascular age-related macular degeneration (AMD) (Figure 1.15).[168] Aptamers may also prove useful for the treatment of other important human maladies, such as infectious diseases, cancer, and cardiovascular disease.[155]

1.8.3.1 SELEX: A Powerful *in vitro* Selection Tool

SELEX[158-160] is often introduced by the preparation of a very large "library" of random sequences, usually containing 10^{14} unique DNA or RNA oligonucleotides that fold into unique three dimensional structures depending on their particular base sequence. The library is then incubated with a target molecule, such as a protein, and a "selection" step is performed. Those oligonucleotides in the library that bind the target are separated and retained while those that do not are discarded. In DNA based SELEX, the retained DNAs are then amplified by a polymerization chain reaction (PCR), directly followed by a second round of selection (Figure 1.14). RNA based SELEX requires that the selected RNAs be converted to complementary DNA (cDNA), prior to PCR and then the amplified cDNAs be transcribed by T7 RNA polymerase to yield the enriched RNA library, finally followed by a second selection step. Typically, these selection and amplification processes are repeated from 8 to 15 rounds until the oligonucleotide aptamers with the highest affinity for the target protein are isolated. The "better fit" aptamers are then cloned and sequenced. Once identified, the aptamer is chemically synthesized

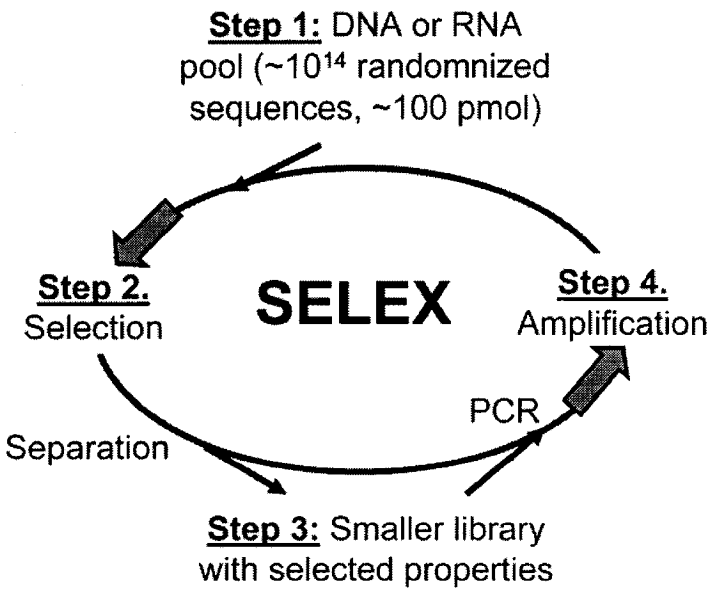


Figure 1.14: SELEX processes

on a larger scale, usually via solid-phase synthesis, and its protein binding properties are confirmed through *in vitro* assays.[169]

1.9. Chemical Modifications in Oligonucleotide Therapeutics and Diagnostics

Although the mechanisms involved in the applications of synthetic oligonucleotides (i.e., antisense, siRNA and aptamers) are different, applications of oligonucleotide-based therapeutics and diagnostics face similar key hurdles *e.g.,* delivery, cellular uptake and biostability.[170] The primary concern is that natural, unmodified DNA and RNA strands are subject to rapid degradation, primarily through the action of 3' exonucleases, with a lesser amount endonuclease activity as well. For example, an unmodified antisense oligonucleotide exhibited a half-life in serum of less than one minute,[171] and an unmodified thrombin-binding aptamer showed an *in vivo* half-life of 108 seconds only.[172] Other concerns such as selectivity (i.e. high binding affinity to the target and low non-specific binding toward non-targets) must also be considered. Ideally, chemically modified nucleosides should allow their oligonucleotides to activate the mechanisms that lead to the selective destruction of mRNA in antisense and siRNA applications, or to maintain the desired three dimensional structures that allow for binding in aptamers applications. Two successful examples of oligonucleotide based therapeutics are shown in Figure 1.15. Vitravene™ is the first and only antisense drug approved for market. It was developed for the treatment of human cytomegalovirus (CMV) induced retinitis,[173] and was approved by the U.S. Food and Drug Administration (FDA) in 1998. It is a phosphorothioate based DNA (PS-DNA), in which a sulfur atom replaces a non-bridging oxygen in the natural phosphodiester backbone (Figure 1.15). The second example, named Macugen®, is a vascular endothelial growth factor (VEGF) RNA aptamer approved by the FDA in 2004 for the treatment of neovascular age-related macular degeneration (AMD).[174] Macugen consists of modified 2'-fluoro pyrimidines and 2'-O-methyl purines, along with the 5'-terminus conjugated to a polyethylene glycol moiety as an anchor and a 3'-terminus protected by with an inverted

thymidine (Figure 1.15).[174] The chemical modifications improve the *in vivo* half-life, bioavailability and pharmacokinetic parameters of this aptamer.[175,176]

1.9.1 Chemical Modifications for siRNA and Antisense

Chemical modifications can be made to the heterocyclic base, the sugar and the phosphodiester backbone. For antisense applications, the structural requirements of RNase H activation[137] pertaining to chemical modifications,[131,177,178] especially 2'-carbohydrate modifications,[179] have been widely discussed in the literature. The experiences gained from applying chemical modifications to antisense may also benefit their use in siRNA based applications, with careful consideration to the differences in mechanisms of action in cells (Figure 1.13). Detailed discussions about chemically modified siRNA and their functions for RNAi can be found in several reviews.[178,180,181]

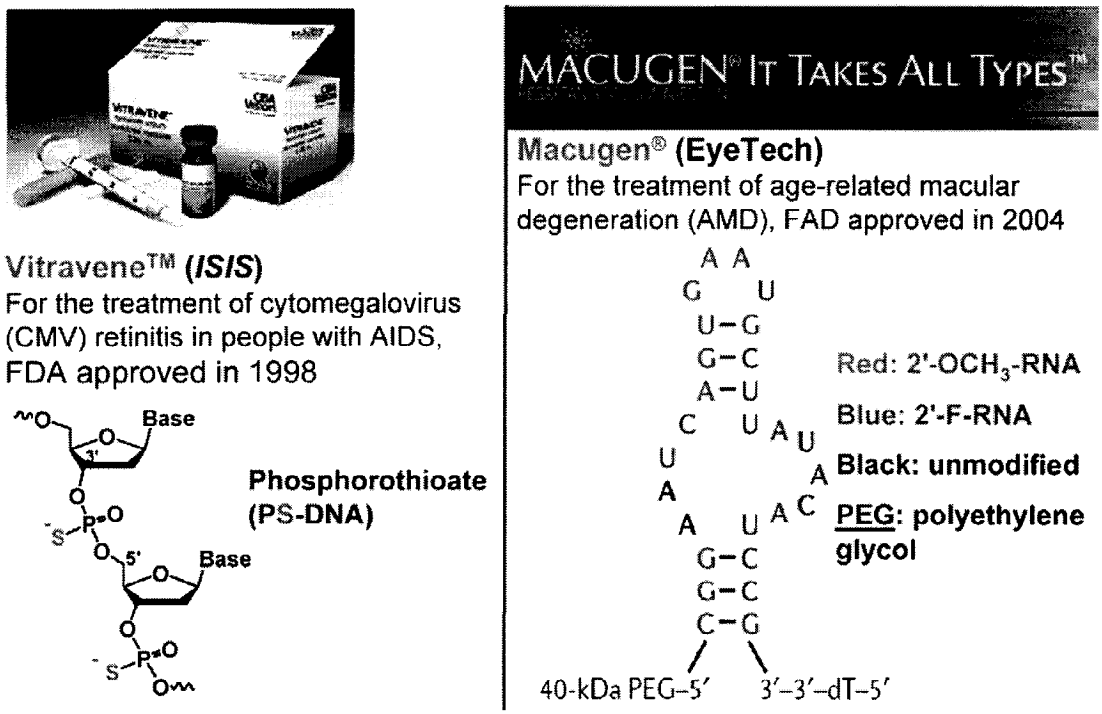


Figure 1.15: Two successful oligonucleotides based therapeutics (right figure is adapted from Ng et al., 2006[174])

For both siRNA and antisense approaches, high thermal denaturation values with mRNA targets are generally required and are often used as the first evaluation criteria for interesting candidates for these applications.[177] 2'-Carbohydrate modifications are extremely important in modulating sugar conformations, impacting mRNA binding.[179]

Generally RNA-like analogues with C3'-*endo* sugar conformation (Figure 1.17) are favored for strong mRNA binding.[182] Substituents at the 2'-*O* position, such as 2'-*O*-methyl (2'-*O*-Me) or 2'-*O*-methoxyethyl (2'-*O*-MOE) (Figure 1.16a & b) or substitution of the hydroxyl group with a fluorine atom at the 2' position, such as 2'-deoxy-2'-fluoro-ribonucleic acids (2'F-RNA) (Figure 1.16d) favor a C3'-*endo* sugar conformation (Figure 1.17). This conformation allows the 2'-electronegative atom to adopt a pesudoaxial position, minimizing energy through a gauche interaction with the O4'-ring oxygen (X-C2'-C1'-O4', X is an electronegative atom).[183] 2'-*O*-Me, 2'-*O*-MOE, and 2'F-RNA raise the T_m value by 1.7, 2, and 3 °C/modification in RNA/RNA duplexes, respectively.[179] An entirely C3'-*endo* conformation obtained for locked nucleic acids (LNA) through a methylene bridge between the 2'-oxygen and the 4'-carbon shows an even higher T_m increase (> 3 °C/modification).[184] Preorganization of A-like (RNA) analogues is believed to reduce the entropy loss during duplex assembly of a complex, thus requiring A-like building blocks so in order to stabilize the complex.[182] Preorganization was first described by Cram in the analysis and design of small organic host-guest complexes.[185] Oligonucleotide duplex formation is an enthalpy-favored process due to the energy lost by the formation of hydrogen bonds and base stacking, while it is an entropy unfavoured process since an ordered structure (duplex) forms from a less ordered structure (single strands). If a monomeric unit preorganizes itself into a conformation required by a higher ordered structure, the entropy loss during duplex formation is reduced since the monomer is already in the required conformation and doesn't lose as many orders of freedom. 2'-Alkyl modifications[177] such as 2'-α-*C*-hydroxymethyl (2'-α-*hm*) (Figure 1.16f) are expected to adopt a C2'-*endo*

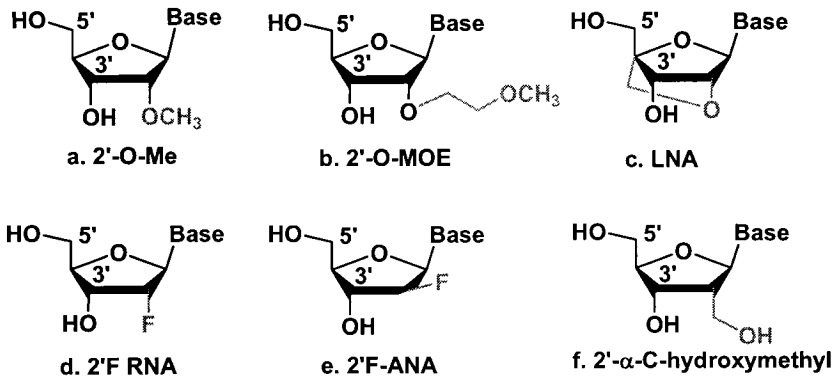


Figure 1.16: Some examples of 2'-carbohydrate modifications

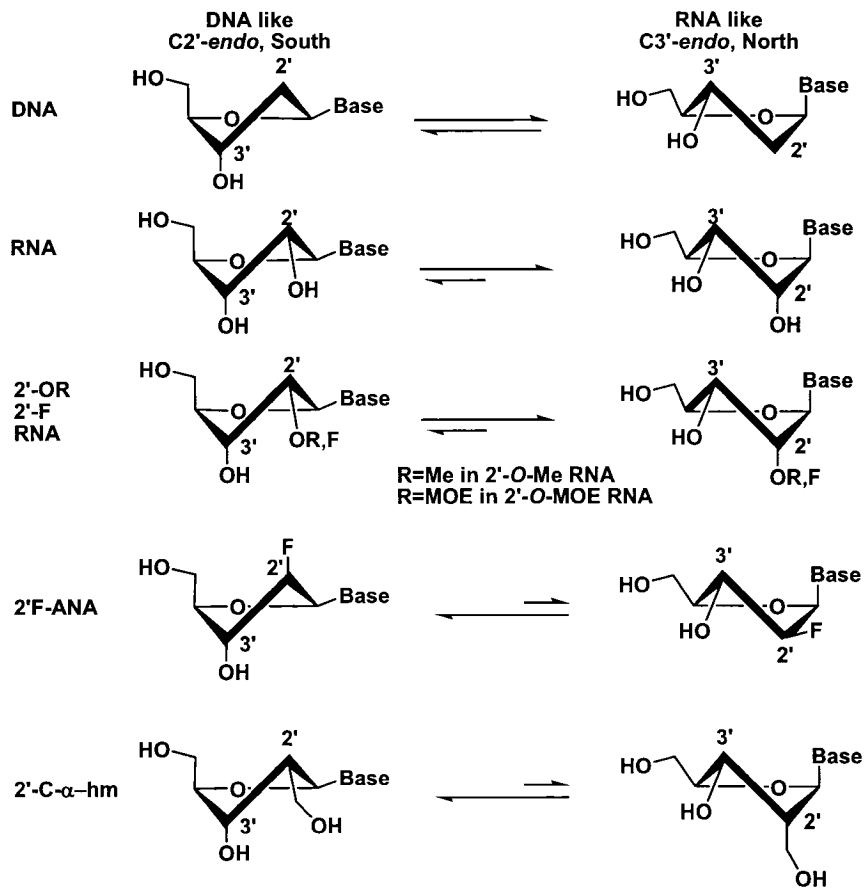


Figure 1.17: Preferred sugar conformations of modified nucleosides

(DNA like) conformation (Figure 1.17), since the gauche effect of 2' group with the sugar ring oxygen is removed and steric concerns become dominant. According to the preorganization theory, chemical modifications leading to C2'-*endo* conformations generally result in lower affinity for mRNA. The 2'-α-*C*-hydroxymethyl modification, for instance, causes a 3°C decrease in a *DNA*:RNA duplex where the DNA strand is modified.[177]

2'-Deoxy-2'-fluoro-D-arabinonucleic acids (2'F-ANA, Figure 1.16e) is an isomer of 2'F-RNA (Figure 1.16d). 2'F-ANA is expected to adopt C2'-*endo* (DNA like) conformation (Figure 1.17) due to the same reason that the most electronegative "up" fluorine prefers in a pseudo axial to the gauche interaction with the O4'-ring oxygen (F-C2'-C1'-O4'). Previous studies have shown that the 2'F-ANA actually favors a *south/east* (C2'-*endo*/O4'-*endo*) conformation.[186] Interestingly, the incorporation of 2'F-ANA units, leads to ~ +1 °C/modification in binding affinity with RNA compared with DNA (i.e. 2'F-ANA/RNA>DNA/RNA),[187] though stabilization caused by 2'F-ANA is much less than what is observed with 2'F-RNA. Pre-organization of 2'F-ANA sugar, or/and increase base stacking on 2'F-ANA sugar could lead to the observed stability difference.[187] 2'F-RNA, and more recently 2'F-ANA, have been shown to be important chemical modifications for antisense and siRNA applications.[180]

1.9.1.1 Phosphodiester Backbone Modification: 2',5'-Linked RNA

Natural phosphodiester (PO) backbone is often modified for antisense and siRNA applications. Much success has been made for phosphorothioate (PS) linkage in which a sulfur atom replaces a non-bridging oxygen in the PO backbone (Figure 1.15) as in the first antisense drug Vitravene™. The PS-DNA modification results in enhanced stability to enzymatic degradation and maintained ability to activate RNase H. However, the PS-DNA duplex formed with the target RNA has a lower T_m (i.e. lower affinity) than the phosphodiester compound (PO-DNA), and phosphorothiates exhibit non-specific binding

with cellular proteins, inducing undesirable effects (toxicity).[188,189] Another interesting RNA analogue is the regioisomeric 2',5'-linked RNA (Figure 1.18). Actually 2',5'-RNA is found in some biological processes in cells such as lariat RNA splicing intermediates (Figure 1.5).[15] 2',5'-RNA are not only interesting from a structural point of view, but also have potential use in the down-regulation of gene expression.[190,191] For instance, previous studies show that DNA/2',5'-RNA chimeras and 2',5'-DNA show less nonspecific binding to plasma and cellular proteins minimizing toxicity in antisense applications.[192] 2',5'-RNA is more nuclease resistant than unmodified DNA.[192] In addition, 2',5'-RNAs are able to associate with complementary single-stranded RNA [190] as well as duplex DNA [193] and, as such, can potentially be used to down-regulate gene expression *via* the antisense and antigene approaches.[191,192,194,195] It is also well-documented that annealing two normal RNA strands is more favorable than annealing a 2′,5′-RNA strand with a normal RNA strand.[190,196-198] Furthermore, mutually complementary 2′,5′-RNA strands have the ability to associate, but exhibit much lower T_m than those of the corresponding RNA:RNA or RNA:2′,5′-RNA duplexes. A comparison of the T_m values of various duplexes of mixed base composition revealed the

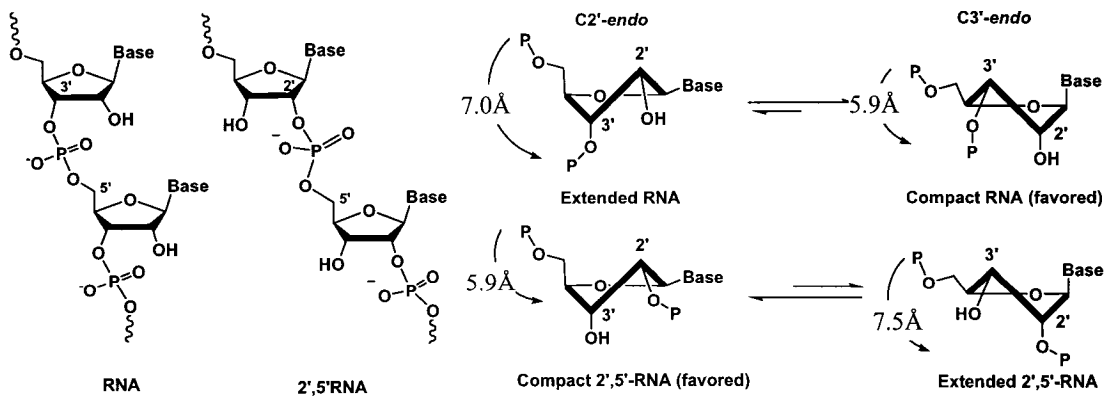


Figure 1.18: RNA, 2',5'RNA and their favored sugar conformations (adapted from Premraj et al. 2001 [199])

following order of duplex thermal stability: RNA:RNA > DNA:DNA ≈ DNA:RNA > RNA:2′,5′-RNA > 2′,5′-RNA: 2′,5′-RNA > DNA: 2',5'-RNA (undetected).[196] Molecular modeling and circular dichroism (CD) studies of these duplexes revealed that RNA: 2',5'-RNA hybrids adopt a continuous A-type helix structure similar to that of native RNA,[196,199] but have smaller interstrand phosphate-phosphate distances (by *ca.* 1 Å). This may account, at least in part, for the lower thermal stability of RNA:2',5'-RNA relative to RNA:RNA and RNA:DNA helices.[196] Another interesting structural feature of 2',5'-RNA is that it favors a "compact" structure which is defined by the intramolecular phosphate-phosphate distance (Figure 1.18). A compact 2',5'-RNA adopts C2'-*endo* sugar pucker, while extended 2',5'-RNA adopts C3'-*endo*. In RNA, a favored compact structure adopts the C3'-*endo* pucker due to the shift of phosphodiester linkage (Figure 1.18).[200]

1.9.2 Chemical Modifications for Aptamers

The target binding activity of aptamers is strongly dependent on the folding or 3D structure of the oligonucleotide strand, so high thermal stability and nuclease resistance is very desirable. Important lessons learned from the chemical modification of AS-ODNs and siRNA have informed aptamer modifications and several methods have been devised to improve the stability of aptamers towards endo/exonucleases. Chemical modifications of aptamers are made after *in vitro* selection, called post-SELEX modifications, as was previously mentioned with siRNA and antisense modifications.[201] Macugen® is an example of post-SELEX modification, composed of 2'-O-Me purines, and other 3',5' terminal modifications (Figure 1.15).[174] Precaution must be taken to keep the correct 3-D structure of the modified oligomer so as to maintain the binding affinity and selectivity to a target. It has been shown that aptamers are often sensitive to chemical modifications. For example, several studies have been directed at modifying the thrombin-binding aptamer d($G_2T_2G_2TGTG_2T_2G_2$) (Figure 1.12G),[97] but very few, if any, have led to an improvement over the original molecule. Heckel and Mayer reported that the introduction of thymidine residues modified with a nitrophenylpropyl moiety (T-NPP)

generally abolished interaction of the aptamer with thrombin.[202]. Di Giusto and King reported the synthesis of circular aptamers targeting thrombin with improved nuclease resistance and anticoagulant potency compared to a classical thrombin DNA aptamer.[203] However, circularization of the aptamers produced a mixture of constructs and required a ligase enzyme, making the method very difficult to scale-up. Other attempts to circularize the thrombin-binding DNA aptamer via chemical methods abolished the anti-thrombin activity [204]. Recently, Seela and coworkers reported the insertion of the hairpin-forming sequence GCGAAG into the central loop of the thrombin-binding aptamer, creating a G-quadruplex fused to a mini-hairpin structure. The T_m data suggested that the mini-hairpin induced a structural change in the aptamer section but binding to thrombin was not investigated.[205] Saccà *et al.* used spectroscopy to study a number of modifications, including the effects of backbone charge, atom size, base substitution and modification at the sugar 2'-position. All the sugar (ribose, 2'-*O*-methylribose) and phosphate (methylphosphonate, phosphorothioate) modifications led to a reduction in the thermal stability of the thrombin binding aptamer. 2'-*O*-methylribose modification also disrupt the G-quadruplex conformation.[206] Another study suggests that incorporation of 2'-*O*-methylribose results in structural changes or loss of activity in RNA aptamers binding to human transcription factor NF-κB.[201]

Another way to modify aptamers is called mirror-design (or "Spiegelmer"),[207] in which standard SELEX methods are used to select a normal RNA aptamer (D-RNA) against the enantiomer of the target protein which is composed of D-amino acids. The resulting RNA aptamer (D-RNA) sequence is then converted to its enantiomeric form, L-RNA, resulting in an aptamer with high binding affinity to the native protein molecule (L-amino acids) and high resistance against cleavage by nucleases. Unfortunately this strategy is limited to cases where an enantiomer of the target molecule is available.

1.9.2.1 Direct evolution with modified nucleoside 5'-triphospahtes (NTPs)

An alternative approach for modifying aptamers is the directed evolution in SELEX procedure with modified deoxyribonucleoside 5'-triphosphates (dNTPs) or ribonucleoside 5'-triphosphates (rNTPs).[208] This method ensures the generation of modified aptamers with desirable binding properties and biostability. Two extra steps are added to the standard SELEX cycle (Figure 1.14), one to generate a library of modified oligonucleotides through the biosynthesis of the initial unmodified oligonucleotides (DNA or RNA) pool by polymerase; the other to convert a smaller library of selected modified oligonucleotides into cDNA through modified template-mediated polymerase biosynthesis (e.g. reverse transcription converts modified RNA into cDNA in RNA-SELEX, Figure 1.14). A successful SELEX procedure using modified NTPs has two main requirements; the first is a polymerase which can incorporate the modified NTPs, and the second is a polymerase which can direct DNA synthesis through modified templates. The biosynthesis of modified dNTPs and polymerase-directed DNA synthesis on a modified template have been investigated in preparation for SELEX aptamers and other purposes (example 10-13 in Table 1.1).

In Table 1.2, a summary of catalytic (RNA or DNA enzymes) or selective binding aptamers that have been successfully generated by SELEX with modified NTPs is presented. Many of these NTPs modified at C5 in pyrimidines or C8 in purines can generate RNA or DNA enzymes with new catalytic functions. Very few modified nucleoside 5'-triphosphates are compatible with the SELEX process since few modified rNTPs are substrates of RNA polymerases and few modified oligonucleotides are faithful templates for reverse transcriptases. It is interesting to note that most RNA aptamers evolved via SELEX for therapeutics utilize 2'-fluoro or 2'-amino modified rNTPs (Table 2). Other successful SELEX processes with modified rNTPs include ribonucleoside 5'-(α-P-borano)-triphosphates (BH_3-RNA),[209] ribonucleoside 5'-(α-thio)triphosphates (S-RNA)[210] and 4'-thio-ribonucleoside 5'-triphosphates (4'S-RNA)[211] as shown in

Table 1.1: Biosynthesis of modified nucleoside 5'-triphosphate and polymerase directed DNA synthesis on a modified template

No.	Modified triphosphates	Polymerases and biosynthesis	Year	Reference
Modified NTPs and DNA polymerases				
1*	Cyclohexenyl adenine 5 -triphosphates (CeATP)	Substrate of Vent (exo-), Taq DNA polymerase and HIV-1 RT; *Taq* and HIV-1 RT were able to incorporate seven consecutive CeNA; all three enzyme can mediate DNA biosynthesis (6 nt) on a CeNA template	**2005**	Kempeneers et al. [212]
2*	α-L-Threofuranosyl nucleoside 5'-triphosphates (TNA-5'-TP)	Therminator™ was able to synthesize 50 nt through a DNA template	**2005**	Horhota et al. [213]
3	TNA templates	DNA polymerase-mediated DNA synthesis on a chimeric DNA/TNA template (9 TNA); trace of full-length product was observed by *Bst*, T7 DNA polymerase (exo-), its mutated commercial version Sequensase, MMLV-RT and its mutated commercial version SuperScript II	**2003**	Chaput et al. [214]
4	TNA-5'-TPs	TNA synthesis by DNA polymerase; Deep Vent (exo-), the most effective enzyme tested, was able to extend three TNA after 24hours	**2003**	Chaput and Szostak [215]
5	2'-5'-linked DNA template	2'-5'-liked DNA serves a template for polymerase-directed DNA synthesis. Only Kf (exo-) and HIV-1 RT could extend 4 nt to give full-length products	**2003**	Sinha et al. [216]
6	2'-OMe	Evolution of a DNA polymerase (*Taq*) variant for 2'-OMe rNTP incorporations	**2003**	Fa et al. [217]
7	All four 2'F-rNTPs	All four 2'F-rNTPs are substrate of human DNA pol γ; 2'F-rUTP, 2'F-rCTP and 2'F-rGTP are substrate of pol α	**2000**	Richardson et al. [218]
8	2'OMe-rUTP and 2'OMe-rCTP	Human DNA pol α and γ could not incorporate 2'OMe-NTPs	**2000**	Richardson et al. [218]
9*	2'-Deoxyribonucleoside 5'-(α-methyl)triphosphates (P-Me dNTPs)	Biosynthesis of modified DNA by *E. coli* DNA polymerase I	**1997**	Dineva [219]
10	All four 2'F-rNTPs	Substrates of Pfu (exo-), Vent (exo-), Deep Vent (exo-), UlTma DNA pol	**1997**	Ono et al. [220]
11	2'F-rCTP	Mouse cell pol *γ* and *Xenopus laevis* pol *α*, wheat germ DNA pol A and C, and AMV reverse transcriptase could incorporate 2'F-dCTP but the elongation was difficult (chain termination)	**1985**	Aoyama et al. [221]

Table 1.1 (continued)

No.	Modified triphosphates	Polymerases and biosynthesis	Year	Reference
Modified NTPs and DNA polymerases				
12	2'F-araTTP, 5-Me-2'F-araCTP and 5-I-2'F-araCTP	Chain terminator of HSV-1 and HSV-2 DNA pol	**1981**	Ruth and Cheng [222]
13	2'F-rUTP, 2'F-rCTP	2'F-rCTP is a substrate of pol II but not III; *E. coli* pol II and III; 2'F-dUTP is not a substrate for both enzymes	**1978**	Helfman et al. [223]
Modified NTPs and RNA polymerases				
15*	Adenosine 5'-(α-P-Seleno) triphosphate (ATPαSe)	T7 RNA polymerase only incorporates one diasteromer of ATPαSe (ATPαSe I, *Sp*)	**2006**	Carrasco et al. [224]
16	2'-OMe	Evolution of a T7 RNA polymerase variant for 2'-OMe rNTP incorporations	**2004**	Chelliserrykattil and Ellington [225]
17*	2'-dexoy-2'-α-C-branched-homo-udine-5'-triphosphate	Substrate of T7 RNA polymerase in the presence of Mn^{2+} (extension of 20-21 nt in RNA strand)	**2004**	Pavey et al. [226]
18	2'-OMe, 2'-N_3-rNTPs	Y639F/H784A T7 RNA polymerase (double mutations) can incorporate 2'-OMe, 2'-N_3-rNTPs	**2002**	Padilla and Sousa [227]
19	2'-deoxy-2'-thio-CTP (2'S-rCTP)	A mutant T7 polymerase (Y639F) can incorporate 2'S-rCTP and replace all CTP in HDV (hepatitis delta virus) ribozymes	**1998**	Raines and Gottleib [228]
20	2'F pyrimidines (2'F-rUTP, 2'F-rCTP); 2'NH_2 pyrimidines (2'NH_2-rUTP, 2'NH_2-rCTP);	They are substrate of T7 RNA pol which was able to synthesis 2500 nt modified RNAs	**1992**	Aurup et al. [229]

*

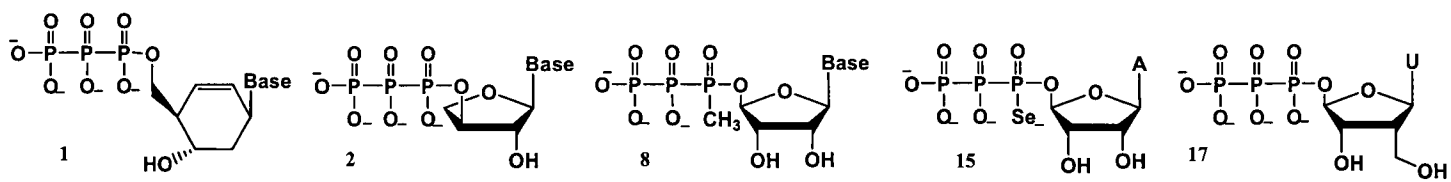


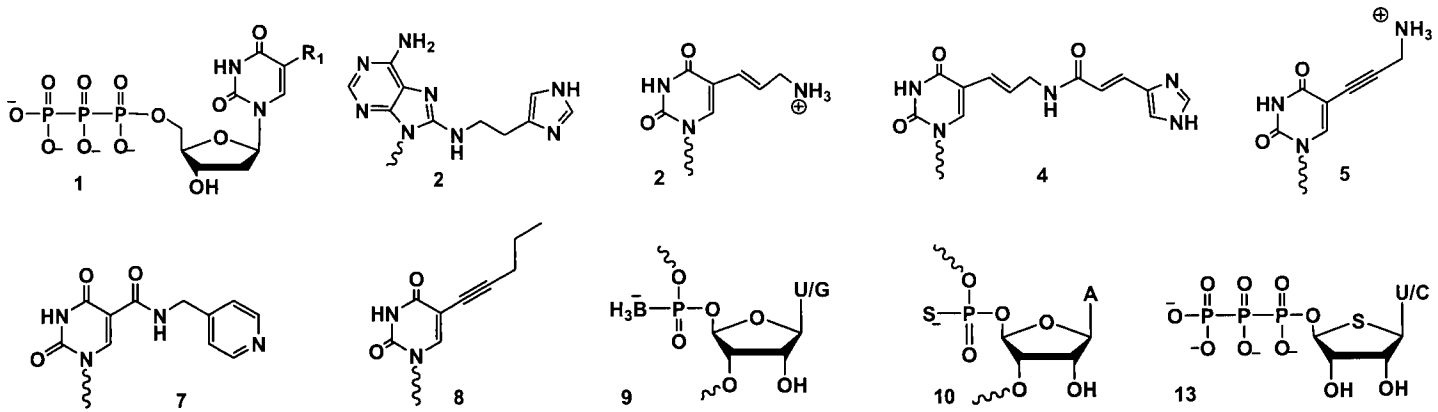
Table 1.2. Successful SELEX reported in the literature

No.	Modified triphosphates	Polymerases used	Class	Target/function**	Year	Reference
Base modifications						
1*	2'-deoxyuridine triphosphates bearing amino acids at the C5 position	KOD Dash DNA polymerase,	DNA aptamer	Protein-like DNA molecules	**2006**	Kuwahara et al. [230]
2*	8-[2-(4-Imidazolyl)ethylamino]deoxyadenosine and 5-(3-aminoallyl)-deoxyuracil	Sequenase 2.0 (genetically engineered variant of T7 DNA polymerase)	Deoxy-ribozyme	RNase (internal rC with no Mg2+)	**2001**	Perrin et al. [231]
3	*N*-6-Aminohexyl adenosine	T7 RNA pol	Ribozyme	Ligation to its 5′ end (250 fold more activie)	**2000**	Teramoto et al [232]
4*	C5-Imidizole-uracil analog (unnamed)	Superscript II Reverse Transcriptase	Deoxy-ribozyme	Sequence directed RNase	**2000**	Santoro et al. [233]
5*	5-(3′-aminopropynyl)-deoxyuracil	Vent DNA polymerase	Deoxy aptamer	ATP	**1999**	Battersby et al.[234]
6	5-Imidizole-uracil	T7 RNA polymerase	Ribozyme	Amide synthase	**1997**	Wiegand et al. [235]
7*	5-Pyridylmethylcarboxamid-uracil	Not mentioned	Ribozyme	Diels-Alderase (800 fold more active)	**1997**	Tarasow et al.[236]
8*	5-(1-Pentynyl)-deoxyuracil	Vent DNA polymerase More info about thrombin binding interaction!!	Deoxy aptamer	Thrombin	**1994**	Latham et al. [237]
Backbone modifications						
9*	Guanosine and Uridine 5' (α-P-borano)-triphosphates (BH_3-G(U)TP)	AmpliScribe™ Transcription Kits	RNA aptamer	ATP	**2002**	Lato et al. [209]
10*	α-S-rNTP, α-S-dNTP (phosphorothioate linked RNA, or DNA) incorporation	T7 RNA pol in transcription with S-rNTPs; HIV-RT in reverse transcription with S-dNTPs	Detailed optimization of the aptamer selection process with S-rNTPs and S-dNTPs; direct PCR amplification with S-dNTPs to have PS-DNA libraries		**2000**	Andreola et al.
11	α-S NTPs (phosphorothioate linked RNA, S-RNA)	T7 RNA polymerase	RNA aptamer	bFGF	**1998**	Jhaveri et al.

Table 1.2 (continued)

12	α-S dNTPs (phosphorothioate linked DNA, or S-DNA)	*Taq* polymerase	Deoxy aptamer	NF-IL6 transcription factor	**1998**	King et al. [238]
Sugar modification						
13*	4'-Thio-ribonucleoside 5'-triphosphates (4'-S-UTP and 4'-S-CTP)	T7 RNA polymerase	RNA aptamer	Thrombin	**2005**	Kato et al. [211]
14	2'NH_2-dCTP	T7 RNA polymerase	Ribozyme	metalation reaction of porphyrin	**2001**	Kawazoe et al. [239]
15	2'NH_2-dUTP, 2'NH_2-dCTP	T7 RNA polymerase	Ribozyme	Trans cleavage of RNA	**2000**	Beaudry et al. [240]
16	2'F-dUTP, 2'F-dCTP	T7 RNA polymerase	Aptamer	VEGF(165), leading to the discovery of Macugen	**1998**	Ruckman et al.
17	2'-Fluoro pyrimidines 2'NH_2-dUTP, 2'NH_2-dCTP	T7 RNA polymerase	Aptamer	hKGF	**1997**	Pagratis et al. [241]
18	2'-Fluoro pyrimidines 2'NH_2-dUTP, 2'NH_2-dCTP	T7 RNA polymerase	Aptamer	IFN-γ	**1997**	Kubik et al. [242]
19	2'NH_2-dUTP, 2'NH_2-dCTP	T7 RNA polymerase	Aptamer	VPF/VEGF	**1995**	Green et al.[243]
20	2'NH_2-dUTP, 2'NH_2-dCTP	T7 RNA polymerase	Aptamer	bFGF	**1995**	Jellinek et al.[244]
21	2'NH_2-dUTP, 2'NH_2-dCTP	T7 RNA polymerase	Aptamer	Human neutrophil elastase	**1994**	Lin et al. [245]

*



** Abbreviations: bFGF, basic fibroblast growth factor; hKGF, human keratinocyte growth factor; IFN-γ, interferon γ; NF-IL6, nuclear factor for human interleukin 6; VEGF, vascular endothelial growth factor; VPF, vascular permeability factor

Table 1.2, 2'-OMe ribonucleosides are an important post-SELEX modification. Unfortunately 2'-OMe rNTPs are not substrates of T7 RNA polymerase, although recent reports have shown that genetically engineered RNA or DNA polymerases actually can incorporate 2'-OMe rNTPs (example 14, 16 in Table 1.1) but the procedure is usually tedious. There are even fewer examples for DNA SELEX as most studies are still at the initial stage of searching for a DNA polymerase that efficiently accepts modified dNTPs as substrates (example 1-5 in Table 1.1). For example, α-L-Threofuranosyl nucleoside 5'-triphosphates (TNA-5'-TP) are not easily accepted by many common DNA polymerases[246] and efficient oligonucleotide (bio)synthesis requires very specialized mutant DNA polymerases.[247] TNA synthesis on a DNA template required at least 1 day incubation with Deep Vent (exo-) for an insertion of three contiguous TNA residues[215] and DNA synthesis on a TNA template by common polymerases such as *Bst* Pol I and MMLV reverse transcriptase required at least 1 hour to obtain full-length extension products.[214]

The general poor enzymatic recognition of modified NTPs underscores the difficulties involved in using DNA/RNA polymerases to synthesize unnatural oligonucleotides. Accordingly, there is a need for modified NTPs that are faithful substrates for DNA and RNA polymerases. Ideally, the resulting modified oligonucleotide would also serve as template for cDNA synthesis, so that the selection and amplification steps of the SELEX protocols could be applied.

2.0 Thesis Objectives

During the past decades, the use of oligonucleotides as therapeutics and diagnostics has drawn great attention. Chemical modifications are generally required to offer flexibility in the design and utility of these molecules, and to improve their "drug-like" properties such as efficacy, selectivity and delivery. These requirements have been clearly seen in the development of antisense, siRNA and aptamer therapeutics. My thesis focused on two

important chemical modifications, namely, 1-(2-deoxy-2-*C*-α-hydroxymethyl-β-D-ribofuranosyl)thymine (or 2'-α-*hm*-dT, abbreviated as "**H**") and 2'-deoxy-2'-fluoro-D-arabinonucleic acid (FANA), and their applications.

Chapter II provides the total synthesis of **H** and the first investigation of oligoribonucleotides containing this chemical modification. The shorter inter-strand phosphate-phosphate distance of 2',5'-RNA duplexes inspired us to consider this 2' methylene modification. We hypothesized that the extended 2',5'-RNA resulting from incorporation of **H** could increase the P-P distance so as to decrease the charge repulsion interactions and stabilize duplex formation. Both the 2'-CH_2O-phosphoramidite and 3'-*O*-phosphoramidite derivatives of **H** were synthesized and incorporated into both 2',5'-RNA and RNA chains. Hybridization was evaluated by T_m and CD experiments. The thermal stability and conformation of a 2',5'-linked RNA hairpin modified with **H** units was also investigated.

Chapter III presents one of the potential applications of oligonucleotides incorporating the **H** modification, namely a novel branched RNA structure for the inhibition of yeast debranching enzyme (yDBR). The preparation, debranching activity and inhibition of several Y-shaped RNAs were evaluated.

Chapter IV deals with the first study of G-quadruplexes modified with 2'-deoxy-2'-fluoro-D-arabinonucleotides (2'F-ANA). The effect of 2'F-ANA in the G-quadruplex was investigated in the thrombin-binding DNA aptamer d($G_2T_2G_2TGTG_2T_2G_2$), an anti-HIV phosphorothioate aptamer PS-d($T_2G_4T_2$), and a DNA telomeric sequence d($G_4T_4G_4$) via T_m and CD experiments. 2'F-ANA-modified thrombin-binding aptamers were further evaluated in terms of their nuclease resistance in 10% fetal bovine serum (FBS) and the binding affinity to thrombin.

Chapter V reports the first study on the biosynthesis of 2'F-ANA with 2'-deoxy-2'-fluoro-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs) catalyzed by DNA polymerases. Primer extension reactions were conducted to evaluate nine different

DNA polymerases for their ability to incorporate 2'F-araNTPs, as well the ability of these polymerases to accept FANA-DNA chimeras as template strands for the synthesis of a complementary DNA or a FANA-DNA chimeric strand. The fidelity of incorporation of 2'F-araNTPs by Deep Vent (exo-) (DV), $9^{o}N_{m}^{TM}$ (9N) and Phusion^{TM} High Fidelity (Ph) DNA polymerases was also evaluated.

Chapter VI reports a follow-up study of Chapter V, showing the first comparison studies on the efficiency of incorporation of 2'F-araNTPs and 2'-deoxy-2'-fluoro-D-ribonucleoside 5'-triphosphates (2'F-rNTPs) by four DNA polymerases, namely, DV, 9N, HIV-1 reverse transcriptase and Moloney Murine Leukemia Virus (MMLV) reverse transcriptase.

References

1. Saenger, W. (**1984**). *Principles of Nucleic Acid Structure.* Oxford University Press: New York, USA.

2. Blackburn, G.M. and Gait, M.J. (**1996**). *Nucleic Acids in Chemistry and Biology.* Oxford University Press: New York, USA.

3. Watson, J.D. and Crick, F.H.C. (**1953**). Molecular structure of nucleic acids - a structure for deoxyribose nucleic acid. *Nature*, 171, 737-738.

4. Watson, J.D. and Crick, F.H.C. (**1953**). Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171, 964-967.

5. International Human Genome Sequencing Consortium. (**2001**). Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.

6. Westheimer, F.H. (**1987**). Why nature chose phosphates. *Science*, 235, 1173-1178.

7. Donohue, J. and Trueblood, K.N. (**1960**). Base pairing in DNA. *J. Mol. Biol.*, 27, 363-371.

8. Hascheme.A.E.V. and Rich, A. (**1967**). Nucleoside confromations - an analysis of steric barriers to rotation about gylcosidic bond. *J. Mol. Biol.*, 27, 369-384.

9. Altona, C., Geise, H.J. and Romers, C. (**1968**). Conformation of non-aromatic ring compounds. 25. Geometry and conformation of ring D in some steroids from X-ray structure determinations. *Tetrahedron*, 24, 13-&.

10. Altona, C. and Sundaralingam, M. (**1972**). Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.*, 94, 8205-8212.

11. Hall, L.D. (**1963**). Conformations of some ribofuranosides. *Chem. Ind.*, 950-951.

12. Jardetzky, C.D. (**1960**). Proton magnetic resonance studies on purines, pyrimidines, ribose nucleosides and nucleotides. 3. Ribose conformation. *J. Am. Chem. Soc.*, 82, 229-233.

13. Chow, L.T., Gelinas, R.E., Broker, T.R. and Roberts, R.J. (**1977**). Amazing sequence arrangement at 5' ends of adenovirus-2 messenger-RNA. *Cell*, 12, 1-8.

14. Berget, S.M., Moore, C. and Sharp, P.A. (**1977**). Spliced segments at 5# terminus of

adenovirus 2 late messenger-RNA. *Proc. Natl. Acad. Sci. U. S. A.*, 74, 3171-3175.

15. Wallace, J.C. and Edmonds, M. (**1983**). Polyadenylated nuclear RNA contains branches. *Proc. Natl. Acad. Sci. USA*, 80, 950-954.

16. Sharp, P.A. (**1987**). Splicing of messenger-RNA precursors. *Science*, 235, 766-771.

17. Murphy, W.J., Watkins, K.P. and Agabian, N. (**1986**). Identification of a novel Y-branch structure as an intermediate in trypanosome messenger-RNA processing - evidence for *trans* splicing. *Cell*, 47, 517-525.

18. Reed, R. and Maniatis, T. (**1988**). The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes Dev.*, 2, 1268-1276.

19. Zhuang, Y., Goldstein, A.M. and Weiner, A.M. (**1989**). UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl. Acad. Sci. U. S. A.*, 86, 2752-2756.

20. Remaud, G., Balgobin, N., Glemarec, C. and Chattopadhyaya, J. (**1989**). Why do all lariat RNA introns have adenosine as the branch point nucleotide - conformational studies on the implication of the branch-point modification by guanine, uracil or cytosine in the naturally-occurring branched tetranucleotide by ^{1}H-NMR and ^{31}P-NMR spectroscopy. *Tetrahedron*, 45, 1537-1548.

21. Ruskin, B. and Green, M.R. (**1985**). An RNA processing activity that debranches RNA lariats. *Science*, 229, 135-140.

22. Jacquier, A. and Rosbash, M. (**1986**). Efficient transsplicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon intron interaction. *Science*, 234, 1099-1104.

23. Chapman, K.B. and Boeke, J.D. (**1991**). Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell*, 65, 483-492.

24. Khalid, M.F., Damha, M.J., Shuman, S. and Schwer, B. (**2005**). Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Res.*, 33, 6349-6360.

25. Jacquier, A. and Rosbash, M. (**1986**). RNA splicing and intron turnover are greatly diminished by a mutant yeast branch point. *Proc. Natl. Acad. Sci. U. S. A.*, 83, 5835-5839.

26. Nam, K., Hudson, R.H.E., Chapman, K.B., Ganeshan, K., Damha, M.J. and Boeke, J.D. (**1994**). Yeast lariat debranching enzyme. Substrate and sequence specificity. *J. Biol. Chem.*,

269, 20613-20621.

27. Carriero, S., Mangos, M.M., Agha, K.A., Noronha, A.M. and Damha, M.J. (**2003**). Branchpoint sugar stereochemistry determines the hydrolytic susceptibility of branched RNA fragments by the yeast debranching enzyme (yDBR). *Nucleosides, Nucleotides & Nucleic Acids*, 22, 1599-1602.

28. Ooi, S.L., Dann, C., III, Nam, K., Leahy, D.J., Damha, M.J. and Boeke, J.D. (**2001**). RNA lariat debranching enzyme. *Methods Enzymol.*, 342, 233-248.

29. Carriero, S. and Damha, M.J. (**2003**). Inhibition of pre-mRNA splicing by synthetic branched nucleic acids. *Nucleic Acids Res.*, 31, 6157-6167.

30. Damha, M.J. and Ogilvie, K.K. (**1988**). Synthesis and spectroscopic analysis of branched RNA fragments: messenger RNA splicing intermediates. *J. Org. Chem.*, 53, 3710-3722.

31. Damha, M.J., Ganeshan, K., Hudson, R.H.E. and Zabarylo, S.V. (**1992**). Solid-phase synthesis of branched oligoribonucleotides related to messenger RNA splicing intermediates. *Nucleic Acids Res.*, 20, 6565-6573.

32. Carriero, S. and Damha, M.J. (**2002**). Solid-phase synthesis of branched oligonucleotides. In Beaucage, S. L. (ed.), *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons., pp. 4.14.11-14.14.32.

33. Merrifield, R.B. (**1963**). Solid phase peptide synthesis .1. Synthesis of a tetrapeptide. *J. Am. Chem. Soc.*, 85, 2149-&.

34. Letsinger, R.L., Kornet, M.J., Mahadevan, V. and Jerina, D.M. (**1964**). Reactions on polymer supports. *J. Am. Chem. Soc.*, 86, 5163-&.

35. Letsinge, R.l. and Mahadeva, V. (**1965**). Oligonucleotide synthesis on a polymer support. *J. Am. Chem. Soc.*, 87, 3526.

36. Beaucage, S.L. and Caruthers, M.H. (**1981**). Deoxynucleoside phosphoramidites - a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.*, 22, 1859-1862.

37. Caruthers, M.H., Barone, A.D., Beaucage, S.L., Dodds, D.R., Fisher, E.F., Mcbride, L.J., Matteucci, M., Stabinsky, Z. and Tang, J.Y. (**1987**). Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method. *Methods Enzymol.*, 154, 287-313.

38. Letsinger, R.L. and Lunsford, W.B. (**1976**). Synthesis of thymidine oligonucleotides byphosphite triester intermediates. *J. Am. Chem. Soc.*, 98, 3655-3661.

39. Letsinger, R.L. and Heavner, G.A. (**1975**). Synthesis of phosphoromonoamidate diester nucleotides via phosphite-azide coupling method. *Tetrahedron Lett.*, 147-150.

40. Damha, M.J., Giannaris, P.A. and Zabarylo, S.V. (**1990**). An improved procedure for derivatization of controlled-pore glass beads for solid-phase oligonucleotide synthesis. *Nucleic Acids Res.*, 18, 3813-3821.

41. Guzaev, A.P. and Manoharan, M. (**2003**). A conformationally preorganized universal solid support for efficient oligonucleotide synthesis. *J. Am. Chem. Soc.*, 125, 2380-2381.

42. Schaller, H., Weimann, G., Khorana, H.G. and Lerch, B. (**1963**). Studies on polynucleotides. 24. Stepwise synthesis of specific deoxyribopolynucleotides (4) - protected derivatives of deoxyribonucleosides and new syntheses of deoxyribonucleoside-3 phosphates. *J. Am. Chem. Soc.*, 85, 3821-3827.

43. Hakimelahi, G.H., Proba, Z.A. and Ogilvie, K.K. (**1982**). New catalysts and procedures for the dimethoxytritylation and selective silylation of ribonucleosides. *Can. J. Chem.*, 60, 1106-1113.

44. Ogilvie, K.K., Beaucage, S.L., Schifman, A.L., Theriault, N.Y. and Sadana, K.L. (**1978**). Synthesis of oligoribonucleotides. II. Use of silyl protecting groups in nucleoside and nucleotide chemistry. VII. *Can. J. Chem.*, 56, 2768-2780.

45. Ogilvie, K.K., Thompson, E.A., Quilliam, M.A. and Westmore, J.B. (**1974**). Selective protection of hydroxyl-groups in deoxynucleosides using alkylsilyl reagents. *Tetrahedron Lett.*, 2865-2868.

46. Corey, E.J. and Venkates.A. (**1972**). Protection of hydroxyl groups as *tert*-butyldimethylsilyl derivatives. *J. Am. Chem. Soc.*, 94, 6190-&.

47. Gasparutto, D., Livache, T., Bazin, H., Duplaa, A.M., Guy, A., Khorlin, A., Molko, D., Roget, A. and Teoule, R. (**1992**). Chemical synthesis of a biologically-active natural transfer-RNA with its minor bases. *Nucleic Acids Res.*, 20, 5159-5166.

48. Damha, M.J. and Zabarylo, S. (**1989**). Automated solid-phase synthesis of branched oligonucleotides. *Tetrahedron Lett.*, 30, 6295-6298.

49. Puglisi, J.D. and Tinoco, I., Jr. (**1989**). Absorbency melting curves of RNA. *Methods*

Enzymol., 180, 304-325.

50. Rich, A. and Tinoco, I., Jr. (**1960**). The effect of chain length upon hypochromism in nucleic acids and polynucleotides. *J. Am. Chem. Soc.*, 82, 6409-6411.

51. Jaeger, J.A., Santalucia, J. and Tinoco, I., Jr. (**1993**). Determination of RNA structure and thermodynamics. *Annu. Rev. Biochem.*, 62, 255-287.

52. Chastain, M. and Tinoco, I. (**1991**). Structural elements in RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, 41, 131-177.

53. Gray, D.M., Ratliff, R.L. and Vaughan, M.R. (**1992**). Circular-dichroism spectroscopy of DNA. *Methods Enzymol.*, 211, 389-406.

54. Tunissch-Schneider, M.J.B. and Maestre, M.F. (**1970**). Circular dichroism spectra of oriented and unoriented deoxyribonucleic acid films - a preliminary study. *J. Mol. Biol.*, 52, 521.

55. Riesner, D., Steger, G., Zimmat, R., Owens, R.A., Wagenhofer, M., Hillen, W., Vollbach, S. and Henco, K. (**1989**). Temperature-gradient gel-electrophoresis of nucleic acids - analysis of conformational transitions, sequence variations, and protein-nucleic acid interactions. *Electrophoresis*, 10, 377-389.

56. Riesner, D., Steger, G., Wiese, U., Wulfert, M., Heibey, M. and Henco, K. (**1992**). Temperature-gradient gel electrophoresis for the detection of polymorphic DNA and for quantitative polymerase chain-reaction. *Electrophoresis*, 13, 632-636.

57. Rich, A. (**1993**). DNA comes in many forms. *Gene*, 135, 99-109.

58. Lebrun, A. and Lavery, R. (**1997**). Unusual DNA conformations. *Curr. Opin. Struct. Biol.*, 7, 348-354.

59. Pohl, F.M. and Jovin, T.M. (**1972**). Salt-induced cooperative conformational change of a synthetic DNA - equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.*, 67, 375-396.

60. Wang, A.H.J., Quigley, G.J., Kolpak, F.J., Crawford, J.L., Vanboom, J.H., Vandermarel, G. and Rich, A. (**1979**). Molecular-structure of a feft-handed double helical DNA fragment at atomic resolution. *Nature*, 282, 680-686.

61. Rippe, K., Fritsch, V., Westhof, E. and Jovin, T.M. (**1992**). Alternating d(G-A) sequences

form a parallel-stranded DNA homoduplex. *EMBO J.*, 11, 3777-3786.

62. Rippe, K. and Jovin, T.M. (**1989**). Substrate properties of 25-nt parallel-stranded linear DNA duplexes. *Biochemistry*, 28, 9542-9549.

63. Kettani, A., Bouaziz, S., Skripkin, E., Majumdar, A., Wang, W.M., Jones, R.A. and Patel, D.J. (**1999**). Interlocked mismatch-aligned arrowhead DNA motifs. *Structure*, 7, 803-815.

64. Liu, K.L., Miles, H.T., Frazier, J. and Sasisekharan, V. (**1993**). A novel DNA duplex - a parallel-stranded DNA helix with hoogsteen base-pairing. *Biochemistry*, 32, 11802-11809.

65. Chastain, M. and Tinoco, I. (**1992**). Poly(rA) binds poly(rG).poly(rC) to form a triple helix. *Nucleic Acids Res.*, 20, 315-318.

66. Gehring, K., Leroy, J.L. and Gueron, M. (**1993**). A tetrameric DNA-structure with protonated cytosine.cytosine base-pairs. *Nature*, 363, 561-565.

67. Gellert, M., Lipsett, M.N. and Davies, D.R. (**1962**). Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.*, 48, 2013-&.

68. Williamson, J.R. (**1994**). G-quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.*, 23, 703-730.

69. Williamson, J.R. (**1993**). Guanine quartets. *Curr. Opin. Struct. Biol.*, 3, 357-362.

70. Suhnel, J. (**2001**). Beyond nucleic acid base pairs: from triads to heptads. *Biopolymers*, 61, 32-51.

71. Gilbert, D.E. and Feigon, J. (**1999**). Multistranded DNA structures. *Curr. Opin. Struct. Biol.*, 9, 305-314.

72. Burkard, M.E., Turner, D.H. and Tinoco Jr, I. (**1999**). Appendix 1: Structures of base pairs involving at least two hydrogen bonds. In Gesteland, R. F., Cech, T. R. and Atkins, J. F. (eds.), *The RNA World*. Cold Spring Harbor Laboratory Press, pp. 675-680.

73. Blackburn, E.H. (**1991**). Structure and function of telomeres. *Nature*, 350, 569-573.

74. Rhodes, D. and Giraldo, R. (**1995**). Telomere structure and function. *Curr. Opin. Struct. Biol.*, 5, 311-322.

75. Phan, A.T. and Mergny, J.L. (**2002**). Human telomeric DNA: G-quadruplex, i-motif and

watson-crick double helix. *Nucleic Acids Res.*, 30, 4618-4625.

76. Pinnavaia, T.J., Marshall, C.L., Mettler, C.M., Fisk, C.I., Miles, H.T. and Becker, E.D. (**1978**). Alkali-metal Ion specificity in solution ordering of a nucleotide, 5'-guanosine monophosphate. *J. Am. Chem. Soc.*, 100, 3625-3627.

77. Davis, J.T. (**2004**). G-quartets 40 years later: From 5 '-GMP to molecular biology and supramolecular chemistry. *Angew Chem Int Edit*, 43, 668-698.

78. Sen, D. and Gilbert, W. (**1988**). Formation of parallel 4-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, 334, 364-366.

79. Henderson, E., Hardin, C.C., Walk, S.K., Tinoco, I. and Blackburn, E.H. (**1987**). Telomeric DNA oligonucleotides form novel intramolecular structures containing quanine quanine base-pairs. *Cell*, 51, 899-908.

80. Blackburn, E.H. (**1994**). Telomeres - no end in sight. *Cell*, 77, 621-623.

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

82. Kilpatrick, M.W., Torri, A., Kang, D.S., Engler, J.A. and Wells, R.D. (**1986**). Unusual DNA structures in the adenovirus genome. *J. Biol. Chem.*, 261, 1350-1354.

83. Fry, M. and Loeb, L.A. (**1994**). The fragile-X syndrome d(CGG)(N) nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci. U. S. A.*, 91, 4950-4954.

84. Makarov, V.L., Hirose, Y. and Langmore, J.P. (**1997**). Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, 88, 657-666.

85. Blackburn, E.H. (**1991**). Telomeres. *Trends Biochem. Sci.*, 16, 378-381.

86. Greider, C.W. and Blackburn, E.H. (**1985**). Identification of a specific telomere terminal transferase-activity in tetrahymena extracts. *Cell*, 43, 405-413.

87. Han, F.X.G., Wheelhouse, R.T. and Hurley, L.H. (**1999**). Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition. *J. Am. Chem. Soc.*, 121, 3561-3570.

88. Perry, P.J., Gowan, S.M., Reszka, A.P., Polucci, P., Jenkins, T.C., Kelland, L.R. and Neidle,

S. (**1998**). 1,4- and 2,6-Disubstituted amidoanthracene-9,10-dione derivatives as inhibitors of human telomerase. *J. Med. Chem.*, 41, 3253-3260.

89. Fang, G.W. and Cech, T.R. (**1993**). The beta-subunit of *Oxytricha* telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell*, 74, 875-885.

90. Fang, G. and Cech, T.R. (**1993**). Characterization of a G-quartet formation reaction promoted by the beta-subunit of the *Oxytricha* telomere-binding protein. *Biochemistry*, 32, 11646-11657.

91. Giraldo, R., Suzuki, M., Chapman, L. and Rhodes, D. (**1994**). Promotion of parallel DNA quadruplexes by a yeast telomere binding protein: a circular dichroism study. *Proc. Natl. Acad. Sci. U. S. A.*, 91, 7658-7662.

92. Kerwin, S.M. (**2000**). G-quadruplex DNA as a target for drug design. *Curr. Pharm. Des.*, 6, 441-471.

93. Shafer, R.H. and Smirnov, I. (**2000**). Biological aspects of DNA/RNA quadruplexes. *Biopolymers*, 56, 209-227.

94. Sun, H., Karow, J.K., Hickson, I.D. and Maizels, N. (**1998**). The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.*, 273, 27587-27592.

95. Fry, M. and Loeb, L.A. (**1999**). Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)(n). *J. Biol. Chem.*, 274, 12797-12802.

96. Hurley, L.H., Wheelhouse, R.T., Sun, D., Kerwin, S.M., Salazar, M., Fedoroff, O.Y., Han, F.X., Han, H.Y., Izbicka, E. and Von Hoff, D.D. (**2000**). G-quadruplexes as targets for drug design. *Pharmacol. Ther.*, 85, 141-158.

97. Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. and Toole, J.J. (**1992**). Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, 355, 564-566.

98. Wyatt, J.R., Vickers, T.A., Roberson, J.L., Buckheit, R.W., Klimkait, T., Debaets, E., Davis, P.W., Rayner, B., Imbach, J.L. and Ecker, D.J. (**1994**). Combinatorially selected guanosine-quartet structure is a potent inhibitor of human-immunodeficiency-virus envelope-mediated cell-fusion. *Proc. Natl. Acad. Sci. U. S. A.*, 91, 1356-1360.

99. Phan, A.T., Kuryavyi, V., Ma, J.B., Faure, A., Andreola, M.L. and Patel, D.J. (**2005**). An

interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 634-639.

100. Jing, N.J. and Hogan, M.E. (**1998**). Structure-activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug. *J. Biol. Chem.*, 273, 34992-34999.

101. Borman, S. (**2006**). Targeting telomerase. *Chem. Eng. News*, 84, 32-33.

102. Pilch, D.S., Plum, G.E. and Breslauer, K.J. (**1995**). The thermodynamics of DNA structures that contain lesions or guanine tetrads. *Curr. Opin. Struct. Biol.*, 5, 334-342.

103. Hardin, C.C., Perry, A.G. and White, K. (**2000**). Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers*, 56, 147-194.

104. Keniry, M.A. (**2000**). Quadruplex structures in nucleic acids. *Biopolymers*, 56, 123-146.

105. Lu, M., Guo, Q. and Kallenbach, N.R. (**1993**). Thermodynamics of G-tetraplex formation by telomeric DNAs. *Biochemistry*, 32, 598-601.

106. Salazar, M., Thompson, B.D., Kerwin, S.M. and Hurley, L.H. (**1996**). Thermally induced DNA center dot RNA hybrid to G-quadruplex transitions: Possible implications for telomere synthesis by telomerase. *Biochemistry*, 35, 16110-16115.

107. Ambrus, A., Chen, D., Dai, J.X., Bialis, T., Jones, R.A. and Yang, D.Z. (**2006**). Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, 34, 2723-2735.

108. Phillips, K., Dauter, Z., Murchie, A.I.H., Lilley, D.M.J. and Luisi, B. (**1997**). The crystal structure of a parallel-stranded guanine tetraplex at 0.95 angstrom resolution. *J. Mol. Biol.*, 273, 171-182.

109. Kang, C., Zhang, X.H., Ratliff, R., Moyzis, R. and Rich, A. (**1992**). Crystal structure of 4 stranded *Oxytricha* telomeric DNA. *Nature*, 356, 126-131.

110. Schultze, P., Smith, F.W. and Feigon, J. (**1994**). Refined solution structure of the dimeric quadruplex formed from the *Oxytricha* telomeric oligonucleotide d(GGGGTTTTGGGG). *Structure*, 2, 221-233.

111. Smith, F.W. and Feigon, J. (**1993**). Strand orientation in the DNA quadruplex formed from the *Oxytricha* telomere repeat oligonucleotide $d(G_4T_4T_4)$ in solution. *Biochemistry*, 32,

8682-8692.

112. Smith, F.W., Lau, F.W. and Feigon, J. (**1994**). $d(G_3T_4G_3)$ forms an asymmetric diagonally looped dimeric quadruplex with guanosine 5'-*syn-syn-anti* and 5'-*syn-anti-anti* *N*-glycosidic conformations. *Proc. Natl. Acad. Sci. U. S. A.*, 91, 10546-10550.

113. Keniry, M.A., Strahan, G.D., Owen, E.A. and Shafer, R.H. (**1995**). Solution structure of the Na^+ form of the dimeric guanine quadruplex $[d(G_3T_4G_3)]_2$. *Eur. J. Biochem.*, 233, 631-643.

114. Strahan, G.D., Keniry, M.A. and Shafer, R.H. (**1998**). NMR structure refinement and dynamics of the K^+-$[d(G_3T_4G_3)]_2$ quadruplex via particle mesh Ewald molecular dynamics simulations. *Biophys. J.*, 75, 968-981.

115. Kettani, A., Kumar, R.A. and Patel, D.J. (**1995**). Solution structure of a DNA quadruplex containing the fragile-X syndrome triplet repeat. *J. Mol. Biol.*, 254, 638-656.

116. Bouaziz, S., Kettani, A. and Patel, D.J. (**1998**). A K cation-induced conformational switch within a loop spanning segment of a DNA quadruplex containing G-G-G-C repeats. *J. Mol. Biol.*, 282, 637-652.

117. Kettani, A., Bouaziz, S., Gorin, A., Zhao, H., Jones, R.A. and Patel, D.J. (**1998**). Solution structure of a Na cation stabilized DNA quadruplex containing G center dot G center dot G center dot G and G center dot C center dot G center dot C tetrads formed by G-G-G-C repeats observed in adeno-associated viral DNA. *J. Mol. Biol.*, 282, 619-636.

118. Schultze, P., Macaya, R.F. and Feigon, J. (**1994**). Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, 235, 1532-1547.

119. Wang, K.Y., Krawczyk, S.H., Bischofberger, N., Swaminathan, S. and Bolton, P.H. (**1993**). The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. *Biochemistry*, 32, 11285-11292.

120. Wang, Y. and Patel, D.J. (**1994**). Solution structure of the tetrahymena telomeric repeat $d(T_2G_4)_4$ G-tetraplex. *Structure*, 2, 1141-1156.

121. Mohanty, D. and Bansal, M. (**1994**). Conformational polymorphism in telomeric structures - loop orientation and interloop pairing in $d(G_4T(N)G_4)$. *Biopolymers*, 34, 1187-1211.

122. Smith, F.W. and Feigon, J. (**1992**). Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature*, 356, 164-168.

123. Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A. and Feigon, J. (**1993**). Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U. S. A.*, 90, 3745-3749.

124. Wang, K.Y., Mccurdy, S., Shea, R.G., Swaminathan, S. and Bolton, P.H. (**1993**). A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA. *Biochemistry*, 32, 1899-1904.

125. Padmanabhan, K., Padmanabhan, K.P., Ferrara, J.D., Sadler, J.E. and Tulinsky, A. (**1993**). The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer. *J. Biol. Chem.*, 268, 17651-17654.

126. Padmanabhan, K. and Tulinsky, A. (**1996**). An ambiguous structure of a DNA 15-mer thrombin complex. *Acta Crystallographica, Section D: Biological Crystallography*, D52, 272-282.

127. Kelly, J.A., Feigon, J. and Yeates, T.O. (**1996**). Reconciliation of the x-ray and NMR structures of the thrombin-binding aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, 256, 417-422.

128. Schultze, P., Hud, N.V., Smith, F.W. and Feigon, J. (**1999**). The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the *Oxytricha nova* telomere repeat oligonucleotide $d(G_4T_4G_4)$. *Nucleic Acids Res.*, 27, 3018-3028.

129. Stoddart, C.A., Rabin, L., Hincenbergs, M., Moreno, M.E., Linquist-Stepps, V., Leeds, J.M., Truong, L.A., Wyatt, J.R., Ecker, D.J. and McCune, J.M. (**1998**). Inhibition of human immunodeficiency virus type 1 infection in SCID-hu Thy/Liv mice by the G-quartet-forming oligonucleotide, ISIS 5320. *Antimicrob. Agents Chemother.*, 42, 2113-2115.

130. Crooke, S.T. (**2004**). Progress in antisense technology. *Annu. Rev. Med.*, 55, 61-95.

131. Kurreck, J. (**2003**). Antisense technologies - improvement through novel chemical modifications. *Eur. J. Biochem.*, 270, 1628-1644.

132. Stephenson, M.L. and Zamecnik, P.C. (**1978**). Inhibition of Rous-Sarcoma viral-RNA translation by a specific oligodeoxyribonucleotide. *Proc. Natl. Acad. Sci. U. S. A.*, 75, 285-288.

133. Itaya, M. and Kondo, K. (**1991**). Molecular-cloning of a ribonuclease-H (RNase HI) gene

from an extreme thermophile *Thermus-thermophilus* HB8 - a thermostable RNase-H can functionally replace the *Escherichia-coli* enzyme *in vivo*. *Nucleic Acids Res.*, 19, 4443-4449.

134. Leonetti, J.P., Degols, G., Clarenc, J.P., Mechti, N. and Lebleu, B. (**1993**). Cell delivery and mechanisms of action of antisense oligonucleotides. *Prog Nucleic Acid Re*, 44, 143-166.

135. Woolf, T.M., Melton, D.A. and Jennings, C.G.B. (**1992**). Specificity of antisense oligonucleotides invivo. *Proc. Natl. Acad. Sci. U. S. A.*, 89, 7305-7309.

136. Crooke, S.T. (**1999**). Molecular mechanisms of action of antisense drugs. *BEB Acta-Gene Structure and Expression*, 1489, 31-44.

137. Mangos, M.M. and Damha, M.J. (**2002**). Flexible and frozen sugar-modified nucleic acids-modulation of biological activity through furanose ring dynamics in the antisense strand. *Current Topics in Medicinal Chemistry (Hilversum, Netherlands)*, 2, 1147-1171.

138. Meister, G. and Tuschl, T. (**2004**). Mechanisms of gene silencing by double-stranded RNA. *Nature*, 431, 343-349.

139. Fire, A., Xu, S.Q., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (**1998**). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391, 806-811.

140. Hamilton, A.J. and Baulcombe, D.C. (**1999**). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, 286, 950-952.

141. Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (**2001**). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, 411, 494-498.

142. Zamore, P.D., Tuschl, T., Sharp, P.A. and Bartel, D.P. (**2000**). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, 101, 25-33.

143. Elbashir, S.M., Lendeckel, W. and Tuschl, T. (**2001**). RNA interference is mediated by 21-and 22-nucleotide RNAs. *Genes Dev.*, 15, 188-200.

144. Ketting, R.F., Fischer, S.E.J., Bernstein, E., Sijen, T., Hannon, G.J. and Plasterk, R.H.A. (**2001**). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C-elegans*. *Genes Dev.*, 15, 2654-2659.

145. Schwarz, D.S., Hutvagner, G., Haley, B. and Zamore, P.D. (**2002**). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. *Mol. Cell*, 10, 537-548.

146. Mello, C.C. and Conte, D. (**2004**). Revealing the world of RNA interference. *Nature*, 431, 338-342.

147. Song, J.J., Smith, S.K., Hannon, G.J. and Joshua-Tor, L. (**2004**). Crystal structure of argonaute and its implications for RISC slicer activity. *Science*, 305, 1434-1437.

148. Far, R.K.K. and Sczakiel, G. (**2003**). The activity of siRNA in mammalian cells is related to structural target accessibility: a comparison with antisense oligonucleotides. *Nucleic Acids Res.*, 31, 4417-4424.

149. Hough, S.R., Wiederholt, K.A., Burrier, A.C., Woolf, T.M. and Taylor, M.F. (**2003**). Why RNAi makes sense. *Nat. Biotechnol.*, 21, 731-732.

150. Carstea, E.D., Hough, S., Wiederholt, K. and Welch, P.J. (**2005**). State-of-the-art modified RNAi compounds for therapeutics. *Idrugs*, 8, 642-647.

151. Zimmermann, T.S., Lee, A.C.H., Akinc, A., Bramlage, B., Bumcrot, D., Fedoruk, M.N., Harborth, J., Heyes, J.A., Jeffs, L.B., John, M. *et al.* (**2006**). RNAi-mediated gene silencing in non-human primates. *Nature*, 441, 111-114.

152. Soutschek, J., Akinc, A., Bramlage, B., Charisse, K., Constien, R., Donoghue, M., Elbashir, S., Geick, A., Hadwiger, P., Harborth, J. *et al.* (**2004**). Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature*, 432, 173-178.

153. Jayasena, S.D. (**1999**). Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin. Chem.*, 45, 1628-1650.

154. Nimjee, S.M., Rusconi, C.P. and Sullenger, B.A. (**2005**). Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.*, 56, 555-583.

155. Brody, E.N. and Gold, L. (**2000**). Aptamers as therapeutic and diagnostic agents. *Reviews in Molecular Biotechnology*, 74, 5-13.

156. White, R.R., Sullenger, B.A. and Rusconi, C.P. (**2000**). Developing aptamers into therapeutics. *J. Clin. Invest.*, 106, 929-934.

157. Pestourie, C., Tavitian, B. and Duconge, F. (**2005**). Aptamers against extracellular targets

for in vivo applications. *Biochimie*, 87, 921-930.

158. Tuerk, C. and Gold, L. (**1990**). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249, 505-510.

159. Ellington, A.D. and Szostak, J.W. (**1990**). I*n vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346, 818-822.

160. Joyce, G.F. (**1989**). Amplification, mutation and selection of catalytic RNA. *Gene*, 82, 83-87.

161. Nimjee, S.M., Rusconi, C.P. and Sullenger, B.A. (**2006**). Aptamers to proteins. In Klussmann, S. (ed.), *The Aptamer Handbook*. Wiley-Vch Verlag GmbH & Co. KGaA, Weinheim, pp. 131-166.

162. Toulme, J.J., Darfeuille, F., Di Primo, C. and Dausse, E. (**2006**). Aptamers to nucleic acid structures. In Klussmann, S. (ed.), *The aptamer Handbook*. Wiley-Vch Verlag GmbH & Co. KGaA, Weinheim, pp. 167-190.

163. Fickert, H., Fransson, I.G. and Hahn, U. (**2006**). Aptamers to small molecules. In Klussmann, S. (ed.), *The Aptamer Handbook*. Wiley-Vch Verlag GmbH & Co. KGaA, Weinheim, pp. 95-116.

164. Morishita, R., Higaki, J., Tomita, N. and Ogihara, T. (**1998**). Application of transcription factor "decoy" strategy as means of gene therapy and study of gene expression in cardiovascular disease. *Circ. Res.*, 82, 1023-1028.

165. Papavassiliou, A.G. (**1998**). Transcription-factor-modulating agents: precision and selectivity in drug design. *Mol. Med. Today*, 4, 358-366.

166. Ulrich, H., Martins, A.H.B. and Pesquero, J.B. (**2004**). RNA and DNA aptamers in cytomics analysis. *Cytom Part A*, 59A, 220-231.

167. Porschewski, P., Grattinger, M.A.M., Klenzke, K., Erpenbach, A., Blind, M.R. and Schafer, F. (**2006**). Using aptamers as capture reagents in bead-based assay systems for diagnostics and hit identification. *J Biomol Screen*, 11, 773-781.

168. Eye Study Group. (**2002**). Preclinical and phase 1A clinical evaluation of an anti-VEGF pegylated aptamer (EYE001) for the treatment of exudative age-related macular degeneration. *Retina*, 22, 143-152.

169. Fitzwater, T. and Polisky, B. (**1996**). A SELEX primer. *Methods Enzymol.*, 267, 275-301.

170. Opalinska, J.B. and Gewirtz, A.M. (**2002**). Nucleic-acid therapeutics: basic principles and recent applications. *Nat Rev Drug Discov*, 1, 503-514.

171. de Smidt, P.C., Doan, T.L., de Falco, S. and van Berkel, T.J. (**1991**). Association of antisense oligonucleotides with lipoproteins prolongs the plasma half-life and modifies the tissue distribution. *Nucleic Acids Res*, 19, 4695-4700.

172. Griffin, L.C., Tidmarsh. George, F., Bock, L.C., Toole, J.J. and Leung, L.L.K. (**1993**). *In vivo* anticoagulant properties of a novel nucleotide-based thrombin inhibitor and demonstration of regional anticoagulation in extracorporeal circuits. *Blood*, 81, 3271-3276.

173. Crooke, S.T. (**1998**). Vitravene^{TM} - another piece in the mosaic. *Antisense Nucleic Acid Drug Dev.*, 8, Vii-Viii.

174. Ng, E.W.M., Shima, D.T., Calias, P., Cunningham, E.T., Guyer, D.R. and Adamis, A.P. (**2006**). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov*, 5, 123-132.

175. Boomer, R.M., Lewis, S.D., Healy, J.M., Kurz, M., Wilson, C. and McCauley, T.G. (**2005**). Conjugation to polyethylene glycol polymer promotes aptamer biodistribution to healthy and inflamed tissues. *Oligonucleotides*, 15, 183-195.

176. Healy, J.M., Lewis, S.D., Kurz, M., Boomer, R.M., Thompson, K.M., Wilson, C. and McCauley, T.G. (**2004**). Pharmacokinetics and biodistribution of novel aptamer compositions. *Pharm. Res.*, 21, 2234-2246.

177. Freier, S.M. and Altmann, K.-H. (**1997**). The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.*, 25, 4429-4443.

178. Chiu, Y.L. and Rana, T.M. (**2003**). siRNA function in RNAi: a chemical modification analysis. *RNA*, 9, 1034-1048.

179. Manoharan, M. (**1999**). 2'-Carbohydrate modifications in antisense oligonucleotide therapy: importance of conformation, configuration and conjugation. *BBA-Gene Structure and Expression*, 1489, 117-130.

180. Manoharan, M. (**2004**). RNA interference and chemically modified small interfering RNAs. *Curr. Opin. Chem. Biol.*, 8, 570-579.

181. Fattal, E. and Bochot, A. (**2006**). Ocular delivery of nucleic acids: antisense oligonucleotides, aptamers and siRNA. *Advanced Drug Delivery Reviews*, 58, 1203-1223.

182. Egli, M. (**1998**). Conformational preorganization, hydration, and nucleic acid duplex stability. *Antisense Nucleic Acid Drug Dev.*, 8, 123-128.

183. Baddley. (**1973**). Delocalisation into *anti* antibond orbitals. 14, 1645-1648.

184. Koshkin, A.A., Nielsen, P., Meldgaard, M., Rajwanshi, V.K., Singh, S.K. and Wengel, J. (**1998**). LNA (locked nucleic acid): an RNA mimic forming exceedingly stable LNA : LNA duplexes. *J. Am. Chem. Soc.*, 120, 13252-13253.

185. Cram, D.J. and Doxsee, K.M. (**1986**). Host-guest complexation. 41. Preorganization of a host enhances Its binding of aryldiazonium salts. *J. Org. Chem.*, 51, 5068-5071.

186. Trempe, J.-F., Wilds, C.J., Denisov, A.Y., Pon, R.T., Damha, M.J. and Gehring, K. (**2001**). NMR solution structure of an oligonucleotide hairpin with a 2'F-ANA/RNA stem: implications for RNase H specificity toward DNA/RNA hybrid duplexes. *J. Am. Chem. Soc.*, 123, 4896-4903.

187. Wilds, C.J. and Damha, M.J. (**2000**). 2'-Deoxy-2'-fluoro-β-D-arabinonucleosides and oligonucleotides (2'F-ANA): synthesis and physicochemical studies. *Nucleic Acids Res.*, 28, 3625-3635.

188. Stein, C.A. (**1996**). Phosphorothioate antisense oligodeoxynucleotides: Questions of specificity. *Trends Biotechnol.*, 14, 147-149.

189. Agrawal, S. and Kandimalla, E.R. (**2000**). Antisense therapeutics: is it as simple as complementary base recognition? *Mol. Med. Today*, 6, 72-81.

190. Giannaris, P.A. and Damha, M.J. (**1993**). Oligoribonucleotides containing 2',5'-phosphodiester linkages exhibit binding selectivity for 3',5'-RNA over 3',5'-ssDNA. *Nucleic Acids Res.*, 21, 4742-4749.

191. Bhan, P., Bhan, A., Hong, M., Hartwell, J.G., Saunders, J.M. and Hoke, G.D. (**1997**). 2',5'-Linked oligo-3'-deoxyribonucleoside phosphorothioate chimeras: thermal stability and antisense inhibition of gene expression. *Nucleic Acids Res.*, 25, 3310-3317.

192. Kandimalla, E.R., Manning, A., Zhao, Q.Y., Shaw, D.R., Byrn, R.A., Sasisekharan, V. and Agrawal, S. (**1997**). Mixed backbone antisense oligonucleotides: Design, biochemical and biological properties of oligonucleotides containing 2'-5'-ribo- and

3'-5'-deoxyribonucleotide segments. *Nucleic Acids Res.*, 25, 370-378.

193. Damha, M.J. and Noronha, A. (**1998**). Recognition of nucleic acid double helices by homopyrimidine 2',5'-linked RNA. *Nucleic Acids Res.*, 26, 5152-5156.

194. Torrence, P.F., Maitra, R.K., Lesiak, K., Khamnei, S., Zhou, A. and Silverman, R.H. (**1993**). Targeting RNA for degradation with a (2'-5')oligoadenylate-antisense chimera. *Proc. Natl. Acad. Sci. U. S. A.*, 90, 1300-1304.

195. Xiao, W., Li, G.Y., Maitra, R.K., Maran, A., Silverman, R.H. and Torrence, P.F. (**1997**). Correlation of selective modifications to a 2',5'-oligoadenylate-3',5'-deoxyribonucleotide antisense chimera with affinity for the target nucleic acid and with ability to activate RNase L. *J. Med. Chem.*, 40, 1195-1200.

196. Wasner, M., Arion, D., Borkow, G., Noronha, A., Uddin, A.H., Parniak, M.A. and Damha, M.J. (**1998**). Physicochemical and biochemical properties of 2',5'-linked RNA and 2',5'-RNA:3',5'-RNA "hybrid" duplexes. *Biochemistry*, 37, 7478-7486.

197. Damha, M.J., Giannaris, P.A. and Khan, N. (**1991**). 2'-5'-linked oligonucleotides form stable complexes with complementary RNA and DNA. *Nucleic Acids Symp. Ser.*, 290.

198. Kandimalla, E.R., Manning, A., Zhao, Q., Shaw, D.R., Byrn, R.A., Sasisekharan, V. and Agrawal, S. (**1997**). Mixed backbone antisense oligonucleotides: design, biochemical and biological properties of oligonucleotides containing 2'-5'-ribo- and 3'-5'-deoxyribonucleotide segments. *Nucleic Acids Research*, 25, 370-378.

199. Premraj, B.J., Patel, P.K., Kandimalla, E.R., Agrawal, S., Hosur, R.V. and Yathindra, N. (**2001**). NMR structure of a 2',5' RNA favors A type duplex with compact C2 ' *endo* nucleotide repeat. *Biochem. Biophys. Res. Commun.*, 283, 537-543.

200. Premraj, B.J., Raja, S. and Yathindra, N. (**2002**). Structural basis for the unusual properties of 2',5' nucleic acids and their complexes with RNA and DNA. *Biophys. Chem.*, 95, 253-272.

201. Lebruska, L.L. and Maher, L.J., III. (**1999**). Selection and characterization of an RNA decoy for rranscription factor NF-kB. *Biochemistry*, 38, 3168-3174.

202. Heckel, A. and Mayer, G. (**2005**). Light regulation of aptamer activity: an anti-thrombin aptamer with caged thymidine nucleobases. *J. Am. Chem. Soc.*, 127, 822-823.

203. Di Giusto, D.A. and King, G.C. (**2004**). Construction, stability, and activity of multivalent

circular anticoagulant aptamers. *J. Biol. Chem.*, 279, 46483-46489.

204. Buijsman, R.C., Schipperijn, J.W.J., Kuyl-Yeheskiely, E., Van Der Marel, G.A., Van Boeckel, C.A.A. and Van Boom, J.H. (**1997**). Design and synthesis of a possible mimic of a thrombin-binding DNA aptamer. *Bioorg. Med. Chem. Lett.*, 7, 2027-2032.

205. Rosemeyer, H., Mokrosch, V., Jawalekar, A., Becker, E.-M. and Seela, F. (**2004**). Single-stranded DNA: replacement of canonical by base-modified nucleosides in the minihairpin 5'-d(GCGAAGC)-3' and constructs with the aptamer 5'-d(GGTTGGTGTGGTTGG)-3'. *Helv. Chim. Acta*, 87, 536-553.

206. Sacca, B., Lacroix, L. and Mergny, J.-L. (**2005**). The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. *Nucleic Acids Res.*, 33, 1182-1192.

207. Nolte, A., Klussmann, S., Bald, R., Erdmann, V.A. and Fuerste, J.P. (**1996**). Mirror-design of L-oligonucleotide ligands binding to L-arginine. *Nat. Biotechnol.*, 14, 1116-1119.

208. Kusser, W. (**2000**). Chemically modified nucleic acid aptamers for *in vitro* selections: evolving evolution. *Reviews in Molecular Biotechnology*, 74, 27-38.

209. Lato, S.M., Ozerova, N.D.S., He, K., Sergueeva, Z., Shaw, B.R. and Burke, D.H. (**2002**). Boron-containing aptamers to ATP. *Nucleic Acids Res.*, 30, 1401-1407.

210. Jhaveri, S., Olwin, B. and Ellington, A.D. (**1998**). *In vitro* selection of phosphorothiolated aptamers. *Bioorg. Med. Chem. Lett.*, 8, 2285-2290.

211. Kato, Y., Minakawa, N., Komatsu, Y., Kamiya, H., Ogawa, N., Harashima, H. and Matsuda, A. (**2005**). New NTP analogs: the synthesis of 4'-thioUTP and 4'-thioCTP and their utility for SELEX. *Nucleic Acids Res.*, 33, 2942-2951.

212. Kempeneers, V., Renders, M., Froeyen, M. and Herdewijn, P. (**2005**). Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. *Nucleic Acids Res.*, 33, 3828-3836.

213. Horhota, A., Zou, K., Ichida, J.K., Yu, B., McLaughlin, L.W., Szostak, J.W. and Chaput, J.C. (**2005**). Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.*, 127, 7427-7434.

214. Chaput, J.C., Ichida, J.K. and Szostak, J.W. (**2003**). DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc.*, 125, 856-857.

215. Chaput, J.C. and Szostak, J.W. (**2003**). TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.*, 125, 9274-9275.

216. Sinha, S., Kim, P.H. and Switzer, C. (**2004**). 2',5'-Linked DNA is a template for polymerase-directed DNA synthesis. *J. Am. Chem. Soc.*, 126, 40-41.

217. Fa, M., Radeghieri, A., Henry, A.A. and Romesberg, F.E. (**2004**). Expanding the substrate repertoire of a DNA polymerase by directed evolution. *J. Am. Chem. Soc.*, 126, 1748-1754.

218. Richardson, F.C., Kuchta, R.D., Mazurkiewicz, A. and Richardson, K.A. (**2000**). Polymerization of 2'-fluoro- and 2'-O-methyl-dNTPs by human DNA polymerase α, polymerase γ, and primase. *Biochem. Pharmacol.*, 59, 1045-1052.

219. Dineva, M.A., Ivanov, I.G. and Petkov, D.D. (**1997**). P-alpha-methyl deoxynucleoside triphosphates as substrates for *E. coli* DNA polymerase I in a template-directed synthesis of DNA. *Nucleosides Nucleotides*, 16, 1875-1882.

220. Ono, T., Scalf, M. and Smith, L.M. (**1997**). 2'-Fluoro modified nucleic acids: polymerase-directed synthesis, properties and stability to analysis by matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Res.*, 25, 4581-4588.

221. Aoyama, H., Cottin, L.S., Tarragolitvak, L., Litvak, S. and Guschlbauer, W. (**1985**). 2'-Fluoro-2'-deoxycytidine triphosphate as a substrate for RNA-dependent and DNA-dependent DNA polymerases. *Biochim. Biophys. Acta*, 824, 218-224.

222. Ruth, J.L. and Cheng, Y.C. (**1981**). Nucleoside analogs with clinical potential in antivirus chemotherapy - the effect of several thymidine and 2'-deoxycytidine analog 5'-triphosphates on purified human (alpha, beta) and herpes-simplex virus (type-1, type-2) DNA-polymerases. *Mol. Pharmacol.*, 20, 415-422.

223. Helfman, W.B., Hendler, S.S., Shannahoff, D.H. and Smith, D.W. (**1978**). *Escherichia-coli* DNA polymerase-II and Polymerase-III - substrate-specificity. *Biochemistry*, 17, 1607-1611.

224. Carrasco, N., Caton-Williams, J., Brandt, G., Wang, S. and Huang, Z. (**2006**). Efficient enzymatic synthesis of phosphoroselenoate RNA by using adenosine 5'(α–P-seleno)triphosphate. *Angewandte Chemie Internal Edition,* 45, 94-97.

225. Chelliserrykattil, J. and Ellington, A.D. (**2004**). Evolution of a T7 RNA polymerase variant that transcribes 2'-O-methyl RNA. *Nat. Biotechnol.*, 22, 1155-1160.

226. Pavey, J.B.J., Lawrence, A.J., O'Neil, I.A., Vortler, S. and Cosstick, R. (**2004**). Synthesis and transcription studies on 5'-triphosphates derived from 2'-C-branched-uridines: 2'-homo-uridine-5'-triphosphate is a substrate for T7 RNA polymerase. *Organic & Biomolecular Chemistry,* 2, 869-875.

227. Padilla, R. and Sousa, R. (**2002**). A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs. *Nucleic Acids Res.,* 30, e138.

228. Raines, C. and Gottleib, P.A. (**1998**). Enzymatic incorporation of 2'-thio-CTP into the HDV ribozyme. *RNA,* 4, 340-345.

229. Aurup, H., Williams, D.M. and Eckstein, F. (**1992**). 2'-Fluoro- and 2'-amino-2'-deoxynucleoside 5'-triphosphates as substrates for T7 RNA polymerase. *Biochemistry,* 31, 9636-9641.

230. Kuwahara, M., Hanawa, K., Ohsawa, K., Kitagata, R., Ozaki, H. and Sawai, H. (**2006**). Direct PCR amplification of various modified DNAs having amino acids: Convenient preparation of DNA libraries with high-potential activities for in vitro selection. *Bioorg. Med. Chem.,* 14, 2518-2526.

231. Perrin, D.M., Garestier, T. and Helene, C. (**2001**). Bridging the gap between proteins and nucleic acids: a metal-independent RNaseA mimic with two protein-like functionalities. *J. Am. Chem. Soc.,* 123, 1556-1563.

232. Teramoto, N., Imanishi, Y. and Ito, Y. (**2000**). *In vitro s*election of a ligase ribozyme carrying alkylamino groups in the side chains. *Bioconjug. Chem.,* 11, 744-748.

233. Santoro, S.W., Joyce, G.F., Sakthivel, K., Gramatikova, S. and Barbas, C.F. (**2000**). RNA cleavage by a DNA enzyme with extended chemical functionality. *J. Am. Chem. Soc.,* 122, 2433-2439.

234. Battersby, T.R., Ang, D.N., Burgstaller, P., Jurczyk, S.C., Bowser, M.T., Buchanan, D.D., Kennedy, R.T. and Benner, S.A. (**1999**). Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog. *J. Am. Chem. Soc.,* 121, 9781-9789.

235. Wiegand, T.W., Janssen, R.C. and Eaton, B.E. (**1997**). Selection of RNA amide syntheses. *Chem. Biol.,* 4, 675-683.

236. Tarasow, T.M., Tarasow, S.L. and Eaton, B.E. (**1997**). RNA-catalysed carbon-carbon bond

formation. *Nature,* 389, 54-57.

237. Latham, J.A., Johnson, R. and Toole, J.J. (**1994**). The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing 5-(1-pentynyl)-2'-deoxyuridine. *Nucleic Acids Research,* 22, 2817-2822.

238. King, D.J., Ventura, D.A., Brasier, A.R. and Gorenstein, D.G. (**1998**). Novel combinatorial selection of phosphorothioate oligonucleotide aptamers. *Biochemistry,* 37, 16489-16493.

239. Kawazoe, N., Teramoto, N., Ichinari, H., Imanishi, Y. and Ito, Y. (**2001**). *In vitro* selection of nonnatural ribozyme-catalyzing porphyrin metalation. *Biomacromolecules,* 2, 681-686.

240. Beaudry, A., DeFoe, J., Zinnen, S., Burgin, A. and Beigelman, L. (**2000**). *In vitro* selection of a novel nuclease-resistant RNA phosphodiesterase. *Chem. Biol.,* 7, 323-334.

241. Pagratis, N.C., Bell, C., Chang, Y.-F., Jennings, S., Fitzwater, T., Jellinek, D. and Dang, C. (**1997**). Potent 2'-amino-, and 2'-fluoro-2'-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor. *Nat. Biotechnol.,* 15, 68-73.

242. Kubik, M.F., Bell, C., Fitzwater, T., Watson, S.R. and Tasset, D.M. (**1997**). Isolation and characterization of 2'-fluoro-, 2'-amino-, and 2'-fluoro-/amino-modified RNA ligands to human IFN-gamma that inhibit receptor binding. *J. Immunol.,* 159, 259-267.

243. Green, L.S., Jellinek, D., Bell, C., Beebe, L.A., Feistner, B.D., Gill, S.C., Jucker, F.M. and Janjic, N. (**1995**). Nuclease-resistant nucleic-acid ligands to vascular-permeability factor vascular endothelial growth-factor. *Chem. Biol.,* 2, 683-695.

244. Jellinek, D., Green, L.S., Bell, C., Lynott, C.K., Gill, N., Vargeese, C., Kirschenheuter, G., Mcgee, D.P.C., Abesinghe, P., Pieken, W.A. *et al.* (**1995**). Potent 2'-amino-2'-deoxypyrimidine RNA inhibitors of basic fibroblast growth-factor. *Biochemistry,* 34, 11363-11372.

245. Lin, Y., Qiu, Q., Gill, S.C. and Jayasena, S.D. (**1994**). Modified RNA sequence pools for *in-vitro* selection. *Nucleic Acids Res.,* 22, 5229-5234.

246. Kempeneers, V., Vastmans, K., Rozenski, J. and Herdewijn, P. (**2003**). Recognition of threosyl nucleotides by DNA and RNA polymerases. *Nucleic Acids Res.,* 31, 6221-6226.

247. Ichida, J.K., Horhota, A., Zou, K., McLaughlin, L.W. and Szostak, J.W. (**2005**). High fidelity TNA synthesis by Therminator polymerase. *Nucleic Acids Res.,* 33, 5219-5225.

Chapter II

Dr. Damha's group has a long-standing interest in 2',5'-linked oligonucleotides which serves as not only a interesting structure to understand why nature chooses 3',5'-linked RNA/DNA interested of 2',5'-linked counterparts, but also a novel chemical modification with potential use in gene silencing (antisense and siRNAs). An interesting finding about 2',5'-linked RNA duplex is its smaller interstrand phosphate-phosphate distances (by *ca.* 1 Å). This may account, at least in part, for the lower thermal stability of RNA:2',5'-RNA relative to RNA:RNA and RNA:DNA helices. The current chemical modification aims to prepare extended 2',5'-RNA by adding extra CH_2 group on 2' position of sugar ring. This work demonstrates the first studies of 2',5'-RNA modification and its effect of incorporation on the stabilization of duplex.

Reprint from Peng, C.G. and Damha, M.J. (2005) Synthesis and hybridization studies of oligonucleotides containing 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)-thymine (2'-α-*hm*-dT). *Nucleic Acids Research*, 33(22), 7019-7028. With permission from Oxford University Press

Chapter II. Synthesis and Hybridization Studies of Oligonucleotides Containing 1-(2-Deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (2'-α-*hm*-dT)

Abstract

We report the first investigation of oligoribonucleotides containing a few 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine units (or 2'-hm-dT, abbreviated in this work as "**H**"). Both the 2'-*CH_2O*-phosphoramidite and 3'-*O*-phosphoramidite derivatives of **H** were synthesized and incorporated into both 2',5'-RNA and RNA chains. The hybridization properties of the modified oligonucleotides have been studied *via* thermal denaturation and circular dichroism studies. While 3',5'-linked **H** was previously shown to significantly destabilize *DNA*:RNA hybrids and *DNA*:DNA duplexes (modification in the *DNA* strand; ΔT_m ~ -3 °C / insert), we find that 2',5'-linked **H** have a smaller effect on *2',5'-RNA*:RNA and *RNA*:RNA duplexes (ΔT_m = -0.3 °C and -1.2 °C, respectively). The incorporation of 3',5'-linked **H** into *2',5'-RNA*:RNA and *RNA*:RNA duplexes was found to be more destabilizing (-0.7 °C and -3.6 °C, respectively). Significantly, however, the 2',5'-linked **H** units confer marked stability to RNA hairpins when they are incorporated into a 2',5'-linked tetraloop structure (ΔT_m = + 1.5 °C/insert). These results are rationalized in terms of the compact and extended conformations of nucleotides.

Introduction

Our group has had a long-standing research interest in the physicochemical and biochemical properties of 2',5'-linked ribonucleic acids (2',5'-RNA, Figure 2.1).[1,2] These regioisomers of standard (i.e. 3',5'-linked) RNA are not only interesting from a structural point of view, but also have potential use in the down-regulation of gene expression.[3,4] For instance, 2',5'-RNAs are able to associate with complementary single-stranded RNA[3] as well as duplex DNA[5] and, as such, can potentially be used to down-regulate gene expression *via* the antisense and antigene approaches (Figure 2.1).[4,6-8] It is also well-documented that annealing two normal RNA strands is more favorable than annealing a 2′,5′-RNA strand with a normal RNA strand.[1,3,8,9] Furthermore, mutually complementary 2′,5′-RNA strands have the ability to associate, but exhibit a transition

temperature (T_m) that is considerably lower than those of the corresponding RNA:RNA or RNA:2′,5′-RNA duplexes. A comparison of the T_m values of various duplexes of mixed base composition revealed the following order of duplex thermal stability: RNA:RNA > DNA:DNA ≈ DNA:RNA > RNA:2′,5′-RNA > 2′,5′-RNA: 2′,5′-RNA > DNA: 2',5'-RNA (undetected).[1]

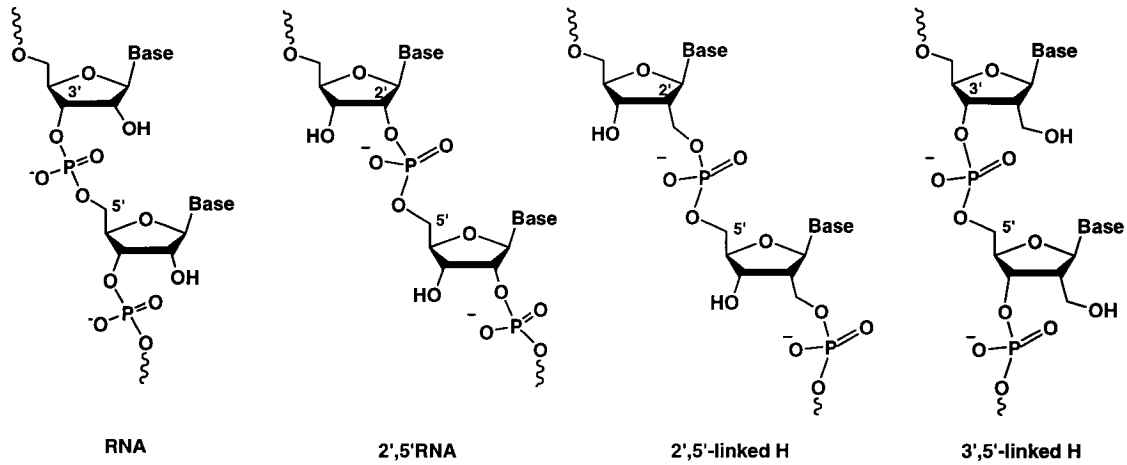
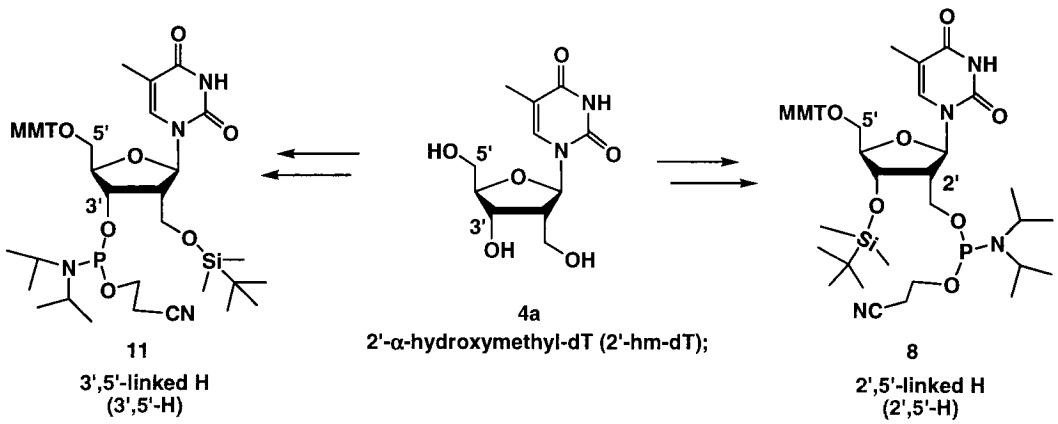


Figure 2.1: Structure of RNA, 2',5'RNA, and strands comprising 2',5'-linked **H** and 3',5'-linked **H** units.



Scheme 2.1: Structures of 2'-α-*hm*-dT ("**H**") and its phosphoramidite derivatives **8** and **11**.

Molecular modeling and circular dichroism (CD) studies of these duplexes revealed that RNA: 2',5'-RNA hybrids adopt a continuous A-type helix structure similar to that of native RNA,[1] but have smaller interstrand phosphate-phosphate distances (by *ca.* 1 Å).

This may account, at least in part, for the lower thermal stability of RNA:2',5'-RNA relative to RNA:RNA and RNA:DNA helices.[1]

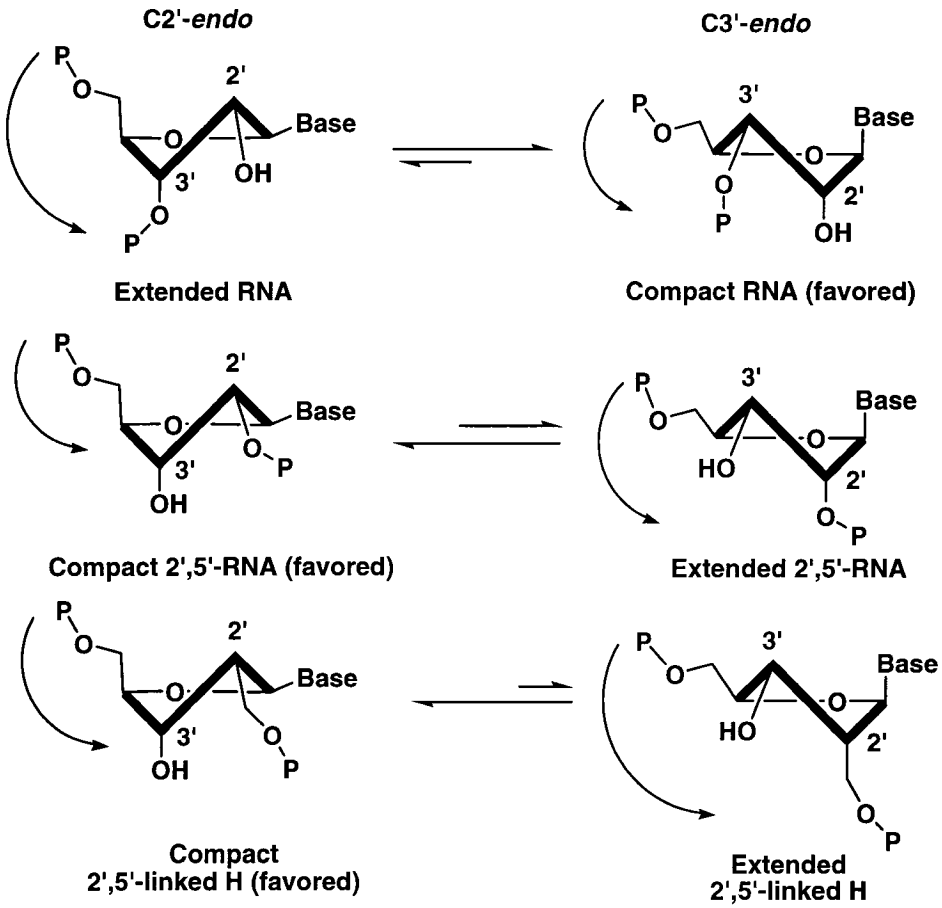


Figure 2.2: The C2'- and C3'-*endo* sugar puckers are favored in 2',5'-linked RNA and 3',5'-linked RNA, respectively.[10,11] The intra-residue P-P distance determines the "compact" or "extended" backbone structure.[11] Lengthening by an extra carbon C2' (**H**) removes the gauche effect between the ring oxygen and O2', which is expected to reinforce the C2'-*endo* or compact conformation for 2',5'-linked **H** units.

As part of our ongoing study of these systems, we now present an investigation of 2',5'-RNA chains containing one or more 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (2'-α-*hm*-dT; abbreviated in this work as "**H**", Figure 2.1 and Scheme 2.1). We anticipated that lengthening the sugar-phosphate backbone by one methylene unit would not only diminish putative P-P repulsions in 2',5'-RNA:RNA duplexes, but would also provide a "compact" nucleotide conformation favoring tighter RNA binding (Figure 2.2).[10,11] This interesting modification was first studied by Schmit

and co-workers[12] and reviewed more recently by Freier and Altmann.[13] Schmit *et al.*[12] reported that 3',5'-linked **H** units (3',5'-**H**) significantly destabilize *DNA*:RNA hybrids and *DNA*:DNA duplexes (modification in the *DNA* strand; ΔT_m = -2.9 °C). Destabilization due to the 2'-α-hydroxymethyl and other 2'-α alkyl groups was explained by the tendency of these substituents to shift the conformational equilibrium of the sugar toward the C2'-*endo* pucker and away from the C3'-*endo* pucker found in RNA, 2'-O-alkyl RNA and 2'F-RNA duplexes.[12,13] To the best of our knowledge, neither 2',5'-linked **H** units (2',5'-**H**) nor incorporation into RNA have ever been examined. The present study revealed that 2',5'-**H** units have a modest destabilizing effect on *2',5'-RNA*:RNA (ΔT_m = -0.3 °C) and *RNA*:RNA duplexes (ΔT_m = -1.2 °C), whereas 3',5'-**H** units are significantly more destabilizing (-0.7 °C and -3.6 °C, respectively), consistent with the results of Schmit *et al.*.[12] In contrast, we find that 2',5'-**H** units confer significant stability to RNA hairpins, particularly when they are placed in the loop structure (ΔT_m = + 1.5 °C).

Materials and Methods

General reagents

All reactions were carried out in oven-dried glassware under a N_2 atmosphere. Dichloromethane (CH_2Cl_2) and acetonitrile (CH_3CN) were dried by refluxing and distilling over calcium hydride (CaH_2) under a N_2 atmosphere. Tetrahydrofuran (THF) was dried by refluxing over sodium and benzophenone under a N_2 atmosphere and collected before use. Anhydrous methanol (MeOH), pyridine (py), *N*-ethyl-*N,N*-diisopropylamine (DIPEA), *N,N*-dimethylformamide (DMF) and 2,6-lutidine were obtained from Aldrich. The following chemicals were used as received from Aldrich: *p*-anisylchlorodiphenyl (MMTr-Cl), benzoyl chloride (Bz-Cl), *tert*-butyldimethylsilyl triflate (TBDMSOTf), *tert*-butyldimethylsilyl chloride (TBDMSCl), 4,5-dicyanoimidazole (DCI), 4-dimethylaminopyridine (DMAP), 1,1,1,3,3,3-hexamethyldisilazane (HMDS), 10 *wt%* palladium on carbon powder (10% Pd/C), triethylamine tris(hydrofluoride) (TREAT HF), *D*-ribose, silver nitrate ($AgNO_3$), 1.0 M tin (IV) tetrachloride solution in dichloromethane ($SnCl_4$), thymine. β-cyanoethyl-N,N-

diisopropylchlorophosphoramidite and other solid-phase synthesis reagents and regular nucleoside monomers were purchased from ChemGenes Corp. (Ashland, MA).

Methyl 2-α-benzoyloxymethyl-2-deoxy-3,5-di-*O*-(2,4-dichlorobenzyl)-α-D-ribo-furanoside (2). Comp (**1**) was obtained following slight modifications of the procedure described by Martin [14]. Under a N_2 atmosphere, 1.5 mL (1.78 g, 12.7 mmol) BzCl was added to a solution of **1** (4.83 g, 9.73 mmol) in 75 mL dry pyridine. The resulting solution was kept stirring at r.t. for 4 h. The reaction progress was monitored by TLC. When the reaction was complete, the reaction mixture was concentrated under reduced pressure, and then washed with saturated $NaHCO_3$ and brine. The aqueous phase was extracted with CH_2Cl_2 three times. The combined organic layers were evaporated to dryness to afford **2** in quantitative yield (5.84 g). R_f (SiO_2) = 0.86 (2:1 EtOAc/hexane); 0.70 (1:2 EtOAc/hexane); ^{1}H NMR (400 MHz, $CDCl_3$): δ7.98-7.09 (m, 11H, ArH), 5.11 (d, $^3J_{1\text{-}2}$ = 5.2 Hz, 1H, H1), 4.73-4.55 (m, 6H, benzyl CH_2 and CH_2-C2), 4.40-4.43 (m, 1H, H4), 4.08 (dd, $^3J_{3\text{-}2}$, $^3J_{3\text{-}4}$ = 5.2, 2.4 Hz, 1H, H3), 3.54-3.64 (AB dd, $^2J_{5,5'}$ = 22.4 Hz, $^3J_{4,5+5'}$ = 4.4, 5.2 Hz, 2H, H5 & H5'), 3.45 (s, 3H, OCH_3), 2.74-2.67 (m, 1H, H2). ESI-MS for $C_{28}H_{26}Cl_4O_6$ [MNa^+] calcd. 621.05, found 621.1.

1-[2-α-Benzoyloxymethyl-2-deoxy-3,5-di-*O*-(2,4-dichlorobenzyl)-D-ribo-furanosyl]thymine (3). Under a N_2 atmosphere, a mixture of dry thymine (1.35 g, 10.7 mmol), ammonium sulfate (0.136 g, 1.029 mmol), and HMDS (2.87 mL, 13.8 mmol) in anhydrous CH_3CN (41 mL) was refluxed at 100 °C for 4 h until the reaction mixture was clear. The reaction mixture was cooled to r.t. and added to a solution of **2** (2.47 g, 4.12 mmol) in dry CH_3CN (19.5 mL) followed by 1.0 M $SnCl_4$ in CH_2Cl_2 (4.12 mL, 4.12 mmol). The resulting solution was warmed to 50 °C and left stirring for an additional 12 h. After the reaction was complete, the reaction mixture was cooled to r.t., washed with saturated $NaHCO_3$ (50 mL) and filtered. The filtrate was extracted three times with the same volume of CH_2Cl_2 and the organic layer was washed with brine, dried over Na_2SO_4 and evaporated to yield a white foam. Purification by silica gel column chromatography with EtOAc/hexane (2:1 v/v) afforded the title compound in 72% yield (2.06 g, β:α = 5:1). R_f (SiO_2) = 0.6 (2:1 EtOAc/hexane); ^{1}H NMR of β-isomer (400 MHz, $CDCl_3$): δ7.97-7.19 (m, 12H, ArH and H-6), 6.35 (d, $^3J_{1',2'}$ = 8.8 Hz, 1H, H1'), 4.74, 4.47 (2dd, $^2J_{CH2\text{-}C2'}$ = 12 Hz, $^3J_{CH2\text{-}C2',\ H2'}$=7.6, 7.2 Hz, 2H, CH_2-C2'), 4.69-4.56 (m, 4H, benzyl CH_2), 4.38 (br,

m, 1H, H4'), 4.31-4.28 (m, 1H, H3'), 3.90, 3.70 (2 dd, $^2J_{5',5''}$ = 10.8 Hz, $^3J_{4',5'+5''}$ = 3.2, 2.4 Hz, 2H, H5' & H5''), 2.99-2.92 (m, 1H, H2'), 1.64 (d, $^4J_{CH3\text{-}C5,H6}$ = 1.2 Hz, 3H, CH_3-C5). ESI-MS for $C_{32}H_{28}Cl_4N_2O_7$ [MNa^+] calcd. 717.39, found 717.0.

1-(2-α-*C*-Benzoyloxymethyl-2-deoxy-β-D-ribofuranosyl)thymine (4). To a solution of **3** (1.065 g, 1.53 mmol) in dry MeOH (50 mL) was added 10% Pd/C (0.39 g, 0.37 mmol). The resulting mixture was kept shaking under 40-50 psi H_2 at r.t. for 6 h. The mixture was filtered and evaporated to dryness, and the residue was purified by silica gel column chromatography (eluent 40:1 to 15:1 CH_2Cl_2/MeOH) to separate the α and β isomers. The desired *β* isomer was obtained in 71% yield (0.41 g). R_f (SiO_2) = 0.51 (6:1 CH_2Cl_2/MeOH); 0.31 (10:1 CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.22 (s, 1H, N-H), 7.78-7.41 (m, 5H, Bz), 7.70 (s, 1H, H6), 6.14 (d, $^3J_{1',2'}$ = 8.8 Hz, 1H, H1'), 5.52 (d, $^3J_{OH,\ H3'}$ = 4.8 Hz, 1H, HO-C3'), 5.10 (t, $^3J_{OH,\ H5'+H5''}$ = 5.2 Hz, 1H, HO-C5'), 4.57, 4.26 (2dd, $^2J_{CH2\text{-}C2}$ = 11.6 Hz, $^3J_{CH2\text{-}C2,\ H2'}$ = 5.6, 8.8 Hz, 2H, CH_2-C2'), 3.88 (br, 1H, H4'), 3.54-3.64 (AB, dd, $^2J_{5',5''}$ = 17.7, $^3J_{4',5'+5''}$ = 5.2, 4.8 Hz, 2H, H5' & H5''), 2.72-2.29 (m, 1H, H2'), 1.74 (s, 3H, CH_3-C5). Four NOESY cross peaks identified the β isomer: H1'-H4', H2'-H6', 2'-CH_2-H4', H5'-H6. ESI-MS for $C_{18}H_{20}N_2O_7$ [MNa^+] calcd. 399.13, found 399.1. Isolated α isomer, ^{1}H NMR (500 MHz, acetone-d_6): δ 10.11 (s, 1H, N-H), 8.02-7.46 (m, 5H, Bz), 7.90 (s, 1H, H-6), 6.55 (d, $^3J_{1',2'}$ = 7.5 Hz, 1H, H1'), 5.13 (d, $^3J_{OH,\ H3'}$ = 3.5 Hz, 1H, HO-C3'), 4.66 (m, 1H, H3'), 4.52-4.30 (m, 2H, CH_2-C2'), 4.48 (m, 1H, H4'), 4.19 (br, 1H, HO-C5'), 3.66 (br, 2H, H5' and H5''), 3.33 (m, 1H, H2'), 1.82 (s, 3H, CH_3-C5). ESI-MS for α $C_{18}H_{20}N_2O_7$ [MNa^+] calcd. 399.13, found 399.1.

1-[2-α-*C*-Benzoyloxymethyl-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]-thymine (5). Compound **4** (0.555 g, 1.47 mmol) and $AgNO_3$ (0.375 g, 2.21 mmol) were dissolved in dry pyridine (10 mL). The flask was purged with N_2 and a solution of MMTrCl (0.683 g, 2.21 mmol) in dry pyridine (5 mL) was added. After stirring for 6h at r.t., the reaction was worked up by evaporating most of the pyridine and taking up the residue in CH_2Cl_2 (20 mL). The solution was washed with saturated $NaHCO_3$ (20 mL), dried, filtered, and finally evaporated to yield the crude product. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (49:1 to 15:1 v/v) afforded the title compound as a white solid (687 mg; 72% yield). R_f (SiO_2) = 0.66 (10:1

CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, pyridine-d_5): δ 12.12 (s, 1H, N-H), 6.95-5.75 (m, 19H, ArH), 6.73 (d, $^4J_{CH3\text{-}C5,\ H6}$ = 1.2 Hz, 1H, H6), 5.94 (d, $^3J_{1',2'}$ = 12 Hz, 1H, H1'), 4.12-3.83 (m, 3H, CH_2-C2' and H3'), 3.43 (br, m, 1H, H4'), 2.48 (s, 3H, OCH_3), 2.45 (m, 2H, H5' & H5''), 2.30 (m, 1H, H2'), 0.50 (s, 3H, CH_3-C5). ^{1}H NMR (400MHz, DMSO-d_6): δ 11.26 (s, 1H, N-H), 7.81-6.88 (m, 19H, ArH), 7.44 (s, 1H, H6), 6.20 (d, $^3J_{1',2'}$ = 8.8 Hz, 1H, H1'), 5.62 (d, $^3J_{3',OH}$ = 5.2 Hz, 1H, OH), 4.70, 4.31 (2dd, $^2J_{CH2\text{-}C2'}$ = 11.2 Hz, $^3J_{CH2\text{-}C2',\ H2'}$ = 5.2, 8.4 Hz, 2H, CH_2-C2'), 4.43 (m, 1H, H3'), 4.0 (br m, 1H, H4'), 3.73 (s, 3H, OCH_3), 3.27, 3.16 (2dd, $^2J_{5',5''}$ = 10.4 Hz, $^3J_{4',5'+5''}$ = 6.4, 3.2 Hz, 2H, H5' & H5''), 3.01-2.94 (m, 1H, H2'), 1.30 (s, 3H, CH_3-C5). ESI-MS for $C_{38}H_{36}N_2O_8$ [MNa^+] calcd. 671.25, found 671.2.

1-[2-*α*-*C*-Benzoyloxymethyl-3-*O*-*tert*-butyldimethylsilyl-2-deoxy-5-*O*-(4-methoxytrityl)-*β*-D-ribofuranosyl]thymine (6). To a solution of **5** (205 mg, 0.316 mmol) in dry CH_2Cl_2 was added 2,6-lutidine (0.15 mL, 1.26 mmol) under a N_2 atmosphere. *tert*-Butyldimethysilyl triflate (TBDMOTf; 0.218 mL, 0.948 mmol) was then added dropwise with stirring at r.t. to yield a clear yellow solution. After stirring at 40°C for 12 h, saturated $NaHCO_3$ (12 mL) was added. The aqueous phase was separated and washed several times with CH_2Cl_2. The combined CH_2Cl_2 layers were dried over Na_2SO_4 and evaporated to dryness. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2:MeOH (100:1 to 20:1 v:v) afforded **6** in 72% yield (0.173 g). When the silylation reaction was conducted at r.t. for 4 h, the yield of the title compound increased to 83%. R_f (SiO_2) = 0.64 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (500 MHz, DMSO-d_6): δ 11.33 (s, 1H, N-H), 7.83-6.88 (19H, ArH), 7.50 (s, 1H, H6), 6.12 (d, $^3J_{1',2'}$ = 8.5 Hz, 1H, H1'), 4.57, 4.30 (2m, 2H, CH_2-C2'), 4.55 (m, 1H, H3'), 3.97 (m, 1H, H4'), 3.71 (s, 3H, OCH_3), 3.24 (m, 2H, H5' & H5''), 3.00 (m, 1H, H2'), 1.42 (s, 3H, CH_3-C5), 0.80 (s, 9H, *t*-Bu-Si), 0.02, -0.06 (2s, 6H, $(CH_3)_2$-Si). ^{1}H NMR (400 MHz, acetone-d_6): δ 9.94 (s, 1H, N-H), 7.96-6.91 (17H, ArH), 7.61 (s, 1H, H6), 6.35 (d, $^3J_{1',2'}$ = 8.8, 1H, H1'), 4.78-4.42 (m, 3H, H3' and CH_2-C2'), 4.15 (br, m, 1H, H4'), 3.79 (s, 3H, OCH_3), 3.44 (m, 2H, H5' & H5''), 3.18 (m, 1H, H2'), 1.44 (s, 3H, CH_3-C5), 0.92 (s, 9H, *t*-Bu-Si), 0.14, -0.08 (2s, 6H, $(CH_3)_2$-Si). ESI-MS for $C_{44}H_{50}N_2O_8Si$ [MNa^+] calcd. 785.33, found 758.3.

1-[3-*O*-*tert*-Butyldimethylsilyl-2-deoxy-2-α-*C*-hydroxymethyl-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (7). A saturated solution of NaOMe in MeOH (4 mL) was added to **6** (0.337 g, 0.442 mmol) at r.t. After stirring for 2-3 h, the reaction mixture was washed with saturated $NaHCO_3$ (10 mL). The aqueous phase was separated and washed several times with CH_2Cl_2. The combined CH_2Cl_2 layers were dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (100:1 to 15:1 v/v) afforded **7** in quantitative yield (290 mg). R_f (SiO_2) = 0.43 (15:1 CH_2Cl_2/MeOH); ^1H NMR (400 MHz, DMSO-d_6): δ 11.30 (s, 1H, N-H), 7.47 (s, 1H, H6), 7.40-6.89 (m, 14H, ArH), 5.93 (d, $^3J_{1',2'}$ = 7.6 Hz, 1H, H1'), 4.57 (d, $^3J_{OH,\ H3'}$ = 4.8 Hz, 1H, OH); 4.45 (br m, 1H, H3'), 3.90 (br, 1H, H4'), 3.73 (s, 3H, OCH_3), 3.68, 3.45 (2m, 2H, CH_2-C2'), 3.19 (m, 2H, H5' and H5''), 2.53 (m, 1H, H2'), 1.51 (s, 3H, CH_3-C5), 0.82 (s, 9H, *t*-Bu-Si), 0.05, -0.04 (2s, 6H, $(CH_3)_2$-Si). ESI-MS for $C_{37}H_{46}N_2O_7Si$ [MNa^+] calcd. 681.31, found 681.2.

1-[2-α-*C*-((β-Cyanoethyl-*N*,*N*-diisopropylphosphoramidic)-hydroxymethyl)-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (8). *N*-ethyl-*N*,*N*-diisopropylamine (0.276 mL, 0.205 g, 1.58 mmol) was added to a solution of **7** (0.29 g, 0.44 mmol) in dry THF (3 mL) under a N_2 atmosphere. The reaction was initiated by addition of β-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite *via* syringe (0.125 g, 0.118 mL, 0.528 mmol). After 2 h, the reaction mixture was passed through a 2-cm layer of silica gel (pre-neutralized with 1% NEt_3), and the desired compound eluted by washing with ice cold hexane followed by CH_2Cl_2. Evaporation of the solution afforded the title compound as a nice white foam (353 mg, 97% yield). R_f (SiO_2) = 0.63 (2:1 EtOAc/hexane); ^{31}P NMR (200 MHz, $CDCl_3$): δ 148.90, 148.24. ESI-MS for $C_{46}H_{63}N_4O_8PSi$ [MNa^+] calcd. 881.42, found 881.2.

1-[2-Deoxy-2-α-*C*-hydroxymethyl-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]-thymine (9). A saturated solution of NaOMe in MeOH (2 mL) was added to **5** (0.208 g, 0.321 mmol) and the resulting mixture stirred at r.t. for 2 h. The reaction mixture was then washed with brine (5 mL) followed by CH_2Cl_2 (20 mL). The combined CH_2Cl_2 layers were dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel flash column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (15:1 v/v) afforded **9**

in quantitative yield (175 mg). R_f (SiO_2) = 0.26 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, acetone-d_6): δ7.6 (d, $^4J_{CH3\text{-}C5,H6}$ = 1.2 Hz, 1H, H6), 7.52-6.91 (m, 14H, ArH), 6.21 (d, $^3J_{1',2'}$ = 8.4 Hz, 1H, H1'), 4.65 (br, m, 1H, H3'), 4.10-4.12 (m, 1H, H4'), 3.98, 3.79 (2dd, $^2J_{CH2\text{-}C2'}$ = 11.2 Hz, $^3J_{CH2\text{-}C2',\ H2'}$ = 10.0, 6.8 Hz, 2H, CH_2-C2'), 3.42-3.35 (AB, dd, $^2J_{5',5''}$ = 9.6 Hz, $^3J_{4',5'+5''}$ = 4.0, 3.2 Hz, 2H, H5' & H5''), 2.64-2.71 (m, 1H, H2'), 1.49 (s, 3H, CH_3-C5). ^{1}H NMR (500 MHz, DMSO-d_6): δ11.30 (s, 1H, N-H); 7.47 (s, 1H, H6); 7.39-6.89 (m, 14H, ArH); 5.97 (d, $^3J_{1',2'}$ = 8.5 Hz, 1H, H1'), 5.30 (d, $^3J_{OH,\ H3'}$ = 5.5 Hz, 1H, 3'-OH); 4.54 (bs, 1H, C2'-C-OH); 4.30 (bs, 1H, H3'); 3.91 (m, 1H, H4'); 3.73 (s, 3H, OCH_3); 3.74-3.43 (2m, 2H, CH_2-C2'); 3.20, 3.12 (2dd, $^2J_{5',5''}$ = 10.5 Hz, $^3J_{4',5'+5''}$ = 4.5, 3.0 Hz, 2H, H5' & H5''), 2.45 (m, 1H, H2'), 1.40 (s, 3H, CH_3-C5); ESI-MS for $C_{31}H_{32}N_2O_7$ [MNa^+] calcd. 567.22, found 567.2.

1-[2-α-*C*-*tert*-Butyldimethylsilyloxymethyl-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (10). To a solution of **9** (0.186 g, 0.342 mmol) in dry pyridine (3 mL) under a N_2 atmosphere was added $AgNO_3$ (0.067 g, 0.393 mmol) followed by TBDMSCl (59 mg, 0.393 mmol). After stirring for 12 h at r.t., the reaction was concentrated and the residue taken up in CH_2Cl_2. The solution was washed with saturated $NaHCO_3$ (10 mL), dried, filtered, and finally evaporated to yield the crude product. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (100:1 to 15:1 v/v) afforded the title compound in 66% yield (148 mg). R_f (SiO_2) = 0.64 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (300 MHz, DMSO-d_6): δ11.25 (s, 1H, N-H); 7.47 (s, 1H, H6); 7.39-6.87 (m, 14H, ArH); 6.08 (d, $^3J_{1',2'}$ = 8.7 Hz, 1H, H1'); 5.32 (d, $^3J_{OH,\ H\text{-}3'}$ = 4.8 Hz, 1H, OH); 4.31(br m, 1H, H3'); 3.98-3.64 (m, 2H, CH_2-C2'); 3.73 (s, 3H, OCH_3); 3.23, 3.13 (2dd, $^3J_{4',\ 5'+5''}$ = 4.2, 3 Hz, 2H, H5' and H5''); 2.57 (m, 1H, H2'); 1.37 (s, 3H, CH_3-C5); 0.79 (s, 9H, *t*-Bu-Si); 0.01- -0.04 (t, 6H, $(CH_3)_2Si$); ESI-MS for $C_{37}H_{46}N_2O_7Si$ [MNa^+] calcd. 681.31, found 681.2.

1-[2-α-*C*-*tert*-Butyldimethylsilyloxymethyl-3-*O*-(β-cyanoethyl-*N*,*N*-diisopropyl-phosphoramidic)-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (11). β-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (32 µL, 0.034 g, 0.145 mmol) was slowly added to a solution of **10** (0.148 g, 0.121 mmol) and *N*-ethyl-*N*,*N*-diisopropylamine (76 µl, 0.056 g, 0.436 mmol) in dry THF (1.2 mL) under a N_2

atmosphere. After stirring for 2.5 h, the crude mixture was passed through a 2-cm layer of silica gel (pre-neutralized with 1% NEt_3), and the desired compound eluted by washing with ice cold hexane followed by CH_2Cl_2. Evaporation of the solution afforded the title compound as a nice white foam (101 mg, 52% yield). R_f (SiO_2) = 0.46 (2:1 EtOAc/hexane); ^{31}P NMR (200 MHz, $CDCl_3$): δ 151.7, 150.3; ESI-MS for $C_{46}H_{63}N_4O_8PSi$ [MNa^+] calcd. 881.42, found 881.2.

Solid Phase Synthesis of Oligonucleotides.

Oligonucleotide syntheses were carried out on a 1 μmol scale using an Applied Biosystems DNA/RNA 381A synthesizer as previously described [1,3,15]. Solutions of phosphoramidites **8** and **11** in acetonitrile (0.09 M) were allowed to react with the solid support for an extended coupling time of 30 min using DCI in acetonitrile (0.5M) as catalyst. These conditions afforded over 99% coupling efficiency. Deprotection was conducted by addition of (a) conc. aq. ammonia/ethanol (3:1 v/v, 1 mL, 48 h, r.t.) followed by evaporation; (b) TEA•3HF (100-200 μL, 48 h, r.t.) followed by evaporation. Typically, 40-80 OD units (A_{260}) of the oligonucleotides were obtained at this point. Oligonucleotides were purified by anion-exchange HPLC (Protein Pak DEAE-5PW column-Waters; 7.5 mm × 7.5 cm), desalted by size-exclusion chromatography on Sephadex G-25 matrix, and characterized by MALDI-TOF mass spectrometry (Kratos Kompact-III instrument; Kratos Analytical Inc., New York). Purity of the isolated oligonucleotides was >95%.

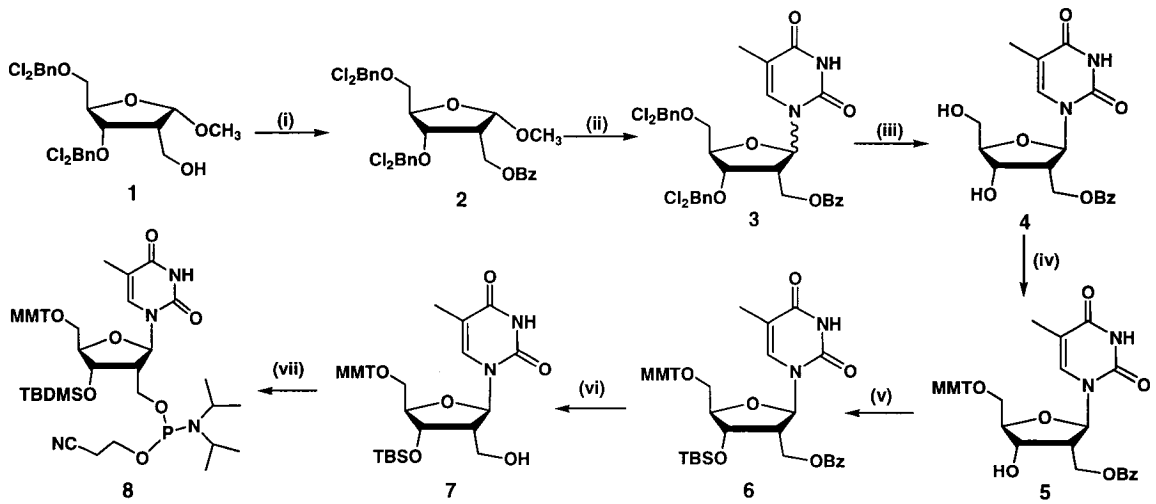
UV Thermal Denaturation Studies.

UV thermal denaturation data was obtained on a Varian CARY 1 spectrophotometer equipped with a Peltier temperature controller (Varian; Mulgrave, Australia). Oligomers and complementary targets were mixed in equimolar ratios in 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4 buffer, pH 7.2, which is representative of intracellular conditions [16]. The total strand concentration was 2.6 μM. Samples were heated to 90 °C for 5 min, cooled slowly to r.t., and refrigerated (4 °C) overnight before measurements. Prior to the thermal run, samples were degassed by placing them in an ultrasound bath for 1 min. Denaturation curves were acquired at 260 nm at a rate of heating of 0.5° C/min. The data

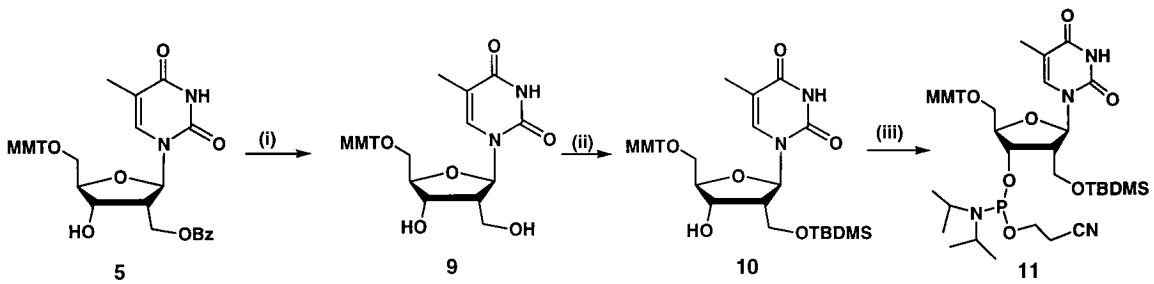
were analyzed with the software provided by Varian Canada and converted to Microsoft Excel. T_m values were calculated as the maximum of the first-derivative plots of absorbance *versus* temperature and have an uncertainty of ±1 °C. Hyperchromicity values (*i.e.* changes in relative absorbance) were calculated using the formula: $H = (A_t-A_i)/A_h$, where H is the hyperchromicity, A_t is the absorbance at any given temperature (t), A_i is the initial absorbance reading, and A_h is the absorbance at the highest temperature.

Circular Dichroism (CD) Spectra.

CD spectra (200-350 nm) were collected on a Jasco J-710 spectropolarimeter at a rate of 100nm/min using fused quartz cells (Hellma, 165-QS). Measurements were carried out in 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4 buffer, pH 7.2 [16] at a duplex concentration of 2.6 μM. The temperature was controlled by an external circulating bath (VWR Scientific) at constant temperature (5 °C). The data were processed on a PC computer using J-700 Windows software supplied by the manufacturer (JASCO, Inc.). To facilitate comparisons, the CD spectra were background subtracted, smoothed and corrected for concentration so that molar ellipticities could be obtained.



Scheme 2.2: Synthesis of **8**. (i) BzCl, py., r.t., 4 h, 100%; (ii) bis(trimethylsilyl) thymine, $SnCl_4$, CH_3CN, 50°C, overnight, 72% (isomer mixture, β/α 5:1); (iii) 10% Pd/C, H_2 (40-50 psi), MeOH, r.t., 6 h, 71% β pure isomer; (iv) MMTr-Cl (1.5eq.), $AgNO_3$, py, r.t., overnight, 72%; (v) TBDMSOTf, 2,6-lutidine, THF, r.t., 4 h, 83%; (vi) NaOMe, MeOH, r.t., 2-3 h, 100%; (vii) iPr$_2$NEt, (iPr$_2$N)(OCH$_2$CH$_2$CN)PCl, THF, r.t., 2 h, 97%.



Scheme 2.3 Synthesis of **11**. (i) NaOMe, MeOH, r.t., 2 h, 100%; (ii) TBDMSCl, $AgNO_3$, py, r.t., overnight, 66%; (iii) iPr$_2$NEt, (iPr$_2$N)(OCH$_2$CH$_2$CN)PCl, THF, r.t., 2.5 h, 52%.

Results and Discussion

Monomer Synthesis

Schemes 2.2 and 2.3 show the synthesis of 2'-α-hm-dT (abbreviated in this work as "**H**") and its conversion to the 2'-*CH$_2$O*- (**8**) and 3'-*O*-phosphoramidites (**11**), using a synthetic strategy analogous to that previously described by Schmit [17] and Li and Piccirilli [18]. We protected the C2'-CH$_2$-OH moiety of *O*-glycoside **1**[14] as the benzoyl ester rather than as Schmit's acetyl ester, since we found that the latter was not stable under the conditions of the subsequent hydrogenolysis reaction (Scheme 2.2). Coupling of **2** with bis(trimethylsilyl)thymine in the presence of $SnCl_4$ catalyst gave anomeric nucleosides **3** in good yield with a β/α stereoselectivity of 5:1. It is interesting to note that the β/α stereoselectivity was found to be higher for the C2-CH$_2$-OAc derivative (10:1), in agreement with Schmit's findings.[12,17] Separation of the anomers by column chromatography was achieved after removal of the 2,4-dichlorobenzyl protecting groups, to provide the desired β anomer in 71% yield from **3**. The anomeric configuration of **4** was unequivocally established using NOESY NMR, which displayed strong H1'-H4', H2'-H6', 2'-CH$_2$-H4', and H5'-H6 cross peaks. Reprotection of hydroxyl groups with MMTrCl followed by TBDMSOTf gave **6** in good yields.[19] Attempts to carry out the silylation reaction with TBDMSCl (DMF/imidazole or $AgNO_3$/py) proved to be problematic.[20] Lastly, we converted **6** to the 2'-CH$_2$O-phosphoramidite derivative **8** *via* consecutive debenzoylation and phosphitylation reactions (Scheme 2.2). To access the 3'-*O*-phosphoramidite derivative, nucleoside **5** was treated with NaOMe to give diol **9**

Table 2.1: Thermal denaturation data (T_m) of duplexes

No.	Designation	Oligonucleotide	T_m[a] (ΔT_m)[b]		
			DNA	**RNA**	2',5'-**RNA**
	2',5'-**RNA containing** 2',5'-**H** (***H***) and 3',5'-**H** (**H**) [c]				
I	2,5RNA	5'-rGUC UGU UGU GUG ACU CUG GUA AC-2'	br [d]	61.6	41.3
II	2',5'-**H**	5'-rGUC UGU ***H***GU GUG ACU CUG GUA AC-2'	br	60.7 (-0.9)	40.5 (-0.8)
III	2'-5'-**H**×2	5'-rGUC UG***H*** ***H***GU GUG ACU CUG GUA AC-2'	br	60.6 (-0.5)	40.0 (-0.7)
IV	2',5'-**H**×3	5'-rGUC UG***H*** ***H***G***H*** GUG ACU CUG GUA AC-2'	br	60.5 (-0.3)	39.1 (-0.6)
V	2',5'-r**A**	5'-rGUC UGU **A**GU GUG ACU CUG GUA AC-2'	br	56.8 (-4.8)	37.3 (-3.5)
VI	2',5'-r**A** ×2	5'-rGUC UG**A** **A**GU GUG ACU CUG GUA AC-2'	br	54.3 (-3.7)	34.0 (-3.7)
VII	2',5'-r**A** ×3	5'-rGUC UG**A** **A**G**A** GUG ACU CUG GUA AC-2'	br	48.0 (-4.5)	28.0 (-4.3)
VIII	3',5'-**H**	5'-rGUC UGU **H**GU GUG ACU CUG GUA AC-2'	br	58.0 (-3.6)	37.4 (-3.9)
	RNA containing 2',5'-**H** (***H***) **and** 3',5'-**H** (**H**)				
IX	RNA	5'-rGUC UGU UGU GUG ACU CUG GUA AC-3'	65.0	78.4	57.1
X	2',5'-**H**	5'-rGUC UGU ***H***GU GUG ACU CUG GUA AC-3'	62.0 (-3.0)	77.2 (-1.2)	52.8 (-4.3)
XI	3',5'-**H**	5'-rGUC UGU **H**GU GUG ACU CUG GUA AC-3'	63.0 (-2.0)	77.7 (-0.7)	54.1 (-3.0)
XII	2',5'-r**U**	5'-rGUC UGU **U**GU GUG ACU CUG GUA AC-3'	62.0 (-3.0)	78.0 (-0.4)	54.1 (-3.0)
XIII	3',5'-r**A**	5'-rGUC UGU **A**GU GUG ACU CUG GUA AC-3'	61.1 (-3.9)	74.1 (-4.3)	51.1 (-6.0)
	DNA control				
XIV	DNA	5'-dGTC TGT TGT GTG ACT CTG GTA AC-3'	68.0	69.0	br

[a] T_m in °C, deviation ±1 °C; Complementary DNA sequence 5'-dGUU ACC AGA GUC ACA CAA CAG AC-3'; complementary RNA sequence: 5'-rGUU ACC AGA GUC ACA CAA CAG AC-3'; complementary 2',5'-RNA sequence: 5'-GUU ACC AGA GUC ACA CAA CAG AC-2'; Buffer: 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4, pH = 7.2; [b] ΔT_m, the T_m change per one modified nucleotide or mismatch relative to entries I or IX; [c] 2',5'-**H** is shown as ***H***; 3',5'-**H** as **H**; 2',5'-r**A** as **A**, 3',5'-r**A** as **A**; 2',5'-r**U** as **U** in the sequence. [d] Broad transition.

quantitatively. Consecutive silylation with TBDMSCl ($AgNO_3$/py) and phosphitylation reactions then afforded the desired compound **11** in moderate yield (52%) (Scheme 2.3).

Oligonucleotide Synthesis.

To investigate the effect of 2',5'- and 3',5'-linked **H** units on the stability of duplexes, monomers **8** and **11** were incorporated into various oligonucleotide sequences by conventional phosphoramidite chemistry (Table 2.1).[1,15,20] 4,5-Dicyanoimidazole (DCI) was used to activate the phosphoramidites[15] and provided average coupling efficiencies of 99% as monitored by the release of the monomethoxytrityl (MMT) cation (see Materials and Methods). Deprotection conditions of oligonucleotides were similar to those employed in RNA synthesis.[20] Oligonucleotides were purified by anion-exchange HPLC and their identity verified by MALDI-TOF mass spectrometry (Supplementary Data). The HPLC chromatograms showed that 2',5'-RNA oligomers elute more rapidly relative to 3',5'-RNA oligomers of the same base composition (Supplementary Data).

Hybridization Studies (T_m and CD analysis).

The binding affinity of various mixed backbone oligonucleotides with one to three **H** units to complementary ssDNA, ssRNA and 2',5'-ssRNA targets was evaluated in a buffer designed to simulate intracellular conditions (Table 2.1). The oligomers, 23 nt in length, were a U5 region of HIV-1 genomic RNA. Oligomers containing 2',5'-**H** substitutions (**II-IV, X**), 3',5'-**H** substitutions (**VIII, XI**) or mismatched bases (**V-VII, XIII**) were prepared in order to ascertain Watson-Crick pairing of the **H** residues. As a comparison, the hybridization properties of unmodified 2',5'-RNA (**I**), RNA (**IX**) and DNA (**XIV**) sequences were also measured, along with a 3',5'-linked oligomer containing a single 2',5'-linked uridine residue (**XII**). Thermal dissociation data (T_m, ΔT_m) for the complexes formed are presented in Table 2.1 and representative melting and circular dichroism (CD) curves are shown in Figure 2.3. The key observations can be summarized as follows:

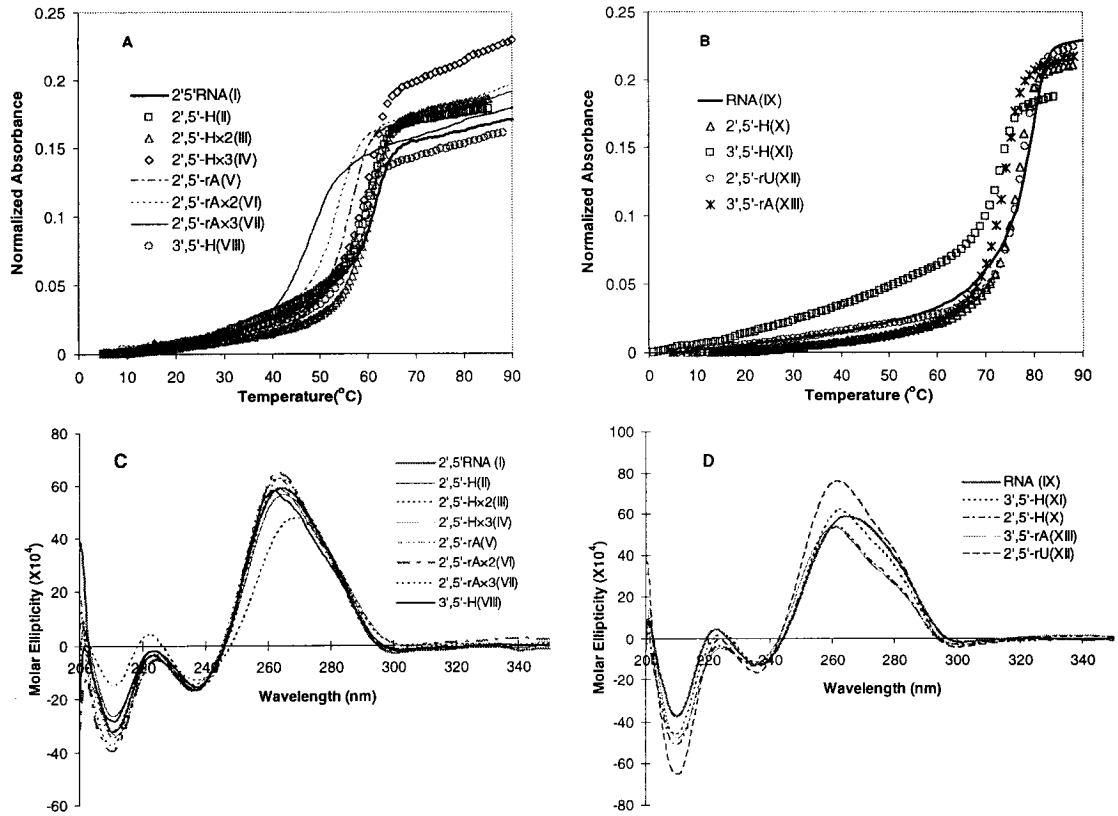


Figure 2.3: Buffer: 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4, pH 7.2. Oligonucleotides were hybridized to complementary RNA (**A** and **C**) Thermal melting curves and CD profile, respectively, of 2',5'-RNA with 2',5'-**H**, 3',5'-**H** and mismatch 2',5'-rA; (**B** and **D**). Thermal melting curves and CD profile, respectively, of RNA with 2',5'-**H**, 3',5'-**H**, 2',5'-rU and mismatch 3',5'-rA.

(a) 2',5'-H Substitutions in the 2',5'-RNA strand: Both 2',5'-RNA:RNA and 2',5'-RNA: 2',5'-RNA duplexes can accommodate a single 2',5'-**H** residue with a small loss of stability (ΔT_m ~ -0.8-0.9 °C). This destabilization is significantly smaller than that created by a mismatch at the same position (ΔT_m ~ -4 °C), suggesting that 2',5'-**H** residues in these duplexes retain classical base-pairing interactions. When the number of 2',5'-**H** units is increased to three (*e.g.*, hybrid **IV**:RNA) the depression in T_m is only -0.3 °C/modification, suggesting a stabilizing effect by the nearly contiguous 2',5'-**H** residues (Figure 2.3A). A nearly superimposable CD signature of the singly and triply substituted hybrids (**II**:RNA and **IV**:RNA) with control duplex **I**:RNA was observed, strongly suggesting that the 2',5'-**H** inserts do not perturb the global morphology of the duplexes (Figure 2.3C). The affinity towards RNA targets is likely to be dependent upon the ratio and intrastrand placement of

normal (2',5'-rN) and 2',5'-**H** units, and based on the above results, a smaller thermal destabilization is to be expected upon increasing the number of consecutive **H** residues. In fact, based on the observed trend from one to three inserts, a fully-modified strand constructed from 2',5'-**H** units might possibly have a binding affinity to RNA targets comparable to that of 2',5'-RNA.

The modest destabilization induced by 2',5'-**H** units is remarkable in view of the one bond chain extension introduced at each of these residues compared with 2',5'-RNA (Figure 2.1 and Scheme 2.1). In addition to the energy (electrostatic) considerations described above, these phenomena may be explained by the anticipated conformation of the 2',5'-**H** units (Figure 2.2). Yathindra *et al.* have shown that the sugars of 2',5'-linked RNA favor a C2'-*endo* conformation, which renders the backbone to be "compact" and of almost equivalent length to that found in the native RNA (*i.e.*, C3'-*endo*).[10,11] With the C3′-*endo* sugar, the 3′-hydroxyl group would sterically interfere with the 2'-*O*-phosphate linkages. The same "compact" conformation is expected for 2',5'-**H** residues, since a strong O4'-C4'-C3'-O3' gauche effect (and lack of an opposing O4'-C1'-C2'-O2' gauche effect) would reinforce the C2'-*endo* pucker (Figure 2.2). Hence, oligonucleotides that are pre-organized in a ''compact'' (or RNA-like) conformation are expected to bind to "compact" RNA strands, and bind weakly, if at all, to ''extended'' oligonucleotide targets (*e.g.* ssDNA). Such conformational compatibility or spatial ''matching'' likely accounts for the ability of oligonucleotides **I** through **VIII** to maintain a stable association with complementary RNA, but not with ssDNA (Table 2.1).[1]

(b) 2',5'-H Substitutions in the RNA strand & 3',5'-H substitutions in the 2',5'-RNA strand: The destabilizing effect of 2',5'-**H** units appears to be greater when incorporated into RNA strands (ΔT_m > -1 °C/modification), regardless of whether the target is complementary RNA or 2',5'-RNA (Sequence **X,** Table 2.1). Similarly, the presence of 3',5'-**H** within 2',5'-RNA leads to reduced duplex stability, with an equal destabilization observed when targeting both RNA and 2',5'-RNA (Sequence **VIII,** Table 2.1). Consistent with this notion, the singly-substituted hybrids **VIII**:RNA and **X**:RNA showed a blue shift in the CD band at 270 nm, whereas hybrids **II**:RNA and **XI**:RNA displayed nearly the same CD profile as the corresponding unmodified controls (Figure 2.3-C and D). As pointed out previously,[1,11] these observations may reflect the major disruption in the normal helical structure

induced by the abrupt displacement of the backbone from the periphery towards the interior of the helix (2'-CH_2OP to 3'-OP) and vice versa (3'-OP to 2'-CH_2OP).

(c) 3',5'-H Substitutions in the RNA strand: Regarding the effects of 3',5'-**H** units within an RNA strand on hybridization affinity for complementary RNA, a significantly less pronounced decrease in duplex stability was observed in *RNA*:RNA duplexes (ΔT_m ~ -0.7 °C) than for Schmit's *DNA*:RNA hybrids (3',5'-**H** in the antisense DNA strand; ΔT_m ~ -3 °C).[12] It may be speculated that this is caused, at least in part, by the narrower minor groove width of *DNA*:RNA hybrids (compared to *RNA*:RNA) leading to more unfavourable interactions involving the large 2'-CH_2-OH group. A much larger decrease in binding affinity was observed for the RNA strand containing a mismatch, confirming the contribution of the thymine base of 3',5'-**H** to binding (**XIII**:RNA *vs* **XI**:RNA; Table 2.1).

Stabilization of Hairpin Structures by 2',5'-H Units.

We next turned our attention towards RNA "hairpins" and studied the effect of incorporating single or multiple 2',5'-**H** units into the loop of these structures (Figure 2.4-A and Table 2.2). We have previously shown that hairpin **P1** retains an A-form conformation, displays ample resistance against nucleases, and inhibits the RNase H activity of HIV reverse transcriptase.[15,21,22] The ribonucleotide residues in the loop are connected by 2',5'-phosphodiester linkages and collectively fold into a distinct rigid structure that is unlike the native 3',5'-tetraloop structure.[15] In view of the anticipated compact conformation of 2',5'-**H** residues (Figure 2.2), it was of interest to discover how these units would influence the stability of the hairpin. This

Table 2.2: Thermal denaturation data of RNA hairpins with modified 2',5'-linked loops

ID	Designation [a]	Hairpin Loop	T_m [b] (ΔT_m)[c]
P1	RRR	5'-GGAC(UUCG)GUCC-3'	64.7
P2	RR(H_5)R	5'-GGAC(***H***UCG)GUCC-3'	63.7 (-1.0)
P3	RR(H_6)R	5'-GGAC(U***H***CG)GUCC-3'	66.2 (+1.5)
P4	RR($H_{5,6}$)R	5'-GGAC(***HH***CG)GUCC-3'	65.1 (+0.2)

[a]Underlined residues are connected *via* 2',5'-phosphodiester linkages; [b] T_m was measured at the wavelength of 260 nm in 0.01 M Na_2HPO_4, 0.1 mM Na_2EDTA, pH 7.0; oligonucleotide concentration ~4.5 μM. Values represent the average of at least five independent measurements. Error in T_m is within ±1 °C. [c]ΔT_m, the T_m change per one 2',5'-**H** (represented as ***H*** in the sequence).

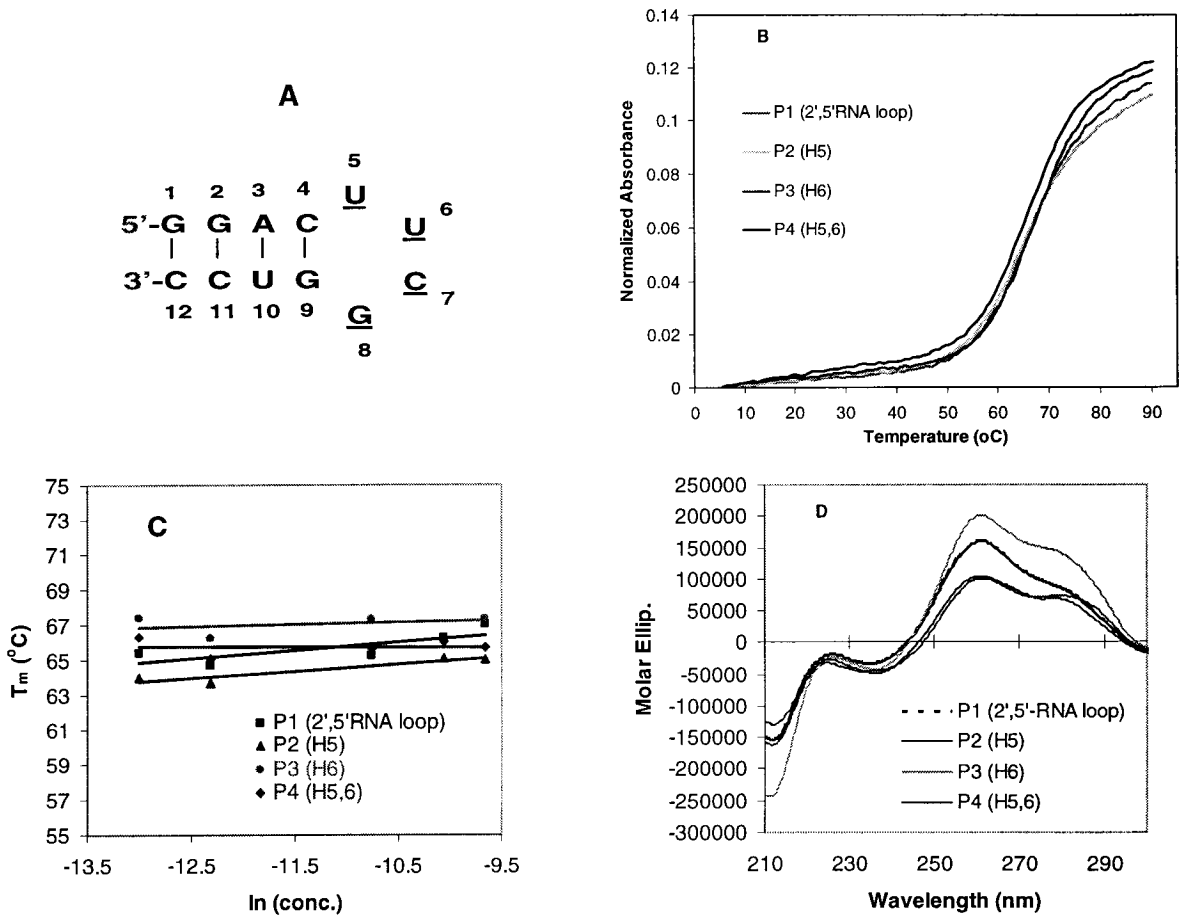


Figure 2.4. (**A**). Hairpin structure and residue numbering: the tetra loop UUCG is connected via 2',5'-phosphodiester linkages, and the modification occurs at the U_5 and/or U_6 positions; (**B**). Thermal melting curves of RNA hairpins with 2',5'-**H** modified loops; (**C**). T_m concentration independence over 30-fold range; (**D**). CD spectra.

was particularly intriguing given that a recent study by Denisov *et al.* showed that the 2',5'-rUCCG loop residues of hairpin **P1** adopt a U5(extended)-U6(compact)-C7(compact)-G8(extended) conformation.[22] Therefore, we anticipated that substitution at position 6 would be better tolerated than one at position 5, particularly if the compact conformation of the 2',5'-**H** unit was preserved at both loop positions. Indeed, we found that a 2',5'-**H** unit at position 6 (**P3**) increases hairpin stability, with an increase in T_m of +1.5 °C relative to the parent hairpin **P1** (Table 2.2). This stabilization is nearly offset by the incorporation of a second 2',5'-**H** unit at position 5 (**P4**, ΔT_m +0.2 °C), which may again be rationalized on the basis of a very strong preference of the 2',5'-**H** unit for a compact (C2'-*endo*) conformation. In view of this, it is not surprising that a single 2',5'-**H** substitution at position 5 (**P2**), where an extended

nucleotide conformation is preferred, is destabilizing (ΔT_m -1.0 °C). Regarding the CD profiles, the solution conformations of all hairpins are similar to that of the control **P1**, with only slight variations in the intensity of the bands. The hairpins containing 2',5'-**H** units at position 5 (**P2** and **P4**) were found to present a reduction in the positive CD band at 260 nm (Figure 2.4-D) relative to **P1** and **P3**, consistent with the expected conformational change at this position, i.e., extended 2',5'-rU → compact 2',5'-**H**.

Conclusions

In summary, we have studied the behavior of oligoribonucleotides (2',5'-RNA and RNA) containing **H** units toward complementary RNA, 2',5'-RNA and DNA. The data obtained demonstrated a destabilization effect upon substituting a rU with **H**, regardless of whether **H** was 2',5' or 3',5'-linked. This destabilization was minimized when the 2',5'-**H** and 3',5'-**H** were incorporated into 2',5'-RNA and RNA strands, respectively (ΔT_m < -1.0 °C). It is expected that longer oligomers containing this modification, particularly if they contain 2',5'-linkages, will show adequate thermal stability, significant nuclease stability [23] and binding to mRNA targets.

In spite of the drop in thermal stability observed when 2',5'-**H** units are incorporated into RNA duplexes, a significant increment in stability was observed when they were incorporated into hairpin loops (ΔT_m +1.5 °C). These results and those described above lend strong support to the notion of 'compact'/'extended' 2',5'/3',5'-backbone structure and its effect on hybrid stability [11]. The ability of hairpins containing stabilizing 2',5'-**H** units to inhibit HIV RT *in vitro* will be the subject of a separate publication.

Acknowledgments

We thank Dr. R.N. Hannoush for contributions during the early stages of this work. We thank J. Watts, D. Sabatino and M. Mangos for helpful feedback during the preparation of this manuscript. We acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a grant to M.J.D. C.G.P. acknowledges support from a FQRNT fellowship and a Clifford Wong McGill Major Fellowship. M.J.D. is

recipient of a James McGill Professorship (McGill University). Funding to pay the Open Access publication charges for this article was provided by NSERC Canada.

Supplementary Data

Synthesis of the intermediate **1**, NMR spectra, extra T_m and CD spectra, and Mass spectrometry analysis are available on pp89-123.

References

1. Wasner, M., Arion, D., Borkow, G., Noronha, A., Uddin, A.H., Parniak, M.A. and Damha, M.J. (**1998**). Physicochemical and biochemical properties of 2',5'-linked RNA and 2',5'-RNA:3',5'-RNA "hybrid" duplexes. *Biochemistry*, 37, 7478-7486.

2. Peng, C.G., Hannoush, R.N. and Damha, M.J. (**2005**). Synthesis of a "taller" 2',5'-linked ribonucleic acid: 2'-α-*C*-hydroxymethyl 2',5'-linked RNA. *J. Biomol. Struct. Dyn.*, 22, 856.

3. Giannaris, P.A. and Damha, M.J. (**1993**). Oligoribonucleotides containing 2',5'-phosphodiester linkages exhibit binding selectivity for 3',5'-RNA over 3',5'-ssDNA. *Nucleic Acids Res.*, 21, 4742-4749.

4. Bhan, P., Bhan, A., Hong, M., Hartwell, J.G., Saunders, J.M. and Hoke, G.D. (**1997**). 2',5'-Linked oligo-3'-deoxyribonucleoside phosphorothioate chimeras: thermal stability and antisense inhibition of gene expression. *Nucleic Acids Res.*, 25, 3310-3317.

5. Damha, M.J. and Noronha, A. (**1998**). Recognition of nucleic acid double helices by homopyrimidine 2',5'-linked RNA. *Nucleic Acids Res.*, 26, 5152-5156.

6. Torrence, P.F., Maitra, R.K., Lesiak, K., Khamnei, S., Zhou, A. and Silverman, R.H. (**1993**). Targeting RNA for degradation with a (2'-5')oligoadenylate-antisense chimera. *Proc. Natl. Acad. Sci. U. S. A.*, 90, 1300-1304.

7. Xiao, W., Li, G.Y., Maitra, R.K., Maran, A., Silverman, R.H. and Torrence, P.F. (**1997**). Correlation of selective modifications to a 2',5'-oligoadenylate-3',5'-deoxyribonucleotide antisense chimera with affinity for the target nucleic acid and with ability to activate RNase L. *J. Med. Chem.*, 40, 1195-1200.

8. Kandimalla, E.R., Manning, A., Zhao, Q.Y., Shaw, D.R., Byrn, R.A., Sasisekharan, V. and Agrawal, S. (**1997**). Mixed backbone antisense oligonucleotides: Design, biochemical and biological properties of oligonucleotides containing 2'-5'-ribo- and 3'-5'-deoxyribonucleotide segments. *Nucleic Acids Res.*, 25, 370-378.

9. Damha, M.J., Giannaris, P.A. and Khan, N. (**1991**). 2'-5'-linked oligonucleotides form stable complexes with complementary RNA and DNA. *Nucleic Acids Symp. Ser.*, 290.

10. Premraj, B.J., Patel, P.K., Kandimalla, E.R., Agrawal, S., Hosur, R.V. and Yathindra, N. (**2001**). NMR structure of a 2',5' RNA favors A type duplex with compact C2 ' *endo* nucleotide repeat. *Biochem. Biophys. Res. Commun.*, 283, 537-543.

11. Premraj, B.J., Raja, S. and Yathindra, N. (**2002**). Structural basis for the unusual properties of 2',5' nucleic acids and their complexes with RNA and DNA. *Biophysical Chemistry*, 95, 253-272.

12. Schmit, C., Bevierre, M.-O., De Mesmaeker, A. and Altmann, K.-H. (**1994**). The effects of 2'- and 3'-alkyl substituents on oligonucleotide hybridization and stability. *Bioorg. Med. Chem. Lett.*, 4, 1969-1974.

13. Freier, S.M. and Altmann, K.-H. (**1997**). The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.*, 25, 4429-4443.

14. Martin, P. (**1995**). New access to 2'-*O*-alkylated ribonucleosides and properties of 2'-O-alkylated oligoribonucleotides. *Helv. Chim. Acta*, 78, 486-504.

15. Hannoush, R.N. and Damha, M.J. (**2001**). Remarkable stability of hairpins containing 2',5'-linked RNA loops. *J. Am. Chem. Soc.*, 123, 12368-12374.

16. Alberts, B. (**1989**). *Molecular Biology of the Cell.* Carland Publishing, Inc. New York, p304.

17. Schmit, C. (**1994**). Efficient synthesis of 2'-deoxy-2'-α-*C*-substituted nucleosides. *Synlett*, 238-240.

18. Li, N.-S. and Piccirilli, J.A. (**2004**). Synthesis of the phosphoramidite derivatives of 2'-deoxy-2'-*C*-α-methylcytidine and 2'-deoxy-2'-*C*-α-hydroxymethylcytidine: Analogues for chemical dissection of RNA's 2'-hydroxyl group. *J. Org. Chem.*, 69, 4751-4759.

19. Kocienski, P.J. (**2004**). *Protecting Group.* Georg Thieme, p215.

20. Damha, M.J. and Ogilvie, K.K. (**1993**). Oligoribonucleotide synthesis - the silyl-phosphoramidite method. In Agrawal, S. (ed.), *Methods Mol. Biol.* The Humana Press Inc., Totowa, New Jersey, pp. 81-114.

21. Hannoush, R.N., Carriero, S., Min, K.-L. and Damha, M.J. (**2004**). Selective inhibition of HIV-1 reverse transcriptase (HIV-1 RT) RNase H by small RNA hairpins and dumbbells. *ChemBioChem*, 5, 527-533.

22. Denisov, A.Y., Hannoush, R.N., Gehring, K. and Damha, M.J. (**2003**). A novel RNA motif based on the structure of unusually stable 2',5'-linked r(UUCG) loops. *J. Am. Chem. Soc.*, 125, 11525-11531.

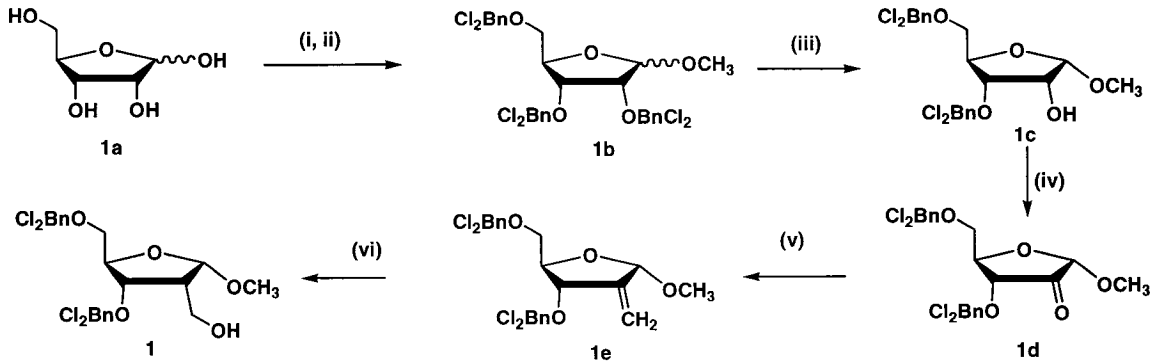
23. Hannoush, R., N., Min, K.-L. and Damha, M., J. (**2004**). Diversity-oriented solid-phase synthesis and biological evaluation of oligonucleotide hairpins as HIV-1 RT RNase H inhibitors. *Nucleic Acids Res.*, 32, 6164-6175.

Chapter II. Synthesis and Hybridization Studies of Oligonucleotides Containing 1-(2-Deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (2'-α-hm-dT)

Supplementary Data

(pp89-123)

Contents



Scheme S2.2a: preparation of **1** in **Scheme 2.2:** (i) MeOH, H_2SO_4, 4°C, overnight, 95%; (ii) 2,4-dichlorobenzyl chloride, NaH, DMF, 40°C, overnight, 73%; (iii) a. $SnCl_4$, 4°C, overnight; b. H_2O, 84%; (iv) Dess-Martin periodinane, CH_2Cl_2, r.t., overnight, 100%; or pyridinium dichromate, acetic acid, CH_2Cl_2, r.t., over night, 68%; or CrO_3, py., Ac_2O, CH_2Cl_2, r.t., 4 h, 42%; (v) methyltriphenylphosphonium bromide (Ph_3PCH_3Br), sodium tert-pentoxide (*t*-C_5H_7ONa), ether, benzene, 4°C, overnight, 95%; (vi) a. 9-bora-bicyclo[3.3.1]nonane (9-BBN), THF, 40°C, overnight; b. $NaBO_3$, H_2O/EtOH, 50°C, 4 h, 82%;

Methyl 2,3,5-tri-*O*-(2,4-dichlorobenzyl)-α,β-D-ribofuranoside (1b). A solution of D-ribose **1a** (25 g, 167 mmol) in 500 mL dry methanol was cooled to 0°C and treated with 2.5 mL concentrated H_2SO_4 under a N_2 atmosphere. The reaction mixture was stirred at 4°C overnight (12-16 h) and then neutralized by passing through a bed of Amberlite IRA-400 (Aldrich) in the free base form. The filtrate was concentrated under vacuum to a colourless syrup (5 g, 30.5 mmol), then redissolved in 30 mL anhydrous DMF. A solution of NaH (4.4 g, 183 mmol) in 50 mL anhydrous DMF was cooled to 0°C under a N_2 atmosphere and then added. The reaction mixture was gradually warmed up and kept at 60°C for one hour until no more gas was produced. The reaction mixture was cooled down to 0°C again and 2,4-dichlorobenzyl chloride (12.7 mL, 91.41 mmol) was added slowly. The reaction solution was gradually warmed up and stirred at 40°C overnight (16 h). The reaction progress was monitored by TLC. The reaction mixture was poured into ice and washed with three portions of CH_2Cl_2. The combined organic phases were dried over Na_2SO_4 and concentrated until a precipitate formed, which was filtered and washed by cold hexanes several times, then dried under vacuum to afford **1b** (14.27 g, 73% yield). The mother liquid after the filtration was evaporated to an oily liquid and still contained some product **1b** (not isolated). R_f (SiO_2) = 0.5-0.54 (3:5 Ether/Hexane); 0.83

(5:2 Ether/Hexane); ^{1}H NMR (200 MHz, $CDCl_3$): δ7.12-7.45 (m, 9H, ArH), 4.98 (s, 1H, H1), 4.59-4.7 (m, 6H, benzyl CH_2), 4.32-4.40 (m, 1H, H4), 4.16 (dd, $^3J_{2,3}$ = 4.6 Hz, $^3J_{3,4}$ = 7.0 Hz, 1H, H3), 3.97 (d, $^3J_{2,3}$ = 4.6 Hz, 1H, H2), 3.60-3.78 (AB dd, $^2J_{5,5'}$ = 16.8 Hz, $^3J_{4,5;4,5'}$ = 3.9, 5.3 Hz, 2H, H5), 3.36 (s, 3H, OCH_3). ESI-MS for $C_{27}H_{24}Cl_6O_5$[MNa^+] calcd. 660.98, found 662.9.

Methyl 3,5-di-*O*-(2,4-dichlorobenzyl)-α-D-ribofuranoside (1c). Under a N_2 atmosphere, a solution of **1b** (14.27 g, 22.2 mmol) in 140 mL CH_2Cl_2 was cooled to 0°C and a 1 M solution of $SnCl_4$ in dry CH_2Cl_2 (34 mL, 34 mmol) was added. The reaction mixture was stirred at 4°C overnight until TLC analysis showed the reaction was completed. The reaction mixture was poured into 140 mL saturated aqueous $NaHCO_3$ and gently mixed to minimize emulsion formation. The aqueous layer was washed with 3 × 150 mL $CHCl_3$. The combined organic phases were dried over Na_2SO_4 and evaporated to an oil. Purification by silica gel column chromatography with ether/hexanes (1:1 v/v) afforded the title compound **1c** as an oil (9 g, 84% yield). R_f (SiO_2) = 0.54-0.58 (5:2 ether/hexanes); ^{1}H NMR (200 MHz, $CDCl_3$): δ7.18-7.42 (m, 6H, ArH), 4.91 (d, $^3J_{1,2}$ = 4.6 Hz, 1H, H1), 4.70, 4.56 (2AB, ^2J = 16.6, 6.6 Hz, 4H, benzyl CH_2), 4.19-4.26 (m, 1H, H4), 4.14 (dd, $^3J_{2,3}$ = 7 Hz, $^3J_{1,2}$ = 4.6 Hz, 1H, H2), 3.84 (dd, $^3J_{3,4}$ = 3.2 Hz, 1H, H3), 3.59 (d, $^3J_{4,5}$ = 4.4 Hz, 2H, H5), 3.48 (s, 3H, OCH_3), 2.90 (d, J = 11.4, 1H, OH). ESI-MS for $C_{20}H_{20}Cl_4O_5$ [MNa+] calcd. 505.18, found 505.0.

Methyl 3,5-di-*O*-(2,4-dichlorobenzyl)-2-oxo-α-D-ribofuranoside (1d). A solution of **1c** (9 g, 18.7 mmol) in 27 mL dry CH_2Cl_2 was added to an ice cold suspension of Dess-Martin periodinane (9.5 g, 22.4 mmol) in 27 mL dry CH_2Cl_2. The resulting mixture was kept stirring at r.t. under a N_2 atmosphere overnight (8-12 h). The reaction progress was monitored by TLC after 4 h. (If necessary, a further 0.2 equiv. (1.58 g, 3.73 mmol) of Dess-Martin periodinane can be added to the reaction mixture to push the reaction to completion.) The reaction mixture was passed through a 2-cm silica gel layer in a sintered glass filter. After washing with ether, the combined filtrates were washed by saturated aqueous $NaHCO_3$ 153 mL, brine 153 mL, dried by Na_2SO_4, and evaporated to dryness. The residue was dissolved in ice-cold ether and filtered to remove insoluble impurities. Evaporation of the ether solution afforded an oil **1d** (8.96g, quantitative yield) which was used for the next reaction without further purification.

Under a N_2 atmosphere, acetic acid (1.27 mL, 2.22 mmol) was added to an ice-cold suspension of pyridinium dichromate (7.22 g, 19.2 mmol) in 20 mL dry CH_2Cl_2. The mixture was stirred at 0°C for 10 min and a solution of **1c** (6.15 g, 12.8 mmol) in 40 mL dry CH_2Cl_2 was added. The reaction mixture turned dark brown, and was stirred at r.t. over night. The reaction progress was monitored by TLC until completion. The reaction mixture was evaporated to dryness, then ether was added and the mixture was filtered through a 2cm silica gel layer in a sintered glass filter to remove most of the chromium salts. After washing with ether, the filtrates were evaporated to an oil **1d** (4.14 g, 68% yield) which was used for the next reaction without further purification.

Under a N_2 atmosphere, acetic anhydride (10.6 mL, 110 mmol) and dry pyridine (17 mL, 210 mmol) were added to an ice-cooled suspension of CrO_3 (10.19 g, 101 mmol) in 90 mL dry CH_2Cl_2 and the reaction mixture was stirred at r.t. for 0.5 h. A solution of **1c** (11.86 g, 24.6 mmol) in 30 mL dry CH_2Cl_2 was added and the mixture was stirred at r.t. for 4 h to get a dark brown solution which was poured into 750 mL EtOAc with vigorous stirring. The resulting mixture was filtered through a 2-cm layer of silica gel in a sintered glass filter, followed by washing with 300 mL EtOAc. The combined filtrates were evaporated (< 30°C) to dryness. The residue was coevaporated with toluene (75 mL) followed by 30 mL CH_2Cl_2 to yield 5 g (42%) **1d** as a syrup, which was used immediately for the next reaction without further purification. R_f (SiO_2) = 0.4 (5:2 ether/hexanes); ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.40-7.20 (m, 6H, ArH), 4.87 (d, 1H, $^4J_{1,3}$ = 0.9 Hz, H1), 4.56-5.05 (m, 4H, benzyl CH_2), 4.37 (ddd, $^3J_{3,4}$ = 9 Hz, $^3J_{4,5;4,5'}$ = 3.3, 2.1 Hz, 1H, H4), 4.22 (d, $^3J_{3,4}$ = 9 Hz, $^4J_{3,1}$ = 1.2 Hz, 1H, H3),3.94, 3.80 (2dd, $^2J_{5,5'}$ = 11.4, $^3J_{4,5;4,5'}$ = 3.3, 2.1 Hz, 2H, H5), 3.50 (s, 3H, OCH_3).

Methyl 3,5-di-*O*-(2,4-dichlorobenzyl)-2-methylene-α-D-ribofuranoside (1e). Under a N_2 atmosphere, a solution of sodium *t*-pentoxide (3.698 g, 33.6 mmol) in 33 mL dry benzene was added to an ice cold suspension solution of methyltriphenylphosphonium bromide (13.335 g, 37.3 mmol) in 853 mL dry ether at 0°C with strong stirring. The resulting mixture was warmed up to r.t. and stirred for 4-6 h. The reaction mixture turned to bright yellow cloudy solution. The reaction mixture then was cooled down to -78°C and a solution of **1d** (8.96 g, 18.7 mmol) in 22 mL dry ether was carefully added. The reaction was then warmed to 4°C and stirred at that

temperature overnight (10-14 h). The reaction progress was monitored by TLC. After the reaction was completed, the reaction mixture was washed with saturated NH_4Cl (1440 mL) and the aqueous phase was extracted by ether three times. The combined organic phases were dried over Na_2SO_4 and evaporated to dryness. Purification by silica gel column chromatography with EtOAc/Hexane (1:4 v/v) afforded **1e** (8.45 g, 94.7%). R_f (SiO_2) = 0.8 (5:2 ether/hexane); 0.55 (1:4 EtOAc/Hexane). ^1H NMR (400 MHz, $CDCl_3$): δ 7.18-7.42 (m, 6H, ArH), 5.46-5.48 (m, 2H, CH_2-C2), 5.26 (t, 1H, H1), 4.56-4.71 (m, 4H, benzyl CH_2), 4.33-4.37 (m, 1H, H4), 4.30-4.34 (m, 1H, H3), 3.70 (d, $^3J_{4,5}$ = 4 Hz, 2H, H5), 3.44 (s, 3H, OCH_3). ESI-MS for $C_{21}H_{20}Cl_4O_4$ [MNa+] calcd. 499.01, found 498.9.

Methyl 2-deoxy-3,5-di-*O*-(2,4-dichlorobenzyl)-2-α-hydroxylmethyl-α-D-ribo-furanoside (1) Under a N_2 atmosphere, **1e** (20.521 g, 42.91 mmol) was added 9-BBN (0.5 M in THF, 250 mL, 125 mmol). The resulting solution was stirred at r.t., then at 40°C overnight (10 h), cooled to r.t., and transferred to a round bottom flask containing a solution of sodium perborate tetrahydrate (39.61 g, 257 mmol) in 257 mL H_2O and 257 mL ethanol. The resulting solution was kept vigorous stirring at 50°C for 4 h and a lot of gas was produced. The reaction mixture was cooled to 0°C, neutralized by acetic acid to pH 8 and concentrated to a small volume which was diluted with 65 ml water and extracted with 436 ml CH_2Cl_2 three times. The combine organic phases were washed by brine, dried over Na_2SO_4 and evaporated to dryness. Purification by silica gel column chromatography with EtOAc/hexane (1:4 v/v) afforded pure **1** (16.414 g, 82% yield). R_f (SiO_2) = 0.54-0.59 (1:4 EtOAc/Hexane). ^1H NMR (400 MHz, $CDCl_3$): δ 7.19-7.40 (m, 6H, ArH), 5.04 (d, $^3J_{1,2}$ = 5.2 Hz, 1H, H1), 4.52-4.60 (m, 4H, benzyl CH_2), 4.31-4.35 (m, 1H, H4), 4.06 (dd, $^3J_{3,2}$ = 7.8 Hz, $^3J_{3,4}$ = 3.4 Hz, 1H, H3), 3.94 (t, 2H, CH2-C2), 3.55-3.64 (AB dd, $^2J_{5,5'}$ = 20.6 Hz, $^3J_{4,5;4,5'}$ = 4.8, 4.4 Hz, 2H, H5), 3.424 (s, 3H, OCH_3), 2.51-2.58 (m, 1H, H2), 1.834 (t, $^3J_{OH,CH2}$ = 5.6, 1H, OH). ESI-MS for $C_{21}H_{22}Cl_4O_5$ [MNa^+] calcd. 517.02, found 517.1.

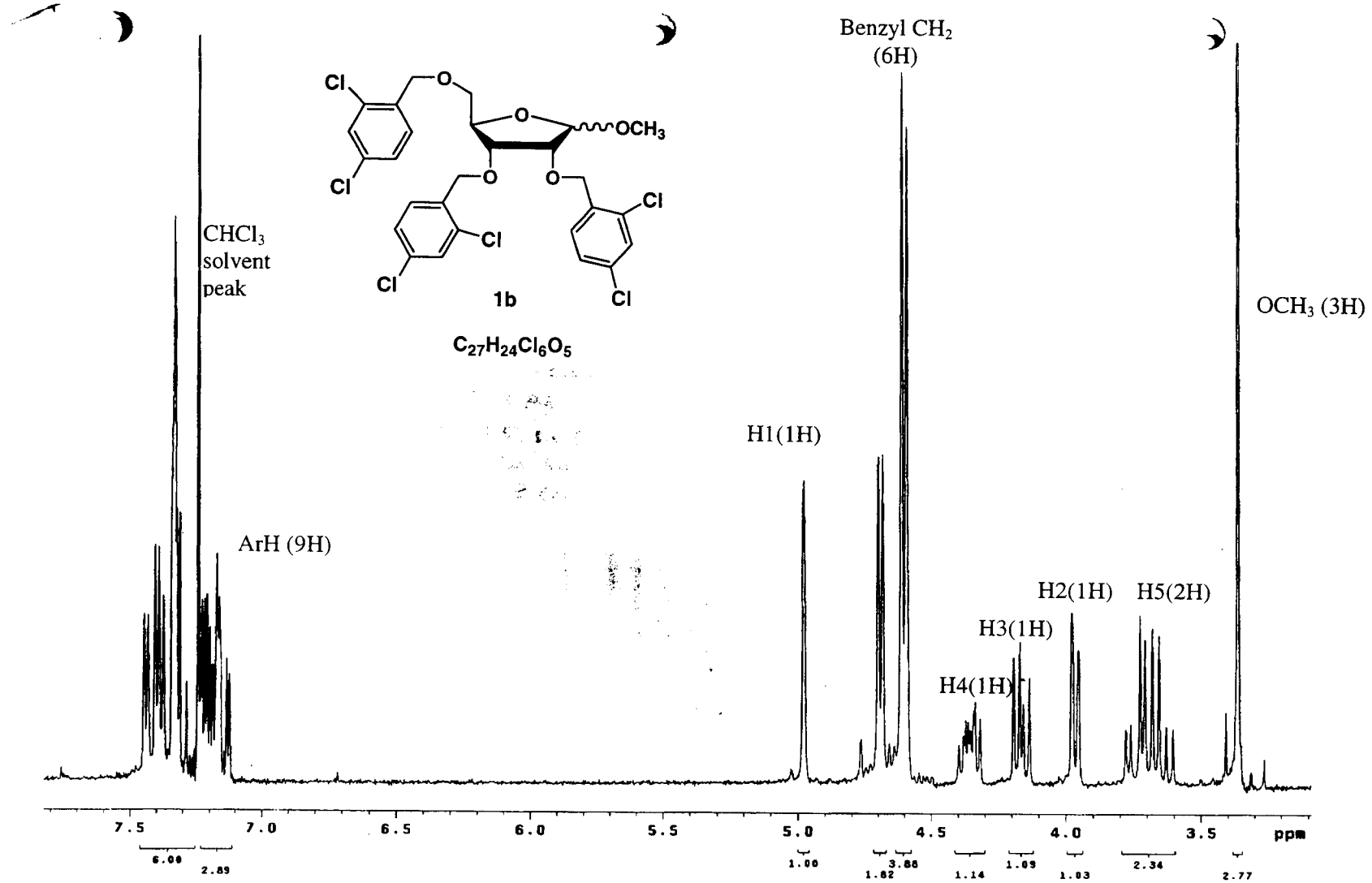


Figure S2.i: ^1H NMR (200MHz, $CDCl_3$) of **1b**

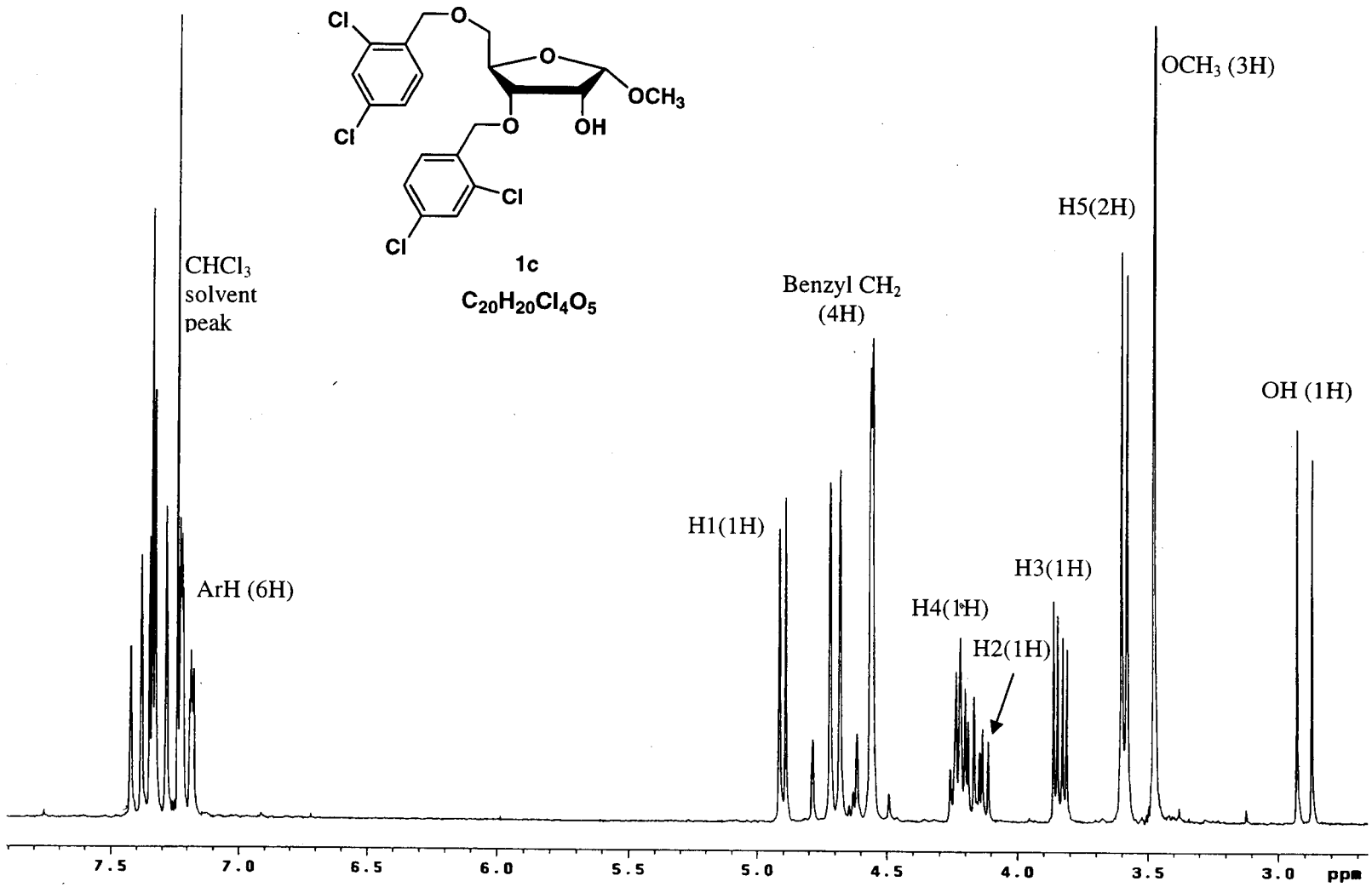


Figure S2.ii: ^{1}H NMR (200MHz, $CDCl_3$) of **1c**

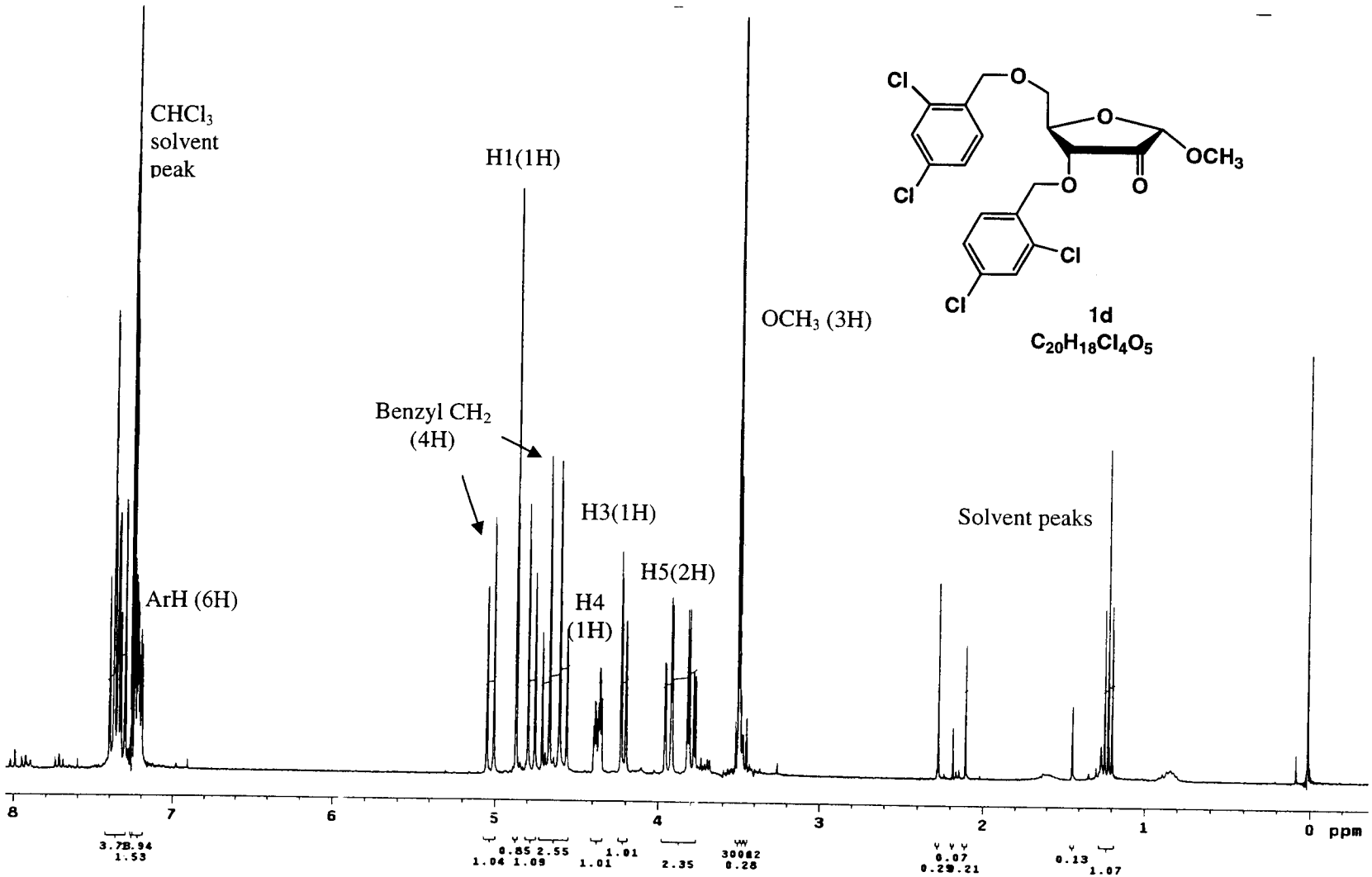


Figure S2.iii: ^{1}H NMR (300MHz, $CDCl_3$) of **1d**

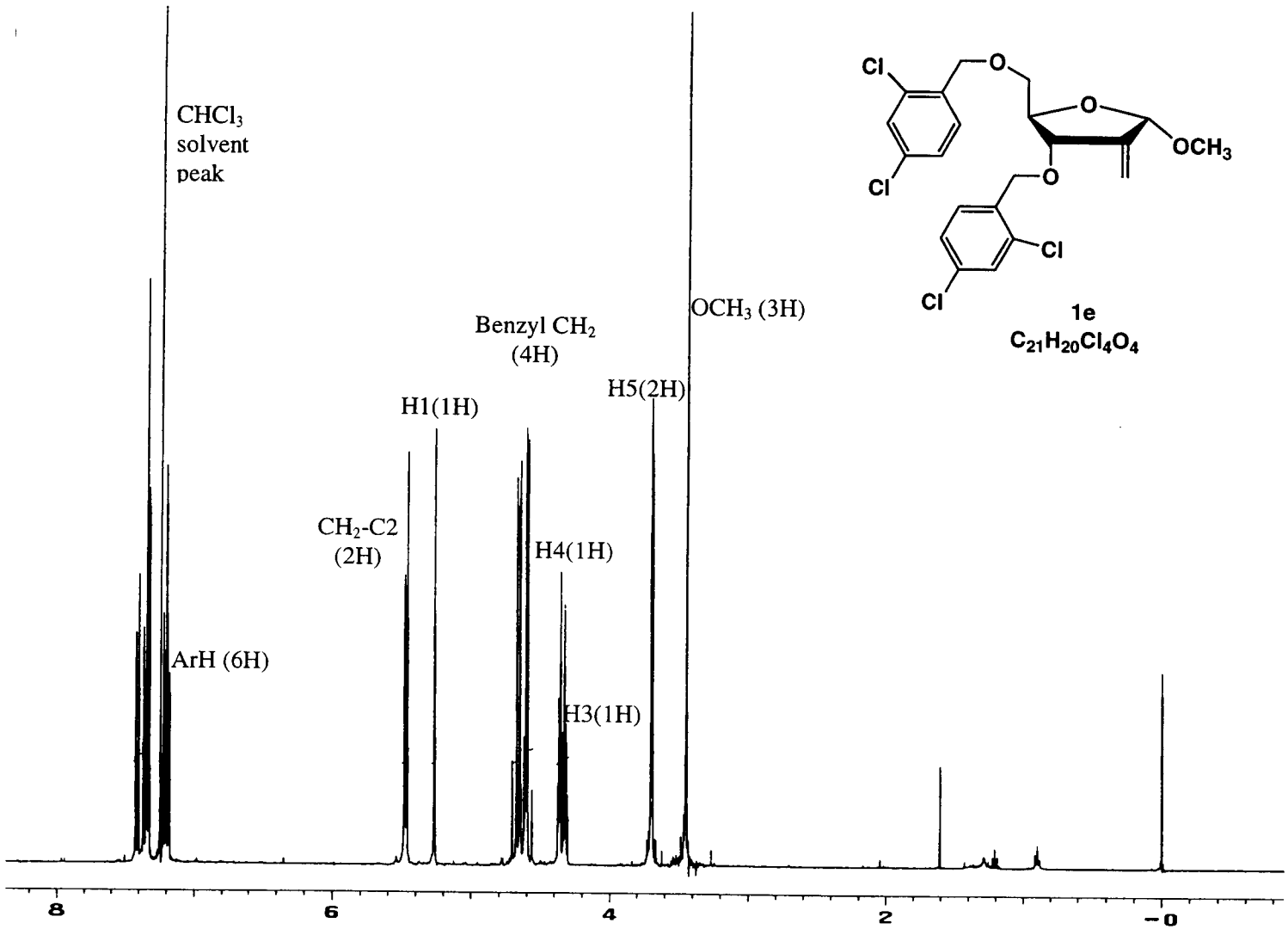


Figure S2.iv: ^{1}H NMR (400MHz, $CDCl_3$) of **1e**

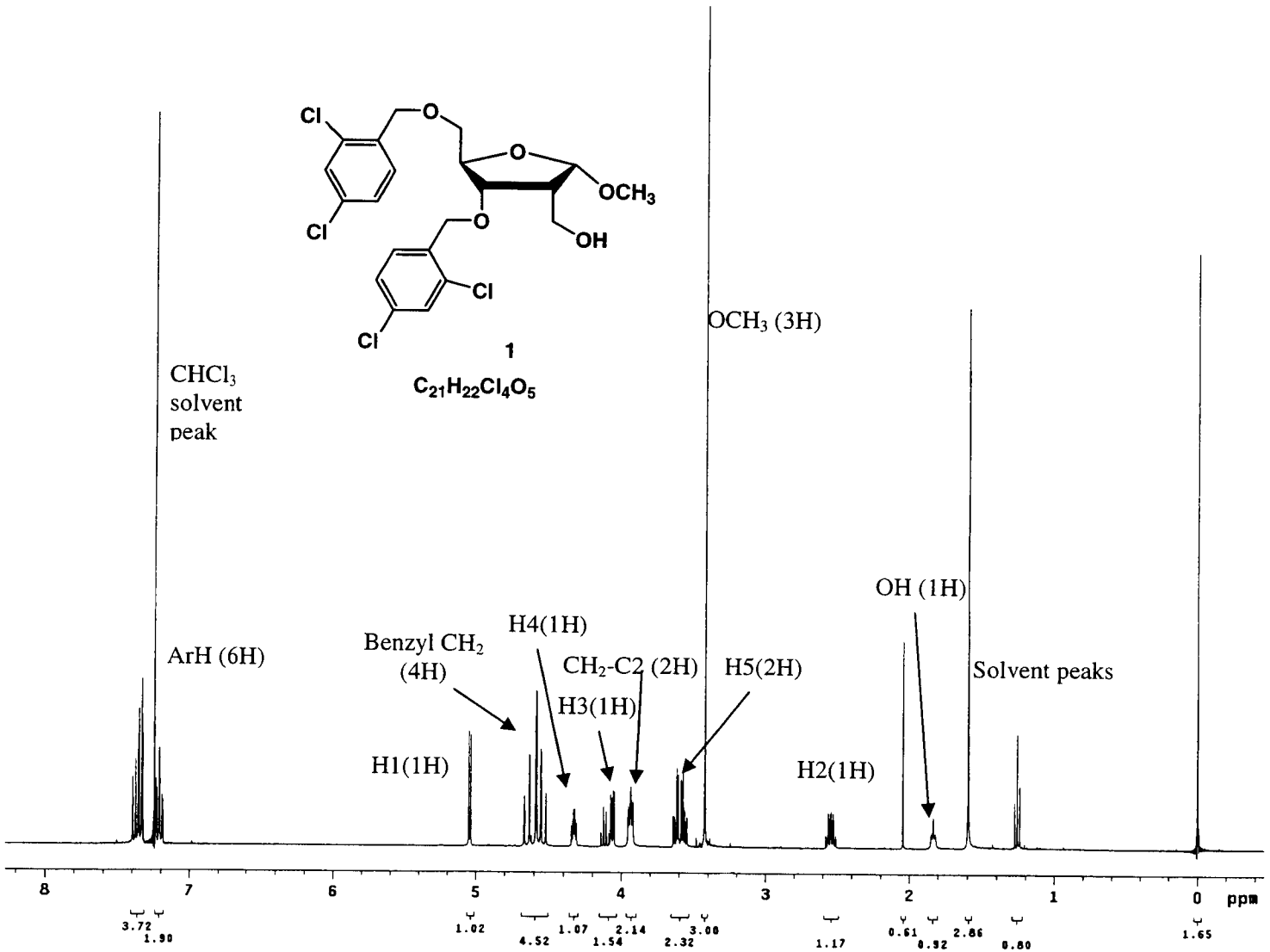


Figure S2.1: ^{1}H NMR (400MHz, $CDCl_3$) of **1**

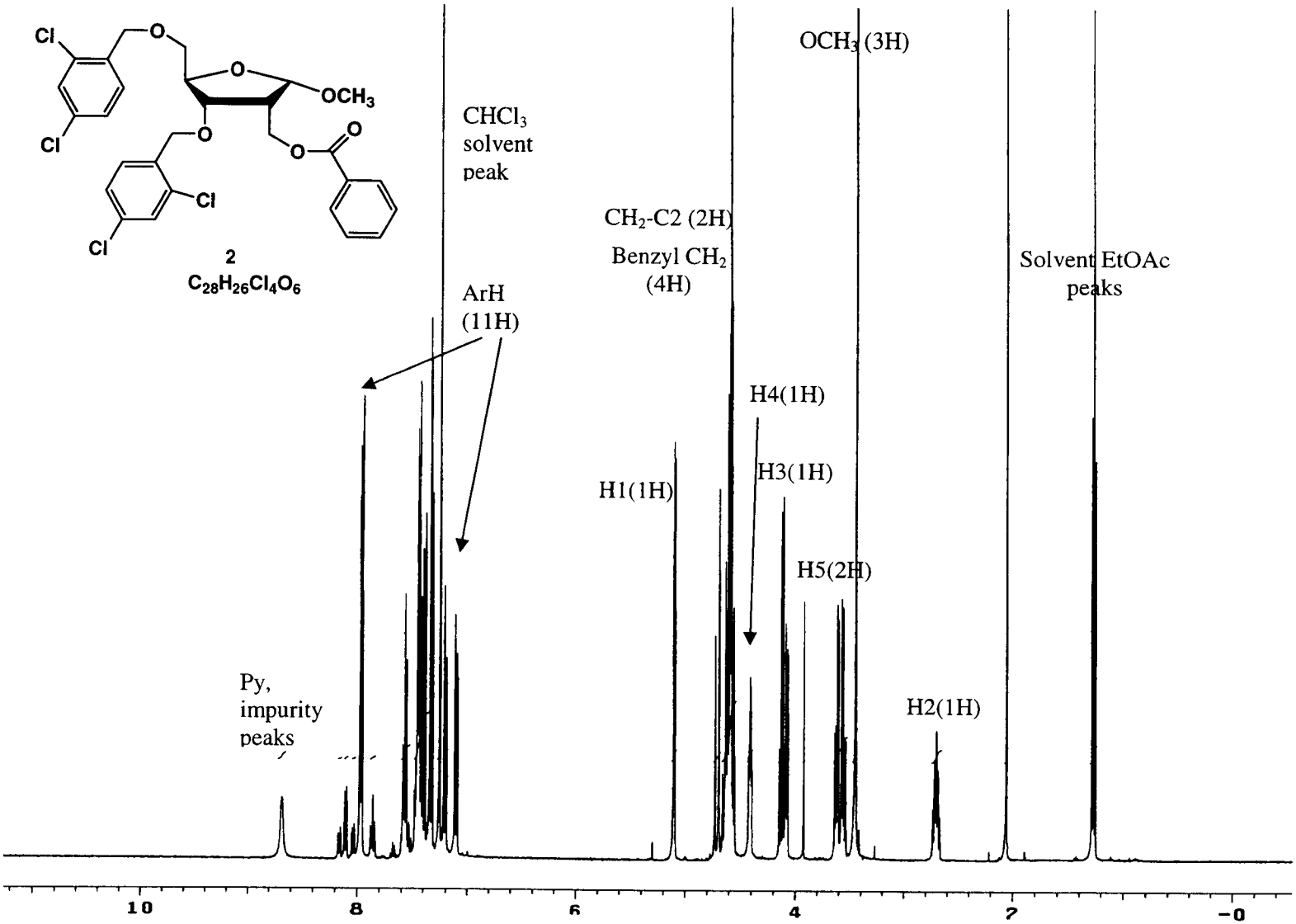


Figure S2.2: ^{1}H NMR (400MHz, $CDCl_3$) of **2**

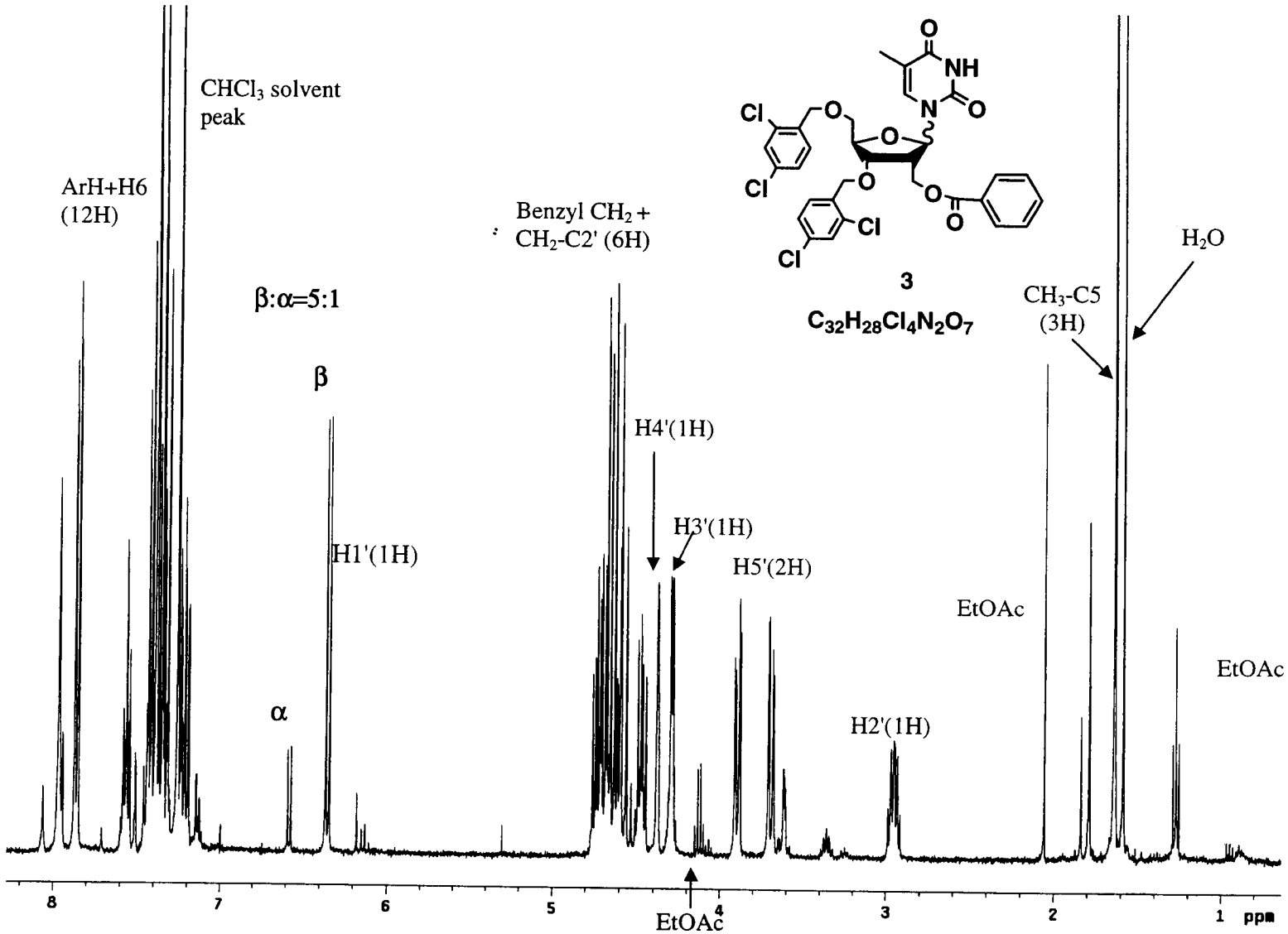
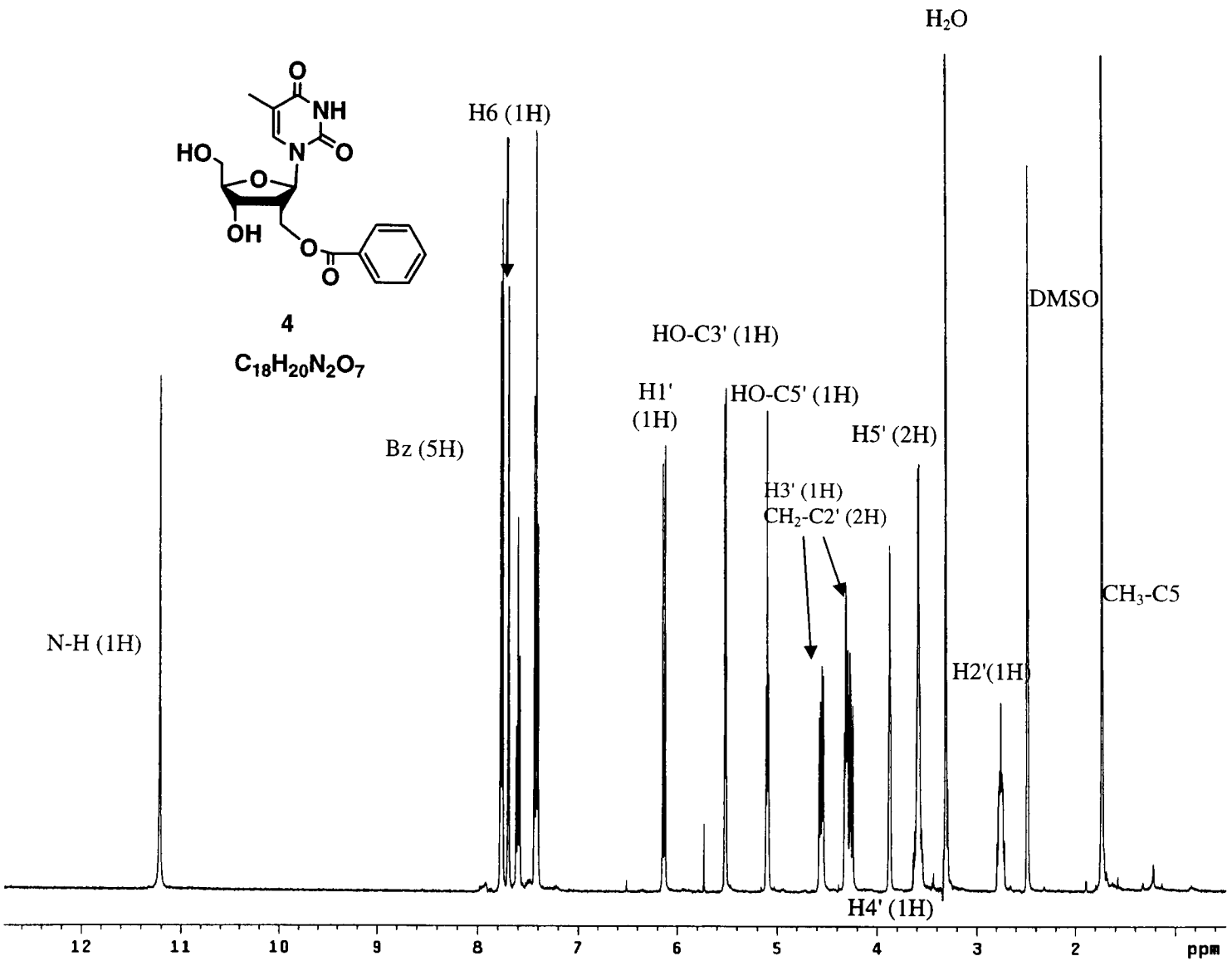


Figure S2.3: ^{1}H NMR (400MHz, $CDCl_3$) of **3**



4

$C_{18}H_{20}N_2O_7$

Figure S2.4a: ^{1}H NMR (400MHz, DMSO-d_6) of **4**

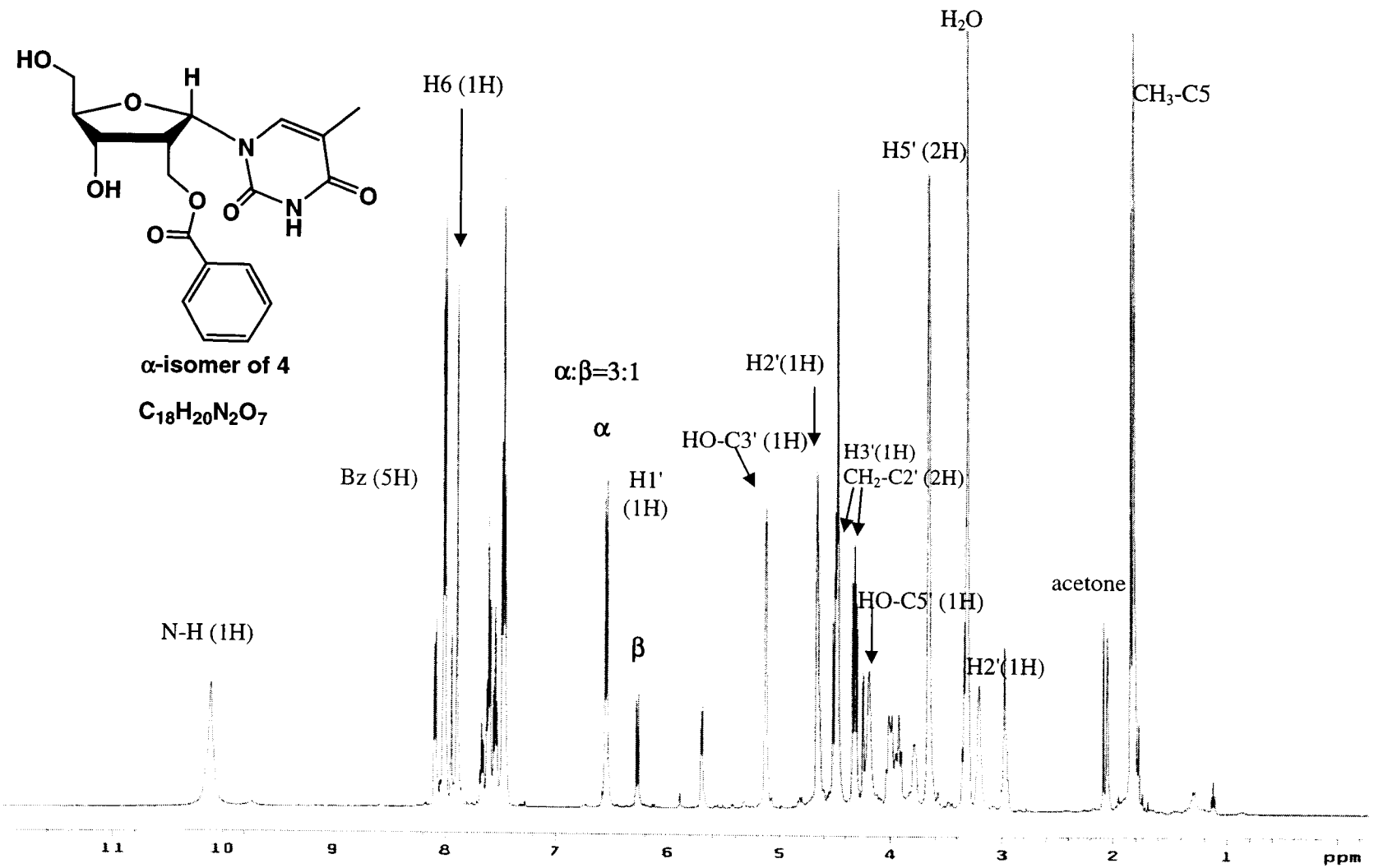


Figure S2.4b: ^{1}H NMR (500 MHz, acetone-d_6) of **4** (α isomer)

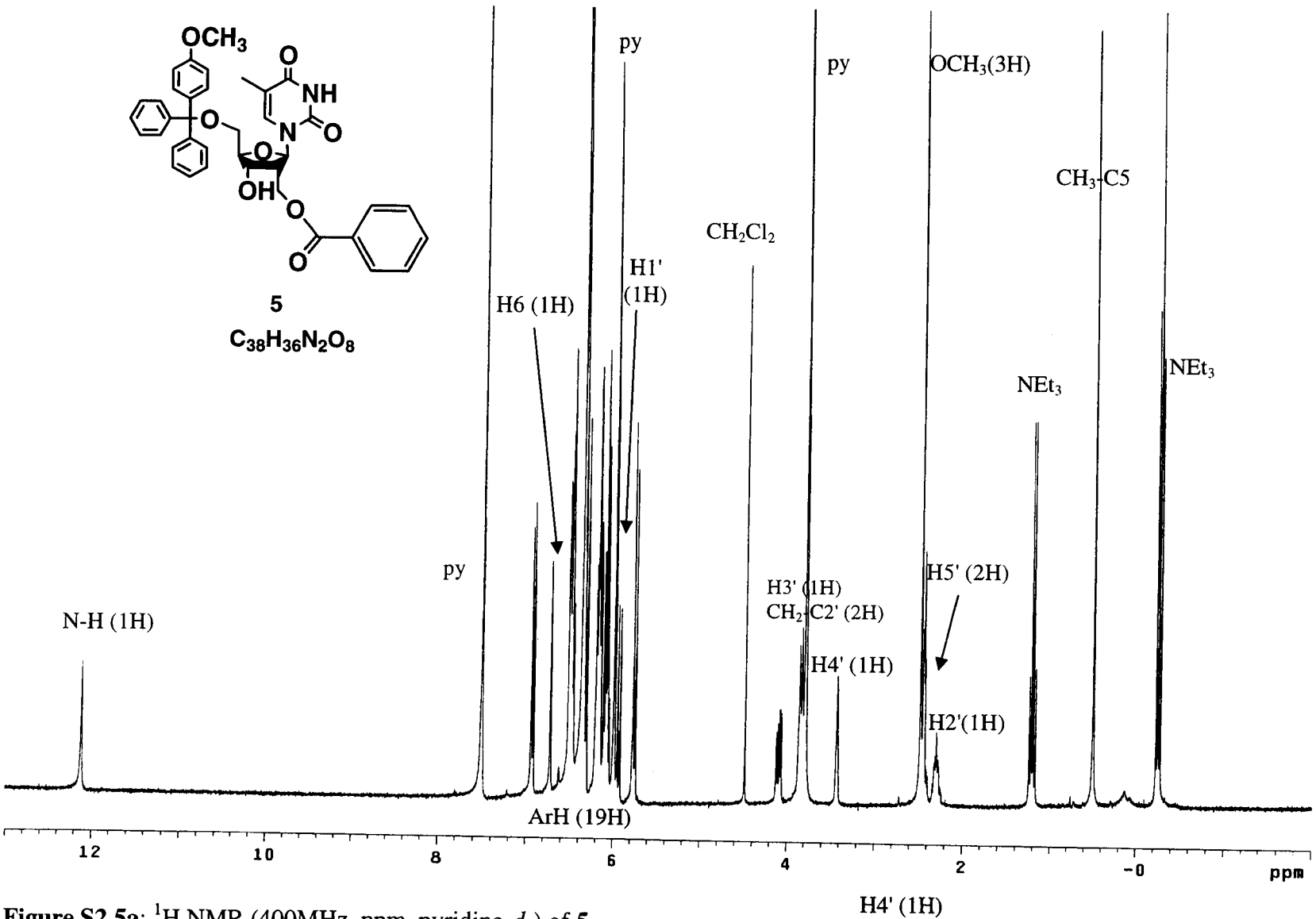


Figure S2.5a: ^{1}H NMR (400MHz, ppm, pyridine-d_5) of **5**.

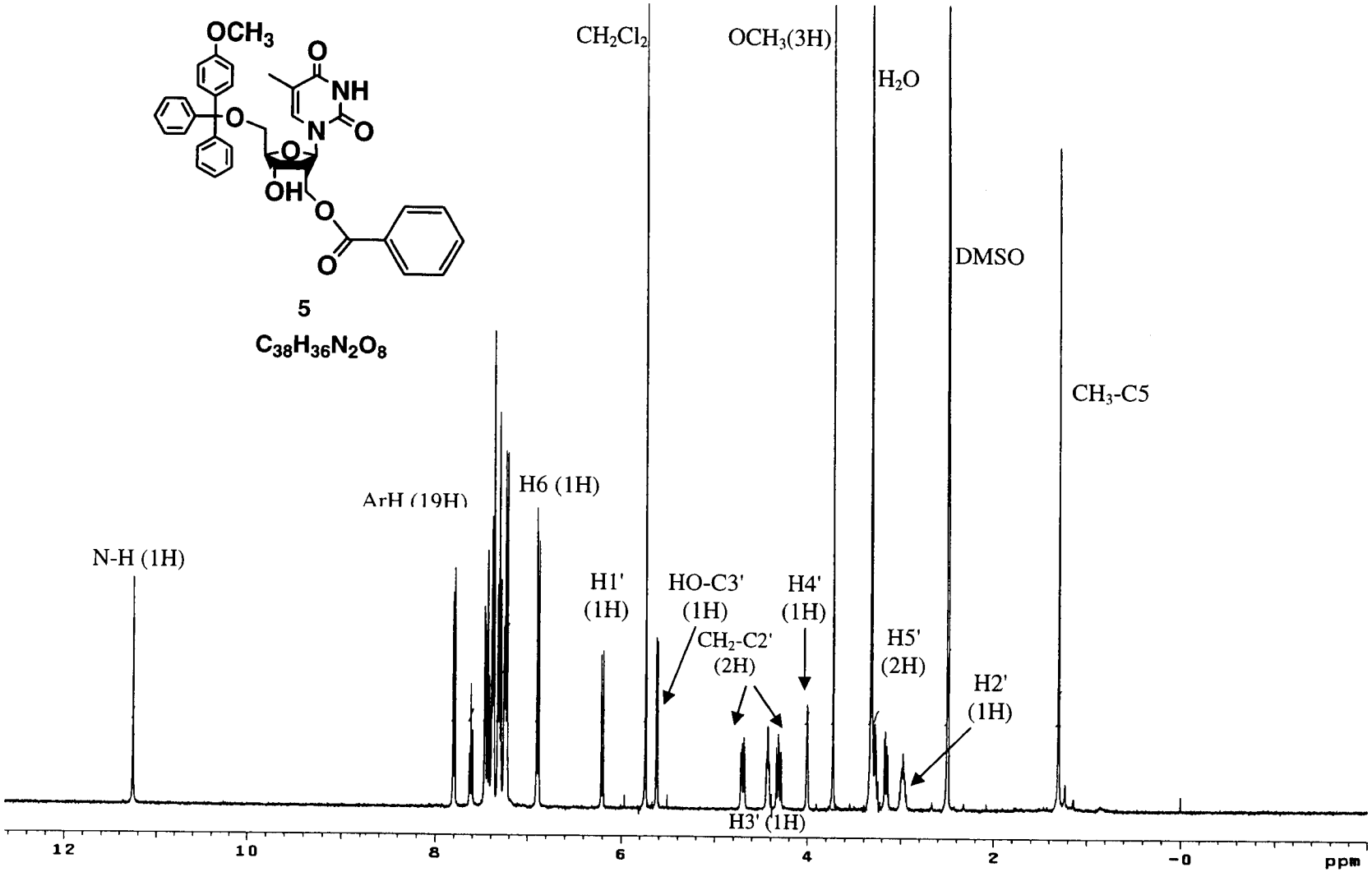


Figure S2.5b: ^{1}H NMR (400MHz, ppm, DMSO-d6) of **5**.

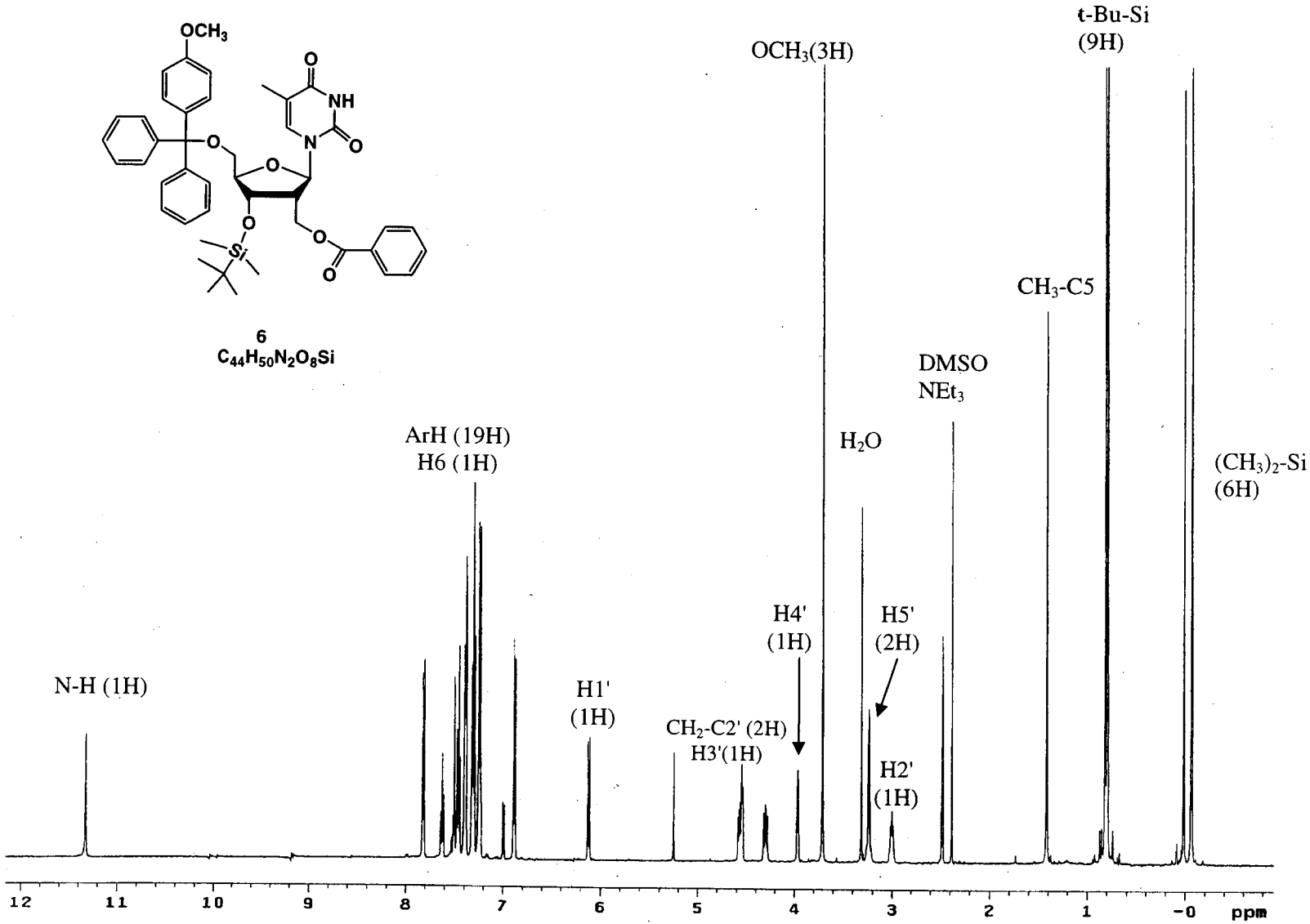


Figure S2.6a: ^{1}H NMR (500MHz, ppm, DMSO-d_6) of **6**.

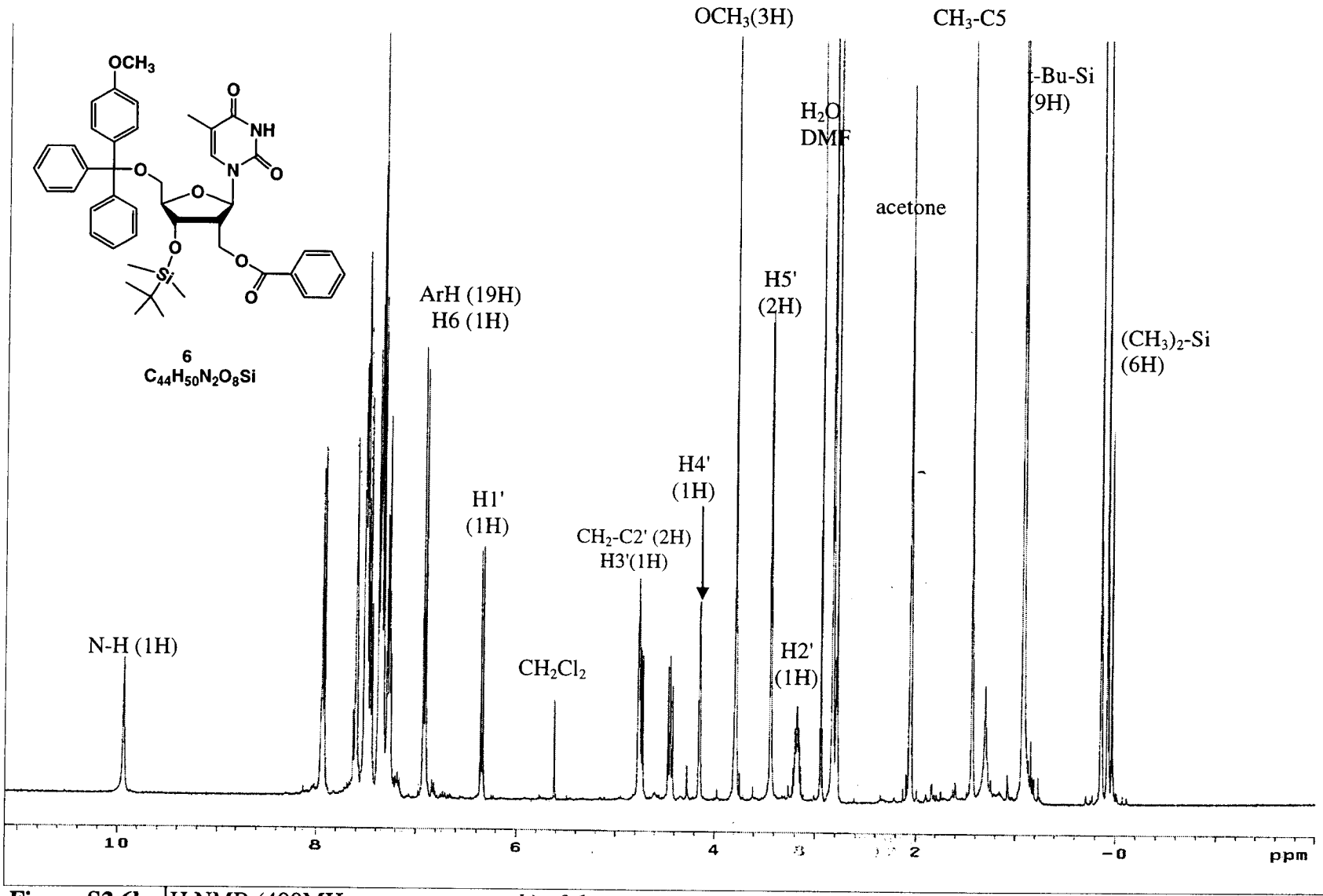


Figure S2.6b: ^{1}H NMR (400MHz, ppm, acetone-d_6) of **6**.

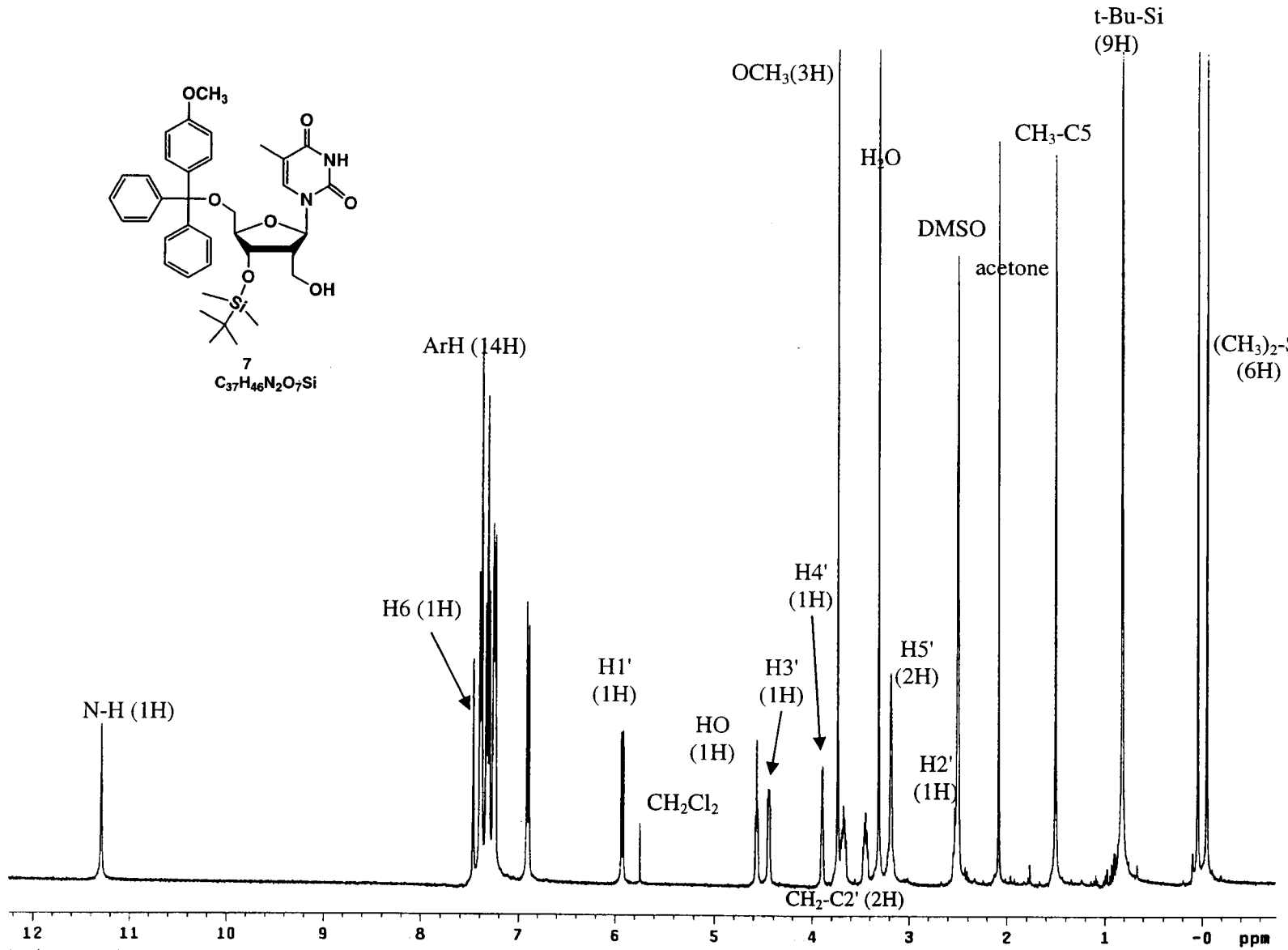


Figure S2.7: ^{1}H NMR (400MHz, ppm, DMSO-d_6) of **7**.

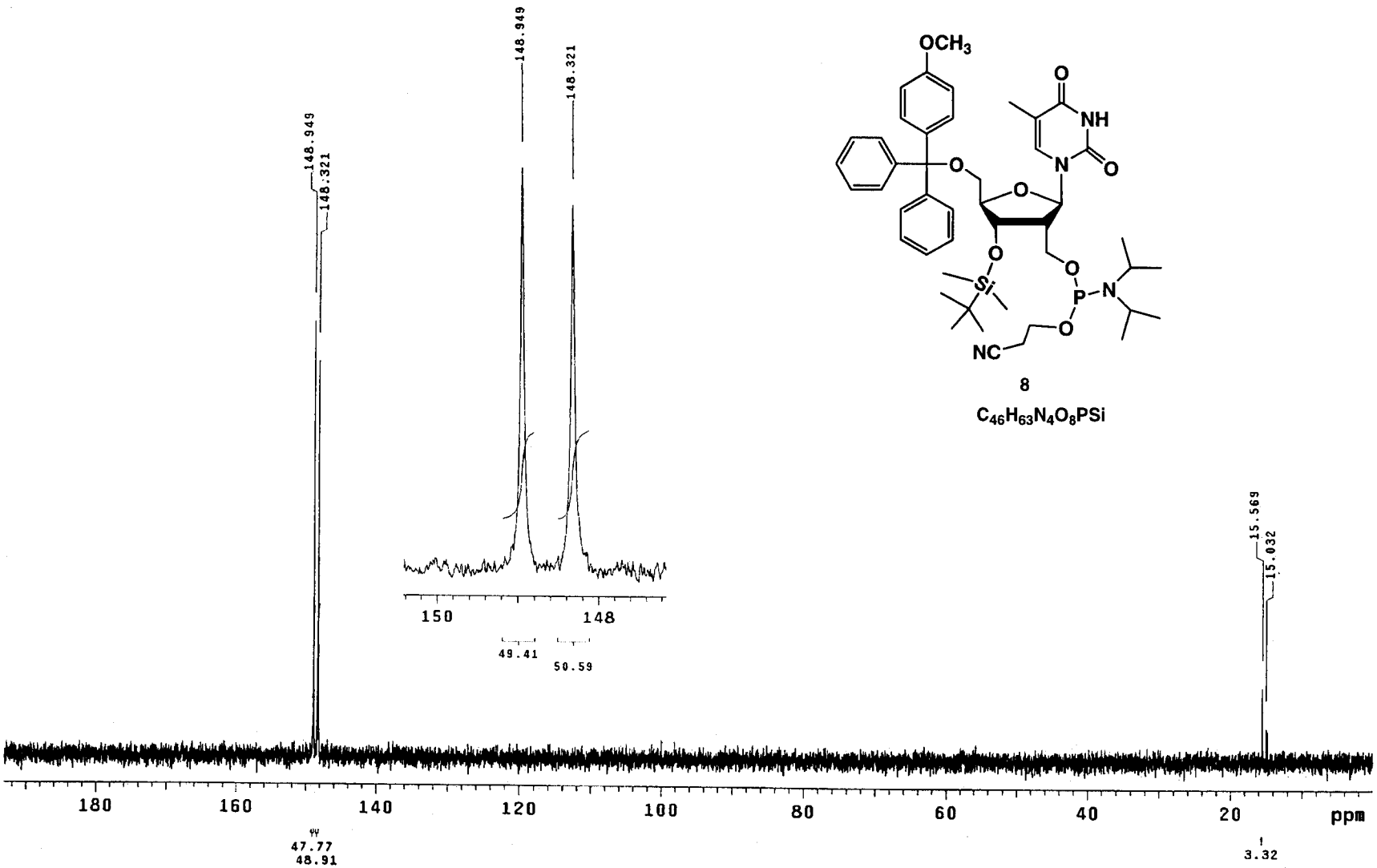


Figure S2.8a: ^{31}P NMR (200 MHz, ppm, $CDCl_3$) of **8**.

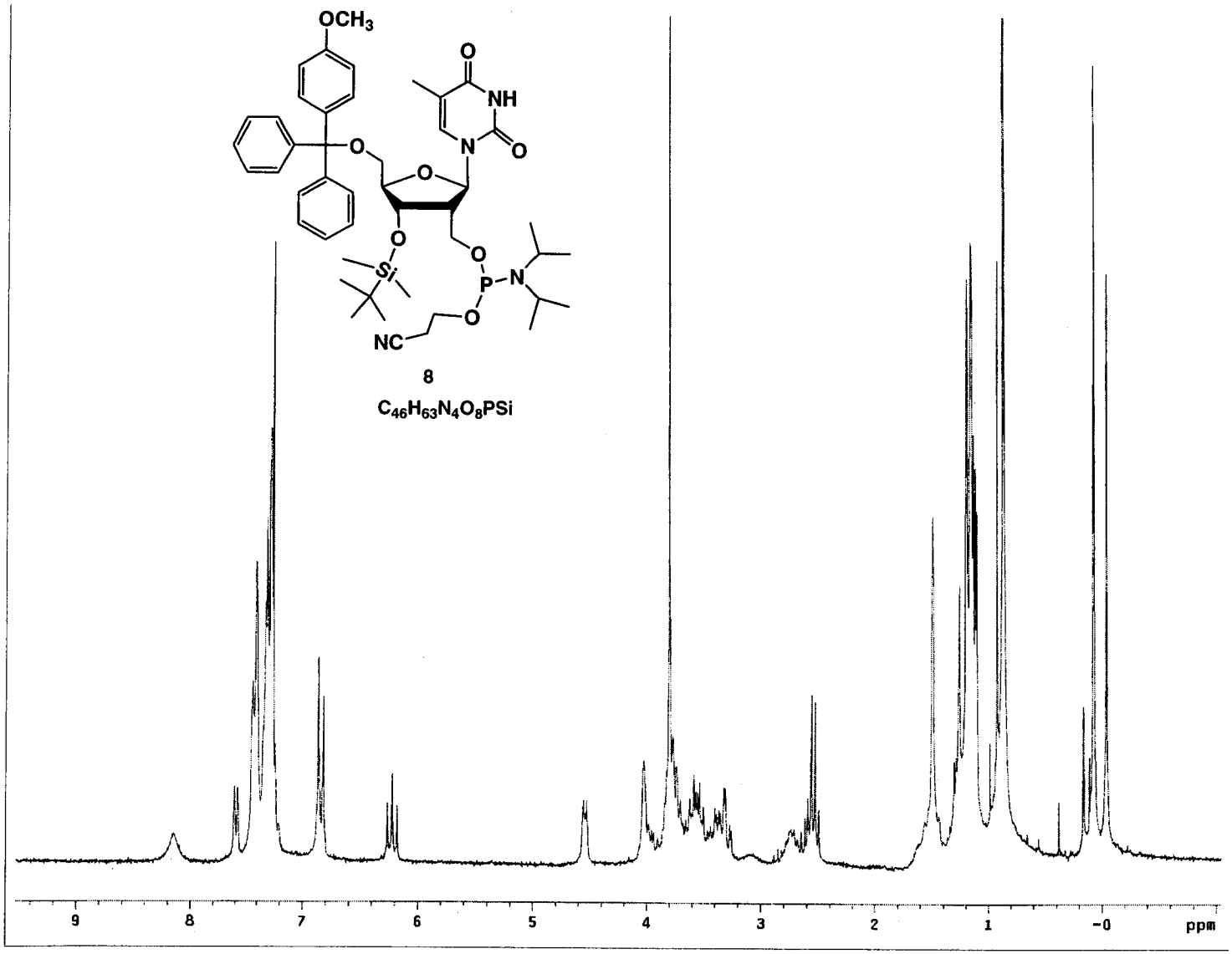


Figure S2.8b. ^{1}H NMR (200 MHz, ppm, $CDCl_3$) of **8**.

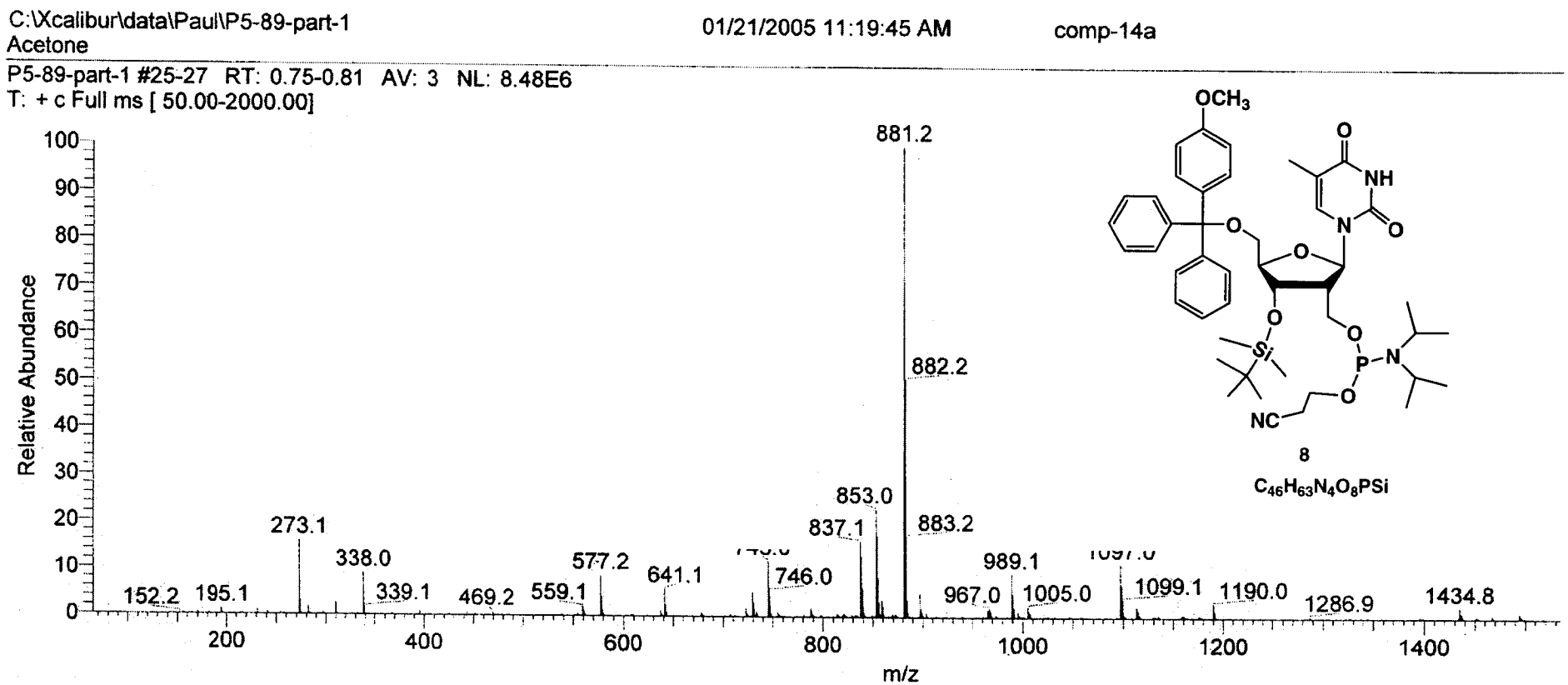


Figure S2.8c: ESI-MS for **8** $C_{46}H_{63}N_4O_8PSi$ [MNa^+] calcd. 881.42, found 881.2.

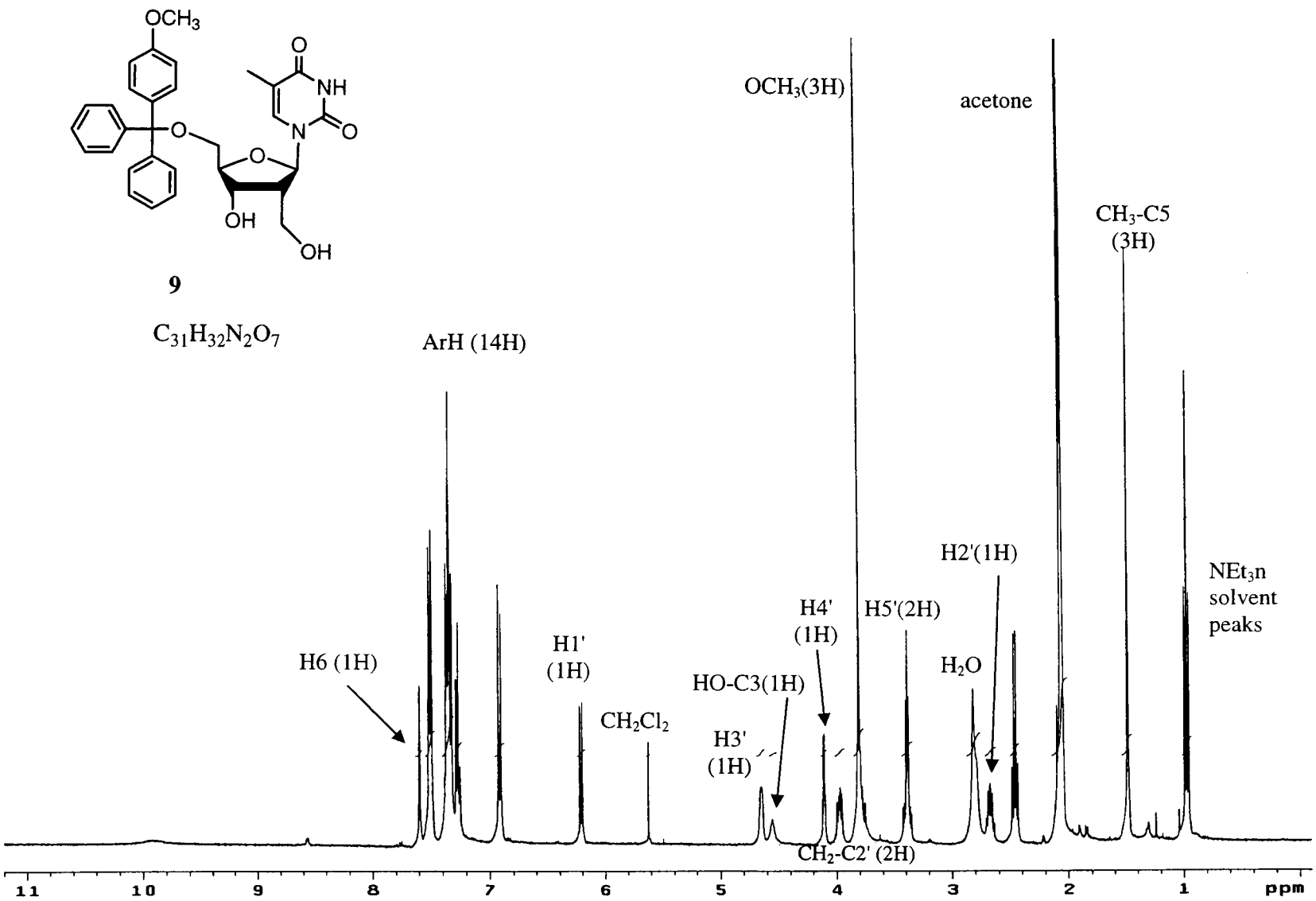


Figure S2.9a: ^{1}H NMR (400MHz, ppm, acetone-d6) of **9**.

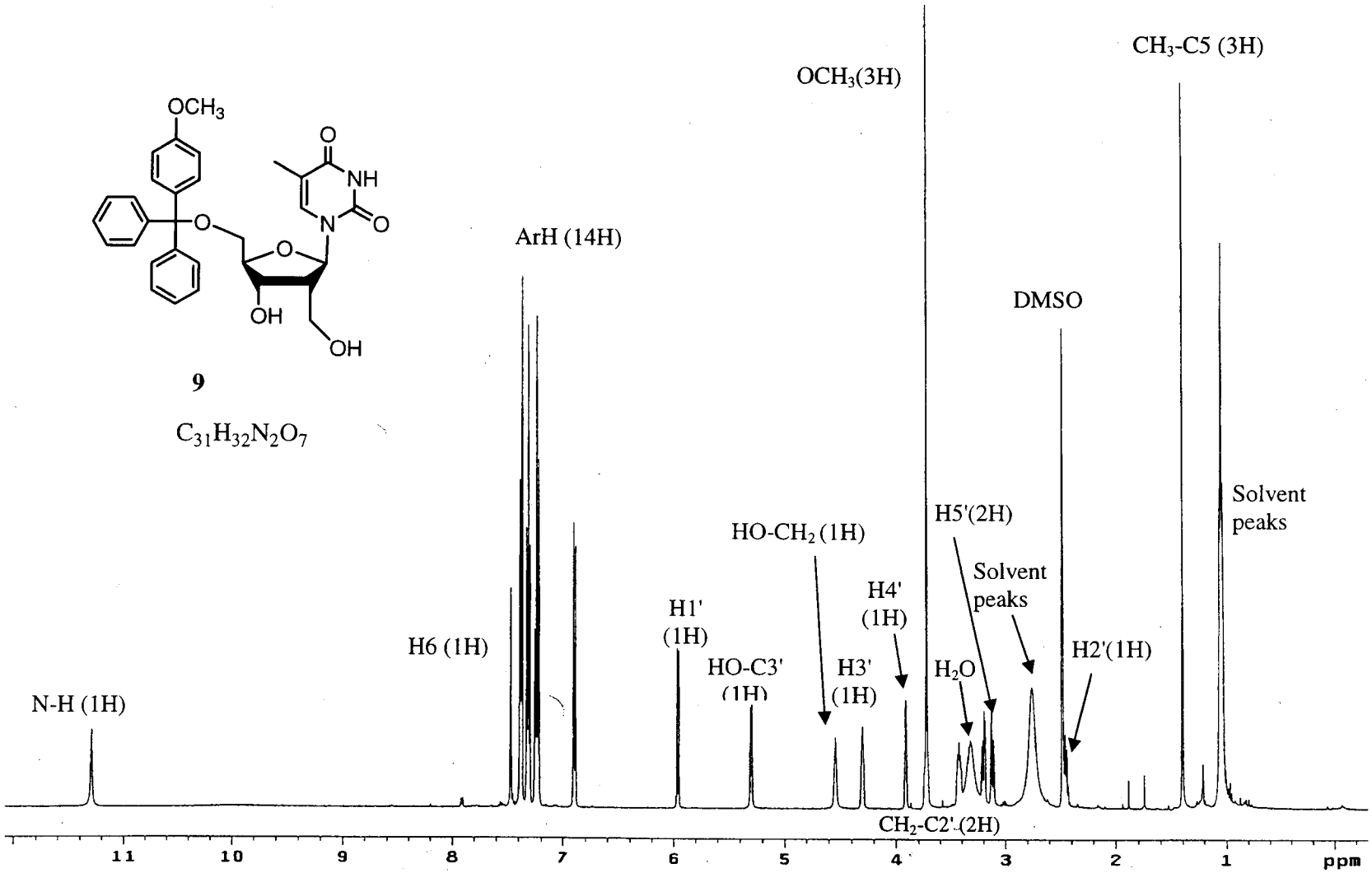


Figure S2.9b: ^{1}H NMR (500MHz, ppm, DMSO-d_6) of **9**.

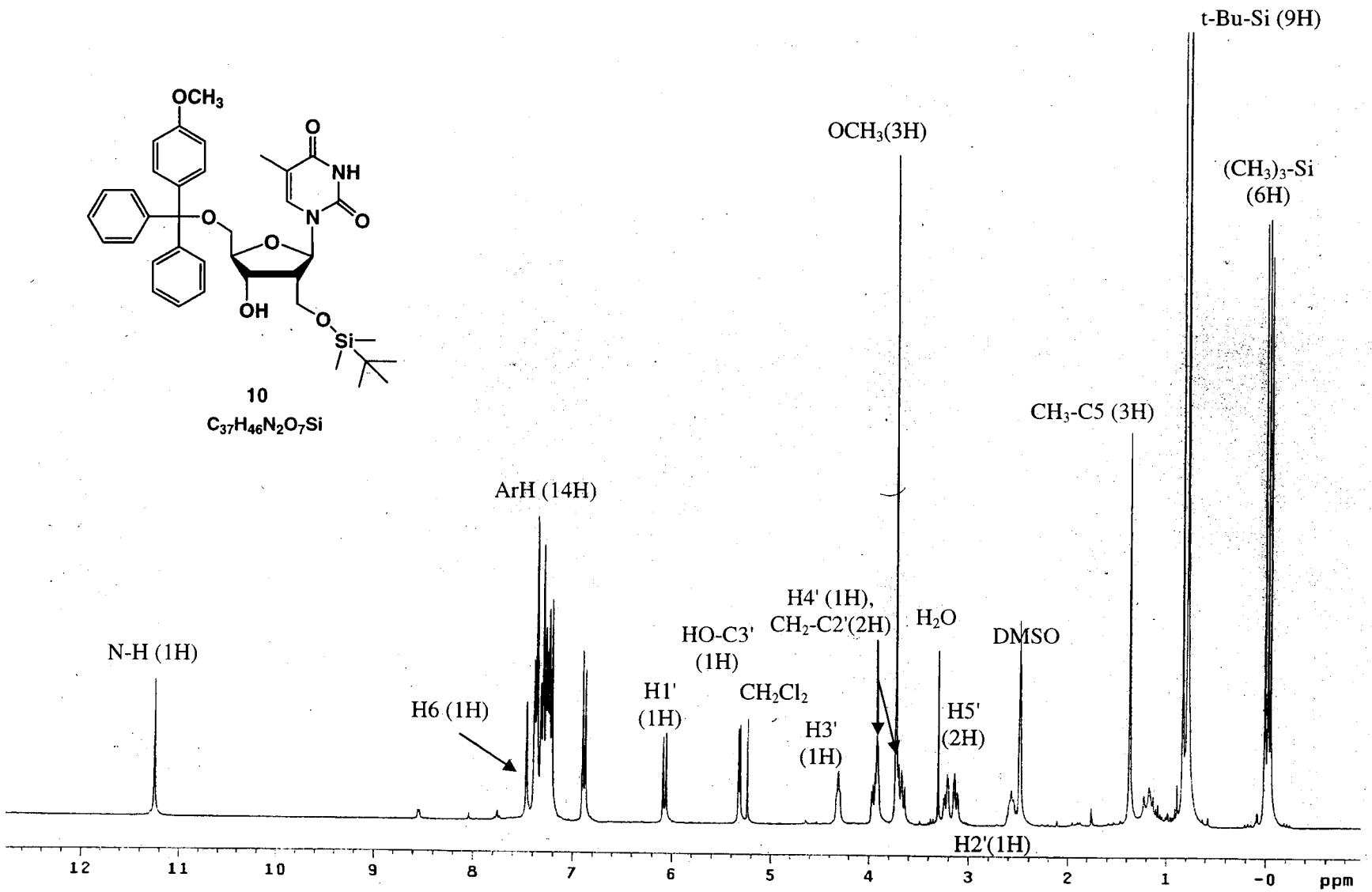
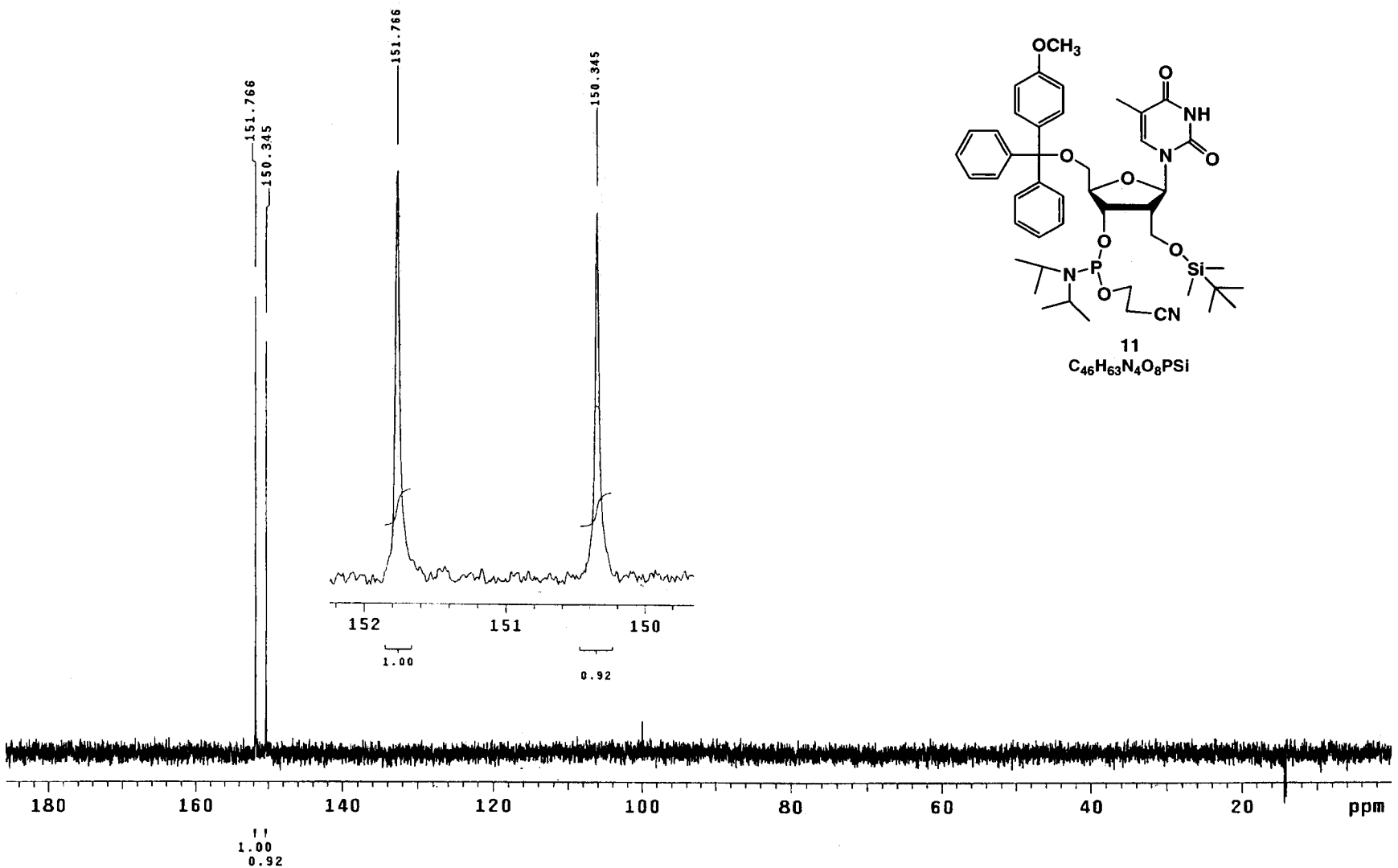


Figure S2.10a. ^{1}H NMR (300MHz, ppm, DMSO-d_6) of **10**.



11
$C_{46}H_{63}N_4O_8PSi$

Figure S2.11a: ^{31}P NMR (200 MHz, ppm, $CDCl_3$) of **11**.

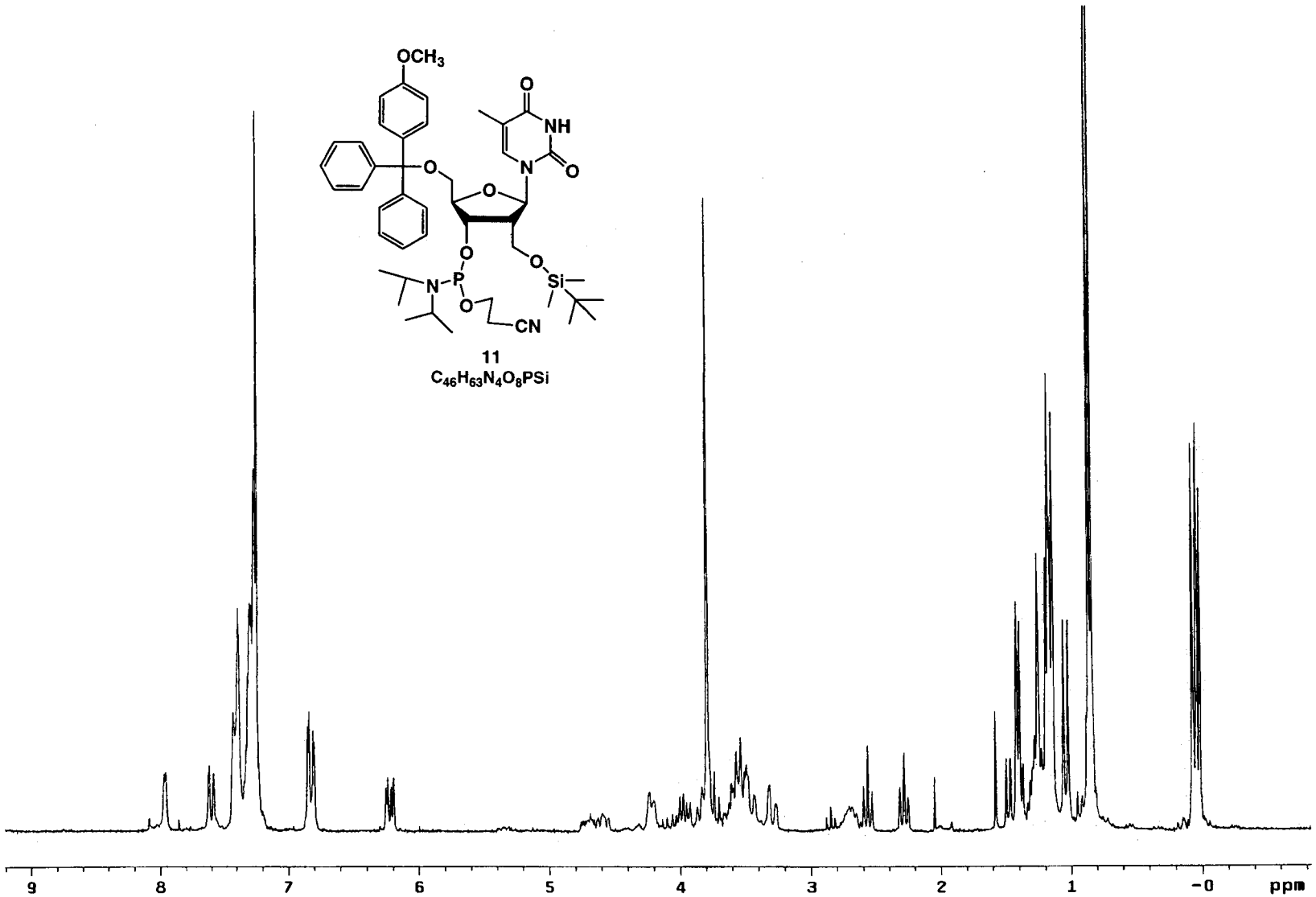


Figure S2.11b: ^{1}H NMR (200 MHz, ppm, $CDCl_3$) of **11**.

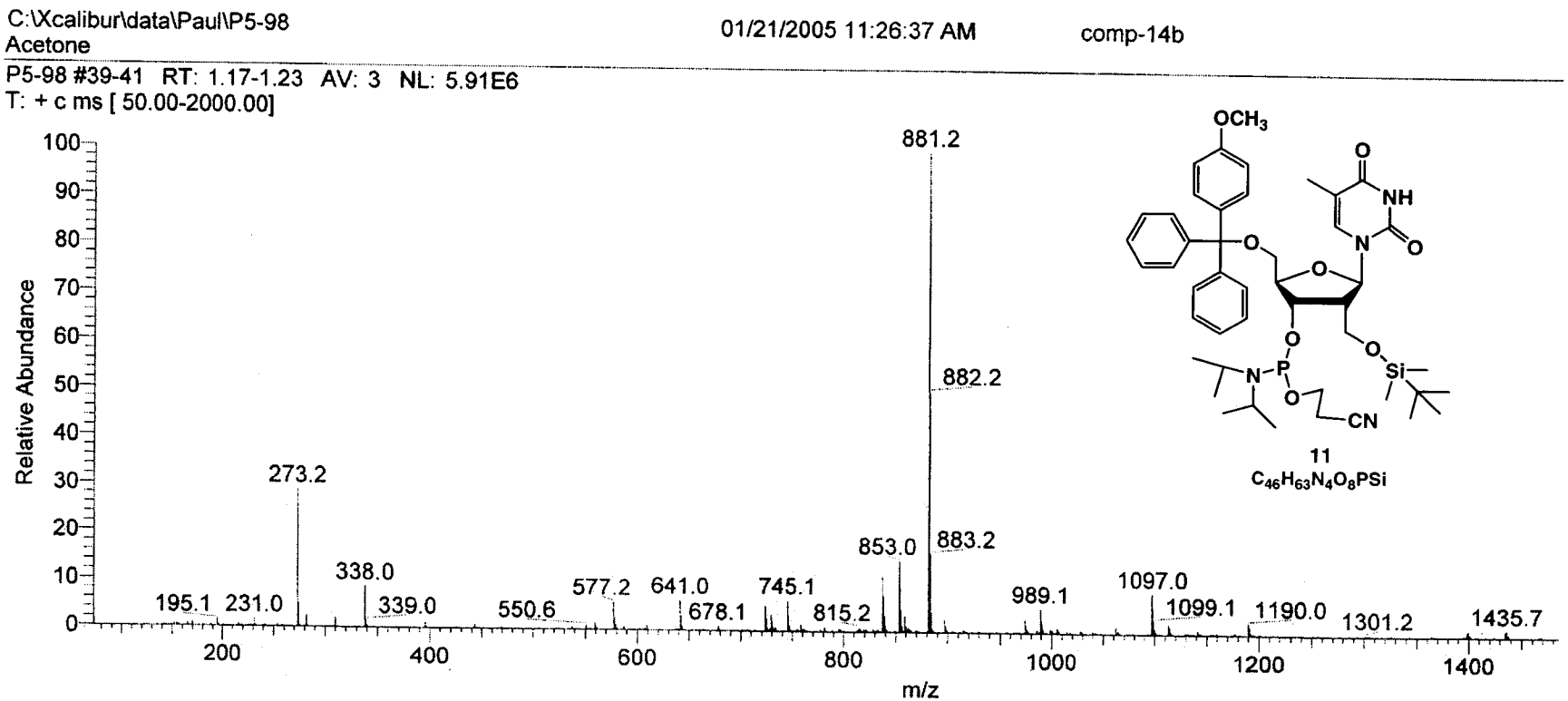


Figure S2.11c: ESI-MS for **11** $C_{46}H_{63}N_4O_8PSi$ [MNa+] calcd. 881.42, found 881.2.

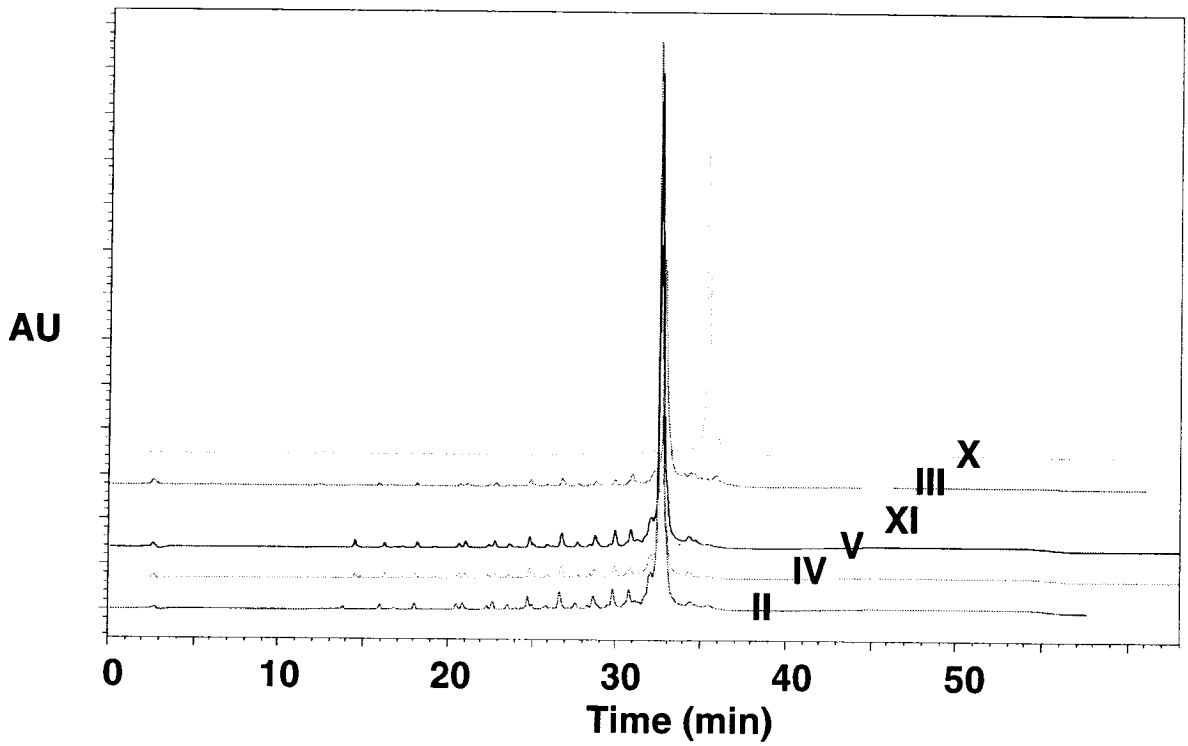


Figure S2.12: HPLC purity test of crude oligos: RNA and 2'5'-RNA comprising 2',5'-**H** or 3',5'-**H** inserts (retention time of oligo **II**, **III**, **IV**, **V**, **X** and **XI** is 32.5, 32.6, 32.6, 32.0, 35.4, 35.7 min, respectively)

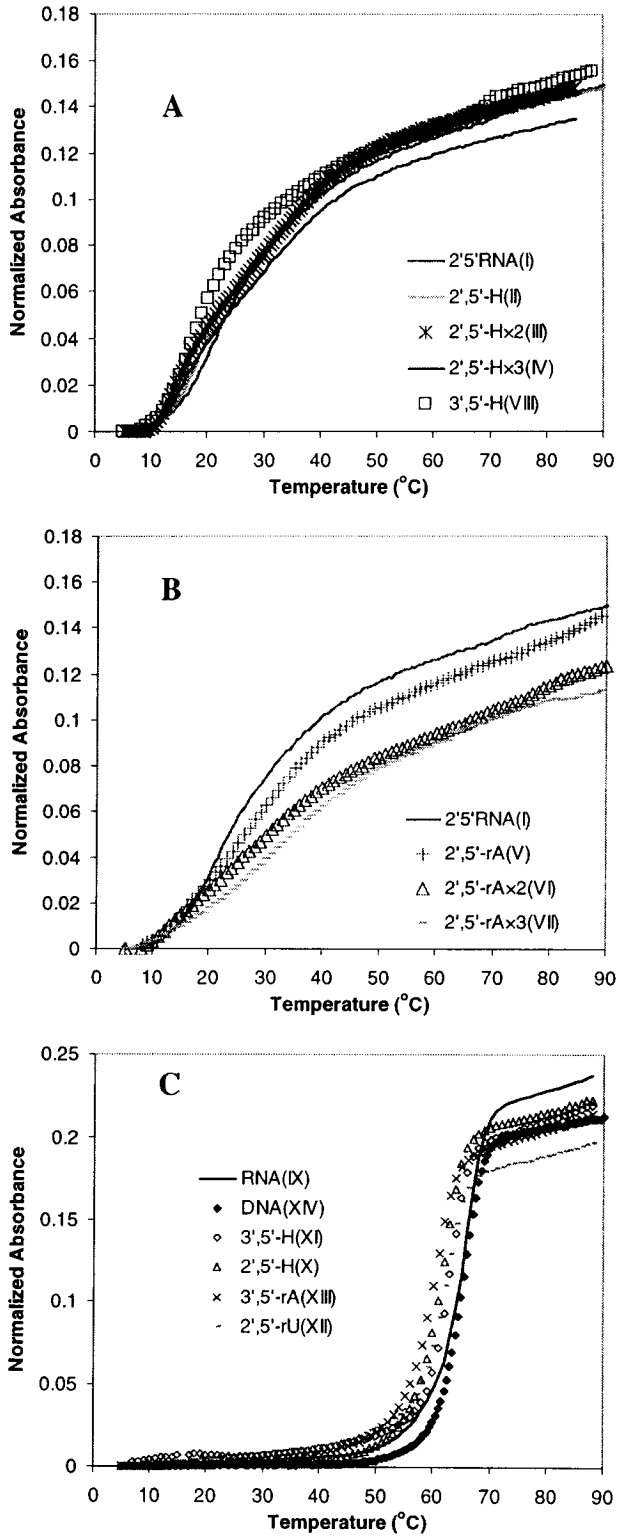


Figure S2.13: Thermal melting curves of 23-bp U5 region duplexes. Buffer: 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4, pH 7.2. Oligonucleotides were hybridized to DNA (A) 2'5'-RNA with 2',5'-**H** and 3',5'-**H** (B) 2'5'-RNA with mismatch 2',5'-rA. (C) RNA with 2',5'-**H**, 3',5'-**H**, 2',5'-rU and mismatch 3',5'-rA (See **Table 2.1**, T_m column 1)

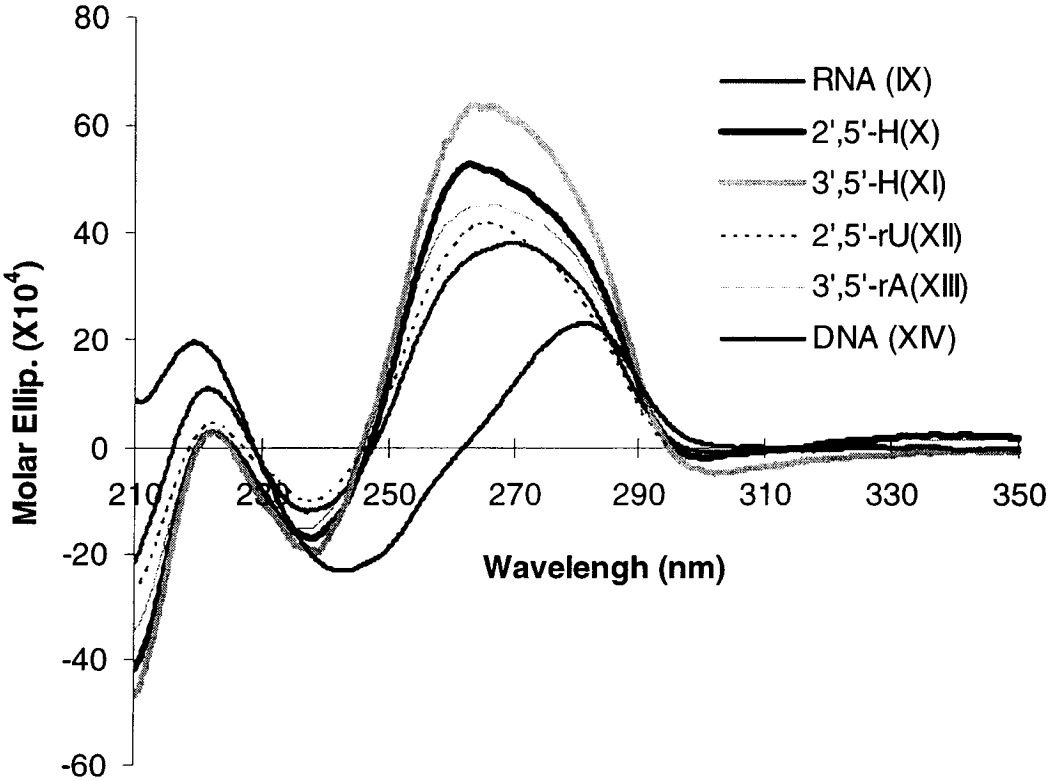


Figure 2.14: CD spectra. Buffer: 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4, pH 7.2. Oligonucleotides were hybridized to DNA. RNA with 2',5'-**H**, 3',5'-**H**, 2',5'-rU and mismatch 3',5'-rA.

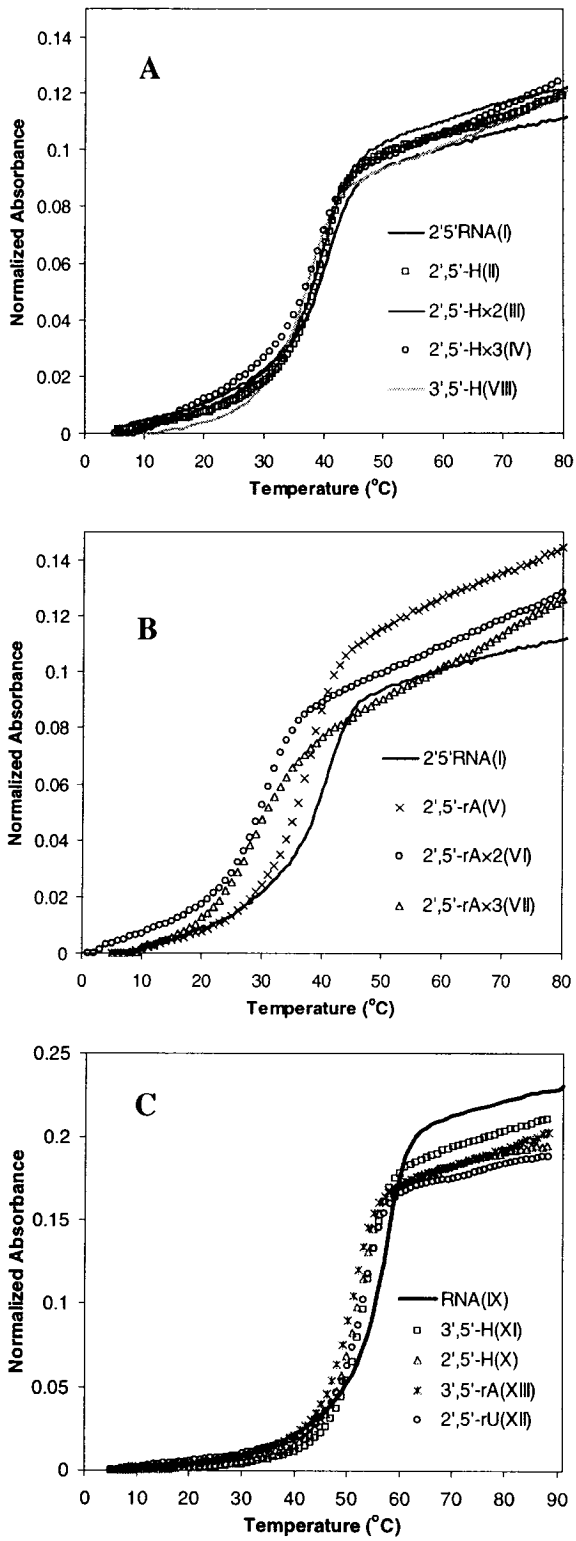


Figure 2.15: Thermal melting curves of 23-bp U5 region duplexes. Buffer: 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4, pH 7.2. Oligonucleotides were hybridized to complementary 2',5'-RNA. (A) 2'5'-RNA with 2',5'-**H** and 3',5'-**H** (B) 2'5'-RNA with mismatch 2',5'-rA. (C) RNA with 2',5'-**H**, 3',5'-**H**, 2',5'-rU and mismatch 3',5'-rA (See **Table 2.1**, T_m column 3)

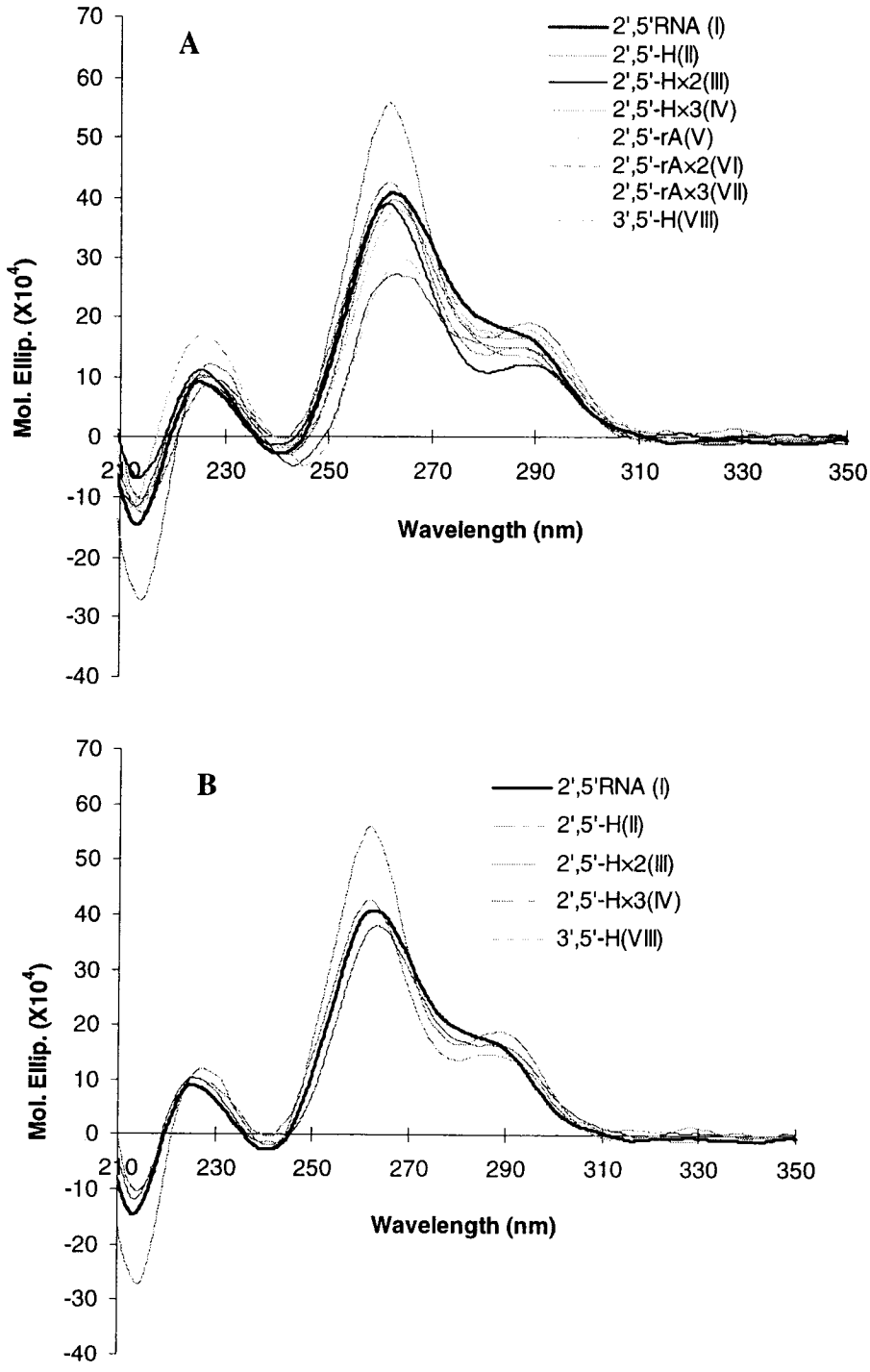


Figure 2.16: CD spectra. Buffer: 140 mM K^+, 1 mM Mg^{2+}, and 5 mM Na_2HPO_4, pH 7.2. Oligonucleotides were hybridized to complementary 2',5'-RNA. (A) 2'5'-RNA with 2',5'-**H**, 3',5'-**H** and mismatch 2',5'-rA. (B). RNA with 2',5'-**H**, 3',5'-**H**, 2',5'-rU and mismatch 3',5'-rA

Table S2.1: Mass spectrometry analysis

NO.	Designation	Oligonucleotide squence (5'->2'/3' direction)	ε (L/(mol.cm))	Calculated Mass (g/mol)	Observed Mass (g/mol)	Observed Molecular Ion
I	2,5RNA	5'-rGUC UGU UGU GUG ACU CUG GUA AC-2'	223500	7318	7319	(M-2H+Na)-
II	2',5'-**H**	5'-rGUC UGU ***H***GU GUG ACU CUG GUA AC-2'	223500	7346	7351	(M-4H+2Na+Li)-
III	2'-5'-**H**×2	5'-rGUC UG***H*** ***H***GU GUG ACU CUG GUA AC-2'	223500	7374	7377	(M-3H-2Li)-
IV	2',5'-**H**×3	5'-rGUC UG***H*** ***HH*** GUG ACU CUG GUA AC-2'	223500	7402	7397	(M-H)-
V	2',5'-r**A**	5'-rGUC UGU **A**GU GUG ACU CUG GUA AC-2'	228000	7341	7345	(M-3H+3Li)-
VI	2',5'-r**A** ×2	5'-rGUC UG**A** **A**GU GUG ACU CUG GUA AC-2'	229300	7364	7364	(M-4H+3Na)-
VII	2',5'-r**A** ×3	5'-rGUC UG**A** **A**G**A** GUG ACU CUG GUA AC-2'	232800	7387	7387	(M-2H+Li)-
VIII	3',5'-**H**	5'-rGUC UGU **H**GU GUG ACU CUG GUA AC-2'	223500	7346	7348	(M-2H+Na)-
IX	RNA	5'-rGUC UGU UGU GUG ACU CUG GUA AC-3'	223500	7318	7319	(M-2H+Na)-
X	2',5'-**H**	5'-rGUC UGU ***H***GU GUG ACU CUG GUA AC-3'	223500	7346	7343	(M-4H+Na+2Li)-
XI	3',5'-**H**	5'-rGUC UGU **H**GU GUG ACU CUG GUA AC-3'	223500	7346	7349	(M-3H+2Li)-
XII	2',5'-r**U**	5'-rGUC UGU UGU GUG ACU CUG GUA AC-3'	223500	7318	7323	(M-4H+3Na)-
XIII	3',5'-r**A**	5'-rGUC UGU **A**GU GUG ACU CUG GUA AC-3'	228000	7341	7342	(M-5H+3Na+Li)-
XIV	DNA	5'-dGTC TGT TGT GTG ACT CTG GTA AC-3'	214100	7076	7076	(M-4H+Na+Li)-
P1	RRR	5'-GGAC(UUCG)GUCC-3'[a]	111600	3787	3783	[M-H]-
P2	RR($\mathbf{H}_5$)R	5'-GGAC(***H***UCG)GUCC-3'	111600	3815	3814.2	[M-H]-
P3	RR($\mathbf{H}_6$)R	5'-GGAC(U***H***CG)GUCC-3'	111600	3815	3813.3	[M-H]-
P4	RR($\mathbf{H}_{5,6}$)R	5'-GGAC(***HH***CG)GUCC-3'	111600	3843	3842.5	[M-H]-

[a] Underlined letters are 2',5'-RNA loop (e.g. UU=$U_{2'p5'}U_{2'p}$); 2',5'-**H** is shown as ***H***; 3',5'-**H** as **H**; 2',5'-r**A** as **A**, 3',5'-r**A** as **A**; 2',5'-r**U** as **U** in the sequence.

Chapter III

The previous Chapter described the synthesis of oligonucleotide chains containing 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine units (or 2'-α-*hm*-dT). In a follow-up project, we incorporated 2'-α-*hm*-dT and other modified nucleosides (arabino and *seco*nucleosides) at the branch point of Y-shaped oligoribonucleotides with the intention of probing the substrate specificity of the yeast lariat RNA debranching enzyme (yDBR). These studies led to the identification of a potent inhibitor of yDBR, making this compound a potential good choice for future enzyme co-crystallization and X-ray analysis of the yDBR active site structure.

Chapter III. Yeast Debranching Enzyme (yDBR) Inhibition by Novel Branched RNAs

Abstract

A series of branched or "Y-shaped" RNAs related to a viral pre-mRNA splicing intermediate with a uridine branch point (**rU**) were synthesized to examine yeast debranching activity and inhibition. Each of these synthetic molecules were modified at the branch point by incorporating both natural (i.e. ribose) and non-natural nucleosides (i.e. arabinose, acyclic nucleoside, and 1-[2-deoxy-2-C-α-hydroxymethyl-β-D-ribofuranosyl]thymine (2'-α-*hm*-dT; denoted as "**H**" in this study). ^{31}P NMR of 2',3'-bis-phosphoramidites revealed an unusual splitting pattern between adjacent phosphorus nuclei in some but not all bis-phosphoramidites suggesting the presence of a through-space phosphorus-phosphorus (P-P) coupling interaction. Debranching assays show that branched ribonucleic acids (bRNAs) with ribose branch points (**rA** and **rU**) are substrates of the debranching enzyme (yDBR) whereas other modified branch points (**sU**, **H**) were not recognized. Debranching inhibition study found **H** branch point linked with cytidine have great inhibition potency (IC_{50} = 70 nM) while that with guanine connected did not show any inhibition.

Introduction

Branched ribonucleic acids (bRNAs) containing vicinal 2',5'- and 3',5'-phosphodiester linkages (Figure 3.1) were first detected in nuclear polyadenylated RNA from HeLa cells by Wallace and Edmonds,[1] and were subsequently shown by others to be mRNA splicing intermediates.[2] Two unique branched RNA configurations have been detected: circles with a tail containing a branch point (i.e., a "lariat" configuration) in monomolecular *cis*-splicing reactions,[2] or branches between two linear RNAs (i.e., a "Y"-like structure) in the case of bimolecular *trans*-splicing reactions.[3] Both in yeast and in mammalian cells, the branch point nucleotide is an adenosine (**rA**),[4,5] which can have stronger base stacking ability within the branch structure than other

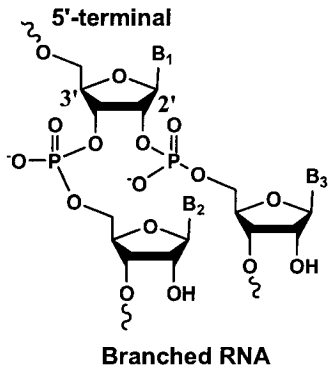


Figure 3.1. Structure of branched RNA

ribonucleosides as branch points. This could be a possible reason why all lariat RNA introns conserve **rA**.[6]

Branched RNAs are hydrolyzed by a specific enzyme which was first identified in HeLa cell extracts and named the RNA lariat debranching enzyme.[7] It is involved in the intron turnover pathway found in mammalian cells[7] and in yeast.[8,9] The yeast lariat debraching enzyme (yDBR) is a unique manganese-dependent phosphodiesterase which selectively cleaves the 2',5'-phosphodiester linkage at the branch site of either lariat or Y-shaped introns, thus converting them into linear molecules.[10] As yDBR does not digest 3',5'-phosphodiester bonds, its activity differs from snake venom phosphodiesterase, which indiscriminately cleaves both 2',5'- and 3',5'-phosphodiester bonds. The debranching efficiency of yDBR is also highly substrate-specific in that an adenosine residue at the branch point position is more rapidly hydrolyzed by yDBR than other natural counterparts,[11,12] and chemically modified branch points with different sugar conformations and phosphodiester configuration inactivate debranching activity altogether.[13] At least one 3',5'- appended nucleotide residue at the branch point is known to be required (i.e. one nucleotide connected at the 3' position) since yDBR cannot hydrolyze a 2',5'- connection when located within a linear nucleic acid substrate.[7] The nucleotide directly connecting to the branch point is also important to the debranching activity and a purine (A) is debranched more efficiently than either of the pyrimidines (C/T).[12,14] With these studies available, however, the real picture of yDBR binding with bRNAs and debranching mechanism is not known. Searching debranching inhibitors of yDBR could be useful for enzyme co-crystallization and X-ray analysis of yDBR active site structure.

Our group has had a long-standing interest in using synthetic Y-shaped bRNAs as intron mimics to study the molecular requirements of RNA splicing[15] and debranching,[10,12,13] largely aided by the development of branched oligonucleotide

synthesis methodology.[16-18] Previous studies indicate that the murine norovirus,[19] MNV-1, seems to possess bRNA structure in its genome with an uridine branch point (**rU**) which could be involved in the viral replication and translation process (Karst, Wobus, and Virgin, unpublished data). As an extension of our attempt to use synthetic bRNAs to inhibit norovirus replication (Karst, Wobus, Damha, Peng, Sabatino, and Virgin, unpublished data), we prepared a series of bRNA mimics of the proposed viral branched structure and applied them in yDBR. We report herein the first synthesis of novel bRNAs containing the 1-[2-deoxy-2-α-C-hydroxymethyl -β-D-ribofuranosyl]thymine (denoted "**H**") which has been recently reported for its incorporation into 2',5'-linked RNA,[20] 2',5'-looped RNA hairpins and siRNA (Peng and Damha, unpublished results). The debranching efficiency and inhibitory properties of these molecules on yDBR activity were investigated, and compared with other bRNAs containing natural and unnatural branch points.

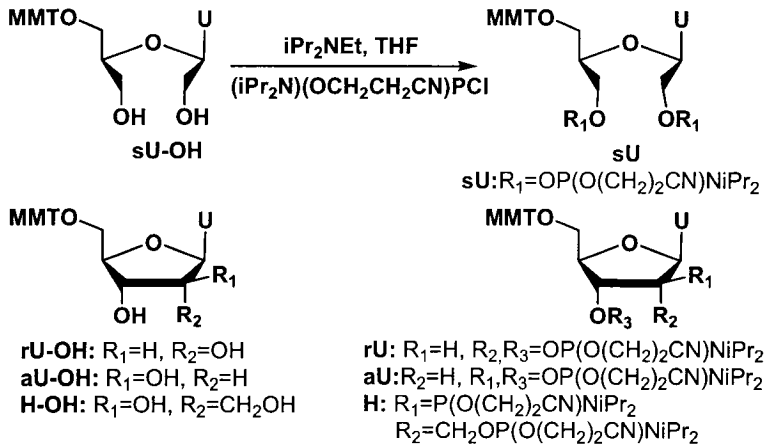
Material and Methods

General Preparation of Branch point Synthons by Phosphitylation

Bis-phosphitylation of nucleosides was conducted by a modified published procedure.[16] *N*-ethyl-*N,N*-diisopropylamine (10 eq.) was added to a solution of 5'-MMT protected nucleoside (1 eq.) containing free 2'- and 3'-OH groups (**rU**-OH, **aU**-OH, **sU**-OH, **H**-OH, Scheme 3.1) in dry THF (3 mL) under a N_2 atmosphere. The reaction was initiated by adding 3 eq. β-cyanoethyl-*N,N*-diisopropylchlorophosphor-amidite dropwise *via* syringe. A white precipitate (diisopropylethylammonium hydrochloride) usually appeared within five minutes indicating a progressing reaction. The reaction was monitored by TLC until complete conversion. After 0.5-2 h, the reaction mixture was washed with 50 mL ethyl acetate (pre-washed with 5% $NaHCO_3$), then with saturated NaCl and dried over sodium sulfate. The crude product was concentrated to dryness and purified via silica gel column chromatography with 1% NEt_3 in CH_2Cl_2/MeOH (100:1 to 20:1 v/v) to afford the corresponding bis-phosphoramidites (**rU**, **aU**, **sU** and **H**) in 50-87% yield.

1-[2,3-*O*-bis-(β-cyanoethyl-*N,N*-diisopropylphosphoramidite)-5-*O*-(4-meth

oxytrityl)-β-D-ribofuranosyl]uracil (rU). 87% isolated yield; R_f (SiO_2) = 0.55 - 0.39 (1:1 Ether/CH_2Cl_2); ^{31}P NMR (200 MHz, ppm, $CDCl_3$): δ 152.41, 152.35, 151.77, 151.70; 151.46, 151.43; 151.39, 151.35, 151.24, 151.20, 151.13, 151.0, 150.89 (13 peaks). ESI-MS (in acetone) for $C_{47}H_{62}N_6O_9P_2$ [MNa+] calcd. 939.98, found 939.3.



Scheme 3.1. Phosphitylation reactions

1-[2,3-*O*-bis-(β-cyanoethyl-*N,N*-diisopropylphosphoramidite)-5-*O*-(4-meth oxytrityl)-β-D-arabinofuranosyl]uracil (aU). 75% isolated yield; R_f (SiO_2) = 0.68 - 0.48 (1:1 Ether/CH_2Cl_2); ^{31}P NMR (200 MHz, $CDCl_3$): δ 152.32, 152.27, 152.05, 151.98, 151.92, 151.85, 151.28, 151.22, 151.18, 151.15, 151.51, 150.44, 150.16. ESI-MS (in acetone) for $C_{47}H_{62}N_6O_9P_2$ [MNa^+] calcd. 939.98, found 939.3.

1-[2,3-*O*-bis-(β-cyanoethyl-*N,N*-diisopropylphosphoramidite)-5-*O*-(4-meth oxytrityl)-2,3-seco)-β-D-ribofuranosyl]uracil (sU). sU-OH (Scheme 1) was obtained by a published procedure.[21] 50% isolated yield; R_f (SiO_2) = 0.71 – 0.63 (4:1:1 Ether/CH_2Cl_2/hexane); ^{31}P NMR (200 MHz, $CDCl_3$): δ 150.39, 150.34, 150.19, 149.86, 149.84, 149.58, 149.53, 149.11. ESI-MS for $C_{49}H_{66}N_4O_9P_2$ [MNa^+] calcd. 940.02, found 941.3.

1-[2-α-*C*-(β-Cyanoethyl-*N,N*-diisopropylphosphoramidic)-hydroxymethyl -3-*O*-(β-Cyanoethyl-*N,N*-diisopropylphosphoramidite)-2-deoxy-5-*O*-(4-methoxyt

rityl)-β-D-ribofuranosyl]thymine (H). **H**-OH (Scheme 3.1) was obtained by a published procedure.[20] 77% isolated yield; R_f (SiO_2) = 0.63 -0.57 (50:45:5 EtOAc/hexane/NEt_3); ^{31}P NMR (200 MHz, $CDCl_3$): δ 151.88, 150.47, 150.34, 148.80, 148.67, 148.06, 147.97. ESI MS (in acetone) for $C_{49}H_{66}N_6O_9P_2$ [MNa^+] calcd. 968.03, found 967.4.

N^6-benzoyl-9-[2,3-*O*-bis-(β-cyanoethyl-*N*,*N*-diisopropylphosphoramidite)-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]adenine (rA) was prepared according to a published procedure.[18] 75% isolated yield; R_f (SiO_2) = 0.71 - 0.57 (1:1 Ether/CH_2Cl_2); 31P NMR;[18] ESI MS (in acetone) for $C_{56}H_{69}N_9O_9P_2$ [M^+] calcd. 1074.15, found 1074.1.

Branched RNA Synthesis

Branched oligonucleotides were synthesized on 500 Å controlled pore glass solid support using an ABI 381A DNA synthesizer (1 μmol scale) according to our well-established convergent solid phase synthetic methodology.[18,22] Briefly, an inverted thymidine (3'-DMT-dT) was first attached to the solid support (90 μmol/g), which is shown to increase nuclease resistance. The introduction of the branch point nucleoside was achieved by coupling of two adjacent CPG-bound nucleotide chains with different 2',3'-bis-phosphoramidites. The coupling conditions and times of branch point addition during solid phase synthesis were identical to those for bRNA synthesis[18] and employed a 0.03 M 2',3'-bis-phosphoramidite concentration, 30 min coupling time and 0.5 M DCI in acetonitrile as coupling activator. The Y-shaped constructs were purified by either 24% denaturing PAGE or anion-exchange HPLC (Protein Pak DEAE-5PW column-Waters; 7.5 mm x 7.5 cm) with proper linear RNA controls. The bRNA products were also analyzed by negative-mode MALDI-TOF MS (Kratos Kompact-III instrument; Kratos Analytical Inc., New York) (Table 3.2).[23,24]

5'-End Labeling of bRNAs

5'-End ^{32}P labeling of bRNAs was based on a published procedure.[25] Briefly, bRNAs were radiolabeled at the 5'-hydroxyl terminus with a radioactive ^{32}P probe and the

enzyme T4 polynucleotide kinase (T4 PNK; MBI Fermentas Life Sciences, Burlington, ON). Incorporation of the ^{32}P label was accomplished in reaction mixtures consisting of branched nucleic acid substrate (100 pmol), 2 μL 10x reaction buffer (500 mM Tris-HCl, pH 7.6 at 25°C, 100 mM $MgCl_2$, 50 mM DTT, 1 mM spermidine and 1 mM EDTA), 1 μL T4 PNK enzyme solution (10 U/1μL in a solution of 20 mM Tris-HCl, pH 7.5, 25 mM KCl, 0.1 mM EDTA, 2 mM DTT and 50% glycerol), 6 μL [γ-^{32}P]-ATP solution (6000 Ci/mmol, 10 mCi/ml; Amersham Biosciences, Inc.) and sterile water to a final volume of 20 μL. The reaction mixture was incubated for about 30-45 min at 37°C, quenched by heating (95°C, 5 min) and purified by 16% denaturing polyacrylamide gel (20%, 7 M urea). The pure labeled product was kept at -20°C for future use.

Debranching Activity Assay

The debranching assay was conducted at room temperature (~ 22°C) with 3–10 pmol of radiolabeled bRNAs, 1-6 μL of yDBR enzyme (28 mg/mL) in Tris-HCl buffer (500 mM, pH 7.4) containing 10% glycerol, 500 mM NaCl, 5 mM DTT, 1 mM EDTA); 2 μL optimized 10x debranching buffer (500 mM Tris-HCl, pH 7.0, 20 mM DTT, 5 mM $MnCl_2$); and autoclaved water to a final volume of 20 μL. Different time points (5, 10, 20, 30, 60 min) were obtained by taking 4 μL aliquots from the reaction mixture and quenching the aliquot by the same volume of a stopping dye solution (98% deionized formamide, 10 mM EDTA, 1 mg/mL bromophenol blue and 1 mg/mL xylene cyanol). The product pattern from each time point was analyzed by 16% denaturing polyacrylamide gel electrophoresis (PAGE) and subsequent autoradiography analysis.

Debranching Inhibition Assay

The debranching inhibition by bRNAs was performed at room temperature (~ 22°C) in the presence of cold bRNA inhibitors (0.01-20 μM), 0.5 μL 10× debranching buffer (500 mM Tris-HCl, pH 7.0, 20 mM DTT, 5 mM $MnCl_2$), the radiolabeled **rAG** substrate (Table 3.2, entry **4**; 0.9 pmol) and sterile water. The reaction was initiated by the addition of 0.6 μL yDBR enzyme to a final volume of 5 μL. The reaction was

maintained for 30 min and quenched by a 5 μL stopping dye solution (98% deionized formamide, 10 mM EDTA, 1 mg/ml bromophenol blue and 1 mg/ml xylene cyanol). The product pattern from each inhibitor concentration was analyzed by 16% denaturing PAGE and subsequent autoradiography analysis.

Results and Discussion

Synthesis of Branch Points

Bis-phosphitylation of nucleosides[16] with excessive phosphitylation reagent usually worked very well to afford moderate to high yields (Scheme 3.1). After phosphitylation, two new chiral centers were generated to give four diastereomers. Consistent with a previous study,[13] the ^{31}P NMR spectra of both ribouridine (**rU**) and arabinouridine (**aU**) show 14 splitting peaks, indicating a through-space P-P coupling. This coupling should give 16 splittings but we could not observe all the peaks due to overlapping signals (Table 3.1 and Figure S3.1). It was previously shown that the xyloadenosine (**xA**) bis-phosphoramidite did not show this coupling interaction (7 peaks observed) for a *trans*-diaxial arrangement of the 2',3'-phosphite groups in a C3'-*endo* conformation which places them sufficiently far from each other and abolishes any through space effect.[13] Previous modeling shows that **aA** branch point adopts C3'-*endo* or O4'-*endo* usual conformation which could allow closer P-P distance for coupling since both phosphorus atoms are not in the trans-diaxial arrangement.[13] Our result of **aU** bisamidite is consistent with **aA** (Table 3.1 and Figure 3.2). The **H** bisamidite has an extra methylene group at the 2'-sugar position relative to the **rU** counterpart and thus accommodates larger P-P distance which tend to be larger to reduce the P-P steric interaction. In ^{31}P NMR we only observed 8 peaks corresponding to the fact that through-space P-P coupling was not observed in **H** (Table 3.1 and Figure S3.4). The acyclic nucleotide (**sU**) bisamidite also allows more space for these two phosphorus atoms and for this reason a through-space P-P coupling was not observed (Table 3.1 and Figure S3.3).

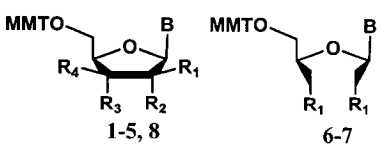
Debranching Activity Assay

The solid-phase synthesis of bRNAs with a convergent method[18,22] offers different yields among which bRNA with **aU** branch point shows the worst coupling efficiency

Table 3.1: Proton-decoupled ^{31}P-NMR splitting patterns

No.	Compound[a]	# of peaks[b]	# of coupled nulcei	Major conformer
1	Bis-ribo-A (**rA**)[13] B = A, R_1, R_4 = H R_2, R_3 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	16	8	C2'-*endo*[13]
2	Bis-ribo-U (**rU**) B = U, R_1, R_4 = H R_2, R_3 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	>14[c]	8	
3	Bis-arabino-A (**aA**)[13] B = A, R_2, R_4= H R_1, R_3 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	16	8	O4'-*endo*[13] C3'-*endo*
4	Bis-arabino-U (**aU**) B = U, R_2, R_4= H R_1, R_3 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	>14[c]	8	
5	Bis-xylo-A (**xA**)[13] B = A, R_1, R_3= H R_2, R_4 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	7	0	C3'-*endo*[13]
6	Bis-seco-A (**sA**) [13] B = A; R_1 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	6	0	N/A
7	Bis-seco-U (**sU**) B = U; R_1 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$	8	0	N/A
8	Bis-hydroxymethyl-T (**H**) B = T, R_1 = H, R_2 = CH$_2$OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$, R_3 = OP(O(CH$_2$)$_2$CN)N*i*Pr$_2$,	8	0	N/A

a



b ^{31}P-NMR on 200 MHz with external reference H_3PO_4.

c Overlapping peaks (see Figure 3.2 and Figure S3.1, Supplementary Data)

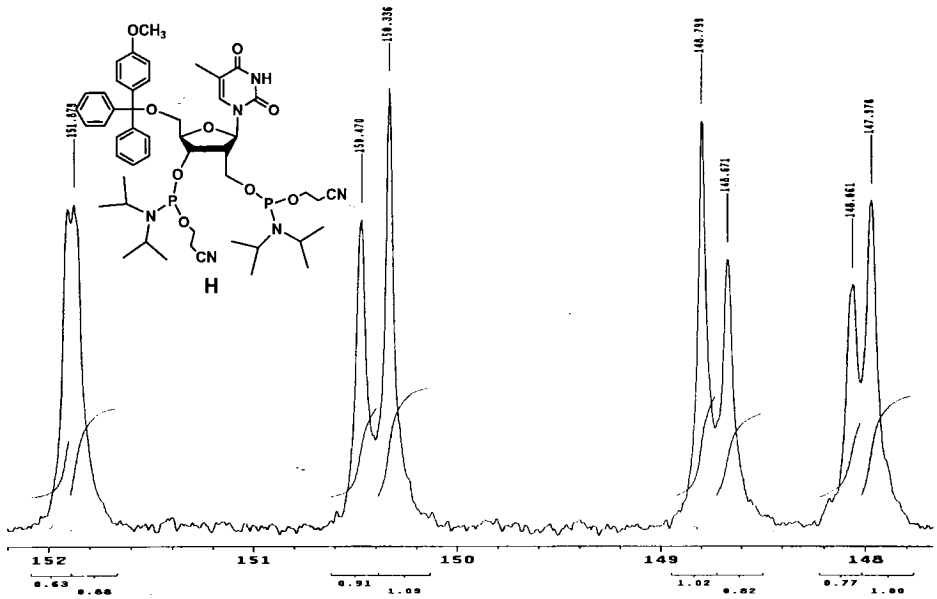


Figure 3.2: ^{31}P NMR of 1-[2-α-*C*-(β-Cyanoethyl-*N,N*- diisopropylphosphoramidic)-hydroxymethyl)3-*O*-(β-Cyanoethyl-*N,N*-diisopropylphosphoramidite)-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (**H**). ^{31}P NMR (200 MHz, $CDCl_3$): δ 151.88, 150.47, 150.34, 148.80, 148.67, 148.06, 147.97 (8 peaks)

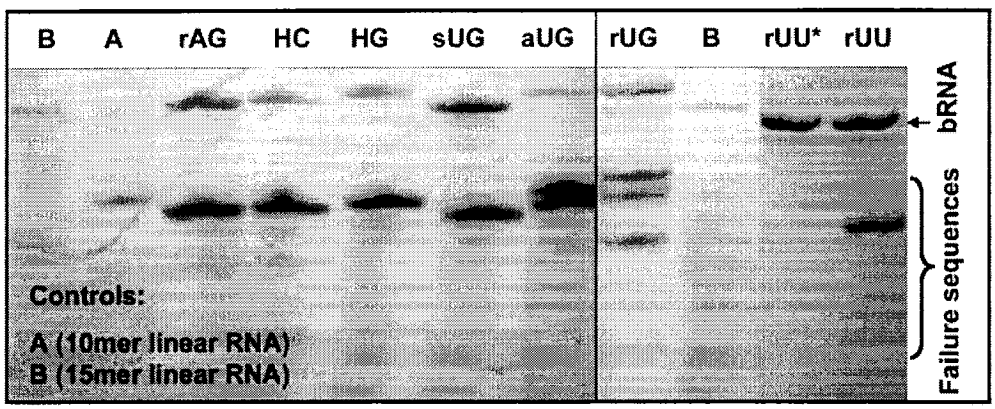


Figure 3.3: Denaturing PAGE analysis of bRNA products (crude) (**rUU***: purified)

Table 3.2: Branched RNA and linear controls synthesized in this study

No.	Code	Branched RNAs and controls[a]	bRNA purity (%) and retention time (min) in HPLC	M.W. (calcd.)	M.W. (MALDI)
1	rUG	GAUCrU(GUGAt)(GUGAt)	30 (33.84)	4774	4772
2	aUG	GAUCaU(GUGAt)(GUGAt)	6 (34.37)	4774	4773
3	sUG	GAUCsU(GUGAt)(GUGAt)	35 (34.10)	4776	4776
4	rAG	GAUCrA(GUGAt)(GUGAt)	25 (34.19)	4797	4796
5	HC	GAUCH(CGGGt)(CGGGt)	17 (33.90)	4832	4832
6	HG	GAUCH(GUGAt)(GUGAt)	16 (32.72)	4802	4800
7	rUU	UUUUrU(UUUUt)(UUUUt)	52 (31.45)	4512.3	4513
8	A	5'-GUUCUGUGAt	60 (30.72)	3149	3151
9	B	5'-GUUCUCGGGGUGAUt	18 (34.07)	4791	4889

[a] 3' termini are capped with an inverted thymidine (3'-3'-dT)

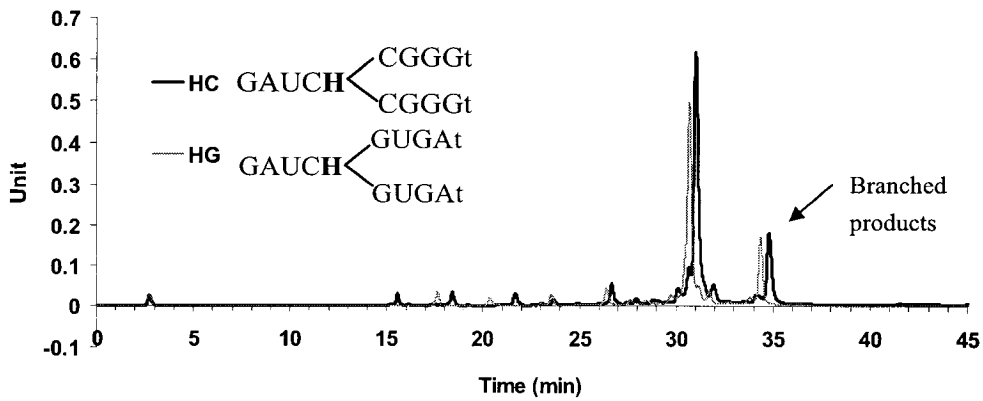


Figure 3.4: HPLC analysis of crude **HC** and **HG**

and **rU** the best yield (Table 3.2). The characteristic slow moving bands shown in Figure 3.3 correspond to Y-shaped bRNA products, compared with linear RNA controls. HPLC also show bRNAs have a longer retention time (Figure 3.4 and Figure S3.5 in Supplementary Data). MALDI experiments confirmed that we obtained the branched oligonucleotides (Table 3.2).

Debranching assays were first conducted to evaluate the ability of yDBR to debranch these bRNAs. As expected, the wild-type **rAG** shows the best debranching activity whereas **rUG**, while also serving as a substrate for yDBR, is debranched much less efficiently based on previous work.[14] After 10 min, **rAG** is completely debranched to generate the detectable linear product as indicated by a linear control, while *ca.* 50% **rUG** still remains intact after 60 min in the same debranching conditions (Figure 3.5). Doubling the amount of yDBR in the reaction

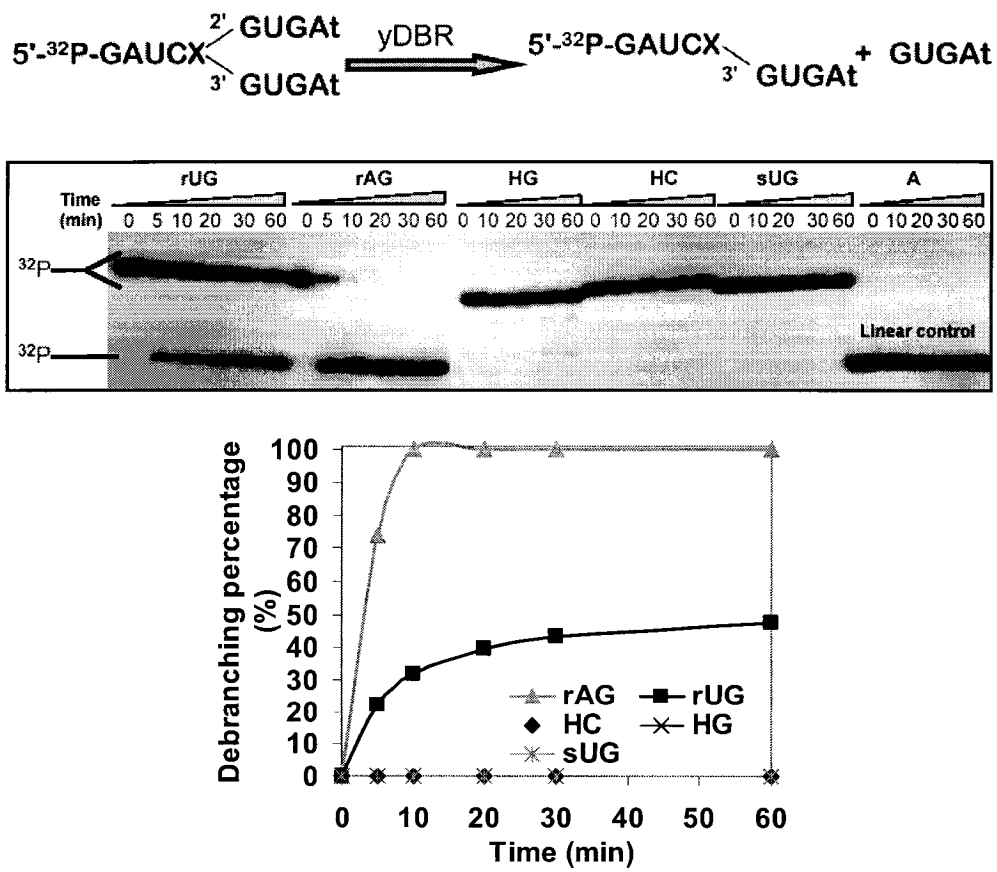


Figure 3.5: Debranching activity of different branched RNAs

medium (2 μL), resulted in complete degradation of **rUG** after 5 min (data not shown). Other non-natural branch points (**sUG**, **HC**, **HG**) did not show any debranching activity indicating they are not a substrate of yDBR, consistent with a previous study.[13] The linear control was not a substrate of yDBR either (Figure 3.5).

Debranching inhibition assay

We next evaluated whether branch point-modified bRNAs found to be inactive towards yDBR branch point hydrolysis could instead serve to inhibit the enzyme's debranching activity toward the wild-type **rAG** substrate. Increasing concentrations (0.01-20 μM) of **sUG**, **HC** and **HG** were added in a debranching reaction with **rAG** and yDBR (Figure 3.6). Of the inhibitors tested, **sUG**, **HG** and the linear control did not show any concentration-dependent inhibition pattern. Only **HC** exhibited excellent debranching inhibition with an IC_{50} of 70 nM. Previous studies also indicate the debranching inhibition by ara-A branched RNA to have an IC_{50} of 450 nM.[13] We currently find **HC** to exhibit much greater potency in debranching inhibition. Interestingly, we also find that while bRNA with **H** linked to cytidine at 2' (i.e., **HC**, Table 3.2 and Figure 3.6) competitively inhibits debranching, **H** linked to guanosine at the 2' position (i.e. **HG**) does not, indicating the importance of the environment around the branch point for enzymatic binding.[14] Early NMR study of branched trinucleotides by Chattopadhyaya and coworkers indicates that the lariat RNA branch point adenosine prefers an unusual C2'-*endo* (*south*) conformation which reduces the charge repulsion of 2',5' and 3',5'-linked phosphate groups and is stabilized by base

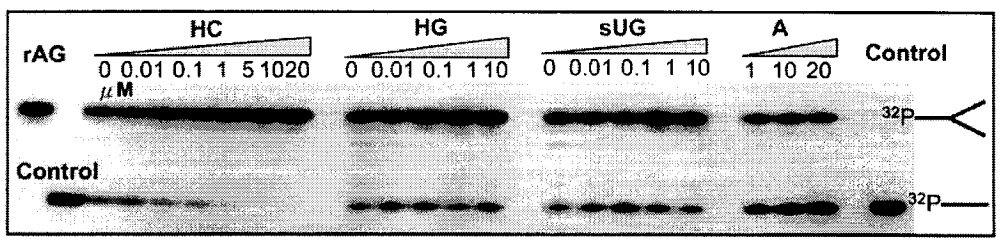


Figure 3.6: Debranching inhibition of **HC**, **HG**, **sUG** and linear RNA control (**A**).

stacking within the branched RNA.[6] The unusual sugar pucker could be a specific conformation for enzymatic recognition.[6] This *south* conformation should be more favored in branch point **H** due to 2' alkyl group[20] and extended 2'-phosphate linkage allow larger P-P distance so as to reduce the P-P repulsion. It is hard to justify based on the present data why cytidine is more favored to inhibit yDBR debranching than guanosine, but again it suggests that yDBR recognition is very sensitive to not only the branch point but also the environment of branch point (i.e. bases connecting to the branch point). Therefore, **HC** could be an interesting analogue for future attempts at co-crystallization of yDBR.

Acknowledgements

This work was supported through a grant from the National Science and Engineering Council (NSERC) of Canada. CGP acknowledges support from a FQRNT fellowship and a Clifford Wong McGill Major Fellowship. We also thank Dr. B. Schwer at Cornell University for kindly donating the yDBR enzyme.

Supplementary Data

^{31}P NMR/^{1}H NMR spectra of branch points and HPLC analysis of selected bRNAs are available on pp141-147.

References

1. Wallace, J.C. and Edmonds, M. (**1983**). Polyadenylated nuclear RNA contains branches. *Proc. Natl. Acad. Sci. USA*, 80, 950-954.

2. Sharp, P.A. (**1987**). Splicing of messenger-RNA precursors. *Science*, 235, 766-771.

3. Murphy, W.J., Watkins, K.P. and Agabian, N. (**1986**). Identification of a novel Y-branch structure as an intermediate in trypanosome messenger-RNA processing - evidence for *trans* splicing. *Cell*, 47, 517-525.

4. Reed, R. and Maniatis, T. (**1988**). The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes Dev.*, 2, 1268-1276.

5. Zhuang, Y., Goldstein, A.M. and Weiner, A.M. (**1989**). UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl. Acad. Sci. U. S. A.*, 86, 2752-2756.

6. Remaud, G., Balgobin, N., Sandstrom, A., Vial, J.M., Koole, L.H., Buck, H.M., Drake, A.F., Zhou, X.X. and Chattopadhyaya, J. (**1989**). Why do all lariat RNA introns have adenosine as the branch-point nucleotide? Conformational study of naturally-occurring branched trinucleotides and its 11 analogs by ^{1}H-NMR, ^{31}P-NMR and CD spectroscopy. *J. Biochem. Biophys. Methods*, 18, 1-35.

7. Ruskin, B. and Green, M.R. (**1985**). An RNA processing activity that debranches RNA lariats. *Science*, 229, 135-140.

8. Jacquier, A. and Rosbash, M. (**1986**). Efficient transsplicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon intron interaction. *Science*, 234, 1099-1104.

9. Chapman, K.B. and Boeke, J.D. (**1991**). Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell*, 65, 483-492.

10. Khalid, M.F., Damha, M.J., Shuman, S. and Schwer, B. (**2005**). Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Res.*, 33, 6349-6360.

11. Jacquier, A. and Rosbash, M. (**1986**). RNA splicing and intron turnover are greatly diminished by a mutant yeast branch point. *Proc. Natl. Acad. Sci. USA*, 83, 5835-5839.

12. Nam, K., Hudson, R.H.E., Chapman, K.B., Ganeshan, K., Damha, M.J. and Boeke, J.D. (**1994**). Yeast lariat debranching enzyme. Substrate and sequence specificity. *J.*

Biol. Chem., 269, 20613-20621.

13. Carriero, S., Mangos, M.M., Agha, K.A., Noronha, A.M. and Damha, M.J. (**2003**). Branchpoint sugar stereochemistry determines the hydrolytic susceptibility of branched RNA fragments by the yeast debranching enzyme (yDBR). *Nucleosides, Nucleotides & Nucleic Acids*, 22, 1599-1602.

14. Ooi, S.L., Dann, C., III, Nam, K., Leahy, D.J., Damha, M.J. and Boeke, J.D. (**2001**). RNA lariat debranching enzyme. *Methods Enzymol.*, 342, 233-248.

15. Carriero, S. and Damha, M.J. (**2003**). Inhibition of pre-mRNA splicing by synthetic branched nucleic acids. *Nucleic Acids Res.*, 31, 6157-6167.

16. Damha, M.J. and Ogilvie, K.K. (**1988**). Synthesis and spectroscopic analysis of branched RNA fragments: messenger RNA splicing intermediates. *J. Org. Chem.*, 53, 3710-3722.

17. Damha, M.J., Ganeshan, K., Hudson, R.H.E. and Zabarylo, S.V. (**1992**). Solid-phase synthesis of branched oligoribonucleotides related to messenger RNA splicing intermediates. *Nucleic Acids Res.*, 20, 6565-6573.

18. Carriero, S. and Damha, M.J. (**2002**). Solid-phase synthesis of branched oligonucleotides. In Beaucage, S. L. (ed.), *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons., pp. 4.14.11-14.14.32.

19. Karst, S.M., Wobus, C.E., Lay, M., Davidson, J. and Virgin, H.W. (**2003**). STAT1-dependent innate immunity to a Norwalk-like virus. *Science*, 299, 1575-1578.

20. Peng, C.G. and Damha, M.J. (**2005**). Synthesis and hybridization studies of oligonucleotides containing 1-(2-deoxy-2-α-C-hydroxymethyl-β-D-ribofuranosyl) thymine (2'-α-*hm*-dT). *Nucleic Acids Res.*, 33, 7019-7028.

21. Mangos, M.M., Min, K.-L., Viazovkina, E., Galarneau, A., Elzagheid, M.I., Parniak, M.A. and Damha, M.J. (**2003**). Efficient RNase H-directed cleavage of RNA promoted by antisense DNA or 2'F-ANA constructs containing acyclic nucleotide inserts. *J. Am. Chem. Soc.*, 125, 654-661.

22. Damha, M.J. and Zabarylo, S. (**1989**). Automated solid-phase synthesis of branched oligonucleotides. *Tetrahedron Lett.*, 30, 6295-6298.

23. Lecchi, P., Le, H.M.T. and Pannell, L.K. (**1995**). 6-Aza-α-thiothymine - a matrix for maldi spectra of oligonucleotides. *Nucleic Acids Res.*, 23, 1276-1277.

24. Asara, J.M. and Allison, J. (**1999**). Enhanced detection of oligonucleotides in UV

MALDI MS using the tetraamine spermine as a matrix additive. *Anal. Chem.*, 71, 2866-2870.

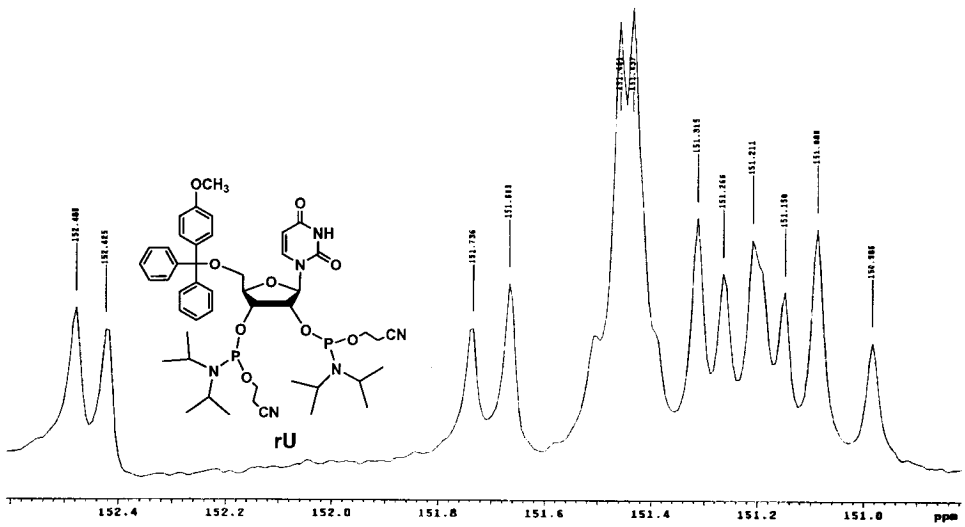
25. Galarneau, A., Min, K.-L., Mangos, M.M. and Damha, M.J. (**2005**). Assay for evaluating ribonuclease H-mediated degradation of RNA-antisense oligonucleotide duplexes. *Methods Mol. Biol.*, Totowa, NJ, United States, Vol. 288, pp. 65-80.

Chapter III. Yeast Debranching Enzyme (yDBR) Inhibition by Novel Branched RNAs

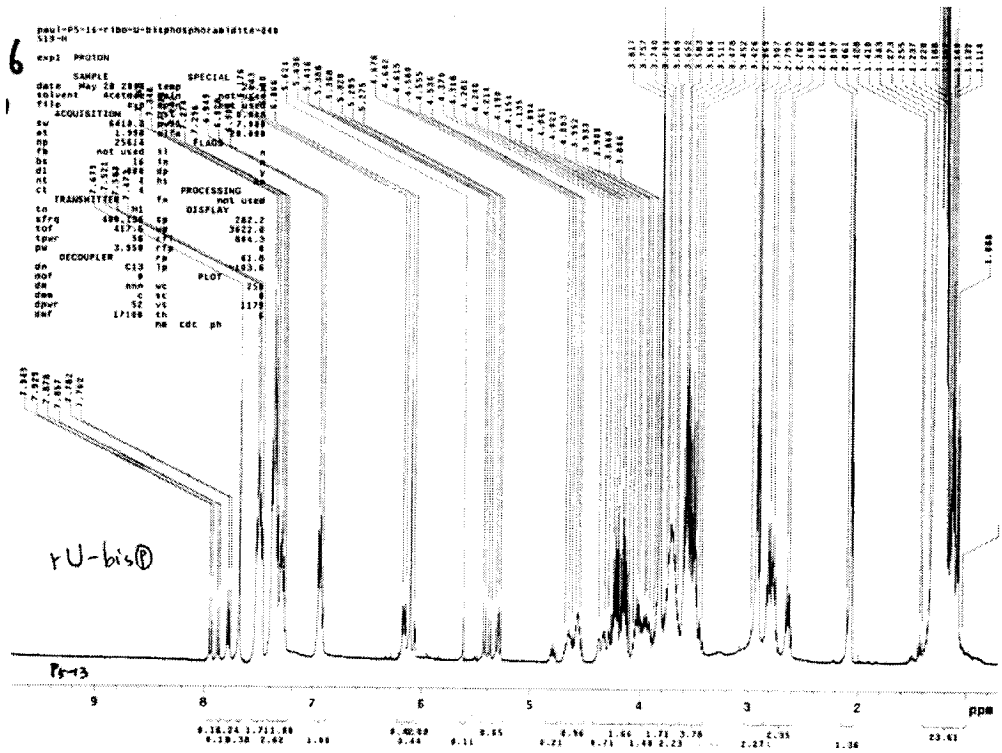
Supplementary Data

(pp141-147)

Contents

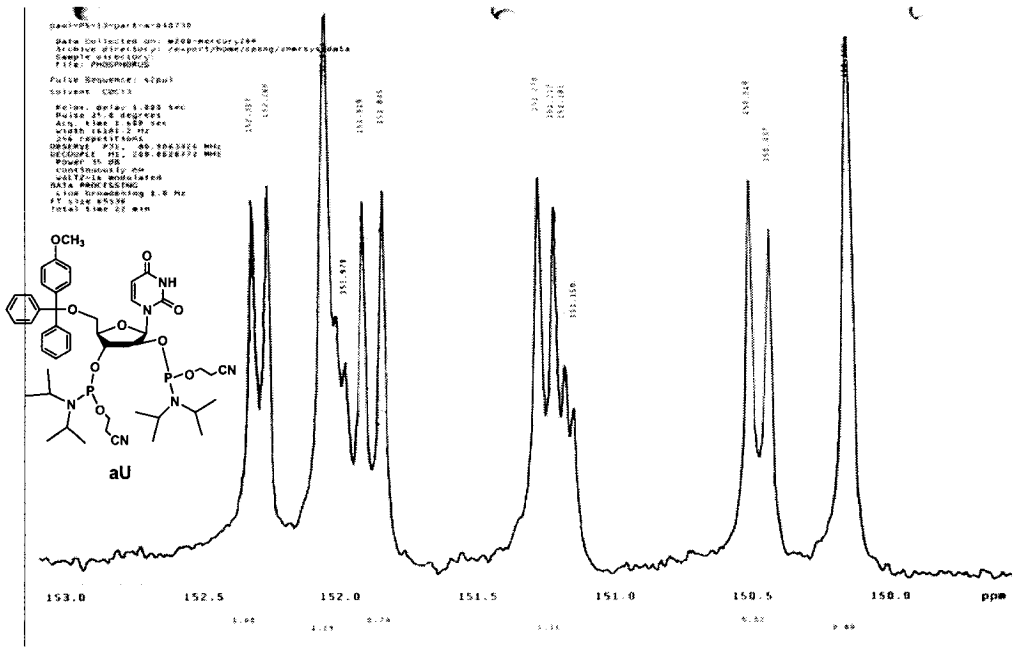


^{31}P NMR (200 MHz, $CDCl_3$): δ 152.41, 152.35, 151.77, 151.70; 151.46, 151.43; 151.39, 151.35, 151.24, 151.20, 151.13, 151.0, 150.89.

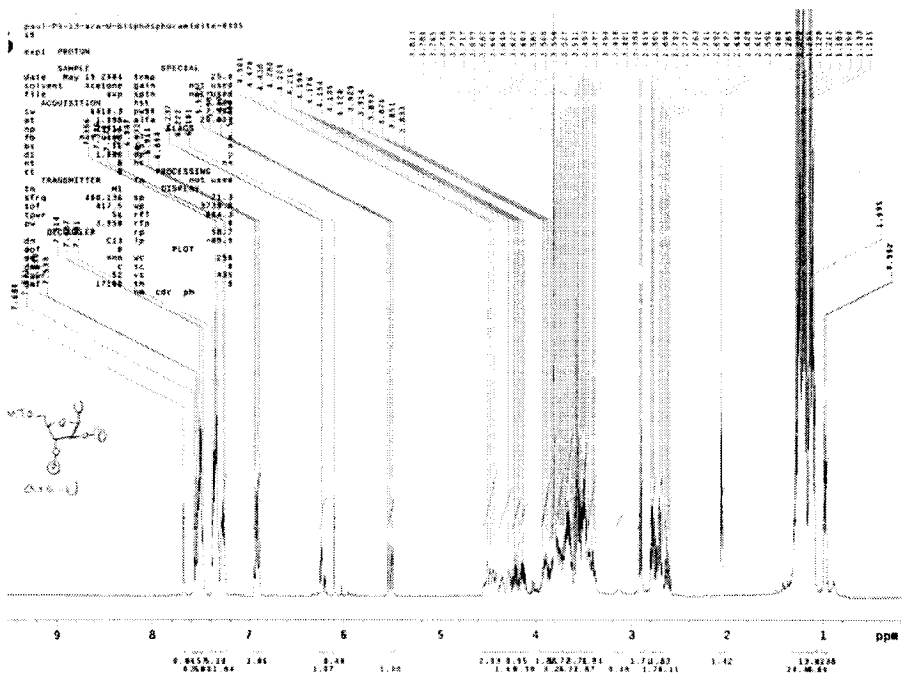


^{1}H NMR (400 MHz, acetone-d_6)

Figure S3.1: ^{31}P NMR and ^{1}H NMR of 1-[2,3-*O*-bis-(β-Cyanoethyl-*N,N*-diisopropylphosphoramidite)-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]uracil (**rU**)

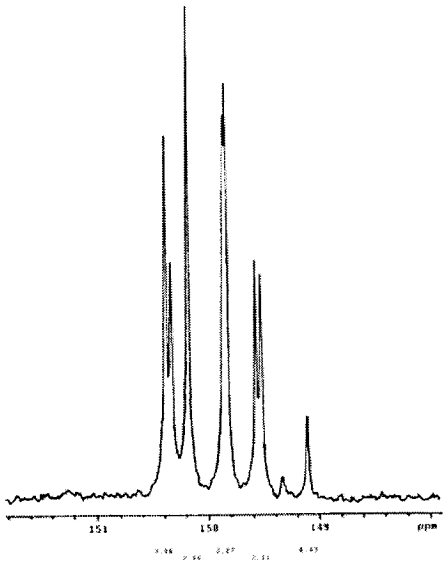


^{31}P NMR (200 MHz, $CDCl_3$): δ 152.32, 152.27, 152.05, 151.98, 151.92, 151.85, 151.28, 151.22, 151.18, 151.15, 151.51, 150.44, 150.16.

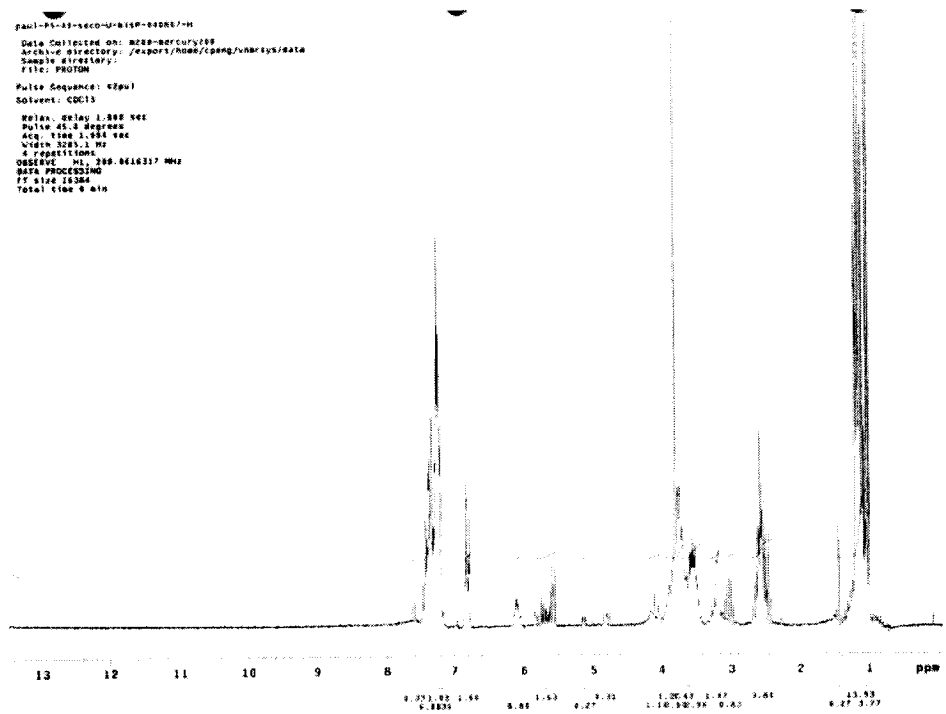


1H NMR (400 MHz, acetone-d_6)

Figure S3.2: ^{31}P NMR and ^{1}H NMR of 1-[2,3-*O*-bis-(β-cyanoethyl-*N,N*-diisopropyl-phosphoramidite) -5-*O*-(4-methoxytrityl-β-D-arabinofuranosyl]uracil (**aU**)

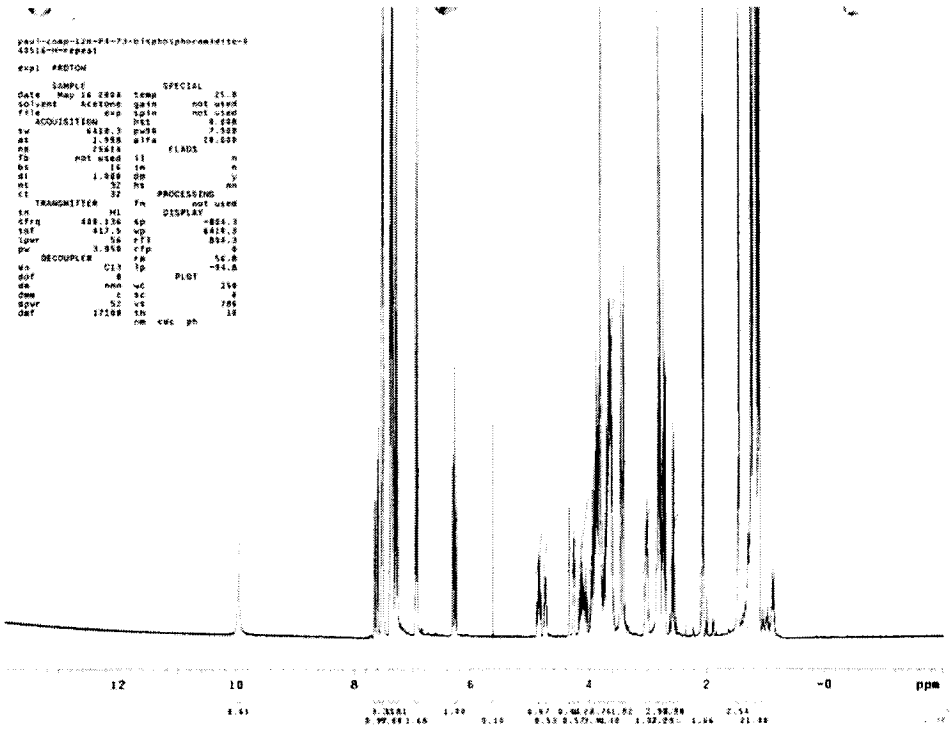


^{31}P NMR (200 MHz, $CDCl_3$): δ 150.39, 150.34, 150.19, 149.86, 149.84, 149.58, 149.53, 149.11.



1H NMR (400 MHz, d6-acetone)

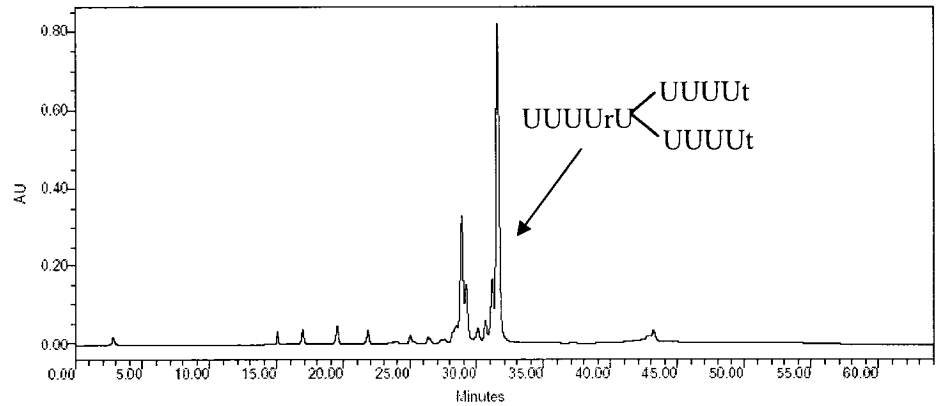
Figure S3.3: ^{31}P NMR and ^{1}H NMR of 1-[2,3-*O*-bis-(β-cyanoethyl-*N,N*-diisopropylphosphoramidite)-5-*O*-(4-methoxytrityl)-2,3-seco)-β-D-ribofuranosyl]uracil (**sU**)



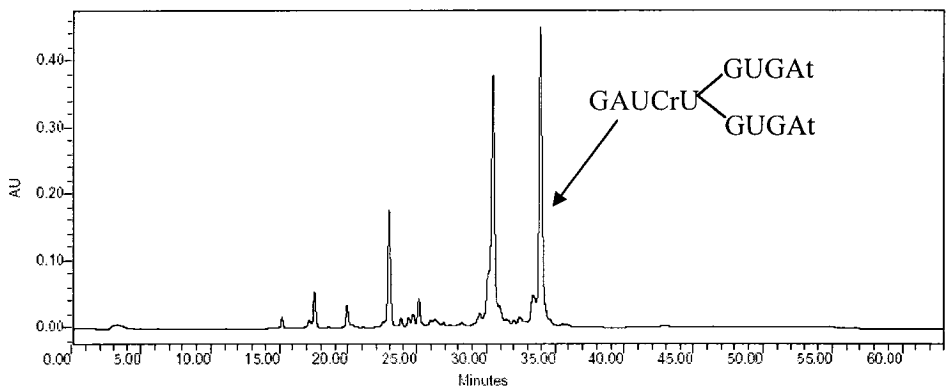
1H NMR (400 MHz, acetone-d_6)

Figure S3.4: ^{1}H NMR of 1-[2-α-*C*-(β-Cyanoethyl-*N,N*-diisopropylphosphoramidic)-hydroxymethyl)-3-*O*-(β-Cyanoethyl-*N,N*-diisopropylphosphoramidite)-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (**H**)

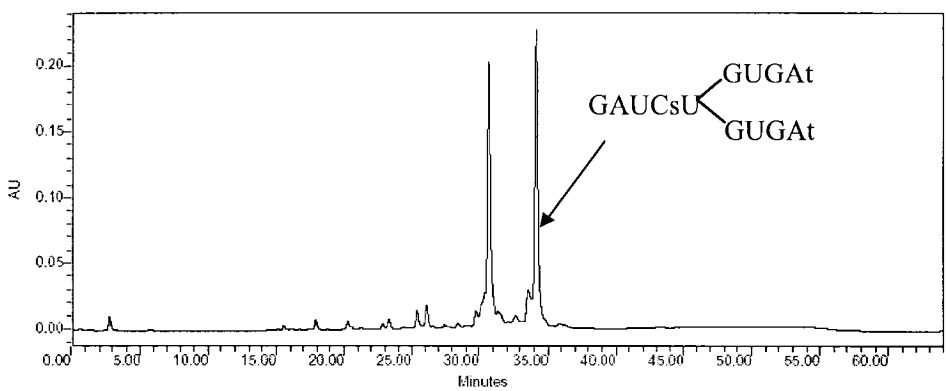
Sample Name:	rUU	Acquired By:	System
Sample Type:	Unknown	Sample Set Name:	
Vial:	1	Acq. Method:	paul_dA_Anal_short
Injection #:	1	Date Acquired:	8/18/04 11:16:37 AM
Run Time:	64.00 Minutes	Injection Volume:	10.00 ul



Sample Name:	rUG	Acquired By:	System
Sample Type:	Unknown	Sample Set Name:	
Vial:	1	Acq. Method:	paul_dA_Anal_short
Injection Volume:	1000.00 ul	Run Time:	63.00 Minutes



Sample Name:	sUG	Acquired By:	System
Sample Type:	Unknown	Sample Set Name:	
Vial:	1	Acq. Method:	paul_dA_Anal_short
Injection Volume:	20.00 ul	Run Time:	64.00 Minutes



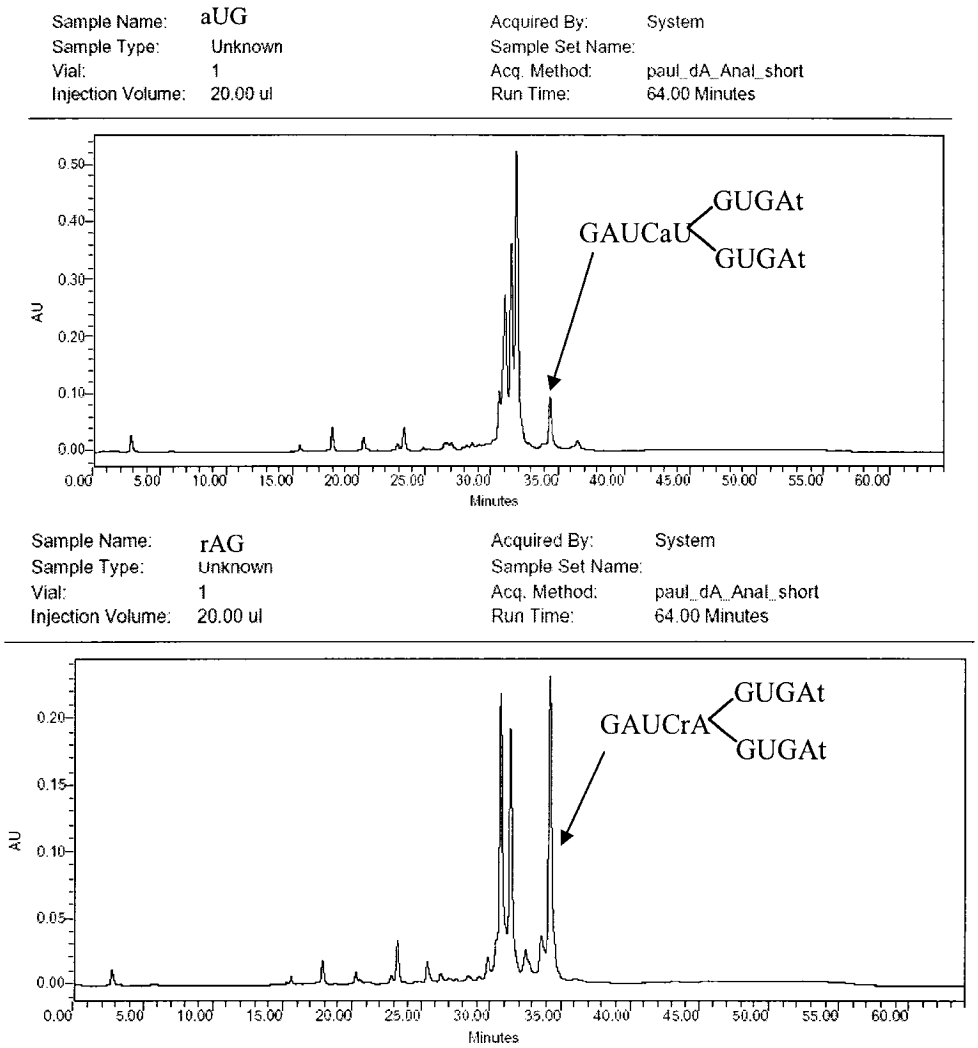


Figure S3.5: HPLC analysis of bRNA crude samples

Chapter IV

G-quadruplexes are interesting structures consisting of a square arrangement of four guanines (a G-quartet) stabilized by Hoogsteen hydrogen bonding and monovalent cations (e.g. K^+) that bind in the center of the quartets. While they can be formed from both DNA and RNA, nothing is known about the capability of arabinose-based oligonucleotides to form quadruplex structures. The Damha group has reported that 2'-deoxy-2'-fluoro-β-D-arabinonucleic acid (2'F-ANA) oligonucleotides enhance the thermal stability of various DNA structures such as duplexes, triplexes, and cytosine-rich quadruplexes. In light of this and other advantageous characteristics of 2'F-ANA, such as synthetic accessibility through conventional solid-phase phosphoramidite chemistry, significant nuclease resistance, and promising antisense properties, we have undertaken the first study concerning the ability of 2'F-ANA to form G-quadruplex structures. The outcome of these and other studies reported here and in subsequent Chapters have opened new perspectives for the biological application of 2'F-ANA as aptamer oligonucleotides.

Chapter IV. G-quadruplex Induced Stabilization by 2'-Deoxy-2'-fluoro-β-D-arabinonucleic Acids (2'F-ANA)

Abstract

The impact of 2'-deoxy-2'-fluoro-β-D-arabinonucleotide residues (2'F-araN) on different G-quadruplexes derived from a thrombin-binding DNA aptamer d($G_2T_2G_2TGTG_2T_2G_2$), an anti-HIV phosphorothioate aptamer PS-d($T_2G_4T_2$), and a DNA telomeric sequence d($G_4T_4G_4$) via UV thermal melting (T_m) and circular dichroism (CD) experiments has been investigated. Generally, replacement of deoxyguanosines that adopt the *anti* conformation (*anti*-guanines) with 2'F-araG can stabilize G-quartets and maintain the quadruplex conformation, while replacement of *syn*-guanines with 2'F-araG is not favored and results in a dramatic switch to an alternative quadruplex conformation. It was found that incorporation of 2'F-araG or T residues into a thrombin-binding DNA G-quadruplex stabilizes the complex (ΔT_m up to ~ + 3°C / 2'F-ANA modification); 2'F-araN units also increased the half-life in 10% fetal bovine serum (FBS) up to 48 fold. Two modified thrombin-binding aptamers (PG13 and PG14) show an approximately four fold increase in binding affinity to thrombin, as assessed via a nitrocellulose filter binding assay, both with increased thermal stability (ca. +1°C / 2'F-ANA modification in T_m) and nuclease resistance (4-7 fold) as well. Therefore, the 2'-deoxy-2'-fluoro-β-D-arabinonucleoic acid (2'-FANA) modification is well suited to tune (and improve) the physicochemical and biological properties of naturally occurring DNA G-quadruplexes.

Introduction

Oligonucleotide aptamers derived from a *Systematic Evolution of Ligands by EXponential enrichment (SELEX)* process,[1-3] have drawn great attention for their potential as therapeutic and diagnostic agents.[4-7] The application of nucleic acid aptamers *in vivo* and their possible use in pharmacotherapy face the same key hurdles as siRNA and antisense based therapeutics, e.g., delivery, cellular uptake and biostability. To date, several methods have been devised to improve the stability of aptamers, most of which make use of SELEX, including a mirror-design of RNA aptamers (or "Spiegelmers");[8] post-SELEX modification by 2'-OMe to increase the stabilization of

RNA aptamers;[9] and direct evolution in SELEX using modified dNTPs or rNTPs (e.g. 2'-F, 2'-NH_2 pyrimidine 5'-triphosphate).[10-14] Macugen®, an aptamer recently approved by the FDA for the treatment of neovascular age-related macular degeneration, is a 2'F/2'-OMe ribose-modified oligomer.[15] There still exists a great demand for new methods and more diverse range of chemistries to create aptamers with more favorable pharmacokinetic properties.

A guanine quartet (G-quartet) structure with four Hoogsteen-paired, coplanar guanines (Figure 4.1A) was first observed more than 40 years ago[16] and later demonstrated to be a unique structural motif of guanine-rich oligonucleotides.[17] These guanine-rich segments are found in nature and also in sequences identified by screening techniques such as SELEX. Nucleic acid G-quadruplexes have also aroused interest as therapeutic agents to inhibit human thrombin,[18,19] HIV infection,[19] HIV-1 integrase activity, [20,21] and as targets to inhibit telomerase activity in anticancer drug design.[22]

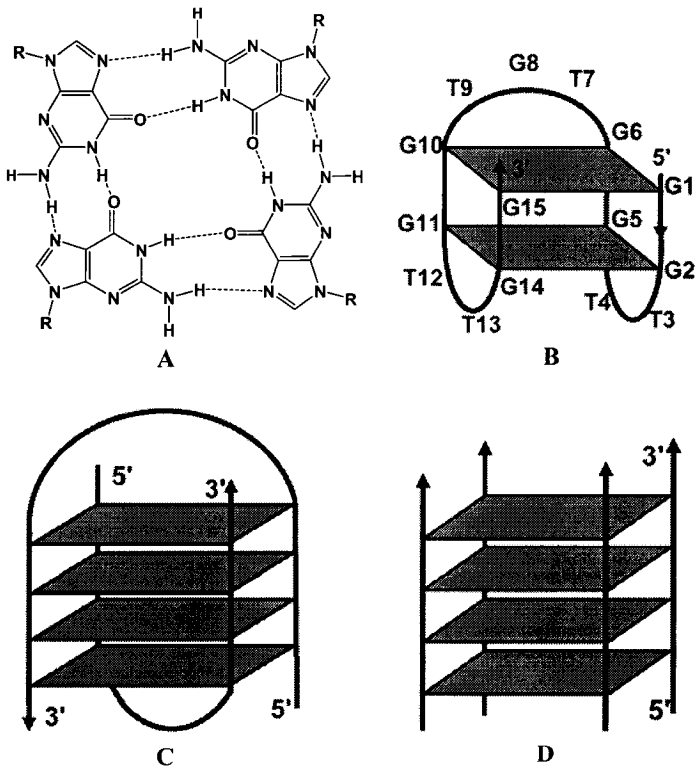


Figure 4.1: Structure of (**A**) G-quartet with cyclic array of four guanines linked by Hoogsteen H-bonds;[16] (**B**) thrombin-binding DNA $d(G_2T_2G_2TGTG_2T_2G_2)$ with an edge-loop (chair-like) unimolecular G-quadruplex in K^+;[23,24] (**C**) a telomeric DNA $d(G_4T_4G_4)$ with diagonal-loop dimeric hairpin complex $[d(G_4T_4G_4)]_2$ in Na^+ [25,26] (**D**) a tetrameric G-quadruplex $[d(PS\text{-}T_2G_4T_2)]_4$ with four parallel strands.[17,19]

G-quartets display extraordinary structural polymorphism.[27-29] The 15-mer DNA sequence $d(G_2T_2G_2TGTG_2T_2G_2)$[18] binds and inactivate thrombin, a key enzyme in the blood clotting cascade. NMR spectroscopy reveals that it adopts a uniquely folded structure with two stacked G-quartets connected through edge-loops (one TGT and two TT loops) involving antiparallel alignment of adjacent strands (see Figure 4.1B).[23,24,30,31] A seemly different crystal structure of this aptamer-thrombin complex[32,33] was later analyzed to be in agreement with the previous NMR structure.[33,34] Potassium ions stabilize the G-quadruplex by coordination to the G residues.[23] The G-quartet displays an alternating *syn* and *anti* conformation of guanine bases (*syn*- and *anti*-Gs) on the same plane and 5'-*syn*-G/3'-*anti*-G along each strand of the quadruplex (i.e. 5'-*syn-anti*-3' connectivity where *syn*-G includes G1, G5, G10 & G14 and *anti*-G includes G2, G6, G11 & G15 in Figure 4.1B) and all the thymidines are in the *anti* orientation.[23]

The telomeric sequence $d(G_4T_4G_4)$ adopts symmetrical dimeric quadruplexes comprising four G-quartets linked through a diagonal loop (Na^+) (Figure 1C)[25,26] or an edge-loop structure (K^+).[35] The diagonal loop configuration (Figure 1C) consists of *5'-syn*-G/*3'-anti*-G along each G-strand, *syn-syn-anti-anti* conformations in a G-quartet and all sugar residues puckered in the *south* (C2'-*endo)* conformation.[26,36]

Another interesting example is the phosphorothioate oligonucleotide PS-$d(T_2G_4T_2)$ identified as an inhibitor of HIV-1 infection *in vitro* by combinatorial screening of a library of DNA strands. This aptamer inhibits HIV envelope-mediated cell fusion *in vitro* and its structure consists of a parallel-stranded tetramer with all guanosines in the *anti*-conformation (Figure 1D).[17,19]

Based on the above, the glycosidic conformations of guanines (i.e. *syn*- and *anti*-Gs) have a strong correlation with strand alignments. Parallel-stranded quadruplexes (Figure 4.1D) support only *anti*-G residues, while antiparallel-stranded quadruplexes favor alternating *syn-anti* Gs engaged in unimolecular (Figure 4.1B) or intermolecular complexes (Figure 4.1C).[27,37]

G-quadruplexes are sensitive to chemical modifications. Several studies aimed at modifying the thrombin-binding aptamer $d(G_2T_2G_2TGTG_2T_2G_2)$ have been reported, but very few, if any, have led to an improvement over the original molecule. For example, Heckel and Mayer reported that the introduction of thymines modified with a nitrophenylpropyl moiety (T-NPP) at certain positions generally abolished interaction of

the aptamer with thrombin.[38] Di Giusto and King reported the synthesis of circular aptamers targeted against thrombin that showed improved nuclease resistance and higher anticoagulant potency than those of the canonical thrombin DNA aptamer. [39] However, circularization of the aptamers produces a mixture of constructs and requires a ligase enzyme, making the method very difficult to scale-up. Other attempts to circularize the thrombin-binding DNA aptamer via chemical methods resulted in a complete loss of the anti-thrombin activity.[40] Recently, Seela and coworkers reported the insertion of a hairpin-forming sequence GCGAAG into the position of the central loop (TGT) of the thrombin-binding aptamer. This construct formed a G-quadruplex fused to a mini-hairpin structure. According to the T_m data, the mini-hairpin induces a structural change in the aptamer section, leading to less stable G-quadruplex. Binding to thrombin was not investigated. [41] Saccà *et al.* studied the effect of backbone charge and atom size, base substitutions as well as the effect of modification at the sugar 2'-position as analyzed by spectroscopy. All sugar (ribose, 2'-*O*-methylribose) and phosphate (methylphosphonate, phosphorothioate) modifications led to a reduction in the thermal stability of the aptamer.[42] In fact, the 2'-*O*-methylribose modification led not only to a destabilization of the structure but to a complete transformation of the G-tetrad conformation, as shown by spectroscopy in potassium buffer.[42] 2'-*O*-methylribose was also shown to causes structural changes in RNA aptamers and often resulted in a loss of activity. [9] Accordingly, there is a need for new chemical modifications to improve the nuclease stability of this and other aptamers. Ideally, these modifications will not alter the subtle binding interactions of the selected native aptamers and the thermal stability of G-quadruplexes.

2'-Deoxy-2'-fluoro-β-D-arabinonucleic acids (2'F-ANA) confer DNA-like (*South/East)* conformations[43] to oligonucleotides while rendering them more nuclease resistant.[44] The incorporation of 2'F-araN units in oligonucleotides also raises the T_m of different systems, *i.e.*, duplexes (~+1 °C/nt),[45] triplexes (~+0.8 °C/nt)[46] and C-rich quadruplexes (~+1 °C/nt, pH<4.0).[47] In light of this and other advantageous characteristics of 2F-ANA, such as synthetic accessibility through conventional solid-phase phosphoramidite chemistry, and promising antisense/siRNA properties, we have undertaken the first study concerning the ability of 2'F-ANA to form various G-quadruplex structures as analyzed by T_m and CD experiments. 2'F-ANA modified

thrombin-binding aptamers were further evaluated by their nuclease resistance and binding affinity to thrombin. The outcome of these studies has opened new perspectives for the application of 2'F-ANA as aptamer oligonucleotides.

Materials and Methods

Chemical Synthesis of Oligonucleotides

The sequence and composition of the oligomers prepared in this study are shown in Table 4.1 and Table 4.2. Arabinose modified aptamer syntheses were carried out at a 1 μmol scale on an Applied Biosystems (ABI) 3400A synthesizer using standard β-cyanoethylphosphoramidite chemistry according to published protocols.[48] The final concentrations of the monomers were 0.10 M for 2'-deoxyribonucleoside phosphoramidites and 0.125 M for the 2'F-arabinose phosphoramidites. The coupling time was extended to 150 seconds for the 2'-deoxyribonucleoside phosphoramidites dC and dG, and 15 min for the 2'F-araG and T phosphoramidites. These conditions gave about 99% average stepwise coupling yields. Oligonucleotides were purified by anion-exchange HPLC (Waters Protein Pak DEAE-5PW column; 7.5 mm × 7.5 cm), desalted by size-exclusion chromatography on Sephadex G-25 resin, and characterized by MALDI-TOF mass spectrometry (Kratos Kompact-III instrument; Kratos Analytical Inc., New York). Purity of the isolated oligonucleotides was >95% by HPLC.

UV Thermal Dissociation Studies (T_m)

UV thermal dissociation data was obtained on a Varian CARY 1 spectrophotometer equipped with a Peltier temperature controller. Thrombin-binding aptamers (PG1-14) were dissolved in T_m buffer (10 mM Tris, pH 6.8, with and without 25 mM KCl) at a final concentration of 8 μM.[49] Thrombin-binding aptamers were annealed in T_m buffer at 80°C for 10 minutes, allowed to cool to room temperature and refrigerated (4°C) overnight before measurements. $dT_2G_4T_2$ and related sequences (PG17-20) were dissolved in phosphate buffered saline (PBS buffer, pH 7.2) composed of 137 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4 at a final concentration of 20 μM.[50] The telomeric DNA $dG_4T_4G_4$ and related sequences (PG21-24) were dissolved in 10 mM sodium phosphate buffer (pH 7, 0.1 mM EDTA, and 200 mM NaCl) at a final concentration of 100 μM.[51] All samples (PG17-24) were annealed at 98°C for 5 minutes,

naturally cooled down to room temperature and refrigerated (4°C) overnight before measurements. The annealed samples were transferred to pre-chilled Hellma QS-1.000 (Cat #114) quartz cell, sealed with a Teflon-wrapped stopper and degassed by placing them in an ultrasonic bath for 1 min. Extinction coefficients were obtained from the following internet site (http://www.idtdna.com/analyzer/applications/oligoanalyzer) based on the nearest-neighbor approach[52] and modified aptamers (phosphorothioates and 2'F-ANA) were assumed to have the same extinction coefficient as the natural DNA aptamer. Denaturation/cooling curves were acquired at either at 295 nm for $d(G_2T_2G_2TGTG_2T_2G_2)$ and related sequences (PG1-14), $dG_4T_4G_4$ and related sequences (PG21-24), or at 260 nm for $dT_2G_4T_2$ and related sequences (PG17-20), at a heating/cooling rate of 0.5°C/min between 10°C and 80°C (for PG1-14), 20°C and 90°C (for PG17-20) or 40°C and 98°C (for PG21-24). The data were analyzed with the software provided by Varian and converted to Microsoft Excel (Table 4.1 & 4.2). The decreases in UV absorbance (hypochromicity) with increasing temperature were normalized between 1 and 0 by the formula: $N = (A_t-A_l)/(A_h-A_l)$, where A_t is the absorbance at any given temperature (t), A_l is the minimum absorbance reading at high temperature, and A_h is the maximum absorbance reading at low temperature. T_m concentration dependence studies were also conducted in the same way at 295 nm using thrombin-binding aptamers (PG1-14) with different concentrations ranging from 4 to 76 μM. Starna quartz cells (Starna Cells, Inc., Cat. # 1-Q-1) with 1mm path length were used at high concentrations to reduce the amount of aptamers required and to avoid exceeding the Absorbance range of the instrument (Figure 4.4).

Circular Dichroism (CD) Spectra

CD spectra (200-320 nm) were collected on a Jasco J-710 spectropolarimeter at a rate of 100nm/min using fused quartz cells (Hellma, 165-QS). Measurements were carried out either in 10 mM Tris, pH 6.8 (with and without 25 mM KCl) at a concentration of 8 μM for thrombin-binding aptamers (PG1-14);[49] in PBS buffer (pH 7.2, 137 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4) for $dT_2G_4T_2$ and related sequences (PG17-20) at a final concentration of 20 μM,[50] or in sodium phosphate buffer (10 mM sodium phosphate buffer, pH 7, 0.1 mM EDTA, and 200 mM NaCl) for $dG_4T_4G_4$ and related sequences (PG21-24) at a final concentration of 100 μM.[51] Temperature was controlled

by an internal circulating bath (VWR Scientific) at constant temperature. The data was processed using J-700 Windows software supplied by the manufacturer (JASCO, Inc.). To facilitate comparisons, the CD spectra were background subtracted, smoothed and corrected for concentration so that molar ellipticities could be obtained (Figure 4.3 C&D and Figure S4.4 in Supplementary Data). Temperature-dependent CD spectra were also conducted for $dT_2G_4T_2$ and related sequences (PG17-20). A 10 min equilibration time was allowed at each temperature before CD scanning. The T_m profile was obtained by plotting the maximum molar ellipticities *vs.* temperature and normalizing (Figure 4.5 and Figure S4.5 in Supplementary Data).

Nuclease Stability Assay

Nuclease stability of anti-thrombin aptamers was conducted in 10% fetal bovine serum (FBS, Wisent Inc., Cat. #080150) diluted with multicell Dulbecco's Modified Eagle's Medium (DMEM, Wisent Inc., Cat. #319005-CL) at 37°C. A single strand DNA (ssDNA) 23mer (P-8), which is unable to form G-quadruplexes, was used as a control. Approximately 8 μmol of stock solution of aptamers and ssDNA control (~1.2 O.D.U) was evaporated to dryness under reduced pressure and then incubated with 300 μl 10% FBS at 37°C. At 0, 0.25, 0.5, 1, 2, 6, and 24 h, 50 μl of samples were collected and stored at -20°C for at least 20min. The samples were evaporated to dryness and then 10μl of gel loading buffer and 10 μl of autoclaved water was added. 10 μl of the mixture was used for polyacrylamide gel electrophoresis (PAGE), which was carried out at room temperature using 20% polyacrylamide gel in 0.5 × TBE buffer (Tris-borate-EDTA). The degradation patterns on the gels were visualized using Stains-All (Bio-Rad) according to the manufacturer's protocol (Figure 4.8 and Figure S4.6 in Supplementary Data).

5'-End Labeling of Synthetic Oligonucleotides

Aptamers (PG1-14) were radiolabeled at the 5'-hydroxyl terminus with a radioactive ^{32}P probe using a T4 polynucleotide kinase (T4 PNK) according to the manufacture's specifications (MBI Fermentas Life Sciences, Burlington, ON). Incorporation of the ^{32}P label was accomplished in reaction mixtures consisting of DNA aptamers substrate (100 pmol), 2 μl 10×reaction buffer (500 mM Tris-HCl, pH 7.6 at 25°C, 100 mM $MgCl_2$, 50

mM DTT, 1mM spermidine and 1 mM EDTA), 1 μl T4 PNK enzyme solution (10 U/1μl in a solution of 20 mM Tris-HCl, pH 7.5, 25 mM KCl, 0.1 mM EDTA, 2 mM DTT and 50% glycerol), 6 μl [γ-^{32}P]-ATP solution (6000 Ci/mmol, 10 mCi/ml; Amersham Biosciences, Inc.) and autoclaved sterile water to a final volume of 20 μl. The reaction mixture was incubated for 30-45 min at 37°C, followed by a second incubation for 5 min at 95°C to thermally denature and deactivate the kinase enzyme. The solution was purified according to a standard protocol[53] and the isolated yield of ^{32}P-5'-DNA following gel extraction averaged 50%. The pure labeled samples were kept at -20°C for future use.

Nitrocellulose Filter Binding Assay

Labeled aptamers (1.25 pmol) were heated to 95°C for 5 minutes in the binding buffer (Tris-Ac, pH 7.4, 140 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$)[18] and immediately placed on ice for 5 minutes before binding to increasing concentrations of thrombin protease (Amersham Biosciences, Inc.) ranging from 10–1000 nM in the binding buffer at 37°C in a final volume of 20 μl for 30 minutes. Mixtures were filtered through a nitrocellulose filter (13 mm Millipore, HAWP, 0.45 μm) pre-wetted with binding buffer in a Millipore filter binding apparatus, and immediately rinsed with 600 μl ice cold washing buffer (Tris-Ac, pH 7.4, 140 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1% sodium pyrophosphate (w/v)), then the filter was air dried and the bound aptamer quantified by scintillation counting. The binding percentage (%) was calculated by the subtraction of the background from the counts in the microtube. K_d was roughly estimated from the concentration (nM) where 50% of the maximum binding percentage was observed with a certain aptamer during the thrombin concentrations studied. The binding curves for various aptamers, including controls, are shown in Figure 4.9, whereas binding data are given on Table 4.1.

Results and Discussion

A Fully 2'F-ANA Modified Thrombin-binding Aptamer

The study began with the replacement of all the deoxynucleotides in the thrombin-binding aptamer d($G_2T_2G_2TGTG_2T_2G_2$) by 2'F-ANA units. Consistent with literature

results,[23,49] it was found through comparisons of the T_m profiles both with and without the presence of K^+, that potassium ions stabilized the unimolecular G-quadruplex of the DNA aptamer (PG1) (Figure 4.3A & B). As shown in the literature, similar hypochromicity (UV absorbance decrease with increased temperature) was observed at 295 nm for all DNA (PG1),[49] which is an indication of G-quartet formation.[28,54] The T_m of the all DNA aptamer obtained in this study was 47.4°C (Table 4.1 and Figure 4.3B), consistent with 46.4°C reported by Smirnov and Shafe[49] and the CD spectra also support the UV melting experiment. Previous studies have shown two typical CD spectra: a CD spectrum with a positive CD band at ~ 265 nm and a negative band at ~ 240 nm (Type I CD spectrum) is related to parallel quadruplex with all *anti*-Gs, while one with a positive band at ~295 nm and a negative band at ~ 260 nm (Type II CD spectrum) corresponds to antiparallel quadruplex with alternating *syn-anti* Gs.[28] The all DNA aptamer shows very low amplitude bands without K^+, whereas it demonstrates characteristic Type II CD when K^+ is present, indicating a G-quadruplex structure with alternating *syn-anti* Gs (Figure 4.3C & D).[27,28] The lack of concentration dependence in the T_m data (Figure 4.4) and the lack of hysteresis in the heating/cooling processes (Figure 4.5) support a unimolecular G-quadruplex structure (Figure 4.1B). Lack of hysteresis at a heating/cooling rate of 0.5 °C/min indicates a fast kinetics for the formation of unimolecular quadruplexes.[54]

The all 2'F-ANA aptamer (PG2) also forms a defined G-quadruplex stabilized by K^+ in the UV melting experiment (Figure 4.3A &B). The G-quadruplex of all 2'F-ANA aptamer (PG2) is more thermally stable than the all DNA aptamer (PG1) (ΔT_m = 0.4 °C/ 2'F-ANA modification in Table 4.1). Further characterization experiments indicated that the G-quadruplex PG2 formed in K^+ is different from that of that of PG1. The effect of K^+ on the CD spectra is consistent with the T_m experiment: very low absorbance in the absence of K^+ and a strong positive absorbance peak at 260 nm in the presence of K^+. Clearly this CD spectrum is type I, which corresponds to a parallel quadruplex with all *anti*-Gs.[27,28] It is a strong indication that the all 2'F-ANA aptamer (PG2) could not maintain the unimolecular G-quadruplex topology (Figure 4.1B). An intermolecular G-quadruplex is thus formed to support a G-quartet structure with all *anti*-Gs. A concentration-dependent T_m profile supports the existence of intermolecular G-quadruplex formation (Figure 4.4). Hysteresis was observed in the heating and cooling

Table 4.1: Sequences, CD, T_m and binding data of 2'F-ANA modified thrombin-binding aptamers

Code	Type	Sequence[a]	T_m (ΔT_m)[b] (°C)	CD Type[c]	Hysteresis in T_m[d]	K_d (nM)[e]	$t_{1/2}$ (h)[f]
PG1	All DNA	d(GGTTGGTGTGGTTGG)	46.4[49] 47.4	II	no	210 200[18]	0.5
PG2	All 2'F-ANA	d(**GGTTGGTGTGGTTGG**)	54.1 (+0.4)	I	yes	500	>24
PG3	2'F-ANA G-*anti*	d(G**G**TTG**G**TGTG**G**TTG**G**)	53.3 (+1.5)	II	no	>700	4.8
PG4	2'F-ANA G-*anti*& loop	d(G**GTTG**G**TGTG**G**TTG**G**)	61.6 (+1.3)	II	no	500	9.4
PG5	2'F-ANA G-*syn*	d(**G**GTTG**G**TGTG**G**TTG**G**)	45.4 (-0.5)	I	yes	>700	0.8
PG6	2'F-ANA G-*syn*&loop	d(**GGTTG**G**TGTG**G**TTG**G**)	48.5 (+0.1)	I	yes	>700	0.6
PG7	2'F-ANA G-quartet	d(**GG**TTG**G**TGTG**G**TTG**G**)	50.2 (+0.4)	I	yes	450	4.0
PG8	2'F-ANA all-loop	d(GG**TTGG**TGTG**GTT**GG)	56.3 (+1.3)	II	no	280	5.9
PG9	2'F-ANA loop	d(GG**TTGG**T**G**TG**GTT**GG)	57.1 (+1.6)	II	no	300	2.8
PG10	2'F-ANA loop	d(GGT**TGG**TGTG**GTT**GG)	51.2 (+0.8)	II	no	250	2.7
PG11	2'F-ANA loop	d(GG**TTGG**TGTGGTTGG)	56.6 (+1.8)	II	no	370	5.1
PG12	2'F-ANA loop	d(GG**TT**GGTGTG**GTT**GG)	59.1 (+2.9)	II	no	310	3.5
PG13	2'F-ANA loop	d(GGTTGG**T**G**T**G**GTT**GG)	51.0 (+0.9)	II	no	58	3.4
PG14	2'F-ANA loop	d(GG**TTGG**T**G**TGGTTGG)	50.6 (+0.8)	II	no	40	2.0
P8	ssDNA control	d(GTCTCTTGTGTGACTCTGGTAAC)	NA	NA	NA	NC	0.5
H1	Hairpin control (RNA)	r(GGACUUCGGUCC)	NA	NA	NA	NC	NA

[a] Capital and bold letter: **2'F-ANA**

[b] ΔT_m is the T_m change per 2'F-ANA between any modified aptamer with the unmodified DNA aptamer (PG1, T_m = 47.4°C); NA: not applicable.

[c] "Type I" CD spectrum refers to a positive CD band at ~265 nm and a negative band at ~ 240 nm that correlates with G-anti conformation in the the G-quartet. "Type II" CD refers to a CD spectrum with positive band at ~ 295 nm and a negative band at ~ 260 nm, which indicates a alternative *anti*-G and *syn*-G conformation in the G-quartet . CD was measured in the buffer of 10 mM Tris, pH 6.8, 25mM KCl.[28]

[d] Hysterisis in T_m refers to the hysterisis existing between a heating and cooling process with 0.5°C/min temperature change during T_m measurements.

[e] Kd was roughly estimated from the concentration (nM) where 50% of the maximum binding percentage was observed with a certain aptamer during the thrombin concentrations studied; NC: not calculated.

processes, again supporting an intermolecular G-quadruplex adopted by PG2 (Figure4.5) It is expected that PG2 can form a tetrameric G-quadruplex, as reported in the literature to exist in many short G-rich sequences including the anti-HIV phosphorothioate PS-d($T_2G_4T_2$) and others.[19,55-58]

Replacement of *Anti*-Gs and *Syn*-Gs with 2'F-ANA

The next phase of the investigation studied the effect of 2'F-araG on *anti*-Gs and *syn*-Gs observed in different G-quadruplexes as discussed in the Introduction. The conformationally biased 2'F-ANA unit[43] (i.e. more *south/east* than its DNA counterpart), featuring a *β*-face fluorine, is expected to stabilize G-quartets requiring *anti*-Gs to avoid steric interaction between guanine and the fluorine.[59] Preorganization of *anti* glycosidic conformation favored by 2'F-araG[59] (Figure 4.2) will reduce the loss of entropy during G-quartet assembly so as to contribute the stabilization of G-quadruplex complex. Conversely, replacement of *syn*-Gs with 2'F-araGs is expected to cause destabilization since 2'F-araG is forced to adopt an unfavorable glycosidic orientation or abandon its preferred G-quartet structure altogether. Previous examples of the stabilizing effect of preorganization can be seen in the application of *north*-conformation locked nucleic acid (LNA), used to stabilize an RNA-like structure[60] and a more compact 2'-deoxy-2'-hydroxymethyl-homouridine (2-α-*hm*-dT) to stabilize a compact 2',5'-RNA hairpin loop structure.[61]

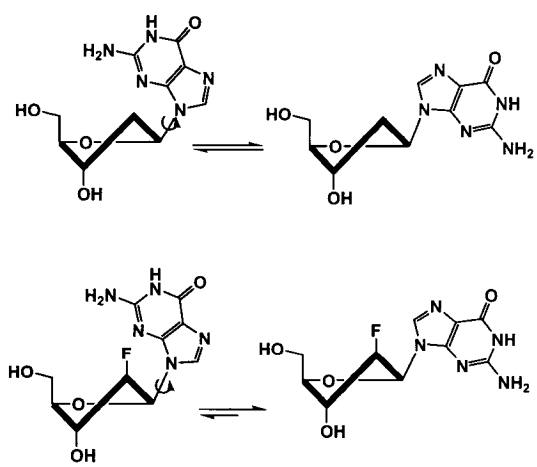


Figure 4.2: Glycosic conformation equilibrium of dG and 2'F-araG with south furanose sugar ring

The obtained T_m and CD data are consistent with the expected results. Modifying the *anti*-Gs with 2'F-ANA (PG3) or modifying both the *anti*-Gs and the loops with 2'F-ANA (PG4) increases thermal stability in the presence of K^+ (ΔT_m = 1.3 and 1.5 °C/ 2'F-ANA modification for PG3 and PG4, respectively). The CD spectrum of PG3 was Type I

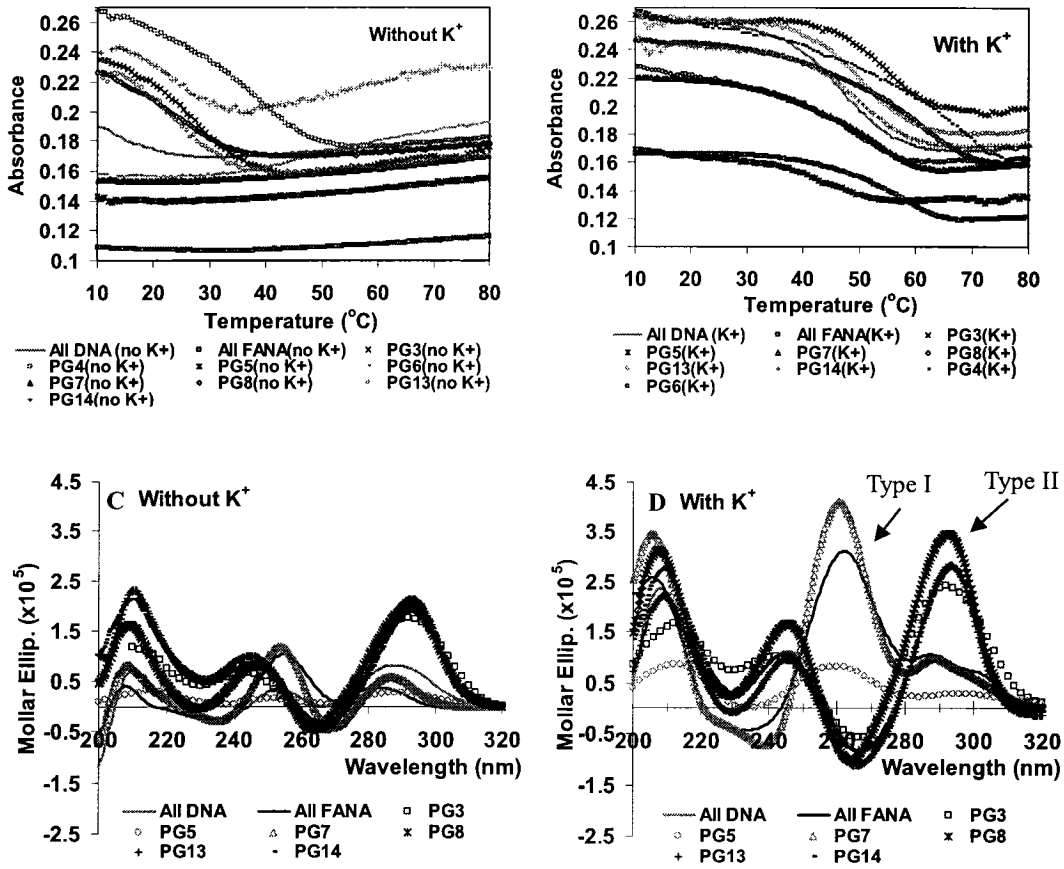


Figure 4.3: T_m profiles of selected 2'F-ANA-modified thrombin-binding aptamers (Table 4.1) measured at 295 nm in buffer 10 mM Tris, pH 6.8 (**A**) without KCl; (**B**) with 25mM KCl, at a final strand concentration of 8 μM. The T_m data are provided in Table 1. CD spectra in the same buffer consisting of 10 mM Tris, pH 6.8 (**C**) without KCl; (**D**) with 25mM KCl at 15°C at a final strand concentration of 8 μM. Type I and Type II CD types are shown in (**D**)

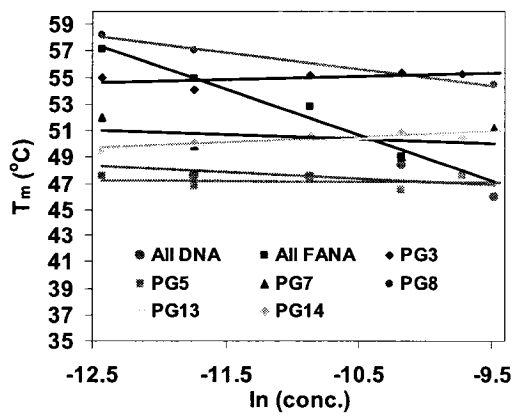


Figure 4.4: T_m *vs* concentration dependence (15-20 folds) study measured at 295 nm in a buffer consisting of 10 mM Tris, pH 6.8, 25 mM KCl for selected 2'F-ANA modified thrombin-binding aptamers (Table 4.1).

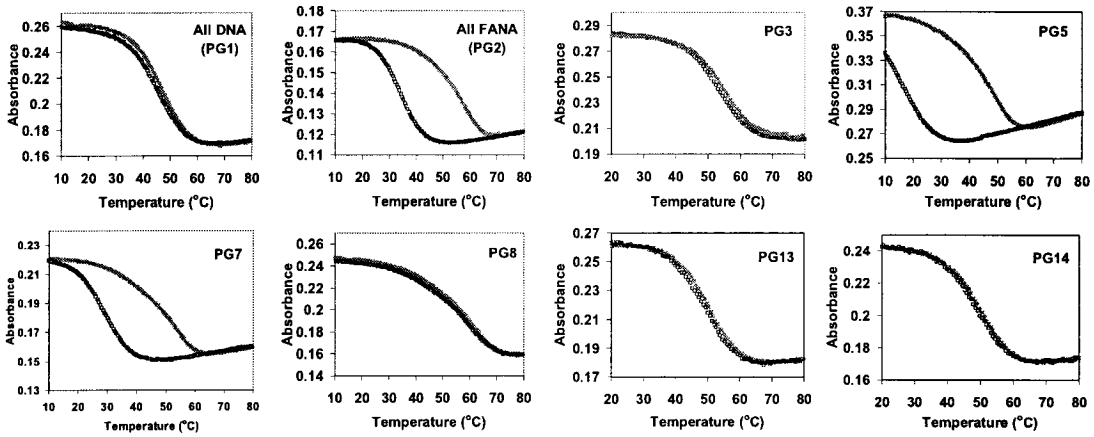


Figure 4.5: Heating and cooling T_m transitions of 2'F-ANA modified thrombin-binding aptamers (Table 1) measured at 295 nm in a buffer of 10 mM Tris, 25 mM KCl, pH 6.8, at a final strand concentration of 8 μM in a heating/cooling rate of 0.5°C/min (red filled square: heating; blue empty square: cooling).

CD, similar to the all DNA aptamer (PG1) (Figure 4.3C & D). PG4 shows a clear Type I CD both without and with K^+, demonstrating that the G-quadruplex adopted by PG4 could form even in the absence of K^+, thus suggesting that replacement of the loop with 2'F-ANA further benefits the G-quadruplex formation. This is consistent with previous work showing that all the thymidines in this thrombin-binding DNA aptamer adopt the *anti* orientation.[23] A concentration-independent T_m profile (Figure 4.4) and lack of hysteresis in heating/cooling processes both support a unimolecular G-quadruplex structure like the DNA control (Figure 4.5 and Figure S4.1 in Supplementary Data). Replacement of *anti*-G with 2'F-ANA is therefore able to stabilize the G-quartet without changing the quadruplex topology.

The results of replacement of the *syn*-Gs with 2'F-ANA also match expectations. Modification of the *syn*-Gs with 2'F-ANA (PG5) and replacement of both *syn*-Gs and loops (PG6) with 2'F-ANA result in less stable complexes compared with PG3 and PG4. PG5 is even less stable than the original all DNA aptamer (ΔT_m = -0.5°C/ 2'F-ANA modification, Table 4.1), and PG6 is only as stable as the all DNA aptamer control (ΔT_m = 0.1°C/ 2'F-ANA modification, Table 4.1). CD experiments show that both PG5 and PG6 display very low amplitude peaks at 260 nm even in the presence of K^+, indicating little structure formed (Figure 4.3C &D and Figure S4.4 in Supplementary Data). A T_m *vs.* concentration study suggested that the T_m of PG5 and PG6 is independent of

concentration so they likely still form unimolecular structures (Figure 4.4) but the formation of these structures seems kinetically slow based on the finding that major hysteresis is observed at a heating/cooling rate of 0.5°C/min (Figure 4.5 and Figure S4.1 in Supplementary Data). Taken together, these results suggest that 2'F-araGs can be forced to adopt a *syn* orientation about the glycosidic bond, but it is certainly not favored. If the unimolecular G-quadruplex is maintained in PG5 and PG6, the defined structure is not present in significant quantity in solution, as indicated by the previous CD experiments. Furthermore, in the case where 2'F-araGs replace both *syn*-Gs and *anti*-Gs in the quartet (PG7), we found that 2'F-ANA brings minimal stabilization (ΔT_m = 0.4°C/ 2'F-ANA modification, Table 4.1). Type I CD profile (i.e. a positive CD band at ~260 nm and a negative band at ~ 240 nm) was observed in the presence of K^+ and not shown in the absence of K^+, indicating a parallel quadruplex with all *anti*-Gs,[28] as in PG2; concentration dependent T_m (Figure 4.4) and significant hysteresis observed in heating/cooling processes (Figure 4.5) support the idea that PG7 could adopt a tetrameric parallel G-quadruplex with all *anti*-Gs.

To verify whether *anti*-Gs can be stabilized by 2'F-ANA in other types of G-quadruplexes, we chose two other aptamers from the literature.

A). Anti-HIV Phosphorothioate PS-d($T_2G_4T_2$) Modified with 2'F-ANA. Both the phosphodiester sequence PO-d($T_2G_4T_2$) and the phosphorothioate sequence PS-d($T_2G_4T_2$) are known to adopt a structure in which all G residues are in the *anti* conformation.[19,50] As a result, this aptamer has a characteristic "Type I" CD signature (Table 4.2 and Figures 4.6A).[28] Thermal denaturation studies at 260 nm yielded noisy curves. The reason for this could be the slow dissociation process of both phosphodiester and phosphorothioate reported in the literature.[50] T_m profiles shown in Figure 4.6B were obtained by plotting the maximum molar ellipticities *vs.* temperature based on temperature-dependent CD spectra (Figure S4.5, Supplementary Data). Although it is impossible to get accurate T_m data for these kinetically slow G-quadruplexes, clearly our apparent T_m data help to confirm the significant thermal stabilization provided by the 2'F-ANA-G units (PG18-20, Table 4.2). Our results indicate that phosphorothioate modifications lead to stabilization, inconsistent with the literature finding.[42] Compared with the controls PG17 (PO-DNA) and PG18 (PS-DNA), 2'F-ANA modified sequences

Table 4.2: CD and T_m data of $d(T_2G_4T_2)$ and $d(G_4T_4G_4)$ related oligonucleotides

Code	Type	Sequence[a]	CD Type[b,c]	T_m [d] (ΔT_m) [e] (°C)
$d(T_2G_4T_2)$ and related sequences				
PG17	All PO-DNA	PO-d(TTGGGGTT)	I	66.0
PG18	All PS-DNA	PS-d(TTGGGGTT)	I	73.5 [50] 74.0
PG19	All PS-2'F-ANA	PS-d(**TTGGGGTT**)	I	83.0 (+1.1)
PG20	PS-2'F-ANA-G	PS-d(TT**GGGG**TT)	I	87.0 (+3.3)
$d(G_4T_4G_4)$ and related sequences				
PG21	All PO-DNA	PO-d(GGGGTTTTGGGG)	II	65 [51] 64.4
PG22	All PO-2'F-ANA	PO-d(**GGGGTTTTGGGG**)	I	90.0 (+2.1)
PG23	G-*syn*	PO-d(**GGG**GTTTT**GGG**G)	I	72.5 (+2.0)
PG24	G-*anti*	PO-d(G**GG**GTTTTG**GG**G)	II	66.2 (+0.5)

[a] Capital and bold letter: **2'F-ANA**; PO: phosphate linkage; PS: phosphorothioate linkage.
[b] CD type I & refer to the note in Table 1.
[c] $dT_2G_4T_2$ and related sequences (PG17-20): phosphate buffered saline (PBS buffer, pH 7.2 at 25°C), 137 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4; strand concentration: 20 μM for both CD and T_m experiments. $dG_4T_4G_4$ and related sequences (PG21-24): 10 mM sodium phosphate buffer, 0.1 mM EDTA, pH 7 and 200 mM NaCl; strand concentration: 10 μM.
[d] $dT_2G_4T_2$ and related sequences (PG17-20): phosphate buffered saline (PBS buffer, pH 7.2 at 25°C), 137 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4; strand concentration: 20 μM; T_m data were generated from concentration-dependent CD spectra (Figure S5 and Materials and Method). $dG_4T_4G_4$ and related sequences (PG21-24): 10 mM sodium phosphate buffer, 0.1 mM EDTA, pH 7 and 200 mM NaCl; strand concentration: 100 μM; T_m measurements were conducted at 295 nm wavelength.
[e] ΔT_m (°C) is the T_m change / 2'F-ANA modification of PG18-20 or PG22-24 relative to the control PG18 (74.0°C) or PG21 (PG64.4°C), respectively.

PG19 and PG20 show large increases in molar ellipticities, indicating a better guanine-guanine interactions within the G-tetrad structure.

B). A Telomeric DNA $d(G_4T_4G_4)$ Modified with 2'F-ANA. Structural analysis has shown that a telomeric DNA sequence $d(G_4T_4G_4)$ (PG21, Table 4.2) can form a symmetrical dimeric quadruplex with four G-quartets and a diagonal loop in the presence of Na^+ (Figure1C).[25,26] The guanine conformation is 5'-d($G_sG_aG_sG_a$-TTTT-$G_sG_aG_sG_a$)-3' where *s* denotes *syn* and *a anti*.[26,36] Several sequences, including a fully 2'F-ANA modified sequence (PG22), a sequence with 2'F-ANA replacing all *syn*-Gs (PG23) and a sequence with 2'F-ANA modifications for all *anti*-Gs (PG24), were tested (Table 4.2). When the *anti*-G residues were replaced by 2'F-ANA G residues, a modest thermal

stabilization of the G-quadruplex structure resulted, without disruption of the 3-D structure (Table 4.2, and Figures 4.7 A&B). When the *syn*-G residues were replaced (PG22 & PG23), a remarkable increase in thermal stabilization of the G-quadruplex structure resulted (up to 25°C; Table 4.2), with a concomitant type II-to-type I conformational change induced by the 2'F-ANA-G(*anti*) replacement (Figure 4.7A).

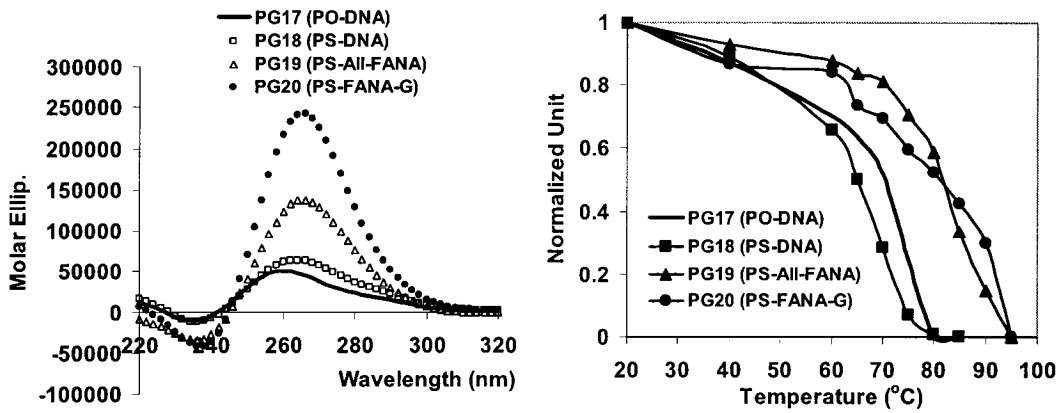


Figure 4.6: (**A**) CD spectra of $dT_2G_4T_2$ and related sequences (PG17-20) obtained in PBS buffer at a strand concentration of 20 μM (single strand). (**B**). T_m curves generated by plotting the maximum absorbance at 265 nm wavelength (normalized) *versus* temperature in a series of CD-temperature dependent measurements (Figure S4.5, Supplementary Data); the corresponding T_m data are shown in Table 2

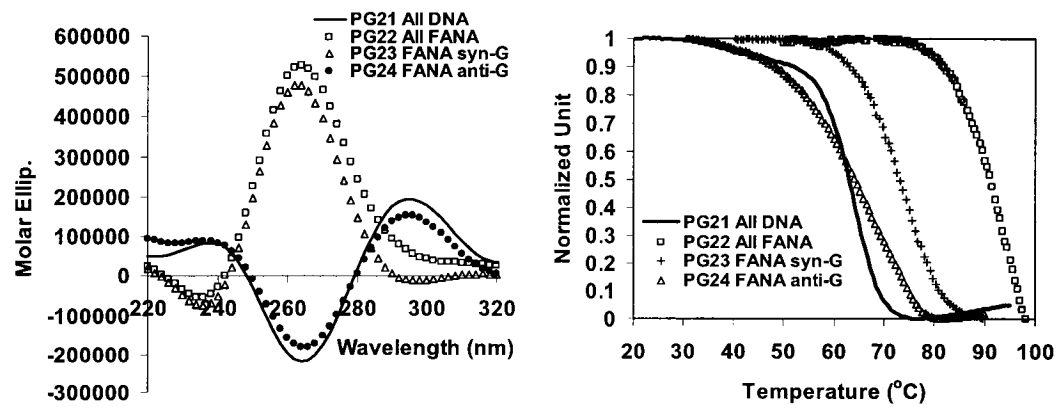


Figure 4.7: (**A**) CD spectra of $dT_4G_4T_4$ and related sequences (PG21-24) in 10 mM sodium phosphate buffer, 0.1 mM EDTA, pH 7 and 200 mM NaCl at a strand concentration of 10 μM. (**B**). T_m profile measured at 295 nm in the same sodium phosphate buffer at a strand concentration of 100 μM

Overall, these two additional examples confirm the results obtained for the thrombin-binding aptamers: the modification of *anti*-Gs with 2'F-ANA can generally stabilize a G-quartet requiring *anti*-Gs, maintaining the overall quadruplex conformation, while the modification of *syn*-Gs with 2'F-ANA is not favored and results in a complete conformational switch to an alternative G-quadruplex structure, as indicated by the change of CD type from II to I for PG2, PG5, PG6, PG7, PG22 & PG23 (Table 4.1 & 4.2, Figure 4.3D, Figure 4.7A and Figure S4.4 in Supplementary Data). The stabilizing effects of 2'F-ANA on G-quadruplexes is related to the conformational preorganization and the small steric effect of the fluorine atom. 2'-O-Methyl modification generates steric problems at the 2' position and is often shown to destabilize a unimolecular G-quadruplex, resulting in significant conformational perturbation in both intra- and intermolecular G-quadruplex systems.[42]

Replacement of Loops with 2'F-ANA

The base conformation of thymidines in the loop region of the thrombin-binding aptamer d($G_1G_2T_3T_4G_5G_6T_7G_8T_9G_{10}G_{11}T_{12}T_{13}G_{14}G_{15}$) (Figure 4.1B) exclusively prefers the *anti* orientation.[23] The literature reports that G_8 shows base stacking interactions with the first G-quartet but the conformation of the nucleotide was not mentioned.[32] The replacement of G_8 with 2'F-araG in this study resulted in an increase of 0.8 °C in T_m (PG8 and PG9 in Table 4.1). It was expected that the replacement of *anti* thymidines in the loops with *anti* conformationally biased 2'F-araT should increase thermal stability. Thermal melting studies support this hypothesis (PG8-14, Table 4.1). Different positions of thymidine contribute to the stability of the G-quadruplex to different degrees (Table 4.1 and Figure S4.2, Supplementary Data). 2'F-ANA replacements in the loops consisting of two thymidines (PG12) bring a greater thermal stabilization (ΔT_m = 2.9°C/ 2'F-ANA modification, Table 4.1) than with modification in the TGT loop (PG9, ΔT_m = 1.6°C/ 2'F-ANA modification, Table 4.1). T_3 and T_{12} are the most sensitive positions to 2'F-araT modification resulting in an almost 6°C difference in T_m (PG9 *vs.* PG10; PG11 *vs.* PG14, Table 4.2), while the replacement of T_{13} and T_4 with 2'F-araT is much less sensitive and results only in slight T_m increases of 1°C (PG9 *vs.* PG11, Table 4.2) and 0.2°C (PG10 *vs.* PG13, Table 4.2), respectively. The crystal structure of this DNA aptamer bound to thrombin suggests that T_3 and T_{12} are different from T_4 and T_{13}. It

appears that the bases of T_3 and T_{12} do not interact with any other moieties within the aptamer, nor do they interact with thrombin, instead, they extend out into the solvent.[32] Modifying T_3 and T_{12} with 2'F-araT might change the conformation of the loops, bringing out stabilization to the whole G-quadruplex structure as suggested by the T_m increase. A detailed picture, however, is difficult to describe at this stage. The replacement of the central nucleotide G_8 with a 2'F-araG is also less important for stabilization (PG8 *vs.* PG9, only 0.8°C T_m difference). Overall, 2'F-ANA modifications of loop deoxynucleotides stabilizes the formation of a unimolecular G-quadruplex (Figure 4.1B), supported by all CD and UV melting analysis. It was found that all loop-modified thrombin-binding aptamers (PG8-14) maintain Type II CD in the presence of K^+, like the DNA aptamer control (Figure S4.4, Supplementary Data). Remarkably, in the absence of K^+, a relatively strong positive peak at ~295 nm is still observed, which indicates significant G-quartet formation. Therefore, replacement of the loop deoxynucleotides with 2'F-ANA helps the formation of G-quadruplex structure. They show concentration-independent T_m data, indicating a unimolecular structure, in agreement with a lack of hysteresis in heating/cooling processes (Figure 4.5 and Figure S4.1, Supplementary Data).

Nuclease Resistance Induced by 2'F-ANA

Incorporation of 2'F-ANA in oligonucleotides leads to an improvement in nuclease resistance.[44] This study further evaluates the nuclease resistance of fully 2'F-ANA-modified thrombin-binding aptamers (PG1-14, Table 4.1). The data clearly shows that 2'F-ANA modified aptamers have enhanced half-lives in 10% FBS (in one case > 48 fold compared with PG1, Figure 4.8, Figure S4.6 in Supplementary Data and Table 4.1) but nuclease stability is found to be dependent on both the position (esp. *syn*-G and *anti*-G) and number of 2'F-ANA residues within the oligonucleotide backbone. For example, replacing *anti*-Gs with 2'F-ANA (PG3 and PG4) increases the half-lives to 4.8 and 9.4 h, respectively, while replacing *syn*-Gs with 2'F-ANA (PG5 and PG6) does not yield much gain in biostability with half-life of only 0.8 and 0.6 h, respectively, comparable to the DNA aptamer (PG1) (Figure 4.8). The contrast could be justified by the previous finding that PG5 and PG6 only form insignificant G-quadruplex structures (Figure 4.3D) due to the unfavoured conformation of the perturbed *syn*-Gs. It is expected that a complex

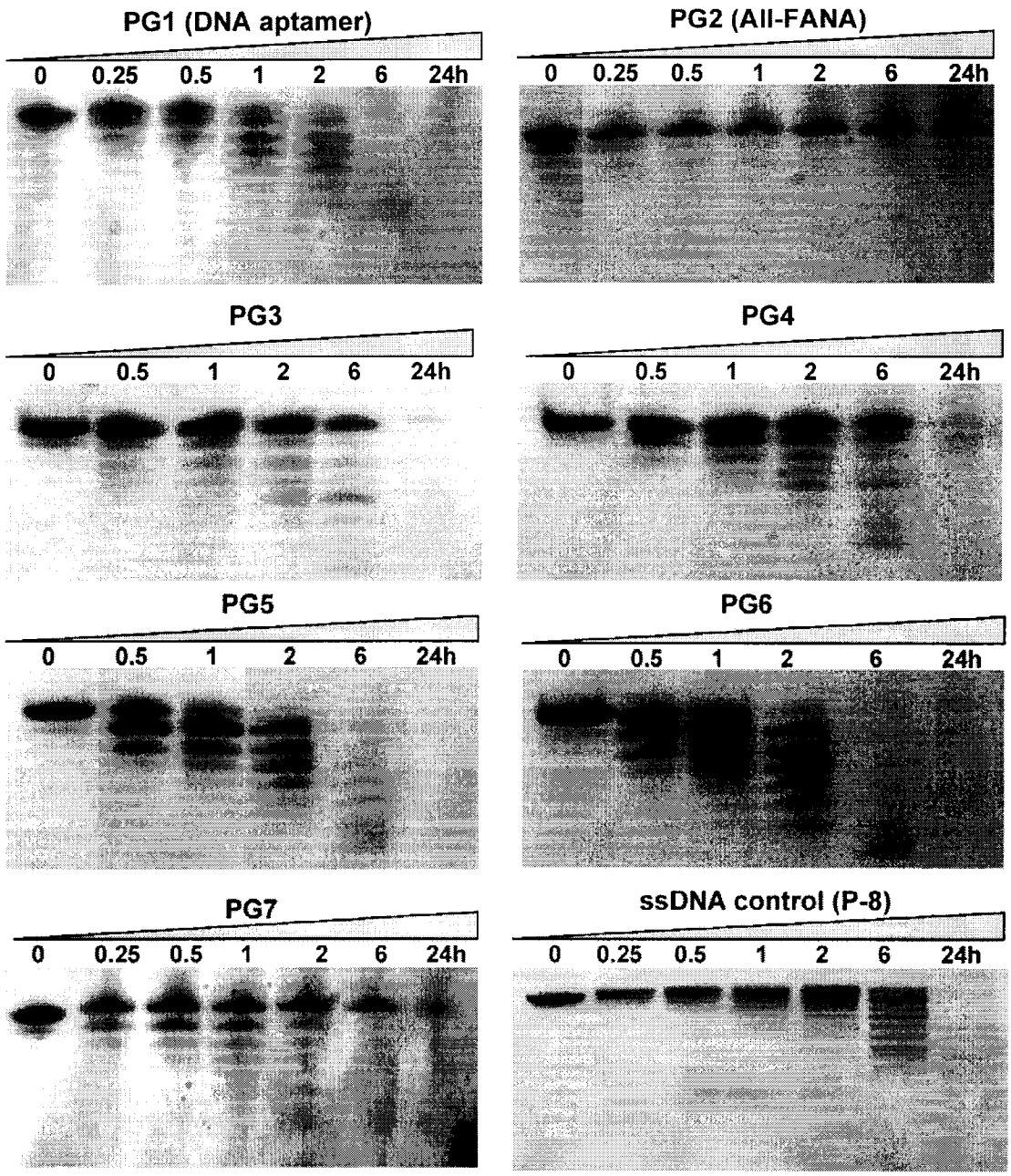


Figure 4.8: Stability of selected FANA modified thrombin-binding aptamers (PG1-7) and ssDNA control to 10% fetal bovine serum (FBS) as monitored by polyacrylamide gel electrophoresis (aptamer sequence in Table 4.2)

structure should be less accessible for nuclease attack compared to dissociated or "open" oligonucleotides. Furthermore, the nuclease stability of PG13 and PG14 is enhanced over PG1 by a factor of 4-7 with only four 2'F-ANA incorporated (Table 4.1). Modification of the loop residues of stable structures (e.g. PG4, PG8) may have a double effect since the

increased stability of the duplex would be combined with protection of the nucleotides most exposed to endonuclease attack.

Binding Affinity of 2'F-ANA-modified Thrombin-binding Aptamers with Thrombin

As mentioned in the introduction, the thrombin-binding aptamer is very sensitive to chemical modifications, which often undermine its thermal stability and/or binding activity.[42]. Thus, a final experiment in this study was conducted using nitrocellulose filter binding assays to assess the effect of 2'F-ANA modifications on the thrombin binding affinity. Data are presented in Table 4.1 and in Figure 4.9. ssDNA (P8) and hairpin (H1) controls show no binding to thrombin as expected.[32] Binding of 2'F-ANA aptamers to thrombin is always adversely affected by 2'F-ANA modifications on G-quartets themselves, whether *syn-* or *anti*-Gs are modified (PG2-PG6). Some loop modifications with 2'F-ANA also disfavor thrombin binding (PG8-PG12). However, two loop-modified aptamers (PG13 and PG14, Table 4.1) show a 4-5 fold enhancement in

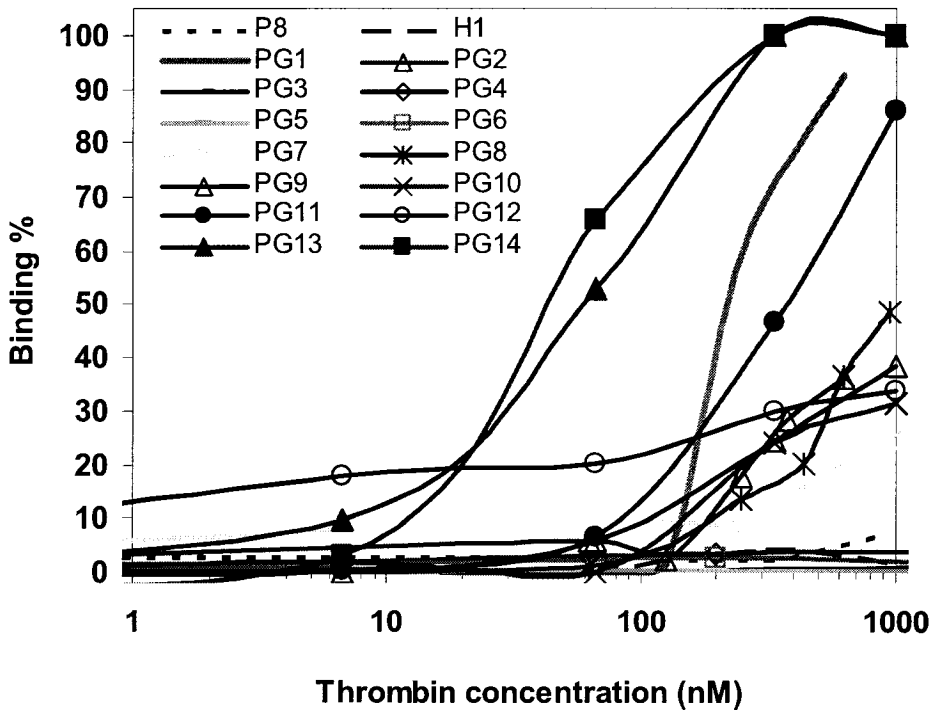


Figure 4.9. Binding affinity with human thrombin assessed by nitrocellulose filter binding assays.

thrombin binding affinity. Therefore, the first 2'F-ANA modified thrombin-binding aptamers to combine enhanced thermal and nuclease stability with stronger binding affinity has been demonstrated. Previous structural studies by NMR and crystallography suggest that the loop moieties would change conformation during binding to thrombin.[32] The TGT loop is involved with the fibrinogen exosite and two TT loops could interact with the heparin exosite. It seems that 2'F-ANA-T modification of one of the TT loops and the two thymidines in the TGT loop enhances thrombin binding.

Conclusions

In the present study 2'F-ANA modified oligonucleotides were investigated based on a thrombin-binding DNA aptamer d($G_2T_2G_2TGTG_2T_2G_2$), an anti-HIV phosphothioate PS-d($T_2G_4T_2$), and a DNA telomeric sequence d($G_4T_4G_4$) (Table 4.1 and Table 4.2) by UV thermal denaturation and CD experiments. Generally, replacement of *anti*-Gs with 2'F-ANA can stabilize a G-quartet requiring *anti*-Gs and maintain the quadruplex conformation, while replacement of *syn*-Gs with 2'F-ANA is not favored and results in complete conformational change of the G-quadruplexes. The data shows that appropriate incorporation of 2'F-ANA residues into G-quadruplexes leads to an increase in the melting temperature of the complex formed (ΔT_m up to ~ + 3°C/ 2'F-ANA modification, Table 4.1 and 4.2). The structure of thrombin-binding aptamers is stabilized by the presence of potassium ions. Nuclease resistance of 2'F-ANA modified thrombin-binding aptamers is increased up to 48 fold in 10% FBS. Two 2'F-ANA-modified thrombin-binding aptamers (PG13 and PG14) show a 4-5 fold enhancement in binding affinity to thrombin along with increased thermal stability and nuclease resistance. Therefore, the impact of 2'F-ANA modifications on G-quartets and loop regions has been demonstrated. The study suggests that 2'F-ANA could have great potential in the application of oligonucleotide-based therapeutics involving the G-quadruplexes.

Acknowledgments

We acknowledge financial support through a grant from the Canadian Institutes of Health Research (CIHR) and Topigen Pharmaceuticals, Inc.. CGP acknowledges support from a FQRNT fellowship and a Clifford Wong McGill Major Fellowship. We thank J. Watts, R.

Donga and J. Lackey for helpful feedback during the preparation of this manuscript.

Supplementary Data

More T_m and CD spectra, stability of FANA modified thrombin-binding aptamers to 10% FBS and MALDI-TOF analysis of thrombin-binding aptamers are available on pp176-184.

References

1. Tuerk, C. and Gold, L. (**1990**). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249, 505-510.

2. Ellington, A.D. and Szostak, J.W. (**1990**). *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346, 818-822.

3. Joyce, G.F. (**1989**). Amplification, mutation and selection of catalytic RNA. *Gene*, 82, 83-87.

4. Nimjee, S.M., Rusconi, C.P. and Sullenger, B.A. (**2005**). Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.*, 56, 555-583.

5. Brody, E.N. and Gold, L. (**2000**). Aptamers as therapeutic and diagnostic agents. *Reviews in Molecular Biotechnology*, 74, 5-13.

6. White, R.R., Sullenger, B.A. and Rusconi, C.P. (**2000**). Developing aptamers into therapeutics. *J. Clin. Invest.*, 106, 929-934.

7. Pestourie, C., Tavitian, B. and Duconge, F. (**2005**). Aptamers against extracellular targets for in vivo applications. *Biochimie*, 87, 921-930.

8. Nolte, A., Klussmann, S., Bald, R., Erdmann, V.A. and Fuerste, J.P. (**1996**). Mirror-design of L-oligonucleotide ligands binding to L-arginine. *Nat. Biotechnol.*, 14, 1116-1119.

9. Lebruska, L.L. and Maher, L.J., III. (**1999**). Selection and characterization of an RNA decoy for rranscription factor NF-kB. *Biochemistry*, 38, 3168-3174.

10. Ruckman, J., Green, L.S., Beeson, J., Waugh, S., Gillette, W.L., Henninger, D.D., Claesson-Welsh, L. and Janjic, N. (**1998**). 2'-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *J. Biol. Chem.*, 273, 20556-20567.

11. Lato, S.M., Ozerova, N.D.S., He, K., Sergueeva, Z., Shaw, B.R. and Burke, D.H. (**2002**). Boron-containing aptamers to ATP. *Nucleic Acids Res.*, 30, 1401-1407.

12. Kato, Y., Minakawa, N., Komatsu, Y., Kamiya, H., Ogawa, N., Harashima, H. and Matsuda, A. (**2005**). New NTP analogs: the synthesis of 4'-thioUTP and 4'-thioCTP and their utility for SELEX. *Nucleic Acids Res.*, 33, 2942-2951.

13. Jhaveri, S., Olwin, B. and Ellington, A.D. (**1998**). *In vitro* selection of phosphorothiolated aptamers. *Bioorg. Med. Chem. Lett.*, 8, 2285-2290.

14. Pagratis, N.C., Bell, C., Chang, Y.-F., Jennings, S., Fitzwater, T., Jellinek, D. and Dang, C. (**1997**). Potent 2'-amino-, and 2'-fluoro-2'-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor. *Nat. Biotechnol.*, 15, 68-73.

15. Ng, E.W.M., Shima, D.T., Calias, P., Cunningham, E.T., Guyer, D.R. and Adamis, A.P. (**2006**). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov*, 5, 123-132.

16. Gellert, M., Lipsett, M.N. and Davies, D.R. (**1962**). Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.*, 48, 2013-&.

17. Sen, D. and Gilbert, W. (**1988**). Formation of parallel 4-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, 334, 364-366.

18. Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. and Toole, J.J. (**1992**). Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature (London, United Kingdom)*, 355, 564-566.

19. Wyatt, J.R., Vickers, T.A., Roberson, J.L., Buckheit, R.W., Klimkait, T., Debaets, E., Davis, P.W., Rayner, B., Imbach, J.L. and Ecker, D.J. (**1994**). Combinatorially selected guanosine-quartet structure is a potent inhibitor of human-immunodeficiency-virus envelope-mediated cell-fusion. *Proc. Natl. Acad. Sci. U. S. A.*, 91, 1356-1360.

20. Phan, A.T., Kuryavyi, V., Ma, J.B., Faure, A., Andreola, M.L. and Patel, D.J. (**2005**). An interlocked dimeric parallel-stranded DNA quadruplex: A potent inhibitor of HIV-1 integrase. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 634-639.

21. Jing, N.J. and Hogan, M.E. (**1998**). Structure-activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug. *J. Biol. Chem.*, 273, 34992-34999.

22. Borman, S. (**2006**). Targeting telomerase. *Chem. Eng. News*, 84, 32-33.

23. Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A. and Feigon, J. (**1993**). Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U. S. A.*, 90, 3745-3749.

24. Wang, K.Y., Mccurdy, S., Shea, R.G., Swaminathan, S. and Bolton, P.H. (**1993**). A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA. *Biochemistry*, 32, 1899-1904.

25. Schultze, P., Smith, F.W. and Feigon, J. (**1994**). Refined solution structure of the dimeric quadruplex formed from the *Oxytricha* telomeric oligonucleotide d(GGGGTTTTGGGG). *Structure*, 2, 221-233.

26. Smith, F.W. and Feigon, J. (**1992**). Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature*, 356, 164-168.

27. Keniry, M.A. (**2000**). Quadruplex structures in nucleic acids. *Biopolymers*, 56, 123-146.

28. Williamson, J.R. (**1994**). G-quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.*, 23, 703-730.

29. Kerwin, S.M. (**2000**). G-quadruplex DNA as a target for drug design. *Curr. Pharm. Des.*, 6, 441-471.

30. Schultze, P., Macaya, R.F. and Feigon, J. (**1994**). Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, 235, 1532-1547.

31. Wang, K.Y., Krawczyk, S.H., Bischofberger, N., Swaminathan, S. and Bolton, P.H. (**1993**).

The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. *Biochemistry*, 32, 11285-11292.

32. Padmanabhan, K., Padmanabhan, K.P., Ferrara, J.D., Sadler, J.E. and Tulinsky, A. (**1993**). The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer. *J. Biol. Chem.*, 268, 17651-17654.

33. Padmanabhan, K. and Tulinsky, A. (**1996**). An ambiguous structure of a DNA 15-mer thrombin complex. *Acta Crystallographica, Section D: Biological Crystallography*, D52, 272-282.

34. Kelly, J.A., Feigon, J. and Yeates, T.O. (**1996**). Reconciliation of the X-ray and NMR structures of the thrombin-binding aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, 256, 417-422.

35. Kang, C., Zhang, X.H., Ratliff, R., Moyzis, R. and Rich, A. (**1992**). Crystal structure of 4 stranded *Oxytricha* telomeric DNA. *Nature*, 356, 126-131.

36. Schultze, P., Hud, N.V., Smith, F.W. and Feigon, J. (**1999**). The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the *Oxytricha nova* telomere repeat oligonucleotide d($G_4T_4G_4$). *Nucleic Acids Res.*, 27, 3018-3028.

37. Hardin, C.C., Perry, A.G. and White, K. (**2000**). Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers*, 56, 147-194.

38. Heckel, A. and Mayer, G. (**2005**). Light regulation of aptamer activity: An anti-thrombin aptamer with caged thymidine nucleobases. *J. Am. Chem. Soc.*, 127, 822-823.

39. Di Giusto, D.A. and King, G.C. (**2004**). Construction, stability, and activity of multivalent circular anticoagulant aptamers. *J. Biol. Chem.*, 279, 46483-46489.

40. Buijsman, R.C., Schipperijn, J.W.J., Kuyl-Yeheskiely, E., Van Der Marel, G.A., Van Boeckel, C.A.A. and Van Boom, J.H. (**1997**). Design and synthesis of a possible mimic of a thrombin-binding DNA aptamer. *Bioorg. Med. Chem. Lett.*, 7, 2027-2032.

41. Rosemeyer, H., Mokrosch, V., Jawalekar, A., Becker, E.-M. and Seela, F. (**2004**). Single-stranded DNA: replacement of canonical by base-modified nucleosides in the minihairpin 5'-d(GCGAAGC)-3' and constructs with the aptamer 5'-d(GGTTGGTGTGGTTGG)-3'. *Helv. Chim. Acta*, 87, 536-553.

42. Sacca, B., Lacroix, L. and Mergny, J.-L. (**2005**). The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. *Nucleic Acids Res.*, 33, 1182-1192.

43. Trempe, J.-F., Wilds, C.J., Denisov, A.Y., Pon, R.T., Damha, M.J. and Gehring, K. (**2001**). NMR solution structure of an oligonucleotide hairpin with a 2'F-ANA/RNA stem: Implications for RNase H specificity toward DNA/RNA hybrid duplexes. *J. Am. Chem. Soc.*, 123, 4896-4903.

44. Kalota, A., Karabon, L., Swider, C.R., Viazovkina, E., Elzagheid, M., Damha, M.J. and

Gewirtz, A.M. (**2006**). 2'-Deoxy-2'-fluoro-beta-D-arabinonucleic acid (2'F-ANA) modified oligonucleotides (ON) effect highly efficient, and persistent, gene silencing. *Nucleic Acids Res.*, 34, 451-461.

45. Wilds, C.J. and Damha, M.J. (**2000**). 2'-Deoxy-2'-fluoro-β-D-arabinonucleosides and oligonucleotides (2'F-ANA): synthesis and physicochemical studies. *Nucleic Acids Res.*, 28, 3625-3635.

46. Wilds, C.J. and Damha, M.J. (**1999**). Duplex recognition by oligonucleotides containing 2'-deoxy-2'-fluoro-D-arabinose and 2'-deoxy-2'-fluoro-D-ribose. Intermolecular 2'-OH-phosphate contacts versus sugar puckering in the stabilization of triple-helical complexes. *Bioconjug. Chem.*, 10, 299-305.

47. Wilds, C.J. (**2000**). Ph.D. thesis, McGill University, Montreal.

48. Viazovkina, E.V., Mangos, M.M., Elzagheid, M.I. and Damha, M.J. (**2002**). Solid-phase synthesis of 2'-deoxy-2'-fluoro-α-D-oligoarabinonucleotides (2'F-ANA) and their phosphorothioate derivatives. *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons, Inc., pp. 4.15.11.

49. Smirnov, I. and Shafer, R.H. (**2000**). Effect of loop sequence and size on DNA aptamer stability. *Biochemistry*, 39, 1462-1468.

50. Wyatt, J.R., Davis, P.W. and Freier, S.M. (**1996**). Kinetics of G-quartet-mediated tetramer formation. *Biochemistry*, 35, 8002-8008.

51. Lu, M., Guo, Q. and Kallenbach, N.R. (**1993**). Thermodynamics of G-tetraplex formation by telomeric DNAs. *Biochemistry*, 32, 598-601.

52. Cantor, C.R., Warshaw, M.M. and Shapiro, H. (**1970**). Oligonucleotide interactions. 3. Circular dichroism studies of conformation of deoxyoligonucleotides. *Biopolymers*, 9, 1059-1077.

53. Galarneau, A., Min, K.-L., Mangos, M.M. and Damha, M.J. (**2005**). Assay for evaluating ribonuclease H-mediated degradation of RNA-antisense oligonucleotide duplexes. *Methods Mol. Biol.*, Totowa, NJ, United States, Vol. 288, pp. 65-80.

54. Mergny, J.L., Phan, A.T. and Lacroix, L. (**1998**). Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, 435, 74-78.

55. Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J.O. and Bates, P.J. (**2003**). Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Res.*, 31, 2097-2107.

56. Lu, M., Guo, Q. and Kallenbach, N.R. (**1992**). Structure and stability of sodium and potassium complexes of dT_4G_4 and dT_4G_4T. *Biochemistry*, 31, 2455-2459.

57. Aboulela, F., Murchie, A.I.H. and Lilley, D.M.J. (**1992**). NMR-study of parallel-stranded tetraplex formation by the hexadeoxynucleotide $d(TG_4T)$. *Nature*, 360, 280-282.

58. Wang, Y. and Patel, D.J. (**1992**). Guanine residues in $d(T_2AG_3)$ and $d(T_2G_4)$ form parallel-stranded potassium cation stabilized G-quadruplexes with antiglycosidic torsion angles in

solution. *Biochemistry*, 31, 8112-8119.

59. Sapse, A.M. and Snyder, G. (**1985**). Ab initio studies of the antiviral drug 1-(2-fluoro-2-deoxy-beta-D-arabinofuranosyl) thymine. *Cancer Invest.*, 3, 115-121.

60. Koshkin, A.A., Nielsen, P., Meldgaard, M., Rajwanshi, V.K., Singh, S.K. and Wengel, J. (**1998**). LNA (locked nucleic acid): An RNA mimic forming exceedingly stable LNA : LNA duplexes. *J. Am. Chem. Soc.*, 120, 13252-13253.

61. Peng, C.G. and Damha, M.J. (**2005**). Synthesis and hybridization studies of oligonucleotides containing 1-(2-deoxy-2-α-C-hydroxymethyl-β-D-ribofuranosyl)thymine (2'-α-hm-dT). *Nucleic Acids Res.*, 33, 7019-7028.

Chapter IV. G-quadruplex Induced Stabilization by 2'-Deoxy-2'-fluoro-β-D-arabinonucleic Acids (2'F-ANA)

Supplementary Data
(pp176-184)

Contents

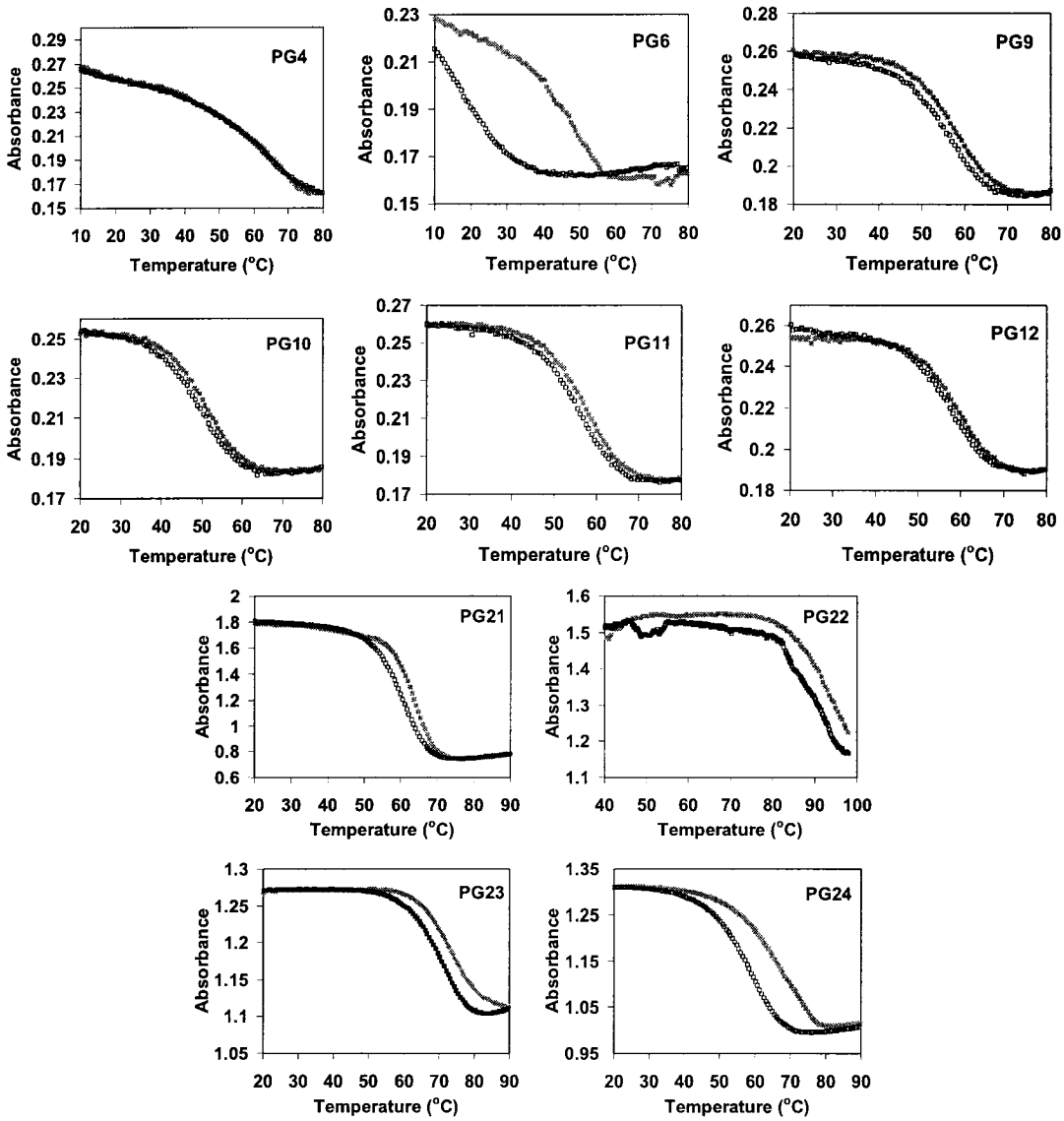


Figure S4.1: Heating and cooling T_m transitions of selected FANA modified thrombin-binding aptamers (Table 4.1) measured at 295 nm in a heating/cooling rate of 0.5°C in a buffer of 10 mM Tris, 25 mM KCl, pH 6.8, at a strand concentration of 8 μM; and $dG_4T_4G_4$ related sequences PG21-24 (Table 4.2) in 10 mM sodium phosphate buffer, 0.1 mM EDTA, pH 7 and 200 mM NaCl at a strand concentration of 100 μM. (red filled square: heating; blue empty square: cooling).

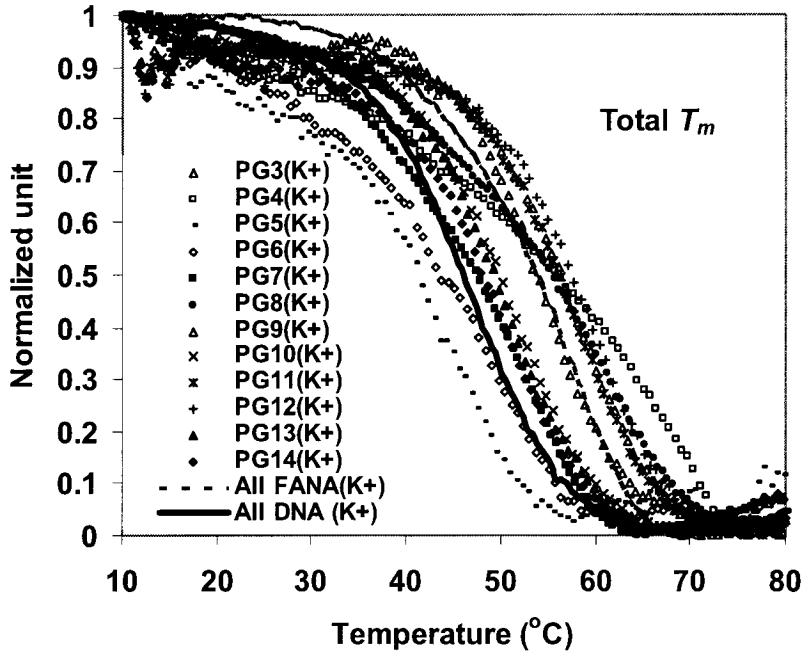


Figure S4.2: Normalized T_m of total FANA modified thrombin-binding aptamers (PG1-14 in Table 4.1) measured at 295 nm in buffer 10 mM Tris, pH 6.8, 25 mM KCl, at a final strand concentration of 8 μM (also refer to Figure 4.3) and T_m data were presented in Table 4.1.

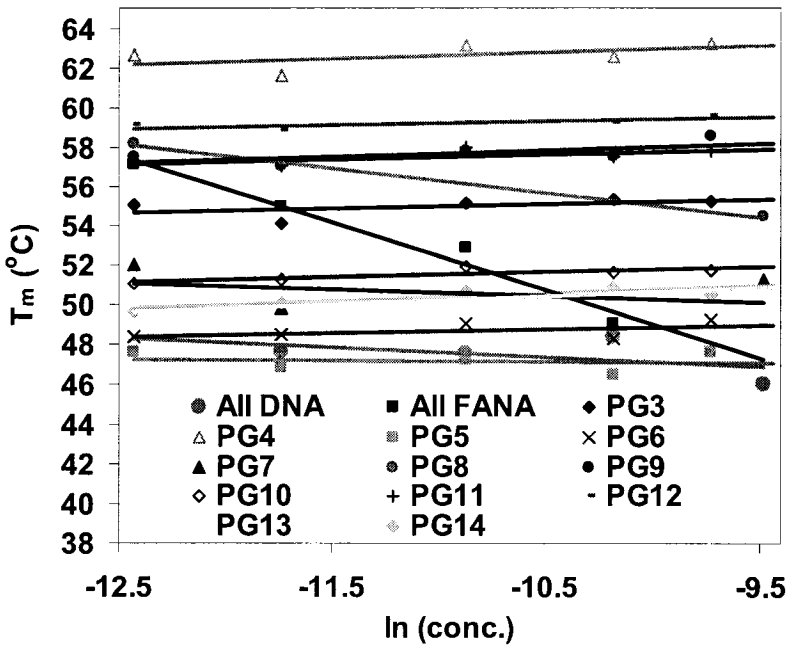


Figure S4.3: T_m versus concentration dependence (15-20 folds) study measured at 295 nm in a buffer consisting of 10 mM Tris, pH 6.8, 25 mM KCl for all FANA modified thrombin-binding aptamers (PG1-14, Table 4.1).

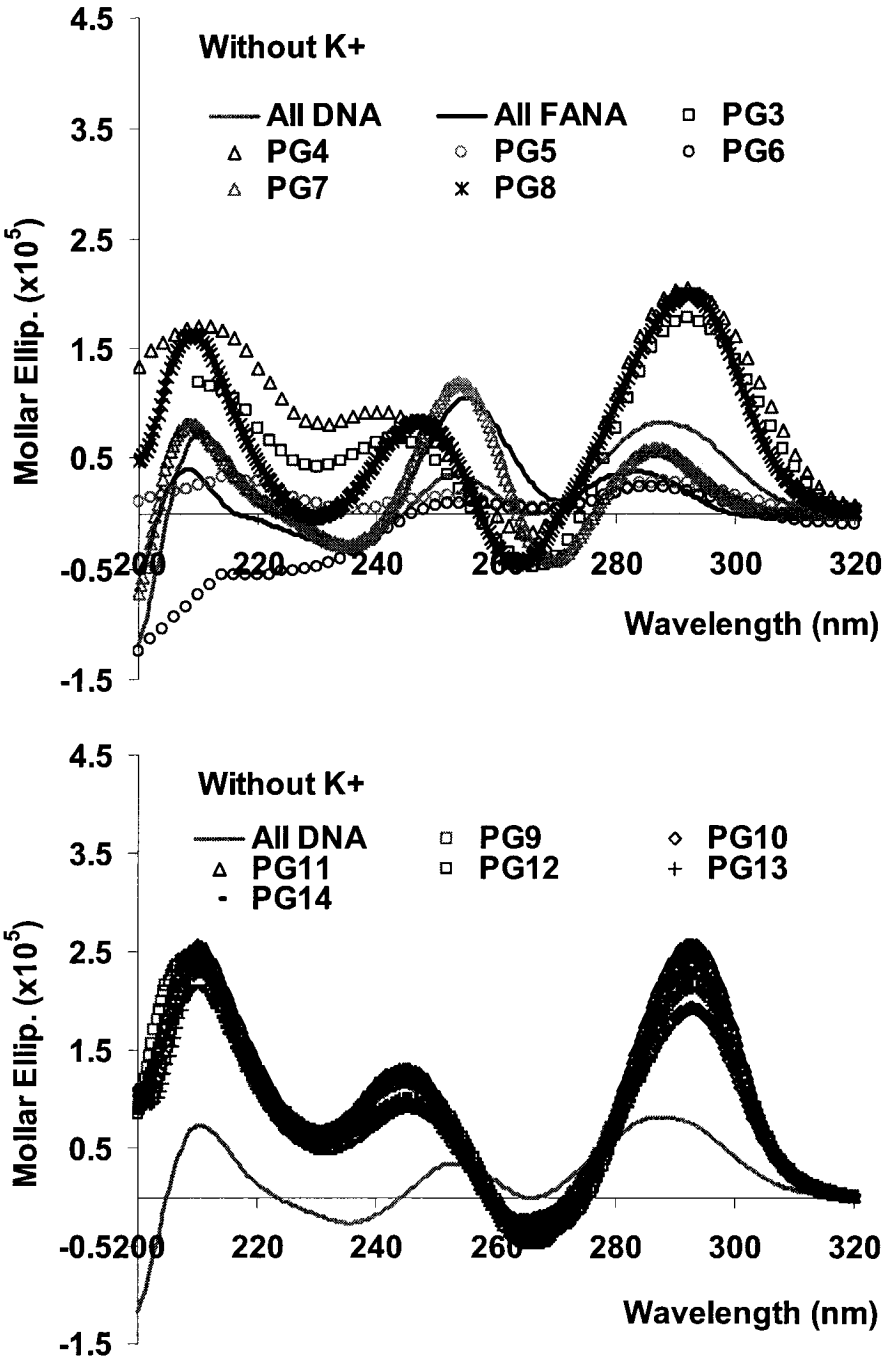

Without K+
All DNA
All FANA
PG3
PG4
PG5
PG6
PG7
PG8
Mollar Ellip. (x10^5)
Wavelength (nm)
Without K+
All DNA
PG9
PG10
PG11
PG12
PG13
PG14
Mollar Ellip. (x10^5)
Wavelength (nm)

Figure S4A

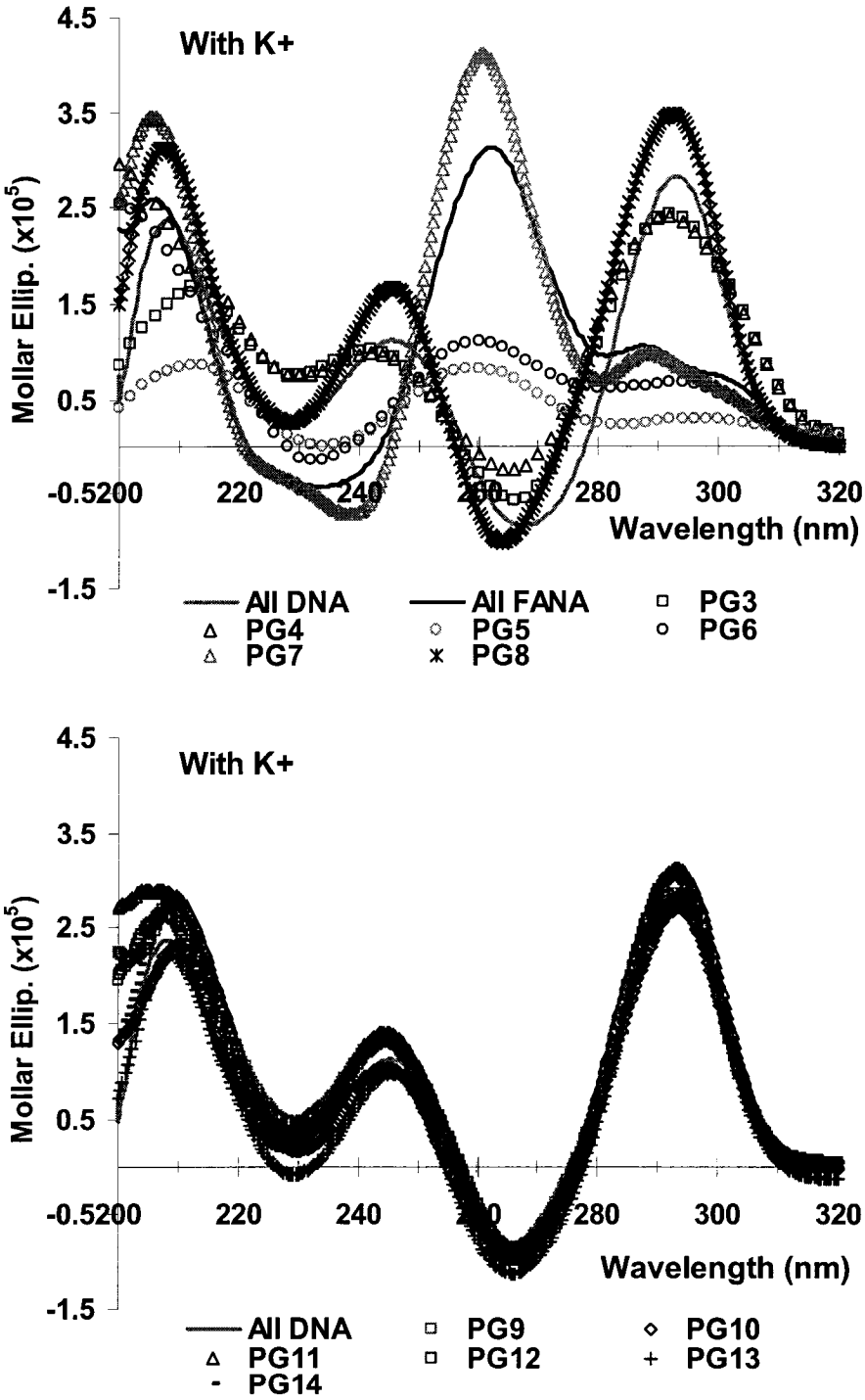


Figure S4B

Figure S4. CD spectra of FANA modified thrombin-binding aptamers (PG1-14, Table 4.1) conducted in the buffer consisting of 10 mM Tris, pH 6.8 (**A**) without KCl; (**B**) with 25mM KCl at 15°C at a final strand concentration of 8 μM.

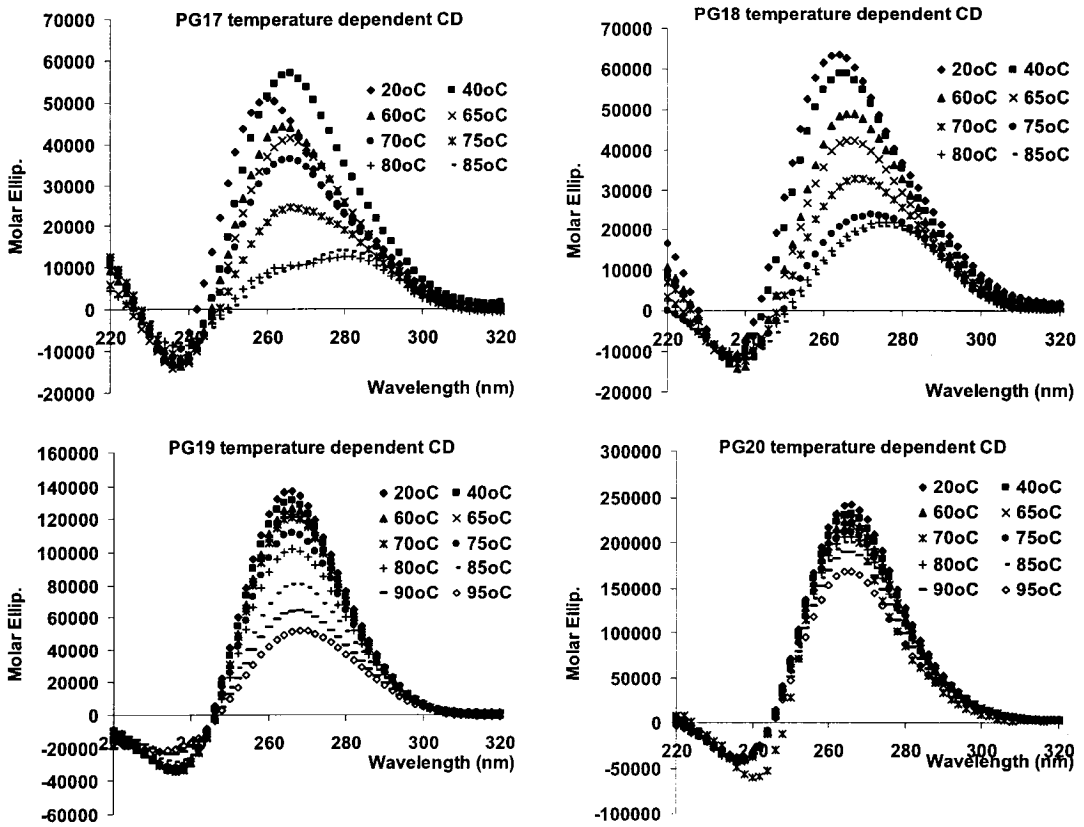


Figure S4.5. Temperature-dependent CD spectra for PG17-20 (Table 4.2) and resulting T_m curves for each complex. PBS buffer, 137 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.2 at 25°C; strand concentration: 20 μM; 10 min equilibrium time was set at each temperature studied. The resulting T_m curves were generated by plotting the maximum absorbance at 265 nm wavelength (normalized) *vs* temperature and shown in Figure 6; the corresponding T_m data are shown in Table 4.2.

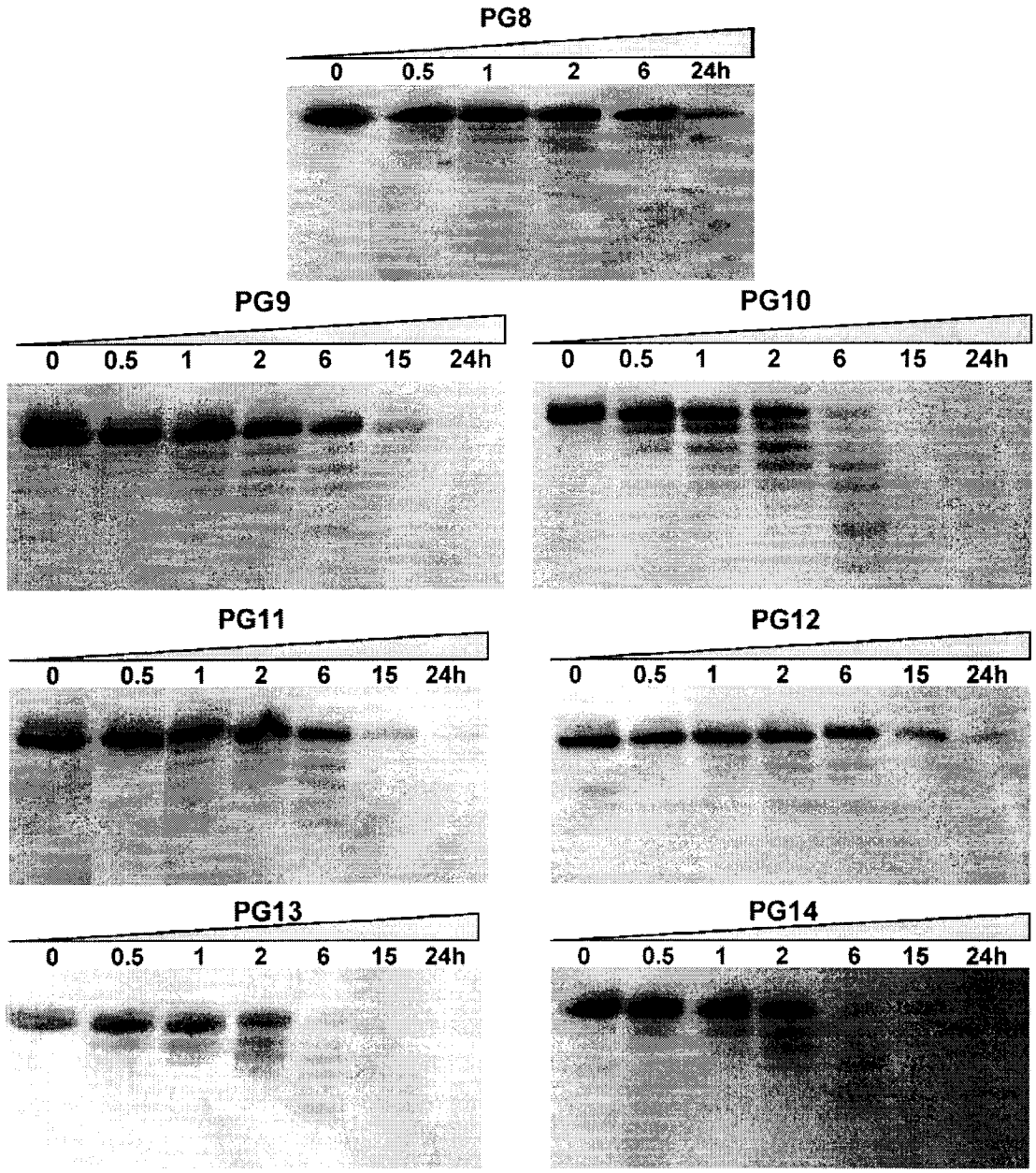

PG8
0 0.5 1 2 6 24h
PG9
0 0.5 1 2 6 15 24h
PG10
0 0.5 1 2 6 15 24h
PG11
0 0.5 1 2 6 15 24h
PG12
0 0.5 1 2 6 15 24h
PG13
0 0.5 1 2 6 15 24h
PG14
0 0.5 1 2 6 15 24h

Figure S4.6-A

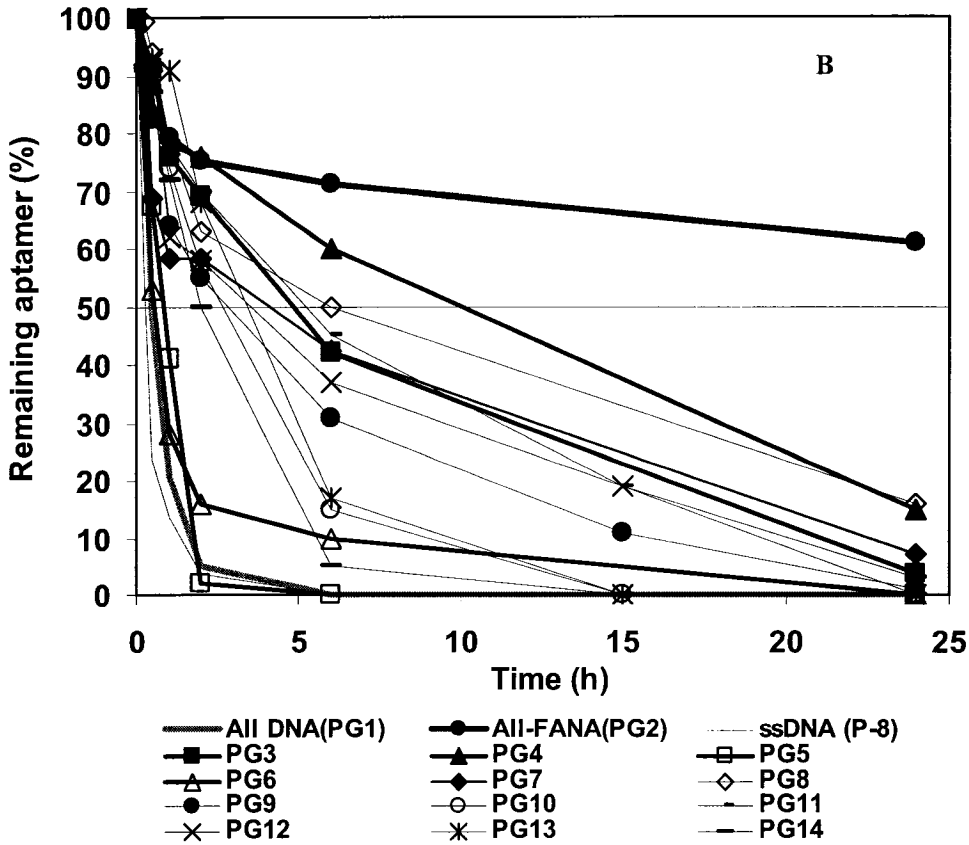


Figure S4.6: (A) Stability of FANA modified thrombin-binding aptamers (PG1-14, Table 4.1) to 10% fetal bovine serum (FBS) as monitored by polyacrylamide gel electrophoresis (time points: 0, 0.25, 0.5, 1, 2, 6, 15h, 24h). **(B)** Stability curve of aptamers to 10% FBS as monitored by PAGE; half life of these aptamers ($t_{1/2}$) is calculated by 50% remaining line and presented in Table 4.1.

Table S4.1: MALDI-TOF analysis of thrombin-binding aptamers (PG1-PG14)

New code	Type	Sequence	Calc. Mass	MALDI
PG1	All DNA	d(GGTTGGTGTGGTTGG)	4726	4726
PG2	All 2'F-ANA	d(**GGTTGGTGTGGTTGG**)	4996	5000
PG3	2'F-ANA G-*anti*	d(G**G**TTG**G**TGTG**G**TTG**G**)	4798	4801
PG4	2'F-ANA G-*anti*& loop	d(G**GTTG**G**TGT**G**GTTGG**)	4906	4905
PG5	2'F-ANA G-*syn*	d(**G**GTT**G**GTGT**G**GTT**G**G)	4798	4797
PG6	2'F-ANA G-*syn*&loop	d(**GGTTG**G**TGTG**G**TTGG**)	4906	4910
PG7	2'F-ANA G-quartet	d(**GG**TT**GG**TGT**GG**TT**GG**)	4870	4871
PG8	2'F-ANA all-loop	d(GG**TT**GG**TGT**GG**TT**GG)	4852	4854
PG9	2'F-ANA loop	d(GG**TT**GG**T**GT**T**GG**TT**GG)	4834	4837
PG10	2'F-ANA loop	d(GGT**T**GG**T**GT**T**GG**TT**GG)	4816	4818
PG11	2'F-ANA loop	d(GG**TT**GG**T**GT**T**GGTTGG)	4816	4818
PG12	2'F-ANA loop	d(GG**TT**GGTGTGG**TT**GG)	4798	4792
PG13	2'F-ANA loop	d(GGTTGG**T**GT**T**GG**TT**GG)	4798	4798
PG14	2'F-ANA loop	d(GG**TT**GG**T**GT**T**GGTTGG)	4798	4796

Chapter V

In Chapter IV, we demonstrated that 2'-deoxy-2'-fluoro-β-D-arabinonucleotides (2'F-araN) can be introduced into the DNA aptamers leading to FANA-DNA chimeras with enhanced thermal stability, nuclease resistance and in some cases improved binding affinity to their intended protein target (*i.e.*, thrombin). This suggests that FANA-DNA aptamers could find use as therapeutic and diagnostic agents. An alternative and more practical way to generate chemically-modified aptamers is through *Systematic Evolution of Ligands by EXponential enrichment* (SELEX), a technique based on the enzymatic synthesis of oligonucleotides starting from chemically modified nucleoside 5'-triphosphates and the native building blocks, namely, deoxyribonucleoside 5'-triphosphates (dNTPs) or ribonucleoside 5'-triphosphates (rNTPs). Due to the high substrate specificity exhibited by DNA (or RNA) polymerases, only a limited number of nucleoside modifications are compatible with the SELEX process. This chapter examines, for the first time, the ability of DNA polymerases to carry out template-directed synthesis of 2'-deoxy-2'-fluoro-β-D-arabinonucleic acids (2'F-ANA).

Reprint from Peng, C.G. and Damha, M.J. (2006) Polymerase-directed synthesis of 2'-deoxy-2'-fluoro-β-D-arabinonucleic acids. *Journal of the American Chemical Society.* (on-line publication on April 10, 2007; doi:10.1021/ja069100g). With permission from ACS Publications.

Chapter V. Polymerase-Directed Synthesis of 2′-Deoxy-2′-Fluoro-β-D-Arabino-Nucleic Acids (2'F-ANA)

Abstract

We have examined the ability of 2'-deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs) to serve as substrates for various DNA polymerases, thereby rendering these modified nucleoside 5'-triphosphates (NTPs) amenable to DNA aptamer selection. We adapted primer extension assays to screen nine DNA polymerases for their ability to incorporate all four 2'F-araNTPs. We found that 2'F-araNTPs are very good substrates for thermophilic DNA polymerases Deep Vent (3'->5' exo-) (DV), $9°N_m^{TM}$ DNA polymerase (9N), Therminator™ (Th) and Phusion™ High Fidelity DNA polymerase (Ph), all of which effectively and selectively incorporated 2'F-araNTPs to yield full-length products. DV, 9N and Ph were shown to synthesize FANA-DNA strands on FANA-DNA templates, and Klenow fragment DNA polymerase (3'->5' exo-) (KF) and *Bst* DNA polymerase (*Bst*) catalyzed FANA template-directed DNA synthesis. While it was not possible to synthesize an all-FANA strand on an all-FANA template, it was possible for DV, 9N and Ph polymerase to drive the formation of multiple FANA:FANA base pairs within a DNA-FANA chimeric duplex. All of these results suggest that it should be possible to evolve FANA-modified aptamers via SELEX.

Introduction

Systematic evolution of ligands by exponential enrichment (SELEX)[1,2] is a powerful method for generating oligonucleotide aptamers with therapeutic utility, particularly when modified nucleoside 5'-triphosphates (NTPs) are used for the selection of aptamers with nuclease-resistant capability.[3] Two key requirements must be met for the effective utilization of modified NTPs in the SELEX process; that is, polymerases must recognize and incorporate the modified nucleoside 5'-triphosphates as building blocks with high fidelity; and the resulting products must be faithfully amplified into cDNA libraries. In addition to the most commonly used 2'-fluoro and 2'-amino rNTPs,[3] only a handful of other modified rNTPs[4-7] and even fewer dNTPs[8] are available for aptamer selection.

2-Fluoro-D-arabinose confers a DNA-like (South/East) conformation[9] to the oligonucleotide while rendering it more nuclease resistant.[10] In addition, 2-fluoro-D-arabinose raises the T_m of duplexes (ca. +0.5-1 °C/nt),[11] triplexes (ca. +0.8 °C/nt),[12] and C-rich (ca. +1

°C/nt, pH < 4.0),[13] and G-rich tetraplexes (ca. +2 °C/nt). (Peng, C. G.; Damha, M.J., unpublished results.) Furthermore, DNA or RNA containing FANA units exhibit gene silencing efficacy (via mRNA targeting) in the low nanomolar range.[14] For all of these reasons, we hypothesized that FANA-modified strands could also have considerable potential as aptamers, and we sought to test this hypothesis by first examining the ability of 2'-deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs) to serve as substrates for various DNA polymerases, thereby rendering these modified NTPs amenable to DNA aptamer selection. To our knowledge, there are no reports on the biosynthesis of FANA in vitro, though 2'F-araNTP derivatives have been studied as potential chain terminators for antiviral and anticancer applications[15] and positron emission tomography.[16]

Results and Discussion

We adapted primer extension assays[7,17] to first screen various DNA polymerases[18] for their ability to incorporate 2'F-araNTPs (N = A, G, T and C, Figure 5.1). Family B polymerases (DV, 9N, Th, Ph) effectively incorporated all four 2'F-araNTPs to yield full-length FANA products (Figure S5.2, Supporting Information), whereas family A polymerases (*Bst*, *Taq*, KF) and MMLV generated premature products containing only two to four 2'F-araNTP residues (for names of enzymes, see note [19]). The divergence in DNA polymerase classification and structure[18] and NTP discrimination has been observed with other modified nucleotides, for example, ddNTPs and acyclic NTPs,[20] suggesting that significant differences exist in the active site of these two families (A and B) of enzymes.

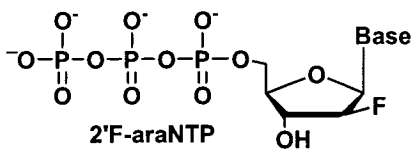


Figure 5.1: Chemical structure of 2'F-araNTPs (*N* = A, G, T and C).

Performing SELEX using modified NTPs requires necessary fidelity in the polymerization reaction so that the selection of functional sequences can be faithfully maintained and enriched from one round to another. Thus, we next conducted "dropout assays"[21] to assess *apparent* fidelity of 2'F-araNTP incorporation by DV and 9N. In one of these assays, 2'F-araATP was removed from the pool, and the ensuing synthesis assessed in comparison with a control reaction containing all four 2'F-araNTPs. For comparison, dropout experiments containing dNTPs were run in parallel and the results obtained are shown in Figure 5.2. Full-length DNA and FANA products were obtained when all four dNTPs (group 1 and 5) and 2'F-

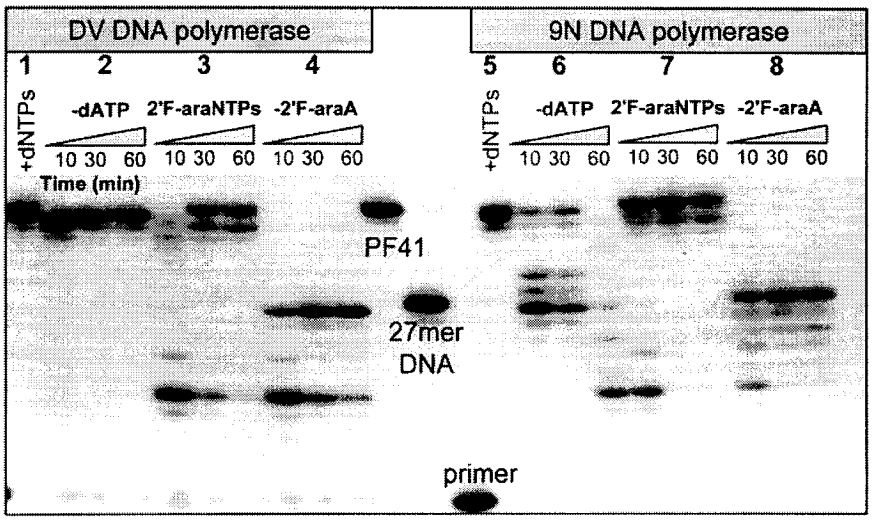


Figure 5.2: Fidelity study on 2'F-araATP incorporation by DV and 9N DNA polymerases on DNA template PF41. Experimental details about this dropout assay are provided in the supporting material. Reaction conditions: 0.8 unit DV and 9N, 25 μM triphosphates at 55oC in 30 μL reaction volume; incubation time: 30 min. Group 1 & 5: dNTPs; 2 & 6: dropout of dATP; 3 & 7: 2'F-araNTPs, 4 & 8: dropout of 2'F-araATP; 27mer DNA is PF39 in Table S5.1, Supporting Information.

araNTPs (groups 3 and 7) were available, although some pausing was evident during synthesis of the arabinose modified oligomers. Compared with dATP, both DV and 9N exhibited excellent selectivity on 2'F-araATP, as demonstrated by efficient chain termination at the position where 2'F-araATP was required (groups 4 and 8). In contrast, dropping out dATP in these assays (groups 2 and 6) afforded full-length in addition to premature products (Table 5.1), suggesting that 2'F-araATP is more strictly selected than dATP under these conditions.

Next, we applied the same dropout experiments to the remaining three triphosphates, namely, 2'F-araGTP, 2'F-araCTP and 2'F-araTTP (Table 5.1; Figures S5.3-S5.4, Supporting Information). Again, DV and 9N demonstrated higher fidelity with respect to 2'F-araGTP incorporation relative to the native dGTP. Selectivity toward the pyrimidine-based 2'F-araNTPs was significantly less, but the same was true for the corresponding pyrimidine-containing dNTPs. Factors governing fidelity of base substitutions by polymerases are complex and not completely understood.[22] Preferential geometry,[23] sugar pucker,[24] and topology and size of the polymerase-active site[23] are all important considerations. In addition, it has also been suggested that dropout essays underestimate the fidelity of base substitution because there would be more opportunity for NTP misincorporation when the normal competition between the correct

and incorrect NTP is lacking in these experiments.[21] Regardless of the mechanisms that operate, the selectivity for 2'F-araNTP incorporation by DV and 9N polymerases appears to be comparable, if not better, than that of dNTPs.

Next, we examined Phusion DNA polymerase (Ph), an enzyme with higher processing capacity and fidelity compared to other thermophilic polymerases (e.g., the error rates for Ph[25] and DV[26] are 4.4×10^{-7} and 2.2×10^{-4}, respectively). All four dropout experiments were conducted with the same DNA template PF21 (Figure 5.3), or PF41 (Figure S5.5, Supporting Information). For the dNTPs, syntheses stopped at the expected termination sites, yielding products lacking the required dNTP at the 3'-terminus. A similar pattern was observed with 2'F-araNTPs, although in the latter, FANA synthesis proceeded with more difficulty (Figure 5.3;

Primer (PF20) 5'-^{32}P-*TAATACGACTCACTATA*-3'
DNA template (PF21) 5'-*ATTATGCTGAGTGATAT*-CCCTCTTCTCACCGTTT-5'
Expected product: ↓ **Ph, triphosphates condition 1-5**
1. 2'F-araNTPs (**All**) 5'-^{32}P-TAATACGACTCACTATA-**GGGAGAAGAGTGGCAAA**-3'
2. Drop out **2'F-araCTP (AGT)** 5'-^{32}P-TAATACGACTCACTATA-**GGGAGAAGAGTGG**-3'
3. Drop out **2'F-araTTP (AGC)** 5'-^{32}P-TAATACGACTCACTATA-**GGGAGAAGAG**-3'
4. Drop out **2'F-araATP (CTG)** 5'-^{32}P-TAATACGACTCACTATA-**GGG**-3'
5. Drop out **2'F-araGTP (CTA)** 5'-^{32}P-TAATACGACTCACTATA-3'

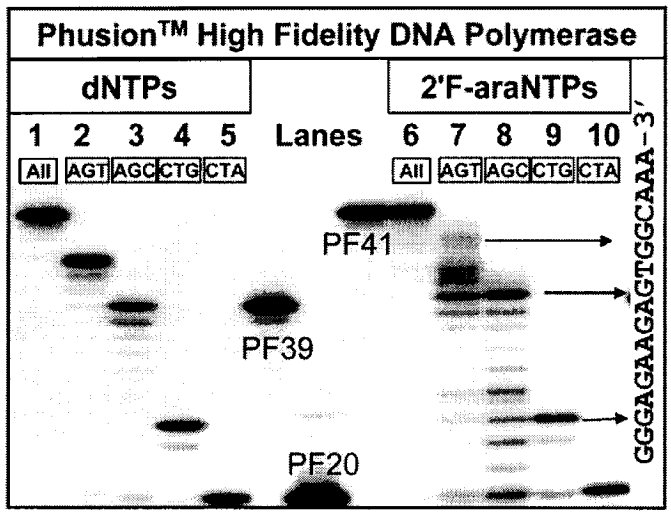


Figure 5.3: Fidelity study on 2'F-araNTP incorporation by Ph on DNA template PF21. Reaction conditions: 0.8 unit Ph, 25 μM triphosphate at 55°C in 30 μL reaction volume for 30min. lane 1: dNTPs; 2: drop out dCTP (i.e. dATP, dGTP and dTTP, shorten as AGT; same coding for lane 3-5), lane 6: 2'F-araNTPs; 7: drop out 2'F-araCTP (i.e. 2'F-araATP, 2'F-araGTP, 2'F-araTTP, shorten as AGT; same coding for lane 8-10). Template PF41 and primer PF20 are shown; PF39 is a 27-nt DNA control.

Table 5.1. Apparent fidelity of FANA synthesis in dropout assay

Polymerase	% Apparent Fidelity[a]		Template used	Figure
	2'F-araNTP	dNTP		
Ph	A: >99%	>99%	PF21	5.3
	G: >99%	>99%		
	T: >99%	>99%		
	C: >99%	>99%		
DV	A: >99%	1%	PF41	5.2
	G: >99%	62%	PF43	S5.3
	T: 3%	19%	PF21	S5.4
	C: 50%	3%	PF23	S5.4
9N	A: >99%	80%	PF41	5.2
	G: 98%	84%	PF43	S5.3
	T: 76%	32%	PF21	S5.4
	C: 43%	82%	PF23	S5.4

[a] Apparent fidelity is calculated according to the equation 1-[(% full-length - 2'F-araNTP)/ (% full-length + 2'F-araNTP)] at 30 min reaction time (sample equation for dNTP values); >99% means that the full-length product is not detectable in our dropout assays.

lanes 6-10). This effect, however, was observed only in the dropout assays, that is, the reaction containing all four 2'F-araNTPs produced excellent yields of the full-length FANA product with virtually no pausing observed. We speculate that in a dropout assay, where polymerization activity is suboptimal, the pausing bands observed may result, at least in part, by the shuttling (cycling) between the strong Ph 3'-5' exonuclease and polymerase activities.[18],[27] Similar results were observed with the DNA template PF41 (Figure S5.5). Of note, apparent fidelity among Ph, DV and 9N is correlated to 3'-5' exonuclease proofreading activity, with Ph having the strongest exonuclease activity and also 2'F-araNTP fidelity, and DV having no exonuclease activity and low 2'F-fidelity. We recently became aware of Wengel's work that Ph can incorporate a few LNA triphosphates in a growing DNA strand.[28] Thus, FANA and LNA NTPs appear to be the first examples of modified NTPs incorporated by Ph.

We are currently testing the limits of these procedures by exploring the synthesis of (a) DNA-FANA on a DNA-FANA template, and (b) DNA on an all-FANA template. Preliminary results obtained indicate that DV, 9N and Ph can synthesize FANA-DNA strands on FANA-DNA templates (Figure S5.6, Supporting Information), and that KF and *Bst* DNA can catalyze FANA template-directed DNA synthesis (Figure S5.7, Supporting Information). While we have not been able to synthesize an all-FANA strand on an all-FANA template, it is possible for DV,

9N and Ph polymerase to drive the formation of multiple FANA:FANA base pairs within a DNA-FANA chimeric duplex (Figure S5.6; Supporting Information).

In summary, family B thermophilic DNA polymerases such as DV, 9N and Ph can utilize 2'F-araNTPs, or a combination of 2'F-araNTPs and dNTPs, to generate FANA or chimeric FANA-DNA strands, respectively. In addition, these polymerases were shown to synthesize chimeric FANA-DNA strands on a chimeric FANA-DNA template. KF and *Bst* (family A) DNA polymerases were able to incorporate dNTPs on a template FANA strand. These results suggest that it should be possible to evolve FANA-modified aptamers via SELEX.

Acknowledgement. This work was supported by the Canadian Institutes of Health Research and Topigen Pharmaceuticals, Inc. Dedicated to Prof. Kelvin K. Ogilvie on the occasion of his 65th birthday.

Supporting Information Available: Experimental details and gel pictures (Figures S5.2-S5.7). This material is available free of charge via the Internet at http://pubs.acs.org (pp194-203).

References

1. Ellington, A.D. and Szostak, J.W. (**1990**). I*n vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346, 818-822.

2. Tuerk, C. and Gold, L. (**1990**). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249, 505-510.

3. Ruckman, J., Green, L.S., Beeson, J., Waugh, S., Gillette, W.L., Henninger, D.D., Claesson-Welsh, L. and Janjic, N. (**1998**). 2'-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *J. Biol. Chem.*, 273, 20556-20567.

4. Lato, S.M., Ozerova, N.D.S., He, K., Sergueeva, Z., Shaw, B.R. and Burke, D.H. (**2002**). Boron-containing aptamers to ATP. *Nucleic Acids Res.*, 30, 1401-1407.

5. Jhaveri, S., Olwin, B. and Ellington, A.D. (**1998**). *In vitro* selection of phosphorothiolated aptamers. *Bioorg. Med. Chem. Lett.*, 8, 2285-2290.

6. Dineva, M.A., Ivanov, I.G. and Petkov, D.D. (**1997**). P-alpha-methyl deoxynucleoside triphosphates as substrates for *E. coli* DNA polymerase I in a template-directed synthesis of DNA. *Nucleosides Nucleotides*, 16, 1875-1882.

7. Kato, Y., Minakawa, N., Komatsu, Y., Kamiya, H., Ogawa, N., Harashima, H. and Matsuda, A. (**2005**). New NTP analogs: the synthesis of 4'-thioUTP and 4'-thioCTP and their utility for SELEX. *Nucleic Acids Res.*, 33, 2942-2951.

8. Kuwahara, M., Hanawa, K., Ohsawa, K., Kitagata, R., Ozaki, H. and Sawai, H. (**2006**). Direct PCR amplification of various modified DNAs having amino acids: Convenient preparation of DNA libraries with high-potential activities for in vitro selection. *Bioorg. Med. Chem.*, 14, 2518-2526.

9. Trempe, J.-F., Wilds, C.J., Denisov, A.Y., Pon, R.T., Damha, M.J. and Gehring, K. (**2001**). NMR solution structure of an oligonucleotide hairpin with a 2'F-ANA/RNA stem: implications for RNase H specificity toward DNA/RNA hybrid duplexes. *J. Am. Chem. Soc.*, 123, 4896-4903.

10. Kalota, A., Karabon, L., Swider, C.R., Viazovkina, E., Elzagheid, M., Damha, M.J. and Gewirtz, A.M. (**2006**). 2'-Deoxy-2'-fluoro-beta-D-arabinonucleic acid (2'F-ANA) modified oligonucleotides (ON) effect highly efficient, and persistent, gene silencing. *Nucleic Acids Res.*, 34, 451-461.

11. Wilds, C.J. and Damha, M.J. (**2000**). 2'-Deoxy-2'-fluoro-β-D-arabinonucleosides and oligonucleotides (2'F-ANA): synthesis and physicochemical studies. *Nucleic Acids Res.*, 28, 3625-3635.

12. Wilds, C.J. and Damha, M.J. (**1999**). Duplex recognition by oligonucleotides containing 2'-deoxy-2'-fluoro-D-arabinose and 2'-deoxy-2'-fluoro-D-ribose. Intermolecular 2'-OH-phosphate contacts versus sugar puckering in the stabilization of triple-helical complexes. *Bioconjug. Chem.*, 10, 299-305.

13. Wilds, C.J. (**2000**). Ph.D. thesis, McGill University, Montreal.

14. Dowler, T., Bergeron, D., Tedeschi, A.L., Paquet, L., Ferrari, N. and Damha, M.J. (**2006**). Improvements in siRNA properties mediated by 2'-deoxy-2'-fluoro-beta-D-arabinonucleic acid (FANA). *Nucleic Acids Res.*, 34, 1669-1675.

15. Wright, G.E. and Brown, N.C. (**1990**). Deoxyribonucleotide analogs as inhibitors and substrates of DNA polymerases. *Pharmacol. Ther.*, 47, 447-497.

16. Shields, A.F. (**2006**). Positron emission tomography measurement of tumor metabolism and growth: Its expanding role in oncology. *Mol Imaging Biol*, 8, 141-150.

17. Chaput, J.C. and Szostak, J.W. (**2003**). TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.*, 125, 9274-9275.

18. Steitz, T.A. (**1999**). DNA polymerases: Structural diversity and common mechanisms. *J. Biol. Chem.*, 274, 17395-17398.

19. Note: DNA Polymerases abbreviations: Deep Vent (3'-5' exo-), DV; $9^{o}N_{m}$, 9N; Therminator, Th; *Bst* large fragment, *Bst*; *Taq*, *Taq*; Phusion High-Fidelity, Ph; Klenow fragment (3'-5' exo-), KF; Moloney Murine leukemia virus reverse transcriptase, MMLV.

20. Gardner, A.F. and Jack, W.E. (**2002**). Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and Taq DNA polymerases. *Nucleic Acids Res.*, 30, 605-613.

21. Ichida, J.K., Zou, K., Horhota, A., Yu, B., McLaughlin, L.W. and Szostak, J.W. (**2005**). An *in vitro* selection system for TNA. *J. Am. Chem. Soc.*, 127, 2802-2803.

22. Kunkel, T.A. (**2004**). DNA replication fidelity. *J. Biol. Chem.*, 279, 16895-16898.

23. Kool, E.T. (**2002**). Active site tightness and substrate fit in DNA replication. *Annu. Rev. Biochem.*, 71, 191-219.

24. Marquez, V.E., Ezzitouni, A., Russ, P., Siddiqui, M.A., Ford Jr., H., Feldman, R.J., Mitsuya, H., George, C. and Barchi Jr., J.J. (**1998**). HIV-1 reverse transcriptase can discriminate between two conformationally locked carbocyclic AZT triphosphate analogs. *J. Am. Chem. Soc.*, 120, 2780-2789.

25. NEB. *manufacturer's instruction.*

26. Huang, H.X. and Keohavong, P. (**1996**). Fidelity and predominant mutations produced by deep vent wild-type and exonuclease-deficient DNA polymerases during in vitro DNA amplification. *DNA Cell Biol.*, 15, 589-594.

27. Kunkel, T.A. and Bebenek, K. (**2000**). DNA replication fidelity. *Annu. Rev. Biochem.*, 69, 497-529.

28. Veedu, R.N., Vester, B. and Wengel, J. (**2006**). Synthesis of an LNA nucleotide triphosphate and its enzymatic incorporation into DNA strands. In Haenni, D. L. (ed.), *Conference Book.* International Roundtable on Nucleosides, Nucleotides and Nucleic Acids, Bern, Switzerland, pp. 292.

Chapter V. Polymerase-Directed Synthesis of 2'-Deoxy-2'-fluoro-β-D-arabinonucleic Acids (2'F-ANA)

Supplementary Data
(pp194-203)

Contents

Material and Method

Nucleotides

Nucleotides. 2'-Deoxyribonucleoside 5'-triphosphates (dNTPs) were purchased from Fermentas. 2'F-araN were synthesized by published procedures (Elzagheid, M. I., et al., Synthesis of protected 2'-Deoxy-2'-fluoro-β-D-arabinonucleosides. In *Current Protocols in Nucleic Acid Chemistry,* John Wiley & Sons: **2002**, pp1.7.1-19). The introduction of the 5'-triphosphate moiety was conducted by Rasayan, Inc. (Encinitas, California, USA) and 2'F-rUTP and 2'F-rCTP were also purchased from Rasayan, Inc. The structures of the fluorinated nucleoside 5'-triphosphates are shown in Figure S5.1.

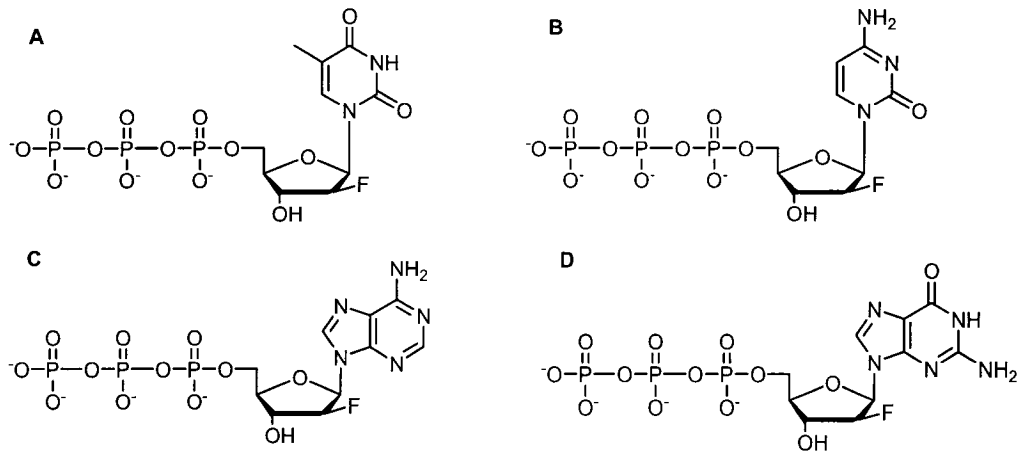


Figure S5.1: Structure of 2'F-araNTPs. **A**: 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)thymine 5'-triphosphate (2'F-araTTP); **B**: 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)cytosine 5'-triphosphate (2'F-araCTP); **C**: 9-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)adenine 5'-triphosphate (2'F-araATP); **D**: 9-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)guanine 5'-triphosphate (2'F-araGTP).

2. Template and Primer Design. DNA and FANA-DNA templates and primers used in this study are shown in Table S5.1. The primer (PF20) is a 17nt DNA sequence 5'-*TAATACGACTCACTATA*-3'. Each template strand comprises three sequence segments: a 17nt long primer binding sequence *5'-TATAGTGAGTCGTATTA-3'*, a 10nt long running start sequence 5'-CTCTTCTCCC-3', and a variable "test sequences" to test 2'F-araNTP incorporation

Table S5.1. Base Sequences of Oligonucleotide Templates and Primer

Code	Type	Sequence*	N-mer
PF20	Primer	5'-*TAATACGACTCACTATA*-3'	17
PF21	DNA temp.	5'-TTTGCC-A-CTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	34
PF23	DNA temp.	5'-TTTACC-G-CTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	34
PF31	FANA-DNA temp.	5'-**TTACCTTT**CTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	35
PF32	DNA temp.	5'-CTCTATGTGCACGCACTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	42
PF33	DNA temp.	5'-TCGGTGGATCATAGACAGTACTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	47
PF34	FANA-DNA	5'-CTCT**ATGTGCACGCA**CTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	42
PF39	DNA control	5'-TAATACGACTCACTATAGGGAGAAGAG-3'	27
PF41	DNA temp.	5'-AGAGCC-T-GAGAAGAGAG*TATAGTGAGTCGTATTA*-3'	34
PF43	DNA temp.	5'-AAATGG-C-GAGAAGAGAG*TATAGTGAGTCGTATTA*-3'	34

*Note: **Bold and capital letter: FANA units;** Capital letters: DNA; *Italized sequence: primer binding sequence and primer (17nt);* Underlined sequence: a running start (10nt).

Table S5.2. Polymerase Information and Reaction Buffer Conditions

Polymerase*	Short name	Activity (units/μL)	Reaction buffer (x1)
Thermophilic DNA polymerase (reaction temperature at 55°C in this study)			
Deep Vent (3'→5' exo-) DNA polymerase	DV	2	20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1 % Triton X-100; 0.2 mM dNTPs
9° N_m DNA polymerase	9N	2	Same as above
Therminator DNA polymerase	Th	2	Same as above
Bst DNA polymerase large fragment	*Bst*	8	Same as above
Taq DNA polymerase	*Taq*	5	Same as above
Phusion High-Fidelity DNA polymerase	Ph	2	5X Phusion™ HF Buffer; 0.2 mM dNTPs
Mesophilic DNA polymerase (reaction temperature at 37°C in this study)			
Klenow fragment (3'→ 5' exo-) DNA polymerase	KF	5	10 mM Tris-HCl (pH 7.5 at 25°C), 5 mM MgCl2, 7.5 mM DTT; 0.033 mM dNTPs
Reverse transcriptase (reaction temperature at 37°C in this study)			
HIV-1 RT (recombinant)	HIV	27.3	50 mM Tris-HCl (pH 7.8 at 25°C), 60 mM KCl, 2.5 mM $MgCl_2$; 0.2 mM dNTPs
Moloney Murine Leukemia Virus RT	MMLV	10	50 mM Tris-HCl (pH 8.3 at 25°C), 50 mM KCl, 4 mM $MgCl_2$, 10mM DTT; 0.5 mM dNTPs

* Enzyme sources: all from New England Biolabs (NEB) except for Ph from Finnzymes (distributed by NEB); HIV-1 RT from Worthington Biochemical Corp.

(PF33), FANA template mediated DNA synthesis (PF31, PF34) and the fidelity of 2'F-araATP (PF41), 2'F-araGTP (PF43), 2'F-araTTP (PF21), 2'F-araCTP (PF23). Other control sequences (PF39, PF33) are also listed in Table S5.1. DNA sequences were obtained from commercial sources (Integrated DNA Technologies). Oligonucleotide FANA-DNA templates were synthesized according to the published procedures (Viazovkina, E. V., et al., Solid-phase synthesis of 2'-deoxy-2'-fluoro-β-D-oligoarabinonucleotides (2'F-ANA) and their phosphorothioate derivatives. In *Current Protocols in Nucleic Acid Chemistry*, John Wiley & Sons, Inc. **2002**, pp 4.15.1-22).

3. General Procedure of Primer Extension Reaction. Primer extension assays were used to evaluate the polymerase activity for incorporation of 2'F-araNTPs, and for template mediated DNA biosynthesis. Unless otherwise noticed in this study, the conditions used in the primer extension assays, such as dNTP concentrations, reaction temperatures, concentration of Mg^{2+}, and reaction buffers (provided by the manufacturers, except for the HIV-RT buffer which was prepared in-house) are shown in Table S5.2. The DNA primer (PF20) was first labeled by a radioactive phosphorous probe (^{32}P) at the 5'-hydroxyl terminus according to published procedure (Galarneau, A., et al., Assay for evaluating ribonuclease H-mediated degradation of RNA-antisense oligonucleotide duplexes. In *Methods Mol. Biol.*, Totowa, NJ, United States, **2005**, 288, pp 65-80). Unlabeled DNA primer was also used to adjust the primer concentration. Then primer to template was mixed together with the final concentration at 100 nM for the primer and the template (1:1 molar ratio). The primer and template were heated at 95°C for 5min and annealed at 4°C at least two hours before use. In an microtube, the mixture of 5x or 10x buffer, triphosphates (dNTPs or 2'F-araNTPs), water, the primer and template mixture was prepared according to pre-calculated reaction volume and concentrations (see buffer conditions in Table S5.2). Mineral oil (20 μL) was used to prevent evaporation. The reaction mixture was incubated at either 37°C or 55°C (see temperature conditions in Table S5.2) and one of the various enzymes tested was added to initiate the primer extension reactions. Time points were obtained by taking 4 μL or 8 μL aliquot from the reaction mixture and quenching the aliquot by the same volume of a stopping dye solution (98% deionized formamide, 10 mM EDTA, 1 mg/ml bromophenol blue and 1 mg/ml xylene cyanol). The product pattern from each time point was analyzed by 12% denaturing polyacryamide gel electrophoresis (PAGE) and subsequent autoradiography analysis.

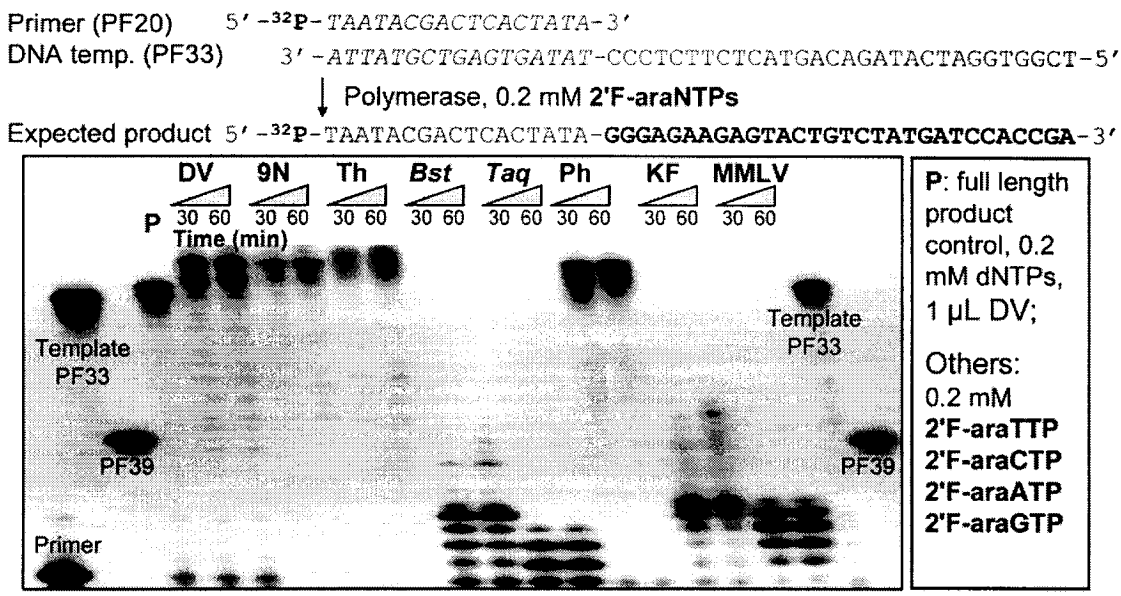


Figure S5.2. Incorporation of 2'F-araN units (7x 2'F-araG, 11x 2'F-araA, 6x 2'F-araT, 6x 2'F-araC) catalyzed by DNA polymerases on a DNA template PF33. Primer extension reaction conditions were performed with 100 nM primer/template (1:1 molar ratio), 0.2 mM 2'F-araNTPs (shown above) and 1 μL DV, 9N, Th, *Bst* and *Taq*, KF, 0.5 μL Ph and MMLV, respectively in 20 μL reaction volume, at 55°C (DV, 9N, Th, *Bst*, *Taq*, Ph) or 37°C (KF and MMLV). Reaction progress over time was analyzed by 12% denaturing PAGE. Lane **P** is the product control obtained in a similar primer extension assay with 1 μL DV, 0.2 mM dNTPs at 55°C for 30min, on the same template PF33; PF39 is a 27-nt DNA control (see Table S5.1).

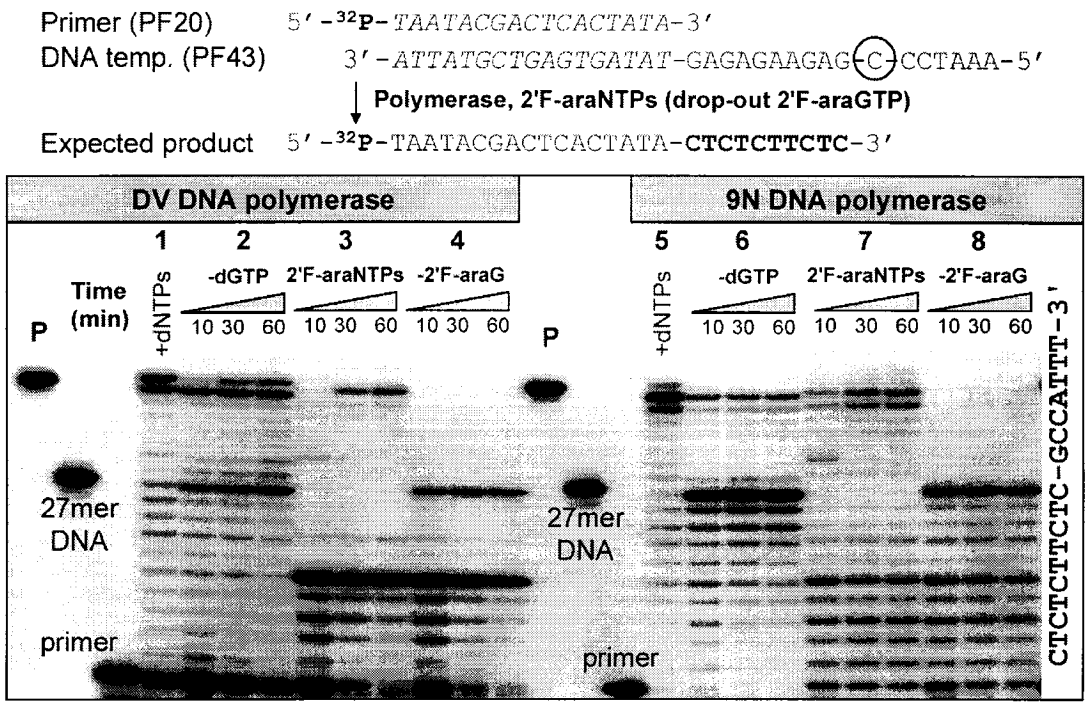


Figure S5.3. Fidelity study on 2'F-araGTP incorporation by DV and 9N DNA polymerases on DNA template PF43. Reaction conditions: primer/template (1:1 molar ratio): 100 nM; 0.4 μL (0.8 unit) DV, 9N, 25 μM triphosphate concentration at 55°C in 30 μL reaction volume. Reaction time for dNTPs (group 1, 5) was 30 min and others were shown above. Group 1 & 5: dNTPs; 2 & 6: drop out dGTP (i.e. dTTP, dCTP and dATP), 3 & 7: 2'F-araNTPs, 4 & 8: drop out 2'F-araGTP; lane P is the template PF41 as the full-length product control; PF39 is a 27-nt DNA control (see Table S5.1). All samples were analyzed by 12% denaturing PAGE.

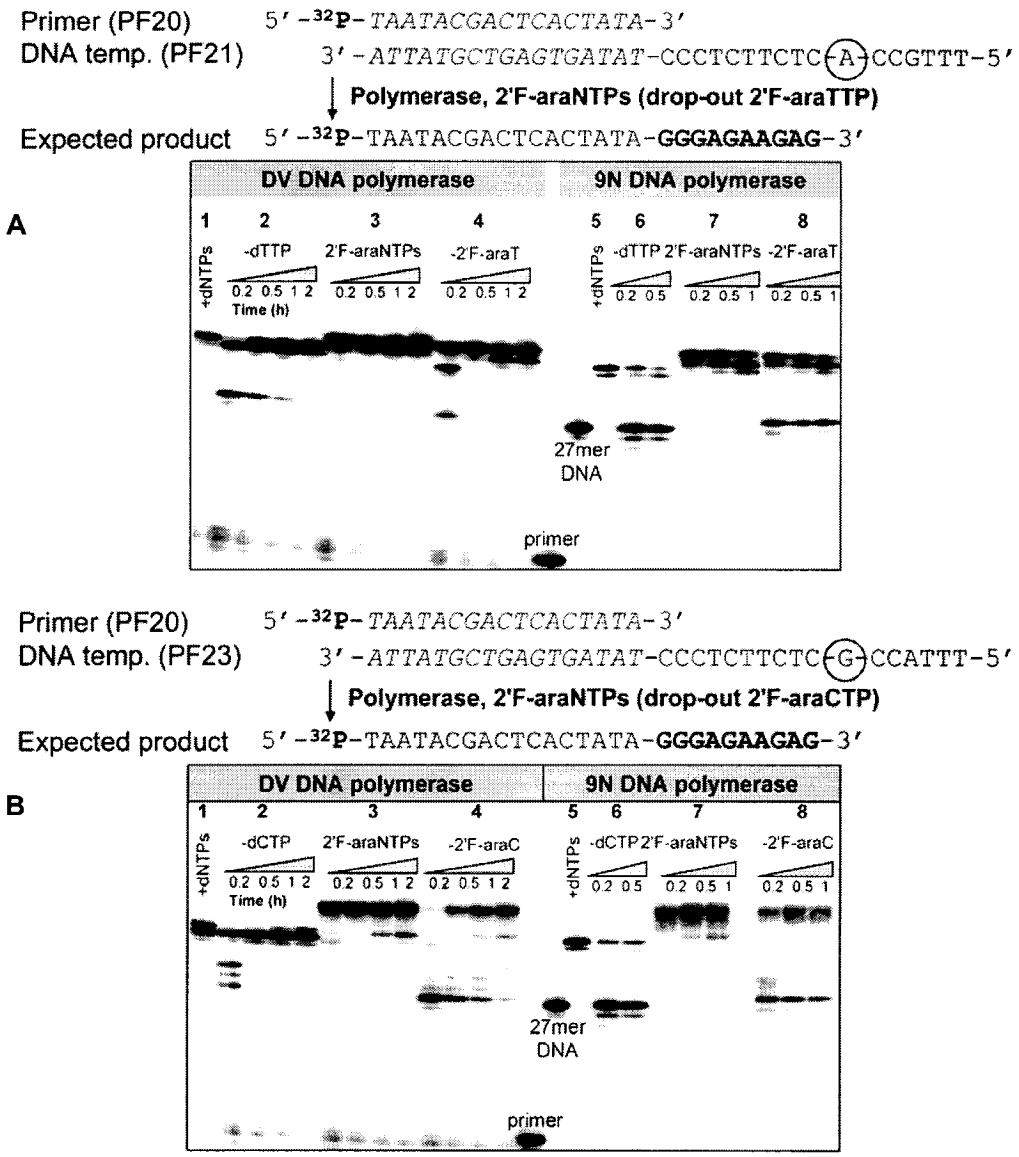


Figure S5.4 A & B. Fidelity study on (**A**) 2'F-araTTP and (**B**) 2'F-araCTP incorporation by DV and 9N DNA polymerases on a DNA template (**A**) PF21 and (**B**) PF23, respectively. Reaction conditions: 100 nM primer/template (1:1 molar ratio); 0.4 μL (0.8 unit) DV or 9N, 25 μM triphosphate concentration at 55 °C in 30 μL reaction volume. Reaction time for dNTPs (group 1, 5 in both **A** and **B**) was 30 min and others were shown above. In **A**: group 1 & 5: dNTPs; 2 & 6: drop out dTTP; 3 & 7: 2'F-araNTPs; 4 & 8: drop out 2'F-araTTPs; in **B**: group 1 & 5: dNTPs; 2 & 6: drop out dCTP; 3 & 7: 2'F-araNTPs; 4 & 8: drop out 2'F-araCTPs; PF39 is a 27-nt DNA control (see Table S1). All samples were analyzed by 12% denaturing PAGE.

Primer (PF20) 5'-^{32}P-*TAATACGACTCACTATA*-3'
DNA template (PF41) 5'-*ATTATGCTGAGTGATAT*-GAGAGAAGAGTCCGAGA-5'

Expected product: ↓ **Phusion, triphosphates condition 1-5**

1. 2'F-araNTPs (**All**) 5'-^{32}P-TAATACGACTCACTATA-**CTCTCTTCTCAGGCTCT**-3'
2. Drop out **2'F-araCTP (AGT)** 5'-^{32}P-TAATACGACTCACTATA-3'
3. Drop out **2'F-araTTP (AGC)** 5'-^{32}P-TAATACGACTCACTATA-**C**-3'
4. Drop out **2'F-araATP (CTG)** 5'-^{32}P-TAATACGACTCACTATA-**CTCTCTTCTC**-3'
5. Drop out **2'F-araGTP (CTA)** 5'-^{32}P-TAATACGACTCACTATA-**CTCTCTTCTCA**-3'

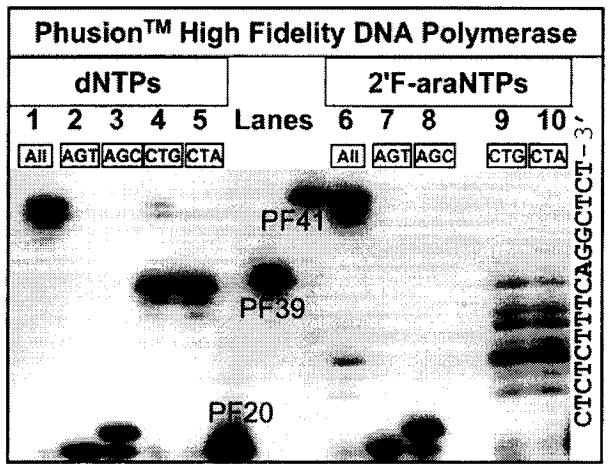


Figure S5.5. Fidelity study on 2'F-araNTP incorporation by Phusion High Fidelity polymerases on purine rich DNA template PF41. The base sequence of the primer, DNA template and resulting terminated products in different dropout conditions (condition 1-5) is shown. Primer extension conditions: 100 nM primer/template (1:1 molar ratio); for dNTPs (lanes 1-5): 0.4 μL (0.8 unit) Ph, 25 μM dNTP concentration; for 2'F-araNTPs (lanes 6-10): 0.8 μL (1.6 unit) Ph, 0.2 mM 2'F-araNTP at 55 °C in 30 μL reaction volume for 30 min. Lane 1: dNTPs; 2: drop out dCTP (i.e. dATP, dGTP and dTTP, shorten as AGT, same coding for lanes 3-5), lane 6: 2'F-araNTPs; 7: drop out 2'F-araCTP (i.e. 2'F-araATP, 2'F-araGTP, 2'F-araTTP, shorten as AGT; same coding for lanes 8-10). The template PF41 and the primer PF20 are shown; PF39 is a 27mer DNA control (see Table S1). All samples were analyzed by 12% denaturing PAGE.

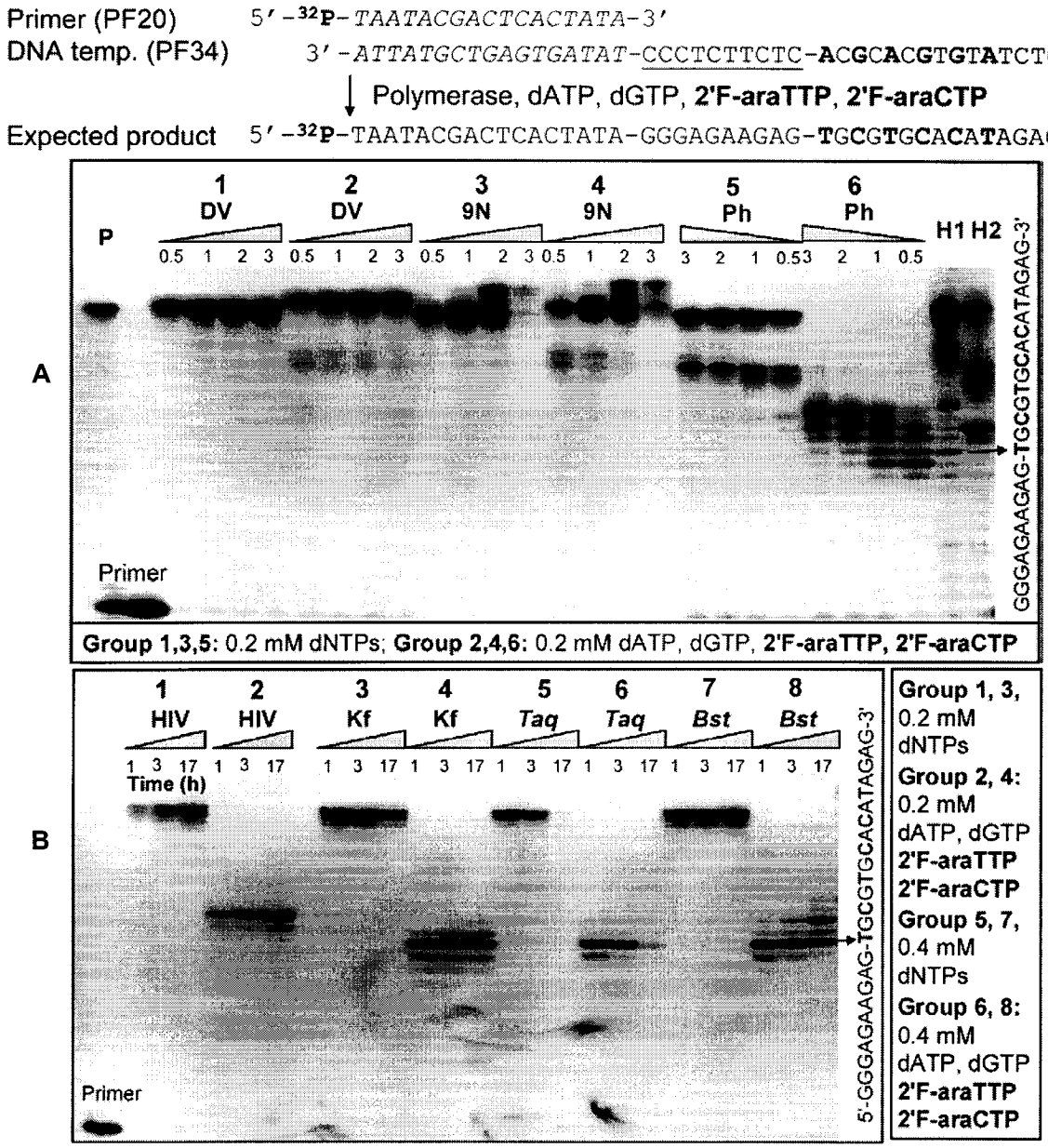


Figure S5.6. Incorporation of 2'F-araNTPs (3x 2'F-araT and 3x 2'F-araC) units catalyzed by DNA polymerases on a chimeric DNA-FANA template PF34. The identity of DNA polymerases, and the base sequence of the primer, chimeric DNA-FANA template (PF34), and resulting chimeric DNA-FANA product are shown. **A:** 20 μL reaction volume for all, 2 μL enzyme used; **B:** 20 μL reaction volume for all, 1.5 μL HIV-RT, 2 μL KF, *Taq*, *Bst*, respectively; H1 were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP + ddTTP (1:1 molar ratio), 0.4 μL HIV-RT in 20 μL reaction volume on a DNA template PF32 with the same sequence as PF34 (See Table S5.1); products shown on lane H2 were formed through a chain termination assay with 0.1 mM dATP, dGTP, dTTP, dCTP + ddCTP (1:1 molar ratio), 0.4 μL HIV-RT in 20 μL reaction volume on a DNA template PF32 with the same sequence as PF34 (See Table S5.1).

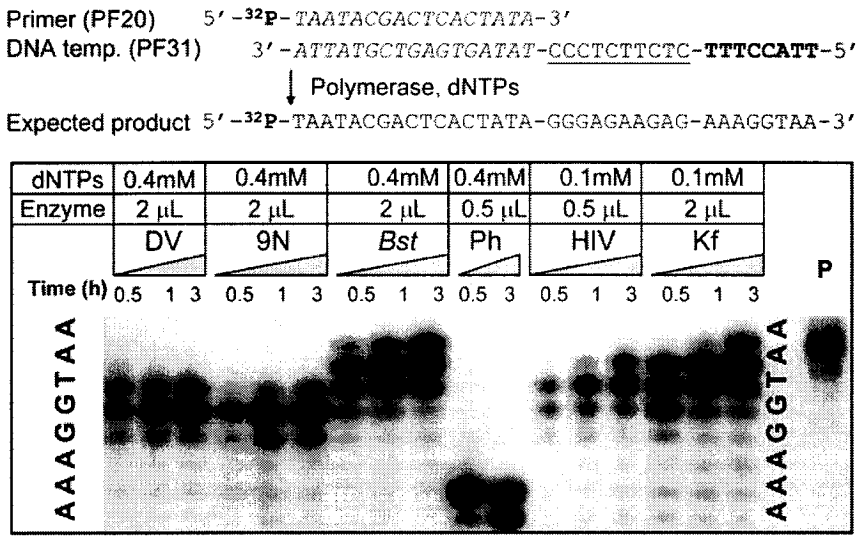


Figure S5.7. Assay to assess the activity of FANA as a template (PF31) for polymerase-directed DNA synthesis. The identity of the DNA polymerases, the base sequence of the primer, chimeric DNA-FANA template (PF31), and resulting DNA oligonucleotide product are shown. Reaction volume was 30 μL, and dNTP and enzyme amounts/concentrations are also shown.

Chapter VI

2'-Deoxy-2'-fluoro-β-D-arabinonucleic acids (2'F-ANA) and 2'-deoxy-2'-fluoro-β-D-ribonucleic acids (2'F-RNA) are diastereomers differing in their configuration at the 2'-carbon of the furanose sugar ring and as such, possess distinct properties that are being exploited in gene silencing (siRNA, antisense aptamers). While 2'-deoxy-2'-fluoro-β-D-ribonucleoside 5'-triphosphates (2'F-rNTPs) are among the most commonly used sugar-modified nucleoside 5'-triphosphates for *in vitro* selection experiments, 2'-deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs) have just been applied to the biosynthesis of 2'F-ANA oligonucleotides (Chapter V). As a follow-up study, here we compare the synthesis efficiency of the two isomeric nucleoside 5'-triphosphates (NTPs), namely, 2'F-araNTPs and 2'-deoxy-2'-fluoro-β-D-ribonucleoside 5'-triphosphate (2'F-rNTPs) by Deep Vent (exo-), $9^{o}N_{m}^{TM}$, HIV-1 reverse transcriptase (RT) and Moloney Murine Leukemia Virus (MMLV) RT DNA polymerases. Incorporation of 2'F-araTTP proceeded more efficiently relative to 2'F-rUTP, while the incorporation of 2'F-araCTP was slightly less efficient than that observed with 2'F-rCTP. Considering the potential advantages of both 2'F-ANA and 2'F-RNA modifications in oligonucleotide-based gene silencing approaches, this study suggests that it should be possible to generate 2'F-ANA/2'F-RNA-based aptamers through an *in vitro* selection process.

Chapter VI. Assessing Polymerase Activity with 2'-Deoxy-2'-fluoro-β-D-ribonucleoside 5'-Triphosphates (2'F-rNTPs) and 2'-Deoxy-2'-fluoro-β-D-arabinonucleoside 5'-Triphosphates (2'F-araNTPs)

Abstract

2'-Deoxy-2'-fluoro-β-D-ribonucleosides (2'F-rN) and 2'-deoxy-2'-fluoro-β-D-arabinonucleosides (2'F-araN) differ solely in the stereochemistry at the 2'-carbon of the furanose sugar ring. While 2'F-rN 5'-triphosphates (2'F-rNTPs) are among the most commonly used sugar-modified nucleoside 5'-triphosphates (NTPs) for *in vitro* selection, the epimeric 2'F-araN 5'-triphosphates (2'F-araNTPs) have only recently been applied to polymerase-directed biosynthesis (Chapter V). In an effort to continue characterizing the polymerase specificity of 2'F-araNTPs, this study assessed the ability of eight different DNA polymerases to incorporate three contiguous pyrimidine 2'F-araNTPs into a growing DNA chain. Furthermore, primer extension assays were conducted to compare, for the first time, the incorporation efficiency of the two isomeric NTPs, i.e., 2'F-araNTPs or 2'F-rNTPs, by four DNA polymerases [Deep Vent (exo-), $9^{o}N_{m}^{TM}$, HIV-1 RT and MMLV-RT]. Under the conditions used, incorporation of 2'F-araTTP proceeded more efficiently relative to 2'F-rUTP, while the incorporation of 2'F-araCTP is slightly less efficient than that observed with 2'F-rCTP. Interestingly, these preferences were observed for all four of the DNA polymerases tested. Considering the potential advantages of each 2'-deoxy-2'-fluoro-D-arabinonucleic acid (2'F-ANA) and 2'-deoxy-2'-fluoro-D-ribonucleic acid (2'F-RNA) modifications in oligonucleotide-based gene silencing approaches, these studies suggest that it should be possible to generate 2'F-ANA/2'F-RNA-based aptamers through an *in vitro* selection (SELEX) process. By examining the structure-activity relationship (SAR) of the ribose and arabinose sugar, this study also contributes to the understanding of how C2' stereochemistry and sugar pucker affect NTP selection (specificity) by DNA polymerases.

Introduction

NTPs are extensively used as probes to elucidate mechanistic details of DNA biosynthesis and they represent the active form of several approved and pre-clinical antiviral and anticancer agents.[1-3] Chemically-modified nucleoside 5'-triphosphates (NTPs) are also of great importance in the development of aptamers as research, therapeutic and diagnostic agents.[4,5] They are usually generated from the process of *Systematic Evolution of Ligands by EXponential enrichment* (SELEX)[6-8] and gained recognition with the recent FDA approval of Macugen®, a 2'F/2'-OMe ribose-modified oligonucleotide indicated for the treatment of neovascular age-related macular degeneration.[9]

Modified NTPs can serve either as substrates or inhibitors of polymerases depending on the structural changes introduced through modifications to the base, phosphate and/or sugar moieties relative to the natural NTPs (i.e. deoxyribonucleoside 5'-triphosphates (dNTPs) and ribonucleoside 5'-triphosphates (rNTPs)).[2] Introducing a fluorine (F) atom into the sugar moiety of nucleosides, particularly at C2', has a strong impact on the sugar ring conformational equilibrium.[10] When the 2'-fluorine atom is "up" (or β) in 2'-deoxy-2'-fluoro-β-D-arabinonucleic acid (2'F-ANA), the arabinofuranose sugar ring predominantly puckers in the *South/East* (C2'/O4'-*endo*) conformation,[11,12] whereas a "down" (or α) fluorine in 2'-deoxy-2'-fluoro-β-D-ribonucleic acid (2'F-RNA) steers the ribofuranose ring into the *North* (C3'-*endo*).[13-15] These C2'-modified sugars in turn dictate the conformation of the corresponding oligonucleotides.

A fluorine atom at C2' significantly impacts the thermodynamic stability and biological activity of duplexes. For example, oligonucleotides modified with 2'F-RNA units have enhanced binding affinity with mRNA (~ +2 °C/nt) relative to the unmodified counterparts.[16-19] The incorporation of 2'F-ANA units also raises the T_m of different systems, *i.e.*, duplexes (~ +1 °C/nt),[20] triplexes (~ +0.8 °C/nt),[21] C-rich tetrads (~ +1 °C/nt, pH<4.0)[22] and G-rich tetrads (~ +2 °C/nt) (Peng and Damha, manuscript in preparation; Chapter IV). Pyrimidine-based 2'F-RNAs were introduced into hammerhead ribozymes[23] or siRNAs[24] and exhibited increased or maintained activity. Antisense DNA or siRNA strands containing a significant 2'F-ANA content displayed gene silencing efficacy in the low nanomolar range.[25,26]

Given that 2'-deoxy-2'-fluoro-β-D-ribonucleoside 5'-triphosphates (2'F-rNTPs) are good substrates for T7 RNA polymerase,[27] they are among the most commonly used in RNA-based SELEX.[28-30] A few studies have shown that 2'F-rNTPs are also substrates for DNA polymerases.[31,32] Richardson *et al.* reported that all four 2'F-rNTPs are substrates of human DNA pol γ, and DNA poly α can accepts 2'F-rUTP, 2'F-rCTP and 2'F-rGTP, but not 2'F-rATP.[33] Furthermore, the incorporation efficiency of 2'F-rCTP by pol α is 17 fold higher compared to that of 2'F-rUTP.[33] Thermostable DNA polymerases [Pfu (exo-), Vent (exo-), Deep Vent (exo-) and UlTma] are also able to incorporate 2'F-rUTP and 2'F-rCTP with reasonable efficiency.[34]

2'-Deoxy-2'-fluoro-β-D-arabinonucleoside analogues have been studied for antiviral and anticancer applications. The recently approved anticancer drug clofarabine[35] is 2-chloro-2'F-araA. 2'-Deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs) and derivatives have long been investigated to understand their virotoxicity or cytotoxicity since they are involved in DNA biosynthesis by viral and eukaryotic DNA polymerases.[2] FMAU (2'F-araT) and other 2'F-arabinonucleosides can be incorporated into eukaryotic as well as viral DNA,[36,37] but other studies show that they are preferred by viral polymerases but not the mammalian polymerases.[37,38] One of clofarabine's mechanisms of action, in fact, is the induction of strand breaks after incorporation into cellular DNA.[39] FMAU[40,41] and other 2'F-araNTPs incorporating ^{18}F radionuclide ($t_{½}$ = 110 min) allows for the non-invasive imaging of tumors through positron emission tomography (PET). The successful application of these NTPs in PET relies on the ability of DNA polymerase to largely (and more selectively) incorporate these NTP in tumor cells.[42] 2'F-araNTPs have not been extensively investigated for *in vitro* selection/amplification processes.[43]

Our group has recently shown that thermostable DNA polymerases such as Deep Vent (exo-) (DV), $9^oN_m^{TM}$ (9N), Therminator™, and Phusion™, accept both purine and pyrimidine 2'F-araNTPs as substrates and are able to synthesize long (30 nt) stretches of 2'F-ANAs with high efficiency and fidelity (Peng and Damha, manuscript submitted, Chapter V). In addition, these enzymes are capable of synthesizing DNA on a 2'F-ANA template, and can drive the formation of 2'F-ANA/2'F-ANA base pairs on chimeric 2'F-

ANA-DNA templates (Peng and Damha, manuscript submitted, Chapter V). This suggests that it should be possible to generate 2'F-ANA-based aptamers through SELEX.

In this follow-up study, we designed a less challenging system to first assess the incorporation of three contiguous pyrimidine 2'F-araNTPs (i.e. 2'F-araTTP or 2'F-araCTP). Then we compared for the first time the incorporation efficiency of pyrimidine 2'F-araNTPs to that of the corresponding 2'F-rNTP stereoisomers by two retroviral DNA polymerases, namely, HIV-1 reverse transcriptase (RT), Moloney Murine Leukemia (MMLV)-RT, and two thermostable DNA polymerases, namely, DV and 9N. Finally, we also compared the incorporation efficiency of various combinations of pyrimidine 2'F-araNTP, 2'F-rNTP, and native rNTP along with dATP and dGTP.

Materials and Methods

Nucleotides.

dNTPs were purchased from Fermentas. 2'-deoxy-2'-fluoro-β-D-arabinonucleosides were obtained from Topigen Pharmaceuticals, Inc. (Montreal, Canada). The introduction of the 5'-triphosphate moiety was conducted by Rasayan, Inc. (Encinitas, California, USA). 2'F-rUTP and 2'F-rCTP were purchased from Rasayan, Inc. The structures of the fluorinated nucleoside 5'-triphosphates are shown in Figure 6.1.

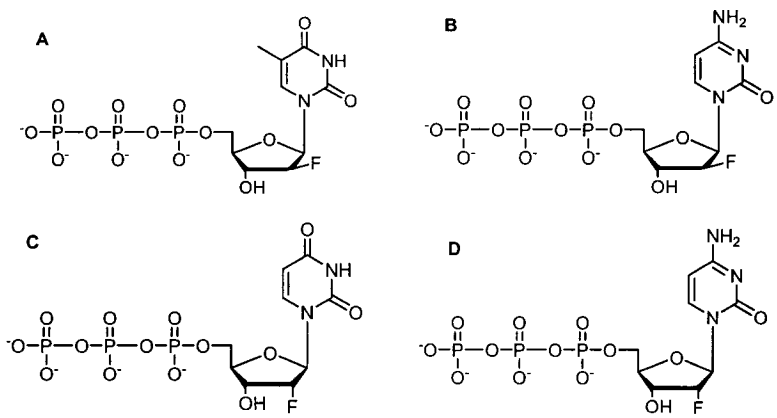


Figure 6.1: A. 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)thymine 5'-triphosphate (2'F-araTTP); B. 1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)cytosine 5'-triphosphate (2'F-araCTP); C. 1-(2-deoxy-2-fluoro-β-D-furanosyl)uracil 5'-triphosphate (2'F-rUTP); D. 1-(2-deoxy-2-fluoro-β-D-ribofuranosyl)cytosine 5'-triphosphate (2'F-rCTP)

Template and Primer Design

DNA and 2'F-ANA-DNA templates and primers used in this study are shown in Table 6.1. The primer (PF20) is a 17nt DNA sequence 5'-*TAATACGACTCACTATA*-3'. Each template strand comprises three sequence segments: a 17nt long primer binding sequence 5'-*TATAGTGAGTCGTATTA*-3', a 10nt running start sequence 5'-CTCTTCTCCC-3',[44,45] and a variable DNA "test sequence". The oligomers were designed to test the incorporation of three 2'F-araT units on the DNA template PF22, three 2'F-araC units on the DNA template PF24, and six mixed 2'F-araT and 2'F-araC units on the DNA template PF32. DNA sequences were obtained from Integrated DNA Technologies (IDT). Oligonucleotide 2'F-ANA-DNA templates were synthesized according to a published procedure.[46]

Table 6.1: Base sequence of oligonucleotide templates and primer

Code	Type	Sequence*	N-mer
PF20	Primer	5'-*TAATACGACTCACTATA*-3'	17
PF22	DNA temp.	5'-TTTGCCAAACTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	36
PF24	DNA temp.	5'-TTTACCGGGCTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	36
PF32	DNA temp.	5'-CTCTATGTGCACGCACTCTTCTCCC*TATAGTGAGTCGTATTA*-3'	42

*Note: *Italized sequence: primer binding sequence and primer (17nt);* Underlined sequence: a running start (10nt)

Primer Extension Reactions

We adapted primer extension assays[44,45] to first screen eight DNA polymerases (Table 6.2) for their ability to incorporate three consecutive 2'F-araT and 2'F-araC units. Enzyme activity and default reaction buffer conditions suggested by the manufacturers are shown in Table S6.1 (Supplementary Data). All reaction buffers were provided by the manufacturers except for the HIV-1 RT buffer which was prepared in-house. The DNA primer (PF20) was first labeled by a radioactive phosphorus probe (^{32}P) at the 5'-hydroxyl terminus according to a published procedure.[47] Unlabeled DNA primer was also used to adjust the primer concentration. The final concentrations were 85 nM for the primer and 255 nM for the template in the screening experiments, and 100 nM for the primer and the

template in the comparison studies. The primer and template were heated at 95°C for 5 min and annealed at 4°C at least two hours before use. In a microtube, the mixture of a buffer, triphosphates (different combinations of four dNTPs, 2'F-araNTPs, 2'F-dNTPs and/or rNTPs), water, primer and template was prepared according to pre- calculated reaction volume and concentrations. Mineral oil (20 μL) was used to prevent evaporation for assays running at 55°C. A polymerase was added to initiate the primer extension reaction. Time points were obtained by removing 4 μL or 8 μL aliquots from the reaction

Table 6.2: Polymerase used and incorporation efficiency of three 2'F-araT(C)TPs

Polymerases *	**Short name**	**Incorporation of three 2'F-araTTP**	**Incorporation of three 2'F-araCTP**
Thermophilic DNA polymerases (assays conducted at 55°C except otherwise noticed)			
Deep Vent (3'→5' exo-) DNA polymerase	**DV**	++++ (37°C) ++++	+++ (37°C) ++++
9° N_m™ DNA polymerase	**9N**	++++ (37°C) ++++	None (37°C) ++++
Bst DNA polymerase large fragment	***Bst***	++ (37°C) +++	None (37°C) ++++
Taq DNA polymerase	***Taq***	++ (37°C) ++++	++ (37°C) ++++
Phusion™ High-Fidelity DNA polymerase	**Ph**	++++	++++
Mesophilic DNA polymerase (assays conducted at 37°C)			
Klenow fragment DNA polymerase (3'→ 5' exo-)	**Kf**	++	++
Reverse transcriptase DNA polymerases (assays conducted at 37°C)			
HIV-1 RT (recombinant)	**HIV-1 RT**	++++	++++
Moloney Murine Leukemia Virus RT	**MMLV-RT**	+++	++++

* Enzyme sources: all from New England Biolabs (NEB) except for Ph from Finnzymes (distributed by NEB); HIV-1 RT from Worthington Biochemical Corp.;

mixture and quenching with an equal volume of a stopping dye solution (98% deionized formamide, 10 mM EDTA, 1 mg/ml bromophenol blue and 1 mg/ml xylene cyanol). The product pattern was analyzed by 12% denaturing polyacryamide gel electrophoresis (PAGE) and subsequent autoradiography analysis.

Results

Incorporation of Three Contiguous 2'F-araTTPs and 2'F-araCTPs

Primer extension assays were used to assess the incorporation of three contiguous 2'F-araTTPs on the DNA template PF22. Under the conditions used, all of the polymerases studied generated full-length products, although to different extents (Figure 6.2). Pausing was observed for *Taq*, MMLV-RT, DV, 9N and *Bst*. Remarkably, the thermophilic enzymes DV, 9N and *Taq* were able to incorporate three 2'F-araTTPs even at 37°C (Section A in Figure 6.2). Ph could afford full-length products without any pausing at 55°C.

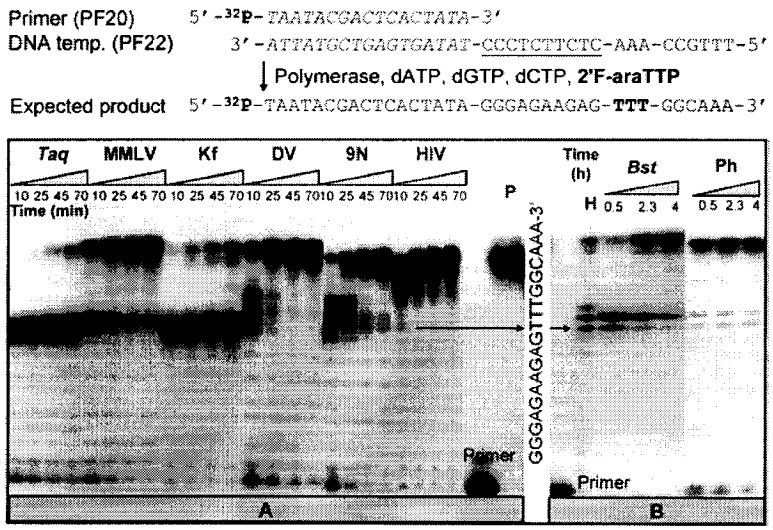


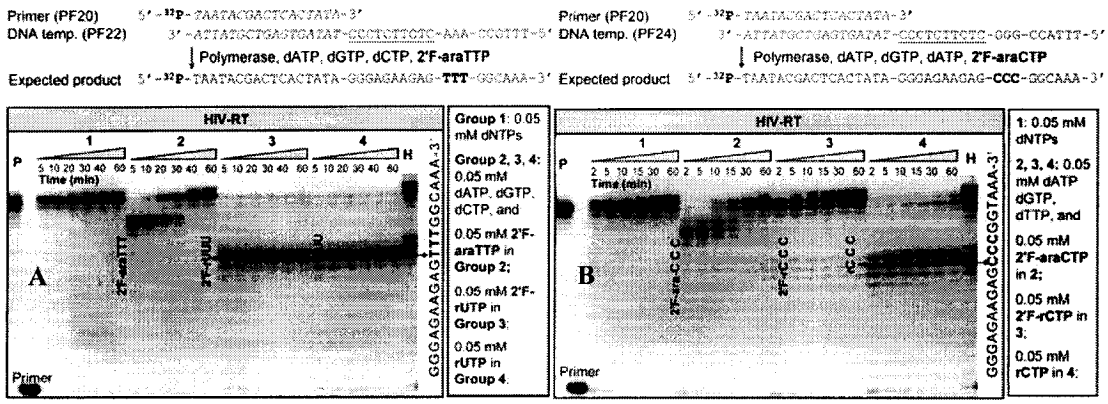
Figure 6.2: DNA-template directed polymerization of three contiguous 2'F-araTs by DNA polymerases. **A**: all reactions at 37°C, 20 μL reaction volume, 1 μL polymerase, dNTPs and 2'F-araTTP concentrations: 0.2 mM for *Taq*, DV, 9N, HIV-1 RT; 0.5 mM for MMLV, 0.033 mM for Kf; **B**: reactions at 55°C, 20 μL reaction volume, 2 μL *Bst* and 0.4 μL Ph, 0.2 mM triphosphates; **P**: full-length product control; **H**: products formed through a chain termination assay with 0.1 mM dATP, dGTP, dTTP, dTTP+ddTTP (1:1 molar ratio), 0.4 μL HIV-1 RT in 20 μL reaction volume, on the same template PF22

The same primer extension assays were used to assess the incorporation of three 2'F-araC units on the DNA template PF24 (Figure S6.1, Supplementary Data). MMLV-RT, Kf and HIV-1 RT all gave full-length products (Section A). Thermophilic enzymes DV and *Taq* could incorporate three consecutive 2'F-araCTPs at 37°C, however, the rate of polymerization was significantly slower and some premature products (pausing) were observed. 9N and *Bst* failed to generate full-length products at 37°C (section A) but were able to incorporate three 2'F-araCTPs and gave the expected full-length products at higher temperature (55°C) (Section B & C). The high-fidelity enzyme Ph also easily provided a full-length product at 55°C. Overall, all the DNA polymerases tested were able to incorporate three 2'F-araTTPs and 2'F-araCTPs into a DNA strand and afforded full-length products. 9N and DV could incorporate 2'F-araTTP better than 2'F-araCTP (Table 6.2).

Incorporation Efficiency of 2'F-araTTP *vs* 2'F-rUTP

Once the conditions of the primer extension assays were established, the incorporation efficiencies of 2'F-araTTP and 2'F-rUTP by HIV-1 RT and MMLV-RT on the DNA template PF22 were compared. As shown in Figure 6.3A, 2'F-araTTP was an excellent substrate of HIV-1 RT and its incorporation proceeded more efficiently compared to 2'F-rUTP and rUTP. Interestingly, different pausing patterns were observed. In the absence of 2'F-modifed NTPs, the expected full-length DNA product was obtained from DNA-primed (PF20) biosynthesis on the DNA template PF22. Addition of an equivalent amount of the dideoxynucleoside 5'-triphosphate (ddTTP) led to the formation of three new bands resulting from chain termination directly opposite the d(AAA) residues present in the template strand (lane H; Figure 6.3). These bands served as "markers" facilitating the identification of products formed in the 2'F-modifed NTPs-based assays. The replacement of dTTP with either 2'F-rUTP or rUTP strongly inhibited RT-catalyzed synthesis of full-length product, as evidenced by the appearance of significant amounts of products resulting from chain termination after incorporation of one or two rNTPs. Even the incorporation of the first rUTP or 2'F-rUTP unit proceeded with difficulty as evidenced by the band patterns observed after 5 min. These truncated products represented more than 94% of the total polymerization products even after 1h incubation

with the enzyme. In sharp contrast, polymerization in the presence of 2'F-araTTP provided a significant amount of the full-length product after 30 min. However, the appearance of a significant amount of two shorter products, 5'-[DNA primer]-2'F-ara(TpTpT)-dGG-3' and 5'-[DNA primer]-2'F-ara(TpTpT)-dGGC-3', were observed during the first 5-20 min of reaction. This suggests that while the arabinose units do not cause significant chain termination as observed with both 2'F-rUTP and rUTP, the efficiency of subsequent steps, particularly the addition of subsequent dNTPs, is somewhat hindered following incorporation of 2'F-araTTPs (delayed pausing).



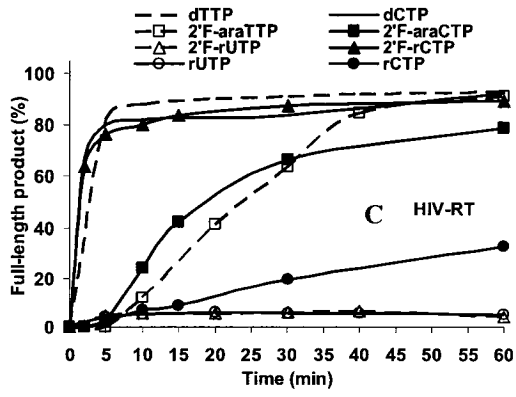


Figure 6.3: Efficiency of incorporation of 2'F-araTTP *vs* 2'F-rUTP through the DNA template PF22 (**A**), and 2'F-araCTP *vs* 2'F-rCTP through the DNA template PF24 (**B**) mediated by HIV-1 RT DNA polymerization. (**C**) is a plot of time-course of full-length product (%). Reactions conducted at 37°C in a 40 μL reaction volume, with 0.66 μL (18 units) enzyme; triphosphate concentrations were shown; products shown on lane **H** were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio) on the template PF22 (**A**), or 0.1 mM dATP, dGTP, dTTP, dCTP+ddCTP (1:1 molar ratio) on the template PF24 (**B**), respectively; 0.4 μL HIV-1 RT in 20 μL reaction volume at 37°C. **P**: full-length product control.

Table 6.3: The order of incorporation efficiency of 2'F-araT(C)TP *vs* 2'F-rU(C)TP

Polymerases	Order of efficiency of incorporation [a]	Figures
HIV-1 RT	dTTP > 2'F-araTTP > 2'F-rUTP > UTP	6.3A
	dCTP ≈ 2'F-rCTP ≥ 2'F-araCTP > CTP	6.3B
MMLV-RT	dTTP > 2'F-araTTP > 2'F-rUTP > UTP	S6.2; S6.3
	dCTP ≈ 2'F-rCTP ≥ 2'F-araCTP > CTP	6.5
DV	dTTP ≈ 2'F-araTTP > 2'F-rUTP	6.4A
	dCTP ≈ 2'F-rCTP ≈ 2'F-araCTP	6.4B
9N	dTTP ≈ 2'F-araTTP > 2'F-rUTP	S6.4A
	dCTP ≈ 2'F-araCTP ≈ 2'F-rCTP	S6.4B

[a] incorporation efficiency is evaluated by the percentage of full-length products from primer extension assays.

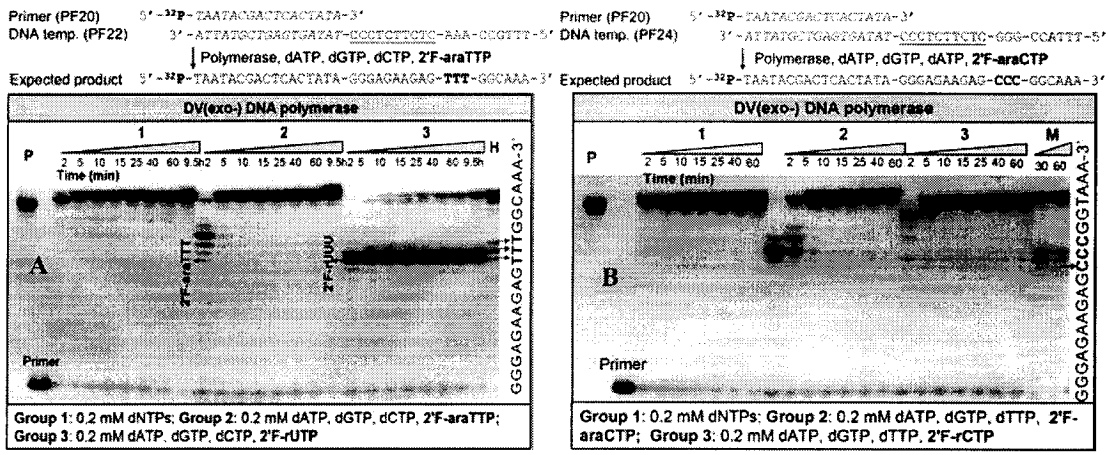
Similar results were obtained with MMLV-RT. This enzyme was able to incorporate 2'F-araTTP more efficiently than 2'F-rUTP and rUTP (Figure S6.2 and S6.3, Supplementary Data). Similarly, pausing was observed *after* the first incorporation of 2'F-araTTP, and *before* the first incorporation of 2'F-rUTP and rUTP. By analyzing the second incorporations of modified triphosphates in Figure S6.3, it is clear that 2'F-araTTP is the best substrate of MMLV-RT.

We also wished to compare the efficiency of incorporation of 2'F-araTTP and 2'F-rUTP by thermostable polymerases DV and 9N, as these enzymes were previously found to be the most effective at polymerizing 2'F-ANA strands (Peng and Damha, manuscript submitted; Chapter V). With DV, incorporation of 2'F-araTTPs proceeded efficiently giving after 5 min nearly quantitative yield of full length product. Accumulation of an intermediate band corresponding to 5'-[DNA primer]-2'F-ara(TpTpT)-dG-3' was observed in the first 2 min (Figure 6.4A). With 2'F-rUTP, however, abrupt pausing was observed after the first and second 2'F-rUTP incorporation, yielding only 20% of the desired full-length product after 9.5h (Figure 6.4A). The same trends were observed with 9N DNA polymerase (Figure S6.4A, Supplementary Data). Incorporation efficiencies of 2'F-araTTP and dTTP were comparable and *ca.* 100% full-length products were generated within 2 min and without pausing, whereas the reaction with 2'F-rUTP paused after one

incorporation producing little of the expected full-length product (*ca.* 20% after one hour).

Incorporation Efficiency of 2'F-araCTP *vs* 2'F-rCTP

The same primer extension assays were used to evaluate the efficiency of DNA-template (PF24) directed incorporation of the cytosine NTPs by HIV-1 RT and MMLV-RT. The incorporation of 2'F modified CTPs seemed to be more complicated than what was previously shown with 2'F modified TTPs (Figure 6.3B and Figure 6.5). Generally, in contrast to what was observed for the thymine NTPs, the incorporation efficiency of 2'F-araCTP by HIV-1 RT and MMLV-RT is comparable to or slightly less efficient than that



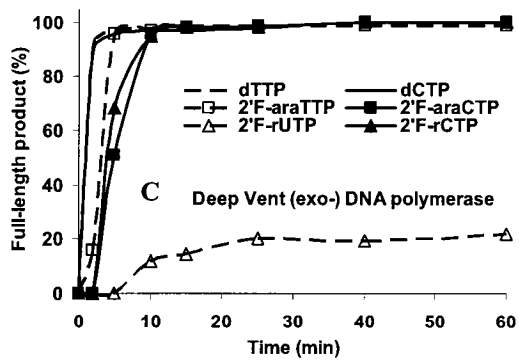


Figure 6.4: Efficiency of incorporation of 2'F-araTTP *vs* 2'F-rUTP through the DNA template PF22 (**A**), and 2'F-araCTP *vs* 2'F-rCTP through the DNA template PF24 (**B**) mediated by DV DNA polymerization. (**C**) is a plot of time-course of full-length product (%). Reactions conducted at 55°C in a 60 μL reaction volume, with 1.5 μL (3 units) enzyme; triphosphate concentrations were shown; products shown on lane **H** were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio), 0.4 μL HIV-1 RT in 20 μL reaction volume on the template PF22 at 37°C; products shown on lane **M** were formed through a chain termination assay with 0.2 mM dATP, dGTP, dTTP, dCTP+ddCTP (1:1 molar ratio), 0.4 μL MMLV-RT in 20 μL reaction volume on the template PF24 at 37°C. **P**: full-length product control.

observed with 2'F-rCTP but better than rCTP. Remarkably, 2'F-rCTP behaved very differently from rCTP, e.g, little or no pausing was observed with these enzymes. On the other hand, assays with rCTP and either HIV-1 RT and MMLV-RT showed significant pausing before the first rCTP incorporation, as observed earlier with rUTP.

Binding of 2'F-araCTP at the active site of MMLV-RT appears to occur more readily over its isomeric counterpart 2'F-rCTP since pausing was observed *after* the first 2'F-araCTP incorporation but *before* the first 2'F-rCTP incorporation (Figure 6.5). Further elongation beyond the 2'F-araC unit, however, seemed very difficult and mainly premature products (*ca.* 90%) were observed after 60 minutes; in contrast, once 2'F-rCTP was incorporated, synthesis proceeded readily giving rise to almost 100% of full-length product (Figure 6.5).

DV and 9N were able to incorporate 2'F-araCTP as well as 2'F-rCTP under the conditions studied here (Figure 6.4B and Figure S6.4B in Supplementary Data).

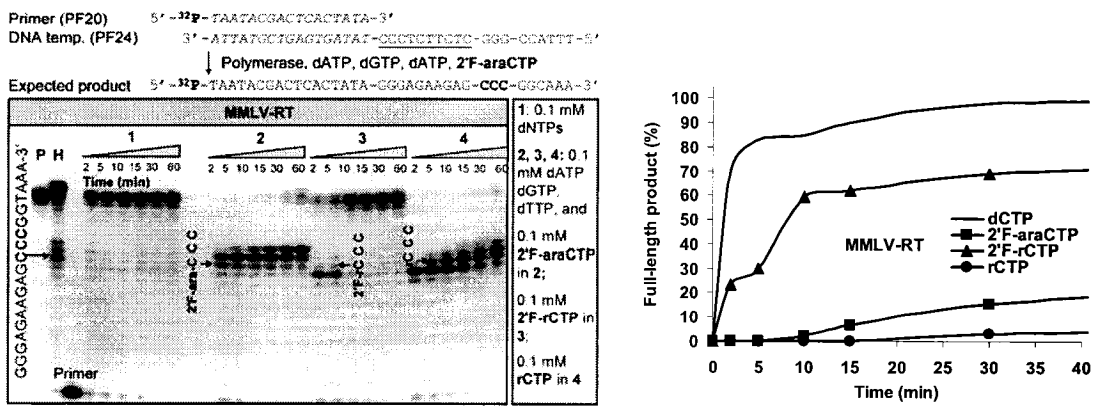


Figure 6.5. Efficiency of incorporation of 2'F-araCTP *vs* 2'F-rCTP through DNA template PF24 mediated by MMLV-RT DNA polymerization. Right figure is a plot of time-course of full-length product (%). Reactions conducted at 37°C, in a 40 μL reaction volume, with 0.6 μL (12 units) enzyme; triphosphate concentrations were shown; products shown on lane **H** were formed through a chain termination assay with 0.1 mM dATP, dGTP, dTTP, dCTP+ddCTP (1:1 molar ratio) on the template PF24, 0.4 μL HIV-1 RT in 20 μL reaction volume at 37°C. **P**: full-length product control.

In summary, incorporation of 2'F-araTTP proceeded more efficiently relative to 2'F-rUTP, while the incorporation of 2'F-araCTP is slightly less efficient than that observed with 2'F-rCTP. While the substrate specificity towards the 2'F-rUTP and rUTP pair were similar, a surprisingly different specificity towards the corresponding cytosine pair (i.e. 2'F-rCTP and rCTP) was observed. The order of incorporation efficiency of 2'F modified pyrimidine NTPs is summarized in Table 6.3.

Incorporation Efficiency of Pairs of Modified Triphosphates

Next, we conducted primer extension assays on the DNA template PF32 in order to synthesize chimeric (ribo/arabino/deoxy) oligonucleotides with 9N, DV and HIV-1 RT polymerases. The monomers consisted of dATP, dGTP and one of five different pyrimidine pairs: 2'F-araTTP/2'F-araCTP (Group 3 in Figure 6.6), 2'F-araTTP/2'F-rCTP (Group 4), 2'F-rUTP/2'F-araCTP (Group 5), 2'F-rUTP/2'F-rCTP (Group 6), and rUTP/rCTP (Group 7). Control reactions with the native dNTPs and rNTPs were carried out as well (Group 1 and 2 in Figure 6.6).

Incorporation efficiency was measured either as the percentage of full-length products, or the number of nucleotides incorporated if no full-length products were obtained. With 9N and DV, incorporation efficiencies followed the same trend (Table 6.3). Interestingly, HIV-1 RT produced opposite results for Groups 3, 4, 6 (Table 6.4). Generally, dNTPs are the best and rNTPs are the worst substrates and the combination of 2'F modified NTPs showed different incorporation efficiencies. The results are summarized in Table 6.4 and possible explanations for these results are discussed below.

Discussion

Previous studies showed that only thermophilic polymerases, namely, DV, 9N, Th and Ph, can generate full-length products of 2'F-ANA with all four 2'F-araNTPs (Chapter V; Peng and Damha, manuscript submitted). The present study showed that all these DNA incorporated three pyrimidine 2'F-araNTPs into a growing DNA strand and afforded full-length products (Table 6.2). These results contrast with those reported by Chaput and Szostak on DNA-template directed synthesis of α-L-threofuranosyl nucleic acids (TNA).[44] Among the various DNA polymerases studied, only DV carried out TNA

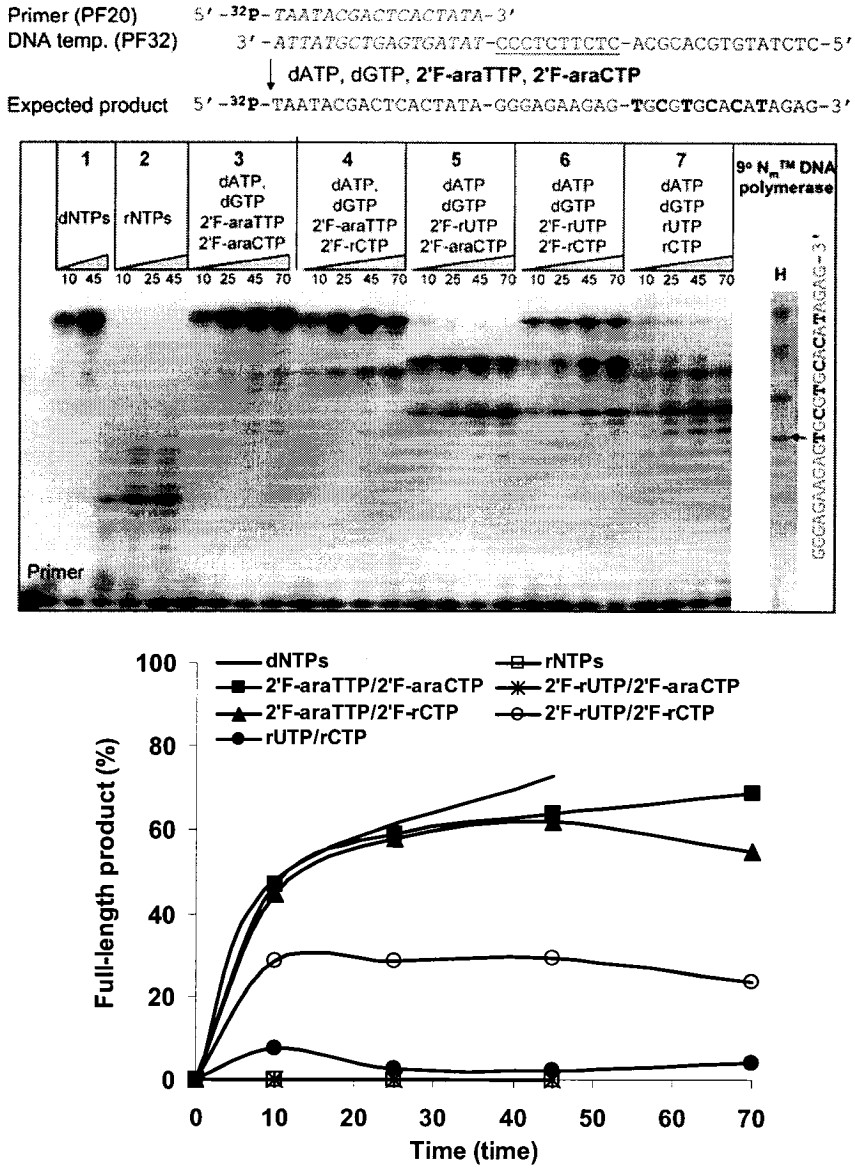


Figure 6.6. DNA template (PF32)-mediated DNA polymerization assay catalyzed by 9N DNA polymerase comparing the efficiency of incorporation of dATP, dGTP in the presence of various pyrimidine NTP combinations, namely: 2'F-araTTP/2'F-araTTP; 2'F-araTTP/2'F-rCTP; 2'F-rUTP/2'F-rCTP; and rUTP/rCTP. Reactions were carried out at 55°C; 40 μL reaction volume, 9N enzyme (2 μL; 4 units); 0.2 mM triphosphates; products shown on lane H1 were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio), 0.4 μL HIV-1 RT in 20 μL reaction volume on the same template PF32; Bottom figure is a plot of time-course of full-length product (%)

Table 6.4: The order of efficiency of incorporation of different triphosphates pairs

Polymerase	Order of efficiency of incorporation	Figures
9N	all four dNTPs (Group 1) > 2'F-araTTP/2'F-araCTP (Group 3) ≈ 2'F- araTTP/2'F-rCTP (Group 4) > 2'F-rUTP/2'F-rCTP (Group 6) > 2'F-rUTP/2'F-araCTP (Group 5) ≈ rUTP/rCTP (Group 7) >> all four rNTPs (Group 2)	Figure 6.6
DV	all four dNTPs (Group 1) ≈ 2'F-araTTP/2'F-araCTP (Group 3) ≈ 2'F-araTTP/2'F-rCTP (Group 4) > 2'F-rUTP/2'F-rCTP (Group 6) > 2'F-rUTP/2'F-araCTP (Group 5) > rUTP/rCTP (Group 7) > all four rNTPs (Group 2)	Figure S6.5; Supplementary data
HIV-1 RT	all four dNTPs (Group 1) >> 2'F-rUTP/2'F-rCTP (Group 6) > 2'F-araTTP/2'F-rCTP (Group 4) > 2'F-araTTP/2'F-araCTP (Group 3) > 2'F-rUTP/2'F-araCTP (Group 5) ≈ rUTP/rCTP (Group 7) > all four rNTPs (Group 2)	Figure S6.6. Supplementary data

synthesis. However incorporation efficiency was poor and at least 1 day incubation was necessary in order for these enzymes to couple three contiguous TNA residues.[44]

Incorporation of 2'F-araTTP proceeded more efficiently relative to 2'F-rUTP, but the opposite trend was observed for the 2'F-araCTP/2'F-rCTP pair (Table 6.3). Furthermore, when comparing polymerization patterns, 2'F-rUTP behaved like rUTP, but the same was not true for the cytosine pair, 2'F-rCTP and rCTP. Below we attempt to rationalize these results.

Co-crystallization of HIV-1 RT with a double-stranded DNA template-primer (binary complex: RT-dsDNA),[48] and with dsDNA and dNTP (ternary complex: RT-dsDNA-dNTP)[49] suggest that the dsDNA adopts an A-like (*north*) structure near the polymerase active site which gradually converts to the B-form conformation (*south*) as this DNA region moves away from polymerase active site. Therefore it is expected that incoming NTPs with an A-like conformation should be favored for binding at the active

site. Other factors such as steric interactions between incoming NTPs and the polymerase amino acid residues also help dictate the specificity of NTP binding. For instance, models based on crystal structures suggest that Phe155 in MMLV-RT (or Tyr-115 in HIV-1 RT) acts as a "steric gate" that discriminates between the 2'H of a dNTP and the 2'OH group of an NTP.[49-51] The elegant work of Marquez and others with conformationally biased NTP analogues and HIV-1 RT fully support this hypothesis.[52,53] In order to explain the contrasting inhibitory behavior of αFddATP and (the more potent) βFddATP they conducted molecular dynamics simulations on these NTPs and showed that αFddATP had unfavorable steric interaction with Tyr115 (HIV-1-RT) (analogous to the interaction of 2'-OH of rNTPs with Tyr115), even though αFddATP adopts the favored *north*-conformation for binding at the polymerase active site[52] (Figure 6.7). Our observation that incorporation of 2'F-araTTP proceeds more efficiently than that of 2'F-rUTP is consistent with the previous studies,[52] suggesting that when the fluorine is on the same face of the ring as the 5′ carbon (*i.e.* β-F) as in 2'F-araTTP, the steric interaction with Tyr115 in HIV-1 RT is avoided. 2'F-araTTP, however, must pay an energy penalty since it must switch to the *north* conformation in order to bind at the active site of RT. Such conformation is possible and precedented for arabinonucleosides[54] The more flexible dTTP thus shows the highest efficiency of incorporation.

Nucleoside analogs that do not block DNA synthesis at the point of incorporation, but only after additional incorporations of normal dNTPs have been also termed "delayed chain terminators", which have been shown to be relatively resistant to excision by HIV-1 RT mutant, one of HIV-1 RT's drug-resistance mechanisms.[55] The noticeable delayed pausing we have observed with 2'F-araNTPs may be due to the inclusion of three contigous 2'F-araT in the dsDNA. This likely induces an unfavorable conformational change in the RT-primer-template complex, which should result in a diminished rate of dNTP incorporation or, in the dissociation of RT from the primer-template duplex resulting in abortive products. The pausing observed *before* the incorporation of 2'F-rUTP or rUTP (Figure 6.3A) indicates that 2'F-rUTP or rUTP has difficulty in binding at the active site of RT, consistent with the proposal of steric gate control. Similar pausing *patterns* are observed for 2'F-rUTP and rUTP, consistent with the fact that 2'F-rUTP is a mimic of rNTPs and are both excellent substrates for T7 RNA polymerase.[27]

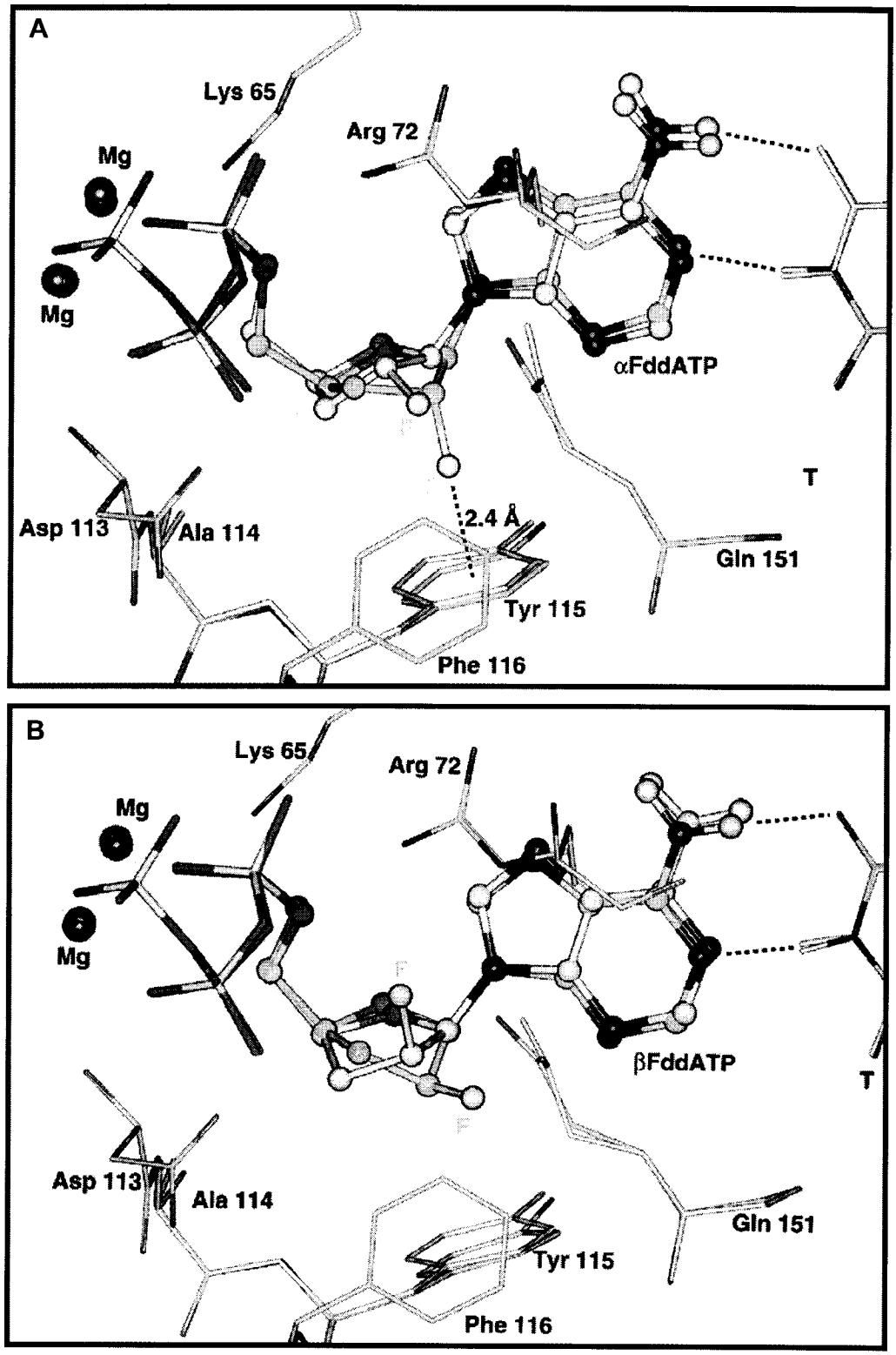


Figure 6.7: Static view of the dNTP binding site of RT in complex with north and south αFddATP in (**A**) and βFddATP in (**B**) (adapted from Mu, Marquez, et al. 2000[52])

Similar results are obtained for MMLV-RT, DV and 9N (Table 6.3). One can speculate that these enzymes could use the same discrimination mechanism (conformational preference and steric gate) to distinguish between dTTP, 2'F-araTTP, 2'F-rUTP and rUTP during incorporation. Our finding that the incorporation efficiency of 2'F-araTTP is higher than that of 2'F-rUTP by HIV-1 RT, MMLV-RT, DV and 9N is not surprising since 2'F-araTTP is a dTTP mimic (and thus favored by DNA polymerases), while 2'F-rUTP is a rUTP mimic, which should be more favored by RNA polymerases.

Surprisingly, a more complex pattern of incorporation efficiency was observed among the cytosine series, dCTP, 2'F-araCTP, 2'F-rCTP and rCTP in the primer extension reactions catalyzed by HIV-1 RT and MMLV RT. Based on the pausing patterns observed, binding of 2'F-araCTP at the active site of MMLV-RT seemed to be more advantageous over its ribose counterpart, 2'F-rCTP (Figure 6.5). This is again consistent with steric gate control for NTP binding.[52] The further elongation observed, however, indicates that other factors play important roles in single base incorporation, aside from the steric gate factor and conformational bias of sugar pucker. Previous studies on the fidelity of base incorporation have shown that the rate-limiting step could be different among multiple molecular checkpoints involved in DNA polymerase-directed polymerization.[56,57] Our experiments simply measure the overall result of all possible factors governing base incorporations. The factors governing the specificity and efficiency of base incorporation by DNA polymerases are complex and not completely understood.[58] Preferential geometry,[1] sugar pucker,[53] topology and size of the polymerase active site[1] are all important considerations. In our screening experiments, we found that the efficiency of incorporation of 2'F-araCTP < 2'F-araTTP in same primer extension conditions with some DNA polymerases (e.g. DV and 9N). For 2'F-rCTP, the literature has shown that the incorporation rate of 2'F-rCTP is much higher than that of 2'F-rUTP by human polymerases α and γ.[33] We expect that the observed incorporation efficiency of 2'F-rCTP $\geq$ 2'F-araCTP is the superposition of two factors: first, the overall incorporation efficiency of 2'F-rCTP may be higher than 2'F-rUTP by HIV-1 RT and MMLV-RT; second, the incorporation efficiency of 2'F-araCTP is lower than that of 2'F-araTTP. Why 2'F-rCTP differs from rCTP remains an unsolved question in this work and requires further study.

Delayed pausing was often observed with the incorporation of 2'F-araTTP and 2'F-araCTP (Figure 6.3, 6.4, 6.5, Figure S6.2 in Supplementary Data). These results are somewhat consistent with studies previously reported by our group, in which a DNA primer strand containing a single 3'-arabinocytidine unit (araC; in this case the 2'-substituent is an OH group) was found to serve as a primer of DNA synthesis. However, significant pausing of HIV-1 RT was observed after extension with 4 × dNTP units.[59]

The order of the incorporation efficiency of the NTP mixtures shown in Table 6.4 is an overall effect of the incorporation efficiencies of individual modified NTPs (Table 6.3). The first modified NTP incorporated is either 2'F-araTTP or 2'F-rUTP, so the rate limiting step could be the incorporation of the first NTP. This seems to be the case for 9N and DV. For example, the order of incorporation efficiency Group 3 ≈ Group 4 > Group 6 > Group 5 (Table 6.4) can be explained by the general trend of incorporation efficiency 2'F-araTTP > 2'F-rUTP and 2'F-rCTP ≥ 2'F-araCTP. The other combinations of rNTPs (Group 2 and Group 7) were not favored by 9N in this primer extension assay, consistent with the poor incorporation of rNTPs by DNA polymerases (Table 6.3).

The order of incorporation efficiency by HIV-1 RT is less predictable by the previous results of incorporation of pyrimidine 2'F-araNTPs *vs* 2'F-rNTPs. As shown in Table 6.4 and Figure S6.6 (Supplementary Data), 2'F-rUTP/2'F-rCTP (Group 6) incorporation is higher than that of 2'F-araTTP/2'F-rCTP (Group 4), therefore it suggests that the efficiency of incorporation of 2'F-rUTP is higher than that of 2'F-araTTP, not consistent with the pervious finding in Figure 6.3A. However, if we compare 2'F-araTTP/2'F-araCTP (Group 3) and 2'F-rUTP/2'F-araCTP (Group 5), the order of efficiency of incorporation tells us that 2'F-araTTP is better incorporated than 2'F-rUTP, consistent with the finding from Figure 6.3A. The efficiency of incorporation of 2'F-araTTP/2'F-rCTP (Group 4) > 2'F-araTTP/2'F-araCTP (Group 3) is consistent with the finding that 2'F-rCTP is a better substrate than 2'F-araCTP (Figure 6.3B). This complication of the efficiency of incorporation by HIV-1 RT indicates that the rate-limiting step in HIV-1 RT should be different from DV and 9N. The incorporation of 2'F modified CTPs seems more important in determining the yield of the full-length products.

Conclusion

2'F-araTTP and 2'F-araCTP are substrates of all the DNA polymerases tested in the present study. DV and 9N were the most efficient at incorporating 2'F-araTTP and 2'F-araCTP and afforded good yields of full length chimeric 2'F-ANA/DNA products. Our comparative studies of the efficiency of incorporation of 2'F-araNTPs *vs* 2'F-rNTPs by four DNA polymerases demonstrate that, under the conditions used, incorporation of 2'F-araTTP proceeded much more efficiently relative to 2'F-rUTP, while the incorporation of 2'F-araCTP was only slightly less efficient than that observed with 2'F-rCTP (Table 6.3). Generally, dNTPs were found to be better substrates than rNTPs; 2'F-rUTP behaves much like rUTP but 2'F-rCTP is strikingly different from rCTP. Furthermore, we demonstrate that mixtures of 2'F-araNTPs and 2'F-rNTPs can be used to generate 2'F-r/2'F-ara-oligonucleotides through template-directed DNA polymerization reactions. The 2'F-araTTP and 2'F-rCTP pair gave excellent results with the DNA polymerases tested. Considering the potential advantages of 2'F-ANA and 2'F-RNA modifications in oligonucleotide-based gene silencing approaches, our studies suggest that it should be possible to generate chimeric 2'F-ANA/2'F-RNA/DNA aptamers through an *in vitro* selection (SELEX) process.

Acknowledgements

This work was supported through a grant from the Canadian Institutes of Health Research (CIHR) and Topigen Pharmaceuticals, Inc. We thank J. Watts, R. Donga and A. Wahba for several helpful discussion, and J. Watts for assistance with the preparation of this manuscript. CGP acknowledges support from a FQRNT fellowship and a Clifford Wong McGill Major Fellowship.

Supplementary Data

Polymerase information, buffer conditions, and extra gel pictures are available on pp230-236.

References

1. Kool, E.T. (**2002**). Active site tightness and substrate fit in DNA replication. *Annu. Rev. Biochem.*, 71, 191-219.

2. Wright, G.E. and Brown, N.C. (**1990**). Deoxyribonucleotide analogs as inhibitors and substrates of DNA polymerases. *Pharmacol. Ther.*, 47, 447-497.

3. Smith, R.A., Sidwell, R.W. and Robins, R.K. (**1980**). Anti-viral mechanisms of action. *Annu. Rev. Pharmacol. Toxicol.*, 20, 259-284.

4. Brody, E.N. and Gold, L. (**2000**). Aptamers as therapeutic and diagnostic agents. *Reviews in Molecular Biotechnology*, 74, 5-13.

5. Nimjee, S.M., Rusconi, C.P. and Sullenger, B.A. (**2005**). Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.*, 56, 555-583.

6. Ellington, A.D. and Szostak, J.W. (**1990**). *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346, 818-822.

7. Tuerk, C. and Gold, L. (**1990**). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249, 505-510.

8. Joyce, G.F. (**1989**). Amplification, mutation and selection of catalytic RNA. *Gene*, 82, 83-87.

9. Ng, E.W.M., Shima, D.T., Calias, P., Cunningham, E.T., Guyer, D.R. and Adamis, A.P. (**2006**). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov*, 5, 123-132.

10. Thibaudeau, C. and Chattopahyaya, J. (**1999**). *Stereoelectronic Effects in Nucleosides and Nucleotides and Their Structural Implications.* Uppsala University Press: Uppsala Sweden.

11. Berger, I., Tereshko, V., Ikeda, H., Marquez, V.E. and Egli, M. (**1998**). Crystal structures of B-DNA with incorporated 2'-deoxy-2'-fluoro-arabino-furanosyl thymines: implications of conformational preorganization for duplex stability. *Nucleic Acids Res.*, 26, 2473-2480.

12. Trempe, J.-F., Wilds, C.J., Denisov, A.Y., Pon, R.T., Damha, M.J. and Gehring, K. (**2001**). NMR solution structure of an oligonucleotide hairpin with a 2'F-ANA/RNA stem: Implications for RNase H specificity toward DNA/RNA hybrid duplexes. *J. Am. Chem. Soc.*, 123, 4896-4903.

13. Guschlbauer, W. and Jankowski, K. (**1980**). Nucleoside conformation is determined by the electronegativity of the sugar substituent. *Nucleic Acids Res.*, 8, 1421-1433.

14. Catlin, J.C. and Guschlbauer, W. (**1975**). Oligonucleotide conformations. 3. Comparative optical and thermodynamic studies of uridylyl-3'-5'-nucleosides containing ribose, deoxyribose, or 2'-deoxy-2'-fluororibose in uridine moiety. *Biopolymers*, 14, 51-72.

15. Ikeda, H., Fernandez, R., Wilk, A., Barchi Jr, J.J., Huang, X. and Marquez, V.E. (**1998**). The

effect of two antipodal fluorine-induced sugar puckers on the conformation and stability of the Dickerson-Drew dodecamer duplex [d(CGCGAATTCGCG)]$_2$. *Nucleic Acids Res.*, 26, 2237-2244.

16. Monia, B.P., Lesnik, E.A., Gonzalez, C., Lima, W.F., Mcgee, D., Guinosso, C.J., Kawasaki, A.M., Cook, P.D. and Freier, S.M. (**1993**). Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene-expression. *J. Biol. Chem.*, 268, 14514-14522.

17. Kawasaki, A.M., Casper, M.D., Freier, S.M., Lesnik, E.A., Zounes, M.C., Cummins, L.L., Gonzalez, C. and Cook, P.D. (**1993**). Uniformly modified 2'-deoxy-2'-fluoro phosphorothioate oligonucleotides as nuclease-resistant antisense compounds with high-affinity and specificity for RNA targets. *J. Med. Chem.*, 36, 831-841.

18. Williams, D.M., Benseler, F. and Eckstein, F. (**1991**). Properties of 2'-fluorothymidine-containing oligonucleotides: interaction with restriction endonuclease *Eco*RV. *Biochemistry*, 30, 4001-4009.

19. Manoharan, M. (**1999**). 2'-Carbohydrate modifications in antisense oligonucleotide therapy: importance of conformation, configuration and conjugation. *BEB-Gene Struct Expr*, 1489, 117-130.

20. Wilds, C.J. and Damha, M.J. (**2000**). 2'-Deoxy-2'-fluoro-β-D-arabinonucleosides and oligonucleotides (2'F-ANA): synthesis and physicochemical studies. *Nucleic Acids Res.*, 28, 3625-3635.

21. Wilds, C.J. and Damha, M.J. (**1999**). Duplex recognition by oligonucleotides containing 2'-deoxy-2'-fluoro-D-arabinose and 2'-deoxy-2'-fluoro-D-ribose. Intermolecular 2'-OH-phosphate contacts versus sugar puckering in the stabilization of triple-helical complexes. *Bioconjug. Chem.*, 10, 299-305.

22. Wilds, C.J. (**2000**). Ph.D. thesis, McGill University, Montreal.

23. Pieken, W.A., Olsen, D.B., Williams, D.M., Aurup, H., Benseler, F. and Eckstein, F. (**1991**). Influence of 2'-modifications on the structure and function of hammerhead ribozymes. *Bio. Chem. Hoppe-Seyler*, 372, 731-731.

24. Capodici, J., Kariko, K. and Weissman, D. (**2002**). Inhibition of HIV-1 infection by small interfering RNA-mediated RNA interference. *J. Immunol.*, 169, 5196-5201.

25. Dowler, T., Bergeron, D., Tedeschi, A.L., Paquet, L., Ferrari, N. and Damha, M.J. (**2006**). Improvements in siRNA properties mediated by 2'-deoxy-2'-fluoro-beta-D-arabinonucleic acid (FANA). *Nucleic Acids Res.*, 34, 1669-1675.

26. Kalota, A., Karabon, L., Swider, C.R., Viazovkina, E., Elzagheid, M., Damha, M.J. and Gewirtz, A.M. (**2006**). 2'-Deoxy-2'-fluoro-beta-D-arabinonucleic acid (2'F-ANA) modified oligonucleotides (ON) effect highly efficient, and persistent, gene silencing. *Nucleic Acids Res.*, 34, 451-461.

27. Aurup, H., Williams, D.M. and Eckstein, F. (**1992**). 2'-Fluoro-2'-deoxynucleoside and 2'-amino-2'-deoxynucleoside 5'-triphosphates as substrates for T7 RNA-polymerase.

Biochemistry, 31, 9636-9641.

28. Kubik, M.F., Bell, C., Fitzwater, T., Watson, S.R. and Tasset, D.M. (**1997**). Isolation and characterization of 2'-fluoro-, 2'-amino-, and 2'-fluoro-/amino-modified RNA ligands to human IFN-gamma that inhibit receptor binding. *J. Immunol.*, 159, 259-267.

29. Ruckman, J., Green, L.S., Beeson, J., Waugh, S., Gillette, W.L., Henninger, D.D., Claesson-Welsh, L. and Janjic, N. (**1998**). 2'-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *J. Biol. Chem.*, 273, 20556-20567.

30. Pagratis, N.C., Bell, C., Chang, Y.-F., Jennings, S., Fitzwater, T., Jellinek, D. and Dang, C. (**1997**). Potent 2'-amino-, and 2'-fluoro-2'-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor. *Nat. Biotechnol.*, 15, 68-73.

31. Aoyama, H., Cottin, L.S., Tarragolitvak, L., Litvak, S. and Guschlbauer, W. (**1985**). 2'-Fluoro-2'-deoxycytidine triphosphate as a substrate for RNA-dependent and DNA-dependent DNA polymerases. *Biochim. Biophys. Acta*, 824, 218-224.

32. Helfman, W.B., Hendler, S.S., Shannahoff, D.H. and Smith, D.W. (**1978**). *Escherichia-coli* DNA polymerase-II and Polymerase-III - substrate-specificity. *Biochemistry*, 17, 1607-1611.

33. Richardson, F.C., Kuchta, R.D., Mazurkiewicz, A. and Richardson, K.A. (**2000**). Polymerization of 2'-fluoro- and 2'-*O*-methyl-dNTPs by human DNA polymerase α, polymerase γ, and primase. *Biochem. Pharmacol.*, 59, 1045-1052.

34. Ono, T., Scalf, M. and Smith, L.M. (**1997**). 2'-Fluoro modified nucleic acids: polymerase-directed synthesis, properties and stability to analysis by matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Res.*, 25, 4581-4588.

35. Montgomery, J.A., Shortnacyfowler, A.T., Clayton, S.D., Riordan, J.M. and Secrist, J.A. (**1992**). Synthesis and biologic activity of 2'-fluoro-2-halo derivatives of 9-beta-D-arabinofuranosyladenine. *J. Med. Chem.*, 35, 397-401.

36. Kong, X.B., Scheck, A.C., Price, R.W., Vidal, P.M., Fanucchi, M.P., Watanabe, K.A., Fox, J.J. and Chou, T.C. (**1988**). Incorporation and metabolism of 2'-fluoro-5-substituted arabinosyl pyrimidines and their selective-inhibition of viral-DNA synthesis in herpes-simplex virus type-1 (HSV-1)-infected and mock-infected vero cells. *Antiviral Res.*, 10, 153-166.

37. Ruth, J.L. and Cheng, Y.C. (**1981**). Nucleoside analogs with clinical potential in antivirus chemotherapy - the effect of several thymidine and 2'-deoxycytidine analog 5'-triphosphates on purified human (alpha, beta) and herpes-simplex virus (type-1, type-2) DNA-polymerases. *Mol. Pharmacol.*, 20, 415-422.

38. Mar, E.C., Chiou, J.F., Cheng, Y.C. and Huang, E.S. (**1985**). Human cytomegalovirus-induced DNA-polymerase and its interaction with the triphosphates of 1-(2'-deoxy-2'-fluoro-beta-D-arabinofuranosyl)-5-methyluracil, 1-(2'-deoxy-2'-fluoro-beta-D-arabino furanosyl)-5-iodocytosine, and 1-(2'-deoxy-2'-fluoro-beta-D-arabinofuranosyl)-5-methylcytosine. *J. Virol.*, 56, 846-851.

39. Faderl, S., Gandhi, V., Keating, M.J., Jeha, S., Plunkett, W. and Kantarjian, H.M. (**2005**). The role of clofarabine in hematologic and solid mal ignancies - development of a next-generation nucleoside analog. *Cancer*, 103, 1985-1995.

40. Sun, H.H., Sloan, A., Mangner, T.J., Vaishampayan, U., Muzik, O., Collins, J.M., Douglas, K. and Shields, A.F. (**2005**). Imaging DNA synthesis with [F-18]FMAU and positron emission tomography in patients with cancer. *Eur J Nucl Med Mol I*, 32, 15-22.

41. Conti, P.S., Alauddin, M.M., Fissekis, J.R., Schmall, B. and Watanabe, K.A. (**1995**). Synthesis of 2'-fluoro-5-[C-11]-methyl-1-beta-D-arabinofuranosyluracil ([C-11]-FMAU) - a potential nucleoside analog for in-vivo study of cellular proliferation with PET. *Nucl. Med. Biol.*, 22, 783-789.

42. Shields, A.F. (**2006**). Positron emission tomography measurement of tumor metabolism and growth: Its expanding role in oncology. *Mol Imaging Biol*, 8, 141-150.

43. Peng, C.G. and Damha, M.J. (**2006**). Probing DNA polymerase activity with 2'-deoxy-2'-fluoroarabinonucleoside and 2'-deoxy-2'-fluororibonucleoside 5'-triphosphates. *Nucleosides, Nucleotides and Nucleic Acids Conference Book*, pp. 287.

44. Chaput, J.C. and Szostak, J.W. (**2003**). TNA synthesis by DNA polymerases. *J. Am. Chem. Soc.*, 125, 9274-9275.

45. Kato, Y., Minakawa, N., Komatsu, Y., Kamiya, H., Ogawa, N., Harashima, H. and Matsuda, A. (**2005**). New NTP analogs: the synthesis of 4'-thioUTP and 4'-thioCTP and their utility for SELEX. *Nucleic Acids Res.*, 33, 2942-2951.

46. Viazovkina, E., Mangos, M.M., Elzagheid, M.I. and Damha, M.J. (**2002**). Synthesis of 2'-fluoroarabinonucleoside phosphoramidites and their use in the synthesis of 2'F-ANA. *In Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons, pp. 4.15.11-14.15.22.

47. Galarneau, A., Min, K.-L., Mangos, M.M. and Damha, M.J. (**2005**). Assay for evaluating ribonuclease H-mediated degradation of RNA-antisense oligonucleotide duplexes. *Methods in Molecular Biology (Totowa, NJ, United States)*, 288, 65-80.

48. Jacobomolina, A., Ding, J.P., Nanni, R.G., Clark, A.D., Lu, X.D., Tantillo, C., Williams, R.L., Kamer, G., Ferris, A.L., Clark, P. *et al.* (**1993**). Crystal-structure of human-immunodeficiency-virus type-1 reverse-transcriptase complexed with double-stranded DNA at 3.0 angstrom resolution shows bent DNA. *Proc. Natl. Acad. Sci. U. S. A.*, 90, 6320-6324.

49. Huang, H., Chopra, R., Verdine, G.L. and Harrison, S.C. (**1998**). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, 282, 1669-1675.

50. Ding, J.P., Das, K., Hsiou, Y., Sarafianos, S.G., Clark, A.D., Jacobo-Molina, A., Tantillo, C., Hughes, S.H. and Arnold, E. (**1998**). Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 angstrom resolution. *J. Mol. Biol.*, 284, 1095-1111.

51. Boyer, P.L., Sarafianos, S.G., Arnold, E. and Hughes, S.H. (**2000**). Analysis of mutations at positions 115 and 116 in the dNTP binding site of HIV-1 reverse transcriptase. *Proc. Natl. Acad. Sci. U. S. A.*, 97, 3056-3061.

52. Mu, L., Sarafianos, S.G., Nicklaus, M.C., Russ, P., Siddiqui, M.A., Ford Jr., H., Mitsuya, H., Le, R., Kodama, E., Meier, C. *et al.* (**2000**). Interactions of conformationally biased north and south 2'-fluoro-2',3'-dideoxynucleoside 5'-triphosphates with the active site of HIV-1 reverse transcriptase. *Biochemistry*, 39, 11205-11215.

53. Marquez, V.E., Ezzitouni, A., Russ, P., Siddiqui, M.A., Ford Jr., H., Feldman, R.J., Mitsuya, H., George, C. and Barchi Jr., J.J. (**1998**). HIV-1 reverse transcriptase can discriminate between two conformationally locked carbocyclic AZT triphosphate analogs. *J. Am. Chem. Soc.*, 120, 2780-2789.

54. Li, F., Sarkhel, S., Wilds, C.J., Wawrzak, Z., Prakash, T.P., Manoharan, O.M. and Egli, M. (**2006**). 2'-Fluoroarabino- and arabinonucleic acid show different conformations, resulting in deviating RNA affinities and processing of their heteroduplexes with RNA by RNase H. *Biochemistry*, 45, 4141-4152.

55. Boyer, P.L., Julias, J.G., Marquez, V.E. and Hughes, S.H. (**2005**). Fixed conformation nucleoside analogs effectively inhibit excision-proficient HIV-1 reverse transcriptases. *J. Mol. Biol.*, 345, 441-450.

56. Radhakrishnan, R., Arora, K., Wang, Y.L., Beard, W.A., Wilson, S.H. and Schlick, T. (**2006**). Regulation of DNA repair fidelity by molecular checkpoints: "Gates" in DNA polymerase beta's substrate selection. *Biochemistry*, 45, 15142-15156.

57. Joyce, C.M. and Benkovic, S.J. (**2004**). DNA polymerase fidelity: Kinetics, structure, and checkpoints. *Biochemistry*, 43, 14317-14324.

58. Kunkel, T.A. (**2004**). DNA replication fidelity. *J. Biol. Chem.*, 279, 16895-16898.

59. Borkow, G., Arion, D., Noronha, A., Scartozzi, M., Damha, M.J. and Parniak, M.A. (**1997**). Inhibitory potency of R-region specific antisense oligonucleotides against in vitro DNA polymerization and template-switching reactions catalyzed by HIV-1 reverse transcriptase. *Int. J. Biochem. Cell Biol.*, 29, 1285-1295.

Chapter VI. Assessing Polymerase Activity with 2'-Deoxy-2'-fluoro-β-D-ribonucleoside 5'-Triphosphates (2'F-rNTPs) and 2'-Deoxy-2'-fluoro-β-D-arabinonucleoside 5'-Triphosphates (2'F-araNTPs)

Supplementary Data

(pp230-236)

Contents

Table S6.1: Polymerase information and default primer extension assay conditions

Polymerase	Activity (units/μL)	Reaction buffer (x1)
Thermophilic DNA polymerase (reaction temperature at 55°C in this study)		
DV	2	20 mM Tris-HCl (pH 8.8 at 25°C), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1 % Triton X-100; 0.2 mM dNTPs
9N	2	Same as above
Th	2	Same as above
Bst	8	Same as above
Taq	5	Same as above
Ph	2	5X Phusion™ HF Buffer; 0.2 mM dNTPs
Mesophilic DNA polymerase (reaction temperature at 37°C in this study)		
KF	5	10 mM Tris-HCl (pH 7.5 at 25°C), 5 mM MgCl2, 7.5 mM DTT; 0.033 mM dNTPs
Reverse transcriptase (reaction temperature at 37°C in this study)		
HIV	27.3	50 mM Tris-HCl (pH 7.8 at 25°C), 60 mM KCl, 2.5 mM $MgCl_2$; 0.2 mM dNTPs
MMLV	10	50 mM Tris-HCl (pH 8.3 at 25°C), 50 mM KCl, 4 mM $MgCl_2$, 10mM DTT; 0.5 mM dNTPs

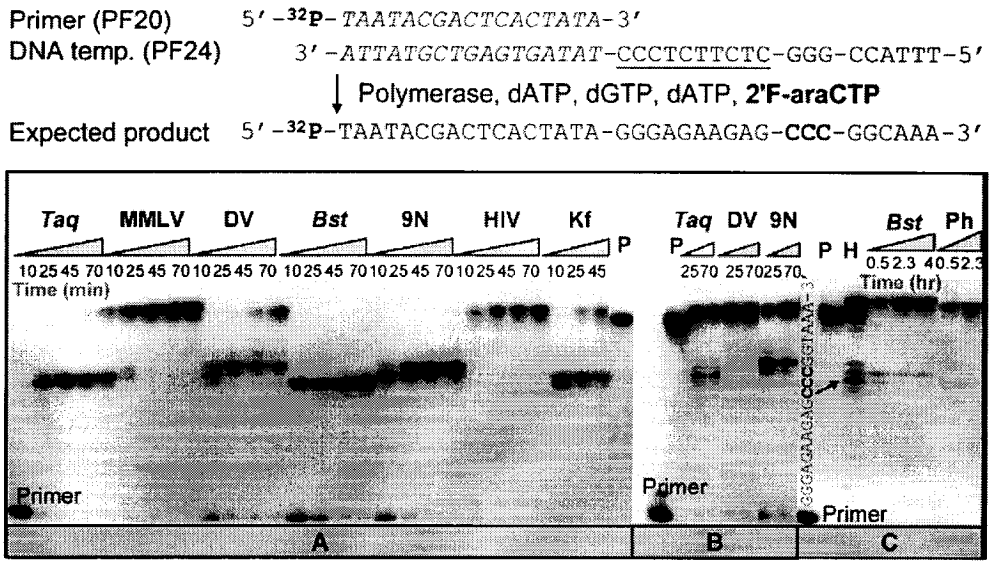


Figure S6.1: DNA-template directed polymerization of three contiguous 2'F-araC nucleotide units catalyzed by DNA polymerases on a DNA-template PF24. **A**: reactions at 37 °C, 20 μL reaction volume, 0.5 μL enzyme, dNTPs and 2'F-araCTP: 0.01 mM for MMLV and 0.05 mM for HIV; reactions at 37 °C, 20 μL reaction volume, 1 μL enzyme, dNTP and 2'F-araCTP: 0.033 mM for Kf and 0.5 mM for MMLV; **B**: reactions at 55 °C, 40 μL reaction volume, 2 μL enzyme added, 0.2 mM triphosphates; **C**: reactions at 55 °C, 20 μL reaction volume, 2 μL *Bst* and 0.4 μL Ph., 0.2 mM triphosphates; products shown on lane H were formed through a chain termination assay with 0.1 mM dATP, dGTP, dTTP, dCTP+ddCTP (1:1 molar ratio), 0.4 μL HIV-1 RT in 20 μL reaction volume, on the same template PF24.

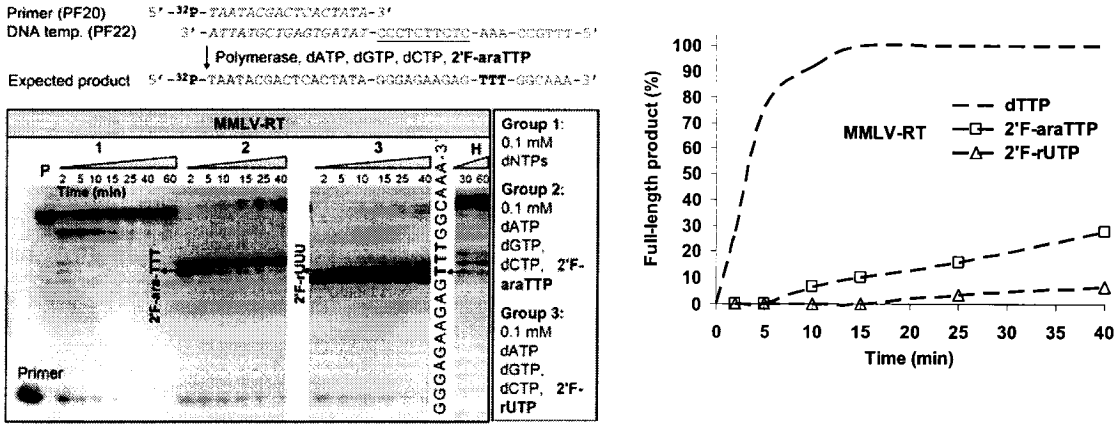


Figure S6.2. Efficiency of incorporation of 2'F-araTTP *vs* 2'F-rUTP through DNA template PF22 mediated by MMLV-RT DNA polymerization. Right figure is a plot of time-course of full-length product (%). Reactions conducted at 37°C, in a 60 µL reaction volume, with 1 µL (20 units) enzyme; triphosphate concentrations were shown; products shown on lane **H** were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio) on the template PF22, 0.4 µL HIV-1 RT in 20 µL reaction volume at 37°C. **P**: full-length product control.

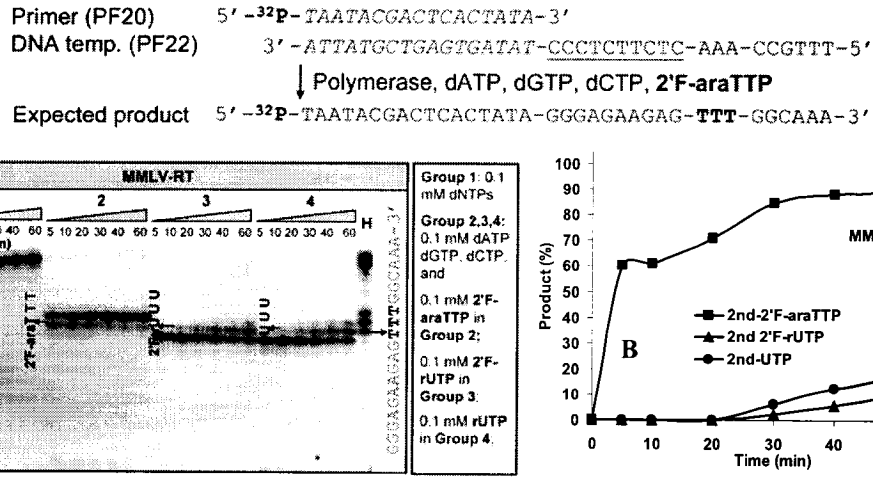
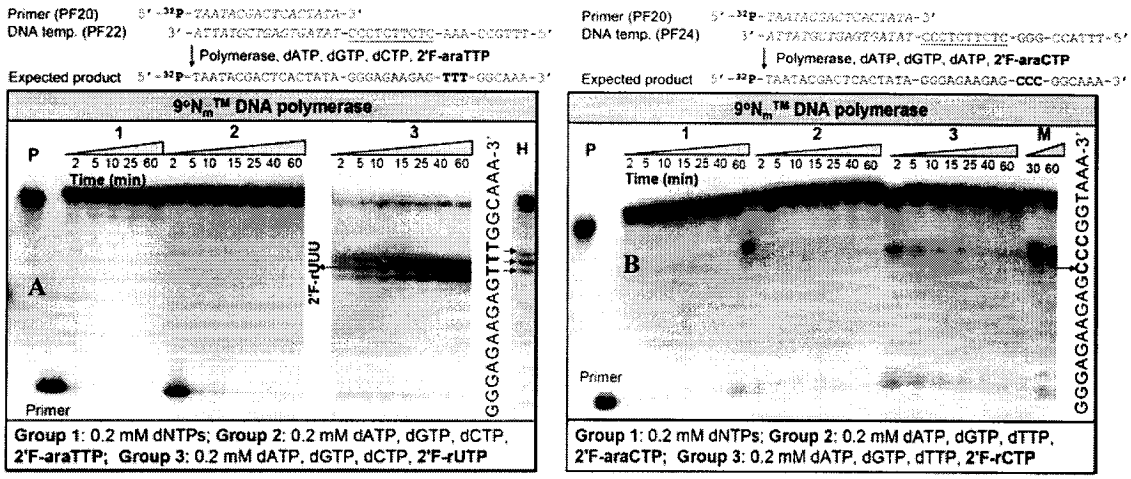


Figure S6.3: Efficiency of incorporation of 2'F-araTTP *vs* 2'F-rUTP through DNA template (PF22)-mediated MMLV-RT DNA polymerization. The base sequence of the primer, DNA template (PF22) and resulting chimeric DNA-2'F-ANA oligonucleotide product is shown. Reactions conducted at 37°C in a 40 µL reaction volume with 0.6 µL (12 units) enzyme, triphosphate concentrations shown; products shown on lane H were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio), 0.4 µL HIV-1 RT (8 units) in 20 µL reaction volume on the same template PF22 at 37°C.; (**B**):Time-course products with the 2nd incorporations of 2'F-araTTP, 2'F-dUTP or UTP (%) based on gel (**A**).



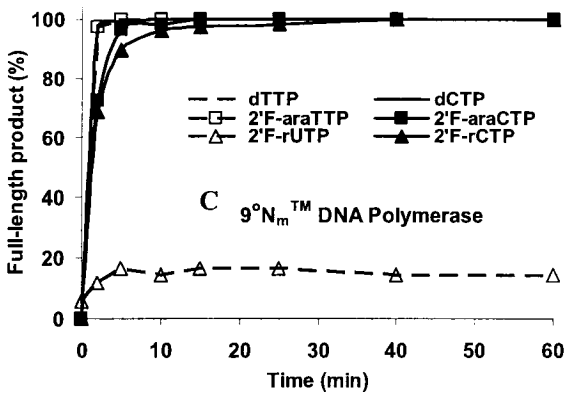


Figure S6.4: Efficiency of incorporation of 2'F-araTTP *vs* 2'F-rUTP through DNA template PF22 (**A**), and 2'F-araC *vs* 2'F-rC through DNA template PF24 (**B**) mediated by 9N DNA polymerization. (**C**) is a plot of time-course of full-length product (%). Reactions conducted at 55°C in a 60 µL reaction volume, with 3 µL (6 units) 9N DNA pol; triphosphate concentrations were shown; products shown on lane **H** were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio), 0.4 µL HIV-1 RT in 20 µL reaction volume on the template PF22 at 37°C. ; products shown on lane **M** were formed through a chain termination assay with 0.2 mM dATP, dGTP, dTTP, dCTP+ddCTP (1:1 molar ratio), 0.4 µL MMLV-RT in 20 µL reaction volume on the template PF24 at 37°C. **P**: full-length product control.

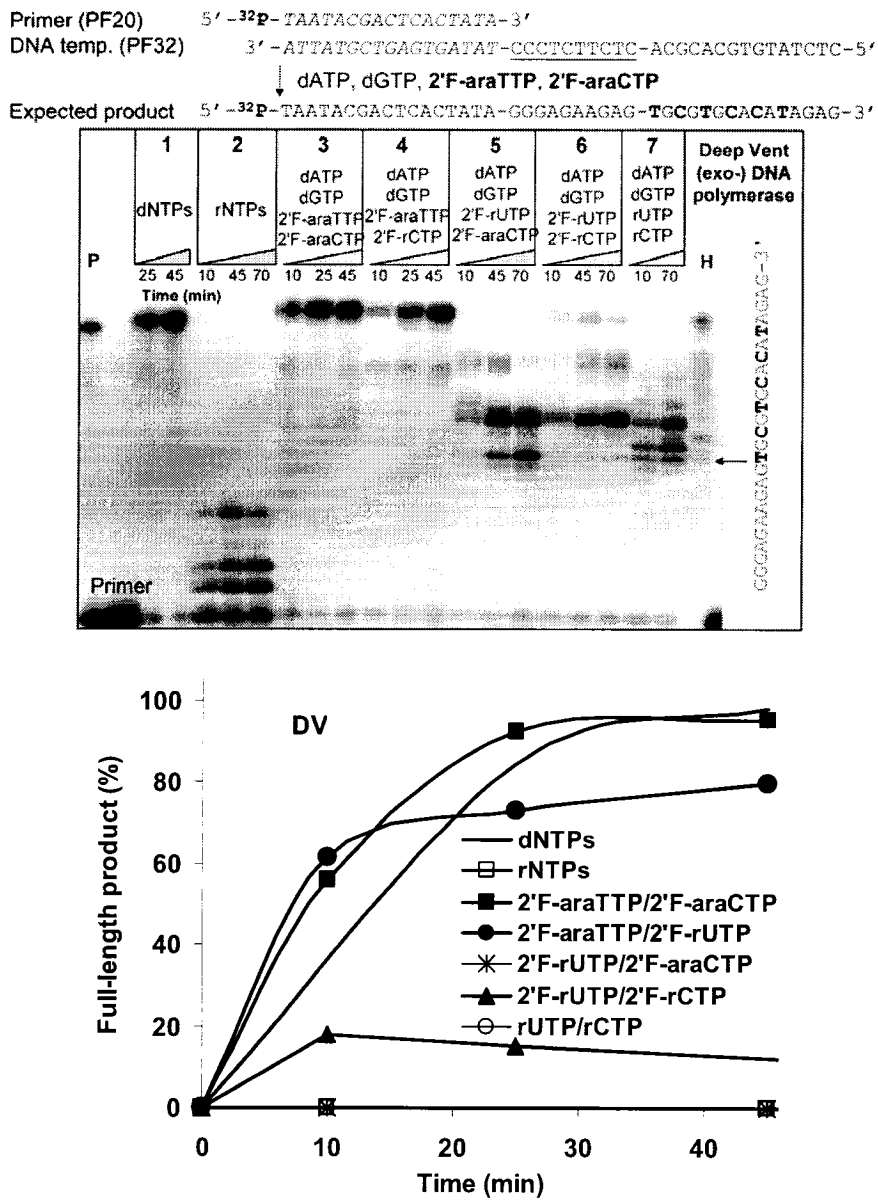


Figure S6.5: DNA template (PF32)-mediated DNA polymerization catalyzed by DV DNA polymerase comparing the efficiency of incorporation of dATP, dGTP in the presence of various pyrimidine NTP combinations, namely: 2'F-araTTP/2'F-araTTP; 2'F-araTTP/2'F-dCTP; 2'F-dUTP/2'F-dCTP; and UTP/CTP. Reactions were carried out at 55°C; 40 μL reaction volume, 9N enzyme (2 μL; 4 units); 0.2 mM triphosphatess; products shown on lane H were formed through a chain termination assay with 0.1 mM dATP, dGTP, dCTP, dTTP+ddTTP (1:1 molar ratio), 0.4 μL HIV-1 RT in 20 μL reaction volume on the same template PF32; The bottom figure is a plot of time-course of full-length product (%).

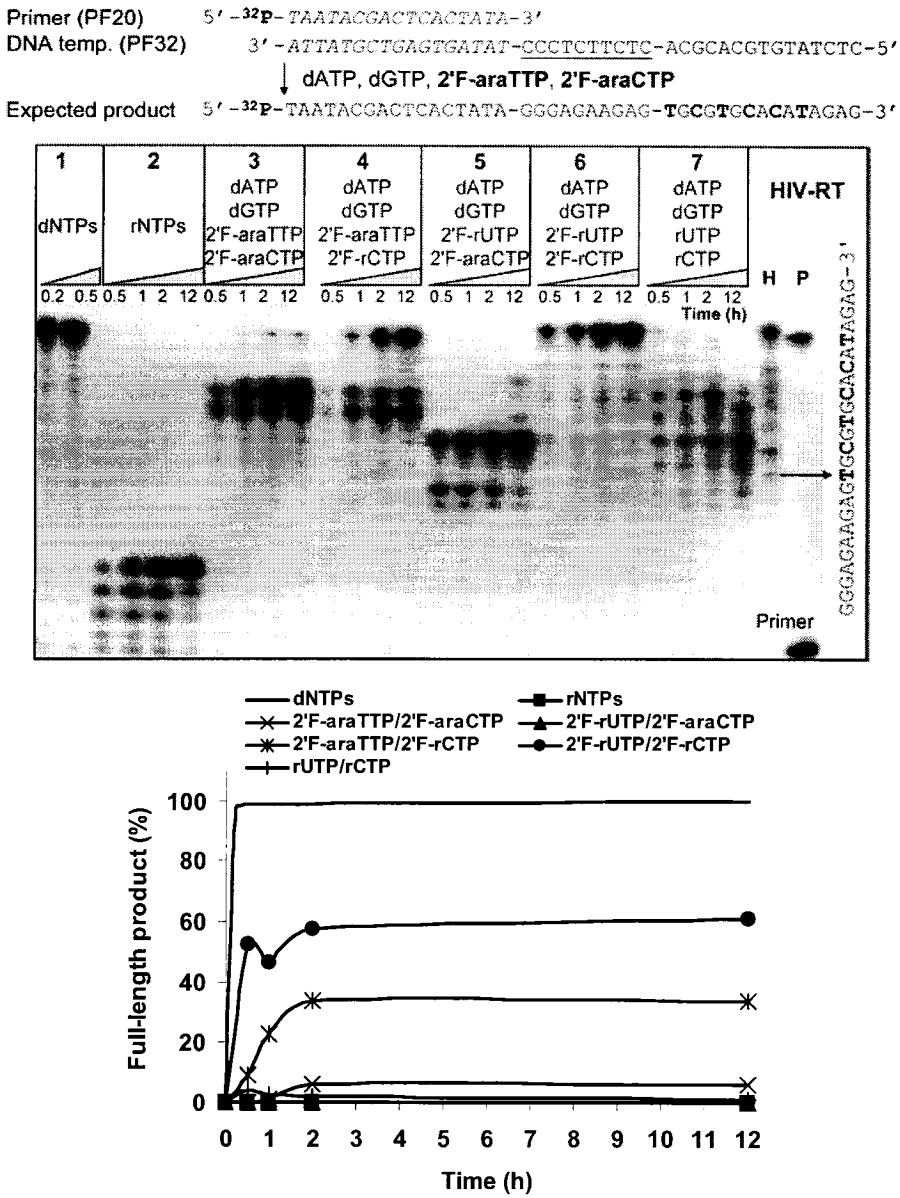


Figure S6.6: DNA template (PF32)-mediated DNA polymerization catalyzed by HIV-1 RT DNA polymerase comparing the efficiency of incorporation of dATP, dGTP in the presence of various pyrimidine NTP combinations, namely: 2'F-araTTP/2'F-araTTP; 2'F-araTTP/2'F-dCTP; 2'F-dUTP/2'F-dCTP; and rUTP/rCTP. Reactions at 37°C, 40 μL reaction volume, 1 μL (27.3 units) HIV-1 RT (except for 0.4 μL for the first lane); 0.2 mM TPs (except for 0.1 mM dNTPs in the first land); H control: 0.1 mM dATP, dGTP, dTTP+ddTTP (1:1), dCTP+ddCTP (1:1), HIV-1 RT, the same template PF32 at 37°C for 15min; The bottom figure is a plot of time-course of full-length product (%).

Chapter VII. Contributions to Knowledge

7.1. Summary and Conclusions

In an effort to develop novel oligonucleotide analogues as potential therapeutics and/or diagnostic reagents, specifically for their use as antisense and aptamer based technologies, the chemical synthesis of two C2'-modified oligonucleotides containing 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine and 2'-deoxy- 2'-fluoro-β-D-arabinonucleosides were investigated. These were generated either through solid-phase or DNA polymerase-mediated oligonucleotide synthesis.

The 3'- and 2'-*O*-phosphoramidite derivatives of 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (or 2'-α-*hm*-dT, abbreviated as "**H**") were synthesized and incorporated into otherwise unmodified DNA, RNA and 2',5'-RNA chains. Thermal denaturation data showed that "**H**" units had a stabilizing as well as a destabilizing effect on duplexes depending on the phosphodiester connectivity (3',5" or 2',5') of the **H** units, and the nature of the accepting oligonucleotide strand (RNA, DNA *vs* 2',5'-RNA). Consistent with previous findings, 3',5'-linked **H** units were found to significantly destabilize *DNA*:RNA hybrids and *DNA*:DNA duplexes (the strand containing a few **H** units is in italics). The incorporation of 3',5'-linked **H** into *RNA*:RNA duplexes was found to be less destabilizing. It was also shown that 2',5'-linked **H** had a small destabilizing effect on *2',5'-RNA*:RNA and *RNA*:RNA duplexes, and in fact, they were shown to stabilize 2',5'-RNA hairpins when incorporated within their constituent 2',5'-linked tetraloop. These results were rationalized in terms of the "compact" and "extended" conformations of nucleotides.

A series of branched RNAs (Y-shaped) related to yeast pre-mRNA splicing intermediates were synthesized incorporating both natural (*i.e.* ribose) and non-natural (*i.e.* "**H**" and acyclic nucleoside) branchpoints in order to examine the effect of sugar and phosphodiester configuration on hydrolytic efficiency of the yeast debranching enzyme (yDBR). Selective hydrolysis of the 2',5'-phosphodiester bond by yDBR was evident only in the wild-type ribo-A and ribo-U branched RNA. On the other hand, the branchpoint analogs were resistant to enzyme hydrolysis even after a 60 min

incubation period, suggesting that the length of the 2',5'-internucleotide linkage (extended in the case of "**H**") dramatically affects debranching efficacy. Moreover, these studies have indicated a concentration dependant inhibition of wild-type Y-RNA debranching in the presence of **H**-containing branched RNA, such that a 50% reduction in debranching efficiency was observed at a concentration of 70 nM of inhibitor. The strength of the inhibition depended on the nature of the 2'-nucleotide residue. In this regard, a 2'-cytidine had the best inhibitory activity. Such an inhibitor would signify a potential choice for enzyme co-crystallization and X-ray analysis of the yDBR active site structure.

2'-Deoxy-2'-fluoro-β-D-arabinonucleic acid (2'F-ANA) has been developed in the Damha group for antisense and siRNA -mediated gene silencing. This thesis explored, for the first time, the potential of this promising modification in aptamer applications. Specifically, the impact of the 2'F-ANA modification on stability and conformation of the well known thrombin-binding aptamer d($G_2T_2G_2TGTG_2T_2G_2$), the anti-HIV phosphorothioate sequence PS-d($T_2G_4T_2$), and a DNA telomeric sequence d($G_4T_4G_4$) were investigated. UV thermal melting (T_m) and circular dichroism (CD) experiments revealed that 2'F-ANA can stabilize G-quartets and maintain the quadruplex conformation provided only the *anti*-deoxyguanosines are replaced with 2'F-ANA residues. Replacement of *syn*-deoxyguanosines with 2'F-ANA resulted in a dramatic switch to an alternative and in some cases very stable, quadruplex conformation. The data also showed that incorporation of 2'F-ANA residues into thrombin-binding DNA stabilizes the complex (ΔT_m up to ~ + 3°C/ 2'F-ANA modification). The 2'F-ANA modification also increased the half-life in 10% fetal bovine serum (FBS) up to 48 fold. Two 2'F-ANA-modified thrombin-binding aptamers (PG13 and PG14) showed enhanced binding affinity to thrombin nearly four fold, as assessed by a nitrocellulose filter binding assay. PG13 and PG14 also exhibited increased thermal stability (~1°C increase in T_m) and nuclease resistance (4-7 fold). Therefore, these studies have opened new perspectives for the biological application of 2'F-ANA as aptamer oligonucleotides.

This thesis work also examined the polymerase-directed synthesis of 2'F-ANA oligonucleotides. A primer extension assays was adapted to screen several DNA polymerases for their ability to incorporate all four 2'F-araNTPs (N = A, G, T and C) and found that these monomers are very good substrates for commonly used thermophilic DNA polymerases. Remarkably, the selectivity (fidelity) for 2'F-araNTP (N= purines) incorporation by DV and 9N polymerases was comparable, if not better, than that of the native dNTPs. Selectivity toward the pyrimidine-based 2'F-araNTPs was significantly less, but the same was true for the corresponding pyrimidine-containing dNTPs. DV, 9N and Ph polymerases catalyzed synthesis of chimeric FANA-DNA strands on FANA-DNA templates, whereas Klenow fragment DNA polymerase (3'->5' exo-) (KF) and *Bst* DNA polymerase (*Bst*) were able to catalyze 2'F-ANA template-directed DNA synthesis. While it was not possible to synthesize an all-2'F-ANA strand on an all-2'FANA template, it was possible for DV, 9N and Ph polymerases to drive the formation of multiple FANA:FANA base pairs within a DNA-FANA chimeric duplex. All of these results suggest that it should be possible to evolve FANA-modified aptamers using *in vitro* selection.

Finally, this thesis provided the first comparison study of incorporation efficiency of the two isomeric NTPs (i.e. 2'F-araNTPs and 2'F-rNTPs) by thermophilic and retroviral DNA polymerases (DV, 9N, HIV-1 RT, and MMLV RT). Under our primer extension conditions, incorporation of 2'F-araTTP proceeded more efficiently than 2'F-rUTP, while the incorporation of 2'F-araCTP is slightly less efficient than that observed with 2'F-rCTP. Considering the potential advantages of 2'F-ANA and 2'F-RNA modifications in oligonucleotide-based gene silencing approaches, our study suggests that it should be possible to generate 2'F-ANA/2'F-RNA-based aptamers through an *in vitro* selection process such as SELEX. By examining the structure-activity relationship (SAR) of the ribose and arabinose sugars, these studies contributed to the understanding of how C2' stereochemistry and sugar pucker affect NTP selection (specificity) by DNA polymerases.

7.2. Publications and Presentations

As a direct result of the studies described herein, the following have been recently published, or are currently in preparation for submission in the near future:

1. Peng, C.G. and Damha, M.J. Assessing polymerase activity with 2'-deoxy-2'-fluoro-β-D-ribonucleoside 5'-triphosphates (2'F-rNTPs) and 2'-deoxy-2'-fluoro-β-D-arabinonucleoside 5'-triphosphates (2'F-araNTPs), manuscript in preparation for *Biochemistry*.

2. Peng, C.G., Mangos M. and Damha, M.J. Yeast debranching enzyme (yDBR) inhibition by novel branched RNAs, manuscript in preparation for *Bioorganic & Medicinal Chemistry Letters.*

3. Peng, C.G. and Damha, M.J. G-quadruplex induced stabilization by 2'-deoxy-2'-fluoro-β-D-arabinonucleic acids (2'F-ANA), manuscript submitted to *Nucleic Acids Research* (Manuscript NAR-00988-2007).

4. Peng, C.G. and Damha, M.J. (**2007**) Polymerase-directed synthesis of 2'-deoxy-2'-fluoro-β-D-arabinonucleic acids. *Journal of the American Chemical Society,* 129, 5310-5311

5. Peng, C.G. and Damha, M.J. (**2006**) DNA polymerase recognition of 2'-deoxy-2'-fluoroarabinonucleoside 5'-triphosphates (2'F-araNTPs). *Nucleosides, Nucleotides and Nucleic Acids* (Conference Proceedings; in press)

6. Peng, C.G. and Damha, M.J. (**2005**) Synthesis and hybridization studies of oligonucleotides containing 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribo-furanosyl)thymine (2'-α-*hm*-dT). *Nucleic Acids Research*, 33(22), 7019-7028.

The results disclosed in this thesis have led to the following provisional patents being filed:

1. Damha, M.J. and Peng, C.G. (**2006**) *2'-Deoxy-2'-Fluoro- β-D-Arabinonucleoside 5'-Triphosphates and Their Use in Enzymatic Nucleic Acid Synthesis.* U.S. Provisional Patent Application US 60/839,417.

2. Damha, M.J. and Peng, C.G. (**2005**) *Aptamers Comprising Arabinose Modified Units*. U.S. Provisional Patent Application US 60/722,973.

Conference and presentations

1. Peng, C.G. and Damha, M.J. (**2006**) Probing DNA polymerase activity with 2'F-araNTPs and 2'F-riboNTPs. The XVII International Roundtable on Nucleosides, Nucleotides and Nucleic Acids (IRTXVII), Bern, Switzerland, September 3-7.

2. Peng, C.G. and Damha, M.J. (**2006**) Synthesis, characterization and applications of novel oligonucleotide structures containing modified 2',5'-RNA. 231st ACS National Meeting, Atlanta, GA, USA, March 26-30. (oral presentation)

3. Peng, C.G. (**2006**) Synthesis, characterization and applications of modified 2',5'-RNA and FANA modified aptamers. Organic Seminar, Chemistry Department, McGill University, Montreal, Jan. 13. (oral presentation)

4. Peng, C.G. and Damha, M.J. (**2005**) Synthesis of 2'-deoxy-2'-α-*C*-hydroxymethyl 2',5'-linked RNA. 16th Quebec-Ontario Minisymposium in Synthetic and Bio-Organic Chemistry, Sainte-Adele, Canada, Nov. 11-13.

5. Peng, C.G. and Damha, M.J. (**2005**) Synthesis of a "taller" 2',5'-linked ribonucleic acid: 2'-α-*C*-hydroxymethyl 2',5'-linked RNA. 14th Conversation in Nucleic Acid Structure/Biology at SUNY, Albany, USA, June 14-18.

6. Peng, C.G. and Damha, M.J. (**2004**) Synthesis and binding properties of 2'-α-*C*-hydroxymethyl-2',5'-linked RNA. 87th CSC Annual meeting, London, Canada, May 29-June 1. (oral presentation)

Appendix I

Reprint of Published Manuscript

Published online December 23, 2005

Nucleic Acids Research, 2005, Vol. 33, No. 22 ***7019–7028***
doi:10.1093/nar/gki1003

Synthesis and hybridization studies of oligonucleotides containing 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (2′-α-hm-dT)

Chang Geng Peng and Masad J. Damha*

Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, QC, Canada H3A 2K6

Received September 8, 2005; Revised and Accepted November 17, 2005

ABSTRACT

We report the first investigation of oligoribonucleotides containing a few 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine units (or 2′-hm-dT, abbreviated in this work as 'H'). Both the 2′-*CH_2O*-phosphoramidite and 3′-*O*-phosphoramidite derivatives of H were synthesized and incorporated into both 2′,5′-RNA and RNA chains. The hybridization properties of the modified oligonucleotides have been studied via thermal denaturation and circular dichroism studies. While 3′,5′-linked H was shown previously to significantly destabilize *DNA*:RNA hybrids and *DNA*:DNA duplexes (modification in the *DNA* strand; $\Delta T_m \sim -3$°C/insert), we find that 2′,5′-linked H have a smaller effect on *2′,5′-RNA*:RNA and *RNA*:RNA duplexes ($\Delta T_m = -0.3$°C and -1.2°C, respectively). The incorporation of 3′,5′-linked H into *2′,5′-RNA*: RNA and *RNA*:RNA duplexes was found to be more destabilizing (-0.7°C and -3.6°C, respectively). Significantly, however, the 2′,5′-linked H units confer marked stability to RNA hairpins when they are incorporated into a 2′,5′-linked tetraloop structure ($\Delta T_m = +1.5$°C/insert). These results are rationalized in terms of the compact and extended conformations of nucleotides.

INTRODUCTION

Our group has had a long-standing research interest in the physicochemical and biochemical properties of 2′,5′-linked ribonucleic acids (2′,5′-RNA, Figure 1) (1,2). These regioisomers of standard (i.e. 3′,5′-linked) RNA are not only interesting from a structural point of view, but also have potential use in the down-regulation of gene expression (3,4). For instance, 2′,5′-RNAs are able to associate with complementary single-stranded RNA (ssRNA) (3) as well as duplex DNA (5) and, as such, can potentially be used to down-regulate gene expression via the antisense and antigene approaches (Figure 1) (4,6–8). It is also well-documented that annealing two normal RNA strands is more favorable than annealing a 2′,5′-RNA strand with a normal RNA strand (1,3,8,9). Furthermore, mutually complementary 2′,5′-RNA strands have the ability to associate, but exhibit a transition temperature (T_m) that is considerably lower than those of the corresponding RNA:RNA or RNA:2′,5′-RNA duplexes. A comparison of the T_m values of various duplexes of mixed base composition revealed the following order of duplex thermal stability: RNA:RNA > DNA:DNA ≈ DNA:RNA > RNA:2′,5′-RNA > 2′,5′-RNA: 2′,5′-RNA > DNA: 2′,5′-RNA (undetected) (1).

Molecular modeling and circular dichroism (CD) studies of these duplexes revealed that RNA: 2′,5′-RNA hybrids adopt a continuous A-type helix structure similar to that of native RNA (1), but have smaller interstrand phosphate–phosphate distances (by ~1 Å). This may account, at least in part, for the lower thermal stability of RNA:2′,5′-RNA relative to RNA: RNA and RNA:DNA helices (1).

As part of our ongoing study of these systems, we now present an investigation of 2′,5′-RNA chains containing one or more 1-(2-deoxy-2-α-*C*-hydroxymethyl-β-D-ribofuranosyl)thymine (2′-α-hm-dT; abbreviated in this work as '**H**', Figure 1 and Scheme 1). We anticipated that lengthening the sugar-phosphate backbone by one methylene unit would not only diminish putative P–P repulsions in 2′,5′-RNA:RNA duplexes, but would also provide a 'compact' nucleotide conformation favoring tighter RNA binding (Figure 2) (10,11). This interesting modification was first studied by Schmit *et al.* (12) and reviewed more recently by Freier and Altmann (13). Schmit *et al.* (12) reported that 3′,5′-linked **H** units (3′,5′-**H**) significantly destabilize *DNA*:RNA hybrids and *DNA*:DNA duplexes (modification in the *DNA* strand; $\Delta T_m = -2.9$°C). Destabilization owing to the 2′-α-hydroxymethyl and other

*To whom correspondence should be addressed. Tel: +1 514 398 7552; Fax: +1 514 398 3797; Email: masad.damha@mcgill.ca

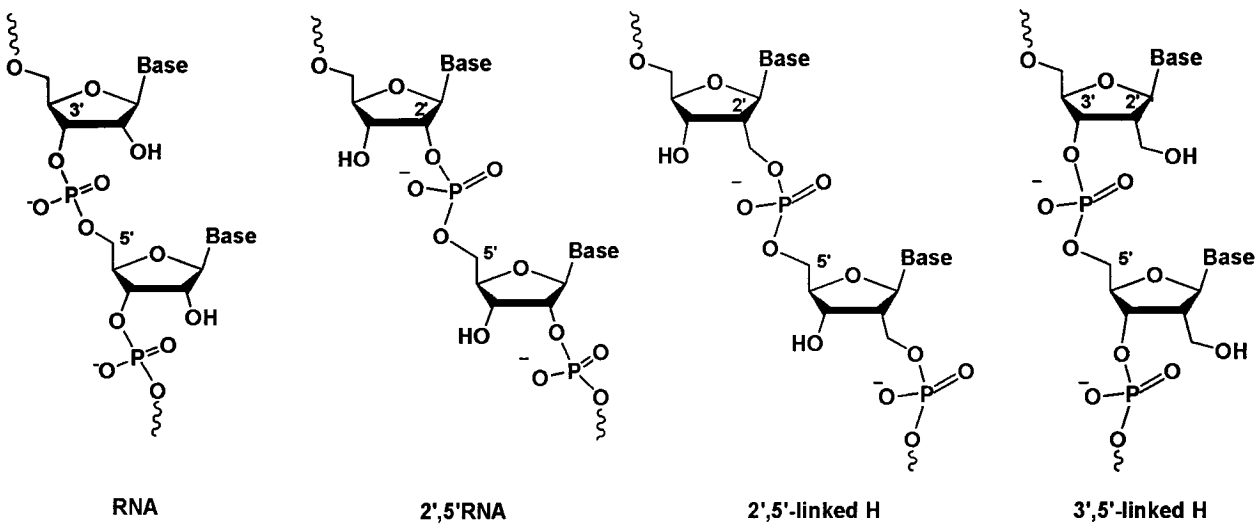
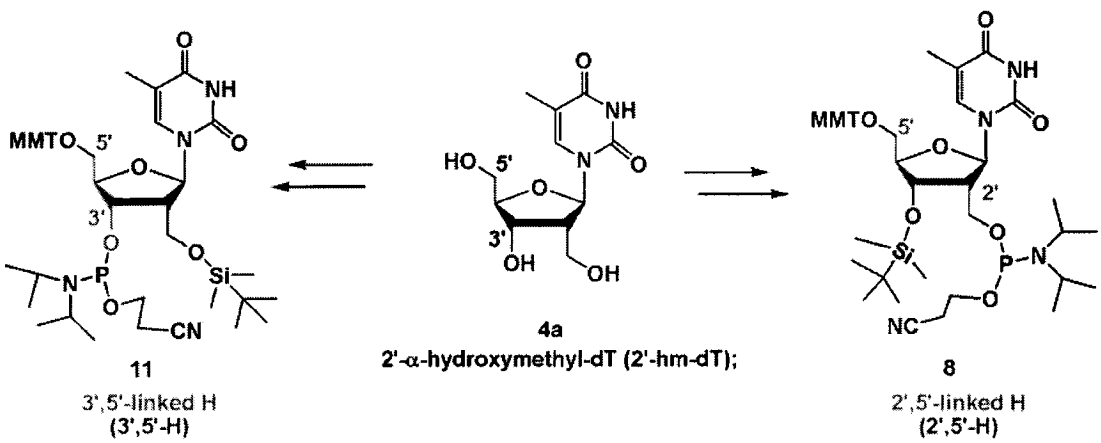


Figure 1. Structure of RNA, 2′,5′-RNA, and strands comprising 2′,5′-linked and 3′,5′-linked **H** units.



Scheme 1. Structures of 2′-α-hm-dT (**H**) and its phosphoramidite derivatives **8** and **11**.

2′-α alkyl groups was explained by the tendency of these substituents to shift the conformational equilibrium of the sugar toward the C2′-*endo* pucker and away from the C3′-*endo* pucker found in RNA, 2′-*O*-alkyl RNA and 2′F-RNA duplexes (12,13). To the best of our knowledge, neither 2′,5′-linked **H** units (2′,5′-**H**) nor incorporation into RNA have ever been examined. The present study revealed that 2′,5′-**H** units have a modest destabilizing effect on *2′,5′-RNA*:RNA ($\Delta T_m = -0.3°C$) and *RNA*:RNA duplexes ($\Delta T_m = -1.2°C$), whereas 3′,5′-**H** units are significantly more destabilizing (−0.7°C and −3.6°C, respectively), consistent with the results of Schmit *et al.* (12). In contrast, we find that 2′,5′-**H** units confer significant stability to RNA hairpins, particularly when they are placed in the loop structure ($\Delta T_m = +1.5°C$).

MATERIALS AND METHODS

General reagents

All reactions were carried out in oven-dried glassware under a N_2 atmosphere. Dichloromethane (CH_2Cl_2) and acetonitrile (CH_3CN) were dried by refluxing and distilling over calcium hydride (CaH_2) under a N_2 atmosphere. Tetrahydrofuran (THF) was dried by refluxing over sodium and benzophenone under a N_2 atmosphere and collected before use. Anhydrous methanol (MeOH), pyridine (py), *N*-ethyl-*N*,*N*-diisopropylamine (DIPEA), *N*,*N*-dimethylformamide (DMF) and 2,6-lutidine were obtained from Aldrich. The following chemicals were used as received from Aldrich: *p*-anisylchlorodipenylmethane (MMTr-Cl), benzoyl chloride (Bz-Cl), *tert*-butyldimethylsilyltriflate (TBDMSOTf), *tert*-butyldimethylsilylchloride (TBDMSCl), 4,5-dicyanoimidazole (DCI), 4-dimethylaminopyridine (DMAP), 1,1,1,3,3,3-hexamethyldisilazane (HMDS), 10 *wt%* palladium on carbon powder (10% Pd/C), triethylamine tris(hydrofluoride) (TREAT HF), *D*-ribose, silver nitrate ($AgNO_3$), 1.0 M tin (IV) tetrachloride solution in dichloromethane ($SnCl_4$), thymine. β-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite and other solid-phase synthesis reagents and regular nucleoside monomers were purchased from ChemGenes Corp. (Ashland, MA).

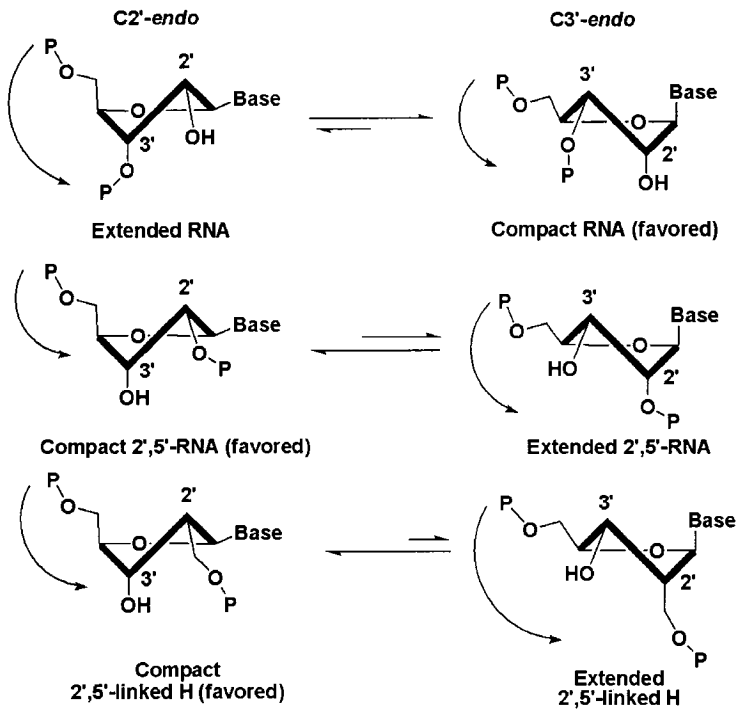


Figure 2. The C2′- and C3′-*endo* sugar puckers are favored in 2′,5′-linked RNA and 3′,5′-linked RNA, respectively (10,11). The intraresidue P–P distance determines the 'compact' or 'extended' backbone structure (11). Lengthening by an extra carbon C2′ (**H**) removes the gauche effect between the ring oxygen and O2′, which is expected to reinforce the C2′-*endo* or compact conformation for 2′,5′-linked **H** units.

Methyl 2-α-benzoyloxymethyl-2-deoxy-3,5-di-*O*-(2,4-dichlorobenzyl)-α-D-ribofuranoside (2)

Compound (**1**) was obtained following slight modifications of the procedure described by Martin [(14); Supplementary Data]. Under a N_2 atmosphere, 1.5 ml (1.78 g, 12.7 mmol) BzCl was added to a solution of **1** (4.83 g, 9.73 mmol) in 75 ml dry pyridine. The resulting solution was kept stirring at room temperature for 4 h. The reaction progress was monitored by TLC. When the reaction was complete, the reaction mixture was concentrated under reduced pressure, and then washed with saturated $NaHCO_3$ and brine. The aqueous phase was extracted with CH_2Cl_2 three times. The combined organic layers were evaporated to dryness to afford **2** in quantitative yield (5.84 g). R_f (SiO_2) = 0.86 (2:1 EtOAc/hexane); 0.70 (1:2 EtOAc/hexane); ^{1}H NMR (400 MHz, $CDCl_3$, δ): 7.98–7.09 (m, 11H, ArH), 5.11 (d, $^3J_{1\text{-}2}$ = 5.2 Hz, 1H, H1), 4.73–4.55 (m, 6H, benzyl CH_2 and CH_2–C2), 4.4–4.43 (m, 1H, H4), 4.08 (dd, $^3J_{3\text{-}2}$, $^3J_{3\text{-}4}$ = 5.2, 2.4 Hz, 1H, H3), 3.54–3.64 (AB dd, $^2J_{5,5'}$ = 22.4 Hz, $^3J_{4,5+5'}$ = 4.4, 5.2 Hz, 2H, H5 and H5′), 3.45 (s, 3H, OCH_3), 2.74–2.67 (m, 1H, H2). ESI-MS for $C_{28}H_{26}Cl_4O_6$ [MNa^+] calcd 621.05, found 621.1.

1-[2-α-Benzoyloxymethyl-2-deoxy-3,5-di-*O*-(2,4-dichlorobenzyl)-D-ribofuranosyl]thymine (3)

Under a N_2 atmosphere, a mixture of dry thymine (1.35 g, 10.7 mmol), ammonium sulfate (0.136 g, 1.029 mmol) and HMDS (2.87 ml, 13.8 mmol) in anhydrous CH_3CN (41 ml) was refluxed at 100°C for 4 h until the reaction mixture was clear. The reaction mixture was cooled to room temperature and added to a solution of **2** (2.47 g, 4.12 mmol) in dry CH_3CN (19.5 ml) followed by 1.0 M $SnCl_4$ in CH_2Cl_2 (4.12 ml, 4.12 mmol). The resulting solution was warmed to 50°C and left stirring for an additional 12 h. After the reaction was complete, the reaction mixture was cooled to room temperature, washed with saturated $NaHCO_3$ (50 ml) and filtered. The filtrate was extracted three times with the same volume of CH_2Cl_2 and the organic layer was washed with brine, dried over Na_2SO_4 and evaporated to yield a white foam. Purification by silica gel column chromatography with EtOAc/hexane (2:1, v/v) afforded the title compound in 72% yield (2.06 g, β:α = 5:1). R_f (SiO_2) = 0.6 (2:1 EtOAc/hexane); ^{1}H NMR of β-isomer (400 MHz, $CDCl_3$, δ): 7.97–7.19 (m, 12H, ArH and H-6), 6.35 (d, $^3J_{1',2'}$ = 8.8 Hz, 1H, H1′), 4.74, 4.47 (2dd, $^2J_{CH2\text{-}C2'}$ = 12 Hz, $^3J_{CH2\text{-}C2',\,H_{2'}}$ = 7.6, 7.2 Hz, 2H, CH_2-C2′), 4.69–4.56 (m, 4H, benzyl CH_2), 4.38 (br, m, 1H, H4′), 4.31–4.28 (m, 1H, H3′), 3.9, 3.7 (2 dd, $^2J_{5',5''}$ = 10.8 Hz, $^3J_{4',5'+5''}$ = 3.2, 2.4 Hz, 2H, H5′ and H5″), 2.99–2.92 (m, 1H, H2′), 1.64 (d, $^4J_{CH3\text{-}C5,H6}$ = 1.2 Hz, 3H, CH_3-C5). ESI-MS for $C_{32}H_{28}Cl_4N_2O_7$ [MNa^+] calcd 717.39, found 717.0.

1-(2-α-*C*-Benzoyloxymethyl-2-deoxy-β-D-ribofuranosyl)thymine (4)

To a solution of **3** (1.065 g, 1.53 mmol) in dry MeOH (50 ml) was added 10% Pd/C (0.39 g, 0.37 mmol). The resulting mixture was kept shaking under 40–50 psi H_2 at room temperature for 6 h. The mixture was filtered and evaporated to dryness, and the residue was purified by silica gel column chromatography (eluent 40:1 to 15:1 CH_2Cl_2/MeOH) to separate the α and β-isomers. The desired β-isomer was obtained in 71% yield (0.41 g). R_f (SiO_2) = 0.51 (6:1 CH_2Cl_2/MeOH); 0.31

(10:1 CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, DMSO-d_6, δ): 11.22 (s, 1H, N-H), 7.78–7.41 (m, 5H, Bz), 7.7 (s, 1H, H-6), 6.14 (d, $^3J_{1',2'}$ = 8.8 Hz, 1H, H1′), 5.52 (d, $^3J_{OH,H3'}$ = 4.8 Hz, 1H, HO-C3′), 5.1 (t, $^3J_{OH,H5'+H5''}$ = 5.2 Hz, 1H, HO-C5′), 4.57, 4.26 (2dd, $^2J_{CH2\text{-}C2}$ = 11.6 Hz, $^3J_{CH2\text{-}C2,H2'}$ = 5.6, 8.8 Hz, 2H, CH_2-C2′), 3.88 (br, 1H, H4′), 3.54–3.64 (AB, dd, $^2J_{5',5''}$ = 17.7, $^3J_{4',5'+5''}$ = 5.2, 4.8 Hz, 2H, H5′ and H5″), 2.72–2.29 (m, 1H, H2′), 1.74 (s, 3H, CH_3-C5). Four NOESY cross peaks identified the β-isomer: H1′-H4′, H2′-H6′, 2′-CH_2-H4′, H5′-H6. ESI-MS for $C_{18}H_{20}N_2O_7$ [MNa^+] calcd 399.13, found 399.1. Isolated α-isomer, ^{1}H NMR (500 MHz, acetone-d_6, δ): 10.11 (s, 1H, N-H), 8.02–7.46 (m, 5H, Bz), 7.9 (s, 1H, H-6), 6.55 (d, $^3J_{1',2'}$ = 7.5 Hz, 1H, H1′), 5.13 (d, $^3J_{OH,H3'}$ = 3.5 Hz, 1H, HO-C3′), 4.66 (m, 1H, H3′), 4.52–4.3 (m, 2H, CH_2-C2′), 4.48 (m, 1H, H4′), 4.19 (br, 1H, HO-C5′), 3.66 (br, 2H, H5′ and H5″), 3.33 (m, 1H, H2′), 1.82 (s, 3H, CH_3-C5). ESI-MS for α $C_{18}H_{20}N_2O_7$ [MNa^+] calcd 399.13, found 399.1.

1-[2-α-*C*-benzoyloxymethyl-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (5)

Compound **4** (0.555 g, 1.47 mmol) and $AgNO_3$ (0.375 g, 2.21 mmol) were dissolved in dry pyridine (10 ml). The flask was purged with N_2 and a solution of MMTrCl (0.683 g, 2.21 mmol) in dry pyridine (5 ml) was added. After stirring for 6 h at room temperature, the reaction was worked up by evaporating most of the pyridine and taking up the residue in CH_2Cl_2 (20 ml). The solution was washed with saturated $NaHCO_3$ (20 ml), dried, filtered, and finally evaporated to yield the crude product. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (49:1 to 15:1, v/v) afforded the title compound as a white solid (687 mg; 72% yield). R_f (SiO_2) = 0.66 (10:1 CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, pyridine-d_5, δ): 12.12 (s, 1H, N-H), 6.95–5.75 (m, 19H, ArH), 6.73 (d, $^4J_{CH3\text{-}C5,H6}$ = 1.2 Hz, 1H, H6), 5.94 (d, $^3J_{1',2'}$= 12 Hz, 1H, H1′), 4.12–3.83 (m, 3H, CH_2-C2′ and H3′), 3.43 (br, m, 1H, H4′), 2.48 (s, 3H, OCH_3), 2.45 (m, 2H, H5′ and H5″), 2.3 (m, 1H, H2′), 0.5 (s, 3H, CH_3-C5). ^{1}H NMR (400MHz, DMSO-d_6, δ): 11.26 (s, 1H, N-H), 7.81–6.88 (m, 19H, ArH), 7.44 (s, 1H, H6), 6.2 (d, $^3J_{1',2'}$ = 8.8 Hz, 1H, H1′), 5.62 (d, $^3J_{3',OH}$ = 5.2 Hz, 1H, OH), 4.7, 4.31 (2dd, $^2J_{CH2\text{-}C2'}$ = 11.2 Hz, $^3J_{CH2\text{-}C2',H2'}$ = 5.2, 8.4 Hz, 2H, CH_2-C2′), 4.43 (m, 1H, H3′), 4.0 (br m, 1H, H4′), 3.73 (s, 3H, OCH_3), 3.27, 3.16 (2dd, $^2J_{5',5''}$ = 10.4 Hz, $^3J_{4',5'+5''}$ = 6.4, 3.2 Hz, 2H, H5′ and H5″), 3.01–2.94 (m, 1H, H2′), 1.3 (s, 3H, CH_3-C5). ESI-MS for $C_{38}H_{36}N_2O_8$ [MNa^+] calcd 671.25, found 671.2.

1-[2-α-*C*-benzoyloxymethyl-3-*O*-*tert*-butyldimethylsilyl-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (6)

To a solution of **5** (205 mg, 0.316 mmol) in dry CH_2Cl_2 was added 2,6-lutidine (0.15 ml, 1.26 mmol) under a N_2 atmosphere. *tert*-Butyldimethysilyl triflate (TBDMOTf; 0.218 ml, 0.948 mmol) was then added dropwise with stirring at room temperature to yield a clear yellow solution. After stirring at 40°C for 12 h, saturated $NaHCO_3$ (12 ml) was added. The aqueous phase was separated and washed several times with CH_2Cl_2. The combined CH_2Cl_2 layers were dried over Na_2SO_4 and evaporated to dryness. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2:MeOH (100:1 to 20:1, v/v) afforded **6** in 72% yield (0.173 g). When the silylation reaction was conducted at room temperature for 4 h, the yield of the title compound increased to 83%. R_f (SiO_2) = 0.64 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (500 MHz, DMSO-d_6, δ): 11.33 (s, 1H, N-H), 7.83–6.88 (19H, ArH), 7.5 (s, 1H, H6), 6.12 (d, $^3J_{1',2'}$ = 8.5 Hz, 1H, H1′), 4.57, 4.3 (2m, 2H, CH_2-C2′), 4.55 (m, 1H, H3′), 3.97 (m, 1H, H4′), 3.71 (s, 3H, OCH_3), 3.24 (m, 2H, H5′ and H5″), 3.0 (m, 1H, H2′), 1.42 (s, 3H, CH_3-C5), 0.8 (s, 9H, *t*-Bu-Si), 0.02, −0.06 (2s, 6H, $(CH_3)_2$-Si). ^{1}H NMR (400 MHz, acetone-d_6, δ): 9.94 (s, 1H, N-H), 7.96–6.91 (17H, ArH), 7.61 (s, 1H, H6), 6.35 (d, $^3J_{1',2'}$ = 8.8, 1H, H1′), 4.78–4.42 (m, 3H, H3′ and CH_2-C2′), 4.15 (br, m, 1H, H4′), 3.79 (s, 3H, OCH_3), 3.44 (m, 2H, H5′ and H5″), 3.18 (m, 1H, H2′), 1.44 (s, 3H, CH_3-C5), 0.92 (s, 9H, *t*-Bu-Si), 0.14, −0.08 (2s, 6H, $(CH_3)_2$-Si). ESI-MS for $C_{44}H_{50}N_2O_8Si$ [MNa^+] calcd 785.33, found 758.3.

1-[3-*O*-*tert*-butyldimethylsilyl-2-deoxy-2-α-*C*-hydroxymethyl-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (7)

A saturated solution of NaOMe in MeOH (4 ml) was added to **6** (0.337 g, 0.442 mmol) at room temperature. After stirring for 2–3 h, the reaction mixture was washed with saturated $NaHCO_3$ (10 ml). The aqueous phase was separated and washed several times with CH_2Cl_2. The combined CH_2Cl_2 layers were dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (100:1 to 15:1, v/v) afforded **7** in quantitative yield (290 mg). R_f (SiO_2) = 0.43 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, DMSO-d_6): δ11.30 (s, 1H, N-H), 7.47 (s, 1H, H6), 7.40–6.89 (m, 14H, ArH), 5.93 (d, $^3J_{1',2'}$ = 7.6 Hz, 1H, H1′), 4.57 (d, $^3J_{OH,H3'}$= 4.8 Hz, 1H, OH); 4.45 (br m, 1H, H3′), 3.90 (br, 1H, H4′), 3.73 (s, 3H, OCH_3), 3.68, 3.45 (2m, 2H, CH_2-C2′), 3.19 (m, 2H, H5′ and H5″); 2.53 (m, 1H, H2′), 1.51 (s, 3H, CH_3-C5), 0.82 (s, 9H, *t*-Bu-Si), 0.05, −0.04 (2s, 6H, $(CH_3)_2$-Si). ESI-MS for $C_{37}H_{46}N_2O_7Si$ [MNa^+] calcd. 681.31, found 681.2.

1-[2-α-*C*-((β-cyanoethyl-*N*,*N*-diisopropylphosphoramidic)-hydroxymethyl)-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (8)

N-ethyl-*N*,*N*-diisopropylamine (0.276 ml, 0.205 g, 1.58 mmol) was added to a solution of **7** (0.29 g, 0.44 mmol) in dry THF (3 ml) under a nitrogen atmosphere. The reaction was initiated by addition of β-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite via syringe (0.125 g, 0.118 ml, 0.528 mmol). After 2 h, the reaction mixture was passed through a 2 cm layer of silica gel (pre-neutralized with 1% NEt_3), and the desired compound eluted by washing with ice cold hexane followed by CH_2Cl_2. Evaporation of the solution afforded the title compound as a nice white foam (353 mg, 97% yield). R_f (SiO_2) = 0.63 (2:1 EtOAc/hexane); ^{31}P NMR (200 MHz, $CDCl_3$, δ): 148.90, 148.24. ESI-MS for $C_{46}H_{63}N_4O_8PSi$ [MNa^+] calcd 881.42, found 881.2.

1-[2-Deoxy-2-α-*C*-hydroxymethyl-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (9)

A saturated solution of NaOMe in MeOH (2 ml) was added to **5** (0.208 g, 0.321 mmol) and the resulting mixture stirred at room temperature for 2 h. The reaction mixture was then washed with brine (5 ml) followed by CH_2Cl_2 (20 ml). The combined CH_2Cl_2 layers were dried over Na_2SO_4, filtered and evaporated to dryness. Purification by silica gel flash column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (15:1, v/v) afforded **9** in quantitative yield (175 mg). R_f (SiO_2) = 0.26 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (400 MHz, acetone-d_6, δ): 7.6 (d, $^4J_{CH3\text{-}C5,H6}$ = 1.2 Hz, 1H, H6), 7.52–6.91 (m, 14H, ArH), 6.21 (d, $^3J_{1',2'}$ = 8.4 Hz, 1H, H1′), 4.65 (br, m, 1H, H3′), 4.1–4.12 (m, 1H, H4′), 3.98, 3.79 (2dd, $^2J_{CH2\text{-}C2'}$ = 11.2 Hz, $^3J_{CH2\text{-}C2',\ H2'}$ = 10.0, 6.8 Hz, 2H, CH_2-C2′), 3.42–3.35 (AB, dd, $^2J_{5',5''}$ = 9.6 Hz, $^3J_{4',5'+5''}$ = 4.0, 3.2 Hz, 2H, H5′ and H5″), 2.64–2.71 (m, 1H, H2′), 1.49 (s, 3H, CH_3-C5). ^{1}H NMR (500 MHz, DMSO-d_6, δ): 11.3 (s, 1H, N-H); 7.47 (s, 1H, H6); 7.39–6.89 (m, 14H, ArH); 5.97 (d, $^3J_{1',2'}$ = 8.5 Hz, 1H, H1′), 5.3 (d, $^3J_{OH,\ H3'}$ = 5.5 Hz, 1H, 3-OH); 4.54 (bs, 1H, C2′-C-OH); 4.3 (bs, 1H, H3′); 3.91 (m, 1H, H4′); 3.73 (s, 3H, OCH_3); 3.74–3.43 (2m, 2H, CH_2-C2′); 3.2, 3.12 (2dd, $^2J_{5',5''}$ = 10.5 Hz, $^3J_{4',5'+5''}$ = 4.5, 3.0 Hz, 2H, H5′ and H5″), 2.45 (m, 1H, H2′), 1.4 (s, 3H, CH_3-C5); ESI-MS for $C_{31}H_{32}N_2O_7$ [MNa^+] calcd 567.22, found 567.2.

1-[2-α-*C*-*tert*-butyldimethylsilyloxymethyl-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (10)

To a solution of **9** (0.186 g, 0.342 mmol) in dry pyridine (3 ml) under a N_2 atmosphere was added $AgNO_3$ (0.067 g, 0.393 mmol) followed by TBDMSCl (59 mg, 0.393 mmol). After stirring for 12 h at room temperature, the reaction was concentrated and the residue taken up in CH_2Cl_2. The solution was washed with saturated $NaHCO_3$ (10 ml), dried, filtered, and finally evaporated to yield the crude product. Purification by silica gel column chromatography with 0.5% NEt_3 in CH_2Cl_2/MeOH (100:1 to 15:1, v/v) afforded the title compound in 66% yield (148 mg). R_f (SiO_2) = 0.64 (15:1 CH_2Cl_2/MeOH); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 11.25 (s, 1H, N-H); 7.47 (s, 1H, H6); 7.39–6.87 (m, 14H, ArH); 6.08 (d, $^3J_{1',2'}$ = 8.7 Hz, 1H, H1′); 5.32 (d, $^3J_{OH,\ H\text{-}3'}$ = 4.8 Hz, 1H, OH); 4.31(br m, 1H, H3′); 3.98–3.64 (m, 2H, CH_2-C2′); 3.73 (s, 3H, OCH_3); 3.23, 3.13 (2dd, $^3J_{4',5'+5''}$ = 4.2, 3 Hz, 2H, H5′ and H5″); 2.57 (m, 1H, H2′); 1.37 (s, 3H, CH_3-C5); 0.79 (s, 9H, *t*-Bu-Si); 0.01- -0.04 (t, 6H, $(CH_3)_2$Si); ESI-MS for $C_{37}H_{46}N_2O_7Si$ [MNa^+] calcd 681.31, found 681.2.

1-[2-α-*C*-*tert*-butyldimethylsilyloxymethyl-3-*O*-(β-cyanoethyl-*N*,*N*-diisopropylphosphoramidic)-2-deoxy-5-*O*-(4-methoxytrityl)-β-D-ribofuranosyl]thymine (11)

β-Cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (32 µl, 0.034 g, 0.145 mmol) was slowly added to a solution of **10** (0.148 g, 0.121 mmol) and *N*-ethyl-*N*,*N*-diisopropylamine (76 µl, 0.056 g, 0.436 mmol) in dry THF (1.2 ml) under a N_2 atmosphere. After stirring for 2.5 h, the crude mixture was passed through a 2 cm layer of silica gel (pre-neutralized with 1% NEt_3), and the desired compound eluted by washing with ice cold hexane followed by CH_2Cl_2. Evaporation of the solution afforded the title compound as a nice white foam (101 mg, 52% yield). R_f (SiO_2) = 0.46 (2:1 EtOAc/hexane); ^{31}P NMR (200 MHz, $CDCl_3$): δ 151.7, 150.3; ESI-MS for $C_{46}H_{63}N_4O_8PSi$ [MNa^+] calcd 881.42, found 881.2.

Solid-phase synthesis of oligonucleotides

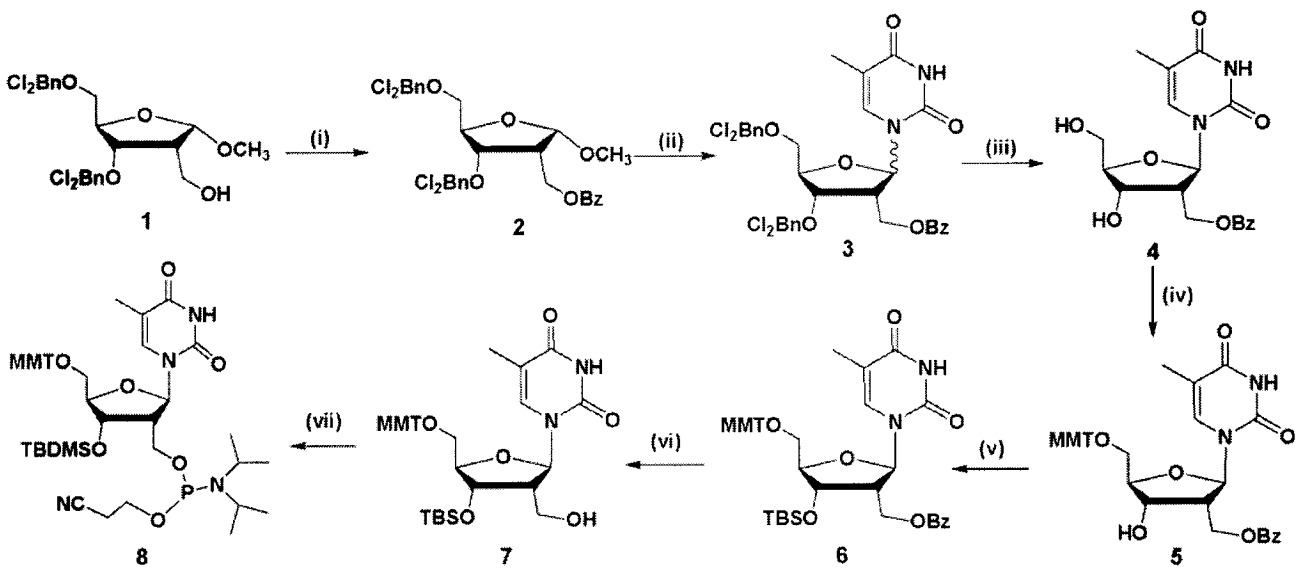
Oligonucleotide syntheses were carried out on a 1 µmol scale using an Applied Biosystems DNA/RNA 381A synthesizer as described previously (1,3,15). Solutions of phosphoramidites **8** and **11** in acetonitrile (0.09 M) were allowed to react with the solid support for an extended coupling time of 30 min using DCI in acetonitrile (0.5M) as catalyst. These conditions afforded >99% coupling efficiency. Deprotection was conducted by addition of (i) conc. aq. ammonia/ethanol (3:1, v/v, 1 ml, 48 h, room temperature) followed by evaporation; (ii) TEA•3HF (100–200 µl, 48 h, room temperature) followed by evaporation. Typically, 40–80 OD units (A_{260}) of the oligonucleotides were obtained at this point. Oligonucleotides were purified by anion-exchange high-performance liquid chromatography (HPLC) (Protein Pak DEAE-5PW column-Waters; 7.5 mm × 7.5 cm), desalted by size-exclusion chromatography on Sephadex G-25 matrix, and characterized by MALDI-TOF mass spectrometry (Kratos Kompact-III instrument; Kratos Analytical Inc., NY). Purity of the isolated oligonucleotides was >95%.

UV thermal denaturation studies

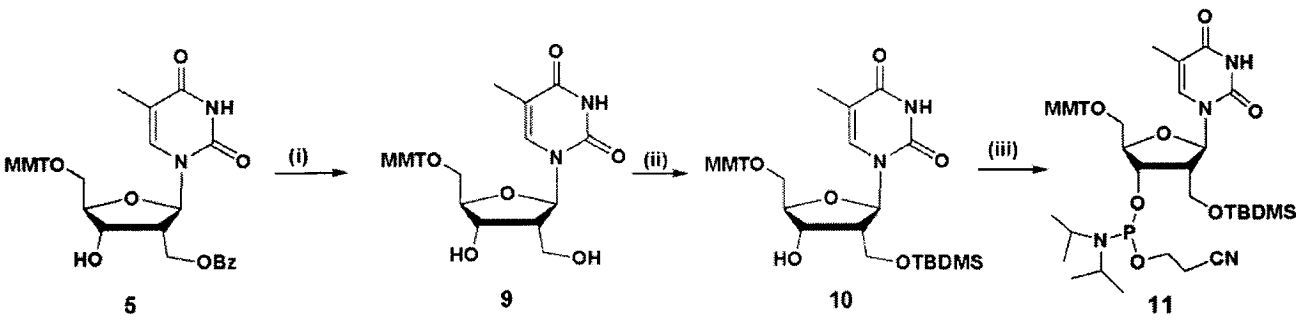
UV thermal denaturation data were obtained on a Varian CARY 1 spectrophotometer equipped with a Peltier temperature controller (Varian, Mulgrave, Australia). Oligomers and complementary targets were mixed in equimolar ratios in 140 mM K^+, 1 mM Mg^{2+} and 5 mM Na_2HPO_4 buffer, pH 7.2, which is representative of intracellular conditions (16). The total strand concentration was 2.6 µM. Samples were heated to 90°C for 5 min, cooled slowly to room temperature, and refrigerated (4°C) overnight before measurements. Prior to the thermal run, samples were degassed by placing them in an ultrasound bath for 1 min. Denaturation curves were acquired at 260 nm at a rate of heating of 0.5°C/min. The data were analyzed with the software provided by Varian Canada and converted to Microsoft Excel. T_m values were calculated as the maximum of the first-derivative plots of absorbance versus temperature and have an uncertainty of ±1°C. Hyperchromicity values (i.e. changes in relative absorbance) were calculated using the formula: $H = (A_t - A_i)/A_h$, where H is the hyperchromicity, A_t is the absorbance at any given temperature (t), A_i is the initial absorbance reading, and A_h is the absorbance at the highest temperature.

Circular dichroism spectra

CD spectra (200–350 nm) were collected on a Jasco J-710 spectropolarimeter at a rate of 100 nm/min using fused quartz cells (Hellma, 165-QS). Measurements were carried out in 140 mM K^+, 1 mM Mg^{2+} and 5 mM Na_2HPO_4 buffer, pH 7.2 (15) at a duplex concentration of 2.6 µM. The temperature was controlled by an external circulating bath (VWR Scientific) at constant temperature (5°C). The data were processed on a PC computer using J-700 Windows software supplied by the manufacturer (JASCO, Inc.). To facilitate comparisons, the CD spectra were background subtracted, smoothed and corrected for concentration so that molar ellipticities could be obtained.



Scheme 2. Synthesis of **8**. (i) BzCl, py., room temperature, 4 h, 100%; (ii) *bis*(trimethylsilyl) thymine, $SnCl_4$, CH_3CN, 50°C, overnight, 72% (isomer mixture, β/α 5:1); (iii) 10% Pd/C, H_2 (40–50 psi), MeOH, room temperature, 6 h, 71% β pure isomer; (iv) MMTr-Cl (1.5 equivalent), $AgNO_3$, py, room temperature, overnight, 72%; (v) TBDMSOTf, 2,6-lutidine, THF, room temperature, 4 h, 83%; (vi) NaOMe, MeOH, room temperature, 2–3 h, 100%; (vii) iPr$_2$NEt, (iPr$_2$N)(OCH$_2$CH$_2$CN)PCl, THF, room temperature, 2 h, 97%.



Scheme 3. Synthesis of **11**. (i) NaOMe, MeOH, room temperature, 2 h, 100%; (ii) TBDMSCl, $AgNO_3$, py, room temperature, overnight, 66%; (iii) iPr$_2$NEt, (iPr$_2$N)(OCH$_2$CH$_2$CN)PCl, THF, room temperature, 2.5 h, 52%.

RESULTS AND DISCUSSION

Monomer synthesis

Schemes 2 and 3 show the synthesis of 2′-α-hm-dT (abbreviated in this work as '**H**') and its conversion to the 2′-CH$_2$O- (**8**) and 3′-*O*-phosphoramidites (**11**), using a synthetic strategy analogous to that described previously by Schmit (17) and Li and Piccirilli (18). We protected the C2′-CH$_2$-OH moiety of *O*-glycoside **1** (14) as the benzoyl ester rather than as Schmit's acetyl ester, since we found that the latter was not stable under the conditions of the subsequent hydrogenolysis reaction (Scheme 2). Coupling of **2** with *bis*(trimethylsilyl)thymine in the presence of $SnCl_4$ catalyst gave anomeric nucleosides **3** in good yield with a β/α stereoselectivity of 5:1. It is interesting to note that the β/α stereoselectivity was found to be higher for the C2-CH$_2$-OAc derivative (10:1), in agreement with Schmit's findings (12,17). Separation of the anomers by column chromatography was achieved after removal of the 2,4-dichlorobenzyl protecting groups, to provide the desired β-anomer in 71% yield from **3**. The anomeric configuration of **4** was unequivocally established using NOESY NMR, which displayed strong H1′-H4′, H2′-H6′, 2′-CH$_2$-H4′, and H5′-H6 cross peaks. Reprotection of hydroxyl groups with MMTrCl followed by TBDMSOTf gave **6** in good yields (19). Attempts to carry out the silylation reaction with TBDMSCl (DMF/imidazole or $AgNO_3$/py) proved to be problematic (20). Lastly, we converted **6** to the 2′-CH$_2$O-phosphoramidite derivative **8** via consecutive debenzoylation and phosphitylation reactions (Scheme 2). To access the 3′-*O*-phosphoramidite derivative, nucleoside **5** was treated with NaOMe to give diol **9** quantitatively. Consecutive silylation with TBDMSCl ($AgNO_3$/py) and phosphitylation reactions then afforded the desired compound **11** in moderate yield (52%) (Scheme 3).

Oligonucleotide synthesis

To investigate the effect of 2′,5′- and 3′,5′-linked **H** units on the stability of duplexes, monomers **8** and **11** were incorporated into various oligonucleotide sequences by conventional

Table 1. Thermal denaturation data (T_m) of duplexes

No.	Designation	Oligonucleotide	T_m[a] (ΔT_m)[b] DNA	RNA	2′,5′-RNA
2′,5′-RNA containing 2′,5′-**H** (***H***) and 3′,5′-**H** (**H**)[c]					
I	2,5RNA	5′-rGUC GUG UGU GUG ACU CUG GUA AC-2′	br[d]	61.6	41.3
II	2′,5′-**H**	5′-rGUC UGU ***H***GU GUG CUG GUA AC-2′	br	60.7 (-0.9)	40.5 (−0.8)
III	2′-5′-**H**×2	5′-rGUC UG***H*** ***H***GU GUG ACU CUG GUA AC-2′	br	60.6 (-0.5)	40.0 (−0.7)
IV	2′,5′-**H**×3	5′-rGUC UG***H*** ***HGH*** GUG ACU CUG GUA AC-2′	br	60.5 (-0.3)	39.1 (−0.6)
V	2′,5′-**rA**	5′-rGUC UGU **A**GU GUG ACU CUG GUA AC-2′	br	56.8 (-4.8)	37.3 (−3.5)
VI	2′,5′-**rA** ×2	5′-rGUC UG**A** **A**GU GUG ACU CUG GUA AC-2′	br	54.3 (-3.7)	34.0 (−3.7)
VII	2′,5′-**rA** ×3	5′-rGUC UG**A** **AGA** GUG ACU CUG GUA AC-2′	br	48.0 (-4.5)	28.0 (−4.3)
VIII	3′,5′-**H**	5′-rGUC UGU **H**GU GUG ACU CUG GUA AC-2′	br	58.0 (-3.6)	37.4 (−3.9)
RNA containing 2′,5′-**H** (***H***) and 3′,5′-**H** (**H**)					
IX	RNA	5′-rGUC UGU UGU GUG ACU CUG GUA AC-3′	65.0	78.4	57.1
X	2′,5′-H	5′-rGUC UGU ***H***GU GUG ACU CUG GUA AC-3′	62.0 (−3.0)	77.2 (−1.2)	52.8 (−4.3)
XI	3′,5′-H	5′-rGUC UGU **H**GU GUG ACU CUG GUA AC-3′	63.0 (−2.0)	77.7 (−0.7)	54.1 (−3.0)
XII	2′,5′-**rU**	5′-rGUC UGU **U**GU GUG ACU CUG GUA AC-3′	62.0 (−3.0)	78.0 (−0.4)	54.1 (−3.0)
XIII	3′,5′-**rA**	5′-rGUC UGU **A**GU GUG ACU CUG GUA′	61.1 (−3.9)	74.1 (−4.3)	51.1 (−6.0)
DNA control					
XIV	DNA	5′-dGTC TGT TGT GTG ACT CTG GTA AC-3′	68.0	69.0	br

[a]T_m in °C, deviation ±1°C; complementary DNA sequence 5′-dGUU ACC AGA GUC ACA CAA CAG AC-3′; complementary RNA sequence, 5′-rGUU ACC AGA GUC ACA CAA CAG AC-3′; complementary 2′,5′-RNA sequence, 5′-GUU ACC AGA GUC ACA CAA CAG AC-2′; buffer, 140 mM K^+, 1 mM Mg^{2+} and 5 mM Na_2HPO_4, pH = 7.2.
[b]ΔT_m, the T_m change per one modified nucleotide or mismatch relative to entries I or IX.
[c]2′,5′-**H** is shown as ***H***; 3′,5′-**H** as **H**; 2′,5′-**rA** as **A**, 3′,5′-rA as **A**; 2′,5′-rU as **U** in the sequence.
[d]Broad transition.

phosphoramidite chemistry (Table 1) (1,15,20). 4,5-Dicyanoimidazole (DCI) was used to activate the phosphoramidites (15) and provided average coupling efficiencies of 99% as monitored by the release of the monomethoxytrityl (MMT) cation (Materials and Methods). Deprotection conditions of oligonucleotides were similar to those employed in RNA synthesis (20). Oligonucleotides were purified by anion-exchange HPLC and their identity verified by MALDI-TOF mass spectrometry (Supplementary Data). The HPLC chromatograms showed that 2′,5′-RNA oligomers elute more rapidly relative to 3′,5′-RNA oligomers of the same base composition (Supplementary Data).

Hybridization studies (T_m and CD analysis)

The binding affinity of various mixed backbone oligonucleotides with one to three **H** units to complementary single-stranded DNA (ssDNA), ssRNA and 2′,5′-ssRNA targets was evaluated in a buffer designed to simulate intracellular conditions (Table 1). The oligomers, 23 nt in length, were complementary to portions of the U5 region of HIV-1 genomic RNA. Oligomers containing 2′,5′-**H** substitutions (**II-IV, X**), 3′,5′-**H** substitutions (**VIII, XI**) or mismatched bases (**V-VII, XIII**) were prepared in order to ascertain Watson–Crick pairing of the **H** residues. As a comparison, the hybridization properties of unmodified 2′,5′-RNA (**I**), RNA (**IX**) and DNA (**XIV**) sequences were also measured, along with a 3′,5′-linked oligomer containing a single 2′,5′-linked uridine residue (**XII**). Thermal dissociation data (T_m, ΔT_m) for the complexes formed are presented in Table 1 and representative melting and CD curves are shown in Figure 3. The key observations can be summarized as follows.

2′,5′-H substitutions in the 2′,5′-RNA strand. Both 2′,5′-RNA:RNA and 2′,5′-RNA: 2′,5′-RNA duplexes can accommodate a single 2′,5′-**H** residue with a small loss of stability ($\Delta T_m \sim -0.8$–0.9°C). This destabilization is significantly smaller than that created by a mismatch at the same position ($\Delta T_m \sim -4$°C), suggesting that 2′,5′-**H** residues in these duplexes retain classical base-pairing interactions. When the number of 2′,5′-**H** units is increased to three (e.g. hybrid **IV**:RNA) the depression in T_m is only −0.3°C/modification, suggesting a stabilizing effect by the nearly contiguous 2′,5′-**H** residues (Figure 3A). A nearly superimposable CD signature of the singly and triply substituted hybrids (**II**:RNA and **IV**:RNA) with control duplex **I**:RNA was observed, strongly suggesting that the 2′,5′-**H** inserts do not perturb the global morphology of the duplexes (Figure 3C). The affinity towards RNA targets is likely to be dependent upon the ratio and intrastrand placement of normal (2′,5′-rN) and 2′,5′-**H** units, and based on the above results, a smaller thermal destabilization is to be expected upon increasing the number of consecutive **H** residues. In fact, based on the observed trend from one to three inserts, a fully-modified strand constructed from 2′,5′-**H** units might possibly have a binding affinity to RNA targets comparable with that of 2′,5′-RNA.

The modest destabilization induced by 2′,5′-**H** units is remarkable in view of the one bond chain extension introduced at each of these residues compared with 2′,5′-RNA (Figure 1 and Scheme 1). In addition to the energy (electrostatic) considerations described above, these phenomena may be explained by the anticipated conformation of the 2′,5′-**H** units (Figure 2). Yathindra and coworkers (10,11) have shown that the sugars of 2′,5′-linked RNA favor a C2′-*endo* conformation, which renders the backbone to be 'compact' and of almost equivalent length to that found in the native RNA (i.e. C3′-*endo*). With the C3′-*endo* sugar, the 3′-hydroxyl group would sterically interfere with the 2′-*O*-phosphate linkages. The same 'compact' conformation is expected for 2′,5′-**H** residues, since a strong O4′-C4′-C3′-O3′ gauche effect (and lack of an opposing O4′-C1′-C2′-O2′ gauche effect) would reinforce the C2′-*endo* pucker (Figure 2). Hence, oligonucleotides that are pre-organized in a 'compact' (or RNA-like)

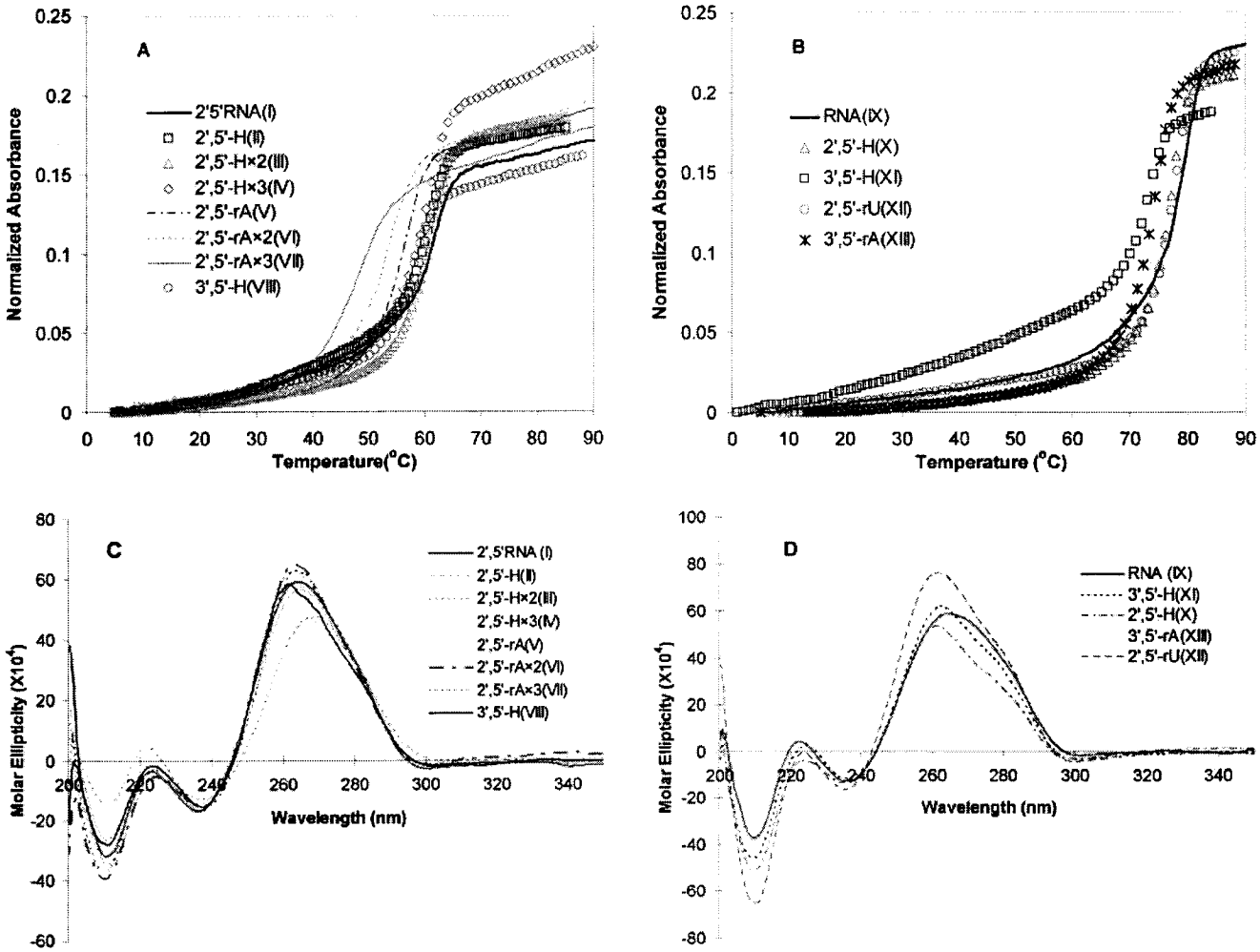


Figure 3. Buffer: 140 mM K^+, 1 mM Mg^{2+} and 5 mM Na_2HPO_4, pH 7.2. Oligonucleotides were hybridized to complementary RNA. (**A** and **C**) Thermal melting curves and CD profile, respectively, of 2′,5′-RNA with 2′,5′-**H**, 3′,5′-**H** and mismatch 2′,5′-rA. (**B** and **D**) Thermal melting curves and CD profile, respectively, of RNA with 2′,5′-**H**, 3′,5′-**H**, 2′,5′-rU and mismatch 3′,5′-rA.

conformation are expected to bind to 'compact' RNA strands, and bind weakly, if at all, to 'extended' oligonucleotide targets (e.g. ssDNA). Such conformational compatibility or spatial 'matching' probably accounts for the ability of oligonucleotides **I** through **VIII** to maintain a stable association with complementary RNA, but not with ssDNA (Table 1) (1).

2′,5′-H substitutions in the RNA strand and 3′,5′-H substitutions in the 2′,5′-RNA strand. The destabilizing effect of 2′,5′-**H** units appears to be greater when incorporated into RNA strands ($\Delta T_m > -1$°C/modification), regardless of whether the target is complementary RNA or 2′,5′-RNA (Sequence **X**, Table 1). Similarly, the presence of 3′,5′-**H** within 2′,5′-RNA leads to reduced duplex stability, with an equal destabilization observed when targeting both RNA and 2′,5′-RNA (Sequence **VIII**, Table 1). Consistent with this notion, the singly-substituted hybrids **VIII**:RNA and **X**:RNA showed a blue shift in the CD band at 270 nm, whereas hybrids **II**:RNA and **XI**:RNA displayed nearly the same CD profile as the corresponding unmodified controls (Figure 3C and D). As pointed out previously (1,11), these observations may reflect the major disruption in the normal helical structure induced by the abrupt displacement of the backbone from the periphery towards the interior of the helix (2′-CH_2OP to 3′-OP) and vice versa (3′-OP to 2′-CH_2OP).

3′,5′-H substitutions in the RNA strand. Regarding the effects of 3′,5′-**H** units within an RNA strand on hybridization affinity for complementary RNA, a significantly less pronounced decrease in duplex stability was observed in *RNA*:RNA duplexes ($\Delta T_m = -0.7$°C) than for Schmit's *DNA*:RNA hybrids (3′,5′-**H** in the antisense DNA strand; $\Delta T_m \sim -3$°C) (12). It may be speculated that this is caused, at least in part, by the narrower minor groove width of *DNA*:RNA hybrids (compared with *RNA*:RNA) leading to more unfavorable interactions involving the large 2′-CH_2-OH group. A much larger decrease in binding affinity was observed for the RNA strand containing a mismatch, confirming the contribution of the thymine base of 3′,5′-**H** to binding (**XIII**:RNA versus **XI**:RNA; Table 1).

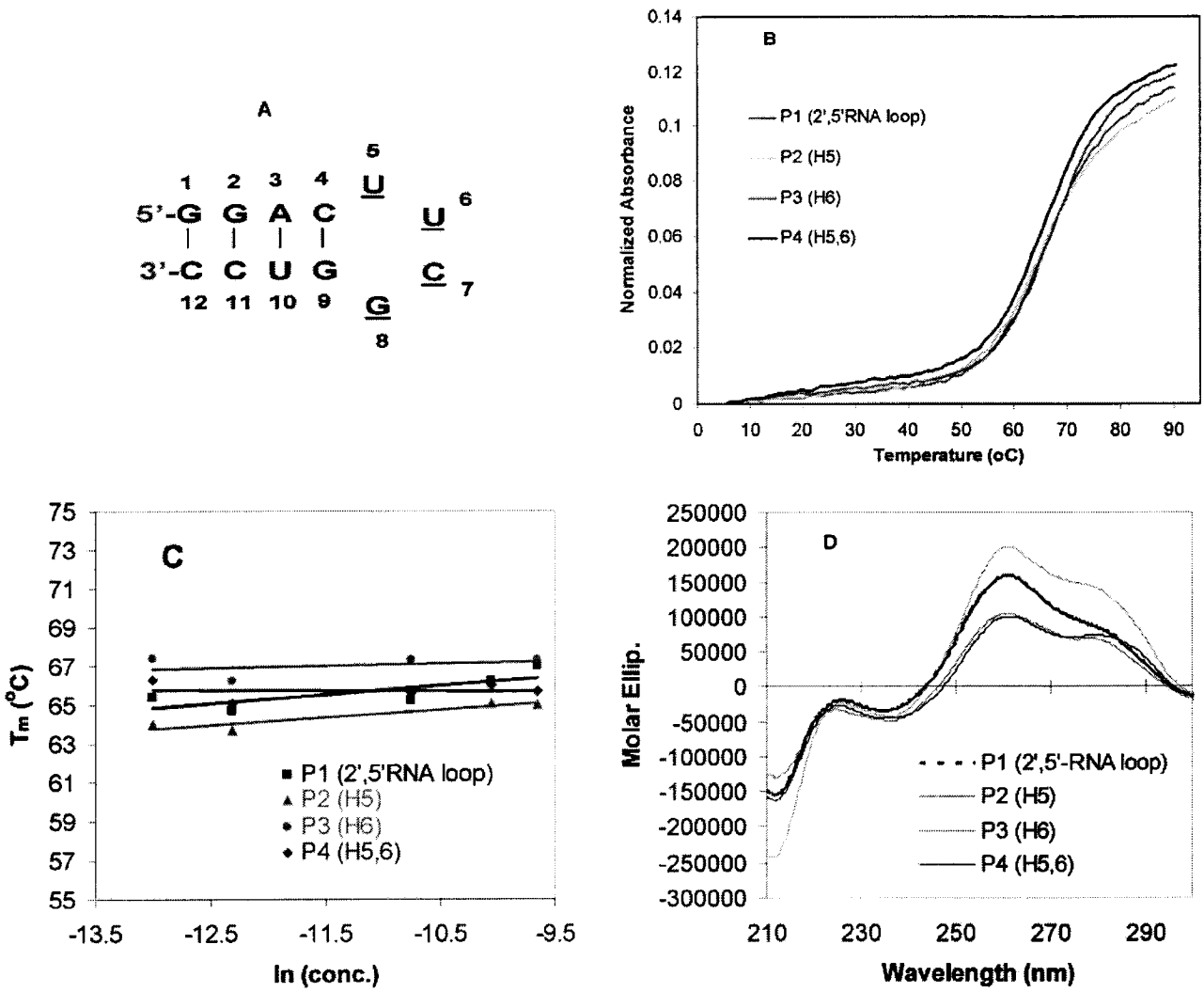


Figure 4. (**A**) Hairpin structure and residue numbering: the tetra loop UUCG is connected via 2′,5′-phosphodiester linkages, and the modification occurs at the $\underline{U}_5$ and/or $\underline{U}_6$ positions. (**B**) Thermal melting curves of RNA hairpins with modified 2′,5′-linked loops. (**C**) T_m concentration independence over 30-fold range. (**D**) CD spectra.

Stabilization of hairpin structures by 2′,5′-H units

We next turned our attention towards RNA 'hairpins' and studied the effect of incorporating single or multiple 2′,5′-**H** units into the loop of these structures (Figure 4A and Table 2). We have shown previously that hairpin **P1** retains an A-form conformation, displays ample resistance against nucleases, and inhibits the RNase H activity of HIV RT (15,21,22). The ribonucleotide residues in the loop are connected by 2′,5′-phosphodiester linkages and collectively fold into a distinct rigid structure that is unlike the native 3′,5′-tetraloop structure (15). In view of the anticipated compact conformation of 2′,5′-**H** residues (Figure 2), it was of interest to discover how these units would influence the stability of the hairpin. This was particularly intriguing given that a recent study by Denisov *et al.* (22) showed that the 2′,5′-rUCCG loop residues of hairpin **P1** adopt a U5(extended)-U6(compact)-C7(compact)-G8(extended) conformation. Therefore, we anticipated that substitution at position 6 would be better tolerated than one at position 5, particularly if the compact conformation of the 2′,5′-**H** unit was preserved at both loop positions. Indeed, we found that a 2′,5′-**H** unit at position 6 (**P3**) increases hairpin stability, with an increase in T_m of +1.5°C relative to

Table 2. Thermal denaturation data of RNA hairpins with modified 2′,5′-linked loops

ID	Designation[a]	Hairpin loop	T_m[b] (ΔT_m)[c]
P1	RRR	5′-GGAC(UUCG)GUCC-3′	64.7
P2	R$\overline{\text{R}}$(**H**$_5$)R	5′-GGAC($\overline{\textit{H}\text{UCG}}$)GUCC-3′	63.7 (−1.0)
P3	RR(**H**$_6$)R	5′-GGAC(U***H***CG)GUCC-3′	66.2 (+1.5)
P4	RR(**H**$_{5,6}$)R	5′-GGAC(***HH***CG)GUCC-3′	65.1 (+0.2)

[a]Underlined residues are connected via 2′,5′-phosphodiester linkages.
[b]T_m was measured at the wavelength of 260 nm in 0.01 M Na_2HPO_4 and 0.1 mM Na_2EDTA, pH 7.0; oligonucleotide concentration ~4.5 μM. Values represent the average of at least five independent measurements. Error in T_m is within ±1°C.
[c]ΔT_m, the T_m change per one 2′,5′-**H**(represented as ***H***in the sequence).

the parent hairpin **P1** (Table 2). This stabilization is nearly offset by the incorporation of a second 2′,5′-**H** unit at position 5 (**P4**, ΔT_m = +0.2°C), which may again be rationalized on the basis of a very strong preference of the 2′,5′-**H** unit for a compact (C2′-*endo*) conformation. In view of this, it is not surprising that a single 2′,5′-**H** substitution at position 5 (**P2**), where an extended nucleotide conformation is preferred, is destabilizing (ΔT_m = −1.0°C). Regarding the CD profiles, the solution conformations of all hairpins are similar to that of the control **P1**, with only slight variations in the intensity of the bands. The hairpins containing 2′,5′-**H** units at position 5 (**P2** and **P4**) were found to present a reduction in the positive CD band at 260 nm (Figure 4D) relative to **P1** and **P3**, consistent with the expected conformational change at this position, i.e. extended 2′,5′-rU→compact 2′,5′-**H**.

CONCLUSIONS

In summary, we have studied the behavior of oligoribonucleotides (2′,5′-RNA and RNA) containing **H** units toward complementary RNA, 2′,5′-RNA and DNA. The data obtained demonstrated a destabilization effect upon substituting a rU with **H**, regardless of whether **H** was 2′,5′ or 3′,5′-linked. This destabilization was minimized when the 2′,5′-**H** and 3′,5′-**H** were incorporated into 2′,5′-RNA and RNA strands, respectively (ΔT_m < −1.0°C). It is expected that longer oligomers containing this modification, particularly if they contain 2′,5′-linkages, will show adequate thermal stability, significant nuclease stability (23) and binding to mRNA targets.

In spite of the drop in thermal stability observed when 2′,5′-**H** units are incorporated into RNA duplexes, a significant increment in stability was observed when they were incorporated into hairpin loops (ΔT_m = +1.5°C). These results and those described above lend strong support to the notion of 'compact'/'extended' 2′,5′/3′,5′-backbone structure and its effect on hybrid stability (11). The ability of hairpins containing stabilizing 2′,5′-**H** units to inhibit HIV RT *in vitro* will be the subject of a separate publication.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr R.N. Hannoush for contributions during the early stages of this work. We thank J. Watts, D. Sabatino and M. Mangos for helpful feedback during the preparation of this manuscript. We acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a grant to M.J.D. C.G.P. acknowledges support from a FQRNT fellowship and a Clifford Wong McGill Major Fellowship. M.J.D. is recipient of a James McGill Professorship (McGill University). Funding to pay the Open Access publication charges for this article was provided by NSERC Canada.

Conflict of interest statement. None declared.

REFERENCES

1. Wasner,M., Arion,D., Borkow,G., Noronha,A., Uddin,A.H., Parniak,M.A. and Damha,M.J. (1998) Physicochemical and biochemical properties of 2′,5′-linked RNA and 2′,5′-RNA:3′,5′-RNA hybrid Duplexes. *Biochemistry*, **37**, 7478–7486.
2. Peng,C.G., Hannoush,R.N. and Damha,M.J. (2005) Synthesis of a 'taller' 2′,5′-linked ribonucleic acid: 2′-α-C-hydroxymethyl 2′,5′-linked RNA. *J. Biomol. Struct. Dyn.*, **22**, 856.
3. Giannaris,P.A. and Damha,M.J. (1993) Oligoribonucleotides containing 2′,5′-phosphodiester linkages exhibit binding selectivity for 3′,5′-RNA over 3′,5′-ssDNA. *Nucleic Acids Res.*, **21**, 4742–4749.
4. Bhan,P., Bhan,A., Hong,M., Hartwell,J.G., Saunders,J.M. and Hoke,G.D. (1997) 2′,5′-Linked oligo-3′-deoxyribonucleoside phosphorothioate chimeras: thermal stability and antisense inhibition of gene expression. *Nucleic Acids Res.*, **25**, 3310–3317.
5. Damha,M.J. and Noronha,A. (1998) Recognition of nucleic acid double helices by homopyrimidine 2′,5′-linked RNA. *Nucleic Acids Res.*, **26**, 5152–5156.
6. Torrence,P.F., Maitra,R.K., Lesiak,K., Khamnei,S., Zhou,A. and Silverman,R.H. (1993) Targeting RNA for degradation with a (2′-5′)oligoadenylate-antisense chimera. *Proc. Natl Acad. Sci. USA*, **90**, 1300–1304.
7. Xiao,W., Li,G., Maitra,R.K., Maran,A., Silverman,R.H. and Torrence,P.F. (1997) Correlation of selective modifications to a 2′,5′-oligoadenylate-3′,5′-deoxyribonucleotide antisense chimera with affinity for the target nucleic acid and with ability to activate RNAase L. *J. Med. Chem.*, **40**, 1195–1200.
8. Kandimalla,E.R., Manning,A., Zhao,Q., Shaw,D.R., Byrn,R.A., Sasisekharan,V. and Agrawal,S. (1997) Mixed backbone antisense oligonucleotides: design, biochemical and biological properties of oligonucleotides containing 2′-5′-ribo- and 3′-5′-deoxyribonucleotide segments. *Nucleic Acids Res.*, **25**, 370–378.
9. Damha,M.J., Giannaris,P.A. and Khan,N. (1991) 2′-5′-Linked oligonucleotides form stable complexes with complementary RNA and DNA. *Nucleic Acids Symp. Ser.*, 290.
10. Premraj,B.J., Patel,P.K., Kandimalla,E.R., Agrawal,S., Hosur,R.V. and Yathindra,N. (2001) NMR structure of a 2′,5′ RNA favors a type duplex with compact C2′ endo nucleotide repeat. *Biochem. Biophys. Res. Commun.*, **283**, 537–543.
11. Premraj,B.J., Raja,S. and Yathindra,N. (2002) Structural basis for the unusual properties of 2′,5′ nucleic acids and their complexes with RNA and DNA. *Biophys. Chem.*, **95**, 253–272.
12. Schmit,C., Bevierre,M.-O., De Mesmaeker,A. and Altmann,K.-H. (1994) The effects of 2′- and 3′-alkyl substituents on oligonucleotide hybridization and stability. *Bioorg. Med. Chem. Lett.*, **4**, 1969–1974.
13. Freier,S.M. and Altmann,K.H. (1997) The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Res.*, **25**, 4429–4443.
14. Martin,P. (1995) New access to 2′-*O*-alkylated ribonucleosides and properties of 2′-*O*-alkylated oligoribonucleotides. *Helv. Chim. Acta*, **78**, 486–504.
15. Hannoush,R.N. and Damha,M.J. (2001) Remarkable stability of hairpins containing 2′,5′-linked RNA loops. *J. Am. Chem. Soc.*, **123**, 12368–12374.
16. Alberts,B. (1989) Molecular Biology of the Cell. Carland Publishing, Inc., NY, pp. 304.
17. Schmit,C. (1994) Efficient synthesis of 2′-deoxy-2′-a-C-substituted nucleosides. *Synlett*, 238–240.
18. Li,N.-S. and Piccirilli,J.A. (2004) Synthesis of the phosphoramidite derivatives of 2′-deoxy-2′-C-a-methylcytidine and 2′-deoxy-2′-C-a-hydroxymethylcytidine: analogues for chemical dissection of RNA's 2′-hydroxyl group. *J. Org. Chem.*, **69**, 4751–4759.
19. Kocienski,P.J. (2004) Protecting group. Georg Thieme, NY, pp. 215.
20. Damha,M.J. and Ogilvie,K.K. (1993) Oligoribonucleotide synthesis. The silyl-phosphoramidite method. *Methods Mol. Biol.*, **20**, 81–114.
21. Hannoush,R.N., Carriero,S., Min,K.-L. and Damha,M.J. (2004) Selective inhibition of HIV-1 reverse transcriptase (HIV-1 RT) RNase H by small RNA hairpins and dumbbells. *ChemBioChem*, **5**, 527–533.
22. Denisov,A.Y., Hannoush,R.N., Gehring,K. and Damha,M.J. (2003) A novel RNA motif based on the structure of unusually stable 2′,5′-linked r(UUCG) loops. *J. Am. Chem. Soc.*, **125**, 11525–11531.
23. Hannoush,R.N., Min,K.-L. and Damha,M.J. (2004) Diversity-oriented solid-phase synthesis and biological evaluation of oligonucleotide hairpins as HIV-1 RT RNase H inhibitors. *Nucleic Acids Res.*, **32**, 6164–6175.

American Chemical Society

Publications Division
Copyright Office

1155 Sixteenth Street, NW
Washington, DC 20036
Phone: (1) 202-872-4368 or -4367
Fax: (1) 202-776-8112 E-mail: copyright@acs.org

VIA FAX: 514-398-2382 DATE: May 2, 2007

TO: Chang Geng Peng, Department of Chemistry, McGill University
3420 University, Rm. 209, Pulp and Paper Bldg., Montreal, Quebec H3A 2K7, Canada

FROM: C. Arleen Courtney, Copyright Associate C. Arleen Courtney

Thank you for your request for permission to include **your** paper(s) or portions of text from **your** paper(s) in your thesis. Permission is now automatically granted; please pay special attention to the implications paragraph below. The Copyright Subcommittee of the Joint Board/Council Committees on Publications approved the following:

Copyright permission for published and submitted material from theses and dissertations
ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).

Publishing implications of electronic publication of theses and dissertation material
Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journal's editor. If an ACS journal editor considers Web posting to be "prior publication", the paper will not be accepted for publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically.

If your paper has **not** yet been published by ACS, we have no objection to your including the text or portions of the text in your thesis/dissertation in **print and microfilm formats**; please note, however, that electronic distribution or Web posting of the unpublished paper as part of your thesis in electronic formats might jeopardize publication of your paper by ACS. Please print the following credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [JOURNAL NAME], in press (or 'submitted for publication'). Unpublished work copyright [CURRENT YEAR] American Chemical Society." Include appropriate information.

If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in **print or microfilm formats**, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information.

Submission to a Dissertation Distributor: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination through the distributor; ACS's copyright credit line should be printed on the first page of the ACS paper.

Use on an Intranet: The inclusion of your ACS unpublished or published manuscript is permitted in your thesis in print and microfilm formats. If ACS has published your paper you may include the manuscript in your thesis on an intranet that is not publicly available. Your ACS article cannot be posted electronically on a publicly available medium (i.e. one that is not password protected), such as but not limited to, electronic archives, Internet, library server, etc. The only material from your paper that can be posted on a public electronic medium is the article abstract, figures, and tables, and you may link to the article's DOI or post the article's author-directed URL link provided by ACS. This paragraph does not pertain to the dissertation distributor paragraph above. 06/07/06

Table 1. Apparent Fidelity of FANA Synthesis in Dropout Assays

		% apparent fidelity[a]			
polymerase	2′F-araNTP		dNTP	template used	figure
Ph	A:	>99%	>99%	PF21	3
	G:	>99%	>99%		
	T:	>99%	>99%		
	C:	>99%	>99%		
DV	A:	>99%	1%	PF41	2
	G:	>99%	62%	PF43	S3
	T:	3%	19%	PF21	S4
	C:	50%	3%	PF23	S4
9N	A:	>99%	80%	PF41	2
	G:	98%	84%	PF43	S3
	T:	76%	32%	PF21	S4
	C:	43%	82%	PF23	S4

[a] Apparent fidelity is calculated according to the equation 1 − [(% full-length-2′F-araNTP)/ (% full-length+2′F-araNTP)] at 30 min reaction time (sample equation for dNTP values); >99% means that the full-length product is not detectable in our dropout assays.

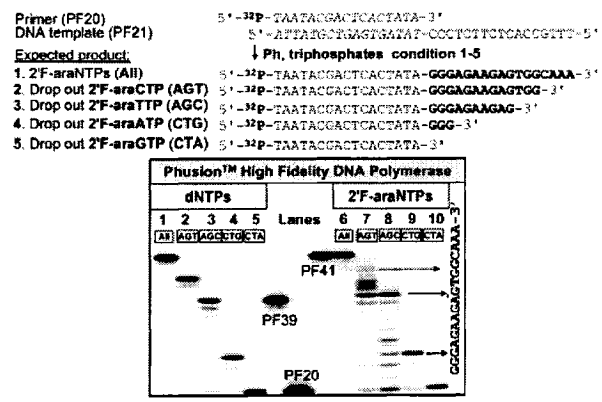


Figure 3. Fidelity study on 2′F-araNTP incorporation by Ph on DNA template PF21. Reaction conditions: 0.8 unit Ph, 25 μM triphosphate at 55 °C in 30 μL of reaction volume for 30 min. Lane 1, dNTPs; lane 2, drop out dCTP (i.e., dATP, dGTP, and dTTP, shorten as AGT; same coding for lanes 3−5), lane 6, 2′F-araNTPs; lane 7, drop out 2′F-araCTP (i.e., 2′F-araATP, 2′F-araGTP, 2′F-araTTP, shorten as AGT; same coding for lanes 8−10). Template PF41 and primer PF20 are shown; PF39 is a 27-nt DNA control.

thermophilic polymerases (e.g., the error rates for Ph[25] and DV[26] are 4.4 × 10^{-7} and 2.2 × 10^{-4}, respectively). All four dropout experiments were conducted with the same DNA template PF21 (Figure 3), or PF41 (Figure S5, Supporting Information). For the dNTPs, syntheses stopped at the expected termination sites, yielding products lacking the required dNTP at the 3′-terminus. A similar pattern was observed with 2′F-araNTPs, although in the latter, FANA synthesis proceeded with more difficulty (Figure 3; lanes 6−10). This effect, however, was observed only in the dropout assays, that is, the reaction containing all four 2′F-araNTPs produced excellent yields of the full-length FANA product with virtually no pausing observed. We speculate that in a dropout assay, where polymerization activity is suboptimal, the pausing bands observed may result, at least in part, by the shuttling (cycling) between the strong Ph 3′-5′ exonuclease and polymerase activities.[18,27] Similar results were observed with the DNA template PF41 (Figure S5). Of note, apparent fidelity among Ph, DV, and 9N is correlated to 3′-5′ exonuclease proofreading activity, with Ph having the strongest exonuclease activity and also 2′F-araNTP fidelity, and DV having no exonuclease activity and low 2′F-fidelity. We recently became aware of Wengel's work that Ph can incorporate a few LNA triphosphates in a growing DNA strand.[28] Thus, FANA and LNA NTPs appear to be the first examples of modified NTPs incorporated by Ph.

We are currently testing the limits of these procedures by exploring the synthesis of (a) DNA−FANA on a DNA−FANA template and (b) DNA on an all-FANA template. Preliminary results obtained indicate that DV, 9N, and Ph can synthesize FANA−DNA strands on FANA−DNA templates (Figure S6, Supporting Information), and that KF and *Bst* DNA can catalyze FANA template-directed DNA synthesis (Figure S7, Supporting Information). While we have not been able to synthesize an all-FANA strand on an all-FANA template, it is possible for DV, 9N, and Ph polymerase to drive the formation of multiple FANA−FANA base pairs within a DNA−FANA chimeric duplex (Figure S6; Supporting Information).

In summary, family B thermophilic DNA polymerases such as DV, 9N, and Ph can utilize 2′F-araNTPs, or a combination of 2′F-araNTPs and dNTPs, to generate FANA or chimeric FANA−DNA strands, respectively. In addition, these polymerases were shown to synthesize chimeric FANA−DNA strands on a chimeric FANA−DNA template. KF and *Bst* (family A) DNA polymerases were able to incorporate dNTPs on a template FANA strand. These results suggest that it should be possible to evolve FANA-modified aptamers via SELEX.

Acknowledgment. This work was supported by the Canadian Institutes of Health Research and Topigen Pharmaceuticals, Inc. Dedicated to Prof. Kelvin K. Ogilvie on the occasion of his 65th birthday.

Supporting Information Available: Experimental details and gel pictures (Figures S2−S7). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818−822.
(2) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505−510.
(3) Ruckman, J.; Green, L. S.; Beeson, J.; Waugh, S.; Gillette, W. L.; Henninger, D. D.; Claesson-Welsh, L.; Janjic, N. *J. Biol. Chem.* **1998**, *273*, 20556−20567.
(4) Lato, S. M.; Ozerova, N. D. S.; He, K.; Sergueeva, Z.; Shaw, B. R.; Burke, D. H. *Nucleic Acids Res.* **2002**, *30*, 1401−1407.
(5) Jhaveri, S.; Olwin, B.; Ellington, A. D. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2285−2290.
(6) Dineva, M. A.; Ivanov, I. G.; Petkov, D. D. *Nucleosides Nucleotides* **1997**, *16*, 1875−1882.
(7) Kato, Y.; Minakawa, N.; Komatsu, Y.; Kamiya, H.; Ogawa, N.; Harashima, H.; Matsuda, A. *Nucleic Acids Res.* **2005**, *33*, 2942−2951.
(8) Kuwahara, M.; Hanawa, K.; Ohsawa, K.; Kitagata, R.; Ozaki, H.; Sawai, H. *Bioorg. Med. Chem.* **2006**, *14*, 2518−2526.
(9) Trempe, J.-F.; Wilds, C. J.; Denisov, A. Y.; Pon, R. T.; Damha, M. J.; Gehring, K. *J. Am. Chem. Soc.* **2001**, *123*, 4896−4903.
(10) Kalota, A.; Karabon, L.; Swider, C. R.; Viazovkina, E.; Elzagheid, M.; Damha, M. J.; Gewirtz, A. M. *Nucleic Acids Res.* **2006**, *34*, 451−461.
(11) Wilds, C. J.; Damha, M. J. *Nucleic Acids Res.* **2000**, *28*, 3625−3635.
(12) Wilds, C. J.; Damha, M. J. *Bioconjug. Chem.* **1999**, *10*, 299−305.
(13) Wilds, C. J. Ph.D. Thesis, McGill University, Montreal, Canada, 2000.
(14) Dowler, T.; Bergeron, D.; Tedeschi, A. L.; Paquet, L.; Ferrari, N.; Damha, M. J. *Nucleic Acids Res.* **2006**, *34*, 1669−1675.
(15) Wright, G. E.; Brown, N. C. *Pharmacol. Ther.* **1990**, *47*, 447−97.
(16) Shields, A. F. *Mol. Imaging Biol.* **2006**, *8*, 141−150.
(17) Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274−9275.
(18) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395−17398.
(19) DNA polymerases abbreviations: deep vent (3′-5′ exo-), DV; $9^{o}N_{m}$, 9N; Therminator, Th; *Bst* large fragment, *Bst*; *Taq*, *Taq*; Phusion high-fidelity, Ph; Klenow fragment (3′-5′ exo-), KF; Moloney Murine leukemia virus reverse transcriptase, MMLV.
(20) Gardner, A. F.; Jack, W. E. *Nucleic Acids Res.* **2002**, *30*, 605−613.
(21) Ilchida, J. K.; Zou, K.; Horhota, A.; Yu, B.; McLaughlin, L. W.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 2802−2803.
(22) Kunkel, T. A. *J. Biol. Chem.* **2004**, *279*, 16895−16898.
(23) Kool, E. T. *Annu. Rev. Biochem.* **2002**, *71*, 191−219.
(24) Marquez, V. E.; Ezzitouni, A.; Russ, P.; Siddiqui, M. A.; Ford, H., Jr.; Feldman, R. J.; Mitsuya, H.; George, C.; Barchi, J. J., Jr. *J. Am. Chem. Soc.* **1998**, *120*, 2780−2789.
(25) NEB manufacturer's instruction.
(26) Huang, H. X.; Keohavong, P. *DNA Cell Biol.* **1996**, *15*, 589−594.
(27) Kunkel, T. A.; Bebenek, K. *Annu. Rev. Biochem.* **2000**, *69*, 497−529.
(28) Veedu, R. N.; Vester, B.; Wengel, J. *Conference Book*; Haenni, D. L., Ed.; International Roundtable on Nucleosides, Nucleotides and Nucleic Acids: Bern, Germany, 2006; p 292.

JA069100G