

**Synthesis and Biophysical Evaluation of Cyclohexenyl  
Nucleic Acids and its analogues**

**Thesis Submitted to  
Savitribai Phule Pune University  
For the degree of**

**Doctor of Philosophy  
in  
Chemistry**

**BY**

**Manojkumar Varada**

**Research Supervisor**

**Dr. (Mrs.)Vaijayanti A. Kumar**

**Division of Organic Chemistry  
CSIR-National Chemical Laboratory  
Pune-411008**

**June 2015**

## **CERTIFICATE**

This is to certify that the work presented in the thesis entitled “**Synthesis and Biophysical Evaluation of Cyclohexenyl Nucleic Acids and its analogues.**” submitted by Mr. Manojkumar Varada, was carried out by the candidate at the CSIR-National Chemical Laboratory Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

**Dr. Vaijayanti A. Kumar**

(Research Supervisor)

Division of Organic Chemistry

CSIR-National Chemical Laboratory

Pune 411008

**June 2015**

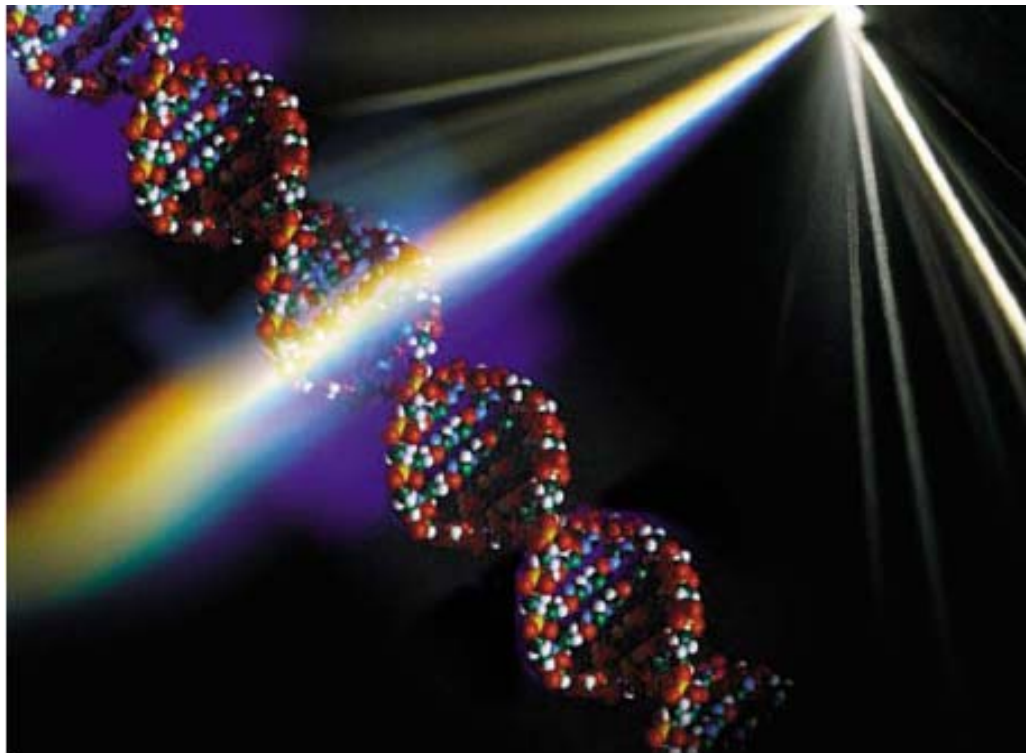
## **CANDIDATE'S DECLARATION**

I hereby declare that the thesis entitled “**Synthesis and Biophysical Evaluation of Cyclohexenyl Nucleic Acids and its analogues.**” submitted for the award of degree of *Doctor of Philosophy* in Chemistry to the Savitribai Phule Pune University has not been submitted by me to any other university or institution. This work was carried out by me at the CSIR-National Chemical Laboratory, Pune, India. Such materials as obtained from other sources have been duly acknowledged in the thesis.

**Manojkumar Varada**

**June 2015**

CSIR-National Chemical Laboratory  
Pune- 411 008



*Dedicated to  
Family members and well wishers*

## **Acknowledgement**

This thesis has been seen through to completion with the support and encouragement of numerous people including my well wishers, friends, and colleagues. At this point of accomplishment I would like to thank all those people who made this thesis possible and an unforgettable experience for me. It is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study.

First and foremost, I would like to express my heartfelt and sincere gratitude to my research supervisor Dr. (Mrs.) Vijayanti A. Kumar for all the advice, guidance, support and encouragement during every stage of this work. The confidence she had in me made to pursue the multiple roles of nucleic acids which always kept me energetic and interesting. Her energy, enthusiasm and scientific discussions are and would be the source of inspiration to realize my dream in this scientific world. I am highly thankful for her patience, motivation, timely advice and the continuous support provided during every stage of my research work. I could not have imagined having a better advisor and mentor for my doctoral study. Although this eulogy is insufficient, I preserve an everlasting gratitude for her.

I would like to express my thanks and appreciation to Dr. Moneesha Fernandes for her professional support and personal care. She was always beside me during the happy and hard moments, to push me and motivate me towards my goal. She was the one along with my mentor.

My special thanks to Dr. (Mrs.) Namrata Erande for her efforts and unconditional support in all possible ways during ups and downs in the course of this work and Amit for help during writing thesis.

I gratefully acknowledge Dr. (Mrs.) Anita Gunjal, and Mrs. Meenakshi Mane and Sunita Kunte for their constant support, encouragement and especially for passing on their years of experience in handling the DNA synthesizer and HPLC analysis.

I take this opportunity to offer my sincere appreciation to Dr. Anil Kumar for his support in all the ways and Dr. (Mrs.) Vandana Pore for her help and valuable suggestions.

My deepest gratitude to my teachers Krishna veni, Prof. H. S. P. Rao, , Dr. C. R. Ramanathan, Dr. N. D. Reddy, for their inspirational teaching, guidance and blessings.

I also extend my sincere thanks to HOD Dr. Pradeep Kumar, Dr. C. V. Ramana, Dr. B. L. V. Prasad, Dr. Hotha Srinivas, Dr. D. Srinivas reddy for their kind help and encouragement during the course of this work. I am also very much thankful to Mrs. Shanta Kumari and their group members for their excellent support in MALDI-TOF, HRMS and LC-MS analysis. The kind support from NMR group and the LC-MS facility, Organic Chemistry Division are also greatly acknowledged.

I have high regards for my seniors Dr. Madhuri, Dr. Sachin, Dr. Seema, Harshit and Namrata for their unconditional support and help during my Ph D. Course. I am indebted to my colleagues Dr. Venu, Dr. Kiran, Dr. Anjan, Tanaya, Govind, Amit, Harsha, Manisha, Ragini and Gandhali for providing a stimulating and fun filled environment. Also I would like to thank Dr. Raman, Dr. Amit Patwa, Dr. Ashwini, Dr. Sridhar Kosgi, Dr. Gitali, Dr. Mahesh, Dr. Roopa, Dr. Manaswani, Pradnya, Tanpreet, Deepak, Satish, Nitin and Vijay for their help and encouragement. I thank Bhumkar for the laboratory assistance.

*Thanks to All my friends*

**Manojkumar Varada**

<b>Contents</b>		
<b>Publication /Symposia</b>		i
<b>Abbreviations</b>		iii
<b>Abstract</b>		vi
<b>Chapter 1</b>		
<b>1</b>	<b>Introduction to nucleic acids</b>	
1.1	Primary structures of DNA and RNA	1
1.2	Base Pairing via Hydrogen bonding	2
1.3	Sugar puckering	3
1.4	Structures of Nucleic acids	4
1.4A	The duplex structure	4
1.4B	Triplex-Forming Oligonucleotides (TFO)	6
1.4C	Quadruplex structure – The ‘G-quadruplex’	7
1.4D	Structural polymorphism in G-quadruplexes	8
1.4E	Thrombin-binding aptamer	9
1.5	Oligonucleotides as Therapeutic Agents	10
1.5.1	Antisense Technology	10
1.5.2	Disruptive antisense approach	11
1.6	Promising chemical modifications of nucleic acids	11
1.6.1	Phosphate linkage modifications (First generation AONs )	12
1.6.2	2’-Modifications of Sugar (Second generation antisense oligonucleotides)	13
1.6.3	Third generation antisense oligonucleotides	14
1.6.4	Threofuranosyl nucleic acid (TNA) and Hexitol nucleic acids (HNA)	15
1.6.5	Six membered carbocyclic analogs	16
1.6.6	Open chain analogs	17
1.7	Tools and techniques for structural studies of nucleic acid complexes, and higher ordered G-quadruplexes	18
1.7.1	UV-spectroscopy	18
1.7.2	Circular Dichroism (CD)	19
1.7.3	Nuclear Magnetic Resonance (NMR)	20
1.7.4	MALDI-TOF Mass Spectroscopy:	20
1.7.5	HPLC	21
1.8	Present work	22
1.9	References	23
<b>Chapter 2</b>		

<b>2</b>	<b>Design and synthesis of cyclohexenyl and <math>\alpha</math>-L-cyclohexenyl nucleosides</b>	
2.1	Introduction	29
2.2	Synthesis of Cyclohexenyl Nucleosides	30
2.3	Our design and rationale	32
2.4	Methodology, Result and Discussion	33
2.5	Enzymatic resolution	36
2.6	Enantiopurity confirmation by chiral shift reagent NMR of rac-2 and ent-2	37
2.7	Synthesis of cyclohexenyl thymine nucleoside	43
2.8	Synthesis of Enantiopure $\alpha$ -L-Cyclohexenyl Adenine, Guanine Nucleosides	44
2.9	Conclusions	46
2.10	Experimental Section	47
2.11	Appendix	56
2.12	References	88
<b>Chapter 3</b>		
<b>SECTION A</b>		
<b>3A</b>	<b>Design, synthesis and biophysical evaluation of Open chain analogues of cyclohexenyl nucleic acids</b>	
3A.1	Introduction	90
3A.2	Design of open chain analogues of cyclohexenyl nucleic acids and rationale	91
3A.3	Synthesis of <i>cis</i> and <i>trans</i> thymine monomers	93
3A.4	Solid phase synthesis of oligonucleotide using phosphoramidite chemistry	94
3A.5	Synthesis of modified oligonucleotides, characterization and UV- $T_m$ studies	95
3A.6	Stability of oligonucleotides to SVPD	98
3A.7	Enzymatic stability studies of modified DNA sequences	99
3A.8	Conclusions	102
3A.9	Experimental Section	103
3A.10	Appendix	110



<b>SECTION B</b>		
<b>3B</b>	<b>Synthesis of <i>cis</i> and <i>trans</i> modified thrombin binding aptamers and their quadruplex formation study</b>	
3B.1	Introduction	145
3B.2	Aptamers: An emerging class of therapeutics	145
3B.3	Discovery, structural features of Thrombin-binding aptamer (TBA)	147
3B.4	Modifications in TBA	148
3B.5	Present work	148
3B.6	Chemical synthesis of <i>cis</i> and <i>trans</i> modified thrombin binding aptamers	149
3B.7	G-quadruplex formation in the presence of monovalent cation using CD spectroscopy	149
3B.8	Duplex stability studies of modified TBA oligomers	151
3B.9	Conclusions	152
3B.10	Experimental Section	153
3B.11	Appendix	156
3B.12	References	161
<b>Chapter 4</b>		
<b>SECTION A</b>		
<b>4A</b>	<b>Design, synthesis and biophysical evaluation of 4'Methoxymethyl threoseoligonucleotides (4'MOM-TNA)</b>	
4A.1	Introduction	165
4A.2	Synthesis of threose nucleosides	165
4A.3	Design of 4'-Methoxymethyl substituted $\alpha$ -L-threose nucleic acid and rationale	167
4A.4	Synthesis of 4'-Methoxymethyl TNA monomers	167
4A.5	Synthesis of modified oligonucleotides, characterization, UV-melting studies	168
4A.6	Experimental Section	170
4A.7	Appendix	177
<b>SECTION B</b>		
<b>4B</b>	<b>Synthesis of 4'-MOM TNA modified thrombin binding aptamer, its quadruplex formation and application as a thrombin inhibitor</b>	
4B.1	Introduction	197
4B.2	Modifications in TBA	198

4B.3	Present work	199
4B.4	Chemical synthesis of 4'-MOM TNA and <i>iso</i> DNA modified TBA sequences	199
4B.5	G-tetraplex formation in presence of thrombin	202
4B.6	CD spectroscopy and $T_m$ measurement of the TBA in presence of water	204
4B.7	Imino proton NMR spectra of TBA sequences	205
4B.8	Anti-thrombin activity measurements	206
4B.9	Stability of quadruplex structure of aptamers to SVPD	207
4B.10	Conclusion	209
4B.11	Experimental Section	210
4B.12	Appendix	211
4B.13	References	216

### **List of Research Publications:**

1. "4'-MOM TNA modified TBA, an upgraded aptamer for high stability, anticoagulation activity and nuclease stability." **ManojkumarVarada**, Namrata D. Erande, Vaijayanti A Kumar\* manuscript under preparation.
2. "Ene-nucleic acids: A different paradigm to DNA chemistry" **ManojkumarVarada**, Namrata D. Erande, Vaijayanti A Kumar\* manuscript submitted.
3. "Glycine-linked nucleoside- $\beta$ -amino acids: Polyamide analogues of nucleic acids." Anjan Banerjee, Seema Bagmare, **ManojkumarVarada**, Vaijayanti A Kumar\* manuscript submitted.
4. "Synthesis of all four nucleoside-based  $\beta$ -amino acids as protected precursors for the synthesis of polyamide-DNA with alternating  $\alpha$ -amino acid and nucleoside- $\beta$ -amino acids." Seema Bagmare, **ManojkumarVarada**, Anjan Banerjee, Vaijayanti A Kumar\*, Tetrahedron. 2013, 69, 1210–1216.
5. "Robust synthesis of enantiopure cyclohexenyl analogues of 2'/3'-deoxyribose sugars as carbocyclic nucleoside precursors." **ManojkumarVarada**, Venubabu kotikam, Vaijayanti A Kumar\* Tetrahedron. 2011, 67, 5744–5749.

### **Symposia Attended/Poster/Oral Presentations:**

1. Invited lecture delivered on '**Alternatives for 3'-5'-sugar-phosphate linkages of DNA Aptamers**' at CRSI symposium , 2015 , NCL.
2. Poster presentation on "**Design, synthesis, duplex binding and G-quadruplex stability studies of novel acyclic olefinic nucleic acids**" Manojkumar varada, Namrata Erande, Vaijayanti A. Kumar.\* Attended XX International Symposium on Bioorganic Chemistry (ISBOC, 2015), IISER, India.

3. Attended **International Meeting on Chemical Biology (ICMB-2013)** Organized by IISER-Pune May 26-28, 2013 Pune, India.
4. Attended the full agenda of **ACS on campus** events at National Chemical Laboratory, Pune, India, October 2012.
5. Attended **Indo-French Conference** On Organic Synthesis, National Chemical Laboratory, Pune, India, Decenber 8-9, 2011.
6. Poster presentation on "Robust synthesis of enantiopureCyclohexenyl analogues of 2'/3' deoxyribose nucleosides" Manojkumarvarada, Venubabu Kotikam, Vaijayanti A. Kumar.\* **Attended XX International Roundtable on Nucleosides,Nucleotides and Nucleic acids (IRT 2012)**, Montreal, Canada.
7. Attended **CRSI Zonal meeting**, conference held at National chemical Laboratory, pune, India, 2011.

## Abbreviations

μL	Microlitre
μM	Micromolar
A	Absorbance
Å	Angstrom
A	Adenine
Ac	Acetate
ACN	Acetonitrile
ANA	Arabinose nucleic acids
AONs	Antisense oligonucleotides
ASO	Antisense oligonucleotide
Bn	Benzyl
bp	Base pair
BSA	Benzenesulfonicacid
Bz	Benzoyl
C	Cytosine
Calcd.	Calculated
cDNA	Complementary DNA
CD	Circular Dichroism
Conc.	Concentrated
dA	Deoxyadenosine
dT	Deoxythymidine
DCM	Dichloromethane
DEPT	Distortionlessenhancement by polarization transfer
DEAD	Diethylazodicarboxylate
DIAD	Diisopropylazodicarboxylate
DI	De-ionized
DIPEA/DIEA	N, N-Diisopropylethylamine
DMAP	4-(N,N-Dimethylamino)pyridine
DMF	<i>N,N</i> -dimethylformamide
DMT-Cl	4,4'-Dimethoxytrityl chloride
DNA	2'-deoxyribonucleic acid
dsDNA	double-strand DNA
Et <sub>3</sub> N	Triethylamine
EtOAc	Ethyl acetate
EtOH	Ethanol
G	Guanine
gm	Gram
GNA	Glycol nucleic acids
h	Hours
HIV	Human Immunodeficiency Virus
HNA	Hexitol nucleic acid
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry

Hz	Hertz
IR	Infrared
L	Liter
LCMS	Liquid Chromatography-Mass Spectrometry
LNA	Locked nucleic acid
M	Molar
MALDI-TOF	Matrix assisted laser desorption ionization-time of flight
MeOH	Methanol
MF	Molecular formula
mg	Milligram
MHz	Megahertz
min	Minutes
miRNA	MicroRNA
mL	Millilitre
mM	Millimolar
mmol	Millimoles
MS	Mass spectrometry
MW	Molecular weight
N	Normal
N-type	North-type
nm	Nanometer
NMR	Nuclear magnetic resonance
Obsd.	Observed
ONs	Oligonucleotides
PAGE	Poly acrylamide gel electrophoresis
PCR	Polymerase chain reaction
Pet-ether	Petroleum ether
pM	Picomolar
ppm	Parts per million
PS	Phosphorothioates
Py	Pyridine
RNase	Ribonuclease H
R <sub>f</sub>	Retention factor
RP-HPLC	Reverse Phase-HPLC
rt	Room temperature
sec	Second
SELEX	<b>S</b> ystematic <b>E</b> volution of <b>L</b> igand by <b>E</b> xponential Enrichment
S-type	South-type
ssDNA	Single-stranded DNA
SVDP	Snake venom phosphodiesterase
T	Thymine
TBA	Thrombin binding aptamer
TBS	<i>tert</i> -Butyldimethylsilyl
TFO	Triplex-forming oligonucleotide
THAP	2',4',6'-Trihydroxyacetophenone monohydrate

THF	Tetrahydrofuran
TLC	Thin layer chromatography
$T_m$	Melting temperature
TMS	Trimethylsilyl
TNA	$\alpha$ -L-Threose nucleic acid
$t_R$	Retention time
U	Uracil
UV-Vis	Ultraviolet-Visible
w/w	Weight by weight

## Abstract

---

The thesis entitled “**Synthesis and Biophysical Evaluation of Cyclohexenyl Nucleic acids and its analogues**” has been divided into four chapters.

---

**Chapter 1:** Introduction to nucleic acids

**Chapter 2:** Design and synthesis of cyclohexenyl and  $\alpha$ -L-cyclohexenyl nucleosides

**Chapter 3:** Open chain analogues of cyclohexenyl nucleic acids

*Section 3A* Design, synthesis and biophysical evaluation of Open chain analogues of cyclohexenyl nucleic acids

*Section 3B:* Synthesis of *cis* and *trans* modified thrombin binding aptamers and their quadruplex formation study

**Chapter 4:** 4'-Methoxymethyl threose Nucleic acids

*Section 4A:* Design, synthesis and biophysical evaluation of 4'-Methoxymethyl threose nucleic acids

*Section 4B:* Synthesis of 4'-MOM TNA modified thrombin binding aptamer, its quadruplex formation and application as a thrombin inhibitor

---

### **Chapter 1: Introduction to nucleic acids**

This chapter briefly describes structure of nucleic acids, and applications of synthetic nucleic acids in antisense therapeutics. The potential of modified oligonucleotides to act as antisense agents, that can inhibit the expression of a target gene in a sequence-specific manner. Besides having a specific binding affinity to a complementary target polynucleotide sequence, antisense oligonucleotide (ASON) desirably should meet the requirements for therapeutic purposes, e.g., potency, bioavailability, low toxicity and low cost. Several modifications have been developed over the last two decades to enhance the effectiveness of antisense oligonucleotides, the sites of the modifications include sugar, base and phosphate groups. In some of the modifications sugar- phosphate backbone was modified or the sugar ring was replaced with other heterocyclic or carbocyclic rings as well as acyclic units. Also ribose sugars in DNA and RNA have a preference for either N or S conformation. This conformational state always has an impact on the binding affinities towards DNA/RNA. This conformational change occurs due to the substituents on the sugar



ring especially 3' and 2' substituents. This substituent effect also has been discussed in this Chapter. Along with the duplex binding affinities, recently discovered biologically important quadruplex forming ONs were also reviewed and describing their important applications in therapeutics.

## Chapter 2: Design and synthesis of cyclohexenyl and $\alpha$ -L-cyclohexenyl nucleosides

Six membered cyclic analogues are one of the several other prominent modifications for antisense therapy. The furanose ring of natural nucleoside was replaced by six membered ring which conferred suitable rigidity to antisense constructs. The 1,5-anhydrohexitol nucleic acid (HNA) modification exhibited the selective and strong RNA recognition, yielding the desired A-type double helix but prohibited RNase-H recognition. The double bond was introduced to increase the flexibility allowing adaptation and enzymatic recognition, hopefully preserving the increased affinity for complementary targets.

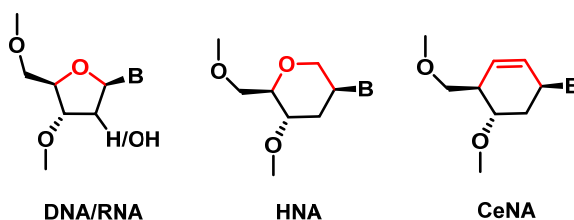


Figure 1 Natural nucleic acids and modified analogues.

The cyclohexenyl nucleosides have demonstrated potent antiviral activity and the CeNA oligomers have been shown to mimic the function of RNA with increased enzymatic and chemical stability. The presence of the double bond allows enough flexibility within a CeNA-RNA double helix, to be recognized by RNaseH. CeNA also been shown to be useful in siRNA applications. We designed the stereo isomer of cyclohexene nucleoside named as  $\alpha$ -L-cyclohexenyl nucleoside. The major advantage of  $\alpha$ -L-cyclohexenyl nucleoside is DNA like(S-type) sugar geometry, which is capable of eliciting the RNase-H activity as well as advantages of both  $\alpha$ -L-LNA and CeNA. We proposed a most efficient common synthetic route for synthesis of both cyclohexenyl and  $\alpha$ -L-cyclohexenyl nucleosides (figure 4).

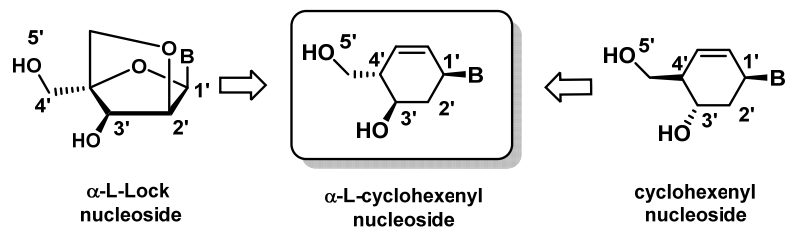
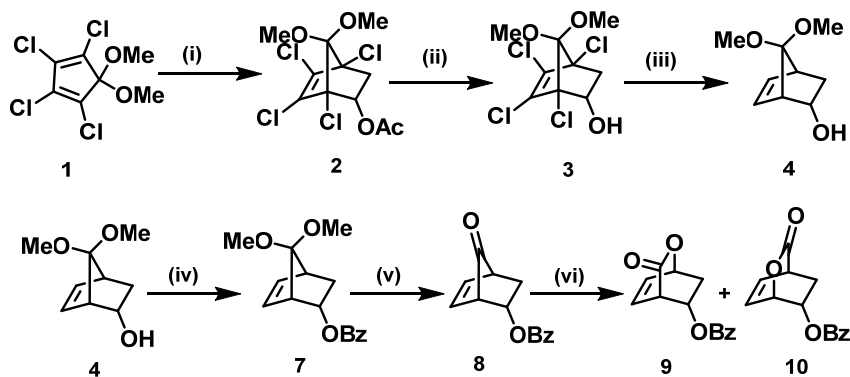


Figure 2: Design of  $\alpha$ -L-cyclohexenyl nucleoside

### Synthesis of cyclohexenyl and $\alpha$ -L-cyclohexenyl nucleosides :

Synthesis started from Diels-Alder reaction between commercially available 5,5-dimethoxy-1,2,3,4-tetrachlorocyclopentadiene **1** and vinyl acetate at 120 °C yields exclusively endo adduct ( $\pm$ )-**2**. Acetate was hydrolyzed followed by dehalogenation gave **4**. Acidic hydrolysis of ketal **4** resulted the rearranged product. To avoid the rearrangement hydroxyl group converted to benzoate. The Bayer-Villiger's oxidation of ketone ( $\pm$ )-**8** gave two regioisomeric inseparable mixture of lactones ( $\pm$ )-**9** and ( $\pm$ )-**10** in 7:3 proportion respectively, in high yield.

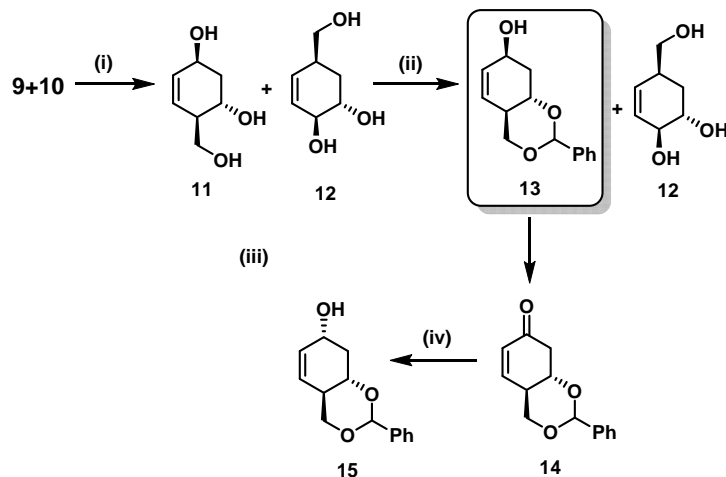
#### Scheme 1 Synthesis of the mixture of lactones **9** and **10**



**Reagents and conditions:** (i) vinyl acetate, 120 °C, 5h, 84% (ii)  $\text{H}_2\text{SO}_4$ , MeOH, 85 °C, 8-10 h, 95% (iii) Na, Liq  $\text{NH}_3$ , THF: EtOH 0.5h, 74%. (iv) BzCl, pyridine, 3h, 93% (v)  $\text{CH}_3\text{COOH}$  :  $\text{H}_2\text{O}$  (6:1), 120 °C, 4 h, 81% (vi) mCPBA, DCM, 0 °C, 5h, 92% ,**9:10** (70:30).

Mixture of triol compounds **11**&**12** obtained by reducing the mixture of lactones, which were further subjected for selective 1,3-diol protection to obtain **13** and **12** which are easily isolable. Allylic hydroxyl group of **13** was oxidized and reduced to get **15**.

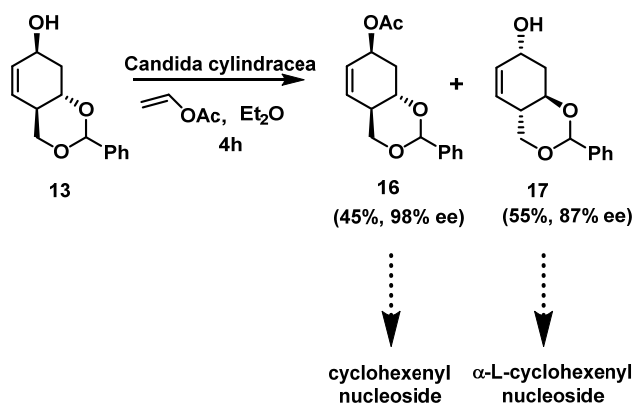
**Scheme 2** Synthesis of 1,3 diol protected key intermediate **13**



**Reagents and conditions:** (i)  $\text{LiAlH}_4$ , dry THF,  $-15\text{ }^\circ\text{C}$ , 2h, 72 % (ii)  $\text{PhCH}(\text{OMe})_2$ , PTSA, dry dioxane, rt, 24 h, 81% (iii)  $\text{CrO}_3$ , dry pyridine,  $\text{Ac}_2\text{O}$ , dry DCM, 2 h, 92% (iv)  $\text{NaBH}_4$ ,  $\text{CeCl}_3 \cdot 7\text{H}_2\text{O}$ , MeOH, 3h, 80%.

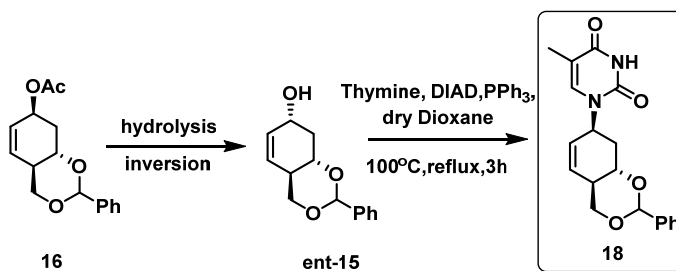
We have introduced an efficient method for enzymatic resolution with good yields as well as good enantiomeric excess. To resolve the enantiomers, compound ( $\pm$ )-**13** was subjected to acylation using vinyl acetate as donor and *Candida cylindracea* lipase (*CCL*) to get **16** in excellent yield and high enantiomeric purity.

**Scheme 3** Synthesis of cyclohexenyl thymine nucleoside



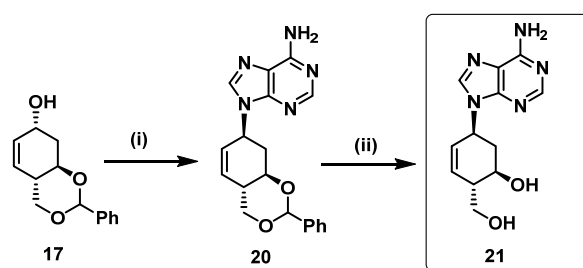
One enantiomer was used for the synthesis of cyclohexenyl nucleosides, another was used for  $\alpha$ -L-cyclohexenyl nucleosides. In compound **16** acetate group was hydrolyzed and inverted the chiral centre to get enantiopure **15**, further subjected for Mitsunobu reaction condition to introduce thymine nucleobase.

**Scheme 4** Synthesis of cyclohexenyl thymine nucleoside



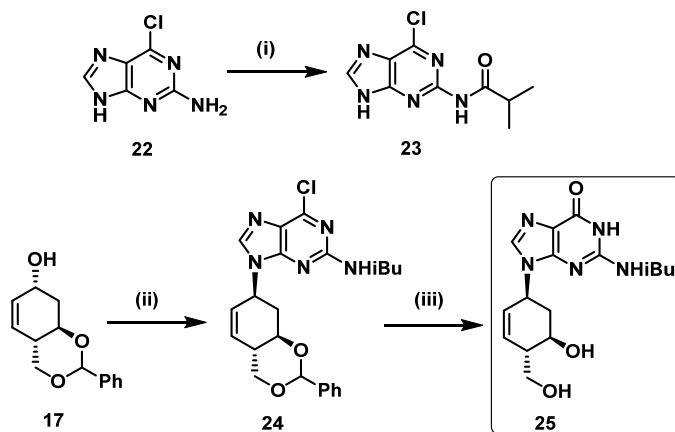
The pure enantiomer **17** was used for synthesis of  $\alpha$ -L-cyclohexenyl adenine and guanine nucleoside synthesis.

**Scheme 5** Synthesis of  $\alpha$ -L-cyclohexenyl adenine nucleoside



**Reagents and conditions:** (i) Adenine, DEAD, PPh<sub>3</sub>, dry THF, rt, overnight, 38 % (ii) 80% aq CH<sub>3</sub>COOH, 60 °C, 8 h, 65%.

**Scheme 6** Synthesis of  $\alpha$ -L-cyclohexenyl guanine nucleoside



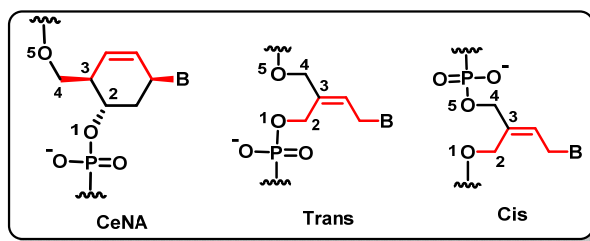
**Reagents and conditions:** (i) Isobutyric anhydride, dry N,N dimethyl acetamide, 150 °C, 2h, 72 % (ii) **23**, DIAD, PPh<sub>3</sub>, dry THF, overnight, 34% (iii) 80% aq CH<sub>3</sub>COOH, 60 °C, 8 h, 68%.

### Chapter 3: Open chain analogues of cyclohexenyl nucleic acids

#### Section 3A: Design, synthesis and biophysical evaluation of Open chain analogues of cyclohexenyl nucleic acids

The evolutionary chemistry with respect to nucleic acids suggested that simple acyclic nucleic acids might be preliminary nucleic acids, which ultimately have evolved as present day carriers of genetic information. To counter the entropic loss in acyclic nucleic acids, an attempt was made by introducing a double bond in the acyclic structure. Incorporation of these acyclic thymidine nucleoside mimics in oligomers was also found to be detrimental to the duplex stability similar to the other acyclic derivatives. We presume that the attachment of nucleobase directly to the double bond in this case may have conferred considerable rigidity, leading to reduced ability of the nucleobase to take part in specific W-C hydrogen bonding.

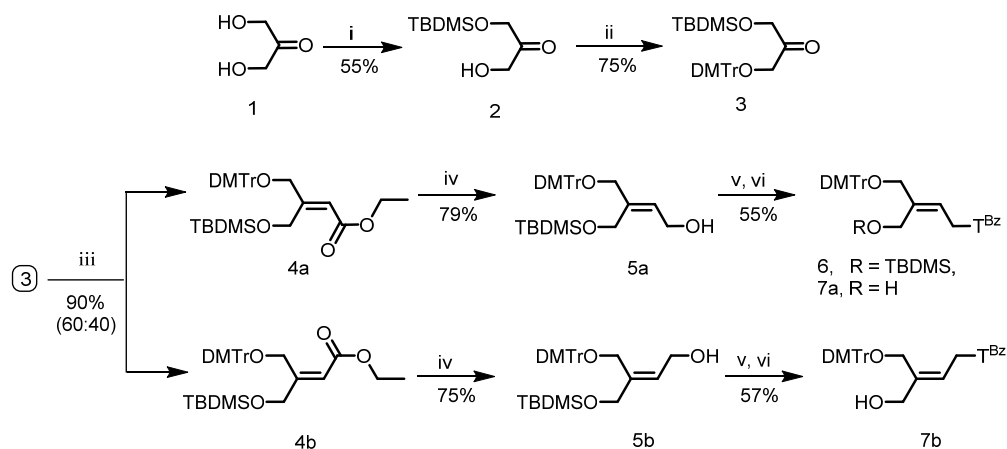
These modifications prompted us to visualize an open chain ene- nucleic acids in which the nucleobase attachment is to a planar double bonded structure through a methylene group, having same number of atoms in the backbone like natural sugar. This would also have a constraint of double bond unsaturation and act as an acyclic version of cyclohexene nucleic acid (Figure 3).



**Figure 3** CeNA and Proposed *cis*- and *trans*-open chain analogs of CeNA

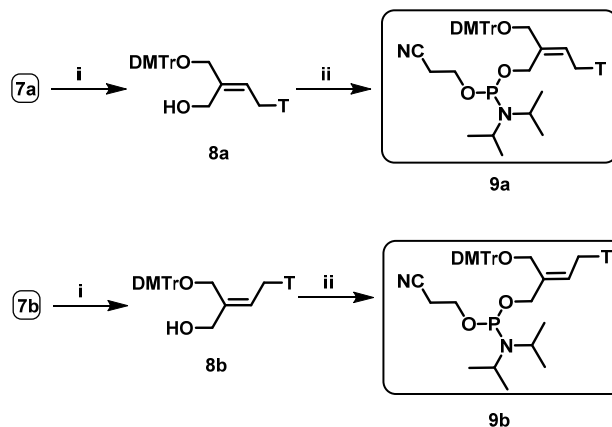
#### Synthesis of *cis* and *trans* thymidine monomers:

We started the synthesis with the mono TBDMS protection of 1,3 dihydroxy acetone. By using simple protections and further Wittig-reaction with two carbon ylide resulted the intermediates **4a**, **4b**, which were isolated carefully and characterized by NOE experiments.

Scheme 7 Synthesis of key intermediates **7a**, **7b**

**Reagents and conditions** (i) TBDMSCl, imidazole, dry DMF (ii) DMTrCl, pyridine, overnight (iii) ethylbromoacetate, PPh<sub>3</sub>, Toluene, reflux, 5h (iv) DIBAL-H, dry DCM, -78 °C (v) T-Bz, PPh<sub>3</sub>, DIAD, dry dioxane, overnight (vi) TBAF in 1M THF, THF, rt, 2h.

Reduction of ester group followed by thymine base introduction by using Mitsunobu reaction condition and TBDMS deprotection gave **7a**, **7b** (scheme 7). N<sup>3</sup> benzoyl group was deprotected and subjected for amidite reaction to obtain *cis* and *trans* thymidine nucleosides (scheme 8)

Scheme 8 Synthesis of *trans* and *cis* phosphoramidite monomersSynthesis of modified oligonucleotides, characterization and UV-*T<sub>m</sub>* studies

Modified oligonucleotides (Table 1) were synthesized on Bioautomation MM4 DNA synthesizer, using phenoxyacetyl (Pac) protected cyanoethyl phosphoramidites and modified amidite building blocks **9a**, **9b**.

**Table 1** Modified DNA sequences, their MALDI-TOF mass, UV- $T_m$  measurements<sup>a</sup>

Name	Sequence <sup>b</sup> 5'→3'	mass cal /obs	UV $T_m$ °C	
			DNA <sup>c</sup>	RNA <sup>d</sup>
DNA1	caccattgtcacactcca	5363/5367	63.5	62.7
DNA1-15T <sup>trans</sup>	caccattgtcacacT <sup>trans</sup> cca	5347/5342	60.2	-
DNA1-9T <sup>trans</sup>	caccattgT <sup>trans</sup> cacactcca	5347/5347	59.6	61
DNA1-15T <sup>cis</sup>	caccattgtcacacT <sup>cis</sup> cca	5347/5344	59.3	59.1
DNA1-9T <sup>cis</sup>	caccattgT <sup>cis</sup> cacactcca	5347/5344	60.8	59.6

<sup>a</sup>UV- $T_m$  values were measured by using 1 $\mu$ M sequences with 1 $\mu$ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is  $\pm 0.5$  °C). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5'tggagtgtgacaatgggtg was the complementary DNA sequence. <sup>d</sup>5'uggagugugacaauggug was the complementary RNA sequence.

### Enzymatic stability studies of modified DNA sequences

Considering nuclease resistance as an important factor, we examined the 3'-exonuclease sensitivity of unmodified homooligomer **t<sub>10</sub>** as well as *cis* and *trans* modified sequences **t<sub>8</sub>T<sup>cis</sup>t** and **t<sub>8</sub>T<sup>trans</sup>t** by using phosphodiesterase I from *Crotalus adamanteus* venom [snake venom phosphodiesterase (SVPD)].

**Table 2** Oligomers used for 3'-exonuclease degradation study

Name	Sequence 5'→3'	mass cal /obs
<b>t<sup>10</sup></b>	ttttttttt	2980/2977
<b>t<sub>8</sub>T<sup>cis</sup>t</b>	tttttttT <sup>cis</sup> t	2964/2961
<b>t<sub>8</sub>T<sup>trans</sup>t</b>	tttttttT <sup>trans</sup> t	2964/2961

The unmodified single strand DNA sequence **t<sub>10</sub>** did not show any 3'-exonuclease resistance. The sequence **t<sub>8</sub>T<sup>trans</sup>** was stable up to 10 min, which is comparatively stable

than unmodified  $t_{10}$ . The sequence  $t_8T^{cis}$  was observed to be stable up to 5h, which was clearly, indicated the very high stability under nuclease conditions (Figure 4).

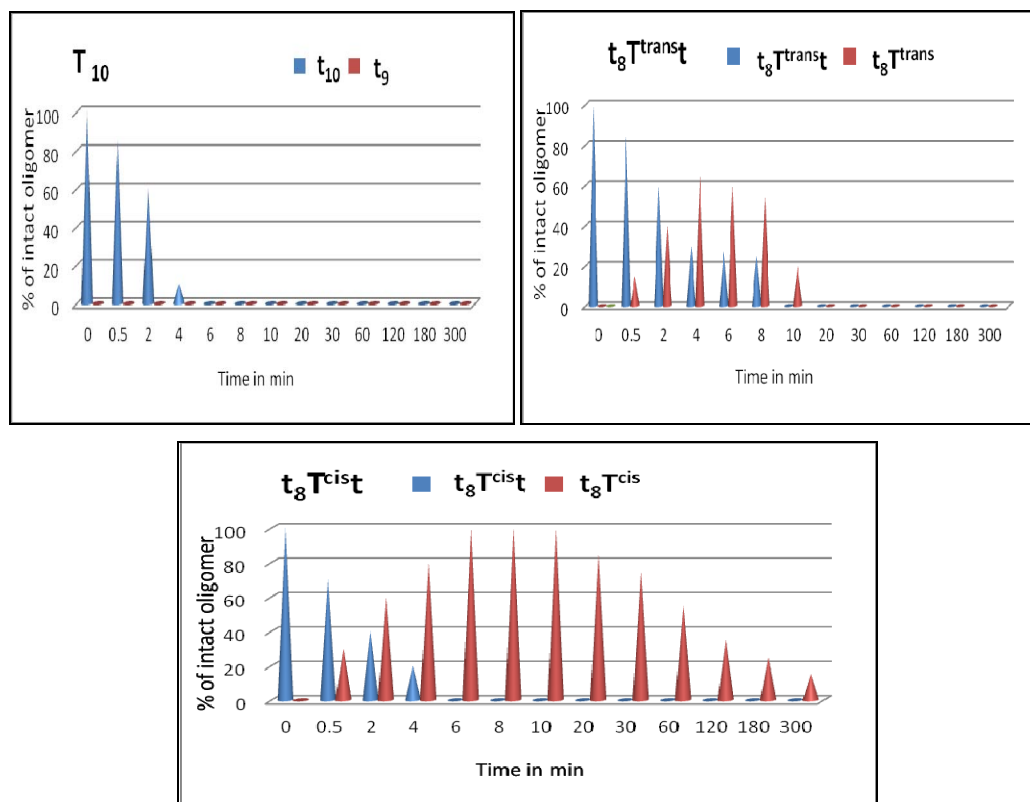


Figure 4 SVPD digestion of native DNA  $t_{10}$  and *cis*, *trans* modified  $t_{10}$  sequences

### Section 3B: Synthesis of *cis* and *trans* open chain analogue modified thrombin binding aptamer, its quadruplex formation

The **thrombin binding aptamer (TBA)** was discovered in 1992 by *in vitro* selection and found to inhibit fibrin- clot formation with high selectivity and affinity. NMR and X-ray structural study reveals that TBA forms an intramolecular, antiparallel G-quadruplex with chair like conformation. This G-quadruplex consists of two G-quartets connected by three edge wise loops (one central TGT loop and two TT loops). The aptamer interacts with two thrombin molecules, inactivating only one of them. Although several DNA/RNA aptamers are able to show interesting and promising pharmacological properties, rarely they could be used as therapeutic agents without

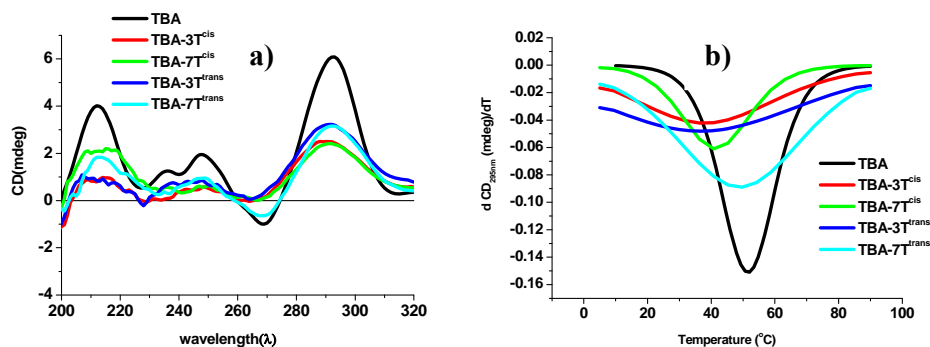


modifications due to less stability. To overcome these problems lot of attempts has been done via chemical and structural modifications.

The UNA analogues were used by Wengel and co-workers. These analogues found to be excellent for stabilizing loop structure in aptamers due to its ability to alleviate strain. We studied the flexibility parameter of our open chain-NA modification by introducing it in the loop region of TBA quadruplex.

**Table 3** Modified TBA sequences, their MALDI-TOF mass analysis and biophysical evaluation by CD- $T_m$  measurements

Name	Sequence <sup>a</sup> 5' → 3'	mass cal / obs	HPLC $t_R$ (min)	CD $T_m$ °C
TBA	ggttggtgtggttgg	4726/4730	9.8	49.5
TBA-3T <sup>cis</sup>	ggT <sup>cis</sup> tggtgtggttgg	4710/4709	10.0	38
TBA-7T <sup>cis</sup>	ggttggT <sup>cis</sup> gtggttgg	4710/4709	10.1	41.4
TBA-3T <sup>trans</sup>	ggT <sup>trans</sup> tggtgtggttgg	4710/4714	9.9	36.1
TBA-7T <sup>trans</sup>	ggttggT <sup>trans</sup> gtggttgg	4710/4708	10.3	43.7



**Figure 5** (a) CD spectra of oligomers TBA, TBA-3T<sup>cis</sup>, TBA-7T<sup>cis</sup>, TBA-3T<sup>trans</sup>, TBA-7T<sup>trans</sup> sequences of 5µM concentration in 10mM potassium phosphate buffer (pH 7.5) containing 100mM KCl at 5°C. (b) Temperature-dependent changes in CD amplitude at 295nm plotted against temperature, first derivative plots at strand concentration 5µM in 10mM potassium phosphate buffer (pH 7.5) containing 100mM KCl.

All four modified sequences showed maxima at 295nm which is characteristic CD signature for formation of stable antiparallel quadruplex (Figure 5a). The stability of the G-quadruplexes was followed by the change in the amplitude of the CD signal at 295nm with temperature. The CD melting results tells that T7 position modifications

are showing less destabilization of antiparallel quadruplex structures as compared to their corresponding T3 position modifications (Figure 5b).

### Duplex stability studies of modified TBA oligomers

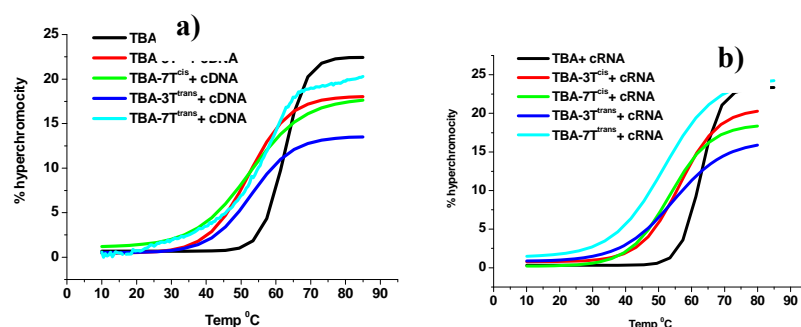
The binding affinity of 15mer TBA ONs TBA, TBA-3T<sup>cis</sup>, TBA-7T<sup>cis</sup>, TBA-3T<sup>trans</sup> and TBA-7T<sup>trans</sup> with complementary DNA and RNA was also investigated by measuring the melting temperatures (UV  $T_m$ ) of the duplexes (Table 4).

**Table 4** UV  $T_m$  (°C)<sup>a</sup> values of TBA and modified TBA: DNA/RNA duplexes

Name	Sequence <sup>b</sup> 5'→3'	mass cal /obs	CDT <sub>m</sub> °C	
			cDNA <sup>c</sup>	cRNA <sup>d</sup>
TBA	ggttggtgtggttgg	4726/4730	62	63
TBA-3T <sup>cis</sup>	ggT <sup>cis</sup> tggtgtggttgg	4710/4709	54	53
TBA-7T <sup>cis</sup>	ggttggT <sup>cis</sup> gtggttgg	4710/4709	53	54
TBA-3T <sup>trans</sup>	ggT <sup>trans</sup> tggtgtggttgg	4710/4714	53	54
TBA-7T <sup>trans</sup>	ggttggT <sup>trans</sup> gtggttgg	4710/4708	56	55

<sup>a</sup>UV- $T_m$  values were measured by using 1 $\mu$ M sequences with 1 $\mu$ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is  $\pm 0.5$  °C). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5'-ccaaccacccaacc was the complementary DNA sequence. <sup>d</sup>5'-ccaaccacccaacc was the complementary RNA sequence.

UV  $T_m$  results showed that all modified and unmodified sequences were forming stable duplexes with cDNA as well as cRNA. Independent to the nucleoside and position of modifications,  $T_m$  values were same for all sequences with cDNA, cRNA (Figure 6). These *cis* and *trans* modifications are destabilizing the duplexes compared to unmodified TBA.

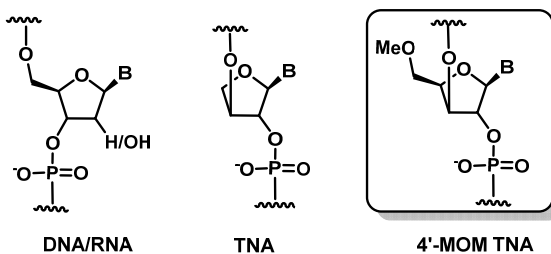


**Figure 6** (a) UV melting profiles of TBA and modified TBA sequences with cDNA, (b) UV melting profiles of TBA and modified TBA sequences with cRNA.

## Chapter 4: 4'-Methoxymethyl threose Nucleic acids

### Section 4A: Design, synthesis and biophysical evaluation of 4'-Methoxymethyl threose nucleic acids

$\alpha$ -L-threose nucleic acid (TNA), was the first synthetic genetic polymer discovered by Eschenmoser and co-workers. Natural five-carbon ribose sugar found in RNA was replaced with an unnatural four-carbon threose sugar in TNA. The successive nucleosidic units were connected through 2' and 3' vicinal phosphodiester linkages. Although the number of methods reported for synthesis of L-Threose in the literature, none of them seemed amenable to large scale preparation, either due to expensive starting materials, laborious workup, or poor yields. We designed a straight forward synthetic route with commercially available cheap starting material D-xylose, which will lead us to 4'-Methoxymethyl modified TNA nucleoside (Figure 7). The design of modified TNA was not only for simplifying the synthesis but also to study the hydration effect due to the methoxymethyl group.

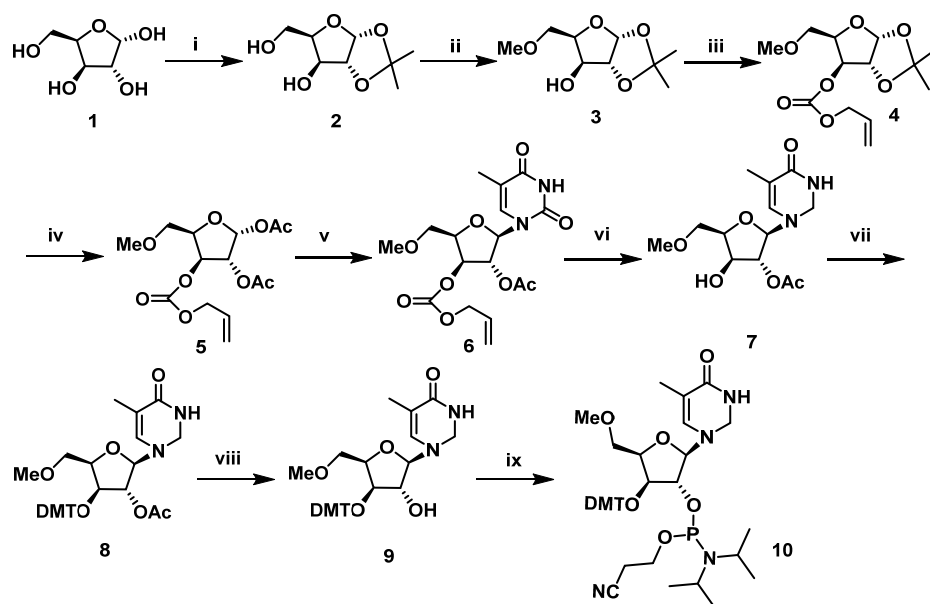


**Figure 7** Structures of natural DNA/RNA, TNA and proposed 4'-MOM TNA

D-xylose was converted to 1,2-*O*-isopropylidene-D-xylofuranose **2** using acetone, conc.H<sub>2</sub>SO<sub>4</sub> and Na<sub>2</sub>CO<sub>3</sub> in one pot reaction. Mono methylation of primary hydroxyl group in presence of methyl iodide and silver oxide yielded **3**. The secondary hydroxyl group was protected with allyloxycarbonyl group to give **4** in very good yield. The acetonide group in **4** was removed and converted into its diacetate **5** by treatment with AcOH and Ac<sub>2</sub>O in presence of catalytic amount of H<sub>2</sub>SO<sub>4</sub>. Compound **5** on treatment with BSA, thymine and TMSOTf under Vorbrüggen conditions afforded exclusively the  $\beta$ -anomer of thymine derivative **6**, in good yield. The alloc group was selectively cleaved using Pd(0) to get **7**. The free 3'- hydroxyl group was protected as

its DMT derivative using DMTr-Cl in dry DCM and 2,4,6-collidine used as base, to get **8**. Compound **8** on ammonolysis gave the free 2'-hydroxyl compound **9**. Phosphitylation of the free 2'-hydroxyl with *N,N*-diisopropylamino-2-cyanoethylphosphino-chloridite afforded the phosphoramidite monomer **10** (scheme 9).

**Scheme 9** Synthesis of 4'-MOM threose thymine nucleoside



**Reagents and conditions** (i) Acetone, conc H<sub>2</sub>SO<sub>4</sub>, Na<sub>2</sub>CO<sub>3</sub> (ii) MeI, Ag<sub>2</sub>O, dry ACN (iii) Alloc-Cl, dry Pyridine, dry DCM, rt, 3h (iv) AcOH: Ac<sub>2</sub>O: H<sub>2</sub>SO<sub>4</sub>(10:1:0.1), rt, overnight (v) Thymine, ACN, BSA, 70 °C, TMS-OTf, 0 °C, reflux, 3h (vi) PPh<sub>3</sub>, Pd(dba)<sub>2</sub>, Piperidine, DCM, rt, 15min (vii) DMTrCl, 2,4,6-collidine, DCM, rt, 24h (viii) 2-cyanoethyl-*N,N*-diisopropylchlorophosphine, DIPEA, dry DCM, rt, 1h.

### Synthesis of modified oligonucleotides, characterization, UV-melting studies

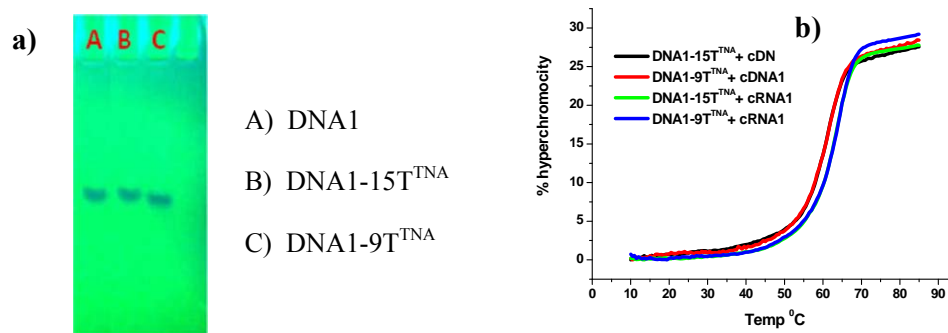
Modified monomer was incorporated in to 18mer DNA sequence and checked the duplex stability with cDNA and cRNA.

**Table 5** Modified DNA sequences, their MALDI-TOF mass analyses and biophysical evaluation by UV-*T<sub>m</sub>* measurements<sup>a</sup>

Name	Sequence <sup>b</sup> 5'→3'	mass cal /obs	UV <i>T<sub>m</sub></i> °C	
			DNA <sup>c</sup>	RNA <sup>d</sup>
DNA1	caccattgtcactcca	5363/5367	63.5	62.7
DNA1-15T <sup>TNA</sup>	caccattgtcacacT <sup>TNA</sup> cca	5393/5393	59.8	61.9

DNA1-9T<sup>TNA</sup> caccattgT<sup>TNA</sup>cacactcca 5393/5387 60 62.2

<sup>a</sup>UV- $T_m$  values were measured by using 1 $\mu$ M sequences with 1 $\mu$ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is  $\pm 0.5$  °C). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5' tggagtgtgacaatggtg was the complementary DNA sequence. <sup>d</sup>5' uggagugugacaauuggug was the complementary RNA sequence.

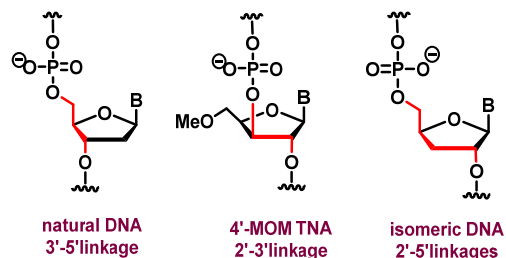


**Figure 8** (a) Gel pictures of purified sequences and (b) UV melting profiles of modified DNA sequences with cDNA/cRNA

It was observed that modified sequences are forming the stable duplex structures with both complementary DNA as well as complementary RNA. Modified TNA:RNA complexes are more stable than modified TNA:DNA complexes (Figure 8).

#### **Section 4B: Synthesis of 4'- MOM TNA modified thrombin binding aptamer, its quadruplex formation and application as a thrombin inhibitor**

Aptamers are rarely used as therapeutic agents due to less stability towards hydrolytic enzymes. Lot of attempts have been done to overcome these problems by introducing chemical and structural modifications. We considered the synthesis of UNA modified TBA and incorporated the 4'-MOM TNA at T7 and T9 positions to study the effect on stability of the quadruplex. TNA is having the 2'-3' backbone and for comparison we also synthesized the TBA sequences with 2'-5' modification at 7<sup>th</sup> and 9<sup>th</sup> position and we studied the tetraplex stability, anticoagulation activity as well as enzymatic stability.



**Figure 9** Different types of backbone linkages

Modified TBA sequences were synthesized on automated Bioautomation MM-4 DNA synthesizer and the modified monomers incorporated at 7<sup>th</sup> and 9<sup>th</sup> positions. All the sequences purified by HPLC, and masses were confirmed by MALDI-TOF().

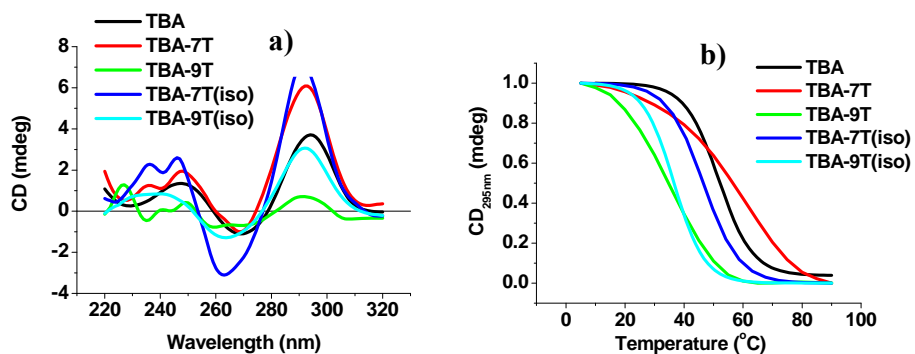
### CD spectroscopy and $T_m$ measurement of the TBA sequences in presence of $K^+$ ion

The G-quadruplex formation of the modified sequences was studied by CD spectroscopy in the presence of monovalent cation  $K^+$  and their stability was determined as a function of temperature dependent change in CD amplitude at 295nm.

**Table 6** Modified DNA sequences, their MALDI-TOF mass analyses and biophysical evaluation by CD- $T_m$  measurements

Name	Sequence <sup>a</sup> 5' → 3'	mass cal /obs	CD $T_m$ °C( $K^+$ )
TBA	ggttggtgtggttgg	4726/4730	49.5
TBA-7T	ggttggT <sup>TNA</sup> gtggttgg	4756/4754	59.7
TBA-9T	ggttggtgT <sup>TNA</sup> ggttgg	4756/4754	34.3
TBA-7T( <i>iso</i> )	ggttggT <sup>iso</sup> gtggttgg	4726/4729	46.5
TBA-9T( <i>iso</i> )	ggttggtgT <sup>iso</sup> ggttgg	4726/4731	36.5

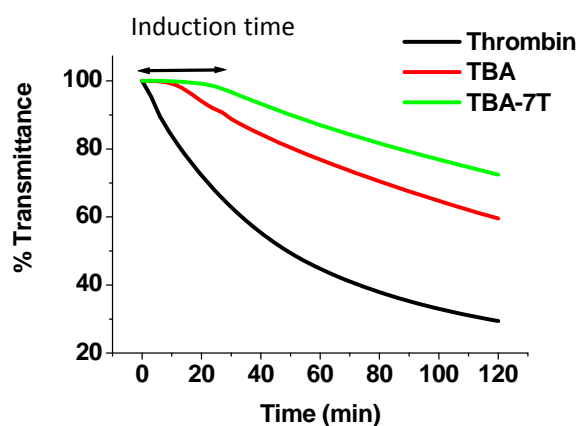
<sup>a</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site



**Figure 10** (a) CD spectra of oligomers **TBA**, **TBA-7T**, **TBA-9T**, **TBA-7T(iso)**, **TBA-9T(iso)** of  $5\mu\text{M}$  concentration in 10mM potassium phosphate buffer (pH 7.5) containing 100mM KCl at  $5^\circ\text{C}$ . (b) CD- $T_m$  in 10mM Na-phosphate buffer (pH 7.5) containing 100mM KCl.

### Anti-thrombin activity measurements

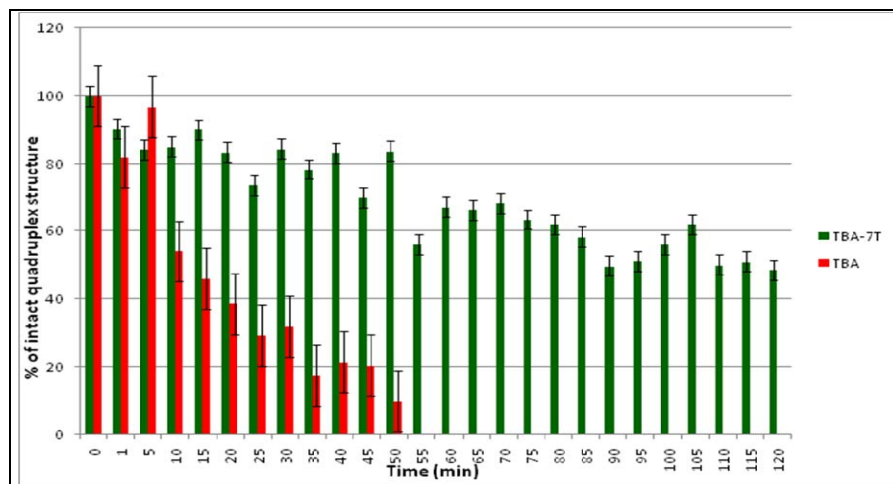
The anti-thrombin activity of the aptamers on thrombin-catalyzed conversion of fibrinogen to fibrin (clotting) was investigated by measuring the percent transmittance with time. **TBA** slowed down the coagulation with an increased induction time ( $t_i$  as coagulation parameter), confirming its reported inhibitory activity. The induction time for the **TBA-7T** was higher than for **TBA**, this high induction time could give a large window to reduce the concentration of **TBA** in acceptable therapeutic range.



**Figure 11** Antithrombin activity measured by % transmittance at 450nm in the presence of **TBA** and **TBA-7T** and % transmittance Vs wavelength plots at different time-points of the study.  $\leftrightarrow$  indicates induction time as coagulation parameter ( $t_i$ ).

### Stability of quadruplex structure of aptamers to SVPD

We studied the stability of **TBA** and **TBA-7T** quadruplex structures against **SVPD** enzyme. The stability of **TBA-7T** was found to be very high compared to the control **TBA**. The reaction was monitored by change in CD amplitude at 295nm.



**Figure 12** Quadruplex stability of the aptamers **TBA** and **TBA-7T** ( $7.5\mu\text{M}$ ) towards Snake venom phosphodiesterase (**SVPD**) enzyme ( $2.5\text{ mg/mL}$ ).



## CHAPTER 1

# Introduction to nucleic acids

## 1 Introduction to Nucleic Acids

### 1.1 Primary structures of DNA and RNA

Nucleic acids (DNA, RNA) are the most important of all biopolymers. DNA carries the hereditary information from generation to generation and RNA functions in converting genetic information from gene into amino acid sequences of proteins. J. D. Watson and F. H. Crick discovered double stranded structure of DNA (Figure 1) in 1953<sup>1</sup>.

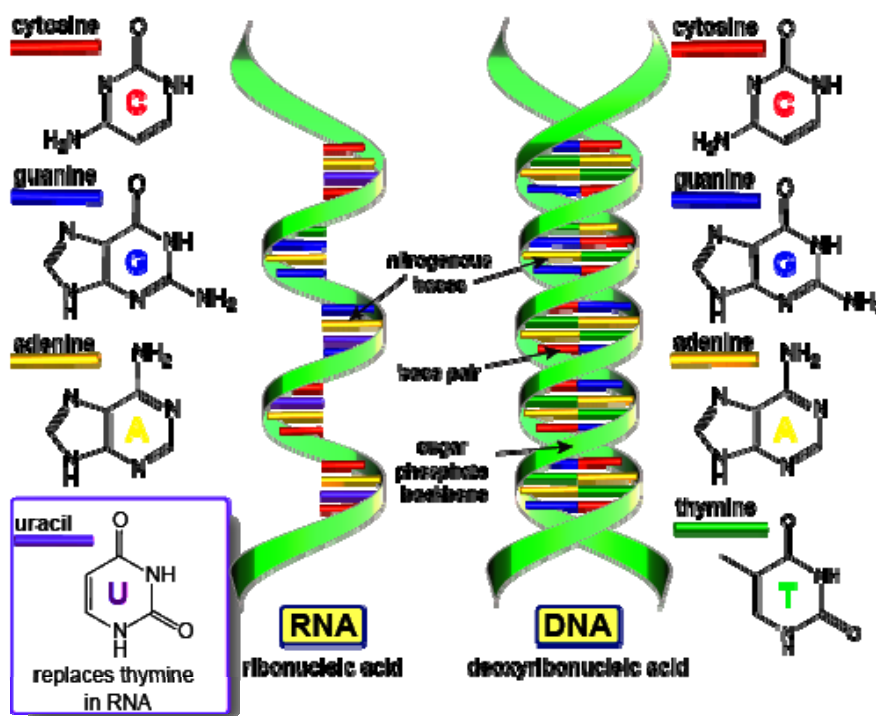


Figure 1: Models of RNA and DNA (Image from: Wikipedia)

Single stranded nucleic acids are composed of repeating smaller units, called nucleotides. Each nucleotide is comprised of three components: a nitrogenous heterocyclic base, which is either purine or pyrimidine, pentose sugar and a phosphate group. DNA and RNA differ in the structure of the sugar in their nucleotide. DNA contains deoxyribose, whereas RNA is of ribose sugars. The nitrogenous bases adenine, cytosine and guanine are found in both DNA and RNA, thymine occurs only in DNA, while uracil occurs in RNA. Each base is connected to sugar *via*  $\beta$ -glycosyl linkage. Adjacent nucleoside units (base + sugar) are connected *via* O3' of one nucleoside to O5' atom of other nucleoside through a phosphodiester linkage.

## 1.2 Base pairing via Hydrogen bonding

Mutual recognition of A by T and C by G use specific hydrogen bonds to establish the fidelity of DNA transcription and translation. The N-H groups of the bases are potent hydrogen donors, while  $sp^2$ -hybridized electron pairs on the oxygens of the base C=O groups and the ring nitrogens are hydrogen bond acceptors. A-T pair forms two hydrogen bonds and C-G forms three hydrogen bonds. RNA usually exists in single stranded form but may fold into secondary and tertiary structures through the formation of specific base pairings (A-U, C-G). Though the Watson-Crick base pairing (Figure 2) is dominant between the nucleobases, other significant pairings are Hoogsteen<sup>2a</sup> (Figure 3) and Wobble base pairs<sup>2b</sup> (Figure 4).

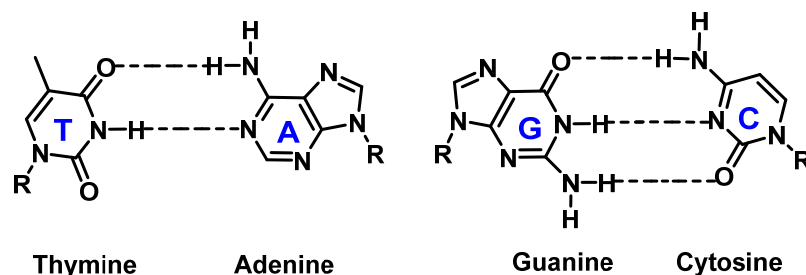


Figure 2: Watson and Crick base pair of A:T and G:C

Hoogsteen A:T base pair applies N7 position of the purine base (as a hydrogen bond acceptor) and C6 amino group (as a donor), which bind the Watson-Crick (N3-O4) face of the pyrimidine base (Figure 3). Hoogsteen pairs have quite different properties from Watson-Crick base pairs. The angle between the two glycosylic bonds (ca.  $80^\circ$  in the A:T pair) is larger and the C1'-C1' distance (ca. 8.6 Å) is smaller than in the regular geometry. In some cases, called reversed Hoogsteen base pairs, one base is rotated  $180^\circ$  with respect to the other. Hoogsteen base-pairing allows sequence specific binding of pyrimidine third strands in the major groove of Watson-Crick purine: pyrimidine duplexes form triple-helical structures (poly(dA):2poly(dT)) and (poly(rG):2poly(rC)).

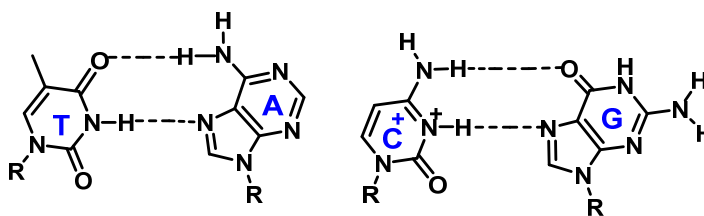
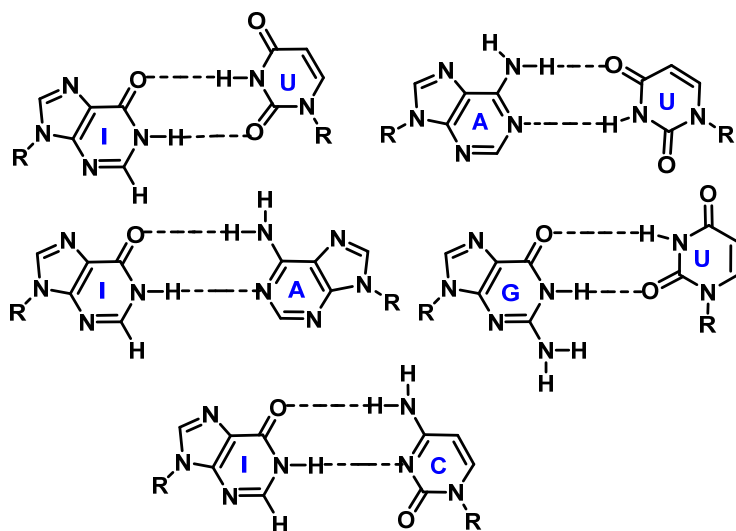


Figure 3: Hoogsteen base pair of A:T and G:C

## Chapter 1

In the wobble base pairing (Figure 4), a single purine base is able to recognize pyrimidines (e.g. G:U, where U= uracil) and have importance in the interaction of messenger RNA (m-RNA) with transfer RNA (t-RNA) on the ribosome during protein synthesis (codon-anticodon interactions). Several mismatched base pairs and anomalous hydrogen bonding patterns have been seen in X-ray studies of synthetic oligodeoxynucleotides.



**Figure 4:** Wobble base pairing of Ionosine and Uracil.

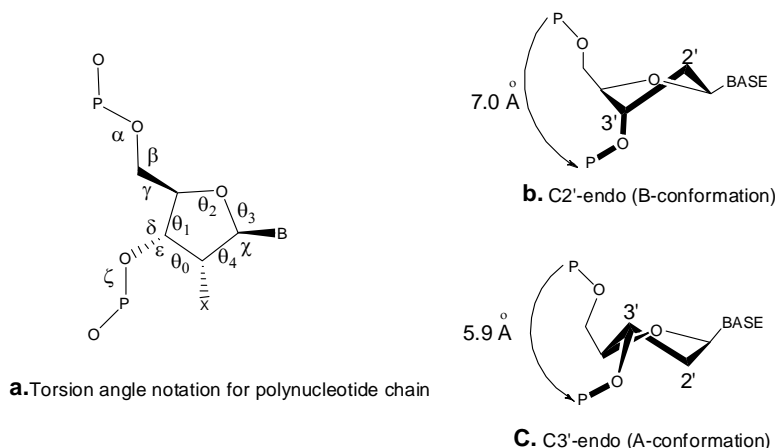
### 1.3 Sugar Puckering

The nucleic acids conformations are stabilized by several forces. Besides base pairing, the conformational features are constrained by pseudo torsional angles for rotation around each bond. The nucleotides conformational structure of are defined by the torsional angles  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  and  $\zeta$  in the phosphate backbone,  $\theta_0$  to  $\theta_4$ , in the furanose ring, and  $\chi$  for the glycosidic bond (Figure 5a).

In nucleic acids the pentose sugar ring is inherently nonplaner. This nonplanarity is termed as puckering. The ring puckering arises in order to minimize non-bonded interactions between substituents. This ‘puckering’ is described by identifying the major displacement of the carbons  $C-2'$  and  $C-3'$  from the median plane of  $C1'-O4'-C4'$ . Thus, if the *endo*-displacement of  $C-2'$  is greater than the *exo*-displacement of  $C-3'$ , the conformation is called  $C2'$ -*endo* and so on for other atoms of the ring (Figure 5, b and c). The *endo*-face of the furanose is on the same side as  $C5'$  and the base; the *exo*-face is on the opposite face to the base. The sugar puckers

## Chapter 1

are located in the north (N) and south (S) domains of the pseudorotation cycle of the furanose ring.<sup>3</sup> In solution, N and S conformations are in rapid equilibrium and are separated by low energy barrier. The average position of the equilibrium is influenced by several factors such as (i) the preference of the electronegative substituents at C2' and C3' for axial orientation, (ii) the orientation of the base, and (iii) the formation of an intra-strand hydrogen-bond from O2' in one RNA residue to O4' in the next, which favors C3'-endo-pucker.



**Figure 5:** (a) Torsion angle notations for polynucleotide chain (b) C2'-endo sugar and (c) C3'-endo sugar conformations

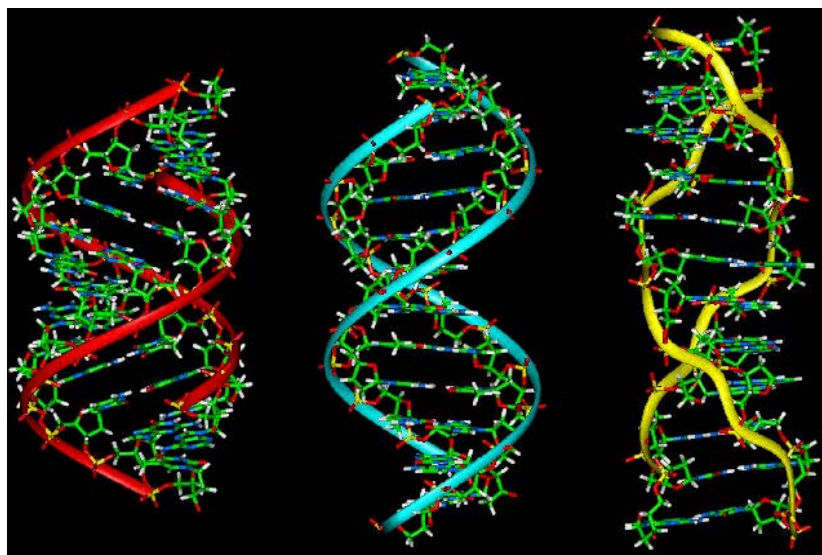
### 1.4 Structures of Nucleic acids

#### 1.4A The duplex structure

DNA can adopt mainly three conformations in biological system, right handed A-form, right handed B-form and left handed Z-form (Figure 6). The preferred structure of DNA molecule depends on both nucleotide sequences, the solvent and salt conditions. DNA mainly adopts B-form conformation under physiological conditions. The major groove is wider than the minor groove in DNA and many sequence specific proteins interact in the major groove, where the bases are perpendicular to the helical axis. The "B" form described by James D. Watson and Francis Crick<sup>4</sup>. On the other hand, A-DNA is favored at relative high salt concentration, especially for GC rich sequences it is thicker right-handed duplex with a shorter distance between the base pairs. A-DNA has 11 base pairs per helical turn, base pairs are tilted to about 20°, with respect to the helical axis, the grooves are not

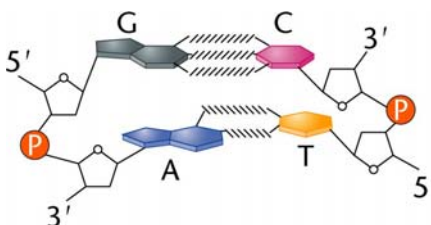
## Chapter 1

as deep as those in B-DNA, the sugar pucker is C3' *endo* compared to C2' *endo* for B-DNA, and the base pairs are shifted to the helix periphery which creates a 9 Å hole in the helix center. In both A and B forms of DNA Watson-Crick base pairing is maintained by *anti* glycosidic conformation of the nucleobases.



**Figure 6:** A, B and Z DNA

Z-DNA is one of the forms of DNA, with the two strands coiling in left-handed helices and a pronounced zig-zag pattern in the phosphodiester backbone<sup>5</sup>. Z-DNA is stabilized by high salt concentrations or polyvalent cations. When the DNA strand is an alternating GCGCGC such as purine-pyrimidine sequence, Z-DNA can form. It has the sugar in the C3' *endo* conformation (like A-form nucleic acid, and in contrast to B-form DNA) and guanine base is in the *syn* conformation. Z-DNA forms excellent crystals. The two single strands are in opposite directions, 5'-end of one strand is always laid against the 3'-end of the other strand of DNA which is called antiparallel orientation (Figure 7).

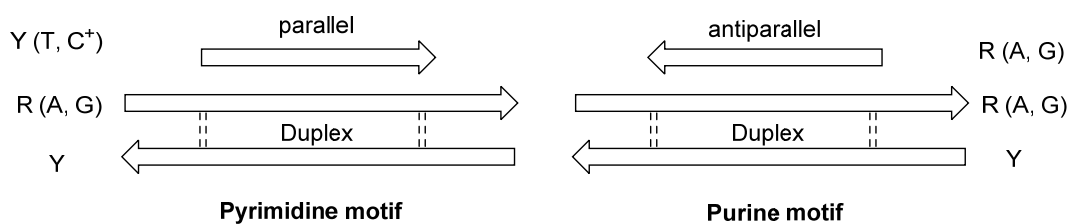


**Figure 7:** Antiparallel orientation of DNA (<http://bio3400.nicerweb.com>)

The chemical structure of RNA is very similar to that of DNA, but differs functionally. The hydroxyl group present at the 2' position of the ribose sugar in RNA but not in DNA, due to this functional group RNA adopt the A-form geometry rather than the B-form. 2'-hydroxyl group is capable of intra molecular nucleophilic attack on adjacent phosphate. One important factor that enables RNA's diverse functionalities is that RNA is transcribed as single strand molecule, which in turn, can fold into secondary and tertiary structure such as stem and bubble. An example of folded RNA structure is *t*-RNA, which is the key RNA involved in the translation of genetic information from *m*-RNA to proteins.

### 1.4B Triplex-Forming Oligonucleotides (TFO)

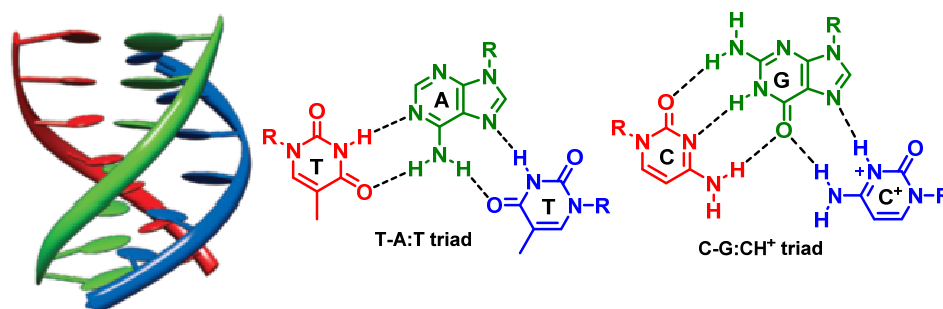
Triple-helical nucleic acid structures are known since 1957.<sup>6</sup> A purine base can form one on Watson-Crick face hydrogen bond and one on the Hoogsteen face hydrogen bond. Thus, in duplex DNA, a purine stretch presents sites in the major groove for Hoogsteen complexation by a third strand.<sup>7</sup> In 1991 it was realized by two different research groups that the single-stranded DNA can also serve as a target for triple helix formation: a purine stretch can be bound on two sides by a molecule carrying both a Watson-Crick complementary domain and a Hoogsteen complementary domain.



**Figure 8** Pyrimidine and purine motifs of triplex formation.

The triplexes formed with synthetic oligonucleotides, remained an obscure part of DNA chemistry until 1987 when it was realized that they offered a means for designing sequence specific DNA targeting agents.<sup>7b</sup> DNA triple helix formation results from the major groove binding of a third strand that is either pyrimidine (Y)- or purine-rich (R), in parallel (p) or antiparallel (ap) orientation respectively to the central purine strand as shown in Figure 8.<sup>8</sup> A purine-rich third strand binds in antiparallel orientation to the central strand, while a pyrimidine-rich strand does so in

a parallel orientation. The specificity in triplex formation is derived from Hoogsteen (HG) hydrogen bonding. Thus, T recognizes A of A-T Watson-Crick base pair to form TpAT triad and protonated C binds to G of G-C base pair to give CH<sup>+</sup> pGC in the pyrimidines motif (Figure 9). Similarly in the purine motif, the third strand A binds to A of A-T, leading to a AapAT triad, while G binds to G of G-C base pair forming GapGC triad in the reverse Hoogsteen mode.<sup>9</sup>



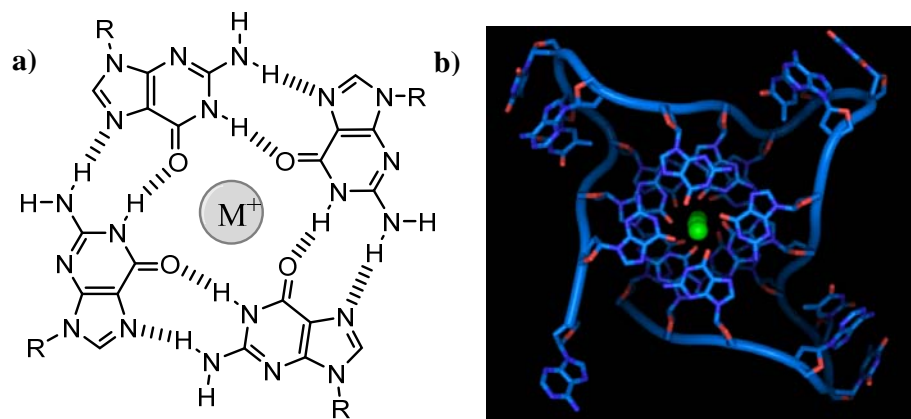
**Figure 9:** Schematic illustration of the parallel triplex and the base pairings involved (Figure adopted from ref.10).

For a given polypurine:polypyrimidine sequence of DNA it is possible to design therapeutic TFO that will specifically bind to it and thereby inhibit the gene expression. Different aspects of triple-stranded structures have been discussed in several reviews.<sup>11</sup> The formation and stabilization of triplexes depends on different types of interactions like electrostatic forces, stacking and hydrophobicity contributions, Hoogsteen hydrogen bonds, and hydration forces. The triplex structure of single DNA molecule was hypothesized as an intermediate in the folding process of some quadruplex forming sequences.<sup>12</sup>

### 1.4C Quadruplex structure – The ‘G-quadruplex’

G-rich DNA and RNA sequences having the ability to self-associate and form novel higher ordered structures called G-quadruplexes. These are the four-stranded structures formed from guanine-rich sequences, held together by hydrogen-bonding between the Watson-Crick face of each guanine with the Hoogsteen face of an adjacent guanine, creating a cyclic arrangement of four guanines in a G-tetrad (Figure 10a).



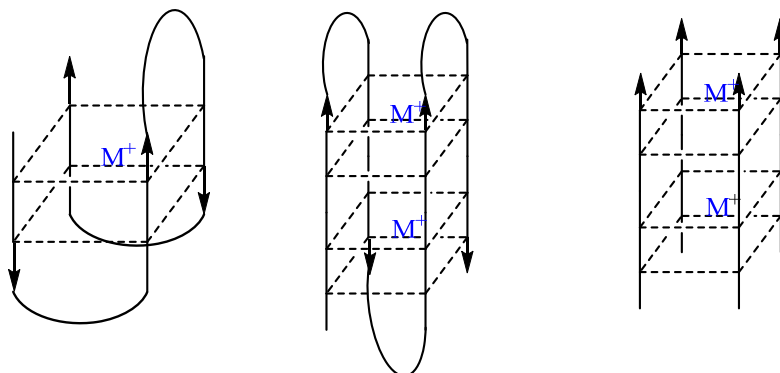


**Figure 10:** (a) The G-tetrad, (b) The G-quadruplex stacked structure arising from G-rich sequences. (<http://www.molecularstation.com/molecular-biology-images>)

These tetrads are stacked in a right-handed helical motif<sup>13</sup> with a helical twist of 30° and a diameter of 25Å, and are stabilized by monovalent cations co-ordinated to the O6 oxygen atoms of the guanines, and sandwiched between the base-stacks. These cations neutralize the electrostatic repulsion between the guanine O6 atoms and thus stabilize the overall structure (Figure10b).

#### 1.4D Structural polymorphism in G-quadruplexes

G-quartets can be formed from the association of one or more G-rich strands of DNA, (Figure 11) i.e., they may be intra- or inter-molecular. Based on loop length, strand polarity and geometry, the presence of metal ions, the oligonucleotide concentration in the solution, these can display a wide variety of topologies etc.<sup>14</sup>



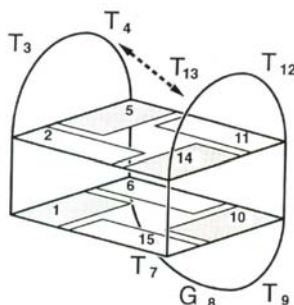
**Figure 11:** Schematic representation of G-quartets, antiparallel, unimolecular and multimolecular parallel DNA G-quadruplex.

The central charged cavity can accommodate a variety of different cations of differing radii that strongly influences not only the final folded topology of the quadruplex, but

also its stability. Potassium is generally found to be optimal and leads to increased G-quadruplex stability. The phosphate backbone generates four grooves that accommodate well-defined networks of ordered water molecules. In RNA and DNA G-quadruplexes, the G-tetrad stacks remain planar with a similar rise and twist, but the direction of the phosphodiester backbone may vary. Unlike in DNA/RNA duplexes, where an antiparallel strand orientation is predominant, in G-quadruplexes, the strands may exist in any combination of parallel and antiparallel orientations.

### 1.4E Thrombin-binding aptamer

The thrombin-binding aptamer is a 15mer DNA oligonucleotide sequence, discovered in 1992<sup>15</sup> by *in vitro* selection and found to inhibit fibrin clotting formation by binding to the thrombin protein with high selectivity and affinity. NMR and X-ray structural studies<sup>16</sup> showed this sequence [d(GGTTGGTGTGGTTGG)] to be an intramolecular antiparallel G-quadruplex with two G-quartets stacked on each other and linked by edge-wise two TT loops and one TGT loop (Figure 12).



**Figure 12:** Thrombin binding aptamer. (Schultze, P., Macaya, R. F., Feigon, J., *J. Mol. Biol.* 1994, 235, 1532)

The quadruplex forming, Thrombin binding aptamer TBA has been innovatively used as a DNA based nano machine that can cyclically bind and release thrombin. Dittermer *et al*<sup>17</sup> have used DNA-fuelled system based on quadruplex-duplex transition to control the the human blood-clotting. When the TBA folded in a quadruplex structure, it strongly binds to thrombin; switching between a quadruplex and a duplex conformation leads to trapping and releasing of thrombin, respectively.

## 1.5 Oligonucleotides as Therapeutic Agents

It is necessary to have a structural knowledge of binding site of the target and the binding forces to design a small organic molecule as a drug against traditional drug which targets protein. The drug discovery process has limitations because of our understanding of protein folding is incomplete. In contrast, the nucleotide sequence in RNA and DNA is universal and the understanding of their structure is much better, nucleic acid targets are very appealing for drug designing.

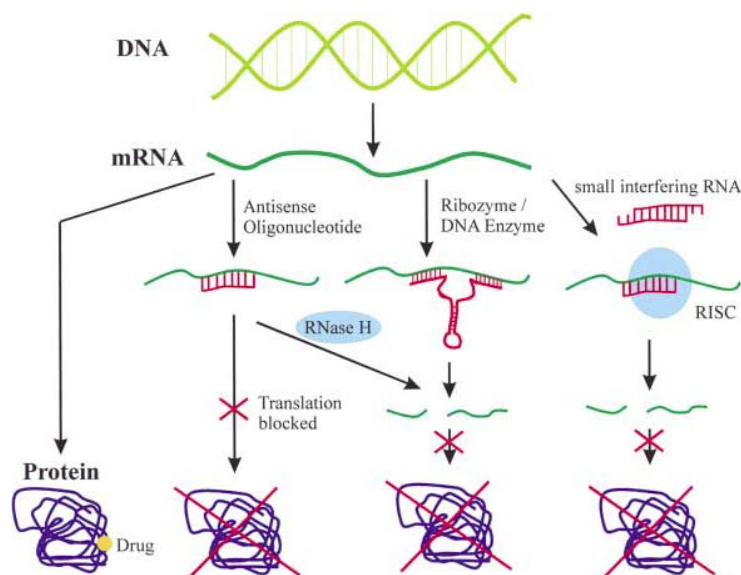
### 1.5.1 Antisense technology

The process of transformation of genetic information from DNA to RNA called transcription and RNA to protein is called translation. Synthesis of RNA is initiated by transcription process which first synthesizes precursor-mRNA. The pre-messenger RNA is composed of exons and introns (intervening sequences). However, introns are not translated into the amino-acid chain of a protein. In order to produce a functional protein these introns are removed from the pre-mRNA by splicing and the coding sequences (exons, carrying the genetic information) are fused together and form mature messenger RNA. The mRNA is then transported to the cytoplasm where it undergoes 'translation' and with the help of other RNA like tRNA (transfer RNA) and rRNA (ribosomal RNA) proteins are produced.

Conventional drugs bind to viral proteins and thereby modulate their function. Antisense technology is an alternative treatment for these disorders. The idea of this technology is inhibition of mRNA translation to proteins by using ssONs (DNA or RNA) sequence complementary to their target mRNA. Complementary sequence binds to mRNA in an anti-parallel, sequence specific manner *via* Watson-Crick hydrogen bonding. The potential of ONs to act as antisense agents that inhibit viral replication in cell culture was discovered by Zamecnik and Stephenson in 1978.<sup>18</sup> Vitravene is the first antisense drug, has been approved for the treatment of patients with cytomegalovirus- induced retinitis.<sup>19</sup> Several AS-ONs have entered phase I–III clinical trials as anticancer agents.

### 1.5.2 Disruptive antisense approach

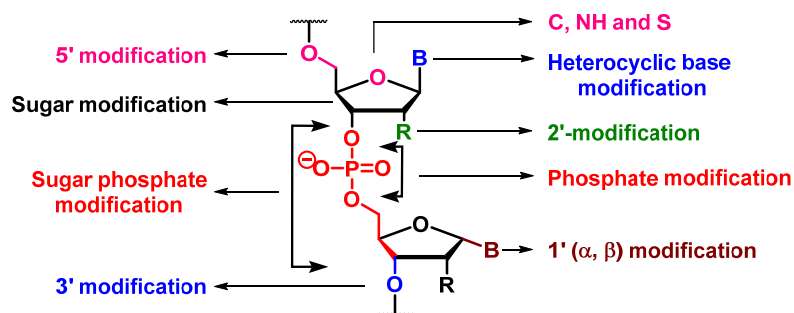
Antisense oligonucleotides can do the down-regulation of the disease-causing functional protein synthesis by disrupting the translation processes at mRNA level. This 'translational arrest' can be done via steric blocking of ribosomal machinery, degradation of the mRNA by activation of RNA cleaving enzymes such as RNase H or RNase L, ribozyme activation and/or RNA interference induced by small interfering RNA molecules like siRNA/miRNA (Figure 13).<sup>20</sup>



**Figure 13:** Various antisense mechanisms for inhibition/controlling of gene expression

### 1.6 Promising chemical modifications of nucleic acids

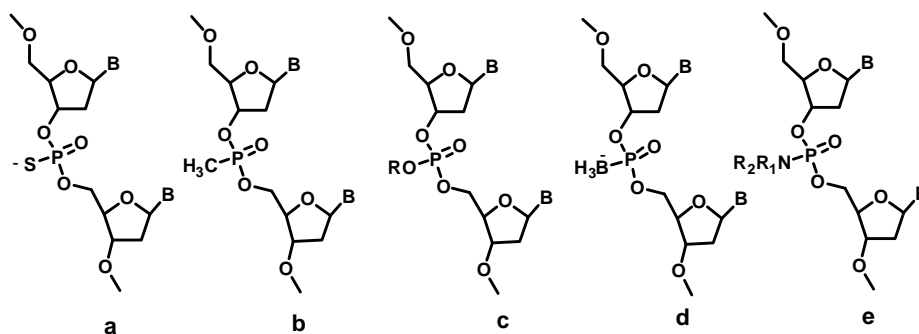
Antisense oligonucleotides should be stable to nucleases and efficiently pass through the cell membrane to the target site without degradation. Natural nucleic acids DNA/RNA are not good antisense drug candidates, because they are rapidly degraded by extra and intracellular nucleases. A vast number of chemical modifications done to improve the potency of the oligonucleotides, which have been described in the recent review articles. The general sites of modification are deoxyribo- or ribo- sugars, phosphate backbones, heterocyclic bases or complete replacement of the sugar phosphate backbone (Figure 14).



**Figure 14:** Schematic showing the extensive chemical modifications of DNA/RNA

### 1.6.1 Phosphate linkage modifications (First generation AONs)

The first generation modifications mainly focused on phosphate linkages. The most widely used backbone modification is a phosphorothioate linkage<sup>21</sup> (PS, Figure 15a), where a non-bridging oxygen on the phosphate linkage is replaced with a sulfur atom. Phosphorothioate-containing antisense oligonucleotides are able to activate RNase-H for target mRNA cleavage. This substitution has only a relatively small effect on the oligonucleotide structure (binding affinity is slightly reduced), but this disadvantage is outweighed by a greater resistance to enzymatic hydrolysis. Vitravene is the only antisense agent approved by FDA so far is based on PS-oligos. Other examples are methyl phosphonates<sup>22</sup> (Figure 15b), phosphotriesters<sup>23</sup> (Figure 15c), boranophosphonates<sup>24</sup> (figure 15d) and phosphoramidates<sup>25</sup> (Figure 15e) replacing the anionic phosphate diester linkages either with anionic or neutral linkages.



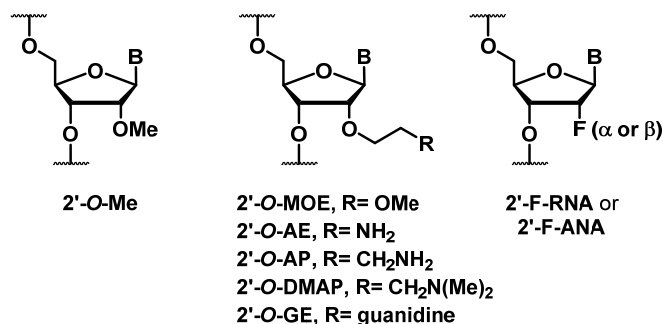
**Figure 15:** First generation modified antisense oligonucleotide

Additional chirality introduced at phosphorus centre because of these modifications. Thus the oligomers are complex mixture of diastereomers and the conformational heterogeneity leads to lowering of the melting temperature. As a consequence, the

search for a stereo controlled synthesis has been undertaken.<sup>26</sup> Synthetic diastereomerically pure phosphorothioate-containing RNA oligomers with the Rp configuration showed higher affinity for complementary RNA than the natural oligomer, whereas Sp oligomers had similar or lower affinity.

### 1.6.2 2'-Modifications of Sugar (Second generation antisense oligonucleotides)

Natural DNA is in C2' *-endo* or southern (S) sugar conformation because of the *gauche effect* between O3' and O4'. RNA is in C3' *-endo* sugar conformation (N-type) conformation which render stability of duplex as the base is in axial orientation, that can form stronger hydrogen bonding as well as stacking. The order of the stability of duplexes was DNA:DNA < DNA:RNA < RNA:RNA. Conformation of the sugar can be locked or frozen by modifying the sugar with electronegative substituent at the C2' and C3' positions to yield oligonucleotides that bind target DNA/RNA with high affinity. In order to improve the RNA binding capacity of AON, second generation modifications focused mimicking RNA or C3' *-endo*-like structures. Electronegative substituents such as fluorine and oxygen shift the ribose conformational equilibrium towards the C3' *-endo* pucker.<sup>27</sup> The 2'-fluoro substitution in both ribo-( $\alpha$ -F) and arabino-( $\beta$ -F) configuration (Figure 16) led to stabilization of duplexes with target RNA, both operating in different mechanism. The observed duplex stability is attributed to the preferential C3' *-endo* sugar puckering of ribofluoro-modified (stabilized by both *anomeric effect* and *gauche effect*)<sup>28</sup> and the favourable inter-residual pseudo-hydrogen bond (2'F...purine H8) in the arabinofluoro-modified nucleic acids.<sup>29</sup>

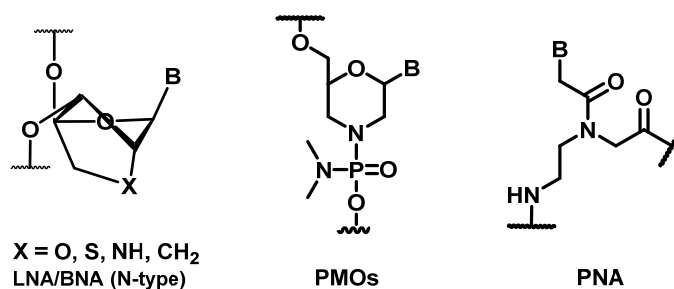


**Figure 16:** Few representative 2'-sugar modified oligonucleotides

The 2'-*O*-Me modification improved thermal stability of an oligonucleotide hybridized to a complementary RNA but did not confer the necessary metabolic stability to antisense oligonucleotides. In combination with phosphorothioate backbone modification, oligonucleotides with these modifications are resistant to metabolic degradation. Oligonucleotides with 2'-*O*-[2-(methoxy)ethyl] modification (2'-*O*-MOE, Figure 16) showed equal or higher binding affinity than 2'-*O*-Me modified oligonucleotides<sup>62</sup> also offers +2° increase in thermal stability ( $T_m$ ) per modification compared to PS-DNA. To improve the antisense properties of 2'-*O*-MOE-modified oligonucleotides, several novel 2'-*O*-modifications analogs of the parent have been synthesized (Figure 16).<sup>30</sup> These 2'-modified oligonucleotides showed excellent binding affinities to complementary RNA and high nuclease stability. Among these 2'-*O*-NMA<sup>31</sup> is a suitable modification for RNase-H based antisense therapeutics.<sup>32</sup>

### 1.6.3 Third generation antisense oligonucleotides

The ribose sugar moiety is locked by an oxymethylene bridge connecting the C(2') and C(4') atoms which conformationally restricts LNA (Figure 17a) monomers into *N*-type sugar pucker.<sup>33</sup> NMR spectroscopy revealed that LNA-containing ONs syntax into A-type duplex geometry.<sup>34</sup> LNA oligonucleotides exhibit unprecedented thermal stability when hybridized to a complementary DNA or RNA strand evidenced by thermal denaturation studies, *i.e.* increase in melting temperature ( $T_m$ ) of +2 to +8 per LNA monomer compared to unmodified duplexes.<sup>35</sup>



**Figure 17:** Structures of (a) LNA, (b)PMOs and (c)PNA

PMOs are the modified oligonucleotides which sugar ring in natural DNA replaced by a morpholine ring, also the internucleosidic ionic phosphate linkage by a non-ionic phosphorodiamidate inter-subunit linkage (Figure 17b). PMOs act by steric

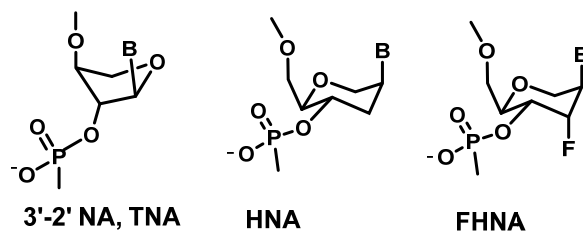
blocking to knock down the gene,<sup>36</sup> leading to the down-regulation of diseased protein synthesis, or they can up-regulate the viable protein synthesis by modifying the splicing of pre-mRNA.<sup>37</sup>

Peptide nucleic acid (PNA) is a DNA mimic, sugar-phosphate backbone replacement by a pseudopeptide backbone. PNAs are achiral, neutral DNA mimics that bind to complementary DNA/RNA sequences with high affinity and sequence specificity.<sup>38</sup> In PNA (Figure 17c) the natural nucleobases are attached via methylene carbonyl linkers to an uncharged pseudopeptide backbone composed of repeating *N*-(2-aminoethyl) glyceryl units. PNA hybridizes to complementary DNA/RNA sequences via specific base complementation to form duplexes for mixed sequences and triplexes for homopyrimidine/homopurine sequences.<sup>39</sup> The main limitations of PNAs are its poor water solubility and lack of cell permeability coupled with ambiguity in DNA/RNA recognition arising from its equally facile binding in a parallel/antiparallel fashion with the target sequence. These limitations are being systematically addressed with rationally modified PNA analogues.<sup>40</sup>

### 1.6.4 L-Threofuranosyl nucleic acid (TNA) and Hexitol nucleic acids (HNA)

Several backbone modifications have been studied in the search for a suitable candidate for antisense/antigene therapies. TNA (Figure 18a) is one of them, containing vicinally connected phosphodiester bridges with four atom backbone undergo base pairing in anti-parallel strand orientation and are capable of cross-pairing with RNA and DNA. It was found that the 3' and 2' phosphate groups preferred to be in quasi diaxial orientation which brought the P...P distance around 5.8 Å. This distance is similar to that between A form of duplexes leading to very stable TNA: RNA structures, crystal structure study also revealed the preferred pairing between TNA and RNA relative to that between DNA and TNA.<sup>41</sup> TNA is structurally the simplest of all potentially natural oligonucleotide-type nucleic acid alternatives studied so far. This, along with the base-pairing properties of TNA, warrants close scrutiny in the context of applications in antisense therapeutics.<sup>42</sup>





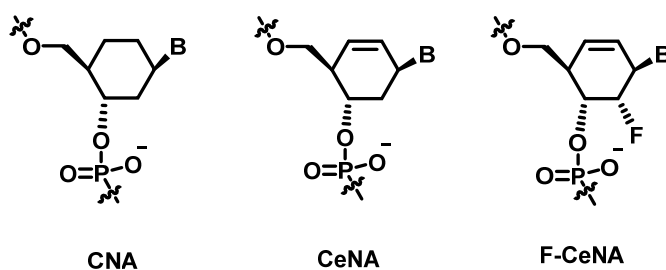
**Figure 18:** Structures of (a) TNA, (b) HNA and (c) FHNA

The replacement of conformationally flexible deoxyribose by the restricted anhydrohexitol ring resulted in a structurally preorganized HNA (Figure 18b) to form A- type helices with efficient base stacking.<sup>43</sup> Molecular associations between HNA and RNA were found to be more stable than those between HNA and DNA. <sup>1</sup>H NMR analysis of a HNA dimer confirmed the axial orientation of the base moiety with respect to the hexitol ring, and this was used as starting conformation for molecular dynamics study of HNA/RNA and HNA/DNA duplexes. Both complexes showed A-type geometry and very similar hydrogen bonding pattern between base pairs. In contrast, the crystal structure of HNA: RNA hybrid suggested less-pronounced rigidity of the backbone. The differences in hydration of the HNA and RNA backbones were also observed in the HNA: RNA duplex crystal structure. The hydration of the HNA strand promoted tighter bridging of the adjacent phosphate groups by water molecules and helped alteration in adjacent P-P distances that were close to RNA than DNA. This reinforcing effect upon hydration could be the second reason for overall duplex stability of HNA: RNA duplex in addition to the conformational preorganization imposed by the six membered rings.<sup>44</sup> Synthesis, biophysical and biological properties of the 3'-fluoro hexitol nucleic acid modified oligonucleotides were reported recently.<sup>45</sup> Axial 3'-fluoro substitution (Figure 18c) boosts the RNA affinity of HNA.

### 1.6.5 Six membered carbocyclic analogs

Furanose ring of natural DNA replaced by a six-membered ring is the basis for CNA (cyclohexyl nucleic acid) and cyclohexenyl nucleic acids (CeNAs), which are characterized by a high degree of conformational rigidity of the oligomers. Cyclohexyl nucleic acid (Figure 19a, CNA) was prepared in both enantiomeric (D/L) forms and D-CNA hybridizes to complementary RNA as compared to DNA with reduced affinity.<sup>46</sup> The cyclohexenyl nucleosides have exhibited potent antiviral

activity and the CeNA oligomers (Figure 19b) have been shown to mimic the function of RNA with increased enzymatic and chemical stability.<sup>47</sup>

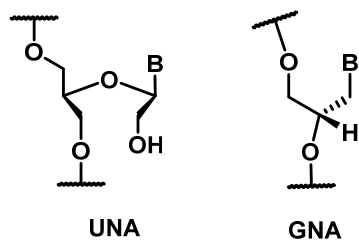


**Figure 19:** Structures of (a) CNA, (b) CeNA and (c) F-CeNA

In addition, the conformational flexibility allowed CeNA oligomers to have tremendous application in antisense therapeutics because of the consequential ability to induce RNase-H activity.<sup>48</sup> Recently 2'-fluoro cyclohexenyl nucleic acid (Figure 19c) modification was reported and it showed slightly lower duplex thermal stability with complementary RNA as compared to that of more rigid 3'-fluoro hexitol nucleic acid.<sup>49</sup> However, F-CeNA modified oligonucleotides were significantly more stable against digestion by snake venom phosphodiesterase as compared to unmodified DNA, 2'-fluoro RNA, 2'-O-MOE and FHNA.

### 1.6.6 Open chain analogs

The Structural difference between Unlocked nucleic acid (UNA) and RNA is absence the C2'-C3' bond normally found in ribonucleosides (Figure 20a), and it is therefore highly flexible.<sup>50</sup> This also mirrored in the destabilizing effects of UNA incorporation in duplexes, with up to 12 °C decrease in T<sub>m</sub> per UNA monomer in RNA:RNA duplexes and 10 °C per UNA monomer in RNA:DNA duplexes. UNA is highly resistant to nucleases.<sup>51</sup> UNA monomers have proven very useful in fine-tuning the specificity and potency of siRNA.<sup>52</sup>



**Figure 20:** Structures of (a) UNA and (b) GNA

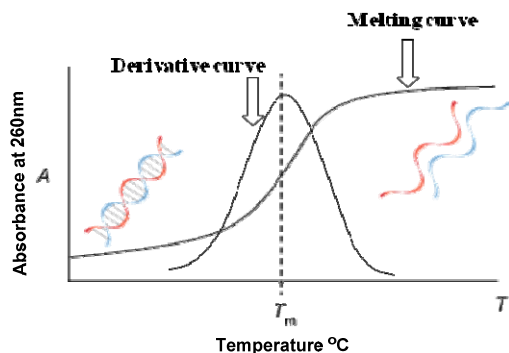
Glycol nucleic acids composed of repeating glycol units in backbone linked by phosphodiester bonds. GNA (Figure 20b) is synthetically easily accessible acyclic nucleic acid which forms highly stable antiparallel helical duplex structures.<sup>53</sup> Watson-Crick base pairing much more stable in GNA than its natural counterparts DNA and RNA. It is possibly the simplest of the nucleic acids, GNA is the most atom economical solution for a functional nucleic acid backbone.

### 1.7 Tools and techniques for structural studies of nucleic acid complexes, and higher ordered G-quadruplexes

#### 1.7.1 UV-spectroscopy

The absorbance of polynucleotide depends on the sum of the absorbance of all the nucleotides and the effect of the interacting nucleotides. The interactions cause a single strand to absorb less than the sum of its nucleotides and a double strand absorbs less than its two component single strand. The effect is called hypochromicity<sup>54</sup> which results from the coupling of the transition dipoles between neighboring stacked bases and is larger in amplitude for A-U and A-T pairs. Conversely, hyperchromicity refers to the increase in absorption when a double stranded nucleic acid is dissociated into single strands. The UV absorption of a DNA duplex increases typically by 10-20% when it is denatured. This transition from a stacked, hydrogen bonded double helix to an unstacked, strand-separated coil has a strong entropic component and is temperature dependent. The mid-point of this thermal transition is known as the **melting temperature** ( $T_m$ ). The converse of melting is the renaturation of two separated complementary strands to reform the duplex.

**Duplex melting:** The UV absorbance value at any given temperature is an average of the absorbance of duplex and single strands according to the ‘all or none model’<sup>55</sup> A plot of absorbance at 260 nm against temperature gives a sigmoidal curve in case of duplexes and the midpoint of the sigmoidal curve (Figure 21) called as the ‘melting temperature’ ( $T_m$ ) (equilibrium point) at which the duplex and the single strands exist in equal proportions.



**Figure 21:** An oligonucleotide duplex melting curve and derivative curve

**Quadruplex melting:** G-quadruplexes melting monitored by increasing temperature at 295 nm<sup>56</sup>. This wavelength maximises hyper/hypochromic shift between the folded and unfolded states. By cooling, the refolding process of the DNA/RNA also likewise, be monitored. Analysis of the UV-melting graphs can provide information of the thermodynamics of the system, which can be combined with CD and calorimetry data such as DSC or ITC to generate a fuller picture of the formation of the stacked complexes.

### 1.7.2 Circular Dichroism (CD)

Circular dichroism (CD) spectroscopy is extensively used to study the secondary structure of DNA in solution, particularly useful for studying chiral molecules and has very special significance in the characterization of biomolecules. This measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. It shows zero CD intensity in absence of regular structure, while an ordered structure results in a spectrum which can contain both positive and negative signals. In current literature commonly used units are the mean residue ellipticity (degree cm<sup>2</sup> dmol<sup>-1</sup>). In the nucleic acid, the heterocyclic bases are principal chromophores as these bases are planar, they do not have any intrinsic CD. It arises from the asymmetry induced by linked sugar group.

The CD signature of the B-form DNA as read from longer to shorter wavelength is a positive band centered at 275 nm, a negative band at 240 nm, with cross over around 258 nm. A-DNA is characterized by a positive CD band centered

## Chapter 1

at 260 nm that is larger than the corresponding B-DNA band, a fairly intense negative band at 210 nm and a very intense positive band at 190 nm (Figure 22).

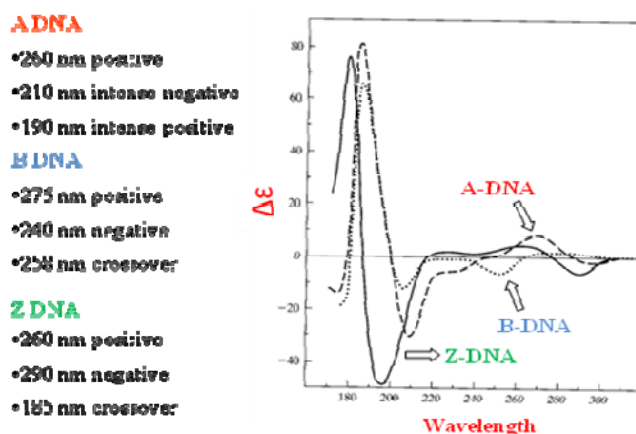


Figure 22: CD spectra of DNA secondary structure

### 1.7.3 Nuclear Magnetic Resonance (NMR)

NMR spectroscopy is a powerful experimental technique that is useful for confirmation of the formation and topology of complexes of DNA and RNA such as duplexes, triplexes and quadruplexes. It also can be used to study hydrogen-bonding between the nucleobases. The information obtainable includes sugar pucker characterization, backbone conformations and also local architecture. In the case of G-tetrad formation, 1D proton NMR spectra provides characteristic imino, amino and aromatic peaks, indicating the presence of Hoogsteen and Watson-Crick base-pairing. Further structure elucidation, of course, necessitates the use of more specialized techniques such as NOESY, TOCSY, HSQC-COSY, etc. The introduction of high field NMR and cryoprobes have in turn, reduced the need for higher concentrations of sample and made the analysis more affordable. This technique has clearly dominated the structure elucidation research of quadruplex folding and topology.

### 1.7.4 MALDI -TOF Mass Spectroscopy:

Matrix-assisted laser desorption/ionization time-of-flight (MALDI TOF) is a soft ionization technique<sup>57</sup> used for analysis of biomolecules, large organic molecules, where these large ions obtain in gas phase producing many fewer multiply charged ions.<sup>58</sup> MALDI is a two step process. First, desorption is triggered by a UV laser (nitrogen lasers 337 nm) beam. Matrix material heavily absorbs UV laser

light, leading to the ablation of upper layer (~micron) of the matrix material. The hot plume produced during ablation contains many species: neutral and ionized matrix molecules, protonated and deprotonated matrix molecules, matrix clusters and nanodroplets. Second, the analyte molecules are ionized (more accurately protonated or deprotonated) in the hot plume. The time-of-flight (TOF) analyzer uses an electric field to accelerate the ions through the same potential, and then measures the time they take to reach the detector. If the particles all have the same charge, the kinetic energies will be identical, and their velocities will depend only on their masses. Lighter ions will reach the detector first.

The matrix is the crystalline, a low molecular weight compound with several conjugated double bonds. The some basic properties are that they should not evaporate during standing in specrometer, often acidic for acting as a proton source to encourage ionization of the analyte molecules, strong optical absorption in either the UV or IR range. Most commonly used matrices are 3, 5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid, SA),  $\alpha$ -cyano-4-hydroxycinnamic acid (alpha-cyano or alpha-matrix, CHCA) and 2, 5-dihydroxybenzoic acid (DHB). 2', 4', 6' trihydroxy acetophenone (THAP). A solution of one of these molecules is made, in a mixture of highly purified water and an organic solvent (normally acetonitrile (ACN) or ethanol). Trifluoroacetic acid (TFA) or Ammonium citrate may also be added.

### 1.7.5 HPLC

HPLC is a chromatographic technique that has been used traditionally for separation of mixtures and also quantification of the components. When coupled to other analytical techniques, such as mass spectrometry, it is a powerful analytical technique. In the present work of this thesis, this technique has been successfully employed to study the degradation of single-stranded nucleic acid oligomers (both modified and unmodified) by analyzing the elution profile of the digestion mixture at different time-points by reverse-phase HPLC. Accordingly, the area under the peak can be used to quantify the amount of intact oligonucleotide. When plotted against time, this gives an idea of the half-life of the oligonucleotide under the given test conditions.

### 1.8 Present work

The thesis describes the design and synthesis of different chemically modified DNA/RNA analogues/mimics towards the ultimate goal of therapeutics. We introduced a robust synthetic route for cyclohexenyl nucleosides, known to exhibit potent antiviral activity and increased enzymatic and chemical stability. We also synthesized the acyclic nucleosides, threose nucleosides and incorporated into DNA sequences, their biophysical studies are presented briefly. We also studied the quadruplex structural studies of acyclic and TNA modified DNA sequences.

The thesis is organized as follows:

**Chapter 2** This chapter describes synthesis of Enantiopure Cyclohexenyl and  $\alpha$ -L-Cyclohexenyl Nucleosides.

**Chapter 3** This Chapter has been divided into two sections

Section A: In this section we describe the design and synthesis of open chain analogue of CeNA (*cis* and *trans*) modified monomers, incorporation into nucleic acid and the evaluation of binding properties with complementary DNA/RNA.

Section B: In this section we describe the synthesis of a *cis* and *trans* modified thrombin binding aptamer, its quadruplex formation.

**Chapter 4** This Chapter has been divided into two sections

Section A: In this section we describe the design and synthesis of 4'-MOM Threofuranosyl modified monomers, their incorporation into nucleic acids and the evaluation of modified NA binding properties with complementary DNA/RNA.

Section B: In this section we describe the synthesis of 4'-MOM Threofuranosyl modified thrombin binding aptamer, its quadruplex formation, and application as a thrombin inhibitor.

## 1.9 References

1. Watson, J. D.; Crick, F. H. C., *Nature* **1953**, 171, (4356), 737-738.
2. (a) Hoogsteen, K., *Acta. Cryst.* **1963**, 65, 907; (b) Crick, F. H. C., *J. Mol. Biol.* **1966**, 19, 548.
3. (a) Saenger, W., Springer-Verlag, New York. **1984** (b) Lescrinier, E.; Froeyen, M.; Herdewijn, P., *Nucleic Acid. Res.* **2003**, 31, 2975.
4. (a) Richmond, T. J.; Davey C. A., *Nature*, **2003**, 423, 145; (b) Leslie, A.G.; Arnott, S.; Chandrasekaran, R.; Ratliff, R.L., *J. Mol. Biol.* **1980**, 143, 49; (c) Wahl, M.; Sundaralingam, M., *Biopolymers*, **1997**, 44, 45; (d) Ghosh. A.; Bansal, M., *Acta Crystallogr D Biol Crystallogr.* **2003**, 59, 620.
5. Wang, A. H. J.; Quigley, G. J.; Kalpaks, F. J.; Vander, M. G.; VanBoom, J. H.; Rich, A., *Science* **1981**, 211, 171-176.
6. Felsenfeld, G.; Davies, D.; Rich, A., *J. Am. Chem. Soc.* **1957**, 79, 2023-2024.
7. (a) Moser, H.; Dervan, P., *Science* **1987**, 238, 645-650. (b) LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhouh, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; He' e'ne, C., *Nucleic Acids Res.* **1987**, 15, 7749-7760.
8. (a) Thuong, N. T.; Helene, C., *Angew. Chem. Int. Ed.* **1993**, 32, 666-690. (b) Beal, P. A.; Dervan, P. B., *Science* **1991**, 251, 1360-1363.
9. Hoogsteen, K., *Acta. Cryst.* **1959**, 12, 822-823.
10. Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; Parrinello, M., *Angew. Chem. Int. Ed.* **2013**, 52, 2269-2273.
11. (a) Wells, R. D.; Harvey, S. C., Springer-Verlag, New York. **1988**. (b) Helene, C., *Anti-Cancer Drug Design* **1991**, 6, 569-584. (c) Frank-Kamenetskii, M. D., *Methods in Enzymology* **1992**, 211, 180-191. (d) Soyfer, V. N.; Potaman, V. N., *Eds* **1996**, Springer-Verlag, New York.



12. (a) Bončina, M.; Lah, J.; Prislán, I.; Vesnaver, G., *J. Am. Chem. Soc.* **2012**, 134, 9657-9663. (b) Gray, R.; Buscaglia, R.; Chaires, J., *J. Am. Chem. Soc.* **2012**, 134, 16834-16844. (c) Mashimo, T.; Yagi, H.; Sannohe, Y.; Rajendran, A.; Sugiyama, H., *J. Am. Chem. Soc.* **2010**, 132, 14910-14918.
13. Gellert, M.; Lipsett, M.N.; Davies, D.R., *Proc. Natl. Acad. Sci.*, **1962**, 48, 2013-2018.
14. (a) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.*, **2006**, 34, 5402., (b) Collie, G. W.; Parkinson, G. N., *Chem. Soc. Rev.*, **2011**, 40, 5867., (c) Mirkin, S. M. *Front Biosci.* **2008**, 13, 1064-1071.
15. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, 355, 564-566.
16. (a) Macaya, R. F.; Schultze, P.; Smith, F.W.; Roe, J.A.; Feigon, J., *Proc. Natl. Acad. Sci., USA* , **1993**, 90, 3745-3749, (b) Schultze, P.; Macaya, R. F.; Feigon, J., *J. Mol. Biol.*, **1994**, 235, 1532-1547, (c) Kelly, J.A.; Feigon, J.; Yeates, T.O., *J. Mol. Biol.*, **1996**, 256, 417-422.
17. Dittmer, W.U.; Reuter, A.; Simmel, F.C.; *Angew. Chem. Int. Ed.*, **2004**, 43, 3550-3553.
18. Zamecnik, P. C.; Stephenson, M. L.; *Proceedings of the National Academy of Sciences USA* **1978**, 75, (1), 280-284.
19. Agrawal, S.; Zhao, Q. Y., Antisense therapeutics. *Current Opinion in Chemical Biology* **1998**, 2, (4), 519-528.
20. (a) Kurreck, J., *European Journal of Biochemistry* **2003**, 270, (8), 1628-1644. (b) Turner, J. J.; Fabani, M.; Arzumanov, A. A.; Ivanova, G.; Gait, M. J., *Biochimica et Biophysica Acta-Biomembranes* **2006**, 1758, (3), 290-300.
21. Stein, C. A.; Cohen, J. S. Phosphorothioate oligodeoxynucleotide analogues. In Cohen, J. S. (ed.): *Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression*. London: Macmillan Press, **1989**, 97.
22. Millar, P. S. Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): *Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression*. London: Macmillan Press, **1989**, 79.

## Chapter 1

---

23. Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G., *Nucleic Acid Res.* **1986**, *14*, 7421.
24. Li, H.; Huang, F.; Shaw, B. R., *Bioorg. Med. Chem.* **1997**, *5*, 787.
25. Froehler, B.; Ng, P.; Matteucci, M., *Nucleic Acid Res.* **1988**, *16*, 4831.
26. (a) Yano, J.; Smyth, G. E., *Adv Polym Sci*, **2012**, *249*, 1; (b) Oka, N.; Kondo, T.; Fujiwara, S., *Org. Lett.* **2009**, *11*, 967.
27. Guschlbauer, W.; Jankowski, K., *Nucleic Acids Res.* **1980**, *8*, 1421-1433.
28. Viazovkina, E.; Mangos, M. M.; Elzagheid, M. I.; Damha, M. J., *Curr. Protoc. Nucleic. Acid Chem.* Chapter 4, unit 415. **2002**.
29. Watts, J. K.; Martin-Pintado, N.; Gomez-Pinto, I.; Schwartzenruber, J.; Portella, G.; Orozco, M.; Gonzalez, C.; Damha, M. J., *Nucleic Acids Res.* **2010**, *38*, 2498-2511.
30. Prakash, T. P., *Chemistry & biodiversity* **2011**, *8*, 1616-1641.
31. Prakash, T. P.; Kawasaki, A. M.; Lesnik, E. A.; Owens, S. R.; Manoharan, M., *Org. Lett.* **2003**, *5*, 403-406.
32. Prakash, T. P.; Kawasaki, A. M.; Wancewicz, E. V.; Shen, L.; Monia, B. P.; Ross, B. S.; Bhat, B.; Manoharan, M., *J. Med. Chem.* **2008**, *51*, 2766-2776.
33. a) Singh, S. K.; Koshkin, A. A.; Wengel, J.; Nielsen, P. *Chem. Commun.* **1998**, 455. b) Koshkin, A. A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998**, *54*, 3607. c) Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T. *Tetrahedron Lett.* **1997**, *38*, 8735.
34. a) Petersen, M.; Bondensgaard, K.; Wengel, J.; Jacobsen, J. P. *J. Am. Chem. Soc.* **2002**, *124*, 5974. b) Nielsen, K. E.; Rasmussen, J.; Kumar, R.; Wengel, J.; Jacobsen, J. P.; Petersen, M. *Bioconjugate Chem.* **2004**, *15*, 449.
35. a) Koshkin, A. A.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V. K.; Singh, S. K.; Wengel, J. *J. Am. Chem. Soc.* **1998**, *120*, 13252. b) Obika, S.; Nanbu, D.;

## Chapter 1

---

- Hari, Y.; Andoh, J.-i.; Morio, K.-i.; Doi, T.; Imanishi, T. *Tetrahedron Lett.* **1998**, *39*, 5401. c) Wengel, J. *Acc. Chem. Res.* **1999**, *32*, 301.
36. Summerton, J., *Biochim. Biophys. Acta* **1999**, 1489, 141-158.
37. Draper, B. W.; Morcos, P. A.; Kimmel, C. B., *Genesis* **2001**, *30*, 154-156.
38. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., *Science* **1991**, *254*, 1497-1500.
39. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordon, B.; Nielsen, P. E., *Nature* **1993**, *365*, 566-568.
40. (a) Ganesh, K. N.; Nielsen, P. E., *Curr. Org. Chem.* **2000**, *4*, 931-943. (b) Kumar, V. A., *Eur. J. Org. Chem.* **2002**, 2002, 2021-2032. (c) Kumar, V. A.; Ganesh, K. N., *Acc. Chem. Res.* **2005**, *38*, 404-412. (d) Uhlmann, E.; Breipohl, G.; Will, D., *Angew. Chem., Int. Ed.* **1998**, *37*, 2796-2823.
41. Pallan, P. S.; Wilds, C. J.; Wawrzak, Z.; Krishnamurthy, R.; Eschenmoser, A.; and Egli, M., *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 5893.
42. Schoning, K. U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser A.; *Science*, **2000**, *290*, 1347-1351.
43. De Winter, H.; Lescrinier, E.; Van Aerschot, A.; Herdewijn, P., *J. Am. Chem. Soc.* **1998**, *120*, 5381.
44. Maier, T.; Przylas, I.; Strater, N.; Herdewijn, P.; Saenger, W., *J. Am. Chem. Soc.* **2005**, *127*, 2937.
45. Egli, M.; Pallan, P. S.; Allerson, C. R.; Prakash, T. P.; Berdeja, A.; Yu, J.; Lee, S.; Watt, A.; Gaus H.; Bhat, B.; Swayze, E. E.; Seth, P. P., *J. Am. Chem. Soc.* **2011**, *133*, 16642.
46. Mourinsh, Y.; Rosemeyer, H.; Esnouf, R.; Mevedovici, A.; Wang, J.; Ceulemans, G.; Lescrinier, E.; Hendrix, C.; Busson, R.; Sandra, P.; Seela, F.; Aerschot, A. V.; Herdewijn, P., *Chem. Eur. J.* **1999**, *5*, 2139.

## Chapter 1

---

47. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A. and Herdewijn, P. *J. Am. Chem. Soc.*, **2000**, *122*, 8595–8602.
48. (a) Wang, J.; Verbeure, B.; Luyten, I.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P. *Nucleos. Nucleot. & Nucleic Acids*, **2001**, *20*, 785-788. (b) Wang, J.; Verbeure, B.; Luyten, I.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P. *Nucleic Acids Res.* 2001, *29*, 4941-4947.
49. Seth, P. P.; Yu, J.; Jazayeri, A.; Pallan, P. S.; Allerson, C. R.; Østergaard, M.E.; Liu, F.; Herdewijn, P.; Egli, M.; Swayze, E. E., *J. Org. Chem.* **2012**, *77*, 5074.
50. Langkjaer, N.; Pasternak, A.; Wengel, J., *Bioorg. Med. Chem.* **2009**, *17*, 5420-5425.
51. Itkes, A.V.; Karpeisky, M.; Kartasheva, O.N.; Mikhailov, S.N.; Moiseyev, G.P.; Pfliederer, W.; Charubala, R.; Yakovlev, G.I., *FEBS Lett.* **1988**, *236*, 325-328.
52. Bramsen, J.B.; Pakula, M.M.; Hansen, T.B.; Bus, C.; Langkjaer, N.; Odadzic, D.; Smicius, R.; Wengel, S.L.; Chattopadhyaya, J.; Engels, J.W., *Nucleic Acids Res.* **2010**, *38*, 5761-5773.
53. Zhang Lili.; Peritz Adam.; Meggers Eric., *J. Am. Chem. Soc.* **2005**, *127*, 4174-4175.
54. Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M., *Nucleic acids in Chemistry and Biology*. RSC publication 3<sup>rd</sup> edition.
55. Cantor, C. R.; Schimmel, P. R. (Eds), *Biophysical Chemistry part III* W. H. Freeman and Company 1971, New York.
56. Mergny, J.-L.; Lacroix, L., *Curr. Protoc. Nucleic Acid Chem.* **2009**, Ch.17, Unit 17.1.

## *Chapter1*

---

57. Vanholde, K. E.; Brahms, J.; Michelson.A., *Journal of Molecular Biology* **1965**, 12, (3), 726.
58. Little, D. P.; Cornish, T. J.; Odonnell, M. J.; Braun, A.; Cotter, R. J.; Koster, H., *Anal. Chem.*, **1997**, 69, (22), 4540-4546.
59. Crain, P. F.; McCloskey, J. A., *Curr. Opin. Biotechnol.*, **1998**, 9, (1), 25-34.

## CHAPTER 2

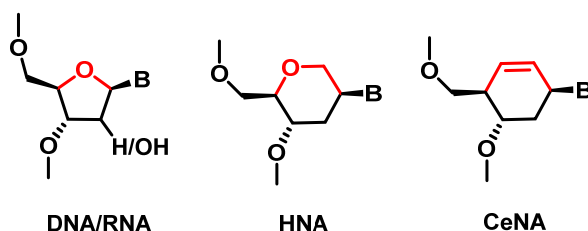
Design and synthesis of  
cyclohexenyl and  $\alpha$ -L-cyclohexenyl  
nucleosides

## 2. Design and synthesis of cyclohexenyl and $\alpha$ -L-cyclohexenyl nucleosides

### 2.1 Introduction

Antisense oligonucleotides have attracted lot of attention for drug designing over last two decades. These are specifically designed to target mRNA. The key role of antisense oligonucleotides in targeting mRNA,<sup>1</sup> siRNA,<sup>2</sup> miRNA,<sup>3</sup> splice correction,<sup>4</sup> etc has further highlight the need for the development of modified oligonucleotides. Natural oligonucleotides cannot be used as therapeutic agents due to poor stability against nucleases. Therefore, modifications have been introduced to enhance the therapeutic efficiency like increasing the binding affinity, water solubility, degradation of RNA by RNase-H enzyme, cellular uptake etc.

Six membered cyclic analogs are one of the several other prominent modifications for antisense therapy.<sup>5</sup> The furanose ring of natural nucleoside was replaced by six membered ring which conferred suitable rigidity to antisense constructs. Therefore upon formation of a double stranded complex loss of entropy was less. The 1,5-anhydrohexitol nucleic acid (HNA)<sup>6</sup> modification exhibited selective and strong RNA recognition, yielding the desired A-type double helix but prohibited RNase H recognition. The double bond was introduced to increase the flexibility allowing adaptation and enzymatic recognition, hopefully preserving the increased affinity for complementary targets.

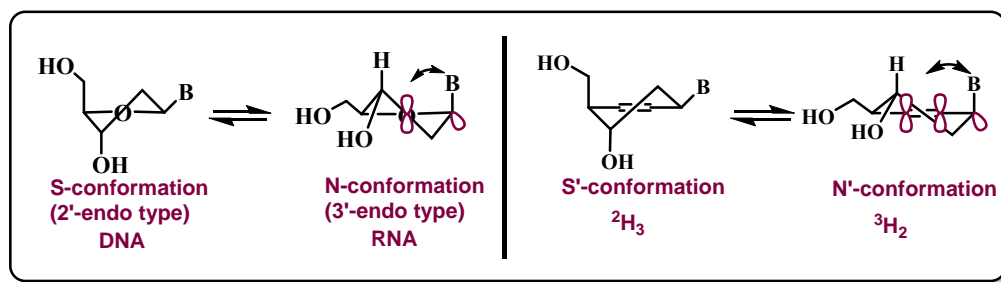


**Figure 1** Natural nucleic acids and modified analogues.

As per structural considerations the cyclohexene nucleosides are one of the best mimics of natural furanose nucleosides in six-membered ring analogues (Figure 1). The thermodynamic parameters of cyclohexenyl nucleoside ( $\Delta G$  1.8 kJ/mol) between S-type ( $^2H_3$ ) and N-type ( $^2H^3$ ) and equilibrium occurs *via* the eastern hemisphere with a barrier of 10.9 kJ/mol) are very similar to natural ribose

## Chapter 2

nucleosides ( $\Delta G$  2 kJ/mol between N-type and S-type, and equilibrium occurs *via* the eastern hemisphere with a barrier of 4– 20 kJ/mol).<sup>7</sup> The  $\Pi$ - $\sigma^*$  interaction of cyclohexenyl nucleoside mimics the anomeric effect in furanose nucleoside.<sup>8</sup> CeNA (cyclohexenyl nucleic acids) exists in both N-type and S-type sugar conformations at the nucleoside level (Figure 2) due to highly flexible cyclohexene ring with low energy barrier between the two forms.



**Figure 2:** Comparison of the conformational equilibrium

The cyclohexenyl nucleosides have demonstrated potent antiviral activity and the CeNA oligomers have been shown to mimic the function of RNA with increased enzymatic and chemical stability.<sup>9</sup> The presence of the double bond allows enough flexibility within a CeNA-RNA double helix, to be recognized by RNase H.<sup>10</sup> CeNA also been shown to be useful in siRNA applications.<sup>11</sup>

### 2.2 Synthesis of Cyclohexenyl nucleosides

Although cyclohexenyl nucleic acids have tremendous applications, the main limitation that has stymied the scope of this highly valuable discovery is the lack of a robust and scalable synthetic strategy of the key intermediate in enantiomerically pure form. Two synthetic routes were reported in the literature for enantiopure synthesis of cyclohexenyl nucleic acids.

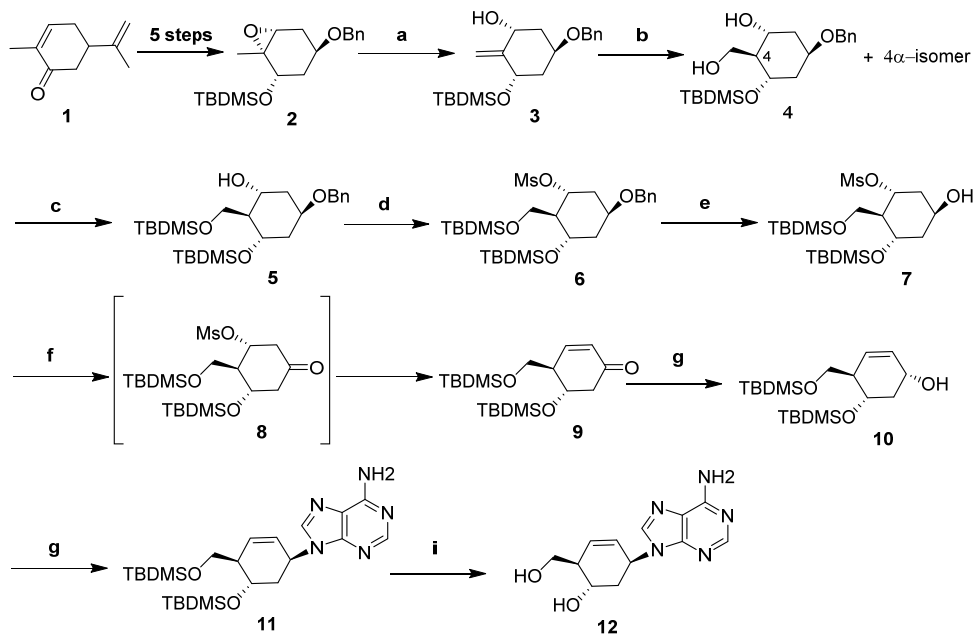
First synthetic route was reported by Herdewijn, P *et al.* (scheme 1) Synthesis started from R-(-)-carvone<sup>12</sup>, which converted into **2** by reported procedures within 5 steps. Epoxide ring in **2** was opened by using LiTMP, Et<sub>2</sub>AlCl followed by hydroboration of the double bond using 9BBN to get **4**.  $\alpha$ -isomer at 4-carbon also was observed as a side product in addition to the required compound **4**. Then TBDMS protection of primary hydroxyl group then mesylation of secondary hydroxyl compound and benzyl group deprotection yields **7**. Compound **7** was further taken for



## Chapter 2

oxidation of hydroxyl group to get **8** followed by elimination to get **9**. Compound **9** was reduced to alcohol **10** which was subjected to Mitsunobu reaction conditions to obtain the required cyclohexenyl adenine nucleoside. The main drawback of this synthesis is the overall yield was low (2–3%), which is not practical for oligonucleotide synthesis.

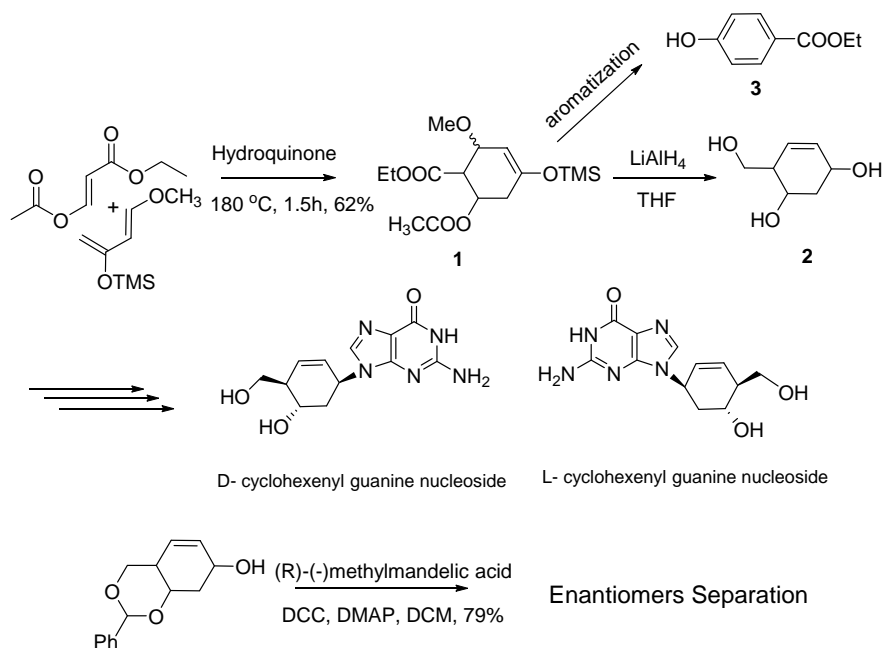
**Scheme 1** Synthesis of cyclohexenyl adenine nucleoside from R-(-)-carvone



**Reagents and conditions:** (a) LiTMP, Et<sub>2</sub>AlCl; (b) 9BBN, THF; (c) TBDMSCl, imidazole, DMF, rt, 70%; (d) MsCl, Et<sub>3</sub>N, DCM, 0 °C, 92%; (e) Pd-C (10%), HCOONH<sub>4</sub>, MeOH, reflux, 76%; (f) MnO<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, rt, 48% and 47% recovery of 7; (g) NaBH<sub>4</sub>, CeCl<sub>3</sub>·7H<sub>2</sub>O, MeOH, 0 °C to rt, 91%; (h) Adenine, DEAD, PPh<sub>3</sub>; (i) TFA/H<sub>2</sub>O(3:1).

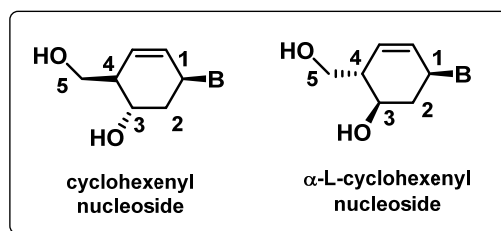
Second synthetic route was also reported by Herdewijn, P *et al.*<sup>13</sup> They started the synthesis with Diels–Alder reaction of ethyl (2E)-3-acetoxyprop-2-enoate as dienophile<sup>14</sup> and Danishefsky's diene<sup>15</sup> to construct the six-membered ring skeleton **1**. The major drawback in this was compound **1** could easily undergo aromatization to give **3** which is not useful for further synthesis (scheme 2).

## Scheme 2 Synthesis of cyclohexenyl guanine nucleoside from Danishefsky's diene



## 2.3 Our design and rationale

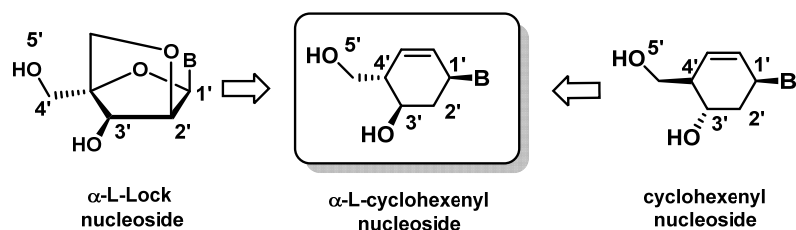
We were interested in developing an easy and straight forward synthetic route for the cyclohexenyl nucleosides. The enzymatic resolution also would give rise to the other stereoisomer corresponding to  $\alpha$ -L-cyclohexenyl nucleoside which is not known in the literature (Figure 3).



**Figure 3:** structures of cyclohexenyl nucleoside and  $\alpha$ -L-cyclohexenyl nucleoside

The unprecedented thermal stability of duplexes involving LNA has inspired to investigate the properties of the stereoisomers of LNA and stereo isomeric analog termed as  $\alpha$ -L-LNA.<sup>16</sup> The main difference between LNA and its isomer was LNA is in N-type conformation whereas  $\alpha$ -L-LNA was in S-type (DNA like) sugar geometry which is capable to show RNase H activity.

We designed the stereo isomer of cyclohexene nucleoside named as  $\alpha$ -L-cyclohexenyl nucleoside. The major advantage of  $\alpha$ -L-cyclohexenyl nucleoside is DNA like (S-type) sugar geometry, which is capable of eliciting the RNase-H activity as well as advantages of both  $\alpha$ -L-LNA and CeNA. We proposed a most efficient common synthetic route for synthesis of both cyclohexenyl and  $\alpha$ -L-cyclohexenyl nucleosides (Figure 4).

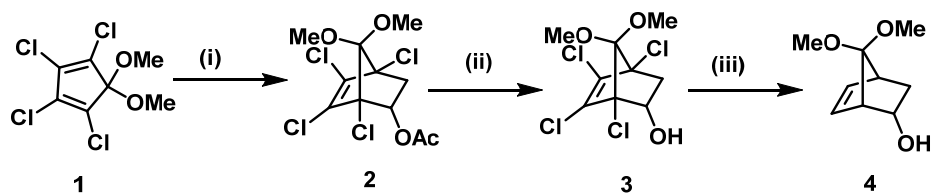


**Figure 4:** Design of  $\alpha$ -L-cyclohexenyl nucleoside

## 2.4 Methodology, Results and Discussion

Diels-Alder reaction between commercially available 5,5-dimethoxy-1,2,3,4-tetrachlorocyclopentadiene **1** and vinyl acetate at 120 °C gave exclusively endo adduct ( $\pm$ )-**2**.<sup>17</sup> The acetate group in ( $\pm$ )-**2** was hydrolysed under acidic conditions to give the free alcohol ( $\pm$ )-**3**. Compound ( $\pm$ )-**3** was then subjected to reductive dehalogenation under Birch conditions to get ( $\pm$ )-**4**. The purpose of the acetate hydrolysis was to improve the yields of the dehalogenation reaction (Scheme 3).<sup>18</sup>

**Scheme 3** Synthesis of dehalogenated bicyclic intermediate **4**



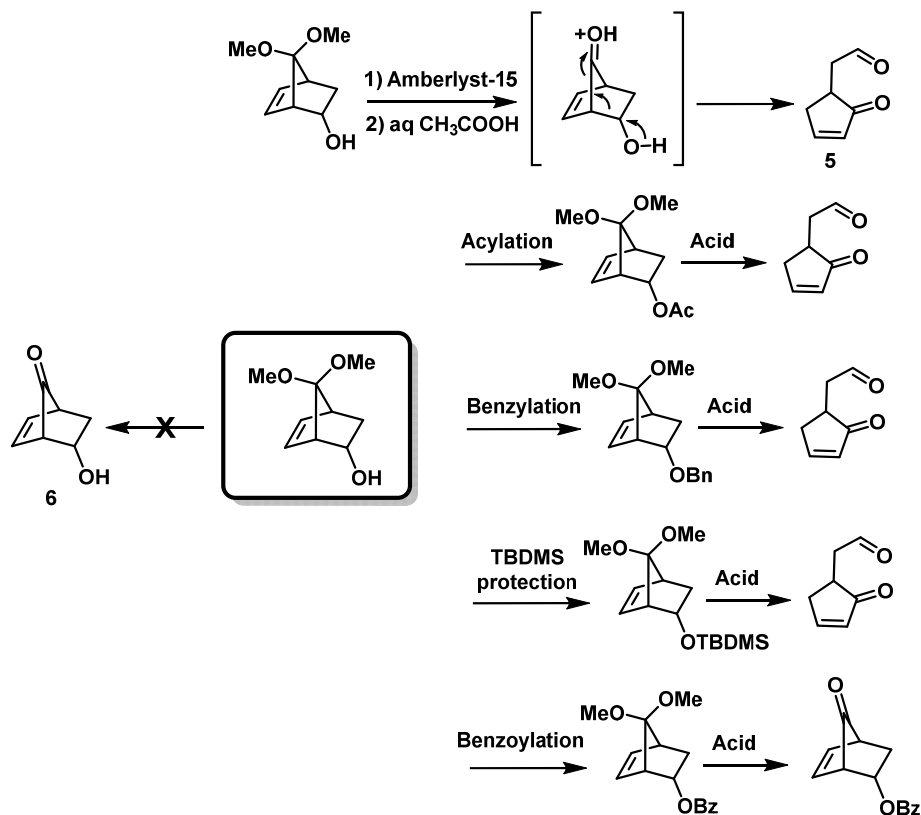
**Reagents and conditions:** (i) vinyl acetate, 120 °C, 5h, 84% (ii) H<sub>2</sub>SO<sub>4</sub>, MeOH, 85 °C, 8-10 h, 95% (iii) Na, Liq NH<sub>3</sub>, THF: EtOH 0.5h, 74%.

Next aim was to hydrolyze the ketal group of **4** to get free ketone ( $\pm$ ) **6** proved to be difficult as we found that the formation of compound ( $\pm$ )-**5** was a major product under variety of ketal hydrolysis conditions. We employed different protecting groups for the secondary hydroxy group in ( $\pm$ )-**4** such as TBS, benzyl or acetate protection<sup>18</sup> but each time compound ( $\pm$ )-**5** was formed as a major product in the acidic

## Chapter 2

hydrolysis. The silyl ether, benzyl ether and the acetate protecting groups probably undergo hydrolytic cleavage (scheme 4). A report was found in the literature where a similar retro aldol type of rearrangement was observed and a similar major product was obtained.<sup>19</sup> Then we tried benzoyl protection as suggested by Sgarbi *et. al.*<sup>20</sup>

**Scheme 4** Different conditions tried for ketal hydrolysis and mechanism for rearranged product.

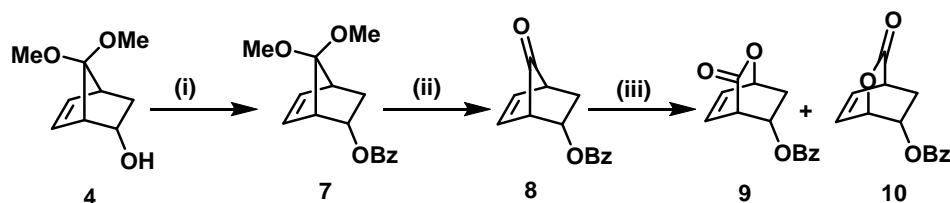


This was found to be stable under these reaction conditions and the rearranged product ( $\pm$ )-**5** was not formed. We observed that the aqueous acetic acid reflux conditions of ( $\pm$ )-**7** gave a clean ketal hydrolysis product ( $\pm$ )-**8** in 81% yield (scheme 5). The Bayer-Villager's oxidation of ketone ( $\pm$ )-**8** gave two regioisomeric inseparable mixture of lactones ( $\pm$ )-**9** and ( $\pm$ )-**10** in 7:3 proportion respectively, in high yield (Scheme 5). The products were confirmed by 2D COSY NMR spectroscopy. The mixture of lactones (**9+10**) was reduced with LiAlH<sub>4</sub> and gave the tri hydroxy substituted cyclohexene derivatives ( $\pm$ )-**11** and ( $\pm$ )-**12** as an inseparable mixture (scheme 6). The *cis* geometry of C1&C4 and *trans* placement of substituent at C3 was fixed at this point in compound **11** and **12**. The mixture of regioisomers

## Chapter 2

obtained was subjected to chemo selective 1,3-diol protection and only compound ( $\pm$ )-**11** was protected to furnish ( $\pm$ )-**13** quantitatively.

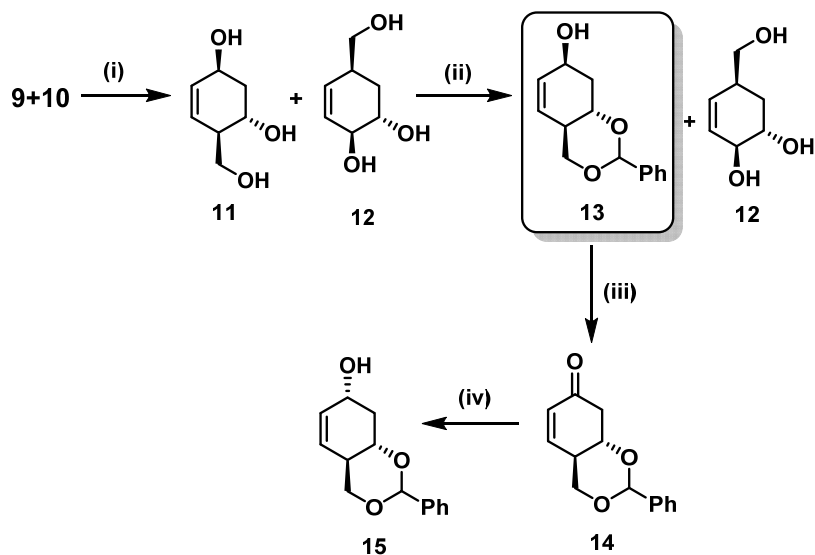
**Scheme 5** Synthesis of the mixture of lactones **9** and **10**



**Reagents and conditions:** (i) BzCl, pyridine, 3h, 93% (ii) CH<sub>3</sub>COOH : H<sub>2</sub>O (6:1), 120 °C, 4 h, 81% (iii) mCPBA, DCM, 0 °C, 5h, 92% ,**9:10** (70:30).

Then the unreacted triol ( $\pm$ )-**12** could be easily separated chromatographically. Successfully we have synthesized key intermediate ( $\pm$ )-**13** with good yield. The compound ( $\pm$ )-**12** obtained as byproduct may be used further to achieve 2'-5' linkages of 3'-deoxy cyclohexenyl nucleosides which may be useful in therapeutic studies. Compound ( $\pm$ )-**13** subjected for oxidation using CrO<sub>3</sub>/pyridine in 3h to get the enone ( $\pm$ )-**14** (scheme 6). The reported MnO<sub>2</sub> oxidation takes very long time and depends on the quality of MnO<sub>2</sub> used in the reaction.

**Scheme 6** Synthesis of 1,3 diol protected key intermediate **13**



**Reagents and conditions:** (i) LiAlH<sub>4</sub>, dry THF, -15 °C, 2h, 72 % (ii) PhCH(OMe)<sub>2</sub>, PTSA, dry dioxane, rt, 24 h, 81% (iii) CrO<sub>3</sub>, dry pyridine, Ac<sub>2</sub>O, dry DCM, 2 h, 92% (iv) NaBH<sub>4</sub>, CeCl<sub>3</sub>·7H<sub>2</sub>O, MeOH, 3h, 80%.

## Chapter 2

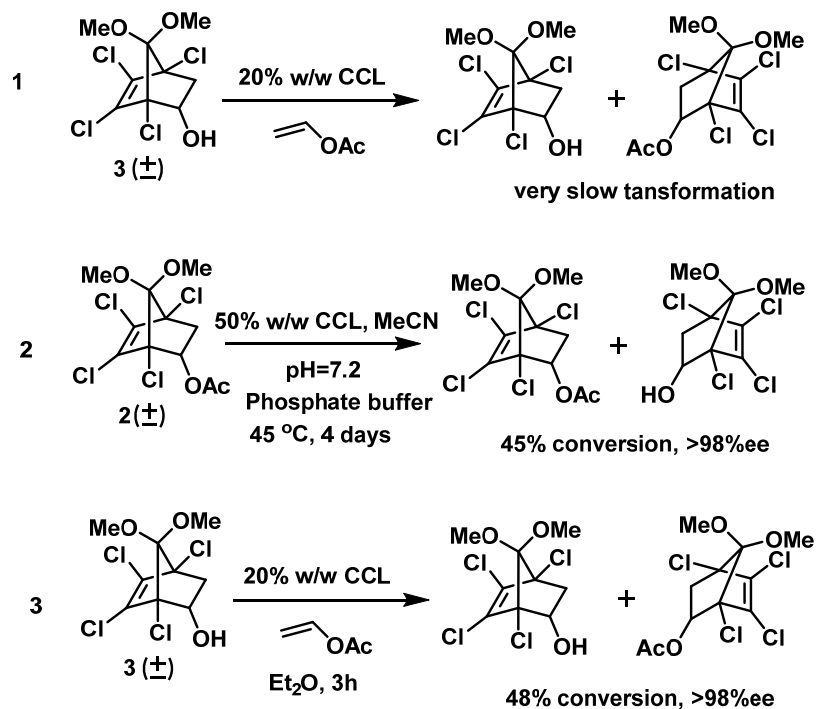
---

The use of CrO<sub>3</sub> reduces the reaction time and is highly reproducible. Compound (±)-**14** was then subjected to Luche reduction to get epimeric alcohol (±)-**15** using NaBH<sub>4</sub> in the presence of CeCl<sub>3</sub>·7H<sub>2</sub>O (Scheme 6)<sup>13</sup> and it can be converted to cyclohexenyl nucleosides units using reported procedures.<sup>21</sup>

### 2.5 Enzymatic resolution

In the field of organic synthesis the use of enzymes has become an interesting area. Since many enzymes have been demonstrated to possess activity and widely used to carry out synthetic transformations. Hydrolases are the most frequently used enzymes due to their considerable stability and broad substrate spectrum. Enzymes are commercially available and they work under mild reaction conditions. Enzyme catalyzed kinetic resolution of racemic alcohols through *trans* esterification is an attractive route among the numerous synthetic methods for asymmetric synthesis.

Enzymatic resolution of (±)-**3** is known in the literature using *Candida cylindracea* lipase (*CCL*) and vinyl acetate as an acyl donor.<sup>22</sup> We followed the reported reaction conditions but reaction was very sluggish, it took almost 7 days to complete and could not be scaled up (scheme 7.1). The acylation was sluggish but enantio specificity was good, then we tried enzymatic hydrolysis of (±)-**2** using the same lipase in phosphate buffer (pH 7.2). The acetate compound was less soluble in buffer due to highly non polar nature of the compound. When acetonitrile was used as co-solvent along with phosphate buffer (pH 7.2), good enantioselective hydrolysis was observed in 4 days (Scheme 7.2, yield 45%, ee 98%). Continuing the hydrolysis of the remaining enantiomerically enriched acetate for 24 h, the other enantiomer as unreacted acetate **2** was obtained in good yield and high ee (yield 42%, ee 98 %). Alternatively, optimization of the resolution time was also be achieved by using *CCL* for esterification using vinyl acetate as acyl donor in diethyl ether solvent. The acylation was enantioselective, giving the acetate (48% yield and >98% ee) in 3h (Scheme 7.3).

Scheme 7 Enzymatic resolution of racemic compounds **2** and **3**

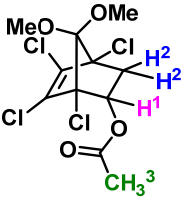
The enantiomeric identity of the resolved acetate **2** was established by <sup>1</sup>H NMR using chiral shift reagent and chiral HPLC analysis.

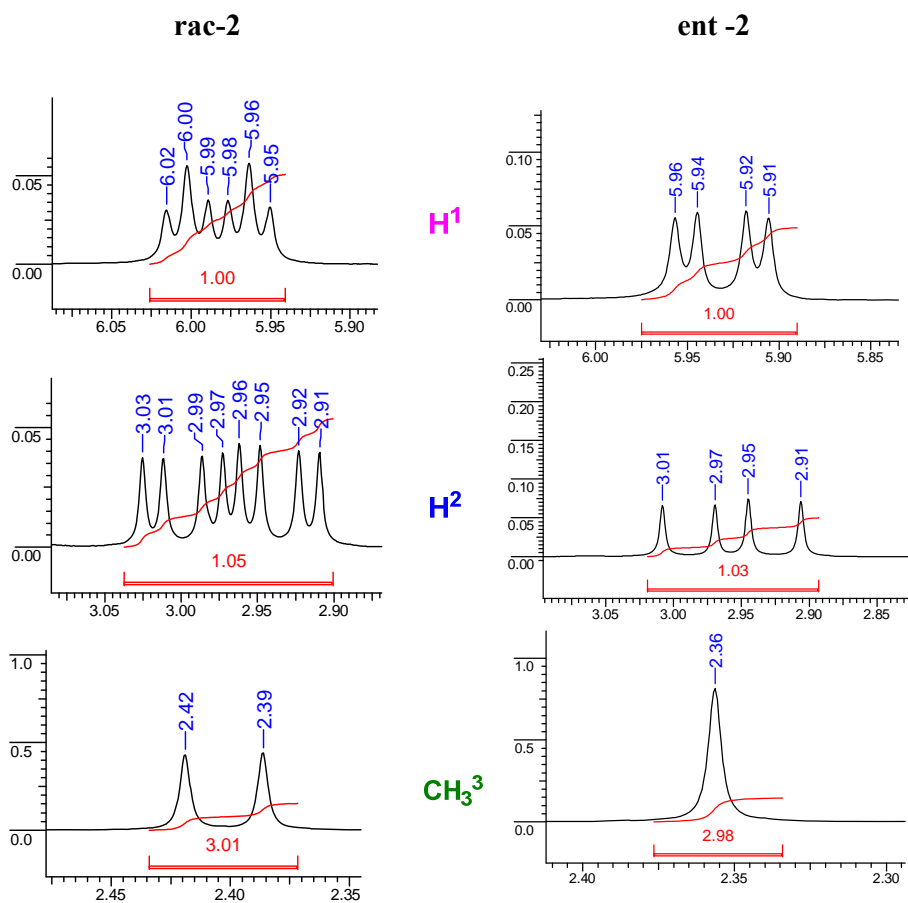
## 2.6 Enantiopurity confirmation by chiral shift reagent NMR of rac-**2** and ent-**2**

In principle, the enantiomeric molecules do not differ in physical properties unless they are placed in a chiral environment. We used Tris[3-(heptafluoropropyl-hydroxymethylene)-d-camphorato] europium(III) derivative as chemical shift reagent which contains chiral ligand and the resulting complexes with chiral molecules are diastereotopic. Protons attached to chiral centers will give rise to separate signals only if the molecule is placed in a diastereotopic environment.

We demonstrated the splitting patterns of the protons of rac-**2** and ent-**2** in chiral shift reagent NMR (Table 1). The splitting pattern of H<sup>1</sup> and H<sup>2</sup> protons in enantiopure compound was dd whereas in racemic compound it was multiplet. The methyl protons (H<sup>3</sup>) of acetate group showing one singlet in ent-**2** and two singlets in rac-**2** (Figure 5 & Table 1).

**Table 1** Chemical shift values of **rac-2** and **ent-2** with NMR shift reagent

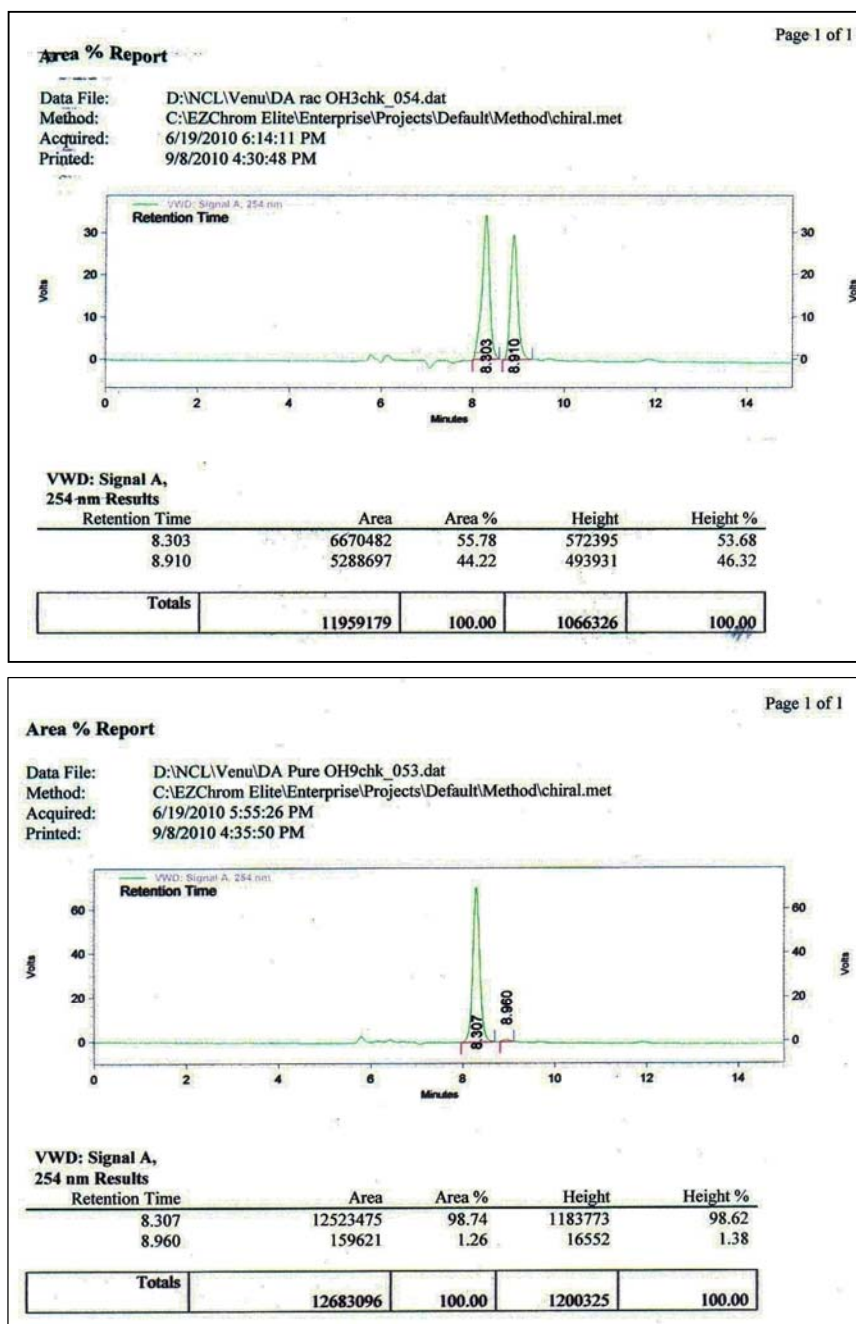
structure	Proton	rac-2		ent-2	
		$\delta$ -value	Splitting pattern	$\delta$ -value	Splitting pattern
	H <sup>1</sup>	6.03	multiplet	5.97	dd
	H <sup>2</sup>	3.04	multiplet	3.02	dd
	CH <sub>3</sub> <sup>3</sup>	2.43	2 singlets	2.37	singlet

**Figure 5:** <sup>1</sup>H NMR splitting patterns of **rac-2** and **ent-2** with chiral shift reagent



## Chapter 2

The enantiopurity was confirmed by chiral HPLC. The rac-2 compound showed two peaks, whereas only one peak observed in HPLC chromatogram for ent-2 (Figure 6).



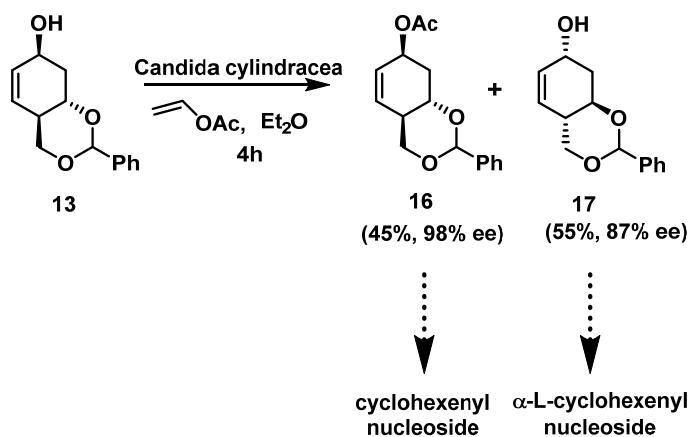
**Figure 6:** Chiral HPLC chromatograms of rac-2 and ent-2 compounds

Successfully we have introduced an efficient method for enzymatic resolution with good yields as well as good enantiomeric excess. The enzymatic resolution at such an early stage to get the different sugar analogues would be time consuming, we

## Chapter 2

decided to resolve the two enantiomers of ( $\pm$ )-**13** using enzymatic acylation. Compound ( $\pm$ )-**13** was subjected to acylation using vinyl acetate as donor and *Candida cylindracea* lipase (CCL) to get **16** in excellent yield and high enantiomeric purity (Scheme 8).

**Scheme 8** Enzymatic resolution of racemic compounds **13**



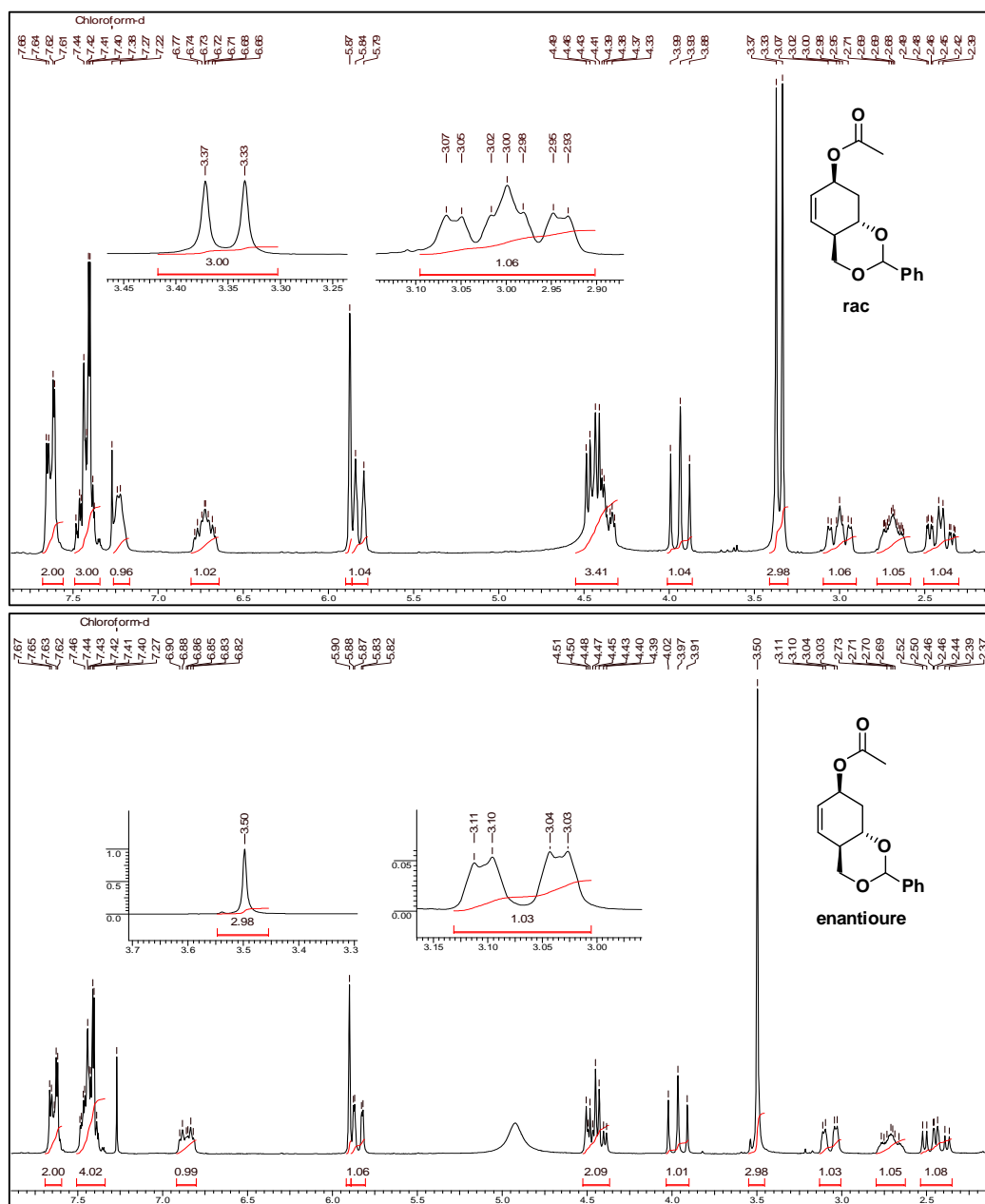
The high enantiomeric purity was established by  $^1\text{H}$  NMR using chiral shift reagent. The enantiomeric identity of the resolved acetate **16** and **17** was established after converting the acetate **16** (hydrolysis of the acetate followed by oxidation and reduction as described in Scheme 6) to the known compound *i.e.* enantiomerically pure D-isomer **15** and comparing with the reported<sup>19</sup> HPLC retention time on chiral HPLC column.

**Table 2** Chemical shift values of *rac*-**15** and *ent*-**15** with NMR shift reagent

structure	Proton	<i>rac</i> - <b>16</b>		<i>ent</i> - <b>16</b>	
		$\delta$ -value	Splitting pattern	$\delta$ -value	Splitting pattern
	$\text{H}^1$	3.07	multiplet	3.11	dd
	$\text{CH}_3^2$	3.37	2 singlets	3.50	singlet

## Chapter 2

In case **rac-16** and **ent-16** also we noticed the difference in splitting patterns of the protons of in chiral shift reagent NMR (Table 2). The splitting pattern of H<sup>1</sup> proton in enantiopure compound was dd, whereas in racemic compound it was multiplet. The methyl protons (H<sup>2</sup>) of acetate group showing one singlet in **ent-16** and two singlets in **rac-16** (Figure 7) and chiral HPLC chromatogram showed in Figure 8



**Figure 7:** <sup>1</sup>H NMR splitting patterns of **rac-16** and **ent-16** with chiral shift reagent

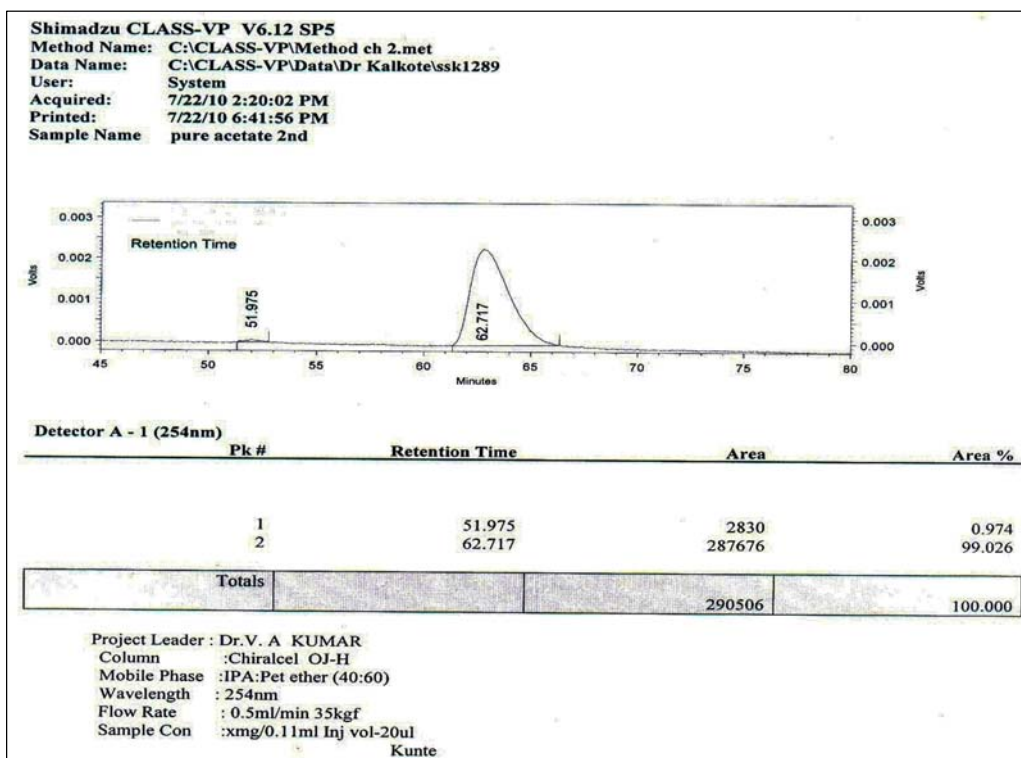
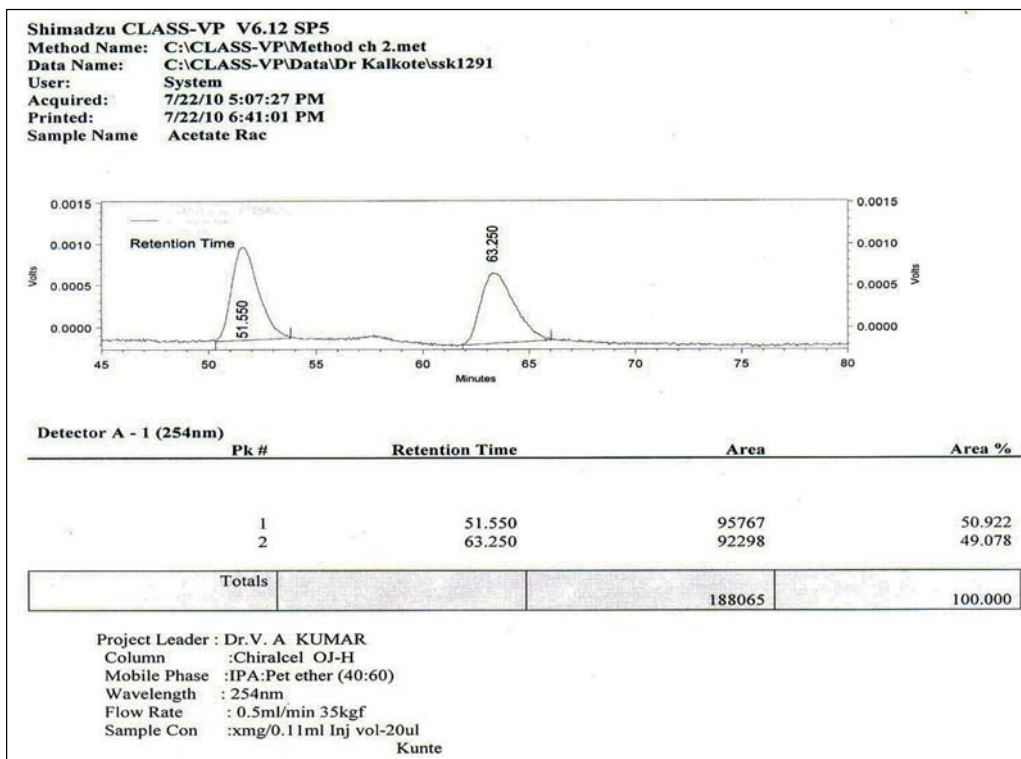


Figure 8: Chiral HPLC chromatograms of rac-16 and ent-16 compounds

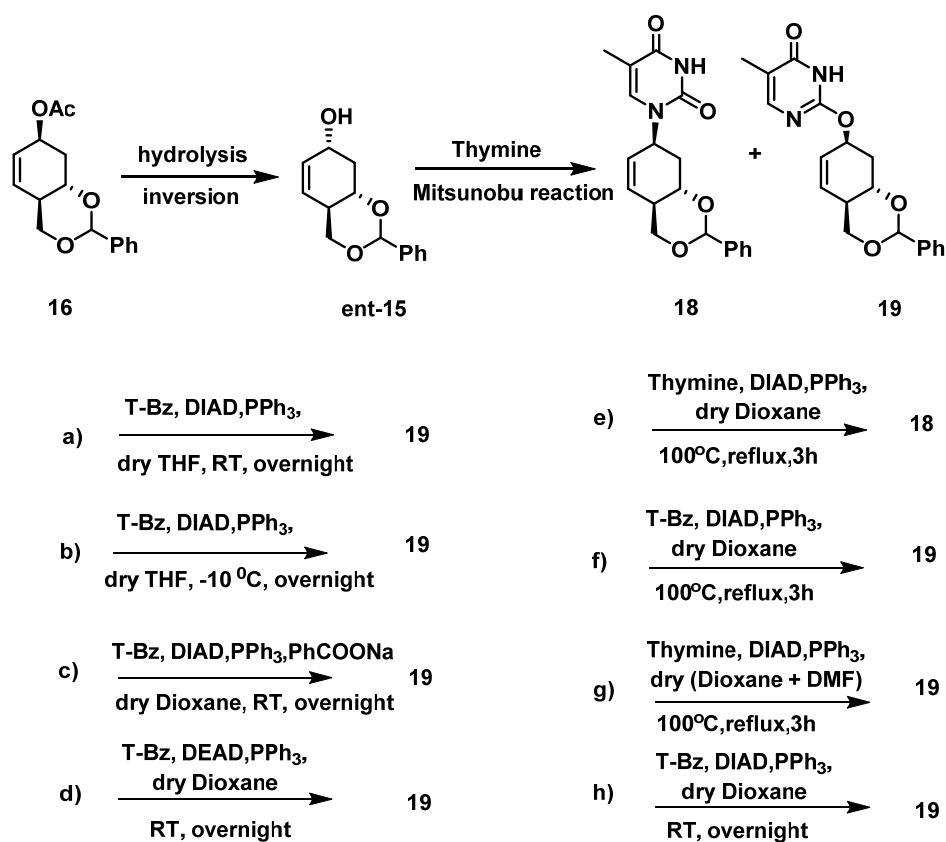
## Chapter 2

The enriched enantiomer **17** was further subjected for same enzymatic acylation conditions and pure enantiomer **17** was isolated by silica gel chromatography, which was not possible to synthesize by earlier reported synthetic route.<sup>13</sup>

### 2.7 Synthesis of cyclohexenyl thymine nucleoside

After successful synthesis of enantiopure intermediates ent-**15** and **17**, ent-**15** was used for the synthesis of cyclohexenyl nucleoside. Compound ent-**15** was subjected for Mitsunobu reaction conditions to introduce thymine base moiety onto the cyclohexenyl ring (scheme 9).

Scheme 9 Synthesis of cyclohexenyl thymine nucleoside



We observed O-alkylated product **19** instead of required N-alkylated product **18**. In literature also we found that in case of thymine there could be a competition between N-nucleophile and O-nucleophile. Initially we tried the reaction with N<sup>3</sup>-benzoyl protected thymine as nucleophile and diisopropyl azodicarboxylate as base in THF at room temperature, we observed the O-alkylated product **19** (scheme 9a). Then we

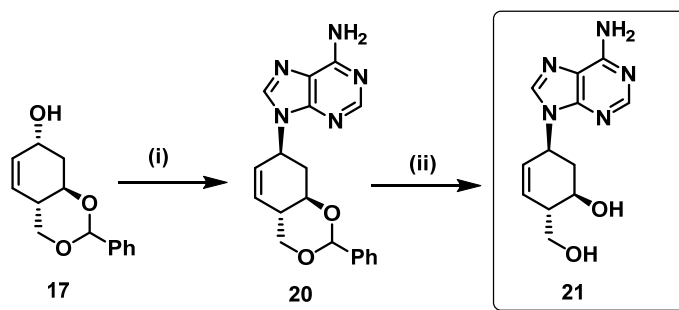
## Chapter 2

used the same reaction conditions with low temperature at  $-10^{\circ}\text{C}$ , was also yielded the undesired product (scheme 9b). We also used the known literature procedure where sodium benzoate used in Mitsunobu reaction was also ended with O-alkylated product (scheme 9c). Diethyl azodicarboxylate was used instead of diisopropyl azodicarboxylate and solvent changed to dioxane which also gave **19** (scheme 9d). Finally we obtained the desired N-alkylated product **18** with 30% yield by changing the reaction conditions to unprotected thymine as nucleophile and reflux at  $100^{\circ}\text{C}$  in dioxane (scheme 9e). To improve the yield, we changed the reaction conditions, which again ended with O-alkylated product (scheme 9f,9g,9h). Compounds **18** and **19** were confirmed by chemical shift value of the allylic proton in  $^1\text{H}$  NMR. The chemical shift value of the proton was 5.3 ppm for N-alkylated product, whereas in case of O-alkylated the value shifted to around 5.6 ppm.

### 2.8 Synthesis of Enantiopure $\alpha$ -L-Cyclohexenyl Adenine, Guanine Nucleosides

Synthesis of  $\alpha$ -L-cyclohexenyl adenine nucleoside was obtained from compound **17**. Enantiopure compound **17** was subjected for Mitsunobu reaction<sup>12</sup> in presence of adenine, DEAD, triphenylphosphine and dioxane was used as a solvent to yield **20** which was further treated with 80% aq acetic acid for deprotection of benzylidene group to give the  $\alpha$ -L-cyclohexenyl adenine **21** in 32% yield (scheme 10).

**Scheme 10** Synthesis of  $\alpha$ -L-cyclohexenyl adenine nucleoside



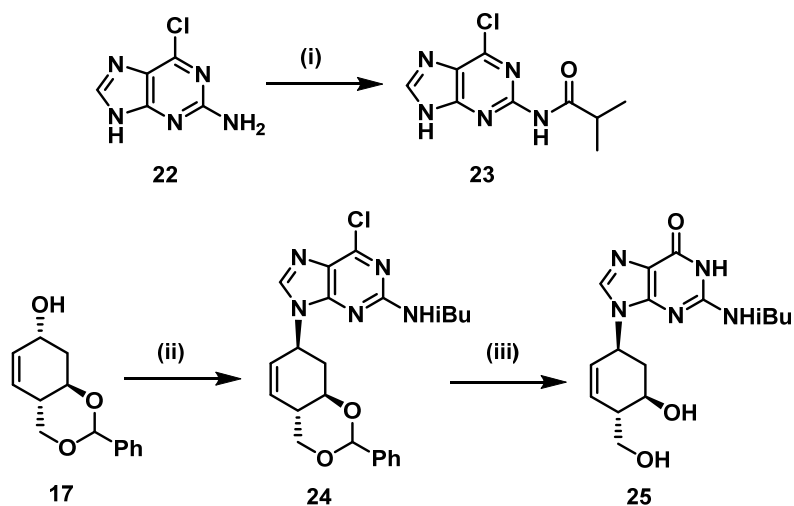
**Reagents and conditions:** (i) Adenine, DEAD, PPh<sub>3</sub>, dry THF, rt, overnight, 38 % (ii) 80% aq CH<sub>3</sub>COOH, 60 °C, 8 h, 65%.

To complete the synthesis of  $\alpha$ -L-cyclohexenyl guanine nucleoside initially we protected the amine group of 2-amino-6-chloropurine with isobutyryl group yielded **23** to avoid the formation of unwanted product with N-7 isomer in further nucleophilic substitution.<sup>23</sup> Compound **17** was subjected with protected 2-amino-6-

## Chapter 2

chloropurine **22** under Mitsunobu reaction condition to get **24** with 34% yield, which was again treated with aq acetic acid to yield the final  $\alpha$ -L-cyclohexenyl guanine nucleoside **25** (scheme 11).

**Scheme 11** Synthesis of  $\alpha$ -L-cyclohexenyl adenine nucleoside



**Reagents and conditions:** (i) Isobutyric anhydride, dry N,N dimethyl acetamide, 150 °C, 2h, 72 % (ii) **23**, DIAD, PPh<sub>3</sub>, dry THF, overnight, 34% (iii) 80% aq CH<sub>3</sub>COOH, 60 °C, 8 h, 68%.

## **2.9 Conclusions:**

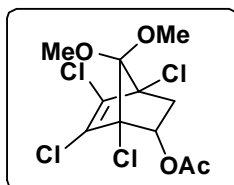
- Introduced a robust method for the synthesis of important cyclohexenyl- 2-deoxyribose sugar analogues in excellent yields from commercially available starting materials.
- Both cyclohexenyl and  $\alpha$ -L-cyclohexenyl nucleosides was achieved by using our designed synthetic route which was not possible with previous reports.
- An excellent enzymatic resolution method was accomplished for non polar compounds which can produce high yields and enantiomeric purity.
- Enantiopurity was established by using chiral shift reagent NMR as well as chiral HPLC.
- The easy access to these sugars as outlined in this chapter would allow further exploitation of the cyclohexenyl as well as cyclohexane nucleic acid analogs. These synthons will have applications not only in the synthesis of nucleoside/oligonucleotide analogues but also in carbohydrate chemistry where modified carbasugars as sugar mimics have potential applications.



## 2.10 Experimental Section

All the non-aqueous reactions were carried out under the inert atmosphere of Nitrogen/ Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). Column chromatographic separations were performed using silica gel 60-120 mesh (Merck) or 200- 400 mesh (Merck) and using the solvent systems EtOAc/Pet ether and MeOH/DCM. IR spectra were recorded on an infrared Fourier Transform spectrophotometer using chloroform or neat.  $^1\text{H}$  and  $^{13}\text{C}$  spectra were obtained using Bruker AC-200, AC-400 and AC-500 NMR spectrometers. The chemical shifts are reported in delta ( $\delta$ ) values and referred to internal standard TMS for  $^1\text{H}$ . Enzyme *Candida cylindracea* lipase was purchased from Ltd. Aldrich Inc.

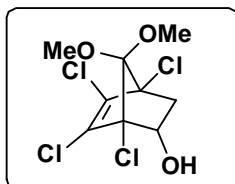
### ( $\pm$ )(1*S*,2*S*,4*R*)-1,4,5,6-tetrachloro-7,7-dimethoxy-bicyclo-[2.2.1]hept-5-en-2-yl acetate (**2**) :



To 5,5-dimethoxy-1,2,3,4-tetrachlorocyclopentadiene (12 mL, 45.45 mmol) was added vinyl acetate (7.82 mL, 90.9 mmol) and allowed to stir at 120°C for 5h. The excess vinyl acetate was removed *in vacuo* and the residue was purified by silica gel chromatography (pet ether:EtOAc = 97:3) to afford exclusively the endo adduct ( $\pm$ )-**3** (13.36 g) in 84% yield.

$^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  5.51 (dd, 1H,  $J = 7.7, 2.3$  Hz, CHOAc), 3.60 (s, 3H, OMe), 3.56 (s, 3H, OMe), 2.83 (dd, 1H,  $J = 12.6, 7.8$  Hz,  $\text{CH}_2$ ), 2.07 (s, 3H, OAc), 1.76 (dd, 1H,  $J = 12.8, 2.6$  Hz,  $\text{CH}_2$ );  $^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  170.2, 131.0, 127.8, 111.8, 76.4, 73.9, 52.6, 51.7, 43.8, 20.6; LCMS: mass calculated for  $\text{C}_{11}\text{H}_{12}\text{Cl}_4\text{O}_4\text{K}$  ( $\text{M} + \text{K}^+$ ) 386.91, observed 385.19.

### ( $\pm$ ) (1*S*,2*S*,4*R*)-1,4,5,6-tetrachloro-7,7-dimethoxybicyclo[2.2.1]hept-5-en-2-ol (**3**):



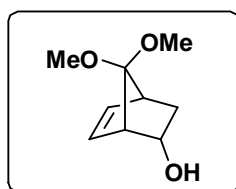
To a solution of ( $\pm$ )-**2** (20 g, 57.14 mmol) in methanol (150 mL) was added 5% aqueous  $\text{H}_2\text{SO}_4$  (40 mL) and stirred at 65°C for 5h. Methanol was removed on rotavapor *in vacuo*. The residue was diluted with EtOAc and water, saturated aqueous  $\text{NaHCO}_3$

## Chapter 2

and brine wash were given. Organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered, concentrated *in vacuo* and purified by silica gel chromatography (pet ether:EtOAc = 92:8) to result **3** (16.8 g) in 95% yield as a white solid,

**$^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):**  $\delta$  4.66 (m, 1H, CHOH), 3.58 (s, 3H, OMe), 3.55 (s, 3H, OMe), 2.67 (dd, 1H,  $J = 12.4, 8.0$  Hz,  $\text{CH}_2$ ), 2.12 (d, 1H,  $J = 4.8$  Hz), 1.79 (dd, 1H,  $J = 12.3, 2.4$  Hz,  $\text{CH}_2$ );  **$^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):**  $\delta$  130.8, 127.3, 112.0, 79.8, 76.4, 74.2, 52.5, 51.6, 44.2.

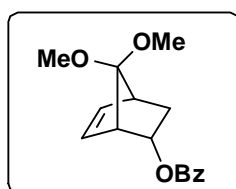
### ( $\pm$ )(1S,2S,4S)-7,7-dimethoxybicyclo[2.2.1]hept-5-en-2-ol (**4**) :



To a suspension of liquid ammonia (700 mL) and sodium (4.78 g, 207.8 mmol) at  $-78$  °C, solution of ( $\pm$ )-**3** (8 g, 25.97 mmol) in 10/1 mixture of THF (100 mL)/EtOH (10 mL) was added dropwise. After completion of addition, stirring was continued for 15 min, reaction mixture was quenched with saturated aqueous  $\text{NH}_4\text{Cl}$  solution and kept at room temperature overnight to allow liquid ammonia to evaporate. THF/EtOH was removed on rotavapor *in vacuo*, residue was diluted with DCM. Water wash and brine wash were given to the organic layer, dried over  $\text{Na}_2\text{SO}_4$ , filtered, concentrated *in vacuo* and purified by silica gel chromatography (pet ether:EtOAc, 85:15) to result ( $\pm$ )-**4** (3.4 g) in 78% yield as a pale yellow thick liquid.

**$^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):**  $\delta$  6.49 (m, 1H, Alkene CH), 6.10 (m, 1H, Alkene CH), 4.85 (s, 1H, OH), 4.56 (m, 1H, CHOH), 3.26 (merged, 1H, CH), 3.18 (s, 3H, OMe), 3.15 (s, 3H, OMe), 2.87 (m, 1H, CH), 2.42 (m, 1H,  $\text{CH}_2$ ), 0.85 (dd, 1H,  $J = 12.3, 2.2$  Hz,  $\text{CH}_2$ )  **$^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):**  $\delta$  137.9, 128.3, 119.2, 76.7, 70.2, 51.6, 50.7, 49.6, 45.6, 36.2.

### ( $\pm$ ) (1S,2S,4S)-7,7-dimethoxybicyclo[2.2.1]hept-5-en-2-yl benzoate (**7**) :



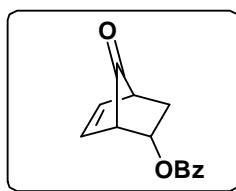
To a solution of ( $\pm$ )-**4** (9 g, 17.6 mmol) in pyridine (45 mL) was added benzoyl chloride (8.1 mL, 58.23 mmol) and the reaction mixture was stirred at rt for 4h. Pyridine was removed *in vacuo* and the residue was diluted with EtOAc. Water wash and brine

## Chapter 2

wash were given to the organic layer, dried over  $\text{Na}_2\text{SO}_4$ , concentrated in vacuo. The residue was purified by silica gel chromatography (pet ether:EtOAc, 89:11) to afford ( $\pm$ )-**7** (13.5 g) in 93% yield.

**$^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):**  $\delta$  7.99-7.94 (m, 2H, Aromatic), 7.55-7.27 (m, 3H, Aromatic), 6.43 (m, 1H, Alkene), 6.11 (m, 1H, Alkene), 5.61 (m, 1H, CHOBz), 3.41 (m, 1H, CH), 3.26 (s, 3H, OMe), 3.19 (s, 3H, OMe), 2.97 (m, 1H, CH), 2.58-2.48 (m, 1H,  $\text{CH}_2$ ), 1.18 (dd, 1H,  $J = 12.5, 2.4$  Hz,  $\text{CH}_2$ );  **$^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):**  $\delta$  166.4, 136.0, 133.6, 132.8, 130.1, 129.5, 128.3, 119.0, 73.8, 51.9, 49.9, 48.7, 45.2, 33.4; **HRMS:** mass calculated for  $\text{C}_{16}\text{H}_{18}\text{O}_4\text{Na}$  ( $\text{M} + \text{Na}^+$ ) 297.1102, observed 297.1113; **IR ( $\text{CHCl}_3$ ):**  $\nu_{\text{max}}$  2950, 1720, 1602  $\text{cm}^{-1}$

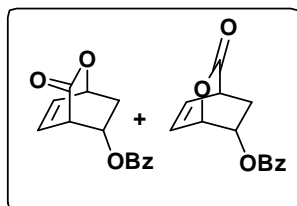
### ( $\pm$ ) (1*S*,2*S*,4*S*)-7-oxobicyclo[2.2.1]hept-5-en-2-yl benzoate (**8**) :



Solution of ( $\pm$ )-**7** (4.5 g, 16.4 mmol) in 180 mL of 6/1 mixture of acetic acid/ water was refluxed for 4h. Solvent was removed *in vacuo*, residue diluted with the EtOAc and water wash, saturated aqueous  $\text{NaHCO}_3$  wash and finally brine wash were given. The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , concentrated in vacuo and purified by silica gel chromatography (pet ether:DCM, 50:50) to give a mixture of **8** (81%).

**$^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):**  $\delta$  8.06 -7.94 (m, 2H, Aromatic), 7.58-7.39 (m, 3H, Aromatic), 6.79 (m, 1H, Alkene CH), 6.50 (m, 1H, Alkene CH), 5.68-5.60 (m, 1H, CHOBz), 3.57 (m, 1H, CH), 3.09 (m, 1H, CH), 2.66-2.53 (m, 1H,  $\text{CH}_2$ ), 1.49 (dd, 1H,  $J = 13.5, 3.0$  Hz,  $\text{CH}_2$ );  **$^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):**  $\delta$  201.0, 166.1, 134.4, 133.3, 133.2, 129.7, 129.5, 128.9, 128.4, 69.0, 51.8, 47.4 32.5; **LCMS:** mass calculated for  $\text{C}_{14}\text{H}_{12}\text{O}_3\text{Na}$  ( $\text{M} + \text{Na}^+$ ) 251.0786, observed 251.4641; **IR ( $\text{CHCl}_3$ ) :**  $\nu_{\text{max}}$  3064, 2947, 1790, 1720, 1600  $\text{cm}^{-1}$ .

### ( $\pm$ ) (1*S*,4*S*)-3-oxo-2-oxabicyclo[2.2.2]oct-5-en-8-yl benzoate (**9** + **10**) :

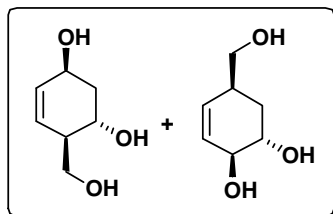


solution of ( $\pm$ )-**8** (5 g, 21.1 mmol) in 100 mL of dry DCM, was added  $\text{Na}_2\text{CO}_3$  (2.23 g, 21.1 mmol), stirred and cooled to  $0^\circ\text{C}$ . *m*-CPBA (5.2 g, 21.1 mmol) was added to this suspension and stirred it for 6h at rt. The

reaction mixture was quenched with 10% aq. solution of Na<sub>2</sub>S<sub>2</sub>O<sub>5</sub> (30 mL). Organic layer was separated and aqueous layer was extracted with DCM. The combined organic layer was washed with saturated aq NaHCO<sub>3</sub> followed by brine, and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The organic layer was concentrated in *vacuo* followed by silica gel chromatography (pet ether:EtOAc, 90:10) to give a mixture of **9** and **10** (92%) in 70:30 ratio.

**<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>):** δ 7.98 (m, 2H, Aromatic), 7.56-7.41 (m, 3H, Aromatic), 6.78 (m, 1H, minor), 6.72 (m, 1H, major) 6.54 (m, 1H), 5.60 (m, 1H, minor), 5.48 (m, 1H, major), 5.42 (m, 1H, minor), 5.32 (m, 1H, major), 4.05 (m, 1H, major), 3.57 (m, 1H, minor), 2.85 (m, 1H, major), 2.60 (m, 1H, minor), 1.75 (d, *J* = 6.0 Hz, 1H, major), 1.59 (m, 1H, minor); **<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>):** δ 172.2, 170.7, 165.5, 134.3, 133.4, 133.3, 132.5, 129.5, 129.3, 129.0, 128.9, 128.4, 128.2, 76.6, 73.5, 73.0, 68.8, 65.5, 46.2, 40.3, 35.0, 29.3; **HRMS:** mass calculated for C<sub>14</sub>H<sub>13</sub>O<sub>4</sub> (M+H<sup>+</sup>) 245.0813, observed 245.0812; **IR (CHCl<sub>3</sub>):** ν<sub>max</sub> 1759, 1716 cm<sup>-1</sup>.

**(±) (1*S*,3*S*,6*R*)-6- (hydroxymethyl) cyclohex -4-ene-1,3-diol (**11**+ **12**) :**



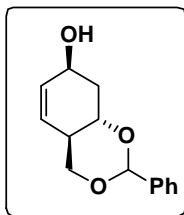
To a solution of mixture **9+10** (3.3g, 13.5mmol) in dry THF (400mL) at -15°C, LAH (1.5g, 40.5 mmol) was added and the resulting mixture was stirred at the same temperature for 2h. The reaction mixture was then quenched cautiously with ethyl acetate (50 mL)

followed by aqueous saturated solution of Na<sub>2</sub>SO<sub>4</sub>, to precipitate out aluminium salts. After the filtration, filtrate was concentrated *in vacuo*, to result inseparable epimeric mixture of **11** and **12** in 85% yield which was used further without separation.

**<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) of compound (**11+12**):** δ 5.86-5.72(m, 2H, Alkene CH), 4.38-4.32 (m, 1H, allylic CH-O), 4.14-4.08 (m, 1H, CH<sub>2</sub>O), 3.87-3.79 (m, 1H, CH<sub>2</sub>O), 3.68-3.60 (m, 1H, CH-O), 2.50 (m, 1H, Allylic CH), 2.11-2.02 (m, 1H, CH<sub>2</sub>), 1.88-1.73 (m, 1H, CH<sub>2</sub>); **<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>):** δ 130.5, 127.9, 66.1, 65.4, 61.7, 41.5, 35.3. **LCMS:** mass calculated for C<sub>7</sub>H<sub>12</sub>O<sub>3</sub>Na (M+ Na<sup>+</sup>) 167.06, observed 167.17.

## Chapter 2

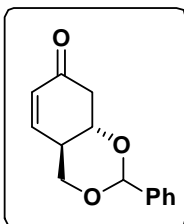
(±)(4a*R*,7*S*,8a*S*) -2-phenyl -4a,7,8,8a -tetrahydro- 4H-benzo[d][1,3]dioxin-7-ol (13):



Mixture of compounds **11** and **12** (3g, 21mmol) was dissolved in dry dioxane, and benzaldehyde dimethylacetal (6 mL, 27 mmol), PTSA (200mg, 1.05mmol) was added to it slowly and the reaction was stirred at rt for 24h. Reaction mixture was quenched with ice and stirred for 30 min. Extracted with EtOAc three times, combined organic layer was washed with water and brine solution, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated in *vacuo*, followed by silica gel chromatography (pet ether:EtOAc 88:12) to afford (±)-**13** in 70 % yield.

**<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>):** δ 7.55-7.50 (m, 2H,Aromatic), 7.41-7.36 (m, 3H,Aromatic), 5.89-5.83 (m, 1H,Alkene CH), 5.65 (s,1H,CHPh), 5.53 (dd, 1H, *J* =9.7, 1.5 Hz, Alkene CH), 4.41 (m, 1H,CHOH), 4.34 (dd, 1H, *J* =10.7, 4.5 Hz,CH<sub>2</sub>O), 3.89-3.83 (m, 1H,CH-O), 3.70 (t, 1H, *J* = 11.4 Hz, CH<sub>2</sub>O ), 2.48-2.46 (m, 1H,OH), 2.25-2.16 (m, 1H,CH), 2.03-1.87 (m, 2H, CH<sub>2</sub>); **<sup>13</sup>C NMR (50 MHz,CDCl<sub>3</sub>):** δ 138.1, 130.2, 128.9, 128.3, 126.8, 126.1, 102.3, 75.3, 70.6, 65.2, 40.4, 37.2; **HRMS:** mass calculated for C<sub>14</sub>H<sub>17</sub>O<sub>3</sub> (M+ H<sup>+</sup>) 233.1177, observed 233.1173.

(±) (4a*R*,8a*S*)-2-phenyl-8,8a-dihydro-4H-benzo[d][1,3]dioxin-7-one (14) :



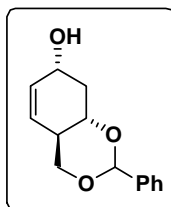
To a solution of CrO<sub>3</sub> (90 mg, 0.9 mmol), Ac<sub>2</sub>O (0.085 mL, 0.9 mmol), Pyridine (0.14 mL, 1.81 mmol) in dry DCM (15 mL), was added 5mL solution of **13** (200 mg, 0.9 mmol) in DCM stirred for 1.5h at rt. Reaction mixture was filtered on celite and purified by silica gel chromatography (pet ether:EtOAc = 91:9) to afford **14** (185 mg) in 92% yield.

**<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>):** δ 7.55 (m, 2H,Aromatic), 7.43-7.38 (m, 3H,Aromatic), 6.63 (dd, 1H, *J*= 9.8, 1.8 Hz, Alkene CH), 6.18 (m,1H, Alkene CH), 5.64 (s, 1H,CHPh), 4.49 (dd, 1H, *J*= 10.9, 4.5, CH<sub>2</sub>O), 4.07 (m, 1H,CH-O), 3.81 (t, 1H, *J* = 11.3 Hz, CH<sub>2</sub>O), 3.02-2.86 (m, 2H, CH<sub>2</sub>,CH), 2.68 (dd,1H, *J*=16.4, 12.8Hz,CH<sub>2</sub>); **<sup>13</sup>C NMR (50MHz,CDCl<sub>3</sub>):** δ 196.8, 145.0, 137.5, 132.1, 129.2, 128.4, 126.1, 101.7,

## Chapter 2

76.5, 69.3, 44.5, 40.0; **LCMS**: mass calculated for  $C_{14}H_{14}O_3$  ( $M^+ K^+$ ) 269.0943, observed 269.1373; **IR** ( $CHCl_3$ ):  $\nu_{max}$  1720, 1610  $cm^{-1}$ .

### (±)(4a*R*,7*R*,8a*S*)-2-phenyl-4a,7,8,8a-tetrahydro-4H-benzo[d][1,3]dioxin-7-ol(**15**):

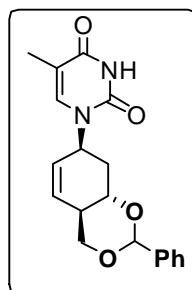


To a solution of **14** (165 mg, 0.717 mmol) in 10 mL dry MeOH, was added  $CeCl_3 \cdot 7H_2O$  (400 mg, 1.07 mmol) stirred for 1h at rt.  $NaBH_4$  (33 mg, 0.86 mmol) was added in portions, stirred for 2h at rt. Reaction was quenched with crushed ice and stirred for 30 min. Reaction mixture concentrated *in vacuo*, residue was dissolved in

EtOAc and washed with water, brine. EtOAc layer was dried over  $Na_2SO_4$  concentrated *in vacuo* and purified by silica gel chromatography (pet ether:EtOAc = 91:9) to afford **15** (135 mg) in 82% yield.

**<sup>1</sup>H NMR (200 MHz,  $CDCl_3$ )**:  $\delta$  7.52 (m, 2H, Aromatic), 7.39 (m, 3H, Aromatic), 5.79 (m, 1H, Alkene CH), 5.61 (s, 1H, CHPh), 5.47 (m, 1H, Alkene CH), 4.55 (m, 1H, CHOH), 4.32 (dd, 1H,  $J = 10.8, 4.4$  Hz,  $CH_2O$ ), 3.76-3.57 (m, 2H, CH-O,  $CH_2O$ ), 2.59-2.49 (m, 2H,  $CH_2, CH$ ), 1.89-1.73 (m, 1H,  $CH_2$ ); **<sup>13</sup>C NMR (50 MHz,  $CDCl_3$ )**:  $\delta$  138.0, 132.8, 129.0, 128.4, 126.2, 124.8, 102.17, 76.5, 70.7, 67.7, 40.0, 38.3; **LCMS**: mass calculated for  $C_{14}H_{16}O_3Na$  ( $M^+ Na^+$ ) 255.15, observed 255.37.

### 5-methyl-1-((4a*R*,7*S*,8a*S*)-2-phenyl-4a,7,8,8a-tetrahydro-4H-benzo[d][1,3]dioxin-7-yl)pyrimidine-2,4(1*H*,3*H*)-dione (**18**):

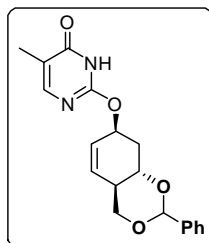


To a solution of pure ent-**15** (300 mg, 1.28 mmol) in 10mL dry dioxane, was added thymine (320 mg, 2.2 mmol) and triphenylphosphine (660 mg, 2.2 mmol), stirred for 15 min at rt. Reaction mixture cooled to ice bath temperature and DIAD added slowly then ice bath removed and reflux the reaction mixture at 100 °C for 3h. Reaction mixture concentrated under reduced pressure and residue was dissolved in EtOAc and washed with water, brine. EtOAc layer was dried over  $Na_2SO_4$  concentrated *in vacuo* and purified by silica gel chromatography (pet ether:EtOAc = 70:30) to afford **18** (138 mg) in 30% yield.

## Chapter 2

**<sup>1</sup>H NMR of 17 (400 MHz, CDCl<sub>3</sub>)** δ 1.97 (s, 3 H), 2.19 - 2.26 (m, 2 H), 2.52 - 2.62 (m, 1 H), 3.64 (s, 1 H), 3.79 (t, *J*=11.17 Hz, 1 H), 4.38 - 4.44 (m, 1 H), 5.31 (br. s., 1 H), 5.63 (s, 1 H), 5.66 - 5.72 (m, 1 H), 5.94 (d, *J*=9.54 Hz, 1 H), 7.18 (s, 1 H), 7.33 - 7.41 (m, 3 H), 7.44 - 7.59 (m, 9 H), 7.69 (d, *J*=7.28 Hz, 3 H), 7.67 (d, *J*=8.03 Hz, 2 H), 8.31 (br. s., 1 H) ppm; **<sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)** δ 12, 14.2, 34.8, 39.7, 53.0, 70.2, 102.1, 111.8, 126.1, 127.7, 128.3, 129.1, 130.1, 130.4, 131.6, 135.0, 136.1, 137.8, 149.8, 162.7, 168.9 ppm; **LCMS:** mass calculated for C<sub>19</sub>H<sub>20</sub>N<sub>2</sub>O<sub>4</sub>Na (M+ Na<sup>+</sup>) 363.1423, observed 363.08.

### **<sup>1</sup>H NMR 19 (500 MHz, CDCl<sub>3</sub>)**

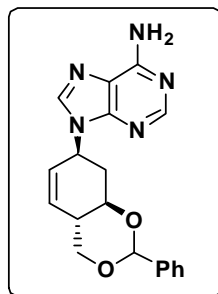


Exact Mass: 340.1423

δ 1.82 - 1.90 (m, 1 H), 1.99 (s, 3 H), 2.59 - 2.66 (m, 1 H), 2.69 - 2.77 (m, 1 H), 3.71 (t, *J*=11.14 Hz, 1 H), 3.85 (ddd, *J*=12.13, 9.08, 3.20 Hz, 1 H), 4.35 (dd, *J*=10.99, 4.58 Hz, 1 H), 5.51 - 5.57 (m, 1 H), 5.64 (s, 2 H), 5.80 (d, *J*=9.77 Hz, 1 H), 7.10 (d, *J*=1.22 Hz, 1 H), 7.35 - 7.41 (m, 3 H), 7.47 - 7.54 (m, 5 H), 7.63 - 7.68 (m, 1 H), 7.93 (dd, *J*=8.39, 1.37 Hz, 2 H) ppm; **LCMS:** mass

calculated for C<sub>19</sub>H<sub>20</sub>N<sub>2</sub>O<sub>4</sub>K (M+ K<sup>+</sup>) 379.1423, observed 379.06.

### **9-((4a*S*,7*S*,8a*R*)-2-phenyl-4a,7,8,8a-tetrahydro-4*H*-benzo[*d*][1,3]dioxin-7-yl)-9*H*-purin-6-amine (20) :**



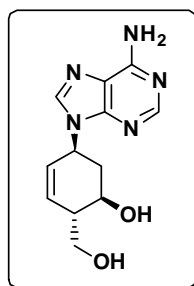
To a solution of pure enantiomer **17** (150 mg, 0.64 mmol) in 10mL dry dioxane, was added adenine (175 mg, 1.28 mmol) and triphenylphosphine (340 mg, 1.28 mmol), stirred for 15 min at rt. Reaction mixture cooled to ice bath temperature and DEAD (0.02 mL, 1.28 mmol) added slowly the removed the ice bath and stirred the reaction mixture at rt for overnight. Reaction

mixture concentrated under reduced pressure and residue was dissolved in EtOAc and washed with water, brine. EtOAc layer was dried over Na<sub>2</sub>SO<sub>4</sub> concentrated *in vacuo* and purified by silica gel chromatography (methanol:DCM = 5:95) to afford **20** (86 mg) in 38% yield.

**<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)** δ 2.04 (d, *J*=7.63 Hz, 1 H), 2.16 (q, *J*=11.90 Hz, 1 H), 2.74 - 2.86 (m, 2 H), 3.75 (t, *J*=11.29 Hz, 1 H), 3.95 (ddd, *J*=12.21, 9.16, 3.05 Hz, 1

H), 4.37 (dd,  $J=10.83, 4.43$  Hz, 1 H), 5.59 (tdd,  $J=8.74, 8.74, 4.20, 2.14$  Hz, 1 H), 5.67 (s, 1 H), 5.74 - 5.83 (m, 2 H), 7.33 - 7.41 (m, 3 H), 7.51 (dd,  $J=7.78, 1.68$  Hz, 2 H), 7.87 (s, 1 H), 8.37 (s, 1 H) ppm;  $^{13}\text{C}$  NMR (126 MHz,  $\text{CDCl}_3$ )  $\delta$  36.3, 39.8, 51.5, 70.4, 102.2, 114.1, 119.6, 126.1, 127.7, 128.4, 128.6, 129.1, 137.8, 138.4, 139.3, 149.7, 153.0, 155.6 ppm; LCMS: mass calculated for  $\text{C}_{19}\text{H}_{20}\text{N}_5\text{O}_2$  ( $\text{M}^+ \text{H}^+$ ) 350.16, observed 350.12.

**(1R,2S,5S)-5-(6-amino-9H-purin-9-yl)-2-(hydroxymethyl)cyclohex-3-en-1-ol (21):**

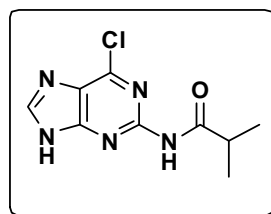


Compound **20** (100 mg, 0.27 mmol) was dissolved in 80% aq  $\text{CH}_3\text{COOH}$  solution and stirred reaction mixture for 8h at 60 °C. Reaction mixture concentrated under reduced pressure and residue was dissolved in EtOAc and washed with water, brine. EtOAc layer was dried over  $\text{Na}_2\text{SO}_4$  concentrated *in vacuo* and purified by silica gel chromatography (methanol:DCM = 10:90) to afford

**21** (47 mg) in 65% yield.

$^1\text{H}$  NMR (200 MHz,  $\text{DMSO}-d_6$ )  $\delta$  1.95 - 2.08 (m, 2 H), 2.23 (br. s., 2 H), 3.61 - 3.82 (m, 3 H), 4.67 (br. s., 1 H), 4.95 (br. s., 1 H), 5.21 (br. s., 1 H), 5.69 (d,  $J=8.46$  Hz, 1 H), 5.92 (d,  $J=8.72$  Hz, 1 H), 7.25 (br. s., 2 H), 8.13 (br. s., 2 H) ppm; LCMS: mass calculated for  $\text{C}_{12}\text{H}_{15}\text{N}_5\text{O}_2\text{K}$  ( $\text{M}^+ \text{K}^+$ ) 300.13, observed 300.98.

**N-(6-chloro-9H-purin-2-yl)isobutyramide (23) :**



To a solution of commercially available 2-amino-6-chloropurine (1g, 6 mmol) in  $\text{N,N}$ -dimethylacetal 30 mL, was added isobutyric anhydride (2.7 mL, 15mmol) and reflux the reaction mixture for 2h at 150 °C, cooled to room temperature and evaporated under reduced pressure to 1/10

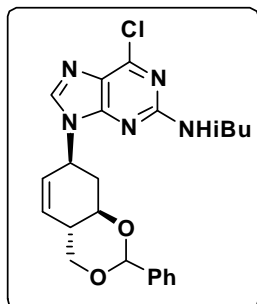
of its volume. The precipitated crude product was collected and crystallized from boiling of ethanol/water (1:1, 200mL) to yield **23** (710 mg).

$^1\text{H}$  NMR (200 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  1.23 (d,  $J=6.95$  Hz, 6 H), 2.71 - 2.88 (m, 1 H), 8.40 (s, 1 H) ppm; LCMS: mass calculated for  $\text{C}_9\text{H}_{10}\text{ClN}_5\text{ONa}$  ( $\text{M}^+ \text{Na}^+$ ) 262.04, observed 261.84.



**N-(6-chloro-9-((4a*S*,7*S*,8a*R*)-2-phenyl-4a,7,8,8a-tetrahydro-4*H*benzo[*d*][1,3]dioxin-7-yl)-9*H*-purin-2-yl)isobutyramide (**24**) :**

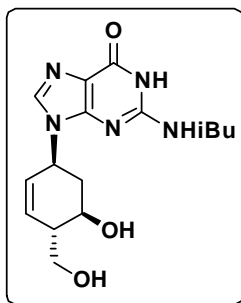
To a solution of **17** (100 mg, 0.43 mmol) in 10mL dry THF, was added **23** (205 mg, 0.86 mmol) and triphenylphosphine (225 mg, 0.86 mmol), stirred for 15 min at rt. Reaction mixture cooled to ice bath temperature and DIAD added slowly the removed the ice bath and reflux the reaction mixture stirred at room temperature for overnight.



Reaction mixture concentrated under reduced pressure and residue was dissolved in EtOAc and washed with water, brine. EtOAc layer was dried over Na<sub>2</sub>SO<sub>4</sub> concentrated *in vacuo* and purified by silica gel chromatography (pet ether:EtOAc = 70:30) to afford **24** (67 mg) in 35% yield.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 1.28 (s, 3 H), 1.32 (s, 3 H), 2.71 - 3.01 (m, 3 H), 3.75 (t, *J*=11.18 Hz, 1 H), 3.96 (ddd, *J*=12.22, 9.19, 3.09 Hz, 1 H), 4.38 (dd, *J*=10.80, 4.48 Hz, 1 H), 5.57 - 5.71 (m, 2 H), 5.79 (s, 2 H), 8.07 (s, 1 H), 8.20 (s, 1 H) ppm; <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>) δ 19.3, 36.2, 39.7, 52.0, 70.3, 76.4, 102.2, 126.1, 127.0, 128.4, 128.6, 129.1, 129.3, 131.5, 132.0, 132.2, 133.6, 137.7, 142.6, 152.0, 152.4, 160.3, 175.5 ppm; LCMS: mass calculated for C<sub>23</sub>H<sub>26</sub>ClN<sub>5</sub>O<sub>2</sub>Na (M+ Na<sup>+</sup>) 462.94.13, observed 463.25.

**N-(9-((1*S*, 4*S*, 5*R*)-5-hydroxy-4-(hydroxymethyl)cyclohex-2-en-1-yl)-6-oxo-6,9 dihydro-1*H*-purin-2-yl)isobutyramide (**25**) :**

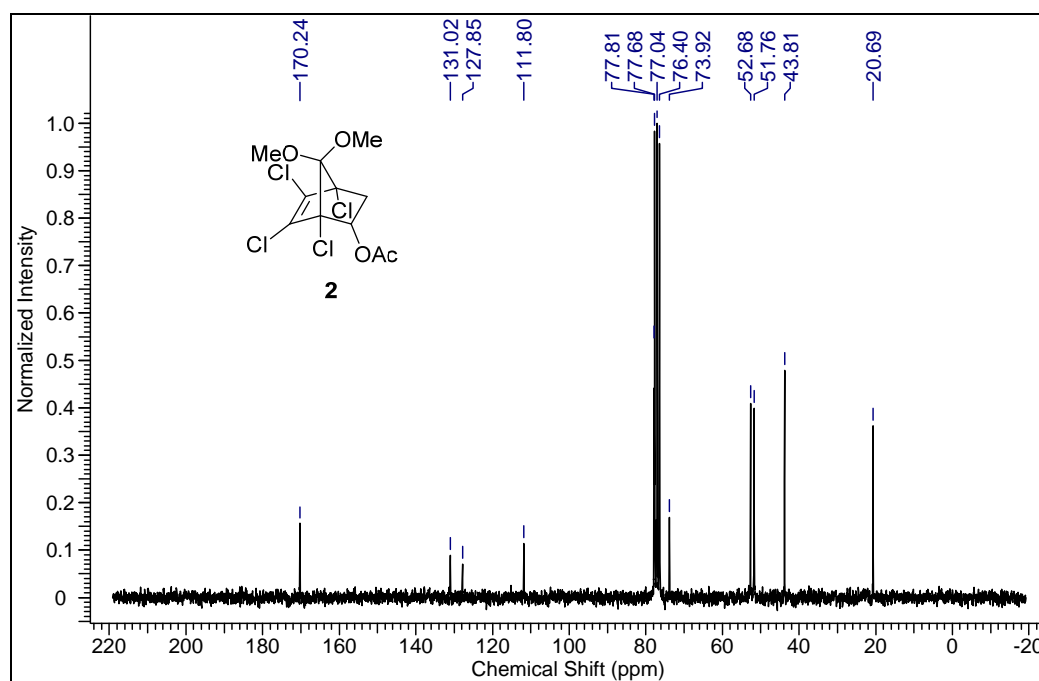
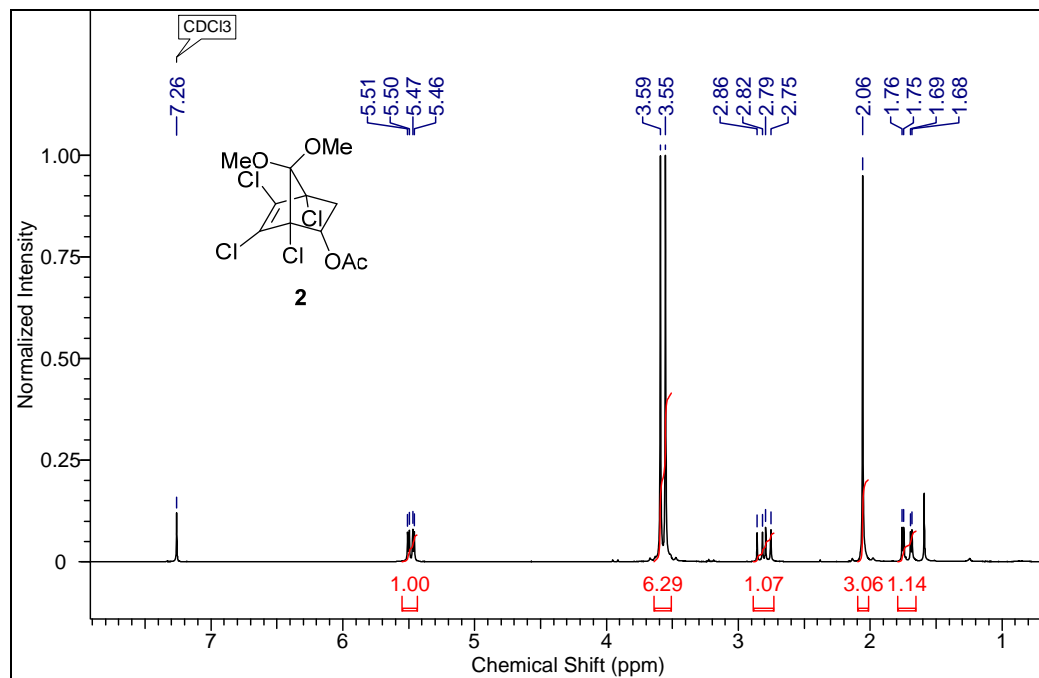


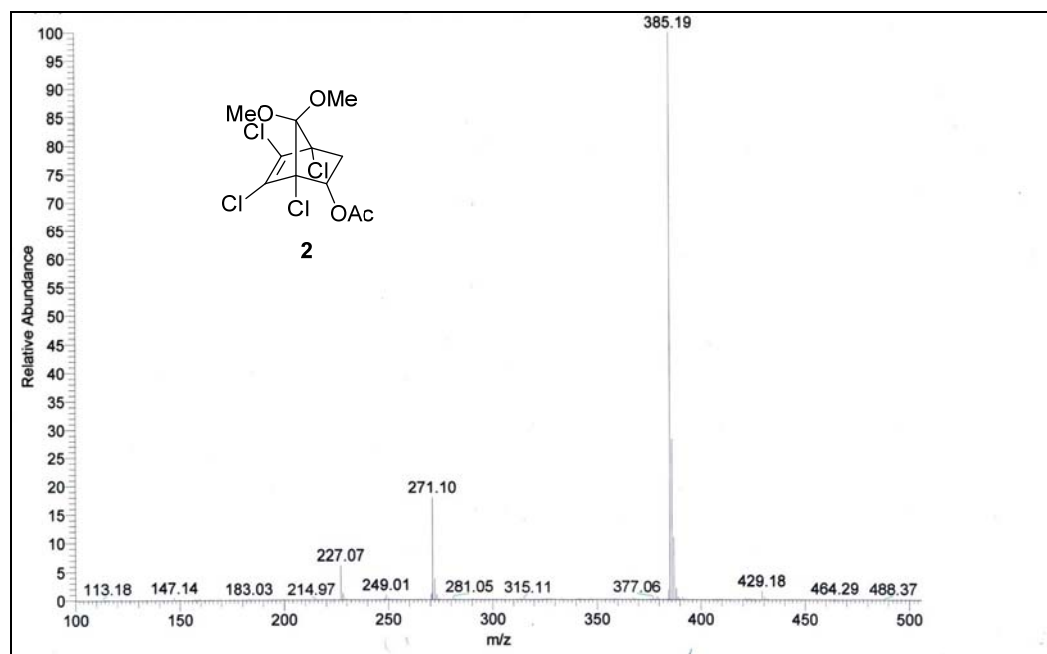
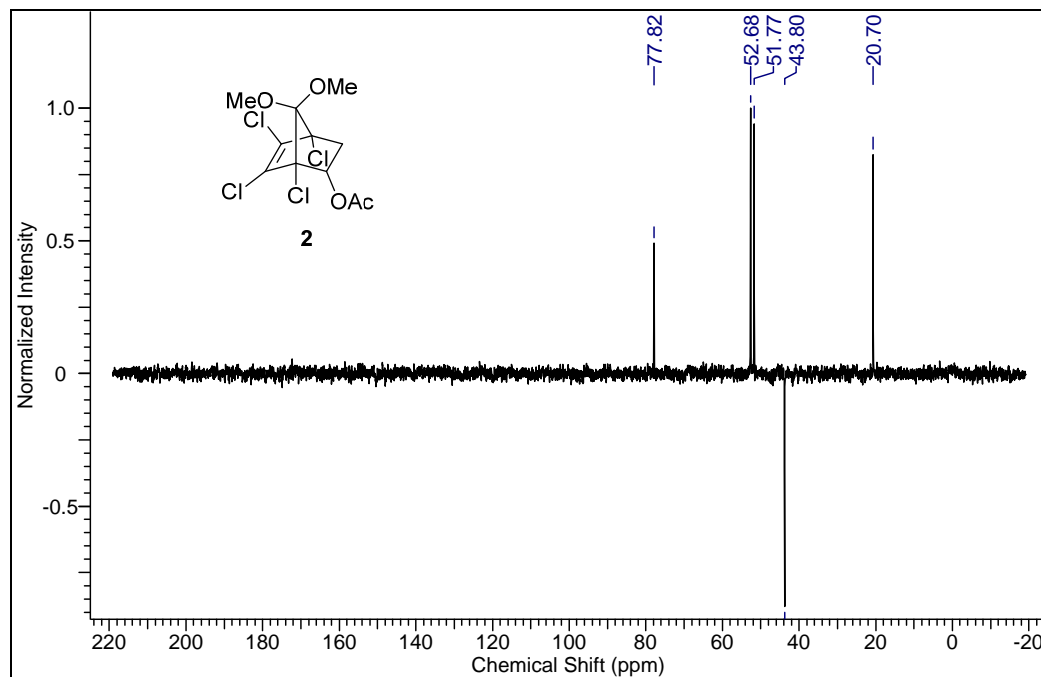
Compound **24** (100 mg, 0.22 mmol) was dissolved in 80% aq CH<sub>3</sub>COOH solution and stirred reaction mixture for 8h at 60 °C. Reaction mixture concentrated under reduced pressure and residue was purified by silica gel chromatography (methanol: DCM = 10:90) to afford **25** (52 mg) in 68% yield.

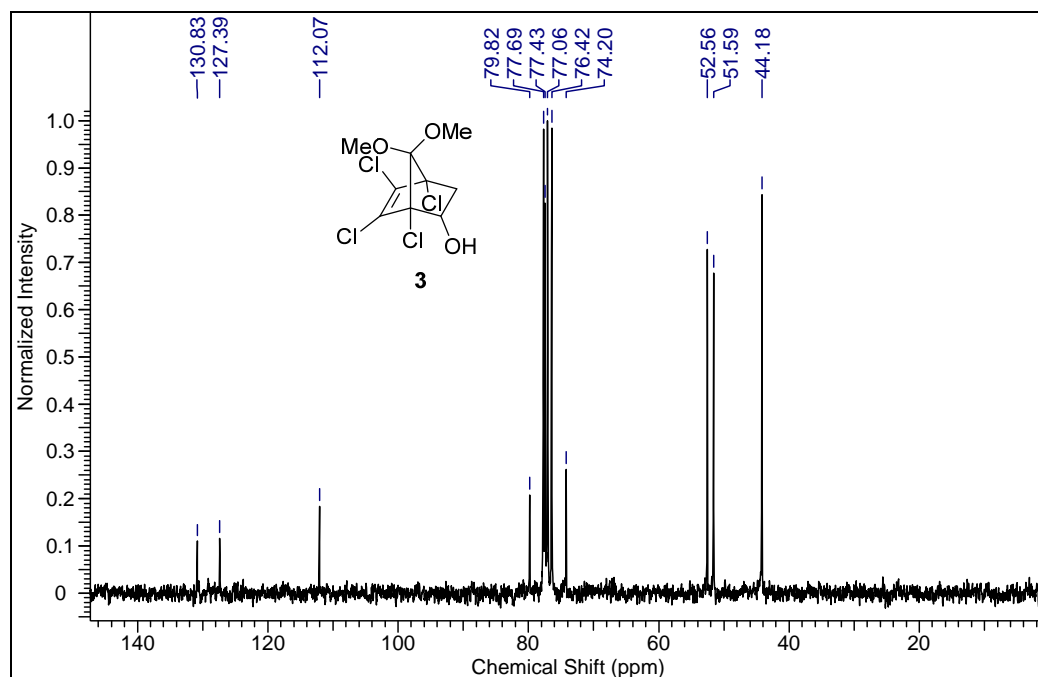
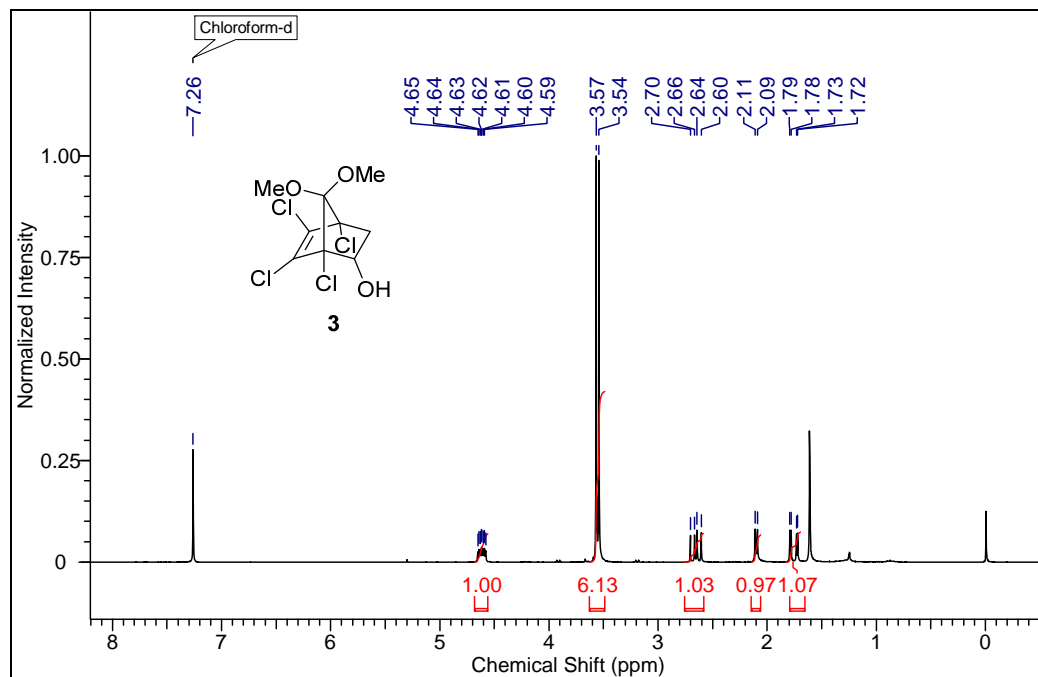
<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 1.28 (s, 7 H), 2.28 - 2.53 (m, 2 H), 2.53 - 2.83 (m, 2 H), 3.60 - 3.96 (m, 3 H), 5.18 - 5.35 (m, 1 H), 5.81 (d, *J*=10.23 Hz, 1 H), 6.01 (dt, *J*=10.11, 2.27 Hz, 1 H), 8.02 (br. s., 1 H) ppm; LCMS: mass calculated for C<sub>16</sub>H<sub>21</sub>N<sub>5</sub>O<sub>4</sub>Na (M+ Na<sup>+</sup>) 370.37, observed 370.11.

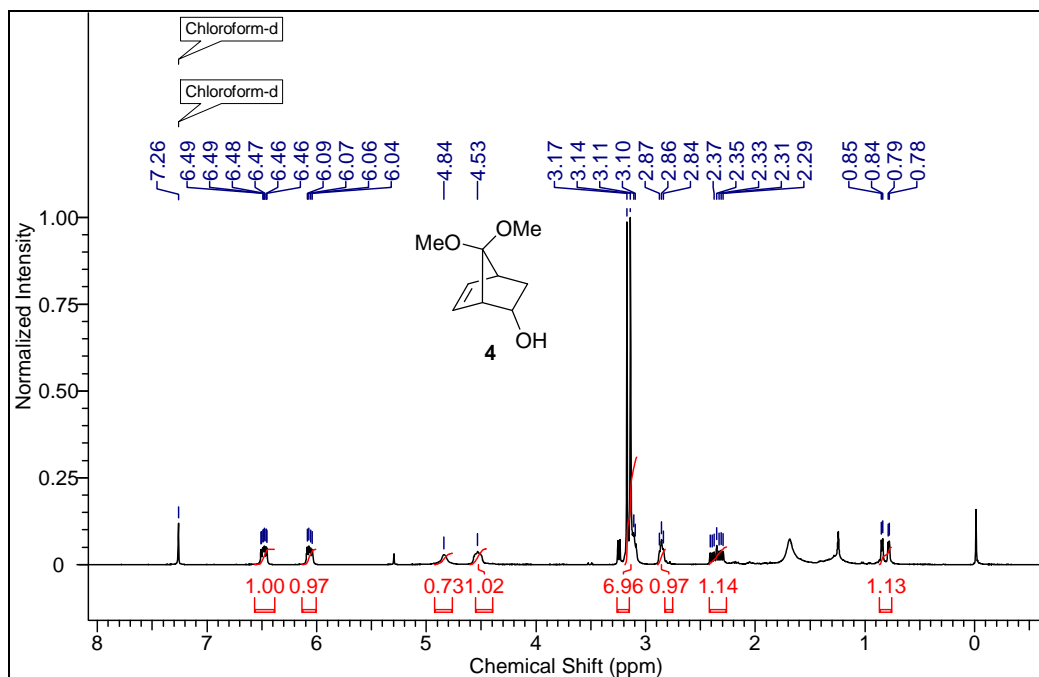
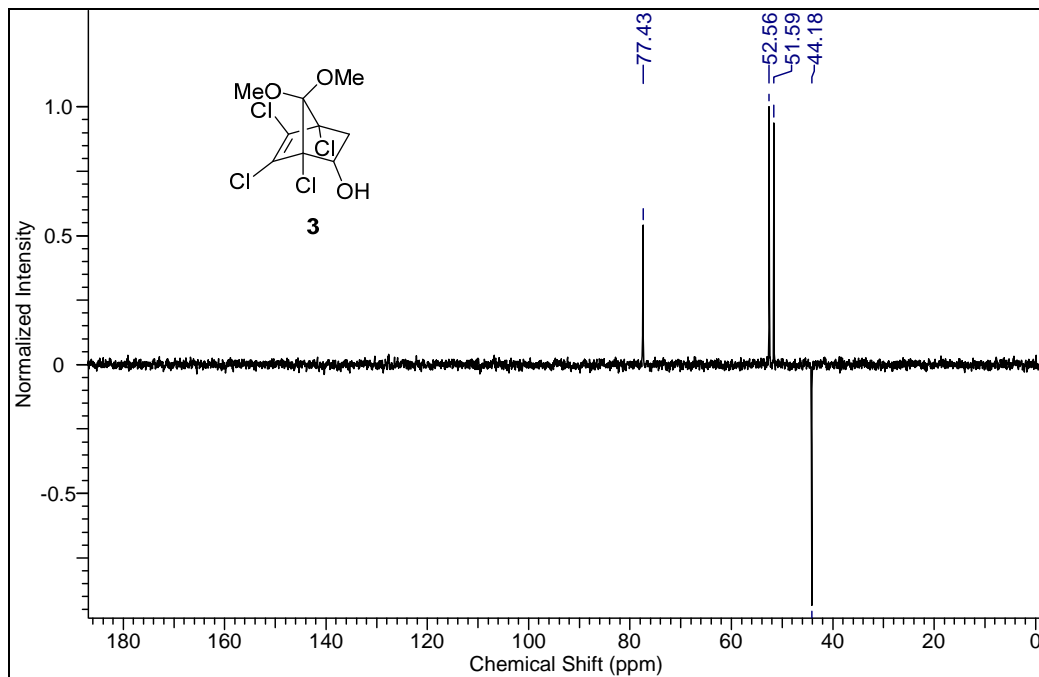
## 2.11 Appendix

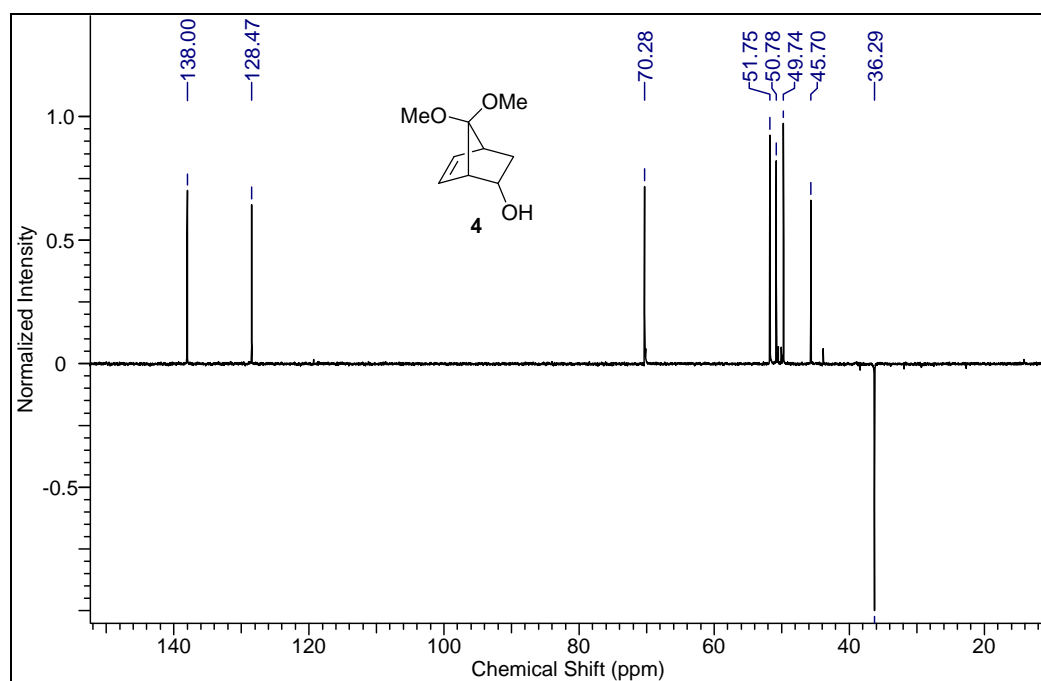
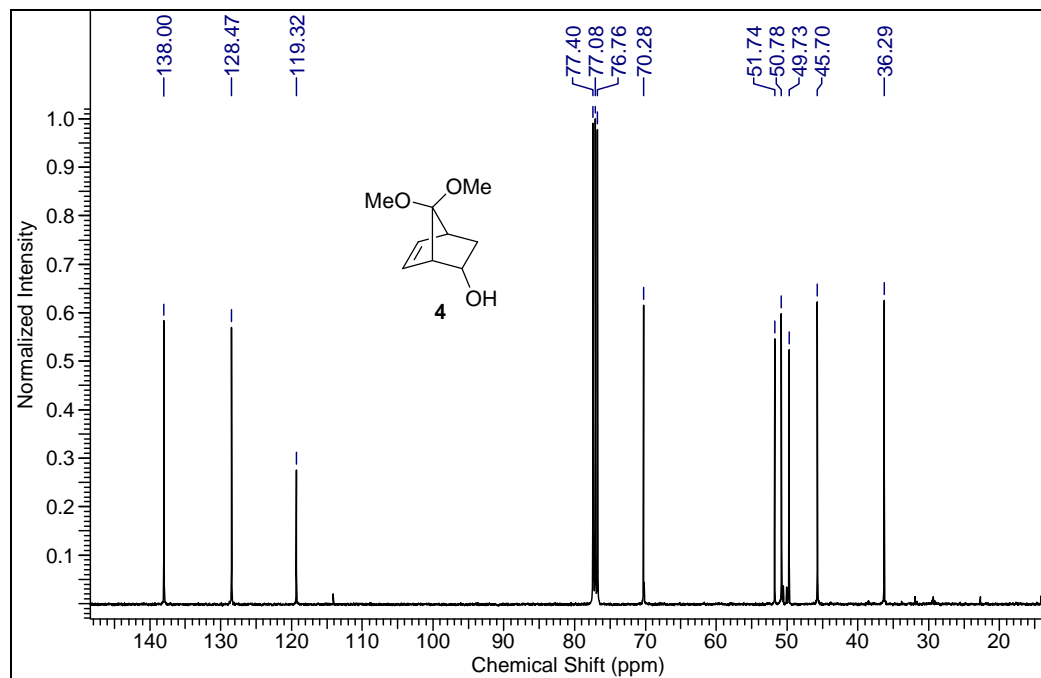
Compounds - Spectral data	Page No.
2 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	57
2 DEPT NMR & LC MS	58
3 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	59
3 DEPT NMR & 4 $^1\text{H}$ NMR	60
4 $^{13}\text{C}$ NMR & DEPT NMR	61
7 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	62
7 DEPT NMR & LC MS	63
8 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	64
8 DEPT NMR & LC MS	65
9 + 10 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	66
9 + 10 COSY	67
9 + 10 LC MS	68
12 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	69
12 DEPT NMR & LC MS	70
13 $^1\text{H}$ NMR & $^{13}\text{C}$ LC MS MR	71
13 DEPT NMR & LC MS	72
14 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	73
14 DEPT NMR & LC MS	74
15 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	75
15 DEPT NMR & LC MS	76
18 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	77
18 DEPT NMR & LC MS	78
19 $^1\text{H}$ NMR & LC MS	79
20 $^1\text{H}$ NMR & $^{13}\text{C}$ NMR	80
20 DEPT NMR & 21 $^1\text{H}$ NMR	81
21 LC MS & 23 $^1\text{H}$ NMR	82
23 LC MS & 24 $^1\text{H}$ NMR	83
24 $^{13}\text{C}$ NMR & DEPT NMR	84
24 LC MS & 25 $^1\text{H}$ NMR	85
25 $^{13}\text{C}$ NMR & DEPT NMR	86
25 LC MS	87

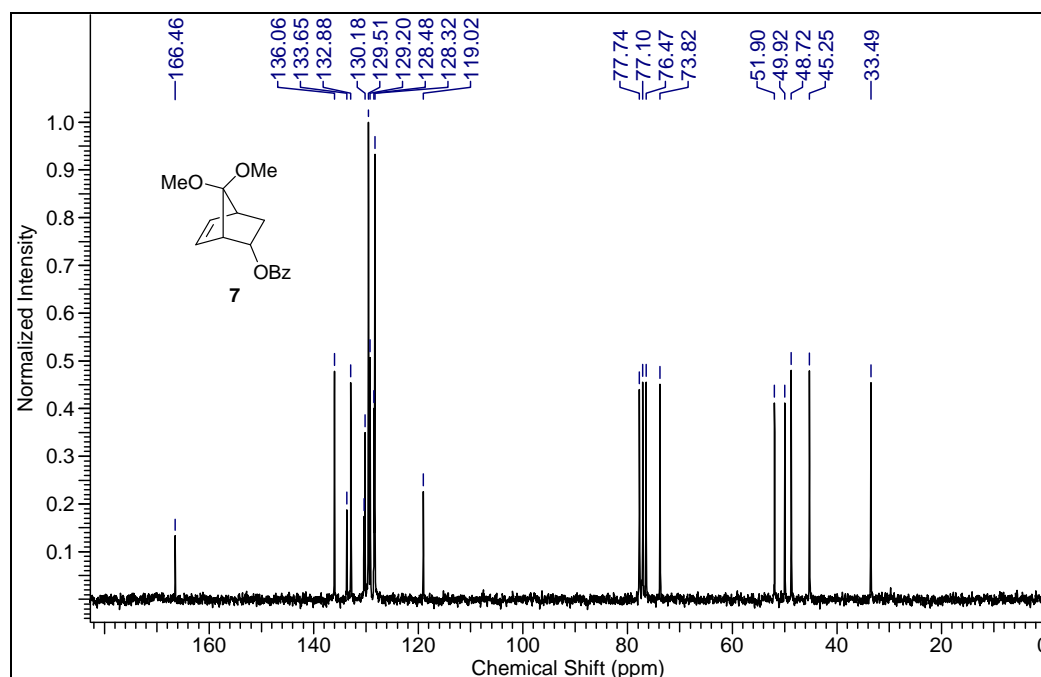
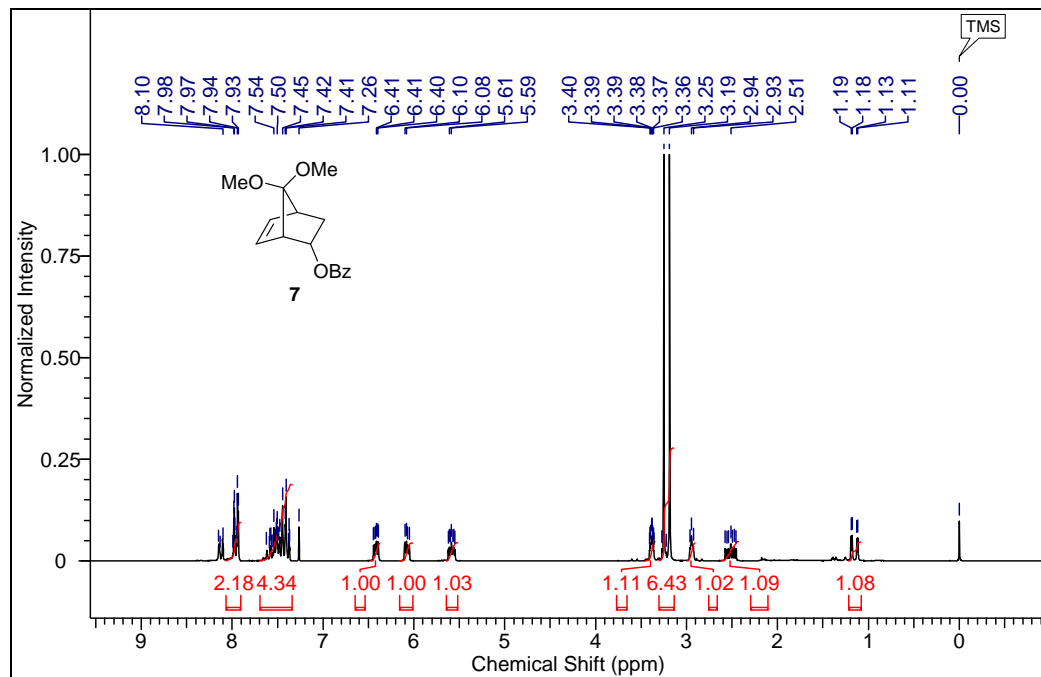




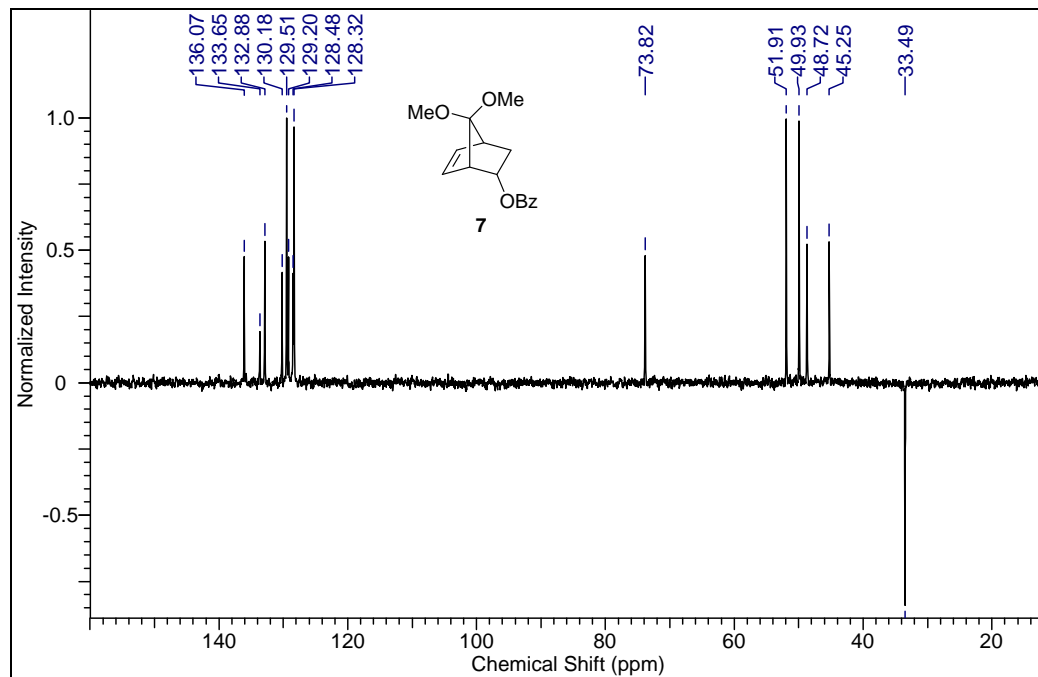




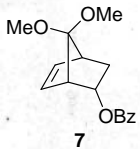








**Elemental composition calculator**



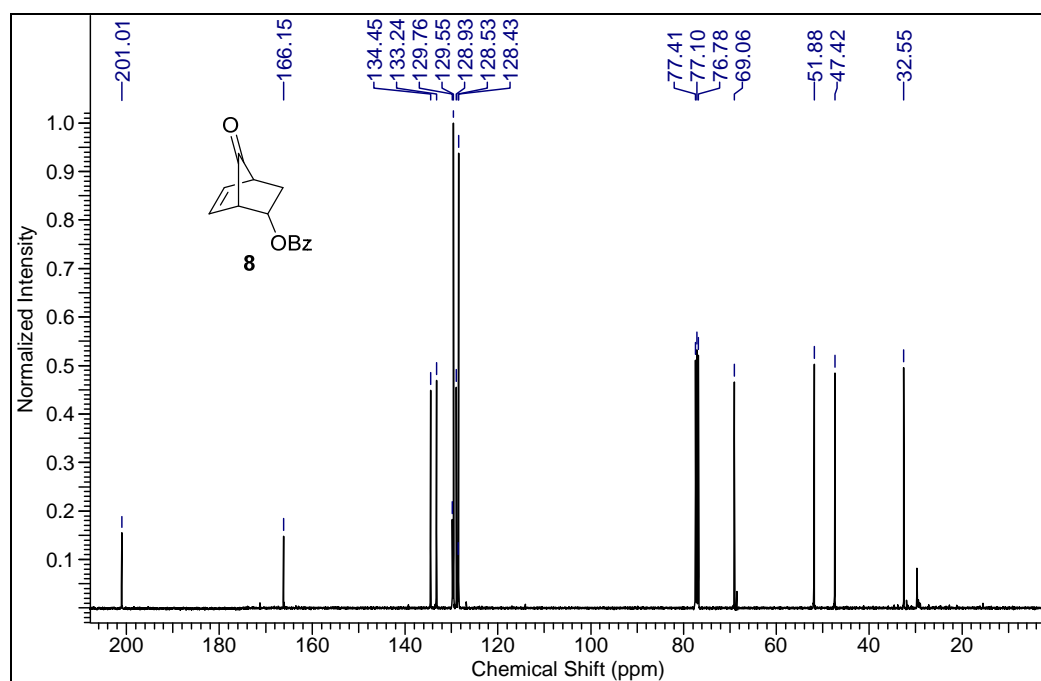
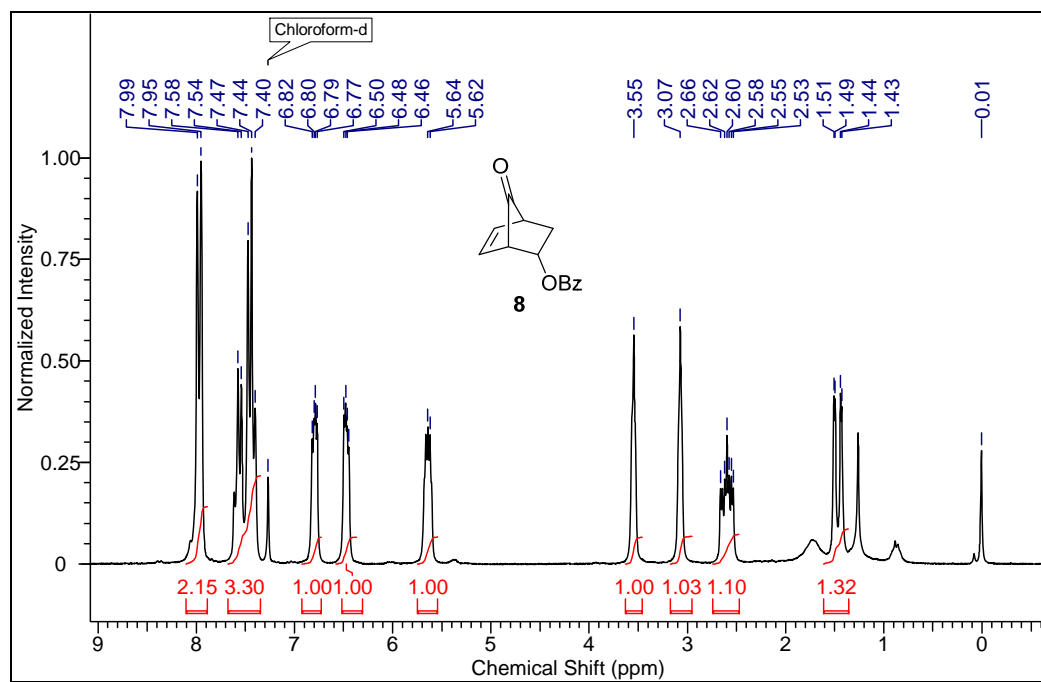
7

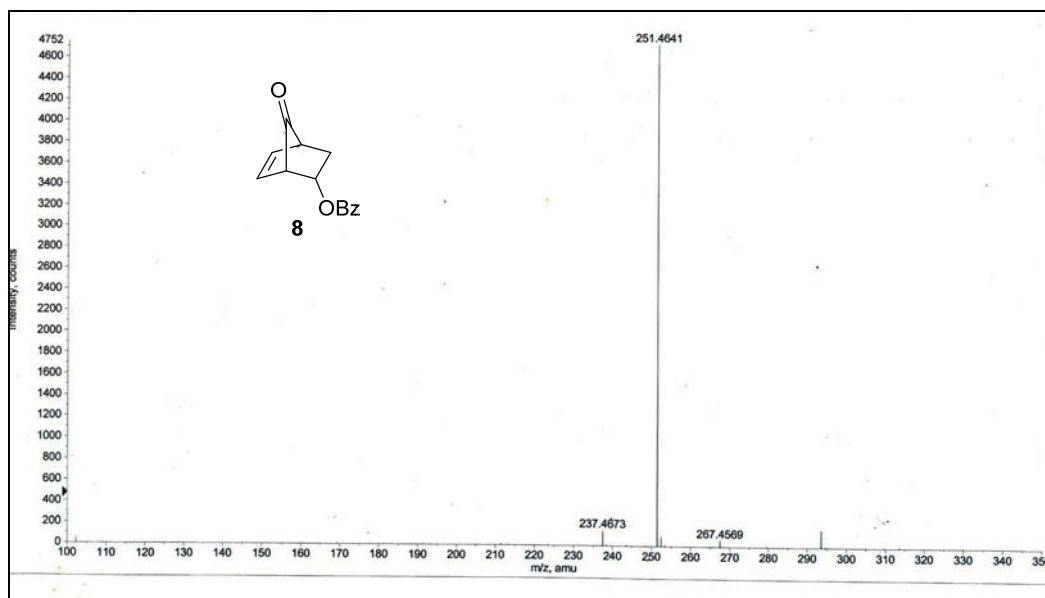
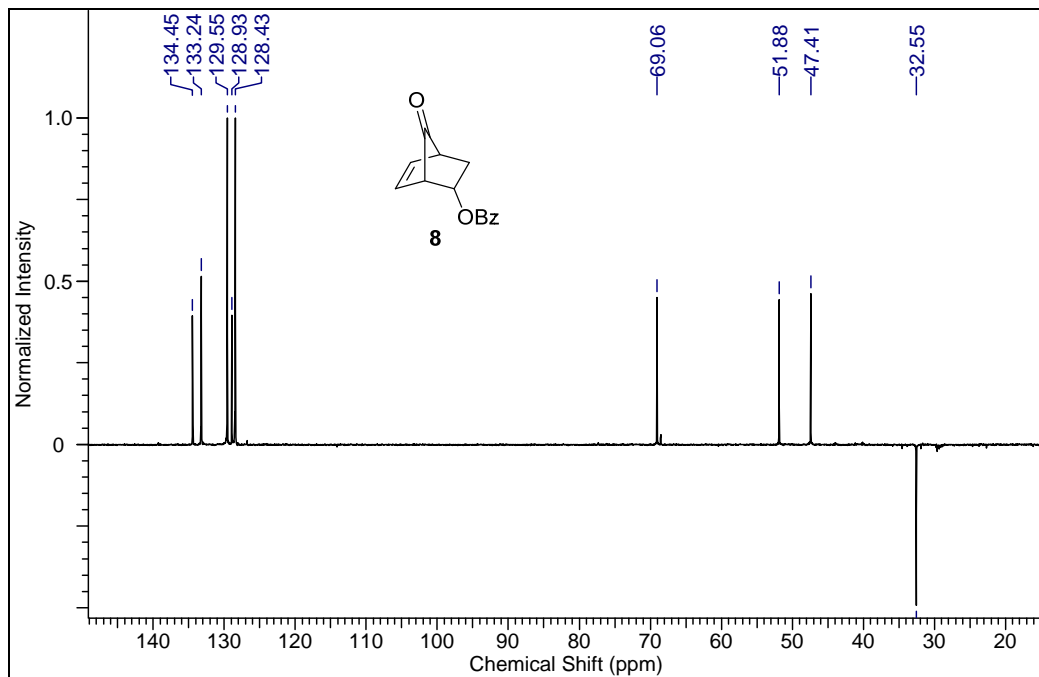
Target m/z: +297.1113 amu  
 Tolerance: +5.0000 ppm  
 Result type: Elemental  
 Max num of results: 100  
 Min DBE: -0.5000 Max DBE: +50.0000  
 Electron state: OddAndEven  
 Num of charges: 0  
 Add water: N/A  
 Add proton: N/A  
 File Name: 10MAR2011.wiff

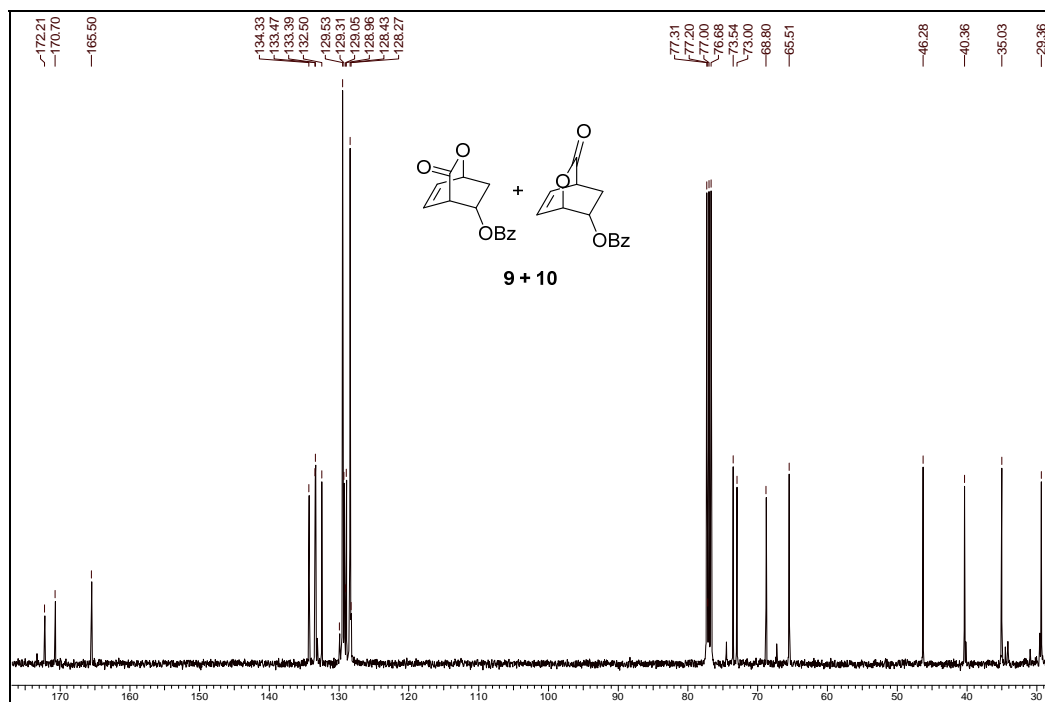
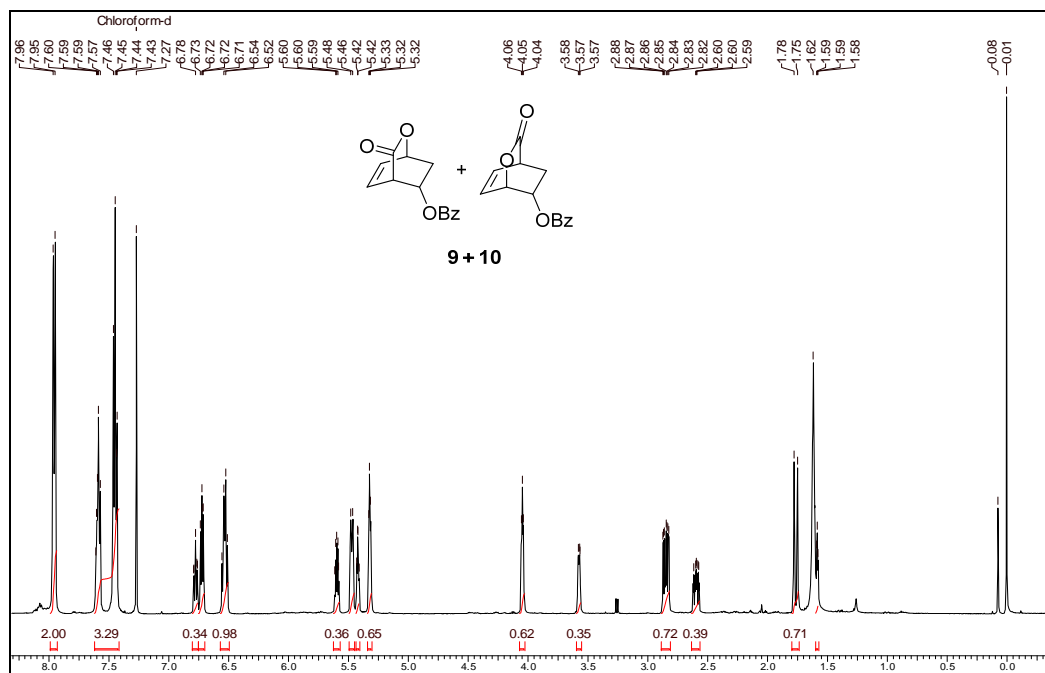
	Elements	Min Number	Max Number
1	C	0	17
2	H	0	20
3	Na	0	1
4	O	0	5

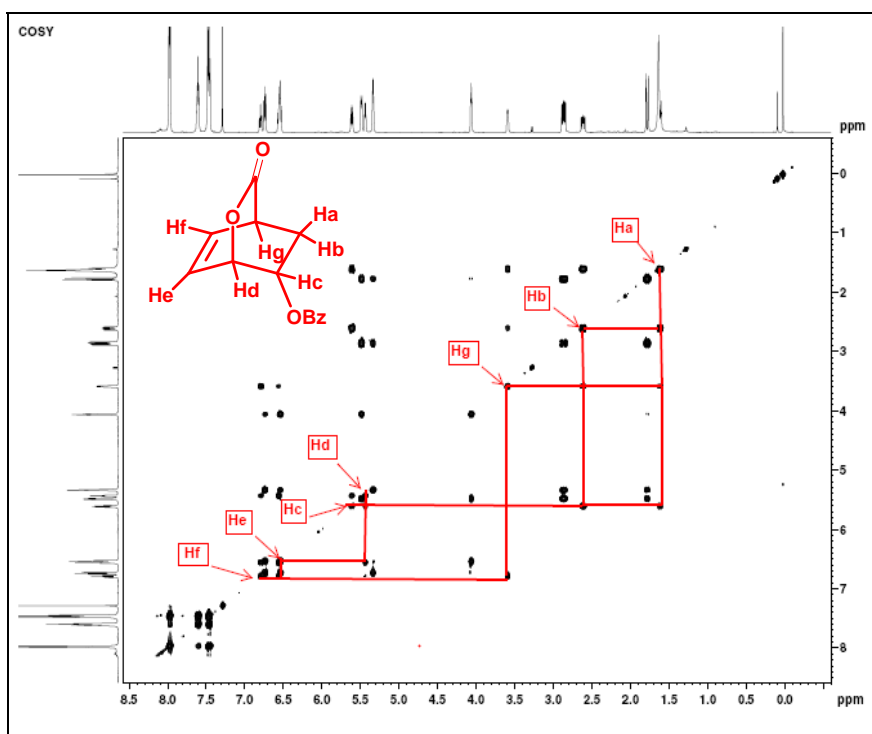
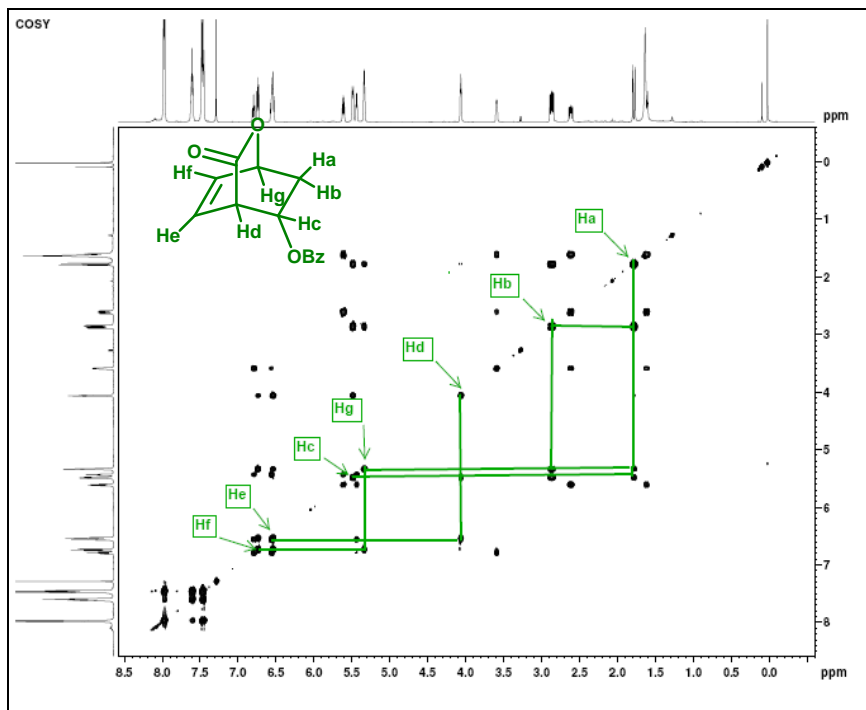
  

	Formula	Calculated m/z (amu)	mDa Error	PPM Error	DBE
1	C16 H18 O4 Na	297.1102	1.0210	3.4364	7.5

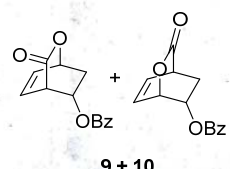








**Elemental composition calculator**



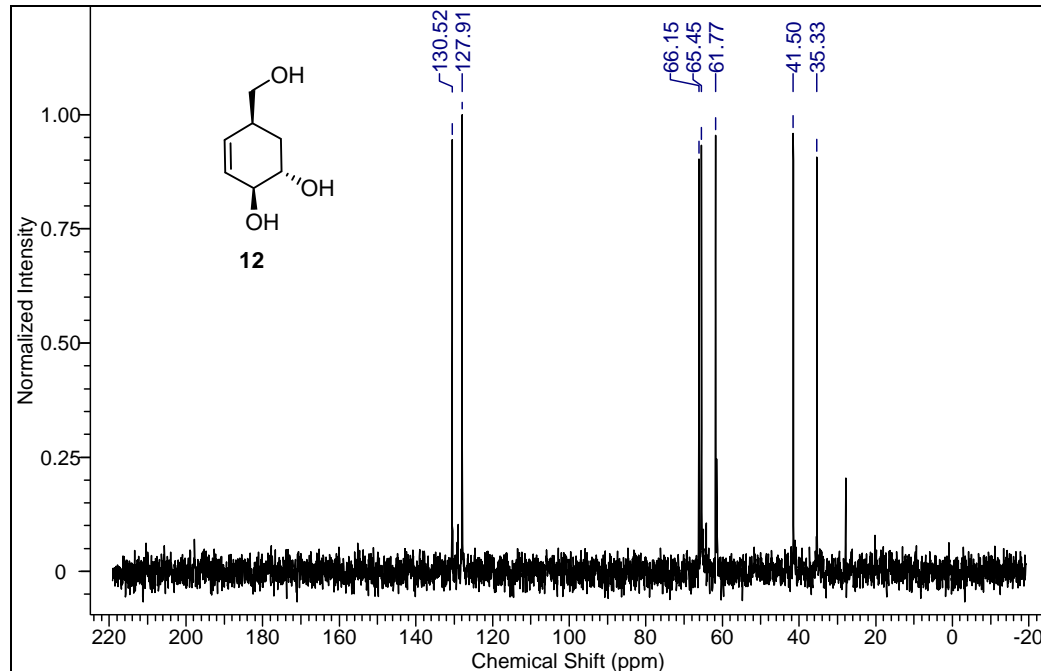
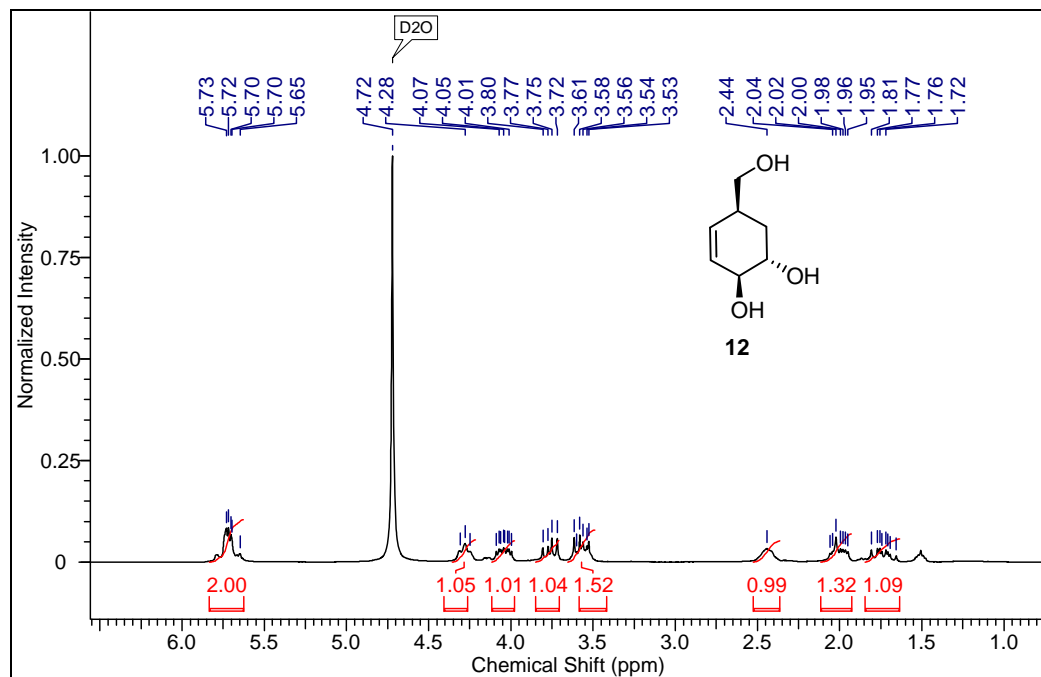
9 + 10

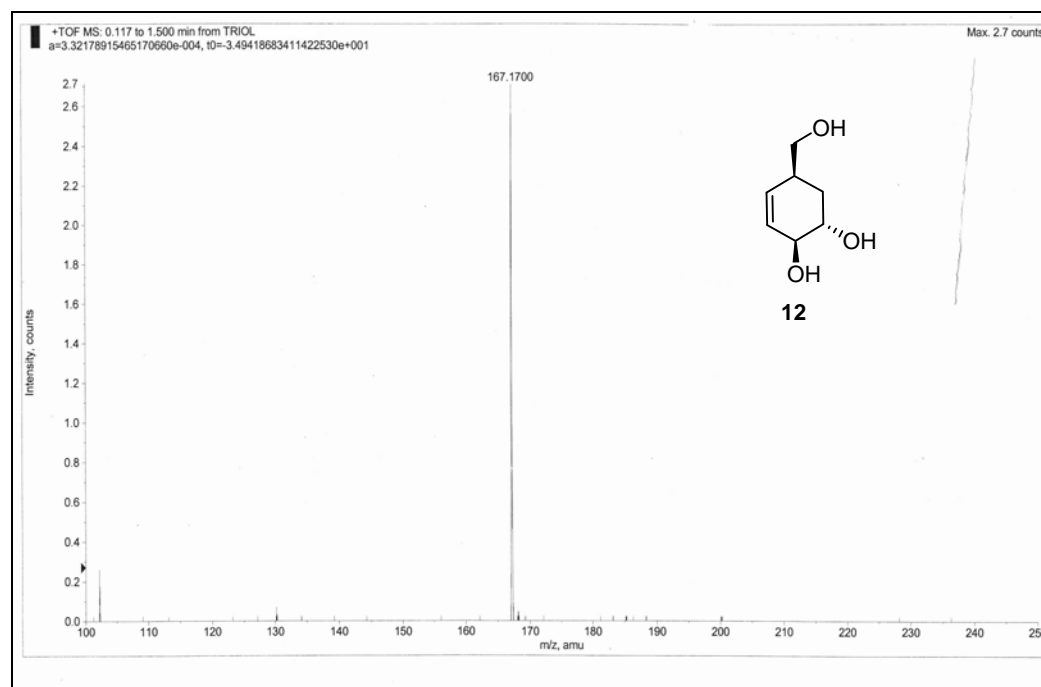
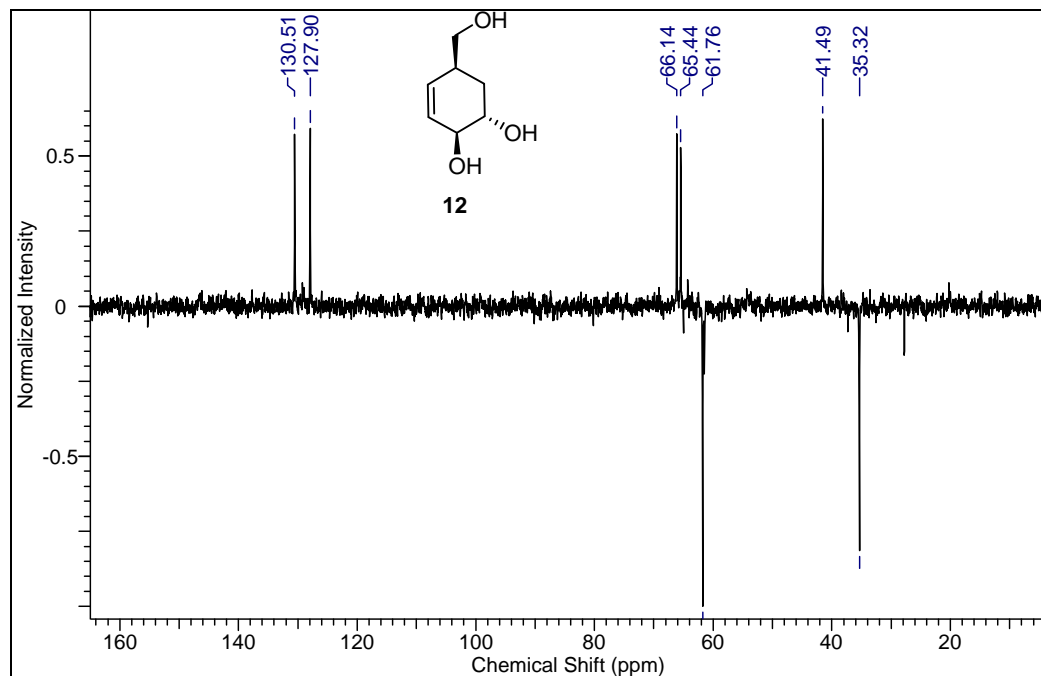
Target m/z: +245.0812 amu  
Tolerance: +5.0000 ppm  
Result type: Elemental  
Max num of results: 100  
Min DBE: -0.5000 Max DBE: +50.0000  
Electron state: OddAndEven  
Num of charges: 0  
Add water: N/A  
Add proton: N/A  
File Name: 10MAR2011.wiff

	Elements	Min Number	Max Number
1	C	0	17
2	H	0	20
3	Na	0	1
4	O	0	5

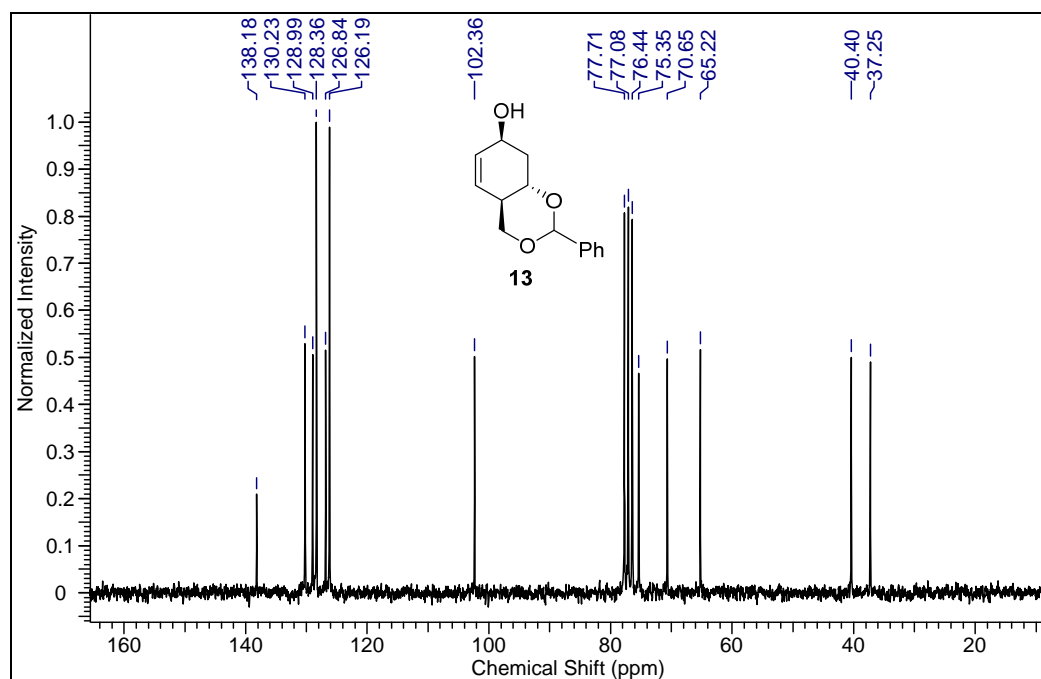
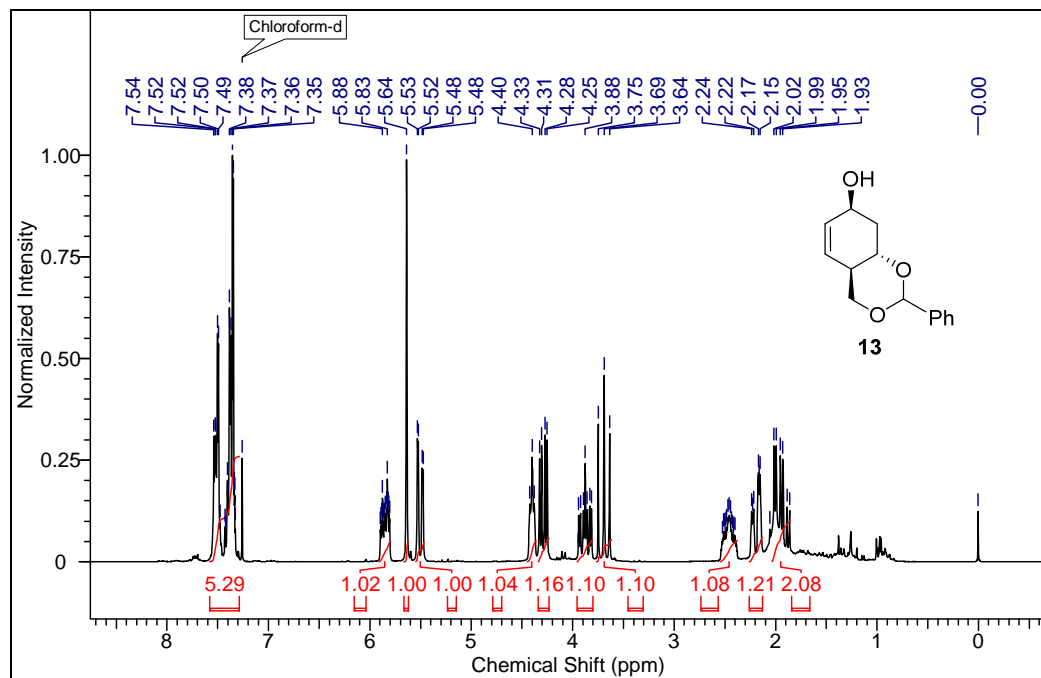
  

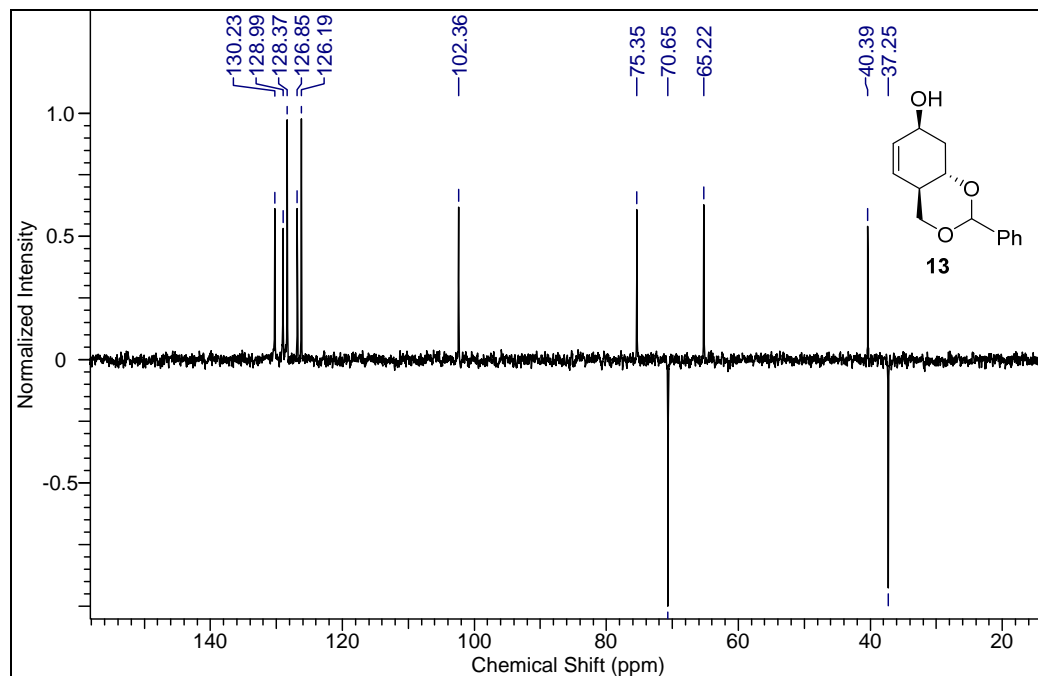
	Formula	Calculated m/z (amu)	mDa Error	PPM Error	DBE
1	C14 H13 O4	245.0813	-0.1840	-0.7510	8.5



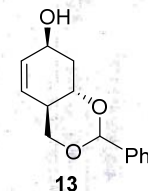








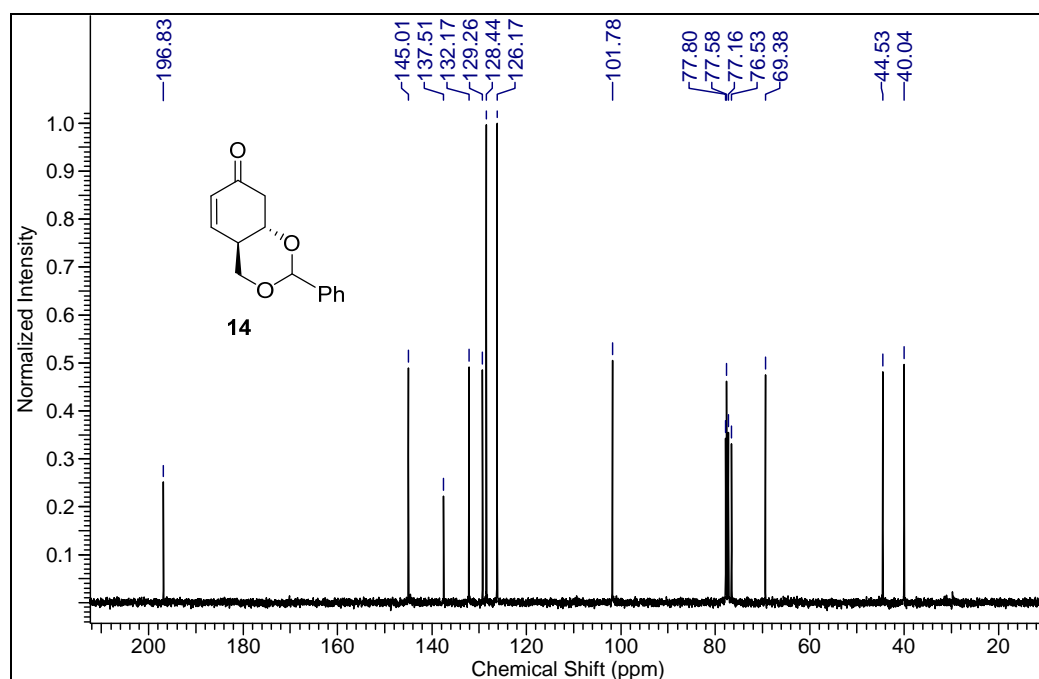
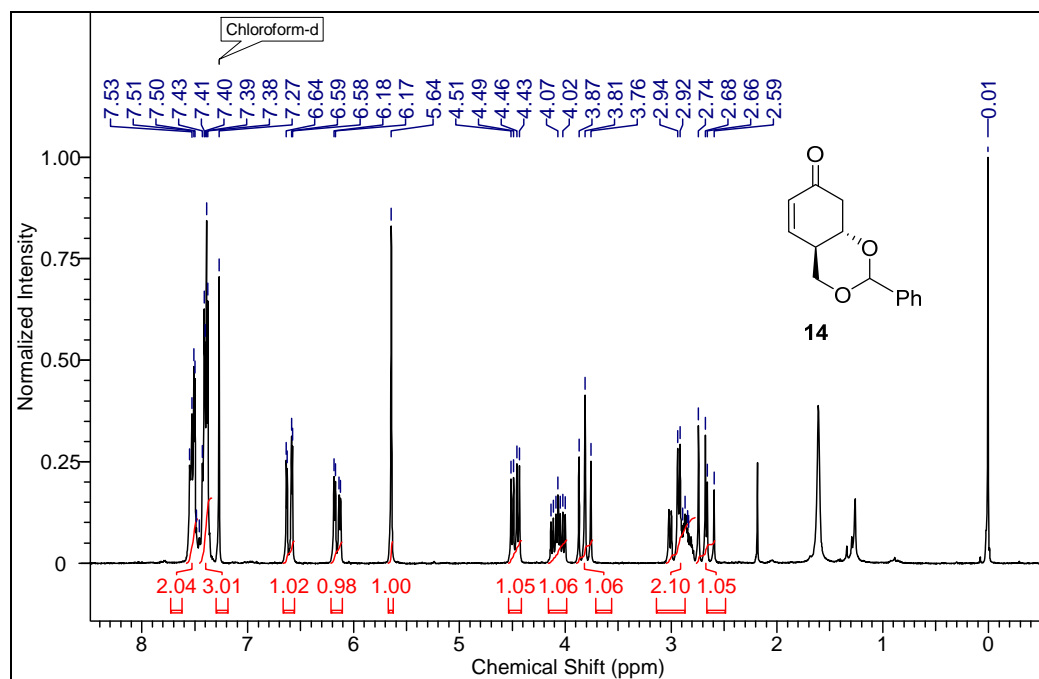
**Elemental composition calculator**

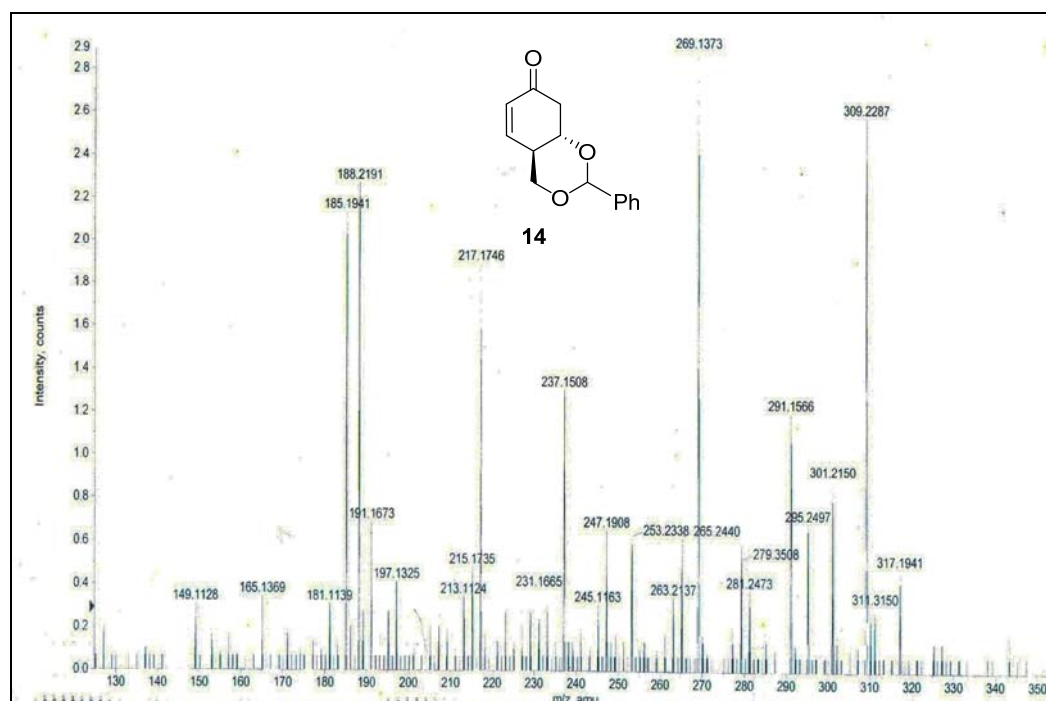
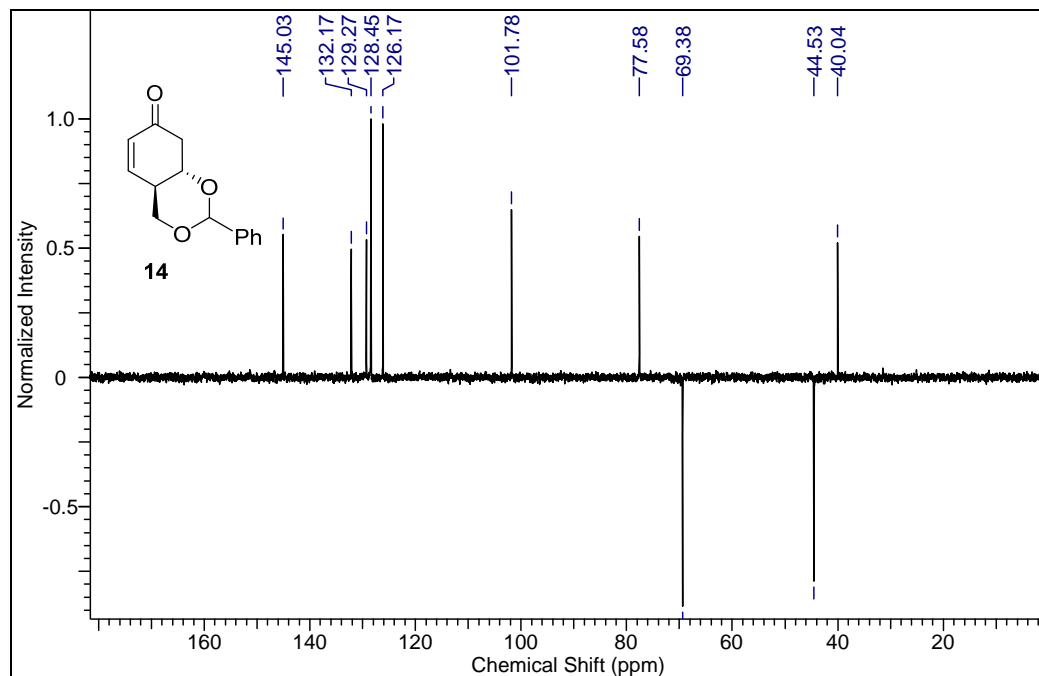


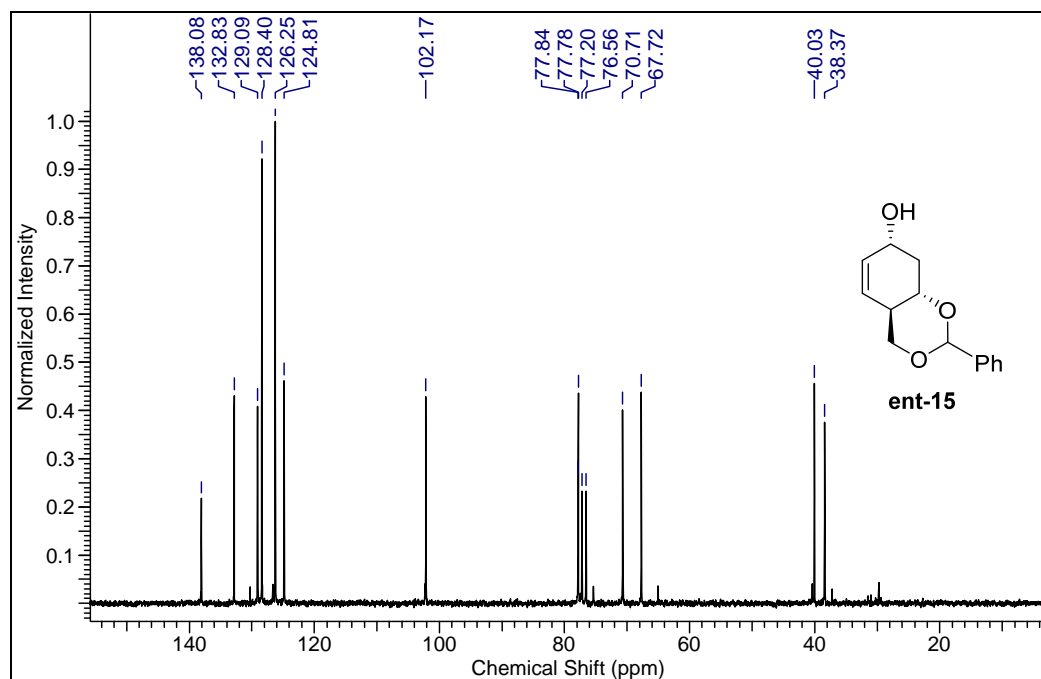
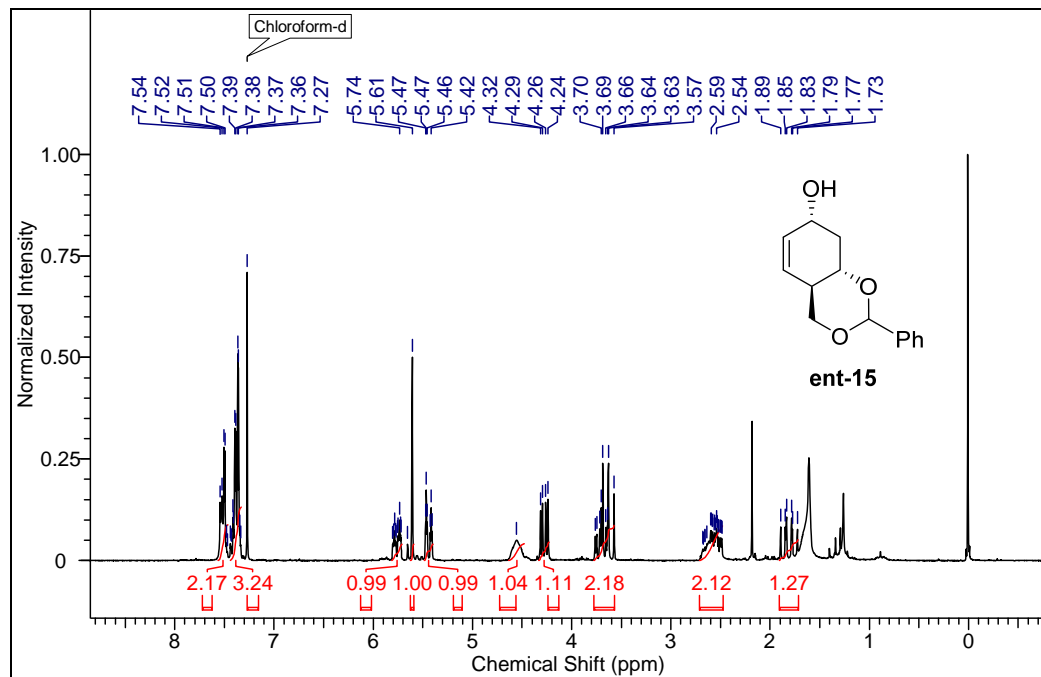
Target m/z: +233.1173 amu  
Tolerance: +5.0000 ppm  
Result type: Elemental  
Max num of results: 100  
Min DBE: -0.5000 Max DBE: +50.0000  
Electron state: OddAndEven  
Num of charges: 0  
Add water: N/A  
Add proton: N/A  
File Name: 10MAR2011.wiff

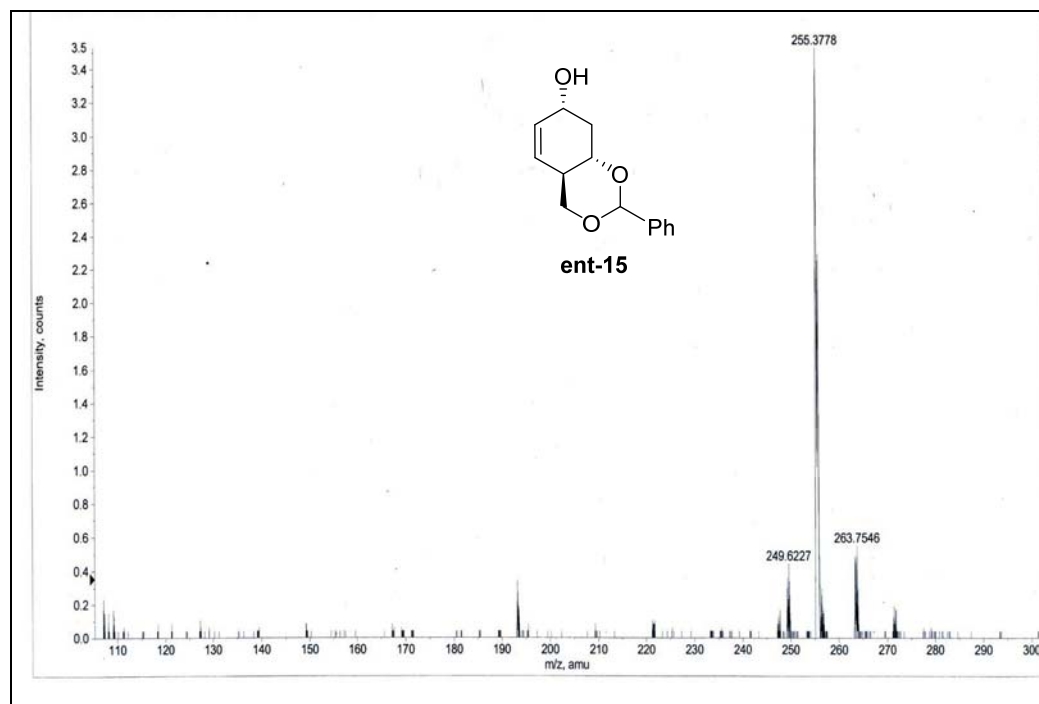
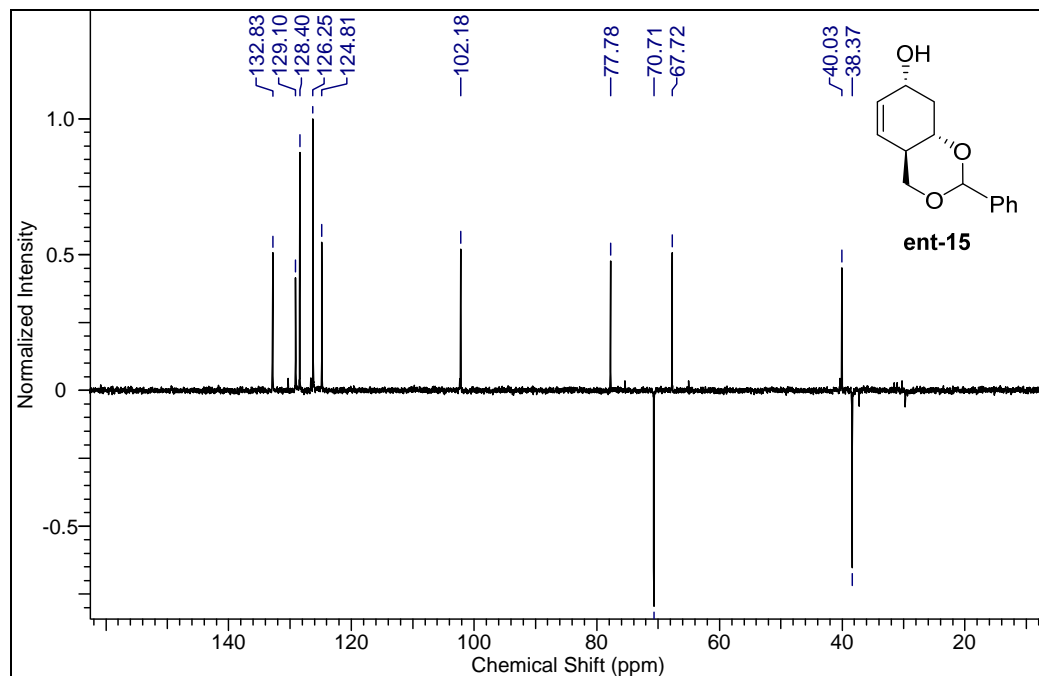
	Elements	Min Number	Max Number:
1	C	0	17
2	H	0	20
3	Na	0	1
4	O	0	5

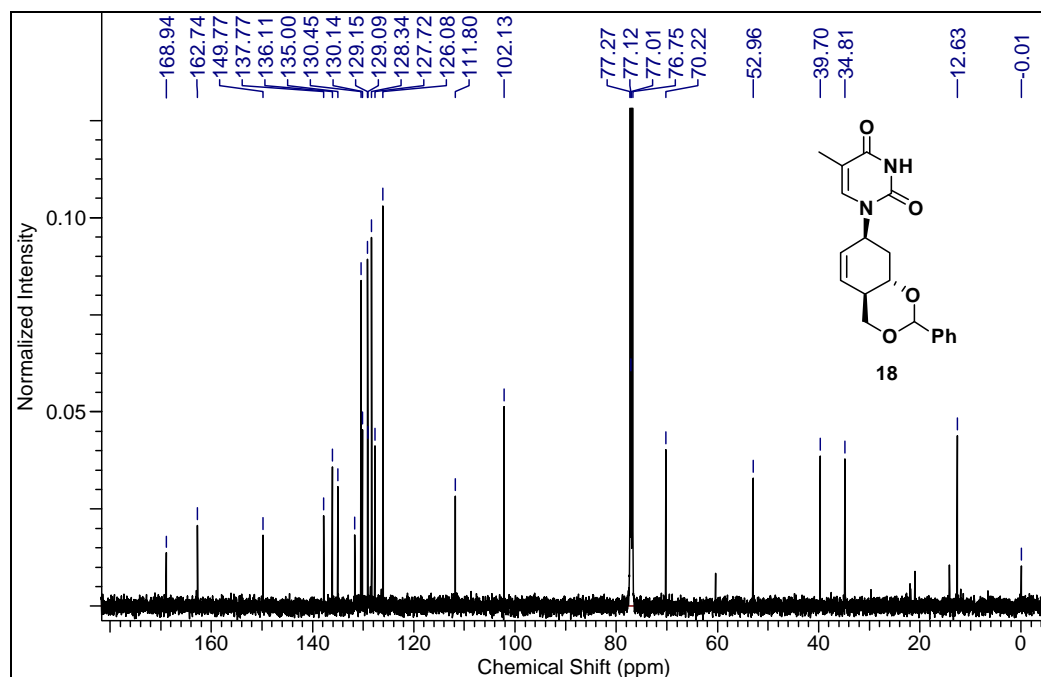
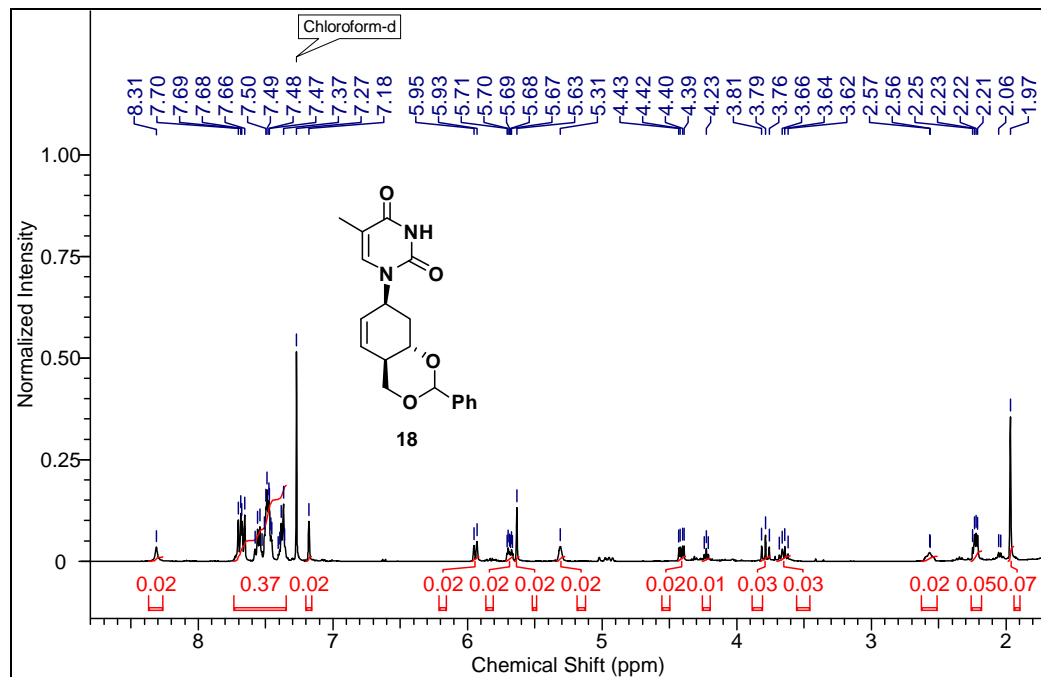
	Formula	Calculated m/z (amu)	mDa Error	PPM Error	DBE
1	C14 H17 O3	233.1177	-0.4696	-2.0144	6.5

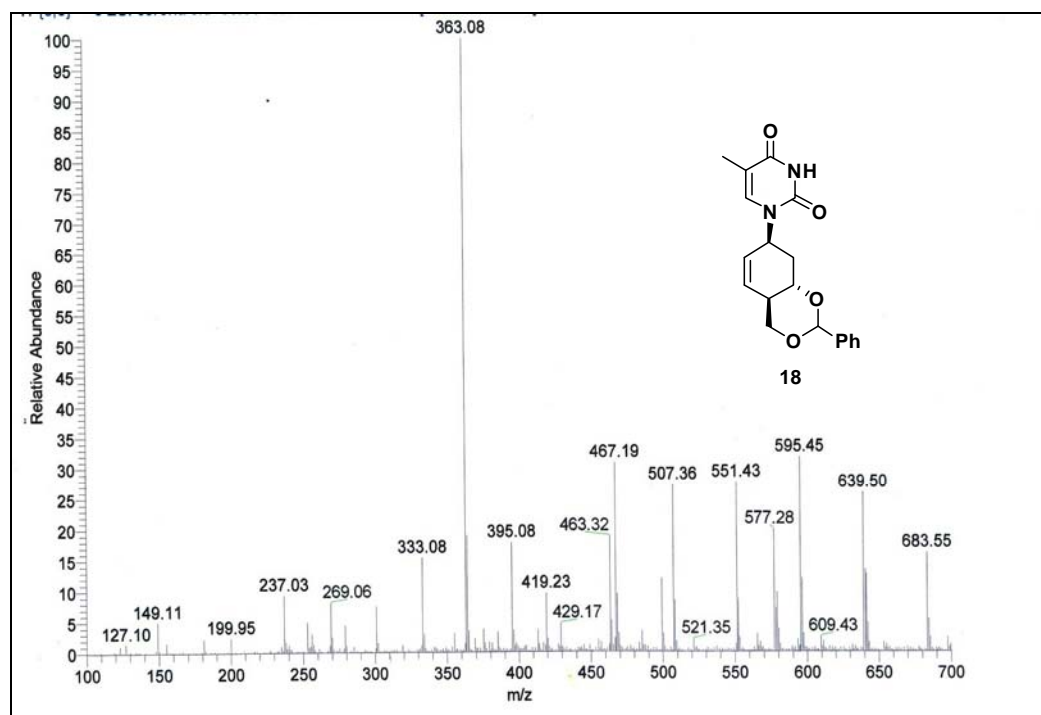
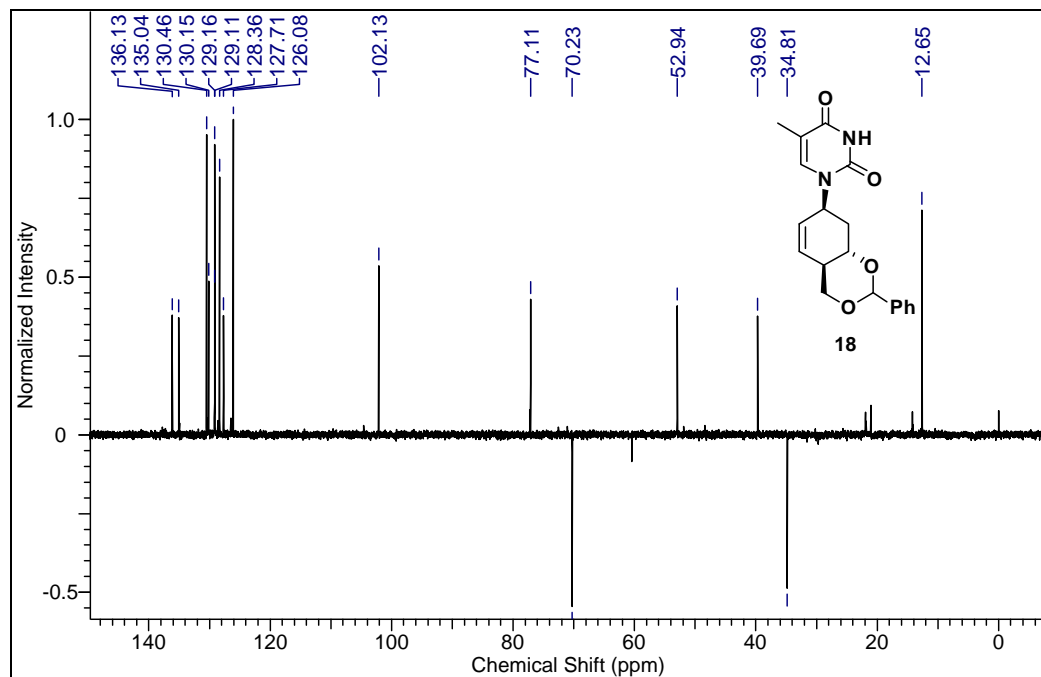




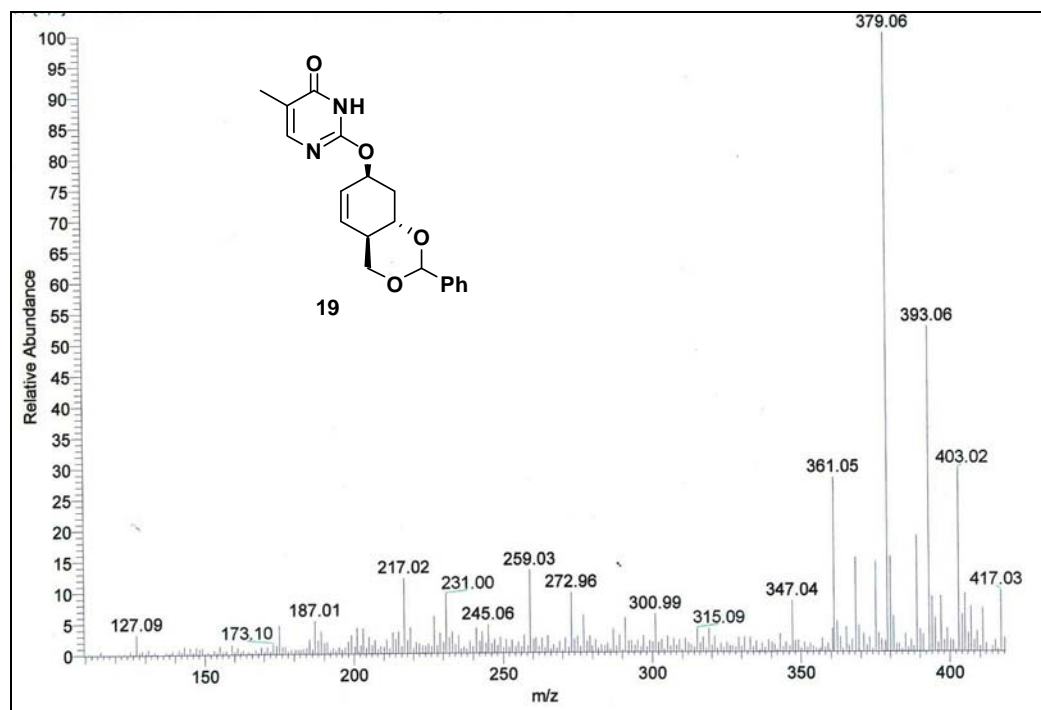
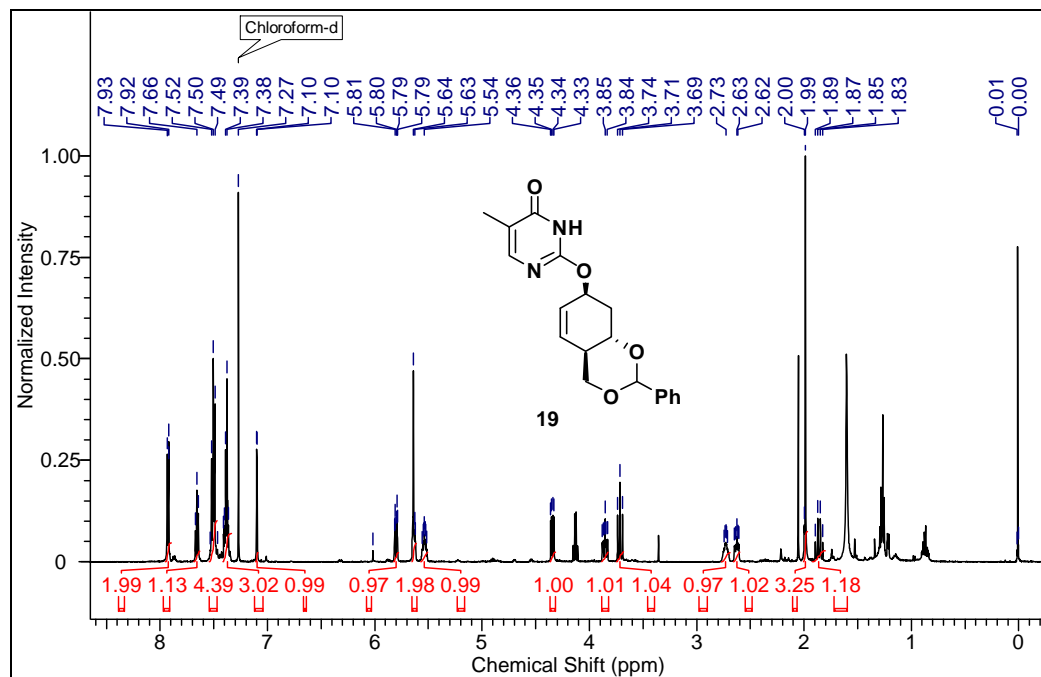


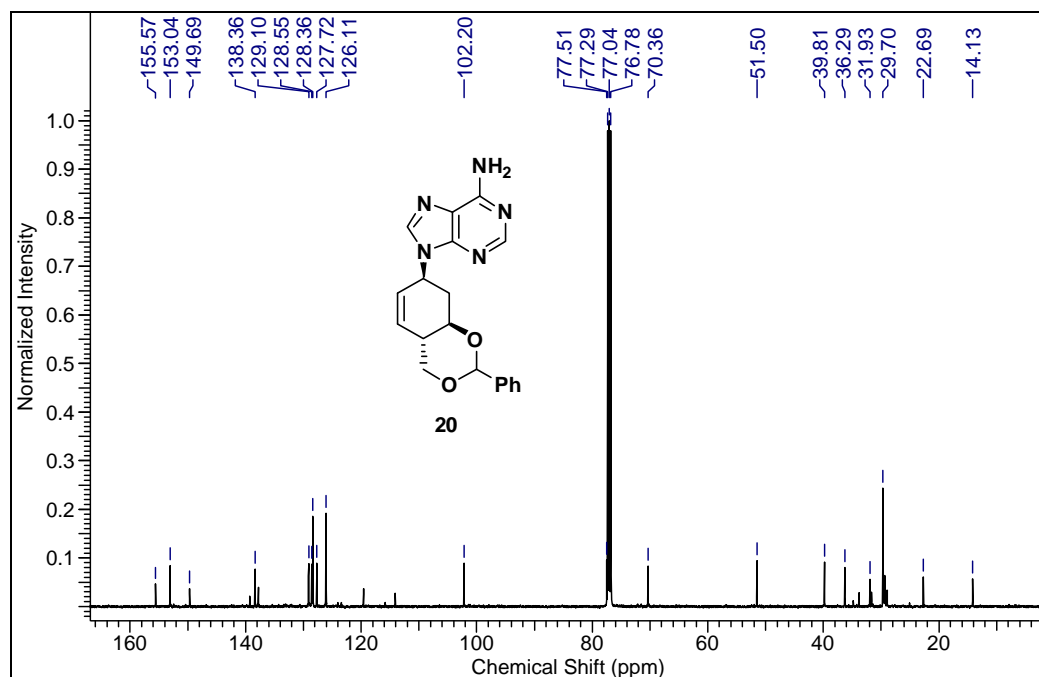
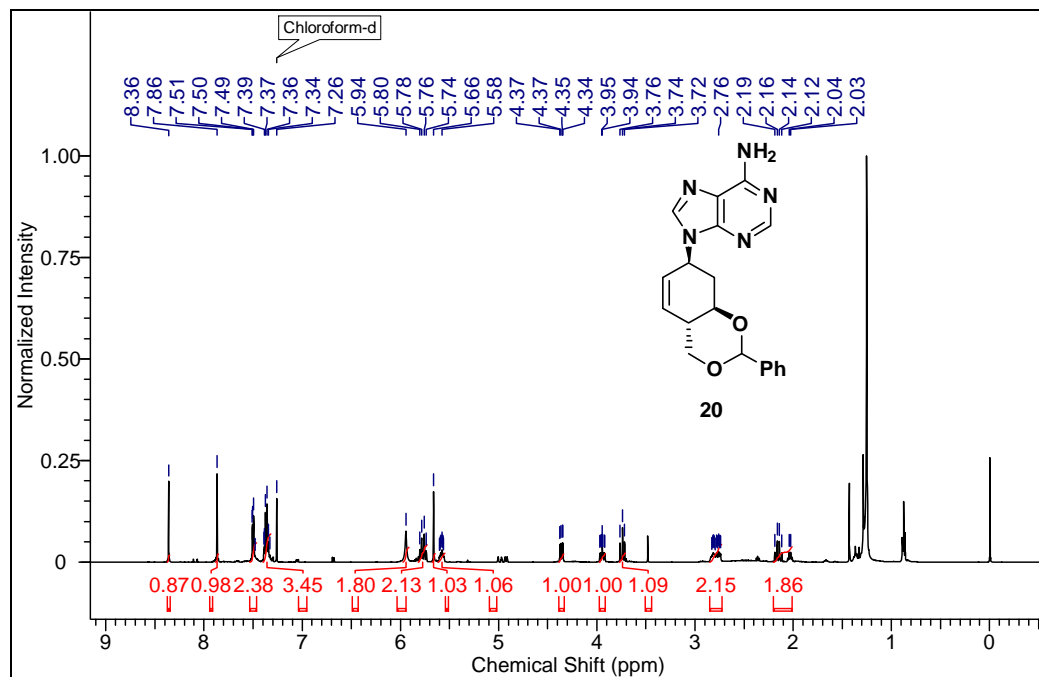


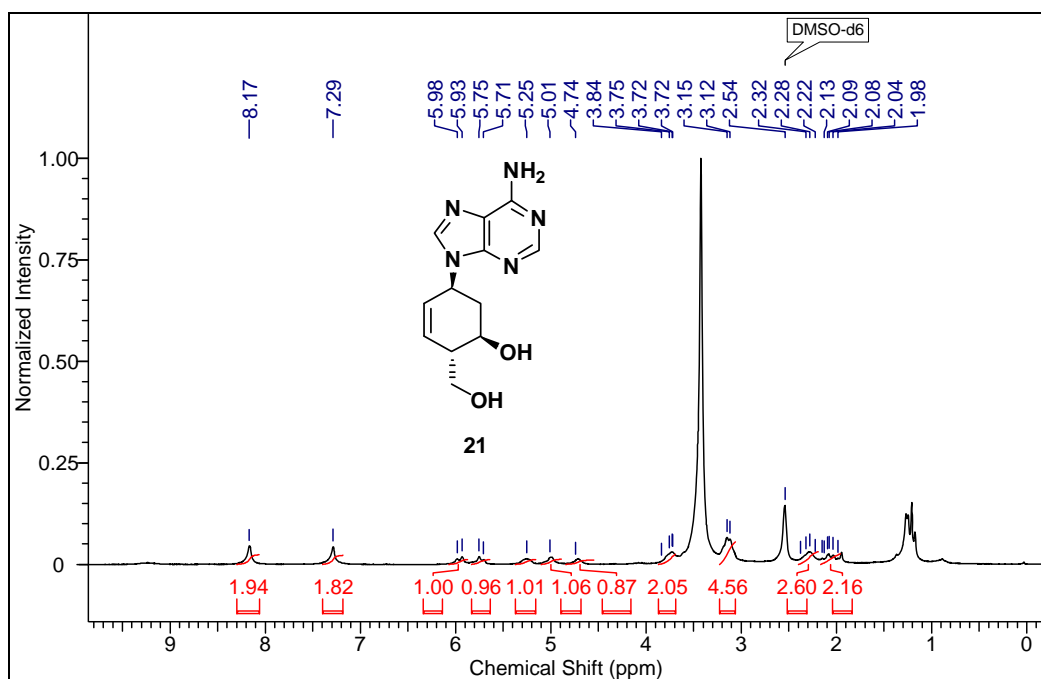
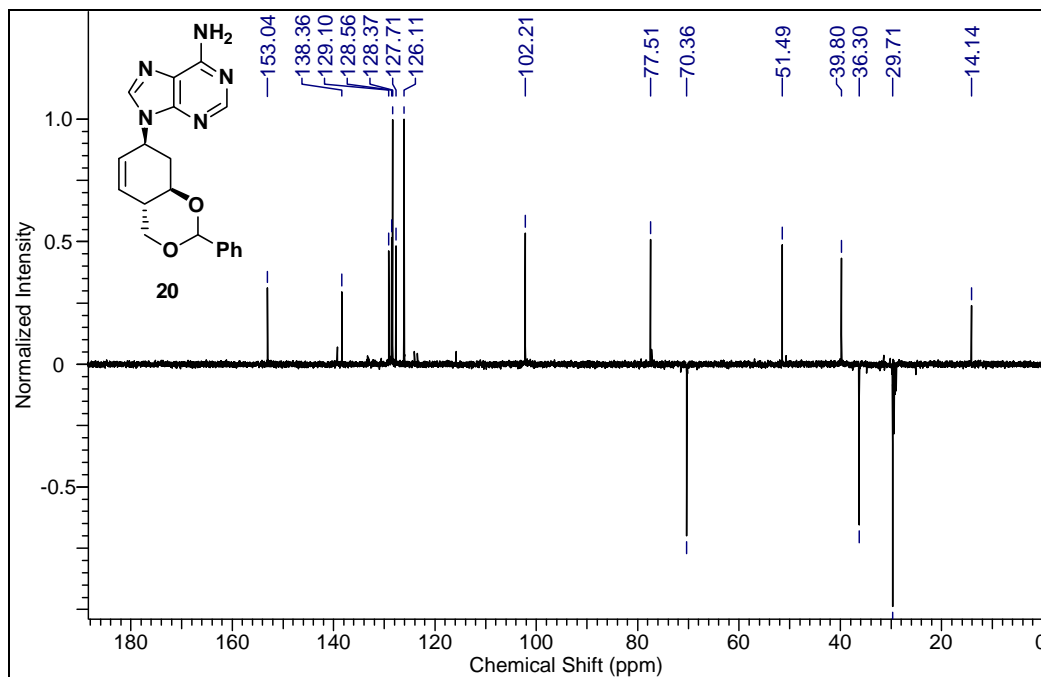


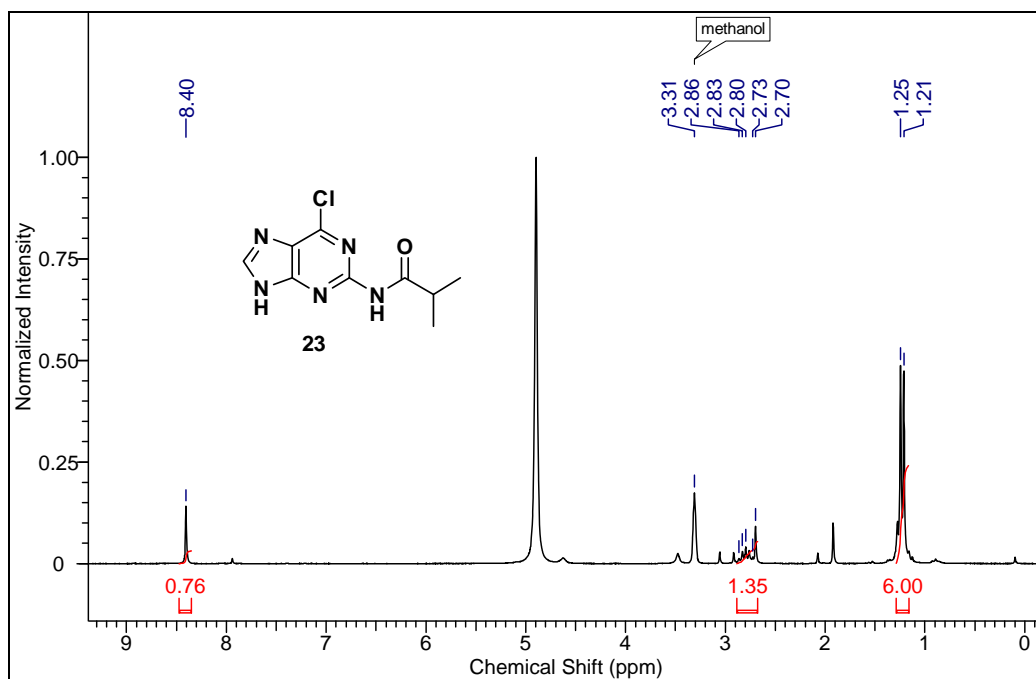
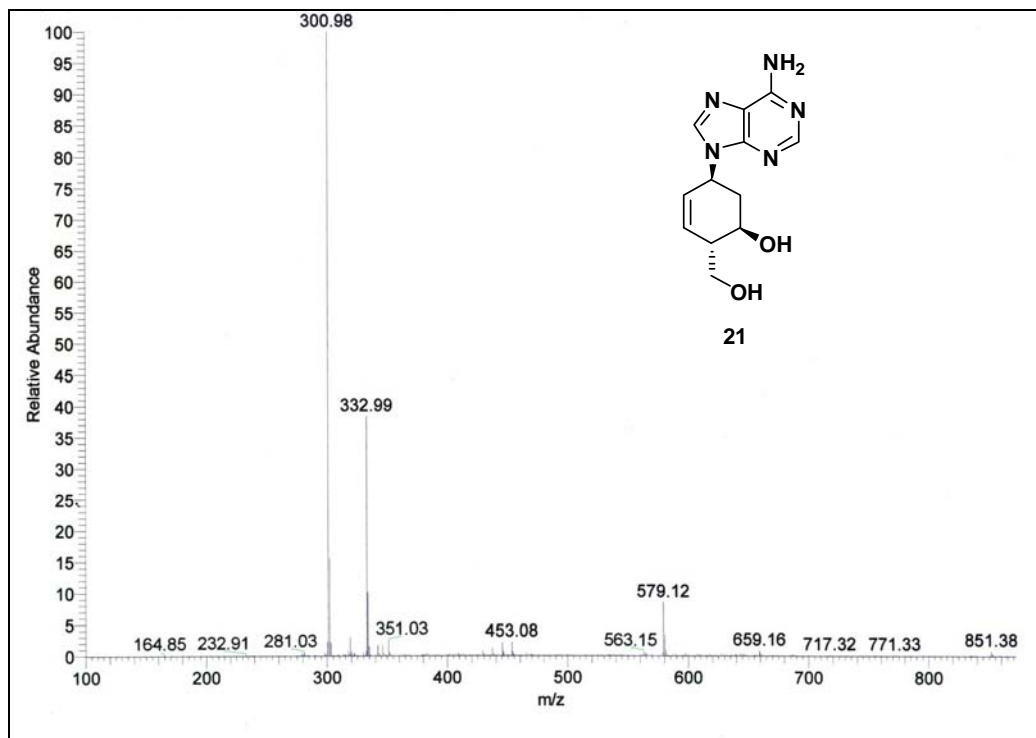


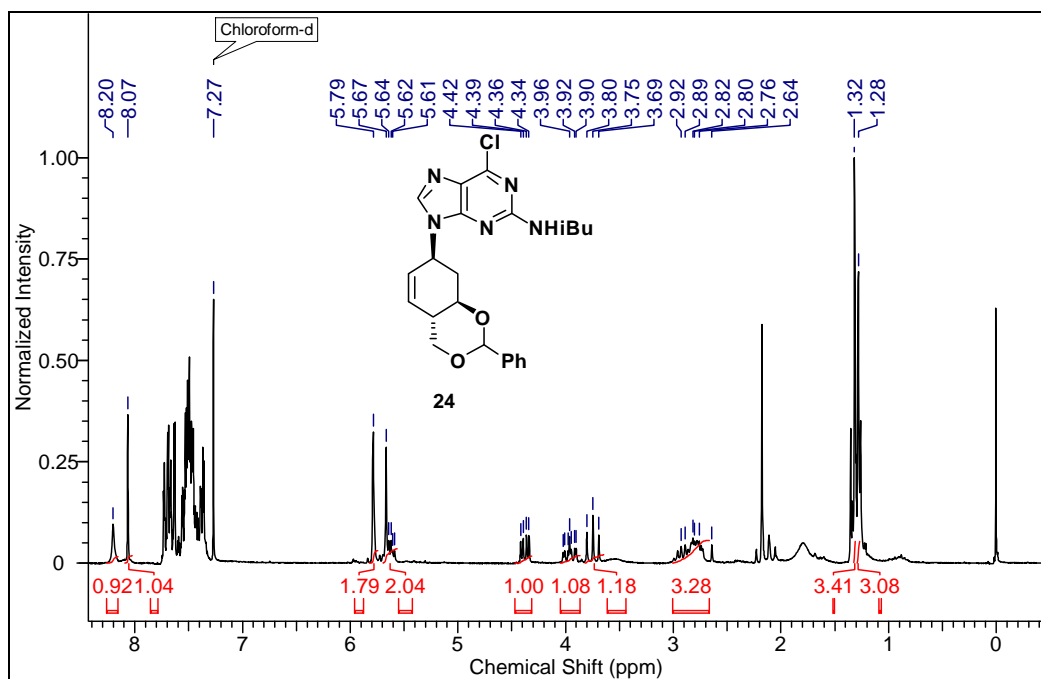
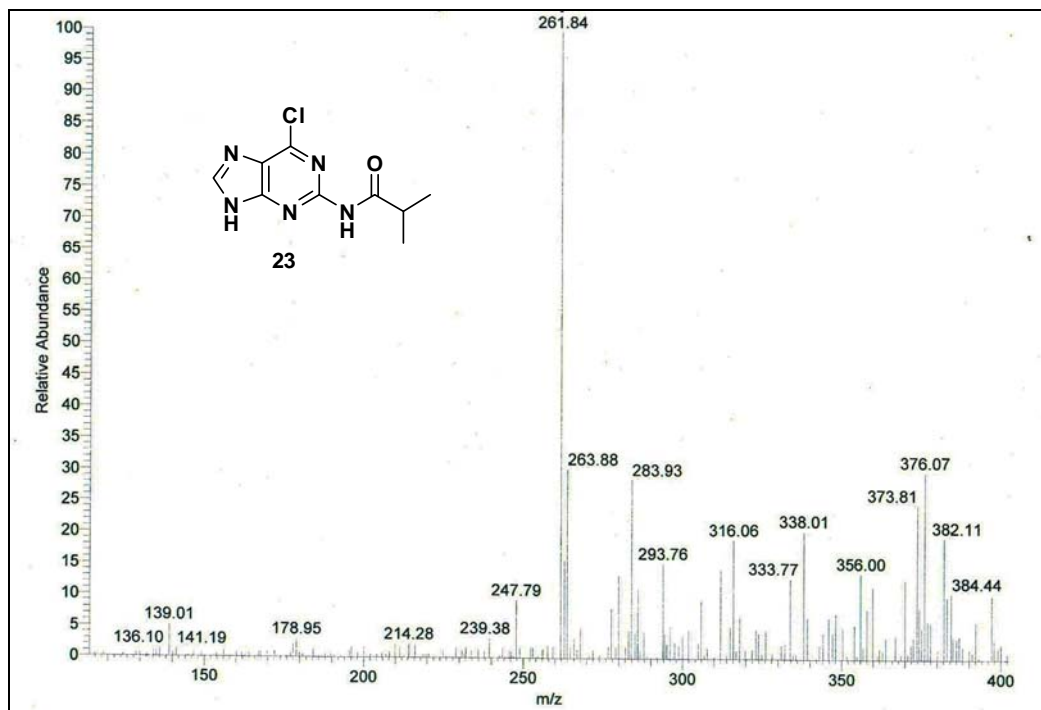


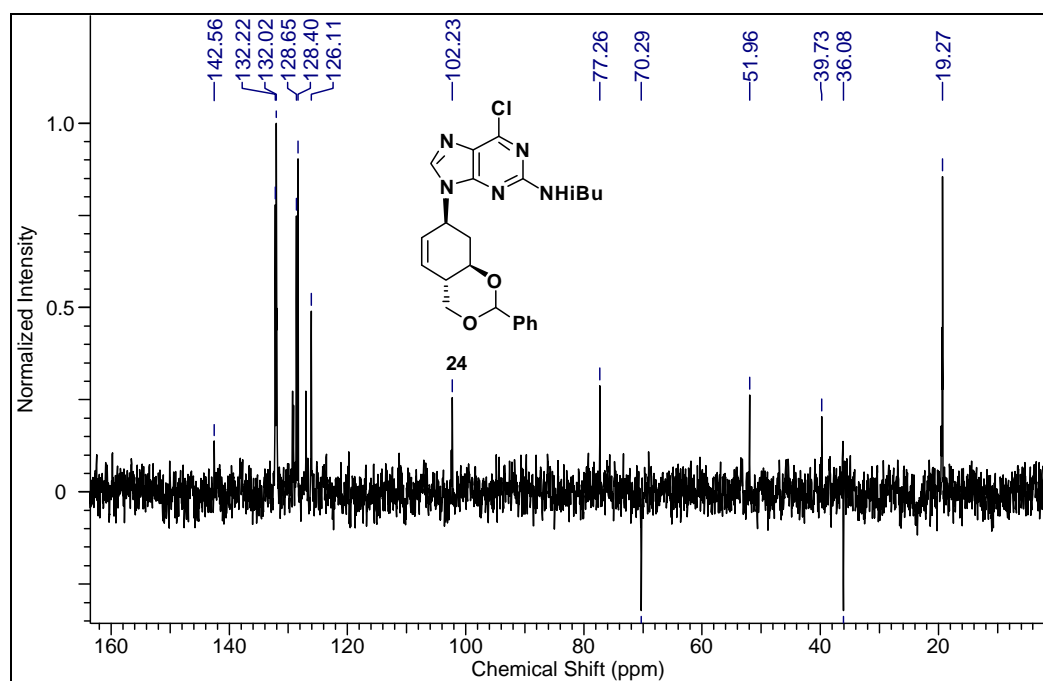
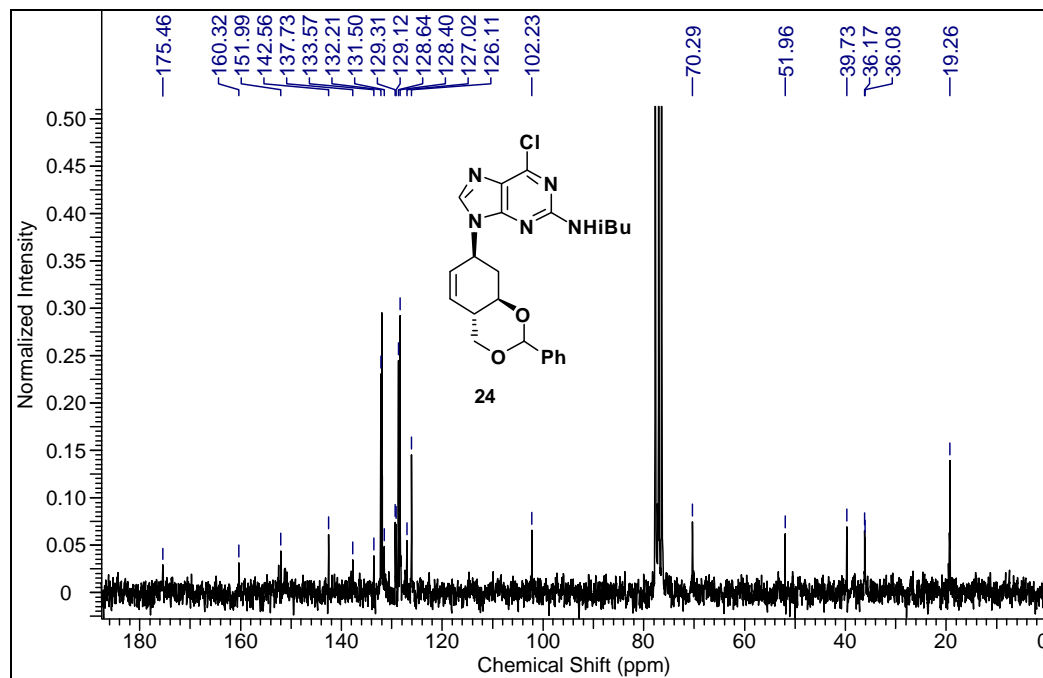


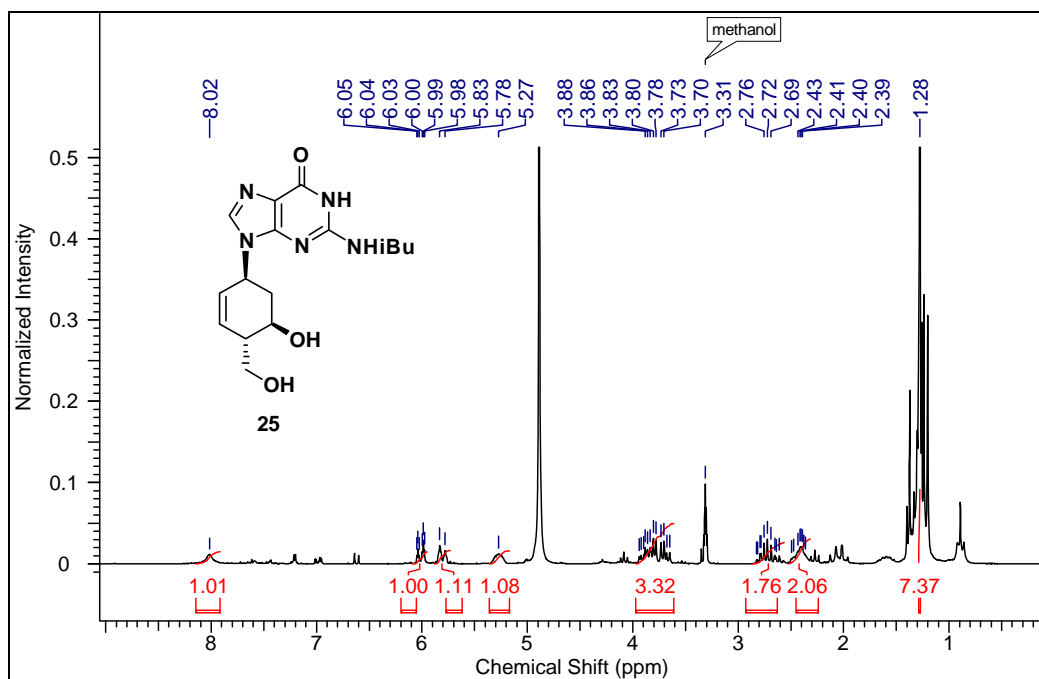
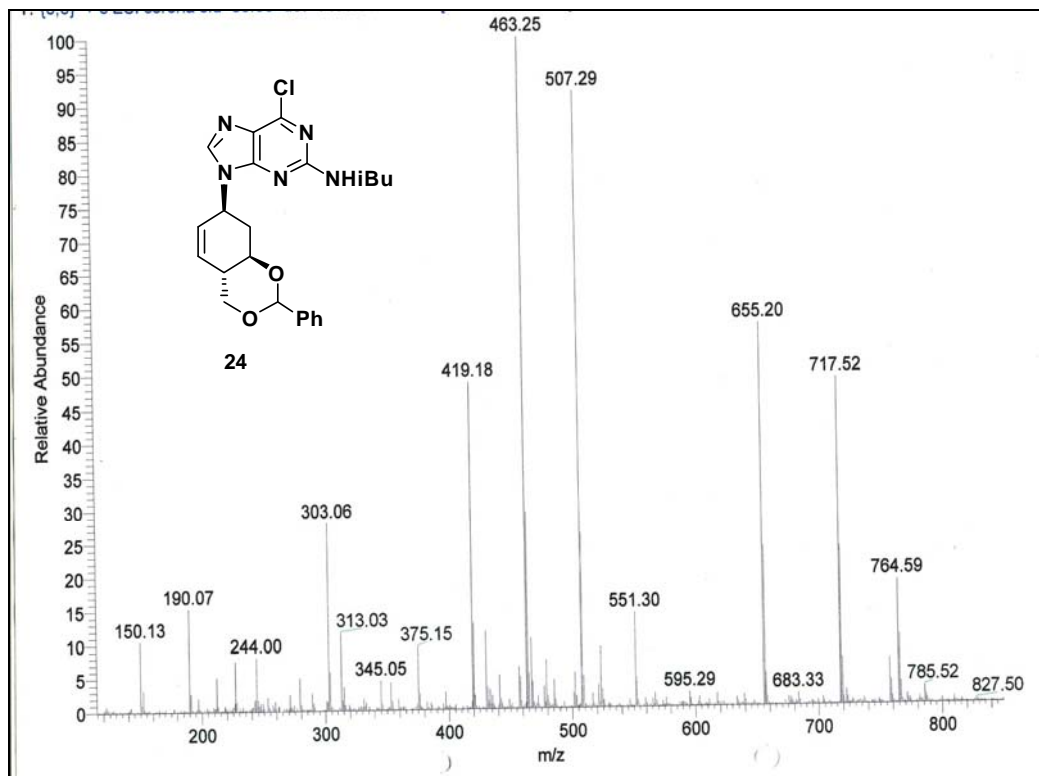


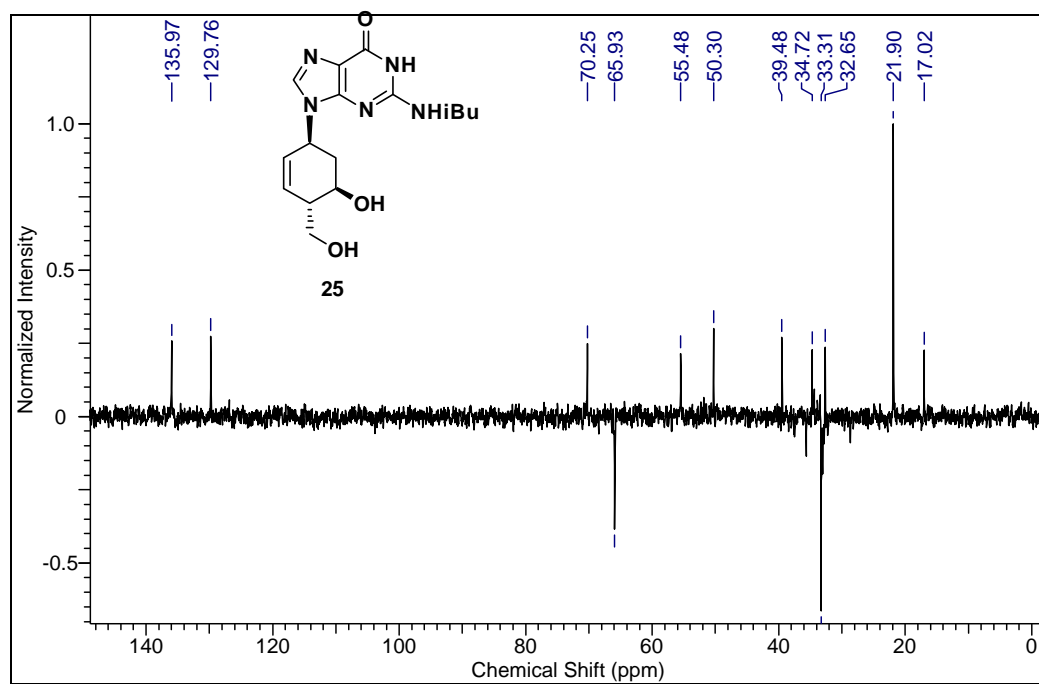
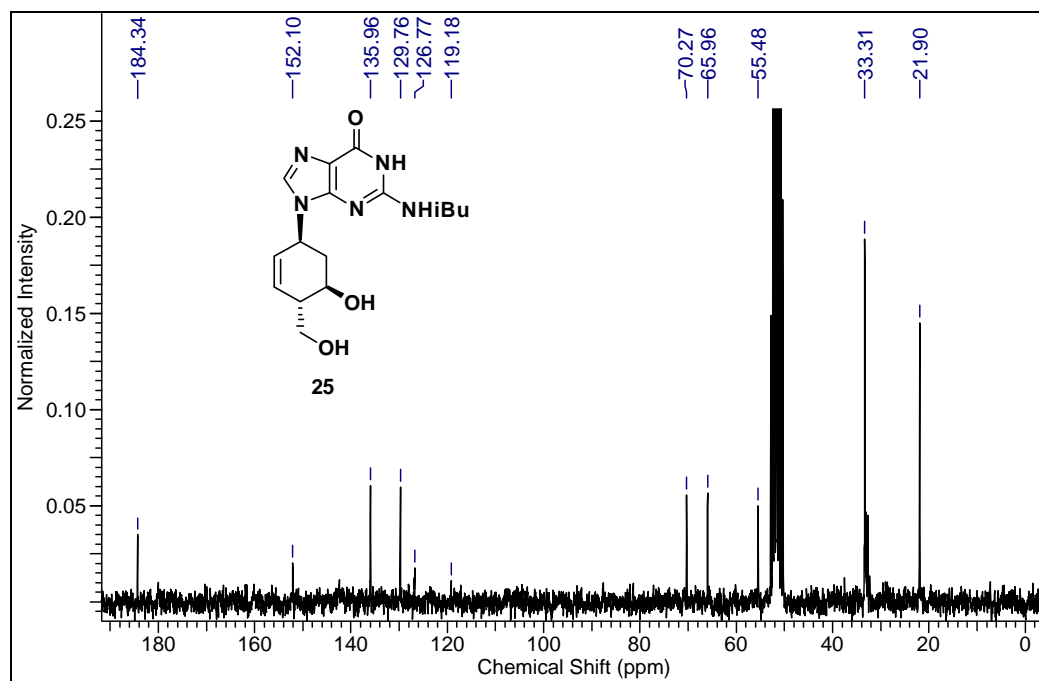




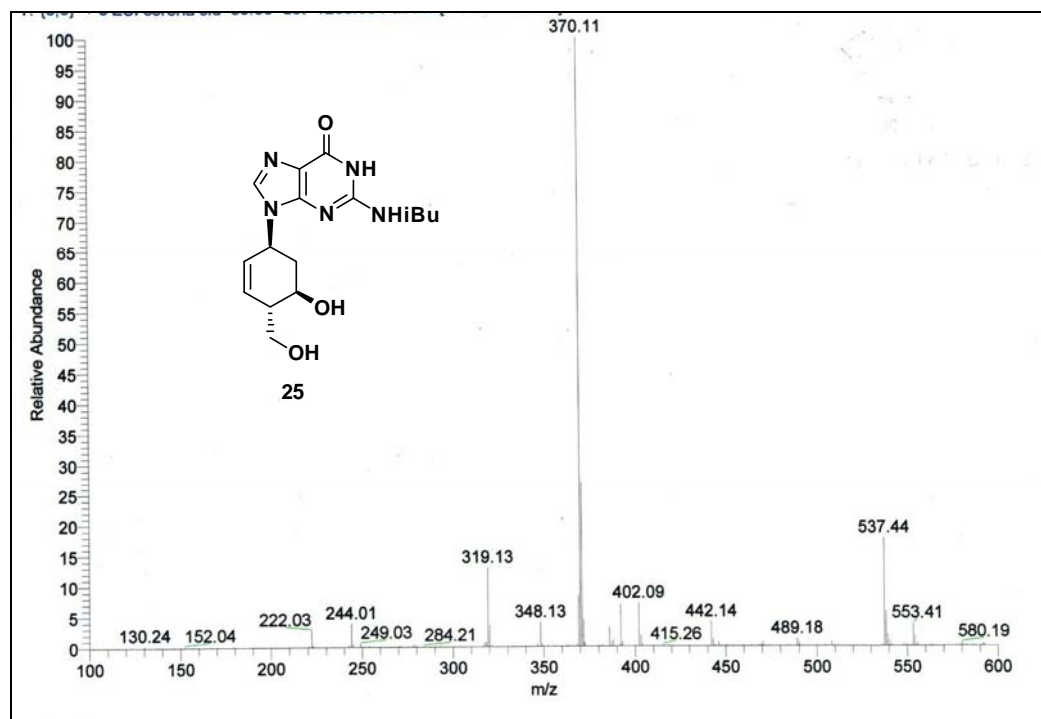












### 2.12 References

1. (a) Kole, R.; Krainer, A.R.; Altman, S., *Nature Rev-Drug Discov.* **2012**, *11*, 125; (b) J. Kurreck., *Eur. J. Biochem.* **2003**, *270*, 1628.
2. (a) Shukla, S.; Sumaria, C. S.; Pradeepkumar, P.I., *Chem. Med. Chem*, **2010**, *5*, 328; (b) Kurreck J., *Angew. Chem., Int. Ed. Engl.* **2009**, *48*, 1378.
3. (a) Sun, W.; Li, Y. S. H.; Huang, H. D.; John Y-J. Shyy.; Chien, S., *Annu. Rev. Biomed. Eng.* **2010**, *12*, 1; (b) Turner, J. J.; Fabani, M.; Arzumano, A. A.; Ivanova, G.; Gait, M. J., *Biochim. Biophys. Acta, Biomembr*, **2006**, *1758*, 290.
4. (a) Bauman J.; Jearawiriyapaisam, N.; Kole R., *Oligonucleotides*, **2009**, *19*, 1; (b) Sazani, P.; Kole, R., *J. Clin. Invest*, **2003**, *112*, 481.
5. Herdewijn, P., *Chemistry Biodiversity*, **2010**, *7*, 1-59.
6. Arthur Van Aerschot.; Ilse Verheggen.; Chris Hendrix.; Piet Herdewijn., *Angew Chem. Inf. Ed. Engl.* **1995**, *34*, 1338-1339.
7. Nauwelaerts, K.; Lescrinier, E.; Sclep, G.; Herdewijn, P., *Nucleic Acids Res.* **2005**, *33*, 2452-2463.
8. Wang, J.; Froeyen, M.; Hendrix, C.; Andrei, G.; Snoeck, R.; Clercq, E. D.; Herdewijn, P., *J. Med. Chem.* **2000**, *43*, 736-745.
9. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A. and Herdewijn, P., *J. Am. Chem. Soc.*, **2000**, *122*, 8595–8602.
10. (a) Wang, J.; Verbeure, B.; Luyten, I.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P., *Nucleos. Nucleot. & Nucleic Acids*, **2001**, *20*, 785-788. (b) Birgit verbeure.; Eveline Lescrinier Wang, J.; Herdewijn, P., *Nucleic Acids Res.* **2001**, *29*, 4941-4947.

## Chapter 2

---

11. Nauwelaerts, K.; Fisher, M.; Froeyen, M.; Lescrinier, E.; Van Aerschot, A.; Xu, D.; DeLong, R.; Kang, H.; Juliano, R. L. and Herdewijn, P., *J. Am. Chem. Soc.*, **2007**, *129* 9340–9348.
12. Wang, J.; Herdewijn, P., *J. Org. Chem.* **1999**, *64*, 7820-7827.
13. Wang, J.; Morral, J.; Hendrix, C.; Herdewijn, P., *J. Org. Chem.* **2001**, *66*, 8478-8482.
14. Wislicenus, W. *Chem. Ber.* **1887**, *20*, 2930-2934.
15. Danishefsky, S; Kitahara, T., *J. Am. Chem. Soc.* **1974**, *96*, 7807-7808.
16. Birgit Verbeure.; Gilles Gaubert.; Piet Herdewijn.; Jesper Wengel., *J. Am. Chem. Soc.* **2002**, *124*, 2164-2176.
17. (a) Mehta, G.; Talukdar, P.; Mohal, N., *Tetrahedron Lett.* **2001**, *42*, 7663-7666.  
(b) Talukdar, P. M.Sc. Thesis, IISc Bangalore, May **2001**.
18. Cotsaris, E.; Paddon-Row, M. N., *J. Chem. Soc., Chem. Commun.* **1984**, 95-96.
19. Palani, N.; Rajamannar, T.; Balasubramanian, K. K., *Synlett.* **1997**, 59-60.
20. Sgarbi, P. W. M. and Clive, D. L. J., *Chem. Commun.* **1997**, 2157-2158.
21. Gu, P.; Griebel, C.; Van Aerschot, A.; Rozenski, J.; Busson, R.; Gais, H.-J.; Herdewijn, P., *Tetrahedron*, **2004**, *60*, 2111-2118.
22. Berger, B.; Rabiller, C. G.; Konigsberger, K.; Faber, K.; Griengl, H., *Tetrahedron asymm.* **1990**, *1*, 541-546.
23. Thomas F, Jenny.; K, Christian Schneider.; Steven A, Benner., *Nucleosides & Nucleotides*, **1992**, *11*, 1257-1261.

## **CHAPTER 3 / SECTION-A**

**Design, synthesis and biophysical  
evaluation of Open chain analogues  
of cyclohexenyl nucleic acids**

## Section 3A: Design, synthesis and biophysical evaluation of Open chain analogues of cyclohexenyl nucleic acids

### 3A.1 Introduction

Nucleic acids store the hereditary information passed from one generation to the next. Chemically modified nucleic acid analogues are required to perform functions in DNA therapeutics. Artificial nucleotide analogues have been widely explored to modulate the properties of nucleic acids, making them promising therapeutic agents<sup>1</sup> and important building blocks for constructing molecular devices.<sup>2</sup> Researchers have investigated various artificial nucleic acid systems containing non-ribose sugars<sup>3</sup> as well as acyclic scaffolds. The acyclic nucleic acids grabbed the attention of chemists and biologists alike for nearly three decades because of their structural simplicity. Early work in the field of nucleoside mimetics led to the discovery that certain acyclic nucleosides such as acyclovir to possess significant therapeutic activity.<sup>4</sup> The acyclic nucleic acids such as FNA<sup>5</sup>, UNA<sup>6</sup>, GNA<sup>7</sup> and *iso*GNA<sup>8</sup> (Figure 1a) have also been examined as progenitor candidates to RNA in the early development of nucleic acids. FNA is an open chain analogue of DNA in which the 2'-carbon of ribose ring system was absent. Wengel designed unlocked RNA (Figure 1a, UNA) system by removing 2'-3' C-C bond in RNA.<sup>6</sup> GNA and FNA are the very interesting acyclic nucleoside candidates, as these molecules are structurally much simpler than natural DNA or RNA. Indeed, both acyclic nucleosides can be obtained under primitive reaction conditions using prebiotically plausible molecules.<sup>9</sup> Wengel also studied the glycerol-based nucleic acids<sup>6</sup> (Figure 1a, GNA) which would be less flexible than FNA or UNA due to less number of flexible bonds in the monomer unit.

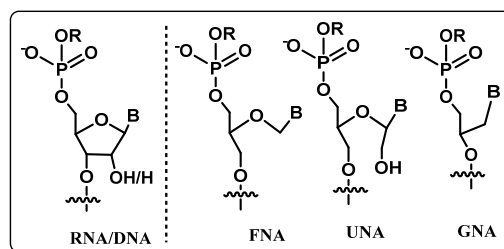


Figure 1a Natural DNA/RNA and acyclic analogues

## Chapter 3

---

All these acyclic structures were considered to be the plausible precursors of DNA/RNA, largely destabilized DNA /RNA duplexes. It was inferred that these acyclic structures would incur large entropic loss while duplex formation which was responsible for the incapability of acyclic nucleic acids to form stable cross-paired duplex structures with cDNA.<sup>5,6</sup>

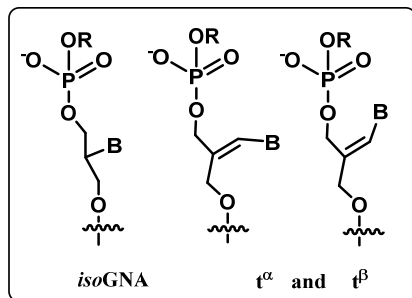
The incorporation of prochiral acyclic nucleosides into oligonucleotides causes the central C-atom of the monomer to become stereogenic, with each monomer adopting either a d-like or l-like orientation in the chain. However, the sugar backbone might, in the case of FNA or related acyclonucleotides, be sufficiently flexible to allow racemic mixtures of d- and l-like monomers to adopt similar conformations such that replication could take place without being affected by the absolute configuration of the stereogenic center. Later, homooligomeric chiral *R*-GNA and *S*-GNA were synthesized by Meggers which could form highly stable antiparallel helical duplex structures.<sup>7</sup> The optically pure (*S*)-GNA could also cross-pair with RNA though with much reduced stability and was proposed to be a potential precursor of RNA as genetic material. This means that the reduced flexibility in GNA as compared to FNA could lead to stable duplex structures when nucleobase attachment is kept flexible through methylene group. As a class of promising RNA progenitors, acyclic nucleic acid systems have been extensively investigated, and experimental evidence has strengthened the hypothesis that structurally simplified analogs might have played a transitional role, which eventually led to the emergence of catalytic RNA molecules.

### 3A.2 Design of Open chain analogues of Cyclohexenyl nucleic acids and rationale

The evolutionary chemistry with respect to nucleic acids suggested that simple acyclic nucleic acids might be preliminary nucleic acids<sup>10</sup> which ultimately have evolved as present day carriers of genetic information. To counter the entropic loss in acyclic nucleic acids, an attempt was made by introducing a double bond in the acyclic structure. Incorporation of these thymidine nucleosides mimics (Figure 1b,  $t^{\alpha}$  and  $t^{\beta}$ ) in oligomers was also found to be detrimental to the duplex stability similar to the other acyclic derivatives.<sup>11</sup> We presume that the attachment of nucleobase directly to the double bond in this case may have conferred considerable rigidity, leading to

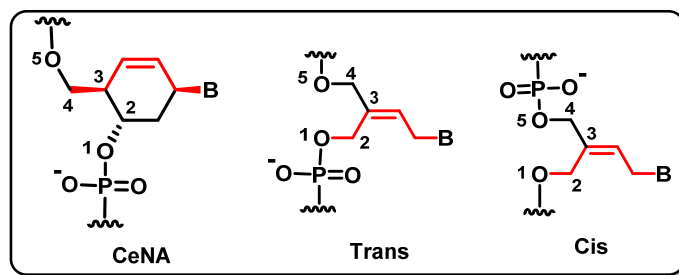
## Chapter 3

reduced ability of the nucleobase to take part in specific W-C hydrogen bonding. The *iso*GNA (Figure 1b, *iso*GNA) later studied by Krishnamurthy *et al* also destabilized duplexes, probably as the nucleobase attachment was directly to the backbone.<sup>8</sup>



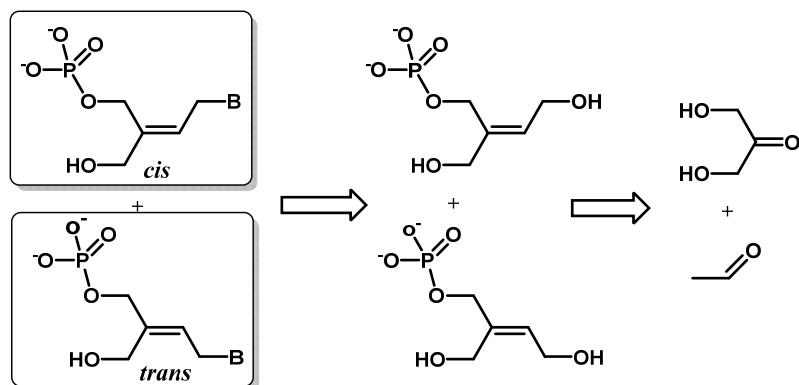
**Figure 1b** GNA and acyclic analogues

These modifications prompted us to visualize an open chain ene- nucleic acids (Figure 2a) in which the nucleobase attachment is to a planar double bonded structure through a methylene group, having same number of atoms in the backbone like natural sugar. This would also have a constraint of double bond unsaturation and act as an acyclic version of cyclohexene nucleic acid.



**Figure 2a** CeNA and Proposed *cis*- and *trans*- open chain analogues of CeNA

The proposed isoprenoid 5-carbon unit could be envisaged to be a part of evolutionary pre-biotic soup containing dihydroxy acetone and acetaldehyde (Figure 2b) similar to acrolein from formaldehyde and acetaldehyde.<sup>12</sup> At this stage, although the specific stereochemistry of prochiral glycerol or other D-sugars may not be explained chemically in the prebiotic world, the possibility of favored *cis* or *trans* geometry of the proposed ene-nucleotides and directionality thus imparted may not be ruled out.

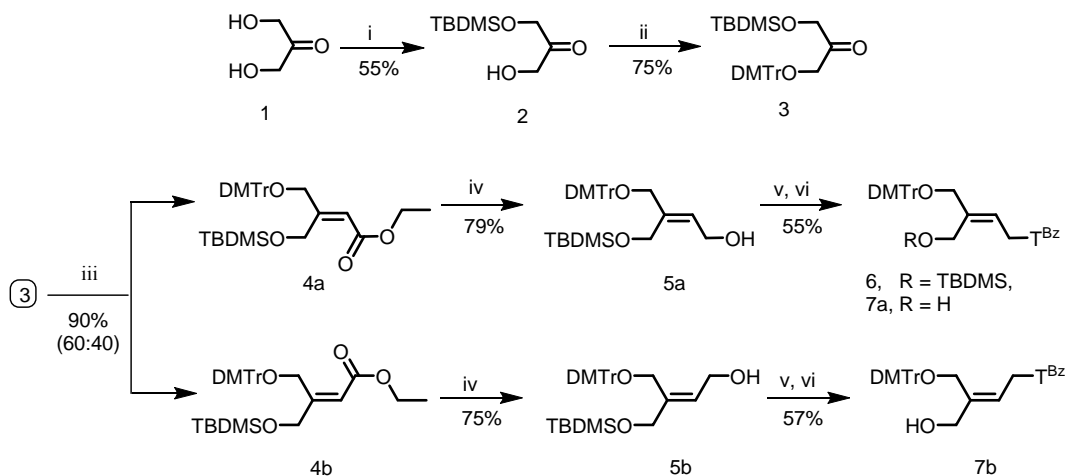


**Figure 2b** Proposed *cis*- and *trans* open chain ene-nucleotides

### 3A.3 Synthesis of *cis* and *trans* thymine monomers

The synthesis started with the mono TBS protection of commercially available dihydroxyacetone<sup>13</sup> to furnish compound **2**. Compound **2** subjected to DMTrCl in dry pyridine to result **3** and further reacted with ethylbromoacetate, PPh<sub>3</sub> under wittig reaction conditions yielded the geometrical isomeric products mixture of *trans* (**4a**) and *cis* (**4b**)  $\alpha$ ,  $\beta$ -unsaturated esters in more than 90% yield in 6:4 ratio. At this stage the two compounds *trans* (**4a**) and *cis* (**4b**) could be separated with very careful column chromatography and identified by nOe experiment. The DMTrO- group is considered to be corresponding to 5'-position and compound with nucleobase on the side of 5'-position is considered as *cis* isomer (scheme 1).

#### Scheme 1 Synthesis of key intermediates **7a**, **7b**



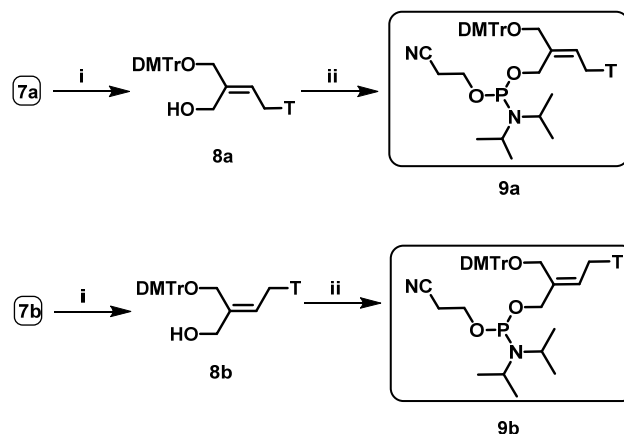


## Chapter 3

**Reagents and conditions** (i) TBDMSCl, imidazole, dry DMF (ii) DMTrCl, dry pyridine, overnight (iii) ethylbromoacetate, PPh<sub>3</sub>, Toluene, reflux, 5h (iv) DIBAL-H, dry DCM, -78 °C (v) T-Bz, PPh<sub>3</sub>, DIAD, dry dioxane, overnight (vi) TBAF in 1M THF, THF, rt, 2h.

Compounds **4a**, **4b** were treated with DIBAL-H at -78 °C in dry DCM individually to obtain the corresponding allylic alcohols **5a**, **5b**. Compound **5a** was subjected for Mitsunobu reaction condition, hydroxy group was substituted with benzoyl protected thymine (T<sup>Bz</sup>) nucleobase **6** which was further treated with TBAF yielded **7a**. In case of **5b** Mitsunobu product was contaminated with triphenylphosphineoxide. Due to the contamination with out purification reaction mixture subjected for silyl group deprotection using TBAF in THF yielded **7b** in 57 %. To obtain the *trans* and *cis* thymine monomers, compounds **7a**, **7b** subjected for ammonia treatment to remove the N<sup>3</sup>-benzoyl group of thymine nucleobase followed by phosphitylation of free hydroxyl group in **8a**, **8b** delivered the required amidite monomers **9a**, **9b** respectively, which could be incorporated into modified oligonucleotides using a DNA synthesizer (scheme 2).

**Scheme 2** Synthesis of *trans* and *cis* phosphoramidite monomers



**Reagents and conditions** (i) 30% aq ammonia solution, dioxane, rt, 4h (ii) 2-cyanoethyl-N,N-diisopropylchlorophosphine, DIPEA, dry DCM.

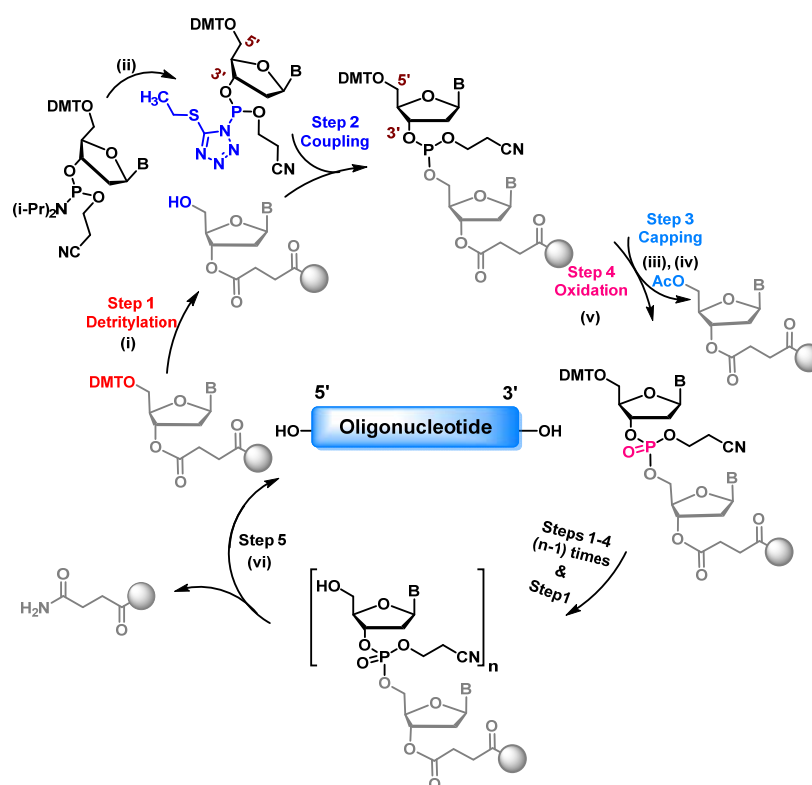
### 3A.4 Solid phase synthesis of oligonucleotide using phosphoramidite chemistry

Oligonucleotide synthesis can be carried out with two types of solid supports. One is the standard support with preloaded nucleosidic unit, which becomes the 3'-terminal residue, as the synthesis is done in 3'→5' direction. To use standard supports one requires four different supports for DNA synthesis carrying the individual canonical nucleobases. The other type is the universal solid support, which

## Chapter 3

carries an abasic sugar unit rather than the 3'-nucleoside unit. The advantage of using the universal support the possibility to synthesize 3'-end modified nucleic acids which is not possible using conventional standard supports. The natural DNA was synthesized in the 3' to 5' direction, on a polystyrene solid support with the required end nucleoside (A, T, C, G) attached to it via a linker. The oligonucleotides were cleaved from solid support by ammonia treatment, their purity ascertained by RP-HPLC on a C18 column, which further used for biophysical studies (scheme 3).

**Scheme 3** The sequence of chemical reactions involved in the solid phase synthesis of oligonucleotide



**Reagents and conditions:** (i) 3% TCA in DCM (ii) 0.25 M 5-(S-ethyltetrazole) in ACN (iii) Ac<sub>2</sub>O/Py (iv) 10% N-methylimidazole in THF 0.1 M (v) I<sub>2</sub>/Py/H<sub>2</sub>O/THF (vi) aq. NH<sub>3</sub>, 55 °C.

### 3A.5 Synthesis of modified oligonucleotides, characterization and UV-T<sub>m</sub> studies

The 18mer DNA sequence chosen for the current study of biological relevance, **DNA1** is used for miRNA down-regulation.<sup>14</sup> Unmodified oligomers were synthesized using a Bioautomation MM4 DNA synthesizer by using commercially available phosphoramidite building blocks. Modified oligonucleotides were

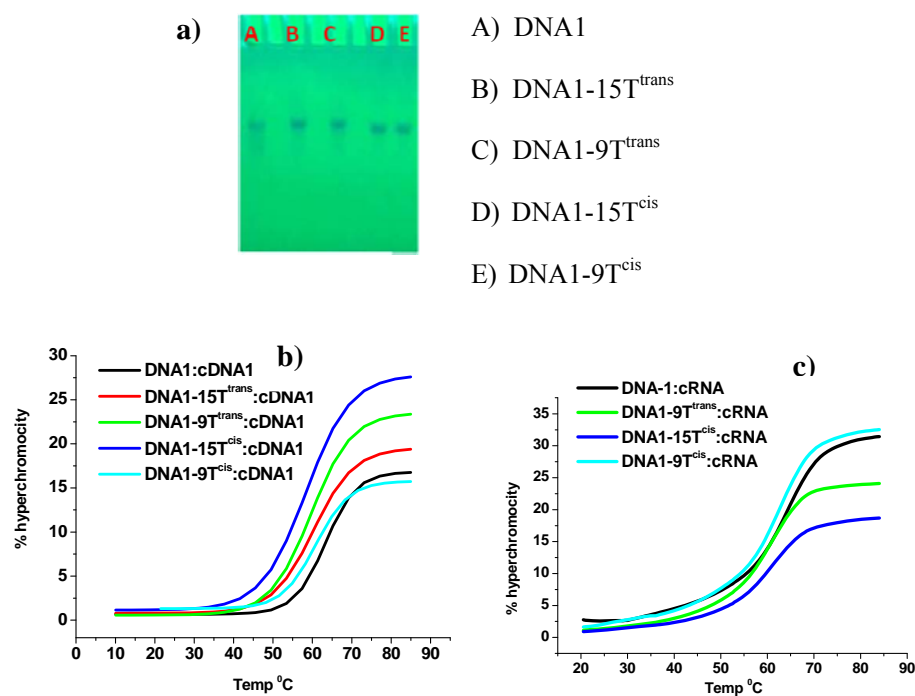
### Chapter 3

synthesized on Bioautomation MM4 DNA synthesizer, using phenoxyacetyl (Pac) protected cyanoethyl phosphoramidites and modified amidite building blocks **9a**, **9b**. The modified phosphoramidites 0.1M in CH<sub>3</sub>CN were manually coupled for 6min, followed by washing step with 10% H<sub>2</sub>O, 0.2% Ac<sub>2</sub>O, 0.2% Lutidine v/v/v in THF done to avoid the unwanted phosphorylation at bases of highly reactive acyclic olefinic monomers. After washing step, capping followed by oxidation with 0.5M tert-butyl hydroperoxide in CH<sub>2</sub>Cl<sub>2</sub>-acetone (1:1) used instead of iodine/water, because it is known that iodine/water cleaved the allylic C-O bond. This is known to occur for other phosphites with allylic or tertiary substituents<sup>12</sup>. Deprotection and cleavage were performed by shaking the support bound oligonucleotide with neat dry diisopropylamine, washing with diethylether followed by shaking with conc aqueous ammonia for 2h at rt<sup>12</sup>. The crude oligomer was purified by RP-HPLC, purity confirmed by gel-electrophoretic mobility studies (Figure 3a) and characterized by MALDI-TOF mass spectrometry (Table 1).

**Table 1** Modified DNA sequences, their MALDI-TOF mass analyses and biophysical evaluation by UV-*T<sub>m</sub>* measurements<sup>a</sup>

Name	Sequence <sup>b</sup> 5'→3'	mass cal /obs	UV <i>T<sub>m</sub></i> °C	
			DNA <sup>c</sup>	RNA <sup>d</sup>
DNA1	caccattgtcacactcca	5363/5367	63.5	62.7
DNA1-15T <sup>trans</sup>	caccattgtcacacT <sup>trans</sup> cca	5347/5342	60.2	-
DNA1-9T <sup>trans</sup>	caccattgT <sup>trans</sup> cacactcca	5347/5347	59.6	61
DNA1-15T <sup>cis</sup>	caccattgtcacacT <sup>cis</sup> cca	5347/5344	59.3	59.1
DNA1-9T <sup>cis</sup>	caccattgT <sup>cis</sup> cacactcca	5347/5344	60.8	59.6

<sup>a</sup>UV-*T<sub>m</sub>* values were measured by using 1μM sequences with 1μM cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is ±0.5 °C). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5'tggagtgtgacaatgggtg was the complementary DNA sequence. <sup>d</sup>5' uggagugugacaaggug was the complementary RNA sequence.



**Figure 3** (a) Gel picture of purified sequences and UV melting profiles of modified DNA sequences with, (b) cDNA, (c) cRNA

It is seen that the sequences modified with *T-cis* as well as *T-trans* are able to form stable duplexes with both DNA as well as RNA independent of the site of modification *i.e* towards 3'-end or in the middle of the sequence. The destabilization observed is 2-4°C in each case (Table 1).

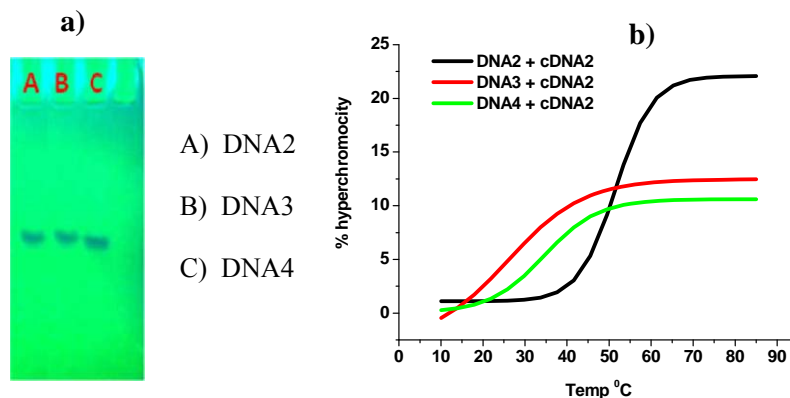
We further studied multiple modifications in the sequence containing continuous stretch of *cis*-thymine units so that the modified units could be inserted continuously or alternately in the sequence. In homothyminyll sequences the acyclic units were seldom tolerated and the duplexes formed were destabilized (Table 2).

**Table 2** Modified DNA sequences, their MALDI-TOF mass analysis and biophysical evaluation by UV- $T_m$  measurements<sup>a</sup>

Name	Sequence <sup>b</sup> 5'→3'	mass cal / obs	UV $T_m$ °C DNA <sup>c</sup>
DNA2	gcgtttttgct	3633/3635	51
DNA3	gcgttT <sup>cis</sup> T <sup>cis</sup> T <sup>cis</sup> tgct	3585/3586	26
DNA4	gcgT <sup>cis</sup> tT <sup>cis</sup> tT <sup>cis</sup> tgct	3585/3582	34

## Chapter 3

<sup>a</sup>UV- $T_m$  values were measured by using 1 $\mu$ M sequences with 1 $\mu$ M cDNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is  $\pm 0.5$  °C). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5'agcaaaaaacgc was the complementary DNA sequence.

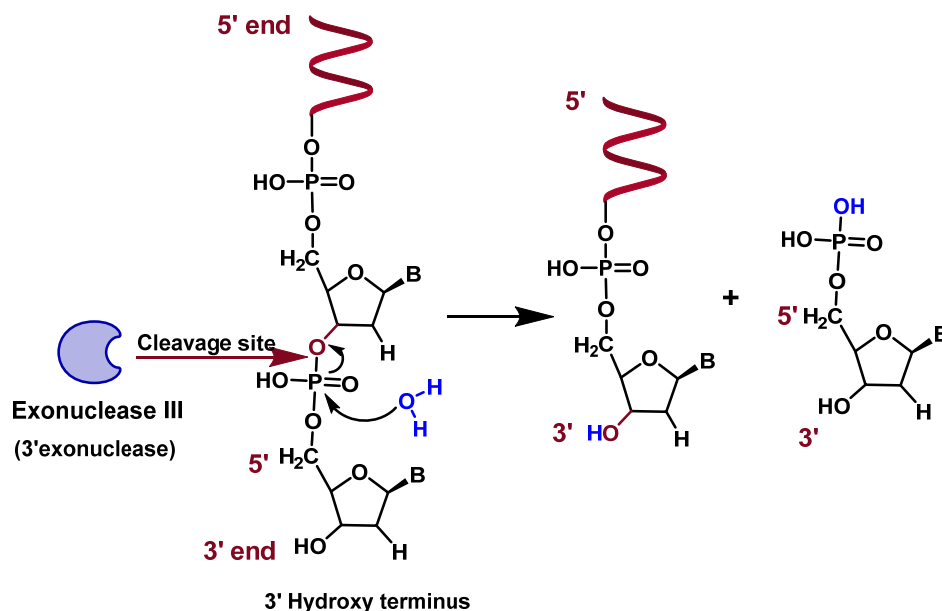


**Figure 4** (a) Gel picture of purified sequences, (b) UV melting profiles of modified DNA sequences with cDNA

### 3A.6 Stability of oligonucleotides to SVPD

The antisense technology AS-ONs must have specific characteristic properties such as sequence specificity, high target-binding affinity, and nuclease resistance. Among all of these, nuclease resistance is key factor, which determines the effectiveness of AS-ONs *in vivo*. The *exo*- and *endo*- nucleases cleave the phosphodiester bonds between the nucleotide subunits of AS-ONs *in vivo*. Exonucleases are enzymes which cleave nucleotides from the ends (*exo*) (either the 3'- or the 5'-end) of a polynucleotide and one at a time, where as endonucleases, cleave phosphodiester bonds in the middle (*endo*) of a polynucleotide chain. 3'-exonuclease catalyzes the degradation of nucleic acid in 3'→5' direction with the removal of a nucleoside-5'-phosphate from the 3'-end of DNA (Figure 5).

Natural DNA/RNA oligonucleotides are not good antisense candidates mainly because of lack of nuclease resistance. This has encouraged enormous chemical efforts to produce therapeutically significant modified ONs to provide improved stability, with maintaining high target affinity and specificity. First generation AS-ONs (phosphorothioate) have shown improved nuclease resistance but reduced affinity towards a complementary RNA sequence.<sup>15</sup> Second generation of AS-ONs involves various 2'-modifications<sup>16</sup> of the ribose moiety, results considerable improvement of binding affinity to the target RNA.



**Figure 5** Exonucleolytic cleavage in 3'- to 5'-direction to yield nucleoside 5'-phosphates

In this case, size, electronegativity, and configuration of the 2'-substituents are very important factors for the target affinity and nuclease resistance.<sup>17</sup> The new third generation of AS-ONs, containing conformationally constrained LNA (also called BNA)<sup>18</sup> shows unusual high affinity towards complementary RNA strand (3-8°C per modification). Unfortunately, though LNA containing AS-ONs are more nucleolytically stable than the native PO AS-ONs, they do not have nuclease resistance as good as that of the PS AS-ONs.<sup>19</sup>

### 3A.7 Enzymatic stability studies of modified DNA sequences

Considering nuclease resistance as an important factor, we examined the 3'-exonuclease sensitivity of unmodified homooligomer  $t_{10}$  as well as *cis* and *trans* modified sequences  $t_8T^{cis}t$  and  $t_8T^{trans}t$  by using phosphodiesterase I from *Crotalus adamanteus* venom [snake venom phosphodiesterase (SVPD)] (Table 3). Enzymatic hydrolysis of the ONs (7.5  $\mu$ M) was carried out at 37 °C in buffer (100  $\mu$ l) containing 100 mM Tris-HCl (pH 8.5), 15 mM MgCl<sub>2</sub>, 100 mM NaCl and SVPD (10 $\mu$ g/mL). Aliquots were removed at several time-points; a portion of each reaction mixture was removed and heated to 90 °C for 2 min to inactivate the nuclease. The amount of intact ONs was analyzed at several time points by RP-HPLC (Figure 6). The

## Chapter 3

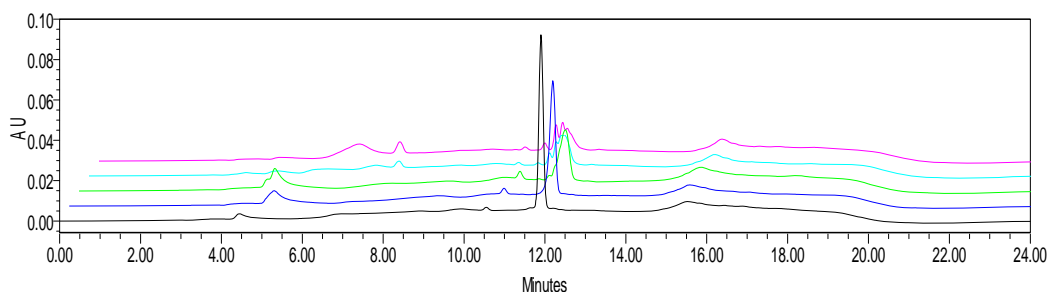
percentage of intact ON was then plotted against the exposure time to obtain the ON degradation curve with time.

**Table 3** Oligomers used for 3'-exonuclease degradation study

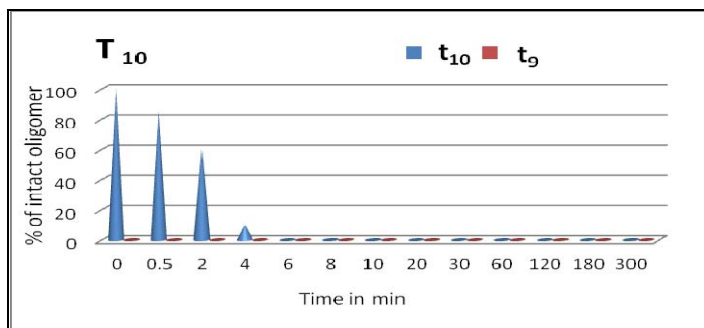
Name	Sequence 5'→3'	mass cal /obs
$t_{10}$	ttttttttt	2980/2977
$t_8T^{cis}t$	tttttttT <sup>cis</sup> t	2964/2961
$t_8T^{trans}t$	tttttttT <sup>trans</sup> t	2964/2961

The unmodified single strand DNA sequence  $t_{10}$  did not show any 3'-exonuclease resistance and was completely degraded within 4min. The synthesis of  $t_8T^{trans}t$  carried out by using standard support with preloaded thymine nucleosidic unit. We incorporated our *trans* modified monomer at 9<sup>th</sup> position from 5'- end of 10mer sequence. 3'- terminal thymidine residue was started cleaving within 0.5min and  $t_8T^{trans}$  fragment was observed simultaneously. The sequence  $t_8T^{trans}$  was stable up to 10min, which is comparatively stable than unmodified  $t_{10}$ . In case of  $t_8T^{cis}t$ , 3'- terminal thymidine residue was completely cleaved within 4min and  $t_8T^{cis}$  was observed to be stable up to 5h, which was clearly indicated the high stability of *cis* modified sequence.

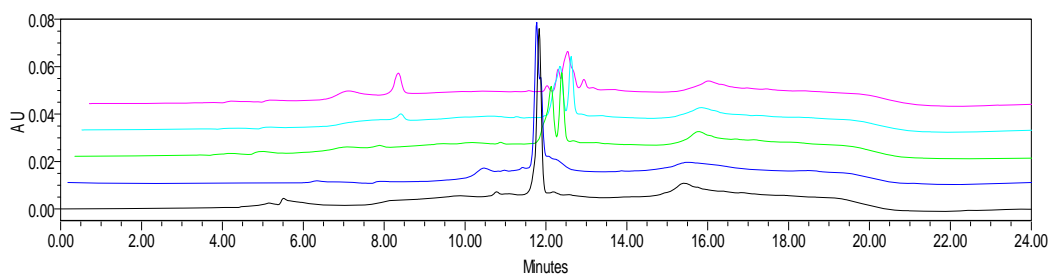
### $t_{10}$ : RP-HPLC at several time-points for SVPD study



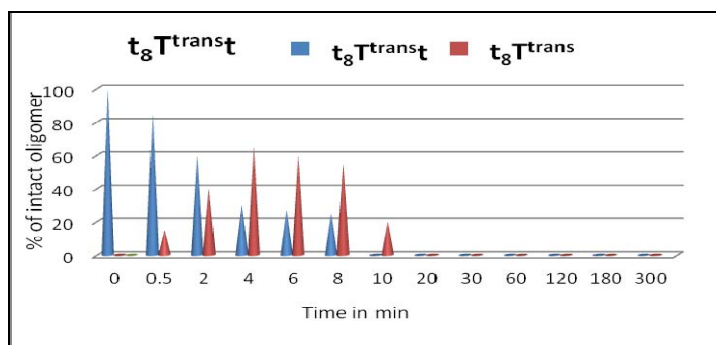
(Black- No SVPD, Blue- 0.5min, Green-2min,cyan-4min, Magenta-6min)



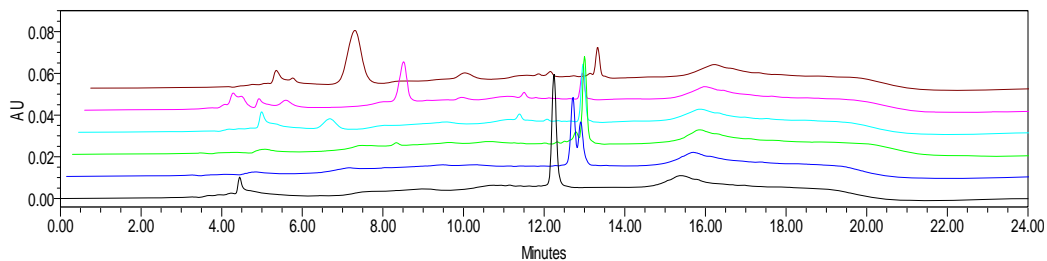
**$t_8T^{trans}$** : RP-HPLC at several time-points for SVPD study



(Black- No SVPD, Blue- 0.5min, Green-4min, Cyan-8min, Magenta-20min)

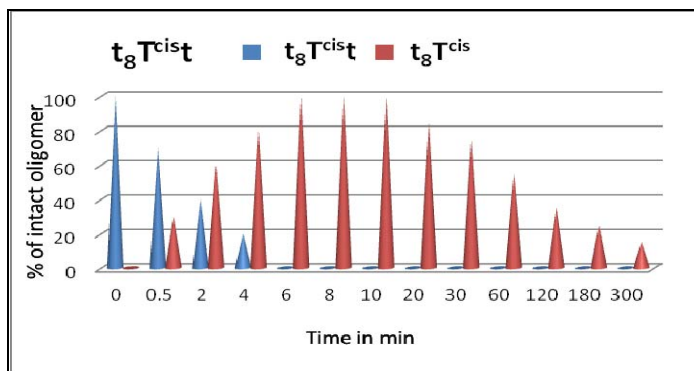


**$t_8T^{cis}$** : RP-HPLC at several time-points for SVPD study



(Black- No SVPD, Blue- 0.5min, Green- 6min, Cyan-1h, Magenta-3h, Brown-5h)





**Figure 6:** RP-HPLC at several time-points for SVPD study for  $t_{10}$ ,  $t_8T^{trans}$ ,  $t_8T^{cis}$

### 3A.8 Conclusions:

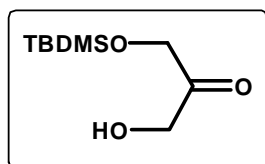
- Successfully *cis* and *trans* open chain analogues of cyclohexenyl nucleosides were synthesized from a prebiotic precursor, dihydroxy acetone.
- These *cis* and *trans* modified nucleosides were incorporated for the first time in mixed Pu/Py DNA sequences and purity was confirmed by gel-electrophoretic mobility studies.
- Duplexes with cDNA and cRNA are quite stable compared to other known acyclic modifications.
- Alternative modifications are stabilizing the duplex structure with cDNA than consecutive positions.
- Nuclease resistance of *cis* and *trans* modified sequences were stable to SVPD enzyme than unmodified  $t_{10}$  and the stability order was  $t_8^{T^{cis}} > t_8^{T^{trans}} > t_{10}$ .

### 3A.9 Experimental section

#### General information:

All the non-aqueous reactions were carried out under the inert atmosphere of Nitrogen/ Argon and the chemicals used were of laboratory or analytical grade. All solvents used were dried and distilled according to standard protocols. TLCs were carried out on precoated silica gel 60 F254 (Merck). Column chromatographic separations were performed using silica gel 60- 120 mesh (Merck) or 200- 400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether and MeOH/DCM. <sup>1</sup>H and <sup>13</sup>C NMR spectra were obtained using Bruker AC-200, AC-400 and AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for <sup>1</sup>H. High resolution mass spectra were recorded on a Thermo Fisher Scientific Q Exactive mass spectrometer.

#### 1-((tert-butyldimethylsilyl)oxy)-3-hydroxypropan-2-one (2)



Compound **1** (9.4 g, 104.4 mmol) was dissolved in dry DMF (100 mL), then TBS-Cl (5.0 g, 33.5 mmol) and imidazole (2.95 g, 43.4 mmol) was added under nitrogen atmosphere. The reaction mixture was stirred at room temperature for 10h then quenched with water (100 mL). Compound was extracted with ethyl acetate from crude reaction mixture and organic layer washed with brine solution, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated on rotavapor *in vacuo*. Crude compound purified through column chromatography (pet ether:EtOAc, 90:10) to result **2** (6.2 g, 55%) as a colour less thick liquid.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 0.10 (s, 6 H), 0.93 (s, 9 H), 3.01 (t, *J*=4.99 Hz, 1 H), 4.32 (s, 2 H), 4.51 (d, *J*=4.93 Hz, 2 H) ppm; <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>) δ -5.7, 18.1, 25.7, 66.6, 67.7, 211.1 ppm; HRMS (EI): Mass calculated for C<sub>9</sub>H<sub>20</sub>O<sub>3</sub>NaSi (M+ Na<sup>+</sup>), 227.1074, found 227.1069.

#### 1-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-((tert-butyldimethylsilyl)oxy)propan-2-one (3)

To a solution of **2** (5 g, 24.5 mmol) in pyridine (15 mL) DMTr chloride (10g, 29.5 mmol) and catalytic amount of DMAP were added, stirred at rt for 6h. Pyridine was



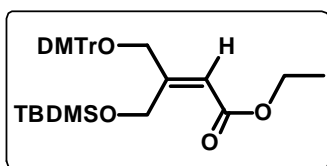
removed *in vacuo* and the residue was diluted with EtOAc. Water wash and brinewash were given to the organic layer, dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated in *vacuo*. The residue was subjected to silica gel column chromatography (pet ether:EtOAc, 95:5) to afford **3** (9.3 g) in 75% yield.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 0.02 (s, 6 H), 0.85 (s, 9 H), 3.80 (s, 7 H), 3.96 (s, 2 H), 4.38 (s, 2 H), 6.79 - 6.89 (m, 5 H), 7.26 - 7.39 (m, 9 H), 7.41 - 7.49 (m, 2 H) ppm; <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>) δ -5.6, 18.2, 25.7, 55.2, 68.2, 68.4, 86.9, 113.3, 127.0, 128.0, 130.0, 135.4, 144.3, 158.7, 206.8 ppm; HRMS (EI): Mass calculated for C<sub>30</sub>H<sub>38</sub>O<sub>5</sub>NaSi (M+ Na<sup>+</sup>) 529.2381, found 529.2369.

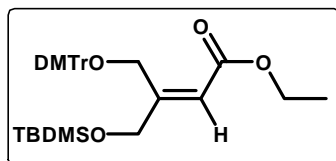
**ethyl(Z)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(((tertbutyldimethylsilyl)oxy) methyl)but-2-enoate (4a & 4b)**

Solution of **3** (10g, 19.7 mmol) and two carbon wittig ylide (9.5 g, 29.6 mmol) in 100 mL toluene was refluxed for 4h. Solvent was removed *in vacuo*, residue diluted with the EtOAc and water wash, saturated aqueous NaHCO<sub>3</sub> wash and finally brine wash were given. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated in *vacuo* followed by column chromatography (pet ether: EtOAc, 98:2) to give a **4a** and **4b** (90%) in 60:40 ratio.

<sup>1</sup>H NMR(**4a**) (200 MHz, CDCl<sub>3</sub>) δ -0.08 (s, 6 H), 0.72 (s, 9 H), 1.34 (t, *J*=7.14 Hz, 4 H), 3.80 (s, 6 H), 3.90 (s, 2 H), 4.21 (q, *J*=7.07 Hz, 2 H), 4.81 (s, 2 H), 6.38 (t, *J*=1.77 Hz, 1 H), 6.84 (d, *J*=8.72 Hz, 4 H), 7.23 - 7.39 (m, 8 H), 7.41 - 7.49 (m, 2 H) ppm; <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ -5.7, 14.4,



18.0, 25.7, 55.2, 59.9, 61.9, 64.0, 86.6, 112.2, 113.2, 126.8, 127.9, 128.0, 129.9, 136.1, 144.8, 158.5, 160.1, 166.7 ppm; HRMS (EI): Mass calculated for C<sub>34</sub>H<sub>44</sub>O<sub>6</sub>NaSi (M+ Na<sup>+</sup>) 599.2799, found 599.2789.

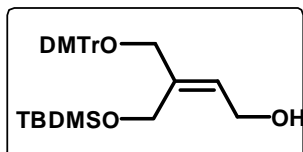


<sup>1</sup>H NMR(**4b**) (200 MHz, CDCl<sub>3</sub>) δ 0.10 (s, 6 H), 0.94 (s, 9 H), 1.19 (t, *J*=7.14 Hz, 4 H), 3.77 - 3.81 (m, 8 H), 4.05 (q, *J*=7.07 Hz, 2 H), 4.39 (s, 2 H), 4.52 (s, 2 H), 5.94 - 5.99 (m, 1 H), 6.82 (d, *J*=8.84 Hz, 5 H), 7.25 - 7.43 (m, 13 H) ppm; <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ -5.4, 14.3, 18.4, 26.0, 55.2,

### Chapter 3

59.8, 62.3, 63.5, 86.5, 113.0, 113.1, 126.8, 127.8, 128.1, 129.1, 129.9, 130.0, 136.0, 144.8, 158.5, 158.8, 166.4 ppm; **HRMS (EI)**: Mass calculated for  $C_{34}H_{44}O_6NaSi$  ( $M+Na^+$ ) 599.2799, found 599.2790.

#### **(Z)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(((tertbutyldimethylsilyl)oxy)methyl) but-2-en-1-ol (5a)**

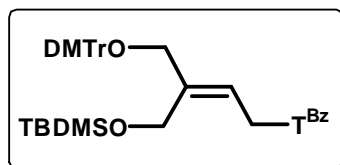


DIBAL-H was added to a solution of ester **4a** (1g, 2.6 mmol) in DCM at  $-78\text{ }^{\circ}\text{C}$ . After 45min at same temperature aq. sodium potassium tartarate and diethyl ether added. The resultant cloudy reaction mixture was then vigorously stirred for 1h at which organic layer appears like clear solution. Organic layer was separated and washed with brine solution and extracted with DCM, dried over  $Na_2SO_4$ . Compound was purified through column chromatography (pet ether:EtOAc, 70:30) to obtain **5a** (0.71g) in 79% yield.

**$^1H$  NMR** (200 MHz,  $CDCl_3$ )  $\delta$  -0.01 (s, 6 H), 0.82 (s, 9 H), 3.62 (s, 2 H), 3.80 (s, 8 H), 4.19 (s, 2 H), 4.22 - 4.31 (m, 2 H), 6.02 (t,  $J=6.57$  Hz, 1 H), 6.83 (d,  $J=8.84$  Hz, 6 H), 7.25 (d,  $J=2.65$  Hz, 2 H), 7.28 - 7.52 (m, 11 H) ppm;  **$^{13}C$  NMR** (50 MHz,  $CDCl_3$ )  $\delta$  -5.5, 18.2, 25.8, 55.2, 58.7, 59.9, 65.3, 86.2, 113.1, 126.2, 126.7, 127.8, 128.1, 130.0, 136.3, 139.5, 145.0, 158.4 ppm; **HRMS (EI)**: Mass calculated for  $C_{32}H_{42}O_5NaSi$  ( $M+Na^+$ ) 557.2694, found 557.2677.

#### **(Z)-3-benzoyl-1-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(((tertbutyldimethylsilyl)oxy)methyl)but-2-en-1-yl)-5-methylpyrimidine-2,4(1H,3H)dione (6)**

To a solution of **5a** (0.5g, 0.93 mmol) in dry dioxane (4ml) was added triphenyl phosphine(0.37g, 1.4mmol) and  $N^3$ -benzoyl protected thymine (0.32g, 1.4mmol) stirred for 15min. DIAD (0.36mL, 1.86mmol) was dissolved in 1mL dry dioxane and added to the reaction mixture, continued the stirring for overnight at room temperature. Dioxane was removed *in vacuo* and the residue was diluted with EtOAc. Water wash and brine wash were given to the organic layer, dried over  $Na_2SO_4$ , concentrated in vacuo. Crude residue was subjected

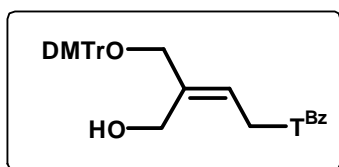


### Chapter 3

to silica gel column chromatography (pet ether:EtOAc, 70:30) to obtain **6** (0.38 g) in 55% yield.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 0.01 (s, 6 H), 0.83 (s, 9 H), 1.97 (s, 3 H), 3.67 (s, 2 H), 3.80 (s, 6 H), 4.23 (s, 2 H), 4.57 (d, *J*=7.58 Hz, 2 H), 5.79 (t, *J*=7.71 Hz, 1 H), 6.84 (d, *J*=8.84 Hz, 5 H), 7.23 (br. s., 2 H), 7.29 - 7.54 (m, 13 H), 7.59 - 7.69 (m, 1 H), 7.91 - 7.98 (m, 2 H) ppm; <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ -5.3, 12.6, 18.3, 25.9, 25.9, 44.3, 55.3, 59.7, 65.4, 86.6, 110.9, 113.2, 120.0, 126.9, 128.0, 128.1, 129.2, 130.0, 130.6, 131.7, 135.0, 136.1, 139.6, 142.9, 144.9, 150.1, 158.6, 163.3, 169.3 ppm, LCMS (EI): Mass calculated for C<sub>44</sub>H<sub>50</sub>N<sub>2</sub>O<sub>7</sub>NaSi (M+ Na<sup>+</sup>) 769.3387, found 769.17.

#### (Z)-3-benzoyl-1-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(hydroxymethyl)but-2-en-1-yl)-5-methylpyrimidine-2,4(1H,3H)-dione (**7a**)



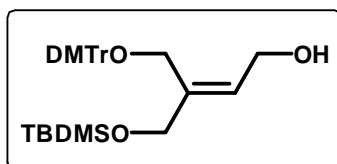
Compound **6** (1g, 1.28mmol) was dissolved in 15mL THF and TBAF (0.394 g, 1.5 mmol) was added. The reaction mixture was stirred for 2h at room temperature. The solvent was removed in vacuo. The residue was dissolved into 50 mL of ethyl acetate, washed with water (3 x 25 mL) and then with brine. The organic layer dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The resulting residue was purified on silica gel column chromatography (pet ether:EtOAc, 60:40) to yield **7a** (0.7g) in 83%.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 1.89 (s, 3 H), 3.68 (s, 3 H), 3.71 (s, 8 H), 4.10 (s, 2 H), 4.41 (d, *J*=7.58 Hz, 2 H), 5.62 (t, *J*=7.64 Hz, 1 H), 6.76 (d, *J*=8.84 Hz, 5 H), 7.12 - 7.24 (m, 9 H), 7.29 - 7.43 (m, 6 H), 7.48 - 7.61 (m, 2 H), 7.79 - 7.88 (m, 3 H) ppm; <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 12.4, 25.8, 44.2, 55.2, 59.6, 65.3, 86.5, 110.8, 113.1, 119.9, 126.8, 127.9, 128.0, 129.1, 129.9, 130, 131.6, 134.9, 136.0, 139.5, 140.0, 142.8, 144.7, 150.0, 158.5, 163, 169.2 ppm; HRMS (EI): Mass calculated for C<sub>38</sub>H<sub>36</sub>O<sub>7</sub>N<sub>2</sub>Na (M+ Na<sup>+</sup>) 655.2415, found 655.2398.

#### (E)-4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(((tertbutyldimethylsilyl)oxy)methyl)but-2-en-1-ol (**5b**)

### Chapter 3

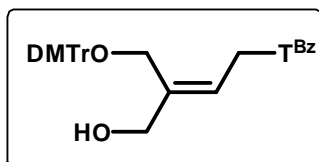
DIBAL-H was added to a solution of ester **4b** (1g, 2.6 mmol) in DCM at -78 °C. After 45min at same temperature aq. sodium potassium tartarate and diethyl ether added. The resultant cloudy reaction mixture was then vigorously stirred for 1h at which organic layer appears like clear solution. Organic layer was separated and



washed with brine solution and extracted with DCM, dried over Na<sub>2</sub>SO<sub>4</sub>. Compound was purified through column chromatography (pet ether:EtOAc, 70:30) to obtain **5b** (0.67g) in 75% yield.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 0.08 (s, 6 H), 0.92 (s, 9 H), 3.66 (s, 2 H), 3.80 (s, 7 H), 4.07 (d, *J*=6.82 Hz, 2 H), 4.22 (s, 2 H), 5.88 (t, *J*=6.82 Hz, 1 H), 6.85 (d, *J*=8.84 Hz, 5 H), 7.24 - 7.50 (m, 12 H) ppm; <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ -5.3, 18.4, 26.0, 55.2, 58.8, 59.6, 65.1, 86.6, 113.2, 113.3, 113.3, 126.4, 126.8, 127.9, 128.0, 128.1, 129.9, 130.0, 130.0, 136.1, 139.3, 144.9, 158.5 ppm; HRMS (EI): Mass calculated for C<sub>32</sub>H<sub>42</sub>O<sub>5</sub>NaSi (M+ Na<sup>+</sup>) 557.2694, found 557.2684.

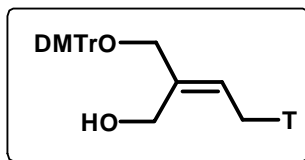
#### (E)-3-benzoyl-1-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(hydroxymethyl)but-2-en-1-yl)-5-methylpyrimidine-2,4(1H,3H)-dione (**7b**)



Compound **5b** was subjected for Mitsunobu reaction and without purification the crude mixture used for TBDMS deprotection to obtain **7b** in 57% yield over two steps.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 1.86 (s, 3 H), 3.78 (s, 7 H), 3.80 (s, 2 H), 4.19 - 4.23 (m, 4 H), 5.65 (t, *J*=7.02 Hz, 1 H), 6.85 (d, *J*=8.85 Hz, 5 H), 7.27 - 7.37 (m, 8 H), 7.42 - 7.50 (m, 5 H), 7.90 (d, *J*=7.32 Hz, 2 H) ppm; <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>) δ 12.3, 44.9, 55.3, 59.6, 65.5, 87.0, 110.9, 113.4, 121.7, 127.1, 128.0, 128.1, 129.1, 130.0, 130.5, 131.7, 135.0, 135.5, 139.5, 142.2, 144.4, 149.8, 158., 163.1, 169.1 ppm; HRMS (EI): Mass calculated for C<sub>38</sub>H<sub>36</sub>O<sub>7</sub>N<sub>2</sub>Na (M+ Na<sup>+</sup>) 655.2415, found 655.2396.

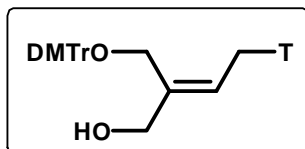
#### (Z)-1-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(hydroxymethyl)but-2-en-1-yl)-5-methylpyrimidine-2,4(1H,3H)-dione (**8a**)



30% aq. ammonia solution (0.5mL) was added to a solution of **7a** (0.5g, 0.76mmol) in 10mL dioxane and stirred for 7h at room temperature. The solvent was removed under reduced pressure. The residue was dissolved into 50 mL of ethyl acetate, washed with water (3 x 25 mL) and then with brine. The organic layer dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The resulting residue was purified on silica gel column. The product was eluted with 50% ethyl acetate in petroleum ether to afford **8a** (0.4 g, 85 %) as a white solid.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 1.83 (s, 3 H), 3.80 (s, 7 H), 3.81 (br. s., 2 H), 4.17 (d, *J*=7.02 Hz, 2 H), 4.23 (br. s., 2 H), 5.62 (t, *J*=7.02 Hz, 1 H), 6.85 (d, *J*=9.16 Hz, 5 H), 7.21 - 7.26 (m, 1 H), 7.28 - 7.36 (m, 7 H), 7.44 (d, *J*=7.02 Hz, 2 H), 8.82 (br. s., 1 H) ppm; <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 12.2, 44.7, 55.3, 59.7, 65.5, 87.0, 110.9, 113.3, 122.1, 127.1, 128.0, 128.0, 130.0, 135.5, 139.8, 141.8, 144.4, 150.8, 158.7, 164.1 ppm; HRMS (EI): Mass calculated for C<sub>31</sub>H<sub>32</sub>O<sub>6</sub>N<sub>2</sub>Na (M+ Na<sup>+</sup>) 551.2153, found 551.2147.

**(E)-1-(4-(bis(4-methoxyphenyl)(phenyl)methoxy)-3-(hydroxymethyl)but-2-en-1-yl)-5-methylpyrimidine-2,4-dione (8b)**



30% aq. ammonia solution (0.5mL) was added to a solution of **7b** (0.5g, 0.76mmol) in 10mL dioxane and stirred for 7h at room temperature. The solvent was removed under reduced pressure. The residue was dissolved into 50 mL of ethyl acetate, washed with water (3 x 25 mL) and then with brine. The organic layer dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The resulting residue was purified on silica gel column. The product was eluted with 50% ethyl acetate in petroleum ether to afford **8b** (0.33 g, 80 %) as a white solid.

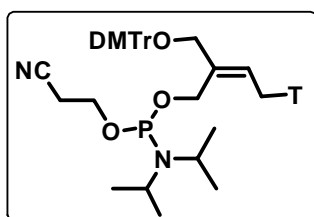
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 1.92 (s, 3 H), 3.76 (s, 2 H), 3.77 (s, 6 H), 4.19 (s, 2 H), 4.43 (d, *J*=7.63 Hz, 2 H), 5.62 (t, *J*=7.63 Hz, 1 H), 6.81 (d, *J*=8.85 Hz, 4 H), 7.21 (d, *J*=7.32 Hz, 1 H), 7.24 - 7.32 (m, 7 H), 7.40 (d, *J*=7.32 Hz, 2 H), 9.41 (br. s., 1 H) ppm; <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 12.3, 45.4, 55.2, 58.7, 66.3, 86.8, 111.3, 113.2, 120.6, 126.9, 127.9, 128.1, 130.0, 135.9, 140.2, 142.6, 144.7, 151.2, 158.6, 164 ppm;

### Chapter 3

**HRMS (EI):** Mass calculated for  $C_{31}H_{32}O_6N_2Na$  ( $M^+ Na^+$ ) 551.2153, found 551.2145.

#### General procedure followed for synthesis of phosphoramidite derivatives 9a, 9b

To the compound **8a, 8b** (100 mg, 0.17 mmol) dissolved in dry DCM (3 mL), DIPEA (0.64 mmol, 0.12 mL) was added. 2-cyanoethyl-*N,N*-diisopropyl-chloro phosphine (0.35 mmol, 0.08 mL) was added to the reaction mixture at 0 °C and stirring continued at room temperature for 1 hour. The contents were diluted with DCM and washed with 5%  $NaHCO_3$  solution. The organic phase was dried over anhydrous sodium sulphate and concentrated to white foam. The residue was re-dissolved in DCM and the compound was precipitated with *n*-hexane to obtain corresponding phosphoramidite derivatives in 70-75 % yield.

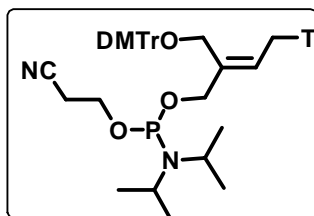


#### phosphoramidite derivative 9a

$^{31}P$  NMR (500MHz,  $CDCl_3$ )  $\delta$  148.82

**HRMS (EI):** Mass calculated for  $C_{40}H_{49}O_7N_4NaP$  ( $M^+ Na^+$ ) 751.3231, found 751.3212.

#### phosphoramidite derivative 9b



$^{31}P$  NMR (500MHz,  $CDCl_3$ )  $\delta$  148.27

**HRMS (EI):** Mass calculated for  $C_{40}H_{49}O_7N_4NaP$  ( $M^+ Na^+$ ) 751.3231, found 751.3212.



## Chapter 3

### 3A.10 Appendix

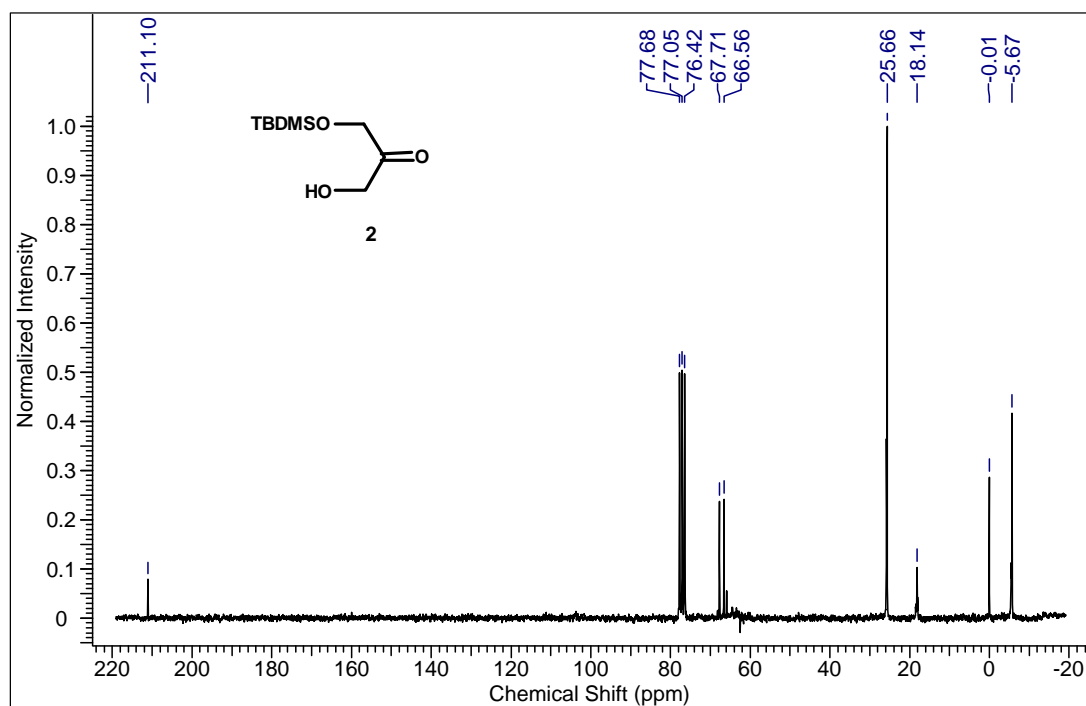
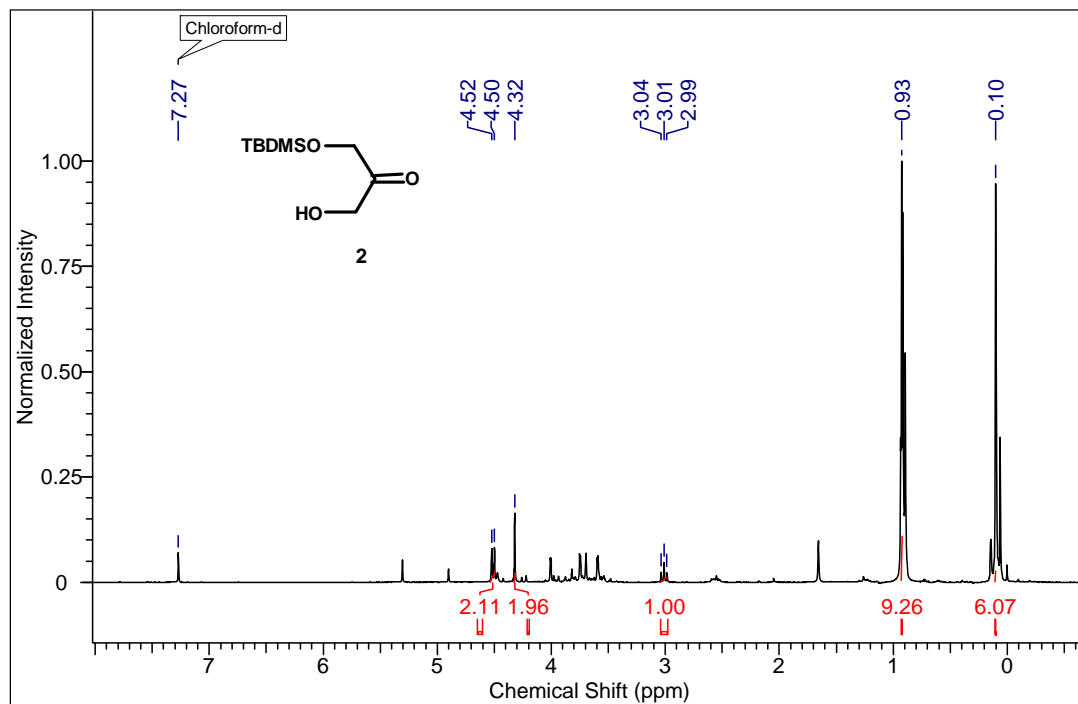
Compounds - Spectral data	Page No.
$2^1\text{H}$ NMR & $^{13}\text{C}$ NMR	112
2- DEPT & HRMS	113
$3^1\text{H}$ NMR & $^{13}\text{C}$ NMR	114
3- DEPT & HRMS	115
$4\text{a}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	116
4a- DEPT & HRMS	117
$4\text{b}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	118
4a DEPT & HRMS	119
4a, 4a NOESY	120
$5\text{a}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	121
5a- DEPT & HRMS	122
$6^1\text{H}$ NMR & $^{13}\text{C}$ NMR	123
6- DEPT & HRMS	124
$7\text{a}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	125
7a- DEPT & HRMS	126
$5\text{b}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	127
5b- DEPT & HRMS	128
$7\text{b}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	129
7b- DEPT & HRMS	130
$8\text{a}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	131
8a- DEPT & HRMS	132
$8\text{b}^1\text{H}$ NMR & $^{13}\text{C}$ NMR	133
8b- DEPT & HRMS	134
$9\text{a}^{31}\text{P}$ NMR & $^{13}\text{C}$ NMR	135
$9\text{b}^{31}\text{P}$ NMR & $^{13}\text{C}$ NMR	136
HPLC & MALDI-TOF of DNA 1	137
HPLC & MALDI-TOF of DNA $15\text{T}^{\text{trans}}$	138
HPLC & MALDI-TOF of DNA $9\text{T}^{\text{trans}}$	139
HPLC & MALDI-TOF of DNA $15\text{T}^{\text{cis}}$	140
HPLC & MALDI-TOF of DNA $9\text{T}^{\text{cis}}$	141

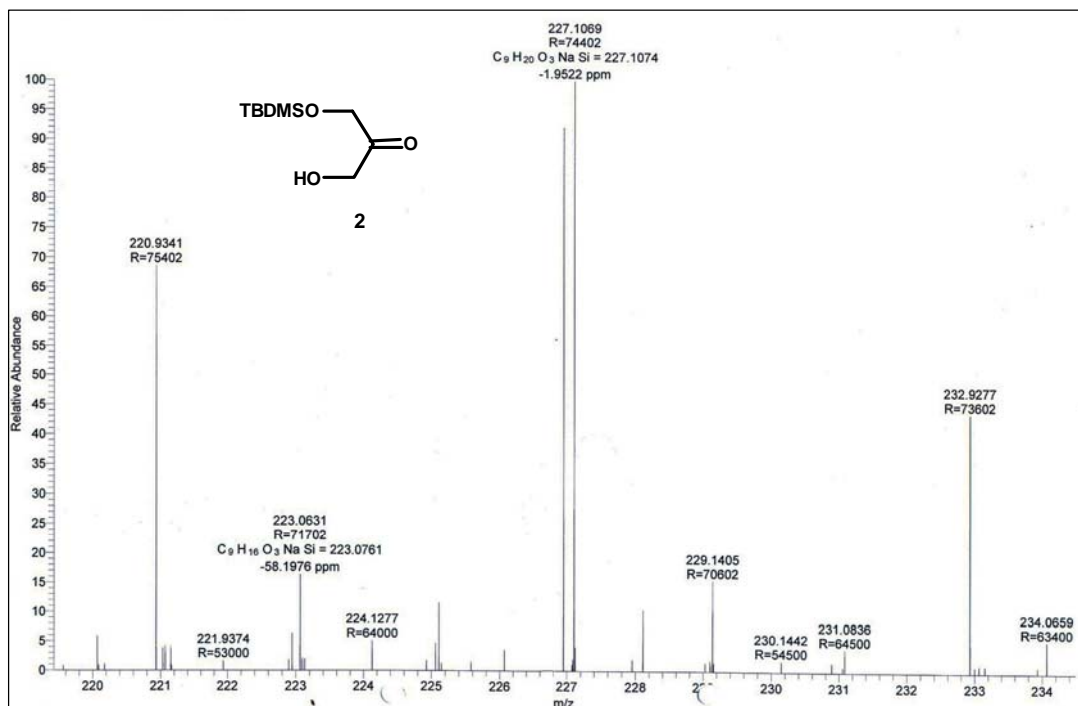
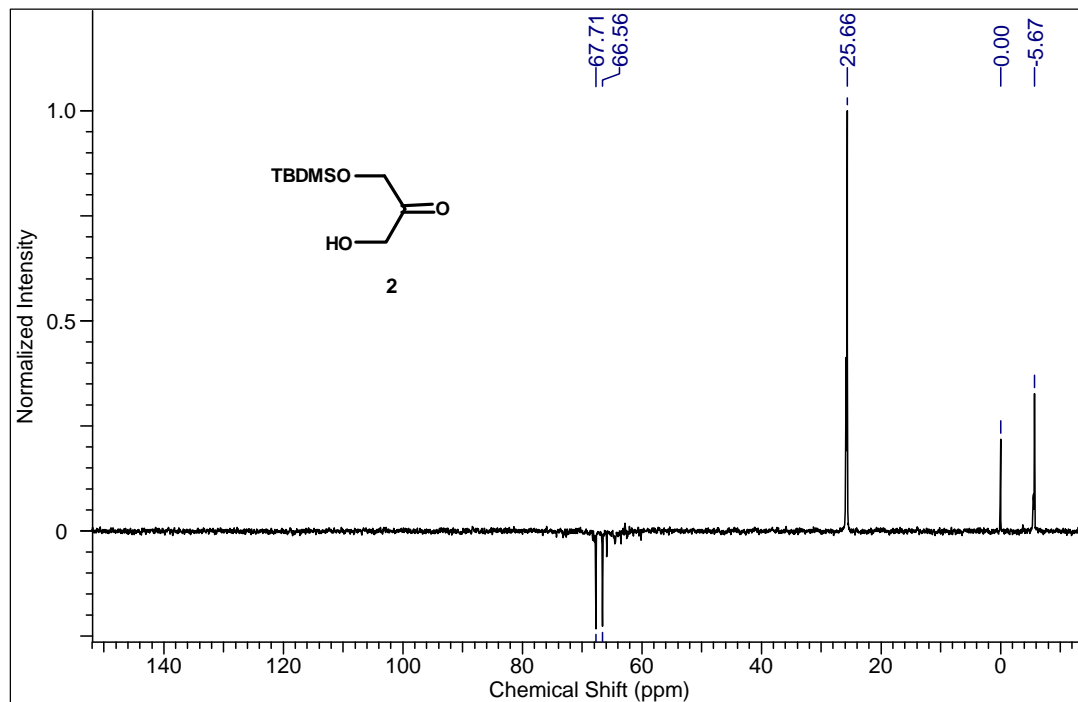
### *Chapter 3*

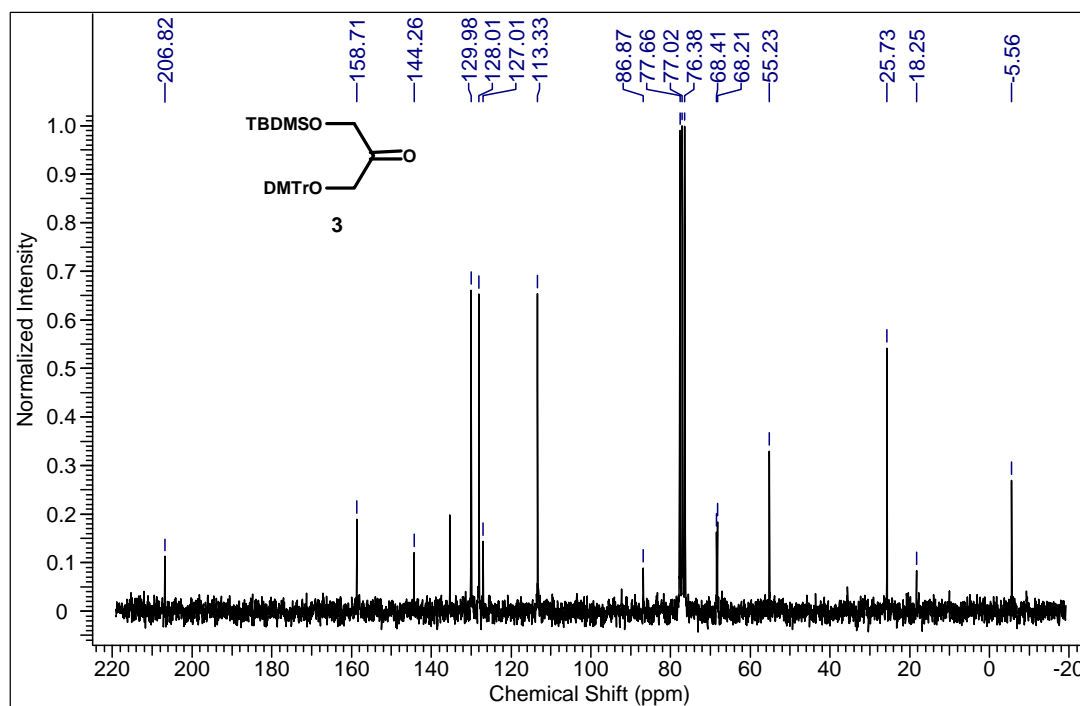
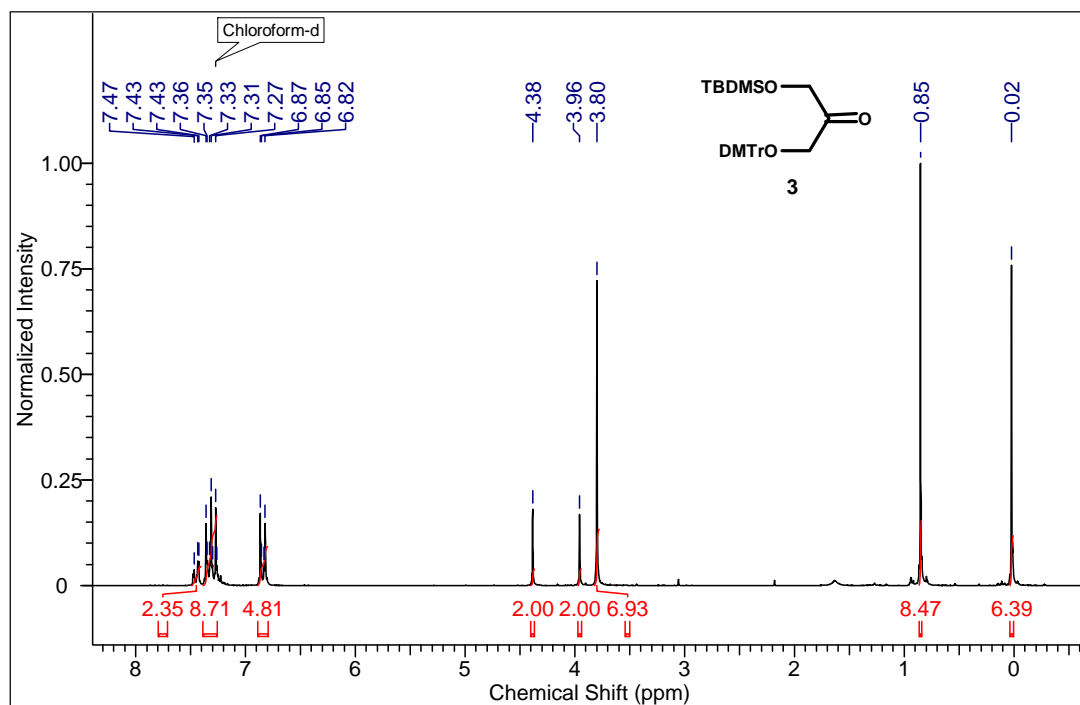
---

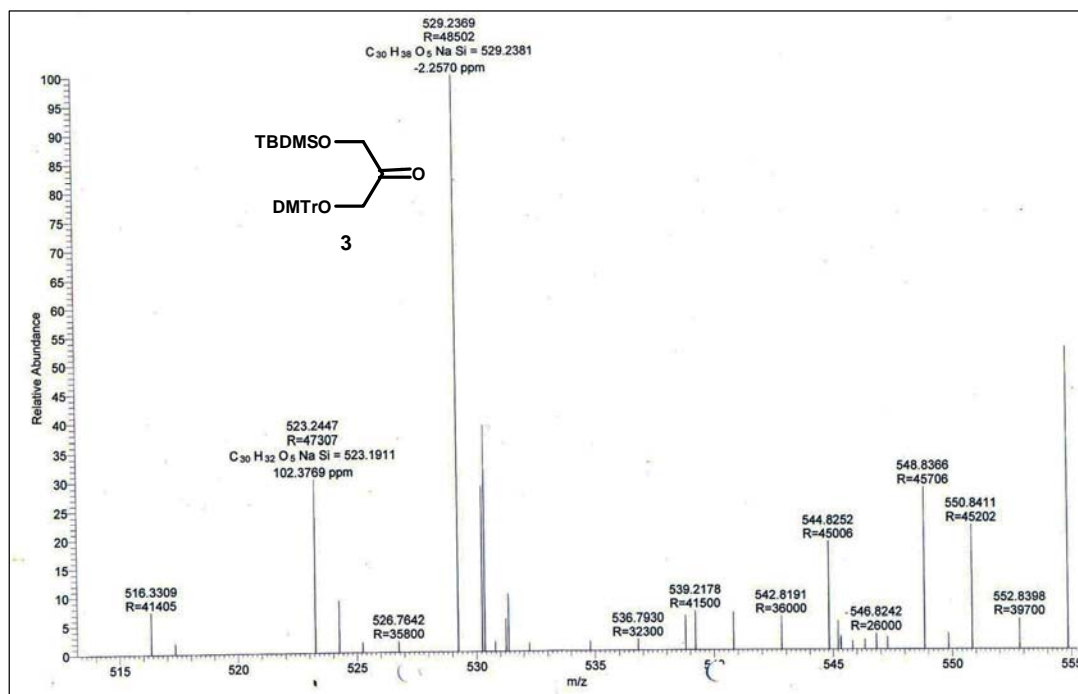
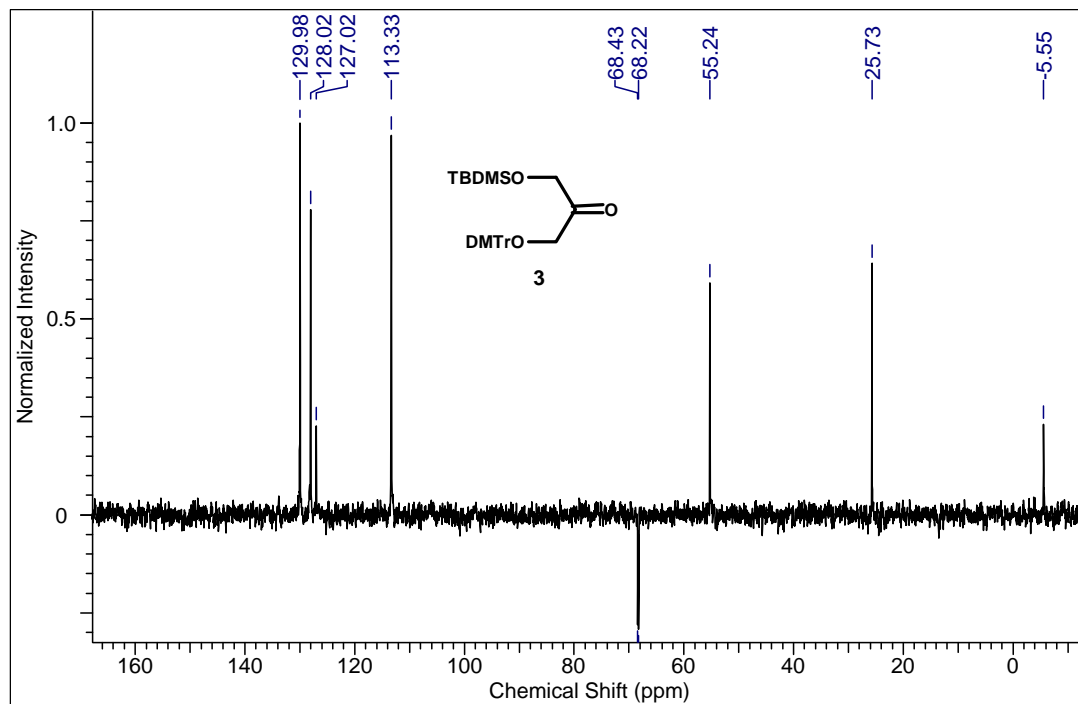
<b>HPLC &amp; MALDI-TOF of DNA 2</b>	142
<b>HPLC &amp; MALDI-TOF of DNA 3</b>	143
<b>HPLC &amp; MALDI-TOF of DNA 4</b>	144

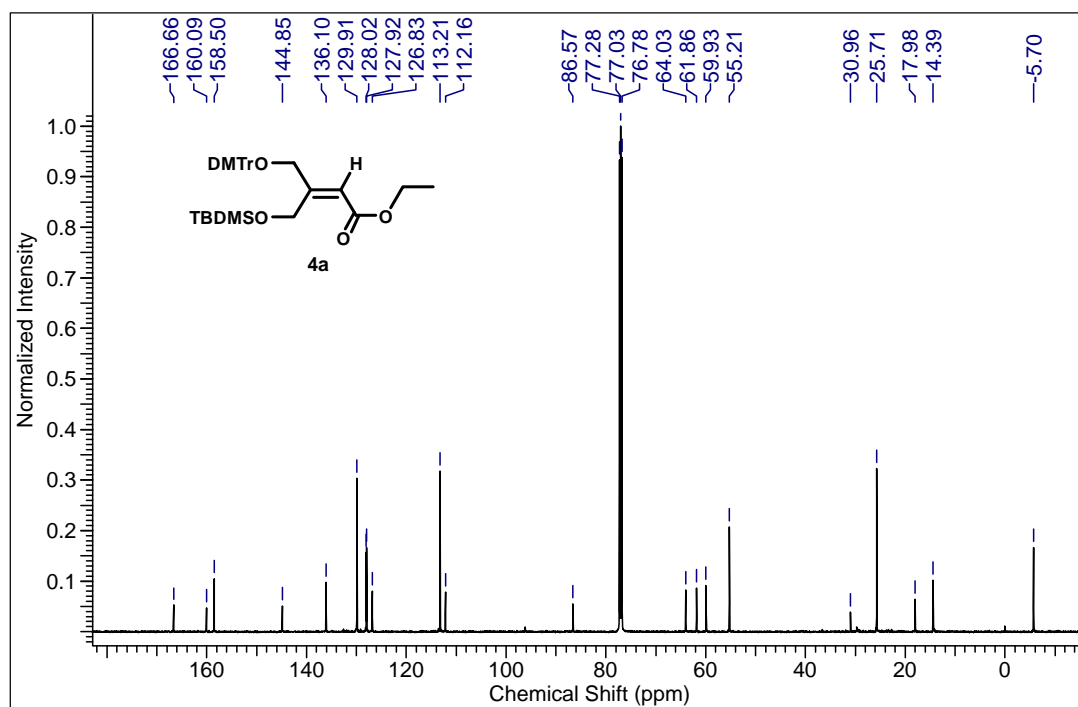
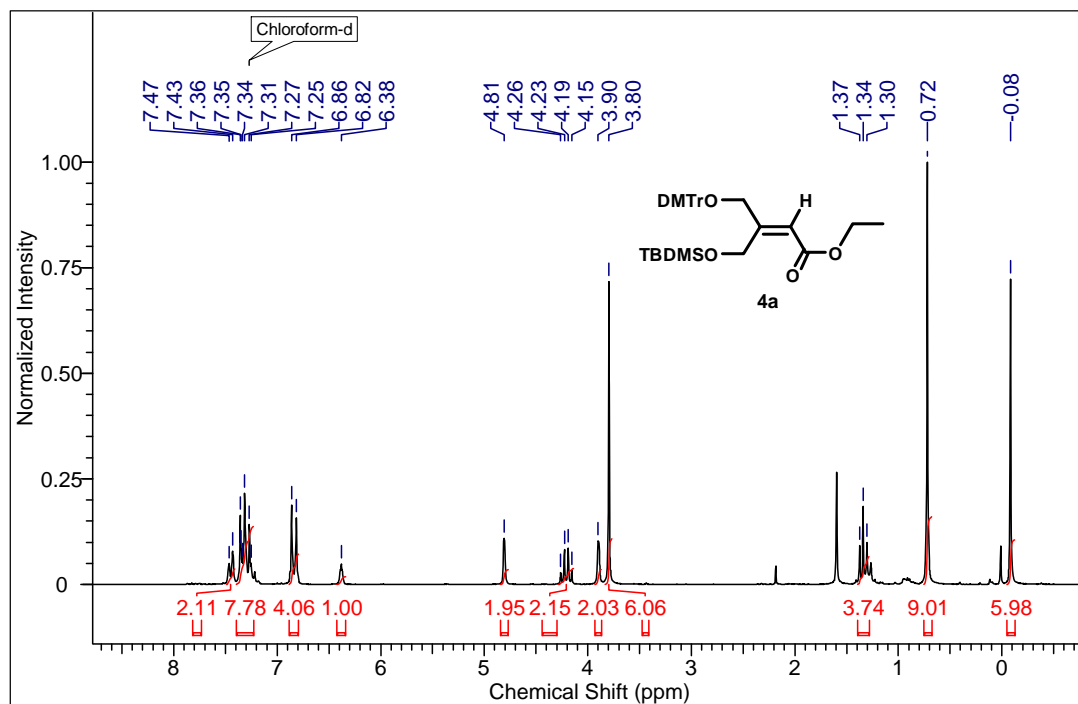
## NMR spectral data:

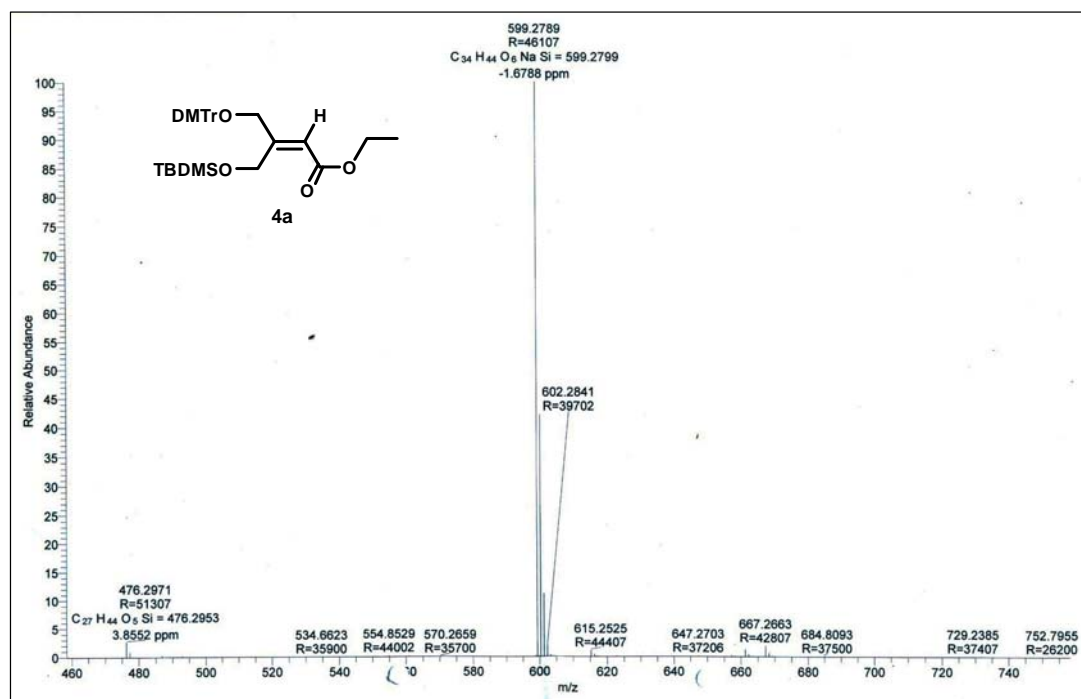
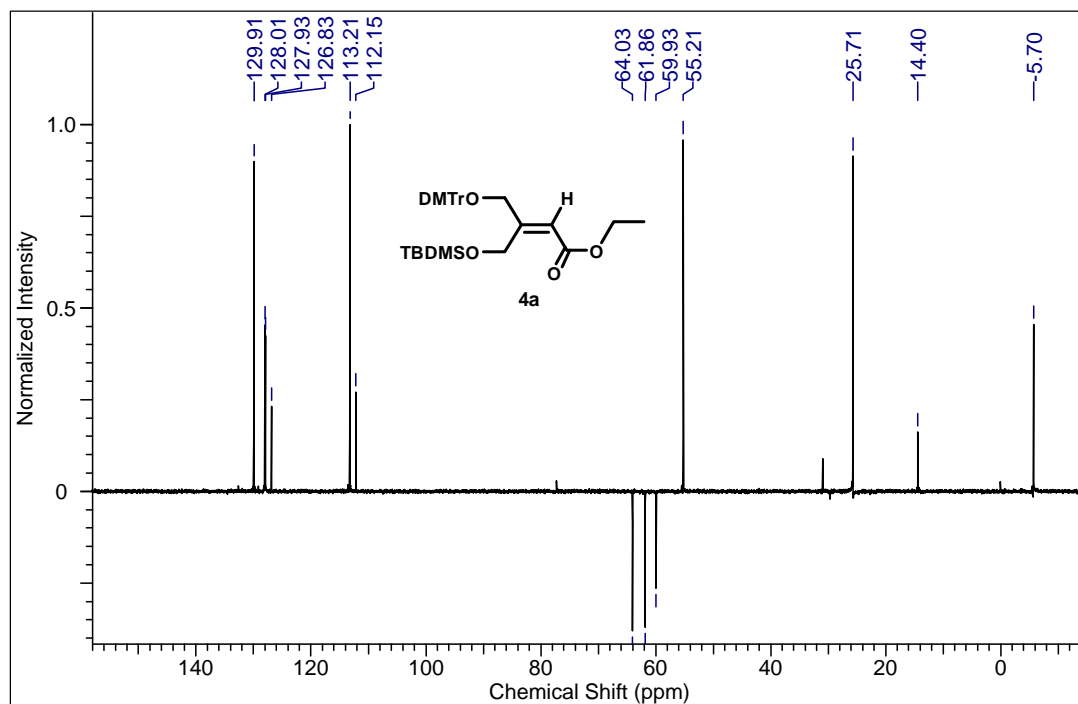




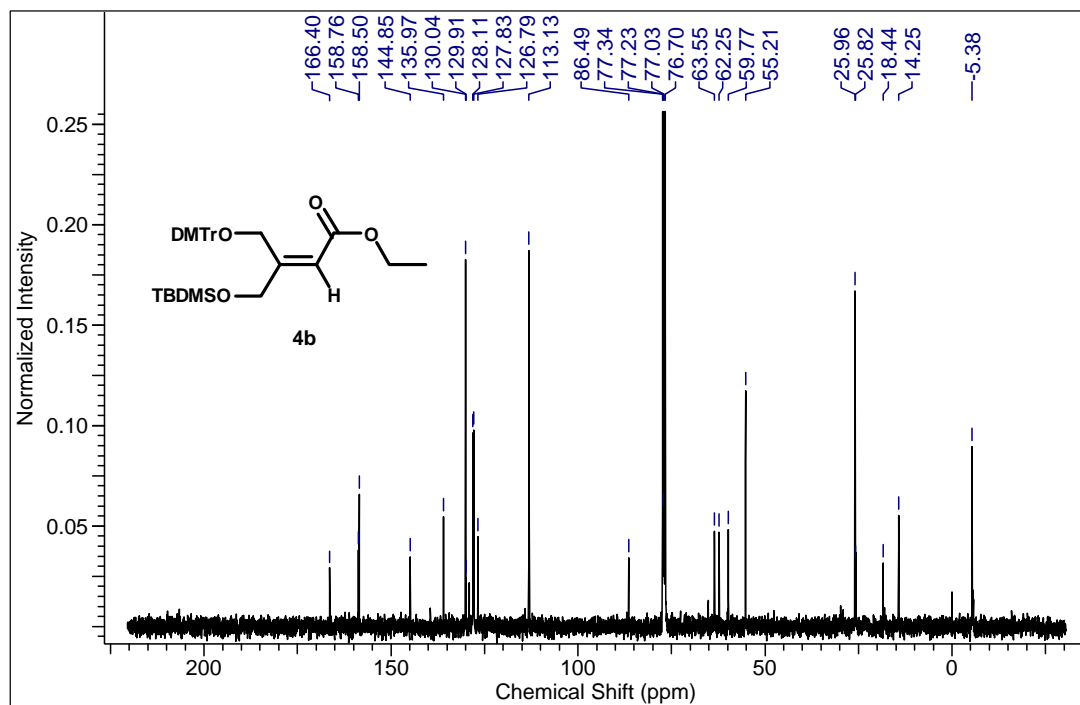
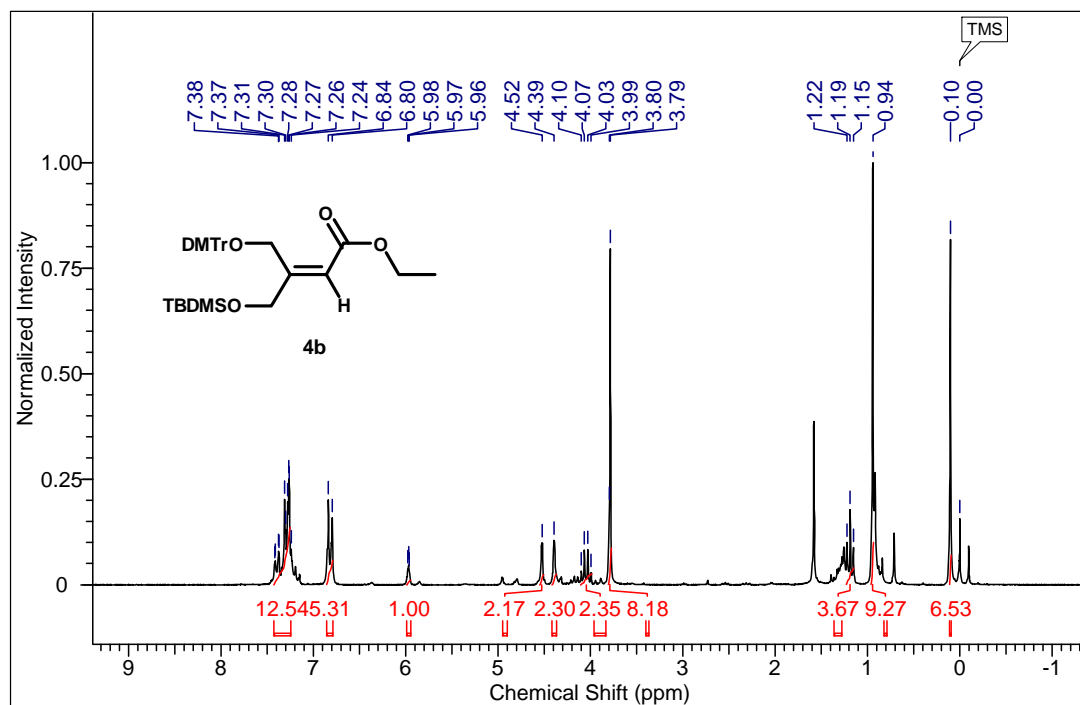


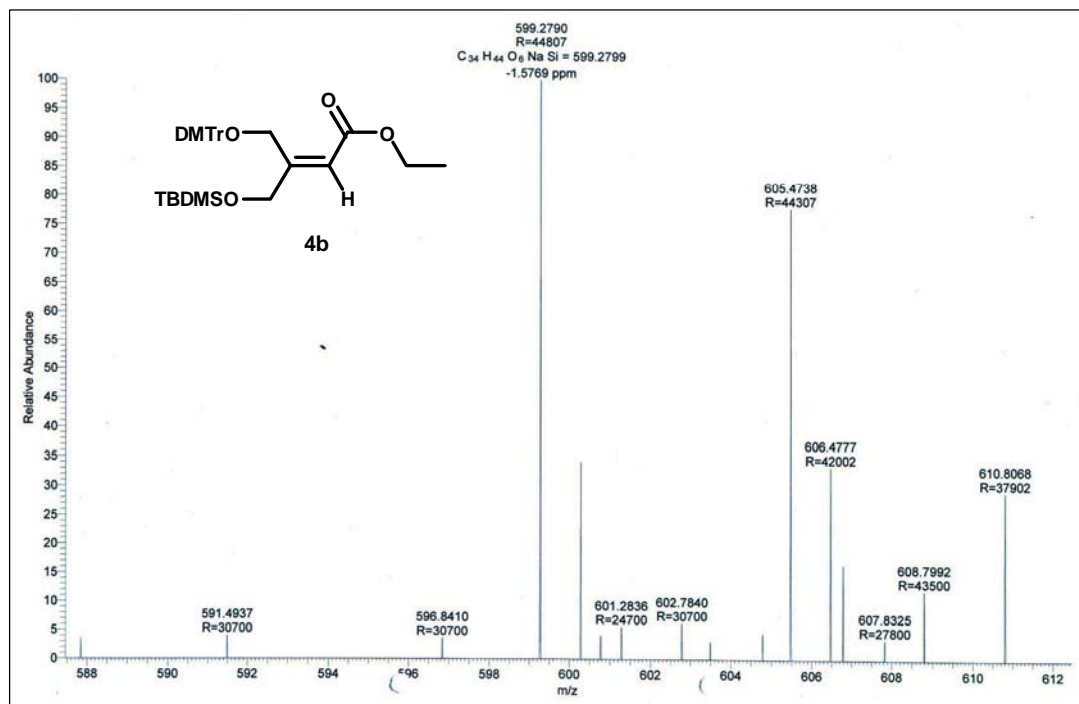
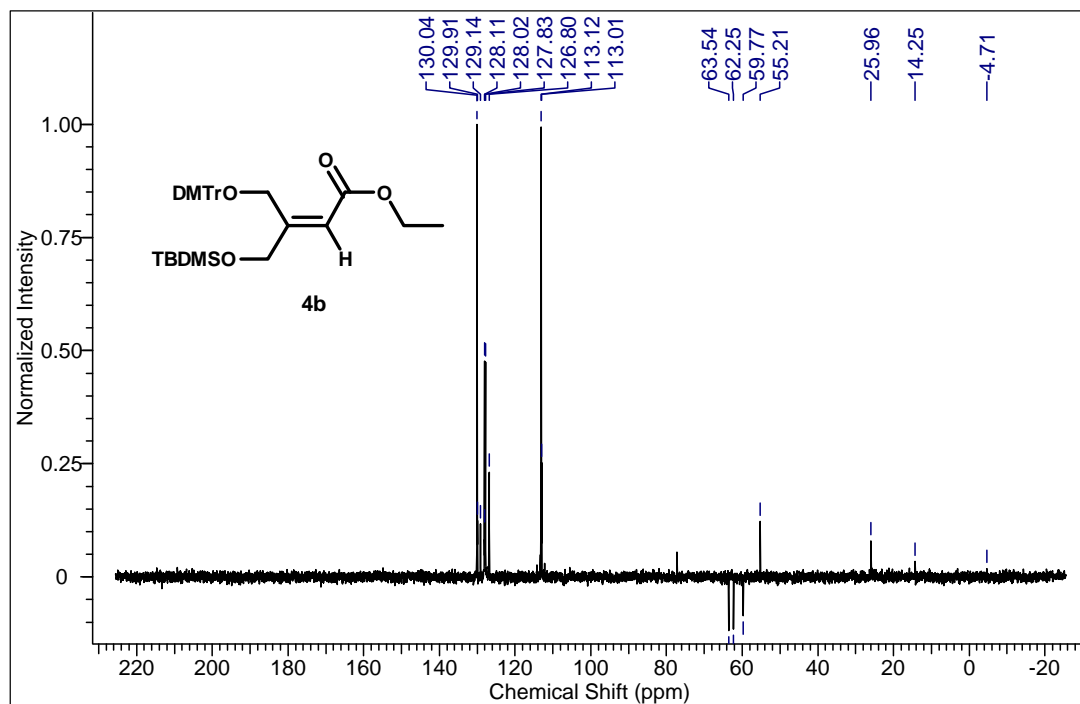


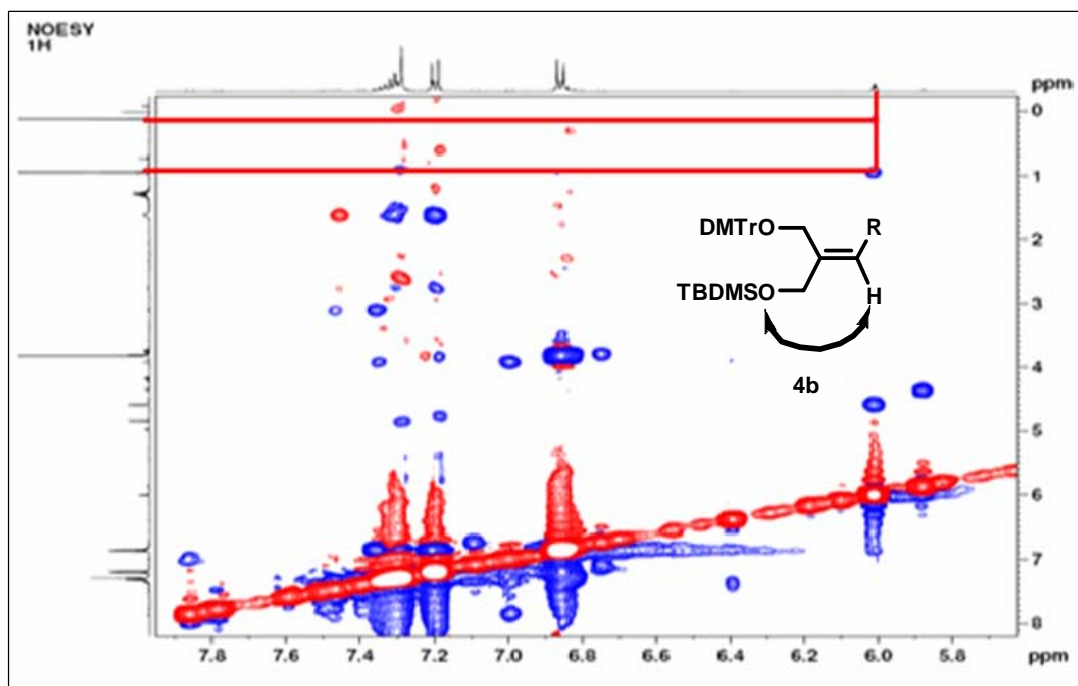
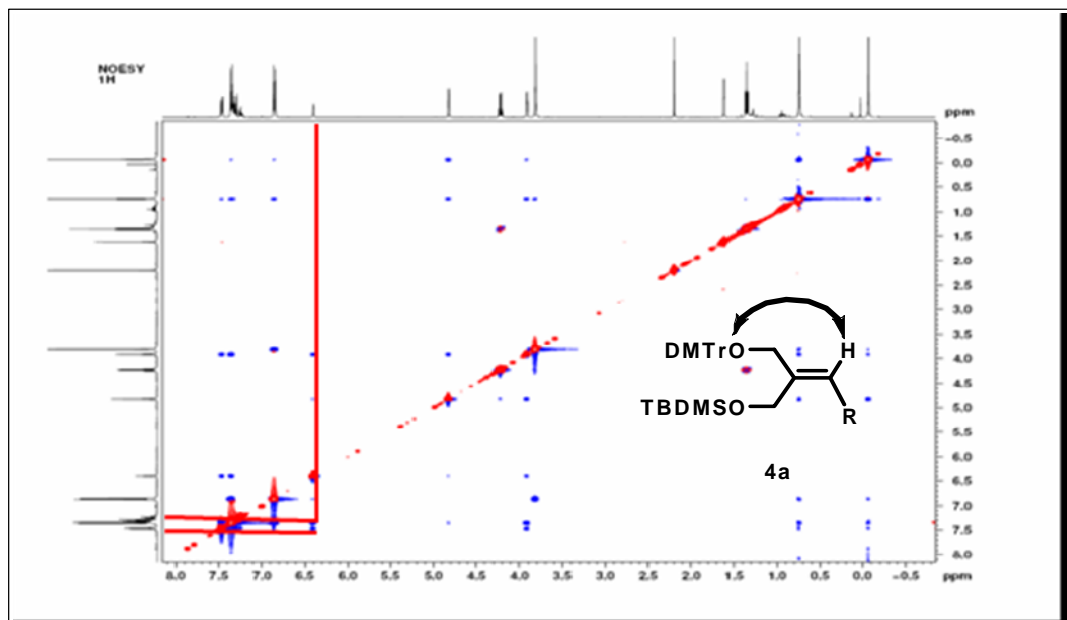


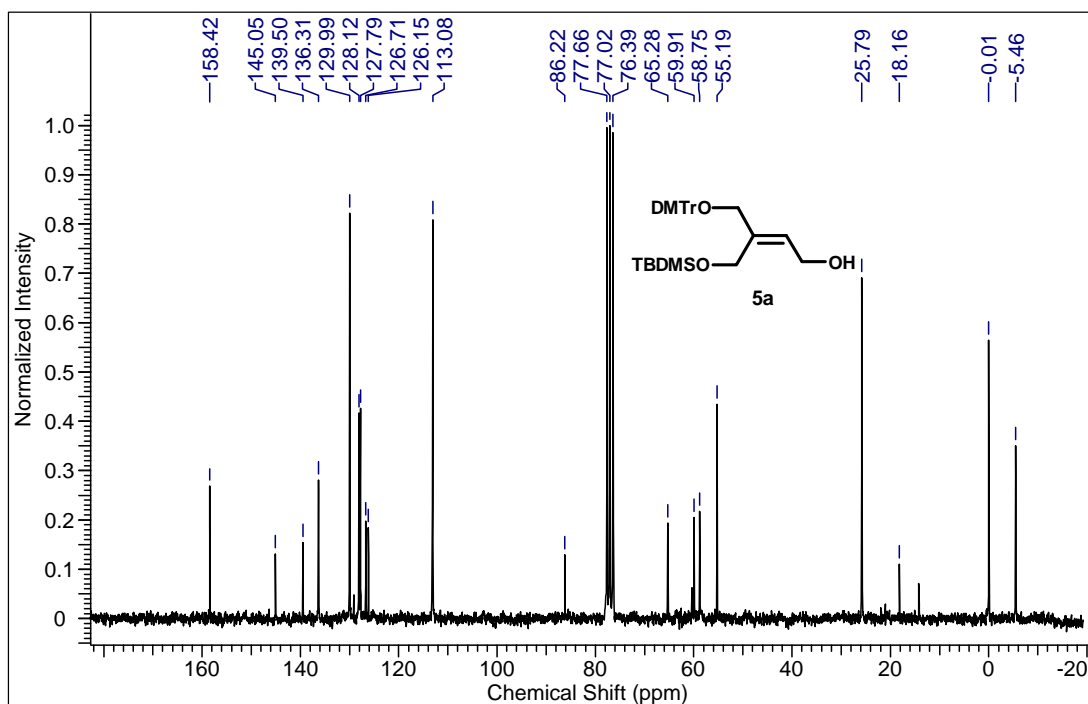
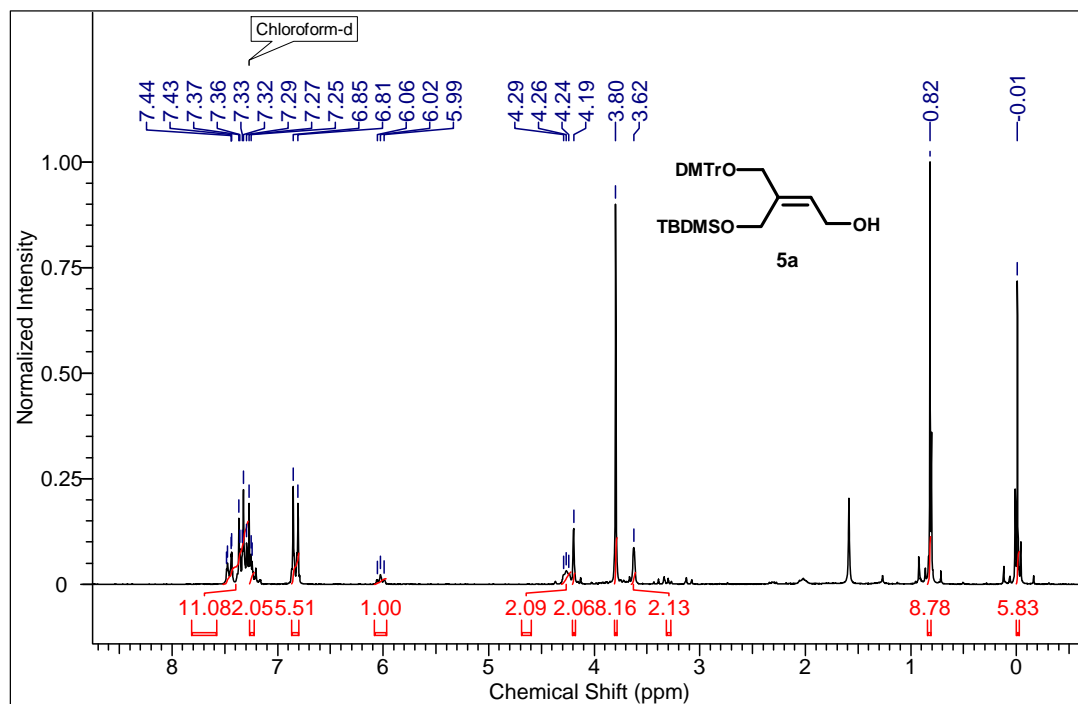


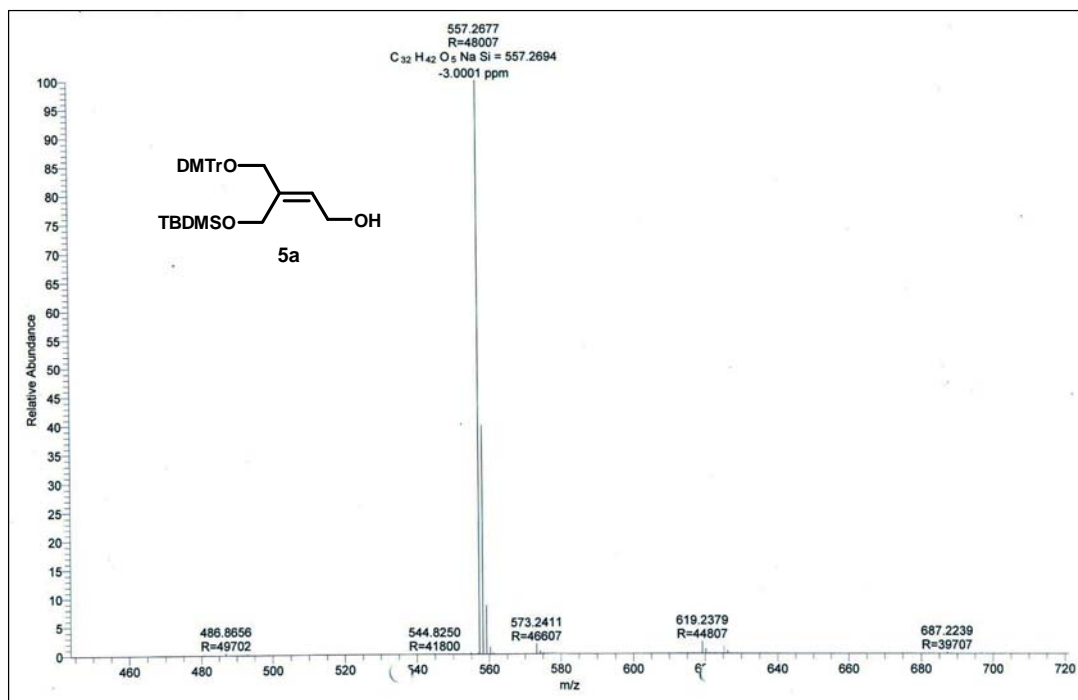
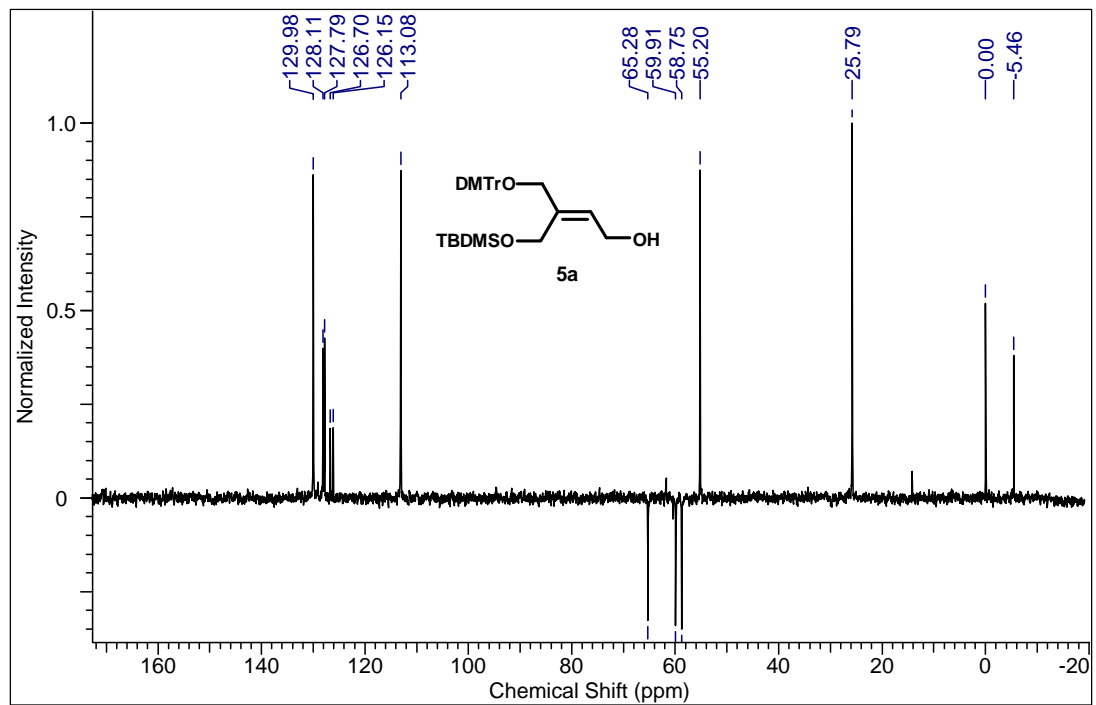


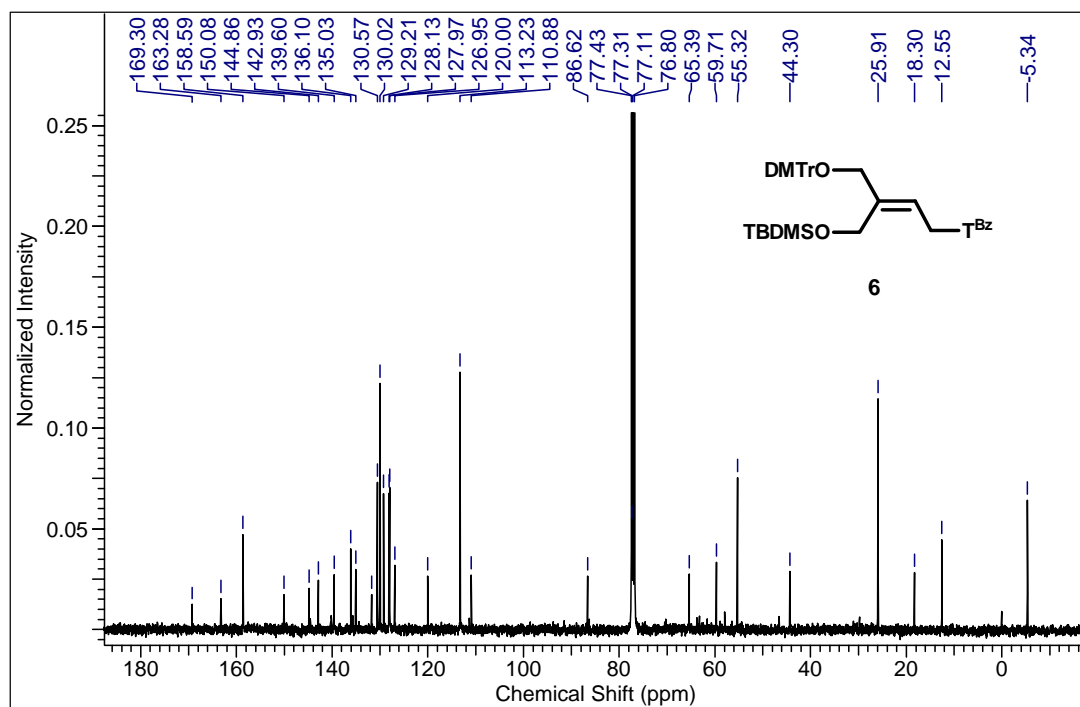
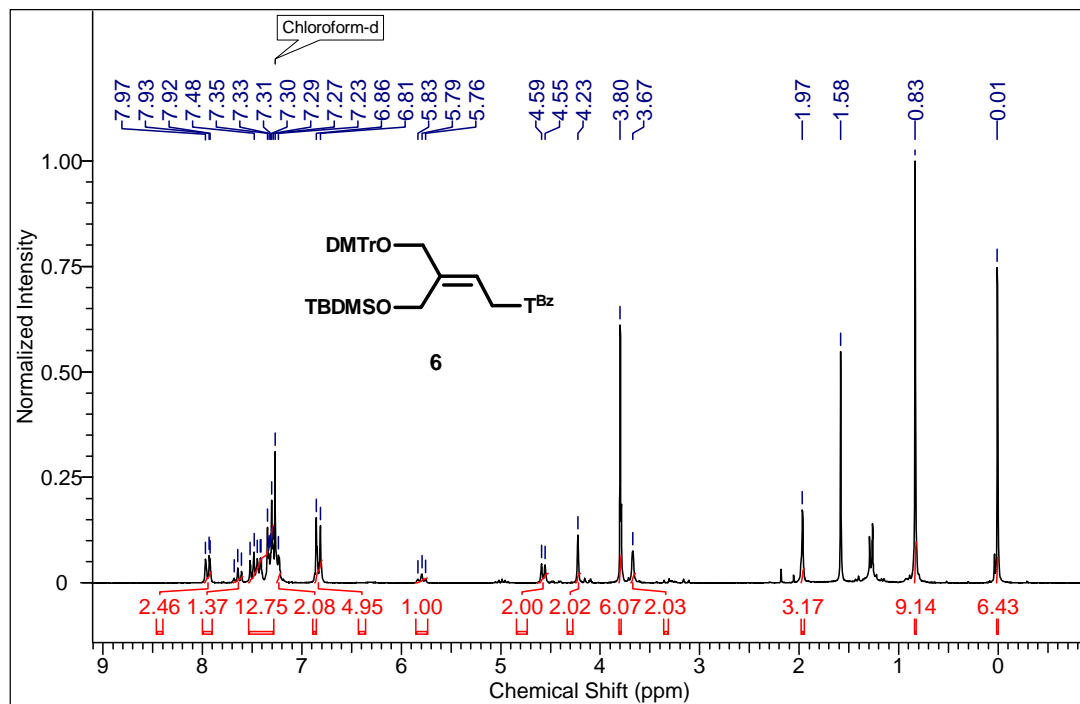


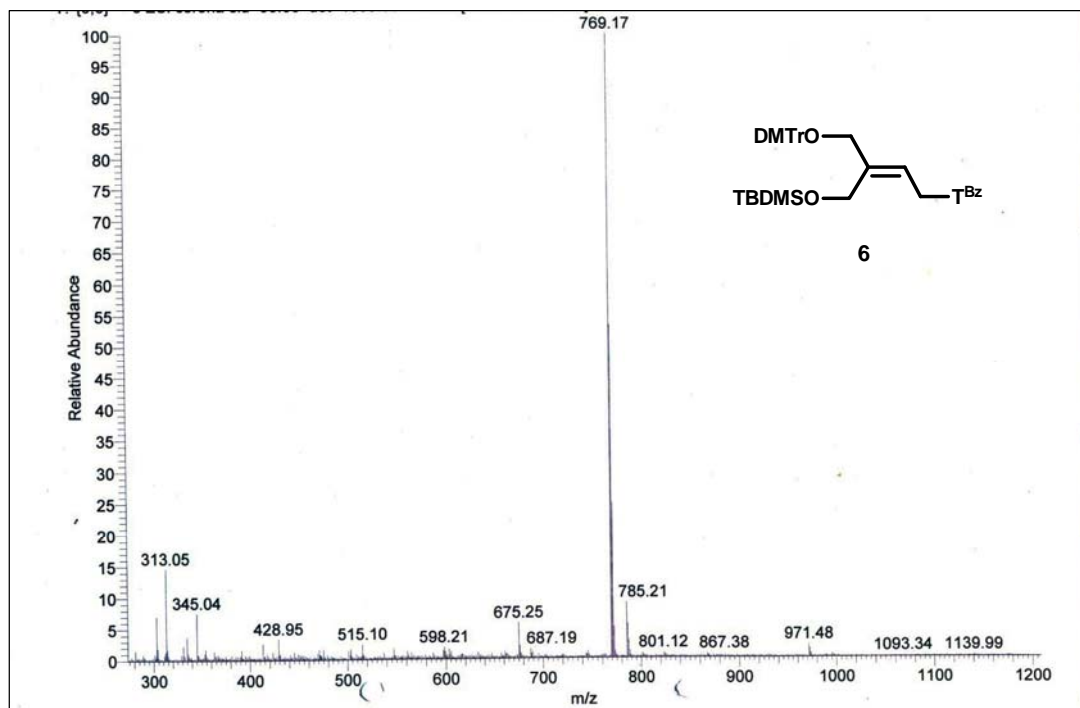
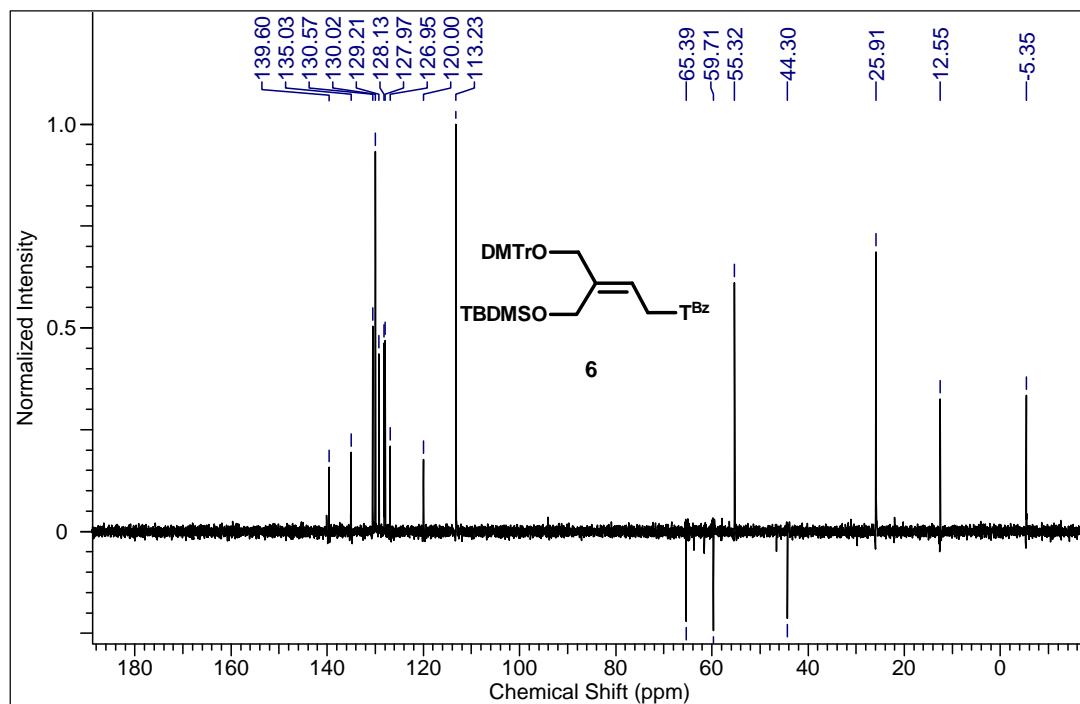


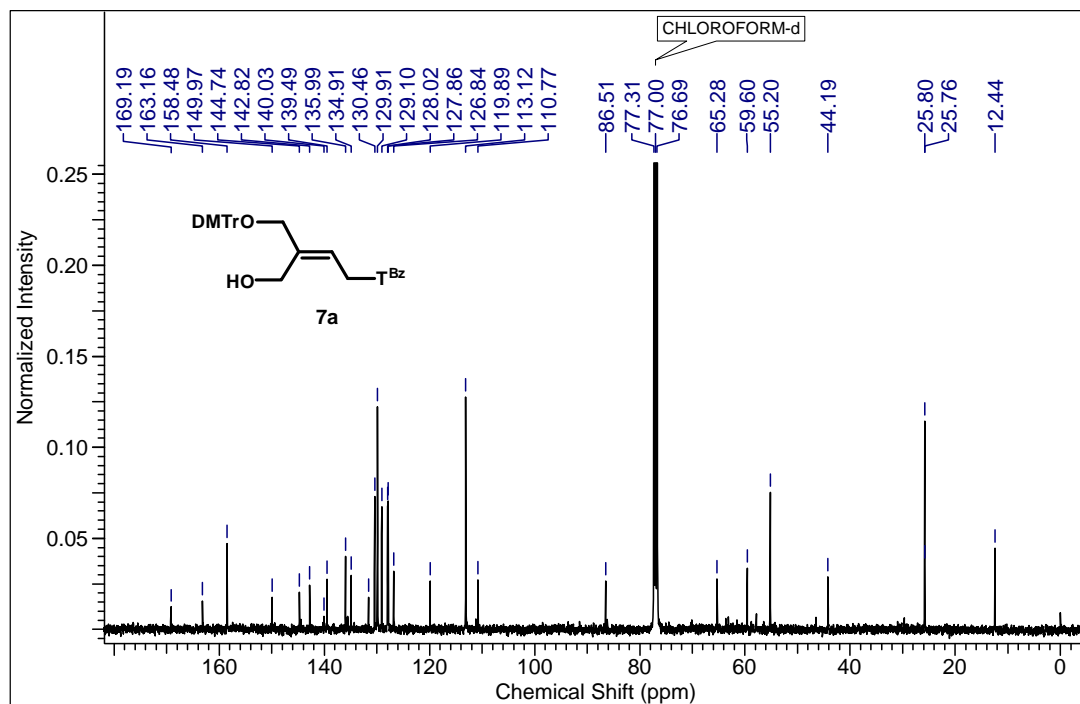
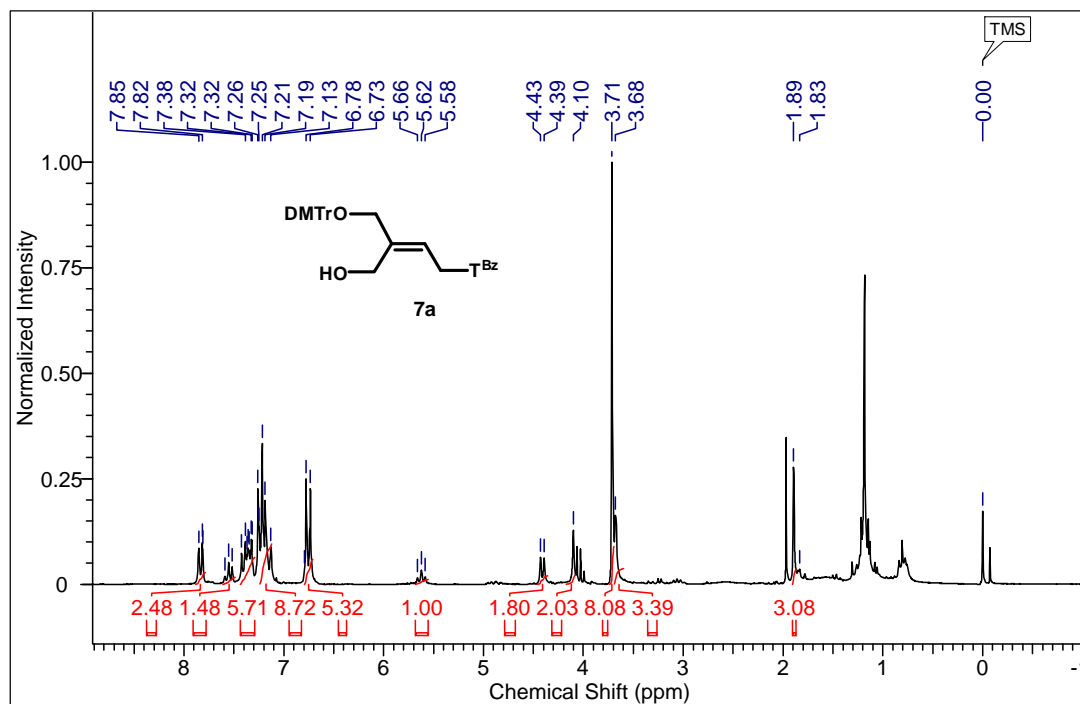




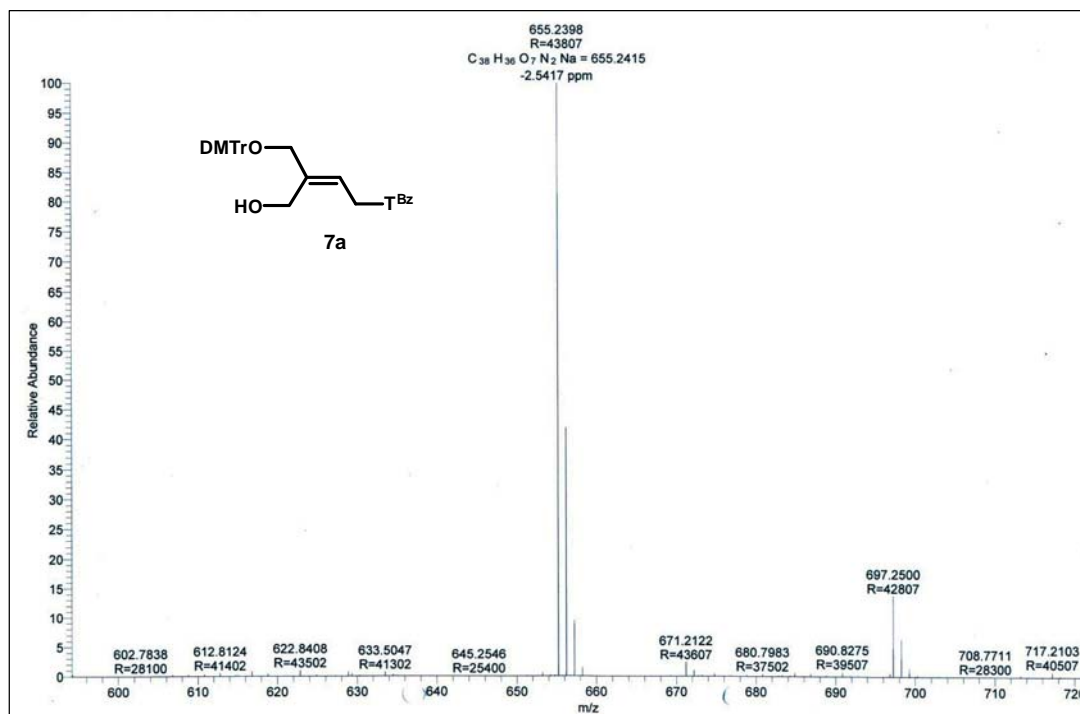
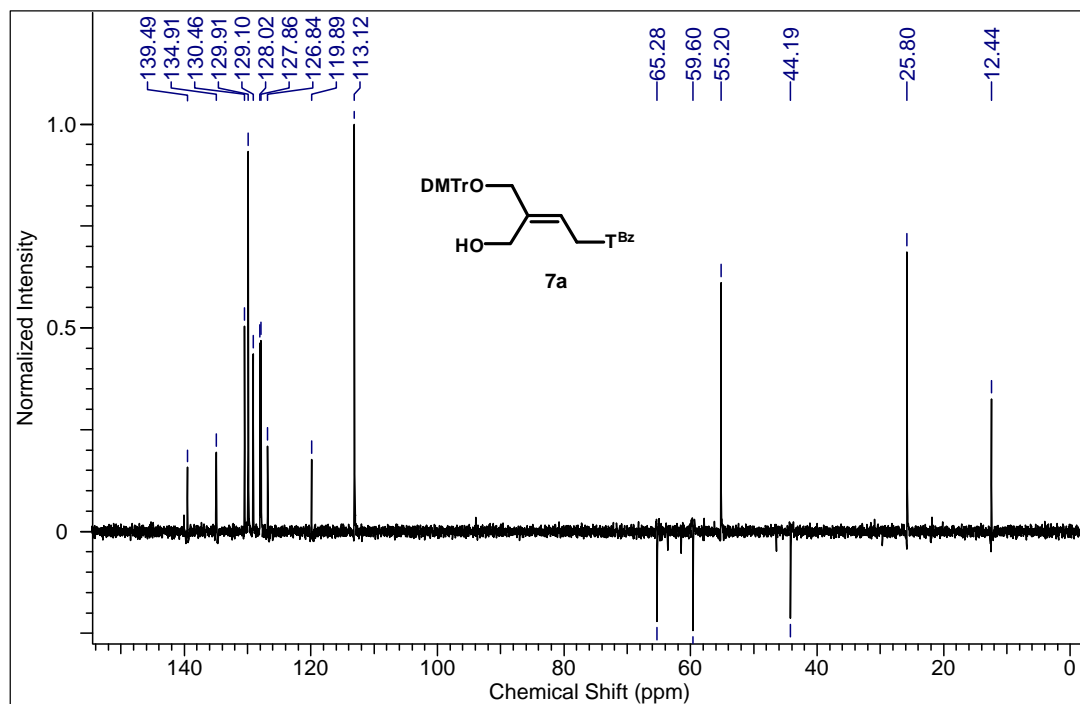


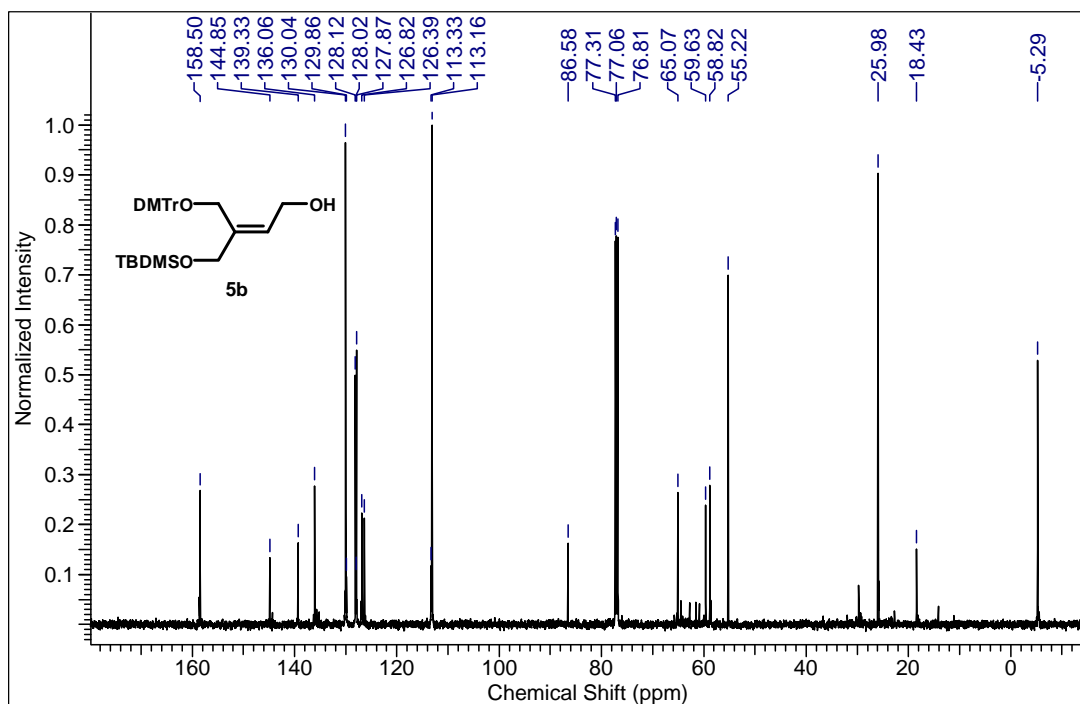
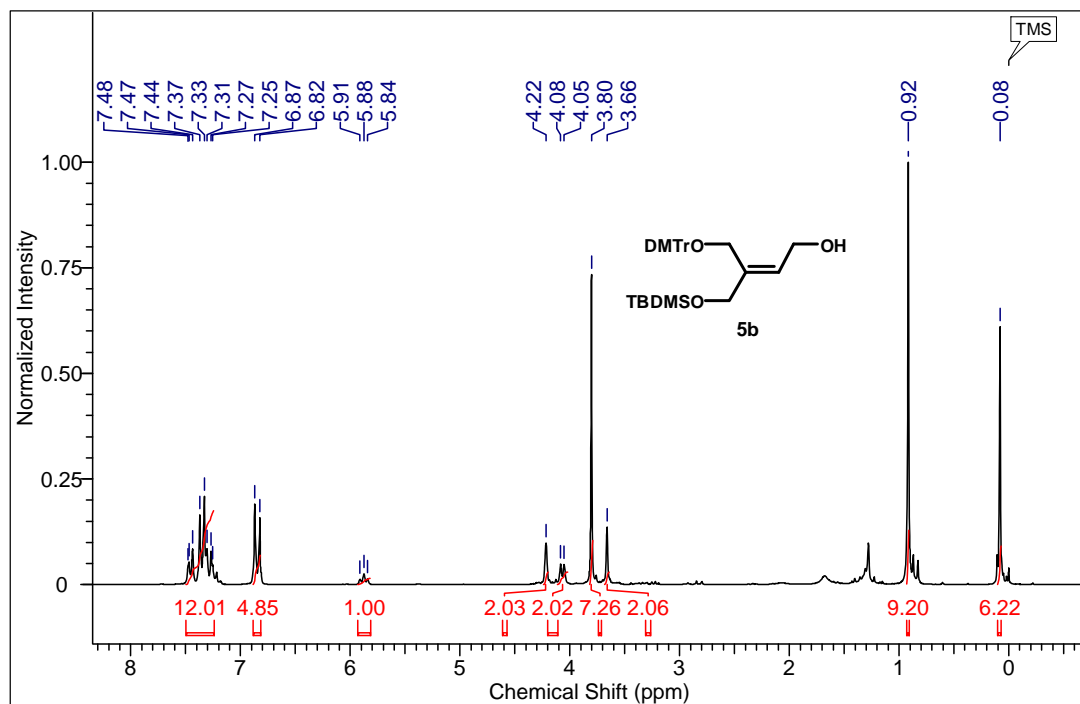




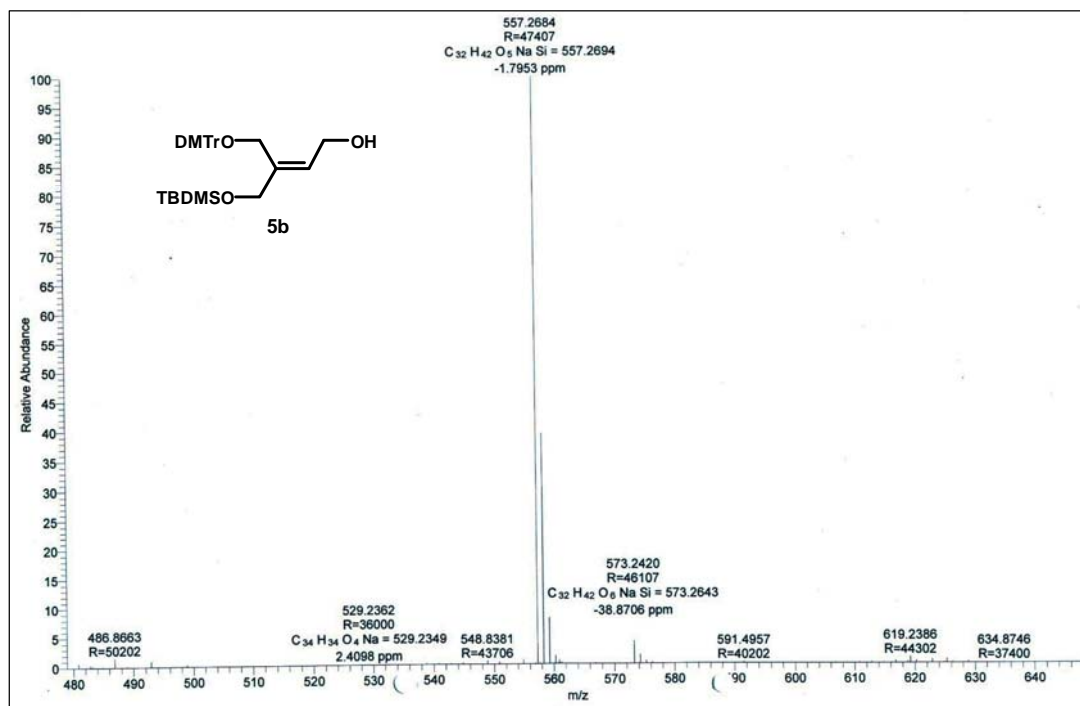
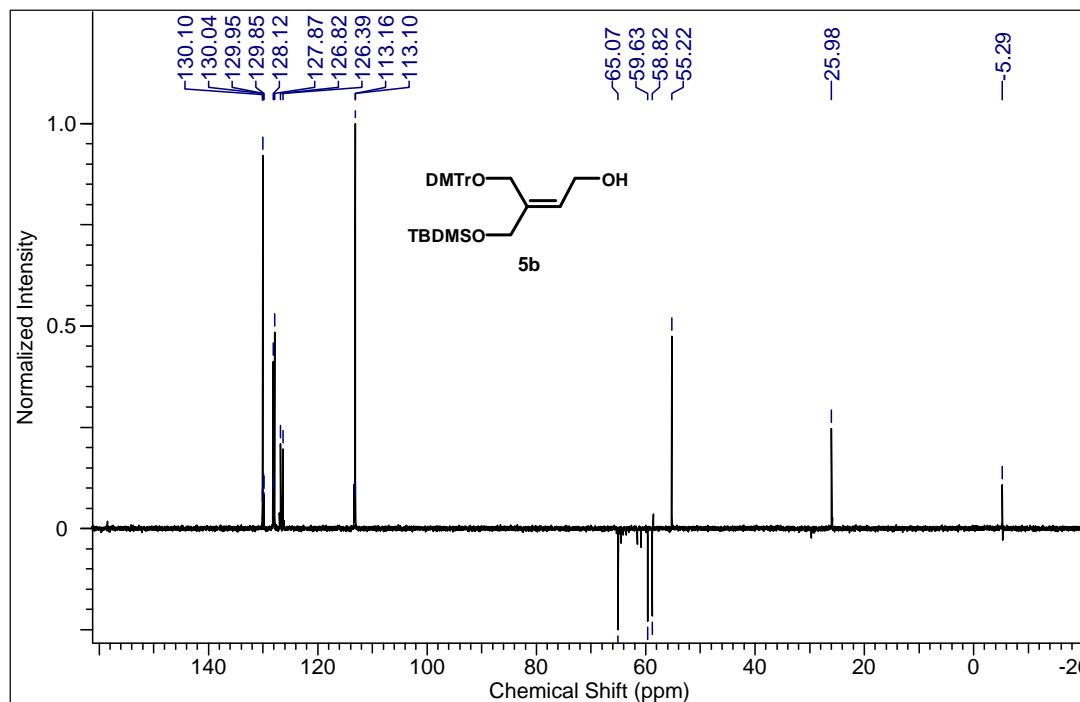


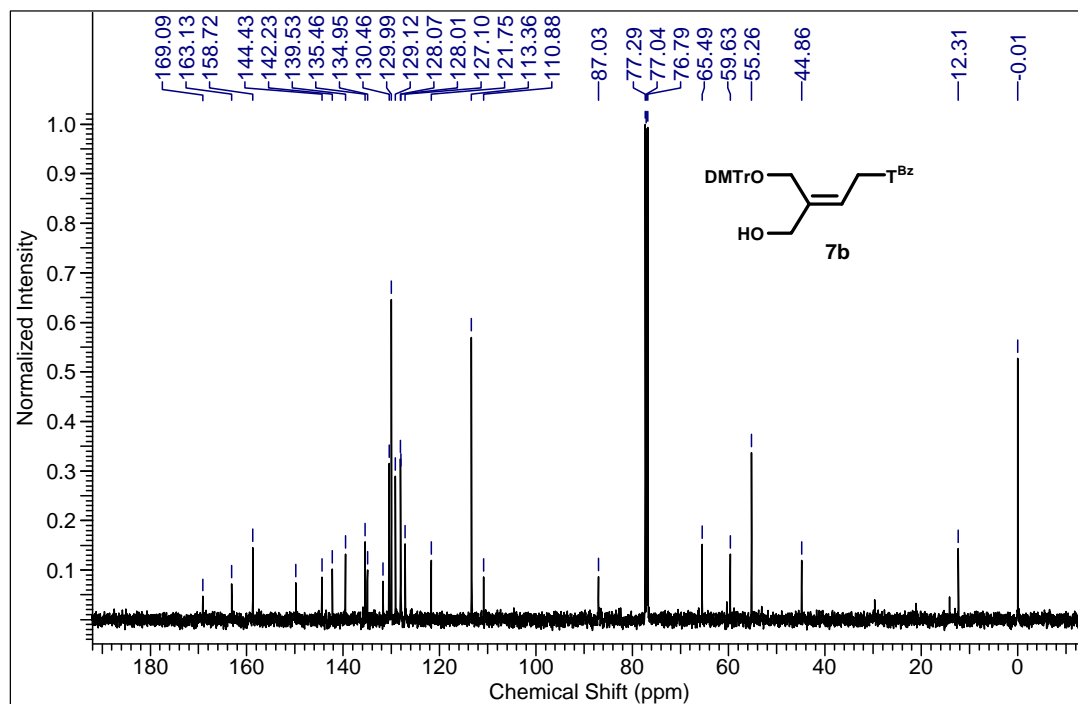
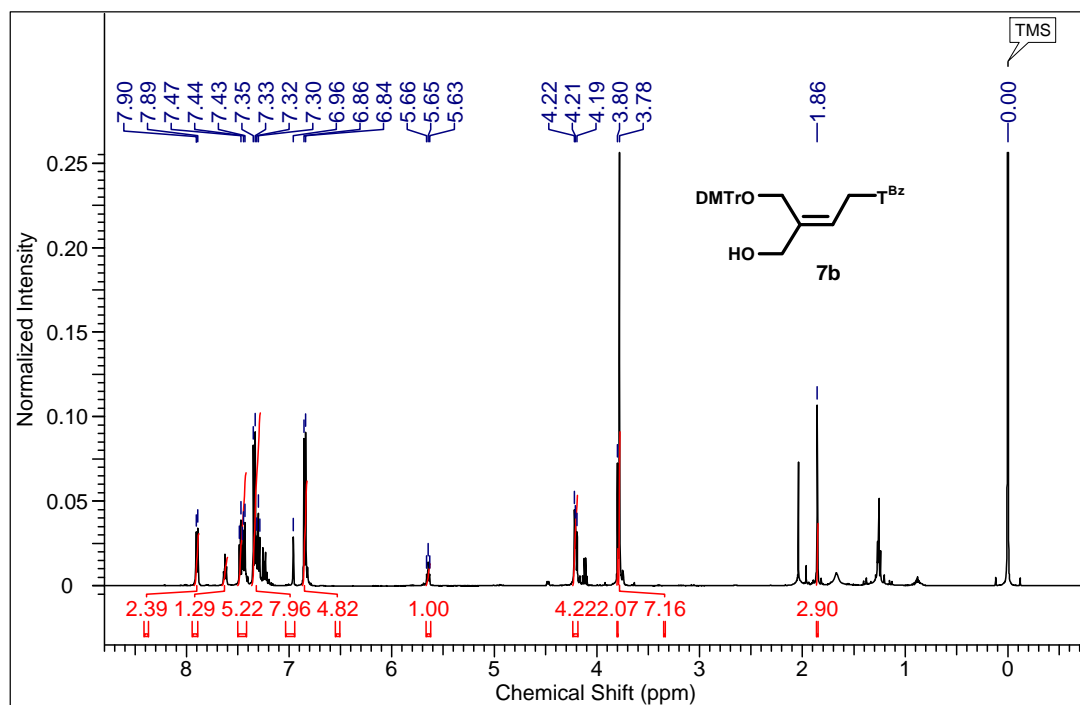


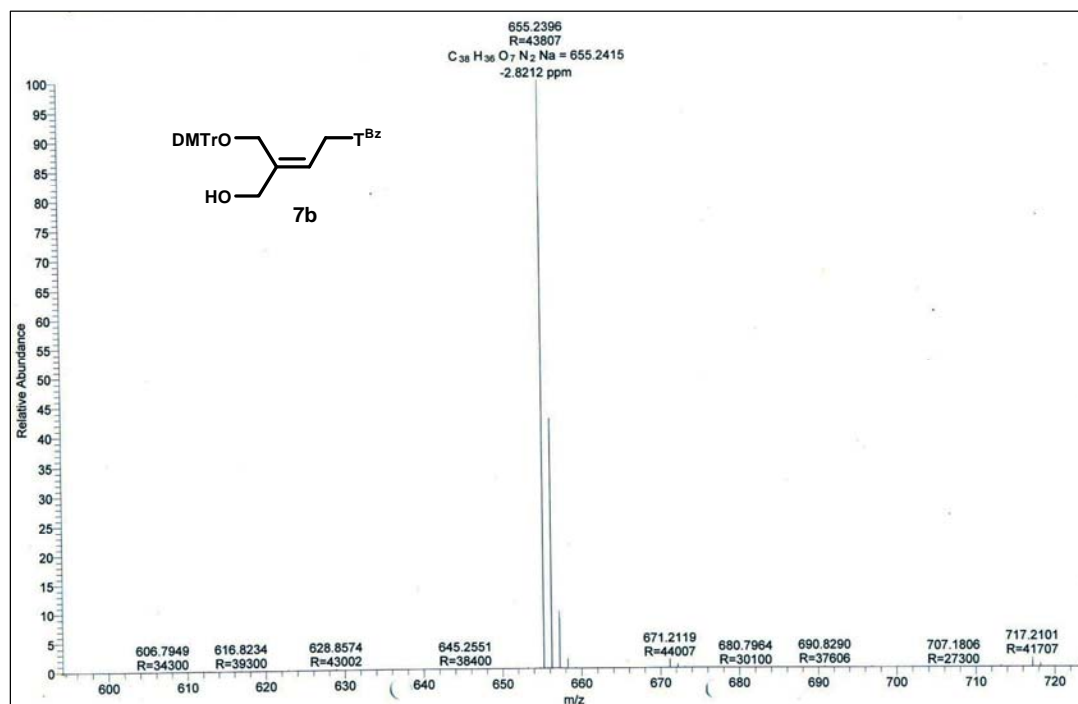
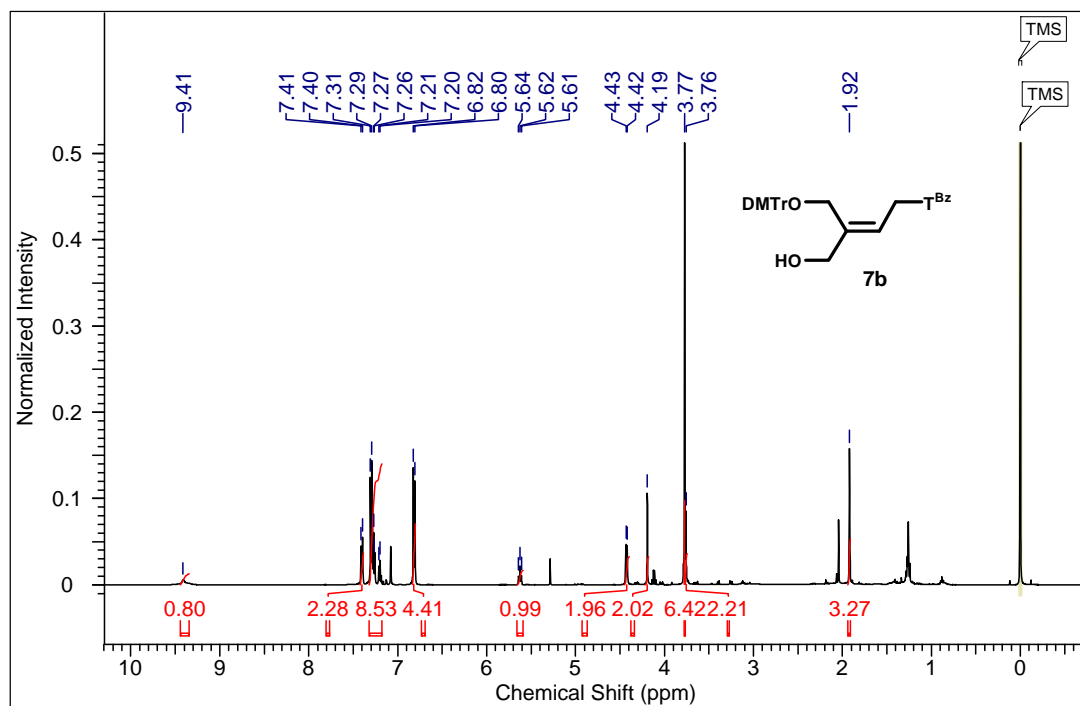


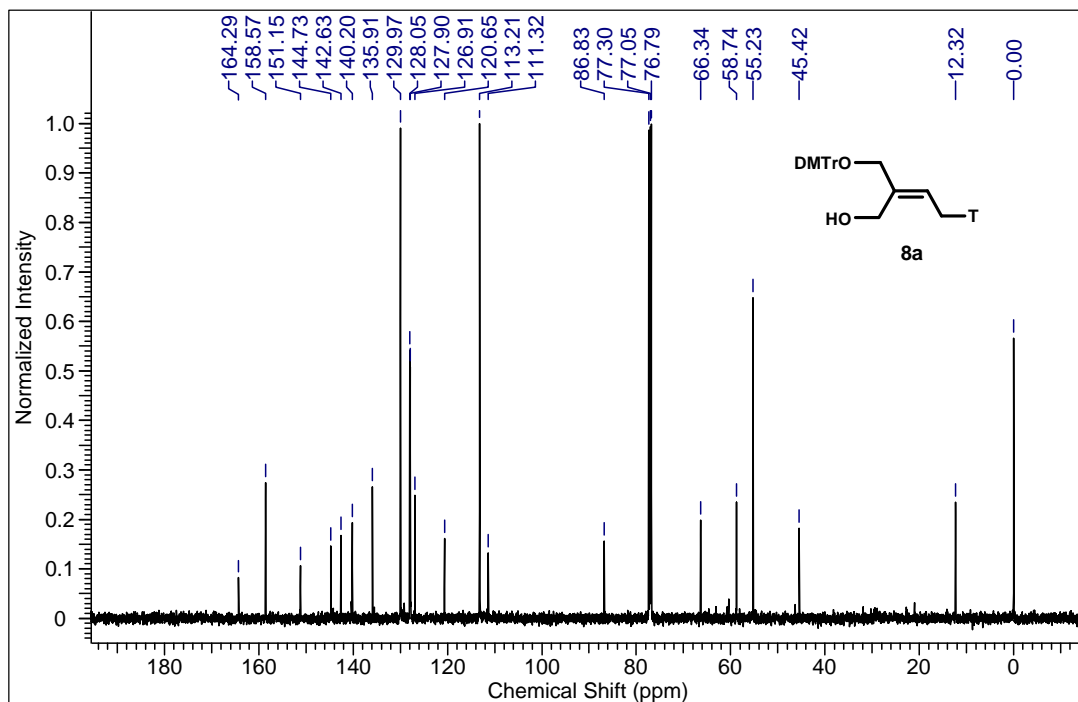
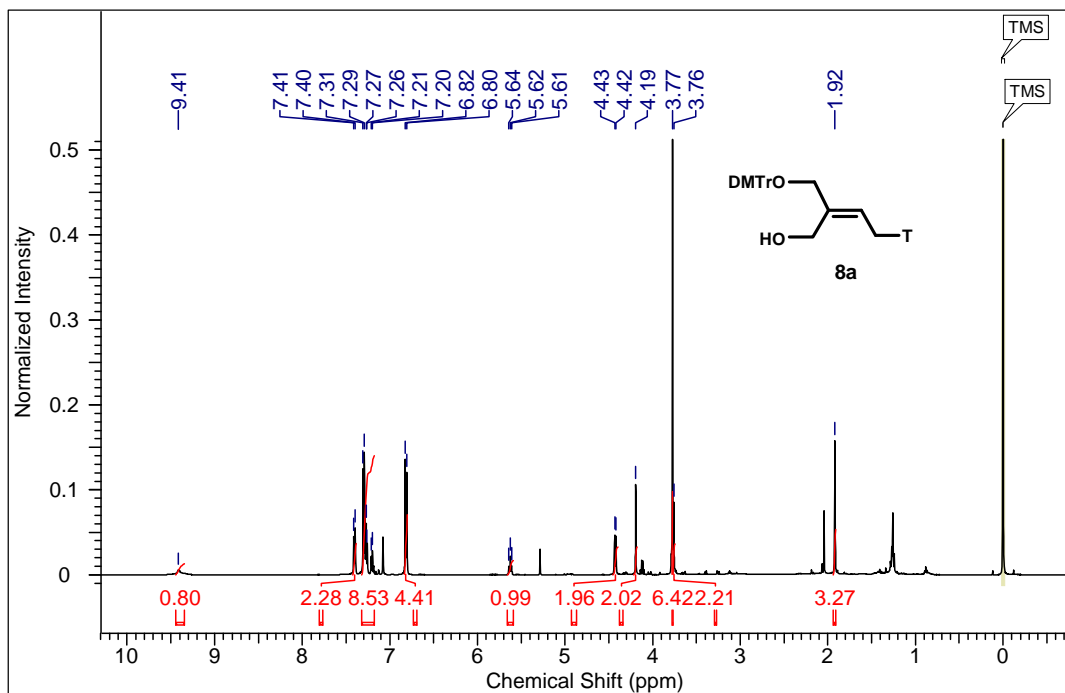


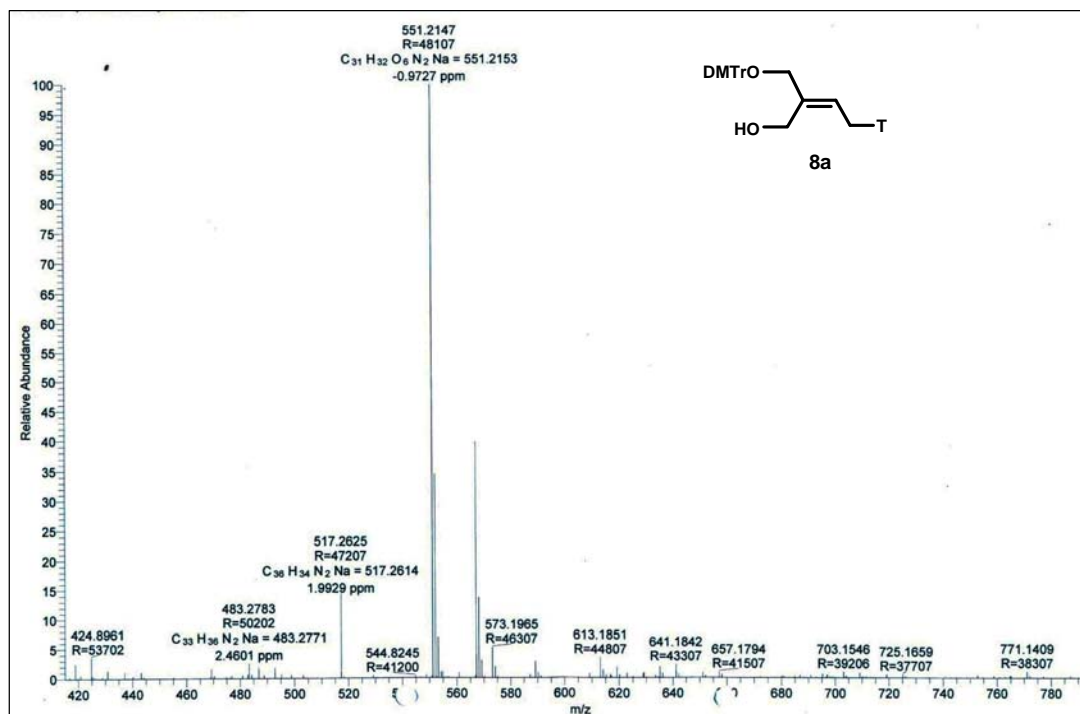
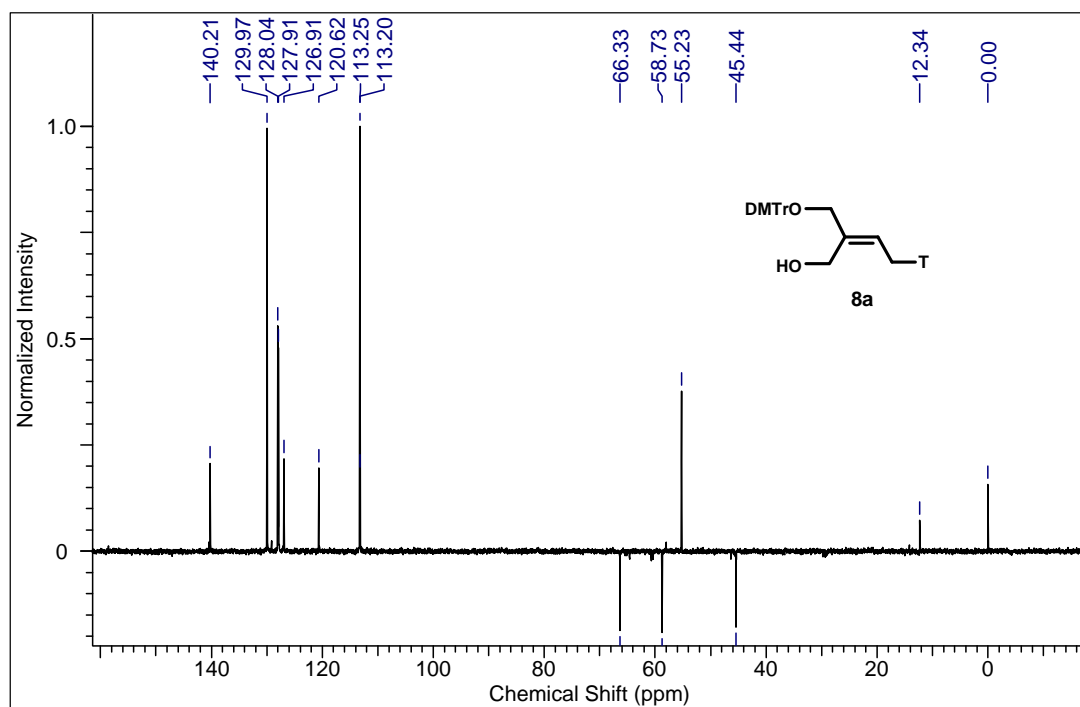
### Chapter 3

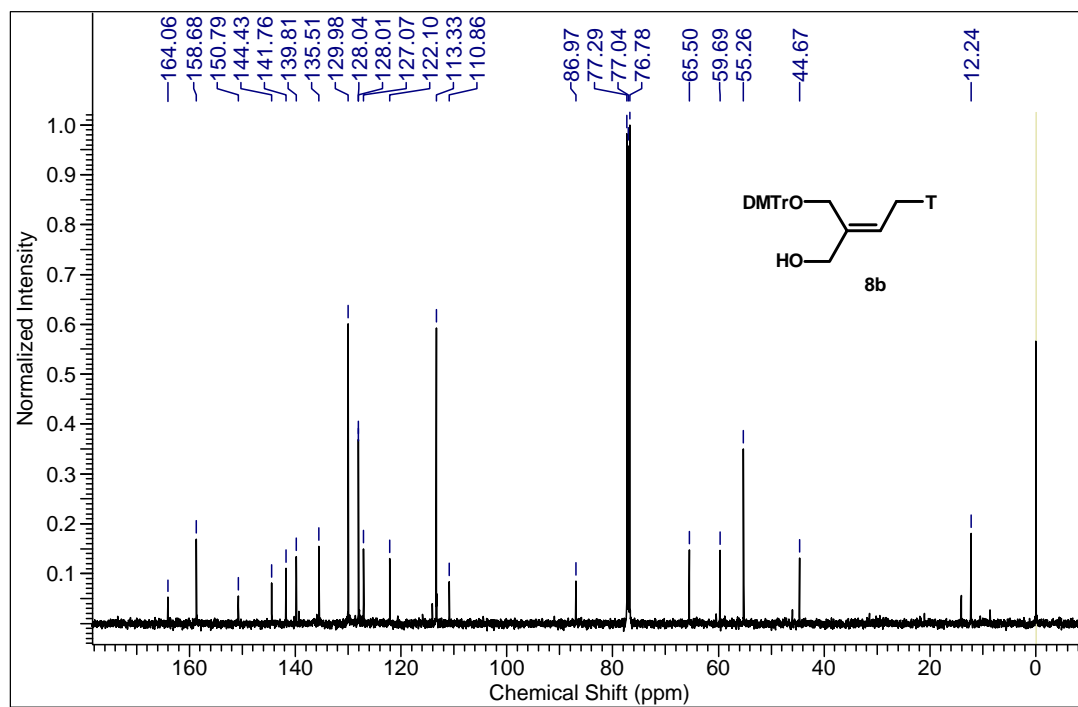
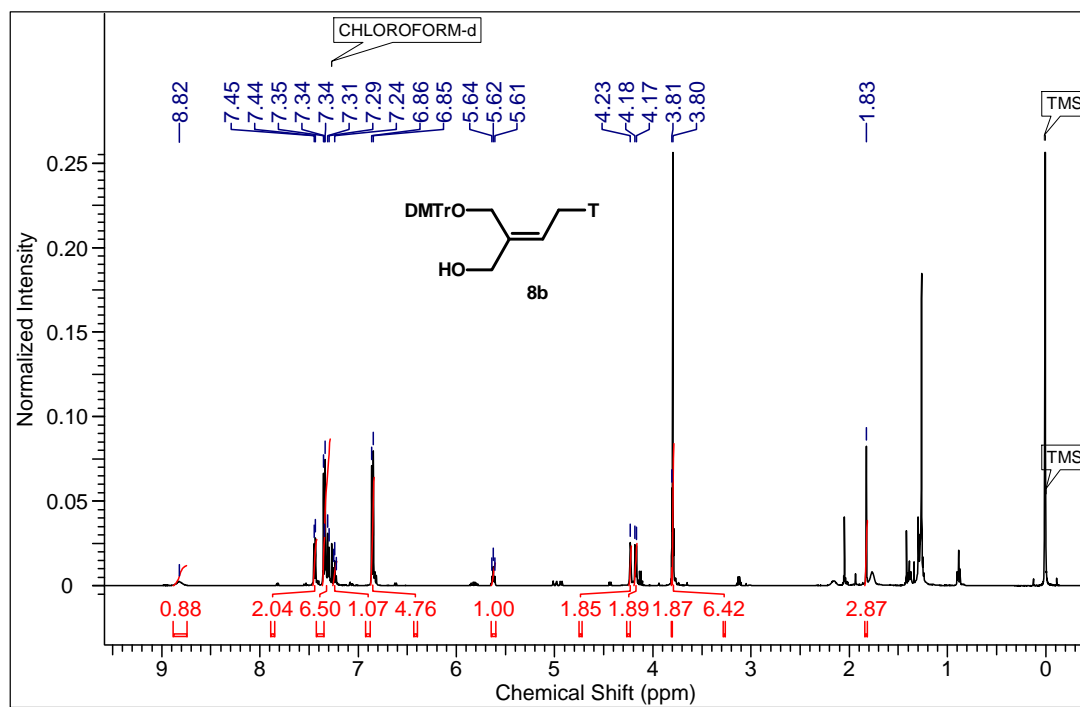




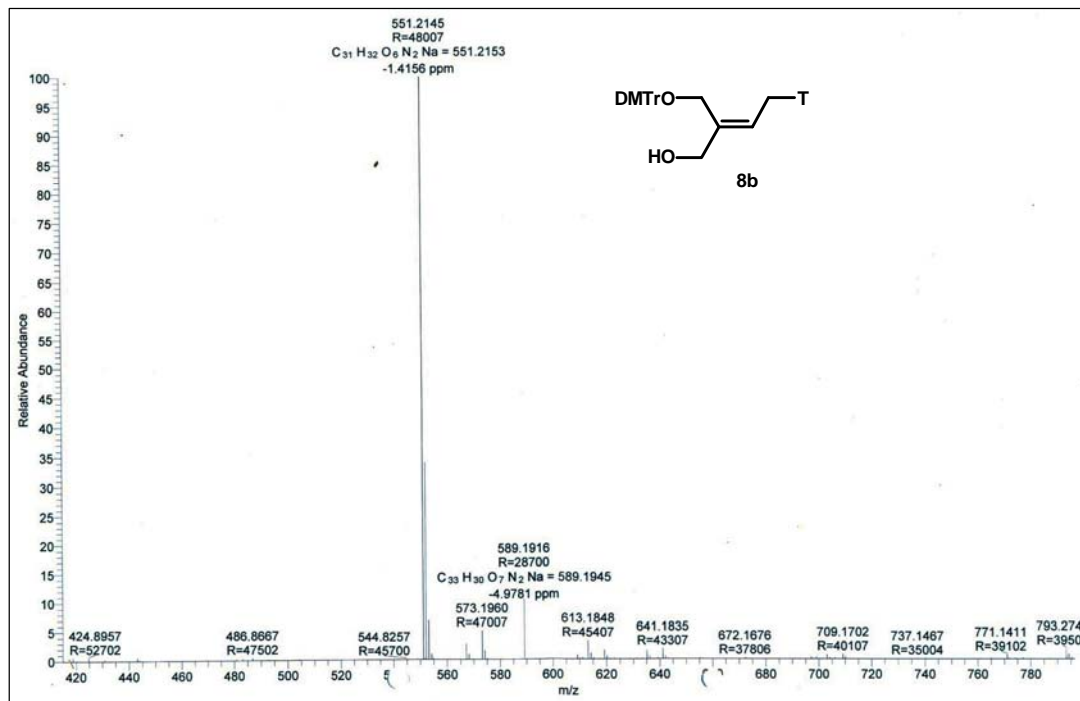
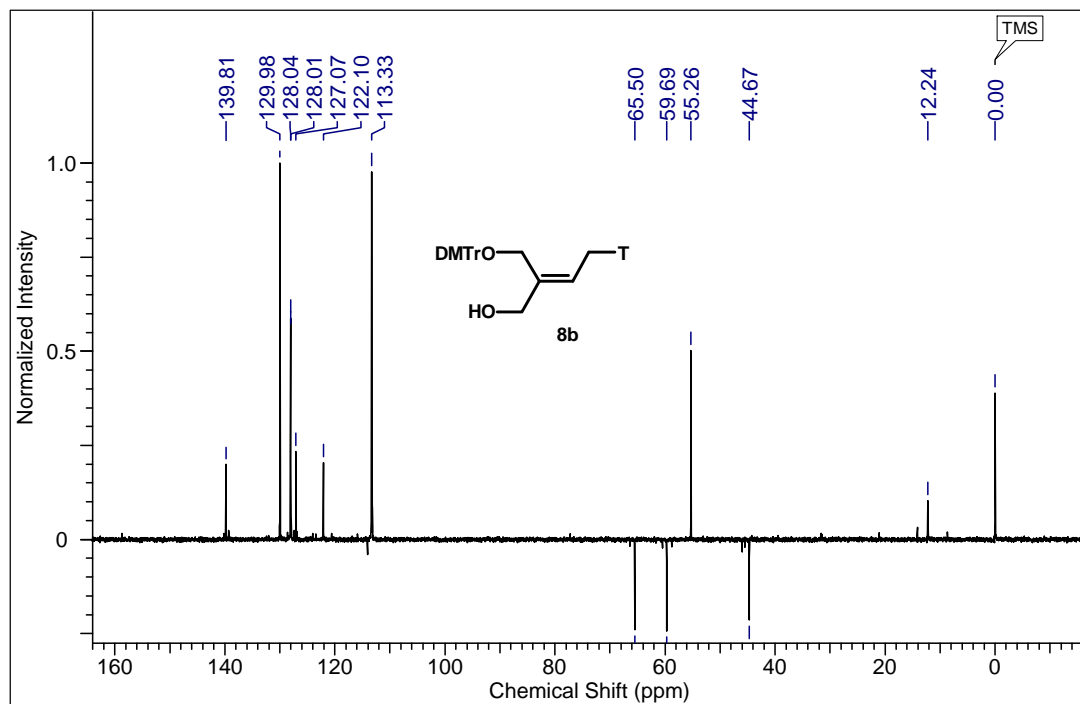


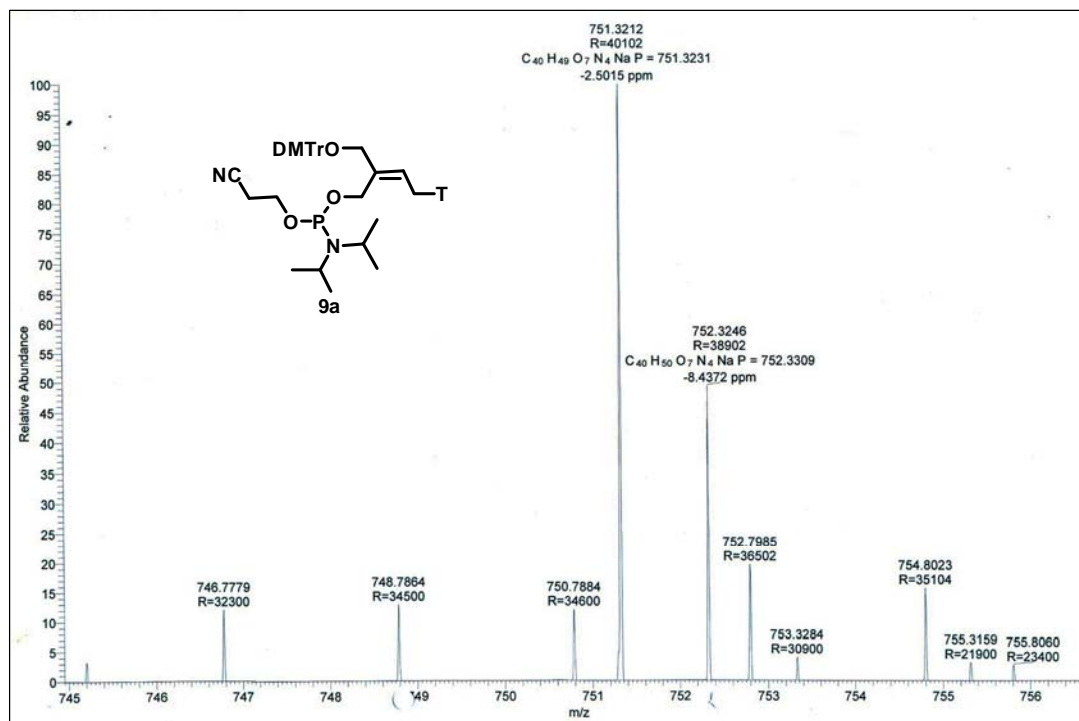
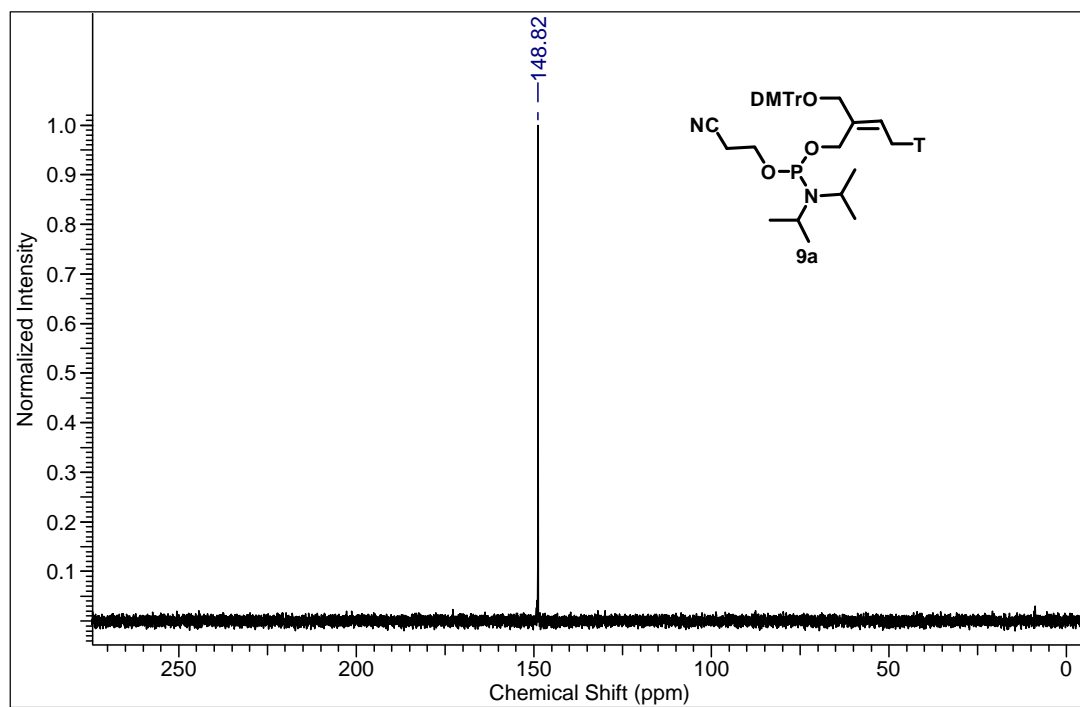


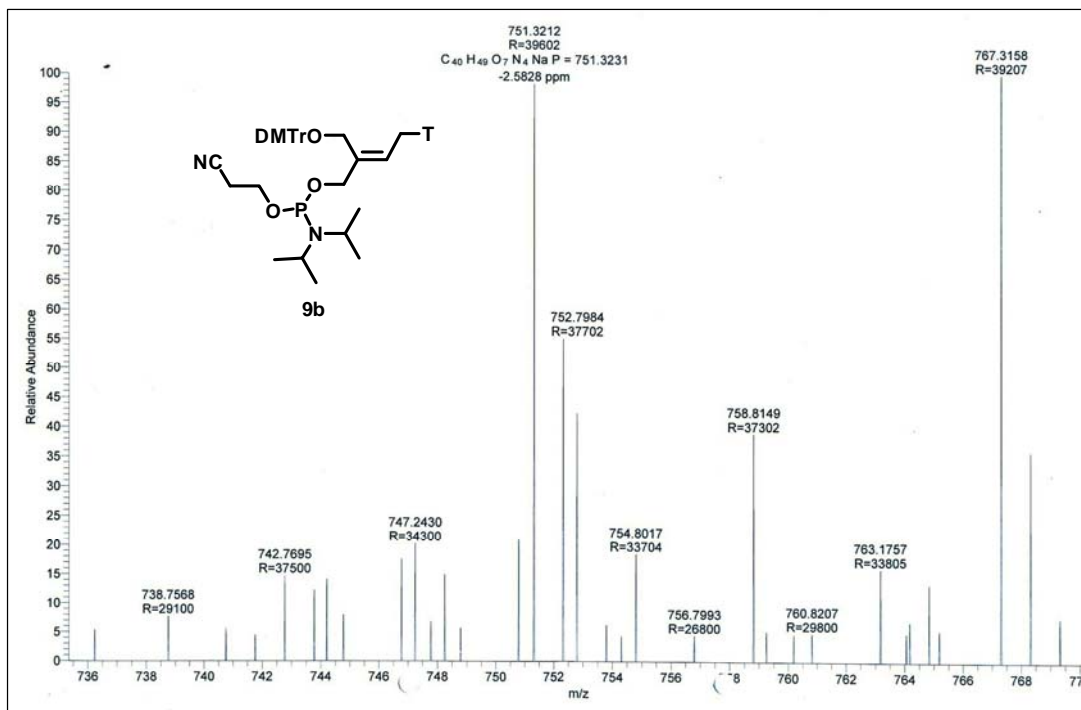
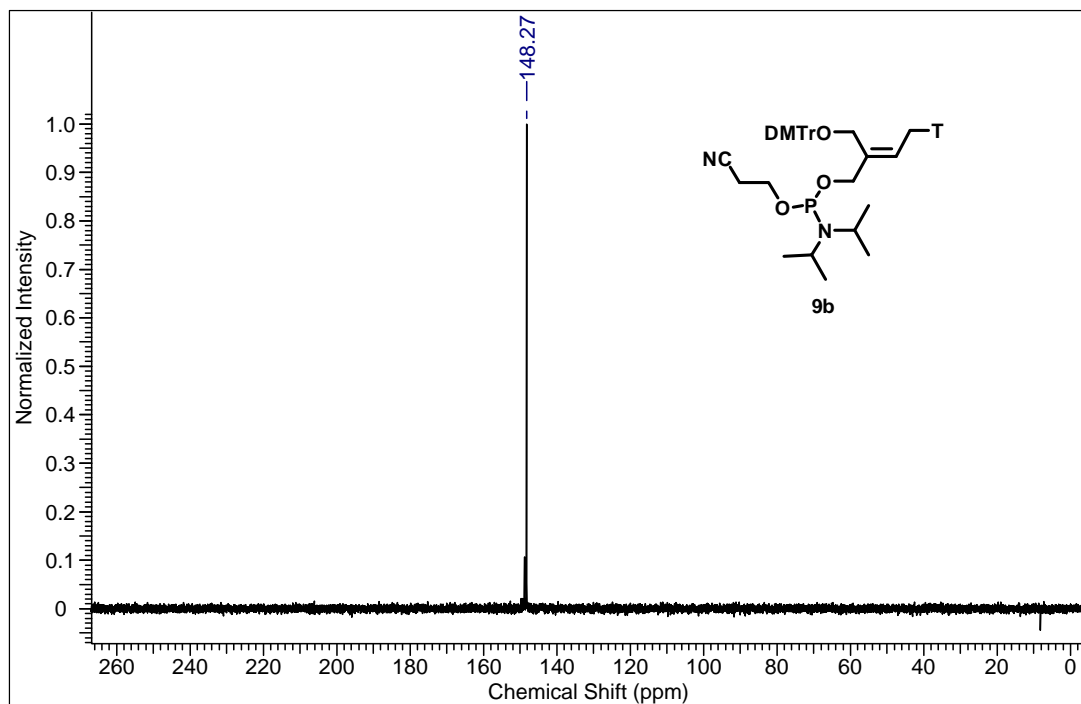




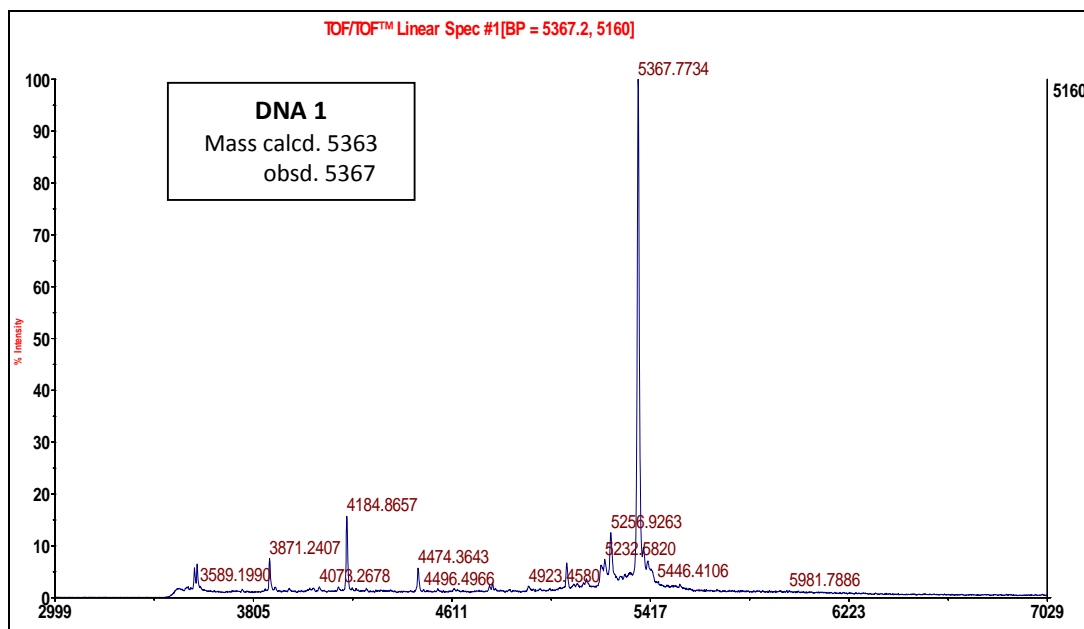
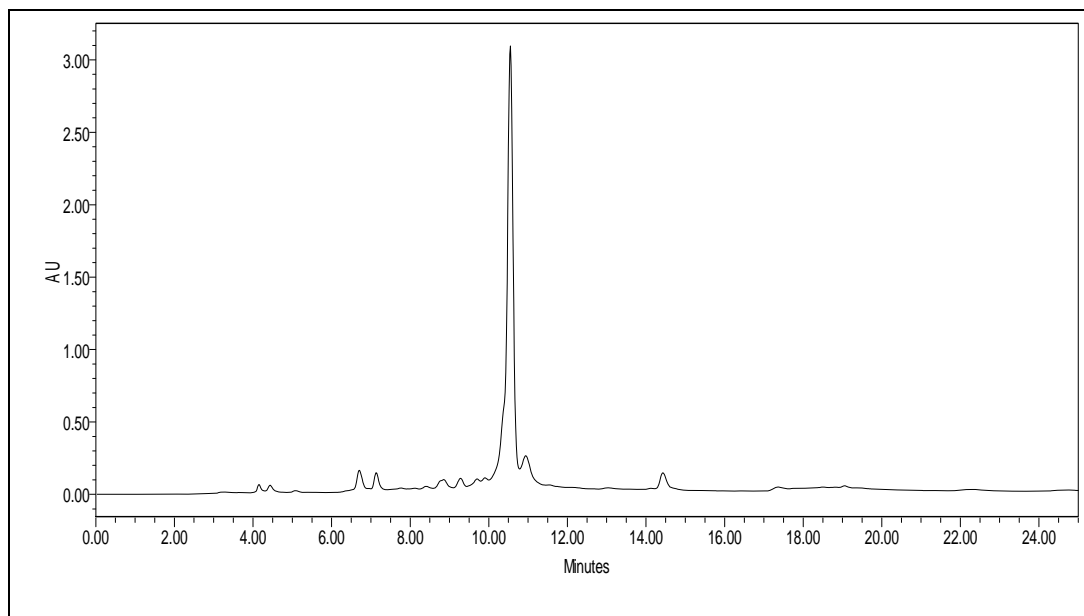




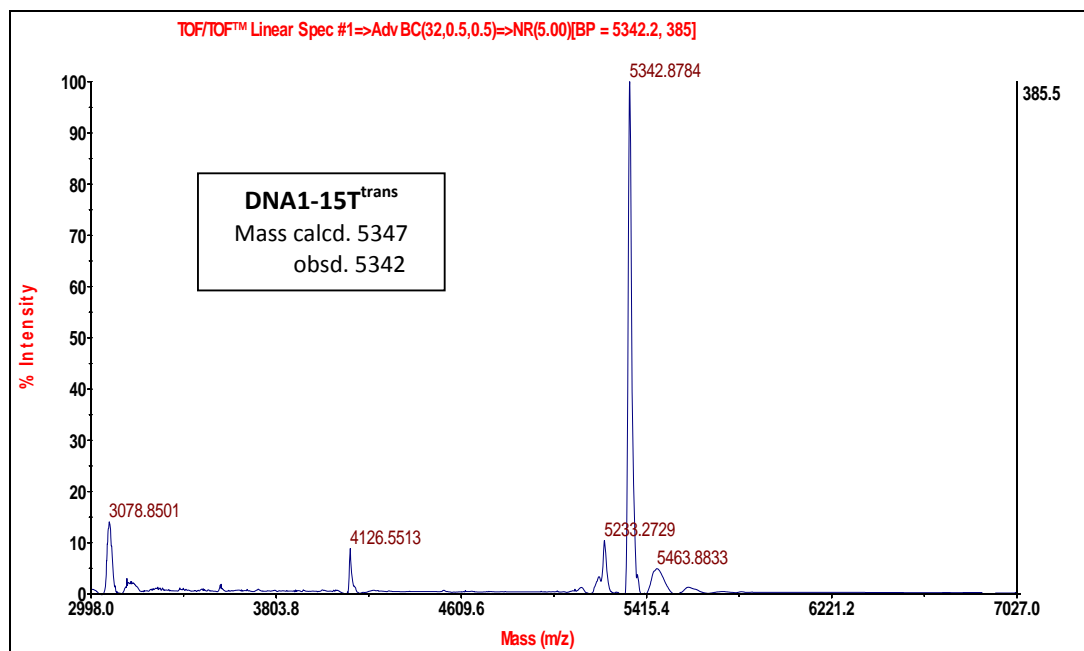
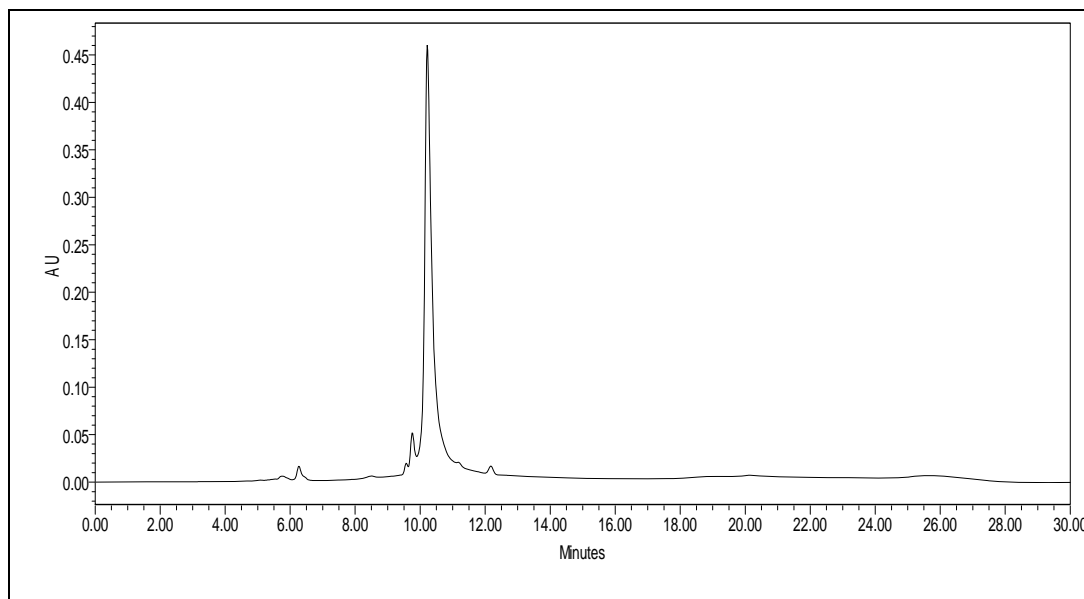




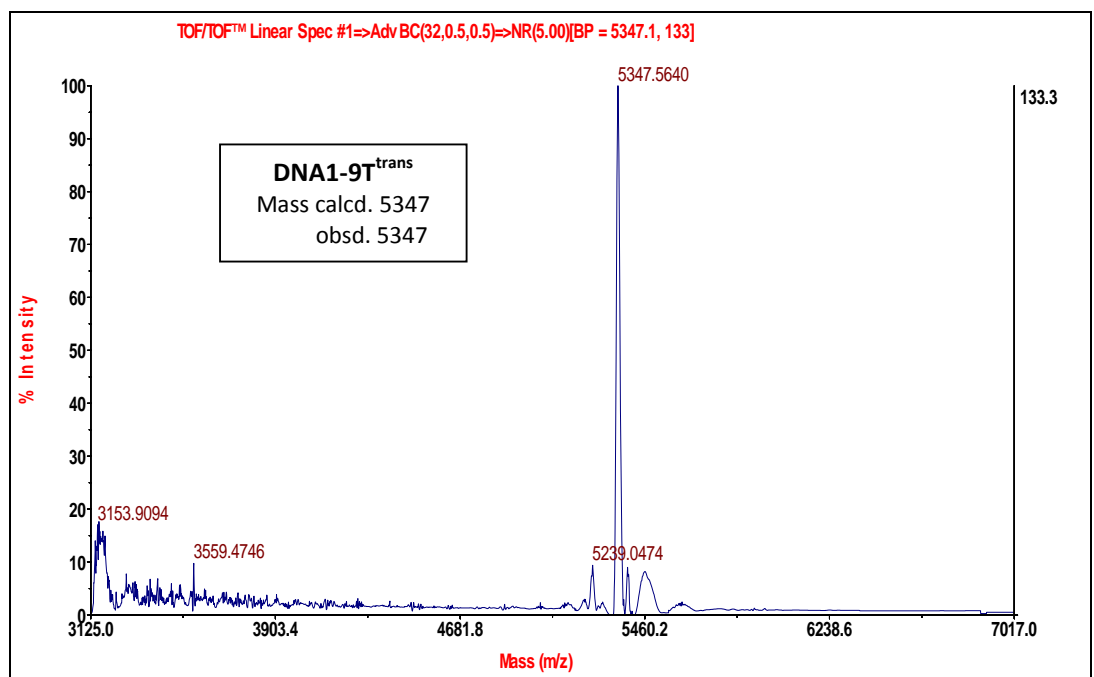
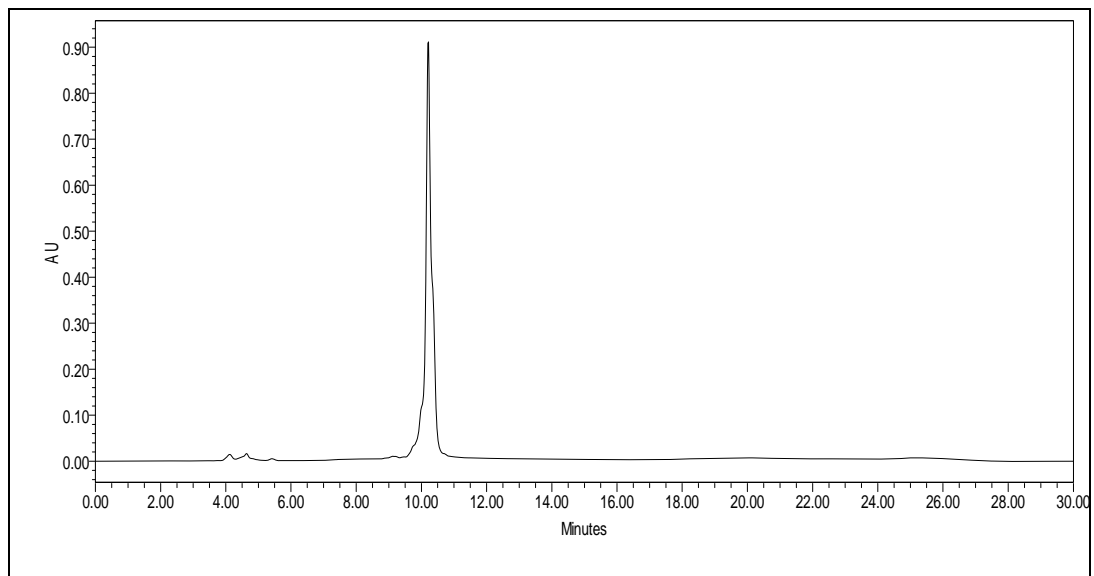
HPLC & MALDI-TOF of DNA 1:



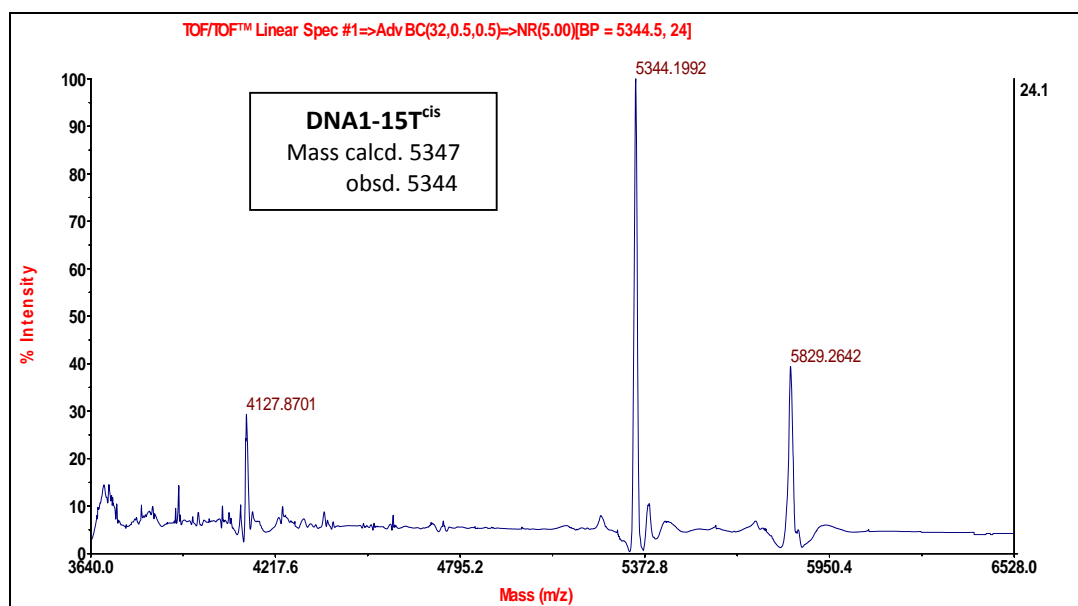
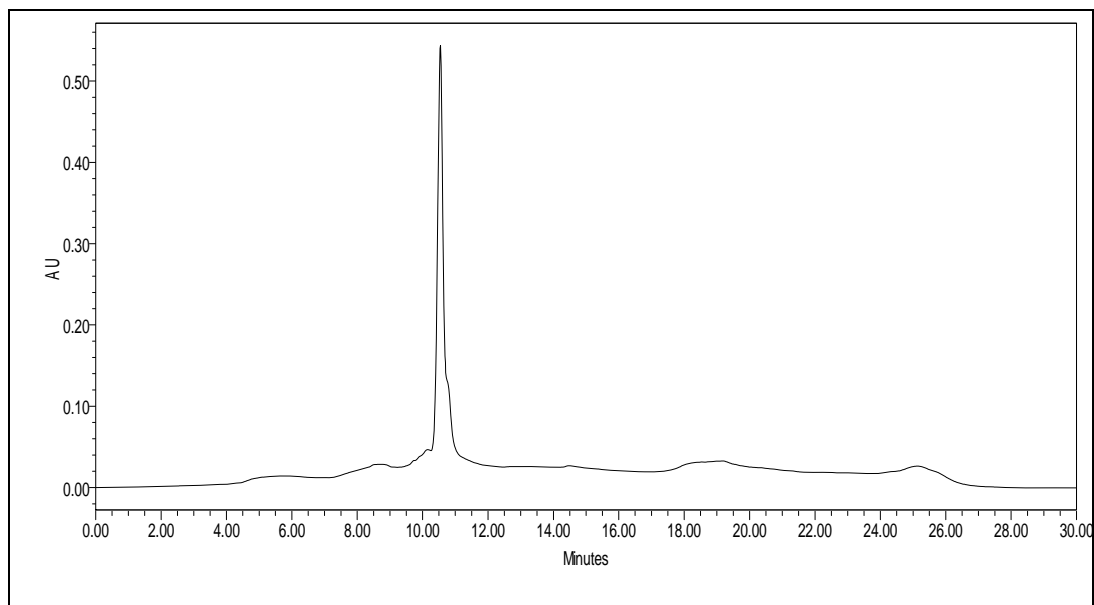
HPLC & MALDI-TOF of DNA1-15T<sup>trans</sup>.



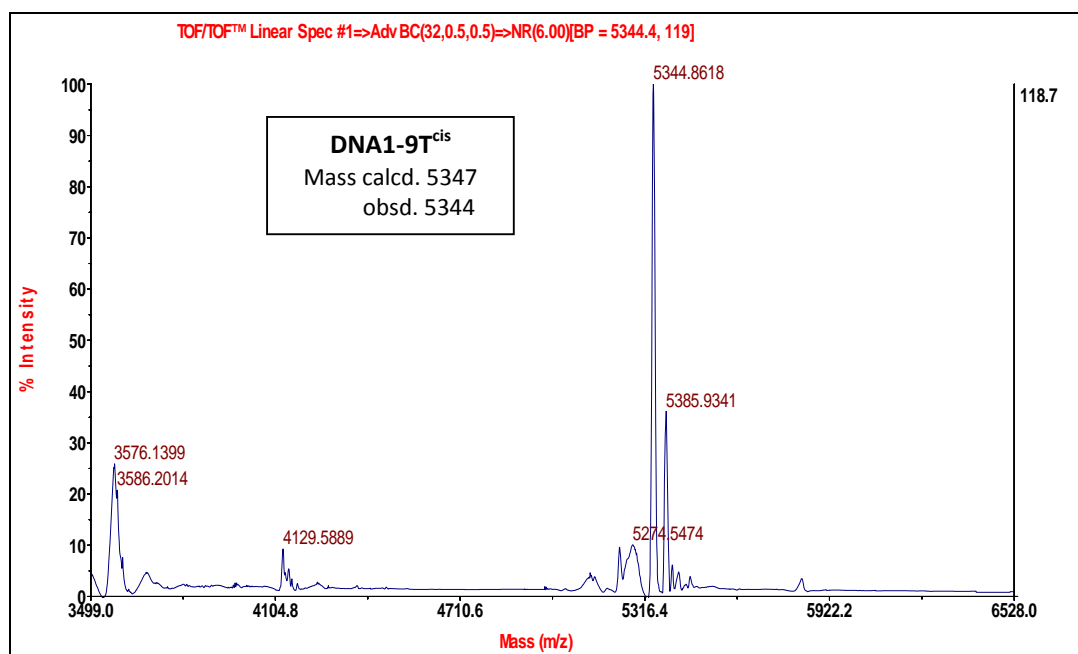
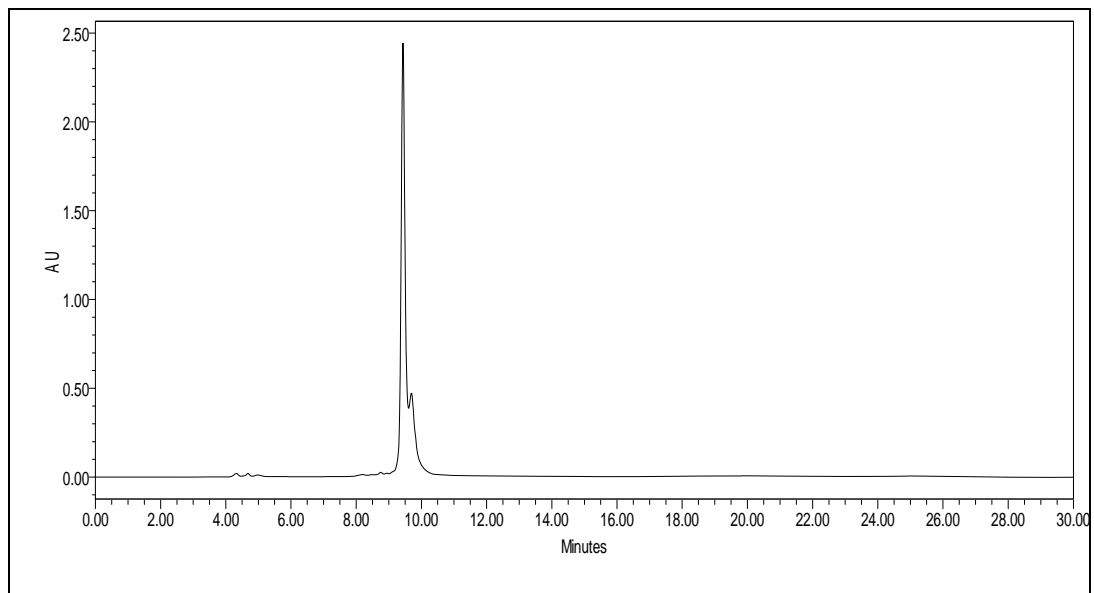
HPLC & MALDI-TOF of DNA1-9T<sup>trans</sup>.



HPLC & MALDI-TOF of DNA1-15T<sup>cis</sup>:



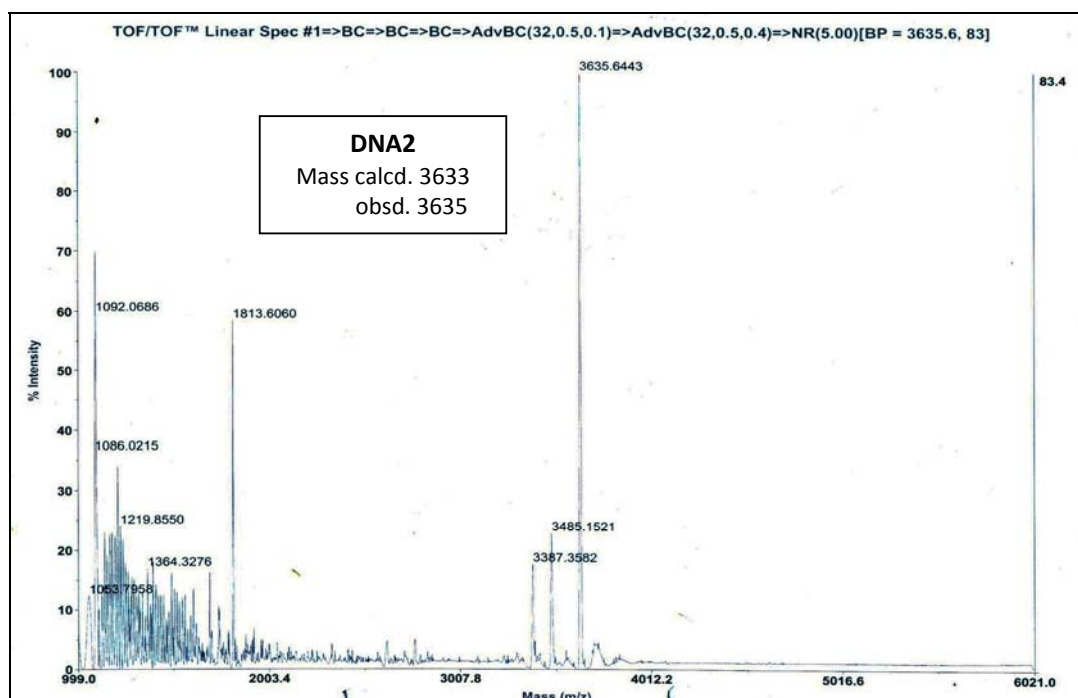
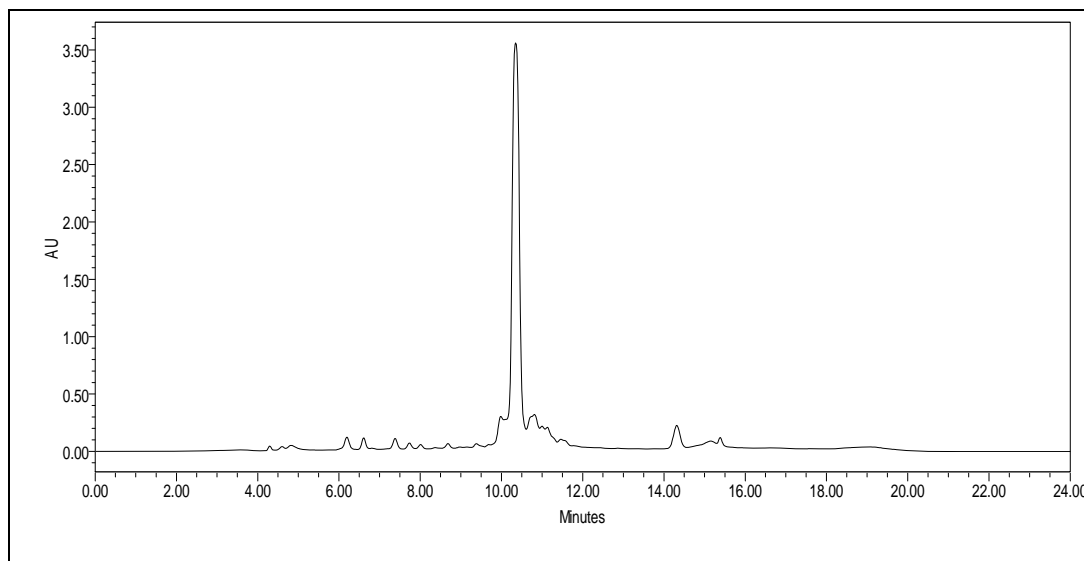
HPLC & MALDI-TOF of DNA1-9T<sup>cis</sup>.





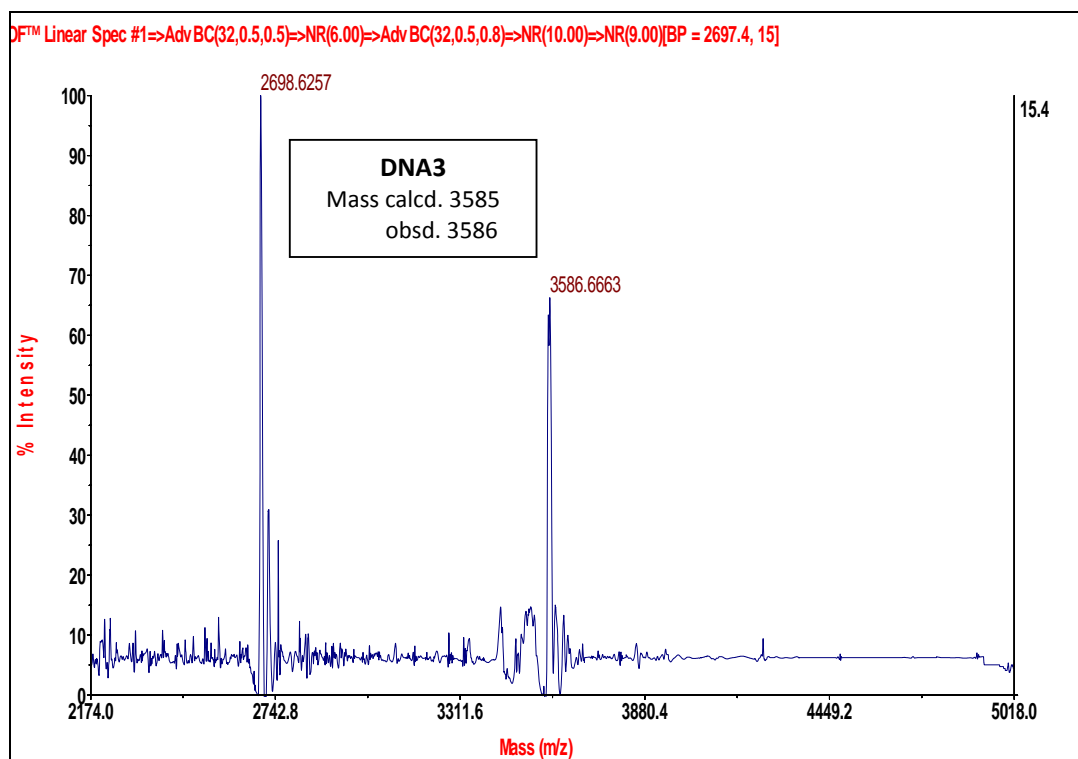
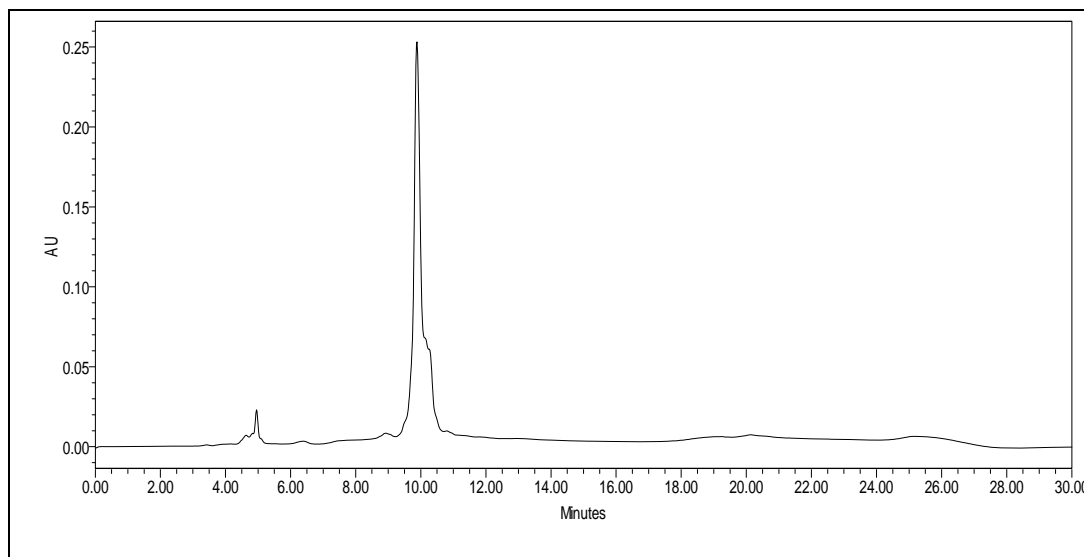
### Chapter 3

HPLC & MALDI-TOF of DNA 2:



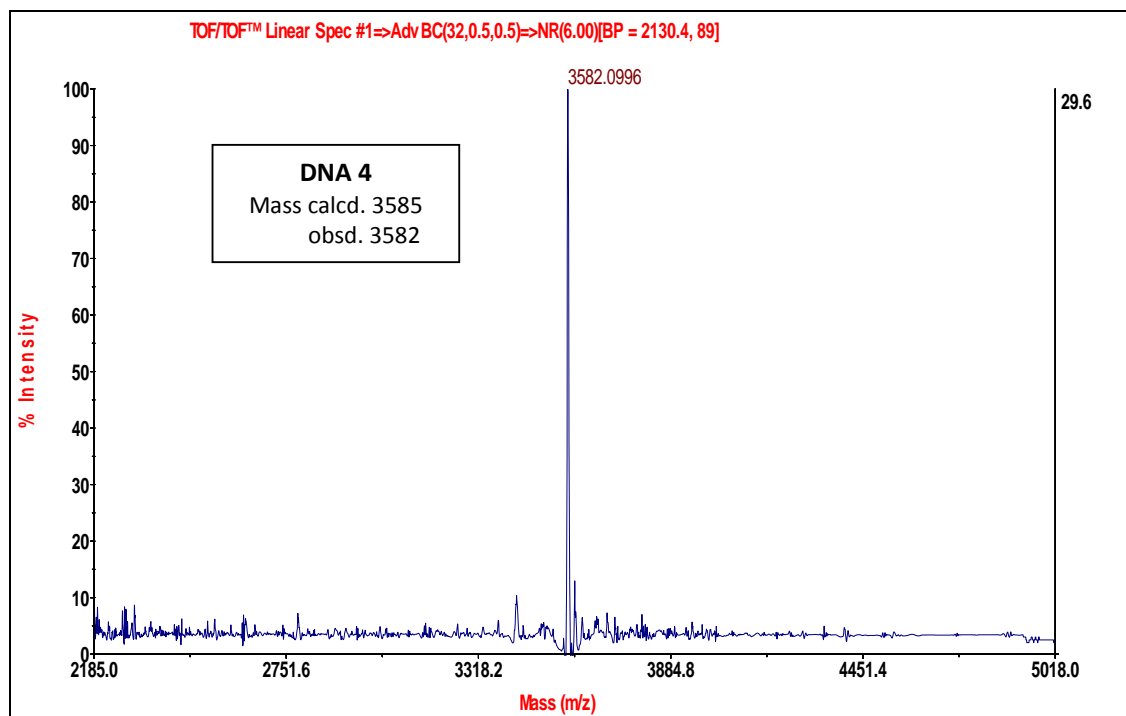
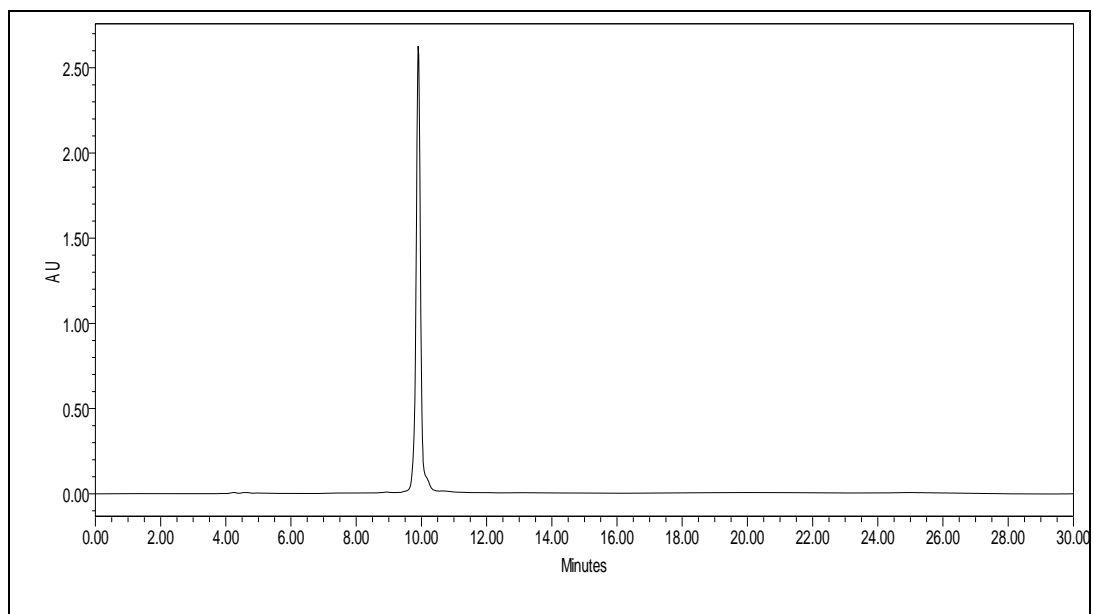
### Chapter 3

#### HPLC & MALDI-TOF of DNA 3:



## Chapter 3

HPLC & MALDI-TOF of DNA 4:



## CHAPTER 3 / SECTION-B

Synthesis of *cis* and *trans*  
modified thrombin binding  
aptamers and their quadruplex  
formation study

## Section 3B: Synthesis of *cis* and *trans* modified thrombin binding aptamers and their quadruplex formation study

### 3B.1 Introduction

G-quadruplexes are a unique class of highly ordered nucleic acid structures, which are formed by four G-rich oligonucleotide strands associate through hydrogen-bonding. The stacked tetrads are stabilized by sandwiched monovalent cations coordinated to the O6 oxygen atoms of the guanines (Figure 7). G-quadruplex structures plays important role of regulatory functions in many biological processes<sup>20</sup> such as telomeres, promoters, centromeres, etc. The d(TTAGGG) repeat sequences, called as telomere, they have recently received great attention because of their potential links to cancer, HIV and other diseases. A unique G-rich DNA sequence in the telomeres was found to protect the chromosomes from recombination, end to end fusion, and degradation through forming G-quadruplexes with highly polymorphic structures in the presence of alkali metal cations. This regular occurrence of quadruplex structures in the genome<sup>21</sup> suggests the designing of potential drug molecules based on DNA/RNA G-quartets.<sup>22</sup>

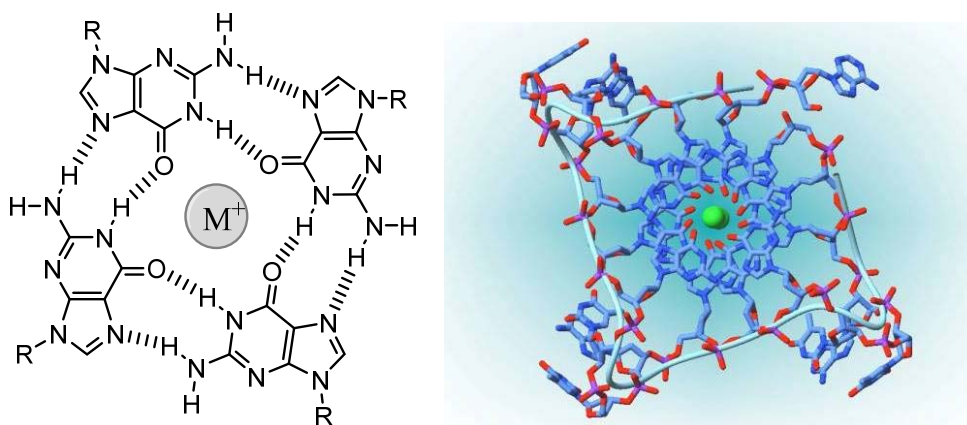


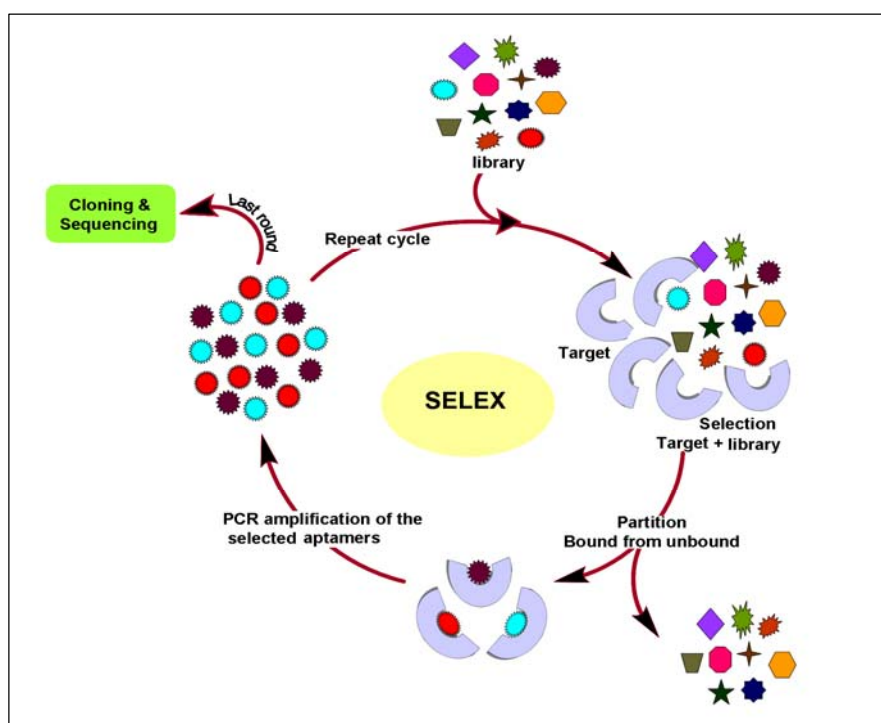
Figure 7: The G-tetrad motif and a G-quadruplex. (<http://www.al-nasir.com>)

### 3B.2 Aptamers: An emerging class of therapeutics

From past 25 years, huge number of nucleic acid ligands also termed aptamers have been developed, which can inhibit the activity of many pathogenic proteins. Aptamers are using for treatment of a variety of human maladies, cancer, infectious diseases, and cardiovascular disease.<sup>23</sup> Aptamers became an attractive class of

### Chapter 3

therapeutic compounds because of their affinity and specificity towards the targets<sup>24</sup>, such as small molecules to peptides, proteins or even whole cells. Their binding affinity and specificity may be equal or even superior to that of antibodies. Natural aptamers also exist in riboswitches. Aptamers can differentiate the small structural differences that may exist as a result of the presence or absence of a small functional group such as methyl or hydroxyl or even as a result of differing chirality. Aptamers were first developed by an *in vitro* selection process termed SELEX (Systematic Evolution of Ligands by EXponential enrichment) independently in the laboratories of Joyce,<sup>25</sup> Szostak<sup>26</sup> and Gold<sup>27</sup> in 1990 (Figure 8). The word ‘aptamer’ was first coined by Ellington and Szostak<sup>26</sup> in 1990.



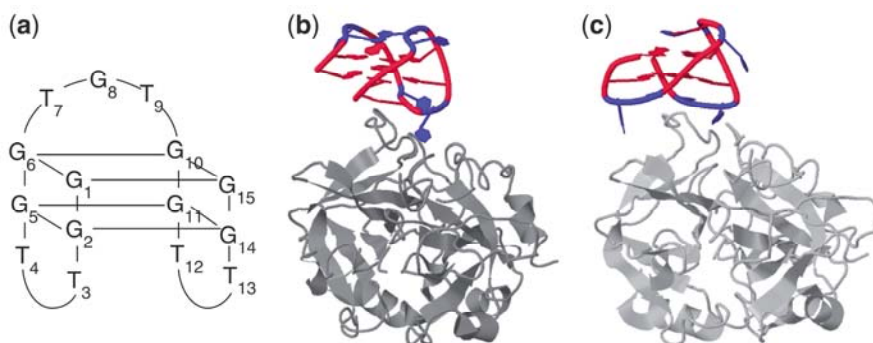
**Figure 8:** SELEX *in vitro* selection protocol ([www.lookfordiagnosis.com](http://www.lookfordiagnosis.com))

Till date, hundreds of aptamers have been introduced against a range of different targets. These can be used for basic research, clinical purposes, in diagnostics, as well as in therapeutics and macromolecular drugs.

### 3B.3 Discovery, structural features of Thrombin-binding aptamer (TBA)

Thrombin is a key regulatory enzyme in the coagulation cascade. It is a serine protease produced from prothrombin by the action of factor Xa. Thrombin will catalyze the conversion of fibrinogen into fibrin, which is the building block of the fibrin matrix of blood clots.<sup>28</sup>

The **thrombin binding aptamer (TBA)** was discovered in 1992 by *in vitro* selection and found to inhibit fibrin- clot formation with high selectivity and affinity. NMR and X-ray structural study reveals that TBA forms an intramolecular, antiparallel G-quadruplex structure with chair like conformation (Figure 9a). This G-quadruplex consists of two G-quartets connected by three edge wise loops, a central TGT loop and two TT loops. The aptamer interacts with two thrombin molecules, inactivating only one of them. X-ray studies indicated that inhibition of fibrinogen-clotting is a result of specific blocking of the thrombin anion exosite I. The central TGT loop (Figure 9b)<sup>29</sup> and two TT loops are involved in ionic interactions with the electropositive heparin binding site of a second thrombin molecule to compensate the residual negative charge of the aptamer. In contrast, NMR studies indicated that the two TT loops interact with the thrombin anion exosite I (Figure 9c),<sup>30</sup> while the TGT loop is in close proximity to the heparin binding site of a neighbouring thrombin molecule.



**Figure 9** (a) Quadruplex structure of the thrombin binding aptamer (TBA), interaction with the thrombin anion exosite I according to (b) X-ray and (c) NMR studies.

(*Nucleic Acids Research*, 2011, Vol. 39, No. 3)

### 3B.4 Modifications in TBA

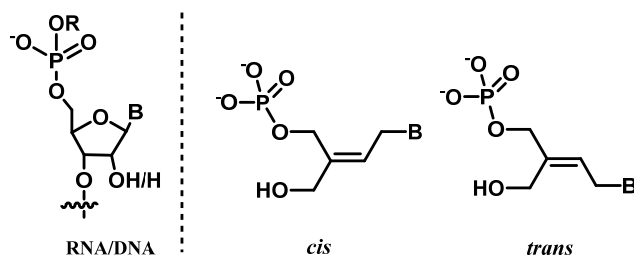
Although several DNA/RNA aptamers are able to show interesting and promising pharmacological properties, rarely used as therapeutic agents without modifications due to less stability towards hydrolytic enzymes. To overcome these problems lot of attempts has been done via chemical and structural modifications.<sup>31</sup>

Unlocked nucleic acids substitution at T-7 position of TGT loop, resulted in enhanced the quadruplex stability, anticoagulation activity as well as enzymatic stability.<sup>32</sup> LNA substitutions showed the reduction of anticoagulation activity.<sup>33</sup> The effect on quadruplex stability of North-nucleoside in the loops of TBA was reported by Eritja and coworkers.<sup>34</sup> The replacement of thymidines in the TGT loop of the TBA quadruplex by uridine (U) and 2'-fluorouridine (FU) induced greater stability to the antiparallel quadruplex structure, determined by UV- $T_m$  experiments and CD spectroscopy. However the presence of North-methanocarbathymidine (NT) in the same positions destabilized the quadruplex structure. Also, substitution of thymidines in the TT loops by U, FU and NT destabilized the antiparallel quadruplex structure. Thus the changes in sugar conformations of the nucleotides in the loop region of TBA are important in determining the stability of TBA quadruplex structure.

### 3B.5 Present work

The acyclic UNA analogues were introduced by Wengel and co-workers to study the stability duplexes.<sup>35</sup> They also introduced UNA analogues in TBA<sup>36</sup>, found to be an excellent application in stabilizing the structure due to its ability to alleviate strain in quadruplex loop structure<sup>31</sup>. We studied the flexibility parameter of our open chain-NA modification (Figure 10) by introducing it in the loop region of TBA quadruplex in comparison with unmodified TBA and with the UNA modification of TBA. The replacement of T3 and T7 positions of thymidine by UNA units were found to stabilize the quadruplex structure of TBA. We chose these two positions for replacing the thymidinyl units of TBA to study its effect on the quadruplex stability.





**Figure 10** The *cis*- and *trans*- Ene-nucleotides

### 3B.6 Chemical synthesis of *cis* and *trans* modified thrombin binding aptamers

Using an automated Bioautomation MM-4 DNA synthesizer and commercially available phenoxyacetyl (Pac) protected 5'-*O*-dimethoxytrityl-2'-deoxy-3'-phosphoramidites four sequences, **TBA-3T<sup>cis</sup>**, **TBA-7T<sup>cis</sup>**, **TBA-3T<sup>trans</sup>** and **TBA-7T<sup>trans</sup>** were synthesized by following the reported protocol which was already discussed in (section-A). The 3'-5'-linked **TBA** was synthesized for control experiments. All modified sequences were purified by HPLC and purity was confirmed by gel-electrophoretic mobility studies (Figure 11) and characterized by MALDI-TOF mass spectrometry (Table 4).

### 3B.7 G-quadruplex formation in the presence of monovalent cation using CD spectroscopy

The G-quadruplex formation for the synthesized sequences was studied by CD spectroscopy<sup>37</sup> in the presence of added monovalent cation such as K<sup>+</sup> and their stability was determined as a function of temperature-dependent change in CD amplitude at 295nm.

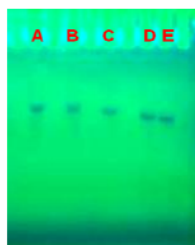
All four modified sequences showed maxima at 295nm which is characteristic CD signature for formation of stable antiparallel quadruplex (Figure 12a). The stability of the G-quadruplexes was studied by the change in the amplitude of the CD signal at 295nm with temperature (Figure 12b, Table 4). The CD melting results tells that T7 position modifications are showing the stabilization of antiparallel quadruplex structures as compared to their corresponding T3 position modifications. The modifications of *cis* and *trans* units at both T3 and T7 positions showed the destabilization of tetraplex structure compared to unmodified TBA. This may indicate that the open chain-NA modification is indeed more strained and is less suitable for quadruplex formation compared to the highly evolved DNA quadruplexes.

## Chapter 3

**Table 4** Modified TBA sequences, their MALDI-TOF mass analysis and biophysical evaluation by CD- $T_m$  measurements

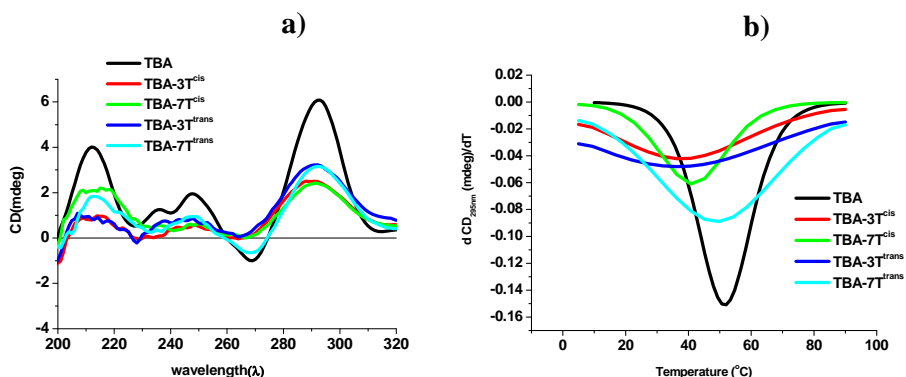
Name	Sequence <sup>a</sup> 5'→3'	mass cal /obs	HPLC $t_R$ (min)	CD $T_m$ °C
TBA	ggttggtgtggttgg	4726/4730	9.8	49.5
TBA-3T <sup>cis</sup>	ggT <sup>cis</sup> tggtgtggttgg	4710/4709	10.0	38
TBA-7T <sup>cis</sup>	ggttggT <sup>cis</sup> gtggttgg	4710/4709	10.1	41.4
TBA-3T <sup>trans</sup>	ggT <sup>trans</sup> tggtgtggttgg	4710/4714	9.9	36.1
TBA-7T <sup>trans</sup>	ggttggT <sup>trans</sup> gtggttgg	4710/4708	10.3	43.7

<sup>a</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. Gel-electrophoretic mobility studies revealed that all the modified sequences were more than 95% pure.



A) TBA  
B) TBA-3T<sup>cis</sup>  
C) TBA-7T<sup>cis</sup>  
D) TBA-3T<sup>trans</sup>  
E) TBA-7T<sup>trans</sup>

**Figure 11** Gel picture of purified sequences



**Figure 12** (a) CD spectra of oligomers TBA, TBA-3T<sup>cis</sup>, TBA-7T<sup>cis</sup>, TBA-3T<sup>trans</sup>, TBA-7T<sup>trans</sup> sequences of 5 μM concentration in 10 mM potassium phosphate buffer (pH 7.5) containing 100 mM KCl at 5 °C. (b) Temperature-dependent changes in CD amplitude at 295 nm plotted against temperature, first derivative plots at strand concentration 5 μM in 10 mM potassium phosphate buffer (pH 7.5) containing 100 mM KCl.

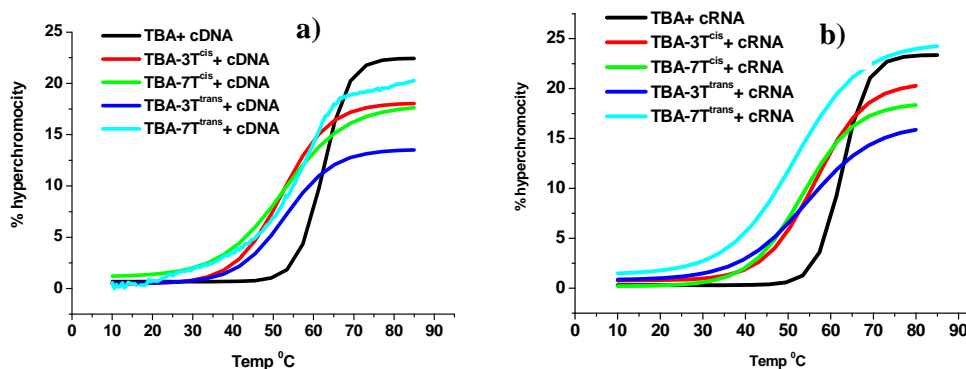
## 3B.8 Duplex stability studies of modified TBA oligomers

The binding affinity of 15mer TBA ONs **TBA**, **TBA-3T<sup>cis</sup>**, **TBA-7T<sup>cis</sup>**, **TBA-3T<sup>trans</sup>** and **TBA-7T<sup>trans</sup>** with complementary DNA and RNA was also investigated by measuring the melting temperatures (UV  $T_m$ ) of the duplexes (Table 5).

**Table 5** UV  $T_m$  ( $^{\circ}\text{C}$ )<sup>a</sup> values of TBA and modified TBA: DNA/RNA duplexes

Name	Sequence <sup>b</sup> 5' → 3'	mass cal / obs	CD $T_m$ $^{\circ}\text{C}$	
			cDNA <sup>c</sup>	cRNA <sup>d</sup>
TBA	ggttggtggttgg	4726/4730	62	63
TBA-3T <sup>cis</sup>	ggT <sup>cis</sup> tggtggttgg	4710/4709	54	53
TBA-7T <sup>cis</sup>	ggttggT <sup>cis</sup> gtggttgg	4710/4709	53	54
TBA-3T <sup>trans</sup>	ggT <sup>trans</sup> tggtggttgg	4710/4714	53	54
TBA-7T <sup>trans</sup>	ggttggT <sup>trans</sup> gtggttgg	4710/4708	56	55

<sup>a</sup>UV- $T_m$  values were measured by using 1 $\mu\text{M}$  sequences with 1 $\mu\text{M}$  cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is  $\pm 0.5$   $^{\circ}\text{C}$ ). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5'ccaaccacccaacc was the complementary DNA sequence. <sup>d</sup>5' ccaaccacccaacc was the complementary RNA sequence.



**Figure 13** (a) UV melting profiles of TBA and modified TBA sequences with cDNA, (b) UV melting profiles of TBA and modified TBA sequences with cRNA.

UV  $T_m$  results showed that all modified and unmodified sequences were forming stable duplexes with cDNA as well as cRNA. Independent to the nucleoside

and position of modifications,  $T_m$  values were same for all sequences with cDNA, cRNA. These *cis* and *trans* modifications are destabilizing the duplexes compared to unmodified TBA.

### 3B.9 Conclusions:

- Successfully synthesized the *cis* and *trans* modified TBA sequences for the first time and purity was confirmed by gel-electrophoretic mobility studies.
- *cis* and *trans* modified TBA sequences are forming stable antiparallel quadruplex in the presence of  $K^+$  ions.
- CD melting results showed that modifications at T7 (TGT loop) position, forming more stable quadruplex structures than T3 (TT loop) position. The stability of both T7 and T3 modifications was less, compared to unmodified TBA.
- The destabilization results may indicate that the open chain-NA modification is indeed more strained compared to UNA and is less suitable for quadruplex formation compared to the highly evolved DNA quadruplexes.
- Duplex stability studies also done for the TBA modified sequences with cDNA/cRNA, modification are destabilizing the duplex structures.

### 3B.10 Experimental Section

#### MALDI-TOF mass

Mass was obtained by MALDI-TOF mass spectrometry. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems). The matrix used for analysis was THAP (2', 4', 6'-trihydroxyacetophenone).

#### CD spectroscopy

CD spectra were recorded on Jasco J-815 CD Spectrometer equipped with a Jasco PTC-424S/15 peltier system. 2 mm path-length quartz cuvettes were used for a sample volume 500  $\mu$ l and strand concentration of 5  $\mu$ M in 10mMol Na/K-phosphate buffer( pH 7.5) containing 100mM NaCl/KCl respectively. Oligomers prepared in buffer were annealed by heating at 95° C for 5 minutes then slowly cooling to room temperature followed by refrigeration for 5 to 6 hours before use. Spectral scans were collected over a wavelength range 200- 320nm at a scanning rate of 100 nm min<sup>-1</sup>. Three scans were averaged for each sample. CD thermal denaturation of the TBA sequences, Thrombin binding studies were performed.

#### Buffers used for CD experiments

##### Phosphate buffer (pH = 7.2, 100 mM NaCl)

Na<sub>2</sub>HPO<sub>4</sub> (110mg, 10mM), NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O (35.3mg,10mM), NaCl (585mg,100mM) was dissolved in minimum quantity of water and the total volume was made 100 mL. The pH of the solution was adjusted 7.2 with aq. NaOH solution in de-ionised water (DI), and stored at 4°C.

##### Phosphate buffer (pH = 7.2, 100 mM KCl)

K<sub>2</sub>HPO<sub>4</sub> (174mg,10mM), KH<sub>2</sub>PO<sub>4</sub> (136mg,10mM), KCl (746mg,100mM) was dissolved in minimum quantity of water the total volume was made 100 mL. The pH of the solution was adjusted 7.2 with aq. KOH solution in DI water, stored at 4°C.

#### UV-T<sub>m</sub> experiments

Thermal denaturation of the *cis* and *trans* modified oligomers was performed using a 10 mm quartz cell in a Cary 300 Bio UV-Visible Spectrophotometer Varian. The

## Chapter 3

---

mixture of 1 $\mu$ M oligomers and 1 $\mu$ M cDNA/cRNA were annealed in a 10 mMol potassium phosphate buffer pH 7.5, 150mMol NaCl. The concentration was calculated on the basis of absorbance from molar extinction coefficients of the corresponding nucleobases of DNA/RNA. Absorbance *versus* temperature profiles were obtained by monitoring the absorbance at 254 nm from 10–85°C at a ramp rate of 0.5°C per minute. A stream of dry nitrogen was gently applied through the sample compartment to prevent condensation of water on the cuvette at low temperatures.

### **Snake venom phosphodiesterase stability experiments**

Enzymatic hydrolysis of the oligonucleotides (7.5  $\mu$ M) was carried out at 37 °C in buffer (100  $\mu$ l) containing 100 mM Tris-HCl (pH 8.5), 15 mM MgCl<sub>2</sub>, 100 mM NaCl and SVPD (10 $\mu$ g/mL). Aliquots were removed at several time-points; a portion of each reaction mixture was removed and heated to 90 °C for 2 min to inactivate the nuclease. The amount of intact ONs was analyzed at several time points by RP-HPLC. The percentage of intact ON was then plotted against the exposure time to obtain the ON degradation curve with time.

### **Polyacrylamide Gel Electrophoresis (PAGE)**

#### **Preparation of Gel and buffer solutions**

**1. 5X TBE buffer (500 ml,pH 8.0)**

Tris (hydroxymethyl)methyl amine (Tris base) 27g, boric acid 13.75g and EDTA 1.462g was dissolved in 500 ml of deionised water (DI water). The pH was adjusted to 8.0 with tris base or HCl.

**2. 1X TBE buffer (500 ml)**

100ml of 5X TBE buffer was diluted to 500ml with DI water.

**3. 29:1 acrylamide : N,N-methylene bis-acrylamide solution.**

29g of acrylamide and 1g of N,N-methylene bis-acrylamide was dissolved in 100ml of DI water.

**4. Bromophenol Blue dye (marker dye)**

40% of glycerol solution in DI water (V/V) + 0.25% Bromophenol Blue dye solution was prepared.

**5. 40% sucrose solution in DI water (w/v).**

This solution was mixed with the Bromophenol Blue dye solution and the sample to be loaded in 1:1 ratio, to make the loading solution viscous and heavy for proper loading into the gel wells.

### 6. 10% ammonium persulfate solution.

50mg ammonium persulfate was dissolved in 500 µl DI water. Fresh solution prepared just before preparing the gel for casting.

### 7. 20% acrylamide gel.(19.314% acrylamide+ 0.6% cross linker)

To prepare 10 ml gel solution, 6.66 ml of acrylamide (29:1) +1.27 ml DI water +2 ml of 5X TBE buffer were mixed and degassed for 15-20 minutes under vacuum. Then 70 µl of ammonium persulfate was added followed by 4.6 µl of TEMED (N,N,N',N'-tetraethylmethylenediamine)

### 8. Gel casting procedure.

The above prepared gel solution after swirling for a minute was immediately poured into the previously cleaned and fixed gel plates. The appropriate gel comb was inserted for the formation of wells. Then the cast gel was allowed to stand for 40 minutes or more till it was set to a proper polymerised gel.

### 9. Gel experiment

**Pre-run:** The set gel was rinsed with DI water to remove any excess of unpolymerised gel after the comb was removed. For a pre-run or blank run, the gel plate assembly was placed in the gel run chamber and completely immersed in the 1X TBE buffer. Each well was loaded with 2 µl of the bromophenol blue dye in 40% sucrose solution(1:1). Then 200V voltage was applied and the gel run was carried out at 4°C for 1hour till the marker dye had travelled down and washed out along with any unpolymerised gel.

#### **Sample run:**

The DNA oligomer samples 2 µl solution (350 µmol concentration) were mixed with equal volume of the 40% sucrose solution and loaded into the appropriately numbered wells. The marker dye was loaded in the first and last wells to monitor the run. The gel was run with the voltage set at 150V for 120min till the marker was visible at 3/4<sup>th</sup> the gel height.

#### **Gel visualization:**

The gels after run were washed with DI water and then were visualized by UV-shadowing.

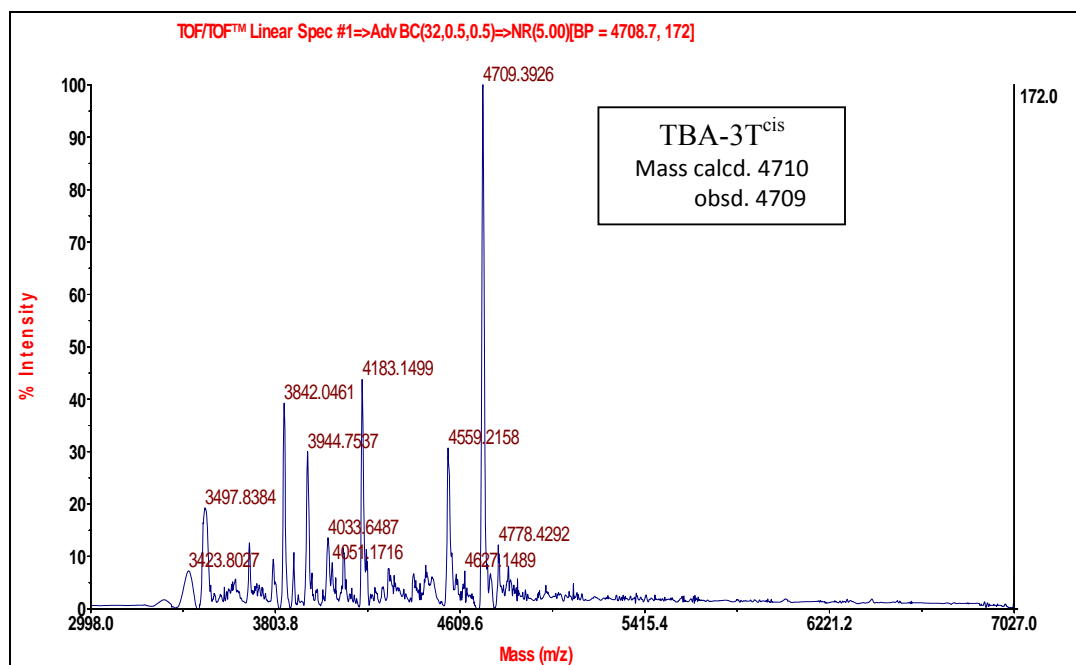
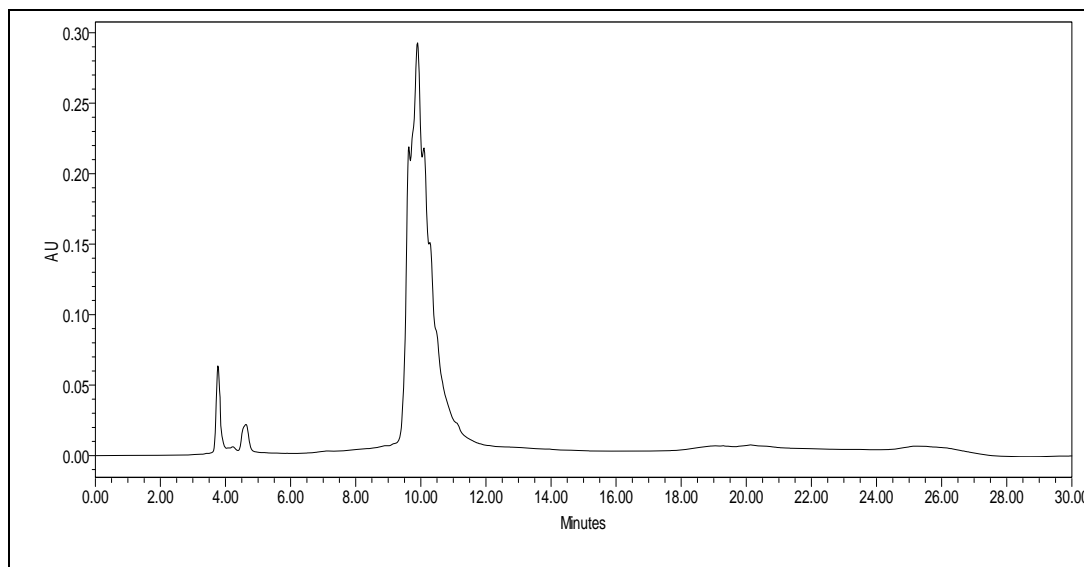
**3B.11 Appendix**

<b>Compounds - Spectral data</b>	<b>Page No.</b>
<b>HPLC &amp; MALDI-TOF of TBA-3T<sup>cis</sup></b>	157
<b>HPLC &amp; MALDI-TOF of TBA-7T<sup>cis</sup></b>	158
<b>HPLC &amp; MALDI-TOF of TBA-3T<sup>trans</sup></b>	159
<b>HPLC &amp; MALDI-TOF of TBA-7T<sup>trans</sup></b>	160



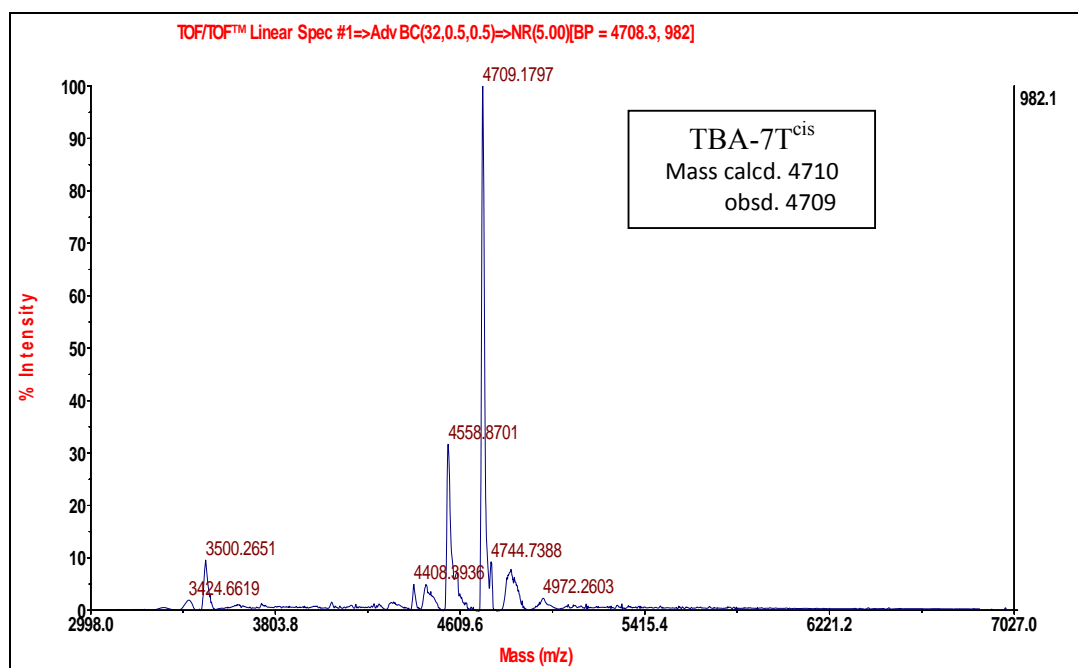
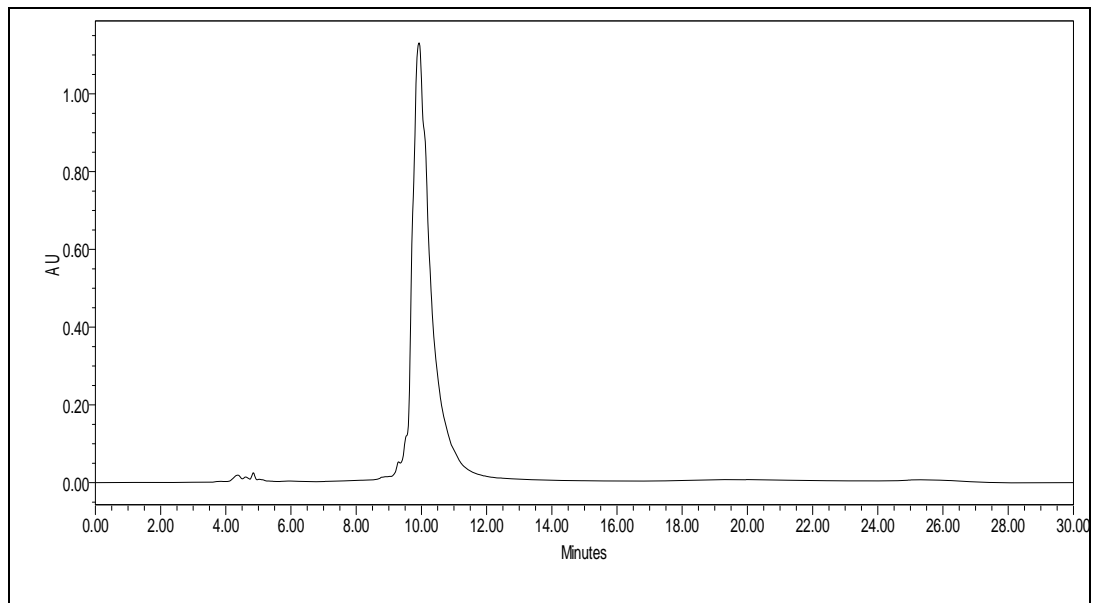
## Chapter 3

HPLC & MALDI-TOF of TBA-3T<sup>cis</sup>:



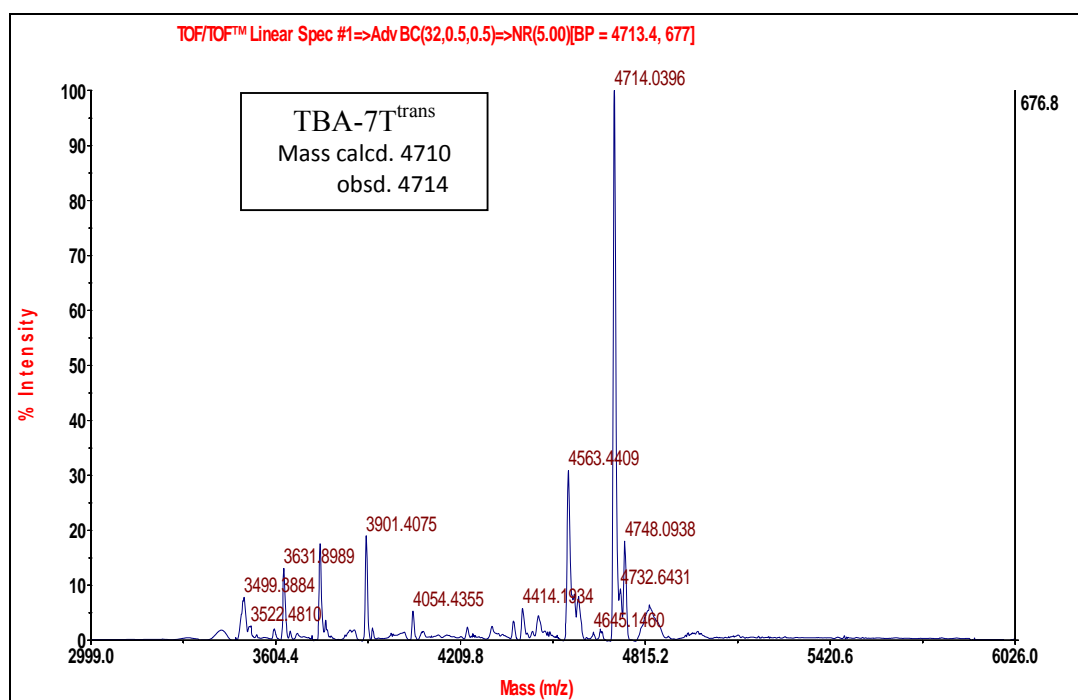
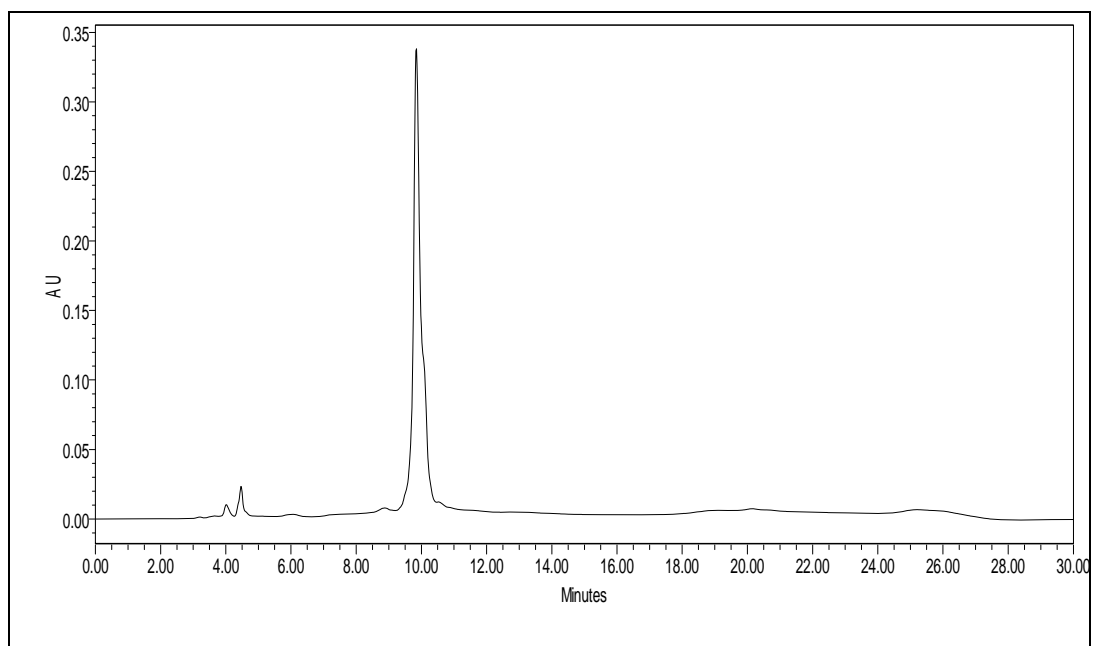
### Chapter 3

HPLC & MALDI-TOF of TBA-7T<sup>cis</sup>:



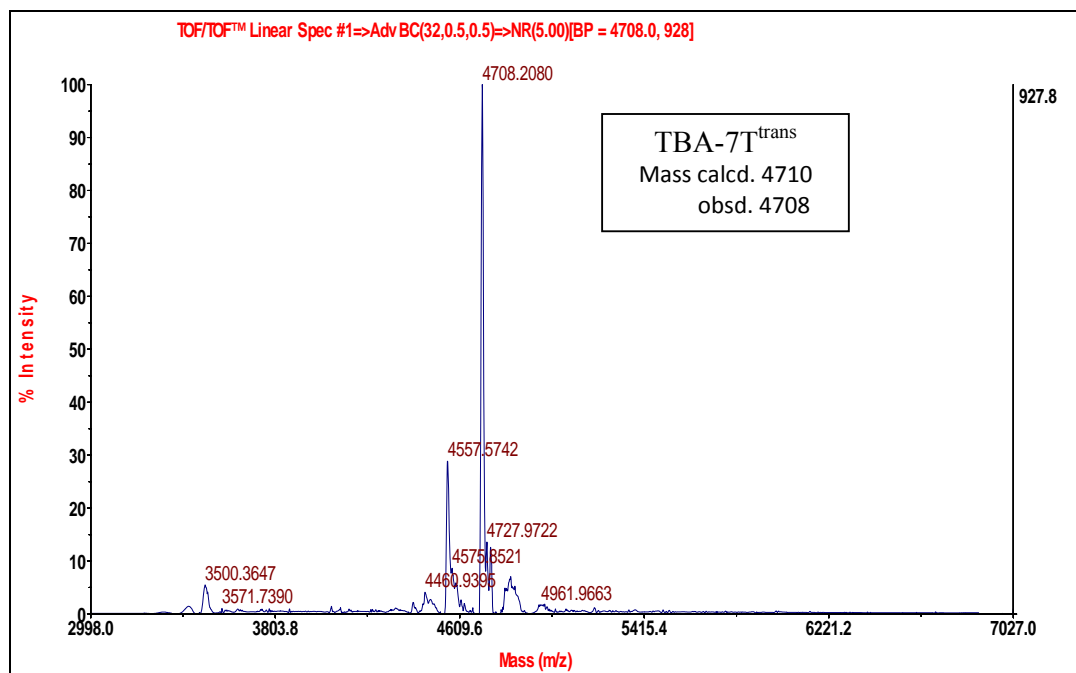
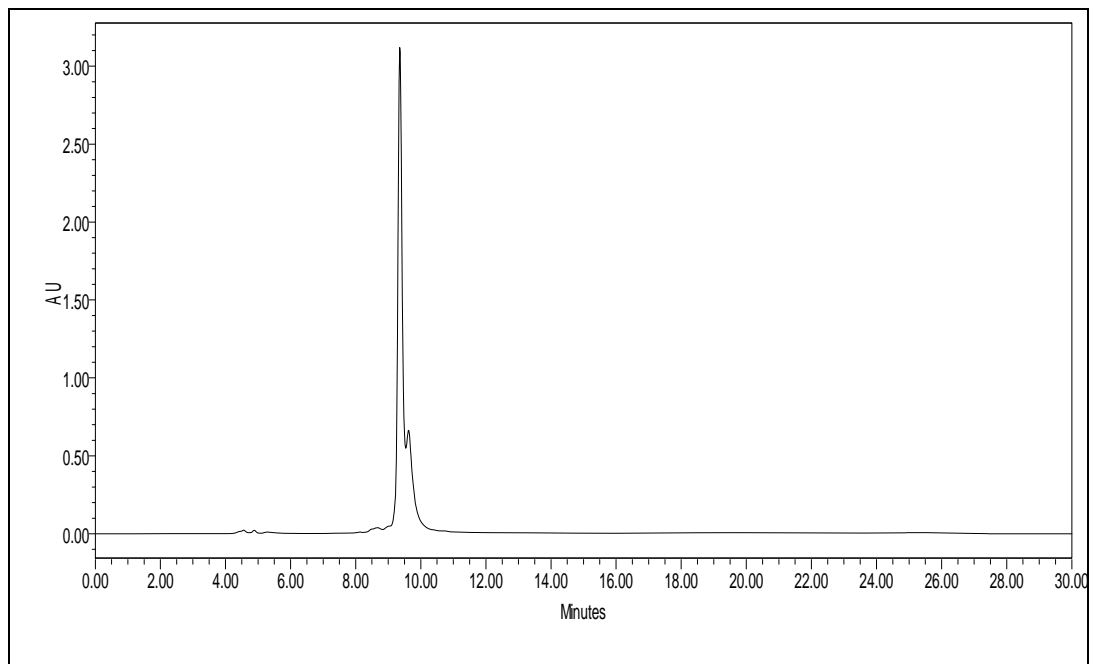
## Chapter 3

### HPLC & MALDI-TOF of TBA-3T<sup>trans</sup>.



### Chapter 3

HPLC & MALDI-TOF of TBA-7T<sup>trans</sup>.



### 3B.12 References

1. (a) Ray, A.; Norden, B., *FASEB J.*, **2000**, *14*, 1041–1060; (b) Braasch, D. A.; Corey, D. R., *Chem. Biol.*, **2001**, *8*, 1–7; (c) Heasman, J., *Dev., Biol.*, **2002**, *243*, 209–214; (d) Dias, N.; Stein, C. A., *Mol. Cancer Ther.*, **2002**, *1*, 347–355; (e) Shukla, S.; Sumaria, C. S.; Pradeepkumar, P. I., *Chem. Med. Chem*, **2010**, *5*, 328–349. (a) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R., *Cell* **1982**, *31*, 147–157. (b) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., *Cell* **1983**, *35*, 849–857.
2. (a) Pasternak, A., Wengel, J., *Bioorg. Med. Chem. Lett.*, **2011**, *21*, 752–755; (b) Jensen, B.; Henriksen, J. R.; Rasmussen, B.; Rasmussen, M.; Andresen, T. L.; Wengel, J.; Pasternak, A., *Bioorg. Med. Chem.*, **2011**, *19*, 4739–4745; (c) Kumar, N.; Nielsen, J. T.; Maiti, S.; Petersen, M., *Angew. Chem., Int. Ed.*, **2007**, *46*, 9220–9222; (d) Datta, B.; Schmitt, C.; Armitage, B. A., *J. Am. Chem. Soc.*, **2003**, *125*, 4111–4118; (e) Datta, B.; Bier, M. E.; Roy, S.; Armitage, B. A., *J. Am. Chem. Soc.*, **2005**, *127*, 4199–4207; (f) Paul, A.; Sengupta, P.; Krishnan Y.; Ladame, S.; *Chem. Eur. J.*, **2008**, *14*, 8682–8689.
3. (a) De Bouvere, B.; Kerreinans, L.; Hendrix, C.; De Winter, H.; Schepers, G.; Van Aerschot, A.; Herdewijn, P., *Nucleosides Nucleotides*, **1997**, *16*, 973–976; (b) Eschenmoser, A., *Science*, **1999**, *284*, 2118–2124; (c) Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; Van Aerschot, A.; Herdewijn, P., *J. Am. Chem. Soc.*, **2000**, *122*, 8595–8602; (d) Pinheiro, V. B.; Holliger, P.; *Curr. Opin. Chem. Biol.*, **2012**, *16*, 245–252.
4. Schaeffer, H. J.; Beauchamp, L.; de Miranda, P.; Elion, G. B.; Bauer, D. J.; Collins, P., *Nature* **1978**, *272*, 583.
5. Schneider, K. C.; Benner, S. A., *J. Am. Chem. Soc.* **1990**, *112*, 453.
6. Nielsen, P.; Dreieø, L.H.; Wengel, J., *Bioorg. Med. Chem. Lett.* **1995**, *3*(1), 19–28.
7. Zhang, L.; Peritz, A.; Meggers, E., *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175.

### Chapter 3

---

8. Karri, P.; Punna, V.; Kim, K.; Krishnamurthy, R., *Angew. Chem. Int. Ed.* **2013**, 52, 5840-5844.
9. Joyce, G. F.; Schwartz, A.W.; Miller, S. L.; Orgel, L. E., *Proc. Natl. Acad. Sci. U.S.A.* **1987**, 84, 4398.
10. Zhanga, S.; Switzer, C.; Chaput, J. C., *Chemistry & Biodiversity* **2010**, 7, 245-258.
11. Boesen, T.; Pedersen, D. S.; Nielsen, B. M.; Petersen, A. B.; Henriksen, U.; Dahl, B. M.; Dahl, O., *Bio.Org Med Chem. Lett.* **2003**, 13, 847-850.
12. a) Hein, J. E.; Blackmond, D. G., *Acc. Chem. Res.* **2012**, 45,12, 2045-2054. b) Trump, J. E. V.; Miller, S. L.; *Science* **1972**, 178, 859-869. c) Cleaves, H., *Monatshefte für Chemie*, **2003**,134, 585–593.
13. Schrger, J.; Welzel, P., *Tetrahedron* **1994**, 50, 6839-6858.
14. Krutzfeldt, J.; Rajewsky, N.; Braich, R.; Rajeev, K. G.; Tuschl, T.; Manoharan, M.; Stoffel, M., *Nature* **2005**, 438, (7068), 685-689.
15. Campbell, J. M.; Bacon, T. A.; Wickstrom, E., *Journal of Biochemical and Biophysical Methods* **1990**, 20, (3), 259-267.
16. Freier, S. M.; Altmann, K. H., *Nucleic Acids Research* **1997**, 25, (22), 4429-4443.
17. Lesnik, E. A.; Guinosso, C. J.; Kawasaki, A. M.; Sasmor, H.; Zounes, M.; Cummins, L. L.; Ecker, D. J.; Cook, P. D.; Freier, S. M., *Biochemistry* **1993**, 32, (30), 7832-7838.
18. (a) Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.; In, Y.; Ishida, T.; Imanishi, T., *Tetrahedron Letters* **1997**, 38, (50), 8735-8738. (b) Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.; Morio, K.; Doi, T.; Imanishi, T., *Tetrahedron Letters* **1998**, 39, (30), 5401-5404.

### Chapter 3

---

19. Morita, K.; Takagi, M.; Hasegawa, C.; Kaneko, M.; Tsutsumi, S.; Sone, J.; Ishikawa, T.; Imanishi, T.; Koizumi, M., *Bioorganic & Medicinal Chemistry* **2003**, 11, (10), 2211-2226.
20. (a) Johnson, J.E.; Smith, J.S.; Kozak, M.L.; Johnson, F.B.; *Biochimie*, **2008**, 90, 1250–1263. (b) Burge, S.; Parkinson, G.N.; Hazel, P.; Todd, A.K.; Neidle, S.; *Nucleic Acids Res.*, **2006**, 34, 5402–5415. (c) Siddiqui-Jain, A.; Grand, C.L.; Bearss, D.J.; Hurley, L.H.; *Proc. Natl Acad. Sci. USA*, **2002**, 99, 11593–11598, (d) Huppert, J.L.; Balasubramanian, S. *Nucleic Acids Res.*, **2005**, 33, 2908–2916 (e) Xu, Y., Kaminaga, K., Komiyama, M.; *J. Am. Chem. Soc.*, **2008**, 130, 11179–11184, (f) Huppert, J.L.; *Chem. Soc. Rev.*, **2009**, 37, 1375–1384.
21. (a) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S.; *Nat. Chem*, **2013**, 5, 182-186. (b) Lam, E Yi Ni.; Beraldi, D.; Tannahill, D.; Balasubramanian, S., *Nature Communications*, **2013**, 4, article no:1796.
22. (a) Bates, P.J.; Kahlon, J.B.; Thomas, S.D.; Trent, J.O.; Miller, D.M.; *J. Biol. Chem.*; **1999**, 274, 26369–26377. (b) Mashima, T.; Matsugami, A.; Nishikawa, F.; Nishikawa, S.; Katahira, M., *Nucleic Acids Res.*, **2009**, 37, 6249–6258. (c) Huppert, J.L., *Philos. Trans. Roy. Soc. Ser. A*, 2007, 365, 2969–2984. (d) Balasubramanian, S.; Neidle, S., *Curr. Opin. Chem. Biol.*, **2009**, 13, 345–353. (e) Tian-miao, O.; Yu-jing, L.; Jia-heng, T.; Zhi-shu, H.; Kwok-Yin, W.; Lian-quan, G.; *Chem. Med. Chem.*, **2008**, 3, 690–713.
23. Nimjee, S.M.; Rusconi, C.P.; Sullenger, B.A.; *Annu. Rev. Med.*, **2005**, 56, 555-583.
24. (a) W. James in *Aptamers in Encyclopedia of Analytical Chemistry* (Ed.: R. A. Meyers), Wiley, Chichester, UK, **2000**, 4848-4871.; (b) *The Aptamer Handbook*. Edited by S. Klussmann. (2006) WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.
25. Robertson, D. L.; Joyce, G. F., *Nature* **1990**, 344, 467-468.
26. Ellington, A. D.; Szostak, J. W., *Nature* **1990**, 346, 818-822.
27. Tuerk, C.; Gold, L., *Science* **1990**, 249, 505-510.

### Chapter 3

---

28. Coughlin, SR., *Nature*, **2000**, 407, 258–64.
29. Padmanabhan, K.; Tulinsky, A., *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **1996** 52, 272–282.
30. Kelly, J.A.; Feigon, J.; Yeates, T.O., *J. Mol. Biol.*, **1996**, 256, 417–422.
31. Veronica, E.; Maria S.; Antonella, C.; Rita, S.; Michela, V, L.; Antonella, V.; Aldo G., *Org. Biomol. Chem.*, **2014**, 12, 8840–8843.
32. Pasternak, A.; Hernandez, F. J.; Rasmussen, L. M.; Vester, B.; Wengel, J., *Nucleic Acids Res.*, **2010**, 39, 1155-1164.
33. Bonifacio, L.; Church, F. C.; Jarstfer, M. B., *Int. J. Mol. Sci.*, **2008**, 9, 422-433.
34. Aviñó, A.; Mazzini, S.; Ferreira, R.; Gargallo, R.; Marquez, V.E.; Eritja, R.; *Bioorg. Med. Chem.*, **2012**, 20, 4186-4193.
35. Campbell, M. A.; Wengel, J., *Chem. Soc. Rev.* **2011**, 40, 5680-5689.
36. Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J., *Nature*, **1992**, 355, 564-566.
37. Karsisiotis, A. I.; Hessari, N. M.; Novellino, E.; Spada, G. P.; Randazzo, A.; da Silva M. W., *Angew. Chem. Int. Ed.*, **2011**, 50, 10645 –10648.



## CHAPTER 4 / SECTION-A

Design, synthesis and biophysical  
evaluation of 4'-Methoxymethyl  
threose nucleic acids

## **Section 4A Design, synthesis and biophysical evaluation of 4'-Methoxymethyl threose oligonucleotides (4'MOM-TNA)**

### **4A.1 Introduction**

$\alpha$ -L-threose nucleic acid (TNA),<sup>1</sup> was the first synthetic genetic polymer discovered by Eschenmoser and co-workers. Natural five-carbon ribose sugar found in RNA was replaced with an unnatural four-carbon threose sugar in TNA. The successive nucleosidic units were connected through 2' and 3' vicinal phosphodiester linkages (Figure 1). Though the repeating backbone unit shortened by one-atom, TNA still underwent informational Watson-Crick base pairing in an antiparallel strand orientation and also cross-paired with complementary strands of DNA and RNA.<sup>1</sup> The chemical simplicity of threose relative to ribose and ability of TNA to cross-pair with RNA provoke the careful consideration of TNA as a possible RNA progenitor.<sup>2</sup> Using the hairpin strategy,<sup>3</sup> it was also demonstrated that after several rounds of *in vitro* selection and amplification, TNA aptamers were evolved with high binding affinity and specificity to human thrombin. This demonstration also showed the TNA's ability to fold into tertiary structures with sophisticated chemical functions.<sup>4</sup>

### **4A.2 Synthesis of threose nucleosides**

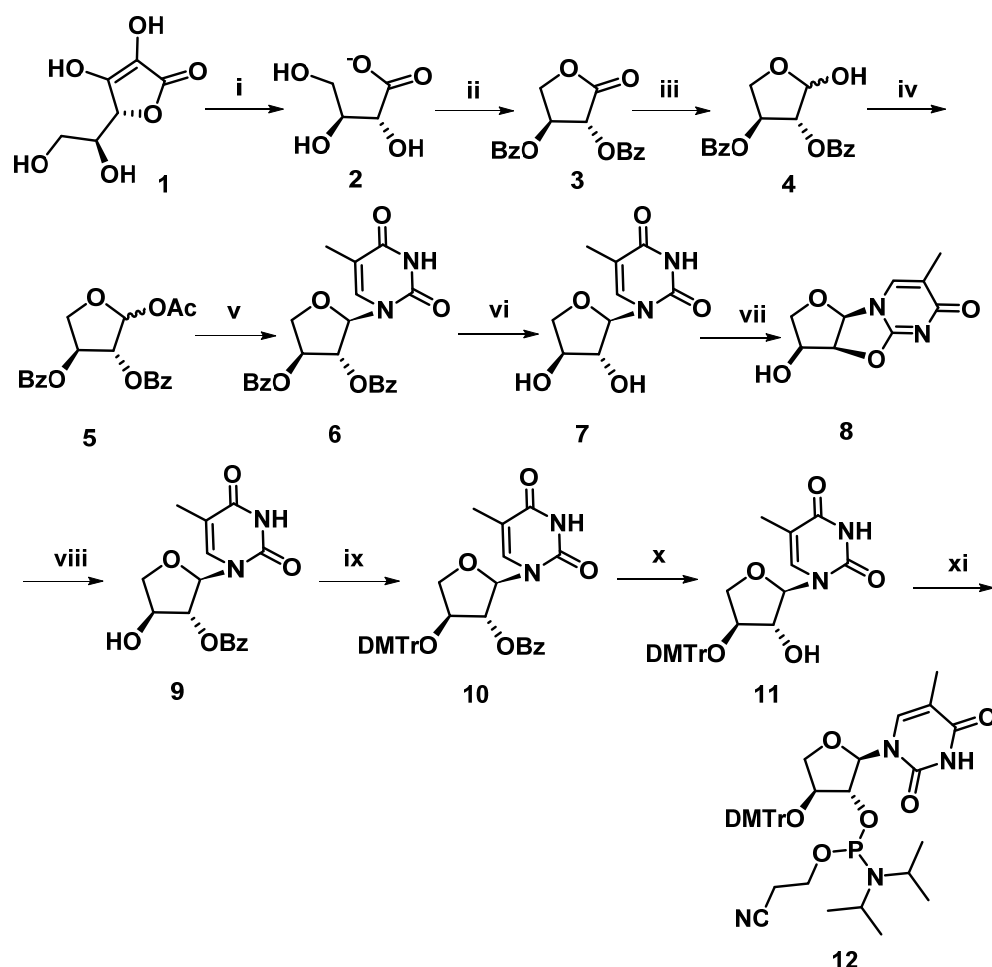
The oligonucleotide chemist required enormous amounts of the monomers to evaluate the biophysical properties of the modified nucleotides. In case of TNA the access of the tetrose sugar L-Threose was less. Although the number of methods were reported for synthesis of L-Threose in the literature, none of them seemed amenable to large scale preparation, either due to expensive starting materials, laborious workup, or poor yields.<sup>5</sup>

Among all of the synthetic routes, an improved method was introduced by Albert Eschenmoser *et al.*<sup>6</sup> started from L-ascorbic acid. **1** was converted into calcium salt of threonic acid **2** by using reported procedures. A one-pot lactonization by treatment with a Dowex resin (H<sup>+</sup> form) yielded threonolactone, which was benzoylated *in situ* with BzCl to give the 2,3-di-O-benzoyl-L-threonolactone **3** in 60% yield. Its reduction with DIBAL-H at -78 °C afforded the 2',3'-di-O-benzoylated lactol **4** as a mixture of  $\alpha/\beta$  isomers in 53% yield. Acetylation of anomeric alcohol yields **5**, which was further subjected for *Vorbruggen-Hilbert-Johnson* nucleosidation to introduce thymine nucleobases **6**. Compound **6** was treated with methanolic ammonia to obtain dihydroxy

## Chapter 4

compound **7**. Next aim was to do selective DMTr protection of 3'-hydroxyl group which was not easy task. **7** was converted to its corresponding 2'- anhydro nucleoside, upon heating with HMPA in presence of sodiumbenzoate yields **9** in 55% yield. DMTr protection of 3'-OH yields **10** followed by benzoate hydrolysis gave **11** which was converted to amidite derivative **12** (scheme 1). The yields were moderate for most of the steps in the synthetic route which is practically difficult for further oligonucleotide synthesis.

**Scheme 1** Synthesis of Threofurnosyl thymine nucleoside from L-Ascorbic acid



**Reagents and conditions** (i) 30% aq H<sub>2</sub>O<sub>2</sub>, CaCO<sub>3</sub>/H<sub>2</sub>O, 15-18 °C, 24h, 65% (ii) Dowex resin, H<sub>2</sub>O, 80 °C, 30min, TsOH, MeCN, reflux, 24h, BzCl, Pyridine, rt, 16h, 60% (iii) DIBAL-H, THF, -78 °C, 10h, 53% (iv) BzCl, pyridine/DCM (1:1), rt, 18h, 70% (v) Thymine, ACN, BSA, 70 °C, TMS-OTf, 0 °C, reflux, 1h, 64% (vi) NH<sub>3</sub> in MeOH, r.t. / 18 h, 77% (vii) Diphenylcarbonate, NaHCO<sub>3</sub>, HMPA / 150° / 1-2 h, 66% (viii) BzONa, BzOH, HMPA / 150° / 2.5 h, 55% (ix) DMTrCl, 2,4,6-collidine DCM / r.t. / 24 h (x) NH<sub>3</sub> in MeOH, r.t. / 25 h, quantitative yield, (xi) (i-Pr)<sub>2</sub>NEt, DCM / r.t. / 1.5-18 h, 70%

### 4A.3 Design of 4'-Methoxymethyl substituted $\alpha$ -L-threose nucleic acid and rationale

We designed a straight forward synthetic route with commercially available cheap starting material D-xylose, which will lead us to 4'-Methoxymethyl modified TNA nucleoside. The design of modified TNA was not only for simplifying the synthesis but also we could see the hydration effect due to the methoxymethyl group (Figure1).

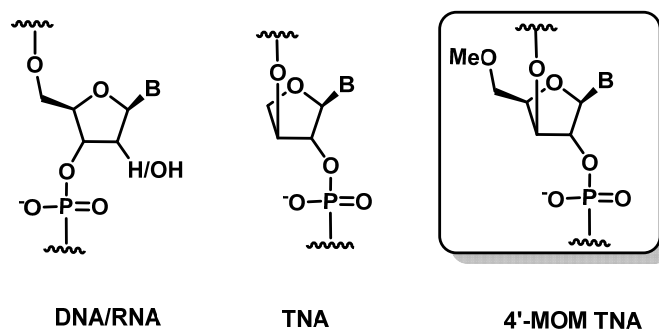
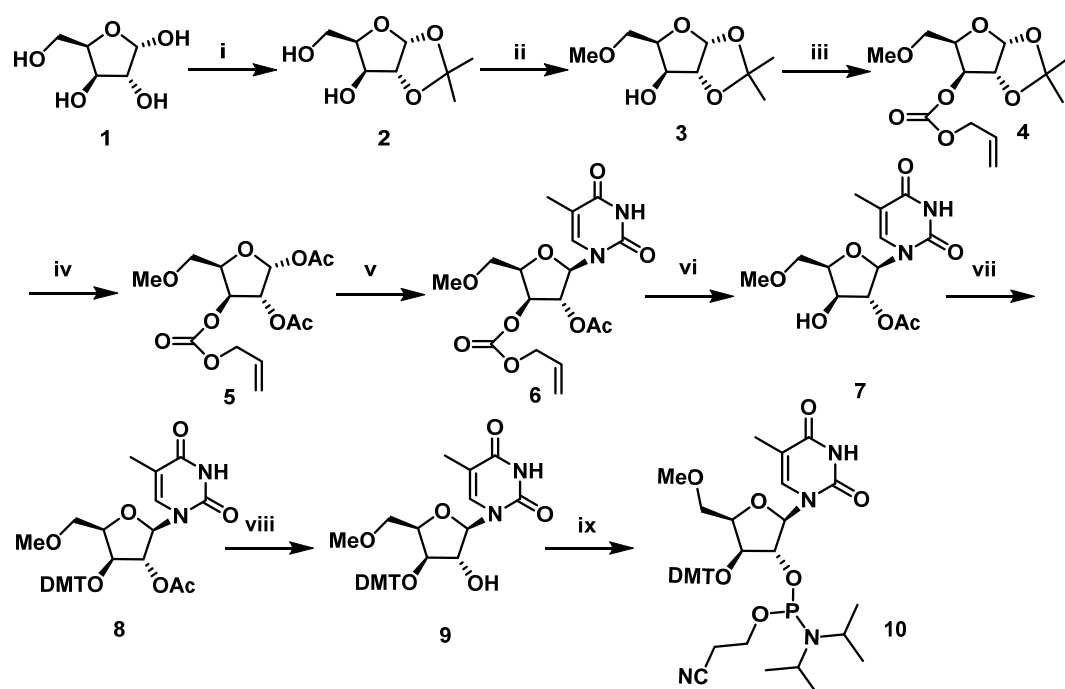


Figure 1 Structures of natural DNA/RNA, TNA and proposed 4'-MOM TNA

### 4A.4 Synthesis of 4'-Methoxymethyl TNA monomers

The synthesis started with commercially available cheap starting material. D-xylose **1** was transformed to 1,2-*O*-isopropylidene- $\alpha$ -D-xylofuranose **2** using acetone, conc.H<sub>2</sub>SO<sub>4</sub> and Na<sub>2</sub>CO<sub>3</sub> in one pot reaction. Mono methylation of primary hydroxyl group in presence of methyl iodide and silver oxide yielded **3**. The secondary hydroxyl group was protected with allyloxycarbonyl group to give **4** in very good yield. The acetonide group in **4** was removed and converted into its diacetate **5** by treatment with AcOH and Ac<sub>2</sub>O in presence of catalytic amount of H<sub>2</sub>SO<sub>4</sub>. Compound **5** on treatment with BSA, thymine and TMSOTf under Vorbrüggen conditions afforded exclusively the  $\beta$ -anomer of thymine derivative **6**, in good yield. The alloc group was selectively cleaved using Pd(0) to get **7**. The free 3'-hydroxyl group was protected as its DMT derivative using DMTr-Cl in dry DCM and 2,4,6-collidine used as base, to get **8**. Compound **8** on ammonolysis gave the free 2'-hydroxyl compound **9**. Phosphitylation of the free 2'-hydroxyl with *N,N*-diisopropylamino-2-cyanoethylphosphino-chloridite afforded the phosphoramidite monomer **10** (Scheme 2).

Scheme 2 Synthesis of 4'-MOM Threofurnosyl thymine nucleoside



**Reagents and conditions** (i) Acetone, conc  $\text{H}_2\text{SO}_4$ ,  $\text{Na}_2\text{CO}_3$  (ii) MeI,  $\text{Ag}_2\text{O}$ , dry ACN (iii) Alloc-Cl, dry Pyridine, dry DCM, rt, 3h (iv) AcOH:  $\text{Ac}_2\text{O}$ :  $\text{H}_2\text{SO}_4$ (10:1:0.1), rt, overnight (v) Thymine, ACN, BSA, 70  $^\circ\text{C}$ , TMS-OTf, 0  $^\circ\text{C}$ , reflux, 3h (vi)  $\text{PPh}_3$ ,  $\text{Pd}(\text{dba})_2$ , Piperidine, DCM, rt, 15min (vii) DMTrCl, 2,4,6-collidine, DCM, rt, 24h (viii) 2-cyanoethyl-N,N-diisopropylchlorophosphine, DIPEA, dry DCM, rt, 1h.

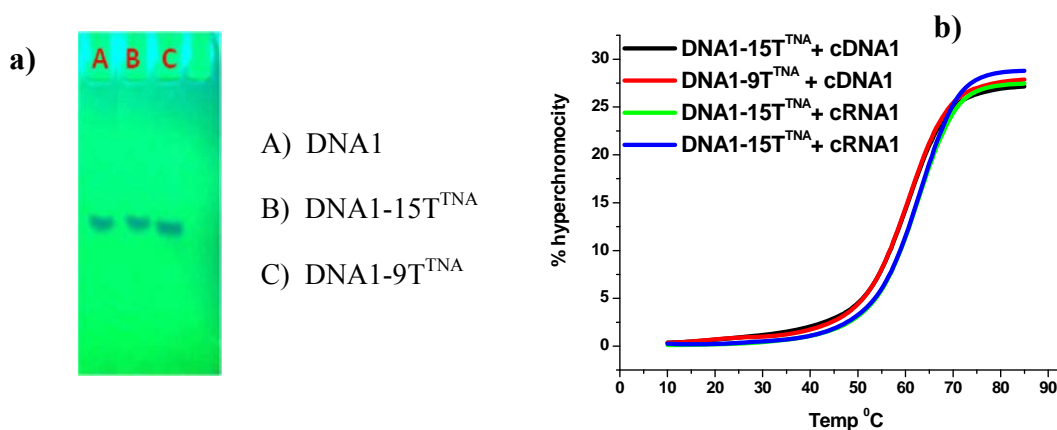
#### 4A.5 Synthesis of modified oligonucleotides, characterization, UV-melting studies

As we discussed in the chapter 3 we chose 18mer DNA sequence and incorporated the modified monomer **10** to investigate the stability offered by the 4'-MOM TNA. Modified oligonucleotides were synthesized on Bioautomation MM4 DNA synthesizer by using regular protocol with increasing the coupling time as 6min. We synthesized two different oligonucleotide sequences containing the modifications at 9<sup>th</sup> and 15<sup>th</sup> position. The crude oligonucleotide was purified by RP-HPLC, purity confirmed by gel-electrophoretic mobility studies (Figure 2a) and characterized by MALDI-TOF mass spectrometry (Table 1).

**Table 1** Modified DNA sequences, their MALDI-TOF mass analyses and biophysical evaluation by UV- $T_m$  measurements<sup>a</sup>

Name	Sequence <sup>b</sup> 5' → 3'	mass cal /obs	UV $T_m$ °C	
			DNA <sup>c</sup>	RNA <sup>d</sup>
DNA1	caccattgtcacactcca	5363/5367	63.5	62.7
DNA1-15T <sup>TNA</sup>	caccattgtcacacT <sup>TNA</sup> cca	5393/5393	59.8	61.9
DNA1-9T <sup>TNA</sup>	caccattgT <sup>TNA</sup> cacactcca	5393/5387	60	62.2

<sup>a</sup>UV- $T_m$  values were measured by using 1 $\mu$ M sequences with 1 $\mu$ M cDNA/cRNA in sodium phosphate buffer (0.01M, pH 7.2) containing 150 mM NaCl and are averages of three independent experiments. (Accuracy is  $\pm 0.5$  °C). <sup>b</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site. <sup>c</sup>5'tggagtgtgacaatggg was the complementary DNA sequence. <sup>d</sup>5' uggagugugacauggug was the complementary RNA sequence.

**Figure 2** (a) Gel pictures of purified sequences and (b) UV melting profiles of modified DNA sequences with cDNA/cRNA

It is observed that single modified sequences are forming the stable duplex structures with both complementary DNA as well as complementary RNA. The duplex stability of modified sequences were as good as control. Modified TNA:RNA complexes are more stable than modified TNA:DNA complexes. Both the modified sequences are showing similar stability independent to the position of modification (Figure 2b & Table 1).

## 4A.6 Experimental Section

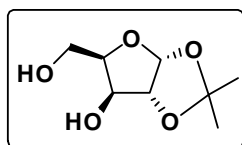
**General remarks:** All the reagents were purchased from Sigma-Aldrich and used without further purification. DMF, ACN, were dried over P<sub>2</sub>O<sub>5</sub> and CaH<sub>2</sub> respectively and stored by adding 4 Å molecular sieves. Pyridine, TEA were dried over KOH and stored on KOH. THF was passed over basic alumina and dried by distillation over sodium metal. Reactions were monitored by TLC. TLCs were run in either Petroleum ether with appropriate quantity of EtOAc or DCM with an appropriate quantity of MeOH for most of the compounds. TLC plates were visualized with UV light and iodine spray and/or by spraying perchloric acid solution and heating. Usual reaction work up involved sequential washing of the organic extract with water and brine followed by drying over anhydrous Na<sub>2</sub>SO<sub>4</sub> and evaporation of the solvent under vacuum. Column chromatographic separations were performed using silica gel 60-120 mesh (Merck) or 200- 400 mesh (Merck) and using the solvent systems EtOAc/Petroleum ether or MeOH/DCM. TLC was run using pre-coated silica gel GF254 sheets (Merck 5554).

<sup>1</sup>H and <sup>13</sup>C NMR spectra were obtained using Bruker AC-200, AC-400 and AC-500 NMR spectrometers. The chemical shifts (δ/ppm) are referred to internal TMS/DMSO-d<sub>6</sub> for <sup>1</sup>H and chloroform-*d*/DMSO-d<sub>6</sub> for <sup>13</sup>C NMR. <sup>1</sup>H NMR data are reported in the order of chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; br, broad; br s, broad singlet; m, multiplet and/ or multiple resonance), number of protons. Mass spectra were recorded on APQSTAR spectrometer, LC-MS on a Finnigan-Matt instrument. DNA oligomers were synthesized on CPG solid support using Bioautomation MerMade 4 synthesizer. The RNA oligonucleotides were obtained commercially (Sigma-Aldrich). RP-HPLC was carried out on a C18 column using either a Varian system (Analytical semi-preparative system consisting of Varian Prostar 210 binary solvent delivery system, a Dynamax UV-D2 variable wavelength detector and Star chromatography software) or a Waters system (Waters Delta 600e quaternary solvent delivery system with 2998 photodiode array detector and Empower2 chromatography software). MALDI-TOF spectra were recorded on a Voyager-De-STR (Applied Biosystems) MALDI-TOF instrument or AB Sciex TOF/TOF™ Series Explorer™ 72085 instrument and the matrix used for analysis was THAP (2', 4', 6'-trihydroxyacetophenone). UV experiments were performed on a Varian Cary 300 UV-VIS spectrophotometer fitted with a Peltier-controlled temperature programmer. CD

spectra were recorded on a Jasco J-715 Spectropolarimeter, with a ThermoHaake K20 programmable water circulator for temperature control of the sample.

### Experimental procedures and spectral data

#### 1,2-*O*-isopropylidene- $\alpha$ -D-xylofuranose (**2**)

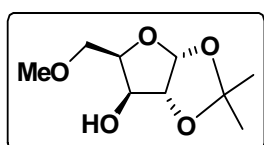


Finely powdered D-xylose compound **1** (10.0g, 67.1mmol) was dissolved in acetone (260mL) containing H<sub>2</sub>SO<sub>4</sub> (10.0mL, 96%, 66.0 mmol) and stirred for 30min. A solution of Na<sub>2</sub>CO<sub>3</sub> (13.0g, 122.65mmol) in H<sub>2</sub>O (112mL) was carefully added under external cooling so as to keep the temperature of the mixture at 20°C, and the mixture was stirred for a further 2.5h. Then, solid Na<sub>2</sub>CO<sub>3</sub> (7g, 66.0 mmol) was added till the pH = 7.0, Na<sub>2</sub>SO<sub>4</sub> was filtered off and washed with acetone, and the combined filtrates were evaporated to yield 13.8g of crude **2** contaminated with 5% of 1,2:3,5-di-*O*-isopropylidene- $\alpha$ -D-xylofuranose, and 5% of starting D-xylose. The crude **2** was purified by silica gel filtration using 30:1 DCM/MeOH and pure **2** was obtained as syrup, which crystallized on standing. Yield 11.36g, 93%.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  1.23 (s, 3 H), 1.40 (s, 3 H), 3.80 - 3.92 (m, 2 H), 4.04 - 4.15 (m, 2 H), 4.19 (br. s., 1 H), 4.44 (d, *J*=3.51 Hz, 1 H), 4.64 (d, *J*=4.27 Hz, 1 H), 5.88 (d, *J*=3.51 Hz, 1 H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  26.1, 26.6, 60.3, 75.7, 76.9, 77.2, 77.5, 79.6, 85.3, 104.7, 111.7 ppm. HRMS: mass calculated for C<sub>8</sub>H<sub>14</sub>O<sub>5</sub>Na (M+Na)<sup>+</sup> 213.0734, observed (M+ Na)<sup>+</sup> 213.0733.

#### 5-*O*-methyl-1,2-*O*-isopropylidene- $\alpha$ -D-xylofuranose (**3**)

To a solution of **2** (1g, 5.2 mmol) and CH<sub>3</sub>I (0.46 mL, 7.8 mmol) in 20 mL acetonitrile, Ag<sub>2</sub>O(1.4g, 6.2 mmol) was added and the reaction mixture was vigorously stirred at room temperature for 12h. The reaction mixture was filtered and the filtrate was concentrated under reduced pressure. The crude compound was purified by column chromatography (eluted in 50% EtOAc in petroleum ether) to give **3** (0.68 g) as a white solid in 64% yield.

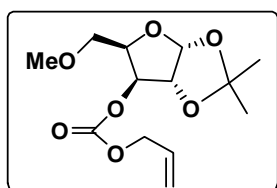


<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>)  $\delta$  1.32 (s, 3 H), 1.49 (s, 3 H), 3.44 (s, 3 H), 3.79 - 3.85 (m, 1 H), 3.89 (t, *J*=3.60 Hz, 2 H), 4.18 - 4.25 (m, 1 H), 4.29 (t, *J*=2.84 Hz, 1 H), 4.52 (d,



$J=3.66$  Hz, 1 H), 5.98 (d,  $J=3.79$  Hz, 1 H) ppm;  $^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ )  $\delta$  26.1, 26.7, 59.9, 71.0, 76.4, 76.6, 85.4, 104.8, 111.6 ppm; HRMS: mass calculated for  $\text{C}_9\text{H}_{16}\text{O}_5\text{Na}$  ( $\text{M}+\text{Na}$ ) $^+$  227.0889, observed ( $\text{M}+\text{Na}$ ) $^+$  227.0890.

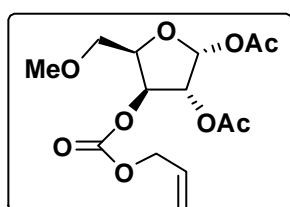
#### 5-*O*-methyl-3-allyloxy 1,2-*O*-isopropylidene- $\alpha$ -D-xylofuranose (4)



Compound **3** (0.8 g, 3.92 mmol) was dissolved in dry dichloromethane (10 mL). Anhydrous pyridine (0.75 mL) was added and the reaction mixture was cooled to 0°C in an ice-bath. Allyloxycarbonyl chloride (0.47 mL, 4.71 mmol) was added dropwise and then the reaction was stirred at room temperature for 3 hrs. TLC showed absence of starting compound. The reaction mixture was extracted with  $\text{CH}_2\text{Cl}_2$ , followed by water wash and drying over sodium sulphate. Removal of solvent yielded a sticky gum having compound **4**, which was purified by column chromatography (eluted in 25% EtOAc in petroleum ether) to give **4** (0.64 g) as a colourless thick liquid in 70% yield.

$^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ )  $\delta$  1.31 (s, 3 H), 1.51 (s, 3 H), 3.38 (s, 3 H), 3.59 - 3.66 (m, 2 H), 4.44 (td,  $J=5.75, 3.03$  Hz, 1 H), 4.59 (d,  $J=3.79$  Hz, 1 H), 4.65 (dt,  $J=5.81, 1.26$  Hz, 2 H), 5.11 (d,  $J=3.03$  Hz, 1 H), 5.26 - 5.44 (m, 2 H), 5.81 - 6.06 (m, 2 H) ppm;  $^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ )  $\delta$  26.3, 26.7, 59.4, 69.0, 69.5, 76.4, 79.9, 83.2, 104.7, 112.2, 119.5, 131.2, 154.0 ppm. HRMS: mass calculated for  $\text{C}_{13}\text{H}_{20}\text{O}_7\text{Na}$  ( $\text{M}+\text{Na}$ ) $^+$  311.1101, observed ( $\text{M}+\text{Na}$ ) $^+$  311.1101.

#### 5-*O*-methyl-3-allyloxy 1,2-di-*O*-acetyl - $\alpha$ -D-xylofuranose (5)

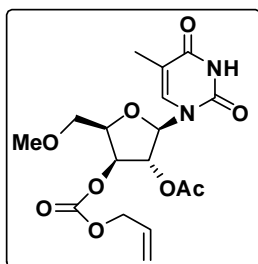


Compound **4** was desiccated (1.20 g, 3.6 mmol) and dissolved in acetic acid (16 mL). Acetic anhydride (1.6 mL) was added, after cooling the reaction flask to 10°C, followed by dropwise and slow addition of concentrated sulfuric acid (0.16 mL). The reaction mixture was stirred overnight at room temperature. TLC indicated complete product formation. The reaction was quenched with ice and 5% aqueous  $\text{NaHCO}_3$ , then extracted with  $\text{CH}_2\text{Cl}_2$ , followed by water wash and drying over sodium sulphate. After solvent removal the crude product was purified by silica gel column chromatography

using petroleum ether and ethyl acetate as eluants. The compound **5** was eluted in 25% ethyl acetate in petroleum ether yield (1.03 g), 72%.

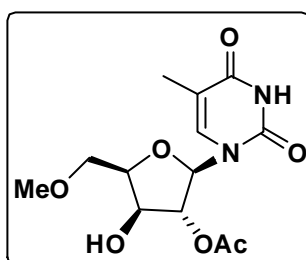
**<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>)** δ 2.07 - 2.15 (m, 6 H), 3.32 - 3.44 (m, 3 H), 3.48 - 3.66 (m, 2 H), 4.52 - 4.73 (m, 3 H), 5.23 - 5.47 (m, 4 H), 5.82 - 6.09 (m, 1 H), 6.44 (d, *J*=4.55 Hz, 1 H) ppm; **<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>)** δ 20.38, 20.64, 20.91, 21.12, 29.69, 59.30, 59.43, 69.17, 70.06, 70.46, 75.49, 76.38, 76.46, 77.09, 77.64, 77.73, 77.92, 79.82, 80.48, 92.84, 98.81, 119.51, 131.08, 131.16, 153.96, 154.31, 169.17, 169.30, 169.59 ppm; **HRMS:** mass calculated for C<sub>14</sub>H<sub>21</sub>O<sub>9</sub> (M+ H)<sup>+</sup> 333.1161, observed (M+ H)<sup>+</sup> 333.1180.

#### 5-*O*-methyl-3-allyloxy 2-*O*-acetyl thymidine (**6**)



Compound **5** (0.55 g, 1.65 mmol), obtained from the previous step was dissolved in anhydrous acetonitrile (10 mL). Reaction flask was flushed with nitrogen and thymine (0.15 g, 1.2 mmol) was added. N,O-Bis(trimethylsilyl)acetamide (BSA) (0.81 mL, 3.3 mmol) was added to the reaction flask under nitrogen atmosphere. Then the reaction mixture was refluxed at 70°C for one hour, followed by cooling in an ice bath. TMSOTf (0.9 mL, 4.95 mmol) was added slowly with a syringe and the reaction mixture was refluxed for three hours. TLC showed disappearance of starting material and appearance of a lower moving UV-positive spot which charred on acid spraying and heating. The reaction mixture was cooled to room temperature, diluted with dichloromethane, washed with NaHCO<sub>3</sub> and water, dried over sodium sulphate followed by solvent removal. The crude product was purified by silica gel column chromatography. Compound **6** eluted in 70% EtOAc in petroleum ether. Yield: 0.53 g, 81%.

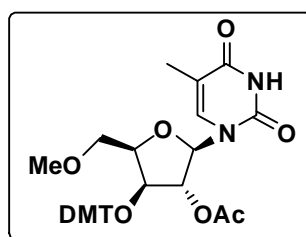
**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** δ 1.93 (s, 3 H), 2.16 (s, 3 H), 3.43 (s, 3 H), 3.68 - 3.84 (m, 2 H), 4.43 (d, *J*=4.02 Hz, 1 H), 4.69 (d, *J*=5.52 Hz, 2 H), 5.25 (br. s., 2 H), 5.31 - 5.45 (m, 2 H), 5.87 - 6.03 (m, 1 H), 6.06 (br. s., 1 H), 7.54 (br. s., 1 H) ppm; **<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)** δ 15.48, 23.42, 51.20, 51.41, 51.63, 51.84, 52.05, 52.27, 52.48, 62.29, 72.45, 72.49, 80.05, 80.37, 80.69, 81.17, 82.19, 82.43, 90.45, 114.43, 122.62, 133.91, 138.96, 153.76, 156.75, 172.75 ppm. **HRMS:** mass calculated for C<sub>17</sub>H<sub>23</sub>O<sub>9</sub>N<sub>2</sub> (M+ H)<sup>+</sup> 399.1395, observed (M+ H)<sup>+</sup> 399.1398.

**5'-O-methyl-3'-hydroxy-2'-O-acetyl thymidine (7)**

Compound **6** (1 g, 2.51 mmol) was dissolved in dichloromethane (25 mL). PPh<sub>3</sub> (0.54 g, 2.06 mmol) was added, followed by piperidine (2.0 mL, 0.02 mmol) and tris(dibenzylidene acetone) dipalladium [Pd<sub>2</sub>(dba)<sub>3</sub>] (0.15 g, 0.16 mmol). The reaction mixture was stirred for 15 minutes.

TLC showed absence of starting compound. Solvent was removed and the crude product was given a wash with solvent ether. Column purification done on a silica gel column yielded the pure compound **7**, which eluted in methanol (3 %) in dichloromethane. Yield: 0.51, 65%.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.91 (s, 3 H), 2.12 (s, 3 H), 3.46 (s, 3 H), 3.79 - 3.91 (m, 2 H), 4.20 (q, *J*=4.02 Hz, 1 H), 4.29 (br. s., 1 H), 4.49 (br. s., 1 H), 5.14 (s, 1 H), 5.79 (d, *J*=2.26 Hz, 1 H), 9.43 (br. s., 1 H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 12.5, 20.7, 59.5, 70.6, 74.7, 76.8, 79.2, 80.4, 81., 82.3, 87.8, 90.2, 111.1, 150.5, 164.1, 170.1 ppm. HRMS: mass calculated for C<sub>13</sub>H<sub>19</sub>O<sub>7</sub>N<sub>2</sub> (M+ H)<sup>+</sup> 315.1184, observed (M+ H)<sup>+</sup> 315.1187.

**5'-O-methyl-3'-dimethoxytrityl-2'-O-acetyl thymidine (8)**

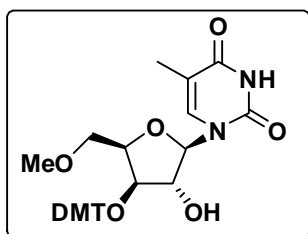
The substrate **7** (0.5 g, 1.6 mmol) dissolved in anhydrous DCM (10 mL). 4, 4'-Dimethoxytritylchloride (1.02 g, 3 mmol) and 2,4,6-collidine (1.3 mL, 9 mmol) was added in one lot. The reaction mixture was stirred 24h at room temperature, when TLC showed a faster moving trityl-positive spot which

charred on acid spraying and heating. The reaction was quenched with methanol, extracted with dichloromethane, washed with NaHCO<sub>3</sub> and water, dried over sodium sulphate, followed by solvent removal. The crude product was purified by silica gel column chromatography using dichloromethane, methanol and TEA (0.5%) as eluants. Compound **8** is eluted in methanol (2 %) in dichloromethane. Yield: 0.74g, 76%.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 1.84 (s, 3 H), 1.97 (s, 3 H), 3.51 (s, 3 H), 3.59 (br. s., 2 H), 3.80 (s, 7 H), 4.09 - 4.16 (m, 1 H), 4.44 - 4.50 (m, 1 H), 4.68 - 4.74 (m, 1 H), 5.78 (d, *J*=4.58 Hz, 1 H), 6.82 - 6.87 (m, 5 H), 7.28 - 7.41 (m, 8 H), 7.46 (d, *J*=8.24 Hz, 5 H), 7.68 (d, *J*=7.32 Hz, 1 H), 7.66 (d, *J*=7.63 Hz, 1 H), 8.16 (br. s., 1 H) ppm; <sup>13</sup>C NMR

(126 MHz, CDCl<sub>3</sub>) δ 12.5, 20.6, 55.3, 59.0, 70.9, 75.0, 76.8, 78.6, 79.1, 85.7, 87.6, 113.4, 127.4, 128.0, 128.1, 128.5, 128.6, 130.4, 135.3, 144.7, 150.2, 159.0, 163.5, 165.9 ppm. **HRMS:** mass calculated for DMTr cleaved fragment C<sub>13</sub>H<sub>18</sub>O<sub>7</sub>N<sub>2</sub>Na (M+ Na)<sup>+</sup> 337.1006, observed (M+ Na)<sup>+</sup> 337.1006.

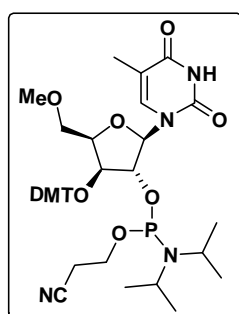
#### 5'-O-methyl-3'-dimethoxytrityl-2'-hydroxy thymidine (9)



The substrate **8** (0.70 g, 1.16 mmol) was dissolved in AR grade methanol (50 mL). Aqueous ammonia (15 mL) was added and the pinkish slightly turbid reaction mixture was stirred for one hour at room temperature. TLC showed the absence of starting compound. The solvents were removed to get a yellowish solid. The solid was redissolved in dichloromethane and given a water wash. The organic layer was dried over sodium sulfate and concentrated to get a pale yellow solid foam. Purification was done by silica gel column chromatography using dichloromethane, methanol and TEA (0.5%) as eluants. Compound **9** eluted in methanol (3%) in dichloromethane. Yield: 0.56g, 87%.

<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 1.84 (s, 3 H), 3.32 - 3.39 (m, 1 H), 3.46 (s, 3 H), 3.52 - 3.60 (m, 1 H), 3.66 - 3.72 (m, 1 H), 3.77 (s, 3 H), 3.75 (s, 3 H), 3.94 - 4.00 (m, 1 H), 4.23 (t, *J*=3.51 Hz, 1 H), 5.51 (d, *J*=2.14 Hz, 1 H), 6.82 (dd, *J*=8.55, 4.88 Hz, 4 H), 7.18 - 7.40 (m, 8 H), 7.41 - 7.50 (m, 3 H), 7.69 (s, 1 H) ppm; <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 12.53, 29.68, 55.25, 59.06, 70.92, 76.77, 78.95, 81.00, 87.39, 90.97, 109.61, 113.42, 127.17, 128.00, 128.05, 130.25, 130, 132.06, 132.14, 135, 144.90, 150.95, 158.91, 164.06 ppm; **HRMS:** mass calculated C<sub>32</sub>H<sub>34</sub>O<sub>8</sub>N<sub>2</sub>Na (M+ Na)<sup>+</sup> 597.2207, observed (M+ Na)<sup>+</sup> 597.2207.

#### 5'-O-methyl-3'-dimethoxytrityl- thymidyl 2'-O-phosphoramidite (10)



Compound **9** (0.08 g, 0.14 mmol) was co-evaporated with dry CH<sub>2</sub>Cl<sub>2</sub>, and then dissolved in dry dichloromethane (3.0 mL). Diisopropylethylamine (DIPEA) (0.1 mL, 0.56 mmol) was added, followed by chloro (2-cyanoethoxy)-N, N-diisopropyl amino)-phosphine (0.06 mL, 0.28 mmol) at 0°C. The reaction mixture was stirred under argon atmosphere at room temperature for 1 hour,

## Chapter 4

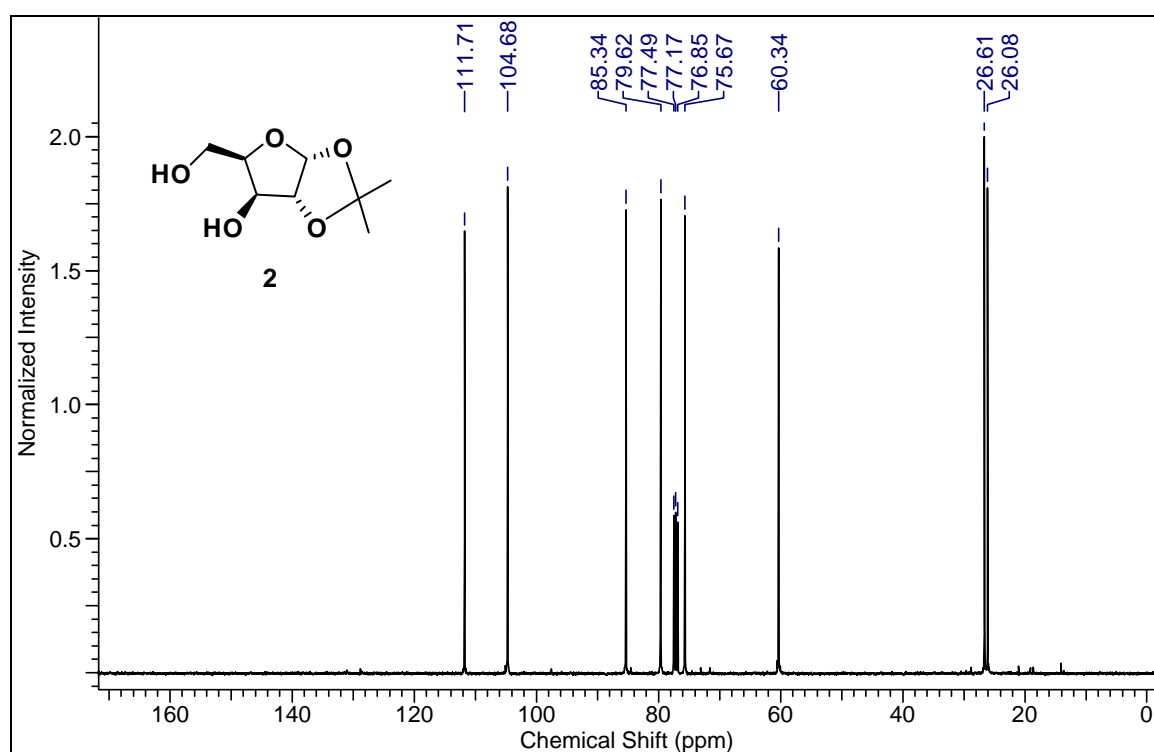
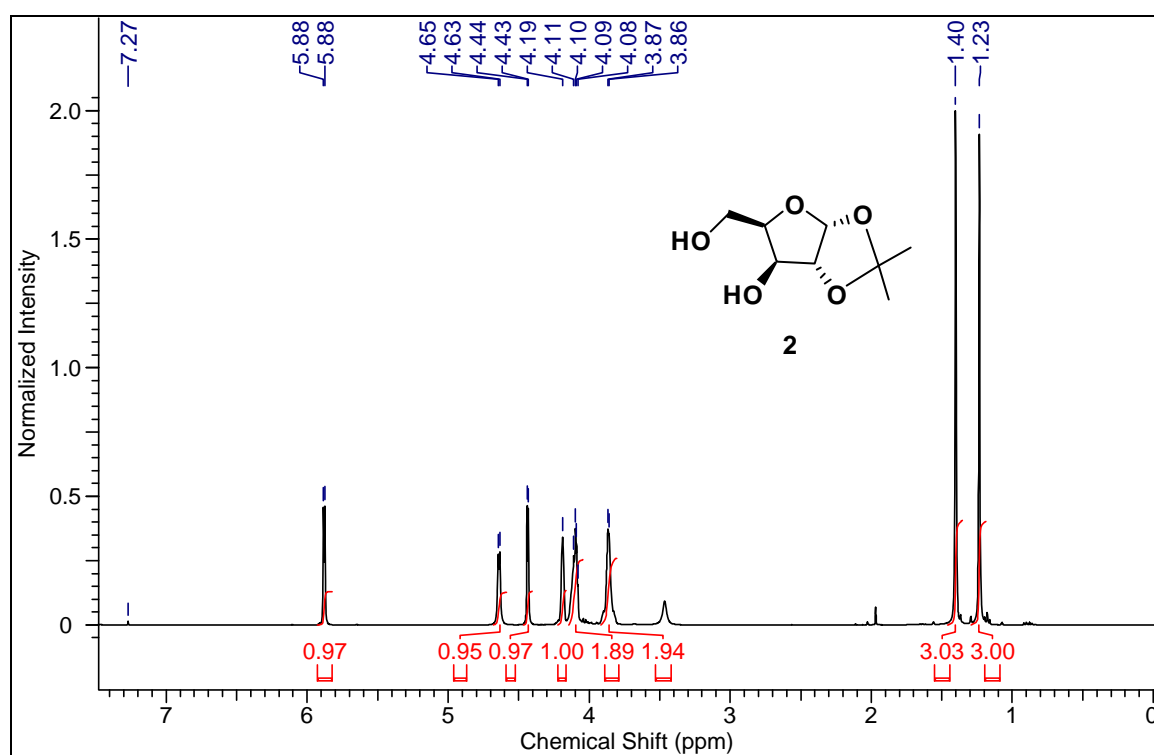
---

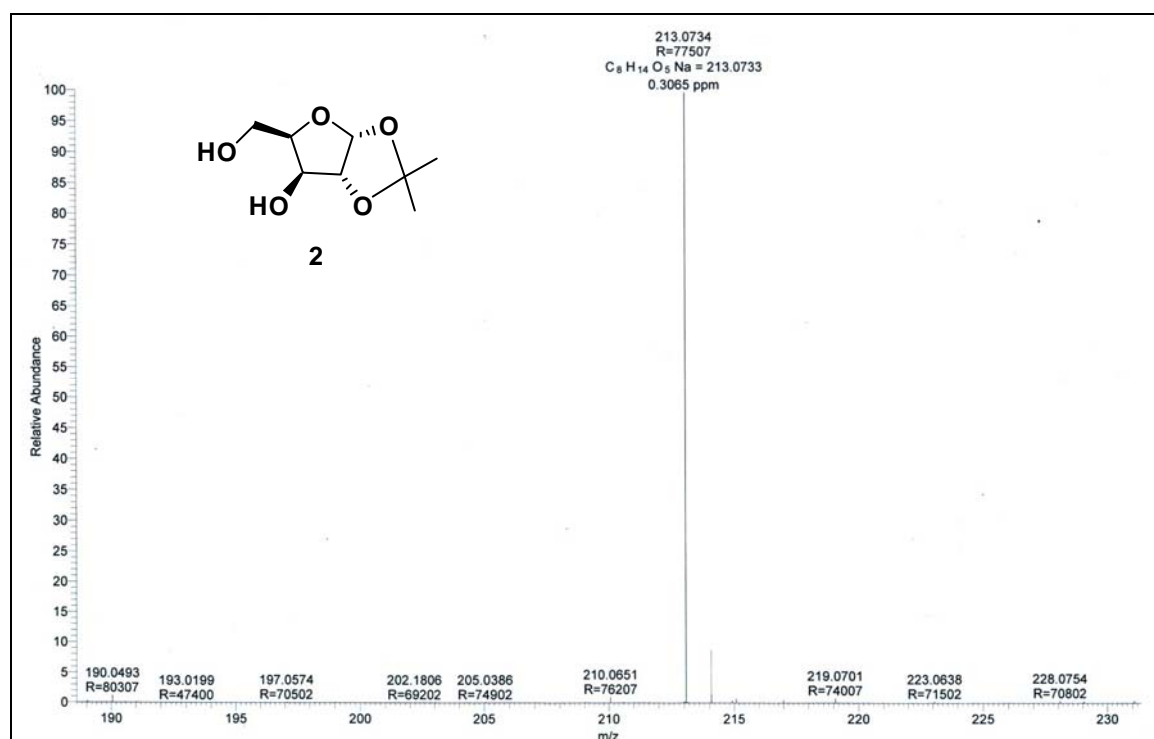
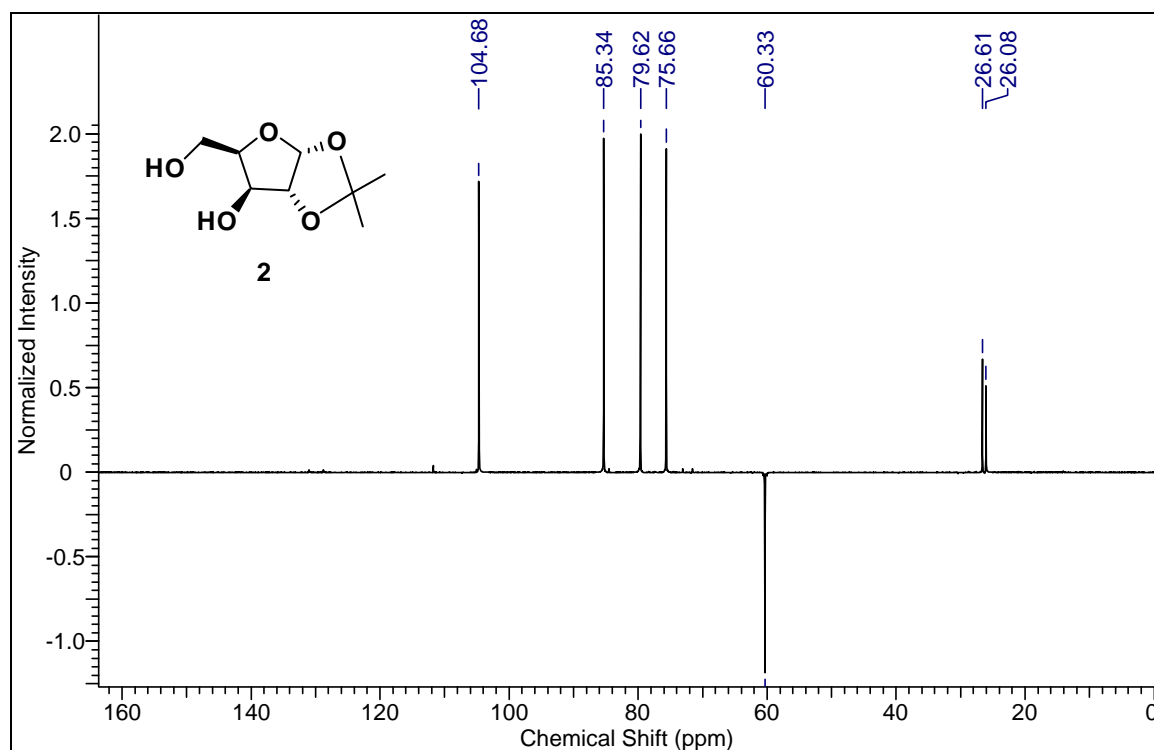
when TLC indicated the absence of starting material. The reaction mass was diluted with dichloromethane, washed with NaHCO<sub>3</sub> and water, dried over sodium sulphate, followed by solvent removal. The crude product was purified by silica gel column chromatography using a 1:1 mixture of dichloromethane:ethylacetate and 1% triethylamine. Yield 0.065 g, 68%.

**<sup>31</sup>P NMR (202 MHz, CDCl<sub>3</sub>)** 150.32, 150.93 δ ppm; **HRMS:** mass calculated C<sub>41</sub>H<sub>51</sub>O<sub>9</sub>N<sub>4</sub>NaP (M+ Na)<sup>+</sup> 797.3289, observed (M+ Na)<sup>+</sup> 797.3286.

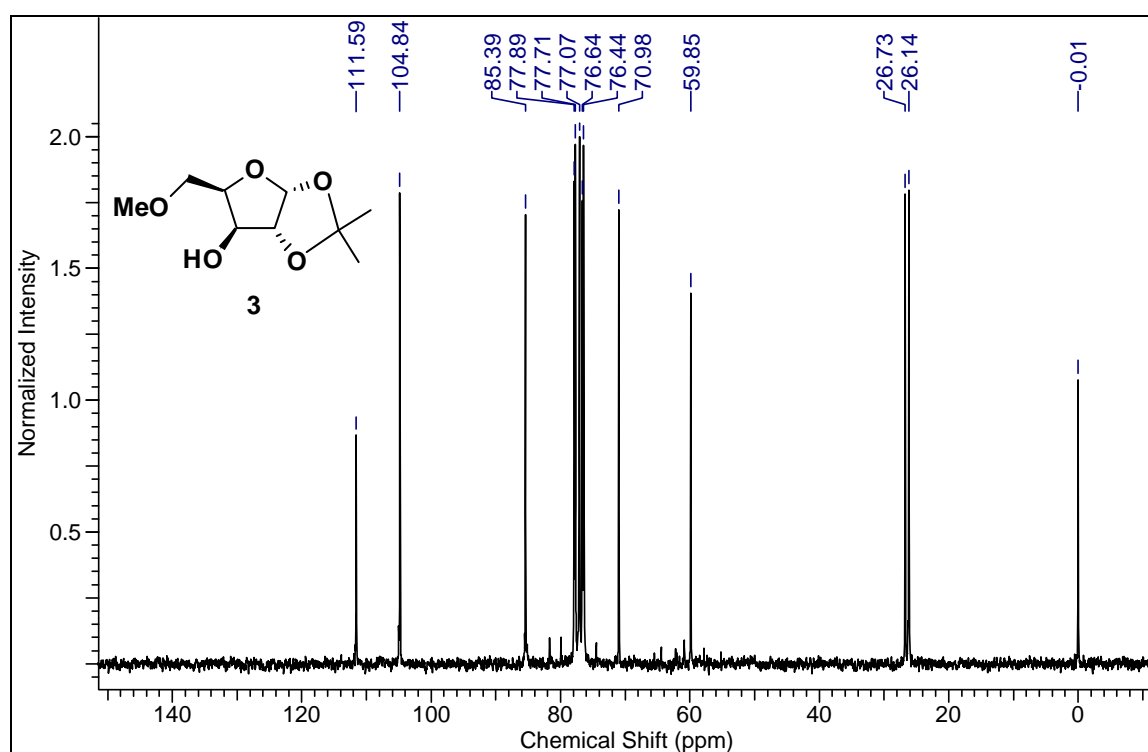
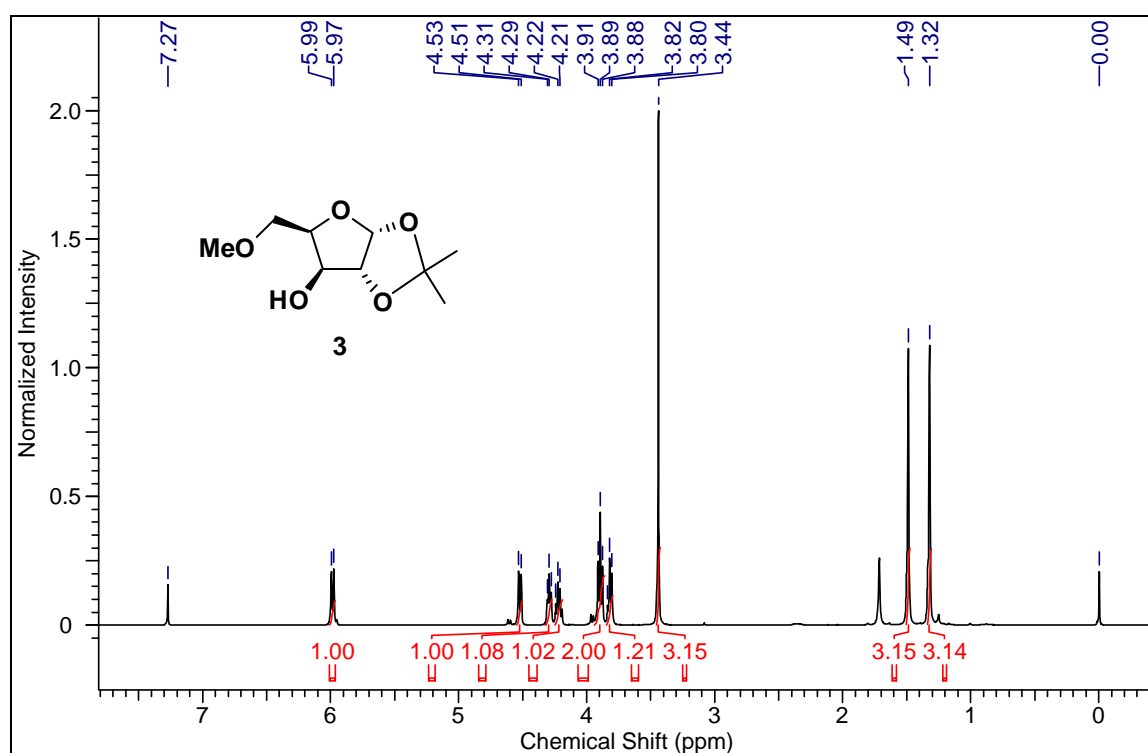
4A.7 Appendix

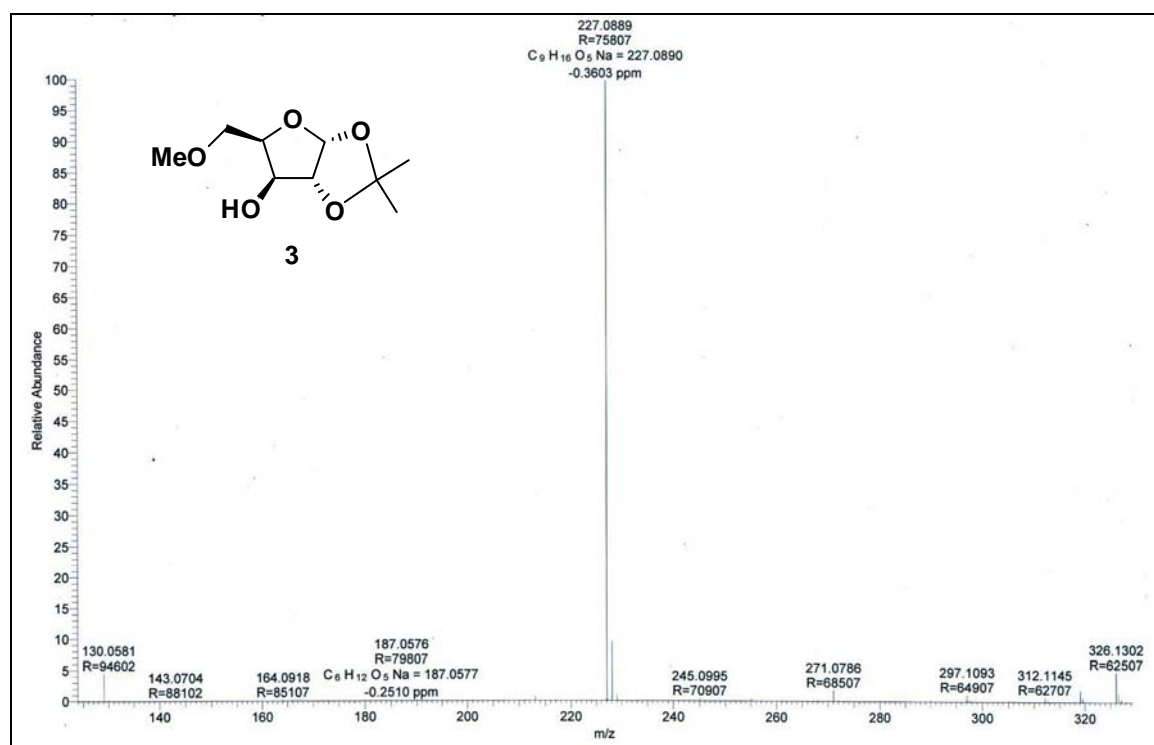
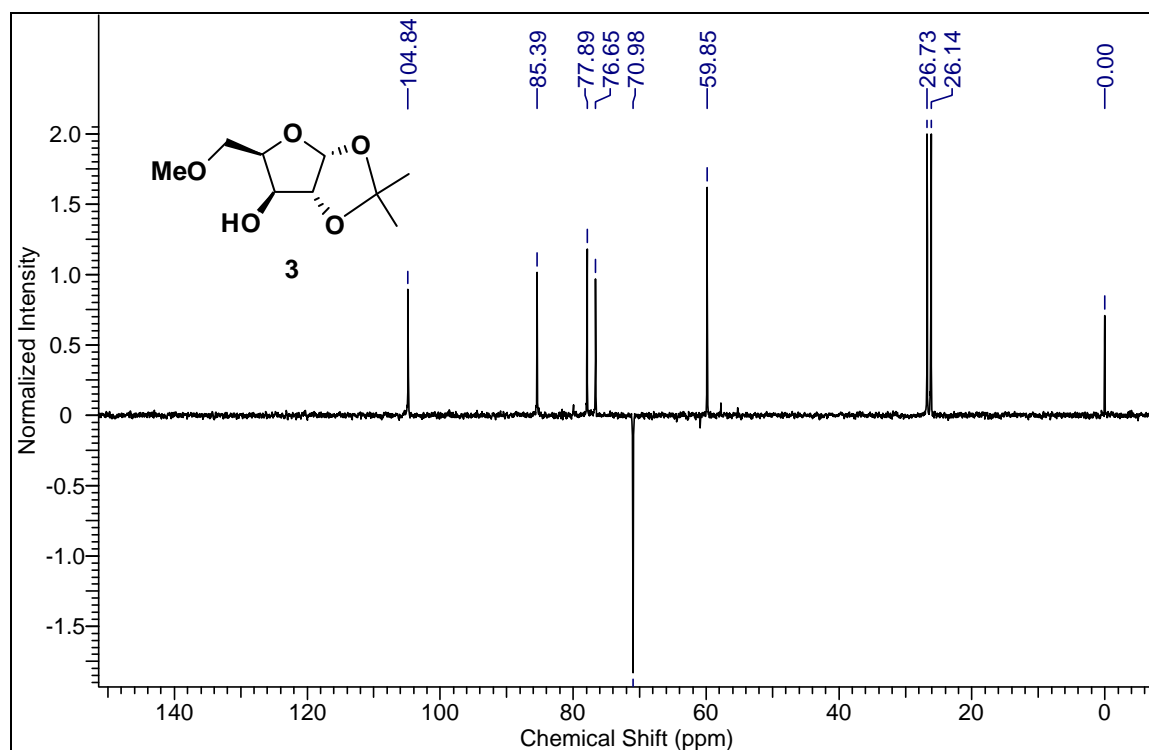
Compounds - Spectral data	Page No.
2 <sup>1</sup> H NMR & <sup>13</sup> C NMR	178
2- DEPT & HRMS	179
3 <sup>1</sup> H NMR & <sup>13</sup> C NMR	180
3- DEPT & HRMS	181
4 <sup>1</sup> H NMR & <sup>13</sup> C NMR	182
4- DEPT & HRMS	183
5 <sup>1</sup> H NMR & <sup>13</sup> C NMR	184
5- DEPT & HRMS	185
6 <sup>1</sup> H NMR & <sup>13</sup> C NMR	186
6- DEPT & HRMS	187
7 <sup>1</sup> H NMR & <sup>13</sup> C NMR	188
7- DEPT & HRMS	189
8 <sup>1</sup> H NMR & <sup>13</sup> C NMR	190
8 -DEPT & HRMS	190
9 <sup>1</sup> H NMR & <sup>13</sup> C NMR	192
9 -DEPT & HRMS	193
10 a <sup>31</sup> P NMR & <sup>13</sup> C NMR	194
HPLC & MALDI-TOF of DNA1-15T <sup>TNA</sup>	195
HPLC & MALDI-TOF of DNA1-9T <sup>TNA</sup>	196

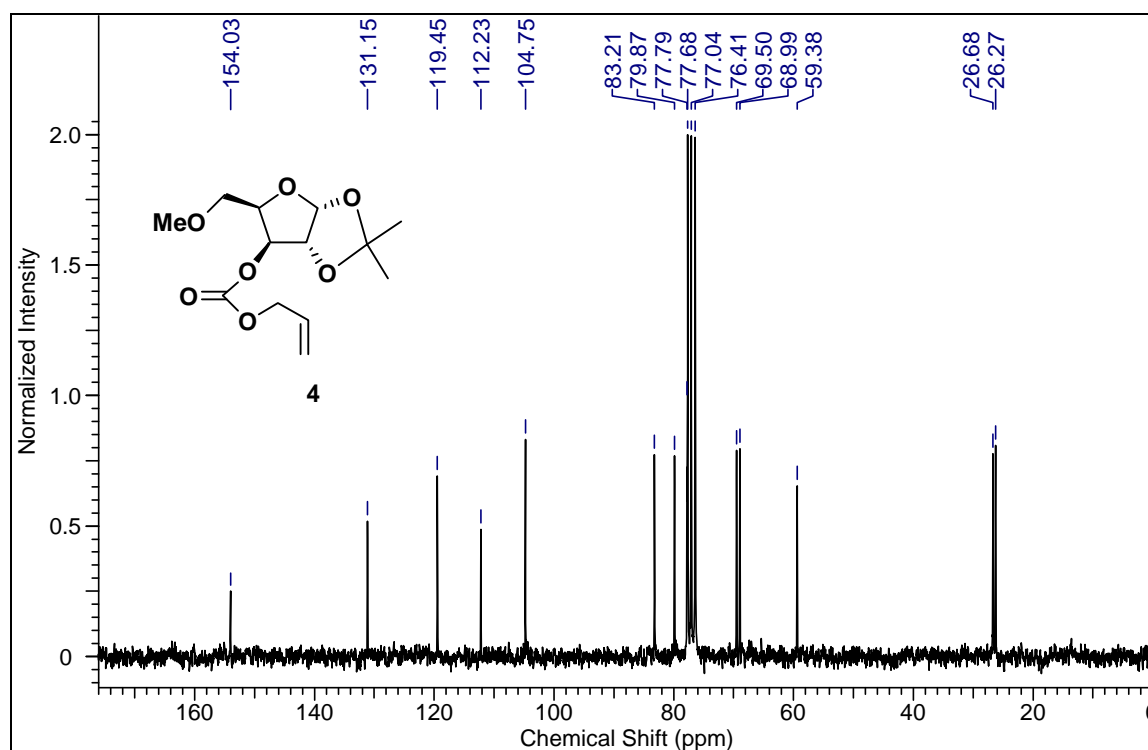
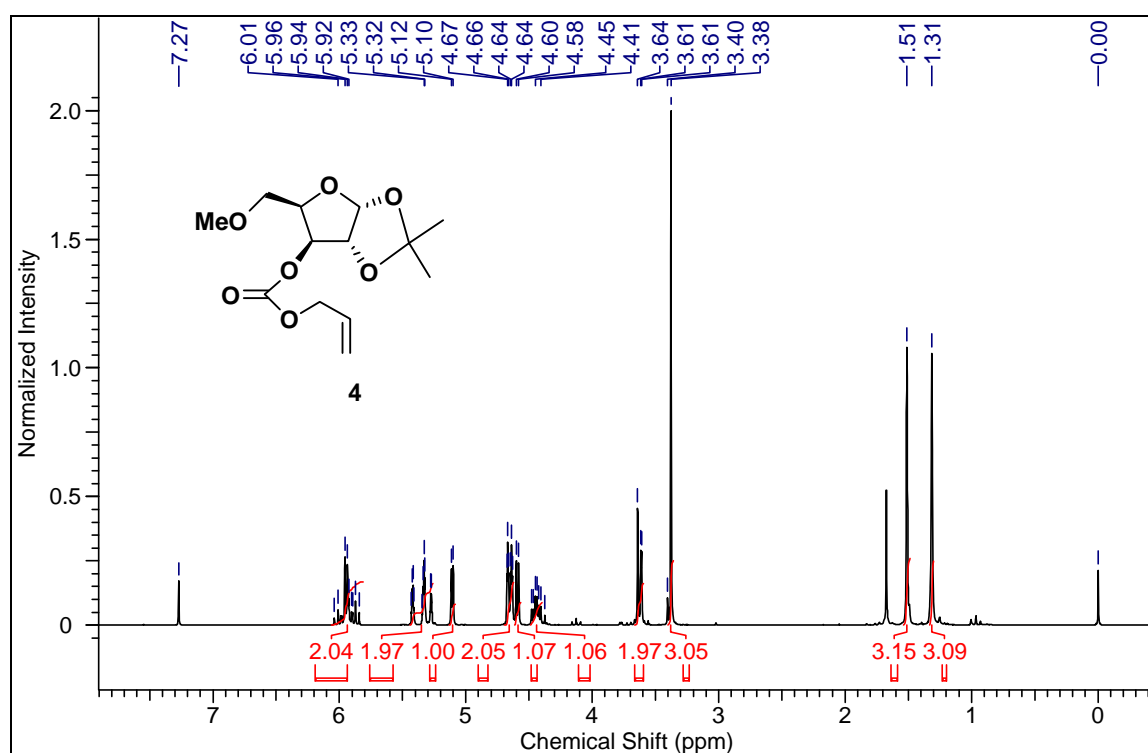


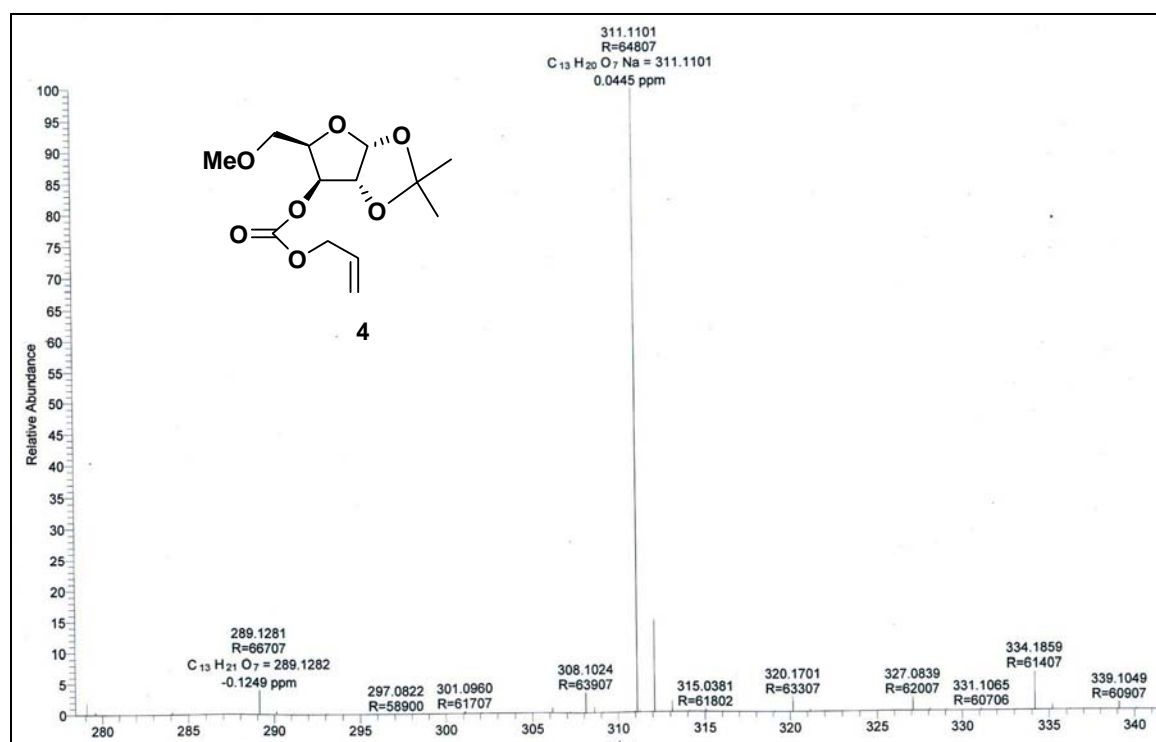
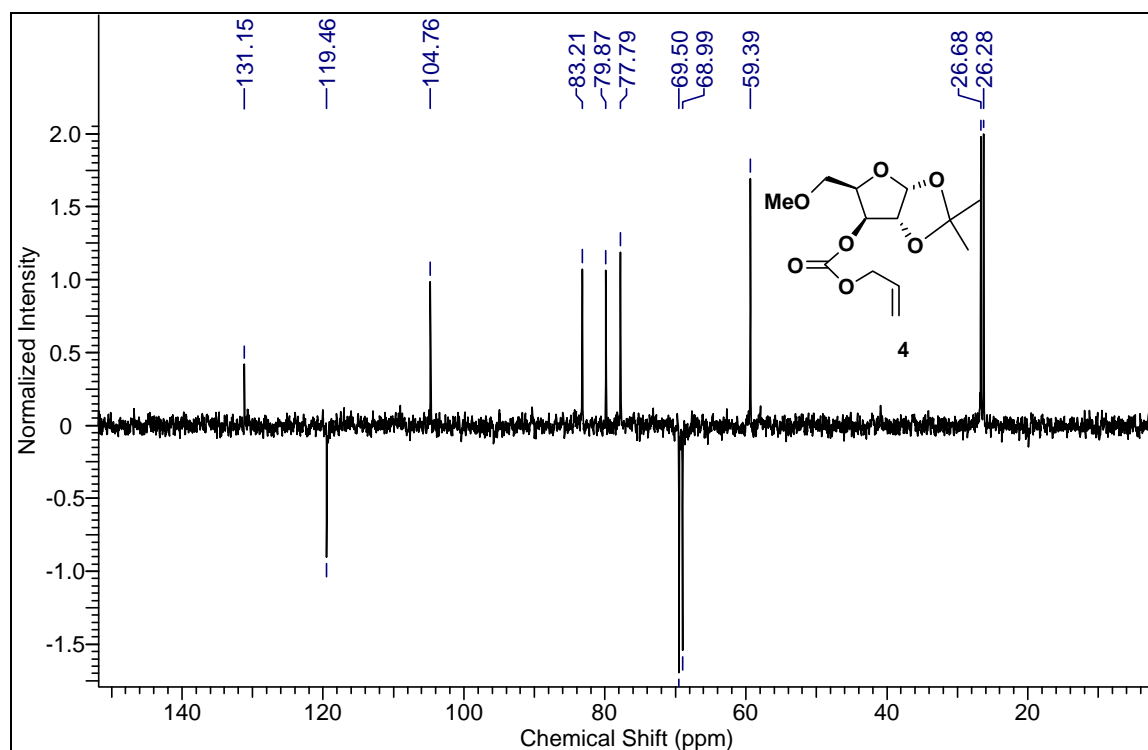


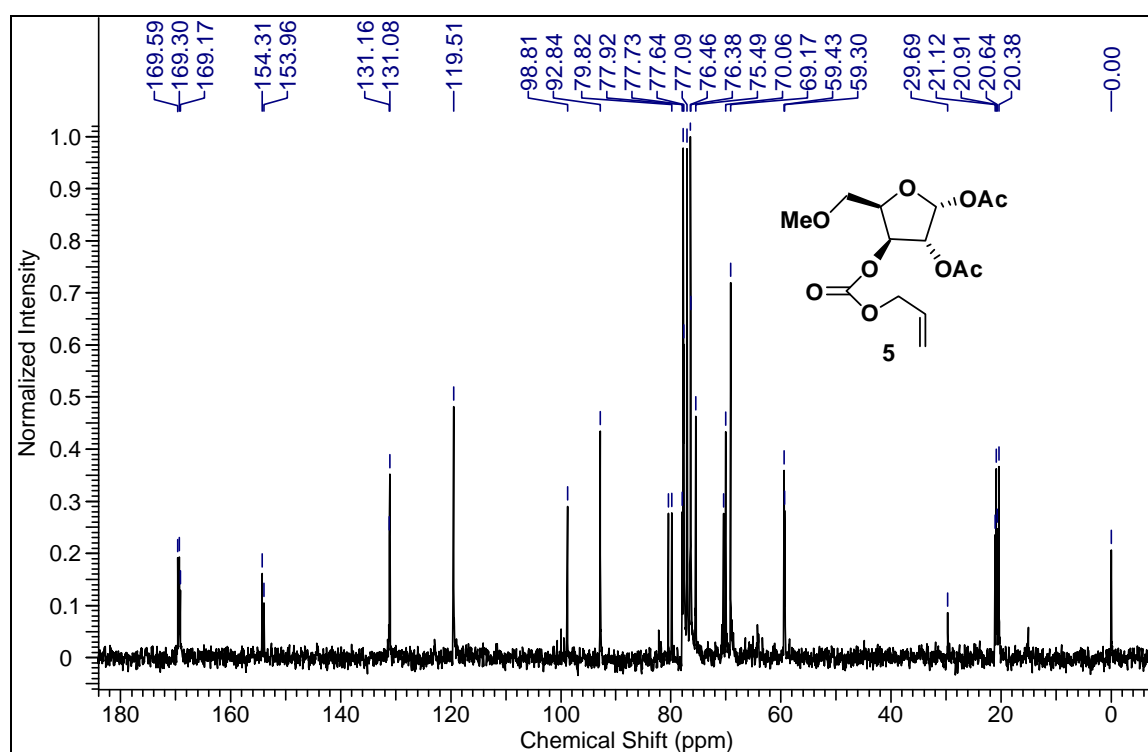
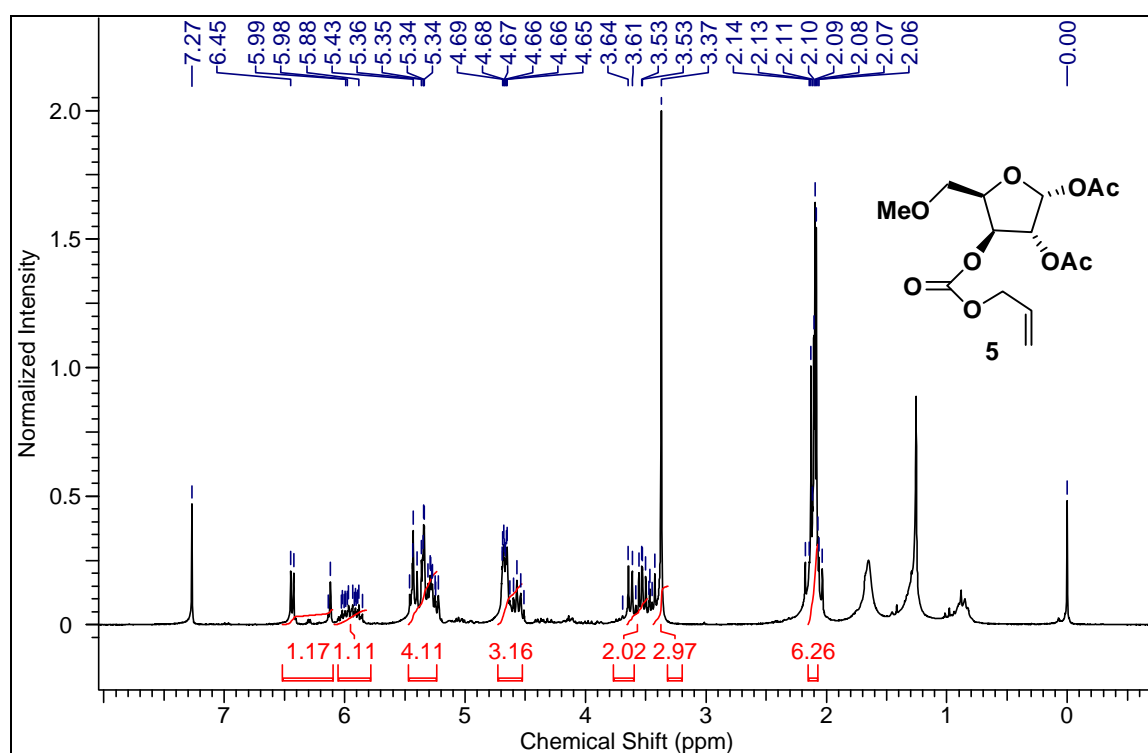


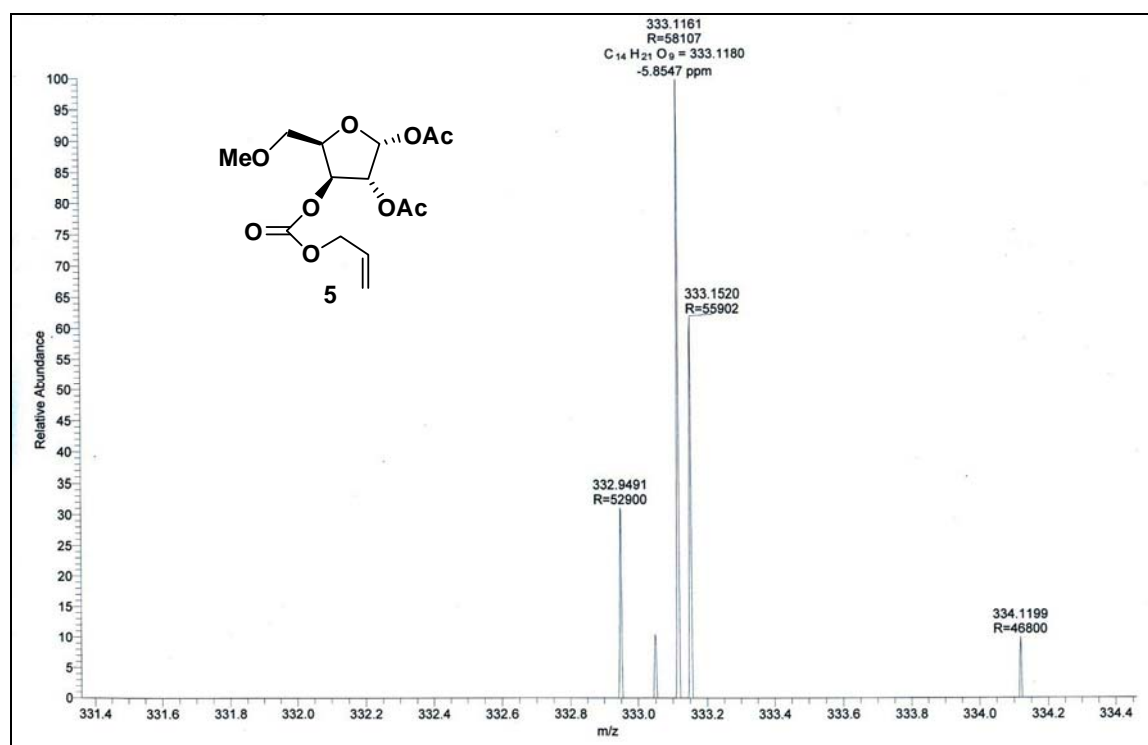
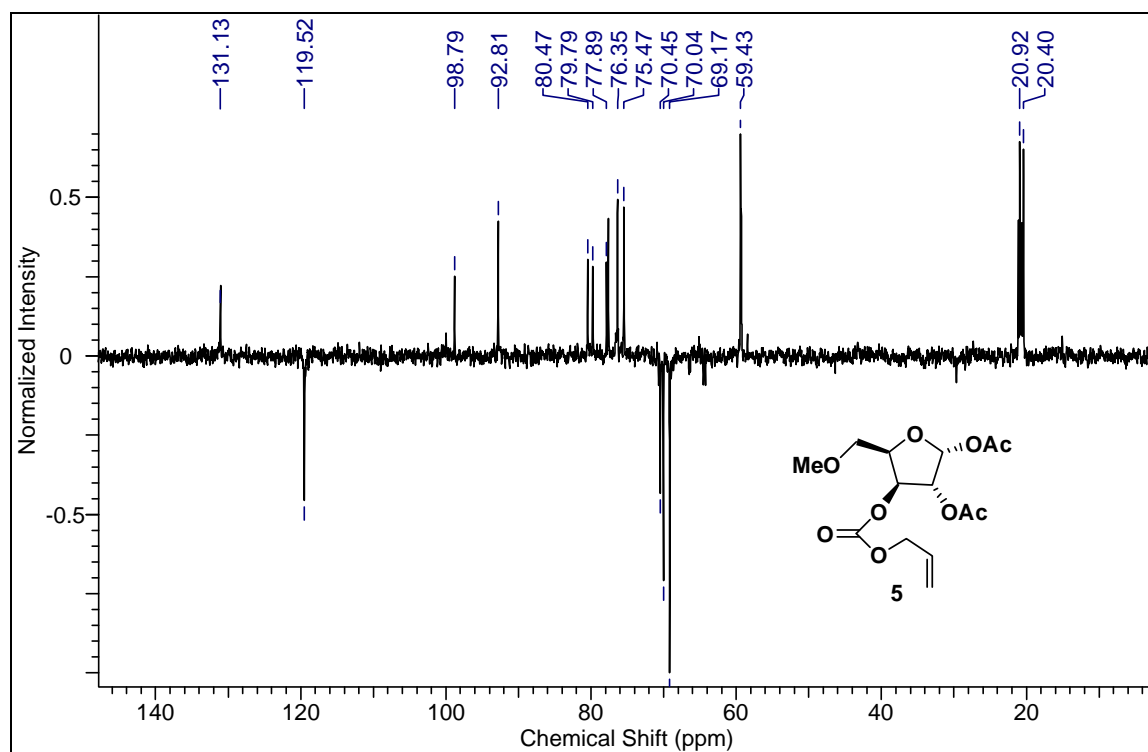


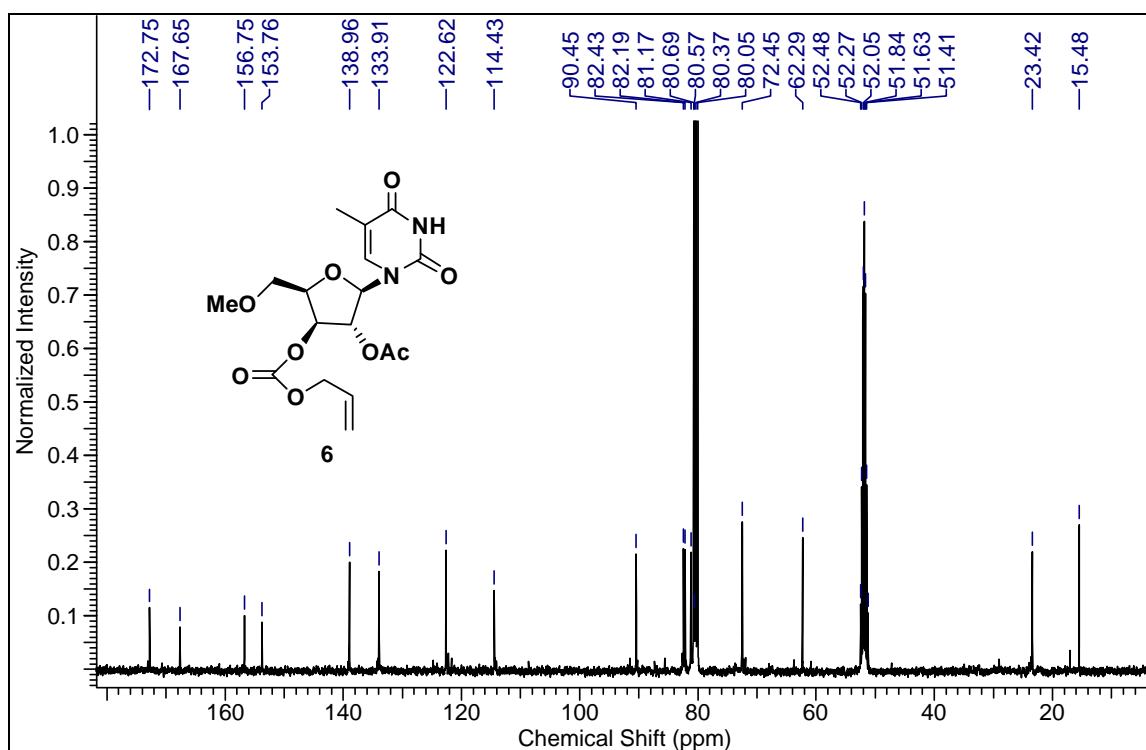
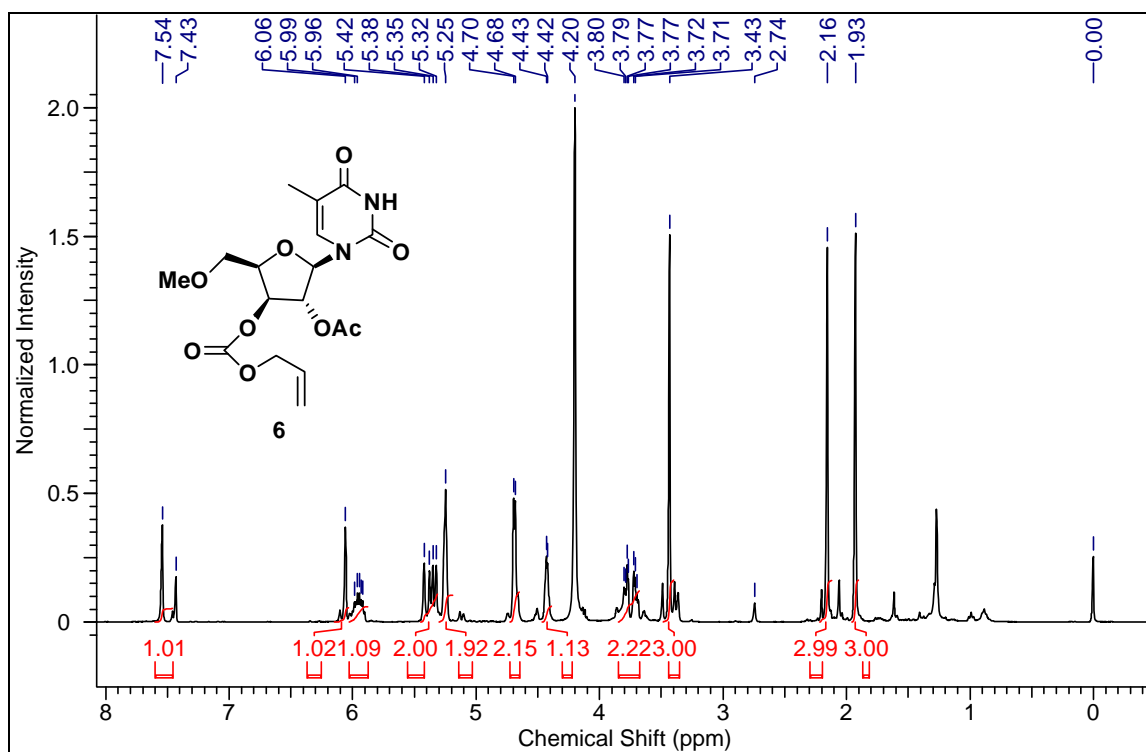


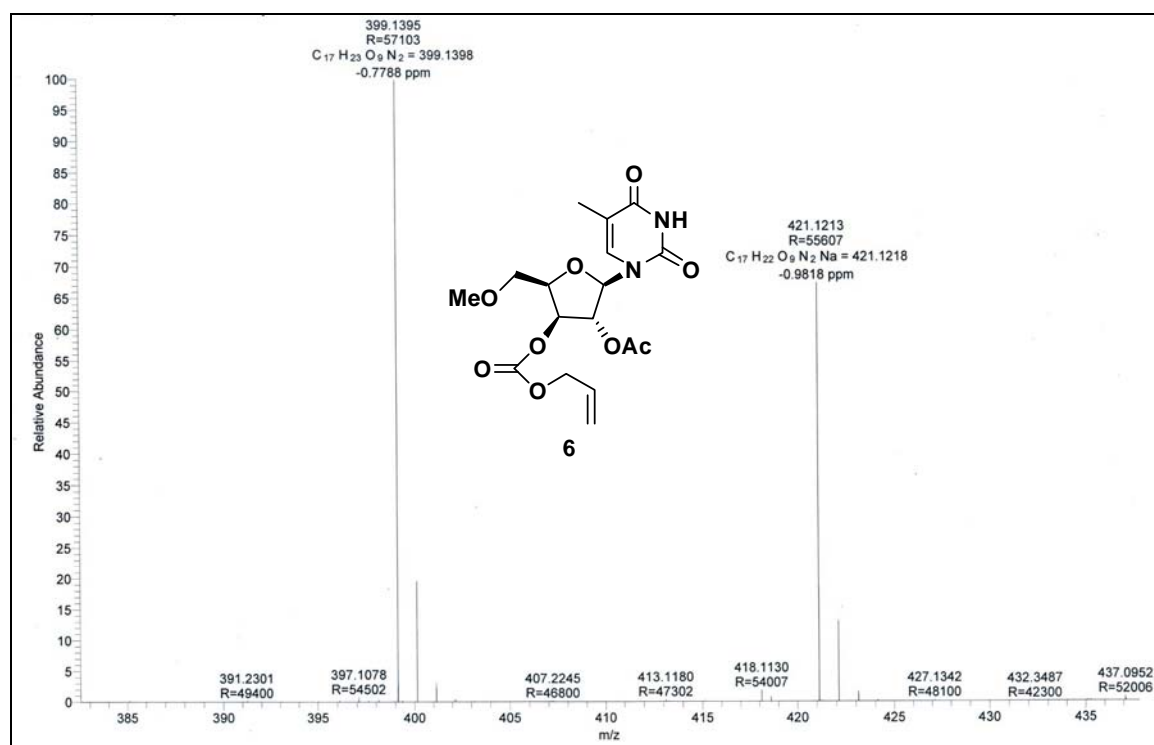
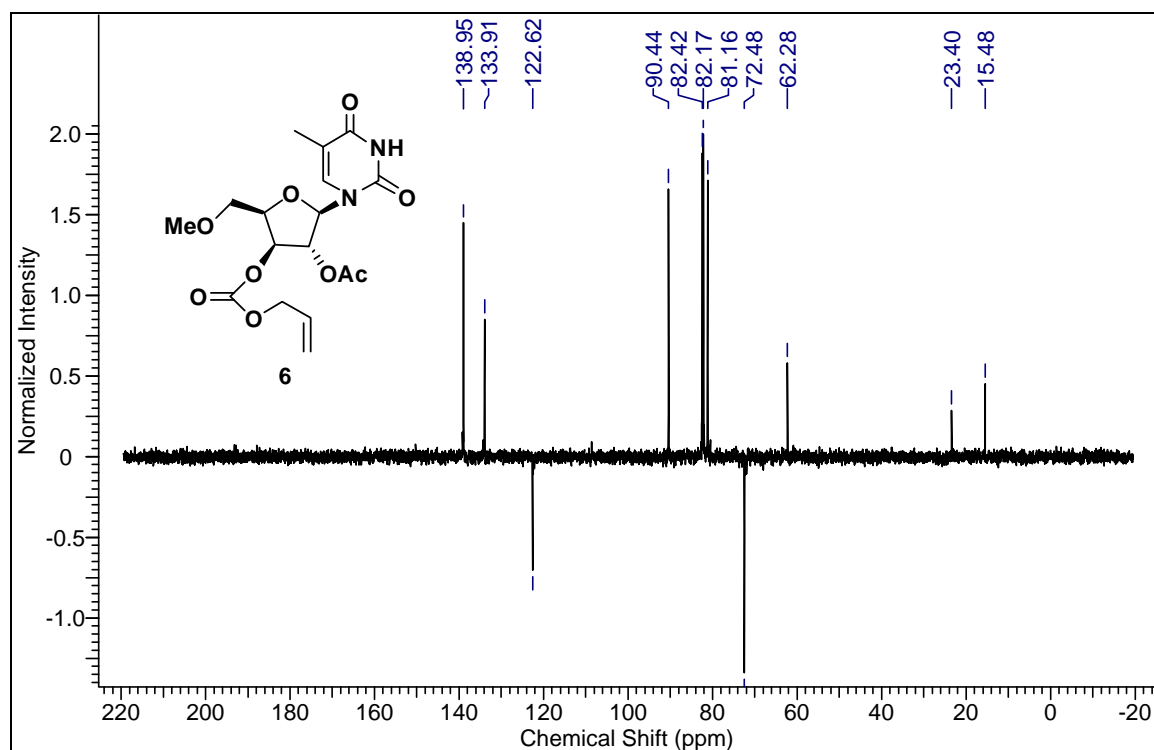




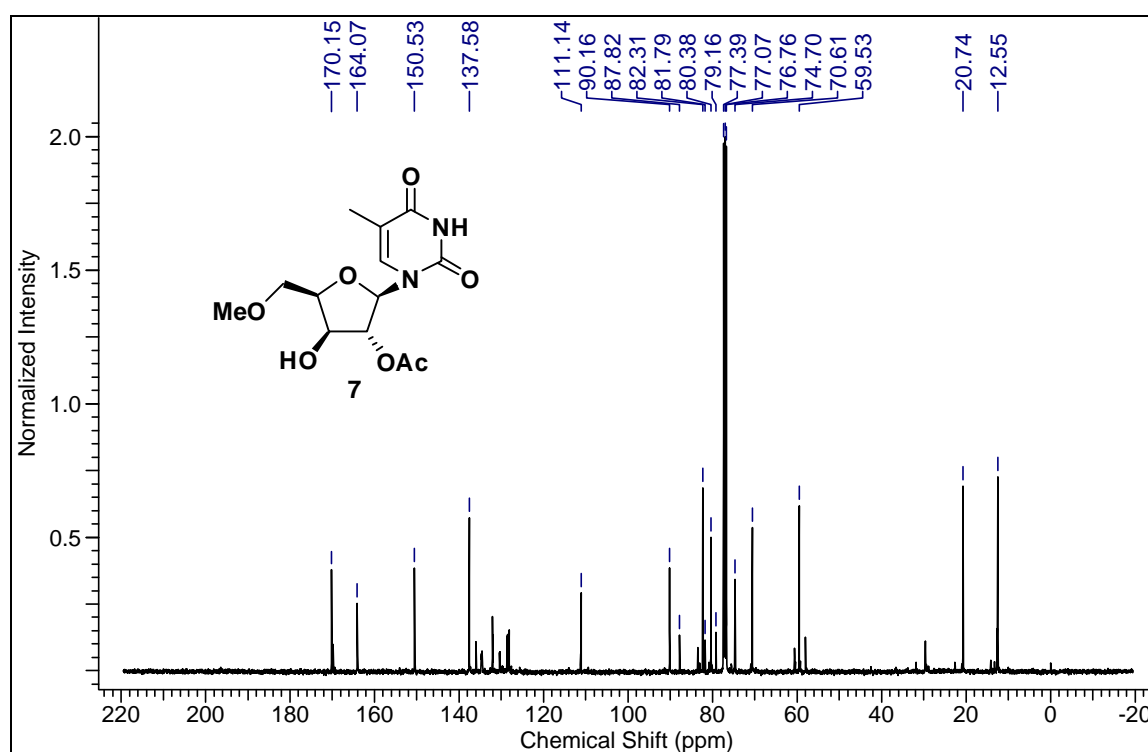
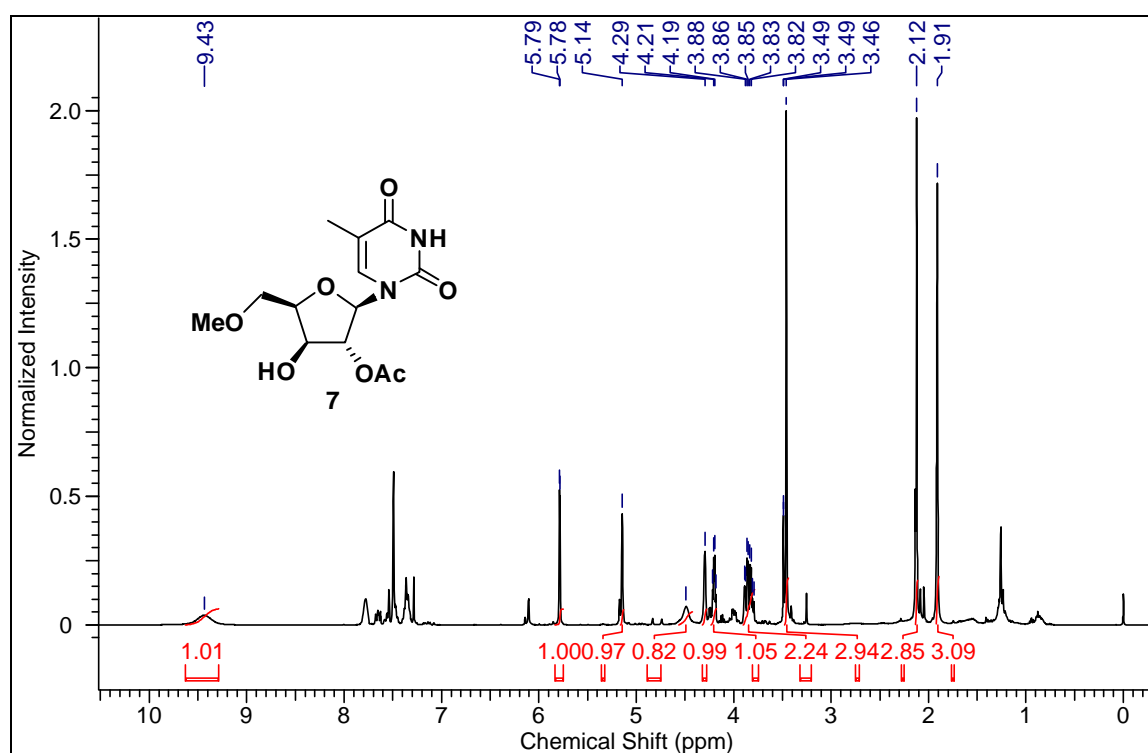


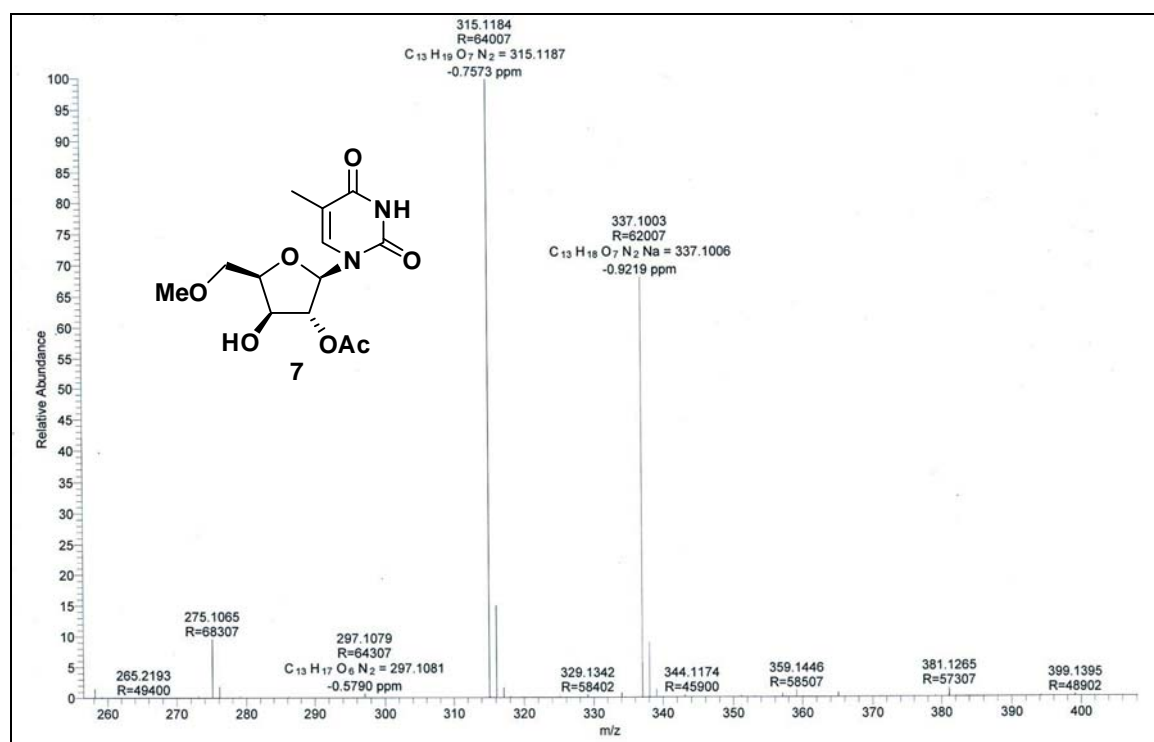
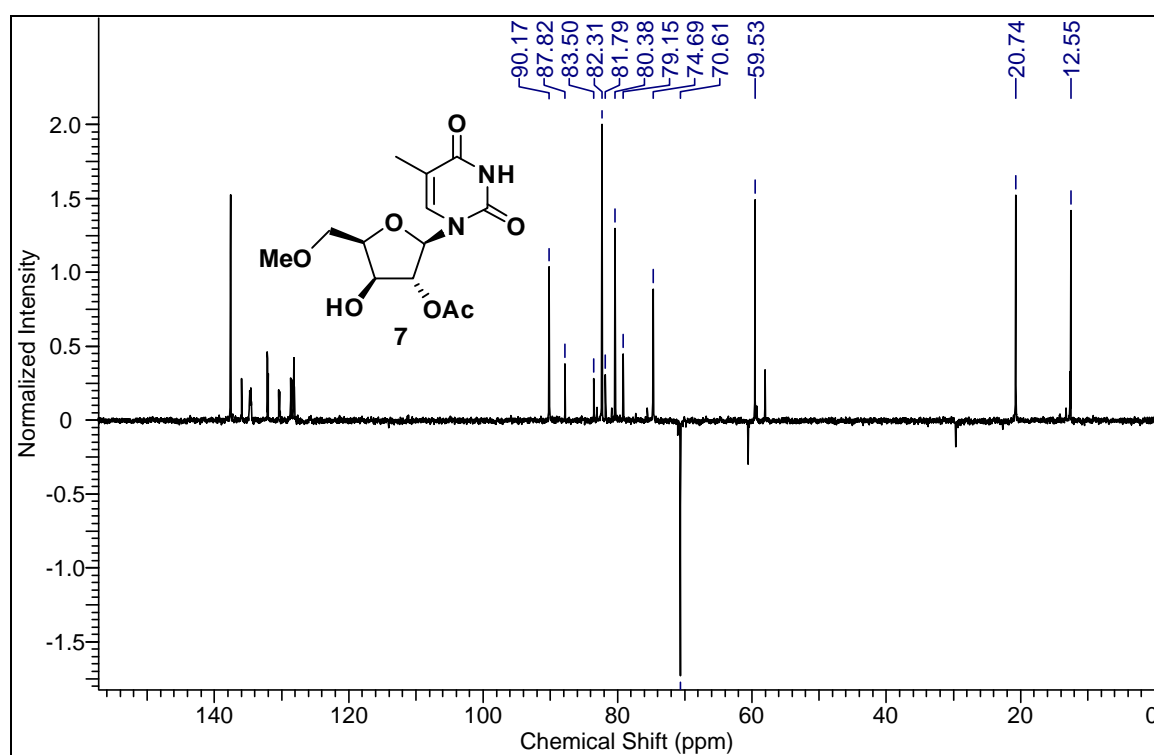


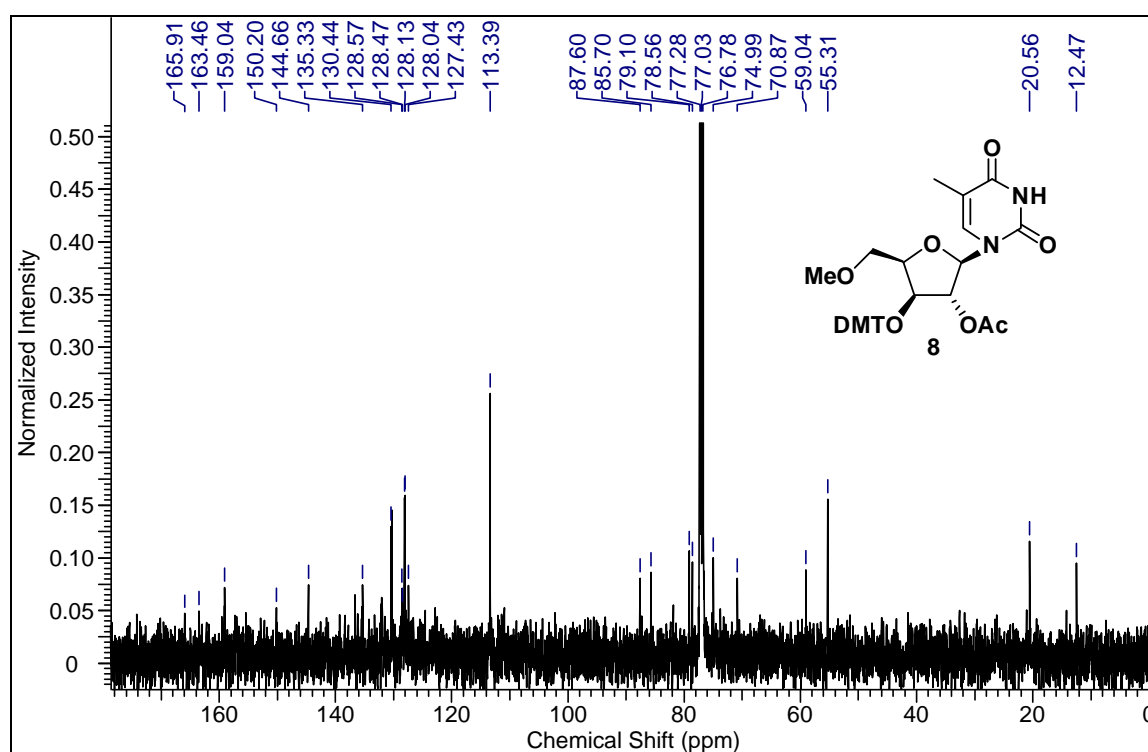
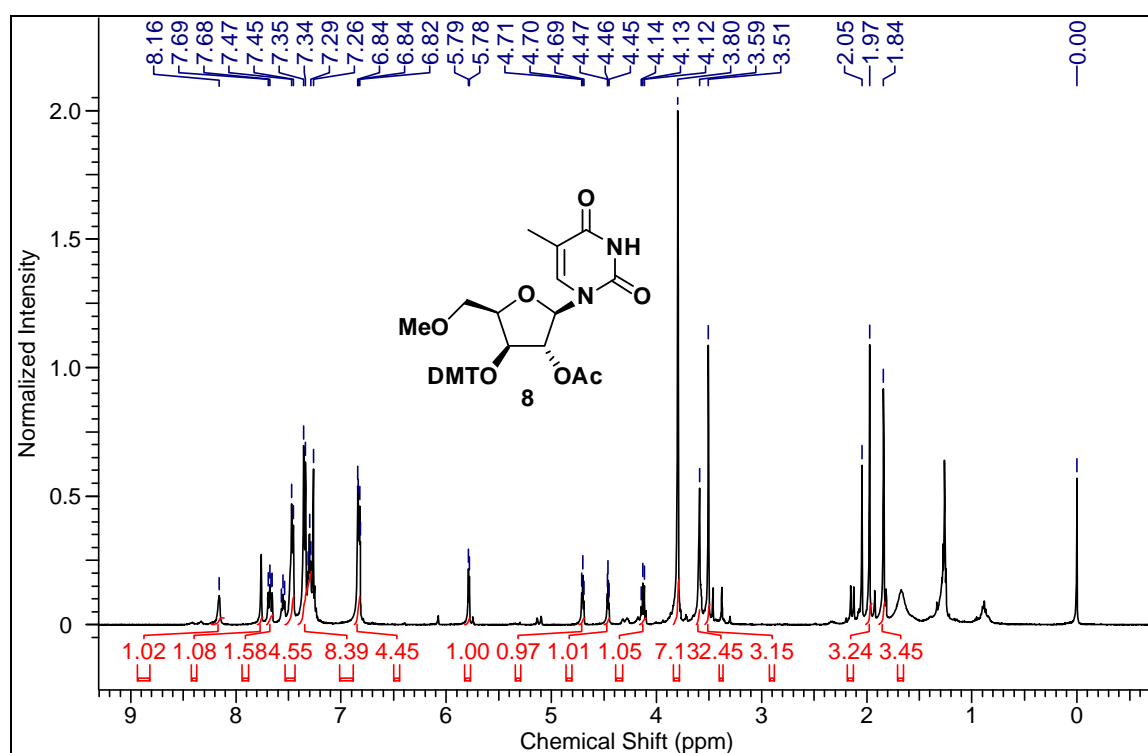


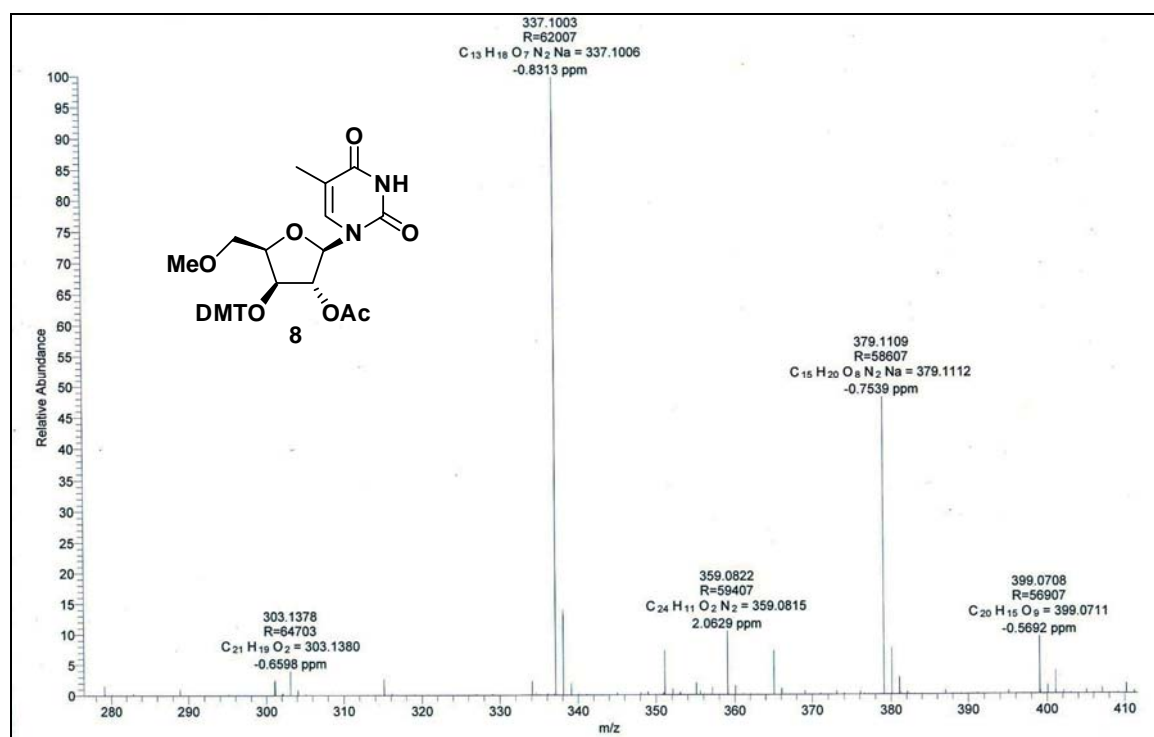
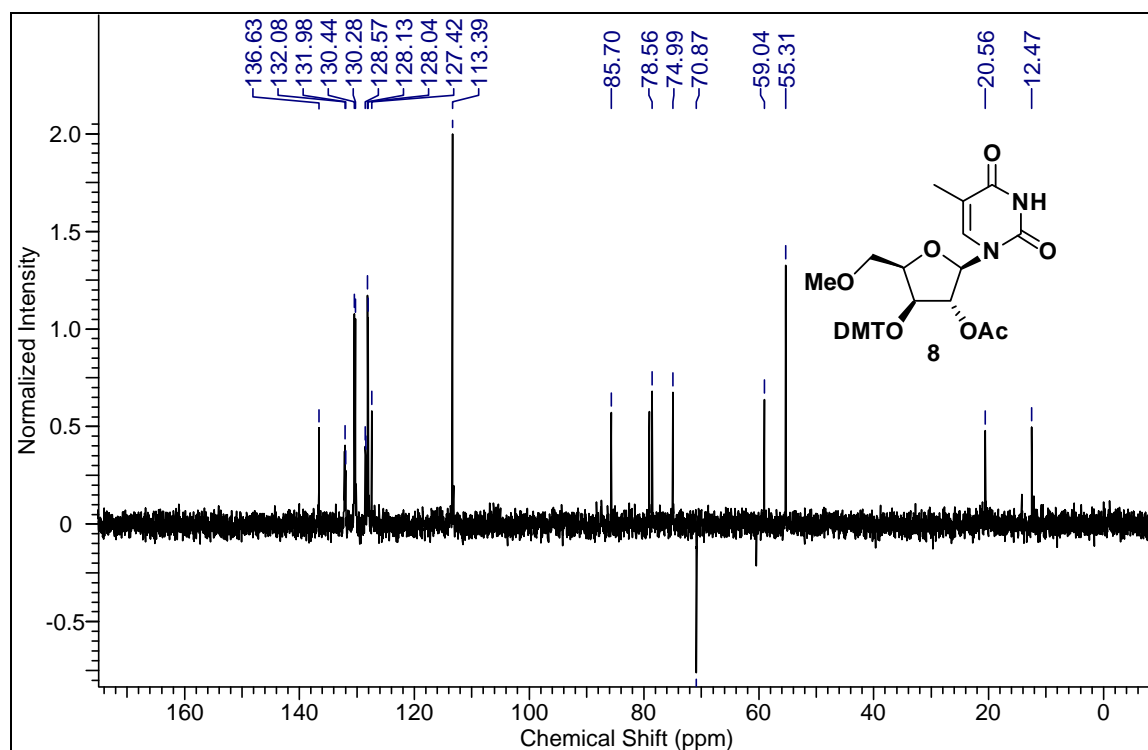


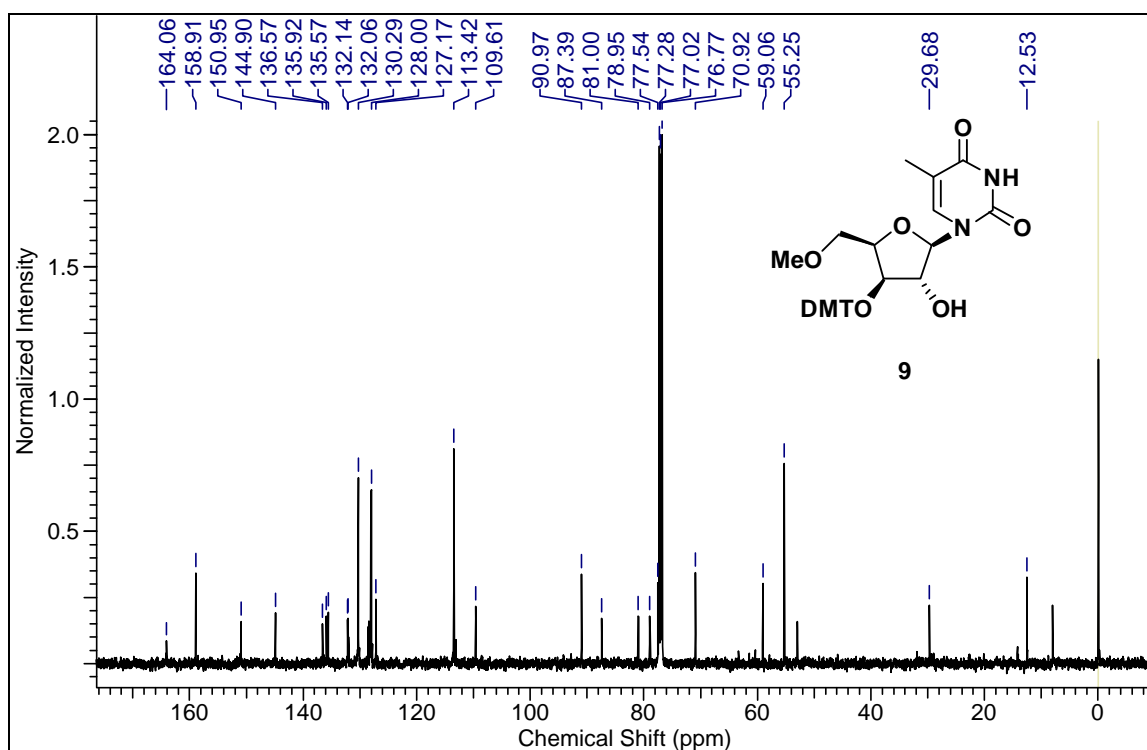
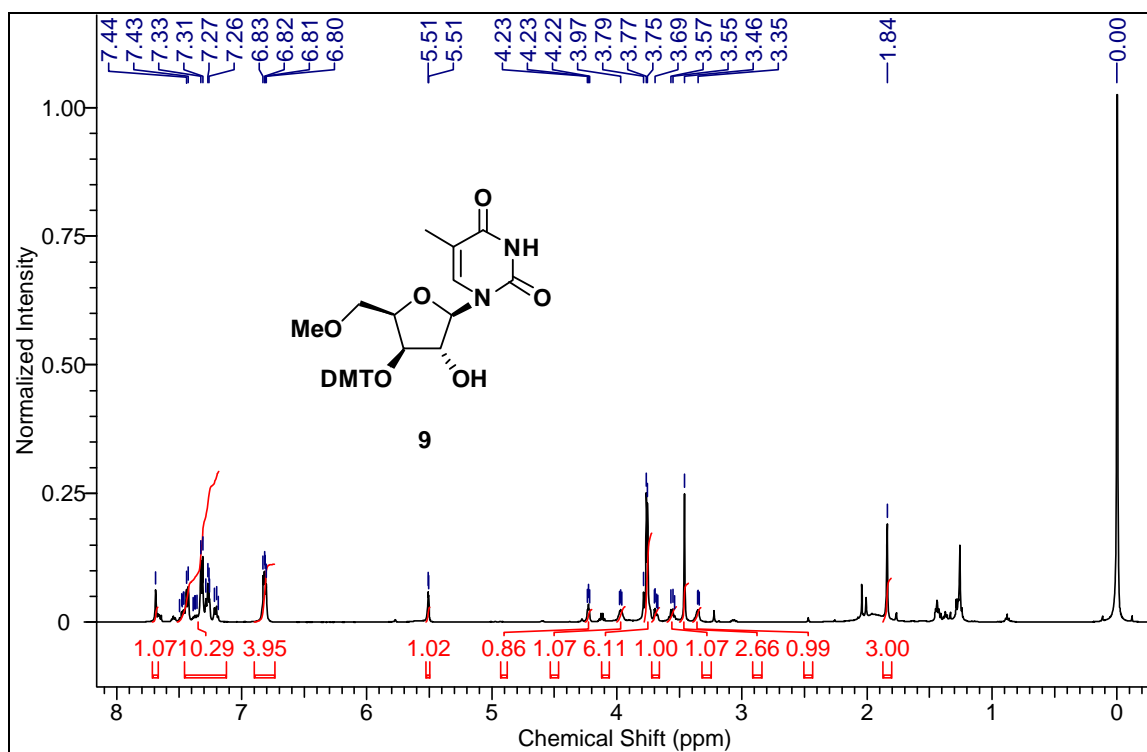


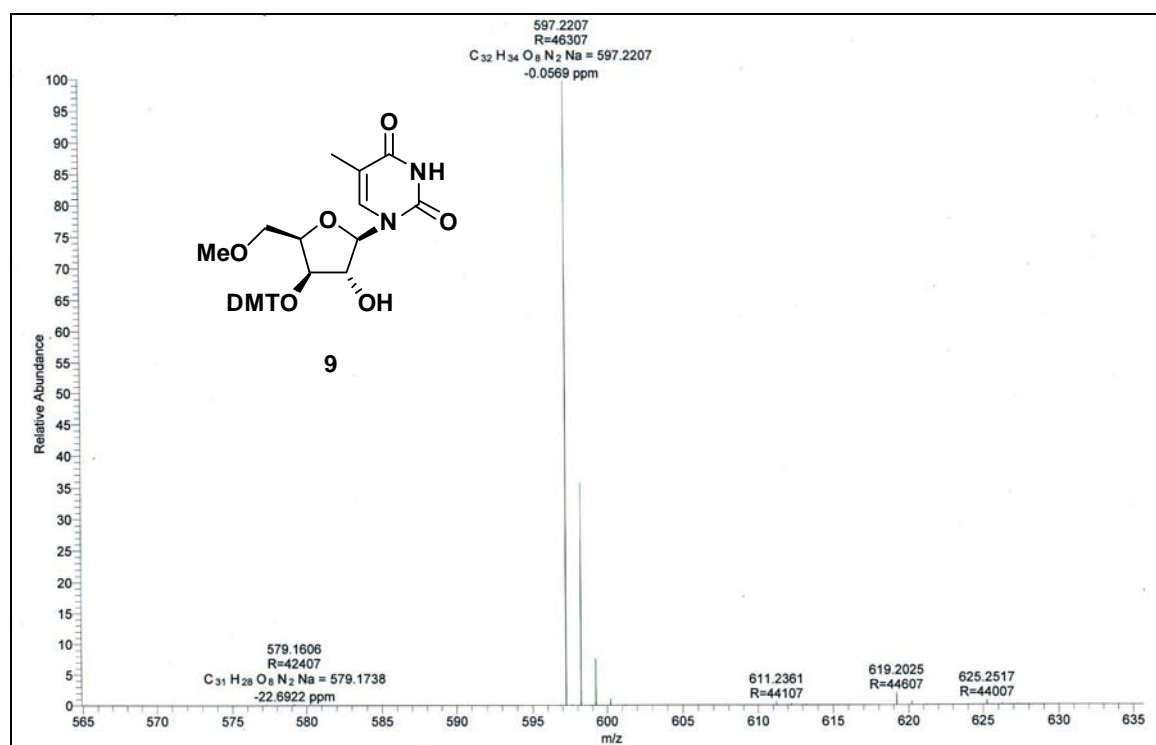
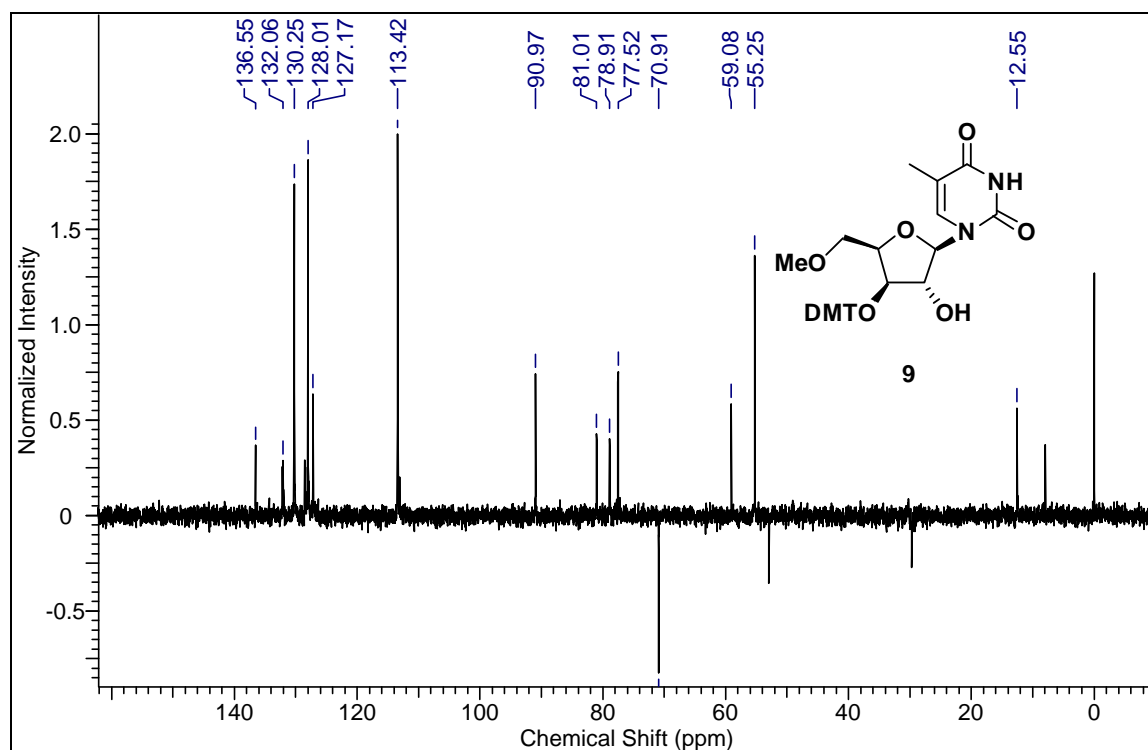


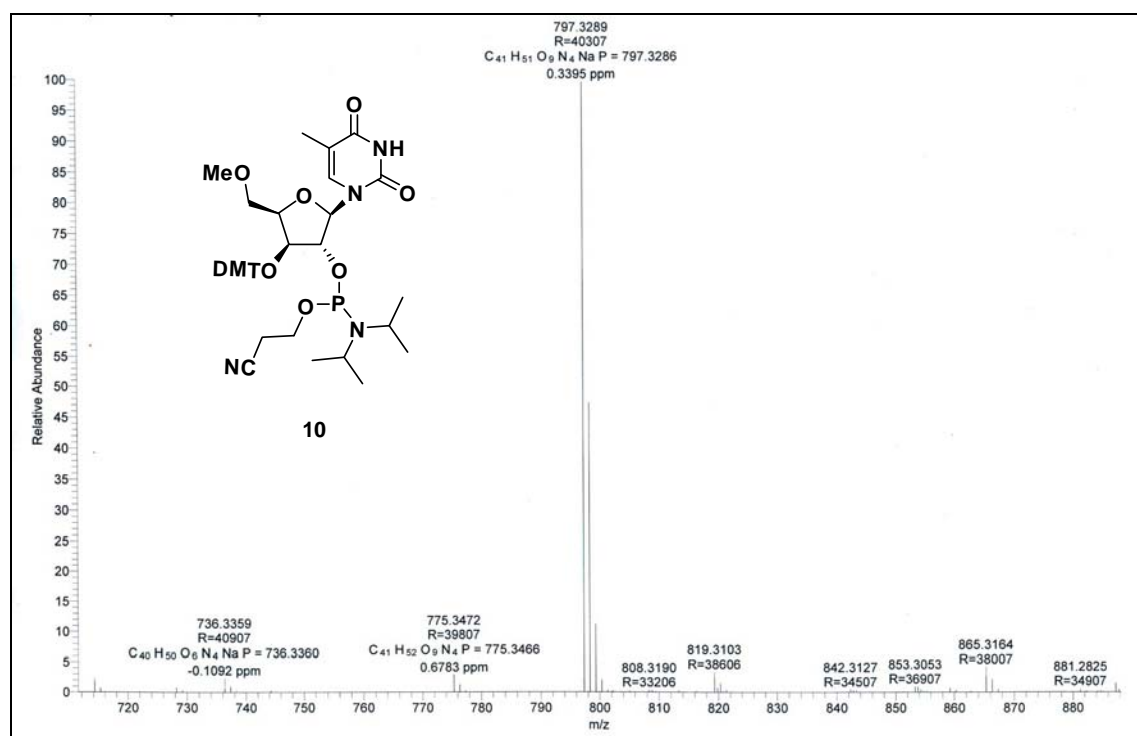
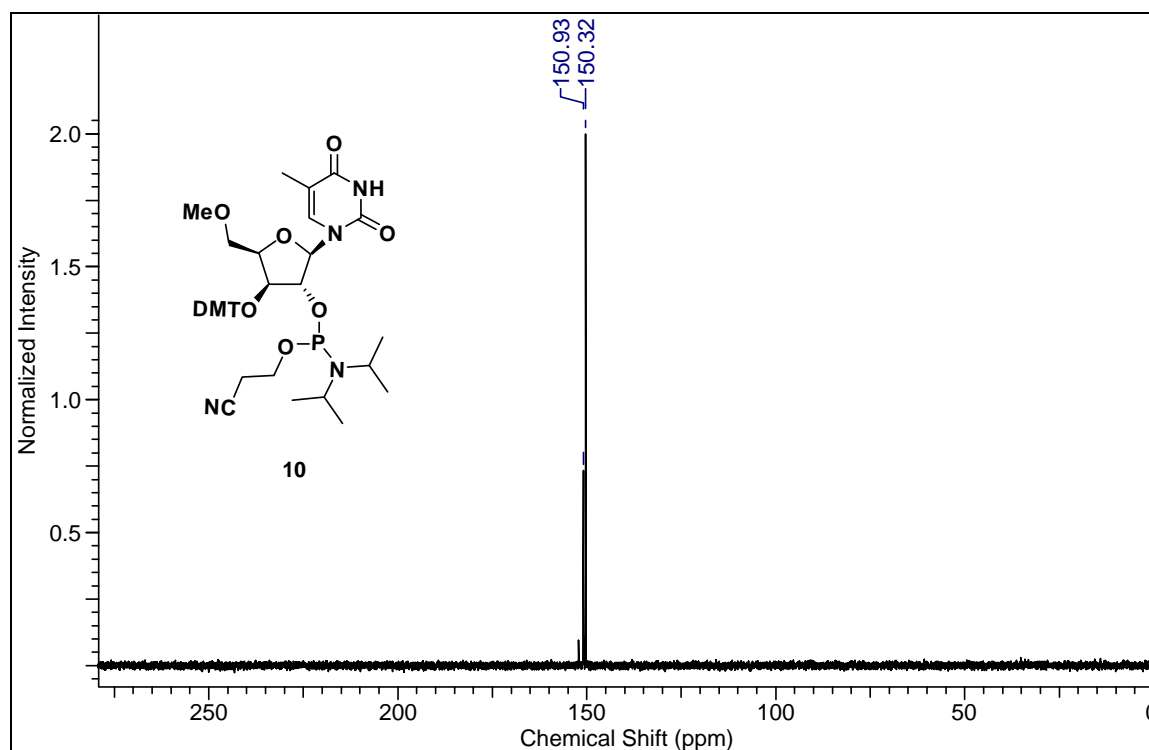






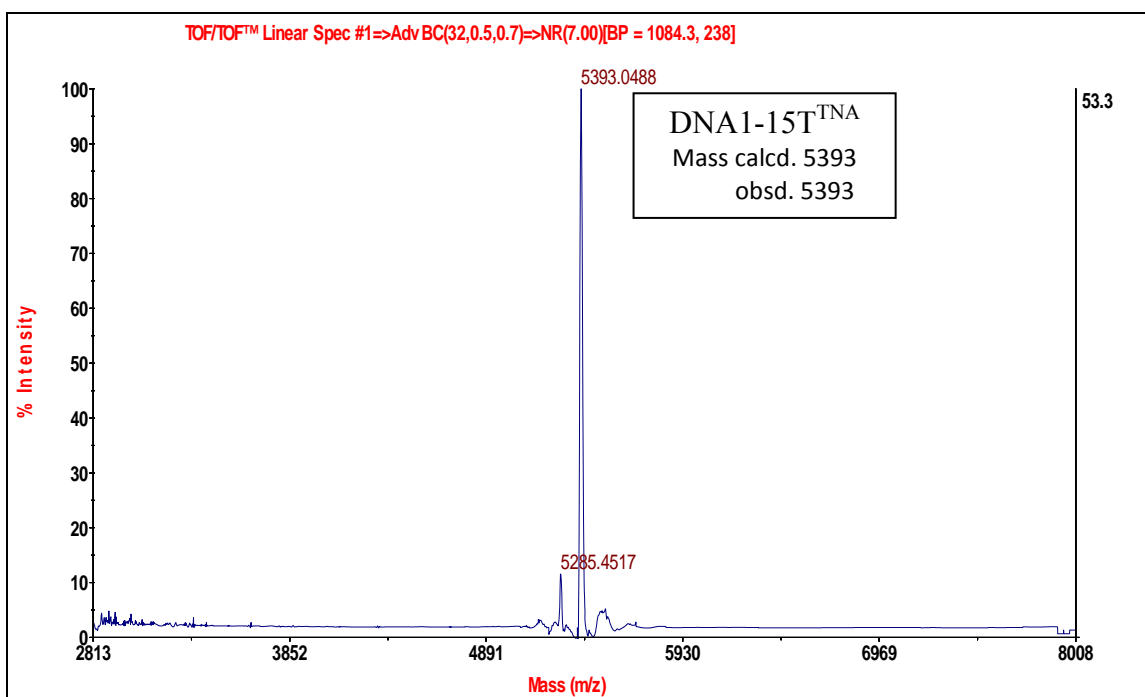
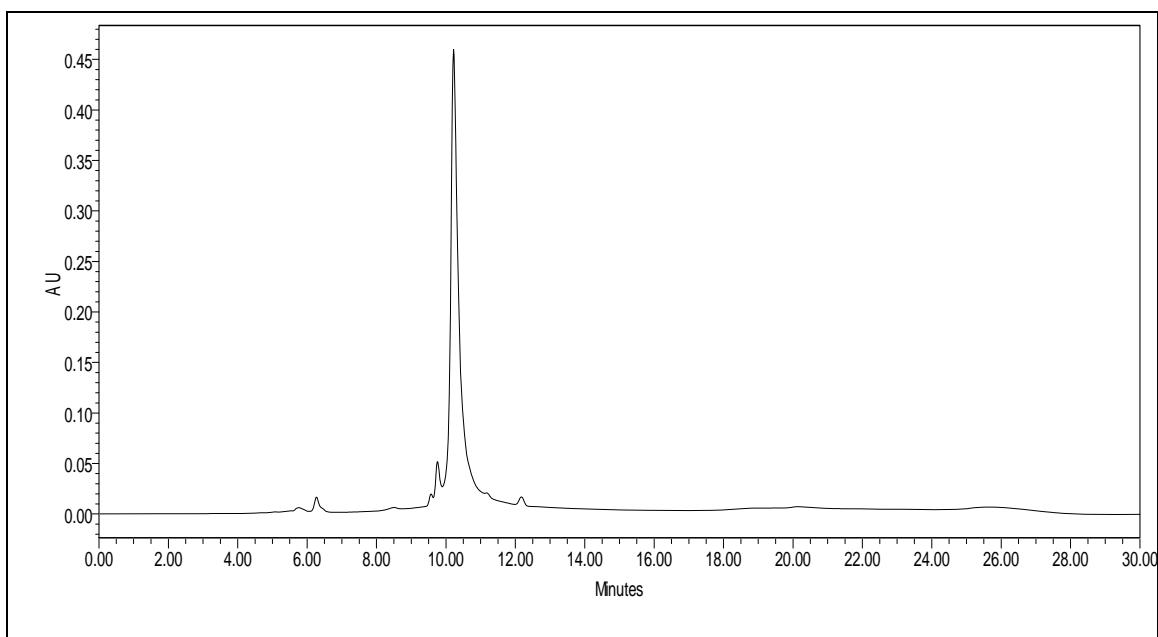






## Chapter 4

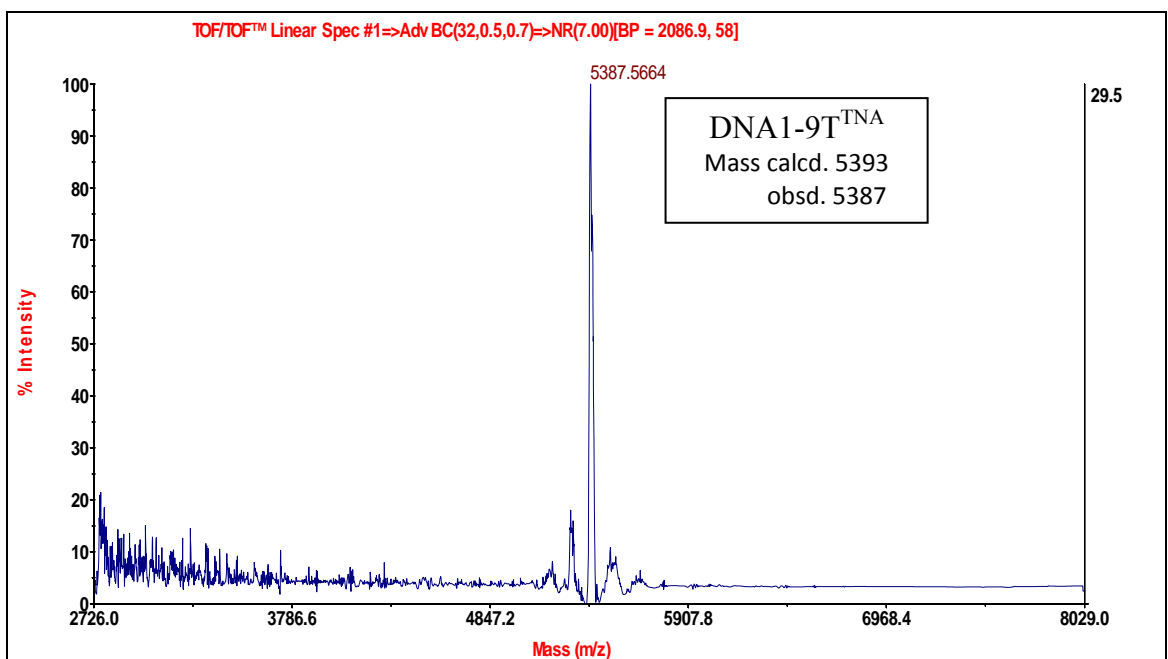
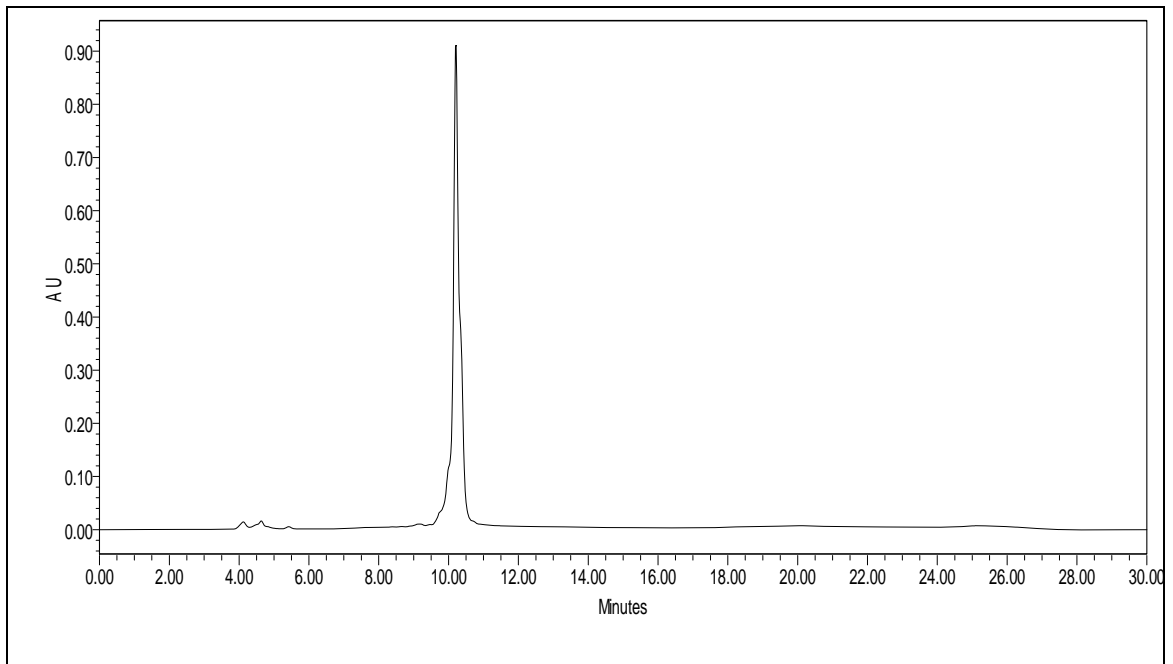
HPLC & MALDI-TOF of DNA1-15T<sup>TNA</sup> :





## Chapter 4

### HPLC & MALDI-TOF of DNA1-9T<sup>TNA</sup> :



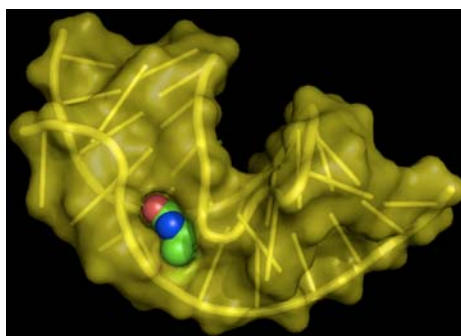
## **CHAPTER 4 / SECTION-B**

**Synthesis of 4'-MOM TNA  
modified thrombin binding aptamer,  
its quadruplex formation and  
application as a thrombin inhibitor**

## **Section 4B Synthesis of 4'-MOM TNA modified thrombin binding aptamer, its quadruplex formation and application as a thrombin inhibitor**

### **4B.1 Introduction**

Nucleic acid aptamers are acting as potential tools for molecular biology and medicinal chemistry. These are the single stranded DNA or RNA oligonucleotides, which can specifically bind to the target molecules. Aptamers can recognize various molecules, from small ligands to proteins due to their fold into complex three-dimensional shapes and form binding pockets for the specific recognition and tight binding of target (Figure 3). Huge number of aptamers to different kinds of targets have been reported to date, and continuously new aptamers are being discovered through an *in vitro* selection process called SELEX (systemic evolution of ligands by exponential enrichment).<sup>7</sup> Along with the designing of new aptamers, modifications of known aptamer work is also in progress to overcome the potential drawbacks, mainly poor stability, or to improve affinity and selectivity of nucleic acid aptamers.



**Figure 3** aptamer (<http://en.wikipedia.org/wiki/Aptamer>)

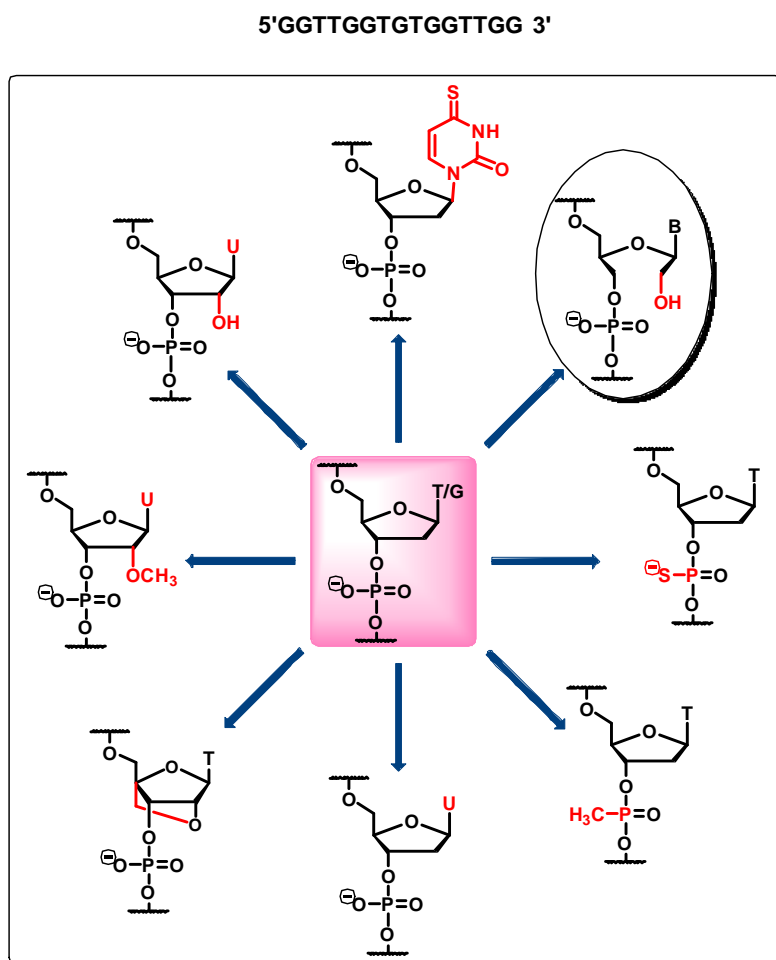
Thrombin is a key regulatory enzyme in the coagulation cascade. It is a serine protease produced from prothrombin by the action of factor Xa. Thrombin will catalyze the conversion of fibrinogen into fibrin, which is the building block of the fibrin matrix of blood clots.<sup>8</sup>

The **thrombin binding aptamer (TBA)** was discovered in 1992 by *in vitro* selection and bind to thrombin with high selectivity and affinity to inhibit fibrin-clot formation. NMR and X-ray structural study reveals that TBA forms an intramolecular,

antiparallel G-quadruplex with chair like conformation. This G-quadruplex consists of two G-quartets connected by three edge wise loops, a central TGT loop and two TT loops. The aptamer interacts with two thrombin molecules, inactivating only one of them.

#### 4B.2 Modifications in TBA

Natural TBA has less stability towards hydrolytic enzymes. Lot of attempts has been done to overcome this problem, by introducing chemical and structural modifications (Figure 4).<sup>9</sup>



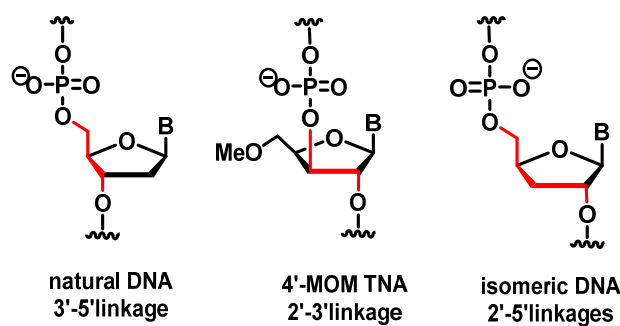
**Figure 4** Different modifications of TBA

The stability of G-quadruplex structure depends not only on the charge, but also the ionic radius of the atom in the oligonucleotide backbone.<sup>10</sup> Substitution of the oxygen atoms of the phosphate backbone with sulfur atom reduces the quadruplex structure stability. The 2'-modified Locked nucleic acid analogue destabilizing the quadruplex structures and Uridine modified analogs are stabilizing.

Unlocked nucleic acid is one of the prominent modifications known in the literature for enzymatic stability, which was introduced by Wengel *et al.* They have incorporated these UNA monomers into TBA and studied the stability of quadruplex structure. According to X-ray data of TBA bound to thrombin (Chapter 3 Figure 9b), T7 is buried in a hydrophobic cluster in the fibrinogen recognition site. The increased affinity may thus be attributed to a favorable change of the orientation with the T7 UNA-modified monomer leading to a better quadruplex-protein complex fit.<sup>11</sup>

#### 4B.3 Present work

We considered the synthesis of modified TBA and incorporated the 4'-MOM TNA at T7 and T9 positions to study the effect on stability of the quadruplex. TNA is having the 2'-3' backbone and for comparison, we also synthesized the TBA sequences with 2'-5' modification<sup>12</sup> at 7<sup>th</sup> and 9<sup>th</sup> position (TGT loop) and we studied the tetraplex stability, anticoagulation activity as well as enzymatic stability (Figure 5).



**Figure 5** Different types of backbone linkages

#### 4B.4 Chemical synthesis of 4'-MOM TNA and *iso* DNA modified TBA sequences

Using an automated Bioautomation MM-4 DNA synthesizer and commercially available phosphoramidites we synthesized unmodified TBA as a control and incorporated the modified monomers at 7<sup>th</sup> and 9<sup>th</sup> positions. All the sequences purified by HPLC, and masses were confirmed by MALDI-TOF (Table 2).

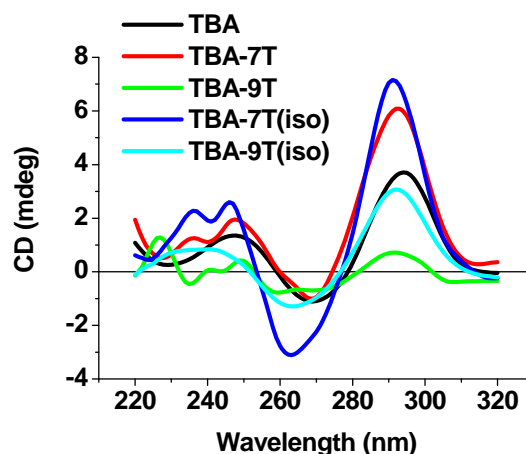
#### CD spectroscopy and $T_m$ measurement of the TBA sequences in presence of $K^+$ ion

The G-quadruplex formation of the synthesized sequences was studied by CD spectroscopy<sup>12</sup> in the presence of added monovalent cation  $K^+$  and their stability was determined as a function of temperature dependent change in CD amplitude at 295nm.

**Table 2** Quadruplex melting data of TBA and modified TBA sequences in presence of  $K^+$ 

Name	Sequence <sup>a</sup> 5'→3'	mass cal /obs	CD $T_m$ °C( $K^+$ )
TBA	ggttggtgtggttgg	4726/4730	49.5
TBA-7T	ggttggT <sup>TNA</sup> gtggttgg	4756/4754	59.7
TBA-9T	ggttggtgT <sup>TNA</sup> ggttgg	4756/4754	34.3
TBA-7T( <i>iso</i> )	ggttggT <sup>iso</sup> gtggttgg	4726/4729	46.5
TBA-9T( <i>iso</i> )	ggttggtgT <sup>iso</sup> ggttgg	4726/4731	36.5

<sup>a</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site

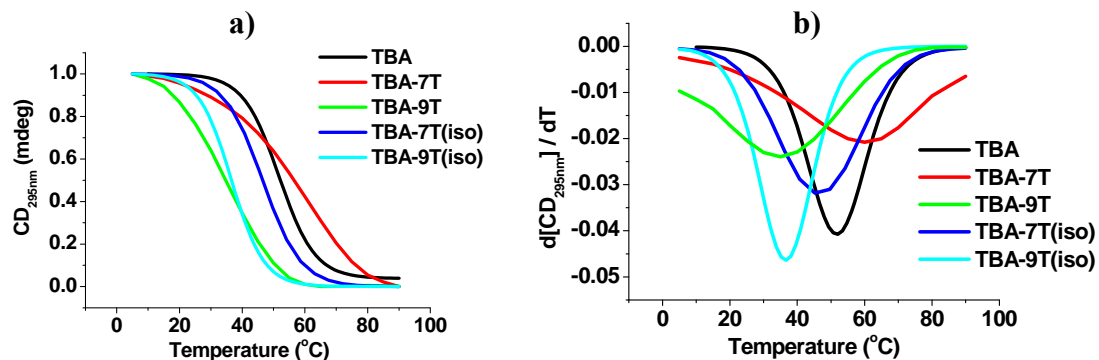


**Figure 6** CD spectra of oligomers **TBA**, **TBA-7T**, **TBA-9T**, **TBA-7T(*iso*)**, **TBA-9T(*iso*)** of  $5\mu\text{M}$  concentration in 10mM potassium phosphate buffer (pH 7.5) containing 100mM KCl at  $5^\circ\text{C}$ .

All sequences are showing the antiparallel G-quadruplex characteristic positive band at 295nm (Figure 6). The band intensity is high for **TBA-7T** and **TBA-7T (*iso*)** compared to others and **TBA-9T** is showing very less intensity band. After conforming of formation of quadruplex, we checked the stability by CD melting.

CD melting results showed that **TBA-7T** sequence was forming very stable antiparallel quadruplex structure and it is stabilizing by  $10^\circ\text{C}$  compared to **TBA**. The stability of **TBA-**

7T (*iso*) is as good as TBA and 9T position modified sequences were destabilizing the tetraplex stability (Figure 7).



**Figure 7** CD- $T_m$  of oligomers TBA, TBA-7T, TBA-9T, TBA-7T(*iso*), TBA-9T(*iso*) having a strand concentration of 5  $\mu$ M: (a) In 10mM Na-phosphate buffer (pH 7.5) containing 100mM KCl. (b) The first derivative curves of TBA, TBA-7T, TBA-9T, TBA-7T(*iso*), TBA-9T(*iso*) in K-phosphate buffers.

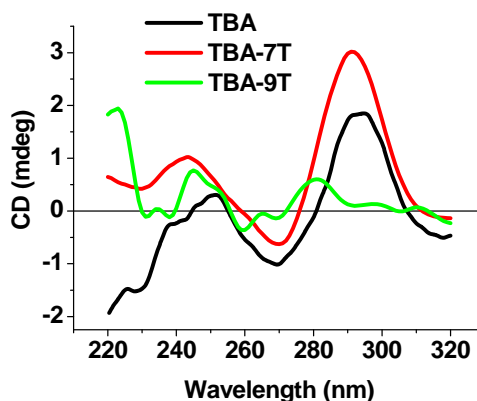
### CD spectroscopy and $T_m$ measurement of the TBA sequences in presence of Na<sup>+</sup> ion

The G-quadruplex formation of the TBA, TBA-7T and TBA-9T sequences in presence of Na<sup>+</sup> cations were studied by CD spectroscopy.

**Table 3** Quadruplex melting data of TBA and modified TBA sequences in presence of Na<sup>+</sup>

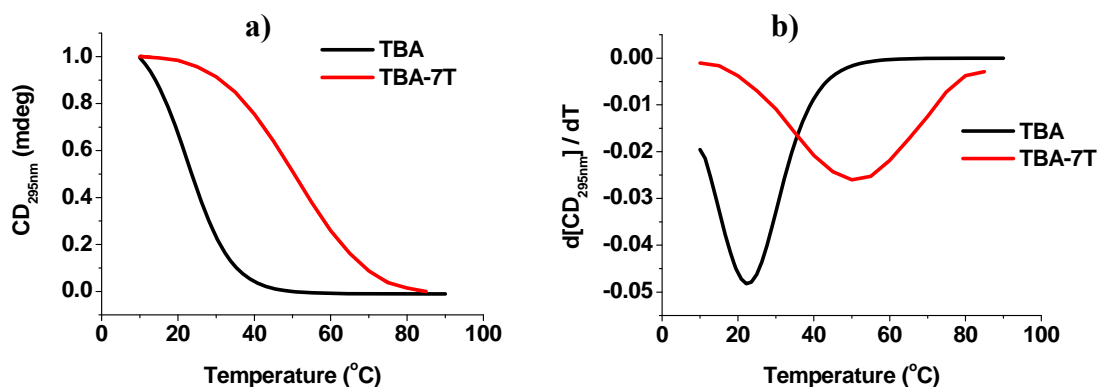
Name	Sequence <sup>a</sup> 5' → 3'	CD $T_m$ °C (Na <sup>+</sup> )
TBA	ggttggtgtggttg	22
TBA-7T	ggttggT <sup>TNA</sup> gtggttg	50.6
TBA-9T	ggttggtgT <sup>TNA</sup> ggttg	nd

<sup>a</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site



**Figure 8** CD spectra of oligomers **TBA**, **TBA-7T**, **TBA-9T** of  $5\mu\text{M}$  concentration in  $10\text{mM}$  potassium phosphate buffer (pH 7.5) containing  $100\text{mM}$  NaCl at  $5^\circ\text{C}$ .

In presence of sodium ions **TBA** and **TBA-7T** were only showing characteristic bands of antiparallel tetraplex (Figure 8). Then we studied the stability of tetraplex in presence of sodium ions. The melting results reveals that **TBA-7T** modification was stabilizing the antiparallel quadruplex structure by  $28^\circ\text{C}$  compared to **TBA** (Figure 9).

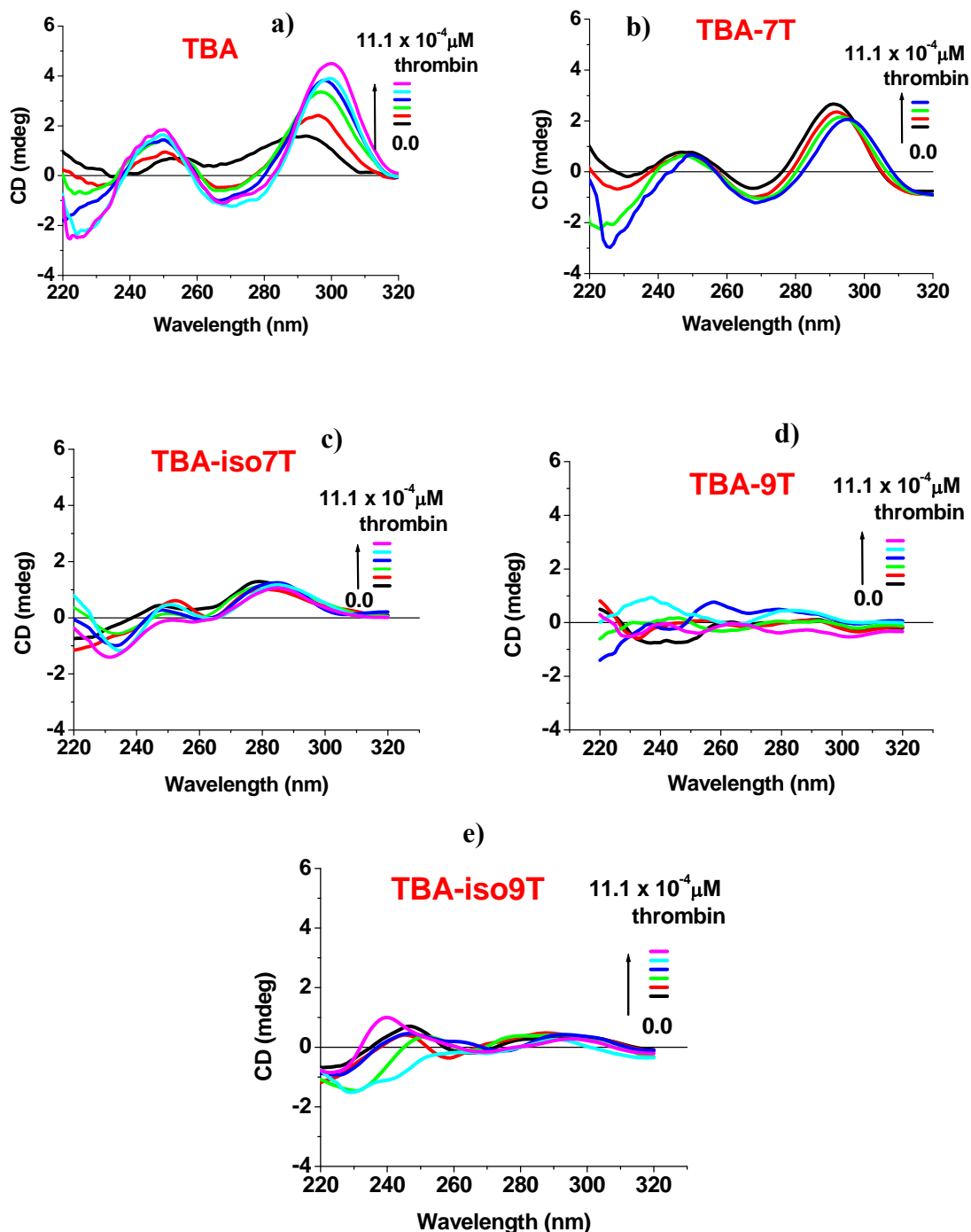


**Figure 9** CD- $T_m$  of oligomers **TBA**, **TBA-7T** having a strand concentration of  $5\mu\text{M}$ : (a) In  $10\text{mM}$  Na-phosphate buffer (pH 7.5) containing  $100\text{mM}$  NaCl. (b) The first derivative curves of **TBA**, **TBA-7T** in Na-phosphate buffers.

#### 4B.5 G-tetraplex formation in presence of thrombin

The **TBA** can fold in to a quadruplex structure in presence of thrombin due to chaperone effect. At low temperature, the quadruplex structure formed because of thrombin-binding, even in the absence of  $\text{K}^+$ . We performed CD experiments with **TBA**, **TBA-7T**, **TBA-9T**, **TBA-7T(iso)**, **TBA-9T(iso)** in the presence of increasing concentrations of thrombin at low temperature.





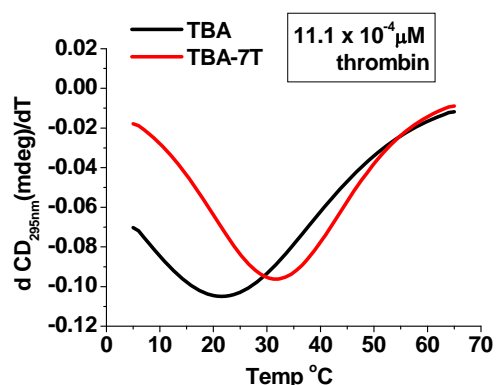
**Figure 10** Changes in CD signal at 295nm upon addition of thrombin to ( a) **TBA**, (b) **TBA-7T**, and (c) **TBA-9T** (d) **TBA-7T(iso)**, (e) **TBA-9T(iso)**

Among all five sequences **TBA**, **TBA-7T**, **TBA-7T(iso)** were forming the quadruplex structure in presence of thrombin and shifting of 295nm band was observed upon addition of thrombin (Figure 10). The stability was checked by temperature

dependent change in CD amplitude at 295nm. **TBA-7T** was stabilizing the quadruplex structure by 10 °C compared to **TBA** (Figure 11 & Table 4).

**Table 4** Quadruplex melting data of **TBA** and **TBA-7T** sequences in presence Thrombin

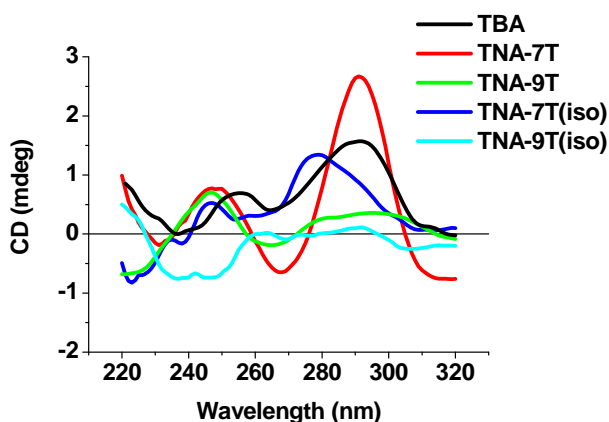
Name	Sequence <sup>a</sup> 5' → 3'	CD $T_m$ °C Thrombin
TBA	ggttggtgtggttgg	22
TBA-7T	ggttggT <sup>TNA</sup> gtggttgg	32



**Figure 11** The first derivative curves of **TBA**, **TBA-7T** in presence of thrombin.

#### 4B.6 CD spectroscopy and $T_m$ measurement of the TBA sequences in presence of water

We also checked the CD pattern of all TBA sequences in water and observed that only **TBA** and **TBA-7T** were showing CD positive band at 295nm (Figure 12). **TBA-7T(iso)** is showing 280nm positive band which is not corresponding to tetraplex. CD melting experiment done to check the stability of quadruplex structures formed by **TBA**, **TBA-7T**. The results showed that **TBA-7T** was stabilized the tetraplex structure by 6 °C compared to **TBA** (Table 5 & Figure 13).

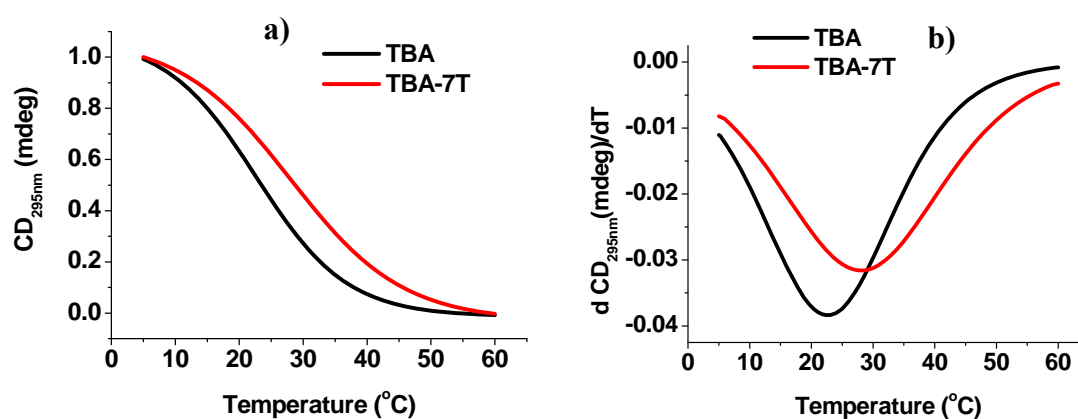


**Figure 12** CD spectra of oligomers **TBA**, **TBA-7T**, **TBA-9T**, **TBA-7T(iso)**, **TBA-9T(iso)** of 5μM concentration in water at 5°C.

**Table 5** Quadruplex melting data of **TBA** and **TBA-7T** sequences in water

Name	Sequence <sup>a</sup> 5' → 3'	CD $T_m$ °C water
TBA	ggttggtgtggttgg	22
TBA-7T	ggttggT <sup>TNA</sup> gtggttgg	28

<sup>a</sup>The lower case letters indicate unmodified DNA and upper case indicate modified site



**Figure 13** CD- $T_m$  of oligomers **TBA**, **TBA-7T** having a strand concentration of  $5\mu\text{M}$ : (a) in water. (b) The first derivative curves of **TBA**, **TBA-7T** in water.

#### 4B.7 Imino proton NMR spectra of TBA sequences

In proton NMR the imino proton signals between 11.5ppm to 12.5ppm range are the characteristic signals of eight H-bonds formed between the guanines of each G-quartet of a quadruplex structure.<sup>13</sup> The imino proton chemical shifts for **TBA** and **TBA-7T** indicate the hydrogen-bonded quadruplex formation in presence of potassium ions (Figure 14a). In absence of potassium ions also **TBA** and **TBA-7T** were forming the quadruplex structure (Figure 14b). The **TBA** NMR spectra in water (Figure 14b) suggesting that the broad peaks between 10.5-11.1ppm corresponds to the protons, which not involved in the H-bonded quadruplex structure. Whereas in case of **TBA-7T** these broad peaks were not observed. It indicates the formation of **TBA-7T** stable quadruplex

structure in absence of  $K^+$  ions. This NMR supports the high stability of TBA-7T quadruplex structure than TBA in water (CD melting, Table 5).

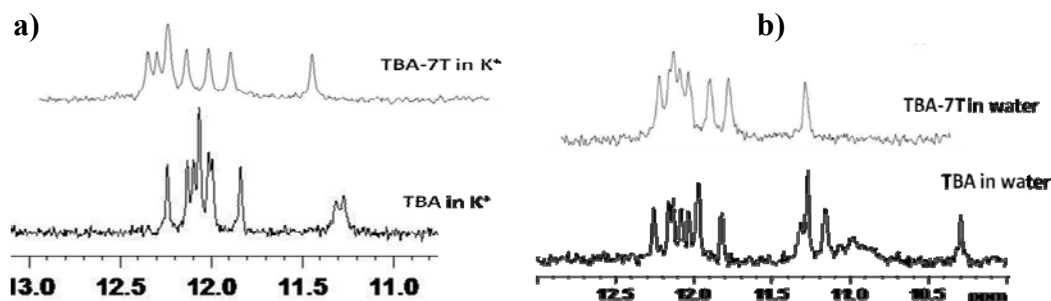
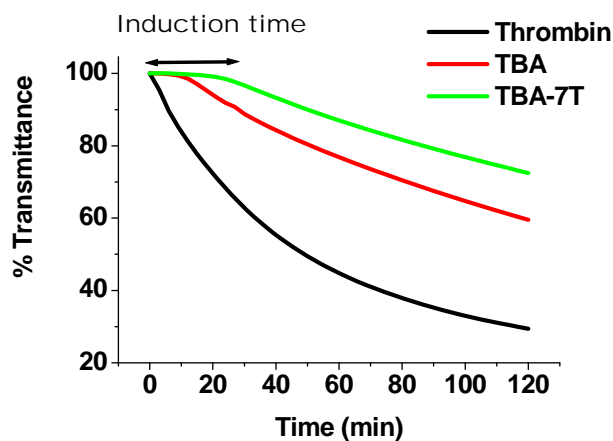


Figure 14 Quadruplex imino proton spectral region for TBA and TBA-7T in (a) presence of  $K^+$  and (b) water at 4°C

#### 4B.8 Anti-thrombin activity measurements

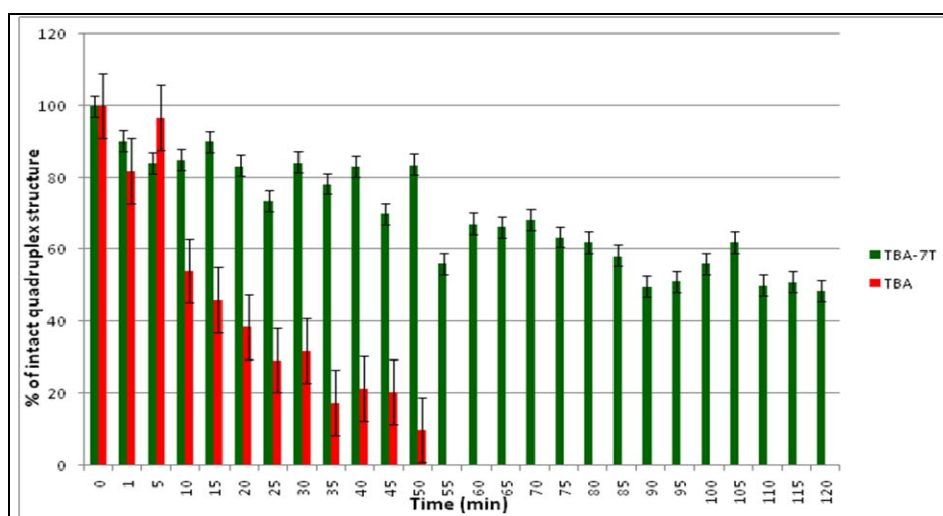
The anti-thrombin activity of all the synthesized TBA sequences on thrombin-catalyzed conversion of fibrinogen to fibrin (clotting) was investigated by measuring the percent transmittance with time. In absence of TBA, fibrinogen was getting coagulated very fast with no induction time. Whereas presence of TBA slowed down the coagulation with an increased induction time ( $t_i$  as coagulation parameter), confirming its reported inhibitory activity.<sup>14</sup> The induction time for the TBA-7T was higher than for TBA, this high induction time could give a large window to reduce the concentration of TBA in acceptable therapeutic range. These experiments provide conclusive evidence that the TBA-7T oligomers are indeed capable of not only forming G-quadruplex structures but also hold similarity in structural topology capable of taking active part in the assigned function of the TBA (Figure 15).



**Figure 15** Antithrombin activity measured by % transmittance at 450nm in the presence of **TBA** and **TBA-7T** and % transmittance Vs wavelength plots at different time-points of the study. ↔ indicates induction time as coagulation parameter ( $t_i$ ).

#### 4B.9 Stability of quadruplex structure of aptamers to SVPD

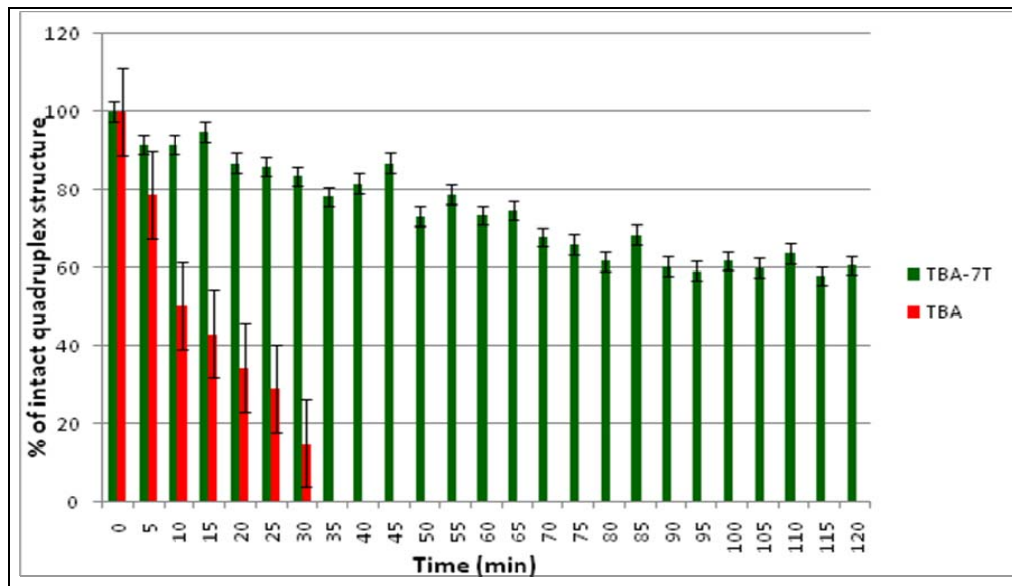
We studied the stability of **TBA** and **TBA-7T** quadruplex structures against **SVPD** enzyme.<sup>15</sup> The stability of **TBA-7T** was found to be very high compared to the control **TBA**. The reaction was monitored by change in CD amplitude at 295nm.



**Figure 16** Quadruplex stability of the aptamers **TBA** and **TBA-7T** (7.5 $\mu$ M) towards Snake venom phosphodiesterase (**SVPD**) enzyme (2.5 mg/mL).

The half-life of **TBA-7T** quadruplex structure was found to be ~120min (figure15) while that of **TBA** was found to be ~15min at 37°C (Figure 16). The observed higher stability of the **TBA-7T** offers obvious advantages for applications in biological systems, where the control unmodified oligomer has a relatively low half-life.

We also studied the stability of **TBA** and **TBA-7T** quadruplex structures against **SVPD** enzyme in presence of thrombin. The quadruplex stability of **TBA** was reduced in presence of thrombin and the half-life was found to be 10min, whereas in case of **TBA-7T** half life was slightly increased compared to the absence of thrombin (figure 17).



**Figure 17** Quadruplex stability of the aptamers **TBA** and **TBA-7T** (7.5 $\mu$ M) towards Snake venom phosphodiesterase (SVPD) enzyme (2.5 mg/mL) in presence of thrombin.

### 4B.10 Conclusion

- We introduced a new synthetic method for chemically modified TNA nucleoside from commercially available cheap starting material.
- Successfully synthesized 4'-MOM TNA modified DNA sequences and stability studies resulted that 4'-MOM TNA: RNA duplexes are more stable than 4'-MOM TNA: DNA duplex.
- The aptamers (**TBA-7T**, **TBA-9T**) studied in this work are the first examples of TNA modified sequences which are able to form G-quadruplexes similar in structure to the well established TBA aptamer which was used for comparison.
- **TBA-7T** resulted in a more stable quadruplex than all other modified as well as unmodified aptamers.
- **TBA-7T** aptamer formed quadruplex not only in the presence of monovalent cations like Na<sup>+</sup> and K<sup>+</sup> but also in water.
- Quadruplex structure formations were confirmed by NMR studies
- It exhibited binding to thrombin in absence of cations at low temperature.

### 4B.11 Experimental Section

All the general experimental procedures were already discussed in (chapter 3).

#### **Quadruplex stability studies in presence of snake venom phosphodiesterase enzyme experiments**

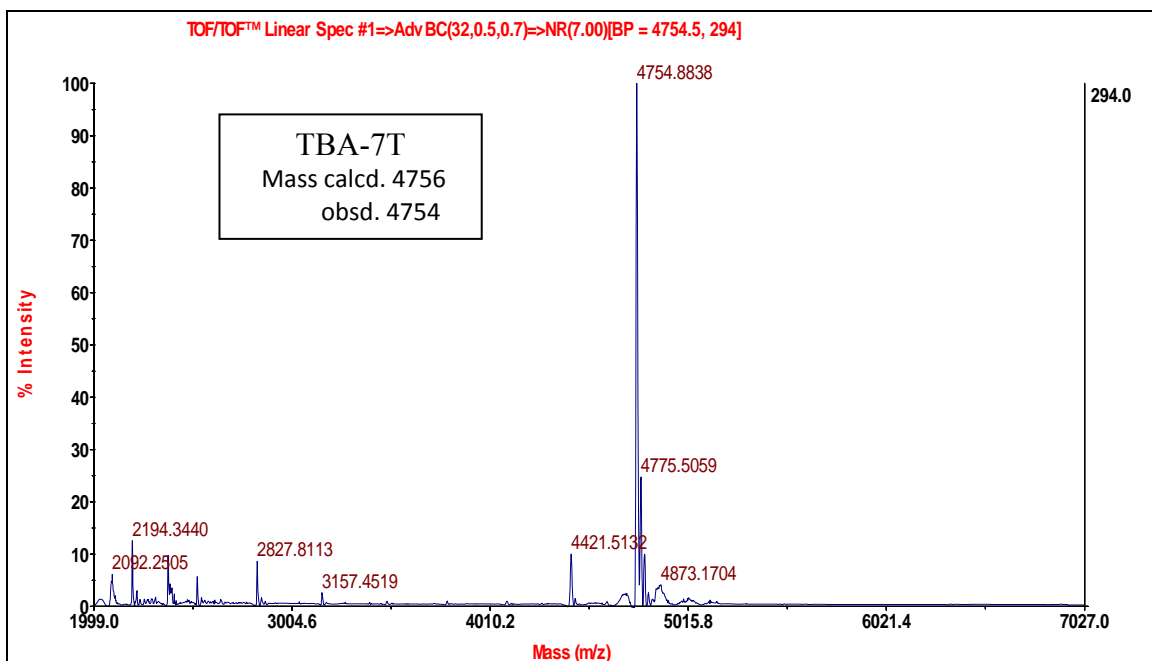
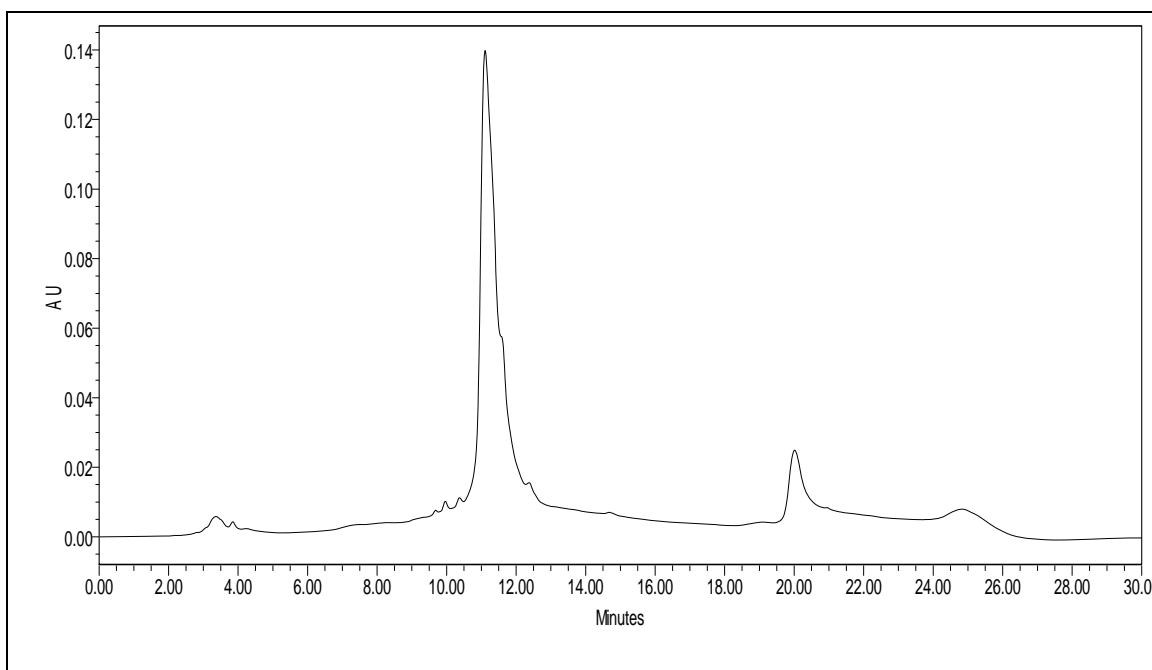
Quadruples stability studies of the **TBA** and **TBA-7T** (7.5  $\mu\text{M}$ ) was carried out at 37 °C in buffer (100  $\mu\text{l}$ ) containing 100 mM Tris-HCl (pH 8.5), 15 mM  $\text{MgCl}_2$ , 100 mM NaCl and SVPD (25 $\mu\text{g}/\text{mL}$ ). CD measured from 320nm to 220nm and the stability of quadruplex was monitored by change in CD amplitude at 295nm positive band at several time points. The percentage of CD amplitude was plotted against the exposure time to obtain the stability with time.



**4B.12 Appendix**

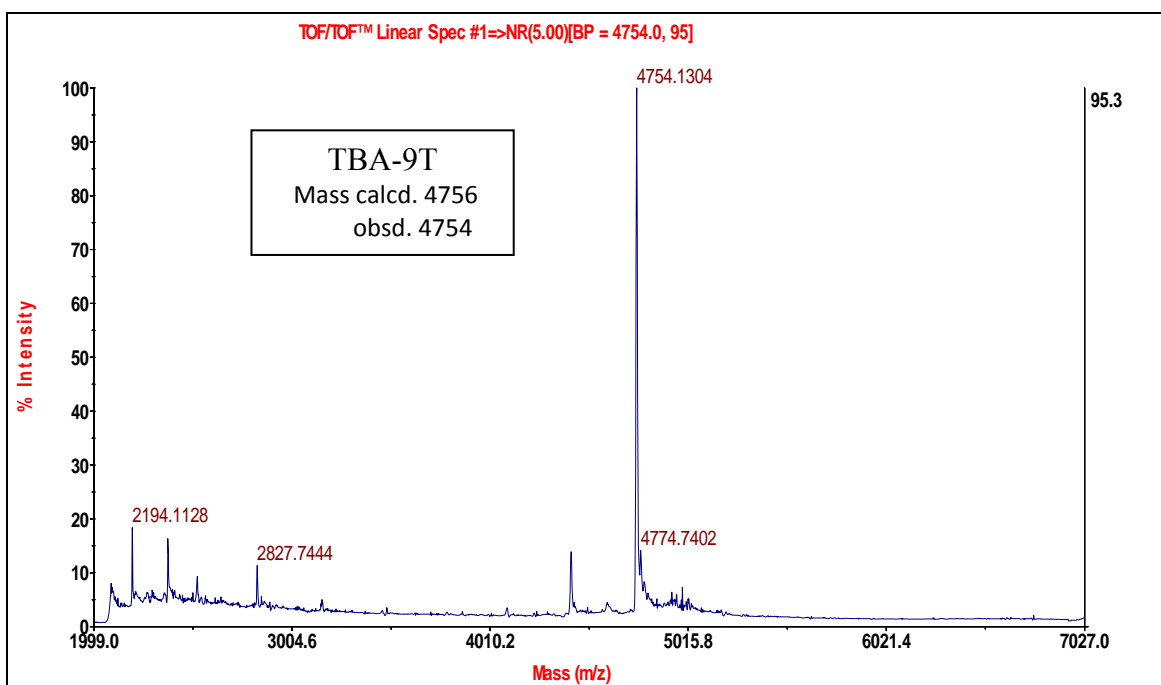
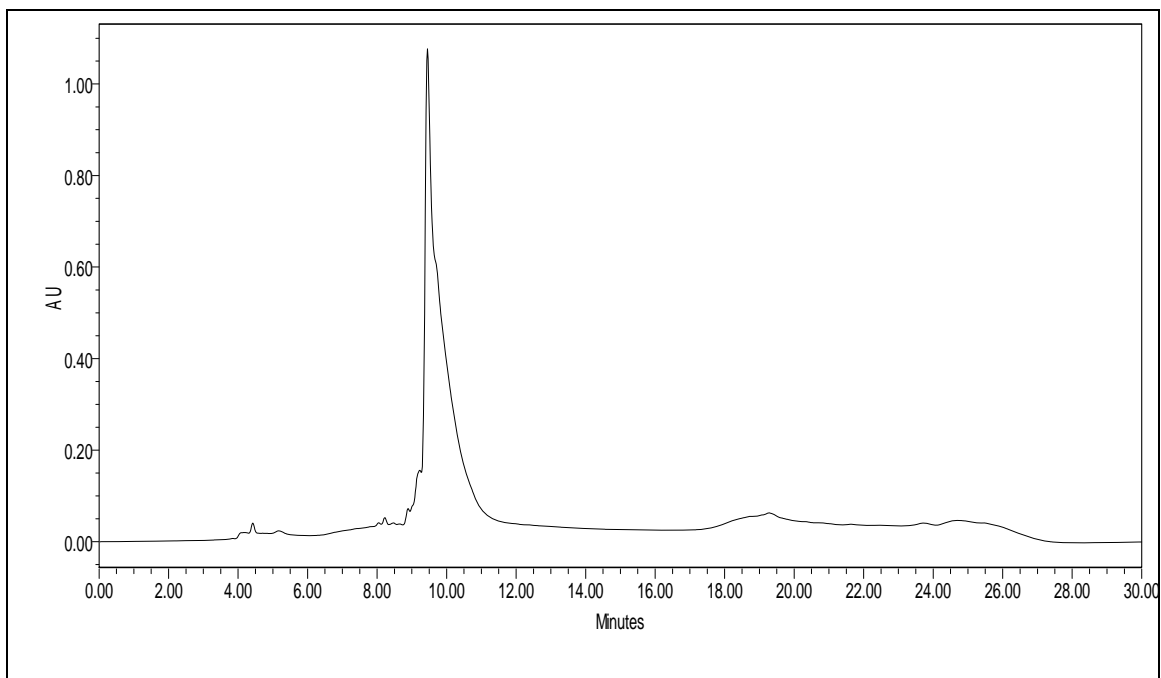
<b>Compounds - Spectral data</b>	<b>Page No.</b>
HPLC & MALDI-TOF of TBA-7T	212
HPLC & MALDI-TOF of TBA-9T	213
HPLC & MALDI-TOF of TBA-7T (iso)	214
HPLC & MALDI-TOF of TBA-9T (iso)	215

HPLC & MALDI-TOF of TBA-7T :



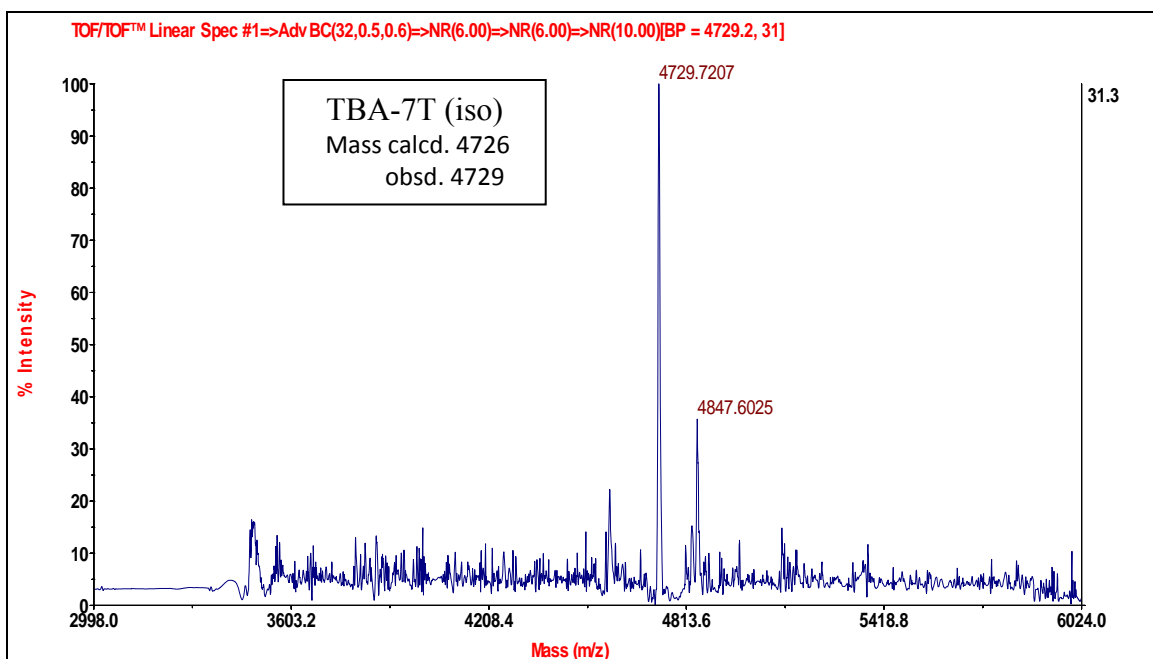
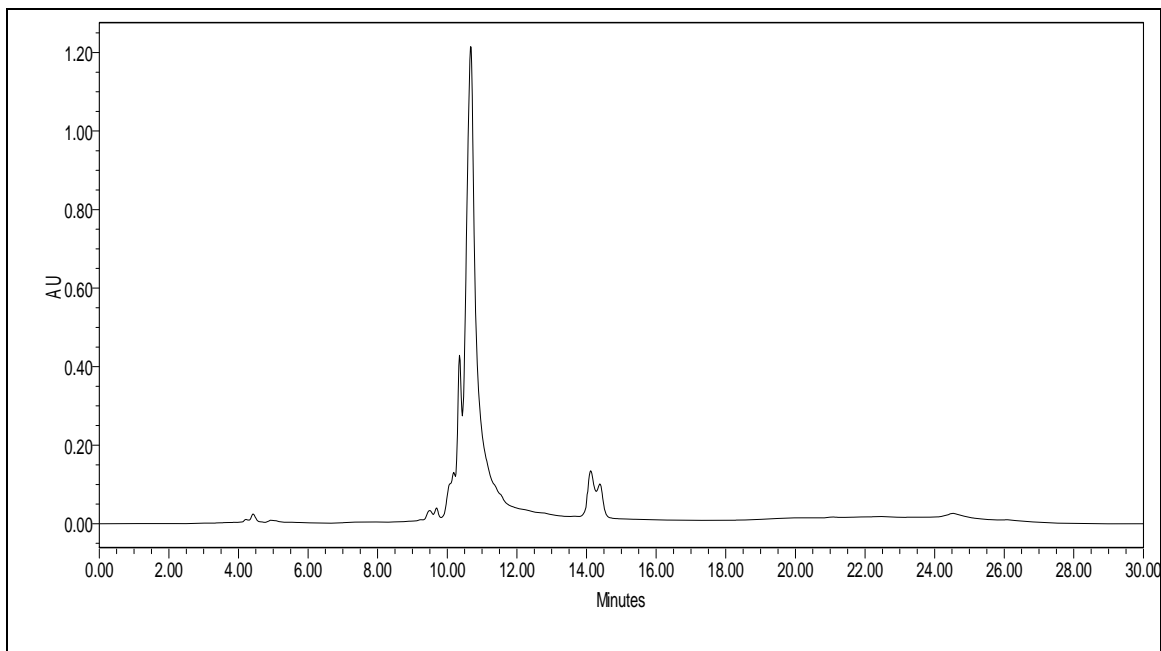
## Chapter 4

HPLC & MALDI-TOF of TBA-9T :



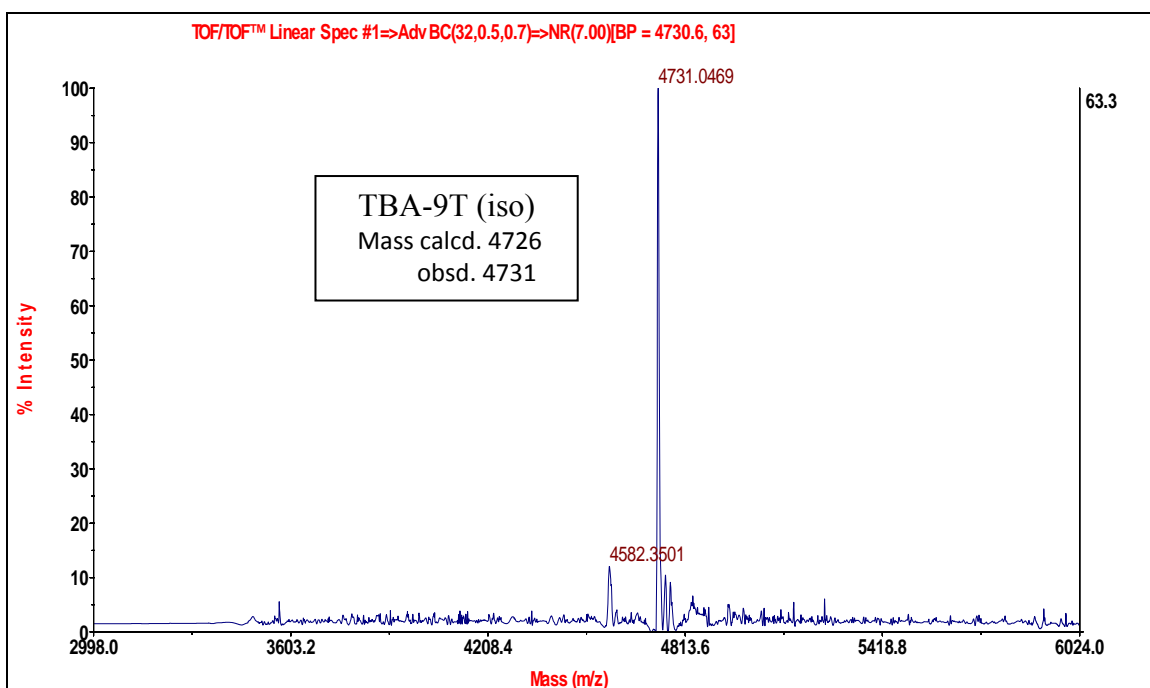
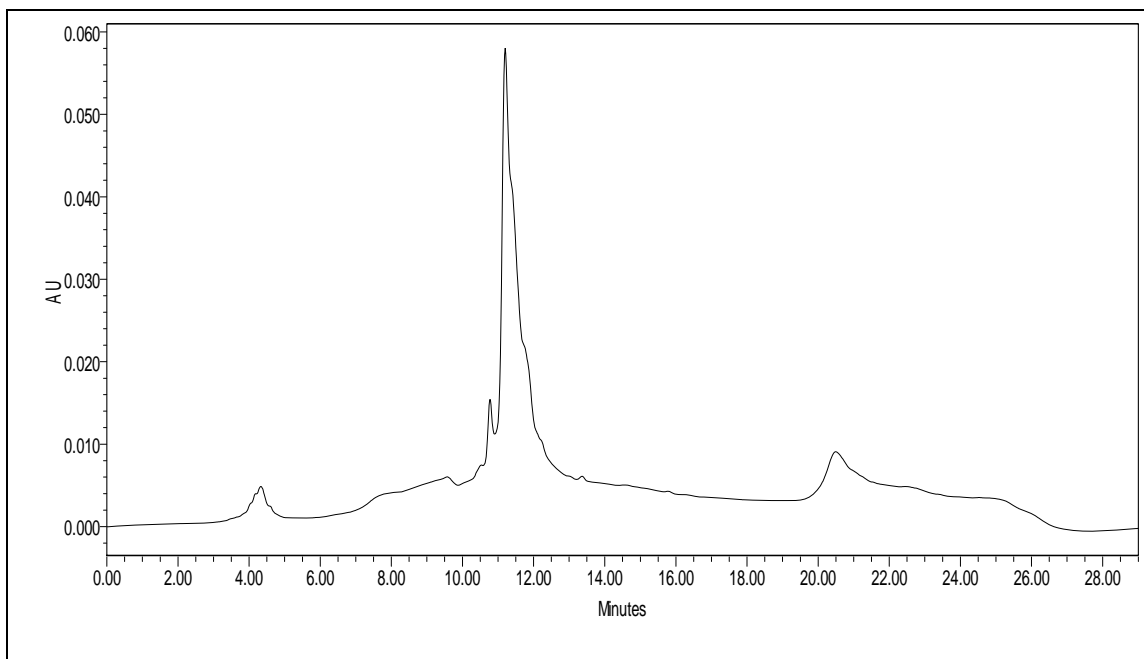
## Chapter 4

HPLC & MALDI-TOF of TBA-7T (iso) :



## Chapter 4

HPLC & MALDI-TOF of TBA-9T (iso) :



### 4B.13 References

1. Schoning, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A., *Science* **2000**, 290, 1347-1351.
2. Orgel, L., *Science* **2000**, 290, 1306-1307.
3. Brudno, Y.; Birnbaum, M. E.; Kleiner, R. E.; Liu, D. R., *Nat. Chem. Biol.* **2010**, 6, 148–155.
4. Yu, H.; Zhang, S.; Chaput, J. C., *Nat. Chem.* **2012**, 4, 183–187.
5. Steiger M., Reichstein T., *Helv. Chim. Acta* **1936**, 19, 1016-19; Katzi, T. Reichstein, *Helv. Chim. Acta* **1938**, 21, 195; Perlin A. S. in –Methods in Carbohydrate Chemistry, Vol. 1, Eds. R. L. Whistler and M. L. Wolfram, Academic Press: New York, **1962**, p. 68, and ref. cit. therein.
6. Kai-Uwe Sch<sup>^</sup>ning, Peter Scholz, Xiaolin Wu, Sreenivasulu Guntha, Guillermo Delgado, Ramanarayanan Krishnamurthy, Albert Eschenmoser, *Helv. Chim. Acta* **2002**, 85, 4111-4153.
7. Ellington, A. D.; Szostak, J. W. *Nature* **1990**, 346, 818-822.
8. Coughlin, SR., *Nature*, **2000**, 407, 258–64.
9. Veronica Esposito, Maria Scutto, Antonella Capuozzo, Rita Santamaria, Michela Varra, Luciano Mayol, Antonella Virgilio, Aldo Galeone., *Org. Biomol. Chem.*, **2014**, 12, 8840–8843.
10. Barbara Sacca, Laurent Lacroix, Jean-Louis Mergny., *Nucleic Acids Res*, **2005**, 33, 1182-1192.
11. Pasternak, A.; Hernandez, F. J.; Rasmussen, L. M.; Vester, B.; Wengel, J., *Nucleic Acids Res.*, 2010, 39, 1155-1164.
12. Gunjal A. D., Fernandes M., Erande N., Rajamohanan P. R., Kumar V. A.. *Chem. Commun*, **2014**, 50, 605-607.
13. Aviñó, A., Mazzini, S., Ferreira, R., Gargallo, R., Marquez, V.E., Eritja, R., *Bioorg. Med. Chem.*, **2012**, 20, 4186-4193.

## Chapter 4

---

14. Uehara, S., Shimada, N., Takeda, Y., Koyama, Y., Takei, Y., Ando, H., Satoh, S., Uno, A., Sakurai K. *Bull. Chem. Soc. Jpn.*, **2008**, *81*, 1485–1491.
15. Erande, N., Gunjal, A. D., Fernandes, M., Gonnade R., Kumar V. A., *Org. Biomol. Chem.*, **2013**, *11*, 746-757.