

**Charged Peptide Nucleic Acid Analogues:
Ethano-locked PNA and β -, γ - bisubstituted PNA**

**Thesis submitted to
Savitribai Phule Pune University
for the degree of**

**Doctor of Philosophy
in
Chemistry**

BY

ANJAN BANERJEE

Research Supervisor

Dr. (Mrs.) Vaijayanti A. Kumar

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

September 2014



सीएसआयआर-राष्ट्रीय रासायनिक प्रयोगशाला

(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)

डॉ. होमी भाभा मार्ग, पुणे - 411 008. भारत

CSIR-NATIONAL CHEMICAL LABORATORY

(Council of Scientific & Industrial Research)

Dr. Homi Bhabha Road, Pune - 411008. India



CERTIFICATE

This is to certify that the work presented in the thesis entitled “**Charged Peptide Nucleic Acid Analogues: Ethano-locked PNA and β -, γ - bisubstituted PNA**” submitted by Mr. Anjan Banerjee, 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

September 2014


(Research Supervisor)

Division of Organic Chemistry

CSIR-National Chemical Laboratory

Pune 411008

Communication
Channels


NCL Level DID : 2590
NCL Board No. : +91-20-25902000
EPABX : +91-20-25893300
: +91-20-25893400

FAX

Director's Office : +91-20-25902601
COA's Office : +91-20-25902660
COS&P's Office : +91-20-25902664

WEBSITE

www.ncl-india.org



CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled “**Charged Peptide Nucleic Acid Analogues: Ethano-locked PNA and β -, γ - bisubstituted PNA**” 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.

Anjan Banerjee

September 2014

CSIR-National Chemical Laboratory

Pune- 411 008

Dedicated to...

My Parents



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.

*It gives me an immense pleasure to express my deep sense of gratitude towards my research guide **Dr. (Mrs.) Vijayanti A. Kumar** for all the advice, guidance, support and encouragement during every stage of this work. She made me realize the importance of doing quality research; she taught me each and every aspect of research, from working table to formulation of ideas to presentation of results. The confidence she had in me, willingness to share new ideas, excitements helped me in a real sense to shape my research career. Although this eulogy is insufficient, I preserve an everlasting gratitude for her.*

Dr. Moneesha Fernandes deserves special thanks. Her encouragement, understanding and suggestions were invaluable during my stay in the lab and went a long way towards the completion of this thesis.

My special thanks to Dr. (Mrs.) Anita Gunjal, Dr. (Mrs.) Vandana S. Pore, Dr. (Mrs.) Radhika S. Kusurkar and Dr. Dilip D. Dhavale for their advice, various kinds of help and encouragement during the course of this work. I am also thankful to Mrs. M. V. Mane for HPLC analysis, and also special thanks to Mrs. Shantakumari for the MALDI-TOF and HRMS experiments. The kind support from NMR group is greatly acknowledged.

I admire the co-operation of my seniors and colleagues who have taught me many things. I thank Dr. Nilkanth Aher, Dr. Sachin Gokhale, Dr. V. Madhuri, Dr. Seema Bagmare, Dr. Namrata Erande, Dr. Kiran Patil, Dr. Venubabu Kotikam, Harshit K. Soni, Manojkumar Varada, Tanaya Bose, Ragini Awachat, Manisha Aher, Govind Bhosle, Amit Kumar Yadav, Harsha C. and Gandhali Kulkarni for their support and help, who made workplace fun, lively and made each other forget their worries even for a while and sometimes forever. I thank Mr. P. Bhumkar for the laboratory assistance.

I wish to thank all my fellow colleagues in NCL, Ramesh Batwal, Kailash Pawar, Dr. Menaka Pandey, Chandani Singh, Ashish Chinchansure, Dr. Swati Kolet, Alok Ranjan, Rohan Erande, Avinash Bansode, Amol Jadav, Janakiram V., Deepak Jadav, Vijay Beniwal, Amit Nagare, Dr. Nagesh Kshupse, Vannuruswamy G.,.....and many more, for their cheerful company and making my NCL life very lively and enjoyable. I also wish thanks to Dr. Tanpreet Kaur, Dr. Deepak Jain, Nitin Bansode, and Vijay Kadam from IISER Pune.

I am indebted to my GJ Hostel Friends, Atul More, Vasudevan N., Manoj Sharma, Sanjay Negi, Tamboli Majid, Pankaj Daramwar, Jeetendra Rout, Raju Nada, Ankush Bhise, Shivakumar KI, R.Lenin, Ashok Kumar, Brijesh Sharma, Bhausaheb Tawade, Balanagulu B., Venugopal E....

I extend huge thanks to my bong friends Pravat Mondal, Kanak Roy, Saikat Halder, Susanta Das, Mrinmoy Chini, Anup Bhunia, Santigopal Mondal, Atanu Patra, Arpan Manna, Achintya Dutta, Arrya Ghosh, Sudip Sasmal, Hriday Agarwal, Himadri Pathak, Arijit Mallick, Subhadip Das, Prathit Chatterjee, Abhik Banerjee, Suman Chandra, Dr. Parthasarathi Chowdhury,

Dr. Sumanta Garai, Dr. Chandan Dey, Dr. Tamas Panda, Binoy Majumder, Dr. Debasis Pati, Dr. Sumantra Bhattacharya and Late Mr. Agnimitra Banerjee.

I am grateful to Council of Scientific and Industrial Research, Government of India, for awarding the junior and senior research fellowships and Dr. S. Sivaram, former Director, and Dr. Sourav Pal, Director, CSIR - National Chemical Laboratory to carry out my research works, extending all infrastructural facilities and to submit this work in the form of a thesis for the award of Ph. D degree.

Finally, it has been a difficult task to capture and express my feelings for my family members. I have no words to express my gratitude to my Parents and relatives, without knowing much what I am doing exactly, just wishing me all the time with no expectations. Their patience and sacrifice were always a main source of inspiration and will remain throughout my life. Their blessings and encouragement have always made me an optimist in any unknown areas I had ventured.

Above all, my acknowledgement would not be completed without thanking the Almighty for giving me the strength and the determination to overcome the hardship faced in my life.

Anjan Banerjee

Table of Contents

List of Publications/ Symposia/ Workshops Attended/ Poster Presentations		<i>i</i>
Abbreviations		<i>iii</i>
Abstract		<i>vi</i>
Chapter 1		
Recent advances in the field of nucleic acids: An Introduction		
1A	Introduction to Nucleic Acids	1
	1A.1 Components and primary structure of nucleic acids	1
	1A.2 Sugar puckering in nucleos(t)ides	2
	1A.3 Base Pairing via Hydrogen bonding	3
	1A.4 Structures of Nucleic acids	4
	1A.4.1 The duplex structure	4
	1A.4.2 Triplex-Forming Oligonucleotides (TFO)	7
	1A.4.3 Quadruplex structure – The ‘G-quadruplex’	8
	1A.5 Spectroscopic technique to study DNA/RNA interactions	9
	1A.5.1 Ultraviolet spectroscopy (UV)	9
	1A.5.2 Stoichiometry	11
	1A.5.3 Circular Dichroism (CD)	11
1B	Applications of Nucleic Acids	12
	1B.1 Introduction	12
	1B.2 Oligonucleotides as antisense therapeutic agent	13
	1B.3 Antisense Mechanisms	13
	1B.3.1 Antisense Oligonucleotides: Disruptive antisense	14
	1B.3.2 Antisense Oligonucleotides: Corrective antisense	16
1C	Chemical Modifications of DNA	17
	1C.1 First generation antisense oligonucleotides	17
	1C.1.1 Alternative phosphate containing linkages	17
	1C.1.2 Non-phosphorus backbone	18
	1C.2 Second generation antisense oligonucleotides	19
	1C.3 Third generation antisense oligonucleotides	20
1D	Peptide Nucleic Acids (PNA)	21
	1D.1 Introduction	21
	1D.2 Chemical and physical properties of PNA	22
	1D.2.1 Duplex formation with comp. oligonucleotides	22
	1D.2.2 Triplex formation of PNA	22

	1D.2.3	Quadruplex formation of PNA	23
1D.3		Structures of PNA-DNA and PNA-RNA complexes	24
1D.4		Applications of PNA	26
	1D.4.1	Antisense effect or inhibition of translation	26
	1D.4.2	Antigene effect or inhibition of transcription	26
	1D.4.3	Inhibition of Replication	27
	1D.4.4	Interactions with Enzymes	27
	1D.4.5	Nucleic acid purification	30
1E		Chemical Modifications of PNA	30
	1E.1	Preorganization through conformational constraints	31
	1E.2	Preorganization through rigid five membered heterocycles	33
	1E.3	Preorganization through rigid six membered heterocycles	36
	1E.4	Modified Nucleobases	36
	1E.5	Stability in cells and cellular uptake of PNA	37
	1E.5.1	PNA- DNA Chimeras	38
	1E.5.2	Conjugation of PNA with Cationic Peptides	39
	1E.5.3	Cationic Backbone Modifications of PNA	41
1E		The Present Work	42
1F		References	44
Chapter 2			
Novel ethano locked PNA (ethano-PNA)			
<u>Section A</u>			
Synthesis of monomer and incorporation in <i>aeg</i> PNA sequences			
2A		Introduction	53
	2A.1	Rationale and objectives of the present work	54
	2A.2	Synthesis of ethano locked monomers	55
	2A.2.1	Synthesis of Boc protected <i>H</i> -ethano-PNA T monomer	55
	2A.2.2	Synthesis of Fmoc/Boc protected <i>Am</i> -ethano-PNA T monomer	56
	2A.3	Solid Phase PNA Synthesis	57
	2A.3.1	Solid Phase Synthesis of oligomers incorporating modified PNA monomers	59
	2A.3.2	Post synthetic conversion of amino to guanidino group on solid phase	60
	2A.3.3	Cleavage of the PNA Oligomers from the Solid Support	61

	2A.3.4	Purification and MALDI-TOF characterization of oligomers	62
	2A.4	Summary	62
	2A.5	Experimental	63
	2A.6	Appendix	76
<u>Section B</u>			
Biophysical studies of ethano-PNA towards comp. DNA/RNA			
2B	Introduction		113
	2B.1	Rationale and objectives of the present work	113
	2B.2	Biophysical spectroscopic techniques for studying PNA-DNA/ PNA-RNA interactions	113
	2B.3	UV-T_m studies on ethano-PNA:cDNA and ethano-PNA:RNA	114
	2B.3.1	Effect of salt concentration on UV- T_m of ethano-PNA:cDNA/RNA duplexes	117
	2B.3.2	UV- T_m studies on ethano-PNA with mismatch RNA	119
	2B.4	Conformational analysis of the ethano-PNA:cDNA and ethano-PNA:RNA	121
	2B.5	Experimentation	123
	2B.6	Summary	124
	2B.7	References	125
Chapter 3			
β-, γ- bisubstituted PNA			
<u>Section A</u>			
Synthesis of monomers and incorporation in <i>aeg</i> PNA sequences			
3A	Introduction		129
	3A.1	Rationale and objectives of the present work	130
	3A.2	Synthesis of β- and γ- bisubstituted PNA T/A^{cbz} monomer	134
	3A.3	Solid Phase PNA Synthesis	135
	3A.3.1	Solid Phase PNA Synthesis with β - and γ - bisubstituted T/A ^{cbz} monomer	135
	3A.3.2	Cleavage of the PNA Oligomers from the Solid Support	135
	3A.3.3	Purification and MALDI-TOF characterization of oligomers	136
	3A.4	Summary	136

	3A.5	Experimental	137
	3A.6	Appendix	147
Section B			
Biophysical studies of modified PNA towards comp. DNA/RNA			
3B	Introduction		175
	3B.1	Rationale and objectives of the present work	175
	3B.2	Biophysical spectroscopic techniques for studying PNA-DNA/ PNA-RNA Interactions	175
	3B.3	UV-T_m studies on modified PNA:cDNA and modified PNA:RNA	176
	3B.3.1	Effect of salt concentration on T_m of modified-PNA:cDNA/RNA duplexes	178
	3B.3.2	UV- T_m studies of modified PNA with a single mismatch RNA	182
	3B.4	Conformational analysis of the modified PNA:cDNA and cRNA	183
	3B.5	Summary	185
	3B.6	Experimentation	186
	3B.7	References	187
Chapter 4			
Polyamide DNA			
Synthesis and biophysical studies of polyamide-DNA with alternating α -amino acid and nucleoside- β -amino acids			
4	Introduction		191
	4.1	Rationale and objective of the present work	191
	4.2	Synthesis of nucleoside-β-amino acid monomers	192
	4.3	Conformations in nucleoside-β-amino acids	194
	4.4	Synthesis of polyamide DNA oligomers	194
	4.5	Biophysical Studies of oligomers with DNA / RNA	196
	4.5.1	UV- T_m Studies of oligomers with antiparallel complementary DNA / RNA	196
	4.5.2	UV- T_m Studies of oligomers with parallel complementary DNA / RNA	198

	4.5.3	UV- T_m Studies of oligomers with single mismatch DNA / RNA	200
	4.6	Binding Stoichiometry: Job's Plot of the TRT1 oligomer	201
	4.7	Conformational analysis of the polyamide DNA:cDNA/RNA	202
	4.7.1	Conformational analysis of the polyamide CT oligomer	203
	4.7.2	Conformational analysis of the mixed purine-pyrimidine oligomer	203
	4.8	Summary	204
	4.9	Experimentation	205
	4.10	Appendix	207
	4.11	References	213

List of Publications

1. *C^{3'}-endo-puckered pyrrolidine containing PNA has favourable geometry for RNA binding: Novel ethano locked PNA (ethano-PNA)* **Anjan Banerjee**, Vaijayanti A. Kumar* *Bioorganic & Medicinal Chemistry*, (2013), **21**, 4092-4101
2. *Synthesis of all four nucleoside-based β -amino acids as protected precursors for the synthesis of polyamide-DNA with alternating α -amino acid and nucleoside- β -amino acids.* Seema Bagmare, Manojkumar Varada, **Anjan Banerjee**, Vaijayanti A. Kumar* *Tetrahedron*, (2013), **69**, 1210-1216
3. *Investigation of the effect of amino acid chirality in the internucleoside linker on DNA:DNA and DNA:RNA duplex stability* Seema Bagmare, **Anjan Banerjee**, Vaijayanti A. Kumar* (Communicated)
4. *New analogue of PNA: bisubstitution at β and γ positions in PNA backbone* **Anjan Banerjee**, Tanaya Bose, Vaijayanti A. Kumar* (manuscript under preparation)

Symposia/ Workshops Attended/ Poster Presentations

1. Presented poster in **20th ISCB International Conference**, at Department of Chemistry, University of Delhi, in March 2014.
2. Attended **Frontiers in Chemistry and Biology of Oligosaccharides** at Indian Institute of Science Education and Research, Pune, India, in January 2014.
3. Presented poster in **14th Tetrahedron Symposium: Challenges in Organic and Bioorganic Chemistry** at Vienna, Austria, in June 2013.
4. Presented poster in **International Meeting on Chemical Biology** at Indian Institute of Science Education and Research, Pune, India, in May 2013.
5. Attended "**5-Days Hands on Workshop on Molecular Biotechnology and Bioinformatics**" at International Centre for Stem Cells, Cancer and Biotechnology, Pune, India, in May 2013.
6. Presented poster during **National Science Day celebrations** at CSIR-National Chemical Laboratory, Pune, India, in February 2013.

-
7. Presented poster during **8th J-NOST Conference** at Indian Institute of Technology Guwahati, Assam, in December 2012.
 8. Attended "**ACS on campus**" event held at CSIR-National Chemical Laboratory, Pune, India, in October 2012.
 9. Presented poster during **National Science Day celebrations** at CSIR-National Chemical Laboratory, Pune, India, in February 2012.
 10. Participated in "**1st CRSI Zonal Meeting**" held at CSIR-National Chemical Laboratory, Pune, India, in May 2011.
 11. Presented poster during **National Science Day celebrations** at CSIR-National Chemical Laboratory, Pune, India, in February 2011.

Abbreviations

A	Adenine
Ac	Acetyl
Ac ₂ O	Acetic anhydride
<i>aeg</i>	Aminoethylglycine
<i>ap</i>	Antiparallel
arg	Arginine
aq.	Aqueous
BAIB	[Bis(acetoxy)iodo] benzene
Bz	Benzoyl
C	Cytosine
Calc.	calulated
Cat	Catalytic/catalyst
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
2'-dA	Deoxyadenosine
DCA	Dichloroacetic acid
DCM	Dichloromethane
dG	2'-Deoxyguanosine
DIPEA/DIEA	Diisopropylethylamine
DMAP	4',4'-Dimethylaminopyridine
DMF	<i>N,N</i> -dimethylformamide
DMSO	<i>N,N</i> -Dimethyl sulfoxide
DNA	2'-deoxyribonucleic acid
ds	Double stranded
EDTA	Ethylenediaminetetraacetic acid
Et	Ethyl
EtOAc	Ethyl acetate
Fmoc	9-Fluorenylmethoxycarbonyl

FT	Fourier Transform
g	gram
G	Guanine
h	Hour(s)
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronum hexafluorophosphate
HOBt	1-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
Hz	Hertz
I ^{bu}	Isobutyryl
IR	Infrared
L-	Levo-
LC-MS	Liquid Chromatography-Mass Spectrometry
Lys	Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-Methyl benzhydryl amine
MF	Molecular formula
mg	milligram
MHz	Megahertz
mins	minutes
μL	Microlitre
μM	Micromolar
mL	millilitre
mM	millimolar
mmol	millimoles
m.p.	melting point
Ms	Methanesulfonyl
MS	Mass spectrometry
MW	Molecular weight/Microwave
N	Normal

nm	nanometre
NMR	Nuclear Magnetic Resonance
<i>p</i>	Parallel
ppm	Parts per million
Pro	Proline
PS-oligo	Phosphorothioate-oligo
Py	Pyridine
PNA	Peptide Nucleic Acid
<i>pet</i>	pyrolidinylethyl
<i>R</i>	Rectus
R _f	Retention factor
RP	Reversed Phase
rt	Room temperature
S	Sinister
SPPS	Solid Phase Peptide Synthesis
<i>t</i> -Boc	<i>tert</i> -Butoxycarbonyl
T	Thymine
TBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate
TEA / Et ₃ N	Triethylamine
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
<i>T_m</i>	Melting Temperature
U	Uridine
UV-Vis	Ultraviolet-Visible

Abstract

The Thesis entitled “**Charged Peptide Nucleic Acid Analogues: Ethano-locked PNA and β -, γ - bisubstituted PNA**” has been divided into four chapters

CHAPTER 1: Recent advances in the field of nucleic acids: An Introduction

CHAPTER 2: Novel ethano locked PNA (ethano-PNA)

Section A: Synthesis of monomer and incorporation in aegPNA sequences

Section B: Biophysical studies of ethano-PNA towards complementary DNA/RNA

CHAPTER 3: β -, γ - bisubstituted PNA

Section A: Synthesis of monomers and incorporation in aegPNA sequences

Section B: Biophysical studies of modified PNA towards complementary DNA/RNA

CHAPTER 4: Polyamide DNA

Synthesis and biophysical studies of polyamide-DNA with alternating α -amino acid and nucleoside- β -amino acids

Chapter 1

Recent advances in the field of nucleic acids: An Introduction

The origin of the concept of antisense and antigene therapy made nucleic acids a target for potential therapeutic interventions. A variety of micro and macro molecules, both natural and synthetic, are capable of interacting with DNA or RNA. These interactions can lead to the inactivation of gene expression. Selective inhibition of disease causing genes is theoretically possible by taking advantage of the known specific hydrogen bonding interactions between complementary base pairs of nucleic acids. Based on these interactions, tailored short nucleic acid sequences or oligonucleotides, could selectively bind to the target DNA/RNA. Oligonucleotides that interact with single stranded RNA are termed as ‘antisense’ whereas those interacting with double stranded DNA are called ‘antigene’ oligonucleotides. However, intracellular enzymes such as nucleases and proteases, rapidly cleave the unmodified oligo deoxy nucleotides ODNs with sugar-phosphate backbone. The development of new classes of modified oligonucleotides has been trigger to allow the industrious application of antisense/antigene principle.

Among all the DNA mimics, PNA introduced by Nielsen *et. al* in 1991 has been found to mimic many properties of DNA. The structure of PNA is remarkably simple, the repeating unit of PNA consisting of N-(1-aminoethyl)-glycine units linked by amide bonds and the nucleobases (A, G, C and T) attached to the backbone through methylene carbonyl linkages (Figure 1). Thus PNA lacks sugar and phosphate groups, making it acyclic, neutral, achiral and homomorphous. PNA is resistant to cellular enzymes and has strong affinity towards complementary DNA/RNA.

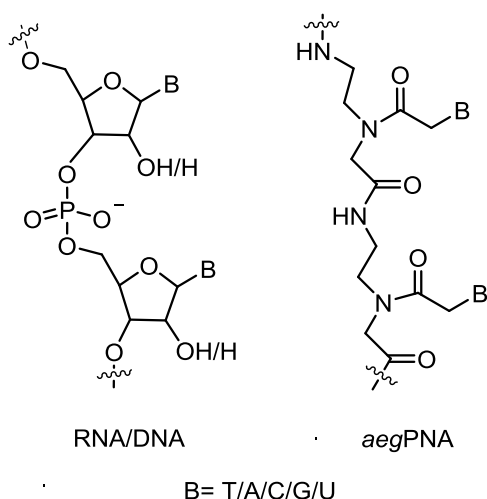


Figure 1 Comparison of DNA and PNA

Though PNAs encompass some disadvantages like poor water solubility, less orientational selectivity between parallel and antiparallel binding modes, lack of selective in binding affinity between DNA and RNA, low membrane penetration and inefficient cellular uptake. Various attempts have been made to address these shortcomings of PNA. Introduction of various modifications/substitutions in the PNA backbone that resulted in

preorganized and chiral PNA, are aimed to achieve directional selective binding to target DNA/RNA.

This chapter gives an outline of the oligonucleotide analogues and various modifications of PNAs carried out to improve their applications as therapeutic agents and diagnostic markers. PNAs suffixed with negatively charged DNA or modified PNAs with positively charged polypeptide sequence at either 'C' or 'N' terminus, resulted in PNA-DNA/PNA peptide chimera and backbone modifications with charged residues, with favorable aqueous solubility, cellular uptake and DNA binding/recognition properties are described.

Chapter 2 : Novel ethano locked PNA (ethano-PNA)

Section A : Synthesis of monomer and incorporation in aegPNA sequences

Some crucial differences in PNA:DNA and PNA:RNA duplexes were pointed out by the NMR and X-ray crystal structural studies. The main difference was in the preferred dihedral angle β in the ethylene diamine segment (N1'-C2'-C3'-N4', Figure

2) of PNA units while binding with RNA and DNA. The preferred dihedral angle β was found to be 60-70° in a PNA:RNA duplex and the same was ~140° in PNA:DNA duplex. We have earlier used the information from the structural studies to design RNA selective *cis*-cyclopentyl (Figure 2, 2b) and *cis*-cyclohexyl (Figure 2, 2c) based PNA in which the dihedral angle β was restricted to 60-70°. The *trans*-cyclohexyl and *trans*-cyclopentyl (Figure 2, 2a) based PNAs were constructed by Nielsen and Appella's groups respectively. The rigid nature of cyclohexyl ring was found to be highly successful in stabilizing the duplexes with RNA in a stereoselective manner when the geometry of substitution was *cis* and the dihedral angle β was restricted to the preferred value (~ 60°) in RNA:PNA duplexes.

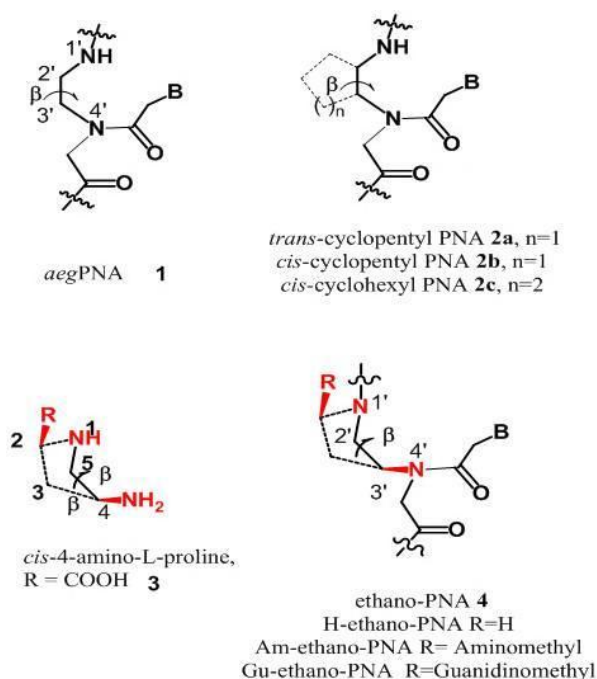


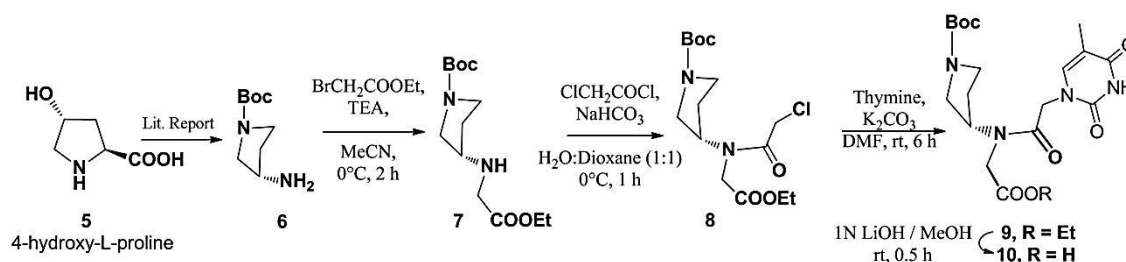
Figure 2: aegPNA, related PNA modifications and the designed ethano-PNA analogues

In this context, we propose to use 3-aminopyrrolidine core structure towards the rational design of conformationally restricted *aeg*PNA. The pyrrolidine ring conformations in substituted prolines are known to be dictated by the R/S stereochemistry of the electronegative substitutions at C-4 centre (Figure 2). In the present PNA design, this N1-C5-C4-N4 segment of the pyrrolidine ring coincides with the aminoethyl segment (*i.e.* N1'-C2'-C3'-N4', Figure 2) of *aeg*PNA and the *gauche*

geometry of the vicinal electronegative substituents would coincide with the preferred dihedral angle β in *aeg*PNA:RNA duplexes. Thus, we present here, the design, synthesis and biophysical studies of a novel ethano-locked PNA (ethano-PNA) that is envisaged to confer conformational constraint on the relatively flexible PNA backbone for favorable RNA binding. The conformational restriction is introduced in the form of an ethylene bridge between aminoethyl linker of *aeg*PNA, hence, termed as ethano-PNA. The ethano group is further amenable for functionalization with aminomethyl (Am-ethano-PNA, Figure 2) and guanidinomethyl (Gu-ethano-PNA, Figure 2) group for better aqueous solubility.

2A.1 Synthesis of Boc protected H-ethano-PNA T monomer

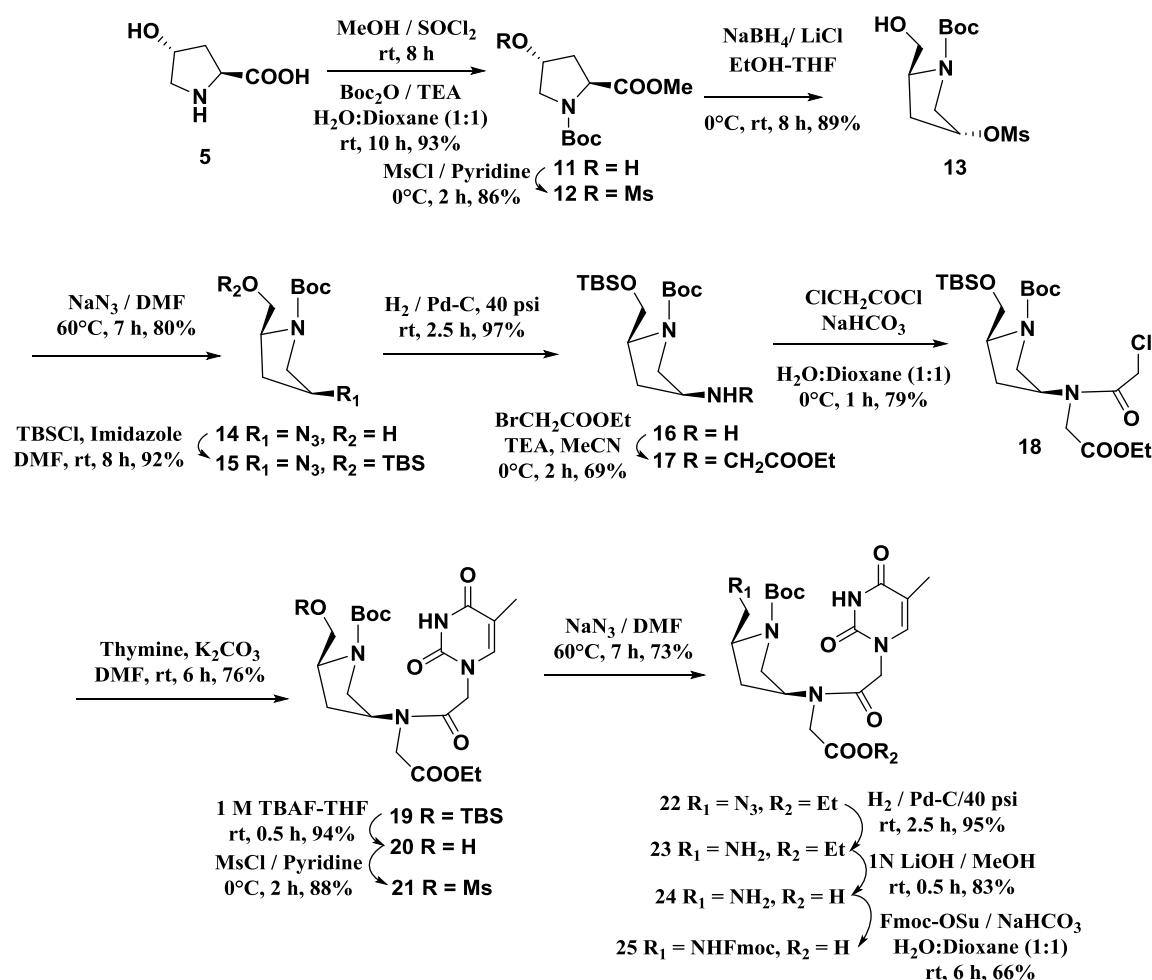
The synthesis of H-ethano-PNA monomer was accomplished as follows, starting from naturally occurring (2*S*, 4*R*) hydroxyproline **5**.



Scheme 1: Synthesis of (*S*) H-ethano-PNA T monomer

2A.2 Synthesis of Fmoc/Boc protected Am-ethano-PNA T monomer

The synthesis of Am-ethano-PNA monomer was also achieved from (2*S*, 4*R*) hydroxyproline **5**.

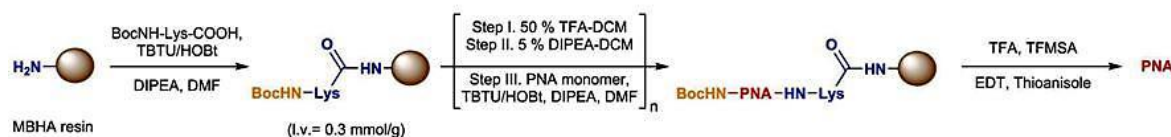


Scheme 2. Synthesis of substituted aminomethyl-ethano-PNA T monomer

2A.3 Solid Phase Synthesis of oligomers incorporating modified PNA monomers

PNA synthesis was undertaken from the C- terminal to N- terminal. PNA oligomers containing modified monomer were assembled by solid phase synthesis on

MBHA resin functionalized with the first residue L-lysine, using TBTU/HOBt activation strategy scheme 3. The sequences are shown in Table 1. PNA sequences which were synthesized were purified by RP- HPLC. Purity of sequences was checked by RP-HPLC and sequences were characterized by MALDI-TOF mass spectrometry.

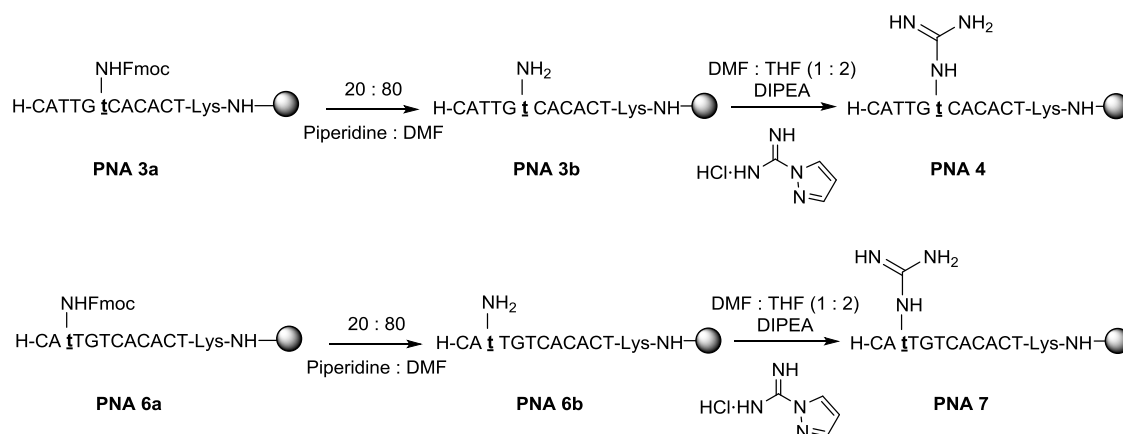


Scheme 3. Schematic representation of solid phase PNA synthesis using the Boc-protection strategy

Table 1. PNA sequences, HPLC retention time and MALDI-TOF mass
 t^H , t^{Am} and t^{Gu} are the modified PNA monomers with H, aminomethyl/guanidinomethyl substituents

Entry No.	Code	Sequences (N' → C')	HPLC t_R (min)	MALDI_Tof Masses	
				Calcd	Obsd
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	11.8	3332.24	3330.76
2.	PNA 2	H- CATTG t^H CACACT-Lys-NH ₂	11.7	3356.40	3355.14
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	11.8	3385.42	3385.43
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	13.0	3427.44	3426.41
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	11.9	3385.42	3354.98
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	11.7	3385.42	3408.36 (M+Na ⁺)
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	12.5	3427.44	3427.39

2A.4 Post synthetic conversion of amino to guanidino group on solid phase



Scheme 4. Conversion of amino to guanidine- functionality on solid support

Scheme 4 represents the global guanidinylation of the Am-ethano-PNA sequences on the solid support. Fmoc protected amino groups in Am-ethano units in the Am-ethano-PNA deprotected by 20% piperidine-DMF solution and resin was then subjected to guanidinylation by suspending in the solution of pyrazole-1-carboxamide hydrochloride (10 equivalents per amino group)/DIPEA in DMF-THF (1:2) mixture. The sequences are shown in Table 1, entries 4 and 7.

Chapter 2 : Novel ethano locked PNA (ethano-PNA)

Section B : Biophysical studies of ethano-PNA towards complementary DNA/RNA

2B.1 UV- T_m studies on ethano-PNA:cDNA and ethano-PNA:RNA

The synthesized PNA sequences were examined for their binding affinity with cDNA/RNA sequences by employing UV-thermal denaturation studies. The sequence chosen for study was a 12mer sequence. The modified oligomers were annealed with the cDNA and RNA and were subjected to temperature dependent UV studies at 260 nm. Unmodified **PNA 1** was used as a control in these experiments. The UV- T_m plots show a single sigmoidal transition, characteristic of PNA:cDNA/RNA duplex melting. The T_m values were obtained by the first derivative and the values corresponding to peaks from such plots for various PNA:cDNA/RNA duplexes of modified and unmodified oligomers are shown in Table 2.

Table 2. PNA sequences and UV- T_m studies. t^H , t^{Am} and t^{Gu} are the modified ethano-PNA monomers with H, aminomethyl and guanidinomethyl substituents respectively

Entry No.	Code	Sequences (N' → C')	UV - T_m (°C)		ΔT_m (°C)	
			cDNA	RNA	Modi-PNA1	Modi-PNA1
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	64.6	66.1	-	-
2.	PNA 2	H-CATTG t^H CACACT-Lys-NH ₂	49.7	59.4	-14.9	- 6.7
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	51.8	60.4	-12.8	- 5.7
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	53.7	62.7	-10.9	- 3.4
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	55.8	65.5	- 8.8	- 0.6
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	57.1	66.3	- 7.5	+ 0.2
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	60.6	68.1	- 4.0	+ 2.0

2B.2 Effect of salt concentration on UV- T_m of ethano-PNA:cDNA/RNA duplexes

In the present studies, a positive charge was introduced in the backbone of PNA and therefore at higher salt concentration some destabilization of the duplexes could be expected, depending upon the electrostatic properties and interactions of the charged

amino- and guanidino- substitutions in the backbone. Thermal melting studies were therefore undertaken at higher salt concentration (100 mM NaCl) Table 3.

Table 3. Effect of salt concentration on UV- T_m studies. t^H , t^{Am} and t^{Gu} are the modified ethano-PNA monomers with H, aminomethyl and guanidinomethyl substituents respectively

Entry No.	Code	Sequences (N' → C')	UV- T_m (°C)			
			cDNA (10mM)	RNA (10mM)	cDNA (100mM)	RNA (100mM)
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	64.6	66.1	60.1	63.7
2.	PNA 2	H- CATTG t^H CACACT-Lys-NH ₂	49.7	59.4	46.2	52.3
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	51.8	60.4	46.6	57.2
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	53.7	62.7	48.8	59.4
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	55.8	65.5	50.0	59.9
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	57.1	66.3	53.8	60.1
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	60.6	68.1	56.3	60.7

2B.3 UV- T_m studies on ethano-PNA with mismatch RNA

To further rule out the possibility of involvement of nonspecific electrostatic interactions at the expense of sequence specificity of binding, we conducted experiments with RNA sequence comprising a single mismatch. Thermal melting experiments were carried out with mismatch RNA sequence (Table 4), to assess the effect on T_m values for H-ethano-PNA, Am-ethano-PNA and Gu-ethano-PNA.

Table 4. UV- T_m studies on ethano-PNA with mismatch RNA. t^H , t^{Am} and t^{Gu} are the modified ethano-PNA monomers with H, aminomethyl and guanidinomethyl substituents respectively

Entry No.	Code	Sequences (N' → C')	UV- T_m (°C)			
			RNA (10mM)	RNA (100mM)	misRNA (10mM)	misRNA (100mM)
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	66.1	63.7	55.1	53.0
2.	PNA 2	H- CATTG t^H CACACT-Lys-NH ₂	59.4	52.3	45.4	42.2
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	60.4	57.2	46.7	44.7
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	62.7	59.4	47.1	46.3
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	65.5	59.9	50.8	43.3
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	66.3	60.1	51.8	45.7
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	68.1	60.7	53.6	46.8

2B.4 Conformational analysis of the ethano-PNA:cDNA and ethano-PNA:RNA

In order to examine if the ethano-PNAs cause any significant differences on the overall conformational features of the duplexes, the CD studies of PNA complexes with cDNA as well as RNA were undertaken. The CD Analysis of single strands was also

carried out for any conformation of the PNAs. No significant change in CD patterns was observed in either the DNA or RNA duplexes and the fact suggests that the overall structural features of the duplexes with either DNA or RNA remained unchanged.

2B.5 Summary

A novel conformationally constrained PNA analogue was designed which would exert restrictions in the aminoethyl segment of PNA such that the dihedral angle β is restricted for RNA selective binding. The modified oligomers containing ethano-PNA monomers did show preferential binding towards RNA than cDNA. The modification also showed positional dependence as well as substitution dependence binding as only **PNA 7** could stabilize the duplex better than the unmodified PNA. The destabilizing effect and preferential binding towards RNA could be accounted by the reduced stacking interactions as observed from the decreased amplitudes of the signals in the CD studies.

Chapter 3 : β -, γ - bisubstituted PNA

Section A : Synthesis of monomers and incorporation in aegPNA sequences

A major barrier to the successful development of antisense technology has been the inability to develop nucleic acid analogues that can traverse the cell-membrane and bind selectively to the intended RNA targets. PNA, developed by Nielsen *et al.* rekindled optimism that the uptake problem would finally be resolved because of the neutral backbone.

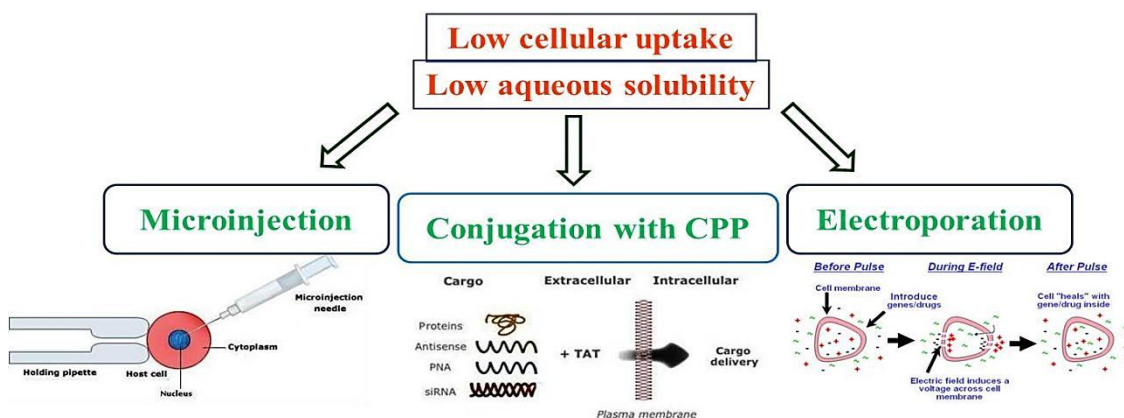


Figure 3. Several strategies to transport cargos into cells

This optimism, however, soon faded with the discovery that PNA is not readily taken up by mammalian cells, despite the neutral backbone. Considerable efforts have been vested in the last decade trying to develop means to transport PNA into cells. Several strategies (Figure 3), including microinjection, electroporation, DNA-mediated

transduction, PNA-DNA chimera and covalent attachment to cell-penetrating peptides (CPPs), have been developed, but for the most part they are limited to small-scale experimental setups. Although the conjugation of positively charged peptides improves the cellular uptake of PNAs, but these charged peptides often are cytotoxic, show off-target effects and may not be useful for *in vivo* studies.

As an alternative to peptide conjugation, the PNAs could be designed to be intrinsically positively charged. We and others have embarked to address this problem and have devised several interesting structural analogues of PNA based on the position of a substituent in the PNA backbone (α -PNA, β -PNA, and γ -PNA).

GPNA and γ GPNA designed by Ly *et al.* obtained by substitution in *aeg*PNA by D-arginine (α - and γ - position resp.), showed efficient binding to the target RNA as *aeg*PNA and was also found to be equipped to cross the cell-membrane barriers (Figure 4B). Recently Sforza *et al.* developed an “Extended Chiral Box” PNA⁵⁸ having four lysine residues attached intrinsically to PNA backbone. The substitutions are at α -, γ and $\alpha + \gamma$ positions of *aeg*PNA monomer unit (Figure 4C). Similarly there are several other reports of having substitution at only one position (α - or γ -) of PNA.

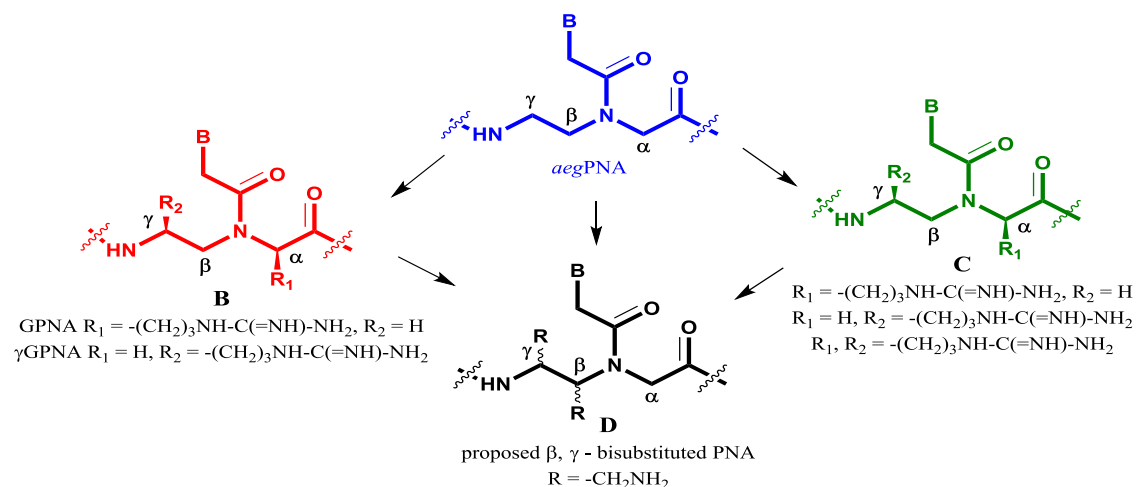
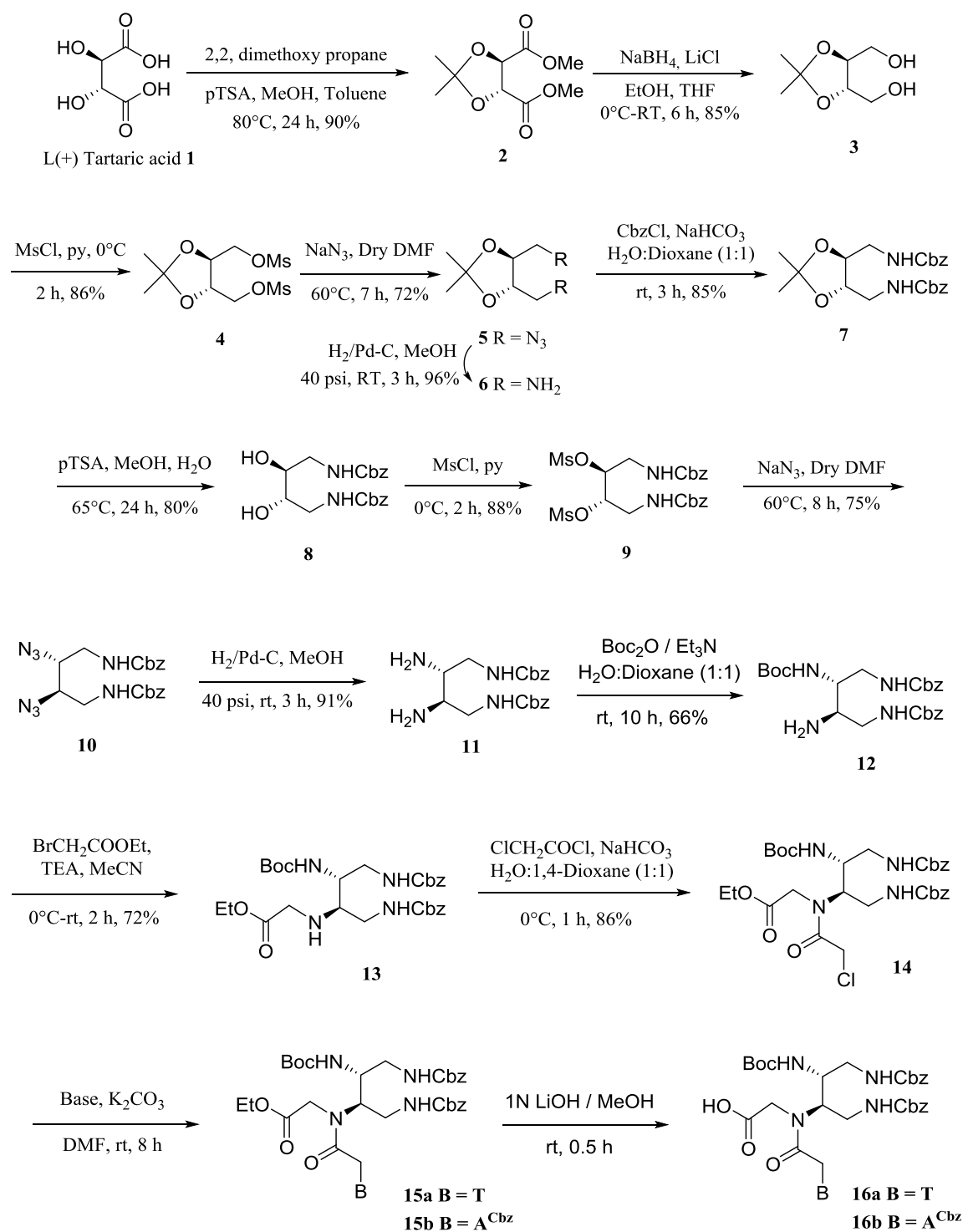


Figure 4. Rationale for the design of proposed β , γ -bisubstituted PNA

Thus in this context we designed an *aeg*PNA monomer unit bearing bisubstitution of aminomethyl group at β - and γ - positions (Figure 4D). The intrinsic amino group will increase the aqueous solubility as well as cellular uptake.

3A.1 Synthesis of β - and γ -bisubstituted PNA T/A^{cbz} monomer

We started our synthesis with commercially available L (-) Tartaric acid which by chemical transformations could lead to our desired monomer. Also by changing to D (+) Tartaric acid we had the other distereoisomer, and effect of stereochemistry was studied.

Scheme 5. Synthesis of β - and γ - bisubstituted PNA T/A^{cbz} monomer

3A.2 Solid Phase PNA Synthesis with β - and γ - bisubstituted T/A^{cbz} monomer

As in the case of solid phase peptide synthesis, PNA synthesis is also done conveniently from 'C' terminal to 'N' terminal. PNA oligomers containing modified monomer was assembled by solid phase synthesis on MBHA (4-methyl benzhydryl amine) resin functionalized with the first residue L-Lysine, using TBTU/HOBt

activation strategy. The modified monomers were incorporated in pre-defined positions by solid phase synthesis to yield aegPNAs. All the PNA oligomers were cleaved from the solid support by using a standard protocol employing TFA-TFMSA to yield PNAs carrying lysine at the carboxyl terminus and aminomethyl groups in the backbone. The PNA oligomers were purified by HPLC on a PepRP column and characterized by MALDI-TOF mass spectrometry, shown in Table 5.

Table 5. PNA sequences, HPLC retention time and MALDI-TOF mass
 t_L^{Am}/a_L^{Am} and t_D^{Am}/a_D^{Am} are the modified monomers derived from L- /D - tartaric acid resp.

Entry No.	Code	Sequences (N' → C')	HPLC t_R (min)	MALDI - ToF Masses	
				Calcd.	Obsd.
1.	PNA 1	H-AACCGATTTCAG-Lys-NH ₂	12.41	3381.35	3384.09
2.	PNA 2	H- (K) ₂ -AACCGATTTCAG-Lys-NH ₂	12.71	3637.70	3639.01
3.	PNA 3	H- (K) ₄ -AACCGATTTCAG-Lys-NH ₂	12.86	3894.05	3895.98
4.	PNA 4	H- AACCGA t_L^{Am} TTCAG -Lys-NH ₂	12.47	3439.43	3439.88
5.	PNA 5	H- AACCGA t_D^{Am} TTCAG -Lys-NH ₂	12.48	3439.43	3440.39
6.	PNA 6	H- A a_L^{Am} CCGATTTCAG -Lys-NH ₂	12.49	3439.43	3439.52
7.	PNA 7	H- A a_D^{Am} CCGATTTCAG -Lys-NH ₂	12.44	3439.43	3439.26
8.	PNA 8	H- $a_L^{Am}a_L^{Am}$ CCGATTTCAG -Lys-NH ₂	13.12	3497.51	3497.40
9.	PNA 9	H- $a_D^{Am}a_D^{Am}$ CCGATTTCAG -Lys-NH ₂	13.27	3497.51	3497.83

Chapter 3 : β -, γ - bisubstituted PNA

Section A : Biophysical studies of modified PNA towards complementary DNA/RNA

The purpose of the designed oligonucleotides is to achieve high binding affinity and specificity to target nucleic acid and to increase aqueous solubility and in turn cellular internalization. All newly designed and synthesized molecules were hence subjected to various biophysical and biochemical studies to evaluate their potential as antisense agents.

3B.1 UV- T_m studies on modified PNA:cDNA and modified PNA:RNA

The synthesized PNA sequences were examined for their binding affinity with complementary DNA/RNA sequences by employing UV-thermal denaturation studies. The sequence chosen for study was a 12mer sequence (antimiR 29a). The modified oligomers were annealed with the cDNA and RNA and were subjected to temperature dependent UV studies at 260 nm. Unmodified **PNA 1** was used as a control in these experiments. The detailed UV- T_m values for the sequences are listed in Table 6.

Table 6. PNA sequences and UV- T_m studies with cDNA and cRNA

Entry No.	Code	Sequences (N' → C')	UV- T_m (°C)		ΔT_m (°C)	
			cDNA	cRNA	cDNA	cRNA
					modi-PNA1	
1.	PNA 1	H - AACCGATTTCAG -Lys-NH ₂	58.7	64.5	-	-
2.	PNA 2	H - (K) ₂ - AACCGATTTCAG -Lys-NH ₂	60.1	66.6	+ 1.4	+ 2.1
3.	PNA 3	H - (K) ₄ - AAC CGATTTCAG -Lys-NH ₂	61.5	68.6	+ 2.8	+ 4.1
4.	PNA 4	H- AACCGA t_L^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	- 58.7	- 64.5
5.	PNA 5	H- AACCGA t_D^{Am} TTCAG -Lys-NH ₂	n.d.	25.9	- 58.7	-38.6
6.	PNA 6	H- A a_L^{Am} CCGATTTCAG -Lys-NH ₂	32.3	36.2	- 26.4	- 28.3
7.	PNA 7	H- A a_D^{Am} CCGATTTCAG -Lys-NH ₂	42.2	48.9	- 16.5	- 15.6
8.	PNA 8	H- $a_L^{Am} a_L^{Am}$ CCGATTTCAG -Lys-NH ₂	35.9	41.7	- 22.8	- 22.8
9.	PNA 9	H- $a_D^{Am} a_D^{Am}$ CCGATTTCAG -Lys-NH ₂	45.7	50.6	- 13.0	- 13.9

t_L^{Am}/ a_L^{Am} and t_D^{Am}/ a_D^{Am} are the modified monomers derived from L- and D - tartaric acid resp.

3B.2 Effect of salt concentration on T_m of modified-PNA:cDNA/RNA duplexes

PNA being charge-neutral, the strength of unmodified PNA:DNA duplexes is known to be little affected by increasing salt concentrations, in contrast to the highly salt concentration dependent duplex strength of native duplexes. In the present studies, a number of positive charges were introduced in the backbone of PNA and therefore at higher salt concentration some destabilization of the duplexes could be expected. Thermal melting studies were therefore undertaken at higher salt concentration (50 mM and 100 mM NaCl).

Table 7. UV- T_m studies with cDNA and cRNA at 50 mM and 100mM salt concentrations

Entry No.	Code	Sequences (N' → C')	UV- T_m (°C)			
			cDNA		cRNA	
			50 mM	100mM	50 mM	100mM
1.	PNA 1	H - AACCGATTTCAG -Lys-NH ₂	56.9	55.3	63.1	60.6
2.	PNA 2	H - (K) ₂ - AACCGATTTCAG -Lys-NH ₂	59.2	57.3	65.0	63.7
3.	PNA 3	H - (K) ₄ - AAC CGATTTCAG -Lys-NH ₂	59.9	58.5	65.7	64.7
4.	PNA 4	H- AACCGA t_L^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	n.d.	n.d.
5.	PNA 5	H- AACCGA t_D^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	23.2	19.7
6.	PNA 6	H- A a_L^{Am} CCGATTTCAG -Lys-NH ₂	29.4	26.1	34.7	30.6
7.	PNA 7	H- A a_D^{Am} CCGATTTCAG -Lys-NH ₂	38.5	35.0	45.7	44.3
8.	PNA 8	H- $a_L^{Am} a_L^{Am}$ CCGATTTCAG -Lys-NH ₂	34.2	30.8	40.4	36.0
9.	PNA 9	H- $a_D^{Am} a_D^{Am}$ CCGATTTCAG -Lys-NH ₂	43.3	42.2	46.7	45.3

t_L^{Am}/ a_L^{Am} and t_D^{Am}/ a_D^{Am} are the modified monomers derived from L- and D - tartaric acid resp.

3B.3 UV-T_m studies of modified PNA with a single mismatch RNA

Table 8. PNA sequences and UV-T_m studies with mmRNA

Entry No.	Code	Sequences (N' → C')	UV -T _m (°C)		ΔT _m (°C)
			cRNA	mmRNA	
1.	PNA 1	H - AACCGATTCAG -Lys-NH ₂	64.5	51.9	+ 12.6
2.	PNA 2	H - (K) ₂ - AACCGATTCAG -Lys-NH ₂	66.6	52.8	+ 13.8
3.	PNA 3	H - (K) ₄ - AAC CGATTCAG -Lys-NH ₂	68.6	53.2	+ 15.4
4.	PNA 4	H- AACCGA \underline{t}_L^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	n.d.
5.	PNA 5	H- AACCGA \underline{t}_D^{Am} TTCAG -Lys-NH ₂	25.9	n.d.	+ 25.9
6.	PNA 6	H- A \underline{a}_L^{Am} CCGATTCAG -Lys-NH ₂	36.2	24.9	+ 11.3
7.	PNA 7	H- A \underline{a}_D^{Am} CCGATTCAG -Lys-NH ₂	48.9	36.6	+ 12.3
8.	PNA 8	H- $\underline{a}_L^{Am}\underline{a}_L^{Am}$ CCGATTCAG -Lys-NH ₂	41.7	30.5	+ 11.2
9.	PNA 9	H- $\underline{a}_D^{Am}\underline{a}_D^{Am}$ CCGATTCAG -Lys-NH ₂	50.6	40.6	+ 10.0

$\underline{t}_L^{Am}/\underline{a}_L^{Am}$ and $\underline{t}_D^{Am}/\underline{a}_D^{Am}$ are the modified monomers derived from L- and D- tartaric acid resp.

Although the melting values of the modified PNAs were low as compared to any of the unmodified ones (PNAs 1, 2 and 3), we conducted experiments with RNA sequence comprising of a single mismatch, considering the T_m values of the PNA 7 and PNA 9 (48.9 °C and 50.6 °C, Table 8, entry 7 and 9 resp.), which is still higher than rest of the oligomers.

3B.4 Conformational analysis of the modified PNA:cDNA and cRNA

In order to examine if the modified PNAs cause any significant differences on the overall conformational features of the duplexes, the CD studies of PNA complexes with cDNA as well as RNA were undertaken. The CD profile of the derived PNA oligomers, with the β- and γ- bisubstituted monomers derived from L- / D- isomers, when duplexed with cDNA/RNA did not show any significant change in the CD pattern corresponding to unmodified duplexes.

3B.5 Summary

A new PNA analogue was designed, having substitutions at β- and γ- positions of the PNA backbone. The oligomers containing derived monomers in the middle of the sequence didn't show binding towards cDNA/RNA. The PNAs showed positional dependence as oligomers having modification towards N - terminal did show some binding, nevertheless lower than the unmodified ones. Change in stereochemistry (from L to D) showed slight preference in binding with cDNA/RNA. The CD studies of the

oligomers with cDNA and rRNA does not show any specific conformation of the modified PNAs. The PNA **9** having $T_m = 50.6^\circ\text{C}$, which is the highest among the modified ones, can be used further for cellular studies.

Chapter 4 :
Synthesis and biophysical studies of polyamide-DNA with alternating α -amino acid and nucleoside- β -amino acids

Natural oligonucleotides (ON) with phosphodiester linkage are susceptible for hydrolysis with various enzymes like nucleases and this limits their use for various biological applications. The replacement of the internucleoside sugar-phosphate linkages were reported in literature. Our initial work was based on the prolyl-ACPC backbone (Figure 5) and in that case the alternating nucleoside- β -amino acids with natural β -amino acids gave rise to similar alternating α/β amino acid backbone scaffolds using well-established peptide chemistry. For our interests, we further studied the effect of chirality of the α -amino acids by synthesizing the thymine dimers linked with L-proline, D-proline and prochiral glycine units on base stacking interactions. Further, we have designed a promising strategy which used the combination of α -amino acids alternating with nucleoside based β -amino acids to this end. The synthesis of primary β -amino acid orthogonally protected building blocks corresponding to all four nucleosides and a model tetrameric DNA sequence in 5'-3' direction was described, where we employed trityl protected nucleoside- β -amino acid and glycine as α -amino acid. We describe herein, the synthesis of oligomers in which the phosphodiester linkages are completely replaced by the 5-atom amide linkages alternately using nucleoside- β -amino acids and glycine and their binding preferences to cDNA/RNA sequences (Figure 5).

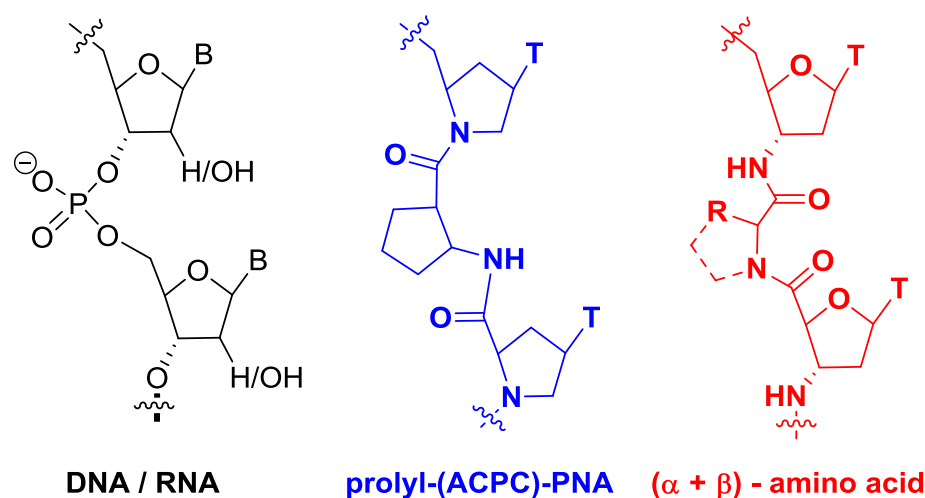


Figure 5. DNA/RNA and backbone alterations. Rationale for nucleoside- β -amino acids

4.1 Synthesis of nucleoside- β -amino acid monomers

We decided to use the trityl protection for 3'-sugar-amino function to be deprotected by trichloroacetic acid at each coupling step. This strategy would accommodate the orthogonal protection of nucleobases (benzoyl for adenine and cytosine and isobutyryl for guanine). The complete synthesis was reported earlier as part of Seema Bagmare's PhD Thesis Dec'13, NCL, Pune.

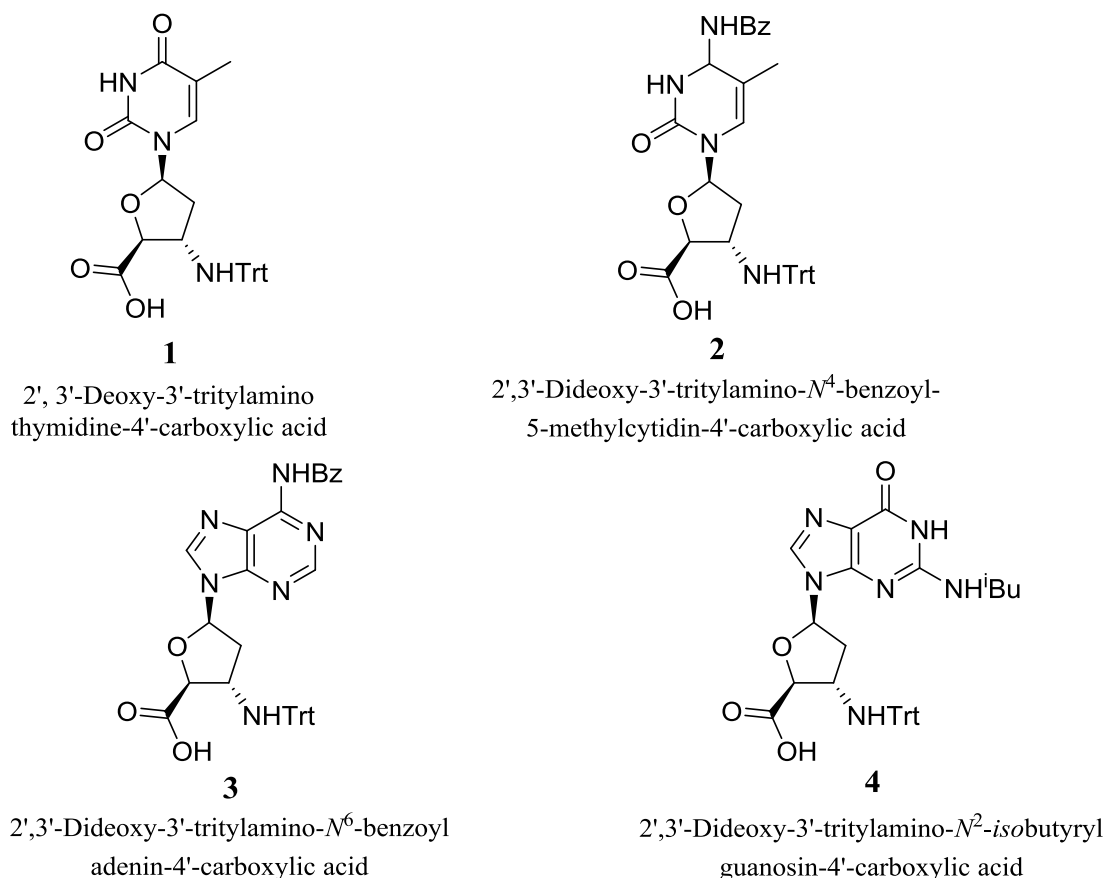


Figure 6. All four protected natural nucleoside based β -amino acids

4.2 Conformations in nucleoside- β -amino acids

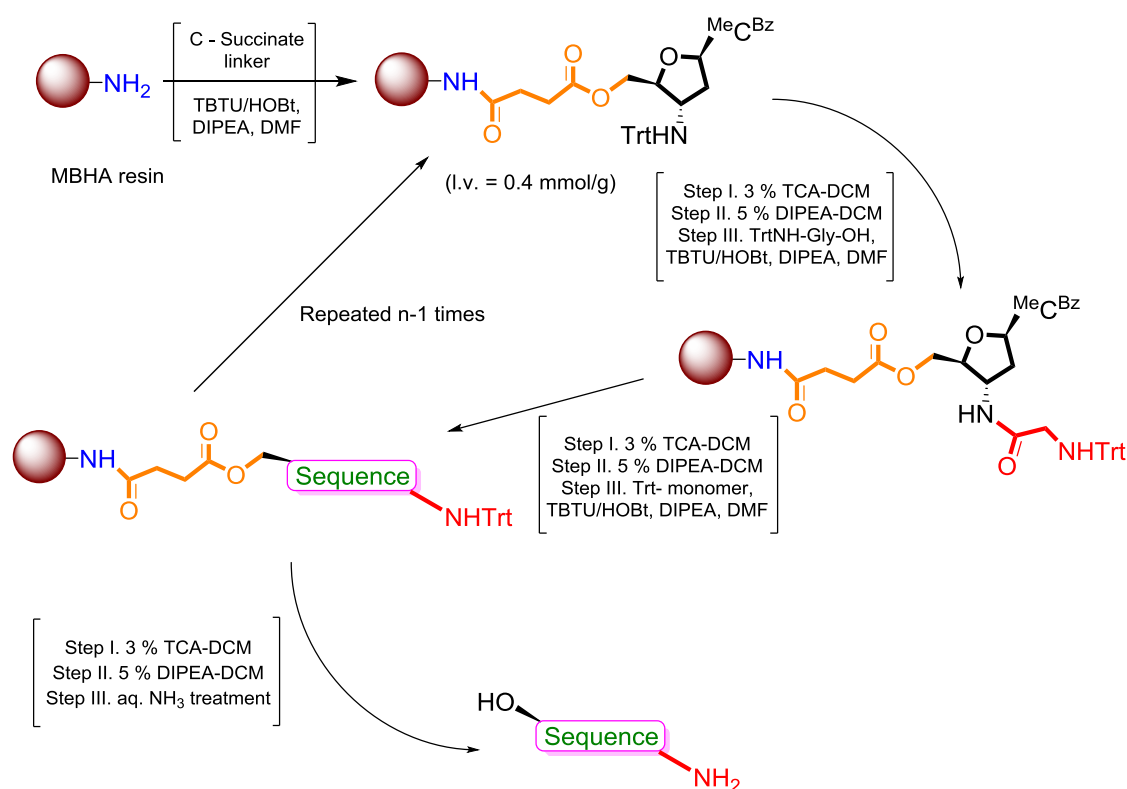
Table 9. S-type preference for the synthesized monomers

Comp. No	Nucleoside- β -amino acid	ΣH_1	% N*
1	2',3'-Deoxy-3'-tritylaminothymidine-4'-carboxylic acid	10.54	85.5
2	2',3'-Dideoxy-3'-tritylamino- <i>N</i> ⁴ -benzoyl-5-methylcytidin-4'-carboxylic acid	9.54	~100
3	2',3'-Dideoxy-3'-tritylamino- <i>N</i> ⁶ -benzoyladenin-4'-carboxylic acid	10.79	83.2
4	2',3'-Dideoxy-3'-tritylamino- <i>N</i> ² -isobutyrylguanosin-4'-carboxylic acid	9.28	~100

* % N = 100 - % S

4.3 Synthesis of polyamide DNA oligomers

The MBHA resin, which has better swelling properties, was chosen for carry out the oligomer synthesis from 3' → 5' direction. By using 3 % TCA in DCM for deprotection of Trt group and TBTU/HOBt activation strategy for coupling of monomers the sequences were synthesized (Scheme 6). The sequences were cleaved from resin via aqueous NH₃ treatment and purified by HPLC on a PepRP column and characterized by MALDI-TOF mass spectrometry. The synthesised sequences are listed in Table 1.



Scheme 6. Solid Phase Synthesis of polyamide DNA oligomers

Table 10. Polyamine sequences and PNA, HPLC retention time and MALDI-TOF mass

Entry No.	Code	Sequences	HPLC t _R (min)	MALDI-ToF Mass (Cald./Obsd.)
1.	TRT 1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T -3'	9.58	2884.04/2906.69 (M + Na ⁺)
2.	TRT 2	5'- C _g A _g C _g T _g G _g A _g T _g T _g T _g C _g A _g A -3'	10.60	3536.28/3558.60 (M + Na ⁺)

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

4.4 UV- T_m Studies of oligomers with antiparallel complementary DNA / RNA

Table 11. Polyamide DNA sequences and UV- T_m studies with *ap* cDNA and *ap* RNA

Entry No.	Code	Sequences	UV - T_m (°C)		ΔT_m (°C) modified-ctrl	
			<i>ap</i> cDNA 1	<i>ap</i> RNA 1	DNA	RNA
1.	DNA1	5'- CTT CTT CCT T - 3'	27.8	36.1	-	-
2.	PNA1	N'- cttcttct t -K -C'	47.3	51.3	+ 19.5	+ 15.2
3.	TRT 1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T - 3'	39.7	48.3	+ 11.9	+ 12.2
			<i>ap</i> cDNA 2	<i>ap</i> RNA 2	DNA	RNA
4.	DNA2	5'- CAC TGA TTT CAA - 3'	35.9	44.7	-	-
5.	PNA2	N'- cactgatttcaa- K -C'	53.8	60.6	+ 17.9	+ 15.9
6.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g T _g C _g A _g A-3'	45.3	54.0	+ 9.4	+ 9.3

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

The lower case bases represent *aeg*PNA monomer units

ap cDNA1: 5'-AAG GAA GAA G : antiparallel complementary to TRT1; *ap* cDNA2: 5'-TTG AAA

TCA GTG : antiparallel complementary to TRT2; *ap* RNA1: 5'-AAG GAA GAA G : antiparallel

complementary to TRT1; *ap* RNA2: 5'- UUG AAA UCA GUG : antiparallel complementary to TRT2

4.5 UV- T_m Studies of oligomers with parallel complementary DNA / RNA

Table 12. Polyamide DNA sequences and UV- T_m studies with *p* cDNA and *p* RNA

Entry No.	Code	Sequences	UV - T_m (°C)		ΔT_m (°C) antiparallel-parallel	
			<i>p</i> cDNA 3	<i>p</i> RNA 3	DNA	RNA
1.	DNA1	5'- CTT CTT CCT T - 3'	n.d.	n.d.	+ 27.8	+ 36.1
2.	PNA1	N'- cttcttct t -K -C'	32.9	38.5	+ 14.4	+ 12.8
3.	TRT1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T - 3'	n.d.	n.d.	+ 39.7	+ 48.3
			<i>p</i> cDNA 4	<i>p</i> RNA 4	DNA	RNA
4.	DNA2	5'- CAC TGA TTT CAA - 3'	n.d.	n.d.	+ 35.9	+ 44.7
5.	PNA2	N'- cactgatttcaa- K -C'	40.4	48.6	+ 13.4	+ 12.0
6.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g T _g C _g A _g A-3'	n.d.	n.d.	+ 45.3	+ 54.0

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

The lower case bases represent *aeg*PNA monomer units

p cDNA3: 5'-GAA GAA GGA A: parallel complementary to TRT1; *p* cDNA4: 5'-GTG ACT AAA

GTT: parallel complementary to TRT2; *p* RNA3: 5'-GAA GAA GGA A: parallel complementary to

TRT1; *p* RNA4: 5'-GUG ACU AAA GUU: parallel complementary to TRT2

4.6 UV- T_m Studies of oligomers with single mismatch DNA / RNA

Table 13. Polyamide DNA sequences and UV- T_m studies with *ap* mmDNA and *ap* mmRNA

Entry No.	Code	Sequences	UV - T_m ($^{\circ}$ C)		ΔT_m ($^{\circ}$ C) mismatched- matched	
			<i>ap</i> mm DNA 5	<i>ap</i> mm RNA 5	DNA	RNA
1.	DNA1	5'- CTT CTT CCT T - 3'	18.6	29.4	- 9.2	- 6.7
2.	PNA 1	N ² - cttcttctct t -K -C'	37.0	41.2	- 10.3	- 10.1
3.	TRT 1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T - 3'	30.8	38.3	- 8.9	- 10.1
			<i>ap</i> mm DNA 6	<i>ap</i> mm RNA 6	DNA	RNA
4.	DNA2	5'- CAC TGA TTT CAA - 3'	29.4	32.8	- 6.5	- 11.9
5.	PNA 2	N ² - cactgatttcaa-K -C'	45.7	48.6	- 8.1	- 12.0
6.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g C _g A _g A-3'	37.5	43.2	- 7.8	- 10.8

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

The lower case bases represent *aeg*PNA monomer units

mm DNA5: 5'-AAG GAT GAA G: complementary to TRT1; mm DNA6: 5'-TTG AAT TCA GTG: complementary to TRT2; mm RNA5: 5'-AAG GAU GAA G: complementary to TRT1; mm RNA6: 5'-UUG AAU UCA GUG: complementary to TRT2

4.7 Binding Stoichiometry: Job's Plot of the TRT1 oligomer

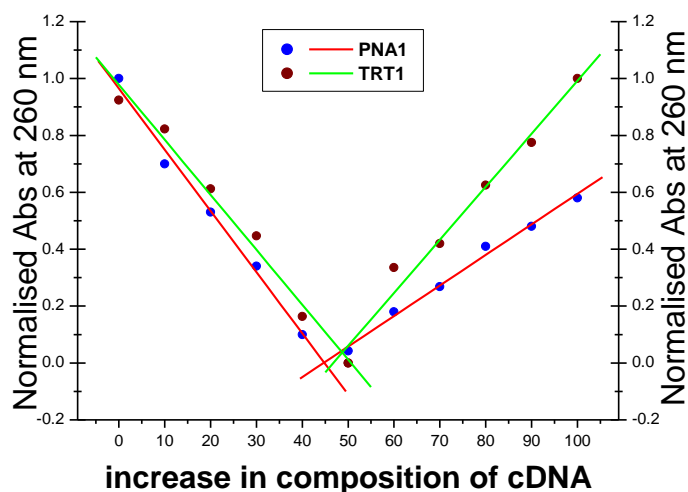


Figure 7. The Job's plot of the oligomer TRT1 and PNA1

Various stoichiometric mixtures of sequence TRT 1 and DNA 1 were made with relative molar ratios of strands 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 100:0, all at the same total strand concentration of 2 μ M in sodium

phosphate buffer, 10 mM NaCl. The samples with the individual strands were annealed and UV absorbance was recorded. A mixing curve was plotted, absorbance at fixed wavelength (λ_{\max} 260 nm) against mole fraction of **TRT 1** (Figure 7).

4.8 Conformational analysis of the polyamide DNA:cDNA/RNA

The CD studies concerning the single strands of the polyamide DNA sequence (Figure 8) did not show any characteristic features of any conformation.

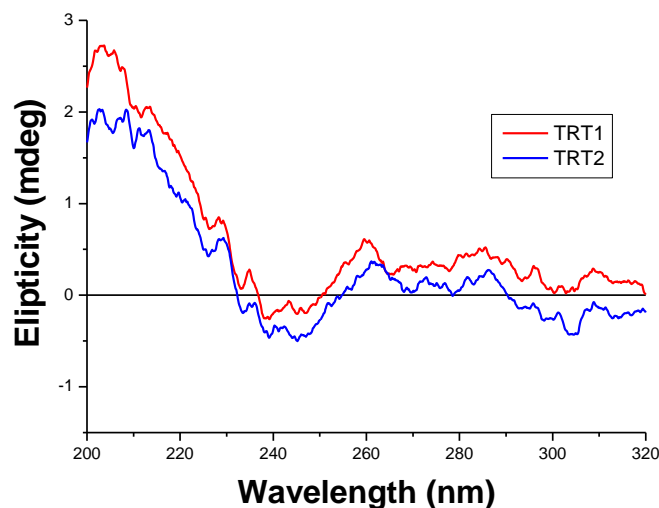


Figure 8. CD Profiles of Single strands of polyamide DNA

The CD profile of the derived DNA oligomers (**TRT 1** and **2**), Figure 9A and 10A with the modification when duplexed with cDNA showed similar CD pattern as that of the unmodified cDNA2:DNA2 duplexes. The CD profile of the derived DNA oligomers (**TRT 1** and **2**) with the modification when duplexed with RNA showed similar CD pattern as that of the unmodified DNA2:RNA2 duplex (Figure 9B and 10B).

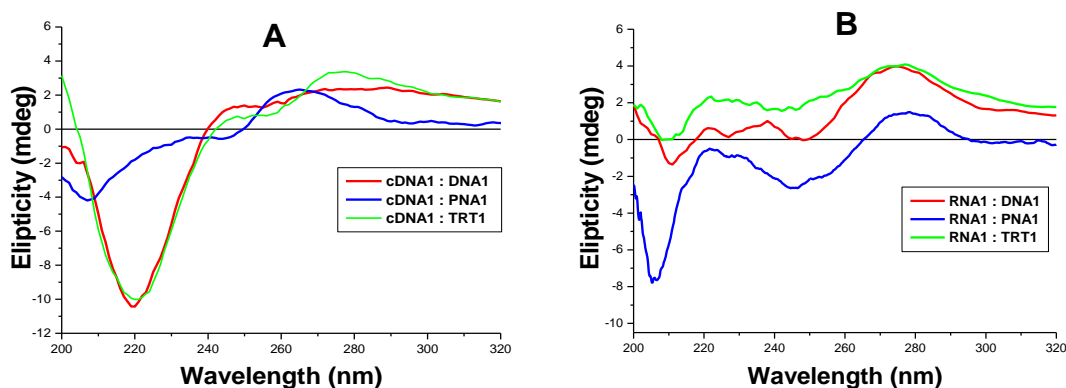


Figure 9. CD Profiles of CT sequence TRT 1 with (A) cDNA and (B) RNA

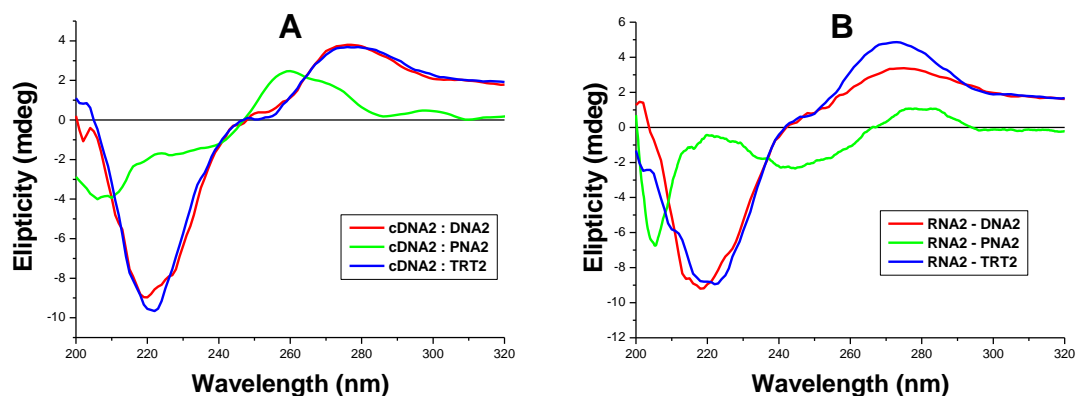


Figure 10. CD Profiles of mixed purine-pyrimidine sequence TRT 2 with (A) cDNA and (B) RNA

4.9 Summary

A new backbone modified DNA analogue had been successfully designed and synthesized, having a polyamide linkage between the sugars in the back bone instead of phosphodiester linkage. The derived polyamide DNA showed higher binding affinity with respect to the unmodified DNA towards the antiparallel complementary DNA and RNA, RNA being preferred. The UV – T_m studies regarding the binding preference of the oligomers towards parallel complementary DNA/ RNA, results in non binding, inferring that the designed oligos were having directional preference. The mismatch studies also revealed that they possess base specific binding property. The CD studies of the derived oligomers did not show specific conformational feature in the duplexes with DNA/RNA.

CHAPTER 1

Recent advances in the field of nucleic acids:

An Introduction

1A Introduction to Nucleic Acids

1A.1 Components and primary structure of nucleic acids

Nucleic acids are the most important class of biopolymers dominating the modern molecular science after the Watson-Crick discovery of the double helical structure of DNA¹ (Figure 1). Their vital roles are fundamental for the storage and transmission of genetic information within cells and contain all information required for transmission and execution of steps necessary to make proteins which are another important class of biopolymers, important for cellular function.

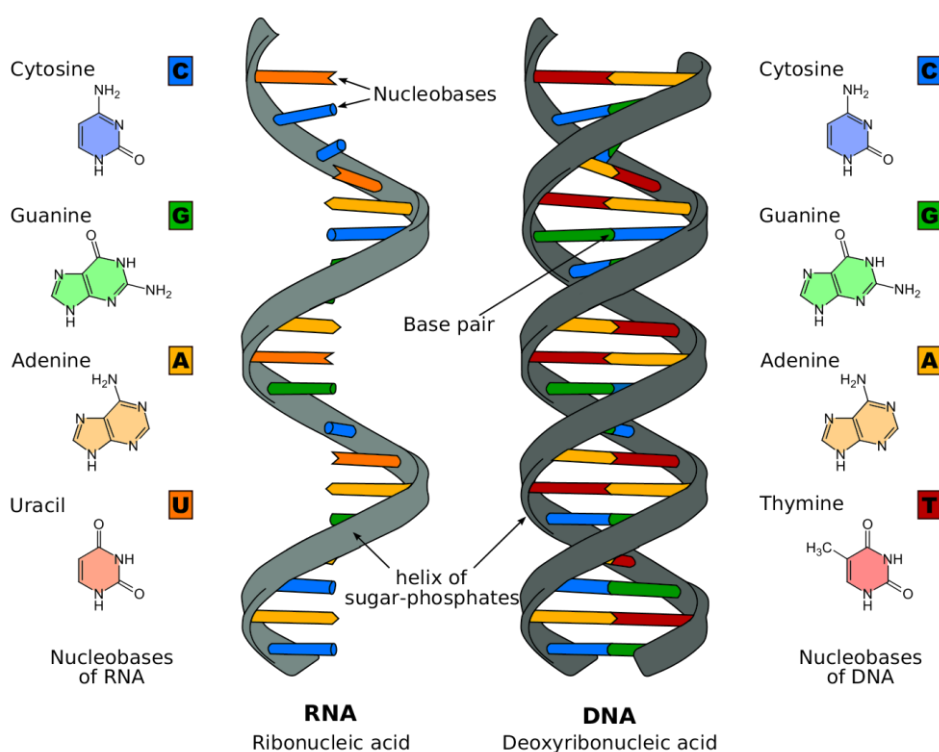


Figure 1. Chemical structure and schematic representation of DNA and RNA (courtesy: Wikipedia)

Nucleic Acids have three main components (i) A pentose sugar ring, which is ribose for RNA and deoxyribose for DNA, (ii) a nitrogenous heterocyclic base (nucleobase), which is either a purine or a pyrimidine, connected to the C1' of the pentose sugar via a β -glycosidic linkage (iii) the 3' 5' phosphodiester linkage joining the individual nucleoside units. The nitrogenous bases adenine (A), cytosine (C), and guanine (G) are found in both RNA and DNA, while thymine (T) only occurs in DNA and uracil (U) only occurs in RNA (Figure 2). Polymerization of the above mentioned relatively simple components leads to the DNA formation. A deoxyribose sugar attached with the nucleobase is called as a nucleoside, a nucleoside phosphorylated at 5'-hydroxyl group is called as a nucleotide monomeric unit and the polymeric chain of these nucleotides is called as DNA.

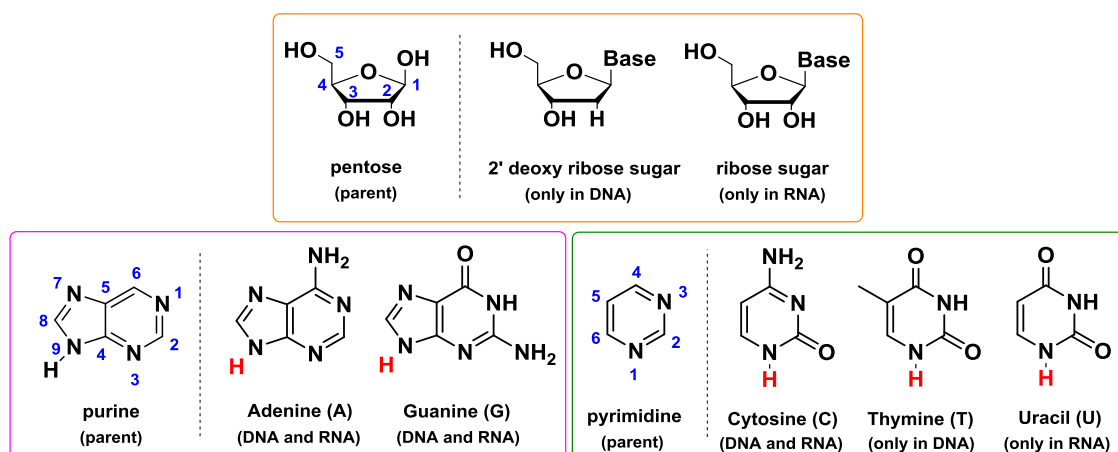


Figure 2. Chemical components of nucleic acids

1A.2 Sugar pucker in nucleos(t)ides

The molecular geometry of an individual nucleotide is very closely related to that of the corresponding nucleotide units in oligomers and nucleic acid helical structure. The details of the conformational structure of nucleotides are accurately defined by the torsional angles α , β , γ , δ , ϵ and ζ in the phosphate backbone, θ_0 to θ_4 , in the furanose ring, and χ for the glycosidic bond (Figure 3). Because many of these torsional angles are interdependent, one can simply describe the shapes of nucleotides in terms of the parameters: the sugar pucker, the syn-anti conformation of the glycosidic bond, the orientation of C4'-C5' bond and the phosphodiester backbone. The pentose sugar rings in nucleosides are twisted or puckered in order to minimize non-bonded interactions between their substituents.

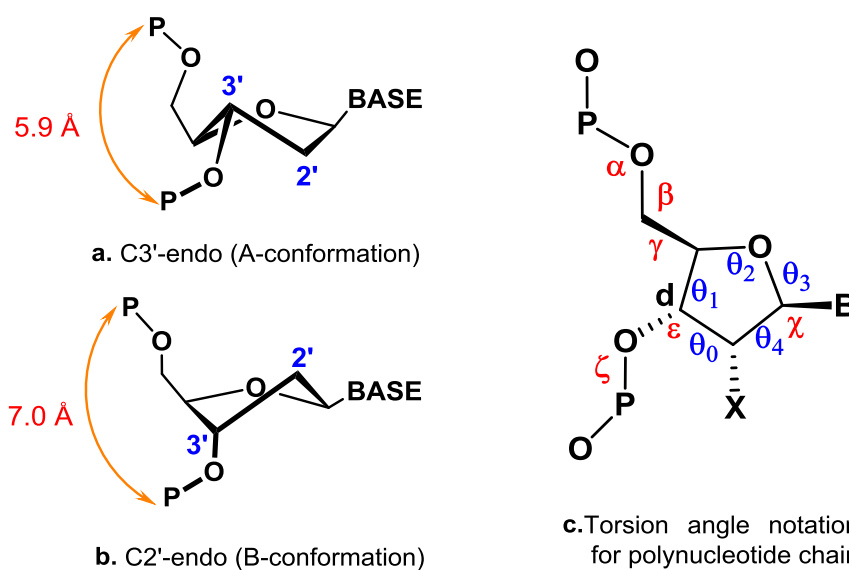


Figure 3. a) Structures of C3'-endo and b) C2'-endo preferred sugar pucker and c) Torsion angle notation for polynucleotide chain

This 'pucker' is described by identifying the major displacement of carbon 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 3a and 3b). 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 are located in the north (N) and south (S) domains of the pseudorotation cycle of the furanose ring.² 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 (i) the preference of the electronegative substituents at C2' and C3' for axial orientation, (as in the case of RNA the C2'-OH interact with backbone phosphate, C3'-endo is preferable conformation), (ii) the orientation of the base (syn goes with C2'-endo), 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. The rise in the sugar phosphate backbone for each monomeric unit is 5.9 Å in case of RNA and 7.0 Å for DNA.

1A.3 Base Pairing via Hydrogen bonding

The N-H groups of the nucleobases are potent hydrogen bond donors, while the sp^2 -hybridized electron pairs on the oxygens of the base C=O groups and that on the ring nitrogens are hydrogen bond acceptors. In Watson-Crick pairing, there are two hydrogen bonds in an A:T base pair and three in a C:G base pair (Figure 4).¹

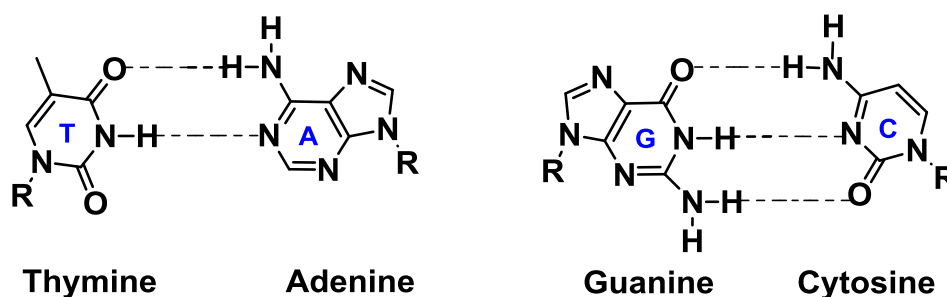


Figure 4. Watson and Crick hydrogen-bonding scheme for A: T and G: C base pair

While Watson-Crick base pairing is then dominant pattern between the nucleobases, other significant pairings are Hoogsteen (HG)³ and Wobble base pairs.⁴

A Hoogsteen A:T base pair (Figure 5) applies the N-7 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. Hoogsteen pairs have quite different properties from Watson-Crick base pairs. The angle between the two glycosidic bonds (ca. 80° 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° with respect to the other.

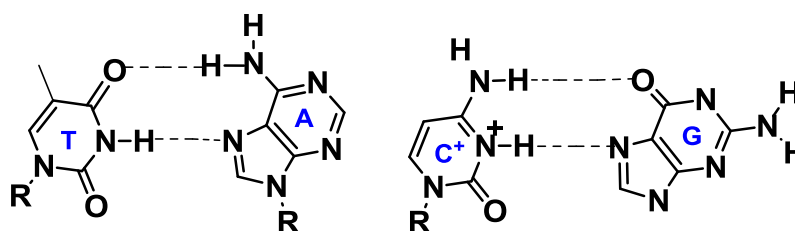


Figure 5. Hoogsteen hydrogen-bonding scheme for A: T and G: C base pair

Hoogsteen base-pairing allows sequence specific binding of the pyrimidine third strands in the major groove of Watson-Crick purine:pyrimidine duplexes to form triple-helical structures (poly (dA):2poly (dT)) and (poly (rG):2poly (rC)). The triplexes are present in three-dimensional structures of transfer RNA.

In the wobble base pairing (Figure 6), 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 (*mRNA*) with transfer RNA (*tRNA*) 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.⁵

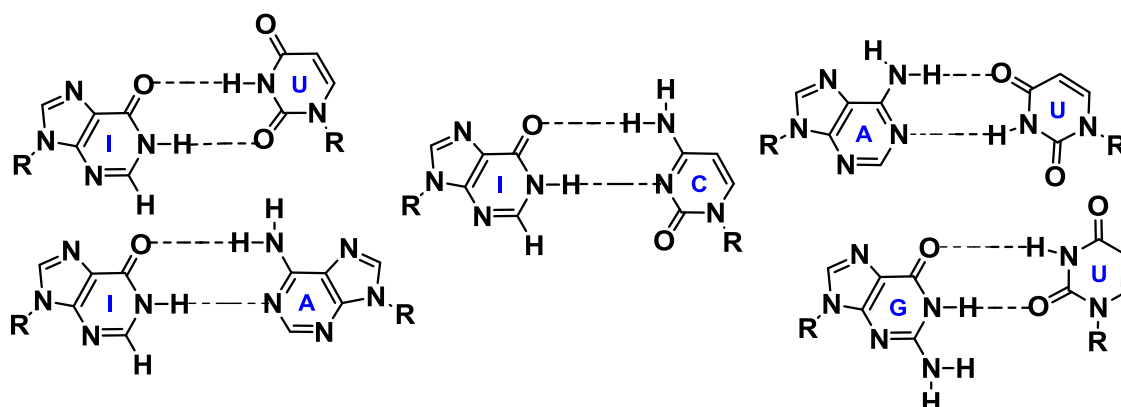


Figure 6. Wobble base pair for Inosine and Uracil

1A.4 Structures of Nucleic acids

1A.4.1 The duplex structure

Three major DNA conformations are believed to be found in nature, *A*- DNA, *B*- DNA, and *Z*- DNA (Figure 7). The "B" form described by James D. Watson and Francis Crick is believed to predominate in cells.⁶⁻⁹ *B*-DNA is a right-handed double helix with a wide major-groove and a narrow minor-groove where the bases are perpendicular to the helical axis. It is 23.7 Å wide and extends 34 Å per 10 bp of sequence. *A*-DNA and *Z*-DNA differ significantly in their geometry and dimensions from *B*-DNA, although still form helical structures. At low humidity and high salt, the favored form is highly crystalline *A*-DNA while at high humidity and low salt, the

dominant structure is *B*-DNA. In both *A* and *B* forms of DNA the Watson-Crick base pairing is maintained by anti glycosidic conformation of the nucleobases. The sugar conformation however, is different in both forms with the *B* form showing *C2'*-*endo* puckered sugar and the *A* form DNA exhibiting *C3'*-*endo* sugar- pucker. A very unusual form of DNA-duplex is the left-handed *Z*-DNA.¹⁰

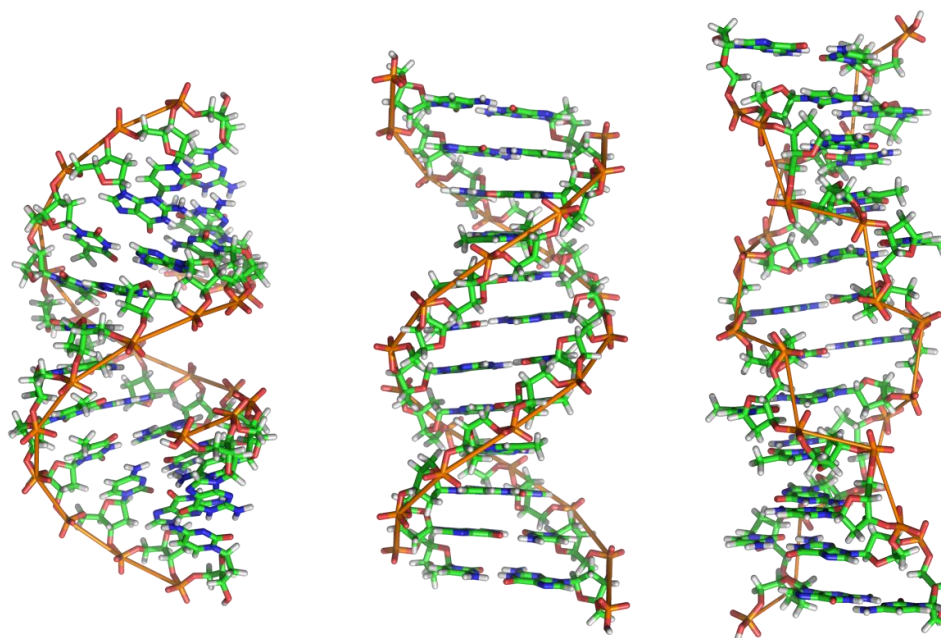


Figure 7. The structures of *A* - DNA, *B* - DNA and *Z* - DNA (courtesy: Wikipedia)

This conformation of DNA is stabilized by high concentrations of $MgCl_2$, $NaCl$ and ethanol, and is favored for alternating G:C/G:C sequences. *Z*-DNA has characteristic *zig-zag* phosphate backbone and the Watson-Crick base pairing is achieved by purines adopting *syn* glycosidic conformation with *C3'*-*endo* sugar-pucker.⁶⁻⁹ The characteristic features of major DNA conformations are summarized in Table 1. Some other possible DNA conformations are *C*-DNA, *D*-DNA, *E*-DNA, *L*-DNA, *P*-DNA and *S*-DNA,¹¹⁻¹⁴ which, have not been observed in naturally occurring biological systems.

RNA is recognized to have greater structural versatility than DNA.¹⁵ RNA can form double stranded duplexes. The most important structural feature of RNA that distinguishes it from DNA is the presence of a hydroxyl group at the 2'-position of the ribose sugar, capable of intramolecular nucleophilic attack on adjacent phosphate. The presence of this functional group enforces the *C3'*-*endo* sugar conformation (as opposed to the *C2'*-*endo* conformation of the deoxyribose sugar in DNA) that causes the helix to adopt the *A*-form geometry and hinders the formation of *B*-type helix. This results in a very deep and narrow major groove and a shallow and wide minor groove.

Table1. Salient features of the three major forms of DNA and RNA duplexes

Geometry attribute	A-DNA	B-DNA	Z-DNA
Helix sense	right-handed	right-handed	left-handed
Repeating unit	1 bp	1 bp	2 bp
Rotation/bp	33.6°	35.9°	60°/2bp
Mean bp/turn	11	10.0	12
Inclination of bp to axis	+19°	-1.2°	-9°
Rise/bp along axis	2.3 Å	3.32 Å	3.8 Å
Pitch/turn of helix	24.6 Å	33.2 Å	45.6 Å
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	anti	anti	C: anti, G: syn
Sugar pucker	C3'-endo	C2'-endo	C: C2'-endo, G: C3'-endo
Diameter	25.5 Å	23.7 Å	18.4 Å
Conditions	Low humidity & High salt	Dilute aqueous solutions	High salt & alternating GC sequences
Major groove	Narrow & deep	Wide and deep	Flat
Minor groove	Wide	Narrow and deep	Narrow and deep

RNA is single stranded but can form complex and unusual structures such as stem and bubble structures, due to the chain folding as a consequence of intramolecular base pairing. 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. *t*RNA contains about 70 bases that are folded such that there are base paired stems, bulges and open loops. The overall shape of the completely folded *t*RNA is L-shaped. At low ionic strength, A-RNA has 11 base pairs per turn in a right-handed, antiparallel double helix. The sugars adopt a C3'-endo pucker and other geometric parameters are all very similar to A-DNA. If the salt concentration is raised above 20%, A-RNA form is observed which has 12 base pairs per turn of the duplex. Both structures have typical Watson-Crick base pairs, which are displaced by 4.4 Å from the helix axis, hence forming a very deep major-groove and rather shallow minor-groove.

The presence/absence of 2'-OH functional group in RNA/DNA controls the pseudorotational equilibrium of C3'-endo (compact backbone) ↔ C2'-endo (extended backbone). This kind of sugar puckering mainly affects the backbone torsion angle δ , which on average is $\sim 80^\circ$ in A-form duplex (RNA:RNA or RNA:DNA) and $\sim 120^\circ$ in B-form duplex (DNA:DNA).¹⁶ The DNA:DNA duplex (C2'-endo conformation for

every sugar unit) is an extended B-form structure whereas RNA:RNA duplex (C3'-endo conformation for every sugar unit) remains in compact A-form. As the conformational energy barriers in the 2'-deoxyribose sugar in DNA are relatively low than in the ribose sugar of RNA, the 2'-deoxyribose sugar has a flexibility to adopt C3'-endo (or O4'-endo) conformation in the DNA strands and allows DNA:RNA duplexes also to be in A-form. The C3'-endo conformation of ribose sugar in RNA, which has the anti-glycosidic nucleobases, projects the nucleobases in axial orientation and assists efficient base-pair stacking and hence strengthens the duplex stability with RNA/DNA. Based on the base-pairing strength, which is further enhanced by axially oriented nucleobases (in C3'-endo conformation), the duplex stability of natural nucleic acids (DNA/RNA) is in the order: RNA:RNA > RNA:DNA > DNA:DNA.

1A.4.2 Triplex-Forming Oligonucleotides (TFO)

A DNA triplex is formed when pyrimidine or purine bases occupy the major groove of the DNA double Helix forming Hoogsteen pairs with purines of the Watson-Crick base pairs. Triple-helical nucleic acid structures are known since 1957.¹⁷ 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.¹⁸ 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.^{19, 20} 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.

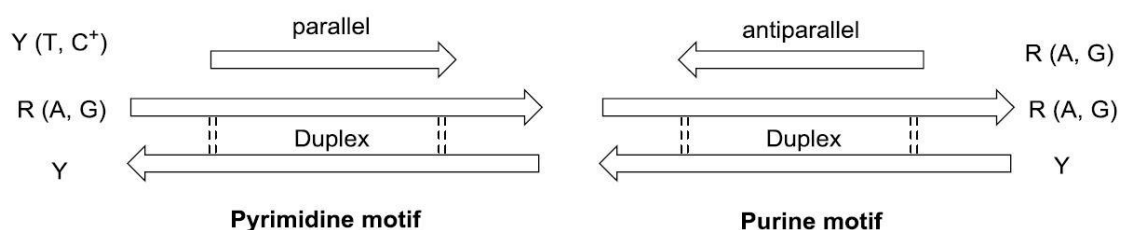


Figure 8. Pyrimidine and purine motifs of triplex formation

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^+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.²¹

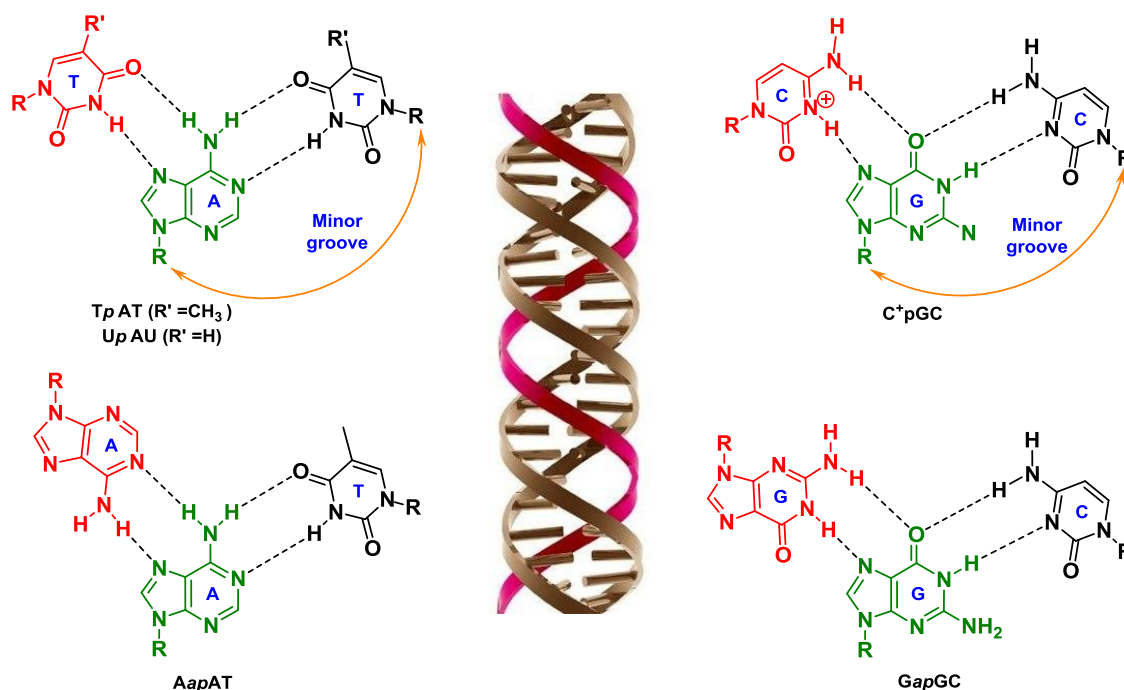


Figure 9. A ribbon model demonstrating the relative position of a TFO in the major groove of DNA, The pyrimidine binding motif in a parallel (*p*) orientation and the purine motif in antiparallel (*ap*) orientation

The widespread occurrence of polypurine/ polypyrimidine tracts in eukaryotic DNA suggest that these sequences may have a biological function and consequently, therapeutic significance. 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.²² 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.

1A.4.3 Quadruplex structure – The ‘G-quadruplex’

Nucleic acids are highly flexible molecules than can adopt different conformational structures. Most of telomeric DNA is double-stranded except for the extreme terminal part where the 3' region of the G-rich strand is single-stranded.²³ Both G-rich and C-rich telomeric strands may form *in vitro* unusual DNA structures. The G-rich strand can adopt a four - stranded G-quadruplex structure involving planar G-quartets,²⁴ while the C-rich strand can form the so-called *i*-motif with intercalated C⁺C⁺ base pairs.²⁵ Four guanine bases can associate through Hoogsteen hydrogen bonding to form a square planar structure called a guanine tetrad, and two or more guanine tetrads can stack on top of each other to form a G-quadruplex (Figure 10). The central cavity of G-quadruplexes is occupied by cations (especially potassium) which neutralize the electrostatic repulsion between guanine ‘O6’ oxygens and thus stabilize the overall

structure. Depending on the direction of the strands or parts of a strand that form the tetrads, structures may be described as parallel or antiparallel. They can be also formed with one (unimolecular), two (bimolecular) or four (tetramolecular) G-rich strands.

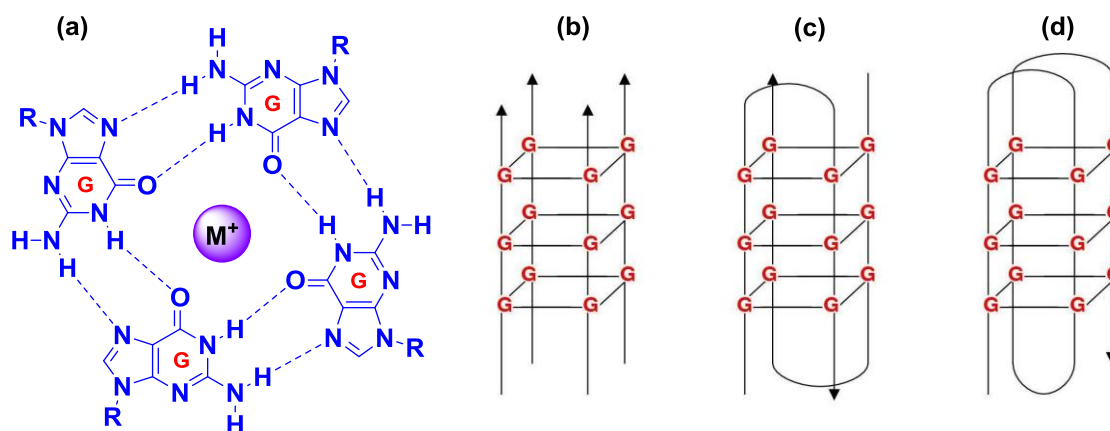


Figure 10. (a) Schematic illustration of G-tetrad, (b) G-quadruplex, parallel arrangement of four DNA strands, tetramolecular; (c) Intermolecular G-quadruplex, bimolecular, and (d) anti-parallel intramolecular G-quadruplex, unimolecular.

1A.5 Spectroscopic technique to study DNA/RNA interactions

The ability of antisense oligonucleotides to bind *in vitro* the target DNA/RNA, can be detected by using various biophysical techniques, *viz.* Ultraviolet absorption (UV), stoichiometry and circular dichroism (CD).

1A.5.1 Ultraviolet spectroscopy (UV)

The absorbance of polynucleotide depends on the sum of the absorbance of the nucleotide plus the effect of the interaction among the nucleotides. The interaction cause a single strand to absorb less than the sum of its nucleotides and a double strand absorbs less than its two component single strands. The effect is called hypochromicity 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.²⁶ In converse terms, 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 20-30 % 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). Such a thermal dissociation of nucleic acid helices in solution to give single stranded DNA/RNA is a function of base composition, sequence, chain length as well as of temperature, salt concentration, and pH of the solvent (buffer). In particular, early observations of the relationship between T_m and base composition for different DNAs showed that A-T pairs are less stable than G-C

pairs, a fact which is now expressed in a linear correlation between T_m and the gross composition of a DNA oligomer by the equation:

$$T_m = X + 0.41 (\% C + G) ^\circ C$$

The constant X is dependent on salt concentration and pH and has a value of 69.3 °C for 0.3 M sodium ions at pH 7.²⁶

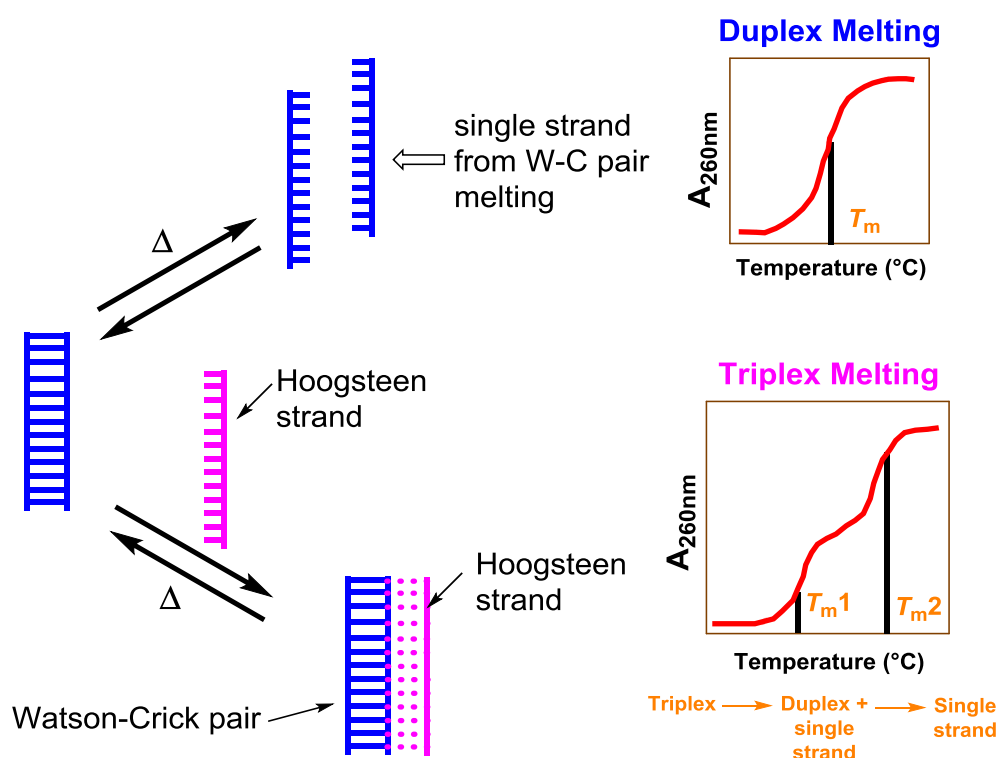


Figure 11. Schematic representation of DNA duplex and triplex meltings

Duplex melting: According to the ‘all or none model’²⁷, the UV absorbance value at any given temperature is an average of the absorbance of duplex and single strands. A plot of absorbance against temperature gives a sigmoidal curve in case of duplexes and the midpoint of the sigmoidal curve (Figure 11) called as the ‘melting temperature’ (T_m) (equilibrium point) at which the duplex and the single strands exist in equal proportions.

Triplex melting: In the case of triplexes, the first dissociation leads to melting of triplex generating the duplex (WC duplex) and third strand (Hoogsteen strand), followed by the duplex dissociation to form two single strands. The DNA triplex melting shows characteristic double sigmoidal transition (Figure 11) and UV melting temperature for each transition is obtained from the first derivative plots. The lower melting temperature (T_{m1}) corresponds to triplex to duplex transition while the second higher melting temperature (T_{m2}) corresponds to the transition of duplex to single strands.²⁸

1A.5.2 Stoichiometry

The binding stoichiometry of oligonucleotide to target DNA/RNA by UV-titration (mixing): This is derived from Job's plot.²⁹ The two components of the complex are mixed in different molar ratios, so that the total concentration of each mixture should be constant, i.e., as the concentration of one strand decreases, concentration of the second strand increases and the UV-absorbance of each mixture is recorded. The absorbance decreases in the beginning to a minimum and then again increases. The molar ratio of the two strands at which absorbance reached minimum indicates the stoichiometry of complexation.

1A.5.3 Circular Dichroism (CD)

Circular dichroism is a spectroscopic technique that depends on the property of optical activity, i.e., chiral molecules that interact with the right and left circularly polarized lights differently. The most common applications of CD include probing the structure of biological macromolecules, in particular determining the α -helical content of proteins and conformational analysis of nucleic acids and interaction with other ligands.

In the nucleic acids, the heterocyclic bases are principal chromophores. As these bases are planar, they don't have any intrinsic CD. CD arises from the asymmetry induced by linked sugar group. CD spectra from dinucleotides exhibit hyperchromicity, being more intense by roughly an order of magnitude than those from monomers. It is when the bases are linked together in polynucleotides, giving rise to many degenerate interactions that they gain additional characteristics associated with the asymmetric feature of secondary structure such as in proteins and nucleic acids.³⁰

The simplest application of CD to DNA structure determination is for identification of polymorph present in sample.³¹ 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 crossover around 258 nm. A-DNA is characterized by a positive CD band centered 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. The 250-230 nm region is also usually fairly flat though not necessarily zero. Naturally occurring RNAs adopt the A-form if they are duplex.

The commonly used units in current literature are the mean residue ellipticity (degree $\text{cm}^2 \text{dmol}^{-1}$) and the difference in molar extinction coefficients called as molar circular dichroism or $\Delta\epsilon$ (litre $\text{mol}^{-1}\text{cm}^{-1}$).

1B Applications of Nucleic Acids

1B.1 Introduction

The double helix of DNA is nature's simple and elegant solution to the problem of storing, retrieving, and communicating the genetic information of living organisms. Recognition of DNA and RNA sequences by complementary oligonucleotides is a central feature of biotechnology and is important for hybridization based biological applications. Currently much research is going on for the development of DNA-based nanostructures for drug delivery, immunotherapy, diagnostics and molecular biology.³²

Designing a drug against a protein target requires a specific understanding of 3D structure of the protein. Although recent advances in X-ray crystallography, NMR and computer modeling of proteins have accelerated the drug design process, this method lacks generality. In contrast, the base sequence in RNA and DNA is universal and this generality is very appealing from the drug-design point of view. In principle, one can design a drug that, like nucleic acids, is repetitive in its primary structure and binds sequence specifically to these drug targets. In order for the sequence specific recognition to happen, such a drug should contain nucleobases that are fundamental units of nucleic acid recognition.

Two innovative strategies are being tested for inhibiting the production of disease-related proteins using such sequence specific DNA fragments as gene expression inhibitors. The first strategy known as *antisense strategy*³³ aims to selectively impede translation by inhibiting the protein synthesis. Antisense oligonucleotides recognize a complementary sequence on target mRNA through Watson-Crick base pairing and form a duplex (RNA:DNA hybrid) that is not processed by the protein synthesis machinery and hence would retard the expression of the corresponding protein. When target proteins are disease related, this will have a therapeutic value.

The second strategy which is known as *antigene strategy*,³⁴ aims to stop the production of an unwanted protein by selectively inhibiting transcription of gene. In this method, oligonucleotides target the major groove of DNA where it winds around the double-helical DNA to form triplex involving HG hydrogen bonds. Thus the double stranded DNA itself can act as a target for the third strand oligonucleotides or their analogue, and the limitation for triplex formation is that it is possible only at homopurine stretches of DNA, since it requires purine to be the central base.

1B.2 Oligonucleotides as antisense therapeutic agent

This is vital to make antisense or antigene based inhibition as a practical approach to therapeutics. Various cellular processes can be inhibited depending on the site at which the antisense oligonucleotide hybridizes to the target nucleic acid. For an 'antisense' oligonucleotide to be able to inhibit translation, it must reach the interior of the cell unaltered. The requirements for this are the stability of the oligonucleotide towards extra and intra-cellular enzymes and equally important is its ability to traverse the cell membrane. After reaching the cytoplasm, it must bind the target mRNA with sufficient affinity and high specificity. In addition, it must possess an adequate half-life in order to elicit its action. The toxicity of the oligonucleotide should also be negligible to the cell. The unmodified antisense oligonucleotides are intended to enter the cell where they can pair with, and so inactivate, the complementary mRNA sequences. But their inability to permeate cell membrane as they carry anionic charge results from repulsive interaction with the cell lipid layer. Further, they are degraded by intracellular enzymes such as exo-nucleases.

1B.3 Antisense Mechanisms

The antisense effect of a synthetic oligonucleotide sequence was first demonstrated in the late 1970s by Zamecnik and Stephenson.³⁵ Using nucleotide sequences from the 5' and 3' ends of the 35S RNA of Rous sarcoma virus (RSV), they identified a repeated sequence of 21 nucleotides (nt) that appeared to be crucial to viral integration. They synthesized a 13-mer oligonucleotide, complement to a portion of this viral sequence. They correctly concluded that the oligonucleotide was inhibiting viral integration by hybridizing to the crucial sequences and blocking them. The specific inhibition is based on the Watson-Crick base-pairing between the heterocyclic bases of the antisense oligonucleotide and of the target nucleic acid. The RNA therapeutics target mRNA sequences through *disruptive* or *corrective* activity with strong and specific interaction of synthetic AON with target mRNA.

The disruptive antisense down-regulates the synthesis of disease-causing functional proteins by 'translational arrest' via **RNase-H activation**, or through **steric-blocking oligonucleotides**, or through **RNA interference** (either siRNA, miRNA or piRNA).

The corrective antisense approach up-regulates the gene expression of viable proteins by the **splice correction** of pre-mRNA. All these approaches involve the binding of complementary oligonucleotides/DNA analogues to target RNA through

base pairing, and therefore all are operating by an antisense mechanism besides their substantial differentiation in mechanisms of action and the functional outcomes.

1B.3.1 Antisense Oligonucleotides: Disruptive antisense approach

The AONs which can disrupt the translation processes at mRNA level cause the disruption or down-regulation of the synthesis of disease-causing functional protein. AONs usually consist of 15-20 nucleotides, which are complementary to their target mRNA. The following are major mechanisms for their antisense activity.

RNase-H activation: The first is that most AONs are designed to activate RNase H, which cleaves the RNA moiety of a DNA/RNA heteroduplex and therefore leads to degradation of the target mRNA (Figure 12a).

Steric-blocking oligonucleotides: In addition, AONs that do not induce RNase H cleavage can be used to inhibit translation by steric blockage of the ribosome. When the AONs are targeted to the 5' terminus, binding and assembly of the translation machinery can be prevented (Figure 12b).

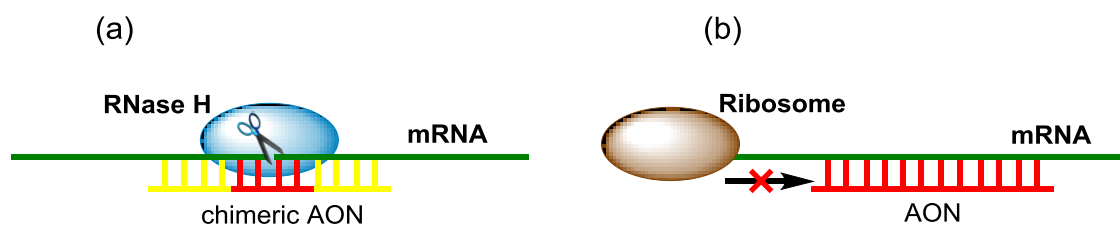


Figure 12. (a) RNase H cleavage induced by AON (b) Translational arrest by blocking the ribosome

RNA interference (RNAi): RNA interference (RNAi), is a technique in which exogenous, double stranded RNAs (dsRNAs) that are complementary to known mRNAs, are introduced into a cell to specifically destroy that particular mRNA, thereby diminishing or abolishing gene expression.³⁶ RNA interference has been first demonstrated in the nematode *Caenorhabditis elegans* to manipulate gene expression.³⁷

siRNA: Formation of small interfering RNA (siRNA) occurs in two steps involving binding of the RNA nucleases to a large double-stranded RNA (dsRNA) and its cleavage by the enzyme Dicer into discrete 21 – 25 nucleotide RNA fragments called siRNA (Figure 13). In the second step, these siRNAs are incorporated into the RNA induced silencing complex (RISC) and binds to the target mRNA by complementary base pairing, which subsequently suppresses gene expression either by cleavage or by translational repression. siRNA delivered by microinjection into the intestine exerts interference effects in tissues in both the injected animal and its progeny. This technology is an extremely useful tool for identifying gene functions and evaluating

potential therapeutic targets. So, an increasing number of biotechnological and pharmaceutical companies are attempting to develop siRNA-based drugs for the prevention and treatment of human disease such as cardiovascular diseases, neurological diseases, viral infections, cancer etc.³⁸⁻⁴⁰

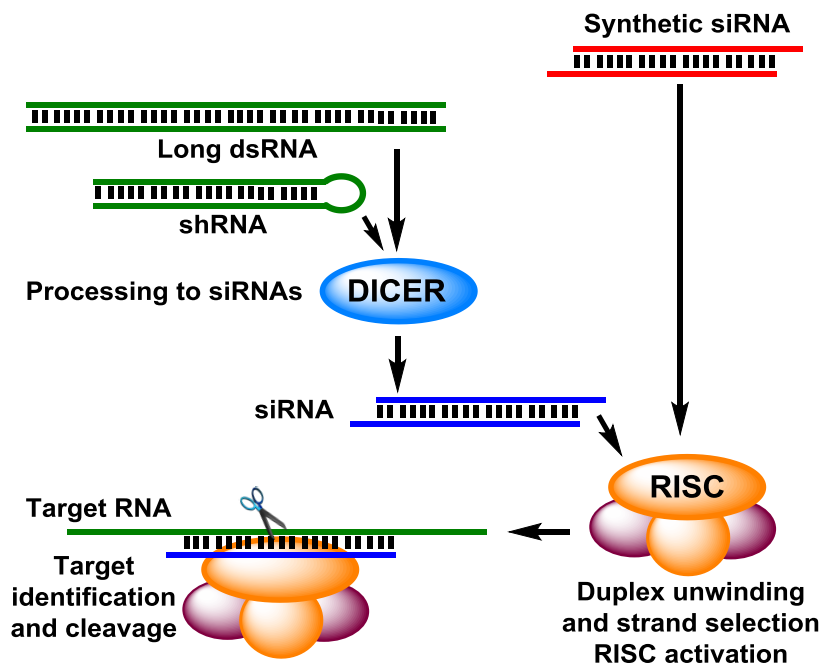


Figure 13. siRNA mediated down regulation of gene expression

miRNA: MicroRNAs (miRNAs),⁴¹ are a class of non-protein-coding RNAs that regulate gene expression post-transcriptionally. They are single-stranded RNAs (ssRNAs) of 19-25 nucleotides in length, generated from endogenous hairpin transcripts.⁴² miRNAs can also function to silence mRNA much like siRNAs by integrating into the RISC and guiding the degradation of target message. miRNAs regulate the gene expression by base-pairing to partially complementary sites on the target messenger RNAs (mRNAs), usually in the 3' untranslated region (UTR). Binding of a miRNA to the target mRNA typically leads to translational repression and exonucleolytic mRNA decay, although highly complementary targets can be cleaved endonucleolytically. miRNAs have been found to regulate more than 30% of mRNAs and have roles in fundamental processes, such as development, differentiation, cell proliferation, apoptosis, and stress responses. Both loss and gain of miRNA function contribute to cancer development through a range of different mechanisms.⁴³ Because miRNAs regulate cancer cell differentiation, proliferation, survival, and metastasis, manipulating miRNA function, either by mimicking or inhibiting miRNAs implicated in cancer, could provide a powerful therapeutic strategy to interfere with cancer initiation and progression.

piRNA: (Piwi-interacting RNA) piRNAs, a largest class of small non-coding RNAs, are 26–32 nucleotides in length and associate with Piwi proteins, such as Mili and Miwi and function in transposon silencing through heterochromatin formation or RNA destabilization via the formation of RISC.⁴⁴ They are distinct from microRNA in size (26–31 nt rather than 21–24 nt), lack of sequence conservation, and increased complexity.

1B.3.2 Antisense Oligonucleotides: Corrective antisense approach

Recent research opens up an exciting possibility for the application of AONs as corrective antisense in the cases of certain diseases caused by genetic mutations. In this approach the AONs can help restoring the viable protein production by acting at pre-mRNA levels for splice corrections^{45,46} or to yield mRNA that is translated into viable proteins. The disruptive antisense effects could involve non-antisense, non-specific interactions leading to stimulation of immune response but corrective antisense depends upon specific and stable interactions with mRNA for the desired corrective action.

Splice correction: Precursor mRNA, more commonly termed pre-mRNA, is an incompletely processed single strand of messenger ribonucleic acid (mRNA), synthesized from a DNA template in the nucleus of a cell by transcription. Once pre-mRNA has been completely processed, it is termed "mature messenger RNA", "mature mRNA", or simply "mRNA". Pre-mRNA includes two different types of segments, exons and introns (Figure 14). Most of exons encode protein, while introns do not and must be excised before translation. This process is called splicing.⁴⁷ Spliceosomes, small proteins found in the nucleus and composed of protein and RNA, perform the excision. An exon is any region of DNA within a gene that is transcribed to the final mRNA molecule, rather than being spliced out from the transcribed RNA molecule. Introns are sections of DNA co-linear to the mRNA sequence that will be spliced out after transcription, but before the mRNA is translated.

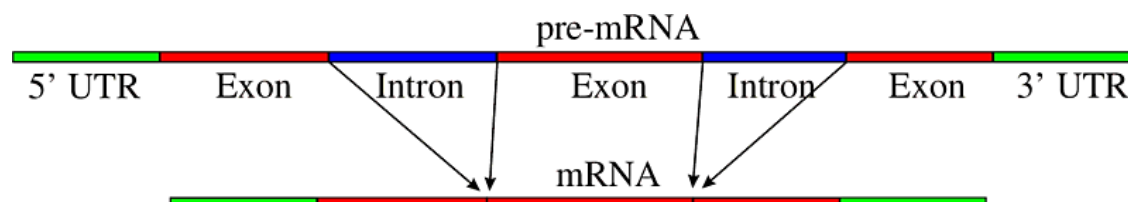


Figure 14. Simple illustration of exons and introns in pre-mRNA. The mature mRNA is formed by splicing. (UTR: Untranslated region)

Splicing can be experimentally modified so that targeted exons are excluded from mature mRNA transcripts by blocking the access of splice-directing small nuclear ribonucleoprotein particles (snRNPs) to pre-mRNA using oligonucleotides.

In order to meet all the requirements of a successful medicinal agent, it is necessary for normal oligonucleotides to be chemically modified in a suitable manner.

1C Chemical Modifications of DNA

One of the major challenges for antisense approaches is the stabilization of ONs, as unmodified oligodeoxynucleotides are rapidly degraded in biological fluids by nucleases. A vast number of chemically modified nucleotides have been used in antisense experiments most of which have been described in review articles and are under different phases of clinical trials.⁴⁸ To address the combined task of improving the rate, affinity or specificity of oligonucleotide recognition, while enhancing membrane permeability and resistance to nuclease digestion, several chemical modifications of DNA have been attempted (Figure 15). These can be broadly distinguished as (i) analogs with unnatural bases, (ii) modified sugars (especially at the 2' position of the ribose) (iii) altered sugar-phosphate backbone.

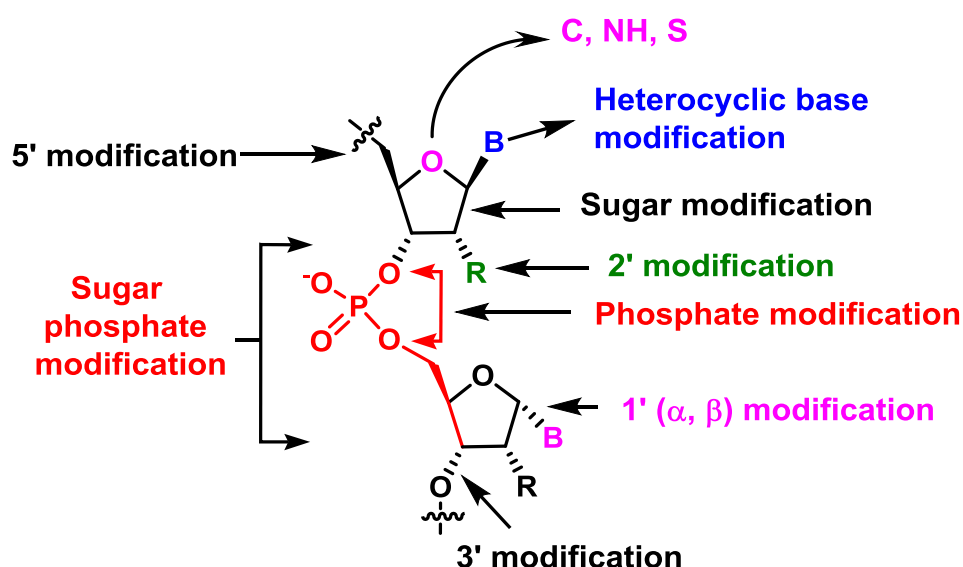


Figure 15. Different sites of modification for DNA

1C.1 First generation antisense oligonucleotides

1C.1.1 Alternative phosphate containing linkages

The modifications of phosphate moiety resulting in phosphorothioates⁴⁹ **1**, methylphosphonates⁵⁰ **2**, phosphoramidates⁵¹ **3** and phosphotriesters⁵² **4** and boranophosphonate⁵³ **5** have led to the first generation 'antisense' oligonucleotides (Figure 16). These have shown promising results with one drug *Fomivirsen* (marketed as *Vitravene* in 1998) based on the phosphorothioate backbone being approved by FDA for cytomegalovirus retinitis⁵⁴, but which was discontinued in 2004.

Phosphorothioates: In phosphorothioates (PS-oligos) negatively charged oxygen has been substituted by a sulfur atom at phosphorus centre, and the linkage thus retains a

formal negative charge, which is however reduced compared to the phosphodiester homologue. PS-oligos are easily synthesized on a commercial DNA synthesizer, as a mixture of diastereomers at the phosphorus atom, and are resistant to nuclease cleavage. PS oligomers can act as substrates to RNase H and these first generation antisense agents have been extensively tested in various human clinical trials against numerous targets. However, these oligomers have a tendency to induce non-specific effects, through binding to extra cellular and cellular proteins as well as cleavage of non-target mRNAs that are only partially complementary.

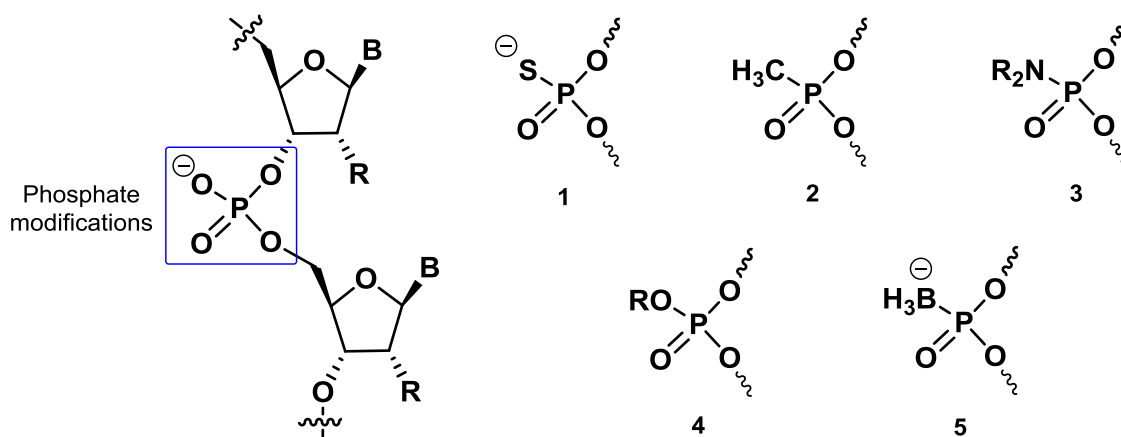


Figure 16. Structure of alternative linkages to phosphate

1C.1.2 Non-phosphorus backbone

Out of all the backbone modifications, those involving the replacement of the phosphodiester group with neutral linkers which are more lipophilic than the unmodified one, might increase the efficiency of passage into the cells, decrease the electrostatic repulsion with target nucleic acids and therefore improve affinity. Affinity can be improved with neutral linkages, provided that the linkage is pre-organized into the correct conformation for binding and is favorably hydrated in the duplex state.

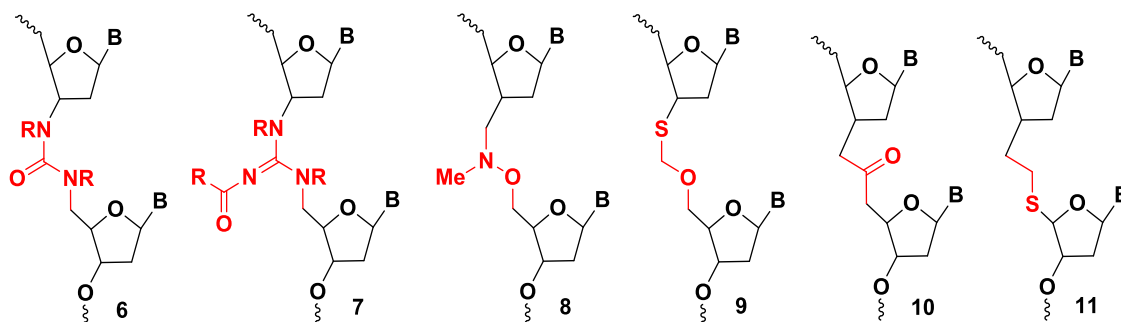


Figure 17. Structure of non-phosphorous backbone containing linkages

In the first group, four atoms are present in the backbone⁵⁵ (Urea **6**, guanidine **7**, Figure 17). Because of the conjugation of lone pairs of electrons from nitrogen and oxygen atoms with a carbonyl (or imine) group, three atoms are maintained in the same

plane. As a consequence, a rather severe destabilization of the duplexes was observed when these modifications were introduced into DNA strand.

The second group of modifications (hydroxylamine⁵⁶ **8**, thioformacetal⁵⁷ **9**; Figure 17), introduce less pronounced conformational restrictions. Interestingly, an increase in the size of the substituents on the nitrogen atom does not substantially destabilize the duplexes, when the amide group is located in the middle of the backbone.

In the third group of modifications (ketone **10**, thioether **11**; Figure 17), the backbone has a high degree of conformational freedom.⁵⁸ Consequently, much lower T_m s were obtained for the corresponding duplexes because of an increased negative entropy contribution to the hybridization process.

1C.2 Second generation antisense oligonucleotides

The instability of oligoribonucleotides under physiological conditions led to the synthesis of their *2'-modified* analogues. The problems associated with phosphorothioate oligodeoxynucleotides are solved to some degree by this type of *O-alkyl* modification at the 2' position of the ribose. *2'-O-methyl* **12** and *2'-O-methoxyethyl* RNA **13** (Figure 18) are the most important members of this class of oligonucleotides.⁵⁹ The oligonucleotides containing these building blocks are less toxic than phosphorothioate DNAs and have a slightly enhanced affinity towards their complementary RNAs. This class of compounds is attractive for antisense applications because ribonucleosides are much less costly as starting materials than their deoxy analogues and because of the greater thermal stability of the hybrids of 2'-*O*-methyl ribonucleotides with complementary RNA than that of corresponding unmodified DNA:RNA duplexes.

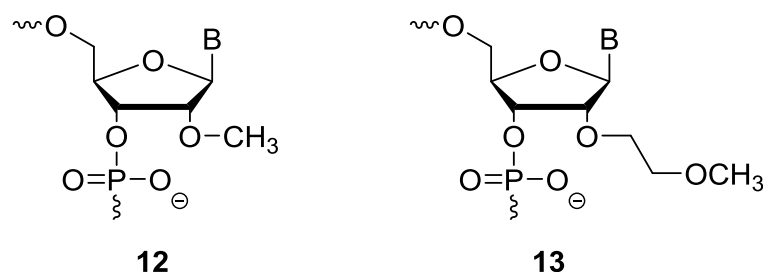


Figure 18. Structures of 2'-modified ribonucleic acid

These have shown promising results with one drug *Mipomersen* (marketed as *Kynamro*) based on the 2'-*O*-methoxyethyl modification with phosphorothioate linkages being approved by the United States FDA for the treatment of homozygous familial hypercholesterolemia in 2013.⁶⁰

1C.3 Third generation antisense oligonucleotides

The replacement of the ribose sugar by hexose or carbocycles has not been very successful in terms of specificity of binding/hybridization. However, morpholino oligomers,^{61, 62} (PMOs) where the monomers are linked through neutral carbamate linkages **14** or through phosphoramidate linkages⁶³ **15** (Figure 19), have shown promising antisense activity as they have superior permeability properties.

The locked nucleic acids (LNAs)⁶⁴ **16** (Figure 19) are oligonucleotides containing one or more 2'-O, 4'-C-methylene- β -D-ribofuranosyl nucleotides. These were found to exhibit unprecedented stability of their complexes with complementary DNA and RNA. They are also stable to 3'-exonucleolytic degradation and possess good water solubility. The conformational preorganization of LNA could be helpful in imparting the enhanced binding affinity to DNA.

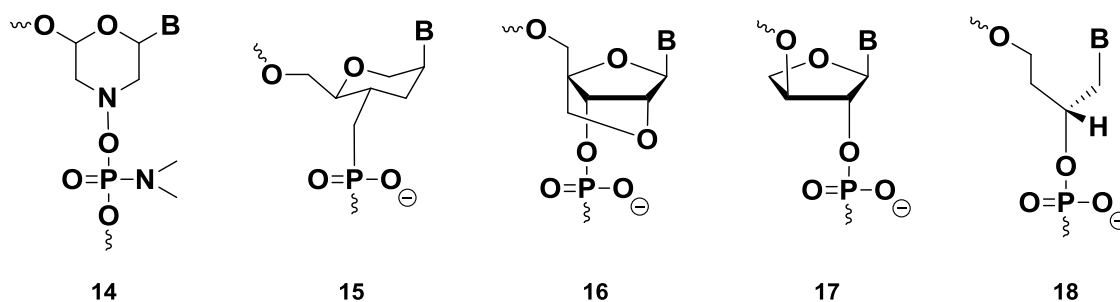


Figure 19. Structures of morpholino oligomer (PMO), phosphonate linkage, locked nucleic acid (LNA), α -threofuranosyl nucleic acid (TNA) and glycol nucleic acid (GNA)

Threofuranosyl Nucleic Acids (TNAs)⁶⁵ **17** (Figure 19) containing vicinally connected (3'→2') phosphodiester bridges undergo informational base pairing in antiparallel strand orientation and are capable of cross-pairing with RNA and DNA. Being derived from a sugar containing only four carbons, TNA is structurally the simplest of all potentially natural oligonucleotide-type nucleic acid alternatives.

Inspired by TNA structure, structurally simplified nucleic acid containing acyclic glycol nucleoside backbones has been synthesized in order to improve synthetic accessibility of artificial duplexes. The resulting Glycol Nucleic Acid (GNA)⁶⁶ **18** (Figure 19) forms highly stable antiparallel helical duplex structures following the Watson-Crick base pairing rules. Being the most atom economical solution for a functional nucleic acid backbone, GNA has been proposed as potential predecessor of RNA as a genetic material and catalyst for Earth's earliest organisms.

The attempts to optimize the properties of oligonucleotides have resulted in the synthesis and analysis of a huge variety of new oligonucleotide derivatives with modifications to the phosphate group, the ribose, or the nucleobases; some of them have

been mentioned in the above section. Most of the analogues being investigated so far are closely related to natural oligonucleotides and only few attempts to radically modify the backbone of DNA have been successful. The most radical change to the natural structure is the replacement of the entire sugar-phosphate backbone, known as peptide nucleic acid (PNA), will be discussed in detail in the following section.

1D Peptide Nucleic Acids (PNA)

1D.1 Introduction

Peptide Nucleic Acids (PNAs) are DNA analogues first introduced by Nielsen *et al.* in 1991 where the sugar-phosphate backbone has been replaced by a polyamide chain composed of *N*-2-aminoethylglycine repeating units, covalently linked to nucleobases through a carboxymethyl spacer⁶⁷ (Figure 20). They contain the natural nucleobases and are therefore able to bind to complementary DNA/RNA through the classical Watson-Crick base pairing rule, although some non-specific interactions were also recently revealed.⁶⁸ The PNA backbone is constituted by six atoms for each repeating unit and a two atom spacer between the backbone and the nucleobase, similar to the natural DNA.

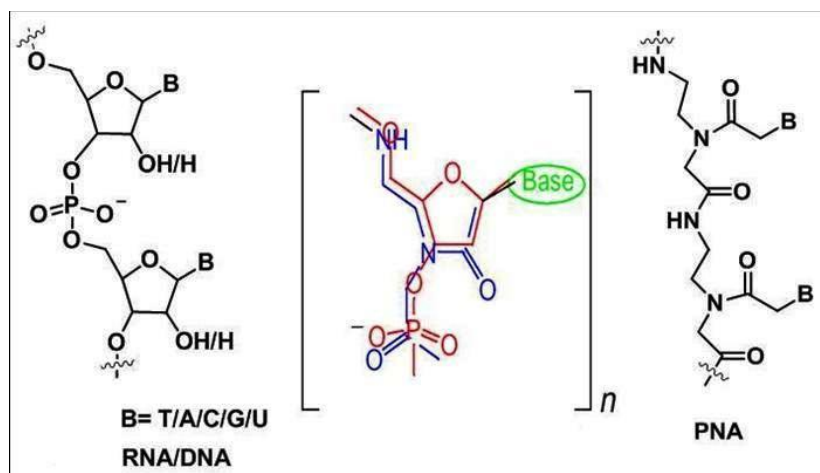


Figure 20. Chemical structures of DNA/RNA and PNA

By convention, PNAs are depicted like peptides, with the N-terminus (corresponding to 5'-end of DNA) at the first (left) position and the C-terminus (3'-end of DNA) at the right. PNA is set apart from DNA in that the backbone of PNA is acyclic, achiral and neutral. PNAs can bind to complementary nucleic acids in both antiparallel and parallel orientation, unlike natural DNA duplexes that always prefer antiparallel orientation. However, the antiparallel orientation is slightly preferred, and the parallel duplex has been shown to have different structure.⁶⁸ Because of very favourable hybridization properties⁶⁹ and high chemical and bio-stability,⁷⁰ it was

regarded as a very promising lead for gene targeted drugs *via* antisense/antigene technology,⁷¹ molecular biology,⁷² biosensors⁷³ and diagnostics markers.^{74, 75}

1D.2 Chemical and physical properties of PNA

1D.2.1 Duplex formation with complementary oligonucleotides

PNAs hybridize to complementary oligonucleotides obeying Watson-Crick base pairing rules. The internucleobase distances in PNA are conserved such that they complement the internucleobase distances in DNA for effective recognition of the base sequences. Therefore PNA is a true DNA mimic in terms of base pair recognition. Though in DNA:DNA duplexes the two strands are always in an antiparallel orientation (with the 5'-end of one strand opposed to the 3' end of the other), PNA:DNA adducts can be formed in two different orientations, arbitrarily termed parallel and antiparallel (Figure 21) both adducts being formed at room temperature, with the antiparallel orientation showing higher stability.⁷⁶ This creates the possibility for PNAs to bind two DNA tracts of opposite sequence.

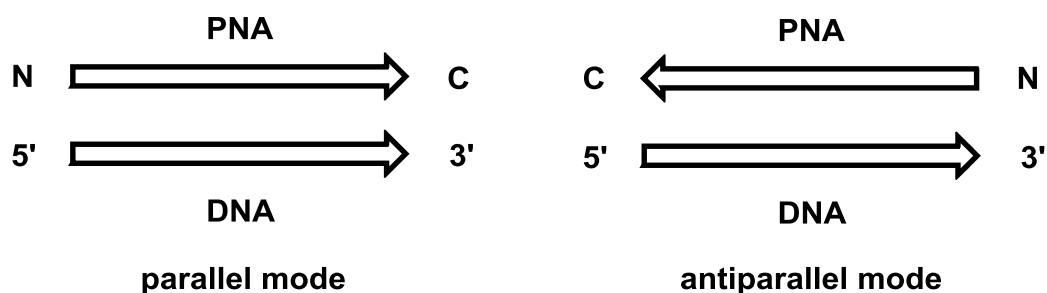


Figure 21. Parallel and antiparallel modes of PNA-DNA binding

An interesting aspect of PNA-DNA duplex formation is, the T_m decreases with increase in salt concentration (ionic strength) which is in contrast to that of DNA-DNA duplex, for which increase in T_m with salt concentration is observed.⁷⁷ Base pair mismatches result in a reduction of the T_m value of 8 – 20 °C.⁶⁸ This discrimination is, in some cases, approximately double that observed for DNA-DNA duplexes.

1D.2.2 Triplex formation of PNA

Polypyrimidine PNAs are able to form stable adducts with complementary polypurine DNA, through the formation of PNA₂:DNA triplexes.⁶⁷ The base pairing in these complexes occurs via Watson-Crick and Hoogsteen hydrogen bonds. If only one PNA sequence is used to form a PNA₂:DNA triple helix then both strands are necessarily either parallel or antiparallel to the DNA strand. When two different homopyrimidine PNA sequences are used the most stable complex is formed when the Watson-Crick PNA strand is oriented antiparallel and the Hoogsteen strand is parallel to the purine strand of the DNA. The sequence specificity of triple helix formation is

based on the selectivity of formation of the intermediate PNA:DNA duplex, whereas binding of the third strand contributes only slightly to selectivity. The stability of these structures enables PNA to perform strand invasion,⁷⁸ a property which is uniquely shown by PNAs. PNAs also form stable PNA₂-RNA triplexes with RNA.⁷⁹

Strand invasion: This is a unique property of PNAs to displace one strand of DNA double helix to form strand invasion complexes^{80, 81} (Figure 22), which is a favorable attribute for their application as antisense/antigene agents. The tendency of homopyrimidine PNAs to form PNA₂:DNA triple helices is so strong that under certain conditions they can bind as a triple helix to one strand of double-stranded DNA (dsDNA), while the second DNA strand is displaced and forms a single stranded loop structure (P loop). The prerequisites for strand displacement by PNA are (i) a DNA duplex that is not too stable and (ii) low salt concentration.

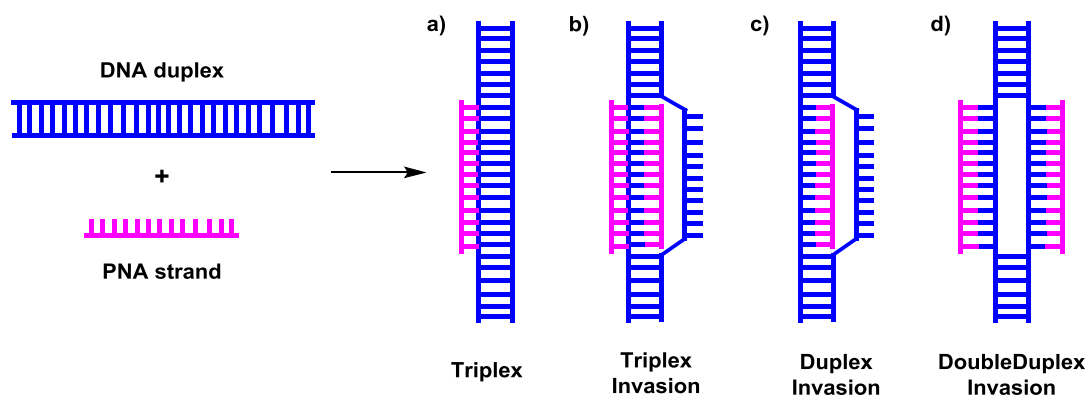


Figure 22. PNA binding modes for targeting dsDNA (a) Triplex (b) Triplex Invasion (c) Duplex Invasion (d) Double Duplex Invasion

For triple helix formation, strand displacement is independent of the orientation (parallel or antiparallel) of the pyrimidine PNA strand in relation to the purine strand of the complementary DNA. Strand invasion is highly sequence specific. The original studies of strand invasion were carried out on homopyrimidine PNAs at low salt concentrations.

1D.2.3 Quadruplex formation of PNA

The novel supramolecular architecture of G-quartets has led to the development of interesting and functional non covalent assemblies such as G-wires,⁸² ion-channels⁸³ and self-assembled ionophores.⁸⁴ PNAs have been developed to mimic Watson-Crick and Hoogsteen base-pairing, and they are expected to be in proper register for participating in G-tetrad formation as well. In an attempt to use this mode of molecular recognition homologous G-rich PNA and DNA oligomers hybridize to form a PNA₂-DNA₂ quadruplex.⁸⁵ Complementary cytosine-rich PNAs are capable of invading folded DNA or RNA quadruplexes to form standard heteroduplex structures.

Specifically, the PNA is designed to have the same (or similar) G-rich sequence as the target, in which case the PNA and DNA or RNA form a *heteroquadruplex* wherein the PNA and the target each contribute guanines to the H-bonded tetrads. The results have implications for the use of G-rich PNA for homologous hybridization to G-rich targets in chromosomal DNA and suggest additional applications in assembling quadruplex structures within lipid bilayer environments.⁸⁶

1D.3 Structures of PNA-DNA and PNA-RNA complexes

Three-dimensional structures have been determined for the major families of PNA complexes (Figure 23). Two duplex structures, a PNA:RNA and a PNA:DNA duplex, were solved by NMR, and a PNA₂:DNA triplex and a PNA:PNA duplex were solved by x-ray crystallography.

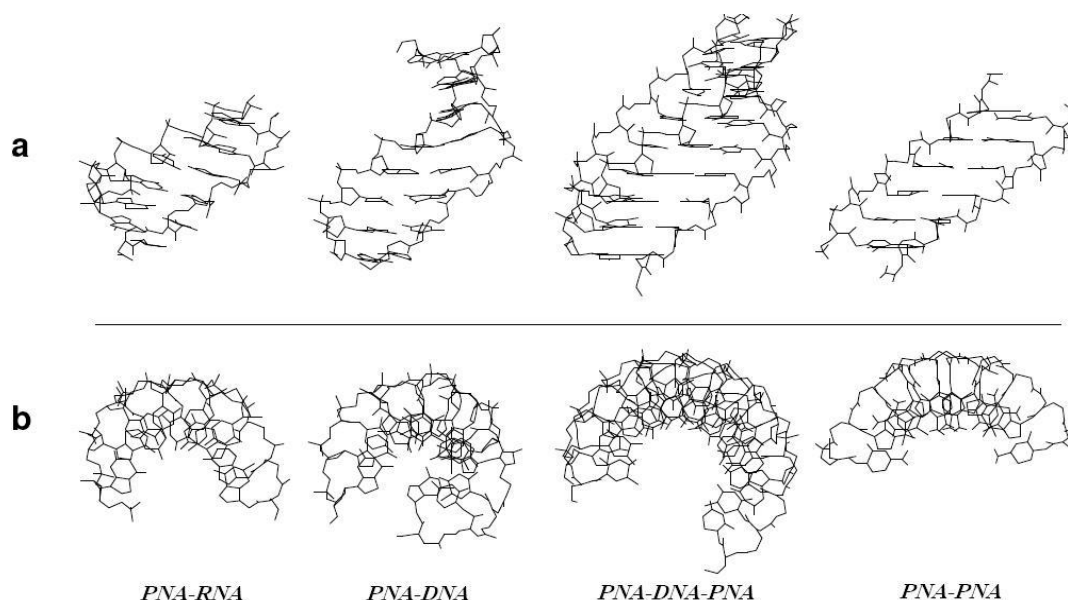


Figure 23. Structures of various PNA complexes in side view (upper panel), top view (lower panel)

PNA:DNA duplexes: The detailed structural information obtained from the NMR spectroscopic study of two antiparallel PNA-DNA duplexes⁸⁷ (8-mer and 10-mer) suggest that the DNA strand is in a conformation similar to the B form, binding through Watson-Crick base-pairing with a glycosidic *anti*-conformation, and the deoxyribose in the C2'-*endo* form. The NMR study of an octameric antiparallel PNA-DNA duplex showed the presence of elements for both A-form and B-form DNA. The right-handed helix contains approximately 13 base-pairs per turn compared to 10 base pairs for B form DNA with widened major groove and shallow and narrow minor groove. The primary amide bonds of the PNA backbone are in the *trans*-conformation and directed towards the solvent while the carbonyl oxygen atoms of the backbone-nucleobase linker point towards the carboxy terminus of the PNA strand. There is no evidence for the

formation of hydrogen bonds between the amide and carbonyl groups in the backbone. CD spectra indicate that the antiparallel PNA-DNA duplexes form a right handed helix,⁶⁸ however parallel PNA-DNA duplexes show a structure that differs considerably from that of either the B or A form.

PNA:RNA duplexes: ¹H NMR spectroscopy for PNA-RNA duplexes⁸⁸ indicate that all bases form Watson-Crick base pairs, the glycosidic torsion angle in the RNA strand indicates an *anti*-conformation, and the ribose sugars are in the 3'-*endo* form which resembles an A-form structure. The tertiary amide bonds of PNA are all in the *cis*-conformation and being the carbonyl group of the tertiary amide in the PNA backbone isosteric to the C2'-hydroxyl group, it increases the solvent contact of the carbonyl oxygen atom. The secondary amide protons of the backbone do not participate in any form of hydrogen bonding. The CD spectra of antiparallel PNA-RNA duplexes indicate the formation of a right-handed helix with geometry similar to the A or B form.

PNA₂:DNA triplexes: The structure of PNA₂:DNA triple helices was obtained from an X-ray crystal structure analysis of the complex formed by a bis-PNA and its complementary antiparallel DNA.⁸⁹ The structure is different from both A-form and B-form DNA, and forms a P helix with 16 bases per turn. The nucleobases of the PNA strand bind to the DNA by Watson-Crick and Hoogsteen base pairing and lie almost perpendicular to the helix axis, which is the characteristic of B-form DNA. The helix is considerably widened as compared to A-form DNA. The phosphate groups of DNA are hydrogen bonded to the PNA backbone amide protons of the Hoogsteen strand. These hydrogen bonds, together with additional van der Waals contacts and the lack of electrostatic repulsion, are the main factors responsible for the enormous stability of the triple helix. The deoxyribose of the DNA strand is in the C3'-*endo* conformation. The CD spectra of PNA₂:DNA triple helices indicate the presence of a right-handed helix and a geometry similar to that of the pure DNA triple helix.⁹⁰

PNA:PNA duplexes: The existence of both right and left handed helices, which are stacked coaxially and alternately in the crystal by forming a continuous pseudohelix, was confirmed by the X-ray crystal structure analysis of a self-complementary PNA:PNA duplex (H-cgtacg-NH₂).⁹⁰ The nucleobase-pairing is of Watson-Crick-type, where the bases lie almost perpendicular to the helix axis, with a propeller twist of 5-9° with 18 bases per turn compared to the 11 and 10 bases per turn in A- and B-form DNA, respectively. The amide groups of the backbone are in the *trans*-conformation

and the carbonyl groups of the linkers point towards the carboxy terminus. Thus, the structure bears a strong similarity to the P form of PNA₂:DNA helices.⁸⁹

1D.4 Applications of PNA

1D.4.1 Antisense effect or inhibition of translation

Unlike antisense oligonucleotides, PNA-RNA hybrids are not substrates for RNase H,⁸¹ and therefore PNAs must exert their antisense effect by other mechanisms, such as direct steric blocking of ribosomes or essential translation factors. The targets around the AUG initiation codon in general seem sensitive.⁹¹ Therefore, with general antisense gene targeting a sensible first choice would be targets close to or overlapping the AUG initiation site. In bacterial systems, there is strong evidence that the mRNA start codon region is susceptible to antisense inhibition.⁹² Translation experiments performed in cell-free extracts showed that duplex-forming PNA blocked translation in a dose-dependent manner when the target was 5'-proximal to the AUG start codon on the RNA, whereas PNAs had no effect when targeted towards sequences in the coding region.⁷⁹ Triplex-forming PNAs were efficient and specific antisense agents with a target overlapping the AUG start codon and caused arrest of ribosome elongation with a target positioned in the coding region of the mRNA. Experiments in progress indicate that antisense PNAs can be applied to a variety of bacterial species. There are many situations where it would be useful to specifically control bacterial gene expression with antisense agents.⁹²

1D.4.2 Antigene effect or inhibition of transcription

Homopyrimidine PNA oligomers bind to the homopurine sequence of the template strand resulting in efficient inhibition of transcription either by triple helix formation, or by a strand invasion. A high thymine content in the pyrimidine PNA shows more efficiency in blocking transcription process. Homopurine regions of eight base pair or more in length in double-stranded DNA (dsDNA) can be targeted by homopyrimidine PNAs via the formation of extremely stable PNA triplex strand invasion complexes.⁹³ For the formation of these complexes two PNAs are required where one of them is Watson-Crick bound and the other Hoogsteen bound, more often bis-PNAs in which the two parts are chemically linked are employed for dsDNA targeting.⁹⁴ The main obstacle to therapeutic use of the strand displacement principle is the stability of the GC-rich natural DNA double strand, under physiological salt conditions. Kinetics of strand invasion during the active transcription process increases

when the PNA oligomer is targeted against a transcriptionally active DNA fragment or DNA polymerase and thereby making PNAs good candidates for antigene reagents.

1D.4.3 Inhibition of Replication

In the DNA replication process, the elongation of DNA primers by DNA polymerases can be inhibited by PNAs *in vitro*. As a consequence, DNA replication can be inhibited by PNAs if the DNA duplex is susceptible to strand invasion under physiological salt conditions, or if the DNA is single-stranded during the replication process. In fact, in the case of extrachromosomal mitochondrial DNA, which is largely single-stranded during replication, efficient inhibition of replication by PNAs is possible.

1D.4.4 Interactions with Enzymes

Ribonuclease H (RNase H): The antisense oligonucleotides with an RNase H activity (e.g., PS-oligos) are considered a better antisense molecule (inhibitor) than one without the activity (methylphosphonates and HNAs).

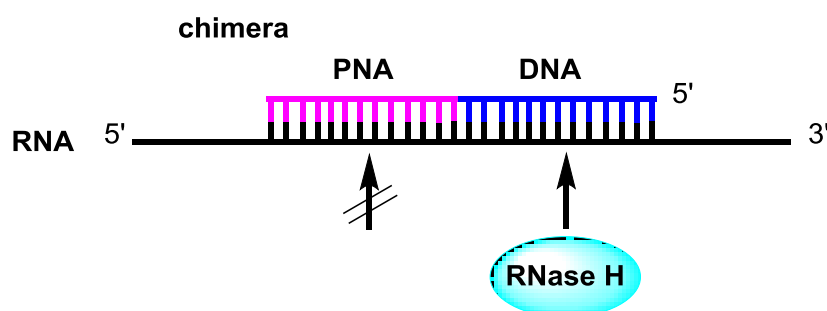


Figure 24. Schematic representation of RNase H-mediated cleavage of activity after the binding of a PNA-DNA chimera to an RNA target

Despite their remarkable DNA binding properties, PNAs generally are not capable of stimulating RNase H activity on duplex formation with RNA. However, recent studies have shown that DNA - PNA chimeras are capable of stimulating RNase H activity. On formation of a chimeric RNA double strand, PNA/DNA can activate the RNA cleavage activity of RNase H (Figure 24). Cleavage occurs at the ribonucleotide parts base-paired to the DNA part of the chimera. Moreover, this cleavage is sequence specific in such a way that certain sequences of DNA/PNA chimerae are preferred over others. They are also reported to be taken up by cells to a similar extent as corresponding oligonucleotides.⁹⁵ Thus, PNA - DNA chimerae appear by far the best potential candidates for antisense PNA constructs.

Inhibition of Telomerase activity: Long (TTAGGG)_n repeats at the 3'-end of DNA strands, are synthesized by human telomerase.⁹⁶ The ribonucleoprotein telomerase is composed of a protein component with DNA polymerase activity, and an

RNA component that acts as a primer binding site. The telomerase activity can be inhibited by PNA sequences that are complementary to the RNA primer binding site better than the corresponding phosphorothioate oligonucleotides because of their higher binding affinity. The activity is dependent upon the binding site and base composition of the PNA oligomer.

PNA directed PCR clamping: The “PCR clamping” method for the detection of point mutations is based on the ability to form strong PNA-DNA complex and inability to function as primers in the polymerase chain reaction (PCR). Targeting the PNA oligomer, even partially, against the primer binding site can block the formation of the PCR product.⁹⁷ Skilful choice of the length of the primer can even allow discrimination of alleles, which differ only in one base pair. Discrimination of point mutations in the *Ki-ras* gene has been carried out by PCR clamping, where the amplification of the wild-type DNA is inhibited, relative to mutated gene, which binds with lower affinity (Figure 25). Demers *et al.* have used PNAs to suppress the preferential amplification of small allelic PCR products during copying of VNTR *loci*, which leads to false genotypic patterns. The quantitation of telomeric repeats is also possible with PNAs. The use of PNAs immobilized on surfaces as sequence specific DNA biosensors has also been reported.⁹⁸

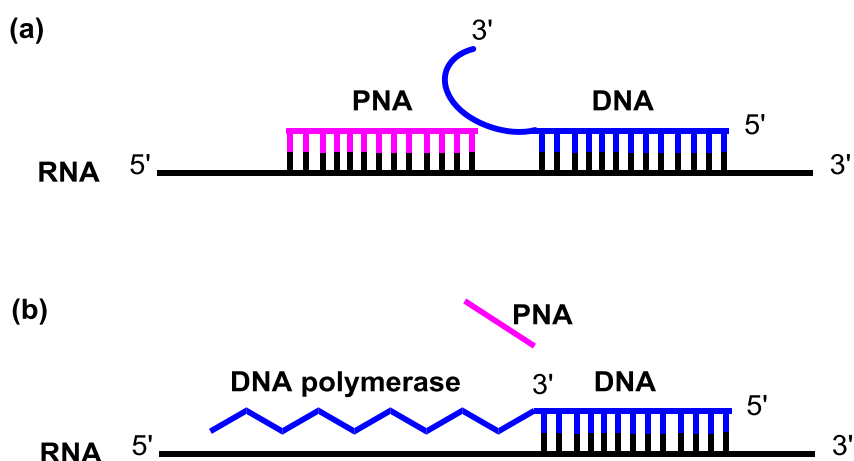


Figure 25. ‘PCR clamping’ technique. (a) inhibition of wild-type DNA amplification and (b) amplification of mutated DNA, by PNA

PNA as artificial restriction enzyme: PNAs in combination with nuclease S1 has been used as an artificial restriction enzyme.⁹⁹ The first model illustrates the hybridization of homopyrimidine PNA decamer to complementary target on double stranded DNA via a strand invasion mechanism leading to the formation of looped-out non complementary DNA strands (Figure 26). The enzyme nuclease S1 can degrade this single-stranded DNA part into well-defined fragments. It was suggested that the

enzyme first digested the displaced strand, and then the gap was enlarged to some extent, thereby allowing the opposite strand to act as a substrate for nuclease S1. When PNAs with a single mismatch were employed very weak cleavage was observed. This difficulty could be eliminated by the use of dimeric targets either in *cis* or *trans*.¹⁰⁰ The binding of PNA to two adjacent targets in *cis* led to an opening of the entire region, thereby provided an easily accessible substrate for the nuclease. This considerably increases the sensitivity to cleave and even more favourable situation arises if two PNA targets are on opposite strands.

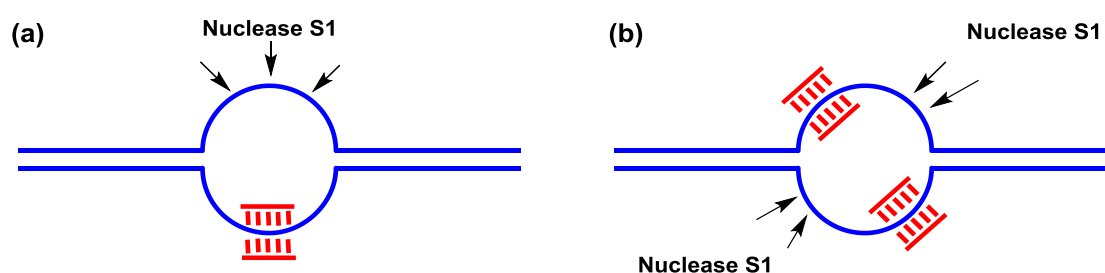


Figure 26. Artificial restriction enzymes (a) single strand cleavage by PNA
(b) double strand cleavage by double PNA clamping

Recently, Komiyama *et al.* have developed a completely chemistry-based artificial restriction DNA cutter (ARCUT) by combining a pair of pseudo-complementary PNA (pcPNA) strands (sequence recognition moiety) and Ce(IV)/EDTA complex (molecular scissors).¹⁰¹ The scission site of ARCUT and its scission specificity can be freely modulated in terms of the sequences and lengths of the pcPNA strands so that even huge genomes can be selectively cut at only one predetermined site. A pair of pcPNAs invades double-stranded DNA at the target site through Watson–Crick base pairing and generates single-stranded portions in both strands of the DNA (Figure 27). These single-stranded portions are preferentially hydrolyzed by Ce(IV)/EDTA complex due to unique substrate specificity of this complex (double-stranded DNA is hardly hydrolyzed).

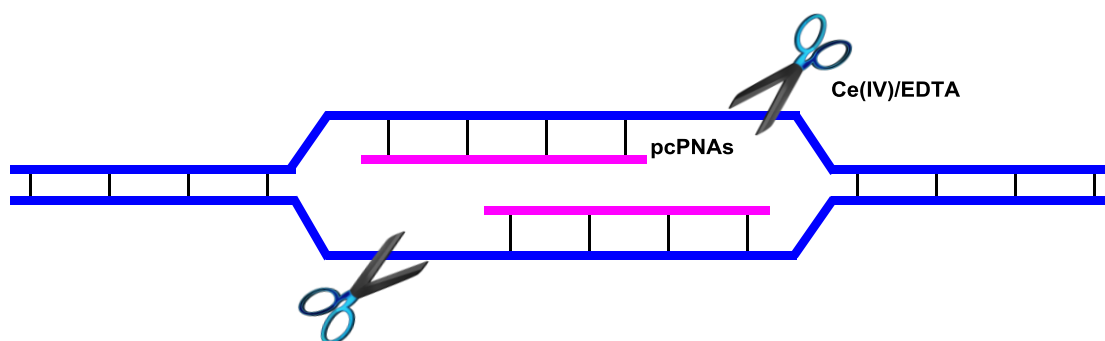


Figure 27. Strategy of site-selective DNA scission using ARCUT

1D.4.5 Nucleic acid purification

Based on its unique hybridization properties, PNAs can also be used to purify target nucleic acids. PNAs carrying six histidine residues have been used to purify target nucleic acids using nickel affinity chromatography. Also, biotinylated PNAs in combination with streptavidin-coated magnetic beads may be used to purify *Chlamydia trachomatis* genomic DNA directly from urine samples. However, it appears that this simple, fast, and straightforward ‘purification by hybridization’ approach has drawbacks. It requires the knowledge of a target sequence and depends on a capture oligomer to be synthesized for each different target nucleic acid. Such target sequences for the short pyrimidines PNA, i.e., the most efficient probe for strand invasion, are prevalent in large nucleic acids. Thus short PNAs can also be used as generic capture probes for purification of large nucleic acids. It has been shown that a biotin-tagged PNA-thymine heptamer could be used to efficiently purify human genomic DNA from whole blood by a simple and rapid procedure.

1E Chemical Modifications of PNA

Since their discovery, many modifications¹⁰²⁻¹⁰⁵ of the basic PNAs structure have been proposed, in order to improve their performances in terms of affinity and specificity towards complementary oligonucleotides. A modification introduced in the PNA structure can improve its pharmacological potential generally in three different ways:- (i) improving DNA binding affinity, (ii) improving sequence specificity, in particular for directional preference (antiparallel vs parallel) and mismatch recognition/discrimination, (iii) improving bioavailability (cell internalization, pharmacokinetics, etc.).

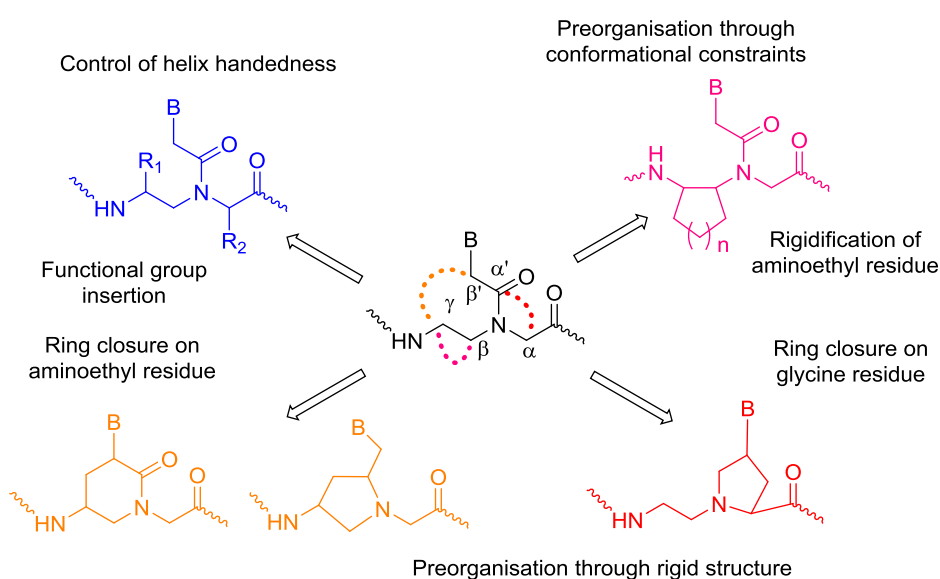


Figure 28. Strategies for inducing preorganization in the PNA monomers

The main strategies which have been used for achieving this goal are summarized in Figure 28. Preorganization was achieved either by cyclization of the PNA backbone (in the aminoethyl side or in the glycine side), by adding substituents at α or γ carbon of the monomer or by inserting the aminoethyl group into cyclic structures. The addition of substituents at α or γ carbon of the monomers can also preorganize the PNA strand, but mainly it has the effect of shifting the PNA preference towards a right-handed or left-handed helical conformation, according to the configuration of the new stereogenic centers, in turn affecting the stability of the PNA:DNA duplex through a control of the helix handedness.

1E.1 Preorganization through conformational constraints

The earliest and simplest of the modifications involved extension of the PNA structure with a methylene group individually in each of the structural sub-units (glycine, aminoethyl and base linker) of the PNA monomer.¹⁰⁶ These resulted in PNAs with *N*-(2-aminoethyl)- β -alanine **19**, *N*-(3-aminopropyl) glycine **20** backbone and ethylene carbonyl linked nucleobases **21**. However, these modifications resulted in a significant lowering of T_m of the derived PNA:DNA hybrids. The deleterious consequences of such subtle changes to the PNA structure suggested the high structural organization to which the original PNA structure is inherently tuned for interaction with DNA.

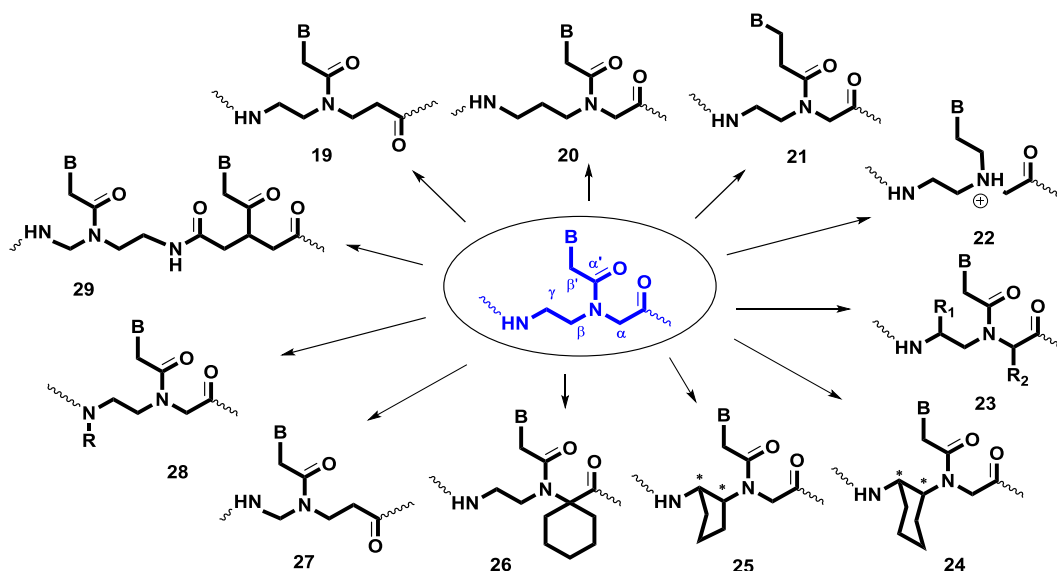


Figure 29. Chemical modifications of PNA

The replacement of the tertiary amide carbonyl by a methylene group leading to a flexible, cationic tertiary amine monomer **22** resulted in a large destabilization of the PNA:DNA hybrids.¹⁰⁷ The necessity of such a pseudo rigid amide group pointed to the importance of constrained flexibility in the backbone. Further rigidification of the PNA

backbone has been attempted by introduction of alkyl substituents individually or simultaneously in the aminoethyl or glycine segments or in both (**23**).¹⁰⁸⁻¹¹⁰ Number of modifications generated by substitution of glycine component by other α - amino acids, leading to chiral PNA **23** having hydrophobic, hydrophilic or charged α, β, γ - substituents have been reported.¹¹¹⁻¹¹³

PNA oligomers incorporating chiral monomers retained the hybridization properties though less efficiently, with tolerance for small and medium substituents at the glycine- α position. Substitution with L/D-alanine showed a slight preference for antiparallel binding with DNA, with D-alanine being slightly better than L-alanine.¹¹⁴ Among other replacements, only those derived from D-lysine exhibited DNA hybridization properties as good as that of original PNA. The incorporation of chiral monomers enhanced the sequence selectivity of PNA oligomers in hybridization, with maximum for D-glutamic acid and D-lysine substitutions. The lysine-modified oligomers were also more readily soluble in aqueous systems. In general, the different substituents caused equal or lower destabilization of PNA:RNA hybrids as compared to PNA:DNA hybrids.

Suitable substitutions may also lead to generation of cyclic structures with 1,2-cyclohexylamino¹¹⁵ **24**, 1,2-cyclopentylamino¹¹⁶ **25** and spirocyclohexyl¹¹⁷ **26** rings in monomers. Several papers have reported a systematic approach in the design of monomers exhibiting DNA/RNA selectivity. The *trans*cyclopentane (*tcyp*-PNA) modified PNA structure **25**, which is equivalent to covalently closing the aminoethyl glycine PNA backbone, has been successfully utilized as a target capture strand to improve the detection limit of a known DNA detection assay, and provided high levels of mismatch discrimination. *cis*-(1*R*,2*S*)-cyclopentyl PNA analogues **25**, (the same as *tcyp*-PNA, except for stereochemistry) hybridize to DNA/RNA without discrimination because of the ring puckering of the cyclopentane ring.

In contrast to the rigid locked chair conformation of a cyclohexane system, in fact the cyclopentane ring is more flexible and can be easily conformationally adjusted. The *ch*PNAs **24** show remarkable differences in duplex stability with their DNA and RNA complexes, depending on number of modifications and stereochemistry. There is a highly significant preference to form a duplex with RNA as compared to DNA. PNAs that bear (*S,S*) cyclohexyl ring **24** in the aminoethyl part hybridize with complementary DNA similar to the unmodified PNA, those derived from (*R,R*) cyclohexyl **24** moiety lack such a property.

Another type of modification involved interchange of various CO and NH groups on the peptide linkages leading to retro inverse¹¹⁸ **27**, peptoid¹¹⁹ **28** and heterodimeric¹¹⁸ **29** analogues. In all these systems, the inter-base residue separations are similar to the unmodified PNA, but accompanied by inversion of intra and inter residue amide bonds. Except for the heterodimer analogue, these exhibited a lower potency for duplex formation with complementary DNA/RNA suggesting that in addition to geometric factors, other subtle requirements such as hydration and dipole-dipole interactions, etc influencing the microenvironment of the backbone, may be involved in effecting efficient PNA:DNA hybridization.

1E.2 Preorganization through rigid five membered heterocycles

Some of the relatively successful conformationally preorganized modifications¹⁰³⁻¹⁰⁵ so far are based on introduction of methylene/ethylene groups to bridge the aminoethyl-glycyl backbone and methylene carbonyl side chain to generate diverse five- or the six- membered nitrogen heterocyclic analogues. The cyclic analogues where the nucleobases are directly attached to the ring have defined nucleobase orientation, overcoming the rotamer problem. It also concomitantly introduces chiral centers, which may impart directional selective binding of PNA with chiral DNA/RNA.

The naturally occurring amino acid *trans*-4-hydroxy-L-proline, a five-membered nitrogen heterocycle, is a versatile, commercially available starting material for creating structural diversity to mimic DNA/PNA structures. From this amino acid a wide variety of chiral, constrained and structurally preorganized PNAs have been synthesized. Depending on the synthetic approach and on the presence of the tertiary amine group in the monomers, the modifications afford either positively charged or uncharged cyclic PNA analogs. The different cyclic PNAs proposed showed that the right stereochemistry and conformation is really important for binding abilities towards nucleic acids and in some cases even for discrimination between RNA and DNA.

For example, in the case of the cyclic PNA analogue *N*-(thymine-1-yl-acetyl)-4-aminoproline **30** (Figure 30), all stereoisomers were synthesized, and the *L-trans*-4-aminoprolyl isomer, was shown to bind to DNA with higher affinity, while the *L-cis* isomer and the *D-trans*, which could not adopt the same spatial arrangement, showed reduced binding performances.¹²⁰ The same model can be applied to the (2*R*, 4*S*)-stereoisomer of the aminoethylprolyl PNA **31** (aepPNA).¹²¹ The aminoethylprolyl-5-one **32** (aepone) thymine monomers¹²² were synthesized and incorporated into

aegPNA-T8 backbone at different positions. The *aepone*-PNAs showed remarkable stabilization of derived PNA₂:DNA triplexes compared to *aegPNA*. An early example of a PNA analogue that induced kinetic selectivity for RNA binding was the POM-PNA, a pyrrolidine-amide oligonucleotide mimic (Figure 30, **33**, $n = 1$).¹²³ The pyrrolidine ring in the *2R, 4R* stereochemistry was selected as it was expected to adopt the sterically less demanding *trans* relative configuration making the system a stereochemical match with native nucleic acids. X-ray crystallographic data for a similar protonated pyrrolidine ring system showed that the alkyl substituent adopted the *trans* relative stereochemistry with close conformational resemblance to native RNA like C3'-endo sugar conformation.¹²⁴ It was found that short pentameric thyminy-POM sequences bind much faster to RNA than to DNA. Our group studied the binding of oligomers of enantiomeric pyrrolidinyl-PNA with *2S, 4S* stereochemistry (Figure 30, **33**, $n = 1$) and found that it leads to destabilization of complexes with either DNA or RNA.¹²⁵ Further our group extended the arguments of a five-atom amide linker between the nucleosides that stabilized the complexes formed with RNA to the *2S, 4S*-pyrrolidinyl-PNA and arrived at the backbone extended pyrrolidinyl PNA (*bepPNA*) (Figure 30, **33**, $n = 2$). Indeed, alternating PNA:*bepPNA* backbone formed very stable complexes with the target RNA.^{126, 127}

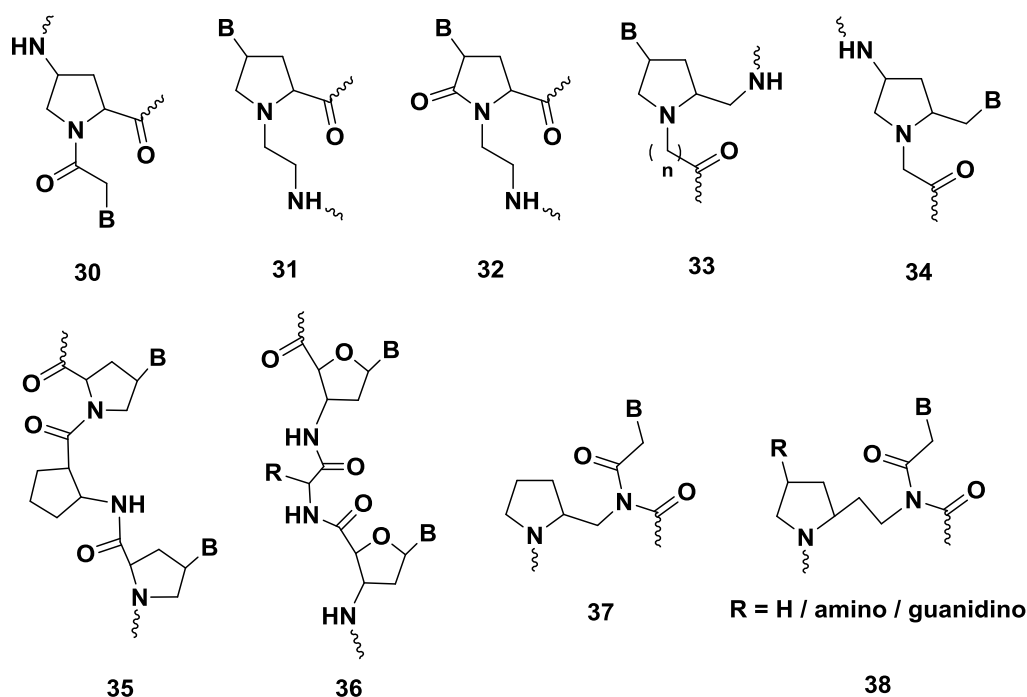


Figure 30. Conformationally preorganized five membered heterocyclic PNA monomers and Oligomers with ($\alpha + \beta$) amino acid backbones

Introduction of a methylene bridge between the β' and β carbon yields another pyrrolidine-PNA **34**.¹²⁸ Diastereomeric monomers bearing T, A, C, G nucleobases were

introduced in PNA oligomers and the complexation with DNA and RNA sequences was studied. It was found that (2*R*,4*S*) homopyrimidine PNA stabilize PNA₂:DNA triplexes, (2*S*,4*R*) stereoisomer in mixed sequences effects enhanced DNA duplex stability, (2*S*,4*S*) and (2*R*,4*R*) enhance PNA:RNA duplex stability.¹²⁹ Conformationally restricted pyrrolidinyl PNAs with an α/β -dipeptide backbone consisting of a nucleobase-modified proline and a cyclic five-membered amino acid spacer such as (1*S*,2*S*)-2-aminocyclopentanecarboxylic acid (ACPC) (Figure 30, **35**, ACPC-PNA) also formed very stable hybrids with DNA with high Watson-Crick base pairing and exhibited high preference for binding to DNA over RNA.¹³⁰ Using this knowledge-based previously our group synthesized a nucleic acid backbone where nucleoside derived β -amino acid was utilized to synthesize oligomers alternating with α -amino acids (Figure 30, **36**).¹³¹ In this case, the effect of chirality of the linker amino acid was found to have significant effect on the stability of duplexes with both RNA and DNA. Such preference of one stereochemistry over the other in the backbone was also observed earlier in the case of (POM)-PNA.¹²³

In most of the studies discussed above it seems that the conformational constraint when applied on two different segments of PNA *i.e.* aminoethyl, glycyll or the nucleobase linker by the bridged atoms did not yield uniform binding patterns in mixed base sequences. Recently in literature a PNA modification based on proline, (Figure 30, **37**) constraint in the PNA structure only in aminoethyl segment was reported.¹³² This modification still destabilized the complexes with DNA/RNA. Further from our group, we extended the backbone by one carbon unit, *pet*-PNA (Figure 30, **38**) and found that the complexes were still destabilizing albeit less compared to **37**. By substituting the pyrrolidine ring in the backbone with amino- and guanidino- groups (Figure 30, **38**), a locked amino/guanidino- substitution could be achieved at extended γ - position. With this modification, we were able to show improved binding with cDNA targets but observed destabilization of the complexes with RNA in duplex forming mixed purine pyrimidine sequences.¹³³ Interestingly, when four modified units of **38** were present in a pyrimidine sequence containing cytosine and thymine bases, a highly stable triplex was formed with RNA target as well.

For the development of PNA towards therapeutic applications, the major drawback has been its inability to cross the cell membrane. The amino-/guanidino-substitutions in PNA discussed above draw attention to the directed methods for

chemical interventions to open up the biological barriers to this important class of DNA mimics.

1E.3 Preorganization through rigid six membered heterocycles

The behaviour of two stereoisomers of a 6-membered ring PNA is the discrimination between the parallel versus antiparallel DNA reported for a single modified (2*S*,5*R*) aminoethyl pipercolic PNA¹³⁴ unit **39** (Figure 31) inserted in the middle of the strand. The effect is highly dependent on the position of the stereocenters, as already seen for flexible chiral PNAs. PNAs derived from piperidinone¹³⁵ **40** were also reported, though the rigidity of the ring and their particular geometry led to a decrease of PNA:DNA duplex stability; however, the best performing was the (3*R*,6*R*)-isomer, which is in line with proper group arrangement.

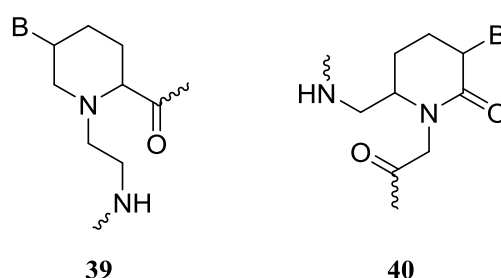


Figure 31. Structures of pipercolic PNA and piperidinonePNA monomers

1E.4 Modified Nucleobases

There is increasing interest in modulating and expanding the recognition motifs of standard base pairs. Employing non-natural nucleobase ligands in place of natural nucleobases would help understand the recognition process in terms of various factors contributing to the event such as hydrogen bonding and internucleobase stacking. Further, new recognition motifs may also have potential applications in diagnostics and nanomaterial chemistry. This when coupled with high affinity and strand invasion properties offered by PNA would add a new dimension to PNA applications. The non standard nucleobases employed so far with PNA are limited, compared to the repertoire of backbone modifications described earlier. 2,6-Diaminopurine¹³⁶ (Figure 32, **41**) offers increased affinity and selectivity for thymine and pseudoisocytosine⁹⁴ **42** which is a very efficient mimic of protonated cytosine for triplex formation. 2-Aminopurine¹³⁷ **43** hydrogen bonds with U and T in reverse Watson-Crick mode and has the advantage of being inherently fluorescent to enable study of kinetic events associated in hybridization. The E-base¹³⁸ **44** was rationally designed for recognition of A:T base pair in the major groove and forms a stable triad with T in the central position.

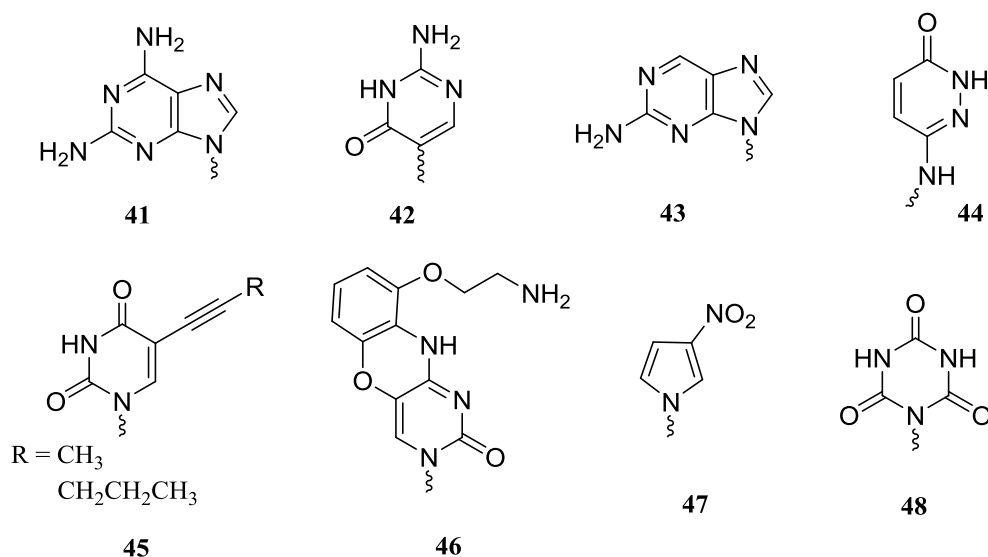


Figure 32. Structures of modified nucleobases

In order to increase the stability of complexes formed by PNA with target nucleic acids, bases that possess a larger surface area (for greater hydrophobic/stacking interactions), make additional H-bonds or that are positively charged have been prepared. A wide variety of 5-substituted uracils were synthesized and their ability for triplex formation has been studied¹³⁹ (Figure 32, **45**). The G-clamp base **46** was developed to build in specific, additional bonding interactions with guanine.¹⁴⁰ Unnatural heterocycles 3-nitropyrrole **47** have been used as potential universal bases in PNA.¹⁴¹ The synthesis of cyanuryl PNA monomer containing cyanuric acid **48** as the base was achieved by direct *N*-monoalkylation of cyanuric acid with *N*-(2-Boc-aminoethyl)-*N'*-(bromoacetyl)glycyl ethyl ester.¹⁴² The monomer was incorporated as a T-mimic into PNA oligomers and biophysical studies on their triplexes/duplex complexes with complementary DNA oligomers indicated unusual stabilization of PNA:DNA hybrids when the cyanuryl unit was located in the middle of the PNA oligomer.

1E.5 Stability in cells and cellular uptake of PNA

The use of PNAs as antisense or antigene therapeutics requires that they show sufficiently high biological stability in serum and in cells. The unmodified PNA oligonucleotides, having the peptide-like structure make them potentially susceptible to degradation by peptidases or proteases in serum. However, PNAs have a remarkably high biostability in both human serum and in cell extracts.⁷⁰ The experiment was carried out with the homopolymer H-(t)₁₀-Lys-NH₂ where no significant degradation could be detected by HPLC after a two hour incubation of the human serum, or in cytoplasmic, or nuclear fractions of mouse tumour cells.

Although PNA binds single-stranded DNA and RNA with superior affinity and selectivity, there are other properties of PNA that can be further improved. Most importantly, *in vivo* applications of unmodified PNA are hindered by poor cellular uptake and endosomal entrapment.¹⁴³ Current methods to enhance the cellular uptake of PNA, such as attachment of lipophilic or other helper groups to PNA that bind selectively to cell surface receptors, by formation of PNA-DNA chimeras, conjugation with cell penetrating peptides (CPP),^{144, 145} are complicated and require high PNA-peptide concentrations that may cause off-target binding and toxicity *in vivo*.

1E.5.1 PNA- DNA Chimeras

The successful applications of the remarkable DNA binding properties of PNAs are sometimes (sequence/length dependent) hampered by their tendency to self-aggregate and poor aqueous solubility. For overcoming these limitations, covalent hybrids or chimeras of PNA with DNA, were designed. Three types of PNA-DNA chimeras (Figure 33) are (i) 5'-DNAlinker X- PNA -*pseudo*-3' **49**¹⁴⁶ (ii) *pseudo*-5'-DNA-linker X- DNA-3' **50**¹⁴⁷ and (iii) *pseudo*-5'-PNA-linker X- DNA-3' **51**.¹⁴⁶ Synthetic protocols have been developed with protecting groups compatible for carrying out on-line synthesis of both PNA and DNA to generate the chimeras. Several interesting properties were noticed in such covalent hybrids such as co-operative stabilizing effects against proteases and nucleases, enhanced water solubility and duplex/triplex stabilities dependent on the structure of chimera and the linker.

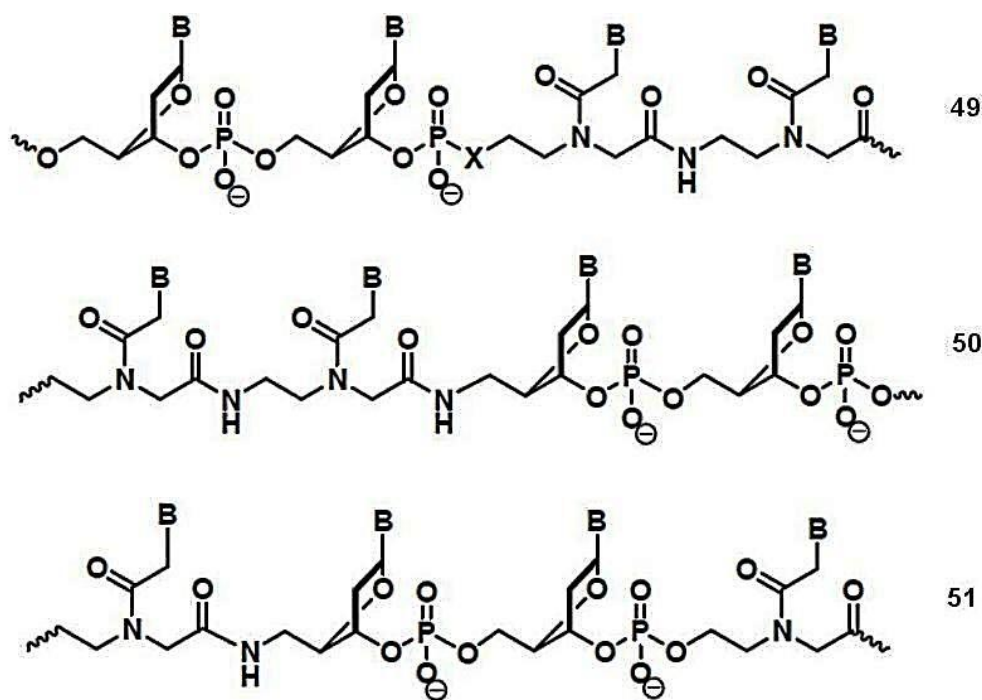


Figure 33. Structures of DNA-PNA chimera

1E.5.2 Conjugation of PNA with Cationic Peptides

Because of the neutral backbone, PNA does not associate with delivery vehicles based on cationic lipids. To use such standard oligonucleotide transfectants as Lipofectamine, PNA needs to be hybridized to complementary oligodeoxynucleotide (ODN) that aids the electrostatic complexation with the positively charged lipids.¹⁴⁸ Recently, a new approach to PNA delivery was developed by Wooley, Taylor and coworkers¹⁴⁹ who used cationic shell-cross-linked knedel-like nanoparticles (cSCKs) to deliver either PNA-ODN hybrid or PNA covalently attached to cSCKs nanoparticles through a biodegradable disulfide linkage. cSCKs nanoparticles have a hydrophobic core and a positively charged cross-linked shell. The latter is highly functionalizable and mediates the cellular delivery through, most likely, an endocytotic mechanism. An elegant extension of this technology was reported by Taylor and coworkers.¹⁵⁰

Perhaps, the most popular approach to enhance cellular delivery has been conjugation of PNA with cell penetrating peptides that deliver the conjugate through the endocytosis pathway.^{144, 145} However, the low ability of PNA-CPP conjugates to escape from endosomes has been the bottleneck of this approach. Various endosomolytic compounds have been explored; unfortunately, most are too toxic for *in vivo* applications.¹⁴⁴ Conjugates with arginine-rich peptides have shown promising activity in HeLa cells in the absence of endosomolytic agents.¹⁵¹ Moreover, CPPs are relatively large peptides, which complicate the preparation and use of PNA-CPP conjugates. Recently, several groups have demonstrated that relatively simple cationic modifications in PNA can substantially improve their cellular uptake and produce effects similar to that of longer and more complex CPPs.

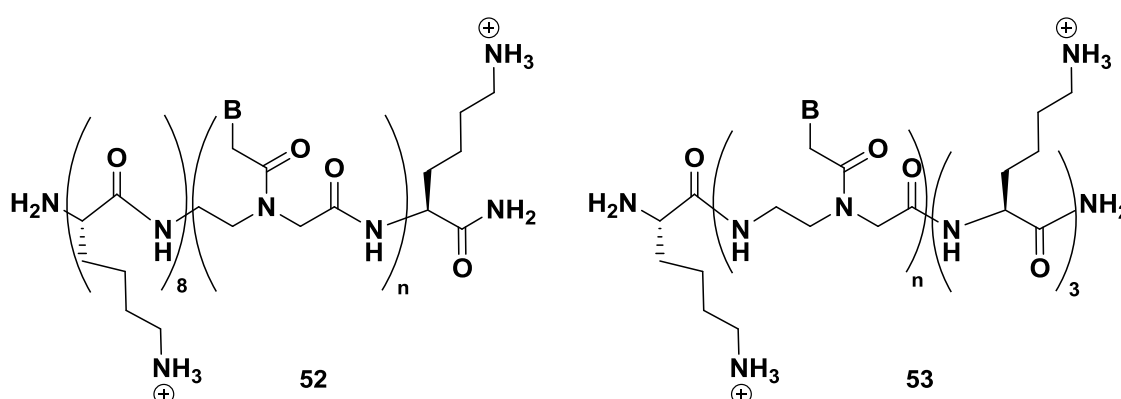


Figure 34. Conjugation of PNA with short oligolysines improves cellular uptake

The groups of Corey¹⁵² and Gait¹⁵³ showed that conjugation of PNA with short oligolysine (Figure 34, **52** and **53**, resp.) enabled efficient delivery in fibroblast and various cancer cell lines (T47D, MCF-7, Huh7, and HeLa). As few as four lysine

residues achieved similar efficiency as R6-Penetratin, a CPP previously optimized for cellular delivery of PNA.¹⁵¹ Using short oligolysine instead of longer CPP significantly reduced the complexity and effort required for PNA use in cell culture. Most recently, Gait and coworkers showed that introduction of a terminal Cys residue further increased the cellular uptake of Cys-Lys-PNA-Lys₃ conjugate.¹⁵⁴ While some studies showed that conjugates built of the unnatural D-lysine were more effective,¹⁵⁵ presumably due to higher biostability, other studies found little difference between the L and D series.¹⁵⁴ In a similar study, Fabbri *et al.* demonstrated that PNA conjugated at the carboxyl terminus with octaarginine was efficiently taken up in human leukemic K562 cells and inhibited activity of the target microRNA-210.¹⁵⁶

Nielsen and coworkers have recently reported on conjugates of PNA with cationic ligands that showed improved cellular delivery and activity. In one study, addition of a lipid domain to the cationic peptides increased the activity of PNA conjugate by two orders of magnitude.¹⁵⁷ The lipophilic fatty acid contributed by promoting both endosomal uptake and endosomal escape of PNA. In another study, conjugation of PNA with polyethylenimine showed significantly higher antisense activity than PNA octaarginine conjugates.¹⁵⁸ Polyethylenimine conjugates had lower toxicity than PNA-octaarginine conjugates. The polyethylenimine conjugate activity did not depend on the presence of lysosomolytic agents, which suggested that these conjugates are able to escape endosomes efficiently. These studies suggest that chemical approaches can be used to tailor cationic modifications that will improve cellular uptake and avoid the problem of endosomal entrapment.

Conjugation of PNA with a lipophilic triphenylphosphonium cation has been shown to increase the cellular delivery. Pandey, Patino and coworkers reported on cyclic and hairpin PNAs conjugated to the triphenylphosphonium cation via a disulfide linkage.¹⁵⁹ The conjugates inhibit HIV replication by targeting the HIV-1 TAR RNA loop.

Most recently, Shiraishi and Nielsen¹⁶⁰ reported on cellular uptake and antisense activity of PNA conjugated with cholesterol and cholic acid in HeLa pLuc705 cells. Although the conjugates alone were inactive, the delivery was dramatically improved by addition of Lipofectamine leading to nanomolar antisense activity.

As the numerous recent studies reviewed above suggest, design and optimization of CPP and other cationic ligands for cellular delivery of PNA is still a vigorous and important area of research.

1E.5.3 Cationic Backbone Modifications of PNA

An alternative approach to conjugation of PNA has been direct modification of PNA's backbone. Several groups have explored cationic modifications of PNA. Ly and coworkers introduced guanidine groups at α -¹⁶¹ and γ -¹⁶² positions of PNA's backbone (Figure 35, **54** and **55** resp.) by custom synthesis of monomers starting from diaminoethane and L- or D- arginine instead of glycine. The α -guanidine-modified PNA (GPNA) derived from the unnatural D-arginine had higher affinity for complementary DNA¹⁶³ and RNA.¹⁶⁴ GPNA was readily taken up by several cell lines (HCT116, human ES, and HeLa), which was attributed to the cationic guanidine modifications. The γ -guanidine-modified PNA had higher affinity for complementary DNA and RNA than α -guanidine-modified PNA, presumably due to favorable preorganization of the γ -modified backbone into a right-handed helix.¹⁶²

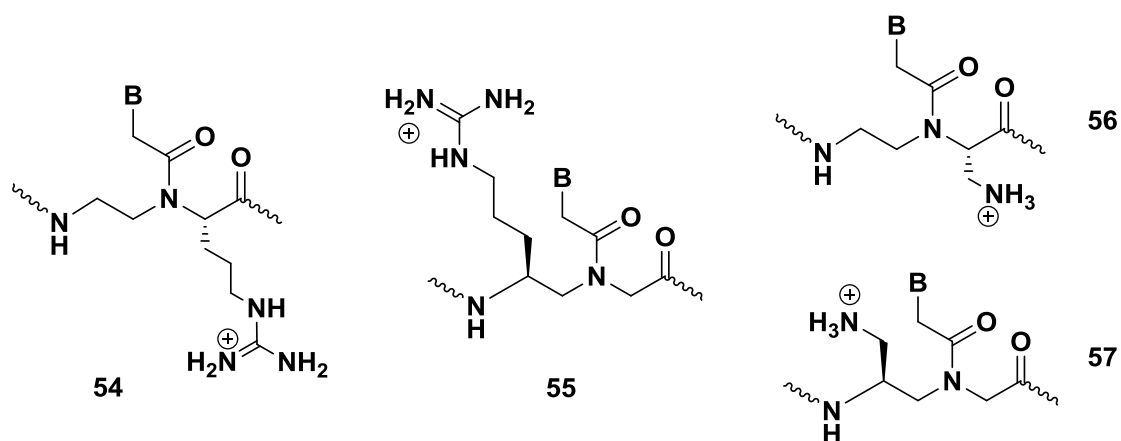


Figure 35. Cationic backbone modifications of PNA

In contrast to α -modified PNA, Englund and Appella found that the *S*-isomer of γ -modified PNA (derived from the natural L- lysine) had higher affinity for complementary DNA than the *R*-isomer.¹¹³ Most recently, Manicardi *et al.* used both α - and γ -modified GPNA 15-mers to inhibit microRNA-210 in K562 cells.¹⁶⁵ Both isomers showed promising though not complete inhibition with the PNAs having eight consecutive γ -modification at the carboxyl terminus performing slightly better than other modification patterns.

Mitra and Ganesh reported similar results on DNA binding and cellular uptake of α - and γ -aminomethylene PNA (Figure 35, **56** and **57** resp.). The amino methylene modification increased PNA binding to DNA, with γ -(*S*)*am*-PNA being significantly better than α -(*R*)*am*-PNA, which, in turn, was better than α -(*S*)*am*-PNA.¹⁶⁶ The cellular uptake was enhanced by these modifications in roughly the same order, with γ -(*S*)*am*-PNA giving the most promising results.

The above section discusses selected recent studies that improve on cellular uptake of PNA. The preliminary results are very encouraging, and it is likely that more improvements and new discoveries will be made in the near future.

1F The Present Work

The preceding sections have given an overview on the concept of peptide nucleic acids (PNAs) which is a successful mimic of DNA. The attractive binding properties of PNAs, both in terms of affinity and specificity, coupled with their strand invasion potential have promoted PNA as a useful tool in molecular biology, diagnostics, and as a possible candidate for antisense/ antigene drug therapy.

The major factors restricting the applications of PNA have been its poor water solubility, insufficient cell uptake, self-aggregation and ambiguity in the binding orientation. In order to overcome these limitations, several modifications of PNA have been carried out. PNAs have also been linked to helper molecules in various chimeras in an endeavour to improve their favourable properties. The work presented in this thesis involves the design, synthesis and biophysical evaluation of these backbone-modified, chiral, charged PNA analogues.

Chapter 2: In this Chapter we have described a novel modified analogue of PNA, ethano-locked PNA (ethano-PNA) and its biophysical studies with DNA and RNA. This Chapter is divided into two sections.

Section A: The rationale for the design of a novel ethano-locked PNA (ethano-PNA) is to envisage a conformational constraint on the relatively flexible PNA backbone for favorable RNA binding. The conformational restriction is introduced in the form of an ethylene bridge between aminoethyl linker of *aeg*PNA, hence, termed as ethano-PNA. The presence of cationic groups like amino or guanidino in the backbone of ethano-PNA may improve the cellular uptake of PNA. This Section deals with the syntheses of H-ethano-PNA and Amino-ethano-PNA monomer, their incorporation in the *aeg*PNA backbone. The simple procedure for conversion of Amino-ethano-PNA to Guanidino-ethano-PNA on solid support also has been described in this Section.

Section B: This section includes the study of the effect of introduction of constrained ethano-PNA monomer into *aeg*PNA backbone on binding with DNA/RNA. It also describes the effect of amino or guanidino groups on the binding efficacy of ethano-PNA with DNA and RNA along with the CD analysis of the duplex with DNA/RNA.

Chapter 3: This chapter discusses about the designing of cationic backbone modification of *aeg*PNA by incorporating bisubstitution at the β - and γ - positions in the

PNA monomer. The biophysical studies with DNA and RNA are also studied. This Chapter has been divided into two sections.

Section A: As an alternative to peptide conjugation, the PNAs could be designed to be intrinsically positively charged for better aqueous solubility and cellular uptake. Thus we introduce substitutions at the β - and γ - positions of PNA with aminomethyl group. This could lead to charge in the backbone. This Section deals with the synthesis of monomers having bisubstitutions and their incorporation in the aegPNA backbone via solid phase PNA synthesis.

Section B: In this section, the effect of introduction of bisubstituted PNA monomers into aegPNA backbone on binding with DNA/RNA is studied. It also analyses the effect of stereoisomeric monomers on the binding efficacy with DNA and RNA, along with the CD analysis of the duplex with DNA/RNA. The cellular uptake of the oligos in cells is also analysed.

Chapter 4: This chapter deals with polyamide-DNA with alternating α -amino acid and nucleoside- β -amino acids, a DNA analogue where phosphate backbone is replaced by amide linkage and the inter nucleic acid distance was substituted with prochiral glycine unit. The chapter discusses about the synthesis of all four natural nucleoside-based β -amino acids and their utility in the synthesis of (α -amino acid + nucleoside- β amino acid) sequence of dephosphono polyamide-DNA. The biophysical studies of these sequences with DNA and RNA and the CD studies were carried out for investigating the effect of this change in the backbone.

1G References

1. J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737-738.
2. E. Lescrinier, M. Froeyen and P. Herdewijn, *Nucleic Acids Research*, 2003, **31**, 2975-2989.
3. K. Hoogsteen, *Acta Crystallographica*, 1963, **16**, 907-+.
4. F. H. C. Crick, *J. Mol. Biol.*, 1966, **19**, 548-&.
5. N. C. Seeman, J. M. Rosenberg and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1976, **73**, 804-808.
6. A. Ghosh and M. Bansal, *Acta Crystallographica Section D-Biological Crystallography*, 2003, **59**, 620-626.
7. A. G. W. Leslie, S. Arnott, R. Chandrasekaran and R. L. Ratliff, *J. Mol. Biol.*, 1980, **143**, 49-72.
8. T. J. Richmond and C. A. Davey, *Nature*, 2003, **423**, 145-150.
9. M. C. Wahl and M. Sundaralingam, *Biopolymers*, 1997, **44**, 45-63.
10. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, G. Vandermarel, J. H. Vanboom and A. Rich, *Science*, 1981, **211**, 171-176.
11. J. F. Allemand, D. Bensimon, R. Lavery and V. Croquette, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 14152-14157.
12. G. H. Hayashi, M.; Nakatani, K., *Nucleic Acids Symp Ser. (Oxf)* 2005, **49**, 261-262.
13. J. M. Vargason, B. F. Eichman and P. S. Ho, *Nat. Struct. Biol.*, 2000, **7**, 758-761.
14. G. Wang and K. M. Vasquez, *Environ. Mol. Mutagen.*, 2006, **47**, 427-427.
15. J. M. Berg, J. L. Tymoczko and L. Stryer, *Biochemistry*, 5th Edition, WH Freeman and Company. 2002, 781-808.
16. A. Eschenmoser and M. Döbler, *Helv. Chim. Acta*, 1992, **75**, 218-259.
17. G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc.*, 1957, **79**, 2023-2024.
18. H. E. Moser and P. B. Dervan, *Science*, 1987, **238**, 645-650.
19. N. T. Thuong and C. Helene, *Angew. Chem.-Int. Edit.*, 1993, **32**, 666-690.
20. P. A. Beal and P. B. Dervan, *Science*, 1991, **251**, 1360-1363.
21. K. Hoogsteen, *Acta Crystallographica*, 1959, **12**, 822-823.
22. V. N. Soyfer and V. N. Potaman, Eds 1996, Springer-Verlag, New York.
23. V. L. Makarov, Y. Hirose and J. P. Langmore, *Cell*, 1997, **88**, 657-666.

24. W. I. Sundquist and A. Klug, *Nature*, 1989, **342**, 825-829.
25. K. Gehring, J. L. Leroy and M. Gueron, *Nature*, 1993, **363**, 561-565.
26. G. M. Blackburn, M. J. Gait, D. Loakes and D. M. Williams, *Nucleic acids in chemistry and biology*, RSC publishing, 3rd edition.
27. C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry part III*, 1971, W. H. Freeman and Company, New York.
28. G. E. Plum, Y. W. Park, S. F. Singleton, P. B. Dervan and K. J. Breslauer, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 9436-9440.
29. C. Y. Huang, *Methods Enzymol.*, 1982, **87**, 509-525.
30. K. E. Vanholde, J. Brahm and Michelso. Am, *J. Mol. Biol.*, 1965, **12**, 726-739.
31. C. R. Calladine, H. R. Drew and Cambridge:Academic Press LTd., 1992.
32. D. Smith, V. Schuller, C. Engst, J. Radler and T. Liedl, *Nanomedicine*, **8**, 105-121.
33. R. W. Wagner, *Nat. Med.*, 1995, **1**, 1116-1118.
34. P. P. Chan and P. M. Glazer, *J. Mol. Med.*, 1997, **75**, 267-282.
35. P. C. Zamecnik and M. L. Stephenson, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 280-284.
36. W. Nellen and C. Lichtenstein, *Trends Biochem.Sci.*, 1993, **18**, 419-423.
37. A. Fire, S. Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello, *Nature*, 1998, **391**, 806-811.
38. J. Q. Yin, J. Gao, R. Shao, W. N. Tian, J. Wang and Y. Wan, *J. Exp. Ther. Oncol.*, 2003, **3**, 194-204.
39. Y. Dorsett and T. Tuschl, *Nat. Rev. Drug Discov.*, 2004, **3**, 318-329.
40. R. De Francesco and G. Migliaccio, *Nature*, 2005, **436**, 953-960.
41. R. C. Lee, R. L. Feinbaum and V. Ambros, *Cell*, 1993, **75**, 843-854.
42. V. N. Kim, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 376-385.
43. M. Garofalo and C. M. Croce, in *Annual Review of Pharmacology and Toxicology, Vol 51, 2011*, Annual Reviews, Palo Alto, pp. 25-43.
44. *Cell*, 2006, **126**, 223-225.
45. Z. Dominski and R. Kole, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 8673-8677.
46. P. Sazani and R. Kole, *J. Clin. Invest.*, 2003, **112**, 481-486.
47. W. Gilbert, *Nature*, 1978, **271**, 501-501.
48. B. L. Davidson and P. B. McCray, *Nat. Rev. Genet.*, 2011, **12**, 329-340.

49. C. A. Stein and J. S. Cohen, *Phosphorothioate oligodeoxynucleotide analogues*, In Cohen J. S. (ed.): *Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression*. London: Macmillan Press, 1989, 97.
50. P. S. Millar, *Non-ionic antisense oligonucleotides*, In Cohen, J. S. (ed.): *Oligodeoxynucleotides-antisense inhibitors of gene expression*. London: Macmillan Press, 1989, 79.
51. B. Froehler, P. Ng and M. Matteucci, *Nucleic Acids Res.*, 1988, **16**, 4831-4839.
52. M. F. Summers, C. Powell, W. Egan, R. A. Byrd, W. D. Wilson and G. Zon, *Nucleic Acids Res.*, 1986, **14**, 7421-7435.
53. B. R. Shaw, J. Madison, S. sood and B. F. Spielvogel, In Agrawal, S. (ed): *Methods in Molecular biology, vol 20: Protocols for Oligonucleotides and Analogs. Synthesis and properties*. Totowa, NJ. Humana Press, Inc, 1993, 225-243.
54. S. Agrawal and Q. Y. Zhao, *Curr. Opin. Chem. Biol.*, 1998, **2**, 519-528.
55. A. Demesmaeker, A. Waldner, J. Lebreton, P. Hoffmann, V. Fritsch, R. M. Wolf and S. M. Freier, *Angew. Chem.-Int. Edit. Engl.*, 1994, **33**, 226-229.
56. M. Matteucci, *Tetrahedron Lett.*, 1990, **31**, 2385-2388.
57. R. J. Jones, K. Y. Lin, J. F. Milligan, S. Wadwani and M. D. Matteucci, *J. Org. Chem.*, 1993, **58**, 2983-2991.
58. J. Lebreton, A. Waldner, V. Fritsch, R. R. Wolf and A. Demesmaeker, *Tetrahedron Lett.*, 1994, **35**, 5225-5228.
59. B. P. Monia, E. A. Lesnik, C. Gonzalez, W. F. Lima, D. McGee, C. J. Guinasso, A. M. Kawasaki, P. D. Cook and S. M. Freier, *J. Biol. Chem.*, 1993, **268**, 14514-14522.
60. F. J. Raal, R. D. Santos, D. J. Blom, A. D. Marais, M. J. Charng, W. C. Cromwell, R. H. Lachmann, D. Gaudet, J. L. Tan, S. Chasan-Taber, D. L. Tribble, J. D. Flaim and S. T. Crooke, *Lancet*, 2010, **375**, 998-1006.
61. J. Summerton and D. Weller, *Antisense Nucleic Acid Drug Dev.*, 1997, **7**, 187-195.
62. J. Summerton and D. Weller, *Nucleosides Nucleotides*, 1997, **16**, 889-898.
63. C. Hendrix, H. Rosemeyer, I. Verheggen, F. Seela, A. VanAerschot and P. Herdewijn, *Chem.-Eur. J.*, 1997, **3**, 110-120.
64. M. Petersen and J. Wengel, *Trends Biotechnol.*, 2003, **21**, 74-81.

65. K. U. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347-1351.
66. L. L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174-4175.
67. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1500.
68. M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566-568.
69. K. K. Jensen, H. Orum, P. E. Nielsen and B. Norden, *Biochemistry*, 1997, **36**, 5072-5077.
70. V. V. Demidov, V. N. Potaman, M. D. Frankkamenetskii, M. Egholm, O. Buchard, S. H. Sonnichsen and P. E. Nielsen, *Biochem. Pharmacol.*, 1994, **48**, 1310-1313.
71. L. Good and P. E. Nielsen, *Nat. Biotechnol.*, 1998, **16**, 355-358.
72. A. Ray and B. Norden, *Faseb J.*, 2000, **14**, 1041-1060.
73. P. E. Nielsen, *Curr. Opin. Biotechnol.*, 1999, **10**, 71-75.
74. B. Metaferia, J. S. Wei, Y. K. Song, J. Evangelista, K. Aschenbach, P. Johansson, X. Y. Wen, Q. R. Chen, A. Lee, H. Hempel, J. S. Gheeya, S. Getty, R. Gomez and J. Khan, *PLoS One*, 2013, **8**, 7.
75. C. M. Micklitsch, B. Y. Oquare, C. Zhao and D. H. Appella, *Anal. Chem.*, 2013, **85**, 251-257.
76. E. Uhlmann, D. W. Will, G. Breipohl, D. Langner and A. Rytte, *Angew. Chem.-Int. Edit. Engl.*, 1996, **35**, 2632-2635.
77. S. Tomac, M. Sarkar, T. Ratilainen, P. Wittung, P. E. Nielsen, B. Norden and A. Graslund, *J. Am. Chem. Soc.*, 1996, **118**, 5544-5552.
78. P. E. Nielsen, M. Egholm and O. Buchardt, *J Mol. Recogn.*, 1994, **7**, 165-170.
79. H. Knudsen and P. E. Nielsen, *Nucleic Acids Res.*, 1996, **24**, 494-500.
80. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Anti-Cancer Drug Des.*, 1993, **8**, 53-63.
81. J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble and L. E. Babiss, *Science*, 1992, **258**, 1481-1485.

82. T. C. Marsh and E. Henderson, *Biochemistry*, 1994, **33**, 10718-10724.
83. S. L. Forman, J. C. Fettinger, S. Pieraccini, G. Gottareli and J. T. Davis, *J. Am. Chem. Soc.*, 2000, **122**, 4060-4067.
84. R. T. West, L. A. Garza, W. R. Winchester and J. A. Walmsley, *Nucleic Acids Res.*, 1994, **22**, 5128-5134.
85. B. Datta, C. Schmitt and B. A. Armitage, *J. Am. Chem. Soc.*, 2003, **125**, 4111-4118.
86. B. Datta, M. E. Bier, S. Roy and B. A. Armitage, *J. Am. Chem. Soc.*, 2005, **127**, 4199-4207.
87. M. Eriksson and P. E. Nielsen, *Nat. Struct. Biol.*, 1996, **3**, 410-413.
88. S. C. Brown, S. A. Thomson, J. M. Veal and D. G. Davis, *Science*, 1994, **265**, 777-780.
89. L. Betts, J. A. Josey, J. M. Veal and S. R. Jordan, *Science*, 1995, **270**, 1838-1841.
90. H. Rasmussen and J. Sandholm, *Nat. Struct. Biol.*, 1997, **4**, 98-101.
91. L. Mologni, P. leCoutre, P. E. Nielsen and C. Gambacorti-Passerini, *Nucleic Acids Res.*, 1998, **26**, 1934-1938.
92. Shan Goh, Jem Stach and L. Good, *Peter E. Nielsen and Daniel H. Appella (eds.), Peptide Nucleic Acids: Methods and Protocols, Methods in Molecular Biology*, 2014, **1050**, 223-236.
93. P. E. Nielsen, M. Egholm and O. Buchardt, *Gene*, 1994, **149**, 139-145.
94. M. Egholm, L. Christensen, K. L. Dueholm, O. Buchardt, J. Coull and P. E. Nielsen, *Nucleic Acids Res.*, 1995, **23**, 217-222.
95. E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem.-Int. Edit.*, 1998, **37**, 2797-2823.
96. J. C. Norton, M. A. Piatyszek, W. E. Wright, J. W. Shay and D. R. Corey, *Nat. Biotechnol.*, 1996, **14**, 615-619.
97. H. Orum, P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt and C. Stanley, *Nucleic Acids Res.*, 1993, **21**, 5332-5336.
98. D. B. Demers, E. T. Curry, M. Egholm and A. C. Sozer, *Nucleic Acids Res.*, 1995, **23**, 3050-3055.
99. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Nucleic Acids Res.*, 1993, **21**, 197-200.

100. V. Demidov, M. D. Frankkamenetskii, M. Egholm, O. Buchardt and P. E. Nielsen, *Nucleic Acids Res.*, 1993, **21**, 2103-2107.
101. K. Ito and M. Komiyama, *Peter E. Nielsen and Daniel H. Appella (eds.), Peptide Nucleic Acids: Methods and Protocols, Methods in Molecular Biology*, 2014, **1050**, 111-120.
102. K. N. Ganesh and P. E. Nielsen, *Curr. Org. Chem.*, 2000, **4**, 931-943.
103. V. A. Kumar, *Eur. J. Org. Chem.*, 2002, 2021-2032.
104. V. A. Kumar and K. N. Ganesh, *Accounts Chem. Res.*, 2005, **38**, 404-412.
105. V. A. Kumar and K. N. Ganesh, *Curr. Top. Med. Chem.*, 2007, **7**, 715-726.
106. B. Hyrup, M. Egholm, P. E. Nielsen, P. Wittung, B. Norden and O. Buchardt, *J. Am. Chem. Soc.*, 1994, **116**, 7964-7970.
107. B. Hyrup, M. Egholm, O. Buchardt and P. E. Nielsen, *Bioorg. Med. Chem. Lett.*, 1996, **6**, 1083-1088.
108. K. L. Dueholm, K. H. Petersen, D. K. Jensen, M. Egholm, P. E. Nielsen and O. Buchardt, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 1077-1080.
109. G. Haaima, A. Lohse, O. Buchardt and P. E. Nielsen, *Angew. Chem.-Int. Edit.*, 1996, **35**, 1939-1942.
110. A. Puschl, S. Sforza, G. Haaima, O. Dahl and P. E. Nielsen, *Tetrahedron Lett.*, 1998, **39**, 4707-4710.
111. T. Sugiyama, Y. Imamura, Y. Demizu, M. Kurihara, M. Takano and A. Kittaka, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 7317-7320.
112. A. Calabretta, T. Tedeschi, R. Corradini, R. Marchelli and S. Sforza, *Tetrahedron Lett.*, 2011, **52**, 300-304.
113. E. A. Englund and D. H. Appella, *Angew. Chem.-Int. Edit.*, 2007, **46**, 1414-1418.
114. S. Sforza, R. Corradini, S. Ghirardi, A. Dossena and R. Marchelli, *Eur. J. Org. Chem.*, 2000, 2905-2913.
115. T. Govindaraju, V. Madhuri, V. A. Kumar and K. N. Ganesh, *J. Org. Chem.*, 2006, **71**, 14-21.
116. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *J. Org. Chem.*, 2004, **69**, 5725-5734.
117. W. Maison, I. Schlemminger, O. Westerhoff and J. Martens, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 581-584.

118. A. H. Krotz, S. Larsen, O. Buchardt, M. Eriksson and P. E. Nielsen, *Bioorg. Med. Chem.*, 1998, **6**, 1983-1992.
119. C. De Cola, A. Manicardi, R. Corradini, I. Izzo and F. De Riccardis, *Tetrahedron*, 2012, **68**, 499-506.
120. B. P. Gangamani, V. A. Kumar and K. N. Ganesh, *Tetrahedron*, 1999, **55**, 177-192.
121. M. D'Costa, V. A. Kumar and K. N. Ganesh, *J. Org. Chem.*, 2003, **68**, 4439-4445.
122. N. K. Sharma and K. N. Ganesh, *Tetrahedron Lett.*, 2004, **45**, 1403-1406.
123. A. I. Khan, T. H. S. Tan and J. Micklefield, *Chem. Commun.*, 2006, 1436-1438.
124. D. T. Hickman, T. H. S. Tan, J. Morral, P. M. King, M. A. Cooper and J. Micklefield, *Org. Biomol. Chem.*, 2003, **1**, 3277-3292.
125. V. Kumar, P. S. Pallan, Meena and K. N. Ganesh, *Org. Lett.*, 2001, **3**, 1269-1272.
126. T. Govindaraju and V. A. Kumar, *Chem. Commun.*, 2005, 495-497.
127. R. J. Worthington, N. M. Bell, R. Wong and J. Micklefield, *Org. Biomol. Chem.*, 2008, **6**, 92-103.
128. M. D'Costa, V. Kumar and K. N. Ganesh, *Tetrahedron Lett.*, 2002, **43**, 883-886.
129. P. S. Lonkar, K. N. Ganesh and V. A. Kumar, *Org. Biomol. Chem.*, 2004, **2**, 2604-2611.
130. C. Vilaivan, C. Srisuwannaket, C. Ananthanawat, C. Suparpprom, K. J., Y. Yamaguchi, Y. Tanaka and T. Vilaivan, *Artificial DNA: PNA & XNA*, 2011, **2**, 50-59.
131. S. Bagmare, M. D'Costa and V. A. Kumar, *Chem. Commun.*, 2009, 6646-6648.
132. A. Slaitas and E. Yeheskiely, *Eur. J. Org. Chem.*, 2002, 2391-2399.
133. S. S. Gokhale and V. A. Kumar, *Org. Biomol. Chem.*, 2010, **8**, 3742-3750.
134. P. S. Shirude, V. A. Kumar and K. N. Ganesh, *Tetrahedron Lett.*, 2004, **45**, 3085-3088.
135. A. Puschl, T. Boesen, T. Tedeschi, O. Dahl and P. E. Nielsen, *J. Chem. Soc.-Perkin Trans. 1*, 2001, 2757-2763.
136. G. Haaima, H. F. Hansen, L. Christensen, O. Dahl and P. E. Nielsen, *Nucleic Acids Res.*, 1997, **25**, 4639-4643.
137. B. P. Gangamani, V. A. Kumar and K. N. Ganesh, *Chem. Commun.*, 1997, 1913-1914.

138. A. B. Eldrup, O. Dahl and P. E. Nielsen, *J. Am. Chem. Soc.*, 1997, **119**, 11116-11117.
139. F. Wojciechowski and R. H. E. Hudson, *Curr. Top. Med. Chem.*, 2007, **7**, 667-679.
140. C. Ausin, J. A. Ortega, J. Robles, A. Grandas and E. Pedroso, *Org. Lett.*, 2002, **4**, 4073-4075.
141. H. Challa, M. L. Styers and S. A. Woski, *Org. Lett.*, 1999, **1**, 1639-1641.
142. R. Vysabhattachar and K. N. Ganesh, *Tetrahedron Lett.*, 2008, **49**, 1314-1318.
143. P. E. Nielsen, *Q. Rev. Biophys.*, 2005, **38**, 345-350.
144. T. Shiraishi and P. E. Nielsen, *Nat. Protoc.*, 2006, **1**, 633-636.
145. F. S. Hassane, A. F. Saleh, R. Abes, M. J. Gait and B. Lebleu, *Cell. Mol. Life Sci.*, 2010, **67**, 715-726.
146. P. J. Finn, N. J. Gibson, R. Fallon, A. Hamilton and T. Brown, *Nucleic Acids Res.*, 1996, **24**, 3357-3363.
147. F. Bergmann, W. Bannwarth and S. Tam, *Tetrahedron Lett.*, 1995, **36**, 6823-6826.
148. D. A. Braasch and D. R. Corey, *Methods in Molecular Biology*, 2002, **208**, 211-223.
149. H. Fang, K. Zhang, G. Shen, K. L. Woolley and J. S. A. Taylor, *Mol. Pharm.*, 2009, **6**, 615-626.
150. Z. Wang, K. Zhang, K. L. Wooley and J. S. Taylor, *Journal of Nucleic Acids*, 2012, **12**, Article ID 962652.
151. S. Abes, J. J. Turner, G. D. Ivanova, D. Owen, D. Williams, A. Arzumanov, P. Clair, M. J. Gait and B. Lebleu, *Nucleic Acids Res.*, 2007, **35**, 4495-4502.
152. J. X. Hu, M. Matsui, K. T. Gagnon, J. C. Schwartz, S. Gabillet, K. Arar, J. Wu, I. Bezprozvanny and D. R. Corey, *Nat. Biotechnol.*, 2009, **27**, 478-484.
153. M. M. Fabani and M. J. Gait, *RNA-Publ. RNA Soc.*, 2008, **14**, 336-346.
154. A. G. Torres, M. M. Fabani, E. Vigorito, D. Williams, N. Al-Obaidi, F. Wojciechowski, R. H. E. Hudson, O. Seitz and M. J. Gait, *Nucleic Acids Res.*, 2012, **40**, 2152-2167.
155. J. X. Hu and D. R. Corey, *Biochemistry*, 2007, **46**, 7581-7589.
156. E. Fabbri, A. Manicardi, T. Tedeschi, S. Sforza, N. Bianchi, E. Brognara, A. Finotti, G. Breveglieri, M. Borgatti, R. Corradini, R. Marchelli and R. Gambari, *ChemMedChem*, 2011, **6**, 2192-2202.

157. U. Koppelhus, T. Shiraishi, V. Zachar, S. Pankratova and P. E. Nielsen, *Bioconjugate Chem.*, 2008, **19**, 1526-1534.
158. P. R. Berthold, T. Shiraishi and P. E. Nielsen, *Bioconjugate Chem.*, 2010, **21**, 1933-1938.
159. G. Upert, A. D. Giorgio, A. Upadhyay, D. Manvar, N. Pandey, V. N. Pandey and N. Patino, *Journal of Nucleic Acids*, 2012, **12**, Article ID 591025.
160. T. Shiraishi and P. E. Nielsen, *Bioconjugate Chem.*, 2012, **23**, 196-202.
161. P. Zhou, M. M. Wang, L. Du, G. W. Fisher, A. Waggoner and D. H. Ly, *J. Am. Chem. Soc.*, 2003, **125**, 6878-6879.
162. B. Sahu, V. Chenna, K. L. Lathrop, S. M. Thomas, G. Zon, K. J. Livak and D. H. Ly, *J. Org. Chem.*, 2009, **74**, 1509-1516.
163. P. Zhou, A. Dragulescu-Andrasi, B. Bhattacharya, H. O'Keefe, P. Vatta, J. J. Hyldig-Nielsen and D. H. Ly, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4931-4935.
164. A. Dragulescu-Andrasi, P. Zhou, G. F. He and D. H. Ly, *Chem. Commun.*, 2005, 244-246.
165. A. Manicardi, E. Fabbri, T. Tedeschi, S. Sforza, N. Bianchi, E. Brognara, R. Gambari, R. Marchelli and R. Corradini, *ChemBioChem*, 2012, **13**, 1327-1337.
166. R. Mitra and K. N. Ganesh, *J. Org. Chem.*, 2012, **77**, 5696-5704.

CHAPTER 2

Novel ethano locked PNA (ethano-PNA)

Section A

Synthesis of monomer and incorporation in aegPNA sequences

Section B

Biophysical studies of ethano-PNA towards comp. DNA/RNA

Section A

Synthesis of monomer and incorporation in *aeg*PNA sequences

2A Introduction

Nucleic acid mimics are currently gaining considerable importance due to the interest in their development in the oligonucleotide (ON) based therapeutics^{1, 2} and diagnostic strategies.³⁻⁶ This application of synthetic nucleic acids is based on the sequence specific recognition of target RNA or DNA sequences by Watson - Crick hydrogen bonding schemes. The effectiveness of the designed ON mimics as lead molecules for such applications primarily depends on their specific hybridization ability with the target sequences. The desired additional attributes of the synthetic ON mimics are aqueous solubility, enzymatic stability and also the ease of synthesis. To achieve these goals, several backbone modifications are being developed that include the 2nd and 3rd generation antisense ONs such as 2'-*O*-alkyl,⁷ 2'-*O*-MTE,⁸ LNA,⁹ HNA,¹⁰ morpholino-NA,¹¹⁻¹³ aminoethylglycyl PNA (*aeg*PNA),¹⁴ or several other PNA modifications reported so far.^{15, 16}

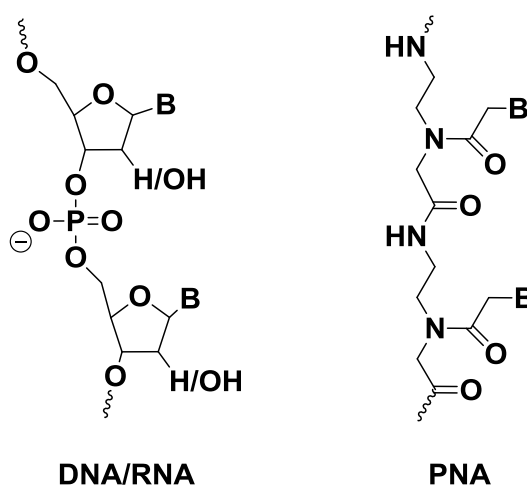


Figure 1. Basic structure of DNA/RNA and PNA

Peptide nucleic acids (*aeg*PNA, Figure 1) are designed oligonucleotide mimics, derived by replacement of the anionic sugar–phosphate backbone with repeating *N*-(2-aminoethyl) glycol units, carrying adenine, guanine, cytosine, and thymine nucleobases via methylene carbonyl linkages.¹⁷ PNA acts as an excellent structural mimic of natural DNA/RNA and exhibits strong sequence specific binding with complementary oligonucleotide sequences following Watson - Crick base-pairing rules.¹⁸ The relatively high binding affinity of PNAs towards the natural oligonucleotides is attributed to the lack of electrostatic repulsions between the uncharged PNA backbones and negatively

charged sugar - phosphate backbone of DNA and RNA. Because of highly favorable hybridization properties and high chemical and bio-stability, PNA is regarded as a very promising lead for developing into efficient antisense agents^{1, 2} and diagnostic markers.³⁻⁶ The positive attributes of PNA are accompanied by some disadvantages which hamper the *in vivo* applications of PNA. The main drawbacks of PNA are its relatively poor water solubility¹⁹ and poor cellular uptake²⁰ due to the uncharged polyamide backbone. The acyclic flexible PNA backbone also binds to the complementary DNA and RNA sequences without any discrimination to form structurally diverse PNA:DNA and PNA:RNA duplexes. These limitations can be systematically overcome with rationally modified PNA analogues.

2A.1 Rationale and objectives of the present work

The NMR²¹ and X-ray crystal structural studies²² pointed out some crucial differences in PNA:DNA and PNA:RNA duplexes. The main difference was in the preferred dihedral angle β in the ethylene diamine segment ($N1'-C2'-C3'-N4'$, Figure 2) of PNA units while binding with RNA and DNA. The preferred dihedral angle β was found to be $60-70^\circ$ in a PNA:RNA duplex and the same was $\sim 140^\circ$ in PNA:DNA duplex. In the literature, a variety of conformationally constrained PNA analogues are known,^{15, 16} which show differential stability and parallel/antiparallel preferences while binding with cDNA and RNA.

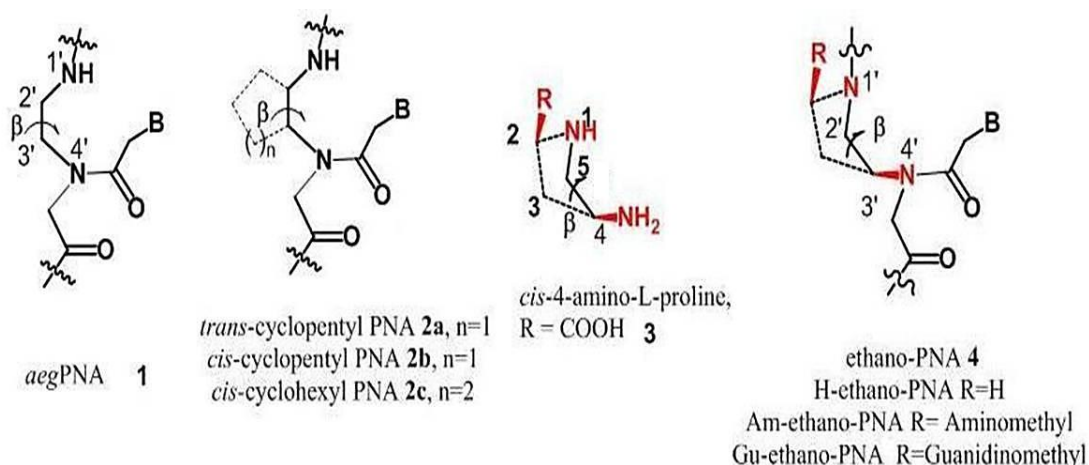


Figure 2. *aeg*PNA, related PNA modifications and the designed ethano-PNA analogues

The designed RNA selective *cis*-cyclopentyl^{23, 24} (Figure 2, **2b**) and *cis*-cyclohexyl²⁵⁻²⁸ (Figure 2, **2c**) based PNAs have restricted the dihedral angle β to $60-70^\circ$. The *trans*-cyclohexyl²⁹ and *trans*-cyclopentyl^{30, 31} (Figure 2, **2a**) based PNAs were constructed by Nielsen and Appella's groups, respectively. The rigid nature of cyclohexyl ring was found to be highly successful in stabilizing the duplexes with RNA

in a stereoselective manner when the geometry of substitution was *cis* and the dihedral angle β was restricted to the preferred value ($\sim 60^\circ$) in RNA:PNA duplexes. The *cis* and *trans*-cyclopentyl based PNAs, however, were more flexible and were able to access the range of dihedral angles that accommodated both the PNA:DNA and PNA:RNA structures.

In this context, we propose to use 3-aminopyrrolidine core structure towards the rational design of conformationally restricted *aeg*PNA. The pyrrolidine ring conformations in substituted prolines are known to be dictated by the R/S stereochemistry of the electronegative substitutions at C-4 center^{32, 33} (Figure 2). The gauche interactions of the substituent amino group and the vicinal ring nitrogen allow either C4-*endo* or C4-*exo* conformations, depending upon the R/S stereochemistry at C4 center, restricting the N1-C5-C4-N4 dihedral angle in the pyrrolidine ring around 60-80° (Figure 2). This preference has been used by several researchers to achieve excellent breakthroughs in designing collagen mimics.³⁴⁻³⁶ In the present PNA design, this N1-C5-C4-N4 segment of the pyrrolidine ring coincides with the aminoethyl segment (i.e., N1'-C2'-C3'-N4', Figure 2) of *aeg*PNA and the gauche geometry of the vicinal electronegative substituents would coincide with the preferred dihedral angle β in *aeg*PNA:RNA duplexes.

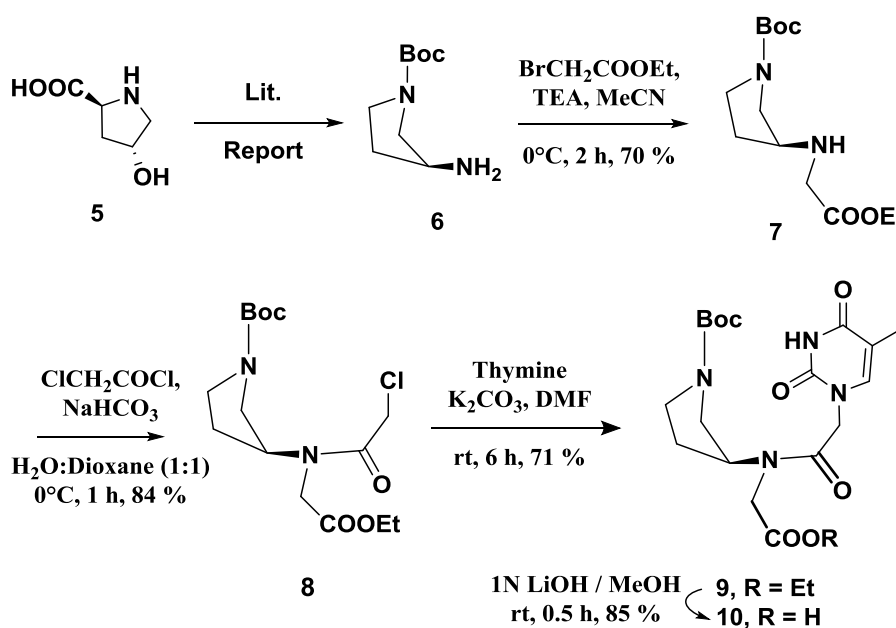
Thus, we present here, the design, synthesis and biophysical studies of a novel ethano-locked PNA (ethano-PNA)³⁷ that is envisaged to confer conformational constraint on the relatively flexible PNA backbone for favorable RNA binding. The conformational restriction is introduced in the form of an ethylene bridge between aminoethyl linker of *aeg*PNA, hence, termed as ethano-PNA. It was envisaged that the easy synthetic methodology developed from naturally occurring *trans*-4-hydroxy-L-proline would be useful compared to the cumbersome resolution steps involved in the synthesis of enantiomerically pure cyclopentyl/cyclohexyl²³⁻²⁸ based PNA analogues. The ethano group is further amenable for functionalization with aminomethyl (Am-ethano-PNA, Figure 2) and guanidinomethyl (Gu-ethano-PNA, Figure 2) group for better aqueous solubility.³⁸ We envisage to confer the modified oligomers with an ability to preferentially bind with RNA.

2A.2 Synthesis of ethano locked monomers

2A.2.1 Synthesis of Boc protected *H*-ethano-PNA T monomer

The synthesis of *H*-ethano-PNA monomer was accomplished as follows, starting from naturally occurring (2*S*, 4*R*) hydroxyproline **5**. The *N*-Boc-3-*S*-aminopyrrolidine **6**

was obtained by decarboxylation using reported procedure,³⁹ followed by conversion of the hydroxy to amino group.⁴⁰ The primary amine was monoalkylated with ethyl bromoacetate in acetonitrile to get **7**. Further, **7** was acylated with chloroacetyl chloride in 1,4-dioxane-water (1:1) maintaining a pH ~8.0 with NaHCO₃ solution during addition of chloroacetyl chloride to obtain **8**. Thymine was derivatized via nucleophilic substitution of chloro group by *N*1 nitrogen of thymine to get the base attached product **9**. The ester was hydrolysed using 1 N LiOH in MeOH to give (*S*) *H*-ethano-PNA thymine monomer **10** (Scheme 1).

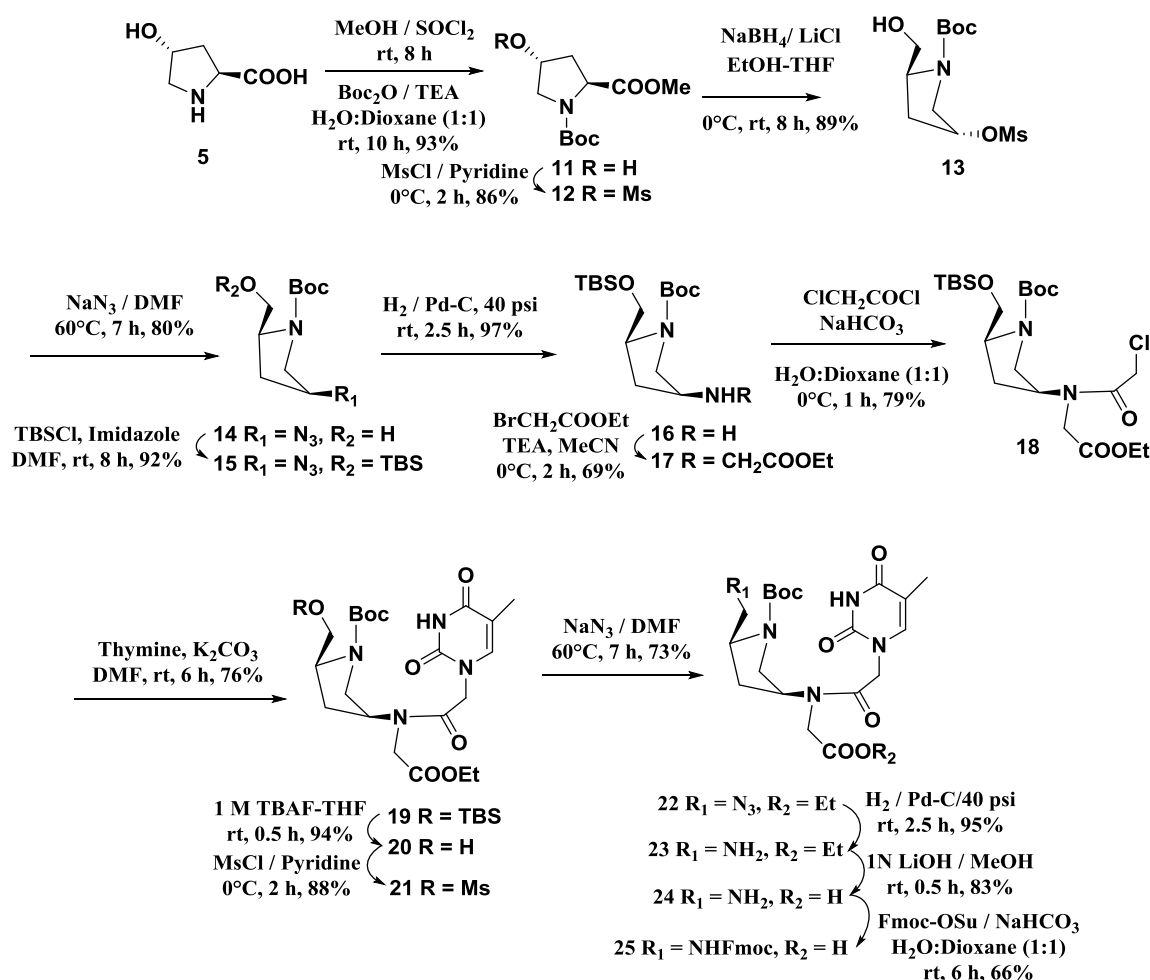


Scheme 1. Synthesis of H-ethano-PNA T monomer

2A.2.2 Synthesis of Fmoc/Boc protected *Am*-ethano-PNA T monomer

The synthesis of *Am*-ethano-PNA monomer was also achieved from (*2S*, *4R*) hydroxyproline. (*2S*, *4R*) hydroxyproline **5** was converted to methyl ester by treatment with methanol and thionyl chloride, which on subsequent protection with Boc anhydride gave *N*-Boc hydroxyproline methyl ester derivative **11**. The hydroxyl group was mesylated using mesyl chloride in pyridine **12** and reduced to **13** with LiBH₄ generated in situ with NaBH₄/LiCl. The mesyl derivative was converted to azide **14** by NaN₃ in DMF and the free hydroxyl was protected as TBS ether by reaction with TBSCl and imidazole in DMF to get **15**. The azido group in **15** was reduced to amine with H₂/Pd-C in methanol. The resultant amine **16** was monoalkylated with ethyl bromoacetate in acetonitrile, to give compound **17**. Acylation of **17** with chloroacetyl chloride in 1,4-dioxane-water (1:1) was carried out, maintaining the pH of the solution at ~8.0 with saturated NaHCO₃ solution during the addition of chloroacetyl chloride, to obtain **18**. Thymine was derivatized *via* nucleophilic substitution of chloro group by *N*1

nitrogen of thymine gave **19**. Further deprotection of TBS group of **19** gave **20** and eventually the hydroxyl group generated was converted to the mesylate derivative **21**. The mesyl group of **21** was displaced by azide **22**, which was subsequently reduced to amine by $\text{H}_2/\text{Pd-C}$ in methanol to obtain **23**. The compound **23** was used without purification for hydrolysis of ester to give the amino acid **24**. The exocyclic amine was protected with orthogonal Fmoc protecting group to give the protected Am-ethano-PNA thymine monomer **25** (Scheme 2). The exocyclic amine in monomer **25** served to provide the amino functionality in PNA and could also be conveniently converted to the highly basic guanidino functionality post synthetically on solid support to get Gu-ethano-PNA sequences.



Scheme 2. Synthesis of substituted aminomethyl-ethano-PNA T monomer

2A.3 Solid Phase PNA Synthesis

General Protocol: Solid phase peptide synthesis protocols can be easily applied to the synthesis of PNAs. The ease of handling and scale-up procedures have made possible the synthesis of PNAs including those incorporating a large number of analogues in an endeavour to improve its favourable binding and biological properties.

As is the case with solid phase peptide synthesis, PNA synthesis is also done conveniently from the 'C' terminus to the 'N' terminus. For this, the monomeric units must have their amino functions suitably protected, and their carboxylic acid functions free. The most commonly used *N*-protecting groups for solid phase peptide synthesis are the *t*-butyloxycarbonyl (Boc) and the 9-fluorenylmethoxycarbonyl (Fmoc) groups. The Fmoc protection strategy has a drawback in PNA synthesis since a small amount of acyl migration has been observed under basic conditions from the tertiary amide to the free amine formed during piperidine deprotection.⁴¹ Hence, the Boc-protection strategy was selected for the present work. The pyrrolidinyl ring nitrogen in all the ethano-PNA were protected with Boc whereas in Am-ethano-PNA monomers, the exocyclic amino function was orthogonally protected with Fmoc and the carboxylic acid function was free to enable coupling with the resin linked monomer. The TBTU/ HOBt activation strategy⁴² was employed for the coupling reaction. MBHA resin (4-Methylbenzhydrylamine resin)⁴³ was selected as solid polymeric matrix on which oligomers were synthesized. The first amino acid is linked to this matrix via benzyl amide linkage. This can be cleaved using standard protocol employing TFMSA-TFA to yield the C-terminal amide.

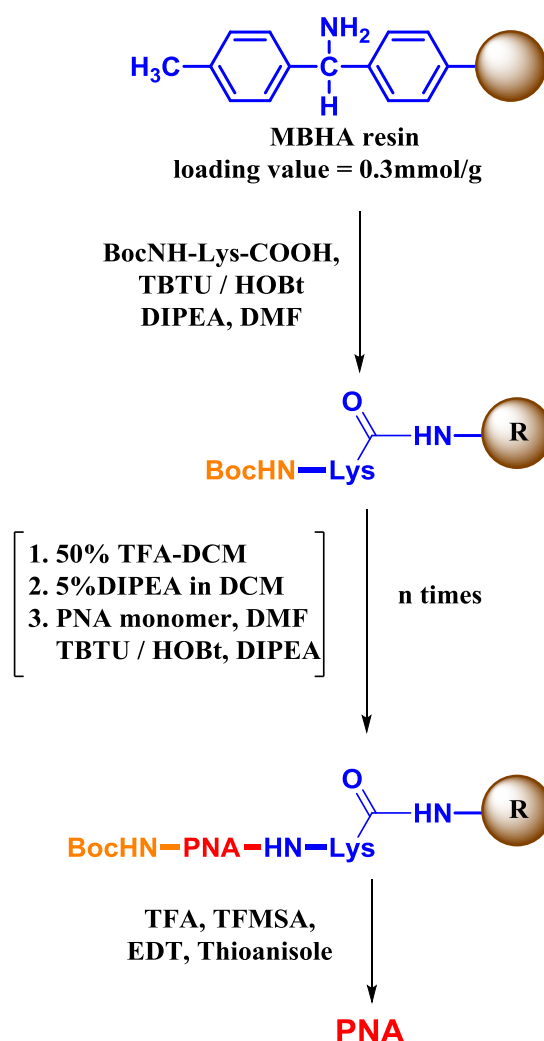
All the oligomers of the present work were synthesized manually on MBHA resin. L-lysine was selected as linker amino acid. Lysine at 'C' terminus is known to help in improving aqueous solubility of PNA. *N* (α)-Boc-*N* (ϵ)-2-Cl-Cbz-*L*-Lysine was linked to the resin through an amide bond. The free amine content on the resin was determined by the picrate assay and was found to be 1.75 mmol/g and loading was suitably lowered to approximately 0.30 mmol/g by partial acetylation of amine content using calculated amount of acetic anhydride. Free -NH₂ groups on the resin available for coupling are again estimated before starting synthesis.

The PNA oligomers were synthesized using repetitive cycles (a typical synthesis cycle is depicted in Scheme 3), each comprising the following steps:

- I. Deprotection of *N*-protecting Boc group using 50% TFA in CH₂Cl₂
- II. Neutralization of the TFA salt formed with diisopropylethyl amine (DIPEA) (5% DIPEA in CH₂Cl₂) to liberate the free amine
- III. Coupling of the free amine with the free carboxylic acid group of the incoming monomer (3-4 equivalents). The coupling reaction was carried out in presence of TBTU and HOBt in DMF or NMP as solvent and DIPEA as base. The deprotection of *N*-Boc amine and subsequent coupling reactions were monitored

by the Kaiser test.⁴⁴ The Boc-deprotection step gives a positive Kaiser's test, wherein the resin beads as well as the solution are blue in colour. On the other hand, upon completion of the coupling reaction, the Kaiser's test is negative, the resin beads remaining colourless.

- IV. Capping of the unreacted amino groups using acetic anhydride in pyridine and CH_2Cl_2



Scheme 3. Schematic representation of solid phase PNA synthesis using the Boc-protection strategy

2A.3.1 Solid Phase Synthesis of oligomers incorporating modified PNA monomers

PNA synthesis was undertaken from the C-terminal to N-terminal. PNA oligomers containing modified monomer were assembled by solid phase synthesis on MBHA resin functionalized with the first residue *L*-lysine, using TBTU / HOBt activation strategy. The sequences are shown in Table 1. The unmodified PNA and the modified ethano-PNA oligomers were designed to test the effect of ethano locked modification of PNA on the stability of the PNA:cDNA and PNA:RNA duplexes. The modified monomers were incorporated in pre-defined positions by solid phase synthesis

to yield a backbone modification at the central position (Table 1, entries 2–4) and near the N-terminal of the aegPNA (Table 1, entries 5–7).

2A.3.2 Post synthetic conversion of amino to guanidino group on solid phase

Guanidino functionality is known to be highly basic, due to its high pKa value (pKa= 12.5) as compared to amino group. It introduces a positive charge that is maintained over a wide pH range. Polyarginine and other guanidine rich peptides exhibit good cell penetrating properties.⁴⁵ Cellular uptake of PNA can be increased by either conjugating PNA to cationic peptide⁴⁶ or introduction of a cell transduction group in the PNA backbone.⁴⁷ There are two ways to obtain the guanidino function in the PNA backbone, either PNA monomer itself with protected guanidino function can be used or guanidinylation can be carried out after completion of solid phase synthesis. The post synthetic conversion of amino to guanidino group circumvents then necessity for prior protection of guanidino groups and simplifies the synthetic procedure; hence this method was chosen for guanidinylation. There are several methods reported for conversion of amino to guanidino functionality in solution, but very few reagents (Figure 3) were used for guanidinylation on solid support.

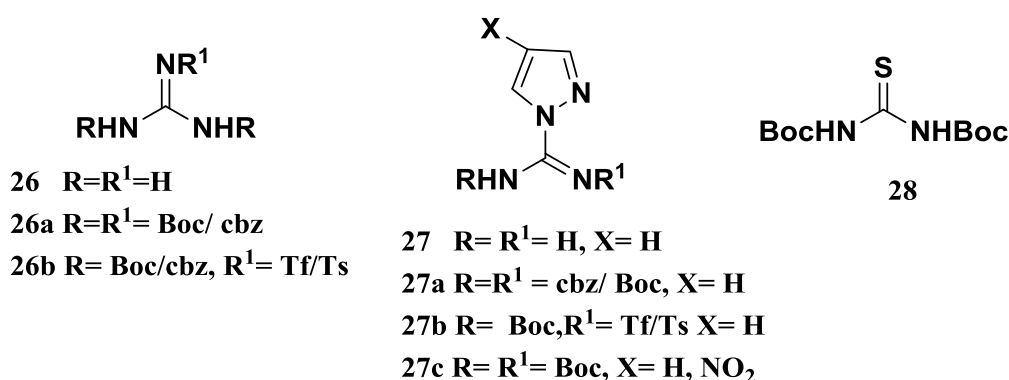
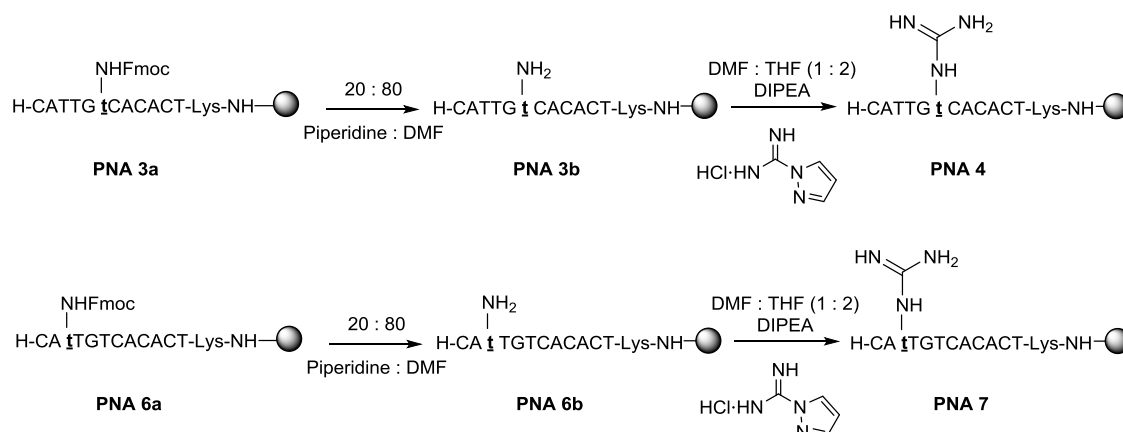


Figure 3. Different Guanidinylation reagents

The most commonly used guanidinylation reagents include guanidine or its derivatives (Figure 3, **26**) with suitable protections,⁴⁸ derivatives of pyrazole-1-carboxamide^{49, 50} derivatives of thiourea in combination with/without HgCl₂/Mukaiyama's reagent.⁵¹ As guanidinylation was carried out at the end of solid phase synthesis, unprotected pyrazole-1-carboxamide hydrochloride was used in present work. Scheme 4 represents the global guanidinylation of the Am-ethano-PNA sequences on the solid support. Fmoc protected amino groups in Am-ethano units in the Am-ethano-PNA deprotected by 20% piperidine-DMF solution and resin was then subjected to guanidinylation by suspending in the solution of pyrazole-1-carboxamide

hydrochloride (10 equivalents per amino group)/DIPEA in DMF-THF (1:2) mixture. The sequences are shown in Table 1, entries 4 and 7.



Scheme 4. Conversion of amino to guanidine- functionality on solid support

2A.3.3 Cleavage of the PNA Oligomers from the Solid Support

The oligomers were cleaved from the solid support, using TFMSA in the presence of TFA (Low, High TFMSA-TFA method),^{41, 52} which yields oligomers with an amide group at their C-terminus. In case of Am-ethano-PNA exocyclic amino-Fmoc protection was removed with 20% piperidine-DMF before cleavage from support and for Gu-ethano-PNA the free amino groups were subsequently converted to guanidino and then oligomers were cleaved with standard cleavage conditions. A cleavage time of 2 h at room temperature was found to be optimum. The side chain protecting groups for nucleobases were also cleaved during this cleavage process. After cleavage reaction, the oligomer was precipitated with dry diethyl ether.

Table 1. PNA sequences, HPLC retention time and MALDI-TOF mass
 t^H , t^{Am} and t^{Gu} are the modified PNA monomers with H, aminomethyl/guanidinomethyl substituents

Entry No.	Code	Sequences (N' → C')	HPLC t_R (min)	MALDI _ ToF Masses	
				Calcd	Obsd
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	11.8	3332.24	3330.76
2.	PNA 2	H- CATTG t^H CACACT-Lys-NH ₂	11.7	3356.40	3355.14
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	11.8	3385.42	3385.43
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	13.0	3427.44	3426.41
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	11.9	3385.42	3354.98
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	11.7	3385.42	3408.36 (M+Na ⁺)
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	12.5	3427.44	3427.39

2A.3.4 Purification and MALDI-TOF characterization of oligomers

The purity of the oligomers was checked by analytical RP-HPLC (C18 column, CH₃CN-H₂O- 0.1% TFA system), which showed more than 85-90% purity. A gradient elution method contained A = 5% Acetonitrile in water + 0.1% trifluoroacetic acid and B = 50% Acetonitrile in water + 0.1% trifluoroacetic acid (A to B = 100% in 30 min with a flow rate of 1.5 mL/min), and the eluent was monitored at 260 nm. The complete conversion of amino groups to guanidino functional group was confirmed by HPLC analysis by co-injection of Am-ethano-PNA and corresponding Gu-ethano-PNA oligomer. These were subsequently purified by reverse phase HPLC on a C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis.

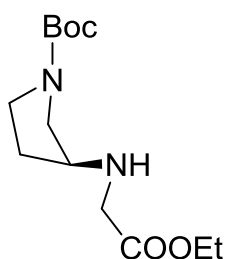
2A.4 Summary

- ❖ H-ethano-PNA and Amino-ethano-PNA thymine monomers have been synthesized and incorporated into *aeg*PNA sequences.
- ❖ The conversion of Amino-ethano-PNA sequences to Guanidino-ethano-PNA sequences was achieved by global guanidination on solid support.

2A.5 Experimental

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 distilled under an inert atmosphere according to the literature procedures. Reactions were monitored by thin layer chromatography. TLC was performed using TLC aluminium sheets pre-coated with silicagel 60 F254 (Merck). TLCs were visualized with UV light and iodine spray and/or by ninhydrin treatment then heating. Silica gel 60-120 mesh was used for column chromatography. ^1H and ^{13}C NMR spectra were recorded on 200 MHz and 50 MHz respectively. Chemical shifts are given in δ (ppm) scale. The specific rotation values were determined from Bellingham + Stanley Ltd. ADP 220 polarimeter. IR spectra were recorded in Perkin-Elmer Spectrum One FTIR Spectrometer. The HRMS data's were obtained from Thermo Fisher Scientific Q Exactive mass spectrometer. The crude PNAs were purified on a semi-preparative C18 column attached to a Waters HPLC system. For all the MALDI-TOF, spectra recorded on AB SCIEX TOF/TOF™ 5800 System; CHCA (α -cyano-4-hydroxycinnamic acid) was used as the matrix.

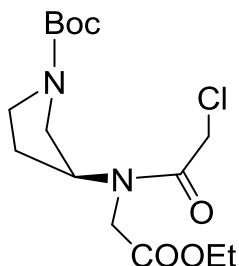
Ethyl-[-N-Boc-pyrrolidin-3-yl-] glycinate (7)



The amino compound **6** (2.7 g, 14.5 mmol) was dissolved in dry CH_3CN (40 mL) and triethylamine (4.1 mL, 29.0 mmol) was added with stirring. The reaction mixture was cooled to $0\text{ }^\circ\text{C}$ and ethyl bromoacetate (1.9 mL, 17.4 mmol) diluted with CH_3CN (20 mL) was added dropwise. The reaction mixture was stirred for 2 h at room temperature. After the completion of the reaction, CH_3CN was removed under reduced pressure and the residue was redissolved in ethyl acetate (150 mL). The organic layer was washed with water (2 X 30 mL), satd NaHCO_3 (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na_2SO_4 . It was then evaporated in vacuo to give the crude product, which on column chromatography purification (3% $\text{MeOH}-\text{CH}_2\text{Cl}_2$) gave ethyl-[-N-Boc-pyrrolidin-3-yl-] glycinate **7** (2.8 g, 70 %). ^1H NMR (CDCl_3) δ : 1.29 (t, 3H), 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.77-2.00 (m, 2H), 1.86 (s, 1H, D_2O exch.), 3.10 (m, 1H), 3.34-3.49 (m, 6H), 4.18 (q, 2H). ^{13}C NMR (CDCl_3) δ : 14.12, 28.43, 31.86, 44.23, 49.08, 51.14, 56.33, 60.91, 79.13, 154.50, 172.17. IR (ν_{max} , cm^{-1}) (CHCl_3): 732, 912, 1170, 1416, 1683, 1739, 2248, 2979, 3433.

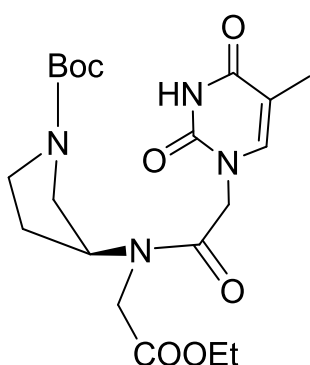
$[\alpha]_D^{27} + 68$ (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{13}\text{H}_{25}\text{N}_2\text{O}_4$: 273.1809, found: 273.1806.

Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-*N*-(2-chloroacetyl)] glycinate (**8**)



Compound **7** (2.6 g, 9.5 mmol) was dissolved in 1:1 solution of 1,4-dioxane-water (70 mL) and NaHCO_3 (8 g, 95.5 mmol) dissolved in water (50 mL) was added. The whole content was stirred for 15 mins at 0°C and chloroacetyl chloride (2.3 mL, 28.6 mmol) was added in two portions maintaining a pH ~ 8.0 . The reaction mixture was kept for 1 h with stirring and then concentrated to 50 mL. It was extracted in CH_2Cl_2 (50 mL) and was washed with brine (2 X 10 mL). The organic layer was dried over anhy. Na_2SO_4 followed by evaporation in vacuo to give the crude product. The product was purified by column chromatography (2% MeOH- CH_2Cl_2) to give the desired product Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-*N*-(2-chloroacetyl)] glycinate **8** (2.8 g, 84 %). ^1H NMR (CDCl_3) δ : 1.28 (t, 3H), 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.97-2.21 (m, 2H), 3.15-3.35 (m, 2H), 3.55-3.65 (m, 2H), 3.94-4.21 (m, 6H), 4.55 (qu, 1H). ^{13}C NMR (CDCl_3) δ : 13.90, 28.24, 29.27, 40.79, 43.63, 45.61, 47.32, 56.54, 61.32, 79.77, 154.01, 166.73, 168.51. IR (ν_{max} , cm^{-1}) (CHCl_3): 732, 914, 1130, 1203, 1416, 1668, 1695, 1747, 2250, 2981. $[\alpha]_D^{27} - 16$ (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{15}\text{H}_{25}\text{ClN}_2\text{O}_5\text{Na}$: 371.1347, found: 371.1396

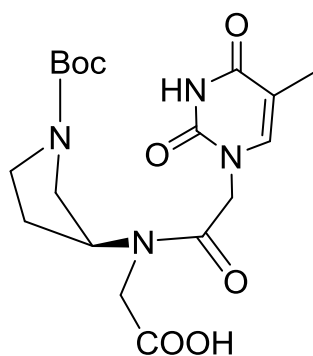
Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-*N*-(*N*1-thyminylacetyl)] glycinate (**9**)



The chloroacetyl derivative **8** (2.0 g, 5.7 mmol), K_2CO_3 (0.8 g, 5.7 mmol) and thymine (0.7 g, 5.7 mmol) was suspended in dry DMF (20 mL) and stirred at room temperature for 6 h. After that DMF was removed under reduced pressure and the residue was redissolved in ethyl acetate (70 mL). The organic layer was washed with water (2 X 25 mL) followed by brine (2 X 10 mL) and dried over anhy. Na_2SO_4 . It was then evaporated in vacuo to give the crude product which was purified by column chromatography (4% MeOH- CH_2Cl_2) to give Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-*N*-(*N*1-thyminylacetyl)] glycinate **9** (1.8 g, 71 %). ^1H NMR (CDCl_3) δ : 1.26 (t, 3H), 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.93 (d, 3H), 2.01 (m, 2H), 3.29-3.37 (m, 2H), 3.52 (m, 1H), 3.73 (q, 1H), 4.00 (d, 1H), 4.14-4.29 (m, 3H), 4.46 (m, 2H), 5.06 (m, 1H), 7.04 (s, 1H), 9.34 (br,

s, 1H). ^{13}C NMR (CDCl_3) δ : 12.23, 13.98, 28.32, 29.56, 43.75, 45.03, 47.44, 48.03, 55.24, 62.30, 79.68, 110.75, 140.90, 151.16, 154.20, 164.37, 167.79, 169.52. IR (ν_{max} , cm^{-1}) (CHCl_3): 667, 761, 1048, 1217, 1409, 1677, 1742, 2931, 3019, 3455. $[\alpha]_{\text{D}}^{27}$ - 14 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{20}\text{H}_{31}\text{N}_4\text{O}_7$: 439.2187, found: 439.2183

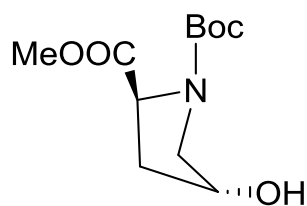
Ethyl-[-*N*-(*N*-Boc-pyrrolidin-3-yl)-*N*-(*N*1-thyminylacetyl)] glycine (**10**)



The amino ester **9** (1.5 g, 4.3 mmol) was dissolved in methanol (10 mL) and to that was added 1N NaOH (10 mL). The reaction mixture was stirred for 30 mins and after the completion of reaction methanol was evaporated under reduced pressure. The resulting aqueous solution was neutralized by Dowex H^+ ion exchange resin. The aqueous layer was washed with ethyl acetate (5 mL) followed by evaporation in vacuo

and dessication to obtain Ethyl-[-*N*-(*N*-Boc-pyrrolidin-3-yl)-*N*-(*N*1-thyminylacetyl)] glycine **10** (1.2 g, 85 %). ^1H NMR (D_2O) δ : 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.84 (s, 3H), 2.04 (m, 2H), 3.31 (m, 2H), 3.49 (m, 2H), 4.03-4.11 (2s, 2H), 4.59 (qu, 1H), 4.84 (m, 2H), 7.34 (s, 1H). ^{13}C NMR (D_2O) δ : 11.34, 27.76, 29.67, 43.84, 44.96, 46.82, 49.58, 55.11, 81.58, 110.82, 143.25, 152.24, 156.11, 166.94, 168.67, 174.87. IR (ν_{max} , cm^{-1}) (CHCl_3): 723, 1377, 1465, 1649, 1702, 2855, 2926, 3433. $[\alpha]_{\text{D}}^{27}$ - 24 (c 1.0, CH_3OH). HRMS (ESI) calcd for $\text{C}_{18}\text{H}_{26}\text{N}_4\text{O}_7\text{Na}$: 433.1694, found: 433.1693

Methyl-(*N*-Boc-4*R*-hydroxypyrrolidine)-2-carboxylate (**11**)

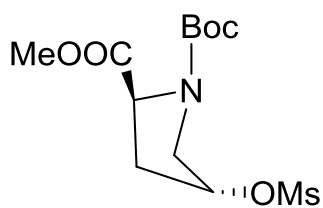


(2*S*, 4*R*)-Hydroxyproline **5** (10 g, 76.3 mmol) was suspended in methanol (100 mL) and stirred for 15 mins at 0 °C. To that added SOCl_2 (8.3 mL, 114.4 mmol) dropwise and the reaction mixture was stirred for 8 h at room temperature. The solvent was removed under reduced pressure and the residue was co-evaporated

thrice with CH_2Cl_2 . The crude product obtained was dissolved in 1:1 solution of 1,4-dioxane-water (200 mL). Triethylamine (31.8 mL, 228.8 mmol) was added slowly to neutralize and to provide basic medium. Boc anhydride (20.0 mL, 91.5 mmol) was added in two portions and the reaction mixture was stirred for 10 h. The resulting solution was concentrated to 75 mL and was extracted with ethyl acetate (100 mL). The organic layer was then washed with sat. NaHCO_3 (2 X 25 mL) followed by brine (2 X 20 mL). It was kept over anhy. Na_2SO_4 and evaporated in vacuo. The crude product

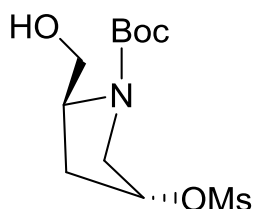
obtained was purified by column chromatography (40% ethyl acetate-petroleum ether) to obtain the desired product Methyl-(*N*-Boc-4*R*-hydroxypyrrolidine)-2-carboxylate **11** (17.4 g, 93 % over two steps). ¹H NMR (CDCl₃) δ: 1.41 (s, 9H, C(CH₃)₃), 2.06 (m, 1H), 2.30 (br, m, 2H), 3.58-3.62 (m, 2H), 3.74 (s, 3H), 4.40-4.49 (t, m, 2H). ¹³C NMR (CDCl₃) δ: 28.22, 39.09, 52.04, 54.68, 57.87, 69.45, 80.37, 153.95, 173.63. IR (ν_{max}, cm⁻¹) (CHCl₃): 773, 1159, 1416, 1682, 1749, 2978, 3447. [α]_D²⁷ - 54 (c 1.0, CHCl₃). HRMS (ESI) calcd for C₁₁H₁₉NO₅Na: 268.1155, found: 268.1158

Methyl-(*N*-Boc-4*R*-*O*-mesyl-pyrrolidine)-2-carboxylate (**12**)



The pure compound **11** (17.1 g, 69.7 mmol) obtained in the previous step was dissolved in dry pyridine (60 mL) and stirred for 15 mins at 0 °C. Mesyl chloride (6.5 mL, 83.7 mmol) was added drop by drop to this solution at 0 °C with stirring and the reaction mixture was left for 2 h. After that pyridine was removed under reduced pressure and the residue was redissolved in ethyl acetate (100 mL). Organic layer was washed with water (2 X 30 mL), sat. NaHCO₃ (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na₂SO₄. The organic layer was concentrated in vacuo to give the crude product which was purified by column chromatography (40% ethyl acetate-petroleum ether) to give Methyl-(*N*-Boc-4*R*-*O*-mesyl-pyrrolidine)-2-carboxylate **12** (19.4 g, 86 %). ¹H NMR (CDCl₃) δ: 1.42 (s, 9H, C(CH₃)₃), 2.19-2.28 (m, 1H), 2.61-2.65 (m, 1H), 3.06 (s, 3H, SCH₃), 3.75 (s, 3H), 3.79 (m, 2H), 4.41-4.45 (m, 1H), 5.27 (qu, 1H). ¹³C NMR (CDCl₃) δ: 28.15, 36.26, 38.71, 52.13, 52.28, 57.37, 77.85, 80.90, 153.29, 172.69. IR (ν_{max}, cm⁻¹) (CHCl₃): 783, 907, 1173, 1402, 1699, 1748, 2978. [α]_D²⁷ - 32 (c 1.0, CHCl₃). HRMS (ESI) calcd for C₁₂H₂₁NO₇SNa: 346.0931, found: 346.0923

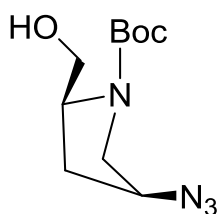
N-Boc-4*R*-*O*-mesyl-prolinol (**13**)



NaBH₄ (2.5 g) was suspended in dry THF (105 mL) and absolute alcohol (140 mL) and stirred for 30 mins in ice bath. At 0 °C LiCl (2.5 g) was added in two portions and the content was stirred for another 1 h in ice bath for in situ generation of LiBH₄. The solution turned milky due to formation of NaCl. To this mixture, the mesylate derivative **12** (6.3 g, 19.5 mmol) dissolved in absolute alcohol (60 mL) was added at 0 °C dropwise. The reaction mixture was stirred for 8 h. NH₄Cl solution was then added for maintaining pH ~7.0 and the solvents were removed under reduced

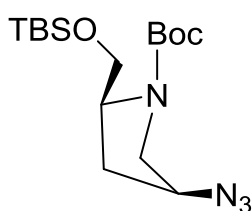
pressure. The residue was dissolved in ethyl acetate (125 mL). Organic layer was washed with water (2 X 30 mL) followed by brine (2 X 20 mL) and dried over anhy. Na_2SO_4 . It was evaporated under vacuo and the crude product obtained was purified by column chromatography (50% ethyl acetate-petroleum ether) to give *N*-Boc-4*R*-O-mesyl-prolinol **13**. The same procedure was repeated thrice with same amount of compound **12** to get compound **13** (15.4 g, 89 %). ^1H NMR (CDCl_3) δ : 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 2.05 (br, m, 1H), 2.33-2.40 (m, 1H), 3.05 (s, 3H, SCH_3), 3.53-3.59 (m, 2H), 3.77-3.92 (m, 2H), 4.14 (m, 1H), 5.20 (qu, 1H). ^{13}C NMR (CDCl_3) δ : 28.28, 35.17, 38.64, 53.46, 58.49, 65.73, 78.32, 81.04, 156.05. IR (ν_{max} , cm^{-1}) (CHCl_3): 903, 1169, 1412, 1676, 2978, 3430. $[\alpha]_{\text{D}}^{27}$ -34 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{11}\text{H}_{21}\text{NO}_6\text{SNa}$: 318.0982, found: 318.0983

N-Boc-4*S*-azido-prolinol (**14**)



The compound **13** (5.0 g, 16.9 mmol) was dissolved in dry DMF (30 mL) and NaN_3 (5.5 g, 84.6 mmol) was added to it. The reaction mixture was stirred at 60 °C for 7 h. After completion of reaction DMF was removed under reduced pressure and the residue was dissolved in ethyl acetate (100 mL). Organic layer was washed with water (2 X 25 mL) followed by brine (2 X 20 mL). The organic layer was kept over anhy. Na_2SO_4 and then concentrated in vacuo. The crude compound obtained was purified by column chromatography (30% ethyl acetate-petroleum ether) to obtain *N*-Boc-4*S*-azido-prolinol **14**. (3.3 g, 80 %). ^1H NMR (CDCl_3) δ : 1.47 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.79-1.87 (br, 2H), 2.26-2.40 (m, 1H), 3.27-3.36 (dd, 1H), 3.68-3.83 (m, 3H), 4.07 (m, 2H). ^{13}C NMR (CDCl_3) δ : 28.29, 33.95, 52.03, 58.65, 59.29, 67.20, 80.97, 156.44. IR (ν_{max} , cm^{-1}) (CHCl_3): 756, 870, 1045, 1167, 1406, 1676, 2104, 2976, 3424. $[\alpha]_{\text{D}}^{27}$ - 12 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{10}\text{H}_{19}\text{N}_4\text{O}_3$: 243.1452, found: 243.1454

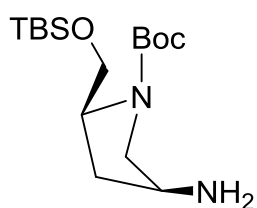
N-Boc-4*S*-azido-2*S*-*O*-TBS-prolinol (**15**)



The pure compound **14** (9.5 g, 39.2 mmol) was dissolved in dry DMF (40 mL). Imidazole (8.0 g, 117.6 mmol) followed by TBSCl (8.3 g, 54.9 mmol) was added to it. The reaction mixture was stirred for 8 h after the completion of the reaction the solvent was removed under reduced pressure and the residue was dissolved in ethyl acetate (150 mL). The organic layer was then washed with water (2 X 30 mL) followed

by brine (2 X 20 mL). It was kept over anhy. Na₂SO₄ and evaporated in vacuo. The crude product obtained was purified by column chromatography (10% ethyl acetate-petroleum ether) to obtain *N*-Boc-4*S*-azido-2*S*-*O*-TBS-prolinol **15** (12.9 g, 92 %). ¹H NMR (CDCl₃) δ: 0.06 (s, 6H), 0.84 (s, 9H), 1.40 (s, 9H, C(CH₃)₃), 2.14 (t, 2H), 3.17 (m, 1H), 3.54-3.75 (m, 4H), 4.00 (m, 1H). ¹³C NMR (CDCl₃) δ: - 3.68, - 5.48, 25.76, 28.34, 31.87, 32.85, 51.73, 57.59, 58.91, 62.54, 79.79, 154.01. IR (ν_{max}, cm⁻¹) (CHCl₃): 775, 837, 1119, 1395, 1699, 2102, 2930. [α]_D²⁷ - 4 (c 1.0, CHCl₃). HRMS (ESI) calcd for C₁₆H₃₃N₄O₃Si: 357.2316, found: 357.2315

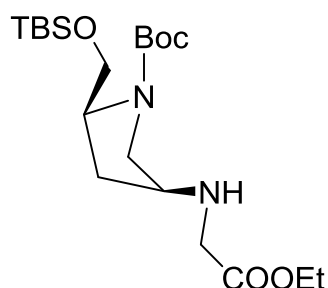
N-Boc-4*S*-amino-2*S*-*O*-TBS-prolinol (**16**)



The azido compound **15** (4.2 g, 11.8 mmol) was dissolved in methanol (35 mL) and was hydrogenated over 10% Pd-C catalyst with 40 psi pressure for 2.5 h under hydrogen atmosphere. The content was then filtered through a bed of celite and the filtrate collected was concentrated to obtain the desired compound *N*-

Boc-4*S*-amino-2*S*-*O*-TBS-prolinol **16**. (3.8 g, 97 %) and was used further without purification. ¹H NMR (CDCl₃) δ: 0.02 (m, 6H), 0.84 (s, 9H), 1.40 (s, 9H, C(CH₃)₃), 2.14 (t, 2H), 3.16 (m, 1H), 3.54-3.75 (m, 4H), 4.00 (m, 1H). ¹³C NMR (CDCl₃) δ: -3.68, -5.48, 25.76, 28.34, 31.87, 32.85, 51.73, 57.59, 58.91, 62.54, 79.79, 154.01. IR (ν_{max}, cm⁻¹) (CHCl₃): 775, 837, 1123, 1416, 1665, 2930, 3433. [α]_D²⁷ - 26 (c 1.0, CHCl₃). HRMS (ESI) calcd for C₁₆H₃₅N₂O₃Si: 331.2411, found: 331.2412

Ethyl-[-*N*-(*N*-Boc-pyrrolidin-3-yl)-2*S*-*O*-TBS-methyl] glycinate (**17**)

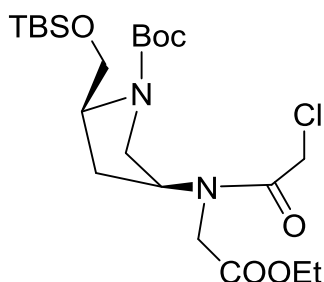


The amino compound **16** (5.5 g, 16.6 mmol) was dissolved in dry CH₃CN (60 mL) and triethylamine (4.6 mL, 33.3 mmol) was added with stirring. The reaction mixture was cooled to 0 °C and ethyl bromoacetate (2.3 mL, 20.0 mmol) diluted with CH₃CN (25 mL) was added dropwise. The reaction mixture was stirred for 2 h at room temperature. CH₃CN was removed under

reduced pressure and the residue was dissolved in ethyl acetate (150 mL). The organic layer was washed with water (2 X 30 mL), sat. NaHCO₃ (2 X 20 mL) followed by brine (2 X 10 mL) and was dried over anhy. Na₂SO₄. It was then evaporated in vacuo to give crude product, which on column chromatography purification (30% ethyl acetate-petroleum ether) gave Ethyl-[-*N*-(*N*-Boc-pyrrolidin-3-yl)-2*S*-*O*-TBS -methyl] glycinate

17. (3.2 g, 69 %). ^1H NMR (CDCl_3) δ : 0.05 (m, 6H), 0.84 (s, 9H), 1.24 (t, 3H), 1.41 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.81-2.00 (m, 2H), 2.15 (br, m, 1H), 2.95 (dd, 1H), 3.18 (qu, 1H), 3.35 (s, 2H), 3.68 (m, 4H), 4.09-4.20 (q, 2H). ^{13}C NMR (CDCl_3) δ : - 5.43, 14.12, 25.85, 28.43, 33.74, 34.96, 49.32, 52.37, 55.64, 57.70, 60.78, 64.43, 79.17, 154.28, 172.14. IR (ν_{max} , cm^{-1}) (CHCl_3): 733, 837, 1103, 1169, 1402, 1697, 1742, 2359, 2930, 3333. $[\alpha]_{\text{D}}^{27}$ - 30 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{20}\text{H}_{41}\text{N}_2\text{O}_5\text{Si}$: 417.2779, found: 417.2789

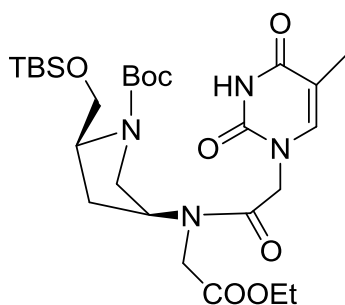
Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-O-TBS-methyl)-N-(2-chloroacetyl)] glycinate (18)



Compound **17** (4.6 g, 11.0 mmol) was dissolved in 1:1 solution of 1,4-dioxane-water (80 mL) and NaHCO_3 (9.3 g, 110.4 mmol) dissolved in water (60 mL) was added. The reaction mixture was stirred for 15 min at 0 $^\circ\text{C}$ and chloroacetyl chloride (4.4 mL, 55.2 mmol) was added in two portions maintaining a pH \sim 8.0. The reaction mixture was kept for 1 h with stirring and then concentrated to 50 mL. It was extracted in CH_2Cl_2 (50 mL) and was washed with brine (2 X 10 mL). The organic layer was dried over anhyd. Na_2SO_4 followed by evaporation in vacuo to give the crude product. The product was purified by column chromatography (20% ethyl acetate-petroleum ether) to give the desired product Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-O-TBS-methyl)-N-(2-chloroacetyl)] glycinate **18**. The same procedure was repeated twice with same amount of compound **17** to get compound **18** (8.7 g, 79 %). ^1H NMR (CDCl_3) δ : 0.05 (s, 6H), 0.89 (s, 9H), 1.27 (t, 3H), 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.64 (s, 2H), 2.05-2.30 (m, 2H), 3.56 (m, 1H), 3.91-4.03 (m, s, 4H), 4.14-4.21 (m, q, 5H). ^{13}C NMR (CDCl_3) δ : - 5.42, 14.05, 18.27, 25.87, 28.45, 29.65, 40.95, 43.56, 48.63, 51.28, 56.25, 61.37, 62.13, 79.96, 153.65, 166.92, 168.75 IR (ν_{max} , cm^{-1}) (CHCl_3): 777, 837, 1113, 1167, 1408, 1667, 1694, 1753, 2361, 2930. $[\alpha]_{\text{D}}^{27}$ - 22 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{22}\text{H}_{42}\text{ClN}_2\text{O}_6\text{Si}$: 493.2495, found: 493.2500

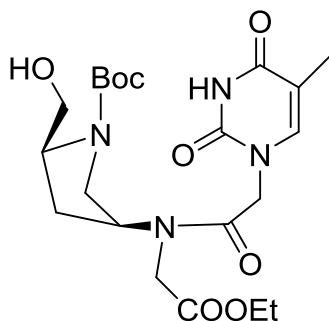
Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-O-TBS-methyl)-N-(N1-thyminylacetyl)] glycinate (19)

The chloroacetyl derivative **18** (5.0 g, 10.1 mmol), K_2CO_3 (1.4 g, 10.1 mmol) and thymine (1.3 g, 10.1 mmol) were suspended in dry DMF (30 mL) and stirred at room temperature for 6 h. After that DMF was removed under reduced pressure and the residue was redissolved in ethyl acetate (80 mL). The organic layer was washed with



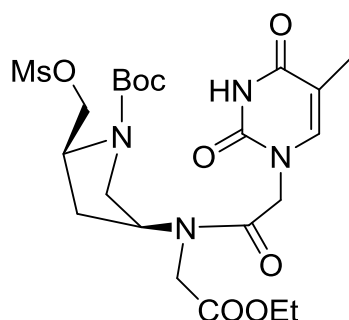
water (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na_2SO_4 . It was then evaporated in vacuo to give crude product. The crude product was purified by column chromatography (3% $\text{MeOH-CH}_2\text{Cl}_2$) to give Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-2*S*-*O*-TBS-methyl)-*N*-(*N*1-thyminylacetyl)] glycinate **19** (4.5 g, 76 %). $^1\text{H NMR}$ (CDCl_3) δ : 0.02 (s, 6H), 0.84 (s, 9H), 1.21 (t, 3H), 1.42 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.87 (d, 3H), 2.10-2.28 (m, 2H), 3.51 (m, 1H), 3.86 (m, 2H), 4.02-4.15 (s, m, 5H), 4.32 (q, 2H), 5.06 (m, 2H), 6.97 (d, 1H), 9.17 (d, 1H). $^{13}\text{C NMR}$ (CDCl_3) δ : -5.58, 12.25, 13.98, 25.81, 28.38, 29.59, 31.55, 43.43, 44.28, 47.70, 51.11, 56.20, 61.44, 62.22, 80.03, 110.70, 140.85, 151.04, 153.57, 164.25, 166.72, 168.89. IR (ν_{max} , cm^{-1}) (CHCl_3): 760, 839, 1117, 1215, 1412, 1680, 1686, 1744, 2400, 2928, 3393. $[\alpha]_{\text{D}}^{27}$ - 16 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{27}\text{H}_{47}\text{N}_4\text{O}_8\text{Si}$: 583.3158, found: 583.3157

Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-2*S*-hydroxymethyl)-*N*-(*N*1-thyminylacetyl)] glycinate (20**)**



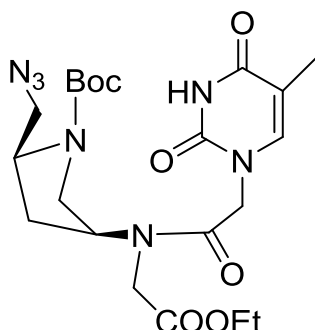
Compound **19** (4.2 g, 7.2 mmol) was dissolved in dry THF (20 mL), 1M TBAF in THF (6.0 mmol) was added to it. The reaction mixture was stirred for 30 min. THF was then removed under reduced pressure and the residue was dissolved in CH_2Cl_2 (50 mL). The organic layer was washed with water (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na_2SO_4 . After evaporation of the solvent in vacuo, the crude product was purified by column chromatography (6% $\text{MeOH-CH}_2\text{Cl}_2$) to Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl)-2*S*-hydroxymethyl)-*N*-(*N*1-thyminylacetyl)] glycinate **20** (3.2 g, 94 %). $^1\text{H NMR}$ (CDCl_3) δ : 1.22-1.32 (t, 3H), 1.44 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.69 (m, 4H) 1.90 (d, 2H), 2.77-3.20 (m, 1H), 3.31-3.39 (m, 4H), 3.68-3.77 (m, 2H), 3.91-4.15 (m, 2H), 4.29 (m, 1H), 4.58-4.92 (m, 2H), 7.19 (s, 1H). $^{13}\text{C NMR}$ (CDCl_3) δ : 12.01, 13.40, 19.44, 23.79, 28.09, 45.16, 48.21, 51.91, 58.61, 61.16, 61.98, 77.20, 80.38, 109.98, 141.28, 150.97, 164.26, 166.88, 168.80. IR (ν_{max} , cm^{-1}) (CHCl_3): 770, 1028, 1163, 1416, 1649, 1686, 1744, 2359, 2965, 3416. $[\alpha]_{\text{D}}^{27}$ - 4 (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{21}\text{H}_{33}\text{N}_4\text{O}_8$: 469.2293, found: 469.2294

**Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-O-mesyl-methyl)-N-(N1-thyminylacetyl)]
glycinate (21)**



The pure compound **20** (2.9 g, 6.2 mmol) obtained in the previous step was dissolved in dry pyridine (20 mL) and stirred for 15 min at 0 °C. Mesyl chloride (0.7 mL, 9.3 mmol) was added dropwise to this solution at 0 °C with stirring and the reaction mixture was left for 2 h. Pyridine was removed under reduced pressure and the residue was dissolved in ethyl acetate (70 mL). Organic layer was washed with water (2 X 30 mL), sat. NaHCO₃ (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na₂SO₄. The organic layer was concentrated in vacuo to give the crude product which was purified by column chromatography (4% MeOH-CH₂Cl₂) to give Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-O-mesyl-methyl)-N-(N1-thyminylacetyl)] glycinate **21** (3.0 g, 88 %). ¹H NMR (CDCl₃) δ: 1.27-1.34 (t, 3H), 1.45 (s, 9H, C(CH₃)₃), 1.92 (d, 3H), 2.28-2.49 (m, 2H), 3.04 (s, 3H), 3.80 (m, 1H), 4.04 (s, 2H), 4.18 (q, 2H), 4.27-4.33 (s, m, 3H), 4.45 (m, 2H), 4.66-4.91 (m, 2H), 7.03 (s, 1H), 9.5 (br, s, 1H). ¹³C NMR (CDCl₃) δ: 12.25, 13.96, 28.24, 37.07, 43.49, 47.81, 48.27, 54.17, 61.60, 62.38, 69.68, 77.20, 80.64, 110.70, 140.93, 151.15, 153.82, 164.35, 167.93, 169.46. IR (ν_{max}, cm⁻¹) (CHCl₃): 731, 914, 1198, 1406, 1647, 1685, 1744, 2361, 2982, 3198. [α]_D²⁷ + 6 (c 1.0, CHCl₃). MS (ESI) calcd for C₂₂H₃₅N₄O₁₀S: 547.6000, found: 547.3760

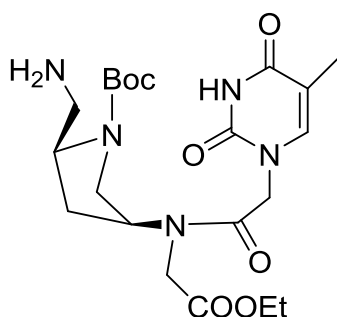
**Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-azidomethyl)-N-(N1-thyminylacetyl)]
glycinate (22)**



The mesylate derivative **21** (2.7 g, 4.9 mmol) was dissolved in dry DMF (30 mL) and NaN₃ (1.6 g, 24.7 mmol) was added to it. The reaction mixture was heated to 60 °C for 7 h. After completion of the reaction DMF was removed under reduced pressure and residue was redissolved in ethyl acetate (80 mL). Organic layer was washed with water (2 X 25 mL) followed by brine (2 X 20 mL). The organic layer was kept over anhy. Na₂SO₄ and then concentrated in vacuo. The crude compound obtained was purified by column chromatography (3% MeOH-CH₂Cl₂) to obtain Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-azidomethyl)-N-(N1-thyminylacetyl)] glycinate **22** (1.8

g, 73 %). ^1H NMR (CDCl_3) δ : 1.28-1.34 (t, 3H), 1.46 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.92 (d, 3H), 2.14-2.35 (m, 2H), 2.95-3.10 (m, 1H), 3.26-3.31 (d, 1H), 3.80 (m, 1H), 3.95-4.05 (m, 2H), 4.21-4.31 (m, 3H), 4.44-4.77 (m, 3H), 4.95-5.04 (m, 1H), 7.03 (s, 1H), 9.57 (d, 1H). ^{13}C NMR (CDCl_3) δ : 12.27, 13.97, 28.29, 43.45, 44.79, 47.68, 48.09, 51.35, 54.83, 61.68, 62.38, 80.46, 110.81, 140.82, 151.13, 153.72, 164.25, 167.96, 169.51. IR (ν_{max} , cm^{-1}) (CHCl_3): 731, 912, 1115, 1206, 1470, 1651, 1694, 1748, 2104, 2251, 2978, 3175. $[\alpha]_{\text{D}}^{27} + 12$ (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{21}\text{H}_{32}\text{N}_7\text{O}_7$: 494.2358, found: 494.2355

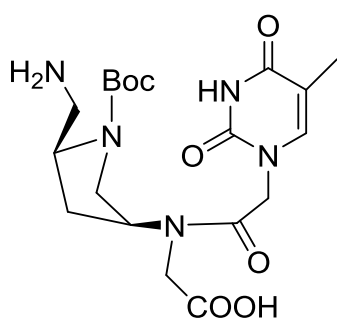
Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-aminomethyl)-N-(N1-thyminylacetyl)] glycinate (23)



The azido compound **22** (1.5 g, 3.0 mmol) was dissolved in methanol (20 mL) and was hydrogenated over 10 % Pd-C catalyst with 40 psi pressure for 3 h under hydrogen atmosphere. The content was then filtered through a bed of celite and the filtrate collected was concentrated to obtain the desired compound Ethyl-[-N (N-Boc-pyrrolidin-3-yl-2S-

aminomethyl)-N-(N1-thyminylacetyl)] glycinate **23** (1.4 g, 95 %) and was used further without purification. HRMS (ESI) calcd for $\text{C}_{21}\text{H}_{34}\text{N}_5\text{O}_7$: 468.2453, found: 468.2450

Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-aminomethyl)-N-(N1-thyminylacetyl)] glycine (24)

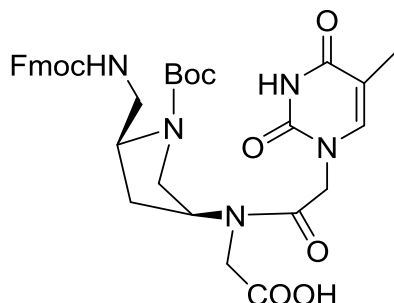


The amino ester **23** (1.2 g, 2.6 mmol) formed above was dissolved in methanol (10 mL) and to that added 1N NaOH (10 mL). The content was stirred for 30 min and after the completion of reaction; methanol was evaporated under reduced pressure. The resulting aqueous solution was neutralized by Dowex H^+ ion exchange resin. The aqueous layer was washed with

ethyl acetate (8 mL) followed by evaporation in vacuo and desiccation to obtain Ethyl-[-N-(N-Boc-pyrrolidin-3-yl-2S-aminomethyl)-N-(N1-thyminylacetyl)] glycine **24** (0.9 g, 83 %). ^1H NMR (D_2O) δ : 1.42 (s, 9H, $\text{C}(\text{CH}_3)_3$), 1.85-1.88 (d, 3H), 3.23-3.31 (m, s 3H), 3.65 (m, 1H), 3.80-3.90 (m, 2H), 3.97 (s, 2H), 4.25 (s, 1H), 4.58 (s, 2H), 7.34 (d, 1H). ^{13}C NMR (D_2O) δ : 11.34, 23.30, 27.60, 45.54, 46.89, 49.62, 52.97, 53.89, 55.50, 82.20, 110.88, 147.23, 152.59, 155.89, 167.31, 169.66, 175.43. IR (ν_{max} , cm^{-1}) (CHCl_3): 729,

904, 1067, 1266, 1404, 1667, 1682, 1692, 1711, 1721, 2400, 3019, 3374, 3451. $[\alpha]_{\text{D}}^{27} - 2$ (*c* 1.0, CH₃OH). HRMS (ESI) calcd for C₁₉H₃₀N₅O₇: 440.2140, found: 440.2141

Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl-2*S*-Fmoc-aminomethyl)-*N*-(*N*1-thyminylacetyl)] glycine (25**)**



The amino acid obtained in the previous step **24** (0.7 g, 1.6 mmol) was dissolved in 1:1 solution of 1,4-dioxane-water (30 mL) and NaHCO₃ (0.8 g, 9.6 mmol) and Fmoc-succinimide (0.6 g, 1.9 mmol) was added to it. The reaction mixture was stirred for 6 h at room temperature. After completion of reaction the solvent was removed

under reduced pressure and aqueous layer was washed with ethyl acetate until impurity of Fmoc succinimide was removed. Then aqueous layer was neutralized with Dowex H⁺ ion exchange resin. The compound was extracted in 5% MeOH-ethyl acetate (30 mL). Organic layer was dried over anhy. Na₂SO₄, concentrated and the crude product was purified by column chromatography (10% MeOH-CH₂Cl₂) to give Ethyl-[*N*-(*N*-Boc-pyrrolidin-3-yl-2*S*-Fmoc-aminomethyl)-*N*-(*N*1-thyminylacetyl)] glycine **25** (0.7 g, 66.5%). ¹H NMR (CDCl₃) δ: 1.41 (s, 9H, C(CH₃)₃), 1.75 (m,br, 3H), 2.87-3.49 (m, br, 9H), 3.71-4.03 (m, br, 4H), 4.18-4.36 (m, br, 4H), 4.71 (br, 1H), 6.92, (s, br, 1H), 7.35 (m, 4H), 7.55-7.72 (d,d, 4H) ¹³C NMR (CD₃OD) δ: 12.12, 14.42, 21.59, 23.66, 28.68, 30.66, 45.17, 48.13, 50.14, 57.17, 69.58, 81.78, 108.36, 111.05, 120.66, 122.03, 128.22, 129.89, 139.23, 141.32, 143.95, 144.85, 153.24, 167.21, 170.47, 175.39. IR (ν_{max}, cm⁻¹) (CHCl₃): 760, 878, 1047, 1215, 1371, 1510, 1622, 1667, 1694, 1732, 1755, 2361, 3019, 3242, 3451. $[\alpha]_{\text{D}}^{27} + 20$ (*c* 1.0, CH₃OH). HRMS (ESI) calcd for C₃₄H₄₀N₅O₉: 662.2821, found: 662.2823

Solid phase oligomer Synthesis

The PNA oligomers were synthesized by standard Boc-solid phase peptide strategy on MBHA resin having initial loading value 1.75 mmol/g. The loading value was lowered to 0.3 mmol/g by capping with acetic anhydride in dry CH₂Cl₂ and pyridine as base. The free amines of the resin was then functionalized by coupling with *N*(α)-Boc-*N*(ε)-2-Cl-Cbz-*L*-Lysine using TBTU/HOBt as coupling reagent and DIPEA as base in DMF.

Picric acid estimation of resin functionalization: The typical procedure for estimation of the loading value of the resin was carried out with 5 mg of the resin which comprises the following steps:

The resin was swollen in dry CH_2Cl_2 for 2 hrs. After soaking with CH_2Cl_2 , it was drained off and a 50% solution of TFA in CH_2Cl_2 was added (1 mL x 3), 15 min each. After washing thoroughly with CH_2Cl_2 , the TFA salt was neutralized with 5% solution of DIPEA in CH_2Cl_2 (1 mL x 5min each). The free amine was treated with a 0.1M picric acid solution in CH_2Cl_2 (1 mL x 2 min each). The excess picric acid was removed by extensively washing the resin with CH_2Cl_2 . The adsorbed picric acid was displaced from the resin by adding a solution of 5% DIPEA in CH_2Cl_2 . The eluant was collected and the volume was made up to 10 mL with CH_2Cl_2 in a volumetric flask. The absorbance was recorded at 358nm in ethanol and the concentration of the amine groups on the resin was calculated using the molar extinction coefficient of picric acid as $14,500 \text{ cm}^{-1}\text{M}^{-1}$ at 358nm.

The lysine functionalized resin was swollen in dry CH_2Cl_2 for 30 min and then treated with 50% TFA in CH_2Cl_2 (3 X 2 mL, 15 min each). After that the resin was washed with CH_2Cl_2 (3 X 2 mL) followed by DMF (3 X 2 mL) and again by CH_2Cl_2 (3 X 2 mL). The neutralization of the TFA salt formed was done by treating the resin with 5% DIPEA in CH_2Cl_2 (3 X 2 mL, 5 min each) and washed further with CH_2Cl_2 (3 X 2 mL). The unmodified and modified PNA monomers were coupled to the resin using TBTU/HOBt as coupling reagent and DIPEA as base in DMF. The efficiency of the Boc deprotection for free amine and the coupling of monomers were monitored by Kaiser's test. At the end of the oligomers synthesis, the oligomers were cleaved from the resin with standard TFA-TFMSA protocol.

Kaiser's Test: Kaiser's test was used to monitor the Boc-deprotection and amide coupling steps in the solid phase peptide synthesis. Three solutions were used, viz. (1) Ninhydrin (5.0 g) dissolved in ethanol (100 mL) (2) Phenol (80 g dissolved in ethanol (20 mL) and (3) KCN (0.001M aqueous solution of KCN in 98 mL pyridine). To a few beads of the resin taken in a test tube, was added 3-4 drops of each of the three solutions described above. The tube was heated for 5 min, and the colour of the beads was noted. A blue colour on the beads and in the solution indicated successful deprotection, while colourless beads were observed upon completion of the amide coupling reaction.

Cleavage of the PNA oligomers from the solid support

In a typical cleavage reaction, the resin bound oligomers (5 mg) were treated with thioanisole (10 μL) and 1,2- ethanedithiol (4 μL) for 10 min in an ice bath followed by TFA (120 μL) and TFMSA (8 μL) at 0 °C for 2 h. After that the resin

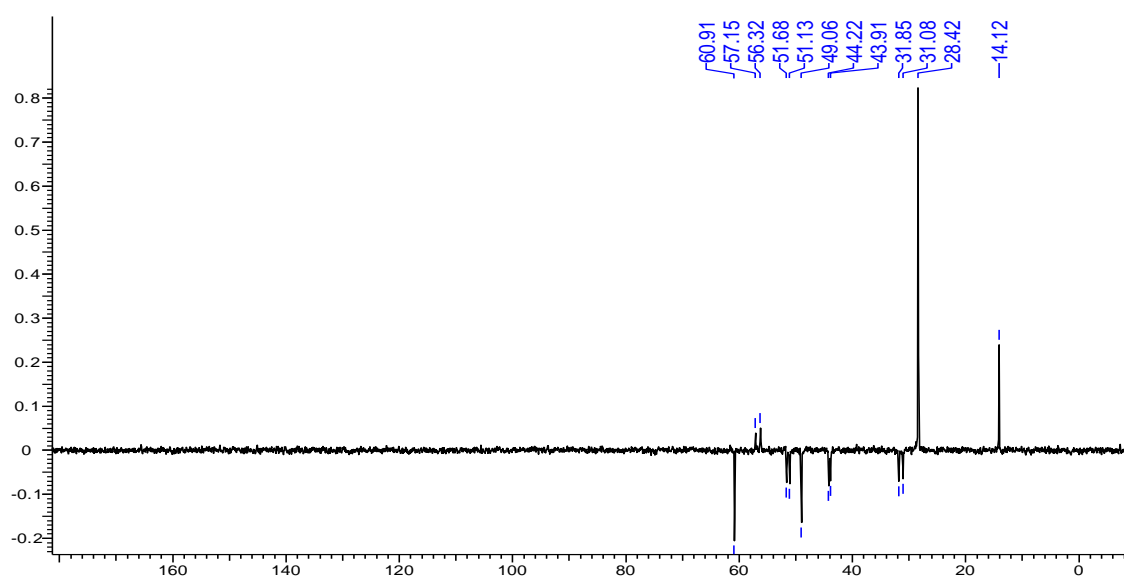
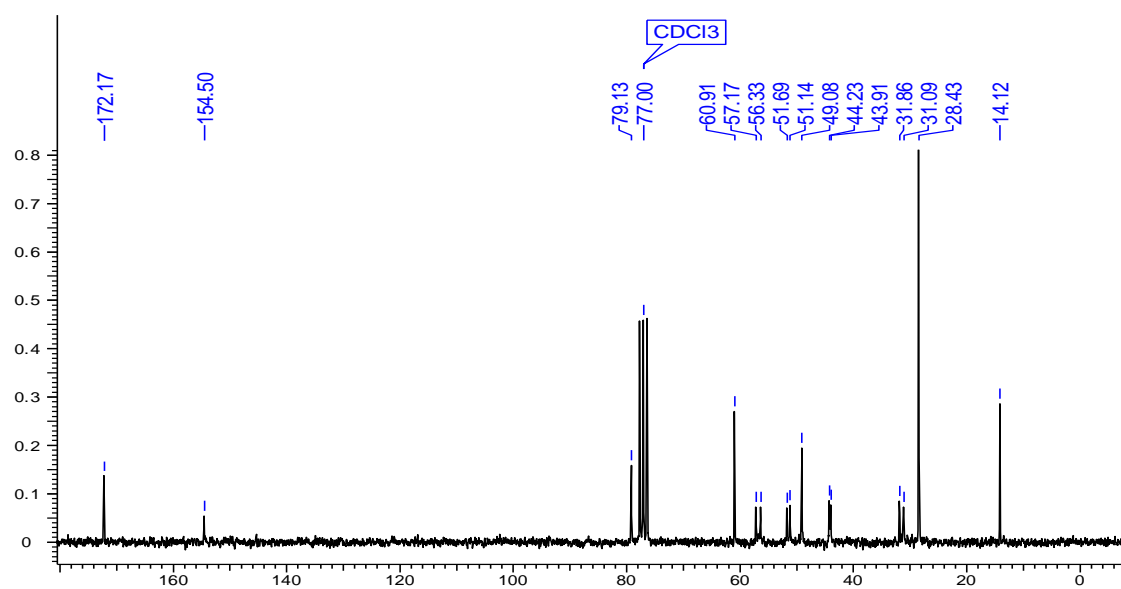
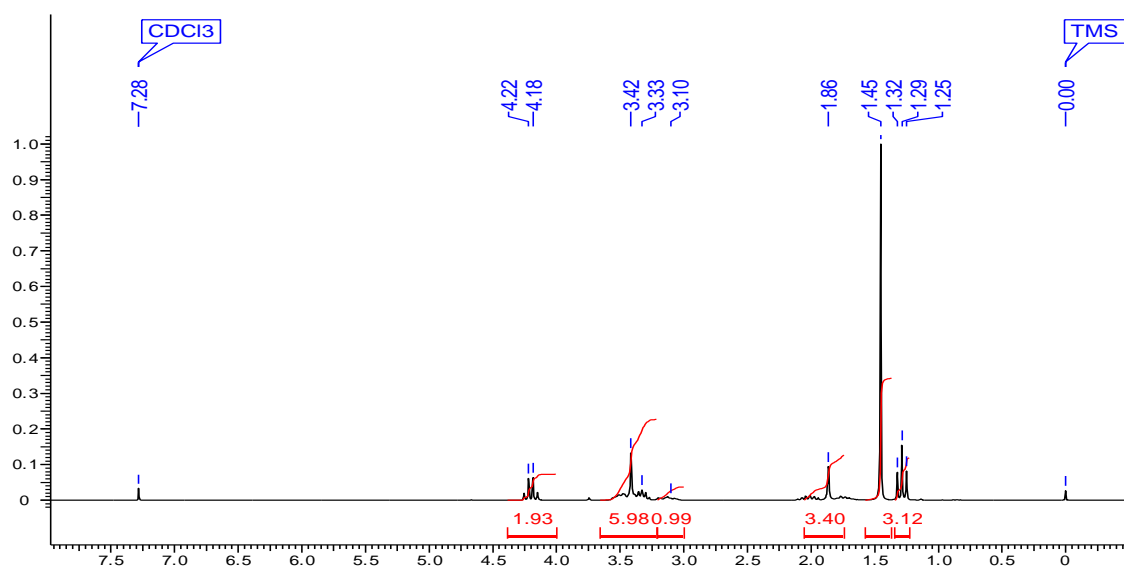
was filtered using TFA (400 μ L) followed by evaporation of the filtrate under vacuo. The oligomers were precipitated by adding cold ether. The precipitated oligomers were washed with cold ether (2 X 300 μ L) and were dissolved in deionised water. Purification of all the PNA oligomers was carried out by HPLC on RP-C18 column with water: CH₃CN-0.1% TFA system and were characterized by MALDI-TOF mass spectrometry. The spectra were acquired in linear mode and the matrix used for analysis was CHCA (α -cyano-hydroxyl-cinnamic acid).

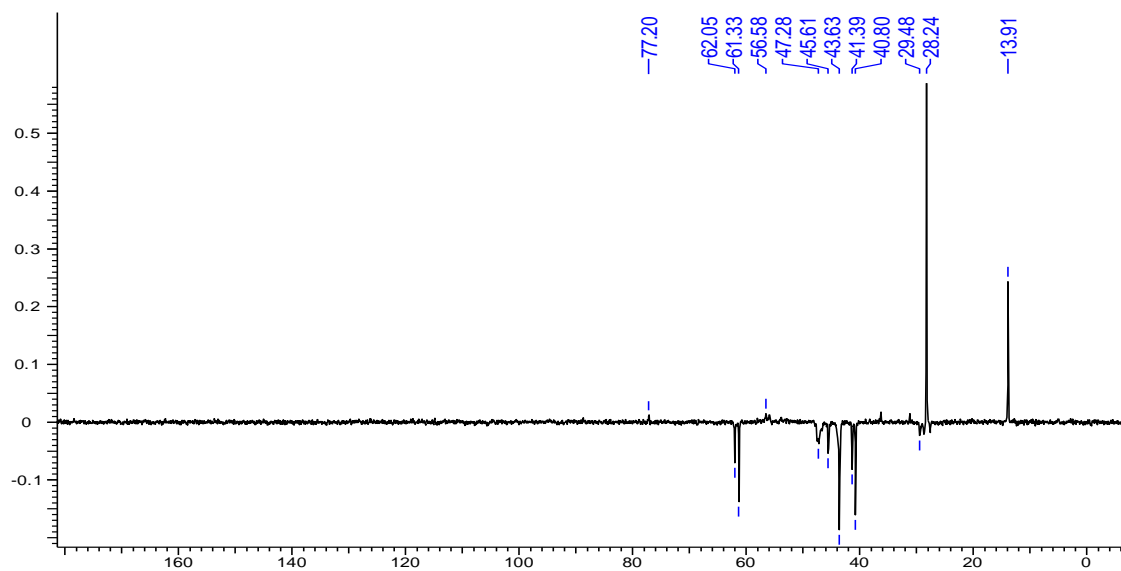
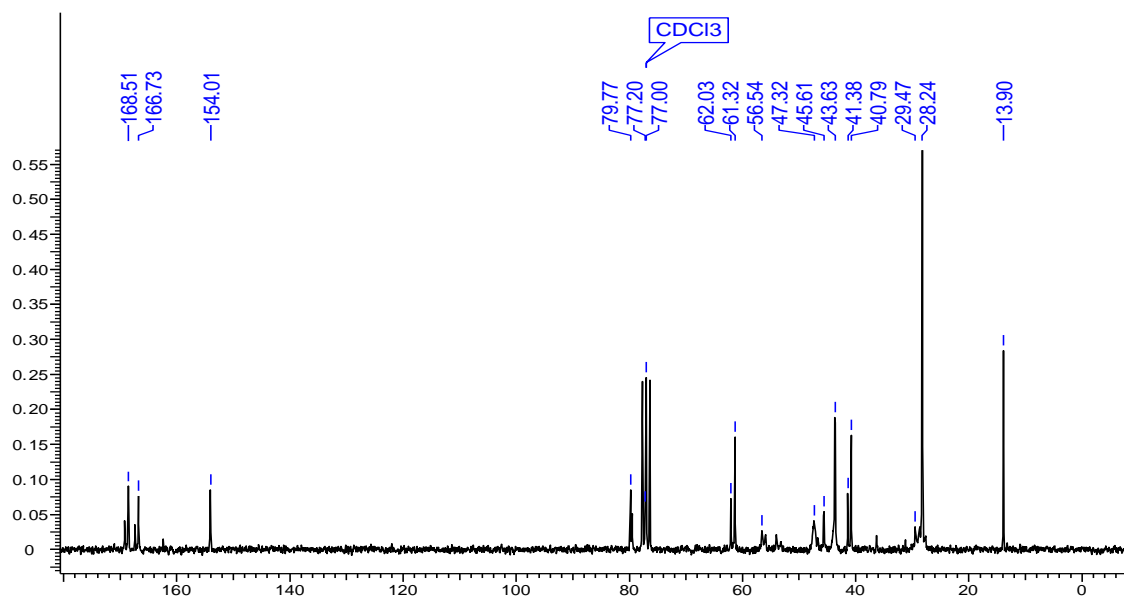
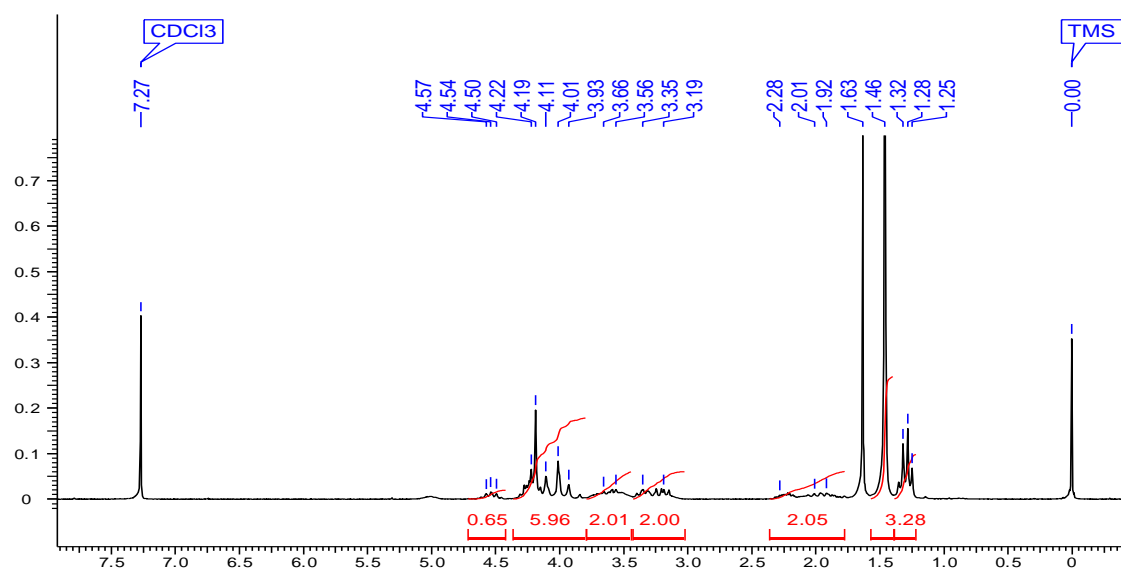
Post synthetic conversion of Am-ethano-PNA to Gu-ethano-PNA

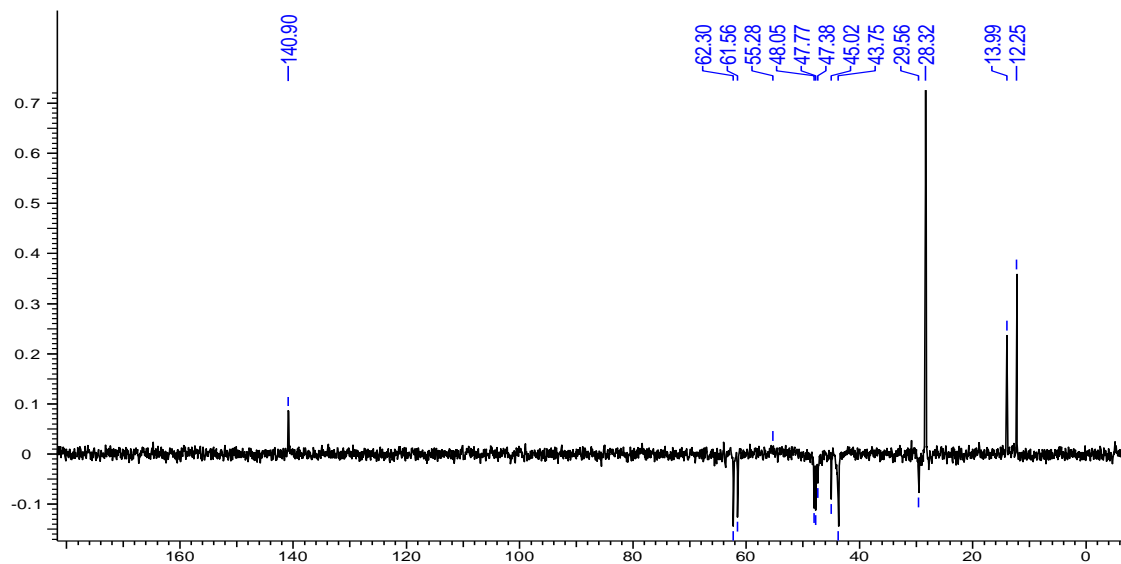
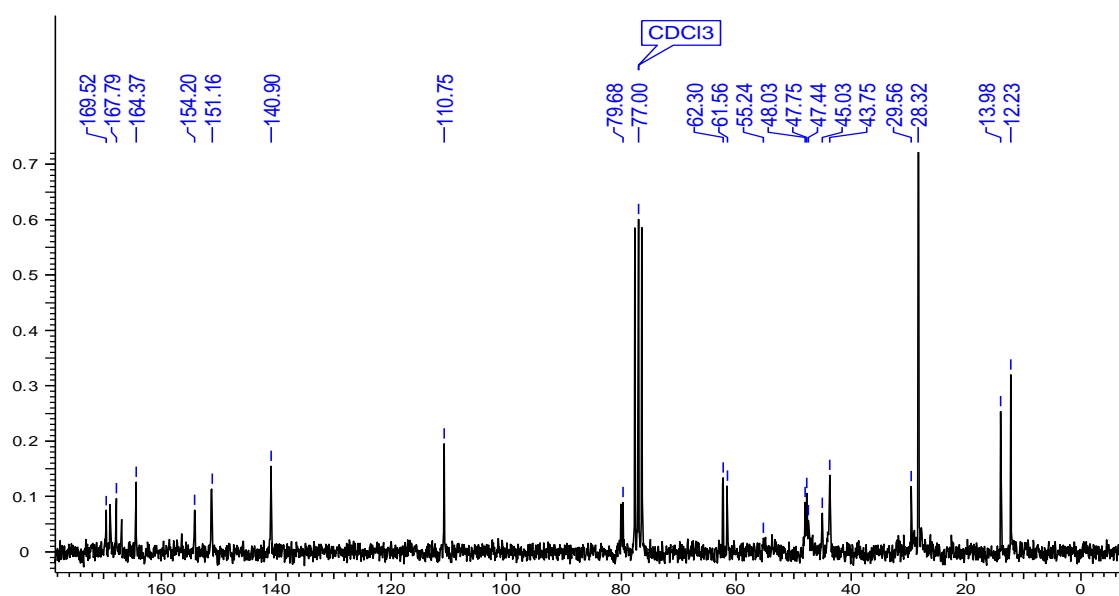
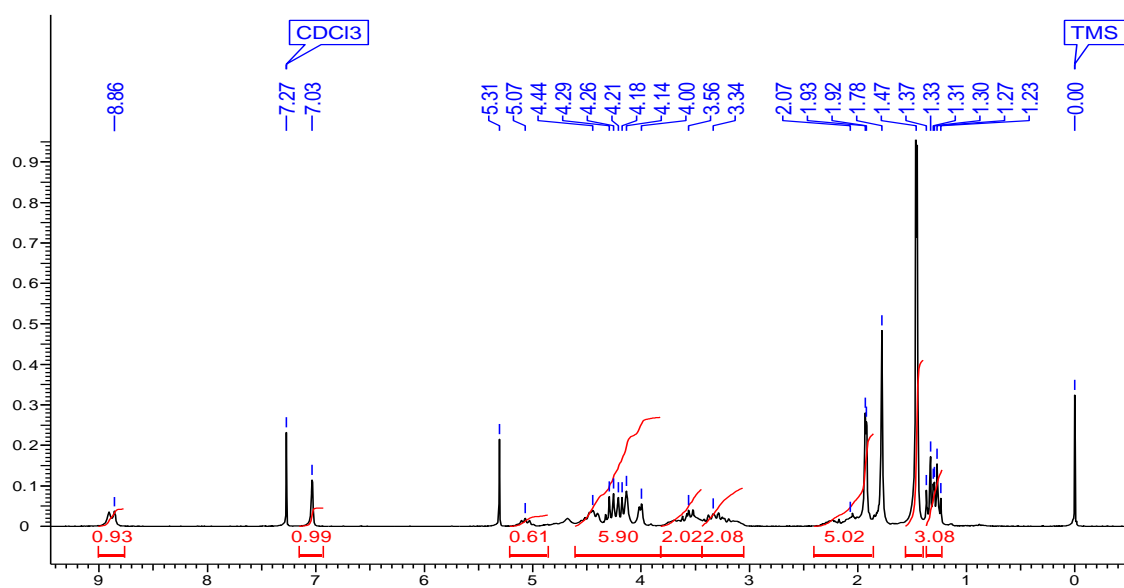
MBHA resin (5 mg) attached to the Fmoc protected Am-ethano-PNA was suspended in dry DMF. Fmoc protected amino group was deprotected by 20% piperidine-DMF solution (3 X 2 mL, 30 min each) and resin was then subjected for guanidinylation by suspending in the solution of 1H-pyrazole-1-carboxamide hydrochloride (10 equivalents)/DIPEA in DMF-THF (1: 2) mixture. After 6–7 h, the resin was washed with dry DMF and treated with the same guanidinylation mixture once again to ensure completion conversion from amino to guanidino functionality. The resin was washed with dry DMF (3 X 2 mL) followed by dry CH₂Cl₂ (3 X 2 mL) and was then dessicated. Gu-ethano-PNA was cleaved from the resin with standard conditions employing TFMSA-TFA and purified by reverse phase HPLC (yield of conversion of amino to guanidino is approx. 95%).

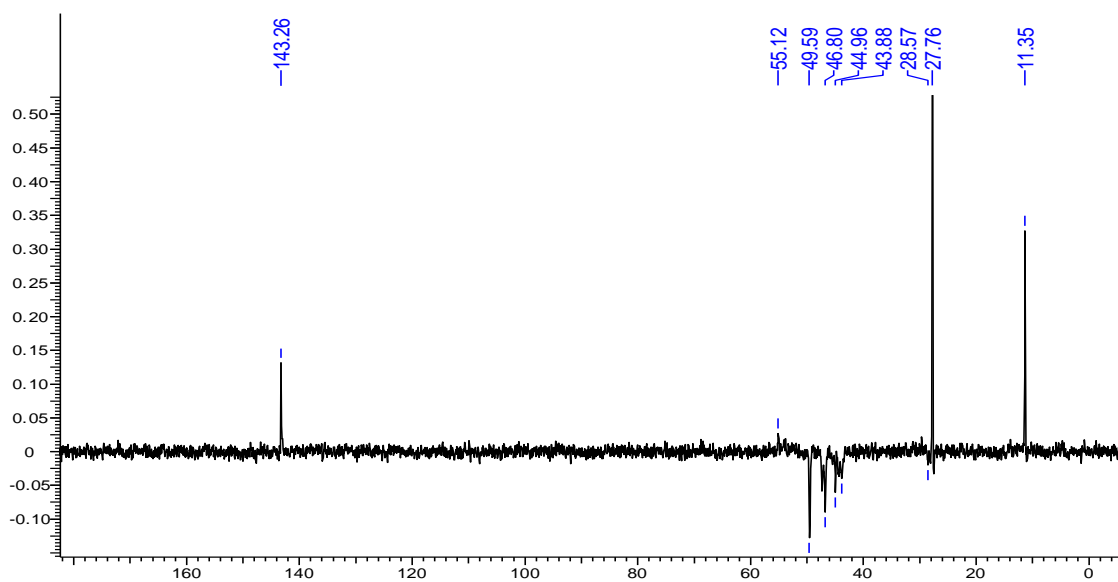
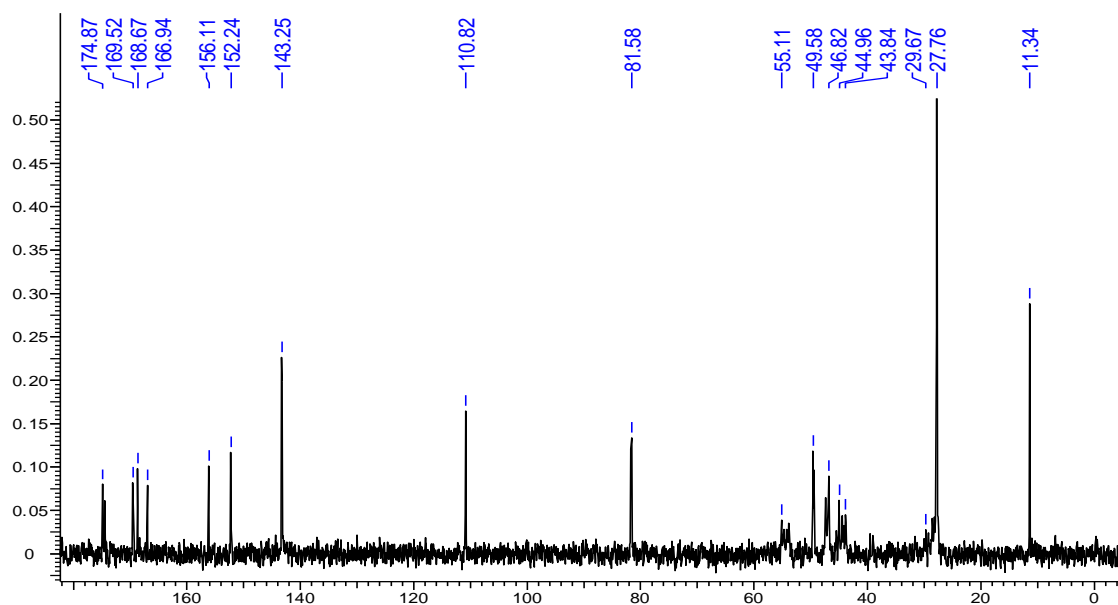
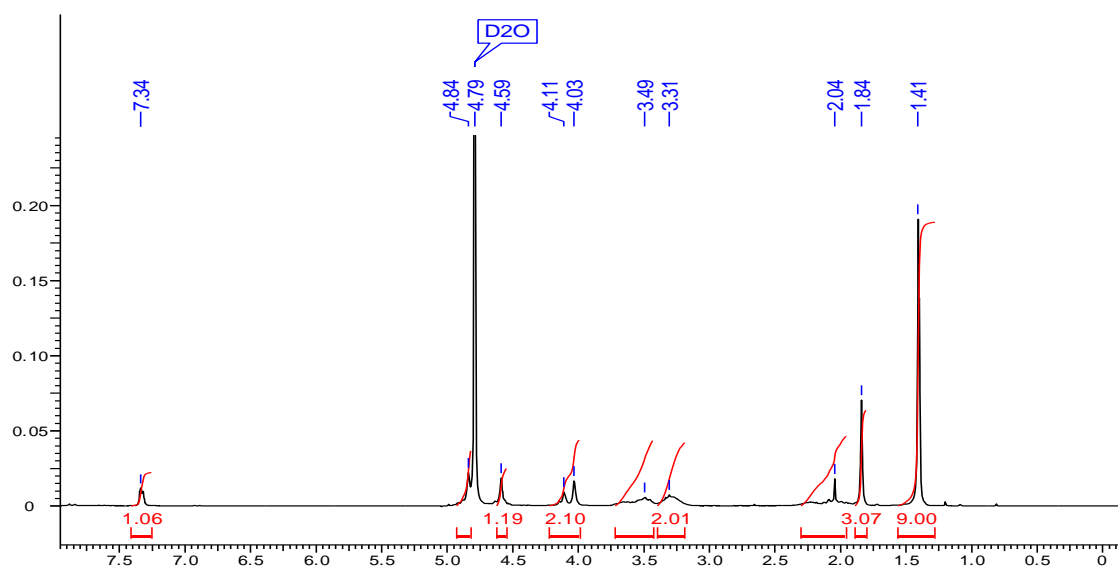
2A.6 Appendix

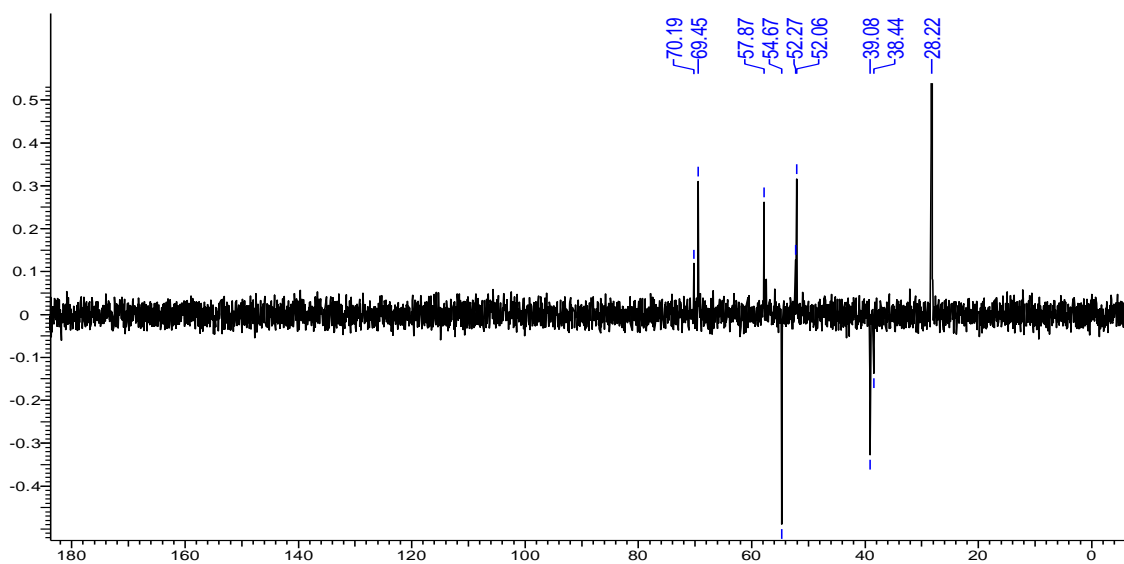
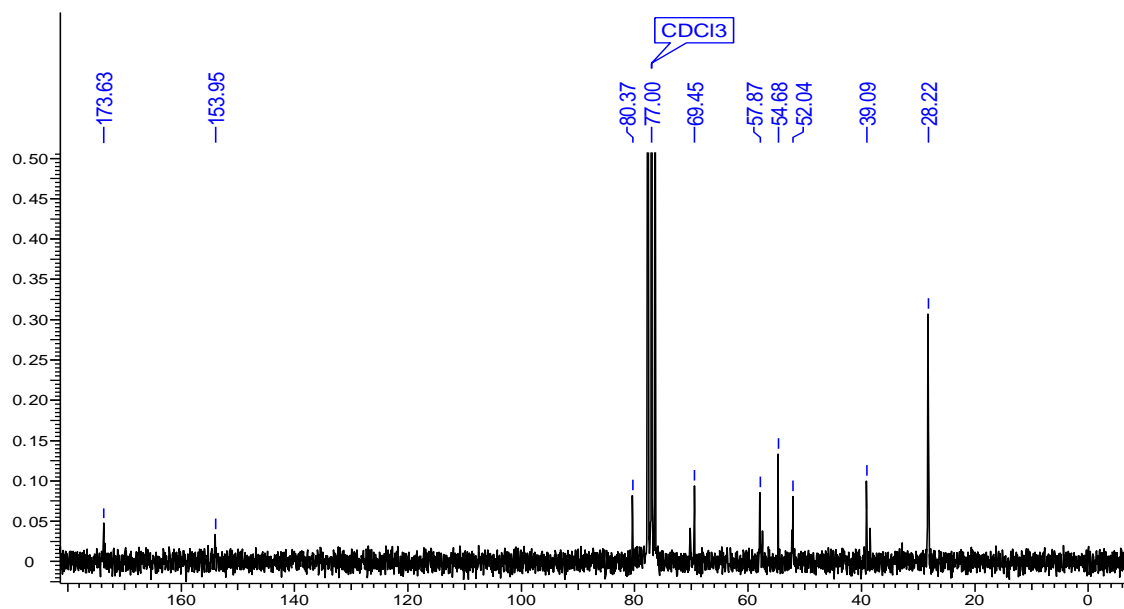
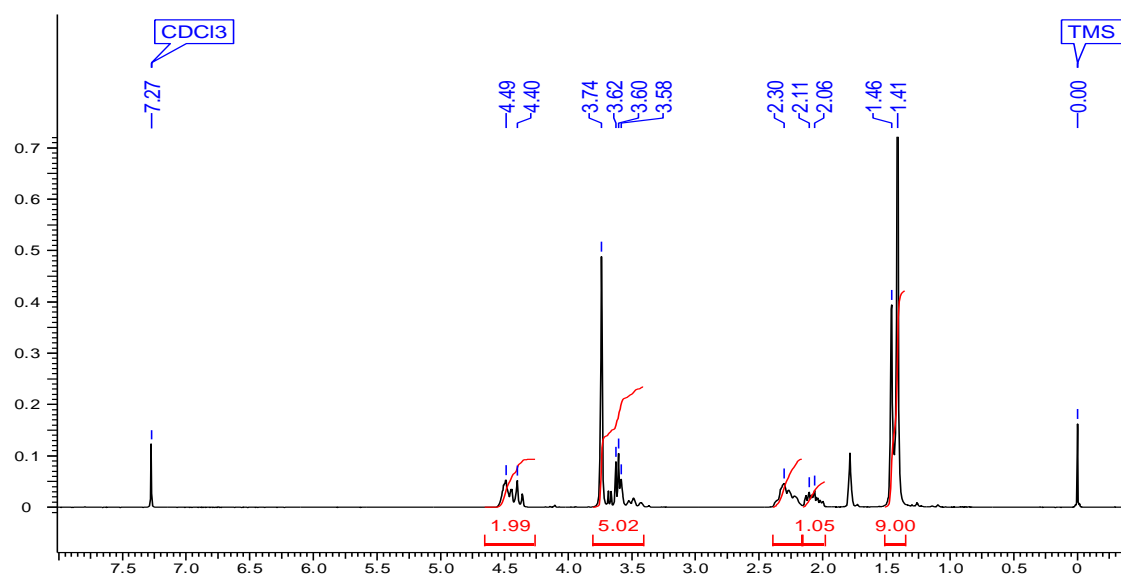
Compounds - Spectral data	Page Nos.
¹ H, ¹³ C and DEPT of compound 7	77
¹ H, ¹³ C and DEPT of compound 8	78
¹ H, ¹³ C and DEPT of compound 9	79
¹ H, ¹³ C and DEPT of compound 10	80
¹ H, ¹³ C and DEPT of compound 11	81
¹ H, ¹³ C and DEPT of compound 12	82
¹ H, ¹³ C and DEPT of compound 13	83
¹ H, ¹³ C and DEPT of compound 14	84
¹ H, ¹³ C and DEPT of compound 15	85
¹ H, ¹³ C and DEPT of compound 16	86
¹ H, ¹³ C and DEPT of compound 17	87
¹ H, ¹³ C and DEPT of compound 18	88
¹ H, ¹³ C and DEPT of compound 19	89
¹ H, ¹³ C and DEPT of compound 20	90
¹ H, ¹³ C and DEPT of compound 21	91
¹ H, ¹³ C and DEPT of compound 22	92
¹ H, ¹³ C and DEPT of compound 24	93
¹ H and ¹³ C of compound 25	94
HRMS of compounds 7 and 8	95
HRMS of compounds 9 and 10	96
HRMS of compounds 11 and 12	97
HRMS of compounds 13 and 14	98
HRMS of compounds 15 and 16	99
HRMS of compounds 17 and 18	100
HRMS of compounds 19 and 20	101
HRMS of compounds 21 and 22	102
HRMS of compounds 23 and 24	103
HRMS of compound 25	104
HPLC and MALDI-TOF of PNA 1	105
HPLC and MALDI-TOF of PNA 2	106
HPLC and MALDI-TOF of PNA 3	107
HPLC and MALDI-TOF of PNA 4	108
HPLC and MALDI-TOF of PNA 5	109
HPLC and MALDI-TOF of PNA 6	110
HPLC and MALDI-TOF of PNA 7	111
HPLC of PNA 3 + PNA 4	112
HPLC of PNA 6 + PNA 7	112

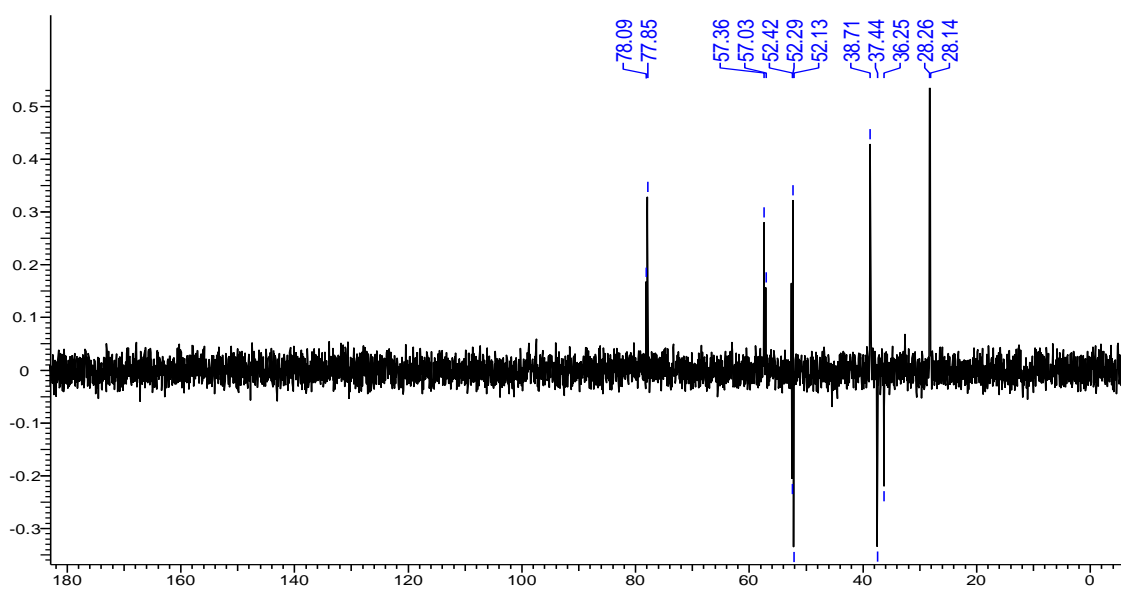
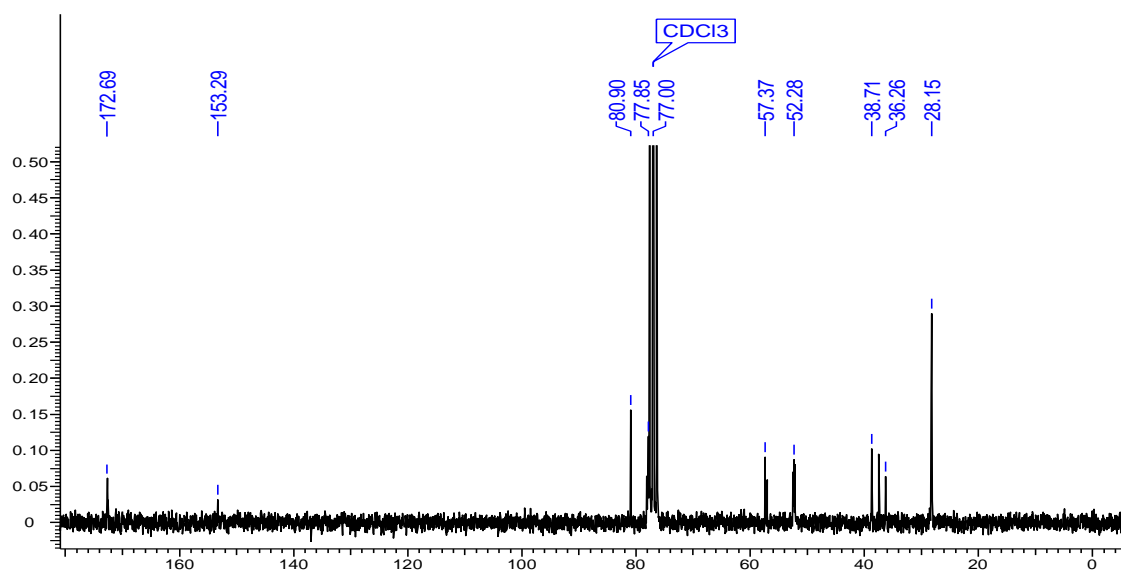
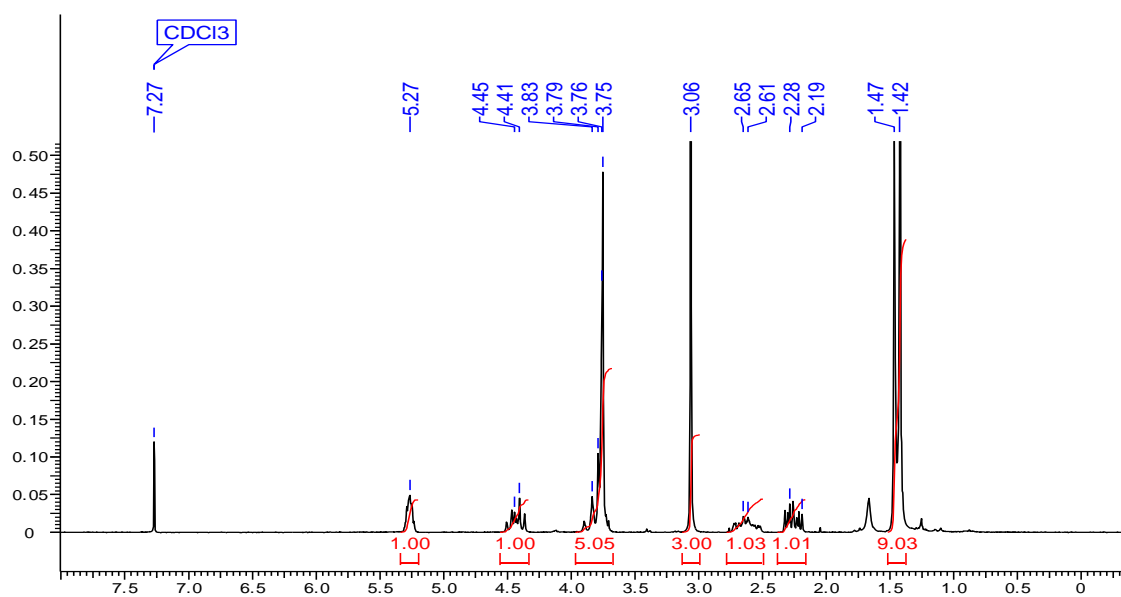
^1H , ^{13}C and DEPT of compound 7

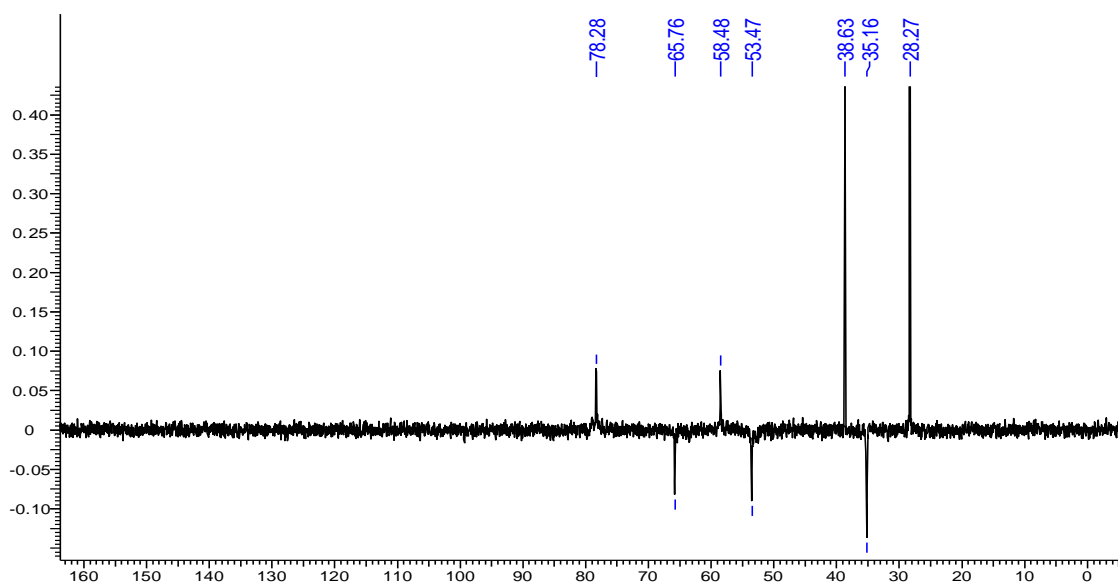
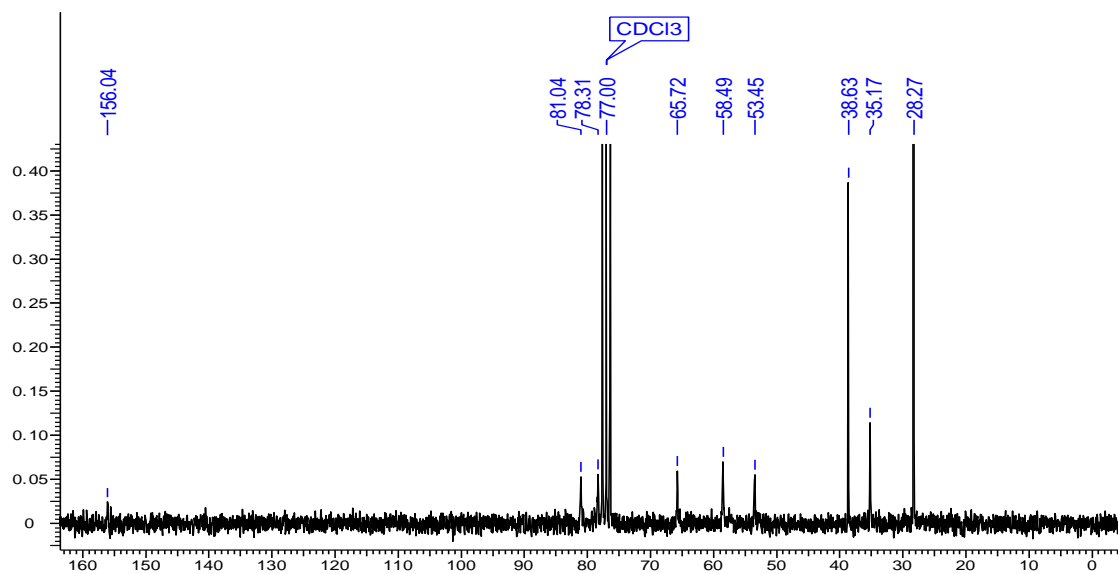
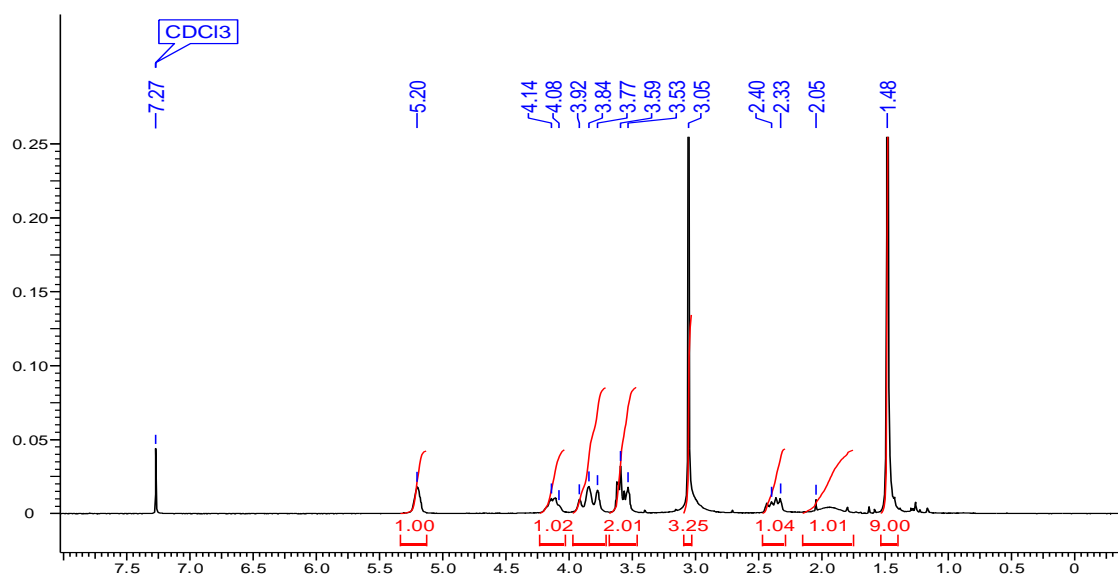
^1H , ^{13}C and DEPT of compound **8**

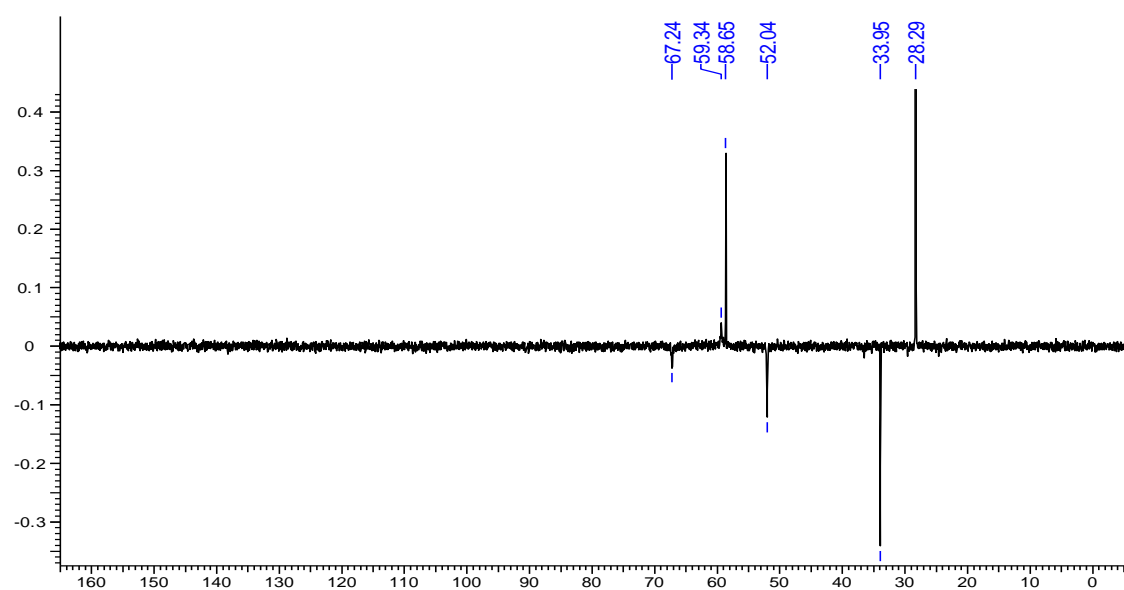
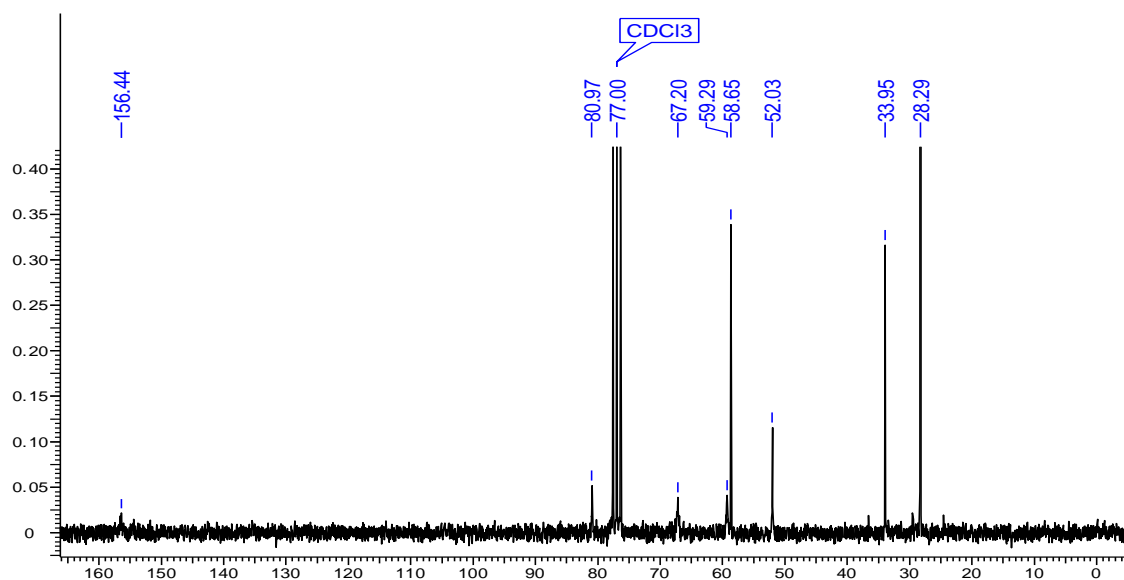
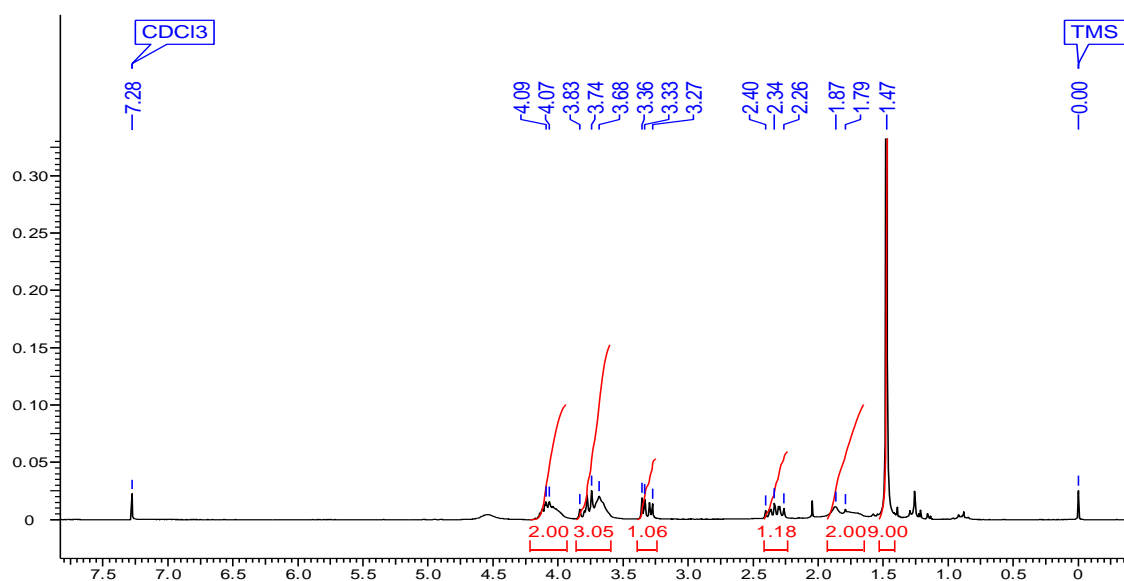
^1H , ^{13}C and DEPT of compound **9**

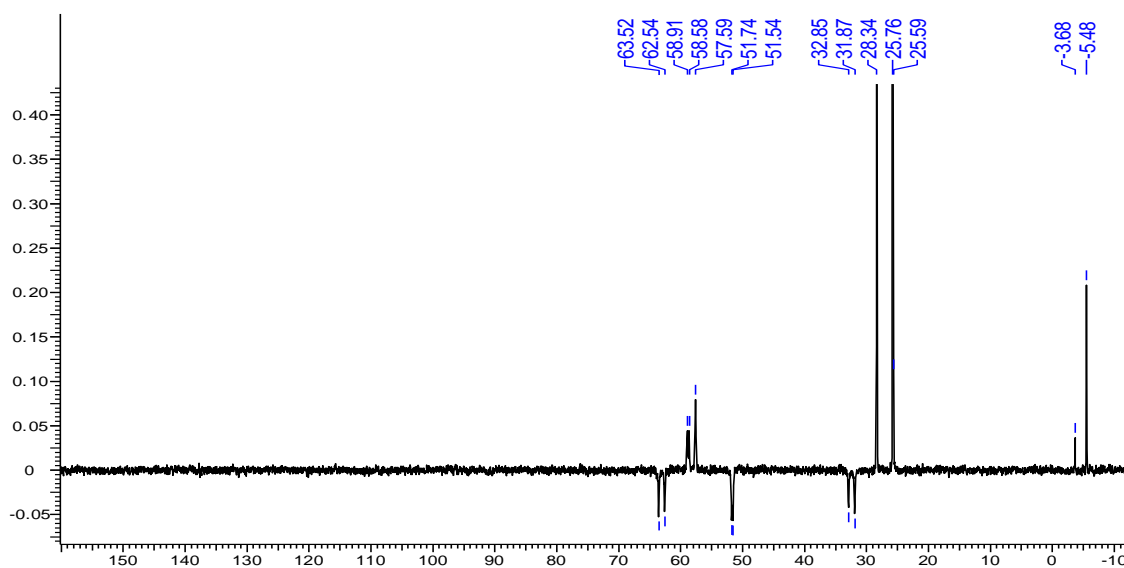
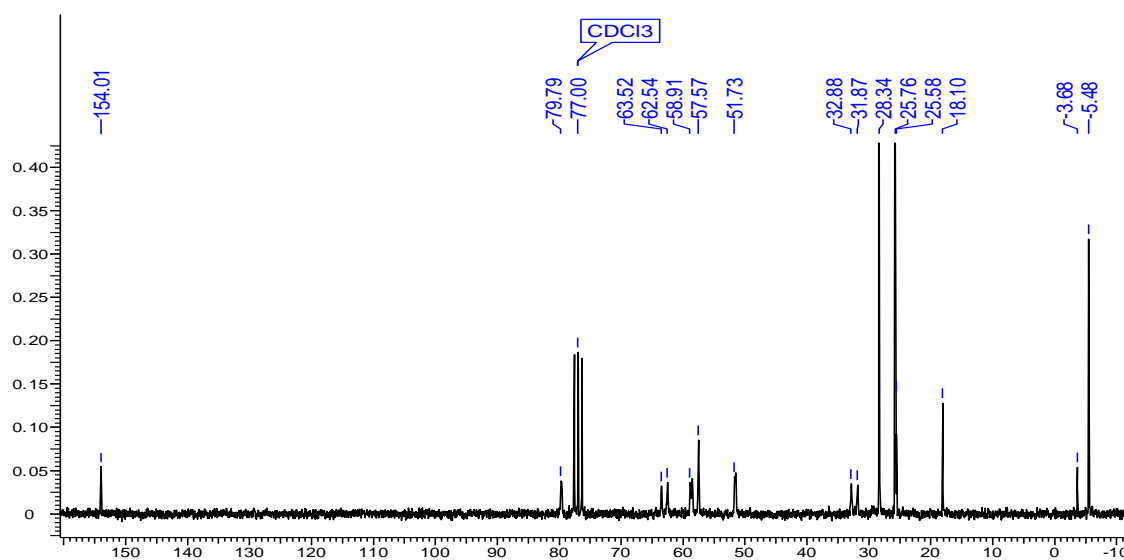
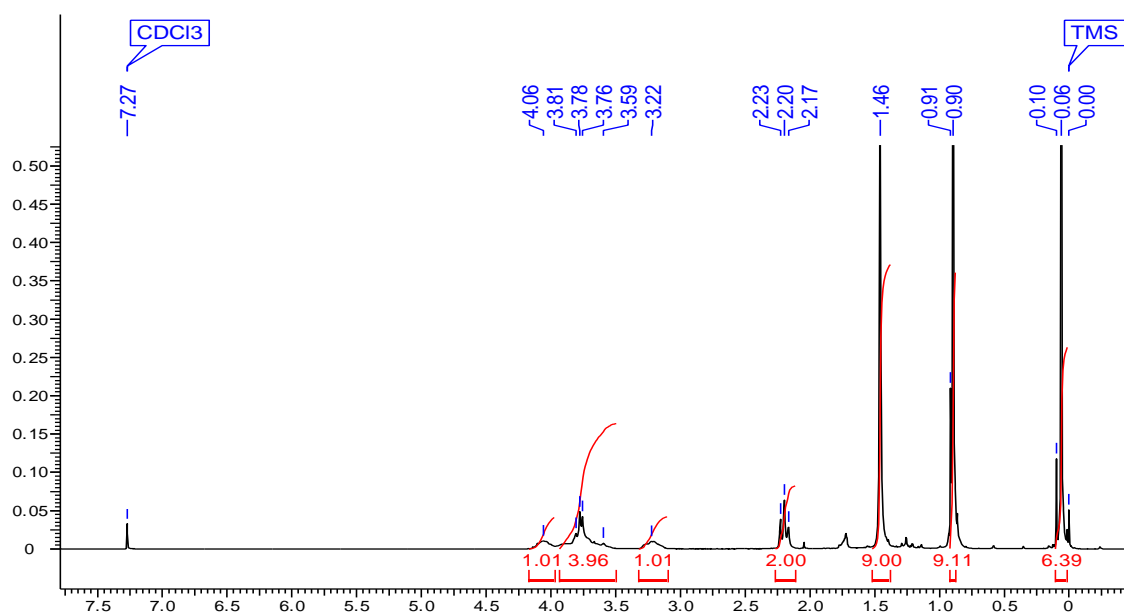
^1H , ^{13}C and DEPT of compound **10**

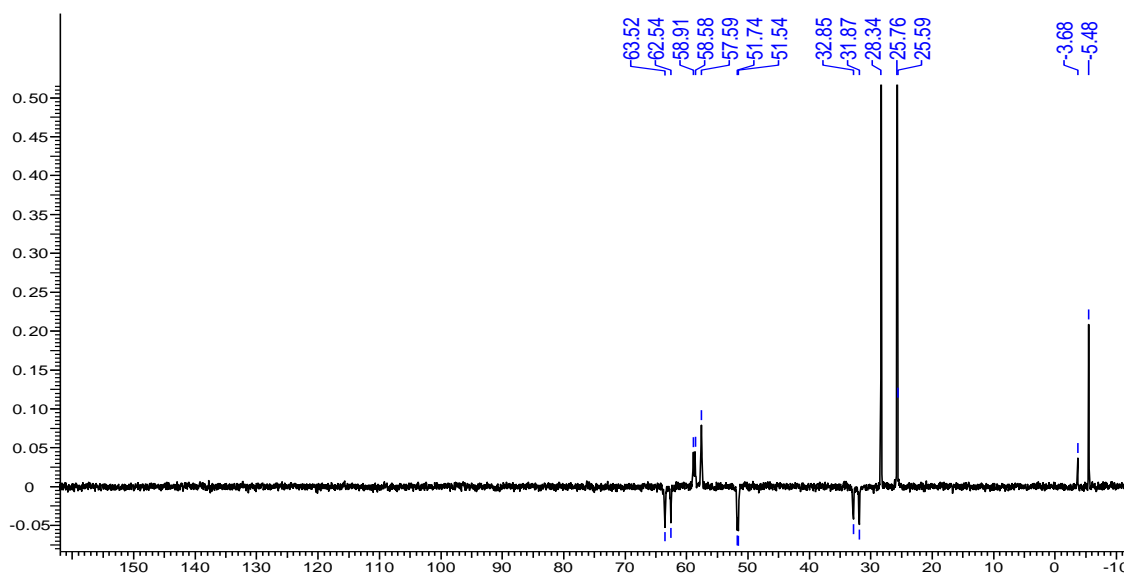
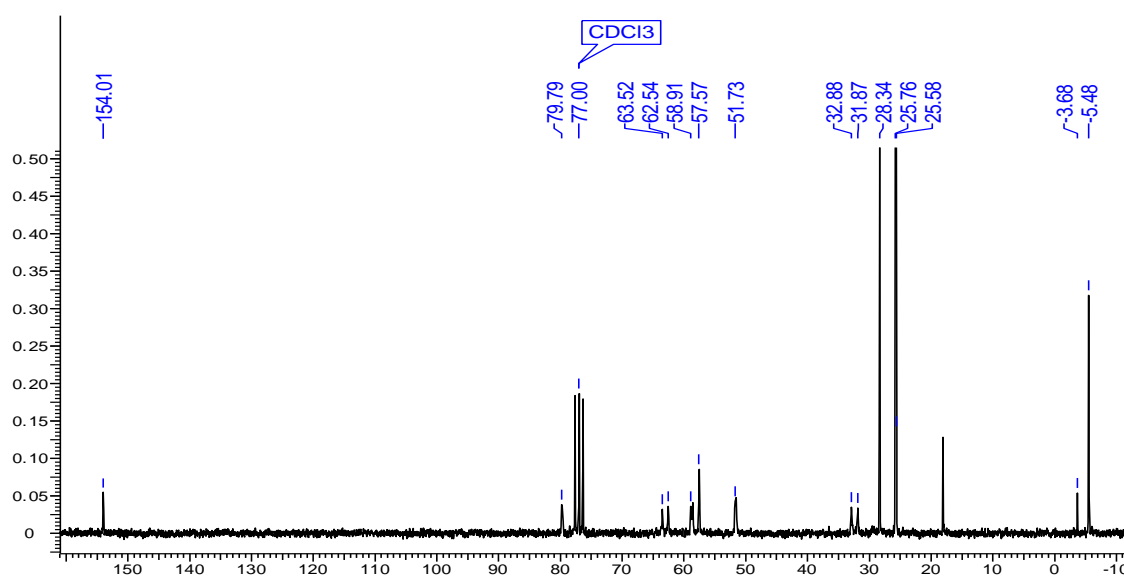
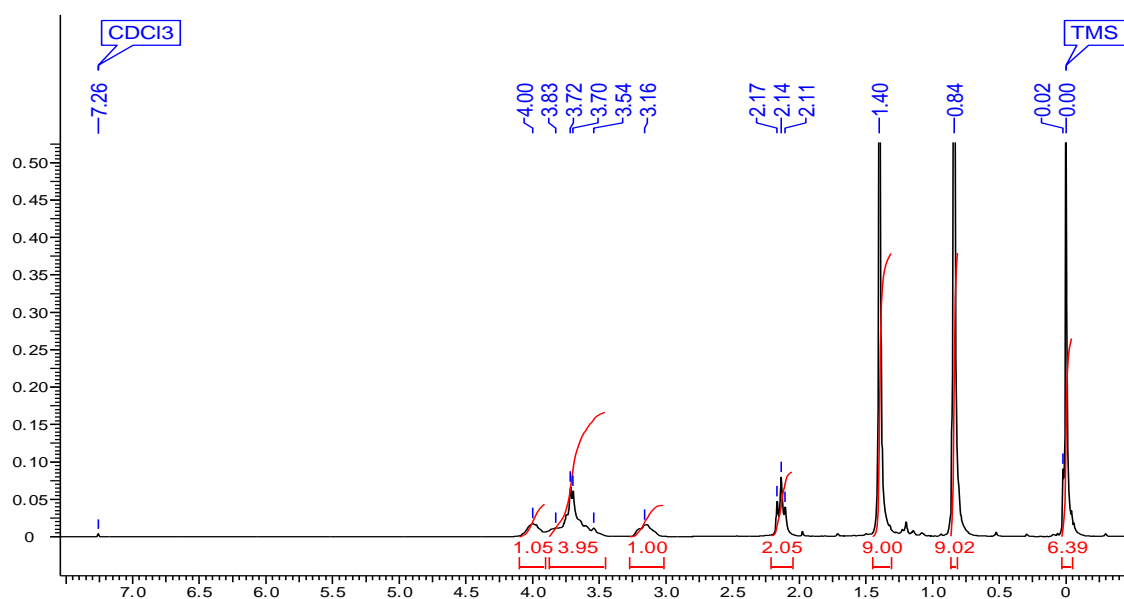
^1H , ^{13}C and DEPT of compound **11**

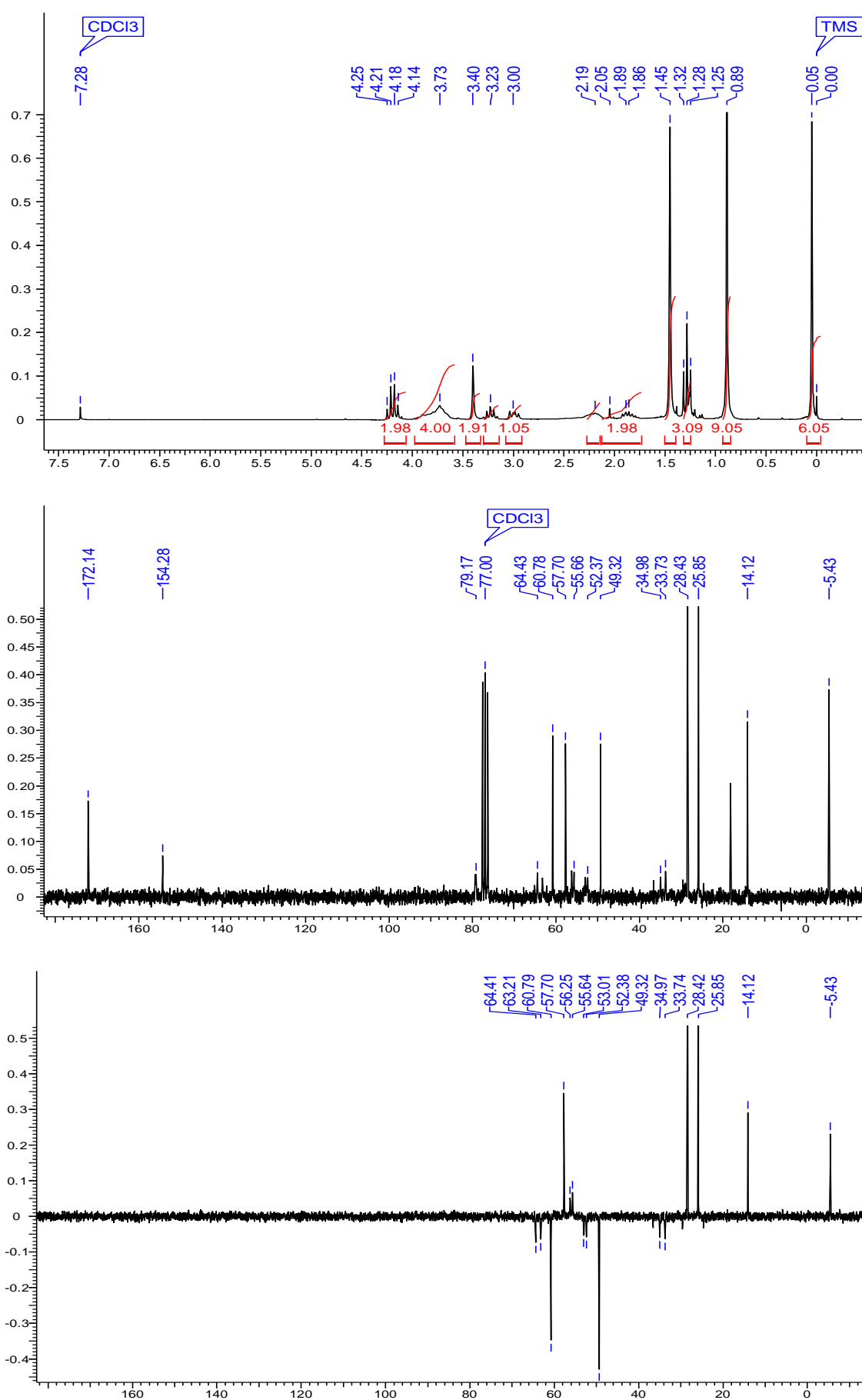
^1H , ^{13}C and DEPT of compound **12**

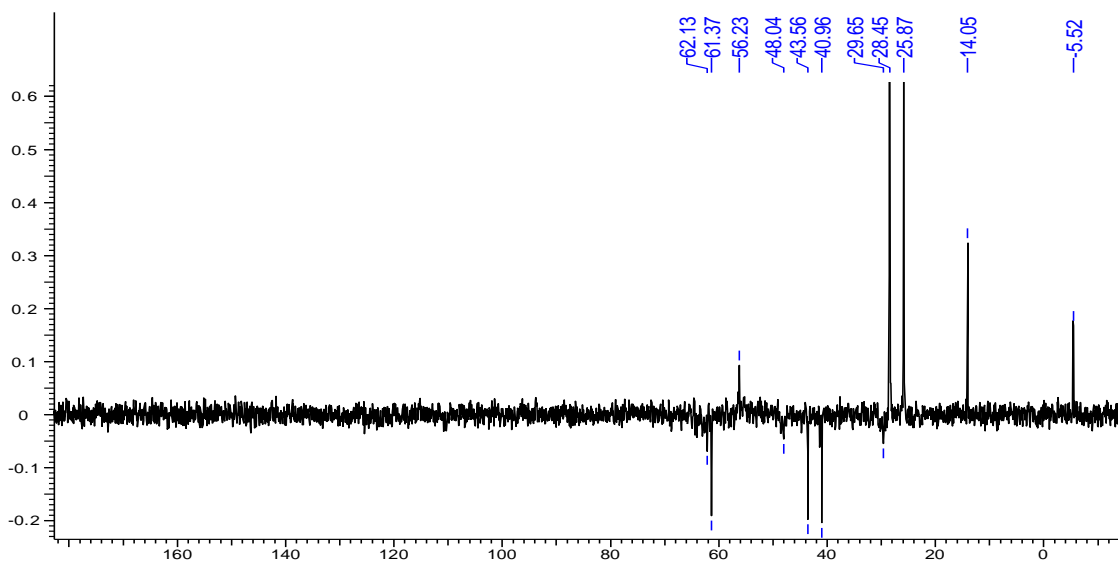
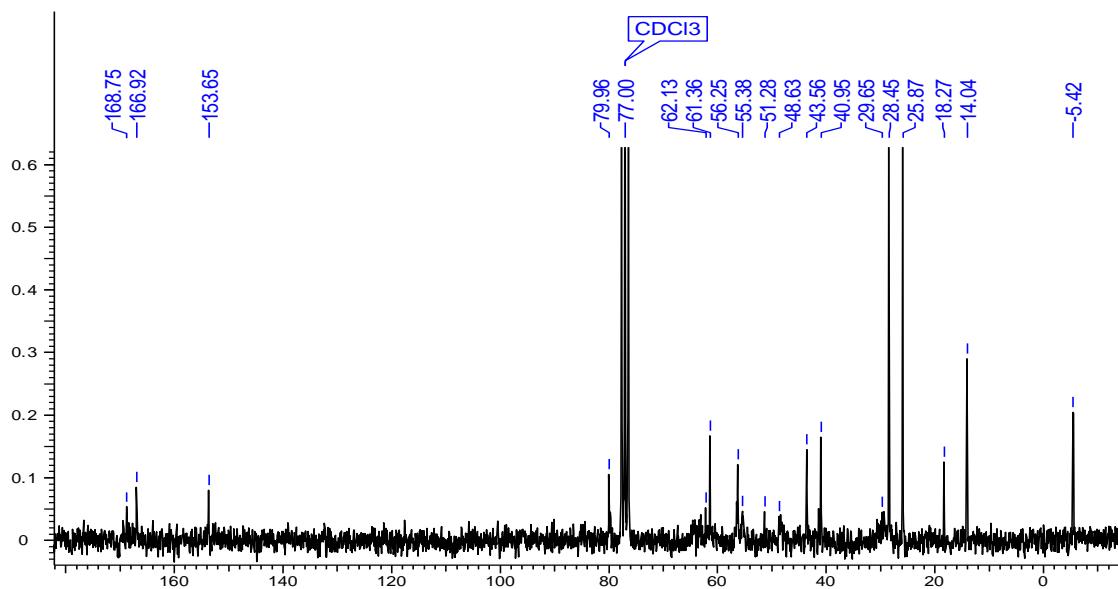
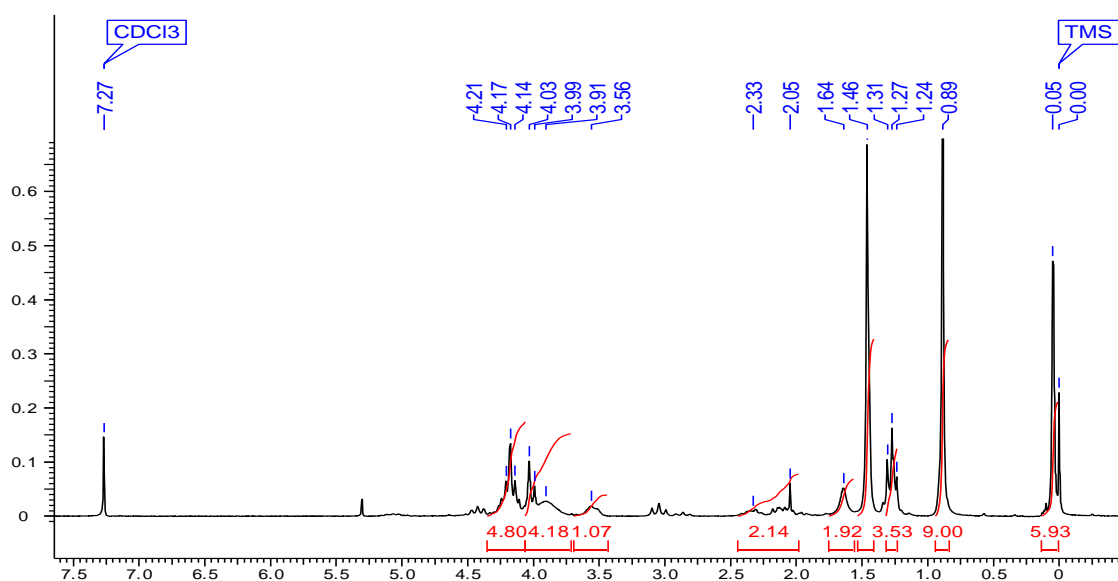
^1H , ^{13}C and DEPT of compound **13**

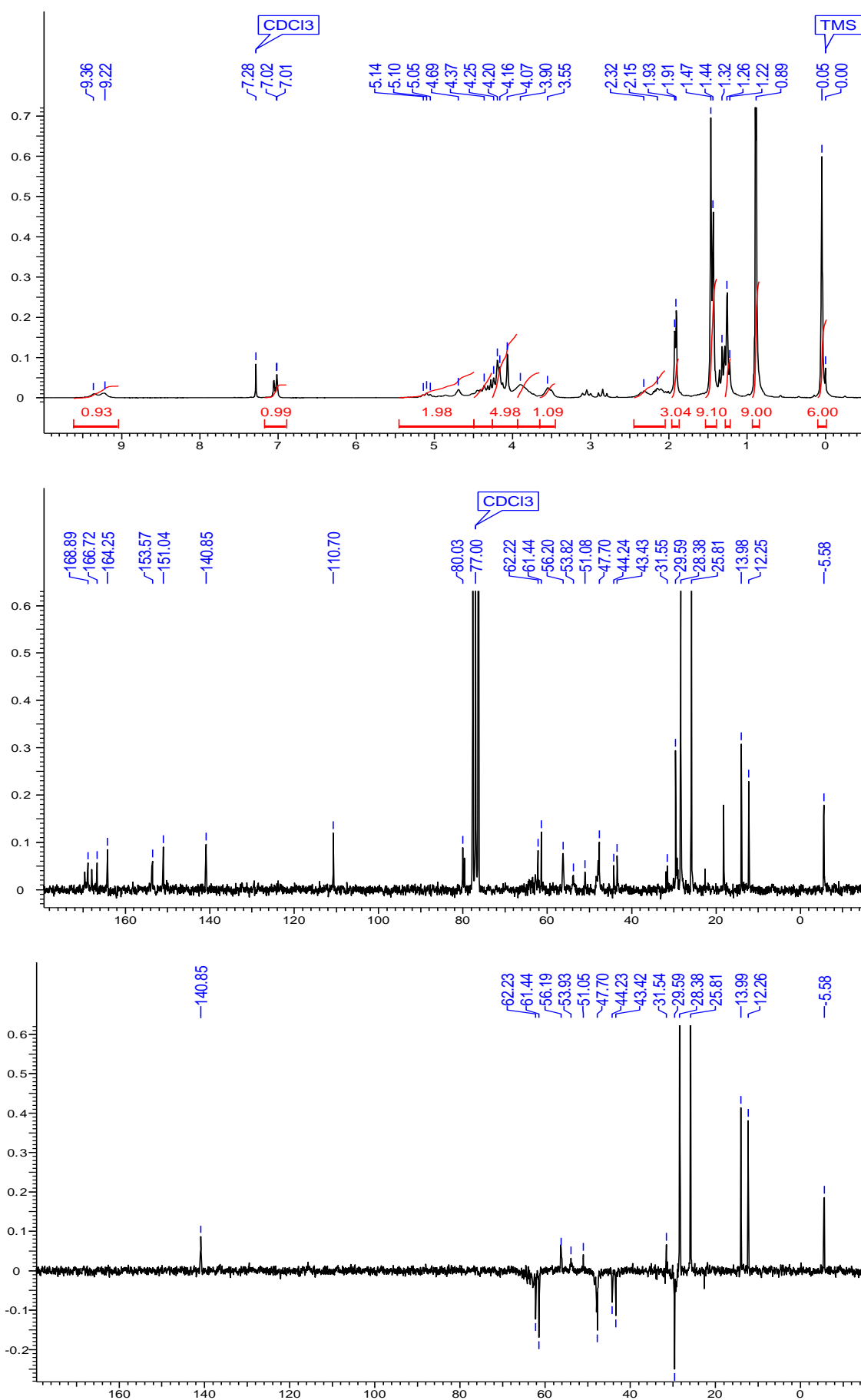
^1H , ^{13}C and DEPT of compound 14

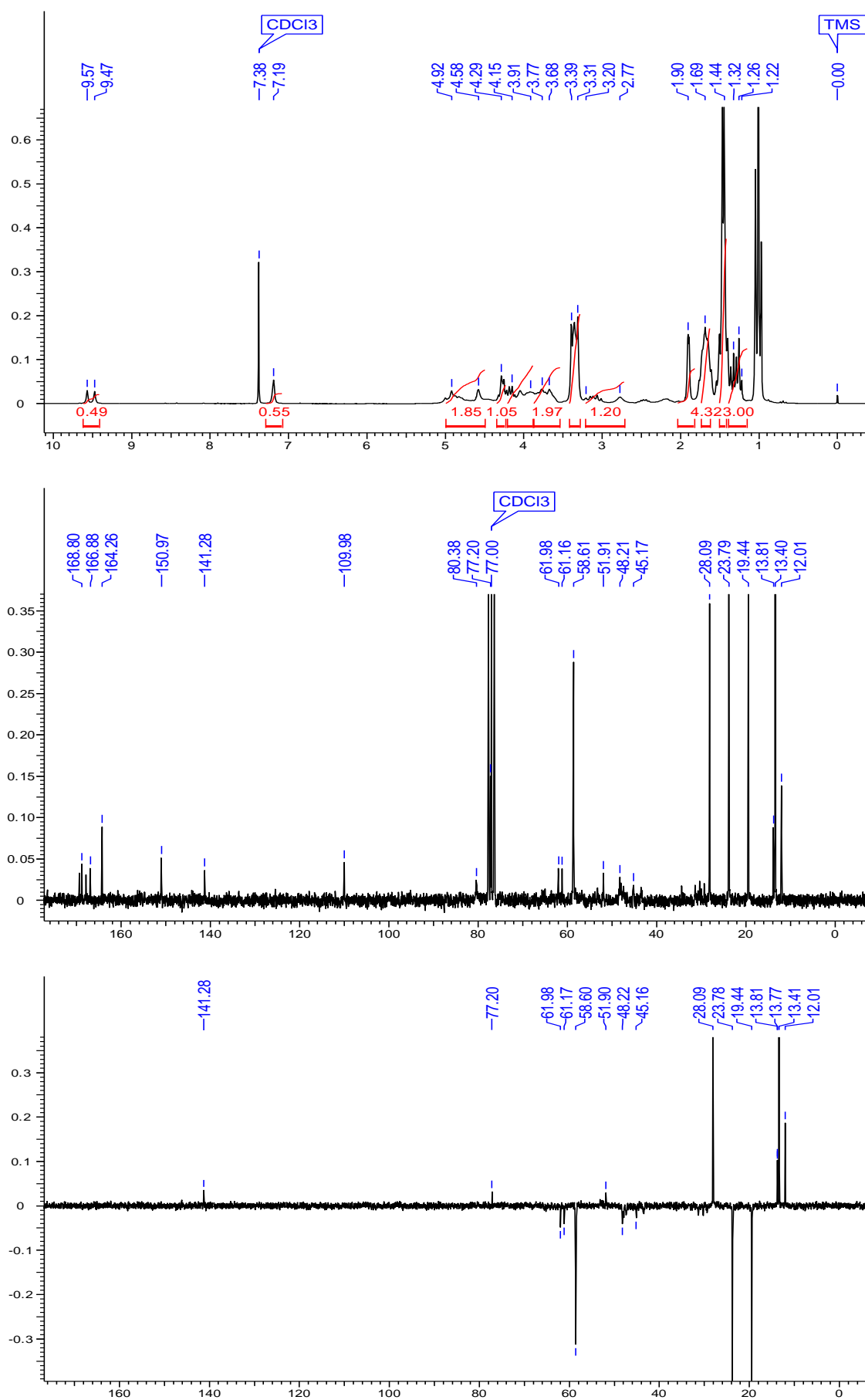
^1H , ^{13}C and DEPT of compound **15**

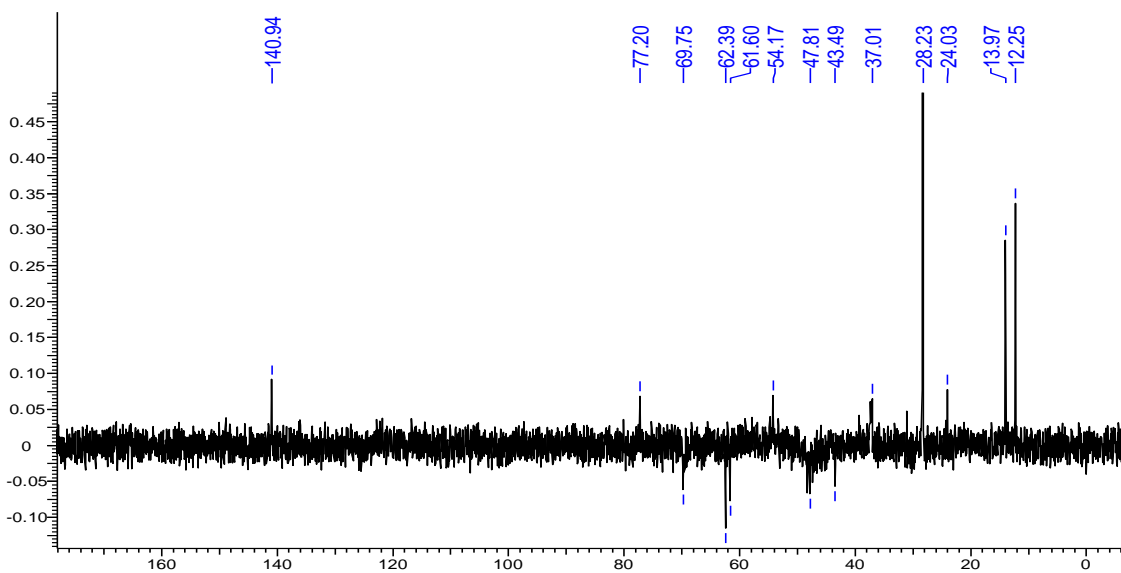
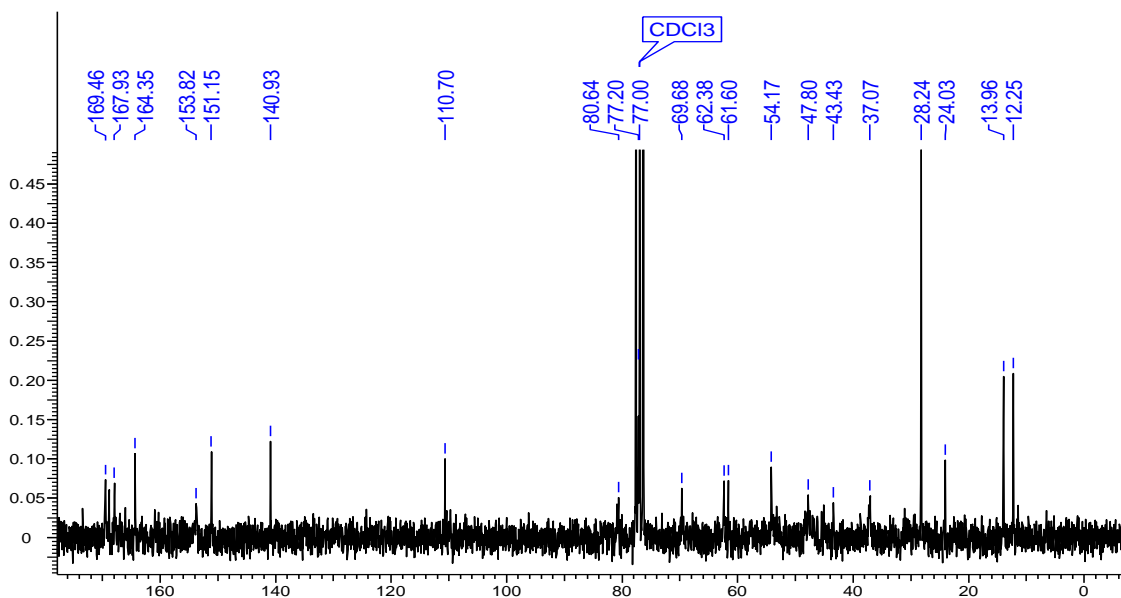
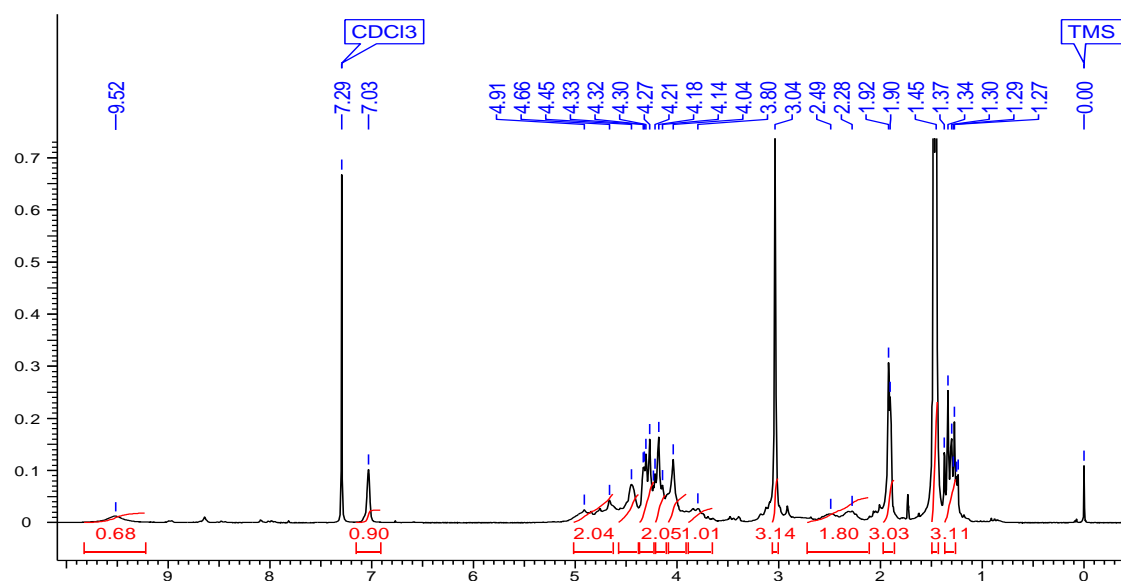
^1H , ^{13}C and DEPT of compound 16

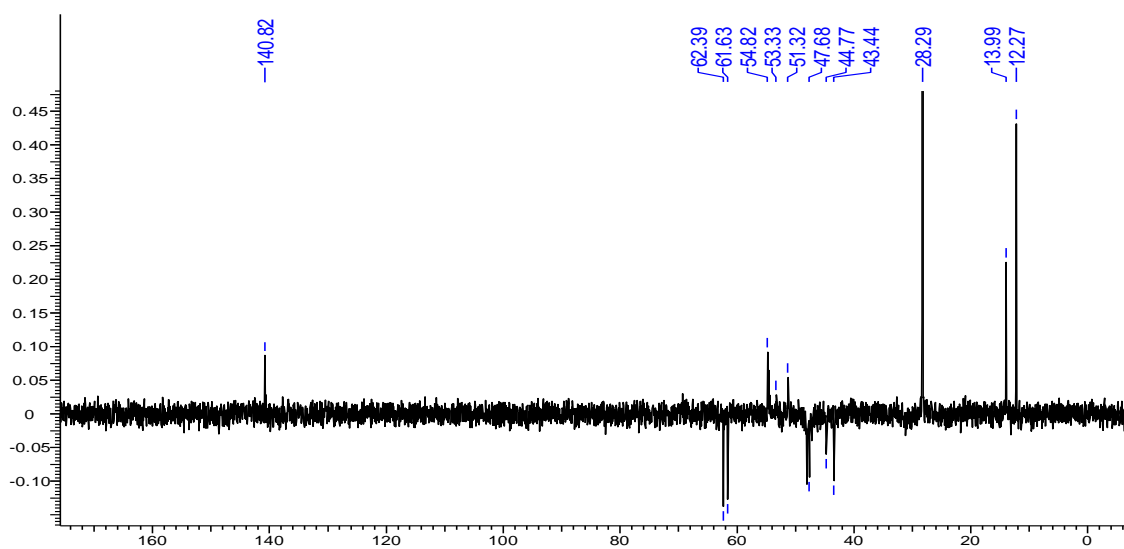
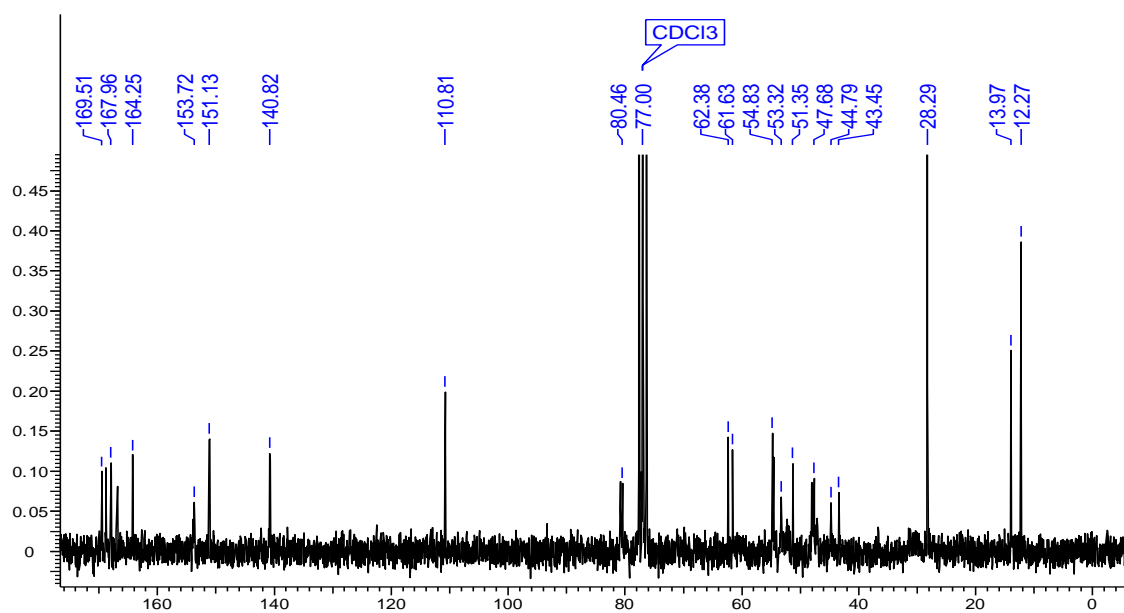
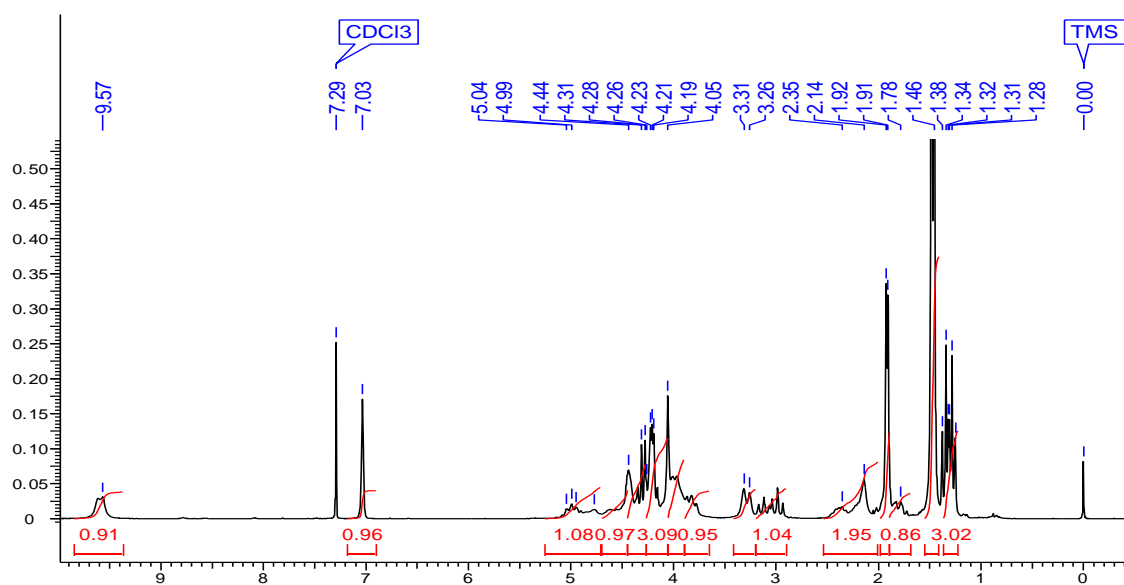
^1H , ^{13}C and DEPT of compound 17

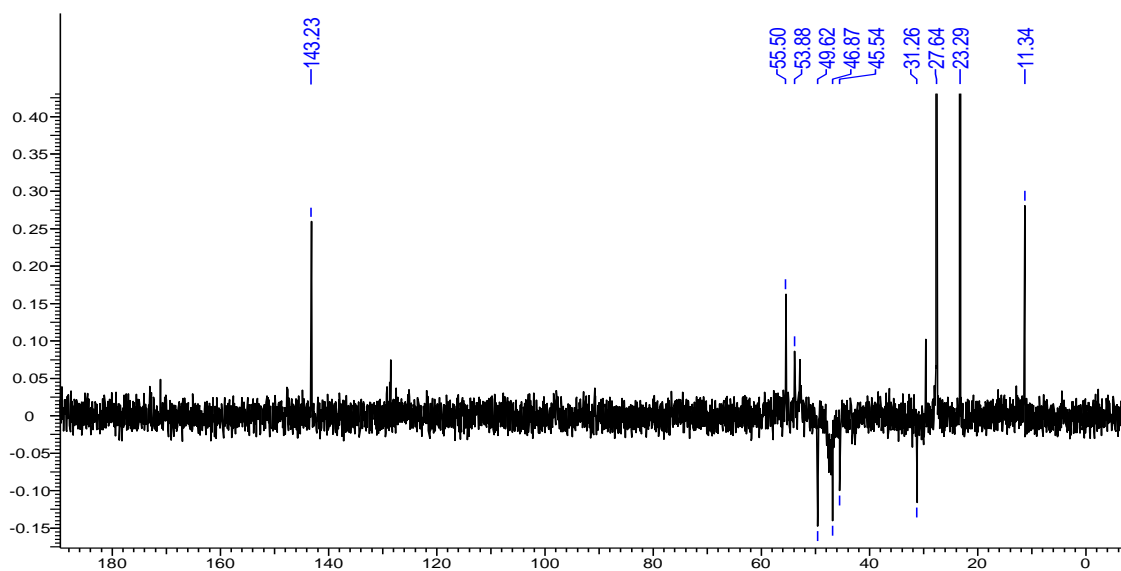
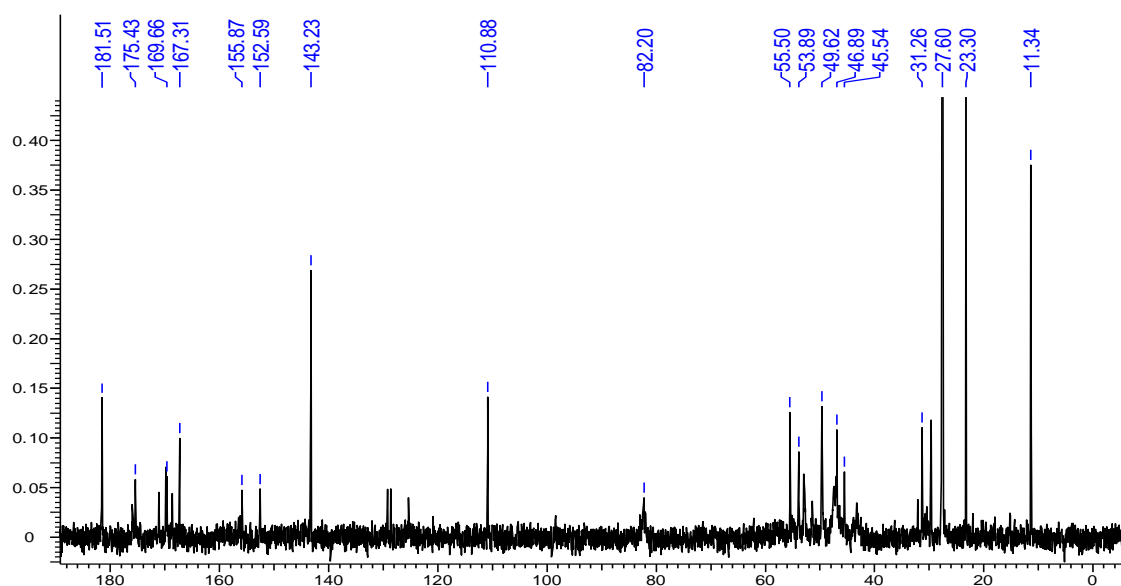
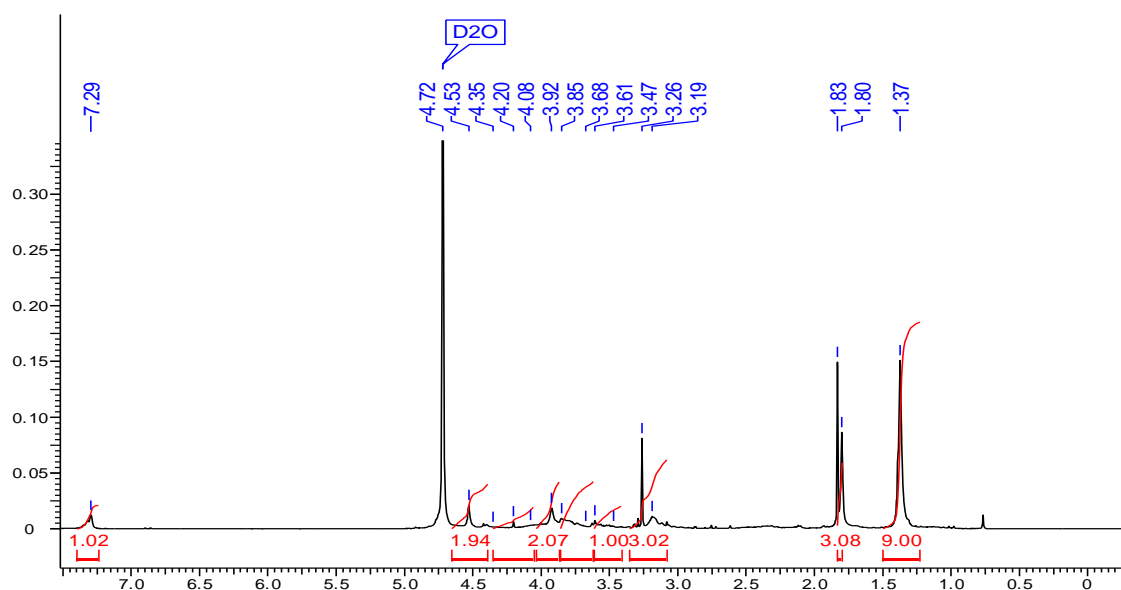
^1H , ^{13}C and DEPT of compound **18**

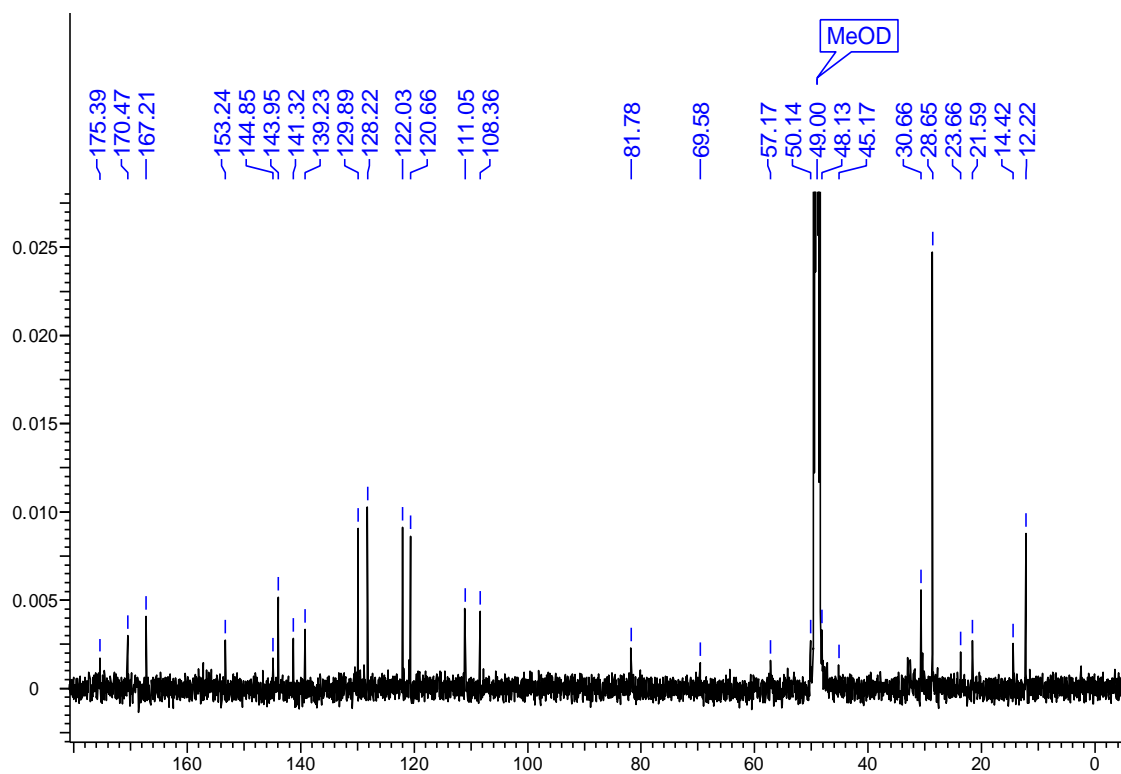
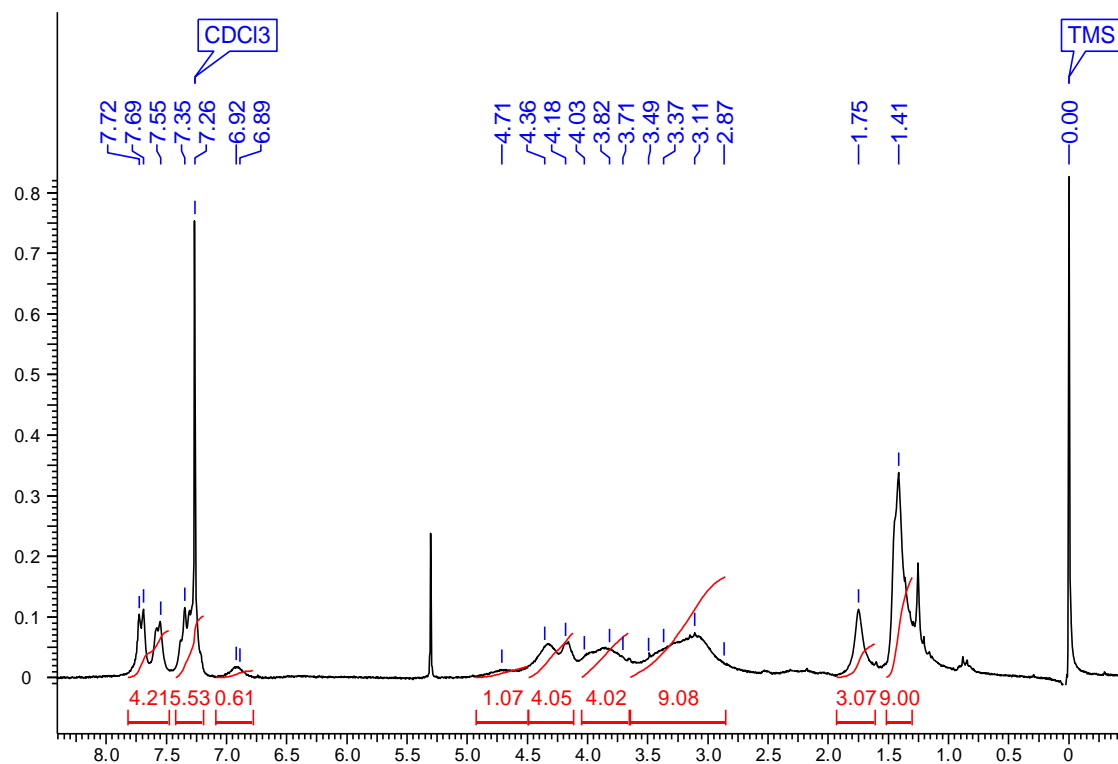
^1H , ^{13}C and DEPT of compound **19**

^1H , ^{13}C and DEPT of compound **20**

^1H , ^{13}C and DEPT of compound **21**

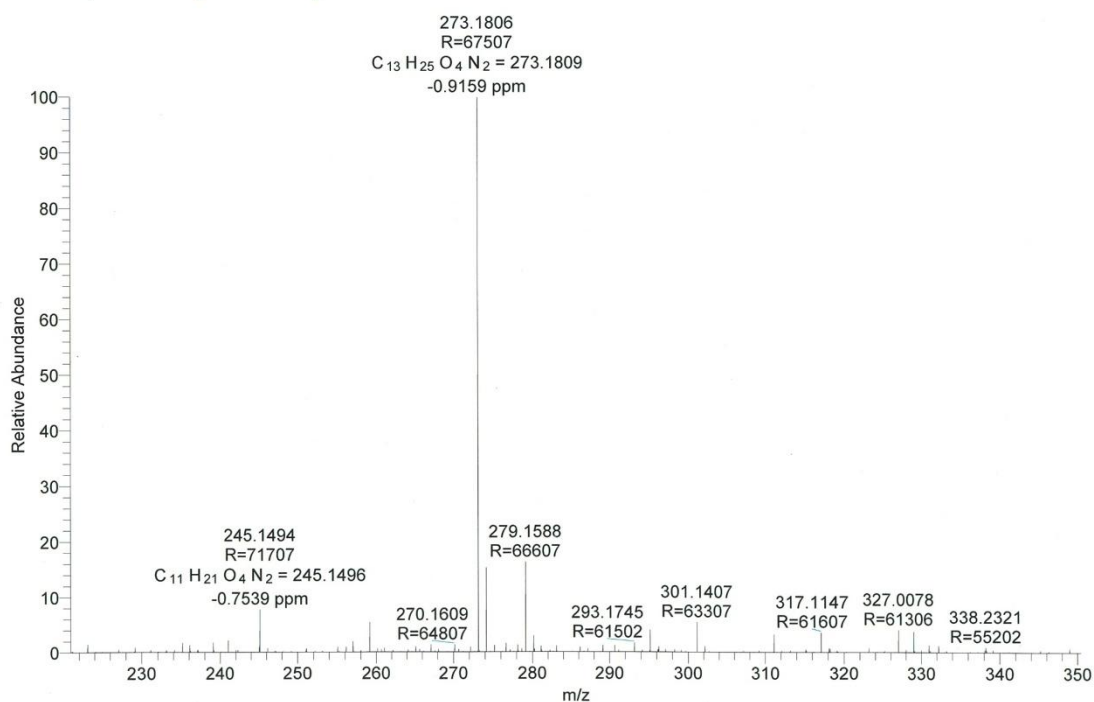
^1H , ^{13}C and DEPT of compound **22**

^1H , ^{13}C and DEPT of compound **24**

^1H and ^{13}C of compound **25**

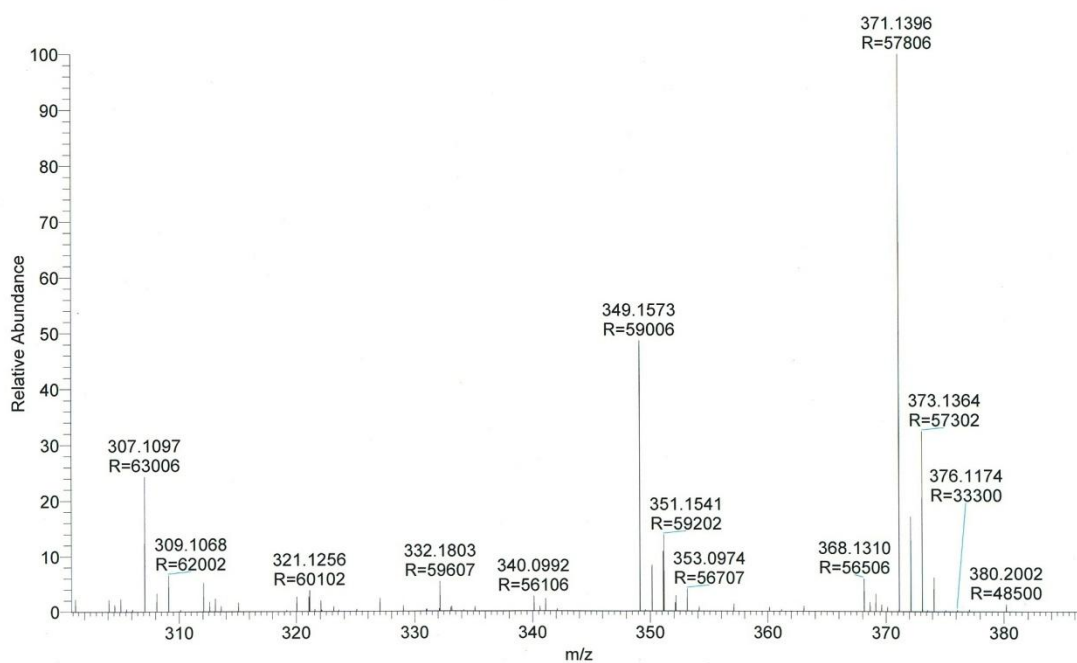
HRMS of compounds 7

ALKYL #184 RT: 0.82 AV: 1 NL: 1.20E8
T: FTMS + p ESI Full ms [100.00-700.00]



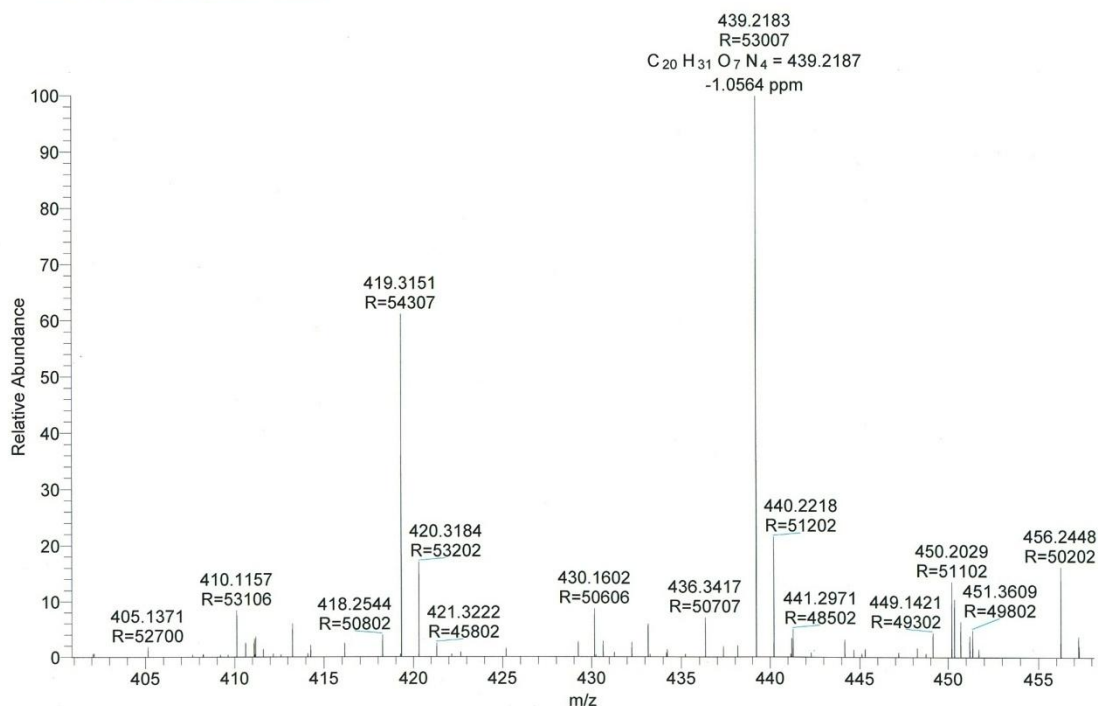
HRMS of compounds 8

ACYL #122 RT: 0.54 AV: 1 NL: 1.18E8
T: FTMS + p ESI Full lock ms [100.00-700.00]

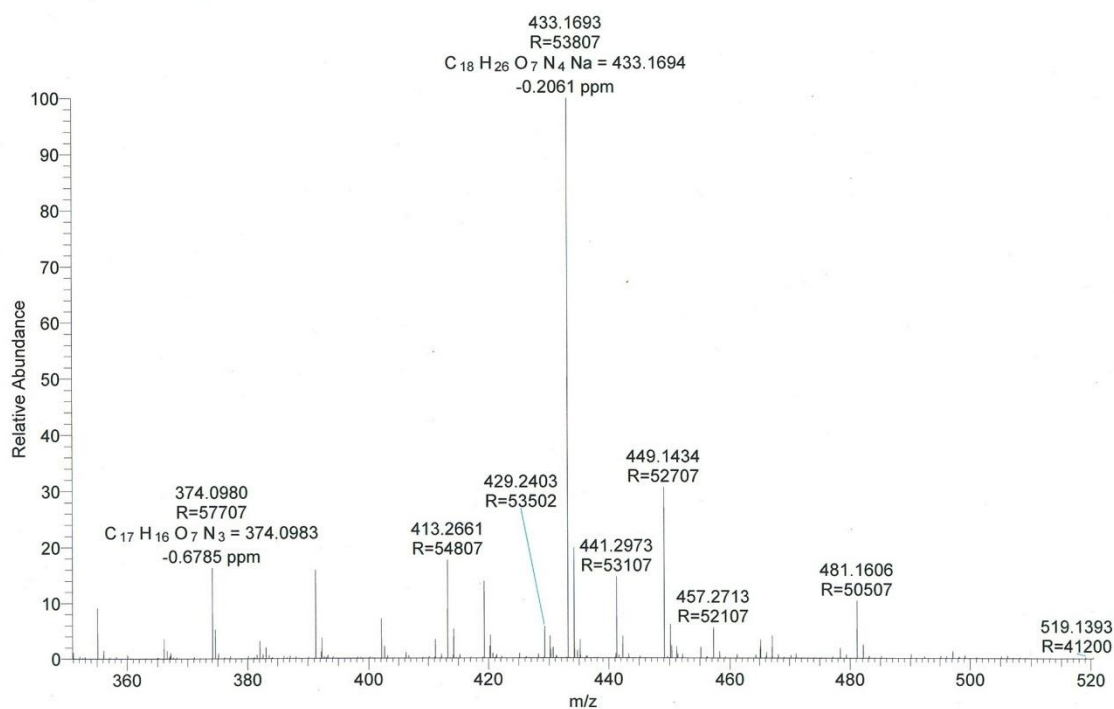


HRMS of compounds **9**

BASET #112 RT: 0.50 AV: 1 NL: 1.63E7
T: FTMS + p ESI Full ms [100.00-700.00]

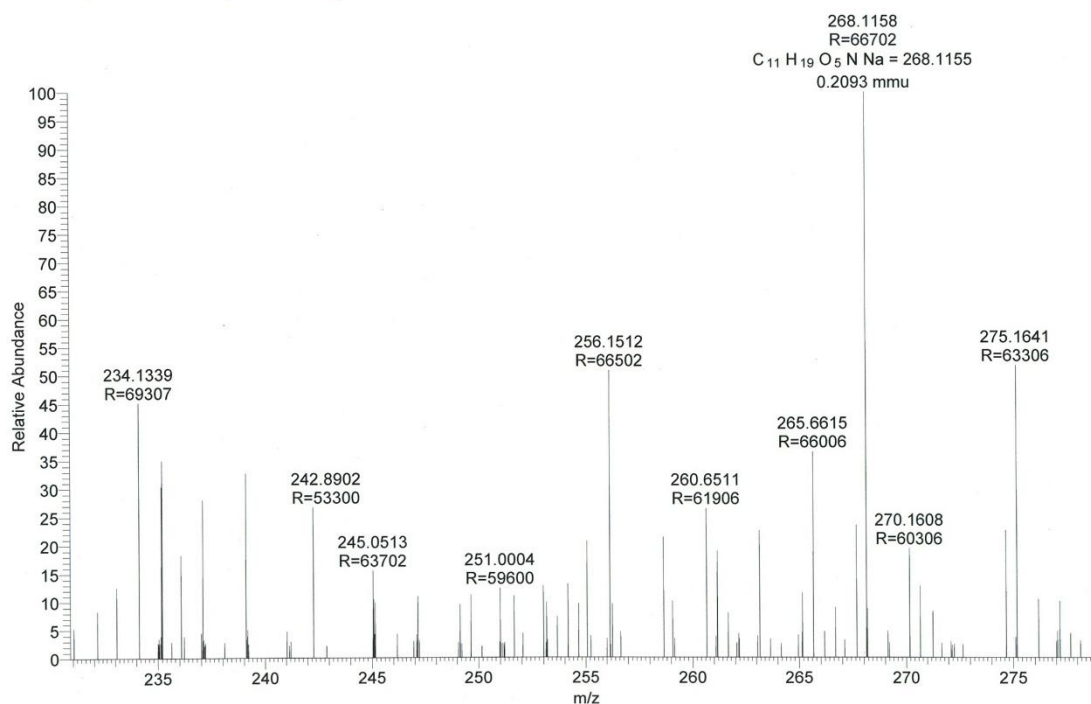
HRMS of compounds **10**

MONOT #106 RT: 0.47 AV: 1 NL: 1.37E8
T: FTMS + p ESI Full ms [100.00-700.00]



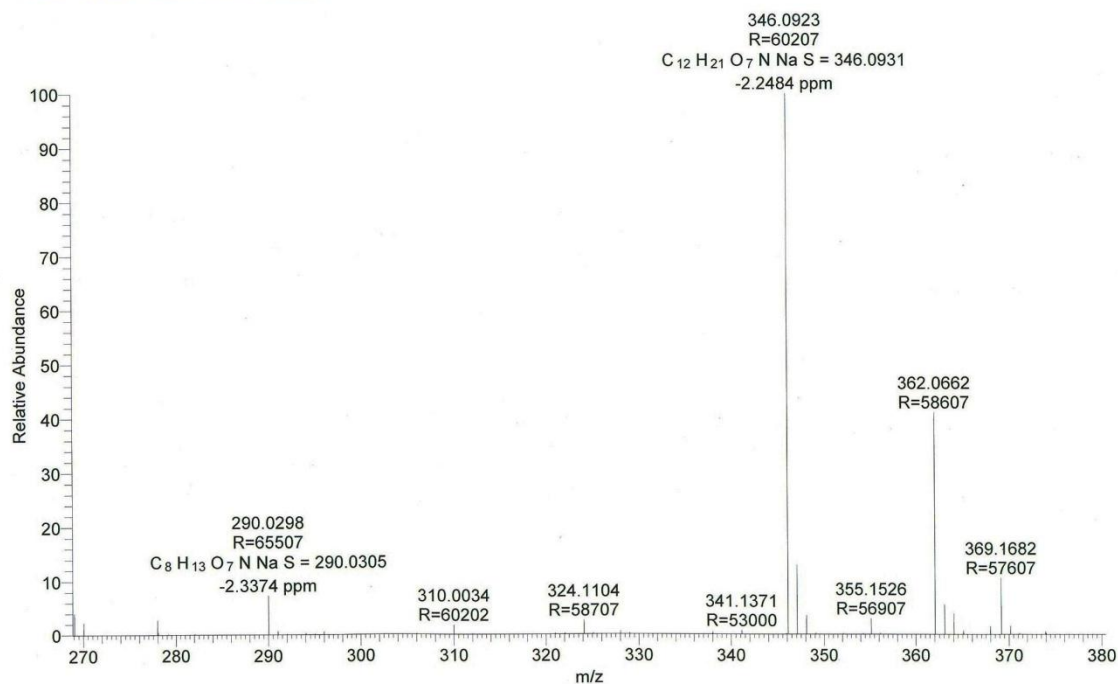
HRMS of compounds 11

AB10 #109 RT: 0.48 AV: 1 NL: 2.09E6
T: FTMS + p ESI Full lock ms [100.00-700.00]



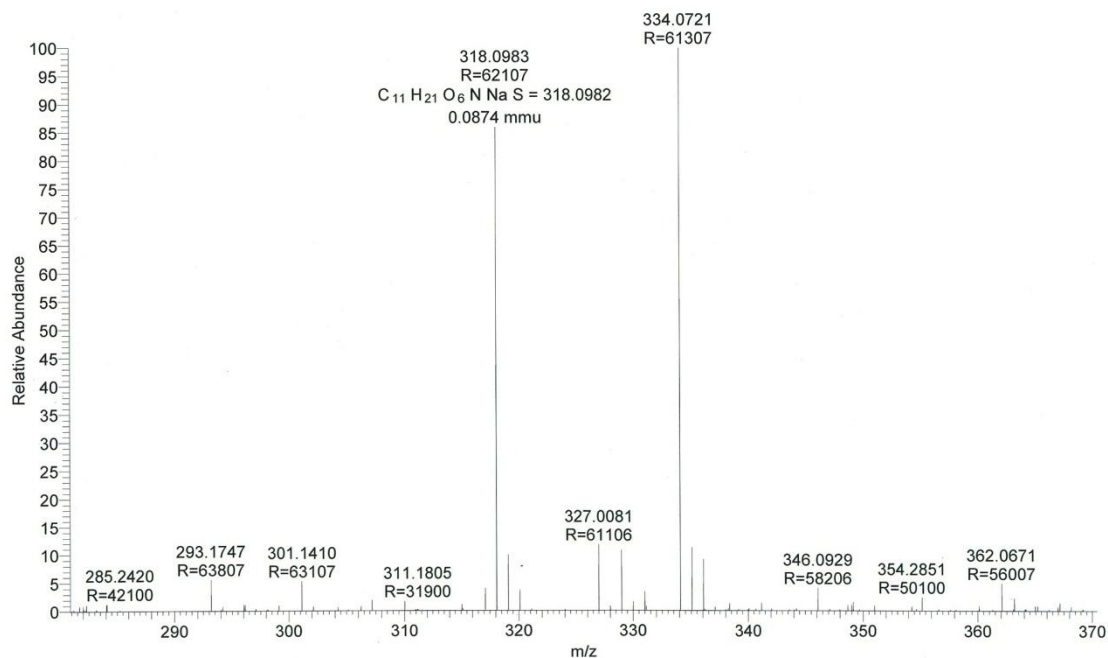
HRMS of compounds 12

AB-11_121017150530 #1514 RT: 6.74 AV: 1 NL: 2.01E8
T: FTMS + p ESI Full ms [100.00-700.00]



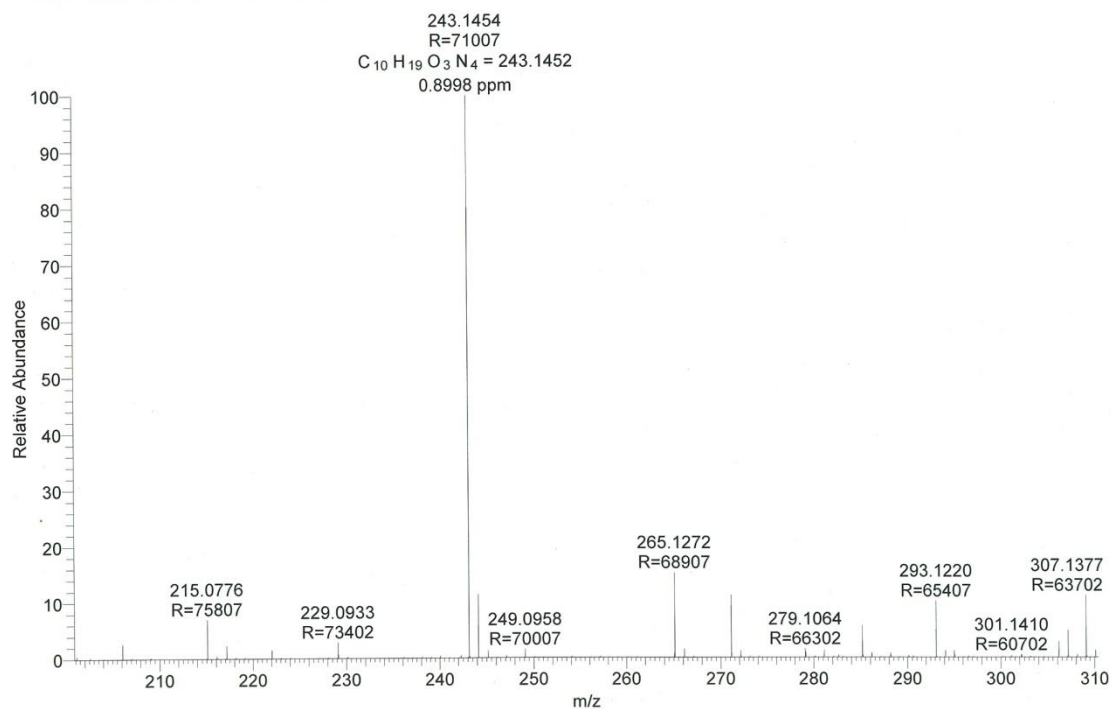
HRMS of compounds 13

AB12 #127 RT: 0.56 AV: 1 NL: 3.60E7
T: FTMS + p ESI Full lock ms [100.00-700.00]



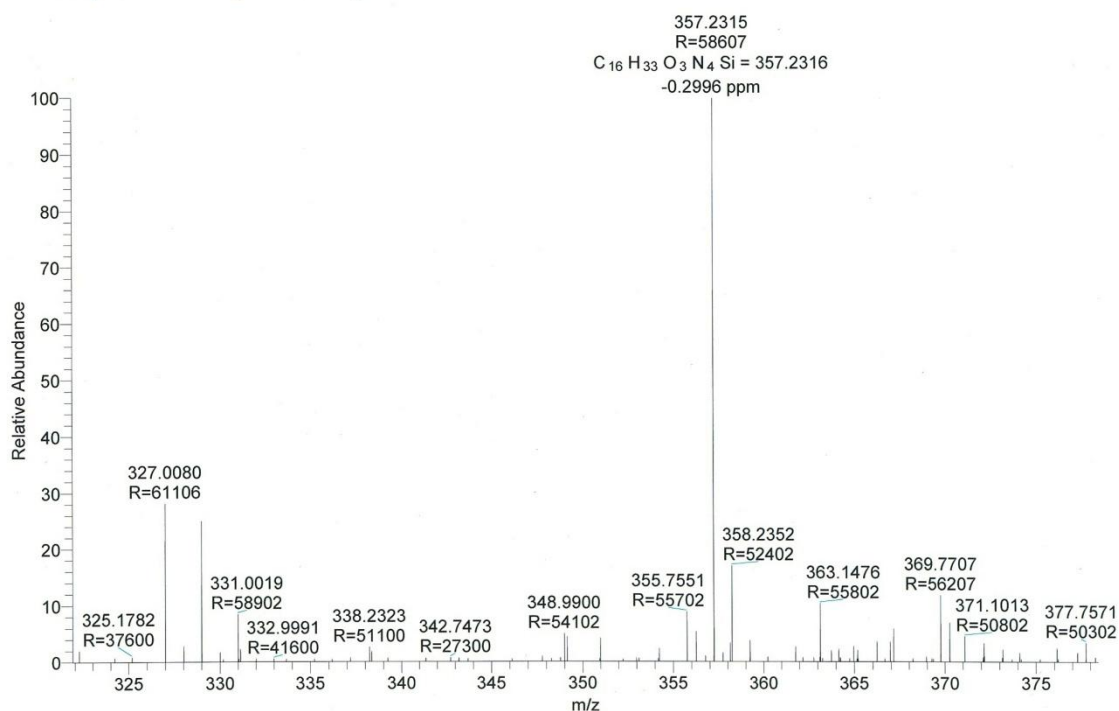
HRMS of compounds 14

AB13 #110 RT: 0.49 AV: 1 NL: 1.64E8
T: FTMS + p ESI Full ms [100.00-700.00]



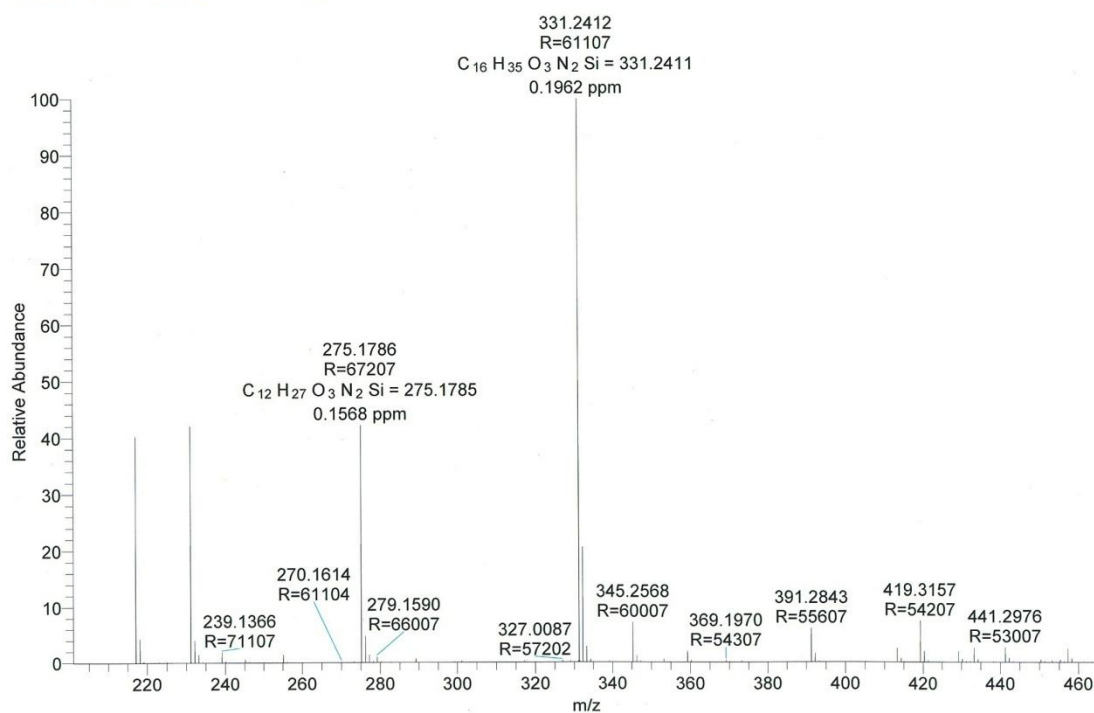
HRMS of compounds 15

AB14 #395 RT: 1.76 AV: 1 NL: 1.10E7
T: FTMS + p ESI Full lock ms [100.00-700.00]



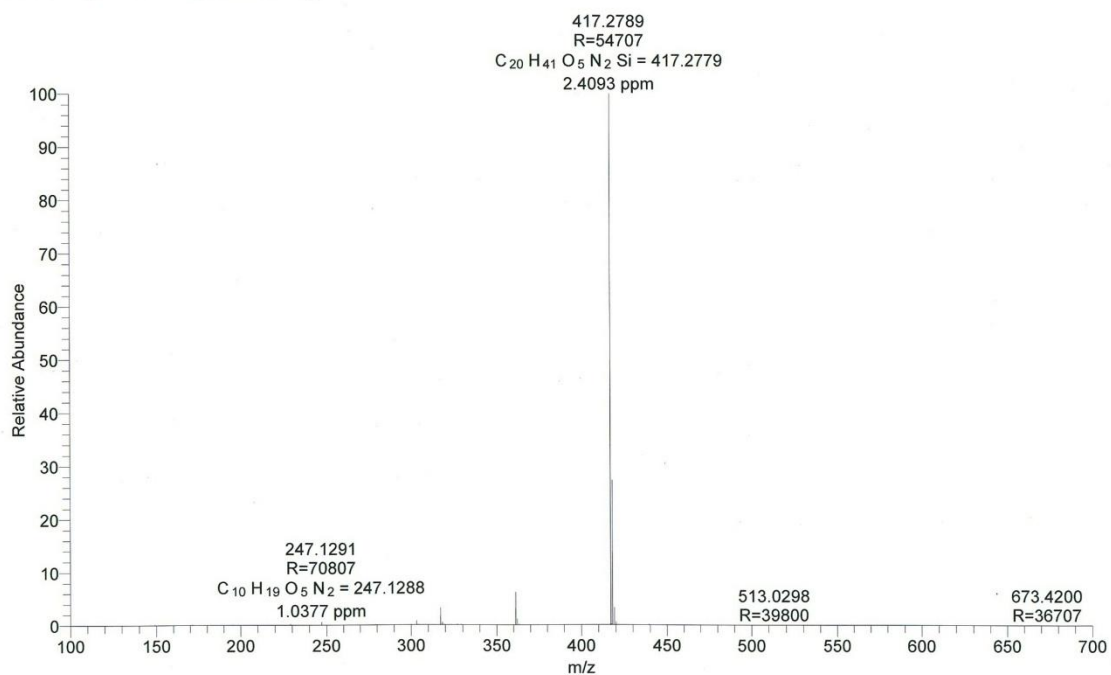
HRMS of compounds 16

AB15 #128 RT: 0.57 AV: 1 NL: 8.94E8
T: FTMS + p ESI Full ms [100.00-700.00]



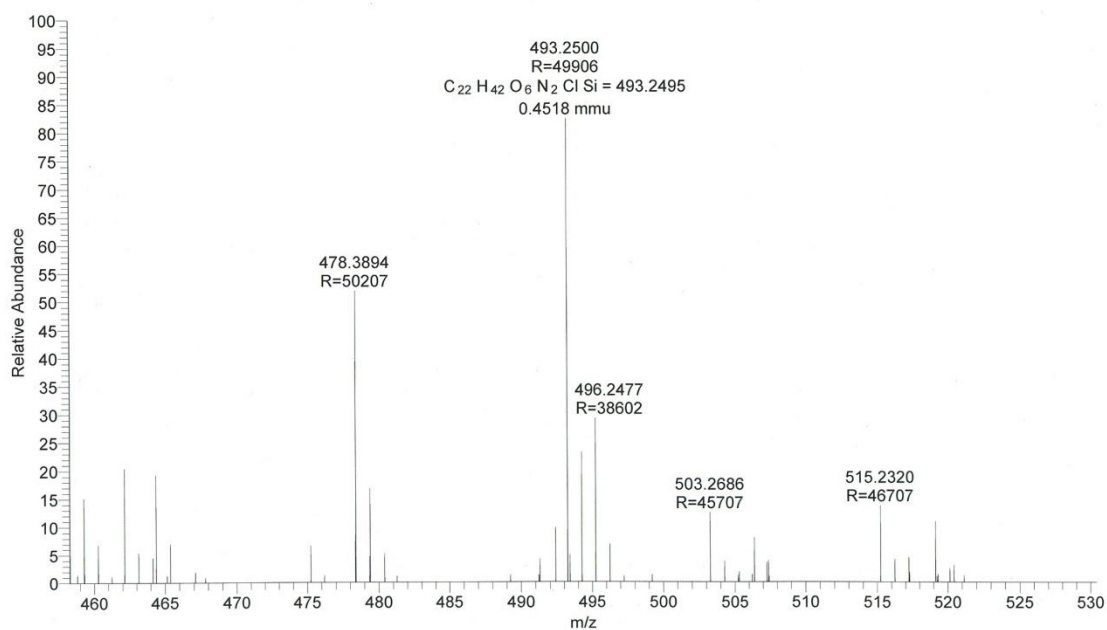
HRMS of compounds 17

ab16_121015110643 #949 RT: 4.23 AV: 1 NL: 9.64E9
T: FTMS + p ESI Full ms [100.00-700.00]



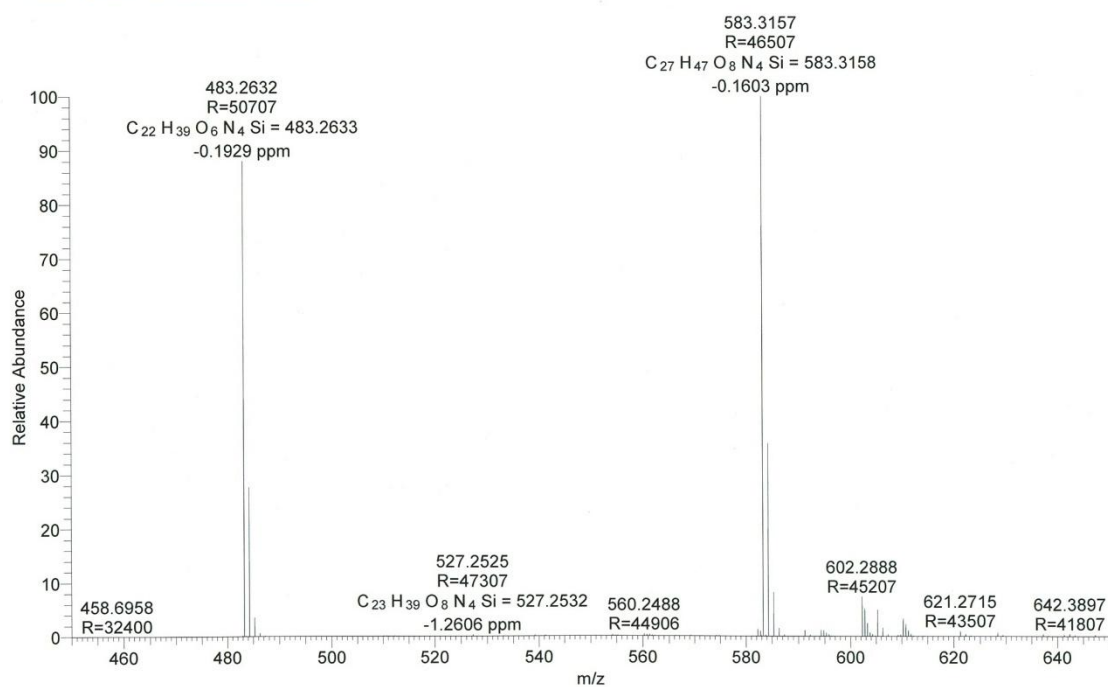
HRMS of compounds 18

AB17 #205 RT: 0.91 AV: 1 NL: 4.19E6
T: FTMS + p ESI Full lock ms [100.00-700.00]

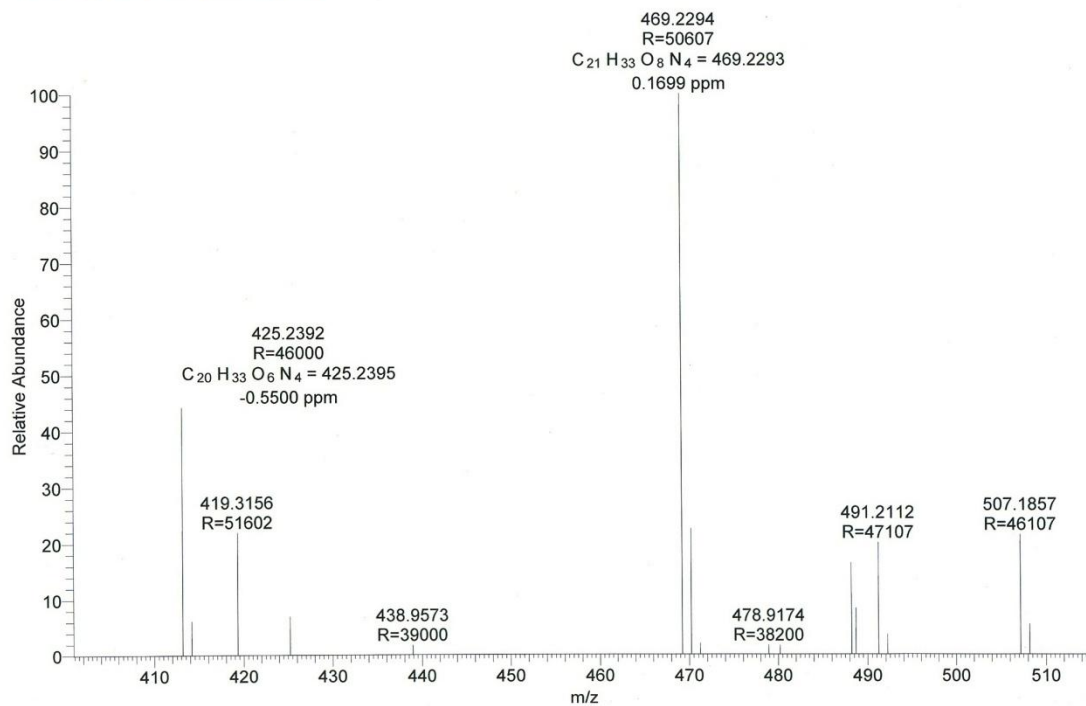


HRMS of compounds **19**

AB18_121015111753 #1920 RT: 8.56 AV: 1 NL: 5.33E8
T: FTMS + p ESI Full ms [100.00-700.00]

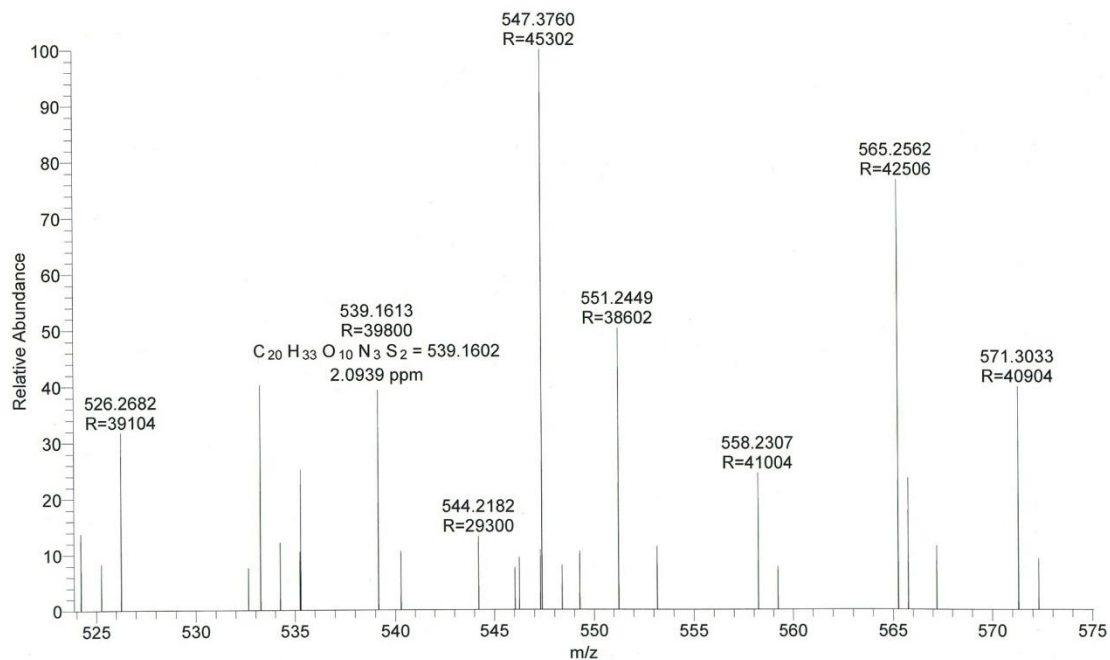
HRMS of compounds **20**

ab19 #110 RT: 0.49 AV: 1 NL: 4.68E7
T: FTMS + p ESI Full ms [100.00-700.00]

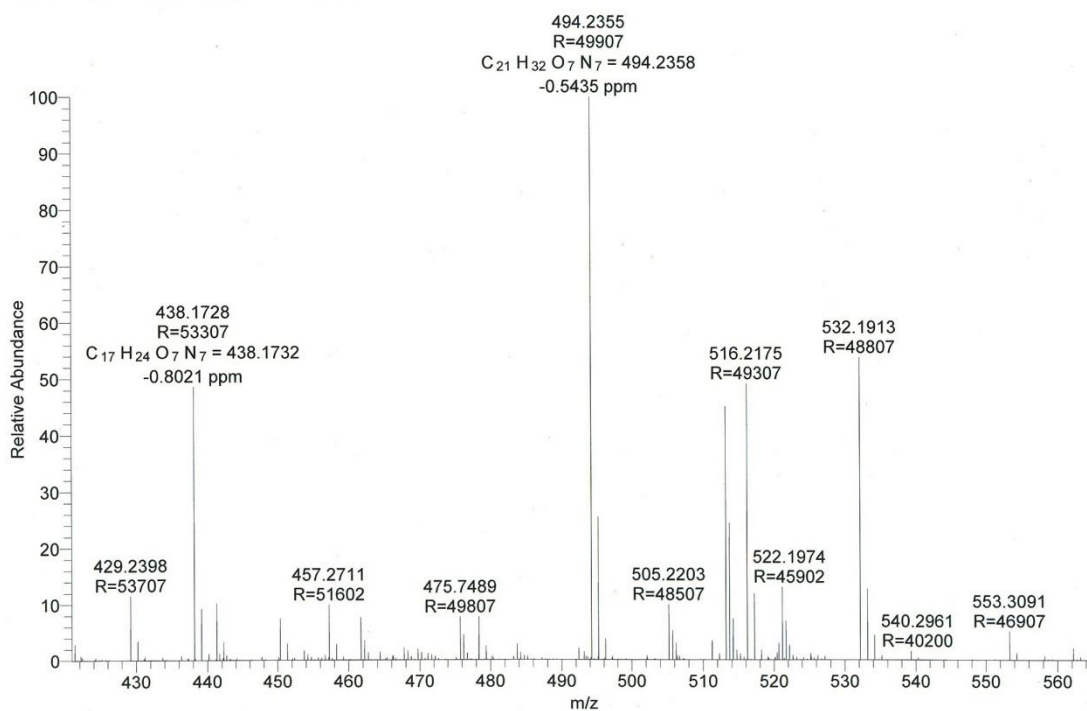


HRMS of compounds **21**

ab20 #129 RT: 0.57 AV: 1 NL: 7.83E5
T: FTMS + p ESI Full lock ms [100.00-700.00]

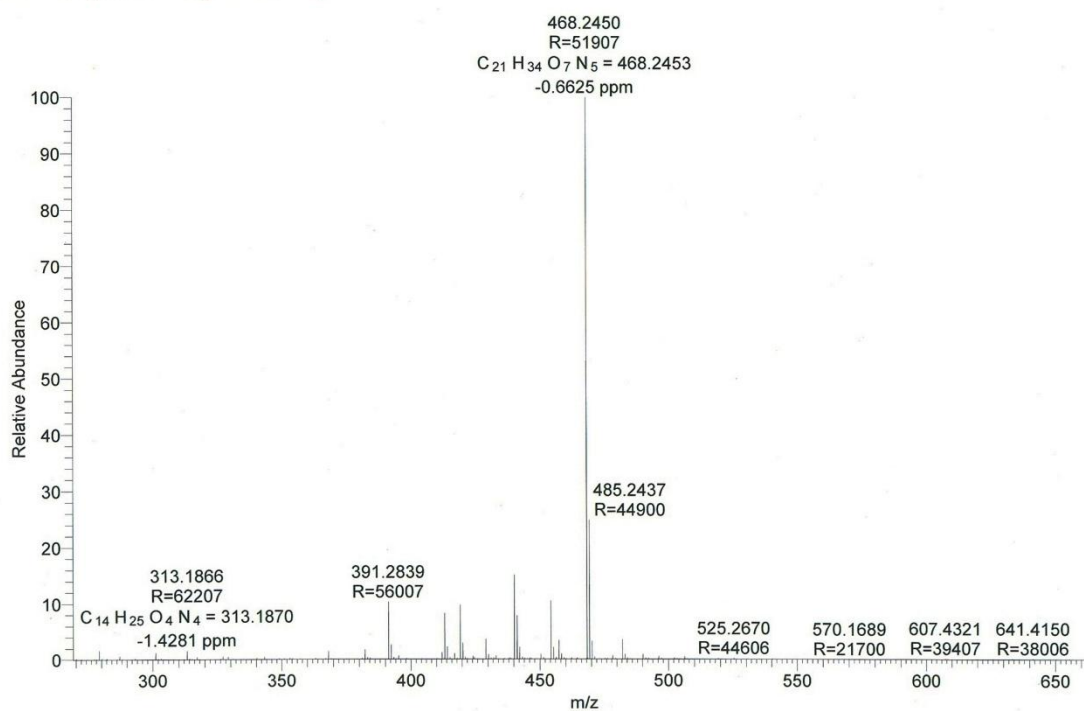
HRMS of compounds **22**

ab21 #124 RT: 0.55 AV: 1 NL: 7.44E7
T: FTMS + p ESI Full lock ms [100.00-700.00]

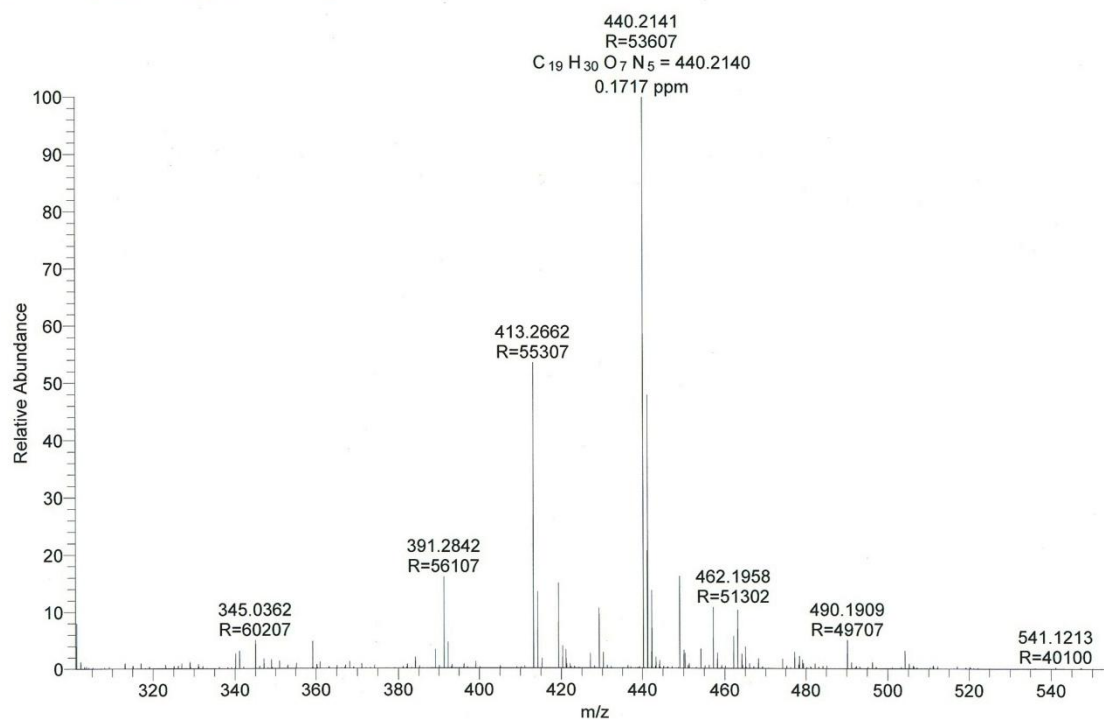


HRMS of compounds **23**

ab22 #99 RT: 0.44 AV: 1 NL: 4.55E8
T: FTMS + p ESI Full ms [100.00-700.00]

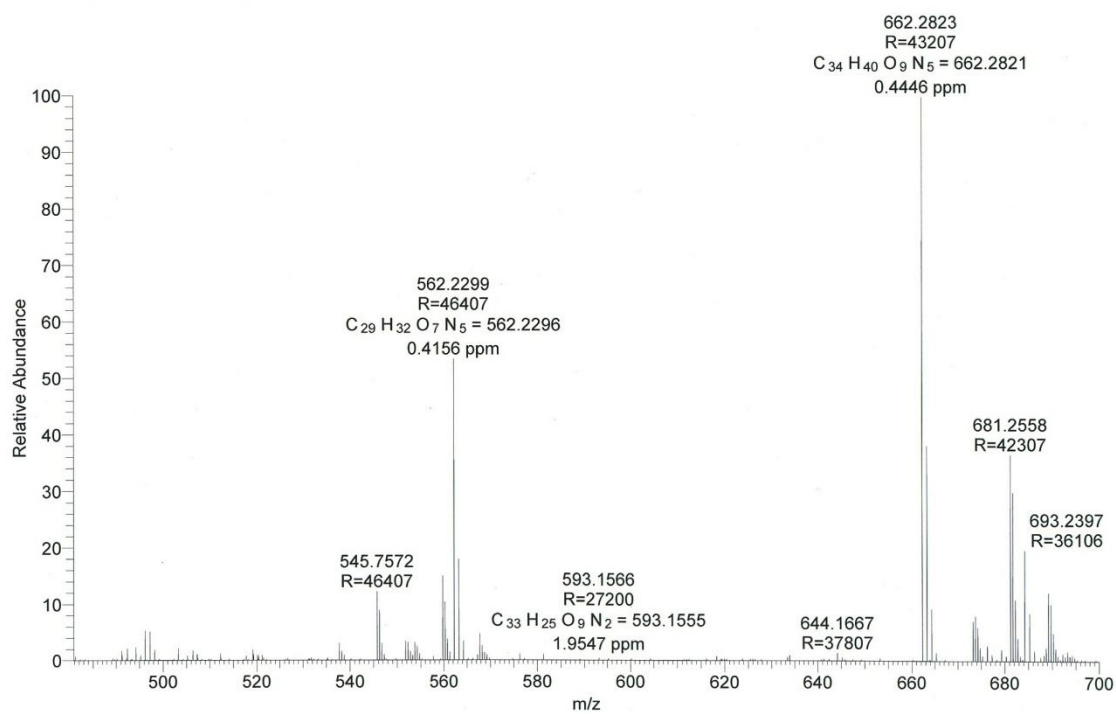
HRMS of compounds **24**

ab23 #106 RT: 0.47 AV: 1 NL: 1.25E8
T: FTMS + p ESI Full ms [100.00-700.00]

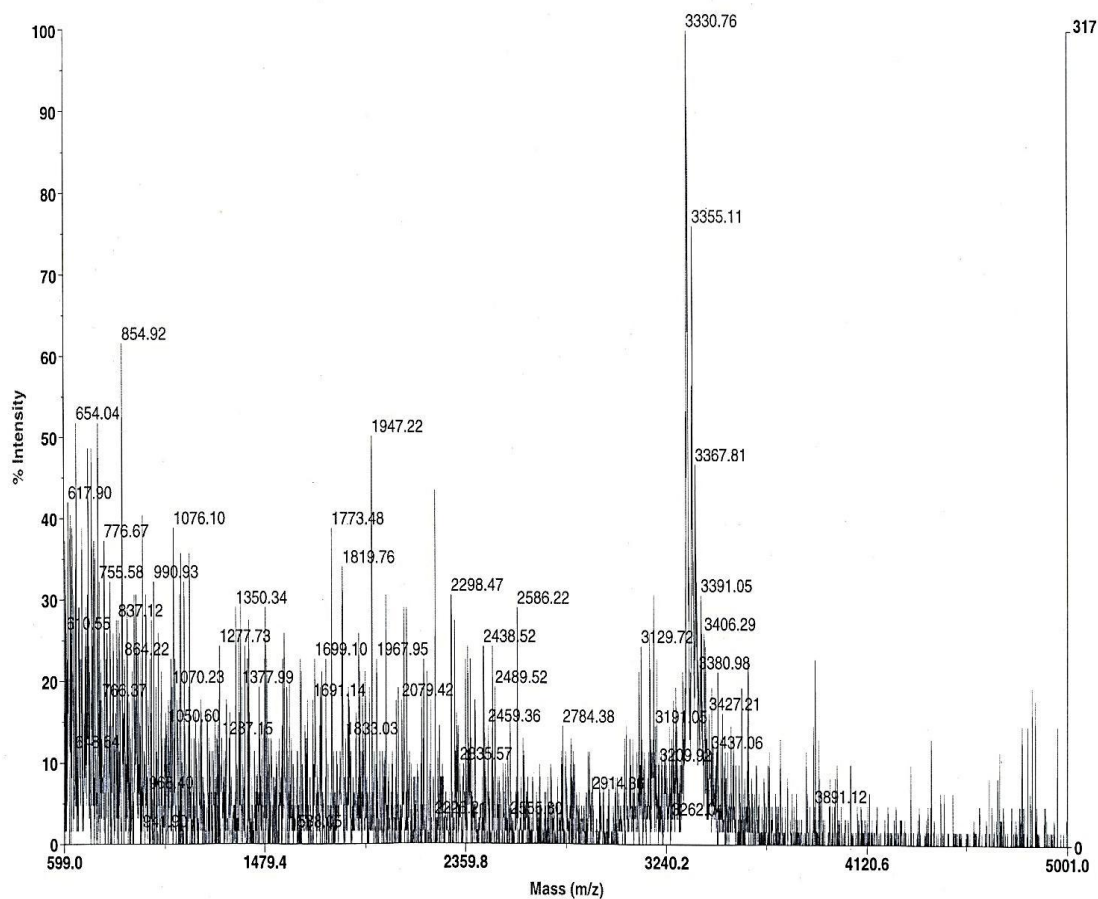
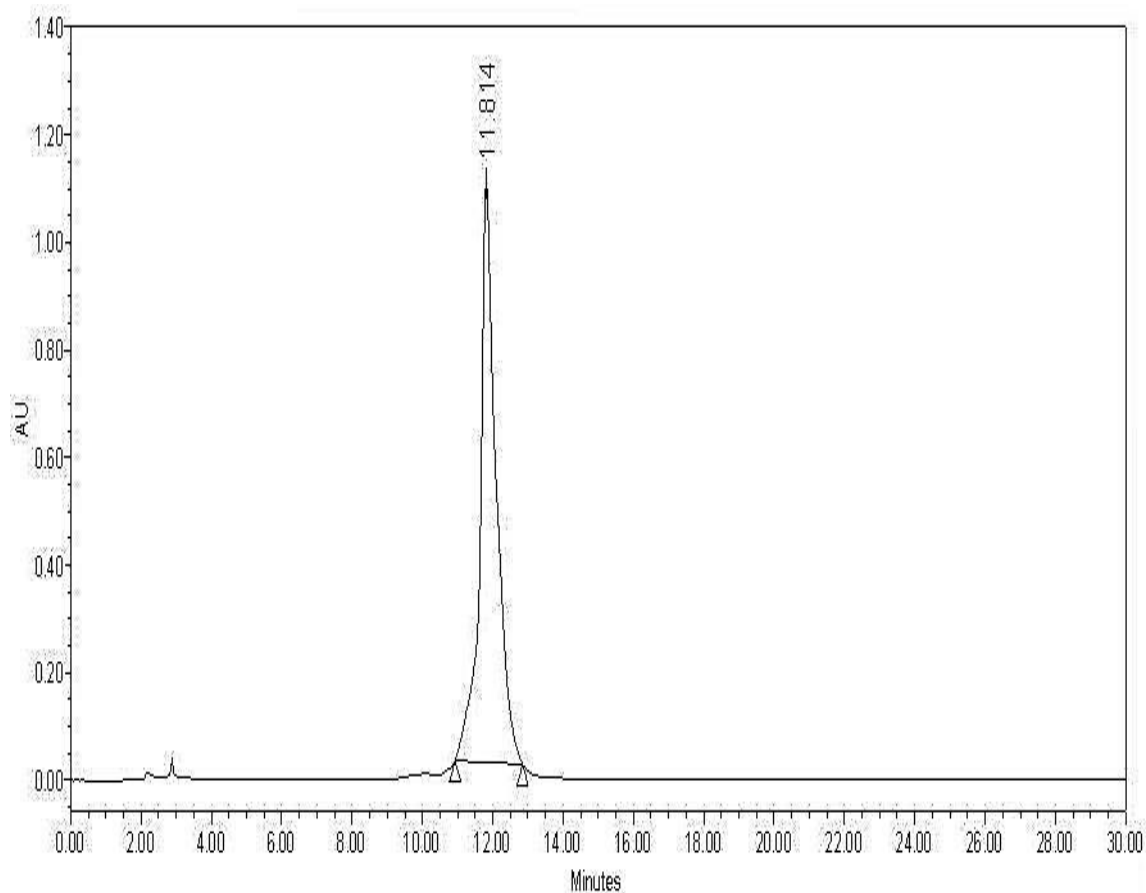


HRMS of compounds 25

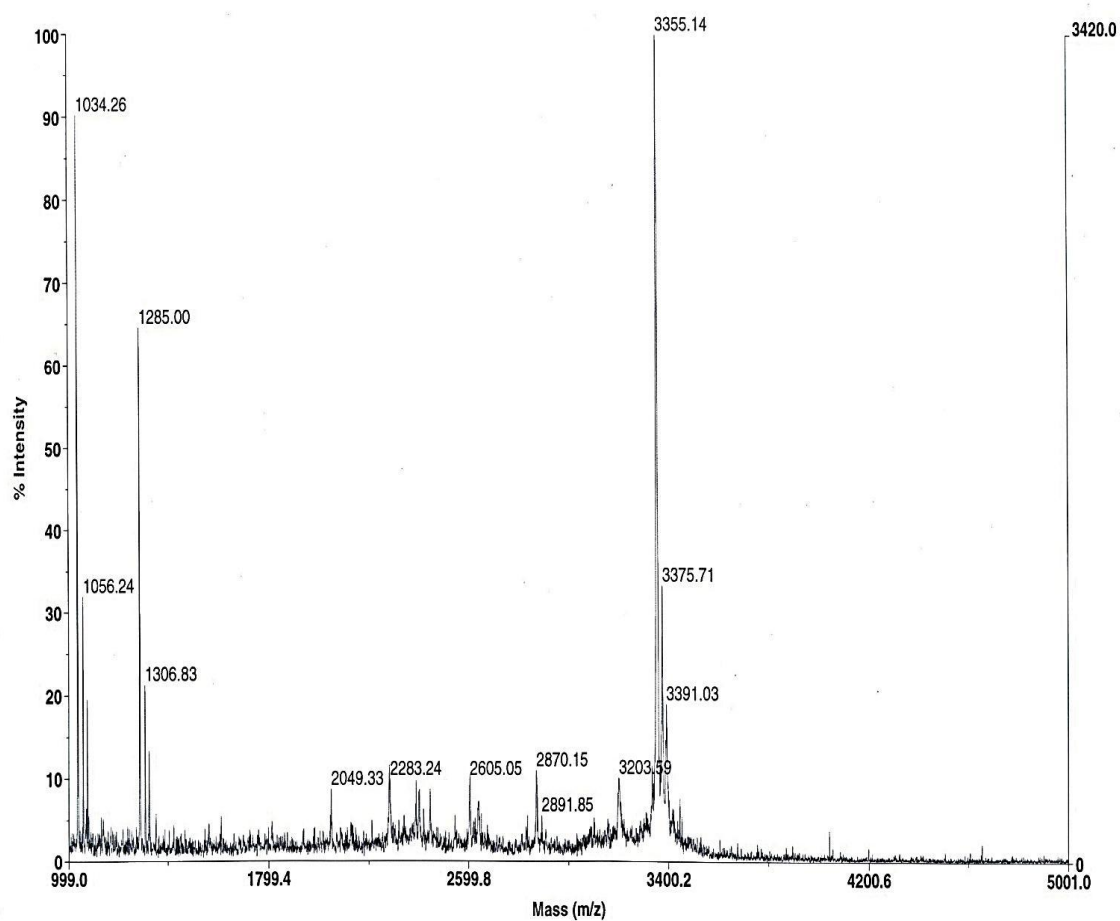
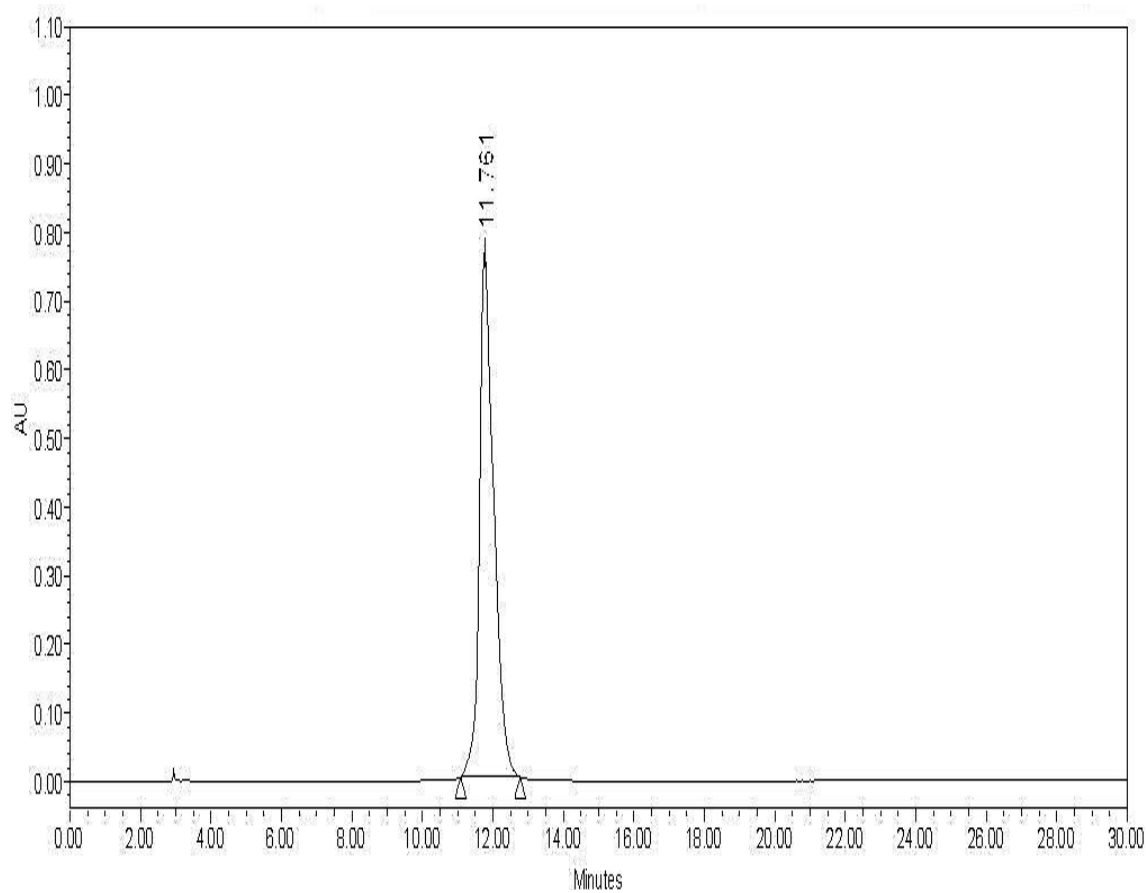
ab24 #134 RT: 0.59 AV: 1 NL: 2.57E7
T: FTMS + p ESI Full lock ms [100.00-700.00]



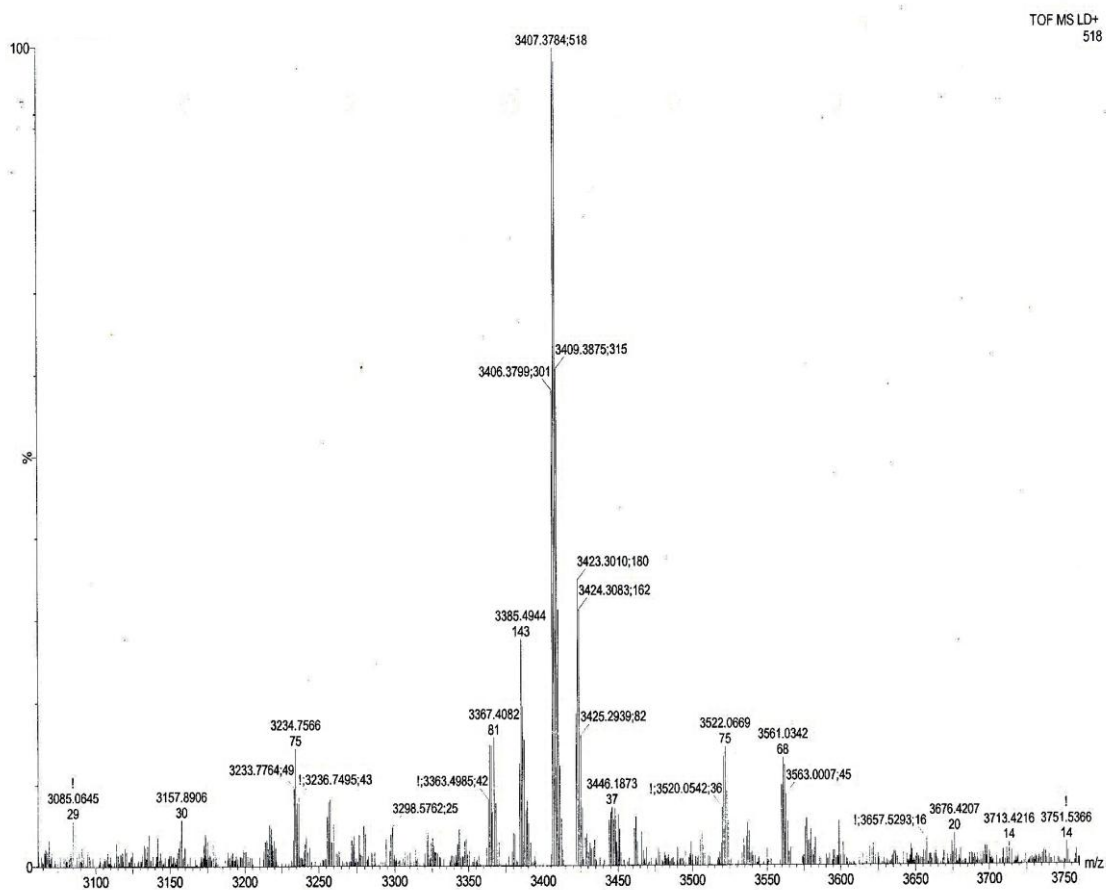
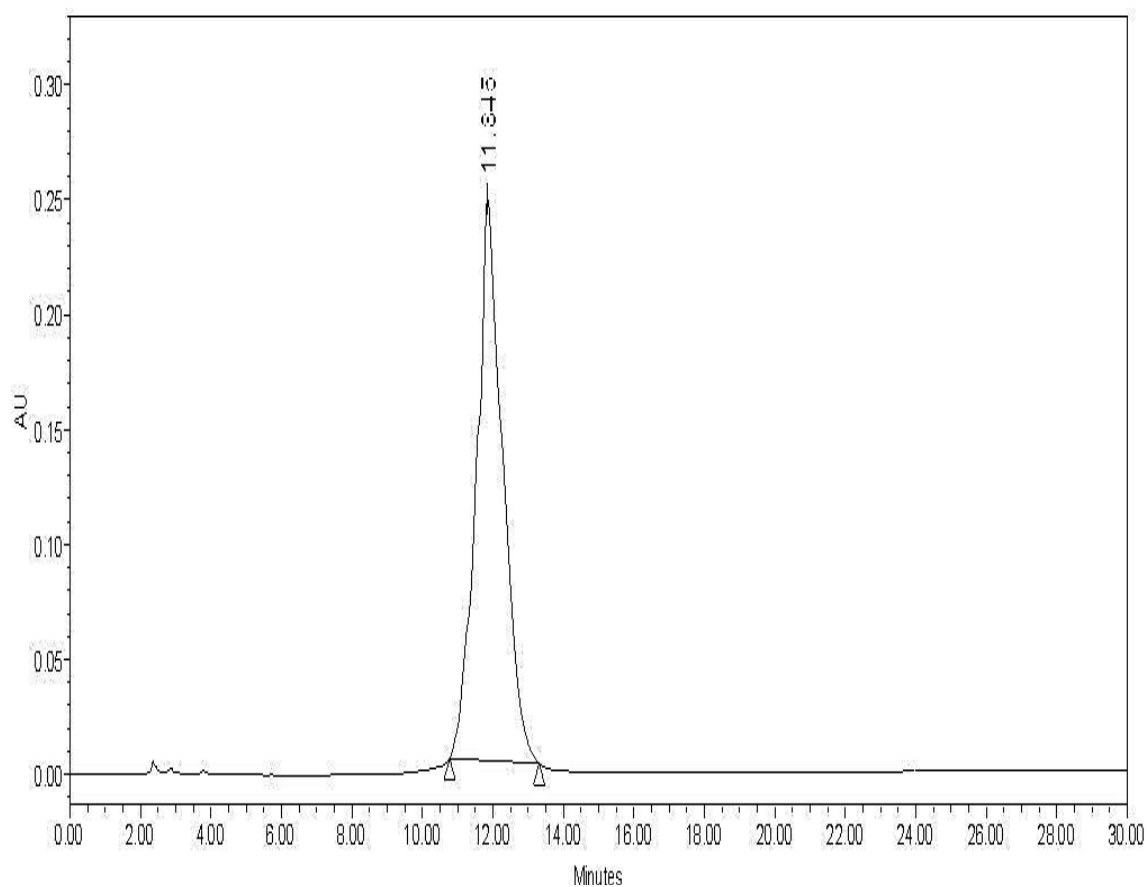
HPLC and MALDI-TOF of PNA 1



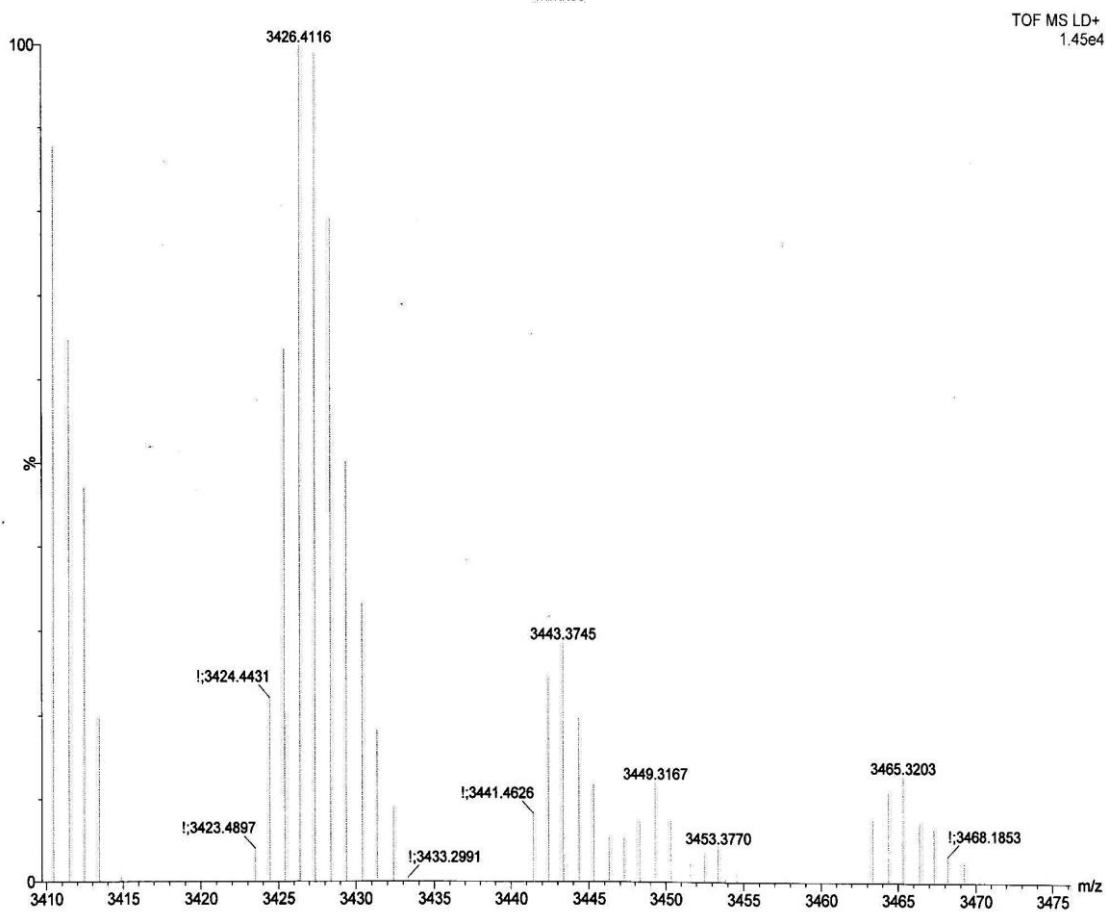
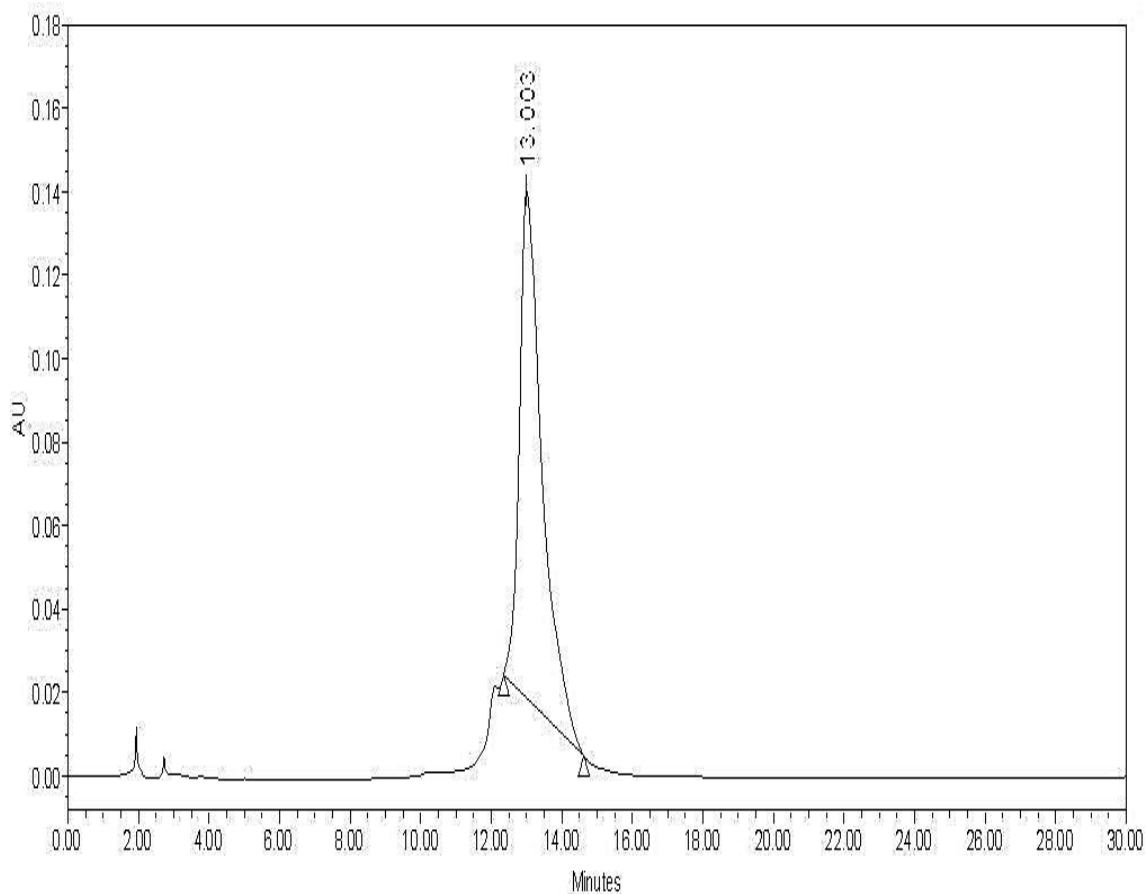
HPLC and MALDI-TOF of PNA 2



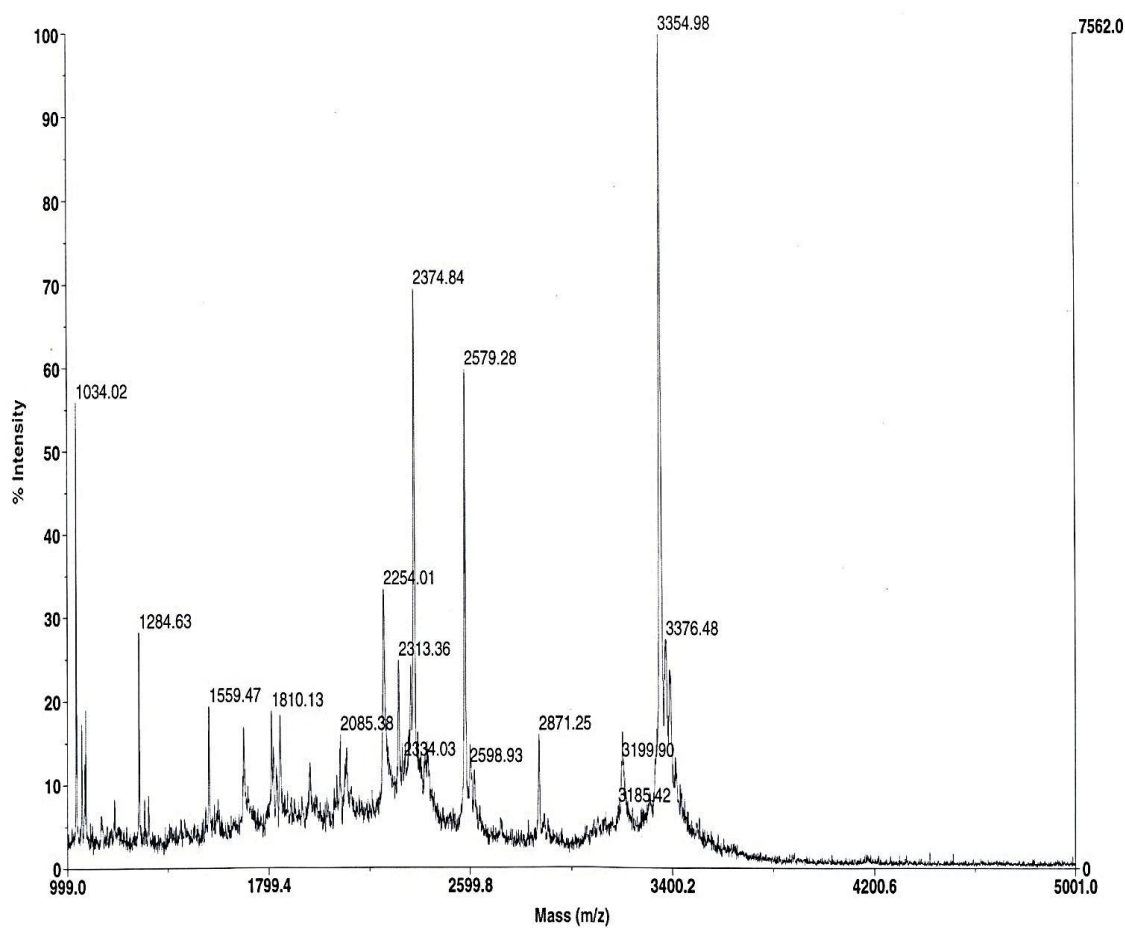
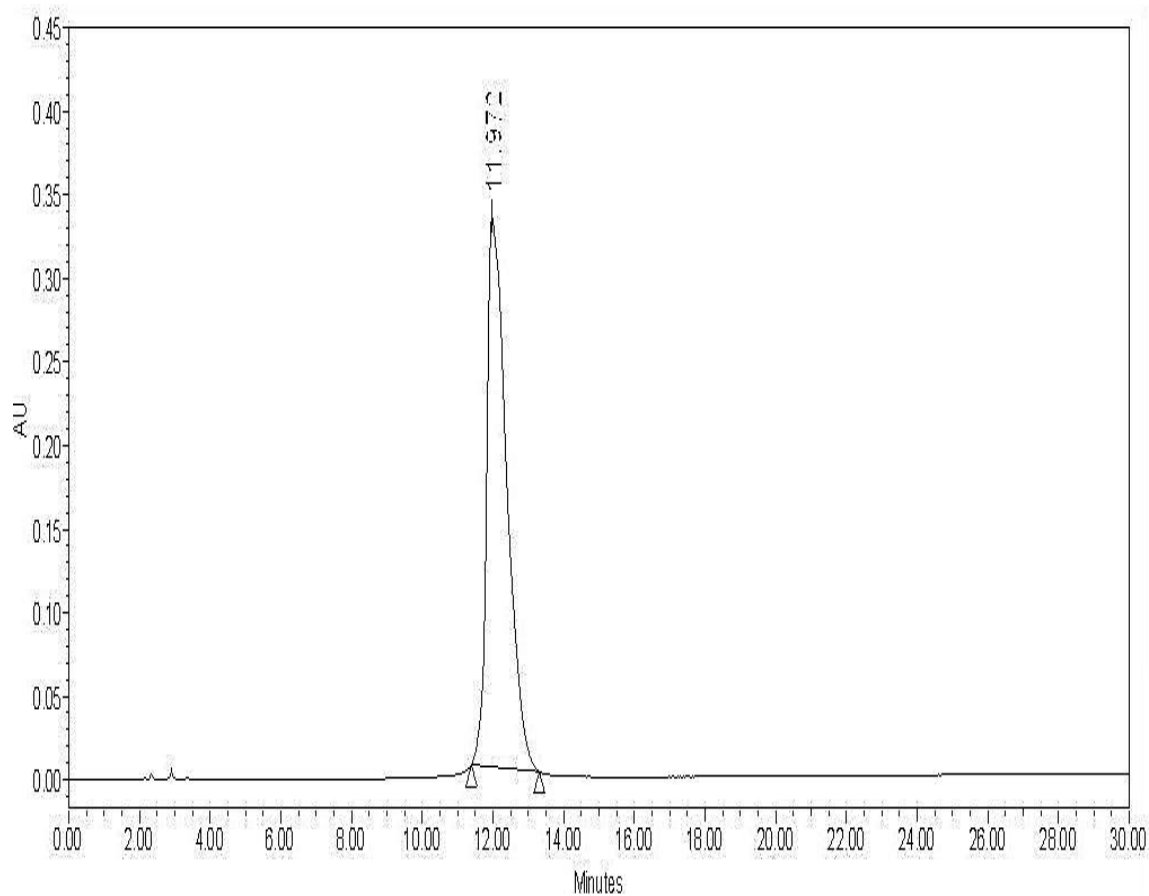
HPLC and MALDI-TOF of PNA 3



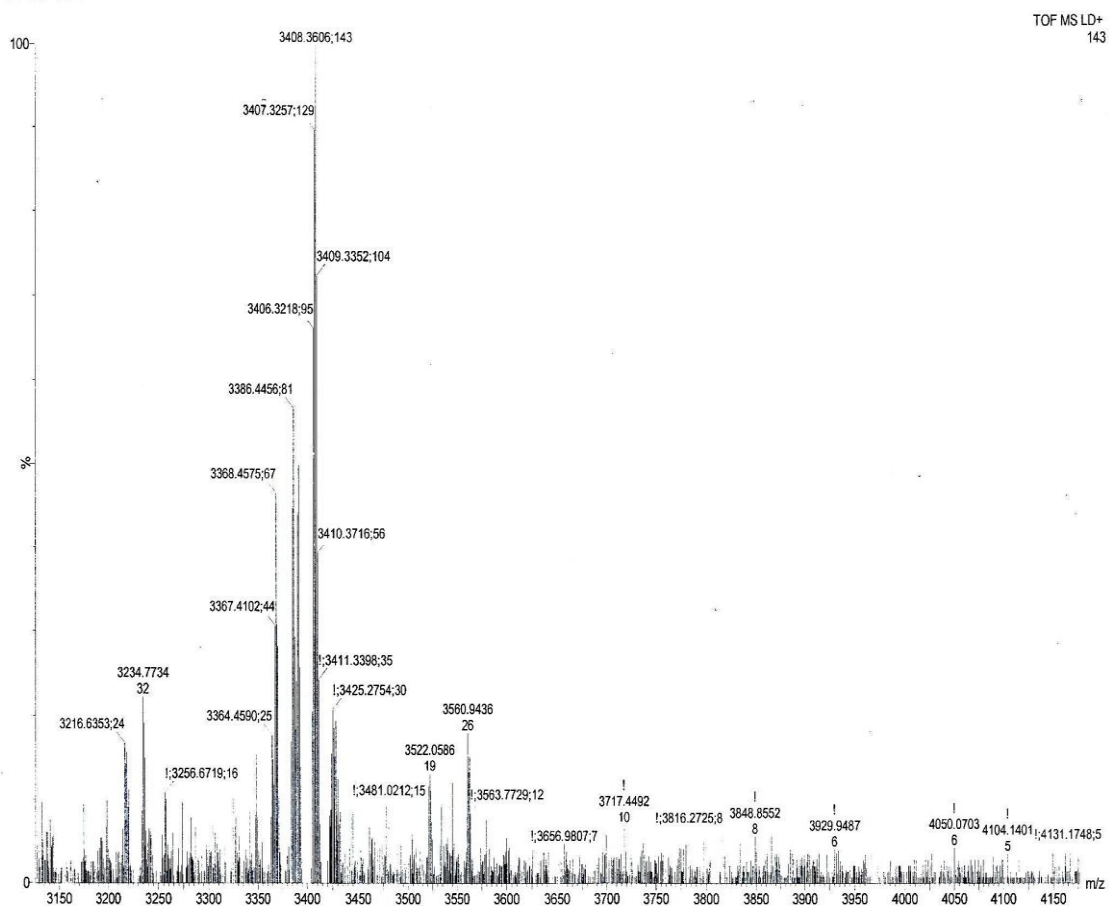
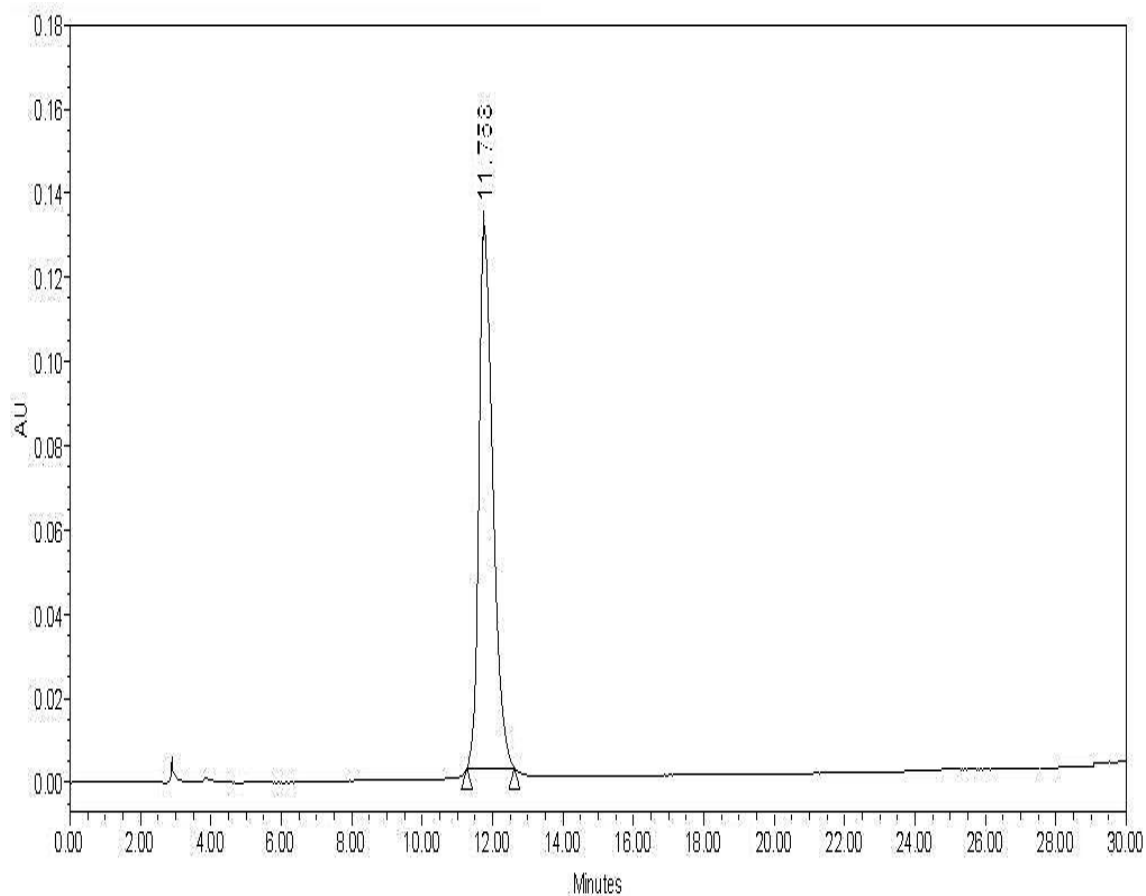
HPLC and MALDI-TOF of PNA 4



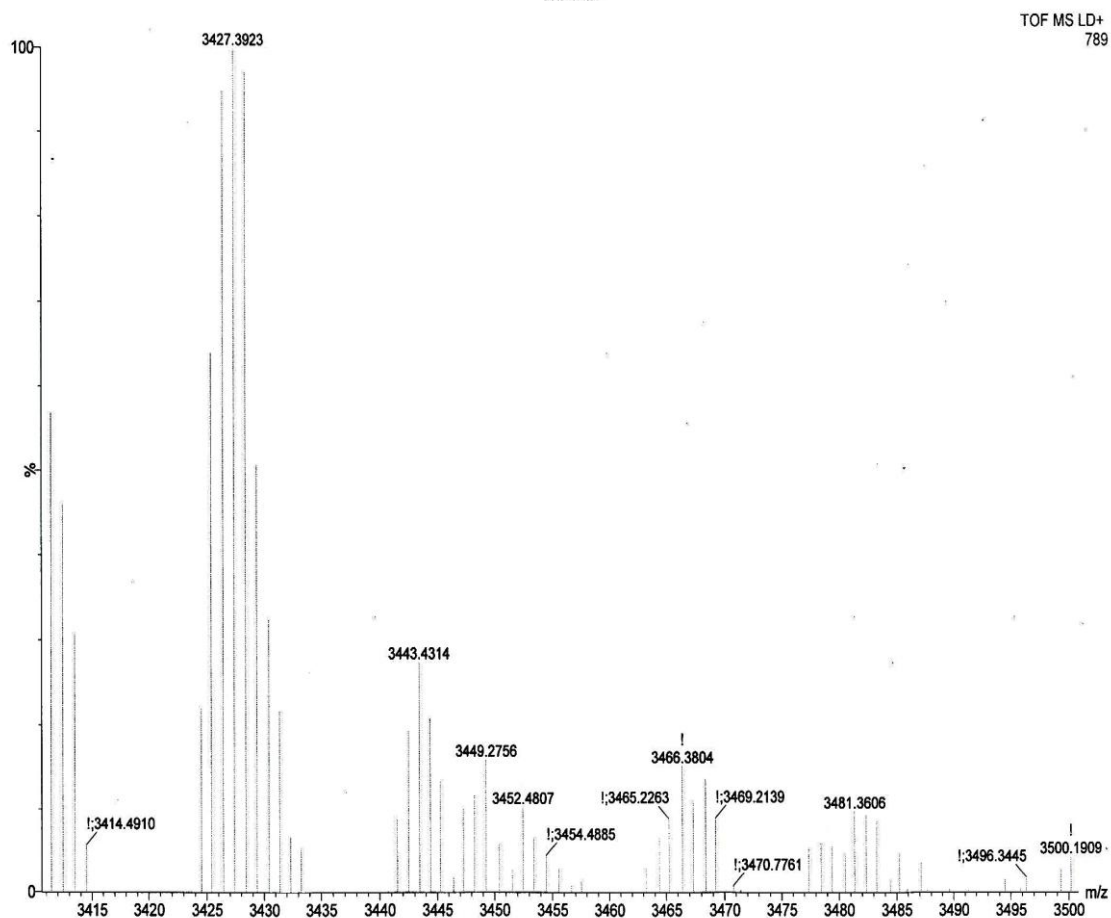
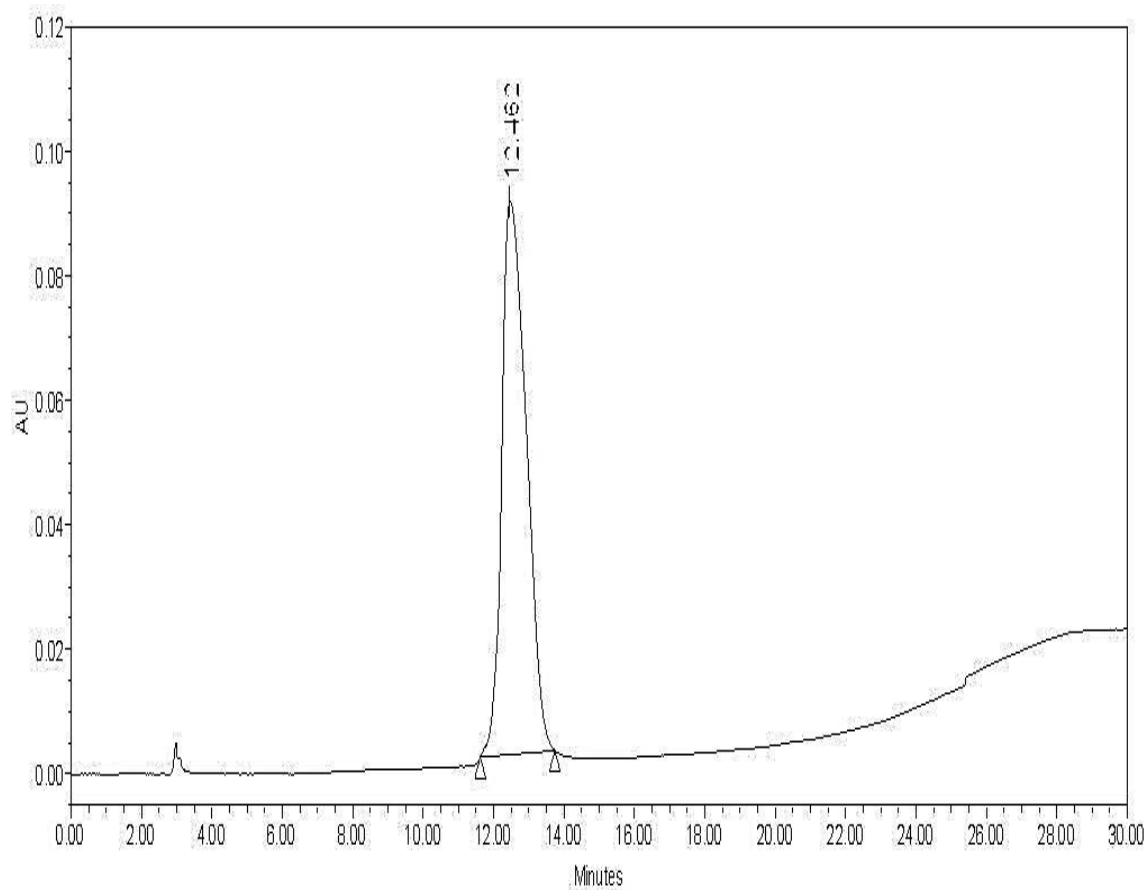
HPLC and MALDI-TOF of PNA 5



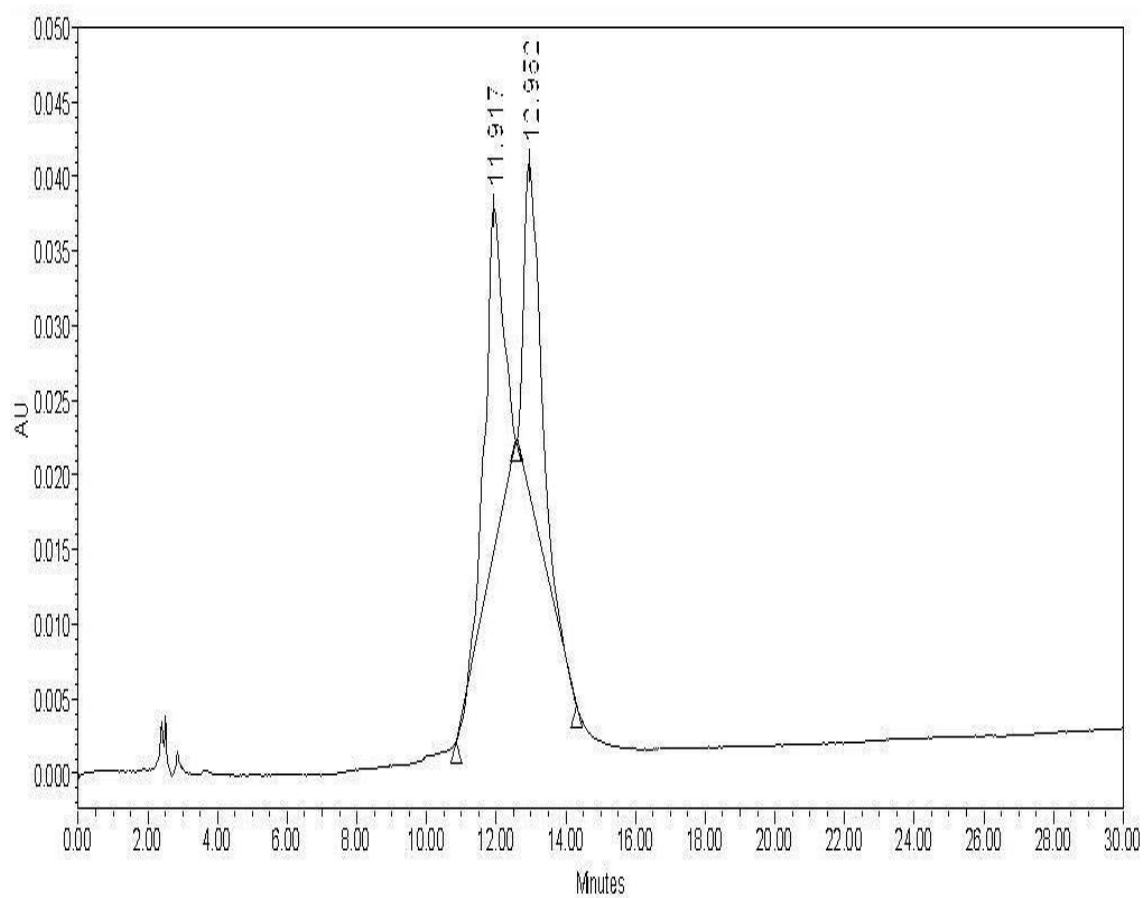
HPLC and MALDI-TOF of PNA 6



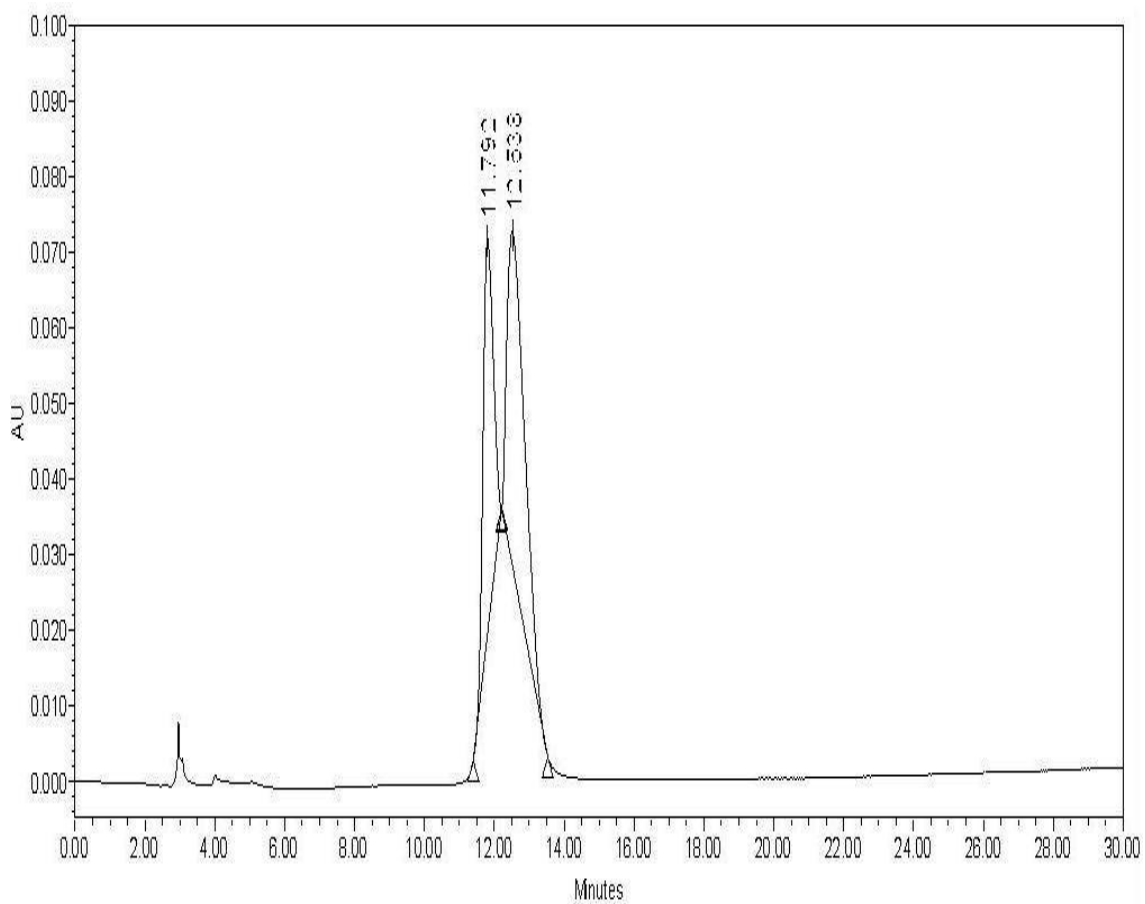
HPLC and MALDI-TOF of PNA 7



HPLC of PNA 3 + PNA 4



HPLC of PNA 6 + PNA 7



Section B

Biophysical studies of ethano-PNA towards comp. DNA/RNA

2B Introduction

The purpose of the designed oligonucleotides is to achieve high binding affinity and specificity to target nucleic acid. All newly designed and synthesized molecules were hence subjected to various biophysical and biochemical studies to evaluate their potential as antisense agents.

2B.1 Rationale and objectives of the present work

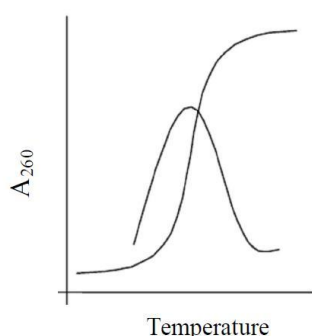
The preceding section reports the synthesis of conformationally constrained PNA analogue ethano-PNA. These ethano-PNA monomer units were introduced at various positions in PNA oligomers in order to study the effect of structural constraint and cationic nature of pyrrolidiny ring functionalized with amino and guanidino groups. In this section the studies that measure the influence of modification on the binding properties to target nucleic acids in the terms of strength and specificity is presented.

In this Section is to present the results of the systematic study of the different, H-ethano-PNA, Am-ethano-PNA and Gu-ethano-PNA backbone modifications in terms of their ability and strength to form specific base paired complexes with DNA/RNA is presented. Also the effect of cationic substituents in the pyrrolidiny ring on the strength of binding with the comp. DNA/RNA was studied. The specificity of binding was addressed by performing experiments with mismatched pairing at a single site in the target DNA/RNA sequences. This Section addresses the biophysical studies of ethano-PNAs and their hybrids with complementary nucleic acids using techniques such as UV spectroscopy and CD spectrophotometry.

2B.2 Biophysical spectroscopic techniques for studying PNA- DNA/ PNA-RNA interactions

Monitoring the UV absorption at 260 nm as a function of temperature has been extensively used to study the thermal stability of nucleic acid duplexes and triplexes and consequently, PNA-DNA hybrids as well. Increasing the temperature perturbs these complexes, inducing a structural transition by causing disruption of hydrogen bonds between the base pairs, diminished stacking interactions between adjacent nucleobases and larger torsional motions in the backbone leading to a loss of secondary and tertiary

structure. This is reflected in the UV absorption at 260 nm, termed as hyperchromicity. The magnitude of hyperchromicity is a measure of the extent of the secondary structure present in nucleic acids. The process is cooperative and the plot of the absorbance at



260 nm vs the temperature is therefore sigmoidal. This also represents a two-state/All or none model for nucleic acid melting *i.e.* the nucleic acids exist in only two states, either in duplexes or single strands at varying temperatures and the relative proportion change. A non-sigmoid (e.g. linear) transition with low hyperchromicity is a consequence of non duplexation (non complementation). In many cases, the transitions are broad and the exact T_m s are obtained from the peak in the first derivative plots. This technique has provided valuable information regarding complementary interactions in nucleic acids hybrid involving DNA, RNA and PNA.⁵³

The fidelity of base pairing in the PNA:DNA/PNA:RNA complexes can be examined by changing the PNA oligomer with a DNA/RNA strand bearing mismatch at a desired site, preferably opposite the site of modification. The base mismatch leads to the absence of or incorrect hydrogen bonding between the bases and causes a drop in the measured melting temperature. A modification of the PNA structure is considered good if it gives a significantly lower T_m with RNA/ DNA sequences containing mismatches as compared to unmodified PNA. It is to be pointed out that in all biophysical experiments described herein; the modified PNAs were always evaluated by comparison with unmodified control PNA.

2B.3 UV- T_m studies on ethano-PNA:cDNA and ethano-PNA:RNA

The synthesized PNA sequences were examined for their binding affinity with cDNA/RNA sequences by employing UV-thermal denaturation studies. The sequence chosen for study was a 12mer sequence. The modified oligomers were annealed with the cDNA and RNA and were subjected to temperature dependent UV studies at 260 nm. Unmodified **PNA 1** was used as a control in these experiments. The UV- T_m plots show a single sigmoidal transition, characteristic of PNA:cDNA/RNA duplex melting. The T_m values were obtained by the first derivative and the values corresponding to peaks from such plots for various PNA:cDNA/RNA duplexes of modified and unmodified oligomers are shown in Table 2. It is seen from the UV-melting data that the introduction of single ethano-PNA-T modification in the middle of the sequence (**PNA 2**) decreased the T_m of derived **PNA 2**:cDNA duplex (ΔT_m - 14.9 °C) over

unmodified **PNA 1**:cDNA. However, the substitution of H-ethano-PNA to get Am-ethano-PNA (**PNA 3**) and Gu-ethano-PNA (**PNA 4**) reduced this change in T_m values (Table 2). The destabilization of the duplexes with amino/guanidino substituted modified PNA oligomers **PNA 3** and **PNA 4** was less compared to the uncharged modified **PNA2**:cDNA duplexes (ΔT_m -12.8 °C for **PNA 3**, ΔT_m - 10.9 °C for **PNA 4**) (Figure 4A). The UV- T_m studies involving complexation with RNA also showed decrease in the T_m values but to a lesser extent as compared to the complexes with cDNA (Figure 4B). In the case of H-ethano-PNA, the complex **PNA 2**:RNA, the ΔT_m was found to be - 6.7 °C compared with the unmodified **PNA1**:RNA (Table 2). This destabilization was less as compared to **PNA2**:cDNA duplex. When the pyrrolidine ring carried the amino and guanidino substituents in **PNA 3** and **4**, respectively, the ΔT_m was - 5.7 °C and - 3.4 °C, compared to the unmodified **PNA 1**:RNA complex.

We then synthesized PNA oligomers **PNA 5- 7**, in which the modification site was shifted towards C-terminal end, rather than in the middle of the sequence, to study the effect of change in the site of modification.^{54, 55} When subjected for thermal denaturation studies, the UV- T_m profiles indeed improved than when the modification was in the middle of the oligomers while binding with both DNA and RNA. In the case of complexation with cDNA, H-ethano-PNA, **PNA 5**, the ΔT_m was - 8.8 °C. The Am-ethano-PNA (**PNA 6**) and Gu-ethano-PNA (**PNA 7**) substitution showed higher stability when compared with H-ethano-PNA oligomer (Figure 5A), but still destabilization was observed (ΔT_m - 7.5 °C and - 4.0 °C respectively). In contrast, the result for the duplex stability with RNA (Figure 5B) showed that the thermal stability of the modified duplexes was as good as the duplex with unmodified PNA (ΔT_m - 0.6 °C and + 0.2 °C for **PNA 5** and **PNA 6**, respectively). When the amino group was converted to guanidino functionality (**PNA 7**), the ΔT_m was found to be positive as the T_m of the duplex **PNA 7**:RNA was higher by + 2.0°C compared with the unmodified **PNA1**:RNA duplex. Thus, the modified **PNA 7** was able bind strongly with RNA. The comparison of the melting data of these modified oligomers while binding with cDNA and RNA, clearly indicated that the difference in T_m ($\Delta T_{m-RNA-DNA}$) is about + 9 to 10°C in all the modified oligomers (**PNA 2 - 7**). In the case of the unmodified achiral **PNA1** there was relatively much less discriminating for RNA binding ($\Delta T_{m-RNA-DNA}$ + 1.5°C). This result is in accordance with the rationale of our design which was aimed to be RNA selective.

Table 2. PNA sequences and UV- T_m studies. t^H , t^{Am} and t^{Gu} are the modified ethano-PNA monomers with H, aminomethyl and guanidinomethyl substituents respectively

Entry No.	Code	Sequences (N' → C')	UV - T_m (°C)		ΔT_m (°C)	
			cDNA	RNA	Modi-PNA1	Modi-PNA1
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	64.6	66.1	-	-
2.	PNA 2	H-CATTG t^H CACACT-Lys-NH ₂	49.7	59.4	-14.9	- 6.7
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	51.8	60.4	-12.8	- 5.7
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	53.7	62.7	-10.9	- 3.4
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	55.8	65.5	- 8.8	- 0.6
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	57.1	66.3	- 7.5	+ 0.2
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	60.6	68.1	- 4.0	+ 2.0

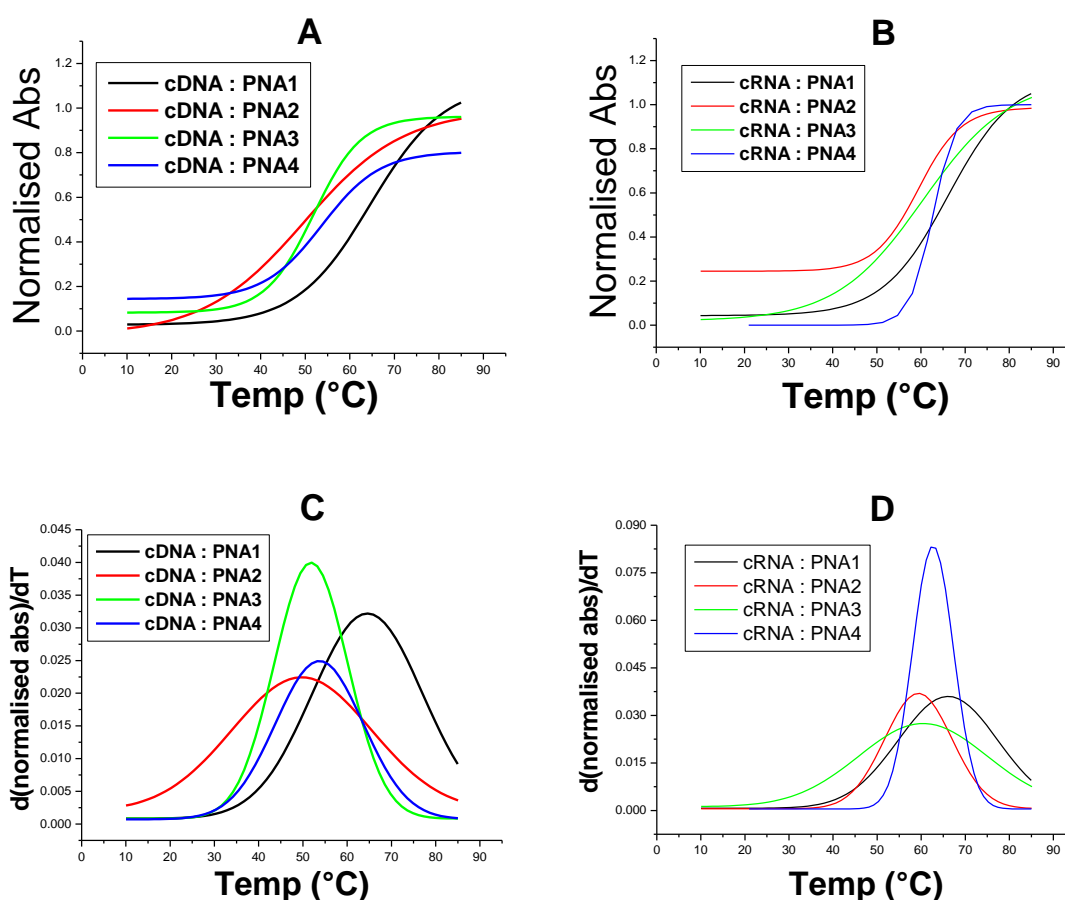


Figure 4. UV-Thermal Melting Profiles. Sigmoidal Curves for (A) cDNA:PNA1/ PNA2/ PNA3/ PNA4 (B) cRNA:PNA1/ PNA2/ PNA3/ PNA4, first derivative curves of the sigmoidal curves for melting values (C) cDNA:PNA1/ PNA2/ PNA3/ PNA4 (D) cRNA:PNA1/ PNA2/ PNA3/ PNA4

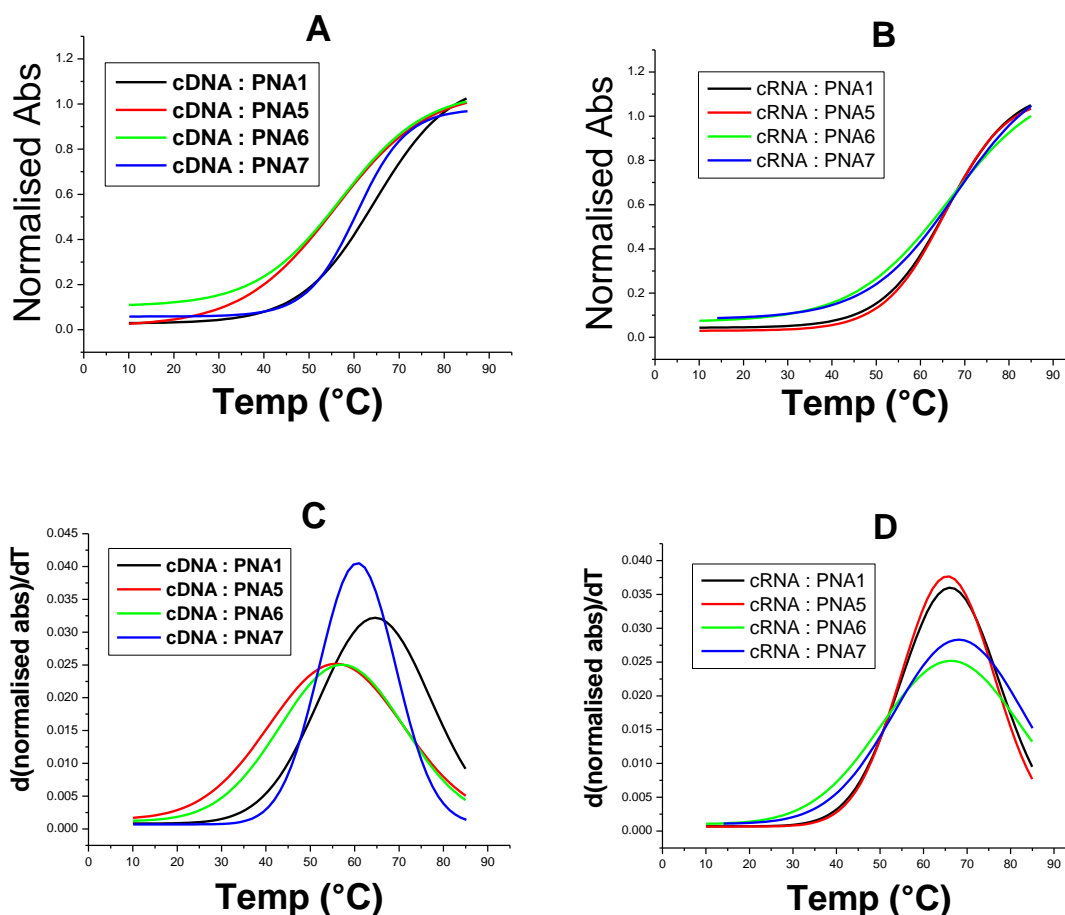


Figure 5. UV-Thermal Melting Profiles. Sigmoidal Curves for (A) cDNA:PNA1/ PNA5/ PNA6/ PNA7 (B) cRNA:PNA1/ PNA5/ PNA6/ PNA7, first derivative curves of the sigmoidal curves for melting values (C) cDNA:PNA1/ PNA5/ PNA6/ PNA7 (D) cRNA:PNA1/ PNA5/ PNA6/ PNA7

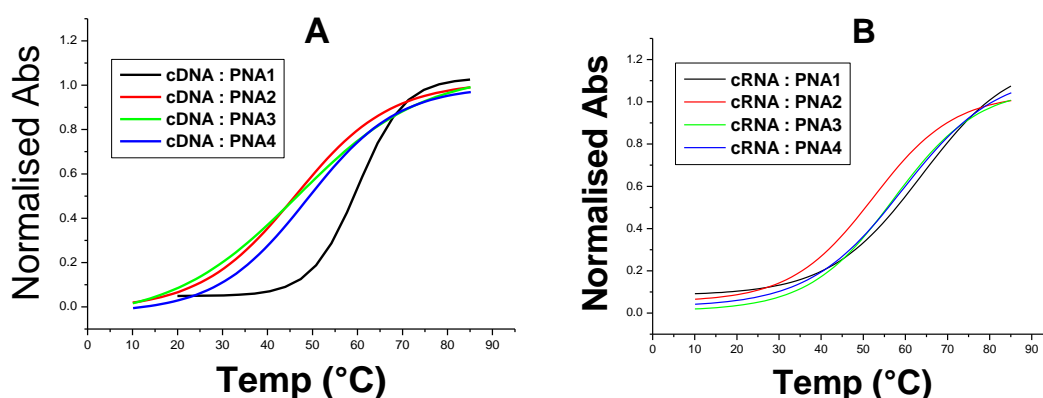
2B.3.1 Effect of salt concentration on UV- T_m of ethano-PNA:cDNA/RNA duplexes

PNA being charge-neutral, the strength of unmodified PNA:DNA duplexes is known to be little affected by increasing salt concentrations, in contrast to the highly salt concentration dependent duplex strength of native duplexes.⁵⁶ In the present studies, a positive charge was introduced in the backbone of PNA and therefore at higher salt concentration some destabilization of the duplexes could be expected, depending upon the electrostatic properties and interactions of the charged amino- and guanidino- substitutions in the backbone.³⁸ Thermal melting studies were therefore undertaken at higher salt concentration (100 mM NaCl). **PNA 1** showed slight destabilization of the duplexes with DNA at higher salt concentrations ($\Delta T_m = - 4.1^\circ\text{C}$, Table 3, Figure 6). The ΔT_m values for uncharged H-ethano-PNA (**PNA 2** and **5**) were similar when modification was present either at the central position or towards C-terminus ($\Delta T_m = - 3\text{-}5^\circ\text{C}$). In the case of Am-ethano-PNA and Gu-ethano-PNA (**PNA 3- 4** and **PNA 6 - 7**), the average destabilization of the duplexes was about 3-4 °C. In

the case of RNA binding at higher salt concentrations, the destabilization was more when the guanidino- group was present in the backbone as compared to amino- or unsubstituted pyrrolidine ring modification at pH 7. The salt dependent destabilization of the duplexes was observed by the earlier researchers also, when γ PNA (corresponding to 2'-position in the present study) was constructed from the lysine residue. In that case, the positively charged amino groups were shown to interact with solvent molecules.³⁸ Involvement of nonspecific electrostatic interactions of amino/guanidino- groups in the side chains was ruled out in these studies, given the fact that PNA:DNA helix parameters did not comply with such interactions.⁵⁷ We envisage that the nonspecific charge-charge interactions with the negatively charged phosphate groups may not be possible in the present studies also, due to additional constraint of the ethano bridge in the backbone.

Table 3. Effect of salt concentration on UV- T_m studies. t^H , t^{Am} and t^{Gu} are the modified ethano-PNA monomers with H, aminomethyl and guanidinomethyl substituents respectively

Entry No.	Code	Sequences (N' \rightarrow C')	UV- T_m ($^{\circ}$ C)			
			cDNA (10mM)	RNA (10mM)	cDNA (100mM)	RNA (100mM)
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	64.6	66.1	60.1	63.7
2.	PNA 2	H-CATTG t^H CACACT-Lys-NH ₂	49.7	59.4	46.2	52.3
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	51.8	60.4	46.6	57.2
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	53.7	62.7	48.8	59.4
5.	PNA 5	H-CAT t^H TGTCACACT-Lys-NH ₂	55.8	65.5	50.0	59.9
6.	PNA 6	H-CAT t^{Am} TGTCACACT-Lys-NH ₂	57.1	66.3	53.8	60.1
7.	PNA 7	H-CAT t^{Gu} TGTCACACT-Lys-NH ₂	60.6	68.1	56.3	60.7



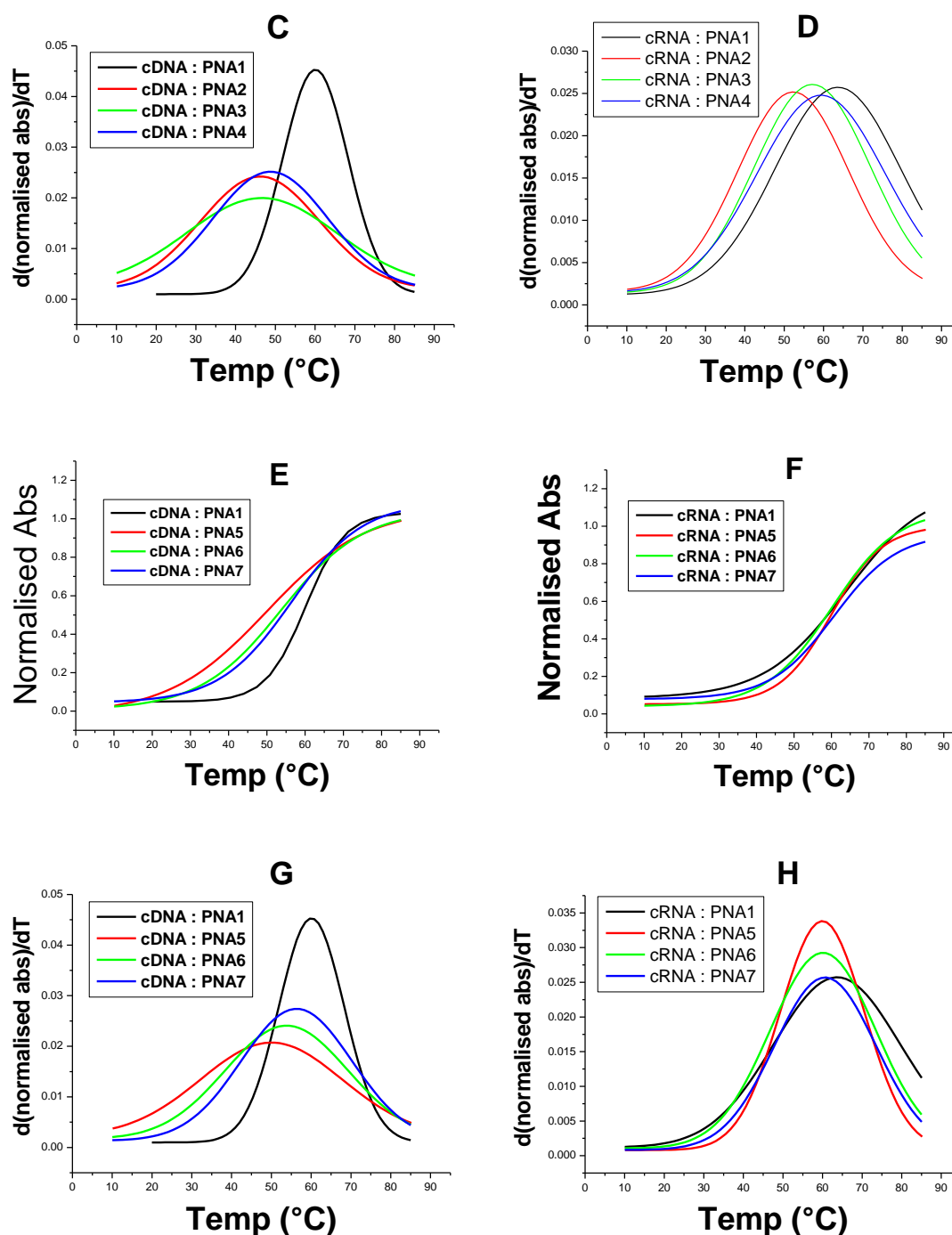


Figure 6. Effect of salt concentration. UV- T_m Sigmoidal Curves for (A) cDNA:PNA1/ PNA2/ PNA3/ PNA4 (B) cRNA:PNA1/ PNA2/ PNA3/ PNA4, first derivative curves of the sigmoidal curves for melting values (C) cDNA:PNA1/ PNA2/ PNA3/ PNA4 (D) cRNA:PNA1/ PNA2/ PNA3/ PNA4 Sigmoidal Curves for (E) cDNA:PNA1/ PNA5/ PNA6/ PNA7 (F) cRNA:PNA1/ PNA5/ PNA6/ PNA7, first derivative curves of the sigmoidal curves for melting values (G) cDNA:PNA1/ PNA5/ PNA6/ PNA7 (H) cRNA:PNA1/ PNA5/ PNA6/ PNA7

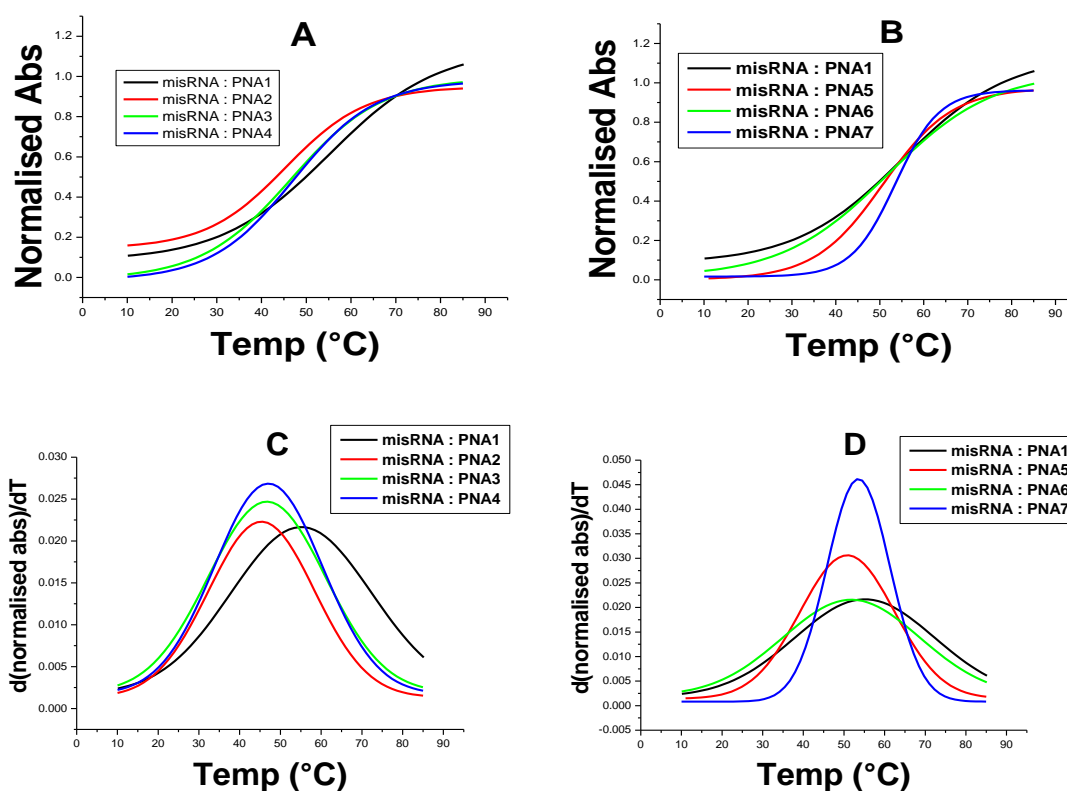
2B.3.2 UV- T_m studies on ethano-PNA with mismatch RNA

To further rule out the possibility of involvement of nonspecific electrostatic interactions at the expense of sequence specificity of binding, we conducted experiments with RNA sequence comprising a single mismatch. Thermal melting

experiments were carried out with mismatch RNA sequence (Table 4, Figure 7), to assess the effect on T_m values for H-ethano-PNA, Am-ethano-PNA and Gu-ethano-PNA. It was found that the all the PNA sequences (**PNA 2 - 7**) showed much lower melting values and the destabilization was slightly more than with the unmodified **PNA 1:mismatch RNA**. (Table 4, $\Delta T_{m\text{matched-mismatched}} - 11.0$ °C). All the modified duplexes formed using **PNA 2- 7** with mismatched RNA were found to be highly destabilized (Table 4, $\Delta T_{m\text{ matched-mismatched}} \sim 12 - 14.0$ °C) at either salt concentration studied. These experiments indicate that although the electrostatic properties of modified PNA may affect their preorganized structure and consequently may contribute to the stability of the duplexes, it is not at the expense of nucleobases sequence fidelity.

Table 4. UV- T_m studies on ethano-PNA with mismatch RNA. t^H , t^{Am} and t^{Gu} are the modified ethano-PNA monomers with H, aminomethyl and guanidinomethyl substituents respectively

Entry No.	Code	Sequences (N' \rightarrow C')	UV - T_m (°C)			
			RNA (10mM)	RNA (100mM)	misRNA (10mM)	misRNA (100mM)
1.	PNA 1	H-CATTGTCACACT-Lys-NH ₂	66.1	63.7	55.1	53.0
2.	PNA 2	H- CATTG t^H CACACT-Lys-NH ₂	59.4	52.3	45.4	42.2
3.	PNA 3	H-CATTG t^{Am} CACACT-Lys-NH ₂	60.4	57.2	46.7	44.7
4.	PNA 4	H-CATTG t^{Gu} CACACT-Lys-NH ₂	62.7	59.4	47.1	46.3
5.	PNA 5	H-CA t^H TGTCACACT-Lys-NH ₂	65.5	59.9	50.8	43.3
6.	PNA 6	H-CA t^{Am} TGTCACACT-Lys-NH ₂	66.3	60.1	51.8	45.7
7.	PNA 7	H-CA t^{Gu} TGTCACACT-Lys-NH ₂	68.1	60.7	53.6	46.8



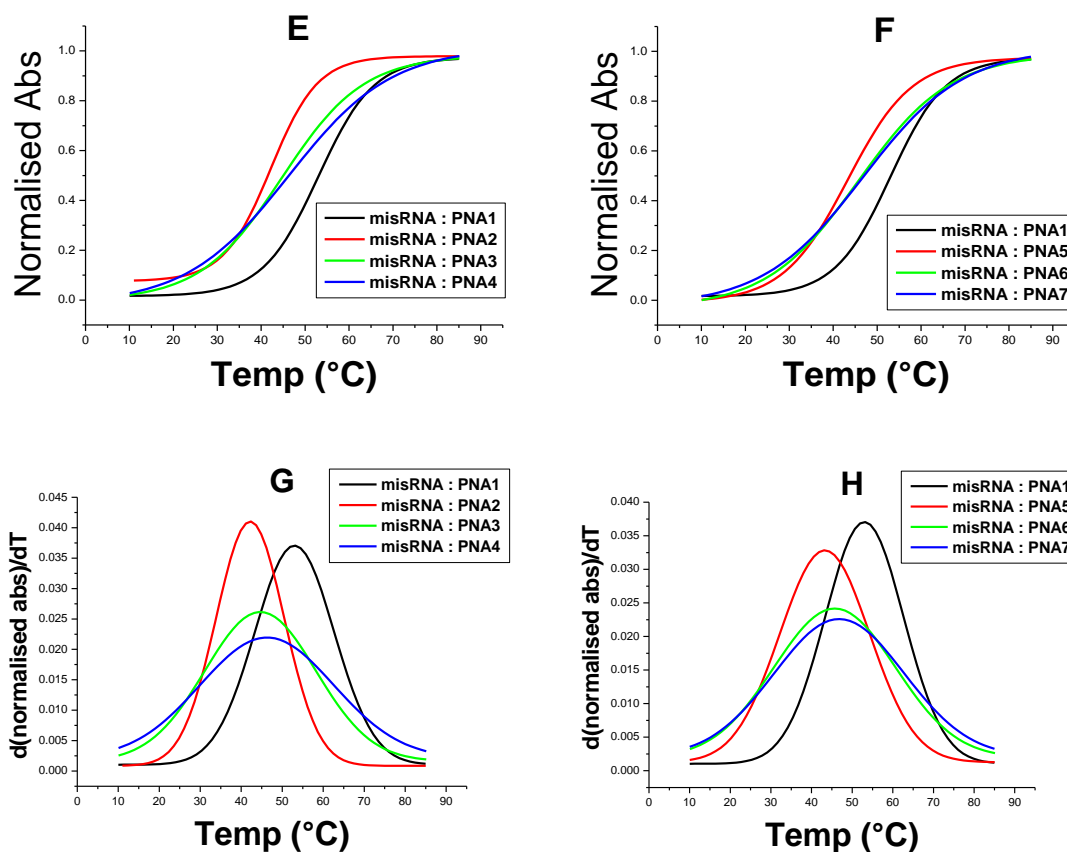


Figure 7. Thermal melting studies with mismatchRNA. Sigmoidal Curves for 10 mM (A) misRNA:PNA1/ PNA2/ PNA3/ PNA4 (B) misRNA:PNA1/ PNA5/ PNA6/ PNA7, first derivative curves of the sigmoidal curves for melting values (C) misRNA:PNA1/ PNA2/ PNA3/ PNA4 (D) misRNA:PNA1/ PNA5/ PNA6/ PNA7. Sigmoidal Curves for 100 mM (E) misRNA:PNA1/ PNA2/ PNA3/ PNA4 (F) misRNA:PNA1/ PNA5/ PNA6/ PNA7, first derivative curves of the sigmoidal curves for melting values (G) misRNA:PNA1/ PNA2/ PNA3/ PNA4 (H) misRNA:PNA1/ PNA5/ PNA6/ PNA7

2B.4 Conformational analysis of the ethano-PNA:cDNA and ethano-PNA:RNA

In order to examine if the ethano-PNAs cause any significant differences on the overall conformational features of the duplexes, the CD studies of PNA complexes with cDNA as well as RNA were undertaken. A maximum at 260 nm and a minimum at 240 nm typically are seen for right handed helical structures.^{47, 58} Base stacking interactions are reflected in the amplitude of the CD signal at 260 nm in PNA:DNA/RNA duplexes.⁴⁷ The CD profile of the derived PNA oligomers (PNA 2-7, Figure 8A and B) with the ethano-PNA modification when duplexed with cDNA showed similar CD pattern as that of the unmodified PNA1:DNA duplex (Figure 8A and B), although the amplitude of the CD-signal was comparatively less in each case. The CD profile of the unmodified PNA 1 when complexed with RNA showed a positive maximum signal at 260 nm and minima at 240 nm and 290 nm (Figure 8C and D). In this case the amplitude of the CD signal at 260 nm was drastically reduced when the ethano-bridged PNA monomers were introduced (PNA 2-7, Figure 8C and D) compared with the

unmodified PNA 1:RNA complex. The reduction observed in the minimum at 240 nm was indicative of tightened helical pitch of PNA:RNA complexes compared with the PNA:DNA complexes.^{47, 59} The CD Analysis of single strands were also undertaken for any conformation of the PNAs (Figure (9)). Thus, no significant change in CD patterns was observed in either the DNA or RNA duplexes and the fact suggests that the overall structural features of the duplexes with either DNA or RNA remained unchanged. The reduced amplitude of the CD signal at 260 nm could be because of reduced base stacking interactions due to the additional constraint in the backbone, in contrast to the open chain charged PNA analogues.⁴⁷ The reduced binding strength of these modified oligomers compared to the unmodified PNA could be therefore because of less effective stacking interactions. The hydrogen bonding interactions were retained as seen by nucleobases sequence fidelity of binding. Nevertheless, the modified oligomers were capable of inducing differential duplex strength while binding with RNA and DNA backbones.

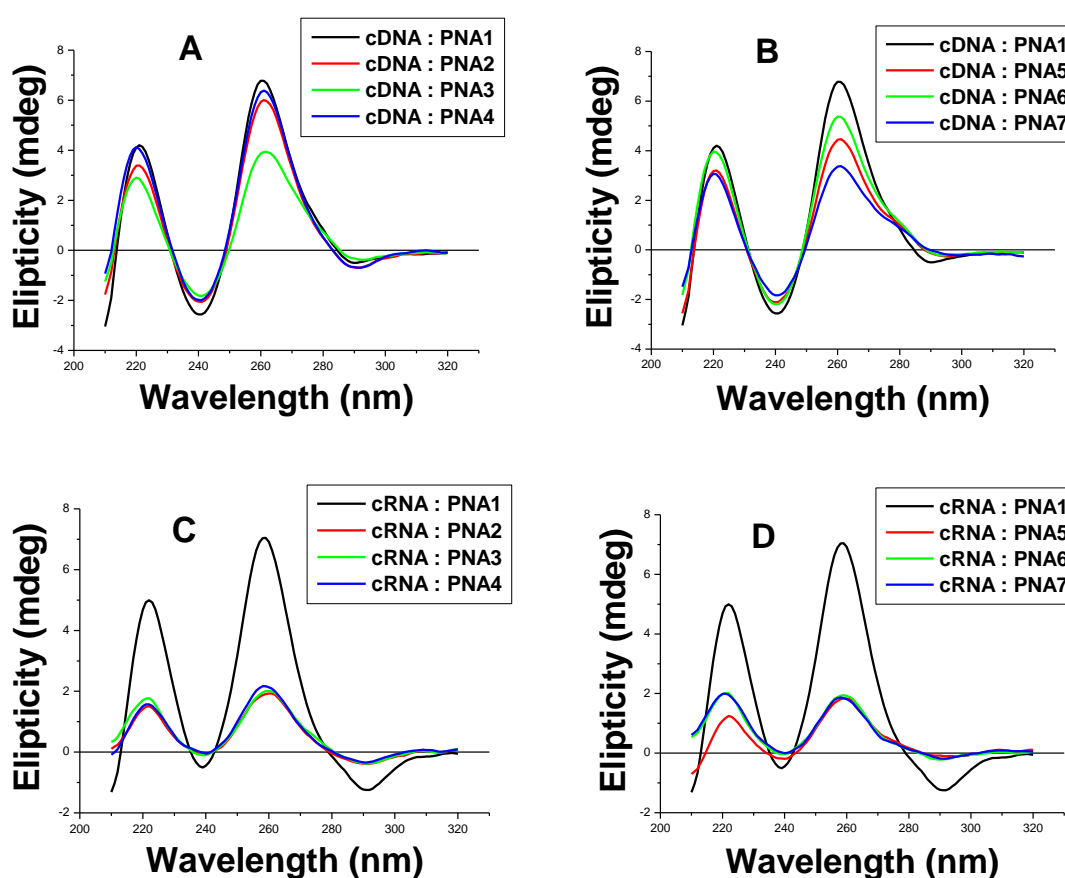


Figure 8. CD Profiles of Duplexes. (A) cDNA:PNA1/ PNA2/ PNA3/ PNA4 (B) cDNA:PNA1/ PNA5/ PNA6/ PNA7 (C) cRNA:PNA1/ PNA2/ PNA3/ PNA4 (D) cRNA:PNA1/ PNA5/ PNA6/ PNA7

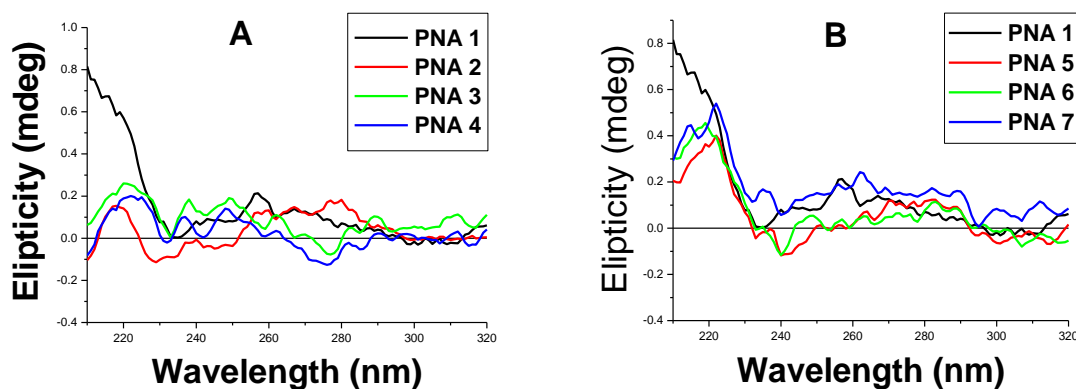


Figure 9. CD Profiles of single strands (A) PNA1/ PNA2/ PNA3/ PNA4 (B) PNA1/ PNA5/ PNA6/ PNA7

2B.5 Experimentation

UV-T_m Measurements: The concentrations were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases of DNA/RNA/PNA. The experiments were performed at 1 μ M concentrations of each strand. The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM and 100 mM) and were annealed by keeping the samples at 90°C for 5 min followed by slow cooling to room temperature and then to 0°C. Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Varian Cary 300 spectrophotometer scanning from 10-85°C at a ramp rate of 0.5°C per minute. Experiments were repeated at least thrice and the data were processed using Origin software 6.1. T_m (°C) values were derived from the first derivative curves and are accurate to within $\pm 0.3^\circ\text{C}$.

CD Analysis of the oligomers: The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM) and were annealed by keeping the samples at 90°C for 5 min followed by slow cooling to room temperature. The experiments were performed at 1 μ M concentrations of each strand. All the CD spectra were recorded at room temperature with Jasco J-715 spectropolarimeter. All spectra represent an average of at least 8 scans recorded from 320 nm to 210 nm at a rate of 100 nm per minute in a 1 cm path length cuvette. All spectra were processed using Origin software 6.1, baseline subtracted and smoothed using a 5 point adjacent averaging algorithm.

2B.6 Summary

- ❖ A novel conformationally constrained PNA analogue was designed which would exert restrictions in the aminoethyl segment of PNA such that the dihedral angle β is restricted for RNA selective binding.
- ❖ The modified oligomers containing ethano-PNA monomers did show preferential binding towards RNA than cDNA.
- ❖ The modification also showed positional dependence as well as substitution dependence binding as only **PNA 7** could stabilize the duplex better than the unmodified PNA.
- ❖ The destabilizing effect and preferential binding towards RNA could be accounted by the reduced stacking interactions as observed from the decreased amplitudes of the signals in the CD studies.

2B.7 References

1. R. Kole, A. R. Krainer and S. Altman, *Nature Reviews Drug Discovery*, 2012, 11, 125-140.
2. J. K. Watts, G. F. Deleavey and M. J. Damha, *Drug Discovery Today*, 2008, 13, 842-855.
3. F. Ahmad and S. A. Hashsham, *Analytica Chimica Acta*, 2012, 733, 1-15.
4. C. D. Chin, T. Laksanasopin, Y. K. Cheung, D. Steinmiller, V. Linder, H. Parsa, J. Wang, H. Moore, R. Rouse, G. Umvilighozo, E. Karita, L. Mwambarangwe, S. L. Braunstein, J. van de Wijgert, R. Sahabo, J. E. Justman, W. El-Sadr and S. K. Sia, *Nature Medicine*, 2011, 17, 1015-U1138.
5. R. de la Rica and M. M. Stevens, *Nature Nanotechnology*, 2012, 7, 821-824.
6. C. M. Micklitsch, B. Y. Oquare, C. Zhao and D. H. Appella, *Analytical Chemistry*, 2013, 85, 251-257.
7. T. P. Prakash, *Chemistry & Biodiversity*, 2011, 8, 1616-1641.
8. T. P. Prakash, M. Manoharan, A. M. Kawasaki, A. S. Fraser, E. A. Lesnik, N. Sioufi, J. M. Leeds, M. Teplova and M. Egli, *Biochemistry*, 2002, 41, 11642-11648.
9. M. Petersen and J. Wengel, *Trends Biotechnol.*, 2003, 21, 74-81.
10. E. Lescrinier, R. Esnouf, J. Schraml, R. Busson, H. A. Heus, C. W. Hilbers and P. Herdewijn, *Chemistry & Biology*, 2000, 7, 719-731.
11. E. P. Stirchak, J. E. Summerton and D. D. Weller, *Nucleic Acids Res.*, 1989, 17, 6129-6141.
12. J. Summerton, *Biochimica Et Biophysica Acta-Gene Structure and Expression*, 1999, 1489, 141-158.
13. J. Summerton and D. Weller, *Antisense & Nucleic Acid Drug Development*, 1997, 7, 187-195.
14. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, 254, 1497-1500.
15. V. A. Kumar, *Eur. J. Org. Chem.*, 2002, 2021-2032.
16. V. A. Kumar and K. N. Ganesh, *Accounts Chem. Res.*, 2005, 38, 404-412.
17. B. Hyrup and P. E. Nielsen, *Bioorg. Med. Chem.*, 1996, 4, 5-23.
18. M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, 365, 566-568.

19. J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble and L. E. Babiss, *Science*, 1992, 258, 1481-1485.
20. P. E. Nielsen, *Q. Rev. Biophys.*, 2005, 38, 345-350.
21. S. C. Brown, S. A. Thomson, J. M. Veal and D. G. Davis, *Science*, 1994, 265, 777-780.
22. M. Eriksson and P. E. Nielsen, *Nat. Struct. Biol.*, 1996, 3, 410-413.
23. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *J. Org. Chem.*, 2004, 69, 5725-5734.
24. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *Chem. Commun.*, 2004, 860-861.
25. T. Govindaraju, R. G. Gonnade, M. M. Bhadbhade, V. A. Kumar and K. N. Ganesh, *Org. Lett.*, 2003, 5, 3013-3016.
26. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *J. Org. Chem.*, 2004, 69, 1858-1865.
27. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *J. Am. Chem. Soc.*, 2005, 127, 4144-4145.
28. T. Govindaraju, V. Madhuri, V. A. Kumar and K. N. Ganesh, *J. Org. Chem.*, 2006, 71, 14-21.
29. P. Lagriffoule, P. Wittung, M. Eriksson, K. K. Jensen, B. Norden, O. Buchardt and P. E. Nielsen, *Chem.-Eur. J.*, 1997, 3, 912-919.
30. M. C. Myers, M. A. Witschi, N. V. Larionova, J. M. Franck, R. D. Haynes, T. Hara, A. Grajkowski and D. H. Appella, *Org. Lett.*, 2003, 5, 2695-2698.
31. J. K. Pokorski, M. A. Witschi, B. L. Purnell and D. H. Appella, *J. Am. Chem. Soc.*, 2004, 126, 15067-15073.
32. C. A. G. Haasnoot, F. Deleeuw, H. P. M. Deleeuw and C. Altona, *Biopolymers*, 1981, 20, 1211-1245.
33. C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber and L. Moroder, *Angew. Chem.-Int. Edit.*, 2001, 40, 923-925.
34. I. R. Babu and K. N. Ganesh, *J. Am. Chem. Soc.*, 2001, 123, 2079-2080.
35. M. Umashankara, I. R. Babu and K. N. Ganesh, *Chem. Commun.*, 2003, 2606-2607.

36. M. Umashankara, M. Nanda, M. Sonar and K. N. Ganesh, *Chimia*, 2012, 66, 936-940.
37. A. Banerjee and V. A. Kumar, *Bioorg. Med. Chem.*, 2013, 21, 4092-4101.
38. N. Tilani, S. De Costa and J. M. Heemstra, *PLoS One*, 2013, 8, 8.
39. M. Hashimoto, Y. Eda, Y. Osanai, T. Iwai and S. Aoki, *Chem. Lett.*, 1986, 893-896.
40. S. S. Gokhale and V. A. Kumar, *Org. Biomol. Chem.*, 2010, 8, 3742-3750.
41. L. Christensen, R. Fitzpatrick, B. Gildea, K. H. Petersen, H. F. Hansen, T. Koch, M. Egholm, O. Buchardt, P. E. Nielsen, J. Coull and R. H. Berg, *Journal of Peptide Science*, 1995, 1, 175-183.
42. R. Knorr, A. Trzeciak, W. Bannwarth and D. Gillessen, *Tetrahedron Letters*, 1989, 30, 1927-1930.
43. G. R. Matsueda and J. M. Stewart, *Peptides*, 1981, 2, 45-50.
44. E. Kaiser, Colescot.Rl, Bossinge.Cd and P. I. Cook, *Anal. Biochem.*, 1970, 34, 595-598.
45. F. Milletti, *Drug Discovery Today*, 2012, 17, 850-860.
46. U. Koppelhus and P. E. Nielsen, *Advanced Drug Delivery Reviews*, 2003, 55, 267-280.
47. B. Sahu, I. Sacui, S. Rapireddy, K. J. Zanotti, R. Bahal, B. A. Armitage and D. H. Ly, *J. Org. Chem.*, 2011, 76, 5614-5627.
48. K. Feichtinger, H. L. Sings, T. J. Baker, K. Matthews and M. Goodman, *J. Org. Chem.*, 1998, 63, 8432-8439.
49. M. S. Bernatowicz, Y. L. Wu and G. R. Matsueda, *J. Org. Chem.*, 1992, 57, 2497-2502.
50. Y. D. Zhang and A. J. Kennan, *Org. Lett.*, 2001, 3, 2341-2344.
51. Y. F. Yong, J. A. Kowalski and M. A. Lipton, *J. Org. Chem.*, 1997, 62, 1540-1542.
52. H. Yajima, N. Fujii, H. Ogawa and H. Kawatani, *Journal of the Chemical Society, Chemical Communications*, 1974, 107-108.
53. J. D. Puglisi and I. Tinoco, *Method Enzymol.*, 1989, 180, 304-325.
54. D. Ittig, A. B. Gerber and C. J. Leumann, *Nucleic Acids Res.*, 2011, 39, 373-380.
55. P. S. Lonkar and V. A. Kumar, *Bioorg. Med. Chem. Lett.*, 2004, 14, 2147-2149.
56. S. Tomac, M. Sarkar, T. Ratilainen, P. Wittung, P. E. Nielsen, B. Norden and A. Graslund, *J. Am. Chem. Soc.*, 1996, 118, 5544-5552.

57. J. I. Yeh, B. Shivachev, S. Rapireddy, M. J. Crawford, R. R. Gil, S. C. Du, M. Madrid and D. H. Ly, *J. Am. Chem. Soc.*, 2010, 132, 10717-10727.
58. W. C. Johnson, N. Berova, K. Nakanishi and R. W. Woody, *Wiley-VCH: New York*, 2000.
59. E. B. Nielsen and Schellma.Ja, *J. Phys. Chem.*, 1967, 71, 2297-2304.

CHAPTER 3

β, γ - bisubstituted PNA

Section A

Synthesis of monomers and incorporation in aegPNA sequences

Section B

Biophysical studies of modified PNA towards comp. DNA/RNA

Section A

Synthesis of monomers and incorporation in *aeg*PNA sequences**3A Introduction**

Antisense technology is a potentially powerful method for controlling gene expression.¹ The concept behind the antisense therapeutics is simple and rational: to inhibit the expression of a specific gene at the mRNA level by using complementary oligonucleotides, called antisense oligonucleotides, thereby blocking the expression of the protein encoded by the target RNA. A major barrier to the successful development of antisense technology has been the inability to develop nucleic acid analogues that can traverse the cell-membrane and bind selectively to the intended RNA targets.²

Peptide nucleic acids (PNA),³ a synthetic analogue of DNA developed by Nielsen *et. al.* in 1991, as described in previous chapters, rekindled optimism that the uptake problem would finally be resolved because of the neutral backbone. It has been generally accepted that neutral molecules can penetrate the cell membrane more effectively than charged species. This optimism, however, soon faded with the discovery that PNA is not readily taken up by mammalian cells, despite the neutral backbone.³ While several PNA-based applications^{4,5} have been successfully developed, the full potential of PNA for *in vivo* applications has not yet been realized. Considerable efforts have been vested in the last decade trying to develop means to transport PNA into cells. Several strategies (Figure 1), including microinjection,⁶ electroporation,⁷ DNA-mediated transduction,⁸ PNA-DNA chimera⁹ and covalent attachment to cell-penetrating peptides (CPPs),^{10, 11} have been developed, but for the most part they are limited to small-scale experimental setups.

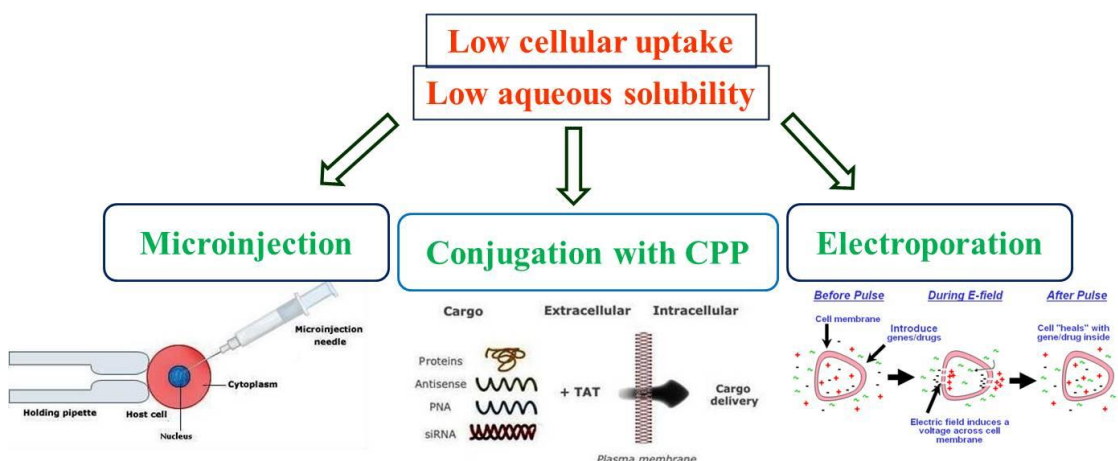


Figure 1. Several strategies to transport cargos into cells

3A.1 Rationale and objectives of the present work

One of the possible strategies for PNA delivery is to link PNAs to polylysine (K) or a polyarginine (R) tails, based on the observation that this cell-membrane-penetrating oligopeptides are able to facilitate uptake of conjugated molecules.¹² Since their discovery, many modifications of the original PNA backbones have been proposed in order to improve performances in terms of affinity and specificity. Modification of the PNA backbone with positively charged groups has also been demonstrated to enhance cellular uptake and thus PNA efficiency.^{13, 14} In respect to the delivery issue Fabani and Gait administered anti-miR PNAs by electroporation.¹⁵ In the second set of experiments, Fabani and Gait showed that miR inhibition can be achieved without the need for transfection or electroporation, by conjugating the PNA to the cell-penetrating peptide (CPP) R6-Penetratin, or merely by linkage to just four Lys residues, highlighting the potential of PNAs for future therapeutic applications as well as for studying miR function.¹⁶ In a parallel work, Oh *et. al.* described the effectiveness of miR targeting by PNA-peptide conjugates, using a series of CPPs as carriers, including R6 pen, Tat, a four-Lys sequence, and transportan.¹⁷ They found that best conditions were obtained with cationic peptides, and in particular the Tat-modified peptide RRRQRRKKRR. The latter approach, although promising since no delivery system is required, is presently limited by uptake efficiency and intracellular distribution. Most PNAs delivered this way are trapped in the endocytic vesicles.⁵ In addition, since these PNA-CPP constructs are amphipathic in nature, they could cause cytotoxic effects since most amphipathic peptides found in nature, such as magainin and mellitin, are known to be cytotoxic.¹⁸

As an alternative to peptide conjugation, the PNAs could be designed to be intrinsically positively charged. We and others have embarked to address this problem and have devised several interesting structural analogues of PNA based on the position of a substituent in the PNA backbone (α -PNA, β -PNA, and γ -PNA).

α -PNA: Nielsen *et. al.* reported the first α -chiral PNA in 1994, in which glycine moiety of the PNA backbone was substituted by alanine.¹⁹ Both the *L*- and *D*- forms of the chiral monomers were synthesized from *L*- or *D*- alanine and incorporated into oligomers. Thermal stability of PNA-DNA duplex containing *D*- form monomers was similar to that of the original PNA with a glycine backbone, whereas a reduction was observed when *L*- form monomers were incorporated. Reductive amination is the most used procedure to obtain a α -chiral aminoethyl glycine backbone and a variety of α -

chiral PNAs bearing side chains from amino acids have been prepared by this method.²⁰⁻²² Alternatively, solid-phase synthesis of chiral monomers has been reported.²³

Introduction of a side chain at the α -position caused moderate destabilization of PNA-DNA duplexes relative to the unmodified PNA.²⁴ However, incorporation of three consecutive D-Lys-based α -PNA monomers greatly improved the sequence selectivity of PNA oligomer with moderate expense to DNA binding.²⁵ The lysine side chain of α -Lys monomer was also used as a handle for site-specific incorporation of functional molecules into PNA oligomers.²⁶ Nielsen *et al.* also reported the synthesis of a series of α -PNAs bearing glycosylated side chains at the α -position and demonstrated their selective biodistribution.²⁷ Metzler-Nolte *et al.* reported the synthesis of C-linked glycosylated α -PNA and its successful incorporation into a PNA oligomer using Fmoc chemistry.²⁸ Czerny *et al.* recently introduced phosphonic ester into the α -side chain (*pePNA*).²⁹ Armitage *et al.* used α -PNAs as a tool for studying interactions between small molecules and the PNA-DNA duplex.³⁰ Recently, Ganesh *et al.* reported that the incorporation of *gem*-dimethyl substituted PNA monomers into PNA oligomers increases the T_m of PNA-DNA duplexes.³¹ Interestingly, incorporation of the homologous aminopropyl-(α,α -dimethyl)glycyl (*apdmg*)-PNA monomers also improved DNA binding. D-arginine-based PNA is a special case. Although the guanidinium group of α -Arg PNA (α -GPNA) is positively charged under physiological pH, incorporation of one unit of α -GPNA into PNA decamer destabilized the PNA-DNA duplex; however, incorporation of multiple units of D-form α -GPNAs at every other position improved DNA binding affinity. By spacing the arginine units, electrostatic attraction between the guanidinium groups and the phosphates seems to overcome the adverse effect of steric repulsion arising from arginine side chains themselves.¹² Recently, α -GPNA was used for PNA microarray and its higher sequence selectivity has been demonstrated.³²

Ly *et al.* introduced a guanidinium group into the PNA backbone³³ based on the finding that the HIV-1 Tat transcription domain (GRKKRRQRRR),³⁴ arginine-rich peptides,³⁵ and a homoarginine peptoid exhibited marked cellular uptake properties.³⁶ Cellular uptake of fully modified α -GPNA decamer derived from L-arginine was evaluated with human HCT116 (colon) and Sao-2 (osteosarcoma) cell lines. α -GPNA permeated the cell membrane and appeared to localize specifically in the nucleus. Since there was no difference in the uptake properties of α -GPNA at 37 °C and 4 °C, the

uptake mechanism of α -GPNA was neither endocytosis driven nor receptor mediated. Membrane flipping has been proposed as mechanism.³³ α -GPNA (Figure 2B) oligomers in which unmodified PNA units and α -GPNA units were alternated at every other position possessed optimal hybridization and cellular uptake properties. Cellular uptake of α -GPNA octamer and dodecamer containing alternate backbone derived from D-arginine was also addressed and found to be efficient for HeLa and ES cells.³⁷

γ -PNA: Although the first γ -chiral PNA monomer was reported in 1994,³⁸ oligomers carrying γ -chiral units did not appear until 2005.³⁹⁻⁴² In 2005, Appella *et al.* reported PNA oligomers carrying γ -Lys PNA as a handle for fluorene incorporation aiming at developing quencher-free molecular PNA beacons.⁴⁰ Incorporation of γ -modified PNA slightly stabilized PNA-DNA duplexes relative to the unmodified duplex. Moreover, the ability of γ -PNA to discriminate single-base mismatches was superior to that of the corresponding unmodified PNA. In the subsequent paper,⁴³ the authors demonstrated that γ -Lys PNA was a versatile scaffold to attach various functional molecules including peptides and fluorophores.

Detailed spectroscopic studies of serine- or alanine-based γ -PNAs carried out by Ly *et al.* revealed that a simple γ -backbone modification preorganized single-stranded PNA oligomers into a right-handed helical structure that was very similar to that of PNA-DNA duplex. The γ -PNAs bound to DNA with very high affinity and high sequence selectivity. Helical induction was sterically driven and stabilized by base stacking. The fully γ -modified decameric PNA formed exceptionally stable PNA-DNA duplex with an increase of 19 °C of the T_m compared to the unmodified PNA.⁴⁴ They have also reported a variety of γ -PNAs bearing side chains derived from amino acids and have studied their cellular internalization.⁴⁵

γ -GPNA,⁴⁶ the second-generation GPNA (Figure 2B), reported by Ly *et al.* was prepared from inexpensive Boc-L-Lysine and had a homo-arginine side chain at the γ -position. Stabilization of PNA-DNA duplexes by incorporation of γ -GPNA units was significant compared to α -GPNAs. The authors attributed the enhancement of thermal stability to conformational preorganization, not to electrostatic interactions between guanidinium and phosphate groups because of the lack of salt dependence of ΔT_m . A fully alternate γ -GPNA decamer was taken up by HeLa cells and the uptake efficiency was comparable to that of the TAT transduction domain. The authors attributed the improved cellular uptake of γ -GPNA to their helical conformation. Manicardi *et al.* pointed out that overall anti-miR activity of GPNAs in leukemic K562 cells is a

combination of cellular uptake and RNA binding.⁴⁷ The substituents in the PNA backbone play a role not only in cellular uptake, but also in the mechanism of miR recognition and inactivation. They also reported modified PNA (α - and/or γ -substitutions) with a nuclear localization signal (NLS) sequence embedded in the PNA backbone (chiral box and extended chiral box, Figure 2C).^{13, 48} Ganesh *et al.* reported the synthesis of chiral PNAs (*am*-PNAs) with cationic aminomethyl groups at the α - or γ -position of the PNA backbone.^{49, 50} The *am*-PNAs could traverse the cell membrane of HeLa cells and localized into the nucleus.

β -PNA: Although some cyclic PNA analogues contain a chiral center at the β -position,⁵¹⁻⁵⁸ PNAs with a single substituent only at the β -position (β -PNAs) have not been synthesized until recently. In 2011, Sugiyama *et al.* reported the first β -PNA bearing a methyl group at the β -position.⁵⁹ A PNA monomer possessing a methyl group at the β -position was designed and both enantiomers, β -(*S*)- and β -(*R*)-configurations, were prepared. The CD spectra of β -PNAs indicated that β -(*S*)-PNA adopted a right-handed helical structure and β -(*R*)-PNA was left-handed; however, the induced right-handed structure of β -(*S*)-PNA did not contribute to the total PNA-DNA duplex stability. Even the β -(*R*)-PNA does not show any binding towards DNA. The benefit of the induced structure might be abolished by unfavorable steric interactions arising from β -methyl groups in the PNA-DNA hybrid duplex. Since the development of β -PNA is a new, unexplored field, more sophisticated designs may be possible.

Thus in this context we designed an *aeg*PNA monomer unit bearing bisubstitution of aminomethyl group at β - and γ - positions (Figure 2D). The intrinsic amino group was envisaged to increase the aqueous solubility as well as cellular uptake.

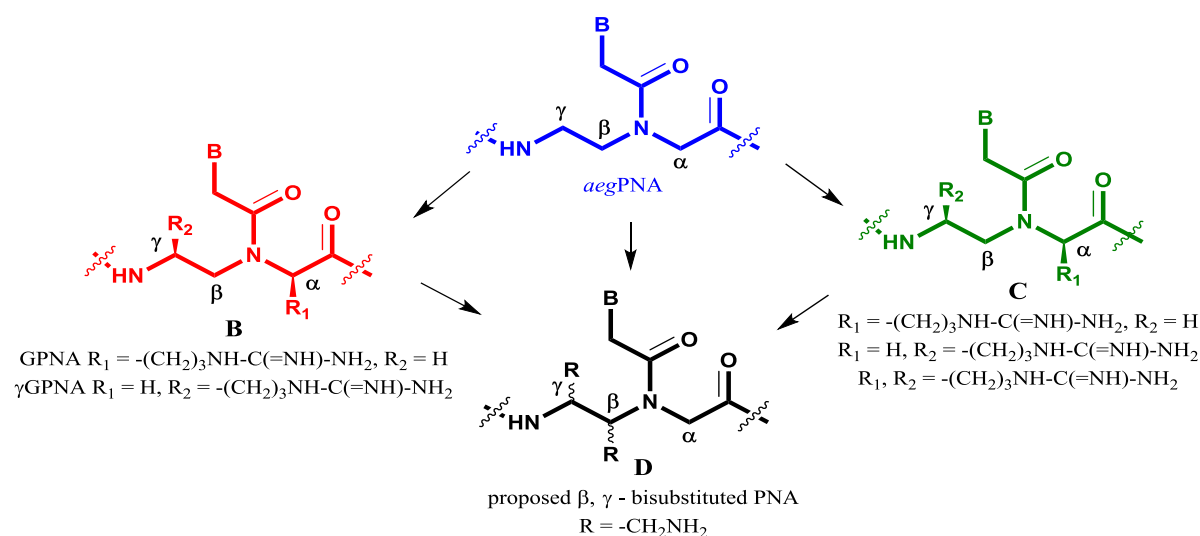
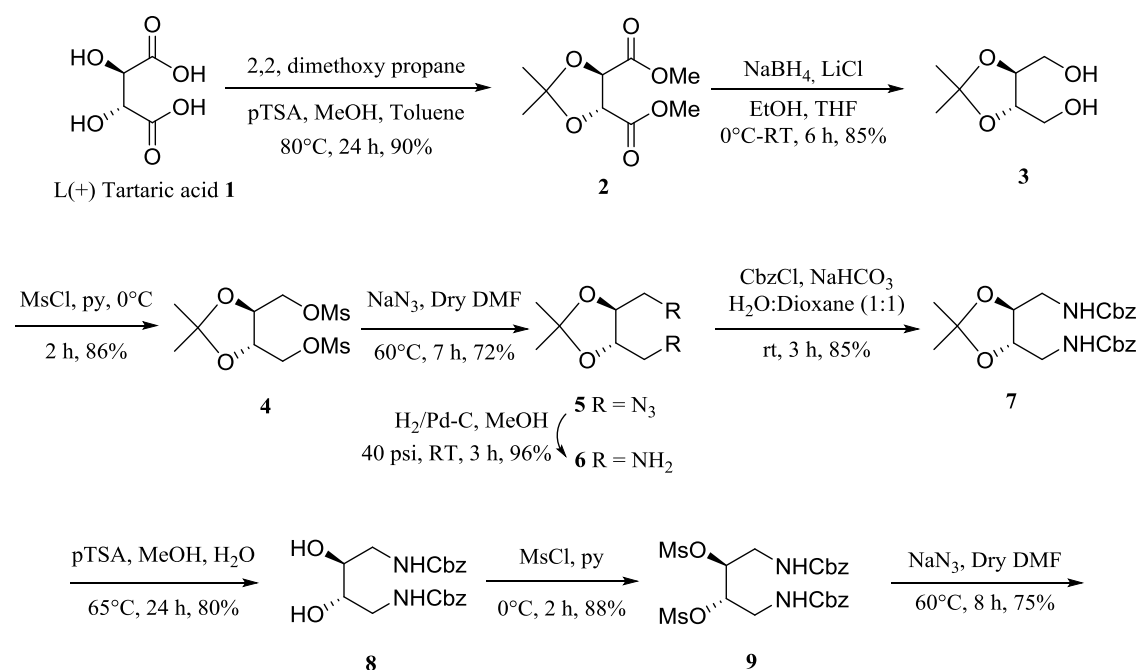
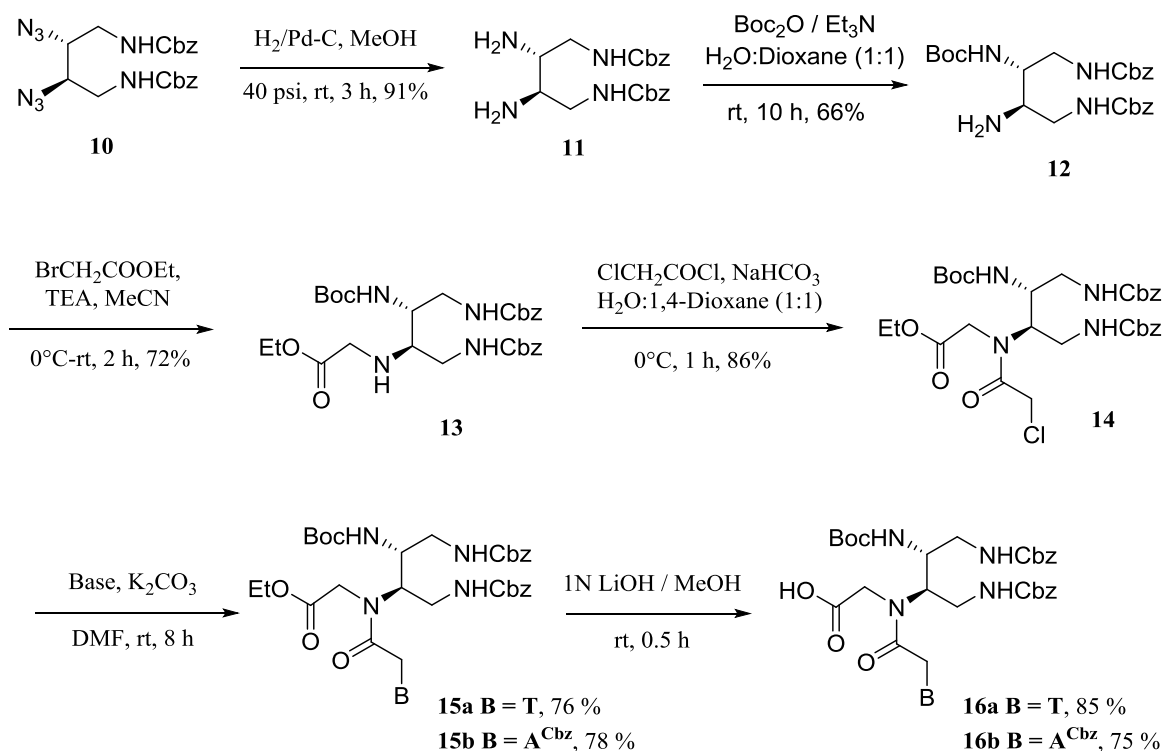


Figure 2. Rationale for the design of proposed β, γ -bisubstituted PNA

3A.2 Synthesis of β - and γ - bisubstituted PNA T/A^{cbz} monomer

The commercially available *L* (-) Tartaric acid **1** was converted to its acetonide protected dimethyl ester **2** according to the literature procedure.⁶⁰ Further, the esters of **2** were reduced to alcohol **3** with LiBH₄ generated *in-situ* with NaBH₄/LiCl. The hydroxyl group was mesylated using mesyl chloride in pyridine to get compound **4** and was converted to azide **5** by NaN₃ in DMF. The azido group in **5** was reduced to amine with H₂/ Pd-C in methanol. The resultant amine **6** was protected with Cbz group by using Cbz - Cl in 1,4-dioxane-water (1:1) and NaHCO₃ to obtain **7**. Deprotection of acetonide in compound **7** was carried out using *p*TSA in methanol to get free hydroxyl group in **8**. Mesylation of hydroxyl groups of **8** was again done by using mesyl chloride in pyridine to get compound **9**, followed by conversion to azide **10** by NaN₃ in DMF. The azido was similarly reduced to amine with H₂/ Pd-C in methanol at 40 psi. The diamine **11** obtained was subjected to mono Boc protection to get **12**. The other amine group was mono alkylated with ethyl bromoacetate in acetonitrile to get **13**. Acylation of **13** with chloroacetyl chloride in 1,4-dioxane-water (1:1) was carried out, maintaining the pH of the solution at ~8.0 with sat. NaHCO₃ solution during the addition of chloroacetyl chloride to obtain **14**. Nucleobases (thymine and adenine - Cbz protected adenine) were derivatized via nucleophilic substitution of chloro group to get the nucleobase attached compounds **15a** and **15b**. The ester was hydrolysed using 1N LiOH in MeOH to give the desired monomers **16a** and **16b**. Similar protocol was followed to get the other stereochemistry, in which synthesis was started from D (+) Tartaric acid.



Scheme 1: Synthesis of β - and γ -bisubstituted PNA T/A^{cbz} monomer

3A.3 Solid Phase PNA Synthesis

3A.3.1 Solid Phase PNA Synthesis with β - and γ -bisubstituted T/A^{cbz} monomer

The general protocol for the solid phase PNA synthesis has been discussed in the previous chapter (section 2A.3). Following the same protocol the PNA oligomers having the bisubstituted PNA monomers were synthesised. PNA oligomers were assembled by solid phase synthesis on MBHA resin functionalized with the first residue α N-Boc- ω N, 2-Cl Cbz Lysine, using TBTU / HOBt activation strategy. The synthesized sequences are shown in Table 1. The modified monomers were incorporated in pre-defined positions by solid phase synthesis to yield a backbone modification at the central position (Table 1, entries 4 - 5) and near the N-terminal of the aegPNA (Table 1, entries 6 - 9). Lysine residues were attached at the N-terminal of the unmodified PNA for comparison with the positive charge incorporated in the oligomers.

3A.3.2 Cleavage of the PNA Oligomers from the Solid Support

The oligomers were cleaved from the solid support, using TFMSA in the presence of TFA (Low, High TFMSA-TFA method), which yields oligomers with an amide group at their C-terminus. A cleavage time of 2 h at room temperature was found to be optimum. The side chain protecting groups for nucleobases were also cleaved during this cleavage process. After cleavage reaction, the oligomer was precipitated with dry diethyl ether.

Table 1. PNA sequences, HPLC retention time and MALDI-TOF mass
 t_L^{Am} / a_L^{Am} and t_D^{Am} / a_D^{Am} are the modified monomers derived from L- /D - tartaric acid resp.

Entry No.	Code	Sequences (N' → C')	HPLC t_R (min)	MALDI - ToF Masses	
				Calcd.	Obsd.
1.	PNA 1	H-AAACCGATTTCAG-Lys-NH ₂	12.41	3381.35	3384.09
2.	PNA 2	H- (K) ₂ -AACCGATTTCAG-Lys-NH ₂	12.71	3637.70	3639.01
3.	PNA 3	H- (K) ₄ -AACCGATTTCAG-Lys-NH ₂	12.86	3894.05	3895.98
4.	PNA 4	H- AACCGA t_L^{Am} TTCAG -Lys-NH ₂	12.47	3439.43	3439.88
5.	PNA 5	H- AACCGA t_D^{Am} TTCAG -Lys-NH ₂	12.48	3439.43	3440.39
6.	PNA 6	H- A a_L^{Am} CCGATTTCAG -Lys-NH ₂	12.49	3439.43	3439.52
7.	PNA 7	H- A a_D^{Am} CCGATTTCAG -Lys-NH ₂	12.44	3439.43	3439.26
8.	PNA 8	H- a_L^{Am} a_L^{Am} CCGATTTCAG -Lys-NH ₂	13.12	3497.51	3497.40
9.	PNA 9	H- a_D^{Am} a_D^{Am} CCGATTTCAG -Lys-NH ₂	13.27	3497.51	3497.83

3A.3.3 Purification and MALDI-TOF characterization of oligomers

The purity of the oligomers was checked by analytical RP-HPLC (C18 column, CH₃CN-H₂O- 0.1% TFA system), which showed more than 85-90% purity. A gradient elution method contained A = 5% Acetonitrile in water + 0.1% trifluoroacetic acid and B = 50% Acetonitrile in water + 0.1% trifluoroacetic acid (A to B = 100% in 30 min with a flow rate of 1.5 mL/min), and the eluent was monitored at 260 nm. Oligomers were subsequently purified by reverse phase HPLC on a C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis.

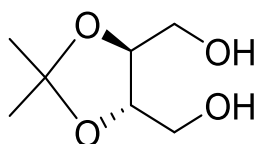
3A.4 Summary

- ❖ A new PNA monomer was designed and synthesized from two isomers, having substitutions at β - , γ - positions of the PNA backbone.
- ❖ The modified monomers were successfully incorporated into *aeg*PNA sequences at specific positions and were well characterized.

3A.5 Experimental

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 distilled under an inert atmosphere according to the literature procedures. Reactions were monitored by thin layer chromatography. TLC was performed using TLC aluminium sheets pre-coated with silica gel 60 F254 (Merck). TLCs were visualized with UV light and iodine spray and/or by ninhydrin treatment then heating. Silica gel 60-120 mesh was used for column chromatography. ^1H and ^{13}C NMR spectra were recorded on 200 MHz and 50 MHz respectively. Chemical shifts are given in δ (ppm) scale. The specific rotation values were determined from Bellingham + Stanley Ltd. ADP 220 polarimeter. IR spectra were recorded in Perkin-Elmer Spectrum One FTIR Spectrometer. The HRMS data's were obtained from Thermo Fisher Scientific Q Exactive mass spectrometer. The crude PNAs were purified on a semi-preparative C18 column attached to a Waters HPLC system. For all the MALDI-TOF, spectra recorded on AB SCIEX TOF/TOF™ 5800 System; CHCA (α -cyano-4-hydroxycinnamic acid) was used as the matrix.

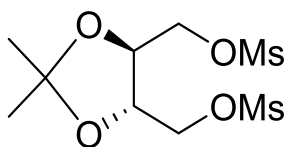
3 ((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) dimethanol



NaBH_4 (3.0 g) was suspended in dry THF (30 mL) and absolute alcohol (40 mL) and stirred for 30 mins in ice bath. At 0 °C LiCl (3.0 g) was added in two portions and the content was stirred for another 1 h in ice bath for *in situ* generation of LiBH_4 . The solution turned milky due to formation of NaCl . To this mixture, compound **2** (2.5 g, 15.4 mmol) dissolved in absolute alcohol (40 mL) was added at 0 °C dropwise. The reaction mixture was stirred for 8 h. NH_4Cl solution was then added for maintaining pH ~7.0 and the solvents were removed under reduced pressure. The residue was dissolved in ethyl acetate (125 mL). Organic layer was washed with water (2 X 30 mL) followed by brine (2 X 20 mL) and dried over anhydrous Na_2SO_4 . It was evaporated under vacuo and the crude product obtained was purified by column chromatography (15 % ethyl acetate-petroleum ether) to give **3** ((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) dimethanol. The same procedure was repeated four times with same amount of compound **2** to get compound **3**. (6.3 g, 85 %) ^1H NMR (CDCl_3) δ : 1.44 (s, 6H), 2.27 (s, 2H, D_2O exch.), 3.67-3.84 (q, 2H; q, 2H), 4.01 (t, 2H). ^{13}C NMR (CDCl_3) δ : 26.95, 61.98, 78.07, 109.25. IR (ν_{max} , cm^{-1}) (CHCl_3): 793, 845, 971, 1053, 1074, 1118, 1157,

1214, 1257, 1372, 1381, 1453, 2937, 2988, 3435. $[\alpha]_D^{27}$ - 26 (from L-isomer). + 28 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₇H₁₅O₄: 163.1850, found: 1.1850.

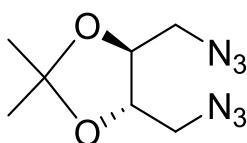
(4) ((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) bis (methylene) dimethanesulfonate



The pure compound **3** (6.3 g, 38.8 mmol) obtained in the previous step was dissolved in dry pyridine (60 mL) and stirred for 15 mins at 0 °C. Mesyl chloride (6.4 mL, 81.6 mmol) was added drop by drop to this solution at 0 °C with stirring and the reaction mixture was left for 2 h.

After that pyridine was removed under reduced pressure and the residue was redissolved in ethyl acetate (100 mL). Organic layer was washed with water (2 X 30 mL), sat. NaHCO₃ (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na₂SO₄. The organic layer was concentrated in vacuo to give, the crude product which was purified by column chromatography (10 % ethyl acetate-petroleum ether) to give **(4) ((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) bis (methylene) dimethanesulfonate**. (10.6 g, 86 %). ¹H NMR (CDCl₃) δ: 1.44 (s, 6H), 3.09 (s, 6H), 4.19 (t, 2H), 4.37 (q, 2H; q, 2H),. ¹³C NMR (CDCl₃) δ: 26.74, 37.60, 67.87, 75.01, 110.86. IR (*v*_{max}, cm⁻¹) (CHCl₃): 793, 907, 1160, 1173, 1217, 1257, 1352, 1456, 1651, 2877, 2929. $[\alpha]_D^{27}$ - 16 (from L -isomer). + 14 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₉H₁₉S₂O₈: 319.3550, found: 319.3551

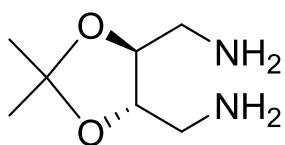
(5) (4*S*, 5*S*) - 4, 5 - bis(azidomethyl) - 2, 2 – dimethyl - 1, 3 - dioxolane



The compound **4** (2.5 g, 7.8 mmol) was dissolved in dry DMF (25 mL) and NaN₃ (5.1 g, 78.5 mmol) was added to it. The reaction mixture was stirred at 60 °C for 7 h. After completion of reaction DMF was removed under reduced pressure and the residue was dissolved in ethyl acetate (100 mL). Organic layer was washed with water (2 X 25 mL) followed by brine (2 X 20 mL). The organic layer was kept over anhy. Na₂SO₄ and then concentrated in vacuo. The crude compound obtained was purified by column chromatography (10 % ethyl acetate-petroleum ether) to obtain **(5) (4*S*, 5*S*) - 4, 5 - bis(azidomethyl) - 2, 2 – dimethyl - 1, 3 - dioxolane**. The same procedure was repeated four times with same amount of compound **3** to get compound **(5)** (5.1 g, 72 %). ¹H NMR (CDCl₃) δ: 1.47 (s, 6H), 3.29-3.60 (q, 2H; q, 2H), 4.04-4.07 (t, 2H). ¹³C

NMR (CDCl₃) δ : 26.77, 51.56, 76.87, 110.31. IR (ν_{\max} , cm⁻¹) (CHCl₃): 667, 756, 870, 1048, 1167, 1257, 1409, 1677, 1742, 2104, 2979, 3019. $[\alpha]_{\text{D}}^{27}$ - 6 (from L -isomer). + 8 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₇H₁₃N₆O₂: 213.2130, found: 213.2131

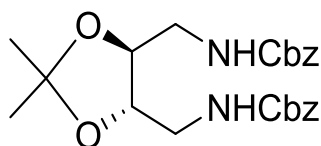
(6) ((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) dimethanamine



The azido compound **5** (5.1 g, 24.0 mmol) was dissolved in methanol (35 mL) and was hydrogenated over 10% Pd-C catalyst with 40 psi pressure for 2.5 h under hydrogen atmosphere. The content was then filtered through a bed of celite and the filtrate collected

was concentrated to obtain the desired compound **(6)** ((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) dimethanamine. (3.7 g, 96 %) and was used further without purification. ¹H NMR (CDCl₃) δ : 1.47 (s, 6H), 1.50 (s, 2H, D₂O exch.), 2.79-2.95 (m, 4H), 3.76-3.81 (t, 2H). ¹³C NMR (CDCl₃) δ : 27.09, 43.97, 80.02, 108.53. IR (ν_{\max} , cm⁻¹) (CHCl₃): 758, 1048, 1216, 1263, 1403, 1638, 1702, 2855, 2926, 3019, 3434. $[\alpha]_{\text{D}}^{27}$ - 12 (from L -isomer). + 12 (from D -isomer) (*c* 1.0, CH₃OH). HRMS (ESI) calcd for C₇H₁₇N₂O₂: 161.1285, found: 161.1285

(7) dibenzyl (((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) bis(methylene)) dicarbamate

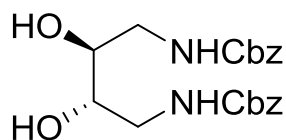


The diamine compound **6** formed in the previous step (3.7 g, 23.1 mmol) was dissolved in water (100 mL) and NaHCO₃ (4.9 g, 57.7 mmol) was added to it and stirred for 15 mins. To that added Cbz-Cl (7.8 mL, 46.2 mmol, 50 % solution in toluene)

dropwise and the reaction mixture was stirred for 8 h at room temperature. The solvent was removed under reduced pressure and the residue was diluted with 30 mL water. The content was the extracted with diethyl ether and the ether layer was kept separately. The aqueous layer was the neutralized using conc. HCl followed by washing with sat. NaHCO₃ (2 X 25 mL) followed by brine (2 X 20 mL). The ethyl acetate and the ether extract were kept over anhy. Na₂SO₄ and evaporated in vacuo. The crude product obtained was purified by column chromatography (30 % ethyl acetate-petroleum ether) to obtain the desired product **(7)** dibenzyl (((4*S*, 5*S*) - 2, 2 – dimethyl - 1, 3 – dioxolane - 4, 5 - diyl) bis(methylene)) dicarbamate (8.4 g, 85 %). ¹H NMR (CDCl₃) δ : 1.36 (s, 6H), 3.43 (br, m, 4H), 3.81 (br, m, 2H), 5.11 (s, 4H), 7.35 (s, 10H). ¹³C NMR

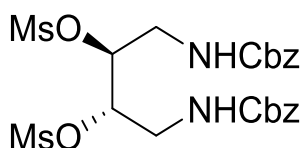
(CDCl₃) δ : 26.96, 42.05, 66.89, 77.20, 109.09, 128.08, 128.48, 136.34, 156.60. IR (ν_{\max} , cm⁻¹) (CHCl₃): 720, 790, 850, 1085, 1120, 1260, 1385, 1490, 1510, 1535, 1670, 1745, 2988, 3447. $[\alpha]_D^{27}$ - 20 (from L -isomer). + 20 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₂₃H₂₈N₂O₆Na: 451.1840, found: 451.1841

(8) dibenzyl ((2*S*, 3*S*) - 2, 3 – dihydroxybutane - 1, 4 - diyl) dicarbamate



The di-Cbz protected compound **7** (4.2 g, 9.8 mmol) was dissolved in methanol (80 mL) and to it added pTSA (0.8 g, 4.9 mmol). The content was refluxed at 60 °C for 24 h. After that methanol was removed under reduced pressure and the residue was directly purified by column chromatography (3 % MeOH-CH₂Cl₂) to give **(8) dibenzyl ((2*S*, 3*S*) - 2, 3 – dihydroxybutane - 1, 4 - diyl) dicarbamate**. The same procedure was repeated twice with same amount of compound **7** to get compound **(8)** (6.1 g, 80 %). ¹H NMR (MeOD) δ : 3.23-3.35 (m, 4H), 3.61 (t, 2H), 5.09 (s, 4H), 7.35 (s, 10H). ¹³C NMR (MeOD) δ : 43.42, 66.20, 70.19, 127.49, 128.13, 136.98, 157.87. IR (ν_{\max} , cm⁻¹) (CHCl₃): 756, 960, 1045, 1172, 1251, 1365, 1525, 1682, 1701, 2361, 2978, 3360. $[\alpha]_D^{27}$ - 14 (from L -isomer). + 16 (from D -isomer) (*c* 1.0, CH₃OH). HRMS (ESI) calcd for C₂₀H₂₄N₂O₆Na: 411.1527, found: 411.1526

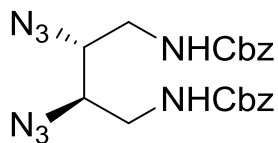
(9) (6*S*, 7*S*) - 3, 10 – dioxo - 1, 12 – diphenyl - 2, 11 – dioxa - 4, 9 – diazadodecane - 6, 7 - diyl dimethanesulfonate



The pure compound **8** (6.1 g, 15.7 mmol) obtained in the previous step was dissolved in dry pyridine (35 mL) and stirred for 15 min at 0 °C. Mesyl chloride (2.7 mL, 34.5 mmol) was added dropwise to this solution at 0 °C with stirring and the reaction mixture was left for 3 h. Pyridine was removed under reduced pressure and the residue was dissolved in ethyl acetate (80 mL). Organic layer was washed with water (2 X 20 mL), sat. NaHCO₃ (2 X 20 mL) followed by brine (2 X 10 mL) and dried over anhy. Na₂SO₄. The organic layer was concentrated in vacuo to give the crude product which was purified by column chromatography (30 % ethyl acetate-petroleum ether) to give **(9) (6*S*, 7*S*) - 3, 10 – dioxo - 1, 12 – diphenyl - 2, 11 – dioxa - 4, 9 – diazadodecane - 6, 7 - diyl dimethanesulfonate** (6.8 g, 88 %). ¹H NMR ((CD₃)₂CO) δ : 2.81 (s, 2H), 3.18 (s, 6H), 3.62-3.68 (m, 4H), 4.92-5.00 (qu, 2H), 5.11 (s, 4H), 7.33-7.39 (m, 10H). ¹³C NMR ((CD₃)₂CO) δ : 38.65, 42.30, 66.97, 79.00, 128.69, 129.20, 137.99, 157.47. IR (ν_{\max} , cm⁻¹) (CHCl₃): 731, 903, 1198, 1412,

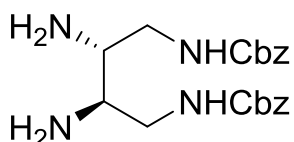
1647, 1685, 1744, 2361, 2978, 3198. $[\alpha]_D^{27}$ - 6 (from L -isomer). + 8 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₂₂H₂₈N₂O₁₀S₂Na: 567.1078, found: 567.1078

(10) dibenzyl ((2*R*, 3*R*) - 2, 3 – diazidobutane - 1, 4 - diyl) dicarbamate



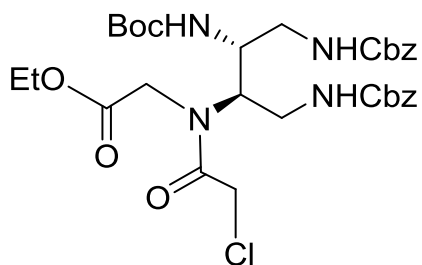
The compound **9** (2.3 g, 4.2 mmol) was dissolved in dry DMF (30 mL) and NaN₃ (2.7 g, 42.2 mmol) was added to it. The reaction mixture was stirred at 60 °C for 7 h. After completion of reaction DMF was removed under reduced pressure and the residue was dissolved in ethyl acetate (100 mL). Organic layer was washed with water (2 X 25 mL) followed by brine (2 X 20 mL). The organic layer was kept over anhy. Na₂SO₄ and then concentrated in vacuo. The crude compound obtained was purified by column chromatography (20 % ethyl acetate-petroleum ether) to obtain **(10) dibenzyl ((2*R*, 3*R*) - 2, 3 – diazidobutane - 1, 4 - diyl) dicarbamate**. The same procedure was repeated thrice with same amount of compound **9** to get compound **(10)** (4.1 g, 75 %). ¹H NMR (CDCl₃) δ: 3.32-3.48 (m, 4H), 3.60-3.62 (m, 2H), 5.11 (s, 4H), 5.19-5.25 (m, 2H), 7.34 (s, 10H). ¹³C NMR (CDCl₃) δ: 42.49, 62.13, 67.18, 128.08, 128.53, 136.06, 156.41. IR (ν_{\max} , cm⁻¹) (CHCl₃): 756, 837, 1119, 1167, 1395, 1470, 1676, 1748, 2102, 2930, 3175. $[\alpha]_D^{27}$ + 12 (from L -isomer). - 10 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₂₀H₂₂N₈O₄Na: 461.1656, found: 461.1656

(11) dibenzyl ((2*R*, 3*R*) - 2, 3 – diaminobutane - 1, 4 - diyl) dicarbamate



The azido compound **10** (4.1 g, 9.4 mmol) was dissolved in methanol (35 mL) and was hydrogenated over 10% Pd-C catalyst with 40 psi pressure for 2.5 h under hydrogen atmosphere. The content was then filtered through a bed of celite and the filtrate collected was concentrated to obtain the desired compound **(11) dibenzyl ((2*R*, 3*R*) - 2, 3 – diaminobutane - 1, 4 - diyl) dicarbamate**. (3.2 g, 91 %) and was used further without purification. ¹H NMR (CDCl₃) δ: 2.75-2.80 (m, 2H), 3.05-3.15 (m, 2H), 3.31-3.40 (m, 2H), 5.09 (s, 4H), 7.34 (s, 10H). ¹³C NMR (CDCl₃) δ: 42.48, 62.11, 67.18, 128.10, 128.54, 136.04, 156.41. IR (ν_{\max} , cm⁻¹) (CHCl₃): 775, 832, 1048, 1123, 1416, 1496, 1646, 1738, 1751, 3018, 3424. $[\alpha]_D^{27}$ + 4 (from L -isomer). - 2 (from D -isomer) (*c* 1.0, CHCl₃). HRMS (ESI) calcd for C₂₀H₂₇N₄O₄: 387.2027, found: 387.2027

(14) ethyl *N* - ((6*R*, 7*R*) - 7 - (((benzyloxy)carbonyl)amino)methyl) -11,11-dimethyl - 3, 9 - dioxo - 1 - phenyl - 2, 10 - dioxo - 4, 8 - diazadodecan - 6 - yl) - *N* - (2 - chloroacetyl) glycinate



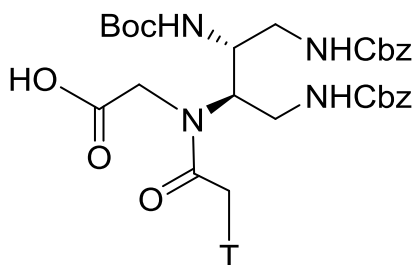
The diamino compound **11** (3.2 g, 8.3 mmol) was dissolved in 1:1 solution of 1,4-dioxane-water (100 mL) and triethylamine (2.3 mL, 16.6 mmol) was added with stirring. The reaction mixture was cooled to 0 °C and kept for 15 minutes. To that Boc-anhydride (2.2 mL, 9.9 mmol) was added dropwise over a period of 30 minutes at 0 °C. The reaction mixture was stirred for 10 h at room temperature. After the completion of the reaction, the content was concentrated to 30 mL and was extracted in ethyl acetate (70 mL). The organic layer was washed with water (2 X 25 mL) followed by brine (2 X 10 mL) and dried over anhy. Na₂SO₄. It was then evaporated in vacuo to give the crude product **12** (2.6 g, 66 %), which was used further alkylation. HRMS (ESI) calcd for C₂₅H₃₅N₄O₆: 487.2551, found: 487.2547

The monoamino compound **12** (2.6 g, 5.3 mmol) was dissolved in dry CH₃CN (60 mL) and triethylamine (1.5 mL, 10.7 mmol) was added with stirring. The reaction mixture was cooled to 0 °C and ethyl bromoacetate (0.7 mL, 6.4 mmol) diluted with CH₃CN (10 mL) was added dropwise. The reaction mixture was stirred for 2 h at room temperature. CH₃CN was removed under reduced pressure and the residue was dissolved in ethyl acetate (50 mL). The organic layer was washed with water (2 X 10 mL), sat. NaHCO₃ (2 X 10 mL) followed by brine (2 X 10 mL) and was dried over anhy. Na₂SO₄. It was then evaporated in vacuo to give crude product **13** (2.2 g, 72 %) for acylation. HRMS (ESI) calcd for C₂₉H₄₀N₄O₈Na: 595.2738, found: 595.2708

Compound **13** (2.2 g, 3.8 mmol) was dissolved in 1:1 solution of 1,4-dioxane-water (30 mL) and NaHCO₃ (3.2 g, 38.4 mmol) dissolved in water (20 mL) was added. The reaction mixture was stirred for 15 min at 0 °C and chloroacetyl chloride (1.5 mL, 19.2 mmol) was added in two portions maintaining a pH ~8.0. The reaction mixture was kept for 1 h with stirring and then concentrated to 20 mL. It was extracted in CH₂Cl₂ (50 mL) and was washed with brine (2 X 10 mL). The organic layer was dried over anhy. Na₂SO₄ followed by evaporation in vacuo to give the crude product. The product was purified by column chromatography (3 % MeOH-CH₂Cl₂) to give the desired product **(14) ethyl *N* - ((6*R*, 7*R*) - 7 - (((benzyloxy)carbonyl)amino)methyl) -**

11,11-dimethyl - 3, 9 – dioxo – 1 – phenyl - 2, 10 – dioxa - 4, 8 – diazadodecan – 6 - yl) – N - (2 - chloroacetyl) glycinate. (2.1 g, 86 %). ^1H NMR (CDCl_3) δ : 1.26-1.30 (t, 3H), 1.45 (s, 9H), 3.32 (m, 5H), 3.53 (t, 2H), 4.02 (s, 2H), 4.16 (m, 4H), 5.09 (s, 4H), 7.34 (s, 10H). ^{13}C NMR (CDCl_3) δ : 14.01, 28.27, 38.56, 40.59, 41.12, 49.07, 49.67, 61.64, 66.72, 79.83, 128.08, 128.45, 136.31, 155.97, 156.77, 167.55, 169.66 IR (ν_{max} , cm^{-1}) (CHCl_3): 777, 837, 1113, 1167, 1408, 1667, 1694, 1753, 2361, 2930, 3017 [α] $_{\text{D}}^{27}$ + 22 (from L -isomer). - 20 (from D -isomer) (c 1.0, CHCl_3). HRMS (ESI) calcd for $\text{C}_{31}\text{H}_{42}\text{N}_4\text{O}_9\text{Cl}$: 649.2640, found: 649.3817

(16a) N - (acetyl – 2 - thyminy) – N - ((6R, 7R) – 7- (((benzyloxy)carbonyl) amino) methyl) 11, 11 – dimethyl - 3, 9–dioxo –1–phenyl -2,10– dioxa - 4, 8 – diazadodecan – 6 - yl) glycine

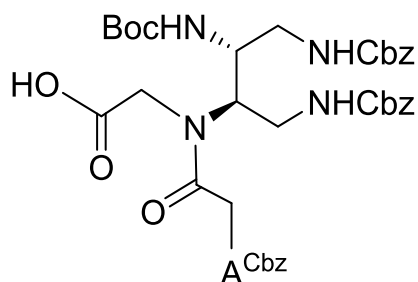


The chloroacetyl derivative **14** (1.0 g, 1.5 mmol), K_2CO_3 (0.21 g, 1.5 mmol) and thymine (0.2 g, 1.5 mmol) were suspended in dry DMF (20 mL) and stirred at room temperature for 6 h. After that DMF was removed under reduced pressure and the residue was redissolved in ethyl acetate (30 mL). The organic layer was washed with water (2 X 10 mL) followed by brine (2 X 10 mL) and dried over anhy. Na_2SO_4 . It was then evaporated in vacuo to give crude product. The crude product **15a** (0.87 g, 76 %) was used further for desired monomer.

The compound **15a** (0.87 g, 1.2 mmol) formed in the above step was dissolved in methanol (10 mL) and to that added 1N LiOH (10 mL). The content was stirred for 30 min and after the completion of reaction; methanol was evaporated under reduced pressure. The resulting aqueous solution was neutralized by Dowex H^+ ion exchange resin. The aqueous layer was washed with ethyl acetate (8 mL) followed by evaporation in vacuo and desiccation to obtain **(16a) N - (acetyl – 2 - thyminy) – N - ((6R, 7R) – 7- (((benzyloxy)carbonyl) amino) methyl) 11, 11 – dimethyl - 3, 9–dioxo –1–phenyl -2,10– dioxa - 4, 8 – diazadodecan – 6 - yl) glycine** (0.71 g, 85 %). ^1H NMR (DMSO-d_6) δ : 1.38 (s, 9H), 1.75 (s, 3H), 3.05-3.07 (dd, 4H), 3.40 (m, 2H), 3.97-4.19 (d, 2H), 4.47-4.64 (d, 2H), 5.01 (s, 4H), 6.94 (m, 1H), 7.34 (s, 10H), 11.30 (s, 1H). ^{13}C NMR (DMSO-d_6) δ : 12.14, 28.38, 38.28, 43.69, 47.02, 47.78, 49.28, 65.46, 69.96, 78.30, 108.37, 127.93, 128.57, 137.44, 142.23, 151.23, 155.99, 156.58, 164.63, 167.41, 170.71. IR (ν_{max} , cm^{-1}) (CHCl_3): 729, 906, 1067, 1262, 1409, 1667, 1685, 1692, 1711,

1728, 2400, 3019, 3374, 3451. $[\alpha]_D^{27} + 2$ (from L -isomer). - 2 (from D -isomer) (*c* 1.0, CH₃OH). HRMS (ESI) calcd for C₃₄H₄₂N₆O₁₁Na: 734.7308, found: 734.8547

(16b) *N* - (acetyl - 2 - ((benzyloxy)carbonyl) - adeninyl) - *N* - ((6*R*, 7*R*) - 7- (((benzyloxy)carbonyl) amino) methyl) 11, 11 -dimethyl - 3, 9-dioxo -1-phenyl - 2,10- dioxo - 4, 8 - diazadodecan - 6 - yl) glycine



The chloroacetyl derivative **14** (1.0 g, 1.5 mmol), K₂CO₃ (0.21 g, 1.5 mmol) and Cbz protected adenine (0.42 g, 1.5 mmol) were suspended in dry DMF (20 mL) and stirred at room temperature for 6 h. After that DMF was removed under reduced pressure and the residue was redissolved in ethyl acetate (30 mL). The

organic layer was washed with water (2 X 10 mL) followed by brine (2 X 10 mL) and dried over anhy. Na₂SO₄. It was then evaporated in vacuo to give crude product. The crude product **15a** (1.06 g, 78 %) was used further for desired monomer.

The compound **15b** (1.0 g, 1.1 mmol) formed in the above step was dissolved in methanol (10 mL) and to that added 1N LiOH (10 mL). The content was stirred for 30 min and after the completion of reaction; methanol was evaporated under reduced pressure. The resulting aqueous solution was neutralized by Dowex H⁺ ion exchange resin. The aqueous layer was washed with ethyl acetate (8 mL) followed by evaporation in vacuo and desiccation to obtain **(16b)** *N* - (acetyl - 2 - ((benzyloxy)carbonyl) - adeninyl) - *N* - ((6*R*, 7*R*) - 7- (((benzyloxy)carbonyl) amino) methyl) 11, 11 -dimethyl - 3, 9-dioxo -1-phenyl -2,10- dioxo - 4, 8 - diazadodecan - 6 - yl) glycine (0.77 g, 75 %). ¹H NMR (DMSO-d₆) δ: 1.36 (s, 9H), 3.05-3.16 (m, 4H), 3.25-3.36 (q, t 2H), 3.47-3.55 (dt, 3H), 3.99-4.22 (m, 3H), 5.02-5.34 (m, 8H), 7.13-7.47 (m, 15H), 8.32 (d, 1H), 9.61 (d, 1H), 10.68 (s, 1H). ¹³C NMR (DMSO-d₆) δ: 28.38, 37.81, 43.69, 47.21, 47.96, 50.34, 65.45, 66.47, 69.95, 78.34, 123.13, 128.05, 128.56, 136.61, 137.44, 145.37, 149.60, 151.66, 152.63, 156.53, 166.68, 167.21, 170.73, 171.24. IR (ν_{max}, cm⁻¹) (CHCl₃): 763, 878, 1044, 1218, 1374, 1519, 1664, 1697, 1735, 1752, 2367, 3016, 3245, 3453. $[\alpha]_D^{27} + 6$ (from L -isomer). - 8 (from D -isomer) (*c* 1.0, CH₃OH). HRMS (ESI) calcd for C₄₂H₄₈N₉O₁₁: 854.8900, found: 854.8900

Solid phase oligomer Synthesis

The PNA oligomers were synthesized by standard Boc-solid phase peptide strategy on MBHA resin having initial loading value 1.75 mmol/g. The loading value was lowered

to 0.3 mmol/g by capping with acetic anhydride in dry CH_2Cl_2 and pyridine as base. The free amines of the resin was then functionalized by coupling with *N* (α)-Boc-*N* (ϵ)-2-Cl-Cbz-*L*-Lysine using TBTU/HOBt as coupling reagent and DIPEA as base in DMF.

Picric acid estimation of resin functionalization: The typical procedure for estimation of the loading value of the resin was carried out with 5 mg of the resin which comprises the following steps:

The resin was swollen in dry CH_2Cl_2 for 2 hrs. After soaking with CH_2Cl_2 , it was drained off and a 50% solution of TFA in CH_2Cl_2 was added (1 mL x 3), 15 min each. After washing thoroughly with CH_2Cl_2 , the TFA salt was neutralized with 5% solution of DIPEA in CH_2Cl_2 (1 mL x 5min each). The free amine was treated with a 0.1M picric acid solution in CH_2Cl_2 (1 mL x 2 min each). The excess picric acid was removed by extensively washing the resin with CH_2Cl_2 . The adsorbed picric acid was displaced from the resin by adding a solution of 5% DIPEA in CH_2Cl_2 . The eluant was collected and the volume was made up to 10 mL with CH_2Cl_2 in a volumetric flask. The absorbance was recorded at 358 nm in ethanol and the concentration of the amine groups on the resin was calculated using the molar extinction coefficient of picric acid as $14,500 \text{ cm}^{-1}\text{M}^{-1}$ at 358 nm.

The lysine functionalized resin was swollen in dry CH_2Cl_2 for 30 min and then treated with 50% TFA in CH_2Cl_2 (3 X 2 mL, 15 min each). After that the resin was washed with CH_2Cl_2 (3 X 2 mL) followed by DMF (3 X 2 mL) and again by CH_2Cl_2 (3 X 2 mL). The neutralization of the TFA salt formed was done by treating the resin with 5% DIPEA in CH_2Cl_2 (3 X 2 mL, 5 min each) and washed further with CH_2Cl_2 (3 X 2 mL). The unmodified and modified PNA monomers were coupled to the resin using TBTU/HOBt as coupling reagent and DIPEA as base in DMF. The efficiency of the Boc deprotection for free amine and the coupling of monomers were monitored by Kaiser's test. At the end of the oligomers synthesis, the oligomers were cleaved from the resin with standard TFA-TFMSA protocol.

Kaiser's Test: Kaiser's test was used to monitor the Boc-deprotection and amide coupling steps in the solid phase peptide synthesis. Three solutions were used, viz. (1) Ninhydrin (5.0 g) dissolved in ethanol (100 mL) (2) Phenol (80 g dissolved in ethanol (20 mL) and (3) KCN (0.001M aqueous solution of KCN in 98 mL pyridine). To a few beads of the resin taken in a test tube, was added 3-4 drops of each of the three solutions described above. The tube was heated for 5 min, and the colour of the beads was noted. A blue colour on the beads and in the solution indicated successful

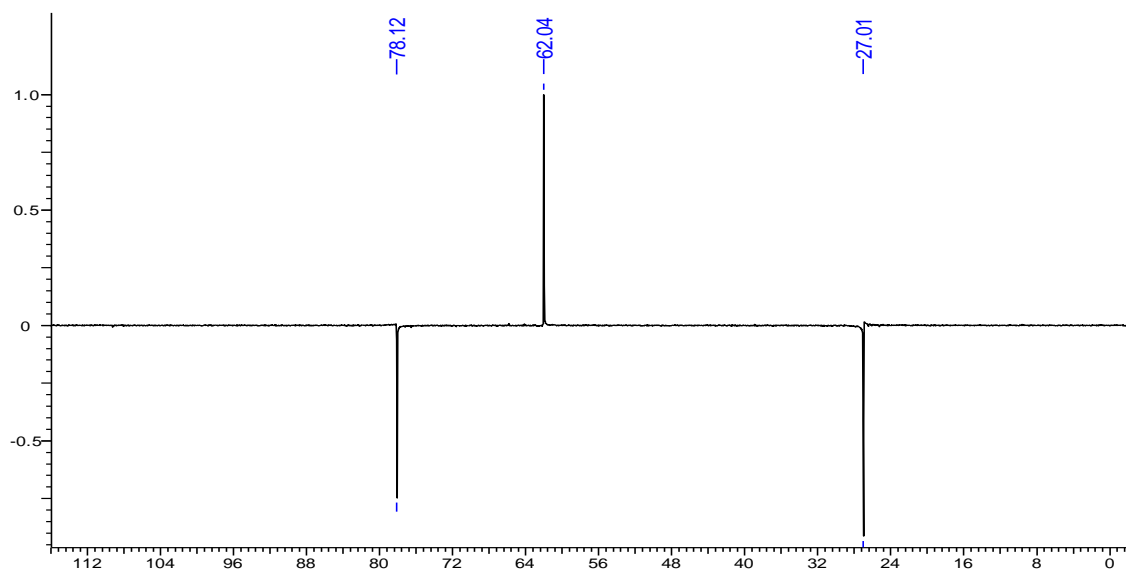
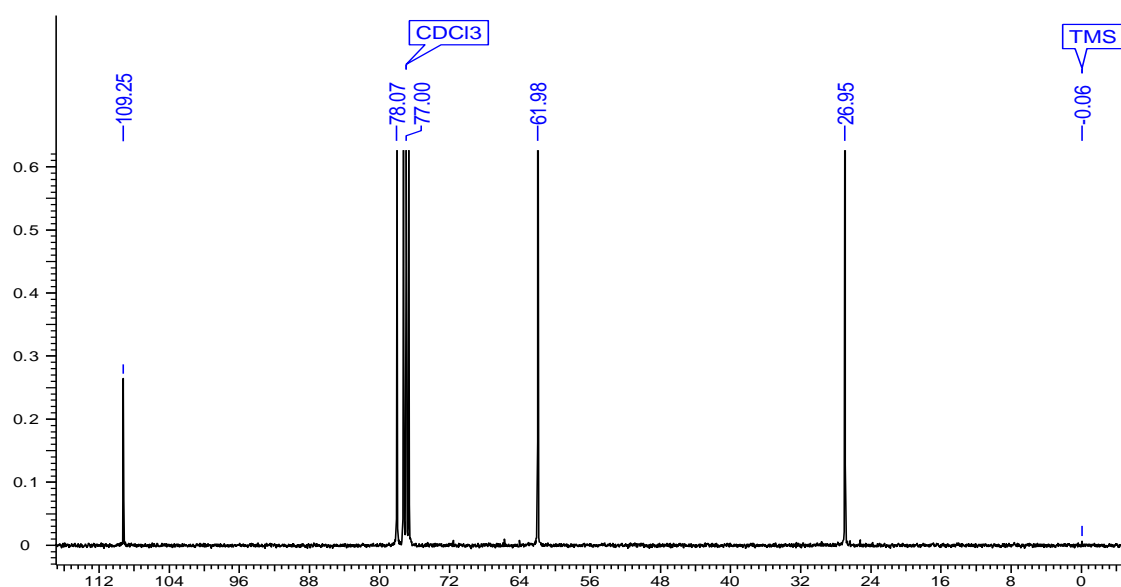
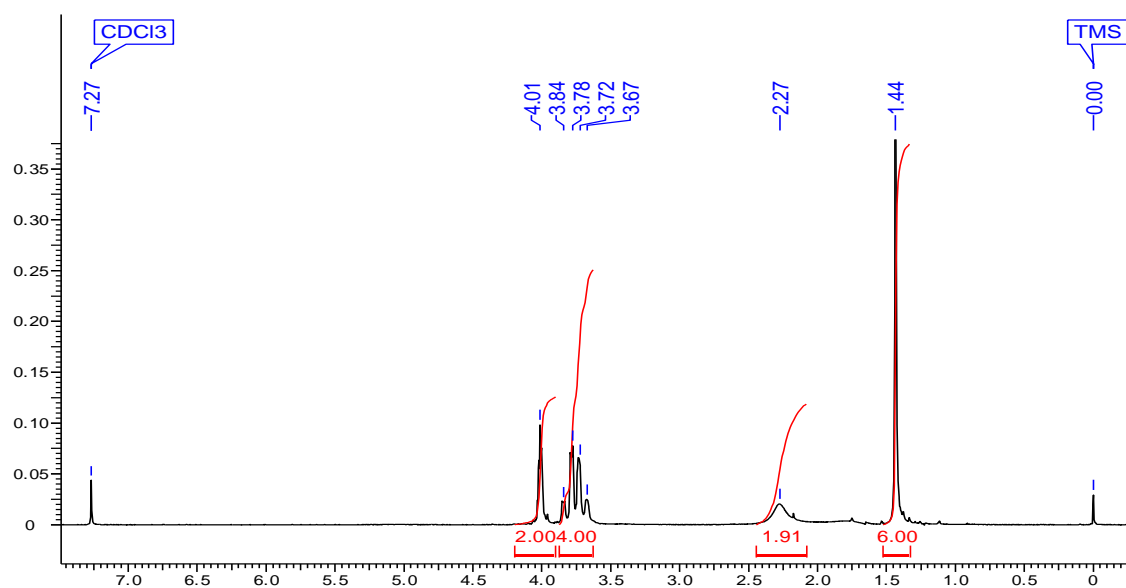
deprotection, while colourless beads were observed upon completion of the amide coupling reaction.

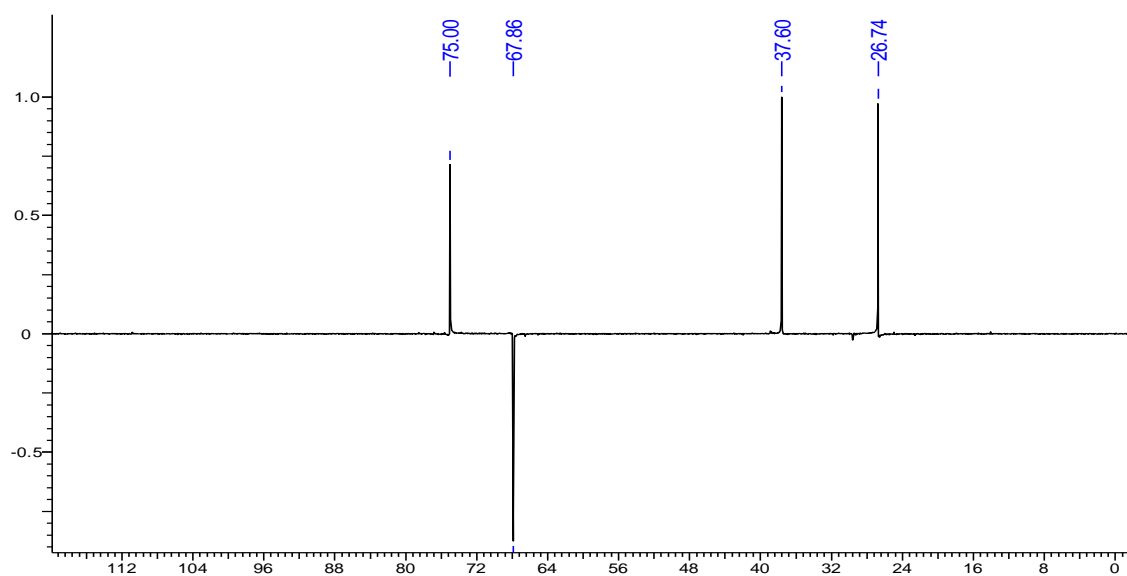
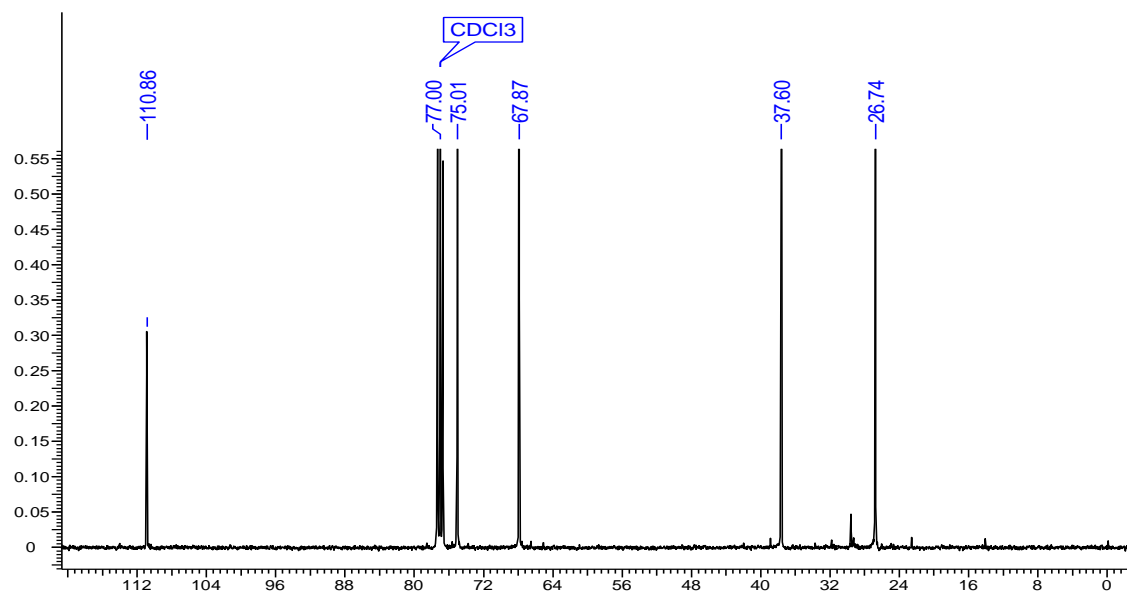
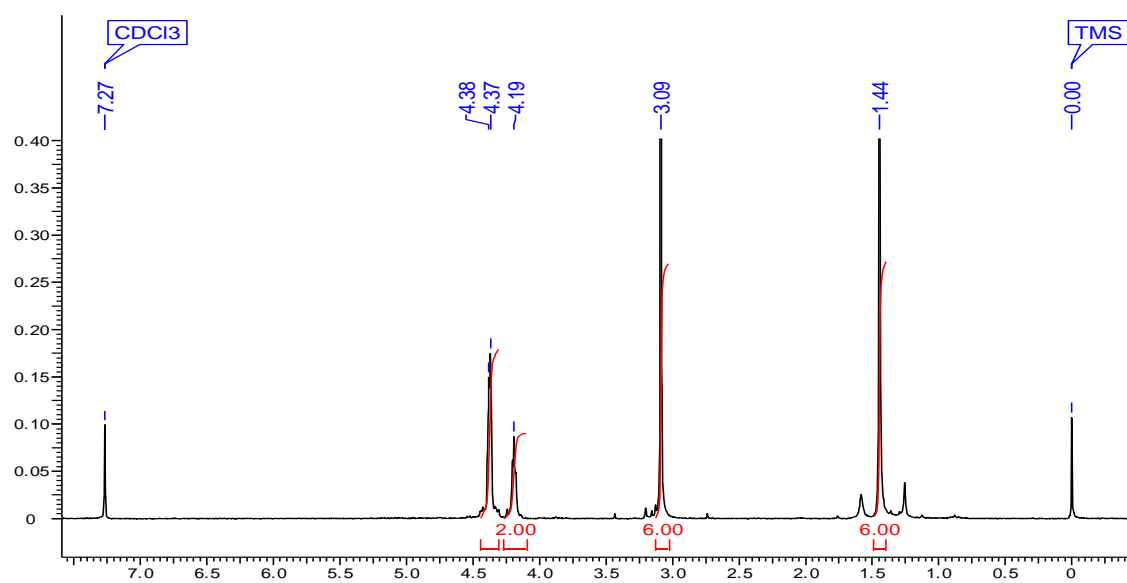
Cleavage of the PNA oligomers from the solid support

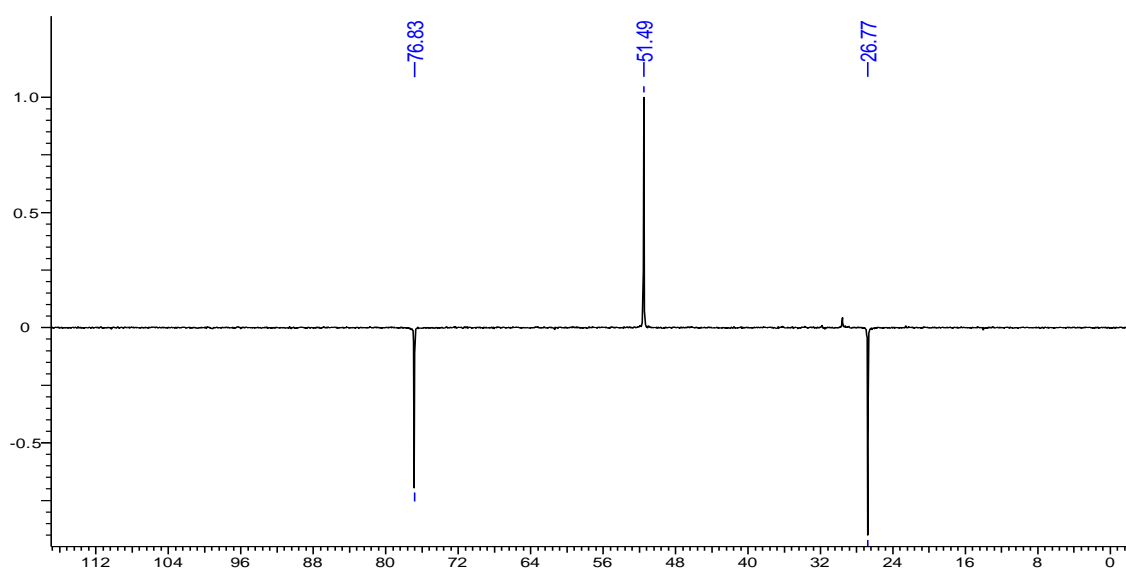
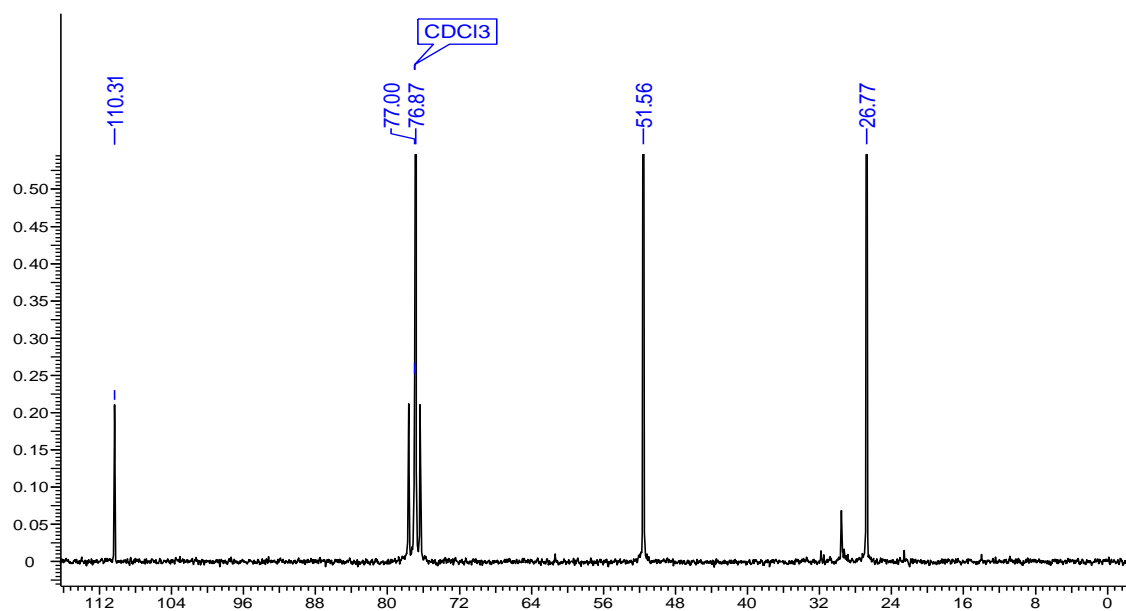
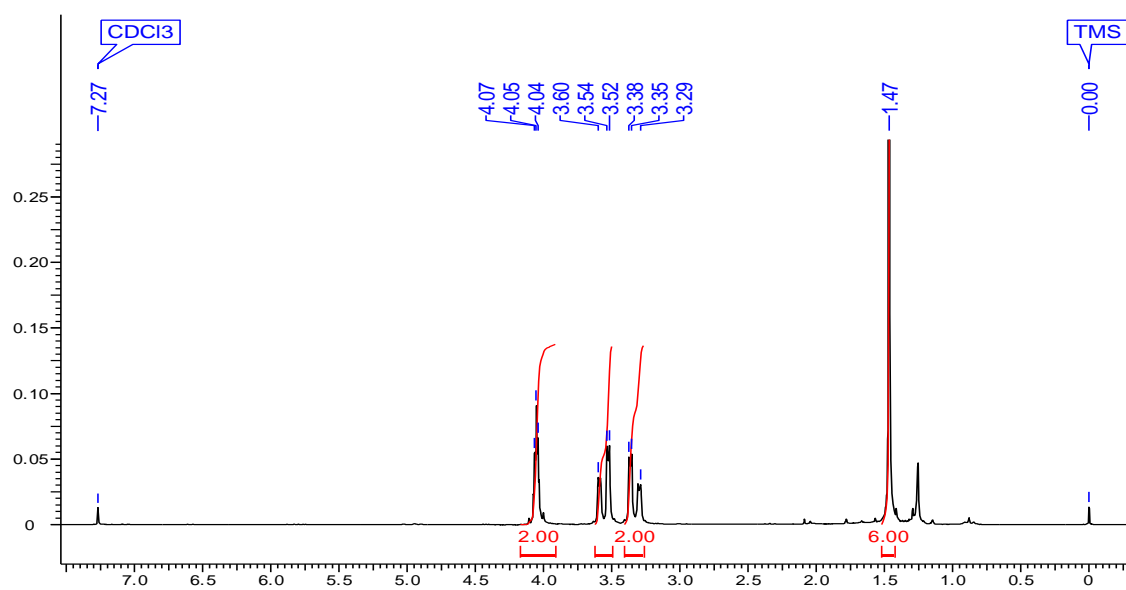
In a typical cleavage reaction, the resin bound oligomers (5 mg) were treated with thioanisole (10 μ L) and 1,2-ethanedithiol (4 μ L) for 10 min in an ice bath followed by TFA (120 μ L) and TFMSA (8 μ L) at 0 °C for 2 h. After that the resin was filtered using TFA (400 μ L) followed by evaporation of the filtrate under vacuum. The oligomers were precipitated by adding cold ether. The precipitated oligomers were washed with cold ether (2 X 300 μ L) and were dissolved in deionised water. Purification of all the PNA oligomers was carried out by HPLC on RP-C18 column with water: CH₃CN-0.1% TFA system and were characterized by MALDI-TOF mass spectrometry. The spectra were acquired in linear mode and the matrix used for analysis was CHCA (α -cyano-hydroxyl-cinnamic acid).

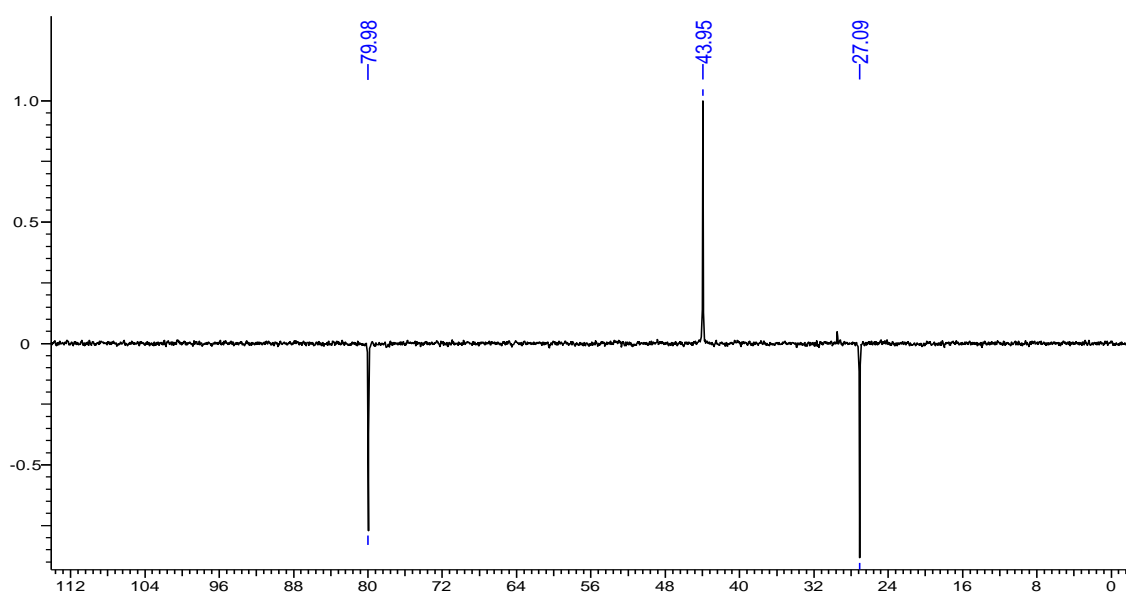
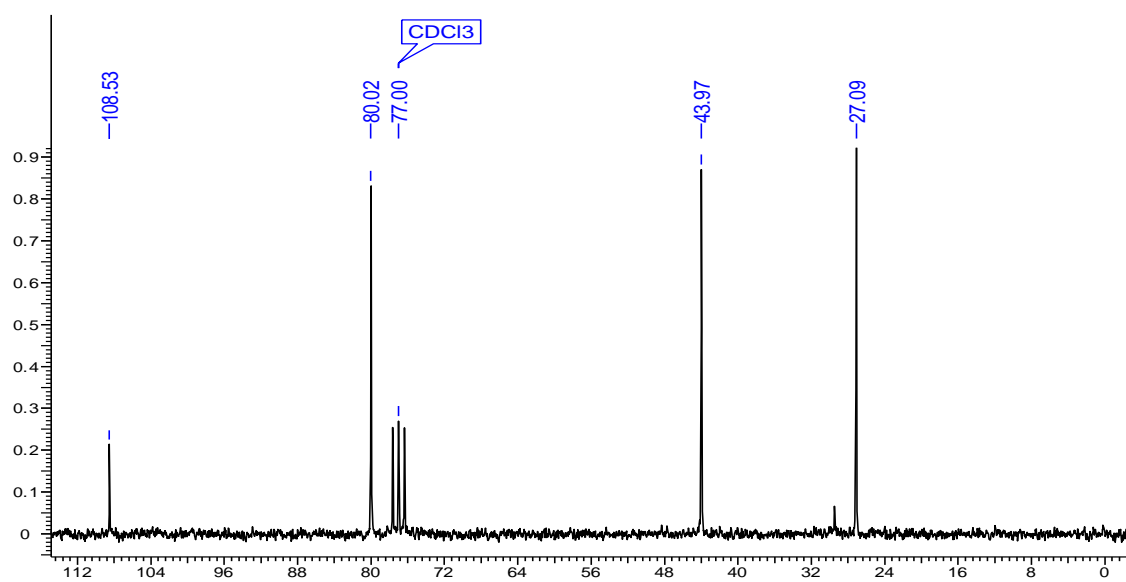
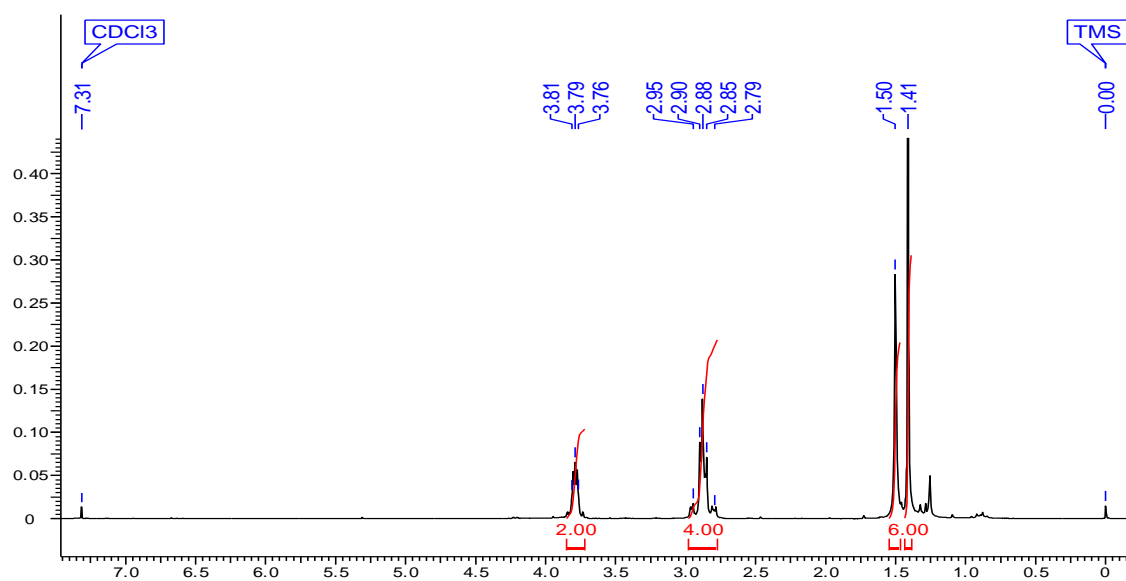
3A.6 Appendix

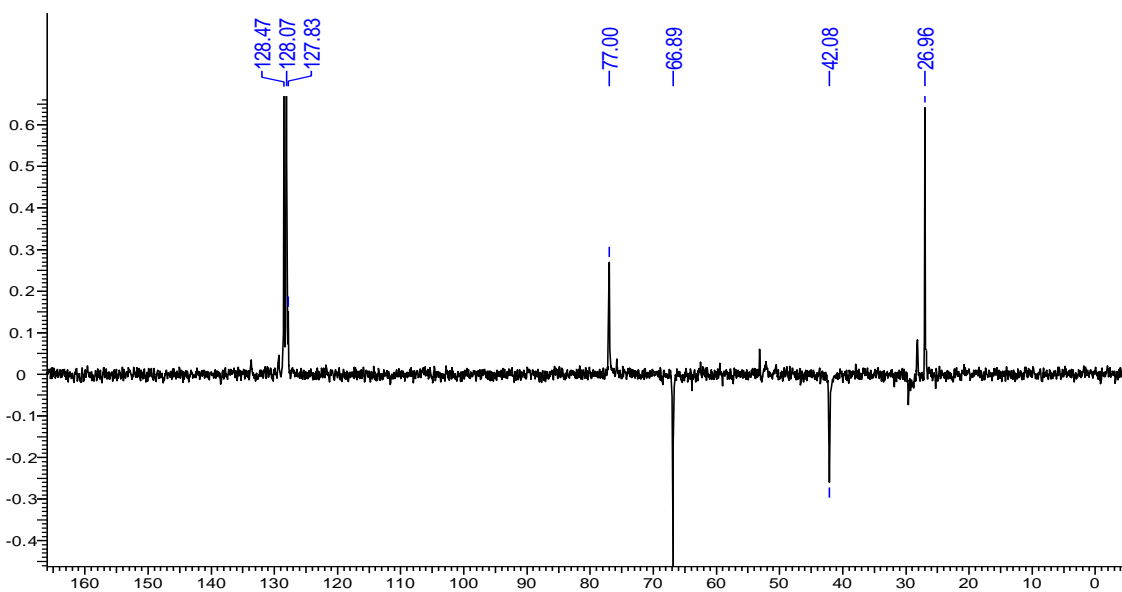
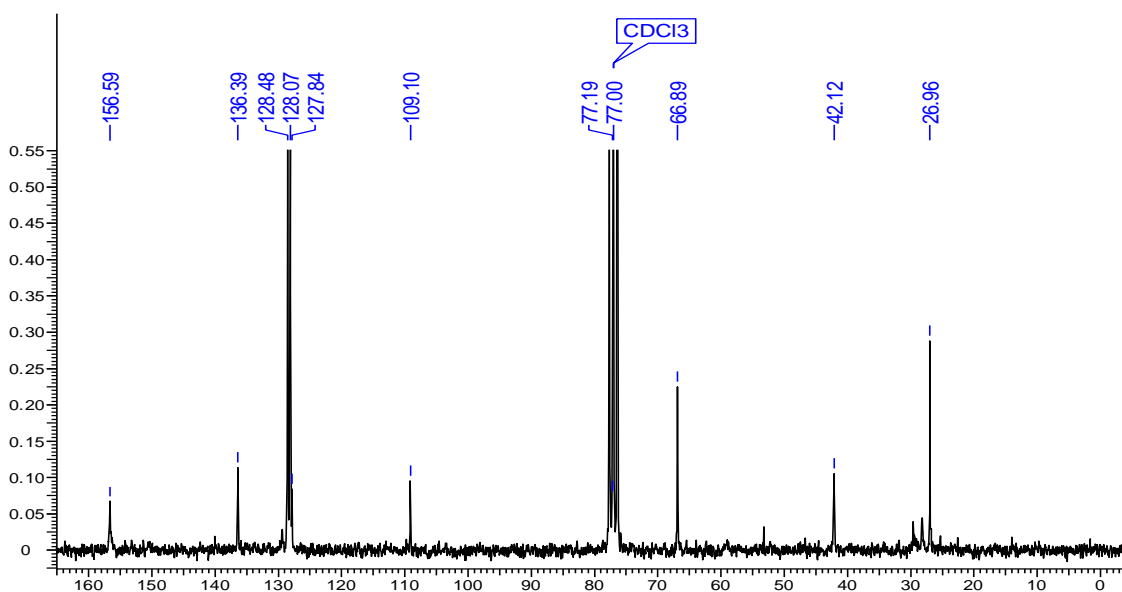
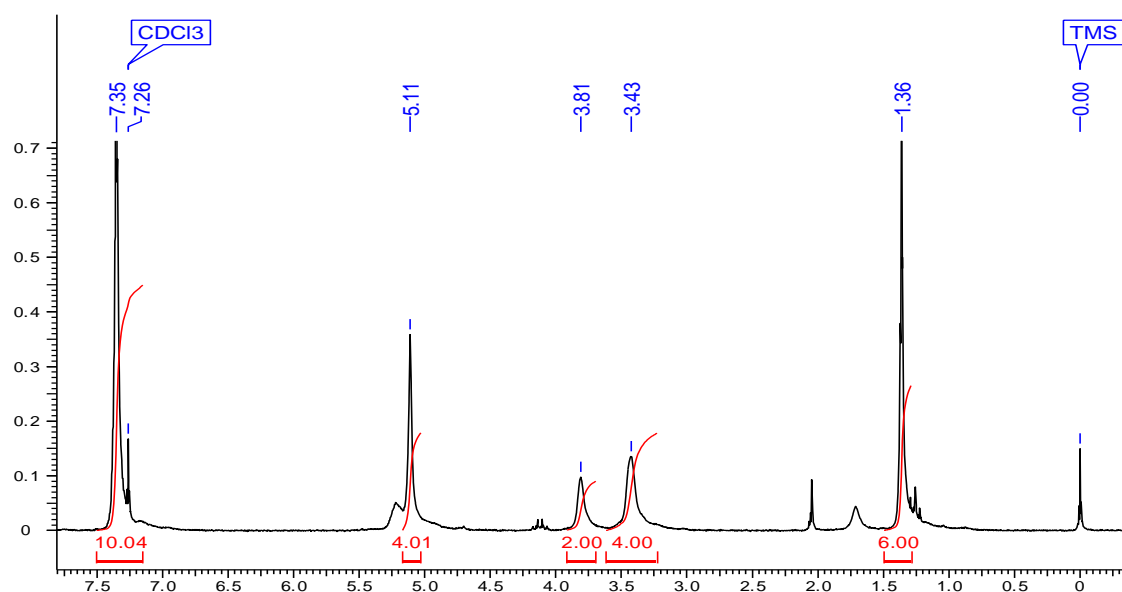
Compounds - Spectral data	Page Nos.
¹ H, ¹³ C and DEPT of compound 3	148
¹ H, ¹³ C and DEPT of compound 4	149
¹ H, ¹³ C and DEPT of compound 5	150
¹ H, ¹³ C and DEPT of compound 6	151
¹ H, ¹³ C and DEPT of compound 7	152
¹ H, ¹³ C and DEPT of compound 8	153
¹ H, ¹³ C and DEPT of compound 9	154
¹ H, ¹³ C and DEPT of compound 10	155
¹ H, ¹³ C and DEPT of compound 11	156
¹ H, ¹³ C and DEPT of compound 14	157
¹ H, ¹³ C and DEPT of compound 16a	158
¹ H, ¹³ C and DEPT of compound 16b	159
HRMS of compound 2 and 6	160
HRMS of compound 7 and 8	161
HRMS of compound 9 and 10	162
HRMS of compound 11 and 12	163
HRMS of compound 13 and 14	164
HRMS of compound 16a	165
HPLC and MALDI-TOF of PNA 1	166
HPLC and MALDI-TOF of PNA 2	167
HPLC and MALDI-TOF of PNA 3	168
HPLC and MALDI-TOF of PNA 4	169
HPLC and MALDI-TOF of PNA 5	170
HPLC and MALDI-TOF of PNA 6	171
HPLC and MALDI-TOF of PNA 7	172
HPLC and MALDI-TOF of PNA 8	173
HPLC and MALDI-TOF of PNA 9	174

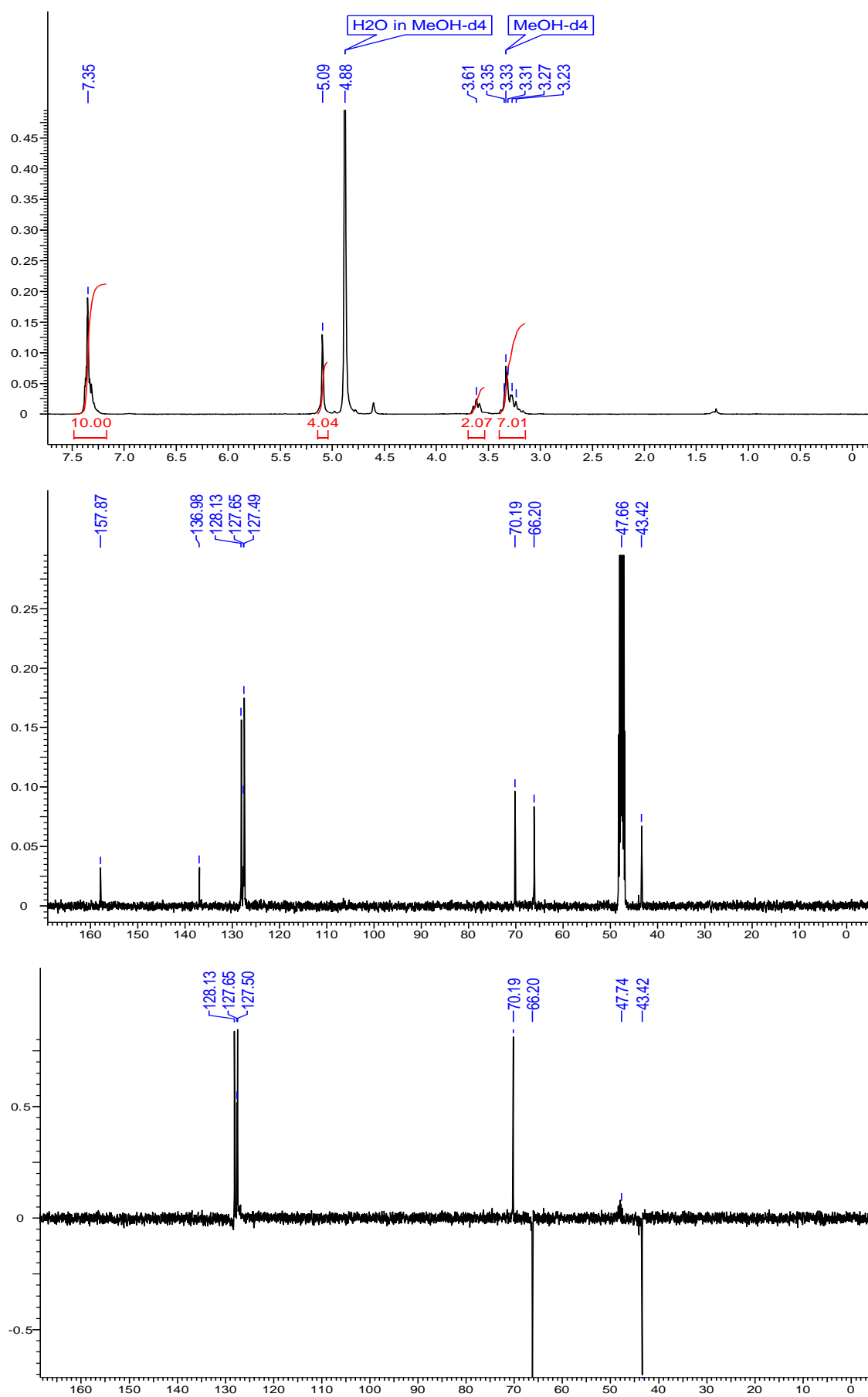
^1H , ^{13}C and DEPT of compound 3

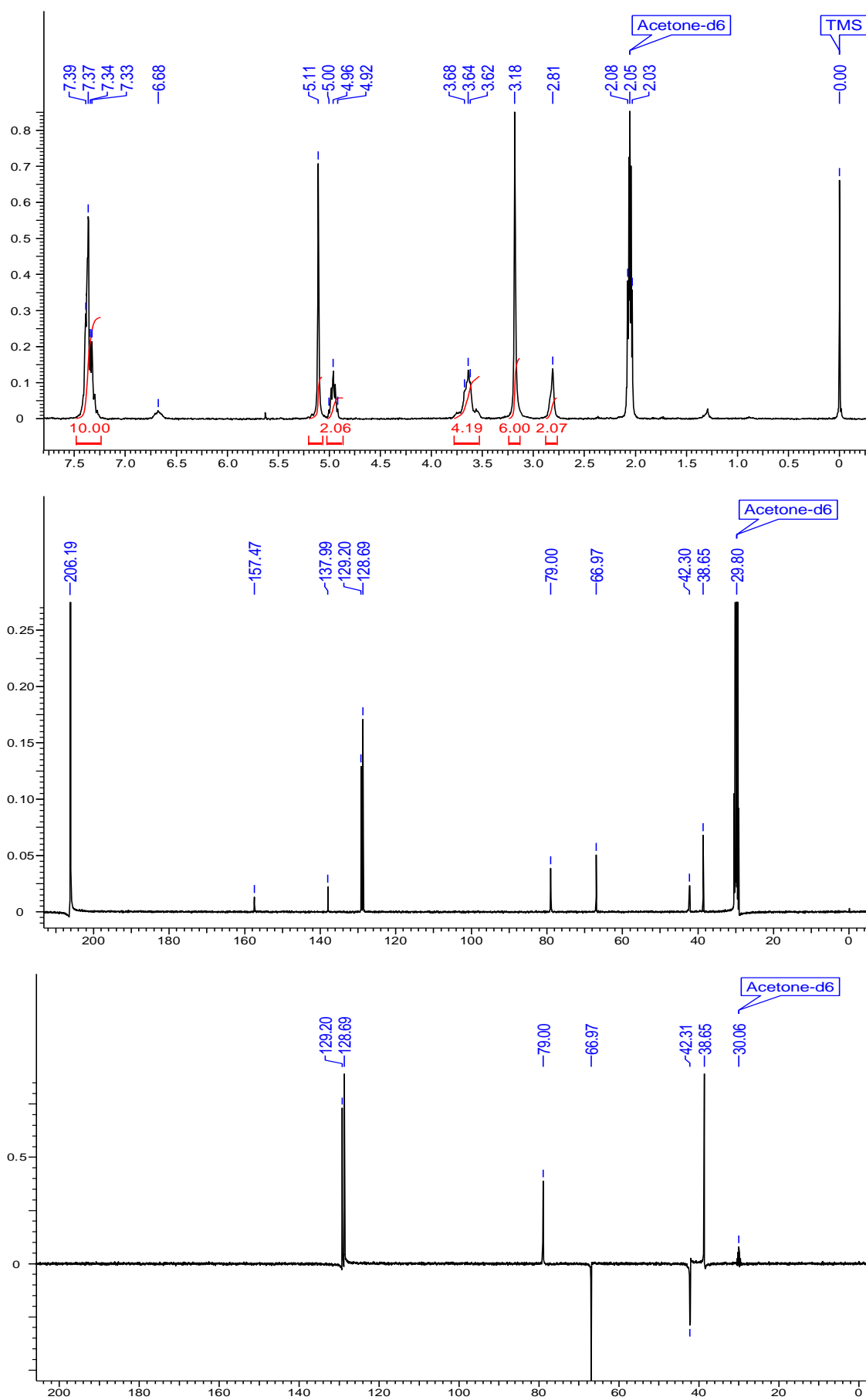
^1H , ^{13}C and DEPT of compound 4

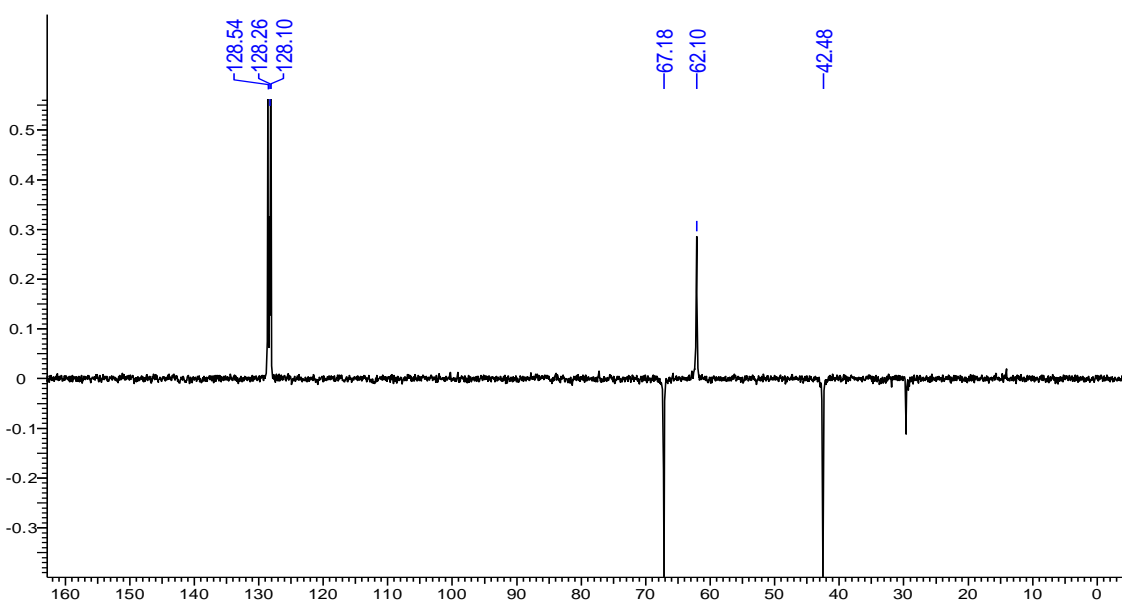
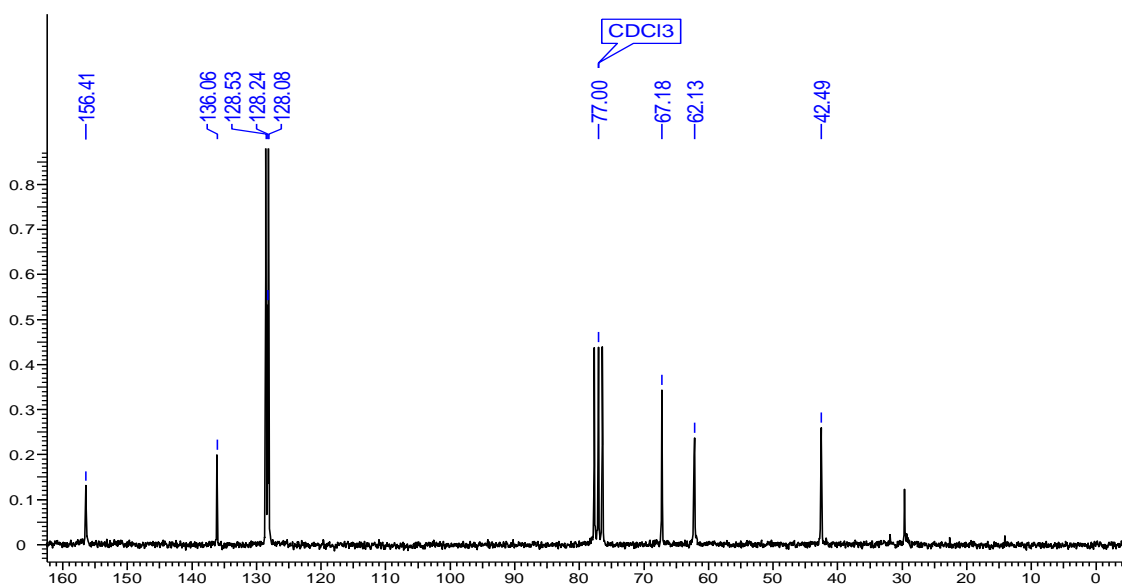
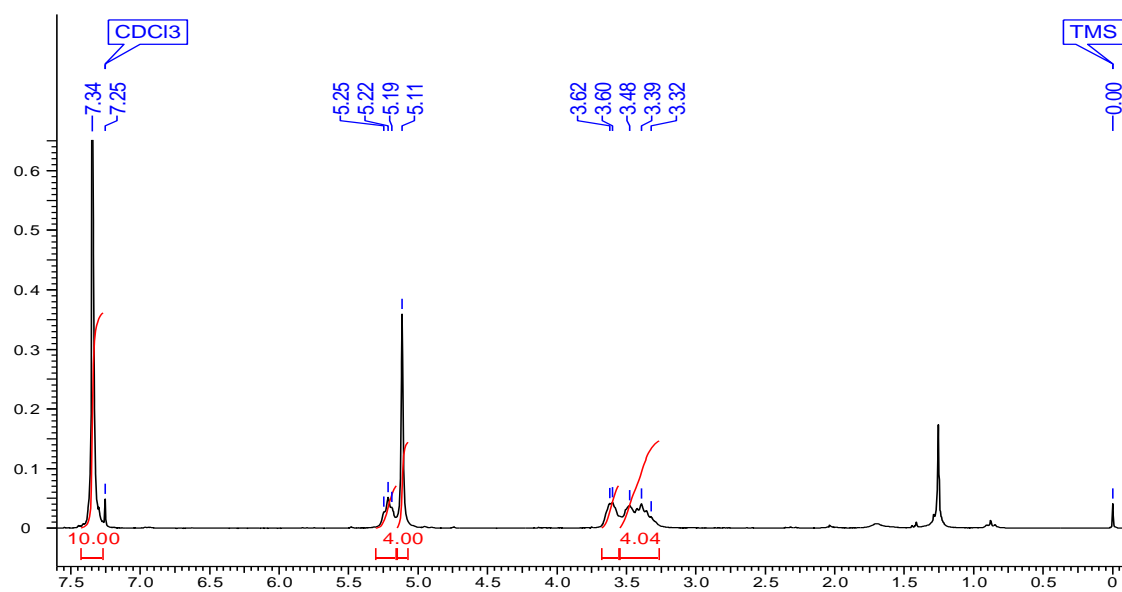
^1H , ^{13}C and DEPT of compound 5

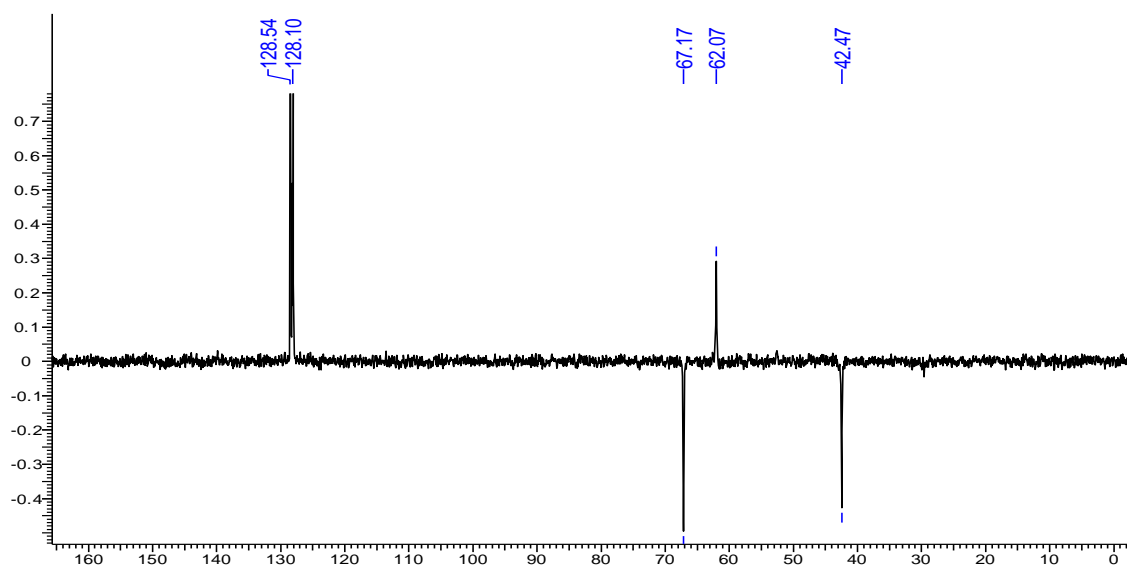
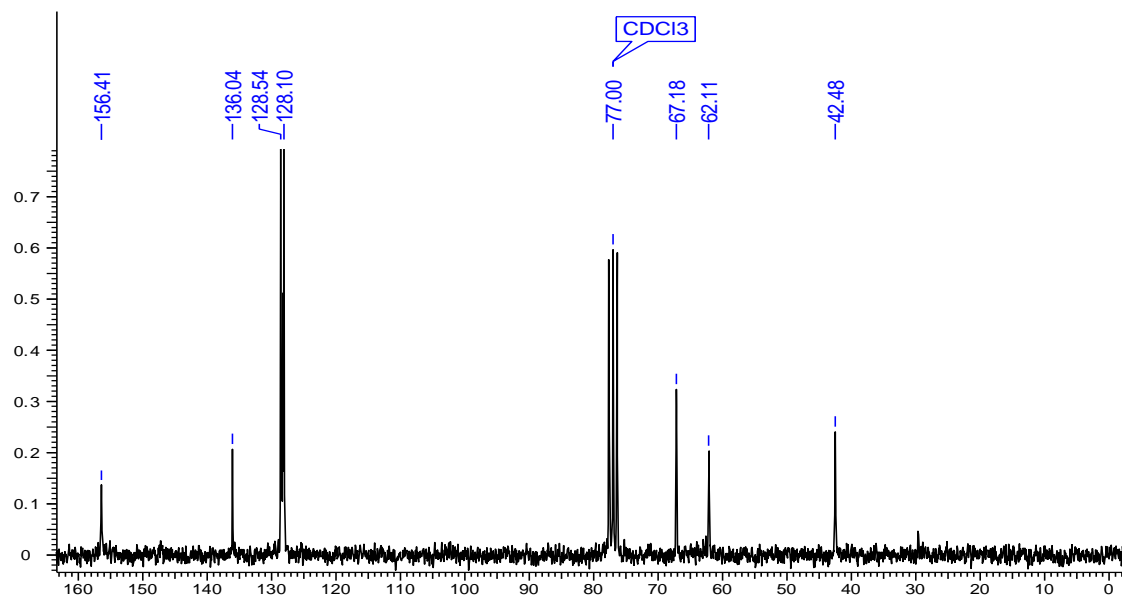
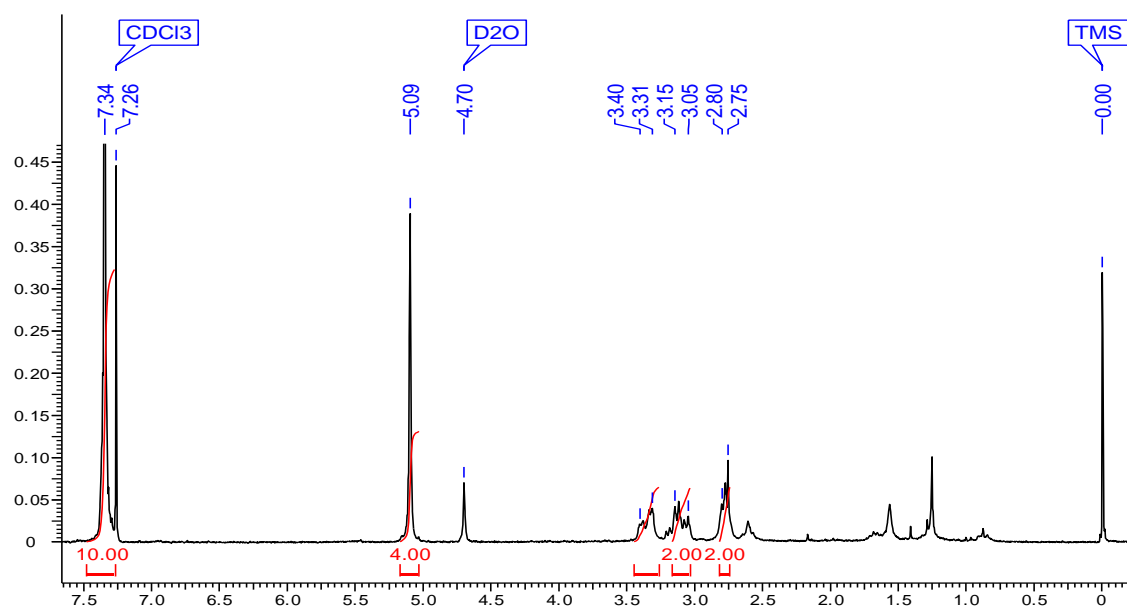
^1H , ^{13}C and DEPT of compound 6

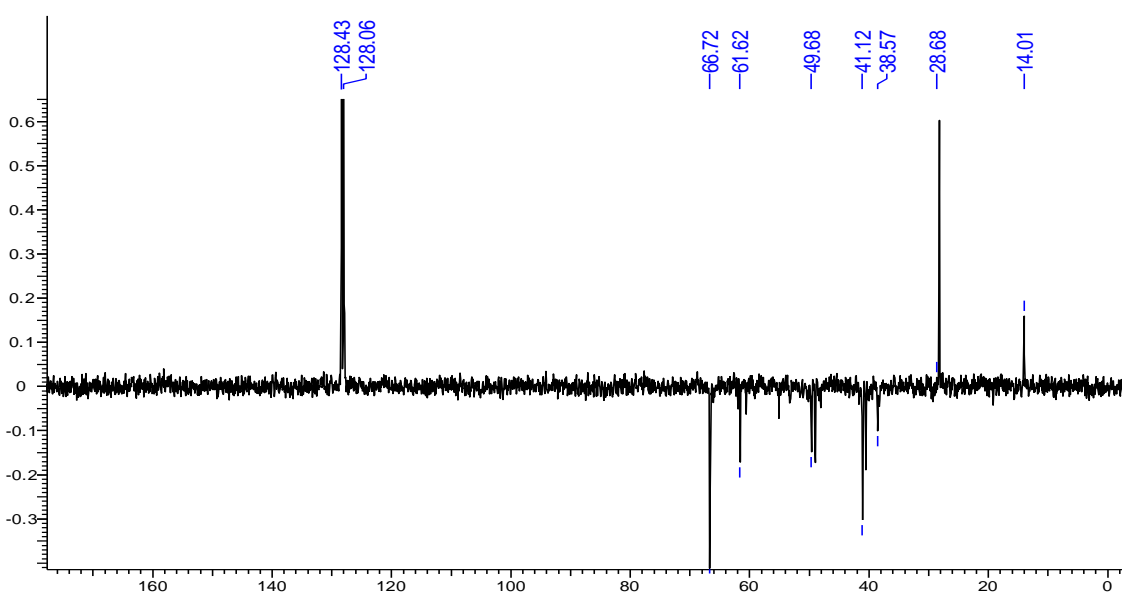
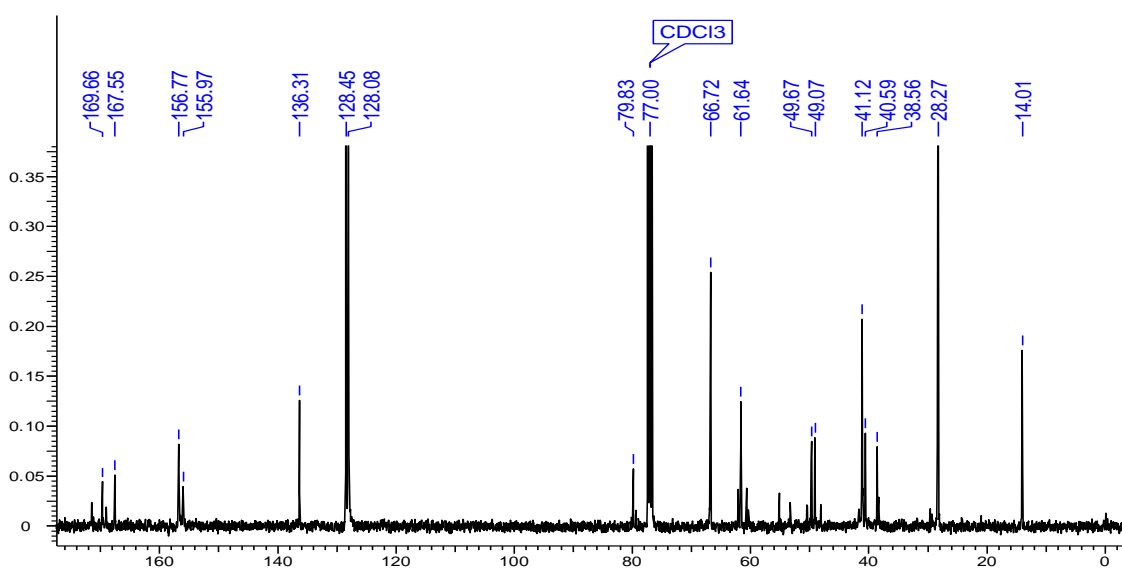
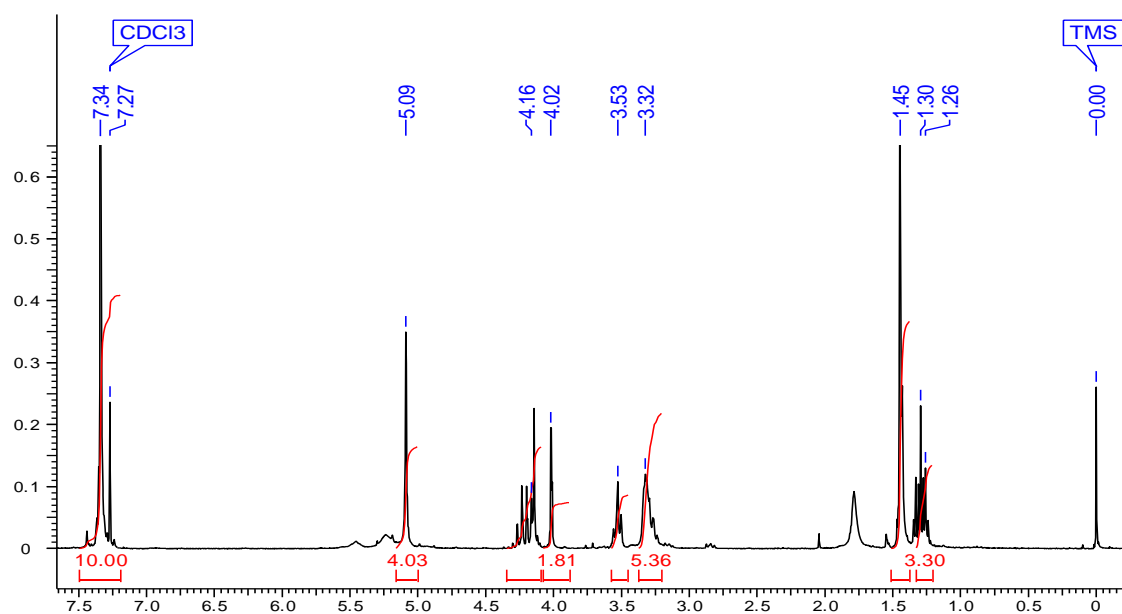
^1H , ^{13}C and DEPT of compound 7

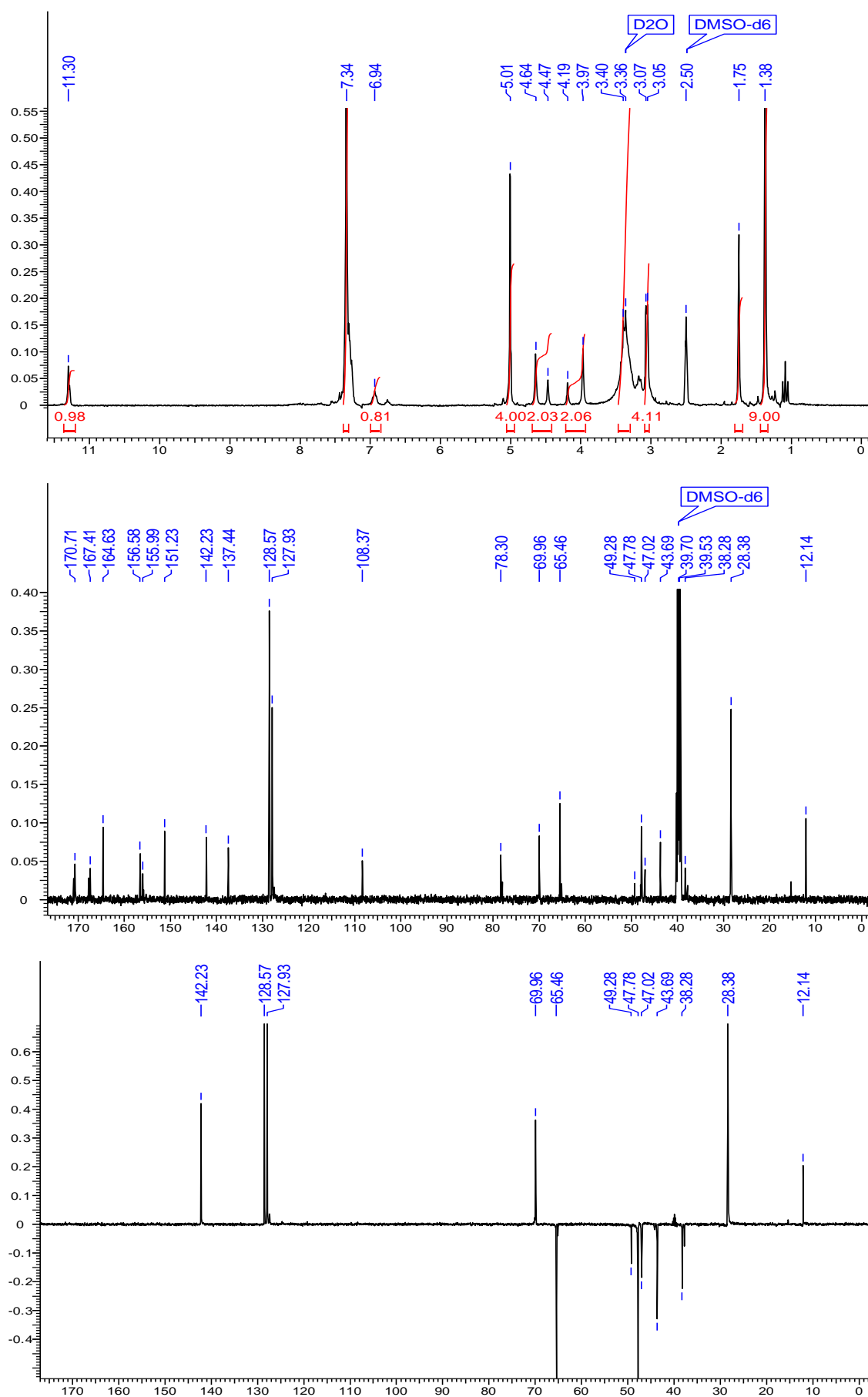
^1H , ^{13}C and DEPT of compound **8**

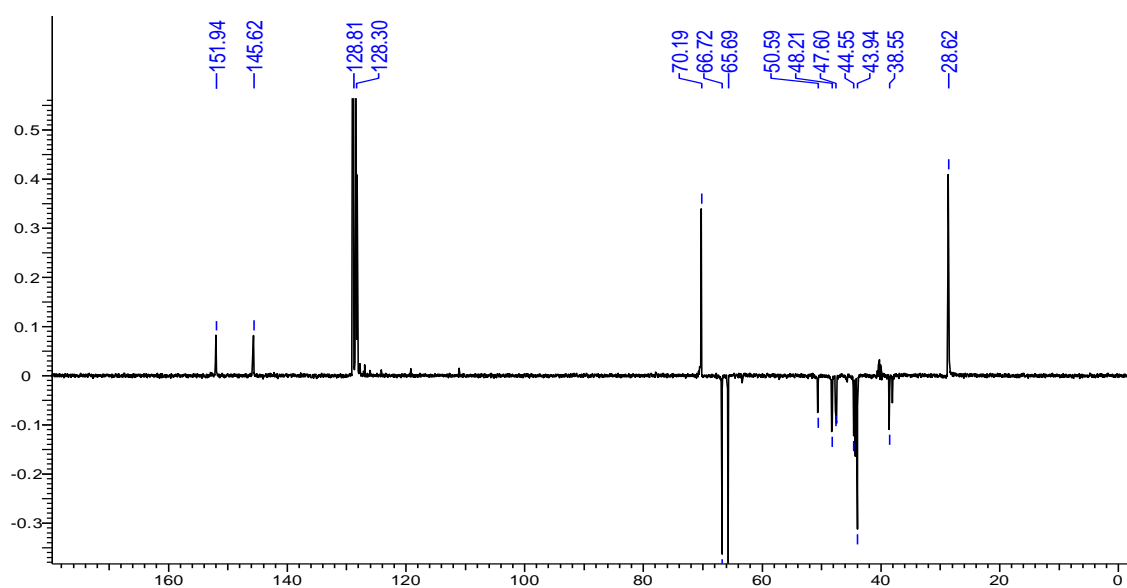
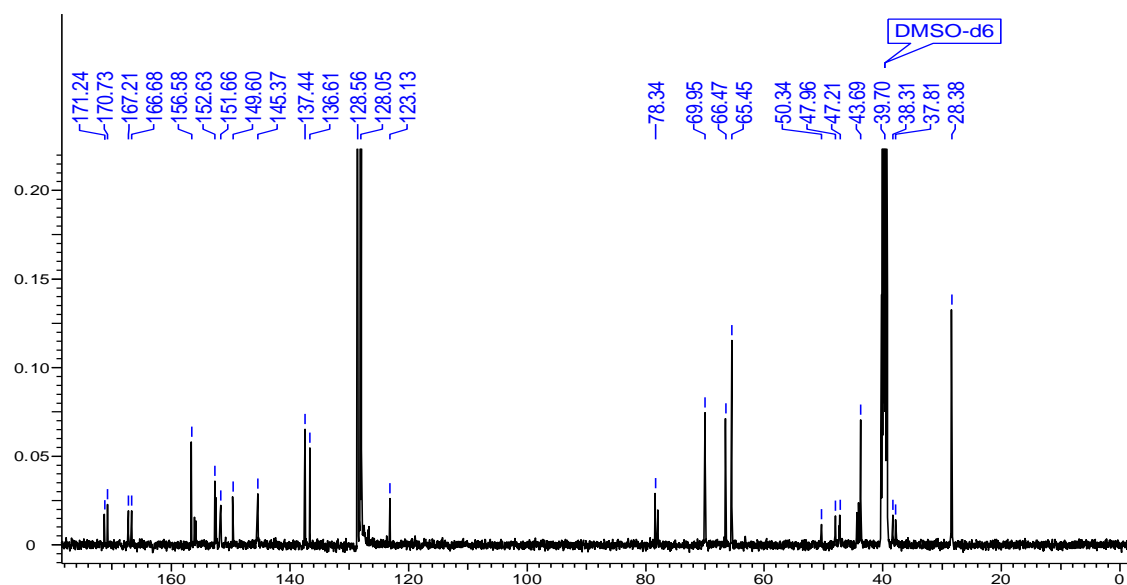
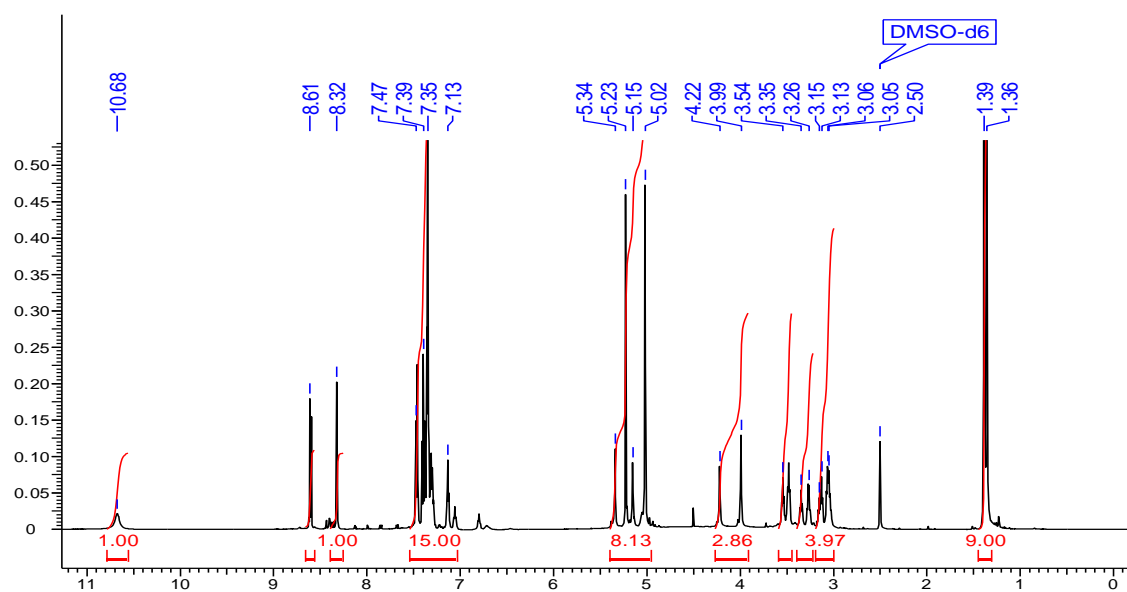
^1H , ^{13}C and DEPT of compound **9**

^1H , ^{13}C and DEPT of compound **10**

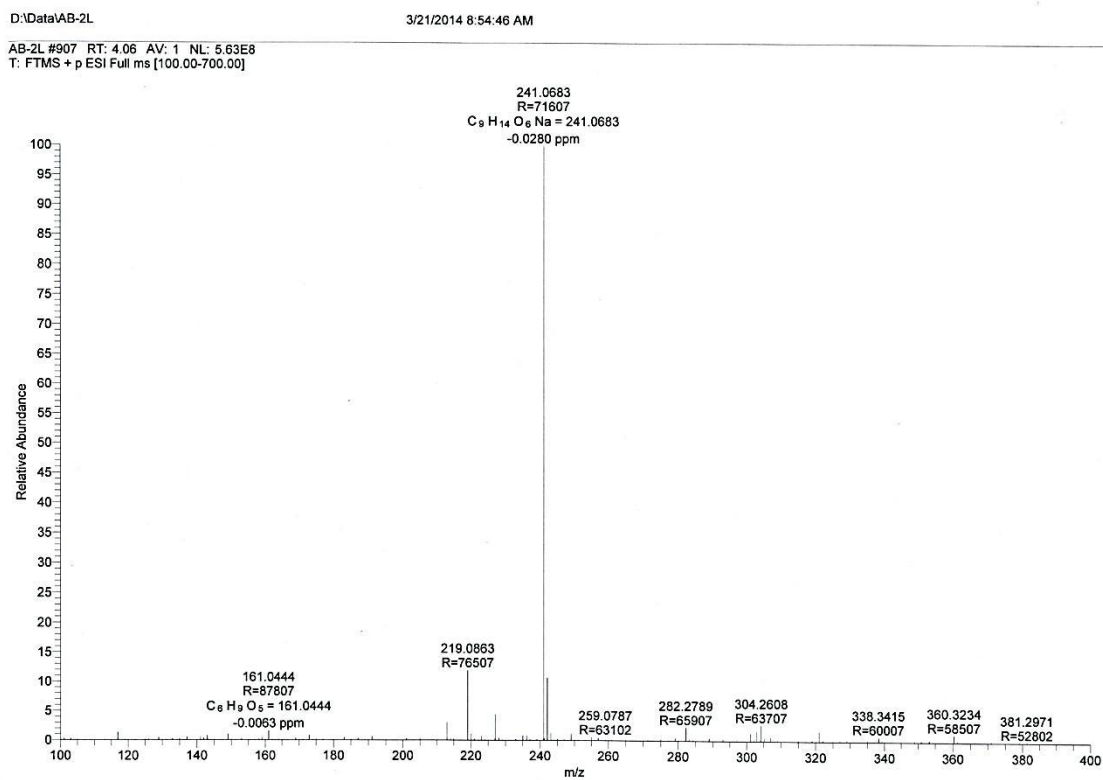
^1H , ^{13}C and DEPT of compound **11**

^1H , ^{13}C and DEPT of compound **14**

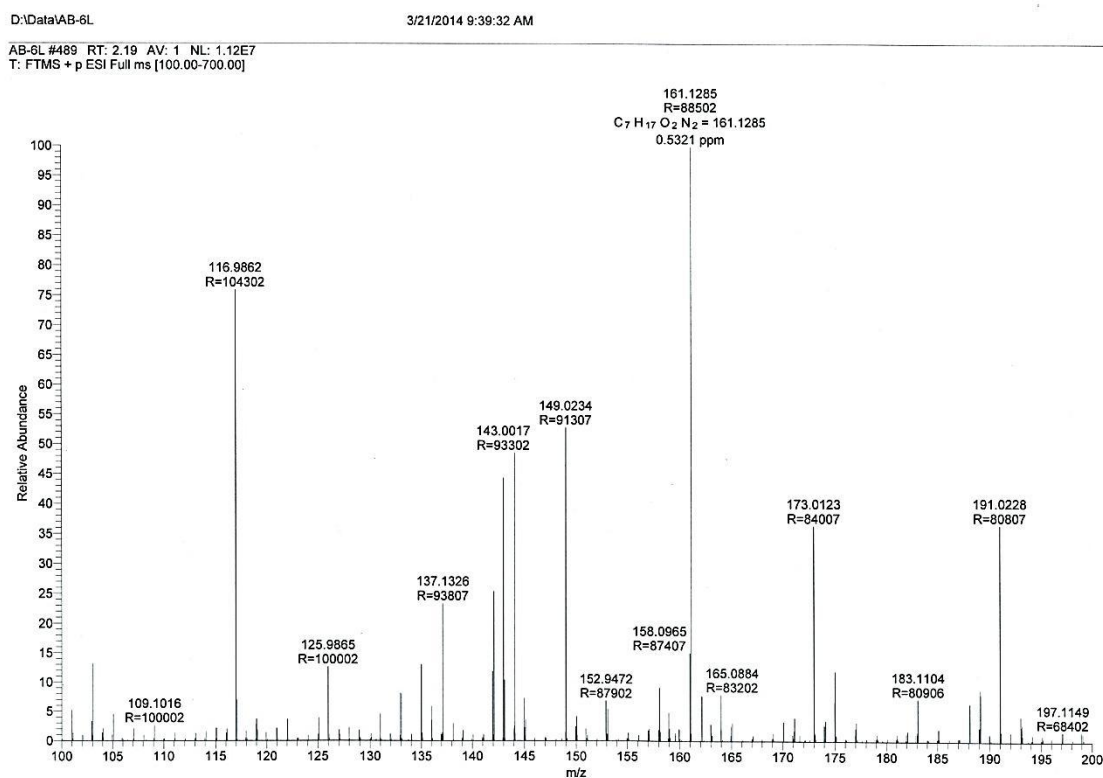
^1H , ^{13}C and DEPT of compound 16a

^1H , ^{13}C and DEPT of compound **16b**

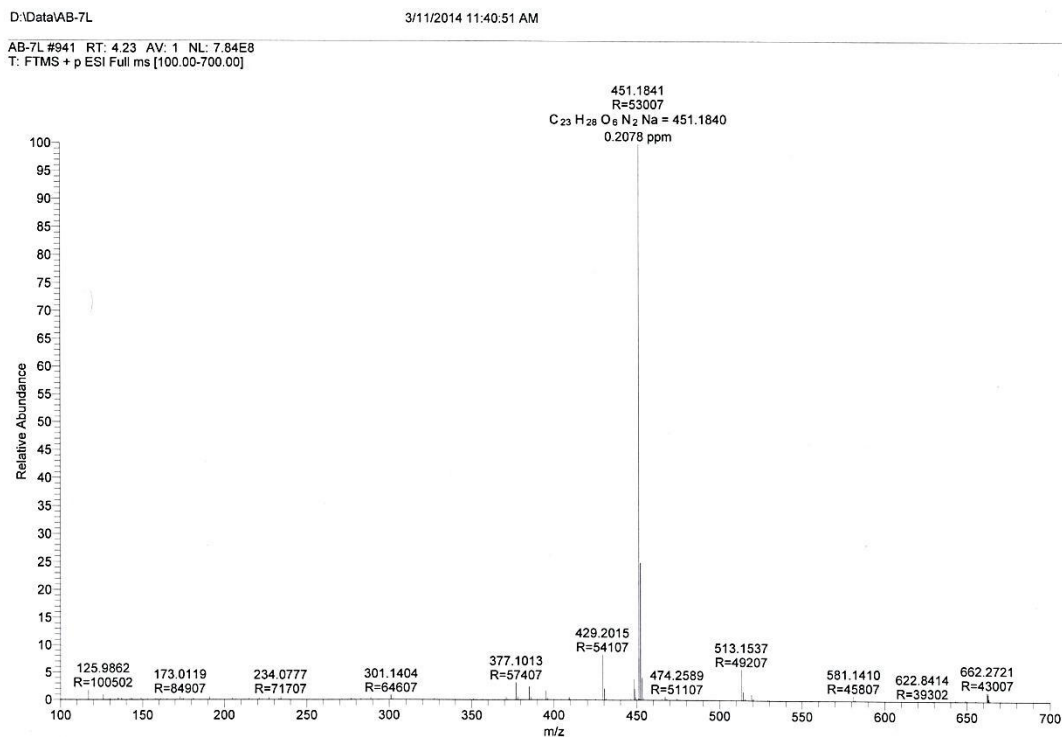
HRMS of compound 2



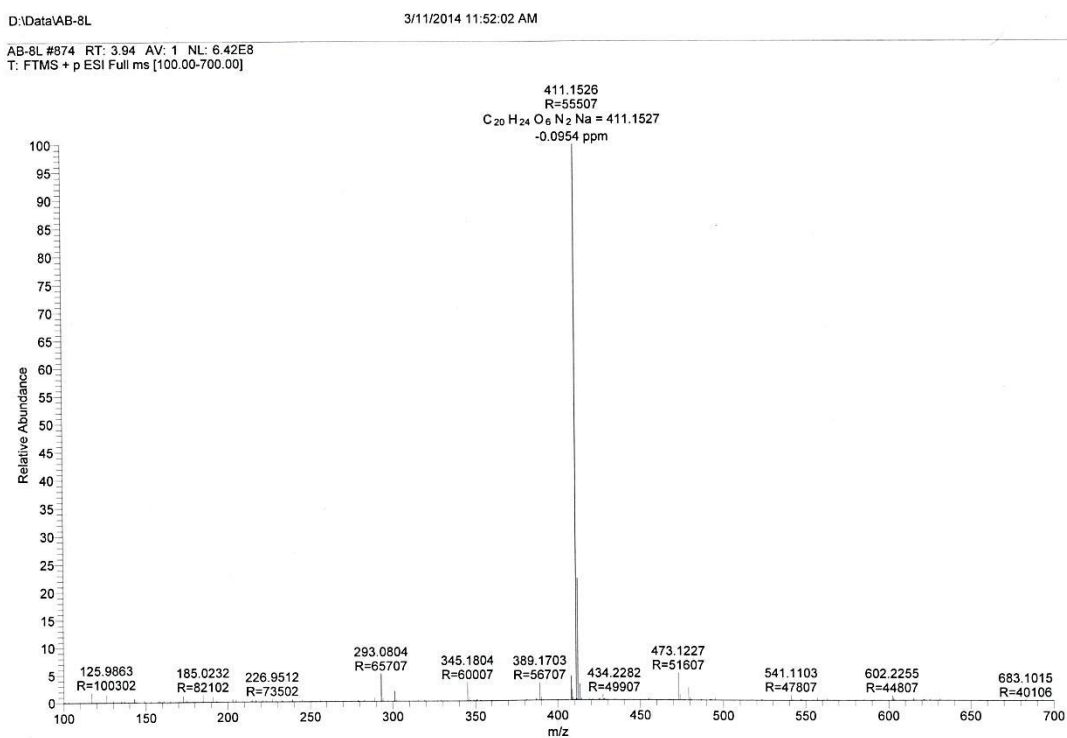
HRMS of compound 6



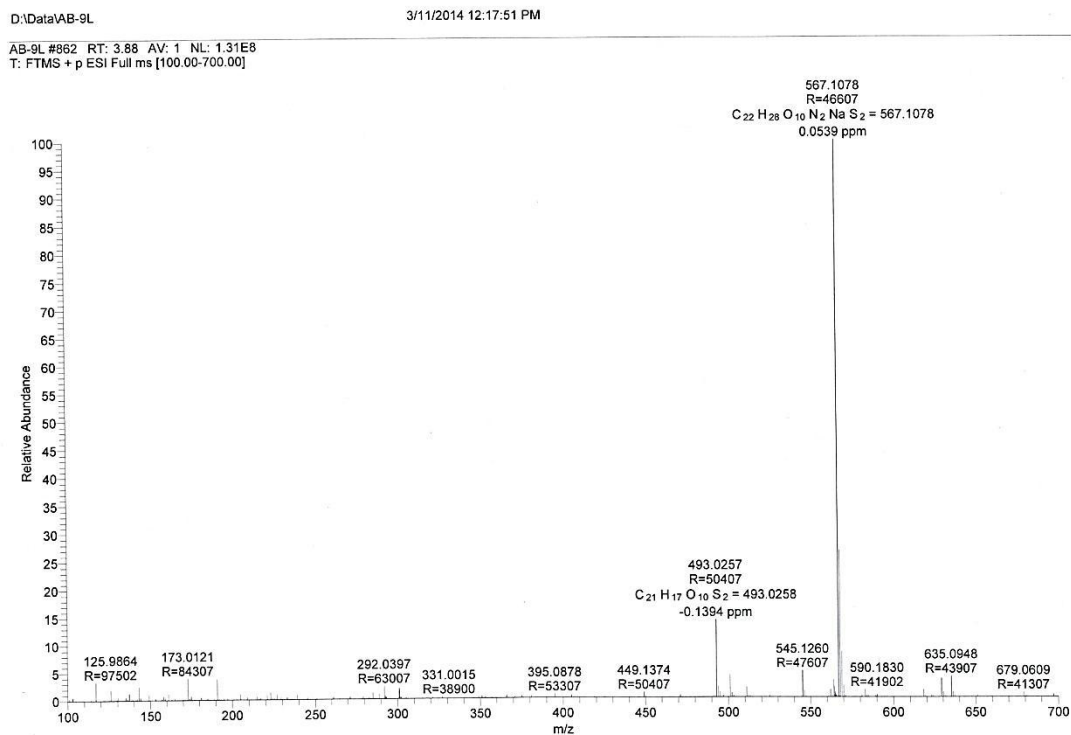
HRMS of compound 7



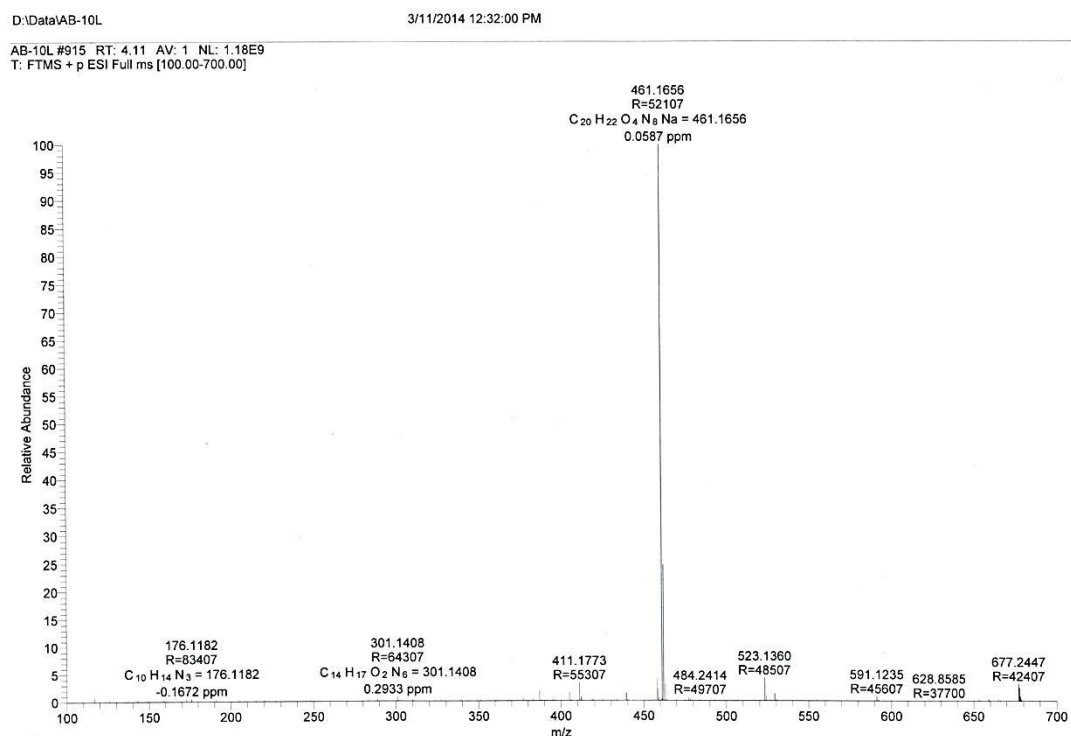
HRMS of compound 8



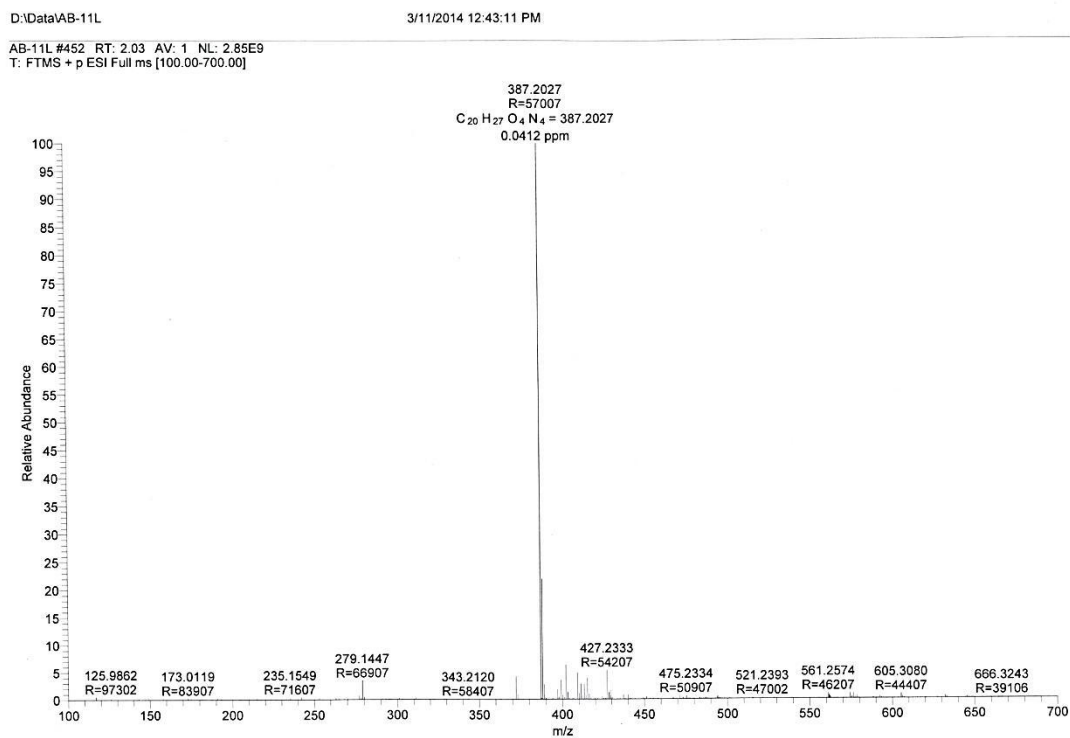
HRMS of compound 9



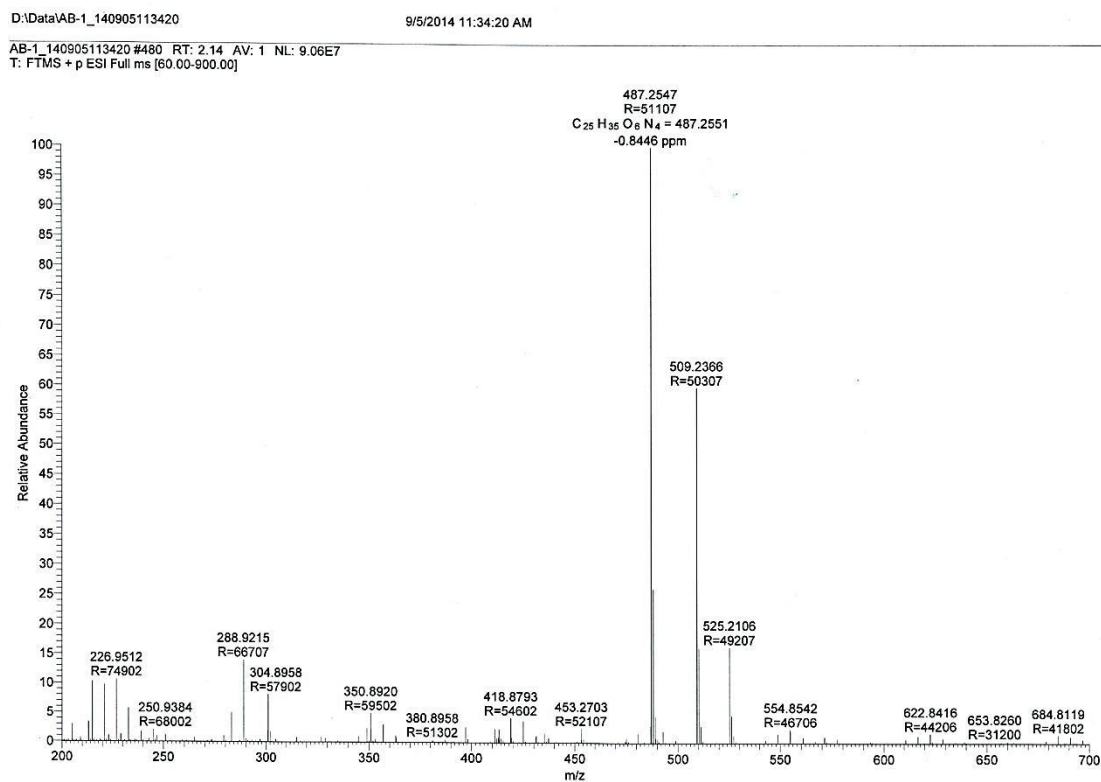
HRMS of compound 10



HRMS of compound 11



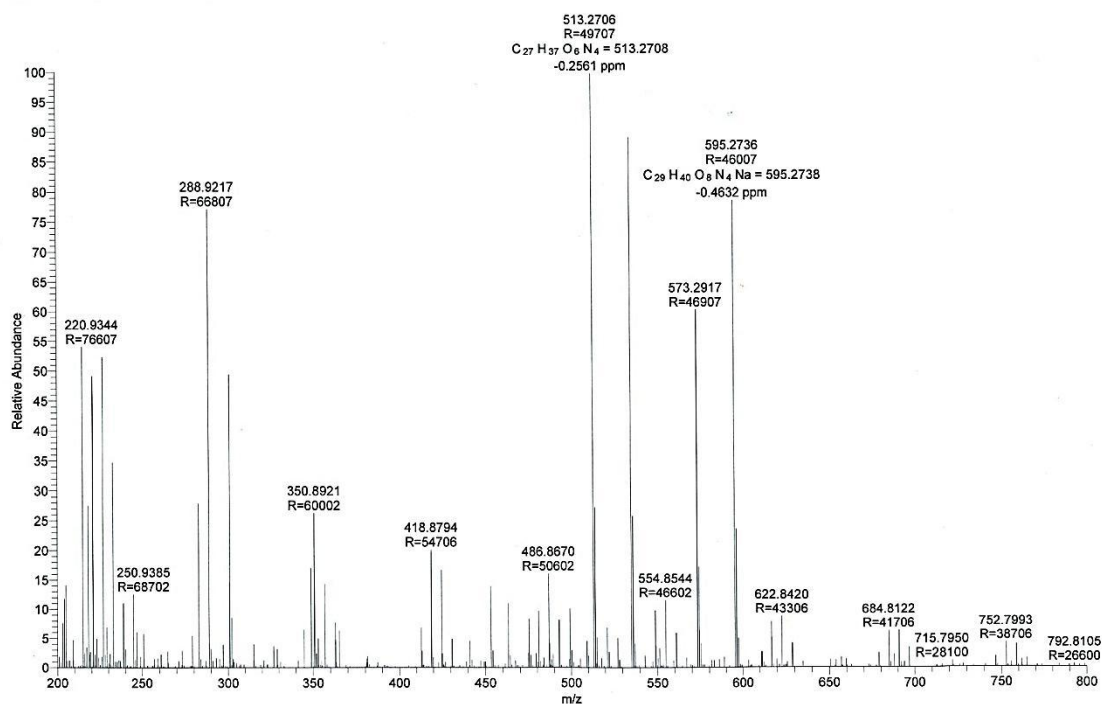
HRMS of compound 12



HRMS of compound 13

D:\Data\AB-2

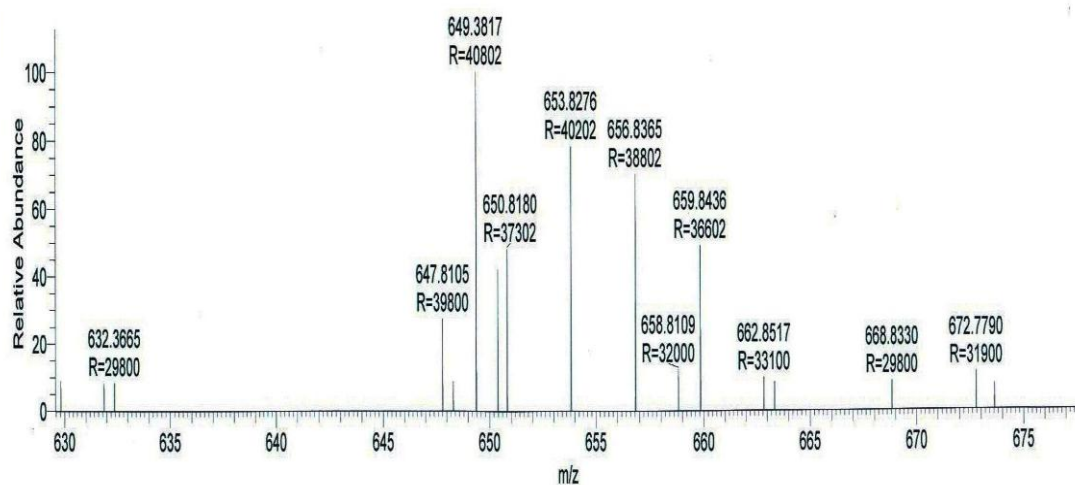
9/5/2014 11:45:32 AM

AB-2 #799 RT: 3.58 AV: 1 NL: 1.70E7
T: FTMS + p ESI Full ms [60.00-900.00]

HRMS of compound 14

D:\Data\AB-3

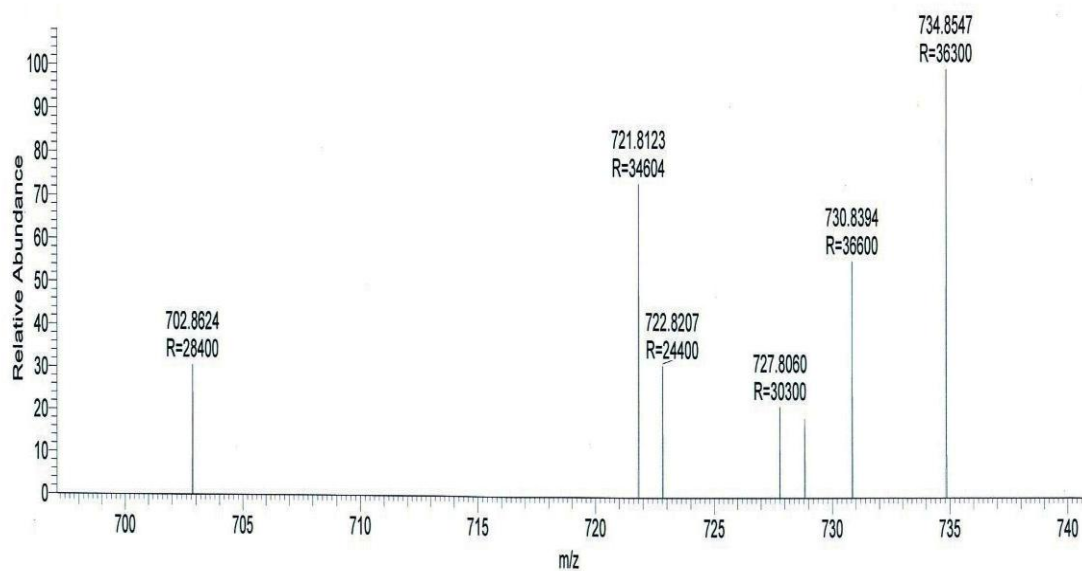
9/5/2014 11:56:40 AM

AB-3 #284 RT: 1.26 AV: 1 NL: 4.45E5
T: FTMS + p ESI Full ms [60.00-900.00]

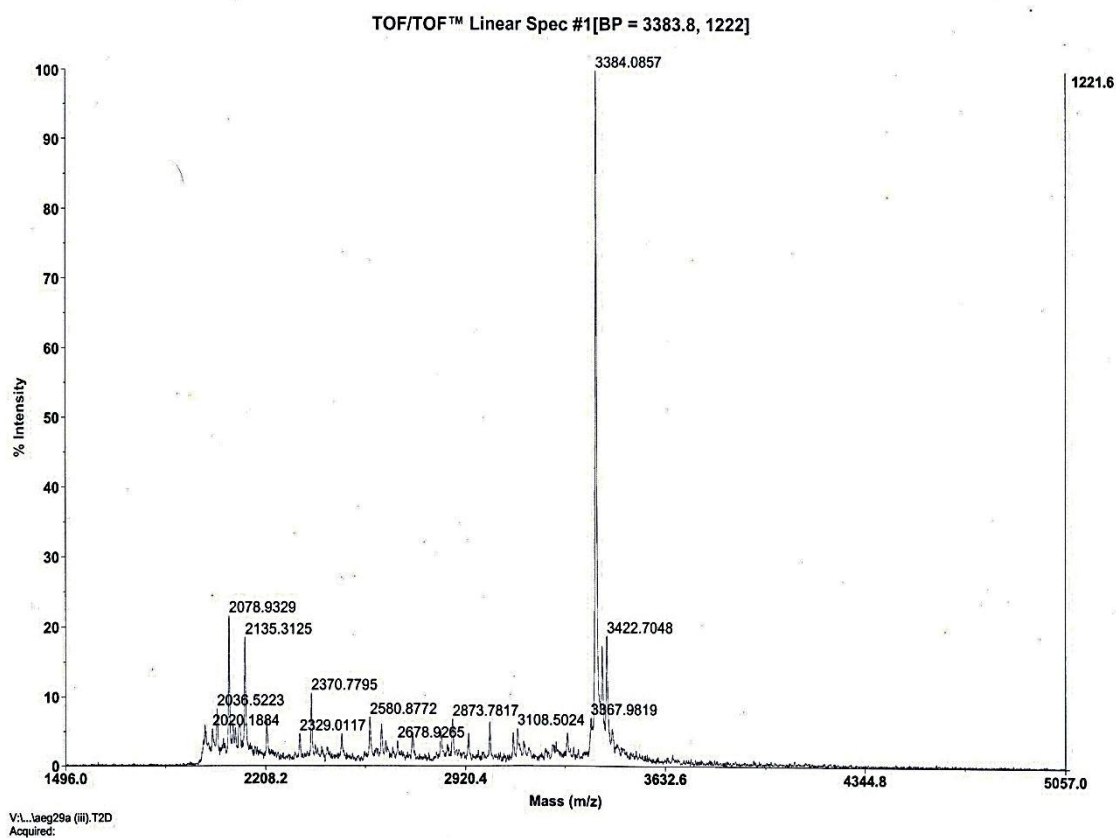
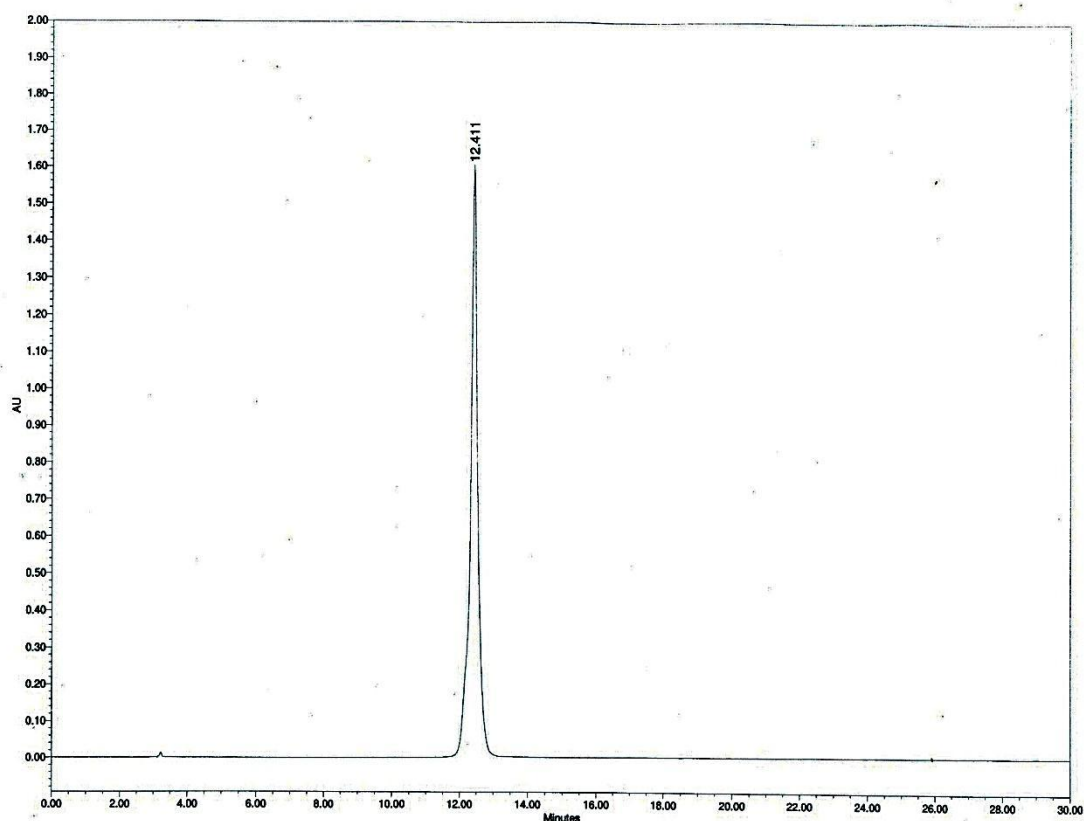
HRMS of compound **16a**

D:\Data\AB-5

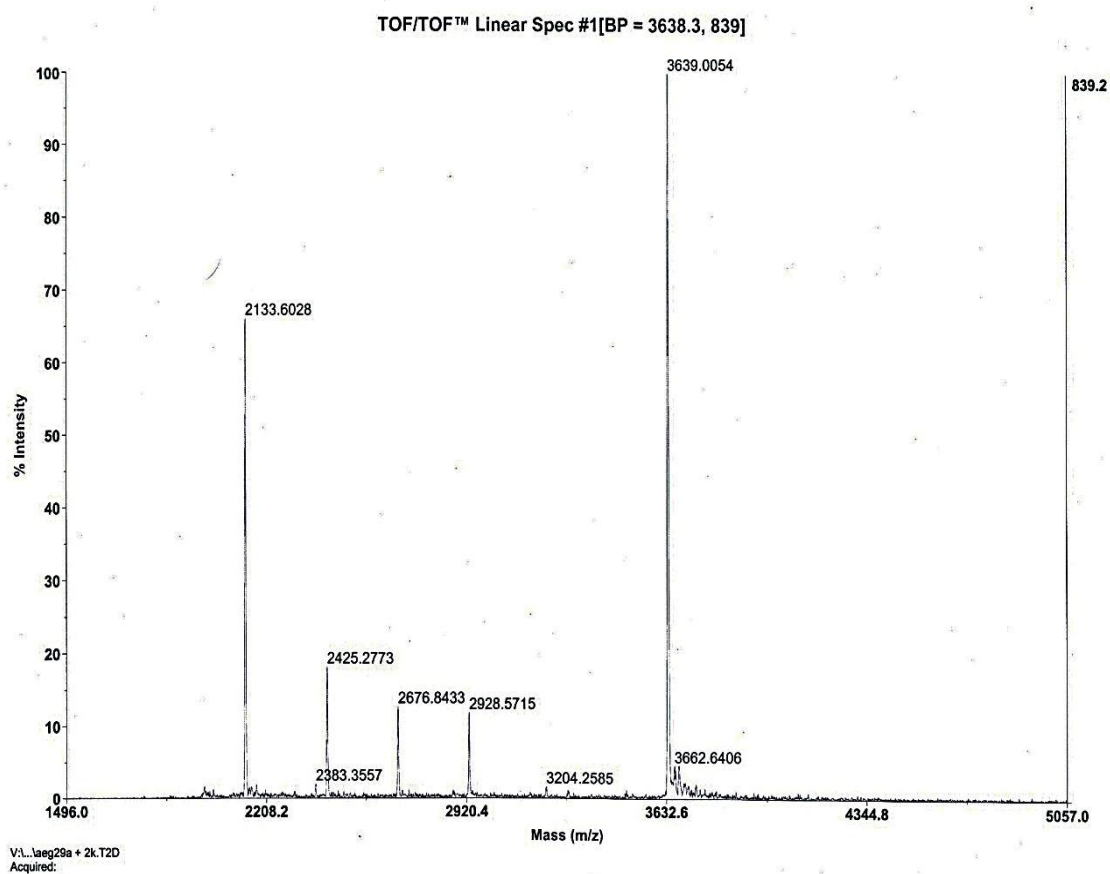
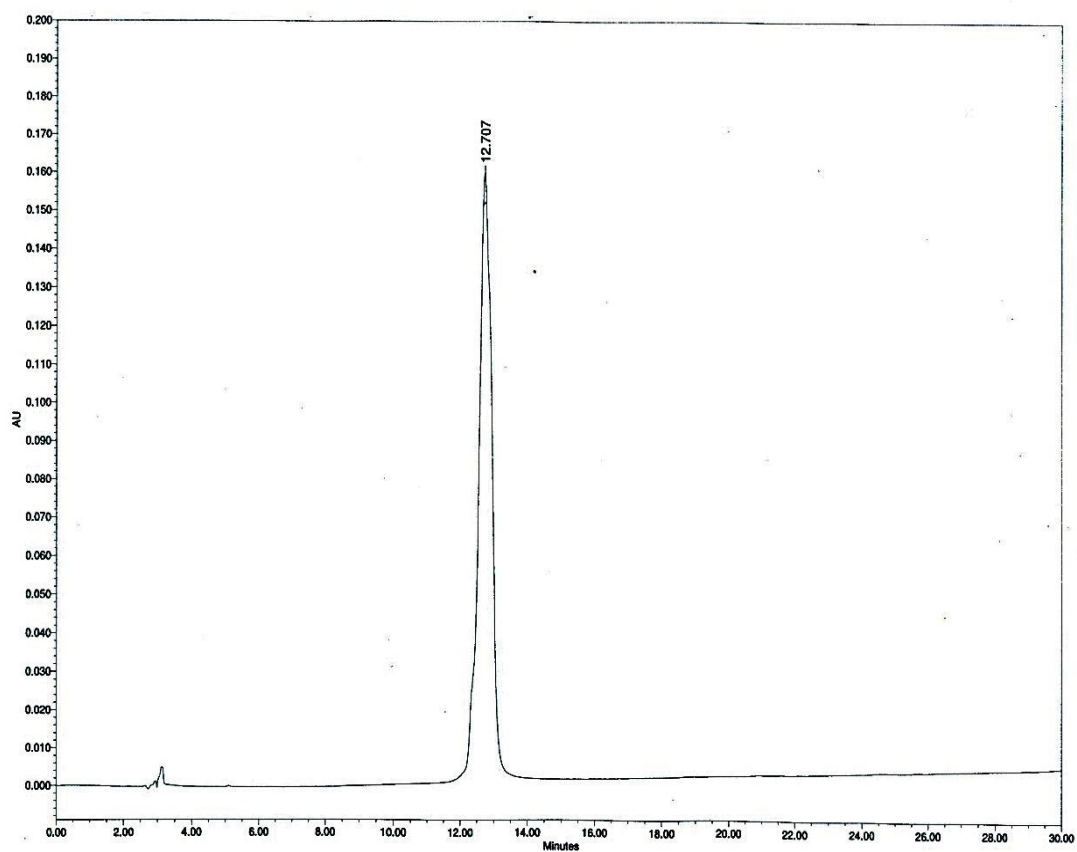
9/5/2014 12:18:56 PM

AB-5 #878 RT: 3.91 AV: 1 NL: 2.09E5
T: FTMS + p ESI Full ms [60.00-900.00]

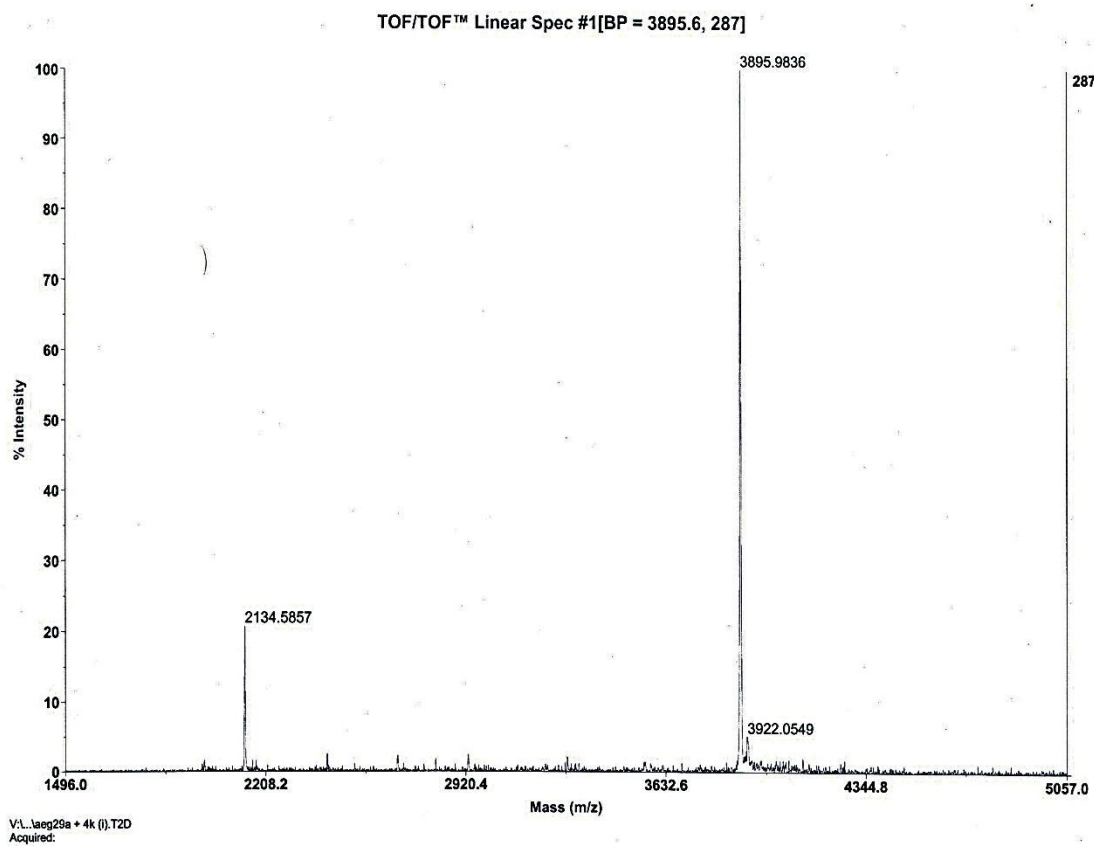
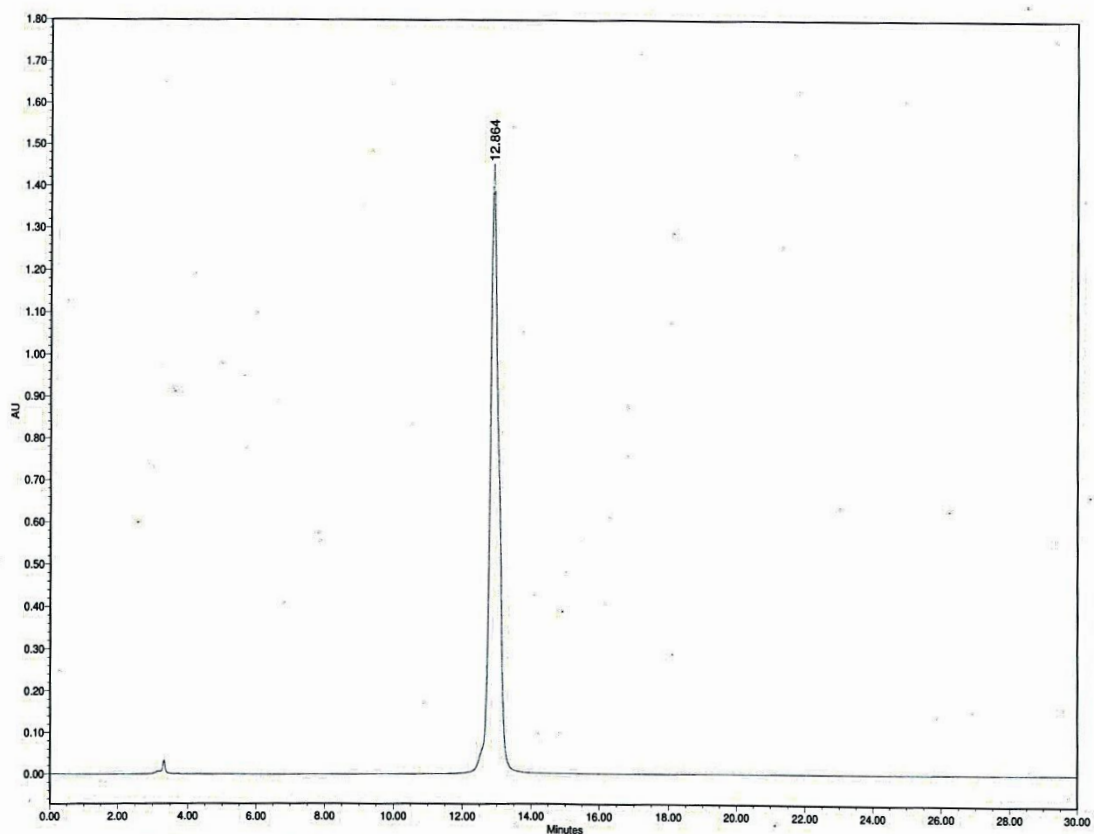
HPLC and MALDI-TOF of PNA 1



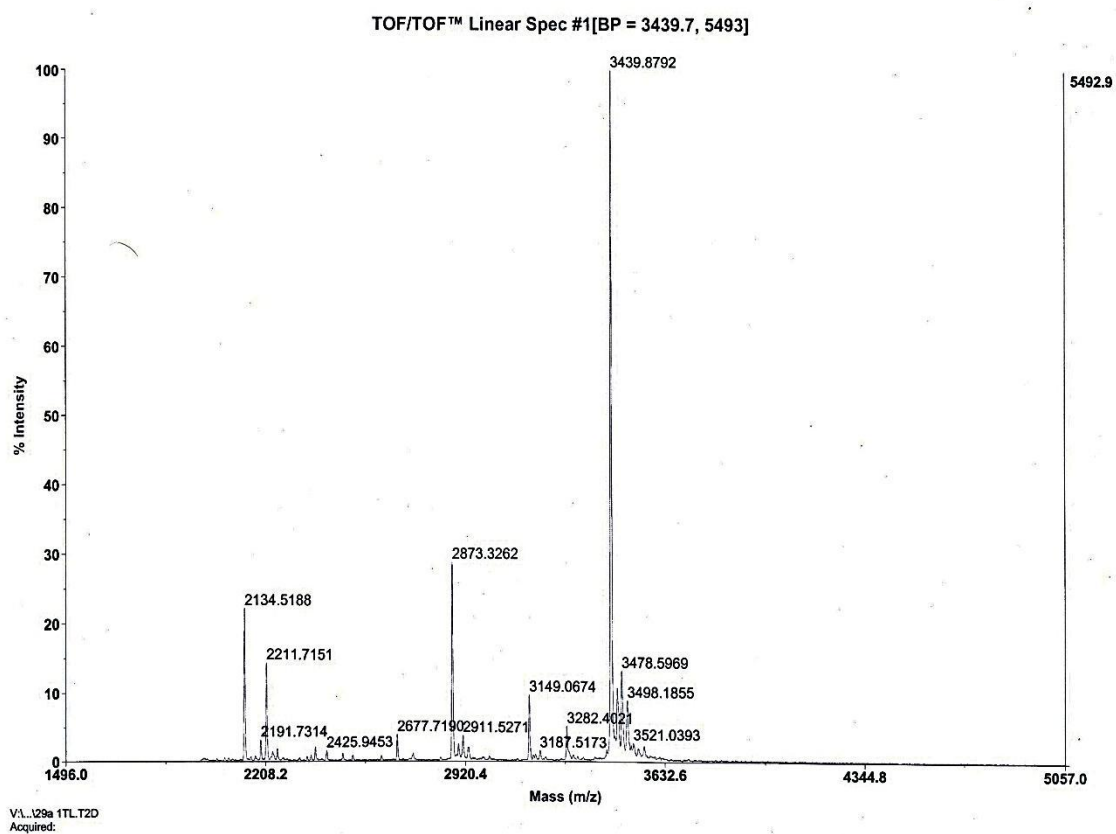
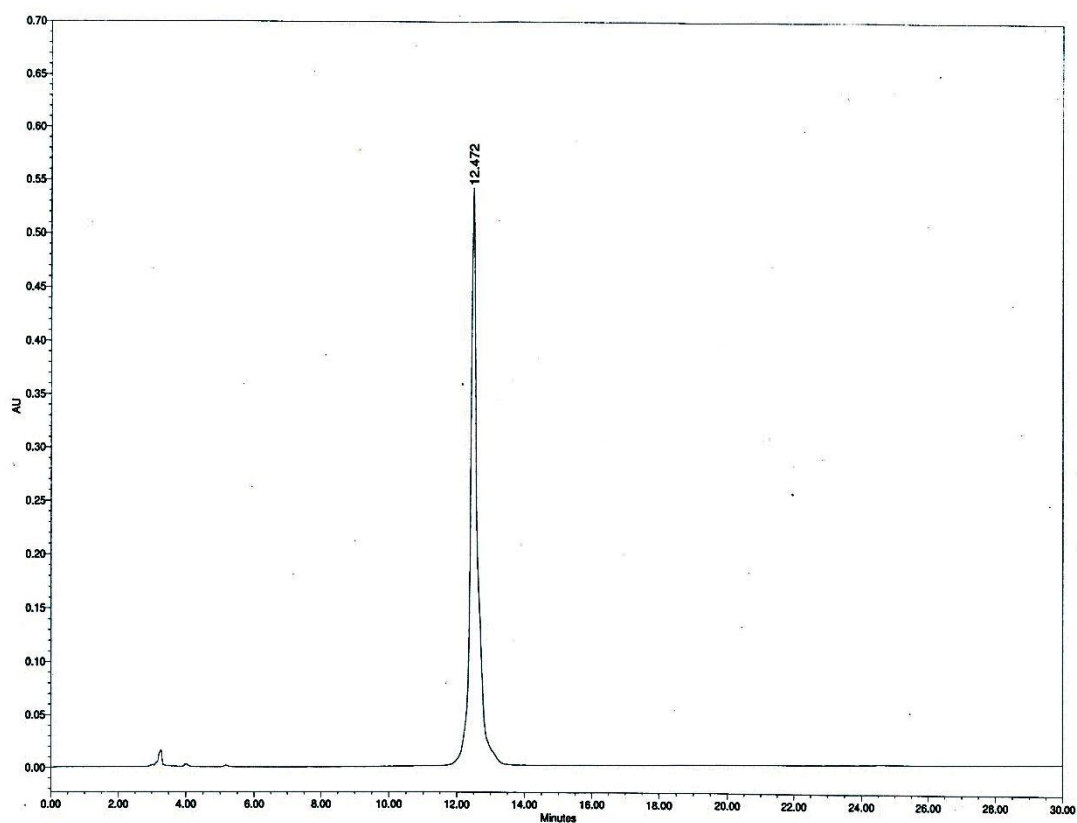
HPLC and MALDI-TOF of PNA 2



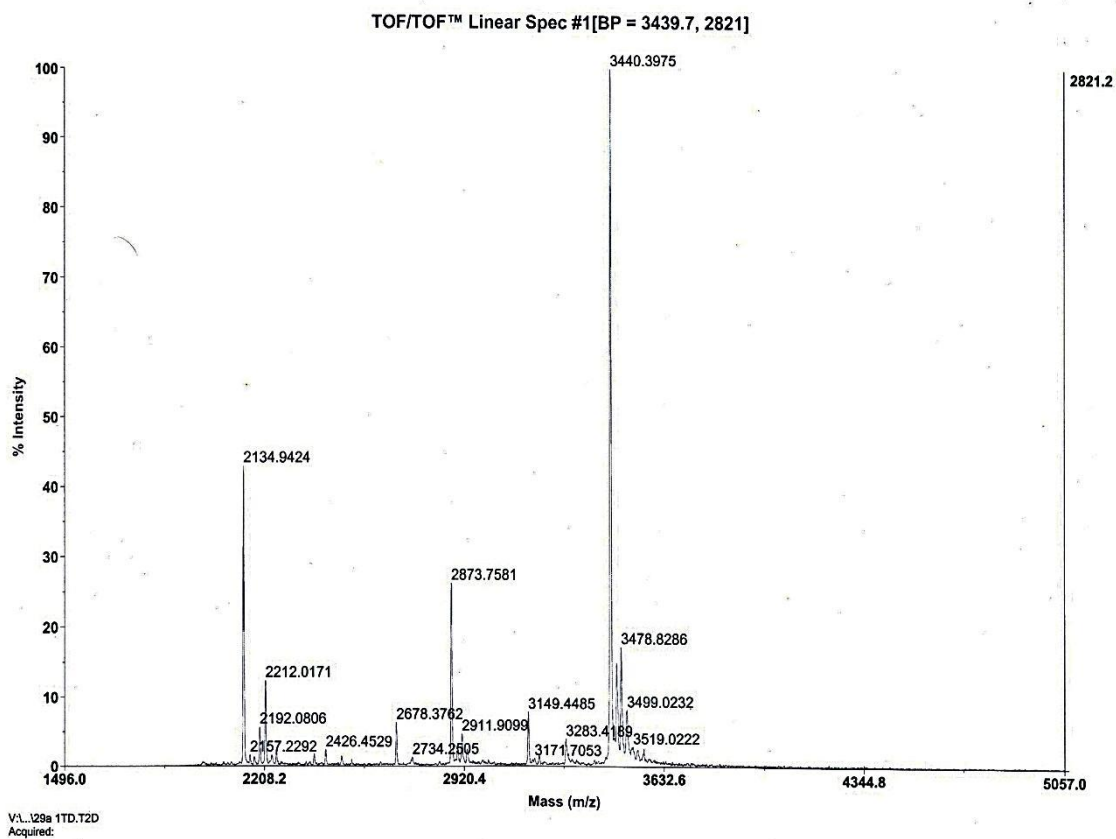
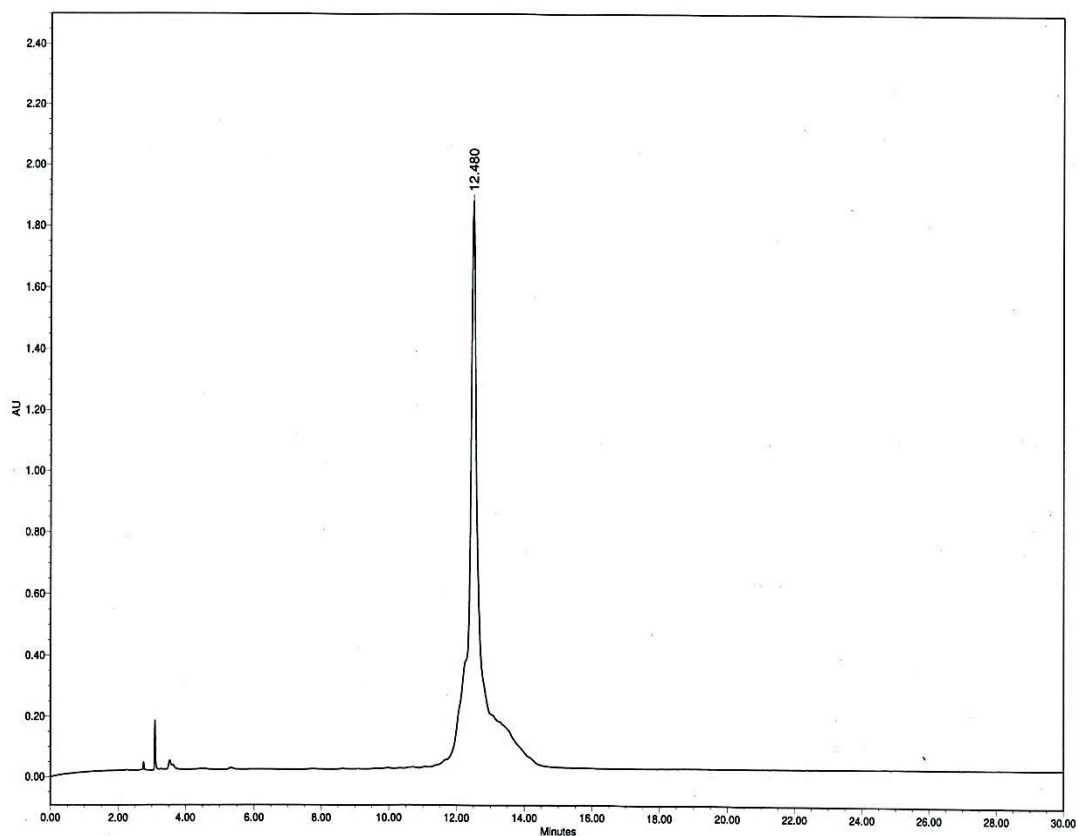
HPLC and MALDI-TOF of PNA 3



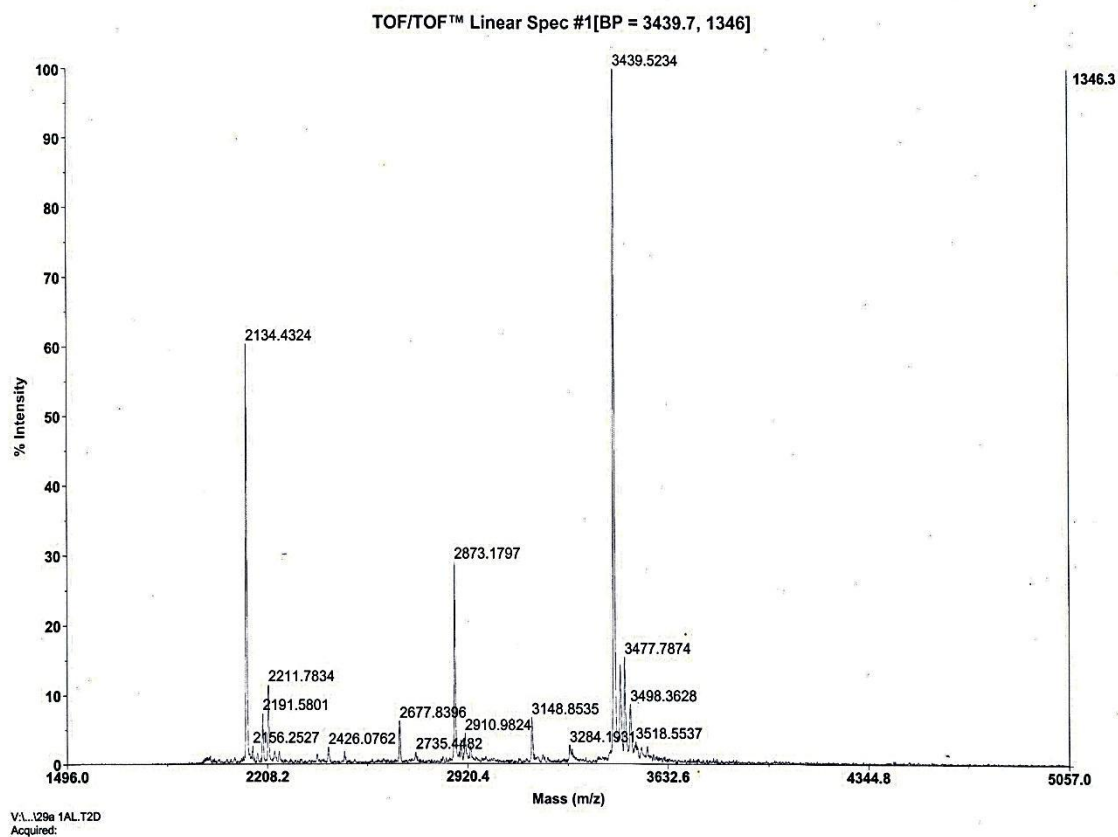
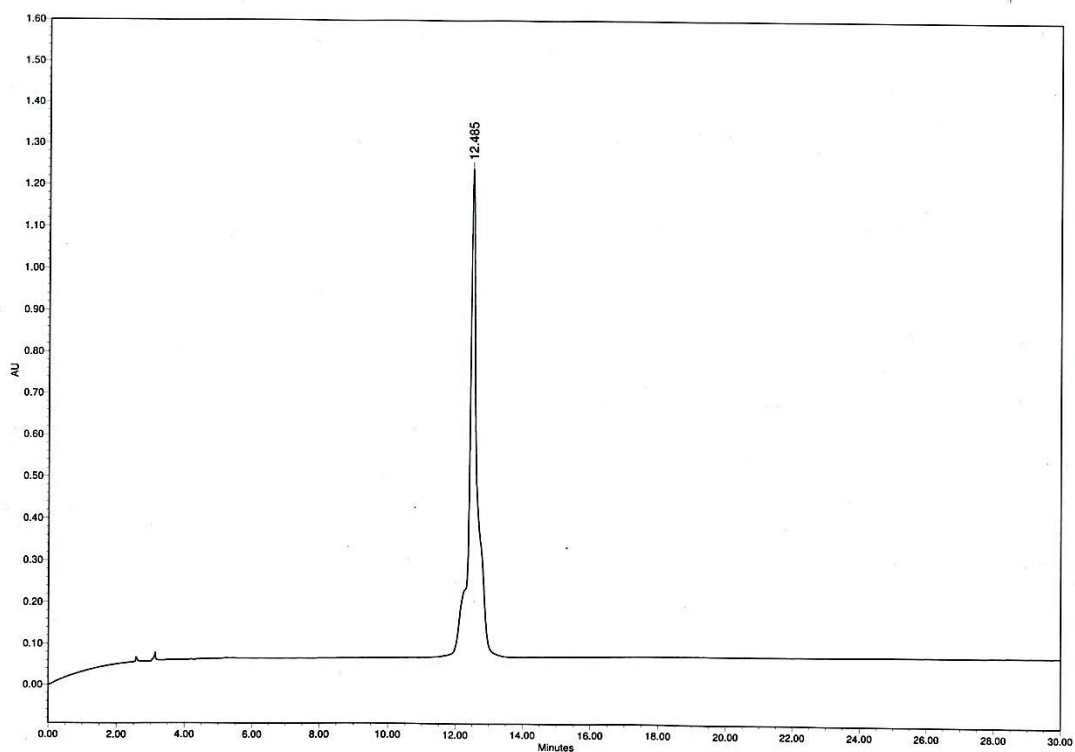
HPLC and MALDI-TOF of PNA 4



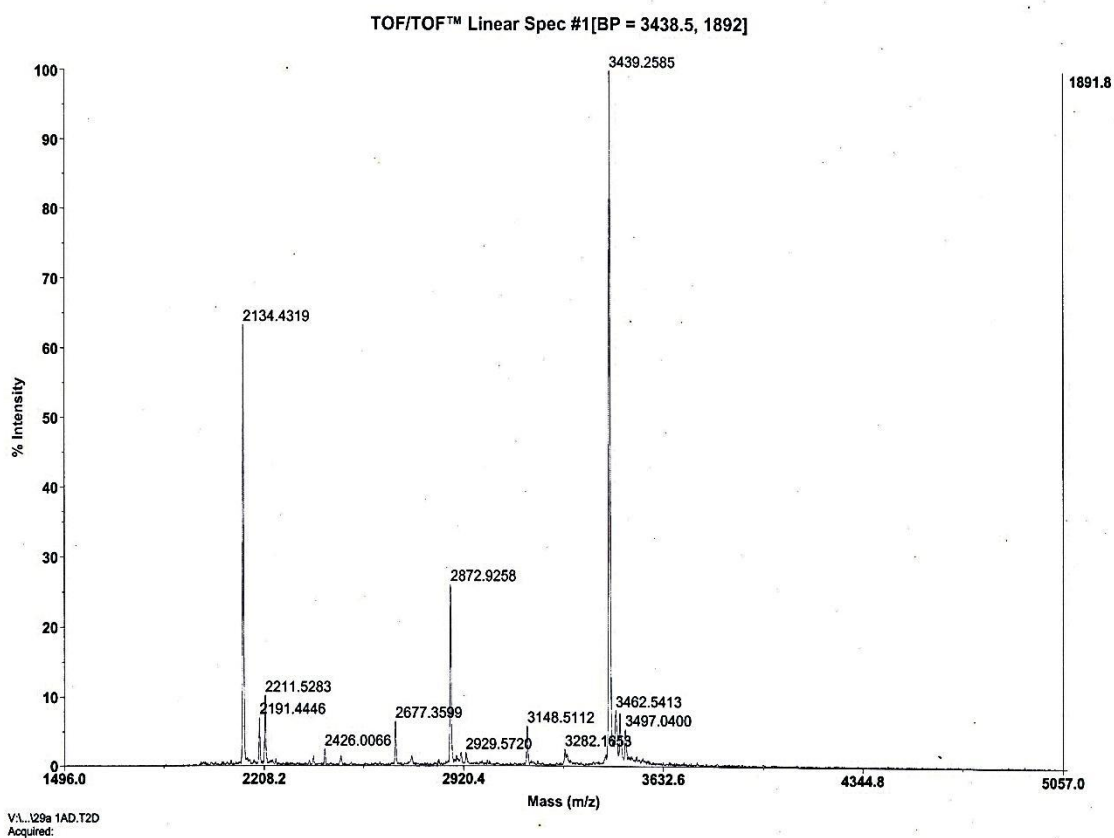
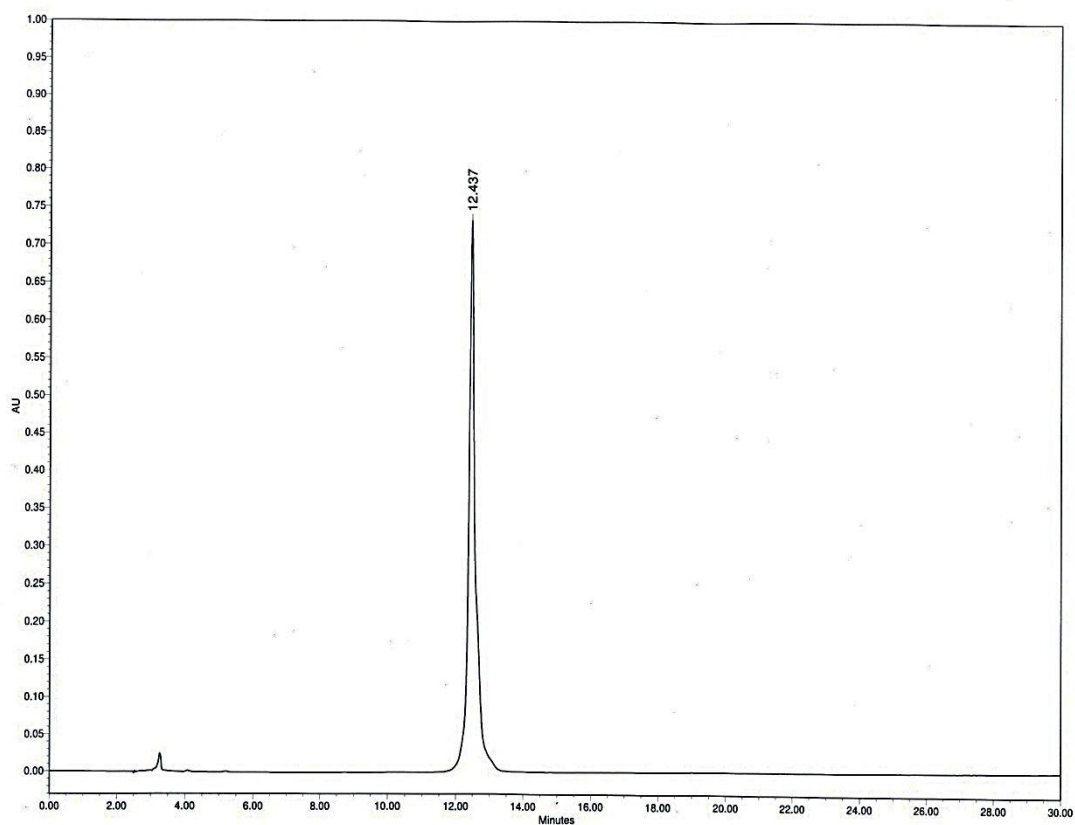
HPLC and MALDI-TOF of PNA 5



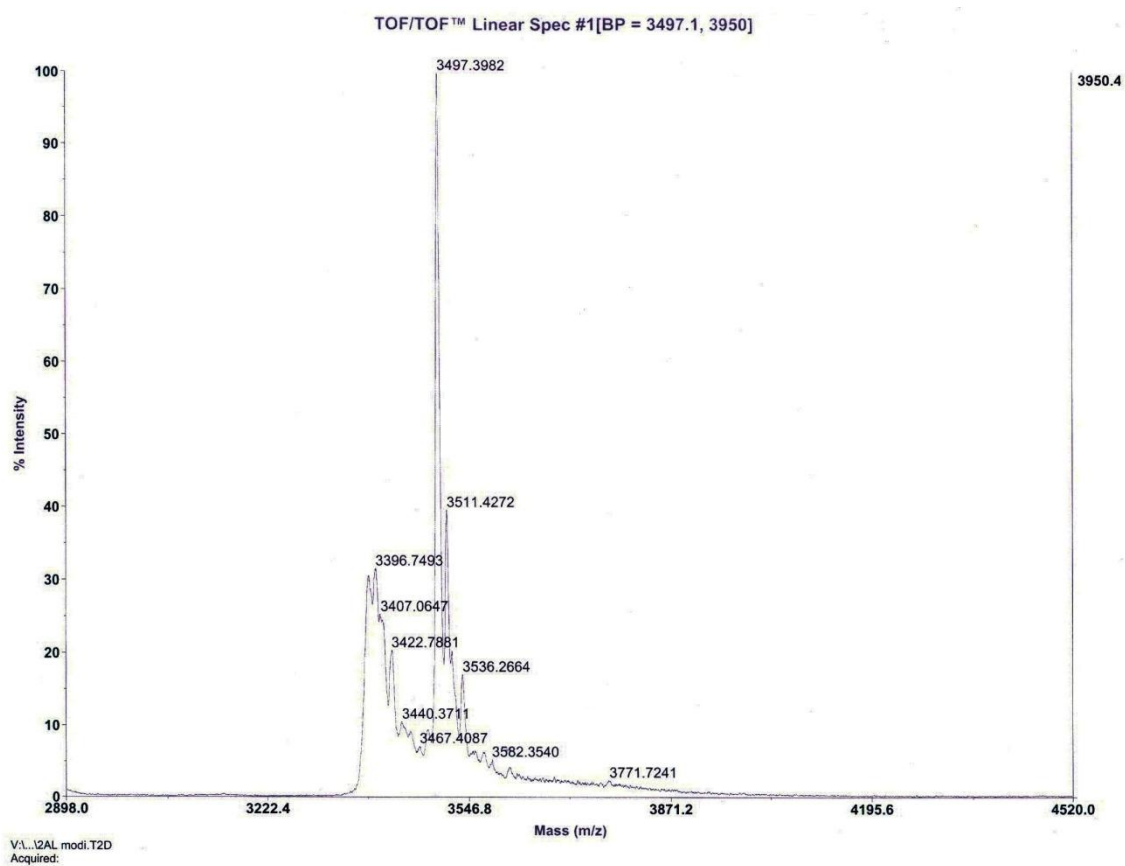
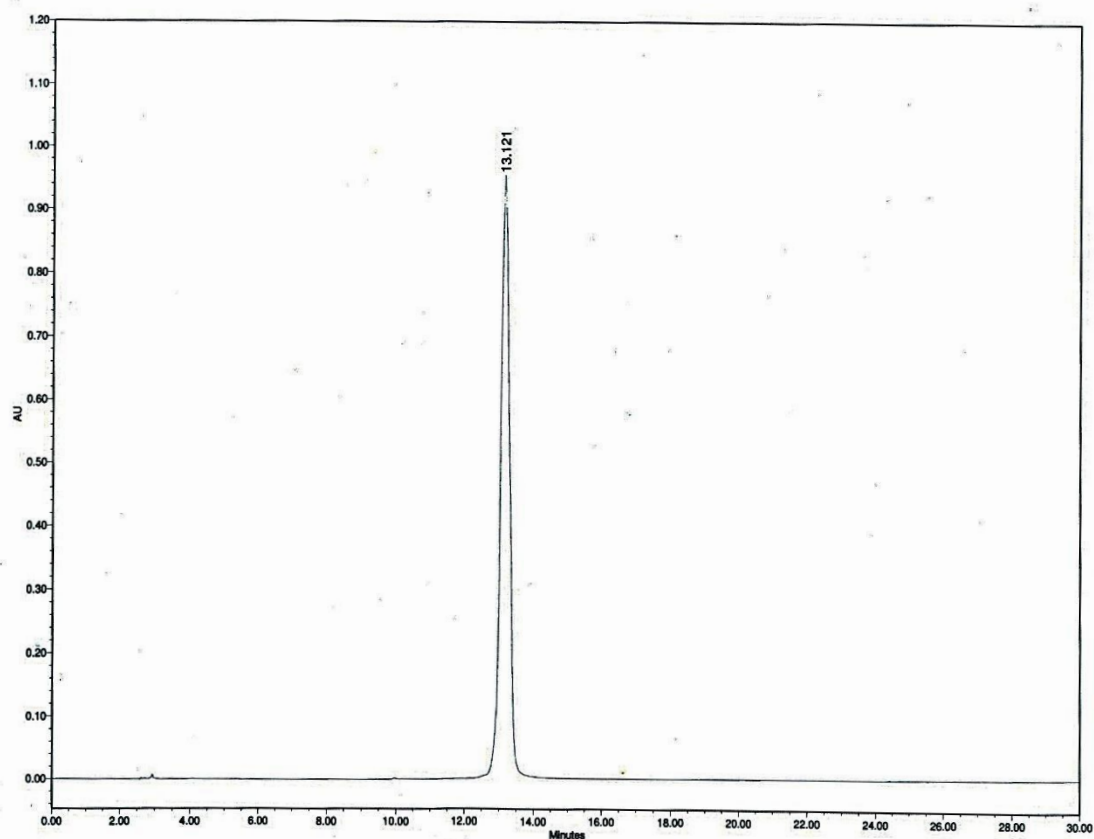
HPLC and MALDI-TOF of PNA 6



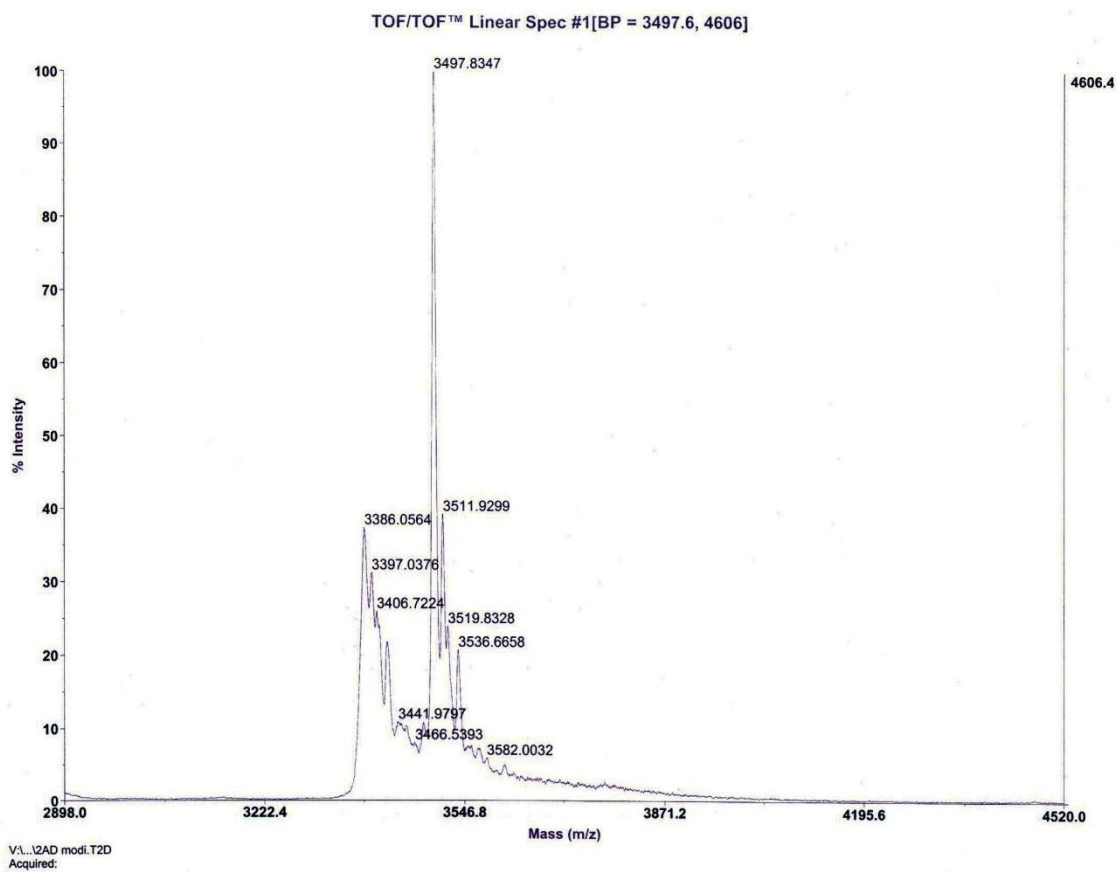
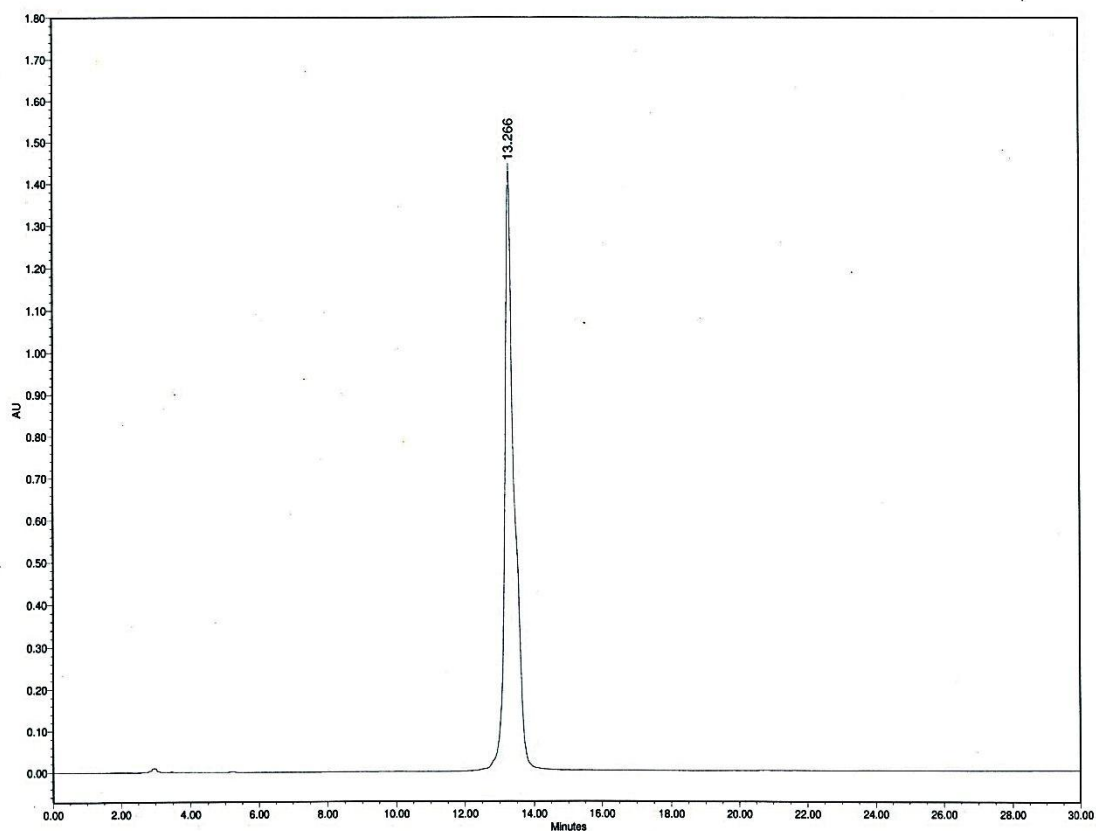
HPLC and MALDI-TOF of PNA 7



HPLC and MALDI-TOF of PNA 8



HPLC and MALDI-TOF of PNA 9



Section B

Biophysical studies of modified PNA towards comp. DNA/RNA

3B Introduction

The purpose of the designed oligonucleotides is to achieve high binding affinity and specificity to target nucleic acid and to increase aqueous solubility and in turn cellular internalization. All newly designed and synthesized molecules were hence subjected to various biophysical and biochemical studies to evaluate their potential as antisense agents.

3B.1 Rationale and objectives of the present work

The preceding section reports the synthesis of β - and γ - bisubstituted T/A^{cbz} monomers. These monomer units were introduced at various positions in PNA oligomers in order to study the effect of cationic nature in the PNA backbone. In this Section the results of the systematic study of β - and γ - bisubstituted backbone modifications in terms of their ability and strength to form specific base paired complexes with DNA/RNA is presented. The specificity of binding was addressed by performing experiments with mismatched pairing at a single site in the target DNA/RNA sequences. This Section addresses the biophysical studies of modified PNA oligomers and their hybrids with complementary nucleic acids using techniques such as UV spectroscopy and CD spectrophotometry.

3B.2 Biophysical spectroscopic techniques for studying PNA- DNA/ PNA-RNA Interactions

As discussed in the chapter 1 as well as in chapter 2, monitoring the UV absorption at 260 nm as a function of temperature has been extensively used to study the thermal stability of nucleic acid duplexes and triplexes and consequently, PNA-DNA hybrids as well. Increasing the temperature perturbs these complexes, inducing a structural transition by causing disruption of hydrogen bonds between the base pairs, diminished stacking interactions between adjacent nucleobases and larger torsional motions in the backbone leading to a loss of secondary and tertiary structure. This is reflected in the UV absorption at 260 nm, termed as hyperchromicity. The magnitude of hyperchromicity is a measure of the extent of the secondary structure present in nucleic acids. The process is cooperative and the plot of the absorbance at 260 nm vs the temperature is therefore sigmoidal. A non-sigmoid (e.g. linear) transition with low

hyperchromicity is a consequence of non duplexation (non complementation). In many cases, the transitions are broad and the exact T_m s are obtained from the peak in the first derivative plots. This technique has provided valuable information regarding complementary interactions in nucleic acids hybrid involving DNA, RNA and PNA.⁶¹

3B.3 UV- T_m studies on modified PNA:cDNA and modified PNA:RNA

The synthesized PNA sequences were examined for their binding affinity with complementary DNA/RNA sequences by employing UV-thermal denaturation studies. The sequence chosen for study was a 12mer sequence (antimiR 29a). The modified oligomers were annealed with the cDNA and RNA and were subjected to temperature dependent UV studies at 260 nm. Unmodified **PNA 1** was used as a control in these experiments. The unmodified **PNA 2** and **PNA 3** were also served as control for comparing the cationic nature in the PNA backbone. The UV- T_m plots show a single sigmoidal transition, characteristic of PNA:cDNA/RNA duplex melting. The T_m values were obtained by the first derivative and the values corresponding to peaks from such plots for various PNA:cDNA/RNA duplexes of modified and unmodified oligomers are shown in Table 2 and Figure 3 and 4.

Table 2. PNA sequences and UV- T_m studies with cDNA and cRNA

Entry No.	Code	Sequences (N' → C')	UV - T_m (°C)		ΔT_m (°C)	
			cDNA	cRNA	cDNA	cRNA
1.	PNA 1	H - AACCGATTTCAG -Lys-NH ₂	58.7	64.5	-	-
2.	PNA 2	H - (K) ₂ - AACCGATTTCAG -Lys-NH ₂	60.1	66.6	+ 1.4	+ 2.1
3.	PNA 3	H - (K) ₄ - AAC CGATTTCAG -Lys-NH ₂	61.5	68.6	+ 2.8	+ 4.1
4.	PNA 4	H- AACCGA t_L^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	- 58.7	- 64.5
5.	PNA 5	H- AACCGA t_D^{Am} TTCAG -Lys-NH ₂	n.d.	25.9	- 58.7	-38.6
6.	PNA 6	H- A a_L^{Am} CCGATTTCAG -Lys-NH ₂	32.3	36.2	- 26.4	- 28.3
7.	PNA 7	H- A a_D^{Am} CCGATTTCAG -Lys-NH ₂	42.2	48.9	- 16.5	- 15.6
8.	PNA 8	H- a_L^{Am} a_L^{Am} CCGATTTCAG -Lys-NH ₂	35.9	41.7	- 22.8	- 22.8
9.	PNA 9	H- a_D^{Am} a_D^{Am} CCGATTTCAG -Lys-NH ₂	45.7	50.6	- 13.0	- 13.9

t_L^{Am}/ a_L^{Am} and t_D^{Am}/ a_D^{Am} are the modified monomers derived from L- and D - tartaric acid resp.

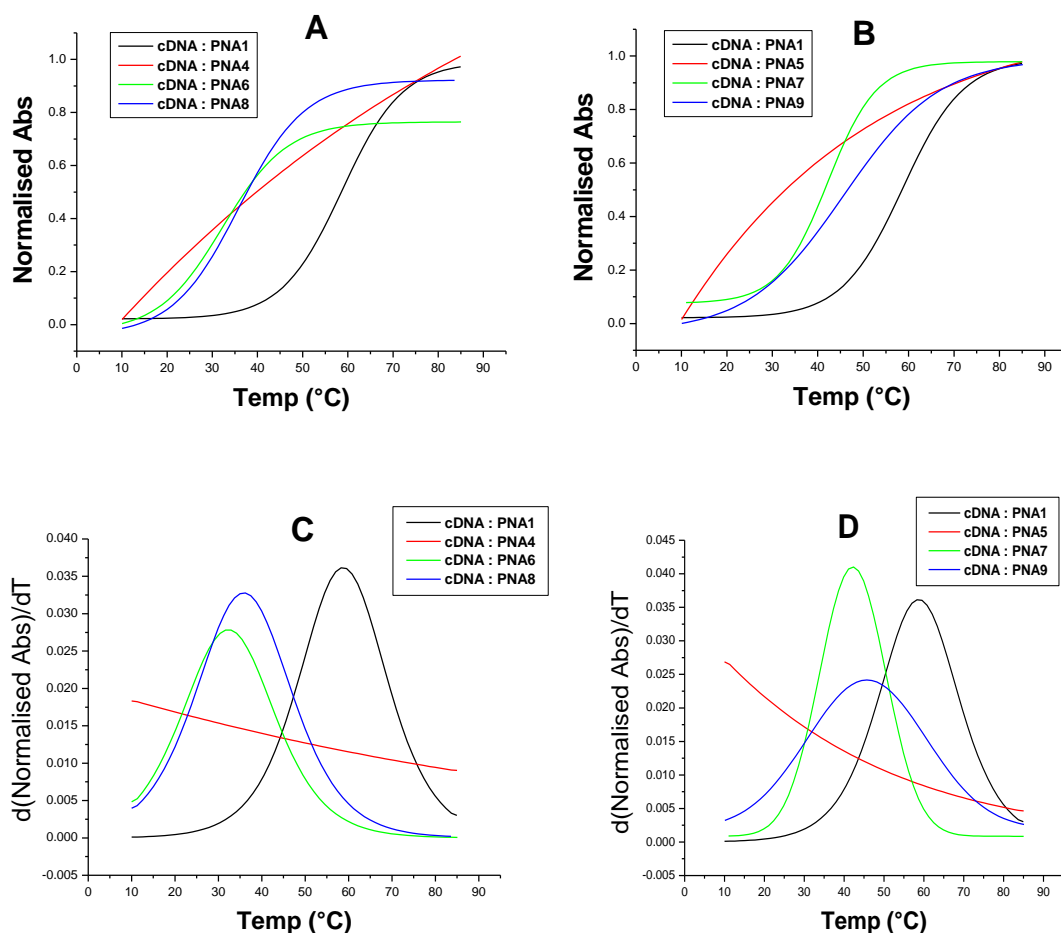


Figure 3. UV-Thermal Melting Profiles. Sigmoidal Curves for (A) cDNA:PNA1/ PNA4/ PNA6/ PNA8 (B) cDNA:PNA1/ PNA5/ PNA7/ PNA9, first derivative curves of the sigmoidal curves for melting values (C) cDNA:PNA1/ PNA4/ PNA6/ PNA8 (D) cDNA:PNA1/ PNA5/ PNA7/ PNA9

It is seen from the UV-melting data that the introduction of single β -, γ -bisubstituted - T modification in the middle of the sequence (Entry 4 and 5, PNA 4 and PNA 5) completely destabilize PNA:DNA duplex over unmodified PNA 1:cDNA. However, the substitution near N - terminal of the PNA, having a single β -, γ -bisubstituted - A modification showed reduced binding with cDNA (Entry 6 and 7, PNA 6 and PNA 7), when compared with unmodified PNA: DNA duplex, the destabilization was found to be - 26.4 °C.

The UV- T_m studies involving complexation with RNA did not show binding towards cRNA for sequence derived from L - isomer, but some binding was seen with the D-isomer derived sequence, but with large ΔT_m (- 38.6 °C) when compared with the unmodified PNA1:RNA (Table 2, Entry 4 and 5). The substitution near N - terminal of the PNA showed better binding than the middle modification, in which ΔT_m - 28.3°C and - 15.6 °C, for L- and D- isomers resp.

We then synthesized PNA oligomers PNA 8 and PNA 9, in which double modification was incorporated towards N-terminal end, to study the effect in binding properties. When subjected for thermal denaturation studies, the UV- T_m profiles indeed improved with both DNA and RNA. In the case of complexation with cDNA, PNA 8, the ΔT_m was -22.8°C and for PNA 9, the ΔT_m was -13.0°C (Table 2, Entry 8 and 9). The result for the duplex stability with RNA showed that the thermal stability of the modified duplexes was lower than the unmodified PNA (ΔT_m -22.8°C and -13.9°C for PNA 8 and PNA 9, resp.)

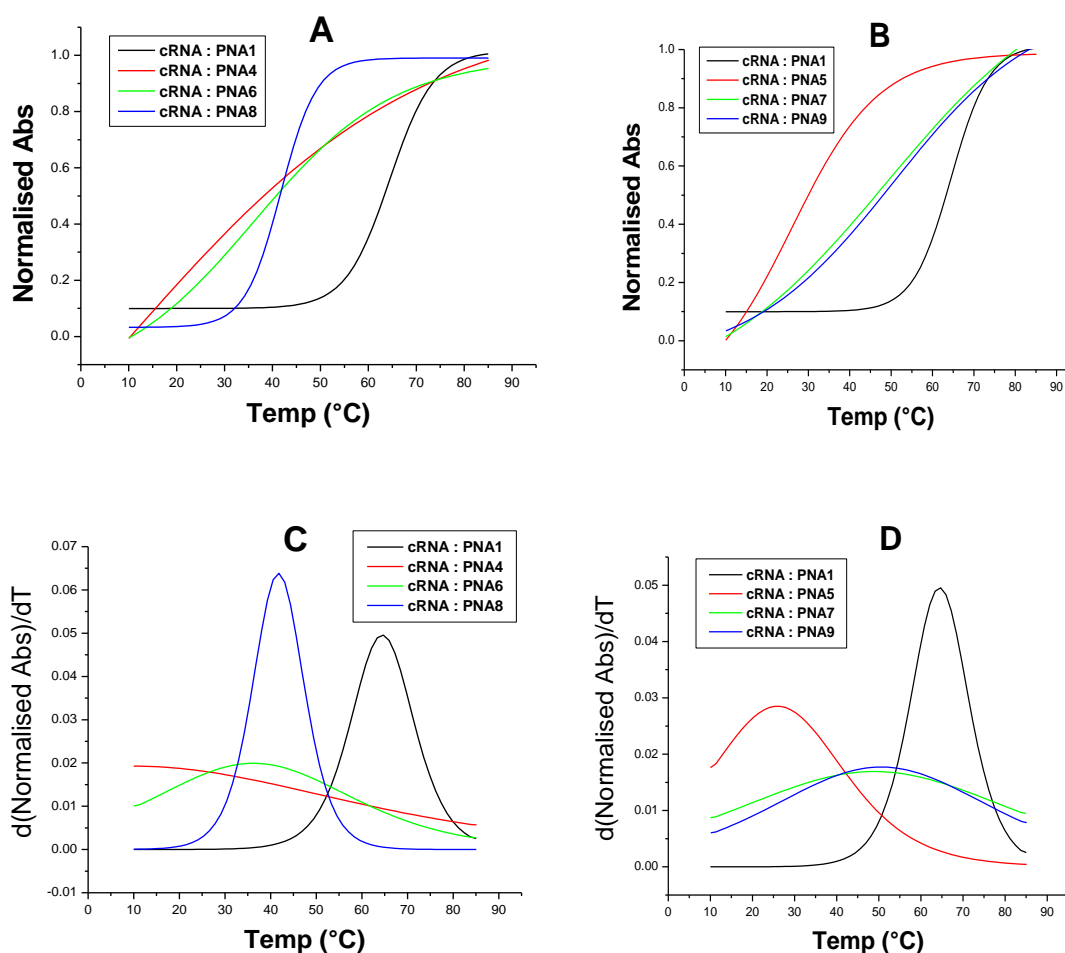


Figure 4. UV-Thermal Melting Profiles. Sigmoidal Curves for (A) cRNA:PNA1/ PNA4/ PNA6/ PNA8 (B) cRNA:PNA1/ PNA5/ PNA7/ PNA9, first derivative curves of the sigmoidal curves for melting values (C) cRNA:PNA1/ PNA4/ PNA6/ PNA8 (D) cRNA:PNA1/ PNA5/ PNA7/ PNA9

3B.3.1 Effect of salt concentration on T_m of modified-PNA:cDNA/RNA duplexes

PNA being charge-neutral, the strength of unmodified PNA:DNA duplexes is known to be little affected by increasing salt concentrations, in contrast to the highly salt concentration dependent duplex strength of native duplexes.⁶² In the present studies, a number of positive charges were introduced in the backbone of PNA and

therefore at higher salt concentration some destabilization of the duplexes could be expected.⁶³ Thermal melting studies were therefore undertaken at higher salt concentration (50 mM and 100 mM NaCl).

Table 3. UV- T_m studies with cDNA and cRNA at 50 mM and 100mM salt concentrations

Entry No.	Code	Sequences (N' → C')	UV - T_m (°C)			
			cDNA		cRNA	
			50 mM	100mM	50 mM	100mM
1.	PNA 1	H - AACCGATTTTCAG -Lys-NH ₂	56.9	55.3	63.1	60.6
2.	PNA 2	H - (K) ₂ - AACCGATTTTCAG -Lys-NH ₂	59.2	57.3	65.0	63.7
3.	PNA 3	H - (K) ₄ - AAC CGATTTTCAG -Lys-NH ₂	59.9	58.5	65.7	64.7
4.	PNA 4	H- AACCGA \underline{t}_L^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	n.d.	n.d.
5.	PNA 5	H- AACCGA \underline{t}_D^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	23.2	19.7
6.	PNA 6	H- A \underline{a}_L^{Am} CCGATTTTCAG -Lys-NH ₂	29.4	26.1	34.7	30.6
7.	PNA 7	H- A \underline{a}_D^{Am} CCGATTTTCAG -Lys-NH ₂	38.5	35.0	45.7	44.3
8.	PNA 8	H- \underline{a}_L^{Am} \underline{a}_L^{Am} CCGATTTTCAG -Lys-NH ₂	34.2	30.8	40.4	36.0
9.	PNA 9	H- \underline{a}_D^{Am} \underline{a}_D^{Am} CCGATTTTCAG -Lys-NH ₂	43.3	42.2	46.7	45.3

\underline{t}_L^{Am} / \underline{a}_L^{Am} and \underline{t}_D^{Am} / \underline{a}_D^{Am} are the modified monomers derived from L- and D- tartaric acid resp.

PNA 1 showed slight destabilization of the duplexes with DNA as well as with RNA at higher salt concentration $T_m = 56.9$ °C at 50 mM and 55.3 °C at 100 mM for DNA and $T_m = 63.1$ °C at 50 mM and 60.6 °C at 100 mM for RNA (Table 3, entry 1). In case of PNA 4, modification of β - and γ - bisubstituted monomers derived from L- isomer does not bind at all even at higher salt concentration. Same is the result with the modification derived from D- isomer (Table 3, entry 4 and 5). But however for the duplex formed with RNA, PNA 5 showed slight binding (Figure 5 and Figure 6).

When the modification is shifted towards the N- terminal, for single modification, the trend is the same as expected with the PNA-DNA duplex at higher concentration of NaCl. Comparing with the 10 mM thermal melting they (PNA 6 and 7) showed destabilisation at 50 mM as well as 100 mM. The PNAs with double modification also showed the similar trend. Thus, we observed that the PNAs having modifications derived from D- isomer of tartaric acid are binding slightly better than the PNAs having the other isomer derived monomers, however as overall, the binding property of all the modified PNAs are very less.

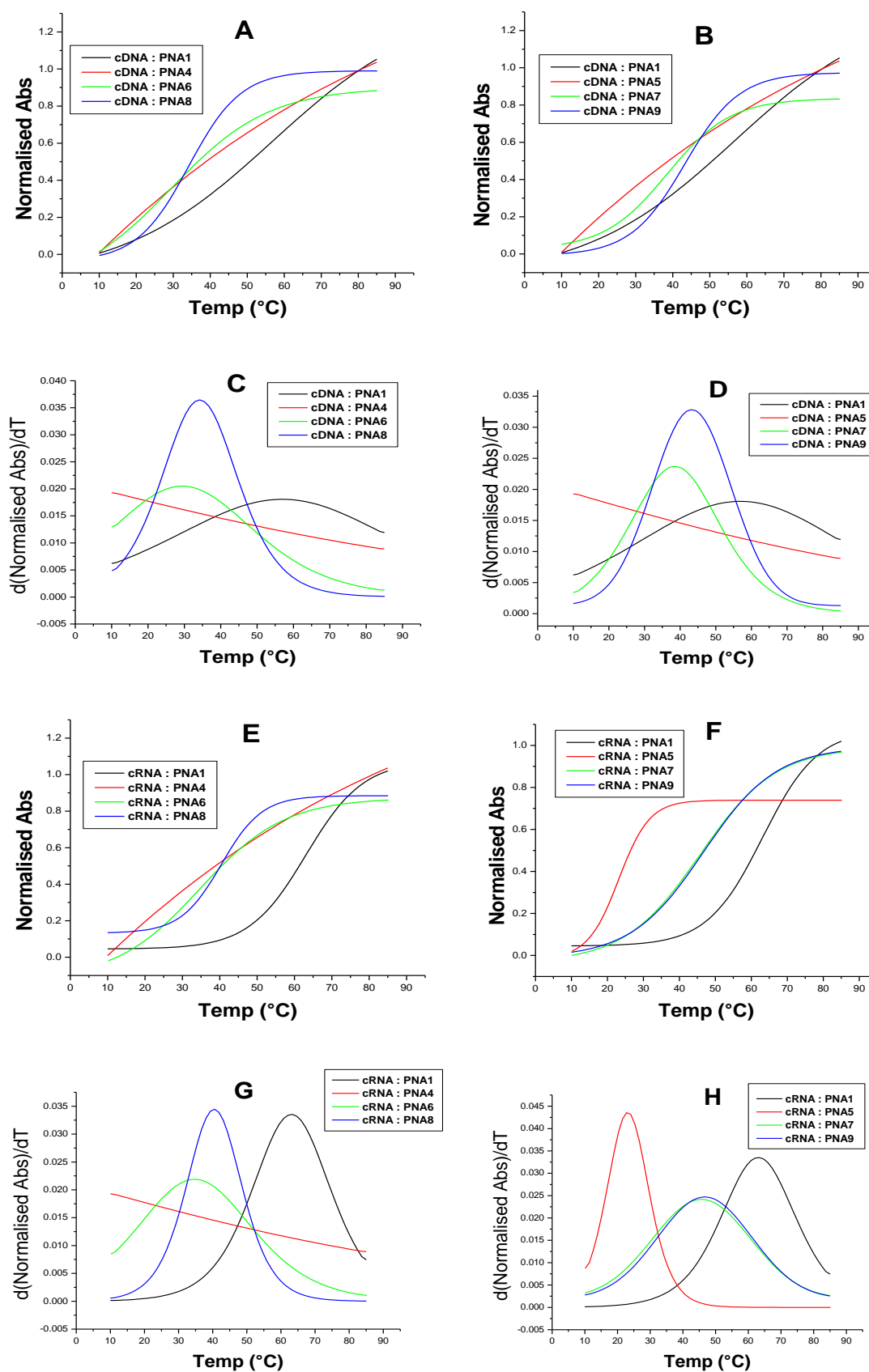


Figure 5. UV-Thermal Melting Profiles at 50 mM NaCl concentration. Sigmoidal Curves for (A) (B) cDNA:PNA1/4/6/5/7/9, (E) (F) cRNA:PNA1/4/6/5/7/9 first derivative curves of the sigmoidal curves for melting values (C) (D) cDNA:PNA1/4/6/5/7/9 (G) (H) cRNA:PNA1/4/6/5/7/9

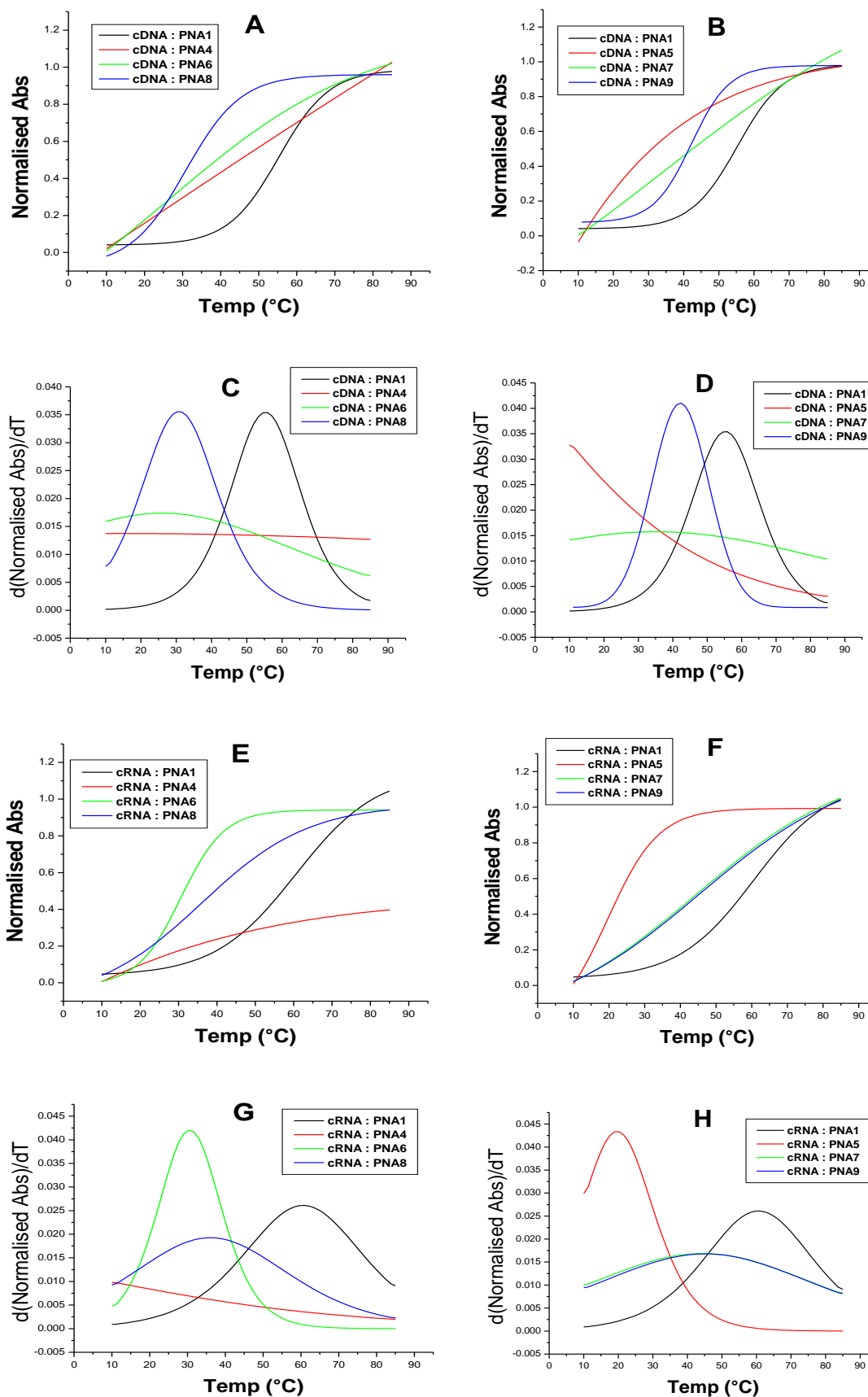


Figure 6. UV-Thermal Melting Profiles at 100 mM NaCl concentration. Sigmoidal Curves for (A) (B) cDNA:PNA1/4/6/5/7/9, (E) (F) cRNA:PNA1/4/6/5/7/9 first derivative curves of the sigmoidal curves for melting values (C) (D) cDNA:PNA1/4/6/5/7/9 (G) (H) cRNA:PNA1/4/6/5/7/9

3B.3.2 UV- T_m studies of modified PNA with a single mismatch RNA

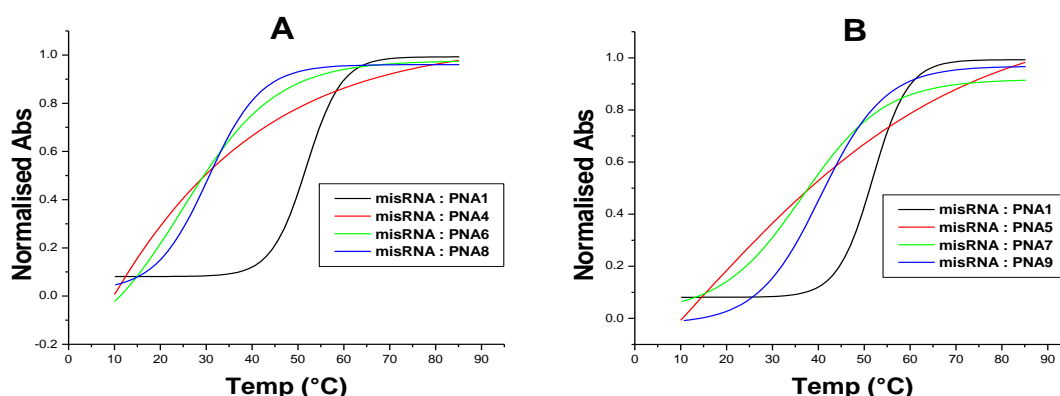
Although the melting values of the modified PNAs were low as compared to any of the unmodified ones (PNAs 1, 2 and 3), we conducted experiments with RNA sequence comprising of a single mismatch, considering the T_m values of the PNA 7 and PNA 9 (48.9 °C and 50.6 °C, Table 2, entry 7 and 9 resp.), which is still higher than rest of the oligomers. The UV- T_m values for the mismatch studies are in Table 4.

Table 4. PNA sequences and UV- T_m studies with mmRNA

Entry No.	Code	Sequences (N' → C')	UV - T_m (°C)		ΔT_m (°C)
			cRNA	mmRNA	
1.	PNA 1	H - AACCGATTTCAG -Lys-NH ₂	64.5	51.9	+ 12.6
2.	PNA 2	H - (K) ₂ - AACCGATTTCAG -Lys-NH ₂	66.6	52.8	+ 13.8
3.	PNA 3	H - (K) ₄ - AAC CGATTTCAG -Lys-NH ₂	68.6	53.2	+ 15.4
4.	PNA 4	H- AACCGA \underline{t}_L^{Am} TTCAG -Lys-NH ₂	n.d.	n.d.	n.d.
5.	PNA 5	H- AACCGA \underline{t}_D^{Am} TTCAG -Lys-NH ₂	25.9	n.d.	+ 25.9
6.	PNA 6	H- A \underline{a}_L^{Am} CCGATTTCAG -Lys-NH ₂	36.2	24.9	+ 11.3
7.	PNA 7	H- A \underline{a}_D^{Am} CCGATTTCAG -Lys-NH ₂	48.9	36.6	+ 12.3
8.	PNA 8	H- \underline{a}_L^{Am} \underline{a}_L^{Am} CCGATTTCAG -Lys-NH ₂	41.7	30.5	+ 11.2
9.	PNA 9	H- \underline{a}_D^{Am} \underline{a}_D^{Am} CCGATTTCAG -Lys-NH ₂	50.6	40.6	+ 10.0

\underline{t}_L^{Am} / \underline{a}_L^{Am} and \underline{t}_D^{Am} / \underline{a}_D^{Am} are the modified monomers derived from L- and D - tartaric acid resp.

It was found that the all the PNA sequences showed much lower melting values and the destabilization with mismatch RNA. All the modified duplexes formed using PNA 4- 9 with mismatched RNA were found to be highly destabilized (Table 4, ΔT_m matched-mismatched ~ 10 - 12.0 °C). PNA 7 has showed a good destabilization of 36.6 °C, same with the PNA 9 40.6 °C, showing $\Delta T_m = + 12.3$ °C and + 10.0 °C resp. From the experiments, we can say that even though the PNAs are not stabilizing the cDNA/RNA, but still showed discrimination in binding even with a single mismatch (Figure 7).



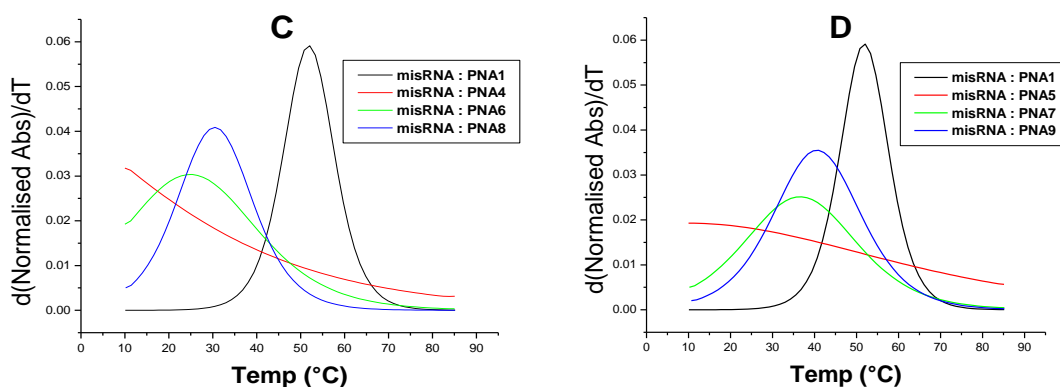


Figure 7. Thermal melting studies with misRNA Sigmoidal Curves for (A) misRNA:PNA1/ PNA4/ PNA6/ PNA8 (B) misRNA:PNA1/ PNA5/ PNA7/ PNA9, first derivative curves of the sigmoidal curves for melting values (C) misRNA:PNA1/ PNA4/ PNA6/ PNA8 (D) misRNA:PNA1/ PNA5/ PNA7/ PNA9

3B.4 Conformational analysis of the modified PNA:cDNA and cRNA

In order to examine if the modified PNAs cause any significant differences on the overall conformational features of the duplexes, the CD studies of PNA complexes with cDNA as well as RNA were undertaken.

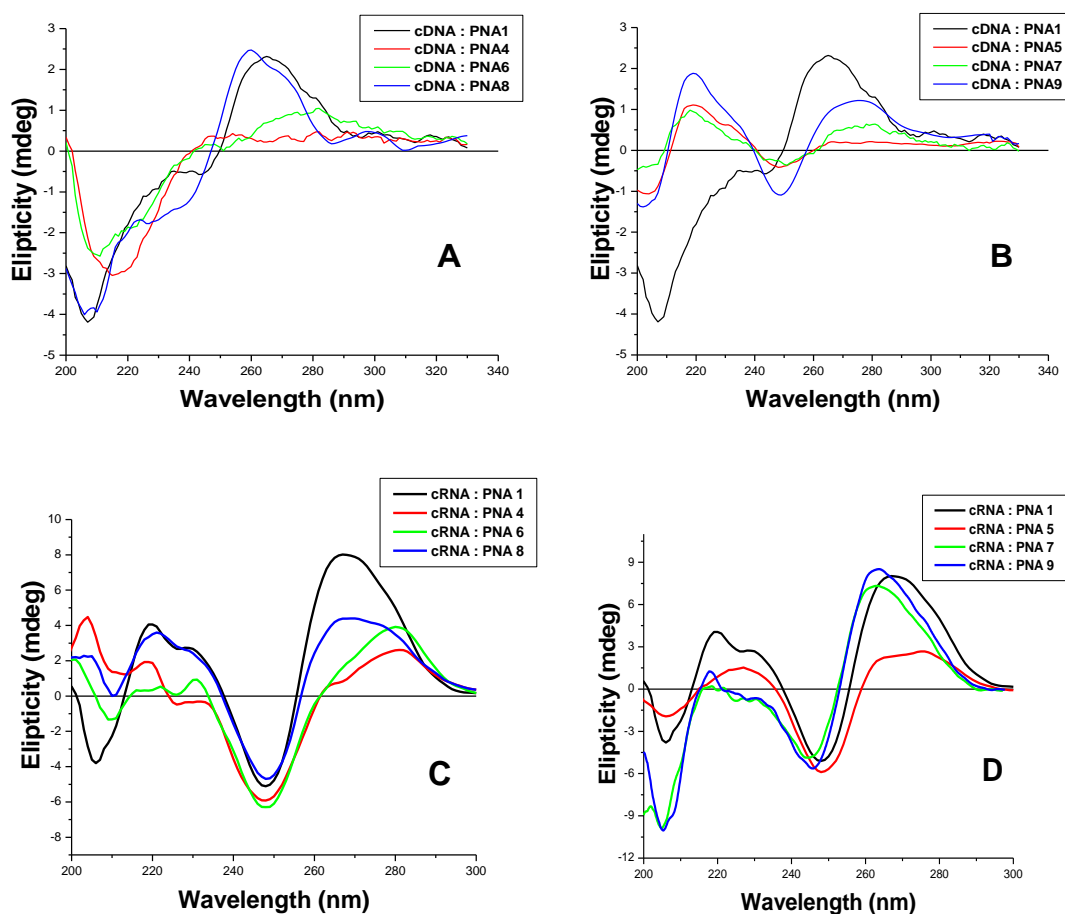


Figure 8. CD Profiles of Duplexes. (A) cDNA:PNA1/ PNA4/ PNA6/ PNA8 (B) cDNA:PNA1/ PNA5/ PNA7/ PNA9 (C) cRNA:PNA1/ PNA4/ PNA6/ PNA8 (D) cRNA:PNA1/ PNA5/ PNA7/ PNA9

A maximum at 260 nm and a minimum at 240 nm typically are seen for right handed helical structures.^{64, 65} Base stacking interactions are reflected in the amplitude of the CD signal at 260 nm in PNA:DNA/RNA duplexes.⁶⁴ The CD profile of the derived PNA oligomers (**PNA 4**), (Figure 8A) with the β - and γ - bisubstituted monomers derived from L- isomer, when duplexed with cDNA did not show any CD pattern, whereas **PNA6** and **PNA 8** showed some kind of pattern as that of the unmodified **PNA1**:DNA duplex. The oligomers derive from D- isomers also showed the same pattern.

The CD studies involving the duplex with RNA showed maximum at 270 nm and a minimum at 250 nm for PNA **4**, **6** and **8**. The modified oligomers PNA **5**, **7** and **9** also showed similar trend. The single strand CD signals for the PNAs have no significant conformation Figure 9.

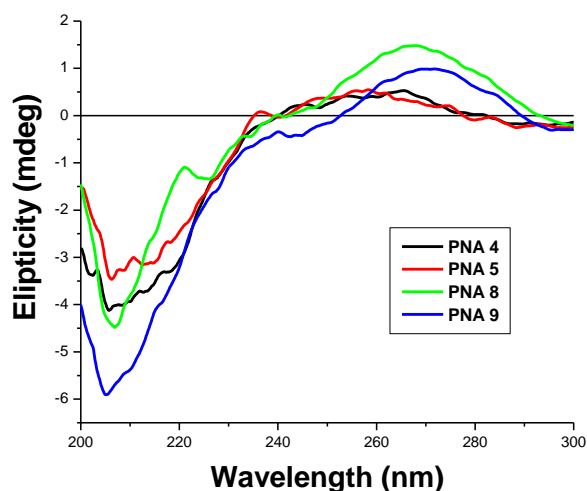


Figure 9. CD Profiles of single strands PNA**4**/ PNA**5**/ PNA**8**/ PNA**9**

3B.5 Summary

- ❖ A new PNA analogue was designed, having substitutions at β - and γ -positions of the PNA backbone.
- ❖ The modified oligomers containing β -, γ - bisubstituted PNA monomers in the middle of the sequence didn't show binding towards cDNA or cRNA.
- ❖ The PNAs showed positional dependence as oligomers having modification towards N – terminal did show some binding, nevertheless lower than the unmodified ones.
- ❖ Change in stereochemistry (from *L* to *D*) showed slight preference in binding with cDNA/RNA.
- ❖ The CD studies of the oligomers with cDNA and cRNA does not show any specific conformation of the modified PNAs.
- ❖ The PNA **9** having $T_m = 50.6^\circ\text{C}$, which is the highest among the modified ones, can be used further for cellular studies.

3B.6 Experimentation

UV-T_m Measurements: The concentrations were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases of DNA/RNA/PNA. The experiments were performed at 1 μ M concentrations of each strand. The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM and 100 mM) and were annealed by keeping the samples at 90°C for 5 min followed by slow cooling to room temperature and then to 0°C. Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Varian Cary 300 spectrophotometer scanning from 10-85°C at a ramp rate of 0.5°C per minute. Experiments were repeated at least thrice and the data were processed using Origin software 6.1. T_m (°C) values were derived from the first derivative curves and are accurate to within $\pm 0.3^\circ\text{C}$.

CD Analysis of the oligomers: The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM) and were annealed by keeping the samples at 90°C for 5 min followed by slow cooling to room temperature. The experiments were performed at 1 μ M concentrations of each strand. All the CD spectra were recorded at room temperature with Jasco J-815 spectropolarimeter. All spectra represent an average of at least 8 scans recorded from 320 nm to 210 nm at a rate of 100 nm per minute in a 1 cm path length cuvette. All spectra were processed using Origin software 6.1, baseline subtracted and smoothed using a 5 point adjacent averaging algorithm.

3B.7 References

1. B. M. Paterson, B. E. Roberts and E. L. Kuff, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 4370-4374.
2. C. A. Stein, *Nat. Biotechnol.*, 1999, **17**, 209-209.
3. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1500.
4. R. Gambari, *Expert Opin. Ther. Patents*, 2014, **24**, 267-294.
5. U. Koppelhus and P. E. Nielsen, *Adv. Drug Deliv. Rev.*, 2003, **55**, 267-280.
6. J. C. Hanvey, N. J. Pepper, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble and L. E. Babiss, *Science*, 1992, **258**, 1481-1485.
7. M. A. Shamma, C. G. Simmons, D. R. Corey and R. J. S. Reis, *Oncogene*, 1999, **18**, 6191-6200.
8. S. E. Hamilton, C. G. Simmons, I. S. Kathiriya and D. R. Corey, *Chem. Biol.*, 1999, **6**, 343-351.
9. L. Moggio, A. Romanelli, R. Gambari, N. Bianchi, M. Borgatti, E. Fabbri, I. Mancini, B. di Blasio, C. Pedone and A. Messere, *Biopolymers*, 2007, **88**, 815-822.
10. G. Cutrona, E. M. Carpaneto, M. Ulivi, S. Roncella, O. Landt, M. Ferrarini and L. C. Boffa, *Nat. Biotechnol.*, 2000, **18**, 300-303.
11. U. Koppelhus, S. K. Awasthi, V. Zachar, H. U. Holst, P. Ebbesen and P. E. Nielsen, *Antisense Nucleic Acid Drug Dev.*, 2002, **12**, 51-63.
12. P. Zhou, A. Dragulescu-Andrasi, B. Bhattacharya, H. O'Keefe, P. Vatta, J. J. Hyldig-Nielsen and D. H. Ly, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4931-4935.
13. S. Sforza, T. Tedeschi, A. Calabretta, R. Corradini, C. Camerin, R. Tonelli, A. Pession and R. Marchelli, *Eur. J. Org. Chem.*, 2010, 2441-2444.
14. A. Dragulescu-Andrasi, S. Rapireddy, G. F. He, B. Bhattacharya, J. J. Hyldig-Nielsen, G. Zon and D. H. Ly, *J. Am. Chem. Soc.*, 2006, **128**, 16104-16112.
15. M. M. Fabani and M. J. Gait, *RNA-Publ. RNA Soc.*, 2008, **14**, 336-346.
16. M. M. Fabani, C. Abreu-Goodger, D. Williams, P. A. Lyons, A. G. Torres, K. G. C. Smith, A. J. Enright, M. J. Gait and E. Vigorito, *Nucleic Acids Res.*, 2010, **38**, 4466-4475.
17. S. Y. Oh, Y. Ju and H. Park, *Mol. Cells*, 2009, **28**, 341-345.

18. M. Zasloff, *Nature*, 2002, **415**, 389-395.
19. K. L. Dueholm, K. H. Petersen, D. K. Jensen, M. Egholm, P. E. Nielsen and O. Buchardt, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 1077-1080.
20. P. Gupta, O. Muse and E. Rozners, *Biochemistry*, 2012, **51**, 63-73.
21. A. Puschl, S. Sforza, G. Haaima, O. Dahl and P. E. Nielsen, *Tetrahedron Lett.*, 1998, **39**, 4707-4710.
22. G. Haaima, A. Lohse, O. Buchardt and P. E. Nielsen, *Angew. Chem.-Int. Edit.*, 1996, **35**, 1939-1942.
23. B. S. Balaji, F. Gallazzi, F. Jia and M. R. Lewis, *Bioconjugate Chem.*, 2006, **17**, 551-558.
24. S. Sforza, G. Haaima, R. Marchelli and P. E. Nielsen, *Eur. J. Org. Chem.*, 1999, 197-204.
25. S. Sforza, R. Corradini, S. Ghirardi, A. Dossena and R. Marchelli, *Eur. J. Org. Chem.*, 2000, 2905-2913.
26. C. Dose and O. Seitz, *Bioorg. Med. Chem.*, 2008, **16**, 65-77.
27. R. Hamzavi, F. Dolle, B. Tavitian, O. Dahl and P. E. Nielsen, *Bioconjugate Chem.*, 2003, **14**, 941-954.
28. R. Hamzavi, C. Meyer and N. Metzler-Nolte, *Org. Biomol. Chem.*, 2006, **4**, 3648-3651.
29. S. Dorn, N. Aghaallaei, G. Jung, B. Bajoghli, B. Werner, H. Bock, T. Lindhorst and T. Czerny, *BMC Biotechnol.*, 2012, **12**, 9.
30. I. Dilek, M. Madrid, R. Singh, C. P. Urrea and B. A. Armitage, *J. Am. Chem. Soc.*, 2005, **127**, 3339-3345.
31. A. Gourishankar and K. N. Ganesh, *Artificial DNA: PNA & XNA*, 2012, **3**, 5-13.
32. S. Rossi, A. Calabretta, T. Tedeschi, S. Sforza, S. Arcioni, L. Baldoni, R. Corradini and R. Marchelli, *Artificial DNA: PNA & XNA*, 2012, **3**, 63-72.
33. P. Zhou, M. M. Wang, L. Du, G. W. Fisher, A. Waggoner and D. H. Ly, *J. Am. Chem. Soc.*, 2003, **125**, 6878-6879.
34. H. Nagahara, A. M. Vocero-Akbani, E. L. Snyder, A. Ho, D. G. Latham, N. A. Lissy, M. Becker-Hapak, S. A. Ezhevsky and S. F. Dowdy, *Nat. Med.*, 1998, **4**, 1449-1452.
35. S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda and Y. Sugiura, *J. Biol. Chem.*, 2001, **276**, 5836-5840.

36. P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman and J. B. Rothbard, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 13003-13008.
37. A. Dragulescu-Andrasi, P. Zhou, G. F. He and D. H. Ly, *Chem. Commun.*, 2005, 244-246.
38. L. Kosynkina, W. Wang and T. C. Liang, *Tetrahedron Lett.*, 1994, **35**, 5173-5176.
39. C. Dose and O. Seitz, *Org. Lett.*, 2005, **7**, 4365-4368.
40. E. A. Englund and D. H. Appella, *Org. Lett.*, 2005, **7**, 3465-3467.
41. S. Ficht, C. Dose and O. Seitz, *Chembiochem*, 2005, **6**, 2098-2103.
42. T. Tedeschi, S. Sforza, R. Corradini and R. Marchelli, *Tetrahedron Lett.*, 2005, **46**, 8395-8399.
43. E. A. Englund and D. H. Appella, *Angewandte Chemie International Edition*, 2007, **46**, 1414-1418.
44. A. Dragulescu-Andrasi, S. Rapireddy, B. M. Frezza, C. Gayathri, R. R. Gil and D. H. Ly, *J. Am. Chem. Soc.*, 2006, **128**, 10258-10267.
45. M. J. Crawford, S. Rapireddy, R. Bahal, I. Sacui and D. H. Ly, *Effect of Steric Constraint at the γ -Backbone Position on the Conformations and Hybridization Properties of PNAs*, 2011.
46. B. Sahu, V. Chenna, K. L. Lathrop, S. M. Thomas, G. Zon, K. J. Livak and D. H. Ly, *J. Org. Chem.*, 2009, **74**, 1509-1516.
47. A. Manicardi, E. Fabbri, T. Tedeschi, S. Sforza, N. Bianchi, E. Brognara, R. Gambari, R. Marchelli and R. Corradini, *Chembiochem*, 2012, **13**, 1327-1337.
48. A. Manicardi, A. Calabretta, M. Bencivenni, T. Tedeschi, S. Sforza, R. Corradini and R. Marchelli, *Chirality*, 2010, **22**, E161-E172.
49. R. Mitra and K. N. Ganesh, *J. Org. Chem.*, 2012, **77**, 5696-5704.
50. R. Mitra and K. N. Ganesh, *Chem. Commun.*, 2011, **47**, 1198-1200.
51. M. C. Myers, M. A. Witschi, N. V. Larionova, J. M. Franck, R. D. Haynes, T. Hara, A. Grajkowski and D. H. Appella, *Org. Lett.*, 2003, **5**, 2695-2698.
52. J. K. Pokorski, M. A. Witschi, B. L. Purnell and D. H. Appella, *J. Am. Chem. Soc.*, 2004, **126**, 15067-15073.
53. J. K. Pokorski, M. C. Myers and D. H. Appella, *Tetrahedron Lett.*, 2005, **46**, 915-917.
54. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *Chem. Commun.*, 2004, 860-861.

-
55. T. Govindaraju, V. A. Kumar and K. N. Ganesh, *J. Am. Chem. Soc.*, 2005, **127**, 4144-4145.
 56. P. Lagriffoule, P. Wittung, M. Eriksson, K. K. Jensen, B. Norden, O. Buchardt and P. E. Nielsen, *Chem.-Eur. J.*, 1997, **3**, 912-919.
 57. S. Bregant, F. Burlina, J. Vaissermann and G. Chassaing, *Eur. J. Org. Chem.*, 2001, 3285-3294.
 58. S. Bregant, F. Burlina and G. R. Chassaing, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 1047-1050.
 59. T. Sugiyama, Y. Imamura, Y. Demizu, M. Kurihara, M. Takano and A. Kittaka, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 7317-7320.
 60. E. A. Mash, K. A. Nelson, S. Vandusen and S. B. Hemperly, *Org. Synth*, 1990, **68**, 92-103.
 61. J. D. Puglisi and I. Tinoco, *Method Enzymol.*, 1989, **180**, 304-325.
 62. S. Tomac, M. Sarkar, T. Ratilainen, P. Wittung, P. E. Nielsen, B. Norden and A. Graslund, *J. Am. Chem. Soc.*, 1996, **118**, 5544-5552.
 63. N. Tilani, S. De Costa and J. M. Heemstra, *PLoS One*, 2013, **8**, 8.
 64. B. Sahu, I. Sacui, S. Rapireddy, K. J. Zanotti, R. Bahal, B. A. Armitage and D. H. Ly, *J. Org. Chem.*, 2011, **76**, 5614-5627.
 65. W.C. Johnson, K. Nakanishi, N. Berova and R. W. Woody, *Circular Dichroism Principles and Applications*, VCH., New York, 2000, 523-540.

CHAPTER 4

polyamide-DNA

Synthesis and biophysical studies of polyamide-DNA with
alternating α -amino acid and nucleoside- β -amino acids

4 Introduction

The discovery of oligonucleotides acting as antisense agents by Zamecnik and Stephenson in 1978 has opened a vast new area of interest in the nucleic acid chemistry.¹ Since then over the decades several synthetic oligonucleotides were developed which led to first generation (Phosphorothioate), second generation (2'-*O*-alkyl) and even third generation (LNA, PNA, PMO etc.) Antisense oligomers to address the RNAs based on different antisense strategies.²

Natural oligonucleotides (ON) with phosphodiester linkage are susceptible for hydrolysis with various enzymes like nucleases and this limits their use for various biological applications. The replacement of the internucleoside sugar-phosphate linkages by the robust amide,³ carbamate,⁴ carbazoyl⁵ or guanidino-⁶ linkages is studied in the literature. The most promising modifications of the sugar-phosphodiester linkage replacement were in achiral and acyclic peptide nucleic acid (PNA) with neutral amide linkages.⁷ The removal of sugars in the backbone of PNA, however lost the sense of 3'→5' directionality which is a very important aspect of natural DNA. The four-atom amide substitution of four-atom phosphodiester linkage led to chimeric ONs with alternate sugar-amide-sugar-phosphodiester groups. Such chimeric ONs bound to RNA with a moderate RNA binding selectivity and with moderately diminished affinity in antiparallel orientation.^{8, 9} Partial replacement of phosphodiester linkages with five-atom amide linkers were found to be useful for RNA recognition.¹⁰

4.1 Rationale and objective of the present work

Pallan *et al.* carried out a systematic study for RNA binding affinity of five-atom amide linked DNA analogues.¹¹ Over the last decade, we and others have been working towards the synthesis of four- and five-atom amide linked nucleic acid analogues for better RNA binding. The five-atom amide-linked oligonucleotides containing thymidine and thymidine-cytosine dimers were reported by our group where thioacetamido backbone (TANA/*iso*TANA) exhibited discrimination in binding affinity towards DNA/RNA.^{12, 13} We also reported (thyminylyl- β -amino acid + α -L-amino acid)₈ sequences where the 5-atom amide linkers completely replaced the anionic phosphodiester linkages, specifically recognized DNA and RNA with high binding affinity.¹⁴ Our initial work was based on the prolyl-ACPC backbone^{15, 16} (Figure 1) and in that case the alternating nucleoside- β -amino acids with natural β -amino acids gave rise to similar alternating α/β amino acid backbone scaffolds using well-established peptide chemistry.¹⁴ For our interests in synthesizing ($\alpha + \beta$) amino acid backbone, we

further studied the effect of chirality of the α -amino acids by synthesizing the thymine dimers linked with L-proline, D-proline and prochiral glycine units on base stacking interactions.¹⁷ Further, we designed a promising strategy which used the combination of α -amino acids alternating with nucleoside based β -amino acids to construct amide-linked nucleoside oligomers. The synthesis of primary β -amino acid orthogonally protected building blocks corresponding to all four nucleosides and a solid phase synthesis of model tetrameric DNA sequence in 5'→3' direction was described, where we employed trityl protected nucleoside- β -amino acid and glycine as α -amino acid.¹⁸ We describe herein, the synthesis of mixed base purine/pyrimidine oligomers in which the phosphodiester linkages are completely replaced by the 5-atom amide linkages alternately using nucleoside- β -amino acids and glycine and their binding preferences to cDNA/RNA sequences (Figure 1).

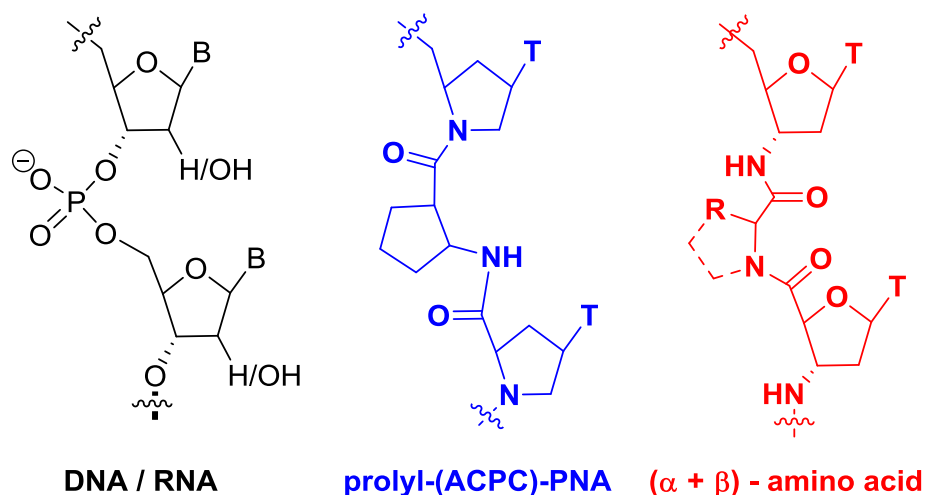


Figure 1. DNA/RNA and backbone alterations. Rationale for nucleoside- β -amino acids

4.2 Synthesis of nucleoside- β -amino acid monomers¹⁸

In our earlier strategy, we used Fmoc protected 3'-deoxy-3'-aminothymidine as nucleoside- β -amino acid monomer and rink amide resin as the solid support.¹⁴ Exocyclic amino protection is not required for thymine nucleobase during oligomer synthesis, and Fmoc strategy worked well. Also cleavage from the rink amide resin requires strongly acidic conditions (TFA:TFMSA / 20 % TFA in DCM),¹⁹ which may not be compatible for mixed base sequences containing purine nucleosides due to likely depurination in strongly acidic medium. We decided to use the trityl protection for 3'-sugar-amino function to be deprotected by trichloroacetic acid at each coupling step.

This altered strategy would accommodate the orthogonal protection of nucleobases (benzoyl for adenine and cytosine and *isobutyryl* for guanine) and the synthesis would now be from 5'→3' direction instead of 3'→5' direction used in DNA

synthesis. We choose the succinate ester at 5'-end as a linker for attachment of the first 3'-amino-nucleoside via succinamide linker to the solid support. The succinate ester can be cleaved concomitantly with the deprotection of exocyclic amino groups, using aqueous/methanolic ammonia at the end of the synthesis, leaving free 5'-OH group.

This trityl strategy has been previously applied successfully for the synthesis of oligonucleotide phosphoramidates on Controlled Pore Glass (CPG) support.^{20, 21} The MBHA-succinate linker strategy has been previously used for the synthesis of glycopeptides²² but has not been employed for the synthesis of polyamide-oligonucleosides. The TEMPO-BAIB method^{23, 24} used earlier by us^{14, 17} and others²⁵⁻²⁸ for oxidation of 5'-primary hydroxyl group in 3'-deoxy-3'-azidothymidine to get the 4'-carboxylic acid would be used as a general procedure for protected trityl aminonucleosides. Thus, the syntheses of all four protected natural nucleoside based β -amino acids were accomplished using simple TEMPO/BAIB method (Figure 2). The complete synthesis was reported earlier as part of Seema Bagmare's PhD Thesis Dec'13, NCL, Pune.¹⁸

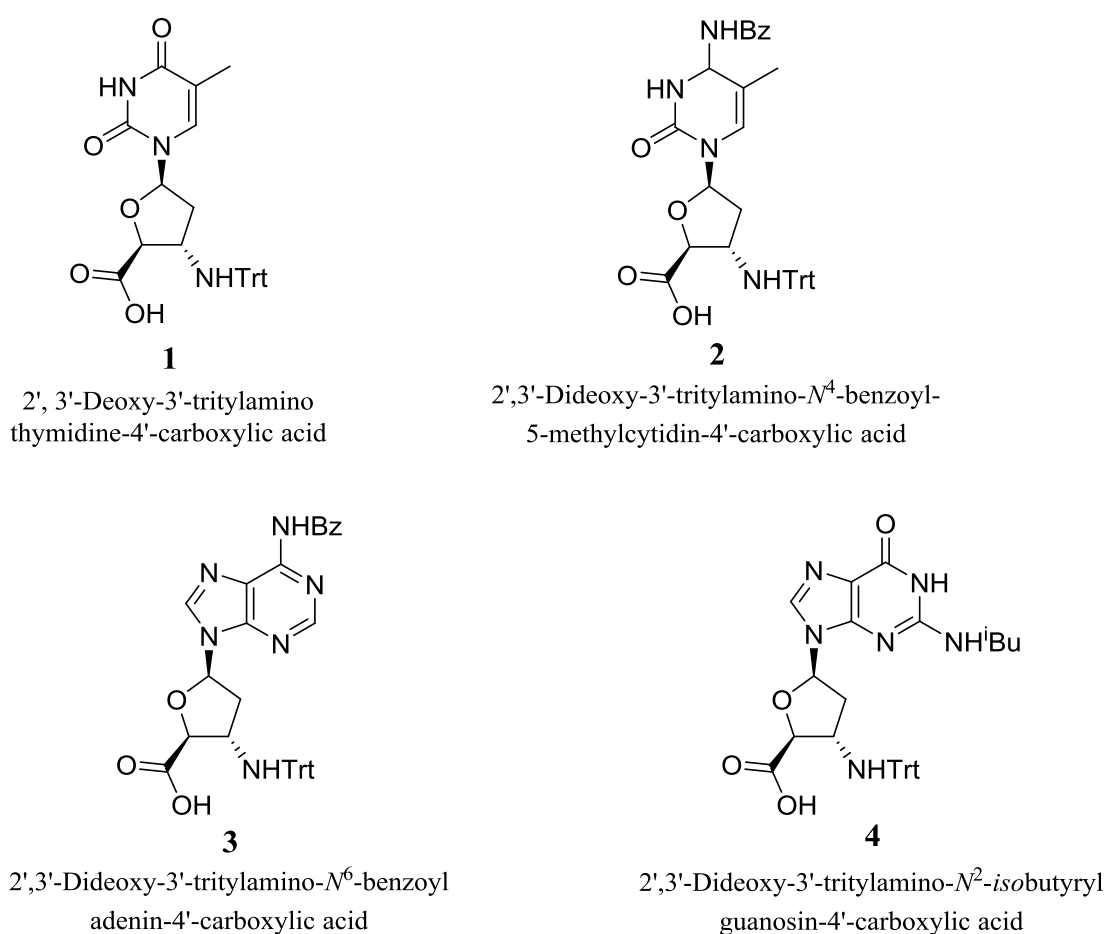


Figure 2. All four protected natural nucleoside based β -amino acids

4.3 Conformations in nucleoside- β -amino acids:

In natural nucleic acids, the pentose sugars are puckered or twisted to give preferred helical conformations. These pentose sugar moieties are puckered in order to minimize non bonded interactions between their substituents. This puckering is described by identifying the major displacement of carbon C-2' and C-3' from the median plane of C1'-O4'-C-4'. The detailed study of X-ray crystal structures of nucleosides and nucleotides revealed that ribo and deoxyribo furanosyl moieties occur preferentially within two distinct major conformations namely north (C-3' *endo*) and south (C-2' *endo*).²⁹ In solution, N-type and S-type conformations are in rapid equilibrium and are separated by a low energy barrier.

A semi-empirical method termed as "Sum Rule"³⁰ can also be employed to have some idea about N \leftrightarrow S equilibrium in the case of sugar conformations. This method was employed in the present work to calculate the conformational preference of nucleoside- β -amino acids. The empirical equation to calculate the percentage of S conformer is as below Table 1:

$$\% S = (\Sigma H1' - 9.8) / 5.9 \times 100 \text{ where } \Sigma H1' = J_{1',2'} + J_{1',2''}$$

Table 1. S-type preference for the synthesized monomers

Comp. No	Nucleoside- β -amino acid	$\Sigma H1'$	% N*
1	2', 3'-Deoxy-3'-tritylaminothymidine-4'-carboxylic acid	10.54	85.5
2	2',3'-Dideoxy-3'-tritylamino- N^4 -benzoyl-5-methylcytidin-4'-carboxylic acid	9.54	~100
3	2',3'-Dideoxy-3'-tritylamino- N^6 -benzoyladenine-4'-carboxylic acid	10.79	83.2
4	2',3'-Dideoxy-3'-tritylamino- N^2 -isobutyrylguanosin-4'-carboxylic acid	9.28	~100

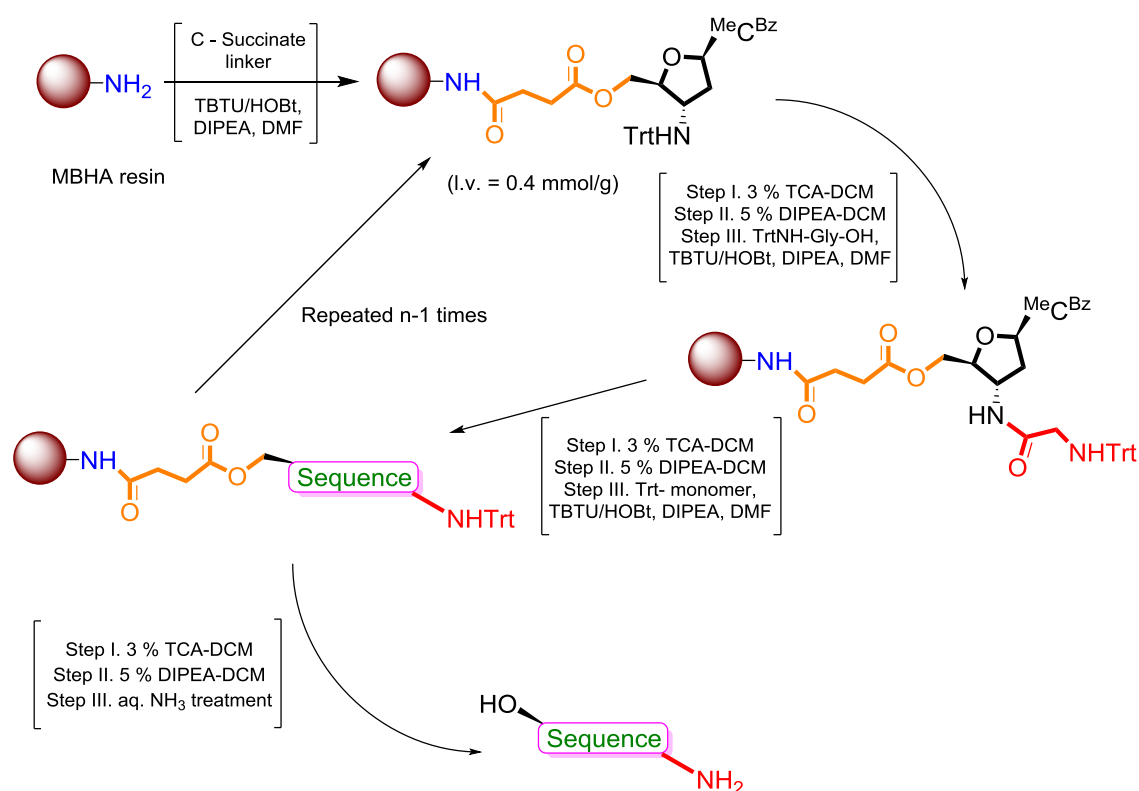
* % N = 100 - % S

It was found that 2', 3'-deoxy-3'aminoribonucleosides, all four nucleoside- β -amino acids prefer to be predominantly in N type sugar conformations.

4.4 Synthesis of polyamide DNA oligomers

Choice of the resin for solid phase synthesis was influenced by the swelling properties of the resin and avoiding harsh acidic deprotection conditions considering sensitivity of the glycosidic bond towards acid. Amino functionalized resin has to be selected such that the final cleavage conditions are compatible towards stability of glycosidic bond. The MBHA resin, which has better swelling properties, was chosen for

carrying out the oligomer synthesis from 5'→3' direction. By using 3 % TCA in DCM for deprotection of trityl group and TBTU/HOBt activation strategy for coupling of monomers, the sequences were synthesized (Scheme 3). The sequences were cleaved from resin via aqueous NH₃ treatment and purified by HPLC on a PepRP column and characterized by MALDI-TOF mass spectrometry. The synthesised sequences are listed in Table 2. For a comparative study of these synthesised polyamide DNA, there corresponding PNA and DNA sequences were also synthesised using the procedure described in previous sections.



Scheme 3. Solid Phase Synthesis of polyamide DNA oligomers

Table 2. Polyamine sequences and PNA, HPLC retention time and MALDI-TOF mass

Entry No.	Code	Sequences	HPLC t_R (min)	MALDI-ToF Mass (Cald./Obsd.)
1.	TRT 1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T -3'	9.58	2884.04/2906.69 (M+Na ⁺)
2.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g T _g C _g A _g A-3'	10.60	3536.28/3558.60 (M+Na ⁺)

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

4.5 Biophysical Studies of oligomers with DNA / RNA

The synthesized polyamide DNA sequences were examined for their binding affinity with cDNA/RNA sequences by employing UV-thermal denaturation studies. The sequence chosen for study was a decamer sequence of CT bases and a mixed sequence of 12mer length (antimiR 29a). The modified oligomers were annealed with the cDNA and RNA and were subjected to temperature dependent UV studies at 260 nm. Unmodified **DNA** and **PNA** were used as a control in these experiments. The UV- T_m plots show a single sigmoidal transition, characteristic of PNA:cDNA/RNA duplex melting. The T_m values were obtained by the first derivative from such.

4.5.1 UV- T_m Studies of oligomers with antiparallel complementary DNA / RNA

It was seen from the UV-melting data that the polyamide-DNA with alternating α -amino acid and nucleoside- β -amino acids was showing stability while comparing with the DNA as well as with RNA (Table 3).

Table 3. Polyamide DNA sequences and UV- T_m studies with *ap* cDNA and *ap* RNA

Entry No.	Code	Sequences	UV - T_m (°C)		ΔT_m (°C) modified-ctrl	
			<i>ap</i> <i>cDNA</i> 1	<i>ap</i> <i>RNA</i> 1	<i>DNA</i>	<i>RNA</i>
1.	DNA1	5'- CTT CTT CCT T - 3'	27.8	36.1	-	-
2.	PNA1	N'- cttcttctct t -K -C'	47.3	51.3	+ 19.5	+ 15.2
3.	TRT 1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T - 3'	39.7	48.3	+ 11.9	+ 12.2
			<i>ap</i> <i>cDNA</i> 2	<i>ap</i> <i>RNA</i> 2	<i>DNA</i>	<i>RNA</i>
4.	DNA2	5'- CAC TGA TTT CAA - 3'	35.9	44.7	-	-
5.	PNA2	N'- cactgatttcaa- K -C'	53.8	60.6	+ 17.9	+ 15.9
6.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g T _g C _g A _g A-3'	45.3	54.0	+ 9.4	+ 9.3

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

The lower case bases represent *aeg*PNA monomer units

ap cDNA1: 5'-AAG GAA GAA G : antiparallel complementary to TRT1; *ap* cDNA2: 5'-TTG AAA

TCA GTG : antiparallel complementary to TRT2; *ap* RNA1: 5'-AAG GAA GAA G : antiparallel

complementary to TRT1; *ap* RNA2: 5'- UUG AAA UCA GUG : antiparallel complementary to TRT2

The CT sequence of the polyamide DNA **TRT 1** (Entry 3) sequence was found to form complex with **cDNA 1** showed better stabilization while comparing with antiparallel (*ap*) cDNA by $\Delta T_m = + 11.9$ °C. Similarly for **TRT 2:cDNA2** mixed

sequence duplex (Entry 6), the increase in melting values was 9.4°C. For further comparison, PNAs of the corresponding sequences were synthesized and melting studies were also carried out. The thermal melting values of the polyamide DNA oligomers were less compared to their corresponding PNA1:cDNA1 and PNA 2:cDNA 2 complexes ($\Delta T_m = - 7.6^\circ\text{C}$ and $\Delta T_m = - 8.5^\circ\text{C}$ for **TRT 1** and **TRT 2** resp.)

When the modified oligomers were subjected to melting studies with the antiparallel (*ap*) cRNA, higher stabilization was observed as compared to DNA:cRNA, but lesser stability than the PNA:cRNA (Table 3). The T_m values of 48.3°C (TRT 1:RNA 1) and 54.0 °C (TRT 2:RNA 2) were obtained for the polyamide DNA:cRNA and 51.3 °C (PNA 1:RNA 1) and 60.6 °C with PNA 2: RNA 2. While comparing the binding affinity of the modified sequences with DNA and RNA, they showed preferential binding towards RNA by $\sim 8^\circ\text{C}$. The melting curves are shown in Figure 2.

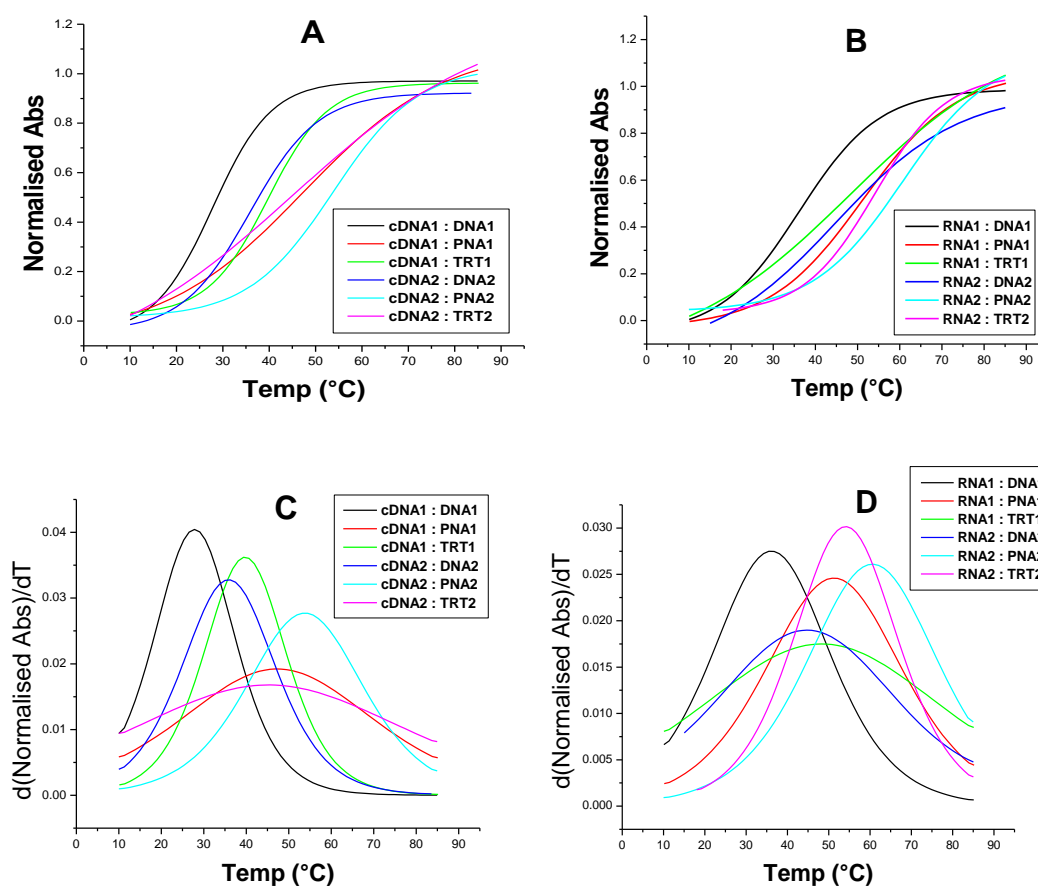


Figure 2.UV-Thermal Melting Profiles. Sigmoidal Curves for (A)*ap* cDNA1:DNA1/ PNA1/ TRT1, *ap* cDNA2/DNA2/ PNA2/ TRT2(B) *ap* RNA1:DNA1/ PNA1/ TRT1,*ap* RNA2:DNA2/ PNA2/ TRT2, first derivative curves of the sigmoidal curves for melting values (C) *ap* cDNA1:DNA1/ PNA1/ TRT1,*ap* cDNA2:DNA2/ PNA2/ TRT2(D)*ap* RNA1:DNA1/ PNA1/ TRT1, *ap* RNA2:DNA2/ PNA2/ TRT2

4.5.2 UV- T_m Studies of oligomers with parallel complementary DNA / RNA

In the polyamide DNA, the phosphate backbone has been replaced by amide bonds. The 5' end of the polyamide DNA monomers are having an acid moiety whereas the 3' end is having an amine residue. The polyamide-sugar backbone comprising these monomers in oligoamide-DNA may deviate from the native phosphodiester-sugar internucleoside linkages, but the conventional 5' → 3' directionality present in DNA/RNA may still be conserved as opposed to that in achiral peptide nucleic acids. PNAs have strong preference to bind DNA in antiparallel orientation but PNAs do bind to cDNA in parallel orientation also, which is sequence dependent. Native DNA does not bind to parallel cDNA at all. For studying the effect of polyamide DNA in directional selective binding to the cDNA strand, melting studies were undertaken using parallel cDNA as well as with parallel complementary RNA.

Thus, UV- T_m Studies of oligomers with parallel (p) complementary DNA / RNA were conducted to see the effect of this backbone modification (Table 4 and Figure 3)

Table 4. Polyamide DNA sequences and UV- T_m studies with p cDNA and p RNA

Entry No.	Code	Sequences	UV- T_m (°C)		ΔT_m (°C) antiparallel-parallel	
			p cDNA 3	p RNA 3	DNA	RNA
1.	DNA1	5'- CTT CTT CCT T - 3'	n.d.	n.d.	+ 27.8	+ 36.1
2.	PNA1	N'- cttcttctct t -K -C'	32.9	38.5	+ 14.4	+ 12.8
3.	TRT1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T - 3'	n.d.	n.d.	+ 39.7	+ 48.3
			p cDNA 4	p RNA 4	DNA	RNA
4.	DNA2	5'- CAC TGA TTT CAA - 3'	n.d.	n.d.	+ 35.9	+ 44.7
5.	PNA2	N'- cactgatttcaa- K -C'	40.4	48.6	+ 13.4	+ 12.0
6.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g T _g C _g A _g A-3'	n.d.	n.d.	+ 45.3	+ 54.0

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

The lower case bases represent *aeg*PNA monomer units

p cDNA3: 5'-GAA GAA GGA A: parallel complementary to TRT1; p cDNA4: 5'-GTG ACT AAA GTT: parallel complementary to TRT2; p RNA3: 5'-GAA GAA GGA A: parallel complementary to TRT1; p RNA4: 5'-GUG ACU AAA GUU: parallel complementary to TRT2

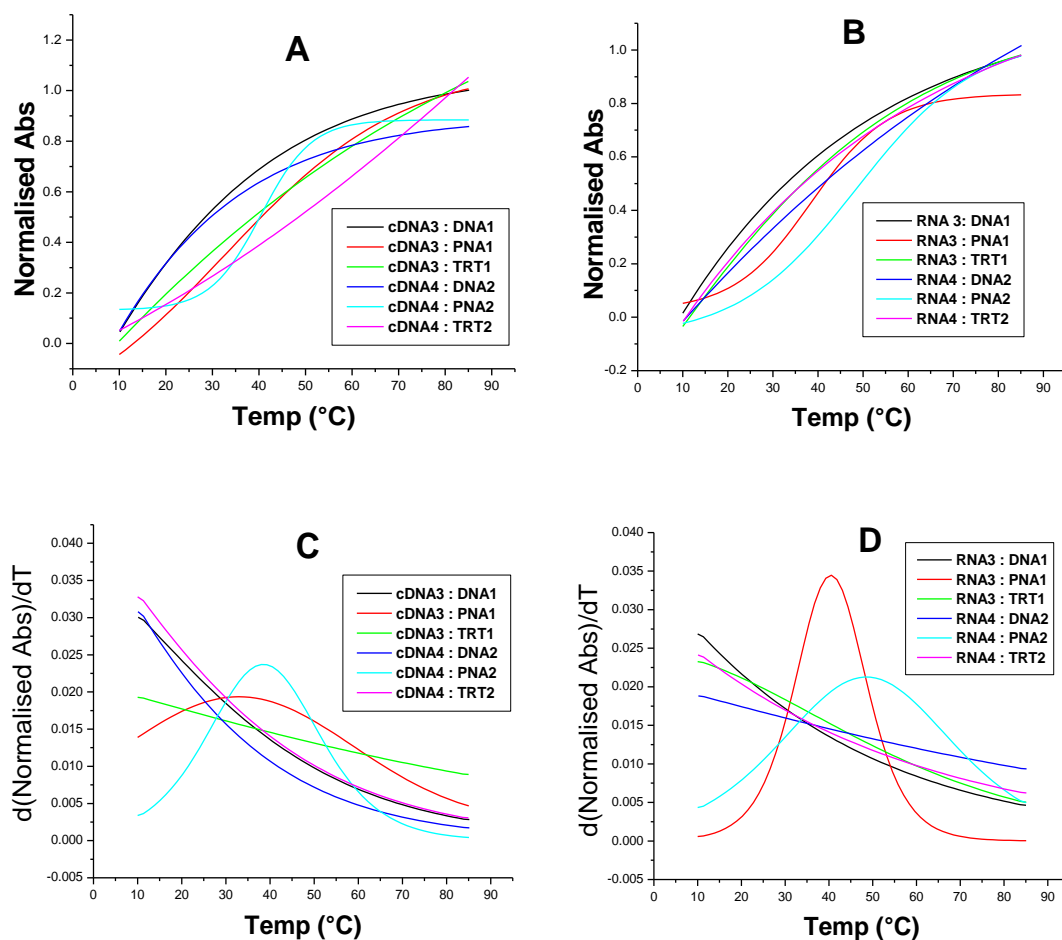


Figure 3. UV-Thermal Melting Profiles. Sigmoidal Curves for (A) *ap* cDNA3:DNA1/ PNA1/ TRT1, *ap* cDNA4/DNA2/ PNA2/ TRT2 (B) *ap* RNA3:DNA1/ PNA1/ TRT1, *ap* RNA4:DNA2/ PNA2/ TRT2, first derivative curves of the sigmoidal curves for melting values (C) *ap* cDNA3:DNA1/ PNA1/ TRT1, *ap* cDNA4:DNA2/ PNA2/ TRT2 (D) *ap* RNA3:DNA1/ PNA1/ TRT1, *ap* RNA4:DNA2/ PNA2/ TRT2

From the results presented above, it is found that indeed, the polyamide DNA backbone exhibited directional specificity of binding with antiparallel complementary DNA/RNA, as none of the two modified oligomers bind to *p* cDNA. For **TRT 1** the binding with *ap* cDNA was 39.7 °C and for **TRT 2** is 45.3 °C, whereas for *p* cDNA it did not show any melting curve in both the cases (entry 3 and 6). However the corresponding PNAs for the sequences show binding towards the *p* cDNA as well as towards *p* RNA, albeit with lower affinity (Table 4, entry 2 and 5)

Similar trend was observed for parallel RNA, where the modified oligomers bind preferentially with the *ap* RNA (Table 3, entry 3 and 6) and not at all bind to *p* RNA (Table 4, entry 3 and 6). Thus, the thermal melting study results indicated that the modification in the backbone of polyamide DNA with sugar units did show directional preferability.

4.5.3 UV- T_m Studies of oligomers with single mismatch DNA / RNA

As it has been observed in the previous section that the polyamide DNA sequences were showing preferential binding towards *ap* cDNA/RNA, melting studies were further conducted to verify the affinity of the sequences to form duplex if a single mismatch is present in the complementary strand.

The detailed T_m values and graphs are shown in Table 5 and Figure 4. From the UV- T_m Studies it was evicted that the **TRT 1** and **TRT 2** sequences had destabilized the duplex by ΔT_m (mismatched-matched) - 8.9 °C and - 7.8 °C for mismatch *ap* cDNA and by ΔT_m (mismatched-matched) - 10.1 °C and - 10.8 °C for mismatch *ap* RNA respectively (from Table 2 and 5).

Table5. Polyamide DNA sequences and UV- T_m studies with *ap* mmDNA and *ap* mmRNA

Entry No.	Code	Sequences	UV - T_m (°C)		ΔT_m (°C) mismatched- matched	
			<i>ap</i> mm DNA 5	<i>ap</i> mm RNA 5	DNA	RNA
1.	DNA1	5'- CTT CTT CCT T - 3'	18.6	29.4	- 9.2	- 6.7
2.	PNA 1	N ² - cttcttct t -K -C'	37.0	41.2	- 10.3	- 10.1
3.	TRT 1	5'- C _g T _g T _g C _g T _g T _g C _g C _g T _g T - 3'	30.8	38.3	- 8.9	- 10.1
			<i>ap</i> mm DNA 6	<i>ap</i> mm RNA 6	DNA	RNA
4.	DNA2	5'- CAC TGA TTT CAA - 3'	29.4	32.8	- 6.5	- 11.9
5.	PNA 2	N ² - cactgattcaa-K -C'	45.7	48.6	- 8.1	- 12.0
6.	TRT 2	5'-C _g A _g C _g T _g G _g A _g T _g T _g C _g A _g A-3'	37.5	43.2	- 7.8	- 10.8

Subscript 'g' between the bases denotes the glycine linker in the backbone of polyamide DNA.

The lower case bases represent *aeg*PNA monomer units

mm DNA5: 5'-AAG GAT GAA G: complementary to TRT1; mm DNA6: 5'-TTG AAT TCA GTG: complementary to TRT2; mm RNA5: 5'-AAG GAU GAA G: complementary to TRT1; mm RNA6: 5'-UUG AAU UCA GUG: complementary to TRT2

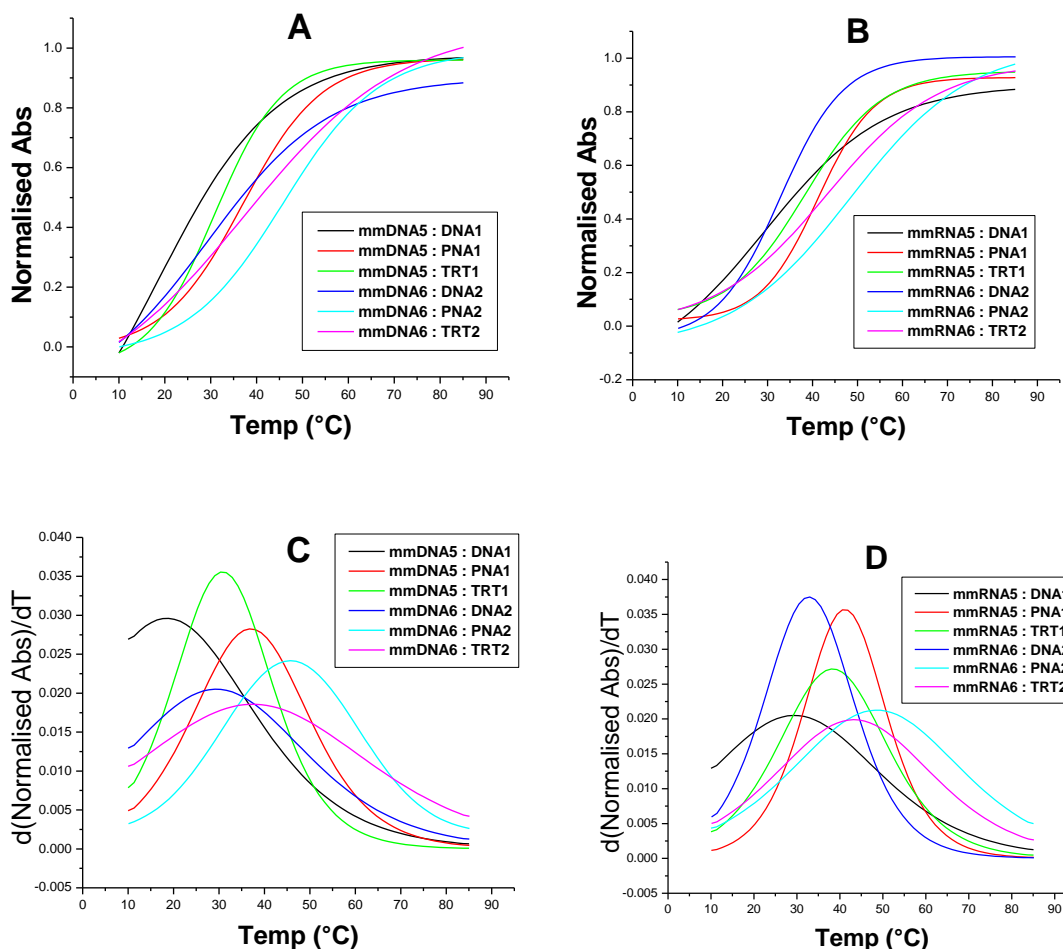


Figure 4. UV-Thermal Melting Profiles. Sigmoidal Curves for Sigmoidal Curves for (A) mmDNA5:DNA1/ PNA1/ TRT1, mmDNA6:DNA2/ PNA2/ TRT2 (B) mmRNA5:DNA1/ PNA1/ TRT1, mmRNA6:DNA2/ PNA2/ TRT2, first derivative curves of the sigmoidal curves for melting values (C) mmDNA5:DNA1/ PNA1/ TRT1, mmDNA6:DNA2/ PNA2/ TRT2 (D) mmRNA5:DNA1/ PNA1/ TRT1, mmRNA6:DNA2/ PNA2/ TRT2

4.6 Binding Stoichiometry: Job's Plot of the TRT1 oligomer

Ultraviolet absorption measurements are useful to determine the stoichiometry of oligonucleotide:DNA/RNA complexes which was described by Job.^{31, 32} In this method, the stoichiometry of the interacting strands could be obtained from the mixing curves, in which the absorbance at a given wavelength is plotted as a function of the mole fraction of each strand which determines the stoichiometry of the complex information.

Various stoichiometric mixtures of sequence TRT 1 and DNA 1 were made with relative molar ratios of strands 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 100:0, all at the same total strand concentration of 2 μ M in sodium phosphate buffer, 10 mM NaCl. The samples with the individual strands were annealed and UV absorbance was recorded. A mixing curve was plotted, absorbance at fixed wavelength (λ_{\max} 260 nm) against mole fraction of TRT 1 (Figure 5).

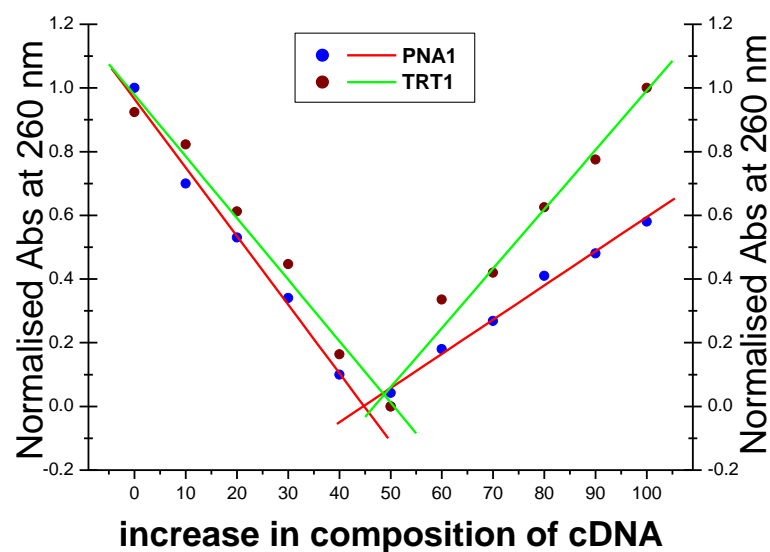


Figure 5. The Job's plot of the oligomer TRT1 and PNA1

4.7 Conformational analysis of the polyamide DNA:cDNA/RNA

The present study includes the replacement of conventional phosphate backbone of DNA by a polyamide DNA backbone. In order to examine if these modified oligomers cause any significant differences on the overall conformational features of the duplexes, the CD studies of polyamide DNA: cDNA as well as RNA were undertaken.

The CD studies concerning the single strands of the polyamide DNA sequence (Figure 6) did not show any characteristic features of any conformation.

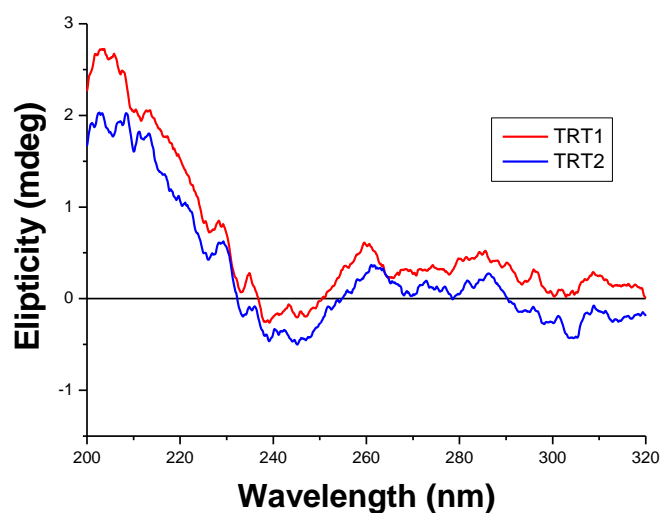


Figure 6. CD Profiles of Single strands of polyamide DNA

4.7.1 Conformational analysis of the polyamide CT oligomer

The CD profile of the derived DNA oligomers (**TRT 1**), Figure 7A with the modification when duplexed with cDNA showed similar CD pattern as that of the unmodified cDNA1:DNA1 duplexes but there was some difference in the pattern with cDNA1:PNA1. The CD profile of the derived DNA oligomers (**TRT 1**) with the modification when duplexed with RNA showed similar CD pattern as that of the unmodified DNA1: RNA1 duplex (Figure 7B).

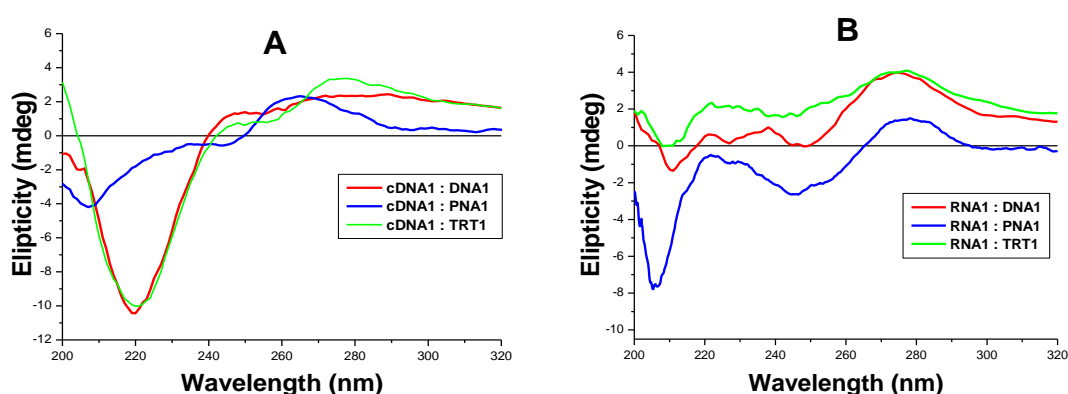


Figure 7. CD Profiles of CT sequence TRT 1 with (A) cDNA and (B) RNA

4.7.2 Conformational analysis of the mixed purine-pyrimidine oligomer

The CD profile of the derived DNA oligomers (**TRT 2**), Figure 8A with the modification when duplexed with cDNA showed similar CD pattern as that of the unmodified cDNA2:DNA2 duplexes. The CD profile of the derived DNA oligomers (**TRT 2**) with the modification when duplexed with RNA showed similar CD pattern as that of the unmodified DNA2: RNA2 duplex (Figure 8B).

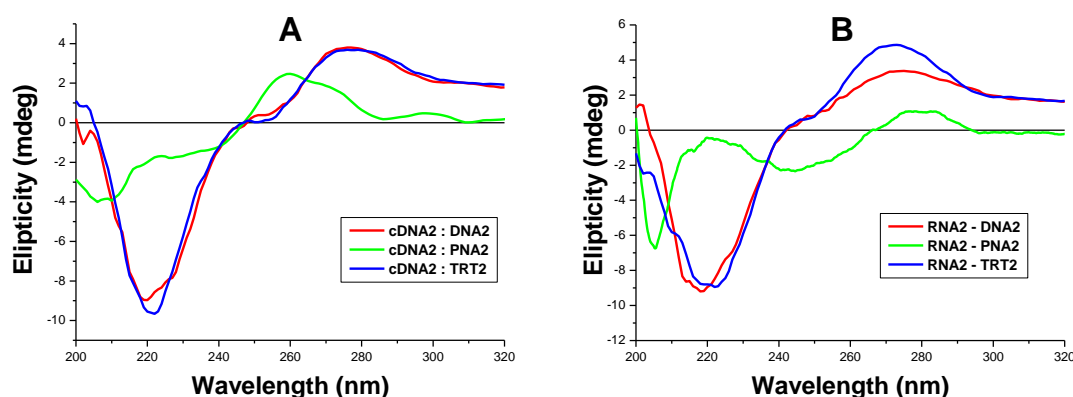


Figure 8. CD Profiles of mixed purine-pyrimidine sequence TRT 2 with (A) cDNA and (B) RNA

4.8 Summary

- ❖ A new backbone modified DNA analogue had been successfully designed and synthesized, having a polyamide linkage between the sugars in the backbone instead of phosphodiester linkage.
- ❖ The derived polyamide DNA showed higher binding affinity with respect to the unmodified DNA towards the antiparallel complementary DNA and RNA, RNA being preferred.
- ❖ The UV – T_m studies regarding the binding preference of the oligomers towards parallel complementary DNA/ RNA, results in non binding, inferring that the designed oligos were having directional preference.
- ❖ The mismatch studies also revealed that they possess base specific binding property.
- ❖ The CD studies of the derived oligomers did not show specific conformational feature in the duplexes with DNA/RNA.

4.9 Experimentation

Solid phase oligomer Synthesis

MBHA resin (60 mg, 1.75 mmol/g, 0.024 mmol) was washed and swelled in DCM for 1 h in solid phase flask. The resin was drained and washed with DMF. To a solution of *N*4-benzoyl-5-methyl 3'-tritylamino-2',3'-dideoxycytidine-5'-succinylate (16.19 mg, 0.024 mmol) in 500 mL DMF, DIPEA (12.47 mL, 0.072 mmol) and TBTU (9.63 mg, 0.028 mmol) was added. To this solution HOBt (3.24 mg, 0.024 mmol) in 100 mL of DMF was added. This activated monomer solution was added to the resin and the suspension was allowed for gentle shaking for 12 h. Resin was drained and washed with DMF (2 X), DCM (2 X) and pyridine (2 X). Unreacted amino groups capped with 10% acetic anhydride in pyridine for 1 h. The nucleoside loading on solid support was determined by spectrophotometric determination of the concentration of trityl cation at 410 nm released after detritylation using 3% TCA in DCM. Calculated loading value for resin using molar extinction coefficient³³ 35,300 M⁻¹ cm⁻¹ at 410 nm for trityl cation was 35.2 mmol/g, which was good enough for next synthesis.

Synthesis of polyamide DNA sequences

Synthesis of polyamide DNA sequences was then undertaken using protected nucleoside β amino acid monomeric units and trityl glycine on preloaded MBHA resin. The following synthetic condition were used, deprotection of trityl group using 3% trichloroacetic acid (3 min X 3)

- (i) Neutralization using 5% DIPEA in DCM (3min X 2)
- (ii) Coupling using 3 *eq.* of monomer, 9 *eq.* DIPEA, 3 *eq.* TBTU and 1.5 *eq.* of HOBt as activator with respect to loading value of resin. All are premixed in DMF prior to addition to resin. This suspension was added to the resin. Reaction time 6 h with occasional swirling
- (iii) Capping of unreacted amino groups using 10% acetic anhydride in pyridine (10 min X 2).

Cleavage of the oligomers from the solid support

In a typical cleavage reaction, the resin bound oligomers (5 mg) were treated with aqueous methanolic ammonia (1.5 mL) at 55 °C for 8 h. After that the resin was filtered followed by evaporation of the filtrate under vacuo. The terminal trityl group was deprotected before cleavage using 3% TCA in DCM. Purification of all the modified oligomers was carried out by reverse phase HPLC on RP-C18 column using 5% acetonitrile in 0.1 M TEAA buffer as eluent system and monitored at 260 nm and

were characterized by MALDI-TOF mass spectrometry. The spectra were acquired in linear mode and the matrix used for analysis was THAP (2',4',6'-trihydroxyacetophenone) – ammonium citrate (2:1).

UV- T_m Measurements

The concentrations were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases of DNA/RNA/PNA. The experiments were performed at 1 μ M concentrations of each strand. The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing 10 mM NaCl (100 mM NaCl for DNA:DNA/RNA) and were annealed by keeping the samples at 90 $^{\circ}$ C for 5 min followed by slow cooling to room temperature and then to 0 $^{\circ}$ C. Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Varian Cary 300 spectrophotometer scanning from 10-85 $^{\circ}$ C at a ramp rate of 0.5 $^{\circ}$ C per minute. Experiments were repeated at least thrice and the data were processed using Origin software 6.1. T_m ($^{\circ}$ C) values were derived from the first derivative curves and are accurate to within $\pm 0.3^{\circ}$ C

CD Analysis of the oligomers

The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (10 mM) and were annealed by keeping the samples at 90 $^{\circ}$ C for 5 min followed by slow cooling to room temperature. The experiments were performed at 1 μ M concentrations of each strand. All the CD spectra were recorded at room temperature with Jasco J-815 spectropolarimeter. All spectra represent an average of at least 8 scans recorded from 320 nm to 210 nm at a rate of 100 nm per minute in a 1 cm path length cuvette. All spectra were processed using Origin software 6.1, baseline subtracted and smoothed using a 5 point adjacent averaging algorithm.

4.10 Appendix

Synthesis of the p and ap c DNA/RNA and mismatch cDNA/RNA

The *p* and *ap* cDNA and mismatch cDNA oligonucleotides were synthesized on BioAutomationMerMade-4 DNA Synthesizer using standard β -cyanoethyl phosphoramidite chemistry.^{34,35} The oligomers were synthesized in the 3'→5' direction on polystyrene solid support (control pore glass (CPG) resin) followed by ammonia treatment to cleave the ester functionality that joins support to the 3'- terminus of the oligomers and deprotects the exocyclic amino protecting groups used during the synthesis.. The oligonucleotides were desalted by gel filtration; their purity ascertained by RP-HPLC on a C18 column (0.1N TEAA: ACN) to be more than 95% and were used without further purification in the biophysical studies. The RNA oligonucleotides were obtained commercially.

Spectral Data of the compounds

(1) 2', 3'-Deoxy-3'-tritylaminothymidine-4'-carboxylic acid

¹H NMR (400 MHz, DMSO-d₆ + drop of D₂O): 8.73 (s, 1H, H6), 7.42-7.44 (m, 6H, aromatic), 7.22-7.26 (m, 6H, aromatic), 7.12-7.16 (m, 3H, aromatic), 5.87 (dd, J= 5.52,5.02 Hz, 1H, H1'), 4.08 (d, J= 5.78 Hz, 1H, H4'), 2.89 (dd, J= 11.54, 5.52 Hz, 1H, H3'), 1.64 (s, 3H, CH₃), 1.21 (m, 1H, H2''), 1.00 (m, 1H, H2).

(2) 2', 3'-Dideoxy-3'-tritylamino-N⁴-benzoyl-5-methylcytidin-4'-carboxylic acid

¹H NMR (400 MHz, DMSO-d₆ + drop of D₂O): 8.84 (br s, 1H, NH), 8.12 (s, 2H, aromatic), 7.57 (m,1H, aromatic), 7.43-7.47 (m, 9H, aromatic), 7.23-7.27 (m, 6H, aromatic), 7.13-7.16 (m, 3H, aromatic), 5.96 (t, J= 4.77 Hz, 1H, H1'), 4.22 (d, J= 5.77 Hz, 1H, H4'), 3.05 (m, 1H, H3'), 1.89 (s, 3H, CH₃), 1.40 (m, 1H, H2''), 1.06 (m, 1H, H2').

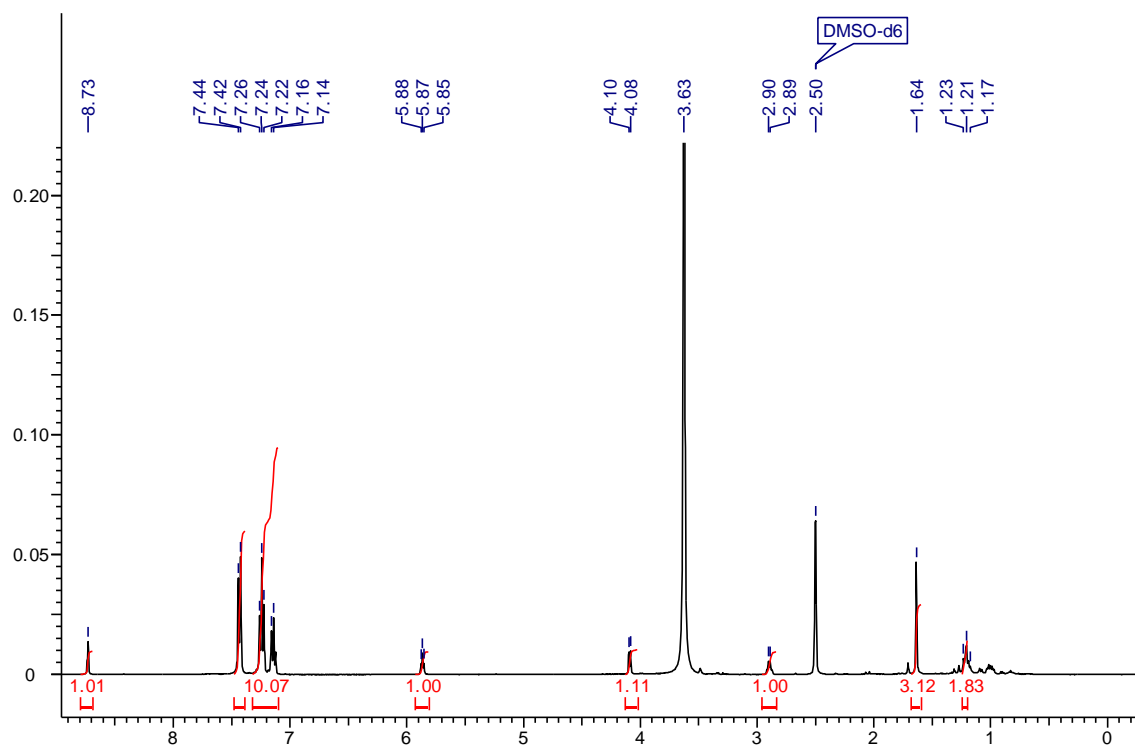
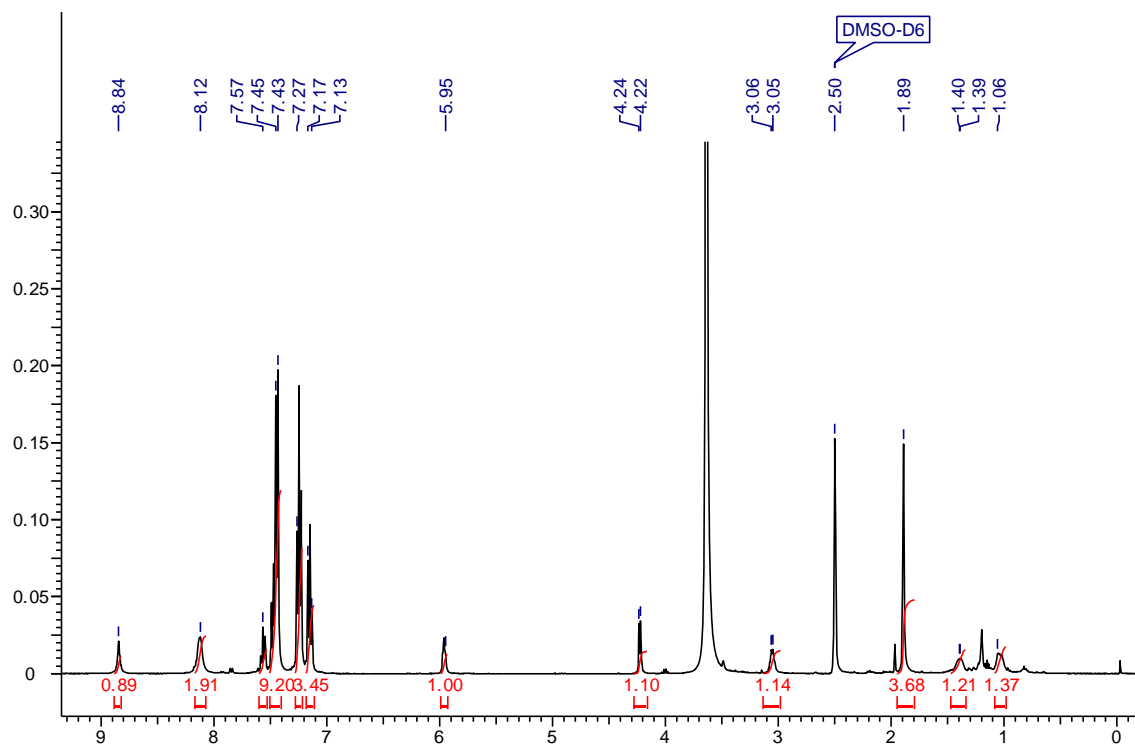
(3) 2',3'-Dideoxy-3'-tritylamino-N⁶-benzoyladenine-4'-carboxylic acid

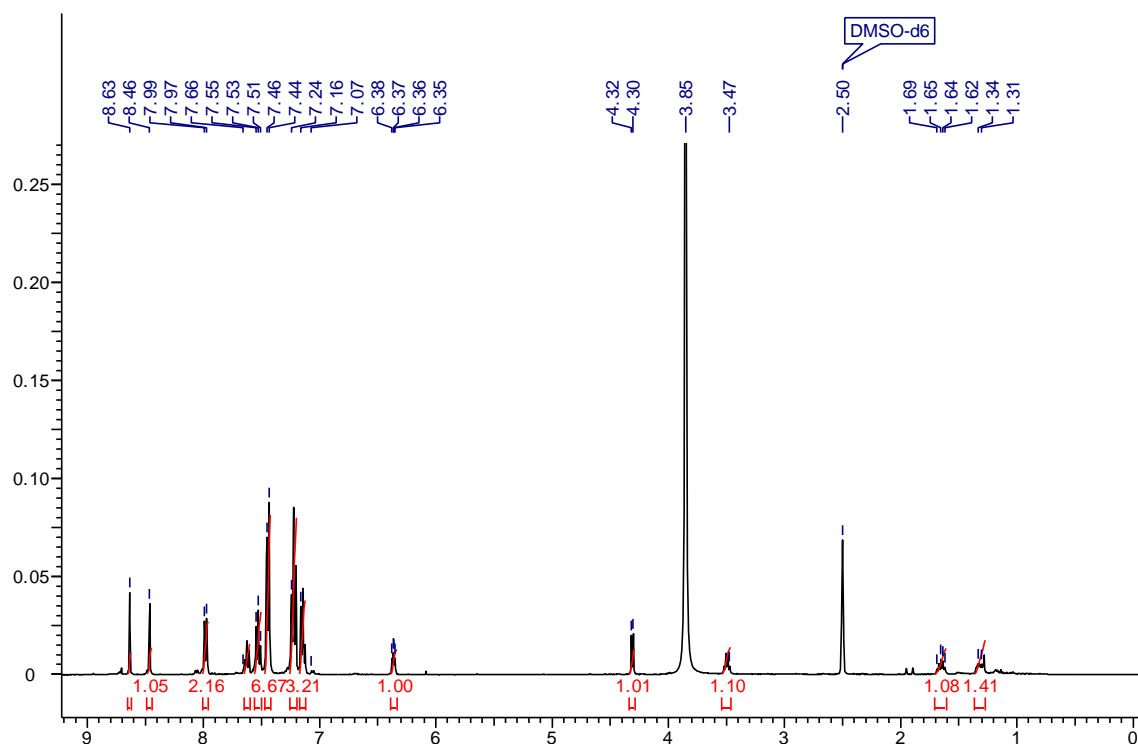
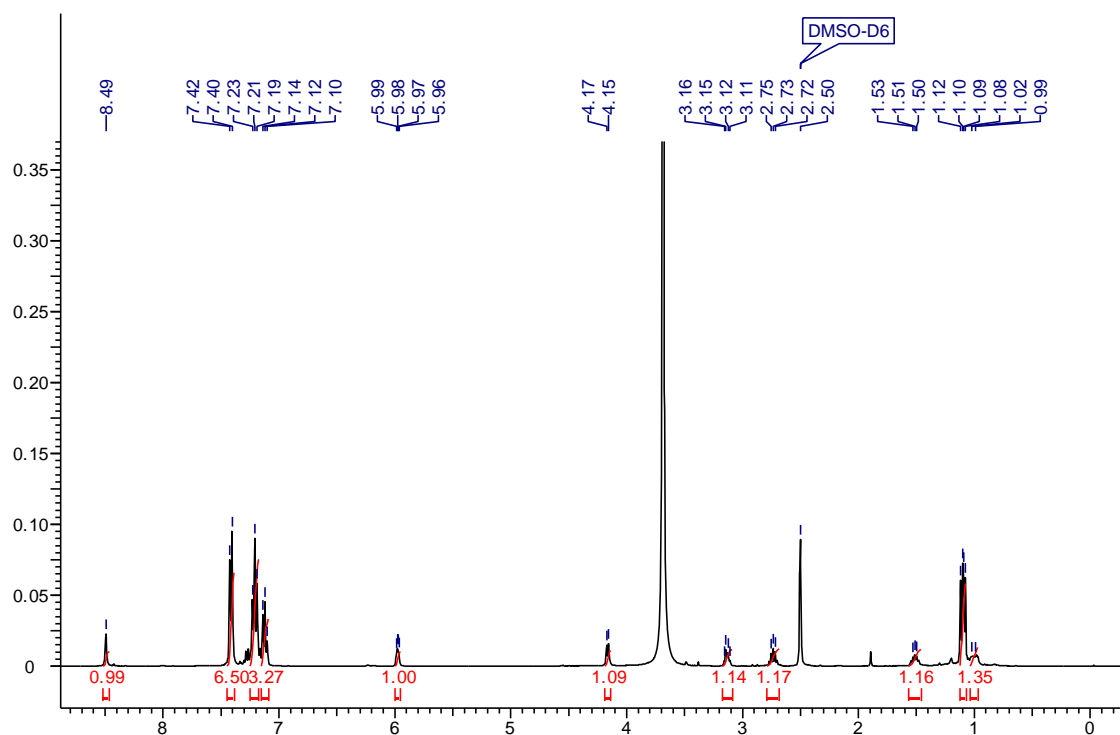
¹H NMR (400 MHz, DMSO-d₆ + drop of D₂O): 8.63 (s, 1H, C2-H), 8.46 (s, 1H, C8-H), 7.97 (d, J= 7.28 Hz, 2H, aromatic), 7.66 (m, 1H, aromatic), 7.51-7.55 (m, 2H, aromatic), 7.44-7.46 (m, 6H, aromatic), 7.24 (m, 6H, aromatic) 7.07-7.16 (m, 3H, aromatic), 6.35-6.38 (dd, J= 6.27, 4.52 Hz, 1H, H10), 4.30 (d, J= 6.27 Hz, 1H, H40), 3.47 (dd, J= 13.05, 6.53 Hz, 1H, H3'), 1.62-1.69 (m, 1H, H2''), 1.31-1.34 (m, 1H, H2').

(4) 2',3'-Dideoxy-3'-tritylamino-*N*²-isobutyrylguanosin-4'-carboxylic acid

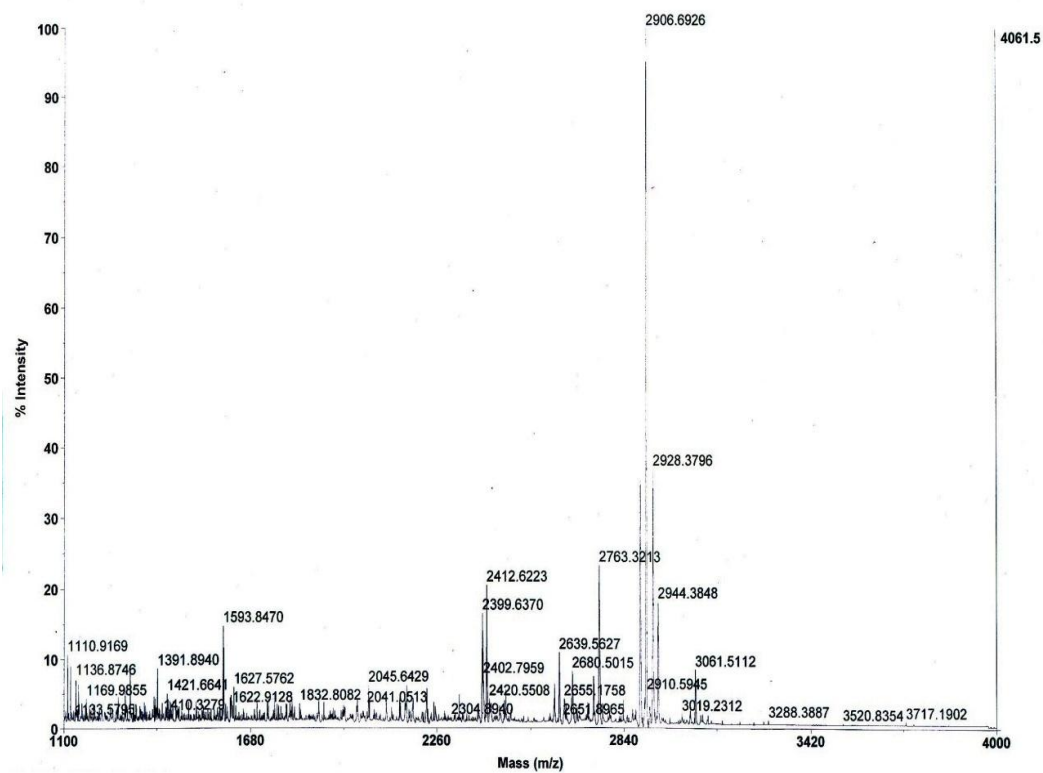
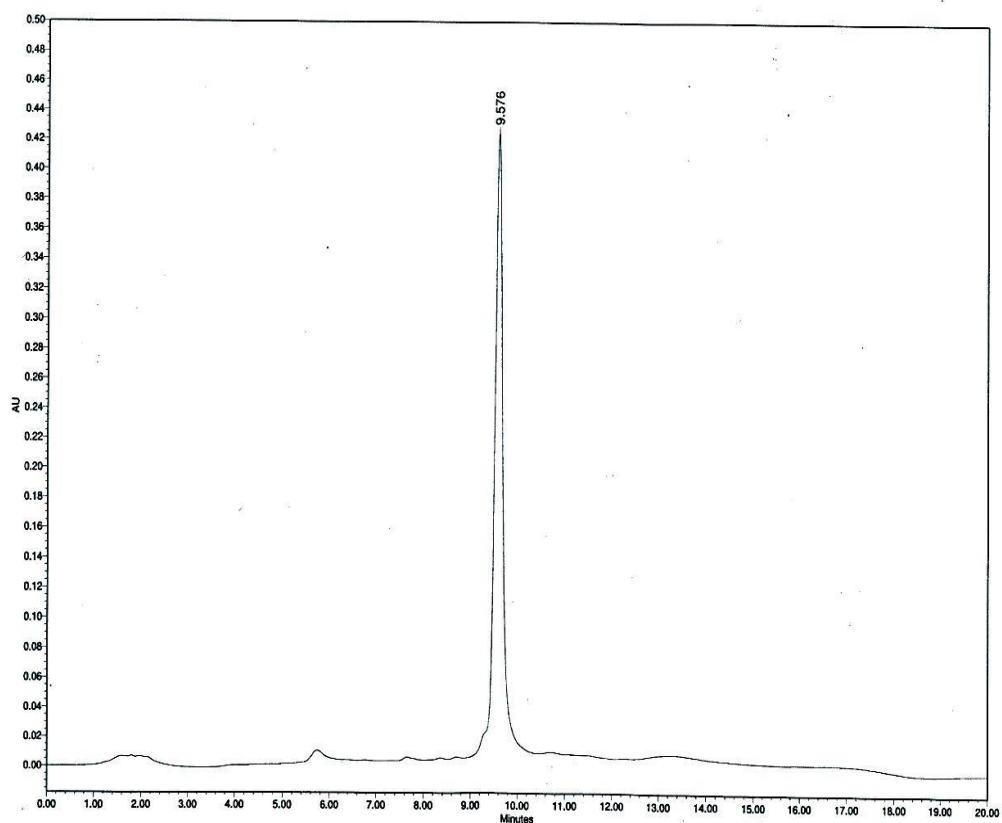
¹H NMR (400 MHz, DMSO-d₆ + drop of D₂O): 8.49 (s, 1H, C8-H), 7.40 (d, J= 7.78 Hz, 6H, aromatic), 7.10-7.23 (m, 10H, aromatic), 5.96-5.99 (dd, J= 5.52, 3.76 Hz, 1H, H1'), 4.17 (d, J= 6.77 Hz, 1H, H4'), 3.11-3.16 (m, 1H, H3'), 2.72-2.75 (m, 1H, *i*Bu-CH), 1.50-1.53 (m, 1H, H2''), 1.08-1.12 (2d, 6H, *i*Bu-CH₃), 0.99 (m, 1H, H2').

Compounds - Spectral data	Page Nos.
¹ H spectrum of compound 1	209
¹ H spectrum of compound 2	209
¹ H spectrum of compound 3	210
¹ H spectrum of compound 4	210
HPLC and MALDI-TOF of TRT 1	211
HPLC and MALDI-TOF of TRT 2	212

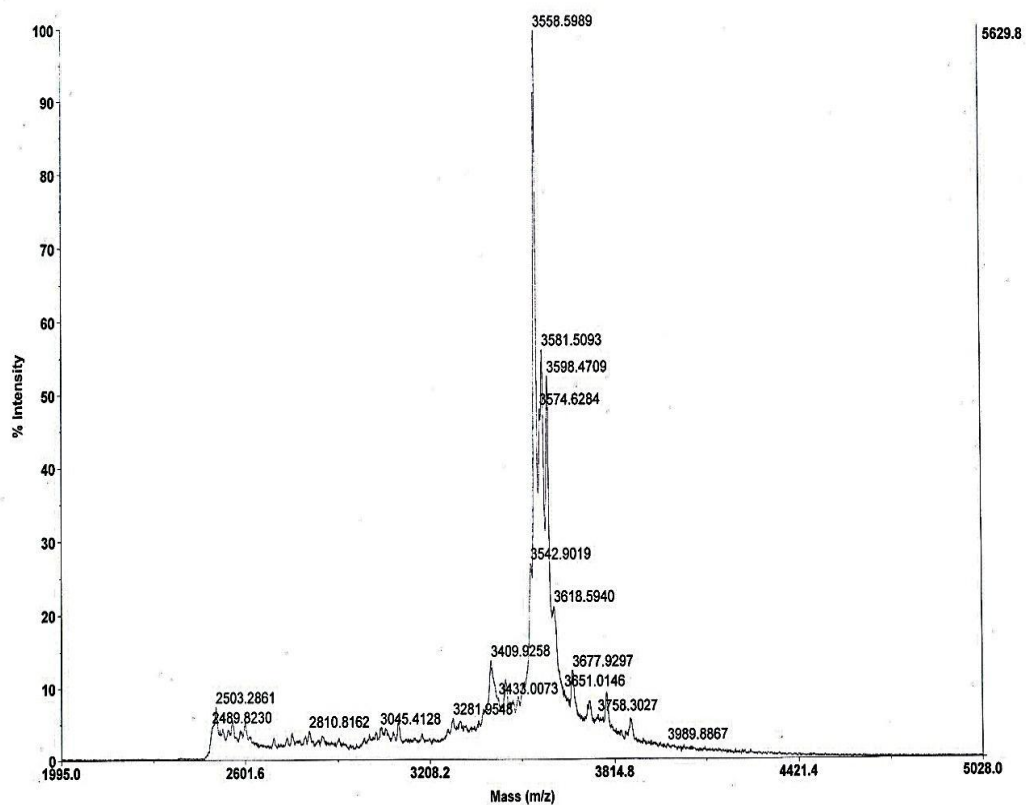
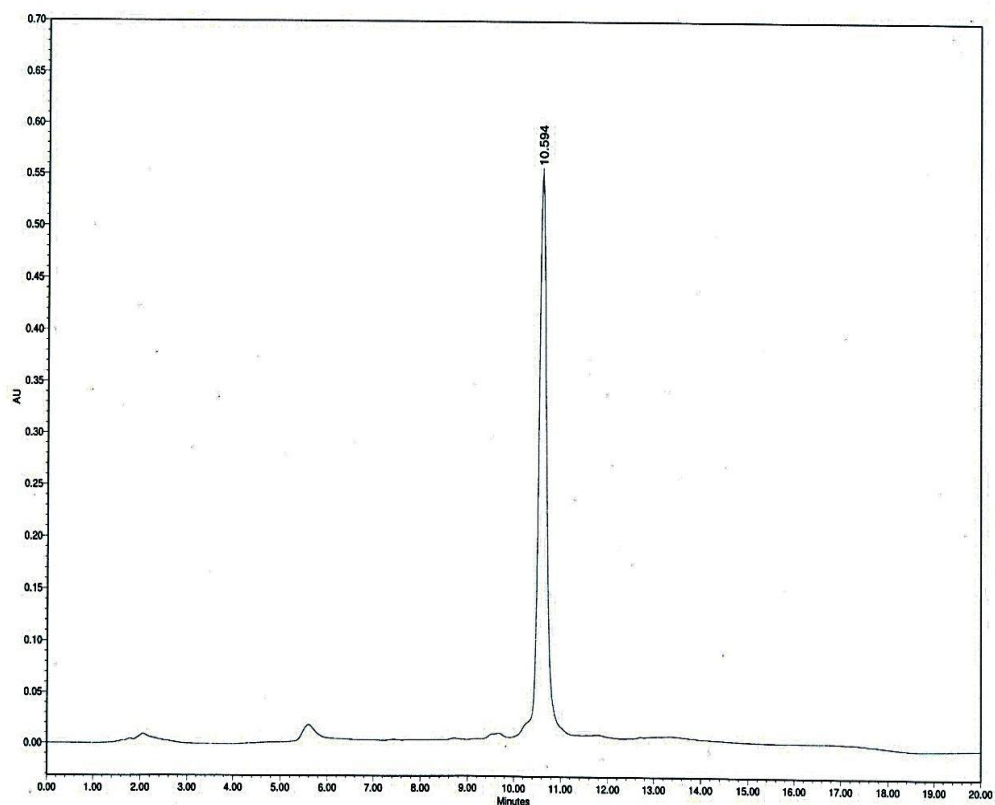
^1H spectrum of compound 1 ^1H spectrum of compound 2

^1H spectrum of compound **3** ^1H spectrum of compound **4**

HPLC and MALDI-TOF of TRT1



HPLC and MALDI-TOF of TRT2



4.11 References

1. P. C. Zamecnik and M. L. Stephenson, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 280-284.
2. R. Kole, A. R. Krainer and S. Altman, *Nat. Rev. Drug Discov.*, 2012, **11**, 125-140.
3. R. Corradini, S. Sforza, T. Tedeschi, F. Totsingan, A. Manicardi and R. Marchelli, *Curr. Top. Med. Chem.*, 2011, **11**, 1535-1554.
4. V. Madhuri and V. A. Kumar, *Org. Biomol. Chem.*, 2010, **8**, 3734-3741.
5. J. Magdalena, S. Fernandez, M. Ferrero and V. Gotor, *Tetrahedron Lett.*, 1999, **40**, 1787-1790.
6. B. A. Linkletter and T. C. Bruice, *Bioorg. Med. Chem.*, 2000, **8**, 1893-1901.
7. P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1500.
8. J. Lebreton, A. Waldner, V. Fritsch, R. R. Wolf and A. Demesmaeker, *Tetrahedron Lett.*, 1994, **35**, 5225-5228.
9. J. Lebreton, A. Demesmaeker, A. Waldner, V. Fritsch, R. M. Wolf and S. M. Freier, *Tetrahedron Lett.*, 1993, **34**, 6383-6386.
10. C. J. Wilds, G. Minasov, F. Natt, P. von Matt, K. H. Altmann and M. Egli, *Nucleosides Nucleotides Nucleic Acids*, 2001, **20**, 991-994.
11. P. S. Pallan, P. von Matt, C. J. Wilds, K. H. Altmann and M. Egli, *Biochemistry*, 2006, **45**, 8048-8057.
12. K. Gogoi, A. D. Gunjal, U. D. Phalgune and V. A. Kumar, *Org. Lett.*, 2007, **9**, 2697-2700.
13. S. S. Gokhale, K. Gogoi and V. A. Kumar, *J. Org. Chem.*, 2010, **75**, 7431-7434.
14. K. Gogoi and V. A. Kumar, *Chem. Commun.*, 2008, 706-708.
15. T. Vilaivan and C. Srisuwannaket, *Org. Lett.*, 2006, **8**, 1897-1900.
16. W. Mansawat, C. Vilaivan, A. Balazs, D. J. Aitken and T. Vilaivan, *Org. Lett.*, 2012, **14**, 1440-1443.
17. S. Bagmare, M. D'Costa and V. A. Kumar, *Chem. Commun.*, 2009, 6646-6648.
18. S. Bagmare, M. Varada, A. Banerjee and V. A. Kumar, *Tetrahedron*, 2013, **69**, 1210-1216.
19. H. Rink, *Tetrahedron Lett.*, 1987, **28**, 3787-3790.
20. K. L. Fearon and J. S. Nelson, *Current Protocols in Nucleic Acid Chemistry*, 2001, 4.7. 1-4.7. 40.

21. K. J. Divakar and C. B. Reese, *J. Chem. Soc.-Perkin Trans. 1*, 1982, 1171-1176.
22. J. P. Malkinson and R. A. Falconer, *Tetrahedron Lett.*, 2002, **43**, 9549-9552.
23. J. B. Epp and T. S. Widlanski, *J. Org. Chem.*, 1999, **64**, 293-295.
24. A. DeMico, R. Margarita, L. Parlanti, A. Vescovi and G. Piancatelli, *J. Org. Chem.*, 1997, **62**, 6974-6977.
25. S. Chandrasekhar, G. P. K. Reddy, M. U. Kiran, C. Nagesh and B. Jagadeesh, *Tetrahedron Lett.*, 2008, **49**, 2969-2973.
26. C. P. Fenna, V. J. Wilkinson, J. R. P. Arnold, R. Cosstick and J. Fisher, *Chem. Commun.*, 2008, 3567-3569.
27. R. Cosstick, R. Threlfall and J. Fisher, in *Nucleic Acids Symposium Series*, Oxford Univ Press, 2008, pp. 313-314.
28. R. Threlfall, A. Davies, N. Howarth and R. Cosstick, *Nucleosides Nucleotides Nucleic Acids*, 2007, **26**, 611-614.
29. Sundaral.M, *Biopolymers*, 1969, **7**, 821-860.
30. L. J. Rinkel and C. Altona, *J. Biomol. Struct. Dyn.*, 1987, **4**, 621-649.
31. P. Job, *Ann Chim.*, 1928, **9**, 113-203.
32. C. R. Cantor and P. R. Schimmel, *Journal of Solid-Phase Biochemistry*, 1980, **5**.
33. R. Rathore, C. L. Burns and I. A. Guzei, *J. Org. Chem.*, 2004, **69**, 1524-1530.
34. R. A. Jones and M. J. Gait, *Oligonucleotide Synthesis—A Practical Approach*, 1984, **1**, 23-34.
35. S. Agrawal, *Protocols for oligonucleotides and analogs: synthesis and properties*, Springer, 1993.

