

**Conformationally restricted PNA analogues;  
Synthesis and biophysical studies of cationic  
pyrrolidyl PNA**

**Thesis submitted to  
The University of Pune for degree of**

**Doctor of Philosophy  
in  
Chemistry**

**By  
Kosgi Sreedhar**

**Research Supervisor  
Prof. Krishna N. Ganesh**

**Division of Organic Chemistry  
National Chemical Laboratory**

**Pune 411008**

**October, 2012**

## **CERTIFICATE**

This is to certify that the work presented in the thesis entitled “**Conformationally restricted PNA analogues; Synthesis and biophysical studies of cationic pyrrolidyl PNA**” submitted by **Kosgi Sreedhar** for the degree of Doctor of Philosophy, was carried out by the candidate at the National Chemical Laboratory, Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

**Prof. Krishna N. Ganesh, FNA, FNASc**  
(Research Supervisor)  
Director, IISER Pune  
Pune- 411008

**October, 2012**

J. C. Bose Fellow  
National Chemical Laboratory,  
Pune-411 008,  
INDIA.

*Dedicated to my beloved friends*

## **CANDIDATE'S DECLARATION**

I hereby declare that the thesis entitled “**Conformationally restricted PNA analogues; Synthesis and biophysical studies of cationic pyrrolidyl PNA**” submitted for the award of degree of *Doctor of Philosophy* in *Chemistry* to the University of Pune has not been submitted by me to any other university or institution. This work was carried out at the National Chemical Laboratory, Pune, India. Such materials as obtained from other sources have been duly acknowledged in the thesis.

**Kosgi Sreedhar**  
National Chemical Laboratory  
Pune 411 008.

**October, 2012**

## *Acknowledgement*

*It gives me a great pleasure to express my deep sense of gratitude towards my research supervisor Prof. K. N. Ganesh for all the advice, guidance and encouragement. He has been a great teacher and guide and has been very patient and tolerant towards my erratic behavior and set me back in the right course. This will surely serve as a good platform for my future endeavors.*

*I am also grateful to Dr. V. A. Kumar, Dr. Moneesha Fernandes, Dr. Srinivas Hotha, Mrs. Anita Gunjal for their help and encouragement during the course of this work.*

*I thank all my seniors and juniors for their kind help and creating a cheerful atmosphere around. I thank all my friends in other labs of this departments and other departments for helping me whenever I was in need. Thanks to Dr. Aswani, Mahesh, roopa, Manaswini, Pradnya, satish, Deepak, Nitin, VDC, Tanpreet and madan. I thank director NCL, for allowing me to work in this premier institute and providing the infrastructure. My sincere thank to IISER, Pune for providing all necessary facilities and CSIR, New Delhi, for the financial support.*

*This page is honor to friends in my life who have already crossed over to begin a new journey. Each one of these people touched my heart and my life and will live on forever in my heart.*

*I thanks D.S. Anna, Sunilanna and Raman for their support and help. Hi ra Bammardhi, yellanna, Praveen ga etc thanx for being in my life and making it complete. Hardest part of living for me is losing you ra mama (Srikanth.L)*

*To all my friends;*

*Thank you for being there for me through some of the most difficult times in my life as if it were your own, forgiving my faults, immature lapses in judgment, tolerating my idiosyncrasies, for the phone calls, notes, e-mails of support when I was overwhelmed*

*with life, for being my friend always, not just when it was convenient, trusting me enough to tell me what's bothering you, for not letting time or distance affect our friendship.*

*Finally thanking you for hugs, smiles, inside jokes, and memories. Understanding that I could never write a long enough blog post to express how much you all mean to me.*

*My friends at hostel have made my stay at NCL a lively and memorable one and I am indebted to them for their support.*

*If I wrote down everything I ever wanted in a wife, I would not have believed I could meet someone better! All of your continued support and urging is deeply appreciated.*

*She has been preserved through my paranoia, insomnia, mood changes.*

*Needless to mention about parents care and sacrificing.*

*Kosgi sreedhar*

# Contents

<b>Abbreviations</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>1 Chapter 1</b>	
<b>Introduction</b>	
1.1 Nucleic Acids: Chemical Structure and Importance	1
1.1.1 Hydrogen bonding between nucleobases	3
1.1.2 Oligonucleotides as Therapeutic Agents	4
1.1.3 Antisense Oligonucleotides	5
1.1.4 Modified oligonucleotides as antisense therapeutic agents	6
1.1.5 Sugar Modifications	12
1.1.6 Sugar-Phosphate modified oligonucleotide	14
1.2 PNA-DNA complex formation and their structures	16
1.2.1 Duplex formation with complementary DNA and RNA	16
1.2.2 Triplex formation	17
1.2.3 PNA targeting double stranded DNA	18
1.2.4 Structure of PNA-DNA duplexes:	19
1.3 Chemical Modifications of PNA	19
1.3.1 Introduction of ionic functional groups into the PNA	20
1.3.2 Construction of bridged structures	22
1.4 BIOLOGICAL APPLICATIONS OF PNA	29
1.5 PRESENT WORK:	29
1.6 References	31
<b>Chapter 2</b>	
<b>2.1 Introduction</b>	<b>37</b>
2.2 Rationale design behind the work	40
2.2.1 Rationale	42
2.3 Objectives	43
2.3.1 Synthesis of <i>cis</i> -(3 <i>R</i> ,4 <i>R</i> )-3,4-diamino pyrrolidinyl derived PNA monomers (16 & 17)	43
2.3.4 Synthesis of (2 <i>S</i> ,3 <i>R</i> ) (1→2) <i>amp</i> PNA Monomer (27)	46
2.3.5 Synthesis of (2 <i>R</i> ,3 <i>R</i> ) (2→3) <i>amp</i> PNA Monomer.	48
2.3.6 Synthesis of <i>t</i> -Boc-protected <i>aeg</i> PNA monomers	49
2.3.7 Synthesis of Fmoc-protected <i>aeg</i> PNA monomer	51

2.4 Conclusion	51
2.5 Experimental Section	51
2.6 References	70
2.7 Appendix	73
<b>Chapter 3</b>	
3.1 Solid Phase Peptide Synthesis	113
3.1.1 Resins for solid phase peptide synthesis	115
3.1.2 Coupling reagents	119
3.1.3 Resin Tests	120
3.1.4 General protocols adopted for PNA synthesis	121
3.1.5 Cleavage of the PNA oligomers from the solid support	121
3.2 Biophysical studies of PNA:DNA complexes	122
3.2.1 UV-Spectroscopic study of PNA:DNA complexes	122
3.3 Results and Discussion	123
3.3.1 Solid Phase synthesis of (3→4) <i>dap</i> PNA (16), (1→3) <i>dap</i> PNA (17), (2→3) <i>amp</i> PNA (39) and (1→2) <i>amp</i> PNA (27) oligomers.	123
3.3.2 Purification of the PNA oligomers	125
3.3.3 Synthesis of complementary DNA oligonucleotides	126
3.3.4 Biophysical studies of PNA:DNA complexes	127
3.3.5 UV thermal studies of (3→4) <i>dap</i> PNA(16) with DNA1 and DNA2	133
3.3.6 UV thermal studies of (1→3) <i>dap</i> PNA (17)	133
3.3.7 UV thermal studies of (1→2) <i>amp</i> PNA (27) with DNA1 and DNA2	133
3.3.8 UV thermal studies of (2→3) <i>amp</i> PNA (39) with DNA1 and DNA2	134
3.4 Factors affecting the binding of PNA to complementary DNA	134
3.4.1 Effect of backbone modification in <i>aeg</i> PNA	134
3.4.2 Effect of backbone extension by one carbon	134
3.4.3 Electrostatic influence in binding	135
3.4.4 Effect of ring puckering	135
3.4.5 Effect of position of modification	138
3.4.6 Effect of primary or secondary amino group	138
3.5 Comparison of UV melting temperatures (0°C) of <i>cis</i> and <i>trans dap</i> (diamino pyrrolidine ) PNAs.	139
3.6 Summary	141
3.7 Conclusion	143
3.8 Experimental	144
3.9 References	147
3.10 Appendix	149



## Chapter 4

4.1 Introduction	164
4.2 Factors effecting the polyproline conformation	169
4.2.1 Solvent effect	169
4.2.2 Stereoelectronic effects	170
4.2.3 Effect of Urea	170
4.2.4 Effect of terminal functional groups	171
4.2.5 Effect of conformer of X <sub>aa</sub> -Pro amide bond	172
4.3 Importance of polyproline II	173
4.3.1 Utility as cell penetrating agents	173
4.3.2 Rigid molecular Spacer	175
4.3.3 Cell motility	176
4.4 Conformation of peptides by circular dichroism (CD)	176
4.5 Rationale for present work	177
4.6 Aim of the current work	179
4.7 Protocols followed for solid phase peptide synthesis	180
4.8 Results and discussions	182
4.8.1 Conformational examination of synthesized polyprolines (1-3) by CD spectroscopy	185
4.9 Conclusion	195
4.10 Experimental section	195
4.11 References	199
4.12 Appendix	202

## Abbreviations

A	Adenine
Ac <sub>2</sub> O	Acetic anhydride
ACN	Acetonitrile
<i>aeg</i>	Aminoethylglycine
<i>aep</i>	Aminoethylpropyl
ala	Alanine
<i>ap</i>	Antiparallel
aq.	Aqueous
Boc	Tert.butylloxycarbonyl
C	Cytosine
cbz	benzyloxy carbonyl
CD	Circular Dichroism
dA	Deoxy adenine
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
dG	2'-deoxyguanine
DIPEA	Diisopropylethylamine
DMF	N,N-Dimethylformamide
DNA	2'-deoxyribonucleic acid
ds	Double stranded
Et	Ethyl
EtOAc	Ethyl Acetate
Fmoc	9-Fluorenylmethoxycarbonyl
g	Gram
G	Guanine
gly	Glycine
h	Hours
HBTU	<i>O</i> -(1H-Benzotriazol-1-yl) <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate
HIV	Human Immuno Deficiency Virus
HOBT	1-Hydroxybenztriazole
HPLC	High Performance Liquid Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry

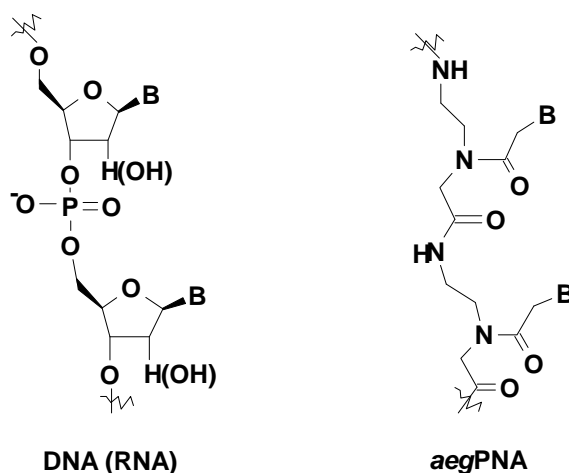
Lys	L-Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-Methyl benzhydryl amine
mg	Milligram
MHz	Megahertz
μM	Micromolar
ml	Milliliter
mM	Millimolar
mmol	Millimoles
M.P.	Melting point
MS	Mass spectrometry
N	Normal
nm	Nanometer
NMP	N-methyl pyrrolidine
NMR	Nuclear Magnetic Resonance
ONs	Oligonucleotides
p	Parallel
PCR	Polymerase Chain Reaction
PPh <sub>3</sub>	Triphenyl phosphine
PNA	Peptide Nucleic Acid
Pro	Proline
Pyr	pyrrolidine
RNA	Ribonucleic acid
r.t.	Room temperature
ss	Single strand/ Single stranded
s	Seconds
T	Thymine
TBTU	<i>O</i> -(1H-Benzotriazol-1-yl) <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate
TEA	Triethylamine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethanesulphonic acid
THF	Tetrahydrofuran
UV-Vis	Ultraviolet- Visible

## ABSTRACT

The thesis entitled “**Conformationally restricted PNA analogues; Synthesis and biophysical studies of cationic pyrrolidyl PNA.**” is divided into 4 chapters as follows:

### Chapter 1: Introduction

There has been considerable interest in the field of modified nucleic acids and oligonucleotide mimics due to their utility as therapeutics or diagnostic agents. However, the limitations of unmodified oligonucleotides are i) they can be rapidly recognised and cleaved by the action of nucleases that primarily hydrolyze the phosphodiester bond of the internucleoside backbone and ii) the ability of the negatively charged DNA to cross phospholipid cell membrane is poor. Hence structural modifications are required to enable oligonucleotides to be used as antisense therapeutic agents in biological systems and also to understand the structural changes that they introduce in the oligonucleotides by incorporating them. Nucleobase, sugar ring and the phosphate backbone are the three possible chemical sites to induce modifications. Of the many mimics that have been developed to date the most widely studied are the Peptide Nucleic Acid (PNA)<sup>1,2</sup> which has been shown to hybridize with complimentary DNAs with higher affinity and fidelity than DNAs.



**Figure 1:** DNA (RNA) and PNA structures

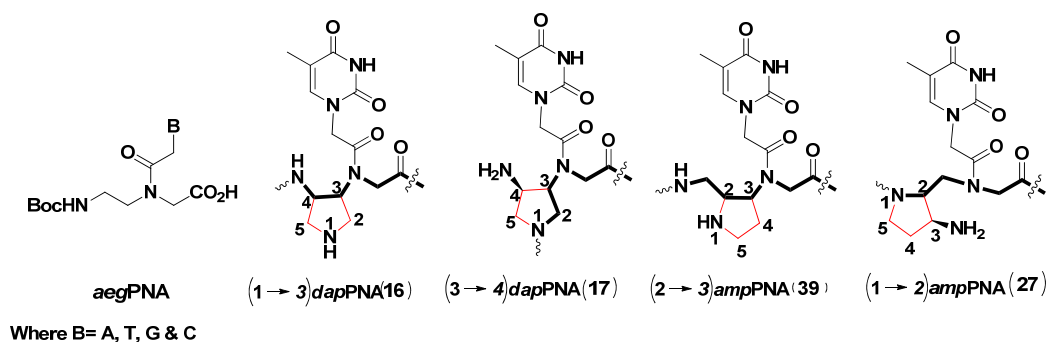
Peptide nucleic acid (PNA) is a class of compound in which the entire negatively charged sugar-phosphate backbone of DNA is replaced by a neutral and achiral polyamide backbone consisting of repeated units of N-(2-Aminoethyl)-glycine. The nucleobase is attached to the backbone through a conformationally rigid tertiary acetamide linkage.

PNA binds to DNA and RNA with high specificity and thermal stability by forming duplexes (Watson-Crick base pair) and triplexes (Watson-Crick base pair and Hoogsteen hydrogen bonding) that are much more stable than the corresponding DNA-DNA hybrids. In contrast to homopyrimidine sequences which form triplexes, mixed sequences of PNA bind to complementary DNA/RNA with 1:1 stoichiometry to form duplexes. Due to these exceptional properties, PNA and its analogues have major applications as a tool in molecular biology, as lead compounds for development of gene targeted drugs through antisense/antigene technology, for diagnostics and biosensors, and as building blocks for designing PNA supramolecular constructs.

Despite the success in many applications, PNA suffers from i) poor aqueous solubility ii) tendency to aggregate in aqueous media and iii) insufficient cellular uptake. In addition, PNA is achiral and can bind in both parallel and antiparallel orientations with nucleic acids, which leads to ambiguity in orientational selectivity. Various modifications on PNA backbone have been made to overcome these problems in an effective manner by many researchers in the recent past, which includes rigidity of the backbone, introduction of chirality into PNA by linking chiral amino acids, introduction of positive charge in the PNA backbone etc. This may lead to the improved aqueous solubility and better binding. This section briefly reviews the recent advancements in the area of chemistry of peptide nucleic acids.

## **Chapter 2: Synthesis of 2,3 and 3,4-disubstituted pyrrolidine based peptide nucleic acid monomers**

**Objective:** It's reported in the literature that the introduction of positive charge in the PNA backbone improves its binding. Considerable interest exists in making positively charged PNAs, as they are expected to possess superior ability to strand invade complementary DNA sequences.



**Figure 2:** Chemical structure of modified PNA monomers

Recently, a five membered carbocyclic analogue of PNA termed as *cp*PNA (*cp* refers to cyclo pentyl) with *cis* chemistry was reported from our group<sup>3</sup> and with *trans* chemistry by Appella *et al*<sup>4</sup> (Figure 2). DNA binding studies on these PNAs showed stabilization of the derived PNA:DNA/RNA complexes. So it may be interesting to make this pre-organised *cp*PNA cationic by the introduction of a nitrogen atom in the five membered cyclopentyl ring in order to generate 2,3 and 3,4-disubstituted pyrrolidine PNA monomers **16**, **17**, **27** & **39** (Figure 2). These involve *dap*PNA (**16** & **17**) derived from 3,4-diaminopyrrolidine and *amp*PNA (**27** & **39**) derived from 3-amino proline. In addition, this would also be interesting to investigate structural-biophysical activity relationship of these analogues. This chapter describes the synthesis of above modified PNA monomers.

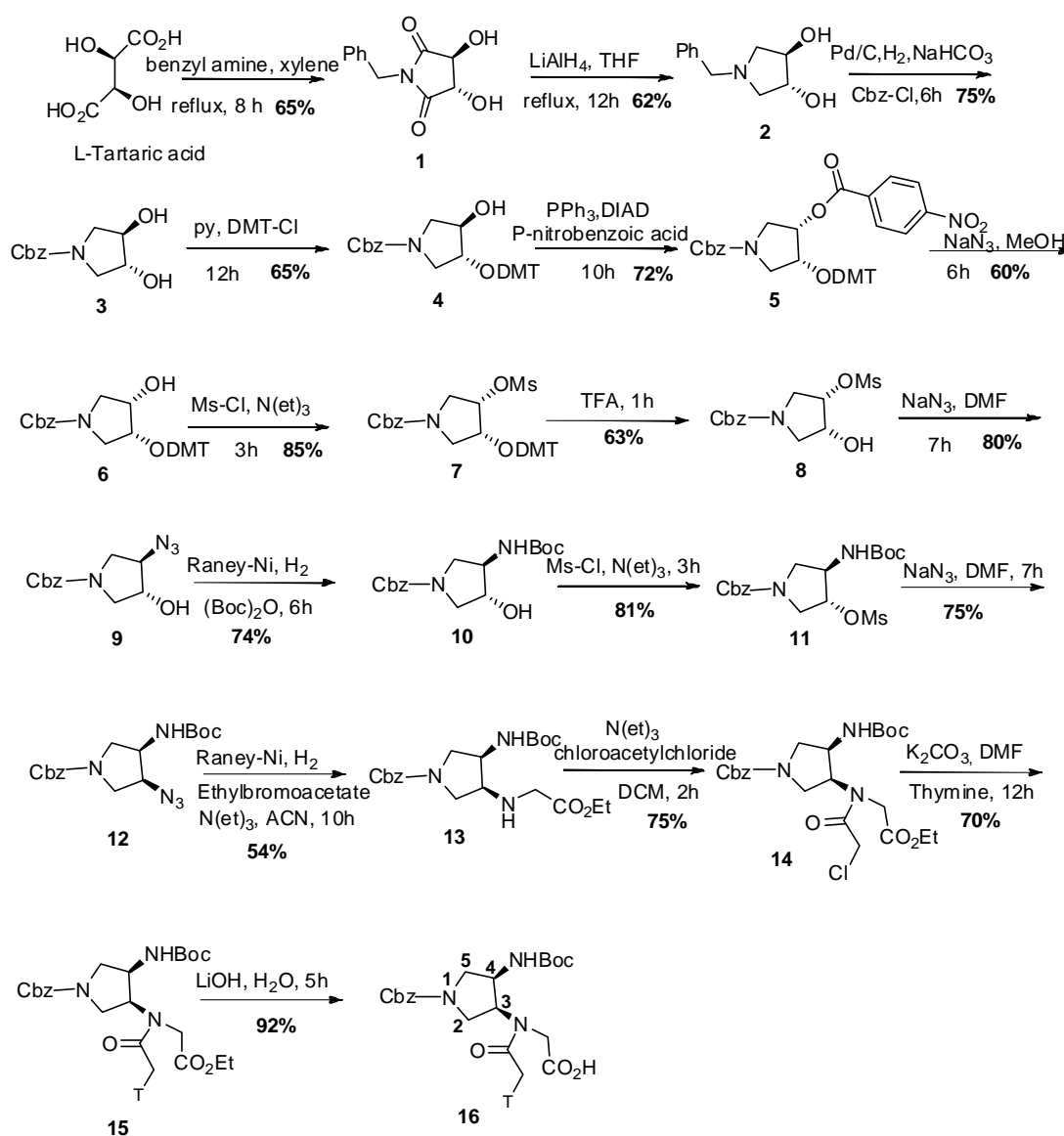
### Synthesis of *cis*-(3*R*,4*R*)-3,4-diamino pyrrolidinyl derived PNA monomers (**16** & **17**)

The synthesis of **16** and **17** were started from the naturally occurring L-tartaric acid which is commercially available and inexpensive. With its two chiral centres, L-tartaric acid has been a material of choice to access multifunctionalized pyrrolidine rings through their reactions with primary amines.<sup>7</sup>

The first step involved the condensation of L-tartaric acid with benzyl amine in xylene, using Dean-Stark apparatus and the reaction progress was monitored by the amount of water collected. A good yield of imide **1** was obtained when the theoretical amount of water was collected and the product was recrystallized from ethanol. This was reduced with LiAlH<sub>4</sub> to get the diol **2**. The reported literature methods such as NaBH<sub>4</sub>/I<sub>2</sub> in THF/diethylether and LiAlH<sub>4</sub> in diethylether<sup>7</sup> were used to reduce the imide **1** but gave low yield of diol **2**. Under refluxing THF

conditions, the two carbonyl groups were incompletely reduced to give mixture of alcohols which being polar could not be easily separated leading to a lower yield of desired product. Finally, the imide **1** was reduced completely to diol **2** by a slight modification of the work-up procedure. Instead of quenching the reaction mixture with aq. NaOH, it was quenched with saturated aqueous solution of sodium sulphate, followed by stirring the filtrate with diethylether to get pure diol **2**. The NMR data obtained was consistent with literature reports.<sup>8</sup>

The N<sup>1</sup>-benzyl (3,4)-diol **2** was hydrogenated in presence of Pd/C followed by *in-situ* protection with benzyloxy carbonyl to yield compound N-benzyloxy carbonyl 3,4-diol **3**. The mono protection of diol **3** with dimethoxytrityl group gave mono O-DMT protected compound **4**. The 4*R*-hydroxyl group of compound **4** was converted to the 4*S*-p-nitro benzoate ester **5**, under Mitsunobu conditions using p-nitrobenzoic acid. This 4*S*-benzoate ester **5** on reaction with sodium azide and methanol yielded the 4*S*-alcohol **6** that was converted to the corresponding O-mesyl derivative **7**. This on treatment with TFA gave the (3*S*,4*R*) O-mesylate alcohol **8** which was transformed to (3*R*,4*R*) azido alcohol **9**. The reduction of azide group in **9** followed by *in-situ* reaction with di-tert-butyl dicarbonate gave the N-boc protected (3*R*,4*R*) alcohol **10** that was mesylated to the O-mesyl derivative **11**. S<sub>N</sub>2 displacement of O-mesyl group in compound **11** by azide group yielded the compound **12** with (3*S*,4*R*) configuration. Upon reduction of azide in **12** followed by N-alkylation with ethylbromoacetate yielded compound **13**, which was N-acylated with chloroacetylchloride to obtain compound **14** with (3*S*,4*R*) configuration. N1-alkylation of thymine with compound **14** using K<sub>2</sub>CO<sub>3</sub> in DMF at 70°C gave the (3*S*,4*R*) ester **15**, which was hydrolysed with LiOH to obtain the monomer **16**. All new compounds were characterized by <sup>1</sup>H, <sup>13</sup>C, and mass spectral data.

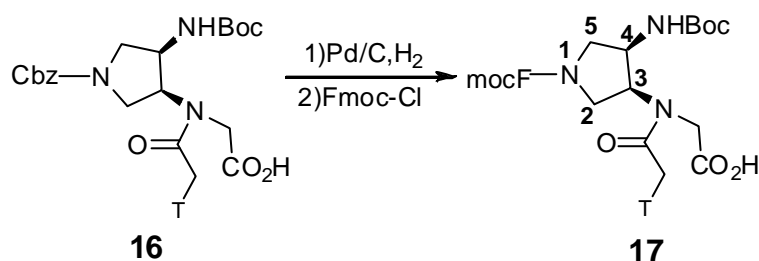


**Scheme 1: Synthesis of 3*S*,4*R* (3→4)-dap PNA Monomer (16)**

### Synthesis of 3*S*,4*R* (1→3)-dap PNA Monomer (17)

Compound **16** (3*S*, 4*R*) on hydrogenolysis with Pd/C and followed by treatment with Fmoc-Cl yielded the desired monomer **17** with (3*S*,4*R*) configuration which was used for the synthesis of PNA oligomers.

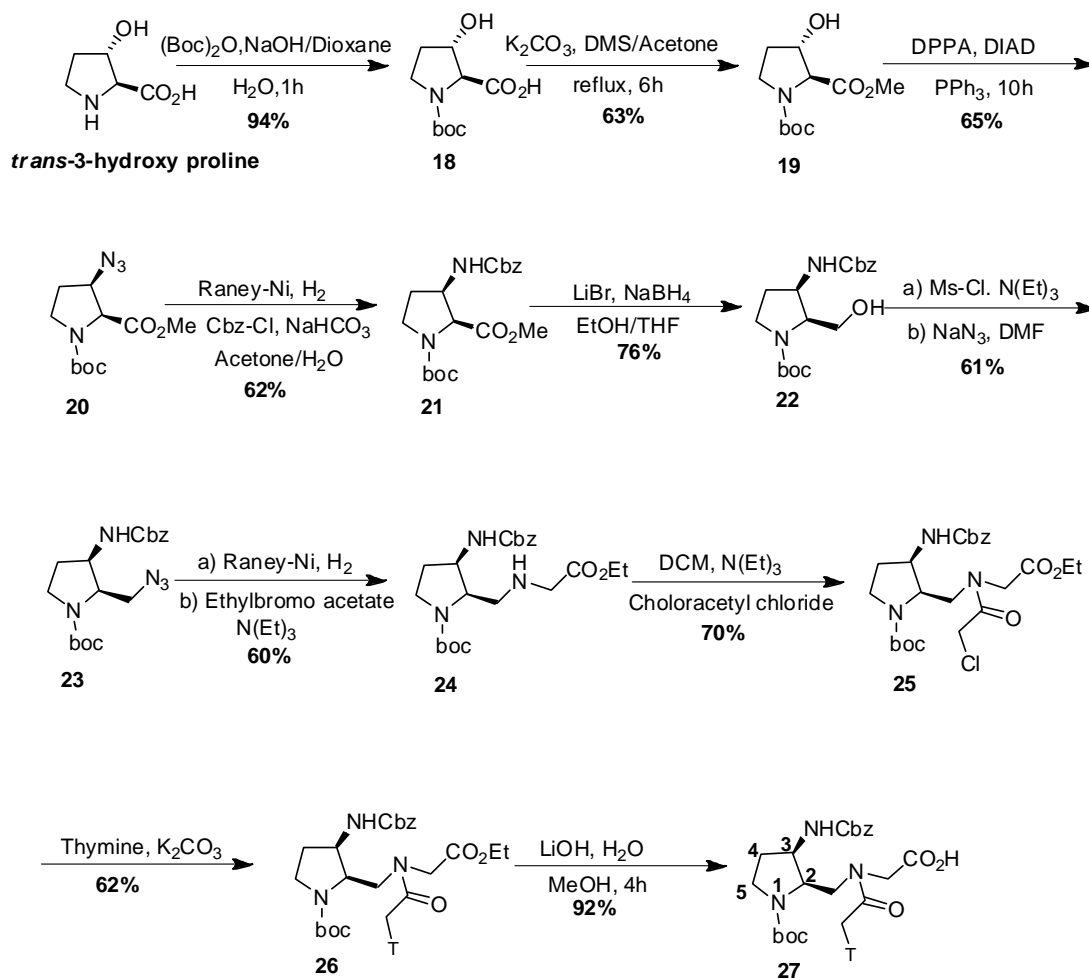




**Scheme 2:** Synthesis of 3*S*,4*R* (*1*→*3*)-*dap* PNA Monomer (**17**)

### Synthesis of 2*S*,3*R* (*1*→*2*)-*amp* PNA Monomer (**27**)

Synthesis of the PNA monomer **27** was started by protecting the ring nitrogen of 3*S*-hydroxy commercially available proline with *tert*-butyloxy carbonyl group to give N-boc protected 3*S*-hydroxy proline **18**. Esterification of the  $\alpha$ -carboxyl group N-boc 3*S*-hydroxy proline with dimethyl sulphate in presence of  $K_2CO_3$  yielded the 3*S*-hydroxy 2*S*-methyl ester **19**. This was followed by reaction of the 3*S*-hydroxyl ester of **19** with diphenyl phosphorylazide under Mitsunobu conditions to yield the 3*R*-azido 2*S*-methyl derivative **20**. The reduction of 3*R*-azido group in **20** with Raney-Ni, followed by *in-situ* protection of the resulting amine with benzyloxycarbonyl gave the fully protected product **21**. The 2*S*-methyl ester group in **21** was reduced to the corresponding alcohol using  $LiBH_4$  to yield the 2*S*-alcohol **22** that was O-mesylated using mesyl chloride followed by  $S_N2$  displacement of O-mesyl group in compound **22** by azide group using sodium azide gave the 3*R*-azide **23**. Upon reduction of 3*R*-azide group in **23** with Raney-Ni followed by N-alkylation with ethylbromoacetate in acetonitrile gave the N-alkylated (2*S*,3*R*) ester **24**, which was acylated with chloroacetylchloride to obtain the chloro methyl derivative **25**. This was used for N-alkylation of thymine to yield the ester **26** which was hydrolysed to the required monomer **27** with (2*S*,3*R*) configuration.

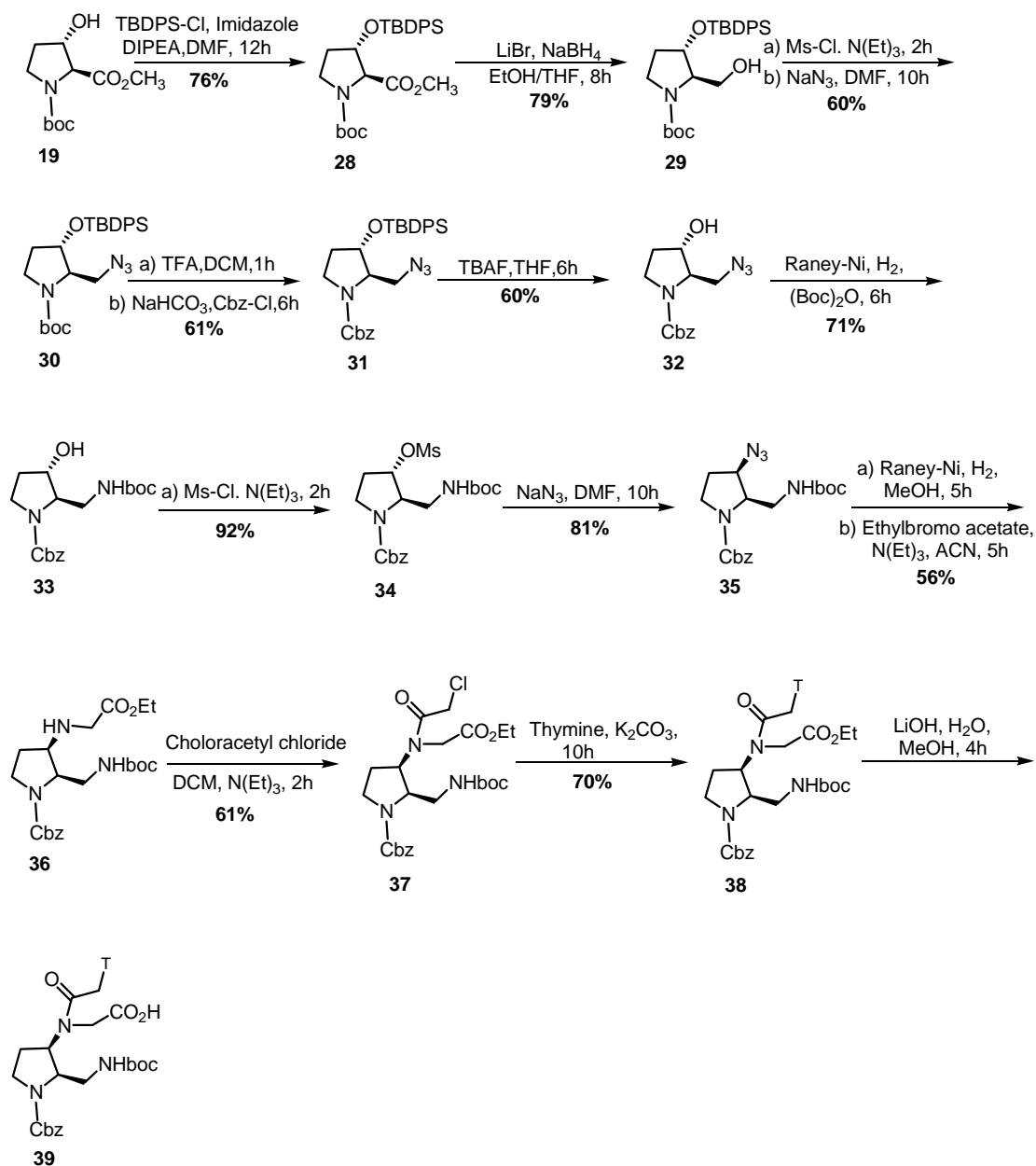


**Scheme 3:** Synthesis of *2S,3R* (*1*→*2*)*amp*PNA Monomer (**27**)

### Synthesis of (*2R,3R*) (*2*→*3*)*amp*PNA Monomer

The *3S*-hydroxyl group of compound **19** was protected as *t*-butyldiphenylsilyl ether and the ester group was reduced with LiBH<sub>4</sub> to obtain the corresponding alcohol **29**. Hydroxyl group of **29** was O-mesylated, followed by azidation yielded the *2R*-azide **30**. The N-*t*-Boc protecting group in **30** was exchanged to N-cbz group by treating **30** with trifluoroacetic acid followed by reaction with benzylchloroformate to yield **31** with (*2S,3S*) configuration. The reaction of **31** with TBAF gave the (*2R,3S*) azido alcohol **32**, that was hydrogenated in presence of Raney-Ni to amine which was *in-situ* protected by *t*-butyloxy carbonyl group with

di-*tert*-butyl dicarbonate yielded the **33** with (2*R*,3*S*) configuration. O-Mesylation of the secondary alcohol group in **33** gave the corresponding O-mesylate **34** in which S<sub>N</sub>2 displacement of mesyl group by azide group was done by reacting with NaN<sub>3</sub> to obtain **35** having (2*R*,3*R*) configuration.



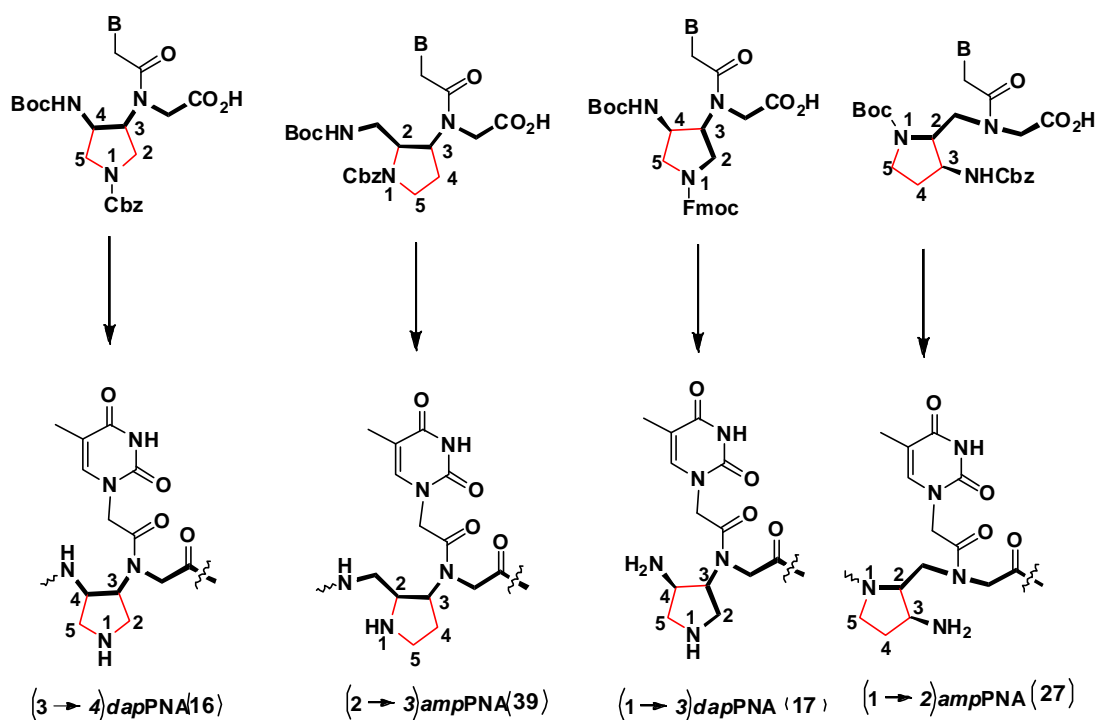
**Scheme 4:** Synthesis of (2*R*,3*R*) (2→3)*amp*PNA Monomer (**39**)

The reduction of 3*R*-azide to amine was followed by N-alkylation with ethylbromoacetate to yield compound **36** having (2*S*,3*R*) configuration that was acylated with chloroacetylchloride to give compound **37** having (2*R*,3*R*) configuration. The chloro compound was used for N-alkylation of thymine base to yield ester (2*R*,3*R*) **38** that was converted to the required monomer acid **39** by hydrolysis of ester using LiOH having (2*R*,3*R*) configuration.

### Chapter 3: Solid phase synthesis and biophysical studies of chiral, cationic *dap*PNA and *amp*PNA derived oligomers.

#### Solid Phase synthesis of *dap*PNA and *amp*PNA derived oligomers

The modified cationic pyrrolidine PNA monomers, **16**, **17**, **27** and **39** (Figure 4), synthesized as described in Chapter-2 were incorporated into standard PNA oligomers by using either Boc or Fmoc chemistry for solid phase synthesis leading to four different PNA backbones (Figure 4).



**Figure 4:** Backbones (shown by thick lines) derived from modified PNA monomers.

The extra primary or secondary amine functions of these oligomers are capable of accepting a proton and help to make PNA cationic in all the cases. The readily available MBHA resin and Rink Amide resin were chosen as the polymer matrix on which the *aeg*PNA oligomers as well as the modified PNA oligomers were assembled using Boc and Fmoc strategy for solid phase synthesis. The PNA oligomers 1–17 (Table 2) after cleavage from the resin were purified by HPLC on RPC-18 column and characterized by MALDI-TOF mass spectrometry. The complementary DNA oligomers were synthesized on an automated DNA synthesizer.

### Synthesis of complementary DNA oligonucleotides

The oligonucleotides DNA1 and DNA2 (Table 1) corresponding to complementary (parallel and antiparallel) PNA sequences were synthesized on Applied Biosystems ABI 3900 DNA Synthesizer using standard  $\beta$ -cyanoethyl phosphoramidite chemistry. The oligonucleotides were synthesized in the 3' to 5' direction on polystyrene solid support, followed by ammonia treatment. These were desalted by gel filtration, and their purity as ascertained by RP HPLC on a C-18 column was found to be more than 98%. They were used without further purification for the biophysical studies of hybridization with PNAs.

**Table-1:** DNA oligonucleotides used in the present work

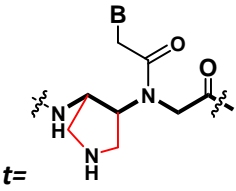
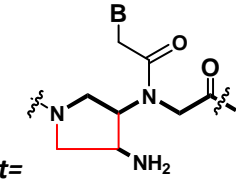
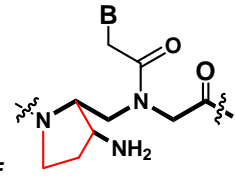
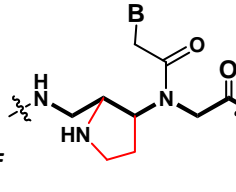
S. No.	Entry	Sequence (5'-3')	Type
1	DNA 1	AGTGATCTAC	Antiparallel
2	DNA 2	CATCTAGTGA	Parallel

### Biophysical Studies of modified PNA oligomers

The binding specificity of modified PNAs towards complementary DNA were investigated using the biophysical technique temperature dependent UV-spectroscopy (UV-T<sub>m</sub> studies). The UV-melting studies were carried out with modified PNA:DNA complexes and analyzed with respect to that of control *aeg*PNA:DNA complexes.

### Thermal stability of PNA:DNA duplexes (antiparallel/parallel)

Table 2: UV melting Temp( $^{\circ}$ C) of PNA:DNA duplexes

Modified unit	S.No	Sequence	UV-Tm	
			DNA 1(ap) ( $\Delta$ Tm <sub>1</sub> )	DNA 2(p) ( $\Delta$ Tm <sub>2</sub> )
Control	PNA 1	H-GTAGATCACT-Lys-NH <sub>2</sub>	52 $^{\circ}$ C	48 $^{\circ}$ C
	PNA 2	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	56 $^{\circ}$ C(+4 $^{\circ}$ C)	49.5 $^{\circ}$ C(+1.5 $^{\circ}$ C)
	PNA 3	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	45.5 $^{\circ}$ C(-0.5 $^{\circ}$ C)	42.5 $^{\circ}$ C(+5.5 $^{\circ}$ C)
	PNA 5	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	60 $^{\circ}$ C(+8 $^{\circ}$ C)	55.5 $^{\circ}$ C(+7.5 $^{\circ}$ C)
	PNA 4	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	53 $^{\circ}$ C(-3 $^{\circ}$ C)	45 $^{\circ}$ C(-3 $^{\circ}$ C)
	PNA 6	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	43 $^{\circ}$ C(-9 $^{\circ}$ C)	46.3 $^{\circ}$ C(-1.7 $^{\circ}$ C)
	PNA 7	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	48 $^{\circ}$ C(-4 $^{\circ}$ C)	52.6 $^{\circ}$ C(+4.6 $^{\circ}$ C)
	PNA 9	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	54 $^{\circ}$ C(+2 $^{\circ}$ C)	47.7 $^{\circ}$ C(-0.3 $^{\circ}$ C)
	PNA 8	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	55 $^{\circ}$ C(+3 $^{\circ}$ C)	49.3 $^{\circ}$ C(+1.3 $^{\circ}$ C)
	PNA 10	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	56 $^{\circ}$ C(+4 $^{\circ}$ C)	45 $^{\circ}$ C(+3 $^{\circ}$ C)
	PNA 11	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	48 $^{\circ}$ C(-4 $^{\circ}$ C)	54.3 $^{\circ}$ C(+6.3 $^{\circ}$ C)
	PNA 13	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	60 $^{\circ}$ C(+8 $^{\circ}$ C)	43.2 $^{\circ}$ C(-4.8 $^{\circ}$ C)
	PNA 12	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	50.5 $^{\circ}$ C(-0.5 $^{\circ}$ C)	50.8 $^{\circ}$ C(+2.8 $^{\circ}$ C)
	PNA 14	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	60.3 $^{\circ}$ C(+8.3 $^{\circ}$ C)	54 $^{\circ}$ C(+6 $^{\circ}$ C)
	PNA 15	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	56 $^{\circ}$ C(+4 $^{\circ}$ C)	58.3 $^{\circ}$ C(+10.3 $^{\circ}$ C)
	PNA 17	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	65.5 $^{\circ}$ C(13.5 $^{\circ}$ C)	47.5 $^{\circ}$ C(-0.5 $^{\circ}$ C)
	PNA 16	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	51 $^{\circ}$ C(-1 $^{\circ}$ C)	48.3 $^{\circ}$ C(+0.3 $^{\circ}$ C)

The mixed purine-pyrimidine PNA:DNA duplexes derived from modified PNAs incorporating one of the four monomers at different positions show stabilization in maximum cases (Table 2). The duplexes formed by (3→4)-dapPNA and (2→3)ampPNA were better stabilized than their analogues (1→3)-dapPNA and (1→2)ampPNA.

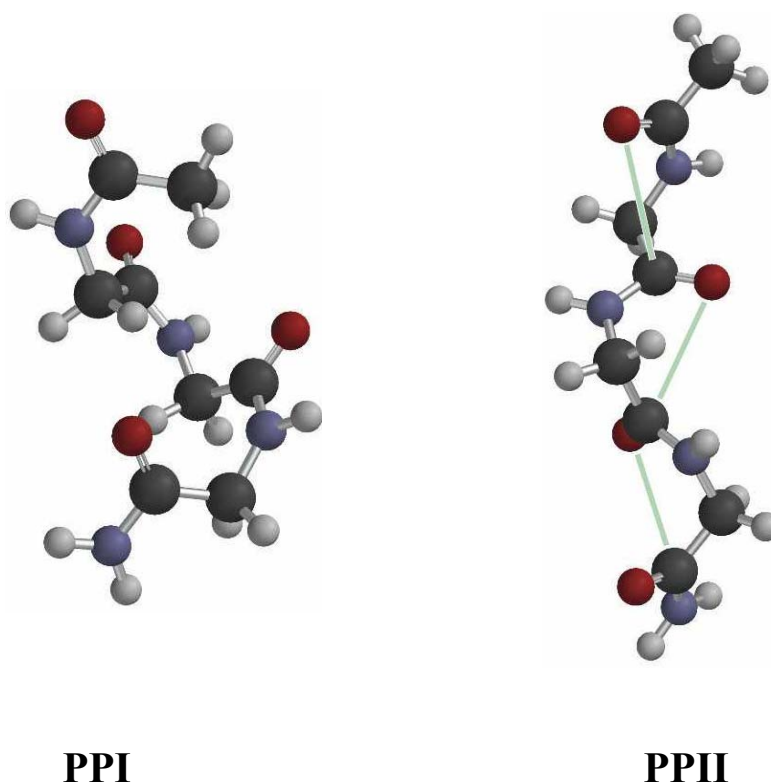
This differentiation in stability of PNA:DNA complexes of different PNA derived units, can help to fine tune the binding properties of PNA:DNA complexes for different applications. The modified PNA monomers presented in this and the preceding chapter constitute a new family of PNA monomers which can tune the binding properties of *aeg*PNAs.

## **Chapter 4: Synthesis and conformational studies of 3*S*-hydroxy substituted poly proline derivatives.**

### **Introduction:**

A protein is a polypeptide chain made up of amino acid residues linked together *via* peptide bond in a definite sequence. Amino acids are chiral (except for glycine, in which the normal asymmetric  $\alpha$ -carbon has two hydrogens), and the naturally occurring proteins contain L-amino acids. The sequence of side chains in protein determines all that is unique about a particular protein, including its biological function and its specific three-dimensional structure.

Repeating units of a single amino acid linked via amide bond (homo oligopeptide) provide model systems that adopt a definite secondary structure in different solutions. The poly(L-proline) may adopt two different conformational forms, known as polyproline I (PPI) and polyproline II (PPII) (Figure 5). Both secondary structures are helical and differ substantially in their physical and spectroscopic properties as well as in their crystallographic structures. PPI is a right-handed helix containing all peptide bonds in “*cis*” disposition and is compact, having 3.3 residues per turn with a helical pitch of  $5.6\text{\AA}/\text{turn}$ . The backbone dihedral angles of PPI are  $(\varphi, \psi, \omega) = (-75, 160, 0)$ . In comparison, the PPII helix is a left handed helix with “*trans*” peptide bonds and is extended with 3 residues per turn and have  $9.3\text{\AA}$  of helical pitch. The backbone dihedral angles of PPII are  $(\varphi, \psi, \omega) = (-75, 145, 180)$ . The kind of conformation PPI or PPII adopted by poly ( L-proline) depends upon the solvent. The analysis of the factors determining the conformation of polyprolines is of intrinsic physio-chemical interest, to understand the molecular basis of conformation in biological macromolecules.



**Figure 5:** Structure of PPI and PPII<sup>6</sup>

**Objective:**

Proline-rich peptide sequences are very important recognition elements that have a significant bias toward the polyproline type II (PPII) conformation. Each strand of collagen triplex with the typical [Pro-Hyp-Gly]<sub>n</sub> repeat unit adopts a PPII-like conformation. Interactions between side chain functional groups and the effect of solvation by water-backbone interactions have been examined as determinants of the PPII conformational stability. To understand the factors that determine the *cis/trans* conformer ratio and the development of concepts for tuning this equilibrium is therefore important and Proline derivatives with a substituent in the  $\gamma$ -position (C4) have proven useful in both respects. Conformational studies have shown that the nature of the substituent at C4 and the absolute configuration at this center critically influence both the pyrrolidine ring pucker and the *cis/trans* conformer ratio of the amide bond. Understanding the different factors that influence the PPII conformation is therefore an important goal. Herein we present conformational studies of polyproline derivatives with hydrogen-bond-donating



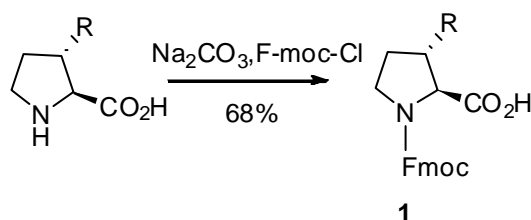
substituent in the  $\beta$ -position(C3) on proline ring. In case of “*cis*” 4*S*-substituted prolines, the intramolecular hydrogen bond between the 4*S*- substituent and the amide carbonyl on the same proline unit is known to influence the PPII conformation.

The specific objectives of this chapter are:

- i) Solid Phase synthesis of homo oligomeric peptides **1-3** (Table 3). Peptide **1** is synthesized using N-fmoc proline, whereas peptides **2&3** were synthesized using N-fmoc L-3-*trans* hydroxy proline.
- ii) Cleavage of the oligomers from the solid support
- iii) purification and characterization
- iv) conformational studies of these peptides as a function of pH, urea, and solvents (buffer and trifluoroethanol, TFE).

### Synthesis of fmoc 3-hydroxy proline **1**

A solution of commercially available *trans*-3-hydroxy proline (1.33 g, 4.62 mmol) in water/acetone (40 mL, 1:1) was cooled to 0°C and sodium bicarbonate (1.36 g, 16.2 mmol) was added.



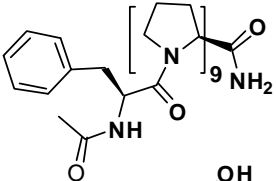
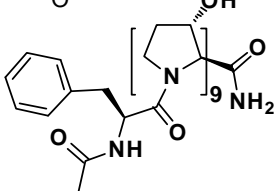
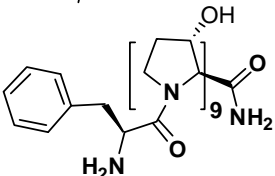
where R = H, OH

### Scheme 5: Synthesis of fmoc 3*S*-hydroxy proline **1**

To this mixture was added Fmoc-OSu (2.03 g, 6.0 mmol), and the reaction was stirred for 1 h at 0 °C and 2 h at 25 °C. The reaction was treated with 10% HCl to a pH of 4 and extracted with CH<sub>2</sub>Cl<sub>2</sub>. The organic layer was dried over anhydrous MgSO<sub>4</sub>, and the solvent was removed in vacuo to get N-Fmoc hydroxyl proline as white solid (yield 62%).

### Synthesized polyproline derivatives for conformational studies:

**Table 3:** Designed polyproline derivatives

Entry no.	Sequence
Peptide 1	
Peptide 2	
Peptide 3	

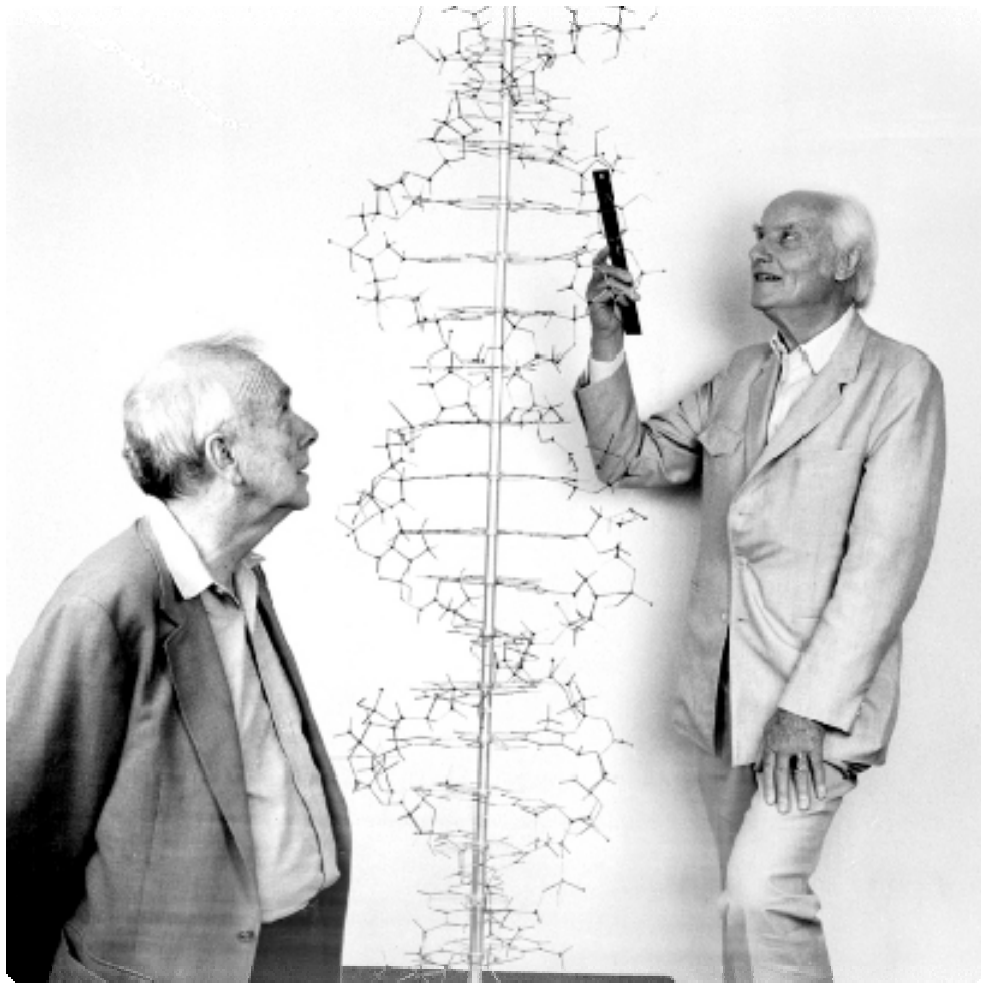
Conformational behaviour of these polyproline oligomers in various solvents, different percentages of n-propanol and the effect of urea were examined. The CD studies of these peptides in different solvents confirmed the presence of PPII conformation and the content of PPII is different in different solvents. With increase in the concentration of urea increase in the ellipticities is observed in the wavelength range of the positive band which is diagnostic of PPII helix content. This indicates that these peptides gain more PPII helix content as the urea concentration is increased. CD spectroscopic studies of oligoprolines in different ratios of phosphate buffer in n-propanol show that the PPII helix with all *trans* amide bonds is stabilized. All the three peptides **1-3** acquired PPII conformation in different solvents and comparative studies of PPII content of peptides were done. N-termini capped *trans*-L-3S-hydroxy proline (3S-Hyp)<sub>9</sub>-Phe-NH-Ac showed higher content of PPII when compared to its charged analogue (3S-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> in n-propanol and urea at any given concentration. All these results can be due to the change in charge dipole interaction which is arised due to N-termini capping.

Furthermore, A new family of  $\beta$  substituted polyprolines were synthesized which can adopt biologically important PPII conformation in different solvents. Further work is required to tune the conformation of polyproline and to know the different factors affecting the conformational behaviour of the peptides.

### **References:**

1. Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624.
2. a.) Nielsen, P. E.; Egholm, M.; Berg, R. M.; Buchardt, O. *Science*. **1991**, *254*, 1497.  
b.) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. M. *J. Am.Chem.Soc.* **1992**, *114*, 1895.
3. Govindaraju, T.; Kumar V. A.; Ganesh K. N. *Chem. Commun.*, **2004**, 860 – 861
4. a) Myers, M. C.; Witschi, M. A.; Larionova, N. V.; Franck, J. M.; Haynes, R. D.; Hara, T.; Grajkowski, A.; Appella, D. H. *Org. Lett.* **2003**, *5*, 2695-2698.
5. Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature*, **1998**, *392*, 666-667.
6. Jia-Cherng Horng, Ronald T. Raines; *Protein Science*,**2006**, *15*, 74–83.
7. (a) Rejman, D.; Kocalka, P.; Budesinsky, M.; Pohl, R.; Rosenberg, I.: *Tetrahedron*, **2007**, *63*, 1243. (b) Gonsalves, A. R.; Serra M. E. S.; Murtinho D.; Silva,V. F.; Beja, A. M.; Paixao, J. A.; Silva, M. R.; Veiga, L. A. *J. Mol. Cat. A: Chem.*, **2003**, *195*, 1.
8. Arakawa, Y; Yoshifuji, S. *Chem. pharm. Buli.* **1991**, *39*, 2219.

# Chapter 1: Introduction



## INTRODUCTION

### 1.1 Nucleic Acids: Chemical Structure and Importance

Nucleic acids are the most important class of biopolymers dominating the modern molecular science after the Watson-Crick discovery of the double helical DNA.<sup>1</sup> Their vital roles are fundamental for the storage and transmission of genetic information within the cells and contain all information required for transmission and execution of steps necessary to make proteins which are another important class of biopolymers and are important for cellular function. These biopolymers are composed of monomers called 'Nucleotides' which consists three components *i.e.* a nitrogenous heterocyclic base (a purine or a pyrimidine); a pentose sugar; and a phosphate group (Figure 1).

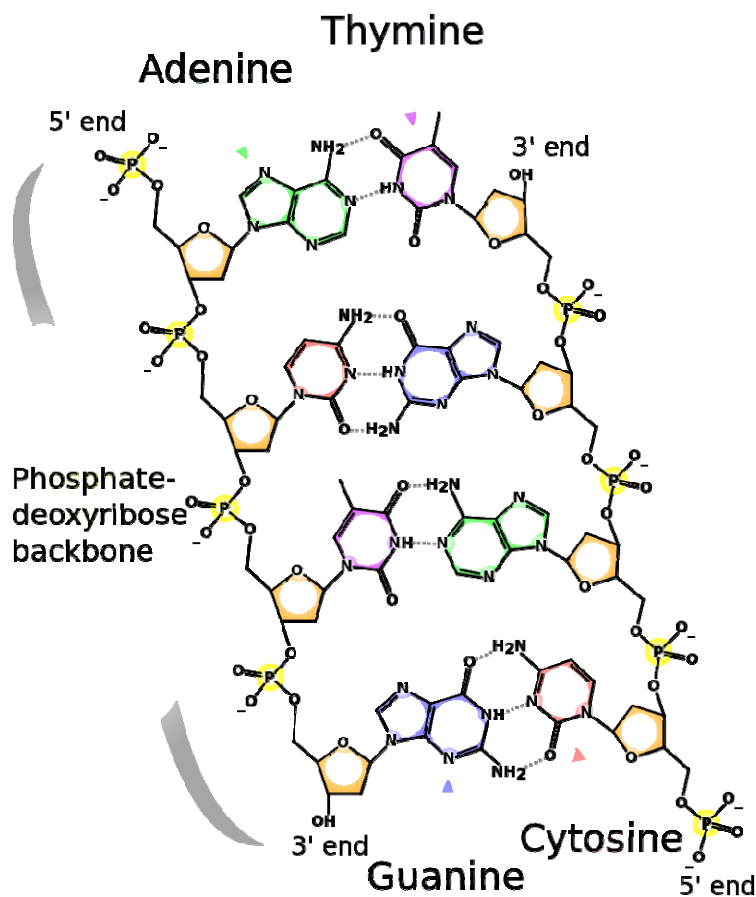
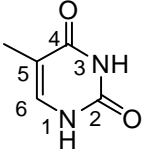
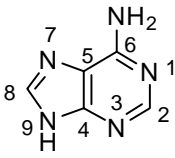
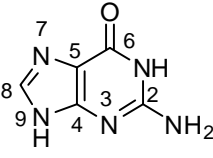
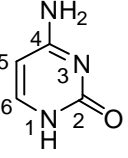
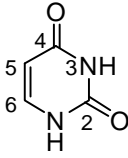
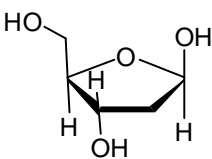
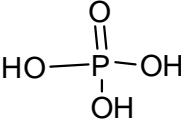
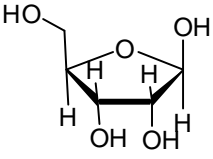


Figure 1: Model structure of DNA<sup>72</sup>

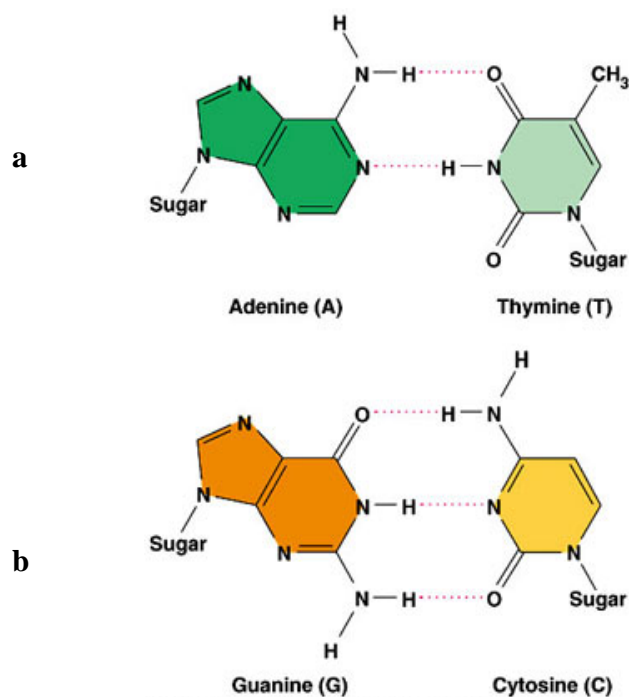
RNA consists ribonucleotides while DNA in contrast is made up of 2'-deoxyribonucleotides. The major bases are monocyclic pyrimidines or bicyclic purines. The major purines are adenine (A) and guanine (G) and are found in both DNA and RNA. The major pyrimidines are cytosine (C), thymine (T) found in DNA and uracil (U) in RNA instead of T. In nucleoside, the purine or pyrimidine base is joined through ring nitrogen to 1-<sup>1</sup>C carbon of a pentose sugar. In ribonucleic acid, the pentose is D-ribose which is locked into a five membered furanose ring by the bond from C-1' of sugar to N-1 of C or U or N-9 of A and G. This bond is on the same side of the sugar ring as the C-5 hydroxymethyl group and is defined as a  $\beta$ -glycosidic linkage. In DNA, the pentose sugar is 2'-deoxy-D-ribose and the four nucleic acids are deoxyadenosine, deoxyguanine, deoxycytidine, and deoxythymidine. In DNA, the methylated pyrimidine base, thymine replaces uracil in RNA (Figure 2).

	DNA only	DNA & RNA		RNA only	
Nucleo bases	 <p>Thymine</p>	 <p>Adenine</p>	 <p>Guanine</p>	 <p>Cytosine</p>	 <p>Uracil</p>
Sugar & phosphate	 <p>2-Deoxyribose</p>	 <p>Phosphate</p>		 <p>Ribose</p>	

**Figure 2:** Components of DNA & RNA

### 1.1.1 Hydrogen bonding between nucleobases

The chemical structure alone cannot be credited for the unique biological properties of nucleic acids. It was found in the first half of the 20<sup>th</sup> century that the amount of adenine in DNA is equal to the amount of thymine, and the amount of cytosine to that of guanine. This important finding made perfect sense after the proposed structure of DNA double helix by Watson and crick in 1953,<sup>1</sup> according to which DNA mainly exists as a double helix formed from two nucleic acid strands (Figure 1). The double helix is stabilized by hydrogen bonds between complementary nucleobase pair A:T and C:G positioned inside the helices and sugar phosphates on the outside. This ability of nucleic acids to form duplexes with their complementary counterparts is the key to the storage and replication of the genetic information.

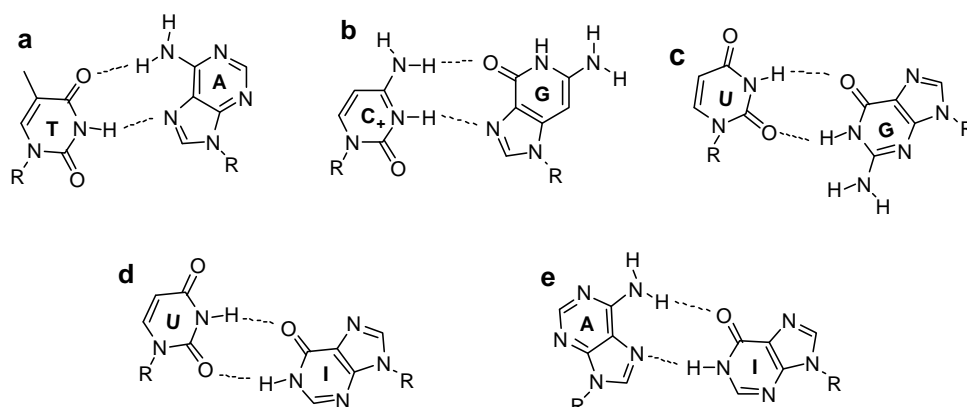


**Figure 3:** Watson-Crick base pairing scheme for (a) A-T base pair and (b) G-C base pair

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 bonding 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 3).<sup>1</sup> While Watson-Crick base pairing is the dominant pattern between the nucleobases, other significant pairings are Hoogsteen (HG)<sup>2a</sup> and Wobble base pairs.<sup>2b</sup>

Hoogsteen base pairing (Figure 4a and 4b) which is not isomorphous with Watson-Crick base pairing has importance in triple helix formation. In wobble base pairing, a single purine is able to recognize non-complementary pyrimidines (e.g. GU, where U = uracil, Figure 4c-e) and have importance in the interaction of messenger RNA (*m*-RNA) with transfer RNA (*t*-RNA) on the ribosome during protein synthesis (codon-anticodon interactions). Several mismatched base pairs and anomalous hydrogen bonding patterns have been seen in X-ray studies of synthetic oligodeoxynucleotides.<sup>3</sup>



**Figure 4:** Hoogsteen (a and b) and Wobble (c-e) hydrogen bonding scheme. I = Inosine

### 1.1.2 Oligonucleotides as Therapeutic Agents

Designing a small organic molecule as a drug against traditional drug target proteins requires structural knowledge of the binding site (target) and the binding forces. Since our understanding of protein folding is incomplete, this way of drug discovery process has serious limitations. In contrast, the nucleotide sequence in RNA and DNA is universal and the understanding of their structure is much better, from the drug-design



point of view nucleic acid targets are very appealing. In principle one can design drugs that, like nucleic acids, is repetitive in its primary structure and bind sequence specifically to these drug targets. In order for the sequence specific recognition to happen, the drug should contain nucleobases, which are fundamental units of nucleic acid recognition; in other words, a short piece of oligonucleotide itself can act as a drug. Two innovative strategies are being tested for inhibiting the production of disease related proteins using such sequence specific DNA fragments as gene expression inhibitors.

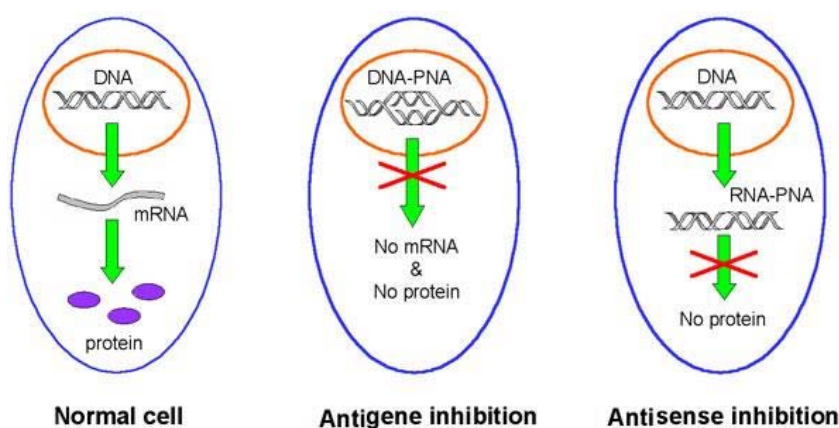
The triplex approach, also called antigene strategy, aims to stall the production of an unwanted protein by selectively inhibiting the transcription of corresponding gene. In this method, the oligonucleotides target the major groove of DNA where it winds around the double helical DNA to form a triplex (Figure 5).

The antisense strategy aims to selectively impede the translation process. The sequences of the bases along a messenger RNA molecule spell out the series of amino acids that must be strung together to make a protein. To obtain a molecule that bind to the sense strand, one must construct a string of nucleotides having the complementary “antisense” sequence. This upon binding to complementary region on *m*-RNA sterically inhibits the protein synthesis machinery.

### 1.1.3 Antisense Oligonucleotides

The concept behind the antisense therapeutics is simple and rational: to inhibit the expression of a specific gene at the *m*-RNA level by using complementary oligonucleotides, called antisense oligonucleotides, thereby blocking the expression of the protein encoded by the target RNA. In theory, hybridization of the antisense oligonucleotide to its complementary *m*-RNA by Watson-Crick base pairing should provide high specificity and affinity. The use of antisense oligonucleotides was originally postulated in the late 1960s by Grineva and co-workers.<sup>4</sup> In the 1970s, Miller and co-workers determined the chemical modification of the native molecule would result in the protection from *exo* and *endo* nucleases.<sup>5</sup> In 1978, Zamecnik and Stephenson first illustrated the idea of antisense therapeutics by demonstrating that antisense oligonucleotides could inhibit replication of Rous *sarcoma* virus in a cellular system.<sup>6,7</sup> Not much progress was made in this field at that time, primarily because the synthetic methodologies for obtaining relatively larger quantity of oligonucleotides were

unavailable. During past decade, numerous reports in the literature have demonstrated the ability of antisense oligonucleotides to block the function of specific genes. Antisense oligonucleotides are currently being investigated as therapeutic agents for the treatment of

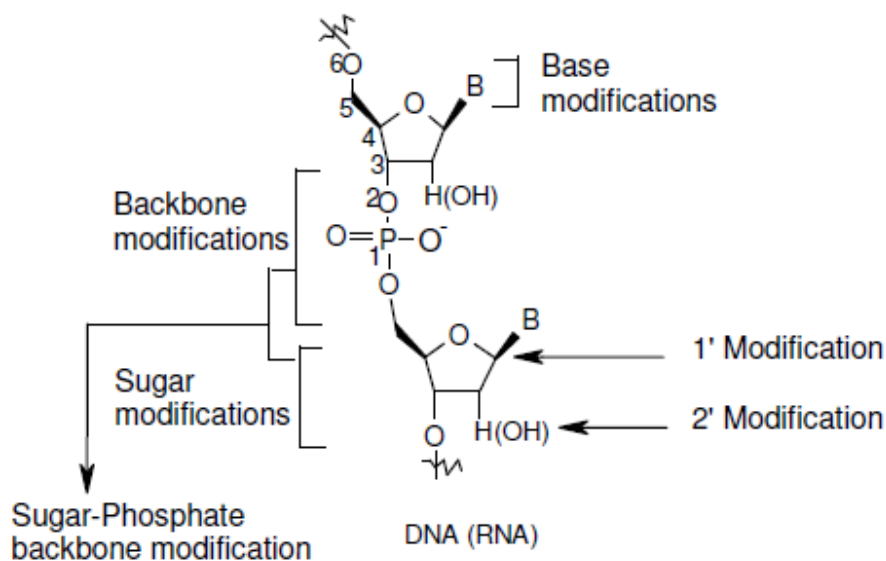


**Figure 5:** Mechanism of action of antisense and antigene oligonucleotides

viral infections, cancers and inflammatory diseases, and several human clinical trials of such are under way.<sup>8</sup> Vitravene developed by ISIS pharmaceuticals is the first and only antisense drug so far approved by the FDA for the treatment of cytomegalovirus (CMV) retinitis in people with AIDS.

#### 1.1.4 Modified oligonucleotides as antisense therapeutic agents

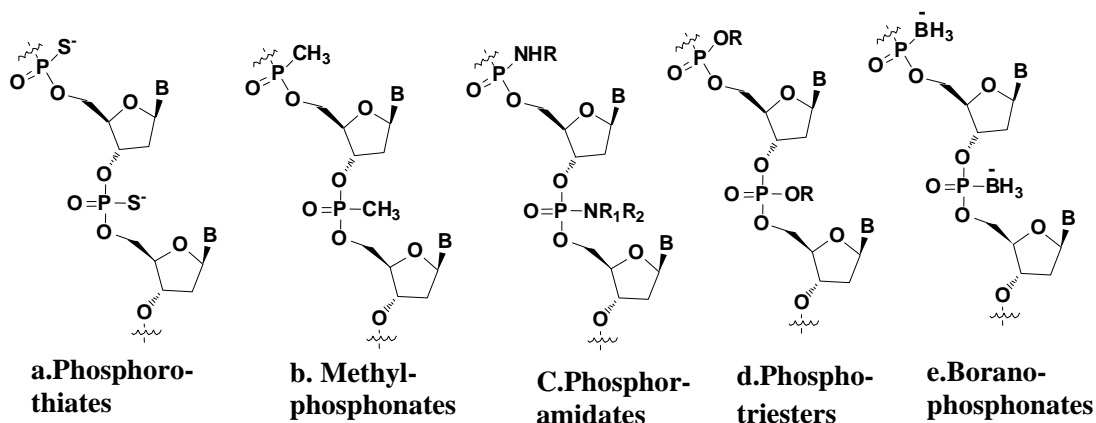
The unmodified antisense oligonucleotides are intended to enter the cell where they can pair with, and so inactivate, the complementary *m*-RNA sequences. But their inability to permeate cell membrane as they carry anionic charge results from repulsive interaction with the cell lipid layer. Further they are degraded by intracellular enzymes such as exonucleases. It became very clear that the nuclease susceptibility of natural oligonucleotides based on an unmodified (deoxy)-ribose phosphodiester (PO) backbone, would preclude profound antisense activity in biological systems at concentrations relevant in the context of potential therapeutic applications. This necessitated the chemical modification of oligonucleotides (Figure 6) and represents the key element in development of antisense therapeutics. The first generation modifications focused attention on oligonucleotides having phosphorothioates (PS, Figure 7a), phosphorodithioates, and phosphoramidates



**Figure 6:** Oligonucleotide modifications

(Figure 7c) replacing the anionic phosphate diester linkages and to 2'-OH protection of the ribonucleotides as second-generation antisense oligonucleotides.

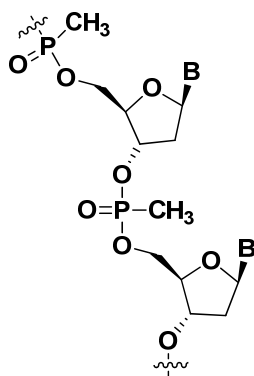
**1.1.4a Phosphorothioates (PS-oligos):** The PS-oligos are oligonucleotides in which one of the non-bridged oxygen was replaced by sulfur and the backbone thus retains the negative charge, but with reduced charge density compared to phosphodiester analogues (Figure 7a).<sup>9-10</sup> PS-oligos can be easily synthesized on a commercial DNA synthesizer, but the avidity of this analog for the complementary DNA was lower than natural oligonucleotide. The non-stereospecific synthesis of phosphorothioates results in the formation of a mixture of diastereomers and the conformational heterogeneity leads to lowering of the melting temperature. As a consequence, the search for a stereocontrolled synthesis has been undertaken and synthesized P-chiral analogues of PS-oligos were used for the biological studies of antisense targets inside living cells. PS-oligos are substrates to RNaseH and these first generation antisense agents have been extensively tested in human clinical trials against numerous targets. The only antisense agent approved by FDA (Vitravene) so far is based on PS-oligos. However, these oligos have tendency to induce non-specific effects,



**Figure 7:** First generation modified antisense oligonucleotide (replacement of 'O' in P-O)

through binding to cellular and extracellular proteins as well as cleavage of non-target *m*-RNAs that are only partially complementary by RNaseH. Replacement of both non-bridging phosphodiester oxygens leads to non-chiral phosphorodithioates. Compared with the corresponding phosphorothioates, the duplex stability of phosphorodithioate modified oligonucleotides with DNA decreased, and they exhibit less non-specific protein binding and more resistance to nuclease.

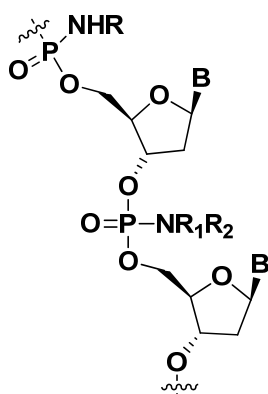
**1.1.4b Methylphosphonates:** Oligonucleoside methylphosphonates are a class of nonionic oligonucleotide analogs (Figure 7b). These oligomers contain deoxyribonucleoside or 2'-



**Figure 8:** Methylphosphonates

O-methylribonucleoside units that are linked in a 3'-5' manner via methylphosphonate groups. The methylphosphonate internucleotide linkage is similar in structure to the natural phosphodiester (PO) internucleotide bond<sup>11</sup>. One of the non esterified oxygens of PO group is replaced with a methyl group, and this substitution results in the loss of the negative charge. Inclusion of methylphosphonate linkage in the oligomer introduces a new center of asymmetry into the oligonucleotide as Rp/Sp configuration and consequently the oligomer preparation can contain up to  $2^n$  diastereomers, where  $n$  is the number of methylphosphonate linkages. Methylphosphonates are resistant to the action of cellular nucleases. The electro neutral methylphosphonate linkage would reduce charge repulsion between methylphosphonate oligos and complementary target nucleic acid and are more readily penetrate cell membranes, thus enhancing their internalization by cells<sup>12</sup>. Although oligodeoxyribonucleoside methylphosphonates form stable duplexes with single stranded DNA, hybrids formed with complementary RNA of the same nucleotide sequence generally have lower stability.<sup>13</sup> The reduction in melting value has been ascribed to the presence of mixture of diastereomers. In contrast to normal DNADNA/RNA duplex, the stability of methylphosphonate-DNA duplexes are not significantly affected by changes in ionic strength. In addition to the poor aqueous solubility, methylphosphonates cannot induce RNase H activity. For these reasons this modification is less useful as antisense molecule but can serve as the platform for the development of effective antisense oligonucleotides.

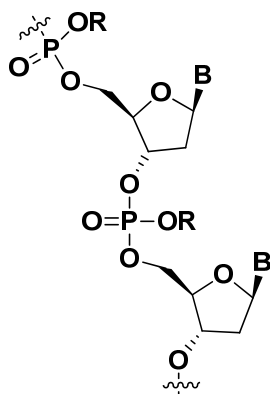
**1.1.4c Oligonucleotide Phosphoramidates:** Phosphoramidates (Figure 7c) are another well-studied class of backbone modifications.<sup>14</sup> They are resistant to hydrolysis by nucleases. With DNA targets, oligonucleotide phosphoramidates (Figure 9, R1 = H, R2 = -CH<sub>3</sub>, -CH<sub>2</sub>CH<sub>2</sub>OMe, (CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>O) exhibit rather poor hybridization characteristics ( $\Delta T_m$  = -0.1 to -2.3°C, at pH 7.2).<sup>15,16</sup> Oligonucleotide phosphoramidates (R1= H, R2 =-CH<sub>2</sub>CH<sub>2</sub>N-(CH<sub>2</sub>CH<sub>2</sub>)<sub>2</sub>O) form weak duplexes at neutral pH and more stable duplexes under acidic conditions (pH 5.6), due to protonation of the terminal amine.



**Figure 9:** Phosphoramidates

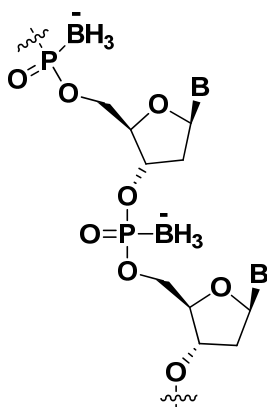
Oligophosphoramidate with ( $R_1 = \text{H}$ ,  $R_2 = \text{CH}_2\text{CH}_2\text{NMe}_2$ ), is however reported to hybridize to DNA targets at both neutral and acidic pH. Hybridization of oligonucleotides consisting of alternating phosphoramidite/phosphodiester linkage to RNA targets is similar to the corresponding interaction with wild type oligonucleotides.<sup>16</sup> The  $T_m$  of these complexes turned out to be independent of the ionic strength of the buffer, in contrast to the behavior of unmodified oligonucleotides with DNA or RNA targets, which shows a strong  $T_m$  increase, with increasing salt concentration.

**1.1.4d Oligo-Phosphotriesters:** Even though phosphotriesters are common intermediates in oligonucleotide synthesis, little data about their hybridization behavior are available. It has been shown that ethyl phosphotriester (Figure 14d,  $R = -\text{CH}_2\text{CH}_3$ ) modified oligonucleotides form substantially less stable self-complementary duplexes compared to wild type.<sup>17</sup> A single modification in a self complementary sequence (which therefore had a backbone modification in each strand of the duplex) resulted in a  $T_m$  decrease of  $-4^\circ\text{C}$  and  $-11^\circ\text{C}$ , depending on whether the Sp or the Rp diastereomer was used.



**Figure 10:** Phosphotriesters

**1.1.4e Oligo-boranophosphonates:** These are derived by replacing one of the non-bridging oxygen atoms in the phosphodiester group of DNA with borane ( $\text{BH}_3$ ).<sup>18</sup> The boron phosphate diester is isoelectronic with phosphodiesters, isosteric with methylphosphonate group and is chiral. These negatively charged oligos are highly water soluble, but more lipophilic than DNA.



**Figure 11:** Boranophosphonates

NMR and CD studies show that replacing the phosphodiester linkage in dinucleotides with the boranophosphodiester results in only a slight change in configurational characteristics,

such as sugar pucker, acyclic torsional angles and base stacking.<sup>19</sup> Boranophosphate DNA is considerably more stable to various nuclease enzymes than native DNA and overall more stable than phosphorothioate DNA. The discovery that oligonucleotide boranophosphodiester can activate *E.coli* RNase H and induce cleavage of RNA is encouraging.<sup>20</sup>

### 1.1.5 Sugar Modifications

**1.1.5a 2'-Modifications:** An important class of second-generation antisense oligonucleotide analogs is 2'-modified oligonucleotides. 2'-O-alkyl substituents (**1-4**, Figure 12) enhance the RNA binding affinity of the corresponding oligonucleotides<sup>21</sup> and also confer a significant degree of nuclease protection to adjacent PO linkages.<sup>22</sup>

The 2'-O-methyl ribose (**1**, Figure 12) enhances binding to complementary ODNs and confers stability to ssRNA from RNases, but is susceptible to degradation by DNases. It is also clear from the studies that larger 2'-alkyl groups lead to greatly enhanced nuclease stability, while RNA binding affinity gradually decreases with increasing size of the 2'-alkoxy substituents.<sup>23</sup> In contrast, ethylene glycol based 2'-O-substituents (**5-7**, Figure 12) promote RNA binding to the same extent as a simple 2'-O-methyl group ( $\sim 1.0^\circ\text{C}/\text{modification}$ ) and provides much higher degree of nuclease protection. Other 2'-modifications comprise the 2'-fluoro substituents (**8**, Figure 12) which has shown to have pronounced favorable effect on RNA binding affinity ( $1.5^\circ\text{C}/\text{modification}$ ) among all 2'-

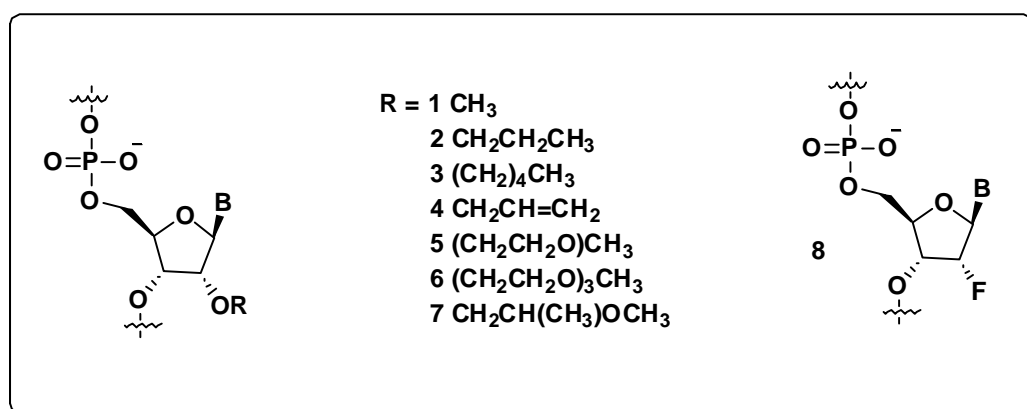


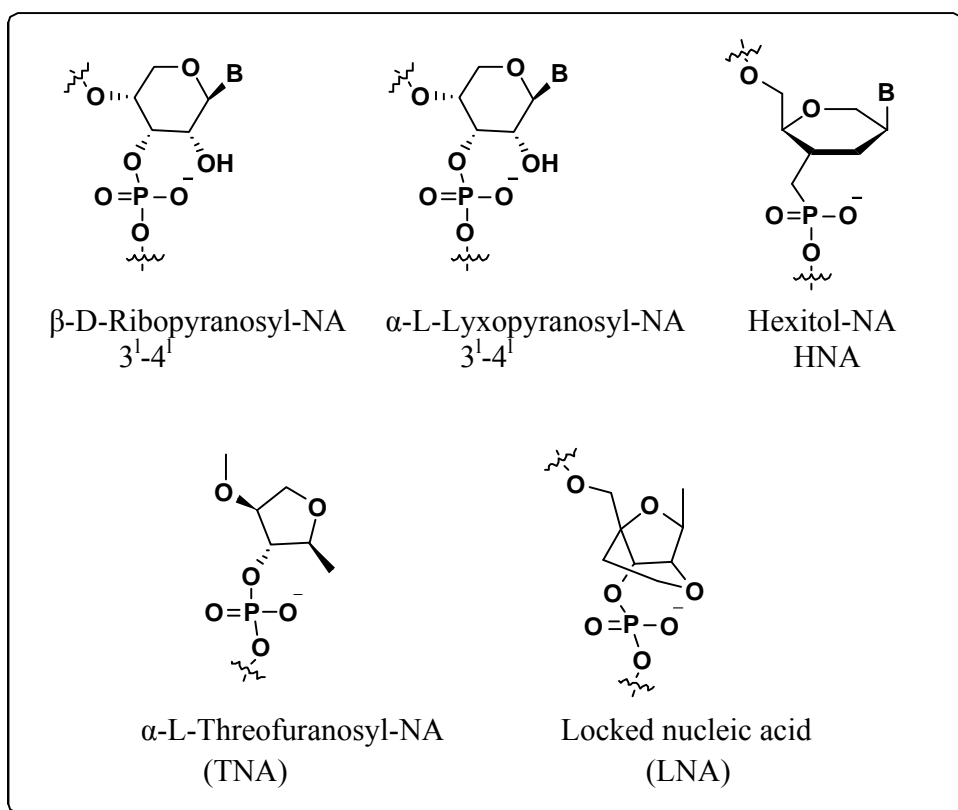
Figure 12 : 2'-Modified oligonucleotides



substituents investigated to date. This phenomenon is generally attributed to a strong preference of the modified sugar moieties for RNA-like-3'-*endo* conformation.<sup>24</sup>

**1.1.5a.1 Pentapyranosyl-NA and threofuranosyl-NA (TNA):** Pentapyranosyl-nucleic acid system was introduced by Eschenmoser group as depicted in the Figure 13. The 3' to 4' ribopyranosyl-series, having the vicinal phosphodiester groups arranged in 3'-axial 4'-equatorial position, does not produce a competent oligonucleotide pairing system. The 3' to 4' lyxopyranosyl oligonucleotides do not self-pair in the form of Watson-Crick duplexes, as well as triplexes in the case of homoadenine and homothymidine strands.<sup>25</sup> The thermal stabilities of these duplexes are typically weaker and more sequence dependent. These findings lead subsequently to the synthesis and evaluation of TNA (Figure 13), in which the diaxiality of the phosphodiester substituents is maintained, while the ring structure has contracted from pyranose to furanose. TNA is an extraordinary oligonucleotidic system, in that it does not only self-pair into antiparallel duplexes of the W-C type with stabilities similar to RNA, but in that it also efficiently cross pairs with target DNA and RNA.<sup>26</sup> As a conformationally restricted oligonucleotide 35 analog, Herdewijn group has developed 1,5-anhydrohexitol-NA (HNA).<sup>27</sup> Not only does HNA form self paired duplexes, but it also cross pairs with natural DNA and RNA. Typically, HNA-DNA duplexes show increased thermal stability ( $\Delta T_m/\text{mod} = +1.3^\circ\text{C}$ ) with DNA as complement and by ( $\Delta T_m/\text{mod} = +3^\circ\text{C}$ ) with RNA as complement.

**1.1.5a.2 Locked Nucleic Acids (LNA):** Another conformationally constrained analog is locked nucleic acids (LNA, Figure 13), introduced in 1998 by Imanishi<sup>28</sup> and Wengel group.<sup>29</sup> Structurally, this modification corresponds to a ribonucleoside in which the 2'-oxygen and the 4'-carbon atoms are linked via methylene unit. This modification locks the ribofuranose subunit into the 3'-*endo* conformation of nucleic acid duplex. Single and multiple LNA substitutions within a mixed sequence DNA oligonucleotide stabilize duplex formation with a complementary RNA typically by  $\Delta T_m/\text{mod} = +4$  to  $8^\circ\text{C}$ , depending on the sequence and the number of LNA units used.

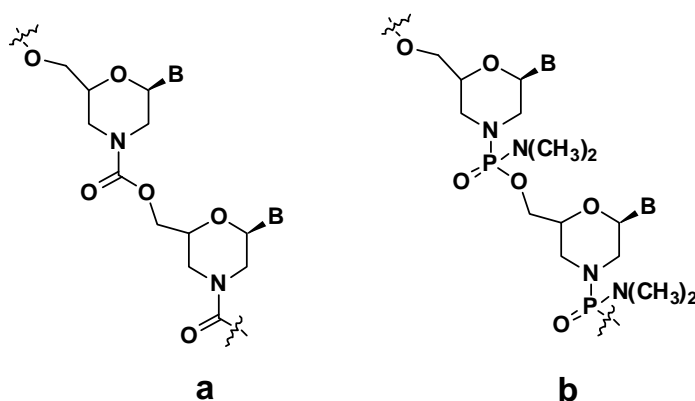


**Figure 13:** Sugar backbone modifications in oligonucleotides

### 1.1.6 Sugar-Phosphate modified oligonucleotide

Morpholinonucleotides and peptide nucleic acids are the most prominent modified nucleic acid mimics that involve the replacement of the sugar-phosphate backbone.

**1.1.6a Morpholino nucleotides:** Summerton *et al.* prepared novel oligonucleotide analogues from ribonucleotide derived morpholine units, linked by carbamate groups (Figure 14a).<sup>30</sup> Cytosine hexamer with stereoregular backbone was prepared in solution phase and was shown to bind poly dG with very high affinity. Solubility characteristics of the resulting oligomer have been improved by terminal conjugation with polyethylene glycol. Fully modified morpholino oligomers where phosphorodiamidate groups are shown to be more effective antisense agents than iso-sequential PS-oligos in cell-free systems and in various cultured cells (Figure 14b). This is attributed to the fact that morpholino oligos,

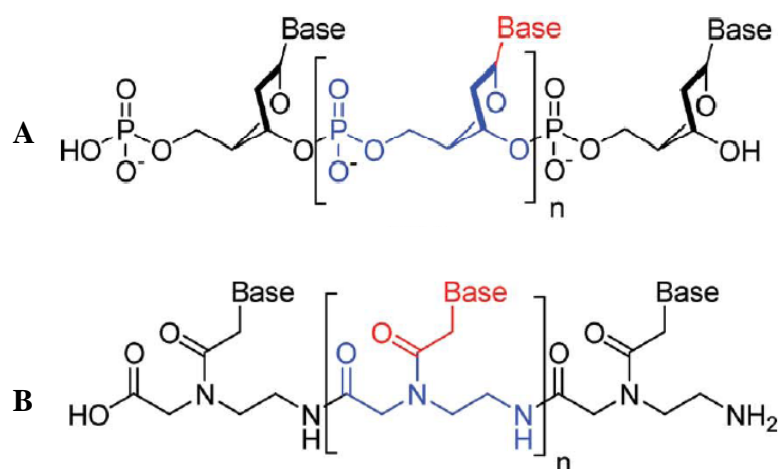


**Figure 14:** Morpholino ODNs linked by **a)** carbamate, **b)** phosphorodiamidate

unlike PS-oligos, do not bind to non-specific proteins and are much more sequence specific in their binding to DNA. In one notable study, morpholino oligos were shown to be more effective than PS-oligos as sequence specific antisense inhibitors of tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) in mouse macrophages, despite poor uptake into these cells.<sup>31</sup>

### 1.1.6b Peptide Nucleic Acids

Peptide nucleic acid (PNA), a DNA mimic resulting from the sugar-phosphate backbone replacement by a pseudopeptide backbone, was introduced by Nielsen *et al.* in 1991.<sup>32</sup> The structure of PNA is remarkably simple consisting of repeating *N*-(1-aminoethyl)-glycine units linked by amide bonds (Figure 15). The purine (A, G) and pyrimidines (C, T) bases are attached to the backbone through methylene carbonyl linkages and hence PNA is hybrid of peptides and nucleic acids in a rare structural combination. PNAs do not contain any sugar moieties or phosphate groups. It was therefore a surprise that PNA in many respects mimicked the behavior of DNA, and in some applications demonstrated superior properties. Homopyrimidine PNAs bind to complementary DNA and RNA to form triplexes with slightly increased affinity to RNA over DNA, while PNAs with both pyrimidine and purine bases form very stable duplexes with DNA/RNA, with equal affinity but more than that of either DNA-DNA or DNARNAs complexes.

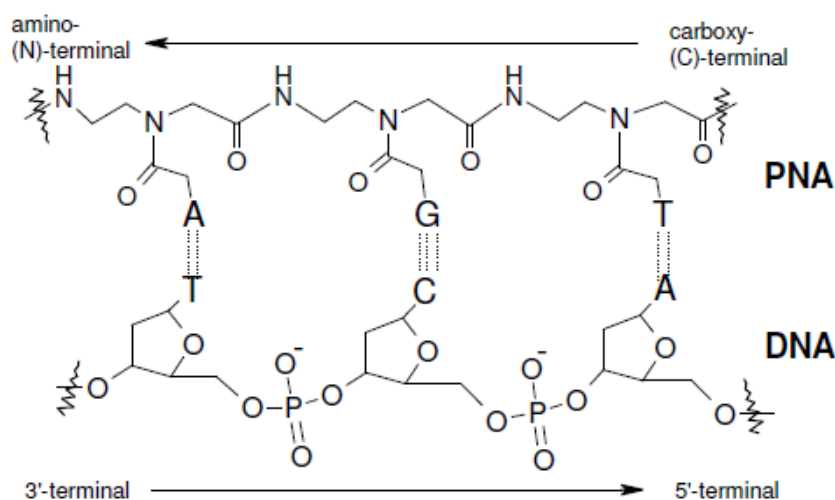


**Figure 15 :** Structure of DNA (A) and PNA (B)

Interestingly PNA can bind to DNA duplex to form triplex or duplex by displacing one of the DNA strand having same sequence of bases. 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 which always prefers antiparallel orientation. However, the antiparallel orientation is strongly preferred, and the parallel duplex has been shown to have different structure.<sup>33</sup> Because of very favorable hybridization properties<sup>33,34</sup> and high chemical and bio-stability,<sup>35</sup> it was regarded as a very promising lead for developing efficient antisense agents and medical drugs.<sup>36</sup>

## 1.2 PNA-DNA complex formation and their structures

**1.2.1 Duplex formation with complementary DNA and RNA:** PNAs obey Watson-Crick rules of hybridization with complementary DNA and RNA. Antiparallel PNA-DNA hybrids are considerably more stable than the corresponding DNA-DNA complexes.<sup>33</sup> The increased stability results in an increase in  $T_m$  of approximately 1°C/base. Antiparallel PNA-RNA duplexes are even more stable compared to DNA-RNA hybrids, and PNA-DNA duplexes.<sup>34</sup> The stability of parallel PNA-DNA and PNA-RNA duplexes is almost

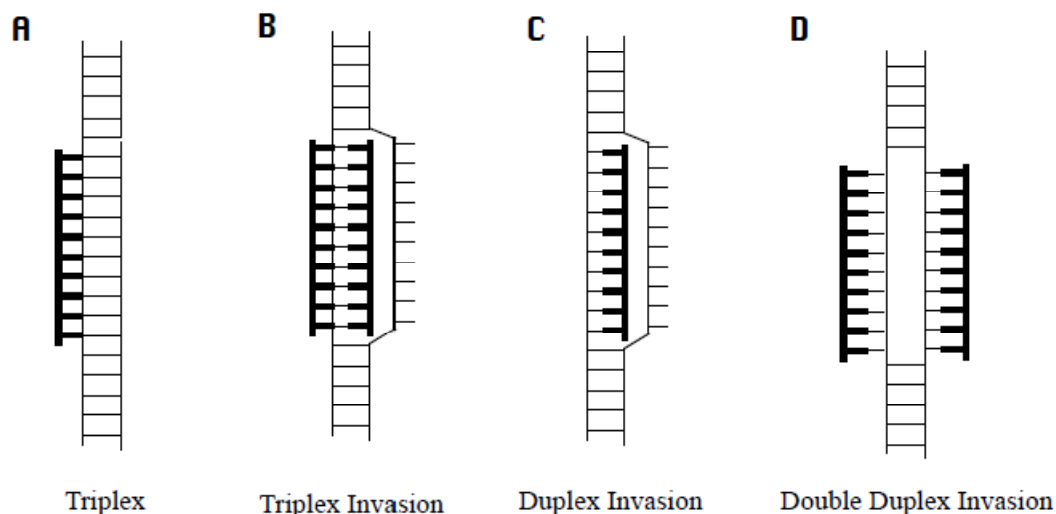


**Figure 16:** An antiparallel PNA:DNA duplex showing chemical structure of PNA (upper strand) as compare to that DNA (lower strand)

exactly the same as that of (antiparallel) DNA-DNA and DNA-RNA duplexes respectively. An interesting aspect of PNA-DNA duplex formation is that the  $T_m$  decreases with increase in salt concentration (ionic strength) which is contrast to that of DNA-DNA duplex, for which increase in  $T_m$  with salt concentration observed.<sup>37</sup> Base pair mismatches result in a reduction of the  $T_m$  value of 8-20°C. This discrimination is, in some cases, approximately double that observed for DNA-DNA duplexes.

**1.2.2 Triplex formation:** Homopyrimidine PNAs and PNAs with high pyrimidine/purine ratio bind to target DNA normally by formation of unusually stable PNA<sub>2</sub>-DNA triplexes.<sup>33</sup> However, in case of C-rich PNAs and GC-rich DNA duplexes, PNA-DNA<sub>2</sub> triplexes are observed. Base pairing mismatches result in a drop in melting temperature of 14-25°C.<sup>38</sup> The sequence specificity of triplex 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. In contrast, the analogous PNA<sub>2</sub>-DNA triplexes are not formed by homopurine PNA.<sup>39</sup> PNAs also form stable PNA<sub>2</sub>-RNA triplexes with RNA.<sup>40</sup>

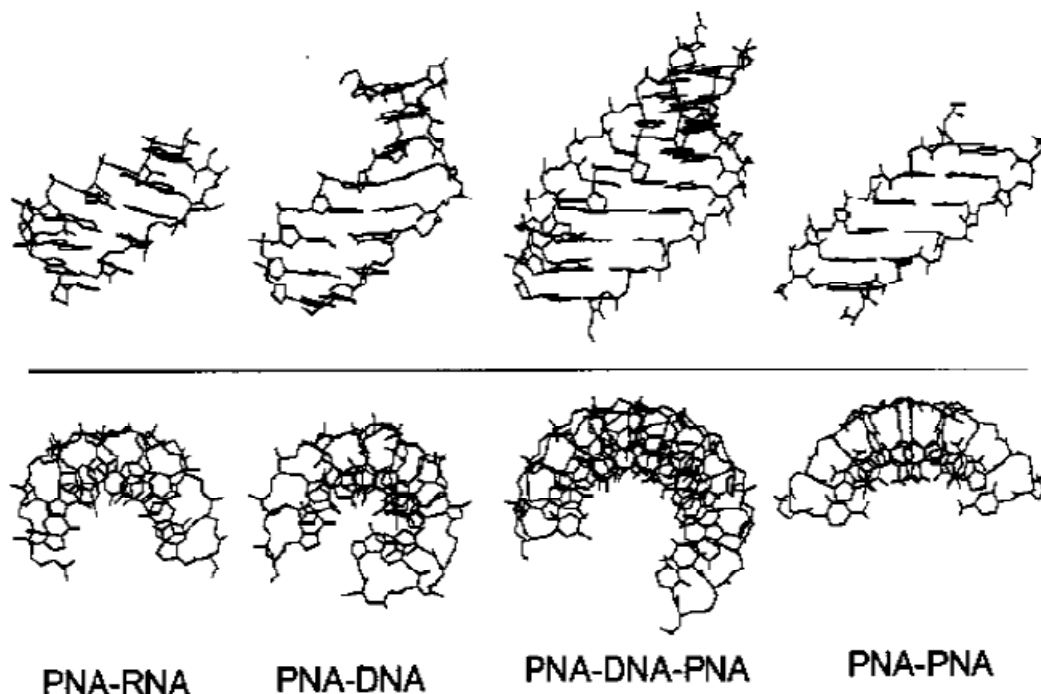
**1.2.3 PNA targeting double stranded DNA:** Homopyrimidine PNAs complex with target DNA duplexes to form triplexes. An interesting and peculiar property of PNA is that, it can bind to duplex DNA to form triplex by displacing one of the strands of DNA duplex. Other



**Figure 17:** Schematic representation of PNA binding modes for targeting double stranded DNA. Thick structures signify PNA.<sup>77</sup>

binding modes for PNA have been demonstrated (Figure 17), where it can bind to DNA duplex through unusual mechanism called strand invasion. Strand invasion based on PNA-DNA formation (Figure 17c) appears to be limited to PNAs that form extremely stable PNA-DNA duplexes, such as very purine rich PNAs.<sup>41</sup> Classical triplex formation with a single PNA Hoogsteen strand has been observed as a kinetic intermediate.<sup>42</sup> However, PNA:DNA<sub>2</sub> triplexes are much less stable than the corresponding triplex invasion complexes (Figure 17b). Double duplex invasion complexes (Figure 17d) can also be formed using pseudo-complementary PNA containing diaminopurine-thiouracil ‘base pairs’ that sterically destabilize the competing PNA-PNA duplex.<sup>43</sup>

**1.2.4 Structure of PNA-DNA duplexes:** Detailed structural information has been obtained from the NMR spectroscopic study of two antiparallel PNA-DNA duplexes (Figure 18).<sup>44</sup>



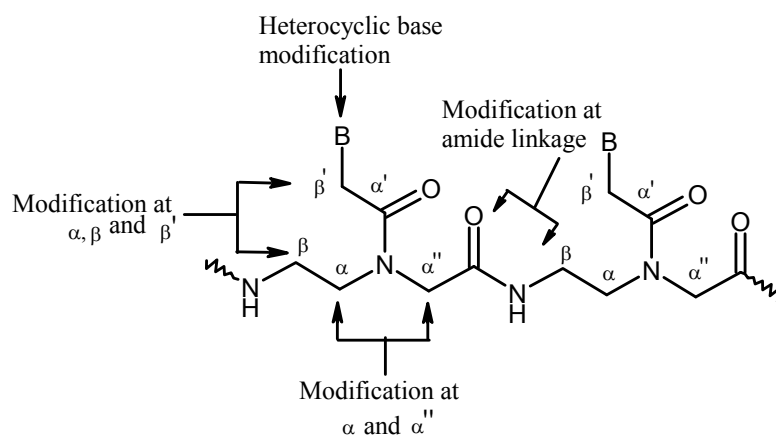
**Figure 18:** Structures of various PNA complexes shown in side view (upper panel) or top view (lower panel).<sup>78</sup>

The DNA strand is in a conformation similar to the B-form, with a glycosidic *anti*-conformation, and the deoxyribose in C2'-*endo* form. A more recent NMR study<sup>45</sup> showed that an octameric antiparallel PNA-DNA duplex contained elements of both A-form and B-form. The primary amide bonds of the backbone are in *trans* conformation and the carbonyl oxygen atoms of the backbone-nucleobase linker point towards the carboxy-terminus of the PNA strand. The CD spectra of antiparallel PNA-DNA complexes are similar to DNA-DNA spectra and indicate the formation of right handed helix.<sup>44</sup>

### 1.3 Chemical Modifications of PNA

The structure of classical PNA monomer has been subjected to a variety of rational modifications with the aim of understanding the structure activity relations in this class of DNA mimics as well as obtaining PNA oligomers with specifically improved properties for various applications within medicine, diagnostics, molecular biology etc. (Figure 19).

The limitations of PNA include low aqueous solubility, ambiguity in DNA binding orientation and poor membrane permeability. Structurally, the analogues can be derived from ethylenediamine or glycine sector of the monomer, linker to the nucleobase, the nucleobase itself or a combination of the above. The strategic rationale behind the modifications<sup>46</sup> are (i) introduction of chirality into the achiral PNA backbone to influence the orientation selectivity in complementary DNA binding, (ii) rigidification of PNA backbone via conformational constraint to pre-organize the PNA structure and to entropically drive the duplex formation, (iii) introduction of cationic functional groups directly in the PNA backbone, in a side chain substitution or at the N or C terminus of PNA, (iv) to modulate nucleobase pairing either by modification of the linker or the nucleobase itself (v) conjugation with ‘transfer’ molecules for effective penetration into cells. In addition to improving the PNA structure for therapeutics, several modifications are directed towards their applications in diagnostics.



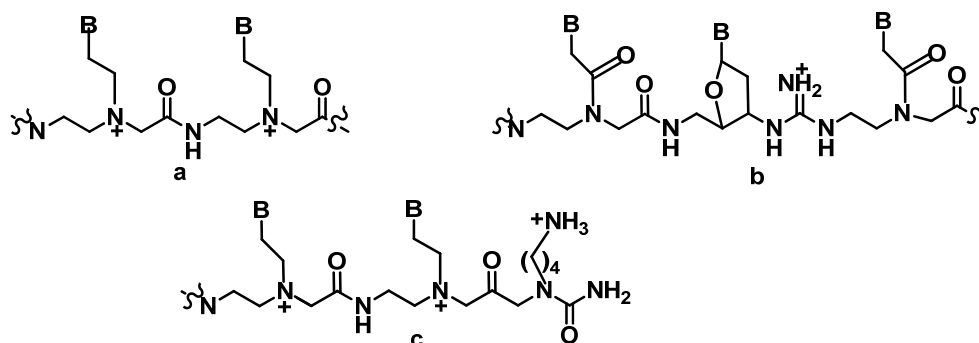
**Figure 19:** Structurally possible various PNA modifications

### 1.3.1 Introduction of ionic functional groups into the PNA

Positive charges were integrated into the PNA by replacing the acetamide linker with a flexible ethylene linker<sup>47</sup> or by the attachment of terminal lysine residues<sup>48</sup> (Figure 20c). Recently, a novel class of cationic PNA (Figure 20b) (*DNG*-PNA) analogs has been reported.<sup>49</sup> In these, alternating PNA /*DNG* chimeras, the O-(PO<sub>2</sub>)-O- linkage of

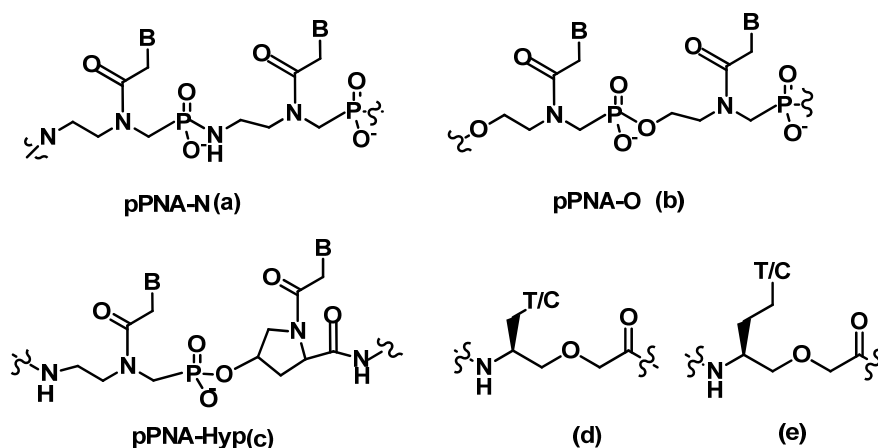


nucleotide was replaced by strongly cationic guanidino  $[N-C(=N^+H)-N]$  function. These analogs with neutral and positive linker showed high binding affinity with DNA/RNA targets.



**Figure 20:** Positively charged PNAs a) Flexible ethylene linker, b) Guanidium linkages, c) Lysine residues

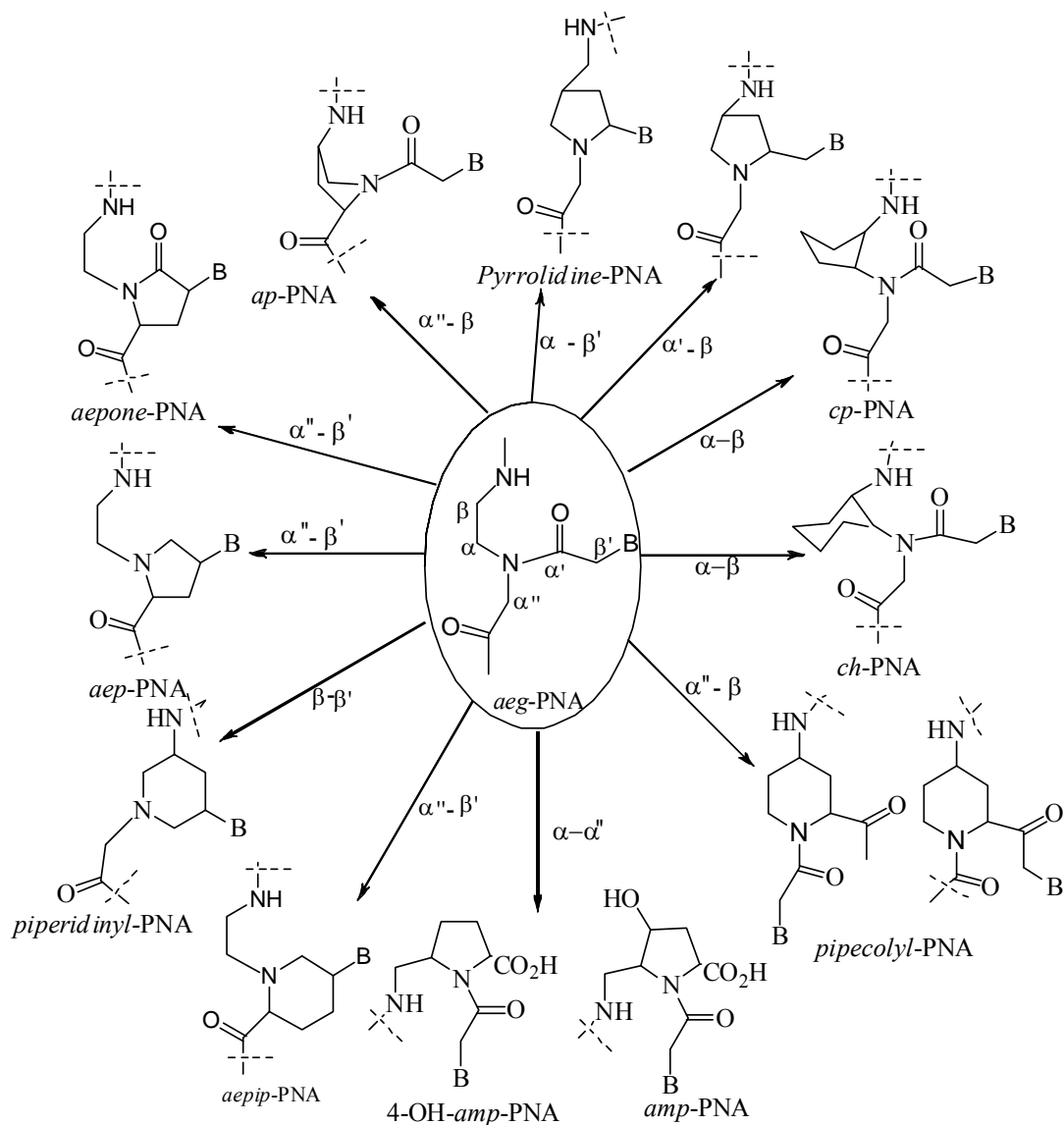
Introduction of negative charges in the PNA backbone (Figure 20 a-c) improved aqueous solubility and showed good binding with both DNA and RNA. These modifications have been well studied and reviewed.<sup>50,51</sup> However, many of these modified complexes were found to have less thermal stability compared to the unmodified PNA sequences. Ether linked PNAs (Figure 21d & e) showed co-operative binding with complementary antiparallel RNA in a sequence specific manner.<sup>52,53</sup>



**Figure 21:** Anionic PNAs (a-c)<sup>73,74</sup> and oxy PNAs (d-e)<sup>75,76</sup>

### 1.3.2 Construction of bridged structures

The *cis* and *trans* rotamers arising from the tertiary amide linkage in each PNA unit lead to a variety of rotameric conformations that cause different PNA:DNA/RNA hybridization kinetics in parallel and antiparallel hybrids.

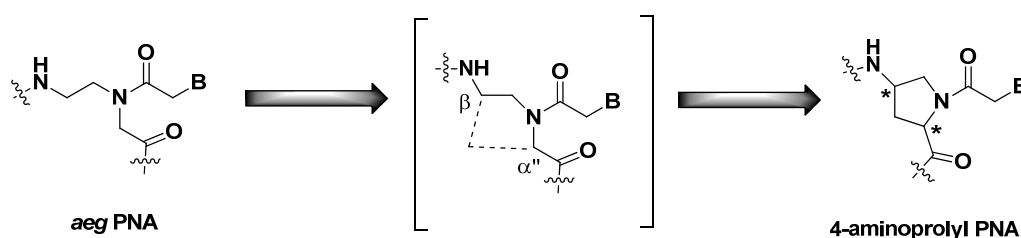


**Figure 22:** Conformationally constrained PNA analogues<sup>54b</sup>

The high rotational energy barrier between *cis* and *trans* rotameric populations makes these rotamers non-interconvertible. Any favorable structural preorganization of PNA may activate a shift in equilibrium towards the preferred complex formation because of the reduced entropy loss upon complex formation, provided that the enthalpic contributions suitably compensate. This may be achieved if the conformational freedom in *aeg*-PNA is reduced by bridging the aminoethyl/glycyl acetyl linker arms to give rise to cyclic analogs with pre-organized structure. Such modifications also restrain the flexible domains of the *aeg*-PNA (glycyl and ethylene diamine) in addition to introducing the chirality into PNA monomeric units. It also offers with the possibility of further fine-tuning the structural features of PNA to mimic DNA. The development in this area of research are described in recent reviews.<sup>54</sup>

### 1.3.2a Aminopropyl PNA (*ap*-PNA)

Introduction of a methylene bridge between the  $\beta$ -carbon atom of the aminoethyl segment and the  $\alpha''$ -carbon of the glycine segment of *aeg*-PNA resulted in 4-aminopropyl PNA (Figure 23) having two chiral centers.<sup>55</sup> All four diastereomeric T-monomers were synthesized from *trans*-4(*R*)-hydroxyproline and incorporated into PNA oligomers. None of the homochiral aminopropyl thyminy PNAs corresponding to any of the diastereomers bound to target DNA sequences.<sup>56</sup> probably due to high rigidity in the backbone resulting in structural incompatibility.

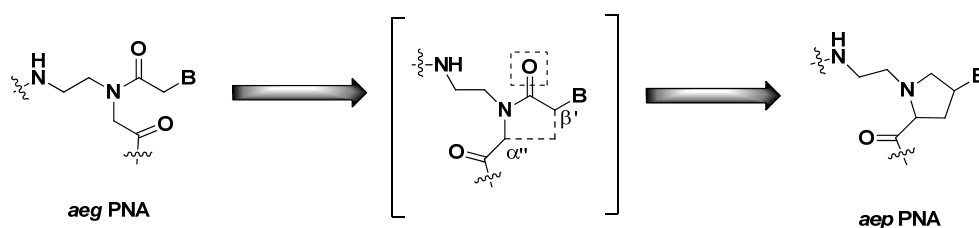


**Figure 23:** Propyl PNA with  $\alpha''$ - $\beta$  methylene bridge

However, incorporation of single chiral *D-trans* or *L-trans* prolyl PNA monomer into *aeg*-PNA at the N-terminus or within the PNA sequence resulted in higher binding of the target DNA with definite preference for a parallel or an antiparallel mode unlike the unmodified PNA.<sup>57</sup> Efforts directed toward releasing the strain by replacing the backbone amide linker with carbamate linkage were not successful.

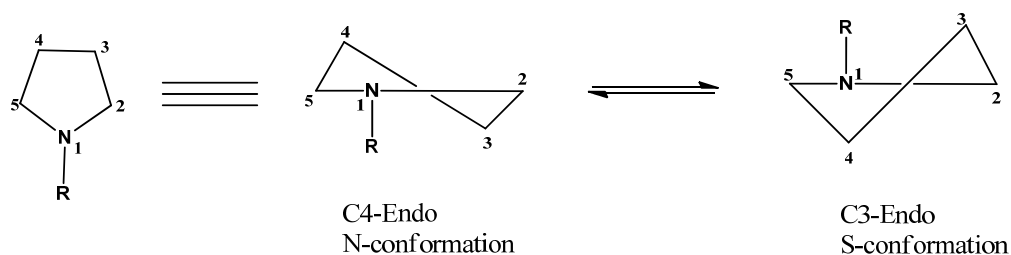
### 1.3.2b Aminoethylprolyl PNA (*aep*-PNA)

Conformationally restricted proline-based PNA in which a pyrrolidine ring replaces the tertiary amide linker to the nucleobase and thus balance flexibility and rigidity in the PNA backbone was derived in the form of *aep* PNA. The  $\alpha'$ -carbon atom of the glycine unit and the  $\beta'$ -carbon atom of the nucleobase linker were joined through a methylene bridge (Figure 24).<sup>58</sup> The flexibility in the aminoethyl segment of *aeg* PNA was retained. The nucleobase attachment to the pyrrolidine ring was fixed by virtue of the chirality of C-4, thus removing the possibility of any rotameric populations. The tertiary amine function in the backbone was found to be at least partially protonated at physiological pH ( $pK_a \sim 6.8$ ). Thus, in *aep* PNA, all elements of the structural freedom of *aeg* PNA were conserved in addition to the restriction of the rotamers. The oligomers comprising 4(*S*), 2(*S/R*) *aep* PNA thymine units showed very favourable binding properties towards the target sequences without compromising the specificity. The mixed pyrimidine hairpin sequences with cytosine and *N*-7 guanine *aep*PNA units exhibited directional discrimination in binding to parallel/antiparallel DNA sequences.<sup>59</sup> The *aep*PNA units carrying the individual nucleobases in a mixed purine/pyrimidine sequence exerted nucleobase-dependent binding efficacies and orientation selectivities towards target oligomers.<sup>60</sup> The adenine homooligomer with (2*S*,4*S*) stereochemistry showed improved binding to the target DNA sequence, whereas the thymine homo-oligomers with same backbone geometry did not show any transition in UV melting experiments. The results so far obtained with *aep* PNA are promising for further biological applications.



**Figure 24:** Aminoethyl prolyl PNA with  $\alpha''$ - $\beta'$  methylene bridge

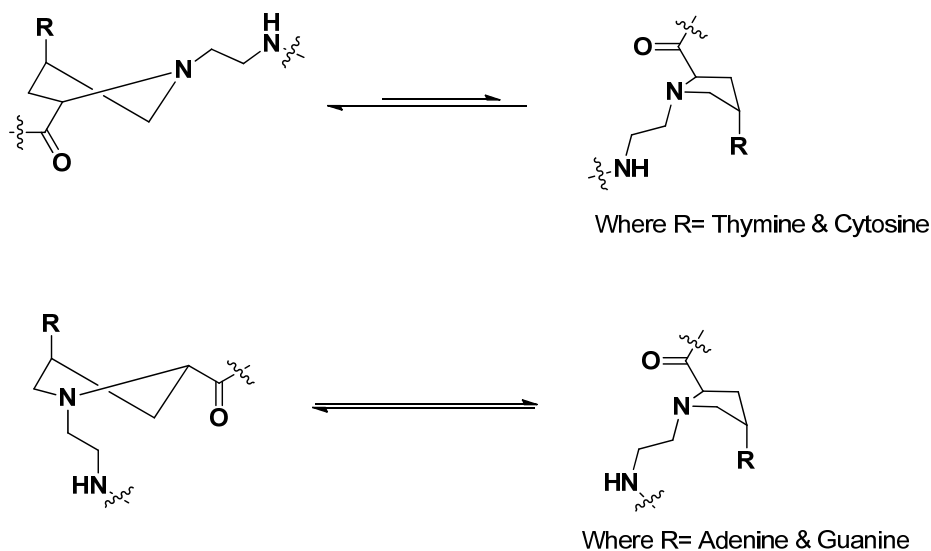
In order to understand the nucleobase dependent properties of *aep* PNA at monomeric level, the structural studies of *aep*PNA (T/C/G/A) is reported from our group.<sup>61</sup> In this study, vicinal coupling constants of pyrrolidine ring protons were used as a function of endocyclic torsion angles and the experimentally measured values were used to derive the pseudo rotation phase angle which provides information about the most puckered region of



**Figure 25:** Conformations of pyrrolidine ring

the ring and the puckering amplitude using PSEUROT programme. This indicates the degree of ring puckering of pyrrolidine ring of various *aep* PNA monomers. The possible ring puckerings of pyrrolidine ring are of two types i.e. C4-endo or N-type and C3 endo or S-type. Conformations of N-type and S-type pyrrolidine ring puckering is showed below. It is found that when C4 of the pyrrolidine ring carries a pyrimidine substitution (Thymine or cytosine), the conformation of ring is remarkably biased to N-type (85-93%) and for

purine (adenine or Guanine) substituted pyrrolidines the conformation is shifted to the S-conformer (55.1%). Similar observations were reported by Altona and Sundaralingam<sup>62</sup> for

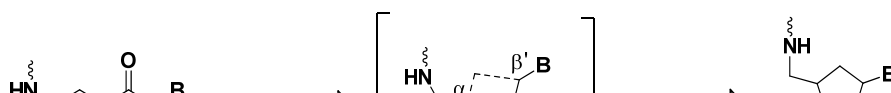


**Figure 26:** The most probable conformation of pyrrolidine ring in *aep*-PNA

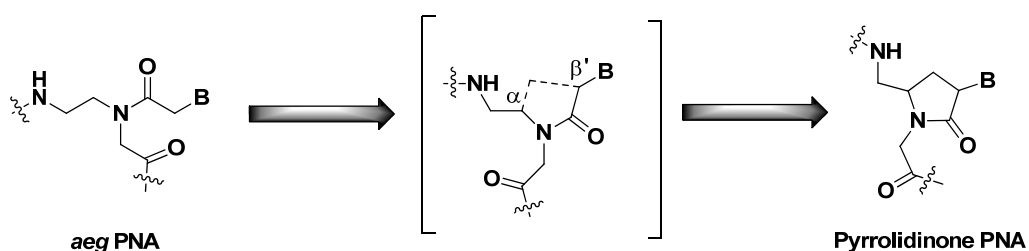
the deoxyribose and ribose rings in nucleotides, where a slight conformational preference for S-type for purine ribosides and N-type conformation for pyrimidines. The change in pyrrolidine ring pucker in *aep*PNA monomers is thus purine/pyrimidine base dependent. This conformational bias in *aep*PNA monomers may have a cumulative influence on the conformation of different base substituted oligomeric backbone with consequence in influencing the final hybridization properties of the derived *aep*PNA oligomers.

### 1.3.2c Pyrrolidine PNA

Another conformationally restricted cyclic PNA analogue was derived from a pyrrolidinone ring system.<sup>63</sup> A methylene bridge was inserted between the  $\alpha$ -carbon atom of the aminoethyl segment and the  $\beta'$ -carbon atom of the acetyl linker to the nucleobase of *aeg*PNA (Figure 27). The carbonyl group of the nucleobase linker was retained and was forced to point towards the carboxy terminus of the backbone. The bridge prevented rotation around the C-N bond of the acetyl segment connecting the nucleobase residue to the backbone, and preorganized PNA in a rotameric conformation prevailing in complexes



of PNA with nucleic acids as studied earlier. The hybridization properties of PNA decamers containing this analogue with complementary DNA, RNA and PNA strands were investigated. The oligomers incorporating the (3*S*,5*R*) isomer were shown to have the highest affinity towards RNA in comparison with DNA.

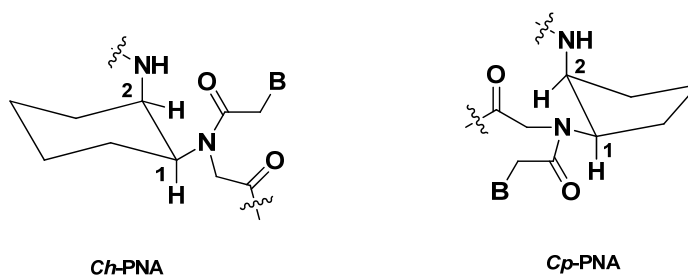


**Figure 27:** Pyrrolidinone PNA with  $\alpha$ - $\beta'$  methylene bridge

### 1.3.2d Carbocyclic PNAs (*Ch*-PNA and *Cp*-PNA)

One of the earliest efforts to introduce conformational constraint in the *aeg*PNA structure resulted in the chiral cyclohexyl-derived backbone (Figure 28).<sup>64</sup> The conformational freedom in the ethylene chain was thus locked in the six-membered cyclic structure. The oligomers with (*S,S*)-cyclohexyl residues were able to hybridize with DNA or RNA, with little effect on the thermal stability depending on their number and the sequence. From the thermal stabilities and molecular modelling based on the solution structure of a PNA-DNA duplex determined by NMR techniques, it was concluded that the right handed hybrid duplex accommodated the (*S,S*) isomer more easily than the (*R,R*) isomer. Thermodynamic measurements also pointed to a decrease in entropy indicating a conformationally constrained structure. In contrast, incorporation of the (*R,R*) isomer resulted in a drastic decrease in the stability of PNA-DNA/RNA complexes. In PNA-PNA duplexes, however, the (*S,S*)- and the (*R,R*)-cyclohexyl residues exerted minor effects on the stability. The crystal structure of both the enantiomers of *cis*-(1*S*,2*R*/1*R*,2*S*)-cyclohexylthyminyl PNA monomers showed the dihedral angle  $\beta$  to be  $-63^\circ$  for (1*S*,2*R*) and  $66^\circ$  for (1*R*,2*S*) isomers,<sup>65</sup> close to those found in the structures of PNA2:DNA triplex and PNA:RNA duplex. The *SR* isomer was more destabilizing than the *RS* isomer in *ch*-PNA:DNA complexes. In the case of RNA complexes a reverse trend was observed, with *RS* being more destabilizing than *SR*. Generally, PNA oligomers with this modification

exhibited a higher affinity toward RNA compared to DNA. The stability studies against DNA oligomer carrying a mismatch indicated that although *Ch*-PNAs formed less stable complexes, the specificity was slightly better than the unmodified *aeg*PNAs. The overall results supported the idea that preorganization of the PNA backbone by chemical modifications such that the torsion angle  $\beta$  is restricted to a range found by NMR and crystal structure data would lead to induction of substantial selectivity in DNA/RNA recognition. A relatively flexible system with a cyclopentyl ring in which the characteristic *endo-exo* puckering that dictates the pseudoaxial/pseudoequatorial dispositions of substituents allows better torsional adjustments to attain the necessary hybridization



**Figure 28:** Cyclohexyl (*Ch*) and Cyclopentyl (*Cp*)PNA

competent conformations. In an attempt to tune the dihedral angle  $\beta$  in this manner the cyclohexyl unit was replaced with *cis*-(1*S*,2*R*/1*R*,2*S*)-cyclopentyl PNA-T (*Cp*-PNA) monomer.<sup>66</sup> The crystal structure data of the monomer indicated that the dihedral angle  $\beta$  was around 24° on the low side of the desired value of 60°. The *RScp*-PNA enantiomer formed higher affinity complexes with DNA as compared to *SRCp*-PNA isomer. Like *Ch*-PNA, *Cp*-PNA also showed higher affinity toward complexation with RNA than DNA. The all modified homooligomers of both enantiomers exhibited significant stabilization of their triplexes with DNA and poly rA. The *Cp*-PNA modification also induced a greater destabilization of the mismatched DNA hybrids than that seen with unmodified *aeg*PNA, indicating that higher affinity was obtained without sacrificing the base specificity. The designs of *Ch*PNA and *Cp*PNA are the outcome of optimized dihedral angles that constrain the PNA backbone for differential DNA/RNA binding and discrimination via preorganization mediation. The inherently rigid *cis*-substituted six membered ring of *Ch*PNAs forbids structural readjustments to bind to DNA (PNA:DNA,  $\beta=140^\circ$ ) and prefers



binding to RNA (PNA:RNA). The flexible (*SR/RS*)-*Cp*-PNA with a relative ease of conformational adjustments in the *cis*-substituted cyclopentyl system allows reorganization of the ring puckering and binding to both DNA and RNA with high affinity with no selectivity.

#### 1.4 BIOLOGICAL APPLICATIONS OF PNA

PNAs are efficient as hybridization probes in fluorescence *in situ* hybridization (FISH) applications<sup>67</sup>, PNA-DNA chimeras can be used as primers for PCR reactions<sup>68</sup> and PNAs can also act as genome cutters<sup>69</sup>. Single stranded DNA detection can be done using PNA with the help of surface enhanced Raman spectroscopy (SERS) method.<sup>70</sup>

Ulf Diederichsen *et al*<sup>71</sup> showed that PNA/peptide conjugates can facilitate vesicle fusion, one of the most fundamental processes in life.

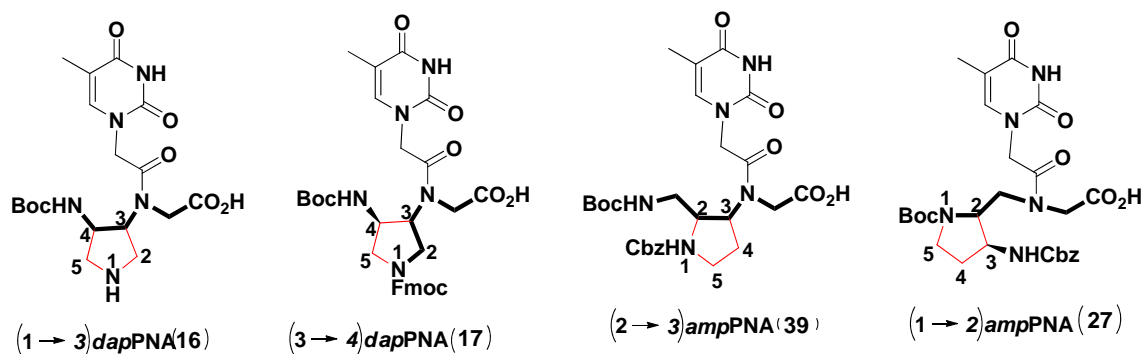
#### 1.5 PRESENT WORK:

The preceding sections give an overview of the peptide nucleic acids (PNAs), the synthetic mimic of natural nucleic acid that is DNA analogues with a homomorphous but chemically different backbone consisting repeating units of *N*-(2-aminoethyl)-glycine (*aeg*) units in contrast to the sugar-phosphate backbone of DNA. The attractive binding properties of PNAs, both in terms of affinity and specificity, coupled with their strand invasion property have promoted PNA as a useful tool in molecular biology, diagnostics, and as a possible candidate for antisense/ antigene drug therapy. The major drawbacks like poor water solubility, inefficient cell uptake, self-aggregation and ambiguity in directionality of binding restrict its applications. Hence various modifications of PNA for improving the activity attempted by different research groups have been described in the previous sections.

### Chapter-2

Chapter 2 involves the synthesis and characterization of novel backbone modified, chiral, ring constrained PNA analogues: *cis*-diamino pyrrolidinyl PNA (*dap*PNA), which is an Aza-analogue of cyclopentyl PNA and *amp*PNA derived from 2-(aminomethyl)-pyrrolidin-3-amine, having one carbon extended backbone at C-2 position of *dap*PNA (Figure 29). This would be interesting to investigate structural-biophysical activity

relationship of these analogues. This chapter describes the synthesis of modified PNA monomers.



**Figure 29:** Structure of modified PNA monomers

### Chapter-3

Chapter 3 deals with the site specific incorporation of synthesized monomers into PNA oligomer sequences by solid phase peptide synthesis. Cleavage of the synthesized oligomers from the solid support, their subsequent purification procedures, followed by suitable characterization and their biophysical studies with complementary DNA were studied through UV-mixing curves.

### Chapter-4

Chapter 4 deals with synthesis and conformational studies of 3-hydroxy substituted polyproline derivatives. Synthesis and conformational studies of polyproline derivatives with naturally occurring 4 substituted proline is done by several research groups. To understand the effect of 3-hydroxy substitution on PPII conformation, in this chapter conformational studies of polyproline derivatives with hydrogen-bond-donating substituent in the  $\beta$ -position(C3) i.e. hydroxyl group on proline ring has done. Conformational behavior of these polyprolines has been examined by CD spectral analyses as a function of pH, urea, and different solvents (buffer and trifluoroethanol).

**1.6 References**

1. Watson, J. D.; Crick, F. H. C: *Nature*, **1953**, *171*, 737-738.
2. (a) Hoogsteen, K. *Acta. Cryst.*, **1963**, *65*, 907. (b) Crick, F. H. C. *J. Mol. Biol.*, **1966**, *19*, 548.
3. Seeman, N. C.; Rosenberg, J. M.; Rich, A: *Proc. Natl. Acad. Sci. USA*. **1976**, *73*, 804-807.
4. Belikova, A. M.; Zarytova, V. F.; Grineva, N. I: *Tetrahedron Lett.* **1967**, *37*, 3557-3562.
5. Miller, P. S.; Braiterman, L. T.; Ts'o, P. O. P: *Biochemistry* **1977**, *13*, 4897.
6. (a) Zamecnik, P. C.; Stephenson, M. L: *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 280-284
7. Stephenson, M. L.; Zamecnik, P. C: *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 285-289.
8. (a) Akhtar, S.; Agrawal, S.: *Trends in Pharmacol Sci.* **1997**, *18*, 12-18. b) Bennet, C. F.: *Biochem Pharmacol.* **1998**, *55*, 9-19. (c) Wickstrom, E: Eds. Marcel Dekker; New York, **1998**.
9. Stein, C. A.; Cohen, J. S In Cohen, J. S. (ed.): *Oligodeoxynucleotides- Antisense Inhibitors of Gene Expression.* (a) Brahms, J.; Mommaerts, W. F. H. London: Macmillan Press, **1989**, p.97
10. De clercq, E, Eckstein, F; Merigan, T. C. *Science*, **1969**, *165*, 1137
11. (a) Checko, K. K.; Linder, K.; Saenger, W.; Miller, P. S: *Nucleic Acid Res.* **1983**, *11*, 2801-2814.
12. Miller, P. S.; Yano, J.; Yano, E.; Carroll, C.; Jayaraman, K.; Ts'o. P. O. P *Biochemistry* **1979**, *18*, 5134-5142.
13. Kean, J. M.; Kipp, S. A.; Miller, P. S.; Kulka, M.; Aurelian, L: *Biochemistry* **1995**, *34*, 14617-14620
14. Froehler, B., Ng, P. and Matteucci, M. *Nucleic Acid Res.*, **1988**, *16*, 4831
15. Letsinger, R. L., Singman, C. N., Histan, G. and Salunke, M. *J. Am. Chem. Soc.*, **1988**, *110*, 4470.
16. Jung, P. M., Histan, G. and Letsinger, R. L. *Nucleosides Nucleotides*, **1994**, *13*, 1597.
17. Summers Zv Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G *Nucleic Acid Res.* **1986**, *14*, 7421-7437.

18. (a) Sood, S., Shaw, B. R. and Spielvogel, B. F. *J. Am. Chem. Soc.*, **1990**, *112*, 9000. (b) Shaw, B. R., Madison, J., Sood, S. and Spielvogel, B. F. Oligonucleotide boranophosphate (borane phosphate). In Agrawal, S. (ed): *Methods in Molecular biology*, **1993**, vol 20: Protocols for Oligonucleotides and Analogs. Synthesis and properties. Totowa, NJ. Humana Press, Inc., pp. 225-243.
19. Li, H.; Huang, F.; Shaw, B. R: *Bioorg. Med. Chem.* **1997**, *5*, 787-795.
20. (a) Higson, A. P.; Sierzchala, A.; Brummel, H.; Zhao, Z.; Caruthers, M. H: *Tetrahedron Lett.* **1998**, *39*, 3899-3902. (b) Rait, V. K.; Shaw, B. R: *Antisense Nucleic Acid Drug Dev.* **1999**, *9*, 53-60.
21. (a) Inoue *et al*: *Nucleic Acid Res.* **1987**, *15*, 6131. (b) Iribarren, A. M. *et al*: *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7747. (c) Lesnik, E. A. *et al*: *Biochemistry* **1993**, *32*, 7832.
22. Moulds, C. *et al*: *Biochemistry*, **1995**, *34*, 5044.
23. Monia, B. P. *et al*: *J. Biol. Chem.* **1993**, *268*, 14514.
24. Kawasaki, A. M. *et al*: *J. Med. Chem.*, **1993**, *36*, 831.
25. Pitsch, S.; Wendeborn, S.; Eschenmoser, A. *Helv. Chim. Acta.* **1993**, *76*, 2161-2183.
26. Schoning, K.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A: Chemical etiology of nucleic acid structure: The  $\alpha$ -threofuranosyl-(3'-2') oligonucleotide system. *Science*, **2000**, *290*, 1347-1351.
27. Schoning, K.; Scholz, (a) Lescrinier, E.; Froeyen, M.; Herdewijn, P. Differences in conformational diversity between nucleic acids with a six –membered ‘sugar’ unit and natural ‘furanose’ nucleic acids. *Nucleic Acids Res.* **2003**, *31*, 2975-2989. (b) Hendrix, C.; Rosemeyer, H.; Verheggen, I.; Seela, F.; Arschot, A. V.; Herdewijn, P: 1',5'-Anhydrohexitol oligonucleotides: Synthesis, base pairing and recognition by regular oligodeoxyribonucleotides and oligoribonucleotides. *Chem. Eur. J.* **1997**, *3*, 110-120.
28. Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J. –I.; Mori, K. –I.; Doi, T.; Imanshi, T: Stability and structural features of the duplexes containing nucleoside analogues with fixed type conformation, 2'-O, 4'-C-methyleneribonucleosides. *Tetrahedron Lett.* **1998**, *39*, 5401-5404.
29. Petersen, M.; Wengel, J. LNA: A versatile tool for therapeutic and genomics. *Trends Biotechnol.* **2003**, *21*, 74-81 (b) Koshkin, A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V.

- K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J: LNA (locked nucleic acid): synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleotide monomers, oligomerization and unprecedented nucleic acid recognition. *Tetrahedron* **1998**, *54*, 3607-3630. (c) Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J: LNA (locked nucleic acid): Synthesis and high affinity nucleic acid recognition. *Chem. Commun.* **1998**, 455-456.
30. (a) Summerton, J., Stein, D., Huang, S. B., Matthews, P., Weller, S. and Partridge, M. *Antisense Nucleic Acid Drug Dev.*, **1997**, *7*, 63. (b) Summerton, J. and Weller, D. *Antisense Nucleic Acid Drug Dev.*, **1997**, *7*, 187-195.
31. Toyler, M. F. *et al Antisense Nucleic Acid Drug Dev.*, **1998**, *8*, 199.
32. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497.
33. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R.H.; Kim, S. K.; Nordon, B.; Nielsen, P. E; **PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules.** *Nature*, **1993**, *365*, 566-568.
34. Jensen, K. K.; Qrum, H.; Nielsen, P. E.; Norden, B: **Hybridization kinetics of peptide nucleic acids (PNA) with DNA and RNA studied with BIAcore Technique.** *Biochemistry* **1997**, *36*, 5072-5077.
35. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.;Sonnichsen, S. H.; Nielsen, P. E: **Stability of peptide nucleic acids in human serum and cellular extracts.** *Biochem. Pharmacol.* **1994**, *48*, 1310.
36. (a) Good, L.; Nielsen, P. E: **Progress in developing PNA as gene targeted drugs.** *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 431. (b) De Mesmaeker, A. et al: *Curr. Opin. Struct. Biol.* **1995**, *5*, 343. (c) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W: **PNA: Synthetic polyamide Nucleic Acids with Unusual Binding Properties.** *Angew. Chem. Int. Ed.* **1998**, *37*, 2796-2823.
37. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graeslund, A:**Ionic effects on the stability and conformation of peptide nucleic acid (PNA) complexes.** *J.Am. Chem. Soc.* **1996**, *118*, 5544-5552.

38. Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H: **Peptide nucleic acids (PNA). Oligonucleotide analogues with an achiral peptide backbone.** *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.
39. Nielsen, P. E.; Christensen, L: **Strand displacement binding of a duplex forming Homopurine PNA to a homopyrimidine duplex DNA target.** *J. Am. Chem. Soc.* **1996**, *118*, 2287-2288.
40. (a) Bonham, M. A. *et al*, *Nucleic Acid Res.* **1995**, *23*, 1197-1203. (b) Knudsen, H.; Nielsen, P. E: **Antisense properties of duplex and triplex forming PNA.** *Nucleic Acid Res.* **1996**, *24*, 494-500.
41. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O: **Peptide nucleic acids (PNA). DNA analogues with a polyamide backbone.** In **Antisense Research and Application.** Crook, S. and Lebleu, B. (eds). CRC Press, Boca Raton, **1993**, pp. 363-373.
42. (a) Praseuth, D. *et al*: **Peptide Nucleic Acids directed to the promoter of the  $\alpha$ -chain of the interleukin-2-receptor.** *Biochim. Biophys. Acta.* **1997**, *1309*, 226-238. (b) Wittung, P.; Nielsen, P. E.; Norden, B: **Extended DNA-recognition repertoire of PNA.** *Biochemistry* **1997**, *36*, 7973-7979.
43. Lohse, J.; Dahl, O.; Nielsen, P. E: **Double duplex invasion by peptide nucleic acid: A general principle for recognition of double stranded DNA.** *Proc. Natl. Acad. Sci. U.S.A.* **1999**,
44. (a) Leijó, M. *et al*: **Structural characterization of PNA-DNA duplexes by NMR. Evidence for DNA in a B-like conformation.** *Biochemistry* **1994**, *22*, 9820-9825. (b) Eriksson, M.; Nielsen, P. E: **Solution structure of a peptide nucleic acid-DNA duplex.** *Nat. Struct. Biol.* **1996**, *3*, 410-413.
45. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G: **NMR Solution structure of a peptide nucleic acid complexed with RNA.** *Science* **1994**, *265*, 777-780.
46. Dueholm, K. L.; Nielsen, P. E. *New J. Chem.* **1997**, *21*, 19.
47. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993**, 518.
48. (a) Haiima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939. (b) Sforza, S.; Haiima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org.*

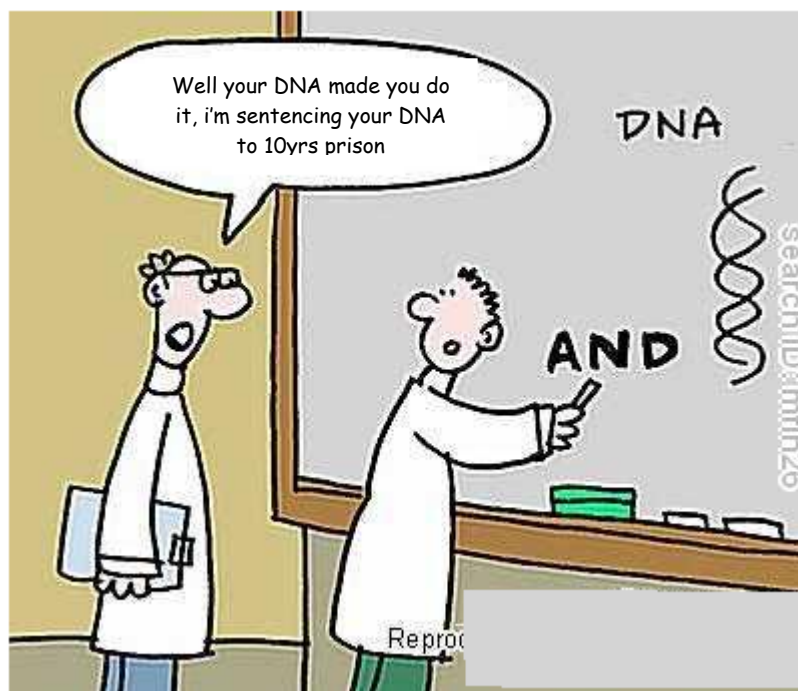
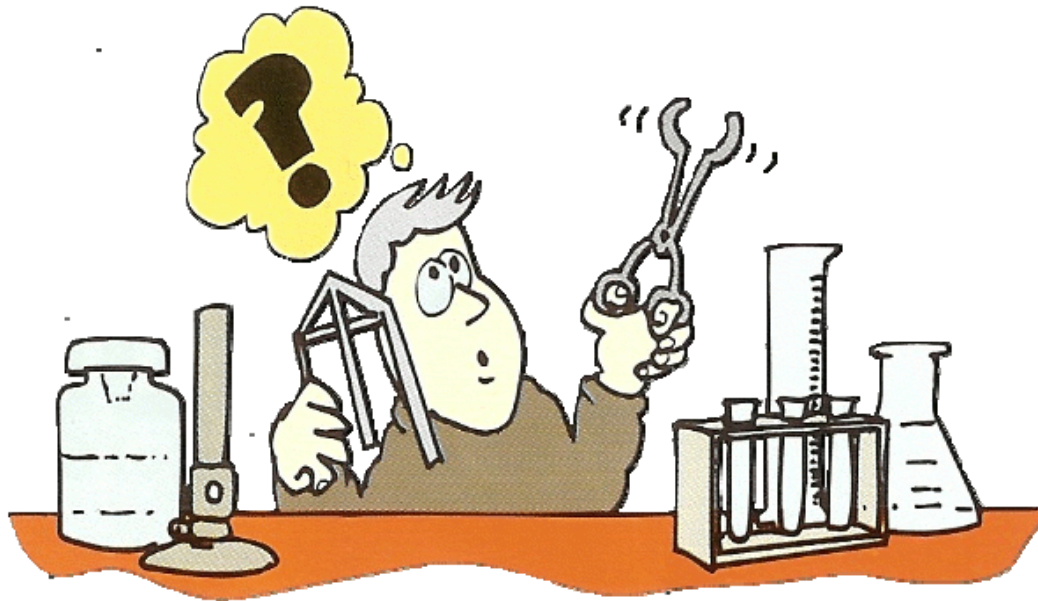
- Chem.* **1999**, 197. (c) Zhang, L.; Min, J.; Zhang, L. *Bioorg. Med. Chem. Lett.* **1999**, 9, 2903.
49. Barawkar, D. A.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, 121, 10418.
50. Efimov, V. A.; Choob, M. V.; Buryakova, A. A.; Chakhmakhcheva, O. G. *Nucleos. Nucleot. Nucl.* **1998**, 17, 1671.
51. Efimov, V. A.; Buryakova, A. A.; Choob, M. V.; Chakhmakhcheva, O. G. *Nucleos. Nucleot. Nucl.* **1999**, 18, 1393.
52. Kuwahara, M.; Arimitsu, M.; Shisido, M. *Tetrahedron* **1999**, 55, 10067.
53. Altmann, K. H.; Chiese, C. S.; Garcia-Echeverria, C. **1997**, 7, 1112.
54. (a) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, 38, 404. (b) Kumar, V. A. *Eur. J. Org. Chem.* **2002**, 2021.
55. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1996**, 52, 15017-15030.
56. Gangamani, B. P.; D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides* **1999**, 18, 1409-1011.
57. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, 55, 177-192.
58. D'Costa, M.; Kumar, V. A.; Ganesh, K. N.; *Org. Lett.* **1999**, 1, 1513-1516.
59. Kumar, V. A.; D'Costa, M.; Ganesh, K. N. *Nucleosides, Nucleotides Nucleic Acids* **2001**, 20, 1187-1191.
60. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **2001**, 3, 1281-1284.
61. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **2001**, 3, 1281-1284.
62. (a) Barchi, J. J., Jr.; Karki, R. G.; Nicklaus, M. C.; Siddiqui, M. A.; George, C.; Mikhailopulo, I. A.; Marquez, V. E. *J. Am. Chem. Soc.* 2008, 130, 9048e9057;
63. Puschl, A.; Boesen, T.; Zuccarello, G.; Dahl, O.; Pitsch, S.; Nielsen, P. E. *J. Org. Chem.* **2001**, 66, 707-712.
64. Lagriffoule, P.; Wittung, P.; Eriksson, M.; Jensen, K. K.; Norden, B.; Buchardt, O.; Nielsen, P. E. *Chem. Eur. J.* **1997**, 3, 912-919.
65. Govindaraju, T.; Gonnade, R.; Badbhade, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **2003**, 5, 3013-3016.
66. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *Chem. Commun.* **2004**, 860-861.
67. Chen, C.; Hong, Y. K.; Ontiveros, S. D.; Egholm, M.; Strauss, W. M. *Mammalian Genome* **1999**, 10, 13.

68. Drewe, L. J.; Brightwell, G.; Hall, E. A. *Mol Cell Probes* **2000**, *14*, 269. Isacson, J.; Cao, H.; Ohlsson, L.; Nordgren, S.; Svanvik, N.; Westman, G. *Molecular and Cellular Probes* **2000**, *14*, 321.
69. Veselkov, A. G.; Demidov, V.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **1996**, *24*, 2483.
70. Laura Fabris, Mark Dante, Gary Braun, Seung Joon Lee, Norbert O. Reich, Martin Moskovits, Thuc-Quyen Nguyen,\* and Guillermo C. Bazan\* *J. Am.chem. soc.* **2007**, *129*, 6086-6087
71. Antonina S. Lygina, Karsten Meyenberg, Reinhard Jahn, and Ulf Diederichsen *Angew. Chem. Int. Ed.* 2011, *50*, 8597–8601
72. [http://en.wikipedia.org/w/index.php?title=File:DNA\\_chemical\\_structure.svg&page=1](http://en.wikipedia.org/w/index.php?title=File:DNA_chemical_structure.svg&page=1)
73. Efomov, V. A.; Buryakova, A. A.; choob, M. V.; chakhmakhcheva, O. G. *Nucleos. Nucleot. Nucl.* **1998**, *17*, 1671
74. Efomov, V. A.; Buryakova, A. A.; choob, M. V.; chakhmakhcheva, O. G. *Nucleos. Nucleot. Nucl.* **1998**, *18*, 1393
75. Kuwahara, M.; Arimitsu, M; Shisido, M. *Tetrahedron*, **1999**, *55*, 10067
76. Altmann, K. H.; Chiese, C.S.; Garcia-Echeverria, C. 1997, *7*, 1112.
77. Peter E. Nielson; *Methods in Molecular Biology*, vol. 208: Peptide Nucleic Acids: Methods and Protocols
78. Zc Thisted, M., Just, T., Pluzek, K. -J., Hyldig-Nielsen, J. J., and Godtfredsen, S. E. Detection of immunoglobulin kappa light chain mRNA in paraffin sections by in situ hybridisation using peptide nucleic acid probes. *Cell Vision*, **1996**, *3*, 358-363



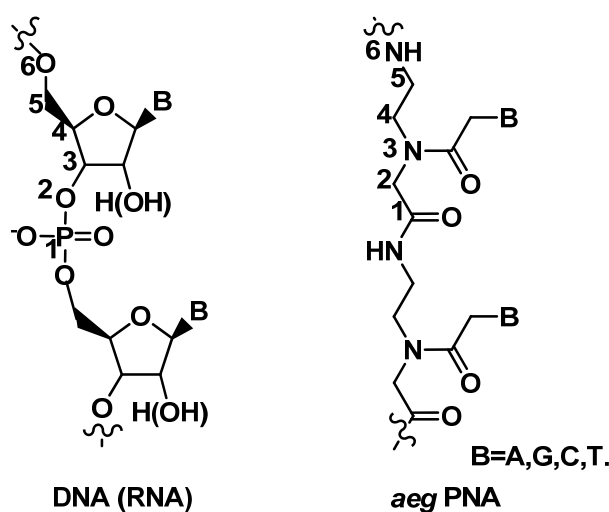
## Chapter 2:

### “Synthesis of 2,3 and 3,4-disubstituted pyrrolidine based peptide nucleic acid monomers”

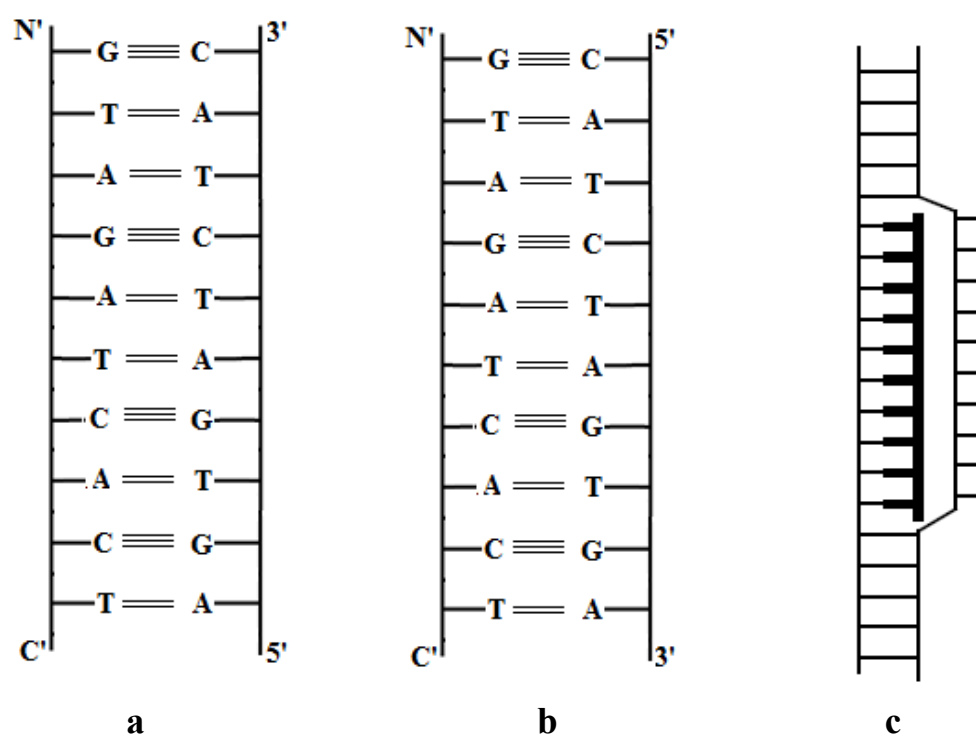


## 2.1 Introduction

Highly stable analogues of DNA or RNA are promising candidates for inhibiting gene expression.<sup>1</sup> When introduced into cells, these analogues seek complementary sequences on target DNA duplex (called as antigene therapy) or single stranded mRNA (called as antisense therapy) and bind to inhibit the production of corresponding coded functional proteins in a cell.<sup>2</sup> Chemical modifications of the sugar, nucleobase or the phosphodiester linkage may improve the entry of these analogues into cells and also prevent their intracellular degradation.<sup>3</sup> Among several modifications, Peptide Nucleic Acid (PNA) designed by Nielsen *et al*<sup>4</sup> using computer modelling has emerged as one of the potentially effective nucleic acid analogues. In PNA, the entire negatively charged sugar-phosphate backbone of DNA is replaced by a neutral and achiral polyamide backbone consisting of repeating *N*-(2-aminoethyl)-glycine (*aeg*) unit. The four nucleobases (A, G, C, T) are attached to the backbone through a conformationally rigid tertiary acetamide linker (Figure 1). The internucleobase distance in PNA are conserved as in DNA and this allows its binding to the target DNA/RNA sequences with high sequence specificity and affinity.<sup>5</sup> Moreover, PNA is stable to cellular enzymes like nucleases and proteases. However, the major limitations for the therapeutic applications of PNA is its poor solubility in aqueous media due to self-aggregation, insufficient cellular uptake and ambiguity in orientational selectivity of binding.<sup>6</sup>



**Figure 1:** DNA (RNA) and PNA structures



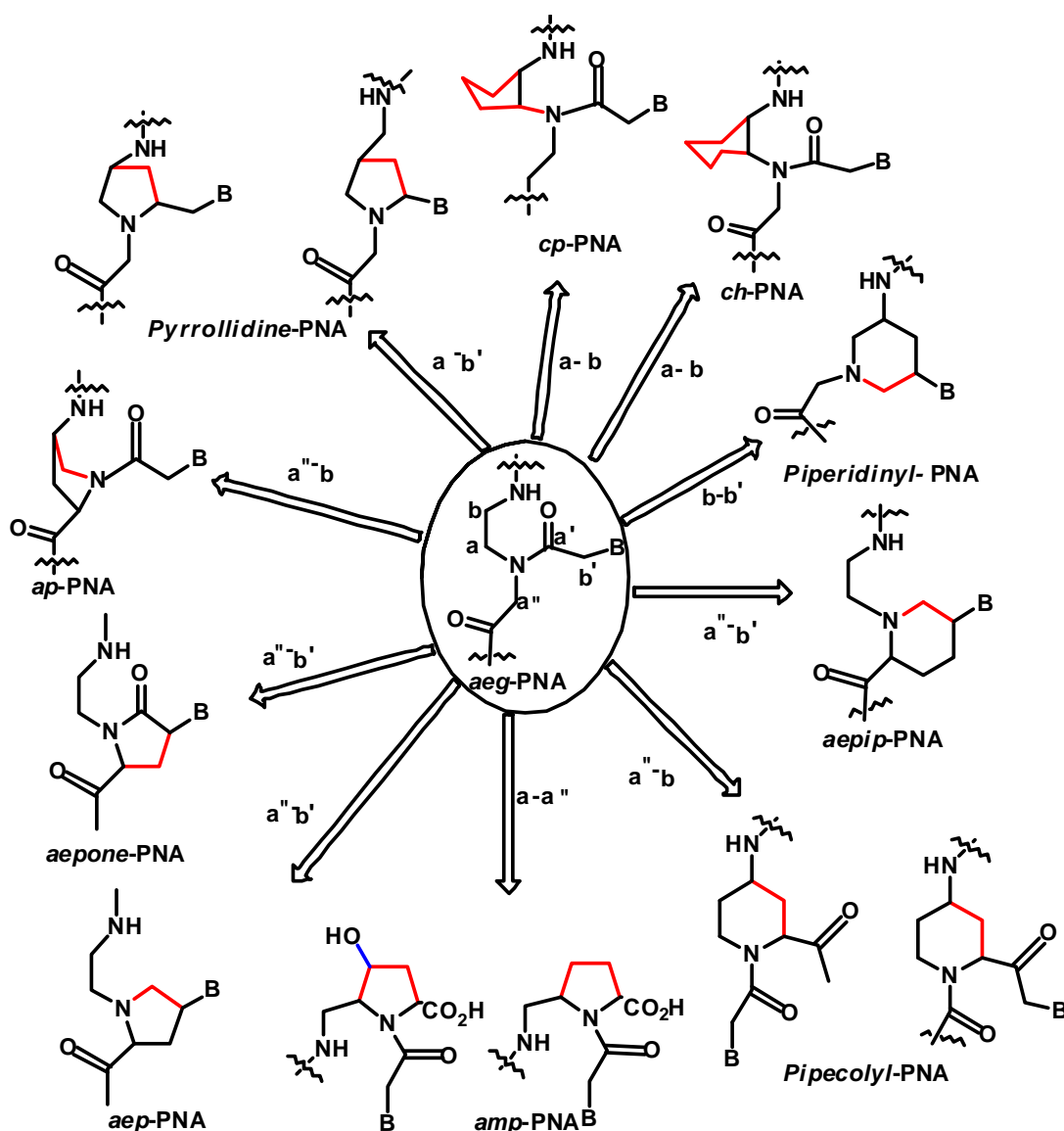
**Figure 2:** (a) Antiparallel mode of binding of PNA:DNA (b) Parallel mode of binding of PNA:DNA (c) PNA binding mode for targeting duplex DNA (Thick structure signifies (PNA)<sup>20</sup>)

The three dimensional structure of PNA:DNA and PNA:RNA duplex<sup>7</sup> demonstrate that PNA has the flexibility to adapt to the A-form or B-form helices preferred by RNA or DNA respectively. The P-form structure preferred by PNA is much wider and slower winding helix. These results suggest that PNA is not a perfect conformational mimic of DNA or RNA, and therefore enough scope exists for chemical modifications for further improvement.

To address the limitations of PNA for biological applications several modifications have been introduced into the classical PNA backbone<sup>8</sup> and many of these modifications have resulted in only marginal effects in terms of hybridization properties. However, these studies have pointed out the importance of rigidity and pre-organization of PNA for effective complexation with complementary DNA. From this laboratory, there have been several reports on the introduction of novel five or six membered rings into PNA backbone to impart chirality as well as

conformational constrain in the PNA backbone<sup>9</sup> resulting in both neutral and positively charged PNA oligomers (Figure 3).

The results indicated that the sterically allowed constrains of the PNA backbone improves its tendency to hybridize with DNA/RNA, and this depends on the nature of stereochemistry of the residue. Introduction of chemical bridges in *aegPNA* to get cyclic structures may help in constraining the backbone and to achieve parallel/antiparallel orientation selective binding by virtue of chirality in the backbone.

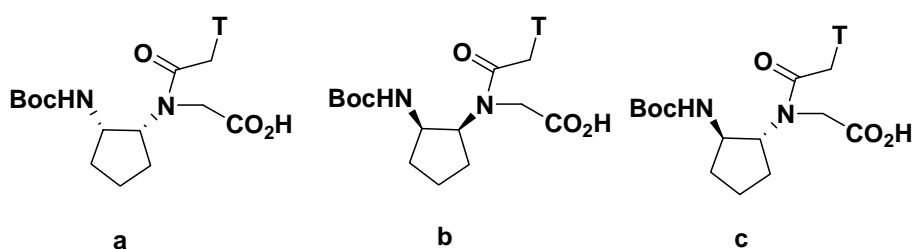


**Figure 3:** Conformationally constrained PNA analogues<sup>9</sup>

It is well known that hybridization of complementary oligomers (DNA, RNA or PNA) is characterized by a large enthalpy gain and a significant entropy loss.<sup>10</sup> The decrease in entropy upon hybrid formation is naturally due to the formation of a highly ordered and fairly rigid duplex structure from two rather flexible and much less ordered single strands. Therefore pre-organising the single stranded PNA in a conformation identical or close to that found in the PNA:DNA/RNA hybrid should greatly reduce the entropy loss and hence increase the free energy upon binding. Restricting the backbone flexibility, e.g. by introducing cyclic structures is thus an attractive strategy in the quest for oligomers with improved hybridization potency.

## 2.2 Rationale design behind the work

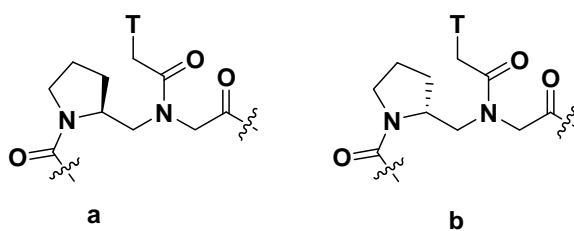
The work presented in this thesis is directed towards the introduction of chirality and conformational constrain in the classical *aeg*PNA backbone to control the orientation selectivity and specificity in binding by preorganization. Recently a five membered carbocyclic analogue of PNA (*cp*-PNA) with *cis* chemistry from our group<sup>11</sup> and *trans* chemistry from Appella *et al*<sup>12</sup> was reported (Figure 4, *cp*-PNAs). Binding studies on *cis* cyclopentyl PNAs showed significant stabilization of the derived PNA:DNA and PNA:RNA complexes and marginal stabilization in *trans* analogues.



**Figure 4:** *cis* Cyclopentyl PNA's (a & b)<sup>11</sup> and *trans* Cyclopentyl PNA (c).<sup>12</sup>

Yehekiely *et al*<sup>13</sup> introduced *N*-(Pyrrolidinyl-2-methyl) glycine based PNA (*Pmg*PNA) derived from L-prolinol and D-prolinol (Figure 5) in which sterically constrained *t*-amide group is part of backbone. UV thermal melting experiments with complementary DNA and RNA showed that oligomers incorporating either (*R*) or the (*S*) *Pmg* PNA units at different positions preferentially bound to complementary RNA with a higher selectivity than control *aeg*PNA. However, they showed poorer

duplex stability towards DNA/RNA than *aeg*PNA which was attributed to lack of positive charge in the PNA.



**Figure 5:** a. (*R*)-Pmg PNA b. (*S*)-Pmg PNA<sup>13</sup>

The major limitations of the therapeutic applications of PNA are their poor solubility in aqueous media and the selectivity of PNA towards DNA /RNA is a valuable concern. The cyclopentyl PNAs show better stabilization of PNA:DNA/RNA complexes than *aeg*PNA, but they are more hydrophobic due to presence of cyclopentane ring. Further absent of positive charge is an important factor for lower solubility. On the other hand, *Pmg*PNAs do not show stabilization of PNA:DNA complexes compared to unmodified PNA, perhaps due to absence of cationic backbone, but show better selectivity towards RNA than DNA as compared to *aeg* PNA.

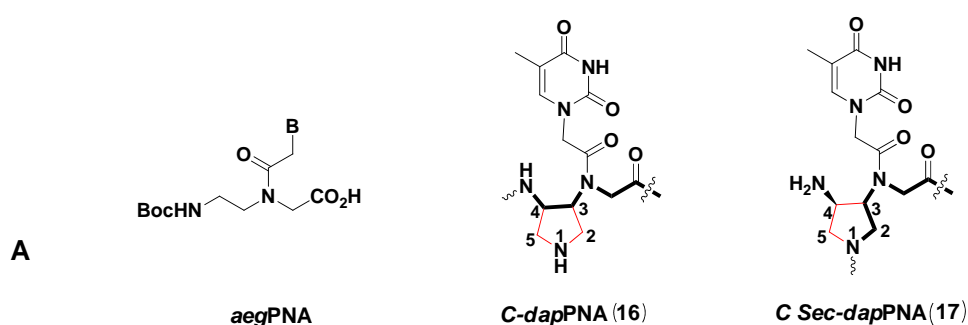
It is reported in the literature that introduction of positive charges in the PNA backbone improves its binding to DNA and RNA.<sup>14</sup> Designing positively charged PNAs is of considerable interest as they may possess superior ability to strand invasion of complementary DNA sequences. The positive charge can arise from the introduction of amino/imino functions in PNA that are protonatable. The introduction of pyrrolidine derivatives in several PNA analogues make the backbone cationic and rigid.<sup>9</sup> But it is important to make sure that the positive charge in the PNA does not compromise the base specificity in binding to DNA/RNA (by electrostatic interactions) that may have detrimental effect on the fidelity of the PNA:DNA base pairing. The mixed results surfacing from different groups indicate that the conformational constraint imposed onto PNA backbone in the form of ring structure may not be uniform for various nucleobases and hence may have a sequence context. Previous studies on DNA/RNA analogues have shown that the number of bonds in the repetitive DNA/RNA backbone can vary from five to seven.<sup>15</sup>

### 2.2.1 Rationale

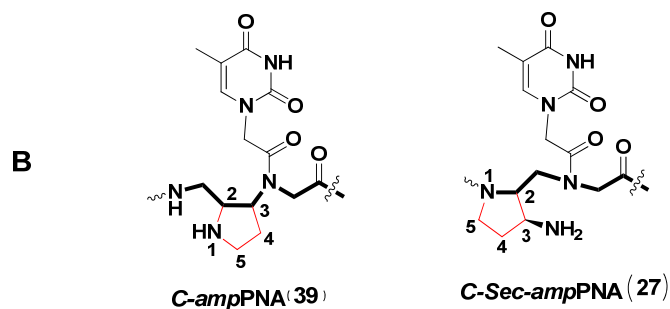
In the light of all these literature findings, the synthesis of 2,3 & 3,4-disubstituted pyrrolidine PNA monomers (**16**, **17**, **27** & **39**) (shown in Figure 6) was initiated. These involved the synthesis of protected monomers

- i) *cis*-(3*R*,4*R*)-3,4-diaminopyrrolidine (*3*→*4*)*dap*PNA(**16**) & (*1*→*3*)*dap*PNA (**17**) with structures derived from introducing a nitrogen atom in the cyclopentyl ring of *cp*-PNA<sup>11</sup> and keeping all the other moieties unchanged in order to get an imino group in the PNA backbone to make it cationic and more hydrophilic than the corresponding cyclopentyl PNA and
- ii) (*2R,3R*)-3-amino-2-aminomethyl pyrrolidine ((*2*→*3*)*amp*PNA **39** & (*1*→*2*)*amp*PNA **27**) similar to monomers 16 & 17 but with one carbon extended in the backbone to create a good balance between the increased rigidity due to pyrrolidine ring and flexibility due to one extra methylene.

Most of the PNA modifications in the literature, utilising a pyrrolidine ring in backbone were either proline derived or have substitution at  $\alpha$ -carbon to ring nitrogen (Figure 3). Eventhough, *cp*PNAs with *cis* stereochemistry showed a significant stabilization of the derived PNA:DNA/RNA complexes but its backbone is hydrophobic. Introduction of nitrogen atom in the backbone of *cp*PNA can make it hydrophilic.



Where B= A, T, G & C



The red bonds represent the derived modified PNA backbone.

**Figure 6: (A) aegPNA and different PNA backbones derived from *dap* Monomers (B) PNA backbones derived from *amp* Monomers**

All analogues carry positive charges either on their primary or secondary amino groups and hence the study of 1,2-*cis* disubstituted pyrrolidine PNA analogues were undertaken. The present target PNAs constitute a new class of PNA analogues having substitution at  $\beta$ -position to the ring nitrogen.

Oligomerization of monomers **17** and **27** (Figure 6) is through ring nitrogen similar to *PmgPNA* monomers (Figure 5) and possess extra chiral centre at C4-position and C3-position respectively that bear amino groups.

### 2.3 Objectives

The specific objectives of this chapter are

- i) synthesis of of *cis*-(3*R*,4*R*)-3,4-diamino pyrrolidinyl derived PNA monomers (**16** & **17**) with N-boc or N-fmoc protecting groups.
- ii) synthesis of *cis*-(2*R*,3*R*)-2-(aminomethyl)pyrrolidine derived PNA monomers (**27** & **39**) with boc as protecting group
- iii) synthesis of boc and fmoc protected aminoethylglycyl (*aeg*) PNA monomers..

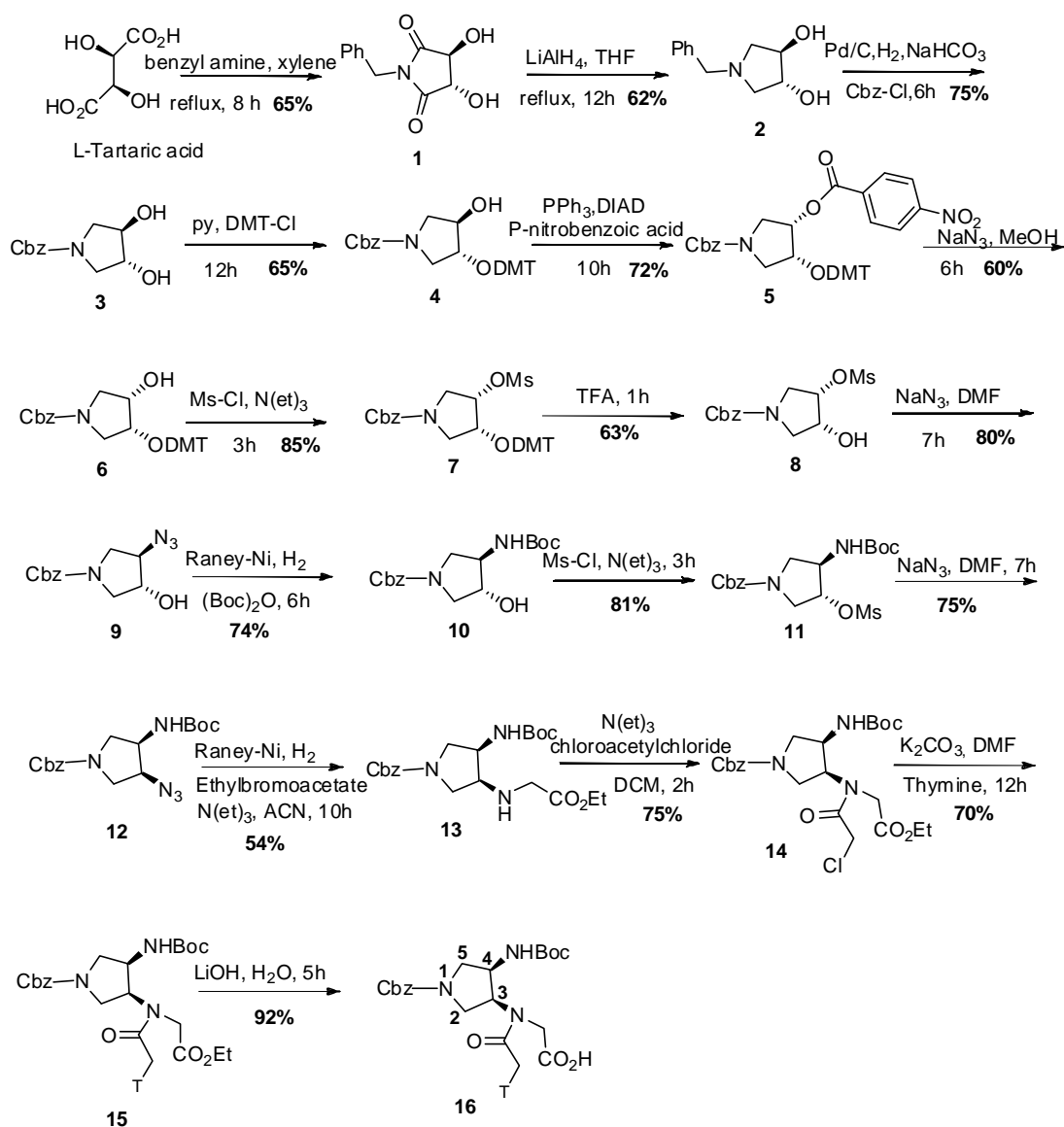
#### 2.3.1 Synthesis of *cis*-(3*R*,4*R*)-3,4-diamino pyrrolidinyl derived PNA monomers (**16** & **17**)

The synthesis of **16** and **17** were started from the naturally occurring L-tartaric acid which is commercially available and inexpensive. With its two chiral centres, L-tartaric acid has been a material of choice to access multifunctionalized pyrrolidine rings through their reactions with primary amines.<sup>16</sup>



The first step involved the condensation of L-tartaric acid with benzyl amine in xylene, using Dean-Stark apparatus and the reaction progress was monitored by the amount of water collected. A good yield of imide **1** was obtained when the theoretical amount of water was collected and the product was recrystallized from ethanol. This was reduced with LiAlH<sub>4</sub> to get the diol **2**. The reported literature methods such as NaBH<sub>4</sub>/I<sub>2</sub> in THF/diethylether and LiAlH<sub>4</sub> in diethylether<sup>16</sup> were used to reduce the imide **1** but gave low yield of diol **2**. Under refluxing THF conditions, the two carbonyl groups were incompletely reduced to give mixture of alcohols which being polar could not be easily separated leading to a lower yield of desired product. Finally, the imide **1** was reduced completely to diol **2** by a slight modification of the work-up procedure. Instead of quenching the reaction mixture with aq. NaOH, it was quenched with saturated aqueous solution of sodium sulphate, followed by stirring the filtrate with diethylether to get pure diol **2**. The NMR data obtained was consistent with literature reports.<sup>17</sup>

The N<sup>1</sup>-benzyl (3,4)-diol **2** was hydrogenated in presence of Pd/C followed by *in-situ* protection with benzyloxy carbonyl to yield compound N-benzyloxy carbonyl 3,4-diol **3**. The mono protection of diol **3** with dimethoxytrityl group gave mono O-DMT protected compound **4**. The 4*R*-hydroxyl group of compound **4** was converted to the 4*S*-p-nitro benzoate ester **5**, under Mitsunobu conditions using p-nitrobenzoic acid. This 4*S*-benzoate ester **5** on reaction with sodium azide and methanol yielded the 4*S*-alcohol **6** that was converted to the corresponding O-mesyl derivative **7**. This on treatment with TFA gave the (3*S*,4*R*) O-mesylate alcohol **8** which was transformed to (3*R*,4*R*) azido alcohol **9**. The reduction of azide group in **9** followed by *in-situ* reaction with di-tert-butyl dicarbonate gave the N-boc protected (3*R*,4*R*) alcohol **10** that was mesylated to the O-mesyl derivative **11**.

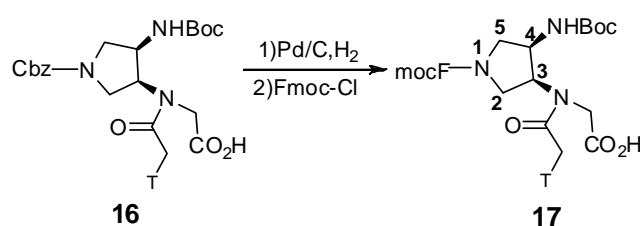
2.3.2 Synthesis of (3*S*,4*R*) (3→4)*dap*PNA Monomer (16)Scheme 1: Synthesis of (3*S*,4*R*) (3→4)*dap*PNA Monomer (16)

$S_N2$  displacement of O-mesyl group in compound **11** by azide group yielded the compound **12** with (3*S*,4*R*) configuration. Upon reduction of azide in **12** followed by N-alkylation with ethylbromoacetate yielded compound **13**, which was N-acylated with chloroacetylchloride to obtain compound **14** with (3*S*,4*R*) configuration. N1-alkylation of thymine with compound **14** using  $K_2CO_3$  in DMF at

70°C gave the (3*S*, 4*R*) ester **15**, which was hydrolysed with LiOH to obtain the monomer **16**. All new compounds were characterized by <sup>1</sup>H, <sup>13</sup>C, and mass spectral data. (see appendix)

### 2.3.3 Synthesis of (3*S*,4*R*) (1→3)*dap*PNA Monomer (17)

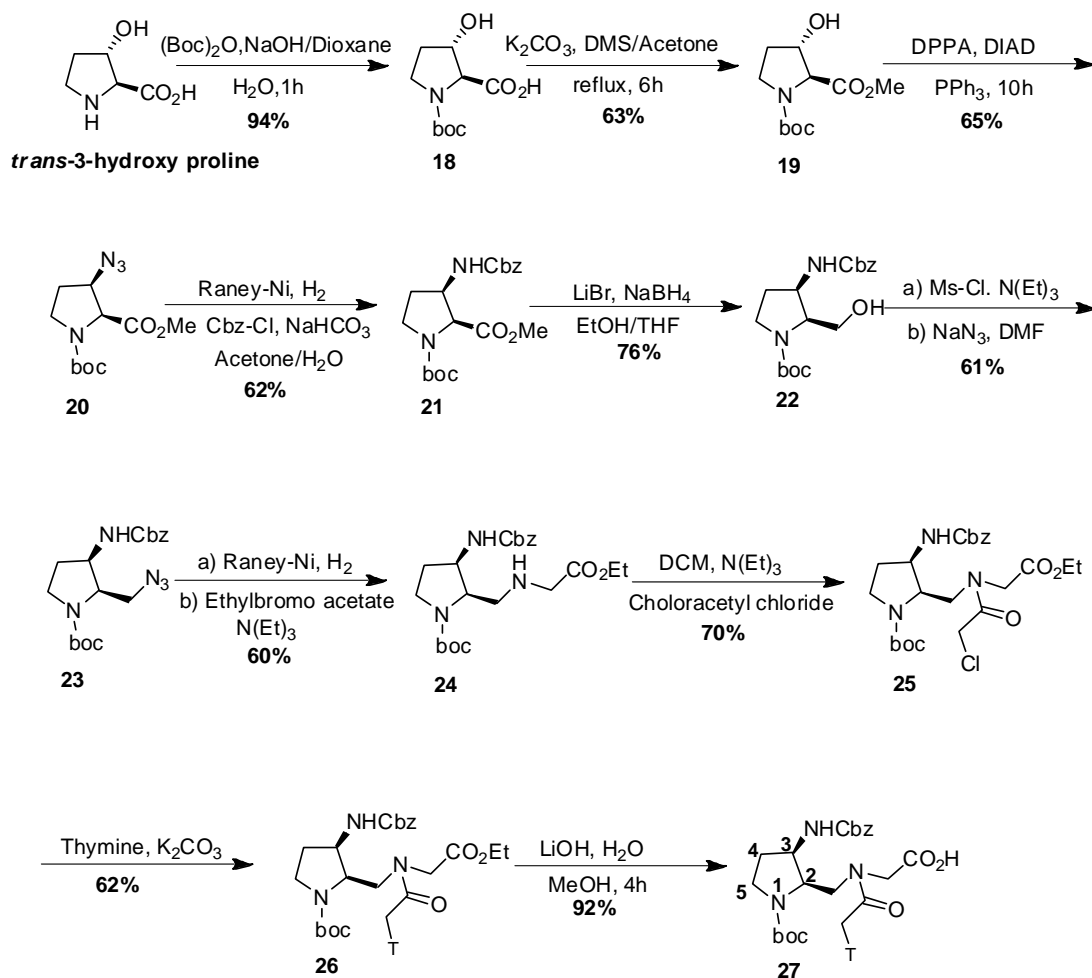
Compound **16** (3*S*,4*R*) on hydrogenolysis with Pd/C and followed by treatment with Fmoc-Cl yielded the desired monomer **17** with (3*S*,4*R*) configuration which was used for the synthesis of PNA oligomers.



**Scheme 2:** Synthesis of (3*S*,4*R*) (1→3)*dap*PNA Monomer (17)

### 2.3.4 Synthesis of (2*S*,3*R*) (1→2)*amp*PNA Monomer (27)

Synthesis of the PNA monomer **27** was started by protecting the ring nitrogen of 3*S*-hydroxy commercially available proline with *tert*-butyloxy carbonyl group to give N-boc protected 3*S*-hydroxy proline **18**. Esterification of the  $\alpha$ -carboxyl group N-boc 3*S*-hydroxy proline with dimethyl sulphate in presence of K<sub>2</sub>CO<sub>3</sub> yielded the 3*S*-hydroxy 2*S*-methyl ester **19**. This was followed by reaction of the 3*S*-hydroxyl ester of **19** with diphenyl phosphorylazide under Mitsunobu conditions to yield the 3*R*-azido 2*S*-methyl derivative **20**. The reduction of 3*R*-azido group in **20** with Raney-Ni, followed by *in-situ* protection of the resulting amine with benzyloxycarbonyl gave the fully protected product **21**. The 2*S*-methyl ester group in **21** was reduced to the corresponding alcohol using LiBH<sub>4</sub> to yield the 2*S*-alcohol **22** that was O-mesylated using mesyl chloride followed by S<sub>N</sub>2 displacement of O-mesyl group in compound **22** by azide group using sodium azide gave the 3*R*-azide **23**.

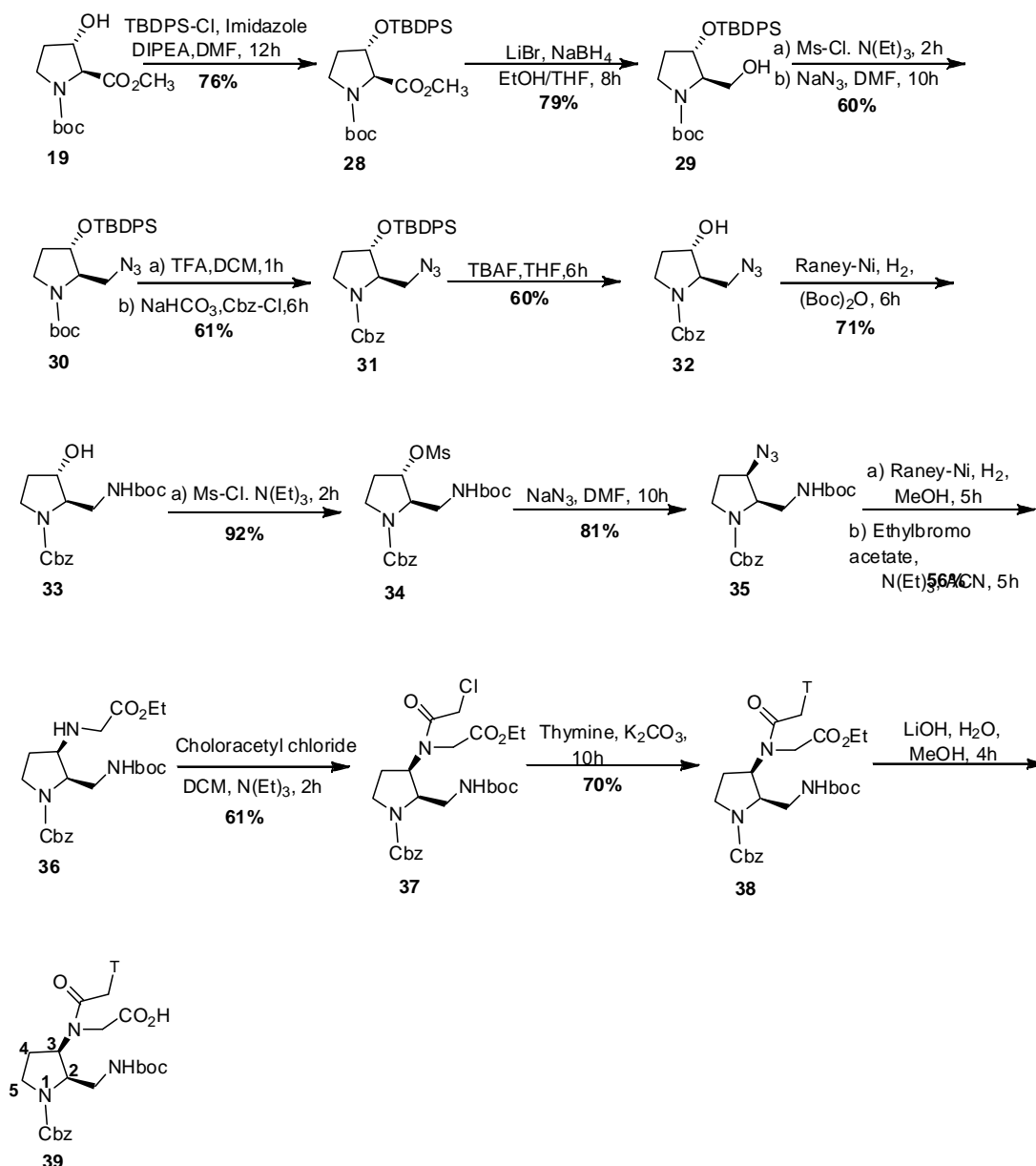


**Scheme 3:** Synthesis of (2*S*,3*R*) (1→2)*amp*PNA Monomer (**27**)

Upon reduction of 3*R*- azide group in **23** with Raney-Ni followed by N-alkylation with ethylbromoacetate in acetonitrile gave the N-alkylated (2*S*,3*R*) ester **24**, which was acylated with chloroacetylchloride to obtain the chloro methyl derivative **25**. This was used for N-alkylation of thymine to yield the ester **26** which was hydrolysed to the required monomer **27** with (2*S*,3*R*) configuration.

### 2.3.5 Synthesis of (2*R*,3*R*) (2→3)*amp*PNA Monomer.

The 3*S*-hydroxyl group of compound **19** was protected as *t*-butyldiphenylsilyl ether and the ester group was reduced with LiBH<sub>4</sub> to obtain the corresponding alcohol **29**. Hydroxyl group of **29** was O-mesylated, followed by azidation yielded the 2*R*-azide **30**.

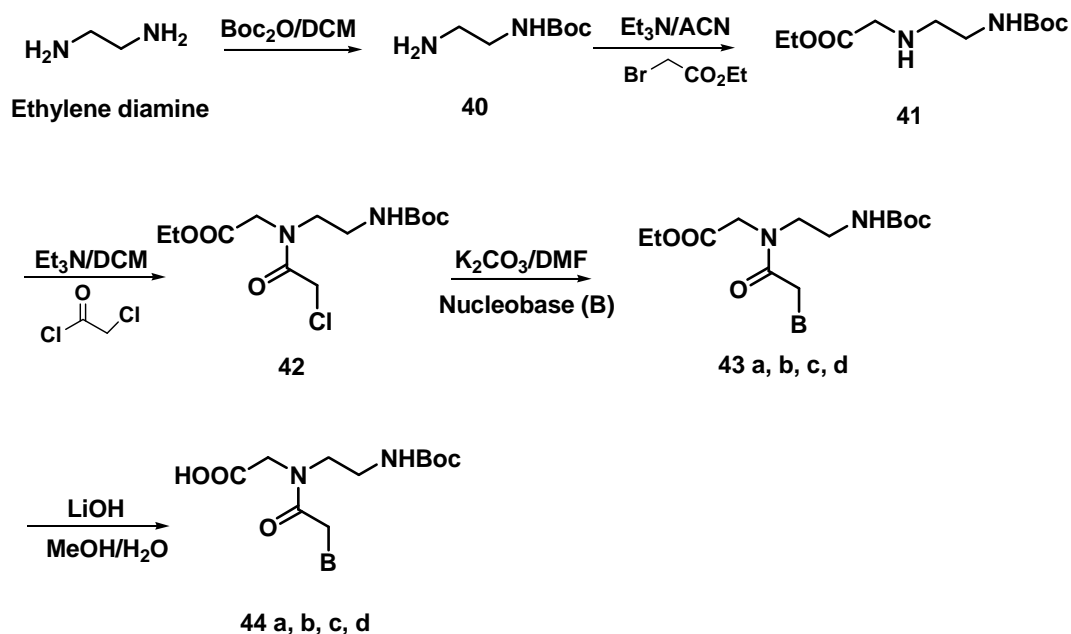


**Scheme 4** : Synthesis of (2*R*,3*R*) (2→3)*amp*PNA Monomer

The N-*t*-Boc protecting group in **30** was exchanged to N-cbz group by treating **30** with trifluoroacetic acid followed by reaction with benzylchloroformate to yield **31** with (2*S*,3*S*) configuration. The reaction of **31** with TBAF gave the (2*R*,3*S*) azido alcohol **32**, that was hydrogenated in presence of Raney-Ni to amine which was *in-situ* protected by *t*-butyloxy carbonyl group with di-*tert*-butyl dicarbonate yielded the **33** with (2*R*,3*S*) configuration. O-Mesylation of the secondary alcohol group in **33** gave the corresponding O-mesylate **34** in which S<sub>N</sub>2 displacement of mesyl group by azide group was done by reacting with NaN<sub>3</sub> to obtain **35** having (2*R*,3*R*) configuration. The reduction of 3*R*-azide to amine was followed by N-alkylation with ethylbromoacetate to yield compound **36** having (2*S*,3*R*) configuration that was acylated with chloroacetylchloride to give compound **37** having (2*R*,3*R*) configuration. The chloro compound was used for N-alkylation of thymine base to yield ester (2*R*,3*R*) **38** that was converted to the required monomer acid **39** by hydrolysis of ester using LiOH having (2*R*,3*R*) configuration.

### 2.3.6 Synthesis of *t*-Boc-protected *aeg*PNA monomers

The synthesis of unmodified *aeg*PNA monomers were carried out following the literature procedures.<sup>18</sup> The synthesis was started from the readily available 1,2-diaminoethane (Scheme 5) which was treated with *t*-butyl pyrocarbonate to give the mono N-protected amine **40** by using a large excess of 1,2-diaminoethane over the reagent in high dilution conditions. The di-Boc derivative obtained as a minor product, being insoluble in water, could be removed by filtration through celite. The *N*1-Boc-1,2-diaminoethane **40** was *N*-alkylated using ethylbromoacetate to aminoethylglycine derivative **41** which was further treated with chloroacetyl chloride to yield the corresponding chloro derivative **42**, this is the common intermediate for the preparation of all the PNA monomers (A, G, C & T) in good yield. Thymine was reacted with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **42** in presence of K<sub>2</sub>CO<sub>3</sub> to obtain the *N*-(Boc-aminoethylglycyl)-thymine ethyl ester **43a** in high yield.



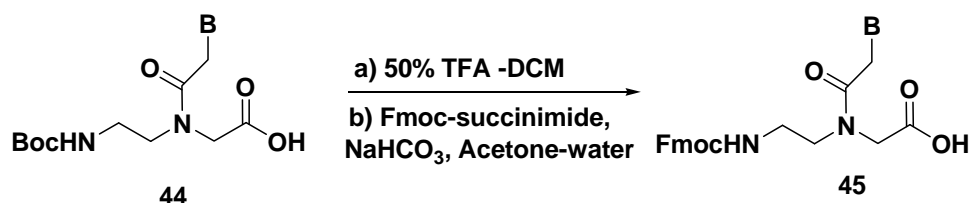
**44a**; B = Thymine (T)  
**44b**; B = Adenine (A)  
**44c**; B = N<sup>4</sup>cbz Cytosine  
**44d**; B = Guanine (G)

**43a**; B = Thymine (T)  
**43b**; B = Adenine (A)  
**43c**; B = N<sup>4</sup>cbz Cytosine  
**43d**; B = 2-amino-6-chloropurine

### Scheme 5: Synthesis of *t*-boc protected *aeg*PNA monomers

In the case of cytosine, the *N*4-amino group was protected as its benzyloxycarbonyl derivative, and was alkylated with *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **42** in presence of  $\text{K}_2\text{CO}_3$  to yield **43b**. Although adenine is known to undergo both *N*7- and *N*9- substitution, *N*7-alkylation was not observed when  $\text{K}_2\text{CO}_3$  was used as a base to get the product **43c**. The *N*-alkylation of 2-amino-6-chloropurine with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate was facile with  $\text{K}_2\text{CO}_3$  as the base and yielded the corresponding *N*-(Boc-aminoethylglycyl)-(2-amino-6-chloropurine)-ethyl ester **43d** in excellent yield. All compounds exhibited  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra consistent with the reported data. The ethyl esters were hydrolyzed in the presence of LiOH to give the corresponding acids **44a-d** ( scheme 5) which were further used for solid phase synthesis. There was no need for protection of the exocyclic amino groups of adenine and guanine, as these have been found to be unreactive under the conditions used for peptide coupling.

### 2.3.7 Synthesis of Fmoc-protected *aeg* PNA monomer



**Scheme 6:** Synthesis of Fmoc protected *aeg*PNA monomers

The *t*-Boc protected *aeg*PNA monomers **44 a-d** (Figure 5) synthesized as per literature procedure as were deprotected using 50% TFA-DCM and the amine formed was protected as Fmoc to give the Fmoc protected *aeg*PNA monomer building blocks **45 a-d** (Scheme 6), necessary for the synthesis of PNA oligomers through solid phase synthesis using Fmoc chemistry.

**2.4 Conclusion:** The design and synthesis of cationic pyrrolidyl PNAs **16, 17, 27** and **39** having chirality has been done.

The synthesis of these modified PNA monomers **16, 17, 27** and **39** (Figure 6) and *aeg* PNA monomers were done by following appropriate synthetic routes with required protecting groups. Monomers were characterised by <sup>1</sup>HNMR, <sup>13</sup>CNMR and with mass spectrometry

## 2.5 Experimental Section

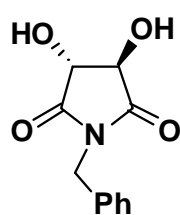
### General

All chemicals and reagents used were of laboratory grade or higher. NaH was obtained from Aldrich as 60% suspension in paraffin oil and was used as such. All the solvents used were purified according to the literature procedures (Perrin, 1989). All reactions were monitored by TLC. Usual work-up implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulphate and evaporation under vacuum. Column chromatography was performed for purification of compounds on 100-200 mesh silica gel (Spectrochem, India). TLCs were carried out on pre-coated silica gel GF<sub>254</sub> sheets (Merck 5554). TLCs were run in either petroleum ether with appropriate quantity of ethyl acetate or



dichloromethane with an appropriate quantity of methanol for most of the compounds. The compounds were visualized with UV light and /or by spraying with ninhydrin reagent subsequent to Boc deprotection (exposing to HCL vapours) and heating. Melting points of samples were determined in open capillary tubes using Buchi Melting point B-540 apparatus and are uncorrected. IR spectra were recorded on a Perkin Elmer 599B instrument. The  $^1\text{H}$  NMR (200 MHz),  $^{13}\text{C}$  NMR (50 MHz) spectra were recorded on Bruker ACF200 spectrometer fitted with an Aspect 3000 computer. The chemical shifts are reported in delta ( $\delta$ ) values and referred to internal standard TMS for  $^1\text{H}$  and deuterated NMR solvent chloroform-*d* itself for  $^{13}\text{C}$ . The optical rotation values were measured on Bellingham-Stanley Ltd, ADP220 polarimeter. Mass spectra were obtained by LCMS techniques.

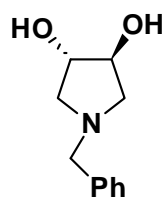
**(3*S*,4*S*)-1-benzyl-3,4-dihydroxypyrrolidine-2,5-dione (1)**<sup>16</sup>



To a stirred suspension of L-tartaric acid (45.0 g, 0.3 mol) in xylene (200 ml) at room temperature, benzyl amine (32.7 ml, 0.3 mol) was added and the resulting suspension was refluxed using a dean stark apparatus till the appropriate amount of water (10.8 ml, 0.6 mol) was collected during 7-8 hrs. After cooling, the resulting solid was filtered and washed with ethanol followed by acetone to obtain product **1** as a light yellow solid which was recrystallized from ethanol to get a white solid. This was used for next step without further purification. Yield: (59.8g, 90%).

**M.P.** = 196-198 °C;  $^1\text{H}$  NMR (200 MHz, DMSO-*d*<sup>6</sup>):  $\delta$  7.26-7.11 (m, 5H, Ph), 6.25-6.17 (m, 2H, 2 x OH), 4.53-4.37 (m, 2H, benzyl CH<sub>2</sub>), 4.32-4.24 (m, 2H, 2 x CH);  $^{13}\text{C}$  NMR (50 MHz, DMSO-*d*<sup>6</sup>):  $\delta$  174.8, 136.2, 128.8, 127.8, 74.8, 41.4.; **MS (EI)** *m/z* 221.07, Found 222.04 [M + H], 244.15 [M + Na]

**(3*S*,4*S*)-1-benzylpyrrolidine-3,4-diol (2)**<sup>17</sup>

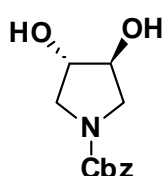


To an ice-cooled solution of **1** (11.0 g, 49.7 mmol) in THF (300 ml) under nitrogen, Lithium aluminium hydride (5.7 g, 149.1 mmol) was added. The mixture was refluxed for 12 hrs after which it was quenched with aqueous sodium sulphate solution. The resulting mixture was filtered and the salt obtained was taken in diethyl ether and stirred for 2-

3 hrs and filtered. This diethyl ether extraction was repeated four times. The combined filtrate was concentrated and the residue was purified by column chromatography to obtain compound **2** as a white solid. Yield: (7.0 g, 73%)

$[\alpha]_D^{25} = +34$  (*c* 1.0, MeOH); **M.P.** = 101-103 °C; **<sup>1</sup>H NMR** (200 MHz, CDCl<sub>3</sub>):  $\delta$  7.32-7.22 (m, 5H, Ph), 4.15-3.95 (m, 4H, 2xOH (2H) and 2xCH (2H)), 3.67-3.47 (m, PhCH<sub>2</sub>-diastereotopic,  $^2J_{HH} = 12.6$  Hz, 2H), 2.89 (dd, *J* = 5.8 Hz, 10.1 Hz, 2H, -CH<sub>2</sub>), 2.41 (dd, *J* = 3.9 Hz, 10.3 Hz, 2H, -CH<sub>2</sub>); **<sup>13</sup>C NMR** (50 MHz, CDCl<sub>3</sub>):  $\delta$  137.0, 129.3, 128.3, 127.4, 78.1, 60.3, 60.0; **MS (EI)** *m/z* 193.11, Found 194.13 [M + H]

### (3*S*,4*S*)-1-benzyloxy carbonyl pyrrolidine-3,4-diol (**3**)

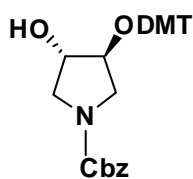


A solution of diol **2** (10gms, 42.19 mmol) in methanol is taken in hydrogenation flask and was added Pd/C. The reaction mixture was hydrogenated in a parr apparatus for 6 hrs at RT and H<sub>2</sub> pressure of 45-50psi. The catalyst was filtered off and the solvent was removed under reduced pressure to yield the residue of amine as light yellow colour oil. This residue was taken in acetone/water (50:50) mixture then added NaHCO<sub>3</sub> followed by benzyl chloroformate (7.2ml, 50.62mmol). The reaction mixture is stirred at RT for overnight. Acetone was evaporated from reaction mixture and more water is added then extracted with ethyl acetate. Ethyl acetate layer is washed with water and brine then dried over sodium sulphate and concentrated then subjected to column chromatographic purification to yield compound **3** as colour less oil. yield : (10.6g, 87%)

$[\alpha]_D^{25} = +42$  (*c* 1.0, MeOH); **<sup>1</sup>H NMR** (200 MHz, CDCl<sub>3</sub>):  $\delta$  7.25-7.32 (m, 5H, Ph), 5.06 (s, 2H, -CH<sub>2</sub>-Ph), 4.06 (s, 2H, O-CH-), 3.85 (br, 1H, OH), 3.58-3.66 (m, 2H, -N-CH<sub>2</sub>-), 3.31-3.37 (m, 2H, -N-CH<sub>2</sub>-); **<sup>13</sup>C NMR** (50 MHz, CDCl<sub>3</sub>):  $\delta$  155.68, 136.37, 128.54, 127.82, 75.23, 74.59, 67.24, 51.81, 51.48; **MS (EI)** *m/z* 237.12, Found 278.25 [M + K].

### (3*S*,4*S*)-1-benzyloxycarbonyl-3-(dimethoxytrityl)pyrrolidin-4-ol(**4**)

To an ice-cooled solution of diol **3** (7g, 12.98 mmol) in pyridine (50 ml) under nitrogen, dimethoxy trityl chloride (5.2gms, 15.5mmol) was added slowly and the reaction mixture was stirred in an ice bath for 15 min. The reaction mixture was

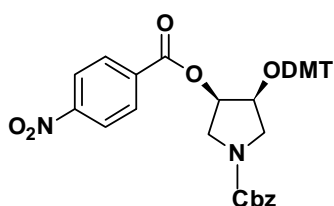


stirred at room temperature for 12 hrs. After the completion of reaction as observed by TLC, pyridine was evaporated and the residue was taken in water, extracted with ethyl acetate. Organic layer was washed with water followed by brine and dried over sodium sulphate. The solvent was evaporated under reduced pressure which upon column chromatographic purification afforded the compound **4** as light reddish oil. Yield : (9.7g , 61.3%)

$[\alpha]_D^{25} = + 125$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.26-7.35 (m, 8H, Ph), 7.15-7.19 (m, 8H, Ph), 6.81-6.85 (m, 2H, Ph), 5.13 (M, 2H,  $-\text{CH}_2\text{Ph}$ ), 3.8 (S, 6H,  $-\text{OMe}$ ), 3.5-3.6(m, 2H,  $-\text{O-CH-}$ ), 2.12-2.2 (m, 2H,  $-\text{N-CH}_2$ ), 1.62-1.64 (m, 2H,  $-\text{N-CH}_2$ )

$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  139.4, 129.1, 127.8, 113.1, 75.5, 67.07, 55.28, 51.91  
**MS (EI)** *m/z* 539, Found 542.59.

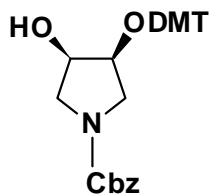
**(3*R*,4*S*)-1-benzyloxycarbonyl-4-(dimethoxytrityl)pyrrolidin-3-yl 4-Nitrobenzoate (5)**



To a stirred solution of **4** (7.5g, 10.9 mmol), triphenylphosphine (3.4gms, 13.08 mmol) and para nitro benzoic acid (2.18gms, 13.08mmol) in dry THF (90 ml) at 0°C under nitrogen, was added Diethyl azodicarboxylate (DIAD) (2.59ml, 13.08mmol) over a period of 15 min. Reaction mixture was stirred at RT for overnight. THF was evaporated under reduced pressure to get crude product **5**, which upon column chromatographic purification yielded compound **5** as light yellow colours solid. Yield: (6.0 g, 63%).

$[\alpha]_D^{25} = + 41$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.18-7.2 (m, 6H, Ar), 7.35-7.42 (m, 6H), 7.2-7.32 (m, 10H, Ar), 6.76-6.85 (m, 2H, Ar), 5.14-5.15 (m, 2H,  $-\text{CH}_2\text{-Ph}$ ), 3.8 (S, 6H,  $-\text{OCH}_3$ ), 3.42-3.54 (m, 2H,  $-\text{OCH}_2$ ), 2.1-2.2 (m, 2H,  $-\text{N-CH}_2$ ), 1.6-1.7 (m, 2H,  $-\text{N-CH}_2$ )

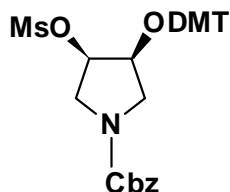
$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  158.5, 155.3, 147.2, 139.4, 130.1, 127.8, 113.2, 81.4, 74.7, 67.0, 55.2, 51.8 **MS (EI)** *m/z* 688.13, Found 688.63

**(3*R*,4*S*)-1-benzyloxycarbonyl-4-(dimethoxytrityl)pyrrolidin-3-ol (6)**

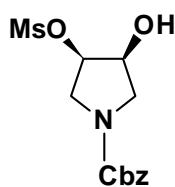
A stirred mixture of benzoate ester **5** (7.0 g, 12.98mmol) and sodium azide (8.4g, 129.8 mmol) in MeOH (200ml) under nitrogen was stirred at 80°C for 12hr. After cooling, MeOH was removed under reduced pressure. The residue was taken in ethyl acetate and washed with water, brine, dried over sodium sulphate and purified by column chromatography to yield **6** as light yellow oil. Yield: (4.19 g, 76%).

$[\alpha]_D^{25} = +38$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.33-7.35 (m, 8H, Ph), 7.27-7.32 (m, 8H, Ph), 5.13 (m, 2H,  $-\text{CH}_2-\text{Ph}$ ), 3.72-3.75 (m, 2H,  $-\text{OCH}-$ ), 3.5 (s, 6H,  $-\text{O}-\text{CH}_3$ ), 1.9-2.0 (m, 2H,  $-\text{CH}_2-\text{N}$ ), 1.3-1.4 (m, 2H,  $-\text{N}-\text{CH}_2$ )

$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  154.5, 136.0, 128.3, 79.3, 69.8, 67.0, 49.0, 48.8, 38.3

**(3*R*,4*S*)-1-benzyloxycarbonyl-4-(dimethoxytrityl)pyrrolidin-3-yl methanesulfonate (7)**

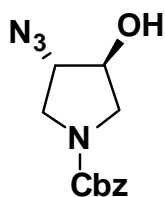
Compound **6** (5gms, 9.27mmol) was suspended in  $\text{CH}_2\text{Cl}_2$  and cooled to 0°C. Triethylamine (2.5ml, 18.54mmol) was added drop wise followed by Methane sulfonyl chloride (1.4ml, 19mmol). After 1h reaction mixture was diluted with  $\text{CH}_2\text{Cl}_2$  and washed with water, brine, dried over sodium sulphate and purified by column chromatography to yield **7** as light yellow oil and which was used further without purification.

**(3*R*,4*S*)-1-benzyloxycarbonyl-4-hydroxypyrrolidin-3-yl methanesulfonate (8)**

To a solution of compound **7** (6g, 9.7 mmol) in DCM (50 ml) was added Trifluoroacetic acid (TFA) (7.4ml, 97 mmol). The reaction mixture was stirred at room temperature for 1hr. TLC indicates the disappearance of starting material **7**. TFA was evaporated under reduced pressure and the residue was taken in water and extracted with ethylacetate. Ethyl acetate layer is washed with water, brine, dried over sodium sulphate and purified by column chromatography to yield product **8** as a white solid.

$[\alpha]_D^{25} = +34$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.29-7.35 (m, 5H, Ph), 5.12-5.14 (m, 2H,  $-\text{CH}_2\text{-Ph}$ ), 5.04-5.09 (m, 1H,  $-\text{CH-OH}$ ), 4.26-4.43 (m, 1H,  $-\text{CH-OMs}$ ), 3.58-3.78 (m, 2H,  $-\text{N-CH}_2$ ), 3.37-3.43 (m, 2H,  $-\text{N-CH}_2$ ), 3.12 (s, 3H,  $-\text{CH}_3$ );  $^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  154.5, 136.0, 128.3, 69.8, 67.0, 49.0, 38.3 **MS (EI)** *m/z* 315.76 Found 338.28 (M+Na)

**(3R,4R)-4-azido-1-benzyloxycarbonylpyrrolidine-3-ol (9)**

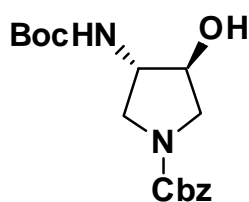


To a stirred solution of O-mesyl compound **8** (3gms, 9.5mmol) in dry DMF (30 ml) was added sodium azide (6.1gms, 95mmol) at room temperature under nitrogen atmosphere. The reaction mixture was heated at  $60^\circ - 70^\circ\text{C}$  for 18 hrs. The solvent was removed under reduced pressure and the residue was taken in water and extracted with ethyl acetate. Ethyl acetate layer is washed with water, brine, dried over sodium sulphate and purified by column chromatography to yield product **8** as a white solid. Yield: (2.0 g, 83%).

$[\alpha]_D^{25} = +39$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.35-7.38 (m, 5H, Ph), 5.13 (m, 2H,  $-\text{OCH}_2\text{-Ph}$ ), 3.95-4.25 (m, 1H,  $-\text{HC-OH}$ ), 3.67-3.7 (1H,  $-\text{HC-N}_3$ ), 3.45-3.49 (m, 4H,  $-\text{CH}_2\text{-N}$ )

$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  155.1, 136.3, 128.5, 127.9, 74.0, 67.3, 65.3, 52.1, 48.8; **MS (EI)** *m/z* 262.36 Found 285.03 (M+Na)

**(3R,4R)-1-benzyloxycarbonyl-4-N-boc-pyrrolidin-3-ol (10)**

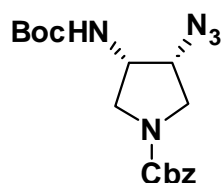


To a solution of azido-alcohol **9** (4g, 15.2 mmol) in ethyl acetate (15 ml) taken in a hydrogenation flask was added boc-anhydride (4.2ml, 18.3 mmol) and Raney-Ni (1.0 g). The reaction mixture was hydrogenated in a Parr apparatus for about 3 hrs at room temperature and at a hydrogen pressure of 40-45 psi during which the TLC indicates the disappearance of starting material. The catalyst was filtered off and the solvent was evaporated under reduced pressure. The crude product obtained was purified by column chromatography to yield product **10**. Yield: (4.0g, 79%)

$[\alpha]_D^{25} = + 53$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.26-7.37 (m, 5H, Ph), 5.12 (m, 2H, - $\text{CH}_2$ - Ph), 4.72-4.9 (br, 1H, - $\text{OH}$ ), 4.18-4.22 (m, 1H, - $\text{CH}$ -OH), 3.82-3.93 (m, 2H, - $\text{CH}_2$ -N), 3.62-3.65 (m, 1H, -CH-N), 3.19-3.23 (m, 2H, - $\text{NCH}_2$ ), 1.43 (m, 9H, *t*-boc)

$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  155.0, 136.5, 128.5, 67.0, 28.3; **MS (EI)** *m/z* 336.43  
Found 335.33

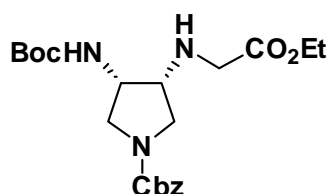
***tert*-butyl ((3*R*,4*S*)-4-azido-1-benzyloxycarbonylpyrrolidine-3-yl)(12)**



To a stirred solution of **10** (6g, 17.85mmol) and triethylamine (3ml, 21.42 mmol) in dry DCM (80 ml) at 0° C under nitrogen, was added methanesulfonyl chloride (1.65ml, 21.42 mmol) in DCM (10 ml) over a period of 15 min. The mixture was stirred for half an hour at 0° C. The reaction mixture was diluted with another 20 ml of DCM and washed successively with water, brine and dried over sodium sulphate to get crude O-mesyl product which was used directly without further purification. To a stirred solution of crude O-mesyl compound in dry DMF (30ml) was added sodium azide (11.6g, 178.5 mmol) at room temperature under nitrogen. The reaction mixture was heated at 60-70° C for 18 hrs. The solvent was removed under reduced pressure and the residue was taken in water and extracted with ethyl acetate two times. The ethyl acetate layer was washed with water, brine, dried over sodium sulphate and purified to get the corresponding azide **12**.

$[\alpha]_D^{25} = + 17$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.32-7.36 (m, 5H, Ph), 5.13-5.15 (m, 2H, - $\text{CH}_2$ -Ph), 4.13-4.15 (m, 1H, -CH- $\text{N}_3$ ), 3.63-3.65 (m, 1H, - $\text{CH}$ -NHboc), 3.62-3.64 (m, 2H, - $\text{CH}_2$ -N), 1.45-1.46 (m, 9H, *t*-boc); **MS (EI)** *m/z* ; 361.14  
Found 384.25 (M+Na)

**ethyl-2-(((3*S*,4*R*)-4-((*tert*-butoxycarbonyl)amino)-1-benzyloxycarbonyl-3-yl)aminoacetate (13)**

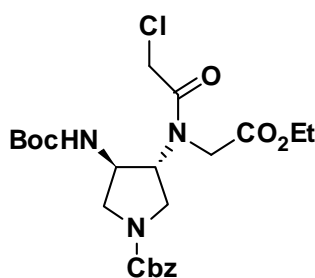


To azide **12** (3.7g, 9.58 mmol), taken in ethanol (20ml) in a hydrogenation flask, was added Raney-Ni (1.4 g) and the mixture was hydrogenated in a parr apparatus at 45 psi for about 3 hrs during which the TLC indicates disappearance of the starting azide. The solvent was filtered over a celite pad and

concentrated to get colourless oil. The oil was taken in dry acetonitrile (50 ml), to which was added DIPEA (2.0ml, 11.5 mmol) and the mixture was stirred under nitrogen at 0° C for 15 min followed by the addition of ethyl bromoacetate (1.05 ml, 9.58 mmol). The mixture was further stirred at room temperature for about 12 hrs. The solvent was evaporated, the residue was taken in ethyl acetate and washed with water, brine, dried over sodium sulphate and purified to get the compound **13** as a light yellow oil.

**ethyl-2-(N-((3*S*,4*R*)-4-((tert-butoxycarbonyl)amino)-1-benzyloxycarbonyl-3-yl)-2-chloroacetamido)acetate (**14**)**

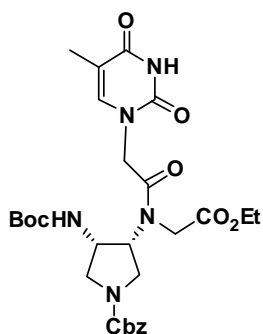
To an ice-cooled stirred solution of **13** (3.7 g, 8.25 mmol) in DCM (30ml) was added triethylamine (1.3ml, 9.9 mmol). The reaction mixture was stirred for 15 min followed by the addition of chloroacetyl chloride (0.8 ml, 9.9 mmol) and the reaction



mixture was continued to stir for another 10 hrs. The reaction mixture was diluted with 100 ml of DCM and washed with water, brine, dried over sodium sulphate and concentrated to obtain chloro compound **14** as a colourless liquid which was used without further purification.

**[(3*S*,4*R*)-1-benzyloxycarbonyl-4-N-boc-aminopyrrolidin-3-yl]-N-(thymine-1-acetyl)-glycine ethyl ester (**15**)**

To compound **14** (1.3 g, 2.48 mmol) taken in 20 ml of DMF was added potassium carbonate (1.34g, 7.4 mmol) followed by thymine (0.468g, 3.72 mmol) at room temperature. The reaction mixture was stirred for 18 hrs at room temperature. DMF was evaporated after which residue was diluted with 100ml of ethyl acetate and



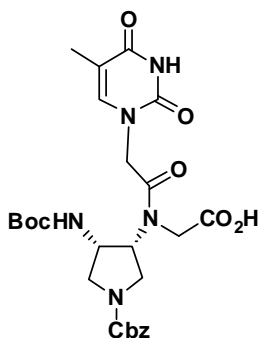
washed with water, brine and dried over sodium sulphate and purified by column chromatography to obtain compound **15** as white solid.

$[\alpha]_D^{25} = -16$  (*c* 1.0, CDCl<sub>3</sub>); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>):  $\delta$  7.32-7.38 (m, 5H, Ph), 7.10-6.95 (m, 1H, Thymine -NH), 5.13-5.15 (m, 2H, -CH<sub>2</sub>-Ph), 4.50-4.53 (m, 2H, -CH<sub>2</sub>-B), 4.13-4.25 (m, 2H, N-CH<sub>2</sub>-CO<sub>2</sub>Et), 3.71-3.78 (m, 2H, -OCH<sub>2</sub>), 2.88-2.91 (m, 4H, -NCH<sub>2</sub>), 2.10-2.15 (m, 1H, -NCH-), 1.93-1.95 (Thymine CH<sub>3</sub>), 1.39 (d, *J* = 9.5 Hz,

9H, *t*-Boc), 1.15-1.30 (m, 3H,  $-\text{CO}_2\text{CH}_2\text{CH}_3$ ).  $^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  169.7, 169.3, 167.7, 166.8, 164.4, 155.6, 151.3, 141.1, 137.9, 128.5, 128.3, 127.3, 127.2, 110.6, 62.1, 62.0, 61.2, 60.9, 60.3, 59.5, 56.6, 55.0, 47.5, 45.0. ; **MS (EI)**  $m/z$ , Cal 587 [M + H], Found; 610.72 [M + Na].

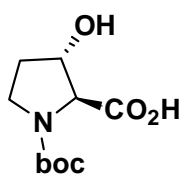
**[(3*S*,4*R*)-1-benzyloxycarbonyl-4-*N*-boc-aminopyrrolidin-3-yl]-*N*-(thymine-1-acetyl)-glycine (16)**

To an ice cooled solution of compound **15** (1g, 1.7 mmol) in 15ml of ethanol was added 2.5 ml of 1N sodium hydroxide solution in water and the reaction mixture was stirred in an ice bath for 30 min. Water (10ml) was added to the reaction mixture and solvent was removed on rotary evaporator to half of its volume. The remaining mixture was neutralised with Dowex  $\text{H}^+$  resin to pH 6 and concentrated to get 130 mg of monomer **16** as white solid.



$[\alpha]_{\text{D}}^{25} = 17$  ( $c$  1.0, DMSO);  $^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  10.26 (br s, 1H,  $-\text{COOH}$ ), 7.80-7.15 (m, 8H, fmoc), 7.10-6.85 (m, 1H, Thymine – NH), 4.60-3.98 (m, 6H,  $-\text{CH}_2$  &  $-\text{CH}$  fmoc,  $-\text{CH}$  ring,  $-\text{CH}_2-\text{CO}_2\text{H}$ ), 3.95-3.55 (m, 3H,  $-\text{CH}_2$ -Thymine,  $-\text{CH}$  ring), 3.45-3.10 (m, 2H, 2 x CH), 2.40 (t,  $J = 8.6$  Hz, 1H,  $-\text{CH}$ ), 2.10-1.90 (m, 1H,  $-\text{CH}$ ), 1.71 (s, 3H, Thymine  $-\text{CH}_3$ ), 1.39 (s, 9H, *t*-Boc).  $^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  175.6, 143.7, 141.1, 127.8, 127.7, 127.1, 125.1, 125.0, 119.9, 49.5, 47.0, 30.6, 29.6, 28.2, 17.5.; **MS (EI)**  $m/z$  559.26, Found 582.73 [M + Na].

***N*-tert butyloxycarbonyl L-*trans* 3-hydroxy proline (18)**

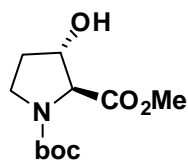


To a stirred solution of *trans*-3*S*-hydroxy-L-proline (3g, 22.9mmol) in dioxane/water (50:50) was added 2N NaOH solution and stirred for 5 min at RT. After cooling the reaction mixture to  $0^\circ\text{C}$ , di *t*-butyldicarbonate (6.3ml, 27.4mmol) was added and stirred for 1h. Dioxane was removed and more water is added to the reaction mixture then extracted with diethylether. The aqueous layer was cooled to  $0^\circ\text{C}$  and added ethylacetate. This mixture was acidified to pH 2 by adding sat.Potassium hydrogen sulphate solution while stirring. Organic layer was separated and aqueous layer was



extracted with ethyl acetate again, dried over sod. Sulphate and solvent was removed to get the desired product

**(2*S*,3*S*)-1-tert-butyloxycarbonyl 3-hydroxypyrrolidine 2-methylcarboxylate (19)**

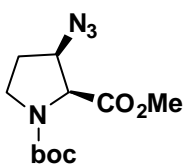


To a stirred solution of **18** (4.5 g, 19.48mmol), N-boc protected *trans*-3-hydroxy-L-proline and potassium carbonate (13.5g, 97.4mmol) in acetone was added dimethylsulphate (5.5ml, 58.44mmol) and was refluxed for 6h. Solvent was evaporated after which residue was diluted with ethyl acetate and washed with water, brine and dried over sodium sulphate and purified by column chromatography to obtain compound **19** as white solid. Yield (2.9g, 61%)

$[\alpha]_D^{25} = 49$  (*c* 1.0, DMSO);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  3.74 (m, s, 3H,  $-\text{OCH}_3$ ), 3.58-3.62 (m,  $-\text{CH}-\text{OH}$ ), 2.94-3.02 (m, 1H,  $-\text{CH}-\text{CO}_2\text{CH}_3$ ), 2.11-2.13 (m, 2H,  $-\text{CH}_2-\text{N}$ ), 1.94-1.96 (m, 2H,  $-\text{CH}_2-\text{N}$ ), 1.41-1.46 (m, 9H, *t*-boc)

$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  171.5, 153.6, 80.2, 75.1, 67.9, 62.2, 32.2, 28.4; **MS (EI)** *m/z* 245.83 Found 285.34 (M+Na)

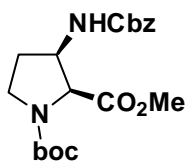
**(2*S*,3*R*)-1-tert-butyloxycarbonyl 3-azidopyrrolidine 2-methylcarboxylate (20)**



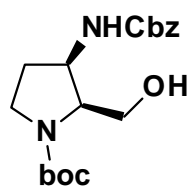
To a stirred solution of **19** (3.5 g, 14.28 mmol), triphenylphosphine (5.6gms, 21.4mmol) in dry THF (80 ml) at 0° C under nitrogen, was added Diethyl azodicarboxylate (DIAD) (4.2ml, 21.4 mmol) over a period of 15 min followed by diphenylphosphoryl azide(4.6ml, 21.4mmol). Reaction mixture was stirred at RT for overnight. THF was evaporated under reduced pressure to get crude product **20**, which upon column chromatographic purification yielded compound **20**. Yield: (2.15 g, 56%).

$^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  4.29-4.53 (m, 2H), 3.79 (s, 3H,  $-\text{OCH}_3$ ), 3.69-3.75 (m, 1H,  $-\text{CH}-\text{CO}_2\text{CH}_3$ ), 3.45-3.6 (m, 1H,  $-\text{CH}-\text{N}_3$ ), 2.06-2.21 (m, 2H,  $-\text{CH}_2$ ), 1.41-1.46 (m, 9H, *t*-boc)

$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  170.0, 153.4, 80.5, 62.0, 61.5, 52.1, 44.5, 30.0, 28.3, **MS (EI)** *m/z* 270.26 Found 270.04

**(2*S*,3*R*)-1-tert-butyloxycarbonyl-3-(((benzyloxy)carbonyl)amino)pyrrolidine-2-methylcarboxylate (21)**

To a solution of azide **20** (3.2gms, 11.85mmol) in methanol in hydrogenation flask was added Pd/C. The reaction mixture was hydrogenated in a parr apparatus for 6 hrs at RT and H<sub>2</sub> pressure of 45-50psi. The catalyst was filtered off and the solvent was removed under reduced pressure to yield the residue of amine as light yellow colour oil. This residue was taken in acetone/water (50:50) mixture then added NaHCO<sub>3</sub> (4.9g, 59.25mmol) followed by benzyl chloroformate (2.19ml, 15.4mmol). The reaction mixture is stirred at RT for overnight. Acetone was evaporated from reaction mixture and more water is added then extracted with ethyl acetate. Ethyl acetate layer is washed with water and brine then dried over sodium sulphate and concentrated then subjected to column chromatographic purification to yield compound **21** as colour less oil.

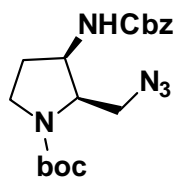
**(2*S*,3*R*)-1-tert-butyloxycarbonyl-3-(((benzyloxy)carbonyl)amino)-2-(hydroxymethyl)pyrrolidine (22)**

In a dried single neck round bottomed flask, ester **21** (3.1g, 8.2mmol) was dissolved in ethanol under nitrogen atmosphere. The solution was cooled to 0°C and was added Lithium borohydride (0.9g, 41.11mmol) over a period of 15min. The reaction mixture was stirred overnight at rt then diluted with ethylacetate. Ethyl acetate layer is washed with water and brine then dried over sodium sulphate and concentrated then subjected to column chromatographic purification to yield compound **22** as colour less oil.

$[\alpha]_D^{25} = +12$  (*c* 1.0, MeOH); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 7.32-7.35 (m, 5H, Ph), 5.09-5.11 (m, 2H, -CH<sub>2</sub>-Ph), 4.29-4.33 (m, 1H), 3.86-3.90 (m, 2H, -CH<sub>2</sub>-OH), 3.68-3.70 (br, 1H, -CH-NHCbz), 3.25-3.42 (m, -CH<sub>2</sub>-N), 2.10-2.13 (m, 2H, -CH<sub>2</sub>), 1.43 (m, 9H, *t*-Boc)

<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 156.3, 136.3, 128.5, 80.0, 66.8, 61.2, 58.7, 52.2, 44.1, 28.4; MS (EI) *m/z* 350.18 Found 373.55 (M+Na)

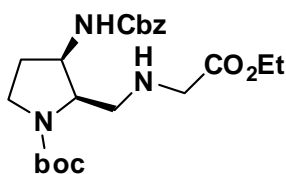
**(2*S*,3*R*)-1-*tert*-butyloxycarbonyl-3-(((benzyloxy)carbonyl)amino)-2-(azidomethyl)pyrrolidine (23)**



To a stirred solution of **22** (2.3 g, 6.5 mmol) and triethylamine (1.9 ml, 13.1 mmol) in dry DCM (30 ml) at 0° C under nitrogen, was added methanesulfonyl chloride (0.6ml, 7.8 mmol) in DCM (5 ml) over a period of 15 min. The mixture was stirred for half an hour at 0° C. The reaction mixture was diluted with another 20 ml of DCM and washed successively with water, brine and dried over sodium sulphate to get crude O-mesyl product which was used directly without further purification. To a stirred solution of crude O-mesyl compound in dry DMF (30ml) was added sodium azide (4.2g, 65 mmol) at room temperature under nitrogen. The reaction mixture was heated at 60 - 70° C for 18 h. The solvent was removed under reduced pressure and the residue was taken in water and extracted with ethyl acetate two times. The ethyl acetate layer was washed with water, brine, dried over sodium sulphate and purified to get the corresponding azide **23**. yield (1.58g, 64.5%).

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 7.44-7.31 (m, 5H, cbzPh), 5.10 (s, 2H, cbz -CH<sub>2</sub>), 5.08-4.97 (m, 1H, -NH), 4.15- 3.95 (m, 1H, -CH-NH), 3.85-3.70 (m, 1H -CH<sub>2</sub>), 3.68-3.27 (m, 3H, -CH<sub>2</sub>-N<sub>3</sub>, 1H -CH<sub>2</sub>), 3.24-3.01 (m, 2H, -CH<sub>2</sub>), 2.38-2.11 (m, 1H, -CH), 1.45 (s, 9H, *t*-boc), <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 155.8, 154.1, 136.0, 128.6, 128.3, 128.2, 79.9, 67.1, 51.7, 29.6, 28.4. ; MS (EI) *m/z* 375.19, Found 375.27

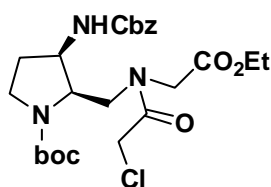
**(2*R*,3*R*)-*tert*-butyloxycarbonyl-3-(((benzyloxy)carbonyl)amino)-2-(((2-ethoxy-2-oxoethyl)amino)methyl)pyrrolidine (24)**



The azide **23** (1.3 g, 3.4 mmol) was taken in ethanol (20ml) in a hydrogenation flask, and was added Raney-Ni (0.7 g) and the mixture was hydrogenated in a parr apparatus at 45 psi for about 3h during which the TLC indicates disappearance of the starting azide. The solvent was filtered over a celite pad and concentrated to get colourless oil. The oil was taken in dry acetonitrile (50 ml), to which was added DIPEA (1.2ml, 6.9mmol) and the mixture was stirred under nitrogen at 0°C for 15 min followed by the addition of ethyl bromoacetate (0.4 ml, 3.4 mmol). The mixture was further stirred at room temperature for about 12 hrs. The solvent was

evaporated, the residue was taken in ethyl acetate and washed with water, brine, dried over sodium sulphate and purified to get the compound **24** as a light yellow oil.

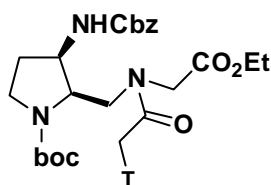
**(2*R*,3*R*)-tert-butylloxycarbonyl-3-(((benzyloxy)carbonyl)amino)-2-((2-chloro-N-(2-ethoxy-2-oxoethyl)amino)methyl)pyrrolidine (25)**



Compound 24 (1.3g, 3mmol) dissolved in DCM (30ml) was cooled to 0°C and added triethylamine (0.83 ml, 6 mmol). The reaction mixture was stirred for 15 min at 0°C followed by the addition of chloroacetyl chloride (0.36ml, 4.5 mmol)

and the stirring was continued for another 10 hrs. The reaction mixture was diluted with 100 ml of DCM and washed with water, brine, dried over sodium sulphate and concentrated to obtain chloro compound **25** as a colourless liquid which was used without further purification.

**[(2*R*,3*R*)-1-tertbutyloxycarbonyl-3-N-benzyloxycarbonyl-aminopyrrolidin-2-yl]-glycine ethylester (26)**

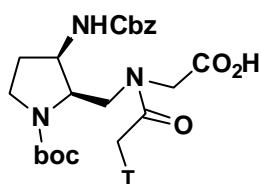


To a 100ml single neck round bottom flask equipped with magnetic pellet and nitrogen inlet were charged with compound **25** (0.9g, 1.76mmol) and dimethyl formamide (15ml) then added K<sub>2</sub>CO<sub>3</sub> (1.2g, 8.8mmol) thymine (0.33g,

2.64mmol). The resulting mixture was stirred at rt for overnight. The reaction mixture was diluted with water and extracted with ethyl acetate. Organic layer is then washed with water, brine, dried over sodium sulphate and concentrated and purified to obtain compound **26** white solid.

<sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 9.63-9.16 (m, 1H, NH), , 7.48-7.18 (m, 5H, cbzPh), 6.96 (s, 1H, thymine), 5.18-4.98 (m, 2H, cbz -CH<sub>2</sub>), 4.43-4.16 (m, 4H, -CH<sub>2</sub>-T, -CH<sub>2</sub>-CO<sub>2</sub>Et), 4.11-3.48 (m, 6H, 3 x CH<sub>2</sub>), 3.42-3.02 (m, 3H, -CH<sub>2</sub> ring, -CH), 2.71-2.42 (m, 1H, -CH), 1.89 (s, 3H, Thymine), 1.44 (s, 9H, *t*-boc), 1.36-1.21 (m, 3H, -CH<sub>3</sub>). <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 169.1, 167.8, 164.1, 156.4, 154.2, 151.1, 140.9, 136.3, 128.5, 128.2, 128.1, 110.8, 80.1, 66.8, 61.7, 50.6, 50.2, 49.9, 48.1, 47.9, 47.3, 28.4, 14.1, 12.2.; MS (EI) *m/z* 601.27, Found 624.48 [M + Na].

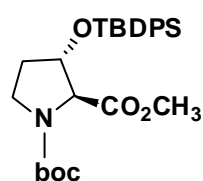
**[(2*R*,3*R*)-1-tertbutyloxyloxycarbonyl-3-*N*-benzyloxycarbonyl-aminopyrrolidin-2-yl]-glycine (27)**



To an ice cooled solution of compound **26** (0.7g, 1.16 mmol) in 10ml of ethanol was added 1.0 ml of 1N sodium hydroxide solution in water and the reaction mixture was stirred in an ice bath for 30 min. Water (10ml) was added to the reaction mixture and solvent was removed on rotary evaporator to half of its volume. The remaining mixture was neutralised with Dowex H<sup>+</sup> resin to pH 6 and concentrated to get corresponding acid as white solid. Yield: (610mg, 92 %).

$[\alpha]_D^{25} = -31$  (*c* 1.0, DMSO); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>):  $\delta$  11.28 (br s, 1H, -COOH), 7.92 (br s, 1H, NHCbz), 7.48-7.25 (m, 5H, cbzPh), 7.21 (s, 1H, thymine), 5.09-4.92 (m, 2H, cbz -CH<sub>2</sub>), 4.66-4.16 (m, 2H, -CH<sub>2</sub>-T), 4.03-3.70 (m, 4H, ring - CH<sub>2</sub>, -CH<sub>2</sub>-CO<sub>2</sub>H), 3.48-2.77 (m, 5H, 2 x CH<sub>2</sub> (4H), -CH), 2.41-2.22 (m, 1H, -CH), 1.70 (s, 3H, Thymine), 1.35 (s, 9H, *t*-boc). <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>):  $\delta$  168.3, 165.0, 156.5, 153.9, 151.5, 142.7, 137.4, 128.9, 128.3, 128.1, 108.6, 79.1, 65.9, 53.4, 52.8, 50.8, 48.7, 48.4, 42.3, 41.4, 28.6, 12.3.; MS (EI) *m/z* 573.24, Found 597.35 [M + Na].

**(2*S*,3*S*)-methyl-3-((tert-butyl-diphenylsilyl)oxy)-1-tertbutyloxycarbonyl pyrrolidine-2-carboxylate (28)**

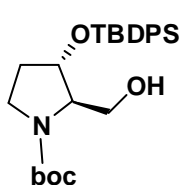


To a solution of compound **19** (3.2g, 13.06mmol) in DMF were added imidazole (1.77g, 26.12mmol), Hunig,s base (4.5ml, 26.12mmol) and catalytic amount of DMAP at 0°C under N<sub>2</sub> atmosphere. After 10min TBDPS-Cl (4.5ml, 4.5mmol) was added and stirred for overnight. Reaction mixture was diluted with water and extracted with EtOAc (x 3), washed with water and brine, dried over sod.sulphate, filtered and concentrated in vacuo. The crude silyl ether was purified by silica gel column chromatography.

$[\alpha]_D^{25} = +18$  (*c* 1.0, MeOH); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>):  $\delta$  7.63-7.66 (m, 4H, Ph), 7.40-7.42 (m, 6H, Ph), 4.38 (m, 1H, -CH-O-), 3.58 (s, 3H, -OCH<sub>3</sub>), 1.79-1.81 (m, 2H, -CH<sub>2</sub>), 1.61-1.63 (m, 2H, -CH<sub>2</sub>-), 1.40-1.48 (m, 9H, *t*-boc), 1.07 (s, 9H, tbdps)

$^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  135.8, 129.9, 127.7, 79.9, 67.9, 51.9, 32.8, 28.4, 28.3, 26.8, 19.1 MS (EI)  $m/z$  483.24 Found 506.46

**((2*R*,3*S*)-3-((*tert*-butyldiphenylsilyloxy)-1-*tert*-butyloxycarbonylpyrrolidin-2-yl)methanol (29)**

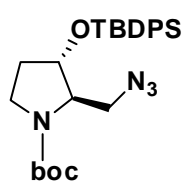


To a solution of compound **28** (4.2g, 8.69mmol) in THF & EtOH (1:1) was added  $\text{LiBH}_4$  (0.95g, 43.47mmol) slowly at  $0^\circ\text{C}$  for 15min under  $\text{N}_2$  atmosphere. The mixture was stirred for 12h. THF and EtOH were evaporated in vacuo. Residue was diluted with water and extracted with EtOAc (x 3), washed with water and brine, dried over sod.sulphate, filtered and concentrated in vacuo. The crude silyl ether was purified by silica gel column chromatography. Yield: (3.4 g, 87%).

$[\alpha]_{\text{D}}^{25} = +31$  ( $c$  1.0, MeOH);  $^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.63-7.69 (m, 4H, Ph), 7.43-7.66 (m, 6H, Ph), 4.38 (m, 1H), 3.58-3.62 (2H, m,  $-\text{CH}_2\text{-OH}$ ), 1.81-1.83 (m, 2H), 1.53-1.55 (m, 2H,  $-\text{CH}_2$ ), 1.43-1.46 (m, 9H, *t*-boc), 1.07 (s, 9H, tbdps)

$^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  135.7, 133.6, 129.9, 127.8, 127.7, 68.4, 64.9, 45.4, 32.9, 28.4, 26.9, 19.1 MS (EI)  $m/z$  455.25 Found 478.67 (M+Na)

**(2*R*,3*S*)-2-(azidomethyl)-3-((*tert*-butyldiphenylsilyloxy)-1-*tert*-butyloxy carbonylpyrrolidine (30)**



To a solution of alcohol **29** (3.3g, 7.25mmol) in  $\text{CH}_2\text{Cl}_2$  were added  $\text{NEt}_3$  (2ml, 14.5mmol) & Mesyl chloride (0.72ml, 9.42mmol) at  $0^\circ\text{C}$  under  $\text{N}_2$  atmosphere. The mixture was stirred for 3h, diluted with  $\text{CH}_2\text{Cl}_2$  washed with water, brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated in vacuo. The crude mesylate was employed directly in the next reaction.

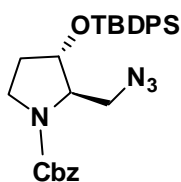
To a solution of crude mesylate in DMF was added  $\text{NaN}_3$  (4.7g, 72.5mmol) at RT. The mixture was heated to  $55^\circ\text{C}$  for 12h. Reaction mixture was diluted with water and extracted with EtOAc (x 3), washed with water and brine, dried over sod.sulphate, filtered and concentrated in vacuo. The crude silyl ether was purified by silica gel column chromatography.

Compound-5:  $[\alpha]_{\text{D}}^{25} = +39$  ( $c$  1.0, MeOH);  $^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.66-7.69 (m, 4H, Ar), 7.30-7.46 (m, 6H, Ar), 4.26 (2H,  $-\text{CH}_2\text{-N}_3$ ), 3.73-3.75 (m, 1H,  $-\text{O}$ -

CH-), 2.90-2.92 (m, 2H, -CH<sub>2</sub>-), 2.31-2.33 (m, 4H), 1.50-1.52 (m, 9H, *t*-boc), 1.10 (s, 9H, tbdps)

<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ; 135.7, 130.0, 128.5, 127.8, 66.9, 51.7, 58.6, 45.2, 32.7, 26.9, 19.2. MS (EI) *m/z* 480.26 Found 480.29

**(2*R*,3*S*)-2-(azidomethyl)-3-((*tert*-butyldiphenylsilyl)oxy)-1-benzyloxy carbonylpyrrolidine (31)**

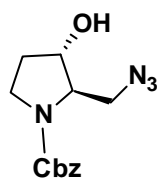


To a stirred solution of compound **30** (2.4g, 5mmol) in CH<sub>2</sub>Cl<sub>2</sub> was added TFA (1ml) at 0°C and the mixture was stirred for 2h at 0°C. CH<sub>2</sub>Cl<sub>2</sub> was evaporated to get the amine as TFA salt. This crude product was taken in acetone/water (1:1) then added NaHCO<sub>3</sub> (2.1g, 25mmol) followed by benzylchloroformate (0.92ml, 6.5mmol). The reaction mixture was stirred for 12h at room temperature. Acetone was evaporated in vacuo and the reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> and washed with water, brine and dried over sodium sulphate, filtered and concentrated in vacuo. The residue was purified by silica gel column chromatography which afforded **31** as colour less oil. Yield (1.46 g, 57%).

[α]<sub>D</sub><sup>25</sup> = + 54 (*c* 1.0, MeOH); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): 7.61-7.65 (m, 4H, Ar), 7.37-7.45 (m, 11H, Ar), 5.17 (s, 2H, -CH<sub>2</sub>-Ph), 3.73-3.75 (m, 1H, -CH-O-), 3.3 (m, 2H, -CH<sub>2</sub>-N<sub>3</sub>), 1.86-1.88 (m, 2H, -CH<sub>2</sub>), 1.6 (2H), 1.05 (s, 9H, tbdps)

<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ; 135.7, 130.0, 128.5, 127.8, 67.0, 51.7, 50.6, 45.5, 32.9, 26.9, 19.2 MS (EI) *m/z* 514.24 Found 537.72

**(2*R*,3*S*)-2-(azidomethyl)-1-benzyloxycarbonylpyrrolidin-3-ol (32)**



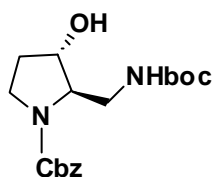
To a solution of compound **31** (3g, 5.8mmol) in THF was added TBAF (2ml) at rt. The mixture was stirred for 4h, quenched with saturated NH<sub>4</sub>Cl and extracted with ethylacetate and washed with water, brine and dried over sodium sulphate, filtered and concentrated in vacuo. The residue was purified by silica gel column chromatography which afforded **32** as colour less oil.

[α]<sub>D</sub><sup>25</sup> = + 9 (*c* 1.0, MeOH); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 7.36-7.20 (m, 5H, Ph), 5.31 (br s, 1H, -NH), 4.30-4.17 (m, 1H, -CH-OH), 4.14-3.97 (m, 1H, -CH-N), 3.59

(s, 2H, -CH<sub>2</sub>-Ph), 3.05 (br s, 1H, -OH), 2.86-2.67 (m, 2H, -CH<sub>2</sub>), 2.65-2.45 (m, 2H, -CH<sub>2</sub>),

<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 155.8, 137.7, 128.8, 128.1, 127.0, 79.2, 69.8, 60.7, 60.0, 58.0, 52.1, 28.2. **MS (EI)** *m/z* 276.12 Found 299.26 (M+Na)

***tert*-butyl-(((2*R*,3*S*)-3-hydroxy-1-benzyloxycarbonylpyrrolidin-2-yl)methyl) carbamate (33)**

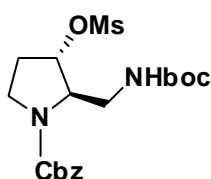


To a solution of azido-alcohol **32** (2.3g, 8.3 mmol) in ethyl acetate (15 ml) taken in a hydrogenation flask was added boc anhydride (2.5ml, 10.8 mmol) and Raney-Ni (1.0 g). The reaction mixture was hydrogenated in a parr apparatus for about 3 hrs at room temperature and at a hydrogen pressure of 40-45 psi during which the TLC indicates the disappearance of starting material. The catalyst was filtered off and the solvent was evaporated under reduced pressure. The crude product obtained was purified by column chromatography to yield product **33**. Yield (2.21g, 76%)

[α]<sub>D</sub><sup>25</sup> = + 37 (*c* 1.0, CHCl<sub>3</sub>); **M.P.** = 130-132 °C; <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 7.40-7.20 (m, 5H, Ph), 5.10-4.87 (m, 2H, -NH), 4.50-4.27 (m, 1H, -CH-NH), 3.65 (s, 2H, -CH<sub>2</sub>-Ph), 3.01 (s, 3H, CH<sub>2</sub>), 2.97-2.80 (m, 2H, -CH<sub>2</sub>), 2.49 (dd, *J* = 7.5 Hz, *J* = 9.2 Hz, 2H, -CH<sub>2</sub>), 1.44 (s, 9H, *t*-Boc).;

<sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 155.3, 137.5, 128.6, 128.3, 127.2, 79.8, 79.0, 59.6, 58.8, 56.5, 51.2, 37.7, 28.2. ; **MS (EI)** *m/z* 350.18 Found 589.13 (M+K)

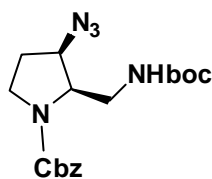
***tert*-butyl-(((2*R*,3*S*)-3-O-mesyl-1-benzyloxycarbonylpyrrolidin-2-yl)methyl) carbamate (34)**



To a solution of alcohol **33** (2.7g, 7.7mmol) in CH<sub>2</sub>Cl<sub>2</sub> were added Et<sub>3</sub>N (2.14ml, 15.4mmol) & mesyl chloride (0.8ml, 10.01mmol) at 0°C under N<sub>2</sub> atmosphere. The mixture was stirred for 3h, diluted with CH<sub>2</sub>Cl<sub>2</sub> washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated in vacuo. The crude mesylate was employed directly in the next reaction.



**tert-butyl-(((2*R*,3*S*)-3-azido-1-benzyloxycarbonylpyrrolidin-2-yl)methyl) carbamate (35)**

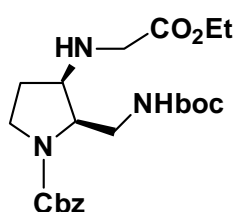


To a stirred solution of crude O-mesyl compound **34** (2.6g, 6.08mmol) in dry DMF (30ml) was added sodium azide (3.95 g, 60.8 mmol) at room temperature under nitrogen. The reaction mixture was heated at 60-70° C for 18 hrs. The solvent was removed under reduced pressure and the residue was taken in water and extracted with ethyl acetate two times. The ethyl acetate layer was washed with water, brine, dried over sodium sulphate and purified to get the corresponding azide **35**.

$[\alpha]_D^{25} = -12$  (*c* 1.0, MeOH);  $^1\text{H NMR}$  (200 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.40-7.20 (m, 5H, Ph), 5.10-4.75 (m, 1H, -NH), 4.15-3.90 (m, 1H, -CH-NH), 3.85-3.70 (m, 1H, -CH-N<sub>3</sub>), 3.60 (s, 2H, -CH<sub>2</sub>-Ph), 3.04 (dd, *J* = 6.8 Hz, *J* = 10.2 Hz, 1H), 2.84 (dd, *J* = 6.6 Hz, *J* = 9.8 Hz, 1H), 2.54-2.32 (m, 2H, -CH<sub>2</sub>), 1.45 (s, 9H, *t*-Boc).;

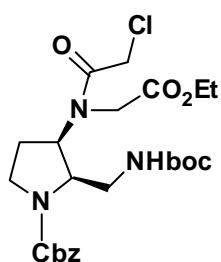
$^{13}\text{C NMR}$  (50 MHz,  $\text{CDCl}_3$ ):  $\delta$  154.9, 137.7, 128.6, 128.2, 127.2, 79.7, 66.6, 59.4, 58.0, 57.8, 56.5, 28.2. ;

**ethyl-2-(((2*R*,3*R*)-2-(((tert-butoxycarbonyl)amino)methyl)-1-benzyloxy carbonylpyrrolidin-3-yl)amino)acetate (36)**



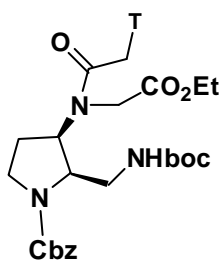
The azide **35** (1.3 g, 3.46 mmol) was taken in ethanol (20ml) in a hydrogenation flask, and was added Raney-Ni (0.8 g) and the mixture was hydrogenated in a Parr apparatus at 45 psi for about 3 hrs during which the TLC indicates disappearance of the starting azide. The solvent was filtered over a celite pad and concentrated to get colourless oil. The oil was taken in dry acetonitrile (50 ml), to which was added DIPEA (1.2ml, 6.93 mmol) and the mixture was stirred under nitrogen at 0° C for 15 min followed by the addition of ethyl bromoacetate (0.4 ml, 3.46 mmol). The mixture was further stirred at room temperature for about 12 hrs. The solvent was evaporated, the residue was taken in ethyl acetate and washed with water, brine, dried over sodium sulphate and purified to get the compound **36** as a light yellow oil.

**2-(N-((2R,3R)-2-(((tert-butoxycarbonyl)amino)methyl)-1-benzyloxycarbonyl pyrrolidin-3-yl)-2-chloroacetamido)acetate (37)**



To compound **36** (1.1 g, 3.16 mmol) in dry DCM (30 ml), to which was added NEt<sub>3</sub> (0.87ml, 6.32 mmol) and the mixture was stirred under nitrogen at 0°C for 15 min followed by the addition of chloro acetyl chloride (0.37ml, 4.74 mmol). The mixture was further stirred at room temperature for about 12 hrs. The solvent was evaporated, the residue was taken in ethyl acetate and washed with water, brine, dried over sodium sulphate and purified to get the compound **37** as a light yellow oil.

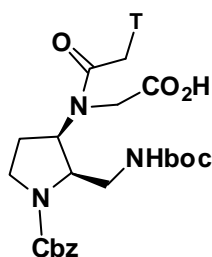
**2-(N-((2R,3R)-2-(((tert-butoxycarbonyl)amino)methyl)-1-benzyloxycarbonyl pyrrolidin-3-yl)-2-thymine-1-acetamido)acetate (38)**



Thymine (0.64 g, 5.08 mmol) was suspended in DMF (30ml) and was added K<sub>2</sub>CO<sub>3</sub> (0.70g, 27.3 mmol). The reaction mixture was stirred for 30 min. Reaction mixture was cooled to 0°C and a solution of compound **37** (1.3g, 2.54mmol) in DMF added. The reaction mixture was continued to stir for another 10 hrs. The reaction mixture was diluted with water and extracted with ethylacetate and the organic layer is washed with water, brine, dried over sodium sulphate and purified to obtain compound **38** as white solid. Yield: (2.6 g, 85 %).

$[\alpha]_D^{25} = -37$  (*c* 1.0, DMSO); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 7.38-7.15 (m, 5H, Ph), 7.10-6.95 (m, 1H, Thymine –NH), 4.90-3.93 (m, 8H, 2 x CH (2H), ester –CH<sub>2</sub>, –CH<sub>2</sub>-CO<sub>2</sub>Et, –CH<sub>2</sub>-Thymine), 3.70-3.40 (m, 2H, –CH<sub>2</sub>-Ph), 3.35-3.10 (m, 1H), 2.90-2.55 (m, 2H, 2 x CH), 2.45-2.15 (m, 2H, 2 x CH), 1.90 (d, *J* = 8.5 Hz, 3H, Thymine CH<sub>3</sub>), 1.41 (d, *J* = 9.5 Hz, 9H, *t*-Boc), 1.30-1.15 (m, 3H, –CO<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 169.7, 169.3, 167.7, 166.8, 164.4, 155.6, 151.3, 141.1, 137.9, 128.5, 128.3, 127.3, 127.2, 110.6, 62.1, 62.0, 61.2, 60.9, 60.3, 59.5, 56.6, 55.0, 47.5, 45.0. MS (EI) *m/z* 601.27, Found 624.52.

**2-(N-((2R,3R)-2-(((tert-butoxycarbonyl)amino)methyl)-1-benzyloxycarbonyl pyrrolidin-3-yl)-2-thymine-1-acetamido)acetic acid (39)**



To an ice cooled solution of compound **38** (0.6g, 0.99 mmol) in 10ml of ethanol was added 1.5 ml of 1N sodium hydroxide solution in water and the reaction mixture was stirred in an ice bath for 30 min. Water (10ml) was added to the reaction mixture and solvent was removed on rotary evaporator to half of its volume. The remaining mixture was neutralised with Dowex H<sup>+</sup> resin to pH 6 and concentrated to get corresponding acid as white solid. Yield: (0.5 g, 85 %).

$[\alpha]_D^{25} = -7$  (c 1.0, DMSO); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>): δ 11.31 (br s, 1H, -COOH), 7.31-7.37 (m, 6H, 5 Ph & 1H Thymine), 4.00-4.62 (m 6H, -CH<sub>2</sub>-N), 3.42-2.85 (m, 5H, 2 -CH<sub>2</sub>, -CH-NH), 2.41-2.25 (m, 1H, -CH), 1.71 (s, 3H, Thymine), 1.36 (s, 9H, *t*-boc). <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>): δ 173.4, 170.8, 168.4, 167.6, 165.0, 156.5, 153.9, 151.4, 144.3, 142.6, 141.2, 128.1, 127.6, 125.6, 120.6, 108.7, 79.1, 65.8, 60.3, 55.3, 53.0, 50.3, 48.3, 47.2, 28.5, 25.6, 21.2, 14.5, 12.3.; MS (EI) *m/z* 573.24, Found 574.41.

## 2.6 References

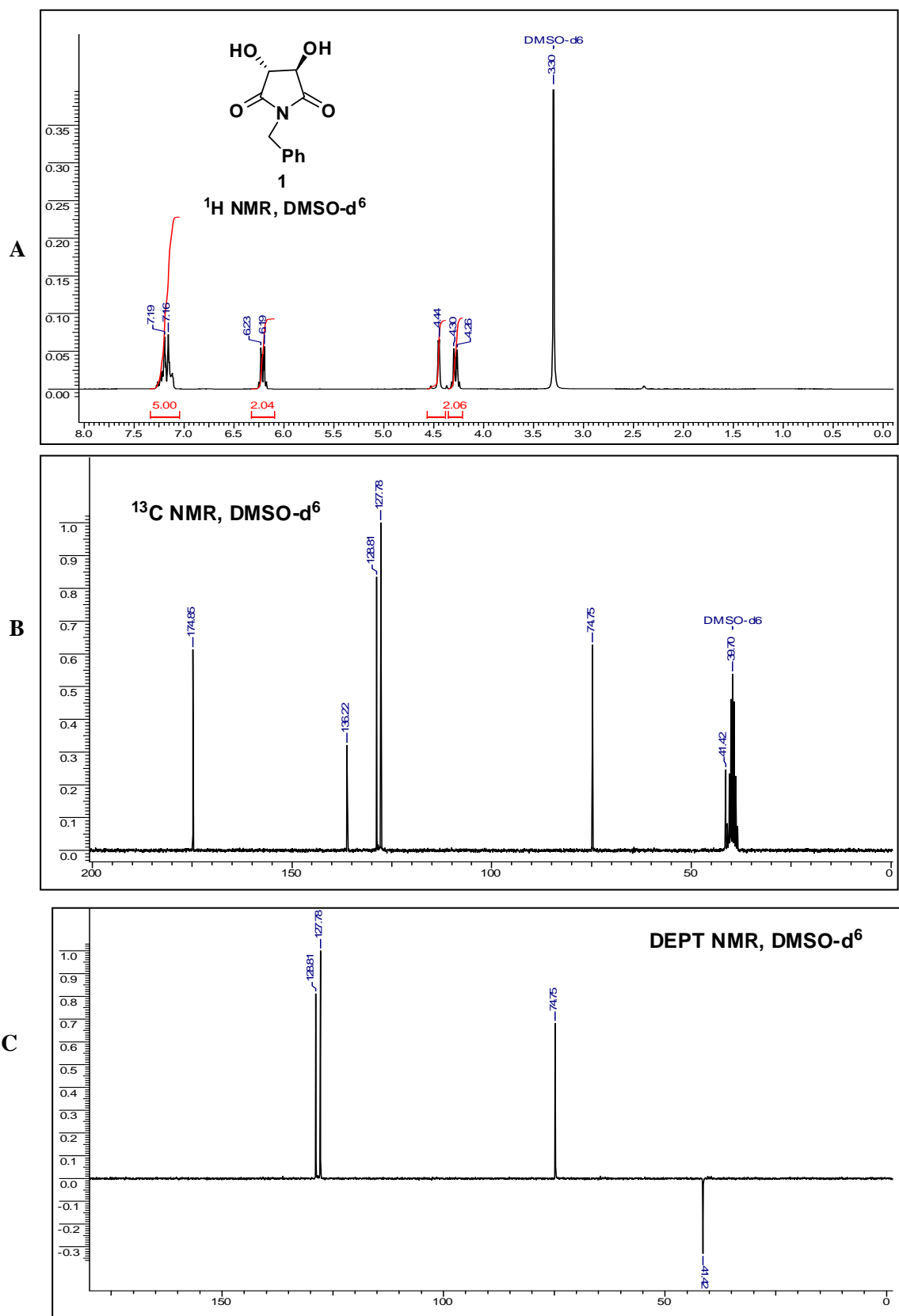
1. (a) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543. (b) Noble, S. A.; Bonham, M. A.; Bisi, J. E.; Bruckenstein, D. A.; Brown P. H.; Brown, S. C. *Drug Dev. Res.* **1995**, *34*, 184.
2. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci.* **1978**, *75*, 280.
3. (a) Mesmaekar, A. D.; Haner, R.; Martin, P.; Moser, E. H. *Acc. Chem. Res.* **1995**, *28*, 366. (b) Milligan, J.; Matteucci, M. D.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 1921. (c) Kurreck, J.; Wyszko, E.; Gillen, C.; Erdmann, V. A. *Nucleic Acid Res.* **2002**, *30*, 1911. (d) Sazani, P.; Shin-Hong Kong, Maier, M.; Weic, D.; Summerton, J.; Manoharan, M.; Kole, R. *Nucleic Acid Res.* **2001**, *29*, 3965.
4. (a) Nielsen P. E.; Egholm, O.; Berg R. H.; Buchardt, O. *Science* **1991**, *254*, 1497. (b) Hanvey, J. C.; Peffer, N. J.; Bissi, J. E.; Thomson, S. A.; Cadilla, R.;

- Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G. *Science* **1992**, 258, 1481. (c) Hamilton, S.; Iyer, M.; Norton, J.; Corey, D. *Bioorg. Med. Chem. Lett.* **1996**, 6, 2897. (d) Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioconjugate Chem.* **1994**, 5, 3.
5. (a) Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Nordén, B.; Gräslund, A. *J. Am. Chem. Soc.* **1996**, 118, 5544. (b) Knudsen, H.; Nielsen, P. E. *Nucleic Acid Res.* **1996**, 24, 494.
6. (a) Orum, H.; Nielsen, P. E.; Jorgensen, M.; Larsson, C.; Stanley, C.; Koch, T.; *Biotechniques* **1995**, 19, 472. (b) Kuhn, H.; Demidov, V. V.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *J Mol. Biol.* **1999**, 286, 13337. Ishihara, T.; Corey, D. R. *J. Am. Chem. Soc.* **1999**, 121, 2012.
7. (a) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science*, **1995**, 270, 1838. (b) Leijon, M.; Graslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry*, **1994**, 33, 9820. (c) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science*, **1994**, 265, 777.
8. Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem.* **2000**, 4, 931 and the references cited there in.
9. Kumar, V.A.; Ganesh, K. N. *Acc. chem. research.* **2005**, 38, 404.
10. Egholm, M.; Buchardt, O.; Christensen, L.; Beherns, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, 365, 566.
11. Govindaraju, T.; Kumar V. A.; Ganesh K. N. *Chem. Commun.* **2004**, 860 .
12. a) Myers, M. C.; Witschi, M. A.; Larionova, N. V.; Franck, J. M.; Haynes, R. D.; Hara, T.; Grajkowski, A.; Appella, D. H. *Org. Lett.* **2003**, 5, 2695. (b) Pokorski, J. K.; Witschi, M. A.; Purnell, B. L.; Appella, D. H. *J. Am. Chem. Soc.* **2004**, 126, 15067.
13. Yeheskiely, E.; Slaitas, A. *Eur. J. Org. Chem.* **2002**, 2391.
14. (a) Tan, T. H.; Hickman, D. T.; Morral, J.; Beadham, I. G.; Micklefield J. *Chem. Commun.* **2004**, 5, 516. (b) Hickman, D.T.; Tan, T. H.; Morral, J.; King, P. M.; Cooper, M. A.; Micklefield, J. *Org. Biomol. Chem.* **2003**, 1, 3277. (c) Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996**, 6, 1083.

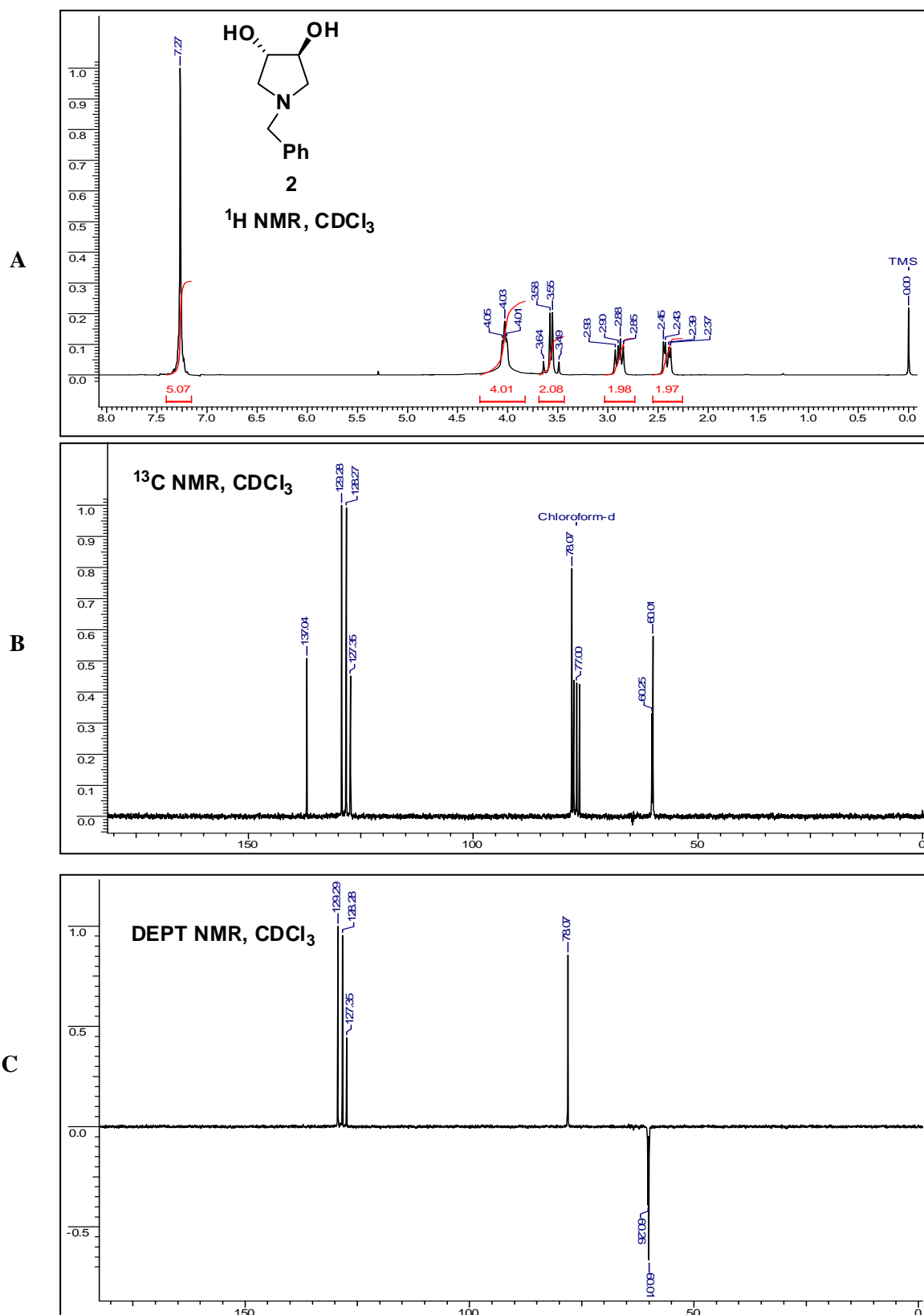
15. (a) Schoning, K.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science*, **2000**, *290*, 1347. (b) Govindaraju, T.; Kumar, V. A. *Tetrahedron*, **2006**, *62*, 2321.
16. (a) Rejman, D.; Kocalka, P.; Budesinsky, M.; Pohl, R.; Rosenberg, I.: *Tetrahedron*, **2007**, *63*, 1243. (b) Gonsalves, A. R.; Serra M. E. S.; Murtinho D.; Silva, V. F.; Beja, A. M.; Paixao, J. A.; Silva, M. R.; Veiga, L. A. *J. Mol. Cat. A: Chem.*, **2003**, *195*, 1.
17. Arakawa, Y; Yoshifuji, S. *Chem. pharm. Buli.* **1991**, *39*, 2219.
18. Meltzer, P. C.; Liang, A. Y.; Matsudaira, P. *J. Org. Chem.* **1995**, *60*, 4305.
19. Peter E. Nielsen and Michael Egholm *Current Issues Molec. Biol.* **1999** *1*(2): 89-104.

**2.7 : APPENDIX**

<b>Entry</b>	<b>Page No</b>
<sup>1</sup> H, <sup>13</sup> C , DEPT NMR and Mass spectra of compounds <b>1-39</b>	74

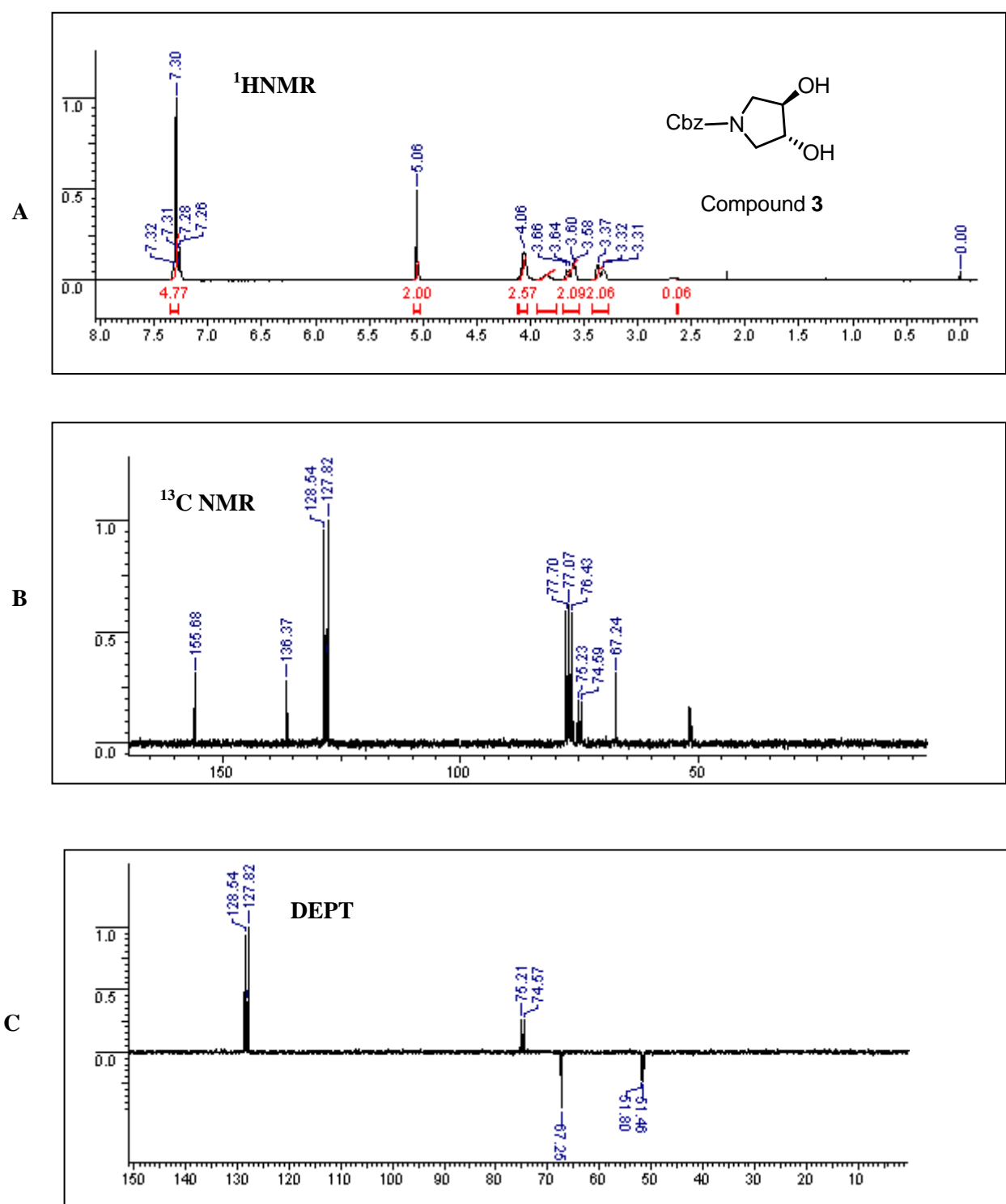


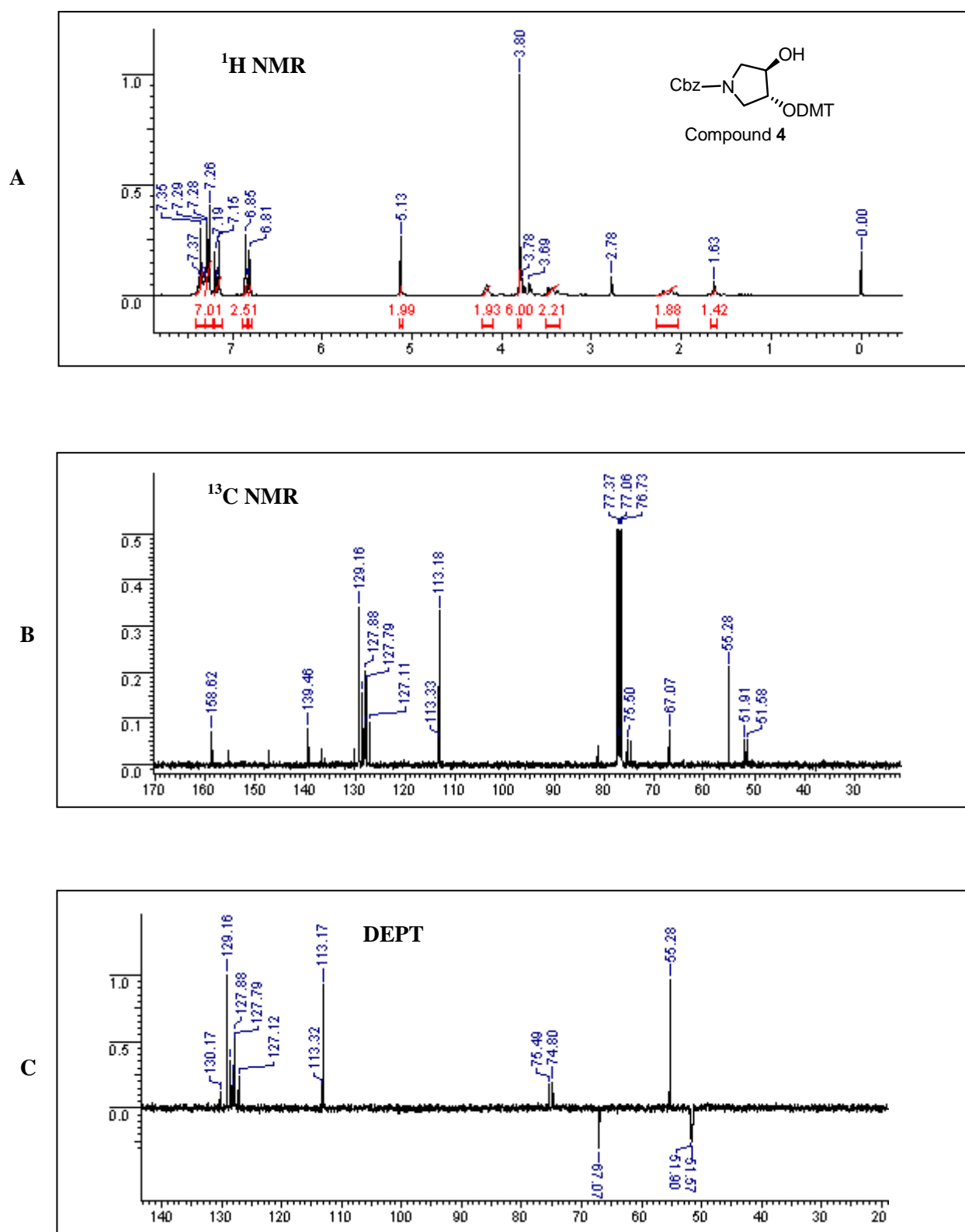
**Figure 1: (A) <sup>1</sup>H NMR of compound 1 (B) <sup>13</sup>C NMR of compound 1 (C) DEPT of compound 1**



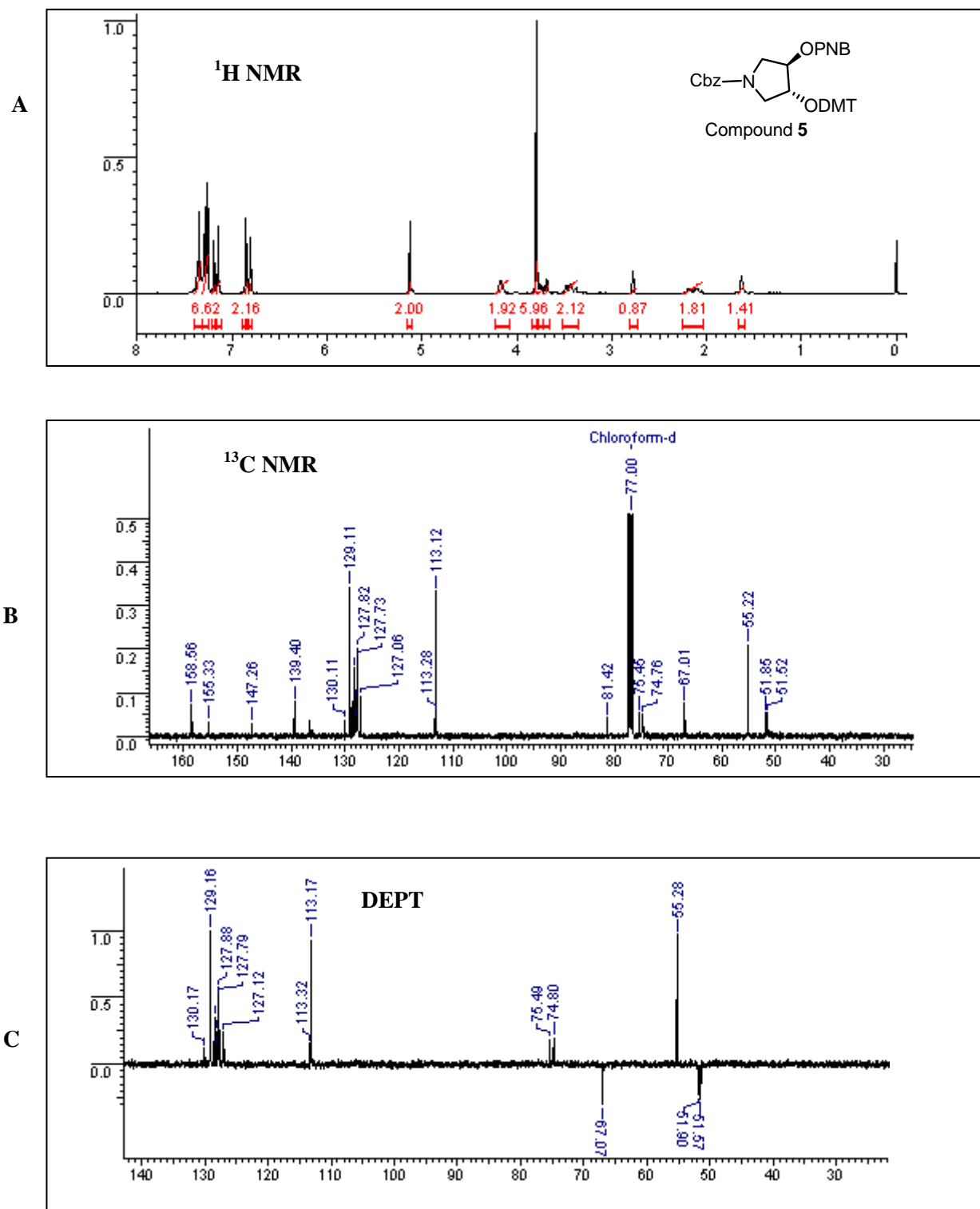
**Figure 2:** (A)  $^1\text{H NMR}$  of compound **2** (B)  $^{13}\text{C NMR}$  of compound **2** (C) DEPT of compound **2**



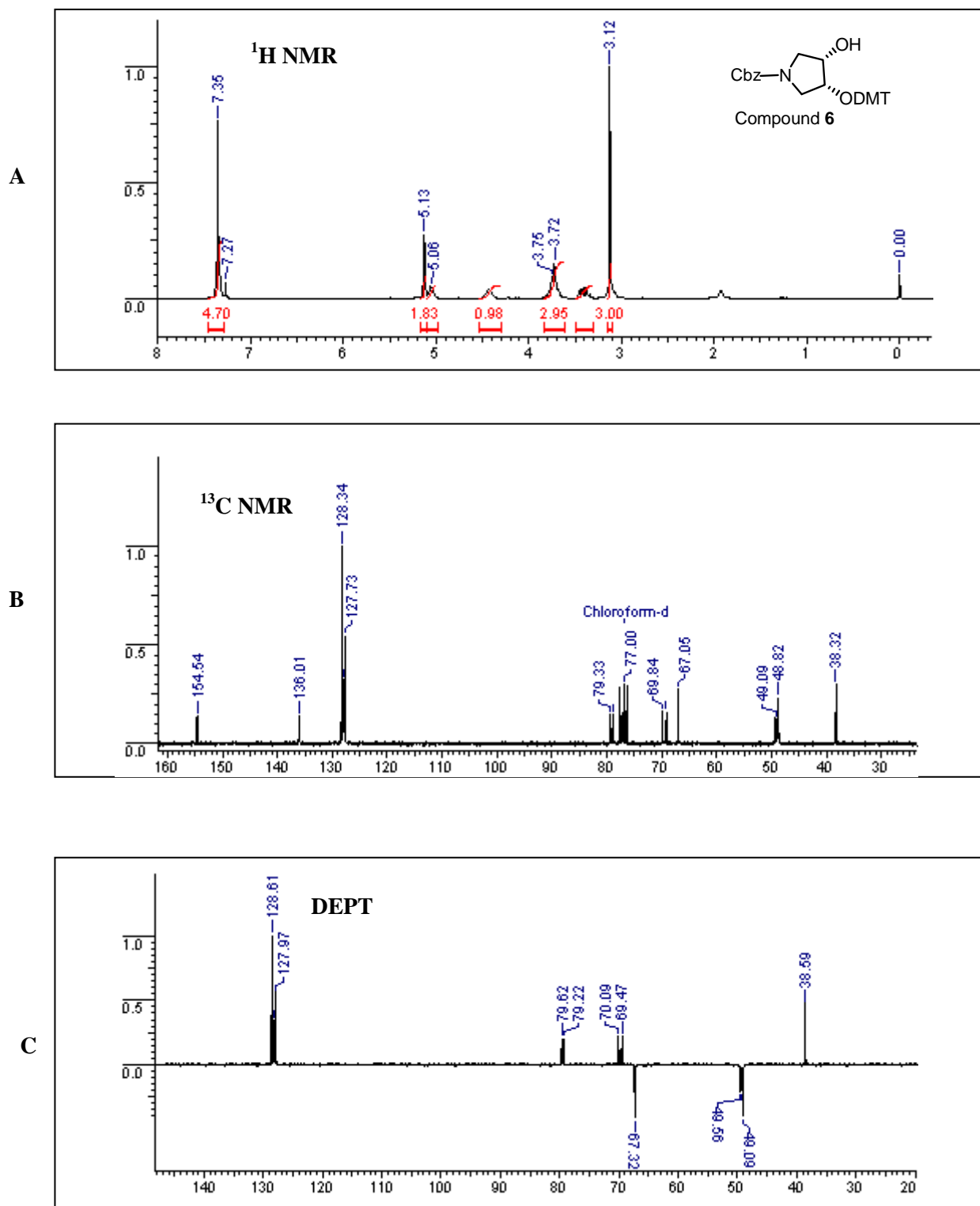




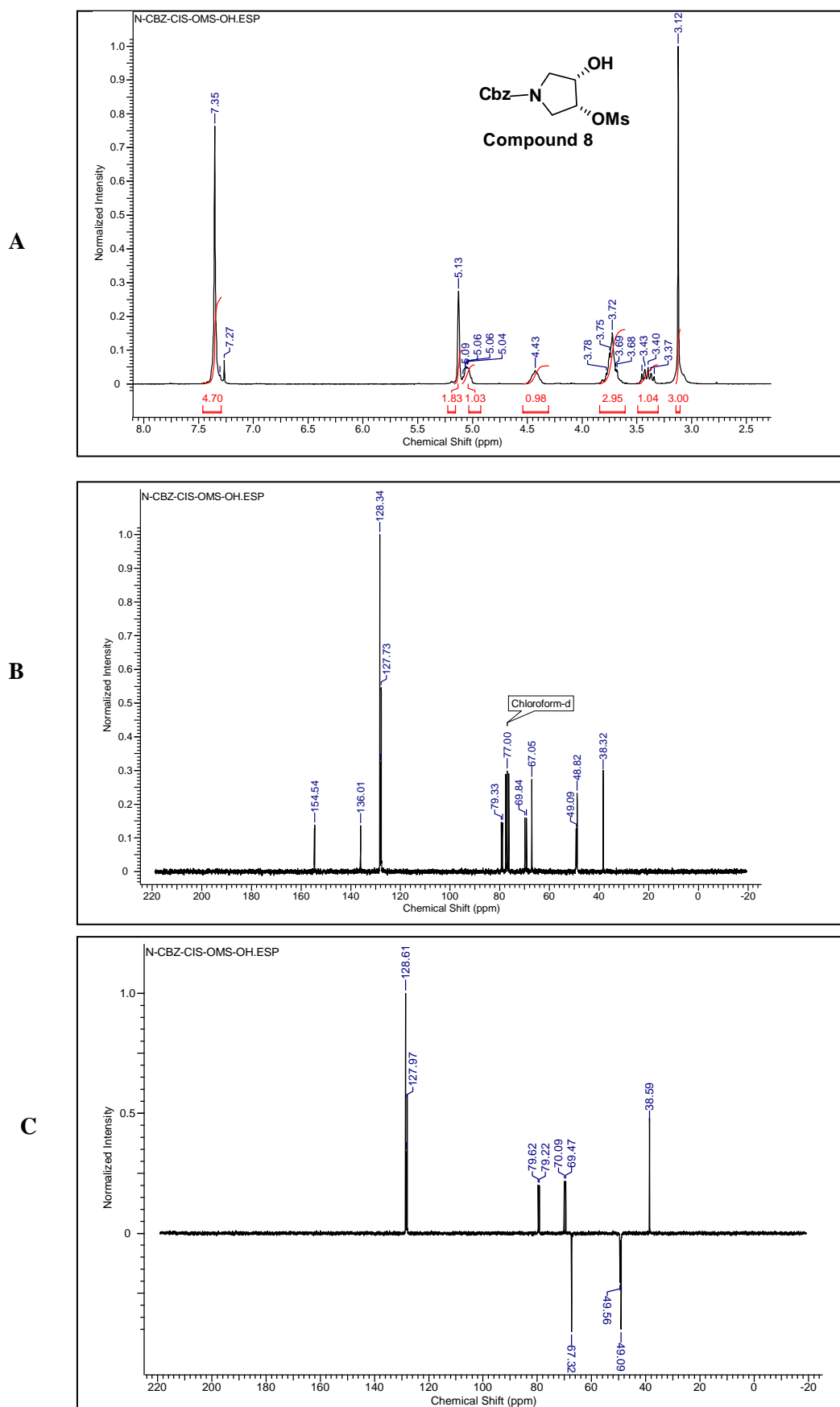
**Figure 4:** (A) <sup>1</sup>H NMR of compound 4 (B) <sup>13</sup>C NMR of compound 4 (C) DEPT of compound 4

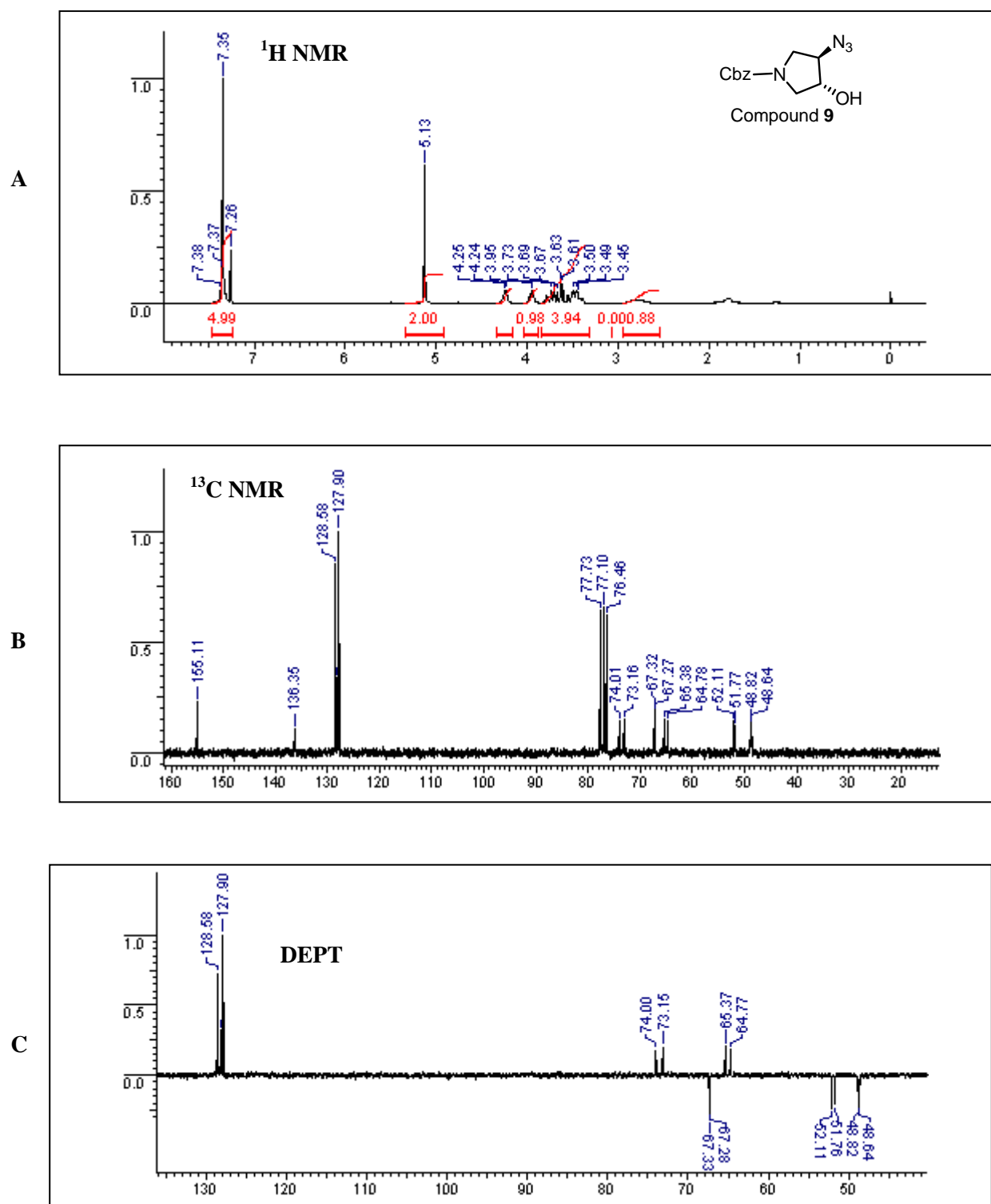


**Figure 5:** (A)  $^1\text{H}$  NMR of compound 5 (B)  $^{13}\text{C}$  NMR of compound 5 (C) DEPT of compound 5

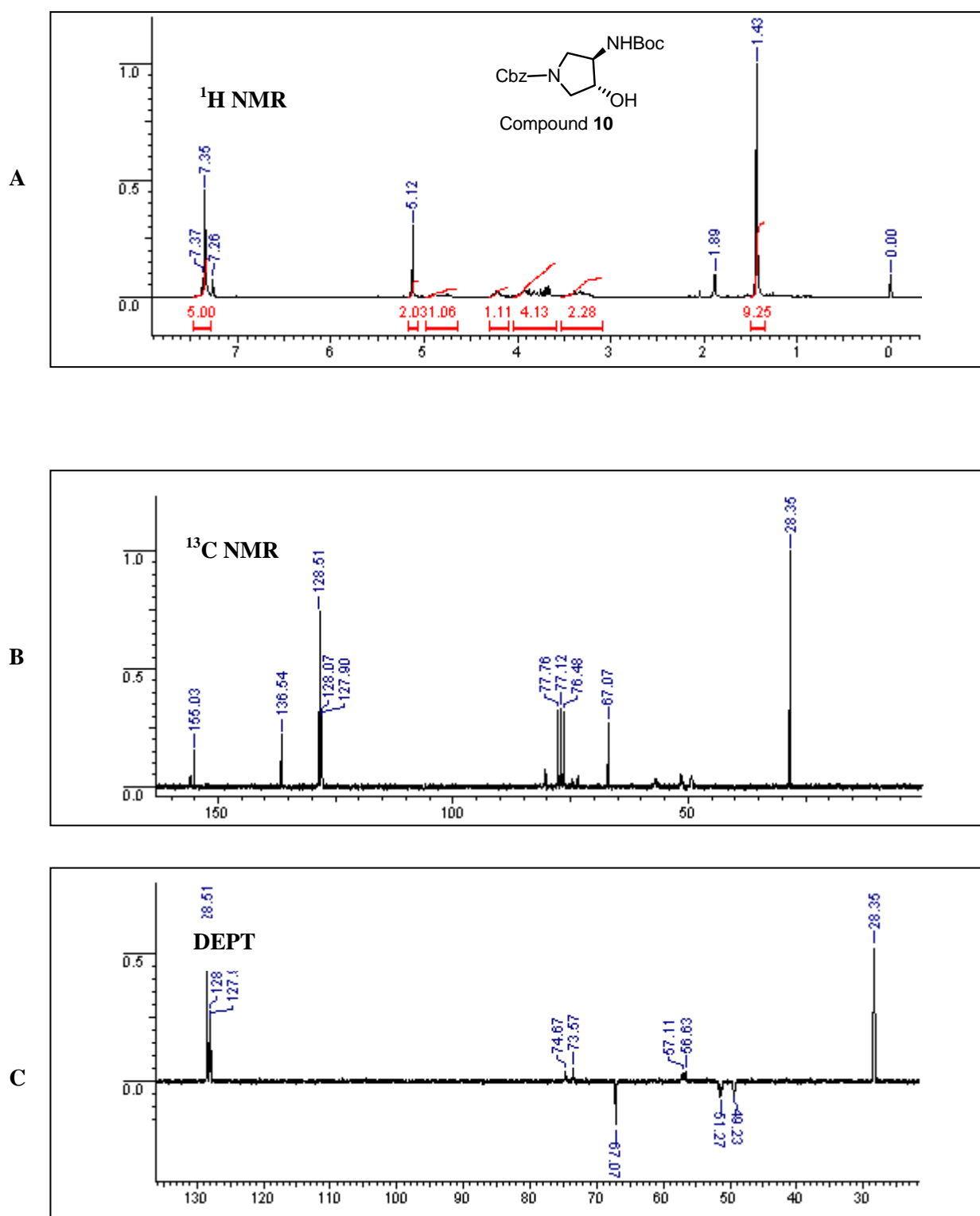


**Figure 6:** (A) <sup>1</sup>H NMR of compound 6 (B) <sup>13</sup>C NMR of compound 6 (C) DEPT of compound 6





**Figure 8:** (A) <sup>1</sup>H NMR of compound **9** (B) <sup>13</sup>C NMR of compound **9** (C) DEPT of compound **9**



**Figure 9:** (A) <sup>1</sup>H NMR of compound **10** (B) <sup>13</sup>C NMR of compound **10** (C) DEPT of compound **10**

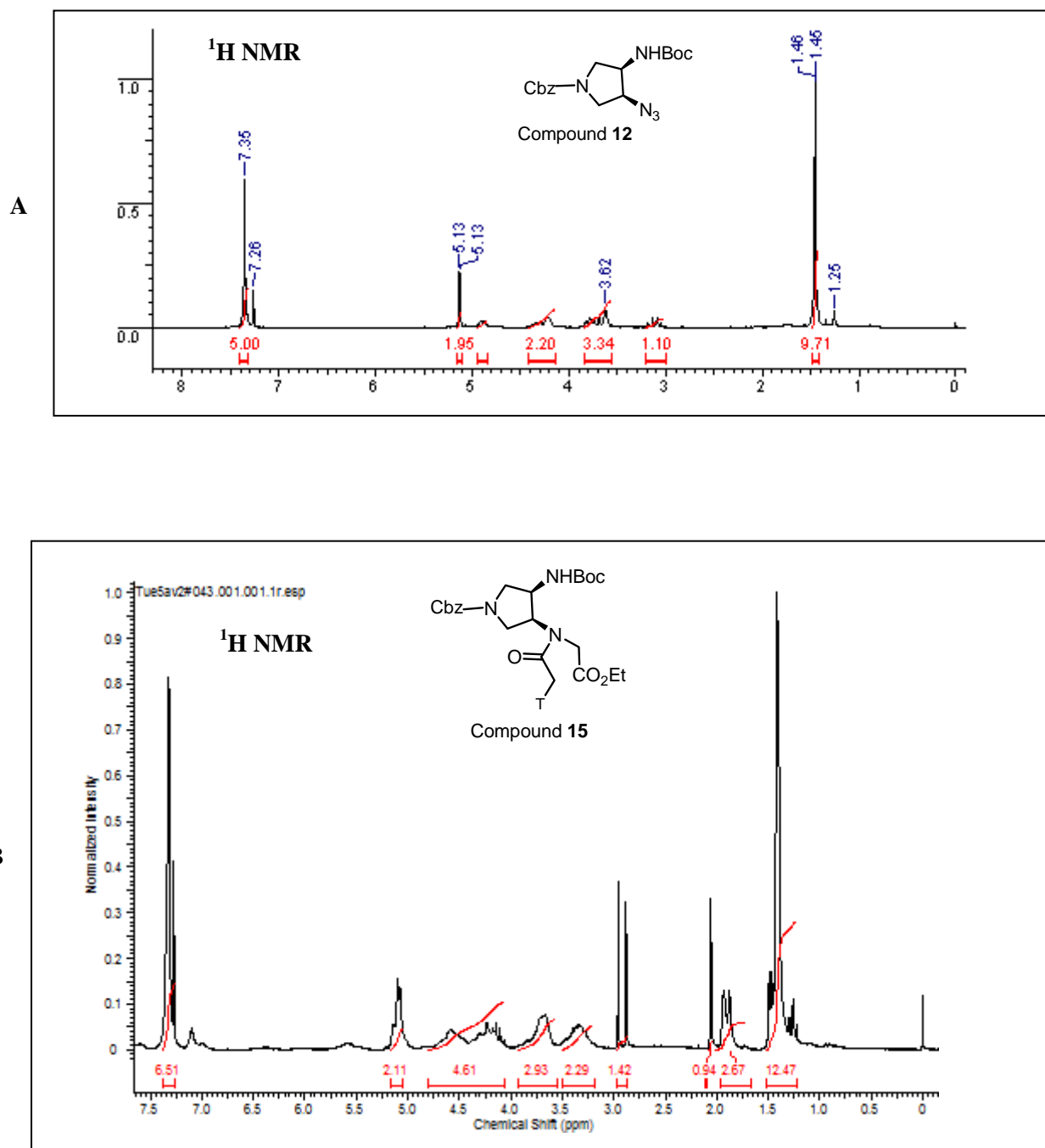
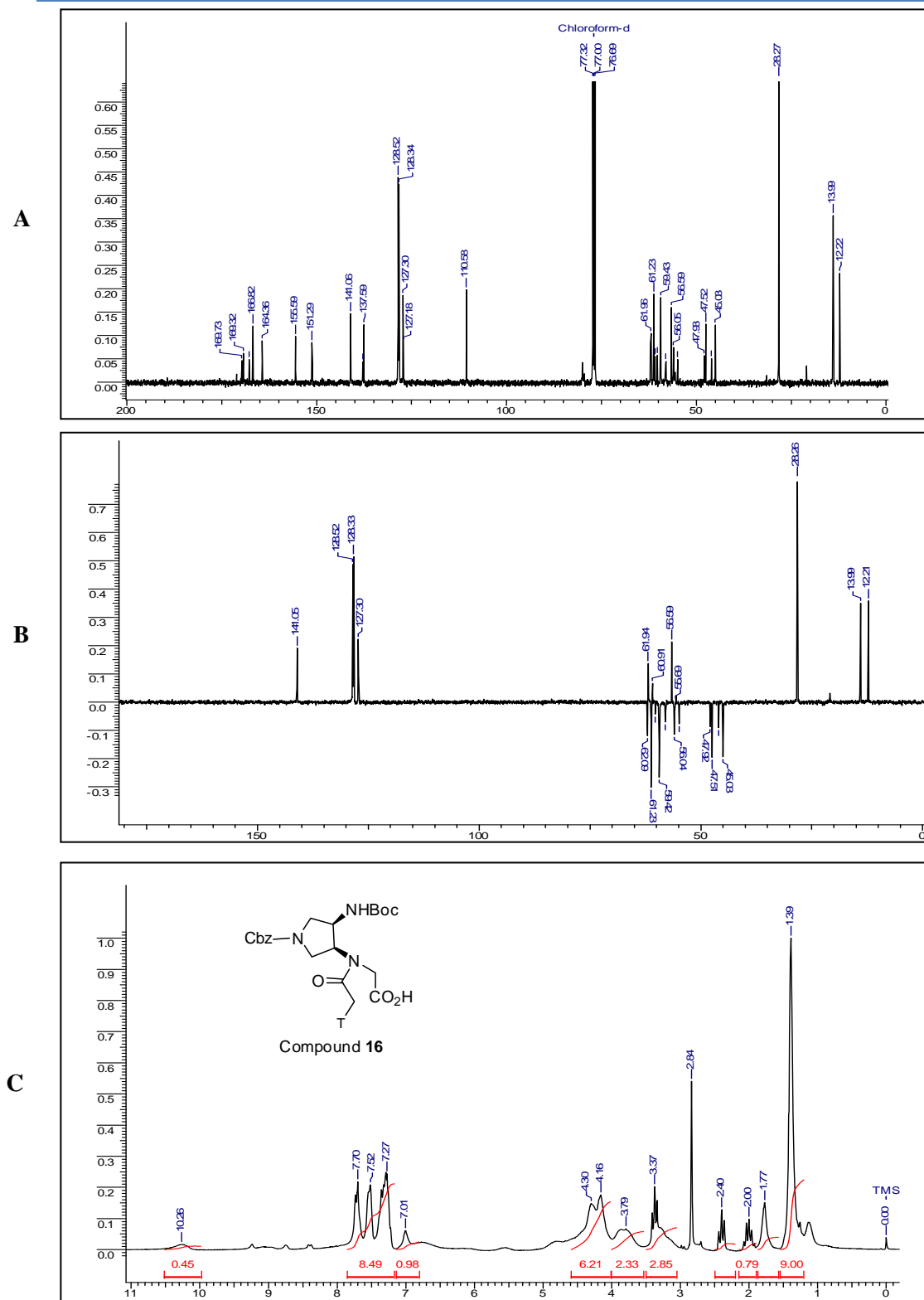
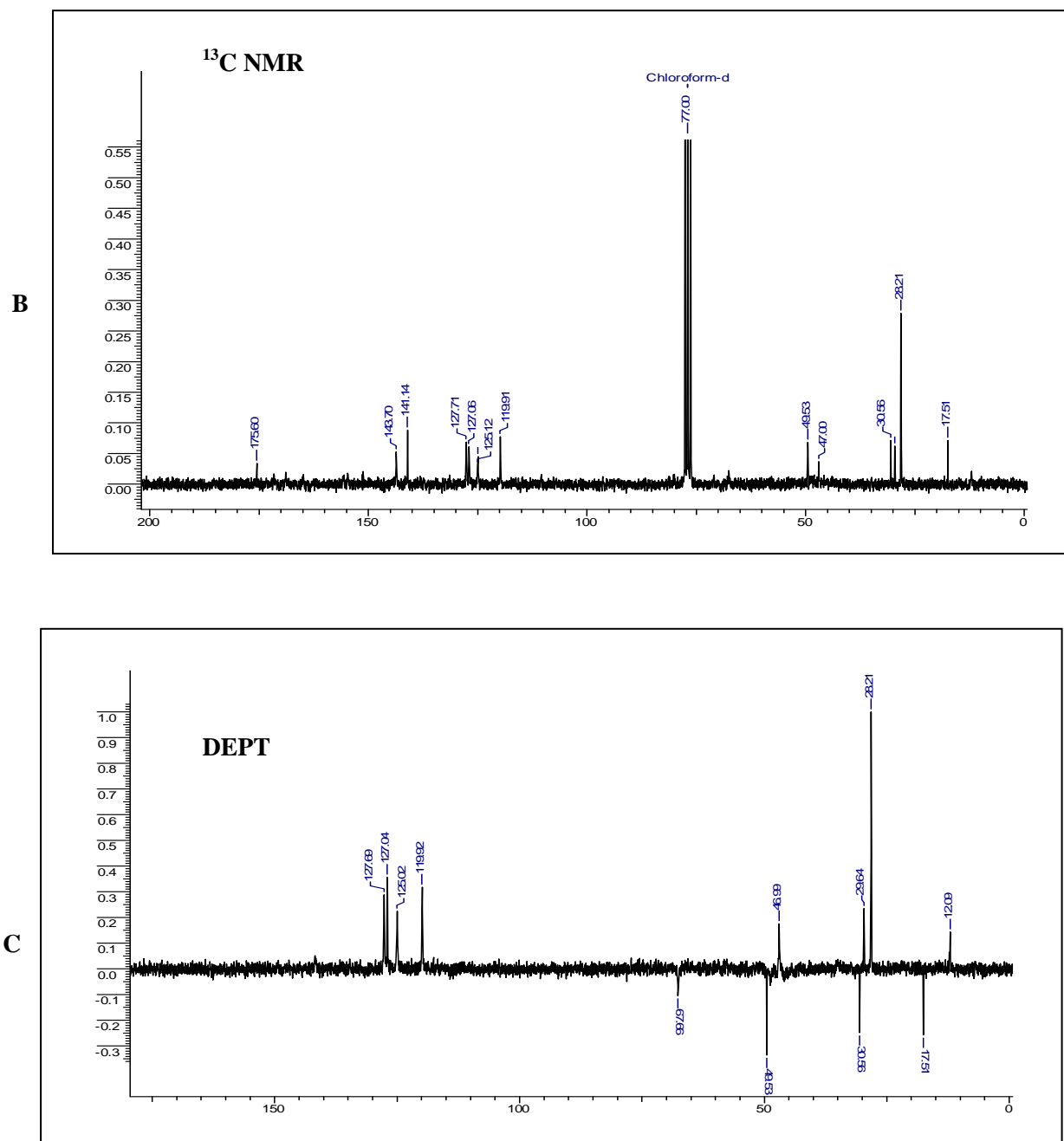


Figure 10: (A) <sup>1</sup>H NMR of compound 12 (B) <sup>1</sup>H NMR of compound 15

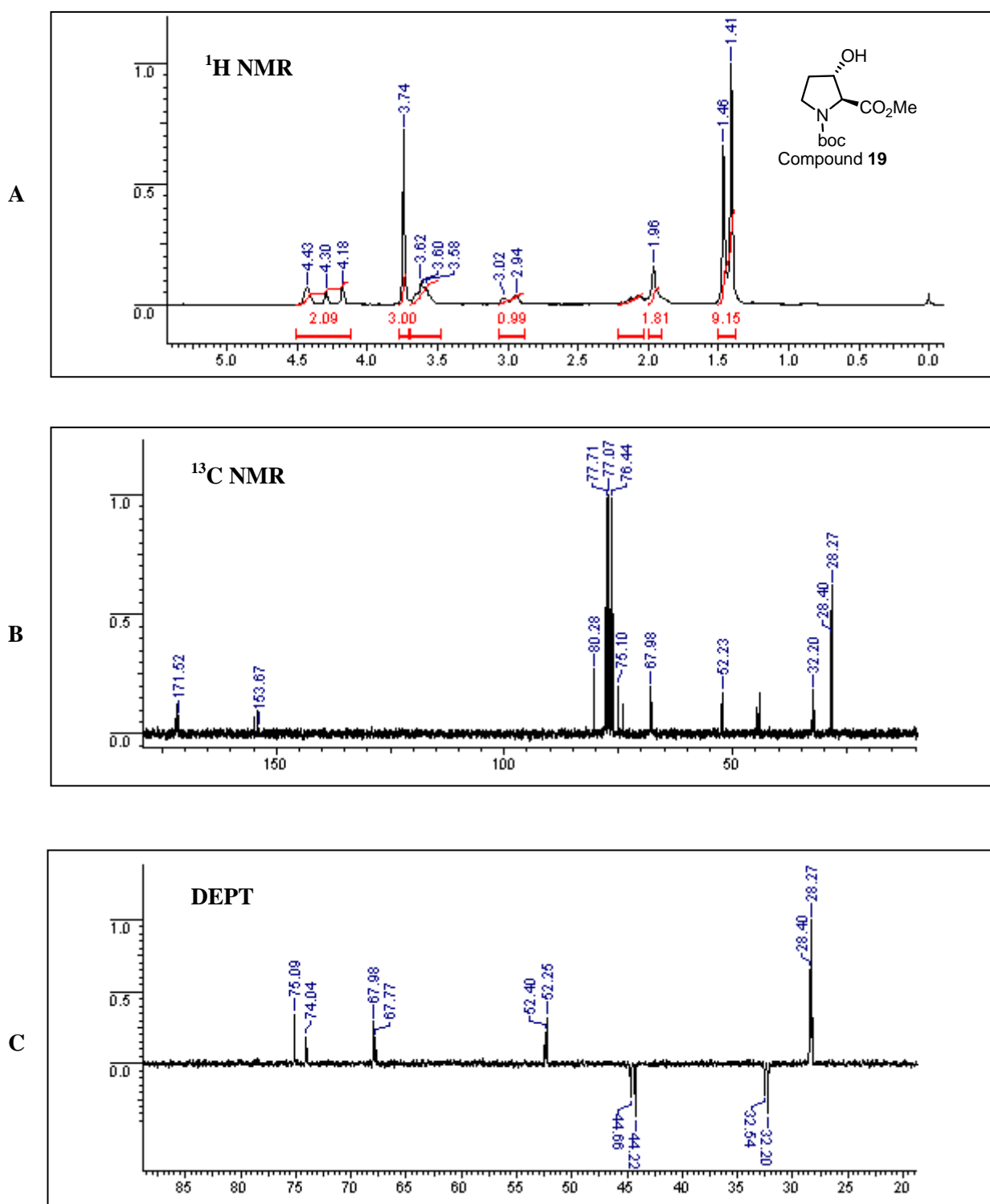


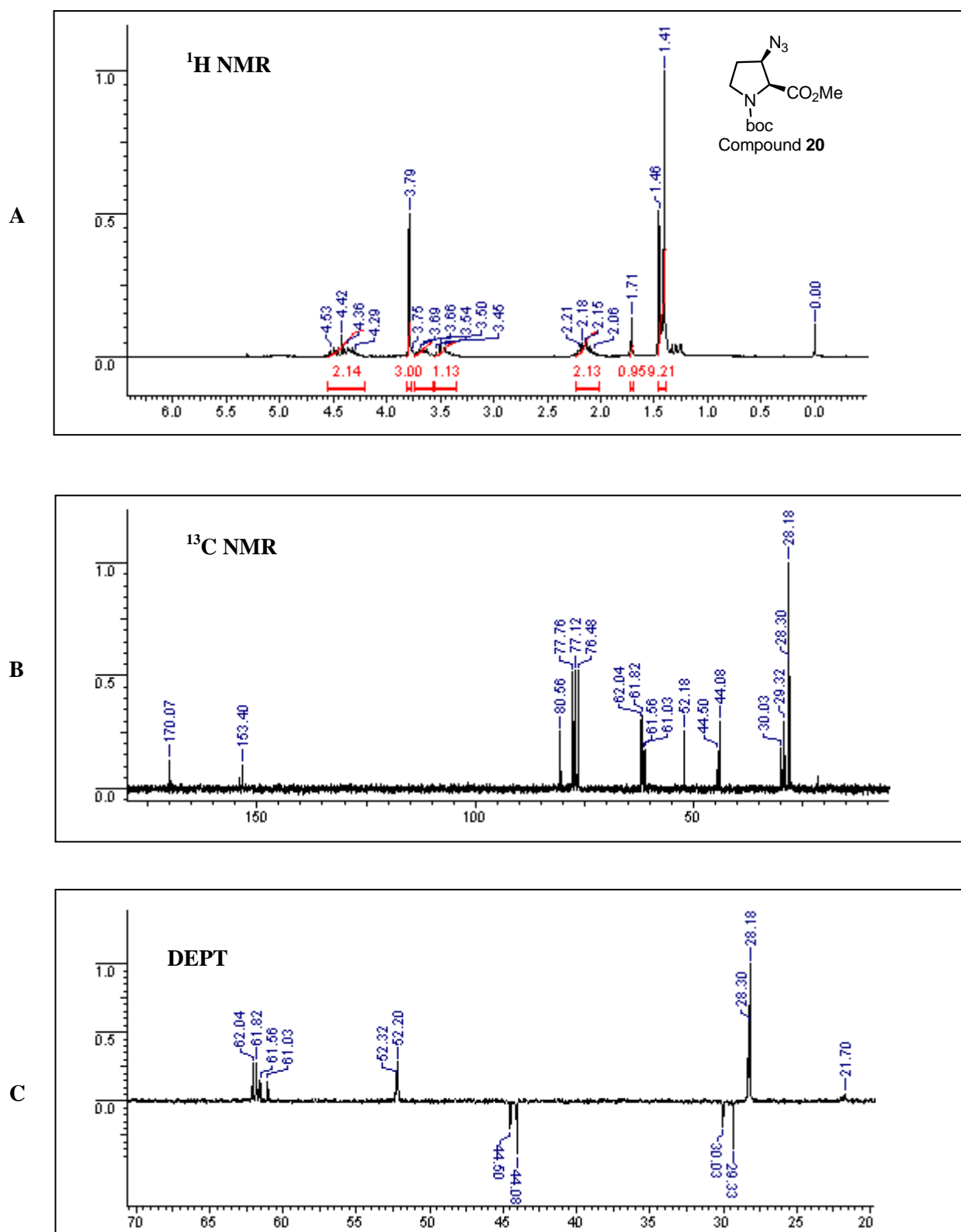


**Figure 11:** (A)  $^{13}\text{C}$ NMR of compound 15 (B) DEPT compound 15 (B)  $^1\text{H}$ NMR of compound 16

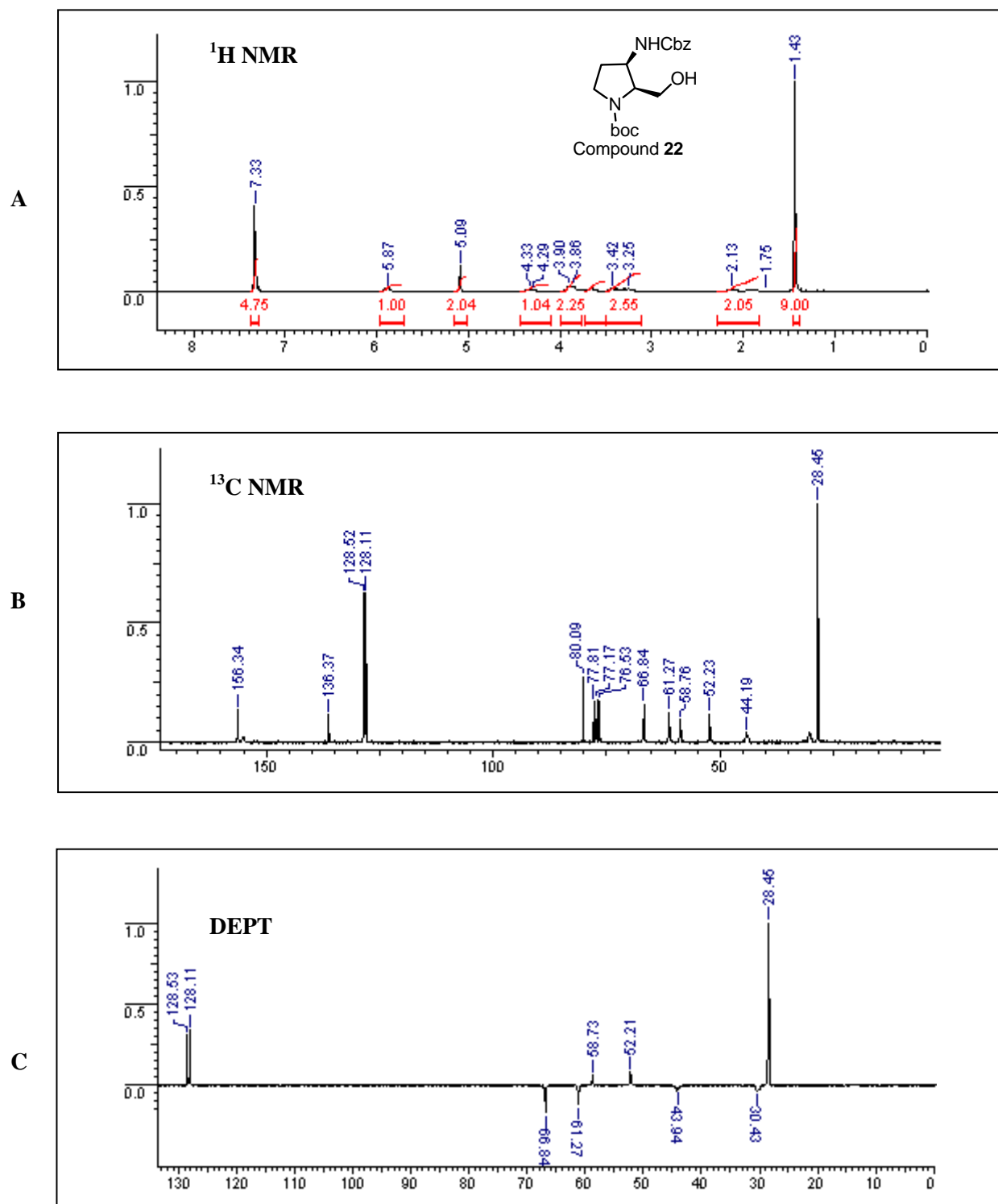


**Figure 12: (B)  $^{13}\text{C}$  NMR of compound 16(C) DEPT of compound 16**

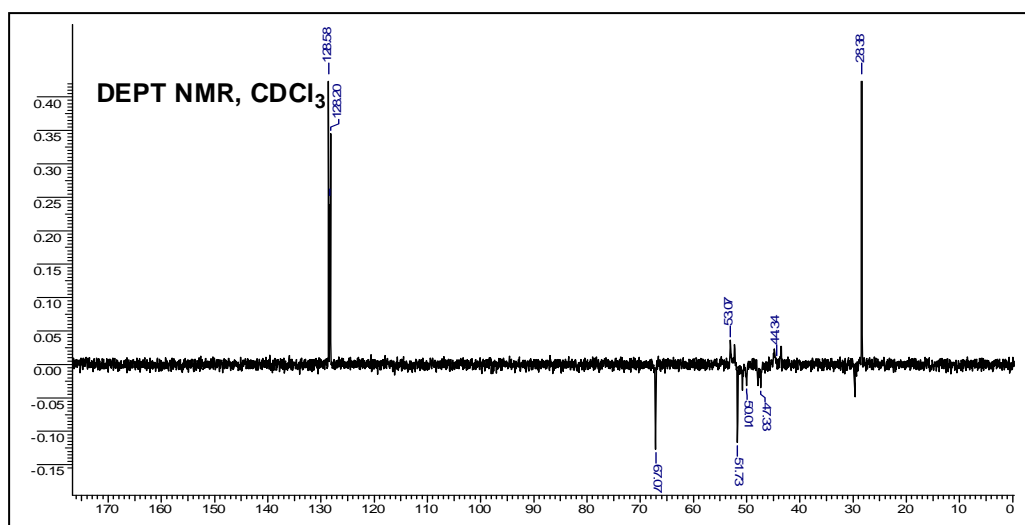
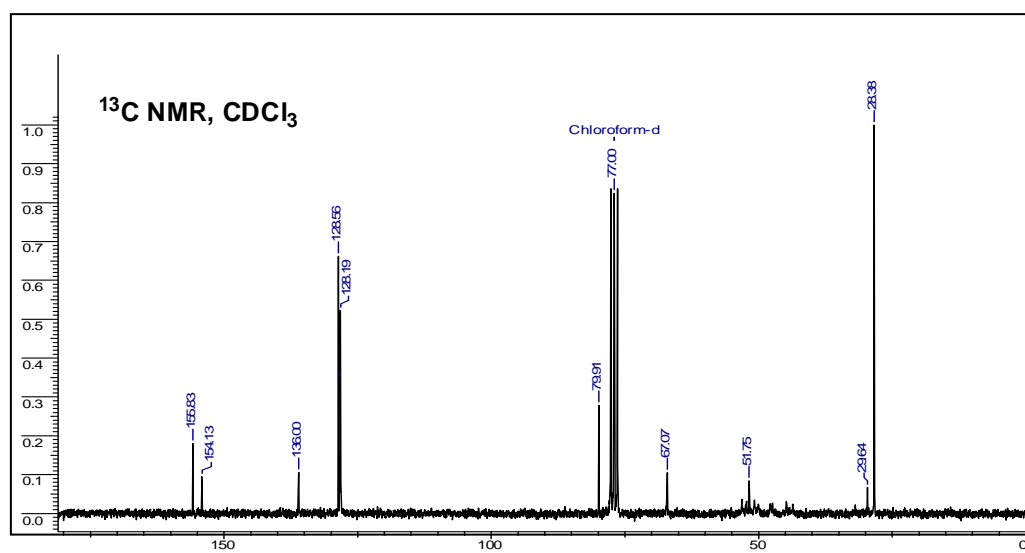
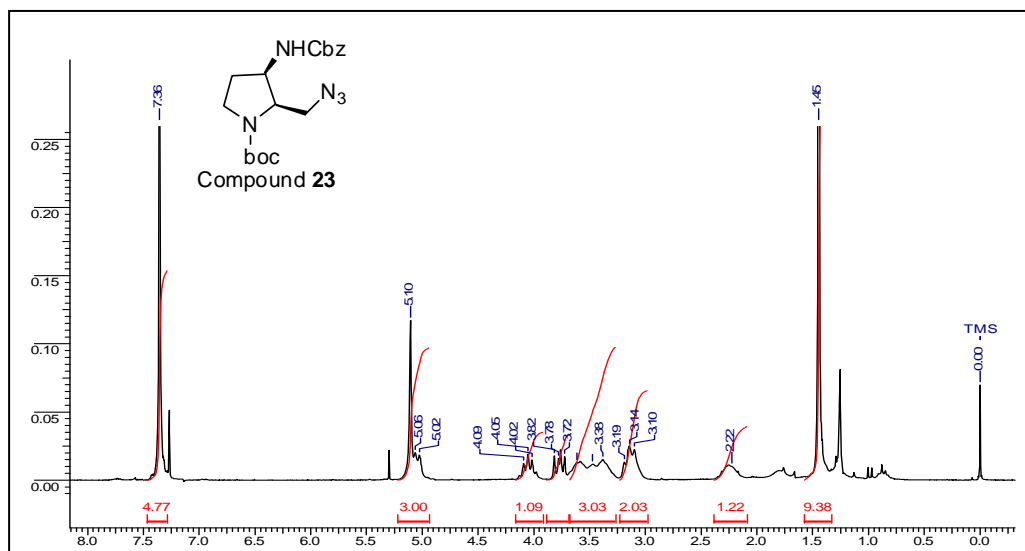


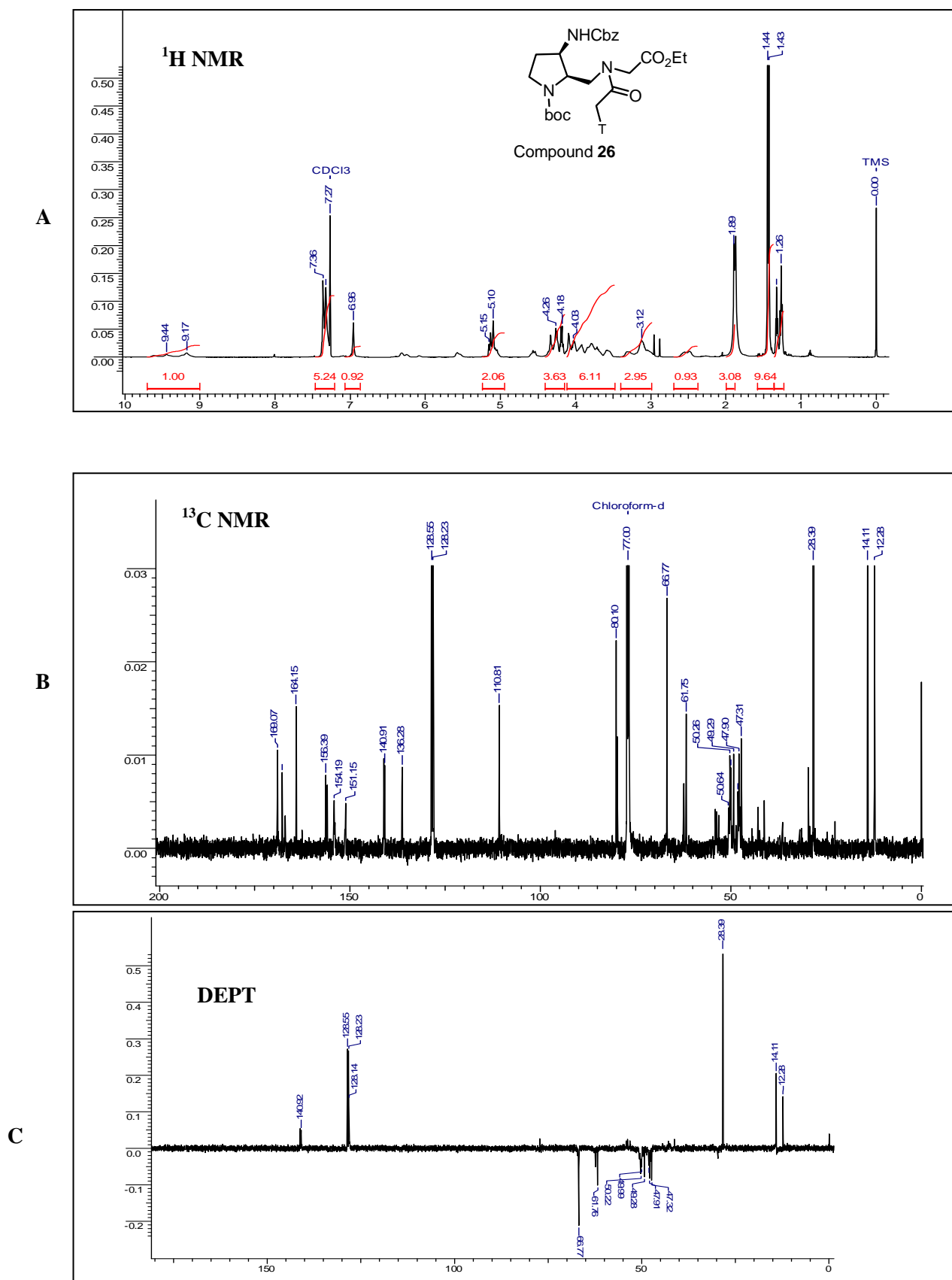


**Figure 14:** (A) <sup>1</sup>H NMR of compound **20** (B) <sup>13</sup>C NMR of compound **20** (C) DEPT of compound **20**

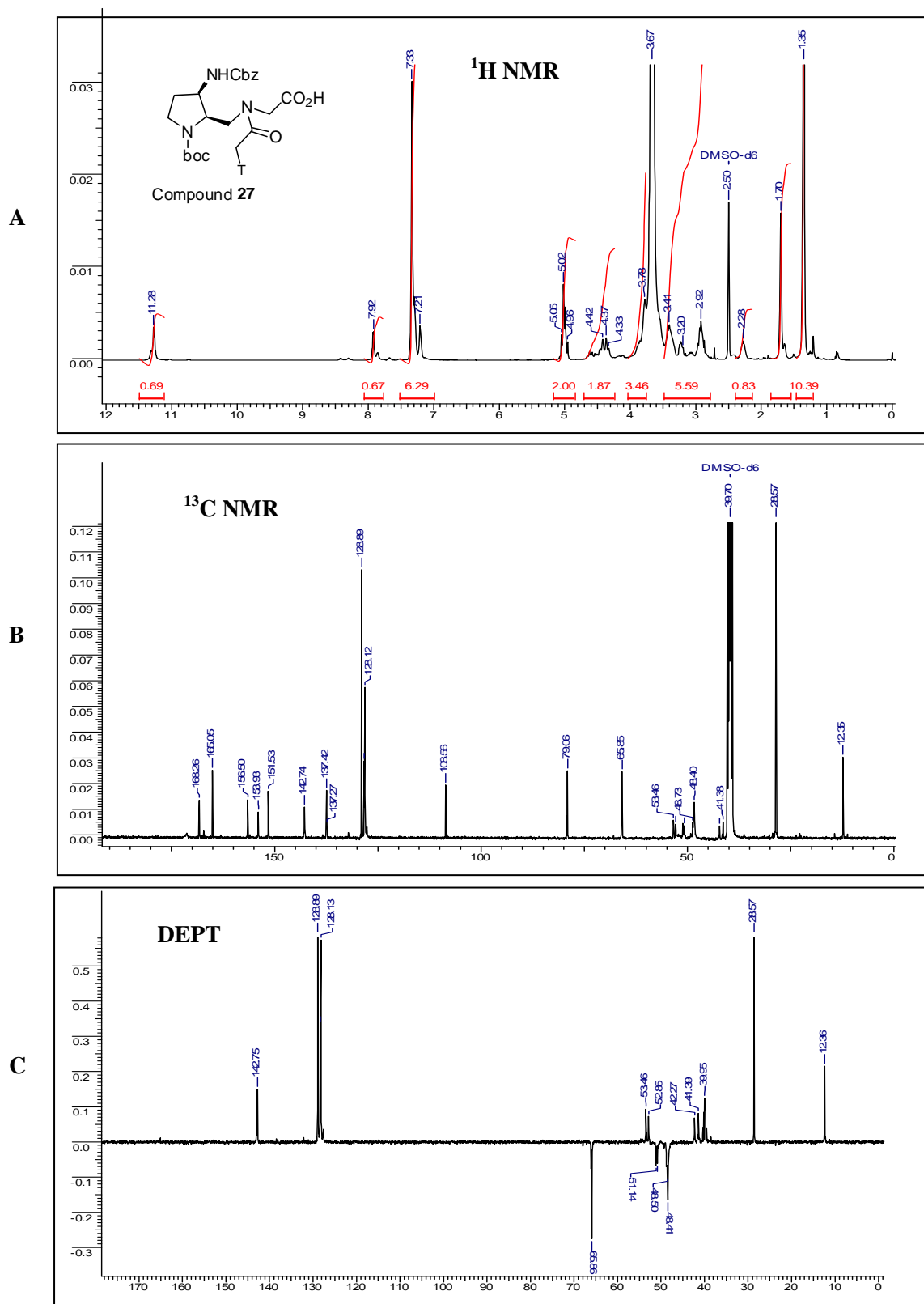


**Figure 15:** (A) <sup>1</sup>H NMR of compound **22** (B) <sup>13</sup>C NMR of compound **22** (C) DEPT of compound **22**



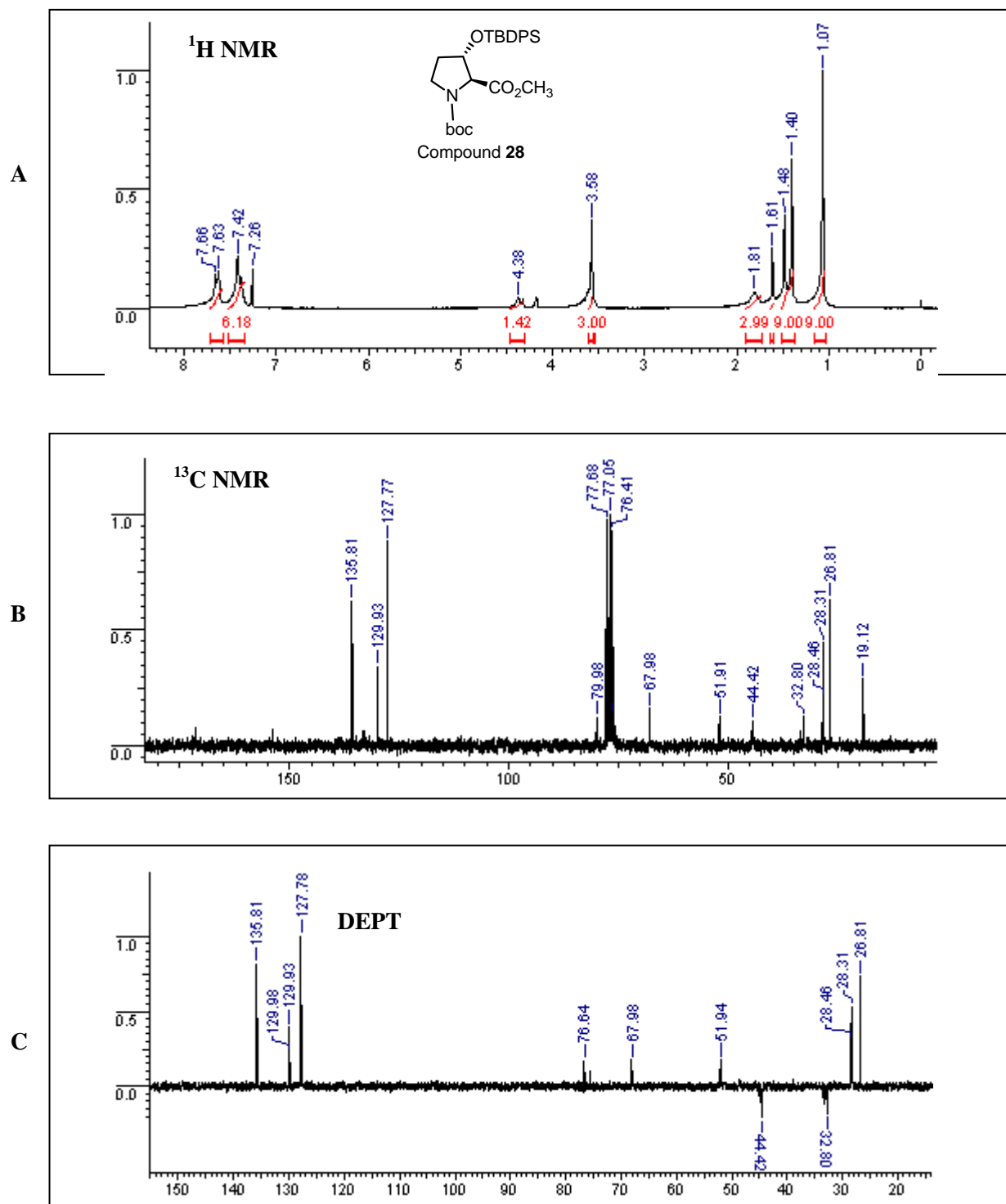


**Figure 16:** (A) <sup>1</sup>H NMR of compound 26 (B) <sup>13</sup>C NMR of compound 26 (C) DEPT of compound 26

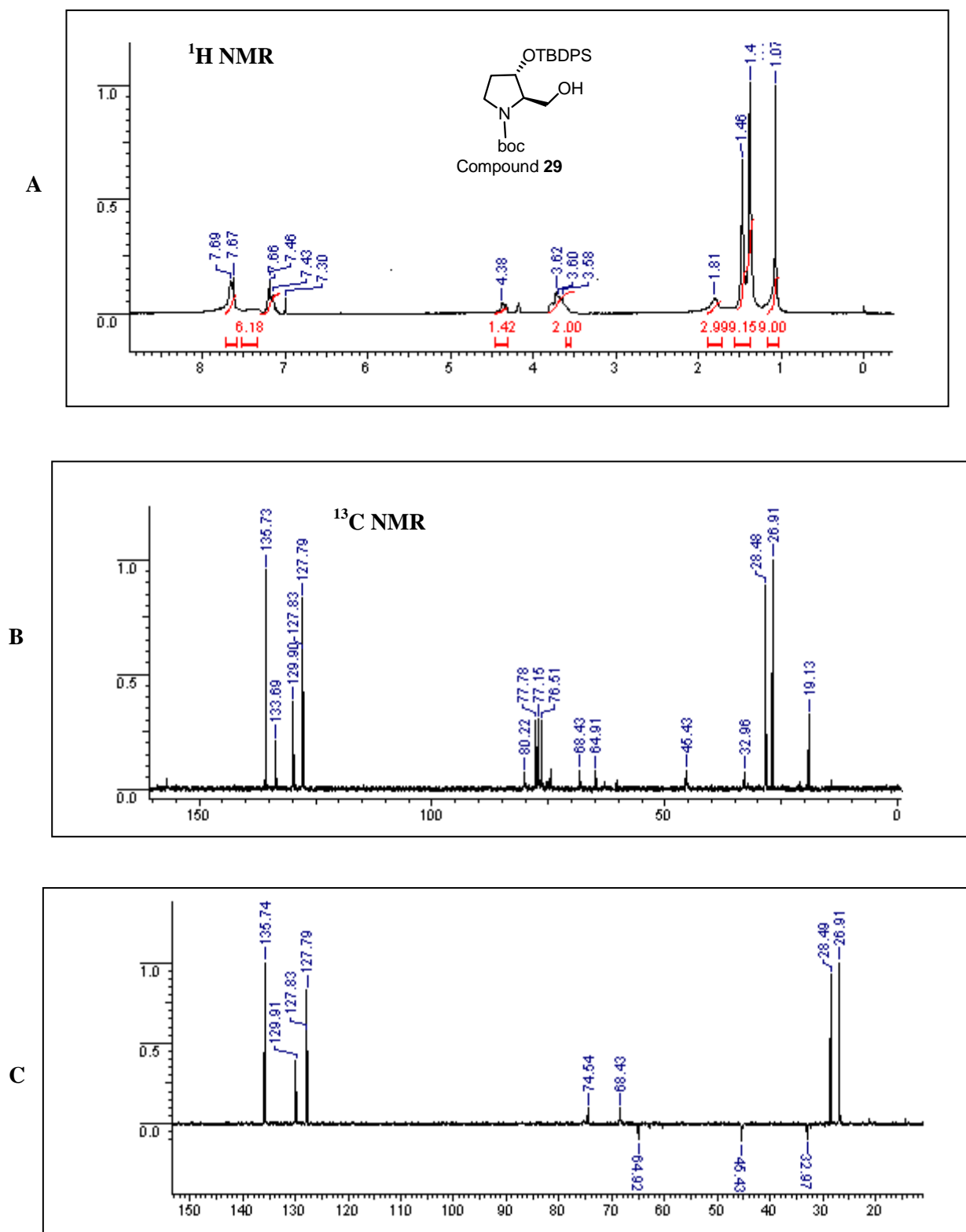


**Figure 17:** (A) <sup>1</sup>H NMR of compound 27 (B) <sup>13</sup>C NMR of compound 27 (C) DEPT of compound 27

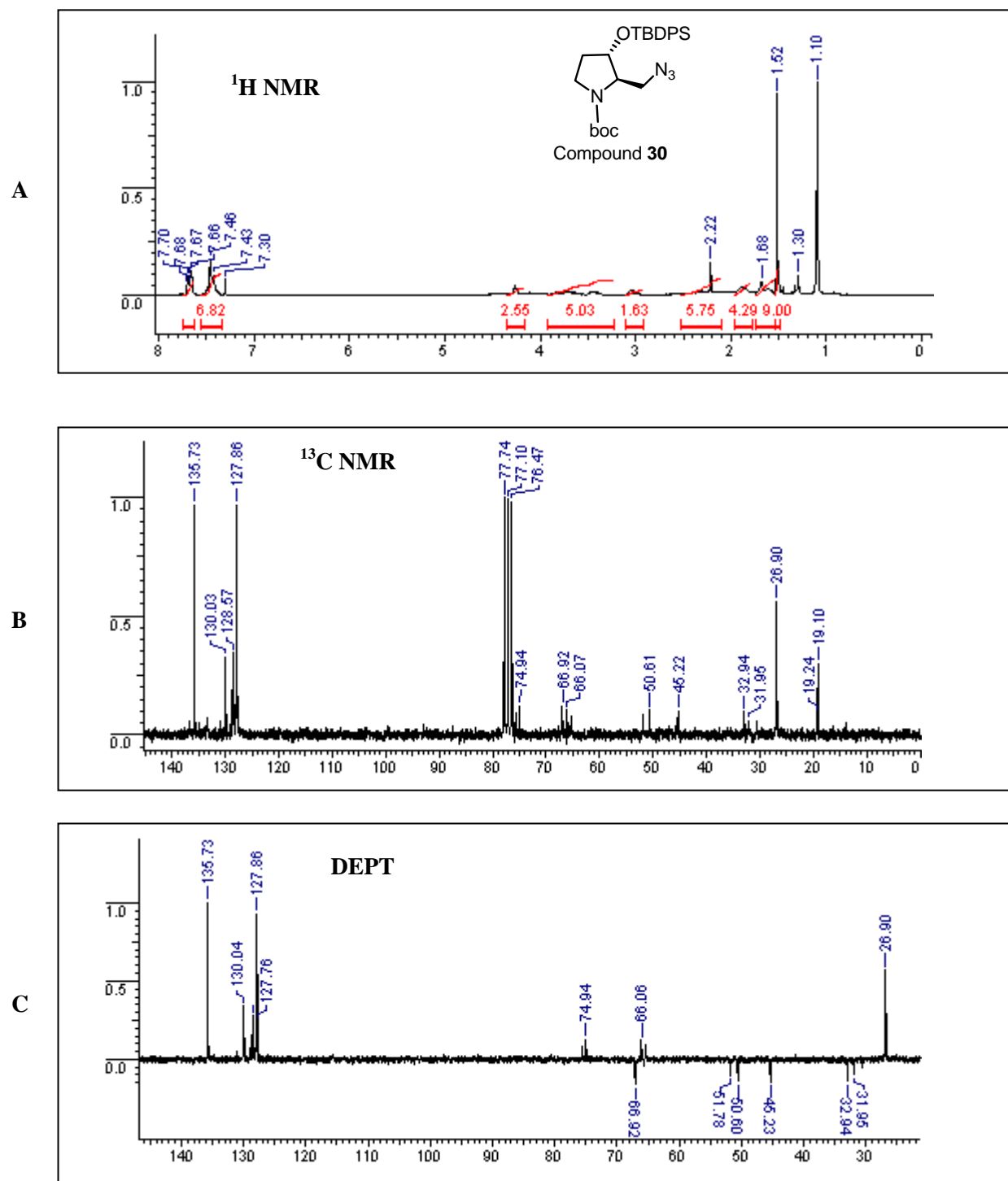




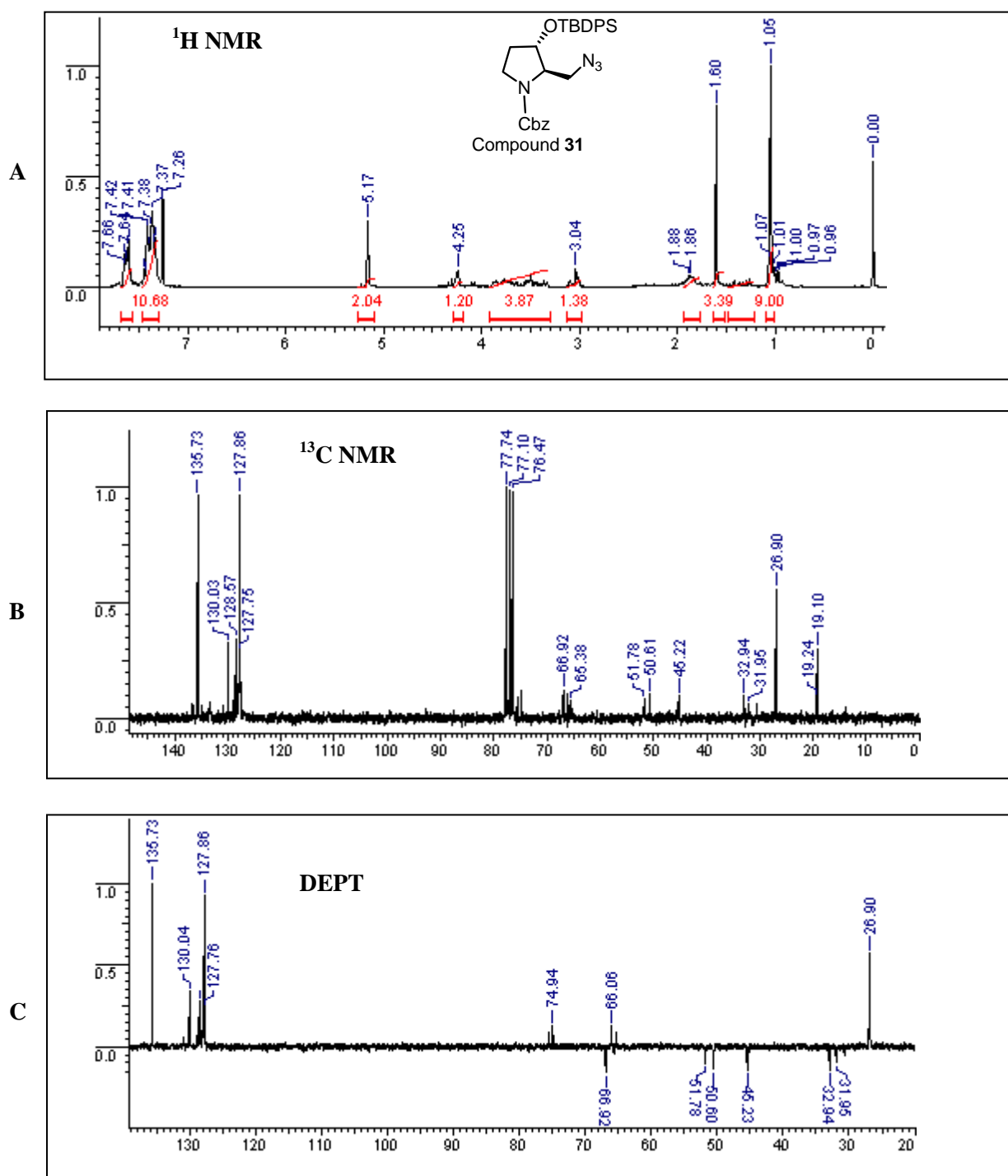
**Figure 18:** (A)  $^1\text{H}$  NMR of compound 28 (B)  $^{13}\text{C}$  NMR of compound 28 (C) DEPT of compound 28



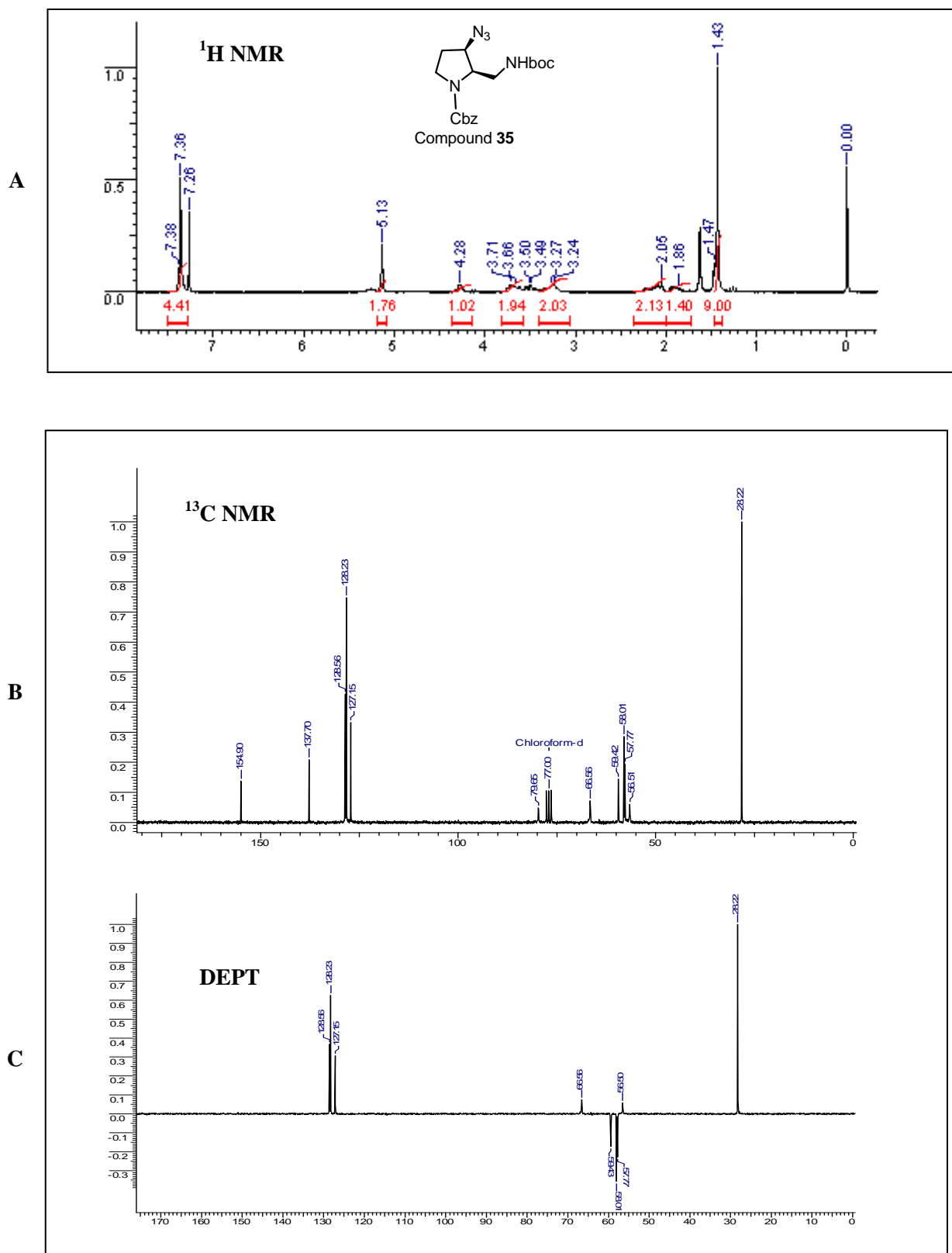
**Figure 19:** (A)  $^1\text{H}$  NMR of compound **29** (B)  $^{13}\text{C}$  NMR of compound **29** (C) DEPT of compound **29**



**Figure 20:** (A)  $^1\text{H}$  NMR of compound 30 (B)  $^{13}\text{C}$  NMR of compound 30 (C) DEPT of compound 30



**Figure 21:** (A) <sup>1</sup>H NMR of compound 31 (B) <sup>13</sup>C NMR of compound 31 (C) DEPT of compound 31



**Figure 22:** (A) <sup>1</sup>H NMR of compound 35 (B) <sup>13</sup>C NMR of compound 35 (C) DEPT of compound 35

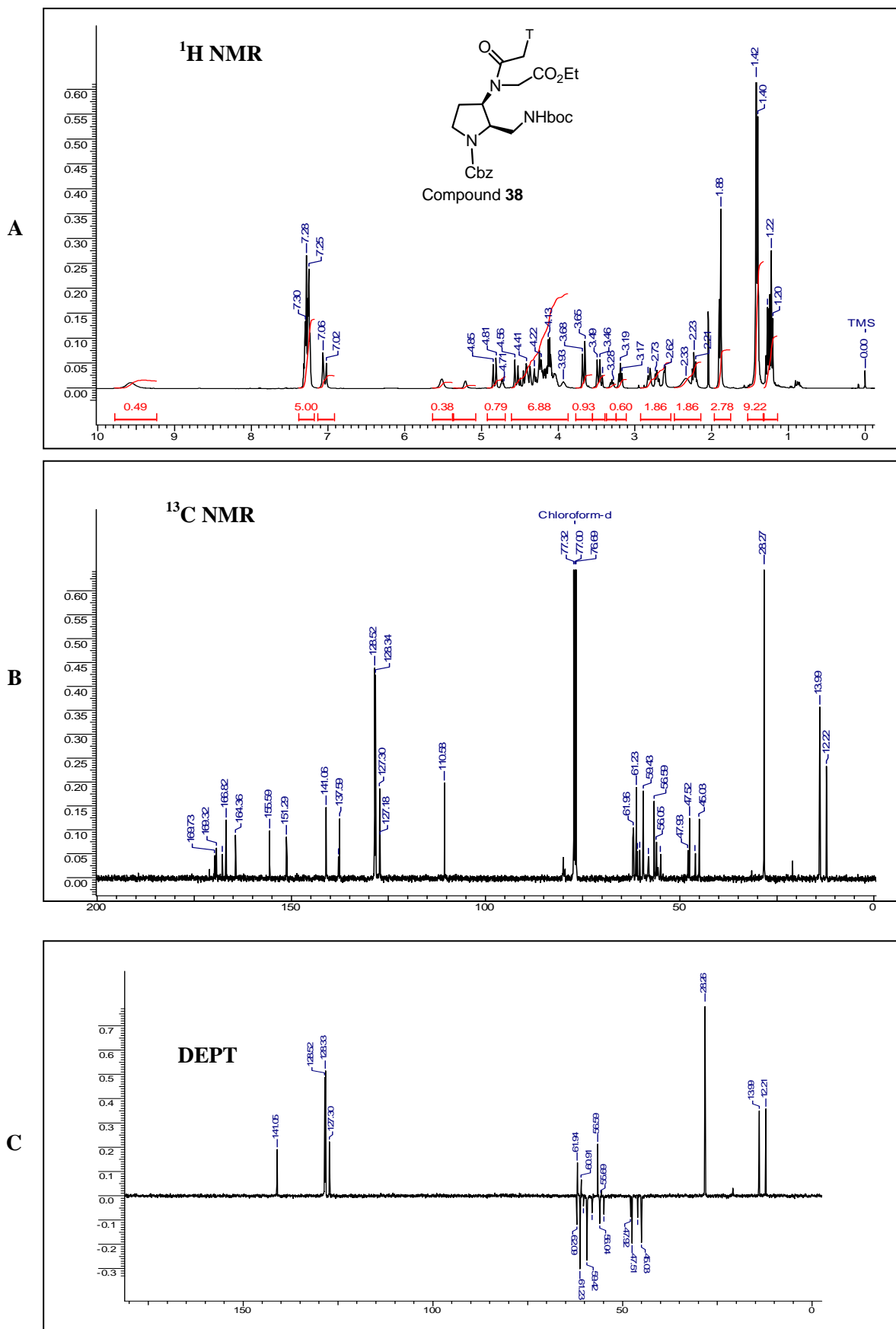


Figure 23: (A) <sup>1</sup>H NMR of compound 38 (B) <sup>13</sup>C NMR of compound 38 (C) DEPT of compound 38

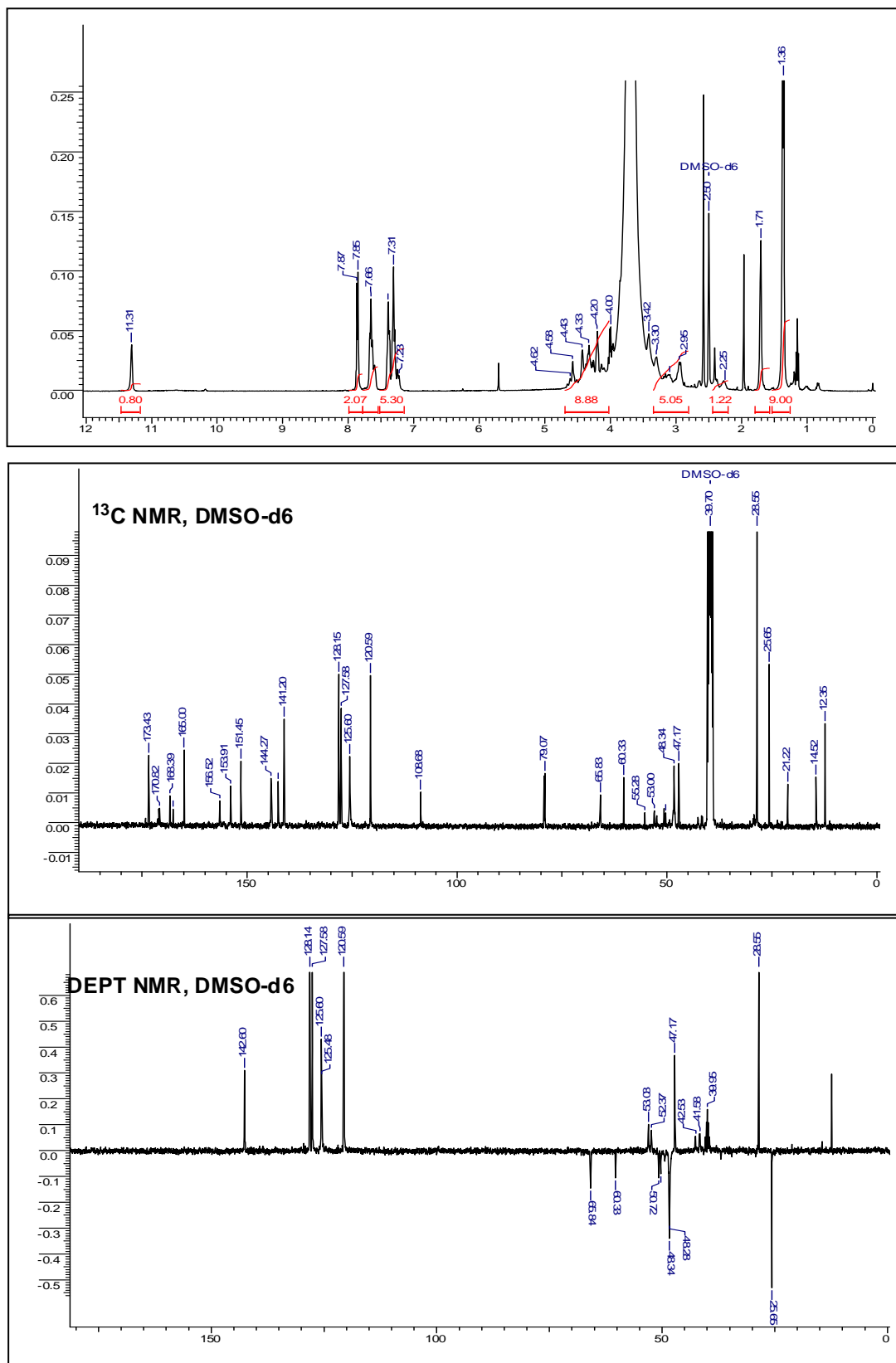
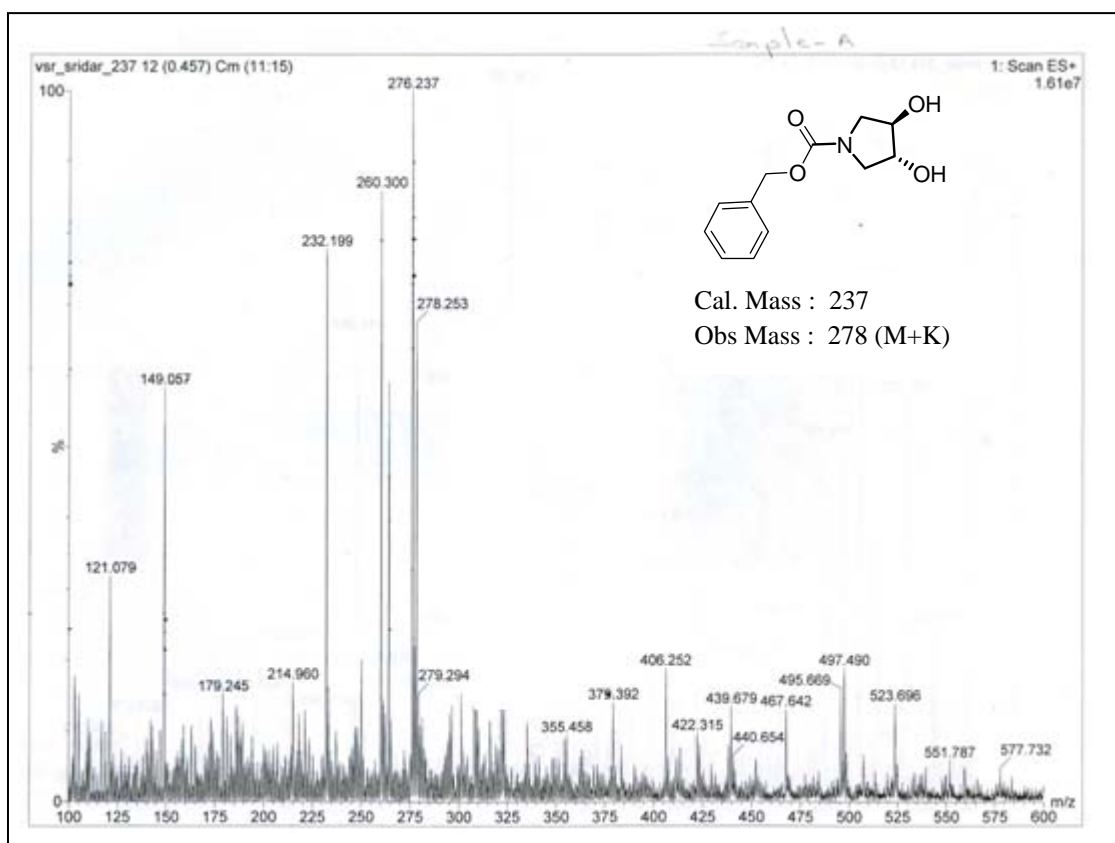
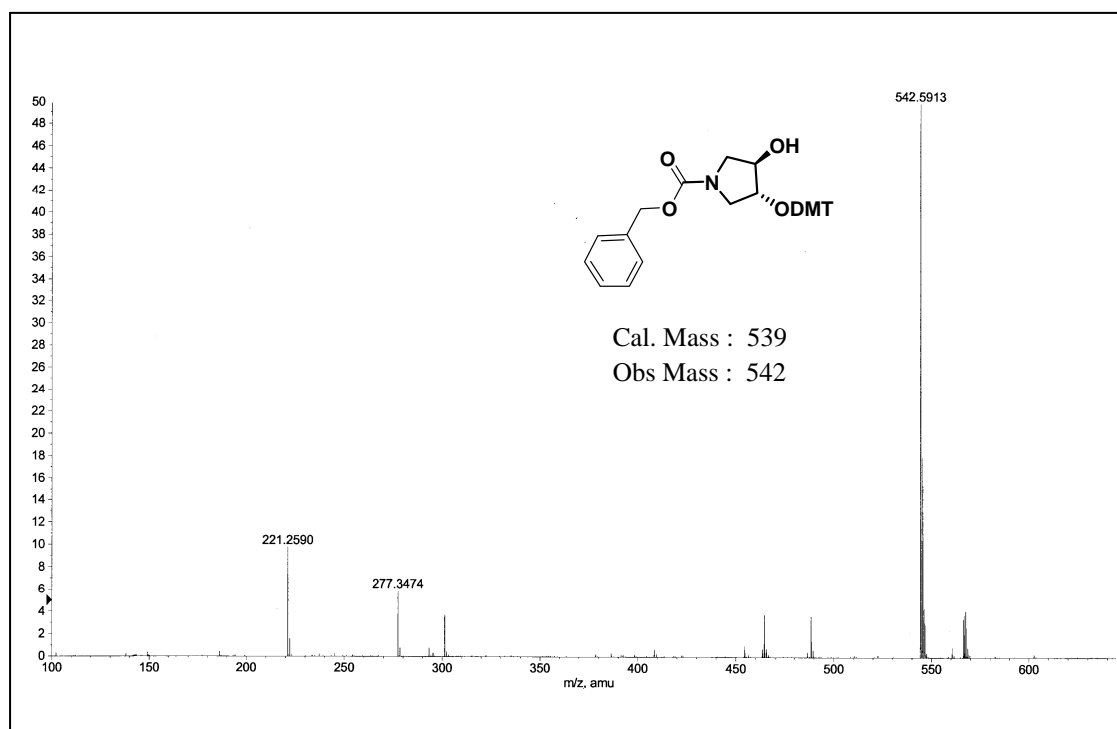


Figure 24: (A) <sup>1</sup>H NMR of compound 39 (B) <sup>13</sup>C NMR of compound 39 (C) DEPT of compound 39

A

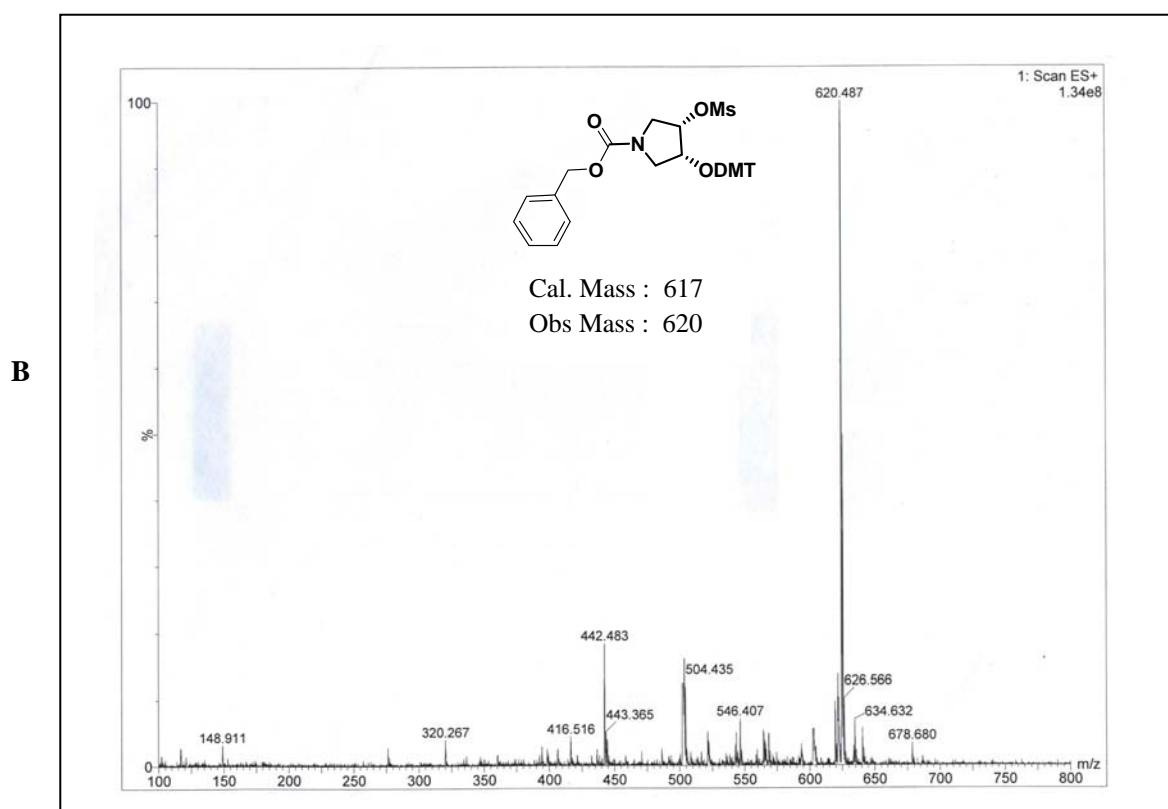
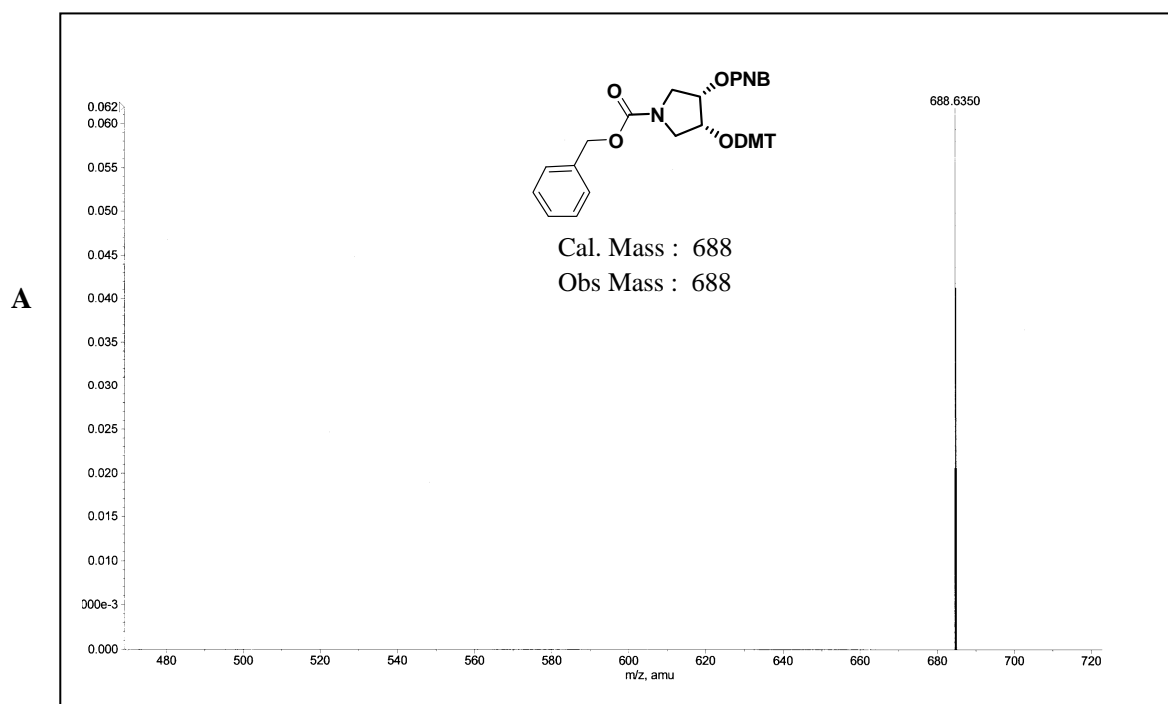


B

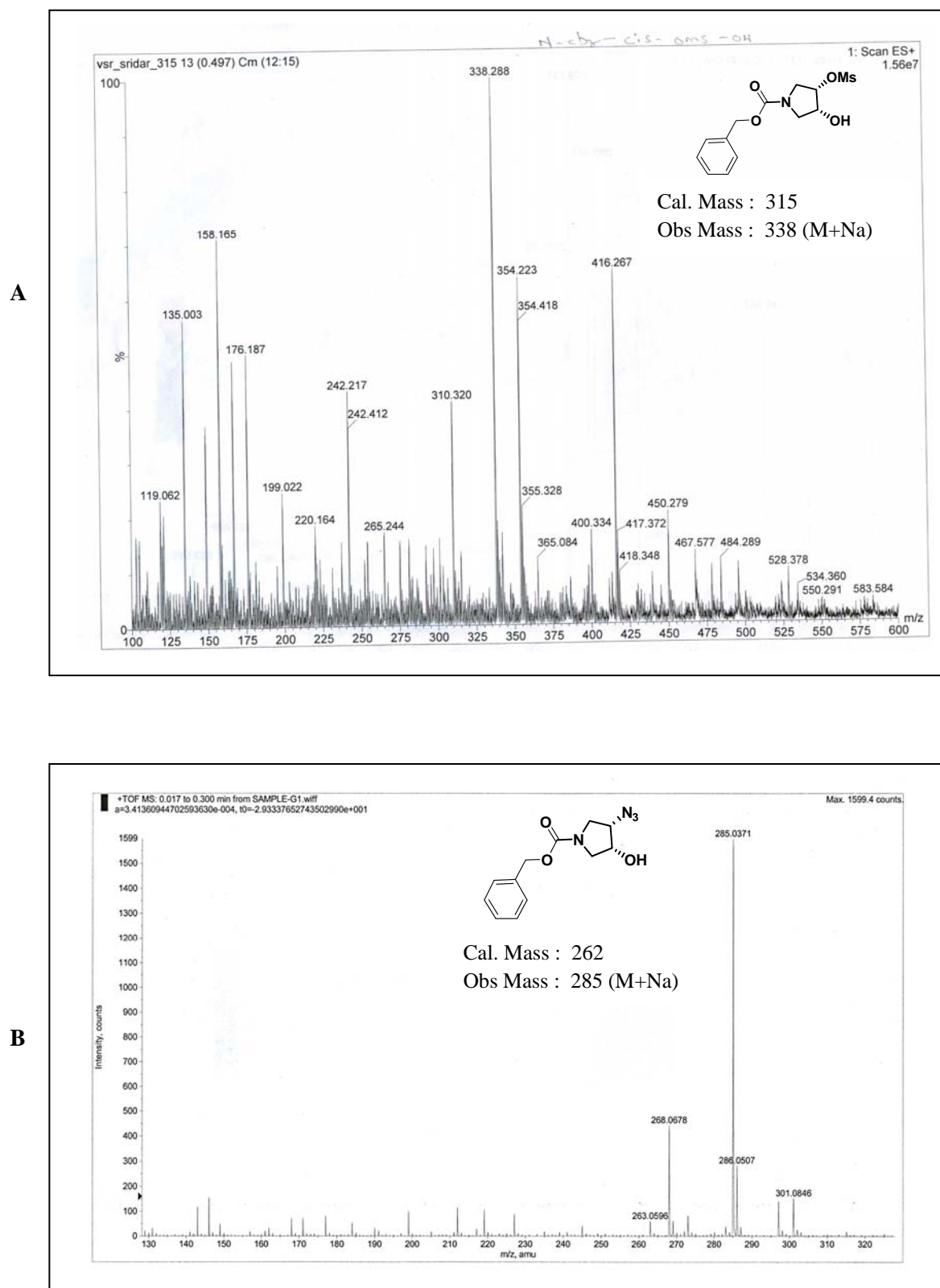


**Figure 25:** (A) Mass spectral analysis of compound 3 (B) Mass spectral analysis of compound 4

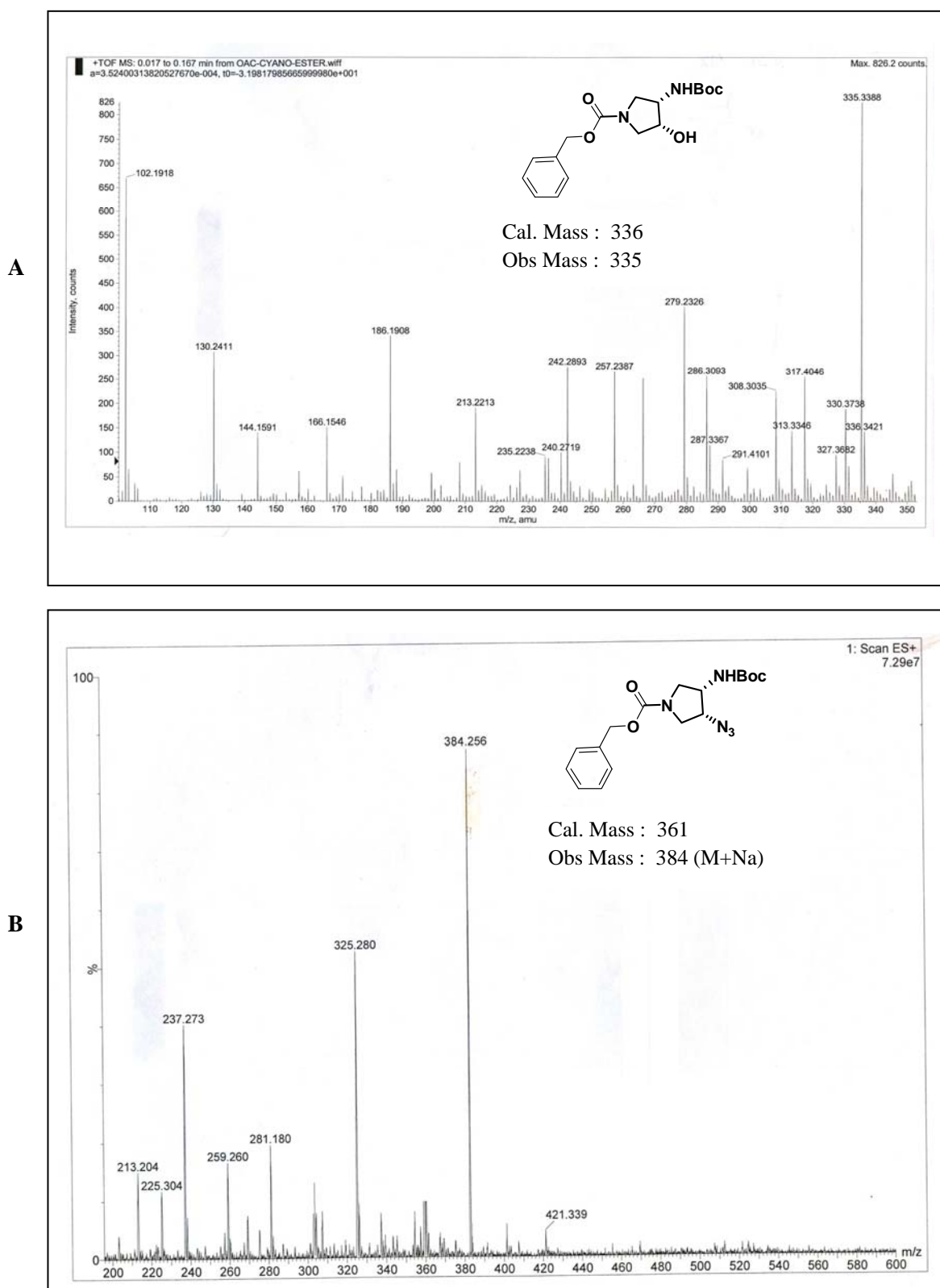


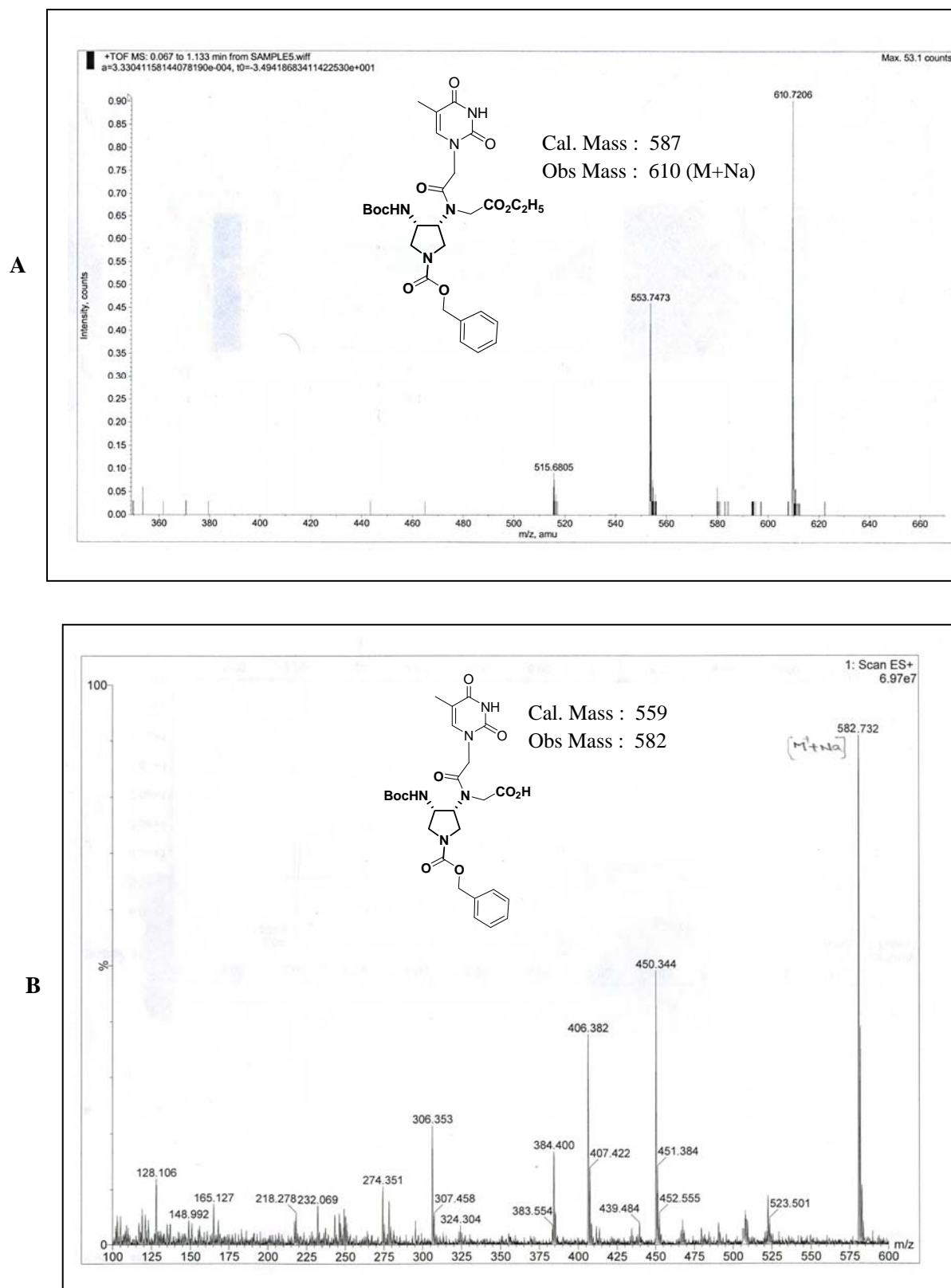


**Figure 26: (A) Mass spectral analysis of compound 5 (B) Mass spectral analysis of compound 7**

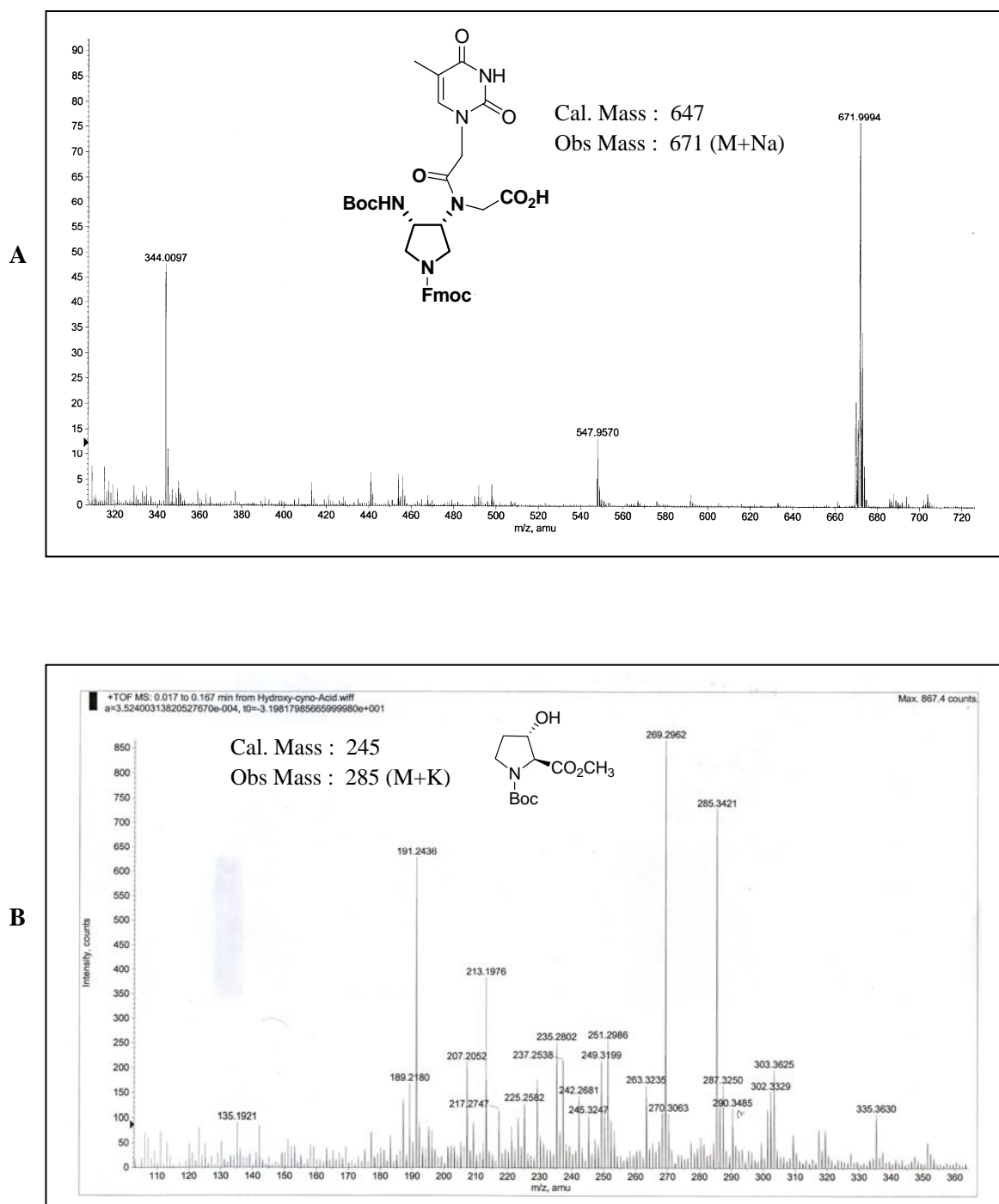


**Figure 27:** (A) Mass spectral analysis of compound **8** (B) Mass spectral analysis of compound **9**

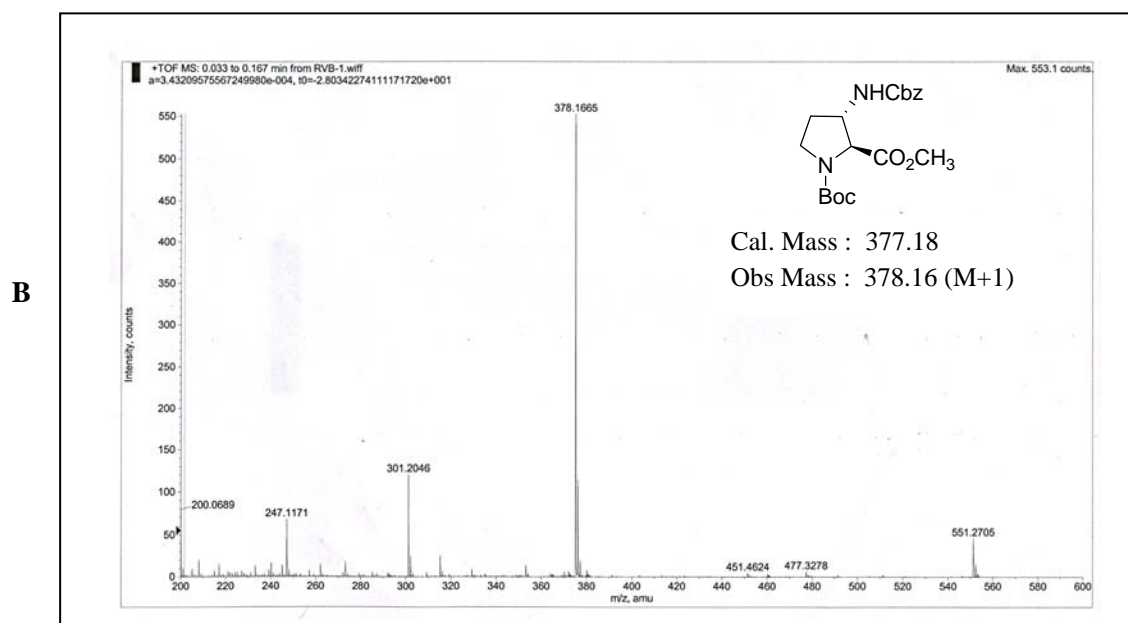
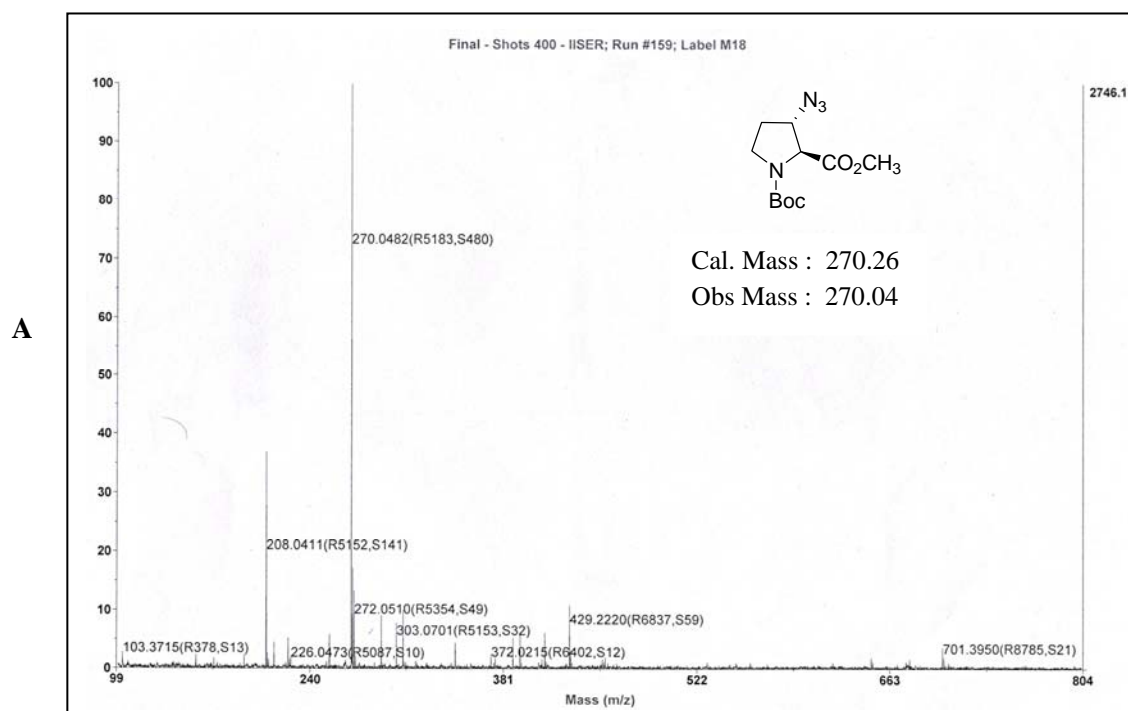




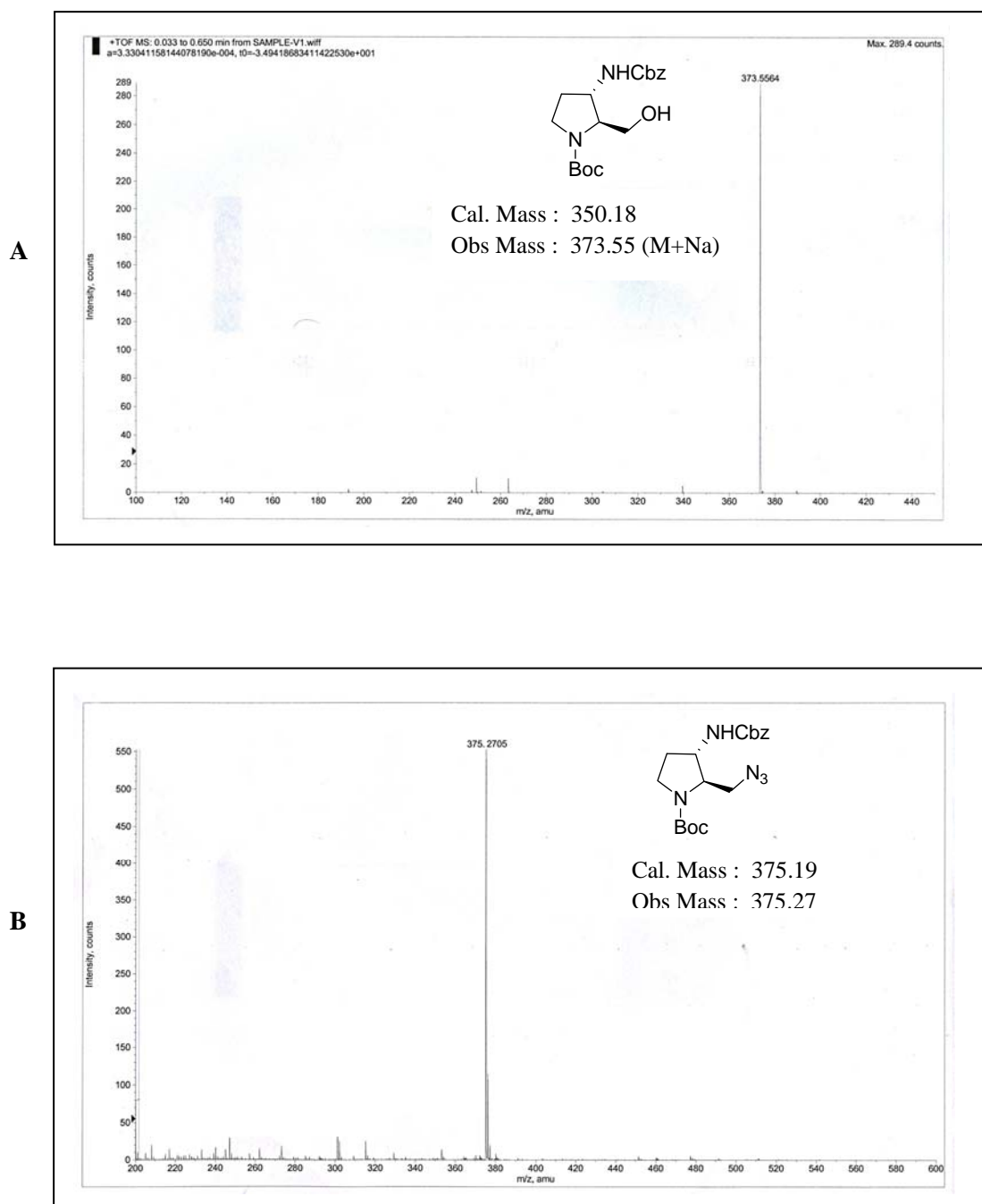
**Figure 29: (A) Mass spectral analysis of compound 15 (B) Mass spectral analysis of compound 16**



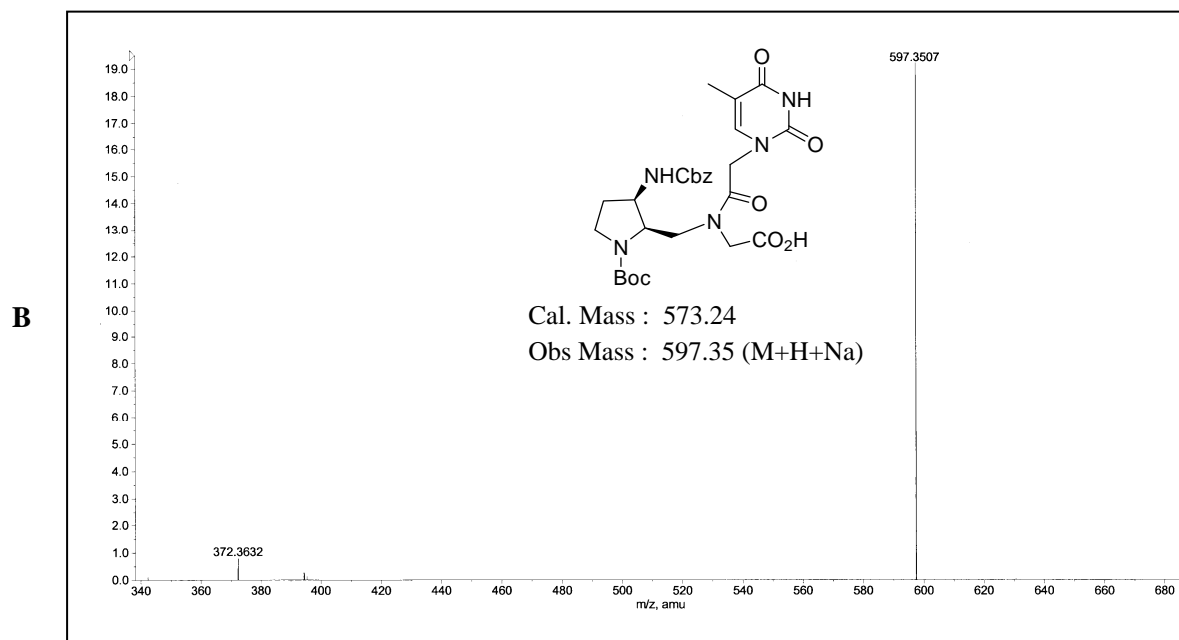
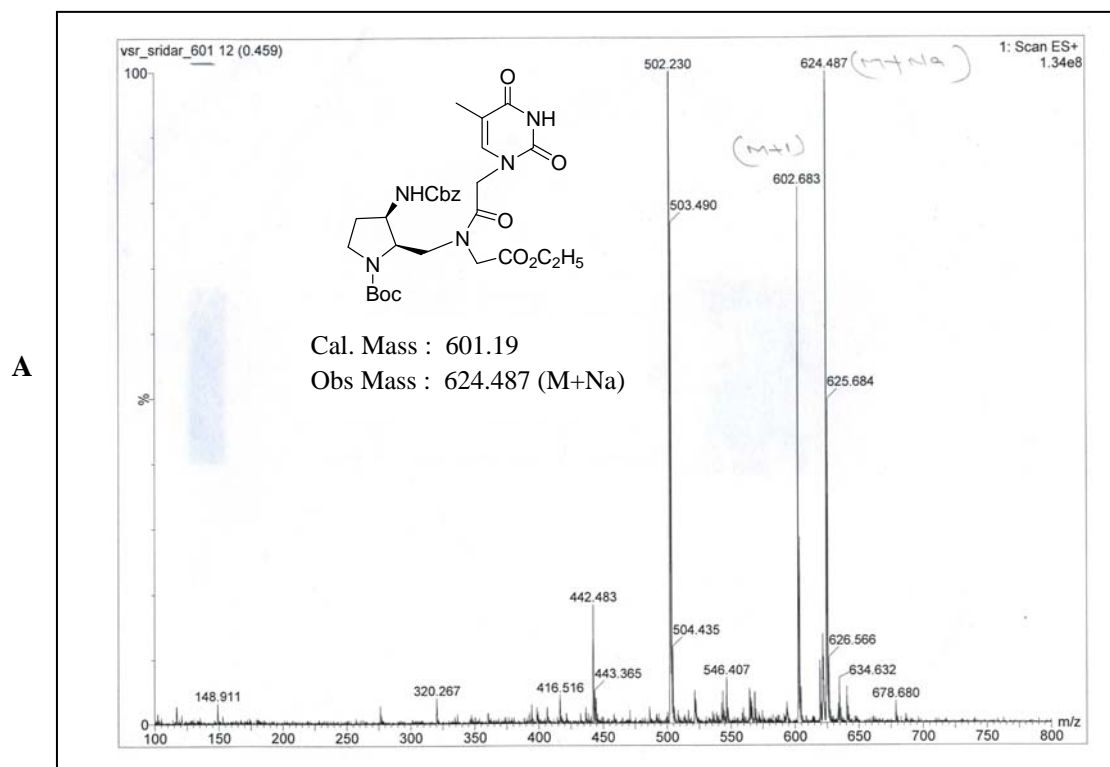
**Figure 30:** (A) Mass spectral analysis of compound **17** (B) Mass spectral analysis of compound **19**



**Figure 31:** (A) Mass spectral analysis of compound **20** (B) Mass spectral analysis of compound **21**



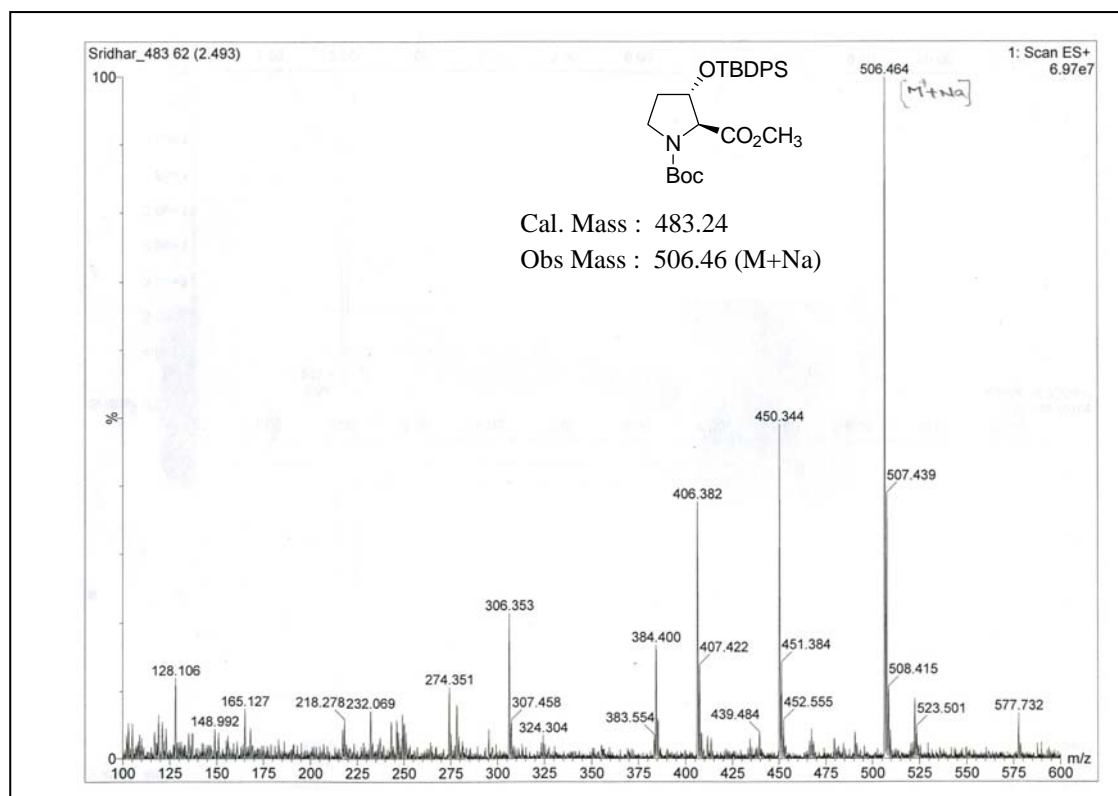
**Figure 32:** (A) Mass spectral analysis of compound **22** (B) Mass spectral analysis of compound **23**



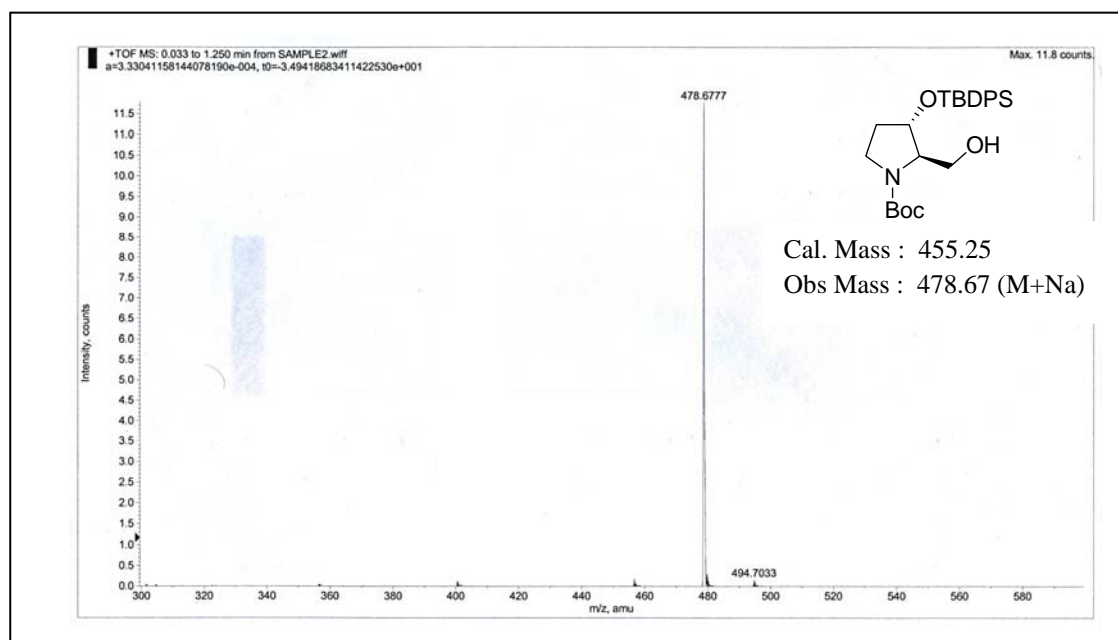
**Figure 33: (A) Mass spectral analysis of compound 26 (B) Mass spectral analysis of compound 27**



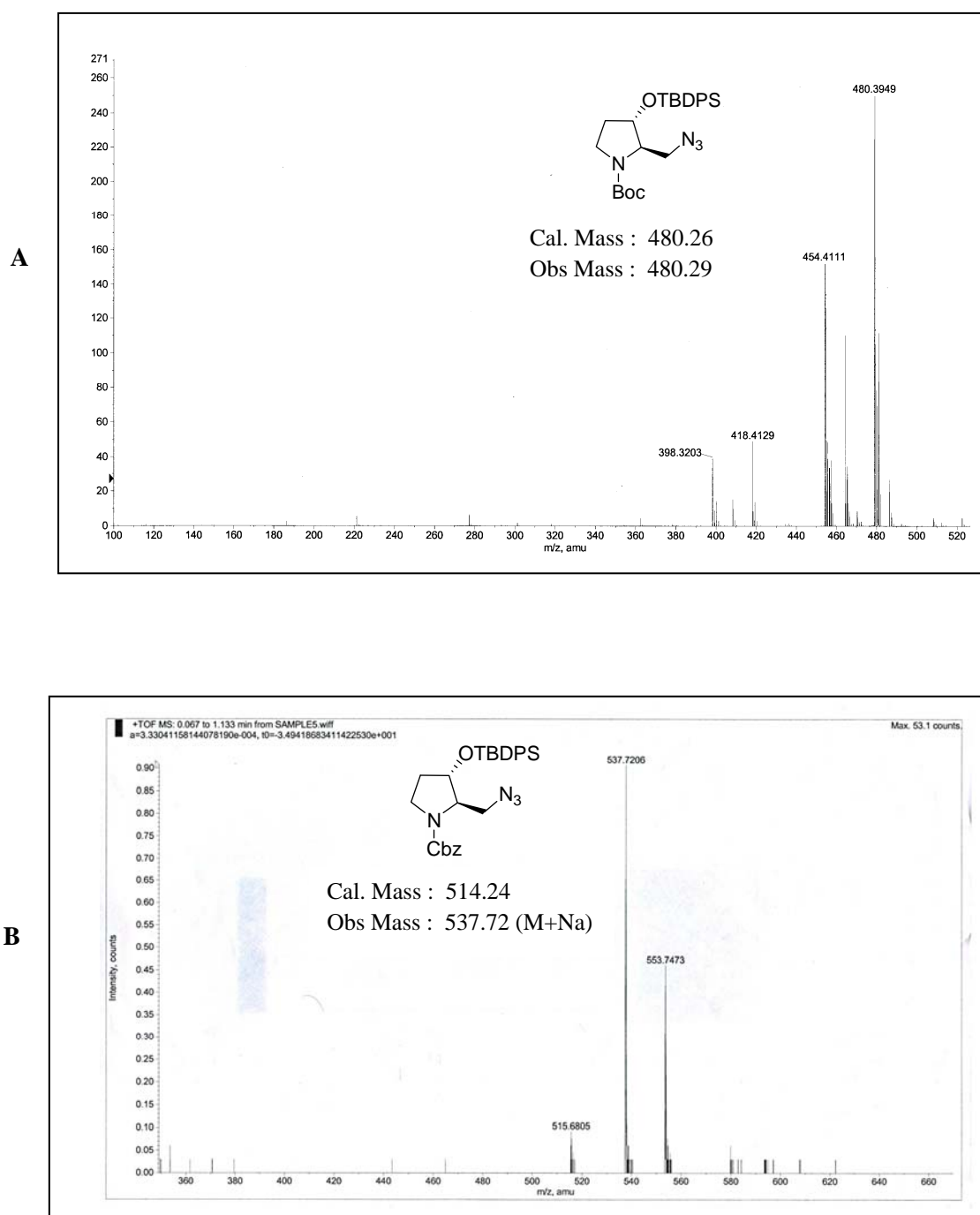
A



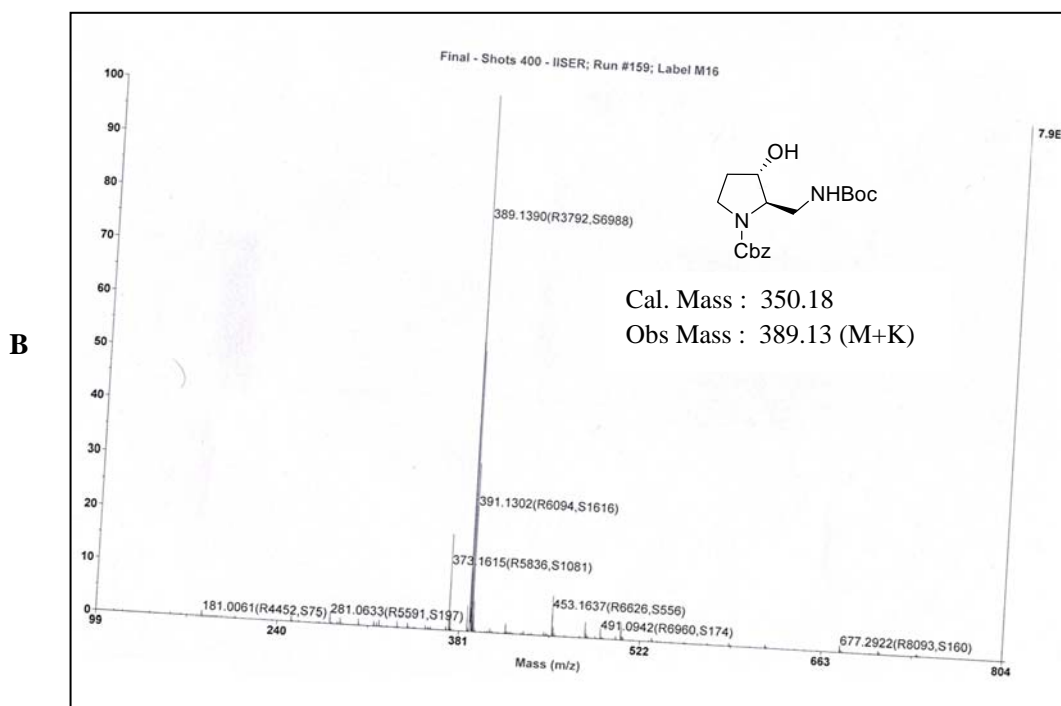
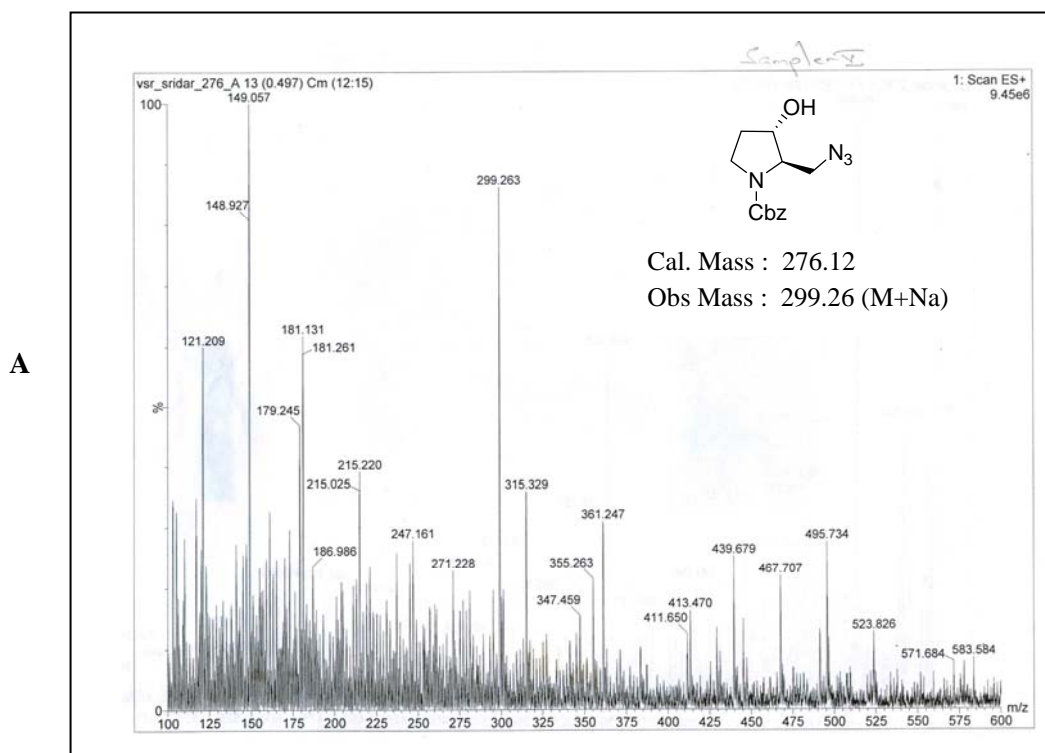
B



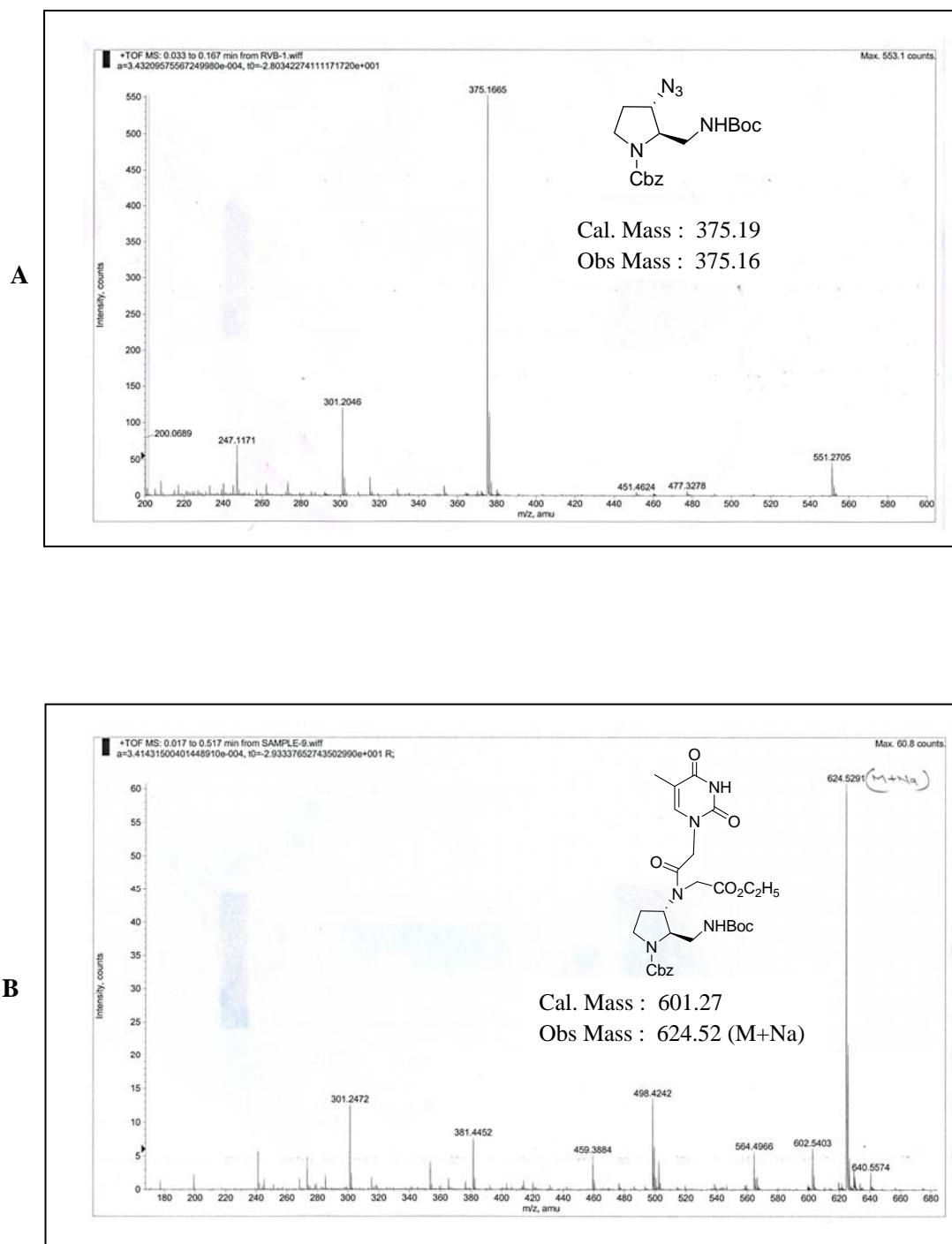
**Figure 34:** (A) Mass spectral analysis of compound 28 (B) Mass spectral analysis of compound 29



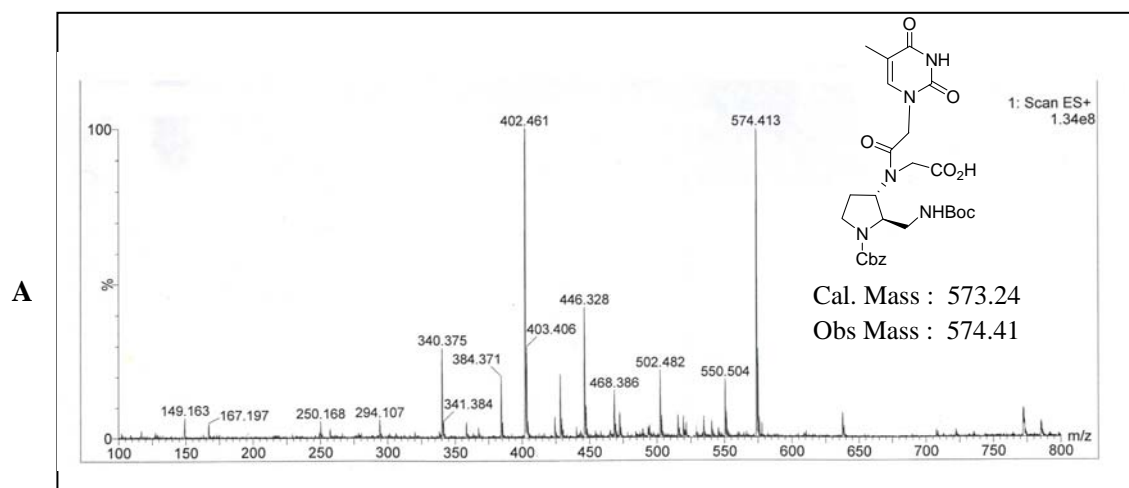
**Figure 35: (A) Mass spectral analysis of compound 30 (B) Mass spectral analysis of compound 31**



**Figure 36:** (A) Mass spectral analysis of compound **32** (B) Mass spectral analysis of compound **33**

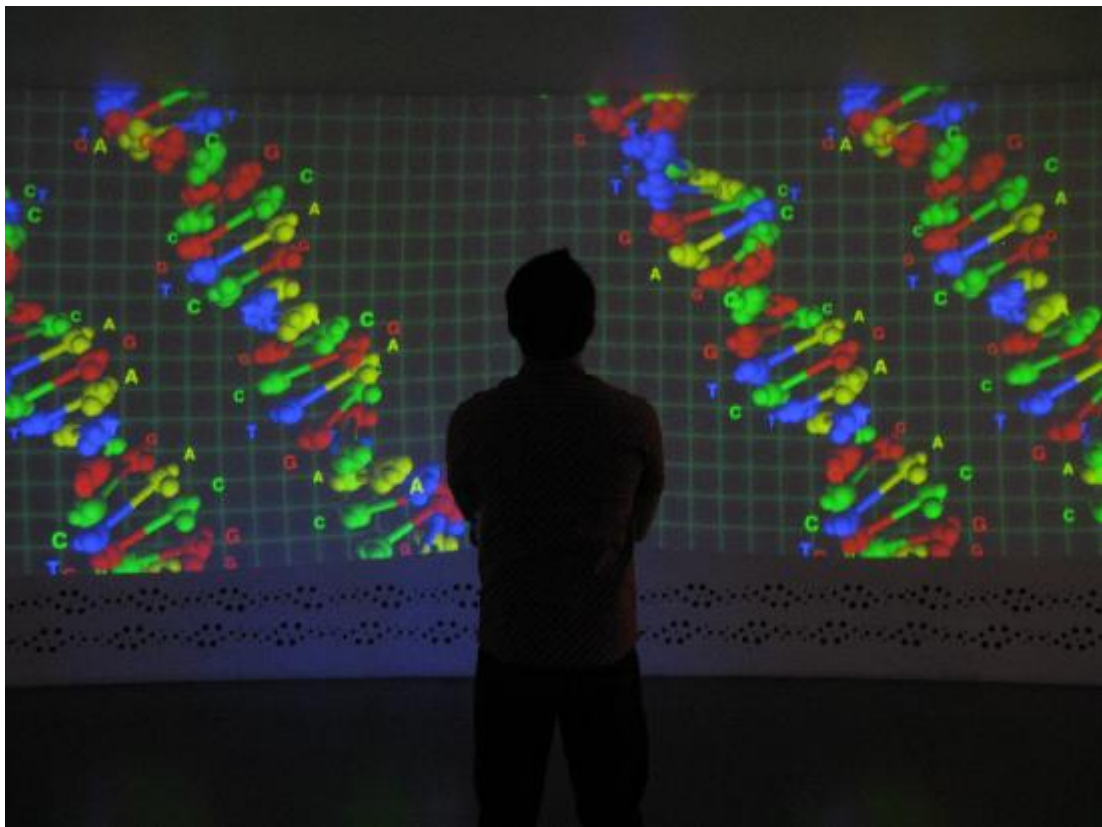


**Figure 37: (A) Mass spectral analysis of compound 35 (B) Mass spectral analysis of compound 38**



**Figure 38: (A)** Mass spectral analysis of compound **39**

## Chapter 3: Solid phase synthesis and biophysical studies of chiral, cationic *dap* and *amp*PNA derived oligomers.



## Introduction

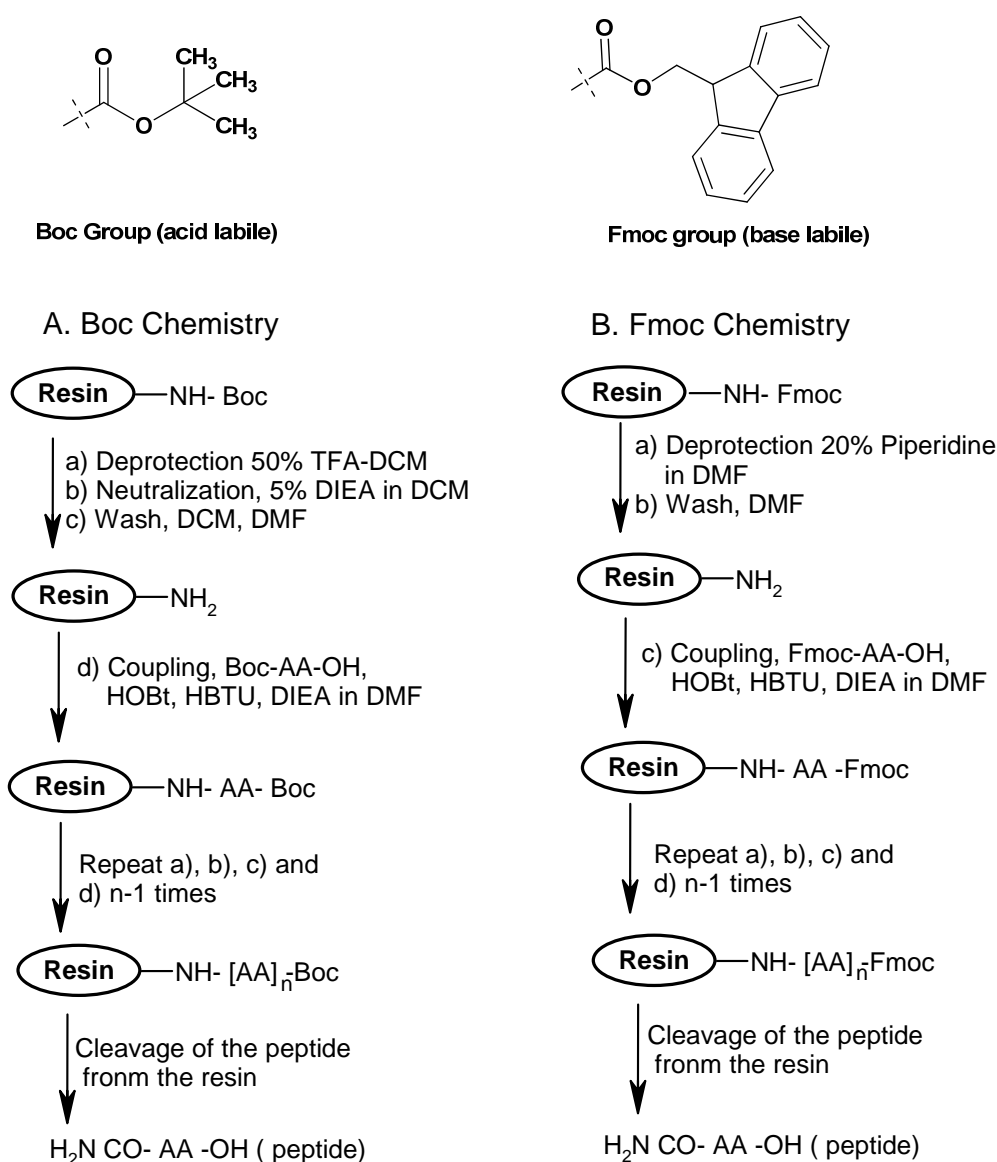
The main goal of designing antisense/antigene oligonucleotides is to achieve selective and high binding affinity of the therapeutic agent to the target nucleic acid and make them stable to cellular enzymes with a long enough half-life within the cells to allow manifestation of its effect. It should not be toxic to the cell in which its activity is required. In this context PNA and their analogues have shown great potential.

In this chapter, the designed and synthesized PNA monomers of chapter 2 were site specifically incorporated into *aeg*PNA oligomers by solid phase peptide synthesis, to study the influence of backbone modified *aeg*PNA monomers on hybridization properties of PNA oligomers. The binding selectivity and specificity of modified PNAs towards complementary DNA was investigated using the biophysical techniques mainly temperature dependent UV-spectroscopy. The UV-melting studies were carried out with all modified PNA:DNA complexes and analyzed with respect to that of control PNA:DNA complexes. These give information on the stability of the complexes of the modified PNAs with DNA. The following sections describe in detail the solid phase peptide synthesis and results of biophysical studies.

### 3.1 Solid Phase Peptide Synthesis

The solid phase peptide synthesis (SPPS)<sup>1</sup> was invented by R. B. Merrifield in 1959. In this method, the peptide is synthesized on an insoluble solid support, in contrast to the solution phase technique and the solid phase method offers great advantages. In this method, the C-terminus amino acid is linked to an insoluble matrix such as polystyrene beads having reactive functional groups, which also act as a protection for the carboxyl group (Figure 1). The desired  $N^\alpha$ -protected amino acid is coupled to the resin bound amino acid either by using an active pentafluorophenyl (pfp) or 3-hydroxy-2,3-dihydro-4-oxo-benzotriazole (Dhbt) ester or by *in situ* activation with carbodiimide reagents or other coupling reagents. The excess amino acid is washed out and the deprotection/coupling reactions are repeated until the desired peptide is achieved. The need to purify the intermediates at every step is obviated. Finally, the resin bound peptide and the side chain protecting groups are cleaved in one step. The advantages of the solid phase synthesis are that i) all the reactions are performed in a single vessel minimizing the loss due to transfer and ii) the excess molar equivalents of monomer i.e

carboxylic acid component can be used to achieve high coupling efficiency. The use of an insoluble polymer support greatly simplifies the synthesis of peptides; after each coupling step, the purification of the growing peptide chain is accomplished by a simple filtration procedure. Its major advantage is that after, each reaction, the corresponding soluble reagents/by products are removed effectively from the retained polymer bound macromolecules and the method is amenable to automation.



**Figure 1.** Schematic representation of SPPS A) Boc-chemistry B) Fmoc-chemistry



Currently, two types of chemistry are available for the routine synthesis of peptides by solid phase method (Figure 1). When Merrifield invented SPPS in 1963, it was according to the *t*-Boc method, in which *t*-butyloxycarbonyl (*t*-Boc) group is used as  $N^\alpha$ -protection that could be removed under acidic conditions such as 50% TFA in DCM.<sup>2</sup> The reactive side chains of amino acids were protected with groups which are stable to *t*-Boc deprotection conditions (Cbz, benzyl etc.), removable under strongly acidic conditions like HF in dimethyl sulfide or TFMSA in TFA. Alternatively, a base labile protecting group like fluorenyl methyloxycarbonyl (Fmoc) can be used for  $N^\alpha$ -protection. This is stable to acidic conditions but can be deprotected efficiently with a secondary base such as piperidine. Fmoc method was introduced by Carpino *et al*<sup>3</sup> in 1972. Coupled with mild acid sensitive side chain protection, this method offers a second strategy for SPPS. In both chemistries, the linker group that joins the peptide to the resin is chosen such that the side chain protecting groups and the linker are cleaved in one step at the end of the peptide synthesis. Although there are other synthetic strategies available for solid-phase peptide synthesis the Fmoc approach has proven the most versatile.<sup>26,27</sup> Lee *et al.* developed a simple, and inexpensive biphasic functionalization approach for preparing an ideal core-shell-type resin<sup>28</sup> and Ernesto Nicolás *et al* introduced a triazine linker for preparing phenylalanine containing peptides.<sup>29</sup>

### 3.1.1 Resins for solid phase peptide synthesis

#### 3.1.1a Merrifield resin

Merrifield resin (entry 1, Table 2) consists of divinylbenzene cross-linked polystyrene beads functionalized with chloromethyl groups. The attachment of N-Boc amino acids is generally achieved by heating the resin, in DMF, with the appropriate carboxylic acid cesium salt in the presence of KI or dibenzo-18-crown-6. The target peptide is released from the resin by treatment with HF or TFMSA.

#### 3.1.1b Wang resin

Wang resin (entry 1, Table 1) is the standard support and most widely used for Fmoc SPPS method. This is synthesized by reacting chloromethyl polystyrene resin with 4-hydroxybenzyl alcohol. The attachment of amino acids to benzyl alcohol functionalized support is normally achieved by DMAP catalyzed esterification with the

appropriate symmetrical anhydride using 2,6-dichlorobenzoyl chloride activation. Wang resin has very good physical properties well suited to SPOS (solid-phase oligonucleotide synthesis). Linking of carboxylic acids, phenols and alcohols can be done directly to wang resin. The release of these functionalities is usually affected by mild acidolysis with TFA.

### 3.1.1c MBHA resin

MBHA resin (entry 4, Table 2) is composed of 1% DVP cross-linked polystyrene functionalized with 4-methylphenyl-aminomethyl group. This resin is the most widely used one for the synthesis of peptides by the *t*-Boc strategy.<sup>4</sup> The cleavage of the peptides from the resin is achieved by treating with HF, TFMSA or HBF<sub>4</sub>. In Fmoc SPPS, this resin is used in conjunction with a low-high acid cleavage, where side-chain protecting groups such as benzyloxycarbonyl, benzoyl etc were first removed prior to cleavage from the resin with TFMSA. The excellent swelling properties of this resin makes it an ideal matrix onto which TFA-labile linkers are attached for combinatorial synthesis.

### 3.1.1d Rink amide resin

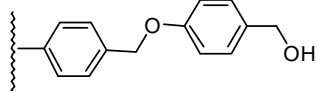
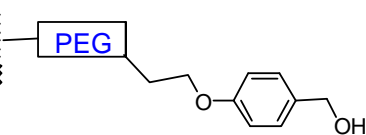
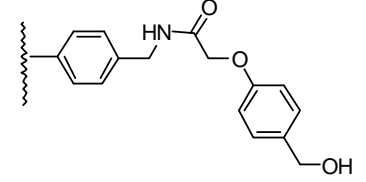
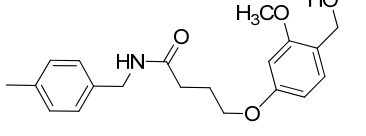
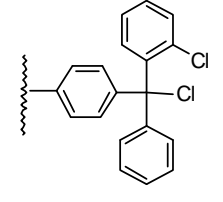
Rink amide resin (entry 6, Table 1) is highly acid sensitive because of the benzhydryl linker joined to the solid support through a benzylic ether bond. The linker during cleavage with high concentration of TFA leading to the formation of highly coloured by-products. This problem is minimized through the use of low concentrations of TFA, or by the addition of trialkylsilanes as the scavengers to the cleavage mixture.<sup>5</sup> Generally, two types of Rink amide resins are available. In the first, the benzhydrylamine handle is linked directly to the polystyrene base matrix via an ether bond (Rink amide resin) and in the second, the Rink amide linker is attached to an appropriate amino-functionalized resin via an amide bond (Rink AM resin, Rink MBHA resin). The later resins are more robust under acidic conditions.

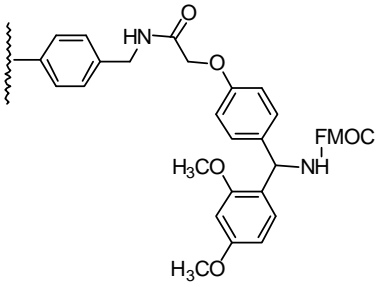
### 3.1.1e Trityl resin

The advantage of trityl functionalized resin (entry 5, Table 1) is that any nucleophile can be linked to the resin under extremely mild conditions. Most of the functional side-chains of the amino acids can be attached to the 2-chlorotrityl resin. Furthermore, symmetrical bifunctional compounds, such as diamines, diols, diphenols

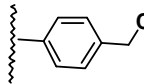
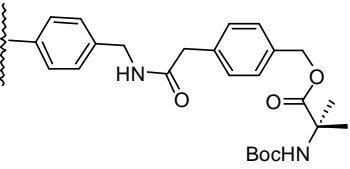
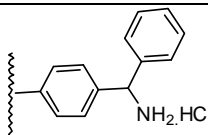
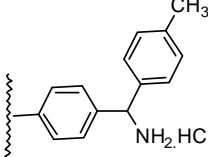
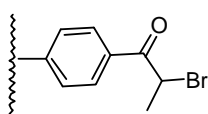
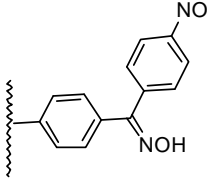
and diacids, are in effect monoprotected by this process, allowing one end of the molecule to be selectively modified. Cleavage of the product from the support takes place under mild base conditions, owing to the high stability of the trityl cations. Furthermore, trityl cations are extremely poor electrophiles and do not undergo alkylation side-reactions.<sup>6</sup>

**Table 1: Resins for boc chemistry**

S.No	Resin	Structure	Final products	Cleavage conditions
1	Wang resin		Peptide acids	TFA (Scavengers)
2	PHB resins		Peptide acids	TFA (scavengers)
3	HMPA resins		Peptide acids	TFA (Scavengers)
4	HMPB resins		Protected peptide acids	Dilute TFA(1-5%)
5	2-chlorotrityl resin		Protected peptide acids	AcOH/TFA/DCM(1,8,8 v/v/v) 0.5TFA in DCM

6	Rink amide ("Knorr") resin		Peptide carboxamides	TFA scavengers
---	----------------------------------	---	-------------------------	-------------------

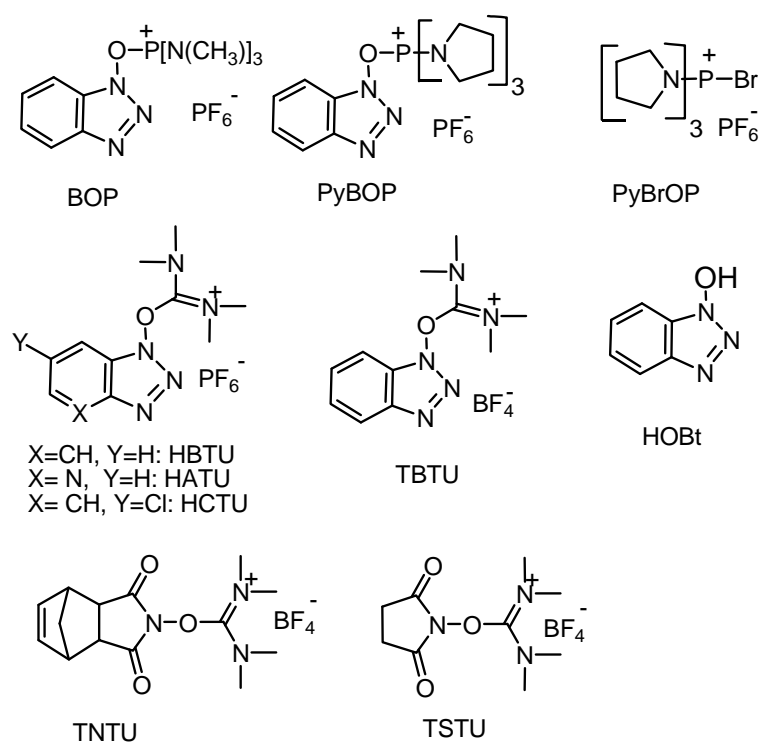
**Table 2: Resins for fmoc chemistry**

S.No	Resin	Structure	Final products	Cleavage conditions
1	Merrifield resin		Peptide acids	HF, TFMSA, TMSOTf, HBr/TFA
2	PAM resin		Peptide acids	HF, TFMSA, TMSOTf
3	BHA resin		Peptide carboxamides	HF, TFMSA
4	MBHA resin		Peptide carboxamides	HF, TFMSA
5	Brominated Wang resin		Protected peptide acids	Hv (350 nm)
6	4-Nitrobenzo phenone oxime(Kaiser) resin		Protected peptide acids Peptide amides Peptide esters	NaOH NH <sub>3</sub> , RNH <sub>2</sub> ROH

### 3.1.2 Coupling reagents

Carbodiimides are some of the most popular *in situ* activating reagents in peptide synthesis. Dicyclohexylcarbodiimide (DCC) is the most popular choice of coupling agent for the apolar polystyrene resins.<sup>7</sup> The principal limitation in using carbodiimides is the dehydration of amide side chains of Asn and Gln residues to nitriles. The addition of HOBt (Hydroxybenzotriazole) to the reaction mixture will prevent such dehydration and has the added benefit of catalysis of amide bond formation. The principal drawback of carbodiimides is the formation of unseparable urea by-products during activation and acylation.

*In situ* activating reagents are widely accepted due to ease of use, fast reactions even with sterically hindered amino acids and free of side reactions. Most reagents used are based on phosphonium or uronium salts which in presence of a tertiary base, can smoothly convert N-protected amino acids to a variety of activated species.



**Figure 2.** Coupling reagents used for Peptide synthesis.

The most commonly employed reagents have benzotriazole moiety e.g. BOP, PyBOP, HBTU and TBTU which generate -OBt (benzotriazole) ester and these have wide application in routine SPPS and solution synthesis for difficult couplings (Figure 2). BOP also leads to highly carcinogenic by-product during. BOP can be substituted by PyBOP, which has pyrrolidine unit instead of N-methyl in BOP, without any loss of performance.

Coupling reagents that generate esters which are more reactive than OBt esters have been developed. Two such reagents are HATU<sup>8</sup> and HCTU<sup>9</sup>, which in presence of a base convert carboxylic acids to corresponding -OAt (7-azabenzotriazole) and -OCt (6-chlorobenzotriazole) esters respectively. These esters are more reactive than their OBt counterparts owing to the lower pKa of HOAt and HOCT compared to HOBt. Furthermore, HOAt has the added advantage of the pyridine nitrogen which provides anchimeric assistance to the coupling reaction. HATU is the most efficient coupling reagent and suitable for difficult and hindered couplings and in the synthesis of long peptides.

The coupling of N-methyl amino acids is a difficult and slow process. PyBOP<sup>10</sup> allows efficient and fast couplings, providing a significant improvement over the existing reagents.<sup>11</sup>

### 3.1.3 Resin Tests

The most widely used procedure for determining the efficiency of coupling is to test for the presence or absence of free amino groups after deprotection or coupling by Kaiser test<sup>12</sup> which is simple and quick (details in experimental section). However some deprotected amino acids do not show the expected dark blue color typical of free primary amino groups (e.g., serine, asparagine, aspartic acid) and proline, being a secondary amino acid, does not yield a positive reaction. Furthermore, occasionally false negative tests are observed, particularly with strongly aggregated sequences, requiring caution while assessing the extent of deprotection or coupling on solid phase. Other methods such as picric acid test<sup>13</sup> or mass spectrometry<sup>14</sup> are also available. For iminoacids like proline, the chloranil test<sup>15</sup> is recommended. Recently a new method for detecting resin-bound primary and secondary amines has been described<sup>16</sup> in which

the resins are treated with 2,3-dichloro-5-nitro-1,4-naphthoquinone in DMF or DCM in the presence of 2,6- di-*tert*-butylpyridine.

### 3.1.4 General protocols adopted for PNA synthesis

As in the case of solid phase peptide synthesis, PNA synthesis is done conveniently from the C- terminus to the N- terminus. For this, the monomeric units must have their amino functions suitably protected, with free carboxylic acid functions. MBHA resin or Rink Amide resin was used as the solid support on which the oligomers were built and the monomers were coupled by *in situ* activation with HBTU/HOBt using either boc or fmoc solid phase strategy. Commercially available MBHA resin and Rink amide resin have a loading value of 2 meq/g and 0.7meq/g respectively, which is not suitable for oligomer synthesis. Since, at this level of loading, PNA oligomers aggregate thereby decreasing the efficiency in successive coupling steps. Hence, it was necessary to lower the loading value to 0.25-0.35 meq/g to avoid the aggregation of growing oligomer. This was achieved by partial acetylation of the amine content with calculated amount of acetic anhydride (capping). The amount of free -NH<sub>2</sub> on the resin available for coupling was estimated by the picrate assay,<sup>17</sup> before starting the PNA synthesis. In the synthesis of all oligomers, orthogonally protected (Boc/Cl-Cbz) L-lysine was selected as the C-terminus spacer-amino acid and it is linked to the resin through amide bond. The introduction of charged groups, for instance a C-terminus lysine amide greatly improves the water solubility of the resulting oligomers. The amino function of the monomers was protected as the corresponding Boc or Fmoc derivative and the carboxylic acid function was free to enable coupling with the resin-linked monomer. Coupling of the resin linked free amine after deprotection was done with 3 to 4 equivalents of free carboxylic function of the incoming PNA monomer using HOBt/HBTU in DMF:NMP(1:1) as solvent and DIPEA as base.

### 3.1.5 Cleavage of the PNA oligomers from the solid support

The oligomers were cleaved from the solid support L-lysine derivatized MBHA resin, using trifluoromethanesulphonic acid (TFMSA) in the presence of trifluoroacetic acid (TFA) (Low, High TFMSA-TFA method)<sup>18</sup> and using 50% TFA in DCM from the solid support L-lysine derivatized Rink Amide resin, which yields oligomer having L-lysine-amide at their C-termini. A cleavage time of 1.5-2 h at room temperature was

found to be optimum. The side chain protecting groups were also cleaved during this cleavage process. For Fmoc protection strategy of solid phase using rink amide resin, N-*fmoc* group at N terminus of PNA sequence was removed using 20% piperidine in DMF followed by cleavage from resin using 50% TFA in DCM. After cleavage reaction, the oligomer was precipitated out by diethyl ether and purified by reverse phase HPLC.

### 3.2 Biophysical studies of PNA:DNA complexes

The binding selectivity and specificity of modified PNAs towards complementary DNA has been investigated using biophysical techniques, temperature dependent UV-spectroscopy. The UV-melting studies were carried out with all modified PNA:DNA complexes and analyzed with respect to that of control PNA:DNA complexes. These give information on the thermal stability of the complexes of modified PNAs with DNA.

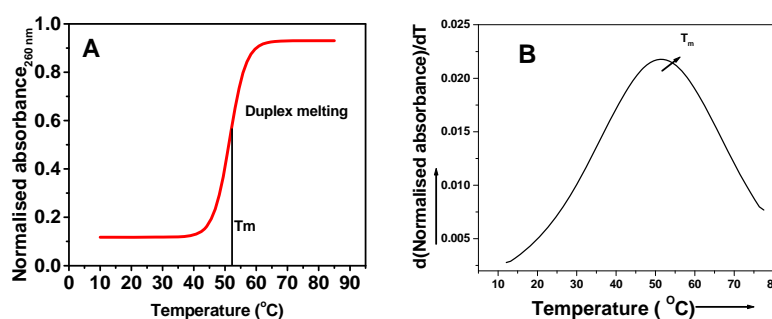
#### 3.2.1 UV-Spectroscopic study of PNA:DNA complexes

Monitoring the UV absorption at 260 nm as a function of temperature has been extensively used to study the thermal stability of nucleic acid complexes and consequently, PNA:DNA/RNA hybrids were investigated by this technique. Increase in temperature induces dissociation of the complex into individual components by rupturing hydrogen bonds between the complementary base pairs. This leads to decrease of stacking between adjacent nucleobases resulting in loss of secondary and tertiary structure. This is evidenced by an increase in the UV absorption at 260 nm with temperature rise, termed as hyperchromicity. A plot of absorbance *versus* temperature gives a sigmoidal curve in case of duplexes/triplexes and the midpoint of transition gives the  $T_m$  (Figure 3A). In case of DNA triplexes, double transition sigmoidal curve is seen, the first lower temperature transition corresponding to triplex melting to the duplex (Watson–Crick duplex), whereas the second higher temperature transition arises from duplex dissociation into two single strands.<sup>19</sup> In case of PNA:DNA complexes, the PNA<sub>2</sub>:DNA triplexes show single sigmoidal curve since the temperature difference between two transition states is too low to distinguish them independently.<sup>20</sup> The two PNA strands dissociate from DNA strand almost simultaneously. The melting temperature  $T_m$  can be obtained from the maxima of the first derivative plots (Figure



**3B).** This technique has provided valuable information regarding complementary interactions in nucleic acid hybrids involving DNA, RNA and PNA.

The fidelity of base-pairing in PNA:DNA complexes can be examined by challenging the PNA oligomer with a DNA strand bearing mismatches. Even one base mismatch with incorrect hydrogen bonding between the bases causes a drop in the melting temperature. A modification of the PNA structure is considered good if it gives a much lower  $T_m$  with DNA sequences containing mismatches as compared to complexes of unmodified PNA. This clearly suggests that specificity arises from complementary hydrogen bonding and not due to other effects.



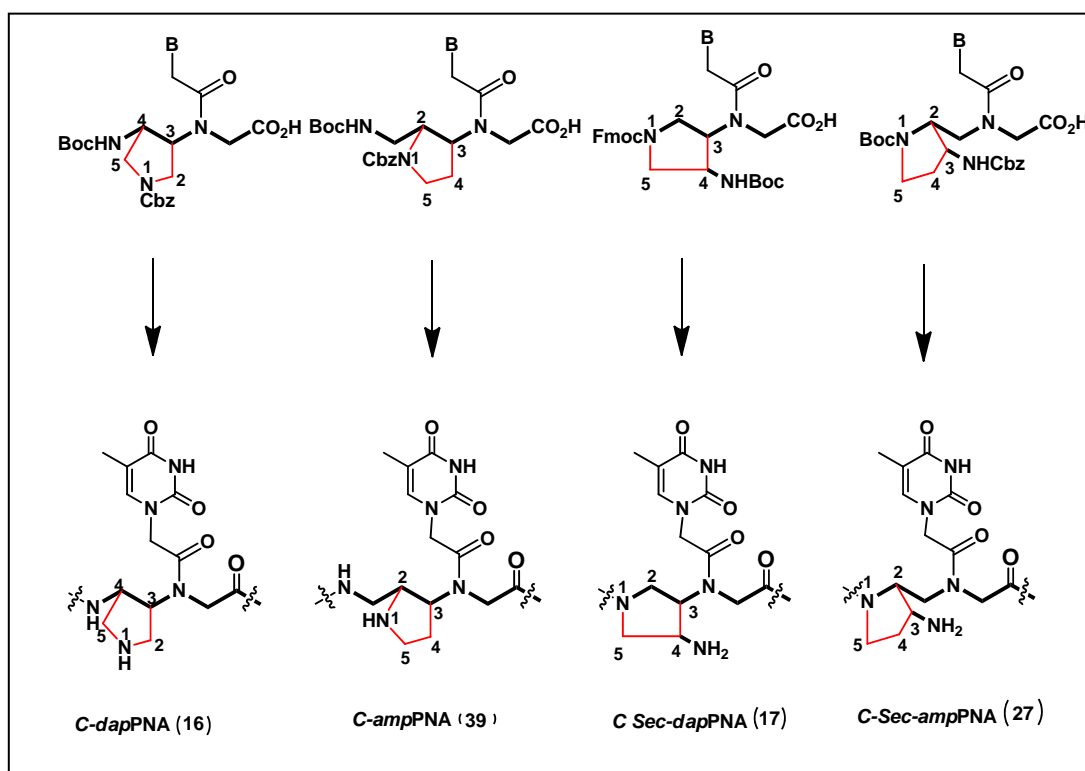
**Figure 3:** Schematic representation of UV melt A) Sigmoidal curve, B) First derivative curve.

### 3.3 Results and Discussion

#### 3.3.1 Solid Phase synthesis of 3→4dap (16), 1→3dap (17), 2→3amp (39) and 1→2ampPNA (27) oligomers.

The modified cationic pyrrolidine PNA monomers, **16**, **17**, **27** and **39** (Figure-4), synthesized as described in Chapter-2 were incorporated into standard PNA oligomers by using either Boc or Fmoc chemistry for solid phase synthesis leading to four different PNA backbones (Figure 4). The extra primary or secondary amine functions of these oligomers are capable of accepting a proton and help to make PNA cationic in all the cases.

These oligomers have “*cis*” chemistry at  $C_2$ - $C_3$  or  $C_3$ - $C_4$  carbon atoms and the term “3→4dap, 1→3dap” represents the linkage of backbone in *dap* derived monomers (**16**, **17**, Figure 4) *i.e.* backbone linkage is between  $C_3$ ,  $C_4$  and  $C_1$ ,  $C_3$  carbons.



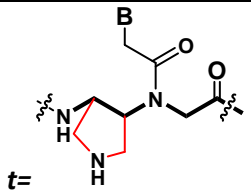
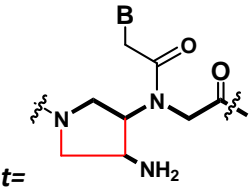
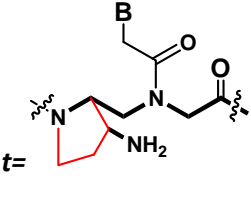
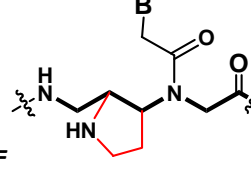
**Figure 4:** Backbones (shown by thick lines) derived from modified PNA monomers.

the term “ $2 \rightarrow 3amp$ ,  $1 \rightarrow 2amp$ ” represents the linkage of backbone in *amp* derived monomers (27, 39, Figure 4) *i.e.* backbone linkage is between C2, C3 and C1, C2 carbons.

The terms “*dap*” and “*amp*” stands for diamino pyrrolidine and amino proline derived monomers respectively.

The synthesis of the designed oligomers was achieved by incorporating the conformationally constrained chiral modified units (Figure 4: 16, 17, 27 and 39) at specific positions in the *aegPNA* oligomer. This was done on the solid support using either boc chemistry or fmoc chemistry procedures as described above. The mixed purine-pyrimidine oligomers (PNA 1-17, Table 2) were synthesized for comparative duplex studies with mixed control *aegPNA* oligomer.

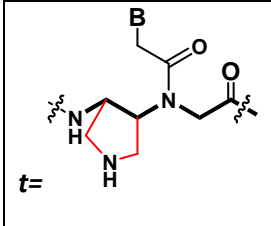
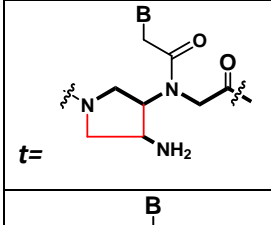
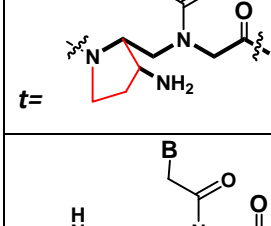
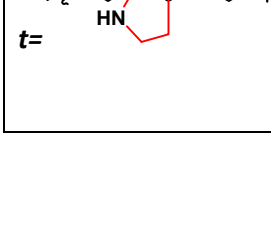
Table 3: Oligomers synthesized

Modified unit	S.No	Sequence
Control	PNA 1	H-GTAGATCACT-Lys-NH <sub>2</sub>
	PNA 2	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>
	PNA 3	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 4	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 5	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>
	PNA 6	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>
	PNA 7	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 8	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 9	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>
	PNA 10	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>
	PNA 11	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 12	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 13	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>
	PNA 14	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>
	PNA 15	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 16	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>
	PNA 17	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>

### 3.3.2 Purification of the PNA oligomers

All PNA oligomers after cleavage were subjected to HPLC purification. The purity of the so obtained oligomers were checked by analytical HPLC (C18 column, CH<sub>3</sub>CN-H<sub>2</sub>O system). The purity of the oligomers were again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis.

**Table 4:** MALDI TOF mass spectral analysis of modified PNAs

Modified unit	S.No	Sequence	Calculated mass	Observed mass
Control	PNA 1	H-GTAGATCACT-Lys-NH <sub>2</sub>	2852.52	2853.37
	PNA 2	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	2893.64	2896.20
	PNA 3	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2893.64	2917.29
	PNA 4	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2934.73	2958.23
	PNA 5	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	2893.64	2896.53
	PNA 6	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	2893.64	2896.88
	PNA 7	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2893.64	2917.30
	PNA 8	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2934.73	2936.28
	PNA 9	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	2893.64	2895.49
	PNA 10	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	2907.51	2911.61
	PNA 11	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2907.51	2909.02
	PNA 12	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2962.85	2961.24
	PNA 13	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	2907.51	2948.17
	PNA 14	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	2907.51	2909.16
	PNA 15	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2907.51	2910.10
	PNA 16	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	2962.85	2963.14
	PNA 17	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	2907.51	2911.65

### 3.3.3 Synthesis of complementary DNA oligonucleotides

The oligonucleotides DNA1 and DNA2 (Table 5) corresponding to complementary (parallel and antiparallel) PNA sequences were synthesized on Applied Biosystems ABI 3900 DNA Synthesizer using standard  $\beta$ -cyanoethyl phosphoramidite chemistry. The oligonucleotides were synthesized in the 3' to 5' direction on polystyrene solid support, followed by ammonia treatment. These were desalted by gel filtration, and their purity as ascertained by RP HPLC on a C-18 column was found to be more than 98%. They were used without further purification for the biophysical studies of hybridization with PNAs.

**Table-5:** DNA oligonucleotides used in the present work

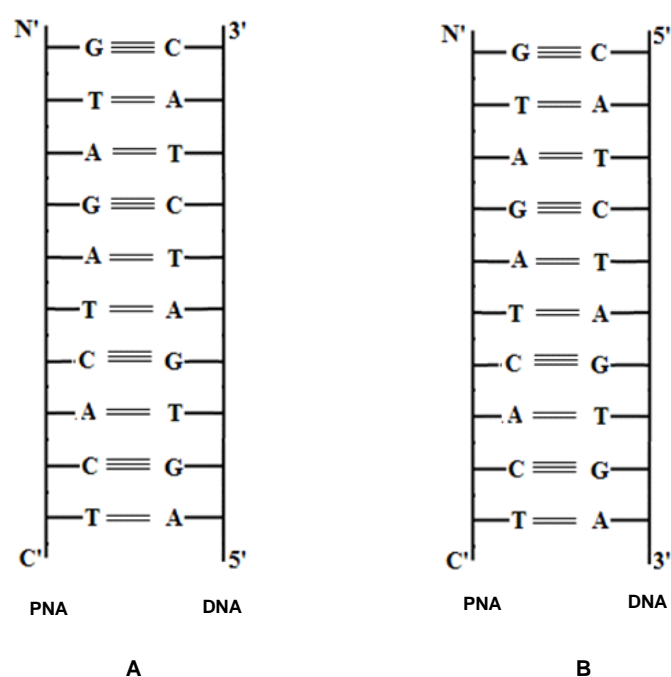
S. No.	Entry	Sequence (5'-3')	Type
1	DNA 1	AGTGATCTAC	Antiparallel
2	DNA 2	CATCTAGTGA	Parallel

### 3.3.4 Biophysical studies of PNA:DNA complexes

To study the effect of backbone modifications on thermal stability, binding specificity and base discrimination of the modified PNAs towards complementary DNA hybridization, biophysical studies were done.

#### 3.3.4a UV-Spectroscopy

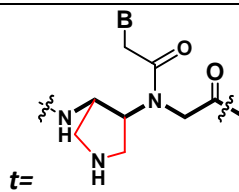
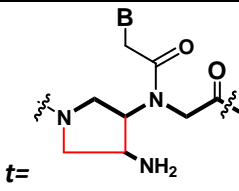
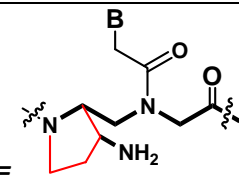
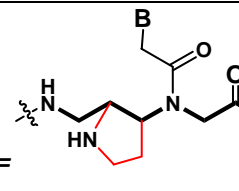
The hybridization studies of modified PNAs with complementary DNA were followed by temperature dependent UV-absorbance experiments. PNA:DNA duplexes were constituted with complementary **DNA1** (antiparallel) and **DNA2** (parallel) for the binding studies.



**Figure 5:** (A) Antiparallel mode of binding of PNA:DNA duplex, (B) Parallel mode of binding of PNA:DNA duplex

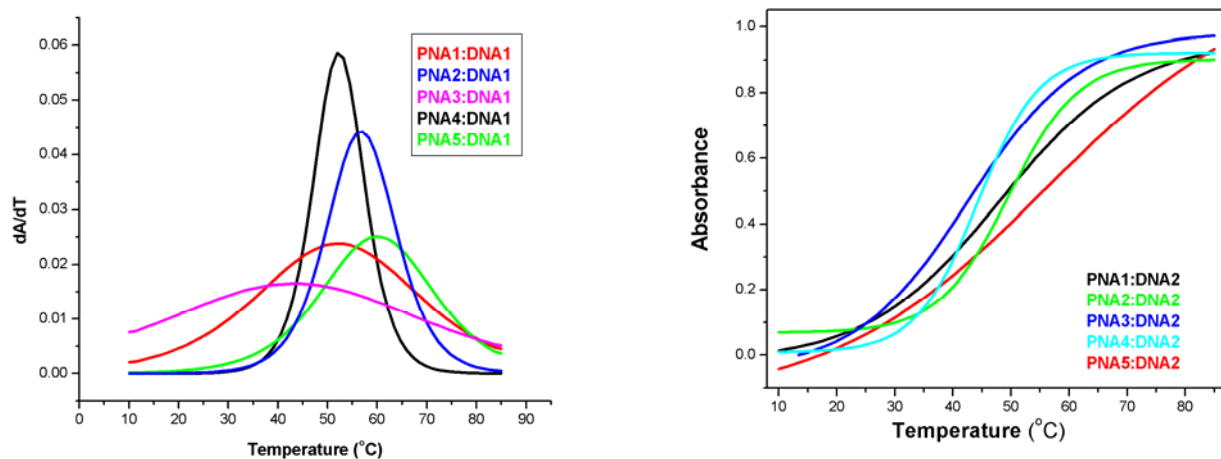
## 3.3.4b Thermal stability of PNA:DNA duplexes (antiparallel/parallel)

Table 6: UV melting Temp( $^{\circ}\text{C}$ ) of PNA:DNA duplexes

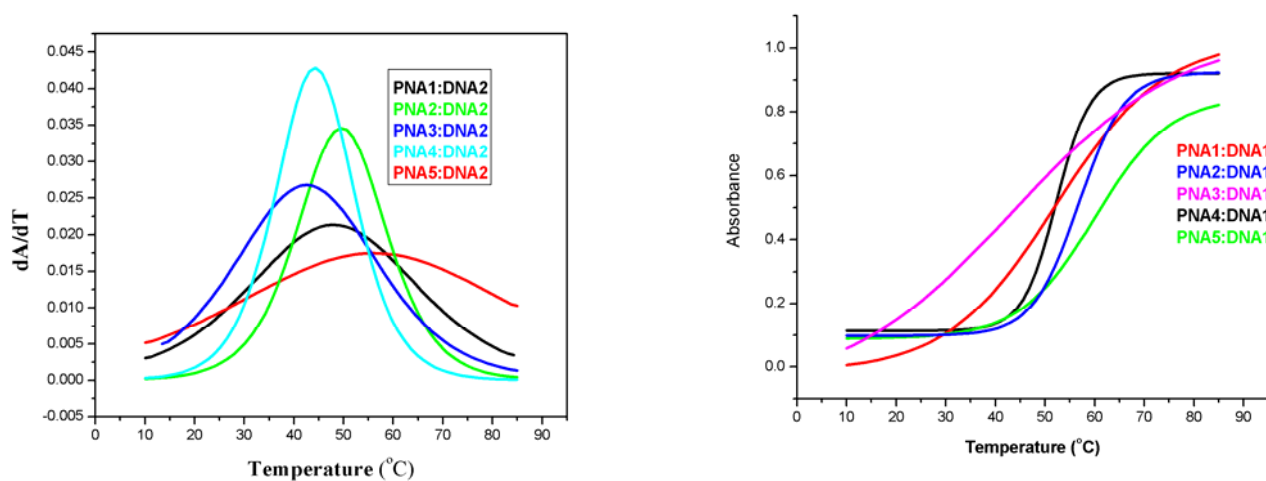
Modified unit	S.No	Sequence	UV-Tm	
			DNA 1(ap) ( $\Delta\text{Tm}_1$ )	DNA 2(p) ( $\Delta\text{Tm}_2$ )
Control	PNA 1	H-GTAGATCACT-Lys-NH <sub>2</sub>	52 $^{\circ}\text{C}$	48 $^{\circ}\text{C}$
	PNA 2	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	56 $^{\circ}\text{C}$ (+4 $^{\circ}\text{C}$ )	49.5 $^{\circ}\text{C}$ (+1.5 $^{\circ}\text{C}$ )
	PNA 3	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	45.5 $^{\circ}\text{C}$ (-0.5 $^{\circ}\text{C}$ )	42.5 $^{\circ}\text{C}$ (+5.5 $^{\circ}\text{C}$ )
	PNA 5	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	60 $^{\circ}\text{C}$ (+8 $^{\circ}\text{C}$ )	55.5 $^{\circ}\text{C}$ (+7.5 $^{\circ}\text{C}$ )
	PNA 4	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	53 $^{\circ}\text{C}$ (-3 $^{\circ}\text{C}$ )	45 $^{\circ}\text{C}$ (-3 $^{\circ}\text{C}$ )
	PNA 6	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	43 $^{\circ}\text{C}$ (-9 $^{\circ}\text{C}$ )	46.3 $^{\circ}\text{C}$ (-1.7 $^{\circ}\text{C}$ )
	PNA 7	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	48 $^{\circ}\text{C}$ (-4 $^{\circ}\text{C}$ )	52.6 $^{\circ}\text{C}$ (+4.6 $^{\circ}\text{C}$ )
	PNA 9	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	54 $^{\circ}\text{C}$ (+2 $^{\circ}\text{C}$ )	47.7 $^{\circ}\text{C}$ (-0.3 $^{\circ}\text{C}$ )
	PNA 8	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	55 $^{\circ}\text{C}$ (+3 $^{\circ}\text{C}$ )	49.3 $^{\circ}\text{C}$ (+1.3 $^{\circ}\text{C}$ )
	PNA 10	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	56 $^{\circ}\text{C}$ (+4 $^{\circ}\text{C}$ )	45 $^{\circ}\text{C}$ (+3 $^{\circ}\text{C}$ )
	PNA 11	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	48 $^{\circ}\text{C}$ (-4 $^{\circ}\text{C}$ )	54.3 $^{\circ}\text{C}$ (+6.3 $^{\circ}\text{C}$ )
	PNA 13	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	60 $^{\circ}\text{C}$ (+8 $^{\circ}\text{C}$ )	43.2 $^{\circ}\text{C}$ (-4.8 $^{\circ}\text{C}$ )
	PNA 12	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	50.5 $^{\circ}\text{C}$ (-0.5 $^{\circ}\text{C}$ )	50.8 $^{\circ}\text{C}$ (+2.8 $^{\circ}\text{C}$ )
	PNA 14	H-GTAGATCAC <sub>t</sub> -Lys-NH <sub>2</sub>	60.3 $^{\circ}\text{C}$ (+8.3 $^{\circ}\text{C}$ )	54 $^{\circ}\text{C}$ (+6 $^{\circ}\text{C}$ )
	PNA 15	H-GTAGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	56 $^{\circ}\text{C}$ (+4 $^{\circ}\text{C}$ )	58.3 $^{\circ}\text{C}$ (+10.3 $^{\circ}\text{C}$ )
	PNA 17	H-G <sub>t</sub> AGATCACT-Lys-NH <sub>2</sub>	65.5 $^{\circ}\text{C}$ (13.5 $^{\circ}\text{C}$ )	47.5 $^{\circ}\text{C}$ (-0.5 $^{\circ}\text{C}$ )
	PNA 16	H-G <sub>t</sub> AGA <sub>t</sub> CACT-Lys-NH <sub>2</sub>	51 $^{\circ}\text{C}$ (-1 $^{\circ}\text{C}$ )	48.3 $^{\circ}\text{C}$ (+0.3 $^{\circ}\text{C}$ )

DNA 1 = 5'-AGTGATCTAC-3'(antiparallel), DNA 2 = 5'-CATCTAGTGA-3'(parallel). A,G,C,T = aegPNA. ( $\Delta\text{Tm}_1$ ) and ( $\Delta\text{Tm}_2$ ) = difference w.r.t. control,

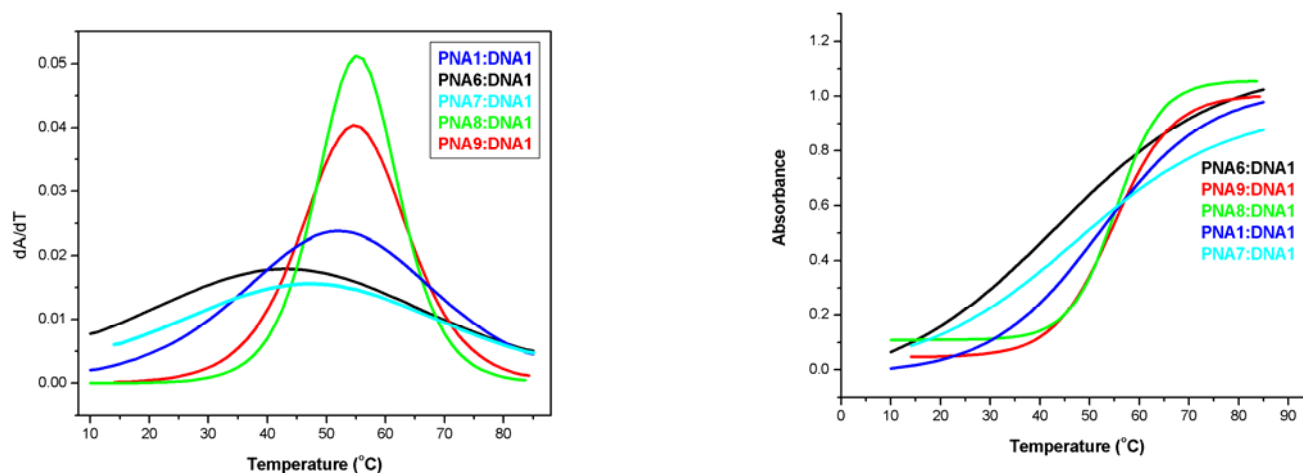
To examine the effect of modified PNA monomers on the corresponding duplexes, hybridization of the mixed purine-pyrimidine PNAs (Table 3, PNA 2-17) incorporating modified thymine monomer units at different positions of *aeg*PNA (PNA 1, Table 3) decamer were studied.



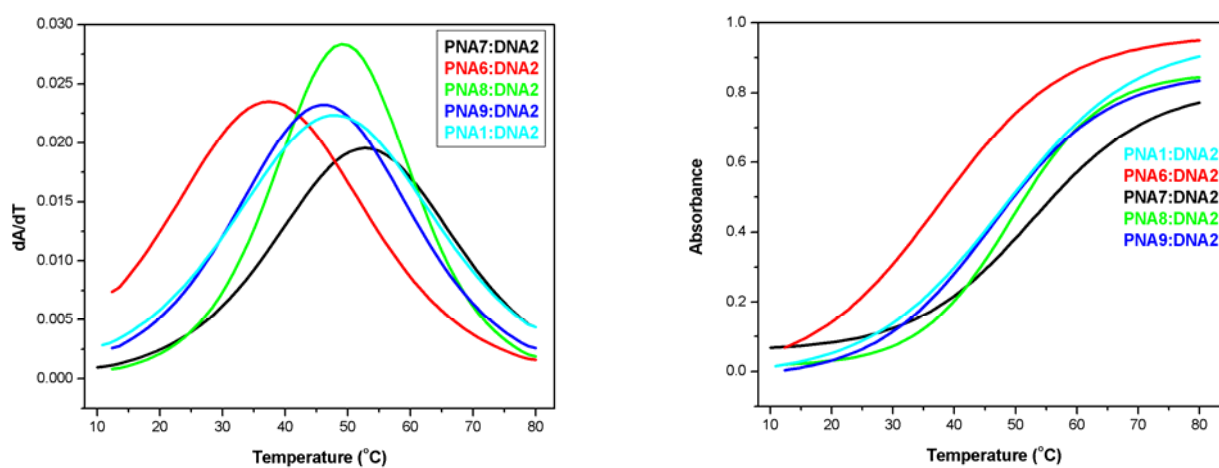
**Figure 6:** Melting curves of  $3 \rightarrow 4dap$  PNAs (PNA1-PNA5) complexed with DNA1 (5' AGTGATCTAC 3') and corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)



**Figure 7:** Melting curves for  $3 \rightarrow 4dap$  PNAs (PNA1-PNA5) complexed with DNA2 (5' CATCTAGTGA 3') and the corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)

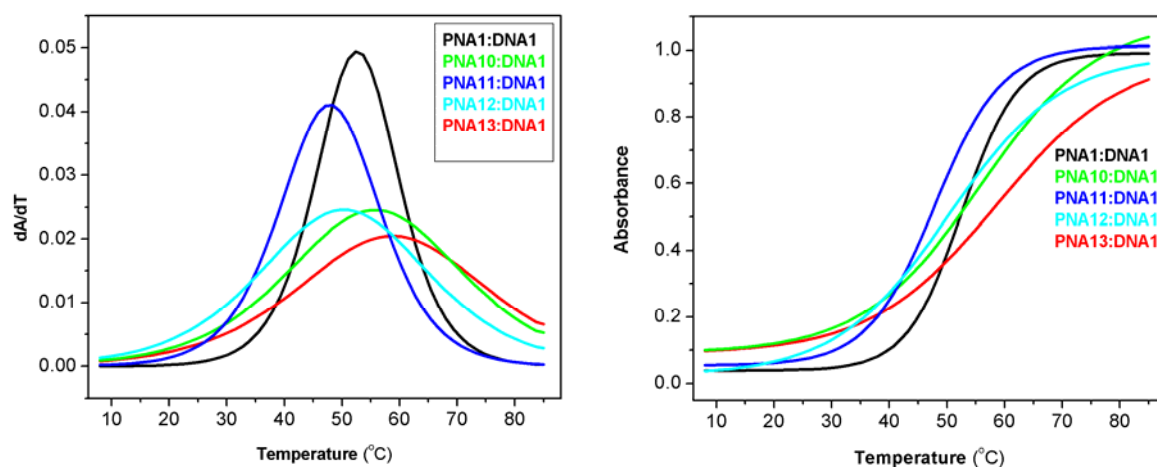


**Figure 8:** Melting curves of  $1\rightarrow 3dap$  PNAs (PNA6-PNA9) complexed with DNA1 (5' AGTGATCTAC 3') and corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)

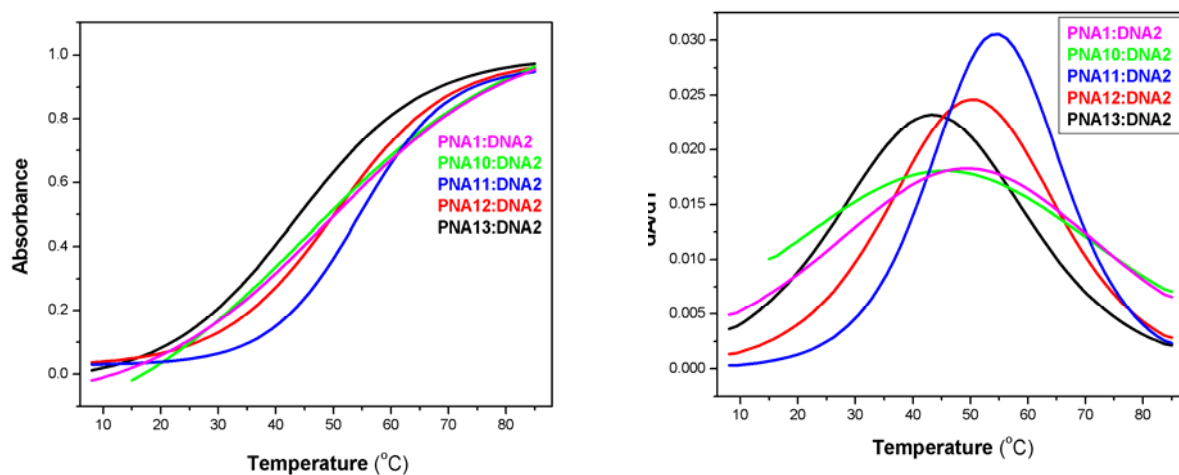


**Figure 9:** Melting curves for  $3\rightarrow 4dap$  PNAs (PNA6-PNA9) complexed with DNA2 (5' CATCTAGTGA 3') and the corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)

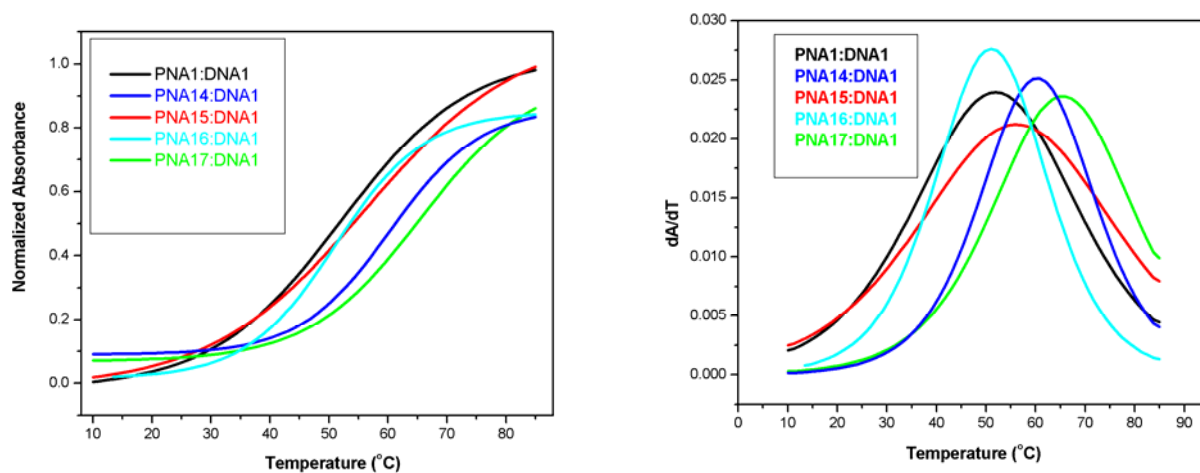




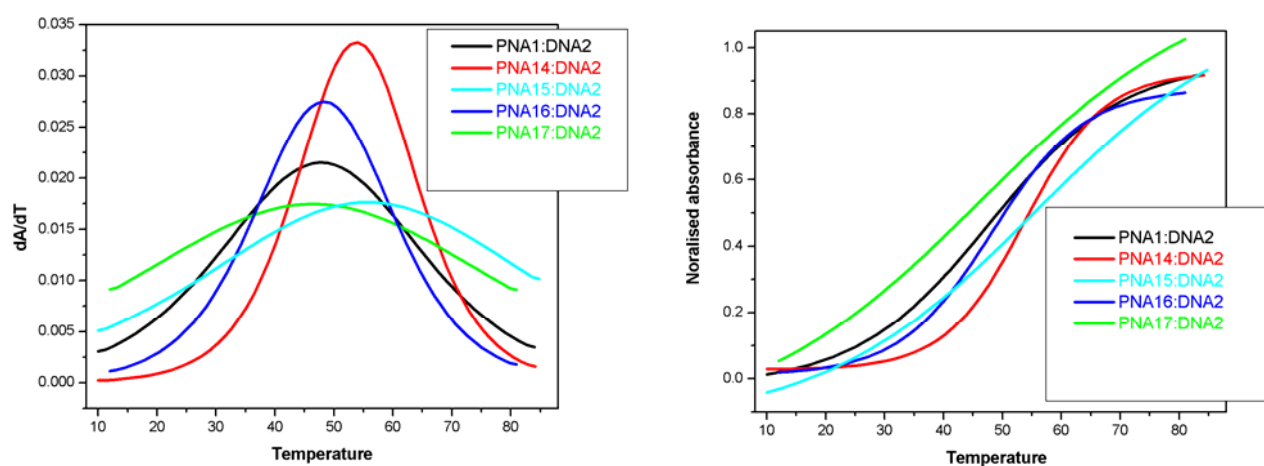
**Figure 10:** Melting curves of  $1 \rightarrow 2amp$  PNAs (PNA10-PNA13) complexed with DNA1 (5' AGTGATCTAC 3') and corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)



**Figure 11:** Melting curves for  $1 \rightarrow 2amp$  PNAs (PNA10-PNA13) complexed with DNA2 (5' CATCTAGTGA 3') and the corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)



**Figure 12:** Melting curves of  $2 \rightarrow 3amp$  PNAs (PNA14-PNA17) complexed with DNA1 (5' AGTGATCTAC 3') and corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)



**Figure 13:** Melting curves for  $2 \rightarrow 3amp$  PNAs (PNA14-PNA17) complexed with DNA2 (5' CATCTAGTGA 3') and the corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl)

The  $T_m$  values of various modified PNAs hybridized with complementary DNAs (DNA1, DNA2) for parallel and antiparallel bindings were determined from temperature dependent UV absorbance and are summarized in Table 6, Figure 6-13.

The unmodified mixed sequence (PNA 1, Table 6) gives a  $T_m$  of 52°C and 48°C for binding with the complementary DNA in antiparallel (with DNA 1) and parallel mode (with DNA 2) respectively.

### 3.3.5 UV thermal studies of 3→4dapPNA(16) with DNA1 and DNA2

The modified 3→4dapPNA with the C-terminus and N terminus modification stabilized antiparallel mode of binding by 4°C and 8°C (entry 2 &5, Table 6), whereas middle and bi modifications are destabilized by -5°C and -3°C respectively.

Binding studies of 3→4dapPNA with parallel DNA showed good binding in all cases (entry 2, 3 &5, Table 6) except in bi modification. The maximum stabilization by 7.5°C(entry 5) is observed in N-terminus modification .

These results conclude that 3→4dapPNA have better binding with both antiparallel and parallel DNA for N-terminus modification when compared to unmodified aegPNA(entry 1, Table 6).

### 3.3.6 UV thermal studies of 1→3dapPNA(17)

1→3dapPNA with antiparallel mode of binding to DNA, C-terminus modification (entry 6, Table 6) showed decrease in  $T_m$  by -9°C and for middle modification (entry 7, Table 6) by -4°C. Whereas for bi and N-terminus modification(entry 8, 9) it showed slight increase in  $T_m$  by 3°C and 2°C respectively.

With parallel DNA 1→3dapPNA showed stabilization by 4.6°C for middle modification and decreases UV melting temperature for N-terminus and C-terminus modifications.

### 3.3.7 UV thermal studies of 1→2ampPNA(27) with DNA1 and DNA2

The 1→2ampPNA with N-terminus modification (entry 13, Table 6) shows stabilization by 8°C and with C-terminus modification (entry 10, Table 6) by 4°C in antiparallel mode when compared to control PNA (entry 1, Table 6). The modified 1→2ampPNA with middle and bi modification destabilized antiparallel mode of binding by -4°C and -5°C respectively.

UV melting studies of *1*→*2amp*PNA with parallel DNA showed increase in temperature by 6.3°C for middle modification and decrease in  $T_m$  for N-terminus modification by -4.8°C.

### 3.3.8 UV thermal studies of *2*→*3amp*PNA(39) with DNA1 and DNA2

The modified *2*→*3amp*PNA with the N-terminus, c-terminus modifications(entry 17 &14, Table 6) stabilized antiparallel mode of binding by 13.5°C and 8.3°C respectively and middle modification showed increase in UV melting temperature by 4°C.

With parallel DNA *2*→*3amp*PNA showed increase in  $T_m$  by 10.3°C for middle modification and by 6°C for c-terminus modification.

These results concluded that the mixed purine-pyrimidine PNA:DNA duplexes derived from different modified PNAs were incorporated at different positions and when these PNAs are incorporated at “C” and “N” terminus antiparallel mode of binding is more stabilised than parallel mode of binding, whereas in case of middle modification parallel mode of binding is more stabilising than antiparallel mode of binding.(Table 6).

## 3.4 Factors affecting the binding of PNA to complementary DNA

### 3.4.1 Effect of backbone modification in *aeg*PNA

The constraints on rotation of single bond were introduced into the *aeg*PNA through formation of rings in *3*→*4dap*PNA, *1*→*3dap*PNA, *2*→*3amp*PNA and *1*→*2amp*PNA units individually, each giving rise to different backbones. The aim of introducing the geometric constraints in the backbone of *aeg*PNA is to pre-organise the flexible backbone of unmodified PNA for entropic advantage in binding with the complementary DNA. Since no conformational/structural details were available, it was thought to systematically synthesise the different possible structures.

### 3.4.2 Effect of backbone extension by one carbon

A comparison of the binding of the *3*→*4dap*PNA, *1*→*3dap*PNA with one carbon extended *2*→*3amp*PNA, *1*→*2amp*PNA oligomers with complementary DNAs show significant difference in the stability with both parallel and antiparallel mode of binding.

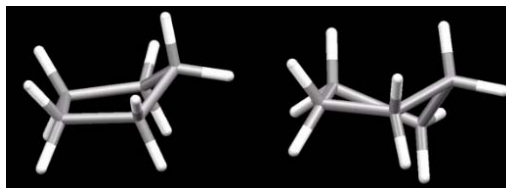
These results indicate that the extension of backbone by one extra carbon atom ( $2\rightarrow 3amp$ PNA) gives the backbone enough flexibility required to adopt a favourable conformation for binding with complementary DNA. The ring through pre-organisation perhaps makes the backbone rigid. Relieving the rigidity through an extra carbon atom forms stabilization of duplex formation.

### 3.4.3 Electrostatic influence in binding

It is reported in the literature that introduction of positive charges in the PNA backbone improves its binding.<sup>21</sup> The cationic PNAs are thus expected to possess superior ability to strand invade the complementary DNA sequences which have negative charges. The positive charge on modified PNA arises due to the the protonation of amino/imino functions in PNA backbone. The PNA sequences incorporating  $2\rightarrow 3amp$ ,  $1\rightarrow 2amp$ ,  $3\rightarrow 4dap$  and  $1\rightarrow 3dap$  units (Table 6) have one free amino group (primary amino group in case  $1\rightarrow 2amp/dap$ PNAs & and the secondary ring amino group in  $3\rightarrow 4dap$ PNA/ $amp$ PNAs) can get protonated making the backbone cationic. The UV-thermal stability studies on these PNA oligomers show better binding with complementary DNA as compared to the *aeg*PNA in which backbone is neutral. PNAs with N-terminus modification showed better stabilisation with antiparallel DNA while parallel DNA is stabilised by middle modification. Sequences involving primary amine in the backbone showed higher melting than the sequences involving secondary amine in the backbone. The results clearly indicate that the electrostatic attraction between positively charged PNA and negatively charged DNA is also a factor in stabilizing the PNA:DNA complexes and is consistent with the literature findings.

### 3.4.4 Effect of ring puckering

The saturated five-membered rings can interconvert their conformations passing through different envelope and twist forms (Figure 14). The barriers for interconversion between the envelopes or half-chairs are extremely slow, the molecules being in a state of "conformational flux".

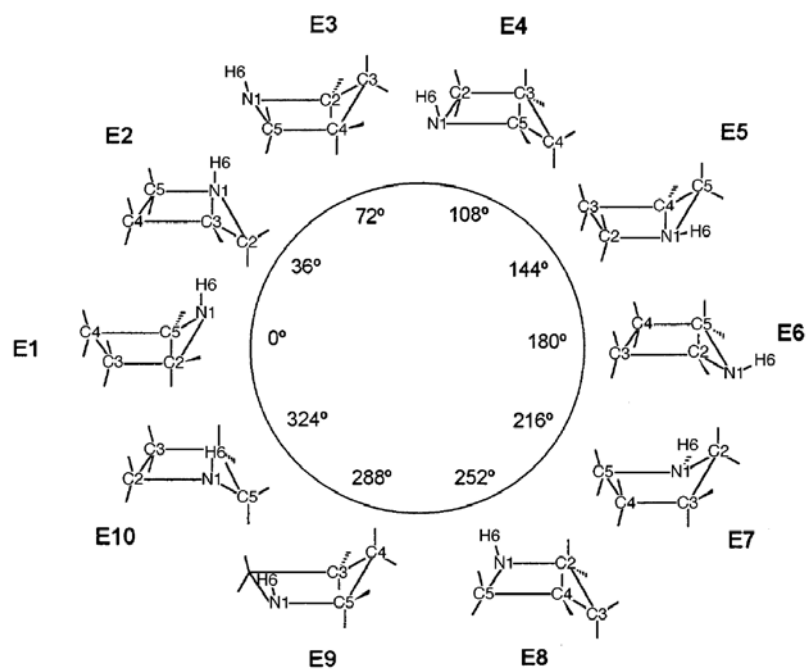


**Figure 14:** Envelope (left) and Twist forms (right) of five membered ring.

The distortion in 5-membered ring arising from out-of-plane atom (atoms) travels around the ring giving rise to the term “ring pseudo rotation”. During such full cycle of pseudo rotation ( $360^\circ$ ), there are a total of 10 different envelopes and 10 different half-chairs (Figure 15). Pseudo rotational barriers for unsubstituted cyclopentane are calculated to be negligible, while for substituted pyrrolidine it is about 1.2 Kcal/mol.<sup>22</sup> This suggests that substituted pyrrolidine ring is rigid and gets locked in stable conformation as compared to cyclopentyl ring. The substituents in the ring system increase this energy barrier and shift the conformational equilibrium of the ring into one or a few of half-chairs or envelopes.

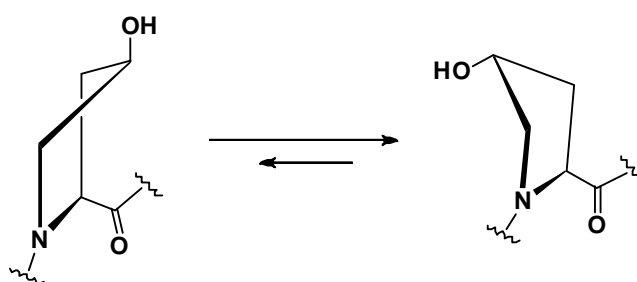
Introduction of a single substitution of (*R,R*) *trans*-cyclopentyl *tcypT*, in the middle of a purine-pyrimidine PNA decamer which was reported earlier by Appella *et al.*,<sup>23a</sup> showed no observable melting transition. This was attributed to the stereochemistry of the the cyclopentyl ring which may not favourably pre-organise the backbone for binding to complementary DNA while its (*S,S*) *trans* analogue showed 6.0°C increase in  $T_m$  than control *aegPNA*. In contrast, both (*R,S*) and (*S,R*) *cis*-cyclopentyl PNAs<sup>23b</sup> bind very strongly to *cDNA*/*RNA* with an increase of  $T_m$  around + 7.5°C per modification.

In the present work, it is seen that by introducing a N-atom in the cyclopentyl ring (*R,S*) 3→4*dapPNA* oligomer shows a melting temperature of 60°C *i.e.*  $\Delta T_m$  is 6°C (PNA 5, Table 6) for the same PNA sequence. This suggests that not only the stereochemistry but the type of the ring pucker also has markable effect on the binding.



**Figure 15:** Envelope conformations of pyrrolidine along the pseudorotational path. The twist conformations are located between adjacent envelope forms.<sup>23</sup>

It is well documented in the literature that the conformations of the five-membered pyrrolidine rings in proline and 4-hydroxyproline are important in controlling the structure and hence the physiological functions of collagen fibrils. The electronegativity of the substituents attached to the ring influences the complex by changing the puckering of the ring (Figure 16). For instance, an electronegative substituent at the 4*R* position of proline has a strong preference to pucker the ring from  $C\gamma$ -endo pucker in proline to  $C\gamma$ -exo pucker.



**Figure 16:** Change in puckering of proline ring by substitution

In *aep*PNA reported earlier from this laboratory, the substituted purines or pyrimidines attached directly when present at the 4-position of the pyrrolidine ring differ in their group electronegativities, causing differential ring pucker effects.<sup>24</sup> This leads to change in backbone conformation and hence causing sequence specific effects.

In *3→4dap*PNA, *2→3amp*PNA proline ring is part of the backbone through a common C-C bond while in *1→3dap*PNA, *1→2amp*PNA the proline ring is main part of the backbone. Thus the effect of pyrrolidine ring puckering on the backbone will be different in these two cases, which may finally influence the stability of derived PNA:DNA complexes.

The better binding in *2→3amp*PNA as compared to *1→2amp*PNA and *3→4dap*PNA over *1→3dap* PNA may be attributed to the puckering of the ring which is influenced by presence of a nitrogen atom in the ring. Thus the necessary torsional adjustment in the ethylene diamine portion of modified PNA is required for good binding. The involvement of the pyrrolidine ring as the side chain to PNA backbone in *3→4dap*PNA and *2→3amp*PNA is favoured over the intervening of the ring in the backbone of *1→3dap*PNA and *1→2amp*PNA for better stability of PNA:DNA complexes.

#### 3.4.5 Effect of position of modification

The position of the modified units derived from *3→4dap* and *2→3amp*PNA monomers present in the PNA sequence also has markable effect on the binding of the PNA:DNA complexes. PNA:DNA duplexes derived from different modified PNAs were incorporated at different positions and when these PNAs are incorporated at “C” and “N” terminus antiparallel mode of binding is more stabilised than parallel mode of binding, whereas in case of middle modification, the parallel mode of binding was found to be more stable than the antiparallel mode of binding.

#### 3.4.6 Effect of primary or secondary amino group

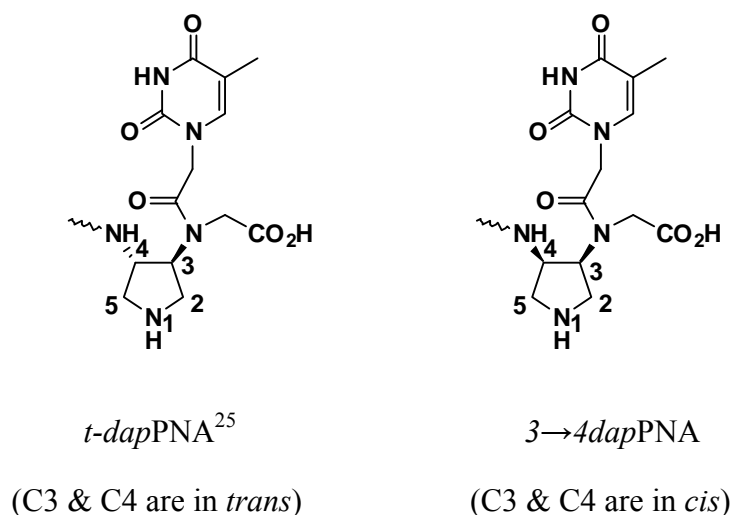
The modified PNA monomers **16**, **17**, **27** and **39** designed to synthesize two different backbones through both primary and secondary amine functions (Figure 4). The UV-melting results of these sequences involving primary amine in the backbone (entry 2-5 & 14-17) shows higher melting than the sequences involving secondary



amine (entry 6-13, table 6) in the backbone. This may be because of more restraints in the backbone in the form of tertiary amide linkage leading to destabilising in case of backbone with secondary amine group as compared to that with primary amine group counterpart.

### 3.5 Comparison of UV melting temperatures ( $0^{\circ}\text{C}$ ) of *cis* and *trans* *dap* (diamino pyrrolidine) PNAs.

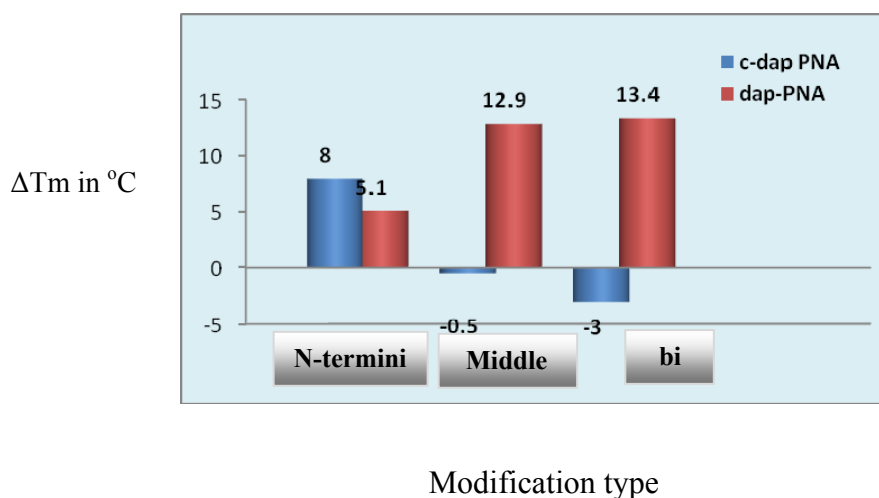
To understand the effect of stereochemistry on binding affinity towards complementary DNA, comparative studies of *t-dap*PNA and *3→4dap*PNA (shown below) have been done with both antiparallel and parallel DNA. For *t-dap*PNA (Figure 17) stereochemistry at C3 and C4 carbons are *trans* whereas for *3→4dap*PNA (Figure 17) stereochemistry at C3 and C4 are *cis*.



**Figure 17:** structures of *t-dap* and *c-dap*PNA

The  $T_m$  values observed for the *t-dap*PNA and *3→4dap*PNA with complementary antiparallel DNA with different positions of modifications are shown in Figure 18 for comparison.

The N-terminus modification by *t-dap*PNA and *3→4dap*PNA showed stabilization towards complementary antiparallel DNA by 5.1 & 8°C respectively (Figure 18) compared to unmodified *aeg* PNA.



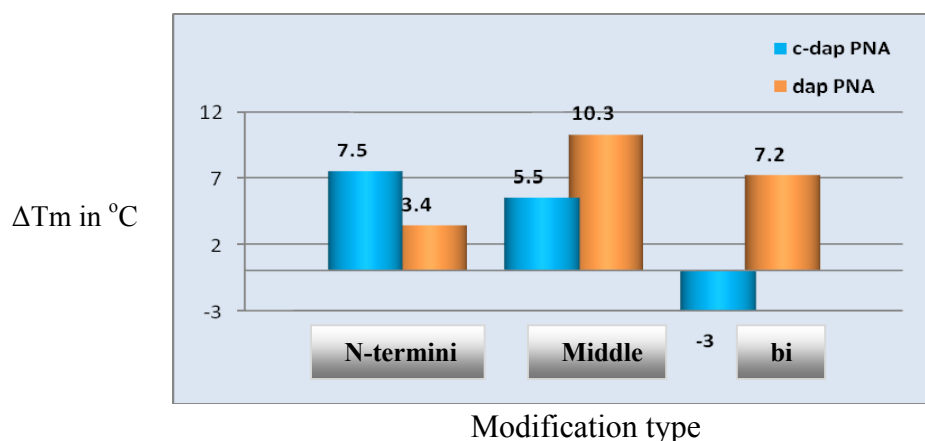
**Figure 18:** Comparative studies (in  $\Delta T_m$  in  $^{\circ}\text{C}$ ) of *t-dap*PNA and *3*→*4dap*PNA with respect to modification with anti parallel DNA

The middle modification by *t-dap*PNA showed stabilization towards DNA by  $12.9^{\circ}\text{C}$  but middle modification by *3*→*4dap*PNA showed no change ( $0.5^{\circ}\text{C}$ ) in UV  $T_m$  was observed as compared to *aeg*PNA.

The bi modification of the oligomer by *t-dap*PNA showed stabilization towards DNA binding by  $13.4^{\circ}\text{C}$  but when the modification was by *3*→*4dap*PNA the destabilization was by  $-3^{\circ}\text{C}$ . These results clearly show the effect of stereochemistry on binding affinity towards DNA and *cis* stereochemistry showed stabilization for N-terminus modification while middle and bi modifications in *trans* stereo chemistry exhibited better stabilization.

The  $T_m$  values observed for the *t-dap*PNA and *3*→*4dap*PNA with complementary parallel DNA with different positions of modifications are shown in Figure 19 for comparison.

The N-terminus modification by *t-dap*PNA and *3*→*4dap*PNA showed stabilization towards complementary antiparallel DNA by  $3.4$  &  $7.5^{\circ}\text{C}$  respectively (Figure 19) compared to unmodified *aeg* PNA. The middle modification by *t-dap*PNA showed stabilization towards DNA by  $10.3^{\circ}\text{C}$  but when middle modification was done by *3*→*4dap*PNA  $5.5^{\circ}\text{C}$  change in UV  $T_m$  was observed when compared to *aeg*PNA.



**Figure 19:** Comparative studies (in  $\Delta T_m$  in  $^{\circ}\text{C}$ ) of *t-dap*PNA and *3→4dap*PNA with respect to modification with parallel DNA

The bi modification of the oligomer by *t-dap*PNA showed stabilization towards DNA binding by  $7.2^{\circ}\text{C}$  but when the same modification was done by *3→4dap*PNA it showed the destabilization by  $-3^{\circ}\text{C}$ . These results indicate the stereochemical effects on derived PNA binding towards parallel DNA, with *cis* stereochemistry showing stabilization for N-terminus modification whereas for middle and bi modifications *trans* stereochemistry showed better stabilization as compared to its *cis* stereo isomer.

Overall, when the stereochemistry of *dap* (diamino pyrrolidine) derived PNA is *cis* at C2 & C3 carbon N-terminus modification showed more stabilization whereas *trans* isomer showed better stabilization for middle and bimoifications.

### 3.6 Summary of results of thermal stability by modified PNAs (2-17, Table 6) with complementary DNA

Summary of modified PNAs thermal stability with complementary DNAs is as follows

- i) extension of PNA backbone by one extra carbon atom can give enough flexibility to the backbone to adopt a favourable conformation for binding with complementary DNA.
- ii) The position of the modified PNA monomers present in the PNA sequence also has markable effect on the binding of the PNA:DNA complexes. PNA:DNA duplexes with modified PNAs were incorporated at different positions and when these PNAs are incorporated at “C” and “N” terminus, antiparallel mode of binding is more stabilised

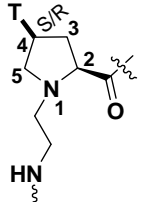
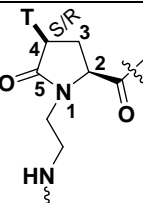
than parallel mode of binding, whereas in case of middle modification, the parallel mode of binding was found to be more stable than the antiparallel mode of binding.

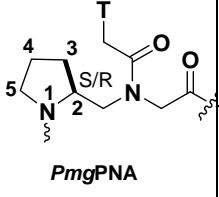
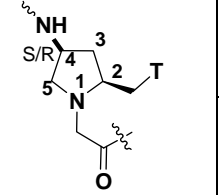
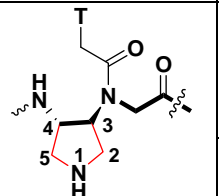
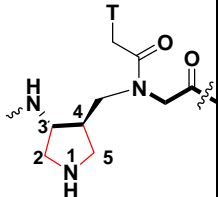
iii) PNA sequences involving primary amine in the backbone shows higher melting than the sequences involving secondary amine in the backbone.

iv) The positive charge on modified PNAs arises due to the the protonation of amino/imino functions in PNA backbone, and the UV-thermal stability studies of these PNA oligomers show better binding with complementary DNA as compared to the *aeg*PNA in which backbone is neutral. PNAs with N-terminus modification showed better stabilisation with antiparallel DNA while parallel DNA is stabilised by middle modification. Sequences involving primary amine in the backbone showed higher melting than the sequences involving secondary amine in the backbone The results clearly indicate that the electrostatic attraction between positively charged PNA and negatively charged DNA is also a factor in stabilizing the PNA:DNA complexes.

v) Incase of all these modified PNAs proline ring is part of the backbone and the effect of pyrrolidine ring puckering on the backbone will be different in each case, which may finally influence the stability of derived PNA:DNA complexes.

**Table 7:** Pyrrolidine based PNA monomers ; effect on thermal stability to complementary DNA per modification compared to *aeg*PNA

PNA modification	Stereochemistry	$\Delta T_m/^\circ\text{C}$ Per modification	Reference
 <i>aeg</i> PNA	2 <i>S</i> , 4 <i>S</i>	12.8 <sup>0</sup> C	Ganesh, K. N. <i>et al</i> <i>Org. Lett.</i> <b>2001</b> 3, 1281–1284
	2 <i>R</i> , 4 <i>S</i>	-10.0 <sup>0</sup> C	
 <i>aepone</i> PNA	2 <i>S</i> , 4 <i>R</i>	14.5 <sup>0</sup> C	Nagendra kumar Sharma thesis submitted to University Of Pune October 2005
	2 <i>S</i> , 4 <i>S</i>	12.5 <sup>0</sup> C	

 <p><b>PmgPNA</b></p>	2 <i>R</i>	-11.0 <sup>0</sup> C	Yeheškiely, E.; Slaitas, A. <i>Eur. J. Org. Chem.</i> <b>2002</b> , 2391
	2 <i>S</i>	-14.0 <sup>0</sup> C	
 <p><b>PyrrolidinePNA</b></p>	(2 <i>S</i> ,4 <i>S</i> )	-13.0 <sup>0</sup> C	Lonkar, P.; Ganesh, K. N.; Kumar, V. A. <i>Org. Biomol. Chem.</i> <b>2004</b> , 2, 2604-2611.
	(2 <i>S</i> ,4 <i>R</i> )	14.0 <sup>0</sup> C	
 <p><b>dapPNA</b></p>	(3 <i>R</i> ,4 <i>S</i> ) (N-terminus modification)	5.1 <sup>0</sup> C	A. K. Sharma; Ph.D thesis (August, 2010), university of pune.
	(3 <i>R</i> ,4 <i>S</i> ) (Middle modification)	12.9 <sup>0</sup> C	
 <p><b>ampPNA</b></p>	(3 <i>R</i> ,4 <i>S</i> ) (N-terminus modification)	6.4 <sup>0</sup> C	A. K. Sharma; Ph.D thesis (August, 2010), university of pune

### 3.7 Conclusion

Four different monomers **16**, **17**, **27**, **39** synthesized in Chapter-2 were used to incorporate site specifically to yield the required oligomers (PNAs 2-17). The PNA oligomers containing one or two modified thymine PNA units were subjected to UV melting studies with complementary DNA. The duplexes formed by 3→4*dap*PNA and 2→3*amp*PNA were better stabilized than their analogues 1→3*dap*PNA and 1→2*amp*PNA. This differentiation in stability of PNA:DNA complexes of different PNA derived units, can help to fine tune the binding properties of PNA:DNA complexes for different applications. Furthermore, there is currently no standard backbone modification which can tune the stabilities of PNA:DNA complexes. The modified PNA monomers presented in this and the preceding chapter constitute a new family of PNA monomers which can tune the binding properties of *aeg*PNAs. This can be achieved has either incorporating the monomer units through pyrrolidine ring

substitutions in the backbone as in  $3 \rightarrow 4dap$  or  $2 \rightarrow 3ampPNA$  or through ring intervening as in  $1 \rightarrow 3dap$  and  $1 \rightarrow 2ampPNA$ . The former has stabilizing effect while the later has destabilizing effect on PNA:DNA complexes, to help adjust the binding properties. Further work is required to tune the stability of PNA:DNA complexes by incorporating both the modified units, in the same PNA sequence .

### 3.8 Experimental

#### 3.8.1 Solid phase peptide synthesis

##### 3.8.1a Functionalization of resin

Commercially available MBHA resin or Rink amide resin (Jupiter Biosciences, West Marredpally, Hyderabad, India) with a loading value of 2.0 m eq/g and 0.7 m eq/g respectively, is not suitable for oligomer synthesis and hence the loading value was minimized to 0.25-0.35 m eq/g to avoid the aggregation of the growing oligomer. The dry resin was taken in solid phase funnel and swelled in DCM for 1 h. The solvent was drained off and the resin was treated with calculated amount of acetic anhydride in 5% DIPEA/DCM solution for about 15 min, solvent was drained off, the resin was thoroughly washed with DCM and DMF to remove the traces of acetic anhydride and dried under vacuum. The dry resin was taken in solid phase funnel and swelled in DCM for about 1h and functionalized with  $N^{\square\square}boc-N^{\square\square}Z$ -Lysine.

##### 3.8.1b Picric acid assay for the estimation of the amino acid loading

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 functionalized dry resin (5 mg) was taken in a sintered funnel and swelled in DCM for 1 h. The solvent was drained off and the resin was treated with 50% TFA/DCM for *boc* functionalized MBHA resin and with 20% Piperidine/DMF for *fmoc* functionalized rink amide resin for 15 min (1 mL x 2) each time. The resin was thoroughly washed with DMF/DCM and the TFA salt in case of MBHA resin was neutralized with 5% diisopropyl ethylamine for 2 min (1 mL x 3). The free amine was treated with 0.1 M picric acid in DCM for 10 min (2 mL x 3) each time. The resin was thoroughly washed with DCM to remove the unbound picric acid. The picrate bound to amino groups was eluted with 5% diisopropyl ethylamine in DCM, followed by

washing with DCM. The elutant was collected into a volumetric flask (10 ml) and made up to 10 ml using DCM. An aliquot (0.2 ml) of picrate eluant was diluted to 2 ml with ethanol and the optical density was measured at 358 nm (picric acid  $\lambda_{\max}$ ), and the loading value of the resin (0.35 meq/g) was calculated using the molar extinction coefficient of picric acid as  $\epsilon_{358}=14,500 \text{ cm}^{-1} \text{ M}^{-1}$  at 358 nm.

### 3.8.1c Kaiser's Test

Kaiser's test was used to monitor the Boc-deprotection and coupling steps in the solid phase peptide synthesis. Three solutions used were (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 pyridine (98 mL)). To a few beads of the resin taken in a test tube, 3-4 drops of each of the three solutions were added. The tube was heated for 5 min, and appearance of blue colour of the beads in the solution indicated successful deprotection, while colourless beads were observed upon completion of the amide coupling reaction.

### 3.8.1d Cleavage of the PNA oligomers from the solid support

**For MBHA resin:** The resin bound PNA oligomer (5 mg) was kept in an ice-bath with thioanisole (10  $\mu\text{L}$ ) and 1,2-ethanedithiol (4  $\mu\text{L}$ ) for 10min, TFA (60  $\mu\text{L}$ ) was added and shaken manually and kept for another 10 min. TFMSA (8  $\mu\text{L}$ ) was added and stirring continued for 1.5 hrs. The reaction mixture was filtered through a sintered funnel. The residue was washed with TFA (3 x 2 mL) and the combined filtrate were evaporated under vacuum and co-evaporated with diethyl ether, avoiding heating during this process. The residue was precipitated using dry diethyl ether and centrifuged. The diethyl ether layer was decanted and the solid part was dissolved in water.

**For Rink amide resin:** The resin bound PNA oligomers (5 mg) was treated with 50% TFA in DCM and kept for 15 minutes and this was repeated three times. The combined filtrate were evaporated under vacuum and co-evaporated with diethyl ether and processed same way as the MBHA resin described above.

### 3.8.2 Purification of PNA oligomers

The purity of the synthesized PNA oligomers was checked by analytical reverse phase HPLC on C18 column using a gradient of 0 to 100% CH<sub>3</sub>CN in water containing 0.1% TFA at a flow rate of 1.5 mL/min. These were subsequently purified by RP-HPLC on a semi preparative C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC.

### 3.8.3 MALDI-TOF Mass Spectrometry

Several matrices have been explored to characterise the PNAs by MALDI-TOF mass spectrometry these include sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid),<sup>23</sup> CHCA ( $\alpha$ -cyano-4-hydroxycinnamic acid)<sup>24</sup> and DHB (2,5-dihydroxybenzoic acid). Among these, CHCA was found to give the best signal to noise ratio and was used for the analysis of the modified PNAs reported here and it was found to give satisfactory results.

### 3.8.4 UV- $T_m$ studies

#### 3.8.4a Sample preparation

The PNA oligomers and the appropriate DNA oligomers were mixed together in stoichiometric amounts (1:1 for the duplex forming PNAs, *viz.*, the mixed base sequences) in sodium phosphate buffer (10mM), NaCl (10mM) at pH 7.4 to achieve a final strand concentration of 0.5 and 1mM. The extinction coefficient  $C = 6.6$ ,  $T = 8.6$ ,  $A = 13.7$  and  $G = 11.7$  [(mmol)<sup>-1</sup> cm<sup>-1</sup>] were used to calculate the concentration of PNA by following Lambert Beer's Law:  $A = \epsilon cl$ .

The antiparallel and parallel complexes were constituted from appropriate mixed sequence PNAs and corresponding complementary DNA. The samples were heated at 85 °C for 5 min followed by slow cooling to room temperature. They were allowed to remain at room temperature for at least half an hour and refrigerated overnight prior to running the melting experiments.

#### 3.8.4b UV melting

All UV melting experiments were performed on Cary 300 Bio UV-Visible Spectrophotometer equipped with a thermal melt system. The samples for  $T_m$  measurement were prepared by mixing calculated amount of appropriate



oligonucleotide and PNA solutions together in 0.01 M sodium phosphate buffer (1ml, pH 7.4). The samples (1 mL) were transferred to quartz cell, sealed with Teflon stopper after degassing with nitrogen gas for 15 min, and equilibrated at the starting temperature for at least 30 min. The OD at 260 nm was recorded in steps from 10-85°C with temperature increment of 0.5°C/min. Nitrogen gas was purged through the cuvette chamber below 20 °C to prevent the condensation of moisture on the cuvette walls. Each melting experiment was repeated at least thrice. The results were normalized and analysis of data was performed using Microcal Origin 6.0 (Microsoft Corp.). The absorbance or the percent hyperchromicity at 260 nm was plotted as a function of the temperature. The  $T_m$  was determined from the peaks in the first derivative plots and is accurate to  $\pm 1^\circ\text{C}$ .

### 3.9 References

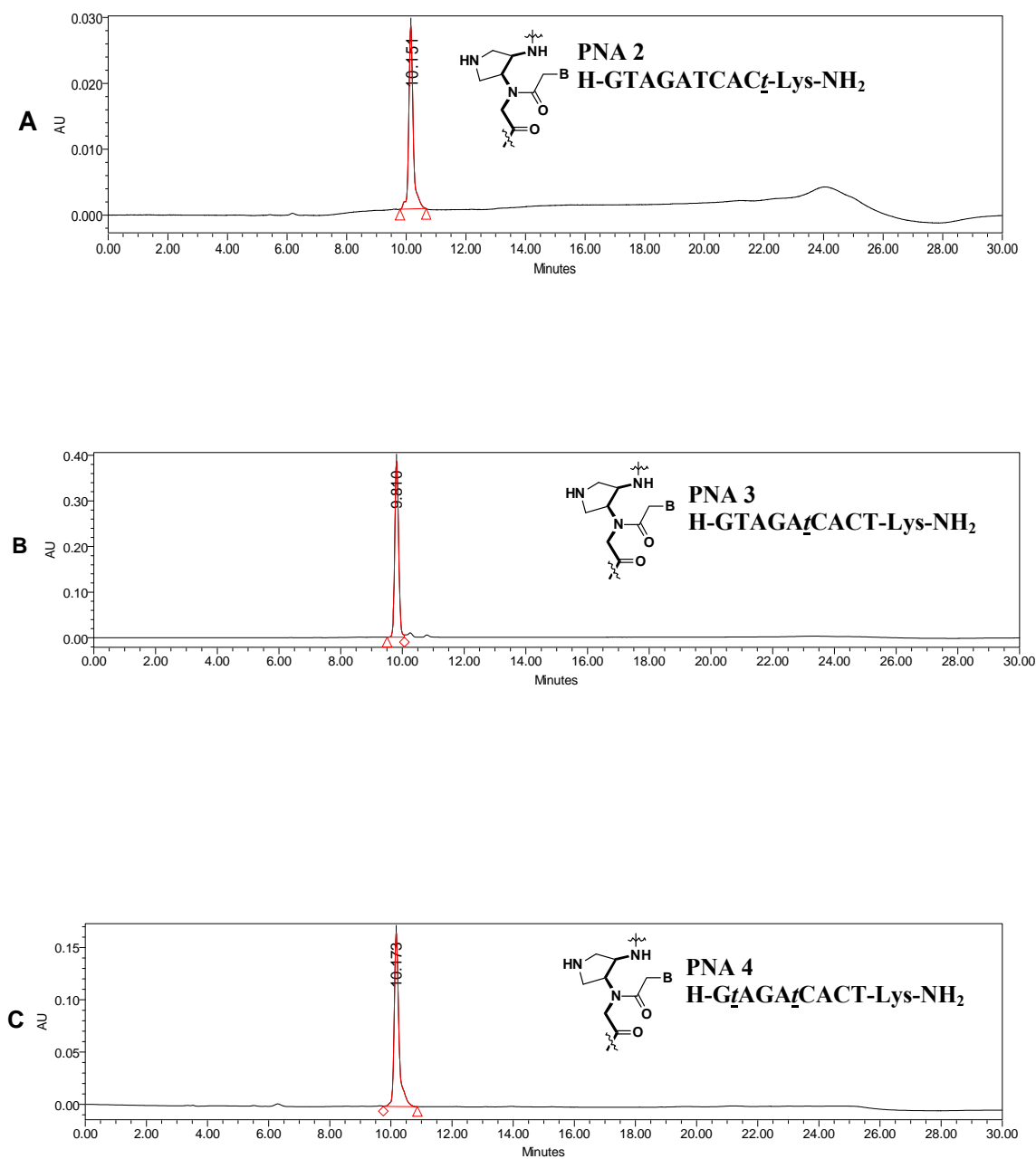
1. Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149.
2. (a) Mckay, F. C.; Albertson, N. F. *J. Am. Chem. Soc.* **1957**, *79*, 4686. (b) Anderson, G. W.; McGreoger, A. C. *J. Am. Chem. Soc.* **1957**, *79*, 6180.
3. Carpino, L.A.; Han, G.Y. *J. Org. Chem.* **1972**, *37*, 3404.
4. Matsueda, G. R.; Stewart, J. M. *Peptides* **1981**, *2*, 45.
5. Rink, H. *Tet. Lett.* **1987**, *28*, 3787.
6. Bernhardt, A.; Drewello, M.; Schutkowski, M. *J. Peptide Res.* **1997**, *50*,143.
7. Sheehan, J. C.; Hess, G. P. *J. Am. Chem. Soc.* **1955**, *77*, 1067.
8. (a) Carpino, L. A. *J. Am. Chem. Soc.* **1993**, *115*, 4397. (b) Albericio, F.; Bofill, J. M.; Faham, A. E.; Kates, S. A. *J. Org. Chem.* **1998**, *63*, 9678.
9. Marder O.; Shvo Y.; Albericio F. *Chim. Oggi* **2002**, *20*, 37.
10. Coste, J.; Dufour, M. N.; Pantaloni, A.; Castro B. *Tet. Lett* **1990**, *31*, 669.
11. Hudson, D. *J. Org. Chem* **1988**, *53*, 617.
12. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.
13. Oded, A.; Houghten, R. A. (1990) *Pept. Res* **1990**, *3*, 42.
14. Fronenot, J. D. *et. al. Pept. Res.* **1991**, *4*, 19.
15. Vojkovsky, T. *Pep. Res.* **1995**, *3*, 42.
16. Blackburn, C. *Tet Lett.* **2005**, *46*, 1405.

17. a.) Merrifield, R. B.; Stewart, J. M.; Jernberg, N. *Anal. Chem.* **1966**, *38*, 1905.  
b.) Kaiser, E.; Bossinger, C. D.; Cplscott, R. L.; Olsen, D. B. *Anal. Chim. Acta.* **1980**, *118*, 149. c) Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem.* **1981**, *117*, 147.
18. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Peptide Sci.* **1995**, *3*, 175.
19. Soyfer, V. N.; Potaman, V. N. *Triple helical Nucleic Acids*; Springer: New York, 1996.
20. Ganesh, K. N.; Kumar, V. A.; Barawkar, K. A. *Supramolecular control of structure and reactivity* (Chapter 6, Ed Hamilton, A. D) **1996**, 263.
21. a.) Tan, T. H.; Hickman, D. T.; Morral, J.; Beadham, I. G.; Micklefield J. *Chem. Commun.* **2004**, *5*, 516. b.) Hickman, D.T.; Tan, T. H.; Morral, J.; King, P. M.; Cooper, M. A.; Micklefield, J. *Org. Biomol. Chem.* **2003**, *1*, 3277. c.) Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1083.
22. a.) Pitzer, K. S.; Donath, W. E. *J. Am. Chem. Soc.* **1959**, *81*, 3213. b.) Carballeira, L.; Perez-Juste, I. *J. chem. Soc., Perkin Trans 2*, **1998**, 1339.
23. a.) Pokorski, J. K.; Witschi, M. A.; Purnell, B.L.; Appella, D. H. *J. Am. Chem. Soc.* **2004**, *126*, 15067. b.) Govindraju, T; Kumar, V. A.; Ganesh, K. N. *J. Am. Chem. Soc.* **2005**, *127*, 4144
24. Nagendra K.Sharma; Krishna N. Ganesh, *Tetrahedron*, Volume 66, Issue 47, **2010**, 9165-9170.
25. A. K. Sharma; Ph.D thesis (August, 2010), university of pune.
26. Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Hames, B. D., Ed.; Oxford University Press: New York, 2000.
27. Kent, S. B. H. *Annu. Rev. Biochem.*; **1988**, *57*, 957–989
28. Hong-Jun Cho, Tae-Kyung Lee, Jung Won Kim, Sang-Myung Lee, and Yoon Sik Lee
29. Carolina Torres-García, Daniel Pulido, Magdalena Carceller, Iván Ramos, Miriam Royo and Ernesto Nicolás.

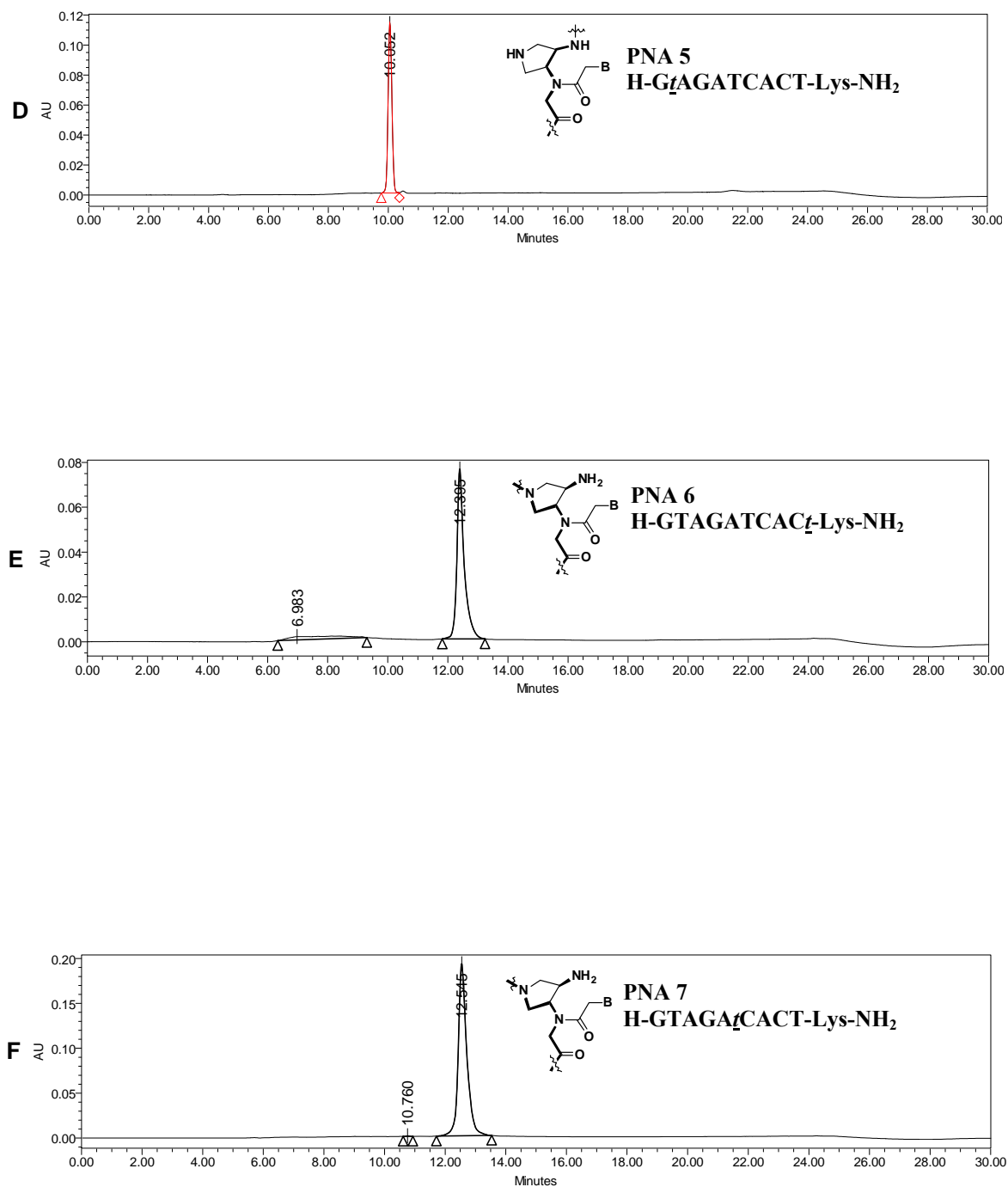
**3.10 : APPENDIX**

<b>Entry</b>	<b>Page No</b>
1) HPLC data for PNA sequences <b>1-17</b>	150
2) MALDI-TOF Spectra of PNAs <b>1-17</b>	155

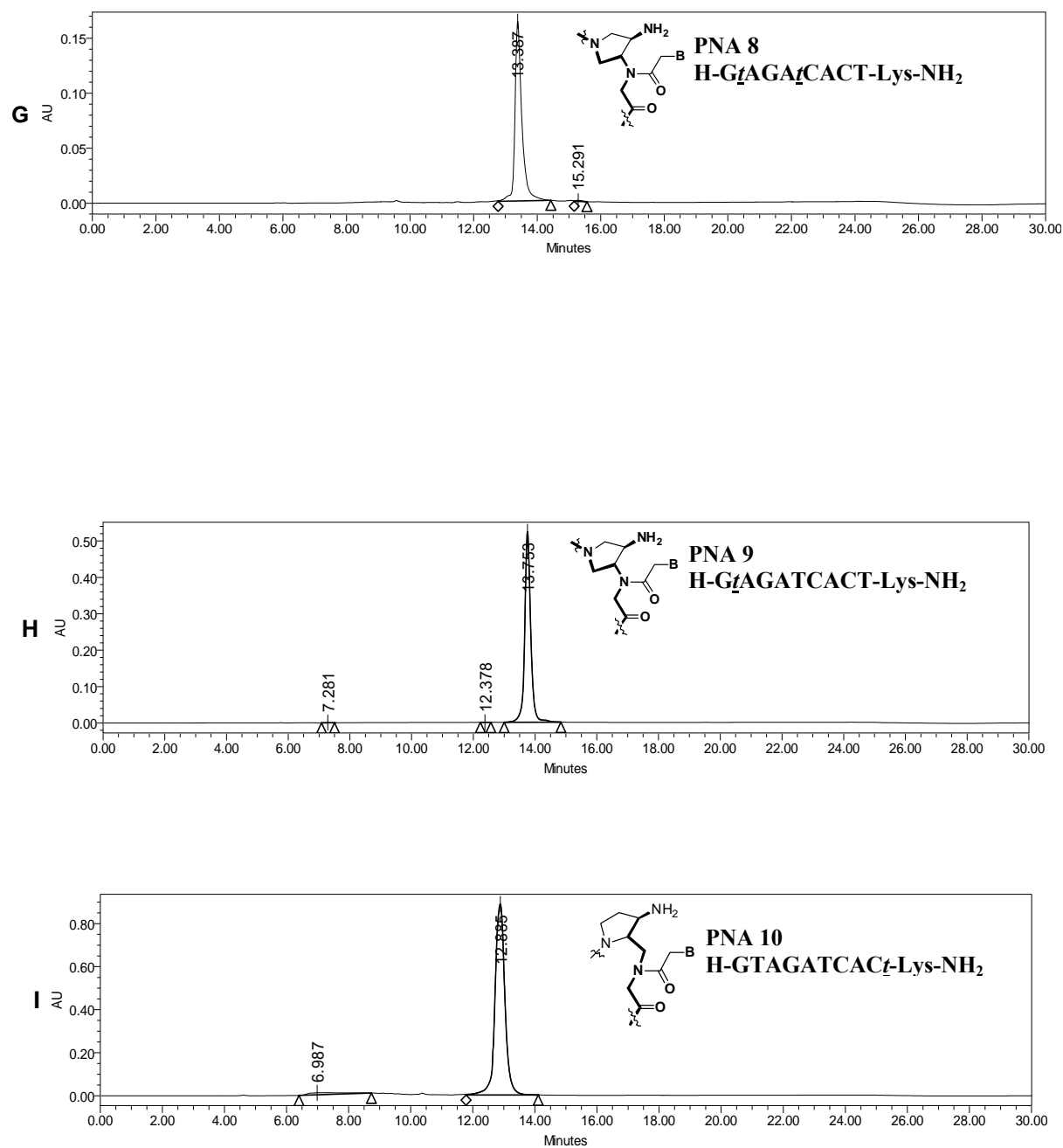
## HPLC data for PNA sequences



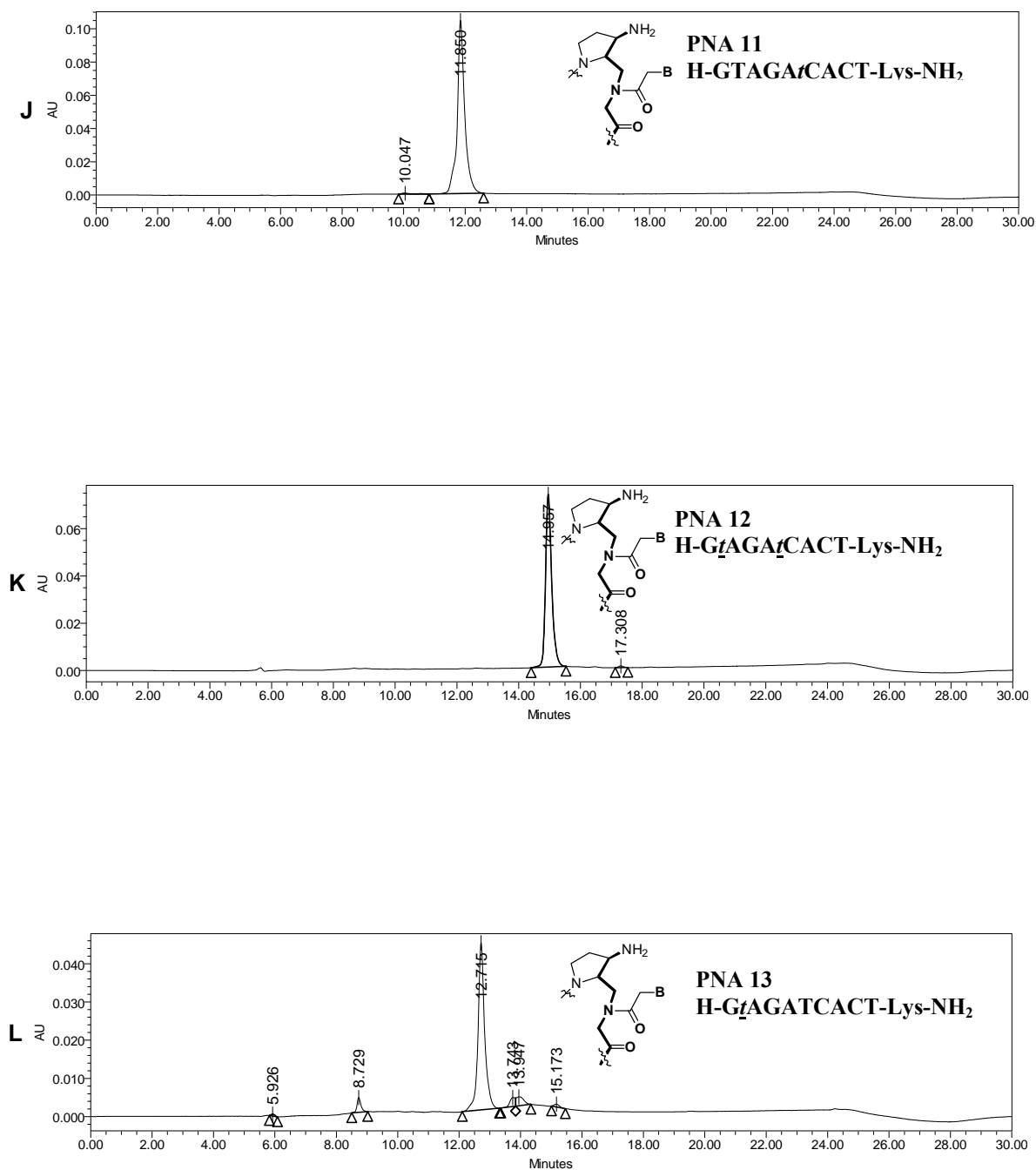
**Figure 20:** (A) Reverse phase HPLC profile of PNA 2 (B) PNA Reverse phase HPLC profile of PNA 3 (C) Reverse phase HPLC profile of PNA 4



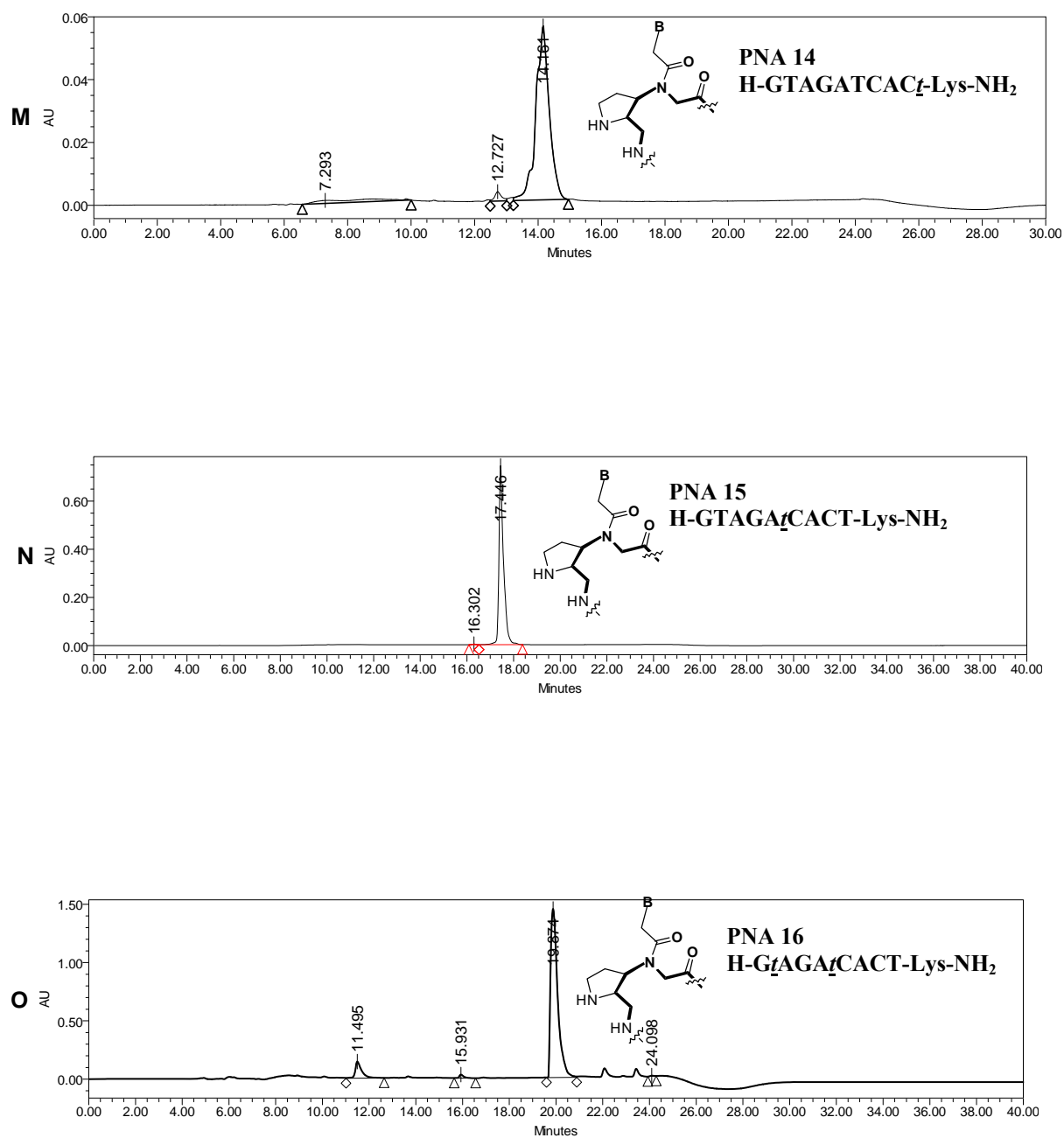
**Figure 21:** (D) Reverse phase HPLC profile of PNA 5 (E) PNA Reverse phase HPLC profile of PNA 6 (F) Reverse phase HPLC profile of PNA 7



**Figure 22:** (G) Reverse phase HPLC profile of PNA 8 (H) PNA Reverse phase HPLC profile of PNA 9 (I) Reverse phase HPLC profile of PNA 10

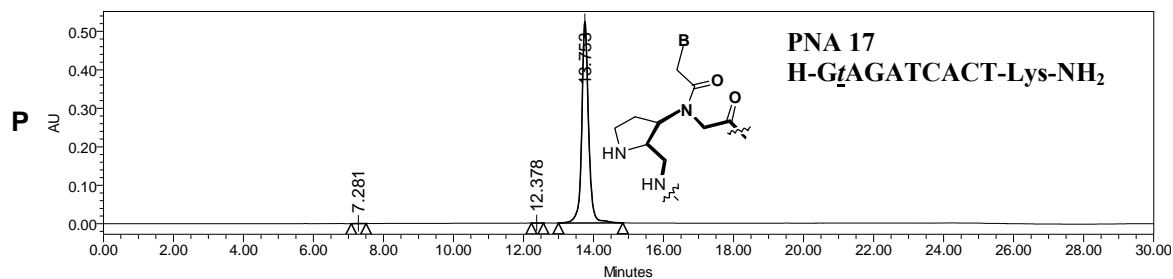


**Figure 23:** (J) Reverse phase HPLC profile of PNA 11 (K) PNA Reverse phase HPLC profile of PNA 12 (L) Reverse phase HPLC profile of PNA 13

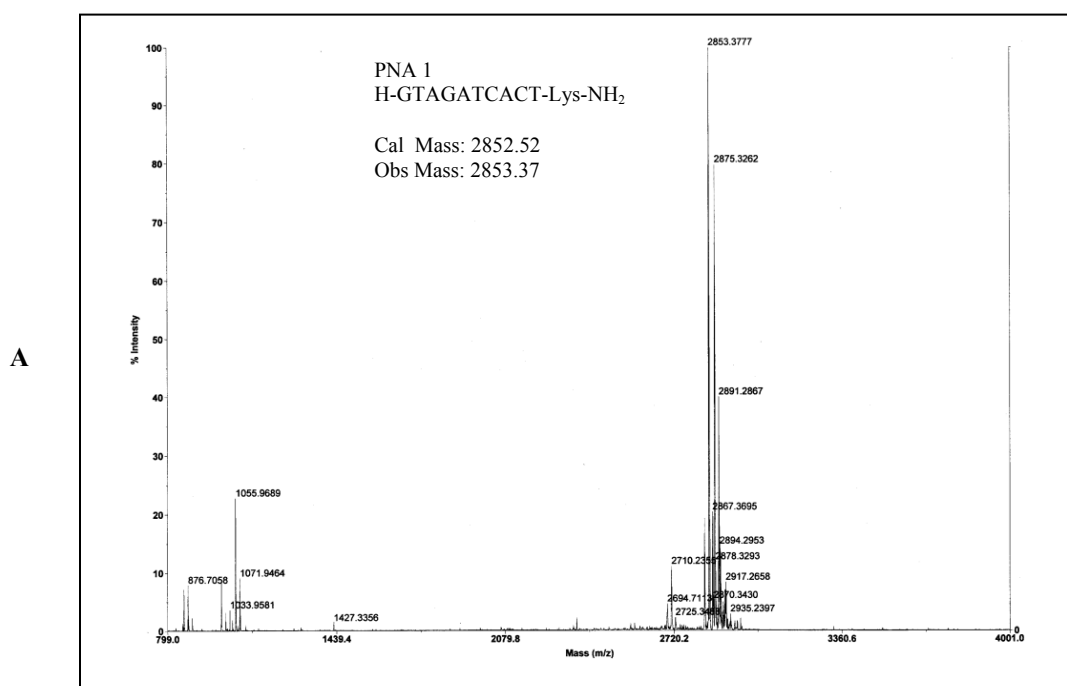


**Figure 24:** (M) Reverse phase HPLC profile of PNA 14 (N) PNA Reverse phase HPLC profile of PNA 15 (O) Reverse phase HPLC profile of PNA 16



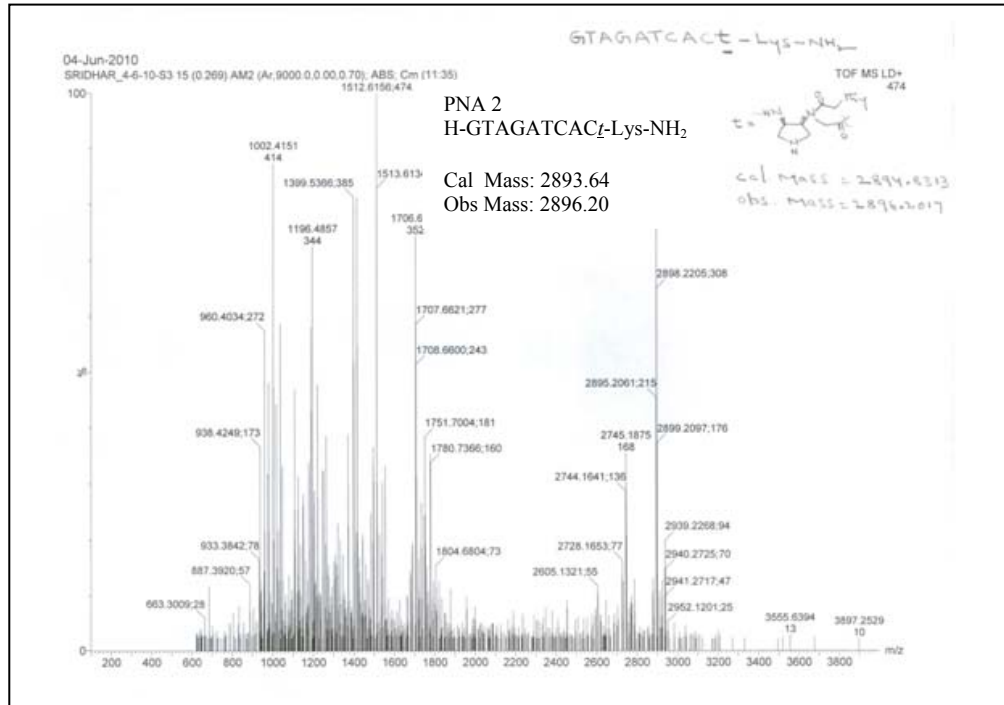


**Figure 25: (P)** Reverse phase HPLC profile of PNA 17

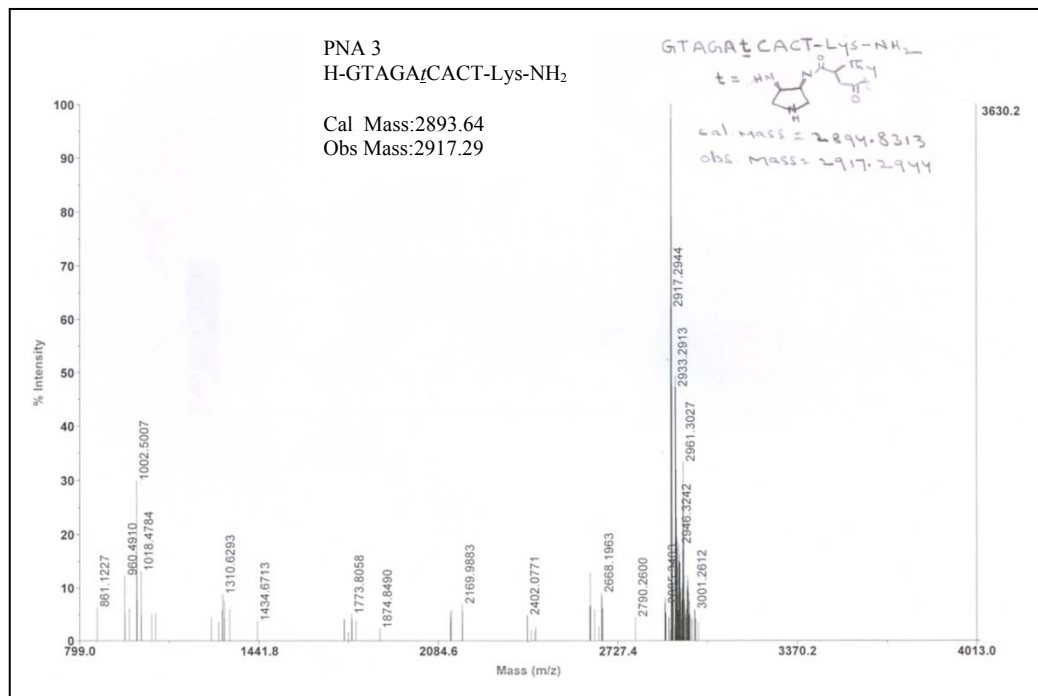


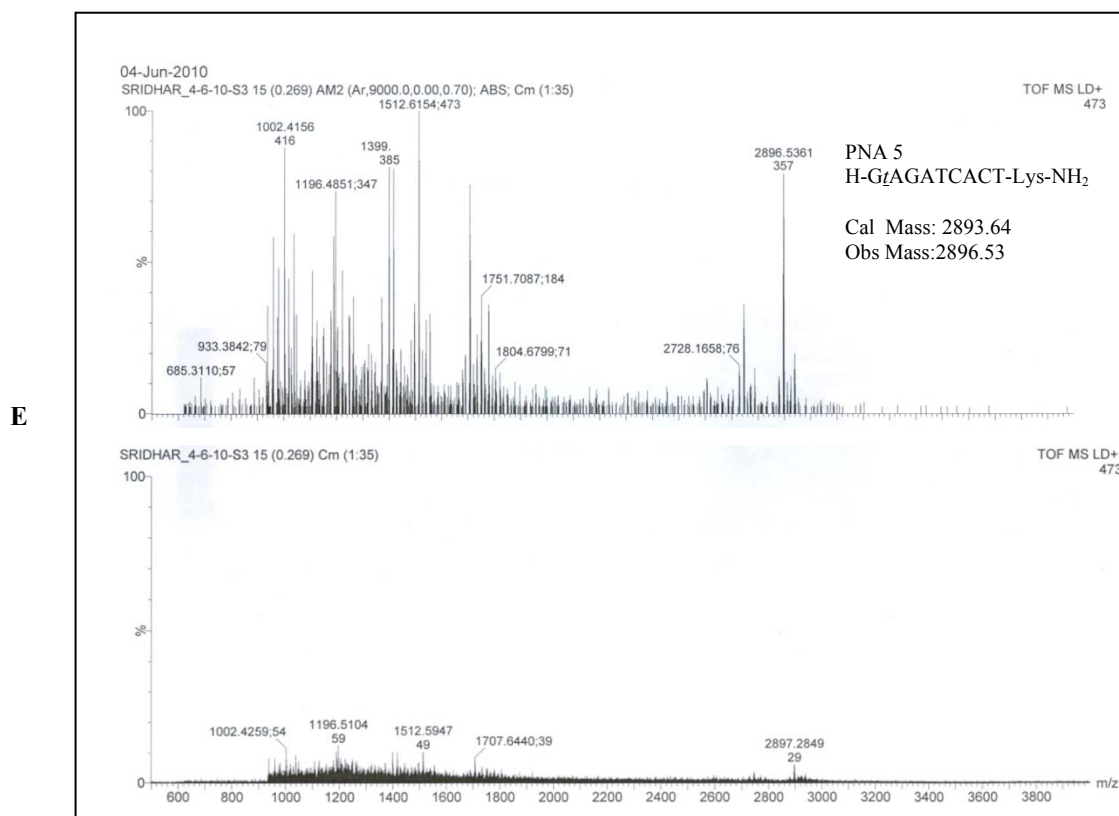
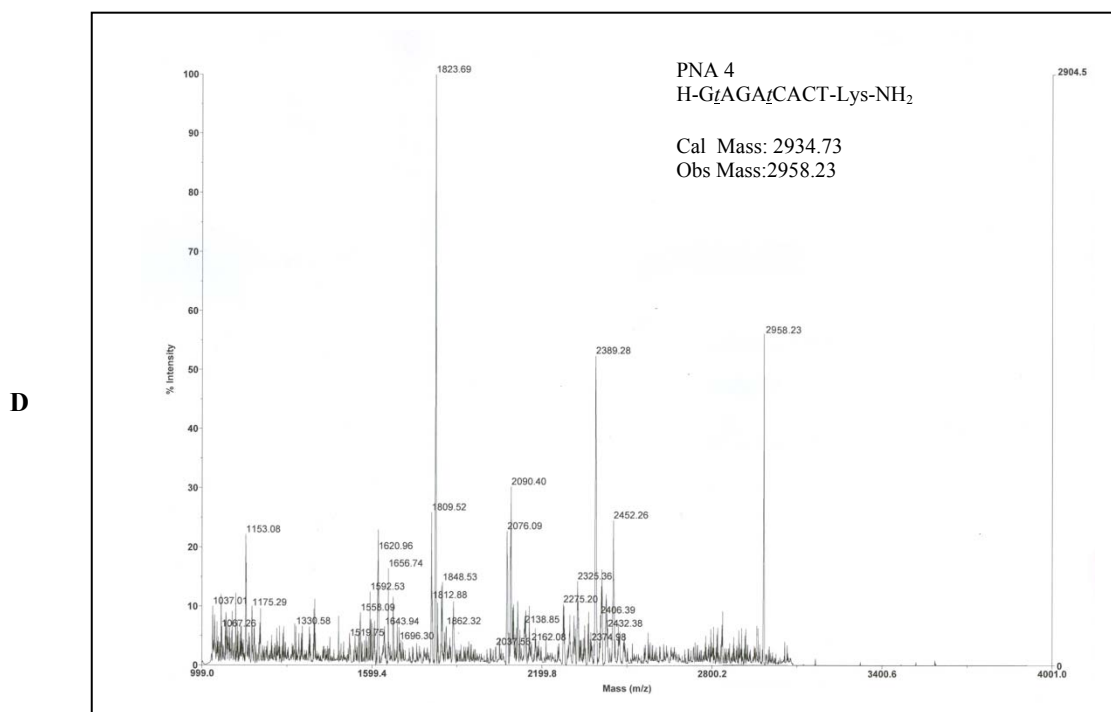
**Figure 26: (A)** MALDI-TOF Spectra of PNA 1

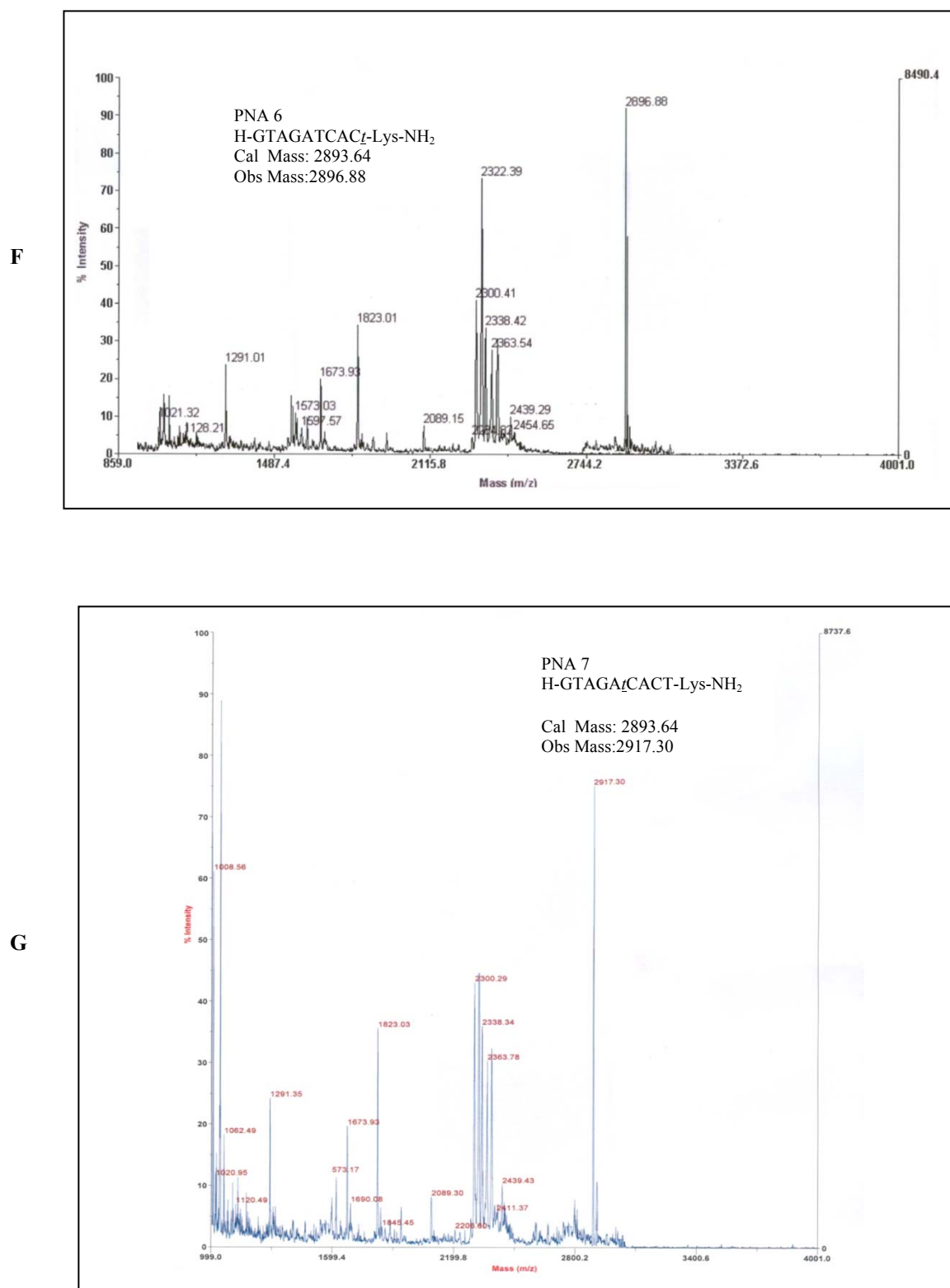
B



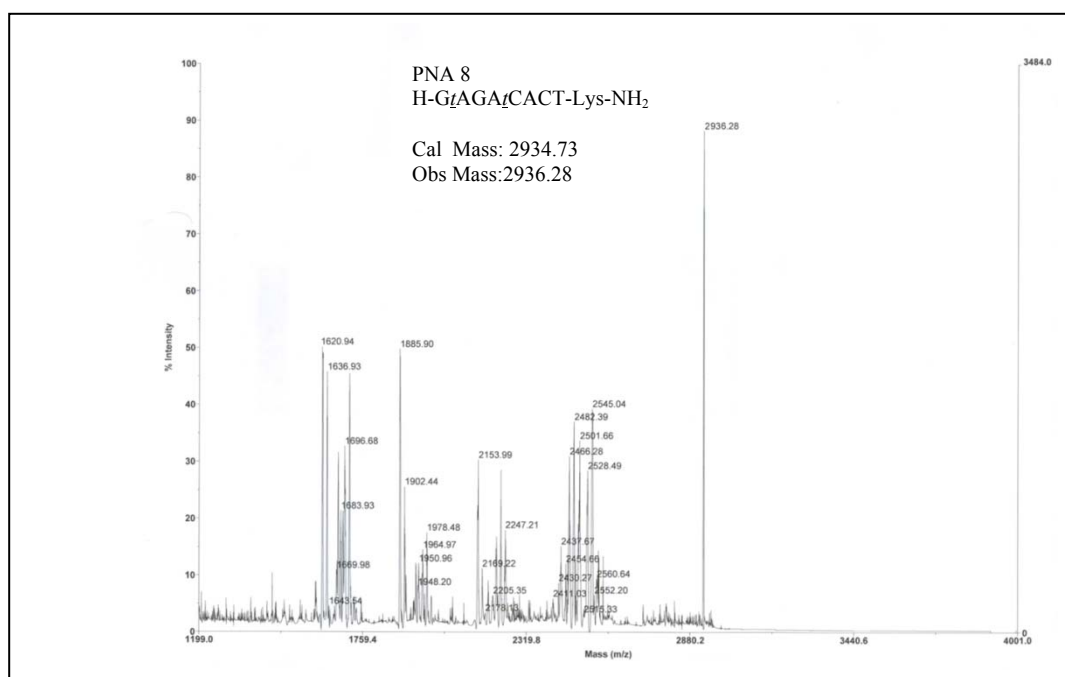
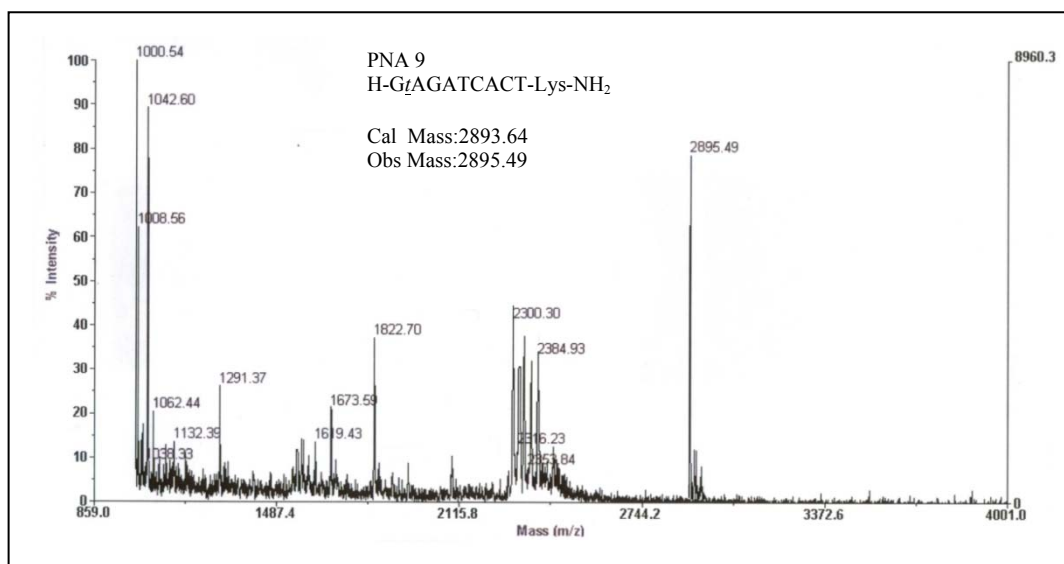
C



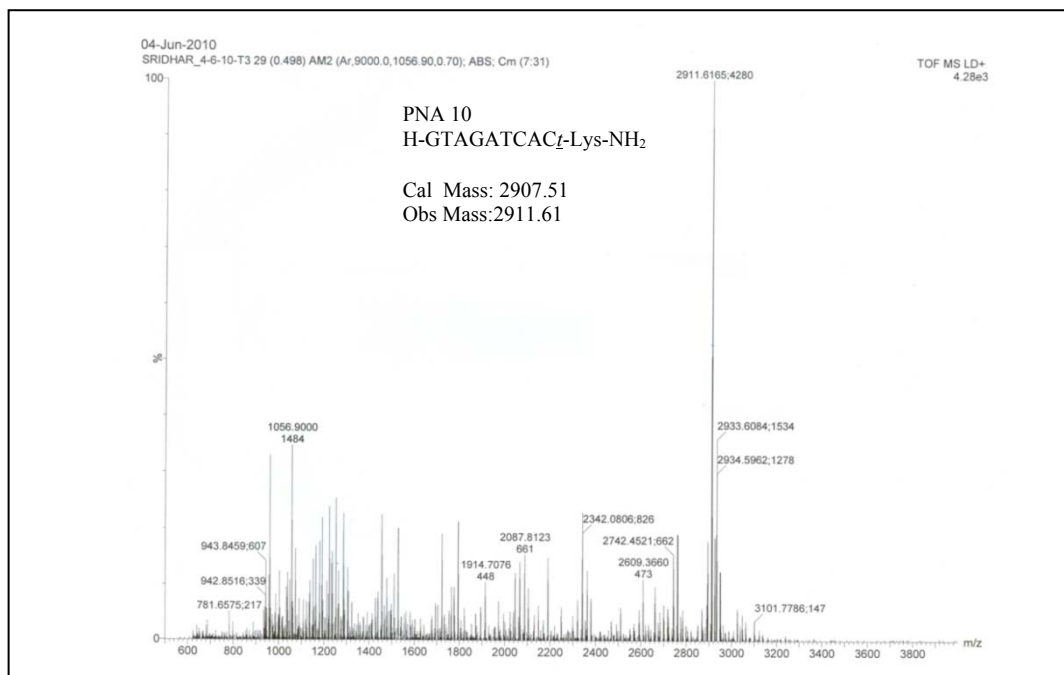
**Figure 27: (B) MALDI-TOF Spectra of PNA 2 (C) MALDI-TOF Spectra of PNA 3****Figure 28: (D) MALDI-TOF Spectra of PNA 4 (E) MALDI-TOF Spectra of PNA 5**



**Figure 29: (F)** MALDI-TOF Spectra of PNA 6 **(G)** MALDI-TOF Spectra of PNA 7

**H****I****Figure 30: (H) MALDI-TOF Spectra of PNA 8 (I) MALDI-TOF Spectra of PNA 9**

J



K

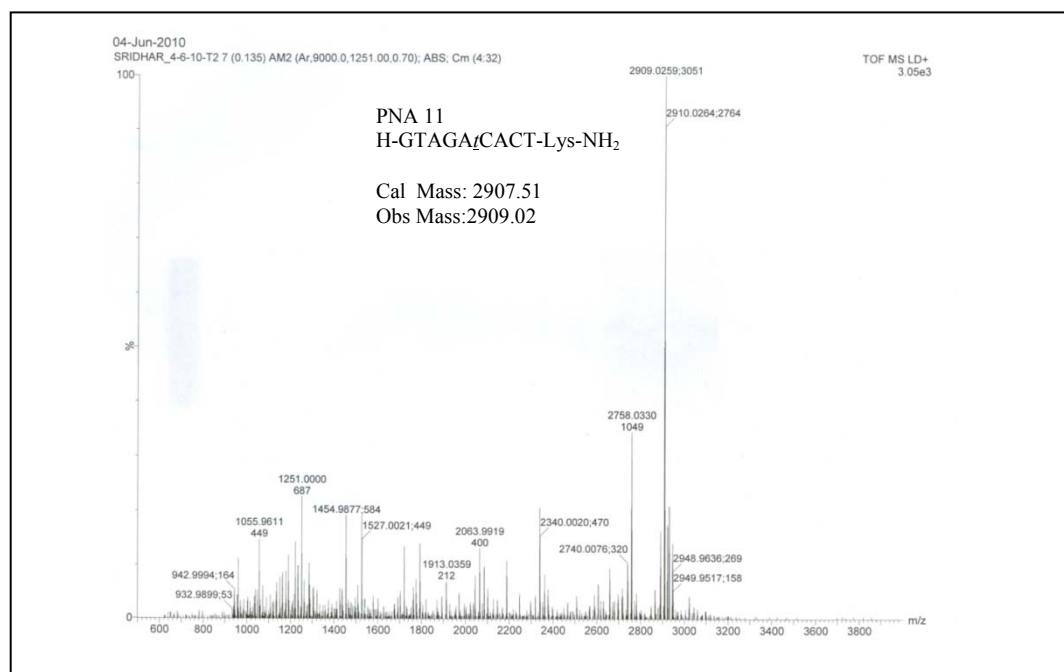
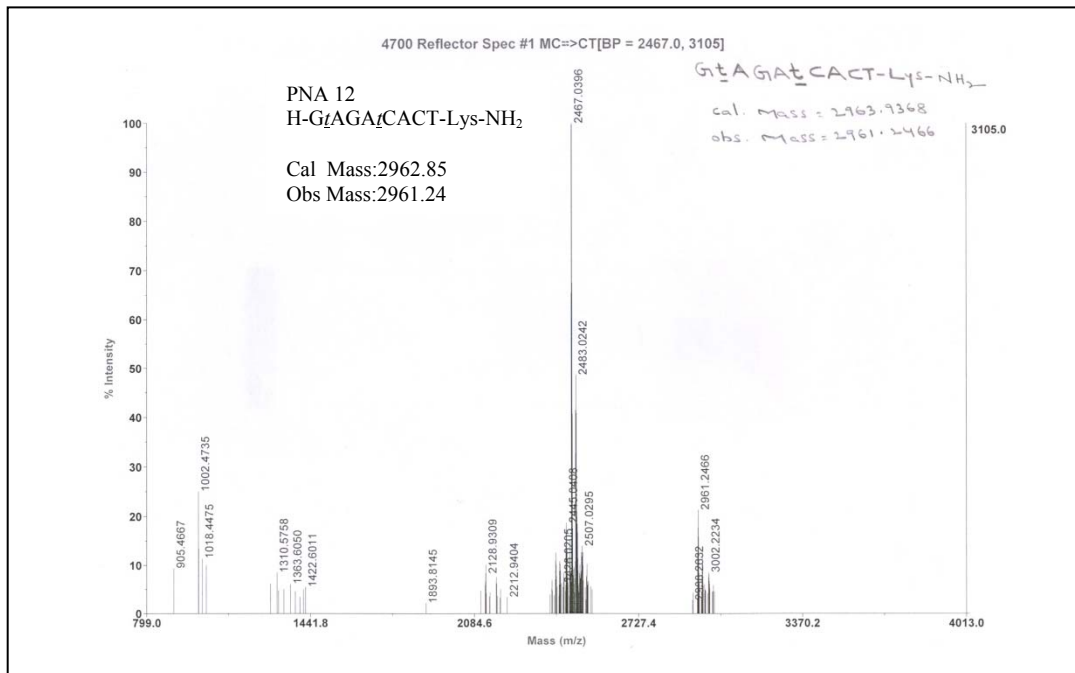


Figure 31: (B) MALDI-TOF Spectra of PNA 10 (C) MALDI-TOF Spectra of PNA 11

L



M

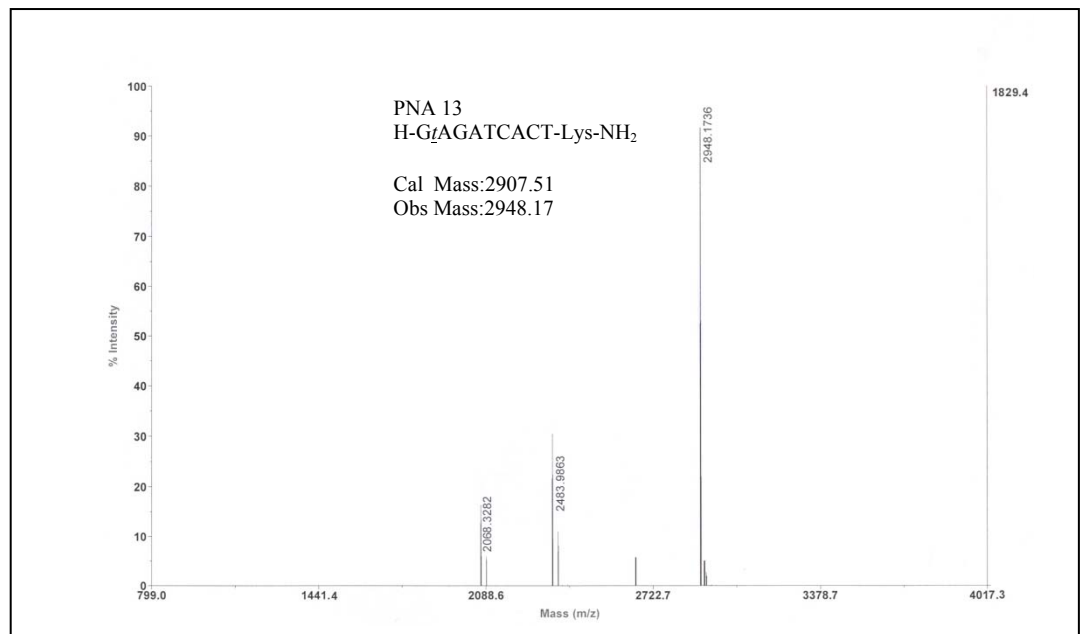
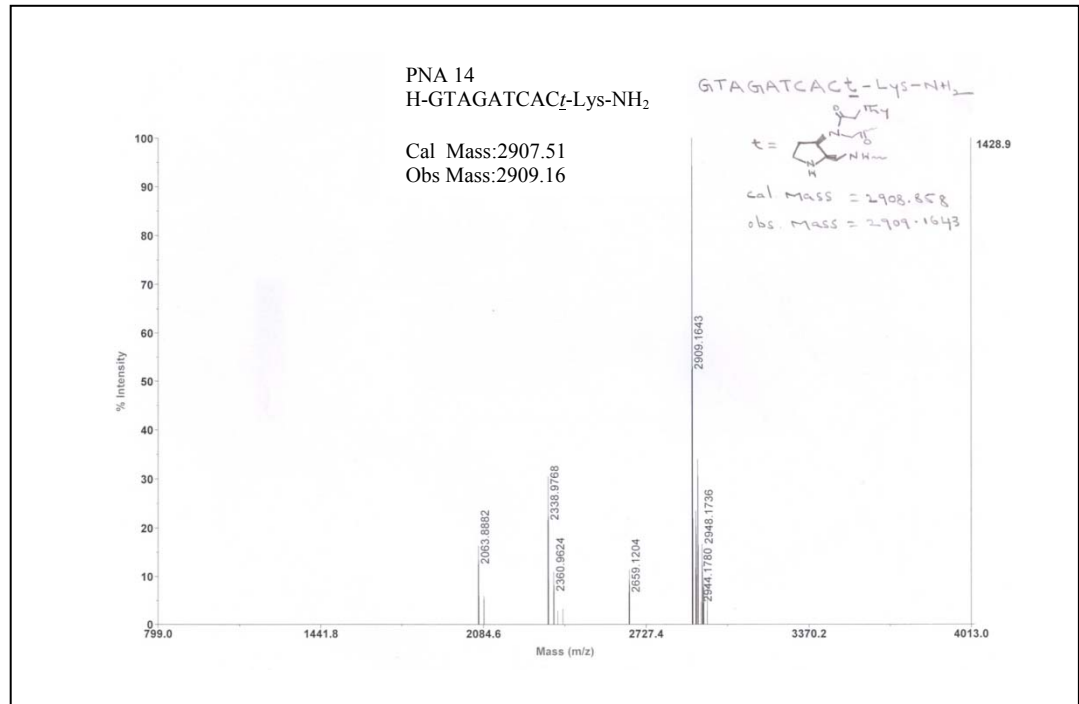


Figure 32: (L) MALDI-TOF Spectra of PNA 12 (M) MALDI-TOF Spectra of PNA 13

N



O

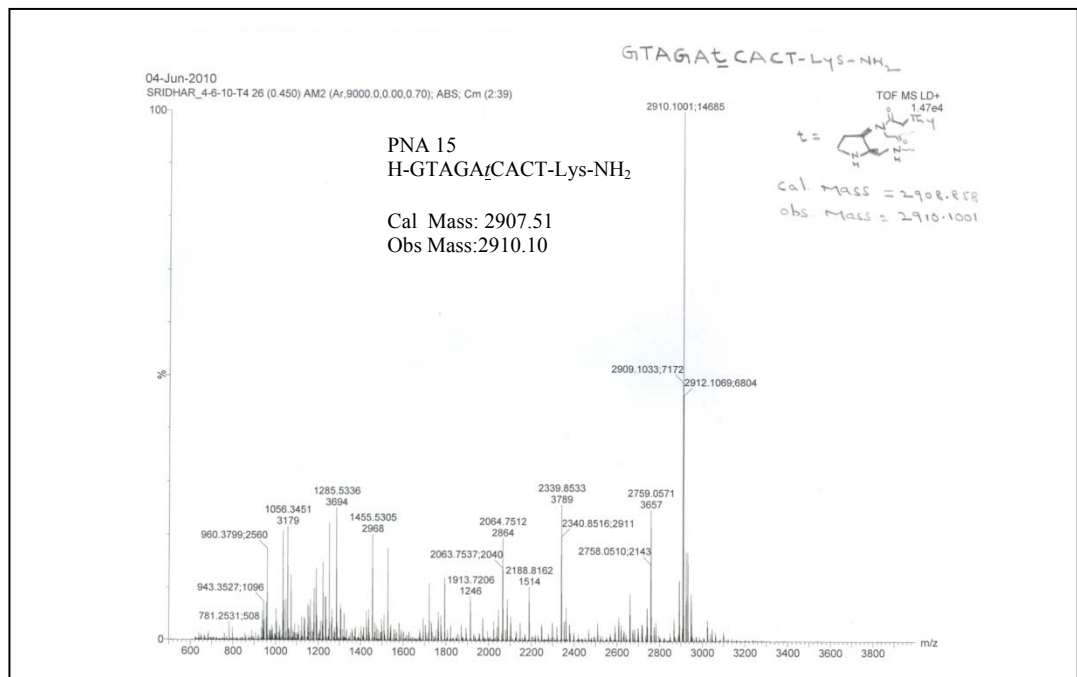


Figure 33: (N) MALDI-TOF Spectra of PNA 14 (O) MALDI-TOF Spectra of PNA 15



P

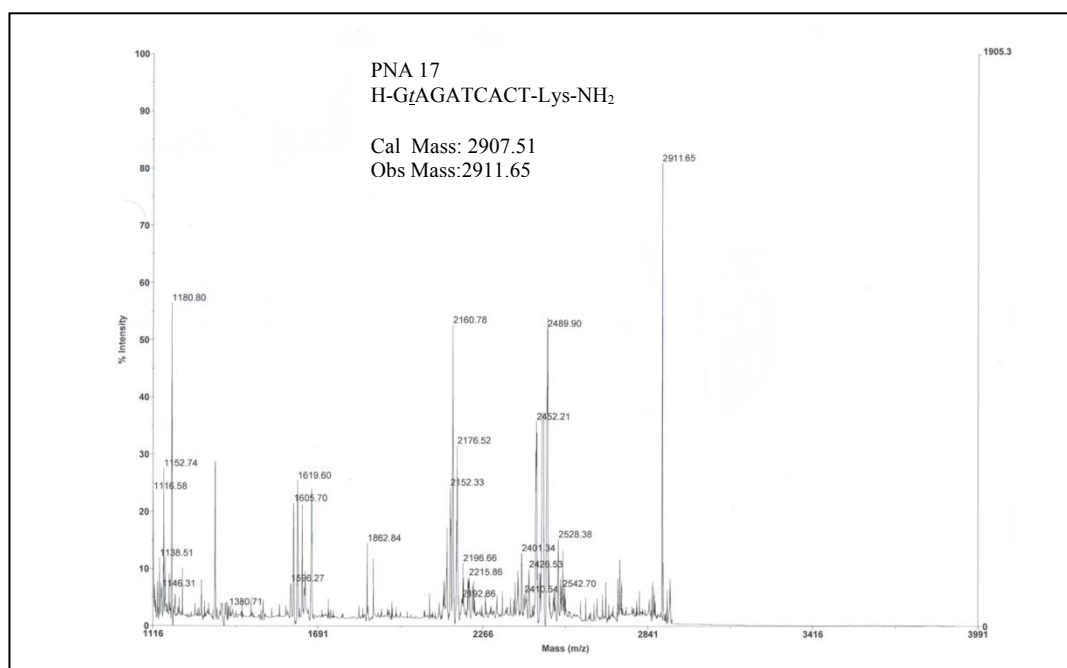
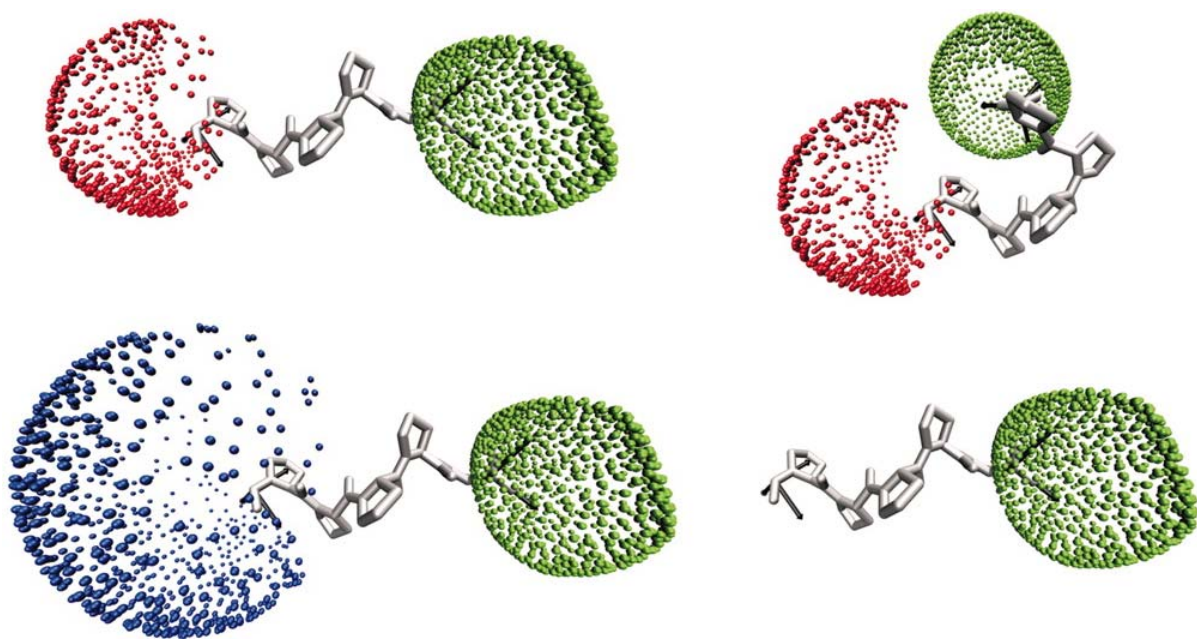


Figure 34: (P) MALDI-TOF Spectra of PNA 17

## Chapter 4: Synthesis and conformational studies of 3*S*-hydroxy substituted polyproline derivatives

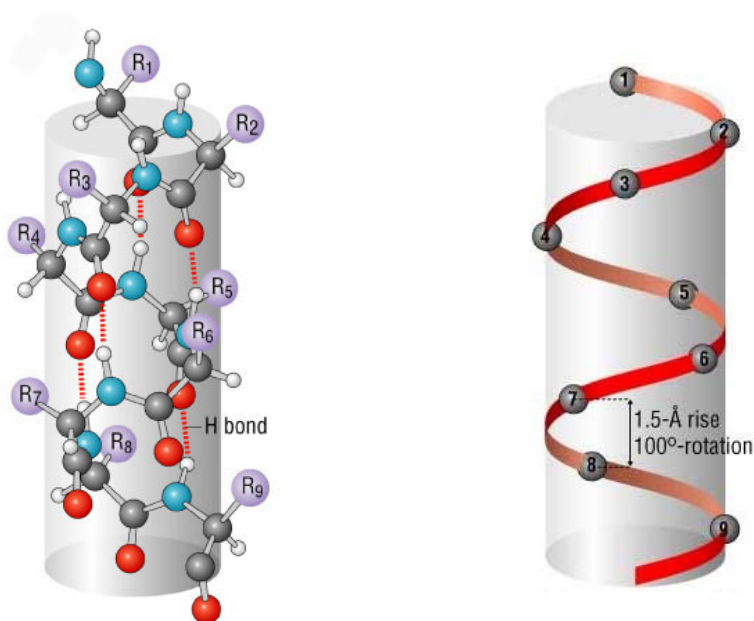


## 4.1 Introduction

The analysis of the factors determining the conformation of polyprolines is of intrinsic physio-chemical interest, to understand the molecular basis of conformation in biological macromolecules. The central features of the present chapter are to understand (i) effect of 3*S*-hydroxy substitution on polyproline conformation and (ii) the effect of terminal functional group on polyproline conformation.

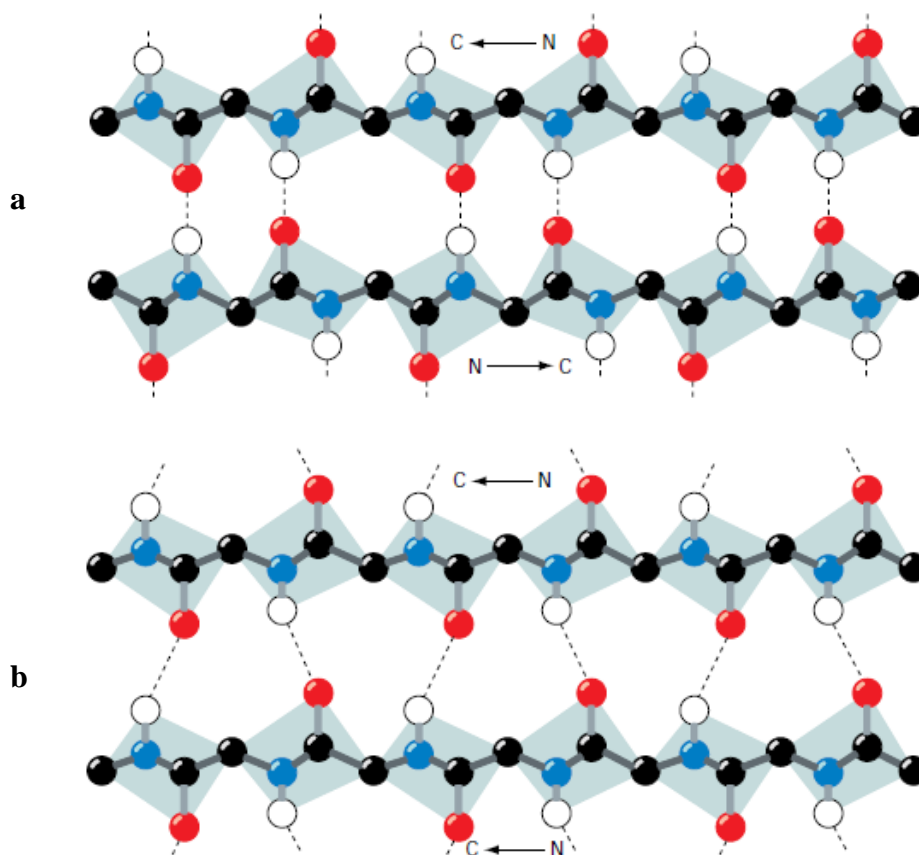
A protein is a polypeptide chain made up of amino acid residues linked together *via* peptide bond in a definite sequence. Amino acids are chiral (except for glycine, in which the normal asymmetric  $\alpha$ -carbon has two hydrogens), and the naturally occurring proteins contain L-amino acids. The sequence of side chains in protein determines all that is unique about a particular protein, including its biological function and its specific three-dimensional structure.

Folded structure of proteins consists of regular and highly periodic arrangements, designated as  $\alpha$ -helix and  $\beta$ -sheet. The key to form both the structures is the hydrogen bonding.  $\alpha$ -helix (Figure 1) is a right-handed spiral stabilized by hydrogen bonds between the amide nitrogen atom of each amino acid and the oxygen atom of the fourth amino acid in the chain. This corresponds to 3.6 amino acids for each turn of the helix. The side chain of the amino acid sticks out from this spiral backbone like the bristles on a bottle brush.



**Figure 1:**  $\alpha$ -helix structure in peptides<sup>1a</sup>

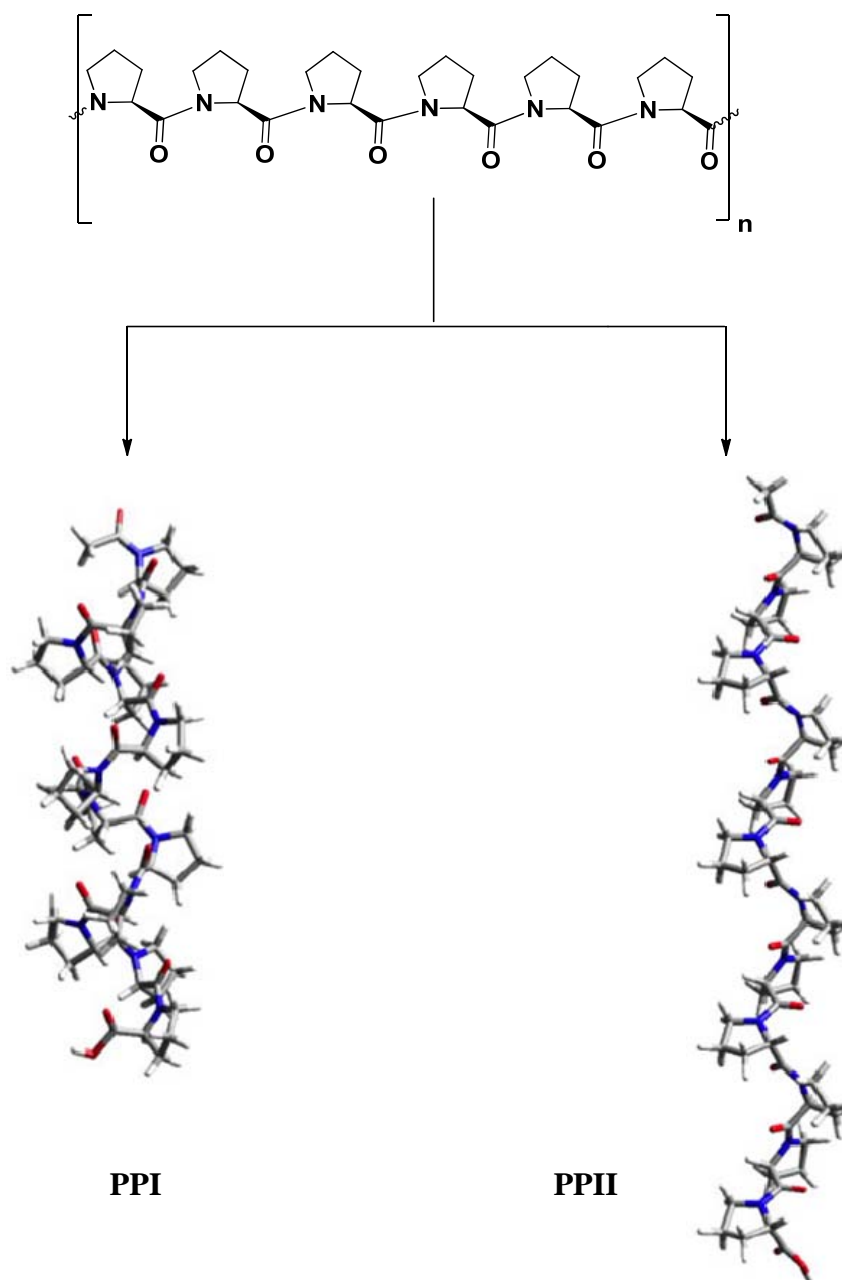
Beta sheets (Figure 2) consist of beta strands connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet. A beta strand (also  $\beta$  strand) is a stretch of polypeptide chain typically 3 to 10 amino acids long with backbone in an almost fully extended conformation. The sheets are “parallel” if all the chains run in the same direction and are “antiparallel” if alternate chains run in opposite directions.



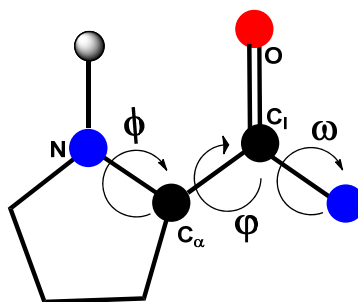
**Figure 2:**  $\beta$ -sheet structure in peptides a) antiparallel  $\beta$ -sheet b) parallel  $\beta$ -sheet<sup>1b</sup>

Repeating units of a single amino acid linked via amide bond (homo oligopeptide) provide model systems that adopt a definite secondary structure in different solutions. Among these polyaminoacids, poly(L-proline) has attracted the interest of several research groups since its preparation in the mid 1950s. The poly(L-proline) may adopt two different conformational forms, known as polyproline I (PPI) and polyproline II (PPII). Both secondary structures are helical and differ substantially in their physical and spectroscopic properties as well as in their crystallographic structures. PPI is a right-handed helix containing all peptide bonds in “*cis*” disposition

and is compact, having 3.3 residues per turn with a helical pitch of  $5.6\text{\AA}/\text{turn}$ . The backbone dihedral angles of PPI are  $(\phi, \psi, \omega) = (-75, 160, 0)$ . In comparison, the PPII helix is a left handed helix with “*trans*” peptide bonds and is extended with 3 residues per turn and have  $9.3\text{\AA}$  of helical pitch. The backbone dihedral angles of PPII are  $(\phi, \psi, \omega) = (-75, 145, 180)$ . The kind of conformation PPI or PPII adopted by poly (L-proline) depends upon the solvent.



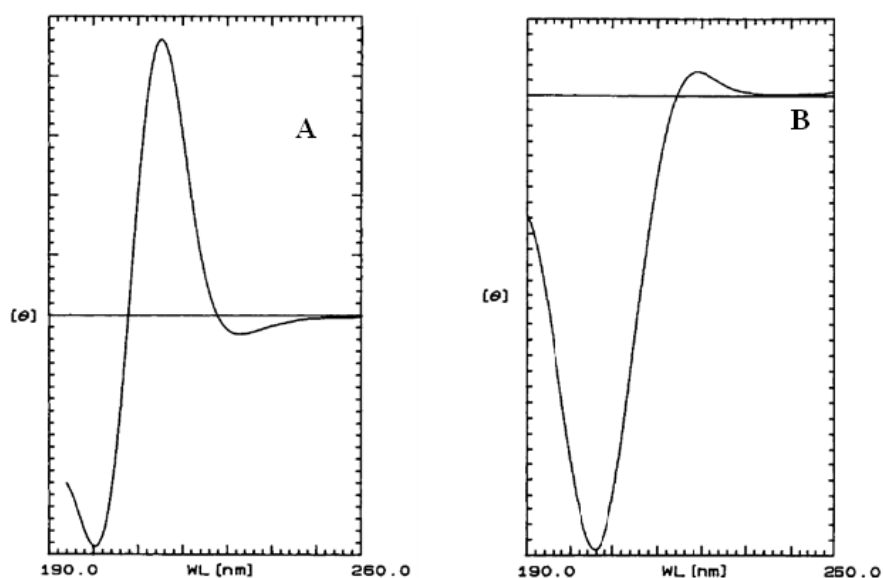
**Figure 3:** Structures of PPI and PPII<sup>2</sup>



**Figure 4:** Backbone dihedral angles of peptide

PPI conformation is favoured when the polyproline is dissolved in aliphatic alcohols, whereas form PPII is favoured in aliphatic acids (AcOH, Propionic acid) or water.

Unlike  $\alpha$ -helices and  $\beta$ -sheets, PPI and PPII helices cannot form intramolecular hydrogen bonds on backbone because of the lack of an amide hydrogen. Thus detecting these secondary structures in folded proteins directly by X-ray crystallography or NMR spectroscopy can be difficult, often leading to their mistaken assignment as random coils.<sup>3</sup> PPI is characterized by the circular dichroism (CD) spectrum with a negative band of medium intensity at 198-200 nm, a strong positive band at 214-215 nm, and a weak negative band at 231-232 nm (Figure 5). PPII shows a strong negative band at 202-206 nm and a weak positive band at 225-229 nm (Figure 5).



**Figure 5:** (A) CD spectral pattern of polyproline I (B) CD spectral pattern of polyproline II

Of both secondary structures, PPII form is biologically relevant in the animal and plant kingdoms. PPII helices are widespread in both folded proteins<sup>4</sup> and unfolded polypeptides<sup>5</sup>, and are known to play important roles in biological signal transduction, transcription, cell motility, and immune responses.<sup>6,7</sup> In addition, the single strands of collagen with the typical [Pro-Hyp-Gly]*n* repeat unit adopt a PPII-like conformation.<sup>8</sup> Whereas PPI helices are rare in biological contexts.

**Table 1:** Conformational properties of PPI and PPII

S.No	Conformer property	PPI	PPII
1	Type of helix present	Right handed	Left handed
2	Conformation of amide bond	“ <i>cis</i> ”	“ <i>trans</i> ”
3	No. Of residues per turn	3.3	3
4	Helical pitch	5.6Å <sup>0</sup> /turn	9.3Å <sup>0</sup> /turn
5	Backbone dihedral angles	$\phi = -75^0$ $\Psi = 160^0$ $\omega = 0$	$\phi = -75^0$ $\Psi = 145^0$ $\omega = 180^0$

The PPII helix has been proposed as an important local conformation in disordered states of proteins that are relevant for protein folding.<sup>9,10</sup> Furthermore, the oligoprolines have been used as molecular spacers in designing of energy or electron transport<sup>11,12</sup> molecules and are attractive as functionalizable molecular scaffolds that allow for switching between the PPII and PPI helix. Hence understanding of the factors that influence the PPII conformation is an important goal.

The biophysical reasons for proline being a common binding motif is, its uniqueness among the 20 common amino acids in having the side-chain cyclised onto the backbone nitrogen atom which leads to limited conformations with backbone  $\phi$  angles of  $\sim 65^\circ$ . Because of the bulky N-substitution on proline, it also restricts the conformation of the residue preceding the proline with a strong preference for a  $\beta$ -sheet conformation. As a consequence, polyproline sequences tend to adopt the PPII helix, which is an extended structure with three residues per turn. The PPII helix is an unusual structure in which the prolines form a continuous hydrophobic strip round the surface of the helix, while the backbone carbonyls are present as ideal hydrogen bonding acceptor sites. Being both conformationally restricted and electron- rich, the PPII helices are present as easily accessible hydrophobic surface with good hydrogen-bonding acceptor sites. The accessibility of PPII helices is greatly enhanced by the fact that they are frequently found either at the amino or carboxyl termini of proteins where they form extended structures described as 'sticky arms'.<sup>13</sup> The relative rigidity of polyproline stretches means that they are pre-organised and lose little conformational entropy on binding and thus bind more favorably than other exposed (i.e., nonglobular) peptide sequences. Proline-rich sequences cannot bind as tightly as globular domains. However, weaker binding can be of great advantage, as it allows the binding of proline-rich regions to be modulated rapidly.

## 4.2 Factors effecting the polyproline conformation

A detailed analysis of the factors determining the conformation of poly L-proline is not only of intrinsic physiochemical interest but may also increase our understanding of the molecular basis of conformation in biological macromolecules. This is especially true since low molecular weight polyprolines exist in only two conformational states. Earlier studies have concluded that the observed secondary structure of polyproline, as well as the nature of the cooperativity, originate from repulsive steric interactions and attractive van der Waals forces. Intramolecular hydrogen bonds cannot occur because of the lack of donor amide hydrogen.

### 4.2.1 Solvent effect

The interaction of solvent with polypeptide is one of the primary factors controlling protein folding and stability. The proline oligopeptides form the polyproline-II helix in water, trifluoroethanol, benzyl alcohol and in organic acids.<sup>14,15,16</sup> However, there can be conformational transition from the polyproline-II to the



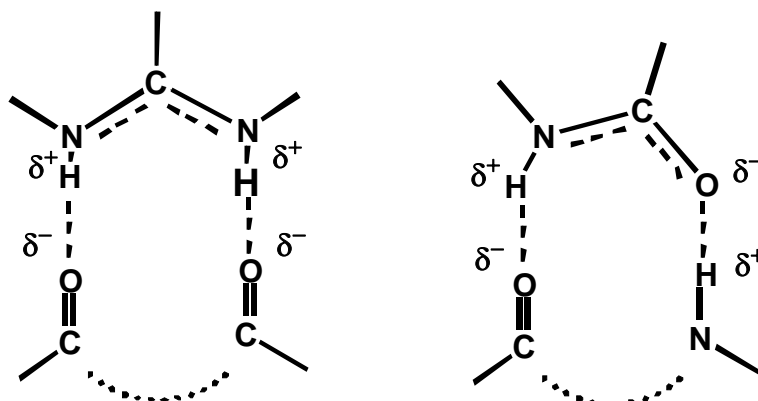
polyproline-I helix in methanol and 1-propanol. Oligopeptides can form the polyproline-I helix in pure aliphatic alcohols such as methanol and 1-propanol<sup>14,15,16</sup>. It was also shown that the propensity of forming polyproline-I helix is more favourable in 1-propanol than in methanol, and the longer the chain-length, the greater is the stability of the polyproline-I helix.

#### 4.2.2 Stereoelectronic effects

To understand the effects of stereoelectronic and electronegativity on polyproline conformation Raines *et al*<sup>17</sup> reported the conformational study of four peptides, (Pro)<sub>10</sub>, (Hyp)<sub>10</sub>, (Flp)<sub>10</sub>, and (flp)<sub>10</sub> where Hyp is (2*S*,4*R*)-4-hydroxyproline, Flp is (2*S*,4*R*)-4-fluoroproline, and flp is (2*S*,4*S*)-4-fluoroproline. They used fluoro group as an electron-withdrawing substituent because organic fluorine does not form hydrogen bonds and has little effect on aqueous solvation. They found that an electron-withdrawing substituent at the 4*R* position of proline can stabilize the PPII conformation, and that an electron-withdrawing substituent at the 4*S* position of proline can stabilize the PPI conformation. These results reveal an important role for stereoelectronic effects in the conformation of polyproline.

#### 4.2.3 Effect of Urea

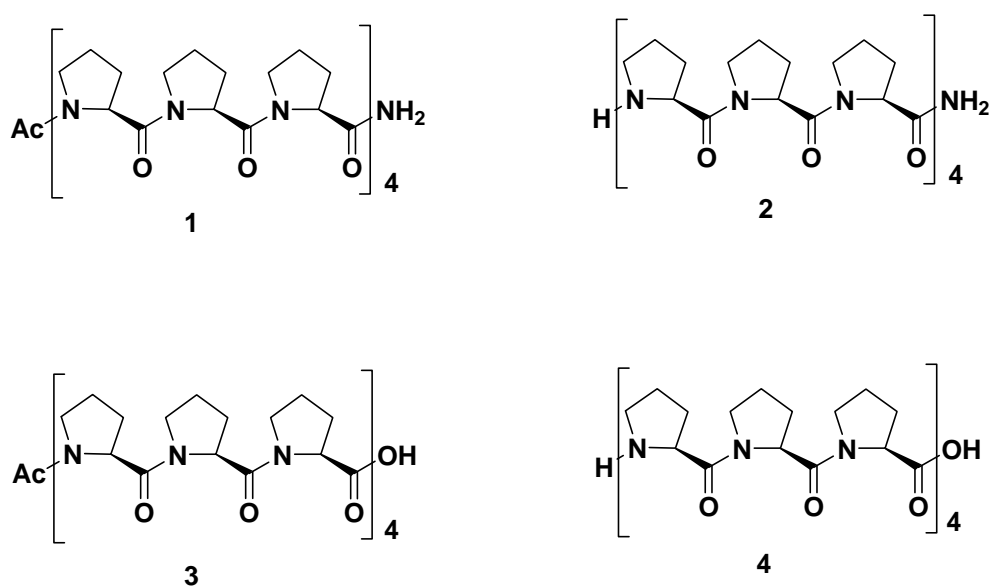
Jencks and co workers<sup>18</sup> examined mechanisms by which urea interacts with the peptide and amide groups of proteins and suggested that urea can form polydentate hydrogen bonding to peptide groups as shown in Figure 6. Polydentate hydrogen bonds have a more favorable entropy of formation and are more stable than single hydrogen bonds, and the structural rigidity of urea will facilitate the formation of such cyclic structures.



**Figure 6:** Polydentate H-bonding in peptide groups<sup>10</sup>

Creamer *et al*<sup>19</sup> demonstrated that urea promotes the formation of left-handed polyproline II (PPII) helical structures in both short peptides and denatured proteins. The observed increase in PPII content is sequence-dependent and the PPII helical content increases with increasing concentrations of urea.

**4.2.4 Effect of terminal functional groups:** Wennemers *et al*<sup>20</sup> examined the conformational stability of the polyproline helix with respect to the functional groups at the C- and N-termini, of the 12-mer oligoprolines with charged and capped termini (Figure 7). The relative ease with which they switch the conformation was used as a measure to analyze their conformational stabilities.



**Figure 7:** Oligoprolines **1 - 4** With different functional groups at the C and N termini

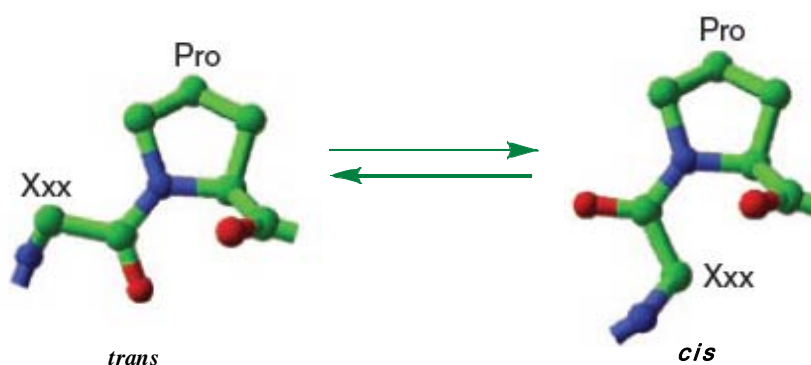
CD spectroscopic studies of these model peptides demonstrated that the positively charged N-terminus and negatively charged C-terminus destabilize the PPII helix and favor the PPI helix, whereas capped termini favor the PPII over the PPI helix. The electrostatic interactions between the terminal charges and the amide dipoles stabilize both helices and such change in dipole interactions are responsible for the above observed relative stabilities. They are significantly stronger in the PPI helix where the amide bonds are oriented almost linear to the helix axis as compared to the PPII helix in which the amides are nearly perpendicular to the axis. It was demonstrated that a negative charge at the C-terminus has a more pronounced effect on the relative stability as compared to a positive charge at the N-terminus. Destabilization of the PPII helix

occurs due to repulsive interaction between the C-terminal carboxylate with the neighboring amide bond.

#### 4.2.5 Effect of conformer of X<sub>aa</sub>-Pro amide bond

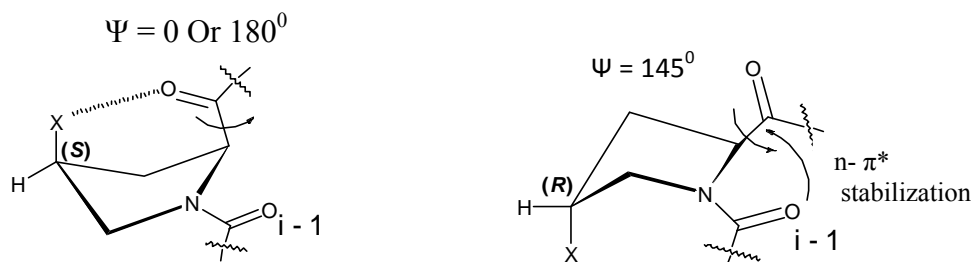
In PPI helix all peptide bonds are in *cis* configuration while in PPII helix all peptide bonds are in *trans* configuration. The isomerization between *cis* and *trans* conformers in X<sub>aa</sub>-Pro amide bonds (X<sub>aa</sub>= any amino acid, Pro= proline) are crucial in many natural processes such as protein folding and signal transduction. Understanding of the factors that determine the *cis/trans* conformer ratio and the development of tools that allow for the tuning of this equilibrium therefore becomes important.

Two important factors that influence the *cis-trans* conformer are the ring puckering and substitution at C4 ( $\gamma$ )-position.



**Figure 8:** Prolyl *cis-trans* isomerisation<sup>21</sup>

Conformational studies have shown that the nature of the substituent at C4 and the absolute configuration at this center critically influence both the pyrrolidine ring pucker and the *cis/trans* conformer ratio of the amide bond in X<sub>aa</sub>-Pro bonds. C4-*exo* ring pucker has a stronger noncovalent interaction between adjacent amide bonds through an  $n-\pi^*$ <sup>22</sup> interaction (Figure 9) thereby favouring the *trans* conformation. In C4-*endo* ring pucker, the weaker noncovalent interaction leads to a higher population of the *cis* conformer.



**Figure 9:** Endo and exo conformation of 4-substituted proline

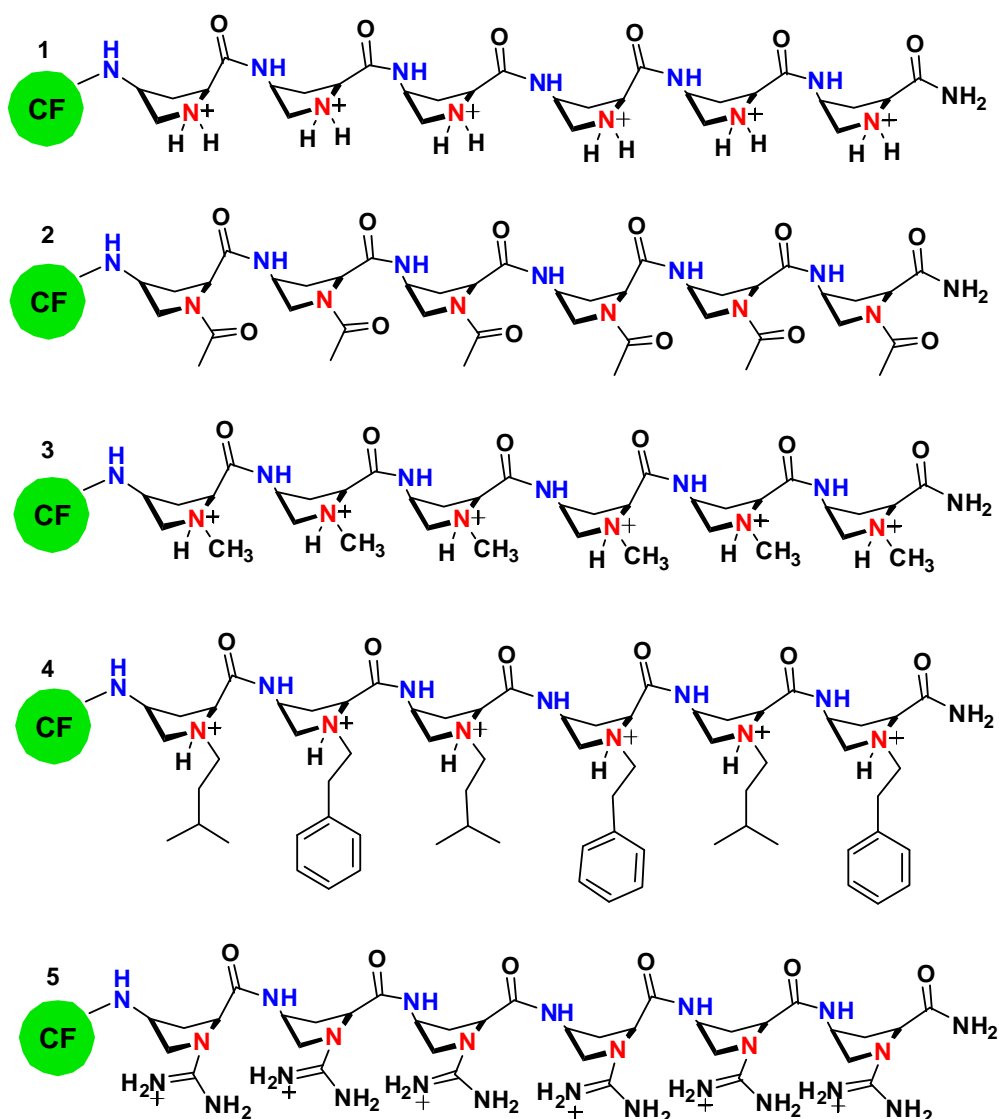
Wennemers *et al*<sup>23</sup> used the concept of intramolecular hydrogen bonding and designed a proline derivative with hydrogen bond donating substituent at  $\gamma$ -position which adopts preferentially C4-endo conformation and favour the trans conformer. It was demonstrated that the ring pucker is not the only factor that influences *cis/trans* ratio of the amide bond. In addition, noncovalent interactions between the substituent at C4 of the proline ring and the amide backbone plays a key role in *cis/trans* conformer ratio in Xaa–Pro bonds.

### 4.3 Importance of polyproline II

#### 4.3.1 Utility as cell penetrating agents

Cell penetrating peptides (CPPs) constitute one of the most promising tools for delivering biologically active molecules into cells and therefore play a key role in the future of disease treatments. CPPs have been shown to efficiently improve intracellular delivery of various biomolecules, including plasmid DNA, oligonucleotides, siRNA (short interfering RNA), PNA (peptide nucleic acid), proteins and peptides, as well as liposome nanoparticles, into cells both *in vivo* and *in vitro*. Factors required for CPPs which can facilitate to interact with target are i) to overcome both extra and intracellular limitations ii) to trigger the movement of cargo across the cell membrane into the cytoplasm iii) improve its intracellular trafficking.

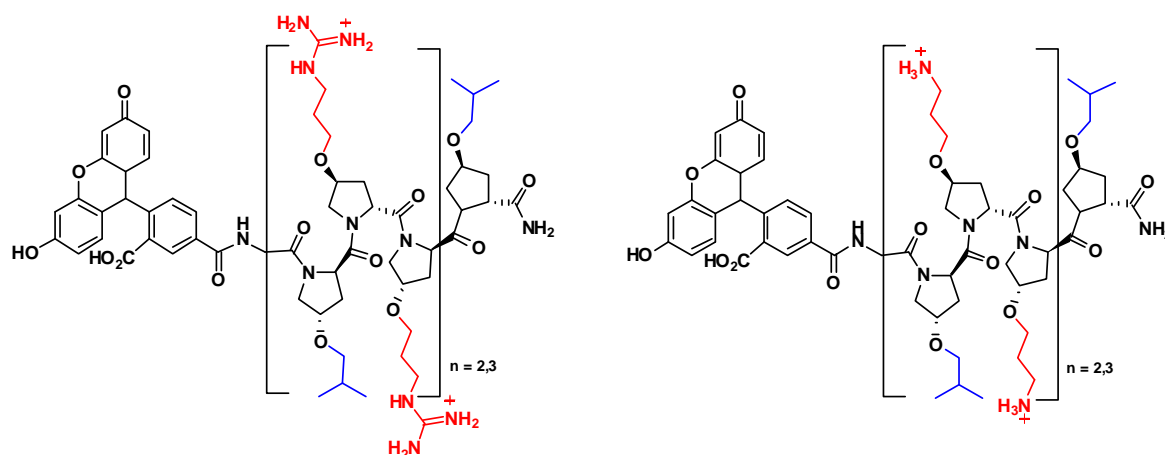
Royo *et al*<sup>24</sup> described three different  $\gamma$ -peptide families derived from *cis*- $\gamma$ -amino-L-proline with a common backbone, but distinct different side chains *i.e.*  $N^\alpha$ -acyl/alkyl/guanidilated- $\gamma$ -peptides (Figure 10) to assay their cellular uptake properties.



**Figure 10:**  $\gamma$ -Aminoproline monomer based  $\gamma$ -peptides labelled with 5(6)-carboxyfluorescein<sup>24</sup>

The modulation of this side chain structure offers compounds with different hydrophobic/hydrophilic character and consequently, with different properties as carriers. The stability imparted by the  $\gamma$ -peptide skeleton of these compounds circumvents problems associated with protease liability, a major limitation of CPPs. These  $\gamma$ -peptides enter into different cell lines (COS-1 and HeLa) via an endocytic mechanism with advantages over the well-known penetrating TAT peptide, such as being less toxic than TAT and protease resistance.

Jean Chmielewski *et al*<sup>25</sup> incorporated different functionality into the PPII helix by O-alkylation of 4-hydroxyproline monomer to yield a amphiphilic scaffold that displays both hydrophobic and cationic moieties. These amphiphilic agents were obtained by



**Figure 11:** Structure of amphiphilic polyproline oligomers<sup>25</sup>

taking advantage of the 3-fold repeating unit of the PPII helix. In their design, one face of the PPII helix was hydrophobic, while the other two faces contained cationic groups. The added groups in side chain have little effect on the backbone structure with minimal cytotoxicity and improved cellular uptake profile. Recently our group reported that the 4(*R/S*)-amino/guanidylprolyl cationic collagen peptides, boost the transfection efficiencies and are functional enhancers in transfecting the gene-encoded plasmids.<sup>32</sup>

#### 4.3.2 Rigid molecular Spacer

Eaton *et al*<sup>26</sup> have done Forster resonance energy transfer (FRET) measurements to determine the quantitative distance information. Single molecule experiments were done on polypeptides and protein studied by FRET between dyes attached to the N and C termini of polyproline of various lengths. Polyproline was used as a spectroscopic reference for single-molecule fluorescence experiments and because of the large number of potential photochemical complications in single molecule studies on biomolecular dynamics, labelled polyproline peptides are an important standard for calibrating the measurements and for testing and refining new experimental methods.

### 4.3.3 Cell motility

Many proline-rich proteins participate in cell motility by delivering actin monomers to specific cellular locations where actin rich membrane protrusions are formed that are necessary for cell motility. Actin monomer is usually delivered to the site of polymerization in the form of profilactin **A** complex of G-actin with profilin **A** polyproline-binding protein. Profilin binds to poly-L-proline and phosphatidylinositol (4,5)-bisphosphate [PtdIns(4,5)P<sub>2</sub>] as well as to G-actin and facilitates the funneling of actin monomers to membrane-associated sites of actin polymerization.

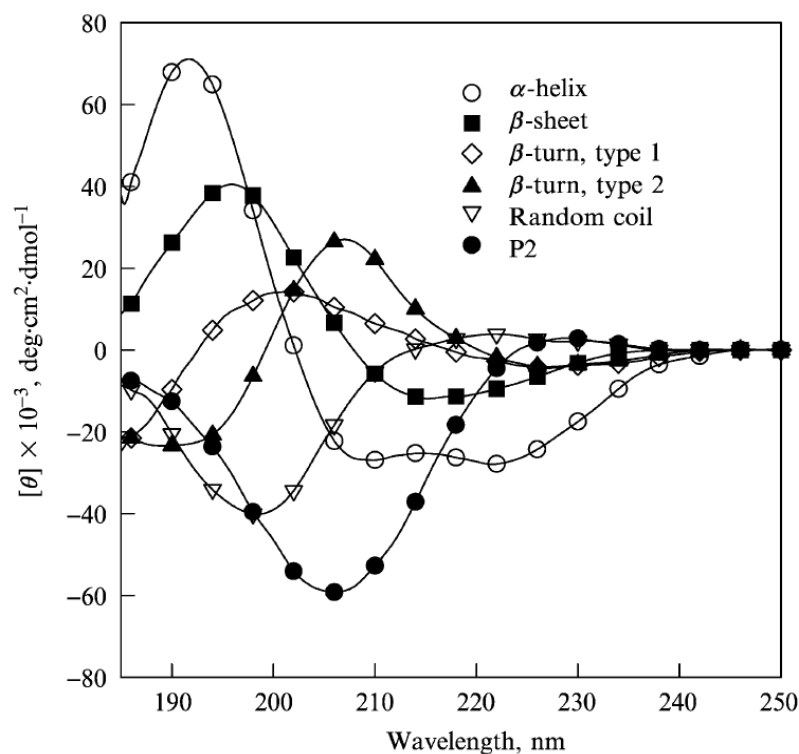
### 4.3.4 Collagen structure

Collagen is an abundant structural protein in all animals. In humans, collagen comprises one third of the total protein, accounts for three quarters of the dry weight of skin and is the most prevalent component of the extracellular matrix (ECM). Twenty-eight different types of collagen composed of at least 46 distinct polypeptide chains have been identified in vertebrates. The defining feature of collagen is an elegant structural motif in which three parallel polypeptide strands in a left-handed, polyproline II-type (PPII) helical conformation coil about each other with a one-residue stagger to form a right-handed triple helix. The tight packing of PPII helices within the triple helix mandates that every third residue be Gly, resulting in a repeating X<sub>aa</sub>Y<sub>aa</sub>Gly sequence, where X<sub>aa</sub> and Y<sub>aa</sub> can be any amino acid. The strands of human collagen contains 22% of Pro or Hyp residues. The abundance of these residues preorganizes the individual strands in a PPII conformation, thereby decreasing the entropic cost for collagen folding.

### 4.4 Conformation of peptides by circular dichroism (CD)

Molecules without a center of symmetry often exhibit the property of circular dichroism (CD), which is the unequal absorption of right-handed and left-handed polarized light. There are many basic reviews of the CD of proteins in the literature.<sup>27–29</sup> Proteins and polypeptides show CD bands in the far ultraviolet region (178–260 nm) which arise mainly from the amide functionalities of the protein backbone and are sensitive to their conformations. In addition, proteins have CD bands in the near ultraviolet (350–260 nm) and visible regions, that arise from aromatic amino acids and prosthetic groups. The CD in these regions depends on the tertiary structure of the protein. The changes in Circular dichroism in both the near and far UV as a function of

temperature, denaturants, and time can be used to estimate the thermodynamic and kinetic parameters of protein and nucleic acid folding.

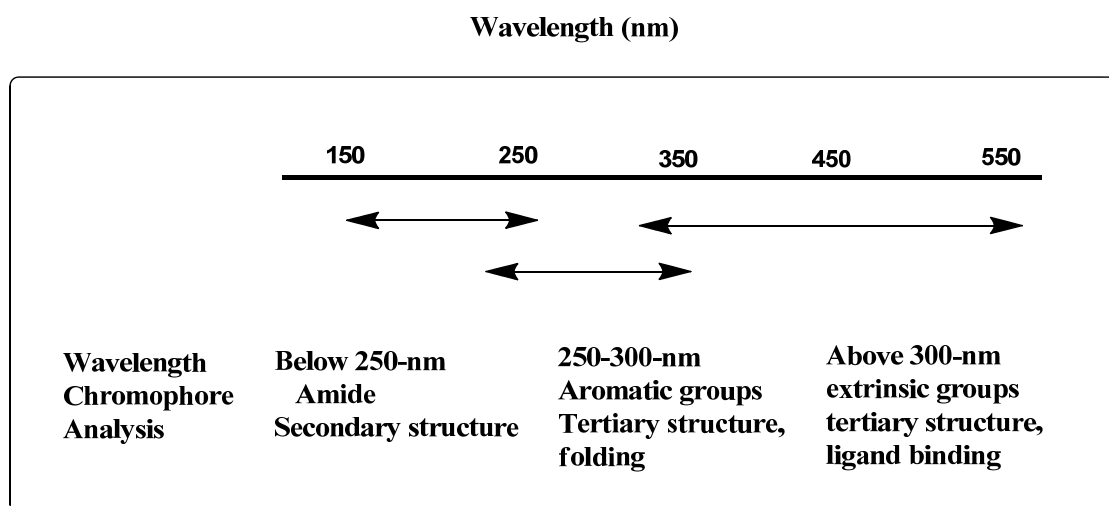


**Figure 12:** Representative circular dichroism curves corresponding to common secondary structures<sup>30</sup>

In addition, there are often changes in both the near and far UV bands when proteins interact with ligands or other proteins. Since CD is a quantitative technique, changes in CD spectra are directly proportional to the amount of the complexes formed, and these changes can be used to estimate the binding constants. The spectra for various common secondary structures of polypeptides is shown in Figure 12.

The CD spectra of proteins are generally divided into three wavelength ranges, based on the energy of the electronic transitions that dominate in the given range (Figure 13). These are (1) the far UV (below 250-nm), where the peptide contributions dominate, (2) the near UV (250–300-nm), where aromatic side chains contribute, and (3) the near UV–visible region (300–700-nm), where extrinsic chromophores contribute. Applications of CD to protein structure and folding have been developed on the basis of these characteristic origins of protein CD spectra.



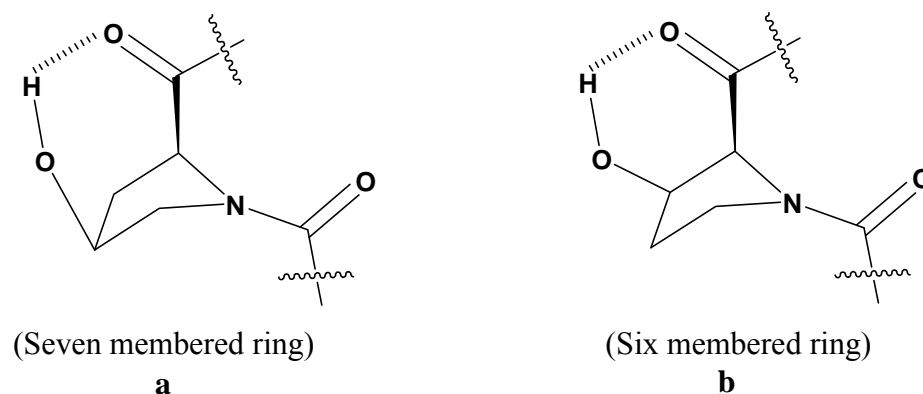


**Figure 13:** CD spectral regions and contributing chromophores in proteins

#### 4.5 Rationale for present work

Proline-rich peptide sequences are very important recognition elements that have a significant bias toward the polyproline type II (PPII) conformation. Each strand of collagen triplex with the typical [Pro-Hyp-Gly]<sub>n</sub> repeat unit adopts a PPII-like conformation. Interactions between side-chain functional groups and the effect of solvation by water-backbone interactions have been examined as determinants of the PPII conformational stability. To understand the factors that determine the *cis/trans* conformer ratio and the development of concepts for tuning this equilibrium is therefore important and Proline derivatives with a substituent in the  $\gamma$ -position (C4) have proven useful in both respects. Conformational studies have shown that the nature of the substituent at C4 and the absolute configuration at this center critically influence both the pyrrolidine ring pucker and the *cis/trans* conformer ratio of the amide bond.

Understanding the different factors that influence the PPII conformation is therefore an important goal. Herein we present conformational studies of polyproline derivatives with hydrogen-bond-donating substituent in the  $\beta$ -position(C3) on proline ring. In case of “*cis*” 4*S*-substituted prolines, the intramolecular hydrogen bond between the 4*S*-substituent and the amide carbonyl on the same proline unit is known to influence the PPII conformation.



**Figure 14:** The plausible picture of hydrogen bonding in 4-hydroxy substituted polyproline (a) & 3-hydroxy substituted polyproline (b)

A substituent on C-3 can also show such potential to form stable intramolecular six membered hydrogen bond (Figure 14). In this context, it would be interesting to study polyprolines derived from substituent at C-3, instead of the naturally occurring C-4 substituent.

*Trans* L-3-hydroxy proline was first isolated by Logan *et al*<sup>31</sup> from cattle Achilles tendon collagen in the year of 1962. This has been found to be 0.26% and is a major component of the free amino acid pool of seeds and vegetative tissues of the tropical legume, *Delonix regia*. 3*S*-hydroxyproline is a major constituent of the soluble nitrogen fraction of cotyledons and also of roots, hypocotyls and leaves. It has the ability to inhibit the growth of mung bean seedlings. 10% of the total hydroxyproline in type IV collagens of basement membranes is 3*S*-hydroxyproline.

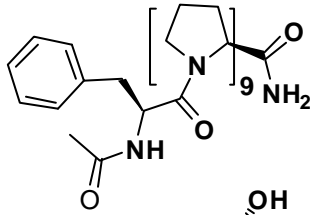
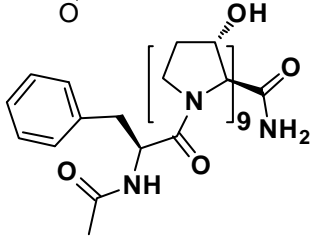
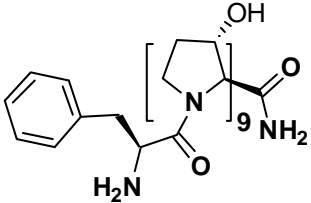
Fibril forming collagens of normal human and bovine tissues consists 3-hydroxyproline residues. Genetic defects inhibiting the formation of 3-hydroxyproline causes recessive osteogenesis imperfect.

#### 4.6 Aim of the current work

The specific objectives of this section are:

- i) Solid Phase synthesis of homo oligomeric peptides **1-3** (Table 2). Peptide **1** is synthesized using N-fmoc proline, whereas peptides **2&3** were synthesized using N-fmoc L-3-*trans* hydroxy proline.

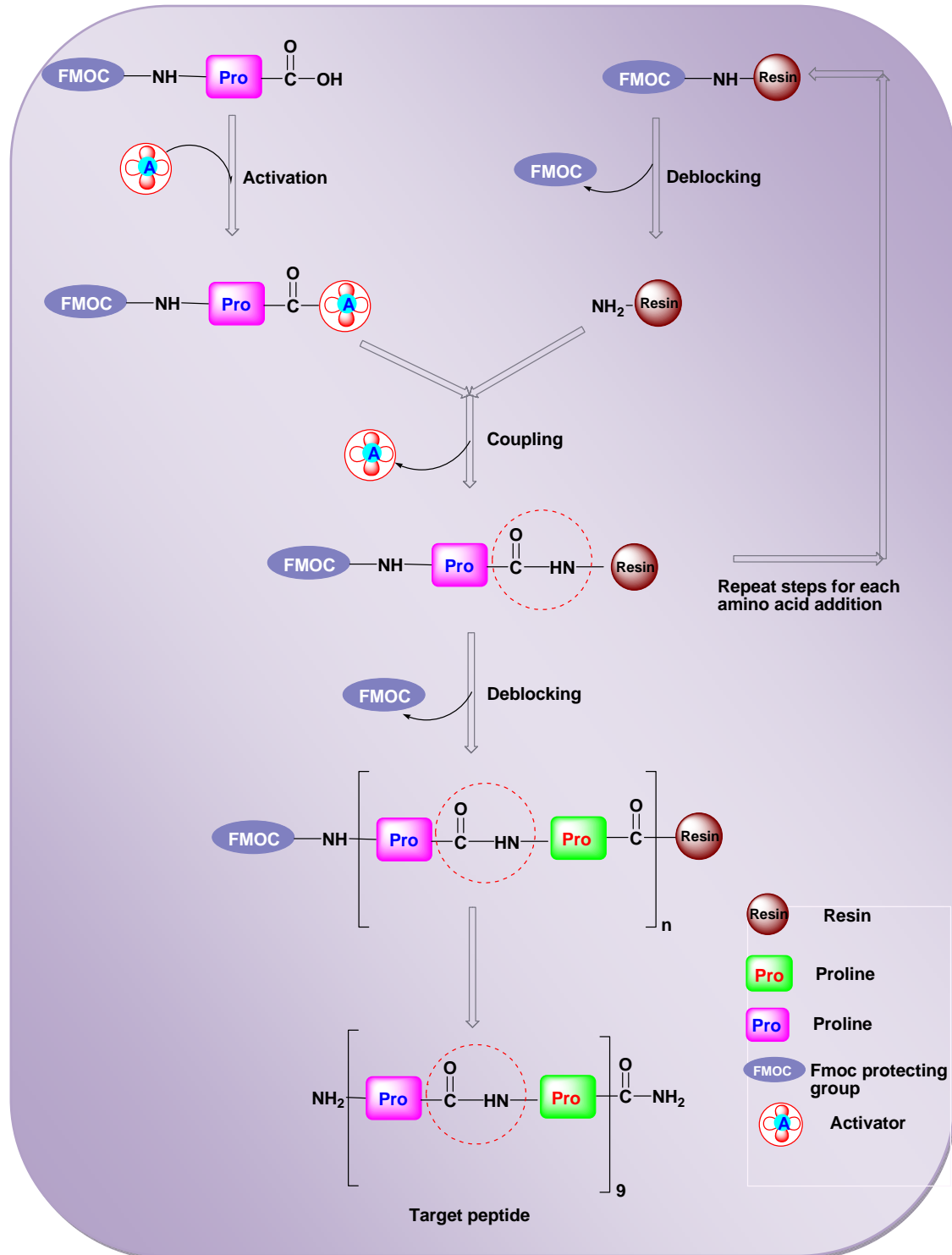
**Table 2:** Designed polyproline derivatives

Entry no.	Sequence
Peptide 1	
Peptide 2	
Peptide 3	

ii) Cleavage of the oligomers from the solid support and iii) purification and characterization iv) conformational studies of these peptides as a function of pH, urea, and solvents (buffer and trifluoroethanol, TFE).

#### 4.7 Protocols followed for solid phase peptide synthesis

Solid phase peptide synthesis (SPPS) protocols can be easily applied to the synthesis of several peptides and substituted polyprolines. The ease of handling and convenient procedures have made possible the synthesis of polyprolines and its analogues to examine the conformational behaviour. In this method, the *C*-terminal of proline/*3S*-hydroxyproline is linked to an insoluble matrix such as polystyrene beads having reactive functional groups. The next *N* $\alpha$ -protected proline/*3S*-hydroxyproline is coupled to the resin bound amino acid by using an active *N,N,N',N'*-Tetramethyl-*O*-(1*H*-benzotriazol-1-yl) (HBTU) ester. The excess amino acid is washed out and the deprotection and coupling reactions are repeated until the desired peptide is achieved. Fmoc strategy was followed for synthesizing the peptides **1-3** and achieved them successfully.

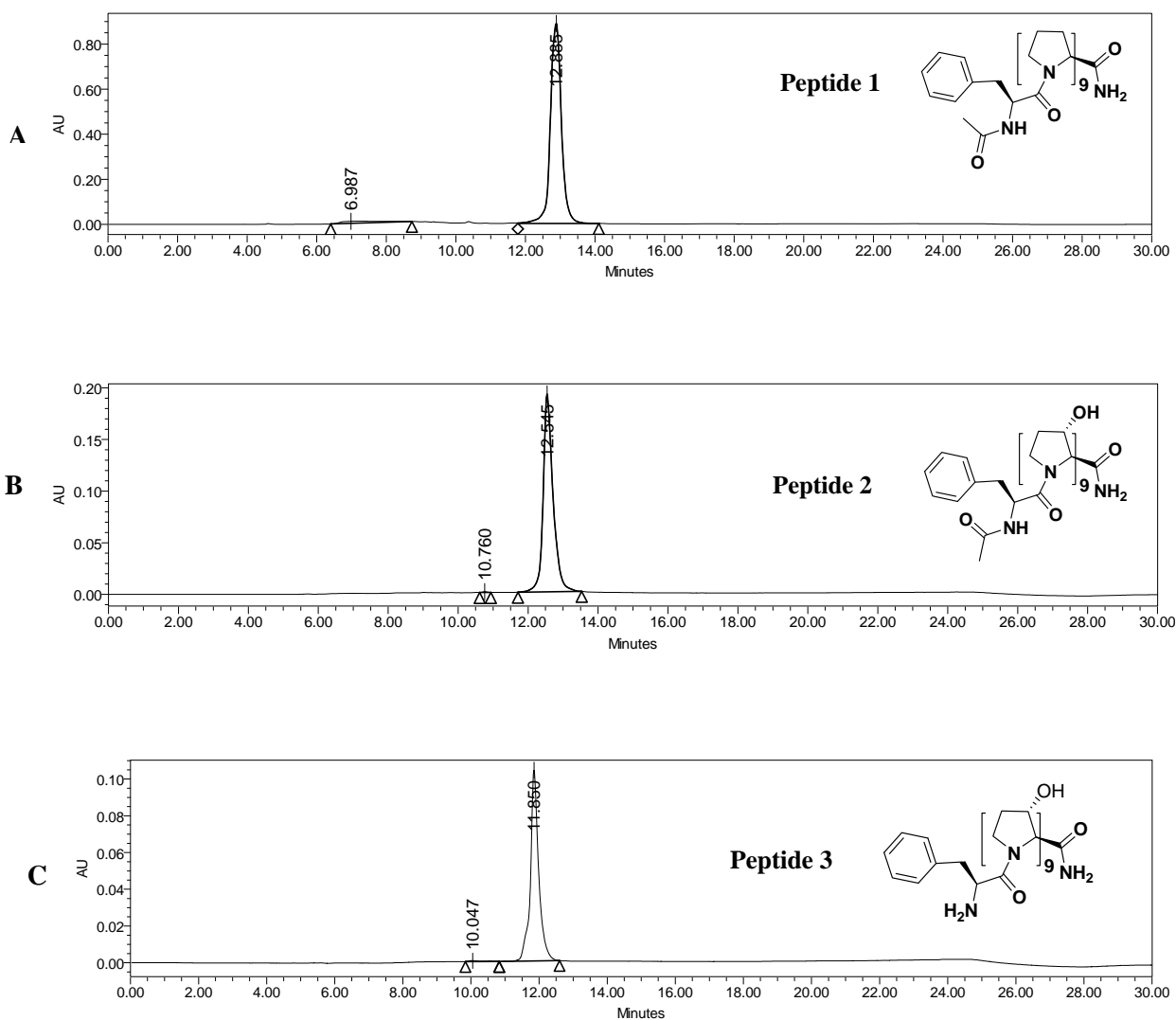


**Figure 15:** Schematic representation of solid phase peptide synthesis

## 4.8 Results and discussions

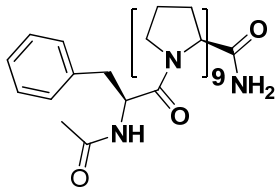
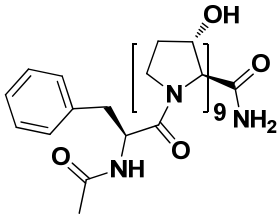
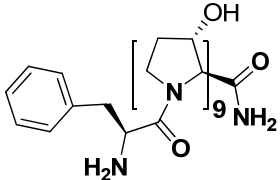
The oligopeptide **1** (Table 2) is synthesized from homo oligomerization of N-fmoc proline whereas peptides **2** and **3** were synthesized from N-fmoc *L-trans* 3*S*-hydroxy proline assembled in the solid phase. These peptides were purified by HPLC, and characterized by MALDI-TOF. The CD spectral analyses of these peptides were carried out as a function of pH, urea, and solvents (buffer and trifluoroethanol, TFE).

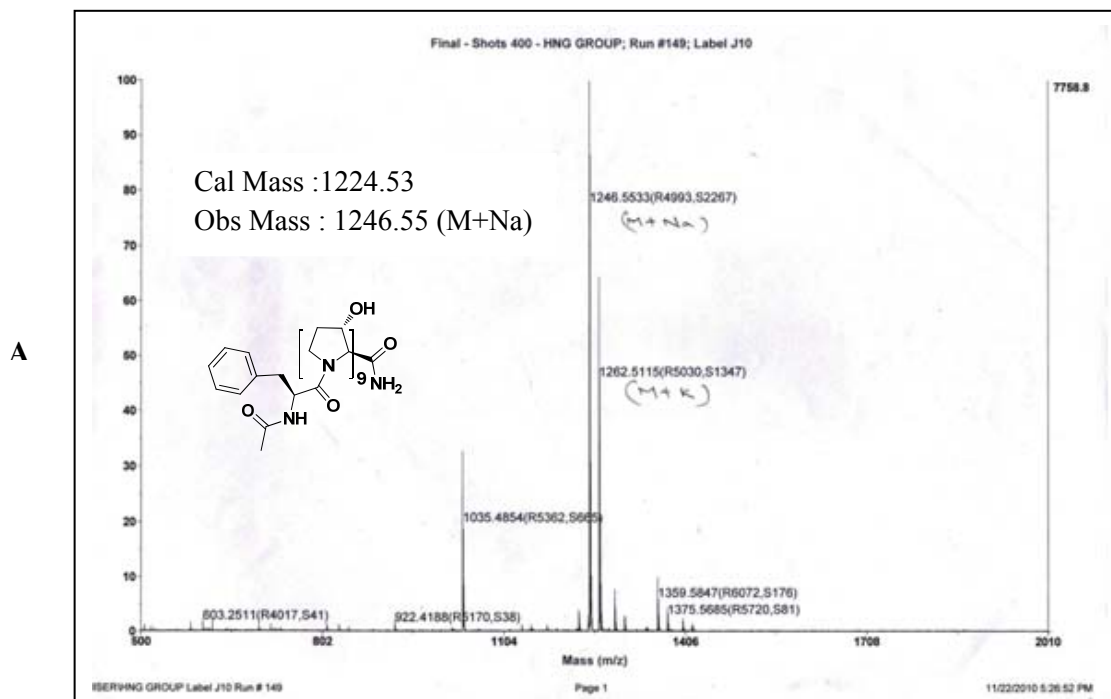
### HPLC data for peptide sequences



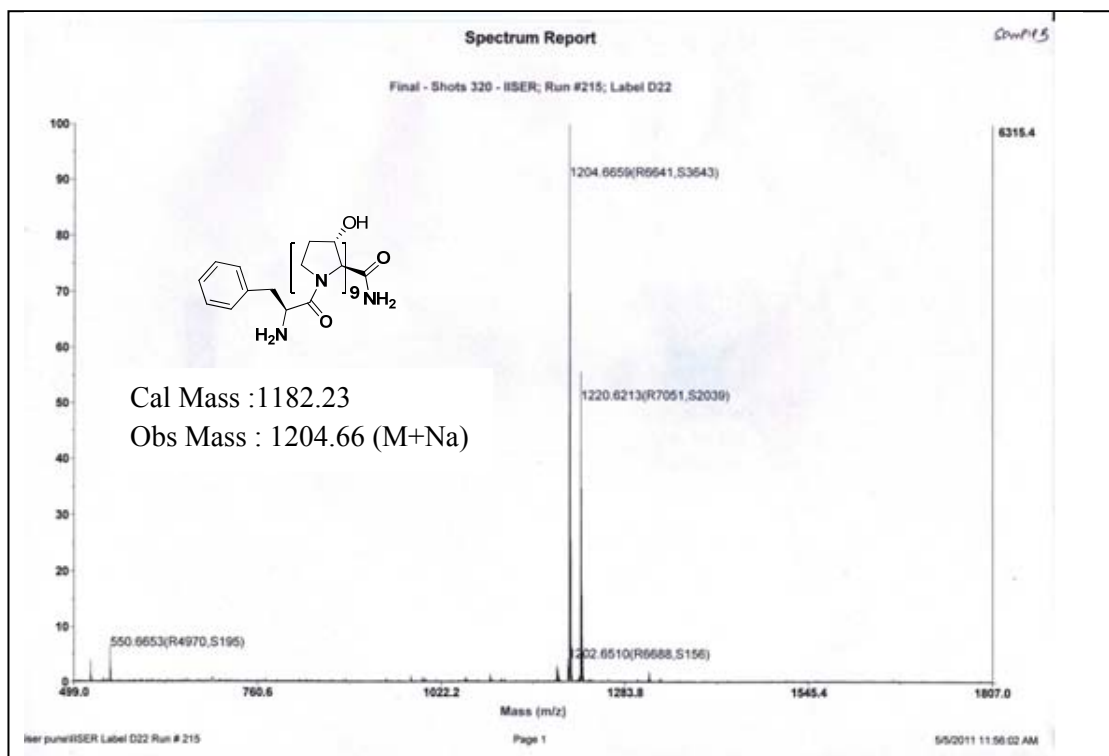
**Figure 16 :** (A) Reverse phase HPLC profile of peptide **1** (B) Reverse phase HPLC profile of peptide **2** (C) Reverse phase HPLC profile of peptide **3**

**Table 3:** Synthesized polyproline derivatives for conformational studies

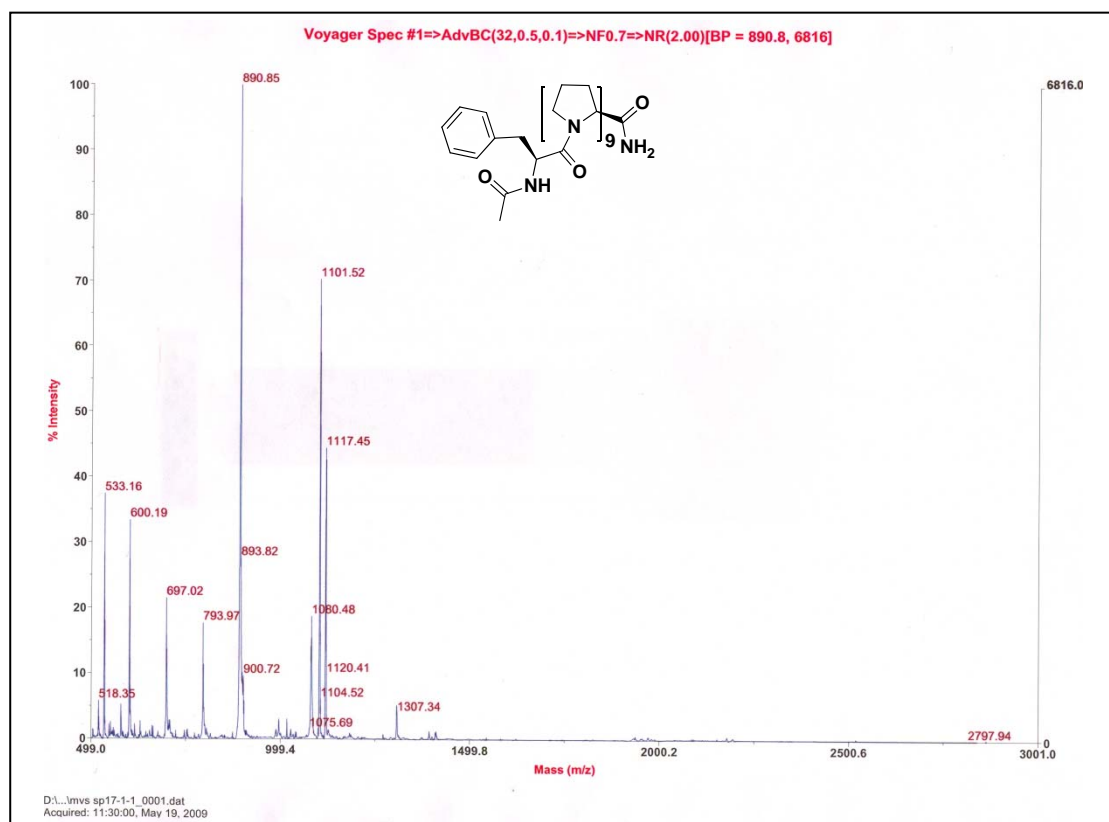
Entry no.	Sequence	Calculated mass	Observed mass
Peptide 1		1080.73	1101.52 (M+Na)
Peptide 2		1224.53	1246.55 (M+Na)
Peptide 3		1182.23	1204.66 (M+Na)

**MALDI TOF-spectras for peptide sequences**

B



C



**Figure 17** : (A) MALDI-TOF spectra of peptide 2 (B) MALDI-TOF spectra of peptide 3 (C) MALDI-TOF spectra of peptide 1

### 4.8.1 Conformational examination of synthesized polyprolines (1-3) by CD spectroscopy

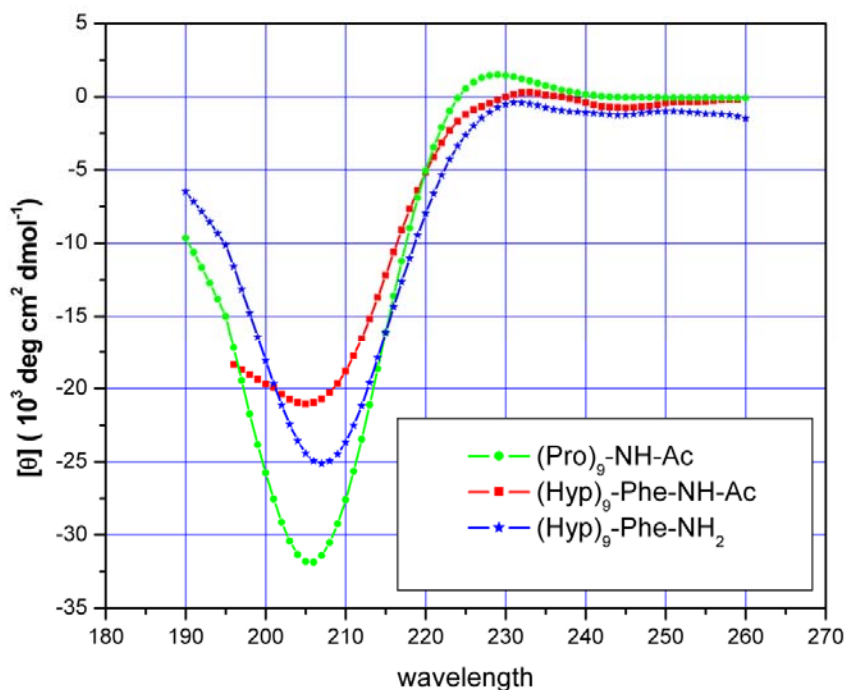
Poly-L-proline is known to undergo a reversible cooperative conformational transition from a righthanded helix, designated form PPI, to a left-handed helix of different geometry, designated form PPII, in certain solvent systems. A detailed analysis of the factors determining the conformation of this synthetic polymer is not only of intrinsic physical-chemical interest but may also increase the understanding of the molecular basis of conformation in biological macromolecules. This is especially true because low molecular weight polyproline, like some proteins, it can exist only in two conformational states.

It has been well explained that the stereoelectronic effects and substitution at 4-position on proline ring plays a major role in determining the puckering of the pyrrolidine ring of proline and hence the conformational stability of polyproline. Conformational studies of polyproline with 3-substitution on proline ring have not been reported till date. In this context synthesis of 3-L-*trans*-hydroxy proline polypeptides and its conformational behavior has been examined by CD spectral analyses as a function of pH, urea, and different solvents (buffer and trifluoroethanol).

#### 4.8.1a Conformational studies of polyproline in Phosphate buffer pH=7.2

Polyproline, (Pro)<sub>9</sub>-Phe-NH-Ac and its 3*S*-hydroxy substituted derivative (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> were used to probe the stereoelectronic effect on the conformation of polyproline where Pro is simple proline and 3*S*-Hyp is (2*S*,3*S*)-3-hydroxyproline. To understand the influence of the terminal functional group on the conformation of polyproline N-terminus of (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> (peptide **3**), is capped to get peptide (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac (peptide **2**). CD spectra of all the three synthesized peptides *i.e.* (Pro)<sub>9</sub>-Phe-NH-Ac (peptide **1**), (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> (peptide **3**) and (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac (peptide **2**) were recorded in aqueous phosphate buffer of pH 7.2 at a concentration of 100  $\mu$ M. The results are shown in Figure **18**. It is seen that CD spectra of all peptides are similar and are showing minima around at 205 nm and maxima at 225-235 nm typical for the PPII helix. The intensity of the positive maximum between 225-235 nm is observed in the decreasing order of (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac and (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> (Figure **18**) and so the content of polyproline II helix.



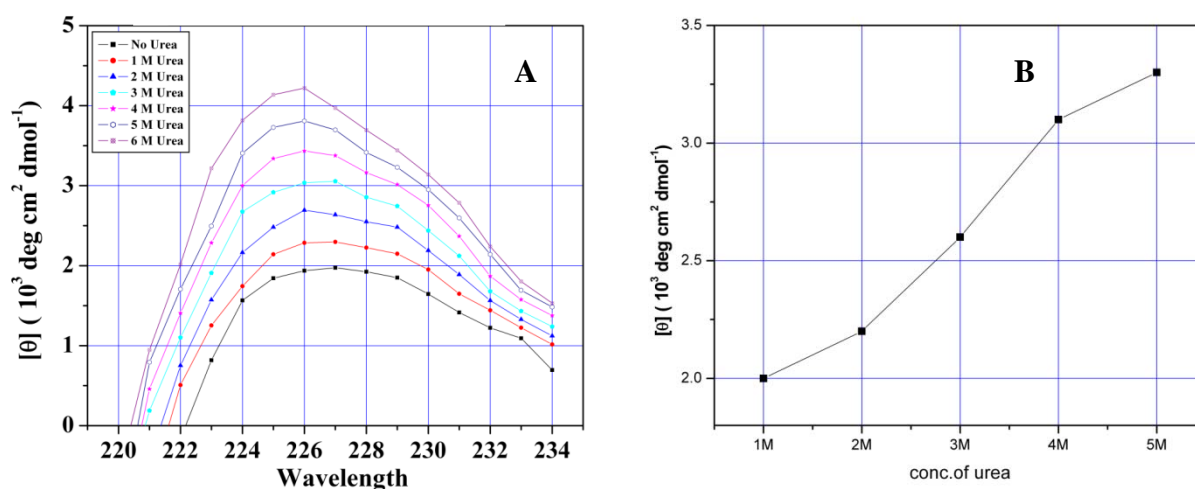


**Figure 18:** CD spectra of peptides **1-3** in phosphate buffer ( $P^H=7.2$ )

This suggests that substitution of *trans* 3*S*-hydroxy group on proline ring decreases the content of PPII. While N-termini capping of 3*S-trans*-L-hydroxy proline resulted in higher content of PPII when compared to its charged termini. The more negative intensity band at 200-210 nm for (Pro)<sub>9</sub>-Phe-NH-Ac indicated a higher random nature than (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac and (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> and is seen to decrease in the order of (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> and (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac. Overall incorporation of the *trans* hydroxyl group on C3 carbon of proline decreased the content of PPII helix in phosphate buffer (pH=7.2).

#### 4.8.1b Effect of urea on polyproline conformation

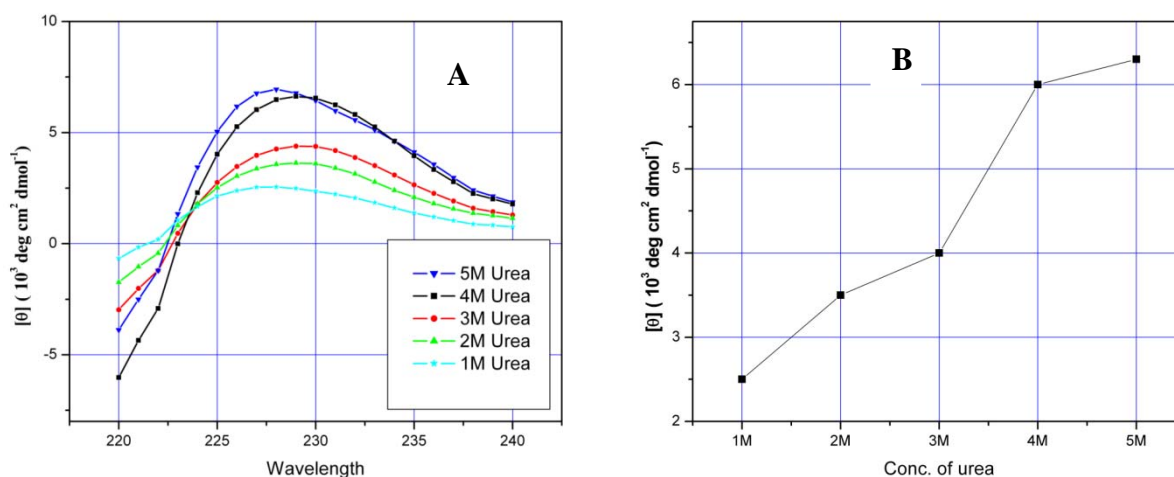
It is seen that proline homopolymers and its derived 3*S-trans*-L-hydroxy proline oligomer formed PPII helices in aqueous solution and possess CD spectra diagnostic for this conformation. It has been reported in the literature<sup>45</sup> that the PPII content of proline polymers and oligopeptides is significantly increased in the presence of urea as which can be confirmed by an increase in the intensity of the positive band in CD spectra.



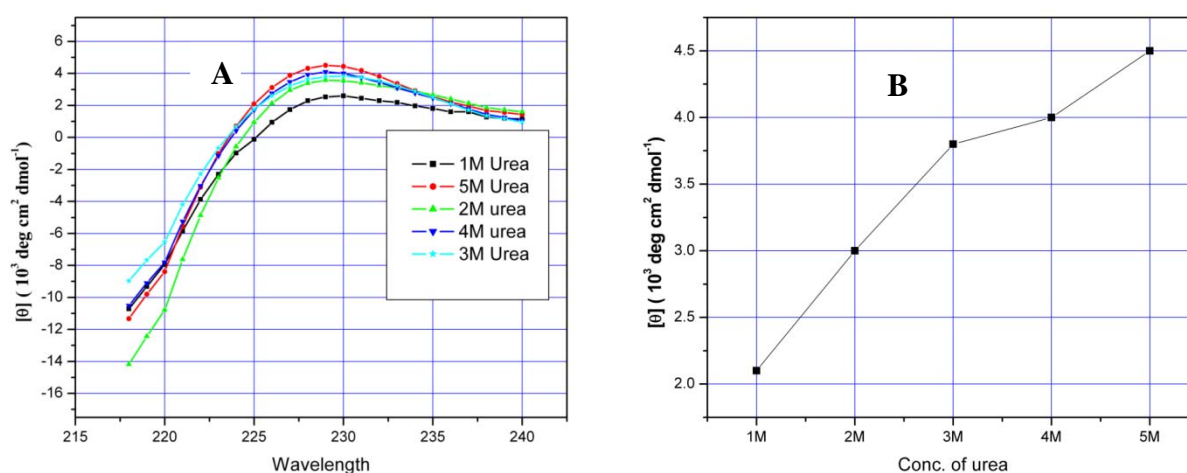
**Figure 19:** (A) CD spectrum of (pro)<sub>9</sub>-Phe-NH-Ac with different conc. of urea (B) Diagrammatic representation of effect of urea on (pro)<sub>9</sub>-Phe-NH-Ac conformation

The synthesized peptides listed in Table 2 were examined to answer the question of what happens to the PPII content of polyproline derived from *trans*-3*S*-hydroxy-L-proline residues in the presence of different concentrations of urea. The CD spectral data were recorded for these peptides by varying the concentration of urea ranging from 1M to 6M in phosphate buffer (pH=7.2) and the results are shown in Figure 19, 20 and 21. All three peptides (Table 2) possess significant PPII helix signals in CD spectra and the ellipticity around 225-235 nm increases with an increasing concentrations of urea.

Increasing ellipticities and the appearance of local maxima in the wavelength range of the positive band diagnostic of PPII helix, indicate that PPII helical content is increased as the urea concentration is increased. The CD data collected for the three peptides are in accordance with previous studies. The mechanism by which urea promotes PPII helical structure is not clear. Nozaki, Tanford<sup>43</sup> and Robinson, Jencks<sup>44</sup> proposed that urea interacts favorably with the polypeptide backbone; such an interaction would influence the conformational behavior of the backbone, by favoring the PPII helical conformation. This negates the idea that the backbone is completely disordered in the presence of high concentrations of urea. It is observed that urea induces PPII helical structure by making the backbone more ordered at higher concentrations.



**Figure 20:** (A) CD spectrum of  $(3S\text{-Hyp})_9\text{-Phe-NH-Ac}$  in different conc. of urea (B) Diagrammatic representation of effect of urea on  $(3S\text{-Hyp})_9\text{-Phe-NH-Ac}$  conformation



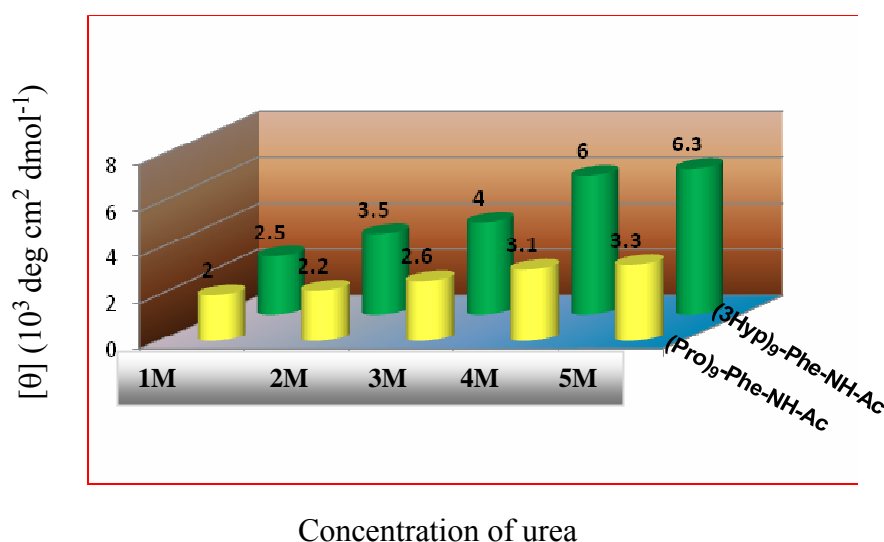
**Figure 2:** (A) CD spectrum of  $(3S\text{-Hyp})_9\text{-Phe-NH}_2$  in different conc. of urea (B) Diagrammatic representation of effect of urea on  $(3S\text{-Hyp})_9\text{-Phe-NH}_2$  conformation

Urea rigidifies the backbone by interacting with the amide bonds and sterically restricts the peptides to form PPII helices. It is observed that  $\beta$ -substitution i.e. C3 substitution by hydroxyl group on polyproline showed (increase in the PPII content upon increasing the concentration of urea) similarly to that of naturally occurring proline. It is observed

that PPII content is more for *trans*-3*S*-hydroxy substituted polyproline when compared to simple polyproline in the presence of urea and is described below *i.e.* in section 4.8.1c

#### 4.8.1c Effect of urea on PPII content

CD spectra of the peptides (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac and (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> showed an increase in PPII content with increasing concentrations of urea. To understand the difference in the PPII content for these peptides at particular concentration of urea comparison studies have done. Firstly the difference in the PPII content is high for (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac when compared to (Pro)<sub>9</sub>-Phe-NH-Ac at any given concentration of urea as shown in Figure 22. These results show that  $\beta$ -substitution by *trans* hydroxyl group increases the PPII content. The difference in the PPII content is more with urea concentration 3M, 4M and 5M.

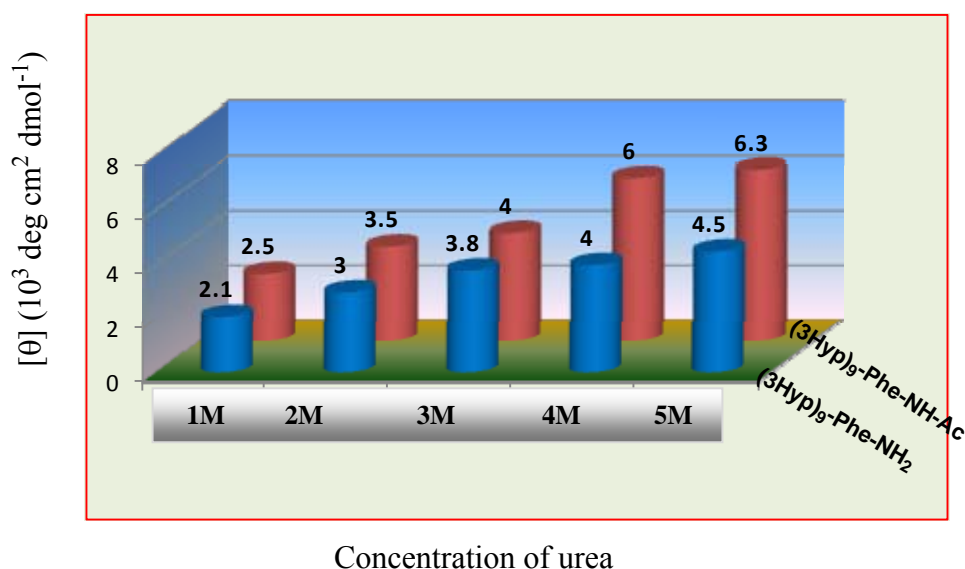


**Figure 22:** Effect of urea on PPII content in peptides (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac.

The difference in the PPII content at 5M is more and is approximately by 3 [θ] (10<sup>3</sup> deg cm<sup>2</sup> dmol<sup>-1</sup>). At 4M and 3M these peptides have difference in PPII content by 2.9 and 1.4 [θ] (10<sup>3</sup> deg cm<sup>2</sup> dmol<sup>-1</sup>) respectively. Least difference by 0.5 [θ] (10<sup>3</sup> deg cm<sup>2</sup> dmol<sup>-1</sup>) is observed for 1M urea concentration.

To know the effect of terminal functional group on PPII content with charged and capped N-termini and the relative ease with which they switch the PPII content in different concentrations urea for peptides (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> (peptide 3), (3*S*-Hyp)<sub>9</sub>-

Phe-NH-Ac (peptide **2**) are shown in Figure **23**. It is observed that N-termini capped peptide showed a considerable increase in PPII content when compared to its charged termini at a given concentration of urea. The difference in the PPII content is more at 4M and 5M concentrations of urea. The difference in the value of  $[\theta]$  ( $10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$ ) is 2, 2.8 for 4M and 5M respectively. A minimal difference in PPII content is observed in 1M-3M urea. The difference in the value of  $[\theta]$  ( $10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$ ) is minimal by approximately 0.2 for 3M urea. These difference in the results may be attributed due to the difference in charge dipole interaction for capped and charged termini.



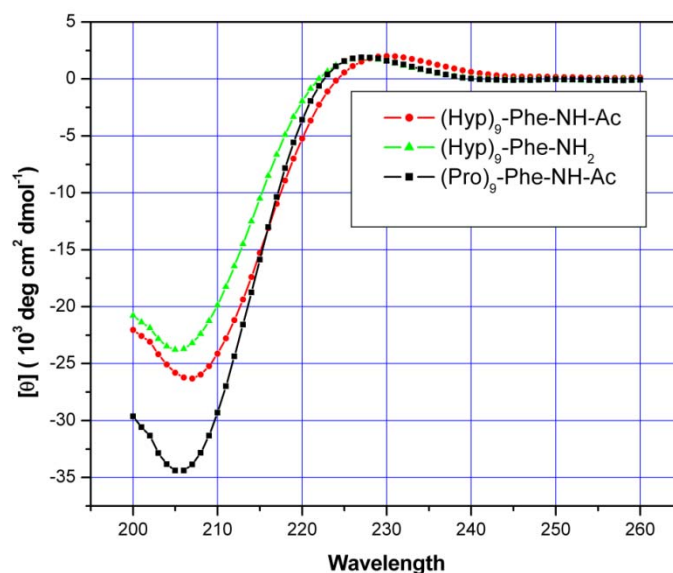
**Figure 23:** Effect of urea on PPII content in peptides (3S-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> (peptide **3**), (3S-Hyp)<sub>9</sub>-Phe-NH-Ac (peptide **2**).

In conclusion peptides **1-3** (Table **2**) showed an increase in PPII content with increase in concentration of urea. The intensity of the PPII content decreases in the order (3S-Hyp)<sub>9</sub>-Phe-NH-Ac, (3S-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> and (Pro)<sub>9</sub>-Phe-NH-Ac). In the presence of urea these experimental findings clearly suggest that  $\beta$ -*trans* hydroxyl substitution on proline ring have positive impact on the PPII content with both N-termini capped and charged. N-termini capped *trans*-L-3-hydroxy polyproline showed higher PPII content when compared to its charged termini due to change in charge dipolar interaction.

#### 4.8.1d Conformational studies of polyproline in 2,2,2 trifluoroethanol (TFE)

The conformational stability of the polyproline II (PPII) helix with respect to the functional groups at the N-termini was examined. Here, a detailed analysis of the influence of charged and neutral functional groups at the N-termini of *trans*-3*S*-hydroxy-L-proline oligomers on the conformational stability of the PPII in trifluoroethanol has done.

CD spectra of all the three peptides were recorded in 2,2,2 trifluoroethanol. The results are shown in Figure 24. Peptides 1-3 (Table 2) exhibit CD spectra with a positive band between 220 and 230 nm and a negative band between 200 and 210 nm (Fig. 24), which are hallmarks of the PPII conformation. The intensity of the positive band of all these peptides i.e. 1-3 (Table 2) at 225-227 nm is almost equal (Figure 24). The more

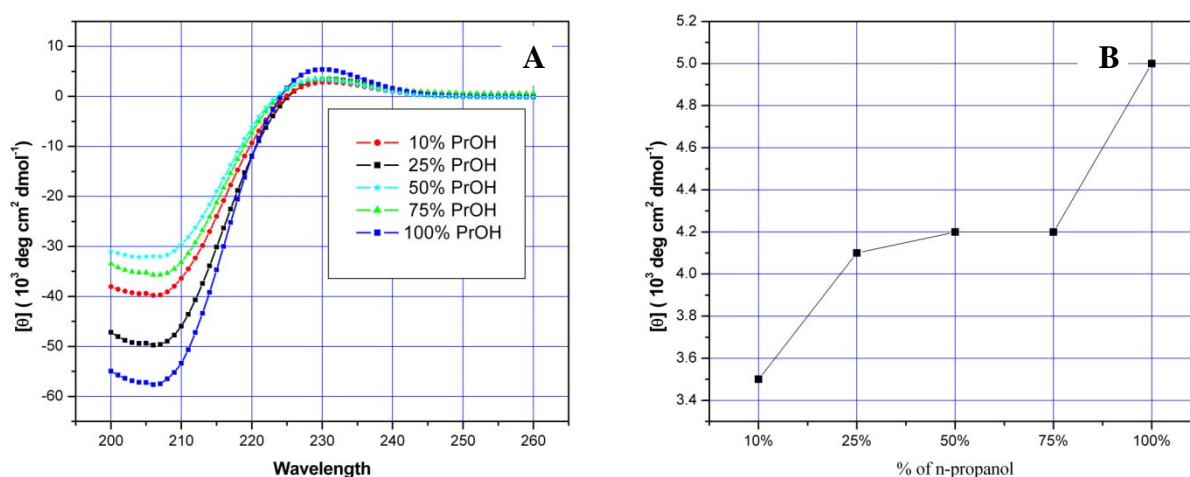


**Figure 24:** CD spectral studies of peptides 1-3 in TFE

intense negative maximum between 200-210 nm in CD spectra by peptide (Pro)<sub>9</sub>-Phe-NH-Ac indicates that this peptide has higher random structure than do (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac and (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub>. The intensity of the negative band at 200-210nm is proportional to the random nature which is seen to decrease in the order of (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac and (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> (Figure 24).

#### 4.8.1e Circular dichroism studies of polyproline in propanol

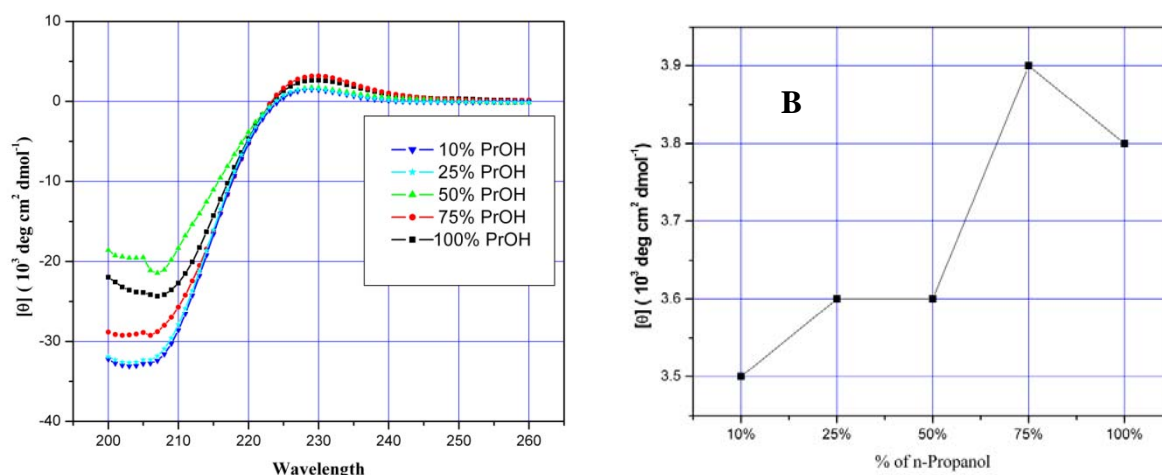
In an effort to understand the effect of n-propanol on the conformational stability of the PPII conformation, the conformational behavior of oligoproline **1-3** in different percentages of n-propanol were studied. CD spectra of these peptides were recorded in different percentages of n-propanol in phosphate buffer (pH=7.2) after incubating the samples at 32°C for 7 days. The CD spectra exhibited a positive band between 225 and 235 nm and a negative band between 200 and 210 nm indicating that



**Figure 25:** (A) CD spectra of (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac in n-Propanol (B) Diagrammatic representation of effect of n-propanol on (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac

all peptides form PPII helices. The PPII helix is the only secondary structure with a positive band in this region. so we use this as a measure of the conformation.

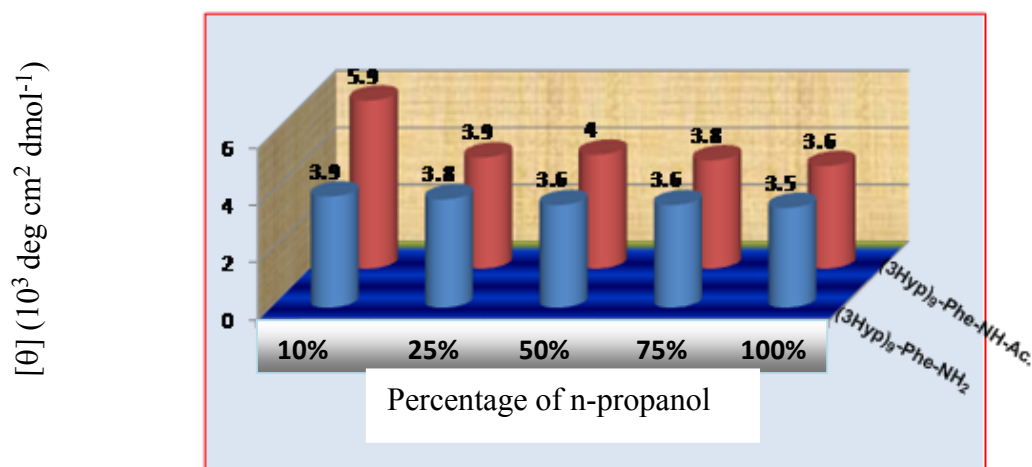
In n-propanol, a solvent that is known to favor the PPI helix, peptides (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac and (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> in contrast exhibited PPII type conformation. N-terminus capped analogue (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac showed an increase in PPII content with increase in the concentration of n-Propanol (Figure 25) whereas its charged termini peptide i.e. (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> showed increase in PPII content by increasing the percentage of n-Propanol from 10-50% and showed less PPII character in 100% Propanol when compared to 75% propanol (Figure 26) in phosphate buffer. Overall peptides **2-3** (Table 2) showed an increase in PPII content with increasing percentage of n-propanol.



**Figure 26:** (A) CD spectra of (3S-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> in n-Propanol (B) Diagrammatic representation of effect of n-propanol on (3S-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> conformation

#### 4.8.1f Effect of n-propanol on PPII content

A detailed analysis of the influence of charged and capped N- termini of *trans*-3S-hydroxy-L-proline homo oligomers on PPII content and the conformational behavior of oligoprolines in different percentages of n-propanol has been done. CD spectra of these peptides were recorded in different percentages of n-propanol in phosphate buffer (pH=7.2) after incubating the samples at 32°C for 7 days and the results indicated that all peptides form PPII helices. The PPII stability is increased by increasing the percentage of n-propanol (Figure 27).



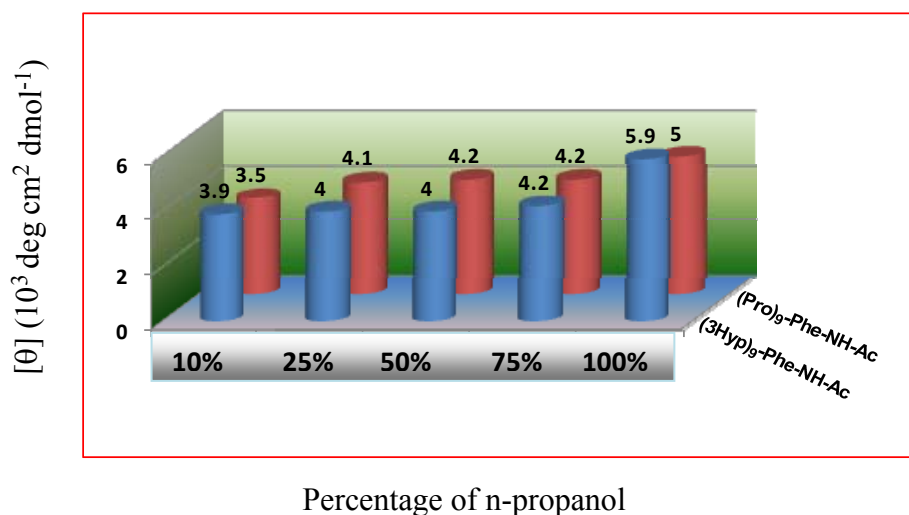
**Figure 27:** Effect of n-propanol on PPII content in peptides (3S-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub>, (3S-Hyp)<sub>9</sub>-Phe-NH-Ac.



It has been observed that the PPII content for (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac is higher as compared to (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> at a given percentage of n-propanol as shown in Figure 27.

In 10% n-propanol these two peptides showed a considerable difference in PPII content with the difference in the value of  $[\theta]$  ( $10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$ ) by 2. In 25% n-propanol and in pure n-propanol these two peptides showed minimal difference in the PPII content and the difference in the value of  $[\theta]$  ( $10^3 \text{ deg cm}^2 \text{ dmol}^{-1}$ ) is by 2.

To understand the effect of *trans* 3*S*-hydroxyl group on PPII content in n-Propanol, comparison studies has done for peptides (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac. No change or minimal change in PPII content is observed for these peptides when percentage of n-propanol is varied from 25%-75%. A considerable increase in PPII content is observed for both the peptides in 100% n-propanol (Figure 25 and 26 respectively). The PPII content for (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac is more than (Pro)<sub>9</sub>-Phe-NH-Ac in 100% n-propanol indicating that substitution at  $\beta$ -position by *trans* hydroxyl group on proline ring shows an increase in PPII content in n-propanol.



**Figure 28:** Effect of n-propanol on PPII content in peptides (Pro)<sub>9</sub>-Phe-NH-Ac, (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac.

In conclusion, N-termini capped *trans*-L-3*S*-hydroxy proline i.e. (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac showed higher value of ellipticity *i.e.* PPII content when compared to its charged analogue (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> in both n-propanol and urea at a given

concentration. All these results can be due to the change in charge dipole interaction which is arised due to N-termini capping.

## 4.9 Conclusion

In conclusion N-fmoc 3*S*-hydroxy proline was oligomerized on solid phase peptide synthesis using fmoc strategy to get the required peptides (Peptides **2-3**). Conformational behaviour of these polyproline oligomers in various solvents, different percentages of n-propanol and the effect of urea were examined. The CD studies of these peptides in different solvents confirmed the presence of PPII conformation and the content of PPII is different in different solvents. With increase in the concentration of urea increase in the ellipticities is observed in the wavelength range of the positive band which is diagnostic of PPII helix content. This indicates that these peptides gain more PPII helix content as the urea concentration is increased. CD spectroscopic studies of oligoprolines in different ratios of phosphate buffer in n-propanol show that the PPII helix with all *trans* amide bonds is stabilized. All the three peptides **1-3** acquired PPII conformation in different solvents and comparative studies of PPII content of peptides were done. N-termini capped *trans*-L-3*S*-hydroxy proline (3*S*-Hyp)<sub>9</sub>-Phe-NH-Ac showed higher content of PPII when compared to its charged analogue (3*S*-Hyp)<sub>9</sub>-Phe-NH<sub>2</sub> in n-propanol and urea at any given concentration. All these results can be due to the change in charge dipole interaction which is arised due to N-termini capping.

Furthermore, A new family of  $\beta$  substituted polyprolines were synthesized which can adopt biologically important PPII conformation in different solvents. Further work is required to tune the conformation of polyproline and to know the different factors affecting the conformational behaviour of the peptides.

## 4.10 Experimental section

### 4.10.1 Synthesis of polyproline oligomers

The synthesis of polyproline oligomers were done by following solid phase peptide synthesis protocols by incorporating the monomers to yield the required length of peptide. Synthesis was done from the *C*-terminus to the *N*-terminus using monomeric units with N-fmoc protected amino and carboxylic acid functions. All the peptides were synthesized using fmoc in which rink amide resin was used as solid support on which the oligomers were built and the monomers were coupled by *in situ*

activation with HBTU / HOBt. In the synthesis of all these oligomers, Phenyl alanine or N-Acetyl phenylalanine was used as *N*-terminal spacer-amino acid. Loading value of Rink amide resin is 0.6m.mol/gm and is directly used for the synthesis of required oligomers without functionalizing. The polyproline oligomers were synthesized using repetitive cycles (Figure 15), each comprising the following steps:

- i) deprotection of the *N*-fmoc group using 20% piperdine in DMF to get free amine
- ii) washing the reaction vessel with 3xDMF followed by 3xDCM and again by 3x DMF
- iii) Coupling of the free amine with the free carboxylic acid group of the incoming monomer (amino acid) using 3 to 4 equivalents. The coupling reaction was carried out in DMF/NMP with HBTU as coupling reagent in the presence of DIPEA and HOBt.

The deprotection of the *N*-fmoc protecting group and the coupling reactions were monitored by Chloranil test. The *N*-fmoc deprotection step shows a positive Chloranil test i.e. where the resin beads as well as the solution turns to blue color. On the other hand, upon completion of the coupling reaction, the Chloranil test shows negative test, i.e. resin beads remains colorless. If reaction does not go to completion capping of the unreacted amino groups can be done using acetic anhydride in pyridine: DCM. Two solutions are made for chloranil test and the procedure is as follows:

- i) 2% acetaldehyde solution in DMF
- ii) 2% Chloranil solution in DMF

Few resin beads to be tested were taken in a test tube and washed several times with ethanol 3-4 drops of 2% acetaldehyde solution in DMF was added followed by 2% solution of Chloranil in DMF and the test tube was kept at ambient temperature for 5-10 minutes and a positive test is indicated by blue resin beads which indicate successful deprotection while colourless beads indicate completion of coupling step.

#### 4.10.2 Cleavage of the polyproline oligomers from the resin

The Rink amide resin (10 mg) with oligomers attached to it was stirred with 20% Trifluoroacetic acid (1ml) in DCM and tri isopropyl silyl chloride (5 $\mu$ L) in an ice bath for 30 min.. The resin was removed by filtration under reduced pressure and washed twice with TFA. The filtrate was combined and TFA was evaporated under vacuum at room temperature then the product was precipitated with cold dry ether. The

peptide was isolated by centrifugation and the precipitate was dissolved in de-ionized water (200  $\mu$ L) and loaded over sephadex G25 column. Fractions of 0.5 mL were collected and the presence of oligomers was detected by measuring the absorbance at 260 nm. The fractions containing oligomers were freeze-dried and the purity of the fractions was assessed by analytical RP-HPLC. If the purity is less than 90%, oligomers were purified by preparative HPLC.

#### 4.10.3 Post-cleavage workup by ether precipitation

Peptide isolation and work-up can be achieved by ether precipitation (step 1) or centrifugation (step 2). For water soluble peptides, work up can be done by following the steps 3-6.

1. **Precipitation:** Filter the precipitated peptide through hardened filter paper in a Hirsch funnel under a light vacuum. Wash the precipitate further with cold ether, dissolve the peptide in a suitable aqueous buffer and lyophilize.
2. **Centrifugation:** Add a small volume of diethyl ether to the residue and triturate thoroughly until a free suspension is obtained. Transfer the suspension to a clean centrifuge tube, seal and centrifuge. It is essential that a spark-free centrifuge is used for this process. Carefully decant the ether from the tube. Repeat the ether wash as necessary. Dissolve the residual solid in a suitable aqueous buffer and lyophilize.
3. **Watersoluble peptides:** After precipitation, add water to the residue and transfer mixture to a separating funnel. A little AcOH may be necessary to aid dissolution.
4. Shake the stoppered funnel well. Release the stopper and allow the two layers to separate by standing. Isolate the lower (aqueous) layer.
5. Add more water to the funnel and repeat step 4 three times. Remove the upper (ethereal) layer and store in a clean flask. Return the combined aqueous extracts to the separating funnel.
6. Add a small amount of fresh diethyl ether and repeat step 4 two or three times, each time removing the ethereal layer and returning the aqueous layer to the separating funnel. Collect the aqueous layer in a clean flask and lyophilize

#### 4.10.4 HPLC (High Performance Liquid Chromatography) purification of polyproline oligomers

Peptide purifications were performed on Waters DELTAPAK-RP semi preparative C18 column attached to Hewlett Packard 1050 HPLC system equipped with

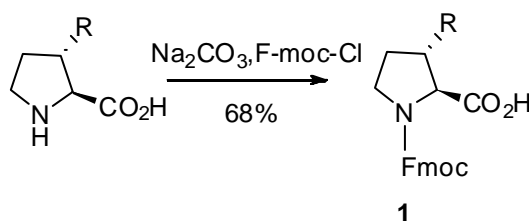
an auto sampler and Jasco-UV970 variable- wavelength detector. An isocratic elution method with 10% CH<sub>3</sub>CN in 0.1% TFA/ H<sub>2</sub>O was used with flow rate 1.5 mL/min (linear gradient from A to B in 20 min) and the eluent was monitored at 225nm. The purity of the oligomers was further assessed by RP-C18 analytical HPLC column (25×0.2 cm, 5 μm) with gradient elution: A to 100% B in 20 min; A= 0.1% TFA in CH<sub>3</sub>CN:H<sub>2</sub>O (5:95); B= 0.1% TFA in CH<sub>3</sub>CN:H<sub>2</sub>O (1:1) with flow rate 1 mL/min. The purities of the hence purified oligomers were found to be > 90%.

#### 4.10.5 MALDI-TOF Mass spectrometry

Literature reports the analysis of polyproline purity by MALDI-TOF mass spectrometry in which several matrices have been explored, *viz.* Sinapinic acid (3, 5-dimethoxy-4- hydroxycinnamic acid), CHCA ( $\alpha$ -cyano-4-hydroxycinnamic acid) and DHB (2,5-dihydroxybenzoic acid). Of these, CHCA was found to give the best signal to noise ratio with all other matrices typically producing higher molecular ion signals. For all the MALDI-TOF spectra recorded for the polyproline oligomers reported in this chapter, CHCA was used as the matrix on Voyager-De-STR (Applied Biosystems) instrument and was found to give satisfactory results.

#### 4.10.6 Synthesis of fmoc 3-hydroxy proline (1)

A solution of commercially available *trans*-3-hydroxy proline (1.33 g, 4.62 mmol) in water/acetone (40 mL, 1:1) was cooled to 0°C and sodium bicarbonate (1.36 g, 16.2 mmol) was added. To this mixture was added Fmoc-OSu (2.03 g, 6.0 mmol), and the reaction was stirred for 1 h at 0 °C and 2 h at 25 °C. The reaction was treated with 10% HCl to a pH of 4 and extracted with CH<sub>2</sub>Cl<sub>2</sub>. The organic layer was dried over anhydrous MgSO<sub>4</sub>, and the solvent was removed in vacuo to get N-Fmoc hydroxyl proline as white solid (yield 62%).



where R = H, OH

**<sup>1</sup>H NMR** (400MHz, CD<sub>3</sub>OD); 7.70-7.74 (t, 2H), 7.54-7.6(m,2H), 7.33-7.38(q, 2H), 7.29-7.26(m,2H), 4.47(s, 1H, -CH-Ar), 4.33-4.35(m,2H, -CH<sub>2</sub>-O), 4.23-4.25(m, 1H, -CH-OH), 3.66-3.72( m,2H, -CH<sub>2</sub>-N), 2.08-2.13 (m, 1H), 1.90-2.07(m, 1H); **<sup>13</sup>CNMR** (400MHz, CD<sub>3</sub>OD); 172.1, 155, 143, 140.6, 127, 126.4, 124.4, 119.2, 74, 73, 67.4, 46.4, 44.3, 31.7, 30.8.

#### 4.11 References

- 1) (a) Biochemistry by stryer, Jeremy M. Berg, John L. Tymoczko, fifth edition Page no 106 (b) Biochemistry by voet & voet 3<sup>rd</sup> edition Page no 130
- 2) Michael Kuemin, Sabine Schweizer, Christian Ochsenfeld, and Helma Wennemers; *J. Am. Chem. Soc.* **2009**, *131*, 15474–15482
- 3) (a) Bochicchio, B. & Tamburro, A. M. *Chirality*, **2002**, *14*, 782–792 (b) Lam, S. L. & Hsu, V. L. *Biopolymers*, **2003**, *69*, 270–281
- 4) Adzhubei AA, Sternberg MJE. Left-handed polyproline II helices commonly occur in globular proteins. *J. Mol. Biol.* **1993**, *229*, 472-493.
- 5) Shi ZS, Chen K, Liu ZG, Kallenbach NR; Conformation of the backbone in unfolded proteins. *Chem Rev*, **2006**, *106*, 1877–1897.
- 6) Kelly MA, et al. Host–guest study of left-handed polyproline ii helix formation. *Biochemistry*, **2001**, *40*,14376–1438
- 7) Rath A, Davidson AR, Deber CM ; The structure of “unstructured” regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. *Biopolymers*, **2005**, *80*, 179–185.
- 8) (a) Shoulders, M. D.; Raines, R. T. *Annu.Rev. Biochem.* **2009**, *78*, 929–958. (b) Brodsky, B.; Thiagarajan, G.; Madhan, B.; Kar, K. *Biopolymers* **2008**, *89*, 345–353. (c) Engel, J.; Bächinger, H. P. *Top. Curr. Chem.* **2005**, *247*, 7–33.
- 9) (a) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. ReV.* **2006**, *106*, 1877–1897. (b) Woody, R. W. *AdV. Biophys.Chem.* **1992**, *2*, 37–79.
- 10) (a) Dukor, R. K.; Keiderling, T. A. *Biopolymers* **1991**, *31*, 1747–1761. (b) Krimm, S.; Tiffany, M. L. *Isr. J. Chem.* **1974**, *12*, 189–200.
- 11) (a) Doose, S.; Neuweiler, H.; Barsch, H.; Sauer, M. *Proc. Natl. AcadSci. U.S.A.* **2007**, *104*, 17400–17405. (b) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Klumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci U.S.A.* **2005**, *102*, 2754–2759. (c) Stryer, L.; Haugland, R. P. *Proc Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719–726.

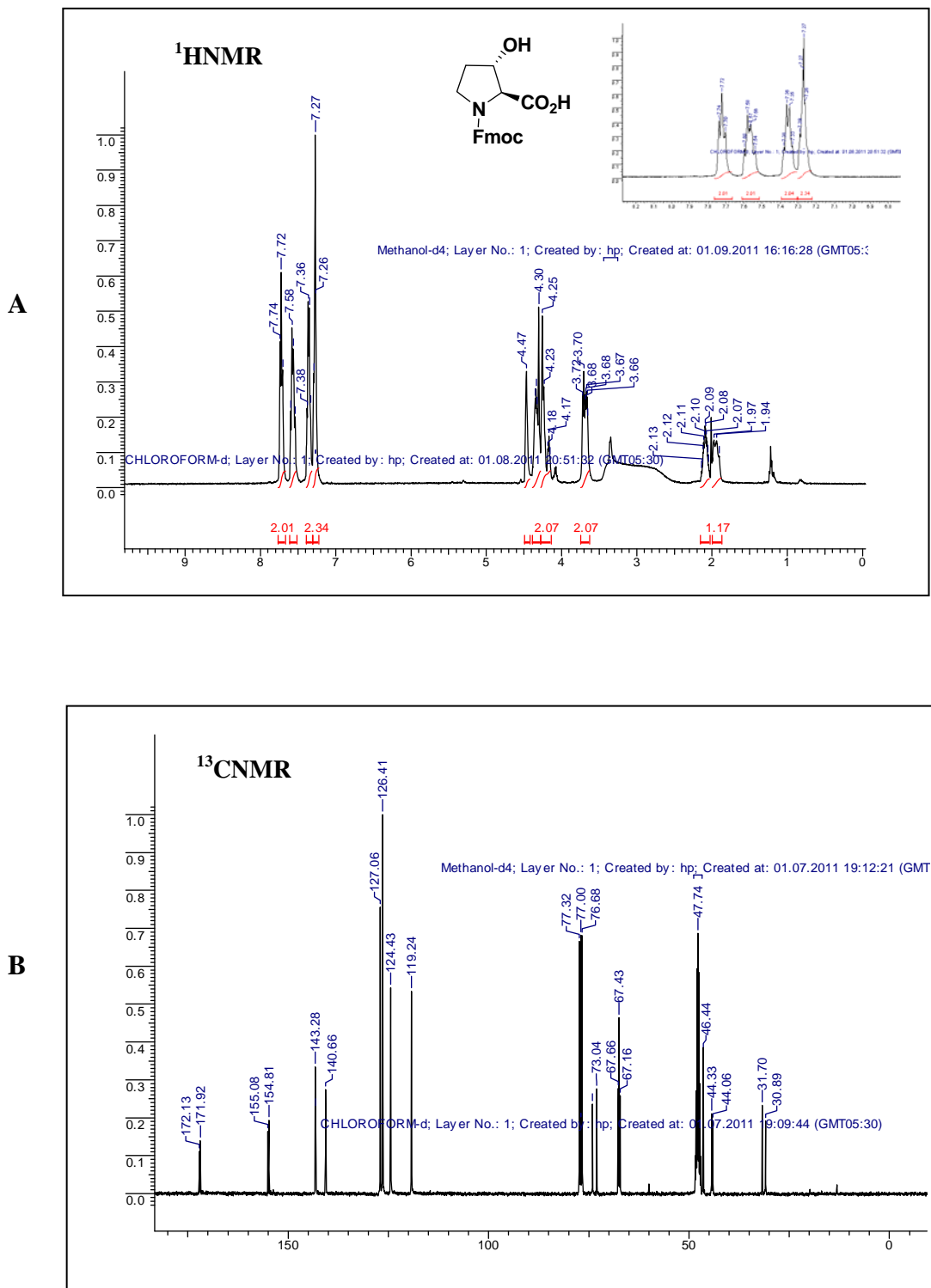
- 12) (a) Cordes, M.; Kottgen, A.; Jasper, C.; Jacques, O.; Boudebous, H.; Giese, B. *Angew. Chem., Int. Ed.* **2008**, *47*, 3461–3463. (b) Serron, S. A.; Iii, A. W. S.; Fleming, C. N.; Danell, R. M.; Baik, M.-H.; Sykora, M.; Dattelbaum, D. M.; Meyer, T. J. *J. Am. Chem. Soc.* **2004**, *126*, 14506–14514.
- 13) Williamson, M. P. The structure and function of proline-rich regions in proteins. *Biochem. J.* , **1994**, *297*, 249–260
- 14) Kakinoki S, Hirano Y, Oka M; On the stability of polyproline-I and II structures of proline oligopeptides. *Polym Bull*; **2005**, *53*, 109–115.
- 15) Knof S, Engel J; Conformational stability, partial specific volumes and spectroscopic properties of poly-L-proline poly-L-hydroxyproline and some of its O-acyl derivatives in various solvent systems. *Isr J Chem*, **1974**, *12*, 165–177.
- 16) Mutter M, Wohr T, Gioria S, Keller M; Pseudoprolines: induction of cis/trans-conformational interconversion by decreased transition state barriers. *Biopolymers*, **1999**, *51*, 121–128.
- 17) Ronald T. Raines; Jia-Cherng horng; *Protein science*, **2006**, *15*, 74-83
- 18) Dwight R. Robinson and William P. Jencks; *JACS*, **1965**, 2462-2470
- 19) Shelly J. Whittington, Brian W. Chellgren, Veronique M. Hermann, and Trevor P. Creamer; *Biochemistry*, **2005**, *44*, 6269-6275
- 20) Helma Wennemers *et al JACS*, **2009**, *131*, 15474-15482
- 21) Lu, k.p., Finn, g; Lee, T.H. and Nicholson L.K. *Nat Chem. Biol*, **2007**, *10*, 619-29
- 22) a) A. Choudhary, D. Gandla, G. R. Krow, R. T. Raines, *J. Am. Chem. Soc.* **2009**, *131*, 7244 – 7246; b) M. P. Hinderaker, R. T. Raines, *Protein Sci.* **2003**, *12*, 1188 – 1194; c) J. A. Hodges, R. T. Raines, *Org. Lett.* **2006**, *8*, 4695 – 4697. See also: d) N. H. Shah, G. L. Butterfoss, K. Nguyen, B. Yoo, R. Bonneau, D. L. Rabenstein, K. Kirshenbaum, *J. Am. Chem. Soc.* **2008**, *130*, 16622 – 16632; e) B. C. Gorske, B. L. Bastian, D. G. Geske, H. L. Blackwell, *J. Am. Chem. Soc.* **2007**, *129*, 8928 – 8929.
- 23) Helma Wennemers *et al; Angew. Chem. Int. Ed.* **2010**, *49*, 6324 – 6327
- 24) Royo *et al JACS*, **2005**, 9459-9468
- 25) Chmielewski *et al J. AM. CHEM. SO*; **2005**, *127*, 11798-11803
- 26) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Klumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.

- 27) S. Beychok, *Science*, **1966**, 154, 1288
- 28) B. L. Vallee, J. F. Riordan, J. T. Johansen, and D. M. Livingston, Cold Spring Harb. *Symp. Quant. Biol.* **1972**, 36, 517
- 29) A. J. Adler, N. J. Greenfield, and G. D. Fasman, *Methods Enzymol.* **1973**, 27, 675
- 30) (a) Whitmore, L; Wallace, B.A. *Biopolymers*, **2008**, 89, 392-400; (b) Greenfield N. J.; *Nature protocols*, **2006**, 1, 2876-2890 (c) Analysis of circular dichroism data, Norma J. Greenfield, *Methods in enzymology*, 383, 282-317
- 31) Ogle, J. D., Arlinghaus, R. B. Logan, M. A. LOGAN; *J. Biol. Chem.* **1962**, 237, 3667.
- 32) Manaswini Nanda and Krishna N. Ganesh, *J. Org. Chem.* **2012**, 77, 4131–4135

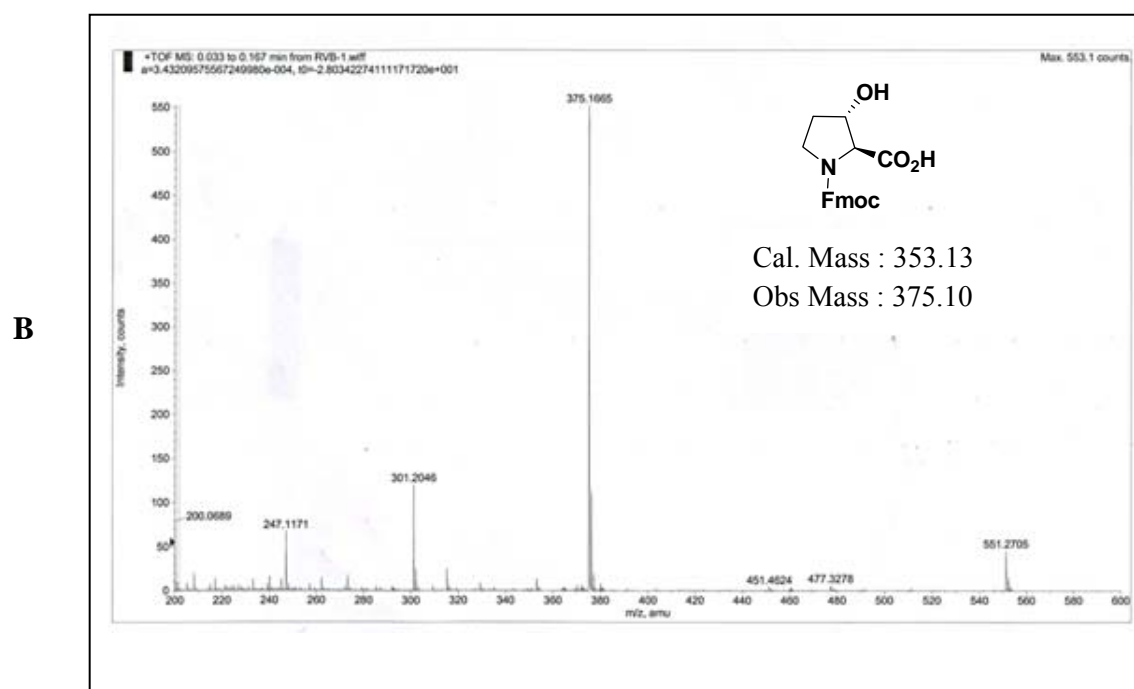
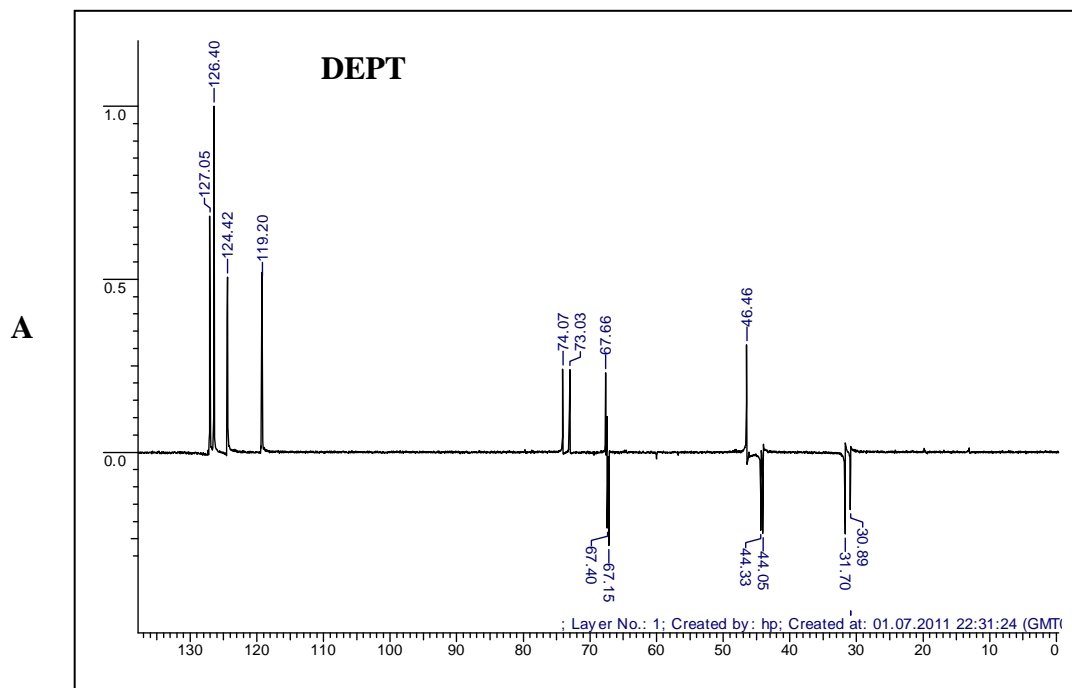


**4.12 : APPENDIX**

<b>Entry</b>	<b>Page No</b>
$^1\text{H}$ , $^{13}\text{C}$ spectra of N-fmoc 3 <i>S</i> -hydroxy proline	203
DEPT and Mass spectra of of N-fmoc 3 <i>S</i> -hydroxy proline	204



**Figure 29:** (A) <sup>1</sup>H NMR of N-fmoc-3S-hydroxy proline (B) <sup>13</sup>C NMR of N-fmoc-3S-hydroxy proline



**Figure 30:** (A) DEPT spectra of N-fmoc-3S-hydroxy proline (B) Mass spectra of N-fmoc-3S-hydroxy proline