Synthesis and biophysical studies of PNA analogues incorporating 3,4-disubstituted pyrrolidine

Thesis submitted to

The University of Pune for degree of

Doctor of Philosophy

in

Chemistry

By

Ashwani Kumar Sharma

Research Supervisor

Prof. Krishna N. Ganesh

Division of Organic Chemistry

National Chemical Laboratory

Pune 411008

September, 2010

CERTIFICATE

This is to certify that the work presented in the thesis entitled "**Synthesis and biophysical studies of PNA analogues incorporating 3,4-disubstituted pyrrolidine**" submitted by **Ashwani Kumar Sharma** 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 **September, 2010**
(Research Supervisor)
Director, IISER Pune
Pune- 411008

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

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled "**Synthesis and biophysical studies of PNA analogues incorporating 3,4-disubstituted pyrrolidine**" 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.

Ashwani Kumar Sharma **September, 2010**
National Chemical Laboratory
Pune 411 008.

Dedicated to my beloved parents
and to my lovely wife

Acknowledgement

It gives me an immense pleasure to express my deep sense of gratitude towards my research supervisor Prof. K. N. Ganesh for all the advice, guidance and encouragement during every stage of this work. I am thankful to him for introducing me to this field of bio-organic chemistry and the freedom he gave me to work and learn new things in this area. 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. H. V. Thulasiram, Dr. Srinivas Hotha, Mrs. Anita Gunjal for their help and encouragement during the course of this work. I am also thankful to Mrs. M. V. Mane for hplc analysis, Mrs. Shantakumari and Dr. Mahesh Kulkarni for mass analysis.

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 department and other departments for helping me whenever I was in need. Thanks to Dr. Roopa, Sreedhar Manaswini, Pradnya, Deepak, Nitin, VDC, Ellipilly. A special thanks to Mahesh Sonar, as he will feel bad if I will not mention him as special (kidding), his wife Yamini and sweet daughter Sayee for never making me feel that I am away from home.

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.

I thank all my family members especially my parents who taught me to always follow the basic principles of life in any condition and to my wife, marry to whom, I always consider my best decision of life.

At last but not least, I thank to all scientific community. Reading their work in literature, books, articles etc. encouraged and will encourage me to do good research.

Ashwani Kumar Sharma

Contents

4 Chapter-4

Synthesis of new pyrrolidine analogues with potential medicinal and biological applications.

Abbreviations

A	Adenine
Ac_2O	Acetic anhydride
ACN	Acetonitrile
aeg	Aminothylglycine
aep	Aminoethylprolyl
ala	Alanine
ap	Antiparallel
aq.	Aqueous
Boc	Tert.butyloxycarbonyl
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	*O*-(1H-Benzotriazol-1-yl) *N*,*N*,*N1*,*N1*-tetramethyl uronium hexafluorophosphate
HIV	Human Immuno Difficiency Virus
HOBt	1-Hydroxybenztriazole
HPLC	High Performance Liquid Chromatography
Hz	Hertz
IR	Infra red
LC-MS	Liquid Chromatography-Mass Spectrometry
Lys	L-Lysine
M	Molar
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-Methyl benzhydryl amine
MF	Merrifield Resin
mg	Milligram

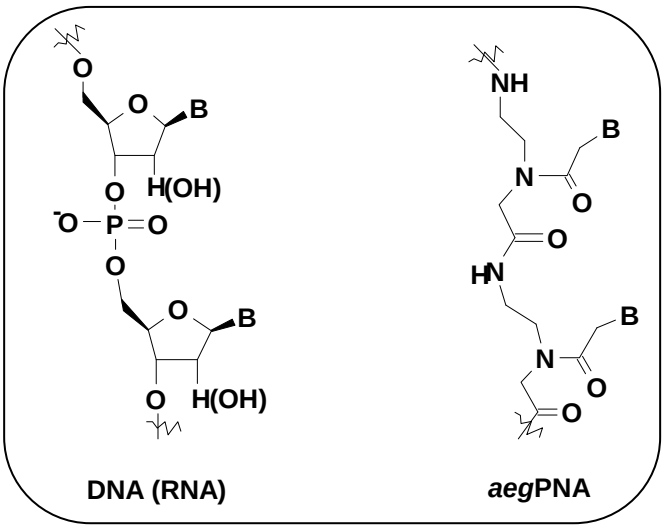
MHz	Megahertz
µM	Micromolar
ml	Milliliter
mM	Millimolar
mmol	Millimoles
M.P.	Melting point
MS	Mass spectrometry
N	Normal
nm	Nanometer
NMP	N-methyl pyrrolidine
NMR	Nuclear Magnetic Resonance
ONs	Oligonucleotides
p	Parallel
PCR	Polymerase Chain Reaction
PPh_3	Triphenyl phosphine
PNA	Peptide Nucleic Acid
Pro	Proline
Pyr	pyrrolidine
RNA	Ribonucleic acid
r.t.	Room temperature
ss	Single strand/ Single stranded
s	Seconds
T	Thymine
TBTU	*O*-(1H-Benzotriazol-1-yl) *N*,*N*,*N1*,*N1*-tetramethyl uronium tetrafluoroborate
TEA	Triethylamine
TFA	Trifluroacetic acid
TFMSA	Trifluromethanesulphonic acid
THF	Tetrahydrofuran
UV-Vis	Ultraviolet- Visible

ABSTRACT

The thesis entitled **"Synthesis and biophysical studies of PNA analogues incorporating 3,4-disubstituted pyrrolidine."** 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 unmodified oligonucleotides are rapidly identified and cleaved by the action of nucleases that primarily hydrolyze the phosphodiester of the internucleoside backbone. Furthermore, the ability of the negatively charged DNA to cross phospholipid cell membrane is poor. Hence structural modifications are needed 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),[1,2] which has been shown to hybridize with complimentary DNAs with higher affinity and fidelity than DNAs.



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 N-(2-Aminoethyl)-glycine units. The nucleobase is attached to the backbone through a conformationally rigid tertiary acetamide linker group.

PNA binds to DNA and RNA with high thermal stability and specificity 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 via antisense/antigene technology, for diagnostics and biosensors, and as building blocks for design of PNA supramolecular constructs.

Despite success in many applications PNA suffers from poor aqueous solubility, with tendency to aggregate in aqueous media and 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 the PNA 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 3,4-disubstituted pyrrolidine based peptide nucleic acid monomers

Objective: It is reported in the literature that introduction of positive charges 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.

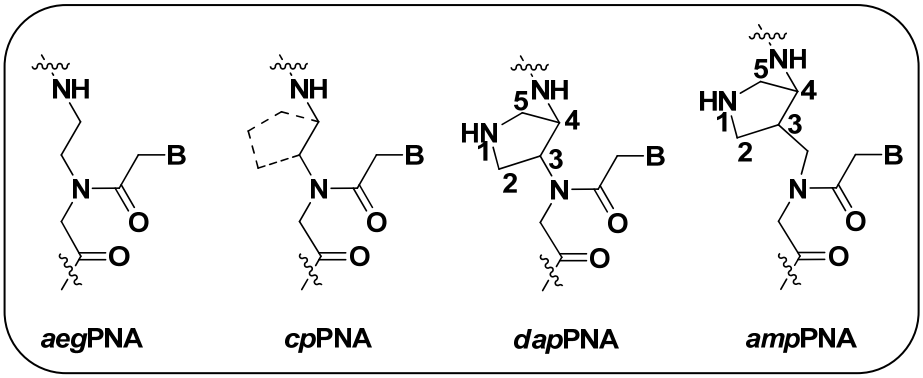


Figure 1: Comparison of backbone in different modified PNAs

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[3] and with *trans* chemistry by Appella *et al*[4] (Figure 1). 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 PNA cationic by the introduction of a nitrogen atom in the five membered cyclopentyl ring in order to generate 3,4-disubstituted pyrrolidine PNA monomers **13** and **26** (Figure 2). These involve *dap*PNA derived from 3,4-diaminopyrrolidine and *amp*PNA derived from 3-(aminomethyl)pyrrolidin-4-amine, having one carbon extended backbone at C-3 position. 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.

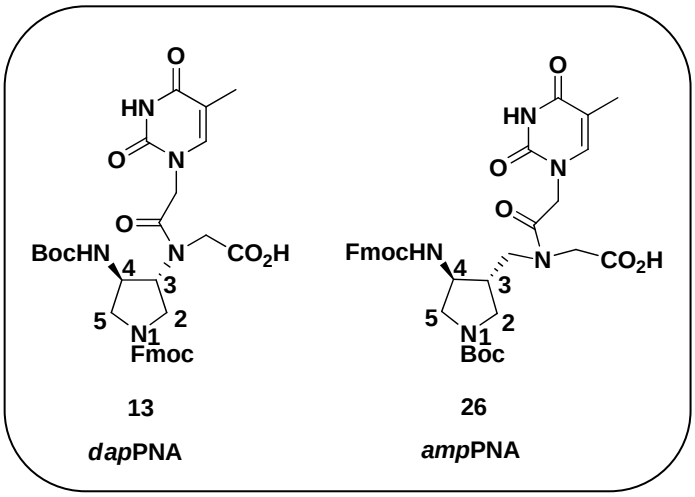
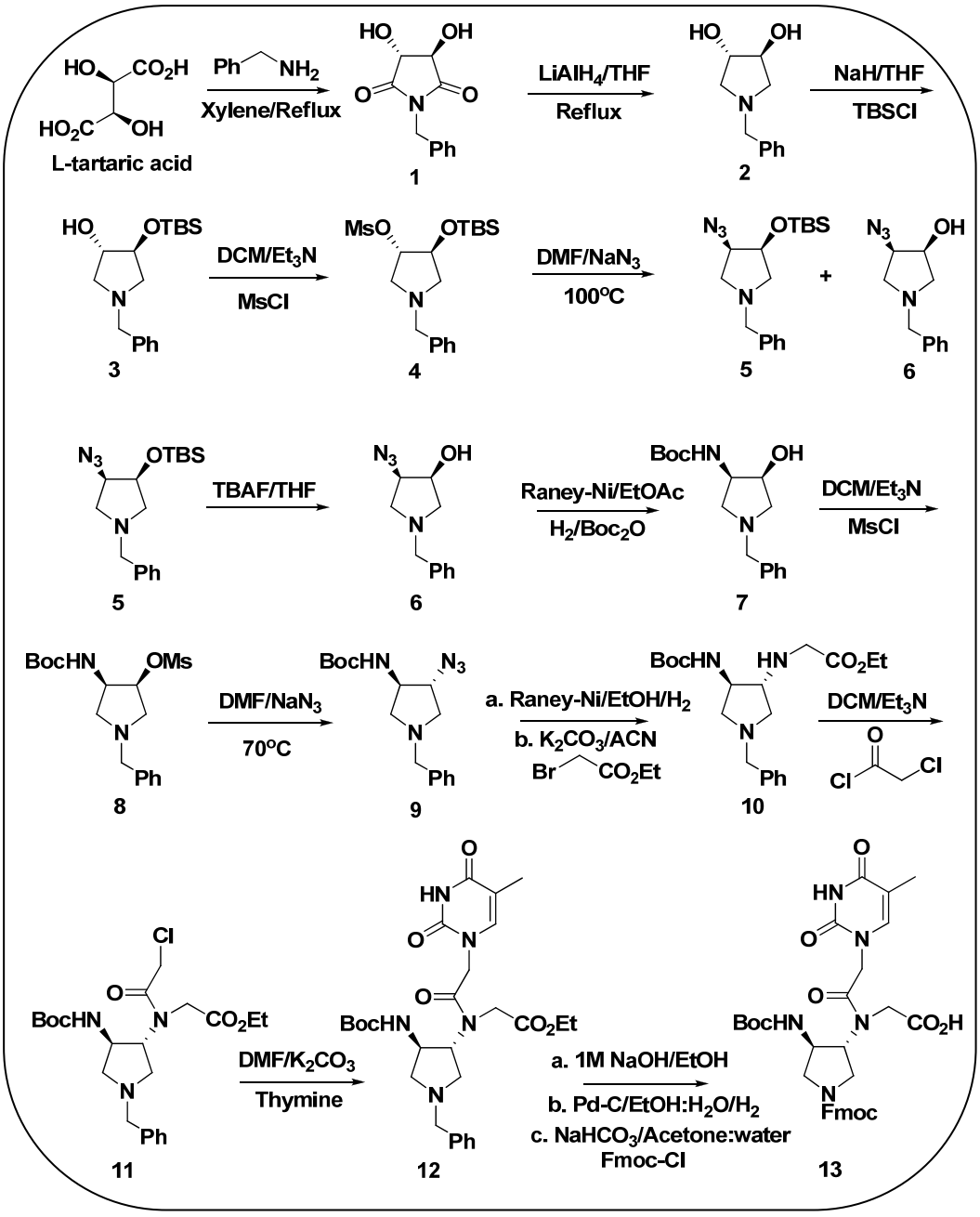


Figure 2: Chemical structure of modified PNA monomers

Synthesis of (3*R*,4*R*) *dap*PNA Monomer (13)

This monomer was synthesized by route shown in Scheme 1 starts from L-Tartaric acid which was reacted with benzylamine to get the corresponding imide **1** that was reduced to the diol **2**. This was mono O-protected with silyl ether to get compound **3** that was mesylated to yield **4**.

Scheme 1: Synthesis of (3*R*,4*R*) *dap*PNA Monomer.

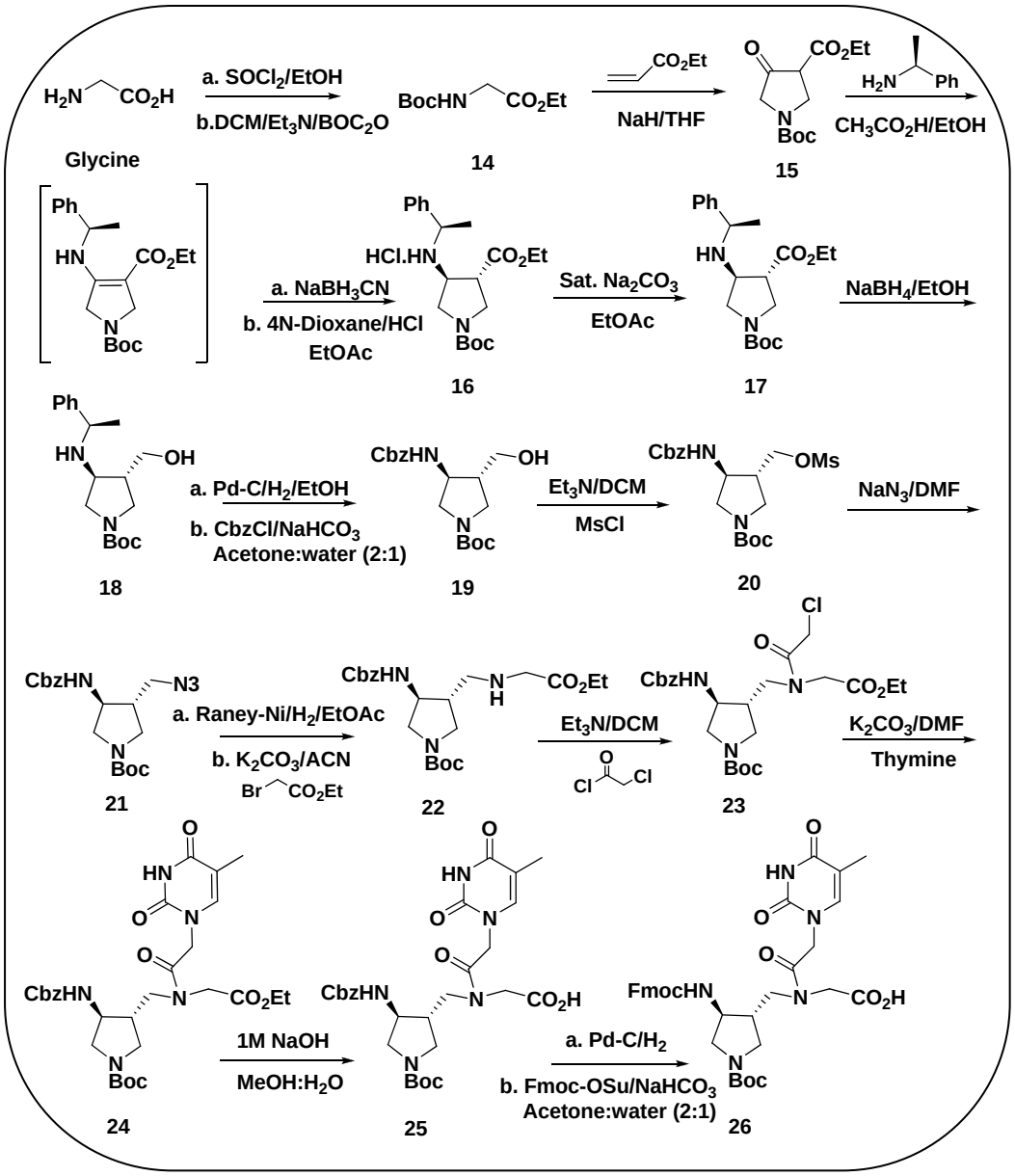


The mesyl derivative **4** was reacted with sodium azide to obtain the azide **5** which was O-desilylated to get the alcohol **6** followed by reduction of azide to amine and *insitu* N-*boc* protection to give **7**. Upon O-mesylation, **7** yielded **8** which on azidation gave **9**. The relative stereochemistry of compounds **8** and **9** were established by crystal structure. On reduction followed by alkylation, azide **9** yielded compound **10** that was N-acylated with chloroacetyl chloride to give the chloro compound **11.** This was used for N1-alkylation of thymine to produce compound **12** which was converted to the required *dap*PNA monomer **13** by hydrolysis of ester followed by N-debenzylation and N-*fmoc* protection.

Synthesis of (3*R*,4*S*) backbone extended *amp*PNA Monomer (26)

The required monomer **26** was synthesized by route as outlined in Scheme 2. Glycine was first protected as ethyl ester followed by the protection of amine with t-boc group to get compound **14** which on aza-Micheal addition followed by the Dieckmann condensation yielded the β-ketoester **15**. This on reductive amination with chiral auxiliary R-(+)-phenylbenzylamine followed by *insitu* reduction yielded the corresponding diastereomeric mixture, from which the desired compound **17** was obtained via its hydrochloride salt **16**. This was reduced to the corresponding alcohol **18** which was debenzylated followed by N-protection with Cbz group to get alcohol **19**. Upon O-mesylation **19** gave the corresponding mesyl derivative **20** which was reacted with sodium azide to yield **21** that was reduced to amine and N-alkylated to **22.** The reaction of **22** with chloroacetyl chloride yielded the chloro compound **23** which was used for N1-alkylation of thymine to obtain ester **24**. This was hydrolysed to the acid **25** and N-decarboxybenzylated to give the free amine which was *insitu* protected as N-*fmoc* derivative to obtain desired *amp*PNA Monomer **26**. All new compounds were characterized by spectral and analytical data.

Scheme2: Synthesis of (3*R*,4*S*) backbone extended *amp*PNA monomer



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

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

The two monomers **13** and **26** can be incorporated into standard PNA oligomers by using either Boc or Fmoc chemistry for solid phase synthesis leading to four different PNA backbones (Figure 3), depending upon the direction of synthesis. All these carry positive charges on their primary or secondary amino groups.

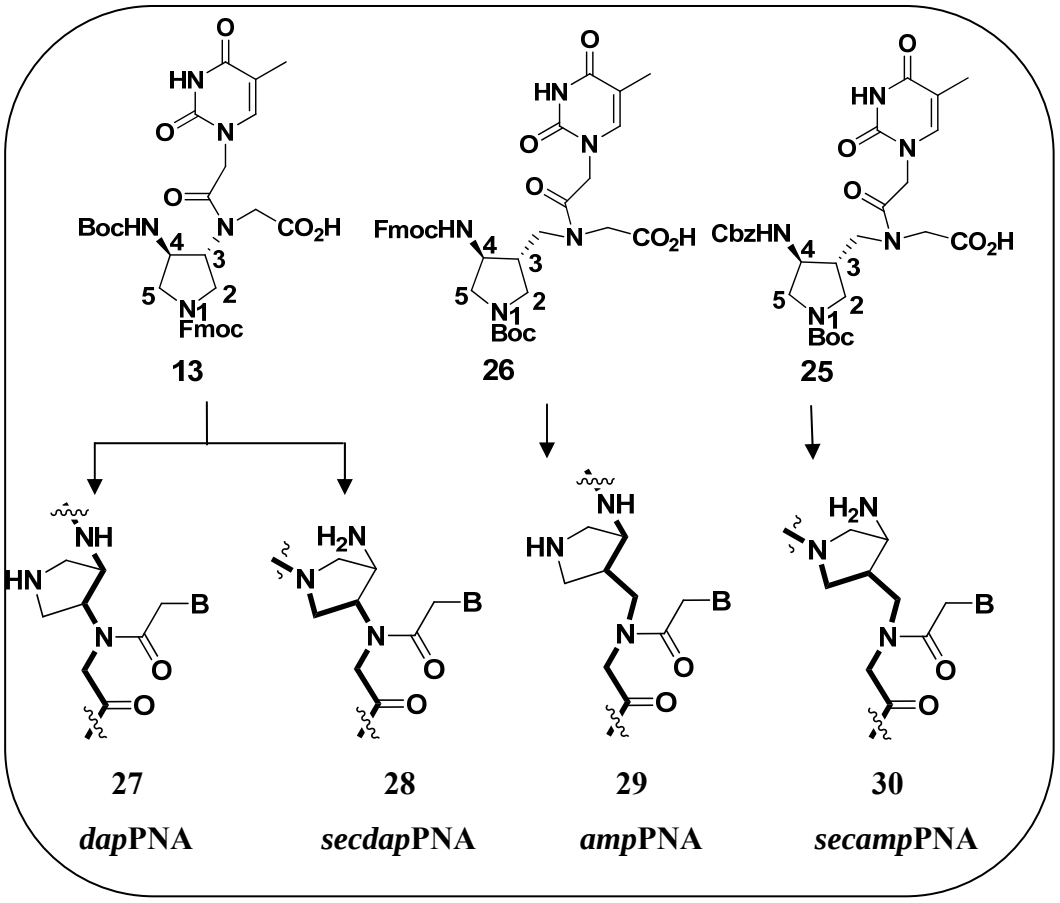


Figure 3: Backbones (shown by thick lines) derived from modified PNA monomers.

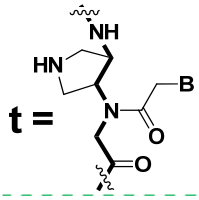
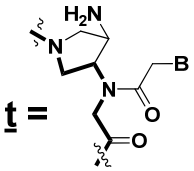
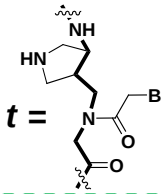
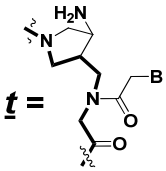
Monomer **13** was used to synthesize PNA oligomers with backbone shown in **27** and **28** by using Boc and Fmoc chemistry respectively. Monomer **26** was used to synthesize PNA oligomer with backbone shown in **29** using Fmoc chemistry. The monomer **25** was used to give PNA oligomer with backbone as shown in **30** using Boc chemistry.

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 – 21** (Table 1, 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.

Biophysical Studies of modified PNA oligomers

The binding specificity of modified PNAs towards complementary DNA were investigated using the biophysical techniques temperature dependent UV-spectroscopy (UV-Tm studies) and Circular Dichroism Spectroscopy (CD spectroscopy). The UV-melting studies were carried out with modified PNA:DNA complexes and analyzed with respect to that of control *aeg*PNA:DNA complexes.

Table 1: Melting studies of PNA_2:DNA triplexes

Modified unit	S. no.	Sequence	UV Tm* DNA 1 (ΔTm_1)
T = *aeg* unit	PNA 1	H-TTTTTTTT-Lys-NH_2	42.1
	PNA 2	H-*t*TTTTTTT-Lys-NH_2	44.8 (+2.7)
	PNA 3	H-TTT*t*TTTT-Lys-NH_2	29.0 (-13.1)
	PNA 4	H-TTTTTTTt-Lys-NH_2	49.8 (+7.7)
	PNA 5	H-$\underline{t}$TTTTTTT-Lys-NH_2	45.6 (+3.5)
	PNA 6	H-TTT$\underline{t}$TTTT-Lys-NH_2	25.1 (-17.0)
	PNA 7	H-TTTTTTT$\underline{t}$-Lys-NH_2	34.3 (-7.8)
	PNA 8	H-*t*TTTTTTT-Lys-NH_2	44.9 (+2.8)
	PNA 9	H-TTT*t*TTTT-Lys-NH_2	39.4 (-2.7)
	PNA 10	H-TTTTTTT*t*-Lys-NH_2	45.8 (+3.7)
	PNA 11	H-$\underline{t}$TTTTTTT-Lys-NH_2	45.8 (+3.7)
	PNA 12	H-TTT$\underline{t}$TTTT-Lys-NH_2	25.0 (-17.1)
	PNA 13	H-TTTTTTT$\underline{t}$-Lys-NH_2	44.6 (+2.5)

***DNA1** is **5'-GCAAAAAAAACG-3'**, **(ΔTm_1)** = difference w.r.t. control.

The UV-*T*m studies for triplexes (with DNA) from modified homopyrimidine sequences, each derived from different monomers indicated stabilization when the modified unit is present at N-terminal or C-terminal and destabilization when present in the middle as shown in Table 1.

Table 2: Melting studies of PNA:DNA duplexes

S. no.	Sequence	UV-Tm	
		DNA *2* (ΔTm_2)	DNA3 (ΔTm_3)
PNA 14	H-GTAGATCACT-Lys-NH_2	44.7	38.1 (-6.6)
PNA 15	H-GtAGATCACT-Lys-NH_2	49.8 (+5.1)	41.5 (-8.3)
PNA 16	H-GTAGAtCACT-Lys-NH_2	57.6 (+12.9)	48.4 (-9.2)
PNA 17	H-GtAGAtCACT-Lys-NH_2	58.1 (+13.4)	45.3 (-12.8)
PNA 18	H-G*t*AGATCACT-Lys-NH_2	nb	nb
PNA 19	H-G*t*AGATCACT-Lys-NH_2	51.1 (+6.4)	49.8 (-1.3)
PNA 20	H-G*t*AGATCACT-Lys-NH_2	49.7 (+5.0)	39.5 (-10.2)
PNA 21	H-G*t*AGA*t*CACT-Lys-NH_2	57.9 (+13.2)	51.1 (-6.8)

*nb represents not binding; **DNA 2** = **5'-AGTGATCTAC-3'(*antiparallel*)**, **DNA 3** = **5'-CATCTAGTGA-3'(*parallel*)**. **A,G,C,T** = ***aeg*PNA**. **(ΔTm_2)** = difference w.r.t. control, **(ΔTm_3)** = difference w.r.t **DNA 2.**

The mixed purine-pyrimidine PNA:DNA duplexes derived from modified PNAs incorporating one of the four monomers at different positions show stabilization in all cases (Table 2). The duplexes and triplexes of *dap* and *amp* PNAs are more stabilised than their analogues *secdap* and *secamp* PNAs. These modified PNA sequences also exhibited more destabilization with single mismatched PNA:DNA duplexes compared to unmodified duplexes.

Chapter 4: Synthesis of new pyrrolidine analogues with potential medicinal and biological applications.

Synthesis of (3*S*,4*S*) Pyrrolidine β-PNA monomer (33)

The past work from this laboratory has resulted in the synthesis of several pyrrolidine PNAs based on proline structures having substitution α- to ring nitrogen, one of the example being pyrrolidine PNA shown in Figure 4 which has a base linker at C-2 position. In this chapter, it is explored to shift the position of base linker from C-2 to C-3 carbon to generate new series of PNA analogues. This modification will allow the introduction of rigidity into the PNA backbone in the form of pyrrolidinyl ring, while simultaneously keeping a certain degree of flexibility in the linker to the

nucleobase. The pyrrolidine ring nitrogen is a tertiary amine and bears a positive charge at physiological pH and may help to improve the solubility of PNA oligomers.

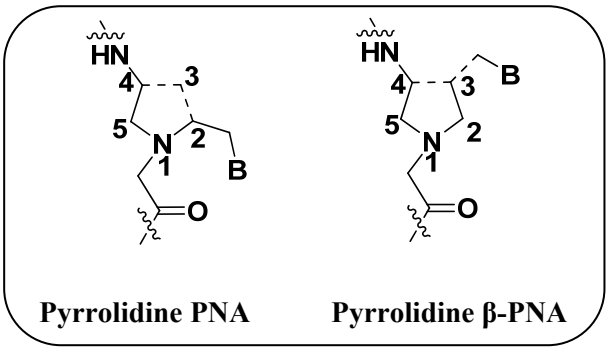
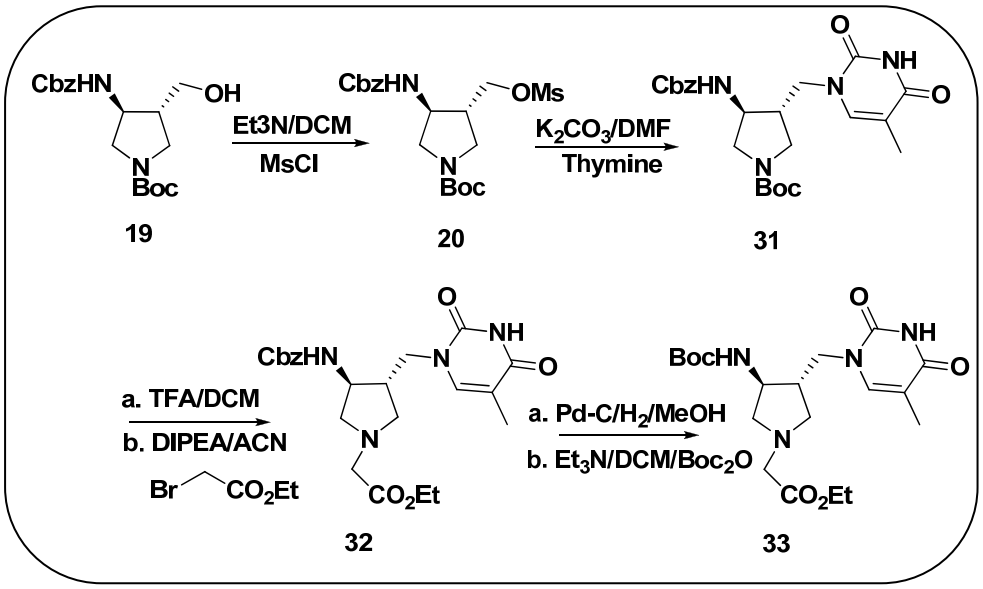


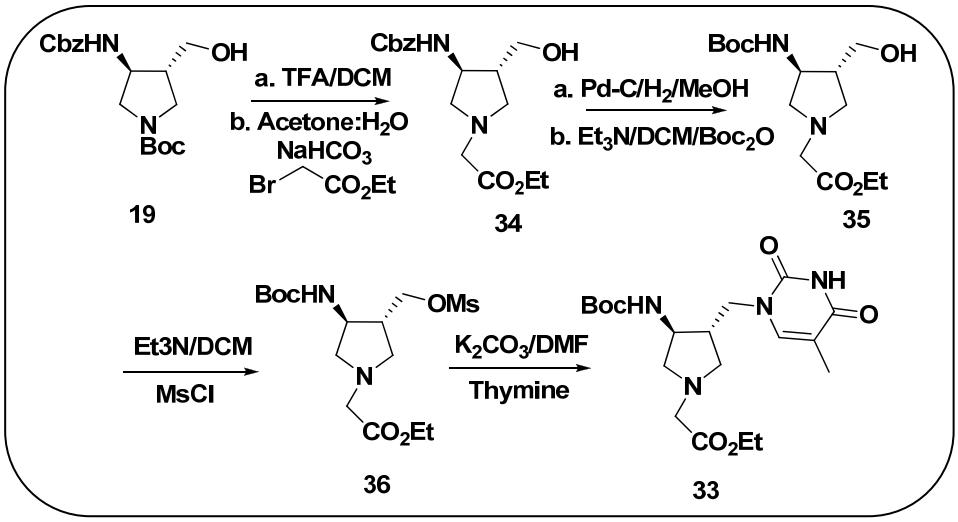
Figure 4: backbone representation of pyrrolidine and Pyrrolidine β-PNA

As shown in Scheme 3, mesyl group of compound **20** was replaced by thymine to get compound **31**. This upon N-boc deprotection gave corresponding amine which was N-alkylated to get compound **32**. The N- Cbz group of compound **32** was deprotected to get free amine which was *insitu* protected with *boc* to yield the Pyrrolidine β-PNA monomer **33**. To improve the overall yield, an alternative route for synthesis of **33** as shown in Scheme 4 has also been tried.

Scheme 3: Synthesis of (3*S*,4*S*) Pyrrolidine β-PNA monomer 33



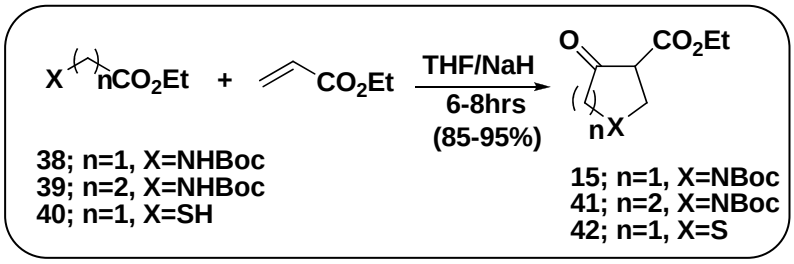
Scheme 4: Alternative route for the synthesis of monomer 33



Convenient and scalable route to heterocyclic β-ketoesters as intermediates for substituted pyrrolidine derivatives

A convenient and scalable route to five and six membered heterocyclic β-ketoesters (Scheme 5) has been carried out. These compounds have previously received attention as precursors for the synthesis of a variety of medicinally and biologically interesting molecules. They are the key compounds in the synthesis of various Carbapenem derivatives that have shown higher activity when compared to imipenem, one of the potent antibacterial agent used as last resort against infections in clinics. They are also key compounds for the synthesis of various unnatural amino acids especially β-amino acids.

Scheme 5

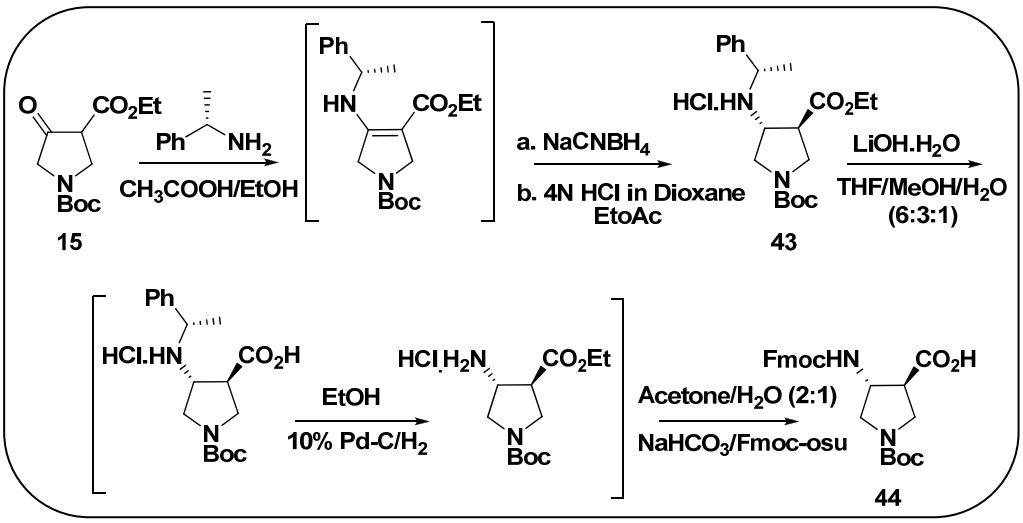


The conventional synthesis consists of two-step procedure involving the Michael addition of glycine ester and ethyl acrylate followed by an intramolecular Dieckmann condensation in the second step. This has now been reduced to a single step by using NaH as a base and THF as a solvent which simultaneously effect both

Michael addition and Dieckmann condensation to yield ketoesters **15, 41-42** in quantitative yield.

The five membered keto ester **15** was converted to the N-protected (3*S*,4*R*) 4-amino-β-proline **44** as shown in Scheme 6 which can be used as a precursor to study the effect on the stability of collagen model peptides.

Scheme 6



Collagen is a group of proteins, present mainly in connective tissues, is made up of three polypeptide strands with Gly-X-Y repeats where X and Y may be proline or Hydroxyproline. Recent studies[5] with $(Pro.4\text{-}Amp.Gly)_n$, $(Pro.Flp.Gly)_n$ and $(Pro.3\text{-}Hyp.Gly)_n$, collagen model peptides in the Y position of the Gly-X-Y repeat have shown stabilizing effect on triple helices compared to their corresponding Hyp peptides. In view of these findings, it may be worthwhile to study the effect of 4-amino-β-proline on collagen stability as the β-peptides are well known for forming interesting secondary structures. Also the 4-amino group may provide a pH dependent switch for the triple-helix stability.

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.

CHAPTER 1

Introduction

1.1 Introduction to Nucleic Acids

Nucleic acids, play an important role in the cell nucleus. These molecules store and transmit genetic information required for the execution of essential steps to make proteins. They are biopolymers composed of monomers called 'Nucleotides' which consists of three components, a nitrogenous heterocyclic base (a purine or a pyrimidine); a pentose sugar; and a phosphate group (Figure 1). The nucleic acids can be differentiated into DNA and RNA on the basis of pentose sugar and the bases present.

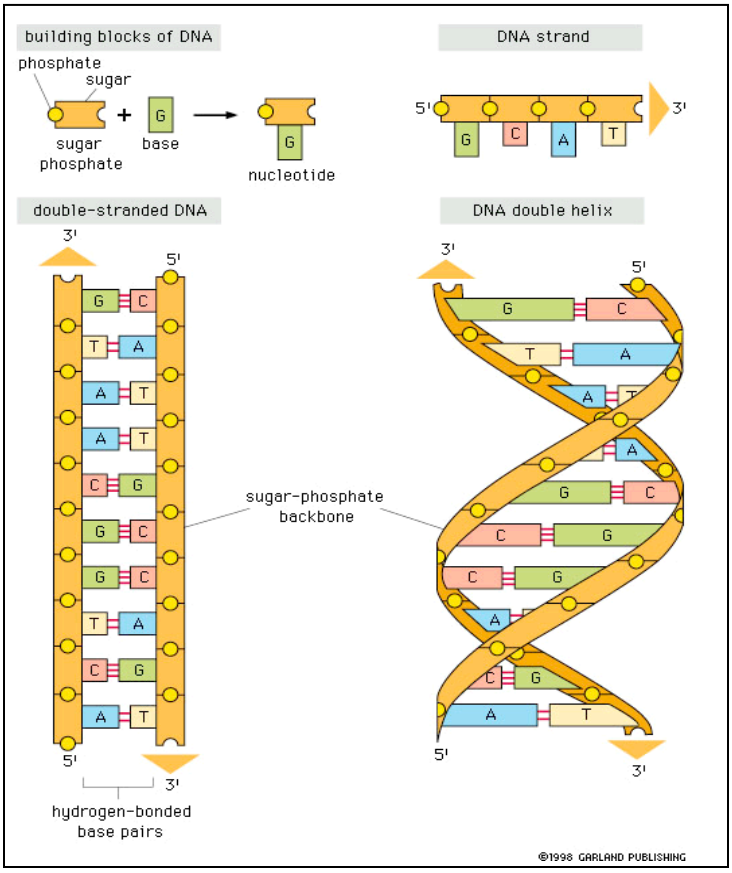


Figure 1: Model structure of DNA

DNA contains 2-deoxyribose while RNA contains ribose. Likewise DNA contains four bases Adenine (A), Guanine (G), Cytosine (C) and Thymine (T), while RNA has A, G and C but instead of T, RNA contains Uracil (U) (Figure 2). A very rare exception is a bacterial virus called PBS1 that contains uracil in its DNA.[1]

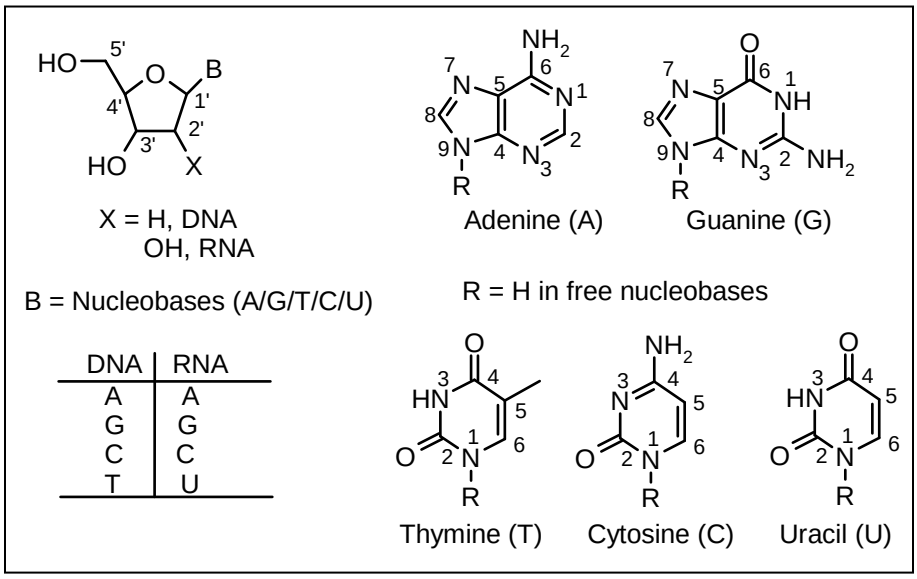


Figure 2: Chemical structures of components of DNA/RNA

1.1.1 Base Pairing via Hydrogen bonding

The chemical structure alone cannot be credited for the unique properties of nucleic acids. It was found in the first half of the 20th 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,[2] 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.

The N-H groups of the nucleobases are potent hydrogen bond donors, while the sp^2-hybridized electron pairs on the oxygen in the carbonyl group of the bases and that on the ring nitrogen are hydrogen bond acceptors. In Watson-Crick pairing, there are two hydrogen bonds in an A:T base pair and three in a C:G base pair (Figure 3).[2] While Watson-Crick base pairing is the dominant pattern between the nucleobases, other significant pairings are Hoogsteen (HG)[3a] and Wobble base pairs.[3b]

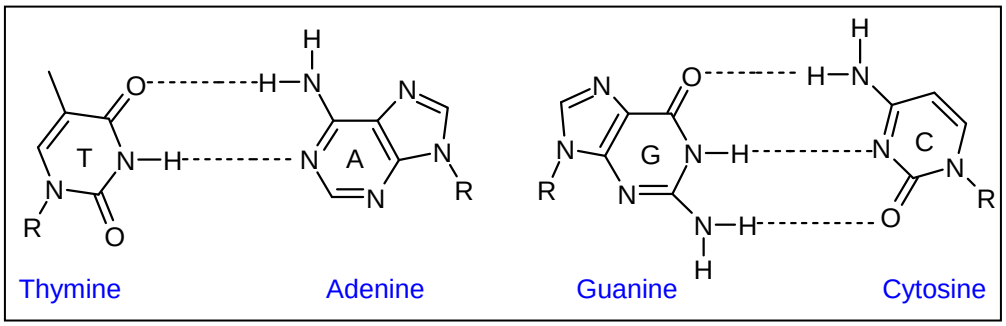


Figure 3: Watson and Crick hydrogen- bonding scheme for A:T and G:C base pair

In Hoogsteen base-pairing (Figure 4), thymine bases can form two hydrogen bonds with adenine already engaged in Watson Crick A:T base pair. Cytosine if protonated can form two hydrogen bonds with guanine of Watson Crick C:G base pair in the parallel orientation and allows sequence specific binding of pyrimidines of third strand in the major groove of Watson-Crick purine:pyrimidine duplexes to form triple-helical structures. The triplexes are present in three-dimensional structures of transfer RNA. In some cases, such as reversed Hoogsteen base pairs, one base is rotated 180° with respect to the other. Guanine rich sequences of nucleic acids form self assembled structures called G-quadruplexes in the presence of monovalent ions (K^+, Na^+ etc) at physiological conditions. These structures are stabilized by cyclic Hoogsteen hydrogen bonding arising from association of four guanines observed in DNA/RNA telomeres. The inventory of sequences forming telomeric structures correspond to d(TTTTGGGG), d(TTGGGG) and d(TTAGGG) present in sequences from *Oxytrieha, Tetrahymena,* and *humans* respectively.

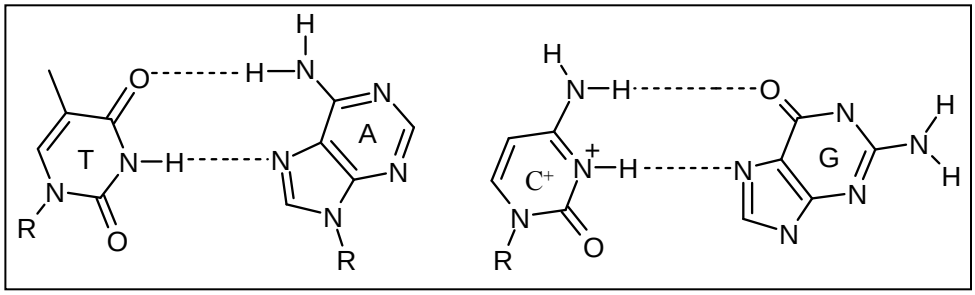


Figure 4. Hoogsteen hydrogen bonding scheme for A:T and G:C base pair

In the wobble base pairing (Figure 5), a single purine base is able to recognize a non-complementary pyrimidine (e.g. G:U) which has 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.[4]

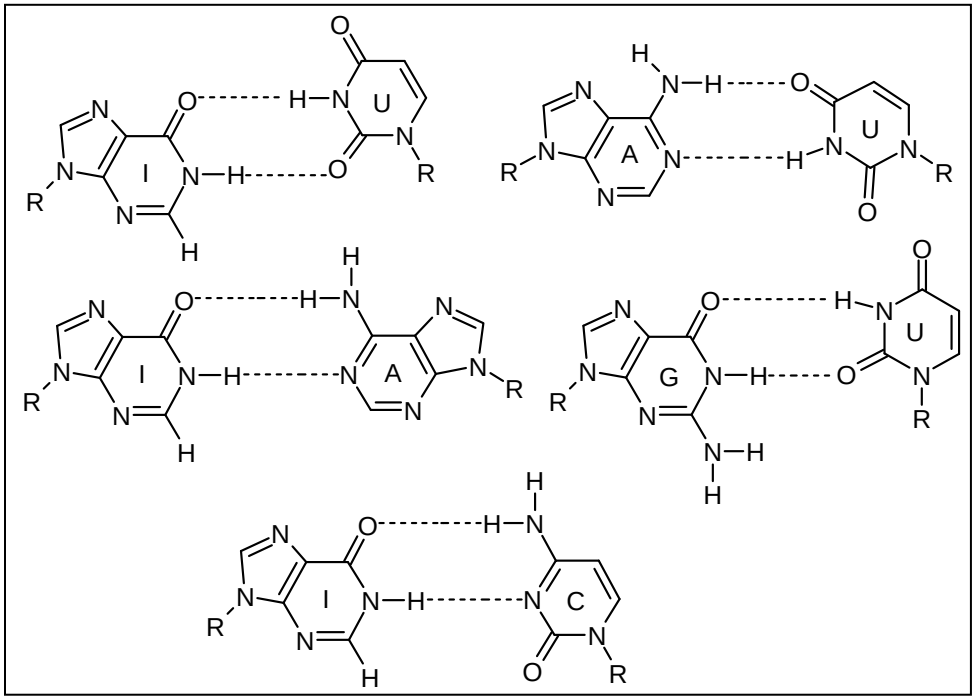


Figure 5. Wobble base pair for Ionosine and Uracil.

1.2 Antigene and Antisense technology

DNA is a double helix, which during the process of transcription splits into individual strands. The coding DNA strand carrying genetic information, then serves as the template for the transcription into *m*RNA sense strand. The concept of `antisense oligonucleotides' as potential therapeutic agents[5] introduced by Zamecnik and Stephenson[6] has arouse much interest in search of potent DNA mimics. Antisense oligonucleotides (Figure 6) recognize a complementary sequence on target m-RNA through Watson-Crick base pairing and form a duplex (RNA-DNA hybrid) that is not processed by the protein synthesis machinery. This would retard the expression of the

corresponding protein and when target proteins are disease related, this will have a therapeutic value.

In another approach, the 'antigene strategy', (Figure 6) obstruction of gene expression can be achieved by binding of oligonucleotides to duplex DNA through Hoogsteen hydrogen bonds leading to the formation of a triple helix.[7] Thus the double stranded DNA itself can act as a target for the third strand oligonucleotides or their analogue, and the limitation for triplex formation is that it is possible only at homopurine stretches of DNA, since it requires purine to be the central base.[7,8]

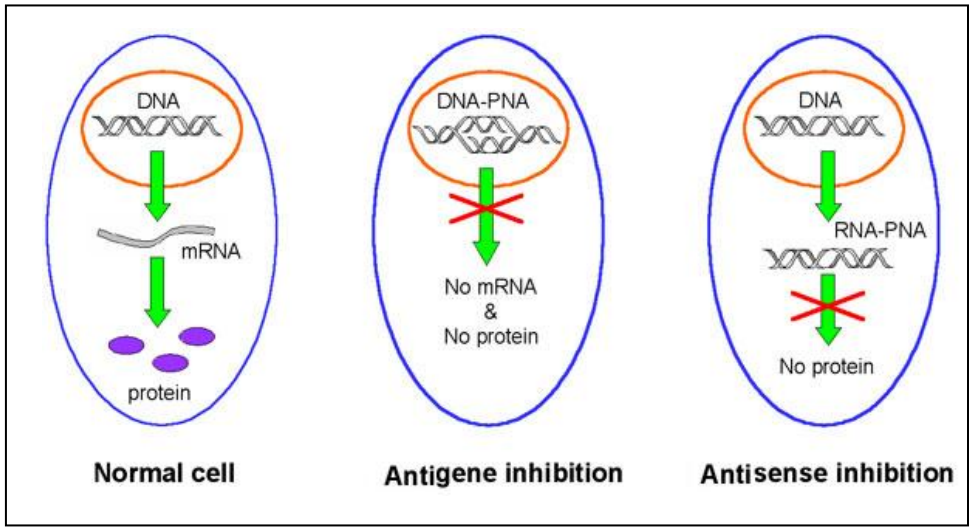


Figure 6: Mechanism of action of antisense and antigene oligonucleotides

While the antisense approach is a convincing strategic model for the inhibition of gene expression, several challenges have to be met in order for synthetic ODNs to become therapeutic agents in biological systems. Unmodified natural ODNs are rapidly degraded in the cellular environment due to the action of nucleases that hydrolyze the phosphodiester (PO) bond. As a consequence, considerable efforts have been focused on the synthesis of ODNs having modifications that render these compounds nuclease-resistant but still allow the formation of stable duplexes/triplexes with target nucleic acids and suppress translation. Additionally, these ODNs should be water-soluble, penetrate cell membranes and reach the cytoplasm or nucleus for their biological effect.

To overcome these drawbacks chemical (structural) modifications of oligonucleotides are needed.

1.3 Chemical modifications of DNA

One of the major challenges for antisense approaches is the stabilization of ODNs, as unmodified oligodeoxynucleotides are rapidly degraded in biological fluids by nucleases and lack the ability to penetrate through the cell membrane. In order to meet all the requirements for successful antisense/antigene oligonucleotides, it is necessary for standard oligonucleotides to be chemically modified in a suitable manner (Figure 7).

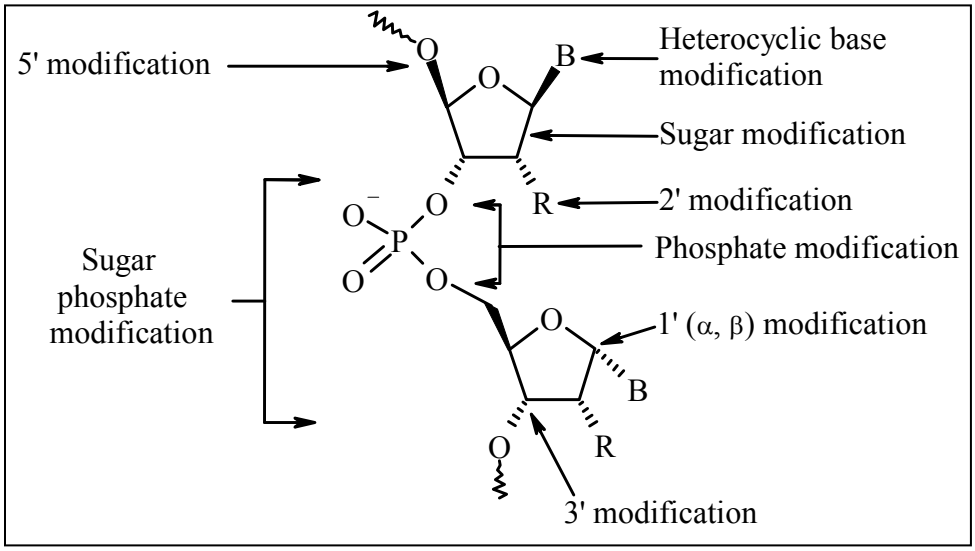


Figure 7: Structurally possible various DNA modifications

Many types of chemically modified antisense ODNs have been produced and tested to date which includes changes in the overall electronic charge of the oligonucleotide, incorporation of non-phosphate oligonucleotide backbones, modification of the nucleobase etc. Several other types of chemically modified nucleotides[9] have been used in antisense experiments. These modifications are commonly grouped into phosphate modifications, nucleobase modifications, sugar modifications and sugar-phosphate backbone modifications.

1.3.1 Modifications in phosphate group

1.3.1a Phosphorothioate (PS-oligos): The PS-oligos are the major representatives of first generation DNA analogs that are the best known and most widely used antisense oligonucleotides to date.[10] They were modified by replacing one of the non-bridging oxygen atoms in the phosphate group with sulphur (Figure 8a)[11,12] thus retaining the negative charge but with reduced charge density as compared to phosphodiester analogues. The main advantages of PS-oligos are resistance against nucleases; ability to recruit RNase H to cleave target mRNA, ease of synthesis and attractive pharmacokinetic properties.

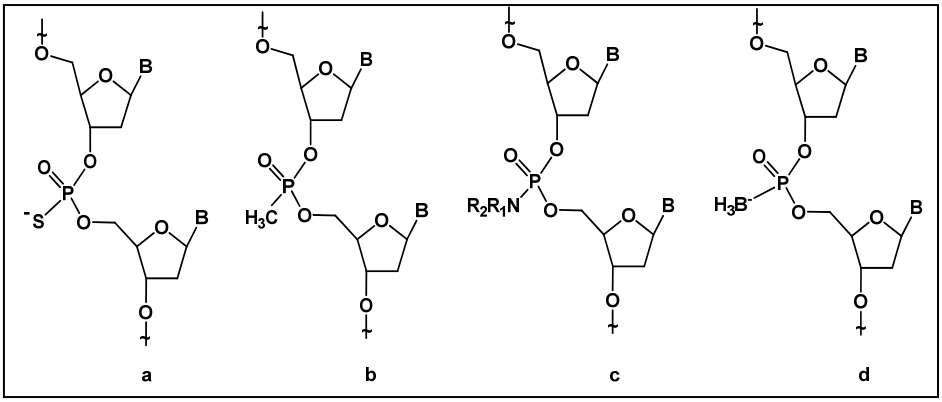


Figure 8: First generation modified antisense oligonucleotides.

The drawbacks of PS-ODNs include lower binding affinity than its corresponding phosphodiester oligonucleotide, which may be attributed to non-stereo specific synthesis of phosphorothioates. In addition, these oligos have a tendency to exhibit generally undesirable sequence-independent properties *in vivo*, such as immune stimulation, complement activation and cellular toxicity.[13-17] The first and the only antisense drug (Vitravene) approved by FDA so far is based on PS-oligos.

1.3.1b Methylphosphonates: Oligonucleoside methylphosphonates are a class of nonionic oligonucleotide analogs (Figure 8b).[18] These oligomers contain deoxyribonucleoside or 2'-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.[19] One of

the non esterified oxygen of PO group is replaced with a methyl group, and this substitution results in the loss of the negative charge. Methylphosphonates are resistant to the action of cellular nucleases. The electroneutral methylphosphonate linkage would reduce charge repulsion in methylphosphonate oligos and complementary target nucleic acid more readily penetrate cell membranes, thus enhancing their internalization by cells (also known by the name MATAGENES, masking tape for gene expression).[20] Although oligodeoxyribonucleoside methylphosphonates form stable duplexes with single stranded DNA, hybrids formed with complementary RNA of the same nucleotide sequence generally have lower stability.[21]

1.3.1c Oligonucleotide Phosphoramidates (Figure 8c): Phosphoramidates are another well-studied class of backbone modifications.[22] They are resistant to hydrolysis by nucleases. With DNA targets, oligonucleotide phosphoramidates (Figure 8c, R_1=H, R_2=CH_3, CH_2CH_2OMe, $(CH_2CH_2)_2O$) exhibit rather poor hybridization characteristics.[23,24] Oligonucleotide phosphoramidates (R_1=H, R_2=CH_2CH_2N-$(CH_2CH_2)_2O$) form weak duplexes at neutral pH and more stable duplexes under acidic conditions (pH 5.6), due to protonation of the terminal amine. Oligophosphoramidate with R_1=H, R_2=$CH_2CH_2NMe_2$ is however reported to hybridize to DNA targets at both neutral and acidic pH.

1.3.1d Oligo-boranophosphonates (Figure 8d): These are derived by replacing one of the non-bridging oxygen atoms in the phosphodiester group of DNA with borane (BH_3).[25] 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 liphophilic than DNA.[26] Boranophosphonate DNA is considerably more stable to various nuclease enzymes than native DNA and overall more stable than phosphorothioate DNA.[27] The discovery that oligonucleotide boranophosphodiesters can activate *E.coli* RNase H and induce cleavage of RNA is encouraging.[28]

1.3.1e Oligonucleotides with backbone replacement not containing phosphorous: A second generation chemical modifications include replacement of the phosphodiester linkage[29] by other atom W-X-Y-Z chain (Figure 9). Among these modifications, common problem is the ineffective permeation of cellular membranes. These ODNs are taken up by endosomes, but are unable to cross the endosomal membrane in the absence of cationic lipids.[30] Although few bind to complementary DNA/RNA fairly well, but none have exhibited the potency to be an effective drug.

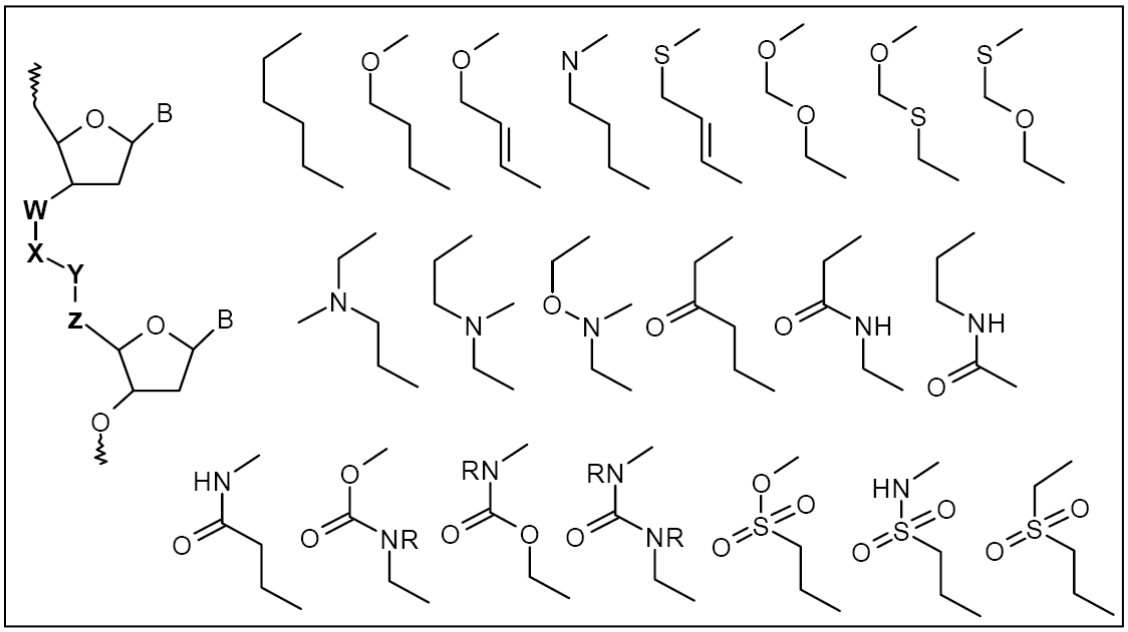


Figure 9: Non-phosphorous neutral backbone modifications

1.3.2 Nucleobase modifications

The binding affinity of oligonucleotides with modification in nucleobases depends upon combination of stacking interactions, effect on Watson-Crick, Hoogsteen/reverse Hoogsteen hydrogen bond formation, hydration forces and hydrophobic interactions.[31-44] The relative contributions are difficult to evaluate in such a system.

1.3.2a Modified nucleobases that stabilize duplex/triplex

The modified nucleobase consisting of phenyl or biphenyl aromatic groups (Figure 10a & b), [31] or the intercalators (Figure 10c) [32] stabilize the duplex through π-stacking interactions. The modifications of pyrimidines at an unpaired site (C^6-position is not involved in Watson-Crick hydrogen bonding) stabilize the *syn*-conformation and leads to duplex destabilization. The oligonucleotides containing modified nucleobases, having fused aromatic system at C^4 and C^5, stabilize the derived duplex (Figure 10d).[33]

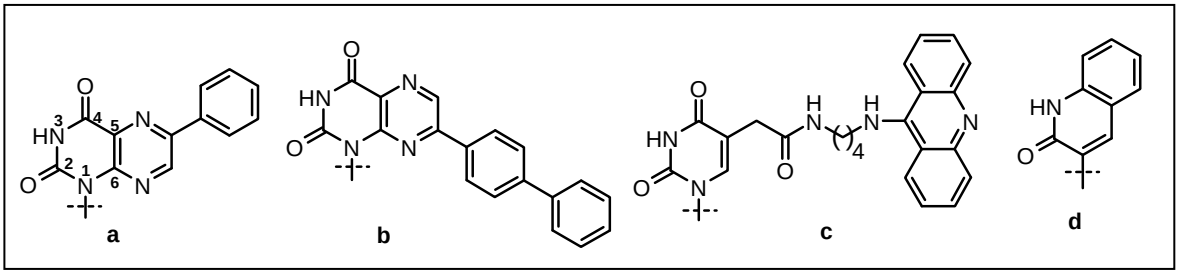


Figure 10: The Duplex-stabilizing modifications a) 6-phenyllumazine, b) 7-(4-biphenyl)lumazine, c) pyrimidine-C^5-intercalator, d) 2-hydroxy quinoline

The pyrimidine nucleosides consisting of unsaturated substituents at *C*-5 position in pyrimidine (Figure 11a, b)[34] and 7-alkynyl-7-deaza-substituent in purines (Figure 11c-f)[35] have a pronounced influence on duplex stabilization. The triple bond stacks with the adjacent base pair and stabilizes the duplex.

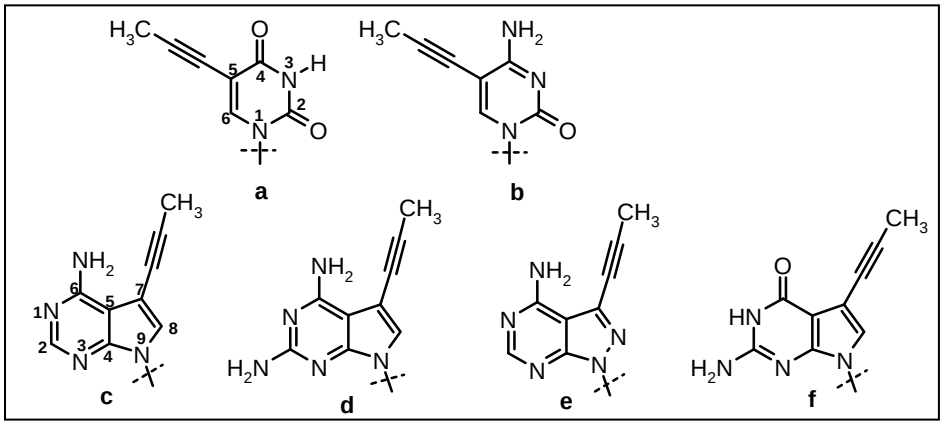


Figure 11: The structures of a) 5-propynyl uracil, b) 5-propynyl cytosine c) 7-propynyl-7-deaza-adenine, d) 7-propynyl-7-deaza-2,6-diamino purine, e) 7-propynyl-7-deaza-8-aza-adenine, f) 7-propynyl-7-deazaguanine

1.3.2b Modifications that form Hoogsteen hydrogen bonding

The incorporation of pseudouracil (Figure 12a) and alkylated pseudouracil (Figure 12b) containing (deoxy) nucleotides are isosteric to thymine analogues. Incorporation of these into oligomers enhances the stability of duplex, whereas 1,3-dialkylpsuedouracil (Figure 12c) containing oligonucleotides destabilize the duplex. The tautomerization in pseudoisocytidine (Figure 12d) has a negative impact on Watson-Crick base pairing and decreases duplex stability.[45-48] 5-aminouracil (Figure 12e) stabilizes triplex when present on central strand.[49]

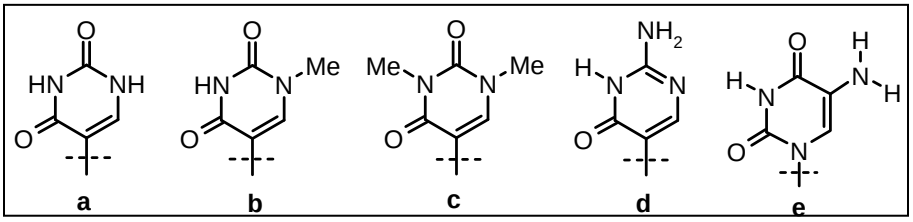


Figure 12 : Structures of a) pseudouracil, b) 1-methyl pseudouracil, c) 1,3-dimethyl pseudouracil, d) pseudoisocytosine, e) 5-aminouracil.

In case of purine, the electron donating substituents such as amino, methylamino, *N*,*N*-dimethylamino, hydroxy at *C*-8, increase the electron density on *N*-7 of purine[50] thereby enhancing its ability to participate as hydrogen bond acceptor in Hoogsteen base pairing (Figure 13).[51-54]

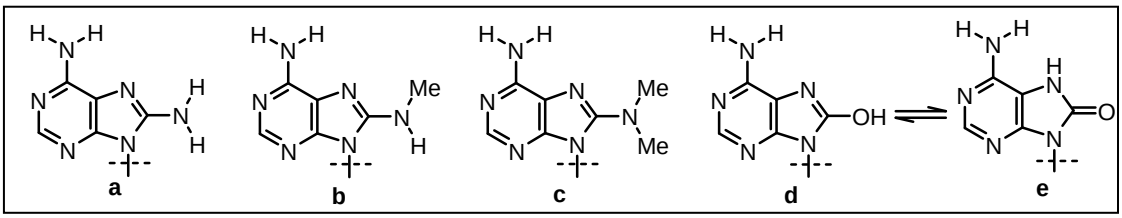


Figure 13: structures of a) 8-amino adenine, b) 8-(*N*-mehyl amino) adenine, c) 8-(*N*,*N*-dimehyl amino) adenine, d) 8-hydroxy adenine, e) 8-oxo-adenine (keto form of 8-hydroxy adenine).

1.3.3 Sugar Modifications

1.3.3a 2'-Modifcations

An important class of second-generation antisense oligonucleotide analogs is 2'-modified oligonucleotides. 2'-O-alkyl substituents (**1**-**4**, Figure 14) enhance the RNA

binding affinity of the corresponding oligonucleotides[55] and also confer a significant degree of nuclease protection to adjacent PO linkages.[56] The 2'-O-methyl sugar 1 (Figure 14) enhances binding to complementary ODNs and confers stability to ssRNA from RNases, but is susceptible to degradation by DNases.

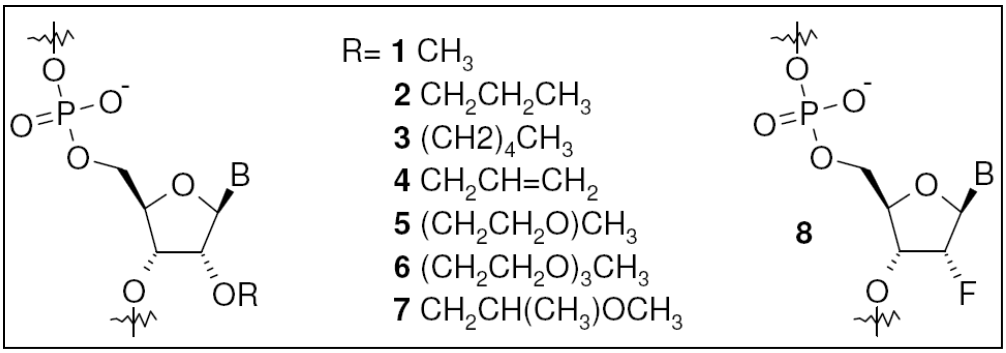


Figure 14: 2'-Modified Oligonucleotides

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.[57] In contrast, ethylene glycol based 2'-O-substituents (**5**-**7**, Figure 14) promote RNA binding to the same extent as a simple 2'-O-methyl group (~1.0^{o}C/modification) and provides much higher degree of nuclease protection. Other 2'-modifications comprise the 2'-fluoro substituents (**8**, Figure 14) which has shown to have pronounced favourable effect on RNA binding affinity (1.5^{o}C/modification) among all 2'- substituents investigated to date. This phenomenon is generally attributed to a strong preference of the modified sugar moieties for RNA-like-3'-*endo* conformation.[58]

1.3.3b Hexitol Nucleic Acids (HNA)

The replacement of conformationally flexible deoxyribose by the restricted anhydrohexitol ring resulted in a structurally pre-organized HNA (Figure 15, I) to form A- type helices with efficient base stacking.[59] Molecular associations between HNA and RNA were found to be more stable than those between HNA and DNA. The presence of 2'OH group could be responsible for the differential hydration patterns in the minor groove of RNA:HNA and DNA:HNA duplexes. The DNA strand in DNA:HNA

complex was found to be more flexible than the RNA strand in RNA:HNA duplexes.[59] In contrast, the crystal structure of HNA:RNA hybrid suggested less-pronounced rigidity of the backbone. The differences in hydration of the HNA and RNA backbones were also observed in the RNA:HNA duplex crystal structure. The hydration of the HNA strand promoted tighter bridging of the adjacent phosphate groups by water molecules and helped alteration in adjacent P-P distances that were close to RNA than DNA.[60]

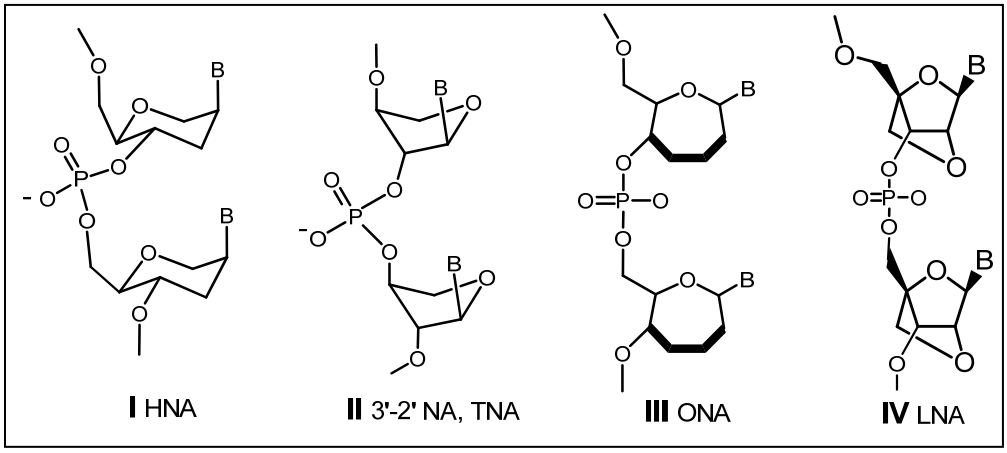


Figure 15: Sugar moiety modified oligonucleotides

1.3.3c L-Threofuranosyl Nucleic Acid (TNA)

3'-2' Nucleic acids (Figure 15, II) cross paired both with RNA and DNA, but the pairing with RNA was found to be stronger in comparison with DNA. The backbone is shorter than both in DNA or RNA as the sugar moiety in TNA contains only four atoms between two adjacent P atoms as compared to five atoms in either DNA or RNA.[61] The detailed crystal structure study of the three dimensional structure provided insight into the origins preferred pairing between TNA and RNA relative to that between DNA and TNA.[61, 62] Cross pairing between TNA and DNA was permissible because of the allowed conformational adjustments of DNA strand in TNA:DNA duplex and not vice-versa. The close and specific distance complementarity could be the structural parameter that determined the capability of TNA ONs for preferential cross pairing.

1.3.3d Oxepane Nucleic Acid (ONA)

Recently Sabatino and Damha[63] have developed a new seven membered DNA analog namely oxepane nucleic acid (ONA) (Figure 15, III). These ONs exhibited cross-pairing with complementary RNA to give a duplex whose conformation evaluated by CD spectroscopy very closely matched that of the natural DNA/RNA hybrid. Furthermore, these ONs confer resistance to nucleases while retaining the ability to direct RNase H-mediated cleavage of a target RNA.

1.3.3e Locked nucleic acids (LNA)

One of the most promising candidates of chemically modified nucleotides developed in the last few years is locked nucleic acid (LNA) (Figure 15, IV), a ribonucleotide containing a methylene bridge that connects the 2'-oxygen of the ribose with the 4'-carbon. Locked Nucleic Acid (LNA) was first described by Wengel[64] and Imanish[65a] et al as a novel class of conformationally restricted oligonucleotide analogues. The design and ability of LNA oligos to bind to supercoiled, double-stranded plasmid DNA in a sequence-specific manner has been described by Catchpole et al[65b] for the first time. LNA oligos are more stably bound to plasmid DNA than the similar peptide nucleic acid (PNA) `clamps' for procedures such as particle mediated DNA delivery. Chimeric DNA:LNA ODNs reveal an enhanced stability against nucleolytic degradation[66] and an extraordinarily high target affinity.[67]

1.3.4 Sugar-Phosphate backbone modification

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

1.3.4a Morpholino nucleotides

Summerton *et al*.[68] prepared novel oligonucleotide analogues from ribonucleotide derived morpholine units, linked by carbamate groups (Figure 16a). 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 16b).

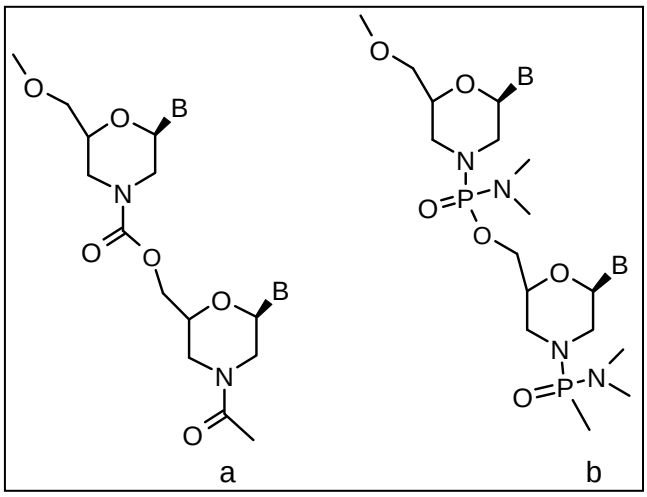


Figure 16: Morpholino ONs linked by **a**) carbamate, **b**) phosphorodiamidate

This is attributed to the fact that morpholino oligos, 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-α (TNF-α) in mouse macrophages, despite poor uptake into these cells.[69]

1.4 Peptide Nucleic Acid

1.4.1 Introduction

PNA refers to peptide nucleic acid (Figure 17) invented by Neilsen et. al 1991[70] with physical properties similar to DNA or RNA but differing in the composition of its "backbone." PNA is unnatural and is chemically synthesized for use in biological research and diagnostics.[71]

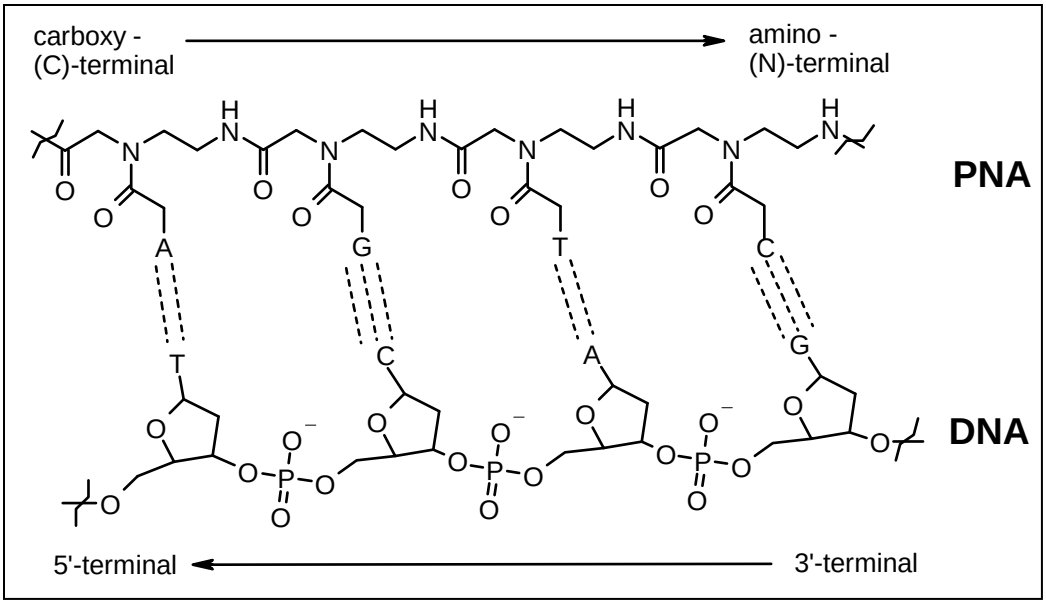


Figure 17: An antiparallel PNA:DNA duplex showing chemical structure of PNA & DNA.

DNA and RNA have backbone containing deoxyribose and ribose sugars respectively, whereas the backbone of PNAs is composed of repeating N-(2-aminoethyl)-glycine units linked by peptide bonds. The various purine and pyrimidine bases are linked to the backbone by methylene carbonyl bonds. PNAs are depicted like peptides, with representation of sequence from N-terminus (left) to C-terminus (right) (Figure 18).

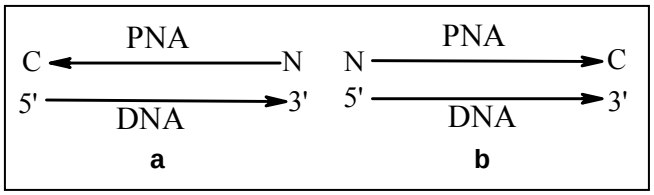


Figure 18: a) anti parallel and b) parallel orientation of PNA in PNA:DNA duplex

Since the backbone of PNA is uncharged, the binding between PNA/DNA strands are stronger than DNA/DNA strands, due to the lack of electrostatic repulsion during hybridization. Early experiments with homopyrimidine strands (strands with repeated pyrimidine bases) have shown that the T_m ("melting" temperature) of a 6-base thymine PNA/adenine DNA double helix was 31°C in comparison to an equivalent 6-base DNA/DNA duplex that denatures at a temperature less than 10 °C.[72]

Synthetic peptide nucleic acid oligomers have been used in recent years in various molecular biology protocols, diagnostic assays and for potential antisense therapeutics. [71] Due to their higher binding strength, even shorter PNA oligomers are apt for use in these roles, which otherwise need longer oligonucleotides (20-25 bases). PNA oligomers also show greater specificity in binding to complementary DNAs, with single PNA/DNA base mismatch being more destabilizing than a similar mismatch in a DNA/DNA duplex. This binding strength and specificity also applies to PNA/RNA duplexes. PNAs are resistant to enzyme degradation,[73,74] and stable over a wide pH range. Finally, their uncharged nature should make them cross through cell membranes easier, improving their therapeutic value.

1.4.2 Triplex formation with complementary DNA and RNA

Homopyrimidine peptide nucleic acids are known to form highly stable and sequence specific triplexes upon binding to complementary homopurine sites of ssDNA and dsDNA. Displacement of the second, homopyrimidine strand takes place upon binding of PNA to dsDNA, forming a P-loop.[75] The extremely high thermal stability of PNA_2/DNA triplexes,[76] is at least partly due to the charge neutrality of the PNA backbone, that excludes electrostatic repulsion from the DNA molecule, and the presence of additional hydrogen bonds between the Hoogsteen strand of the PNA and the oxygen atoms of the DNA backbone.[77] A unique property of PNAs is their ability to displace one strand of a DNA double helix to form strand invasion complexes, which is absent in DNA or any other DNA analogues studied so far.

1.4.3 Duplex formation with complementary DNA and RNA

PNAs obey Watson- Crick rules of hybridization with complementary DNA (Figure 19) and RNA. Antiparallel PNA-DNA hybrids are considerably more stable than the corresponding DNA-DNA complexes.[70] Antiparallel PNA-RNA duplexes are even more stable than both DNA-RNA hybrids and PNA-DNA duplexes. Base pair mismatches result in reduction of the T_m value by 8-20 ^oC.[72]

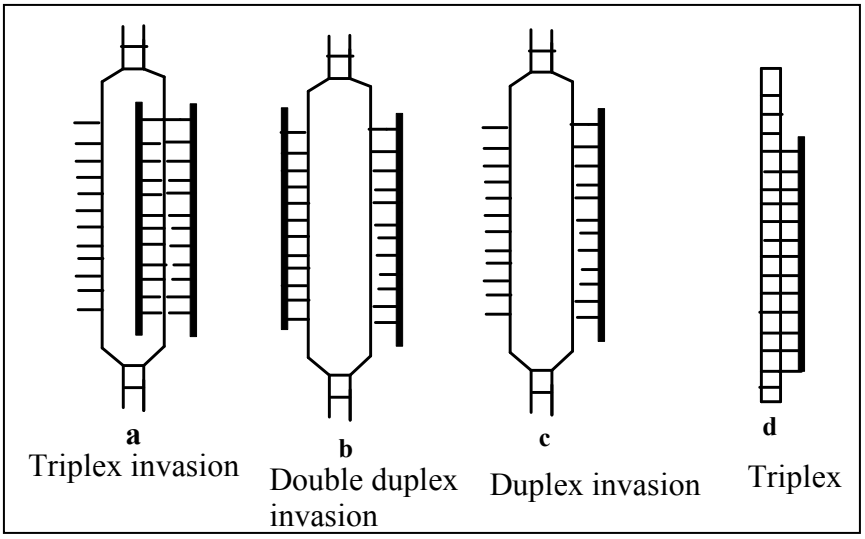


Figure 19: Schematic representation of PNA binding for targeting *ds* DNA

1.4.4 Structure of PNA complexes

Till date, the three-dimensional structures of four PNA complexes have been established. The PNA-RNA[77] and PNA-DNA[78] duplex structures were determined by NMR methods, while the structures of a PNA_2:DNA triplex[79], PNA-PNA duplex[80] and PNA-RNA duplexes were solved by X-ray crystallography (Figure 20).

1.4.4a Structure of PNA-DNA/PNA-RNA duplexes

Almost complete structural information has been deduced from the NMR spectroscopic study of two antiparallel PNA-DNA duplexes.[79] The DNA strand is in a conformation similar to the B-form, with a glycosidic *anti*-conformation, and the deoxyribose in *C*2'-*endo* form.

A more recent NMR study[78,81] 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. In case of RNA, the study revealed that the glycosidic torsion angle in the RNA strand indicates an *anti*-conformation, and the ribose sugars are in the 3'-*endo* form. The RNA strand thus resembles an A-form structure.

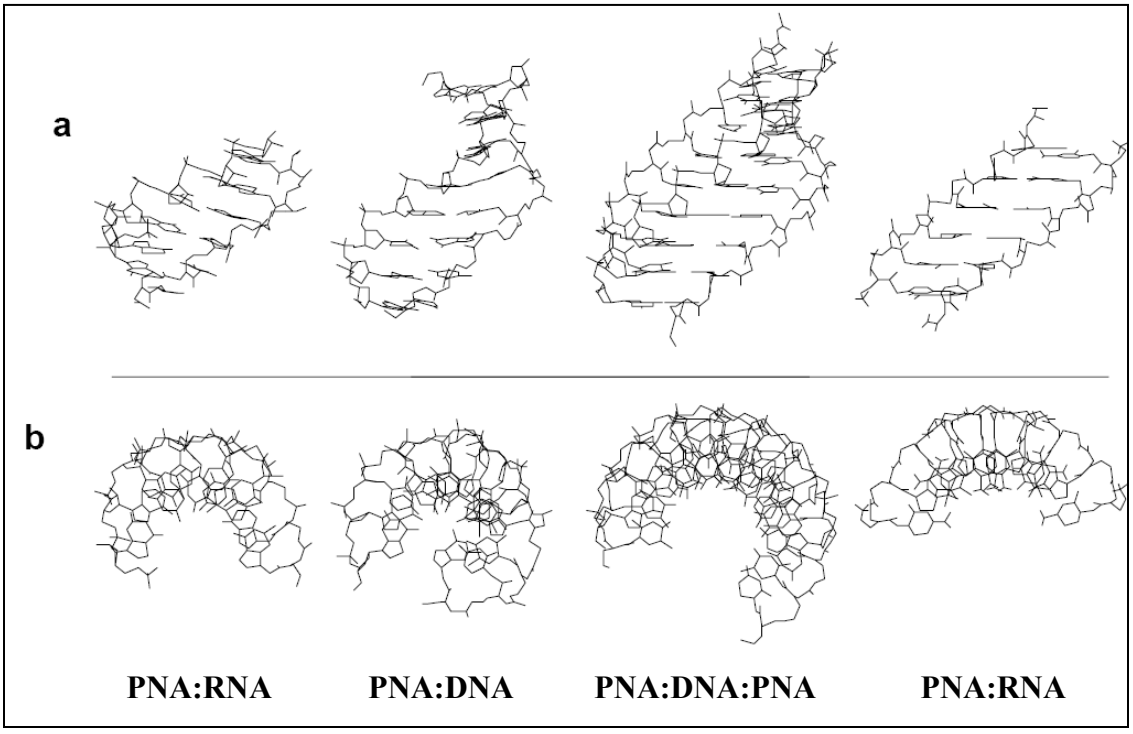


Figure 20: Structures of various PNA complexes shown in side view (upper panel) or top view (lower panel)

The CD spectra of antiparallel PNA-DNA complexes are similar to DNA-DNA spectra and indicate the formation of right handed helix.[82] The CD spectra of antiparallel PNA-RNA duplexes also indicate the formation of a right-handed helix with geometry similar to the A- or B-form.

1.4.4b Structure of PNA_2-DNA triplexes

The structure of PNA_2-DNA triple-helix was resolved by the X-ray crystal structure analysis of the complex formed by bis-(PNA) and its complementary antiparallel DNA.[79] The nucleobases of the PNA strand bind to the DNA through Watson-Crick pairing and Hoogsteen hydrogen bonding. The structure exhibited, both A-form and B-form DNA, and forms a "P-helix" with 16 bases per turn. The DNA phosphate groups are hydrogen bonded to the PNA backbone amide protons of the Hoogsteen strand. These hydrogen bonds, together with additional Van der Waal contact and the lack of electrostatic repulsion are the main factors responsible for the enormous stability of the triplex. The deoxyribose sugar adopted C3'-*endo* conformation like A-form and bases lie almost perpendicular to the helix axis, which is characteristic of B-

form DNA. The crystal structure is in agreement with the CD spectra of PNA_2-DNA triple-helices measured in solution.[82] The X-ray structure of self-complementary PNA-PNA duplex bears a strong similarity to the P-form of PNA_2-DNA triplex.[79]

1.4.5 Antigene and antisense applications of PNA

Peptide nucleic acids are promising as candidates for designing gene therapeutic drugs.[83] They require well identified targets and a well-characterized mechanism for their cellular delivery. In principle, two general strategies can be adapted to design gene therapeutic drugs.[84] Oligonucleotides or their analogs are designed to recognize and hybridize to complementary sequences in a gene wherein they would interfere with the transcription of that particular gene (antigene strategy).[85] Alternatively, nucleic acid analogs can be designed to recognize and hybridize to complementary sequences in mRNA and thereby inhibit its translation (antisense strategy).[86] PNAs are chemically and biologically stable molecules and have significant effects on replication, transcription, and translation processes, as revealed from *in vitro* experiments. Moreover, no general toxicity of PNA has so far been observed. As we shall see, PNA can interfere with the translation process, and PNA:dsDNA strand displacement complexes can inhibit protein binding and block RNA polymerase elongation.[83b]

1.4.5a Inhibition of transcription

Strand displacement complexes with DNA can create a structural hindrance to block the stable function of RNA polymerase and are thus capable of working as antigene agents. Nielsen *et al*.[87] have demonstrated that even an 8-mer PNA-T8 is capable of blocking phage T3 polymerase activity. The presence of a PNA target within the promoter region of IL-2Ra[88] gene has been used to understand the effect of PNA binding to its target on this gene expression. The PNA_2-DNA triplex arrests transcription *in vitro* and is capable of acting as an antigene agent.[83] But one of the major obstacles to applying PNA as an antigene agent is that the strand invasion or the formation of strand displacement complex is rather slow at physiological salt concentrations. Several modifications of PNA have shown improvement in terms of binding.[89]

1.4.5b Inhibition of translation

The basic mechanism of the antisense effects by oligodeoxynucleotides is considered to be either a ribonuclease H (RNase H)-mediated cleavage of the RNA strand in oligonucleotide-RNA heteroduplex or a steric blockage in the oligonucleotide–RNA complex of the translation machinery (Figure 21).[90] Oligodeoxynucleotide analogs such as phosphorothioates activate RNase H and thus hold promise of working as antisense agents. However, they also exhibit some nonspecificity in their action. PNA/RNA duplexes, on the other hand, cannot act as substrates for RNase H. The antisense effect of the peptide nucleic acid is based on the steric blocking of either RNA processing, or translation.[91] Triplex-forming PNAs are able to hinder the translation machinery at targets in the coding region of mRNA. However, translation elongation arrest requires a PNA_2:RNA triplex and thus needs a homopurine target of 10–15 bases. In contrast, duplex-forming PNAs are incapable of this. Triplex-forming PNAs can inhibit translation at initiation codon targets and ribosome elongation at codon region targets.

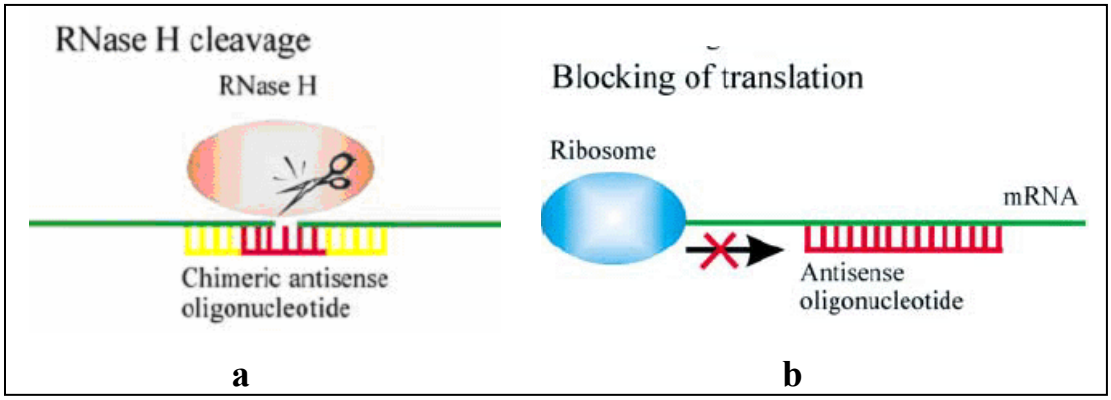


Figure 21. a) RNase H cleavage induced by (chimeric) antisense-oligonucleotides
b) Translational arrest by blocking the ribosome.

1.4.5c Inhibition of replication

It is also possible for PNA to inhibit the elongation of DNA primers by DNA polymerase. Further, the inhibition of DNA replication is feasible when the DNA duplex is subjected to strand invasion by PNA under physiological conditions or when the DNA is single stranded during the replication process.[92] Efficient inhibition of

extrachromosomal mitochondrial DNA, which is largely single-stranded during replication, has been demonstrated by Taylor et al.[93] The PNA-mediated inhibition of the replication of mutant human mitochondrial DNA is a novel (and also potential) approach towards the treatment of patients suffering from ailments related to the heteroplasmy of mitochondrial DNA. Here wild-type and mutated DNA are both present in the same cell. Experiments have shown that PNA is capable of inhibiting the replication of mutated DNA under physiological conditions without affecting the wild-type DNA in mitochondria.

1.4.6 Cellular uptake of PNA

PNAs do not readily enter cultured cells unless present at high concentrations in the media and unlike DNA/RNA, they cannot be complexed directly with cationic lipids because they are uncharged. However, Corey *et. al.*[94] have reported a novel method for *in vitro* cellular delivery of PNAs using cationic lipid. The cationic lipid is capable of associating with the negatively charged phophodiester backbone of DNA and RNA and fusion with the cell membrane allows the oligonucleotides to enter into the cell. Desired PNA oligomers are hybridized to overlapping oligonucleotides and the complex is mixed with cationic lipid. The cationic lipid-DNA-PNA complex thus formed can be internalized and the partially hybridized PNA is imported into the cell. Cellular uptake of PNAs can also be achieved by the attachment of peptide sequence that promotes translocation across cell membranes. Peptides such as Trojan peptide and penetratin have been used as carriers for cellular delivery of PNA.[95] Another strategy that has been adapted to improvise the delivery of PNA *in vitro* is to incorporate it into delivery vehicles (vesicles), e.g. liposomes. Nucleic acid- PNA chimeras are reported to be taken up even at lower extracellular concentration (1μM), so PNA-DNA chimera may be better antisense agent.[96] At higher concentrations of PNA, cytotoxic effects could also be observed.

1.4.7 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 22). The limitations of PNA include low aqueous solubility, ambiguity in DNA binding orientation and poor membrane permeability.

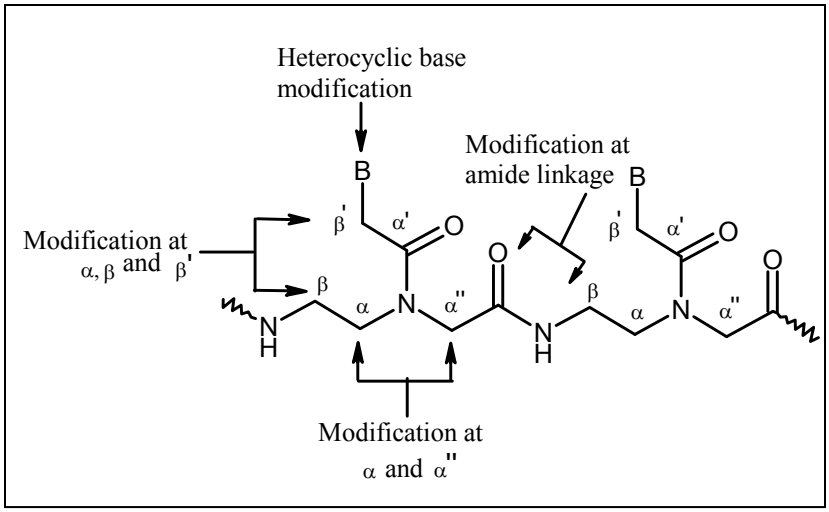


Figure 22: Structurally possible various PNA modifications

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[97] 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.

1.4.7a 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[98] (Figure 23a) or by the attachment of terminal lysine residues[99] (Figure 23c). Recently, a novel class of cationic PNA (Figure 23b) (*DNG*-PNA) analogs has been reported.[100] In these, alternating PNA /DNG chimeras, the O-(PO_2)-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.

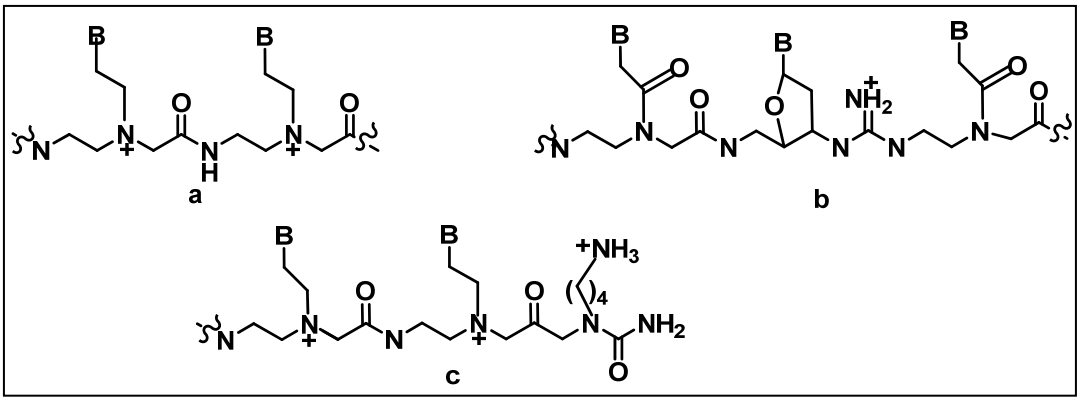


Figure 23: Positively charged PNAs a) Flexible ethylene linker, b) Guanidium linkages, c) Lysine residues

Introduction of negative charges in the PNA backbone (Figure 24 a-c) improved aqueous solubility and showed good binding with both DNA and RNA. These modifications have been well studied and reviewed.[101,102] However, many of these modified complexes were found to have less thermal stability compared to the unmodified PNA sequences. Ether linked PNAs (Figure 24d & e) showed co-operative binding with complementary antiparallel RNA in a sequence specific manner.[103,104]

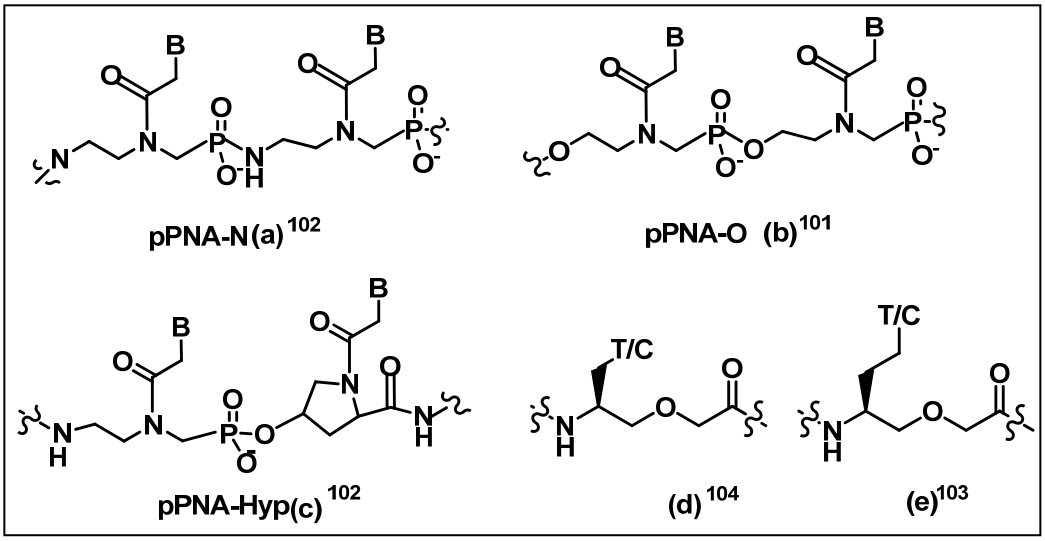


Figure 24: Anionic PNAs (a-c) and oxy PNAs (d-e)

1.4.7b 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. 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 (Figure 25). The development in this area of research are described in recent reviews.[105]

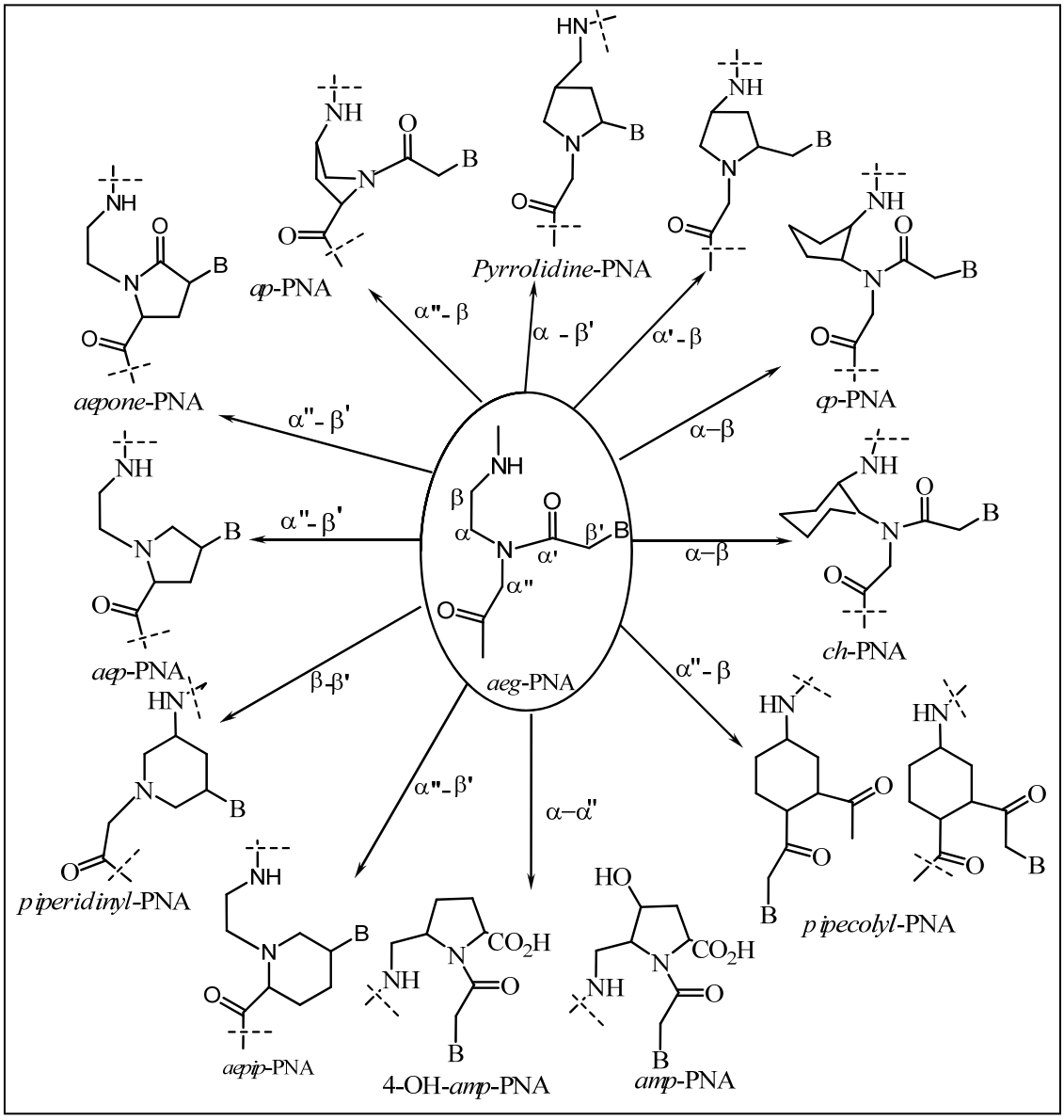


Figure 25: Conformationally constrained PNA analogues [105b]

1.4.8 BIOLOGICAL APPLICATIONS OF PNA

1.4.8a In situ hybridization (PNA-FISH)

The efficiency of PNAs as hybridization probes has also been demonstrated in fluorescence *in situ* hybridization (FISH) applications[106] (Figure 26). Because of their neutral backbone, PNA probes present *in situ* show a high specificity, require low concentrations and short hybridization times. The PNA-FISH technique was first developed for quantitative telomere analysis. Using a unique fluorescein-labelled PNA

probe, Lansdorp *et al.*[107] performed the *in situ* labelling of human telomeric repeat sequences and the data obtained allowed accurate estimates of telomere lengths. Subsequently, telomeric PNA probes were used in several *in situ* studies of cancer and ageing.

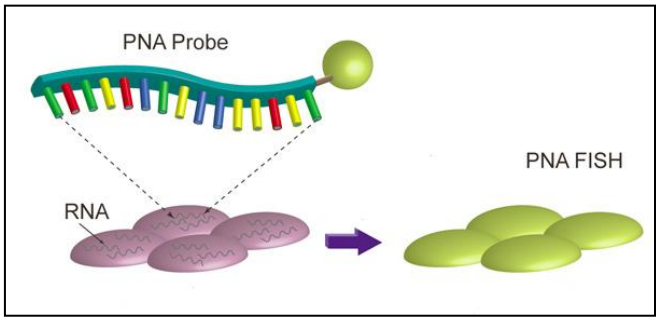


Figure 26: PNA FISH Technology

1.4.8b Solid-phase hybridization techniques

PNAs can be used in many of the same hybridization applications as natural or synthetic DNA probes but with the added advantages of tighter binding and higher specificity. This leads to faster and easier procedures in most standard hybridization techniques.[108]

1.4.8c PCR and Q-PNA PCR

PNA probes have no direct interaction with DNA polymerase but PNAs can terminate the elongation of oligonucleotide primers by binding to the template or competing with the primers.[109] Furthermore, PNA-DNA chimeras can be recognized by the DNA polymerase and can thus be used as primers for PCR reactions.[110] The high affinity binding of PNAs has also been used for detecting single base pair mutations by PCR. This strategy, named PNA directed PCR clamping (Figure 27), uses PNAs to inhibit the amplification of a specific target by direct competition of the PNA targeted against one of the PCR primer sites and the conventional PCR primer. This PNA-DNA complex formed at one of the primer sites effectively blocks the formation of the PCR product. The procedure is so powerful that it can be used to detect single base-pair gene variants for mutation screening and gene isolation. More recently, novel automated real-time PCR has been developed using PNAs. In this method, named Q-PNA PCR, a

generic quencher labelled PNA (QPNA)[111] is hybridized to the 5'-TAG sequence of a fluorescent dye-labelled DNA primer in order to quench the fluorescence of the primer. During PCR, the Q-PNA is displaced by incorporation of the primer into amplicons and the fluorescence of the dye label is liberated.

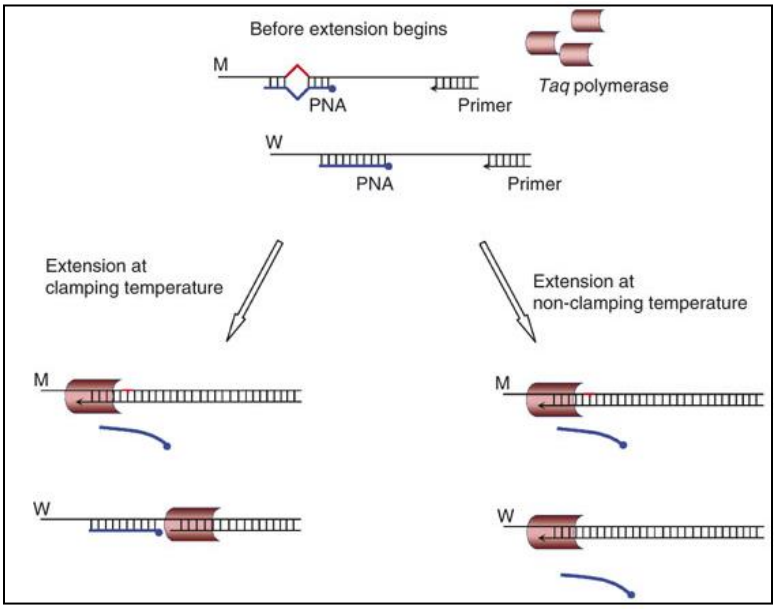


Figure 27: PNA directed PCR clamping

1.4.8d Anti-cancer agent

PNA-peptide:DNA duplexes, which can penetrate into cells, have been used in anti-cancer applications.[112] In this manner, telomerase activity in human melanoma cells and tumour specimens was inhibited by PNA conjugated with Antennapedia derived peptide (Antp) at nanomolar concentrations. Since telomerase is almost ubiquitously expressed in human tumours, the data point out the potential use of PNAs as anticancer drugs.[113] Applications of PNAs as anticancer agents were also reported with PNA complementary to various sequences of bcl-2.

1.4.8e PNA as delivery agents

A major limitation of non-viral gene therapy is the low efficiency of gene transfer into target cells. PNAs can be used as adapter to link peptides, drugs or molecular tracer to plasmid vectors.[114] According to the binding site, the coupling of PNAs to plasmids has no effect either on the transcription of genes included in the

plasmid or on the plasmid's physiological activities. Thus, this approach allows circumventing such barrier to gene transfer and fixing drugs to plasmid in order to enhance the gene delivery or tissuespecific targeting. Using a triplex forming PNA as linker, Brandén *et. al*.[115] observed an eight times higher nuclear localization signal than did the free oligonucleotide.

1.4.9 BIOTECHNOLOGICAL APPLICATIONS OF PNA

1.4.9a Enhanced PCR amplification

The polymerase chain reaction (PCR) has been widely used for various molecular genetic applications including the amplification of variable number of tandem repeat (VNTR)[116] loci for the purpose of genetic typing. PNA has been used to achieve an enhanced amplification of VNTR locus D1S80. Small PNA oligomers are used to block the template, and the latter becomes unavailable for intra- and interstrand interaction during reassociation. On the other hand, the primer extension is not blocked; during this extension, the polymerase displaces the PNA molecules from the template and the primer is extended toward completion of reaction. This approach shows the potential of PNA application for PCR amplification where fragments of different sizes are more accurately and evenly amplified. Since the probability of differential amplification is less, the risk of misclassification is greatly reduced.

1.4.9b PNA-assisted rare cleavage

Peptide nucleic acids, in combination with methylases and other restriction endonucleases, can act as rare genome cutters. The method is called PNA-assisted rare cleavage (PARC) technique[117] (Figure 28). It uses the strong sequence-selective binding of PNAs, preferably bis-PNAs, to short homopyrimidine sites on large DNA molecules. The PNA target site is experimentally designed to overlap with the methylation / restriction enzyme site on the DNA, so a bound PNA molecule will efficiently shield the host site from enzymatic methylation whereas the other, unprotected methylation/restriction sites will be methylated. After the removal of bis-PNA, followed by restriction digestions, it is possible to cleave the whole DNA by enzymes into limited number of pieces.

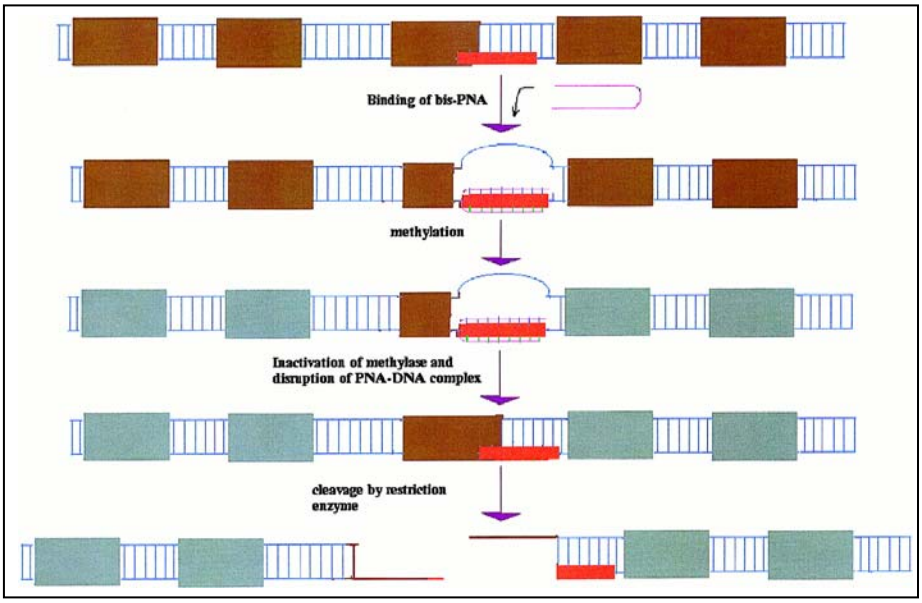


Figure 28: PNA assisted rare cleavage

1.4.9c Determination of telomere size

The conventional method for the determination of telomere length involves Southern blot analysis of genomic DNA and provides a range for the telomere length of all chromosomes present. The modern approach uses fluorescein-labeled oligonucleotides and monitor *in situ* hybridization to telomeric repeats. However, Lansdorp *et al*.[118] shown a more delicate approach resulting in better quantitative results is possible by using fluorescein-labeled PNAs. This PNA-mediated approach permits accurate estimates of telomeric length. *In situ* hybridization of fluorescein-labeled PNA probes to telomeres is faster and requires a lower concentration of the probe compared to its DNA counterpart.

1.4.9d Nucleic acid purification

Based on its unique hybridization properties, PNAs can also be used to purify target nucleic acids. PNAs carrying six histidine residues have been used to purify target nucleic acids using nickel affinity chromatography.[119] Thus, short PNAs can also be used as generic capture probes for purification of large nucleic acids. It has been shown that a biotin tagged PNA-thymine heptamer could be used to efficiently purify human genomic DNA from whole blood by a simple and rapid procedure.

1.4.9e PNAs as artificial restriction enzymes

PNAs in combination with a non-specific nuclease, such as S1 nuclease, have been used as artificial restriction enzymes to cut target DNA at desired positions.[120] Double stranded DNA is cleaved at a site created by PNA strand displacement (Figure 29). This cleavage efficiency is enhanced more than 10 fold when a tandem PNA site is targeted, and additionally if this site is in *trans* rather than *cis* orientation. Thus, the single strand specific nuclease S1 behaves like a pseudo restriction endonuclease. The tethering of an artificial nuclease like Gly-Gly-His to PNA exhibits a similar effect. Binding to complementary DNA and in the presence of Ni, cleaves the DNA duplex in its *proximity.*

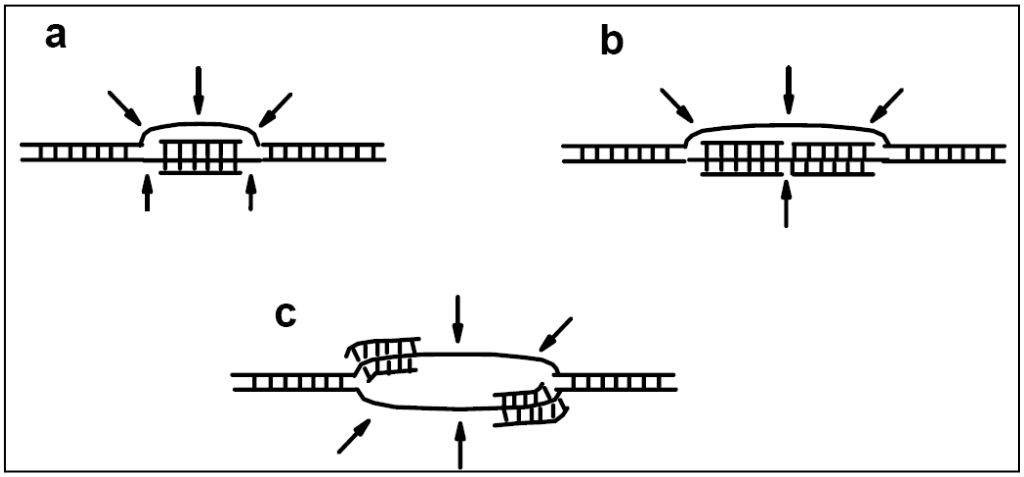


Figure 29: Schematic model for PNA-targeted S1 nuclease ds cleavage of DNA. **a** Single target . **b** Double target in *cis* orientation. **c** Double target in *trans* orientation. Arrows indicate S1 attack.

1.4.9f MALDI-TOF mass spectrometry

MALDI-TOF mass spectrometry has been used successfully in PNA-based diagnostic research to study discrimination of single-nucleotide polymorphisms (SNPs) in human DNA. Human genomic and mitochondrial DNA contains many SNPs that may be linked to diseases. Rapid and accurate screening of important SNPs, based on high-affinity binding of PNA probes to DNA, is possible by using MALDI-TOF mass spectroscopy.[121]

1.4.9g PNA microarray

The basic requirement for the array system is the ability of all different probes to hybridize to their target sequences with high specificity at single temperature. As PNAs are neutral, they can hybridize with nucleic acids in absence of counter ions needed to stabilize pure nucleic acid duplexes. It has been demonstrated that PNA probes can effectively discriminate between single base mismatches in the target sequence at temperature near the optimum for hybridization.[122] They can also be used to analyze and roughly quantify the amount of target molecules over a considerable concentration range. PNA arrays can be reused much more often than conventional oligonucleotide array, since the PNA molecules are extremely stable under conditions, which natural DNA cannot withstand.

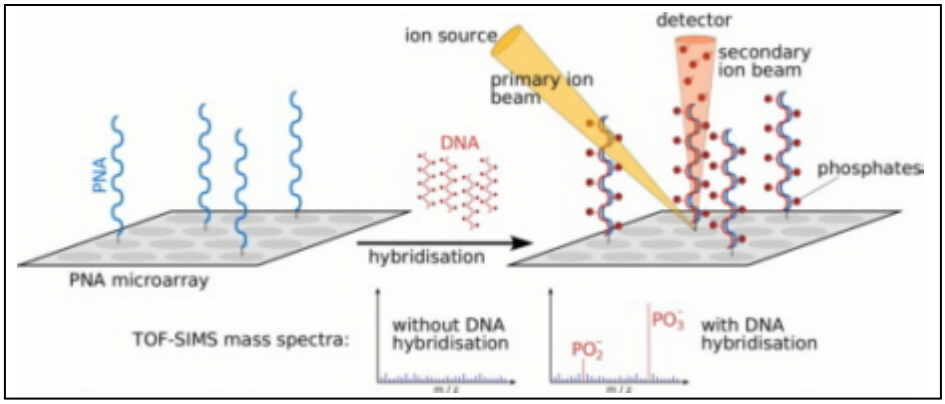


Figure 30: PNA microarray

1.4.9h Antiviral PNAs

One of the key enzymes in the life cycle of retroviruses (such as HIV), reverse transriptase, is very sensitive to PNA "Antisense Inhibition'. Reverse transcription of the RNA template is effectively arrested by PNA oligomers bound to the template.[123] This finding has raised hope that DNA antiviral drugs could be developed with the demonstration that HIV replication in cell culture can be inhibited by PNA targeting *agpol* gene. However, very high PNA concentrations were required emphasizing the need of an efficient cell delivery system for PNA.[124]

1.4.9i Antibacterial PNAs

Inhibiting translation through directly interfering with the ribosome will shut down all protein synthesis, providing a very potent antibiotic strategy. Good and Nielsen[125] demonstrated that PNAs targeted to two regions of ribosomal RNA, the alpha-sarcin loop and the peptidyl transferase center, were capable of causing cell death in bacterial colonies. Homopyrimidine bisPNAs were more effective than the monoPNAs, indicating that triplex formation was important for the observed antibiotic activity. It has been reported that 9-12 mer PNAs attached to the cell wall/membrane–active peptide (KFF)3K,[126] provide improvement in antisense potency in *E. Coli* amounting to two orders of magnitude while retaining target specificity. Peptide-PNA conjugates targeted to rRNA and to mRNA encoding the fatty acids biosynthesis protein Acp, prevented cell growth of *E. coli* K12 without any apparent toxicity to human cells. This indicates that the peptide can be used to carry antisense PNA agents into bacteria. Such peptide-PNA conjugates open exciting possibilities for anti-infective drug development and provides new tools for microbial genetics. These results bear promise that PNA could be developed as 'generic antibiotics'.

1.5 PRESENT WORK:

The preceding sections give an overview of the peptide nucleic acids (PNAs), the synthetic imitator of natural nucleic acids that is DNA analogues with a homomorphous but chemically different backbone consisting of *N*-(2-aminoethyl)-glycine 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 potential 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 involves the synthesis and characterization of novel backbone modified, chiral, ring constrained PNA analogues: *diamino pyrrolidinyl* PNA (*dap*PNA), which is an Aza-analogue of cyclopentyl PNA and *amp*PNA derived from 3-(aminomethyl)-pyrrolidin-4-amine, having one carbon extended backbone at C-3 position of *dap*PNA (Figure 31). This would be interesting to investigate structural-biophysical activity relationship of these analogues. This chapter describes the synthesis of modified PNA monomers.

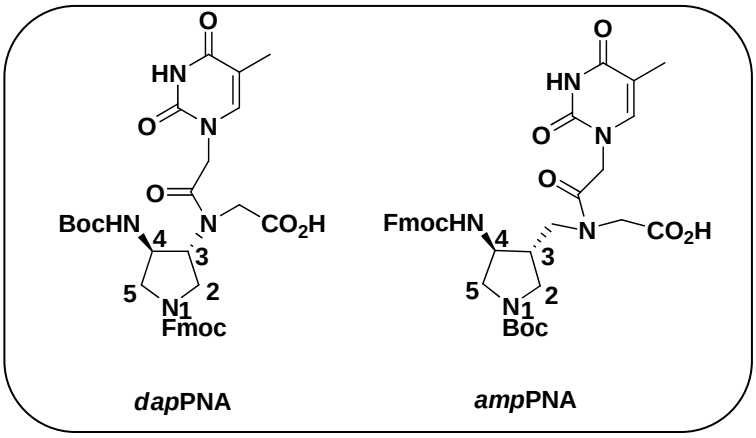


Figure 31: Chemical structure of modified PNA monomers

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 circular dichroism spectroscopy (CD spectroscopy) and UV-mixing curves.

Chapter-4 deals with synthesis of new pyrrolidine analogues with potential medicinal and biological applications. The past work from this laboratory has resulted in the synthesis of several pyrrolidine PNAs based on proline structures having substitution α-to ring nitrogen, one of the examples being pyrrolidine PNA shown in Figure 32 which has a base linker at C-2 position. In this chapter, it is explored to shift the position of base linker from C-2 to C-3 carbon to generate new series of PNA analogues.

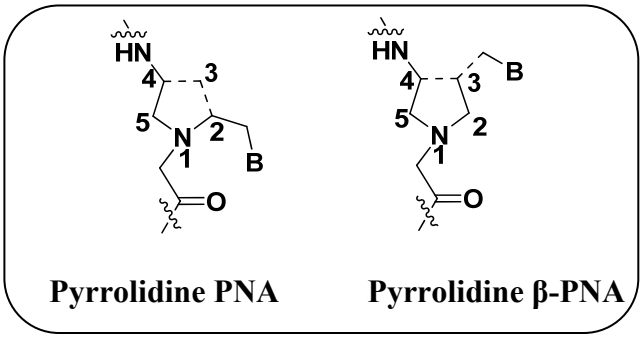


Figure 32: Pyrrolidine and Pyrrolidine β-PNA

In addition, A convenient and scalable route to five and six membered heterocyclic β-ketoesters (Figure 33a) has been carried out. These compounds have previously received attention as precursors for the synthesis of a variety of medicinally and biologically interesting molecules. The five membered keto ester (Figure 33a, n=1) was converted to the N-protected (3*S*,4*R*) 4-amino-β-proline (Figure 33b) which can be used as a precursor, instead of hydroxy proline, to study the effect on the stability of collagen model peptides.

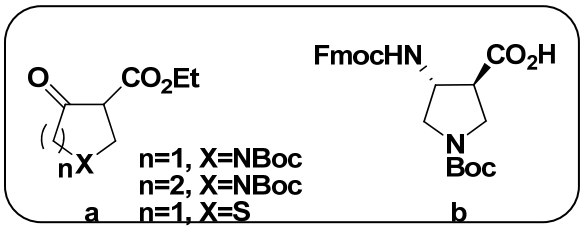


Figure 33: a) β-ketoesters b) Unnatural amino acid 4-amino-β-proline

1.6 Reference

1. Takahashi, I. and Marmur, J. *Nature,* **1963**, *197*, 794.
2. Watson, J. D. and Crick, F. H. C. *Nature*, **1953**, *171*, 737.
3. (a) Hoogsteen, K. *Acta. Cryst*., **1963**, *65*, 907. (b) Crick, F. H. C. *J. Mol. Biol*., **1966**, *19*, 548.
4. Seeman, N. C., Rosenberg, J. M. and Rich, A. *Proc. Natl. Acad. Sci. USA*., **1976**, *73*, 804.
5. Zamechnik, P.C.; Stephenson, M. L. *Proc*. *Natl*. *Acad*. *Sci*. **1978**, *75*, 280
6. (a) Felsenfeld, G.; Davies, D. R.; Rich, A. *J*. *Am*. *Chem*. *Soc*. **1957**, *79*, 2023. (b) Moser, H. E.; Dervan, P. B. *Science* **1987**, *288*, 645. (c) Beal, P.A.; Dervan, P. B. *Science* **1991**, *251*, 1360.
7. Soyfer, V. N.; Potamann, V. N. *Triple Helical Nucleic Acids* **1996**, Springer-Verlag, NY.
8. Chan, P. P.; Glazer, P. M. *J. Mol*. *Med*. **1997**, *75*, 267.
9. Herdewijn, P. *Antisense Nucleic Acid Drug Dev.* **2000**, *10*, 297.
10. Eckstein, F. *Antisense Nucleic Acids Drug Dev.* **2000**, *10*, 117.
11. a.) Stein, C. A. and Cohen, J. S Phoshorothioate oligodeoxynucleotide analogues. In Cohen, J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 97. b.) Ghosh, M. K.; Ghosh, K.; Dahl, O.; Cohen, J. S. *Nucleic Acids Res.* **1993**, *21*, 5761.
12. De Clercq, E.; Eckstein, F.; Merigan, T. C. *Science,* **1969**, *165*, 1137.
13. Dueholm, K. L.; Egholm,M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius T.;Petersen, K. H.; Berg, R.; H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767.
14. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.
15. Puglisi, J. D.; Tinoco, I. *Methods Enzymol.* **1989**, *180*, 304.
16. Stein, C. A. *Nat. Med*., **1995**, *1*, 1119.
17. Summerton, J., Stein, D., Huang, S. B., Matthews, P., Weller, S., Partridge, M. *Antisense Nucleic Acid Drug Dev*., **1997**, *7*, 63.

18. Millar, P. S. Non-ionic antisense oligonucleotides, In Cohen, J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 79.
19. Checko, K. K., Linder, K., Saenger, W. and Miller, P. S. *Nucleic Acid Res.*, **1983**, *11*, 2801.
20. Miller, P. S., Yano, J., Yano, E., Carroll, C., Jayaraman, K. *Biochemistry*, **1979**, *18*, 5134.
21. Kean, J. M., Kipp, S. A., Miller, P. S., Kulka, M., Aurelian, L. *Biochemistry*, **1995**, *34*, 14617.
22. Froehler, B., Ng, P. and Matteucci, M. *Nucleic Acid Res*., **1988**, *16*, 4831.
23. Letsinger, R. L., Singman, C. N., Histand, G. and Salunke, M. *J. Am. Chem. Soc*., **1988**, *110*, 4470.
24. Jung, P. M., Histand, G. and Letsinger, R. L. *Nucleosides Nucleotides*, **1994**, *13*, 1597.
25. (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.
26. Li, H., Huang, F., Shaw, B. R. *Bioorg. Med. Chem*., **1997**, *5*, 787.
27. Sergueev, D. S. and Shaw, B. R. *J. Am. Chem. Soc*., **1998** , *120*, 9417.
28. a.) Higson, A. P., Sierzchala, A., Brummel, H., Zhao, Z. and Caruthers, M. H. *Tet. Lett*., **1998**, *39*, 3899. (b) Rait, V. K. and Shaw, B. R *Antisense Nucleic Acid Drug Dev*., **1998**, *9*, 53.
29. Mahajna, M.; Quistad, G. B.; Casida, J. E. *Chem. Res. Toxicol*. **1996**, *9*, 1202.
30. (a) Chin, D. J.; Green, G. A.; Zon, G.; szoka, Jr., F. C.; Straubinger, R. M. *New Biologist* **1990**, *2*, 1091. (b) Leonetti, J. P.; Mechti, N.; Degols, G.; Gagnor, C.; Lebleu, B. *Proc. Natl. acad. Sci. USA.* **1991**, *88*, 2702.
31. Rosler, A.; Peleiderer, W. *Helv. Chim. Acta*. **1997**, *80*, 1869.
32. Ozaki, H.; Ogawa, Y., Mine, M., Sawai, H. *Nucleosides Nucleotides*, **1998,** *17*, 911.
33. Herdewijn, P. *Antisense Nucleic Acid Drug Dev.* **2000**, *10*, 297

34. (a) Wanger, R. W.; Matteucci, M. D.; Lewis, J. G.; Gutierrez,A. J.; Moulds, C.; Froehler, B.C. *Science*, **1993,** *260*, 1510. (b) Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S.R. *Tetrahedron Lett*. **1992**, *33*, 5307.
35. Buhr, C. A.; Wanger, R. W.; Grant, D.; Froehler, B. C. *Nucleic Acids Res.* **1996**, *24,* 2974.
36. Gutierrez, A. J.; Froehler, B. C. *Tetrahedron Lett.* **1996**, *37*, 3959.
37. Gutierrez, A. J.; Terjprst, T. J.; Matteuci, M. D.; Froehler, B.C. *J. Am. Chem. Soc.* **1994**, *116*, 5540.
38. (a) Inoue, H.; Imura, A.; Ohtsuka, E. *Nucleic Acids Res*. **1985**,*13*, 7119. (b) Lin, K. Y.; Jones, R. J.; Matteucci, M. *J. Am. Chem. Soc.* **1995**, *117*, 3873.
39. Buhr, C. A.; Matteucci, M. D.; Froehler, B. C. *Tet*. *Lett*. **1999**, *40*, 8969.
40. Lee, J. S.; Woodsworth, M. L.; Latimer, L. J. P. Morgan, A. R. *Nucleic Acids Res.* **1984**, *12,* 6603.
41. Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059.
42. Schmid, N.; Behr, J. P. *Tetrahedron Lett.* **1995**, *36*, 1447.
43. Haginoya, N.; Ono, A.; Nomura, Y.; Ueno, Y.; Matsuda, A. *Bioconj. Chem.* **1995**, *8*, 271.
44. (a). Barawkar, D. A.; Rajeev, K. G.; Kumar, V. A.; Ganesh, K. N. *Nucleic Acids Res.* **1996**, *24*, 1229. (b) Seela, F.; Zulauf, M. *Helv. Chim. Acta.* **1999**, *82*, 1878.
45. Adam V.; Sproat B.S. *J. Chem. Soc. Perkin Trans.* **1994**, *1*, 3423.
46. Rosenberg. I.; Tocik. Z.; Watanabe. K. A. *Nucleic Acids Res*. **1991,** *24*, 43.
47. Rosenberg. I.; Farras J. S.; Tocik Z.; Ren W. Y.; Ciszewski. L. A.; Kois. P.; Pankiewicz. K. W.; Spassova. M.; Watanabe. K. A. *Nucleosides Nucleotides* **1993**, *12*, 38 l.Ono. A.; Tso. P.O.P.; Kan L. S. *J. Org. Chem.* **1992**, *57*, 3225.
48. Ono. A.; Tso. P.O.P.; Kan L. S. *J. Org. Chem.* **1992**, *57*, 3225.
49. Rana. V. S.; Ganesh. K. N. *Nucleic Acids Res*. **2000**, 28, 1162.
50. Chattopadhyaya, J. *Tetrahedron*, **1991,** *47,* 4693.
51. Kumar R. K.; Gunjal A. D.; Ganesh K. N. *Biochem. Biophys. Res. Comm*. **1994**, *204*,788.
52. Kawai, K.; Saito, I.; Sugiyama, H**.** *Tetrahedron Lett.* **1998,** *39*, 5221.

53. Soliva, R.; Garcia, G. R.; Blas, J. R.; Eritja, R.; Asensio, J. L.; Gonza´ lez, C.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **2000**, *28*, 4531.
54. Cubero, E.; Avino, A.; de la Torre, B. G.; Frieden, M.; Eritja, R.; Luque, F. J.; Gonzalez, C.; Orozco, M. *J. Am. Chem. Soc.* **2002**, *124,* 3133.
55. (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.
56. Moulds, C. *et al*: *Biochemistry*, **1995**, *34*, 5044.
57. Monia, B. P. *et al*: *J. Biol. Chem*. **1993**, *268*, 14514.
58. Kawasaki, A. M. *et al*: *J. Med. Chem*, **1993**, *36*, 831.
59. De Winter, H., Lescrinier, E., Van Aerschot, A. and Herdewijn, P. *J. Am. Chem. Soc.*, **1998**, *120*, 5381.
60. Maier, T., Przylas, I., Strater, N., Herdewijn, P. and Saenger, W. *J. Am. Chem. Soc.*, **2005**, *127*, 2937.
61. Schöning, K.-U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R. and Eschenmoser, A. *Science*, **2000**, 1347.
62. Pallan, P. S., Wilds, C. J., Wawrzak, Z., Krishnamurthy, R., Eschenmoser, A. and Egli, M. *Angew. Chem Int. Ed. Engl*., **2003**, *42*, 5893.
63. Sabatino, D., Damha, M. J. *J. Am. Chem. Soc.*, **2007**, *129,* 8259.
64. (a) Petersen, M.; Wengel, J. *Trends Biotechnol*. **2003**, *21*, 74. (b) Koshkin, A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998**, *54*, 3607. (c) Singh, S. K.; Nielsen, P.; Koshkin, A. A.; Wengel, J. *Chem. Commun*. **1998**, 455.
65. (a) Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J. I.; Mori, K. I.; Doi, T.; Imanshi, T. *Tetrahedron Lett*. **1998**, *39*, 5401. (b) Hertoghs, K. M.; Ellis, J. H.; Catchpole, I. R. *Nucleic Acid Res*. **2003**, *20*, 5817.
66. Kurreck, J.; Wyszko, E.; Gillen,C.; Erdmann, V. A. *Nucleic Acids Res*. **2002**, *30,*1911.
67. Bondensgaard, K.; Petersen,M.; Singh,S .K.; Rajwanshi,V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P. *Chem. Eur. J.* **2000**, *6*, 2687.

68. (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.
69. Toyler, M. F. *et al Antisense Nucleic Acid Drug Dev*., **1998**, *8*, 199.
70. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497.
71. (a) Ray. A.; Norden. B. *FASEB J*, **2000**, *14*, 1041. (b)Corradini. R.; forza, S. S.; Tedeschi, T.; Totsingan. F.; Marchelli. R. *Curr. Top. Med. Chem.* **2007**, *7,* 681.
72. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R.H.; Kim, S. K.; Nordon, B.; Nielsen, P. E. *Nature*, **1993**, *365*, 566.
73. Jensen, K. K.; Qrum, H.; Nielsen, P. E.; Norden, B. *Biochemistry* **1997**, *36*, 5072.
74. Protozanova, E.; Demidov, V. V.; Nielsen, P. E.;Kamenetskii. M. D. F. *Nucleic Acids Res.* **2003**, *31*, 3929.
75. Møllegaard, N. E.; Buchardt, O.; Egholm, M.; Nielsen. P. E. *Proc. Natl. Acad. Sci.* **1994**, USA *91*, 3892.
76. (a) Jensen, K. K.; Qrum, H.; Nielsen, P. E.; Norden, B. *Biochemistry*, **1997**, *36*, 5072. (b) Kuhn, H.; Demidov, V. V; Nielsen, P. E.; Frank-Kamenetskii, M. D. *J. Mol. Biol*.**1999**, *286*, 1337.
77. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis. D. G. *Science* **1994**, *265*, 777.
78. Eriksson, M.; Nielsen, P. E. *Nature Struct. Biol*. **1996**, *3*, 410.
79. Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838.
80. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P.E. *Nature Struct. Biol*. **1997**, *4*, 98.
81. Leijon. M.; Gräslund. A.; Nielsen, P. E.; Buchardt, O.; Nordén, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry* **1994**, *33*, 9820.
82. Kim, S. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Nordén, B. *J. Am. Chem. Soc.* 1993, *115*, 6477.
83. (a) Knudsen, H.; Nielsen, P. E. *Nucleic Acid Res*. **1996**, *24*, 494. (b) Bentin, T.; Larsen, H. J.; Nielsen, P. E. *Biochemistry* **2003**, *42*, 13987.
84. Jensen, K. K.; Orum, H.; Nielsen, P. E., Nordén, B. *Biochemistry* **1997**, *36*, 5072.

85. (a) Praseuth, D.; Guieysse, A. L.; Helene, C. *Biochim. Biophys. Acta*. **1999**, *489*, 181. (b) Helene, C.; Thuong, N. T.; Harel-Bellan, A. *Ann. N. Y. Acad. Sci*. **1992**, *660*, 27.

86. Noble, S. A.; Bonham, M. A.; Bisi, J. E.; Bruckenstein, D. A.; Brown P. H.; Brown, S. C *et al*.: *Drug Dev*. *Res*. **1995**, *34*, 184.

87. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Anticancer Drug Des*. **1993**, *8*, 53.

88. Dieci, G.; Corradini, R; Sforza, S.; Marchelli, R.; Ottonello, S. Inhibition of RNA polymerase III elongation by a T10 peptide nucleic acid. *J*. *Biol*. *Chem*. **2001**, *276*, 5720.

89. (a) Uhlmann, E.; Pcyman, A.; Brcipohl, G.; Will, D. W. *Angew Chem. Int. Ed*. **1998**, *37*, 2796. (b) Mesmaeker, A.; Waldner A.; Lebreton, J.; Hoffmann, P.; Fritsch, V.; Wolf, R. M. *Angew. Chem.* **1994**, *106,* 237.

90. Walder, R. Y.; Walder, J. A. *Proc. Natl. Acad. Sci.U.S.A,* **1988**, *85*, 5011.

91. Bonham, M. A.; Brown, S.; Boyd, A. L. Brown, P. H.; Bruckenstein, D. A.; Hanvey, J. C. *Nucleic Acids Res* **1995**, *23*, 1197.

92. Diviacco, S.; Rapozzi, V.; Xodo, L.; Helene, C.; Quadrifoglio, F.; Giovannangeli, C. *FASEB J*. **2001,** *15,* 2660.

93. Taylor, R. W.; Chinnery, P. F.; Turnbull, D. M.; Lightowlers, R. N. *Nat. Genet.* **1997**, 212.

94. Hamilton, S. E.; Simmons, C. G, Kathiriya I. S, Corey D. R. *Chem. Biol*. **1999**, *6*, 343.

95. Igloi, G. L. *Proc. Natl. Acad. Sci. USA* **1998**, *95***,** 8562.

96. (a) Rose, D. J. *Anal. Chem.* **1993**, *65***,** 3545. (b) Hoffmann R, Ertl HCJ, Pease AM, He Z, Rovera G, Otvos L, Jr.: Low scale multiple array synthesis and DNA hybridization of peptide nucleic acids. In: *Peptides: Frontiers of Peptide Science, Proceedings of the American Peptide Symposium, 15th, Nashville, June 14-19, 1997*.

97. Dueholm, K. L.; Nielsen, P. E. *New J. Chem.* **1997**, *21*, 19.

98. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993**, 518.

99. (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.

100. Barawkar, D. A.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 10418.

101. Efimov, V. A.; Choob, M. V.; Buryakova, A. A.; Chakhmakhcheva, O. G. *Nucleos. Nucleot. Nucl.* **1998**, *17*, 1671.

102. Efimov, V. A.; Buryakova, A. A.; Choob, M. V.; Chakhmakhcheva, O. G. *Nucleos. Nucleot. Nucl.* **1999**, *18*, 1393.

103. Kuwahara, M.; Arimitsu, M.; Shisido, M. *Tetrahedron* **1999**, *55*, 10067.

104. Altmann, K. H.; Chiese, C. S.; Garcia-Echeverria, C. **1997**, *7*, 1112.

105. (a) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404. (b) Kumar, V. A. *Eur. J. Org. Chem.* **2002**, 2021.

106. Chen, C.; Hong, Y. K.; Ontiveros, S. D.; Egholm, M.; Strauss, W. M. *Mammalian Genome* **1999**, *10*, 13.

107. Lansdorp, P. M. Close encounters of the PNA kind. *Nature Biotechnology* **1996**, *14*, 1653.

108. a.) Stender, H.; Mollerup, T. A.; Lund, K.; Petersen, K. H.; Hongmanee, P.; Godtfredsen, S. E. *Int. J. Tuberc. Lung. Dis.* **1999**, *3*, 830. b.) Larsen, R. D.; Schonau, A.; Thisted, M.; Petersen, K. H.; Lohse, J.; Christensen, B.; Philip, J.; Pluzek, K. J. *Prenat Diagn*. **2003**, *23*, 52.

109. Ørum, H. PCR clamping. *Curr. Issues. Mol. Biol.* **2000**, *2*, 27.

110. Drewe, L. J.; Brightwell, G.; Hall, E. A. *Mol Cell Probes* **2000**, *14*, 269. Isacsson, J.; Cao, H.; Ohlsson, L.; Nordgren, S.; Svanvik, N.; Westman, G. *Molecular and Cellular Probes* **2000**, *14*, 321.

111. Fiandaca, M. J.; Hyldig-Nielsen, J. J; Gildea, B. D.; Coull, J. M. *Genome Res.* **2001**, *11,* 609.

112. Kelland, L. R. *Anti-Cancer Drugs* **2000**, *11*, 503.

113. (a) Knudsen H, Nielsen PE. *Anticancer Drugs.* **1997**, *8*, 113. (b) Taneja, K. L.; Chavez E. A.; Coull, J.; Lansdorp, P. M. *Genes, Chromosomes & Cancer* **2001**, *30*, 57.

114. Kim, D-. H.; Hong, Y-. K.; Egholm, M.; Strauss, W. M. *Bio. Techniques* **2001**, *31*, 475.

115. Brandén, L. J.; Christensson, B.; Smith, C. I. E. *Gene Therapy* **2001**, *8*, 84.

116. Demers, D. B.; Curry, E. T.; Egholm, M.; Sozer, A. C. *Nucleic Acids Res*. **1995**, *23*, 3050.

117. Veselkov, A. G.; Demidov, V.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **1996**, *24,* 2483.

118. Lansdorp, P. M.; Verwoerd, N. P.; Van de Rijke, F. M.; Dragowska, V.; Little, M. T.; Dirks, R. W. *Human Molecular Genetics* **1996**, *5*, 685.

119. Isacsson, J.; Cao, H.; Ohlsson, L.; Nordgren, S.; Svanvik, N.; Westman, G. *Molecular and Cellular Probes* **2000**, *14*, 321.

120. Buryakova, A.; Phelan, D.; Chakhmakhcheva, O. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 419.

121. Ross, P. L.; Lee, K.; Belgrader, P. *Anal. Chem.* **1997** *69,* 4197.

122. Jain, K. K. *Pharmacogenomics* **2000**, *1*, 289.

123. Yamasaki, T.; Abdel-Aziz, M.; Kiyota, N.; Maruyama, T.; Otsuka, M. *Heterocycles* **2003**, *60*, 1561.

124. Igloi, G. L. *Proc. Natl. Acad. Sci. USA* **1998**, *95***,** 8562.

125. Good, L.; Nielsen, P. E. *Proc. Natl. Acad. Sci.* USA **1998**, *95*, 2073.

126. Romanelli, A.; Pedone, C.; Saviano, M.; Bianchi, N.; Borgatti, M.; Mischiati, C.; Gambari, R. *Eur. J. Biochem*. **2001**, *268,* 6066.

CHAPTER 2

Synthesis of 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.[1] When introduced into cells, these analogues seek complementary sequences on target DNA duplex (antigene) or single stranded mRNA (antisense) and bind to inhibit the production of corresponding coded functional proteins in a cell.[2] 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.[3] Among several modifications peptide nucleic acid designed by Nielsen *et al* [4] using computer modelling have emerged as one of the potentially effective nucleic acid analogue. In Peptide nucleic acid (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 which allows its binding to the target DNA/RNA sequences with high sequence specificity and affinity.[5] Moreover, PNA is stable to cellular enzymes like nucleases and proteases. However, major limitations of 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.[6]

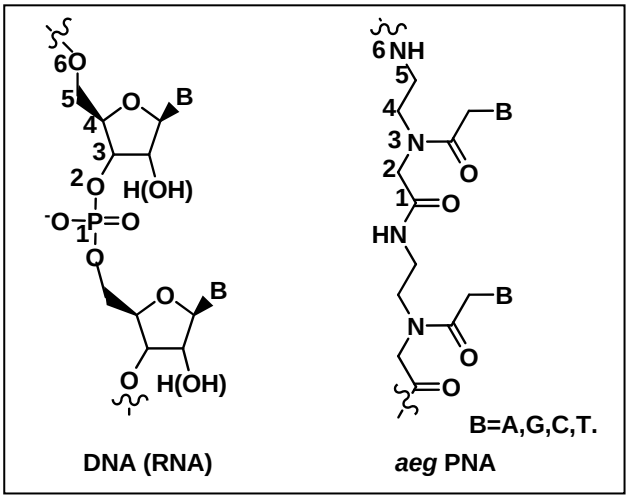


Figure 1: DNA (RNA) and PNA structures

The three dimensional structure of PNA:DNA and PNA:RNA duplex/triplexes[7]demonstrate that PNA have the flexibility to enable the adaption to the A-form and B-form helices preferred by RNA or DNA, but the results also clearly show that the P-form structure preferred by PNA is much wider and more slowly winding helix. These results imply that PNA is not a perfect structural 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 monomer[8] 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[9] resulting in both neutral and positively charged PNA oligomers (Figure 2).

The results suggest that the sterically allowed constrains of the PNA backbones improve its tendency to hybridize with DNA/RNA, but this also depends on the stereochemistry of the residue. Introduction of chemical bridges in *aeg*PNA to get cyclic structures may help in constraining the backbone and thus achieving parallel/antiparallel orientation selective binding by virtue of chirality in the backbone. 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.[10] 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 constraining the single stranded PNA in a conformation identical or close to that found in the 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 obvious strategy in the quest for oligomers with improved hybridization potency.

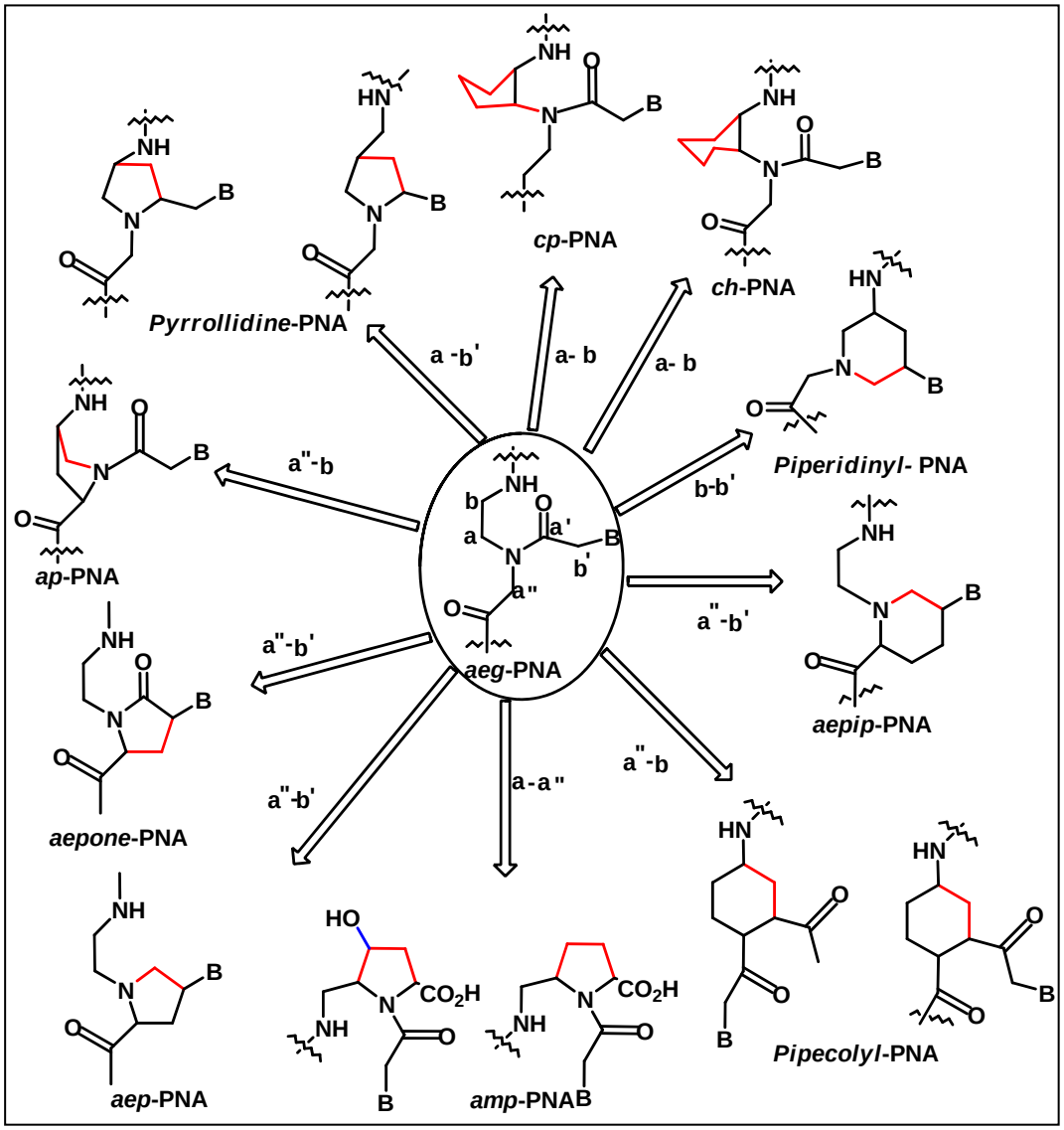


Figure 2: Conformationally constrained PNA analogues[9]

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 PNA backbone to control the orientation selectivity and specificity in binding by preorganization. Recently a five membered carbocyclic analogue of PNA (*cp*-PNA) with both (1*R*,2*S*) and (1*S*,2*R*) *cis* chemistry from our group[11] and (1*S*,2*S*) *trans* chemistry from Appella *et al* [12] was reported (Figure 3, *cp*-PNAs). DNA binding studies on these cyclopentyl PNAs showed stabilization of the derived PNA:DNA and PNA:RNA complexes.

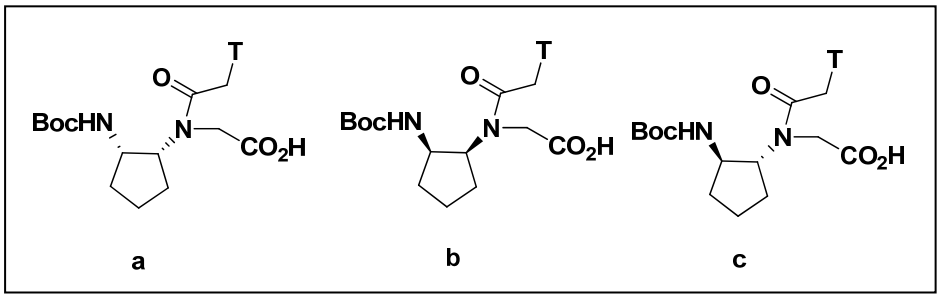


Figure 3: *cis* (**a** & **b**) and *trans*(**c**) Cyclopentyl PNA`s.

Yeheskiely et al [13] introduced *N*-(Pyrrolidinyl-2-methyl) glycine based PNA (*Pmg*PNA) derived from L-prolinol and D-prolinol (Figure 4). 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 selectivity higher than unmodified control *aeg*PNA, However, they showed poor duplex stability towards DNA/RNA than *aeg*PNA which was attributed to lack of positive charge in the PNA.

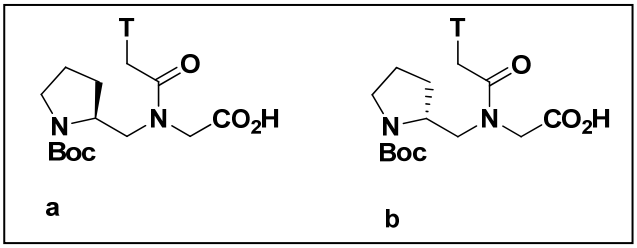


Figure 4: a. (*R*)- Pmg PNA b. (*S*)-Pmg PNA

The major limitations of the therapeutic applications of PNA are their poor solubility in aqueous media and also the selectivity of PNA towards DNA /RNA is of 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 and also do not carry any positive charge which is an important factor for solubility and better duplex stability. 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.[14] Considerable interest has emerged in making positively charged PNAs as they are expected to possess superior ability to strand invasion of complementary DNA sequences. The positive charge can arise from the

introduction of amino/imino functions in PNA which can accept a proton. Several pyrrolidine derivatives were introduced in the PNA to make the backbone cationic and rigid.[9] But it is important to see that the positive charge brought in the PNA should not compromise the sequence specificity of binding of PNA (by electrostatic interactions) which may have detrimental effect on fidelity of the PNA:DNA complexes. 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 or may have 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.[15]

2.2.1 Rationale

In the light of all these literature findings, the synthesis of 3,4-disubstituted pyrrolidine PNA monomers **13** and **26** (Figure 5) was initiated. These involve monomers derived from (3*R*,4*R*)-3,4-diaminopyrrolidine (*dap*PNA **13)** by introducing a nitrogen atom in the cyclopentyl ring of *cp*-PNA and keeping all the other moieties unchanged in order to get an extra amino nitrogen atom in the PNA backbone which will make it cationic and more hydrophilic than the corresponding cyclopentyl PNA. The synthesis of its one carbon extended homologue derived from (3*R*,4*S*)-3-(aminomethyl)-pyrrolidin-4-amine (*amp*PNA **26)** was also achieved. The extended backbone by one carbon at C-3 may create a good balance between the increased rigidity due to pyrrolidine ring and flexibility due to one extra methyl group in the backbone of *amp*PNA.

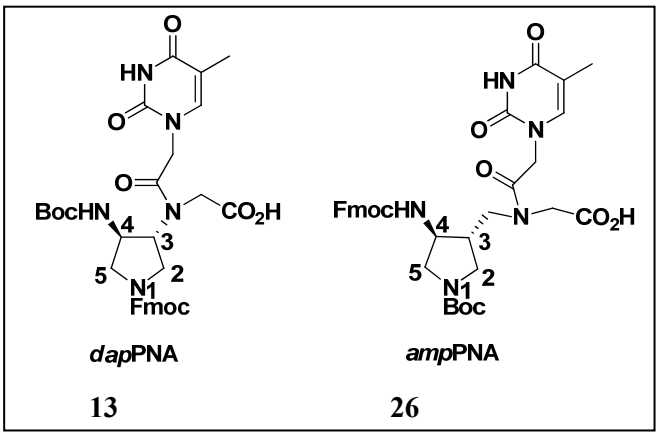


Figure 5: Chemical structure of modified PNA monomers

Most of the PNA modifications in the literature, utilising a pyrrolidine ring in backbone were either proline derived or have substitution at α-carbon to ring nitrogen. The *dap* and *amp*PNA constitute a new class of PNA analogues having β-substitution to ring nitrogen. It will also have an extra advantage that the two monomers **13** and **26** can be incorporated into standard PNA oligomers by using either Boc or Fmoc chemistry for solid phase synthesis leading to four different PNA backbones (Figure 6), depending upon the direction of synthesis. All these carry positive charges on their primary or secondary amino groups.

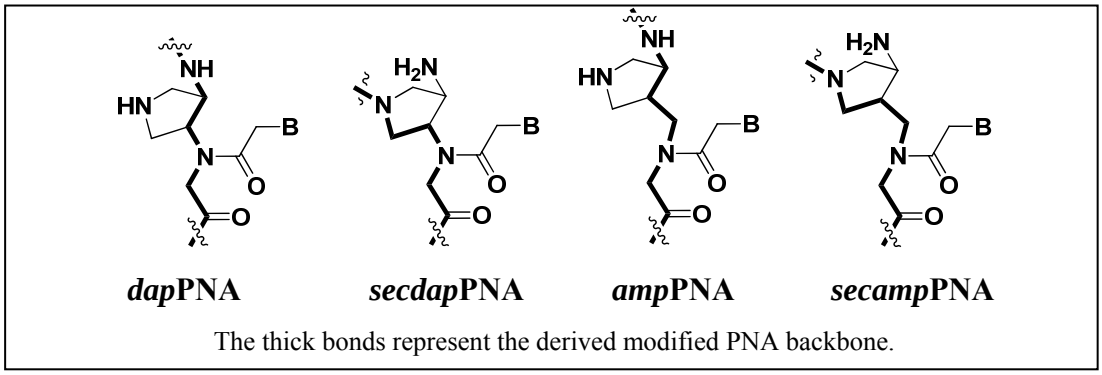


Figure 6: Four different PNA backbones derived from *dap* and *amp* PNA Monomers.

The secondary backbones derived from both the monomers will be like as in *PmgPNA* monomers in terms of oligomerization through ring nitrogen and will have an extra chiral centre in the form of amino group at 4-position.

2.3 Objectives

The specific objectives of this chapter are

i) Synthesis of *trans*-(3*R*,4*R*)-3,4-diamino pyrrolidinyl PNA (*dap*PNA) monomer with Boc and Fmoc as protecting groups.

ii) Synthesis of *trans*-(3*R*,4*S*)-3-(aminomethyl)pyrrolidin-4-amine PNA (*amp*PNA) monomer with Boc and Fmoc as protecting groups.

iii) Syntheis of Boc and Fmoc protected aminoethylglycyl (*aeg*) PNA monomers.

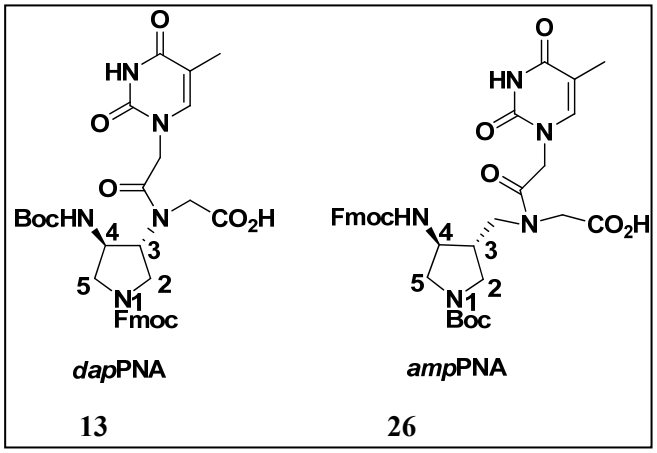
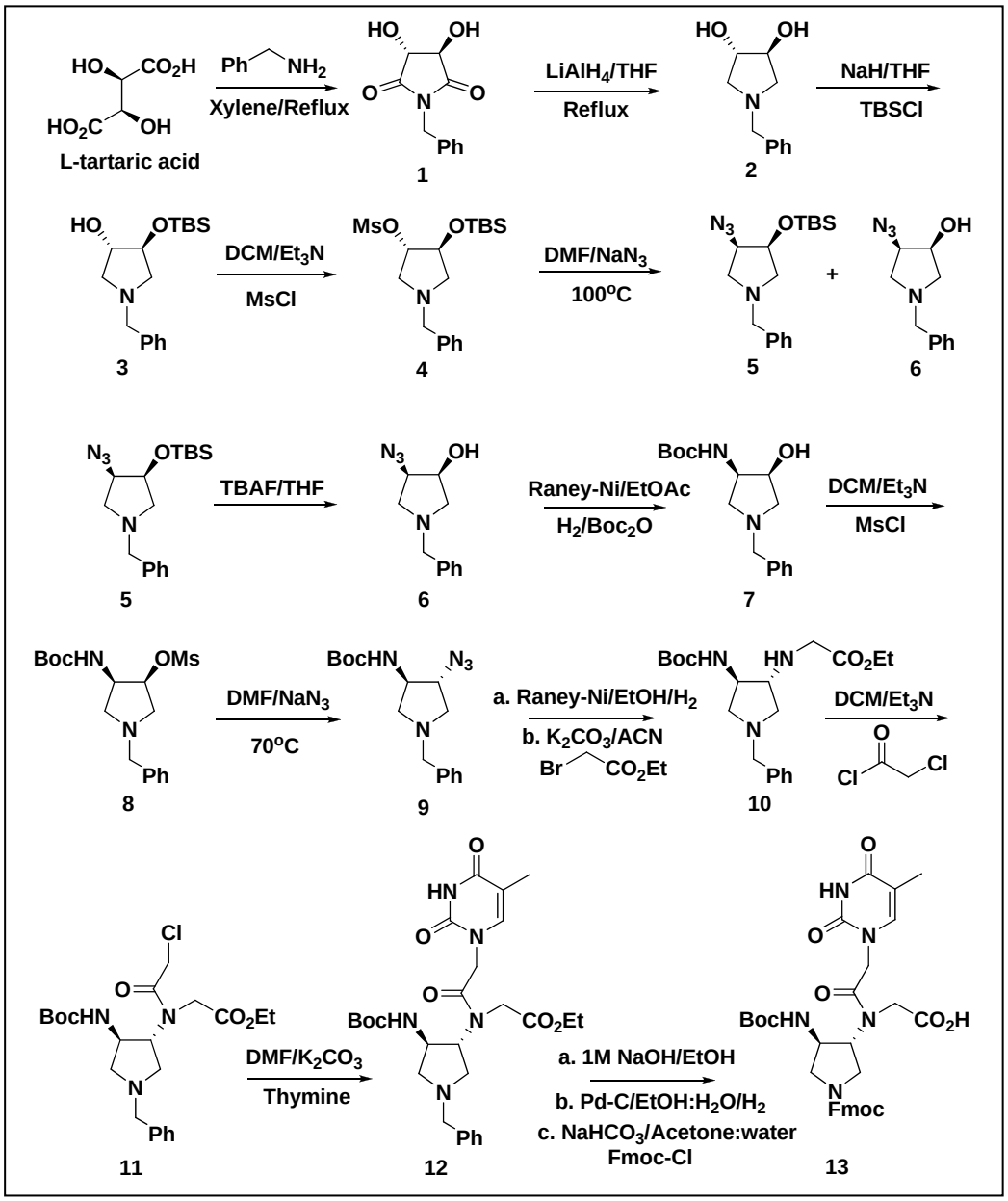


Figure 7: Chemical structure of modified PNA monomers

2.3.1 Synthesis of (3*R*,4*R*)-diaminopyrrolidinyl (*dap*) PNA monomer 13

The synthesis was started with naturally occurring L-tartaric acid which is commercially available and inexpensive. With its two chiral centers, L-tartaric acid has been a material of choice to access multifunctionalized pyrrolidine rings through their reactions with primary amines.[16] Thus, the synthesis of *trans*-(3*R*,4*R*)-*dap*PNA monomer **13** was achieved in good yield starting from the naturally occurring L-tartaric acid (Scheme 1). The first step involved is the condensation of L-tartaric acid with benzyl amine in boiling xylene, using Dean–Stark apparatus. The reaction progress was monitored by the amount of water collected and was mostly complete when it reached the theoretical amount. A good yield of imide **1** was obtained which was purified by recrystallization from ethanol. This was reduced using $LiAlH_4$ to get the diol **2**. The reported literature methods such as $NaBH_4/I_2$ in THF/diethylether and lithium aluminium hydride in diethyl ether[16] were used to reduce the imide **1** but the diol **2** could not be obtained in good yield. Under the reaction conditions the two carbonyl groups were incompletely reduced to give a mixture of various alcohols, which being polar were difficult to separate leading to low yields. Finally the imide **1** was reduced completely to diol **2** by a slight modification of the procedure during the work up in the method using $LiAlH_4$ in THF.[17] Instead of quenching the reaction with NaOH solution in work up, the reaction mixture was quenched with saturated solution of Na_2SO_4, followed by stirring the filtrate with diethyl ether to get pure diol **2**. The NMR data obtained for diol **2** was in consistent with literature data[17].

Scheme 1: Synthesis of (3*R*,4*R*) *dap*PNA Monomer.



The mono O-protection of diol **2** was attempted using benzoyl choride or *t*-butyldimethylsilyl chloride (TBSCl), but in both cases a mixture of mono O- and di O-protected compounds were obtained in almost equal ratio. The reagent used in 0.7-1.0 equivalents to give mono O-protected derivative as a major product, but failed to give good yield of mono O-protected alcohol. Finally reaction with TBSCl as reagent in THF and sodium hydride as the base gave mono O-protected alcohol **3** in very good yield without any di O-protected compound. The alcohol **3** was converted

to its O-mesyl derivative **4** using mesyl chloride and the conversion was confirmed by the appearance of ^{1}H NMR peak at δ 3.04 ppm for $–CH_3$ group of mesyl in the product. This was converted to corresponding azide **5** using sodium azide in DMF at 100°C. The azidation reaction was inefficient at normal temperature of 70° C, but at 100° C the reaction also gave O-desilylated azide **6**. Since the next step involved is removal of TBS group, the formation of desilylated compound **6** was not a problem. The presence of azide group in compounds **5** and **6** was confirmed by their corresponding IR spectra showing azide peak at 2104 cm^{-1} and 2105 cm^{-1} respectively. The compound **5** was O-desilylated to azido alcohol **6** using TBAF in THF and the azide group was hydrogenated using Raney-Ni to the amine followed by *insitu* N-Boc protection using Boc-anhydride to obtain the compound **7**. For hydrogenation Raney Ni instead of Pd-C was found to be a better catalyst as the later also deprotected N-benzyl group.

Compound **7** was converted to O-mesyl derivative **8** which was transformed to azide **9** using sodium azide in DMF at 70° C. The stereochemistry of compounds **8** and **9** were confirmed by their X-ray crystal structures as shown (Figure 8) which indicated complete inversion of configuration due to SN^2 attack by azide group. The product azide **9** was hydrogenated using Raney-Ni as catalyst to obtain the crude amine which without further purification was N-alkylated using ethyl bromoacetate in the presence of K_2CO_3. This leads to compound **10** which upon acylation with chloroacetyl chloride gave the chloro compound **11**. This was employed for N1-alkylation of thymine using K_2CO_3 in DMF to yield the ester **12**. This was first hydrolysed to the acid followed by N-deprotection of benzyl group using 10% Pd-C to obtain crude amine that was finally protected as Fmoc group using Fmoc-Cl or Fmoc-succinimide to get the desired *dap*PNA monomer **13.** This was used to synthesize the desired PNA oligomers on solid phase. All new compounds were characterized by ^{1}H, ^{13}C NMR and mass spectral data. (See Appendix)

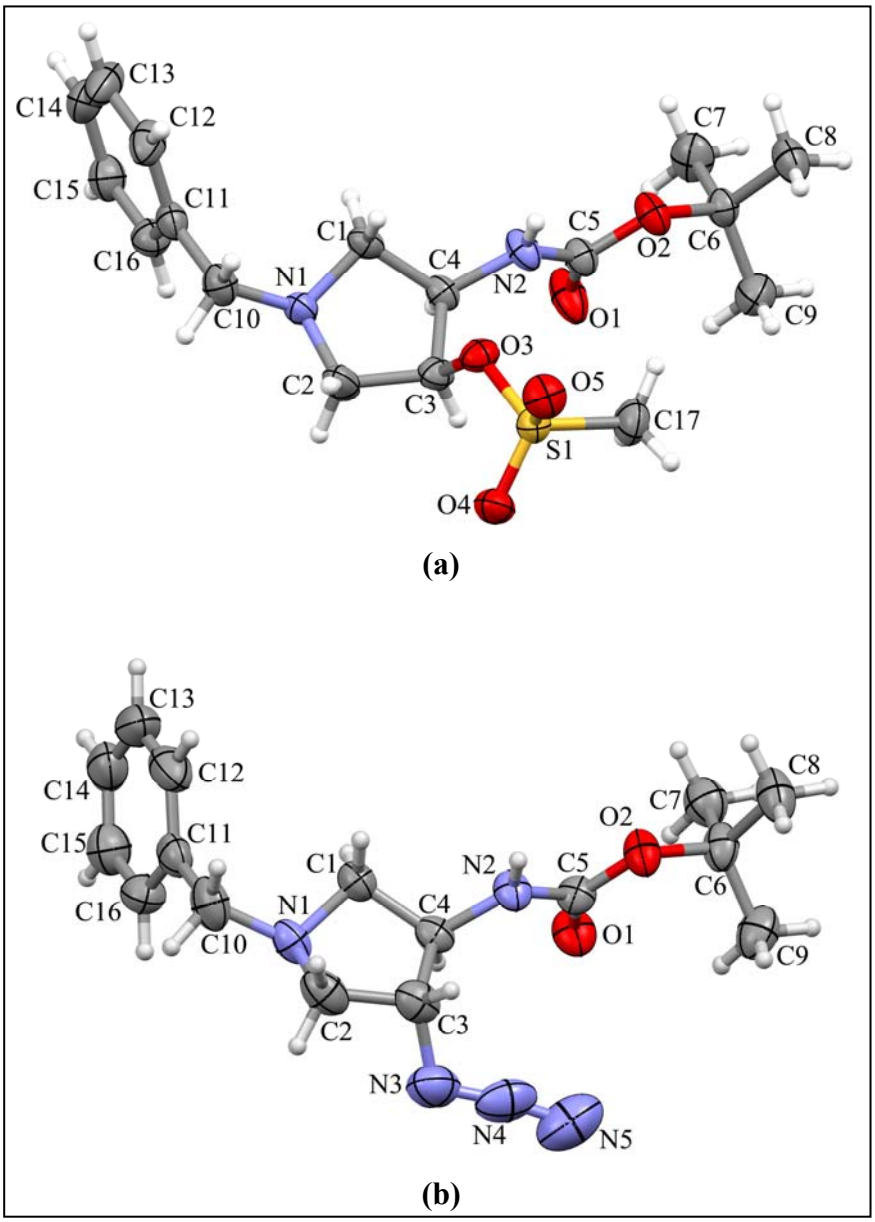
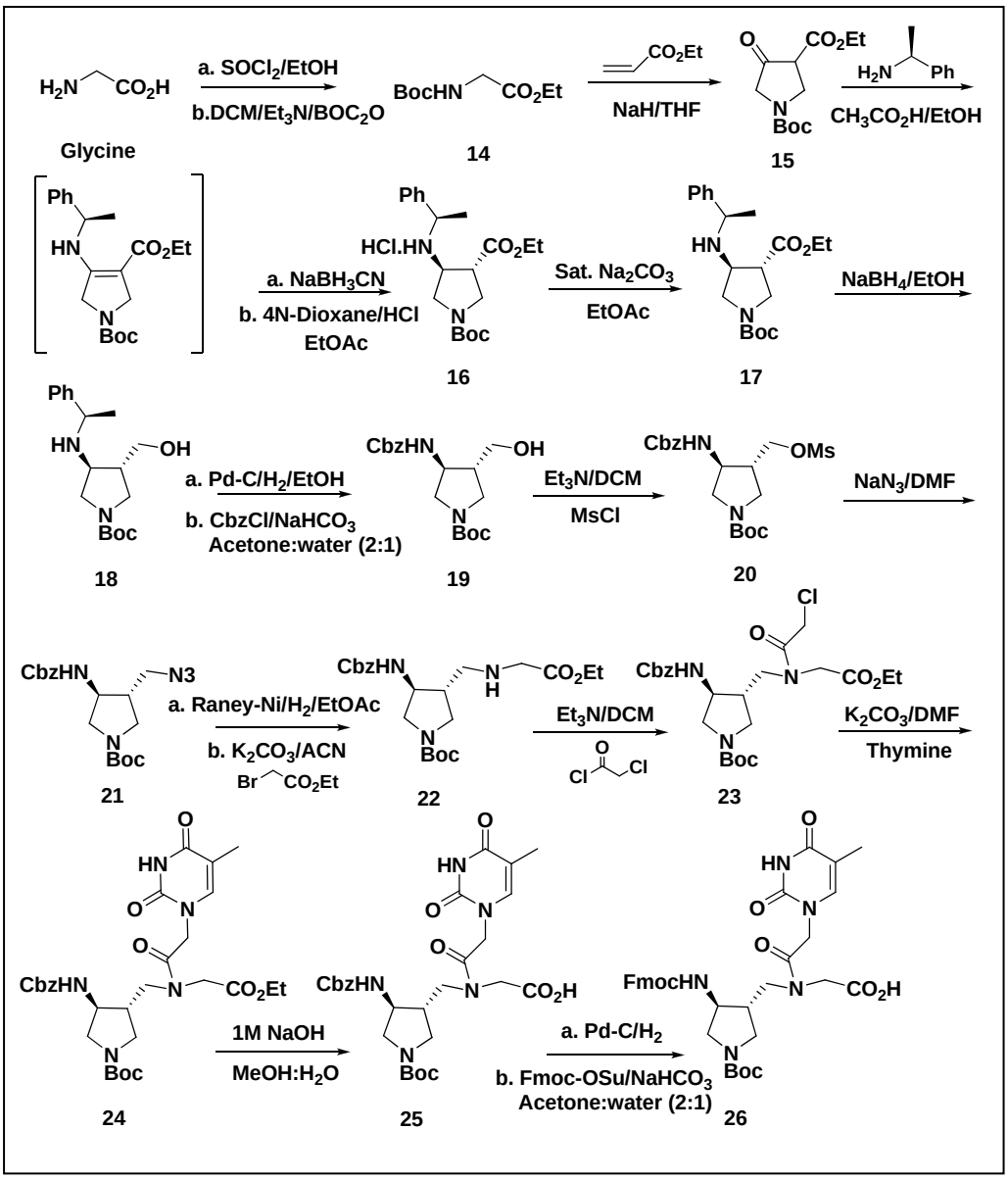


Figure 8: Crystal structure for mesyl **8** (a) and azide **9 (b)**

2.3.2 Synthesis of *trans*-(*3R,4S*)-3-(aminomethyl)pyrrolidin-4-amine (*amp*) PNA monomer 26

The synthesis of *amp*PNA monomer **26** (Scheme 2) was carried out starting from glycine converted to its ethyl ester by treating with thionyl chloride in ethanol. The reaction initially was stirred at room temperature for around half an hour to complete the formation of acid chloride *insitu* and then refluxed during which ethoxide replace the chloro group to get glycine ethylester as its hydrochloride salt. This was N-Boc protected using Boc-anhydride to get N-protected glycine ester **14.**

Scheme 2: Synthesis of (3*R*, 4*S*) backbone extended *amp*PNA Monomer.



A tandem aza-Michael addition followed by a Dieckmann condensation was carried out, using NaH in THF, on **14** to get pyrrolidine β-ketoester **15** in good yield. This reaction is the basis of a new methodology to construct various kind of heterocyclic rings (see Chapter-4). The β-ketoester **15** on reductive amination with chiral auxiliary *R*-(+)-phenylbenzylamine in ethanol followed by *insitu* reduction using $NaCNBH_3$ yielded the corresponding diastereomeric mixture of four possible products, from which the required diastereomer **16** precipitated out selectively in 4 N dioxane/HCl as the hydrochloride salt. The overall yield of the reaction was low as

compound **16** selectively precipitated out as one of the diastereomer among the four possible ones formed in the reaction mixture. Upon recrystallization of **16** from acetonitrile, the chiral purity was confirmed to be > 99% by comparing its specific rotation with that reported in literature.[18] The hydrochloride salt was neutralised to get the N-protected amine **17**, which had optical purity > 99% as determined by GC analysis. The ester group of **17** was reduced to alcohol **18** using sodium borohydride in ethanol and the N-protected chiral auxiliary was removed by hydrogenation using 10% Pd-C to get the corresponding free amine. This was N-protected with benzyloxy carbonyl group (*Cbz*) to the alcohol **19** that was reacted with mesyl chloride in DCM/triethylamine to give O-mesyl derivative **20**. The conversion was confirmed by the appearance of ^{1}H NMR peak at δ 3.04 ppm for $-CH_3$ group of mesyl in compound **20** which was converted to azide **21** using sodium azide in DMF. The presence of azide group was confirmed by IR spectra showing peak for azide at 2101 cm^{-1}.

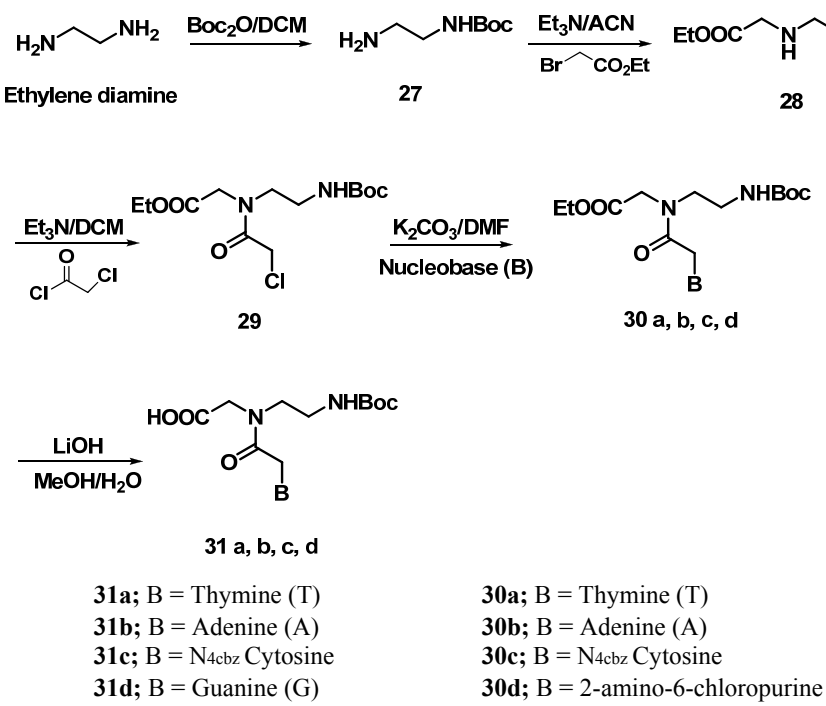
The azide **21** was reduced to amine using Raney-Ni, since Pd/C also cleaves *N-cbz* group. The crude amine was N-alkylated using ethylbromoacetate in acetonitrile containing K_2CO_3 to get **22** that was N-acylated using chloroacetyl chloride to the chloro mesyl compound **23**. The chloro group in compound **23** was used to N-alkylate thymine using K_2CO_3 in DMF to obtain the thymine ester **24**. The appearance of peaks at δ 7.28 ppm and 1.68 ppm in ^{1}H NMR shows the presence of thymine in desired product **24**. This was hydrolysed with 1N sodium hydroxide in methanol to get the PNA monomer **25** required for the solid phase synthesis of PNA oligomers using Boc-chemistry in soild phase. The N-cbz group in **25** was removed by hydrogenation using 10 % Pd-C as catalyst and the free 4-amino group was protected with Fmoc group using Fmoc succinimide to get alternative PNA monomer **26** required for the solid phase synthesis of isomeric PNA oligomers using Fmoc-chemistry in soild phase. All new compounds were characterized by ^{1}H, ^{13}C NMR and mass spectral data.

2.3.3 Synthesis of Boc-protected *aeg*PNA monomers

The synthesis of unmodified *aeg*PNA monomers was carried out following the literature procedures.[19] The synthesis was started from the easily available 1,2-

diaminoethane (Scheme 3) which was treated with *t*-butyl pyrocarbonate to give the mono N-protected amine **27** 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 **27** was subjected to *N*-alkylation using ethylbromoacetate and triethylamine in dry acetonitrile. The aminoethylglycine **28** was further treated with chloroacetyl chloride to yield the corresponding chloro derivative **29,** which is the common intermediate for the preparation of all the PNA monomers in good yield.

Scheme 3: Synthesis of Boc-protected *aeg*PNA monomers



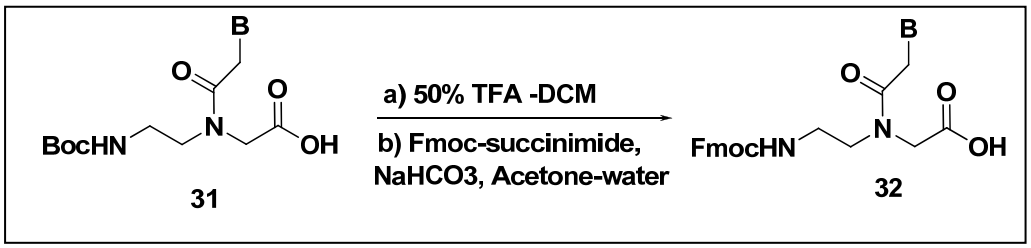
31a; B = Thymine (T)
31b; B = Adenine (A)
31c; B = N_{4cbz} Cytosine
31d; B = Guanine (G)

30a; B = Thymine (T)
30b; B = Adenine (A)
30c; B = N_{4cbz} Cytosine
30d; B = 2-amino-6-chloropurine

Thymine was reacted with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **29** in presence of K_2CO_3 to obtain the *N*-(Boc-aminoethylglycyl)-thymine ethyl ester **30a** in high yield. In the case of cytosine, the *N*4-amino group was protected as its benzyloxycarbonyl derivative, and used for alkylation in presence of K_2CO_3 to provide the *N*1 substituted product **30b**. Although adenine is known to undergo both *N*7- and *N*9- substitution, *N*7-alkylation was not observed when K_2CO_3 was used as a base to get the product **30c**. The *N*-alkylation of 2-amino-6-

chloropurine with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate was facile with K_2CO_3 as the base and yielded the corresponding *N*-(Boc-aminoethylglycyl)-(2-amino-6-chloropurine)-ethyl ester **30d** in excellent yield. All compounds exhibited ^{1}H and ^{13}C NMR spectra consistent with the reported data. The ethyl esters were hydrolyzed in the presence of LiOH to give the corresponding acids **(31a-d)**, 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.

2.3.4 Synthesis of Fmoc-protected *aeg* PNA monomer



Scheme 4: Synthesis of Fmoc *aeg* PNA monomer

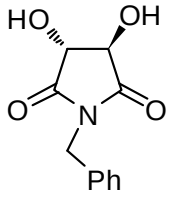
The Boc protected *aeg*PNA monomers **31(a,b,c,d)** were synthesized following literature procedure as discussed earlier. The Boc group was deprotected using 50% TFA-DCM and amine formed was protected as Fmoc to give Fmoc protected *aeg*PNA monomer building block **32 (a,b,c,d)** (Scheme 4), necessary for synthesis of PNA oligomers through solid phase synthesis using Fmoc chemistry.

2.4 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). 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_{254} 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}H NMR (200 MHz), ^{13}C NMR (50 MHz) spectra were recorded on Bruker ACF200 spectrometer fitted with an Aspect 3000 computer. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^{1}H and deuterated NMR solvent chloroform-*d* itself for ^{13}C. The optical rotation values were measured on Bellingham-Stanley Ltd, ADP220 polarimeter. Mass spectra were obtained by LCMS techniques.

(3*R*,4*R*)-1-benzyl-3,4-dihydroxypyrrolidine-2,5-dione[16] (1)

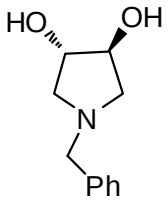


To a stirred suspension of L-tartaric acid (45.0 g, 0.3 mol) in xylene (200 ml) at room temperature, benzyl amine (32.7 ml, 0.3 mol) was added and the resulting suspension was refluxed using a dean stark apparatus till the appropriate amount of water (10.8 ml, 0.6 mol) was collected during 7-8 hrs. After 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%).

$[\alpha]_D^{25}$ = + 138 (*c* 1.0, MeOH); **M.P.** = 196-198 °C; **^{1}H NMR** (200 MHz, DMSO-d^6): δ 7.26-7.11 (m, 5H, Ph), 6.25-6.17 (m, 2H, 2 x OH), 4.53-4.37 (m, 2H, benzyl CH_2), 4.32-4.24 (m, 2H, 2 x CH); **^{13}C NMR** (50 MHz, DMSO-d^6): δ 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]

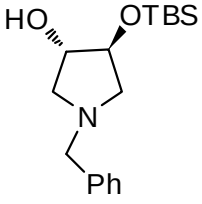
(3*S*,4*S*)-1-benzylpyrrolidine-3,4-diol[17] (2)



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; **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.32-7.22 (m, 5H, Ph), 4.15-3.95 (m, 4H, 2 x OH (2H) and 2 x CH (2H)), 3.67-3.47 (m, $PhCH_2$-diastereotopic, ^2J = 12.6 Hz, 2H), 2.89 (dd, J = 5.8 Hz, 10.1 Hz, 2H, -CH_2), 2.41 (dd, J = 3.9 Hz, 10.3 Hz, 2H, -CH_2); **^{13}C NMR** (50 MHz, $CDCl_3$): δ 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-benzyl-4-(*tert*-butyldimethylsilyloxy)pyrrolidin-3-ol (3)

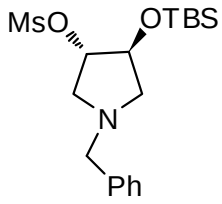


To an ice-cooled solution of diol **2** (3.3 g, 17.1 mmol) in THF (50 ml) under nitrogen, was added NaH (0.49 g, 20.5 mmol) and the reaction mixture was stirred in an ice bath for 15 min followed by the addition of tert-butyldimethylsilyl chloride (2.58 g, 17.1 mmol). The reaction mixture was stirred at room temperature for 12 hrs. After the completion of reaction as observed by TLC, solvent was evaporated to dryness and the residue was taken in water, extracted with ethyl acetate and dried over sodium sulphate. The solvent was evaporated under reduced

pressure which on column chromatographic purification afforded the compound **3** as light reddish oil. Yield: (4.2 g, 80%)

$[\alpha]_D^{25}$ = + 26 (*c* 1.0, MeOH); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.34-7.27 (m, 5H, Ph), 4.18-4.08 (m, 1H, -OH), 4.00-3.92 (m, 1H, -C<u>H</u>-OH), 3.74-3.56 (m, $PhCH_2$ diastereoscopic, ^2J = 2.3 Hz, 2H), 3.21 (dd, *J* = 6.2 Hz, 10.1 Hz, 1H, -C<u>H</u>-OTBS), 2.80-2.60 (m, 3H), 2.20 (dd, *J* = 4.6 Hz, 10.0 Hz, 1H), 0.87 (s, 9H, TBS), 0.05 (d, *J* = 3.9 Hz, 6H, TBS); **^{13}C NMR** (50 MHz, $CDCl_3$): δ 137.5, 128.9, 128.2, 127.2, 79.1, 78.4, 61.0, 60.4, 60.1, 25.7, 17.9, - 5.0 ; **MS (EI)** *m/z* 307.20, Found 308.19 [M + H].

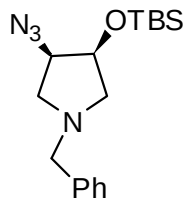
(3*S*,4*S*)-1-benzyl-3-*O*-mesyl-4-(*tert*-butyldimethylsilyloxy)pyrrolidine (4)



To a stirred solution of **3** (4.5 g, 14.6 mmol) and triethylamine (8.3 ml, 58.6 mmol) in dry DCM (40 ml) at 0° C under nitrogen, was added methanesulfonyl chloride (2.3 ml, 29.2 mmol) in DCM (10 ml) over a period of 15 min. The mixture was stirred for another 20 min. at room temperature, diluted with DCM (20ml), washed successively with water, brine and dried over sodium sulphate. The organic layer was concentrated and purified to get mesyl **4** as a colourless oil. Yield: (4.2 g, 75%).

$[\alpha]_D^{25}$ = + 38 (*c* 1.0, MeOH); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.37-7.25 (m, 5H, Ph), 4.89-4.80 (m, 1H, -C<u>H</u>-OMs), 4.41 (dt, *J* = 6.2 Hz, 1H, -C<u>H</u>-OTBS), 3.75-3.51 (m, $PhCH_2$-diastereoscopic, ^2J = 14.0 Hz, 2H) 3.10 (dd, *J* = 6.6 Hz, 9.6 Hz, 1H), 3.00 (s, 3H, mesyl CH_3), 2.95-2.80 (m, 2H), 2.32 (dd, *J* = 5.9 Hz, 9.7 Hz, 1H), 0.88 (s, 9H, TBS), 0.07 (d, *J* = 5.7 Hz, 6H, TBS); **^{13}C NMR** (50 MHz, $CDCl_3$): δ 137.7, 128.6, 128.2, 127.1, 86.7, 76.6, 60.1, 59.8, 57.6, 38.1, 25.6, 17.8, -5.0. ; **MS (EI)** *m/z* 385.17, Found 385.73 [M + H], 407.69 [M + Na]

(3*R*,4*S*)-3-azido-1-benzyl-4-(*tert*-butyldimethylsilyloxy)pyrrolidine (5)

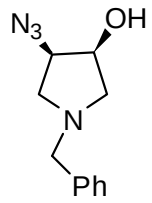


A stirred mixture of mesyl **4** (17.0 g, 44.2 mmol) and sodium azide (23.0 g, 353.8 mmol) in DMF (200ml) under nitrogen was heated at 100-110° C for 12hrs. After cooling the solvent was removed under reduced pressure. The residue was taken in ethyl acetate and washed

with water, brine, dried over sodium sulphate and purified by column chromatography to yield **5** as light yellow oil. Yield: (7.0 g, 48%).

Compound-5: $[\alpha]_D^{25}$ = + 50 (*c* 1.0, MeOH); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.30-7.15 (m, 5H, Ph), 4.33 (q, *J* = 6.2 Hz, 1H, -C<u>H</u>-OTBS), 3.68-3.45 (m, 3H, C<u>H</u>-N_3 and -CH_2-Ph), 3.05-2.88 (m, 2H, -CH_2), 2.54-2.36 (m, 2H, -CH_2), 0.86 (s, 9H, TBS), 0.35 (d, *J* = 7.2 Hz, 6H, TBS); **^{13}C NMR** (50 MHz, $CDCl_3$): δ 138.3, 128.4, 128.2, 127.0, 72.8, 61.2, 60.3, 60.1, 56.3, 25.6, 18.0, -5.2. ; **MS (EI)** *m/z* 332.20, Found 333.85 [M + H].

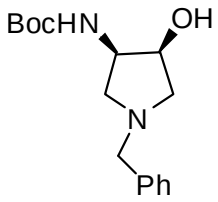
(3*S*,4*R*)-4-azido-1-benzylpyrrolidin-3-ol (6)



A 1M solution of tetrabutyl ammonium fluoride in THF (23 ml) was added to a stirred solution of **5** (5.0 g, 15.1 mmol) in THF (50 ml) under nitrogen at 0° C. The reaction mixture was stirred at room temperature for 2 hrs, Solvent was removed, the residue was extracted with ethyl acetate, washed with water, dried over sodium sulphate and purified by column chromatography to get compound **6** as a colourless oil. Yield: (2.8 g, 86%)

$[\alpha]_D^{25}$ = + 66 (*c* 1.0, MeOH); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.38-7.20 (m, 5H, Ph), 4.38-4.28 (m, 1H, -C<u>H</u>-OH), 3.94 (q, *J* = 6.1 Hz, 1H, -C<u>H</u>-N_3), 3.65 (s, 2H, -CH_2-Ph), 2.96-2.84 (m, 2H,-CH_2), 2.75-2.55 (m, 2H, -CH_2), 2.39 (br s, 1H, -OH); **^{13}C NMR** (50 MHz, $CDCl_3$): δ 137.7, 128.7, 128.2, 127.2, 71.2, 62.0, 60.1, 59.9, 56.2.; **MS (EI)** *m/z* 218.12, Found 219.11 [M + H].

(3*R*,4*S*)-1-benzyl-3-Boc-amino-4-hydroxypyrrolidine (7)

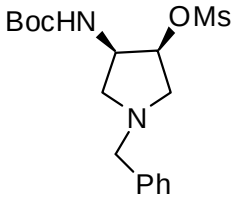


To a solution of azido-alcohol **6** (2.3 g, 10.6 mmol) in ethyl acetate (15 ml) taken in a hydrogenation flask was added boc-anhydride (2.8 g, 12.7 mmol) and Raney-Ni (1.0 g). The reaction mixture was hydrogenated in a parr apparatus for about 3 hrs at room temperature and at a hydrogen pressure of 40-45 psi during which the TLC indicates the disappearance of starting azide. 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 **7** as a white solid. Yield: (2.3 g, 74%).

$[\alpha]_D^{25}$ = + 9 (*c* 1.0, MeOH); **M.P.** = 79-81 °C; **^{1}H NMR** (200 MHz, $CDCl_3$): 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-NH), 3.59 (s, 2H, -CH_2-Ph), 3.05 (br s, 1H, -OH), 2.86-2.67 (m, 2H, -CH_2), 2.65-2.45 (m, 2H, -CH_2), 1.42 (s, 9H, *t*-Boc); **^{13}C NMR** (50 MHz, $CDCl_3$): 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* 292.18, Found 293.75 [M + H], 316.19 [M + Na].

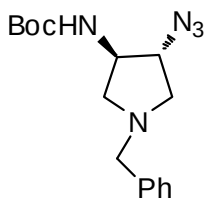
(3*S*,4*R*)-1-benzyl-4-Boc-amino-3-*O*-mesylpyrrolidine (8)



To a stirred solution of **7** (2.3 g, 7.9 mmol) and triethylamine (4.4 ml, 31.6 mmol) in dry DCM (20 ml) at 0° C under nitrogen, was added methanesulfonyl chloride (1.2 ml, 15.8 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. The organic layer was concentrated and purified to get mesylate **8** as a white crystalline solid. Yield: (2.4 g, 83%).

$[\alpha]_D^{25}$ = + 37 (*c* 1.0, $CHCl_3$); **M.P.** = 130-132 °C; **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.40-7.20 (m, 5H, Ph), 5.10-4.87 (m, 2H, -NH and –CH-OMs), 4.50-4.27 (m, 1H, -CH-NH), 3.65 (s, 2H, -CH_2-Ph), 3.01 (s, 3H, mesyl CH_3), 2.97-2.80 (m, 2H, -CH_2), 2.49 (dd, *J* = 7.5 Hz, *J* = 9.2 Hz, 2H, -CH_2), 1.44 (s, 9H, *t*-Boc).; **^{13}C NMR** (50 MHz, $CDCl_3$): 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* 370.16, Found 372.09 [M + H], 394.13 [M + Na].

(3*S*,4*R*)-4-azido-1-benzyl-3-Boc-aminopyrrolidine (9)

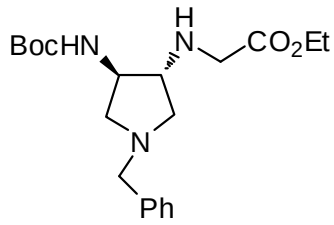


To a stirred solution of mesyl **8** (2.5 g, 6.7 mmol) in dry DMF (30ml) was added sodium azide (3.5 g, 53.8 mmol) at room temperature under nitrogen. The reaction mixture was heated at 60 - 70° C for 18 hrs. The solvent was removed under reduced pressure and the residue was taken in water and extracted 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 **9** as a white solid. Yield (1.85 g, 86%).

$[\alpha]_D^{25}$ = - 12 (*c* 1.0, MeOH); **M.P.** = 70-72 °C; **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.40-7.20 (m, 5H, Ph), 5.10-4.75 (m, 1H, -NH), 4.15-3.90 (m, 1H, -C<u>H</u>-NH), 3.85-3.70 (m, 1H, -C<u>H</u>-N_3), 3.60 (s, 2H, -CH_2-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_2), 1.45 (s, 9H, *t*-Boc).; **^{13}C NMR** (50 MHz, $CDCl_3$): δ 154.9, 137.7, 128.6, 128.2, 127.2, 79.7, 66.6, 59.4, 58.0, 57.8, 56.5, 28.2. ; **MS (EI)** *m/z* 317.19, Found 318.90 [M + H], 340.93 [M + Na].

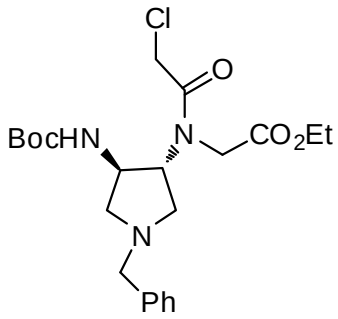
N-[(3*R*,4*R*)-1-benzyl-4-Boc-aminopyrrolidin-3-yl]-glycine ethyl ester (10)



To azide **9** (3.3 g, 10.4 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 K_2CO_3 (2.5 g, 18.1 mmol) and the mixture was stirred under nitrogen at 0° C for 15 min followed by the addition of ethyl bromoacetate (1.4 ml, 12.5 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 **10** as a light yellow oil. Yield: (2.8 g, 71%).

$[\alpha]_D^{25}$ = + 8.3 (*c* 1.2, $CHCl_3$); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 7.38-7.15 (m, 5H, Ph), 5.13-4.93 (m, 1H, -N<u>H</u>Boc), 4.17 (q, *J* = 7.1 Hz, 2H, -CO_2CH_2), 3.93-3.72 (m, 1H, -NH), 3.59 (s, 2H, -CH_2-Ph), 3.49 (s, 2H, -C<u>H</u>$_2$-NH), 3.22-3.00 (m, 2H, 2 x CH), 2.83-2.68 (m, 1H), 2.64-2.48 (m, 1H), 2.20-2.05 (m, 2H, -CH_2), 1.43 (s, 9H, *t*-Boc), 1.26 (t, *J* = 7.1 Hz, 3H, -CO_2CH_2C<u>H</u>$_3$). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 172.3, 155.2, 137.9, 128.7, 128.2, 127.1, 79.2, 65.2, 60.6, 59.9, 59.7, 58.7, 56.1, 48.9, 28.2, 14.1. ; **MS (EI)** *m/z* 377.23, Found 377.13 [M + H], 399.02 [M + Na].

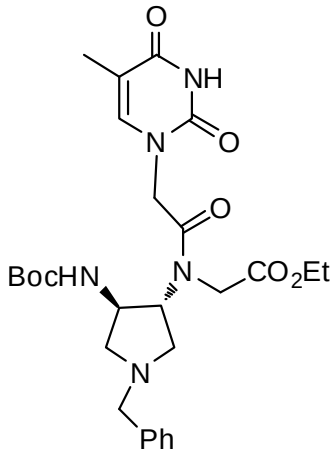
N-[(3*R*,4*R*)-1-benzyl-4-Boc-aminopyrrolidin-3-yl]-N-(chloroacetyl)-glycine ethyl ester (11)



To an ice-cooled stirred solution of **10** (2.55 g, 6.8 mmol) in DCM (30ml) was added triethylamine (3.8 ml, 27.3 mmol). The reaction mixture was stirred for 15 min followed by the addition of chloroacetyl chloride (1.1 ml, 13.8 mmol) and 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 **11** as a colourless liquid which was used without further purification. Yield: (2.6 g, 85 %).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.38-7.18 (m, 5H, Ph), 5.23 (br s, 1H, -NHBoc), 4.51-4.05 (m, 6H, $-CO_2\underline{CH_2}$, $-\underline{CH_2}-CO_2Et$, 2 x CH ring (2H)), 4.03-3.90 (m, 2H, $-CH_2Cl$), 3.71-3.38 (m, 2H, $-CH_2$-Ph), 2.91-2.57 (m, 2H, $-CH_2$), 2.52-2.17 (m, 2H, $-CH_2$), 1.43 (s, 9H, *t*-Boc), 1.24 (t, J = 7.1 Hz, 3H, $-C\underline{H_3}$). **MS (EI)** *m/z* 453.20, Found 454.41 [M + H], 476.39 [M + Na].

N-[(3*R*,4*R*)-1-benzyl-4-Boc-aminopyrrolidin-3-yl]-N-(thymin-1-acetyl)-glycine ethyl ester (12)

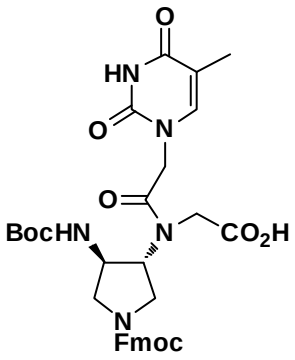


To compound **11** (0.7 g, 1.5 mmol) taken in 10 ml of DMF was added potassium carbonate (0.258 g, 1.9 mmol) followed by thymine (0.236 g, 1.9 mmol) at room temperature. The reaction mixture was stirred for 18 hrs at room temperature after which it was diluted with 100 ml of ethyl acetate and washed with water, brine and dried over sodium sulphate and purified by column chromatography to obtain compound **12** as an off white solid. Yield: (550 mg, 85 %).

$[\alpha]_D^{25}$ = - 12 (*c* 1.0, DMSO); **^{1}H NMR** (200 MHz, $CDCl_3$): 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 $–CH_2$, $-CH_2-CO_2Et$, $-CH_2$-Thymine), 3.70-3.40 (m, 2H, $-CH_2$-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_3), 1.41 (d, *J* = 9.5 Hz, 9H, *t*-Boc), 1.30-1.15 (m, 3H, -$CO_2CH_2C\underline{H}_3$). **^{13}C NMR** (50 MHz, $CDCl_3$): 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* 543.27, Found 544.59 [M + H], 566.57 [M + Na].

N-[(3*R*,4*R*)-1-Fmoc-4-Boc-aminopyrrolidin-3-yl]-N-(thymin-1-acetyl)-glycine (13)

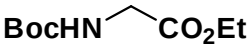


To an ice cooled solution of compound 12 (150mg, 0.28 mmol) in 5ml of ethanol was added 1.5 ml of 1N sodium hydroxide solution in water and the reaction mixture was stirred in an ice bath for 30 min. Water (10ml) was added to the reaction mixture and solvent was removed on rotary evaporator to half of its volume. The remaining mixture was neutralised with Dowex H^+ resin to pH 6 and concentrated to get 130 mg of corresponding acid as white solid. This was taken in 10 ml of ethanol: water (8:2), added 10% Pd-C (65mg, 50% by wt.) and was hydrogenated at 60 psi for 18 hrs. The reaction mixture thus obtained was filtered on celite bed and was concentrated to get 100 mg of crude amine. The crude amine was dissolved in 10 ml of acetone: water (2:1), cooled in an ice bath and sodium bicarbonate (24 mg, 0.28 mmol) was added followed by the addition of Fmoc-chloride (94mg, 0.36 mmol). The reaction mixture was stirred at room temperature for 12 hrs after which the acetone was removed and the water layer was neutralised with saturated solution of potassium hydrogen sulphate and extracted with ethyl acetate three times. The combined ethyl acetate extracts were washed with water, brine, dried over sodium sulphate and purified by column chromatography to obtain monomer 13 as white solid. Yield: (100mg, 57 %).

$[\alpha]_D^{25}$ = **-** 9 (*c* 1.0, DMSO); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 10.26 (br s, 1H, -COOH), 7.80-7.15 (m, 8H, fmoc), 7.10-6.85 (m, 1H, Thymine –NH), 4.60-3.98 (m, 6H, -CH_2 & -CH fmoc, -CH ring, -$\underline{CH_2}$-CO_2H), 3.95-3.55 (m, 3H, -$\underline{CH_2}$-Thymine, -CH ring), 3.45-3.10 (m, 2H, 2 x CH), 2.40 (t, J = 8.6 Hz, 1H, -CH), 2.10-1.90 (m, 1H, -CH), 1.71 (s, 3H, Thymine -CH_3), 1.39 (s, 9H, *t*-Boc). **^{13}C NMR** (50 MHz,

$CDCl_3$): δ 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* 647.26, Found 671.92 [M + Na].

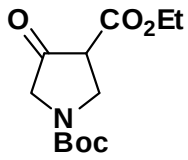
N-Boc glycine ethyl ester (14)



To an ice cooled solution of Glycine (20g, 0.27mol) in 200 ml of absolute ethanol was added thionyl chloride (24ml, 0.32mol) in 50 ml of ethanol dropwise. The reaction mixture was stirred in ice bath for around 30 min. after which it was refluxed for 12 hrs. The solvent was removed to obtain 38 g of crude HCl salt of glycine ethyl ester as white solid. The white solid obtained was taken in 300 ml of dry dichloromethane under nitrogen and to it was added triethylamine (123ml, 0.89mol) portion wise at 0° C followed by boc anhydride (61g, 0.28 mol) dropwise in 100 ml of DCM. The reaction mixture was stirred at room temperature for 48 hrs and was diluted with 200 ml of DCM and washed with water, brine, dried over sodium sulphate and purified to obtain compound **14** as light yellow oil. Yield: (45g, 85%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 5.65-5.50 (m, 1H, -NH), 4.20 (q, J = 7.1 Hz, 2H, -CH_2), 3.88 (d, J = 5.8 Hz, 2H, -CH_2), 1.45 (s, 9H, *t*-Boc), 1.28 (t, J = 7.1 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 170.6, 155.9, 80.0, 61.4, 42.6, 28.4, 14.3. ; **MS (EI)** *m/z* 203.12, Found 204.13 [M + H], 226.12 [M + Na].

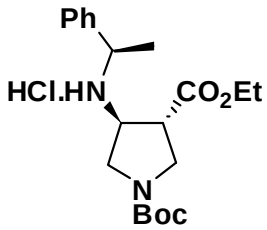
Ethyl N-Boc-4-oxopyrrolidine-3-carboxylate (15)



To an ice cooled solution of compound **14** (20g, 0.10 mmol) in 200 ml of THF was added 60% oil coated sodium hydride (4.7g, 0.11 mmol) portion wise. The reaction mixture was stirred in ice bath for another 20 min followed by addition of Ethylacrylate (13ml, 0.12 mmol) in 50 ml of THF. The reaction mixture was stirred at room temperature for 12 hrs. The solvent was removed under vacuum, the residue was dissolved in 150 ml of water and was neutralised to PH 6 using saturated solution of $KHSO_4$. It was extracted with 200 ml of ethyl acetate thrice and combined organic extracts were washed with water, brine, dried over sodium sulphate and purified to yield β-ketoester **15** as light yellow oil. Yield: (24g, 95 %).

^{1}H NMR (200 MHz, $CDCl_3$): δ 4.35-4.10 (m, 4H, ester -CH_2, -CH_2), 3.90-3.75 (m, 2H, -CH_2), 3.65-3.45 (m, 1H, -CH), 1.40 (s, 9H, *t*-Boc), 1.21 (t, *J* = 7.1 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 168.4, 168.1, 167.1, 166.9, 154.3, 154.2, 81.1, 77.6, 62.4, 60.9, 51.5, 51.2, 48.9, 48.5, 28.6, 28.5, 14.5, 14.4. ; **MS (EI)** *m/z* 257.1, Found 258.0 [M + H], 279.0 [M + Na]

(3*S*,4*R*)-1-N-Boc-3-((R)-1-phenylethylamino)-4-ethoxycarbonylpyrrolidine Hydrochloride[18] (16)

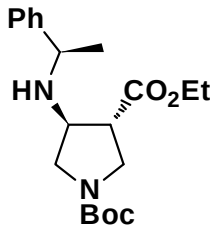


To a stirred solution of β-ketoester **15** (16.0 g, 62.3 mmol) in absolute ethanol (250 mL) under N2 were added (R)-(+)-α-methylbenzylamine (16.0 mL, 124.5 mmol) and glacial acetic acid (7.1 mL, 124.5 mmol) to obtain a cloudy solution. The reaction mixture was stirred at room temperature until the formation of the enamine was complete (6 h, monitored by TLC). Sodium cyanoborohydride (16.5 g, 249.2 mmol) was added to the reaction mixture at room temperature, and the resulting solution was heated to 75 °C and stirred for 14 h under N2. The ethanol was removed via rotary evaporation. Water (250 mL) was added. The mixture was extracted three times with diethyl ether. The combined organic extracts were washed with brine, dried over sodium sulphate and concentrated to give a colorless oil. The oil was filtered through a silica gel column to obtain a colorless oil as a mixture of diasteromers. The oil was dissolved in ethyl acetate (250 mL), and 4 N HCl in dioxane (15.6 mL) was added dropwise at room temperature. The resulting solution was cooled to 0 °C and allowed to stand for 3 h at 0 °C. A precipitate formed during this time. The solid was filtered and washed two times with 100 mL portions of ethyl acetate to provide the desired compound **16** as white solid which was recrystallized with acetonitrile. Yield: (6.1 g, 25%).

$[\alpha]_D^{25}$ = +4.6 (*c* 1.0, MeOH); M. P. = 187- 188°C. **^{1}H NMR** (200 MHz, DMSO-d^6): δ 10.15 (br s, 2H, -NH_2), 7.68-7.58 (m, 2H, Ph), 7.48-7.35 (m, 3H, -Ph), 4.58-4.44 (m, 1H, C<u>H</u>-Ph), 4.02 (q, J = 7.0 Hz, 2H, ester –CH_2), 3.85-3.65 (m, 3H, ring –CH_2 & -CH), 3.55-3.25 (m, 3H, ring –CH_2 & -CH), 1.62 (d, J = 6.5 Hz, 3H, -CH_3) 1.34 (s, 9H, *t*-boc), 1.10 (t, *J* = 7.0 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, DMSO-d^6): δ

170.9, 153.4, 136.9, 129.6, 129.4, 128.5, 79.8, 61.8, 57.0, 56.1, 48.4, 47.1, 45.4, 28.4, 20.4, 14.2. ; **MS (EI)** *m/z* 362.22 (without HCl), Found 363.41 [M + H].

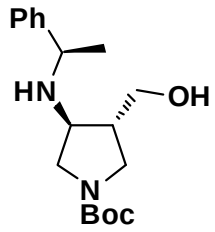
(3*S*,4*R*)-N-Boc-3-((R)-1-phenylethylamino)-4-ethoxycarbonylpyrrolidine[18] (17)



Compound **16** (6.0 g, 15.1 mmol) was mixed with excess saturated Na_2CO_3 solution, extracted into ethyl acetate, dried over sodium sulphate and concentrated in vacuum to get compound **17** as colourless liquid**.** Yield: (5.1 g, 93%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.34-7.28 (m, 5H, Ph), 4.21- 4.11 (q, J = 7.0 Hz, 2H, -CH_2), 3.82 (q, J = 6.5 Hz, 1H, -CH-Ph), 3.71-3.53 (m, 1H, -CH), 3.53-3.22 (m, 3H, ring –CH_2 & -CH), 2.94-2.78 (m, 2H, -CH_2), 1.41 (s, 9H, *t*-boc), 1.35 (d, J = 6.6 Hz, 3H, -CH_3), 1.33-1.18 (m, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 172.4, 154.1, 145.2, 128.4, 127.2, 126.6, 79.4, 60.9, 59.2, 58.5, 57.0, 56.8, 52.2, 51.8, 49.6, 48.7, 46.8, 46.5, 28.3, 24.5, 14.1. ; **MS (EI)** *m/z* 362.22, Found 363.38 [M + H], 385.40 [M + Na].

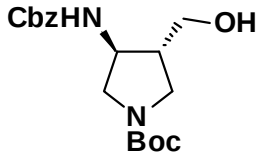
(3*S*,4*R*)-1-N-Boc-3-((*R*)-1-phenylethylamino)-4-hydroxymethylpyrrolidine (18)



Compound **17** (6.3 g, 17.4 mmol) was taken in absolute ethanol and cooled to 0° C. To it was added sodium borohydride (2.0 g, 52.6 mmol) and the reaction mixture was stirred at room temperature for 12 hrs. Ethanol was evaporated and the residue was taken in water and extracted with ethyl acetate twice. The combined ethyl acetate extracts were washed with water, brine, dried over sodium sulphate and purified to get compound **18** as colourless liquid. Yield: (4.5g, 81%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.45-7.15 (m, 5H, Ph), 3.95- 3.60 (m, 3H, -$\underline{CH_2}$-OH, -$\underline{CH}$-Ph), 3.55-3.10 (m, 3H, ring -CH_2, -CH), 2.95-2.55 (m, 3H, ring -CH_2, -CH), 2.15 (br s, 1H, -OH), 1.45-1.25 (m, 12H, -CH_3, *t*-boc), **^{13}C NMR** (50 MHz, $CDCl_3$): δ 154.2, 145.3, 128.6, 127.5, 126.4, 79.4, 65.1, 60.6, 57.4, 51.9, 46.2, 45.6, 28.3, 23.2. ; **MS (EI)** *m/z* 320.21, Found 321.27 [M + H].

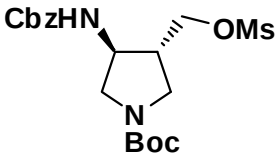
(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-4-hydroxymethyl pyrrolidine (19)



To compound **18** (2g, 6.3 mmol) taken in 10 ml of absolute ethanol in a hydrogenation flask was added 10%-Pd on Carbon (1g, 50% by wt.) and was hydrogenated using parr apparatus for 24 hrs at 60 psi hydrogen pressure. The reaction mixture was filtered on a celite pad and concentrated to obtain crude amine which was dissolved in 30 ml of acetone: water (2:1) and cooled in an ice bath. Sodium bicarbonate (3.1 g, 36.9 mmol) was added to it followed by 50% cbz-chloride in toluene (2.5 ml, 7.5 mmol) and the reaction mixture was stirred at room temperature for 14 hrs. The acetone was removed and water layer was extracted with ethyl acetate twice. The combined ethyl acetate extracts were washed with water, brine, dried over sodium sulphate and purified to yield alcohol **19** as colourless liquid. Yield: (1.8 g, 82%)

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.43-7.32 (m, 5H, cbzPh), 5.11 (s, 2H, cbz -CH_2), 5.08-4.97 (m, 1H, -NH), 4.24- 3.98 (m, 1H, -C<u>H</u>-NH), 3.90-3.43 (m, 4H, -C<u>H</u>$_2$-OH, 1H ring -C<u>H</u>$_2$, -OH), 3.31-2.61 (m, 3H, ring -<u>CH</u>$_2$, 1 H ring –CH_2), 2.24-2.04 (m, 1H, -CH ring), 1.44 (s, 9H, *t*-boc), **^{13}C NMR** (50 MHz, $CDCl_3$): δ 156.4, 154.3, 136.0, 128.3, 128.0, 127.9, 79.5, 66.7, 61.3, 52.0, 51.0, 49.6, 46.9, 45.8, 28.2. ; **MS (EI)** *m/z* 350.18, Found 373.39 [M + Na].

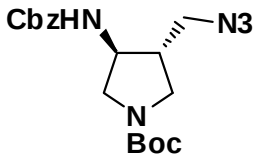
(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-4-((methylsulfonyl)oxy) methylpyrrolidine (20)



To a solution of hydroxyl compound **19** (2.0 g, 5.7 mmol) in 20 ml of dry DCM at 0°C, triethylamine (3.1 mL, 22.2 mmol) was added followed by mesyl chloride (0.9 ml, 11.5 mmol) dropwise with constant stirring. After the addition was complete, the ice bath was removed and stirring continued at ambient temperature for 2 hrs. The reaction mixture was diluted with 50 ml of DCM and washed with water, followed by brine and dried over anhydrous Na_2SO_4. The organic layer was evaporated which on purification gave the mesyl **20** as colourless liquid. Yield: (1.5 g, 62%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.48-7.32 (m, 5H, cbzPh), 5.10 (s, 2H, cbz -CH_2), 5.08-4.97 (m, 1H, -NH), 4.37- 4.05 (m, 3H, -C<u>H</u>-NH, -CH_2), 3.90-3.58 (m, 2H, -C<u>H</u>$_2$-OMs), 3.31-3.08 (m, 2H, -CH_2), 3.02 (s, 3H, -OMS), 2.60-2.35 (m, 1H, -CH), 1.45 (s, 9H, *t*-boc), **^{13}C NMR** (50 MHz, $CDCl_3$): δ 155.8, 154.0, 136.0, 128.5, 128.2, 128.1, 80.0, 68.4, 67.0, 37.2, 29.6, 28.3. ; **MS (EI)** *m/z* 428.16, Found 451.65 [M + Na].

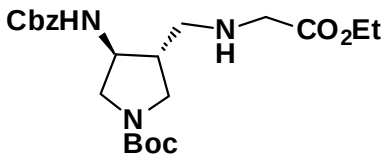
(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-4-azidomethyl pyrrolidine (21)



A stirred mixture of the mesylate **20** (2.0 g, 4.7 mmol) and NaN_3 (2.4 g, 37.2 mmol) in 20 ml of DMF under nitrogen was heated at 65-70°C for 12 h. After cooling, the solvent was evaporated under reduced pressure and the residue was taken in ethyl acetate and washed with water, dried over sodium sulphate and purified by column chromatography to afford azide **21.** Yield: (1.6 g, 91%)

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.44-7.31 (m, 5H, cbzPh), 5.10 (s, 2H, cbz -CH_2), 5.08-4.97 (m, 1H, -NH), 4.15- 3.95 (m, 1H, -C<u>H</u>-NH), 3.85-3.70 (m, 1H –CH_2), 3.68-3.27 (m, 3H, -C<u>H</u>$_2$-N_3, 1H –CH_2), 3.24-3.01 (m, 2H, -CH_2), 2.38-2.11 (m, 1H, -CH), 1.45 (s, 9H, *t*-boc), **^{13}C NMR** (50 MHz, $CDCl_3$): 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 398.18 [M + Na].

N-[(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-(methylpyrrolidin-4-yl)]-glycine ethyl ester (22)

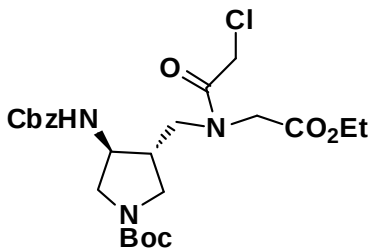


To azide **21** (2.0g, 5.3 mmol), taken in methanol (20ml) in a hydrogenation flask, was added Raney-Ni (500 mg) and the mixture was hydrogenated in a parr apparatus at 45 psi for about 3 hrs during which the TLC indicates disappearance of the starting azide. The solvent was filtered over a celite pad and concentrated to get a colourless oil. The oil was taken in dry acetonitrile (30 ml), to which was added K_2CO_3 (1.47 g, 10.6mmol) and the mixture was stirred under nitrogen at 0° C for 15 min followed by the addition of ethyl bromoacetate (0.71 ml, 6.4 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 **22** as a light yellow oil. Yield: (1.6 g, 69%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.46-7.28 (m, 5H, cbzPh), 5.10 (s, 2H, cbz -CH_2), 4.18 (q, J = 7.2 Hz, 2H, ester-CH_2), 4.07- 3.71 (m, 2H, -C$\underline{H}_2$-NH), 3.68-3.46 (m, 2H –CH-NH, ring –CH_2), 3.38 (s, 2H, -C$\underline{H}_2$-CO_2Et), 3.22-2.98 (m, 2H, -CH_2), 2.80-2.48 (m, 1H, ring –CH_2), 2.24-2.02 (m, 1H, -CH), 1.45 (s, 9H, *t*-boc), 1.27 (t, J = 7.2 Hz, 3H, ester -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 172.2, 156.0, 154.2, 136.1, 128.4, 128.1, 128.0, 79.5, 66.7, 60.7, 50.9, 50.5, 28.3, 14.1. ; **MS (EI)** *m/z* 435.24, Found 436.38 [M + H], 458.38 [M + Na].

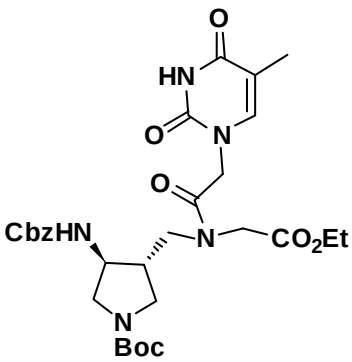
N-[(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-(methylpyrrolidin-4-yl)]-N-(chloroacetyl)-glycine ethyl ester (23)



To an ice-cooled stirred solution of **22** (1.5 g, 3.4 mmol) in 15 ml of DCM was added triethylamine (1.9 ml, 13.8mmol). The reaction mixture was stirred for 15 min followed by the addition of chloroacetyl chloride (0.55 ml, 6.8 mmol) and 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 **23** as a colourless liquid which was used without further purification. Yield: (1.3 g, 74 %).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.48-7.18 (m, 5H, cbzPh), 5.09 (s, 2H, cbz -CH_2), 4.36-3.71 (m, 8H, ester –CH_2, –CH_2-Cl, -$\underline{CH}_2$-CO_2Et, ring –CH_2), 3.68-3.46 (m, 2H, ring –CH_2), 3.42-3.22 (m, 1H, -CH), 3.22-2.92 (m, 2H, -CH_2), 2.60-2.25 (m, 1H, ring –CH), 1.44 (s, 9H, *t*-boc), 1.35-1.14 (m, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 169.1, 167.8, 156.2, 154.5, 136.4, 128.7, 128.3, 79.9, 67.1, 62.3, 53.8, 51.1, 50.2, 47.8, 41.5, 41.2, 40.7, 28.5, 14.3.; **MS (EI)** *m/z* 511.21, Found 534.30 [M + Na].

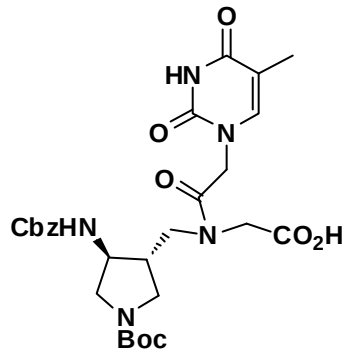
N-[(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-(methylpyrrolidin-4-yl)]-N-(thymin-1-acetyl)-glycine ethyl ester (24)



To compound **23** (1.23g, 2.4 mmol) taken in 10 ml of DMF was added potassium carbonate (0.365 g, 2.90 mmol) followed by thymine (0.4 g, 2.90mmol) at room temperature. The reaction mixture was stirred for 18 hrs at room temperature after which it was diluted with 100 ml of ethyl acetate and washed with water, brine and dried over sodium sulphate and purified by column chromatography to obtain compound **24** as a colourless liquid. Yield: (1.1 g, 76 %).

^{1}H NMR (200 MHz, $CDCl_3$): 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_2), 4.43-4.16 (m, 4H, –CH_2-T, -$\underline{CH_2}$-CO_2Et), 4.11-3.48 (m, 6H, 3 x CH_2), 3.42-3.02 (m, 3H, -CH_2 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_3). **^{13}C NMR** (50 MHz, $CDCl_3$): 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 624.57 [M + Na].

N-[(3*S*,4*R*)-1-N-Boc-3-Cbz-amino-(methylpyrrolidin-4-yl)]-N-(thymin-1-acetyl)-glycine (25)

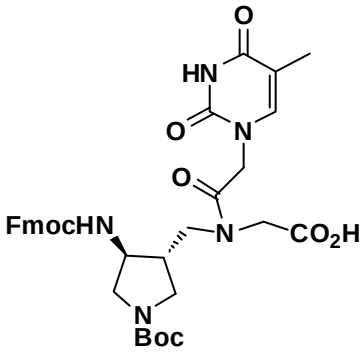


To a solution of **24** (1.0 g, 1.7 mmol) in ethanol (4 mL) was added 1M aq. LiOH (5 mL). The reaction was stirred at room temperature for 2.0hrs. The excess of alkali was neutralized by Dowex H+ resin, which was filtered off. The filtrate was evaporated to get the PNA monomer **25** as white foam with good yield. Yield: (0.82 g, 86%).

$[\alpha]_D^{25}$ = - 10 (*c* 1.0, DMSO); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 11.28 (br s, 1H, -COOH), 7.92 (br s, 1H, N$\underline{H}$Cbz), 7.48-7.25 (m, 5H, cbzPh), 7.21 (s, 1H, thymine), 5.09-4.92 (m, 2H, cbz -CH_2), 4.66-4.16 (m, 2H, –CH_2-T), 4.03-3.70 (m, 4H, ring –CH_2, -$\underline{CH_2}$-CO_2H), 3.48-2.77 (m, 5H, 2 x CH_2 (4H), -CH), 2.41-2.22 (m, 1H, -CH), 1.70 (s, 3H, Thymine), 1.35 (s, 9H, *t*-boc). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 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 597.54 [M + Na].

N-[(3*S*,4*R*)-1-N-Boc-3-Fmoc-amino-(methylpyrrolidin-4-yl)]-N-(thymin-1-acetyl)-glycine (26)



The acid **25** (450 mg, 0.78 mmol) was dissolved in 10 ml of methanol and was hydrogenated at 60 psi for 8 hrs after the addition of 10% Pd-C (200 mg). The reaction mixture thus obtained was filtered on celite bed and was concentrated to get crude amine which was dissolved in 10 ml of acetone: water (2:1), cooled in an ice bath and sodium bicarbonate (400 mg, 4.76 mmol) was added followed by the addition of Fmoc-succinimide (400 mg, 1.19 mmol). The reaction mixture was stirred at room temperature for 12 hrs. The acetone was removed and the water layer was neutralised with saturated solution of potassium hydrogen sulphate and extracted with ethyl acetate three times. The combined ethyl acetate extracts were washed with water, brine, dried over sodium sulphate and purified by column chromatography to obtain PNA monomer **26** as white solid. Yield: (350 mg).

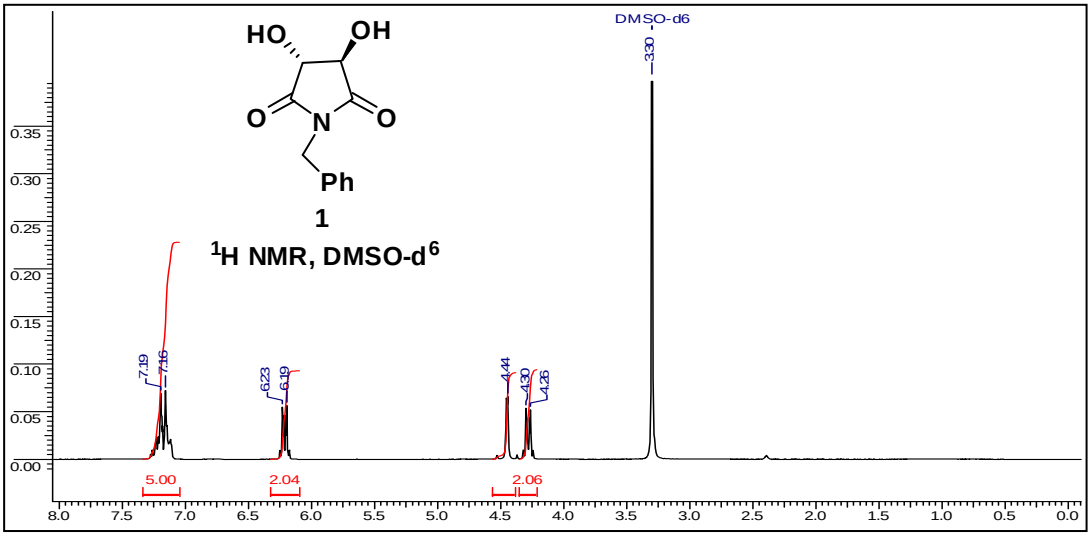
$[\alpha]_D^{25}$ = - 7 (*c* 1.0, DMSO); **^{1}H NMR** (200 MHz, $CDCl_3$): δ 11.31 (br s, 1H, -COOH), 7.93-7.78 (m, 2H, fmoc), 7.73-7.54 (m, 3H, fmoc), 7.48-7.19 (m, 5H, fmoc (3H), -CH thymine, -NH), 4.70-3.96 (m, 9H, 4 x CH_2 (8H), fmoc (1H)), 3.42-2.85 (m, 5H, 2 –CH_2, -C<u>H</u>-NH), 2.41-2.25 (m, 1H, -CH), 1.71 (s, 3H, Thymine), 1.36 (s, 9H, *t*-boc). **^{13}C NMR** (50 MHz, $CDCl_3$): 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* 661.27, Found 684.63 [M + Na].

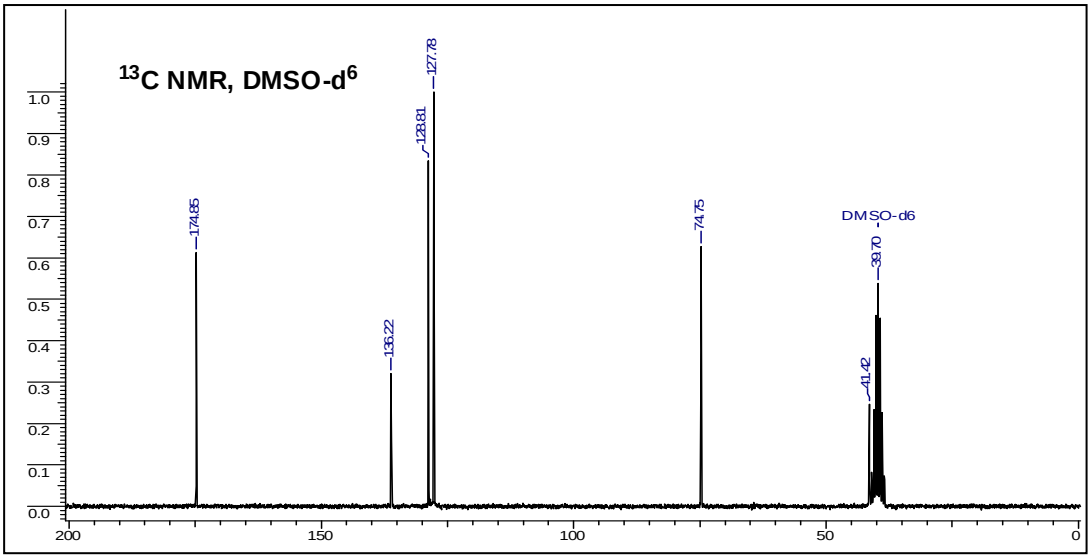
2.5 References

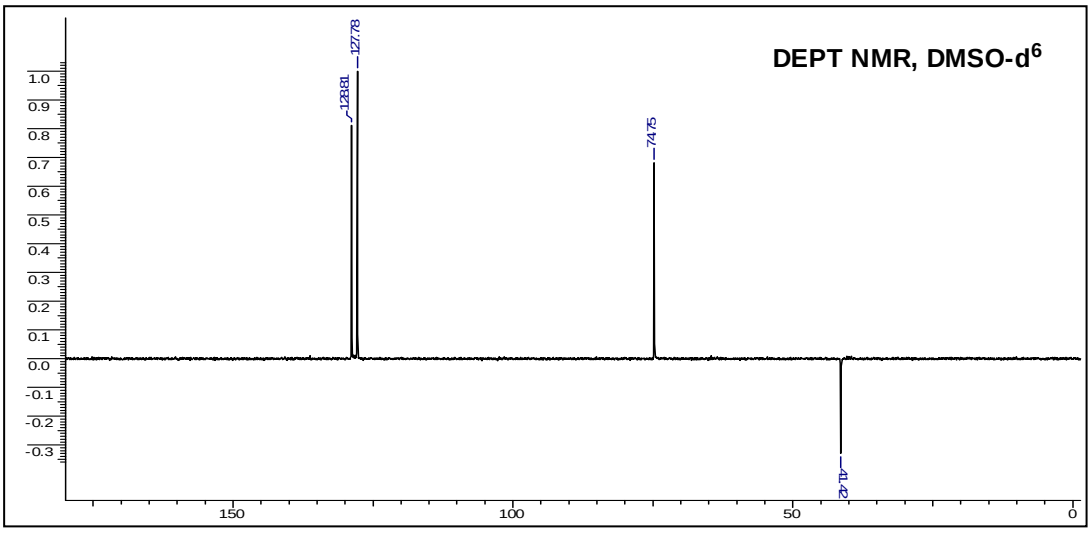
1. (a) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990,** *90*, 543. (b) Noble, S. A.; Bonham, M. A.; Bisi, J. E.; Bruckenstein, D. A.; Brown P. H.; Brown, S. C. *Drug Dev*. *Res*. **1995**, *34*, 184.
2. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci.* **1978**, *75*, 280.
3. (a) Mesmaekar, A. 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 sited there in.
9. Kumar, V.A.; Ganesh, K. N. *Acc. chem. research.* **2005**, 38, 404.

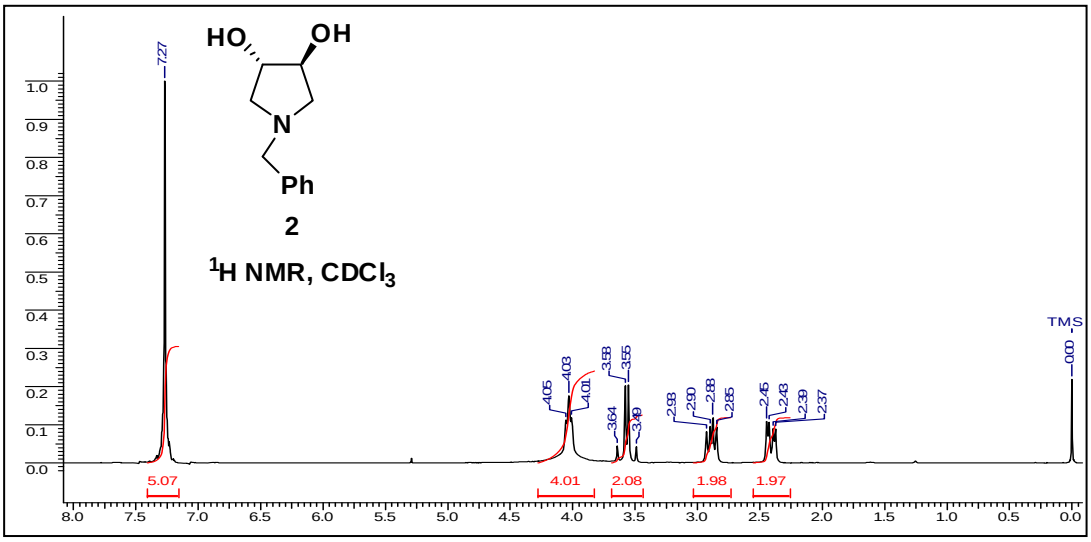
10. Egholm, M.; Buchardt, O.; Christensen, L.; Beherns, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566.
11. Govindaraju, T.; Kumar V. A.; Ganesh K. N. *Chem. Commun.* **2004**, 860 .
12. a) Myers, M. C.; Witschi, M. A.; Larionova, N. 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. Lee, H.; LePlae, P. R.; Porter, E. A.; Gellman, S. H. *J. Org. Chem.* **2001,** *66,* 3597.
19. Meltzer, P. C.; Liang, A. Y.; Matsudaira, P. *J. Org. Chem.* **1995,** *60,* 4305.

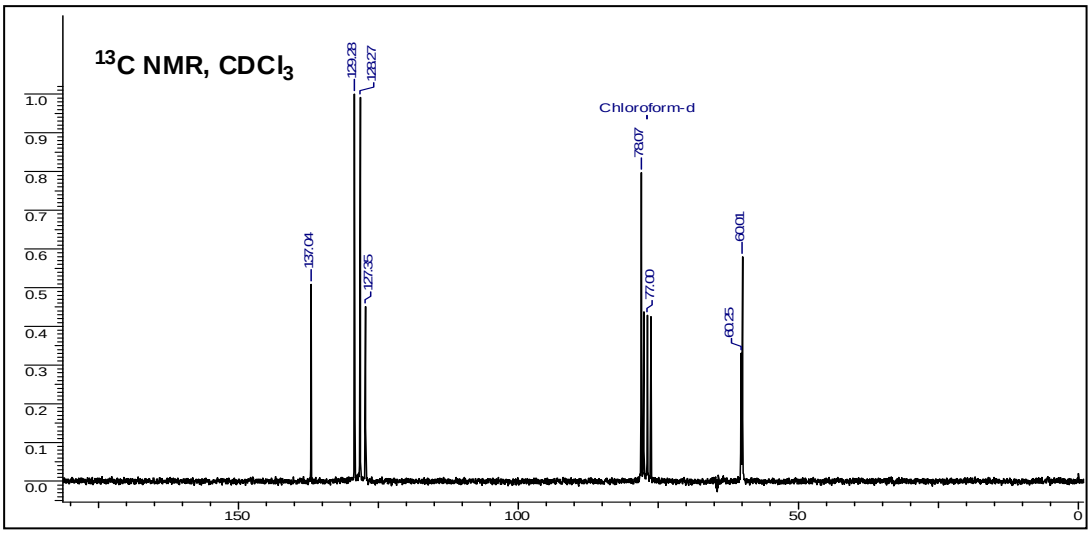
2.6 APPENDIX

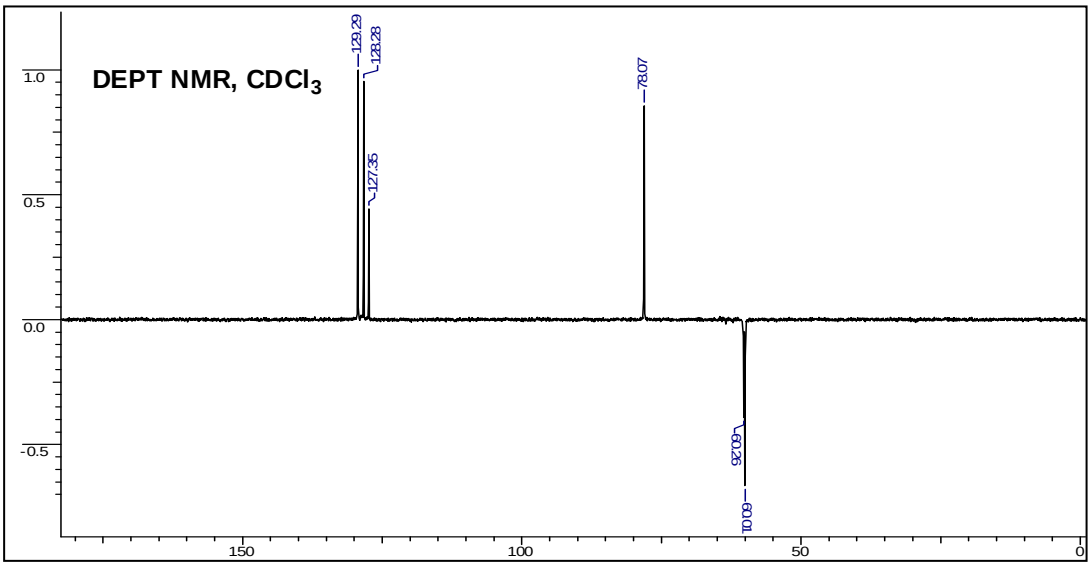

HO
OH
O
N
O
Ph
1
^{1}H NMR, DMSO-d^{6}
DMSO-d6
3.30
7.19
7.16
6.23
6.19
4.44
4.30
4.26
5.00
2.04
2.06

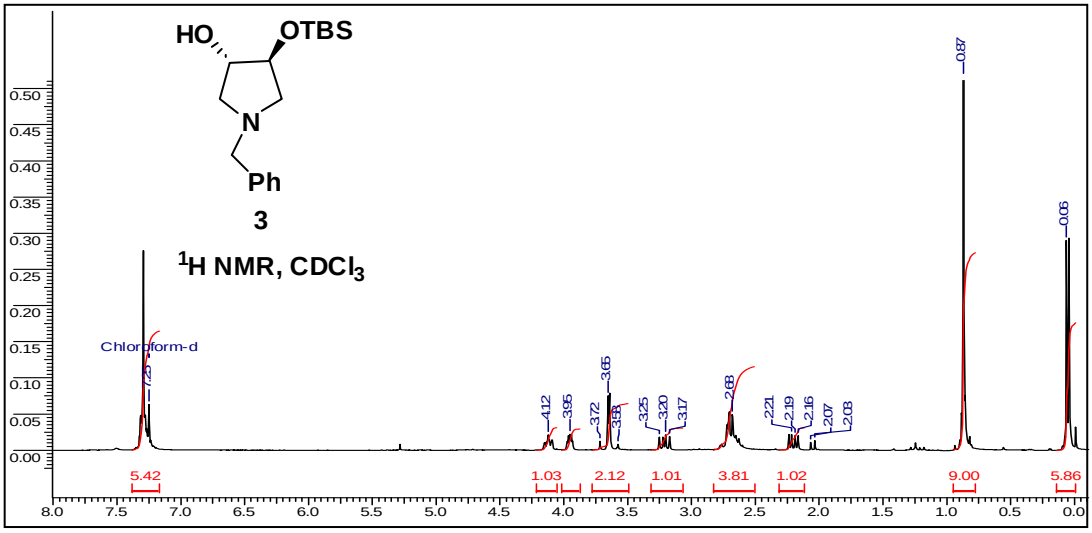

^{13}C NMR, DMSO-d^{6}
127.78
128.81
174.85
136.22
74.75
DMSO-d6
39.70
41.42

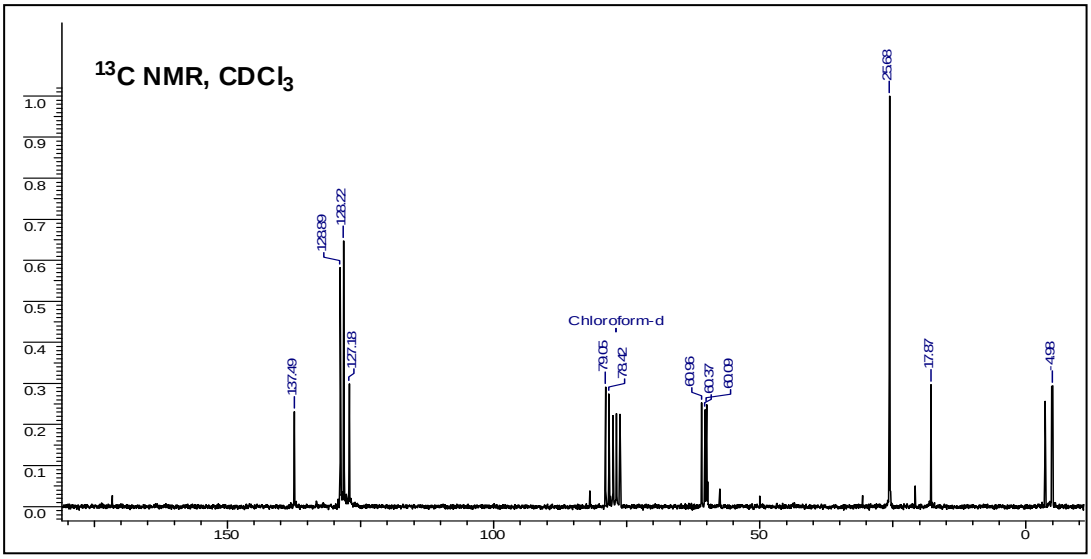

DEPT NMR, DMSO-d^{6}
127.78
128.81
74.75
41.42

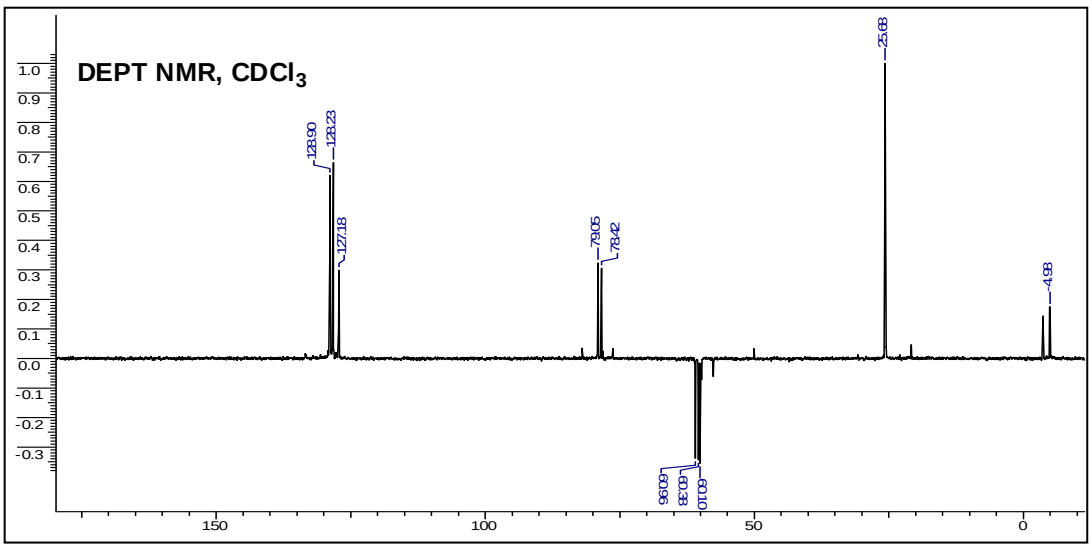

HO
OH
N
Ph
2
^{1}H NMR, CDCl$_3$
7.27
TMS
0.00
5.07
4.01
2.08
1.98
1.97

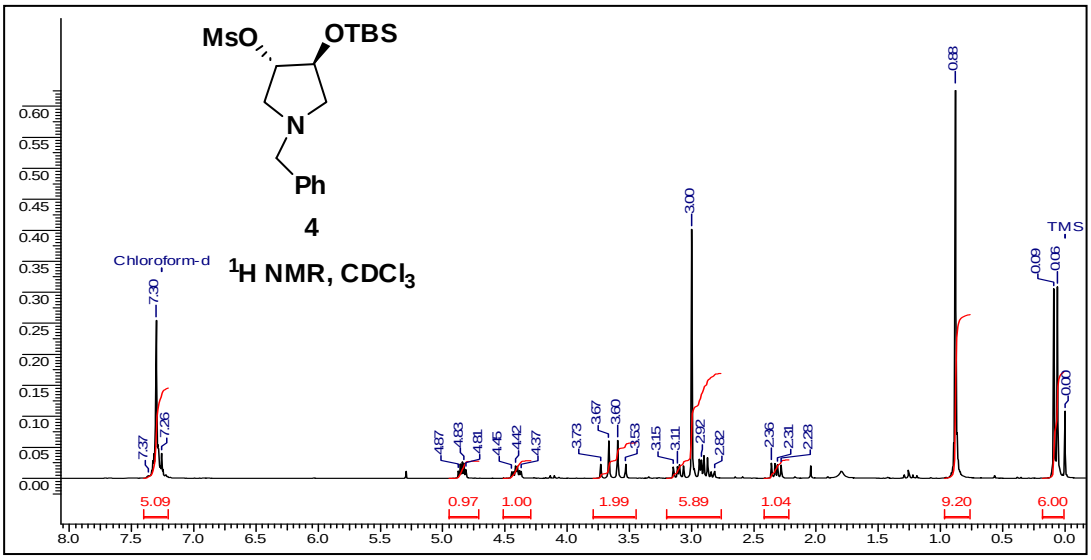

^{13}C NMR, CDCl$_3$
Chloroform-d
137.04
129.28
128.27
127.35
78.07
77.00
60.25
60.01

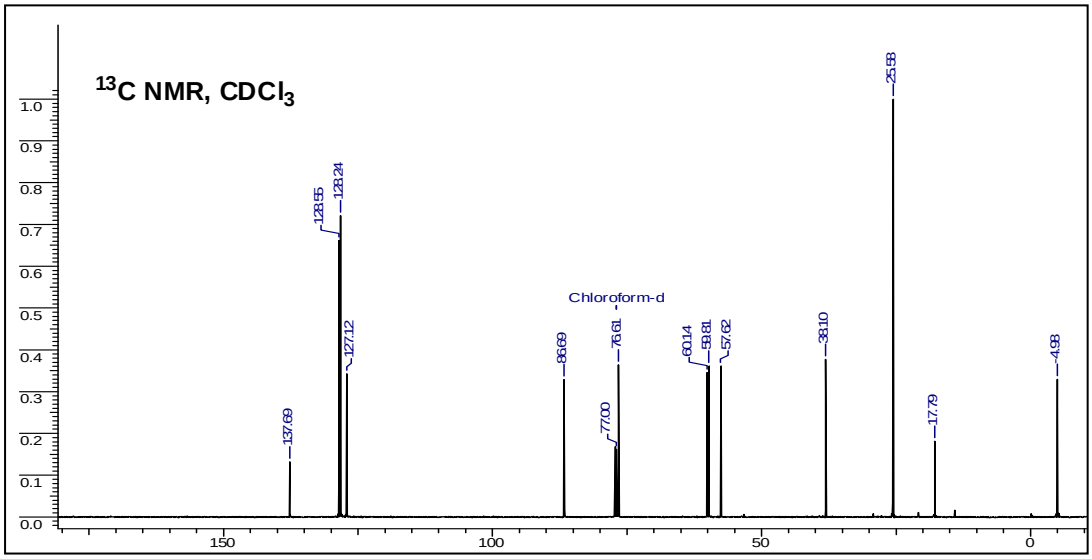

DEPT NMR, CDCl$_3$
129.29
128.28
127.35
78.07
60.26
60.01

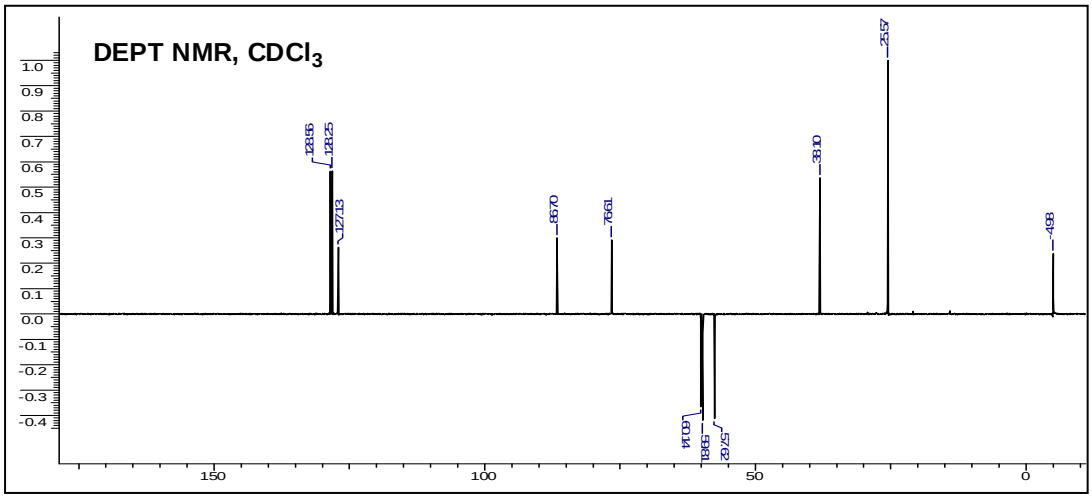

HO
OTBS
N
Ph
3
^{1}H NMR, CDCl$_3$
Chloroform-d
0.87
0.06
7.25
4.12
3.95
3.72
3.65
3.33
3.25
3.20
3.17
2.68
2.21
2.19
2.16
2.07
2.08
5.42
1.03
2.12
1.01
3.81
1.02
9.00
5.86

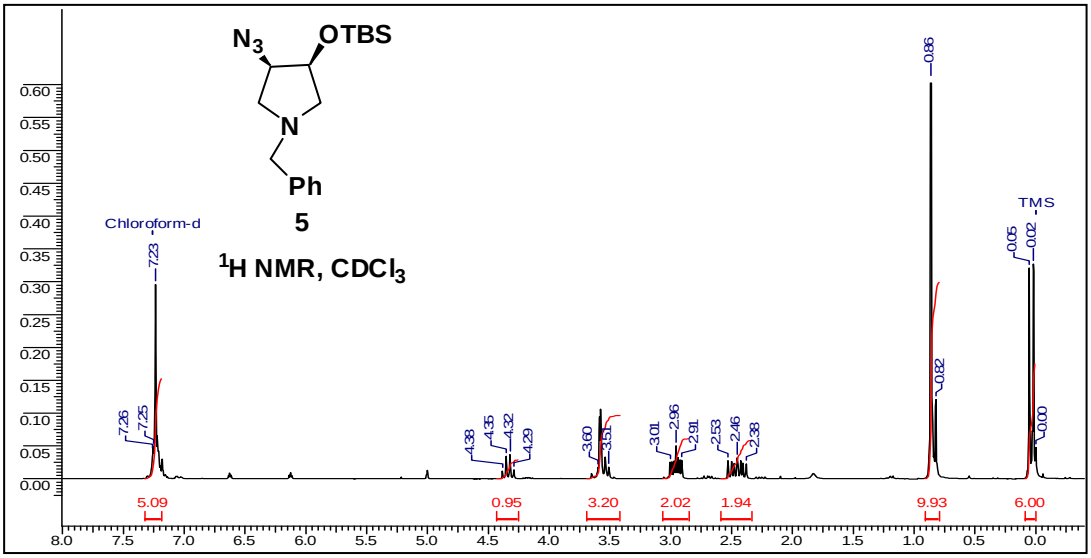

^{13}C NMR, CDCl$_3$
25.68
128.89
128.22
137.49
127.18
Chloroform-d
79.05
78.42
60.96
60.37
60.09
17.87
-4.98

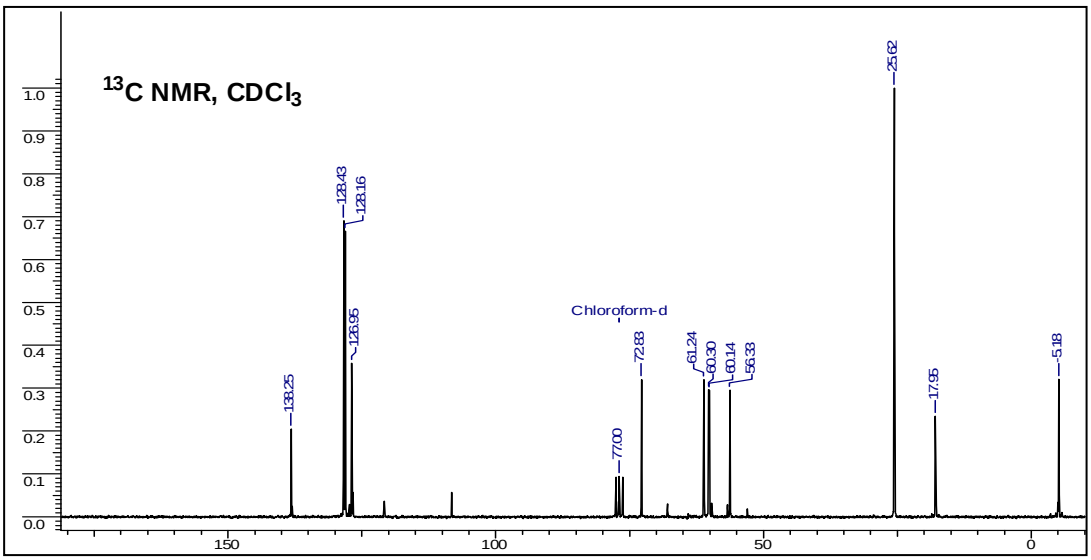

DEPT NMR, CDCl$_3$
25.68
128.90
128.23
127.18
79.05
78.42
-4.98
60.96
60.38
60.10

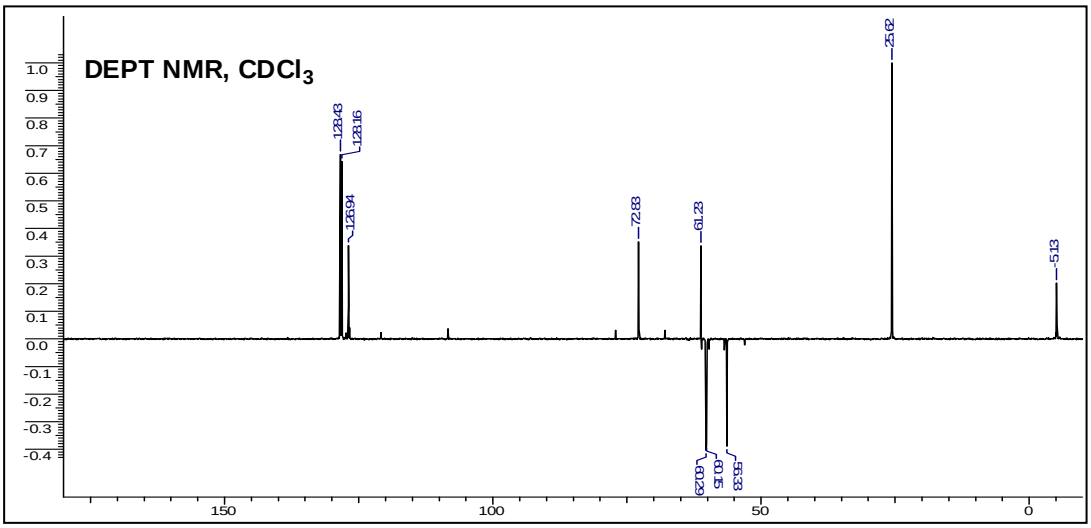

MsO
OTBS
N
Ph
4
^{1}H NMR, CDCl$_3$
Chloroform-d
TMS
5.09
0.97
1.00
1.99
5.89
1.04
9.20
6.00

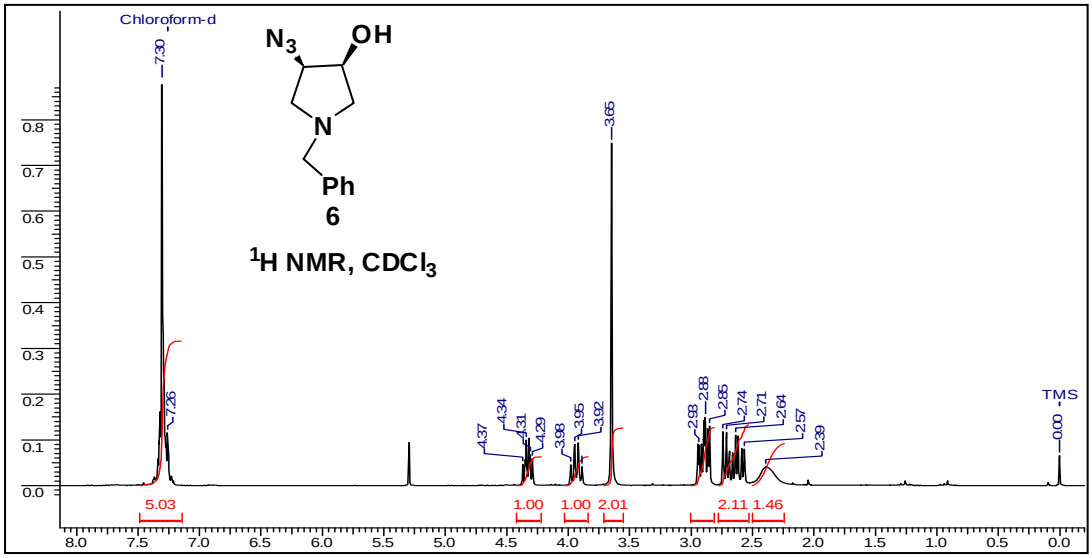

^{13}C NMR, CDCl$_3$
Chloroform-d

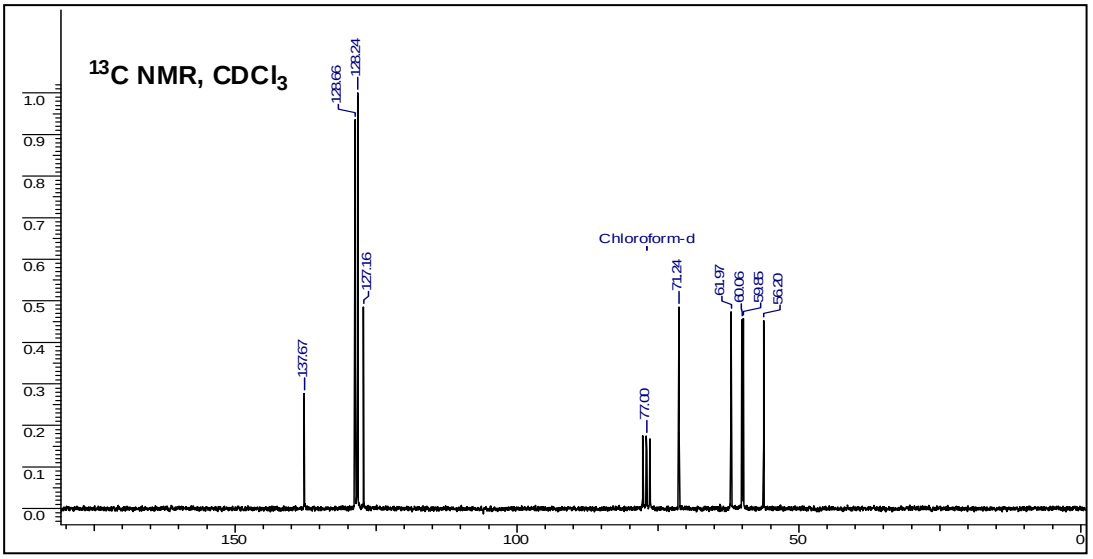

DEPT NMR, CDCl$_3$

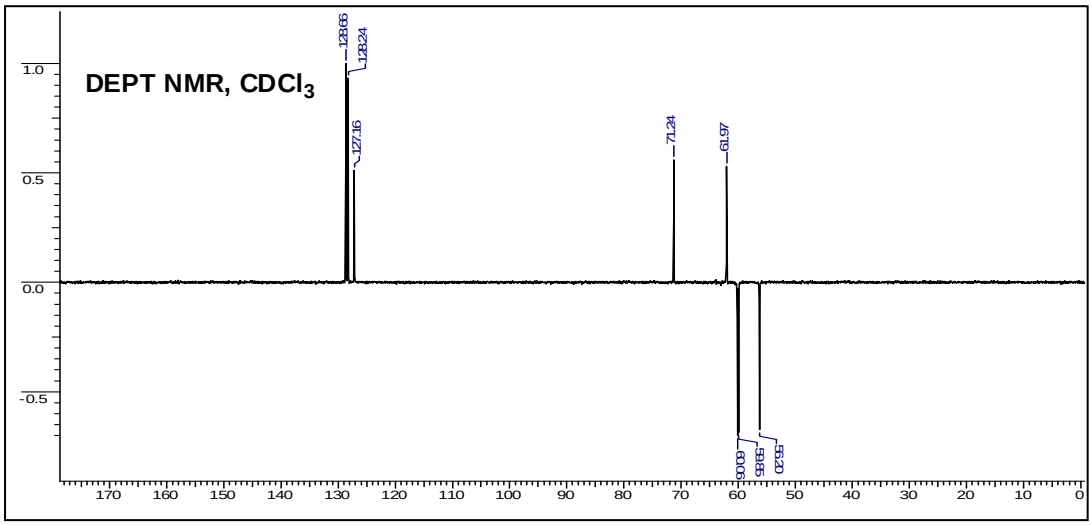

N3
OTBS
N
Ph
5
1H NMR, CDCl3
Chloroform-d
TMS

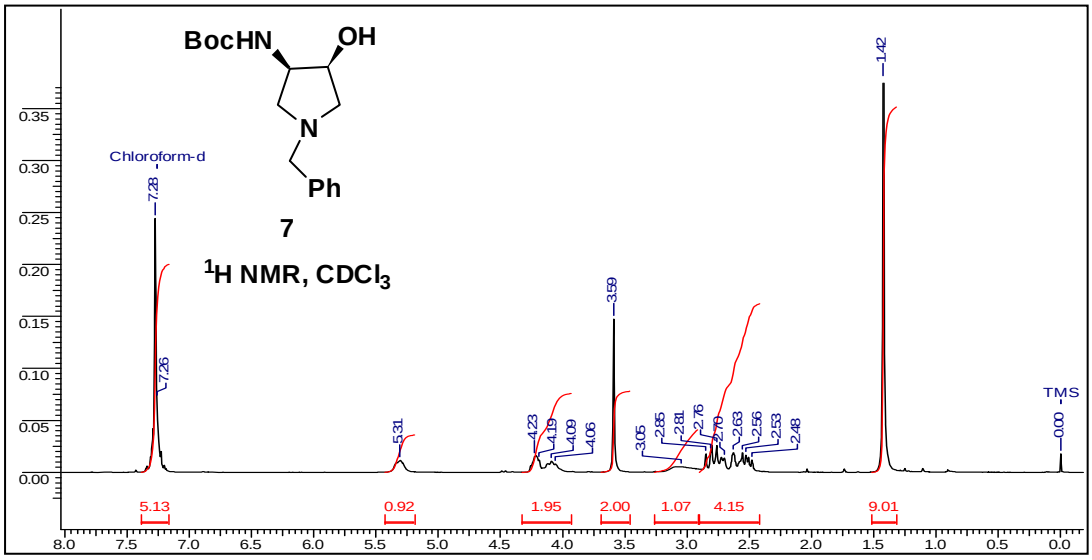

13C NMR, CDCl3
Chloroform-d

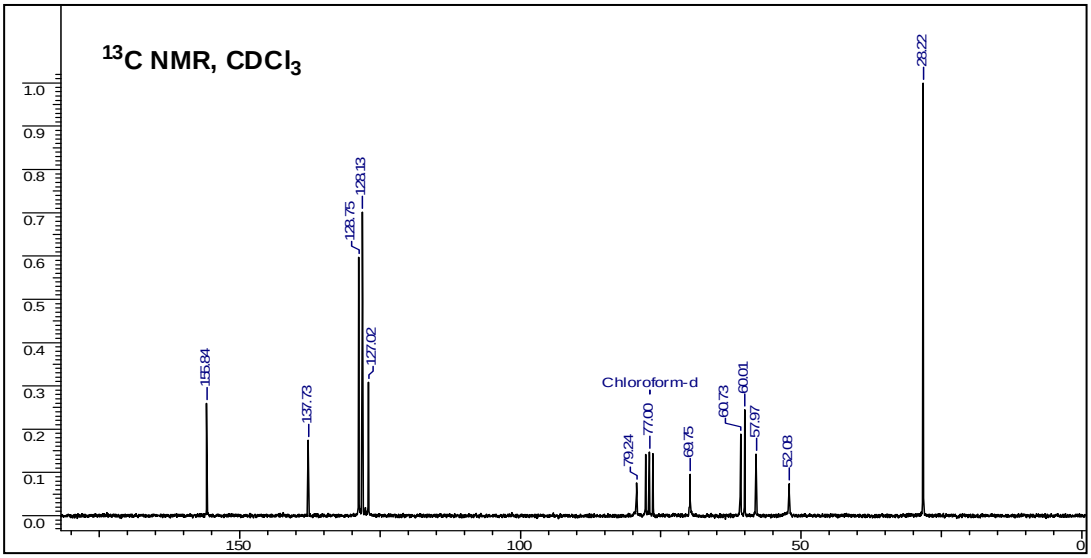

DEPT NMR, CDCl3

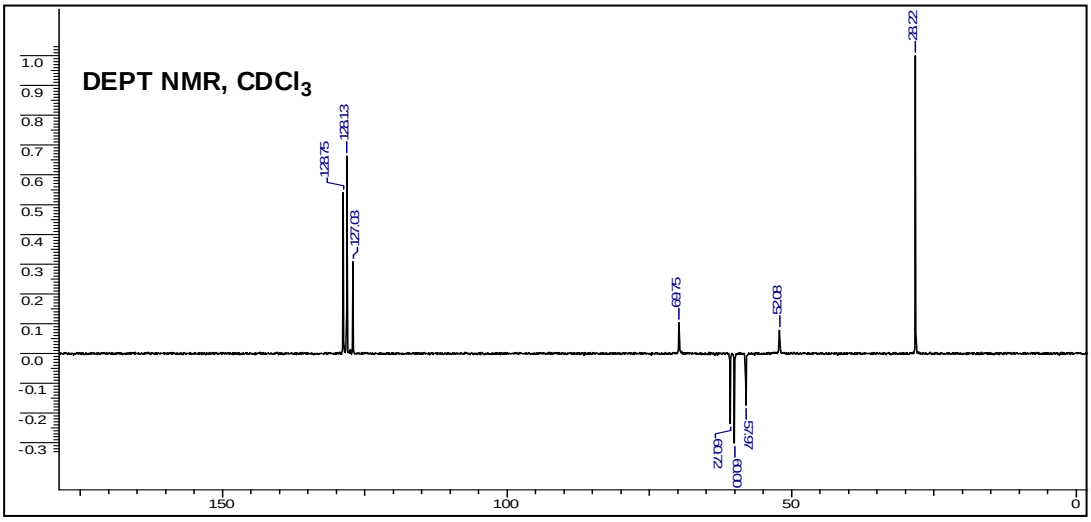

Chloroform-d
7.30
7.26
3.65
4.37
4.34
4.31
4.29
3.98
3.95
3.92
2.93
2.88
2.85
2.74
2.71
2.64
2.57
2.39
TMS
0.00
5.03
1.00
1.00
2.01
2.11
1.46
N3
OH
N
Ph
6
1H NMR, CDCl3

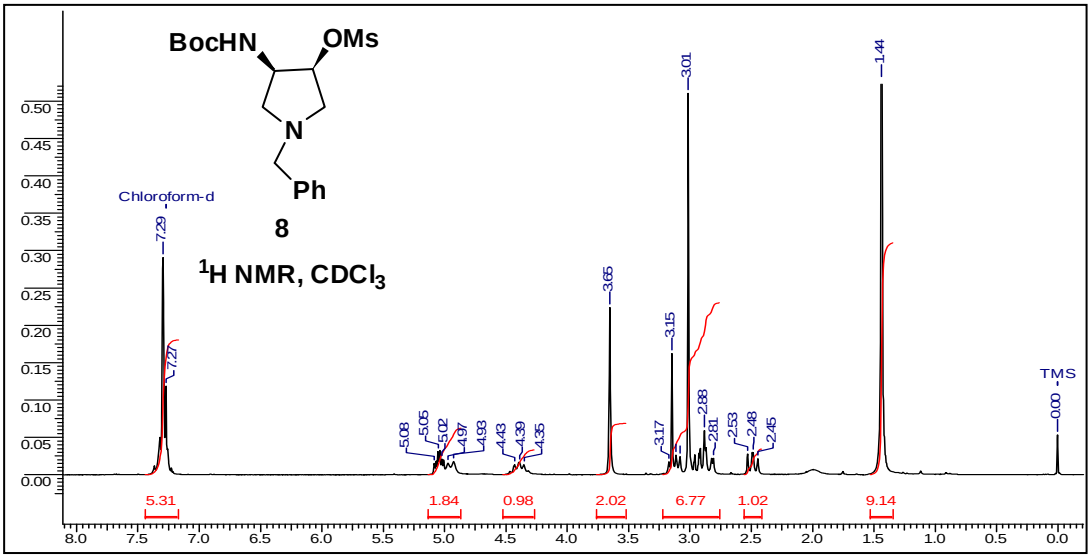

13C NMR, CDCl3
128.66
128.24
127.16
137.67
Chloroform-d
77.00
71.24
61.97
60.06
59.85
56.20

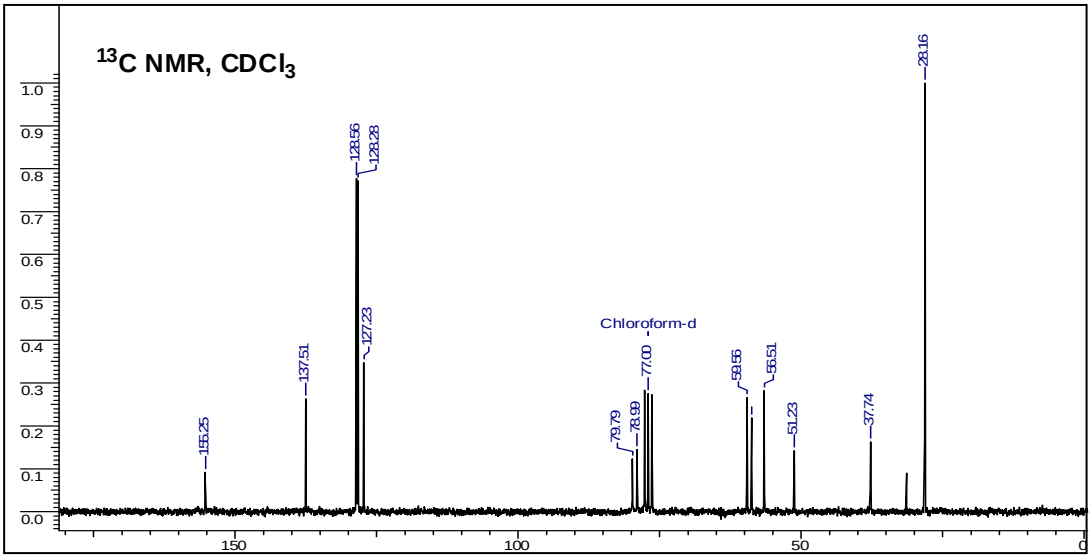

DEPT NMR, CDCl3
128.66
128.24
127.16
71.24
61.97
60.06
59.85
56.20

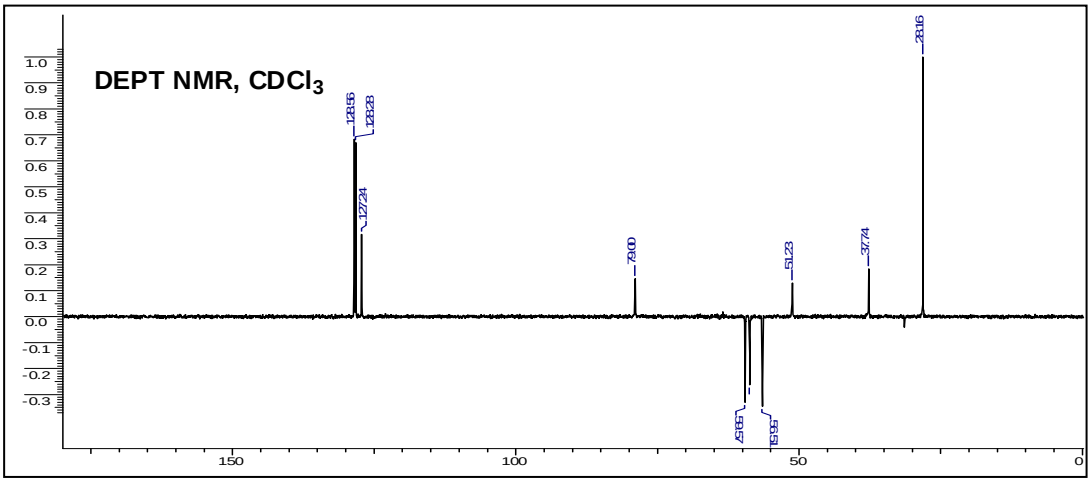

BocHN
OH
N
Ph
7
^{1}H NMR, CDCl$_3$
Chloroform-d
TMS

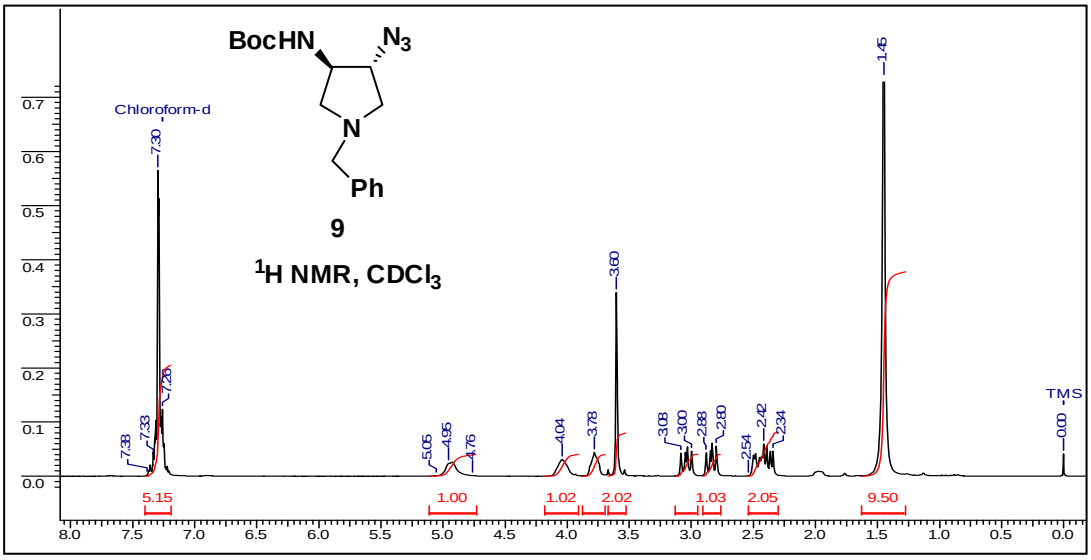

^{13}C NMR, CDCl$_3$
Chloroform-d

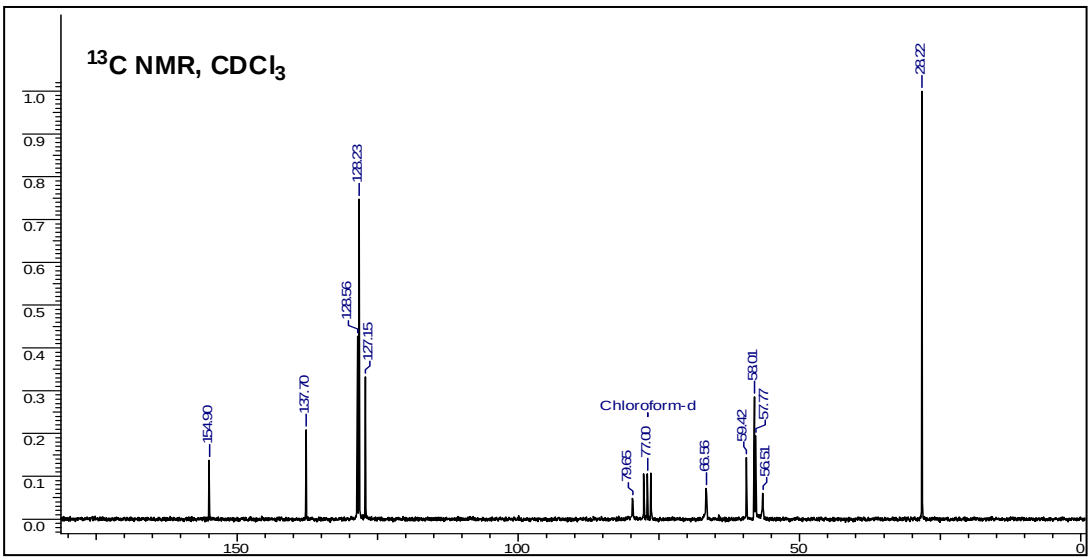

DEPT NMR, CDCl$_3$

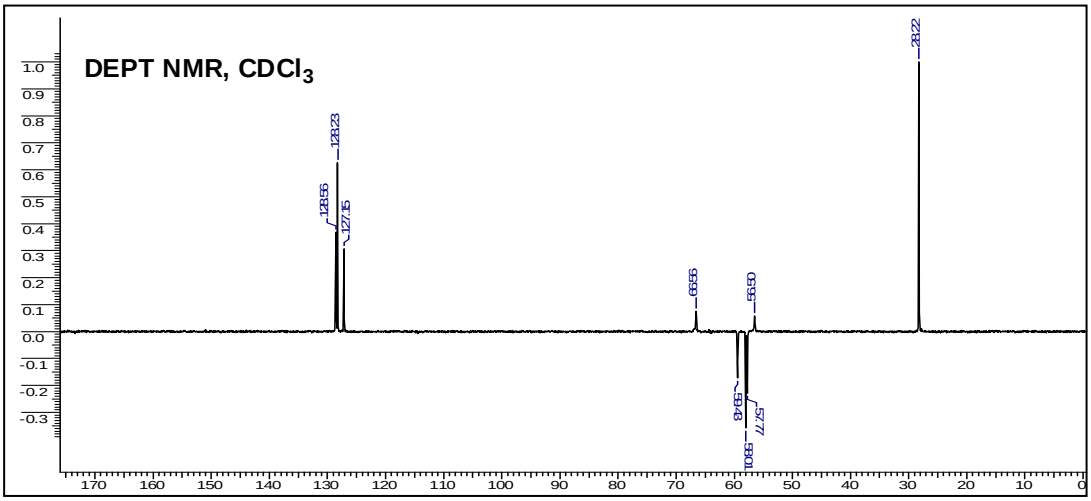

BocHN
OMs
N
Ph
8
Chloroform-d
^{1}H NMR, CDCl$_3$
TMS

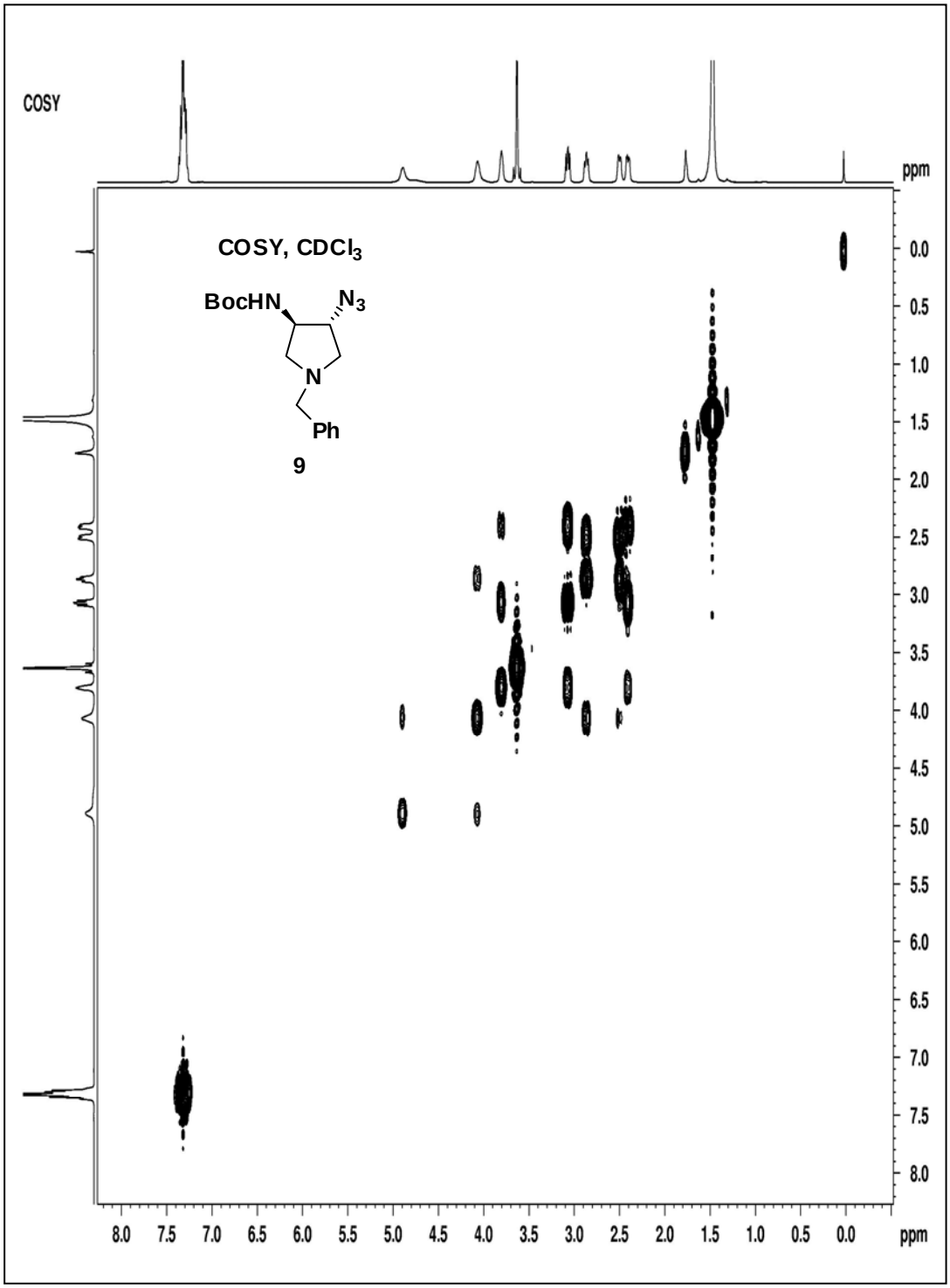

^{13}C NMR, CDCl$_3$
Chloroform-d


DEPT NMR, CDCl$_3$

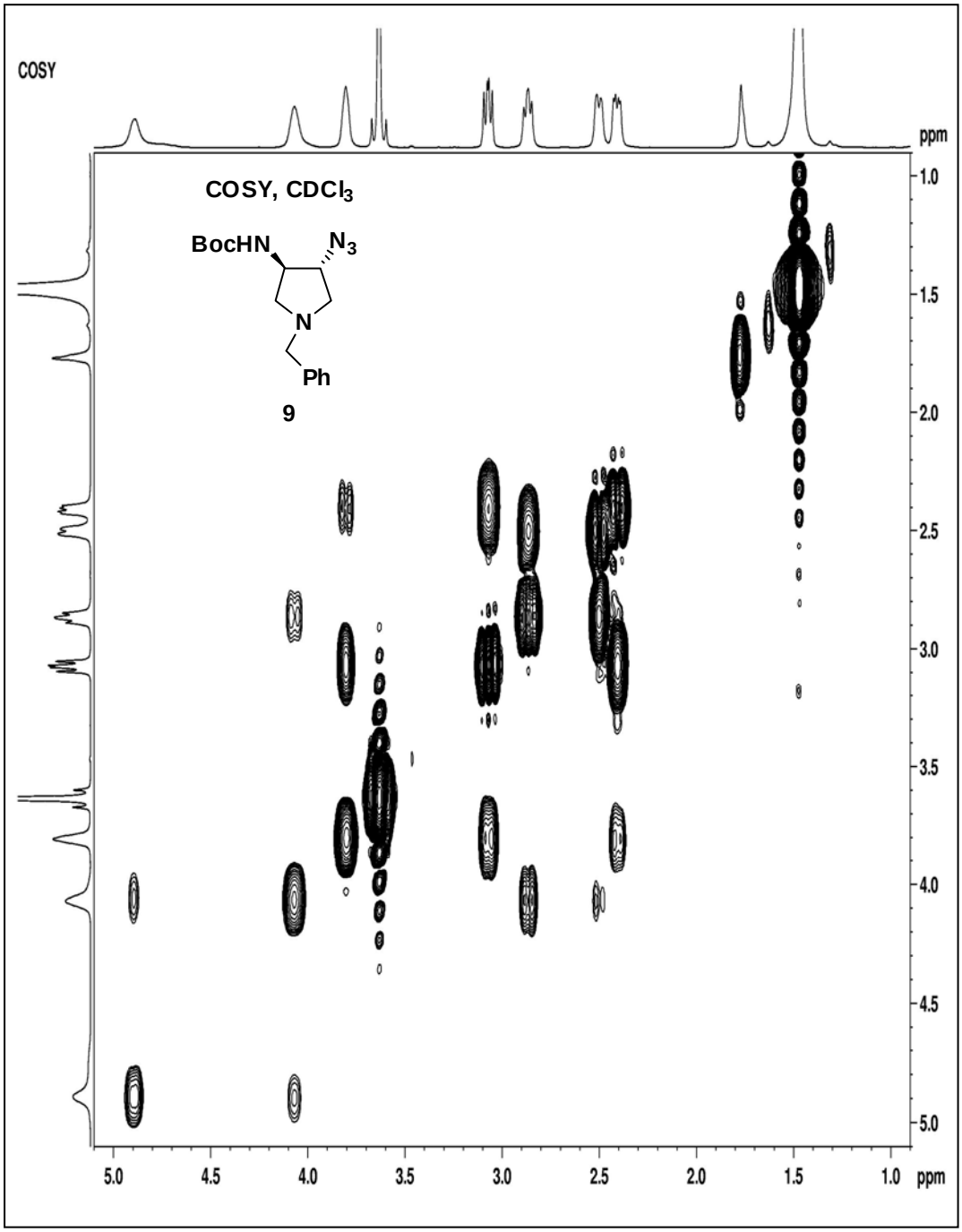

BocHN
N3
N
Ph
9
1H NMR, CDCl3
Chloroform-d
TMS
7.30
7.38
7.33
7.26
5.05
4.95
4.76
4.04
3.78
3.60
3.08
3.00
2.88
2.60
2.54
2.42
2.34
1.45
0.00
5.15
1.00
1.02
2.02
1.03
2.05
9.50

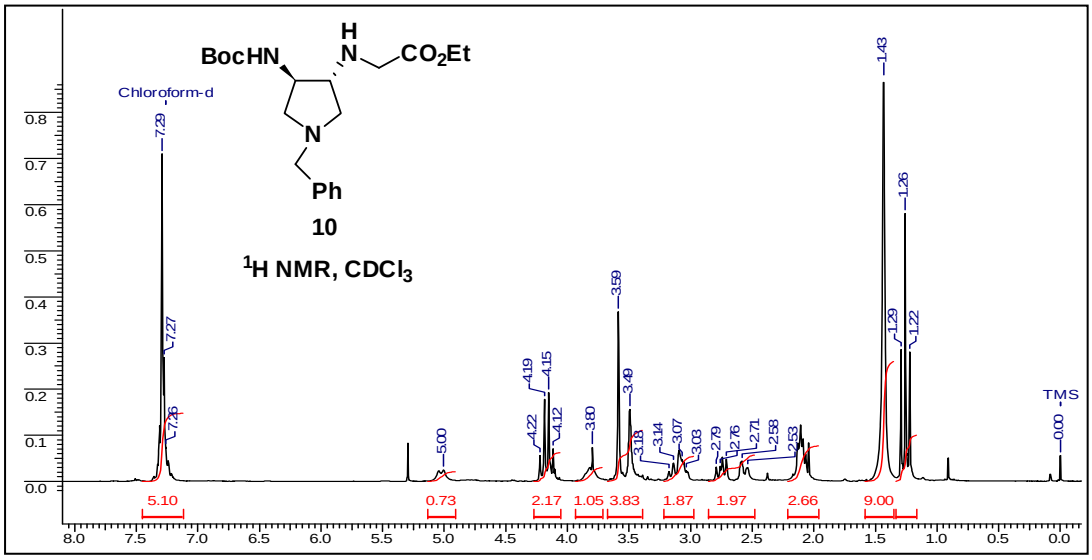

13C NMR, CDCl3
154.90
137.70
128.23
128.56
127.15
Chloroform-d
79.65
77.00
66.55
59.42
53.01
57.77
56.51
28.22

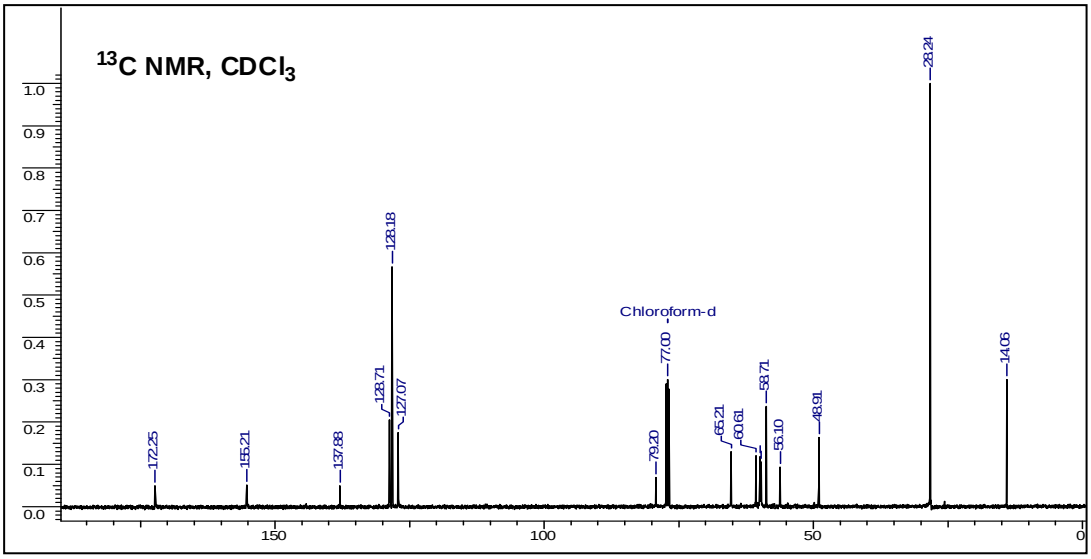

DEPT NMR, CDCl3
128.23
128.56
127.15
66.55
59.42
53.01
28.22

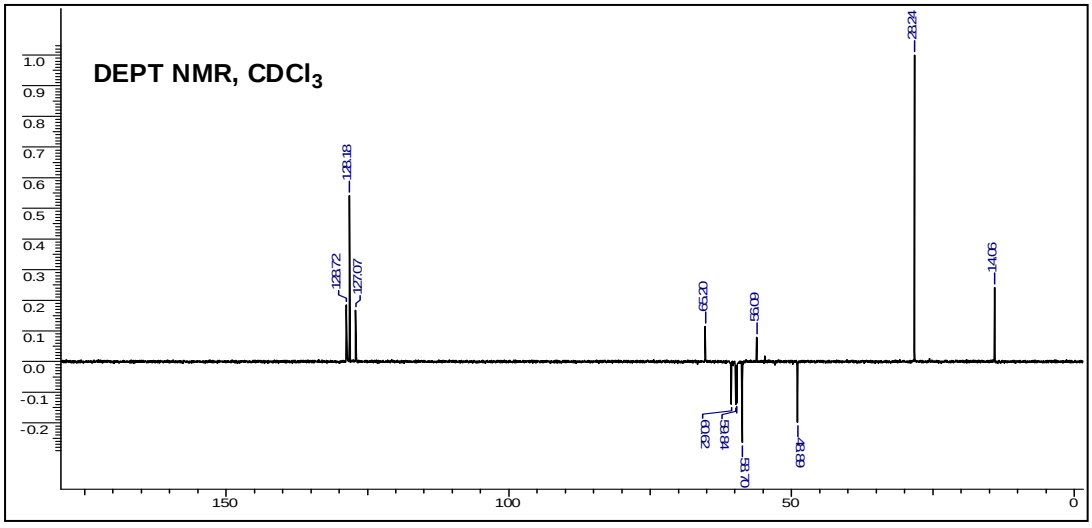

COSY
ppm
COSY, CDCl3
BocHN
N3
N
Ph
9
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
ppm

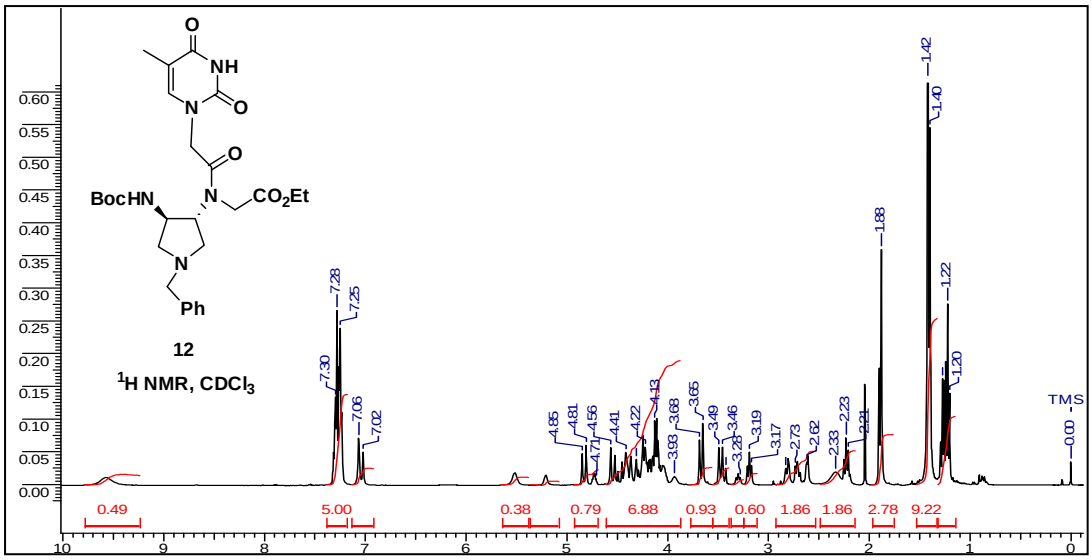
Zoomed region for compound 9, cosy

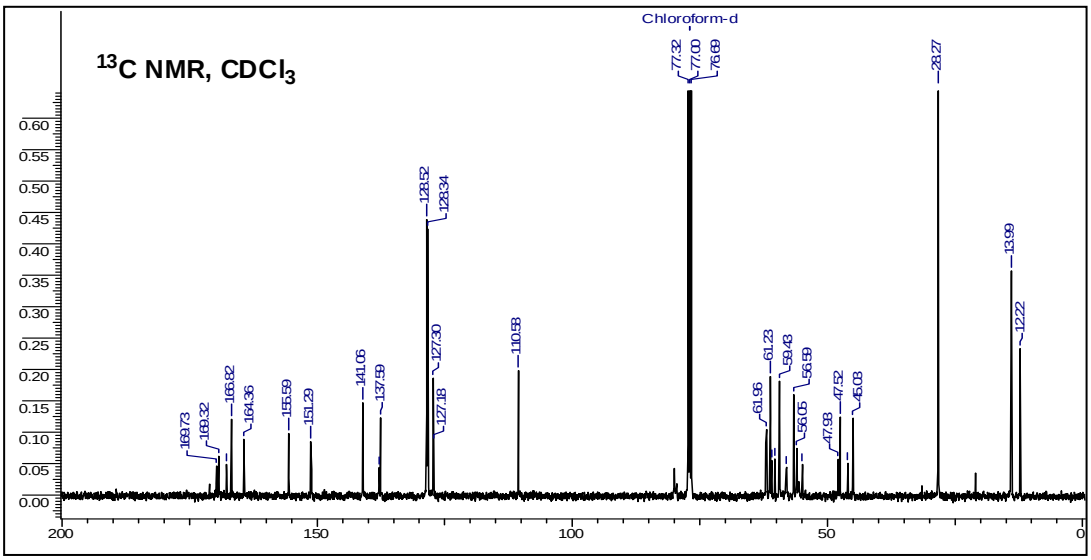


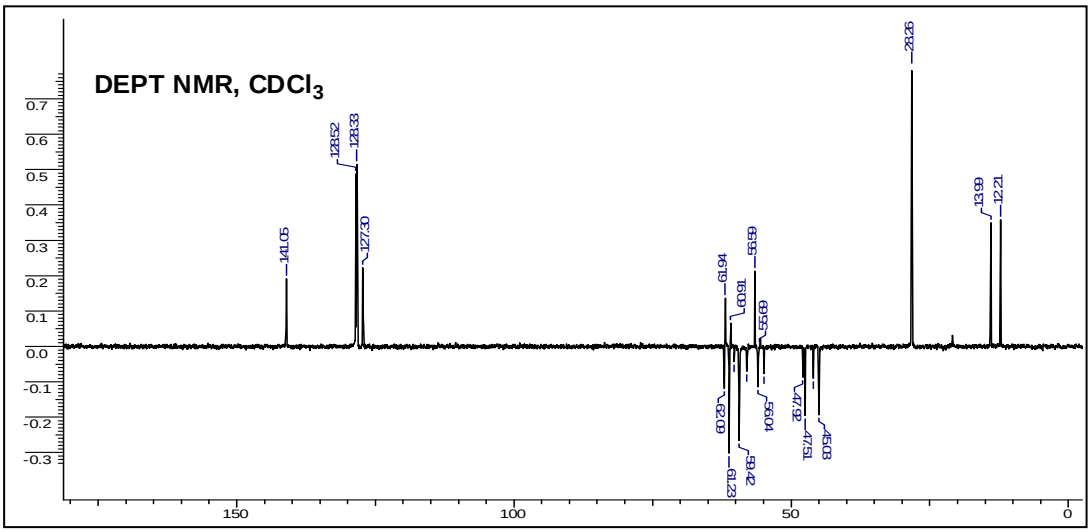

BocHN
H
N
CO2Et
N
Ph
10
1H NMR, CDCl3
Chloroform-d
TMS

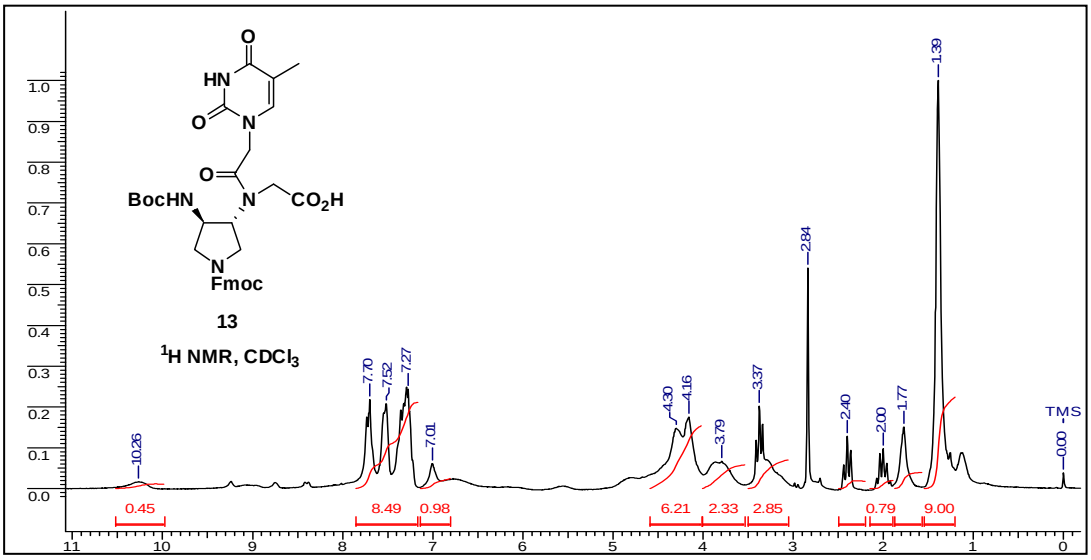

13C NMR, CDCl3
Chloroform-d

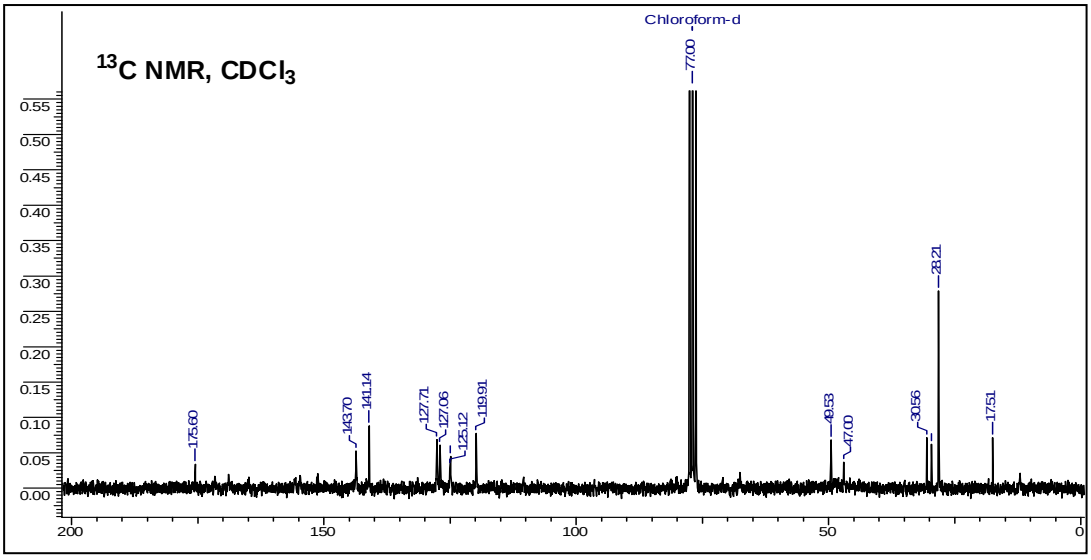

DEPT NMR, CDCl3

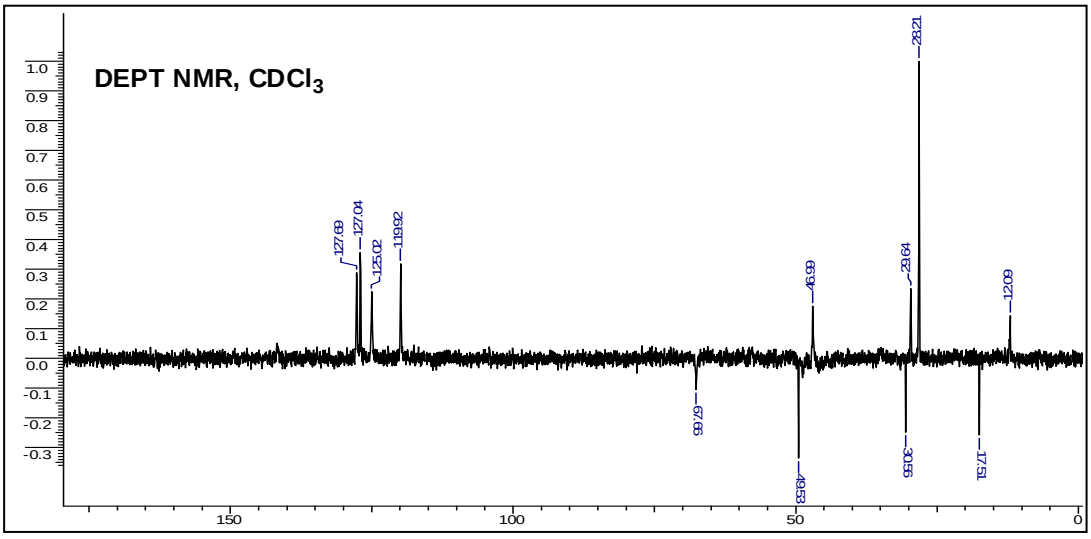

12
^{1}H NMR, CDCl$_3$
TMS

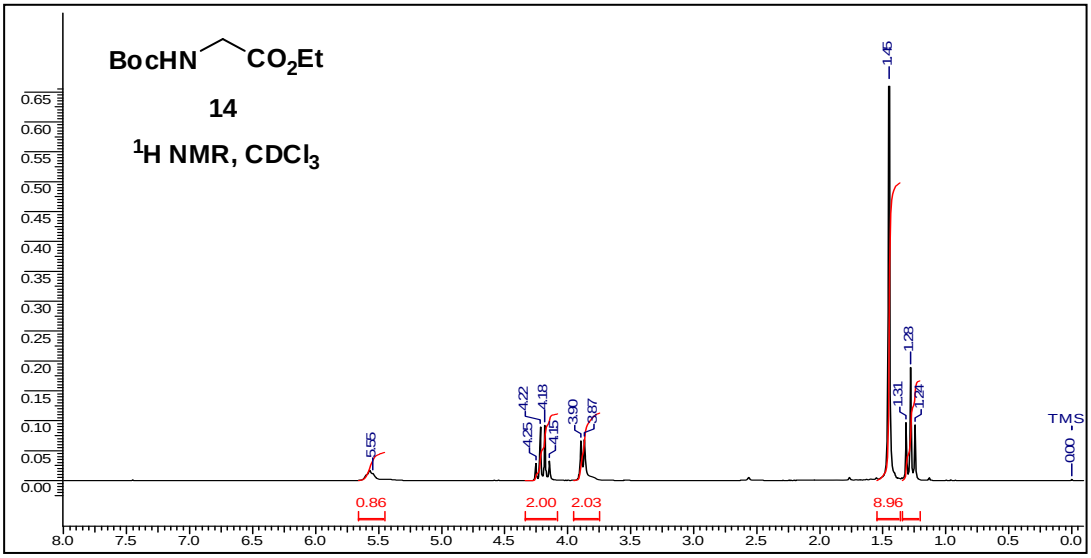

^{13}C NMR, CDCl$_3$
Chloroform-d

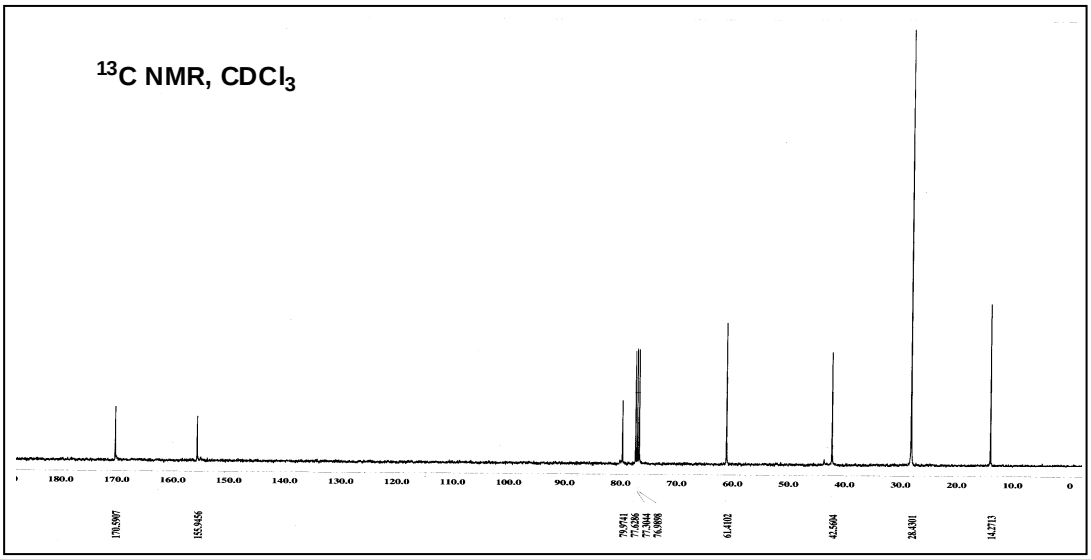

DEPT NMR, CDCl$_3$

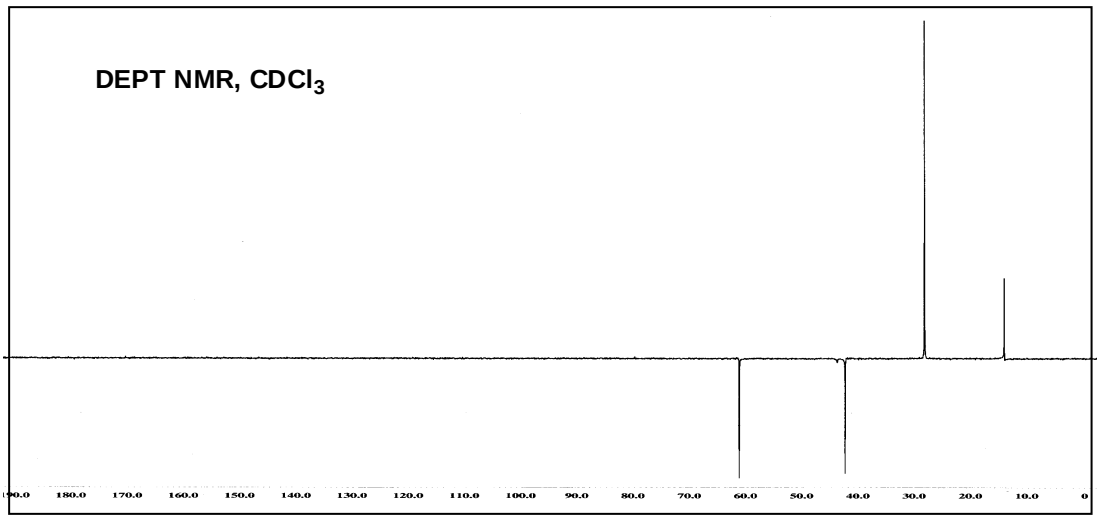

HN
O
N
O
BocHN
N
CO2H
N
Fmoc
13
1H NMR, CDCl3
10.26
7.70
7.52
7.27
7.01
4.30
4.16
3.79
3.37
2.84
2.40
2.00
1.77
1.39
TMS
0.00
0.45
8.49
0.98
6.21
2.33
2.85
0.79
9.00

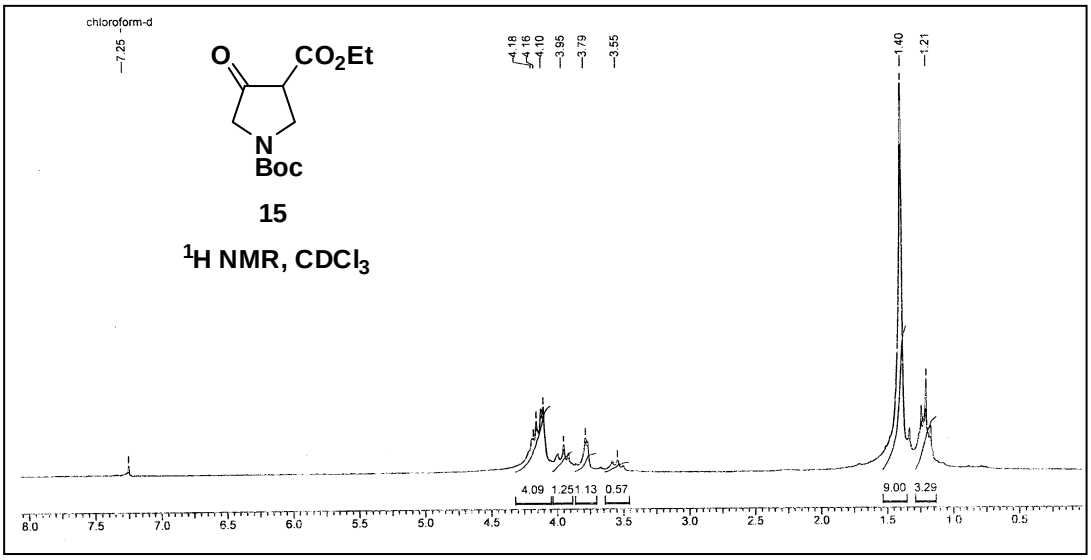

13C NMR, CDCl3
Chloroform-d
77.00
175.60
143.70
141.14
127.71
127.06
125.12
119.91
49.53
47.00
30.56
28.21
17.51

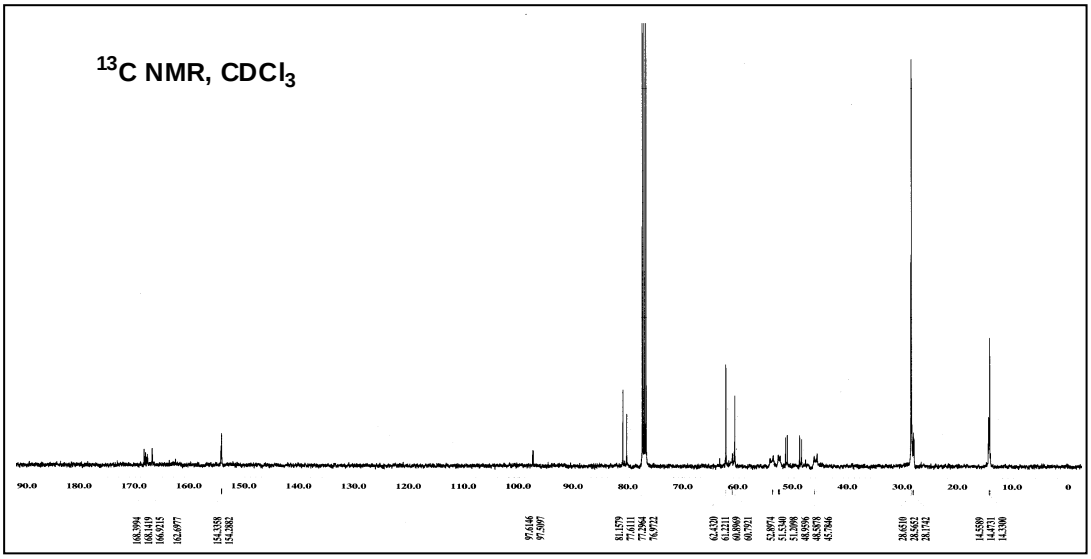

DEPT NMR, CDCl3
127.69
127.04
125.02
119.92
49.53
46.99
30.56
28.21
23.64
17.51
12.09

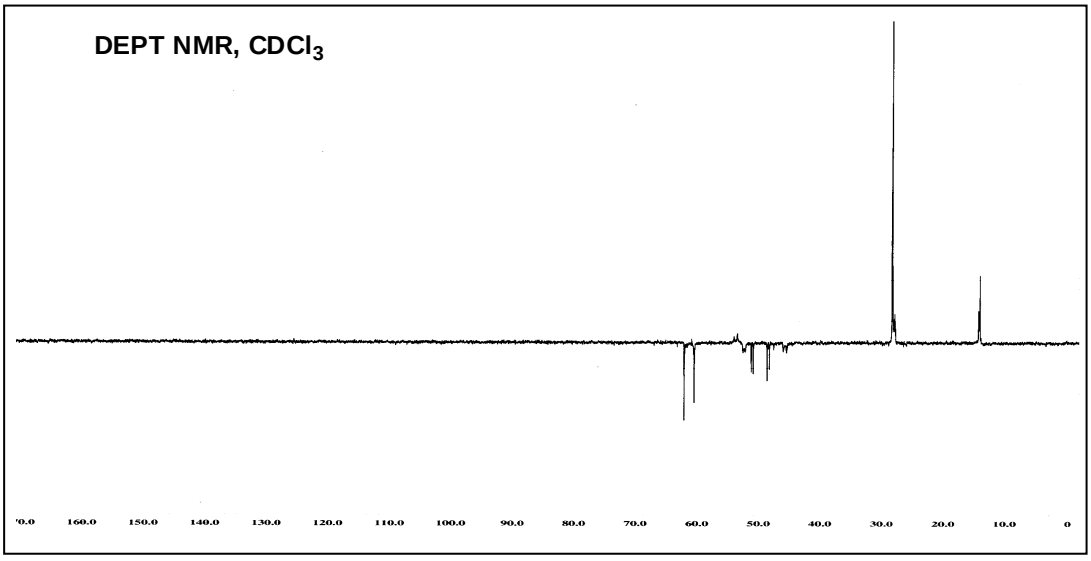

BocHN CO2Et
14
1H NMR, CDCl3
5.55
4.22
4.18
4.25
4.15
3.90
3.87
1.45
1.28
1.31
1.24
TMS
0.00
0.86
2.00
2.03
8.96

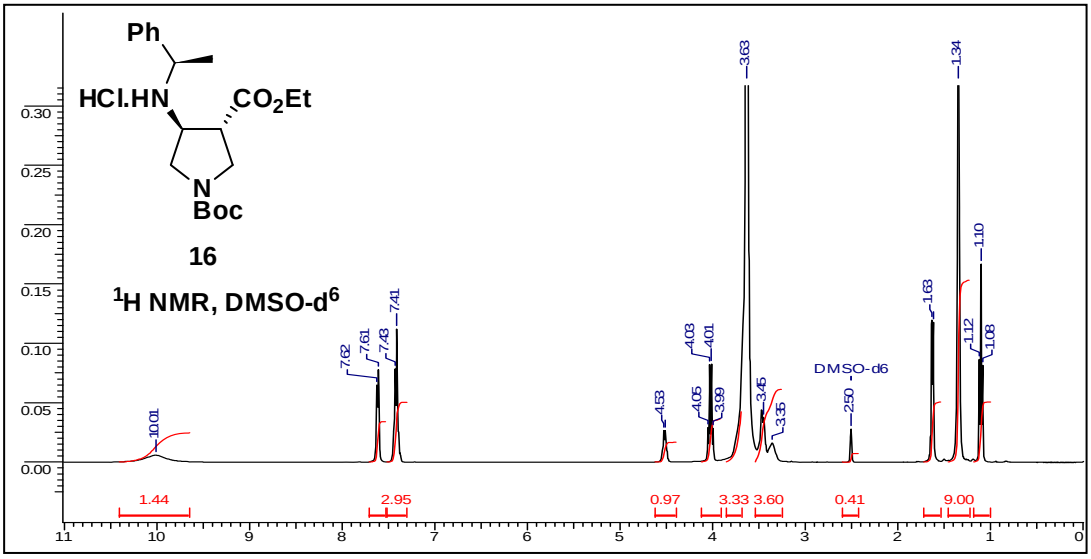

13C NMR, CDCl3

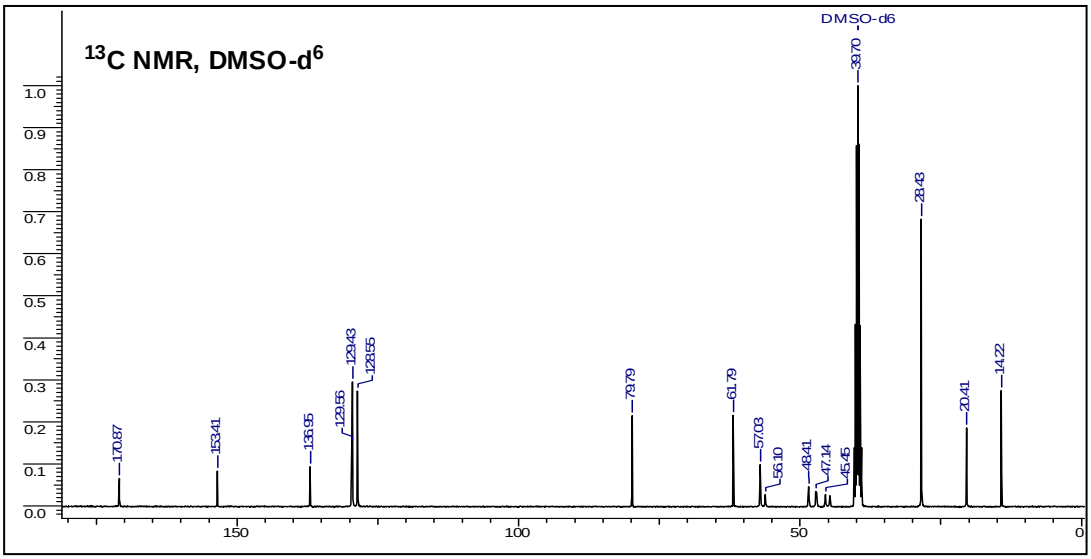

DEPT NMR, CDCl3

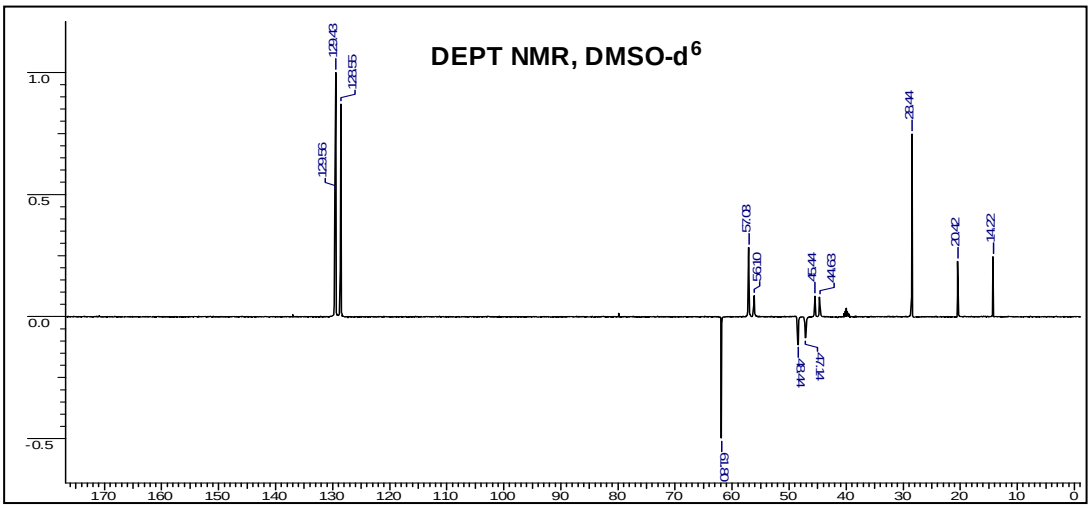

chloroform-d
7.25
O
CO2Et
N
Boc
15
1H NMR, CDCl3
4.18
4.16
4.10
3.95
3.79
3.55
1.40
1.21
4.09
1.25
1.13
0.57
9.00
3.29

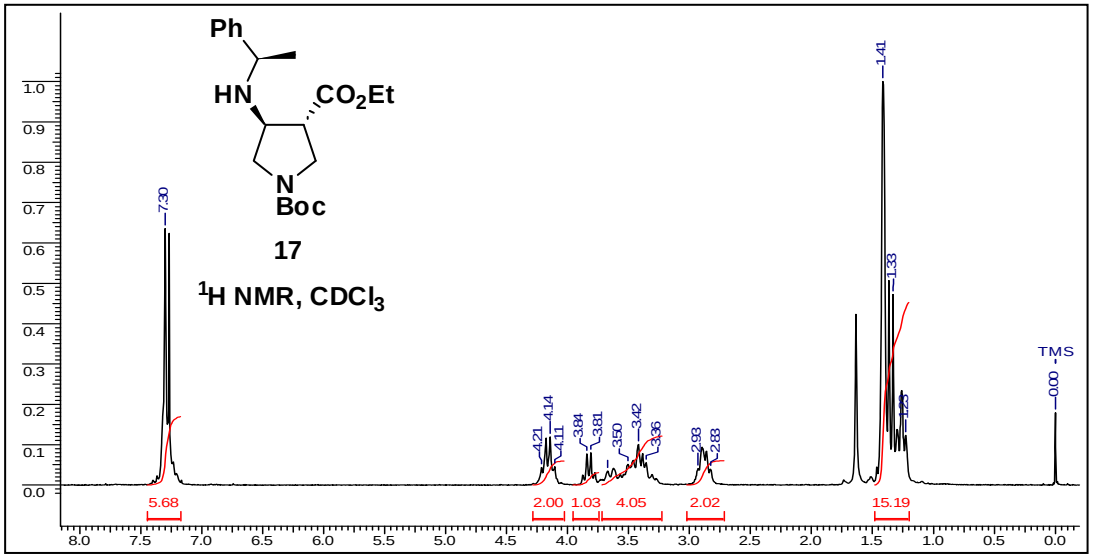

13C NMR, CDCl3

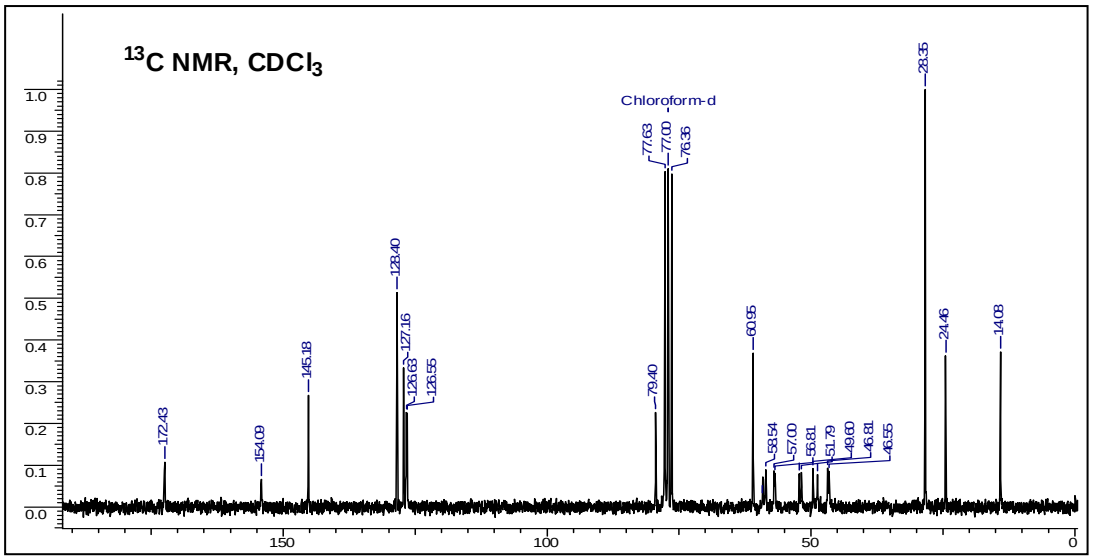

DEPT NMR, CDCl3

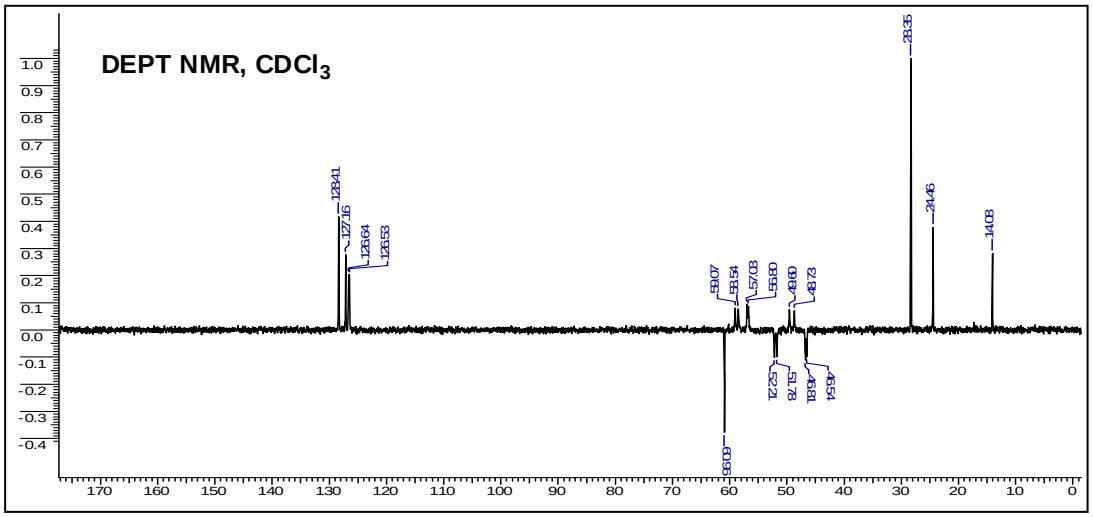

Ph
HCl.HN
CO2Et
N
Boc
16
1H NMR, DMSO-d6
DMSO-d6
7.62
7.61
7.43
7.41
4.03
4.01
3.99
3.63
3.45
3.35
4.53
4.05
2.50
1.63
1.34
1.12
1.10
1.08
10.01
1.44
2.95
0.97
3.33
3.60
0.41
9.00

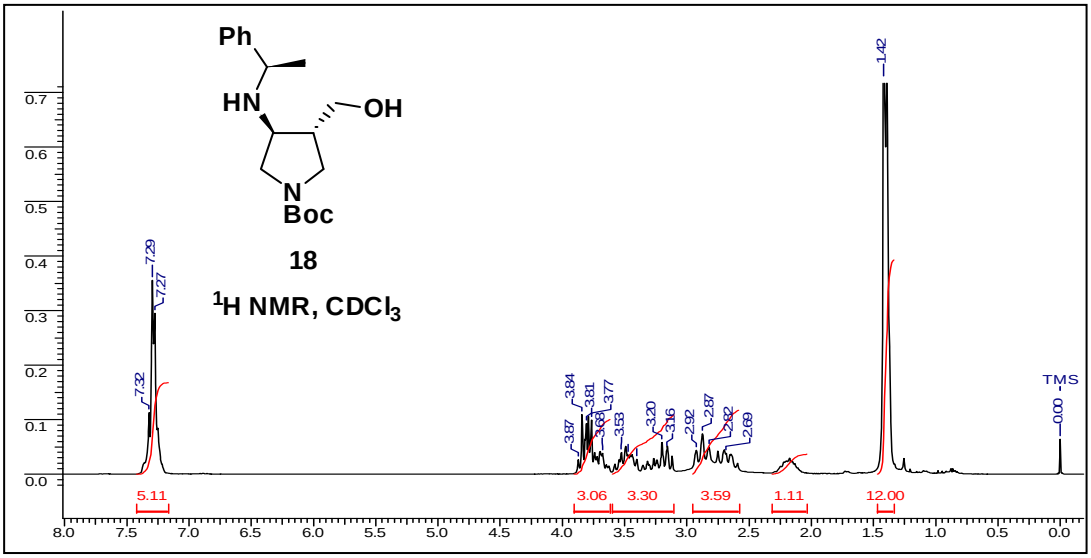

13C NMR, DMSO-d6
DMSO-d6
39.70
28.43
170.87
153.41
133.95
129.56
129.43
128.55
79.79
61.79
57.03
56.10
48.41
47.14
45.45
20.41
14.22

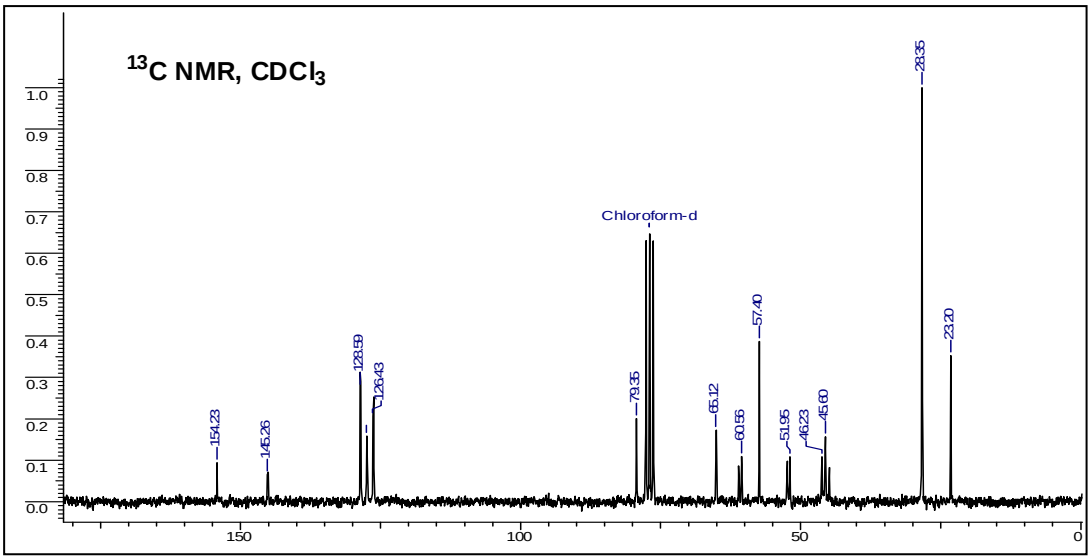

DEPT NMR, DMSO-d6
129.43
128.55
129.56
28.44
57.03
56.10
45.44
44.63
20.42
14.22
61.80

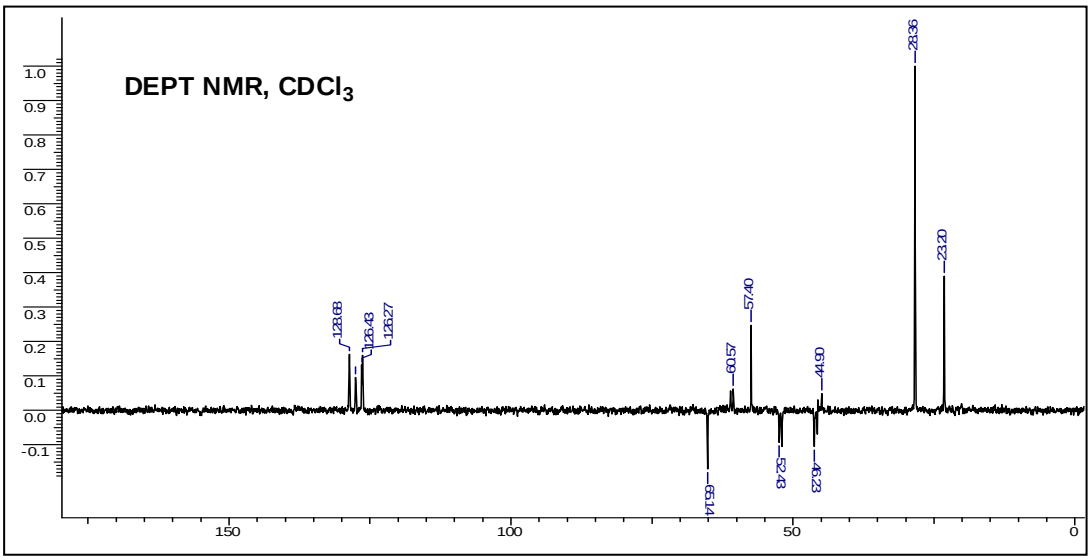

Ph
HN
CO_2Et
N
Boc
17
^{1}H NMR, $CDCl_3$
7.30
1.41
1.33
TMS
0.00
5.68
2.00
1.03
4.05
2.02
15.19

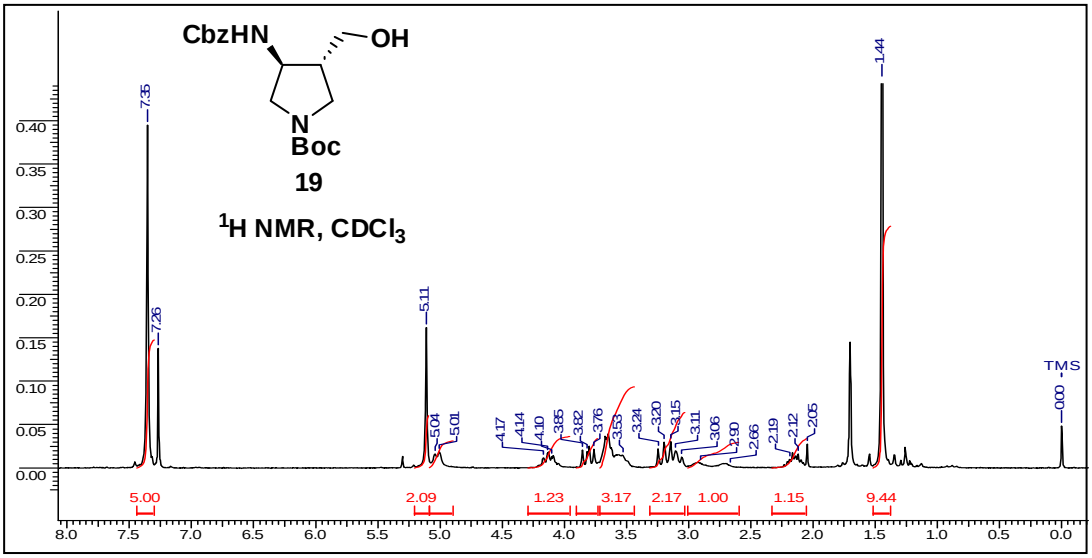

^{13}C NMR, $CDCl_3$
Chloroform-d
77.63
77.00
76.36
172.43
154.09
145.18
128.40
127.16
126.63
126.55
79.40
60.95
28.35
24.46
14.08

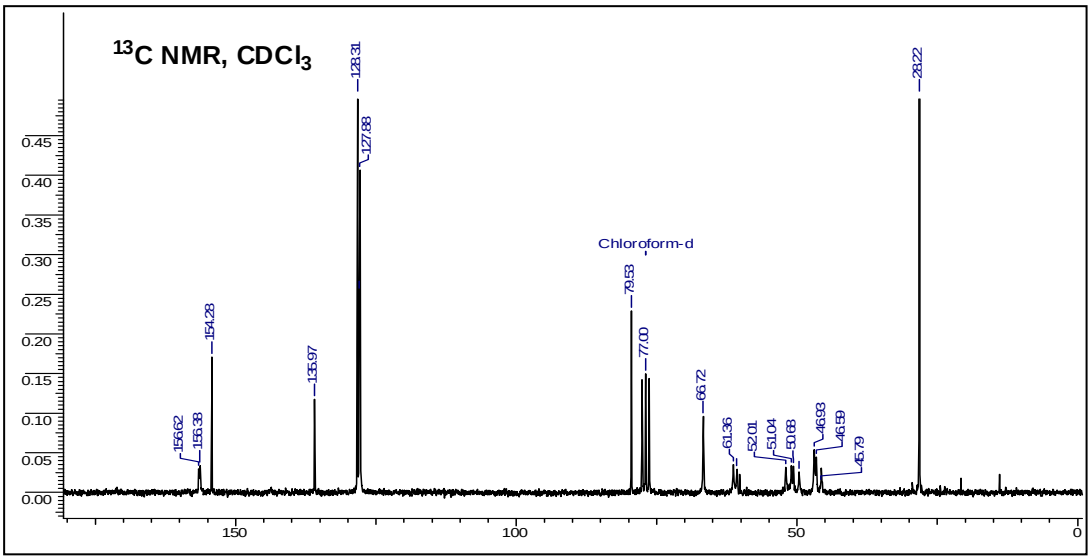

DEPT NMR, $CDCl_3$
28.35
24.46
14.08
128.41
127.16
126.64
126.53
66.09

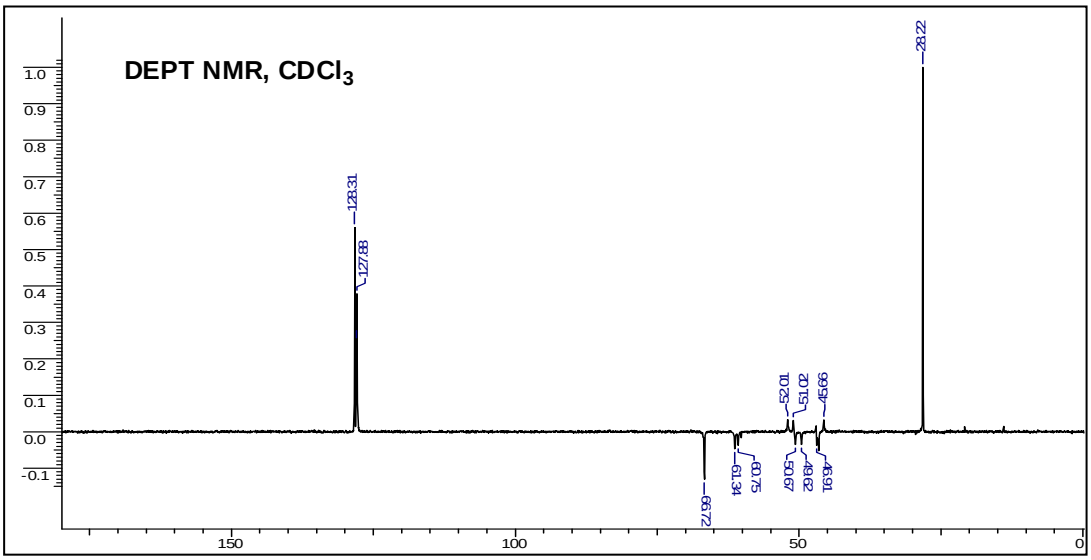

Ph
HN
OH
N
Boc
18
1H NMR, CDCl3
7.29
7.27
7.32
3.84
3.81
3.77
3.67
3.63
3.53
3.20
3.16
2.92
2.87
2.82
2.69
1.42
0.00
TMS
5.11
3.06
3.30
3.59
1.11
12.00

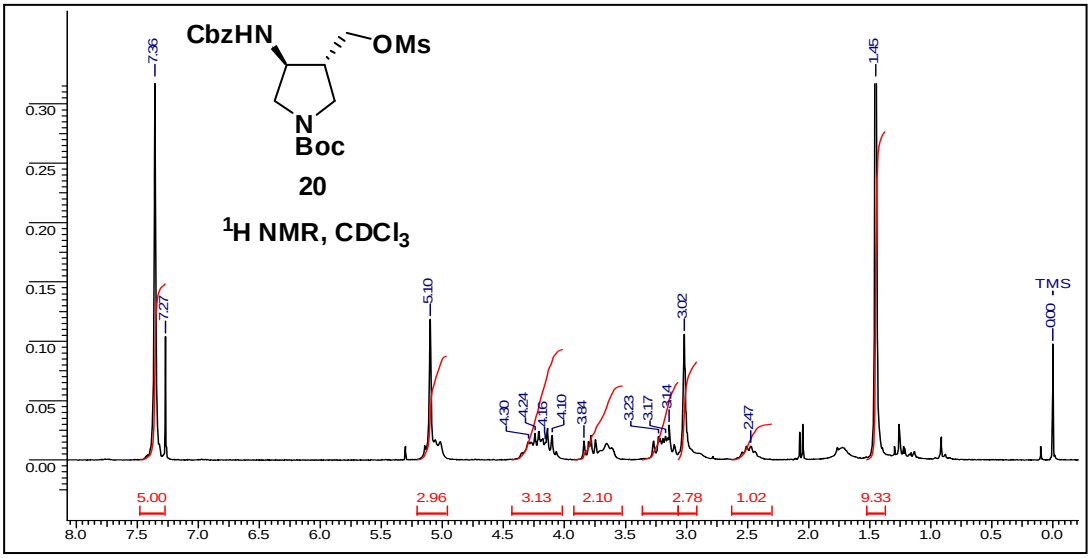

13C NMR, CDCl3
Chloroform-d
154.23
145.26
128.59
126.43
79.35
66.12
60.55
57.40
51.95
46.23
45.60
28.35
23.20

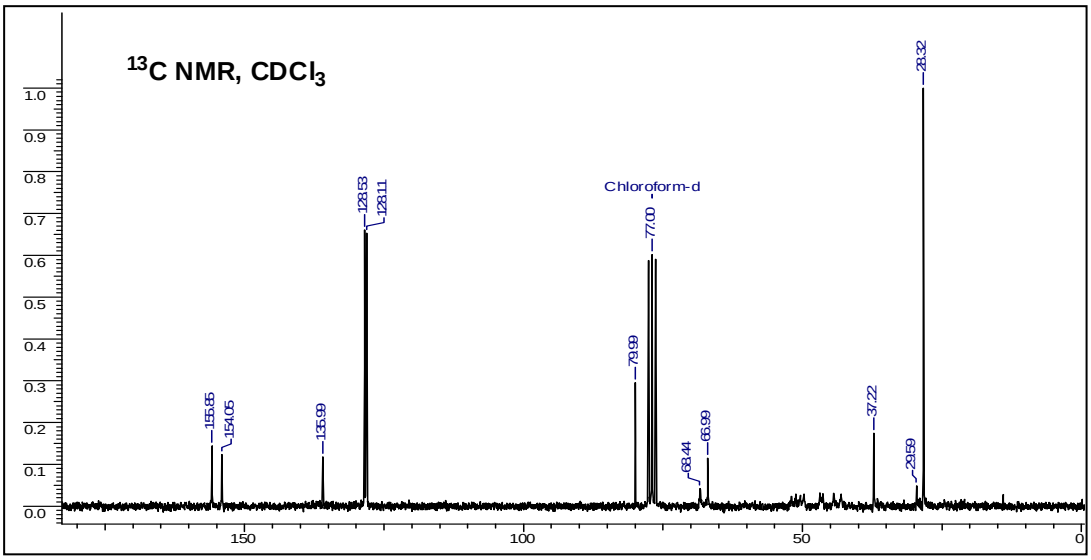

DEPT NMR, CDCl3
128.68
126.43
126.27
66.14
60.57
57.40
52.49
46.23
44.90
28.36
23.20

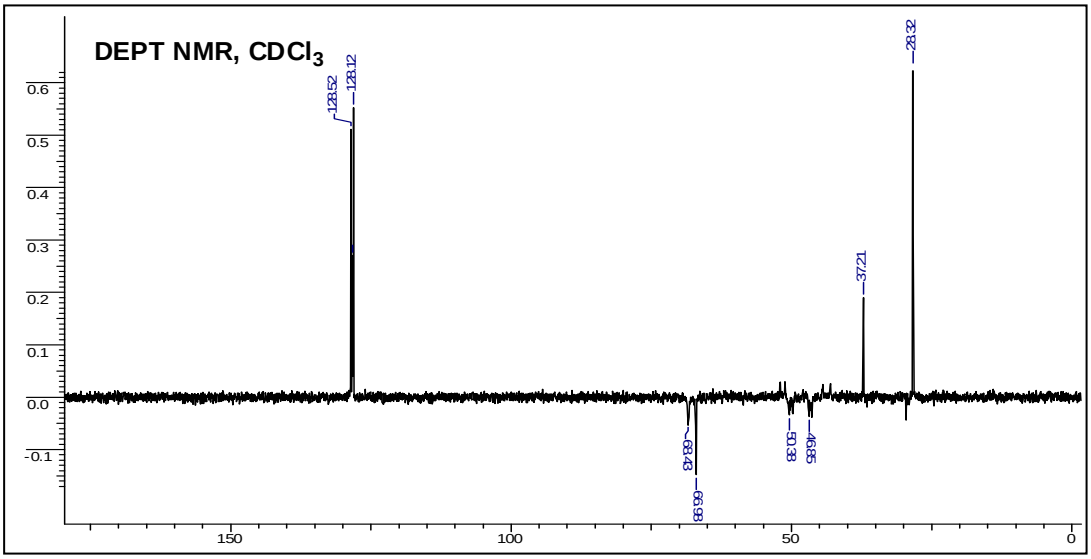

CbzHN
OH
N
Boc
19
1H NMR, CDCl3
TMS

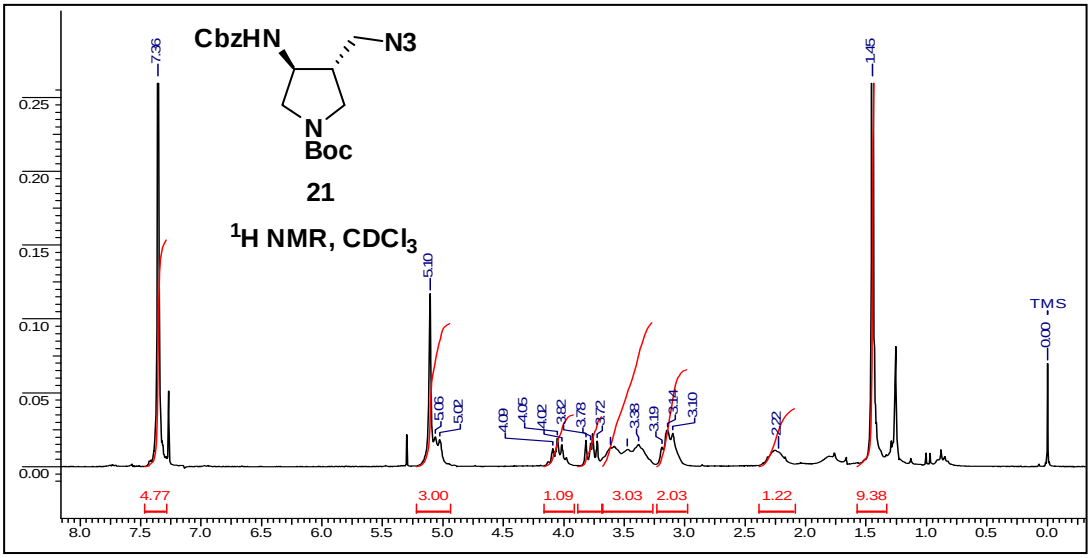

13C NMR, CDCl3
Chloroform-d

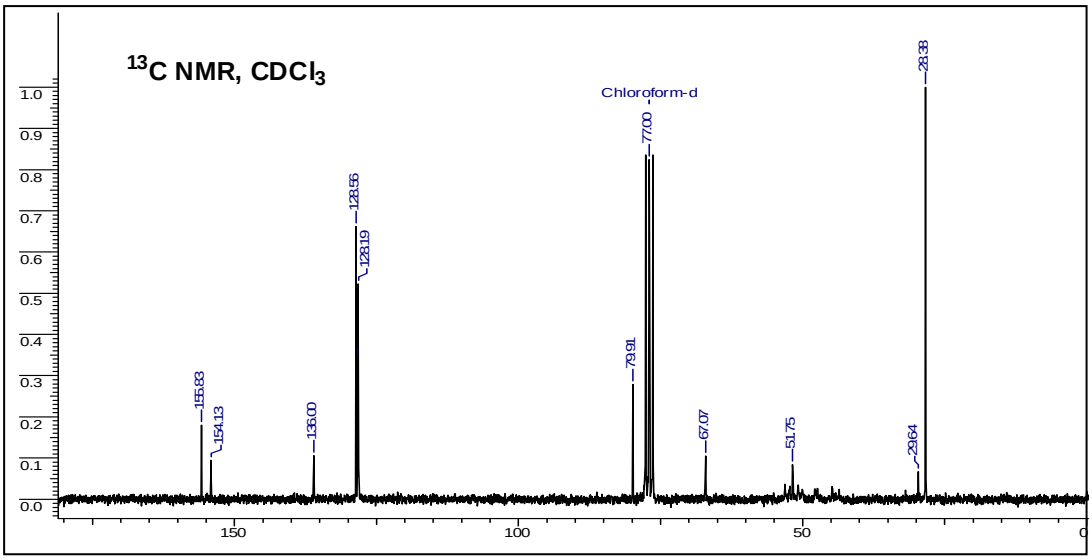

DEPT NMR, CDCl3

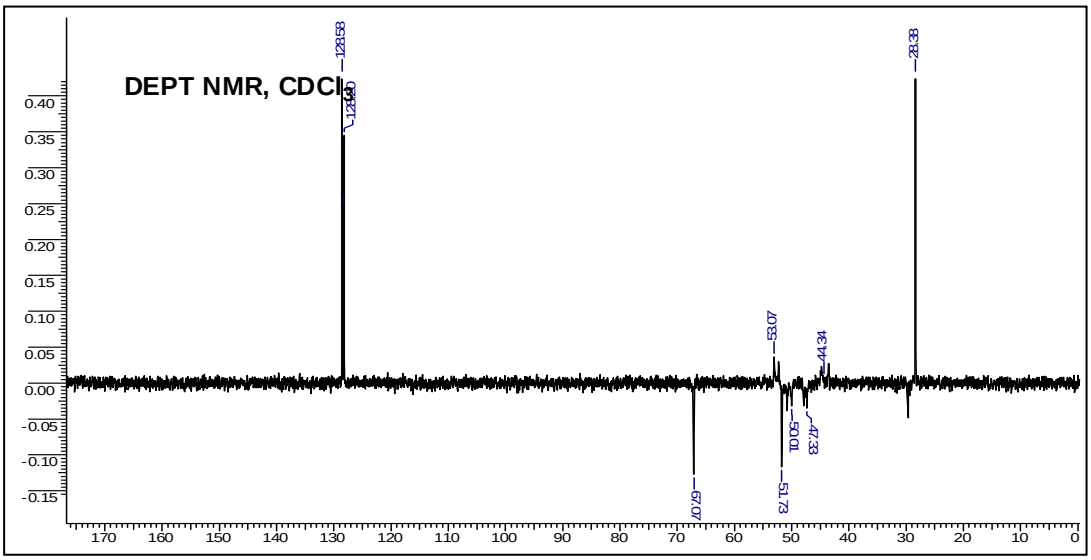

CbzHN
OMs
N
Boc
20
^{1}H NMR, CDCl$_3$
TMS

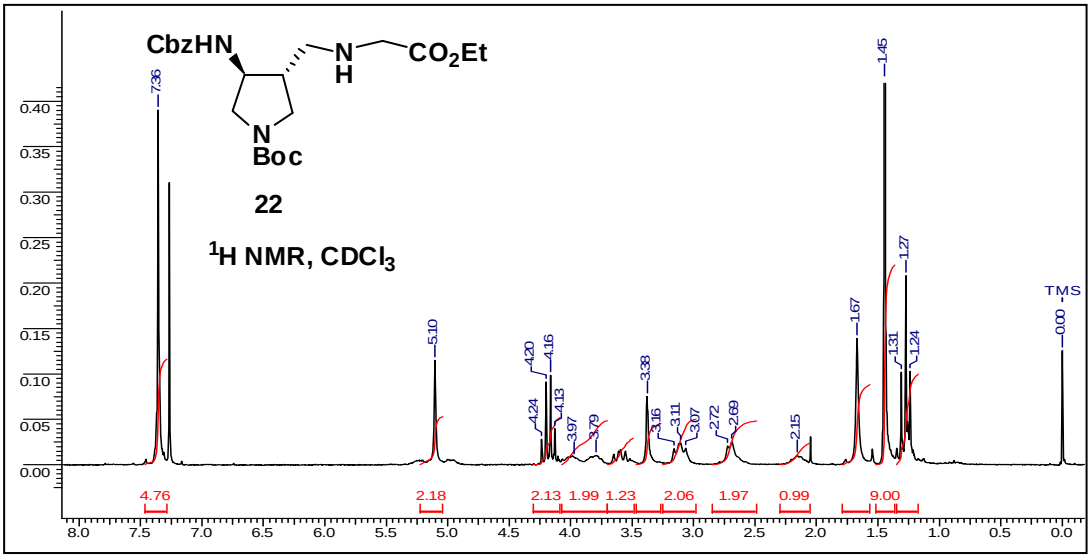

^{13}C NMR, CDCl$_3$
Chloroform-d

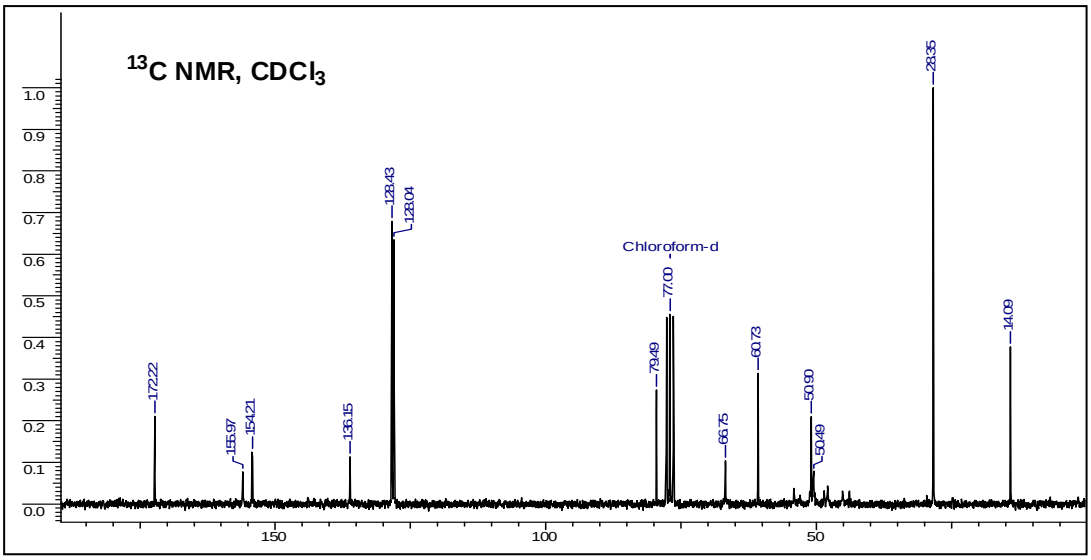

DEPT NMR, CDCl$_3$

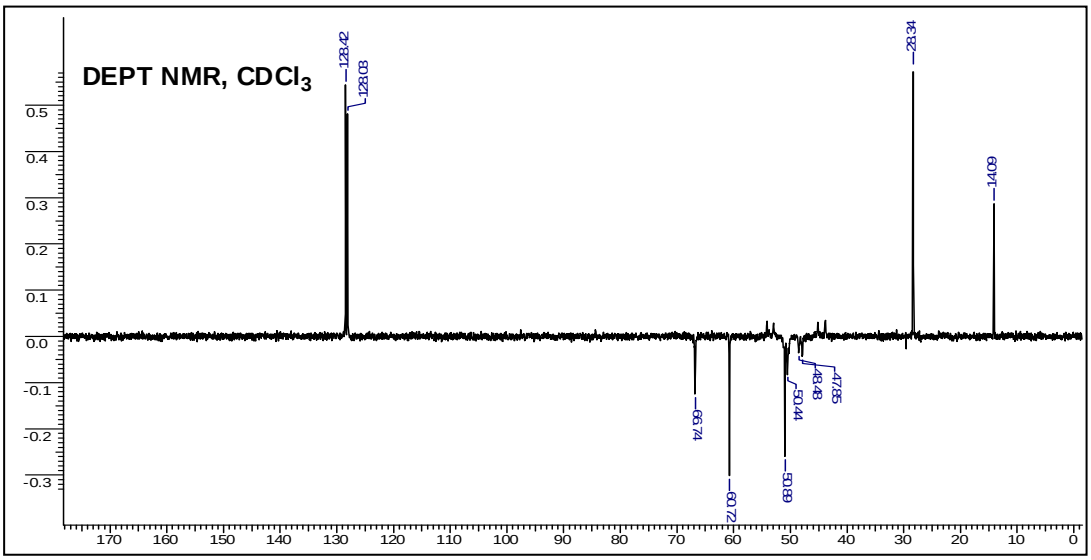

CbzHN
N3
N
Boc
21
^{1}H NMR, CDCl$_3$
TMS

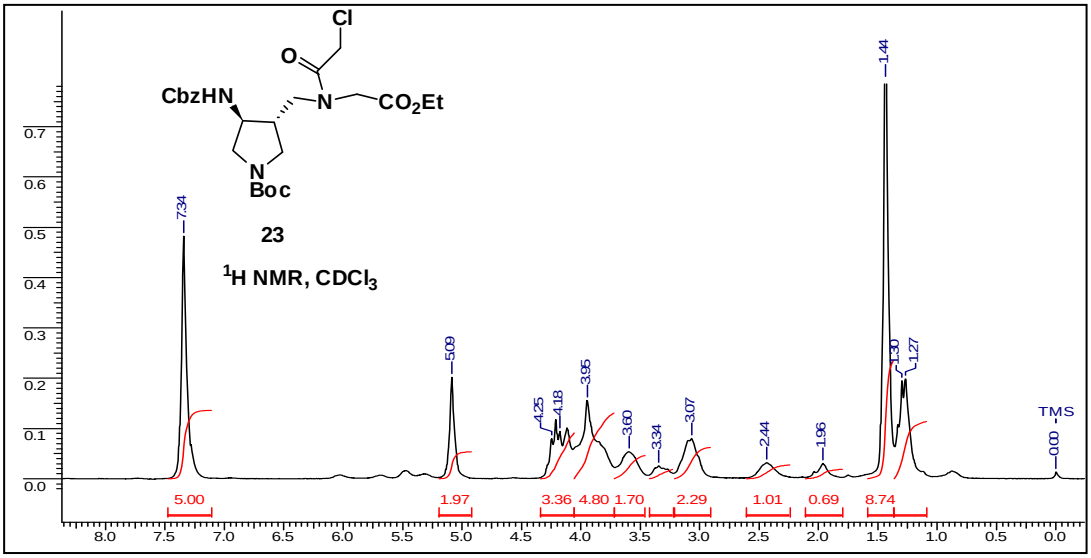

^{13}C NMR, CDCl$_3$
Chloroform-d

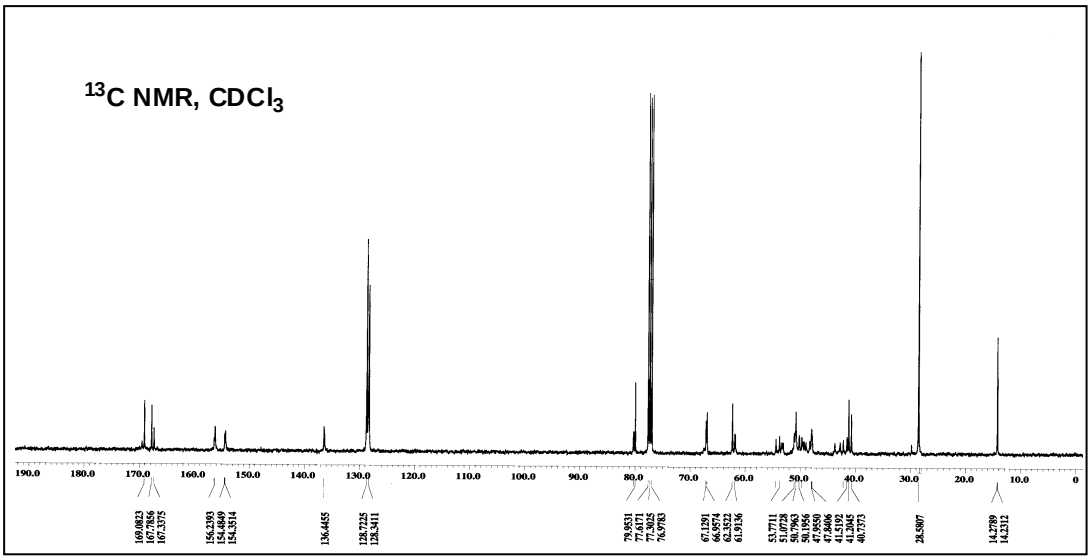

DEPT NMR, CDCl

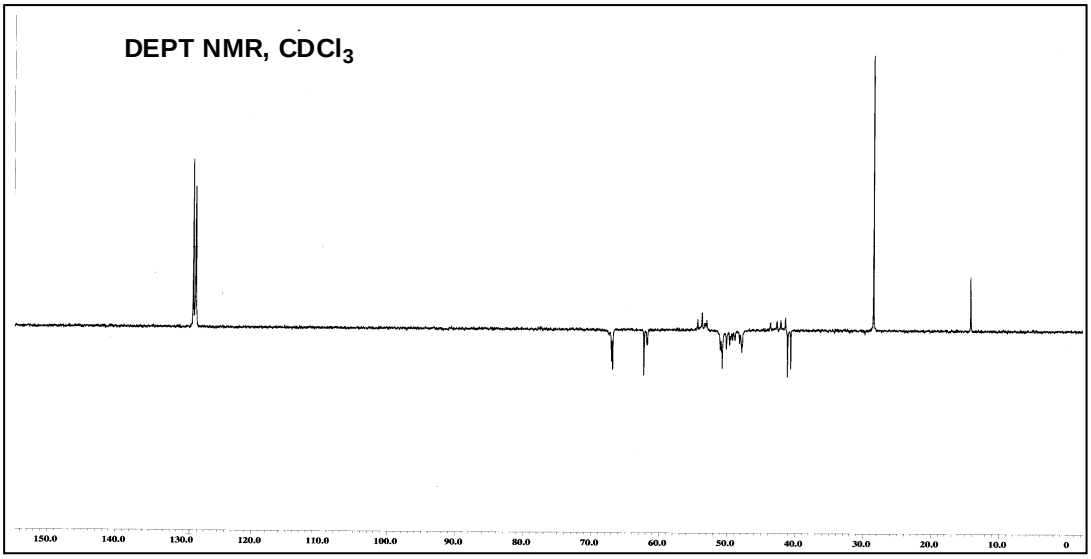

CbzHN
N
H
CO2Et
N
Boc
22
1H NMR, CDCl3
TMS

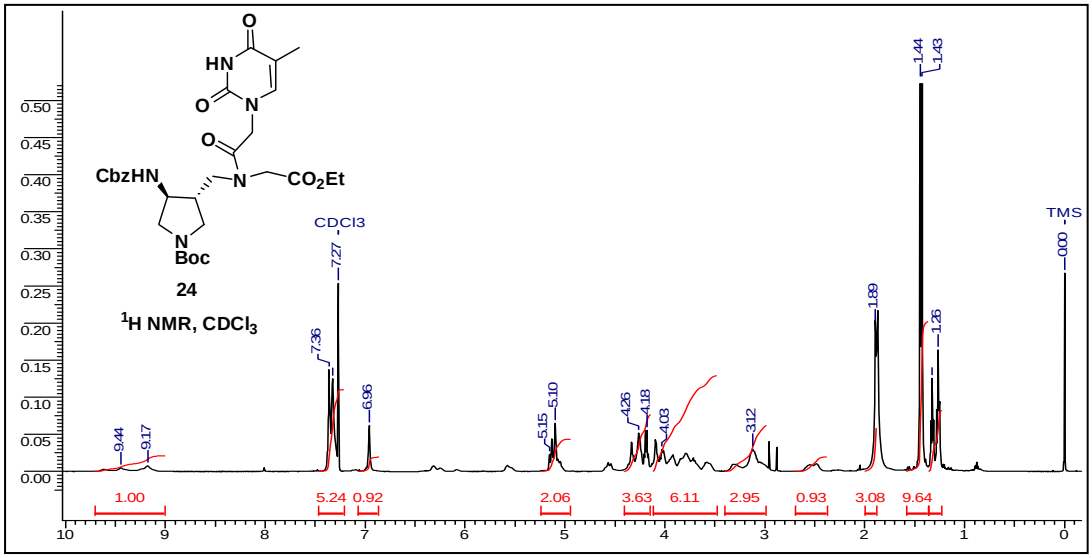

13C NMR, CDCl3
Chloroform-d

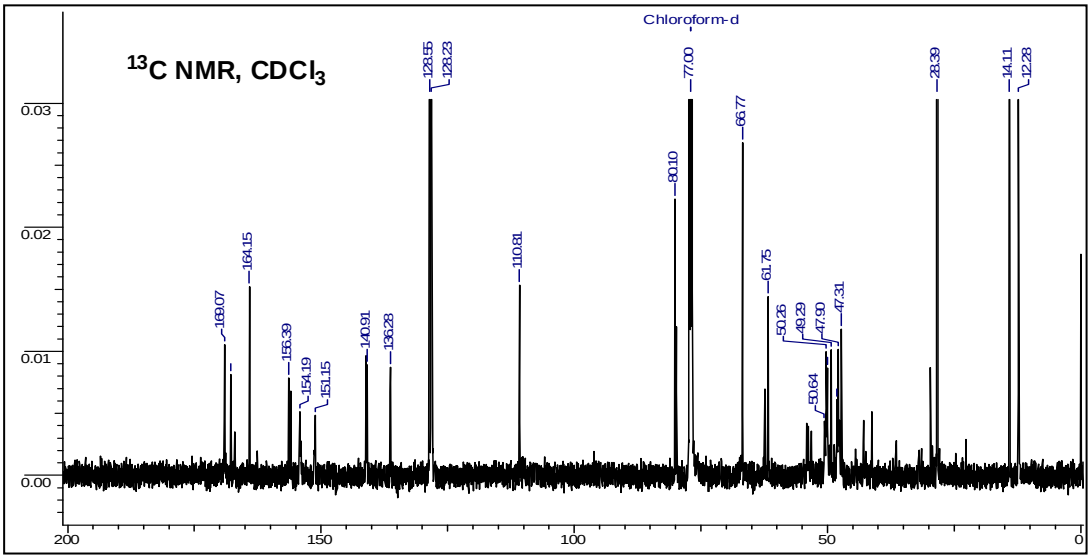

DEPT NMR, CDCl3

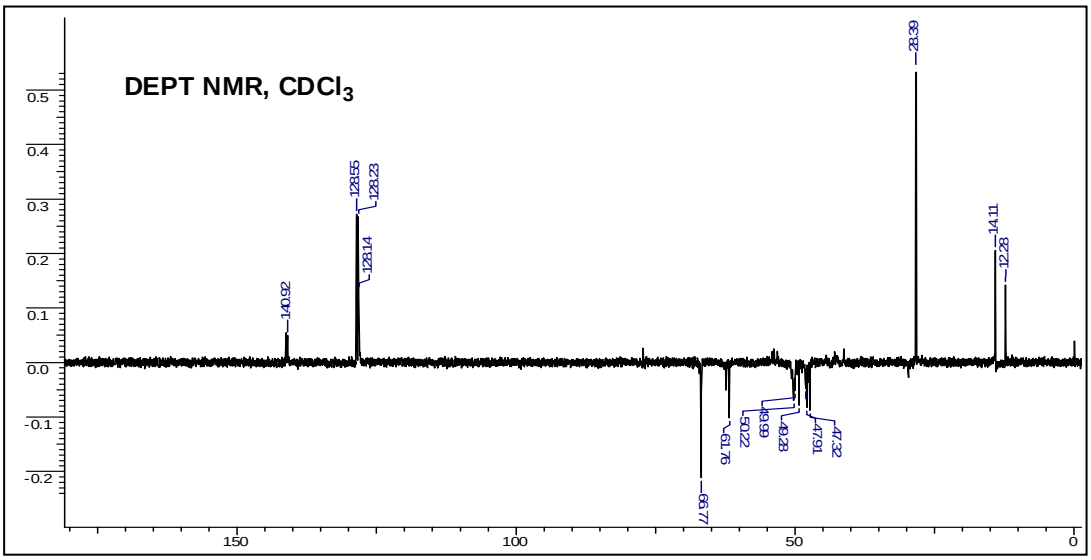

Cl
O
CbzHN
N
CO2Et
N
Boc
23
1H NMR, CDCl3
7.34
5.09
4.25
4.18
3.95
3.60
3.34
3.07
2.44
1.96
1.44
1.30
1.27
0.00
TMS
5.00
1.97
3.36
4.80
1.70
2.29
1.01
0.69
8.74


13C NMR, CDCl3

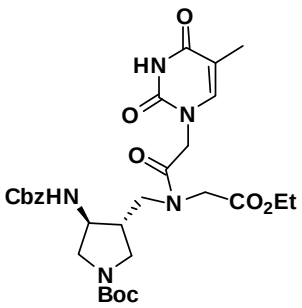

DEPT NMR, CDCl3


HN
O
N
O
CbzHN
N
CO2Et
N
Boc
24
1H NMR, CDCl3
CDCl3
TMS
1.00
5.24
0.92
2.06
3.63
6.11
2.95
0.93
3.08
9.64

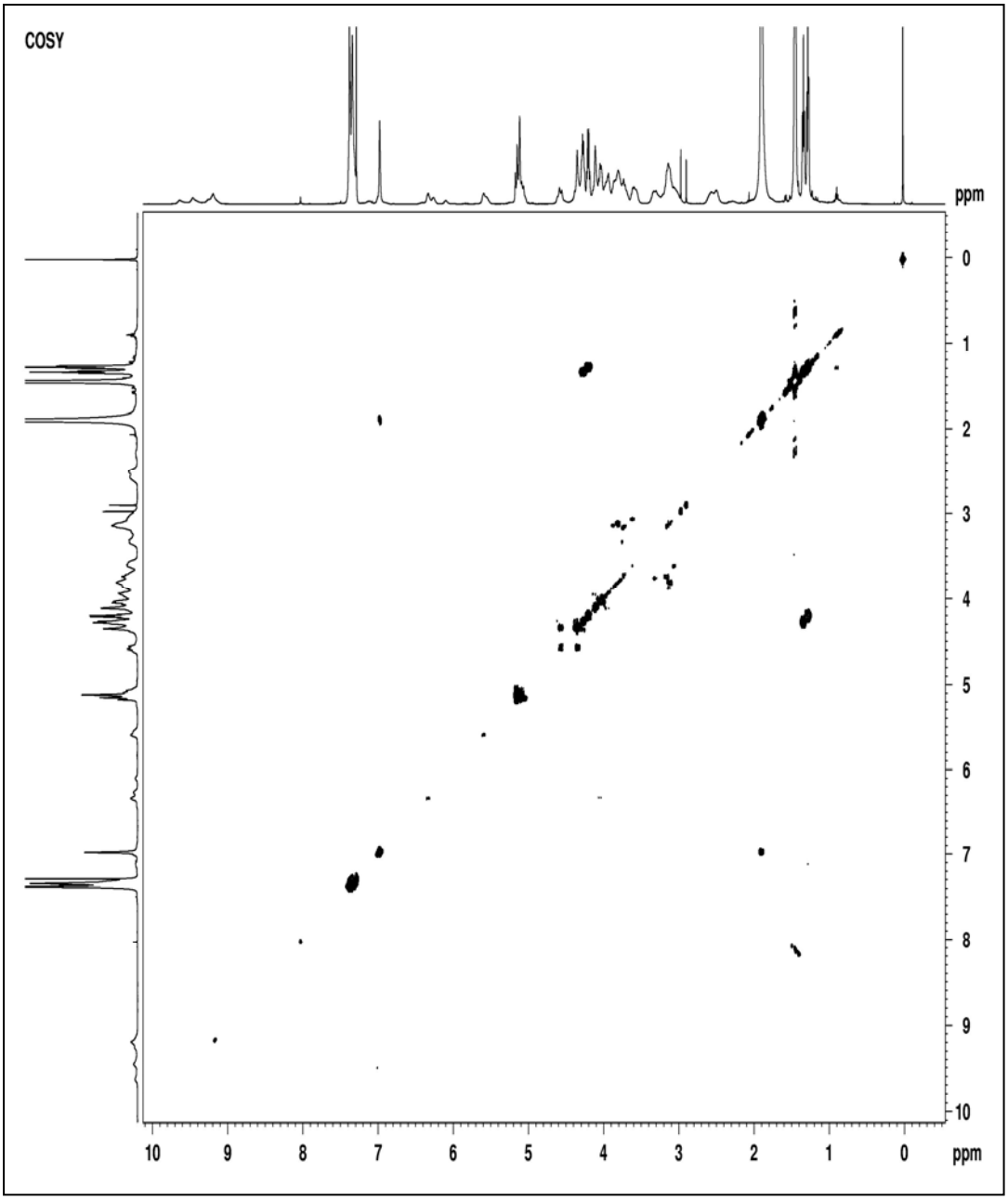

13C NMR, CDCl3
Chloroform-d


DEPT NMR, CDCl3

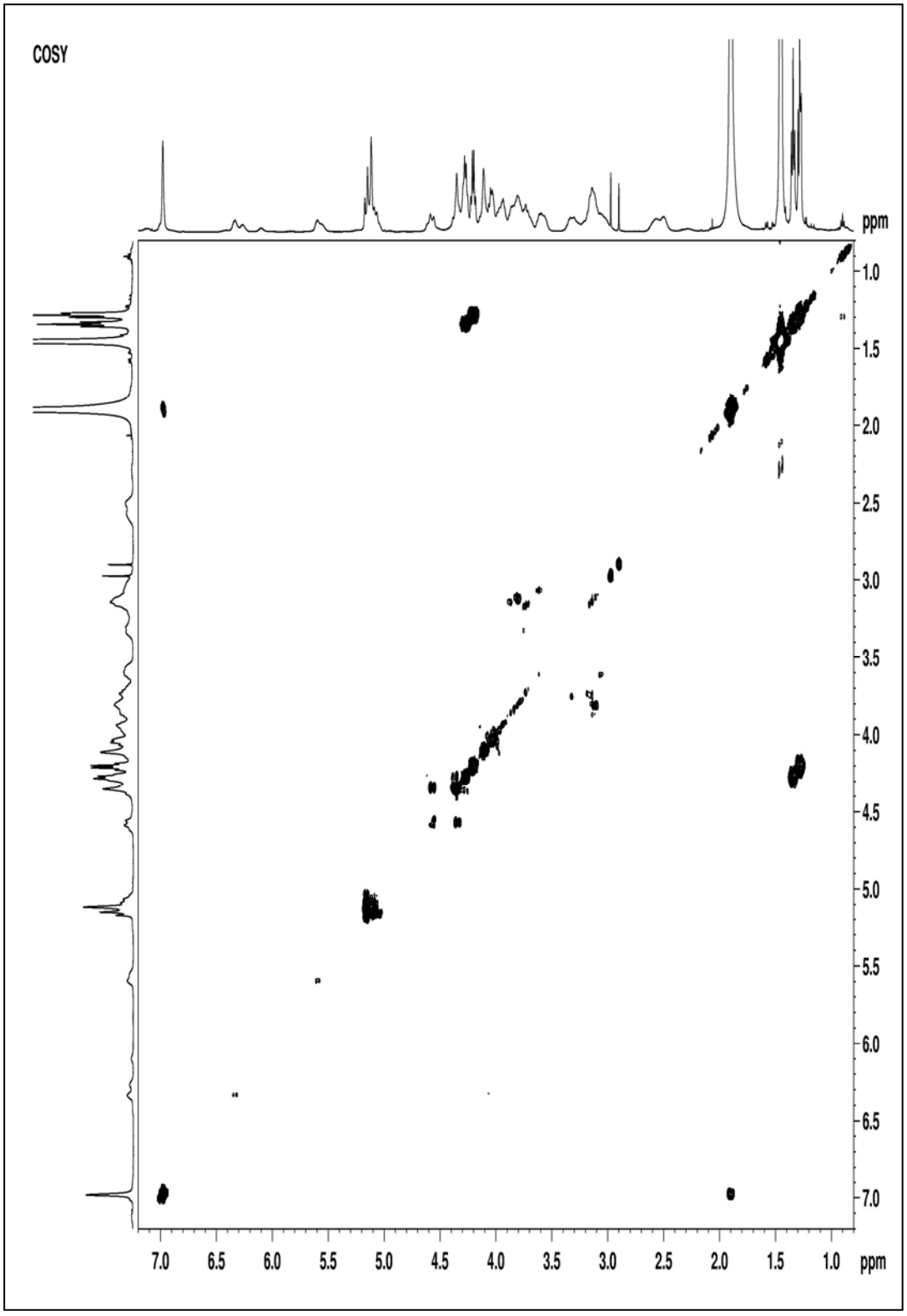
Cosy, $CDCl_3$

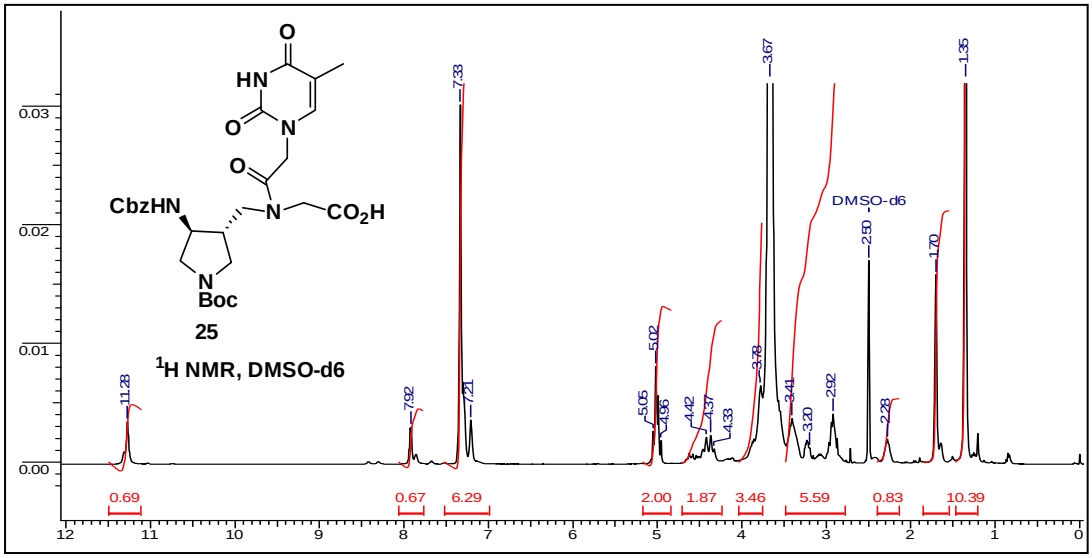


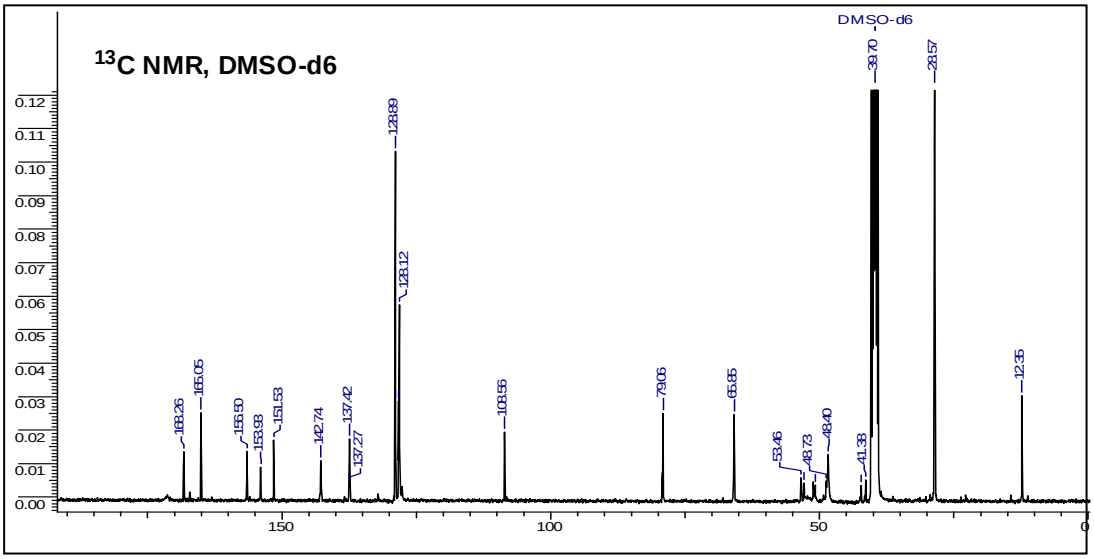
24

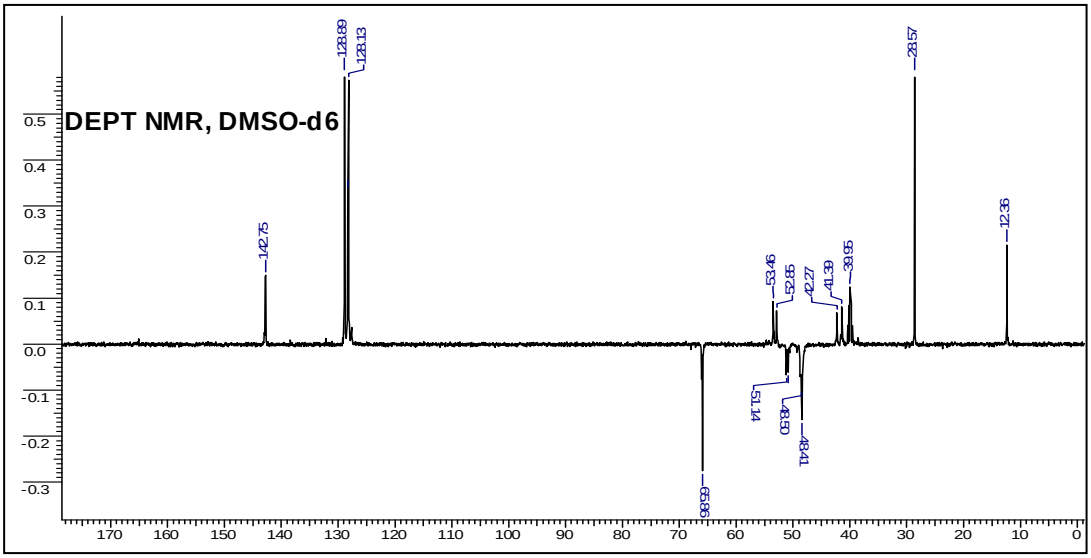


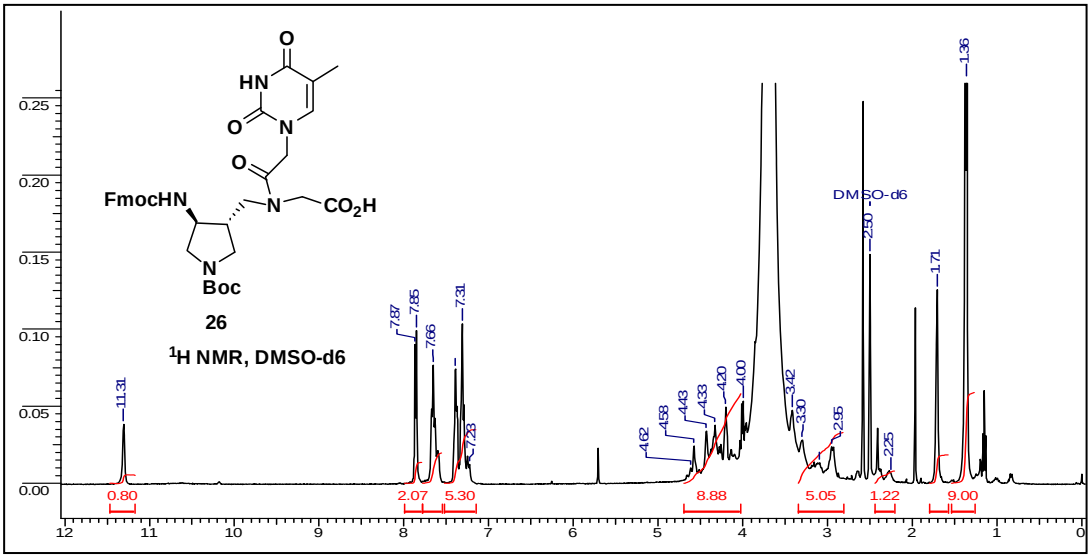
Zoomed region of compound 24, Cosy

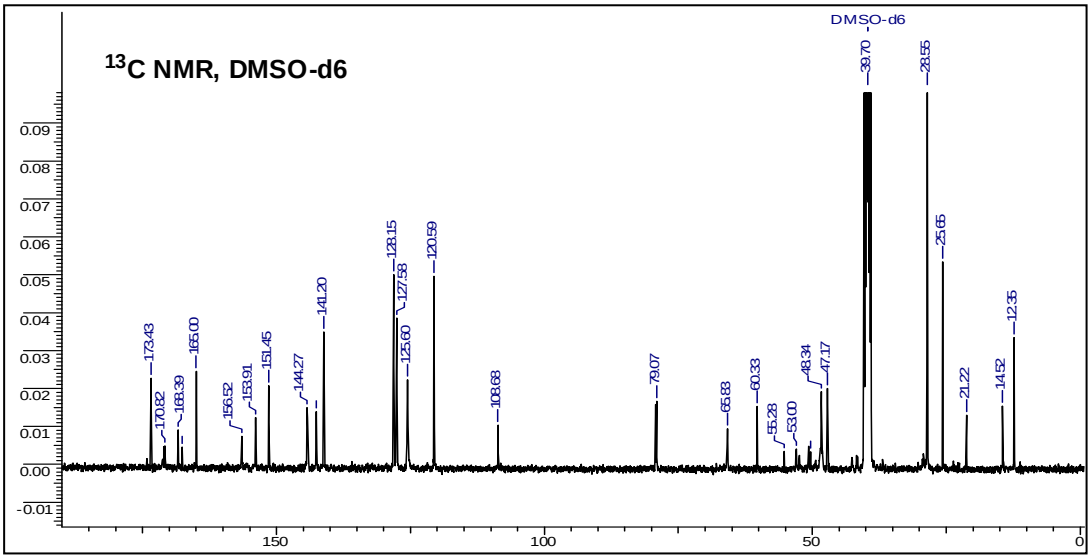


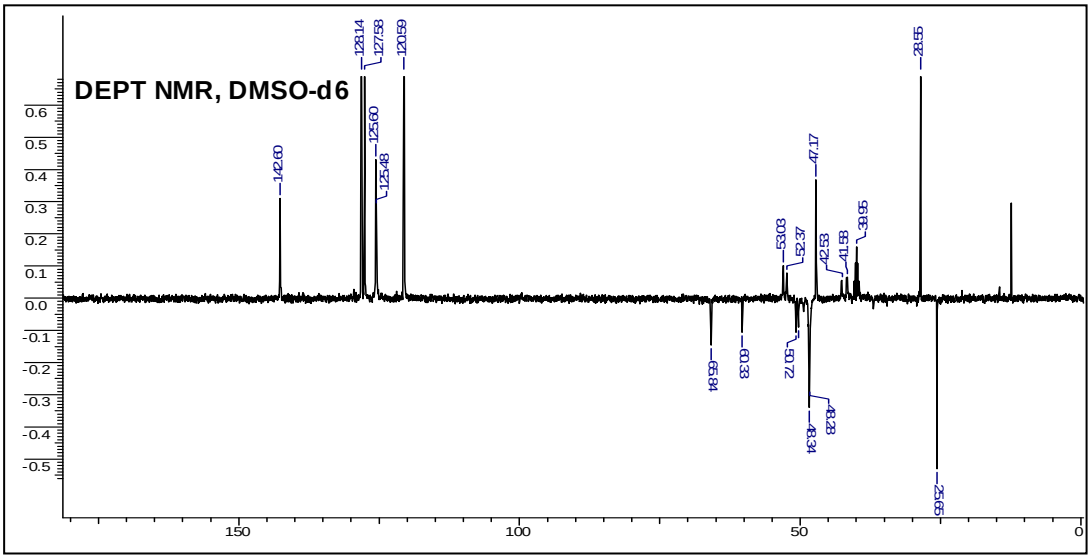

HN
O
N
O
O
CbzHN
N
CO2H
N
Boc
25
1H NMR, DMSO-d6
DMSO-d6

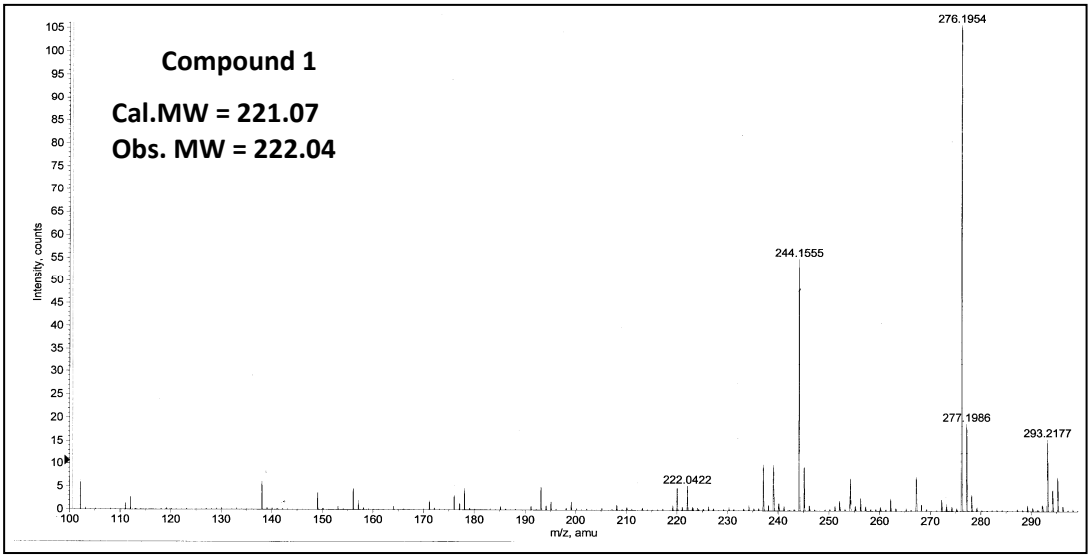

13C NMR, DMSO-d6
DMSO-d6

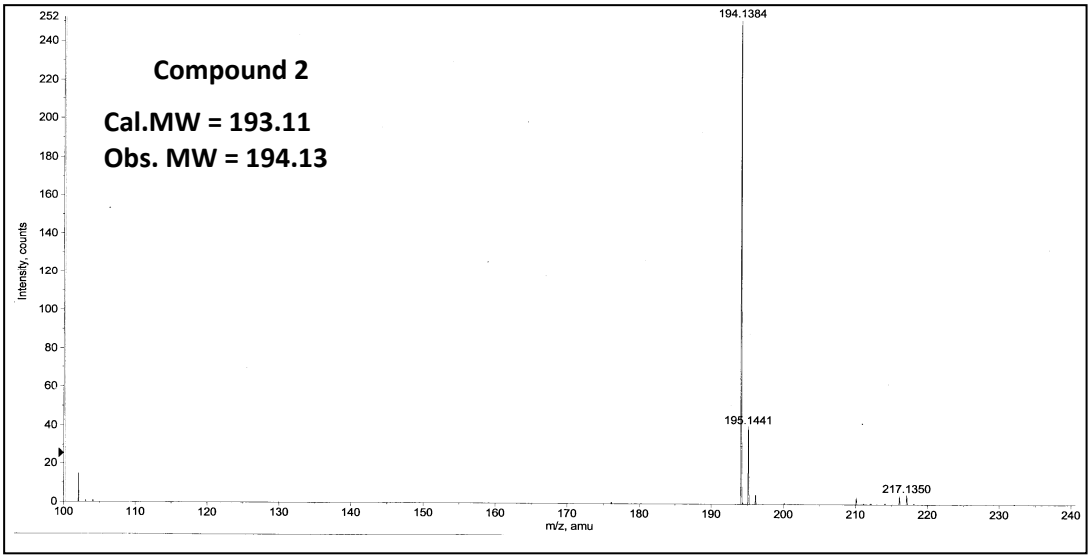

DEPT NMR, DMSO-d6

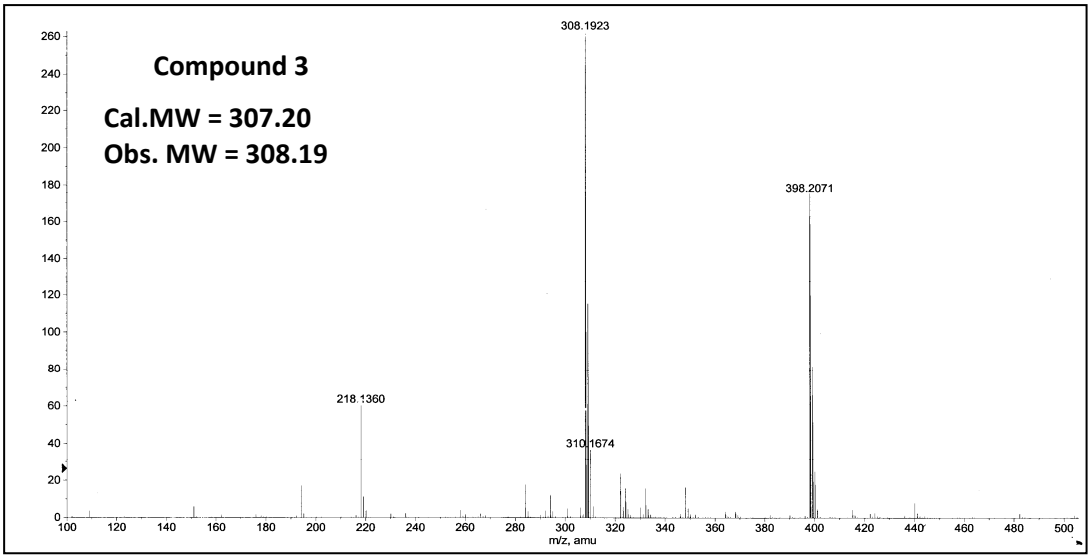

FmocHN
HN
O
N
CO2H
Boc
26
1H NMR, DMSO-d6
DMSO-d6

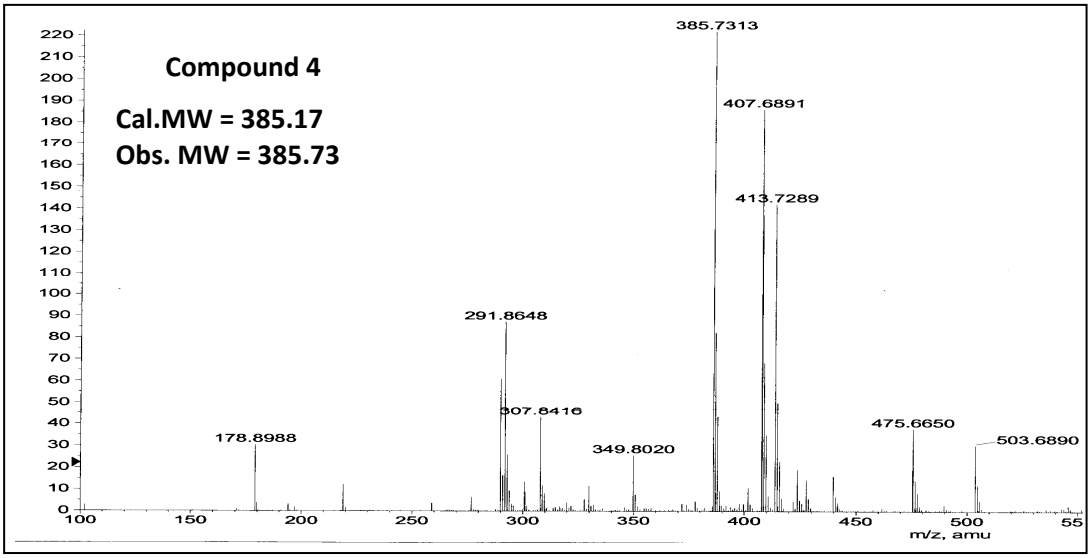

13C NMR, DMSO-d6
DMSO-d6

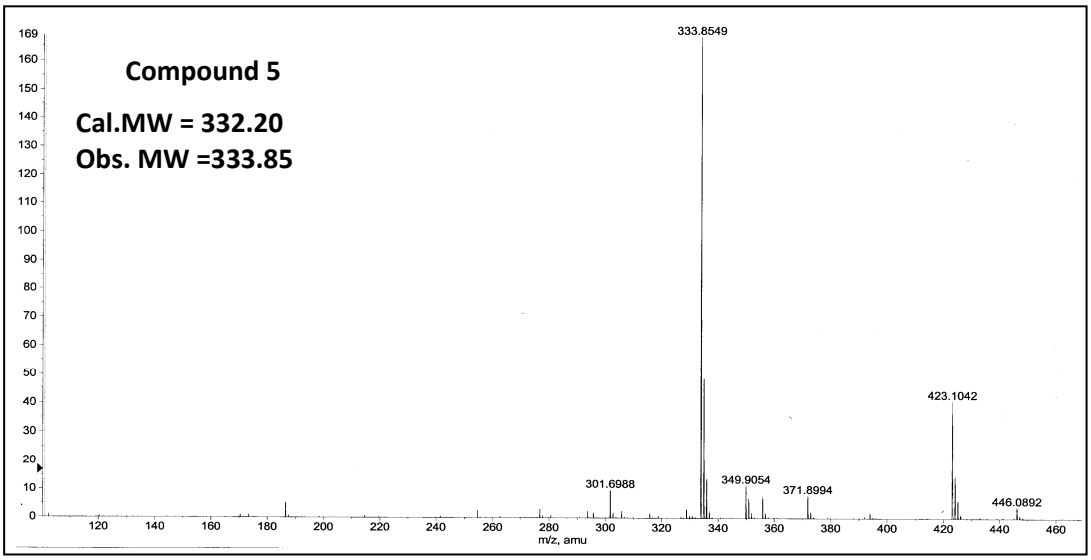

DEPT NMR, DMSO-d6

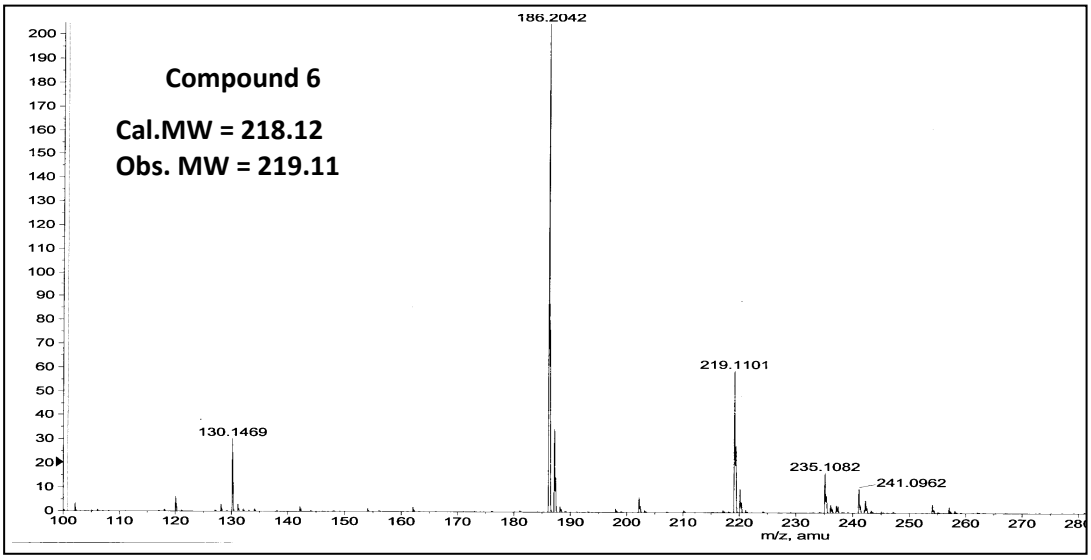

Compound 1
Cal.MW = 221.07
Obs. MW = 222.04
Intensity, counts
m/z, amu
276.1954
244.1555
277.1986
293.2177
222.0422

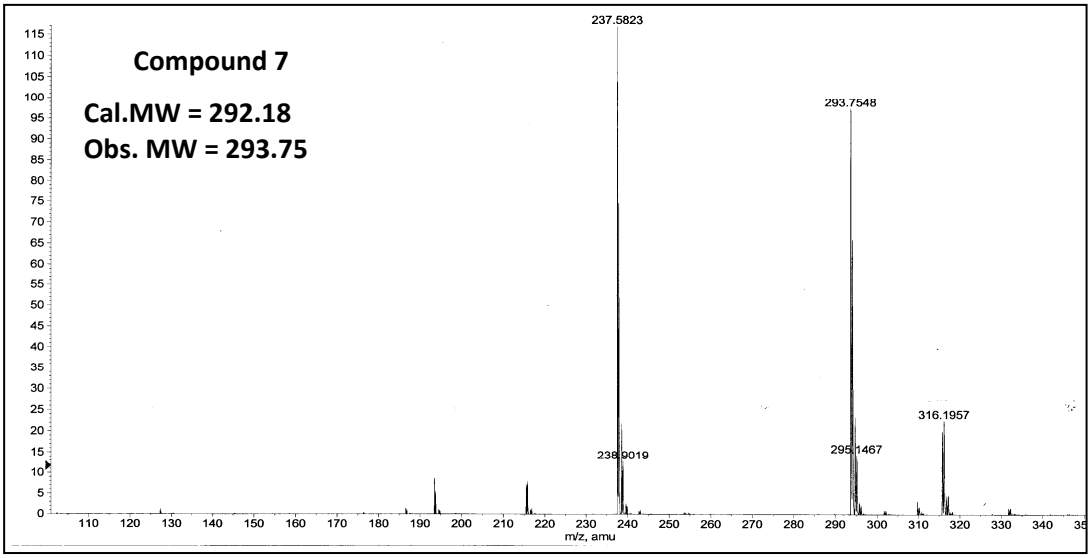

Compound 2
Cal.MW = 193.11
Obs. MW = 194.13
Intensity, counts
m/z, amu
194.1384
195.1441
217.1350

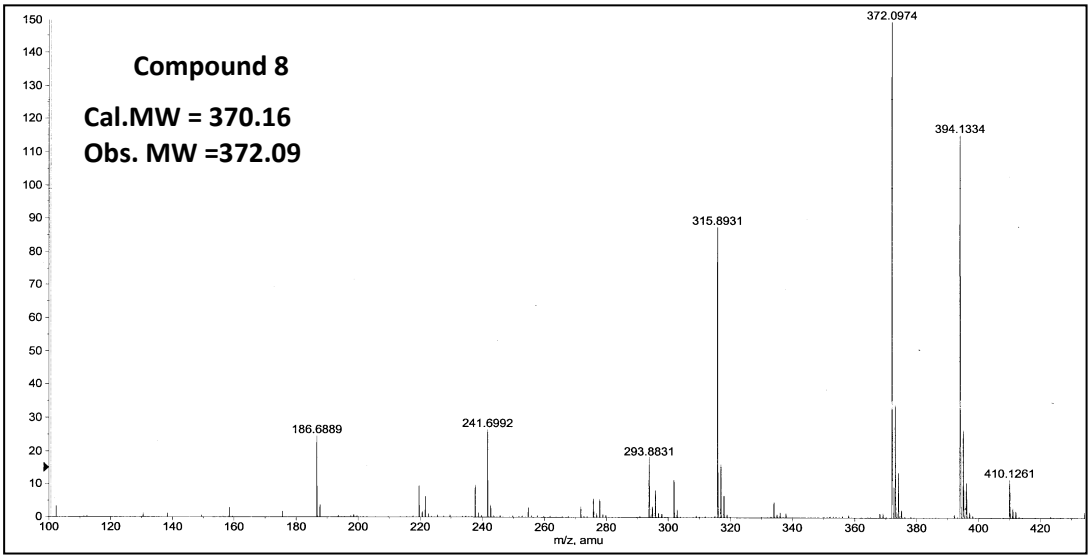

Compound 3
Cal.MW = 307.20
Obs. MW = 308.19
m/z, amu
308.1923
398.2071
218.1360
310.1674

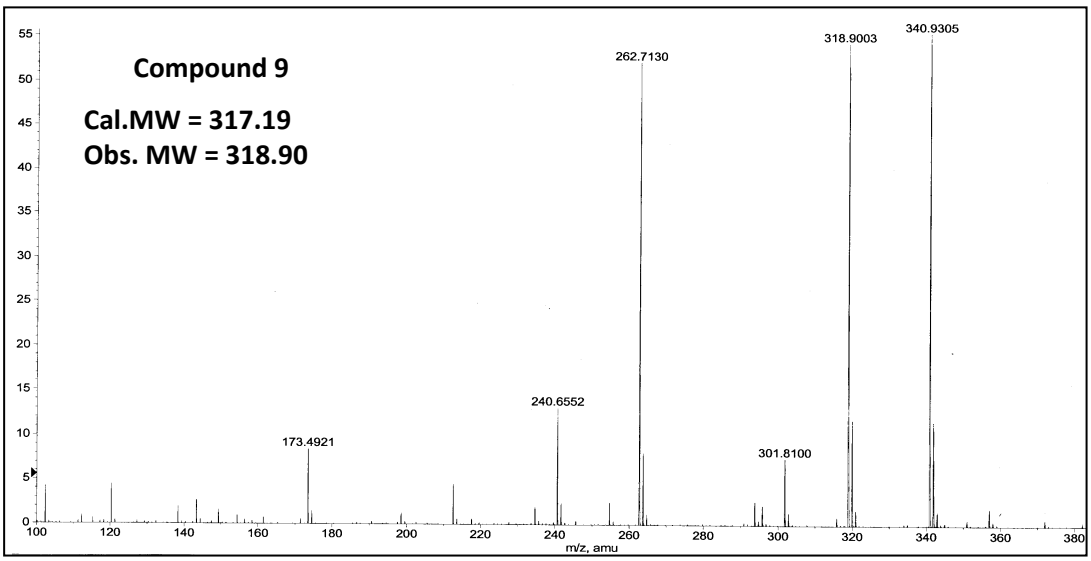

Compound 4
Cal.MW = 385.17
Obs. MW = 385.73
385.7313
407.6891
413.7289
291.8648
307.8416
178.8988
349.8020
475.6650
503.6890
m/z, amu

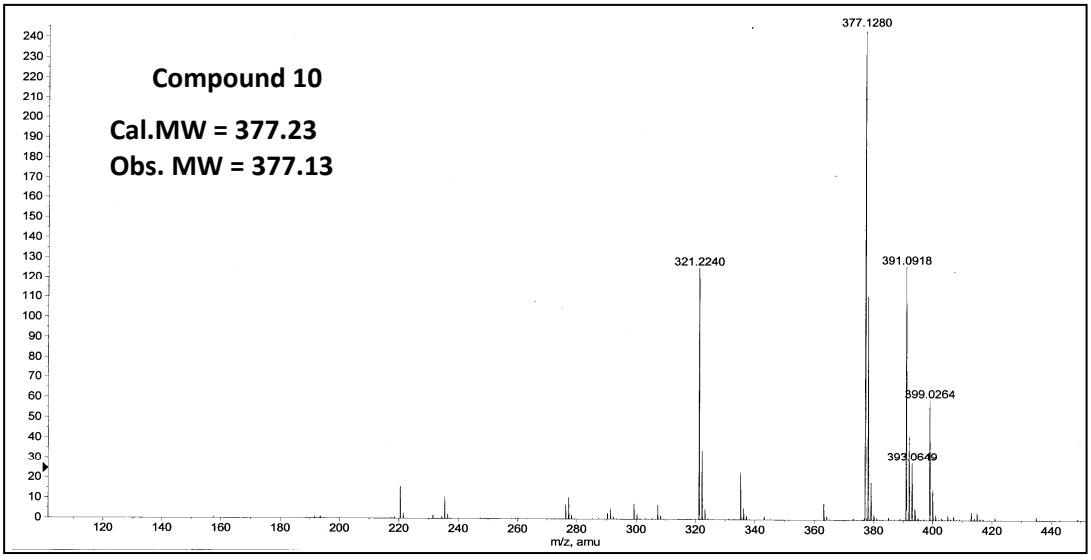

Compound 5
Cal.MW = 332.20
Obs. MW =333.85
333.8549
423.1042
301.6988
349.9054
371.8994
446.0892
m/z, amu

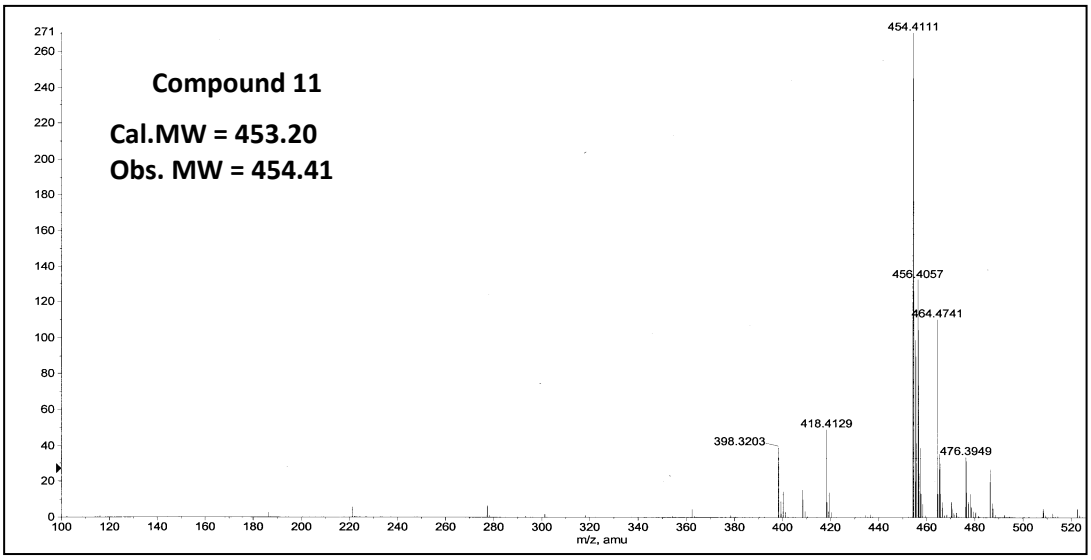

Compound 6
Cal.MW = 218.12
Obs. MW = 219.11
186.2042
219.1101
130.1469
235.1082
241.0962
m/z, amu

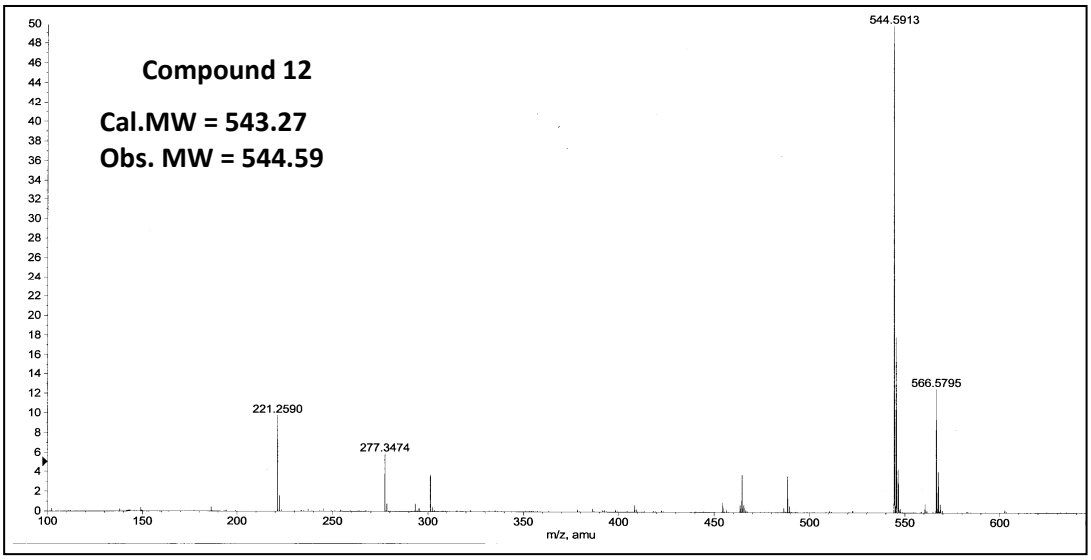

Compound 7
Cal.MW = 292.18
Obs. MW = 293.75
237.5823
293.7548
316.1957
238.9019
295.1467
m/z, amu

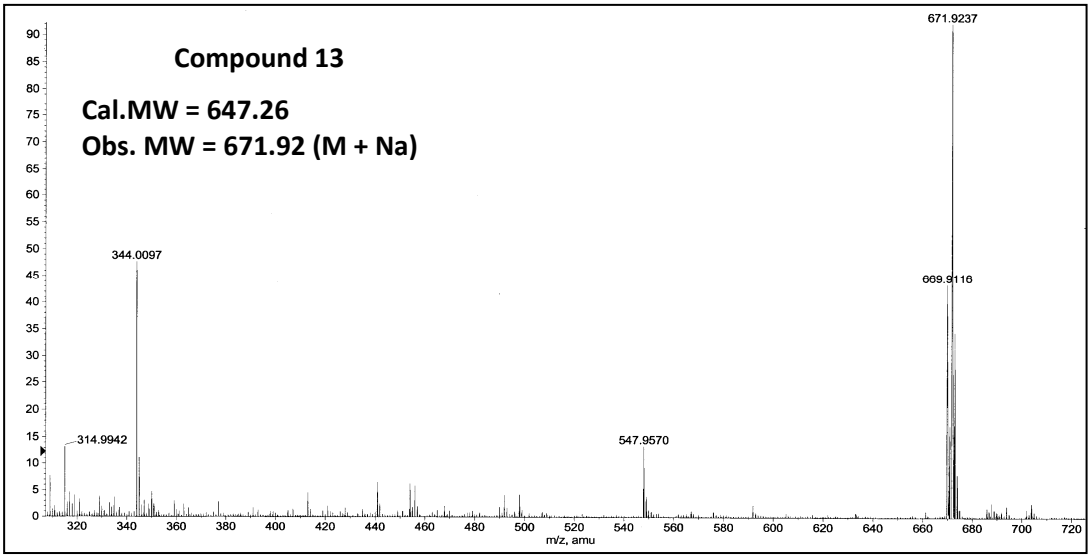

Compound 8
Cal.MW = 370.16
Obs. MW =372.09
372.0974
394.1334
315.8931
186.6889
241.6992
293.8831
410.1261
m/z, amu

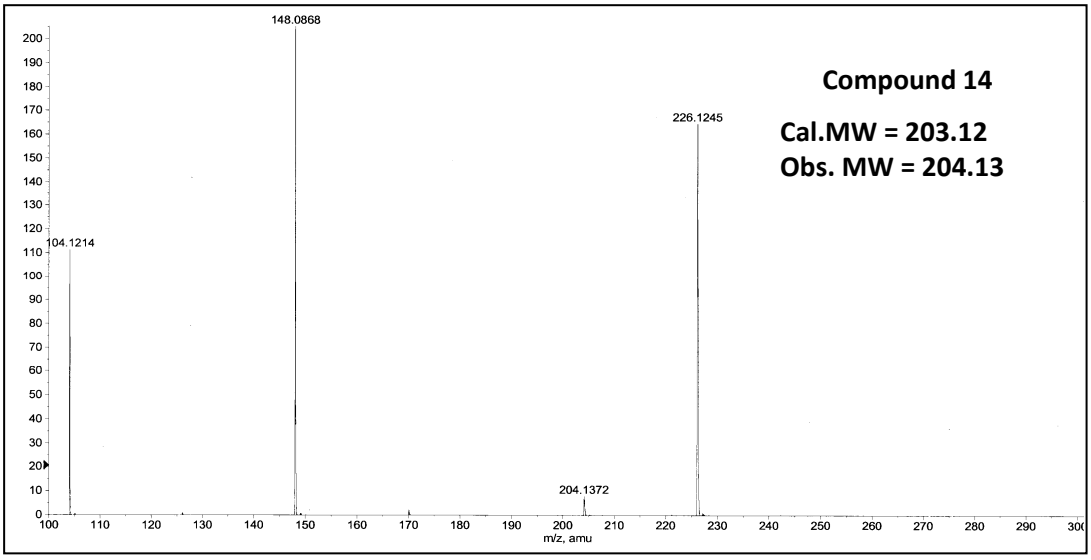

Compound 9
Cal.MW = 317.19
Obs. MW = 318.90
262.7130
318.9003
340.9305
240.6552
173.4921
301.8100
m/z, amu

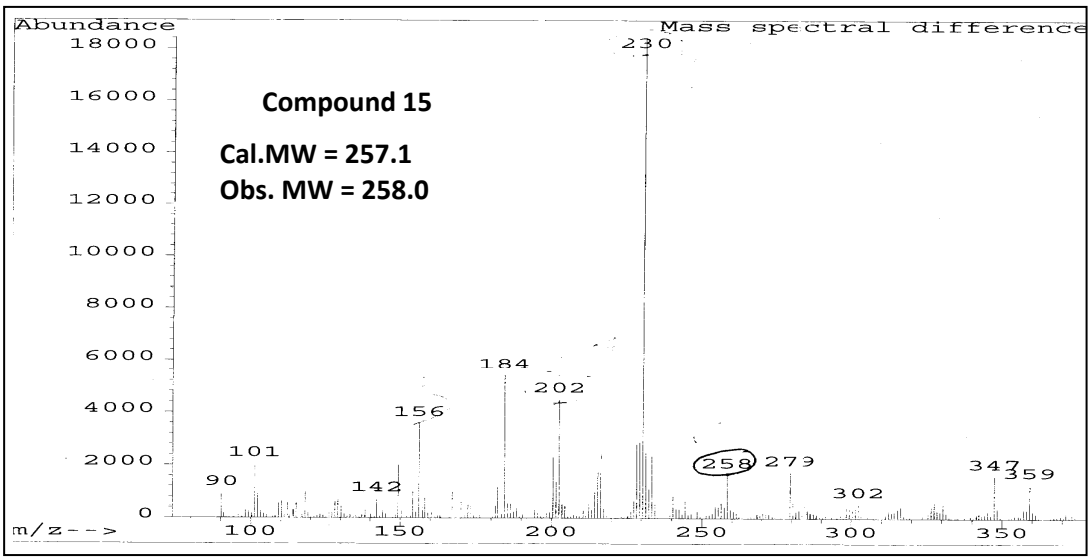

Compound 10
Cal.MW = 377.23
Obs. MW = 377.13
377.1280
321.2240
391.0918
399.0264
393.0649
m/z, amu

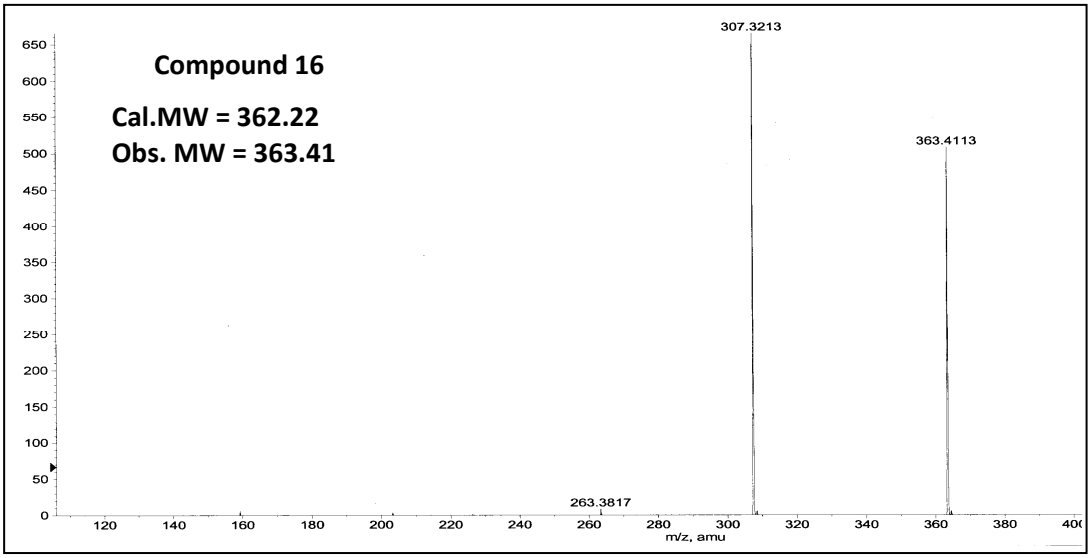

Compound 11
Cal.MW = 453.20
Obs. MW = 454.41
454.4111
456.4057
464.4741
418.4129
398.3203
476.3949
m/z, amu

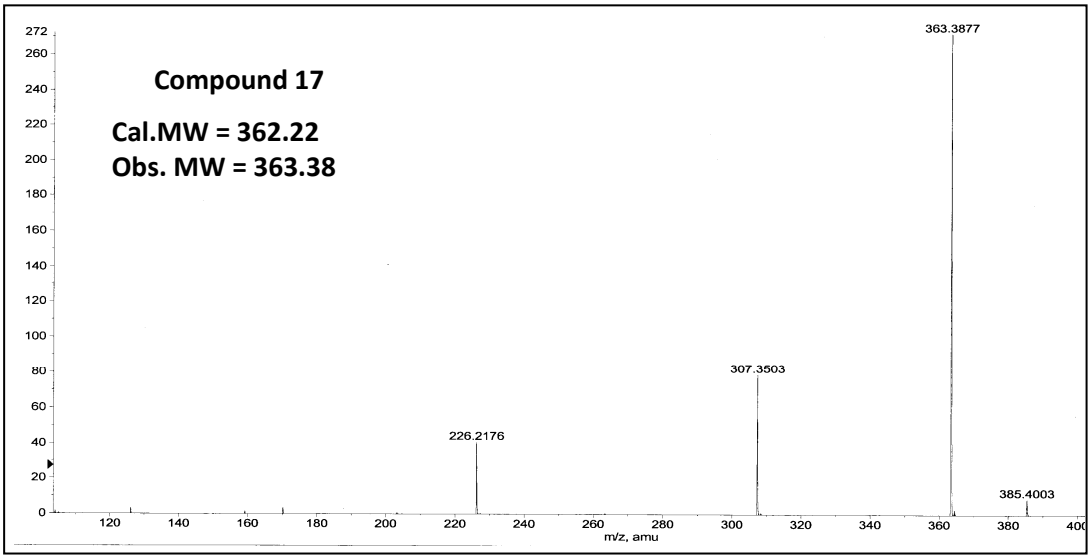

Compound 12
Cal.MW = 543.27
Obs. MW = 544.59
544.5913
566.5795
221.2590
277.3474
m/z, amu

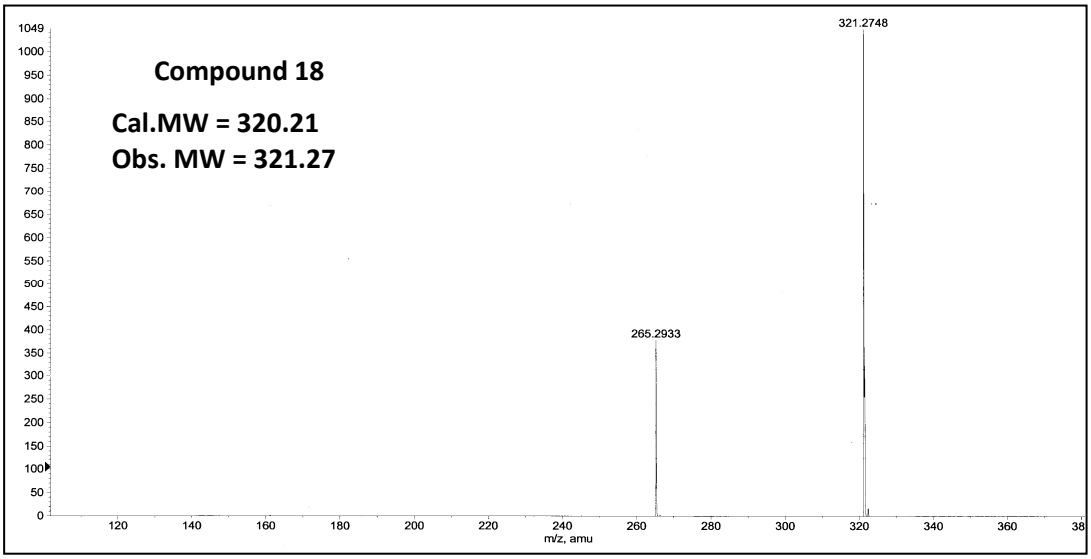

Compound 13
Cal.MW = 647.26
Obs. MW = 671.92 (M + Na)
671.9237
344.0097
669.9116
314.9942
547.9570
m/z, amu

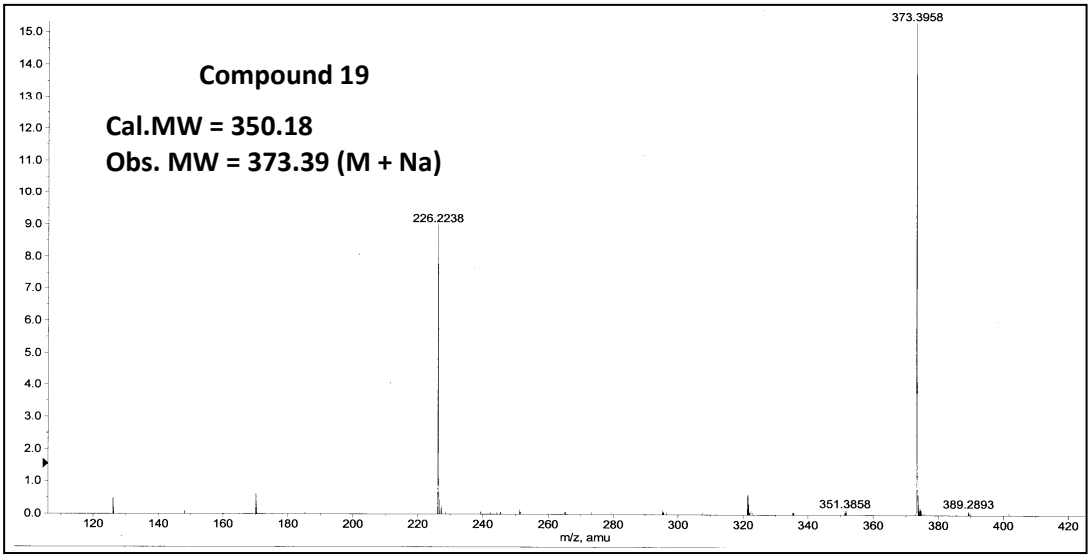

Compound 14
Cal.MW = 203.12
Obs. MW = 204.13
148.0868
226.1245
104.1214
204.1372
m/z, amu

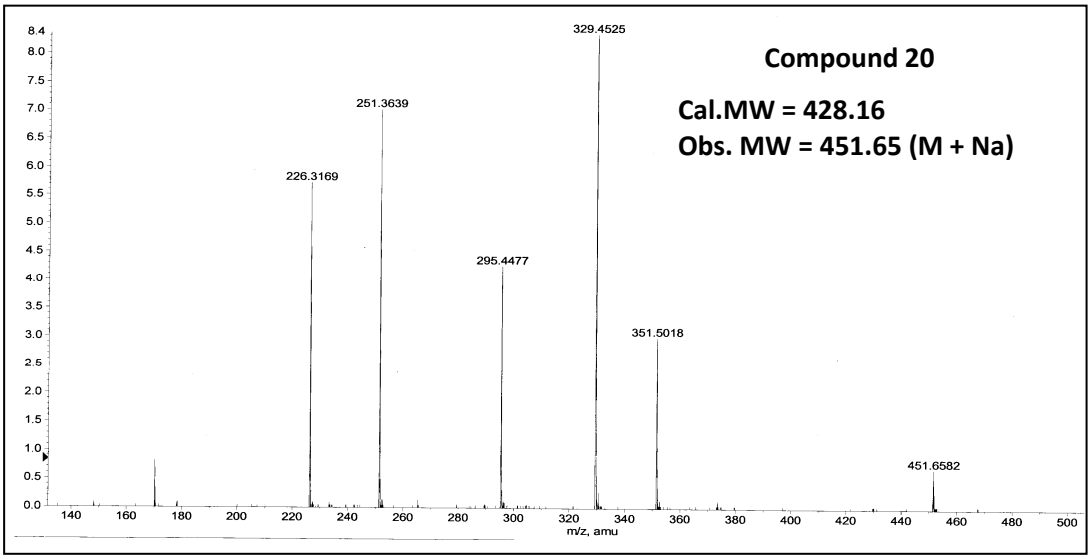

Abundance
Mass spectral difference
Compound 15
Cal.MW = 257.1
Obs. MW = 258.0
230
184
202
156
101
90
142
258
279
302
347
359
m/z-->
100
150
200
250
300
350

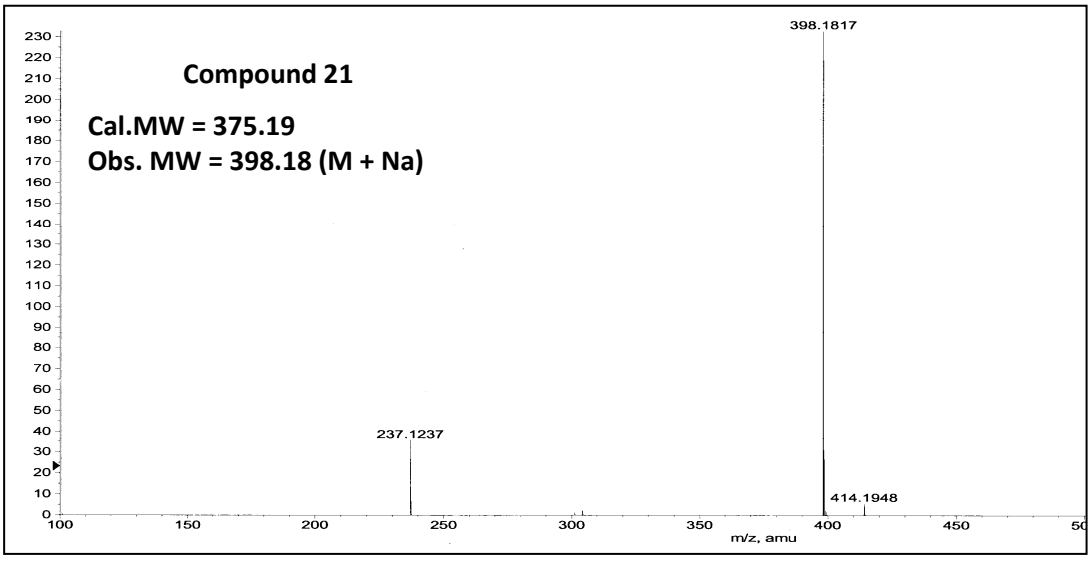

Compound 16
Cal.MW = 362.22
Obs. MW = 363.41
307.3213
363.4113
263.3817
m/z, amu

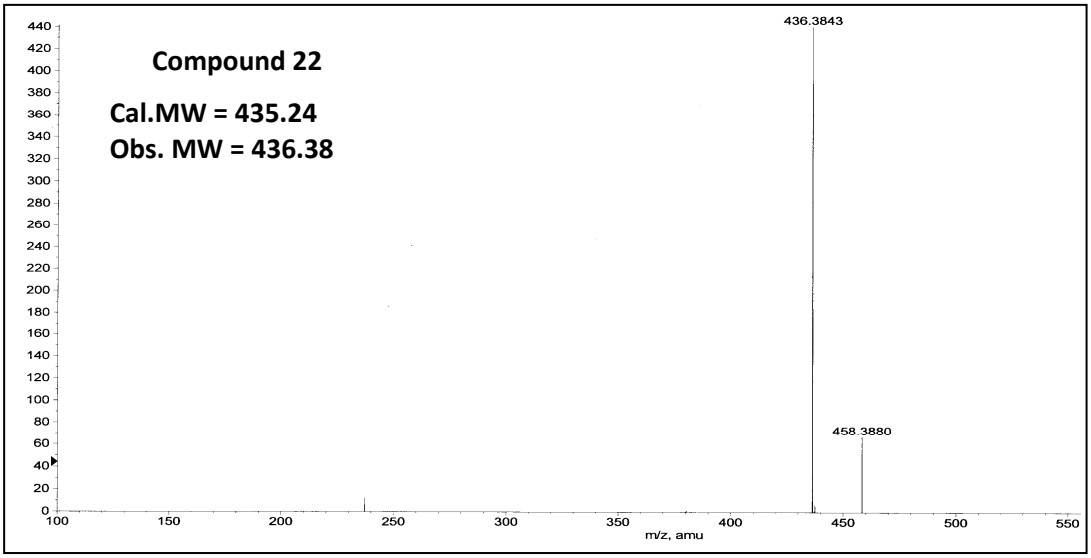

Compound 17
Cal.MW = 362.22
Obs. MW = 363.38
363.3877
307.3503
226.2176
385.4003
m/z, amu

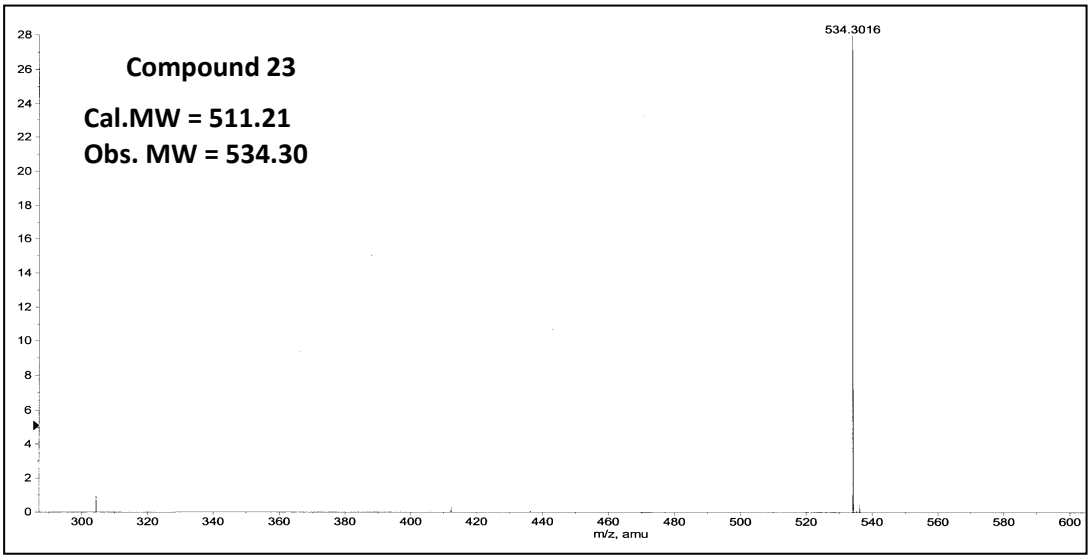

Compound 18
Cal.MW = 320.21
Obs. MW = 321.27
321.2748
265.2933
m/z, amu

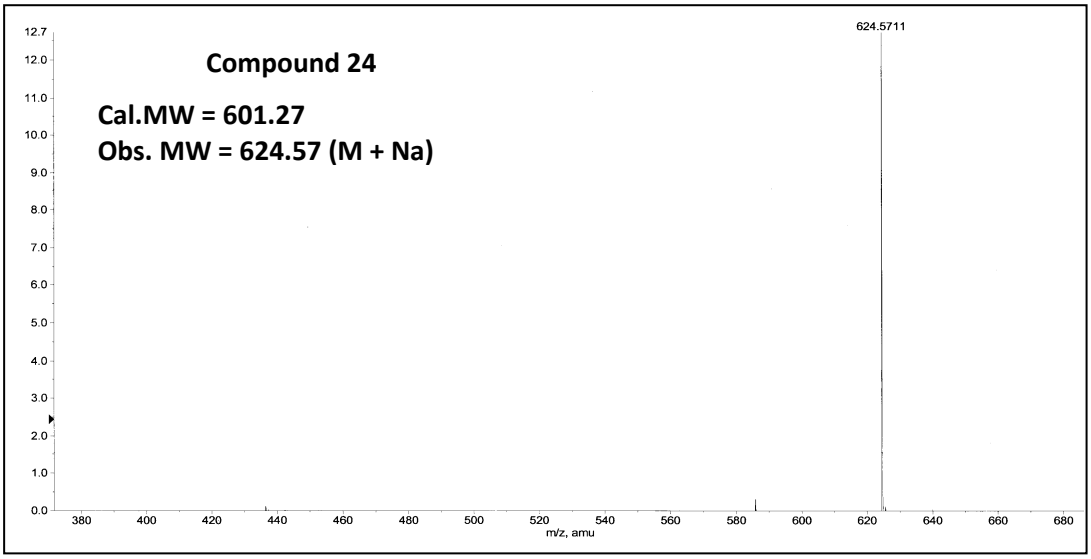

Compound 19
Cal.MW = 350.18
Obs. MW = 373.39 (M + Na)
373.3958
226.2238
351.3858
389.2893
m/z, amu

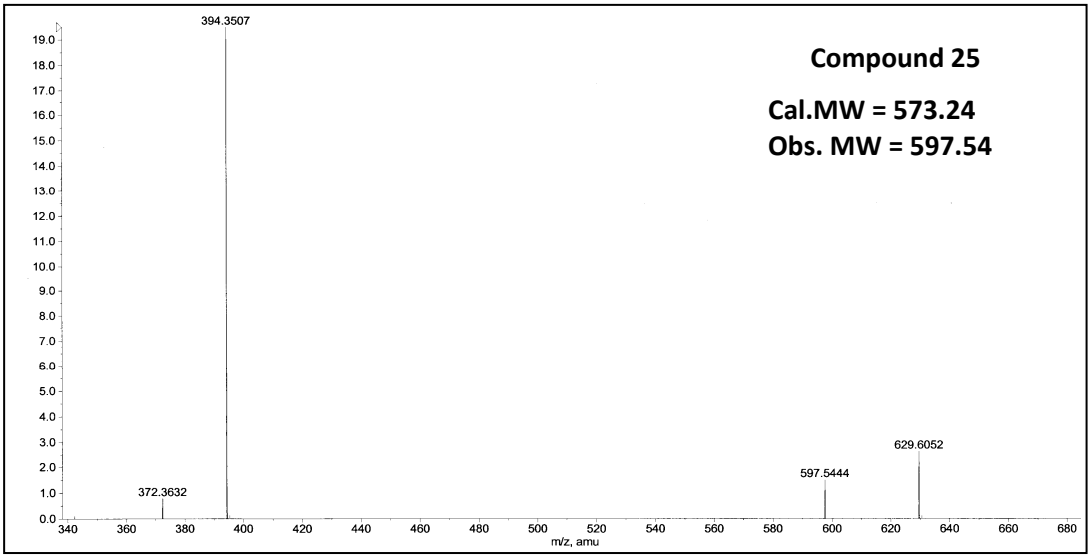

Compound 20
Cal.MW = 428.16
Obs. MW = 451.65 (M + Na)
329.4525
251.3639
226.3169
295.4477
351.5018
451.6582
m/z, amu

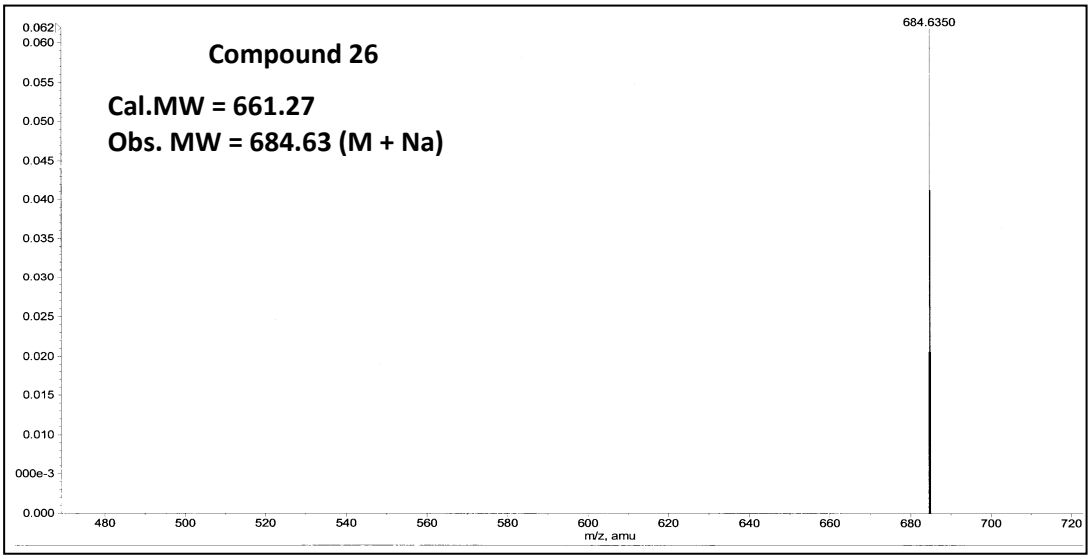

Compound 21
Cal.MW = 375.19
Obs. MW = 398.18 (M + Na)
398.1817
237.1237
414.1948
m/z, amu

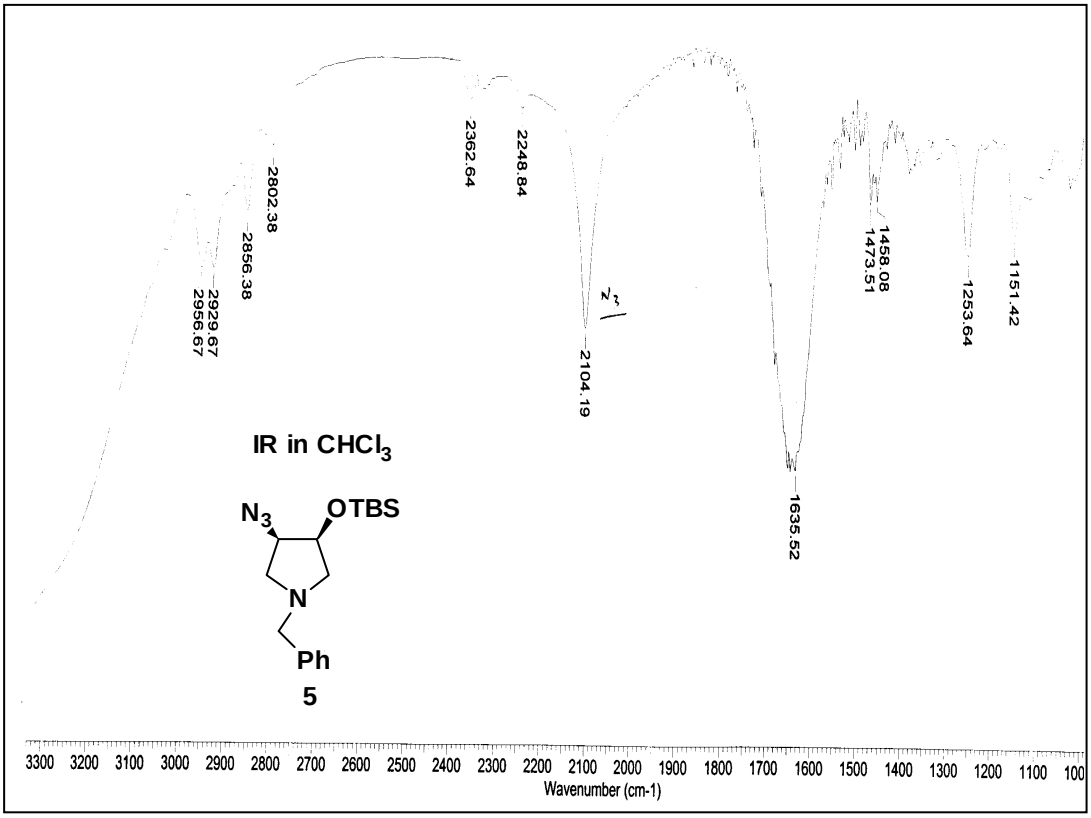

Compound 22
Cal.MW = 435.24
Obs. MW = 436.38
436.3843
458.3880
m/z, amu

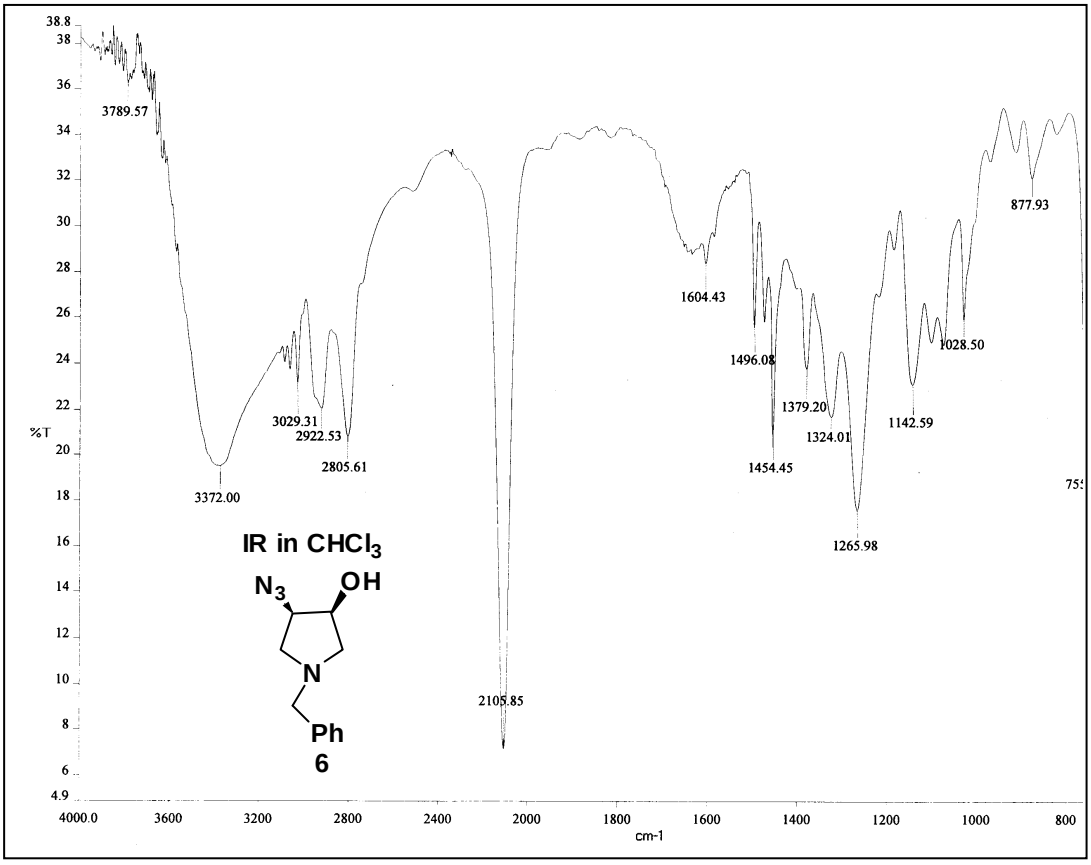

Compound 23
Cal.MW = 511.21
Obs. MW = 534.30
534.3016
m/z, amu

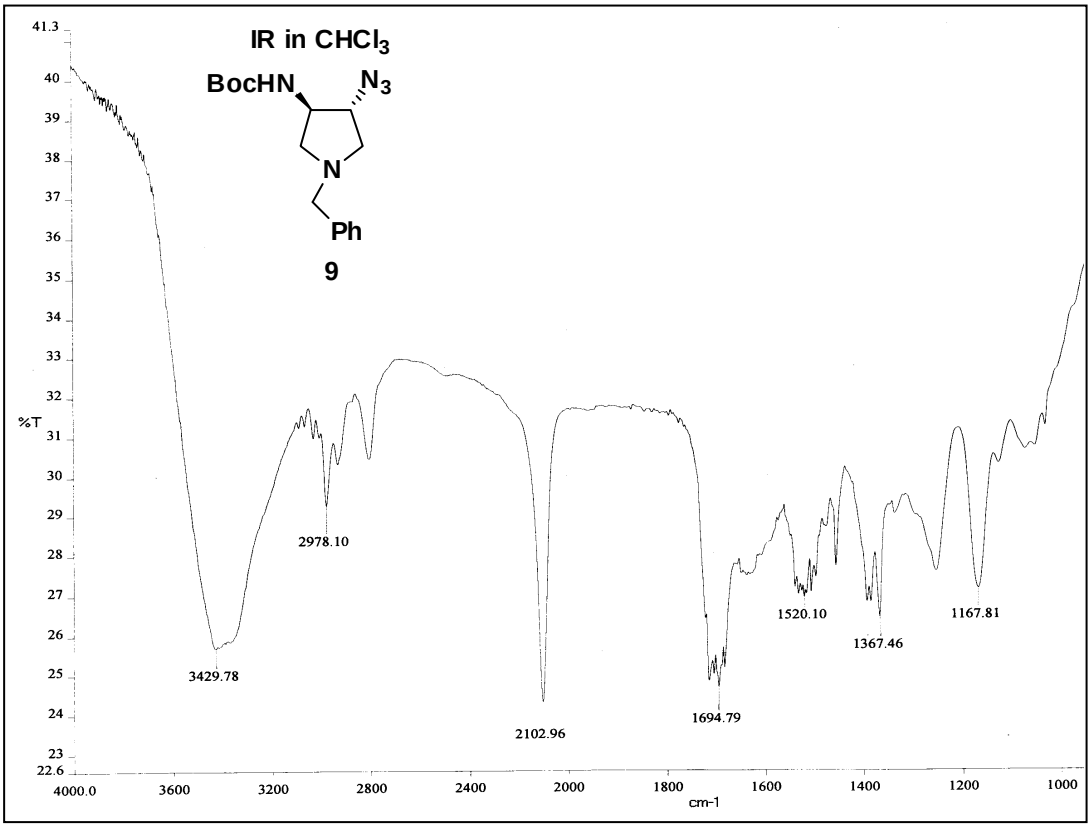

Compound 24
Cal.MW = 601.27
Obs. MW = 624.57 (M + Na)
624.5711
m/z, amu

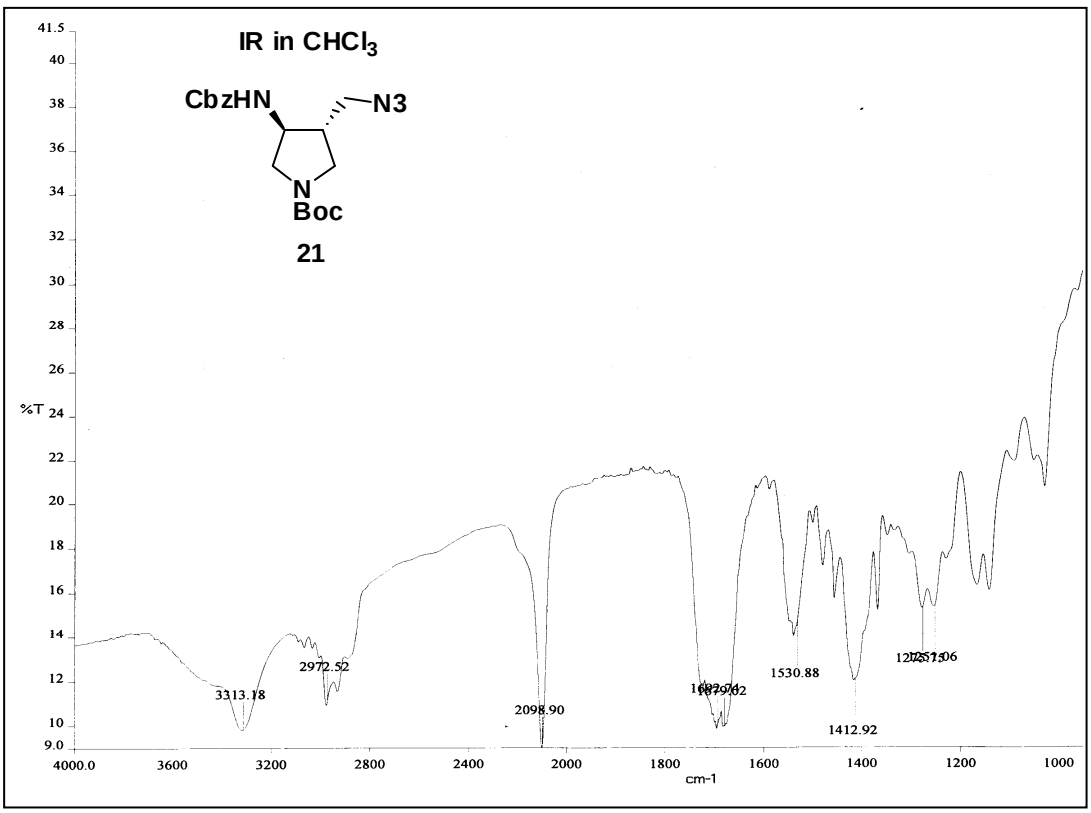

394.3507
Compound 25
Cal.MW = 573.24
Obs. MW = 597.54
372.3632
597.5444
629.6052
m/z, amu


684.6350
Compound 26
Cal.MW = 661.27
Obs. MW = 684.63 (M + Na)
m/z, amu

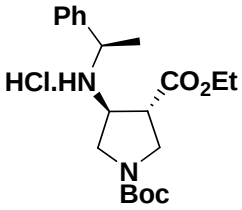

2956.67
2929.67
2856.38
2802.38
2362.64
2248.84
2104.19
1635.52
1473.51
1458.08
1253.64
1151.42
IR in CHCl3
N3
OTBS
N
Ph
5
Wavenumber (cm-1)


%T
3789.57
3372.00
3029.31
2922.53
2805.61
2105.85
1604.43
1496.08
1454.45
1379.20
1324.01
1265.98
1142.59
1028.50
877.93
IR in CHCl3
N3
OH
N
Ph
6
cm-1


IR in $CHCl_3$
BocHN
N_3
N
Ph
9
%T
cm-1
3429.78
2978.10
2102.96
1694.79
1520.10
1367.46
1167.81

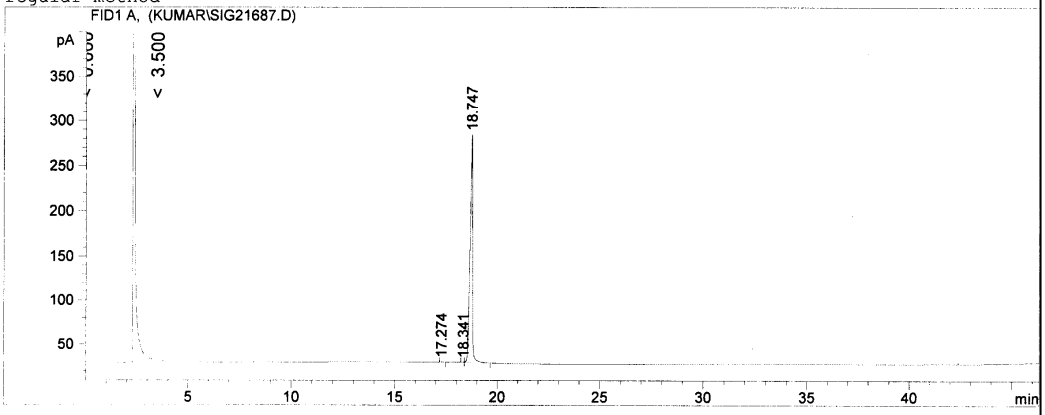

IR in $CHCl_3$
CbzHN
N3
N
Boc
21
%T
cm-1
3313.18
2972.52
2098.90
1530.88
1412.92

GC Analysis to check diastereomeric excess



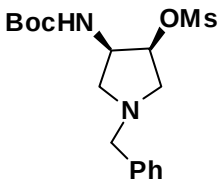
Compound 16

```
150 DEG 2 MINS, 150-220 OC 10 DEG/MIN. 220-250 OC 0.9 D
EGREE/MIN, 250-280 10 Oc/MIN; 280 DEG 1 MINS.ON AGILENT
 6890N COL.HP-5 INJ.280 DET 300 FID.

======================================================================
Injection Date  : 3/19/2008 12:53:42 PM
Sample Name     : AKS-2                          Location : Vial 2
Acq. Operator   : BORIKAR                             Inj :   1
                                                 Inj Volume : Manually
Method          : C:\HPCHEM\1\METHODS\REGULAR.M
Last changed    : 3/19/2008 12:43:57 PM by BORIKAR
                  (modified after loading)
regular method
```




```
======================================================================
                        Area Percent Report
======================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000


Signal 1: FID1 A,

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [pA*s]      [pA]       %
----|-------|----|-------|----------|----------|--------|
   1  17.274 PB    0.0895    2.57291 3.97581e-1  0.11613
   2  18.341 BV    0.0887    3.03087 4.92881e-1  0.13681
   3  18.747 VB    0.1123 2209.84375  254.44867 99.74706

Totals :                  2215.44753  255.33913
```

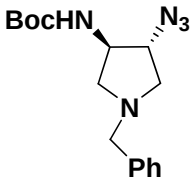
Crystal data and structure refinement for compound 8



8

Identification code	compound **8**
Empirical formula	C17 H26 N2 O5 S
Formula weight	370.46
Temperature	297(2) K
Wavelength	0.71073 A
Crystal system, space group	Orthorhombic, P212121
Unit cell dimensions	a = 6.1261(6) A alpha = 90 deg. b = 9.1598(8) A beta = 90 deg. c = 34.254(3) A gamma = 90 deg.
Volume	1922.1(3) A^3
Z, Calculated density	4, 1.280 Mg/m^3
Absorption coefficient	0.197 mm^{-1}
F(000)	792
Crystal size	0.48 x 0.26 x 0.15 mm
Theta range for data collection	3.26 to 25.00 deg.
Limiting indices	-7<=h<=7, -10<=k<=10, -40<=l<=40
Reflections collected / unique	13995 / 3375 [R(int) = 0.0211]
Completeness to theta = 25.00	99.7 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.9718 and 0.9108
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	3375 / 0 / 250
Goodness-of-fit on F^2	1.061
Final R indices [I>2sigma(I)]	R1 = 0.0350, wR2 = 0.0867
R indices (all data)	R1 = 0.0382, wR2 = 0.0884
Absolute structure parameter	-0.05(7)
Largest diff. peak and hole	0.200 and -0.128 $e.A^{-3}$

Crystal data and structure refinement for compound 9



9

Identification code	Compound **9**
Empirical formula	C16 H23 N5 O2
Formula weight	317.39
Temperature	297(2) K
Wavelength	0.71073 A
Crystal system, space group	Monoclinic, P2(1)
Unit cell dimensions	a = 5.2488(16)A alpha = 90 b = 11.090(3)A beta = 90.723(6) c = 15.462(5)A gamma = 90
Volume	900.0(5) A^3
Z, Calculated density	2, 1.171 Mg/m^3
Absorption coefficient	0.080 mm^{-1}
F(000)	340
Crystal size	0.88 x 0.68 x 0.22 mm
Theta range for data collection	2.26 to 25.00 deg.
Limiting indices	-6<=h<=6, -13<=k<=13, -16<=l<=18
Reflections collected / unique	5022 / 3037 [R(int) = 0.0184]
Completeness to theta = 25.00	97.6 %
Max. and min. transmission	0.9822 and 0.9323
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	3037 / 1 / 211
Goodness-of-fit on F^2	1.076
Final R indices [I>2sigma(I)]	R1 = 0.0404, wR2 = 0.1102
R indices (all data)	R1 = 0.0453, wR2 = 0.1143
Absolute structure parameter	-0.2(15)

CHAPTER 3

Solid phase synthesis and biophysical studies of chiral, cationic *dap*PNA and backbone extended *amp*PNA oligomers.

Introduction

The main goal of designing antisense/antigene molecules is to achieve high binding affinity/specificity 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. Non-toxicity to the cell in which its activity is required is also necessary.[1-2]

In this chapter, the synthesized modified PNA monomers were site specifically incorporated into oligomers along with the unmodified *aeg*PNA monomers 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 has been investigated using the biophysical techniques CD and 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 modified PNAs with DNA. The following sections describe the solid phase peptide synthesis and biophysical techniques in detail.

3.1 Solid Phase Peptide Synthesis

The solid phase peptide synthesis (SPPS) was devised by R. B. Merrifield in 1959. In this method, the peptide is synthesized on an insoluble solid support. In contrast to the solution phase method, the solid phase method devised by Merrifield,[3] 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 acts as a permanent protection for the carboxyl group (Figure 1). The desired N^{α}-protected amino acid is coupled to the resin bound amino acid either by using an active pentaflurophenyl (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 advantage of the solid phase synthesis are (i) all

the reactions are performed in a single vessel minimizing the loss due to transfer, (ii) large excess of molar equivalent of monomer carboxylic acid component is used to achieve high coupling efficiency, (iii) excess reagents can be removed by simple filtration and washing steps and (iv) the method is amenable to automation.

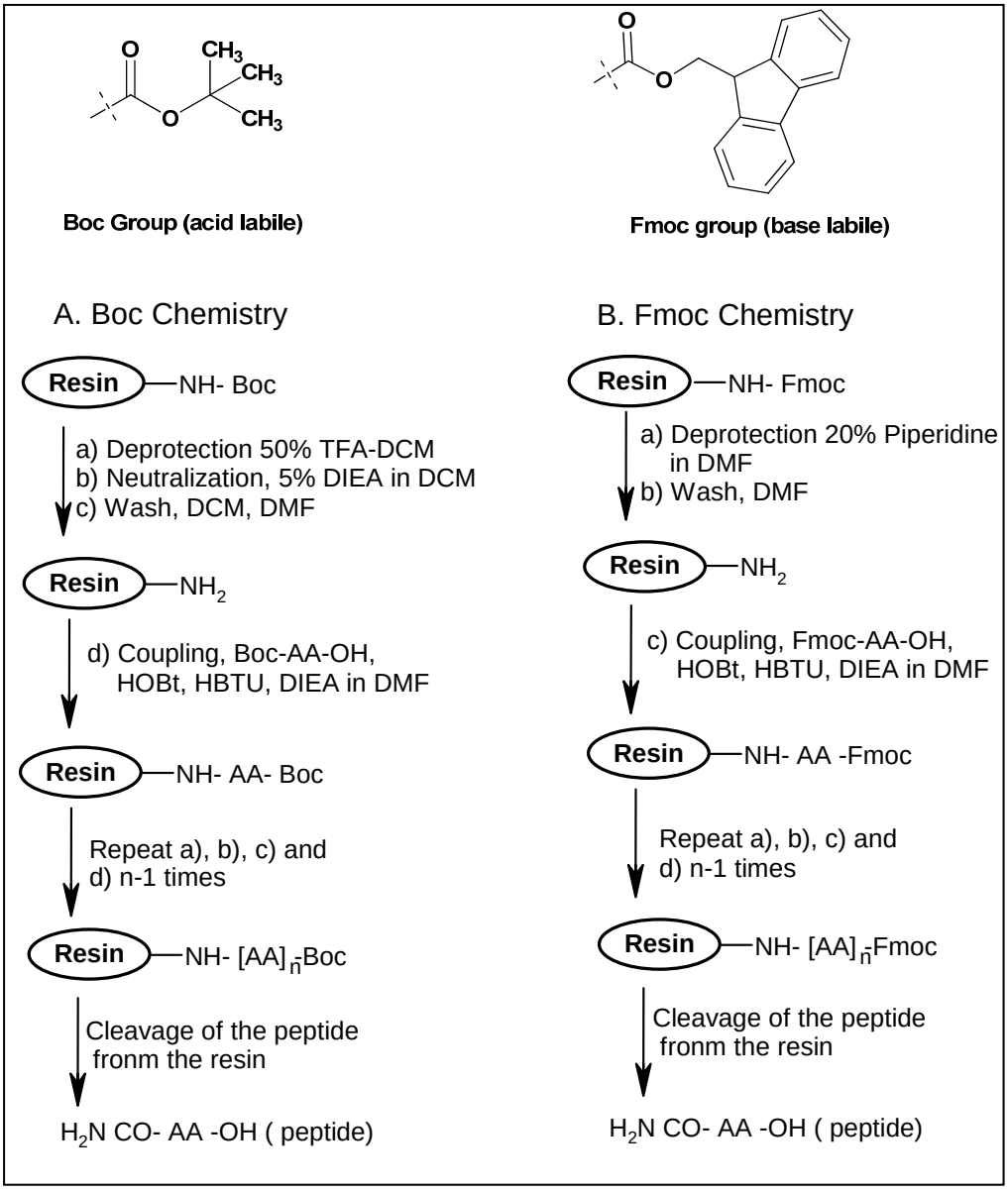


Figure 1. Schematic representation of SPPS A) Boc-chemistry B) Fmoc-chemistry

Currently, two types of chemistries 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. This method involves the use of *t*-butyloxycarbonyl (*t*- Boc) group as N^{α}-protection that could be removed under

acidic conditions such as 50% TFA in DCM.[4] The reactive side chains of amino acids were protected with groups stable to *t*-Boc deprotection conditions (Cbz, benzyl etc.) and removable under strongly acidic conditions using HF in dimethyl sulfide or TFMSA in TFA. Alternatively a base labile protecting group could be used involving fluorenyl methyloxycarbonyl (Fmoc) group for N^{α}- protection, which is stable to acidic conditions but can be deprotected efficiently with a secondary base such as piperidine. This method was introduced by Carpino in 1972.[5] 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.

3.1.1 Resins for solid phase peptide synthesis

3.1.1a Merrifield resin

Merrifield resin (Figure 2a) has been used for more than forty five years for the synthesis of small or medium sized peptides using the Boc strategy. The resin consists of divinylbenzene cross-linked polystyrene beads functionalized with chloromethyl groups. Attachment of 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. Release of peptide or carboxylic acid is normally affected by treatment of the resin with HF or TFMSA.

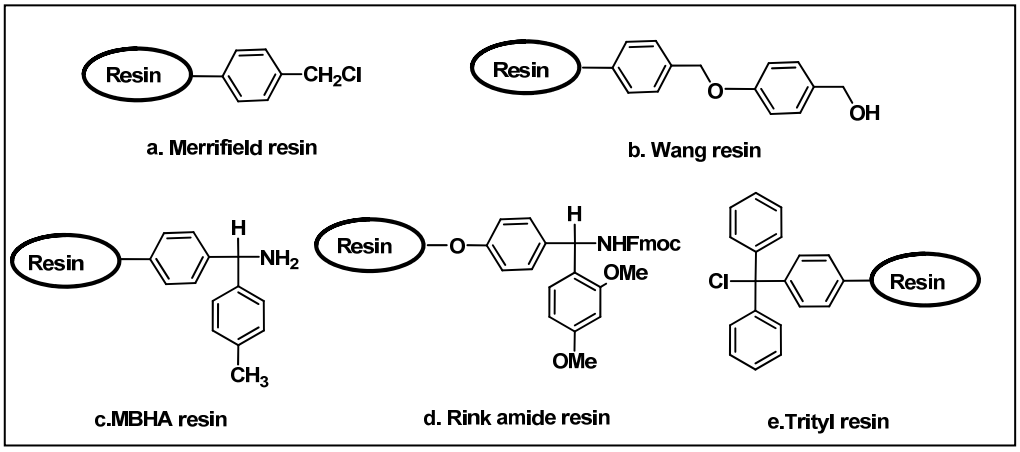


Figure 2. Resins commonly used for solid phase peptide synthesis

3.1.1b Wang resin

Wang resin (Figure 2b) is the standard support for Fmoc SPPS by a batch-wise method. The resin synthesized from chloromethyl polystyrene reacted with 4-hydroxybenzyl alcohol. Attachment of amino acids to benzyl alcohol functionalized supports 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). This support is extremely useful in SPOS for the immobilization of carboxylic acids, alcohols and phenols. The release of these functionalities is usually affected by mild acidolysis with TFA.

3.1.1c MBHA resin

MBHA resin (Figure 2c) is based on 100-200 mesh, 1% DVP cross-linked polystyrene functionalized with 4-methylphenyl-aminomethyl group. These resins are the most widely used ones for the synthesis of peptides by the Boc strategy.[6] Cleavage of the peptides from the resin can be achieved by treatment with HF, TFMSA or HBF_4. In Fmoc SPPS, this resin has been used in conjunction with a low-high acid cleavage, where side-chain protecting groups were first removed prior to cleavage from the resin with TFMSA. The excellent swelling properties of this resin make it an ideal matrix onto which TFA-labile linkers are attached for combinatorial synthesis.

3.1.1d Rink amide resin

Rink amide resin (Figure 2d) is highly acid sensitive because of the benzhydryl linker joined to the support through a benzylic ether bond. Breakdown of the linker during cleavage with high concentration of TFAs can however occur, leading to the formation of highly colored 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.[7] Generally, two types of Rink amide resin 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 latter resins are generally regarded as the more robust, particularly under acidic conditions.

3.1.1e Trityl resin

The advantage of trityl functionalized resin (Figure 2e) is that any nucleophile can be linked to these resins under extremely mild conditions. Most of the side-chains of the trifunctional 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 products from the supports takes place under very mild conditions, owing to the high stability of the trityl cations. Furthermore, trityl cations are extremely poor electrophiles and as such do not undergo alkylation side-reactions.[8]

3.1.2 Coupling reagents

Carbodiimides have been some of the most popular *in situ* activating reagent in peptide synthesis. Dicyclohexylcarbodiimide (DCC) remains the most popular and a particularly appropriate choice for the apolar nature of polystyrene resins.[9] The principal limitation in using carbodiimides is the dehydration of Asn and Gln residues. The addition of HOBt (Hydroxybenzotriazole) to the reaction mixture will prevent dehydration and has the added benefit of catalysis. The principal drawback for use of Carbodiimides is the formation of the DCM-insoluble urea by-products during activation and acylation.

In situ activating agents are widely accepted because they are easy to use, effect fast reactions even between sterically hindered amino acids and generally free of side reactions. Most are based on phosphonium or uronium salts which in presence of a tertiary base, can smoothly convert protected amino acids to a variety of activated species. The most commonly employed reagents have benzotriazole moiety e.g. BOP, PyBOP, HBTUand TBTU, thus generate -OBt (benzotraizole) ester, and these have wide application in routine SPPS and solution synthesis for difficult couplings (Figure 3). BOP should be handled with great care as the by-product formed during activation is highly carcinogenic. BOP can be substituted by

PyBOP, which has pyrrolidine unit instead of N-methyl in BOP, without loss of performance.

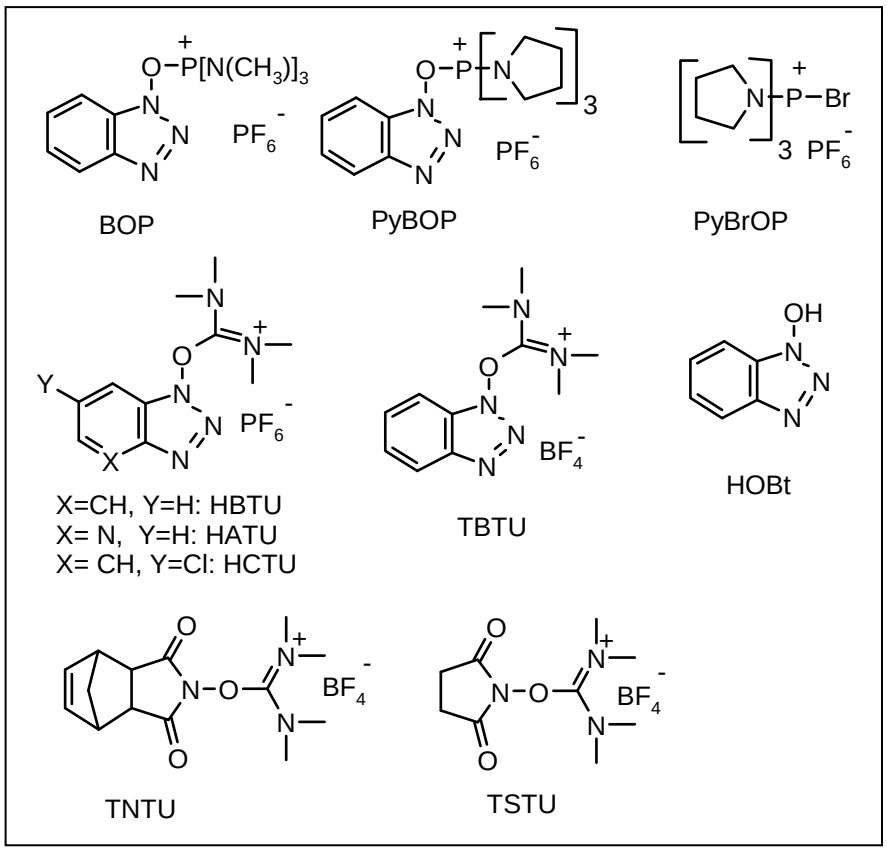


Figure 3. Coupling reagents used for Peptide synthesis.

Coupling reagents which generate esters that are more reactive than OBt ester have been known. Two such reagents are HATU [10] and HCTU [11], which in presence of base converts carboxylic acids to corresponding -OAt (7-azabenzotriazole) and –OCt (6-chlorobenzotraizole) esters respectively. Such 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.[10] HATU is generally regarded as the most efficient coupling reagent. However, owing to its cost, its use is generally recommended for difficult and hindered couplings and the synthesis of long peptides.

The coupling of N-methyl amino acids is a difficult and slow process. PyBOP[12] allows efficient and fast couplings, providing a significant improvement over existing reagents.[13]

3.1.3 Resin Tests

The most widely used procedure for determining efficiency of coupling is to test for the presence or absence of free amino groups (deprotection/coupling) by Kaiser test[14] which is simple and quick (details in experimental section). However, it should be noted that some deprotected amino acids do not show the expected dark blue color typical of free primary amino groups (e.g., serine, aspargine, aspartic acid) and that proline, being a secondary amino acid, does not yield a positive reaction. Furthermore, occasionally false negative tests are observed, particularly with strongly aggregated sequences, which show that one has to be very careful when assessing the extent of deprotection or coupling on solid phase. Other methods such as picric acid monitoring [15] or mass spectrometry[16] are also available. For iminoacids like proline, the chloranil test[17] is recommended. Recently a new method for detecting resin-bound primary and secondary amines has been described.[18] 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 their carboxylic acid functions free. 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).[19] The amount of free -NH_2 on the resin available for coupling was estimated by the picrate assay,[19] 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)[20] 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 CD and temperature dependent UV-spectroscopy. CD spectroscopy has utility in recognizing DNA/RNA/PNA triplexes and duplexes, and complexes formed between protein-DNA[21]. 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 modified PNAs with DNA.

3.2.1 Circular Dichroism Spectroscopy (CD spectroscopy)

CD spectroscopy, a class of absorption spectroscopy, can provide information on the structures of many types of chiral biological macromolecules that possess UV-Vis absorbing chromophores. Circular dichroism is the difference between the absorption of left and right handed circularly-polarized light and is measured as a function of wavelength. The absence of chirality in a regular structure results in zero CD intensity, while an ordered chiral structure results in a spectrum which contains either positive or negative signals.

Although detailed structural information at atomic level as obtained from X-ray crystallography or NMR spectroscopy is not available from CD spectra, it can provide a reliable estimation of the overall conformational state of biopolymers and any structural changes induced by modification when compared that of reference samples. In case of nucleic acids, the sugar units of the backbone possess chirality and the bases attached to sugars are the chromophores.

3.2.1a CD spectroscopy of PNA:DNA triplexes

CD spectroscopy is a useful technique to monitor the structural changes of nucleic acids in solutions and for diagnosing whether new or unusual structures are formed by particular polynucleotide sequences. PNA is non-chiral and shows weak CD signature. CD is predominantly an effect of coupling between the transition moments of the nucleobases as a result of their helical stacking. The reliance on CD spectroscopy to study nucleic acid conformations has stemmed from the sensitivity of technology, ease of CD measurements, its non-destructive nature and the fact that conformations can be studied in solution.

In the CD spectrum of *ss*-DNA, deep intensive negative band is observed between 200-210nm, a high intensive positive band in the region 210nm-240nm, a low intensity negative band between 240-260nm and a moderately intensive positive band in the region 260-300nm (Figure 4A).

PNA being achiral, shows insignificant CD signature. However being a polyamide, it can form left or right handed coil structures with equal facility. Due to these type of structures, very low intense negative band between 205-240nm is seen

due to amide bonds present in PNA oligomer and no absorptions of consequence were seen in between 240-300nm range.

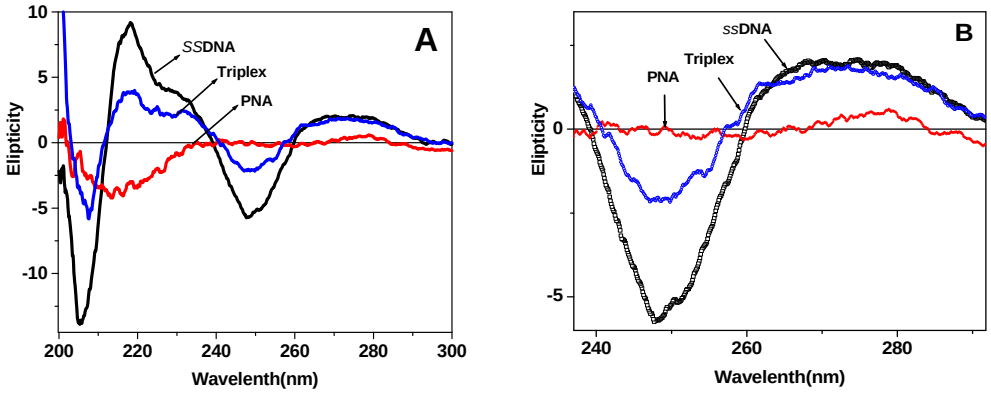


Figure 4: The representative graph A) CD signatures of *ss*DNA, PNA and triplex B) the characteristic CD signature of triplex.

Even though PNA shows low profile CD signals inspite of its achiral nature, PNA:DNA complexes exhibit characteristic CD signature since, PNA adopts helicity when it binds to DNA and forms the chiral PNA-DNA complex. In case of PNA_2:DNA triplex, the CD pattern is observed with a negative band between 205-210nm, followed by high intensity positive band in the region 210nm-240nm, low intense negative band between 240-255nm and a positive band in 260-300nm range. The PNA_2:DNA triplex shows characteristic band between 260nm-285nm (Figure 4B). The elipticity of triplex in PNA_2:DNA triplex in the region of 255-265 nm was more than both single stranded DNA and PNA.

3.2.2 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 was investigated by this technique. Increase in temperature dissociates the complexes, inducing dissociation into individual components by disruption of hydrogen bonds between the base pairs. This results in decrease of stacking between adjacent nucleobases leading to a 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 midpoint of

transition gives the T_m (Figure 5A). In case of DNA triplexes, double transition sigmoidal curve is seen, the first transition corresponding to triplex melting to the duplex (Watson–Crick duplex), whereas the second transition is due to duplex dissociation into two single strands.[22] In case of PNA:DNA complexes, the PNA_2:DNA triplexes show single sigmoidal curve since the temperature difference between two transition states is too low to distinguish them independently.[23] The two PNA strands dissociate from DNA strand simultaneously. The melting temperature T_m can be obtained from the maxima of the first derivative plots (Figure 5B). 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 a single base mismatch leads to incorrect hydrogen bonding between the bases causing 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.

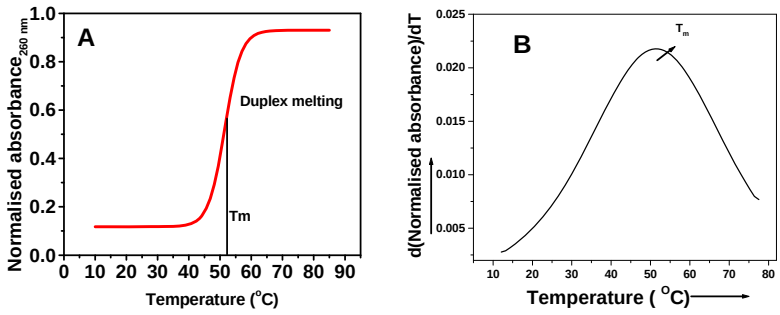


Figure 5: Schematic representation of UV melt A) Sigmoidal curve, B) First derivative curve.

3.3 Results and Discussion

3.3.1 Solid Phase synthesis of *dap*PNA 13 and *amp*PNA 26 oligomers.

The two modified PNA monomers, diaminopyrrolidine derived *dap*PNA monomer **13** and aminomethylpyrrolidine derived *amp*PNA monomer **26** (Figure-6), synthesized as described in Chapter-2. These can be incorporated into standard PNA oligomers by using either Boc or Fmoc chemistry for solid phase synthesis leading to four different PNA backbones (Figure 6), depending upon the direction of synthesis.

The extra primary or secondary amine functions in **33-36** are capable of accepting a proton and help to make PNA cationic in all the cases.

The *dap*PNA monomer **13** was used to synthesize PNA oligomers with modified backbone shown in **33** (*dap*PNA unit) and **34** (*secdap*PNA unit) by using Boc and Fmoc chemistry respectively. The term *sec* represents the modified PNA backbone involves secondary ring Nitrogen. The *amp*PNA monomer **26** was used to synthesize PNA oligomer with modified backbone shown in **35** (*amp*PNA unit) using Fmoc chemistry. The monomer **25** was used to give PNA oligomer with modified backbone involving secondary ring nitrogen as shown in **36** (*secamp*PNA unit) using Boc chemistry.

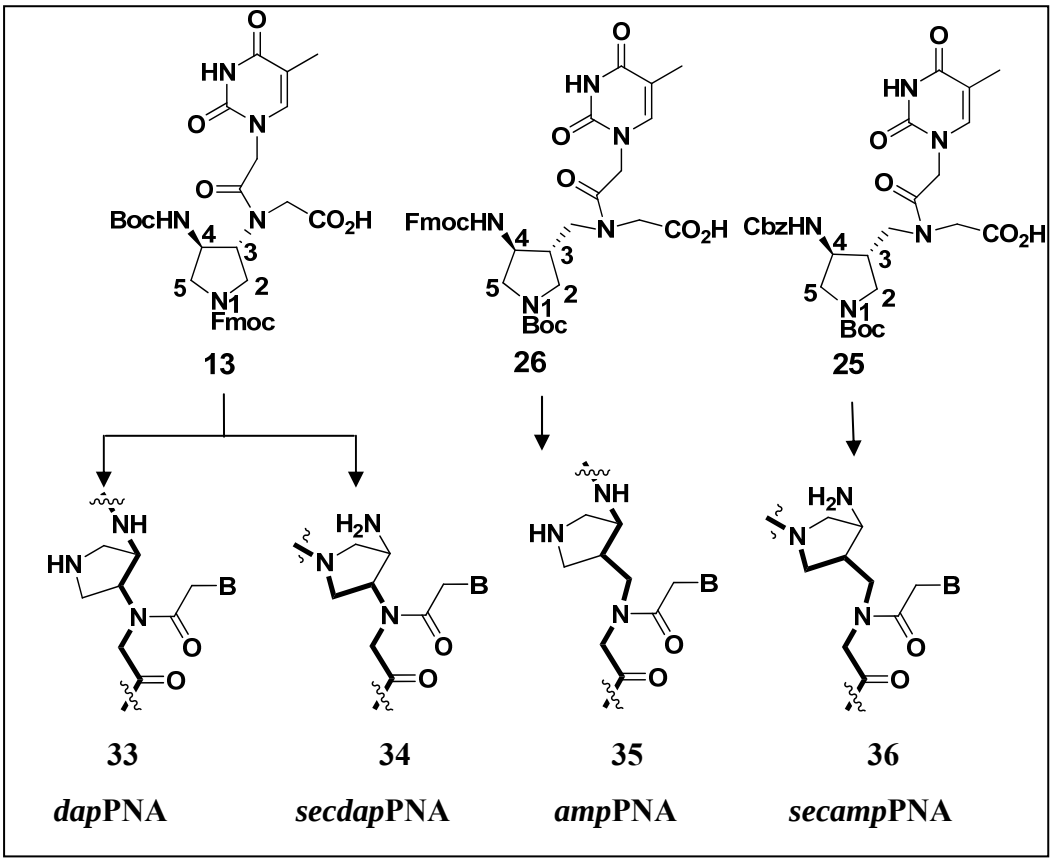


Figure 6: Backbones (shown by thick lines) derived from modified PNA monomers.

The synthesis of the oligomers incorporating the conformationally constrained chiral modified units **13, 25 & 26** at specific positions in the *aeg*PNA oligomer was done on the solid support using the procedures described above. Oligo thymine octamers (**PNA1-PNA13**, Table-1) were synthesized to study the effect of the modified PNA backbone for their triplex forming ability. The mixed purine-

pyrimidine oligomers (**PNA14-PNA21**, Table-2**)** were synthesized for comparative duplex studies with mixed control *aeg*PNA oligomer.

Table 1: Homopyrimidine sequences synthesized by SPPS

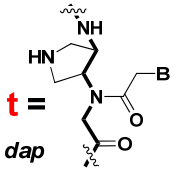
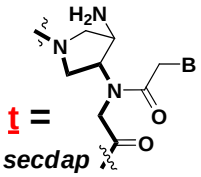
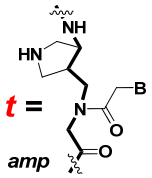
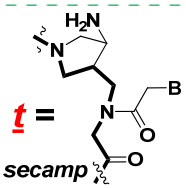
Modified unit	Entry	Sequence
T = *aeg* unit	PNA 1	H-TTTTTTTT-Lys-NH_2
 t = *dap*	PNA 2	H-tTTTTTTT-Lys-NH_2
	PNA 3	H-TTTtTTTT-Lys-NH_2
	PNA 4	H-TTTTTTTt-Lys-NH_2
 t = *secdap*	PNA 5	H-tTTTTTTT-Lys-NH_2
	PNA 6	H-TTTtTTTT-Lys-NH_2
	PNA 7	H-TTTTTTTt-Lys-NH_2
 t = *amp*	PNA 8	H-tTTTTTTT-Lys-NH_2
	PNA 9	H-TTTtTTTT-Lys-NH_2
	PNA 10	H-TTTTTTTt-Lys-NH_2
 t = *secamp*	PNA 11	H-tTTTTTTT-Lys-NH_2
	PNA 12	H-TTTtTTTT-Lys-NH_2
	PNA 13	H-TTTTTTTt-Lys-NH_2

Table 2: Mixed purine-pyrimidine sequences synthesized

S. No.	Entry	Sequence
1	PNA 14	H-GTAGATCACT-Lys-NH_2
2	PNA 15	H-GtAGATCACT-Lys-NH_2
3	PNA 16	H-GTAGAtCACT-Lys-NH_2
4	PNA 17	H-GtAGAtCACT-Lys-NH_2
5	PNA 18	H-GtAGATCACT-Lys-NH_2
6	PNA 19	H-GtAGATCACT-Lys-NH_2
7	PNA 20	H-GtAGATCACT-Lys-NH_2
8	PNA 21	H-GtAGAtCACT-Lys-NH_2

3.3.2 Purification of the PNA oligomers

All the cleaved oligomers were subjected to HPLC purification. The purity of the so obtained oligomers was re-checked by analytical HPLC (C18 column, CH_3CN-H_2O system). These were subsequently purified by reverse phase HPLC on a semi preparative C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis (Table-3).

Table-3: MALDI-TOF mass spectral analysis of modified PNAs

Entry no.	Sequence	Calculated mass	Observed mass
PNA 1	H-TTTTTTTT-Lys-NH_2	2272.31	2275.70
PNA 2	H-tTTTTTTT-Lys-NH_2	2313.61	2314.32
PNA 3	H-TTTtTTTT-Lys-NH_2	2313.61	2312.59
PNA 4	H-TTTTTTTt-Lys-NH_2	2313.61	2315.92
PNA 5	H-tTTTTTTT-Lys-NH_2	2313.61	2316.76
PNA 6	H-TTTtTTTT-Lys-NH_2	2313.61	2316.66
PNA 7	H-TTTTTTTt-Lys-NH_2	2313.61	2316.08
PNA 8	H-*t*TTTTTTT-Lys-NH_2	2327.78	2352.47 (M^++ Na)
PNA 9	H-TTT*t*TTTT-Lys-NH_2	2327.78	2329.61
PNA 10	H-TTTTTTT*t*-Lys-NH_2	2327.78	2328.37
PNA 11	H-*t*TTTTTTT-Lys-NH_2	2327.78	2331.53
PNA 12	H-TTT*t*TTTT-Lys-NH_2	2327.78	2329.29
PNA 13	H-TTTTTTT*t*-Lys-NH_2	2327.78	2330.04
PNA 14	H-GTAGATCACT-Lys-NH_2	2852.52	2853.38
PNA 15	H-GtAGATCACT-Lys-NH_2	2893.64	2916.46 (M^++ Na)
PNA 16	H-GTAGAtCACT-Lys-NH_2	2893.64	2895.83
PNA 17	H-GtAGAtCACT-Lys-NH_2	2934.73	2937.60
PNA 18	H-GtAGATCACT-Lys-NH_2	2893.64	2895.27
PNA 19	H-G*t*AGATCACT-Lys-NH_2	2907.51	2909.77
PNA 20	H-G*t*AGATCACT-Lys-NH_2	2907.51	2909.73
PNA 21	H-G*t*AGA*t*CACT-Lys-NH_2	2962.85	2964.65

3.3.3 Synthesis of complementary oligonucleotides

The oligonucleotides DNA**1-4** (Table 4) corresponding to complementary DNA sequences were synthesized on Applied Biosystems ABI 3900 DNA Synthesizer using standard *β*-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-4: DNA oligonucleotides used in the present work

S. No.	Entry	Sequence (5'-3')	Type
1	DNA 1	GCAAAAAAAACG	Match
2	DNA 2	GCAAAACAAACG	Mismatch
3	DNA 3	AGTGATCTAC	Antiparallel
4	DNA 4	CATCTAGTGA	Parallel

3.3.4 Biophysical studies of PNA:DNA complexes

For the study of binding selectivity, specificity and discrimination of the modified PNA towards complementary DNA, the stoichiometry of the modified PNA:DNA complexes was first determined using Job's method. The binding stoichiometry can be determined by UV spectroscopy measurement at a particular wavelength by plotting absorbance vs mole fractions. The discontinuation point is taken as the binding stoichiometry of DNA:PNA complexes.

The UV-job`s plot experiments were carried out by mixing the appropriate oligomers in different molar ratios keeping the total concentration constant. In this experiment, the absorbance of different molar ratios (100:0, 90:10, 80:20, 70:30, 60:40, 50:50 to 0:100) of **PNA6:DNA1** and **PNA12:DNA1** complexes were recorded at 260 nm by maintaining the total concentration constant.

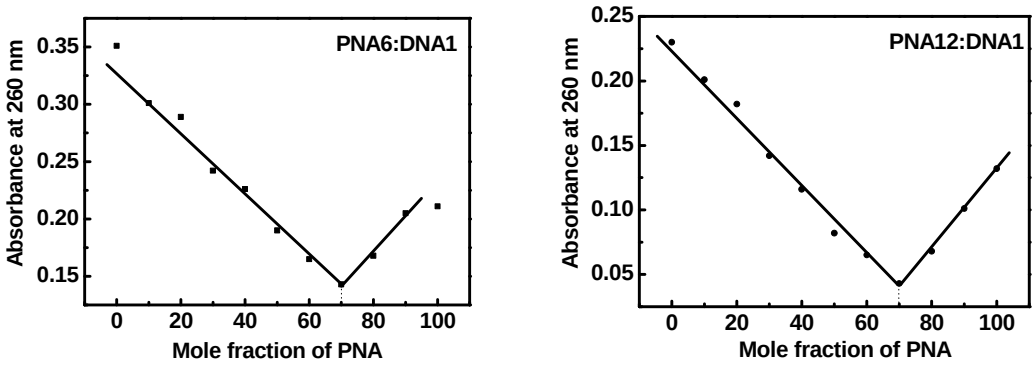


Figure 7. UV job`s plot for **PNA 6** and **PNA 12** with complementary **DNA1,** (CGA8GC) in molar ratios of (a)100:0 (b) 90:10 (c) 80:20 (d) 70:30 (e) 60:40 (f) 50:50 (g) 40:60 (h) 30:70 (i) 20:80 (j) 10:90 (k) 0:100 (Buffer, 10 mM Sodium phosphate pH 7.4, 10 mM NaCl) at 260nm.

The profile showed a breakpoint around 2:1 stoichiometry for homopyrimidine **PNA6:DNA1** and **PNA12:DNA1** complexes at wavelength 260 nm, confirming the formation of PNA_2:DNA triplex (Figure 7).

3.3.4a CD Spectroscopy of PNA:DNA complexes

The **PNA 1-21** (Table 1 & 2) single strand oligomers exhibited very low induced CD signals due to presence of chiral modified units. The CD for single stranded PNA oligomers and their complex with complementary DNA (Figure 8) was recorded. The homopyrimidine PNA oligomers (**PNA 1-PNA 13**) when complexed with complementary **DNA 1** show a very low induced CD with positive maxima at 270 - 285 nm, a negative minimum at 240 -250 nm. A positive band in the region of 260 to 290 nm as seen in the CD spectra of the complexes is characteristic of the *poly*(dA) $[PNA\text{-}T_8]_2$ complex.[24] The UV-Job`s plot was also done for **PNA 2** and **PNA 11** (see experimental section) to confirm that homopyrimidine PNA oligomers bind to **DNA 1** in 2:1 stoichiometry i.e. PNA_2:DNA. Similarly CD spectra have been recorded for the single stranded mixed purine/pyrimidine sequences (**PNA 14-PNA 21)** for the complexes with complementary **DNA 3** for the antiparallel binding mode. These spectra show positive maxima between 260 to 280 nm and negative maxima between 240 to 260 nm.

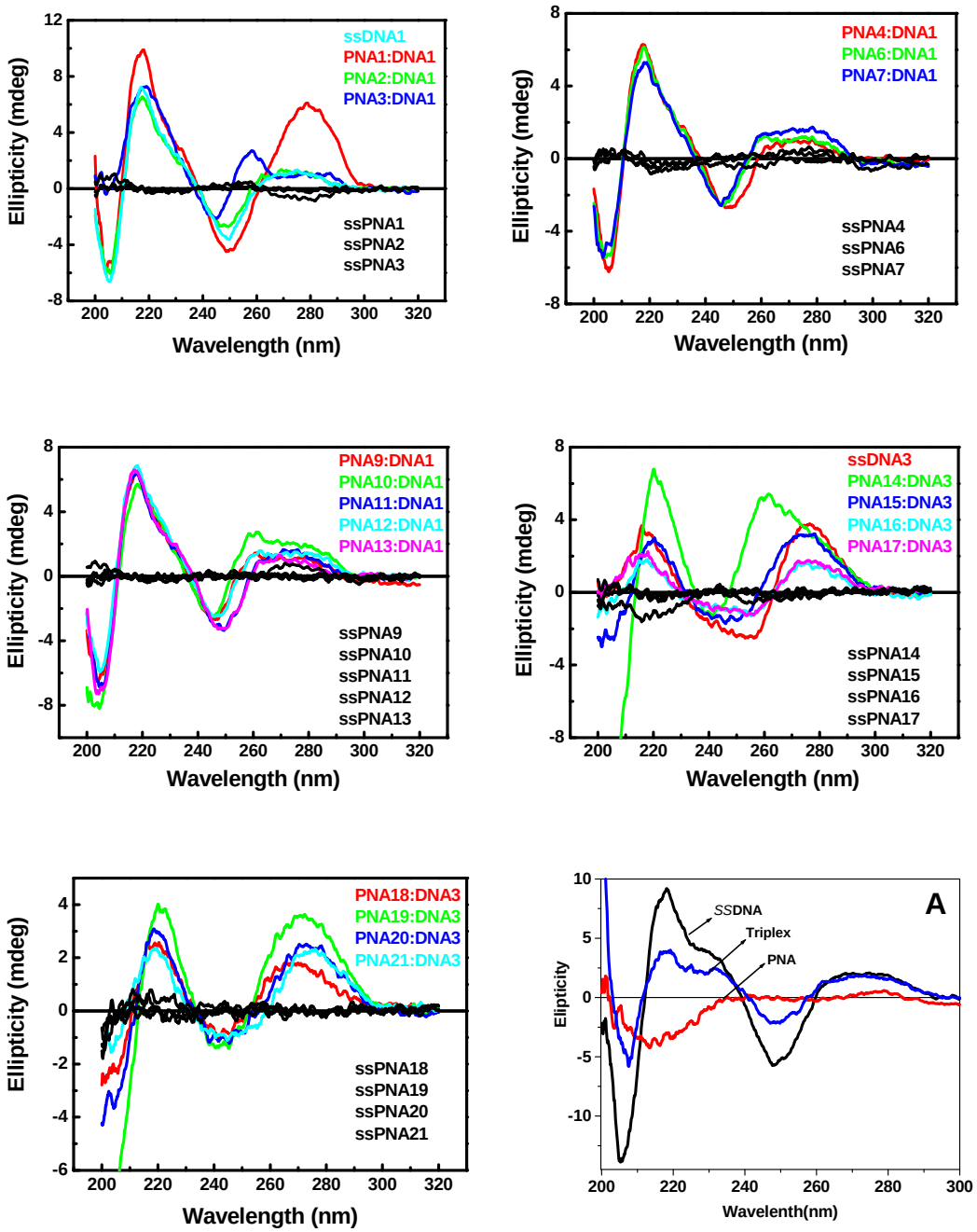


Figure 8: CD patterns for single stranded and hybridized **PNA1**-**PNA21** with complementary DNAs

3.3.4b UV-Spectroscopy

The hybridization studies of modified PNAs with complementary DNA were done by temperature dependent UV-absorbance experiments. The thermal stabilities (T_m) of PNA_2:DNA triplexes were obtained for different PNA modifications with

complementary **DNA1** and mismatch **DNA2** (Figure 9-10, Table 5) and PNA:DNA duplexes with complementary **DNA3** (Antiparallel) and **DNA4** (Parallel) (Figure 11, Table 6).

3.3.4c Thermal stability of triplexes

Unlike DNA triplexes, which show two distinct transitions corresponding to triplex to duplex melting and later duplex to single strands melting, modified PNA_2:DNA triplexes show single transition similar to *aeg*PNA_2:DNA triplex melting.

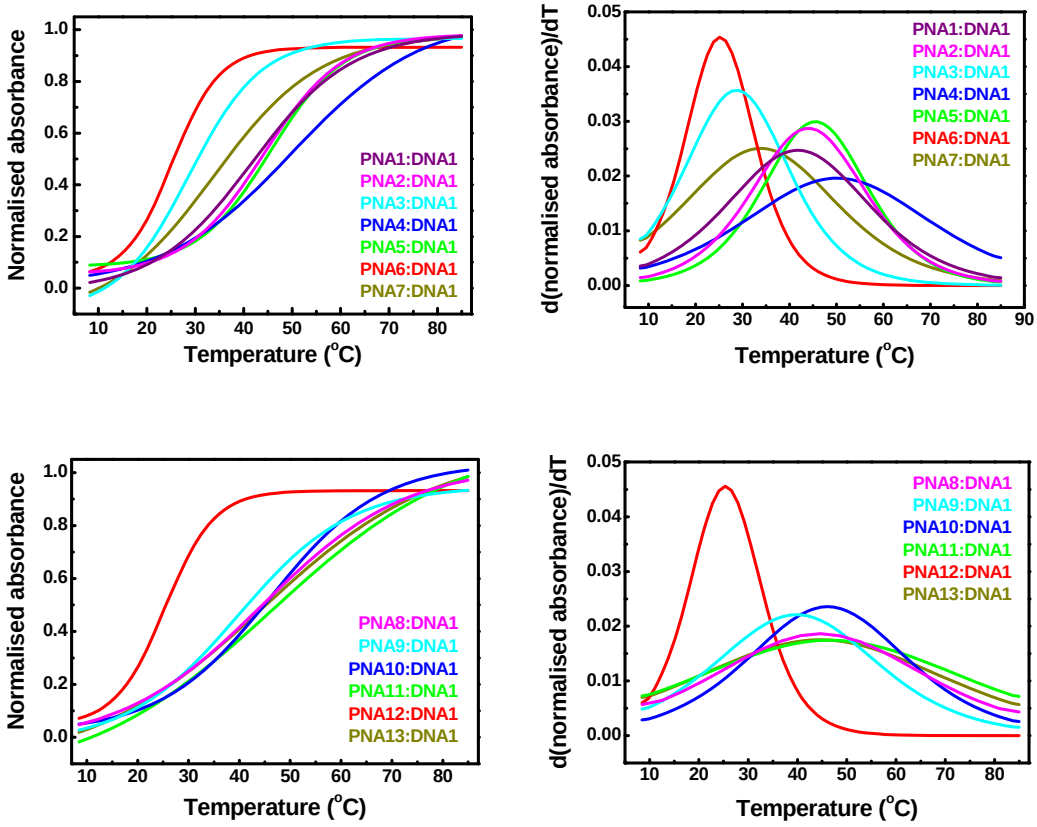
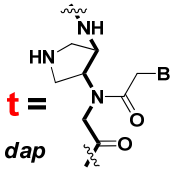
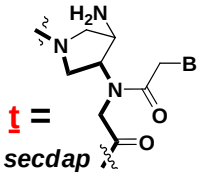
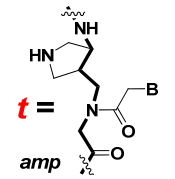
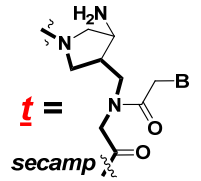


Figure 9: Melting curves for homopyrimidine PNAs (**PNA1-PNA13**) complexed with **DNA1** (CGA8GC) (10 mM sodium phosphate buffer, pH=7.4, 10 mM NaCl).

Table 5: Melting studies of PNA_2:DNA triplexes

Modified unit	S. No	Entry	Sequence	UV T_m* DNA 1 (ΔT_{m1})	UV T_m* DNA2 (Mismatch) (ΔT_{m2})
T = *aeg* unit	1	PNA 1	H-TTTTTTTT-Lys-NH_2	42.1	31.5 (-10.6)
 t = *dap*	2	PNA 2	H-tTTTTTTT-Lys-NH_2	44.8 (+2.7)	34.0 (-10.8)
	3	PNA 3	H-TTTtTTTT-Lys-NH_2	29.0 (-13.1)	nd
	4	PNA 4	H-TTTTTTTt-Lys-NH_2	49.8 (+7.7)	nd
 t = *secdap*	5	PNA 5	H-tTTTTTTT-Lys-NH_2	45.6 (+3.5)	34.1 (-10.5)
	6	PNA 6	H-TTTtTTTT-Lys-NH_2	25.1 (-17.0)	nd
	7	PNA 7	H-TTTTTTTt-Lys-NH_2	34.3 (-7.8)	22.3 (-12.0)
 t = *amp*	8	PNA 8	H-tTTTTTTT-Lys-NH_2	44.9 (+2.8)	36.7 (-8.2)
	9	PNA 9	H-TTTtTTTT-Lys-NH_2	39.4 (-2.7)	nd
	10	PNA 10	H-TTTTTTTt-Lys-NH_2	45.8 (+3.7)	32.8 (-13.0)
 t = *secamp*	11	PNA 11	H-tTTTTTTT-Lys-NH_2	45.8 (+3.7)	38.0 (-7.8)
	12	PNA 12	H-TTTtTTTT-Lys-NH_2	25.0 (-17.1)	nd
	13	PNA 13	H-TTTTTTTt-Lys-NH_2	44.6 (+2.5)	22.4 (-22.2)

***DNA1** is **5'-GCAAAAAAAACG-3'**, **(ΔT_{m1})** = difference w.r.t. unmodified *aeg*PNA (**PNA1**). **DNA2** is **5'-GCAAAACAAACG-3' (mismatch)**, **(ΔT_{m2})** = difference in match & mismatch experiments, nd= no detectable melting transition.

The UV-melting profiles of all the homopyrimidine sequences (**PNA2-PNA13**, entry 2-13) are shown in Figure 9. The T_m data observed for the homopyrimidine modified PNA:DNA complexes is summarized in Table 5**.** The N-terminus modified *dap***PNA 2** and *secdap***PNA 5** (entry 2 and 5) show stabilization towards DNA binding by 2.7°C and 3.5°C respectively compared to unmodified *aeg*PNA. The middle modification in *dap***PNA 3** and *secdap***PNA 6** (entry 3 and 6) decreases the UV melting temperature (T_m) by 13.1°C and 17.0°C respectively, indicating destabilization towards DNA binding. The C-terminus modification of the oligomer shows, stabilization in complementary DNA binding by 7.7°C for the *dap***PNA 4** (entry 4) while destabilization towards DNA binding for *secdap***PNA 7** (entry 7) by 7.8°C was seen as compared to unmodified *aeg*PNA.

In case of modified PNAs, *amp*PNA and *secamp*PNA, the N-terminus modification shows stabilization towards DNA binding by 2.8°C and 3.7°C (**PNA8** and **PNA11**, entry 8 and 11), middle modification shows destabilization by 2.7°C and 17.1°C (**PNA9** and **PNA12**, entry 9 and 12) and C-terminus modification shows stabilization by 3.7°C and 2.5°C (**PNA10** and **PNA13**, entry 10 and 13) respectively, as compared to unmodified *aeg*PNA.

3.3.4d UV-Melting studies of PNA complexes with mismatch DNA

The sequence specificity of PNA hybridization was examined through studying hybridization properties of PNA with DNA having single mismatch at middle **DNA 2**. The PNA_2:DNA complexes comprising control sequence **PNA 1** and the modified PNAs (**PNA2-PNA13**) were individually subjected for UV-melting with mismatch **DNA 2**. The melting profiles are shown in Figure 10 and the T_m data is summarised in Table-5.

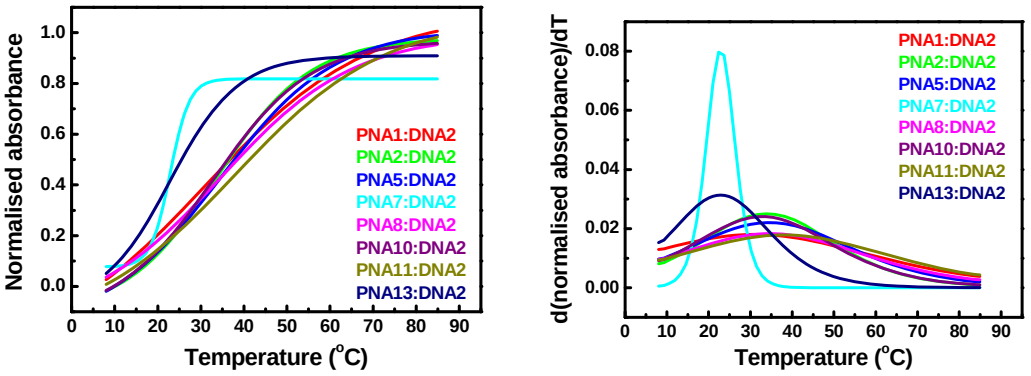


Figure 10: Melting curves for homopyrimidine PNAs (**PNA1-PNA13**) complexed with mismatch **DNA2** (CGAAAACAAAGC) (10 mM sodium phosphate buffer, pH=7.4, 10 mM NaCl). For the melts of sequences not showing any detectable transition, see appendix.

The control sequence **PNA 1** formed triplex with single mismatch sequence (**DNA 2**) with T_m lower by 10.6°C (entry 1, Table 5). PNA with modified unit **t** incorporated at N-terminus (entry 2, Table 5) showed T_m lower by 10.8°C with mismatch **DNA 2** as compared with complimentary **DNA 1**. PNA with modified unit at middle or C-terminus does not show any detectable melting with mismatch **DNA 2** (entry 3 and 4, Table 5). The N-terminus and C-terminus modification with modified unit **t**, ***t*** and ***t*** incorporated individually in the PNA sequences (entry 5, 8 &

11-N-terminus, entry 7, 10 & 13-C-terminus, Table 5) lowers the T_m by 10.5°C, 8.2°C & 7.8°C and 12.0°C, 13.0°C & 22.2°C respectively with mismatch sequence **DNA 2**, while PNAs with corresponding middle modification (entry 6, 9 &12, Table 5) does not show any detectable melting transition with **DNA 2**.

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

To examine the effect of modified PNA monomers on the corresponding duplexes, hybridization of the mixed purine-pyrimidine PNAs (Table 6, **PNA14-21**) incorporating modified thymine monomer units derived from *dap* and *amp*PNA at different positions of *aeg***PNA 14** decamer were studied.

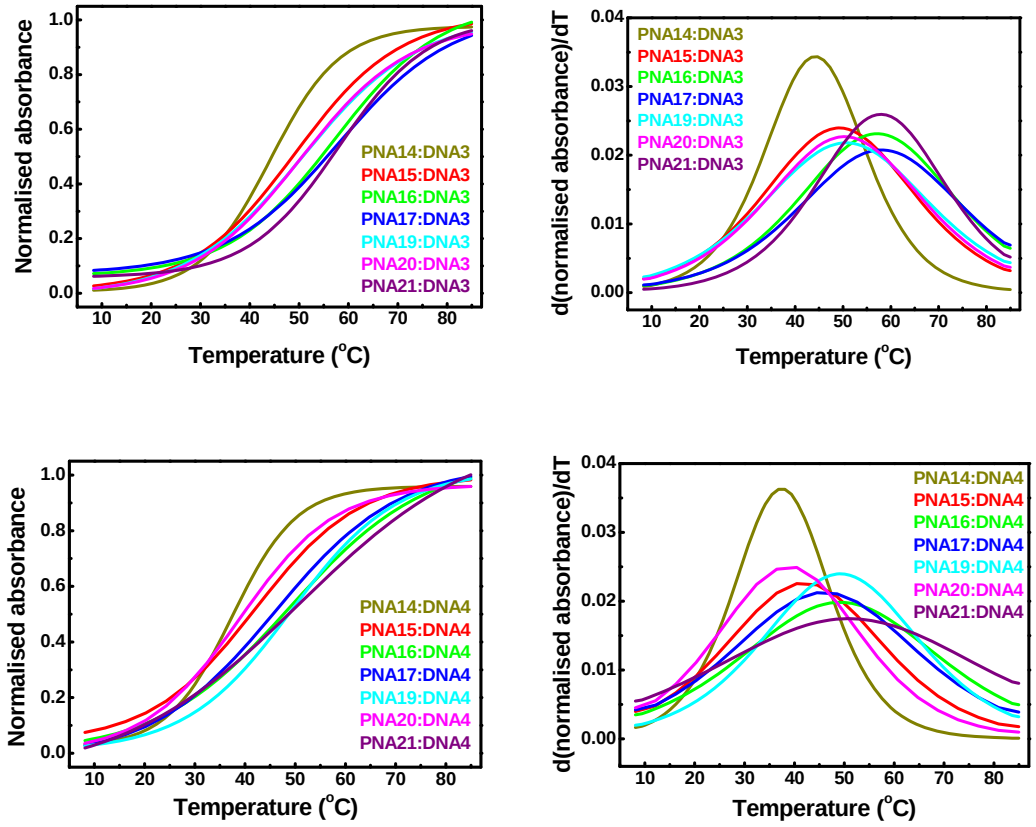
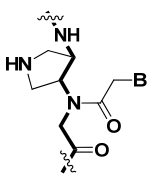


Figure 11: Melting curves for purine-pyrimidine mixed PNAs (**PNA14-PNA21**) complexed with **DNA3 (**5' AGTGATCTAC 3') and **DNA4** (5' CATCTAGTGA 3') and the corresponding first derivative curves. (10mM sodium phosphate buffer, pH=7.4, 10 mM NaCl). For the melts of PNA 18 sequences, which did not show any detectable melting transition, see appendix.

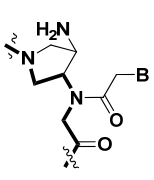
The T_m values of various PNAs hybridized with complementary DNA for parallel and antiparallel bindings were determined from temperature dependent UV absorbance and summarized in Table 6 (Figure 11). The unmodified mixed sequence (**PNA 14**, Table 6) gives a T_m of 44.7°C and 38.1°C for binding with the complementary DNA in antiparallel (with **DNA 3**) and parallel mode (with **DNA 4**). The modified *dap***PNA 15** with the N-terminus modification stabilized antiparallel and parallel mode of binding by 5.1°C and 3.4°C (entry 2, Table 6), that the middle modification *dap***PNA 16** by 12.9°C and 10.3°C (entry 3), the double modified (N-terminus and middle, *dap***PNA 17**) stabilized by13.4°C and 7.2°C respectively as compared to the control unmodified sequence **PNA 14**. The N-terminus modified *secdap***PNA 18** did not bind to complementary DNA in either antiparallel or parallel orientations..

Table-6: UV-T_m values for mixed purine-pyrimidine oligomers

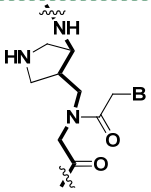
S. No.	Entry	Sequence	UV-T_m*	
			DNA 3 (*ap*) (ΔT_{m2})	DNA 4 (*p*) (ΔT_{m3})
1	PNA 14	H-GTAGATCACT-Lys-NH_2	44.7	38.1
2	PNA 15	H-GtAGATCACT-Lys-NH_2	49.8 (+5.1)	41.5 (+3.4)
3	PNA 16	H-GTAGAtCACT-Lys-NH_2	57.6 (+12.9)	48.4 (+10.3)
4	PNA 17	H-GtAGAtCACT-Lys-NH_2	58.1 (+13.4)	45.3 (+7.2)
5	PNA 18	H-G<u>t</u>AGATCACT-Lys-NH_2	nd	nd
6	PNA 19	H-G*t*AGATCACT-Lys-NH_2	51.1 (+6.4)	49.8 (+11.7)
7	PNA 20	H-G<u>*t*</u>AGATCACT-Lys-NH_2	49.7 (+5.0)	39.5 (+1.4)
8	PNA 21	H-G<u>*t*</u>AGA<u>*t*</u>CACT-Lys-NH_2	57.9 (+13.2)	51.1 (+13.0)



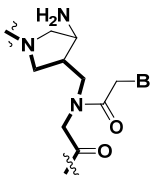
t (*dap*)



<u>t</u> (*secdap*)



t (*amp*)



<u>*t*</u> (*secamp*)

***DNA 2** = 5'-AGTGATCTAC-3'(*antiparallel*), **DNA 3** = 5'-CATCTAGTGA-3'(*parallel*). A,G,C,T = ***aeg*PNA**. (ΔT_{m2}), (ΔT_{m3}) = T_m difference w.r.t. unmodified *aeg*PNA (**PNA14**) binding in antiparallel (with **DNA 3**) or parallel (with **DNA 4**) mode respectively. nd represents no detectable melting transition.

The *amp***PNA 19** with N-terminus modification shows stabilization by 6.4^{o}C in antiparallel mode and 11.7^{o}C in parallel mode of binding (entry 6), compared to control. The *secamp***PNA 20** with modification at N-terminus shows stabilization by 5.0^{o}C and 1.4^{o}C respectively for antiparallel and parallel mode of binding while the double modified *secamp***PNA 21** (at N-terminus and middle) shows almost equal stabilization by 13.2^{o}C and 13.0^{o}C for antiparallel/parallel bonding mode as compared to control *aeg***PNA 14**The modified PNA oligomers *dap*PNA & *amp*PNA and their analogues *secdap*PNA & *secamp*PNA were synthesized by SPPS method using monomers (**13**, **25**, **26**) and employing Boc or Fmoc and mixed Boc/Fmoc strategy. The oligomers obtained after cleavage from solid support were purified by RP-HPLC and subjected for temperature dependent UV and CD spectroscopic studies.

3.4 Discussion

The temperature dependent UV absorbance for PNA_2:DNA triplexes shows a stabilization towards complementary DNA binding for N-terminus and C-terminus homopyrimidine *dap*PNA and *amp*PNA oligomers in almost all the cases (Table 5). One exception is for modified *dap***PNA 7** with C-terminus modification which shows destabilization by 7.8^{o}C compared to control *aeg*PNA. It should be noted here that one *ss*PNA modification corresponds to two modifications in a triplex because of PNA:DNA 2:1 stoichiometry in triplexes. The middle modification in homopyrimidine sequence for all the modified PNAs destabilizes the complementary DNA binding. The destabilization is maximum for *secdap*PNA and *secamp*PNA (**PNA 6** & **PNA 12**). More importantly, the specificity is retained in all these cases. This is evident from the mismatch studies, wherein the presence of a single C-T mismatch (**DNA 2**) in the centre of the DNA sequence made the *aeg*PNA_2:DNA complexes less stable compared to triplexes with complimentary **DNA 1**. The ΔT_m (difference in T_m between complimentary & mismatch) observed for modified PNA triplexes was also lower compared to *aeg*PNA_2:DNA triplexes in most of the cases. Thus higher binding is achieved without sacrificing the specificity of binding.

The mixed purine-pyrimidine PNA oligomers incorporating one of the four modified PNA units individually at different positions show stabilization of duplexes

with both antiparallel and parallel complementary DNA in all the cases. The *secdap***PNA 18** is an exception and does not show any detectable melting transition in either antiparallel or parallel mode. In all the cases the antiparellel mode of binding is preferred to the parallel mode. These results obtained may help to understand the factors which influence the binding of PNA with complementary DNA. The results indicate that the differential conformative effect of pyrrolidine ring induced on backbone upon incorporation has major effects in pre-organizing the backbone for favourable interactions with cDNA.

3.4.1 Factors affecting the binding of PNA to complementary DNA

Effect of backbone modification in *aeg*PNA

The ring constraints were introduced in the unmodified *aeg*PNA in four different ways using *dap*- and *amp*PNA monomers individually, each giving rise to two different backbones depending upon the direction of solid phase peptide synthesis as discussed earlier in this chapter. By introducing the ring constraints in the form of *dap* and *amp*PNA into the unmodified *aeg*PNA stabilizes the binding with complementary DNA in both triplex and duplex mode except when modified unit is present at the middle position in case of triplexes.

The modified *secdap*PNAs comparatively do not show good binding in both duplexes and triplexes with complementary DNA. This may be due to the presence of three members (two carbon atoms and one nitrogen atom) of six membered backbone in the ring. This does not provide enough flexibility to the backbone to bind to complementary DNA.

Effect of backbone extension by one carbon

A comparison of the binding of the *dap*PNA and one carbon extended *amp*PNA oligomers with complementary DNAs shows no significant relative difference in the stability of derived triplexes and duplexes. When the backbone is derived from secondary amino group as in *secdap*PNAs extended by one carbon atom to *secamp*PNAs, improved binding is observed in both triplexes and duplexes. For instance, no melting transition could be observed for *secdap***PNA 18** duplex with

complimentary DNA whereas the one carbon backbone extended *secamp***PNA 20** shows a T_m of 49.7°C (Table 6).

These results indicate that one carbon backbone extension is favoured when the backbone is derived from inclusion of ring nitrogen in main chain as for *secdap*PNA. This may be because extension of the backbone by one extra carbon gives the backbone enough flexibility required to adopt favourable conformation for binding with complementary DNA. This is very much restricted and rigid in the modified *secdap*PNA because of ring intervening in the backbone.

Electrostatic influence in binding

It is reported in the literature that introduction of positive charges in the PNA backbone improves its binding.[26] The cationic PNAs are expected to possess superior ability to strand invade complementary DNA sequences which have negative charges. The positive charge on PNA arises from the introduction of amino/imino functions in PNA which accepts a proton. The PNA sequences incorporating modified *amp* and *dap* units (Table-1) have one free amino group (primary amino group in *secdap*PNA & *secamp*PNA and secondary ring amino group in *dap*PNA & *amp*PNA) which can get protonated making the backbone cationic. The UV-melting stability studies on these PNA oligomers show better binding with complementary DNA as compared to the uncharged control *aeg*PNA. In addition, the binding of cationic *dap*PNA with cDNA (**PNA 16**, Table 6) shows melting temperature of 57.6°C, while with same stereochemistry of the ring, uncharged cyclopentyl PNA (*cp*PNA) introduced into the same PNA sequence at same position (middle) did not show any binding, as reported in literature.[25] These results clearly indicate that the electrostatic attraction between positively charged PNA and negatively charged DNA is also a factor stabilizing the PNA:DNA complexes, consistent with the literature findings.[26]

Effect of ring puckering

The saturated five-membered rings can interconvert their conformations passing through different envelope and twist forms (Figure 12). The barriers for interconversion between the envelopes or half-chairs are extremely slow, the molecules being in a state of "conformational flux".

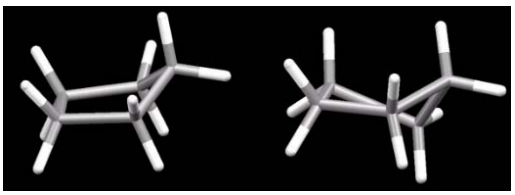


Figure 12: Envelope (left) and Twist forms (right) of five membered ring.

The distortion made by the out-of-plane atom (atoms) appears to be in a motion around the ring giving rise to the term "ring pseudorotation". During such full cycle of pseudorotation (360°), there are a total of 10 different envelopes and 10 different half-chairs (Figure 13). Pseudorotational barriers for cyclopentane are calculated to be negligible, while for pyrrolidine it is about 1.2 Kcal/mol[27] This shows 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.

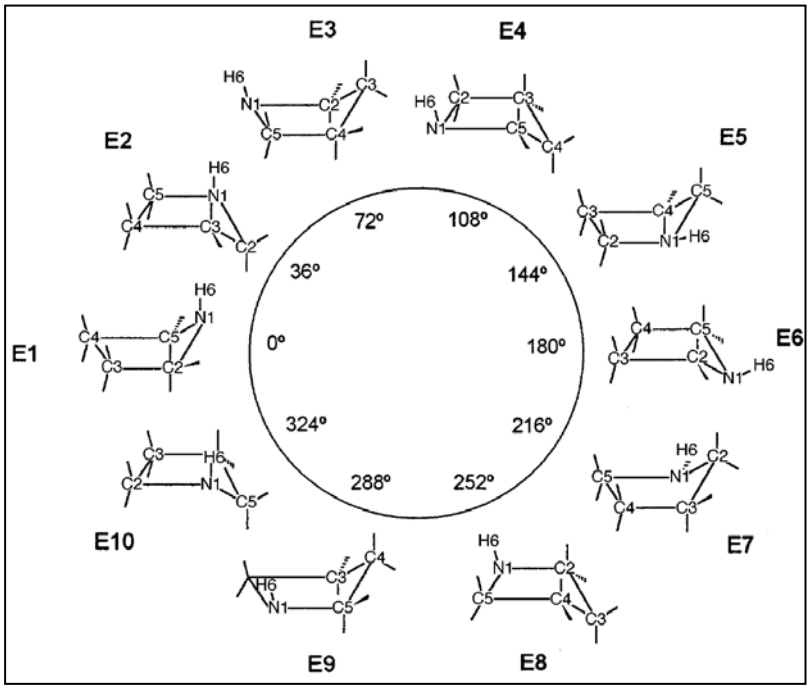


Figure 13: Envelope conformations of pyrrolidine along the pseudorotational path. The twist conformations are located between adjacent envelope forms.[27]

Introduction of a single unit of (*R*,*R*) *trans*-cyclopentyl *tcyp*T, which was reported earlier by Appella et al,[25a] placed in the middle of a purine-pyrimidine PNA decamer showed no observable melting transition. This was attributed to the

stereochemistry of the the cyclopentyl ring which does not favourably pre-organise the backbone for binding to complementary DNA as its (*S*,*S*) *trans* analogue showed 6.0°C increase in T_m than control *aeg*PNA. In contrast, both (*R*,*S*) and (*S*,*R*) *cis*-cyclopentyl PNAs[25b] bind very strongly to *c*DNA/RNA with an increase of T_m around + 7.5°C per modification.

In present work, It is seen that by introducing a N-atom in the cyclopentyl ring (*R*,*R*) *dap*PNA oligomer shows a melting temperature of 57.6 ^oC (**PNA 16**, Table 6) for the same PNA sequence. This shows that not only the stereochemistry but type of the ring pucker also has markable effect on the binding.

It is well documented in the literature that the conformations of the five-membered pyrrolidine rings in proline and 4-hydroxy proline are important in controlling the stability and physiological functions of collagen fibrils. The electronegativity of the substituents attached to the ring influence the stability of such complexes by changing the puckering of the ring (Figure 14). For instance, an electronegative substituent at the 4*R* position of proline has a strong preference to pucker the ring from Cγ-endo pucker in proline to Cγ-exo pucker.

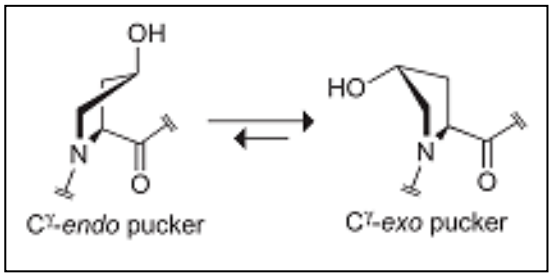


Figure 14: Change in puckering of proline ring by substitution

In *aep*PNA reported earlier from this laboratory, the individual purines or pyrimidines that differ in their group electronegativities, when present at the 4-position of the pyrrolidine ring cause differential ring pucker effects (Nagendra Sharma, Phd Thesis, University of Pune, 2005). This may lead to backbone conformational changes causing sequence specific effects.

In (*R*,*R*) *dap*PNA, proline ring is part of backbone through a common C-C bond while in *secdap*PNA, the proline ring is main part of the backbone. Thus the effect of puckering on the backbone is different in these two cases, which finally influence the stability of derived PNA:DNA complexes.

The better binding in (*R*,*R*) *dap*PNA as compared to (*R*,*R*) trans cyclopentyl 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. This may explain the better stabilization of *dap*PNA as compared to its analogue *secdap*PNA. The involvement of the pyrrolidine ring as the side chain to PNA backbone in *dap*PNA is favoured over the intervening of the ring in the backbone of *secdap*PNA, for better stability of PNA:DNA complexes.

Effect of position of modification

The position of the modified units derived from *dap* and *amp* PNA monomers present in the PNA sequence also has markable effect on the binding of the PNA:DNA complexes. In PNA_2:DNA triplexes, the homothymine octamer with modified unit present at terminus positions (C or N) showed higher melting than the modified unit present at the middle of the sequence (Table-5). However for duplex with the *dap*PNA, the modified unit present at middle is 7.8 ^{o}C more stabilised than when present at N-terminus (entry 2 and 3, Table-6). This may be probably because the triplex involves two PNA strands in PNA_2:DNA complex so one ssPNA modification corresponds to two modifications in triplex.

Effect of primary or secondary amino group

The modified monomers **13** and **26** were designed so as to synthesize two different backbones through both primary and secondary amine, using Boc/Fmoc strategy in solid phase, from a single modified PNA monomer (Figure-6). The UV-melting results on the sequences incorporating the backbones utilizing primary amine group of modified units (in case of *dap* and *amp*PNA) or the secondary amine group of modified units (in case of *secdap* and *secamp*PNA) shows that the backbones utilizing primary amine group have higher melting than their secondary amine analogue (Table 5 & 6). This may be because of more restraints in the backbone in the form of tertiary amide linkage leading to destability in case of backbone utilizing secondary amine group of the modified units as compared to primary amine group counterpart.

3.5 Conclusion

Both *dap* and *amp* monomer building blocks **13** and **26** synthesized in Chapter-2 were used to assemble four different PNA oligomers *dap*, *secdap*, *amp* & *secamp*PNA (each monomer give rise to two modified units). The PNA oligomers containing one or two modified thymine PNA units were subjected to hybridization studies. The triplexes and duplexes formed by *dap*PNA and *amp*PNA were better stabilized than their analogues *secdap*PNA and *secamp*PNA. This differentiation in stability of PNA:DNA complexes, for two different PNA units derived from a single monomer building block, 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 preceding chapter constitutes a new family of PNA monomers which can tune the binding properties of *aeg*PNAs. This can be achieved either incorporating the monomer unit through pyrrolidine ring substitutions in the backbone as in *dap* or *amp*PNA or through ring intervening as in *secdap* and *secamp*PNA. The former has stabilizing effect while the later has destabilizing effect on PNA:DNA complexes, thus help to adjust the binding properties. Further work is required to tune the stability of PNA:DNA complexes by incorporating both the modified units, derived from a single monomer, in the same PNA sequence

3.6 Experimental

3.6.1 Solid phase peptide synthesis

Functionalization of resin

Commercially available MBHA resin or Rink amide resin (Jupiter Biosciences, West Marredpally, Hyderabad, India) with loading value of 2.0 m eq/g and 0.7 m eq/g respectively, is not suitable for oligomer synthesis 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 dried resin was taken in solid phase funnel and swelled in DCM for about 1h and functionalized with N^{α} boc- N^{ω} Z-Lysine.

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 CH_2Cl_2 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 CH_2Cl_2 to remove the unbound picric acid. The picrate bound to amino groups was eluted with 5% diisopropyl ethylamine in CH_2Cl_2, followed by washing with CH_2Cl_2. The elutant was collected into a 10 ml volumetric flask and made up to 10 ml using CH_2Cl_2. 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 λ_{max}), and the loading value of the resin (0.35 meq/g) was calculated

using the molar extinction coefficient of picric acid as ε_{358}=14,500 cm^{-1} M^{-1} at 358 nm.

Kaiser's Test

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

Cleavage of the PNA oligomers from the solid support

For MBHA resin: The resin bound PNA oligomer (5 mg) was kept in an ice-bath with thioanisole (10 μL) and 1,2-ethanedithiol (4 μL) for 10min, TFA (60 μL) was added and shaken manually and kept for another 10 min. TFMSA (8 μL) 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.

Purification of PNA oligomers

The purity of the so obtained PNA oligomers was checked by analytical reverse phase HPLC on C18 column using a gradient of 0 to 100% CH3CN 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.

MALDI-TOF Mass Spectrometry

Literature reports the analysis of PNA purity by MALDI-TOF mass spectrometry in which several matrices have been explored, *viz*. sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid),[28] CHCA (α-cyano-4-hydroxycinnamic acid)[29] and DHB (2,5-dihydroxybenzoic acid). Out of these, CHCA was found to give the best signal to noise ratio. For all the MALDI-TOF spectra recorded for the modified PNAs reported in this chapter, CHCA was used as the matrix and was found to give satisfactory results.

3.6.2 UV-T_m studies

Sample Preparation

The PNA oligomers and the appropriate DNA oligomers were mixed together in stoichiometric amounts (2:1 PNA:DNA for oligothymine-T8 PNAs or 1:1 for the duplex forming PNAs, *viz*., the mixed base sequences) in 10mM sodium phosphate buffer, 10mM NaCl, pH 7.4 to achieve a final strand concentration of either 0.5 or 1mM each strand. Extinction coefficient C = 6.6, T = 8.6, A = 13.7 and G = 11.7 [(mmol)-1 cm-1] was used to calculate the concentration of PNA by following Lambert Beer's Law: $A = \varepsilon cl$.

The antiparallel and parallel complexes involving the PNAs containing all the four nucleobases were constituted using corresponding DNA. The samples were heated at 85 ^{o}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.

UV-Jobs plot

To a solution of DNA in 0.01M sodium phosphate at pH 7.4, were added portions of the complementary PNA oligomer to make 10 different fraction with mole ratio from 0-100% with fixed concentration. Then UV of each fraction was scanned at the temperature of the circulating water was maintained at 10 ^{o}C (i. e.,

well below the melting temperature of the complexes) and the absorbance at 260 nm was recorded. This was plotted as a function of the mole fraction of PNA.

UV melting

UV melting experiments were performed on Cary 300 Bio UV-Visible Spectrophotometer equipped with a thermal melt system. The sample for T_m measurement was prepared by mixing calculated amount of stock oligonucleotide and PNA solutions together in 1 mL of 0.01 M sodium phosphate buffer (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^{o}C with temperature increment of 0.5^{o}C/min. Nitrogen gas was purged through the cuvette chamber below 20 ^{o}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 ±1^{o}C.

3.6.3 Circular dichorism (CD)

CD spectra were recorded on a Jasco J-715 spectropolarimeter. The CD spectra of the PNA:DNA complexes and the relevant single strands were recorded in 0.01M sodium phosphate buffer, pH 7.4. The temperature of the circulating water was kept below the melting temperature of the PNA:DNA complexes, i. e., at 10 ^{o}C.

The CD spectra of the oligothymine T_8 single strands and mixed sequence PNAs and their complexes with DNA were recorded with accumulation of 5 scans from 320 to 195 nm using a 1 cm cell, a resolution of 0.1 nm, band-width of 1.0 nm, sensitivity of 2 mdegrees, response 2 sec and a scan speed of 50 nm/min. The PNA:DNA complexes were constituted by mixing appropriate strands in a 2:1 stoichiometry for triplexes and 1:1 stoichiometry for duplexes in buffer followed by heating to 90 ^{o}C and annealed by slow cooling to 4 ^{o}C.

3.7 References

1. Milligan, J.; Matteucci, M. D.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 1921.
2. Woolf, T.; Jennings, C.; Rebagliati, M.; Melton, D. *Nucleic Acid Res.* **1990**, *18*, 1763.
3. Merrifield, R. B. *J. Am. Chem. Soc*. **1963**, *85*, 2149.
4. (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.
5. Carpino, L.A.; Han, G.Y. *J. Org. Chem.* **1972**, *37*, 3404.
6. Matsueda, G. R.; Stewart, J. M. *Peptides* **1981**, *2*, 45.
7. Rink, H. *Tet. Lett*. **1987**, *28*, 3787.
8. Bernhardt, A.; Drewello, M.; Schutkowski, M. *J. Peptide Res.* **1997**, *50*,143.
9. Sheehan, J. C.; Hess, G. P. *J. Am. Chem. Soc*. **1955**, *77*, 1067.
10. (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.
11. Marder O.; Shvo Y.; Albericio F. *Chim. Oggi* **2002**, *20*, 37.
12. Coste, J.; Dufour, M. N.; Pantaloni, A.; Castro B. *Tet. Lett* **1990**, *31*, 669.
13. Hudson, D. *J. Org. Chem* **1988**, *53*, 617.
14. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.
15. Oded, A.; Houghten, R. A. (1990) *Pept. Res* **1990**, *3*, 42.
16. Fronenot, J. D. *et. al. Pept. Res.* **1991**, *4*, 19.
17. Vojkovsky, T. *Pep. Res*. **1995**, *3*, 42.
18. Blackburn, C. *Tet Lett*. **2005**, *46*, 1405.
19. a.) Merrifield, R. B.; Stewart, J. M.; Jernberg, N. *Anal. Chem.* **1966**, *38*, 1905. b.) Kaiser, E.; Bossinger, C. D.; Cplescott, 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.
20. 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.
21. Cantor, C. R.; Warshaw , M. W.; Shapiro, H. *Biopolymers*, **1970**, *9*, 1059.

22. Soyfer, V. N.; Potaman, V. N. *Triple helical Nucleic Acids*; Springer: New York, 1996.

23. Ganesh, K. N.; Kumar, V. A.; Barawkar, K. A. *Supramolecular control of structure and reactivity* (Chapter 6, Ed Hamilton, A. D) **1996**, 263.

24. Kim, S. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

25. 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.

26. 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.

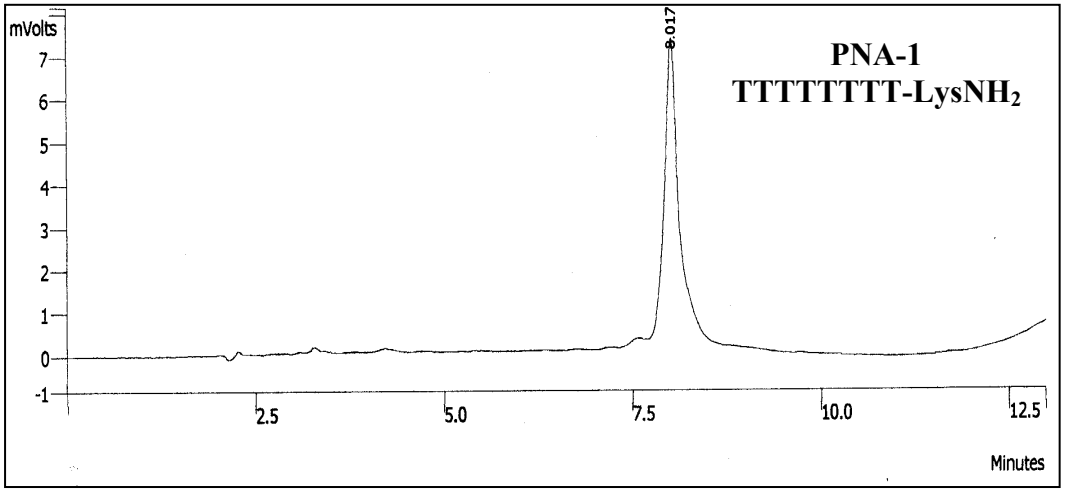
27. a.) Pitzer, K. S.; Donath, W. E. *J. Am. Chem. Soc.* **1959**, *81*, 3213. b.) Carballeira, L.; Perez-Juste, I. *J. chem. Soc., Perkin Trans 2*, **1998**, 1339.

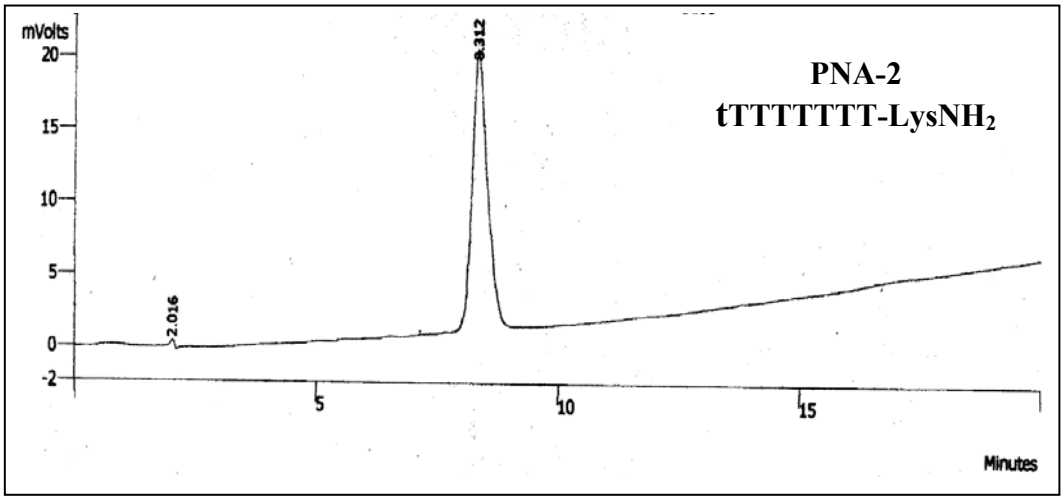
28. Beavis R. C.; Chait B. T. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 436.

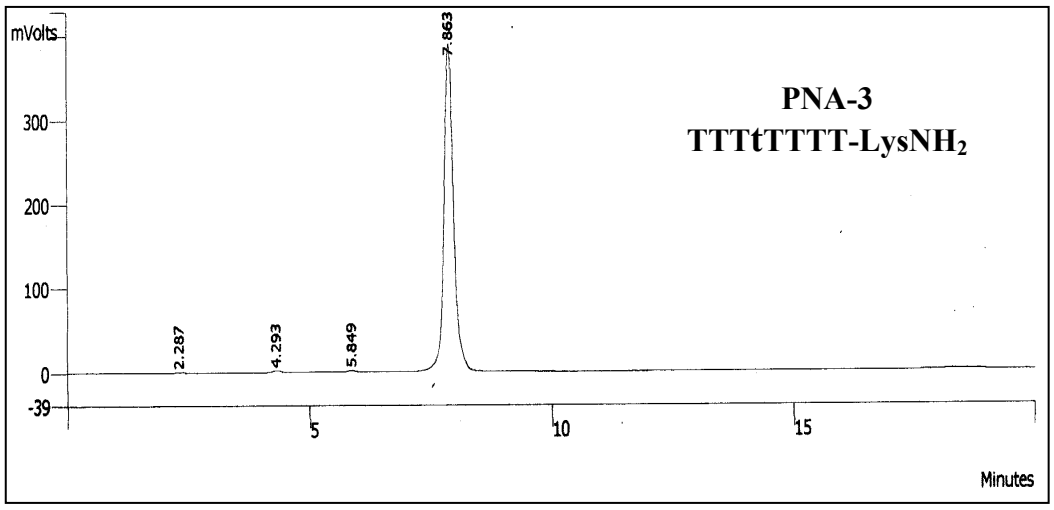
29. Beavis, R. C. *Org. Mass Spectrom.* **1992**, *27*, 156.

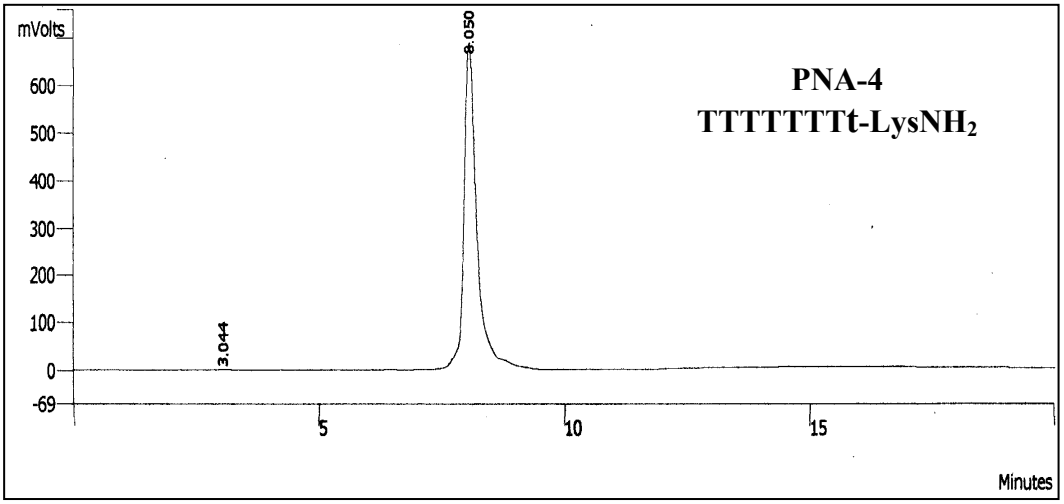
3.8 Appendix

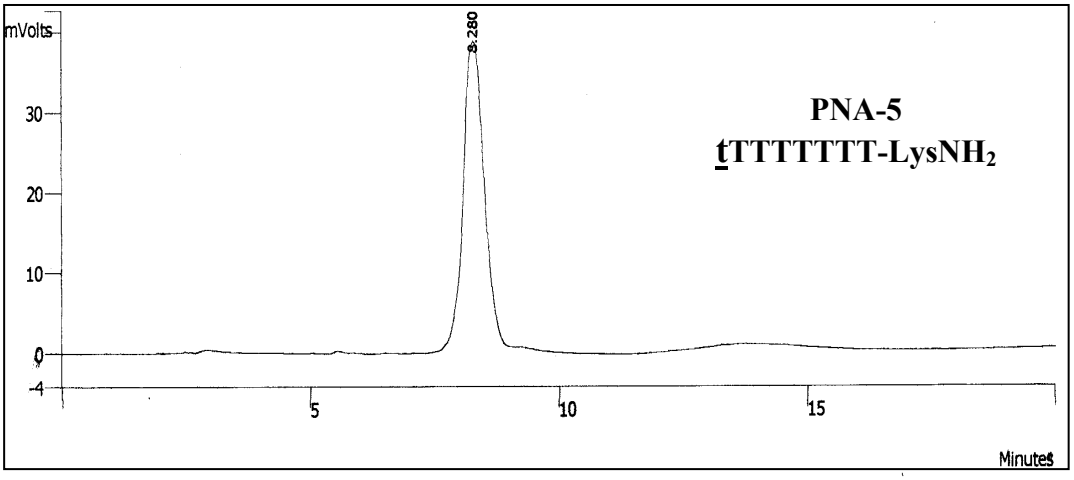
HPLC data for PNA sequences

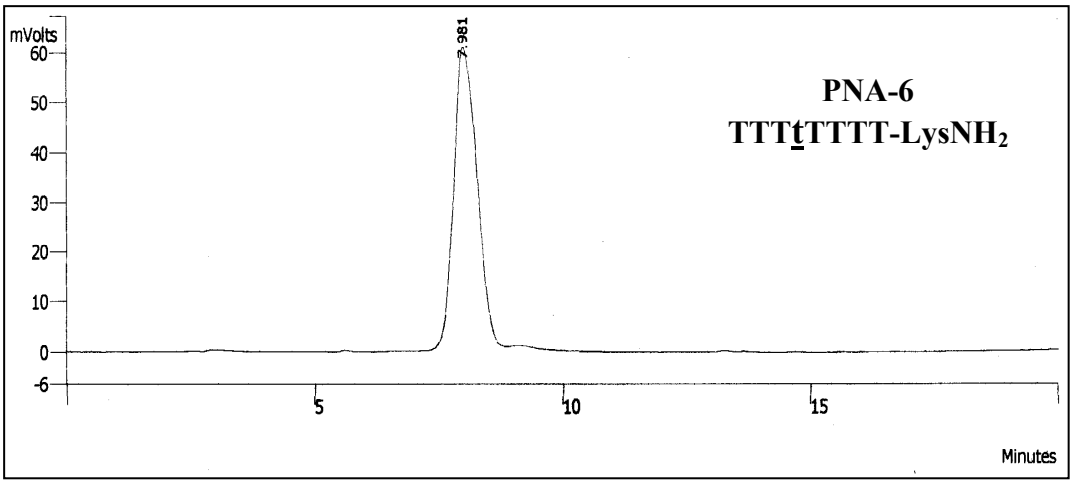


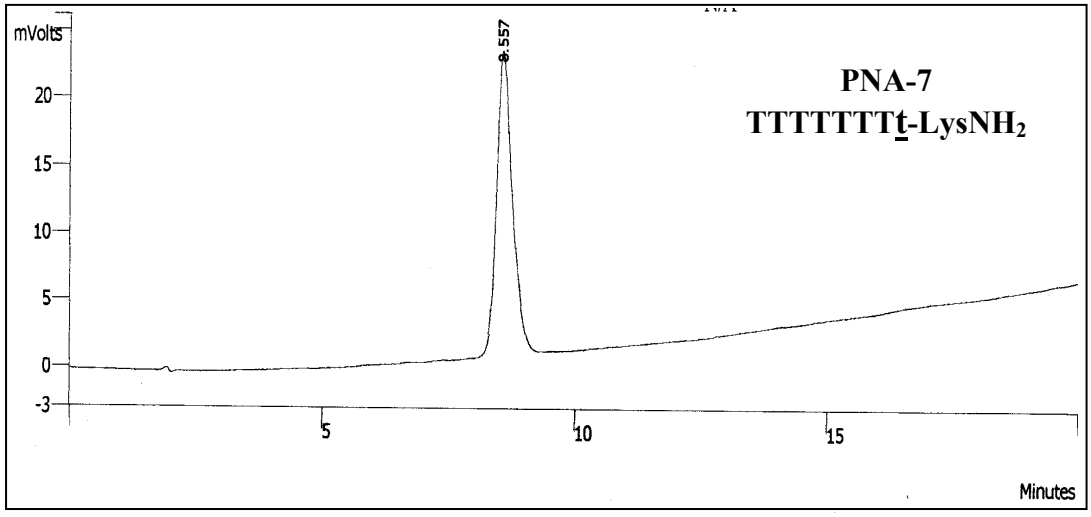


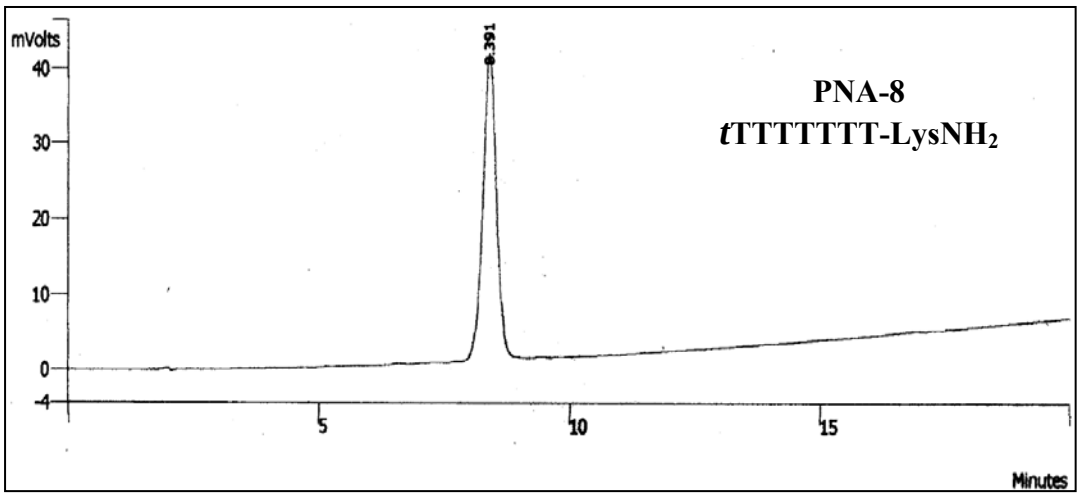


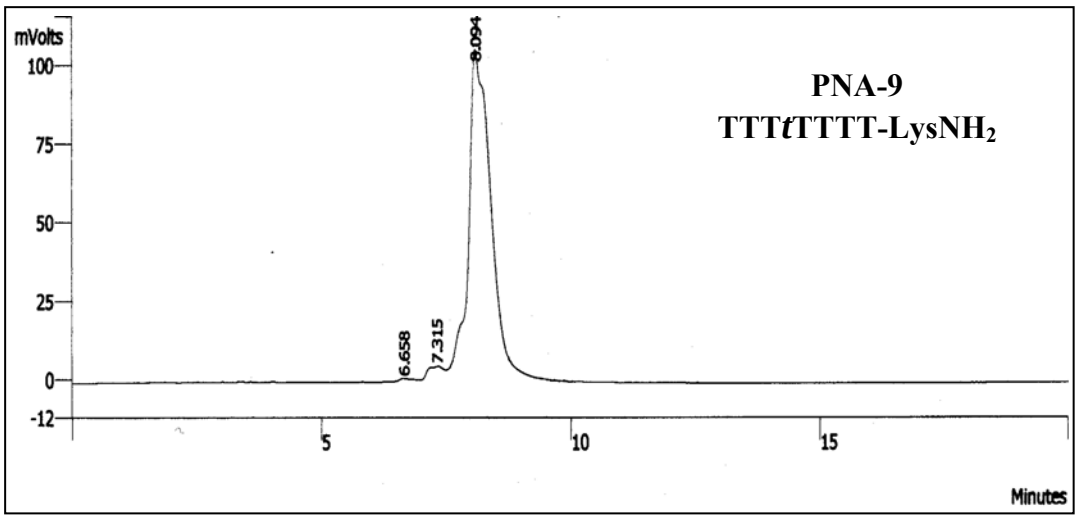

mVolts
600
500
400
300
200
100
0
-69
3.044
8.050
5
10
15
Minutes
PNA-4
TTTTTTTt-LysNH2

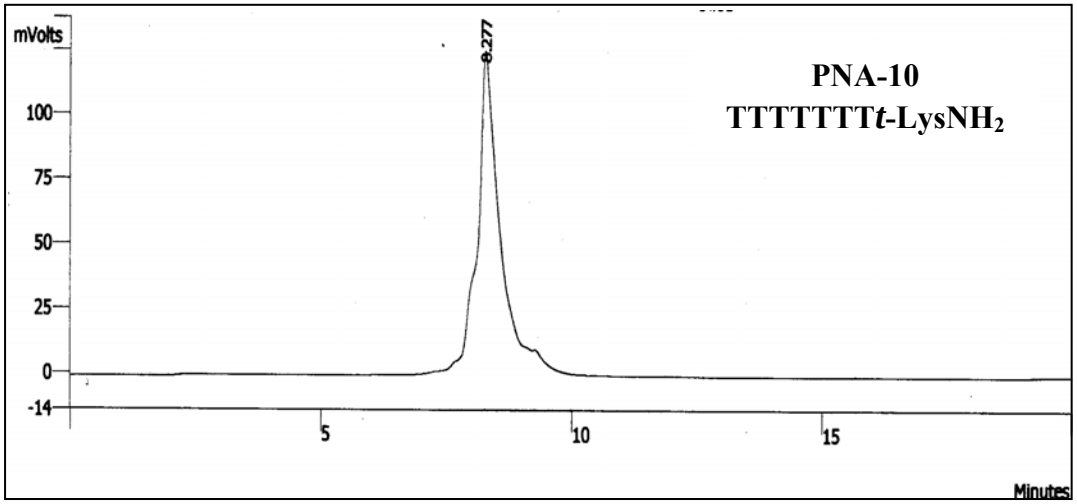

mVolts
30
20
10
0
-4
8.280
5
10
15
Minutes
PNA-5
tTTTTTTT-LysNH2

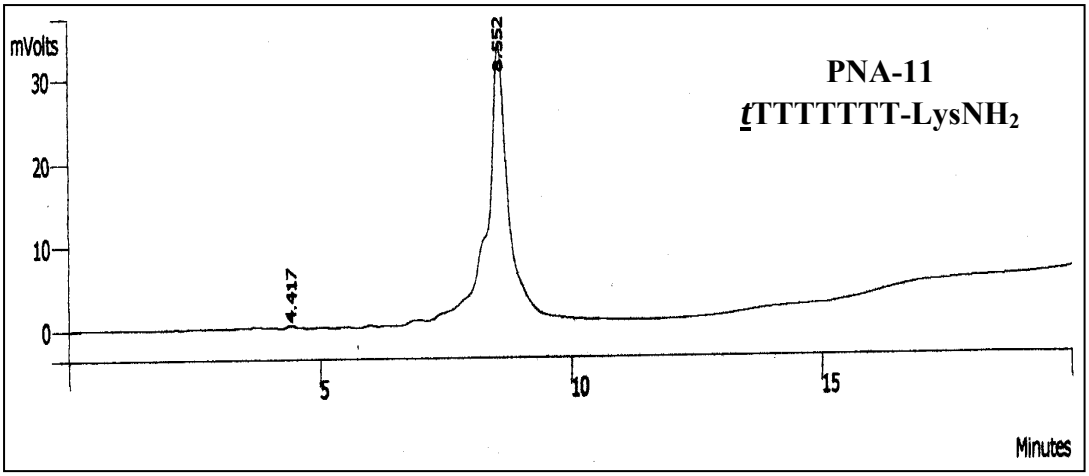

mVolts
60
50
40
30
20
10
0
-6
7.981
5
10
15
Minutes
PNA-6
TTTtTTTT-LysNH2

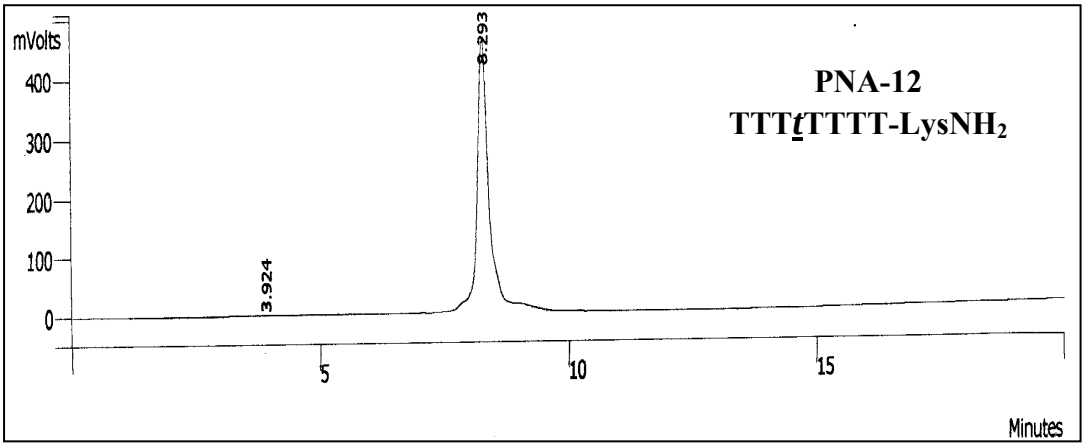

mVolts
20
15
10
5
0
-3
8.557
5
10
15
Minutes
PNA-7
TTTTTTTt-LysNH2

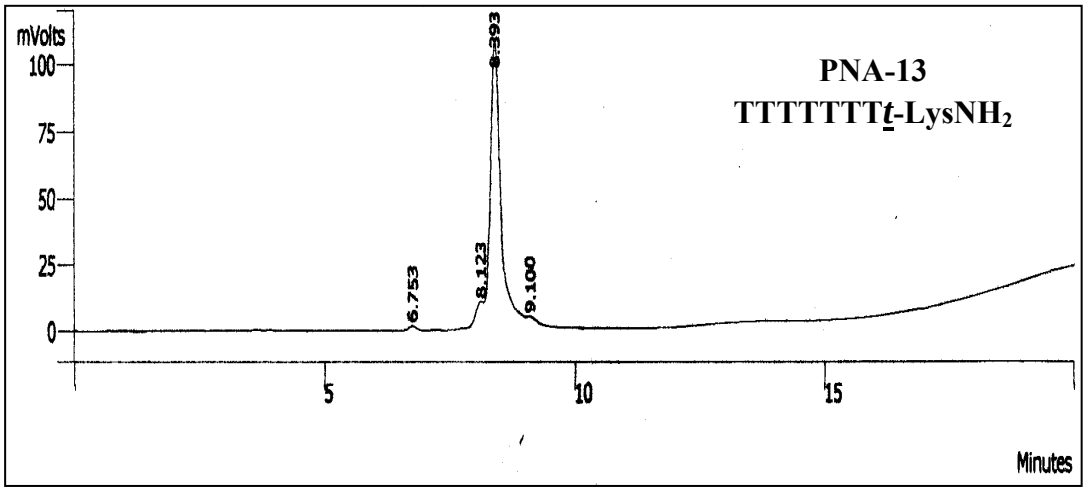

mVolts
40
30
20
10
0
-4
8.391
5
10
15
Minutes
PNA-8
tTTTTTTT-LysNH2

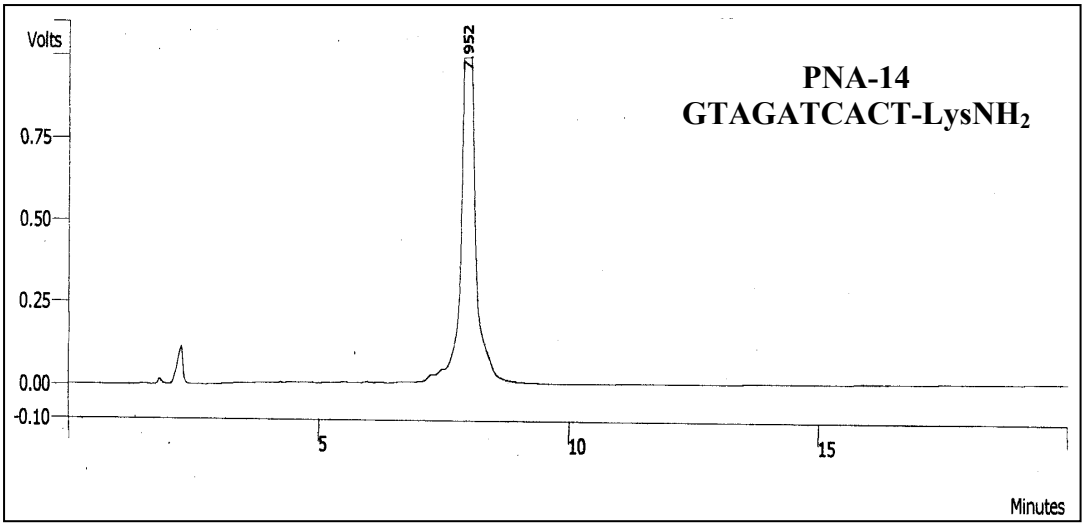

mVolts
100
75
50
25
0
-12
8.094
6.658
7.315
5
10
15
Minutes
PNA-9
TTTtTTTT-LysNH2

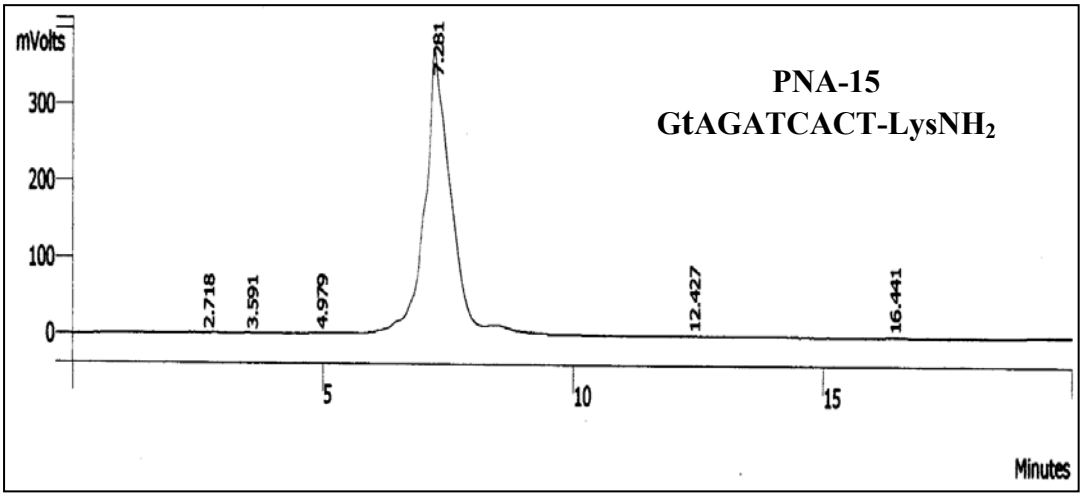

mVolts
100
75
50
25
0
-14
8.277
5
10
15
Minutes
PNA-10
TTTTTTTt-LysNH2

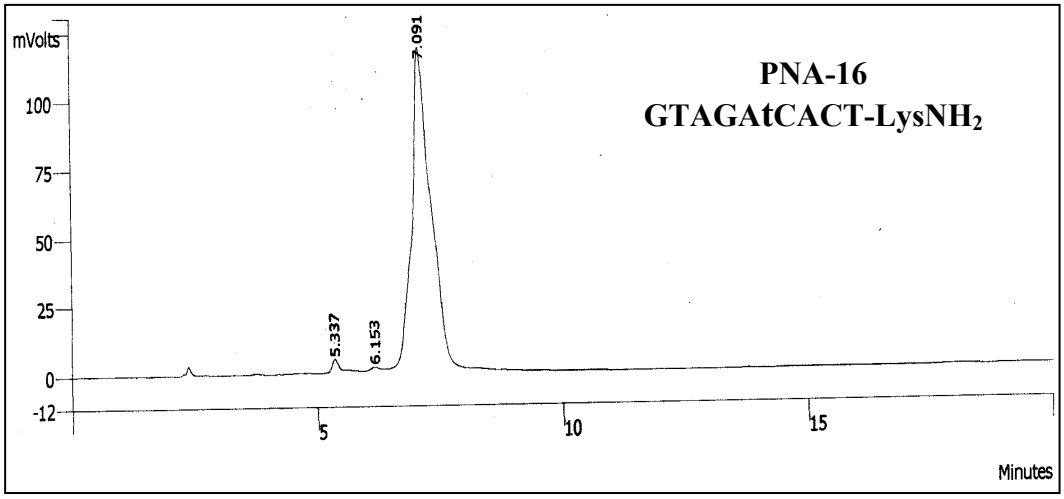

mVolts
30
20
10
0
4.417
8.552
5
10
15
Minutes
PNA-11
tTTTTTTT-LysNH2

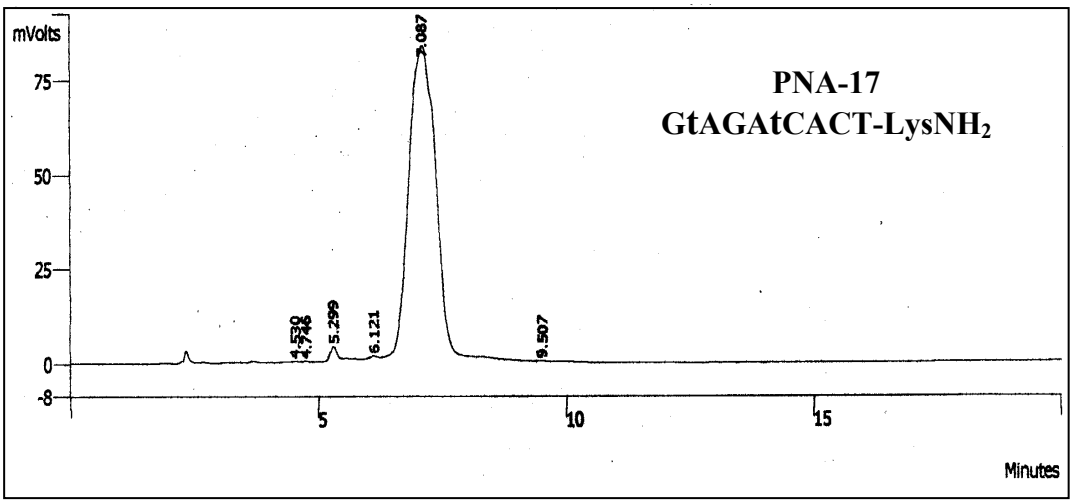

mVolts
400
300
200
100
0
3.924
8.293
5
10
15
Minutes
PNA-12
TTTtTTTT-LysNH2

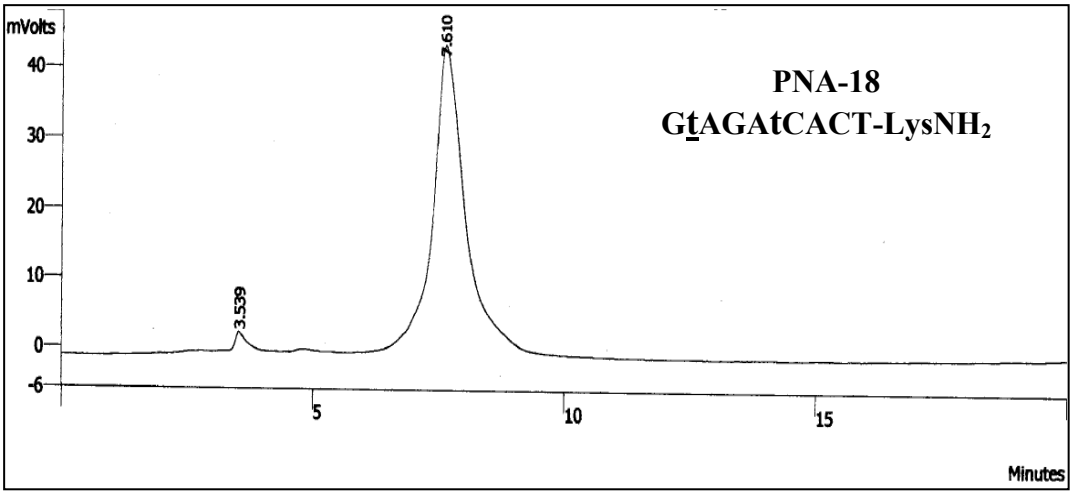

mVolts
100
75
50
25
0
6.753
8.123
8.393
9.100
5
10
15
Minutes
PNA-13
TTTTTTTt-LysNH2

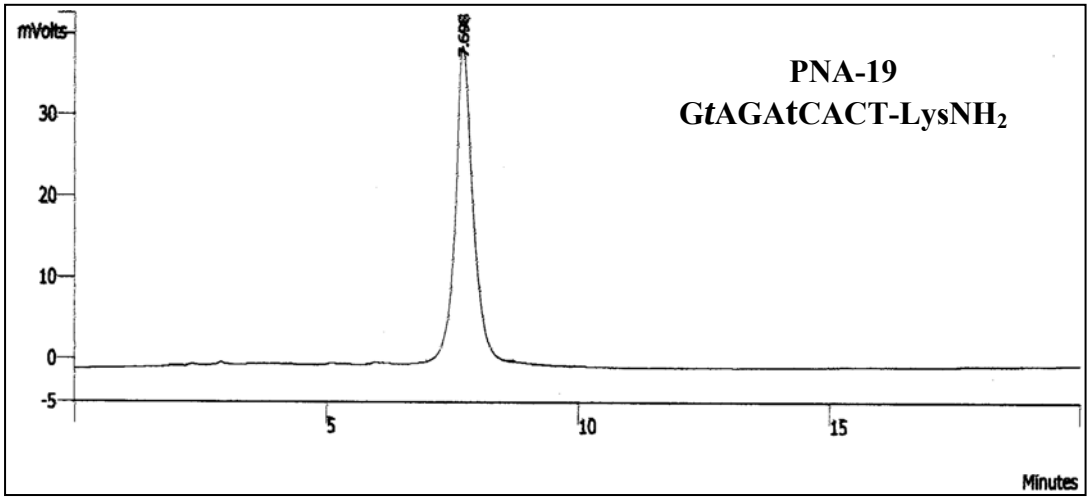

Volts
0.75
0.50
0.25
0.00
-0.10
7.952
5
10
15
Minutes
PNA-14
GTAGATCACT-LysNH2

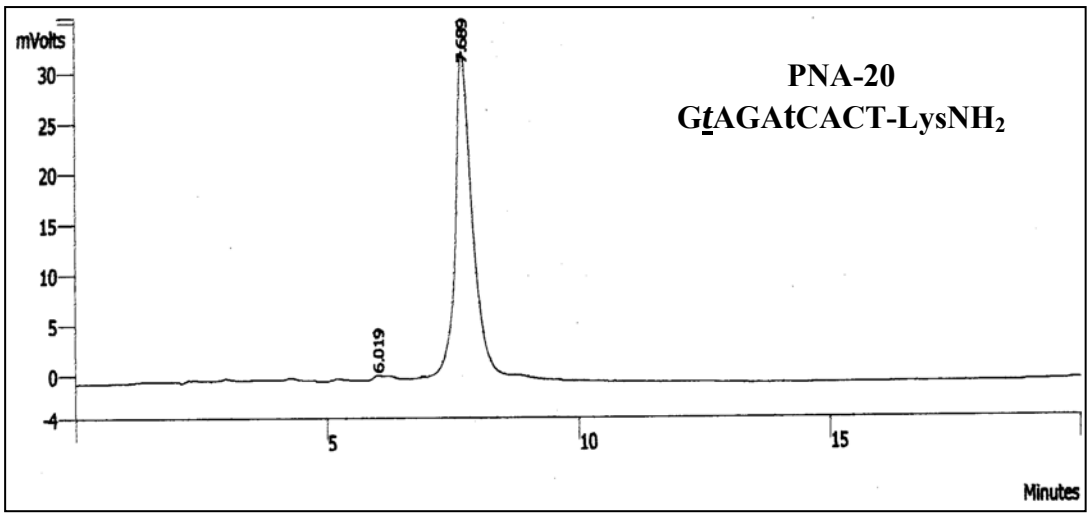

mVolts
300
200
100
0
2.718
3.591
4.979
7.281
12.427
16.441
5
10
15
Minutes
PNA-15
GtAGATCACT-LysNH2

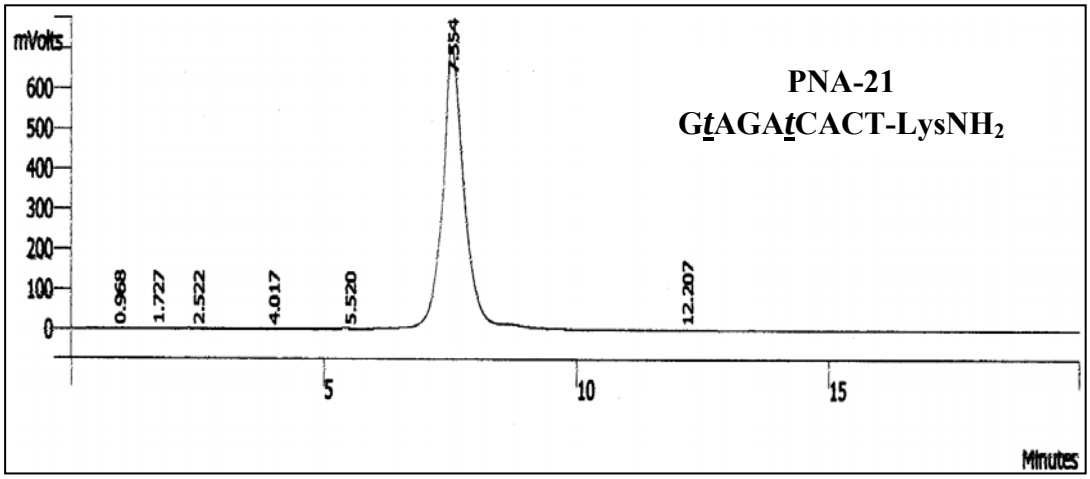

mVolts
100
75
50
25
0
-12
7.091
5.337
6.153
5
10
15
Minutes
PNA-16
GTAGAtCACT-LysNH2


mVolts
75
50
25
0
-8
7.087
4.530
4.746
5.299
6.121
9.507
5
10
15
Minutes
PNA-17
GtAGAtCACT-LysNH2

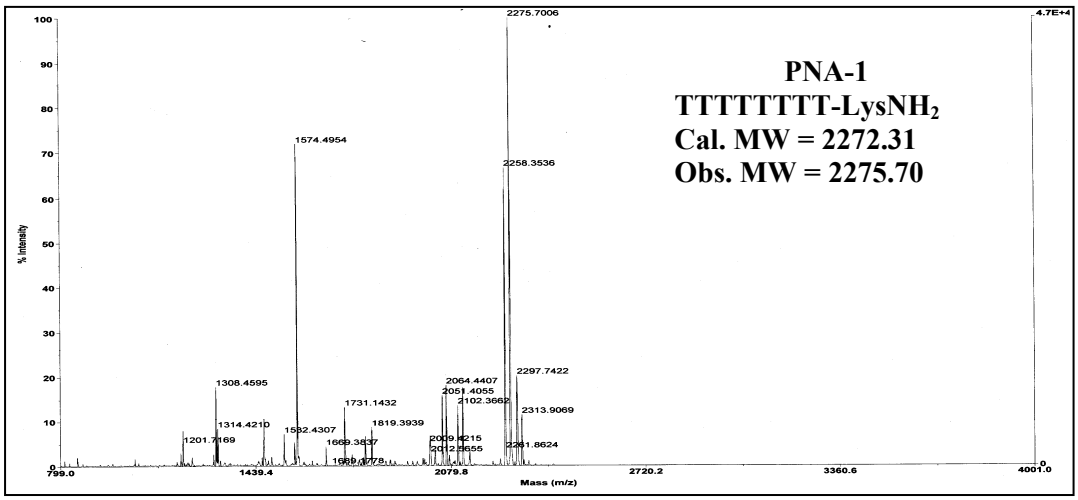

mVolts
40
30
20
10
0
-6
7.610
3.539
5
10
15
Minutes
PNA-18
GtAGAtCACT-LysNH2

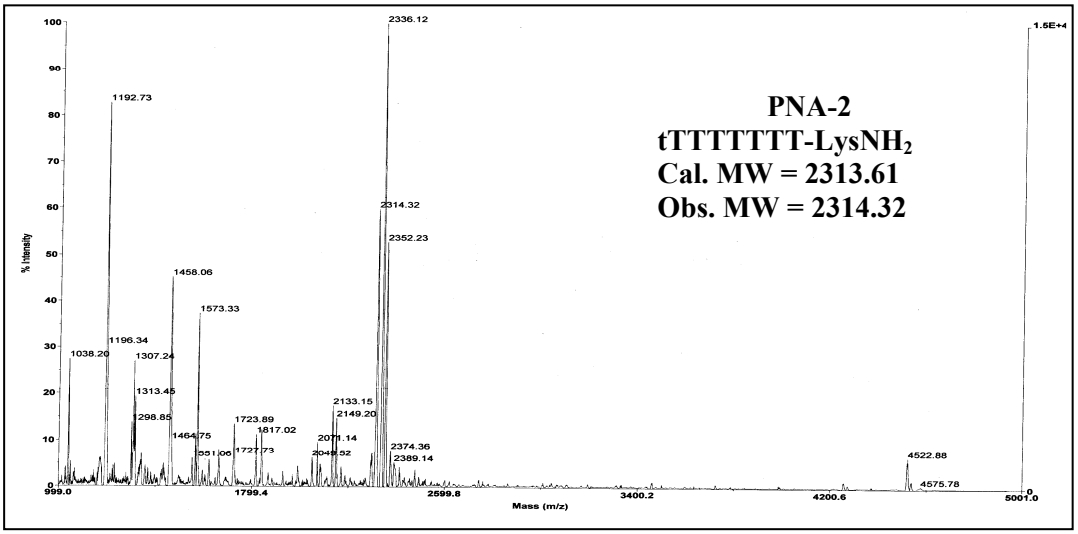

mVolts
30
20
10
0
-5
7.698
5
10
15
Minutes
PNA-19
GtAGAtCACT-LysNH2

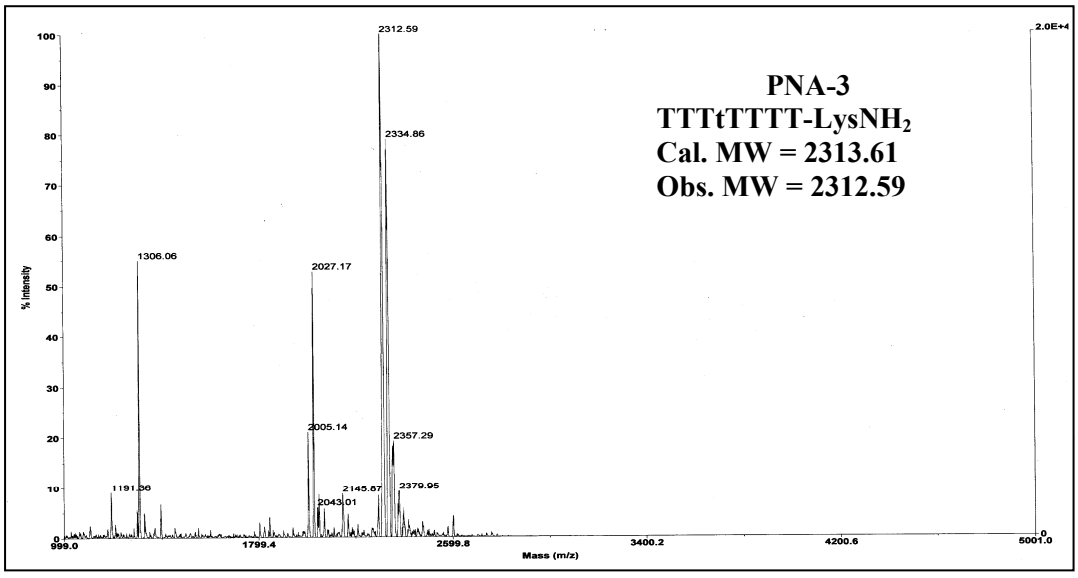

mVolts
30
25
20
15
10
5
0
-4
7.689
6.019
5
10
15
Minutes
PNA-20
GtAGAtCACT-LysNH2

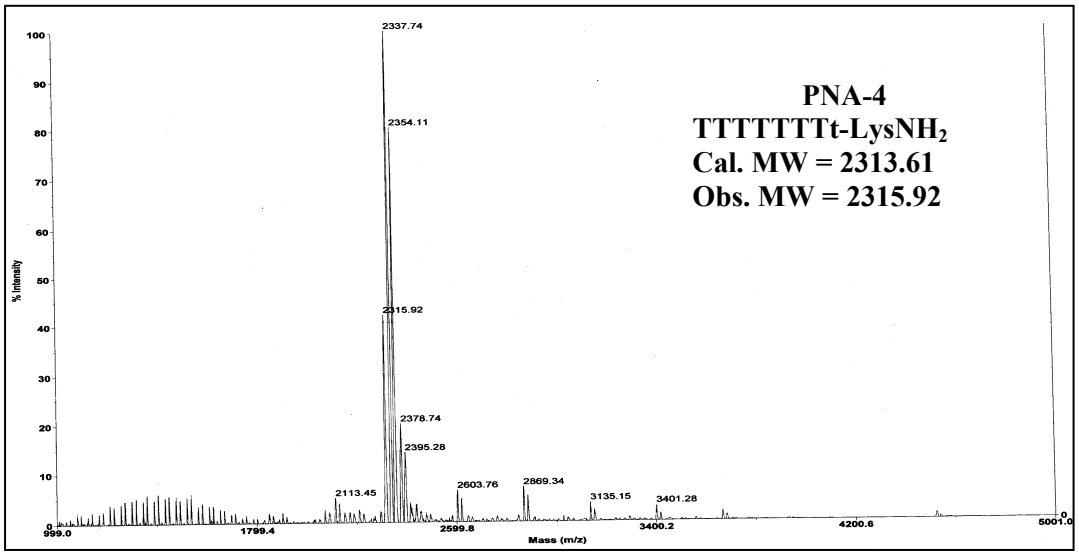

mVolts
600
500
400
300
200
100
0
7.554
0.968
1.727
2.522
4.017
5.520
12.207
5
10
15
Minutes
PNA-21
GtAGAtCACT-LysNH2

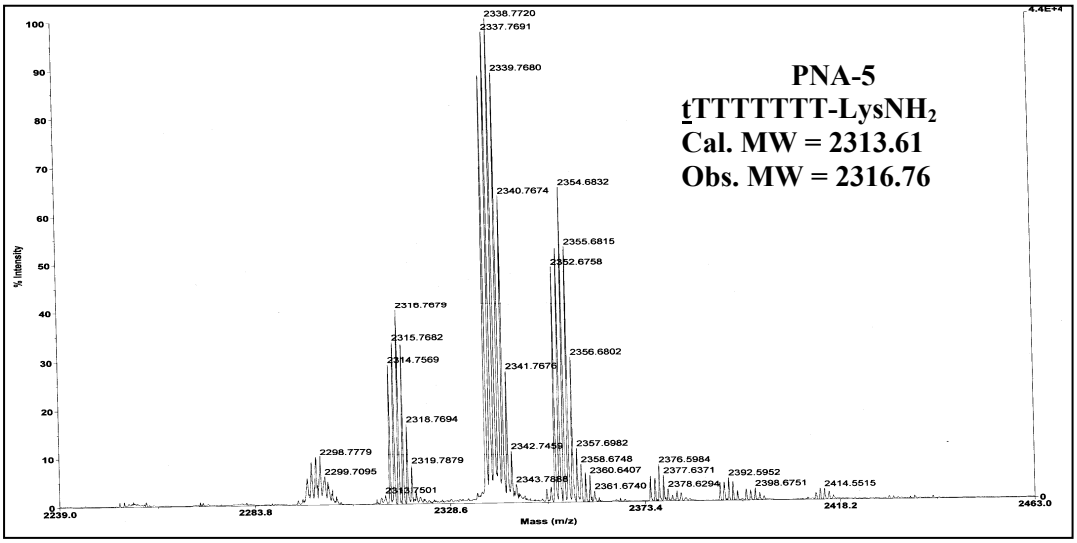
MALDI TOF-spectras for PNA sequences

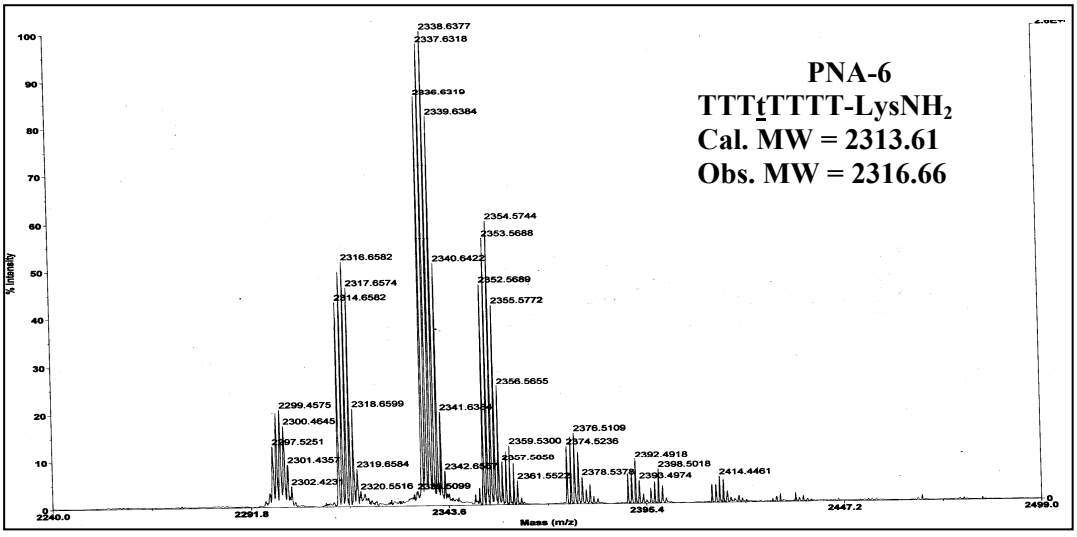


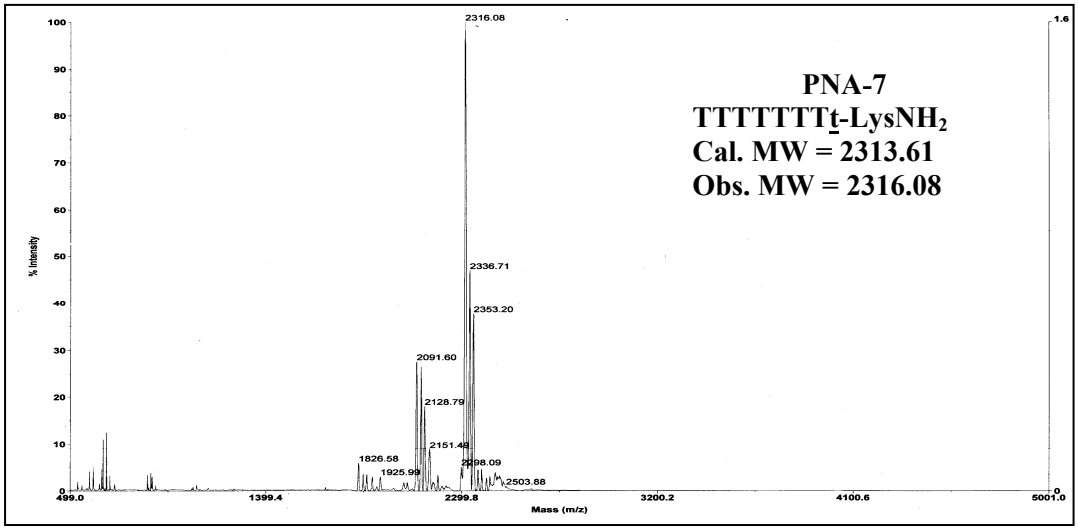


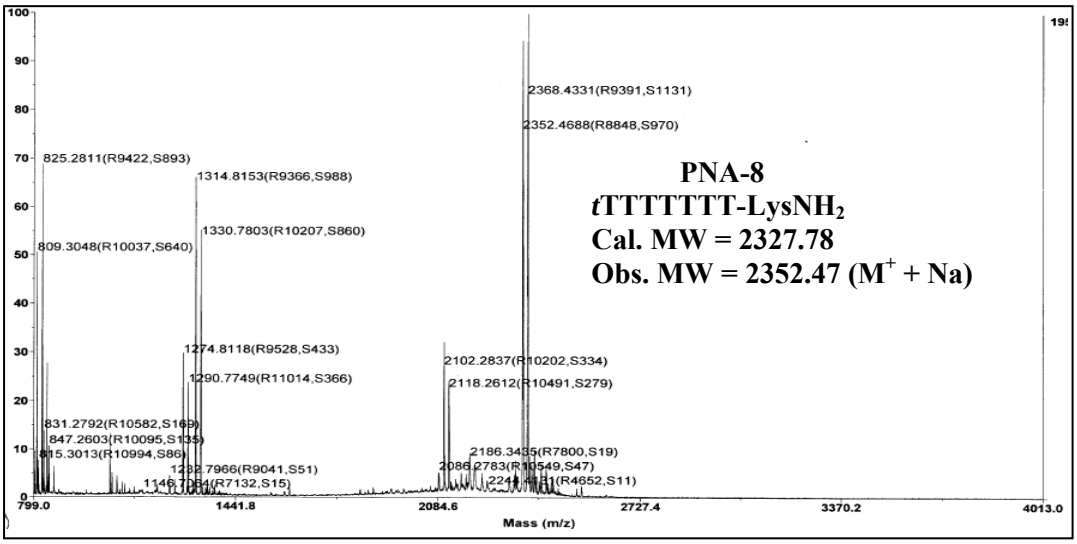


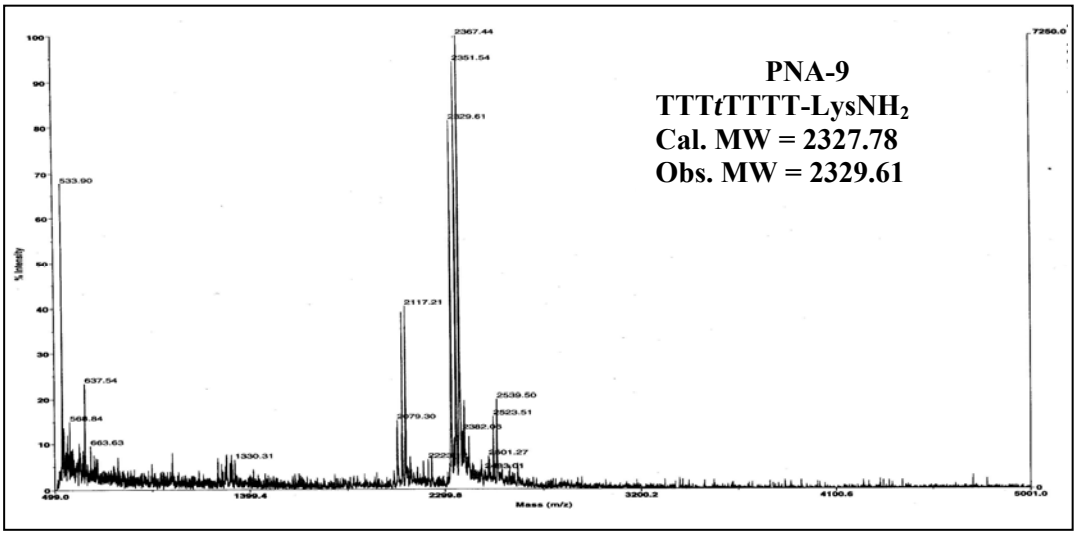

PNA-4
TTTTTTTt-LysNH2
Cal. MW = 2313.61
Obs. MW = 2315.92
% Intensity
Mass (m/z)

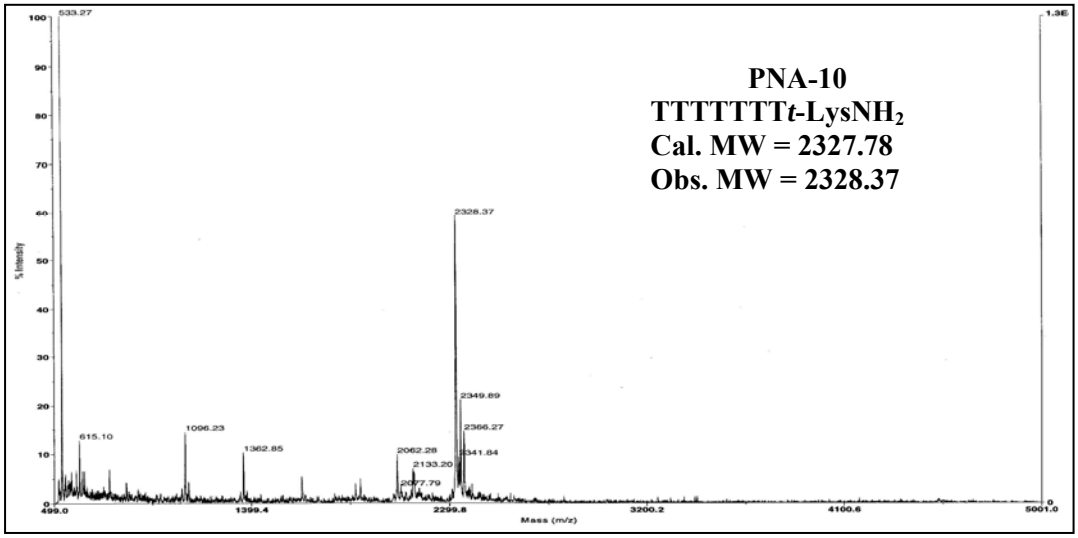

PNA-5
tTTTTTTT-LysNH2
Cal. MW = 2313.61
Obs. MW = 2316.76
% Intensity
Mass (m/z)

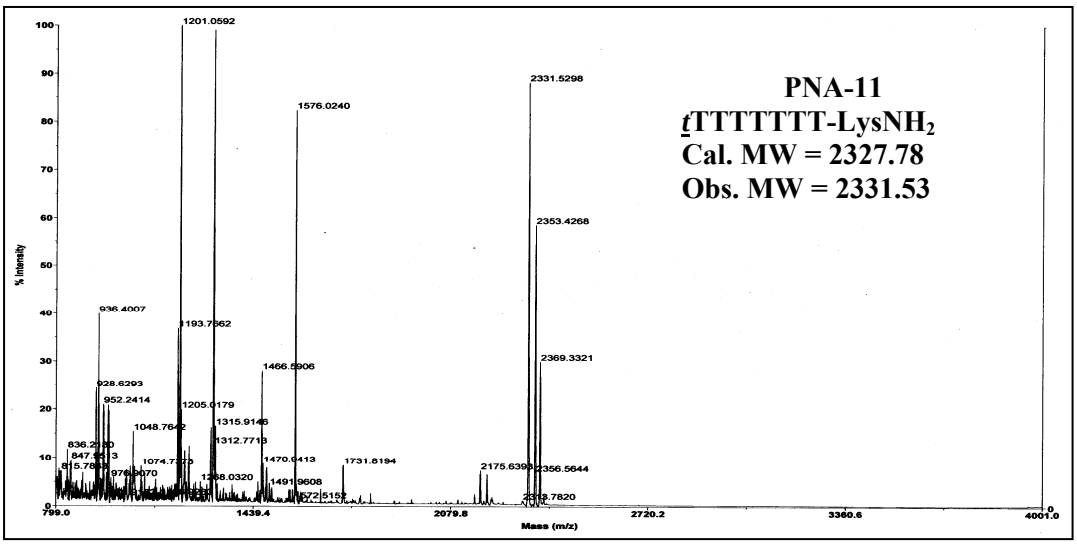

PNA-6
TTTtTTTT-LysNH2
Cal. MW = 2313.61
Obs. MW = 2316.66
% Intensity
Mass (m/z)

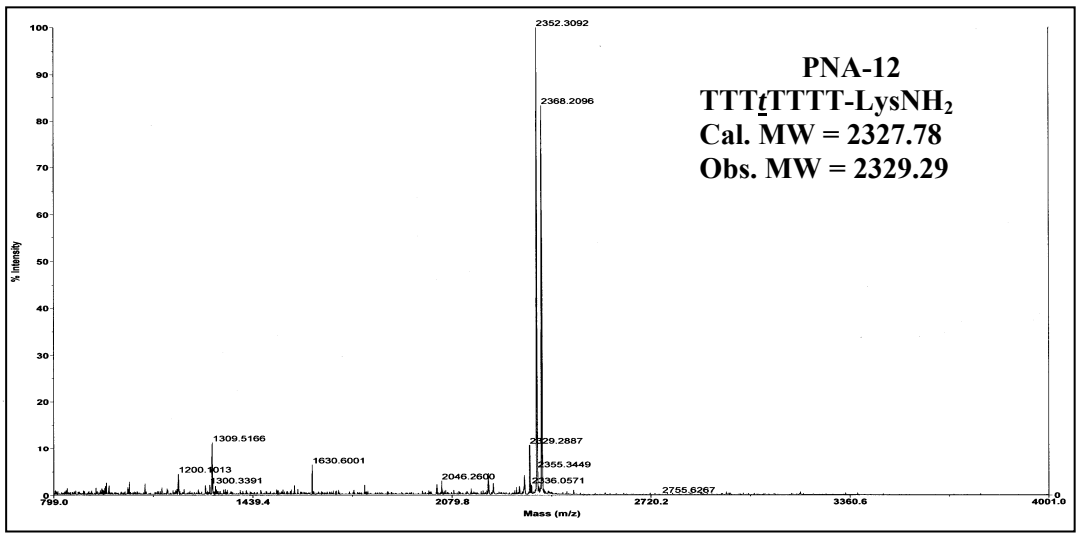

PNA-7
TTTTTTTt-LysNH2
Cal. MW = 2313.61
Obs. MW = 2316.08
% Intensity
Mass (m/z)
2316.08
2336.71
2353.20
2091.60
2128.79
2151.48
1826.58
1925.99
2298.09
2503.88

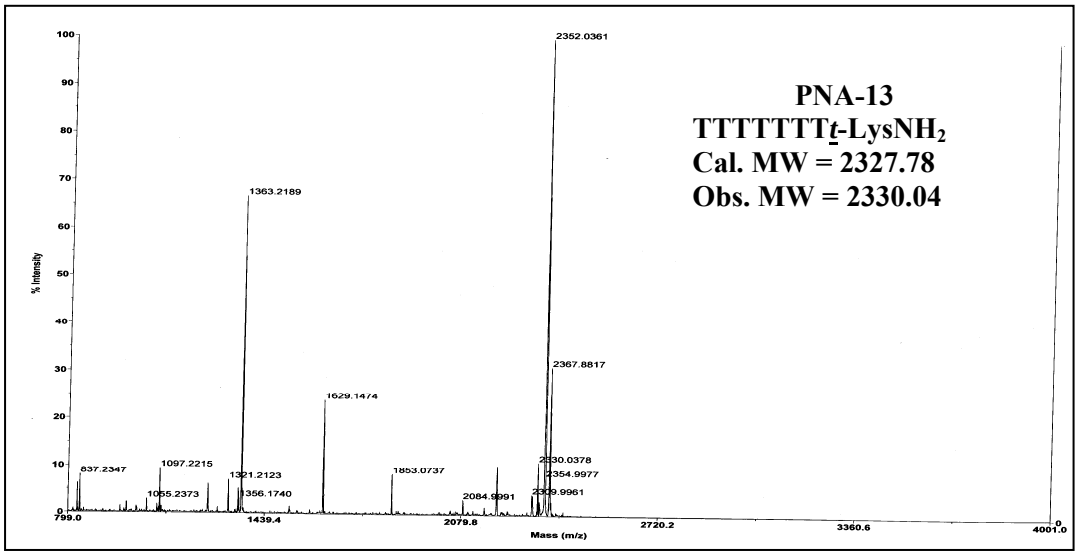

PNA-8
tTTTTTTT-LysNH2
Cal. MW = 2327.78
Obs. MW = 2352.47 (M+ + Na)
Mass (m/z)
2368.4331(R9391,S1131)
2352.4688(R8848,S970)
825.2811(R9422,S893)
1314.8153(R9366,S988)
1330.7803(R10207,S860)
809.3048(R10037,S640)
1274.8118(R9528,S433)
2102.2837(R10202,S334)
1290.7749(R11014,S366)
2118.2612(R10491,S279)
831.2792(R10582,S169)
847.2603(R10095,S135)
815.3013(R10994,S86)
2186.3435(R7800,S19)
1232.7966(R9041,S51)
2086.2783(R10549,S47)
1146.7064(R7132,S15)

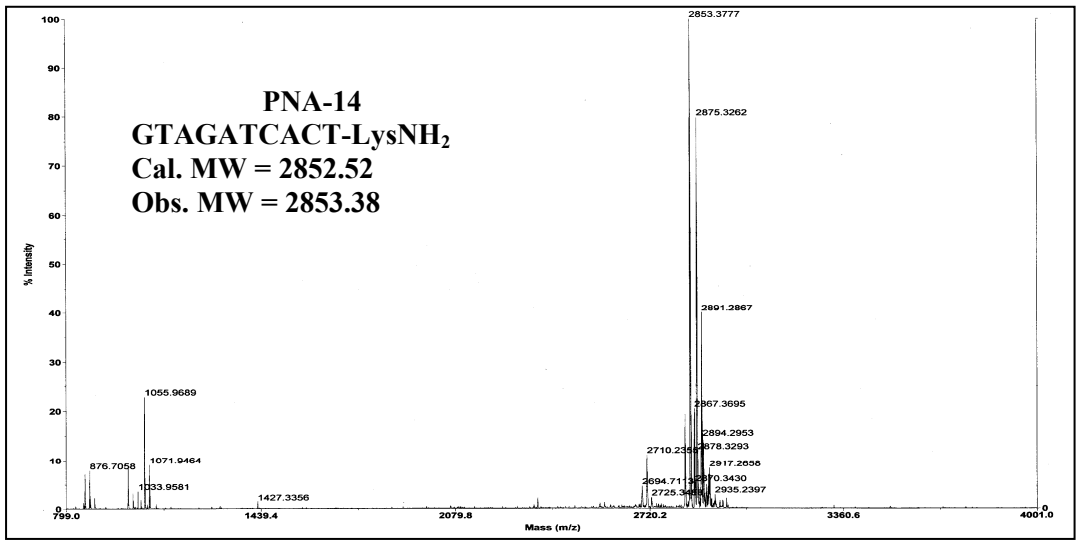

PNA-9
TTTtTTTT-LysNH2
Cal. MW = 2327.78
Obs. MW = 2329.61
% Intensity
Mass (m/z)
2367.44
2351.54
2329.61
533.90
2117.21
637.54
568.84
663.63
1330.31
2539.50
2523.51
2501.27

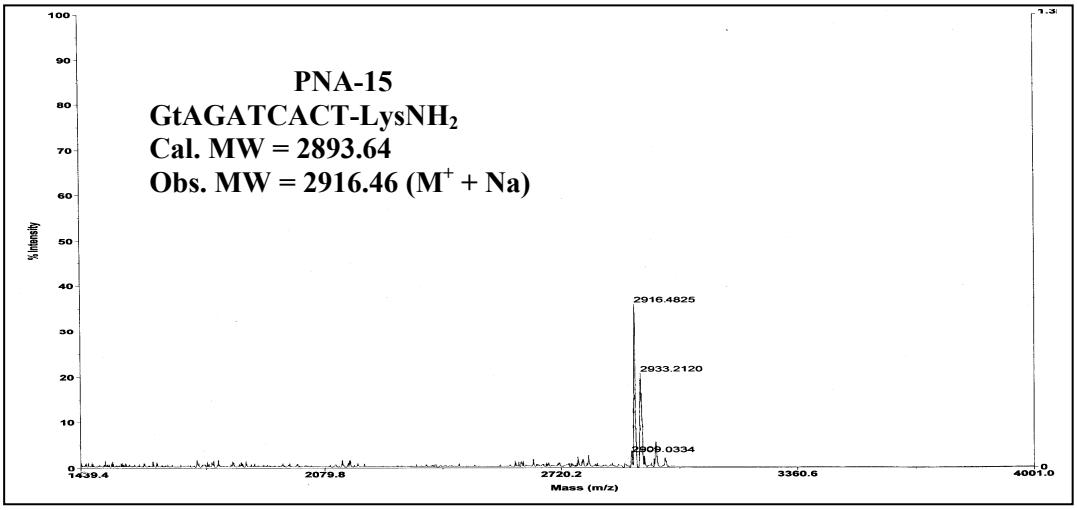

PNA-10
TTTTTTTt-LysNH2
Cal. MW = 2327.78
Obs. MW = 2328.37
% Intensity
Mass (m/z)

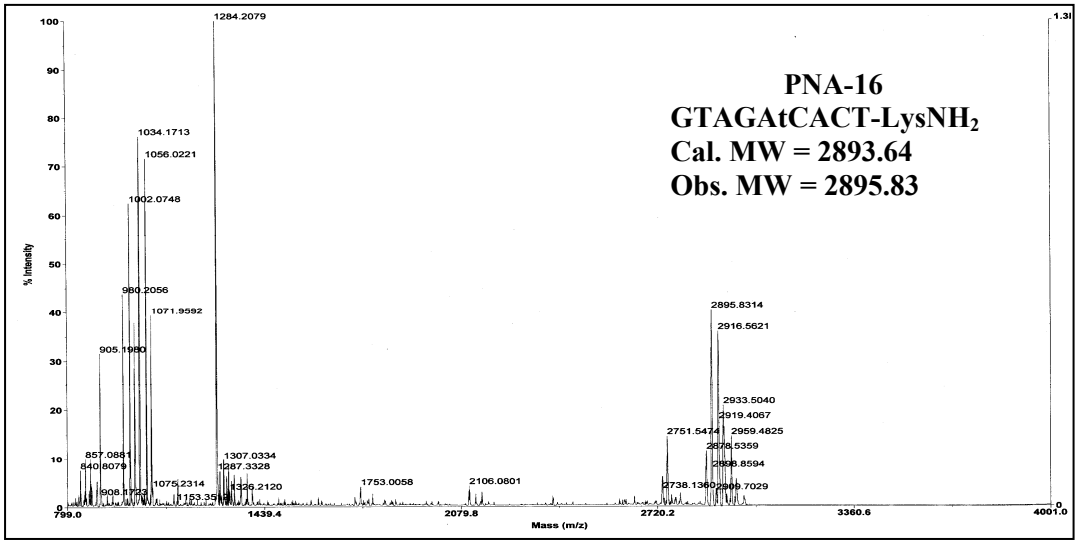

PNA-11
tTTTTTTT-LysNH2
Cal. MW = 2327.78
Obs. MW = 2331.53
% Intensity
Mass (m/z)

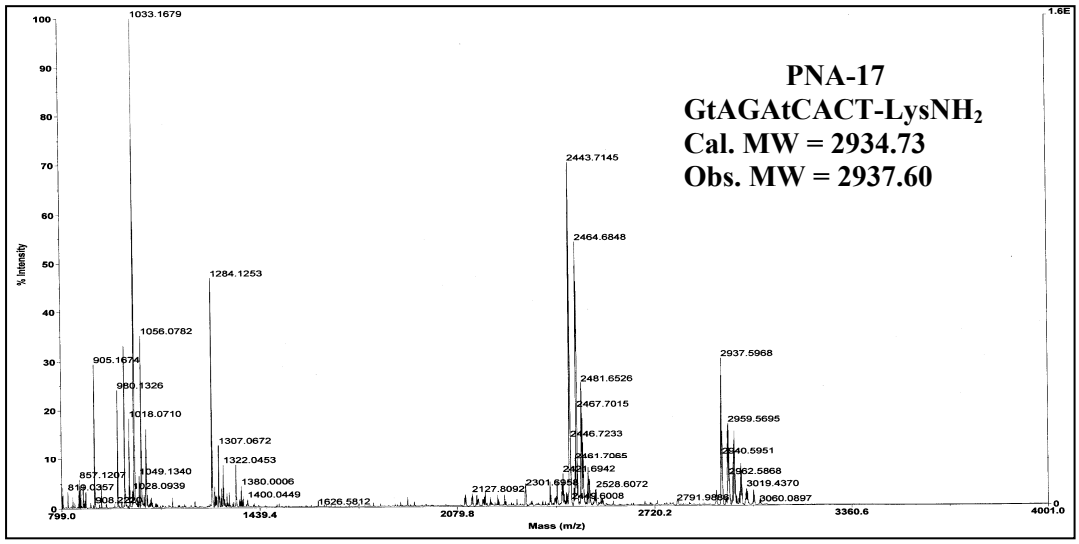

PNA-12
TTTtTTTT-LysNH2
Cal. MW = 2327.78
Obs. MW = 2329.29
% Intensity
Mass (m/z)

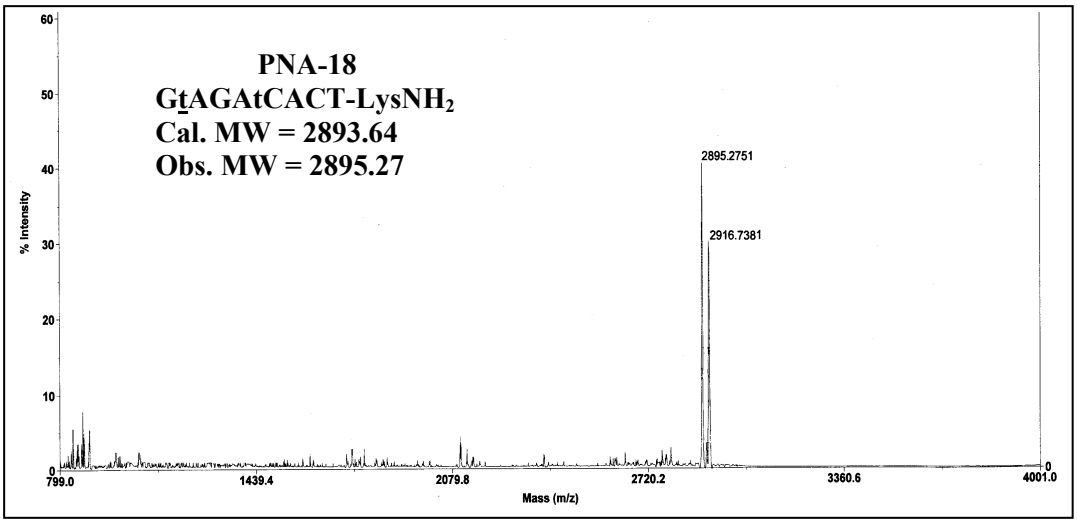

PNA-13
TTTTTTTt-LysNH2
Cal. MW = 2327.78
Obs. MW = 2330.04
% Intensity
Mass (m/z)
2352.0361
1363.2189
2367.8817
1629.1474
637.2347
1097.2215
1321.2123
1055.2373
1356.1740
1853.0737
2084.9991
2330.0378
2354.9977
2309.9961
799.0
1439.4
2079.8
2720.2
3360.6
4001.0

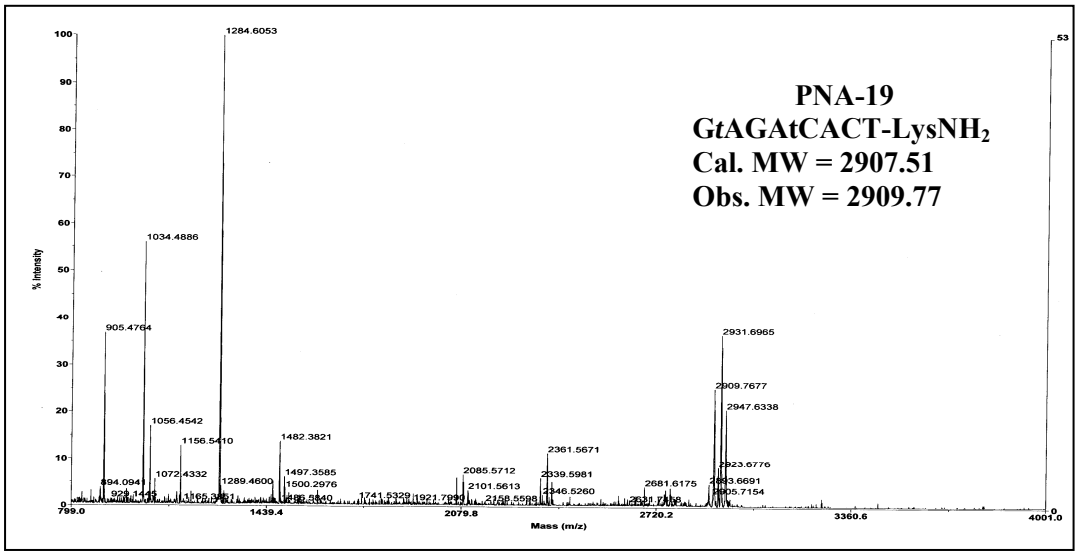

PNA-14
GTAGATCACT-LysNH2
Cal. MW = 2852.52
Obs. MW = 2853.38
% Intensity
Mass (m/z)
2853.3777
2875.3262
2891.2867
1055.9689
2367.3695
2894.2953
2878.3293
2917.2656
2935.2397
876.7058
1071.9464
1033.9581
1427.3356
799.0
1439.4
2079.8
2720.2
3360.6
4001.0

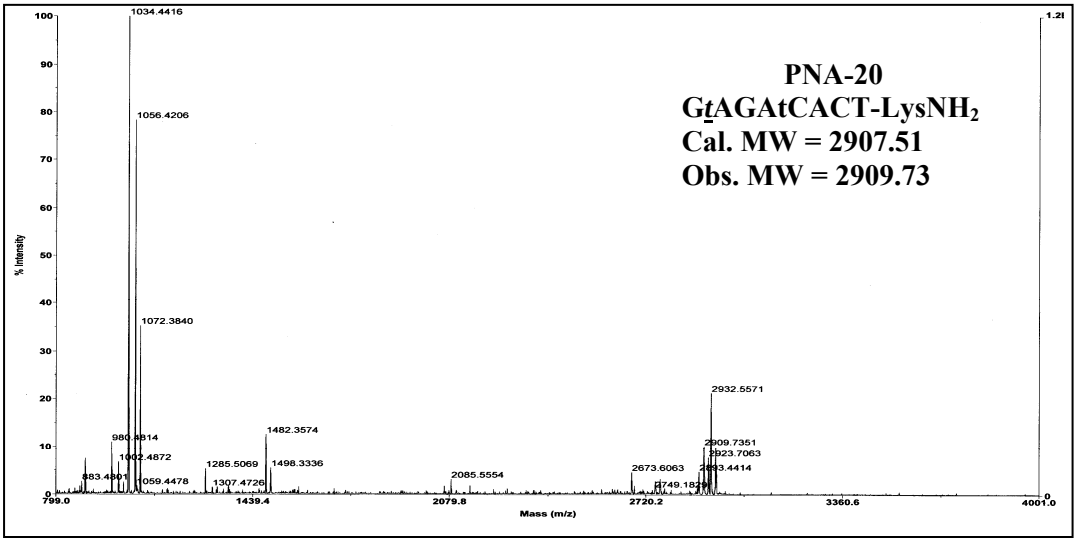

PNA-15
GtAGATCACT-LysNH2
Cal. MW = 2893.64
Obs. MW = 2916.46 (M+ + Na)
% Intensity
Mass (m/z)
2916.4825
2933.2120
1439.4
2079.8
2720.2
3360.6
4001.0

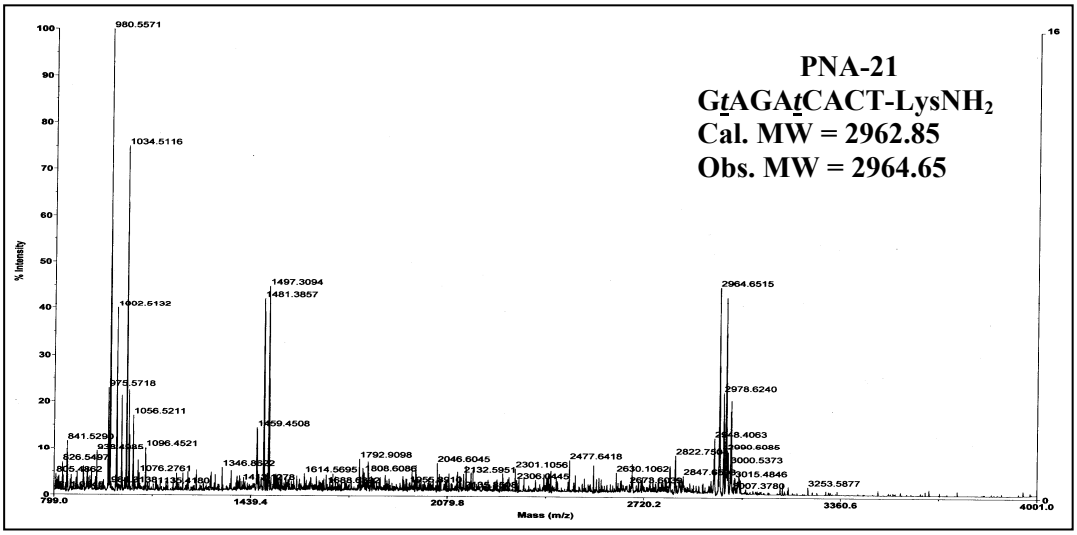

PNA-16
GTAGAtCACT-LysNH2
Cal. MW = 2893.64
Obs. MW = 2895.83
% Intensity
Mass (m/z)
1284.2079
1034.1713
1056.0221
1002.0748
980.2056
1071.9592
905.1980
857.0881
840.8079
908.1723
1075.2314
1153.3514
1307.0334
1287.3328
1326.2120
1753.0058
2106.0801
2738.1360
2895.8314
2916.5621
2933.5040
2919.4067
2751.5474
2959.4825
2876.5359
2898.8594
2909.7029


PNA-17
GtAGAtCACT-LysNH2
Cal. MW = 2934.73
Obs. MW = 2937.60
% Intensity
Mass (m/z)
1033.1679
2443.7145
2464.6848
1284.1253
1056.0782
905.1674
980.1326
1018.0710
2481.6526
2467.7015
2959.5695
2446.7233
1307.0672
1322.0453
2461.7065
2937.5968
2840.5951
2962.5868
857.1207
1049.1340
819.0357
1028.0939
1380.0006
1400.0449
2127.8092
2528.6072
3019.4370
3060.0897
1626.5812
2791.9869

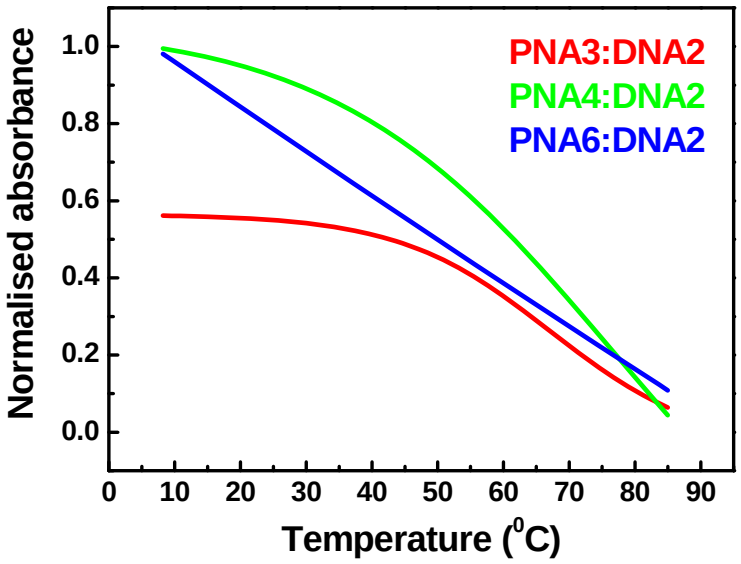

PNA-18
GtAGAtCACT-LysNH2
Cal. MW = 2893.64
Obs. MW = 2895.27
% Intensity
Mass (m/z)
2895.2751
2916.7381

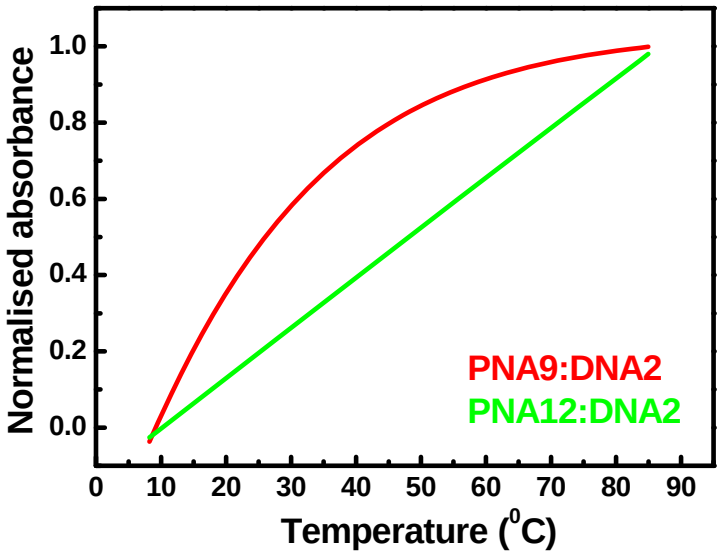

PNA-19
GtAGAtCACT-LysNH2
Cal. MW = 2907.51
Obs. MW = 2909.77
% Intensity
Mass (m/z)

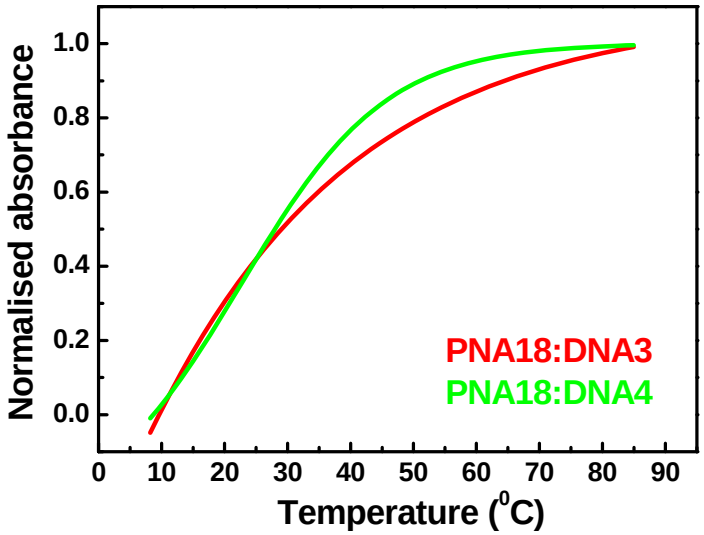

PNA-20
GtAGAtCACT-LysNH2
Cal. MW = 2907.51
Obs. MW = 2909.73
% Intensity
Mass (m/z)


PNA-21
GtAGAtCACT-LysNH2
Cal. MW = 2962.85
Obs. MW = 2964.65
% Intensity
Mass (m/z)

UV melts of sequences not showing any detectable melting transition.






Normalised absorbance
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
Temperature (0C)
PNA18:DNA3
PNA18:DNA4

CHAPTER 4

Synthesis of new pyrrolidine analogues with potential medicinal and biological applications.

Part-A: Synthesis of Pyrrolidine-β PNA Monomer

4.1 Introduction

The favourable binding affinity and specificity of PNA for complementary DNA/RNA[1] has sparked off the quest for compounds that would further improve its applicability by enhancing water-solubility and permeability across cell membranes. The basic aminoethylglycyl structure of PNA has been variously modified with this aim with varying degrees of success, which includes backbone substitution, cyclic backbones or modifications in the nucleobase.[2]

The backbone substitution by changing the glycine unit of PNA by α-amino acids does not alter the hybridization properties of PNA to a great extent in most of the cases.[3] But the introduction of charged cationic amino acids like lysine or arginine has shown a markable effect not only on the hybridization properties of PNA, but this also facilitates the cellular uptake of PNA.[3-5] On other hand, the substituents present in the ethylene diamine segment of PNA, especially the γ-substitution in PNA was proved to be quite successful and interesting. In particular the γ-methyl PNA (Figure 1), which adopts a favourable conformation in single strand PNA, binds with improved stability to complementary DNA/RNA. The reduced entropy loss seems to favour the enhanced hybridization properties.[6,7] The efforts to see the effect of unnatural α-amino acid instead of glycine in PNA, the substituent on the γ-position of PNA and their cellular uptake studies are ongoing projects in this laboratory. The preliminary results have shown stabilization of these PNAs with improved cellular uptake.

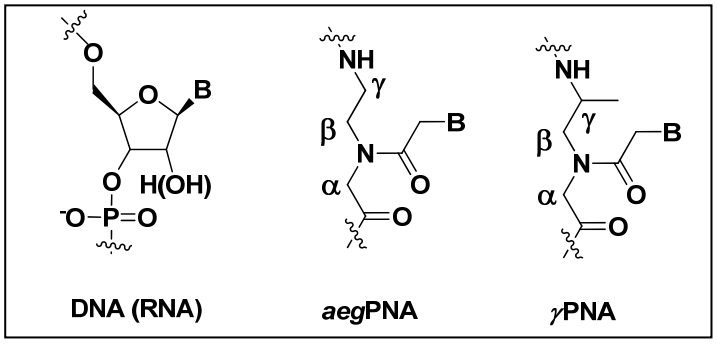


Figure 1: Chemical structures of DNA (RNA) and PNAs

Many attempts have been made to constrain the flexibility of the *aeg*PNA structure by cyclic derivatives, which show beneficial effect on hybridization

properties.[8] This has been discussed in detail in the preceding chapters. In continuing efforts to further tune the hybridization properties of modified PNAs and to determine the effects of different ring constraints to the PNA backbone, synthesis of pyrrolidine β-PNA monomer has been described.

4.2 Rationale for the present work

Efforts to achieve optimum fine-tuning of the PNA structure initiated in the preceding chapters are herewith extended in another direction. The encouraging results from the *dap* and *amp* PNAs discussed in chapter **2** and **3**, which have substitution at β-carbon to pyrrolidine ring nitrogen rather than the widely studied α-substituted or proline derived PNAs, led us to the design of a pyrrolidine β-PNA monomer **4**. The design of the pyrrolidine β-PNA monomer is such that the nucleobase attachment to the backbone is through a flexible methylene linker and not directly to the pyrrolidine ring. The position of the base linker in pyrrolidine β-PNA is changed, from α-position as in the pyrrolidine PNA to β-position, retaining the same PNA backbone (Figure 2). This shift may have influence on the puckering of the pyrrolidine ring and can help in tuning the modified PNAs for better stability with selectivity.

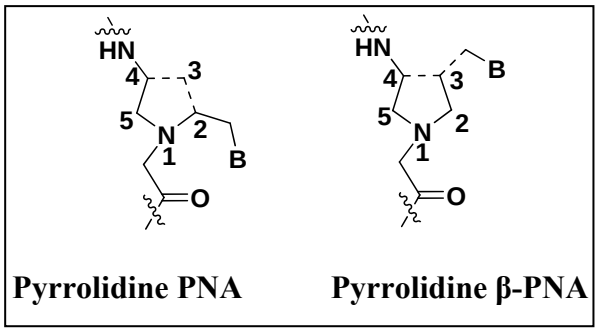


Figure 2: Backbone representation of pyrrolidine and Pyrrolidine β-PNA

It was found earlier from this laboratory that chimeric *aeg*-pyrrolidine PNA T_8 oligomers derived from either (2*R*,4*S*) and (2*S*,4*S*) stereochemistry on pyrrolidine PNA backbone stereochemistry of the ring exhibited stereochemistry-dependent thermal stability of the resulting PNA_2:DNA complexes.[9] The introduction of a single unit of (2*R*,4*S*) pyrrolidine PNA in the T_8 sequence stabilized the PNA_2:DNA complexes ($\Delta Tm=16^oC$). The introduction of the (2*S*,4*S*) isomer on the other hand destabilized the complex ($\Delta Tm= -16^oC$). The binding efficiency of pyrrolidine PNA

oligomers thus seems to be dictated by the stereochemistry of the pyrrolidine backbone that is a part of the acetyl linker connecting the nucleobase in the parent *aeg*PNA.

In the present work, It was chosen to synthesize pyrrolidine β-PNA with *trans* stereochemistry, consistent with the more stabilizing (2*R*,4*S*) pyrrolidine PNA. The only difference in the design of the pyrrolidine β-PNA monomer is the position of the base linker which shifted from C2 to C-3 position (Figure 2). This will have one additional carbon in the base linker as compared to *aeg*PNA. The involvement of the two carbon atoms of the base linker in the pyrrolidine ring is supposed to compensate the effect of one carbon extension towards the stability of pyrrolidine PNA:DNA complexes. This was evident from the results of *secdap* and *secamp* PNA (Figure-3) discussed in preceding chapter in which the backbone having two carbon atoms in pyrrolidine ring structure (*secdap*PNA) destabilized the PNA:DNA complexes. The stability of the PNA:DNA complexes is remarkably enhanced by the insertion of one extra carbon as in *secamp*PNA.

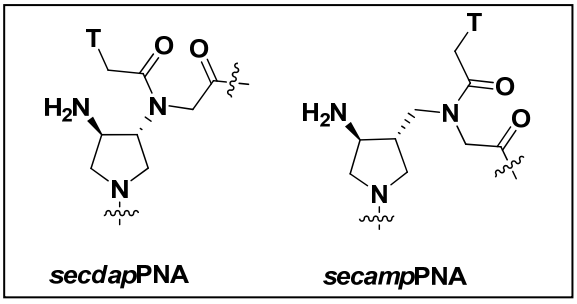


Figure 3: Chemical structures of *secdap* and backbone extended *secamp* PNAs.

Moreover, in case of *secdap* and *secamp*PNA, the extra carbon (as methylene) was introduced in the backbone of PNA while in pyrrolidine β-PNA, the extra carbon is inserted in the base linker. Hence it would be interesting to see if it follows the same trend as for *secdap* and *secamp* PNA. This modification will also allow the introduction of rigidity into the PNA backbone in the form of pyrrolidinyl ring, while simultaneously keeping a certain degree of flexibility in the linker to the nucleobase. The pyrrolidine ring nitrogen is a tertiary amine and bears a positive charge at physiological pH and may help to improve the solubility of PNA oligomers.

The specific objective of this part of the chapter is to carry out the synthesis of pyrrolidine β-PNA monomer.

4.3 Results and Discussion

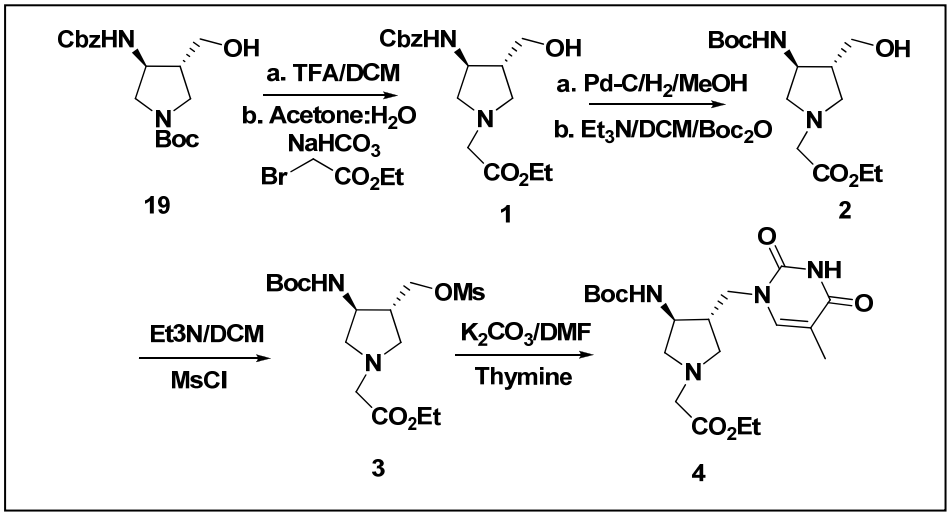
The N-*cbz*, N-*boc* protected alcohol intermediate **19** (Scheme 2, chapter 2) in the synthesis of *amp*PNA monomer was utilised for the synthesis of pyrrolidine β-PNA monomer that was readily obtained starting from glycine.

4.3.1 Synthesis of (3*S*,4*S*) Pyrrolidine β-PNA monomer

Synthetic approach 1

In the first approach, N-*boc* group of the alcohol **19** was deprotected with triflouroacetic acid and the free amine was alkylated with ethybromoacetate using excess of base (Scheme 1) to get N-alkylated alcohol **1.** The N-*cbz* group of compound **1** was deprotected by hydrogenation using Pd/C to give free amine which was protected with *t-boc* group to get compound **2**. The formation of the compound **2** was confirmed by the appearance of peaks due to *t-boc* protons in ^{1}H NMR at δ1.43 ppm and disappearance of signals due to *cbz* phenyl protons from aromatic region and *cbz* methylene protons at δ5.09 ppm. The primary alcohol group in compound **2** was mesylated to afford corresponding mesyl **3** which was reacted with thymine in the presence of base. But too many close spots were seen in the TLC and the desired monomer **4** could be separated in very low yield.

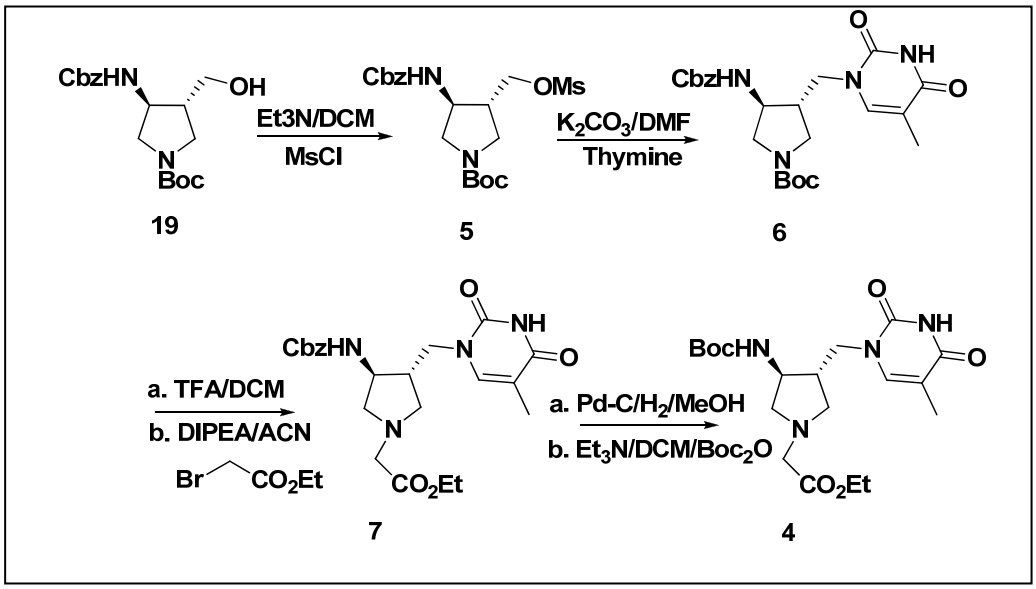
Scheme 1: Synthesis of (3*S*,4*S*) Pyrrolidine β-PNA monomer



Synthetic approach 2

In the second approach (Scheme-2), the alcohol **19** was converted to its mesyl derivative **5** followed by replacement by N^1-thymine to get compound **6**. The reaction did not proceed at room temperature and so was heated to 80°C, the temperature at which the reaction was not clean and compound **6** was obtained in low yield. The identity of product was confirmed by the appearance of peak for thymine methyl protons at 1.89 ppm and the disappearance of signals due to mesyl proton at 3.02 ppm. Compound **6** upon N-boc deprotection gave the corresponding amine that was N-alkylated with ethylbromoacetate to get compound **7.** This compound was unstable on silica gel for column purification and hence used without any further purification. The N- Cbz group of compound **7** was deprotected to get free amine which was *insitu* protected with *boc* to get the desired pyrrolidine β-PNA monomer **4** in moderate yield.

Scheme 2: Synthesis of (3*S*,4*S*) Pyrrolidine β-PNA monomer



4.4 Conclusion

The synthesis of pyrrolidine β-PNA monomer **4** has been achieved. The incorporation of the PNA monomer in the *aeg*PNA sequences might be useful to achieve good binding with selectivity towards DNA/RNA complexes. This will also constitute a new generation of PNA monomers which are derived from five

membered pyrrolidine ring but with substitution at β-position rather than the α-substitution in most pyrrolidinyl PNAs. However there is a need to increase the overall yield of the reactions for the synthesis of the pyrrolidine β-PNA monomer.

Part-B: Convenient and Scalable Route to Heterocyclic β-Ketoesters & synthesis of N-*fmoc*-protected (3*S*,4*R*) 4-aminopyrrolidine-3-carboxylic acid (18) as a precursor for collagen peptides.

4.5 Introduction

The interest in pyrrolidine PNA monomers demanded development of convenient synthetic routes to them. The five and six membered heterocyclic β-ketoesters **8-10** (Figure 4) have previously received attention as precursors for the synthesis of variety of medicinally and biologically interesting molecules. They are the key compounds in the synthesis of various carbapenem derivatives[10] which have shown higher activity when compared to imipenem,[11] one of the most potent type of antibacterial agents and is among those carbapenems used as last resort against infections in the clinical field. The five membered pyrrolidinyl β-ketoester **8** can be easily converted to a novel compound, 4-aminopyrrolidine-3-carboxylic acid **12**, in both racemic[12] and enantiomerically pure form.[13,14] The racemic form of this compound has been used to probe the structure of GABA receptor.[15] The *cis*-form of **12** is a modestly active inhibitor of Influenza Neuraminidase and has been used as a lead structure to discover the trisubstituted pyrrolidine carboxylic acid A-192558 **11** as a potent inhibitor of Influenza Neuraminidase.[12,16] Its *trans*-form has been incorporated in short chain water soluble peptides which are shown to adopt distinct helical secondary structures in solution.[17]

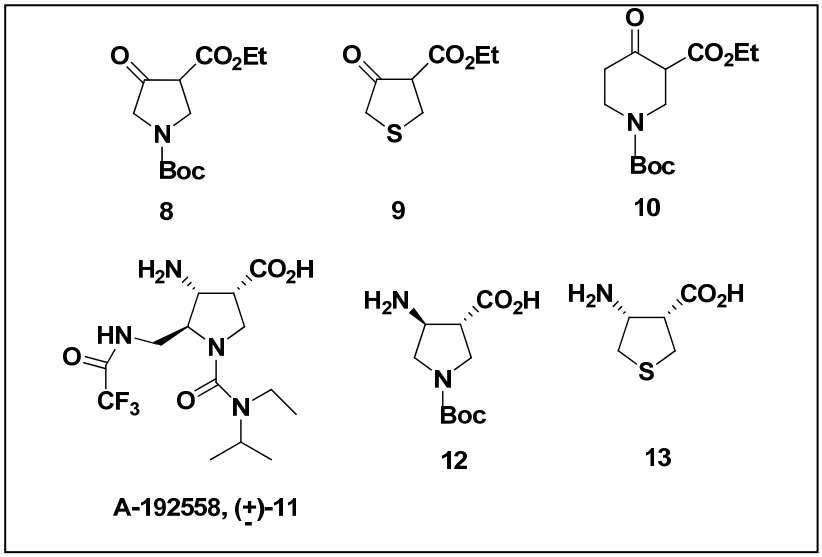


Figure 4: Hetrocyclic β-ketoesters and their biologically important derivatives.

On the other hand, the six membered piperidinyl β-ketoester could be reduced to the corresponding disubstituted piperidine, which is useful addition to the chiral pool. The completely deprotected derivatives of the hydroxy ester display significant activity on the functioning of the central neurotransmitter system through γ-aminobutyric acid (GABA) receptor and therefore of interest in therapies for various psychiatric and neurological disorders.[18] The six membered piperidinyl β-ketoester is also a potent substrate-competitive inhibitor of the neuronal GABA uptake process.[19]

Based on the importance of these molecules it is interesting to develop methodology to construct various heterocyclic rings using the common synthetic precursors and procedures. It was thought of utilizing the five membered pyrrolidine β-ketoester as the starting material for the synthesis of N-*fmoc* protected (3*S*,4*R*) 4-aminopyrrolidine-3-carboxylic acid **18** which may also be a useful precursor for collagen peptide mimetics.

4.6 Results and discussion

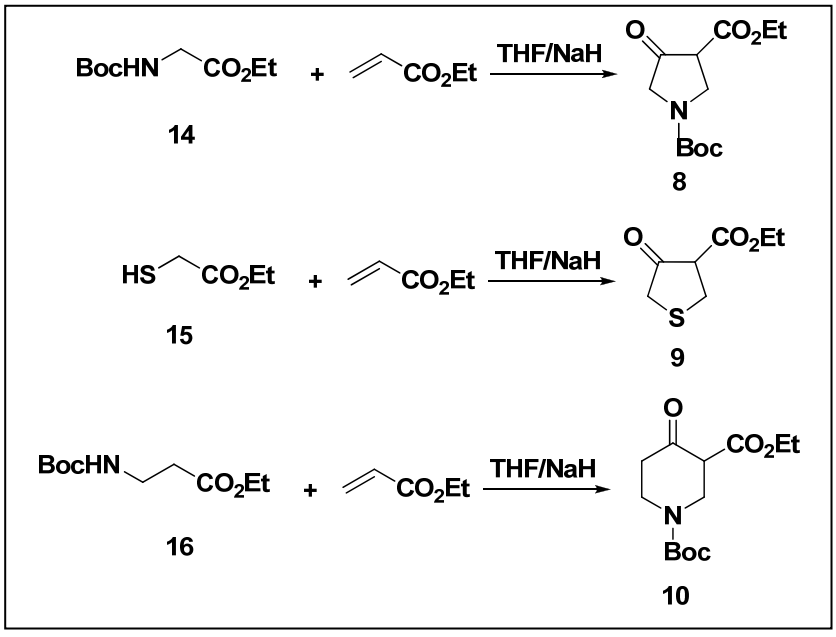
4.6.1 Convenient and Scalable Route to Heterocyclic β-Ketoesters

Despite recent interest in the preparation of this structural class, it was surprising to find that only limited methods are available for the synthesis of these heterocyclic β-keto esters. For instance, the five membered pyrrolidinyl β-ketoester **8**

can be obtained by using the route described by Kuhn and Osswald,[20] which consists of a two step procedure involving the Michael addition of glycine ester to ethyl acrylate followed by an intramolecular Dieckmann condensation in the second step. The yields obtained by this method were poor because the condensation leads to a mixture of isomers of Dieckmann product. Modifications of the procedure of Kuhn and Osswald were also done by other authors in the past and in recent times.[21] There are only a few reports for the synthesis of six membered piperidinyl β-ketoester **9** and five membered tetrahydrothiophene β-ketoester **10**, which involves two steps: addition reaction as first and dieckmann condensation as second.[22] The experimental procedure for preparing both five and six membered is very cumbersome involving sodium sand and the toxic solvents as well as reagents. However there is no simple methodology existing in literature which allows the synthesis of these heterocyclic β-ketoesters.

Herein, a simple, convenient, and scalable synthetic route to both heterocyclic β-ketoesters is described involving a tandem Michael addition followed by a Dieckmann condensation in a single step using THF as solvent and NaH as base (Scheme 3). The N-Boc protected ethyl ester of glycine **14** and β-alanine **16** was obtained from the corresponding amino acid ester hydrochloride salt followed by N-*boc* protection in DCM/Et_3N.

Scheme-3: Synthesis of heterocyclic β-ketoesters 8, 9 and 10.

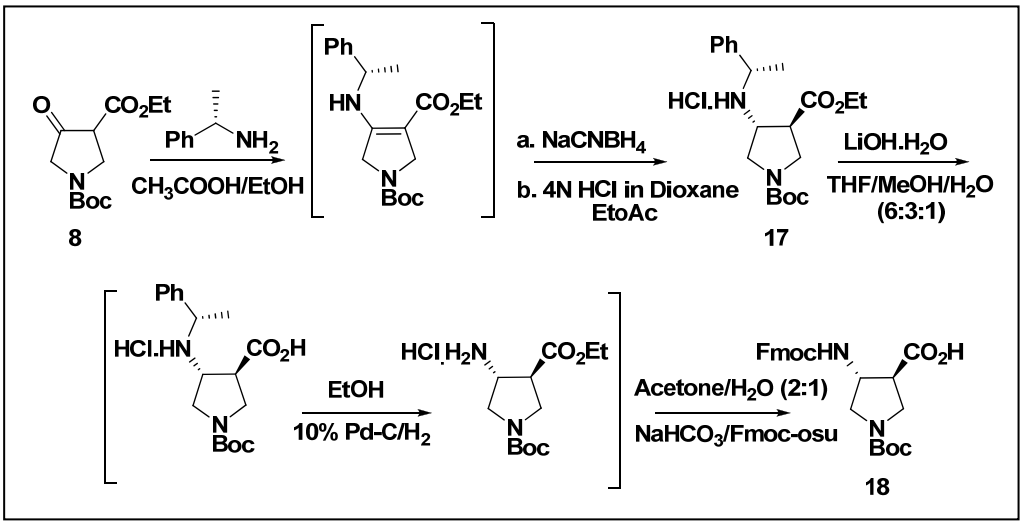


All the reactions proceeded with a quantitative yield. All the compounds synthesized were characterized by ^{1}H, ^{13}C NMR and through mass spectroscopy.

4.6.2 Synthesis of N-*fmoc*-protected (3*S*,4*R*) 4-aminopyrrolidine-3-carboxylic acid 18

The five membered ketoester **8** synthesized by following the above method was converted to N-*fmoc* protected (3*S*,4*R*) 4-amino-β-proline[13] **18** which can be a useful precursor for collagen model peptides.

Scheme 4: Synthesis of N *fmoc*- (3*S*,4*R*) 4-aminopyrrolidine-3-carboxylic acid[13] (18)



The β-ketoester **8** (Scheme 4) was reacted with (*R*)-α-methylbenzylamine in the presence of acetic acid, and the resulting enamine is reduced in situ with $NaBH_3CN$. This reduction produces a mixture of four diastereomeric aminoesters in which **17** is the major product which was precipitated out by treatment with 4 N HCl in dioxane as its hydrochloride salt in diastereomerically pure form. A single trans isomer was obtained in relatively pure form after this treatment (≥ 98% *de*), although there is contamination from the other trans isomer. Recrystallization from acetonitrile yields a very pure form of **17** (≥ 99% *de*) in 30% overall yield from **8**. The diastereomeric excess for compound **17** was determined by GC-MS analysis. Compound **17** on alkaline hydrolysis afforded the corresponding acid which was hydrogenated using Pd-C. The free amine obtained was N-protected with fmoc group to give N *fmoc*- (3*S*,4*R*) 4-aminopyrrolidine-3-carboxylic acid **18** in good yield,

which can be incorporated in collagen model peptides to determine its ability to form stable triple helix.

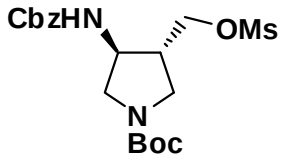
4.7 Conclusion

In conclusion, a very simple, convenient, single step, less expensive and high yielding method for the synthesis of five and six memebered heterocyclic β-ketoesters has been achieved. The five membered pyrrolidinyl β-ketoester **8** was further converted to N-*fmoc* protected (3*S*,4*R*) 4-amino-β-proline **18** which can be used in collagen model peptides to see its effect on collagen stability. It will also antagonize the possible n→π* transition proposed in the collagen peptides and will help reveal its importance in collagen stability. The 4-amino group may help in tuning the cis-trans conformer ratio of prolyl peptide bond and also provide a pH dependent switch for the triple-helix stability. Further work to produce chiral substituted pyrrolidinyl β-keto esters using chiral pool of α-amino acids and the synthesis of 3-hydroxy proline using this methodology is in progress.

4.8 Experimental section

General Procedure: Melting points of samples were determined in open capillary tubes on Bruker melting point apparatus B-540 and are uncorrected. IR spectra were recorded on an infrared Fourier Transform spectrophotometer using KBr pellets or as neat. Column chromatographic separations were performed using silica gel 60-120 mesh, solvent systems EtOAc/Pet ether and MeOH/DCM. ^{1}H and ^{13}C were obtained using Bruker 200 MHz and 500 MHz NMR spectrometers. The chemical shifts are reported in delta (δ) values. Mass spectra were obtained either by MS (EI) techniques.

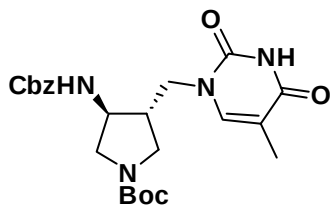
(3*S*,4*R*)-1-N-Boc-3-*cbz* amino-4-((methylsulfonyl)oxy) methylpyrrolidine (5)



To a solution of hydroxyl compound **19** (2.0 g, 5.7 mmol) in 20 ml of dry DCM at 0°C, triethylamine (3.1 mL, 22.2 mmol) was added followed by mesyl chloride (0.9 ml, 11.5 mmol) dropwise with constant stirring. After the addition was complete, the ice bath was removed and stirring continued at ambient temperature for 2 hrs. The reaction mixture was diluted with 50 ml of DCM and washed with water, followed by brine and dried over anhydrous Na_2SO_4. The organic layer was evaporated which on purification gave the mesyl **20** as colourless liquid. Yield: (1.5 g, 62%).

^{1}H NMR (200 MHz, $CDCl_3$): *δ* 7.48-7.32 (m, 5H, cbzPh), 5.10 (s, 2H, cbz -CH_2), 5.08-4.97 (m, 1H, -NH), 4.37- 4.05 (m, 3H, -C<u>H</u>-NH, -CH_2), 3.90-3.58 (m, 2H, -C<u>H</u>$_2$-OMs), 3.31-3.08 (m, 2H, -CH_2), 3.02 (s, 3H, -OMS), 2.60-2.35 (m, 1H, -CH), 1.45 (s, 9H, *t*-boc), **^{13}C NMR** (50 MHz, $CDCl_3$): *δ* 155.8, 154.0, 136.0, 128.5, 128.2, 128.1, 80.0, 68.4, 67.0, 37.2, 29.6, 28.3. ; **MS (EI)** *m/z* 428.16, Found 451.65 [M + Na].

(3*S*,4*S*)-1-N-Boc-3-*cbz* amino-4-thyminylmethyl pyrrolidine (6)

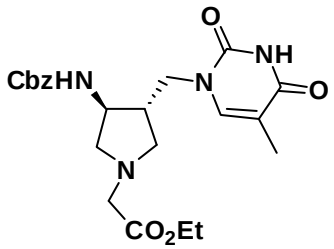


To compound **5** (250 mg, 0.58 mmol) taken in 10 ml of DMF was added potassium carbonate (145 mg, 1.05 mmol) followed by thymine (110 mg, 0.87 mmol) at room temperature. The reaction mixture was heated at

80°C for 18 hrs after which it was diluted with 100 ml of ethyl acetate and washed with water, brine and dried over sodium sulphate and purified by column chromatography to obtain compound **12** as a colourless liquid. Yield: (100 mg, 37 %).

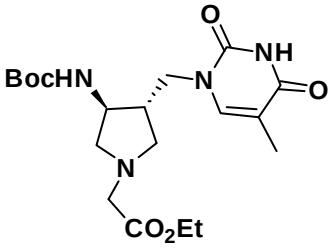
^{1}H NMR (200 MHz, $CDCl_3$): δ 7.46-7.01 (m, 6H, cbzPh, -CH Thymine), 5.25-4.65 (m, 3H, cbz -CH_2, -NH), 4.27- 3.40 (m, 5H, -CH-NH, 2 x ring -CH_2), 3.35-2.83 (m, 2H, -CH_2-T), 2.65-2.35 (m, 1H, -CH), 1.89 (s, 3H, -CH_3 Thymine), 1.45 (s, 9H, *t*-boc), **^{13}C NMR** (50 MHz, $CDCl_3$): δ 163.7, 156.2, 154.3, 151.8, 139.5, 136.1,128.5, 128.3, 128.2, 109.6, 80.0, 67.0, 53.4, 51.2, 47.8, 42.2, 32.0, 28.5, 12.9. ; **MS (EI)** *m/z* 457.22, Found 458.38 [M + 1].

Ethyl 2-[((3*S*,4*S*)-3-*cbz* amino-4-thyminylmethyl) pyrrolidin-1-yl] acetate (7)



To a solution of compound **6** (0.12 mg, 0.26 mmol) in dry DCM (1ml) was added TFA (1ml) and the reaction was stirred at room temperature for 10min, when TLC indicated a complete absence of starting material. The solvents were removed under vacuum. The resulting free amine was immediately subjected to alkylation using ethylbromoacetate (0.036 ml, 0.31 mmol) in dry THF (5 ml) in the presence of DIPEA (0.4 ml, 2.31 mmol). The amine was completely consumed within 12h, upon which, the solvents were removed and water (50ml) was added. This was extracted with ethylacetate three times and the combined ethylacetate layer was dried over Na_2SO_4 and concentrated to get compound **7** in crude form. This was used without column purification as it decomposes on silica gel column. Yield (crude): (80 mg, 68%)

Ethyl 2-[((3*S*,4*S*)-3-*boc* amino-4-thyminylmethyl) pyrrolidin-1-yl] acetate (4)

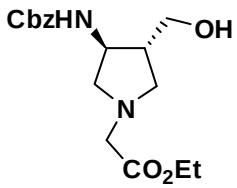


The crude compound **7** (80 mg, 0.18 mmol) was dissolved in 5 ml of methanol and was hydrogenated at 60 psi for 8 hrs after the addition of 10% Pd-C (20 mg). The reaction mixture thus obtained was filtered on celite bed and was concentrated to get crude amine which was dissolved in 5 ml of DCM, cooled in an ice bath and triethylamine (0.038ml, 0.27

mmol) was added followed by the addition of boc anhydride (43 mg, 0.20 mmol). The reaction mixture was stirred at room temperature for 12 hrs. The reaction mixture was diluted with 10 ml of DCM and washed with water, brine, dried over sodium sulphate and concentrated to obtain monomer ester **4** as a colourless liquid. Yield: (40 mg, 54%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.10-7.01 (s, 1H, -CH Thymine), 5.63-5.03 (m, 1H, -NHBoc), 4.17 (q, J = 7.1 Hz, 2H, -CH_2 ester), 4.07-3.73 (m, 2H, -$\underline{CH_2}$-CO_2Et), 3.43-2.74 (m, 5H, -CH_2-Thymine, -CH_2 ring, -$\underline{CH}$-NHBoc), 2.75-2.15 (m, 3H, -CH_2 ring -CH), 1.90 (s, 3H, -CH_3 Thymine), 1.39 (s, 9H, *t*-boc), 1.27 (t, J = 7.2 Hz, 3H) **^{13}C NMR** (50 MHz, $CDCl_3$): δ 170.3, 155.4, 151.8, 146.9, 139.1, 109.6, 79.4, 60.6, 59.4, 55.5, 53.1, 51.7, 28.3, 14.1, 12.9. ; **MS (EI)** *m/z* 410.22 Found 411.75 [M + 1], 433.95 [M + Na].

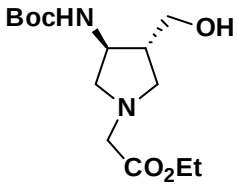
Ethyl 2-[((3*S*,4*R*)-3-*cbz* amino-4-hydroxymethyl) pyrrolidin-1-yl] acetate (1)



Compound **19** (200 mg, 0.57 mmol) was taken in DCM (2 ml) and to it was added TFA (2 ml) dropwise on cooling. The reaction was stirred at room temperature for 30 min. after which the TLC indicates the complete deprotection of *t-boc* group. The reaction mixture was concentrated and resulting mixture was co-evaporated with DCM (7 ml) three times to remove the TFA. The mixture obtained was taken in acetone:water (2:1, 10 ml) and to it was added sodium bicarbonate (143.5 mg, 1.7 mmol) on cooling followed by ethybromoacetate (104.6 mg, 0.62mmol). The reaction was stirred at room temperature for 18 hrs after which the acetone was removed and the mixture in water was extracted with ethylacetate three times. The combined ethyl acetate layer was drived over sodium sulphate and purified by column chromatography to obtain compound **1** as light yellow oil. Yield: (120mg, 63%)

^{1}H NMR (200 MHz, $CDCl_3$): δ 7.34 (s, 5H, *cbz*Ph), 5.76 (br s, 1H, -$\underline{NH}$*cbz*), 5.09 (s, 2H, *cbz* -CH_2), 4.17 (q, J = 7.2 Hz, 2H, -CH_2 ester), 4.06-3.92 (m, 1H, -OH), 3.61 (d, J = 5.6 Hz, 2H, -$\underline{CH_2}$-OH), 3.26 (d, J = 6.9 Hz, 2H, -$\underline{CH_2}$-CO_2Et), 3.18-3.07 (m, 1H, -$\underline{CH}$-NH*cbz*), 2.95-2.70 (m, 3H, -CH_2 ring, -CH), 2.35-2.04 (m, 2H, -CH_2 ring), 1.27 (t, J = 7.1 Hz, 3H); **MS (EI)** *m/z* 336.17 Found 337.51 [M + 1], 359.49 [M + Na].

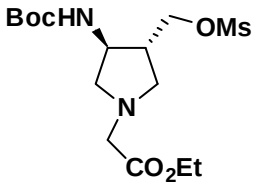
Ethyl 2-[((3*S*,4*R*)-3-*boc* amino-4-hydroxymethyl) pyrrolidin-1-yl] acetate (2)



To compound **1**, (120 mg, 0.36 mmol) in methanol (10 ml), 10% Pd-C (25mg) was added to make a suspension. The reaction mixture was subjected to hydrogenation under 60 *psi* pressure of H_2 gas for 8 hrs in parr-hydrogen apparatus. The reaction mixture was filtered through celite and the filtrate was concentrated under reduced pressure. The reaction mixture was then taken in dry DCM (5 ml) and to it was added Et_3N (54.1 mg, 0.53 mmol) followed by Boc anhydride (86mg, 0.39 mmol). The reaction was stirred at room temperature for 12 hrs after which this was washed with water, followed by saturated brine solution. The organic layer was dried over anhydrous Na_2SO_4 and concentrated under vacuum. The crude material was purified by silica gel chromatography afford N-*boc* protected alcohol compound **2** as a colourless thick oil. Yield: (60 mg, 56%)

^{1}H NMR (200 MHz, $CDCl_3$): δ 5.17-5.02 (br s, 1H, -NH*boc*), 4.18 (q, J = 7.2 Hz, 2H, -CH_2 ester), 3.91 (br s, 1H, -OH), 3.67-3.51 (m, 2H, -CH_2-OH), 3.42-3.07 (m, 3H, -CH_2-CO_2Et, -CH-NH*boc*), 2.95-2.70 (m, 2H, -CH_2 ring), 2.42-1.91 (m, 3H, -CH_2 ring, -CH), 1.44 (s, 9H, *boc*), 1.31-1.19 (m, 3H, -CH_3).

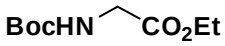
Ethyl 2-[((3*S*,4*R*)-3-*boc*amino-4-methyl sulfonyloxymethyl) pyrrolidin-1-yl] acetate (3)



The solution of alcohol **2** (100 mg, 0.33 mmol) and triethyl amine (133 mg, 1.32mmol) in dry dichloromethane (5 ml) was cooled to 0^{o}C on ice bath under Argon. While stirring methanesulfonyl chloride (75.8 mg, 0.66 mmol) was added dropwise over a period of 3 hrs at 0^{o}C. The reaction was diluted with DCM (20ml) and washed with water and brine. The organic layer was dried over sodium sulphate and purified by column chromatography to obtain mesyl compound **3**. Yield: (56mg, 45%).

^{1}H NMR (200 MHz, $CDCl_3$): δ 4.20 (q, J = 7.2 Hz, 2H, ester –CH_2), 3.34-3.31 (m, 2H, -CH_2-CO_2Et), 3.29-3.10 (m, 1H, -CH-NH), 3.05 (s, 3H, -CH_3 mesyl), 2.95-2.80 (m, 2H, -CH_2-OMs), 2.65-2.35 (m, 2H, -CH_2 ring), 2.20-1.90 (m, 2H, -CH_2 ring), 1.44 (s, 9H, *boc*), 1.31-1.19 (m, 3H, -CH_3).

Synthesis of N-*boc* glycine ethyl ester 14

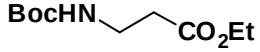


The Glycine (2 g, 27 mmol) was taken in ethanol (20ml) and to it was added thionyl chloride (2.4 ml, 32 mmol) dropwise on cooling under nitrogen. The reaction was stirred at room temperature for 30 min and then was refluxed for 12 hrs. The ethanol was removed under vacuum and the white solid thus obtained was taken in diethyl ether and stirred for 15 minutes. The diethyl ether was then decanted off to remove unreacted thionyl chloride to obtain glycine ethylester hydrochloride salt as a white solid.

The glycine ethylester hydrochloride salt was suspended in 30 ml of DCM and to it was added triethylamine (12.3 ml, 89 mmol) dropwise on cooling under nitrogen followed by boc-anhydride (6.1g, 28 mmol). The reaction mixture was stirred for 48 hrs at room temperature and after then was washed with water, brine and finally purified on silica gel column to obtain N-boc glycine ethylester **14** in excellent yield as colourless oil. Yield: (4.1 g, 78 %)

^{1}H NMR (200 MHz, $CDCl_3$): δ 5.65-5.50 (m, 1H, -NH), 4.20 (q, J = 7.1 Hz, 2H, -CH_2), 3.88 (d, J = 5.8 Hz, 2H, -CH_2), 1.45 (s, 9H, *t*-Boc), 1.28 (t, J = 7.1 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 170.6, 155.9, 80.0, 61.4, 42.6, 28.4, 14.3. ; **MS (EI)** *m/z* 203.12, Found 204.13 [M + H], 226.12 [M + Na].

Synthesis of N-*boc* *β*-alanine ethyl ester 16



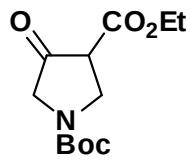
The synthesis of N-boc β-alanine ethyl ester **15** was achieved from commercially available β-alanine in a similar reaction sequence as for compound **14.** The compound **15** was obtained in quantitative yield as colourless oil.

^{1}H NMR (200 MHz, $CDCl_3$): δ 5.03 (br s, 1H, -NH), 4.10 (q, J = 7.1 Hz, 2H, -CH_2), 3.33 (br s, 2H, -CH_2), 2.45 (t, J = 6.1 Hz, 2H, -CH_2), 1.38 (s, 9H, *t*-Boc), 1.20 (t, J = 7.1 Hz, 3H, -CH_3).

β-keto ester general procedure

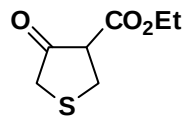
In a typical experiment, the ethyl esters (**14**, **15**, **16**) were taken in dry THF and exactly 1 eq of sodium hydride (60%) was added. The mixture was stirred at 0°C for half an hour followed by the addition of 1.5 eq of ethylacrylate in THF and was stirred at room temperature for 6-8 hrs. The solvent was evaporated under reduced pressure, the mixture obtained was taken in water and acidified with $KHSO_4$ to pH 6 and was extracted with ethyl acetate. The purification of the reaction mixture by flash chromatography resulted in the isolation of ketoesters (**8**, **9**, **10**) in good yields (85-90%). All the reactions can be scaled up to grams scale.

1-tert-butyl-3-ethyl-4-oxopyrrolidine-1,3-dicarboxylate 8



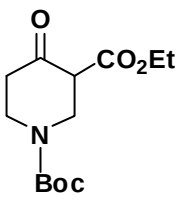
^{1}H NMR (200 MHz, $CDCl_3$): δ 4.35-4.10 (m, 4H, ester -CH_2, -CH_2), 3.90-3.75 (m, 2H, -CH_2), 3.65-3.45 (m, 1H, -CH), 1.40 (s, 9H, *t*-Boc), 1.21 (t, *J* = 7.1 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 168.4, 168.1, 167.1, 166.9, 154.3, 154.2, 81.1, 77.6, 62.4, 60.9, 51.5, 51.2, 48.9, 48.5, 28.6, 28.5, 14.5, 14.4. ; **MS (EI)** *m/z* 257.1, Found 258.0 [M + H], 279.0 [M + Na]

Ethyl-4-oxotetrahydrothiophene-3-carboxylate 9



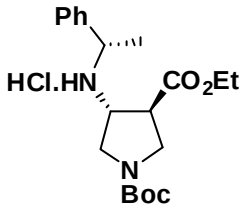
^{1}H NMR (200 MHz, $CDCl_3$): (keto:enol tautomer 3:2 approximately) δ 11.03 (s, 0.5 H, enolic –OH), 4.26 (dq, *J* = 7.1 Hz, *J* = 7.1Hz, 2H, ester -CH_2), 3.98-3.72 (m, 2H, -CH_2), 3.66-3.05 (m, 2H, -CH_2), 1.31 (dt, *J* = 7.1 Hz, *J* = 7.1 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 205.7, 172.2, 169.1, 167.6, 99.3, 61.9, 60.6, 55.4, 37.4, 35.9, 31.3, 28.9, 14.1, 14.0; **MS (EI)** *m/z* 174.04, Found 197.23 [M + Na], 213.24 [M + K]

1-tert-butyl-3-ethyl-4-oxopiperidine-1,3-dicarboxylate 10



^{1}H NMR (200 MHz, $CDCl_3$): δ 4.19 (q, *J* = 7.1 Hz, 2H, ester -CH_2), 4.02 (s, 2H, -CH_2), 3.52 (t, *J* = 6.0 Hz, 2H, -CH_2), 2.51-2.25 (m, 2H, -CH_2), 1.43 (s, 9H, *t*-Boc), 1.26 (t, *J* = 7.2 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, $CDCl_3$): δ 203.0, 170.8, 170.2, 154.6, 80.2, 60.6, 40.4, 39.2, 28.9, 28.4, 14.3; **MS (EI)** *m/z* 271.14, Found 272.26 [M + H], 294.24 [M + Na].

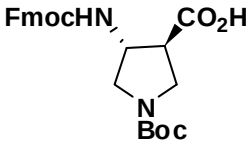
(3*S*,4*R*)-1-N-*boc*-3-((*R*)-1-phenyl ethyl amino)-4-ethoxy carbonyl pyrrolidine Hydrochloride[13] 17



To a stirred solution of β-ketoester **15** (16.0 g, 62.3 mmol) in absolute ethanol (250 mL) under N2 were added (R)-(+)-α-methylbenzylamine (16.0 mL, 124.5 mmol) and glacial acetic acid (7.1 mL, 124.5 mmol) to obtain a cloudy solution. The reaction mixture was stirred at room temperature until the formation of the enamine was complete (6 h, monitored by TLC). Sodium cyanoborohydride (16.5 g, 249.2 mmol) was added to the reaction mixture at room temperature, and the resulting solution was heated to 75 °C and stirred for 14 h under N2. The ethanol was removed via rotary evaporation. Water (250 mL) was added. The mixture was extracted three times with diethyl ether. The combined organic extracts were washed with brine, dried over sodium sulphate and concentrated to give a colorless oil. The oil was filtered through a silica gel column to obtain a colorless oil as a mixture of diasteromers. The oil was dissolved in ethyl acetate (250 mL), and 4 N HCl in dioxane (15.6 mL) was added dropwise at room temperature. The resulting solution was cooled to 0 °C and allowed to stand for 3 h at 0 °C. A precipitate formed during this time. The solid was filtered and washed two times with 100 mL portions of ethyl acetate to provide the desired compound **16** as white solid which was recrystallized with acetonitrile. Yield: (6.1 g, 25%).

$[\alpha]_D^{25}$ = +4.6 (*c* 1.0, MeOH); M. P. = 187- 188°C. **^{1}H NMR** (200 MHz, DMSO-d^6): δ 10.15 (br s, 2H, -NH_2), 7.68-7.58 (m, 2H, Ph), 7.48-7.35 (m, 3H, -Ph), 4.58-4.44 (m, 1H, C<u>H</u>-Ph), 4.02 (q, J = 7.0 Hz, 2H, ester –CH_2), 3.85-3.65 (m, 3H, ring –CH_2 & -CH), 3.55-3.25 (m, 3H, ring –CH_2 & -CH), 1.62 (d, J = 6.5 Hz, 3H, -CH_3) 1.34 (s, 9H, *t*-boc), 1.10 (t, *J* = 7.0 Hz, 3H, -CH_3). **^{13}C NMR** (50 MHz, DMSO-d^6): δ 170.9, 153.4, 136.9, 129.6, 129.4, 128.5, 79.8, 61.8, 57.0, 56.1, 48.4, 47.1, 45.4, 28.4, 20.4, 14.2. ; **MS (EI)** *m/z* 362.22 (without HCl), Found 363.41 [M + H].

(3*R*,4*S*)-1-N-*boc*-3-N-*fmoc*-4-hydroxycarbonylpyrrolidine (18)



Compound 2 (1.39 g, 3.49 mmol) was dissolved in THF/MeOH/H2O (6/3/1, v/v/v, 40 mL), added LiOH.H2O (732 mg, 17.4 mmol) on cooling and the mixture was stirred at

0 °C for 3 h. Aqueous HCl (1 N, 18 mL) was added at 0 °C. The solvent was then removed on a vacuum rotary evaporator to give a white solid which was dissolved in 150 mL of 95% ethanol in a hydrogenation flask. Pd-C (10%, 1.1 g) was added and the resulting mixture was shaken under H_2 (45 psi) for 24 h. After the reaction was complete (disappearance of starting material, as monitored by TLC), the mixture was filtered through Celite, and the filtrate was concentrated to obtain a white solid. This solid was dissolved in acetone/H2O (2/1, v/v, 150 mL) and cooled to 0 °C, Fmoc-Osu (1.53 g, 4.54 mmol) and NaHCO3 (2.93 g, 34.9 mmol) were added. The reaction mixture was stirred at 0 °C for 1 h and then allowed to stir at room temperature for 12 hrs. Water (50 mL) was added to the reaction mixture. The acetone was removed under reduced pressure. The aqueous layer was stirred for 1 h with diethyl ether (50 mL), the layers were separated, and the aqueous layer was acidified with 1 N aqueous HCl, extracted with ethyl acetate, dried over Na_2SO_4, and concentrated to give a foamy white solid. Yield: (0.9 g, 60%).

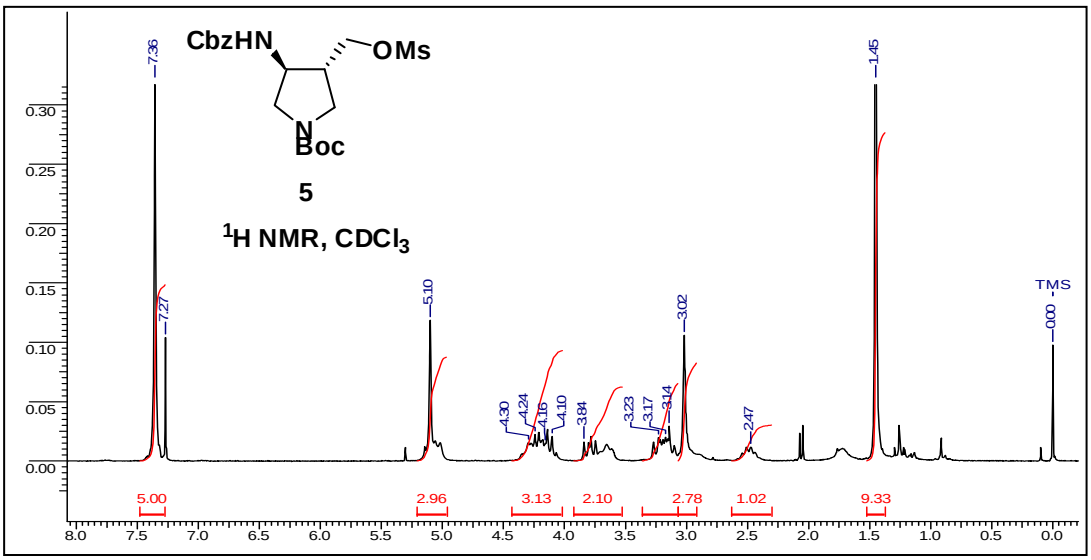
$[\alpha]_D^{25}$ = +18.0 (*c* 1.2, MeOH); M. P. = 110-112°C. **^{1}H NMR** (200 MHz, DMSO-d^6): *δ* 7.68 (d, *J* = 7.2 Hz, 2H, fmoc), 7.57-7.41 (m, 2H, fmoc), 7.32 (t, J = 7.2 Hz, 2H, fmoc), 7.28-7.15 (m, 2H, fmoc), 5.85-5.50 (m, 1H, -NH), 4.48-4.00 (m, 4H, fmoc 3H & -CH-NH), 3.83-2.62 (m, 5H, 2 x ring –CH_2, -CH), 1.41 (s, 9H, *t*-boc). **^{13}C NMR** (50 MHz, DMSO-d^6): *δ* 175.5, 156.2, 154.4, 143.8, 143.6, 141.2, 128.8, 127.7, 127.0, 125.0, 119.9, 80.2, 66.9, 53.4, 53.0, 46.9, 28.3.; **MS (EI)** *m/z* 452.19, Found 475.82 [M + Na].

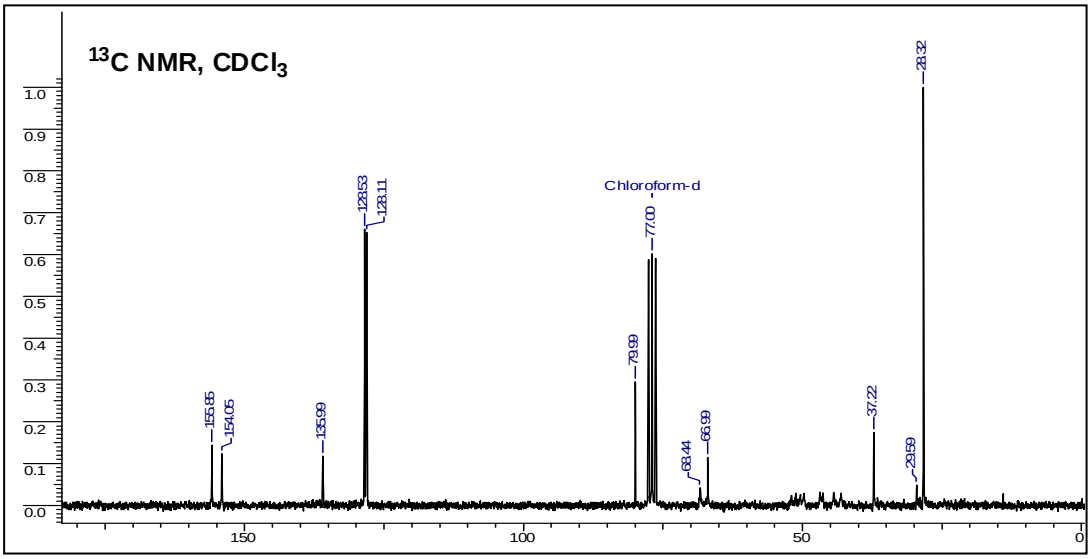
4.9 References

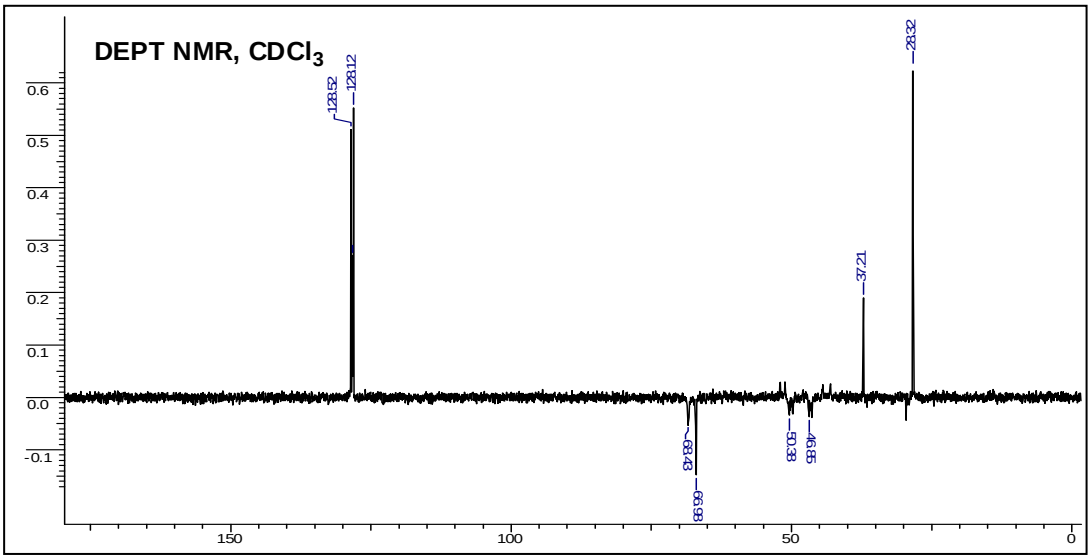
1. (a) Nielsen, P. E.; Egholm, M.; Buchardt, O. *Science* **1991**, *254*, 1497. (b) Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem*. **2000**, *4*, 931.
2. Nielsen, P. E. *Chem. & Biodiv.* **2010**, *7*, 786.
3. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed.* **1996**, *35*, 1939.
4. Dragulescu-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. Commun.* **2005**, 244.
5. Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; Keefe, H. O.; Vatta, P.; Hyldig-Nielsen, J. J.; Ly, D. H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4931.
6. Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J Am. Chem. Soc.* **2006**, *128*, 10258.
7. Rapireddy, S.; He, G.; Roy, S.; Armitage, B. A.; Ly, D. H. *J. Am. Chem. Soc.* **2007**, *129*, 15596.
8. Kumar, V.A.; Ganesh, K. N. *Acc. chem. res.* **2005**, 38, 404.
9. Lonkar, P. S.; Kumar V. A.; Ganesh K. N. *Organic & Bio. Chem*. **2004**, *2*, 2604.
10. Jung M.; Hong, J. H.; Cho, J.; Oh, C. *Arch. Pharm. Chem. Life Sci.* **2008,** *341*, 780.
11. a) Leanza, W.; Wildonger, K.; Miller, T. W.; Christensen, B. G. *J. Med. Chem.* **1979**, *22*, 1435. b.) Birnbaum, J.; Kahan, F. M.; MacDonald J. S. *Am. J. Med.* **1985**, *78*, (Suppl. 6A), 3.
12. Wang, G. T.; Chen, Y.; Wang, S.; Gentles, R.; Sowin, T.; Kati, W.; Muchmore, S.; Giranda, V.; Stewart, K.; Sham, H.; Kempf, D.; Laver, W. G. *J. Med. Chem.* **2001**, *44*, 1192.
13. Lee, H. S.; LePlae, P. L.; Porter, E. A.; Gellman S. H. *J. Org. Chem.* **2001**, *66*, 3597.
14. Bunnage, M. E.; Davies, S. G.; Roberts, P. M.; Smith, A. D.; Withey, J. M. *Org. Biomol. Chem.*, **2004**, *2*, 2763.
15. Thorbeck, P.; Hjeds, H.; Schaumburg, K. *Acta chem. scand.,* **1981**, *B 35*, 473.

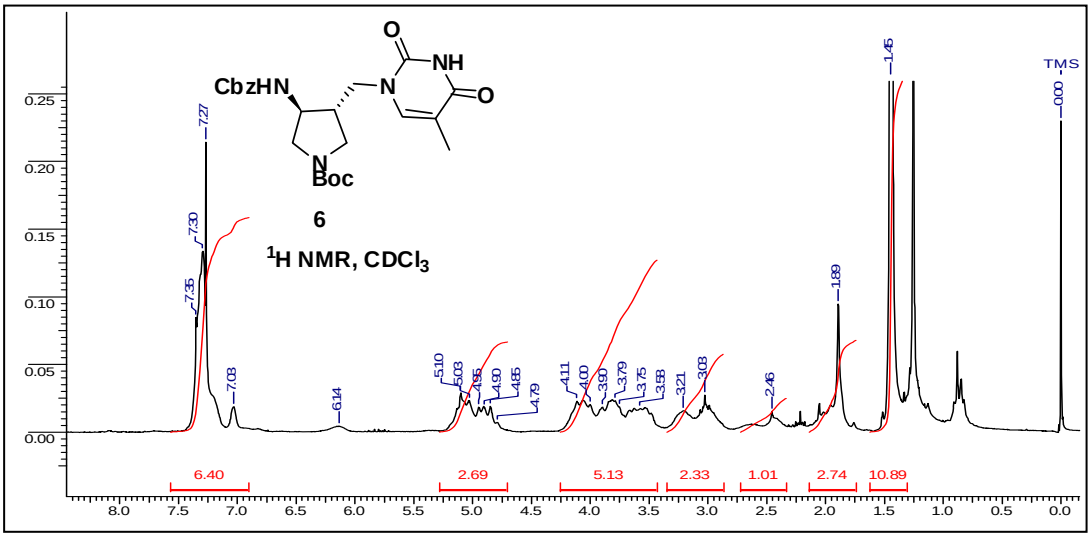
16. Stoll, V.; Stewart, K. D.; Maring, C. J.; Muchmore, S.; Giranda, V.; Gu, Y. Y.; Wang, G.; Chen, Y.; Sun, M.; Zhao, C.; Kennedy, A. L.; Madigan, D. L.; Xu, Y.; Saldivar, A.; Kati, W.; Laver, G.; Swoin, T.; Sham, H. L.; Greer J.; Kempf D. *Biochemistry* **2003,** *42,* 718.
17. Wang, X.; Espinosa, J. F.; Gellman S. H. *J. Am. Chem. Soc.*, **2000**, *122*, 4821.
18. *GABA in Nervous System Function, ed.* E. Roberts, T. N. Chase and D. B. Tower, Raven Press, New York, **1976**; *GABA Neurotransmitters. Pharmacochemical, Biochemical and Pharmacological Aspects, ed.* P. K. Larsen, S. Kruger and H. Kofod, Munksgaard, Copenhagen, **1979**.
19. *Amino Acids as Neurotransmitters, ed.* P. K. Larsen and F. I. Fonnum, Plenum, New York, **1978**.
20. Kuhn, R.; Osswald, G. *Chem. Ber.*, **1956**, *89*, 1423.
21. Sunagawa, M.; Masanoriitoh; Kubota, K.; Sasaki, A.; Ueda, Y.; Angehrn, P.; Bourson, A.; Goetschi, E.; Hebeisen, P.; Then, R. L. *Journal of Antibiotics*, **2002**, *55*, 722.
22. Mcelvain, S. M.; Stork, G. *J. Am. Chem. Soc.*, **1946**, *68*, 1049.

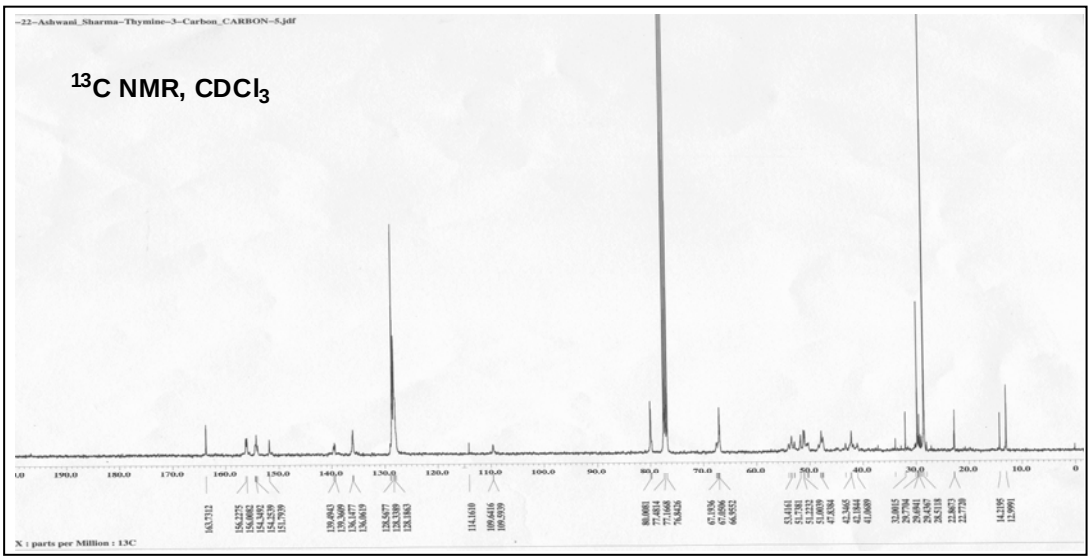
4.10 APPENDIX

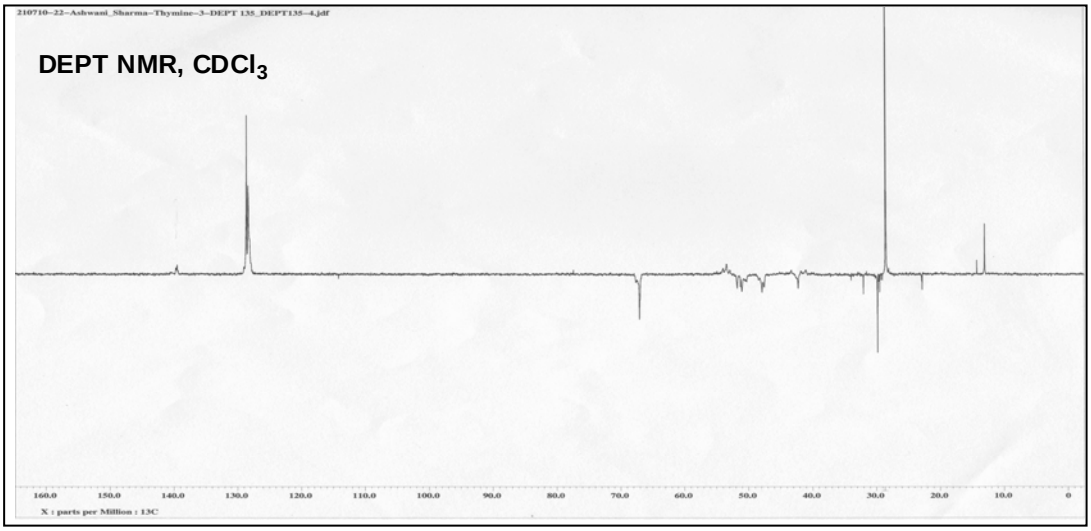

CbzHN
OMs
N
Boc
5
1H NMR, CDCl3
TMS
5.00
2.96
3.13
2.10
2.78
1.02
9.33

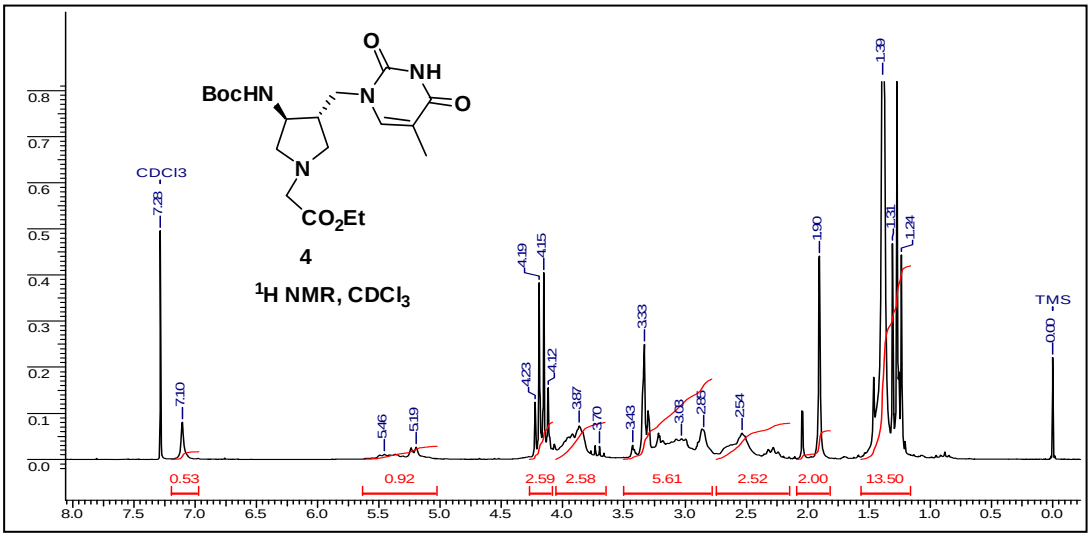

13C NMR, CDCl3
Chloroform-d

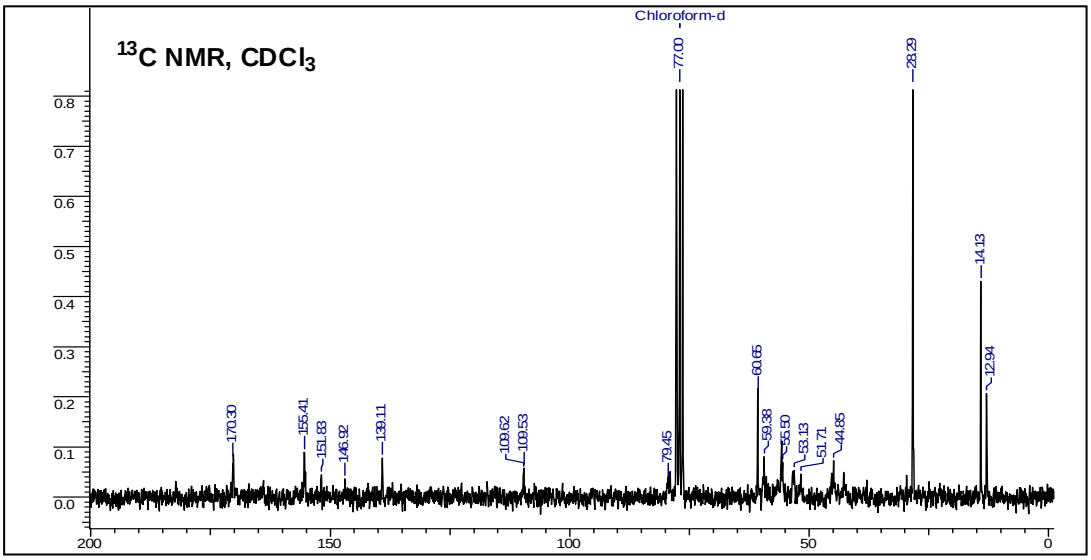

DEPT NMR, CDCl3

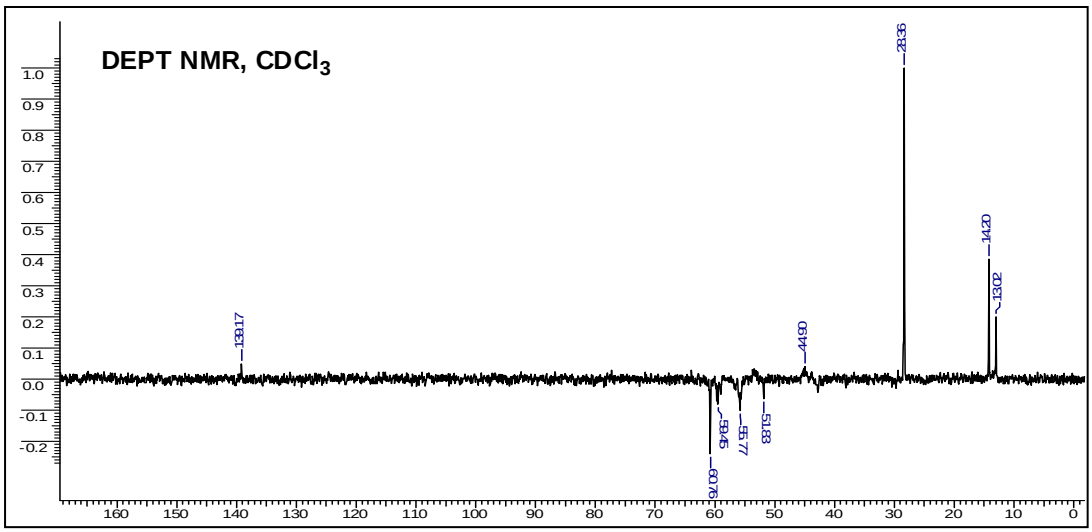

CbzHN
NH
N
O
O
N
Boc
6
^{1}H NMR, CDCl$_3$
TMS
6.40
2.69
5.13
2.33
1.01
2.74
10.89

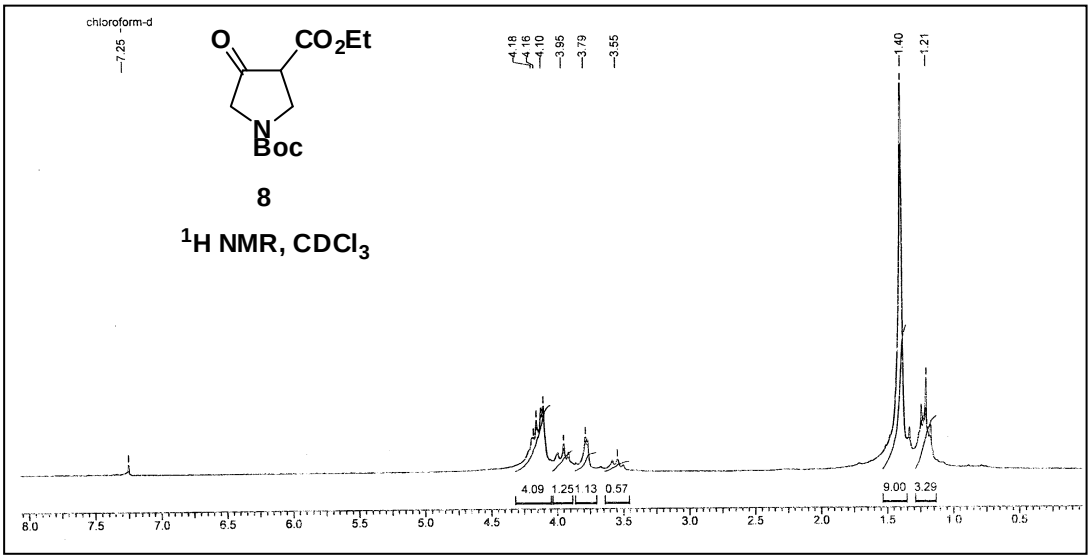

^{13}C NMR, CDCl$_3$


DEPT NMR, CDCl$_3$


BocHN
NH
O
O
N
N
CO2Et
4
1H NMR, CDCl3
CDCl3
TMS
0.53
0.92
2.59
2.58
5.61
2.52
2.00
13.50

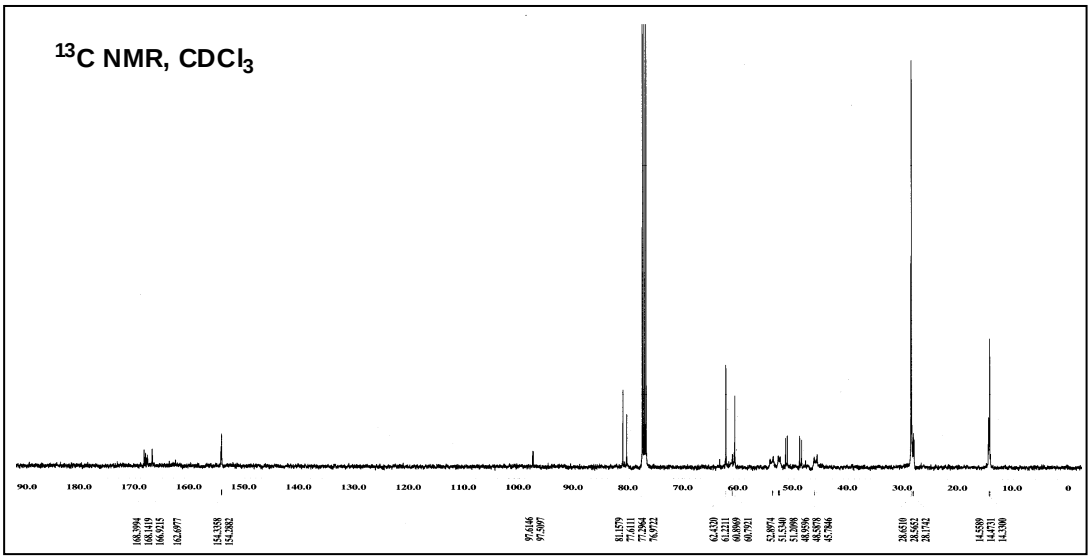

13C NMR, CDCl3
Chloroform-d


DEPT NMR, CDCl3

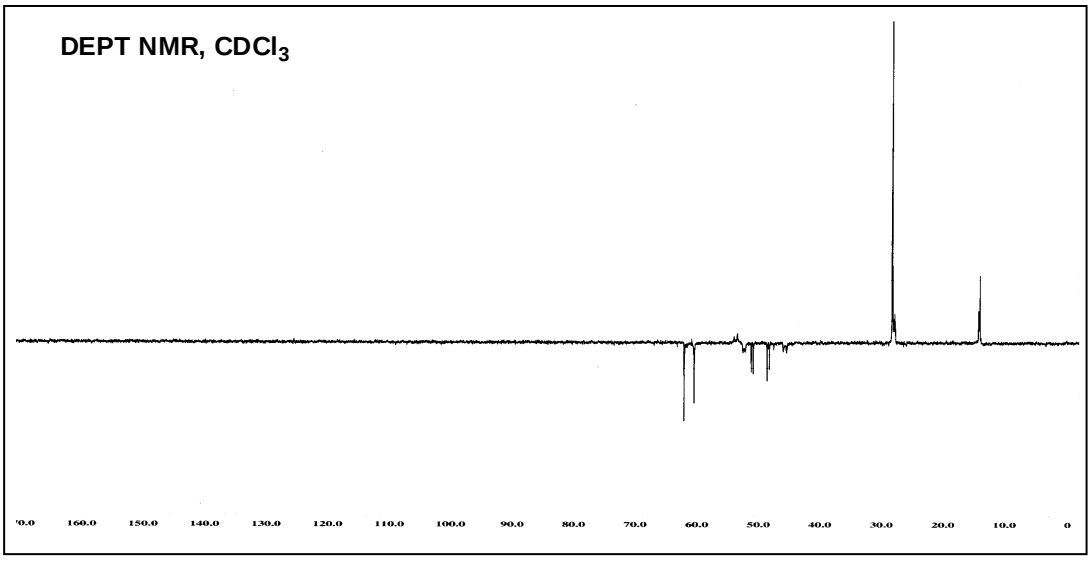


8

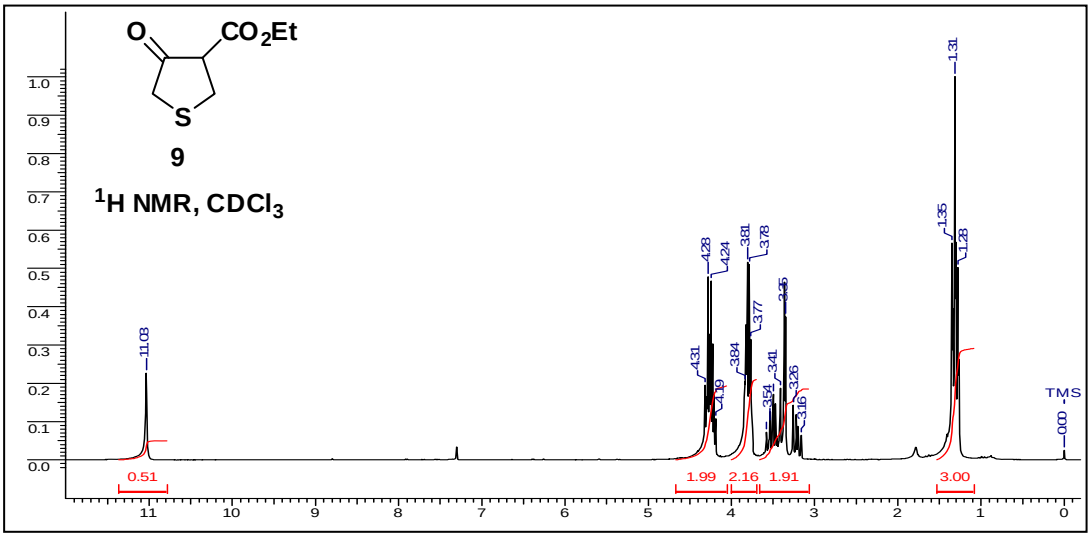
^{1}H NMR, $CDCl_3$

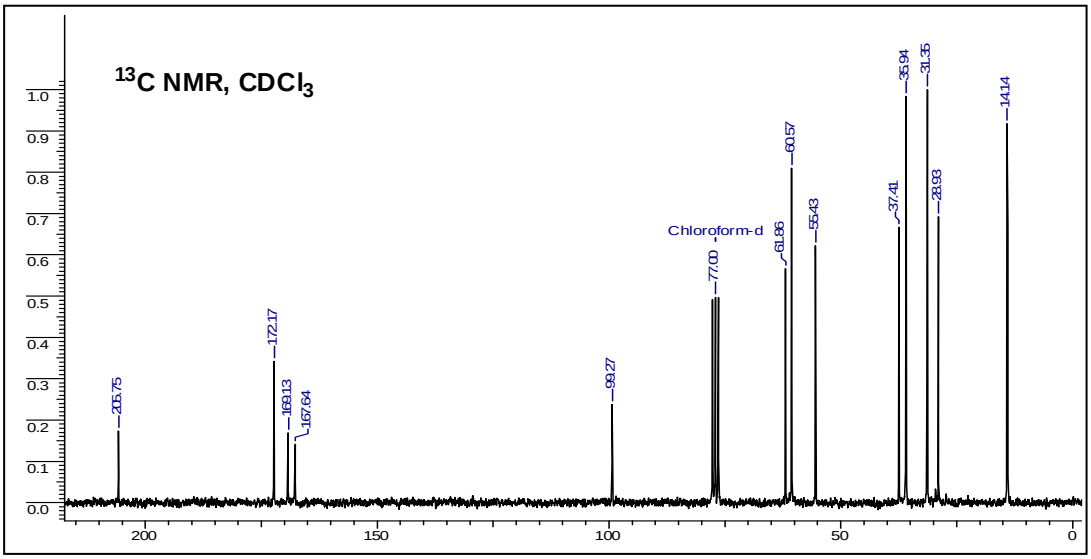


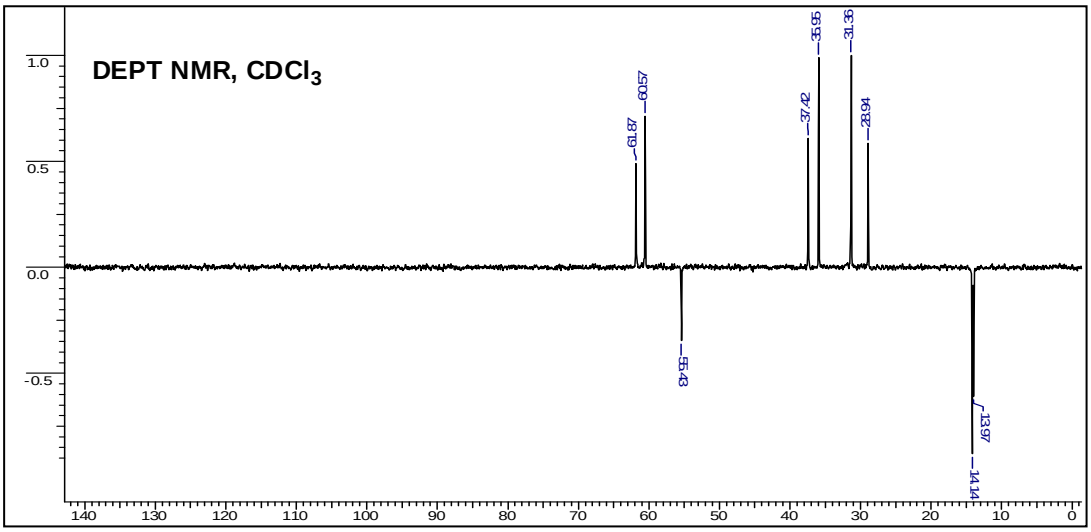
^{13}C NMR, $CDCl_3$

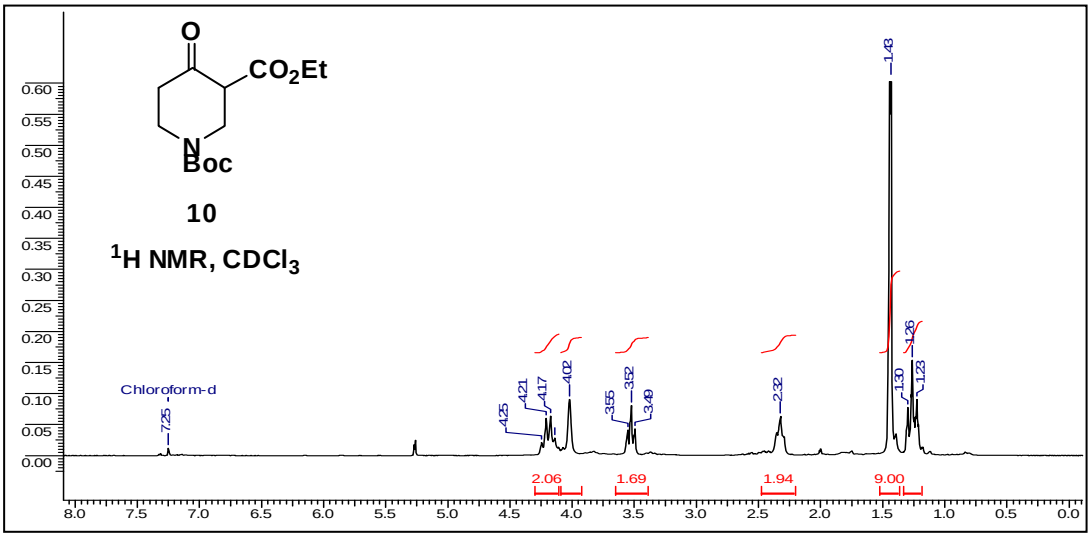


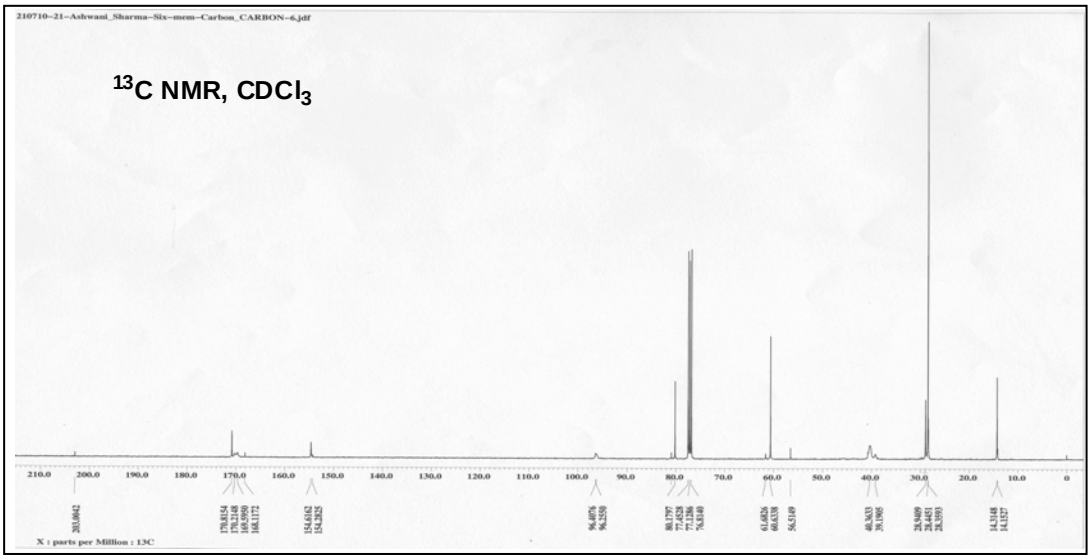
DEPT NMR, $CDCl_3$

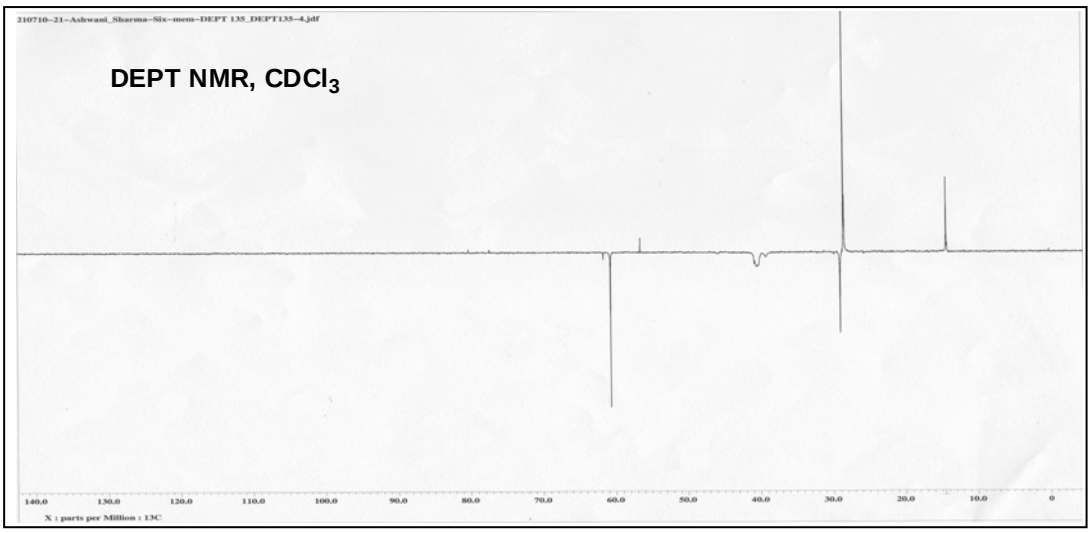

O
CO2Et
S
9
1H NMR, CDCl3
11.03
4.28
4.24
3.81
3.78
3.77
3.41
1.35
1.28
1.31
TMS
0.00
0.51
1.99
2.16
1.91
3.00

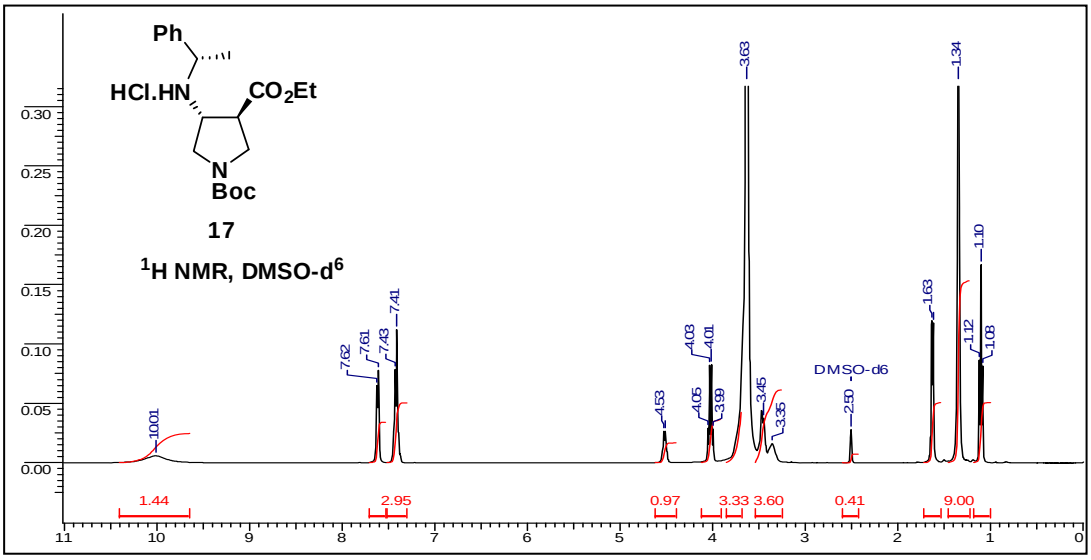

13C NMR, CDCl3
205.75
172.17
169.13
167.64
99.27
Chloroform-d
77.00
61.86
60.57
55.43
37.41
35.94
31.35
28.93
14.14

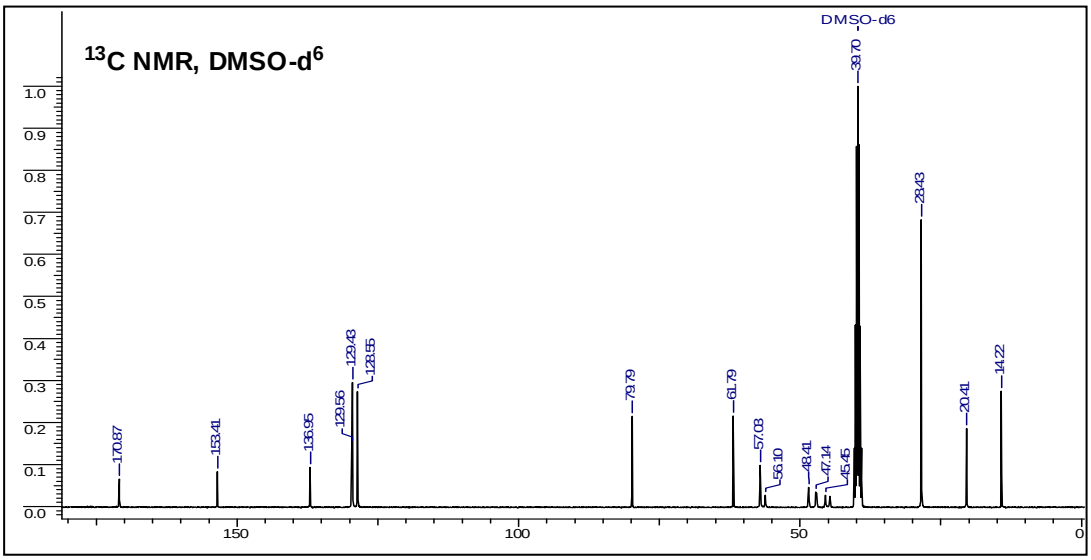

DEPT NMR, CDCl3
61.87
60.57
55.43
37.42
35.95
31.36
28.94
13.97
14.14

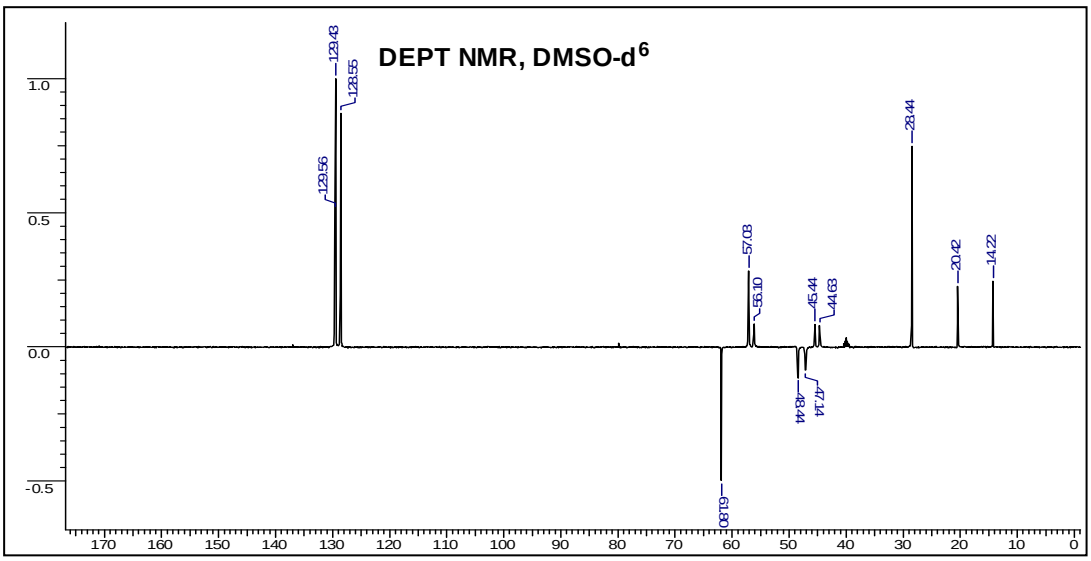

O
CO2Et
N
Boc
10
1H NMR, CDCl3
Chloroform-d
7.25
4.25
4.21
4.17
4.02
3.55
3.52
3.49
2.32
1.43
1.26
1.30
1.23
2.06
1.69
1.94
9.00

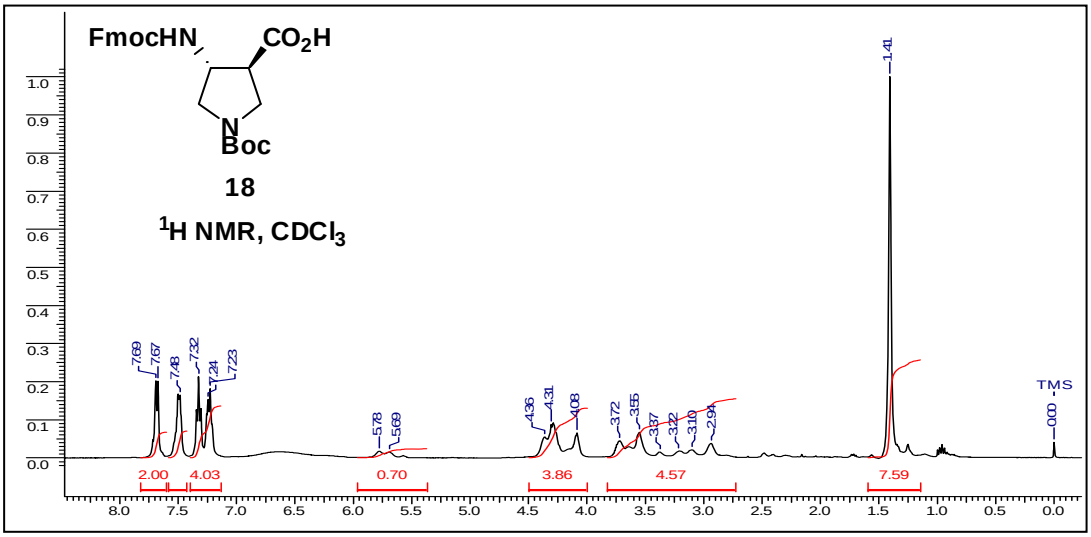

13C NMR, CDCl3

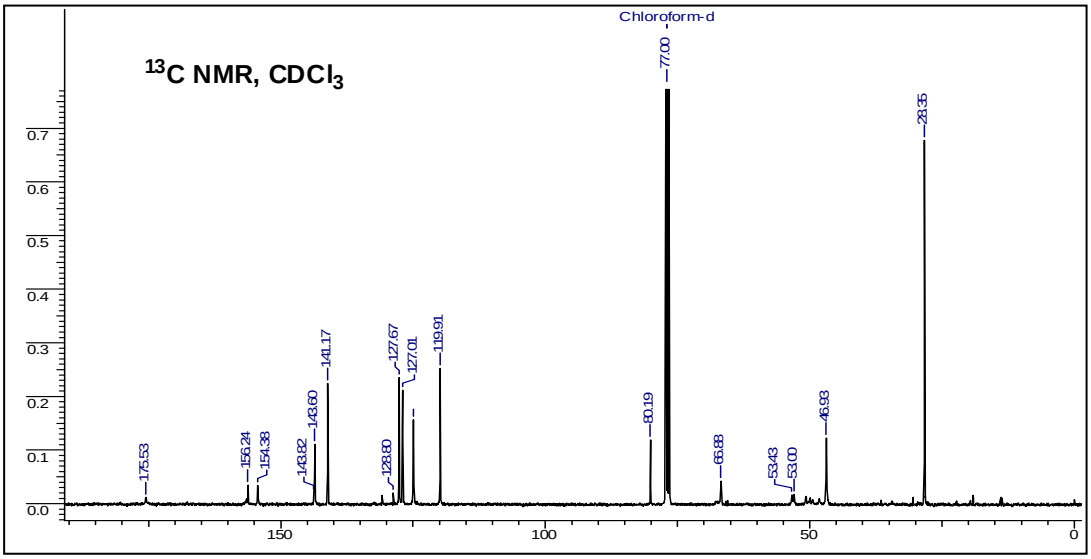

DEPT NMR, CDCl3

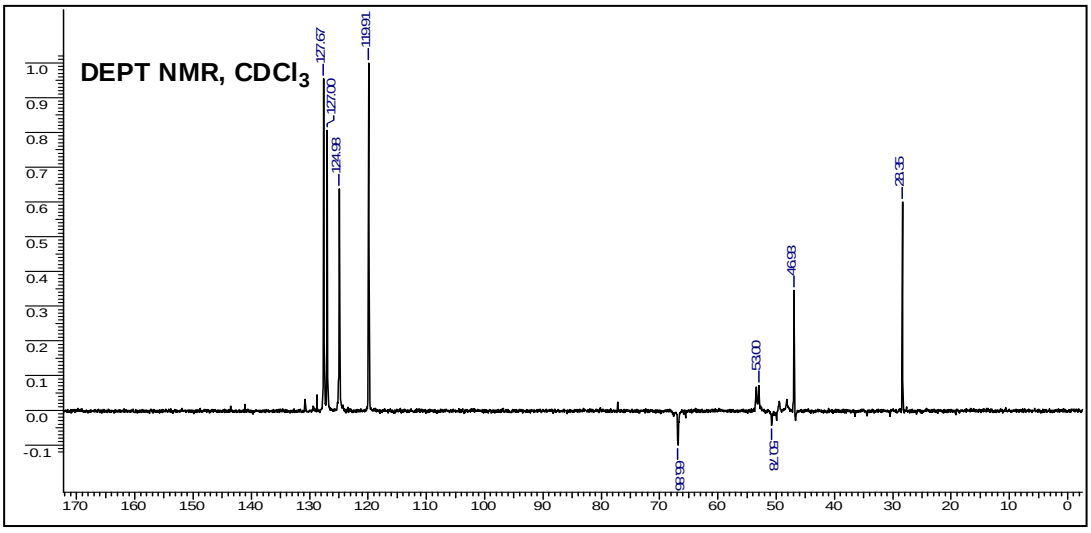

Ph
HCl.HN
CO2Et
N
Boc
17
1H NMR, DMSO-d6
DMSO-d6
7.62
7.61
7.43
7.41
4.03
4.01
3.99
4.05
4.53
3.63
3.45
3.35
2.50
1.63
1.34
1.12
1.10
1.08
1.00
1.44
2.95
0.97
3.33
3.60
0.41
9.00

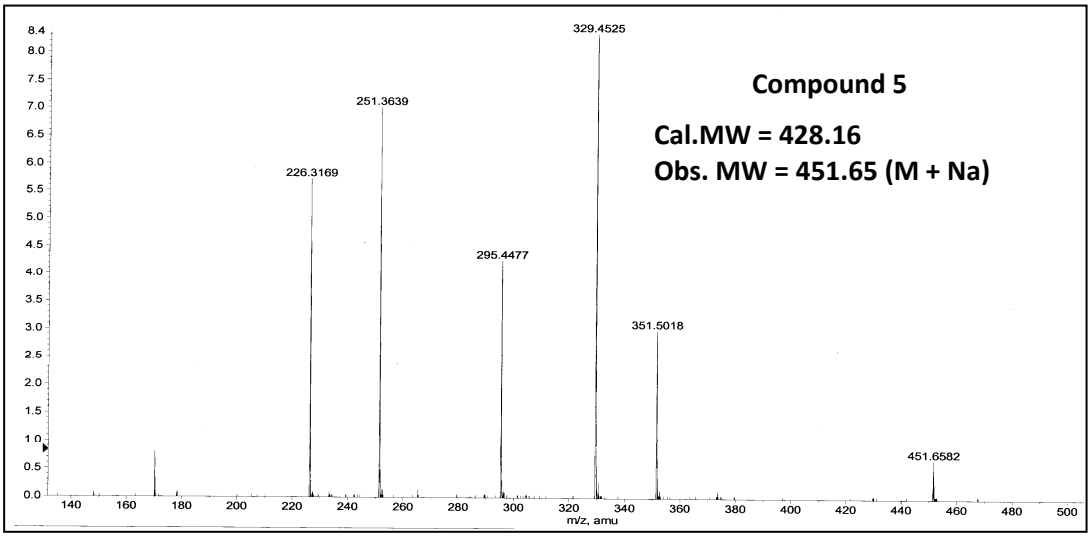

13C NMR, DMSO-d6
DMSO-d6
39.70
28.43
170.87
153.41
136.95
129.56
129.43
128.55
79.79
61.79
57.08
56.10
48.41
47.14
45.45
20.41
14.22

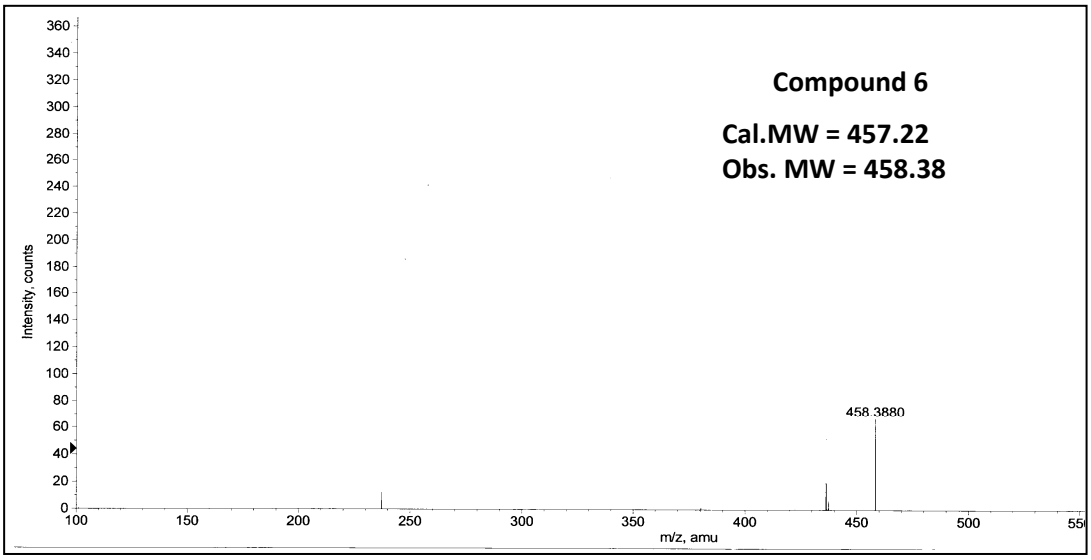

DEPT NMR, DMSO-d6
129.43
128.55
129.56
28.44
57.08
56.10
45.44
44.63
20.42
14.22
48.41
47.14
61.80

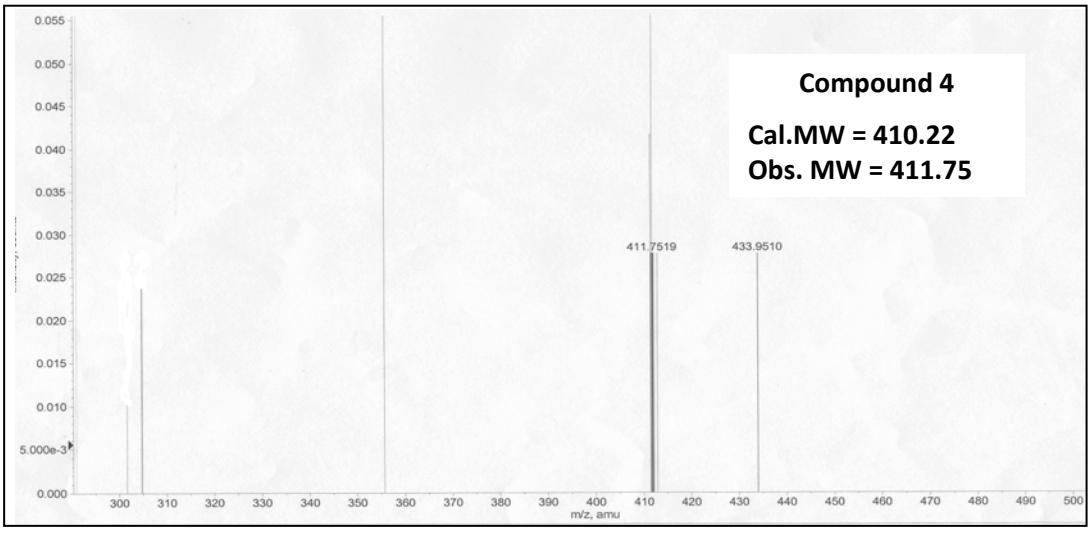

FmocHN
CO2H
N
Boc
18
1H NMR, CDCl3
TMS
2.00
4.03
0.70
3.86
4.57
7.59

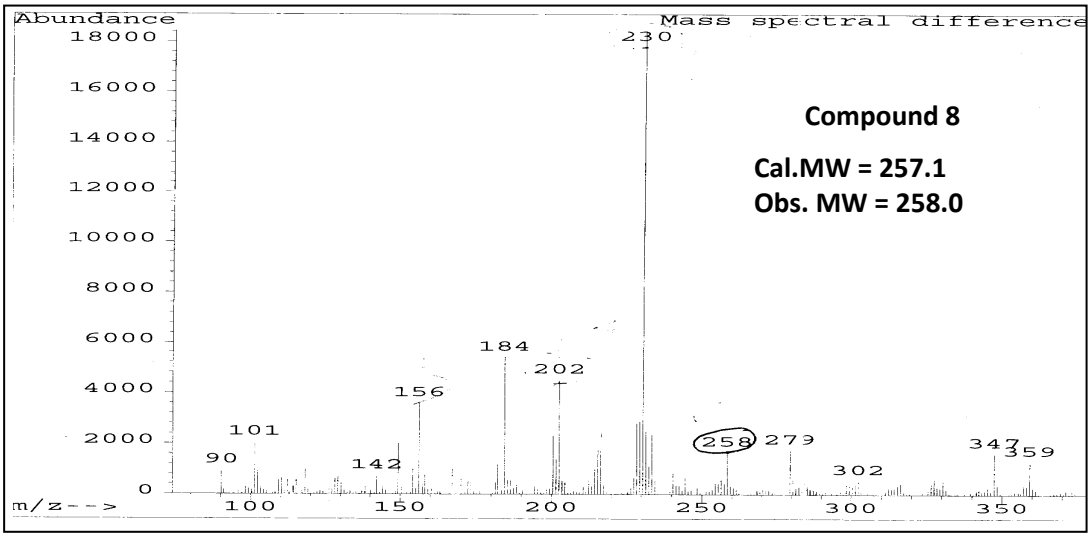

13C NMR, CDCl3
Chloroform-d

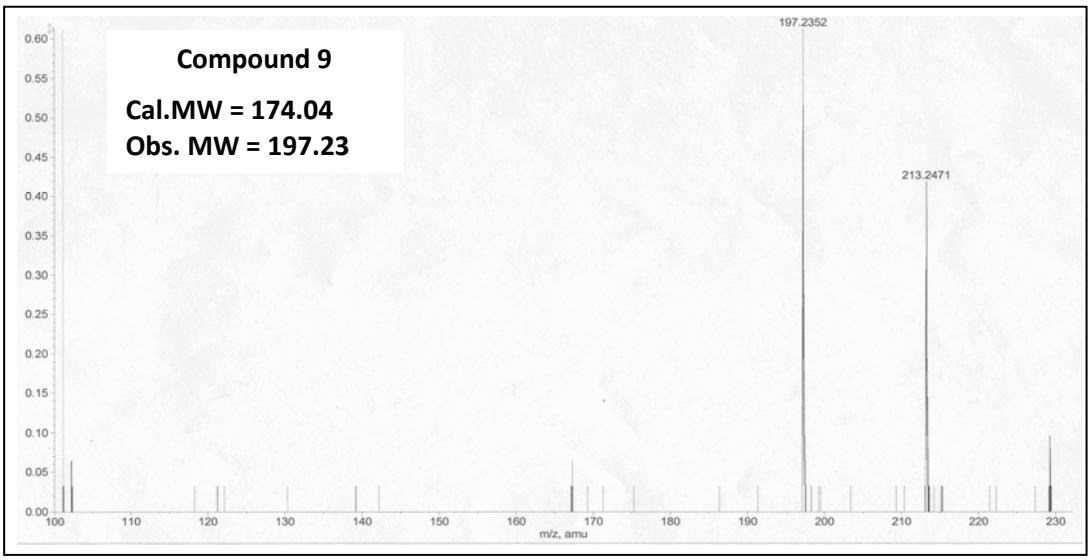

DEPT NMR, CDCl3

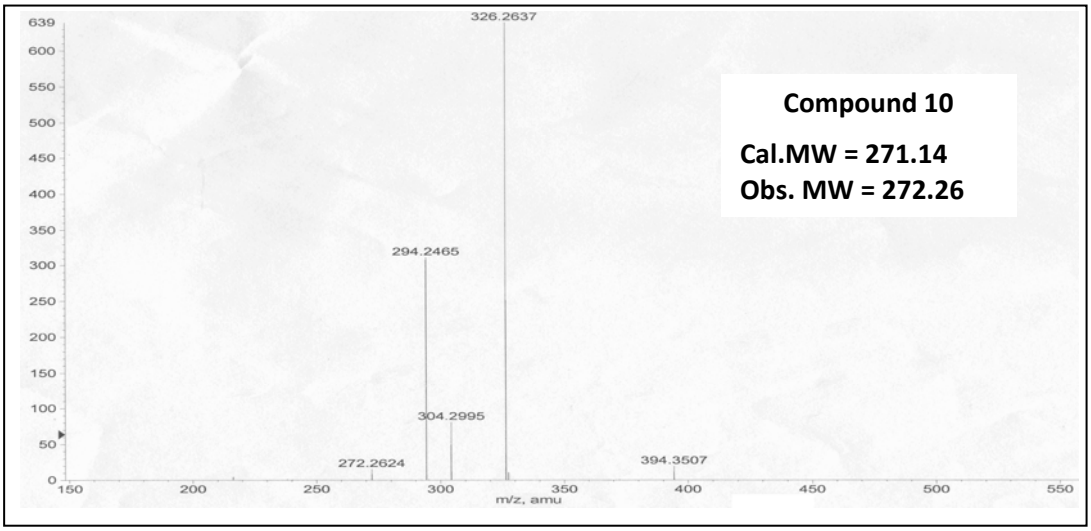

Compound 5
Cal.MW = 428.16
Obs. MW = 451.65 (M + Na)
329.4525
251.3639
226.3169
295.4477
351.5018
451.6582
m/z, amu

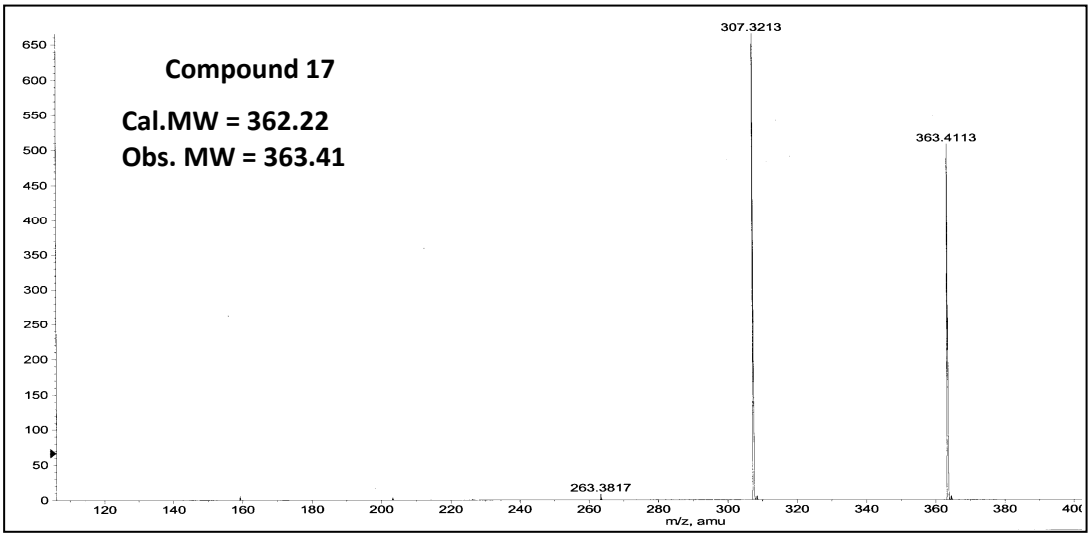

Compound 6
Cal.MW = 457.22
Obs. MW = 458.38
Intensity, counts
458.3880
m/z, amu

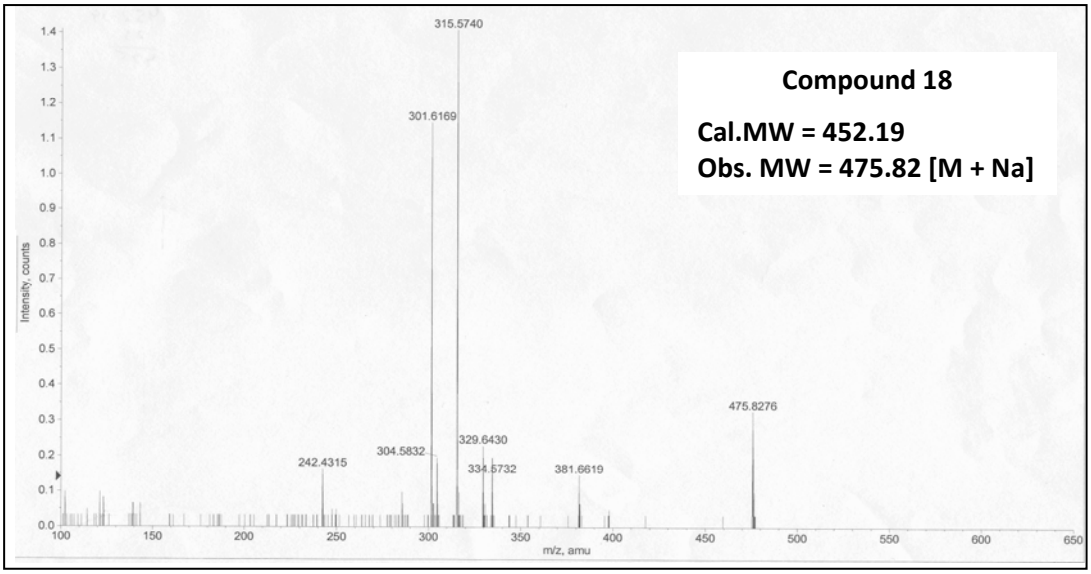

Compound 4
Cal.MW = 410.22
Obs. MW = 411.75
411.7519
433.9510
m/z, amu


Abundance
Mass spectral difference
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
m/z-->
100
150
200
250
300
350
230
184
202
156
101
90
142
258
279
302
347
359
Compound 8
Cal.MW = 257.1
Obs. MW = 258.0

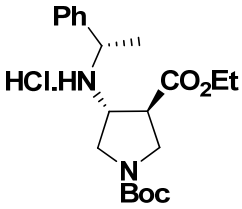

Compound 9
Cal.MW = 174.04
Obs. MW = 197.23
197.2352
213.2471
m/z, amu


Compound 10
Cal.MW = 271.14
Obs. MW = 272.26
326.2637
294.2465
304.2995
272.2624
394.3507
m/z, amu


Compound 17
Cal.MW = 362.22
Obs. MW = 363.41
307.3213
363.4113
263.3817
m/z, amu

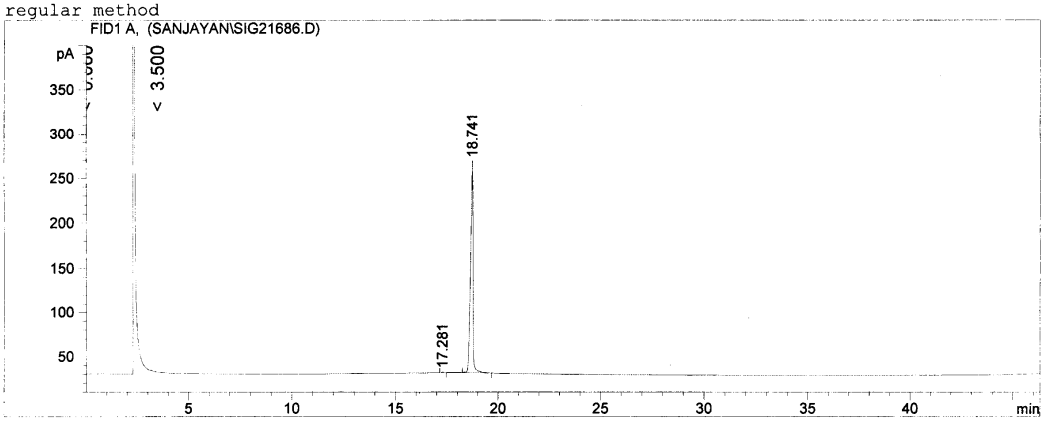

Compound 18
Cal.MW = 452.19
Obs. MW = 475.82 [M + Na]
315.5740
301.6169
475.8276
329.6430
304.5832
242.4315
334.5732
381.6619
Intensity, counts
m/z, amu

GC Analysis to check diastereomeric excess



Compound 17

```
Injection Date  : 3/19/2008 11:52:14 AM
Sample Name     : AKS-1                              Location : Vial 2
Acq. Operator   : BORIKAR                                 Inj :   1
                                                    Inj Volume : Manually
Acq. Method     : C:\HPCHEM\1\METHODS\REGULAR.M
Last changed    : 3/19/2008 11:50:33 AM by BORIKAR
                  (modified after loading)
Analysis Method : C:\HPCHEM\1\METHODS\REGULAR.M
Last changed    : 3/19/2008 12:39:20 PM by BORIKAR
                  (modified after loading)
regular method
```




```
================================================================================
                            Area Percent Report
================================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000


Signal 1: FID1 A,

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [pA*s]      [pA]        %
----|-------|----|-------|----------|----------|--------|
   1  17.281 BB    0.1048    3.94946 5.84629e-1  0.19332
   2  18.741 BB    0.1126 2038.98938  236.54092 99.80668

Totals :                  2042.93884  237.12555
```