γ-C-Substituted Multifunctional Peptide Nucleic Acids: Design, Synthesis and Bioevaluation

A thesis

Submitted in partial fulfillment of the requirements

Of the degree of

Doctor of Philosophy

By

Deepak R. Jain

ID: 20093026

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE

2013

This Thesis is dedicated to...

My parents, who got me started...

And

My wife, who let me go on.

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान, पुणे

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH (IISER), PUNE

(An Autonomous Institution, Ministry of Human Resource Development, Govt. of India)

900 NCL Innovation Park, Dr. Homi Bhabha Road, Pune 411008

Prof. Krishna N. Ganesh
Professor and Coordinator, Chemistry
Director, IISER Pune
J. C. Bose Fellow (DST), NCL Pune

CERTIFICATE

Certified that the work incorporated in the thesis entitled "*γ-C-Substituted Multifunctional peptide Nucleic Acids: Design, Synthesis and Bioevaluation*" submitted by **Mr. Deepak R. Jain** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Date: 27th Dec. 2013, Pune

Prof. Krishna N. Ganesh

(Research Supervisor)

DECLARATION

I declare that, this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea / data / fact/ source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: 27th Dec. 2013, Pune   **Mr. Deepak R. Jain**

                                                                              ID: 20093026

Acknowledgements

It gives me an immense pleasure to express my deep sense of gratitude towards my research supervisor Prof. Krishna N. Ganesh for his support and guidance. He has been the constant source of inspiration throughout my doctoral research. He has always taken efforts to instill independent thinking, writing skills, presentation skills and many more things to develop a researcher in me. He has been the 'real mentor' in my life; he has guided not only in scientific problems but also in personal difficulties.

I sincerely thank Director, IISER Pune for providing fantastic facilities to carry out doctoral research. I thank IISER, Pune for financial assistance and graduate research fellowship. I would also like to thank Directors of NCL, Dr. Sivaram and Dr. Sourav Pal for allowing me to work in NCL lab in early research period.

I am very grateful to my research advisory committee members Dr. H. N. Gopi and Dr. G. J. Sanjayan for their fruitful suggestions, comments and encouragement through various RAC meetings. I would like to specially thank Dr. Gopi who always gave me solution whenever I had problem regarding research. I would like to thank Dr. Mayurika Lahiri who agreed to carry cell permeation studies in her lab. I must acknowledge her student Libi for her help in biology experiments. I am very thankful to Amit and Hemangini from NCCS for their help in FACS analysis. I sincerely thank Dr. S. G. Srivatsan for his encouragement and support by one or the other way. I would also like to acknowledge Dr. Raghavendra Kikeri for his help and discussion regarding various ideas or aspects of research. I thank my friend Maroti for helping me in PAGE and Swati for MALDI-TOF. I sincerely thank Mayuresh (admin) for his constant support throughout my phD.

I admire the co-operation of my seniors, colleague and juniors for cheerful atmosphere in the lab. I thank my seniors Dr. Mahesh, Dr. Roopa, Dr. Shridhar, Dr. Ashwani, Dr. Gitali, Dr. Amit, Dr. Manaswini, and Dr. Tanpreet for their help in initial years of research in NCL. I thank Nitin, my colleague with whom I shared almost everything regarding the ups and downs of research. I thank my juniors cum colleagues Vijay, Satish, Madhan, Prabhakar, Shahaji and Pramod for always keeping healthy atmosphere around. My all labmates actually gave the relaxation needed in tough times of research.

I had a wonderful company with my friends Shekhar, Amar, Sachin, Prakash, Maroti, Sandip, Anupam, Arun, Pramod S., Dnyaneshwar, Sharad, Susheel, Alchemist team and Lab 101 who helped me to release the sort of tension in difficult times of research in one way or other. I cannot forget the time spent in

HR-2. I enjoyed tea sessions by Murthy with Somu and Amar. I thank JP (his mutter-paneer & rice) and Senthil for their time and help in initial days of HR-2 stay.

I shall always remain indebted to my parents and my entire family, for their unconditional love, blessings, sacrifices, patience and support. The values and virtues they have instilled in me have made me achieve whatever I have achieved so far. I would like to thank a very very special person my lovely wife, Jayshree for her continuous support and faith. Finally, I would like to thank my cute son Arpit. His smile will make me forget all the tension & frustration coming from the research. I hope with my hard work and dedication, I would be able to translate their dreams into reality.

Deepak

Table of Contents

Section B: Solid Phase Synthesis, Purification and Characterization of PNA Oligomers

Chapter 3: Biophysical Evaluation of Modified PNA Oligomers

Chapter 4: Cell Permeation Studies of PNA Oligomers

Abbreviations

A	Adenine
Abs.	Absolute
Ac_2O	Acetic anhydride
AcOH	Acetic acid (glacial)
ACN	Acetonitrile
aeg	Aminoethylglycine
aep	Aminoethylprolyl
ap	Antiparallel
APS	Ammonium persulphate
aq.	Aqueous
azb	Azidobutylene
azm	Azidomethylene
$(Boc)_2O$	Boc anhydride
BPB	Bromophenol blue
Bz	Benzoyl
C	Cytosine
Calcd	Calculated
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
CF/5(6)-CF	5(6)-Carboxyfluorescein
ch	Cyclohexyl
CHCA	α-cyano-4-hydroxycinnamic acid
cp	Cyclopentyl
CuAAC	Copper mediated azide-alkyne cycloaddition
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DCU	Dicyclohexyl urea
DHB	2,5-dihydroxybenzoic acid
DIC	N,N'-diisopropylcarbodiimide
DIPEA/DIEA	N,N-Diisopropylethylamine
DMAP	N,N-Dimethyl-4-aminopyridine
DMEM	Dulbecco's Modified Eagle Medium
DMF	N,N-dimethylformamide
DMSO	N,N-Dimethyl sulfoxide
DNA	2'-deoxyribonucleic acid
ds	Double stranded
eam	Ethyleneamino
EBA	Ethylbromo acetate
EDTA	Ethylene diamine tetraacetic acid
egd	Ethyleneguanidino

EMSA	Electrophoretic mobility shift assay
Et	Ethyl
EtBr	Ethidium Bromide
EtOAc	Ethyl acetate
FACS	Fluorescence Activated Cell Sorter
FBS	Fetal bovine serum
Fmoc	9-Fluorenylmethoxycarbonyl
g	gram
G	Guanine
gly	Glycine
h	Hours
his	Histidine
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3 tetramethyl-uronum-hexafluoro-phosphate
HIV	Human Immuno Difficiency Virus
HOBt	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
in situ	In the reaction mixture
in vivo	Within the living
IR	Infra red
$KBrO_3$	Potassium bromate
L-	Levo-
LC-MS	Liquid Chromatography-Mass Spectrometry
Lys	Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-Methyl benzhydryl amine
mg	milligram
MHz	Megahertz
min	minutes
Ni	Nickel
μL	Microliter
μM	Micromolar
mL	milliliter
mM	millimolar
mmol	millimoles
mp	melting point
MS	Mass spectrometry
MsCl	Mesyl Chloride
MW	Molecular weight
N	Normal
nm	Nanometer
NMR	Nuclear Magnetic Resonance

ONs	Oligonucleotides
p	Parallel
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
Pd	Palladium
PIDA	iodosobenzene diacetate (phenyliodonium diacetate)
ppm	Parts per million
PNA	Peptide Nucleic Acid
PS-oligo	Phosphorothioate-oligo
R	Rectus
R_f	Retention factor
RNA	Ribonucleic Acid
RP	Reverse Phase (-HPLC)
rt	Room temperature
RT	Retention time
S	Sinister
SAR	Structure Activity Relationship
SPPS	Solid Phase Peptide Synthesis
ss	Single strand/single stranded
T	Thymine
TEA/Et$_3$N	Triethylamine
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T_m	Melting temperature
UV-Vis	Ultraviolet-Visible

Abstract of Thesis

The thesis entitled "**γ-C-Substituted Multifunctional Peptide Nucleic Acids: Design, Synthesis and Bioevaluation**" is comprised of studies towards the design, synthesis, biophysical and biological evaluation of modified peptide nucleic acid (PNA) analogs. The modifications consist of incorporation of chiral, cationic (amine/guanidine) and charge-neutral azide functionality at γ-position in the PNA backbone. The PNA oligomers obtained by solid phase peptide synthesis were investigated for their binding to target DNA by various biophysical techniques. The work also describes the cell permeation properties of synthesized PNA oligomers.

The thesis is presented as four chapters:

Chapter 1: **Introduction to Peptide Nucleic Acids**

Chapter 2: **Design, Syntheses and Characterization of γ-C-Substituted Peptide Nucleic Acid Monomers and Their Oligomerization**

A: **Syntheses and Characterization of γ-(*S*-ethyleneamino), γ-(*S*-ethylene guanidino), γ-(*S*-azidomethylene) and γ-(*S*-azidobutylene) *aeg* PNA Monomers**

B: **Solid Phase Synthesis, Purification and Characterization of PNA Oligomers**

Chapter 3: **Biophysical Evaluation of Modified PNA Oligomers**

Chapter 4: **Cell Permeation Studies of PNA Oligomers**

Chapter 1: Introduction to peptide nucleic acids

This chapter gives an overview on the background literature for the undertaking research work emphasizing recent advancements in the field of peptide nucleic acids and their applications.

Oligonucleotides (ONs) capable of sequence-specific recognition of nucleic acids (DNA/RNA) are becoming more important for research, diagnosis and therapy. This is vital to make antisense or antigene based inhibition as practical approach to therapeutics. Sequence-specific binding of oligodeoxynucleotides (ODNs) to single

stranded RNA or duplex DNA through triple-helix formation provides a way to modulate gene expression. The specific inhibition is based on the Watson-Crick base-pairing between the heterocyclic bases of the antisense ONs and the target nucleic acid. In order to meet all the requirements of a successful medicinal agent, it is necessary for natural ONs to be chemically modified in a suitable manner. The chemical modifications of ONs have resulted in the synthesis and analysis of a large variety of oligonucleotide derivatives with modifications to the phosphate, the ribose or the nucleobase. This chapter presents a survey of the literature relevant to the area of nucleic acid therapeutics.

Peptide Nucleic Acids (PNA) (Figure **1**), first introduced by Nielsen et al, have emerged as a novel development in the field of oligonucleotide analogs. A brief introduction to the world of PNA is presented in this chapter followed by discussion on the recent trends in this area of research.
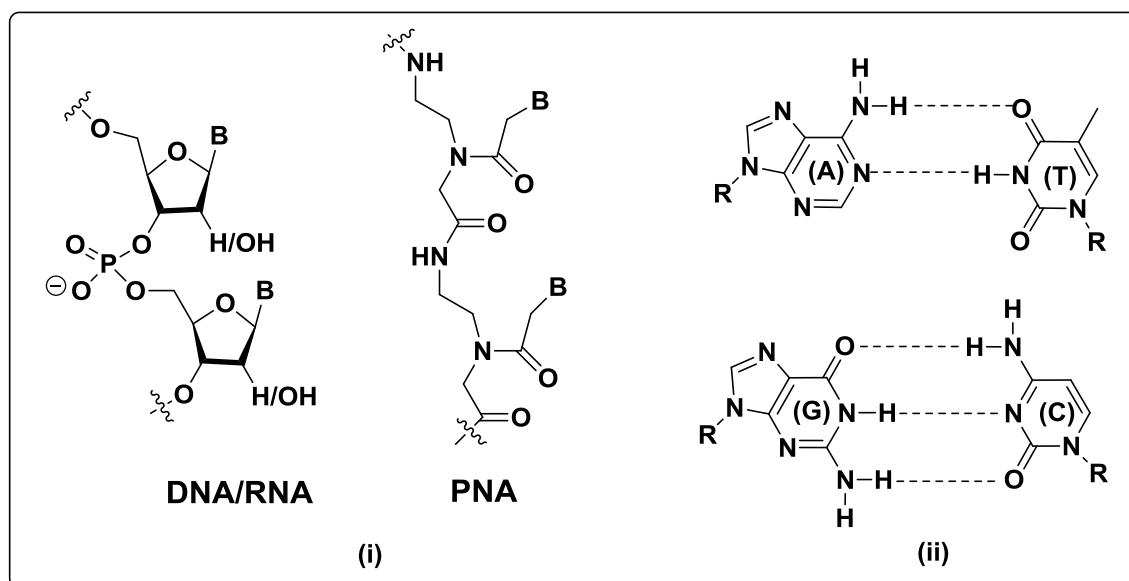

FIGURE 1 (i) Structure of DNA/RNA and PNA (ii) Watson-Crick hydrogen-bonding of nucleobases

The replacement of sugar-phosphate backbone of DNA by *N*-2-aminoethyl-glycyl backbone carrying nucleobases through methylene-carbonyl linker leads to PNA. The charge-neutral and achiral peptide nucleic acids were originally designed as oligonucleotide analogs that could be used for sequence-specific targeting of double-stranded DNA. PNAs bind with higher affinity to complementary DNA/RNA than their natural counterparts obeying Watson-Crick base pairing rule. PNAs and their analogs

are resistant to hydrolytic enzymes like proteases and nucleases. Due to these exceptional properties PNAs have major applications as a tool in molecular biology, as lead compounds for the development of gene targeted drugs via antigene/antisense technology, for diagnostics and biosensors. Peptide nucleic acids suffer from few drawbacks like poor water solubility, lack of efficient cell permeability and ambiguity in binding (parallel vs antiparallel) orientation. These limitations are being systematically addressed by various rationally designed PNA analogs that have improved PNA properties to different extents. The effect of the different structural modifications on biophysical and biochemical properties of PNA is overviewed to draw directions for the present work, aiming towards exploring new pathways of PNA-design and applications.

Chapter 2: Design, syntheses and characterization of γ-C-substituted peptide nucleic acid monomers and their oligomerization

The present work aims to address some of the limitations like poor water solubility and lack of efficient cell permeability with rationally designed γ-C-substituted PNA analogs. This chapter deals with the introduction of side chain carrying cationic (amine and guanidine) or neutral (azide) functional goups at γ-position of the PNA backbone. The presence of cationic side chains may enhance the cellular uptake as well as the PNA:DNA duplex stability. The incorporation of chiral, cationic D/L-Lysine unit (spacer is 4 x CH_2 group) in place of glycine has shown promising properties in terms of improved solubility and higher binding of single strand DNA (Figure **2 A**). It has been reported that cationic guanidinium groups grafted on PNAs (GPNA) enhance the cellular uptake of the substrates (Figure **2 B**). Recently it has also been shown that PNAs grafted with (α/γ, *R/S*)-aminomethylene pendants (spacer is 1 x CH_2 group) show regio and stereospecific effects on DNA binding and improve the cell permeation (Figure **2 C** & **D**).
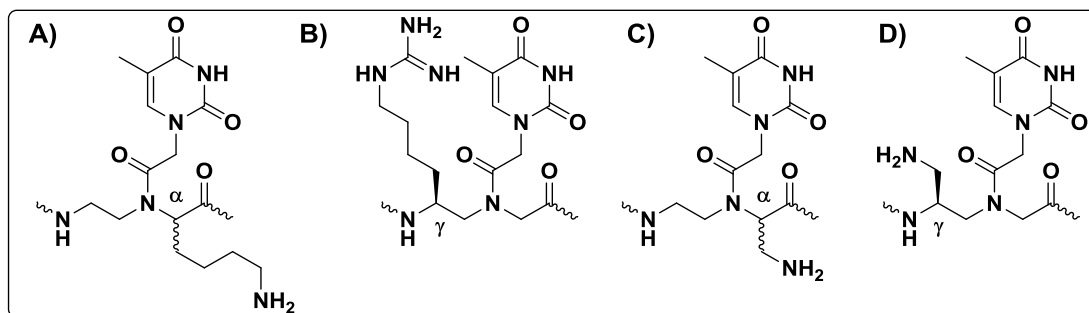
FIGURE 2 (A) D/L-Lysine substituted PNA analogs **(B)** GPNA **(C)** PNAs grafted with (α-*R/S*)-aminomethylene pendants **(D)** PNAs grafted with (γ-*S*)-aminomethylene pendants

In view of this, the present work demonstrates the incorporation of cationic amine or guanidine functionality through a shorter side chain (2 x CH$_2$) at γ-position in the form of γ-(*S*-ethyleneamino) aminoethylglycyl PNA [γ-(*S-eam*) *aeg* **PNA**] (**I**) and γ-(*S*-ethyleneguanidino) aminoethylglycyl PNA [γ-(*S-egd*) *aeg* **PNA**] (**II**) in order to investigate its potential as an effective conjugation point (Figure **3**). The presence of cationic functionality in the side chain of PNA backbone may improve the aqueous solubility and hence may aid in cellular penetration. The presence of cationic functional groups may also increase PNA–DNA duplex stability due to electrostatic interactions of cationic functionality with negatively charged DNA phosphate. Also the detailed studies of these modified PNAs would provide insight into the structure activity relationship (SAR) of PNA modifications at γ-position.
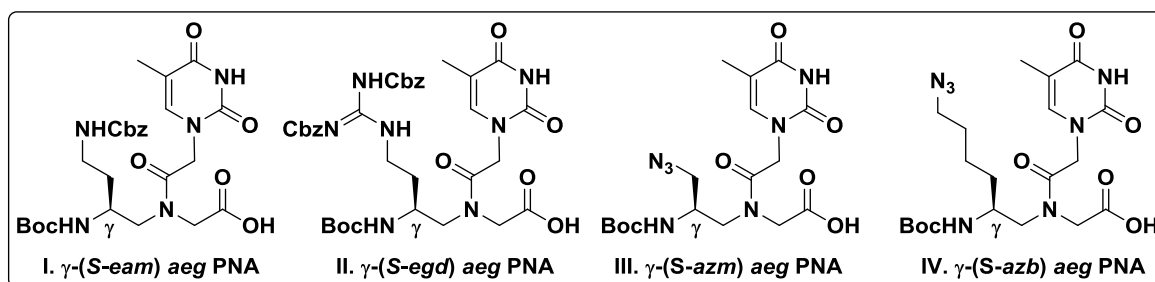

FIGURE 3 Target PNA monomers

The incorporation of azide in the backbone will give a chiral PNA with a neutral functional group which can be directly correlated with an unmodified *aeg* PNA that is also neutral. Also, the azide functional group can be utilized to attach fluorophore via click reaction with fluorescent alkyne derivative. In order to make fluorescently labeled PNAs, we have designed γ-C-substituted PNA analogs incorporating azide in the side chain at various distances (spacer 4 x CH$_2$ and 1 x CH$_2$) from the backbone in the form

of γ-(*S*-azidobutylene) aminoethylglycyl PNA [γ-(*S-azm*) *aeg* **PNA**] (**III**) and γ-(*S*-azidobutylene) aminoethylglycyl PNA [γ-(*S-azb*) *aeg* **PNA**] (**IV**) (Figure **3**).

Section 2A: Syntheses and characterization of γ-(*S-eam*), γ-(*S-egd*), γ-(*S-azm*) and (*S-azb*) *aeg* **PNA** monomers

In an attempt towards the synthesis of target **γ-(*S-eam*) *aeg* PNA** monomer, the commercially available L-Glutamine **1** was treated with di-tert-butyl dicarbonate [(Boc)$_2$O] to obtain Boc-protected L-Glutamine **2** and further converted to final monomer with ethyleneamino (*eam*) side chain at γ-position following the reaction sequence mentioned in the Scheme **1**.

SCHEME 1 Schematic representation of synthesis of γ-(*S-eam*) *aeg* PNA monomer
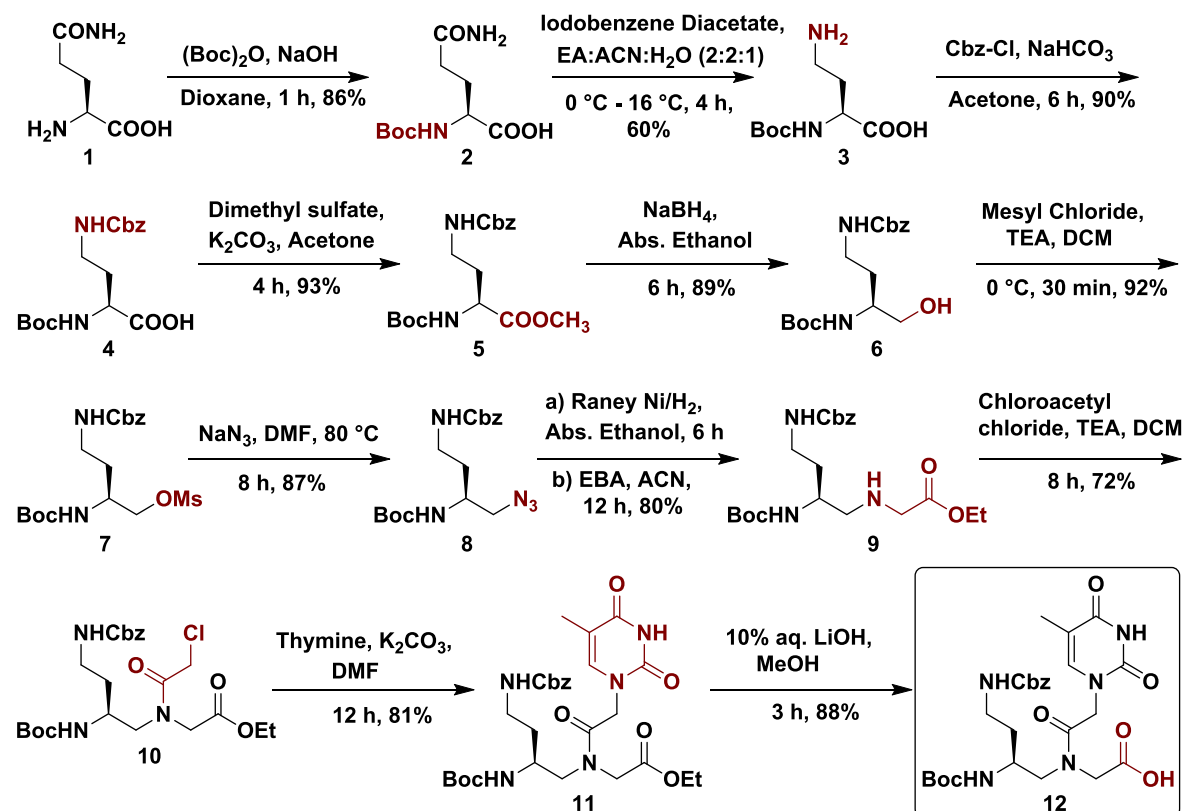
To synthesize **γ-(*S-egd*) *aeg* PNA** monomer, the Cbz-group from compound **11** was deprotected under hydrogenation conditions to give free amine which was guanidinylated using N,N'-bis-di-Cbz-S-Me-isothiourea to give compound **13** as *γ-(S-egd) aeg* PNA ester and further converted to final monomer with ethyleneguanidino (*egd*) side chain at γ-position (Scheme **2**).

**SCHEME 2 Schematic representation of synthesis of γ-(*S-egd*) *aeg* PNA monomer**
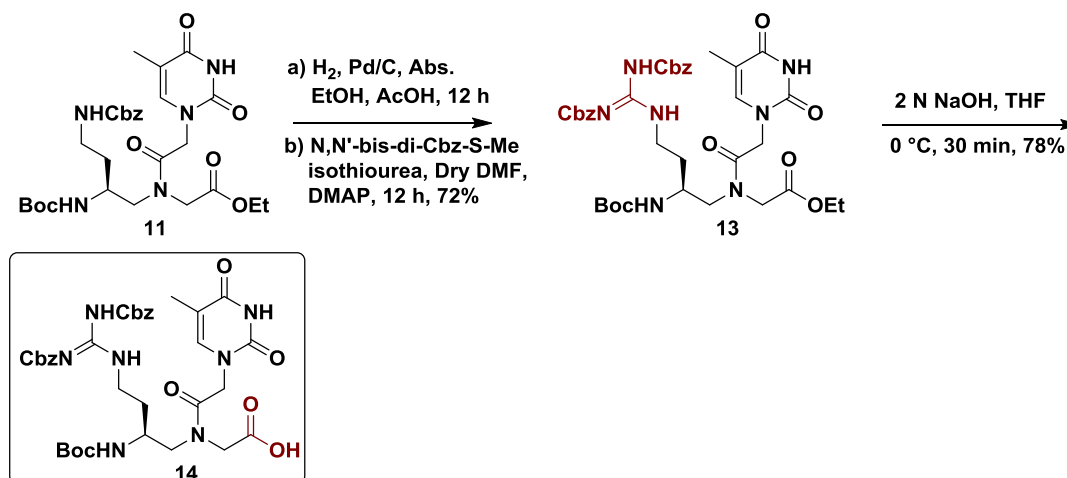
The synthesis of **γ-(*S-azm*) *aeg* PNA** (Scheme **3**) was started from commercially available L-Asparagine which was treated with di-tert-butyl dicarbonate, [(Boc)$_2$O] to obtain Boc-protected L-Asparagine **16**.

SCHEME 3 Schematic representation of synthesis of γ-(*S-azm*) *aeg* PNA monomer
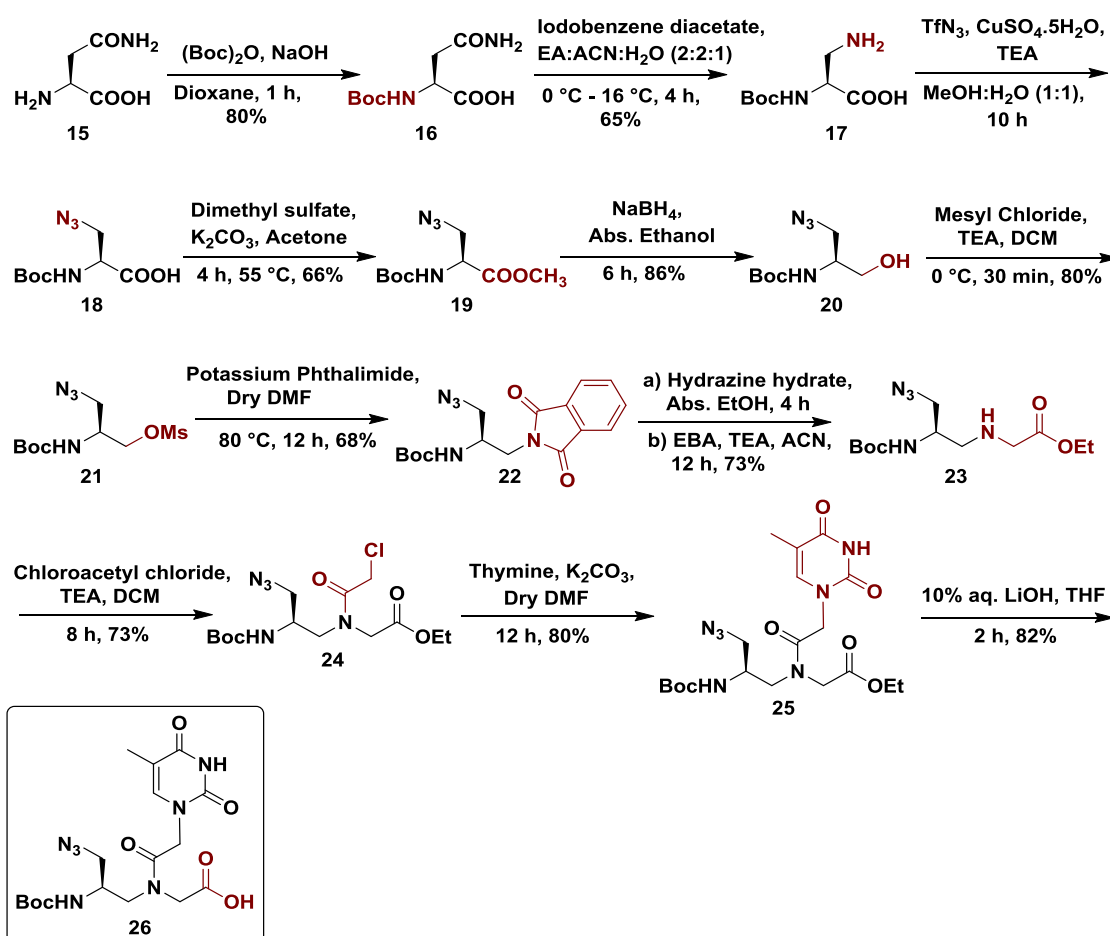
The side chain amide of compound **16** was converted to primary amine under Hoffman rearrangement conditions. This was followed by the conversion of amine to azide using triflyl azide to get azidomethylene (*azm*) side chain and further converted to final monomer following the sequence mentioned in scheme **3**.

To synthesize γ-(**S-azb**) *aeg* **PNA** monomer, side chain amine of L-Lysine was converted to azide using triflyl azide and further converted to final monomer following the sequence mentioned in the scheme **4**.

SCHEME 4 Schematic representation of synthesis of γ-(*S-azb*) *aeg* PNA monomer
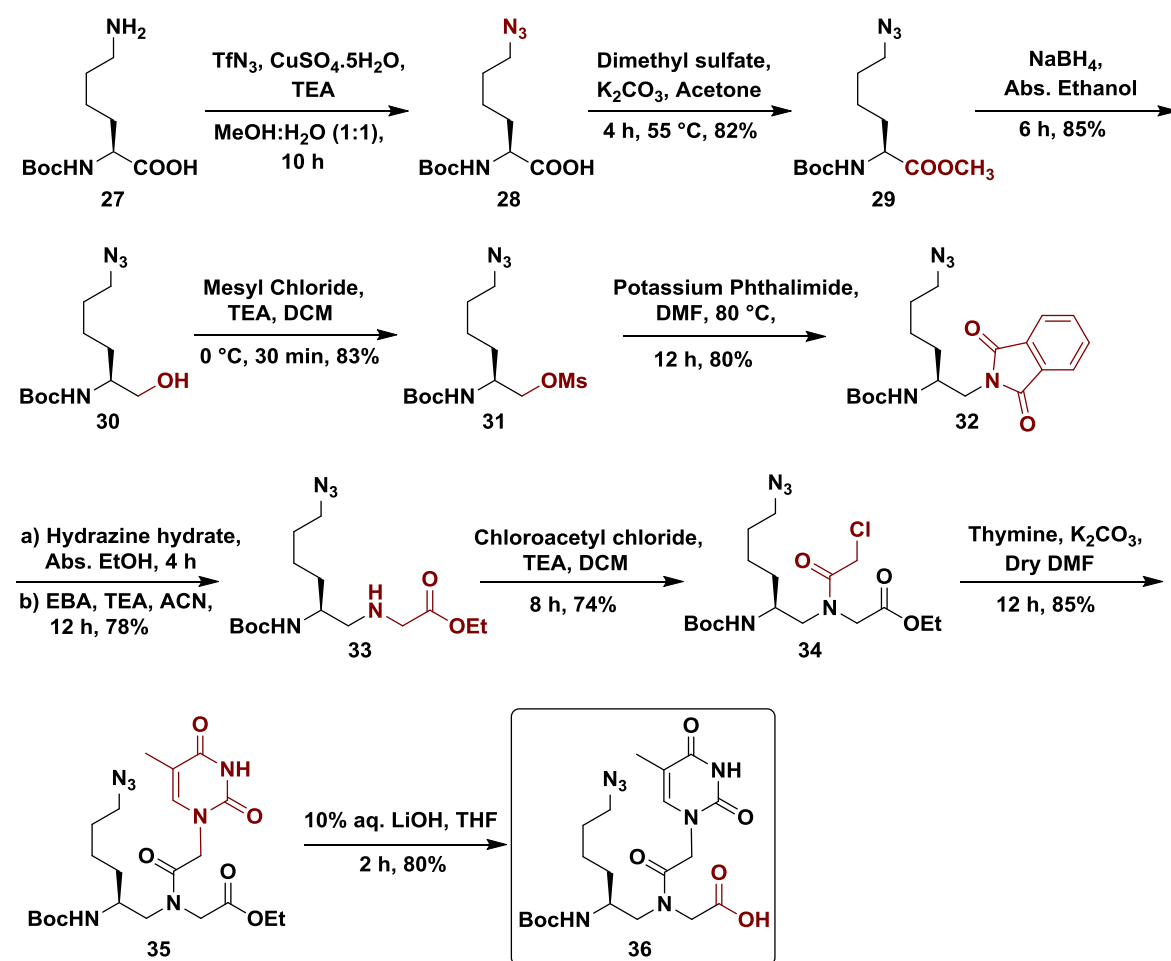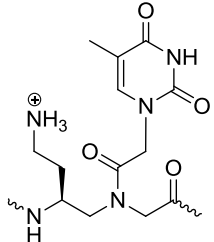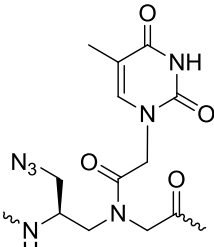

Section 2B: Solid phase synthesis, purification and characterization of PNA oligomers
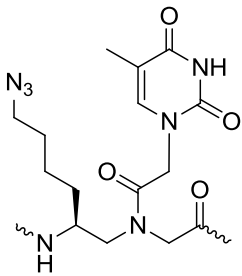
The PNA oligomers were synthesized by solid phase peptide synthesis protocol using *Boc* strategy. The modified PNA monomers were incorporated at desired

positions in the unmodified *aeg* PNA sequence using MBHA (4-methyl–Benzhydryl amine) resin as solid support. In order to study the potential of duplex formation and mismatch discrimination of the γ-C-substituted multifunctional PNA oligomers, the mixed purine pyrimidine sequences were synthesized.

The PNA sequences bearing modified (T) and unmodified (A/T/G/C) monomeric units were synthesized to improve the duplex stability and their cell permeation abilities (Table **1**). The unmodified *aeg* PNA sequence was also synthesized as control for comparisons.

TABLE 1 PNA oligomers with modified/unmodified monomers at various positions

Entry	Sequence Code	PNA sequences	Monomers used
1	*aeg* PNA **1**	H-T T A C C T C A G T-LysNH$_2$ 1 2 3 4 5 6 7 8 9 10	A/G/C/T = *aeg* PNA
2	*eam*-t$_2$ PNA **2**	H-T **t** A C C T C A G T-LysNH$_2$	
3	*eam*-t$_6$ PNA **3**	H-T T A C C **t** C A G T-LysNH$_2$	
4	*eam*-t$_{10}$ PNA **4**	H-T T A C C T C A G **t**-LysNH$_2$	
5	*eam*-t$_{2,6}$ PNA **5**	H-T **t** A C C **t** C A G T-LysNH$_2$	**t** = γ-(*S-eam*) *aeg*
6	*egd*-t$^γ_2$ PNA **6**	H-T **t^γ** A C C T C A G T-LysNH$_2$	
7	*egd*-t$^γ_6$ PNA **7**	H-T T A C C **t^γ** C A G T-LysNH$_2$	
8	*egd*-t$^γ_{10}$ PNA **8**	H-T T A C C T C A G **t^γ**-LysNH$_2$	
9	*egd*-t$^γ_{2,6}$ PNA **9**	H-T **t^γ** A C C **t^γ** C A G T-LysNH$_2$	
10	*egd*-t$^γ_{2,6,10}$ PNA **10**	H-T **t^γ** A C C **t^γ** C A G **t^γ**-LysNH$_2$	**t^γ** = γ-(*S-egd*) *aeg*
11	*azm*-T$^γ_2$ PNA **11**	H-T **T^γ** A C C T C A G T-LysNH$_2$	
12	*azm*-T$^γ_6$ PNA **12**	H-T T A C C **T^γ** C A G T-LysNH$_2$	
13	*azm*-T$^γ_{10}$ PNA **13**	H-T T A C C T C A G **T^γ**-LysNH$_2$	
14	*azm*-T$^γ_{2,6}$ PNA **14**	H-T **T^γ** A C C **T^γ** C A G T-LysNH$_2$	**T^γ** = γ-(*S-azm*) *aeg*

15	*azb*-T$_2$ PNA **15**	H-T **T** A C C T C A G T-LysNH$_2$
16	*azb*-T$_6$ PNA **16**	H-T T A C C **T** C A G T-LysNH$_2$
17	*azb*-T$_{10}$ PNA **17**	H-T T A C C T C A G **T**-LysNH$_2$
18	*azb*-T$_{2,6}$ PNA **18**	H-T **T** A C C **T** C A G T-LysNH$_2$

T = *γ-(S-azb) aeg*

eam = ethyleneamino, *egd* = ethyleneguanidino, *azm* = azidomethylene and *azb* = azidobutylene

The fluorescently labeled PNA oligomers were synthesized by on-resin click reaction between the polymer supported azido component (azide containing PNA oligomers) and the alkyne derivative of 5(6)-carboxyfluorescein (Figure **4**, Table **2**)
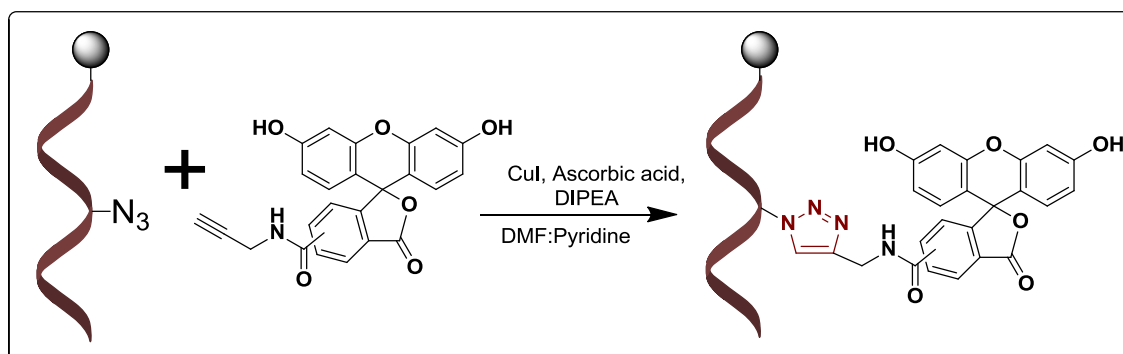

FIGURE 4 Synthesis of fluorescent PNA oligomers on solid support

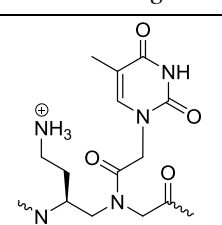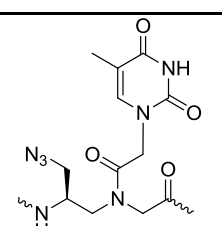**TABLE 2 Fluorescent PNA oligomers synthesized by click reaction**

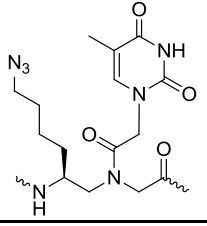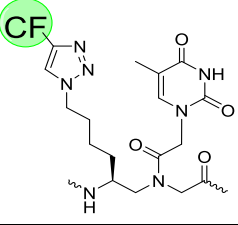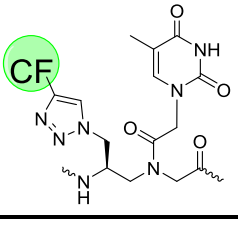| Entry	Sequence Code	Fluorescent PNA Sequences
19	*azb*-T$^f_6$ PNA **16**	H-T T A C C **T^f** C A G T-LysNH$_2$
20	*azb*-T$^f_{10}$ PNA **17**	H-T T A C C T A G **T^f**-LysNH$_2$
21	*azb*-T$^f_{2,6}$ PNA **18**	H-T **T^f** A C C **T^f** C A G T-LysNH$_2$
22	*azm*-t$^f_2$ PNA **11**	H-T **t^f** A C C T C A G T-LysNH$_2$
23	*azm*-t$^f_6$ PNA **12**	H-T T A C C **t^f** C A G T-LysNH$_2$
24	*azm*-t$^f_{10}$ PNA **13**	H-T T A C C T C A G **t^f**-LysNH$_2$
25	*azm*-t$^f_{2,6}$ PNA **14**	H-T **t^f** A C C **t^f** C A G T-LysNH$_2$

All the modified PNA oligomers were purified by RP-HPLC and characterized by MALDI-TOF mass spectroscopy. These purified PNA oligomers were used for biophysical studies discussed in the next chapter.

Chapter 3: Biophysical Evaluation of Modified PNA Oligomers

This chapter deals with the temperature dependent UV-spectroscopic studies to determine the thermal stability of all modified PNA oligomers with complementary **DNA 1**. The results of thermal stabilities for all PNA oligomers are tabulated in Table **3**. The specificity of these PNA oligomers was also investigated by challenging them to bind to **DNA 2** carrying single base mismatch in the middle of the sequence. The incorporation of cationic amine or guanidine in the *aeg* PNA system remarkably increased the PNA:DNA duplex stability. The guanidino and amino modified PNA stabilizes the duplex by ~ 5 °C and ~ 3 °C per modification respectively. Incorporation of neutral azide functionality in the *aeg* PNA backbone at γ-position gives better stabilizing effect when connected with a longer chain (4 x CH_2) to the backbone as compared to azide at shorter distance (1 x CH_2).

Table 3 UV-T_m and ΔT_m (°C) values of duplexes of modified PNA oligomers with complementary DNA 1

Sequence Code	PNA Sequence	DNA 1 5' ACTGAGGTAA 3' UV-T_m (ΔT_{m1})	Modification Incorporated
aeg PNA **1**	H-TTACCTCAGT-LysNH$_2$	43.4	A/T/G/C = *aeg* PNA
eam-t$_2$ PNA **2**	H-T**t**ACCTCAGT-LysNH$_2$	47.3 (+ 3.9)	
eam-t$_6$ PNA **3**	H-TTACC**t**AGT-LysNH$_2$	47.1 (+ 3.7)	
eam-t$_{10}$ PNA **4**	H-TTACCTCAG**t**-LysNH$_2$	49.2 (+ 5.8)	
eam-t$_{2,6}$ PNA **5**	H-T**t**ACC**t**CAGT-LysNH$_2$	49.7 (+ 6.3)	
egd-t$^{\gamma}_2$ PNA **6**	H-T**t$^{\gamma}$**ACCTCAGT-LysNH$_2$	45.9 (+ 2.5)	
egd-t$^{\gamma}_6$ PNA **7**	H-TTACC**t$^{\gamma}$**CAGT-LysNH$_2$	48.8 (+ 5.4)	
egd-t$^{\gamma}_{10}$ PNA **8**	H-TTACCTCAG**t$^{\gamma}$**-LysNH$_2$	52.0 (+ 8.6)	
egd-t$^{\gamma}_{2,6}$ PNA **9**	H-T**t$^{\gamma}$**ACC**t$^{\gamma}$**CAGT-LysNH$_2$	51.9 (+ 8.5)	
egd-t$^{\gamma}_{2,6,10}$ PNA **10**	H-T**t$^{\gamma}$**ACC**t$^{\gamma}$**CAG**t$^{\gamma}$**-LysNH$_2$	57.4 (+ 14.0)	
azm-T$^{\gamma}_2$ PNA **11**	H-T**T$^{\gamma}$**ACCTCAGT-LysNH$_2$	43.9 (+ 0.5)	
azm-T$^{\gamma}_6$ PNA **12**	H-TTACC**T$^{\gamma}$**CAGT-LysNH$_2$	49.7 (+ 6.3)	
azm-T$^{\gamma}_{10}$ PNA **13**	H-TTACCTCAG**T$^{\gamma}$**-LysNH$_2$	45.8 (+ 2.4)	
azm-T$^{\gamma}_{2,6}$ PNA **14**	H-T**T$^{\gamma}$**ACC**T$^{\gamma}$**CAGT-LysNH$_2$	43.6 (+ 0.2)	

azb-T$_2$ PNA **15**	H-T**T**ACCTCAGT-LysNH$_2$	44.9 (+ 1.5)	
azb-T$_6$ PNA **16**	H-TTACC**T**CAGT-LysNH$_2$	49.2 (+ 5.8)	
azb-T$_{10}$ PNA **17**	H-TTACCTCAG**T**-LysNH$_2$	50.9 (+ 7.5)	
azb-T$_{2,6}$ PNA **18**	H-T**T**ACC**T**CAGT-LysNH$_2$	50.6 (+ 7.2)	
azb-T$^f_6$ PNA **16**	H-TTACC**T^f**CAGT-LysNH$_2$	46.4 (+ 3.0)	
azb-T$^f_{10}$ PNA **17**	H-TTACCTAG**T^f**-LysNH$_2$	47.8 (+ 4.4)	
azb-T$^f_{2,6}$ PNA **18**	H-T**T^f**ACC**T^f**CAGT-LysNH$_2$	40.0 (- 3.4)	
azm-t$^f_2$ PNA **11**	H-T**t^f**ACCTCAGT-LysNH$_2$	43.8 (+ 0.4)	
azm-t$^f_6$ PNA **12**	H-TTACC**t^f**CAGT-LysNH$_2$	42.0 (- 1.4)	
azm-t$^f_{10}$ PNA **13**	H-TTACCTCAG**t^f**-LysNH$_2$	42.0 (- 1.4)	
azm-t$^f_{2,6}$ PNA **14**	H-T**t^f**ACC**t^f**CAGT-LysNH$_2$	40.0 (- 3.4)	

Δ **T_{m1}** = indicates the difference in T_m with control *aeg* PNA, nb = not binding, Δ T_m values are accurate to ± 0.5 °C

The effect of modifications incorporated at γ-position on the conformation of PNA:DNA duplexes was investigated by Circular Dichroism (CD) studies. The CD analysis showed that all these modifications incorporated γ-position does not alter the overall conformation of PNA:DNA duplex which gives a right handed helix like B-DNA form. These modified PNA oligomers were also tested to displace intercalated ethidium bromide (EtBr) from the duplex DNA. The comparative EtBr displacement studies showed that charge-neutral azido modified PNAs are better at displacing EtBr from duplex DNA than that of cationic amino/guanidino modified PNAs (Figure **5**).
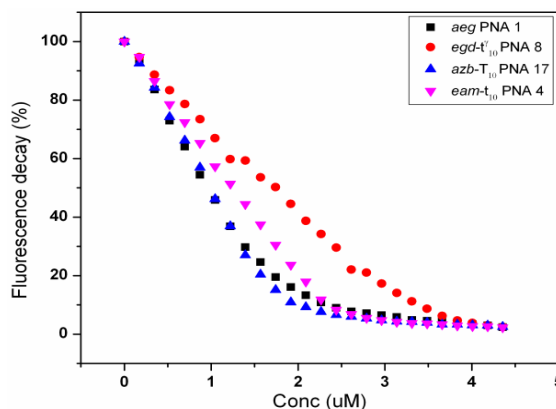

FIGURE 5 Comparative ethidium bromide displacement studies from *ds*DNA

Chapter 4: Cell permeation studies of PNA oligomers

The novel γ-C-substituted peptide nucleic acids have shown improved thermal stabilities with complementary DNA as compared to control *aeg* PNA. In order to gain insight on their cellular uptake properties, PNA oligomers were tagged with 5(6)-carboxyfluorescein at the N-terminus for visualization in the cells (Figure **6**, Table **4**). The cell permeation ability of these modified PNA oligomers was investigated by live cell imaging in NIH 3T3 and MCF-7 cell lines using confocal microscopy.
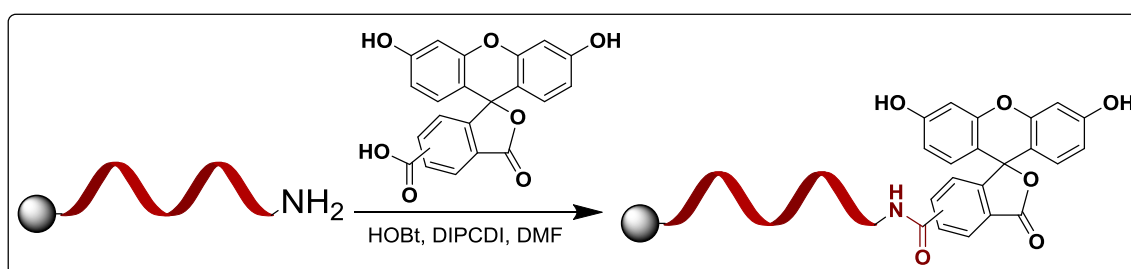
FIGURE 6 Synthesis of 5(6)-carboxyfluorescein tagged PNA oligomers

TABLE 4 5(6)-carboxyfluorescein tagged PNA oligomers

Entry	Sequence code	5(6)-CF tagged PNA oligomers
1	*aeg* PNA **1-CF**	**CF**-T T A C C T C A G T-LysNH$_2$
2	*eam*-t$_2$ PNA **2-CF**	**CF**-T **t** A C C TC A G T-LysNH$_2$
3	*eam*-t$_{2,6}$ PNA **5-CF**	**CF**-T **t** A C C **t** C A G T-LysNH$_2$
4	*egd*-t$^\gamma_2$ PNA **6-CF**	**CF**-T **t$^\gamma$** A C C T C A G T-LysNH$_2$
5	*egd*-t$^\gamma_{2,6}$ PNA **9-CF**	**CF**-T **t$^\gamma$** A C C **t$^\gamma$** C A G T-LysNH$_2$
6	*egd*-t$^\gamma_{2,6,10}$ PNA **10-CF**	**CF**-T **t$^\gamma$** A C C **t$^\gamma$** C A G **t$^\gamma$**-LysNH$_2$
7	*azb*-T$_{10}$ PNA **17-CF**	**CF**-T T A C C T C A G **T**-LysNH$_2$
8	*azm*-T$^\gamma_{10}$ PNA **13-CF**	**CF**-T T A C C T C A G **T$^\gamma$**-LysNH$_2$

aeg = aminoethylglycine, *eam* = ethyleneamino, *egd* = ethyleneguanidino, *azb* = azidobutylene, *azm* = azidomethylene and **CF** = 5(6)-carboxyfluorescein

The qualitative data obtained from confocal microscopy showed that modified PNA as well as unmodified *aeg* PNA can penetrate into both cell lines used for the study and the PNA oligomers were observed to be localized in the vicinity of nucleus in the cytoplasm. The representative confocal microscopy image for cell permeation of three guanidino modified PNA in NIH 3T3 cells is shown in Figure **7**. The

quantification of differential cellular uptake was carried out using Fluorescence Activated Cell Sorter (FACS) analysis. FACS gives the quantitative recording of fluorescent signals from each individual cell as well as it provides important information about the positive and negative cells in terms of their fluorescent signal. Thus, FACS data would provide quantitative estimation of cell permeability of all PNA oligomers which can be differentiated from their percent positive cells and mean fluorescence intensity. FACS studies showed that cationic functional groups (amine/guanidine) incorporated in the side chain at γ-position remarkably increase the efficiency of cell penetration in both cell lines used for the study.
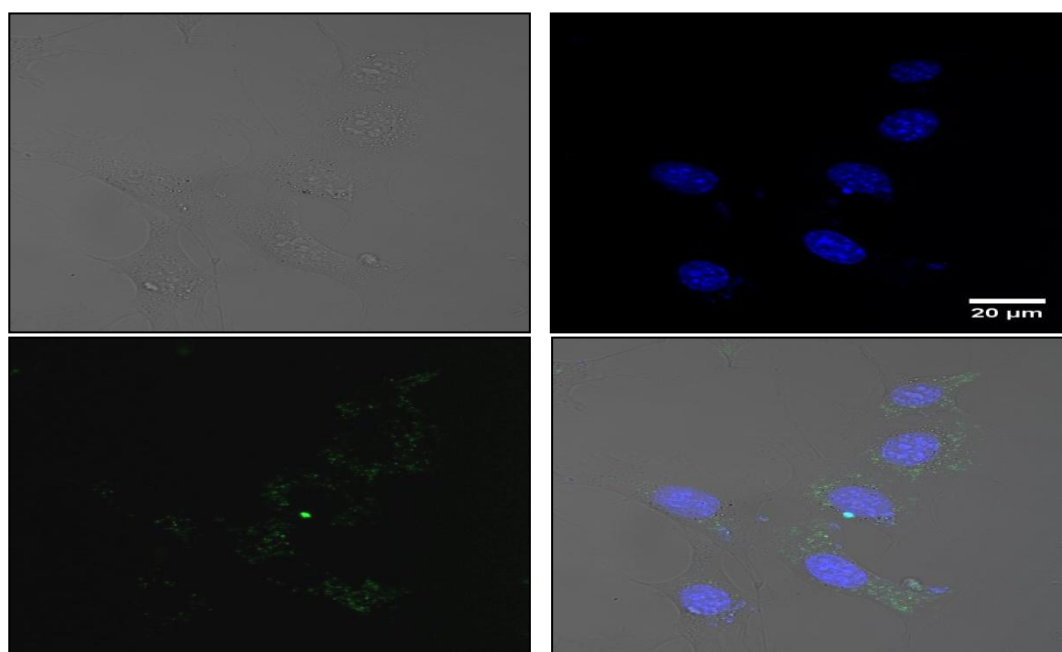
FIGURE 7 Confocal microscopy images for 3-guanidino modified *egd*-t$^{\gamma}_{2,6,10}$ PNA 10-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

The mean fluorescence intensities were found to be minimally changed for control *aeg* PNA and modified PNA oligomers in both cell lines. The FACS analysis data for NIH 3T3 cells is shown in Figure **8**. Whereas, % positive cells (which gives the number of cells that have taken up the PNA inside) for modified PNA oligomers was remarkably increased than that of unmodified *aeg* PNA (Figure **9**). The guanidino modified PNA was found to be better among all modifications to increase the efficiency of cell penetration. The charge-neutral azido modified PNA oligomers were equally efficient in cell penetration like control *aeg* PNA.
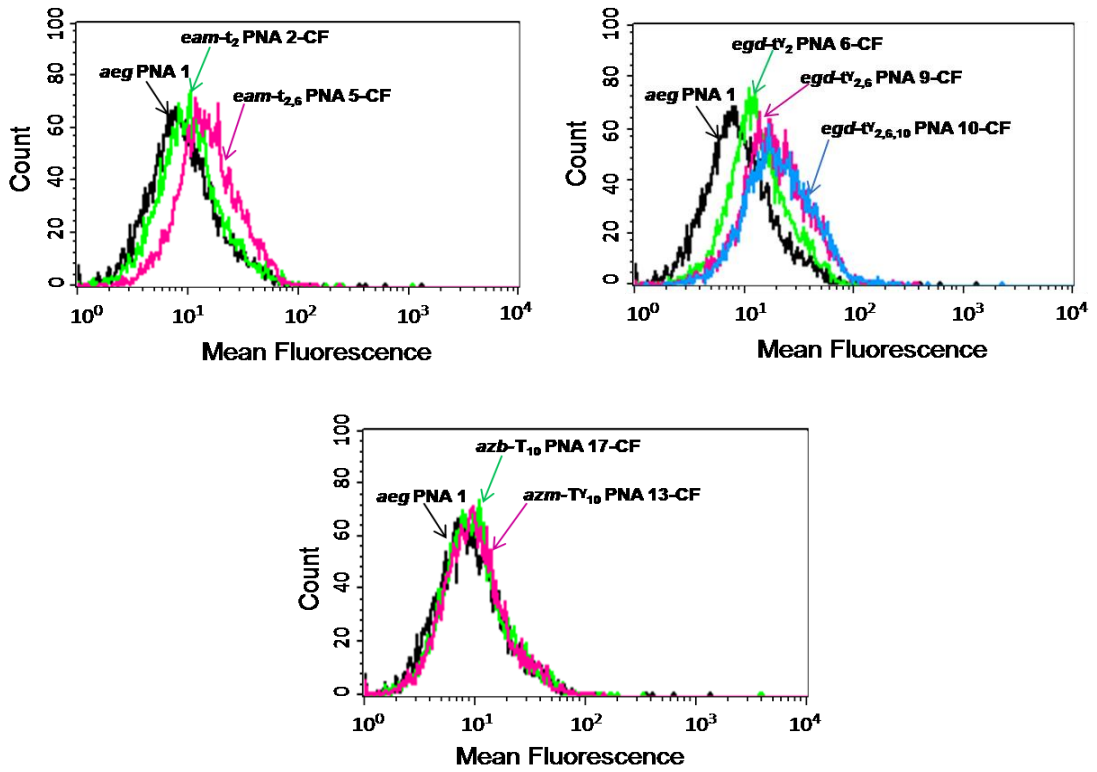
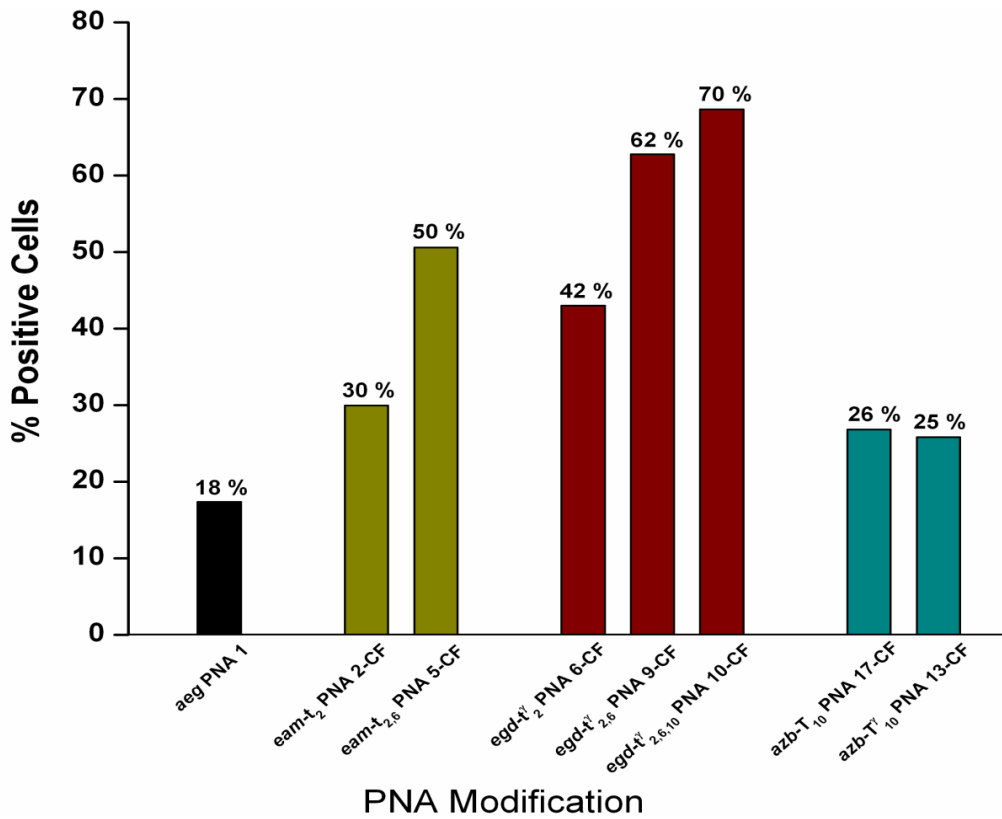FIGURE 8 Quantification of PNA cell penetration in NIH 3T3 cells

FIGURE 9 FACS analysis showing the % positive cells for quantification of cell permeation in NIH 3T3 cells

Summary of thesis

- ➤ Rationally designed PNA monomers by incorporating amino, guanidino and azido functionalities at γ-position have been synthesized using easily available natural amino acid precursors.

- ➤ Modified PNA monomers were incorporated into *aeg* PNA sequence at desired positions by solid phase peptide synthesis protocol using Boc strategy. The synthesized PNA oligomers were purified by RP-HPLC and characterized by MALDI-TOF spectrometry.

- ➤ Azido modified PNA oligomers were utilized to synthesize fluorescently labeled PNAs by on-resin click reaction between the solid supported azide and the alkyne derivative of 5(6)-carboxyfluorescein.

- ➤ The thermal stability of chiral, cationic (amine/guanidine) and charge-neutral azide modified PNA oligomers with complementary DNA was investigated by temperature dependent UV-visible spectroscopy. The specificity of these PNA oligomers was also investigated by challenging them to bind to DNA carrying single base mismatch in the middle of the sequence.

- ➤ The effect of substitution at γ-position in the PNA backbone on the conformation of PNA:DNA duplexes was studied by CD spectroscopy.

- ➤ The differential binding ability of various modified PNA oligomers towards the complementary DNA was examined by the displacement of ethidium bromide intercalated into the duplex DNA.

- ➤ The cell permeation ability of 5(6)-carboxyfluorescein tagged PNA oligomers was investigated by live cell imaging using confocal microscopy.

- ➤ The quantification of cell penetration of modified PNA oligomers was carried out by fluorescence activated cell sorter (FACS) analysis.

Chapter 1

Introduction to Peptide Nucleic Acids

A brief introduction to nucleic acids especially peptide nucleic acids followed by the recent literature trends are presented in this chapter. Important PNA properties and its structural features are discussed to understand the various functions of PNA. The different structural modifications to overcome the drawbacks of PNA has been overviewed to draw directions for the present work, targeted towards exploring new pathways of PNA design and applications.

1.1 Introduction to nucleic acids

The nucleic acids are biological macromolecules essential for all known forms of life. Together with proteins, 2'-deoxyribonucleic acids (DNA) and ribonucleic acids (RNA) are the most important biological macromolecules of all living things. Their important functions are encoding, transmitting and expressing the genetic information within a biological system which is explained by the central dogma of molecular biology (Figure **1.1**).
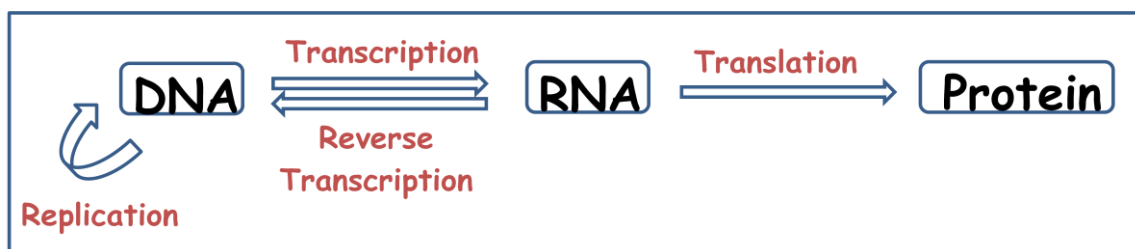
FIGURE 1.1 Central dogma of molecular biology displaying the flow of genetic information

The nucleic acids are constructed from the repeating units of nucleotides which are made up of a nitrogenous base (purines or pyrimidines), a pentose sugar and a phosphate group. The substructure nucleoside is a combination of a nitrogenous base and pentose sugar.[1] Adenine (A) and guanine (G) are the two purine nucleobases whereas thymine (T), cytosine (C) and uracil (U) are the pyrimidine nucleobases. The nucleic acid structures differ in the sugar part of nucleotides where DNA contains 2'-deoxyribose while RNA contains ribose sugar. The two nucleic acids are differentiated on the basis of nucleobase composition. The heterocyclic bases adenine, guanine and cytosine are common in both DNA and RNA while thymine is present in DNA and uracil occurs in RNA. In 1953, Watson and Crick proposed that the molecular architecture of DNA consists of two helical chains each coiled round the same axis with a right handed twist.[2] The phosphodiester group is joined to the *β*-D-deoxyribofuranose residues through 3'-5' linkages forming a helical chain where they are pointed to outside of the helix (Figure **1.2**). The two DNA strands are held together by specific hydrogen bonds between Watson-Crick base pairs (A:T & G:C) to form antiparallel double helical structure.
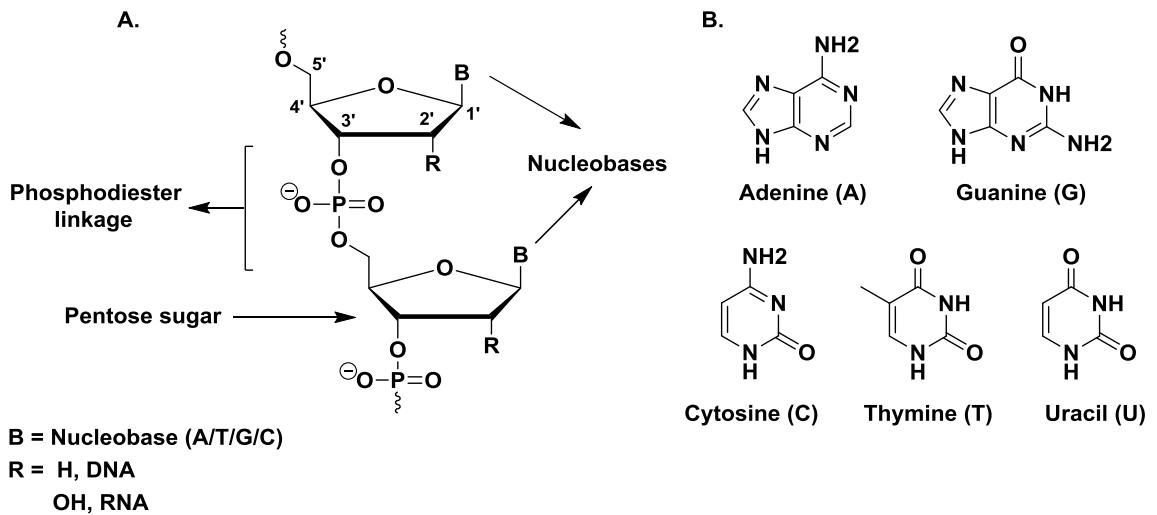
FIGURE 1.2 Chemical structures of A) DNA and RNA B) Nucleobases

1.2 Hydrogen bonding

Self recognition of nucleic acids through hydrogen bonding is one of the important principles of molecular recognition in complex systems. The hydrogen bonds are formed between the major *amino-keto* tautomers of the bases to establish the fidelity of DNA transcription and translation. The N-H groups of the bases are potent hydrogen bond donors while the sp^2 hybridized electron pairs on the oxygens of the carbonyl (C=O) groups and on the ring nitrogens are very good hydrogen bond acceptors. Watson-Crick hydrogen bonding consists of two hydrogen bonds in A:T base pair and three hydrogen bonds in G:C base pair (Figure **1.3**).
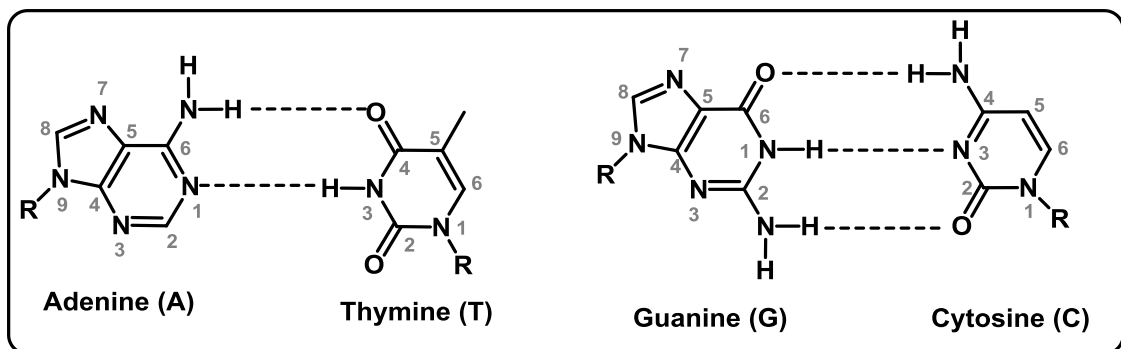
FIGURE 1.3 Watson-Crick hydrogen bonding for A:T and G:C base pairs

Other important hydrogen bonding pairs are the Hoogsteen[3] (HG) and Wobble[4] base pairs. Hoogsteen base pairing is not isomorphous with Watson-Crick base pairing

because they have an 80° angle between the glycosidic bonds and 8.6 Å separation of the anomeric carbons (Figure **1.4 a, b**). Hoogsteen hydrogen bonding has importance in triple helix formation and in protein-DNA complexes. Wobble base pairing is non-Watson-Crick base pairing between two nucleotides in RNA molecules. Wobble base pairs are fundamental in RNA secondary structure and are important for the proper translation of genetic code (Figure **1.4 c, d**).
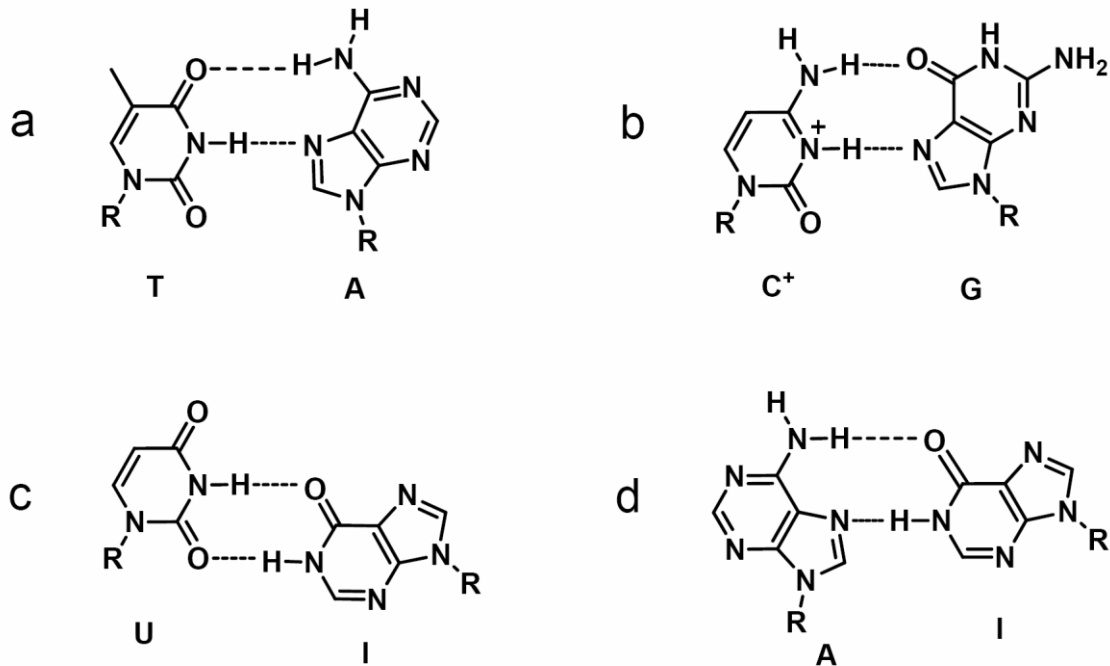

FIGURE 1.4 (a & b) Hoogsteen base pairing (c & d) Wobble base pairing

1.3 Secondary structures of nucleic acids

DNA exists in various possible conformations like A-DNA, B-DNA and Z-DNA, although, the most common is the B-DNA which is a right-handed double helix.[5] It has a wide and deep major-groove with a narrow and shallow minor-groove wherein the bases lie perpendicular to the helical axis. A-DNA also forms a right handed helix where major groove is deep and narrow while the minor groove is broad and shallow. In both A and B forms of DNA, the Watson-Crick base pairing is maintained along with *anti*-glycosidic conformation. Z-DNA is a left-handed double helix and mostly favoured in alternating G-C sequences. The left handed helix for Z-DNA is a result of a switch in the glycosidic bond with a *syn* conformation (Figure **1.5**). The sugar

conformation is however different in both forms, with the B-DNA form showing *C2'-endo* puckering and the A-DNA form exhibiting *C3'-endo* sugar-pucker (Figure **1.6**).
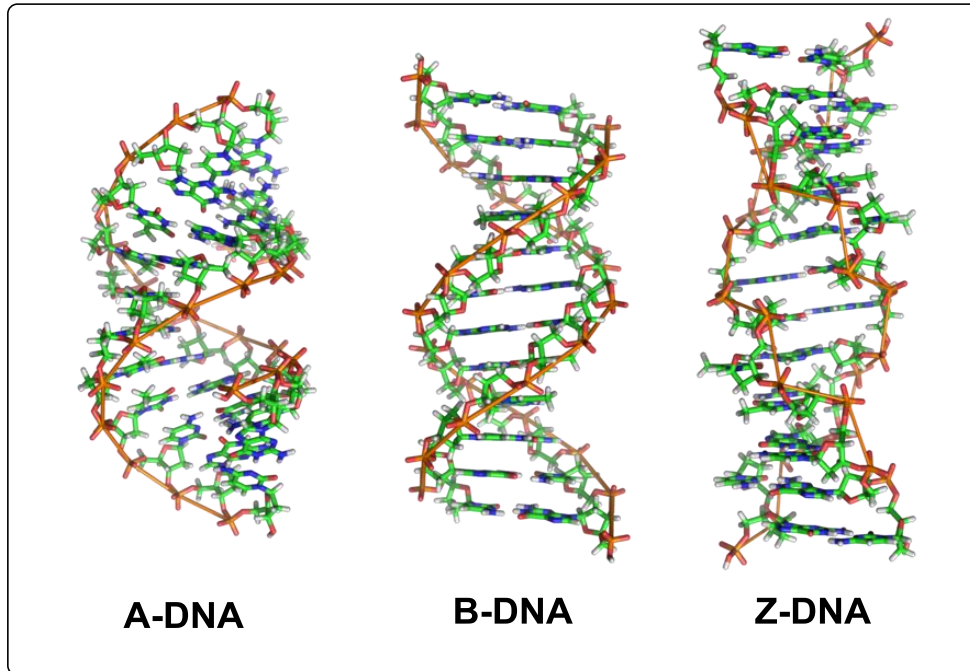

A-DNA   **B-DNA**   **Z-DNA**

FIGURE 1.5 Molecular models of A, B and Z-DNA forms of DNA[5]

RNA has a greater structural versatility than DNA in the variety of its species, in its diversity of conformations and in its chemical reactivity. The presence of the 2'-hydroxyl group in RNA hinders the formation of a B-form helix and it acquires the A-type helix showing *C3'-endo* sugar puckering. More commonly, RNA is single stranded, can form complex and unusual shapes such as stem and bubble structures, which occur due to the intrachain base pairing. An example is *t*-RNA, the key molecule involved in the translation of genetic information to proteins.
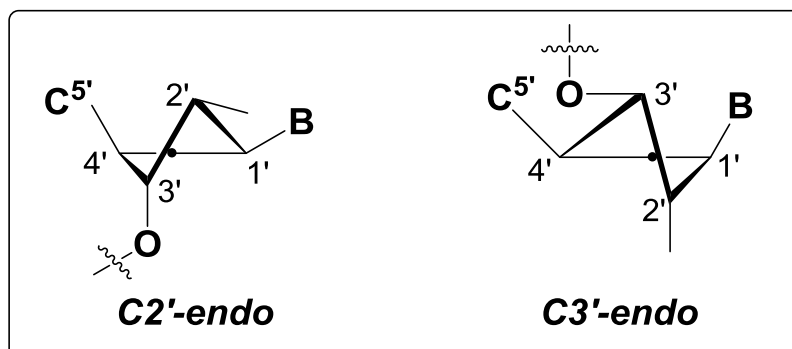
C2'-endo   *C3'-endo*

FIGURE 1.6 Structures of *C2'-endo* and *C3'-endo* sugar puckering

1.4 Applications of nucleic acids

Sequence specific binding of oligonucleotides to single stranded RNA or duplex DNA through triple helix formation provides a way to modulate the gene expression. The central feature of biology is the recognition of DNA and RNA sequences by complementary oligonucleotides and it is important for hybridization based biological applications. The study of such complementary recognition is possible with widely used experimental techniques and diagnostic protocols. The conjugation or attachment of reactive functional groups to oligonucleotides has been investigated to cleave nucleic acids in a sequence specific way which is described as artificial nucleases.[6] The combination of specificity and simplicity makes oligonucleotides highly attractive as diagnostic and research tools.

Oligonucleotide analogs as therapeutic agents

The classical approach of designing drug molecules against protein target requires a deep understanding of 3D structure of the protein, binding site and the binding forces. Since very little is known about the process of protein folding and the lack of generality, this approach of drug discovery has various drawbacks. On the other hand, the nucleotide sequence in RNA and DNA is universal and the understanding of their structures is much better to design drugs for nucleic acid targets. In principle, one can design drugs that are like nucleic acids, are repetitive in its primary structure and bind sequence specifically to the nucleic acid drug targets. In order for the sequence specific recognition to happen, the drug should contain nucleobases which are fundamental units of nucleic acid recognition. In short, a small piece of oligonucleotide containing 15-25 nucleotides can act as a drug which would bind to target nucleic acid and stop further processing of protein production. Two innovative strategies are being tested for inhibiting the production of disease related proteins using such sequence specific DNA fragments as gene expression inhibitors.

The first strategy known as *antisense strategy*[7] aims to selectively impede translation by inhibiting the protein synthesis. Antisense oligonucleotides recognize a complementary sequence on target mRNA through Watson-Crick base pairing and form a duplex which either leads to cleavage of the RNA by RNase H or is not recognized for further processing by protein synthesis machinery. Thus, it will retard

the expression of corresponding disease related proteins (Figure **1.7 b**). In the second strategy which is known as *antigene strategy*, the production of unwanted proteins can be stopped by selectively inhibiting the transcription process. In this strategy, oligonucleotides target the major groove of DNA where it winds around the double helical DNA to form triplex involving Hoogsteen hydrogen bonds[8] (Figure **1.7 a**).
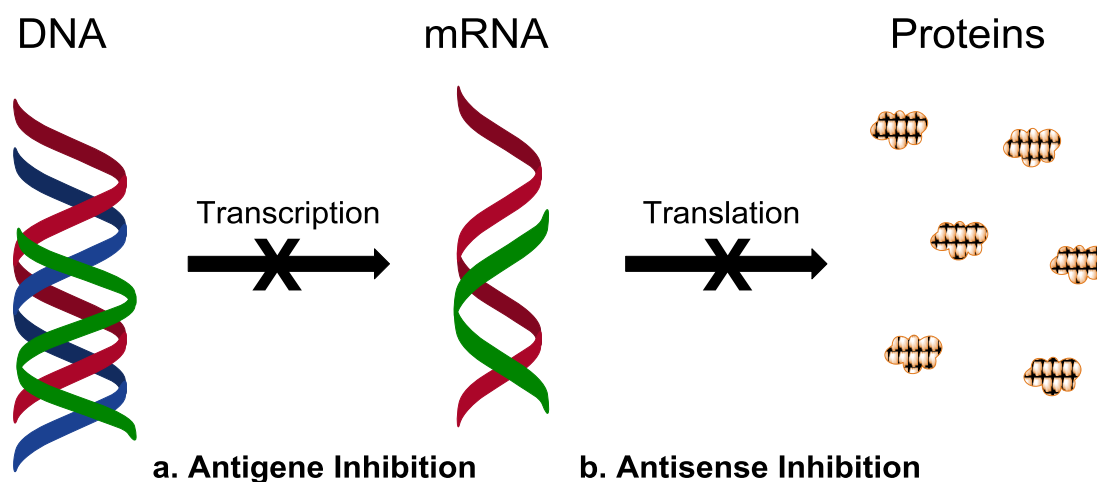

DNA   **mRNA**   **Proteins**

Transcription   Translation

a. Antigene Inhibition   **b. Antisense Inhibition**

FIGURE 1.7 Principle of antisense and antigene oligonucleotides

Gene silencing by RNA interference

The impact in the field of antisense research got triggered after the discovery of RNA interference (RNAi). This naturally occurring phenomenon as a potent sequence-specific mechanism for post-transcriptional gene silencing was first described for the nematode worm *Caenorhabditis elegans*.[8c] RNA interference is initiated by long double-stranded RNA molecules, which are processed into 21-23 nucleotides long RNAs by the Dicer enzyme (Figure **1.8**). This RNase III protein is thought to act as a dimer that cleaves both strands of *ds*RNAs and leaves two-nucleotide, 3´ overhanging ends. These small interfering RNAs (siRNAs) are then incorporated into the RNA-induced silencing complex (RISC), a protein RNA complex, and guide a nuclease, which degrades the target RNA.

1.5 Antisense oligonucleotides

The principle of antisense drugs is based on the direct utility of sequence information and sequence-specific recognition of nucleic acids and hence such drugs
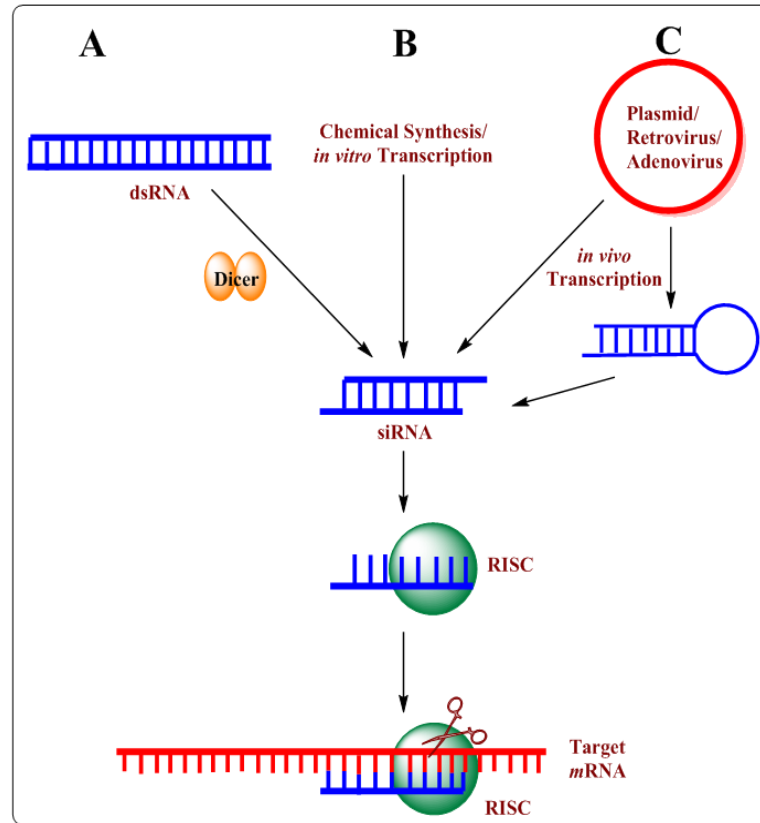
Figure 1.8 Gene silencing by RNA interference

have the potential for treatment of several viral and bacterial infections, cancerous outgrowths and inflammatory or genetic disorders. Various cellular processes can be inhibited depending on the site at which the antisense oligonucleotide hybridizes to the target nucleic acid. Zamecnik and Stephenson discovered that viral replication in cell culture was inhibited by oligonucleotides (ONs) and these ONs have the potential to act as antisense agents for therapeutics.[9] There are several pre-requisites for oligonucleotides to serve the purpose of inhibiting translation and thereby act as antisense therapeutic agents.[10] These include:

i)   Ease of synthesis in bulk quantities

ii)   *In-vivo* stability to cellular degrading enzymes

iii)   Ability to penetrate the cell membrane

iv)   Retention by the target cells

v)   Ability to interact with their cellular targets (DNA/RNA)

vi)   No non-specific interaction with other macromolecules

Although the requirements for specificity and binding affinity are satisfactorily met by the unmodified oligonucleotides, their susceptibility to hydrolytic enzymes (nucleases) and poor cell permeability are the major obstacles for their use as nucleic acid therapeutic agents. In order to meet all the above mentioned requirements, it is necessary for ONs to be chemically modified in a suitable manner.

1.5.1 Chemical modifications of DNA

To improve the rate, affinity and specificity of oligonucleotide recognition, while enhancing the cell membrane permeability and resistance to nuclease degradation, several chemical modifications of DNA have been attempted[11] (Figure **1.9**) in literature.
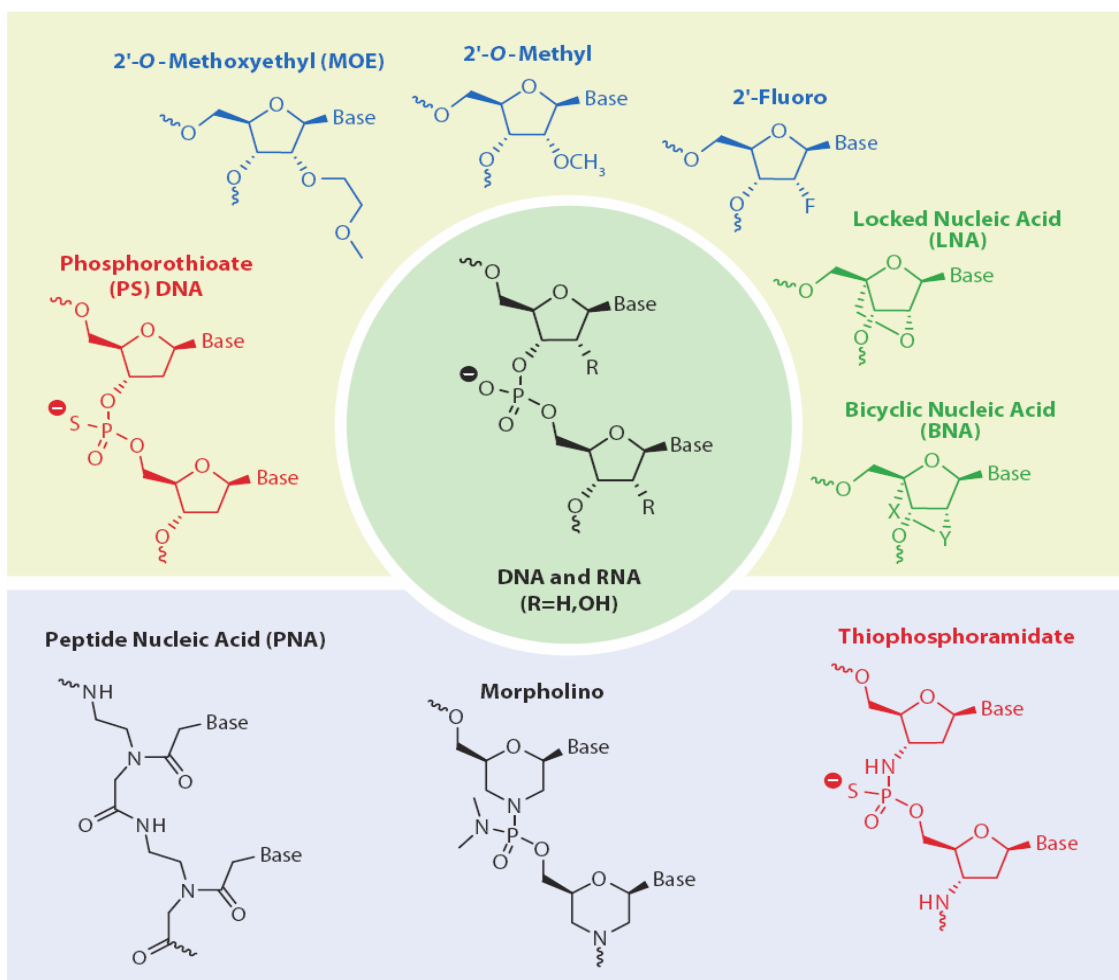
FIGURE 1.9 Examples of chemical modifications of oligonucleotides[11]

In general, three types of chemical modifications of DNA can be distinguished as

    i)   Modified phosphate linkages

ii)   Modified sugars and

iii)   Altered sugar-phosphate backbone

1.5.1a Alternative phosphate containing linkages

The modifications of phosphate moiety resulted in the development of phosphorothioates[12] **1**, methyl phosphonates[13] **2**, phosphoroamidates[14] **3**, phosphotriesters[15] **4** and boranophosphonates[16] **5** (Figure **1.10**).

Phosphorothioate (PS)-containing oligonucleotides differ from natural nucleic acids in that one of the nonbridging phosphate oxygen atoms is replaced with a sulfur atom. Phosphorothioate oligonucleotides are one of the earliest and most widely used backbone modifications for antisense drugs. The substitution of sulfur for oxygen in PS-oligos greatly increases stability to nucleolytic degradation. PS-oligos are able to efficiently elicit the RNase H cleavage of target mRNA, which is critical in the mechanism of action of antisense oligonucleotides.[17] Their binding to plasma proteins protects them from rapid renal excretion and is responsible for increased serum half-life. Moreover, Vitravene (Fomiversen) is one of the two FDA approved drugs, which is based on PS-oligos. Most of the antisense drugs which are under various stages of clinical trials incorporate PS-modifications. The binding of PS-oligos to certain proteins, particularly those which interact with polyanions such as hairpin-binding proteins, prove to be their major drawback.[18,19,20]
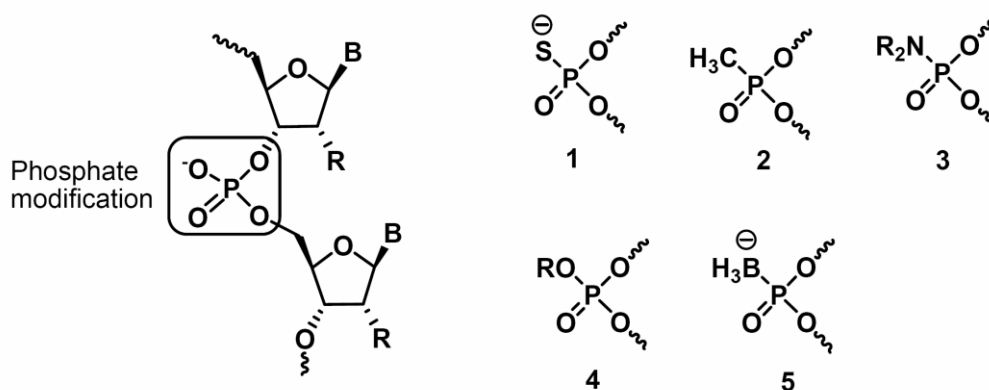
FIGURE 1.10 Structure of alternative phosphate linkages

Other backbone modifications (methylphosphonates, phosphoroamidates etc.) of oligonucleotides have been less successful at improving the oligonucleotide properties.

In methylphosphonates, one of the nonbridging oxygen atoms is replaced with a methyl group and is neutral in charge.[21] Although it provides high nuclease resistance, it does not induce RNase H activity. Additionally, increasing the number of methylphosphonate units in an oligomer leads to loss of affinity towards the target mRNA and to poor water solubility.

1.5.1b Sugar modifications

This class of ONs is represented by nucleotides with alkyl substitutions at the 2'-position of ribose sugar (Figure **1.11**). Organization of the sugar into RNA-like *C3'-endo* pucker increases the binding affinity of these ONs towards the complementary RNA.[22] Furthermore, the 2'-substituent in an oligonucleotide increases the steric bulk near the 3'-phosphate and makes it difficult for degrading enzymes to cleave the phosphodiester bond. The increase in binding affinity of 2'-modified ONs is energetically driven by the electronegative substituent at the 2'-position. Among all 2'-modified ONs, 2'-fluoro modification (Figure **1.11 a**) imparts the highest binding affinity towards the target RNA. 2'-O-methyl (Figure **1.11 b**) group enhances the binding affinity to a lesser extent than do the 2'-fluoro modification but impart a substantial degree of nuclease resistance to the corresponding oligonucleotide. 2'-O-methoxyethyl[23] (MOE) (Figure **1.11 c**) is currently the most advanced 2'-modification of the 2'-modified series and many antisense drugs having this modification have entered clinical trials. Unfortunately, the lack of RNase H activity restricts the use of 2'-modified oligonucleotides for antisense purpose. To overcome this drawback, the gapmer strategy has been used where regions of 2'-modified residues flank a central DNA region of the oligonucleotide.[21]
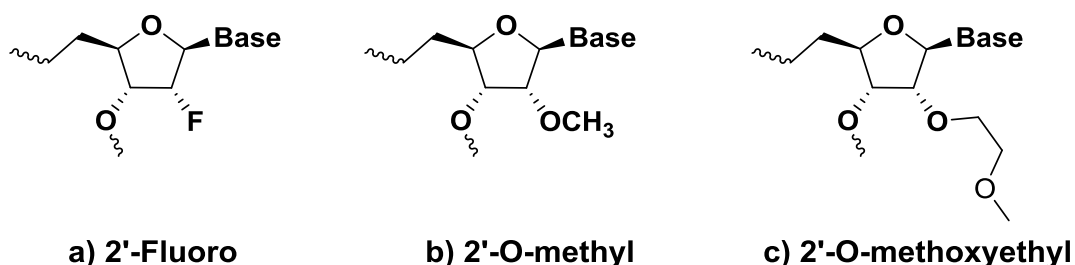

a) 2'-Fluoro   b) 2'-O-methyl   c) 2'-O-methoxyethyl

FIGURE 1.11 2'-modified oligonucleotides

1.5.1c Sugar-phosphate backbone modifications

In addition to phosphodiester and sugar modifications, replacement of the sugar phosphate backbone with isosteric structures has been devised. The concept of conformational restriction has been used widely to enhance binding affinity and biostability. Some of the DNA and RNA analogs developed with modified sugar-phosphate backbones are described below.

Locked nucleic acid (LNA)

These oligonucleotides developed by Jesper Wengel *et al*. in 1998 are one of the most promising class of chemically modified ONs. These analogs are bicyclic systems that contain a methylene bridge that connects the 2'-O- of the ribose with the 4'-C, locking the ribose in *C3'-endo* conformation (Figure **1.12**).[24] Introduction of LNA into a DNA sequence induces a conformational change of the DNA:RNA duplex towards the A-type helix[25] but prevents the RNase H cleavage of the target RNA.
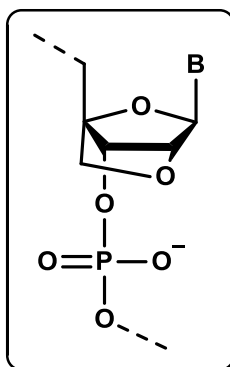
FIGURE 1.12 Locked nucleic acid (LNA)

LNAs and LNA:DNA chimeras have been shown to efficiently inhibit gene expression when targeted to a variety of regions within the luciferase mRNA.[26] LNA shows remarkably increased hybridization properties relative to a DNA:RNA duplex and improves nuclease resistance. Few analogs of LNA have been reported, which have improved activity and/or toxicity profiles in animals.[27,28] Thus, LNA offers attractive properties, such as nuclease stability, high target affinity, potent biological activity and lack of acute toxicity.

Morpholino oligonucleotides (MF)

Phosphorodiamidate morpholino oligonucleotide has morpholine ring replacing the furanose ring in DNA/RNA. It has the phosphorodiamidate linkage which connects the morpholine nitrogen atom with the hydroxyl group of the 3'-side residue (Figure **1.13**).[21]
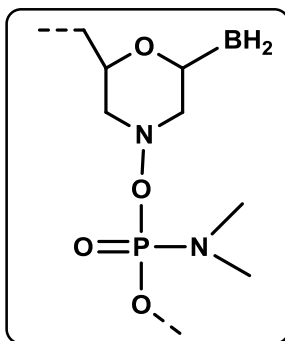
FIGURE 1.13 Morpholino oligonucleotide (MF)

Morpholino ONs are non-ionic, and therefore unlikely to have unwanted electrostatic interactions with nucleic acid binding proteins. These ONs are stable to nucleases and have similar affinity as in DNA:DNA duplexes. However, morpholino ONs do not activate the RNase H and are primarily used in translation arrest or in other steric blocking mechanisms, such as alteration of splicing.[29,30]

Cyclohexene nucleic acids (CeNA)

Cyclohexene nucleic acids (CeNA) contain a six membered cyclohexene ring as a replacement for five membered furanose ring which increases the conformational rigidity of the oligomers (Figure **1.14**). They form stable duplexes with complementary DNA/RNA and are resistant to nucleases.[31] In addition, CeNA:RNA hybrids have been reported to activate RNase H.[32]
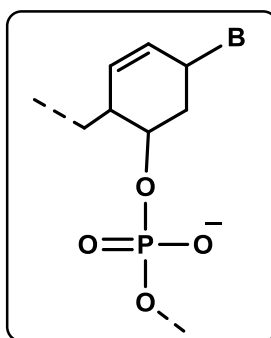
FIGURE 1.14 Cyclohexene nucleic acid (CeNA)

12

N3'-P5' phosphoroamidates (NPs)

N3'-P5' phosphoroamidate (NPs) is a modified phosphate backbone, in which the 3'-oxygen of the deoxyribose ring is substituted with a amino group (Figure **1.15**). NPs are resistant to nucleases and have high affinity towards a complementary RNA strand but do not activate the RNase H.[33] The sequence specificity of phosphoroamidate-mediated antisense effects by steric blocking of translation initiation has been demonstrated in cell culture and *in vivo*.[34]
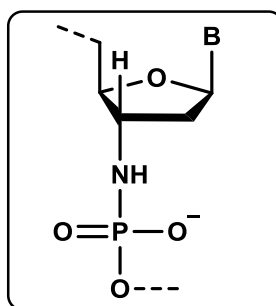

FIGURE 1.15 N3'-P5' phosphoroamidates (NPs)

1.6 Peptide nucleic acids (PNA)

Peptide nucleic acids (PNAs) are a radically different class of oligonucleotide modifications that contain a peptide link instead of phosphate link. PNA was first introduced by Peter Nielsen and coworkers in 1991.[35] PNAs are highly resistant to degrading enzymes like proteases and nucleases, exhibit high binding affinity towards target DNA/RNA, but do not activate RNase H and, as such, have been used primarily in translation inhibition and splicing modulation antisense mechanisms. PNAs seem to be non-toxic, as they are uncharged molecules with low affinity for proteins that normally bind nucleic acids.

PNAs are DNA analogs where the sugar phosphate backbone is replaced with a pseudopeptide backbone in the form of 2-aminoethyl-glycine linkage. Nucleobases are attached through a methylene carbonyl linker to this backbone at the amino nitrogen. The PNA backbone is constituted by six atoms for each repeating unit and a two atom spacer between the backbone and the nucleobase, similar to the natural DNA (Figure **1.16**).[36] PNA was originally designed and developed as a mimic of a DNA recognizing, major groove binding, triplex-forming oligonucleotide.[36,37] However, the polyamide

backbone of PNA has proven to be a surprisingly good structural mimic of the sugar phosphate backbone of nucleic acids.
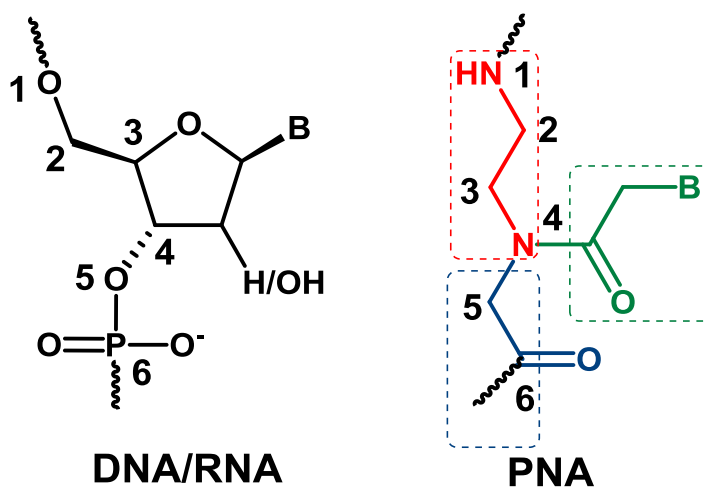

FIGURE 1.16 Chemical structures of DNA/RNA and PNA

The advantages of PNA over the conventional antisense oligonucleotides are numerous, partially due to the high flexibility and the absence of charge in the backbone. PNAs being neither peptide nor nucleic acids are resistant to both proteases and nucleases and consequently have a longer life span in the cellular environment. PNA hybridizes with complementary DNA/RNA with thermal stability superior to DNA:DNA or DNA:RNA duplex. This results from a decrease in electrostatic repulsion between the two negatively charged strands in DNA/RNA duplexes. They have higher mismatch discrimination and form selective duplexes upon binding to complementary DNA or RNA sequences. Therefore, PNA has attracted wide attention in medicinal chemistry for development of gene therapeutics, especially antisense or antigene drugs.

1.6.1 PNA structure

PNA binds to a complementary DNA/RNA through a classical Watson-crick base pairing mechanism. The PNA bases form a helical π-stack similar to DNA but the smaller twist of the PNA double helix and the larger π-overlap between the neighboring bases makes it different from the DNA. NMR methods and X-ray crystallography have been used to derive the three-dimensional structures of the major families of PNA complexes including PNA:RNA,[38] PNA:PNA,[39] PNA:DNA duplexes[40] and PNA$_2$:DNA

triplex[41] (Figure **1.17**). Although PNA oligomer, to some extent is able to structurally adapt to the oligonucleotide complement, it also has a preferred structure of its own, termed the '*P-form*' helix.

a.

b.

PNA:RNA   **PNA:DNA**   **PNA₂:DNA**   **PNA:PNA**
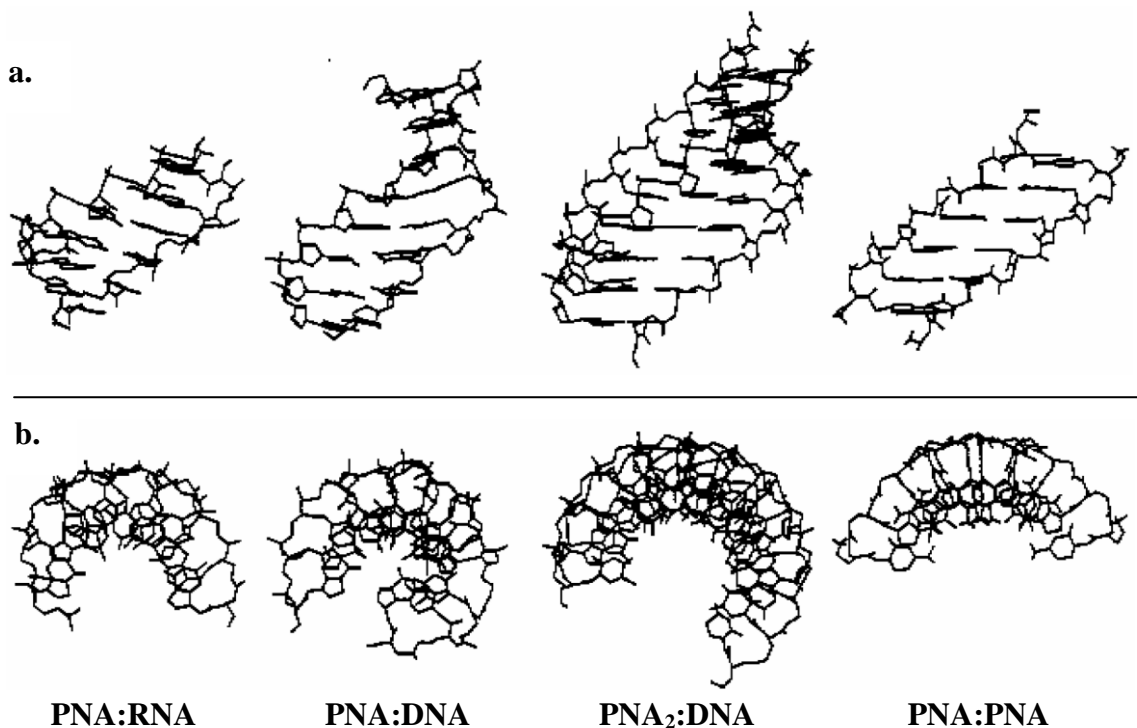
FIGURE 1.17 Structures of various PNA complexes shown in a) side view and b) top view[42]

The *P-form* helix constitutes a very wide PNA duplex (28 Å diameter) with an accordingly large base pair helical displacement and a very large pitch (18 bp). A canonical *B-form* helix which is typical for DNA duplexes has a diameter of 20 Å and a pitch of 10 bp per turn. A canonical *A-form* helix, typical of RNA duplexes, also has a diameter of 20 Å but a pitch of 11 bp per turn, and the base pairs are tilted 20° relative to the helix axis. Also, the base pairs are displaced away from the helix leaving a central 'tunnel' in the helix, analogous to that seen in the *P-form*. These structures suggest that PNA can adapt well to its nucleic acid partner, as the RNA strand in the PNA:RNA duplex is essentially in A-form conformation, whereas PNA:DNA duplex adopts a B-form conformation.

1.6.2 Chemical and physical properties of PNA

PNA has proved itself a promising antisense or antigene agent on the basis of its superior properties, such as highly sequence specific binding to the complementary

DNA/RNA targets, high biological and chemical stability, high mismatch discrimination etc.

1.6.2a Duplex formation with complementary oligonucleotides

PNA was originally designed for sequence specific targeting of double stranded DNA via major groove recognition. PNA hybridizes to complementary oligonucleotides obeying Watson-Crick base pairing rule. Though in DNA:DNA duplexes, the two strands are always in antiparallel orientation (with the 5'-end of one strand opposed to the 3'-end of the other), PNA:DNA adducts can be formed in two different orientations, arbitrarily termed *parallel* and *antiparallel* (Figure **1.18**). Both adducts are formed at room temperature, with the antiparallel orientation showing higher stability.[43] This creates the possibility for PNAs to bind two DNA tracts of opposite sequence.
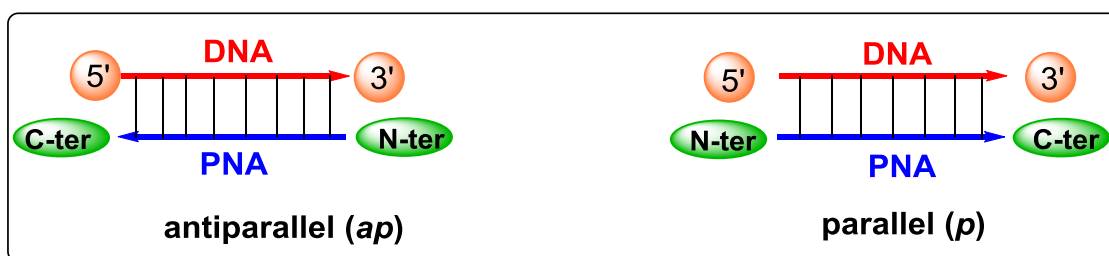
FIGURE 1.18 Antiparallel and parallel modes of PNA:DNA duplex formation

1.6.2b Triple helix formation of PNA

Polypyrimidine PNAs are able to form stable adducts with complementary polypurine DNA, through the formation of PNA_2:DNA triplexes.[44] The base pairing in these complexes occurs via Watson-Crick and Hoogsteen hydrogen bonds. When only one PNA strand is used to form a PNA_2:DNA triplex, both strands are necessarily either antiparallel or parallel to DNA strand. Whereas, when two different homopyrimidine PNA sequences are used, Watson-Crick PNA strand is oriented in antiparallel and the Hoogsteen strand is in parallel orientation to form a stable triplex with homopurine strand of DNA. The sequence specificity of triple helix formation is based on the selectivity of formation of the intermediate PNA:DNA duplex, whereas binding of the third strand contributes only slightly to selectivity. The stability of these

structures enables PNA to perform strand invasion[44,45], a property which is uniquely shown by PNAs.

1.6.2c Strand invasion by PNA

The unique property of PNAs to displace one strand of DNA double helix to form strand invasion complexes,[46] is a favorable attribute for their application as antigene agents. Four modes of binding for sequence specific targeting of double stranded DNA by PNA have been identified (Figure **1.19**), three of which involve invasion of the DNA duplex by PNA strands.

| Triplex | Triplex Invasion | Duplex Invasion | Double Duplex Invasion |
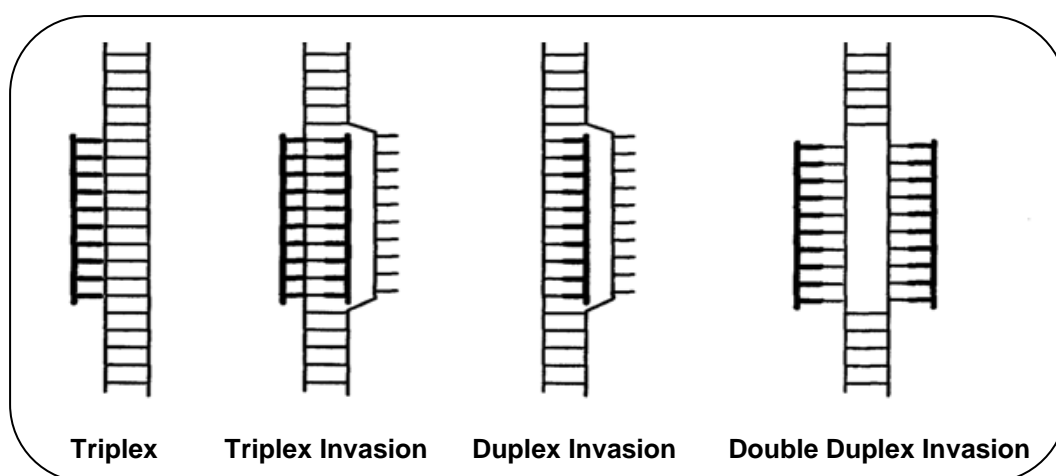
FIGURE 1.19 Various modes for binding of PNA to sequence specific targets in *ds*DNA[47]

It is possible for PNA (homopurine) single strand to either invade (duplex invasion) via Watson-Crick base pairing, or alternatively, invasion may be accomplished by two pseudo-complementary PNA strands, each of which binds to one of the DNA strands of the target (double duplex invasion). These pseudo-complementary PNAs contain modified adenine and thymine nucleobases[47] (Figure **1.20**) that do not allow stable hybridization between the two complementary sequence PNAs, but does permit good binding to the DNA.

The 'triplex invasion' requires a homopurine DNA target and complementary homopyrimidine PNAs that bind the purine DNA strand through combined Watson-Crick- Hoogsteen base paring *via* formation of a very stable PNA_2:DNA triplex. For most applications, the two PNA strands are connected in a bis-PNA designed such that the one strand is antiparallel (W-C strand) and the other strand is parallel (H-strand) to the DNA target.
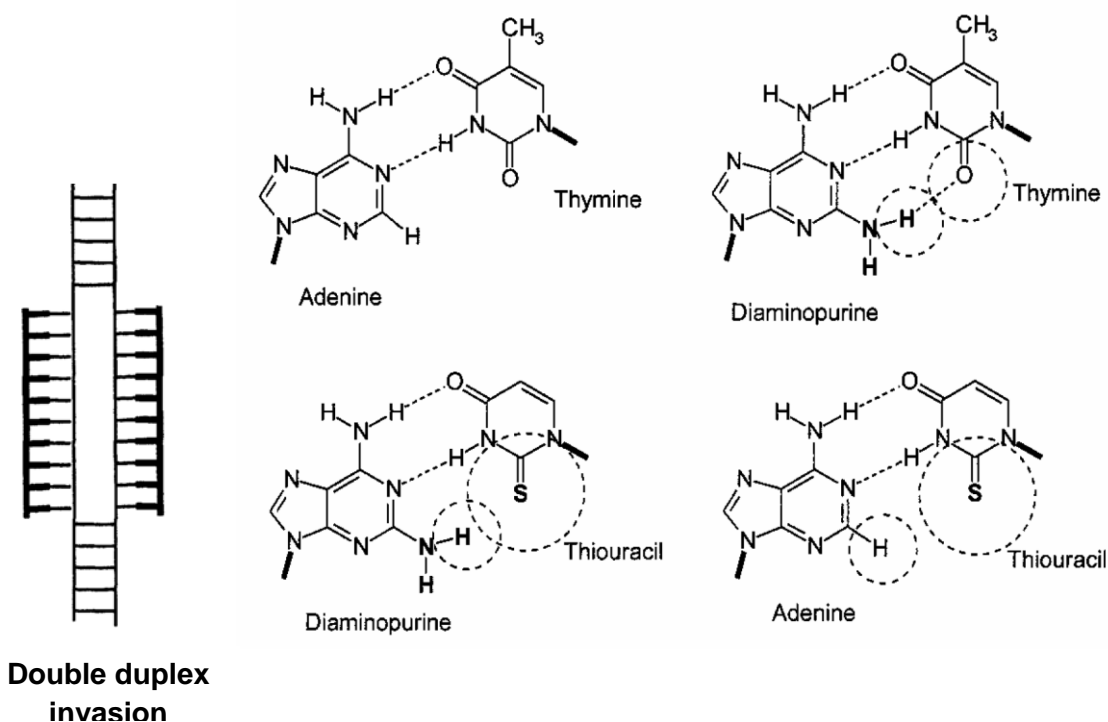
FIGURE 1.20 Double duplex invasion of peudocomplementary PNAs[47]

1.6.2d Quadruplex formation by PNA

The novel supramolecular architecture of G-quartets has led to the development of interesting and functional non-covalent assemblies such as G-wires,[48] ion-channels[49] and self-assembled ionophores.[50] PNAs have been developed to mimic Watson-Crick and Hoogsteen base-pairing, and they are expected to participate in G-tetrad formation as well. In an attempt to use this mode of molecular recognition, homologous G-rich PNA and DNA oligomers hybridize to form a $PNA_2:DNA_2$ quadruplex. The hybrid quadruplex exhibits high thermodynamic stability and expands the range of molecular recognition motifs for PNA beyond duplex and triplex formation. Krishnan-Ghosh *et al.*[50] have also shown the quadruplex formation composed entirely of PNA (Q-PNA).

1.6.2e Higher specificity towards target nucleic acid

PNA shows greater specificity in binding to complementary DNA. A PNA/DNA mismatch is more destabilizing than a mismatch in a DNA/DNA duplex. A single mismatch in mixed PNA:DNA (15-mer) sequence, destabilizes the duplex by 8-

20 °C (15 °C on average). Whereas, in the corresponding DNA:DNA duplex, a single mismatch destabilizes the duplex by 4-16 °C (11 °C on average).[51]

1.6.2f Stronger binding independent of salt concentration

Another important consequence of the neutral backbone is that the T_m values of PNA:DNA duplexes are practically independent of salt concentration. In contrast, the T_m values of DNA:DNA duplexes are highly dependent on ionic strength.[52]

1.6.2g Solubility of PNA

PNAs are charge-neutral compounds and hence have poor water solubility compared with DNA. Neutral PNA molecules have a tendency to aggregate to a degree that is dependent on the sequence of the oligomer. PNA solubility is also related to the length of the oligomer and to the purine/pyrimidine ratio.[52,53] Some of the recent modifications, including the incorporation of positively charged lysine or arginine residues (carboxy-terminal during solid phase synthesis or backbone modification in place of glycine), have shown improvements in solubility of PNA.

1.6.2h Cellular uptake of PNA

Although, PNA binds to complementary DNA/RNA with high affinity, specificity and stability in biological fluids, the progress in the exploration of PNA as antisense/antigene agents and gene expression regulation has been hampered by their poor cellular uptake. Thus efficient cellular delivery systems for PNAs are required if these are to be developed into antisense and antigene agents. However, a number of transfection protocols for PNA have been established like microinjection, electroporation, co-transfection with DNA, conjugation to lipophilic moieties, conjugation to cell penetrating peptides etc.[54] To address the issues like poor cell penetration, solubility and ambiguity in binding orientation, various chemical modifications of PNA have been employed.

1.6.3 Chemical modifications of PNA

Since their discovery, many modifications of the basic *aeg* PNA structure have been proposed, in order to improve their performance in terms of affinity and

specificity towards complementary oligonucleotides, solubility and cell permeation abilities etc.

1.6.3a Extension of PNA backbone

The first approach to modify *aeg* PNA, was extension of the PNA backbone by one methylene group at ethylenediamine side, glycine side or attachment of nucleobase though ethylenecarbonyl instead of methylenecarbony linker (Figure **1.21**).[55,56] The extension of PNA backbone by one methylene group drastically decreased the stability of derived PNA:DNA duplexes compared to unmodified PNA:DNA duplex. These studies revealed that the observed detrimental effect on the duplex stability resulted from the disturbance of inter-nuclobase distance which is an important criterion to have a stable PNA:DNA duplex.
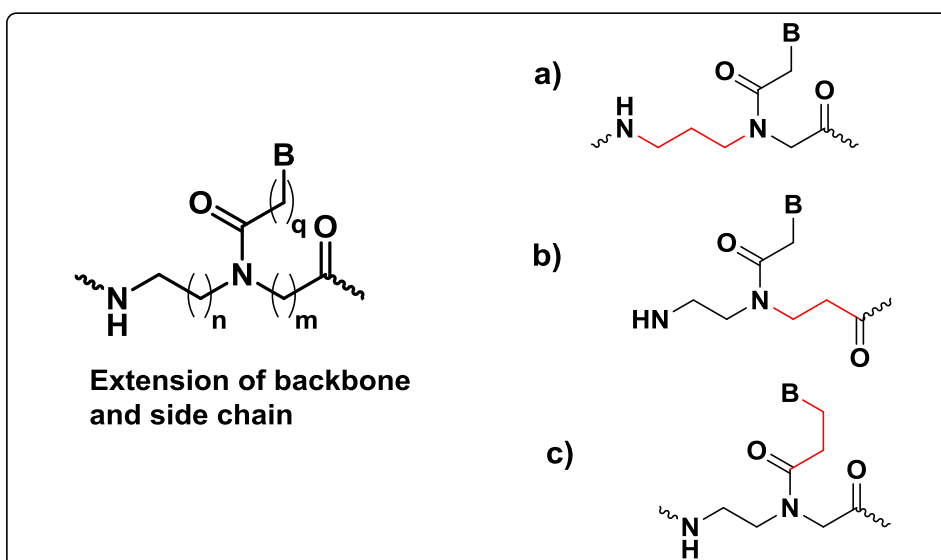
FIGURE 1.21 Extension of PNA backbone at a) ethylenediamine side b) glycine side and c) nucleobase attachment side

A structural feature of PNA that interferes with hybridization is the presence of *cis* and *trans* rotamers around the tertiary amide linkage in each PNA monomer. Attempts to control the rotameric populations have resulted in the construction of olefinic polyamide nucleic acids, **OPAs**, through the introduction of a defined (*E*)- or (*Z*)-C-C double bond (Figure **1.22**) instead of the tertiary amide linker.[57]
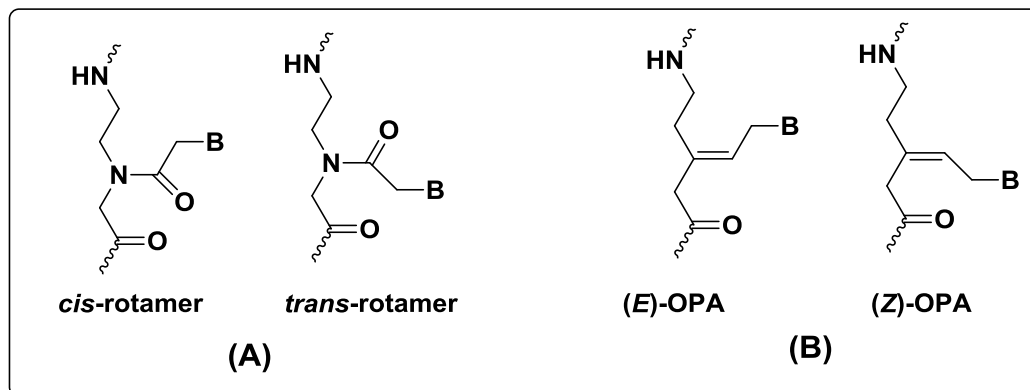
FIGURE 1.22 (A) cis and trans PNA rotamers (B) E- and Z- olefinic PNA

As a strategy to improve binding affinity and solubility of PNA, rigid and uncharged tertiary amide group was replaced with tertiary amine which is positively charged at physiological pH[58] (Figure **1.23**). This cationic PNA was highly soluble in water but it reduced the binding affinity towards target nucleic acids due to unrestricted conformational freedom.
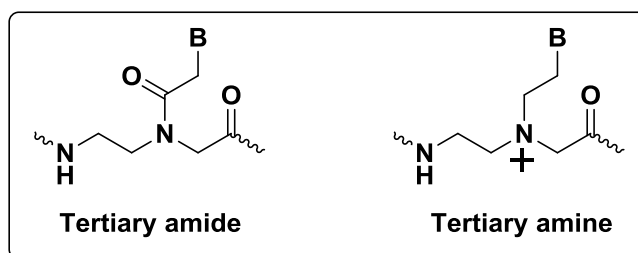
FIGURE 1.23 Replacement of tertiary amide with tertiary amine

1.6.3b Preorganization of PNA

To have a good DNA binding, a 'constrained flexibility' is necessary since too flexible or too rigid structures could not give efficient PNA:DNA binding. A highly flexible backbone would require a larger loss of entropy for efficient binding whereas a very rigid backbone would prevent DNA binding due to a difficult fit to adopt to target structure.[59] Preorganization of PNA structure, i.e. the ability to adopt a conformation which is most suitable for DNA binding, is required to minimize the entropy loss of the binding process. The main strategies which have been used for achieving this goal are summarized in Figure **1.24**.
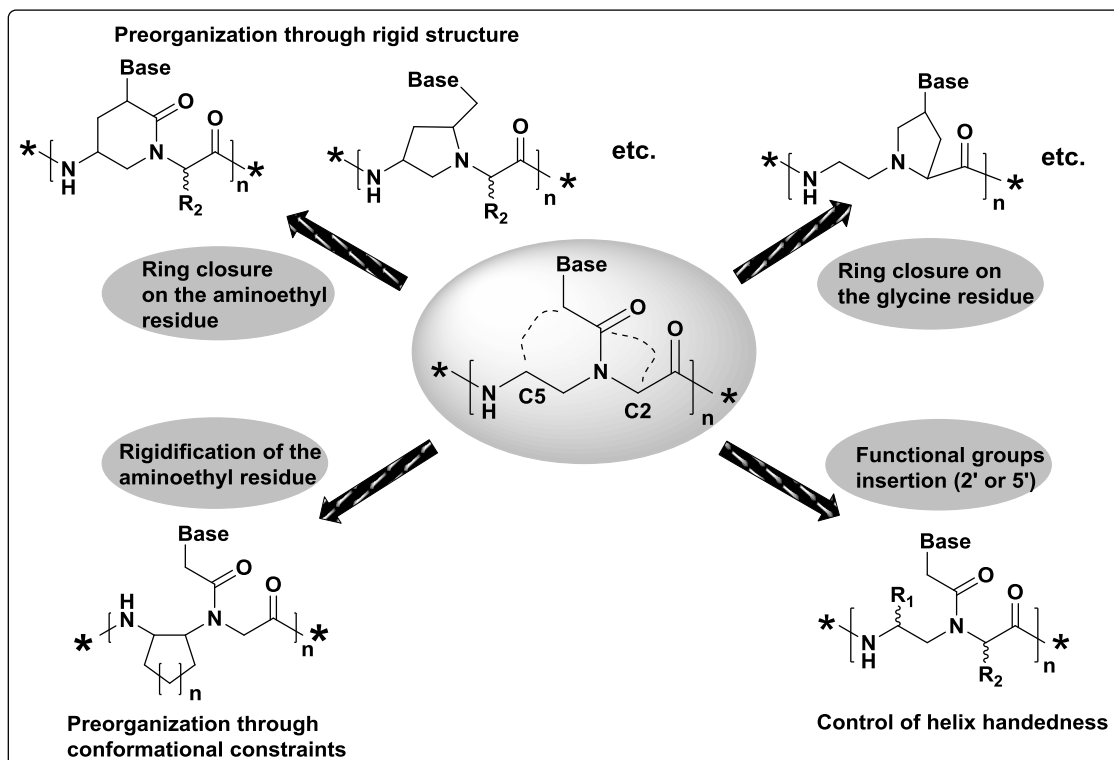
FIGURE 1.24 Strategies for inducing preorganization in the PNA structure

Preorganization can be achieved either by cyclization of the PNA backbone (in the aminoethyl side or in the glycine side), by adding substituents in the *C*2 or *C*5 carbon of the monomer or by inserting the aminoethyl group into cyclic structures. The addition of substituents at *C*2 or *C*5 carbon of the monomers can also in principle preorganize the PNA strand. But it has the effect of shifting the PNA preference towards a right-handed or left-handed helical conformation, according to the configuration of the new stereogenic centers. This in turn affects the stability of the PNA-DNA duplex through interference in the helix handedness. The two classes of PNA with conformationally constrained backbones which have shown improved properties are acyclic PNAs with functional backbones and cyclic PNAs, in which preorganization is achieved through conformational constraint or through rigidification of the backbone.

A) Preorganization through cyclic PNAs

Systematic efforts have been carried out in rational design and synthesis of conformationally constrained PNA analogues towards evolving PNAs for selective and

directional recognition of DNA or RNA. Conformational preorganization has been addressed which are based on introduction of methylene or ethylene groups to bridge the aminoethyl glycyl backbone and methylene carbonyl side chain to generate diverse five or the six membered nitrogen heterocyclic analogues (Figure **1.25**).[59] The cyclic analogs where the nucleobases are directly attached to the ring have defined nucleobase orientation, overcoming the rotamer problem. It also concomitantly introduces chiral centers, which may impart directional selective binding of PNA with chiral DNA/RNA.
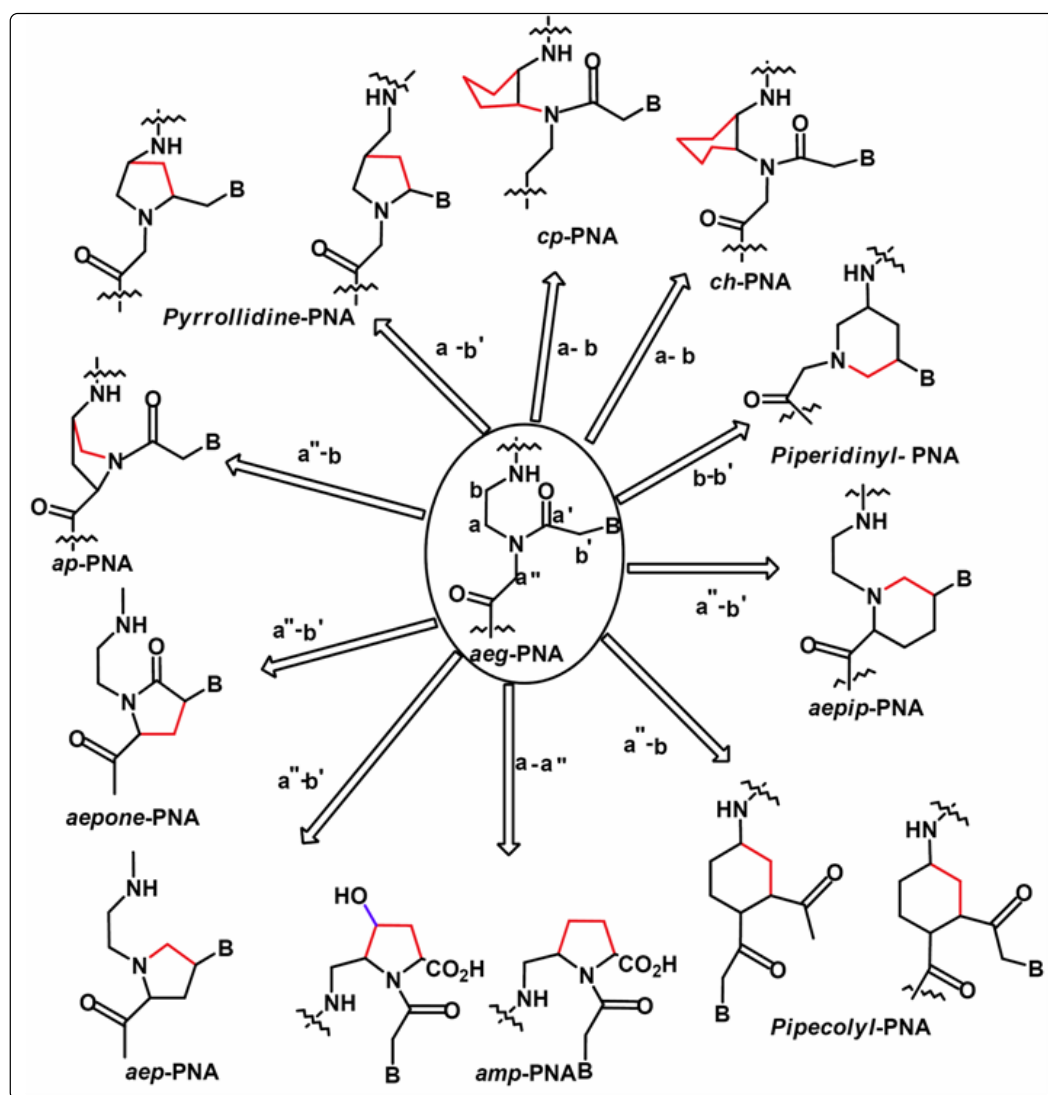
FIGURE 1.25 Cyclic and conformationally constrained PNA analogs[59]

Conformationally restricted proline-based PNA in which a pyrrolidine ring replaces the tertiary amide linker to the nucleobase and thus balance flexibility and rigidity in the PNA backbone was derived in the form of **aminoethylprolyl (*aep*) PNA**. The α"-carbon atom of the glycine unit and the β'-carbon atom of the nucleobase linker

were joined through a methylene bridge.[60] The flexibility in the aminoethyl segment of *aeg* PNA was retained. The nucleobase attachment to the pyrrolidine ring was fixed by virtue of the chirality of *C*-4, thus removing the possibility of any rotameric populations. The tertiary amine function in the backbone was found to be at least partially protonated at physiological pH (p*K*a ~ 6.8). Thus, in *aep* PNA, all elements of the structural freedom of *aeg* PNA were conserved in addition to the restriction of the rotamers. The oligomers comprising (4*S*, 2 *R/S*) *aep* PNA (T) units showed very favorable binding properties towards the target sequences without compromising the specificity.

Introduction of a methylene bridge between the *β*-carbon atom of the aminoethyl segment and the α"-carbon of the glycine segment of *aeg*- PNA resulted in **4-aminoprolyl (*ap*) PNA** having two chiral centers.[61] None of the homochiral aminoprolyl thyminyl PNAs corresponding to any of the diastereomers bound to target DNA sequences,[62] probably due to high rigidity in the backbone resulting in structural incompatibility. However, incorporation of single chiral D-trans or L-trans prolyl PNA monomer into *aeg*-PNA at the N-terminus or within the PNA sequence resulted in higher binding of the target DNA with definite preference for a parallel or an antiparallel mode unlike the unmodified PNA.[63]

One of the earliest efforts to introduce conformational constraint in the *aeg* PNA structure resulted in the chiral cyclohexyl-derived backbone (***ch*-PNA**).[64] The conformational freedom in the ethylene chain was thus locked in the six-membered cyclic structure. The oligomers with (*S*,*S*)-cyclohexyl residues were able to hybridize with DNA or RNA, with little effect on the thermal stability depending on their number and the sequence. A decrease in entropy was observed from thermodynamic measurements indicating the conformationally constrained structures for these cyclic PNAs. From the thermal stabilities and molecular modeling based on the solution structure of a PNA-DNA duplex determined by NMR techniques, it was concluded that the right handed hybrid duplex accommodated the (*S*,*S*) isomer more easily than the (*R*,*R*) isomer.

A relatively flexible system with a cyclopentyl ring in which the characteristic *endo-exo* puckering that dictates the pseudoaxial/pseudoequatorial dispositions of substituents allows better torsional adjustments to attain the necessary hybridization-competent conformations. In an attempt to tune the dihedral angle β in this manner the

cyclohexyl unit was replaced with *cis*-(1*S*,2*R*/ 1*R*,2*S*)-cyclopentyl PNA-T (**cp-PNA**) monomer.[65] The crystal structure data of the monomer indicated that the dihedral angle β was around 24° on the low side of the desired value of 60°. The *RS*-*cp*-PNA enantiomer formed higher affinity complexes with DNA as compared to *SR*-*cp*-PNA isomer.

The designs of *ch*-PNA and *cp*-PNA are the outcome of optimized dihedral angles (Figure **1.26**)[59] that constrain the PNA backbone for differential DNA/RNA binding and discrimination via preorganization. The inherently rigid *cis*-substituted six-membered ring of *ch*-PNAs forbids structural readjustments to bind to DNA (PNA: DNA, β=140°) and prefers binding to RNA (PNA:RNA). The flexible (*SR/RS*)-*cp*-PNA with a relative ease of conformational adjustments in the *cis*-substituted cyclopentyl system allows reorganization of the ring puckering and binding to DNA/RNA with high affinity and no selectivity.

Peptide Nucleic acid (PNA), n=0 *aeg*PNA; n=3, *cp*PNA; n=4,*ch*PNA
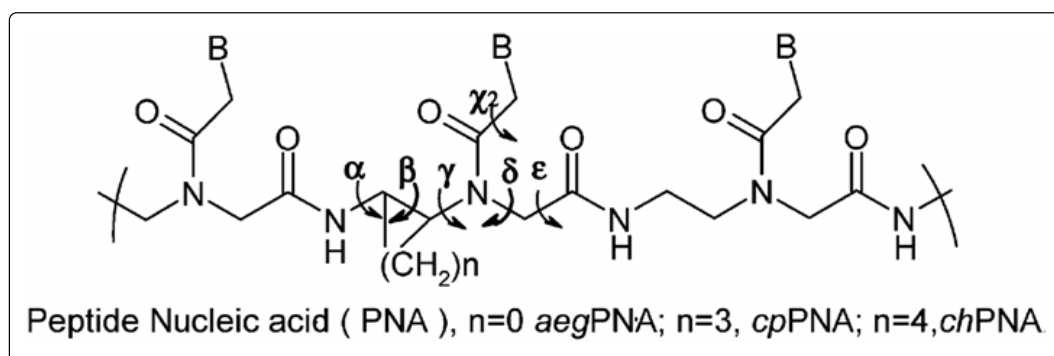
FIGURE 1.26 Dihedral angles between *aeg* PNA, *cp*PNA and *ch*PNA[59]

B) Preorganization through acyclic PNAs

Standard PNA oligomers seem to be a convenient compromise, being both flexible (since they are acyclic and with a very flexible aminoethyl moiety) and rigid (two amide bonds per monomeric residue). It was shown that the insertion of a methylene group in any position of the PNA backbone or in the linker between the backbone and the nucleobase resulted in a net loss of DNA binding properties.

Using the linear *N*-(2-aminoethyl) glycine as a starting point, several PNA derivatives were obtained by insertion of side chains either at the *C*2 (α) or *C*5 (γ) carbon atoms (Figure **1.27**). These modifications have the effect of shifting PNA "natural" constrained flexibility towards more constraints and less flexibility.

R' = R" = H: standard achiral PNA
R'= side chain, R" = H :   acyclic chiral PNA substituted at position 2
R' = H, R" = side chain :   acyclic chiral PNA substituted at position 5
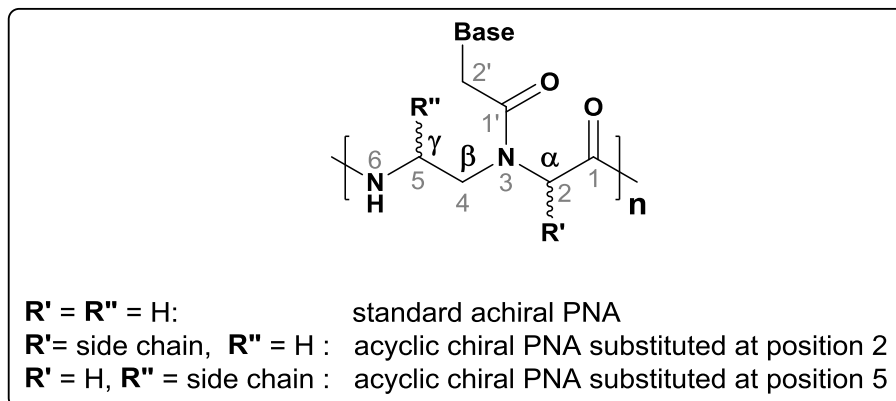
FIGURE 1.27 Achiral and chiral acyclic PNAs

L/D-amino acid synthons have been used to synthesize $C2$-substituted and $C5$-substituted chiral PNAs with desired stereochemistry (Figure **1.28**). Chiral, C2-substituted PNAs derived from alanine,[66] arginine and lysine[67] side chains showed enhanced PNA:DNA stability due to the small steric hindrance and the electrostatic interaction with negatively charged phosphates in DNA.

Apolar amino acids, such as, valine, glutamine, asparagine, tryptophan and phenylalanine showed decreased binding affinity towards target nucleic acid; while negatively charged amino acids such as, glutamic acid and aspartic acid drastically reduced the derived PNA:DNA duplex stability due to the charge repulsion.[68a] PNA:DNA duplex stability is dependent on stereochemistry. When the binding affinity of chiral PNAs including L and D stereochemistry of alanine, lysine, serine, aspartic acid, leucine was considered, PNAs carrying the D-amino acid derived monomers bound complementary antiparallel DNA strands with higher affinity than the corresponding L-monomers.[67] Gourishankar *et al.*[68b] have recently reported the synthesis of sterically constrained new analogs of PNA having gem-dimethyl substitution at glycine. They show superior binding to DNA than isosequential RNA.

The insertion of stereogenic centers in the PNA strand results in a predominant helix handedness. As revealed from CD spectroscopy, PNAs containing D-amino acid derived monomers with the stereogenic center at $C2$-position (R stereochemistry) induced a preference for a right-handed conformation in PNA-PNA duplexes, whereas PNAs containing L-amino acid derived monomers with the stereogenic center in the same position (S stereochemistry) induced an opposite preference for a left-handed double helix. However, at $C5$-position, L-amino acid derived PNAs (S stereochemistry) prefer a right-handed helical conformation and D-amino acid derived PNAs (R

stereochemistry) induce a left-handed conformation in the duplex. PNAs preferring a right-handed helical conformation have higher DNA binding affinity than their mirror images.[69,70]

FIGURE 1.28 Various modified acyclic PNAs

To address the issue of poor cell permeability of *aeg* PNA, Zhou *et al.*[71] have introduced guanidinium group at C2-position of PNA backbone (α-GPNA) (Figure **1.29 a**). Thermal stability studies showed that although the guanidinium group of α-GPNA is positively charged at physiological pH, incorporation of one modified unit

into a decamer PNA sequence destabilized the duplex by 2-4 °C; however, incorporation of multiple α-D-GPNA units at alternate positions in the sequence enhanced the stability of PNA:DNA duplex in antiparallel orientation.[72] Cellular uptake of fully modified α-L-GPNA was evaluated in human HCT116 (colon) and Sao2 (osteosarcoma) cell lines and it was found that these cationic PNAs permeated the cell membrane and appeared to localize specifically in the nucleus.[71] Cell penetration of α-GPNA especially in ES cells was very important because these cells are extremely difficult to transducer, even with best transfecting agents.[73]

Sahu *et al.*[74] have synthesized the second generation guanidino modified PNA analogs (γ-L-GPNA) based on the homo-arginine side chain at γ-position of PNA backbone (Figure **1.29 b**). The PNA:DNA duplex stability was found to be superior for γ-L-GPNA compared to α-GPNA. Cell permeation studies showed that γ-L-GPNAs are easily taken up by HeLa cells and the uptake efficiency was comparable to that of the TAT transduction domain.
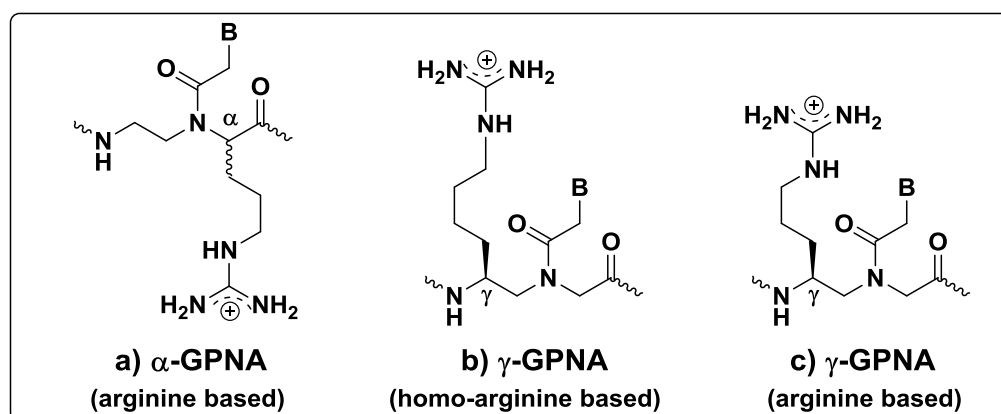
FIGURE 1.29 guanidino modified α- and γ-GPNA analogs

Manicardi *et al.*[75] have recently reported the inhibition of micro-RNA by GPNAs based on arginine side chain (Figure **1.29 c**). Anti-miR-210 activity of various 18-mer PNAs (*aeg* PNA, α-GPNA and γ-GPNAs) in leukemic K562 cells was examined. The placement of GPNA was either alternate in the sequence or consecutive at N-terminus. All modified PNAs were efficiently internalized and the fluorescence signals were found in the cytoplasm. The best anti-miR-210 activity was exhibited by γ-GPNAs with consecutive placement.

Mitra *et al.*[76,77] have reported the synthesis of chiral and cationic PNAs with aminomethyl substitution (*am*-PNAs) at α- or γ-position (Figure **1.30**). Aminomethyl PNAs formed more stable PNA:DNA duplexes compared to unmodified *aeg* PNA and

the order of stabilization was γ-(*S*)-am PNA > α-(*R*)-am PNA > α-(*S*)-am PNA. The aminomethyl PNAs were easily taken up by HeLa cells and they were found to be localized in the nucleus. The degree of cellular uptake efficiency for *am*-PNAs was found to be same as PNA:DNA duplex stabilization order.
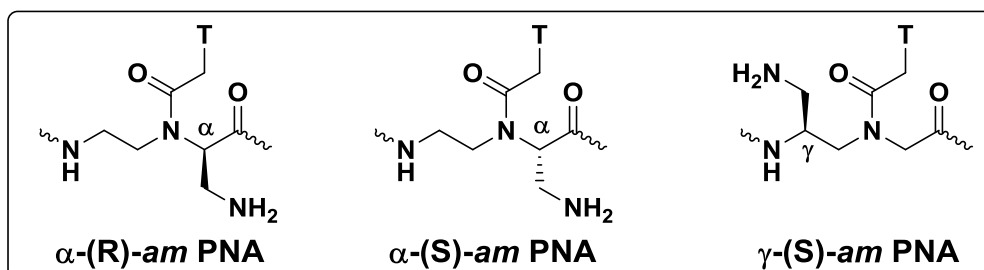
FIGURE 1.30 Structures of aminomethyl PNAs

1.6.3c Nucleobase modifications

Employing non-natural nucleobase ligands in place of natural nucleobases would help understand the recognition process in terms of various factors contributing to the event such as hydrogen bonding and internucleobase stacking. Various modified nucleobase structures are summarized in Figure **1.31**. 2,6-diaminopurine[78] offers increased affinity and selectivity for thymine while pseudoisocytosine[79] is a very efficient mimic of protonated cytosine for triplex formation.
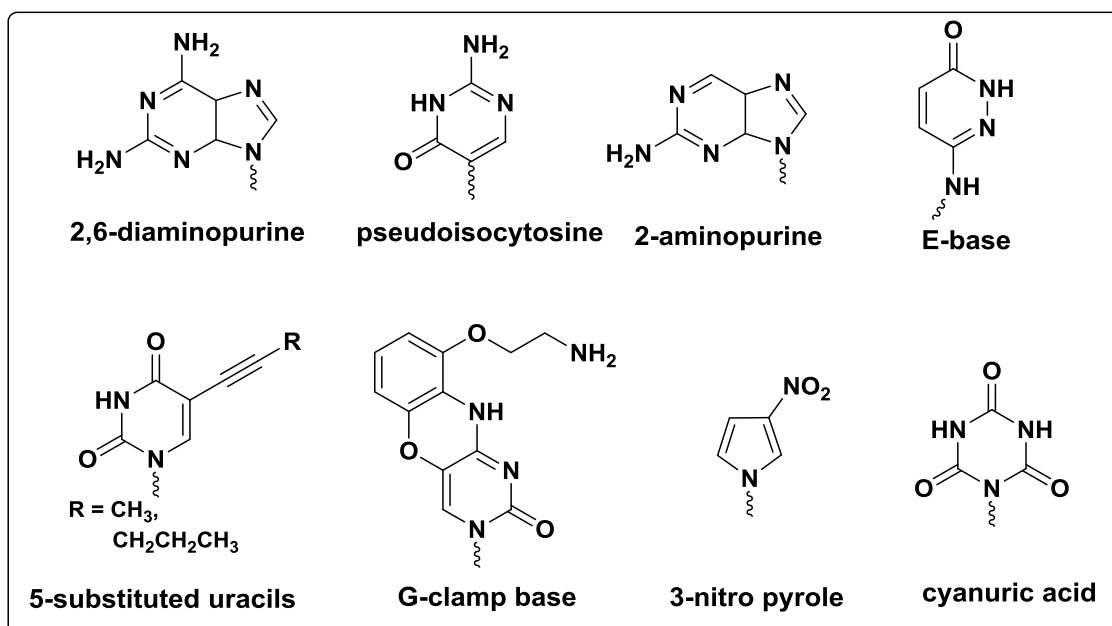
FIGURE 1.31 Non-standard nucleobases used in PNA

2-aminopurine[80] hydrogen bonds with U and T in reverse Watson-Crick mode and has the advantage of being inherently fluorescent to enable the study of kinetic events associated in hybridization. The E-base[81] was rationally designed for recognition of A:T base pair in the major groove and form a stable triad with T in the central position. A wide variety of 5-substituted uracils were synthesized and their ability for triplex formation has been studied.[82] Cyanuryl PNA[83] containing cyanuric acid as a base has been synthesized and incorporated into a PNA sequence as a mimic of T. The biophysical studies on their triplex/duplex complexes with complementary DNA oligomers indicated unusual stabilization of PNA:DNA hybrids when the cyanuryl unit was located in the middle of the PNA sequence.

1.6.4 Applications of PNA

This section describes diverse PNA applications in gene therapeutics, molecular biology, functional genomics etc.

1.6.4a PNA as antisense and antigene agents

PNAs can be used to design gene therapeutic agents due to their unique strand invasion property and their chemical as well as biological stability. There are basically two strategies involved in using PNAs as therapeutic drugs, namely antigene and antisense methods. Moreover, no sign of any general toxicity of PNA has so far been observed.

A) Inhibition of transcription

PNAs are capable of arresting transcriptional processes by virtue of their ability to form a stable triplex structure or a strand invasion or strand displacement complex with DNA. Such complexes can create a structural hindrance to block the stable functioning of RNA polymerase and thus are capable of working as antigene agents (Figure **1.32**).[84] PNA targeted against the promoter region can form a stable PNA/DNA complex that restricts the DNA access of the corresponding polymerase. Nielsen *et al*. have demonstrated that even an 8-mer PNA (T_8) is capable of blocking phage T_3 polymerase activity.[85]
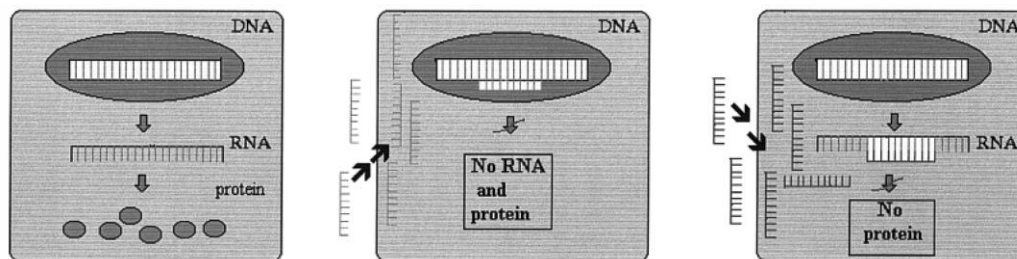
FIGURE 1.32 Antigene and antisense inhibition strategy[84]

B) Inhibition of translation

In the case of antisense strategy the nucleic acid analogues can be designed to recognize and hybridize to complementary sequences in *m*RNA and thereby inhibit its translation (Figure **1.32**). Normally, the PNA antisense effect is based on the steric blocking of either RNA processing, transport into cytoplasm, or translation. It has been concluded from the results of *in vitro* translation experiments involving rabbit reticulocyte lysates that both duplex-forming (mixed sequence) and triplex-forming (pyrimidine rich) PNAs are capable of inhibiting translation at targets overlapping the AUG start codon.[86a] Recently PNAs and PNA-peptide conjugates have been used for redirecting RNA splicing in therapeutic applications such as Duchenne Muscular Dystrophy in cells as well as animal models.[86b] In another example, PNA has also shown to inhibit microRNA, miR-155 which is expressed in haematopoietic systems in cultured B cells and in mice.[86c]

C) Inhibition of replication

PNA can also inhibit the elongation of DNA primers by DNA polymerase. Further, the inhibition of DNA replication should be possible if the DNA duplex is subjected to strand invasion by PNA under physiological conditions or if the DNA is single stranded during the replication process. Efficient inhibition of extra-chromosomal mitochondrial DNA, which is largely single-stranded during replication, has been demonstrated by Taylor *et al*.[87]

1.6.4b Interaction of PNA with enzymes

A) RNase H enzyme

Depending upon the chemical structure of RNase H stimulating oligonucleotides, the intracellular enzyme RNase H can be activated by oligonucleotides to cleave RNA bound to deoxyribonucleic acid oligomers. The

antisense oligonucleotide with an RNase H activity (e.g., phosphorothioate oligomers) is considered a better antisense molecule or inhibitor than one without the activity (methylphosphonates and hexitol nucleic acids).[88] Despite their remarkable nucleic acid binding properties, PNAs are not capable of stimulating RNase H activity on duplex formation with RNA except some DNA/PNA chimeras which are capable of stimulating RNase H activity. On formation of a chimeric RNA double strand, PNA/DNA can activate the RNA cleavage activity of RNase H (Figure **1.33**). Cleavage occurs at the ribonucleotide parts base paired to the DNA part of the chimera. This cleavage is sequence specific in such a way that certain sequences of DNA/PNA chimeras are preferred over others.

B) Polymerase and reverse transcriptase

In general, there is no direct interaction of PNA with either DNA polymerase or reverse transcriptase. However, different groups have shown indirect involvement of PNA in inhibiting these enzyme activities under *in vitro* conditions.
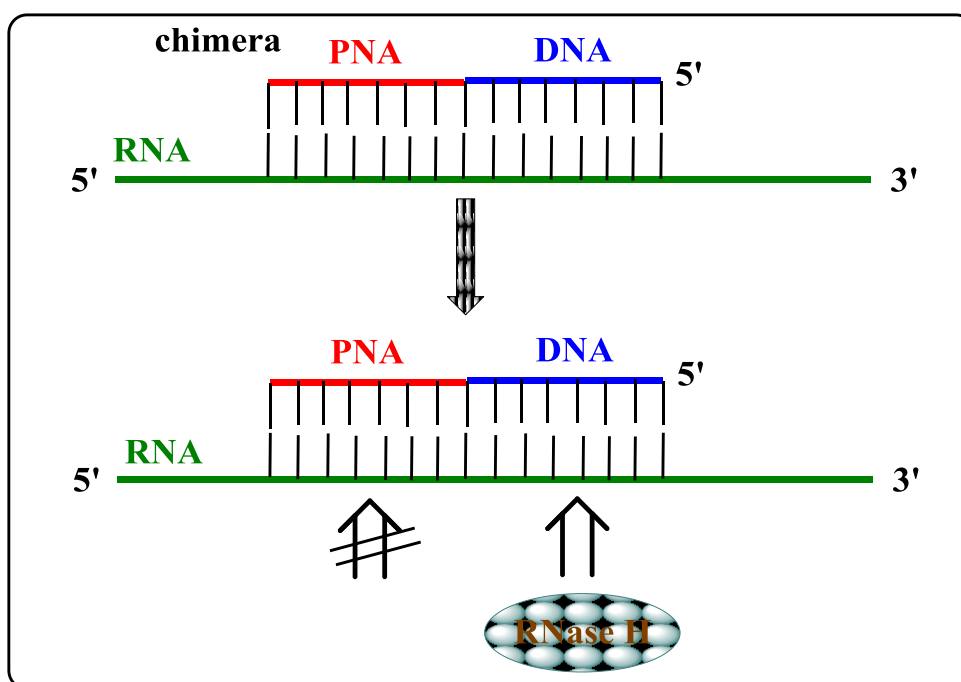
FIGURE 1.33 Schematic representation of RNase H-mediated cleavage activity after the binding of a PNA:DNA chimera to an RNA target

Primer extension by MMLV reverse transcriptase has been shown to be inhibited by introducing a PNA oligomer.[89] In another experiment, it was demonstrated

that the primer extension catalyzed by *Taq*-polymerase can be terminated by incorporating a PNA (PNA-H (t)$_{10}$) in to the system.[90] The reverse transcription of *gag* gene of HIV-I is also inhibited *in vitro* by PNAs. The inhibition has been achieved by using a bis-PNA construct, which is more efficient than the corresponding mono PNA construct.[91]

1.6.4c Tool for molecular biology and functional genomics

A) Enhanced PCR amplification

Molecular genetics applications such as amplification of VNTR (variable number tandem repeat) loci for genetic typing make large use of PCR (polymerase chain reaction), but preferential amplification of small allelic products relative to large allelic products presents a problem. This may result in incorrect typing in a heterozygous sample. By using PNA, enhanced amplification of a specific VNTR product is possible, and this has been done in the case of VNTR locus D1S80.[92] For PCR amplification, the template is blocked using a small PNA and becomes unavailable for intra- and inter-strand interactions during the re-association step. Although re-association is blocked by PNA, primer extension can occur. During extension the polymerase displaces the PNA molecules from the template, and the primer is extended towards completion of the reaction. This approach shows the potential of PNA application for PCR amplification where fragments of different sizes are required to be more accurately and evenly amplified. Since the probability of differential amplification is less, the risk of misclassification is greatly reduced.

B) Artificial restriction enzyme

S1 nuclease enzyme cleaves single-stranded nucleic acids releasing 5'-phosphoryl mono- or oligonucleotides. It removes the single-stranded overhangs of DNA fragments and can be used in RNA transcript mapping and construction of unidirectional deletions. PNAs in combination with nuclease S1 have been reported as the artificial restriction enzyme. Homopyrimidine PNA oligomers hybridize to the complementary targets on *ds*DNA via a strand invasion mechanism, leading to the formation of looped-out non-complementary DNA strands. The enzyme nuclease S1 can degrade this single-stranded DNA part into well defined fragments. If two PNAs

are used for this purpose and allowed to bind to two adjacent targets on either the same or opposite DNA strands, it will essentially open up the entire region, making the substrate accessible for the nuclease digestion and thereby increasing the cleavage efficiency.[93]

1.6.4d Single base pair mutation analysis using PNA-directed PCR clamping

Single-base-pair mutation or single-nucleotide polymorphism (SNP) analysis is possible using the PCR technique if PNA is synthesized targeting the primer binding site (Figure **1.34**).[94] Basically, in the PNA directed PCR clamping technique, at the annealing step the PNA is targeted against one of the PCR primer sites. The temperature set for this step is higher than that for normal PCR primer annealing where the PNA is selectively bound to the DNA molecule. The PNA, which binds to the primer binding site instead of the primer, effectively blocks the formation of a PCR product. PNA is able to discriminate between fully complementary and single-mismatch targets (mutations) in a mixed target PCR. Hence the binding of primer will be favored to out-compete PNA annealing. Consequently, mutated sequences will be preferentially amplified. This PNA clamping is able to discriminate three different point mutations at a single position.
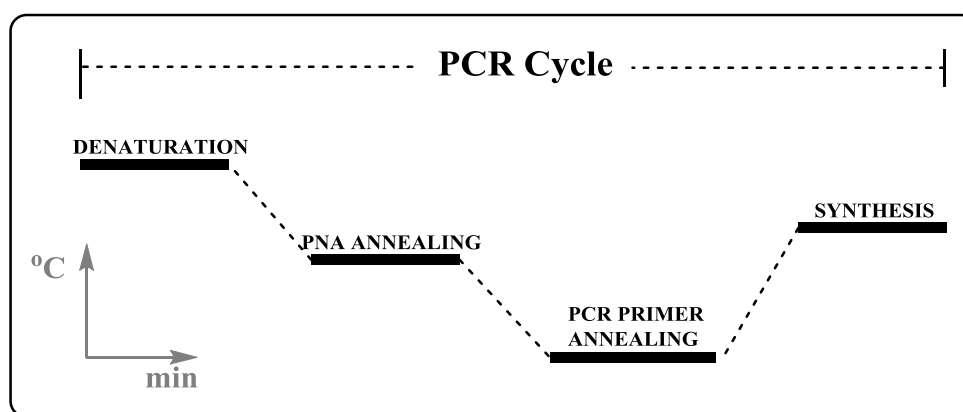
FIGURE 1.34 Schematic representation of the PCR cycle profile used in PNA-directed clamping

1.6.4e PNA as molecular beacons

Molecular beacons are oligonucleotide hybridization probes that can report the presence of particular nucleic acids in homogeneous solution. Molecular beacons are

hairpin shaped molecules with a fluorophore at one end and a quencher at the other (Figure **1.35**).[95] If molecular beacon (PNA) is in single stranded form, because of stem-loop structure of beacon, the fluorescence is quenched. However, upon hybridization with complementary nucleic acid, the fluorophore and quencher are separated resulting in the fluorescence signal.
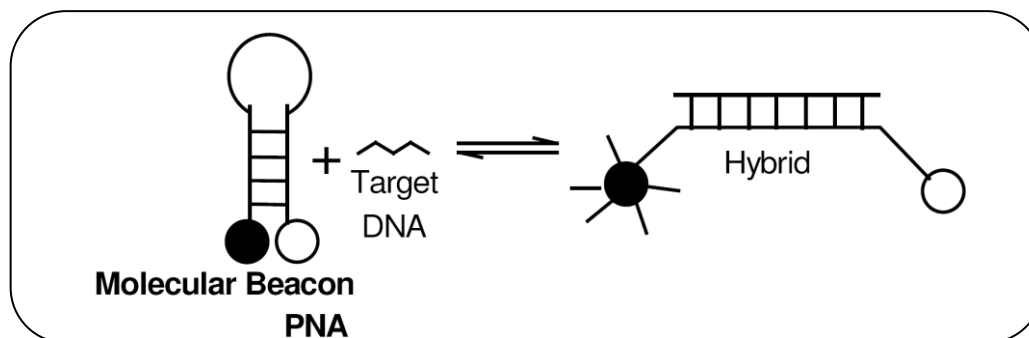
FIGURE 1.35 Schematic representation of molecular beacon

1.7 Scope of present work

The preceding section has given an overview on the concept of nucleic acids especially on peptide nucleic acid (PNA) which is a promising DNA mimic. PNA has properties of natural oligonucleotides and are important for the development of therapeutics in the form of antisense/antigene oligonucleotides. The major drawbacks of PNA like poor water solubility, inefficient cellular uptake, self aggregation and ambiguity in directional selectivity of binding restrict its applications in terms of designing PNA-based gene targeted drugs. Hence, various modifications of PNA to overcome these limitations have been employed by different research groups as described in previous section.

Chapter 2: Design, syntheses and characterization of γ-C-substituted peptide nucleic acid monomers and their oligomerization

This chapter aims to address some of the limitations like poor cellular uptake faced by standard *aeg* PNA with rationally designed PNA analogs having side chains carrying cationic (amine/guanidine) or charge-neutral azide substitutions at γ-position. The chapter is divided into two sections. The first section (**Section 2A**) deals with the design, synthesis and characterization of novel γ-C-substituted multifunctional PNA monomers (Figure **1.36**).
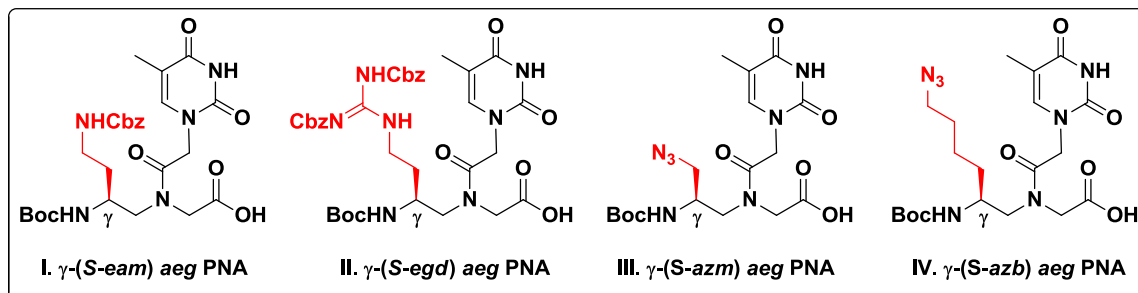
FIGURE 1.36 Chemical structures of modified γ-C-substituted PNA monomers

The presence of cationic functional groups may increase PNA–DNA duplex stability due to electrostatic interactions of cationic functionality with negatively charged DNA phosphate. The presence of cationic functionality in the side chain may improve the solubility and hence may help to enhance the cell penetration of these PNAs. The azide functional group can be utilized to attach fluorphore via click reaction with fluorescent alkyne derivative.

The incorporation of modified PNA units at various positions in the mixed purine-pyrimidine sequence using solid phase peptide synthesis is described in **Section 2B**. The cleavage of peptide from the solid support, purification by RP-HPLC and characterization by MALDI-TOF spectrometry is also described in this section.

Chapter 3 Biophysical evaluation of modified PNA oligomers

This chapter gives the biophysical evaluation of various modified PNA oligomers. Temperature dependent UV absorbance studies give the determination of thermal stabilities of PNA:DNA duplexes with complementary DNA. Mismatch base-pair specificity is also determined from these studies. The effect of various substitutions at γ-position on the conformation of PNA:DNA duplexes is investigated by Circular Dichroism (CD) studies. The ability of modified PNAs to displace ethidium bromide from duplex DNA is also described.

Chapter 4 Cell permeation studies of PNA oligomers

This chapter deals with cell permeation studies of the fluorescently labeled modified PNA analogs into NIH 3T3 (normal fibroblast) and breast cancer causing MCF-7 cell lines. The differential cell permeation ability of various PNA oligomers is quantitatively studied by fluorescence activated cell sorter (FACS) analysis.

1.8 References

1. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Wlater, P. *Molecular Biology of the Cell* (4th Ed.). Garland Science. **2002**, pp. 120-121.

2. Watson, J. D.; Crick, F. H. C. *Nature*, **1953**, *171*, 737-738.

3. Hoogsteen, K. *Acta. Crystal.* **1963**, *16*, 907-916.

4. (a) Crick, F. H. C. *J. Mol. Biol.* **1966**, *19*, 548-555. (b) Soll, D.; Cherayil, J. D.; Bock, R. M.; *J. Mol. Biol.* **1967**, *29*, 97-112.

5. (a) Brahms, J.; Mommaerts, W. F. H. M. *J. Mol. Biol.* **1964**, *10*, 73-85. (b) Fuller, W.; Wilkins, M. H. F.; Wilson, H. R.; Hamilton, L. *J. Mol. Biol.* **1965**, *12*, 60-80. (c) Dickerson, R. E. *Methods in Enzymol.* **1992**, *211*, 67-111. (d) Saenger, W. Principles of Nucleic Acids structure. Springer-Verlag, New York, **1984**. (e) Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Van der M. G.; Van Boom, J. H.; Rich, *Science*, **1981**, *211*, 171-176.

6. (a) Strobel, S. A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D.E.; Dervan, P.B.; *Science* **1991**, *254*, 1639-1642. (b) Huesken, D.; Deichert, A.; Hall, J.; Haner, R. *Nucleosides & Nucleotides* **1999**, *18*, 1507-1511. (c) Haner, R.; Hall, J.; *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 423-430.

7. (a) Toule, J-J. in Antisense oligonucleotides and antisense RNA: Novel Pharmacological and Therapeutic Agents. (Weiss, B., ed), CRC press, **1997**, 1-16. (b) Wagner, R. W. *Nat. Med.* **1995**, *1*, 1116-1118.

8. (a) Soyfer, V. N.; Potamann, V. N. *Triple Helical Nucleic Acids*, **1996**, Springer-Verlag, NY. (b) Chan, P. P.; Glazer, P. M. *J. Mol. Med.* **1997**, *75*, 267. (c) Fire, A.; Xu, S. Q.; Montgomery, M. K.; Kostas, S. A.; Driver, S.E.; Mello, C. C. *Nature* **1998**, *391*, 806-811.

9. Stephenson, M. L.; Zamecnik, P.C. *Proc. Natl. Acad. Sci. U.S.A* **1978**, *75*, 285-288.

10. Uhlmann, E.; Peyman, A. *Chem Rev.* **1990**, *90*, 4, 543-584.

11. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259-293.

12. Stein, C. A.; Cohen, J. S. Phoshorothioate oligodeoxynucleotide analogues. In Cohen J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 97.

13. Millar, P. S. Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): Oligodeoxynucleotides-antisense ihibitors of gene expression. London: Macmillan Press, **1989**, p. 79.

14. Froehler, B.; Ng, P.; Matteucci, M. *Nucleic Acids Res.* **1988**, *16*, 4831-4839.

15. Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G. *Nucleic Acids Res.* **1986**, *14*, 7421-7437.

16. (a) Sood, S.; Shaw, B. R.; Spielvogel, B. F.; *J. Am. Chem. Soc.* **1990**, *112*, 9000-9001. (b) Shaw, B. R.; Madison, J.; sood, S.; Spielvogel, B. F; In Agrawal, S. (ed): *Methods in Molecular biology*, vol 20: Protocols for Oligonucleotides and Analogs. Synthesis and properties. Totowa, NJ. Humana Press, Inc. **1993**, 225-243. (c) Sergueev, D. S.; Shaw, B. R. *J. Am. Chem. Soc.* **1998**, *120*, 9417-9427.

17. Uhlmann, E.; Peyman, A. *Chem Rev.* **1990**, *90*, 4, 543-584.

18. Brown, D. A.; Kang, S. H.; Gryaznov, S. M.; De Dionisio, L.; Heidenreich, O.; Sullivan, S.; Xu, X.; Neerenberg, M. I. *J. Biol. Chem.* **1994**, *43*, 26801-26805.

19. Guvakova, M. A.; Yakubov, L. A.; Vlodavsky, I.; Tonkinson, J. L.; Stein, C. A. *J. Biol. Chem.* **1995**, *270*, 2620-2627.

20. Rockwell, P., O'Connor, W.; King, K.; Goldstein, N. I.; Zhang, L. M.; Stein, C. A. *Proc. Natl Acad. Sci. USA,* **1998**, *94*, 6523-6528.

21. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259-293.

22. Altona, C.; Sundaralingam, M.; *J. Am. Chem. Soc.* **1972**, *94*, 8205-8212.

23. Teplova, M.; Minasov, G.; Tereshka, V.; Inamati, G. B.; Cook, P. D. *et al.* *Nat. Struct. Biol.* **1999**, *6*, 535-539.

24. (a) Braasch, D. A.; Corey, D. R. *Chem. Biol.* **2001**, *8*, 1-7. (b) Orum, H.; Wengel, J. *Curr. Opinion Mol. Ther.* **2001**, *3*, 239-243.

25. Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P. *Chem. Eur. J.* **2000**, *6*, 2687-2695.

26. Braasch, D. A.; Liu, Y.; Corey, D. R. *Nucleic Acids Res.* **2002**, *30*, 5160-5167.

27. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S. *et al.* *J. Med. Chem.* **2009**, *52*, 10-13.

28. Koizumi, M. *Curr. Opin. Mol. Ther.* **2006**, *8*, 144-149.

29. Sazani, P.; Gemignani, F.; Kang, S. H.; Maier, M. A.; Manoharan, M. et al. *Nature Biotechnol.* **2002**, *20*, 1228-1233.

30. Alter, J.; Lou, F.; Rabinowitz, A.; Yin, H.; Rosenfeld, J. *et al.* *Nat. Med.* **2006**, *12*, 175-177.

31. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; van Aerschot, A.; Herdewijn, P. *J. Am. Chem. Soc.* **2000**, *122*, 8595-8602.

32. Verbeure, B.; Lescrinier, E.; Wang, J.; Herdewijn, P. *Nucleic Acids Res.* **2001**, *29, 4*941-4947.

33. Gryaznov, S.; Chen, J. K. *J. Am. Chem. Soc.* **1994**, *116*, 3143-3144.

34. Faira, M.; Spiller, D. G.; Dubertret, C.; Nelson, J. S.; White, M. R. H.; Scherman, D.; Helene, C.; Giovannangeli, C. *Nat. Biotechnol.* **2001**, *19*, 40-44.

35. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science,* **1991**, *254*, 1497-1500.

36. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.

37. Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.

38. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science,* **1994**, *265*, 777-780.

39. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98-101.

40. (a) Leijon, M.; Graeslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry,* **1994**, *22*, 9820-9825. (b) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410-413.

41. Bets, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science*, **1995**, *270*, 1838-1841.

42. (a) Eriksson, M.; Nielsen, P. E.; *Q. Rev. Biophys.* **1996**, *29*, 369-394. (b) Nielsen, P. E.; Egholm, M. *Curr. Iss. Molec. Biol*. **1999**, *1*, 89-104.

43. Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Ryte, A. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2632- 2635.

44. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497-1501.

45. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497-1501. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Mol. Recogn*. **1994**, *7*, 165-170.

46. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Anti-Cancer Drug Design.* **1993**, *8*, 53 - 63. (b) J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G, Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble, L. E. Babiss, *Science*, **1992**, *258*, 1481- 1485.

47. Nielsen, P. E. *Acc. Chem. Res*. **1999**, *32*, 624-630.

48. Marsh, T. C.; Vesenka, J.; Henderson, E. *Biochemistry* **1994**, *33*, 10718-10724.

49. Forman, S. L.; Fettinger, J. C.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4060-4067.

50. (a) West, R. T.; Garza, L. A.; Winchester, W. R.; Walmsley, J. A. *Nucleic Acids Res*. **1994**, *22*, 5128-5134. (b) Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc*. **2004**, *126*, 5944-5945.

51. Demidov, V. V.; Frank-Kamenetskii, M. D.; *Trends Biochem. Sci.* **2004**, *29,* 62-71.

52. Nielsen, P. E, *Pure Appl. Chem.* **1998**, *70*, 105-110.

53. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5-23.

54. Koppelhus, U.; Nielsen, P. E. *Adv. Drug Delivery Rev*. **2003**, *55*, 267-280.

55. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 2796-2823.

56. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc. Chem. Commun*. **1993**, 519.

57. Schutz, R.; Cantin, M.; Roberts, C.; Greiner, B.; Uhlmann, E.; Leumann, C. *Angew. Chem. Int. Ed.* **2000**, *39*, 1250-1253.

58. Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 1083.

59. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res*. **2005**, *38*, 404-412.

60. D'Costa, M.; Kumar, V. A.; Ganesh, K. N.; *Org. Lett.* **1999**, *1*, 1513-1516.

61. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1996**, *52*, 15017-15030.

62. Gangamani, B. P.; D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides* **1999**, *18*, 1409-1011.

63. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, *55*, 177-192.

64. Lagriffoule, P.; Wittung, P.; Eriksson, M.; Jensen, K. K.; Norden, B.; Buchardt, O.; Nielsen, P. E. *Chem. Eur. J.* **1997**, *3*, 912-919.

65. Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *Chem. Commun*. **2004**, 860-861.

66. Dueholm, K. L.; Petetrsen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1077-1080.

67. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed*. **1996**, *35*, 1939-1942.

68. (a) Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998**, *39*, 4707-4710. (b) Gourishankar, A.; Ganesh, K. N. *Artificial DNA: PNA & XNA* **2012**, *3*, 5-13.

69. Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. Eur. *J. Org. Chem*. **1999**, 197-204.

70. Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. Eur. *J. Org. Chem.* **2007**, 5879-5885.

71. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6878-6879.

72. Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; Okeefe, H.; Vatta, P.; Hyldig-Nielsen, J. J.; Ly, D. H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4931-4935.

73. Dragulescu-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. commun.* **2005**, *41*, 244-246.

74. Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas, S. M.; Zon, G.; Livak, K. J.; Ly, D. H. *J. Org. Chem.* **2009**, *74*, 1509-1516.

75. Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Bianchi, N.; Brognara, E.; Gambari, R.; Marcgelli, R.; Corradini, R. *ChemBioChem* 2012, 13, 1327-1337.

76. Mitra, R.; Ganesh, K. N. *Chem. commun.* **2011**, *47*, 1198-1200.

77. Mitra, R.; Ganesh, K. N. *J. Org. Chem.* **2012**, *77*, 5696-5704.

78. Haaima, G.; Hansen, H. F.; Christensen, L.; Dahl, O.; Nielsen, P. E. *Nucleic Acids Res.* **1997**, *25*, 4639-4643.

79. Egholm, M.; Christensen, L.; Deuholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, *23*, 217-222.

80. (a) Gangamani, B. P.; Kumar V. A. *JCS Chem. Commun.* **1997**, 1913-1914. (b) Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Biochem. Biophys. Res. Commun.* **1997**, *240,* 778-782.

81. Eldrup, A. B.; Dahl, O.; Nielsen, P. E. *J. Am. Chem. Soc.* **1997**, *119*, 11116-11117.

82. (a) Wojciechowski. F.; Hudson. R. H. E. *Cur. Top. Med. Chem.* 2007, *7, 667-679.* (b) Bajor, Z.; Sagi, G.; Tegyey, Z.; Kraicsovits, F. *Nucleosides & Nucleotides*, **2003**, *22*, 1963-1983. (c) Hudson, R. H. E.; Li, G.; Tse, J. *Tetrahedron Lett*. **2002**, *43*, 1381-1386.

83. Vysabhattar, R.; Ganesh, K. N. *Tetrahedron Lett*. **2008**, *49*, 1314-1318.

84. Ray, A.; Norden, B. *FASEB J.* **2000**, *14*, 1041-1060.

85. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Gene* **1994**, *149*, 139-145.

86. (a) Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.* **1996** 24, 494-500. (b) Ivanova, G. B.; Arzumanov, A.; Abes, R.; Yin, H.; Wood, M. J. A.; Lebleu, B.; Gait, M. J.

Nucleic Acids Res. **2008**, *36*, 6418-6428. (c) Fabani, M. M.; Goodger, C. A.; Williams, D.; Lyons, P. A.; Torres, A. G.; Smith, K. G. C.; Enright, A. J.; Gait, M. J.; Vigorito, E. *Nucleic Acids Res*. **2010**, *38*, 4466-4475.

87. Taylor, R. W.; Chinnery, P. F.; Turnbull, D. M.; Lightowlers, R. N. *Nature Genet.* **1997**, *15*, 212-215.

88. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 2796-2823.

89. Hanvey, J. C.; Peffer, N. C.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481-1485.

90. Nielsen, P. E.; Egholm, M.; Berg, R. H; Buchardt, O. *Anti-Cancer Drug Design,* **1993**, *8,* 53-63.

91. Mologni, L.; leCoutre, P.; Nielsen, P. E.; Gambacorti-Passerini, C. *Nucleic Acids Res.* **1998**, 26, 1934-1938.

92. Demers, D.B.; Curry, E.T.; Egholm, M.; Sozer, A. C. *Nucleic Acids Res.* **1995**, *23*, 3050-3055.

93. Demidov, V.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Nucleic Acids Res.* **1993**, *21*, 2103-2107.

94. Orum, H.; Nielsen, P.E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res.* **1993**, *21*, 5332-5336.

95. Kuhn, H.; Demidov, V. V.; Gildea, B. D.; Fiandaca, M. J.; Coull, J. M.; Frank Kamenetskii, M. D.; *Antisense Nucleic Acid Drug Dev.* **2001**, *11*,265-270.

Chapter **2**

Design, Syntheses and Characterization of γ-C-Substituted PNA Monomers and their Oligomerization

PNA modifications by introducing various functionalities at γ-position in the backbone have been described in this chapter. Introduction of chirality and cationic moiety into the achiral PNA backbone should influence the orientation selectivity in complementary DNA binding. Modifications have been designed to improve hybridization efficiency, poor aqueous solubility and low cellular uptake of the achiral *aeg* PNA through various γ-C-substituted PNA analogs.

2.1 Introduction

Elucidation of the double helical structure of DNA, a genetic material and its remarkable properties has fascinated chemists. Tremendous efforts have been devoted to understanding its structure and function in biology as well as from a chemical and physicochemical standpoint.[1] Investigations of oligonucleotides as potential therapeutics that target nucleic acids has led the search for nucleic acid mimetics with improved properties. However, some problems remain to be solved towards the medicinal applications of such functional nucleic acids and their derivatives. These include degradation by cellular nucleases, impermeability through cell membrane, low hybridization efficiency etc.
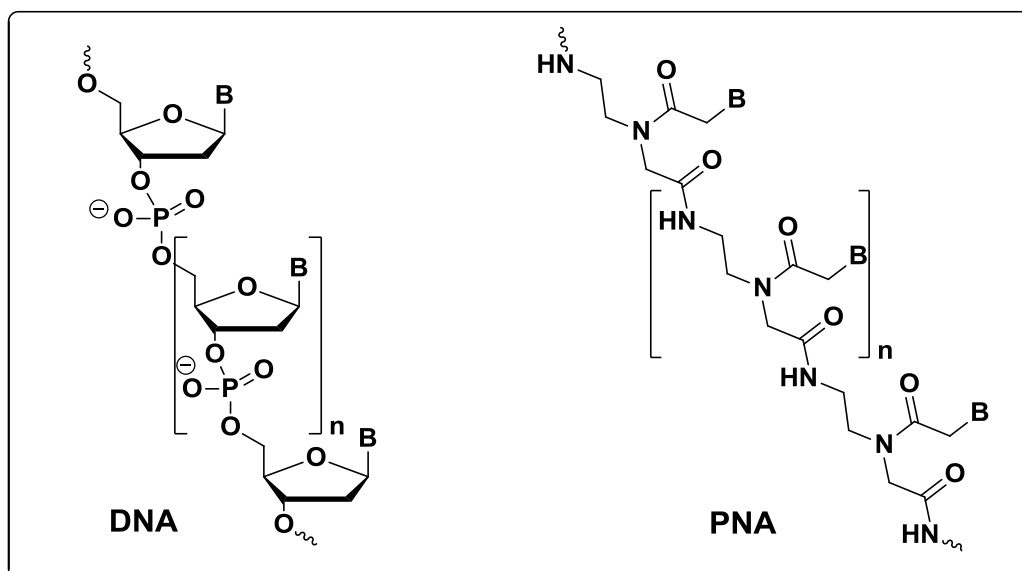
FIGURE 2.1 Structural comparison of DNA and PNA (B = nucleobase)

For two decades, peptide nucleic acids (PNA)[2] have emerged as one of the most promising oligonucleotides for the recognition of nucleic acids (DNA/RNA). The sugar-phosphate backbone from DNA was completely replaced by repeating units of (2-aminoethyl)glycine units and the nucleobases are attached through a methylene carbonyl linker to the backbone (Figure **2.1**). The optimal number of bonds between the nucleobases was found to be six which corresponds to that found in DNA thereby providing the correct inter-nucleobase spacing.[3] Also the unmodified PNAs are uncharged at neutral pH.

The original idea behind designing PNA was to recognize and bind to double stranded DNA in the major groove via Hoogsteen hydrogen bonding to form triple helix.[2] Later it was shown that PNA oligomers can bind to single stranded nucleic acids via Watson-Crick hydrogen bonding to form duplexes with higher affinity.[4] The stability of the PNA interaction with DNA is such that strand invasion of DNA by PNA is thermodynamically favored, and can take place via duplex, triplex or double duplex formation. Besides their unique binding properties, PNAs are both chemically and biologically stable.[5] Peptide nucleic acids are resistant to cellular enzymes like nucleases and proteases thereby having impressive *in vivo* stability.

Though PNA is a promising molecule in terms of recognition, binding specificity, affinity and resistance to cellular enzymes the PNA oligomers have poor water solubility and extremely low cellular uptake.[6] These are the major obstacles for PNA applications in biomedicine. Also ambiguity in the orientation selectivity between parallel and antiparallel binding to complementary oligonucleotide hampers their applications. To overcome these limitations various modifications were attempted which have improved PNA properties to different extents.

Functionalization of the PNA backbone may dramatically change the physicochemical properties of PNA and these changes would also significantly influence the bio-distribution and affect pharmacokinetic profile. A lot of efforts have been made to synthesize conformationally constrained cyclic PNA analogs using the concept of structural preorganization[7] to address interesting attributes of PNA:DNA/RNA hybridization. Further, modifications in the backbone to make chiral, acyclic PNA analogs having the stereocentre at α or γ carbon[8] have led to improvements in PNA properties (Figure **2.2**). Recently, Sugiyama *et al.*, have introduced the methyl substitution at the least studied β carbon position.[9] The functionalization at these positions of the monomer may also preorganize the PNA strand but importantly it has the effect of shifting the PNA preference towards a right handed or left handed conformation, depending on the configuration of the new stereogenic centres. This, in turn, affects the stability of the PNA:DNA duplex through the control of helix handedness.[10]
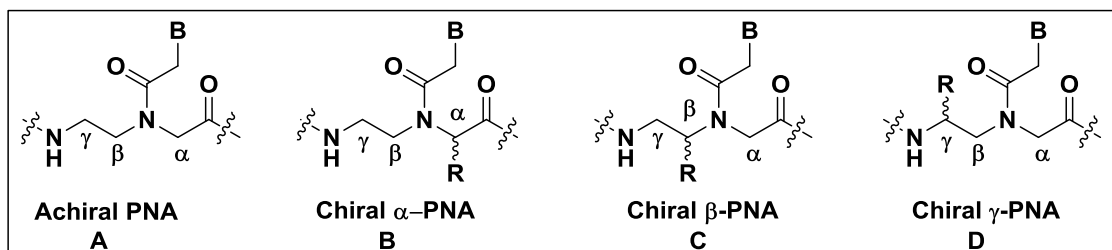
FIGURE 2.2 Structures of achiral PNA and chiral α, β- and γ-modified PNA (R = amino acid side chain or other modifications, B = nucleobase)

In order to be a good DNA binding molecule, a modified backbone should neither too flexible nor too rigid. A highly flexible backbone would require a very high loss of entropy for efficient binding whereas a very rigid backbone would prevent DNA binding due to a difficult fit to adopt to target structure.[11] The ability of unmodified aminoethyl glycyl (*aeg*) PNA to strongly bind DNA has been interpreted as a result of the 'constrained flexibility' in its structure.[12] The earliest and simplest modification involved extension of PNA backbone with a methylene group individually in each of the structural subunits (aminoethyl[13] and glycine[14]) of the PNA monomer. However, such modifications resulted in a significant decrease in melting temperatures of the derived PNA:DNA duplexes. These studies suggested that the inter-residue distance between nucleobases should not be disturbed for efficient binding with complementary oligonucleotides.

Several PNA derivatives were obtained by insertion of side chains either at α or γ-carbon atoms. These included hydroxymethyl (serine), aminobutyl (lysine), guanidinium (arginine) and many functional groups from amino acid side chains. With these acyclic molecules, there was no real preorganization, but constrained flexibility due to the presence of a chiral centre. PNA oligomers incorporating chiral monomers retained the hybridization properties though less efficiently, with tolerance for small and medium sized substituents at the glycine α-position. The side chains of the amino acids are valuable for controlling the binding affinity, specificity, hydrophobicity and attachment of ligands to PNA. The next section deals with the incorporation of cationic (amine/guanidine) and neutral (azide) substitutions at γ-position of PNA monomers.

Chapter 2A

Synthesis and Characterization of γ-(*S-eam*), γ-(*S-egd*), γ-(*S-azm*) and γ-(*S-azb*) *aeg* PNA Monomers

γ-(*S-eam*) *aeg* PNA

γ-(*S-egd*) *aeg* PNA

γ-(*S-azm*) *aeg* PNA

γ-(*S-azb*) *aeg* PNA
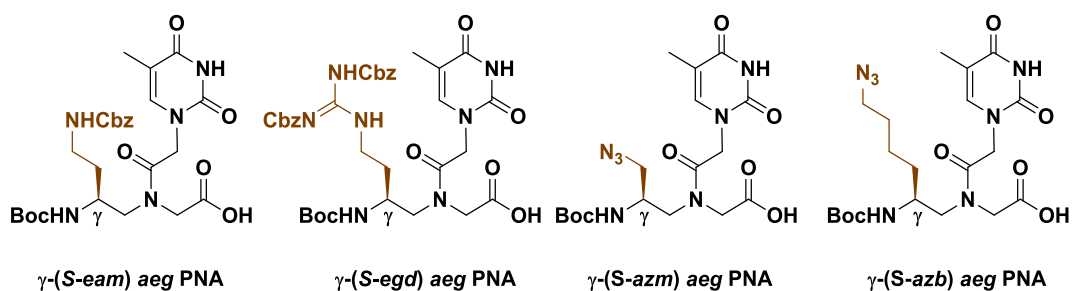
2A.1 Rationale behind the work

The chemical modifications of the *aeg* PNA affords considerable advantages in the development of a pharmacologically efficacious antigene agent.[15] Many strategies have focused on the covalent integration of chiral cyclic ring systems in designs that gives constrained flexibility to the PNA and maintains the base separation from and along the backbone by the same number of bonds as in DNA.[16]

Although there are so many possible modifications in the PNA backbone, one challenging area of research has been the strategic placement of side chains on the *aeg* PNA scaffold. Chiral amino acids have been used extensively to improve the PNA properties. It has been shown that incorporation of D-Lysine carrying positively charged functional group at α-position has a stabilizing effect on the PNA:DNA duplex. Also it has significantly improved the solubility of PNA oligomer in water[17] (Figure **2.3 A**). The presence of such positively charged side chains can help in improving the cellular uptake. However, incorporation of amino acids with negatively charged side chains have shown deleterious effect on binding perhaps due to electrostatic repulsion.
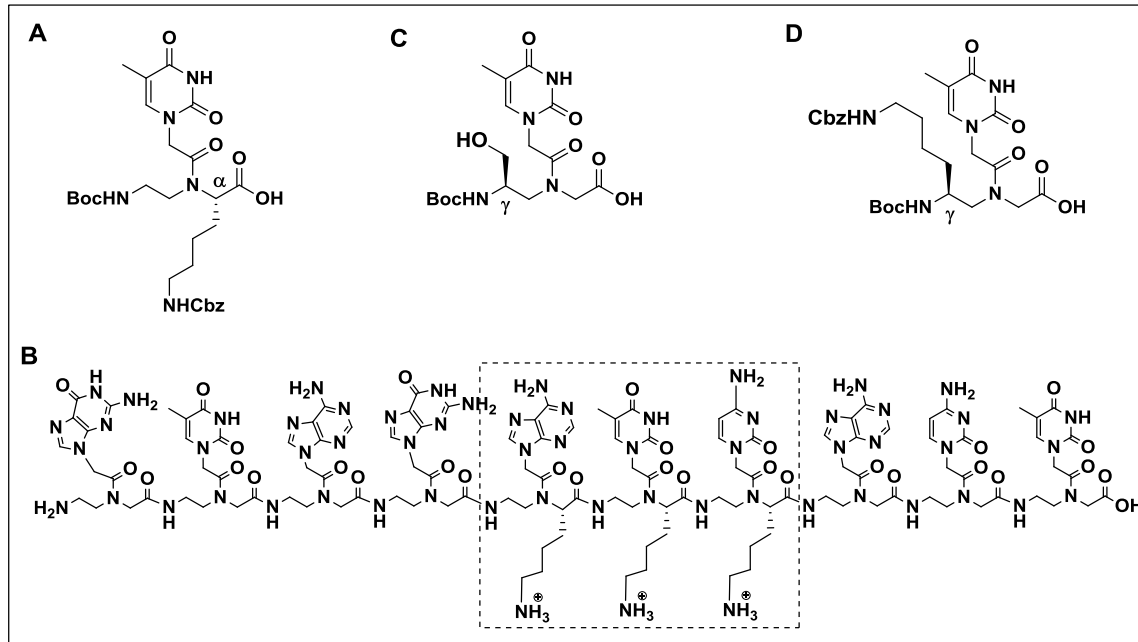
FIGURE 2.3 (A) α-D-Lysine PNA monomer (B) Decamer PNA incorporating three adjacent chiral monomers (chiral box) (C) & (D) γ-L-Serine & γ-L-Lysine PNA monomer

A recent crystallographic study of anti-parallel mixed-sequence PNA-DNA decamer heteroduplex, comprising a three-residue unit D-lysine ''chiral box'' (Figure **2.3 B**), provides evidence that increased rigidity of polyamide backbone conformation

can be obtained through the introduction of chiral centres.[18] Reports have also indicated that L-lysine[19,20,21] and L-serine[22] derived γ-substituted PNA units within unmodified PNA oligomers act as helical director due to conformational preorganization (Figure **2.3 C** and **D**). It has also been reported that cationic guanidium groups grafted on PNAs (GPNA)[23,24,25] and phosphate backbone of oligonucleotides[26] enhance the cell uptake of the substrates (Figure **2.4 A**). Recently it has been shown that PNAs grafted with (α/γ, *R/S*)-aminomethylene pendants show regio and stereospecific effects on DNA binding and improve the cell permeation[27,28] (Figure **2.4 B, C** and **D**).
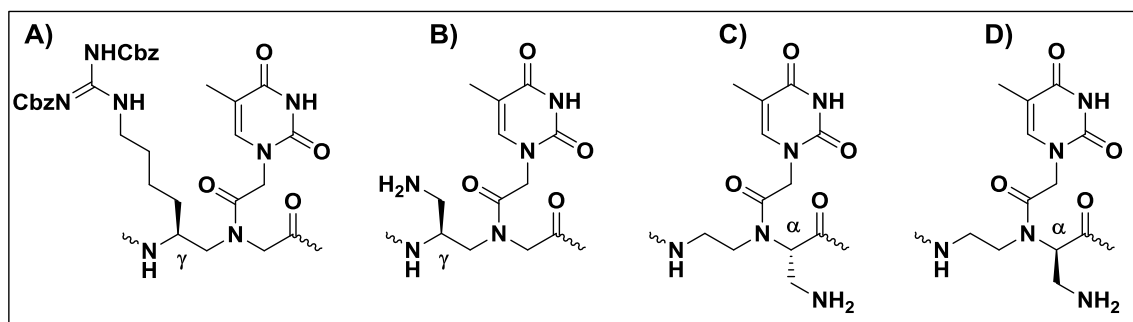
FIGURE 2.4 (A) GPNA (B) PNAs grafted with (γ-*S*)-aminomethylene pendants (C) and (D) PNAs grafted with (α-*R/S*)-aminomethylene pendants

Based on this survey, it was aimed to synthesize PNA monomers having shorter aliphatic side chains compared to lysine based PNA. These modifications may be less prone for nonspecific secondary interactions with nucleotide phosphate groups in other DNA molecules. Our objective is to incorporate various functional groups like amine, guanidine and azide at γ-position to give various modified chiral PNAs having cationic and neutral side chains (Figure **2.5**). These modified PNA monomers have been incorporated in PNA oligomers to improve its hybridization properties as well as to improve cell permeation abilities of PNA. The rationale behind the design of each monomer is discussed further.

2A.1.1 γ-(*S*-ethyleneamino) aminoethylglycyl PNA [γ-(*S-eam*) *aeg* PNA] and γ-(*S*-ethyleneguanidino) aminoethylglycyl PNA [γ-(*S-egd*) *aeg* PNA]

As far as the configuration of the chiral monomer is concerned, PNAs having monomers derived from L-amino acids (*S*-stereochemistry) bind to the complementary DNA with greater stability than those derived from the corresponding D-amino acids

(*R*-stereochemistry) at γ-position. The γ- GPNA oligomers have been shown to be preorganized into a right handed helix and bind to complementary DNA/RNA with high affinity.[29,28]
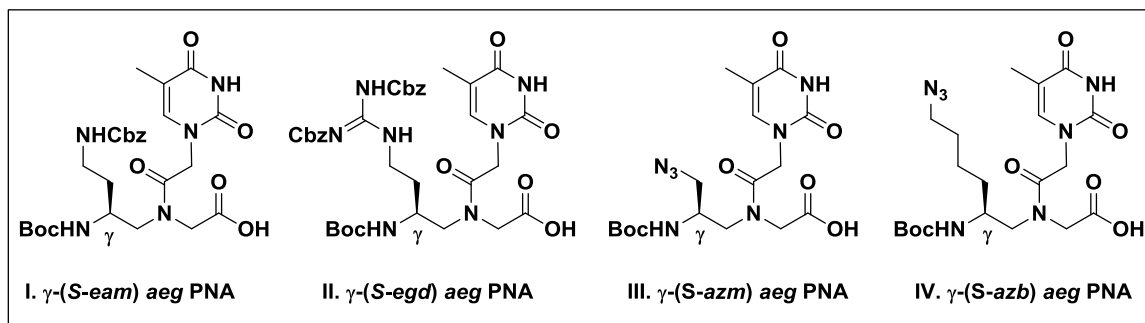
FIGURE 2.5 Target PNA monomers modified at γ-position

These results inspired the design and synthesis of a γ-C-substituted PNA incorporating ethyleneamino and ethyleneguanidino (spacer is 2xCH$_2$ groups) groups in the *aeg* PNA backbone in the form of γ-(*S*-ethyleneamino) aminoethylglycyl PNA [γ-(*S*-eam) *aeg* PNA] (Figure **2.5**, No. **I**) and γ-(*S*-ethyleneguanidino) aminoethylglycyl PNA [γ-(*S*-egd) *aeg* PNA] (Figure **2.5**, No. **II**) in order to investigate its potential as an effective conjugation point (Figure **2.6**). Thus, the incorporation of cationic groups like amine and guanidine at γ-position will improve the hybridization properties as well as cellular uptake of PNA oligomers. Also the detailed studies of these modified PNAs would provide insight into the structure activity relationship (SAR) of PNA modifications at γ-position.
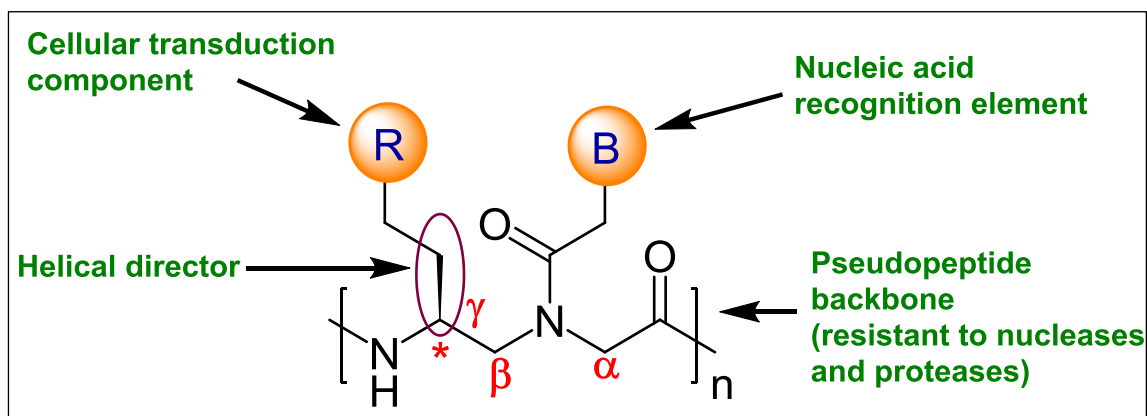
FIGURE 2.6 Rationale behind incorporation of cationic (amine/guanidine) functional groups in PNA

2A.1.2 γ-(*S*-azidobutylene) aminoethylglycyl PNA [γ-(*S-azb*) *aeg* PNA] and γ-(*S*-azidomethylene) aminoethylglycyl PNA [γ-(*S-azm*) *aeg* PNA]

The azide functional group has been incorporated in the PNA either at 5C(U) nucleobase[30] to attach various functional groups or the amide bond has been replaced by triazole moiety via azide and alkyne incorporation in the backbone.[31] Recently the azide functional group has been incorporated in place of nucleobase to attach various fluorophores via solid phase CuAAC strategy.[32] The incorporation of azide in the backbone will give a chiral PNA with a neutral functional group which can be directly correlated with an unmodified *aeg* PNA that is also neutral. None of the reports discuss the incorporation of azide functional group in the side chain at α/γ-position. Here, the objective is to incorporate fluorophore via click reaction between azide modified PNA and fluorescent alkyne in the side chain at γ-position.

The above aims inspired the design and synthesis of a γ-C-substituted PNA incorporating flurophore (via copper mediated azide-alkyne cycloaddition) in the side chain at various distances (spacer 4XCH$_2$ and 1XCH$_2$) from the backbone in the *aeg* PNA in the form of γ-(*S*-azidobutylene) aminoethylglycyl PNA [γ-(*S-azb*) *aeg* PNA] and γ-(*S*-azidomethylene) aminoethylglycyl PNA [γ-(*S-azm*) *aeg* PNA] (Figure **2.5**, No. **III** & **IV**).

The utility of designed PNA monomers has been given below in Figure **2.7**.
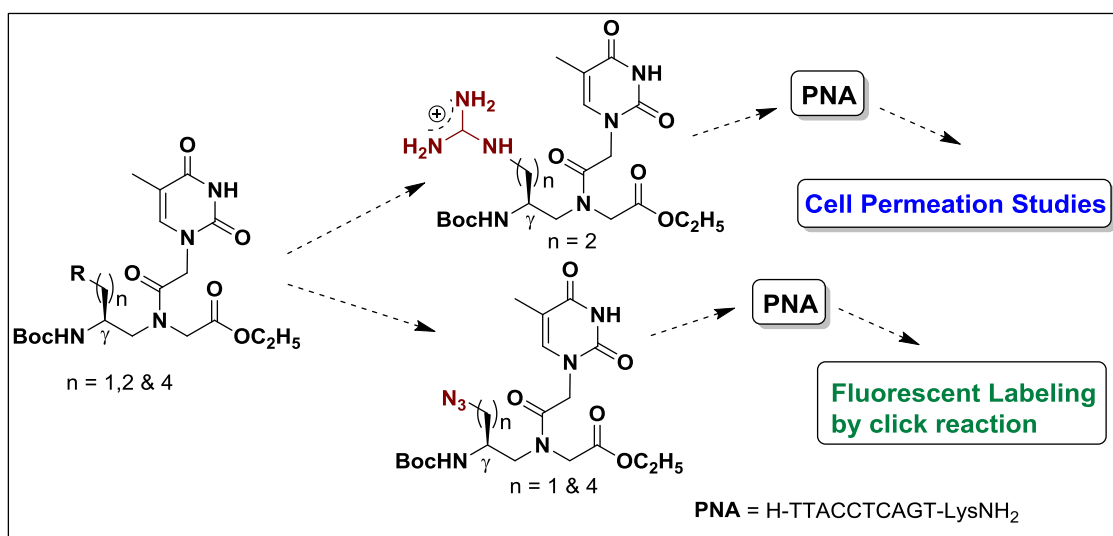
FIGURE 2.7 Utility of modified PNA monomer

2A.2 Aim of the present work

The specific objectives of this section are

- ➢ Synthesis of functionally modified, γ-C-substituted PNA T-monomers
- ➢ Synthesis of alkyne derivative of 5(6)-carboxyfluorescein for click reaction with azide modified PNAs
- ➢ Characterization of rationally synthesized PNA monomers and their intermediates by various spectroscopic techniques

2A.3 Synthesis of modified PNA monomers

This section describes the synthesis of rationally designed γ-C-substituted cationic (amine/guanidine) and charge neutral azide modified PNA monomers.

2A.3.1 Synthesis of γ-(*S-eam*) *aeg* PNA monomer

In an attempt towards the synthesis of target monomer, the commercially available L-Glutamine **1** was treated with di-tert-butyl dicarbonate [(Boc)$_2$O] in aq. NaOH-dioxane to obtain Boc-protected L-Glutamine **2**. This was followed by the reaction of N-α-Boc-L-Glutamine **2** in EtOAc:ACN:H$_2$O with iodobenzene diacetate (PIDA)[33] under cooling which led to the formation of 4-amino-2-(N-Boc-amino)butanoic acid **3** as the product in 4 h. Compound **3** was treated with benzylchloroformate in toluene in the presence of NaHCO$_3$ to obtain orthogonally protected 4-(Cbz-amino)-2-(Boc-amino) butanoic acid **4** in quantitative yield. Compound **4** was converted to methyl ester derivative **5** using dimethyl sulfate and activated K$_2$CO$_3$ and the conversion was confirmed by the appearance of ^{1}H NMR peak at δ 3.73 ppm for –CH$_3$ group of methyl ester in the product. The methyl ester was reduced to give alcohol derivative **6** using sodium borohydride in absolute ethanol. Mesylation of the primary hydroxyl group by controlled addition of mesyl chloride in DCM using triethyl amine gave the mesylate derivative **7** which was immediately treated with sodium azide in dry DMF to obtain the azido compound **8**. The presence of azide group was confirmed by IR spectrum showing peak for azide at 2098 cm^{-1}. The reduction of azide derivative using Raney Ni under hydrogenation conditions yielded the free amine which was then alkylated with ethyl bromoacetate in acetonitrile using triethyl amine to obtain the alkylated compound **9**. This was then acylated with

chloroacetyl chloride using triethyl amine to yield the chloro compound **10**. The condensation of the compound **10** with nucleobase thymine afforded the γ-(*S*-ethyleneamino) aminoethylglycyl ethyl ester **11** in quantitative amount. The appearance of peaks at δ 7.05 ppm and 1.87 ppm in 1H NMR shows the presence of thymine in the desired product **11**. The γ-(*S-eam*) *aeg* PNA ester was hydrolysed using aq. lithium hydroxide solution to get γ-(*S-eam*) *aeg* PNA acid **12** (Scheme **2.1**) which was then used for incorporation by solid phase peptide synthesis to obtain the desired PNA sequences.

All intermediates were purified by column chromatography and characterized by ^{1}H, ^{13}C NMR and mass spectral analysis (Appendix I).

SCHEME 2.1 Schematic representation of synthesis of γ-(*S-eam*) *aeg* PNA monomer
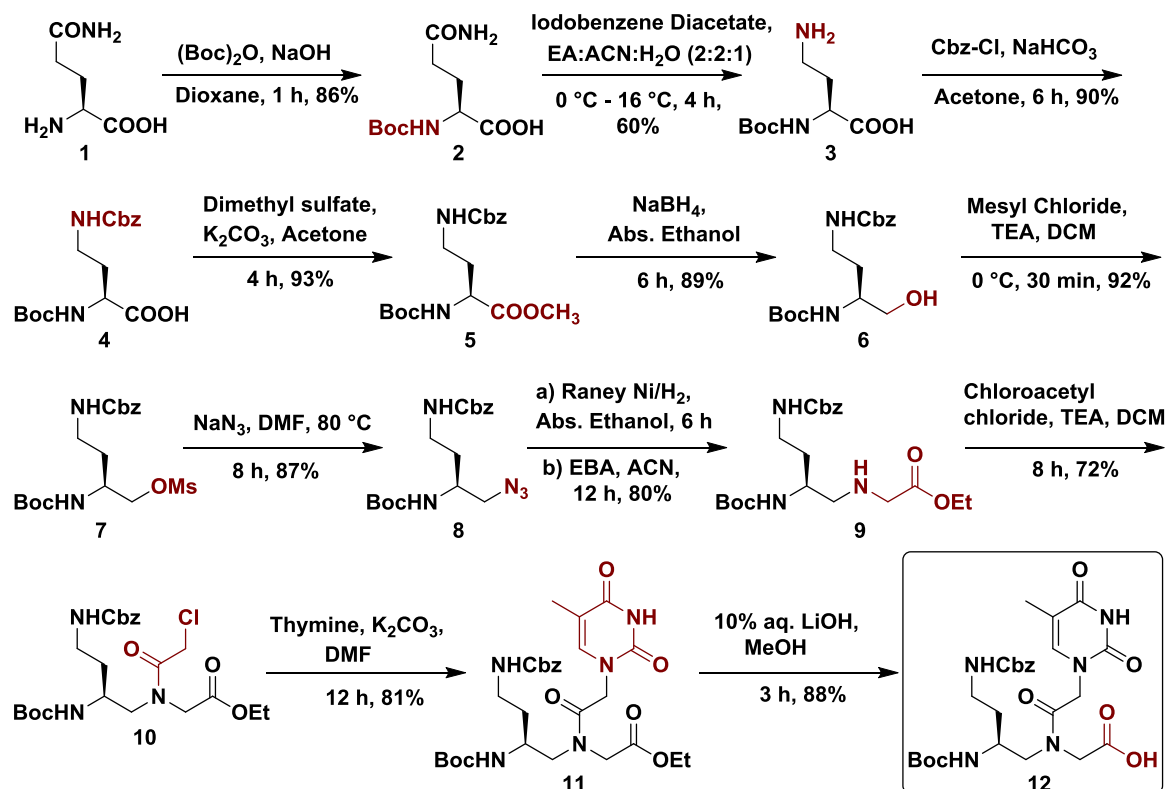
2A.3.2 Synthesis of γ-(*S-egd*) *aeg* PNA monomer

As shown in scheme **2.2**, Cbz-group in compound **11** was deprotected under hydrogenation conditions using Pd/C and glacial acetic acid to give free amine which was guanidinylated using N,N'-bis-di-Cbz-S-Me-isothiourea in dry DMF and catalytic amount of DMAP to give compound **13** as γ-(*S-egd) aeg* PNA ester. The

guanidinylation was confirmed by the appearance of peaks at δ 7.39-7.28 ppm and 5.22-5.09 ppm in ^{1}H NMR. This was then hydrolysed using aq. NaOH (2 N) in THF (1:1) to get *γ-(S-egd) aeg* PNA acid **14** which was used for incorporation by solid phase peptide synthesis to obtain the desired PNA sequences.

SCHEME 2.2 Schematic representation of synthesis of γ-(*S-egd*) *aeg* PNA monomer
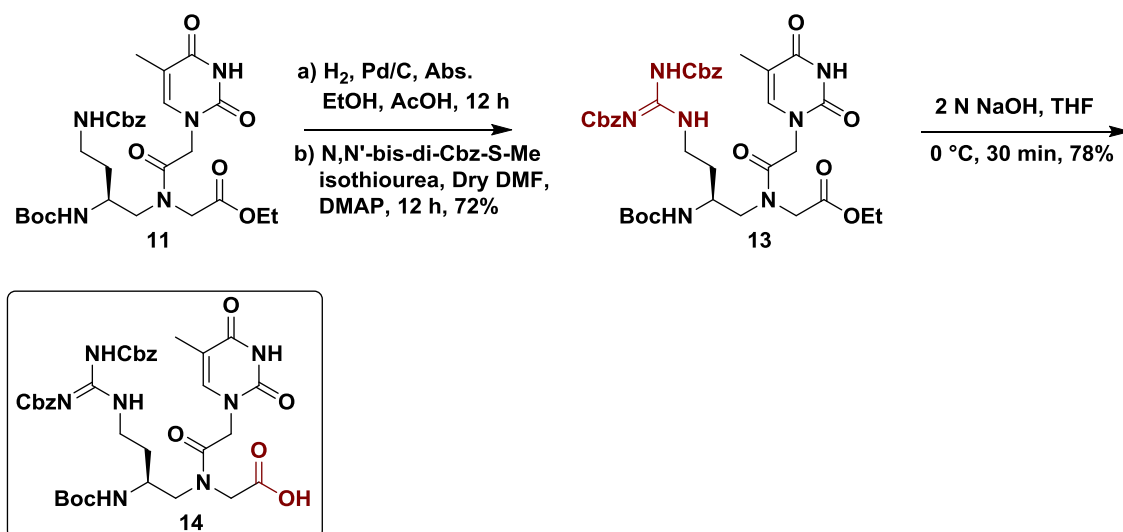
2A.3.3 Synthesis of γ-(*S-azm*) *aeg* PNA monomer

As shown in scheme **2.3**, commercially available L-Asparagine **15** was treated with di-tert-butyl dicarbonate, [(Boc)₂O] in aq. NaOH-dioxane to obtain Boc-protected L-Asparagine **16**. This was followed by the reaction of N-α-Boc-L-Asparagine **16** in ethyl EtOAc:ACN:H₂O with iodobenzene diacetate (PIDA)[34] under cooling which led to the formation of 3-amino-2-(Boc-amino)propanoic acid **17** as the product in 4 h. This was followed by the diazo transfer reaction for the conversion of primary amine to azide derivative. As reported in literature,[34] triflic anhydride was reacted with sodium azide in cooling conditions to obtain triflyl azide which was used for the diazo transfer reaction. The primary amine was treated with triflyl azide using CuSO₄.5H₂O and triethyl amine to give azido compound **18**. The presence of azide group was confirmed by IR spectrum showing peak for azide at 2109 cm^{-1}. This compound was treated with dimethyl sulfate in acetone using activated K₂CO₃ to give ester derivative **19** which was confirmed by the appearance of peak at δ 3.79 ppm in ^{1}H NMR for –CH₃ group of methyl ester. The methyl ester was reduced using sodium borohydride in absolute ethanol to give alcohol derivative **20**. The mesylation of primary hydroxyl group by

controlled addition of mesyl chloride in DCM using triethyl amine gave the mesylate derivative **21**. This mesyl derivative was immediately reacted with potassium phthalimide in dry DMF to yield phthalimido protected compound **22**. Deprotection of phthalimide with hydrazine hydrate in absolute ethanol gave free amine which was then alkylated with ethyl bromoacetate in acetonitrile using triethyl amine to yield alkylated product **23**.

SCHEME 2.3 Schematic representation of synthesis of γ-(*S-azm*) *aeg* PNA monomer
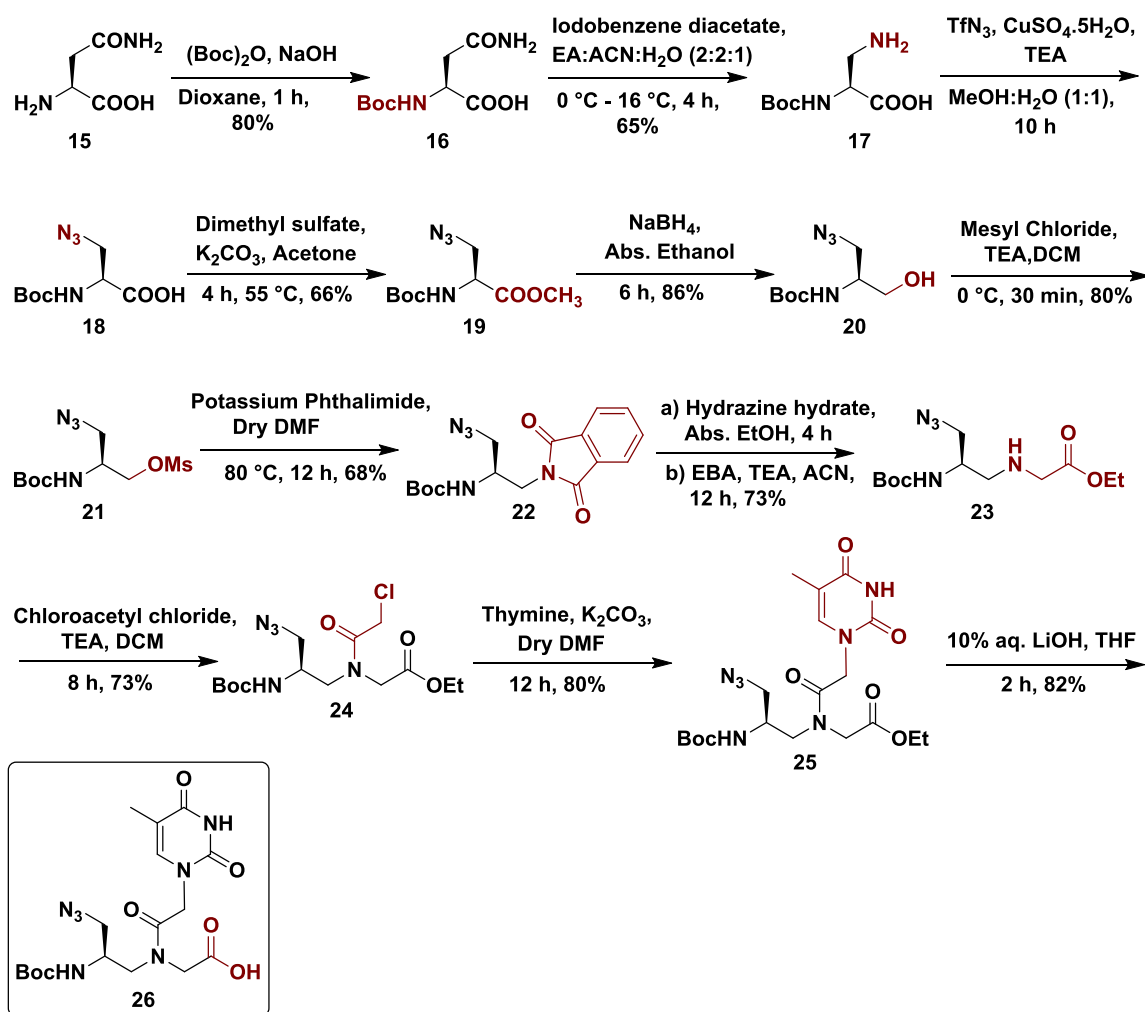
The compound **23** was acylated using chloroacetyl chloride in presence of triethyl amine to get chloro derivative **24**. The chloro compound was reacted with nucleobase thymine in presence of activated K_2CO_3 to give γ-(*S*-azidomethylene) aminoethylglycyl ethyl ester **25** in quantitative amount. The presence of thymine was confirmed by the appearance of peaks at δ 7.30 ppm and 1.88 ppm in ^{1}H NMR. The γ-(*S-azm*) *aeg* PNA ester was hydrolysed using aq. lithium hydroxide to get γ-(*S-azm*)

aeg PNA acid **26**. This was then used for incorporation by solid phase peptide synthesis to obtain the desired PNA sequences. All the intermediates were purified by column chromatography and characterized by ^{1}H, ^{13}C NMR and mass spectral analysis (Appendix I).

2A.3.4 Synthesis of γ-(*S-azb*) *aeg* PNA monomer

A similar strategy was used for the synthesis of aminoethyl γ-(*S*-azidobutylene) glycine PNA monomer **36** using α-Boc protected L-Lysine as starting material. As explained above, the primary amine was converted to azide derivative using diazo transfer reaction (Scheme **2.4**).

SCHEME 2.4 Schematic representation of synthesis of γ-(*S-azb*) *aeg* PNA monomer
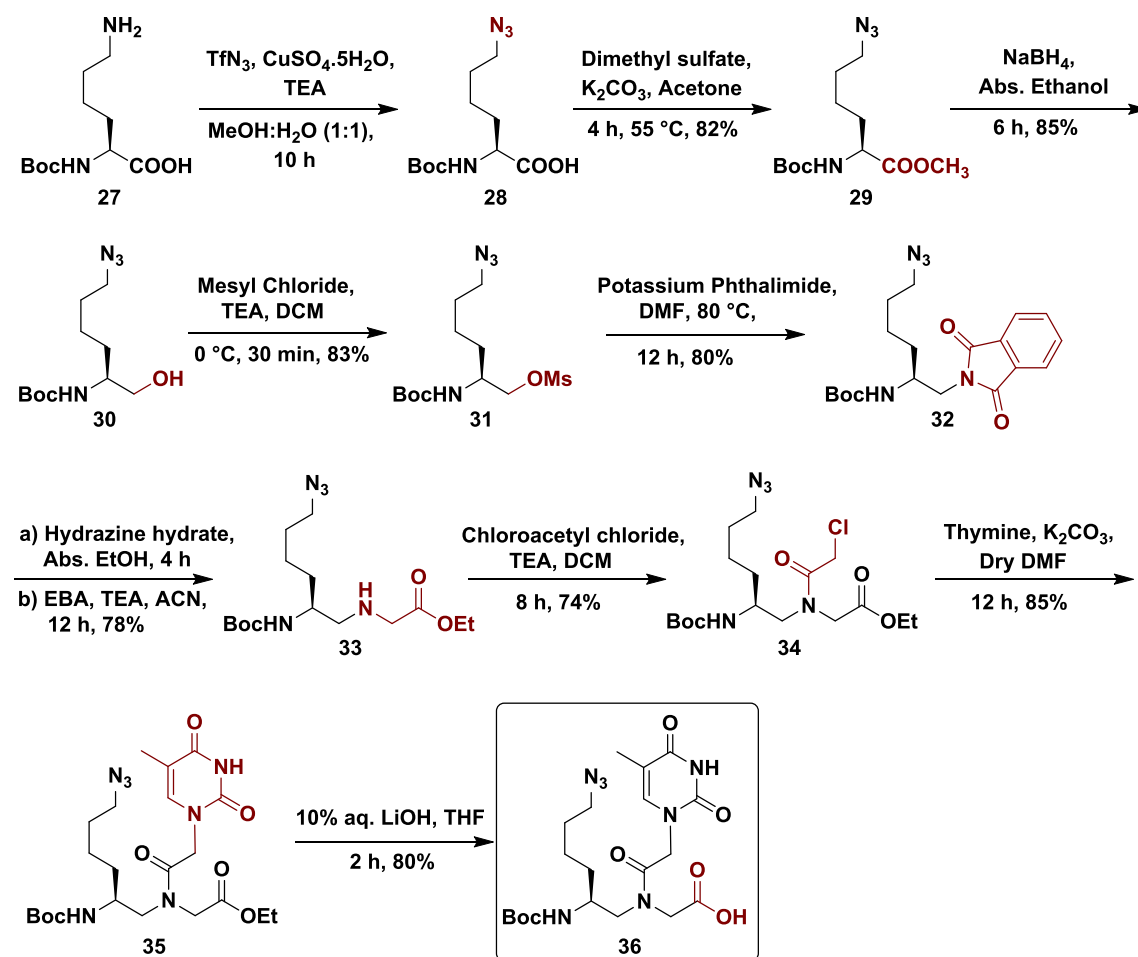
The primary amine of compound **27** was treated with triflyl azide using CuSO$_4$.5H$_2$O and triethyl amine to give azido compound **28**. The appearance of peak at 2106 cm^{-1} in IR spectrum confirms the formation of azide in the product. This

compound was treated with dimethyl sulfate in acetone using activated K_2CO_3 to give the methyl ester **29** which was confirmed by the appearance of peak at δ 3.72 ppm for −CH_3 group of methyl ester in ^1H NMR. The methyl ester was reduced using sodium borohydride in absolute ethanol to give primary alcohol **30**. The mesylation of primary hydroxyl group using mesyl chloride and triethyl amine in DCM gave the mesylate derivative **31**. This mesyl derivative was immediately reacted with potassium phthalimide in dry DMF to yield phthalimido protected compound **32**. Deprotection of phthalimide with hydrazine hydrate in absolute ethanol gave free amine which was then alkylated with ethyl bromoacetate in acetonitrile using triethyl amine to yield the alkylated product **33**. This compound was acylated using chloroacetyl chloride in presence of triethyl amine to get chloro derivative **34**. This was reacted with nucleobase thymine in presence of activated K_2CO_3 to give γ-(*S*-azidobutylene) aminoethylglycyl ethyl ester **35** in good yield. The appearance of peaks at δ 7.09 ppm and 1.91 ppm in ^1H NMR show the presence of thymine in desired product **35**. The *γ-(S-azb) aeg* PNA ester was hydrolysed using aq. lithium hydroxide to get *γ-(S-azb) aeg* PNA acid **36** which was then used for incorporation by solid phase peptide synthesis to obtain the desired PNA sequences. All intermediates were purified by column chromatography and characterized by ^1H, ^{13}C NMR and mass spectral analysis (Appendix I).

2A.3.5 Synthesis of alkyne derivative of 5(6)-carboxyfluorescein

As shown in scheme **2.5**, commercially available N-hydroxysuccinimide ester of 5(6)-carboxyfluorescein **37** was reacted with propargyl amine to give fluorescent alkyne derivative of 5(6)-carboxyfluorescein **38** which was used for click reaction. The presence of terminal alkyne group in **38** was confirmed by IR spectrum showing peaks at 2123 cm^{-1} and 3289 cm^{-1}. The fluorescence spectrum recorded for alkyne derivative of 5(6)-carboxyfluorescein is shown in Figure **2.8**.

SCHEME 2.5 Synthesis of alkyne derivative of 5(6)-carboxyfluorescein
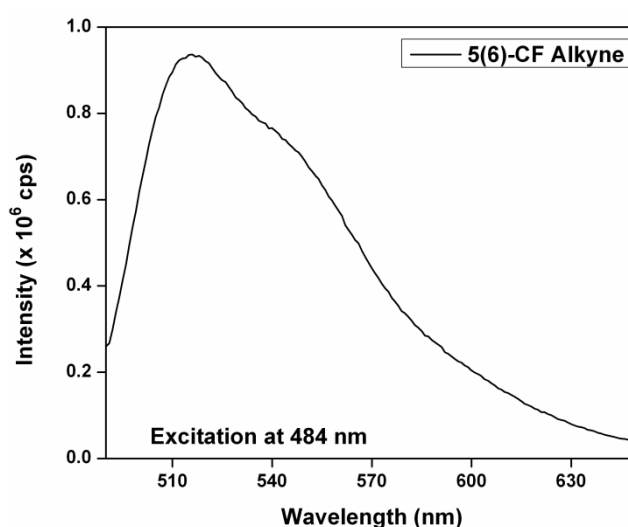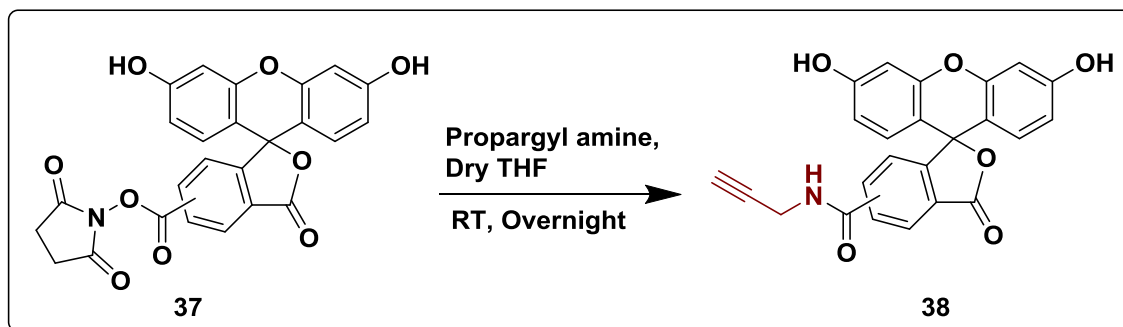
FIGURE 2.8 Fluorescence spectrum for alkyne derivative of 5(6)-carboxyfluorescein

2A.4 Summary

To summarize, this section describes the synthesis and characterization of rationally designed PNA monomers by incorporating amino, guanidino and azido functionalities at γ-position. All the intermediates have been characterized by ^{1}H & ^{13}C NMR spectroscopy, mass spectral analysis and other appropriate analytical data (Section **2.2**). The fluorescent alkyne derivative of 5(6)-carboxyfluorescein was also synthesized and characterized to prepare fluorescent PNAs by click reaction. The next section deals with the incorporation of these modified monomer units into oligomers at various desired positions using solid phase peptide synthesis.

Chapter 2B

Solid Phase Synthesis, Purification and Characterization of PNA Oligomers

PNA Monomer

Solid Phase
PNA Synthesis

PNA = H-TTACCTCAGT-LysNH$_2$
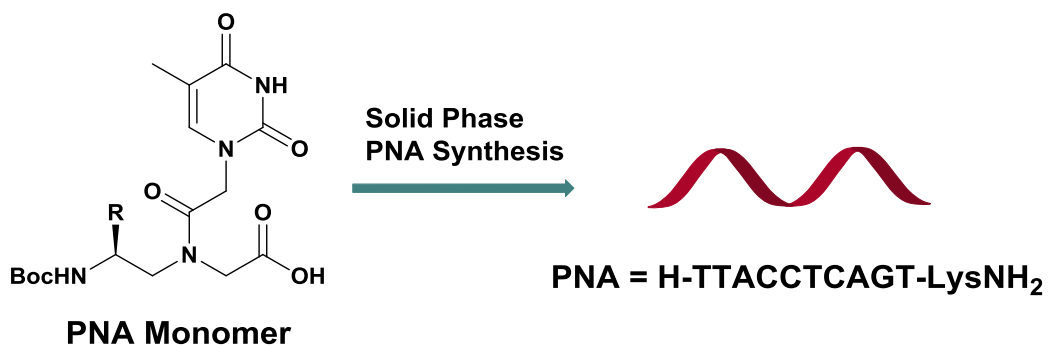
2B.1 Solid phase PNA synthesis

Peptides can be synthesized either by solution phase or by solid phase synthesis techniques.[35] The solution phase method of peptide synthesis can be used efficiently only for short chain peptides moreover, it requires a tedious separation and purification step after each coupling reaction. On the other hand, solid phase peptide synthesis can efficiently be used in the synthesis of several short and long chain peptides as well as in the synthesis of PNA oligomers.
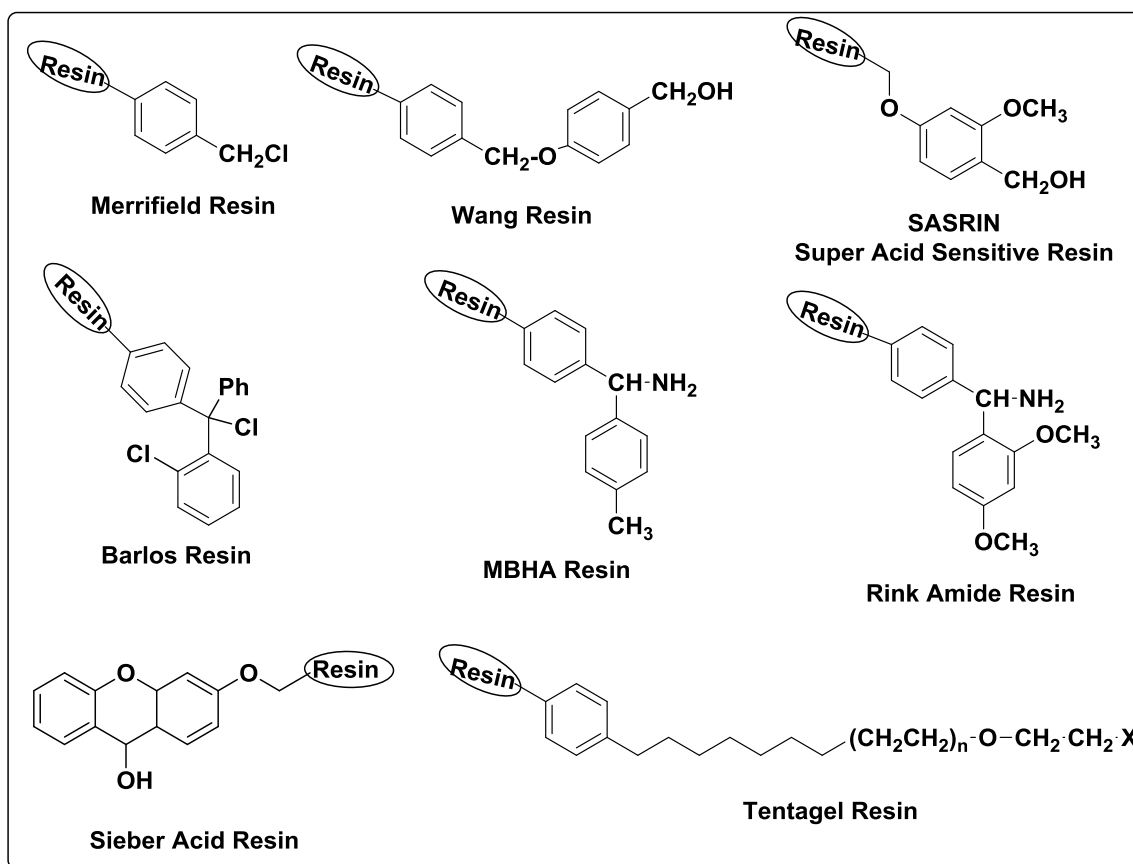
FIGURE 2.9 Representative structures of resin used in SPPS

The solid phase peptide synthesis first devised by Merrifield,[36] involves the combination of reagents with functional groups that are located on the surface and on the inside of beaded polymers.[37] Small solid beads are porous and insoluble in organic solvents. These beads are treated with functional units or linkers on which peptide chains can be built. The beads are immersed in solvent containing the reagents which approach the solvated sites by diffusion. In contrast to solution phase, the solid phase method offers several advantages. In solid phase synthesis, the *C*-terminal amino acid

is linked to an insoluble support that also acts as a permanent protection for the carboxylic acid (Figure **2.9**). Then the next N^α-protected amino acid is coupled to the resin bound amino acid either by using an active pentafluorophenyl (pfp) or 3-hydroxy-2,3-dihydro-4-oxo-benzotriazole (DHBt) ester, or by *in situ* activation with carbodiimide reagents. The excess amino acid is washed out and the deprotection and coupling reactions are repeated until the desired peptide sequence is achieved. The tedious purification step of intermediates after each coupling reaction is omitted. Finally, the resin bound peptide and the side chain protecting groups are cleaved in global deprotection step.

There are two routinely followed strategies of solid phase peptide synthesis-*Fmoc* strategy and *Boc* strategy which use base labile and acid labile protecting groups respectively (Figure **2.10**). Unlike ribosome protein synthesis, solid-phase peptide synthesis proceeds in a C-terminal to N-terminal fashion. The N-termini of amino acid monomers is protected by these two groups and added onto a deprotected amino acid chain. First protocol uses the *t*-butoxycarbony (*t*-Boc) group as N^α-protection that is removed by mild acidic conditions such as 50% TFA in DCM. The reactive side chains are protected with groups that are stable to *t*-Boc deprotection conditions and can be removed under strongly acidic conditions using HF in dimethylsufide or TFMSA in TFA. While in second protocol fluorenylmethyloxycarbonyl (Fmoc) group is used as N^α-protection which is extremely stable to acidic conditions but can be cleaved efficiently with a base such as piperidine. The final peptide and side chain protecting groups can be cleaved with acid (50% TFA).
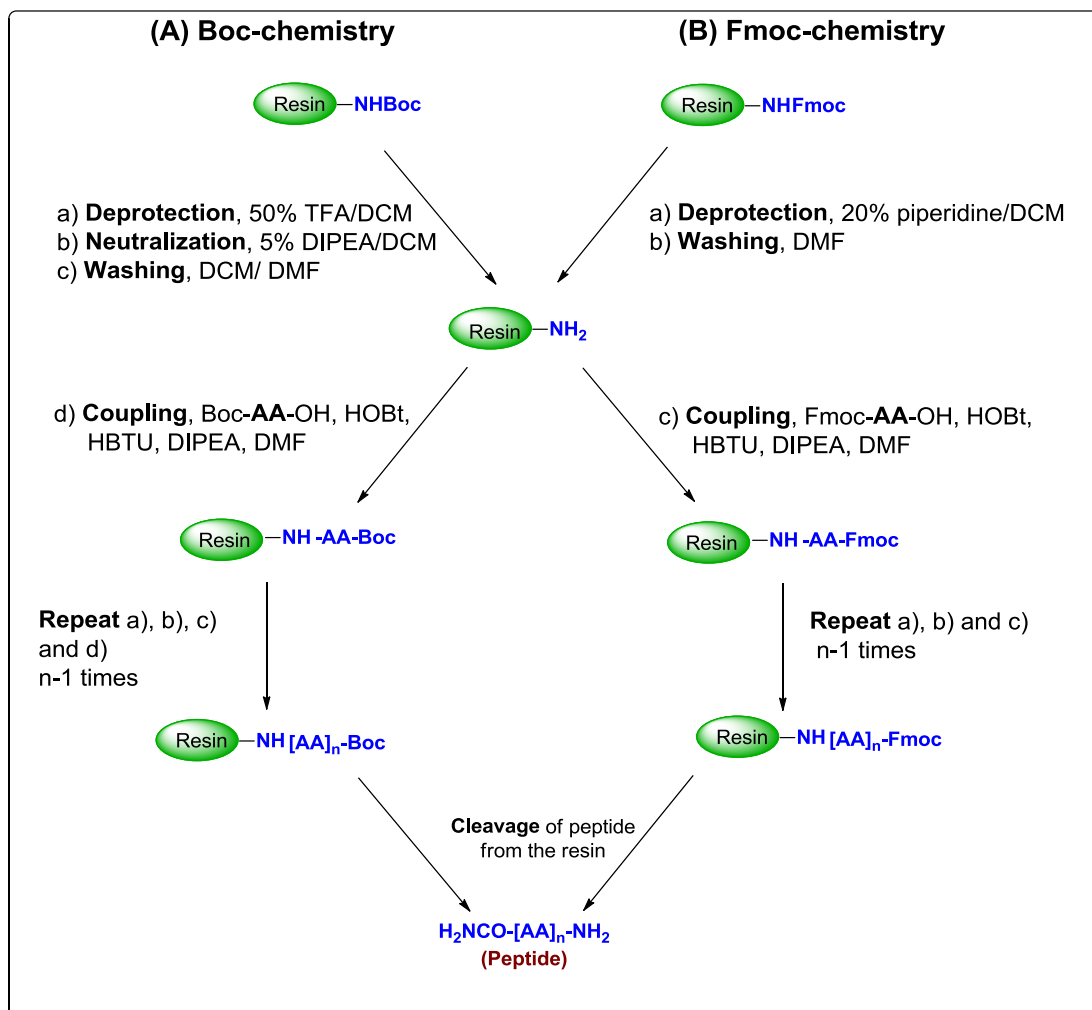
FIGURE 2.10 General protocols for SPPS via (A) Boc-chemistry (B) Fmoc-chemistry

2B.2 Aim of the present work

The specific objectives of this section are

➢ Incorporation of *aeg* PNA monomers (A/T/G/C) and γ-C-substituted PNA (T) monomers at various positions by solid phase synthesis

➢ Synthesis of fluorescent PNA oligomers by click chemistry

➢ Cleavage of oligomers from the solid support, their purification by RP-HPLC and characterization by MALDI-TOF spectrometry

2B.3 Results and discussion

The synthesis, purification and characterization of PNA oligomers incorporating modified (amino/guanidino/azido) as well as unmodified *aeg* PNA monomers in the PNA sequence has been discussed in this section.

2B.3.1 Synthesis of PNA oligomers

The PNA oligomers were synthesized by solid phase peptide synthesis protocol using *Boc* strategy. The modified PNA monomers were incorporated at desired positions in the unmodified *aeg* PNA sequence. PNA oligomerization was performed from C-terminus to the N-terminus end using monomeric units with protected amino and carboxylic acid functions maintaining the orthogonality. The Fmoc strategy was avoided because the terminal free amine generated has tendency to attack on the tertiary amide leading to acyl migration.[38]

MBHA resin (4-methyl–benzhydryl amine resin) was chosen as the solid support on which the oligomers were built and the monomers were coupled by *in situ* activation with HBTU / HOBt. In the synthesis of all oligomers, orthogonally protected (Boc/Cl-Cbz) L-lysine was selected as the *C*-terminal spacer-amino acid and it is linked to the resin through amide bond. The amine content on the resin was suitably lowered from 2 mmol/g to 0.35 mmol/g by partial acylation of amine content using calculated amount of acetic anhydride.[38] The free amine groups on the resin available for coupling was confirmed before starting synthesis by Kaiser's test.

The deprotection of the *N-t*-Boc protecting group and the completion of coupling reaction were monitored by Kaiser's test.[39] The *t*-Boc deprotection leads to a positive Kaiser's test, where the resin beads show blue color (Rheumann's purple). On the other hand, after completion of coupling reaction the resin beads were colorless which means a negative Kaiser's test. It is the most widely used qualitative test for the presence or absence of free amino group (deprotection/coupling).

Using the standard solid phase synthesis protocol (Figure **2.11**), the PNA oligomers of desired length incorporating modified as well as unmodified PNA monomers at desired positions were synthesized.
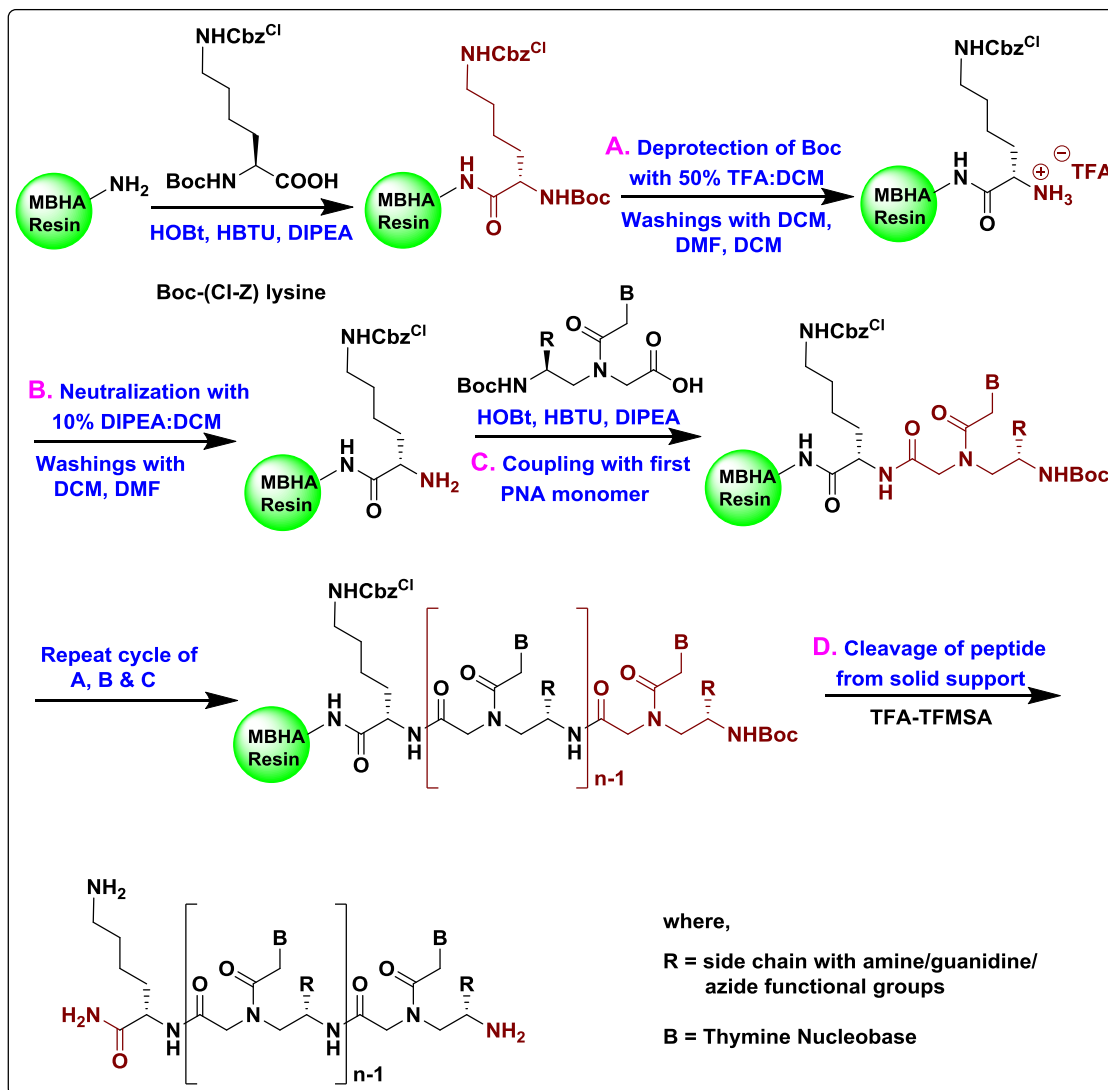
FIGURE 2.11 Solid phase PNA synthesis protocol by *Boc* strategy

2B.3.2 Synthesis of mixed purine-pyrimidine PNA oligomers

Mixed purine pyrimidine base sequences form duplexes of antiparallel or parallel orientations depending on their complementarity. By convention, in antiparallel orientation of PNA:DNA duplexes the *N*-terminal of the PNA faces the 3'-end of the DNA and *C*-terminal faces the 5'-end of DNA. While in parallel orientation, the *C*-terminal of PNA faces the 3'-end of the DNA and the *N*-terminal faces the 5'-end of the DNA (Figure **2.12**).
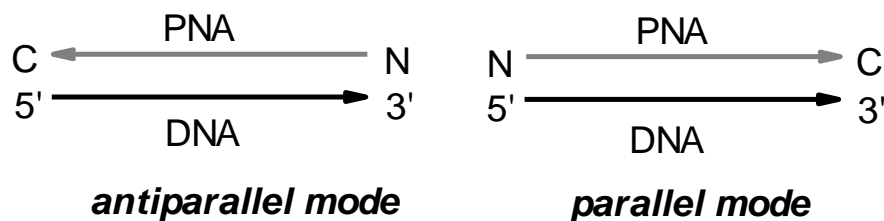
Figure 2.12 Antiparallel and parallel binding of PNA

In order to study the potential of duplex formation and the discrimination in DNA/RNA binding stability of the γ-C-substituted multifunctional PNA oligomers, the mixed purine pyrimidine sequences were synthesized. The PNA sequences bearing modified (T) and unmodified (A/T/G/C) monomeric units were synthesized to improve the duplex stability and their cell permeation abilities.

Choice of sequence

The transcription of the information in DNA into mRNA molecule is a two step process. First, the two strands of DNA structure unwind to allow the transcription machinery to come in and create a RNA strand which contains both the exons (coding region) and the introns (non-coding regions) called pre-mRNA. Secondly, the introns which does not code for any gene are cut and removed while the exons are brought together to produce mature mRNA as a transcript in a process known as splicing. However, genetic mutations that alter pre-mRNA splicing frequently cause an important segment of the RNA to be skipped out of the mature mRNA. In fact, various single nucleotide mutations that cause human disease affect splicing. Antisense ONs has been reported as corrective antisense in the cases of certain diseases caused by genetic mutations. In this approach, the antisense ONs act at pre-mRNA levels for splice corrections and yield mRNA that is translated into viable proteins. The antisense (18-mer) ON 705 corrects the aberrant splicing of mutated intron which interrupts the coding sequence of a luciferase gene (Figure **2.13**). They target the aberrant splice sites and restore the correct splicing pattern by causing the splicing machinery to use only the correct splice sites.
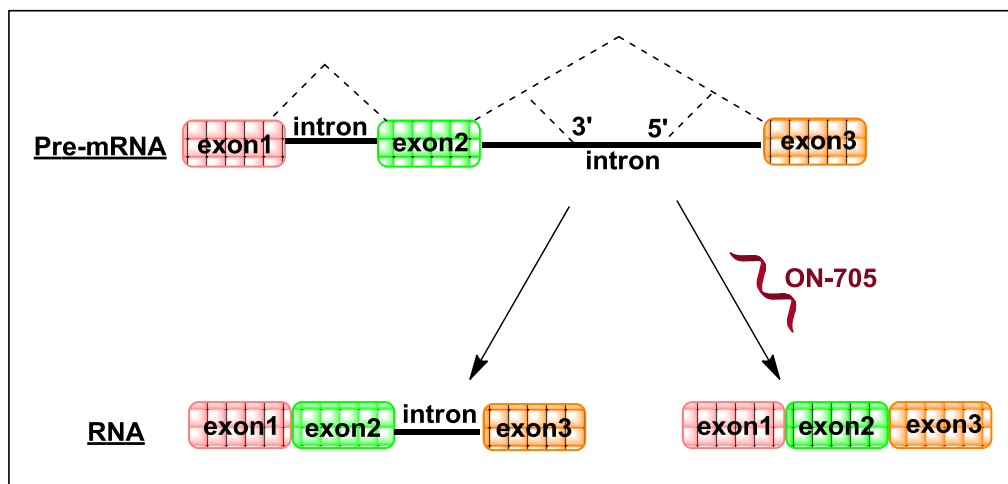
FIGURE 2.13 Correction of splicing of IVS2-705 β-globin pre-mRNA by ON-705

A middle portion of this antisense ON-705 was taken for consideration to carry out the hybridization and other biophysical studies (Figure **2.14**). PNA sequence **1** contains all unmodified *aeg* PNA (A/T/G/C) units and was used as a control sequence for comparative studies. Similarly, the mixed sequences were synthesized incorporating various modified PNA monomer units as *γ-(S-eam) aeg* **PNA 2-5**, *γ-(S-egd) aeg* **PNA 6-10**, *γ-(S-azm) aeg* **PNA 11-14** and *γ-(S-azb) aeg* **PNA 15-18** (Table **2.1**).
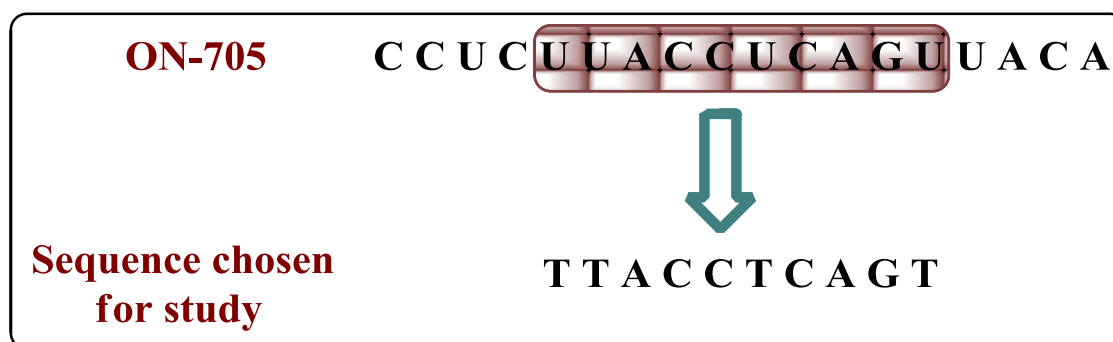
FIGURE 2.14 ON-705 sequence and corresponding 10-mer PNA sequence chosen for study

2B.3.3 Synthesis of fluorescent PNA oligomers by click chemistry

The fluorescent PNA oligomers were synthesized by solid phase click reaction[40] between the polymer supported azido component (azide containing PNA oligomers) and the acetylenic fluorescent component synthesized in solution. The fluorophore was incorporated in the side chain of PNA backbone at γ-position with various distances in order to investigate its potential as an effective conjugation point. The effect of bulky

fluorophore on the binding affinity and stability of PNA oligomer with its complementary DNA was also investigated using various biophysical techniques.

The solid supported azide containing PNA oligomers were treated with alkyne component in presence of CuI, ascorbic acid and DIPEA to obtain click product as fluorescent PNA oligomers (Figure **2.15**). The fluorescent PNA sequences synthesized by copper mediated azide-alkyne cycloaddition reaction are summarized in Table **2.2**
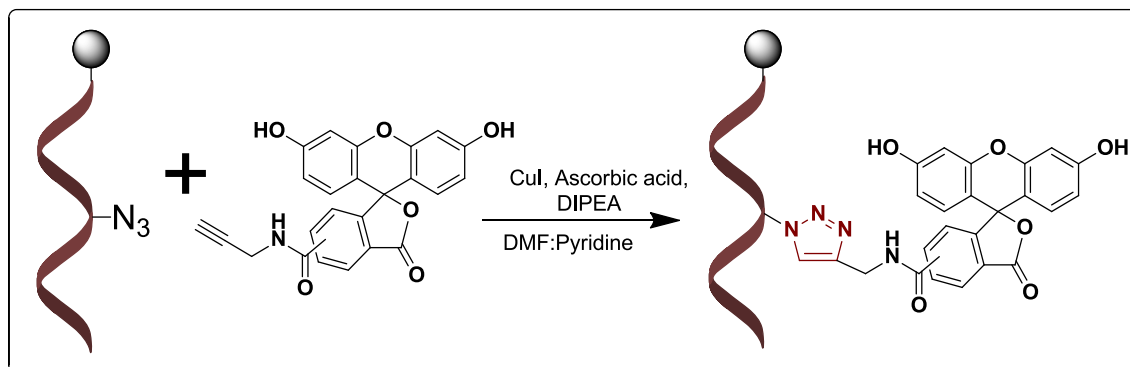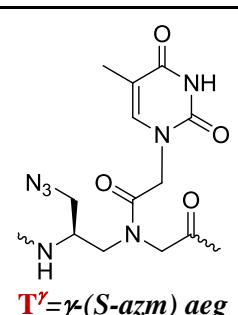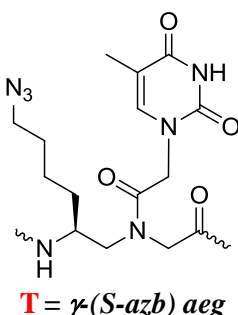
FIGURE 2.15 Synthesis of fluorescent PNA oligomers on solid support

TABLE 2.1 PNA oligomers with modified/unmodified monomers at various positions

Entry	Sequence Code	PNA sequences	Monomers used
1	*aeg* PNA **1**	H-T T A C C T C A G T-LysNH$_2$ 1 2 3 4 5 6 7 8 9 10	A/G/C/T = *aeg* PNA
2	*eam*-t$_2$ PNA **2**	H-T **t** A C C T C A G T-LysNH$_2$	
3	*eam*-t$_6$ PNA **3**	H-T T A C C **t** C A G T-LysNH$_2$	
4	*eam*-t$_{10}$ PNA **4**	H-T T A C C T C A G **t**-LysNH$_2$	
5	*eam*-t$_{2,6}$ PNA **5**	H-T **t** A C C **t** C A G T-LysNH$_2$	**t** = γ-(S-eam) aeg
6	*egd*-t$^γ_2$ PNA **6**	H-T **t^γ** A C C T C A G T-LysNH$_2$	
7	*egd*-t$^γ_6$ PNA **7**	H-T T A C C **t^γ** C A G T-LysNH$_2$	
8	*egd*-t$^γ_{10}$ PNA **8**	H-T T A C C T C A G **t^γ**-LysNH$_2$	
9	*egd*-t$^γ_{2,6}$ PNA **9**	H-T **t^γ** A C C **t^γ** C A G T-LysNH$_2$	
10	*egd*-t$^γ_{2,6,10}$ PNA **10**	H-T **t^γ** A C C **t^γ** C A G **t^γ**-LysNH$_2$	**t^γ** = γ-(S-egd) aeg

11	*azm*-T^{γ}_2 PNA **11**	H-T **T^{γ}** A C C T C A G T-LysNH$_2$	
12	*azm*-T^{γ}_6 PNA **12**	H-T T A C C **T^{γ}** C A G T-LysNH$_2$	
13	*azm*-T^{γ}_{10} PNA **13**	H-T T A C C T C A G **T^{γ}**-LysNH$_2$	
14	*azm*-$T^{\gamma}_{2,6}$ PNA **14**	H-T **T^{γ}** A C C **T^{γ}** C A G T-LysNH$_2$	**T^{γ}**=γ-(S-azm) aeg
15	*azb*-T_2 PNA **15**	H-T **T** A C C T C A G T-LysNH$_2$	
16	*azb*-T_6 PNA **16**	H-T T A C C **T** C A G T-LysNH$_2$	
17	*azb*-T_{10} PNA **17**	H-T T A C C T C A G **T**-LysNH$_2$	
18	*azb*-$T_{2,6}$ PNA **18**	H-T **T** A C C **T** C A G T-LysNH$_2$	**T** = γ-(S-azb) aeg

eam = ethyleneamino, *egd* = ethyleneguanidino, *azm* = azidomethylene and *azb* = azidobutylene

2B.3.4 Cleavage of the PNA oligomers from the solid support

The oligomers were cleaved from the solid support (L-lysine derivatized MBHA resin), using trifluoromethane sulphonic acid (TFMSA) in the presence of trifluoroacetic acid (TFA), which yielded PNA oligomers having L-lysine amide at their C-termini.[41] Side chain protecting groups of lysine as well as other nucleobase protections were also removed during this process. After cleavage reaction was over, the PNA oligomers obtained in solution were precipitated by addition of cold diethyl ether and the PNA oligomers were dissolved in de-ionized water.

TABLE 2.2 Fluorescent PNA oligomers synthesized by click reaction

Entry	Sequence Code	Fluorescent PNA Sequences
19	*azb*-T^f_6 PNA **16**	H-T T A C C **T^f** C A G T-LysNH$_2$
20	*azb*-T^f_{10} PNA **17**	H-T T A C C T A G **T^f**-LysNH$_2$
21	*azb*-$T^f_{2,6}$ PNA **18**	H-T **T^f** A C C **T^f** C A G T-LysNH$_2$
22	*azm*-t^f_2 PNA **11**	H-T **t^f** A C C T C A G T-LysNH$_2$
23	*azm*-t^f_6 PNA **12**	H-T T A C C **t^f** C A G T-LysNH$_2$
24	*azm*-t^f_{10} PNA **13**	H-T T A C C T C A G **t^f**-LysNH$_2$
25	*azm*-$t^f_{2,6}$ PNA **14**	H-T **t^f** A C C **t^f** C A G T-LysNH$_2$

2B.3.5 Purification and characterization of the PNA oligomers

After cleavage from the solid support, PNA oligomers were purified by reverse phase high performance liquid chromatography (RP-HPLC). The purification of PNA oligomers was carried out on a semi-preparative C18 column using a gradient system of acetonitrile and water. The purity of PNA oligomers was checked by reinjecting the sample on the same C18 semi-preparative column. All HPLC chromatograms are shown in Appendix I.

The integrity of these synthesized PNA oligomers was confirmed by MALDI-TOF mass spectrometry. In literature, various matrices like Sinapic acid (3,5-dimethoxy-4-hydroxycinnamic acid), picolinic acid (PA), 2,5-dihydroxybenzoic acid (DHB), α-cyano-4-hydroxycinnamic acid (CHCA) etc. have been reported to record MALDI-TOF spectrum. Of these, DHB was used as matrix to record MALDI-TOF spectra for all synthesized PNAs. The calculated and observed molecular weights for all PNAs with their molecular formulas and HPLC retention time in minutes are mentioned in Table **2.3**. The MALDI-TOF data for confirmation of mixed purine-pyrimidine PNA oligomers are shown in Appendix I.

TABLE 2.3 MALDI-TOF spectral analysis of the synthesized PNA oligomers

Sr. No.	PNA sequence code	Molecular Formula	Calculated Mass	Observed Mass	Retention Time (min)
1	*aeg* PNA **1**	$C_{113}H_{150}N_{55}O_{33}$	2805.1750 $[M + H]^+$	2805.0652	8.92
2	*eam*-t_2 PNA **2**	$C_{115}H_{155}N_{56}O_{33}$	2848.2172 $[M + H]^+$	2848.3081	10.02
3	*eam*-t_6 PNA **3**	$C_{115}H_{155}N_{56}O_{33}$	2848.2172 $[M + H]^+$	2848.2515	9.71
4	*eam*-t_{10} PNA **4**	$C_{115}H_{155}N_{56}O_{33}$	2848.2172 $[M + H]^+$	2848.5977	10.08
5	*eam*-$t_{2,6}$ PNA **5**	$C_{117}H_{159}N_{57}NaO_{33}$	2913.2414 $[M + Na]^+$	2913.3682	9.80
6	*egd*-t^{γ}_2 PNA **6**	$C_{116}H_{158}N_{58}NaO_{33}$	2913.8384 $[M + Na]^+$	2917.2097	10.40
7	*egd*-t^{γ}_6 PNA **7**	$C_{116}H_{158}N_{58}NaO_{33}$	2913.8384 $[M + Na]^+$	2917.1689	10.38
8	*egd*-t^{γ}_{10} PNA **8**	$C_{116}H_{159}N_{58}O_{33}$	2890.2390 $[M + H]^+$	2890.6897	9.82
9	*egd*-$t^{\gamma}_{2,6}$ PNA **9**	$C_{119}H_{164}N_{61}O_{33}$	2975.3030 $[M + H]^+$	2975.4937	10.07
10	*egd*-$t^{\gamma}_{2,6,10}$ PNA**10**	$C_{122}H_{171}N_{64}O_{33}$	3060.3670 $[M + H]^+$	3060.1923	10.21
11	*azm*-T^{γ}_2 PNA **11**	$C_{114}H_{150}N_{58}O_{33}$	2859.1842 $[M]^+$	2835.2915*	9.74
12	*azm*-T^{γ}_6 PNA **12**	$C_{114}H_{150}N_{58}O_{33}$	2859.1842 $[M]^+$	2835.3149*	10.08
13	*azm*-T^{γ}_{10} PNA **13**	$C_{114}H_{150}N_{58}O_{33}$	2859.1842 $[M]^+$	2835.2590*	9.63

14	*azm*-T$^\gamma_{2,6}$ PNA **14**	C$_{115}$H$_{151}$N$_{61}$O$_{33}$	2914.2013 [M]$^+$	2865.3467*	10.12
15	*azb*-T$_2$ PNA **15**	C$_{117}$H$_{156}$N$_{58}$O$_{33}$	2901.2312 [M]$^+$	2900.5166	10.17
16	*azb*-T$_6$ PNA **16**	C$_{117}$H$_{156}$N$_{58}$O$_{33}$	2901.2312 [M]$^+$	2899.7097	10.33
17	*azb*-T$_{10}$ PNA **17**	C$_{117}$H$_{156}$N$_{58}$O$_{33}$	2901.2312 [M]$^+$	2899.7190	10.30
18	*azb*-T$_{2,6}$ PNA **18**	C$_{121}$H$_{163}$N$_{61}$O$_{33}$	2998.2952 [M]$^+$	2949.0596*	10.26
19	*azb*-T$^f_6$ PNA **16**	C$_{141}$H$_{172}$N$_{59}$O$_{39}$	3315.3289 [M + H]$^+$	3318.2996	12.89
20	*azb*-T$^f_{10}$ PNA **17**	C$_{141}$H$_{172}$N$_{59}$O$_{39}$	3315.3289 [M + H]$^+$	3315.9607	13.88
21	*azb*-T$^f_{2,6}$ PNA **18**	C$_{169}$H$_{194}$N$_{63}$O$_{45}$	3825.4829 [M + H]$^+$	3828.1565	15.79
22	*azm*-t$^f_2$ PNA **11**	C$_{138}$H$_{166}$N$_{59}$O$_{39}$	3273.2820 [M + H]$^+$	3275.2666	13.67
23	*azm*-t$^f_6$ PNA **12**	C$_{138}$H$_{166}$N$_{59}$O$_{39}$	3273.2820 [M + H]$^+$	3276.1621	12.95
24	*azm*-t$^f_{10}$ PNA **13**	C$_{138}$H$_{166}$N$_{59}$O$_{39}$	3273.2820 [M + H]$^+$	3276.2029	13.17
25	*azm*-t$^f_{2,6}$ PNA **14**	C$_{163}$H$_{182}$N$_{63}$O$_{45}$	3741.3890 [M + H]$^+$	3744.2063	15.31

* Mass reported after subtracting $-N_2$ from the actual mass

2B.4 Summary

The rationally designed modified PNA monomers have been incorporated into 10-mer mixed purine-pyrimidine PNA sequence. These modified monomers were inserted into *aeg* PNA sequence at various positions using HOBt, HBTU and DIPEA as coupling reagent by solid phase peptide synthesis protocol. All the modified and unmodified PNA oligomers obtained by solid phase synthesis were cleaved from solid support using appropriate protocol. The PNA oligomers after cleavage were purified by RP-HPLC and characterized by MALDI-TOF spectrometry. Detailed experimental procedures and spectral data of all intermediates are discussed in section **2.2**. The next chapter deals with the investigation of biophysical properties of PNA oligomers.

2.2 Experimental

This section describes the detailed synthetic procedures and spectral characterization of the rationally designed PNA monomers.

2.2.1 General

The chemicals used were of laboratory or analytical grade. All the solvents used were distilled or dried to carry out various reactions. Reactions were monitored using thin layer chromatography (TLC). Usual workup involved sequential washing of the organic extract with water and brine followed by drying the organic layer over anhydrous sodium sulphate and evaporation of solvent under vacuum. TLCs were carried out on pre-coated silica gel GF_{254} sheets (Merck 5554). TLCs were analysed under UV lamp, by Iodine spray and by spraying with Ninhydrin solution, followed by heating of the plate. Column chromatographic separations were performed using silica gel (60-120 or 100-200 mesh). Melting points of solid compounds were determined in open capillary tubes using Veego's VMP-D melting point apparatus. IR spectra were recorded on Bruker's fourier transform infra red spectrometer as neat samples.

^{1}H and ^{13}C NMR spectra were recorded using Bruker AC-200 (200 MHz) or JEOL 400 MHz NMR spectrometers. The delta (δ) values for chemical shifts are reported in ppm and are referred to internal standard TMS or deuterated NMR solvents. The optical rotation values were obtained on Rudolph Research Analytical Autopol V polarimeter. Mass spectra for reaction intermediates were obtained by Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using TiO_2 or 2,5-dihydroxybencoic acid (DHB) and the integrity of PNA oligomer was checked on the same instrument using DHB or CHCA as matrx. High resolution mass spectra for final PNA monomers were recorded on Synapt G2 High Definition Mass Spectrometry. PNA oligomers were purified on Dionex ICS 3000 HPLC system using semi-preparative BEH130 C18 (10X250 mm) column.

2.2.2 Procedures and spectral data

2: 5-amino-2-((tert-butoxycarbonyl)amino)-5-oxopentanoic acid[42]
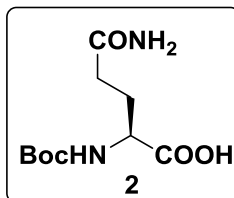
A solution of di-*tert*-butyl dicarbonate [(Boc)$_2$O] (16.4 g, 17.26 mL, 75.3 mmol) in dioxane (100 mL) was added in portion to an ice-cold solution of L-glutamine (10 g, 68.5 mmol) in 1 N NaOH (100 mL). The reaction mixture was further stirred at 0 $^{\circ}$C for 1 h. After completion of reaction dioxane was removed completely under vacuum from the reaction mixture. The aqueous layer was washed with diethyl ether to remove excess [(Boc)$_2$O]. The aqueous layer was cooled in ice-water bath, acidified to pH 2-3 by slow addition of saturated KHSO$_4$ solution and then extracted with ethyl acetate (3 x 150 mL). The combined organic extracts were dried over an. Na$_2$SO$_4$, filtered and concentrated to give compound **2** as a white powder which was used without further purification (14.5 g, 86% yield). mp = 119-121 $^{\circ}$C; R$_f$ = 0.39 EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ - 2.960 (*c* 0.5, Methanol); IR (neat) 3341, 3211, 2979, 2937, 1674, 1517 cm^{-1}; MS (MALDI-TOF) *m/z* calcd for C$_{10}$H$_{18}$N$_2$O$_5$ [M + K]$^+$ 285.0853, found 285.0501.

3: 4-amino-2-((tert-butoxycarbonyl)amino)butanoic acid[43]
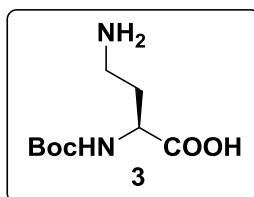
A Slurry of compound **2** (5 g, 20.3 mmol), ethyl acetate (24 mL), acetonotrile (24 mL), water (12 mL) and iodobenzene diacetate (7.87 g, 24.4 mmol) was cooled and stirred at 16 $^{\circ}$C for 30 min. The temperature was allowed to reach 20 $^{\circ}$C and the reaction mixture was stirred until completion (approximately 4 h). The reaction mixture was cooled to 0 $^{\circ}$C and filtered under vacuum. The filter cake was washed with ethyl acetate and dried in vacuum to obtain compound **3** (2.65 g, 60% yield). mp = 200-201 $^{\circ}$C; R$_f$ = 0.2 EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ + 13.6 (*c* 0.5, Methanol); IR (neat) 2976,

2900, 2832, 1740, 1692, 1645, 1563, 1546 cm^{-1}; MS (MALDI-TOF) m/z calcd for C$_9$H$_{18}$N$_{42}$O$_4$ [M + K]$^+$ 257.0904, found 257.0740.

4: 4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino) butanoic acid
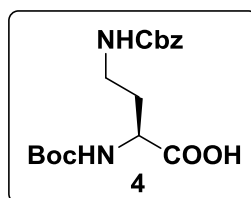
The solution of NaHCO$_3$ (2.3 g, 25 mL, 27.4 mmol) in water was added to an ice-cold solution of compound **3** (2 g, 9.17 mmol) in acetone (25 mL) and stirred for 10 min at 0 $^{\circ}$C. To this, benzylchloroformate (3.75 g 3.7 mL, 11 mmol) as 50 % solution in toluene was added and the reaction mixture was stirred overnight at room temperature. Acetone was removed completely under vacuum and the aqueous layer was washed with diethyl ether (2 x 20 mL). The aqueous layer was acidified to pH 2-3 with sat. KHSO$_4$ solution and extracted with ethyl acetate (3 x 50 mL). The combined organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated to give compound **4** as sticky oil (2.9 g, 90% yield). R$_f$ = 0.67 EtOAc/MeOH (50:50); IR (neat) 2931, 2354, 2318, 1792, 1770, 1705, 1647, 1626 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.40-7.31 (m, 5H), 5.66-5.64 (m, 1H), 5.48-5.43 (m, 1H), 5.14-5.04 (m, 2H), 4.35-4.32 (m, 1H), 3.50-3.06 (m, 2H), 2.07-1.76 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 175.7, 157.0, 156.0, 136.3, 128.5, 128.1, 80.3, 66.9, 51.1, 37.2, 33.3, 28.3; MS (MALDI-TOF) m/z calcd for C$_{17}$H$_{24}$N$_2$O$_6$ [M + K]$^+$ 391.1271, found 391.1075.

5: Methyl-4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino) butanoate
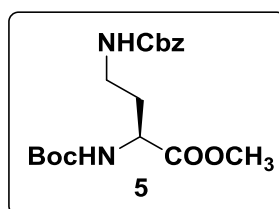
To a stirred solution of compound **4** (5.13 g, 14.5 mmol), K$_2$CO$_3$ (5.02 g, 36 mmol) in acetone (80 mL) was added dimethyl sulfate (1.7 mL, 17.4 mmol) and reaction mixture was heated to 55 $^{\circ}$C for 4 h under reflux condenser. Acetone was evaporated completely and water (100 mL) was added to the concentrate, which was then extracted

with ethyl acetate (3 x 60 mL). The combined organic layer was washed with brine, dried over an. Na_2SO_4, filtered and concentrated. The residue was purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **5** as white solid (5 g, 93% yield). mp = 67-69 $^\circ$C; R_f = 0.5 petroleum ether/EtOAc (70:30); $[\alpha]^{25}_D$ - 18.8 (*c* 0.5, Methanol); IR (neat) 3341, 2975, 1698, 1518, 1446 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 7.39-7.33 (m, 5H), 5.61 (br, 1H), 5.38 (app d, *J* = 8 Hz, 1H), 5.2-5.11 (m, 2H), 4.44-4.34 (m, 1H), 3.73 (s, 3H), 3.55-3.07 (m, 2H), 2.15-1.65 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 172.9, 156.3, 155.7, 136.5, 128.3, 127.9, 80.0, 66.5, 52.3, 50.8, 37.0, 33.2, 28.1; MS (MALDI-TOF) *m/z* calcd for C$_{18}$H$_{26}$N$_2$O$_6$ [M + K]$^+$ 405.1428, found 405.1199.

6: Benzyl tert-butyl (4-hydroxybutane-1,3-diyl)dicarbamate
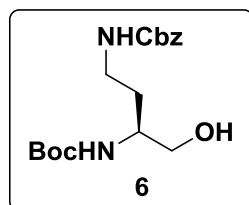
To a stirred solution of compound **5** (5 g, 13.6 mmol) in absolute ethanol (60 mL) was added sodium borohydride (1.56 gm, 41 mmol) and reaction mixture was stirred for 6 h under nitrogen atmosphere at RT. Ethanol was evaporated completely and water (100 mL) was added to the concentrate which was extracted with ethyl acetate (3 x 60 mL). The combined organic layer was washed with brine, dried over an. Na_2SO_4, filtered and concentrated. The residue was then purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **6** as white solid (4.1 g, 89% yield). mp = 80-82 $^\circ$C; R_f = 0.4 petroleum ether/EtOAc (50:50); $[\alpha]^{25}_D$ - 24.4 (*c* 0.5, Methanol); IR (neat) 3334, 2974, 1689, 1519, 1453 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.36-7.29 (m, 5H), 5.64 (br, 1H), 5.13-5.05 (m, 2H), 5.01 (br, 1H), 3.68-3.66 (app d, *J* = 8 Hz, 2H), 3.59-3.42 (m, 2H), 3.06-3.0 (m, 1H), 1.76-1.55 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 155.6, 136.5, 128.4, 128.0, 79.7, 66.6, 65.2, 49.7, 37.6, 32.0, 28.3; MS (MALDI-TOF) *m/z* calcd for C$_{17}$H$_{26}$N$_2$O$_5$ [M + K]$^+$ 377.1479, found 377.1080.

7: 4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino)butyl methane sulfonate

NHCbz

BocHN   OMs

7
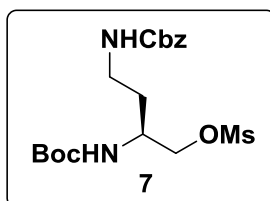
To an ice-cold solution of compound **6** (4 g, 12 mmol), triethyl amine (4.18 mL, 30 mmol) in dry DCM (50 mL) was added mesyl chloride (1.21 mL, 15.6 mmol) and reaction mixture was stirred for 30 min at 0 $^\circ$C under nitrogen atmosphere. To the reaction mixture DCM (20 mL) was added which was washed with water (40 mL) and brine (30 mL). The organic layer was dried over an. Na_2SO_4, filtered and concentrated on rota evaporator to give compound **7** (4.5 g, 92% crude yield). R_f = 0.53 petroleum ether/EtOAc (50:50). This compound was used for next step without further purification.

8: Benzyl tert-butyl (4-azidobutane-1,3-diyl)dicarbamate

NHCbz

BocHN   N$_3$

8
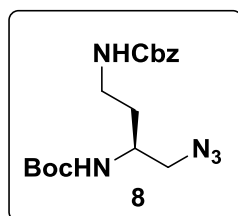
The solution of compound **7** (4.5 g, 11 mmol) and sodium azide (10.86 g, 165 mmol) in dry DMF (50 mL) was heated to 80 $^\circ$C for 8 h. To the reaction mixture water (100 mL) was added which was extracted with ethyl acetate (3 x 70 mL). The ethyl acetate layer was washed with water (50 mL) and brine (50 mL). The combined organic layer was dried over an. Na_2SO_4, filtered and concentrated. The residue obtained was then purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **8** as sticky yellowish oil (3.53 g, 87% yield). R_f = 0.73 petroleum ether/EtOAc (50:50); IR (neat) 3332, 2974, 2931, 2098, 1691, 1517 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.32-7.28 (m, 5H), 5.53 (br, 1H), 5.09-5.01 (m, 2H), 4.82 (br, 1H), 3.80-3.76 (m, 1H), 3.45-3.38 (m, 2H), 3.35-2.97 (m, 2H), 1.68-1.50 (m, 2H), 1.40 (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 156.4, 155.7, 136.5, 128.4, 128.0, 19.8, 66.5,

54.9, 47.6, 37.4, 32.8, 28.1; MS (MALDI-TOF) m/z calcd for $C_{17}H_{25}N_5O_4$ $[M + K]^+$ 402.1544, found 402.1709.

9: Ethyl 2-((4-(((benzyloxy)carbonyl)amino)-2-((tertbutoxycarbonyl)amino)butyl) amino)acetate
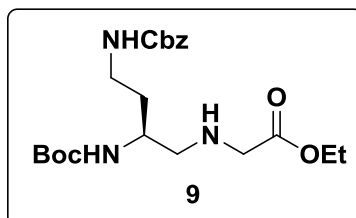
To a solution of compound **8** (750 mg, 2.06 mmol) in absolute ethanol (15 mL) taken in hydrogenation flask was added Raney Nickel (2 mL). The reaction mixture was hydrogenated in a Parr apparatus for 6 h at room temperature and H_2 pressure of 50-55 psi. The catalyst from reaction mixture was filtered off and the solvent was removed under reduced pressure to yield a residue of amine as yellowish oil. The amine compound (626 mg, 1.8 mmol) was treated with ethylbromo acetate (0.18 mL, 1.7 mmol) in acetonitrile (20 mL) using triethyl amine (0.77 mL, 5.5 mmol) and the reaction mixture was stirred at room temperature for 12 h. Acetonitrile was evaporated completely under vacuum and water (50 mL) was added to the concentrate. The aqueous layer was extracted with ethyl acetate (3 x 40 mL). The combined organic layer was washed with sat. $NaHCO_3$, brine, dried over an. Na_2SO_4, filtered and concentrated on rota evaporator. The residue obtained was purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **9** as yellowish oil (700 mg, 80%). R_f = 0.48 petroleum ether/EtOAc (20:80); IR (neat) 3329.89, 2929.46, 2354.43, 2318.10, 1738.10, 1705.25, 1694.13, 1647.56, 1529.97, 1516.80 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 7.25-7.20 (m, 5H), 5.76 (br, 1H), 5.06-4.93 (m, 3H), 4.14-4.03 (q, J = 8 Hz, 2H), 3.65-3.63 (m, 1H), 3.42-3.22 (m, 3H), 2.95-2.86 (m, 1H), 2.71-2.55 (m, 3H), 1.64-1.42 (m, 2H), 1.34 (s, 9H), 1.21-1.14 (t, J = 7 Hz, 3H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 171.9, 156.4, 136.6, 128.3, 127.9, 127.8, 79.4, 66.3, 60.8, 52.8, 50.4, 47.5, 37.4, 33.5, 29.5, 28.2, 14.0; MS (MALDI-TOF) m/z calcd for $C_{21}H_{33}N_3O_6$ $[M + K]^+$ 462.2006, found 426.2247.

10: Ethyl 2-(N-(4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino) butyl)-2-chloroacetamido)acetate
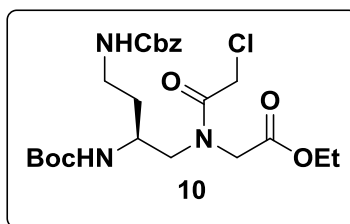
To an ice-cold solution of compound **9** (3.1 g, 7.3 mmol) and triethyl amine (2.96 g; 4 mL, 29.2 mmol) in dry DCM (50 mL) was added chloroacetyl chloride (0.82 g; 0.58 mL, 7.3 mmol) and reaction mixture was stirred for 8 h. To the reaction mixture DCM (20 mL) was added and washed with water (60 mL) and brine (60 mL). The organic layer was dried over an. Na_2SO_4, filtered and concentrated. The residue was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **10** as colorless sticky oil (2.63 g, 72%). $R_f = 0.59$ petroleum ether/EtOAc (40:60); ^{1}H NMR (200 MHz, $CDCl_3$) δ 7.33-7.26 (m, 5H), 5.61 (maj.) & 5.35 (min.) (br, 1H), 5.15-5.0 (m, 2H), 4.27-4.12 (m, 4H), 3.99 (s, 2H), 3.88-3.65 (m, 2H), 3.52-3.38 (m, 2H), 3.25-2.95 (m, 2H), 1.70-1.64 (m, 2H), 1.41 (min.) 7 1.40 (maj.) (s, 9H), 1.31-1.21 (m, 3H); ^{13}C NMR (50 MHz, $CDCl_3$) δ 169.1, 168.7, 168.3, 167.4, 156.5, 136.5, 128.4, 127.9, 79.4, 66.4, 62.0, 61.4, 52.8, 50.7, 49.8, 48.6, 46.8, 40.9, 37.3, 33.1, 32.3, 28.2, 14.0; MS (MALDI-TOF) m/z calcd for $C_{23}H_{34}ClN_3O_6$ [M + K]$^+$ 538.1722, found 538.1591.

11: Ethyl 2-(N-(4-(((benzyloxy)carbonyl)amino)-2-((tertbutoxycarbonyl)amino) butyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)yl)acetamido)acetate
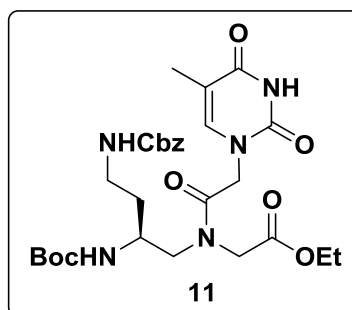
The solution of compound **10** (1 g, 2 mmol), K_2CO_3 (0.33 g, 2.4 mmol) and thymine (0.3 g, 2.4 mmol) in dry DMF (20 mL) was stirred at room temperature for 12 h. To the reaction mixture water (40 mL) was added and extracted with ethyl acetate (3

x 50 mL). The ethyl acetate layer was washed with water (50 mL) and brine (30 mL). The combined organic layer was dried over an. Na_2SO_4, filtered and concentrated. The residue obtained was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **11** as white solid (0.95 g, 81%). mp = 92-94 $^{\circ}$C; R_f = 0.47 EtOAc (100); $[\alpha]^{25}_D$ - 8.1 (*c* 1, Methanol); ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.86 (maj.) 7 9.56 (min.) (br, 1H), 7.37-7.27 (m, 5H), 7.05 (min.) & 6.98 (maj.) (s, 1H), 5.67-5.63 (maj.) & 5.58-5.56 (min.) (comp, 1H), 5.11-5.04 (m, 2H), 4.82 (min.) & 4.78 (maj.) (br, 1H), 4.47-4.37 (m, 1H), 4.28-4.13 (m, 3H), 3.94-3.51 (m, 3H), 3.43-3.02 (m, 3H), 2.17 (br, 1H), 1.89 (min.) & 1.87 (maj.) (s, 3H), 1.69-1.67 (comp, 1H), 1.41 (maj.) & 1.39 (min.) (s, 9H), 1.31-1.23 (m, 3H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 169.1, 167.5, 164.4, 156.7, 151.5, 140.9, 136.5, 128.4, 128.0, 110.8, 79.6, 66.6, 62.3, 52.1, 51.5, 49.8, 48.6, 47.8, 47.2, 37.6, 32.9, 28.3, 14.0, 12.3; MS (MALDI-TOF) *m/z* calcd for $C_{28}H_{39}N_5O_9$ [M + K]$^+$ 628.2385, found 628.2361.

12: 2-(N-(4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino) butyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid
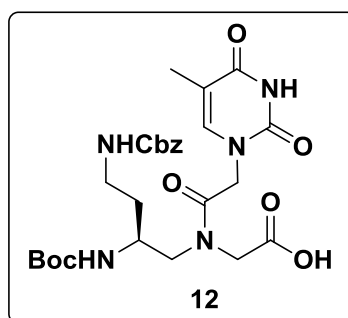
To a stirred solution of compound **11** (500 mg, 0.8 mmol) in methanol was added 10 % aq. LiOH and reaction mixture was stirred at room temperature for ~ 3-4 h. Methanol was removed under vacuum and the aqueous layer was washed with diethyl ether. The aqueous layer was then neutralized with activated Dowex H$^+$ resin till pH of the solution turned ~ 5-6. The resin was removed by filtration and the filtrate was concentrated to obtain the resulting compound **12** as white solid (0.42 g, 88%). mp = 241-245 $^{\circ}$C; R_f = 0.5 EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ - 3.0 (*c* 0.5, Methanol); ^{1}H NMR (400 MHz, DMSO-d^6) δ 11.30 (min.) & 11.25 (maj.) (br, 1H), 7.35-7.26 (m, 5H), 7.21-7.15 (m, 1H), 6.91-6.89 (min.) & 6.78-6.76 (maj.) (d, *J* = 8 Hz, 1H), 5.02-4.96 (m, 2H), 4.76-4.72 (maj.) & 4.59-4.55 (min.) (d, J = 16 Hz, 1H), 4.42 (s, 2H), 3.97-3.74 (m, 4H), 3.47-3.21 (m, 2H), 3.12-2.91 (m, 3H), 1.75 (min.) & 1.73 (maj.) (s, 3H), 1.58-1.41 (m,

2H), 1.37 (maj.) & 1.36 (min.) (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 170.6, 167.9, 167.0, 164.5, 156.0, 155.5, 151.1, 142.1, 137.3, 108.0, 77.7, 65.1, 51.9, 51.1, 48.4, 47.7, 47.1, 46.5, 37.7, 32.1, 28.3, 12.0; MS (MALDI-TOF) *m/z* calcd for C$_{26}$H$_{35}$N$_5$O$_9$ [M + K]$^+$ 600.2072, found 600.1787.

13. Ethyl 5-(((benzyloxy)carbonyl)amino)-9-((tert-butoxycarbonyl)amino)-11-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)-3-oxo-1-phenyl-2-oxa-4,6,11-triazatridec-4-en-13-oate
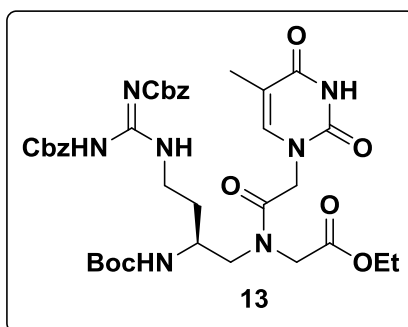
To a solution of compound **11** (1 gm, 1.7 mmol) in absolute ethanol (10 mL), was added 10% Pd/C (100 mg) and catalytic amount of glacial acetic acid (2 drops). The resulting reaction mixture was subjected to hydrogenation conditions on Parr Hydrogenator at 60 psi for 12 h. Reaction mixture was filtered and concentrated to get free amine (0.76 gm, 1.7 mmol). The free amine compound was dissolved in dry DMF (10 mL) to which N,N'-bis-di-Cbz-S-Me isothiourea (0.62 gm, 1.7 mmol) and catalytic amount of DMAP were added under N$_2$ atmosphere and the reaction mixture was stirred at RT for 12 h. To the reaction mixture water (20 mL) was added which was extracted with ethyl acetate (3 X 40 mL). Ethyl acetate layer was washed once each with water (40 mL) and brine (30 mL). The organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The residue obtained was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **13** as white solid (0.93 g, 72% yield). mp = 99-102 $^\circ$C; R$_f$ = 0.48 EtOAc (100); [α]$^{25}_D$ - 1.1 (*c* 1, Methanol); ^{1}H NMR (400 MHz, CDCl3) δ 11.70 (s, 1H), 9.24 (br, 1H), 8.52-8.51 (app d, *J* = 4 Hz, 1H), 7.39-7.28 (m, 10H), 7.02 (min.) & 6.98 (maj.) (s, 1H), 5.78 (d, *J* = 8 Hz, 1H), 5.22-5.09 (m, 4H), 4.84-4.80 (app d, *J* = 16 Hz, 1H), 4.45-4.13 (m, 5H), 3.96-3.69 (m, 2H), 3.64-3.31 (m, 4H), 2.09 (br, 1H), 1.90 & 1.88 (s, 3H), 1.82-1.58 (m, 2H), 1.41 (maj.) & 1.38 (min.) (s, 9H), 1.30 (min.) & 1.25 (maj.) (t, *J* = 8 Hz, 3H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 169.1, 168.4, 167.5, 164.2,

163.5, 156.1, 153.6, 151.2, 141.1, 136.6, 134.5, 128.8, 128.4, 128.0, 110.6, 79.8, 68.2, 67.1, 62.2, 61.4, 52.2, 51.3, 50.0, 48.8, 47.7, 37.9, 31.9, 28.3, 14.0, 12.3; MS (MALDI-TOF) m/z calcd for $C_{37}H_{47}O_7N_{11}$ 788.3231 [M + Na]$^+$, found 788.4692.

14. 5-(((benzyloxy)carbonyl)amino)-9-((tert-butoxycarbonyl)amino)-11-(2-(5- methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)-3-oxo-1-phenyl-2-oxa-4,6,11-triazatridec-4-en-13-oic acid
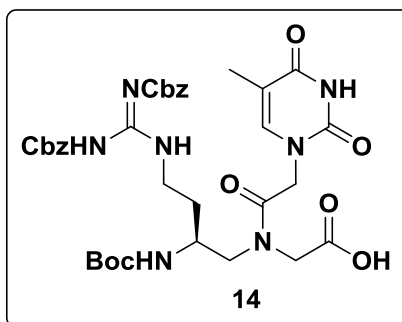
To a stirred solution of compound **13** (0.5 g, 0.65 mmol) in THF was added 2 N aqueous NaOH solution (4 mL) at 0 $^\circ$C and reaction mixture was stirred further for 30 mins at 0 $^\circ$C until complete consumption of starting material. THF was evaporated from the reaction mixture and the aqueous layer was washed with ethyl acetate (2 X 20 mL). The aqueous layer was neutralized at 0 $^\circ$C using 5 N HCl till the pH of solution becomes 3-4 and it was then extracted with ethyl acetate (3 X 30 mL). Combined organic layer was dried over an. Na2SO4, filtered and evaporated to give compound **14** as white solid (0.37 g, 78% yield). mp = 95-97 $^\circ$C; R$_f$ = 0.73 EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ + 3.7 (c 1, Methanol); ^{1}H NMR (400 MHz, CDCl$_3$) δ 11.74 (br , 1H), 10.10 (br, 1H), 8.53 (br, 1H), 7.37-7.28 (m, 10H), 6.99 (s, 1H), 6.62 (s, 1H), 5.76 (br, 1H), 5.30-5.10)m, 5H), 4.70-4.01 (m, 4H), 3.80-3.29 (m, 6H), 1.85 (maj.) & 1.82 (min.) (s, 3H), 1.40 (min.) & 1.38 (maj.) (s, 9H); MS (MALDI-TOF) m/z calcd for $C_{34}H_{41}O_7N_{11}$ 776.2658 [M + K]$^+$, found 776.3631.

16: 4-amino-2-((tert-butoxycarbonyl)amino)-4-oxobutanoic acid

A solution of di-*tert*-butyl dicarbonate [(Boc)$_2$O] (8.0 g, 8.42 mL, 41.6 mmol) in dioxane (50 mL) was added in portion to an ice-cold solution of L-asparagine (5 g, 37.8 mmol) in 1 N NaOH (50 mL). The reaction mixture was further stirred at 0 $^\circ$C for 1 h. After completion of reaction dioxane was removed completely from the mixture on rota evaporator.
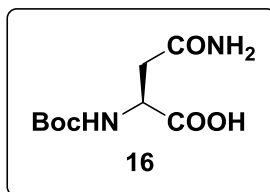
16

The aqueous layer was washed with diethyl ether to remove excess [(Boc)$_2$O]. The aqueous layer was cooled in ice-water bath, acidified to pH 2-3 by slow addition of saturated KHSO$_4$ solution and then extracted with ethyl acetate (3 x 60 mL). The combined organic extracts were dried over an. Na$_2$SO$_4$, filtered and concentrated to give compound **16** as a white powder which was used without further purification (11.2 g, 80% yield). R$_f$ = 0.4 EtOAc/MeOH (50:50); IR (neat) 3345, 3210, 2979, 2936, 1673, 1511 cm^{-1}; MS (MALDI-TOF) *m/z* calcd for C$_9$H$_{16}$N$_2$O$_5$ [M + K]$^+$ 271.0696, found 271.0638.

17: 3-amino-2-(Boc-amino)propanoic acid

17
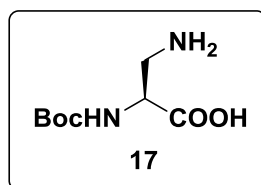
A Slurry of compound **16** (5 g, 21.5 mmol), ethyl acetate (24 mL), acetonotrile (24 mL), water (12 mL) and iodobenzene diacetate (8.32 g, 25.8 mmol) was cooled and stirred at 16 °C for 30 min. The temperature was allowed to reach 20 °C and the reaction mixture was stirred until completion (approximately 4 h). The reaction mixture was cooled to 0 °C and filtered under vacuum. The filter cake was washed with ethyl acetate and dried in vacuum to obtain compound **17** (2.87 g, 65% yield). R$_f$ = 0.25 EtOAc/MeOH (50:50); MS (MALDI-TOF) *m/z* calcd for C$_8$H$_{16}$N$_2$O$_4$ [M + K]$^+$ 243.0747, found 243.0456.

18: 3-azido-2-(Boc-amino)propanoic acid

18
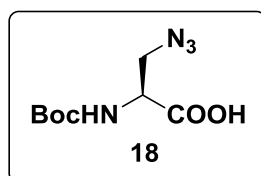
To a stirred solution of compound **17** (1.2 g, 5.9 mmol) in methanol:water (1:1) was added triethyl amine (2.5 mL, 17.9 mmol) and $CuSO_4.5H_2O$ (145 mg, 0.58 mmol) solution in water (0.5 mL). To this triflyl azide solution in DCM was added through dropping funnel and reaction mixture was stirred for 10 h. Methanol was evaporated completely under vacuum and sat. $NaHCO_3$ was added to the reaction mixture which was extracted with ethyl acetate (2 x 20 mL). The aqueous layer was acidified to pH 2-3 by slow addition of sat. $KHSO_4$ solution and extracted with ethyl acetate (3 x 30 mL). The combined ethyl acetate extracts were dried over an. Na_2SO_4, filtered and concentrated to give compound **18** as yellowish sticky liquid which was used for next step without further purification. R_f = 0.2 EtOAc (100); IR (neat) 3187, 2983, 2937, 2109, 1691, 1513 cm^{-1}; MS (MALDI-TOF) m/z calcd for $C_8H_{14}N_4O_4$ $[M + K]^+$ 269.0652, found 269.0127.

19: Methyl 3-azido-2-(Boc-amino)propanoate
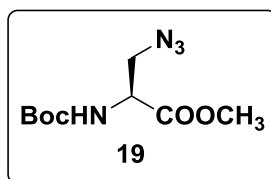
To a stirred solution of compound **18** (1 g, 4.34 mmol), K_2CO_3 (1.8 g, 13 mmol) in acetone (15 mL) was added dimethyl sulfate (0.65 g; 0.5 mL, 5.2 mmol) and reaction mixture was heated to 55 $^\circ$C for 4 h under reflux condenser. Acetone was evaporated completely and water (30 mL) was added to the concentrate, which was then extracted with ethyl acetate (3 x 40 mL). The combined ethyl acetate extracts were washed with brine, dried over an. Na_2SO_4, filtered and concentrated. The residue was purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **19** as colorless liquid (0.7 g, 66% yield). R_f = 0.55 petroleum ether/EtOAc (75:25); IR (neat) 3361, 2925, 2855, 2105, 1711, 1506 cm^{-1}; ^1H NMR (200 MHz, $CDCl_3$) δ 5.38 (app d, J = 6 Hz, 1H), 4.50-4.43 (m, 1H), 3.79 (s, 3H), 3.72 (app d, J = 2 Hz, 2H), 1.45 (s, 9H); ^{13}C NMR (50 MHz, $CDCl_3$) δ 170.2, 155.0, 80.4, 53.4, 52.7, 52.5, 28.1; MS (MALDI-TOF) m/z calcd for $C_9H_{16}N_4O_4$ $[M + K]^+$ 283.0809, found 283.0178.

20: 2-[(azidomethylene)-N-α-Boc-amino]ethanolamine
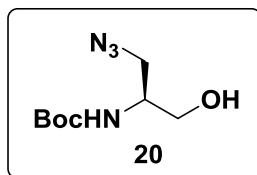
To a stirred solution of compound **19** (3.8 g, 15.6 mmol) in absolute ethanol (50 mL) was added sodium borohydride (1.78 g, 46.8 mmol) and reaction mixture was stirred for 6 h under nitrogen atmosphere at room temperature. Ethanol was evaporated completely and water (100 mL) was added to the concentrate which was extracted with ethyl acetate (3 x 75 mL). The combined organic layer was washed with brine, dri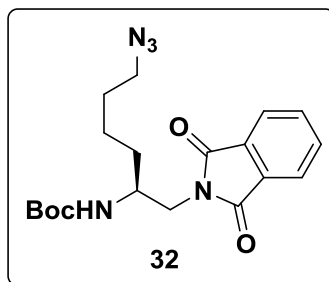ed over an. Na_2SO_4, filtered and concentrated. The residue was then purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **20** as colorless liquid (2.9 g, 86% yield). R_f = 0.35 petroleum ether/EtOAc (70:30); IR (neat) 3331, 2977, 2933, 2097, 1687, 1514 cm^{-1}; ^1H NMR (200 MHz, $CDCl_3$) δ 5.06 (app d, J = 6 Hz, 1H), 3.81-3.70 (m, 3H), 3.52-3.43 (m, 2H), 2.66 (br, 1H), 1.44 (s, 9H); ^{13}C NMR (50 MHz, $CDCl_3$) δ 155.7, 80.2, 62.2, 51.5, 28.3; MS (MALDI-TOF) m/z calcd for $C_8H_{16}N_4O3$ [M + K]$^+$ 255.0859, found 255.0220.

21: 3-azido-2-(Boc-amino)propyl methanesulfonate
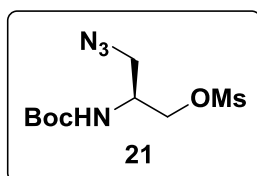
To an ice-cold solution of compound **20** (2.1 g, 9.72 mmol), triethyl amine (2.44 g, 3.38 mL, 24.3 mmol) in dry DCM (40 mL) was added mesyl chloride (1.48 g, 1 mL, 12.63 mmol) and reaction mixture was stirred for 30 min at 0 $^{\circ}$C under nitrogen atmosphere. To the reaction mixture DCM (20 mL) was added which was washed with water (20 mL) and brine (20 mL). The organic layer was dried over an. Na_2SO_4, filtered and concentrated on rota evaporator. to give compound **21** (2.29 g, 80% crude yield). R_f = 0.69 petroleum ether/EtOAc (50:50). This compound was used for next step without further purification.

22: tert-butyl 1-azido-3-(1,3-dioxoisoindolin-2-yl)propan-2-ylcarbamate
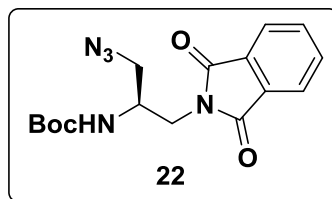
To a stirred solution of compound **21** (2.25 g, 7.65 mmol), K_2CO_3 (2.11 g, 15.3 mmol) in dry DMF (30 mL) was added potassium phthalimide (5.66 g, 30.6 mmol) and reaction mixture was heated to 80 $°$C for 12 h. To the reaction mixture water (60 mL) was added and extracted with ethyl acetate (3 x 50 mL). The organic layer was washed with water (50 mL) and brine (30 mL). The organic layer was then dried over an. Na_2SO_4, filtered and concentrated on rota evaporator. The residue obtained was purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **22** as white solid (1.8 g, 68% yield). mp = 114-117 $°$C; R_f = 0.72 petroleum ether/EtOAc (50:50); $[\alpha]^{25}_D$ + 50.0 (*c* 0.5, Methanol); IR (neat) 3370, 2977, 2936, 2102, 1773, 1711, 1512 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.88-7.86 (m, 2H), 7.74-7.72 (m, 2H), 5.04 (app d, *J* = 8 Hz, 1H), 4.12 (br, 1H), 3.89-3.73 (m, 2H), 3.58-3.55 (m, 2H), 1.30 (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 168.4, 155.3, 134.1, 131.9, 123.5, 79.9, 52.7, 49.4, 39.7, 28.1; MS (MALDI-TOF) *m/z* calcd for $C_{16}H_{19}N_5O_4$ [M + K]$^+$ 384.1074, found 384.0857.

23: Ethyl 2-((3-azido-2-((tert-butoxycarbonyl)amino)propyl)amino)acetate
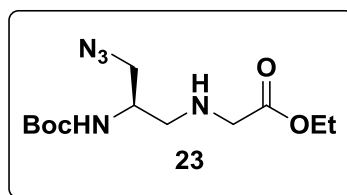
To a stirred solution of compound **22** (1.1 g, 3.18 mmol) in absolute ethanol (20 mL) was added hydrazine hydrate (3.09 g, 3 mL, 63.6 mmol) and reaction mixture was stirred for 4 h at room temperature. After completion of reaction ethanol was evaporated completely and water (30 mL) was added to the concentrate which was extracted with ethyl acetate (3 x 50 mL). The organic layer was dried over an. Na_2SO_4, filtered and concentrated to give free amine (0.68 g) which was taken in acetonitrile (20 mL). To this, triethyl amine (0.8 g, 1.1 mL, 7.9 mmol) was added and stirred at 0 $°$C

for 10 min. Ethylbromo acetate (0.46 g, 0.31 mL, 2.86 mmol) was then added dropwise and reaction mixture was stirred for 12 h at room temperature. Acetonitrile was evaporated completely under vacuum and water (50 mL) was added to the concentrate. The aqueous layer was extracted with ethyl acetate (3 x 40 mL). The combined organic layer was washed with sat. NaHCO$_3$, brine, dried over an. Na$_2$SO$_4$, filtered and concentrated on rota evaporator. The residue obtained was purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **23** as yellowish sticky oil (0.7 g, 73% yield). R$_f$ = 0.79 EtOAc (100); IR (neat) 3337, 2975, 2927, 2857, 2099, 1737, 1702, 1514 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.09 (br, 1H), 4.18 (q, J = 8 Hz, 2H), 3.76 (br, 1H), 3.49-3.33 (m, 4H), 2.81-2.67 (m, 2H), 2.24 (br, 1H), 1.44 (s, 9H), 1.27 (t, J = 8 Hz, 3H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 172.2, 155.4, 79.8, 60.9, 52.5, 50.7, 50.0, 29.6, 28.4, 14.1; MS (MALDI-TOF) m/z calcd for C$_{12}$H$_{23}$N$_5$O$_4$ [M + K]$^+$ 340.1387, found 340.0984.

24: Ethyl 2-(N-(3-azido-2-((tert-butoxycarbonyl)amino)propyl)-2-chloroacetamido) acetate
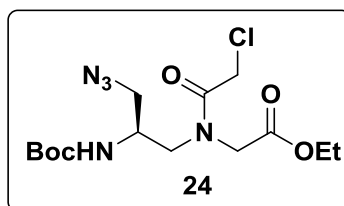
To an ice-cold solution of compound **23** (1.1 g, 3.65 mmol) and triethyl amine (1.1 g; 1.5 mL, 10.9 mmol) in dry DCM (20 mL) was added chloroacetyl chloride (0.45 g; 0.32 mL, 4 mmol) and reaction mixture was stirred for 8 h. To the reaction mixture DCM (20 mL) was added and washed with water (40 mL) and brine (40 mL). The organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated. The residue was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **24** as pale yellow sticky oil (1 g, 73% yield). R$_f$ = 0.65 petroleum ether/EtOAc (50:50); IR (neat) 3340, 2977, 2928, 2860, 2102, 1742, 1704, 1659, 1515 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.14-5.12 (maj.) 5.06-5.05 (min.) (app d, J = 8 Hz, 4 Hz, 1H), 4.27-4.18 (m, 4H), 4.02 (s, 2H), 3.89-3.65 (m, 2H), 3.59-3.41 (m, 3H), 1.44 (min.) 1.43 (maj.) (s, 9H), 1.33-1.30 (maj.) 1.29-1.27 (min.) (t, J = 6 Hz, 4 Hz, 3H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 168.8, 168.4, 167.6, 80.0, 62.2, 52.3, 51.3, 50.4, 49.1, 48.7,

40.9, 28.3, 14.1; MS (MALDI-TOF) m/z calcd for $C_{14}H_{24}ClN_5O_5$ $[M + K]^+$ 416.1103, found 416.1190.

25: Ethyl 2-(N-(3-azido-2-((tert-butoxycarbonyl)amino)propyl)-2-(5-methyl- 2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetate
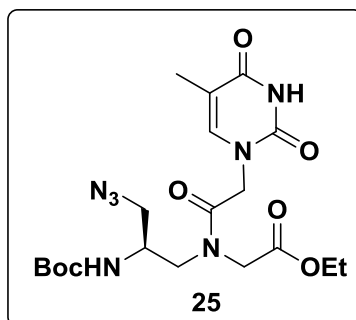

The solution of compound **24** (1 g, 2.65 mmol), K_2CO_3 (0.44 g, 3.18 mmol) and thymine (0.4 g, 3.18 mmol) in dry DMF (15 mL) was stirred at room temperature for 12 h. To the reaction mixture water (40 mL) was added and extracted with ethyl acetate (3 x 50 mL). The organic layer was washed with water (50 mL) and brine (30 mL). Ethyl acetate layer was dried over an. Na_2SO_4, filtered and concentrated. The residue was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **25** as white solid (1 g, 80% yield). R_f = 0.6 EtOAc (100); mp = 102-105 °C; $[\alpha]^{25}_D$ + 9.4 (*c* 1, Methanol); IR (neat) 3331, 2977, 2101, 1737, 1686, 1518, 1467, 1417 cm^{-1}; ^{1}H NMR (400 MHz, CD_3OD) δ 7.30-7.28 (comp, 1H), 4.79-4.50 (m, 2H), 4.4-4.1 (m, 4H), 4.1-3.8 (m, 2H), 3.75-3,55 (m, 2H), 3.45-3.35 (m, 1H), 3.31-3.24 (m, 1H), 1.88 (s, 3H), 1.45 (s, 9H), 1.34-1.31 (min.) 1.29-1.27 (maj.) (t, J = 6 Hz, 4 Hz, 3H); ^{13}C NMR (100 MHz, CD_3OD) δ 171.0, 170.3, 167.4, 158.2, 153.3, 144.0, 117.2, 111.3, 81.1, 62.9, 53.2, 51.4, 51.0, 50.4, 29.1, 14.8, 12.6; MS (MALDI-TOF) m/z calcd for $C_{19}H_{29}N_7O_7$ $[M + K]^+$ 506.1766, found 506.2396.

26: 2-(N-(3-azido-2-((tert-butoxycarbonyl)amino)propyl)-2-(5-methyl-2,4-dioxo-3,4-di hydropyrimidin-1(2H)-yl)acetamido)acetic acid

To a stirred solution of compound **25** (1 g, 2.14 mmol) in THF, a solution of 10% LiOH in water was added and the reaction mixture was stirred for 2 h. THF was removed completely under vacuum and the aqueous layer was washed with diethyl ether.

26
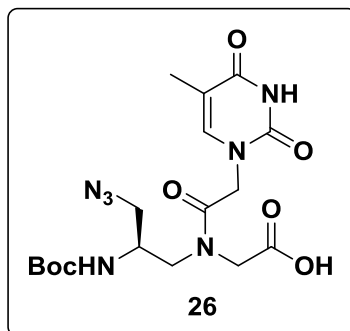
The aqueous layer was acidified in ice-water bath with sat. $KHSO_4$ to pH 3-4 and extracted with ethyl acetate (3 x 40 mL). The combined organic layer was dried over an. Na_2SO_4, filtered and concentrated to give final acid monomer **26** as white solid (0.85 g, 90% yield). mp = 150-152 $^{\circ}$C; R_f = 0.5 EtOAc/MeOH (50:50); $[\alpha]^{25}_D$ + 11.2 (c 0.5, Methanol); IR (neat) 3338, 2975, 2822, 2371, 2317, 2099, 1733, 1682, 1670, 1525 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d^6) δ 11.31 (maj.) & 11.28 (min.) (s, 1H), 7.27-7.26 (comp, 1H), 7.15-7.13 (maj.) & 7.01-6.98 (min.) (d, J = 8 Hz, 12 Hz, 1H), 4.73-4.41 (m, 2H), 4.27-4.21 (m, 1H), 4.05-3.86 (m, 2H), 3.60-3.40 (m, 3H), 3.23-3.0 (m, 1H), 1.75 (s, 3H), 1.39 (min.) & 1.38 (maj.) (s, 9H); ^{13}C NMR (100 MHz, DMSO-d^6) δ 170.3, 167.4, 164.4, 155.2, 151.0, 141.9, 108.2, 78.4, 51.2, 49.4, 49.2, 48.5, 47.7, 39.5, 28.2, 11.9; MS (MALDI-TOF) m/z calcd for $C_{17}H_{25}N_7O_7$ [M + K]$^+$ 478.1453, found 478.2229.

28: 6-azido-2-((tert-butoxycarbonyl)amino)hexanoic acid

28
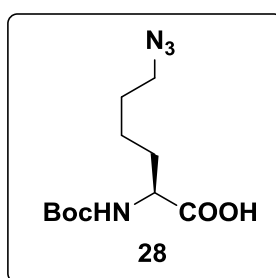
To a stirred solution of N-α-(tert-Butoxycarbonyl)-L-lysine (246 mg, 1 mmol) in methanol:water (1:1) was added triethyl amine (0.3 g, 0.42 mL, 3 mmol) and aq. $CuSO_4.5H_2O$ (12.5 mg, 0.05 mmol) solution (0.5 mL). To this triflyl azide solution in DCM was added through dropping funnel and reaction mixture was stirred for 10 h at room temperature. Methanol was evaporated completely under vacuum and sat. $NaHCO_3$ was added to the reaction mixture which was extracted with ethyl acetate (2 x 20 mL). The aqueous layer was acidified to pH 2-3 by slow addition of sat. $KHSO_4$

solution and extracted with ethyl acetate (3 x 30 mL). The combined organic layer was dried over an. Na_2SO_4, filtered and concentrated to give compound **28** which was used for next step without further purification. $R_f = 0.22$ EtOAc (100); IR (neat) 3207, 2982, 2106, 1691, 1613 cm^{-1}; MS (MALDI-TOF) *m/z* calcd for $C_{11}H_{20}N_4O_4$ [M + K]$^+$ 311.1122, found 311.0664.

29: Methyl 6-azido-2-((tert-butoxycarbonyl)amino)hexanoate
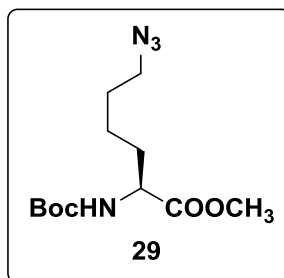
To a stirred solution of compound **28** (4.4 g, 16.2 mmol), K_2CO_3 (5.6 g, 40.4 mmol) in acetone (80 mL) was added dimethyl sulfate (2.44 g; 1.88 mL, 19.4 mmol) and reaction mixture was heated to 55 $^{\circ}$C for 4 h under reflux condenser. Acetone was evaporated completely and water (30 mL) was added to the concentrate, which was then extracted with ethyl acetate (3 x 40 mL). The combined organic layer was washed with brine, dried over an. Na_2SO_4, filtered and concentrated. The residue was purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **29** as yellowish liquid (3.8 g, 82% yield). $R_f = 0.8$ petroleum ether/EtOAc (60:40); IR (neat) 2925, 2856, 2095, 1707, 1512 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 5.06 (d, *J* = 8 Hz, 1H), 4.33-4.23 (m, 1H), 3.72 (s, 3H), 3.25 (t, *J* = 7 Hz, 2H), 1.76-1.53 (m, 6H), 1.42 (s, 9H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 173.1, 155.3, 80.0, 53.1, 52.3, 51.1, 32.2, 28.3, 28.2, 22.5; MS (MALDI-TOF) *m/z* calcd for $C_{12}H_{22}N_4O_4$ [M + K]$^+$ 325.1278, found 325.0638.

30: tert-butyl (6-azido-1-hydroxyhexan-2-yl)carbamate

To a stirred solution of compound **29** (1.6 g, 5.6 mmol) in absolute ethanol (30 mL) was added sodium borohydride (0.64 g, 16.8 mmol) and reaction mixture was stirred for 6 h under nitrogen atmosphere at room temperature. Ethanol was evaporated completely and water (100 mL) was added to the concentrate which was extracted with ethyl acetate (3 x 50 mL). The combined organic layer was washed with brine, dried over an. Na_2SO_4, filtered and concentrated.

30
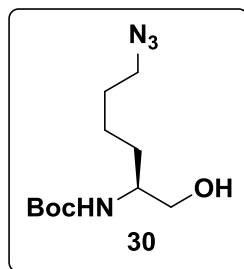
The residue was then purified on silica gel (60-120 mesh) using petroleum ether and ethyl acetate to give compound **30** as colorless liquid (1.2 g, 85% yield). R_f = 0.5 petroleum ether/EtOAc (50:50); IR (neat) 3341, 2937, 2868, 2093, 1684, 1511 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 4.84 (d, J = 4 Hz, 1H), 3.56-3.47 (m, 2H), 3.49-3.47 (m, 1H), 3.24 (t, J = 6 Hz, 2H), 3.12 (br, 1H), 1.72-1.52 (m, 6H), 1.40 (s, 9H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 156.3, 79.5, 65.2, 52.3, 51.2, 31.0, 28.6, 28.3, 23.1; MS (MALDI-TOF) m/z calcd for C$_{11}$H$_{22}$N$_4$O$_3$ [M + K]$^+$ 297.1329, found 297.0787.

31: 6-azido-2-((tert-butoxycarbonyl)amino)hexyl methanesulfonate

31
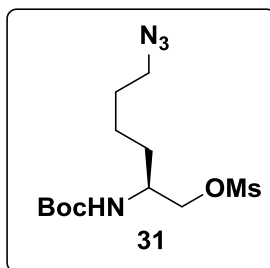
To an ice-cold solution of compound **30** (3.25 g, 12.6 mmol), triethyl amine (3.19 g, 4.4 mL, 31.5 mmol) in dry DCM (60 mL) was added mesyl chloride (1.87 g, 1.27 mL, 16.38 mmol) and reaction mixture was stirred for 30 min at 0 °C under nitrogen atmosphere. To the reaction mixture DCM (20 mL) was added which was washed with water (20 mL) and brine (20 mL). The organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated on rota evaporator to give compound **31** (3.5 g, 83% yield). R_f = 0.62 petroleum ether/EtOAc (50:50). This compound was used for next step without further purification.

32: tert-butyl (6-azido-1-(1,3-dioxoisoindolin-2-yl)hexan-2-yl)carbamate

To a stirred solution of compound **31** (3.5 g, 10.4 mmol), K_2CO_3 (2.87 g, 20.8 mmol) in dry DMF (50 mL) was added potassium phthalimide (7.7 g, 41.6 mmol) and reaction mixture was heated to 80 °C for 12 h. To the reaction mixture water (60 mL) was added and extracted with ethyl acetate (3 x 75 mL). The organic layer was washed with water (50 mL) and brine (30 mL). The organic layer was then dried over an. Na_2SO_4, filtered and concentrated on rota evaporator. The residue obtained was purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **32** as white solid (3.2 g, 80% yield). mp = 94-96 °C; R_f = 0.79 petroleum ether/EtOAc (60:40); $[\alpha]^{25}_D$ + 20.6 (c 0.5, Methanol); IR (neat) 3371, 2936, 2867, 2093, 1772, 1704, 1614, 1511 cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.84-7.82 (m, 2H), 7.70-7.68 (m, 2H), 4.62 (app d, J = 12 Hz, 1H), 3.97-3.93 (m, 1H), 3.70-3.60 (m, 2H), 3.27 (t, J = 8 Hz, 2H), 1.64-1.43 (m, 6H), 1.21 (s, 9H); ^{13}C NMR (100 MHz, CDCl$_3$) δ 168.5, 155.7, 133.9, 132.0, 123.2, 79.2, 51.2, 49.6, 42.2, 32.3, 28.5, 28.0, 23.0; MS (MALDI-TOF) m/z calcd for $C_{19}H_{25}N_5O_4$ [M + K]$^+$ 426.1544, found 426.0712.

33: Ethyl 2-((6-azido-2-((tert-butoxycarbonyl)amino)hexyl)amino)acetate
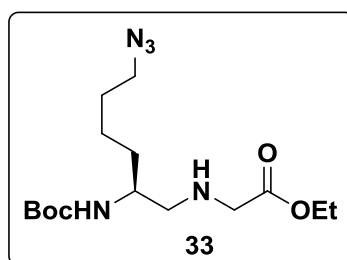
To a stirred solution of compound **32** (0.8 g, 2.06 mmol) in absolute ethanol (20 mL) was added hydrazine hydrate (2.06 g, 2 mL, 41.2 mmol) and reaction mixture was stirred for 4 h at room temperature. After completion of reaction ethanol was evaporated completely and water (30 mL) was added to the concentrate which was extracted with ethyl acetate (3 x 40 mL). The organic layer was dried over an. Na_2SO_4,

filtered and concentrated to give free amine (0.53 g) which was taken in acetonitrile (20 mL). To this, triethyl amine (0.5 g, 0.7 mL, 5 mmol) was added and stirred at 0 °C for 10 min. Ethylbromo acetate (0.3 g, 0.2 mL, 1.8 mmol) was then added dropwise and reaction mixture was stirred for 12 h at room temperature. Acetonitrile was evaporated completely under vacuum and water (50 mL) was added to the concentrate. The aqueous layer was extracted with ethyl acetate (3 x 40 mL). The combined organic layer was washed with sat. NaHCO$_3$, brine, dried over an. Na$_2$SO$_4$, filtered and concentrated on rota evaporator. The residue obtained was purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **33** as yellowish oil (0.55 g, 78% yield). R$_f$ = 0.5 EtOAc (100); IR (neat) 2926, 2859, 2094, 1741, 1672, 1517 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 4.74-4.71 (app d, J = 6 Hz, 1H), 4.24-4.13 (q, J = 8 Hz, 2H), 3.65-3.63 (m, 1H), 3.42-3.38 (d, J = 8 Hz, 2H), 3.30-3.24 (t, J = 6 Hz, 2H), 2.74-2.60 (m, 2H), 1.93 (br, 1H), 1.64-1.52 (m, 6H), 1.44 (s, 9H), 1.31-1.24 (t, J = 7 Hz, 3H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 172.4, 155.8, 79.2, 60.8, 52.9, 51.3, 50.9, 32.8, 28.7, 28.4, 23.1, 14.2; MS (MALDI-TOF) m/z calcd for C$_{15}$H$_{29}$N$_5$O$_4$ [M + K]$^+$ 382.1857, found 382.1853.

34: Ethyl 2-(N-(6-azido-2-((tert-butoxycarbonyl)amino)hexyl)-2-chloroacetamido) acetate
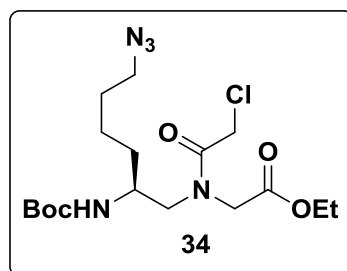
To an ice-cold solution of compound **33** (1 g, 2.9 mmol) and triethyl amine (0.88 g; 1.21 mL, 8.7 mmol) in dry DCM (20 mL) was added chloroacetyl chloride (0.326 g; 0.23 mL, 2.9 mmol) and reaction mixture was stirred for 8 h. To the reaction mixture DCM (20 mL) was added and washed with water (40 mL) and brine (40 mL). The organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated. The residue was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **34** as yellowish sticky oil (0.9 g, 74% yield). R$_f$ = 0.67 petroleum ether/EtOAc (50:50); IR (neat) 3333, 2977, 2938, 2868, 2095, 1742, 1694, 1655, 1515 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 4.76-4.72 (min.) & 4.62-4.5 (maj.) (d, J = 8 Hz,

6Hz, 1H), 4.29-4.08 (m, 4H), 4.01 (s, 2H), 3.91-3.60 (m, 2H), 3.53-3.13 (m, 3H), 1.74-1.49 (m, 6H), 1.43 (maj.) & 1.42 (min) (s, 9H), 1.34-1.25 (comp, 3H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 168.7, 167.3, 155.5, 79.1, 61.8, 53.1, 51.0, 49.5, 48.7, 40.9, 32.0, 28.4, 28.1, 23.2, 13.9; MS (MALDI-TOF) *m/z* calcd for C$_{17}$H$_{30}$ClN$_5$O$_5$ [M + K]$^+$ 458.1573, found 458.1383.

35: Ethyl 2-(N-(6-azido-2-((tert-butoxycarbonyl)amino)hexyl)-2-(5-methyl-2,4-dioxo -3, 4-dihydropyrimidin-1(2H)-yl)acetamido)acetate
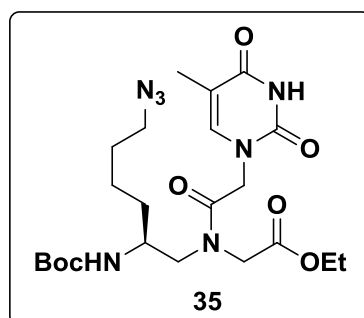
The solution of compound **34** (1 g, 2.38 mmol), K$_2$CO$_3$ (0.4 g, 2.9 mmol) and thymine (0.36 g, 2.9 mmol) in dry DMF (15 mL) was stirred at room temperature for 12 h. To the reaction mixture water (40 mL) was added and extracted with ethyl acetate (3 x 50 mL). The ethyl acetate layer was washed with water (50 mL) and brine (30 mL). The combined organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated. The residue obtained was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **35** as white sticky solid (1 g, 85% yield). R$_f$ = 0.63 EtOAc (100); [α]^{25}$_D$ + 1.7 (*c* 1, Methanol); IR (neat) 2936, 2354, 2317, 2097, 1792, 1769, 1740, 1705, 1677, 1658 cm^{-1}; ^{1}H NMR (200 MHz, CDCl$_3$) δ 9.52 (maj.) 9.32 (min.) (br, 1H), 7.09, 7.03 (s, 1H); 5.33-5.29 (d, *J* = 8 Hz, 1H), 5.02-4.71 (comp, 1H), 4.54-4.12 (m, 5H), 3.97-3.77 (comp, 1H), 3.56-3.22 (m, 5H), 1.91 (min.) 1.90 (maj.) (s, 3H), 1.62-1.48 (m, 6H), 1.43 (maj.) 1.39 (min.) (s, 9H), 1.31-1.23 (comp, 3H); ^{13}C NMR (50 MHz, CDCl$_3$) δ 174.8,169.1, 167.4, 164.5, 155.9, 151.2, 141.3, 141.0, 110.6, 79.5, 61.4, 52.6, 51.1, 49.3, 47.6, 31.7, 28.5, 28.3, 23.3, 14.1, 12.3; MS (MALDI-TOF) *m/z* calcd for C$_{22}$H$_{35}$N$_7$O$_7$ [M + K]$^+$ 548.2235, found 548.3024.

36: 2-(*N*-(6-azido-2-((tert-butoxycarbonyl)amino)hexyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid
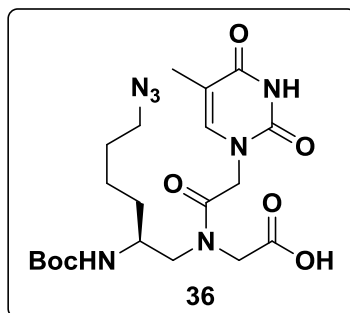
To a stirred solution of compound **35** (0.5 g, 0.98 mmol) in THF, a solution of 10% aq. LiOH was added and the reaction mixture was stirred for 2 h. THF was removed completely under vacuum and the aqueous layer was washed with diethyl ether. The aqueous layer was acidified in ice-water bath with sat. KHSO$_4$ to pH 3-4 and extracted with ethyl acetate (3 x 40 mL). The combined organic layer was dried over an. Na$_2$SO$_4$, filtered and concentrated to give final acid monomer **36** as off white solid (0.4 g, 84% yield). mp = 167-169 $^\circ$C; R$_f$ = 0.48 EtOAc/MeOH (50:50); [α]$^{25}_{D}$ + 1.4 (*c* 1, Methanol); IR (neat) 3361, 2937, 2318, 2097, 1677, 1530 cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d^6) δ 11.29 (maj.) & 11.27 (min.) (br s, 1H), 7.22 (s, 1H), 6.82 (maj.) & 6.67 (min.) (d, *J* = 8 Hz, 1H), 4.75-4.45 (m, 2H), 4.21-3.86 (m, 3H), 3.69-3.52 (m, 2H), 3.2-2.96 (m, 2H), 1.75 (s, 3H), 1.57-1.42 (m, 6H) 1.38 (min.) & 1.36 (maj.) (s, 9H); ^{13}C NMR (100 MHz, DMSO-d^6) δ 170.4, 167.3, 164.4, 155.7, 151.0, 141.9, 115.7, 108.1, 78.0, 51.5, 50.9, 50.6, 49.2, 48.5, 47.7, 30.8, 28.2, 28.0, 22.9, 12.0; MS (MALDI-TOF) *m/z* calcd for C$_{20}$H$_{31}$N$_7$O$_7$ [M + K]$^+$ 520.1922, found 520.2502.

38. Alkyne derivative of 5(6)-carboxyfluorescein[44]
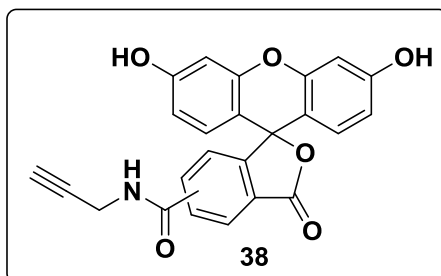
To a 5(6)-carboxyfluorescein-N-hydroxysuccinimidyl ester **37** (97 mg, 0.21 mmol) in THF (30 mL) was added propargylamine (32 μL, 0.45 mmol), and the reaction was stirred overnight under nitrogen atmosphere at room temperature. After

completion of reaction (monitored by TLC), THF was evaporated and the residue was washed with DCM and then with diethylether to afford 5(6)-carboxyfluorescein alkyne **38** as orange solid in a quantitative yield (87 mg). IR (neat) 3414, 3289, 3065, 2926, 2594, 2123, 1759, 1709, 1602 cm^{-1}; MS (MALDI-TOF) *m/z* calcd for $C_{24}H_{15}NO_6$ [M + K]$^+$ 452.0536, found 452.0947.

2.2.3 Synthesis of PNA oligomers by solid phase PNA synthesis

The modified and unmodified PNA monomers were incorporated into 10-mer PNA oligomers using standard solid phase protocol on L-lysine derivatized MBHA resin (50 mg) having 0.35 mmol/g loading value. All PNA oligomers were synthesized manually using the sintered glass column. The PNA monomers were coupled one after another to make a PNA sequence using HOBt, HBTU and DIPEA in DMF/NMP as coupling reagents. The PNA oligomers were synthesized using repetitive cycles, each comprising the following steps:

➢ Deprotection of the *N-t*-Boc group using 50% TFA in DCM (3 x 15 min)

➢ Washing of beads with DCM, DMF and again DCM (thrice each)

➢ Neutralization of the TFA salt of amine using 10% DIPEA in DCM to liberate free amine (3 x 10 min)

➢ Washing of beads with DCM and DMF (thrice each)

➢ Coupling of the free amine with the free carboxylic acid group of the incoming monomer (3 equivalents). The coupling reaction was carried out in DMF/NMP with HBTU as coupling reagent in the presence of DIPEA and HOBt.

➢ Capping (when needed) of the unreacted amino groups using acetic anhydride in pyridine:DCM

After each coupling and deprotection steps Kaiser's test was carried out for confirmation of PNA chain elongation which involved following steps:

➢ Ninhydrine (5.0 g) dissolved in ethanol (100 mL)

➢ Phenol (80.0 mg) dissolved in ethanol (20 mL)

➢ Potassium cyanide (2 mL, 0.001 M aq. Solution) added to 98 mL pyridine

➢ Few resin beads to be tested were taken in a test tube and washed with ethanol

➢ 3-4 drops from each of the above mentioned solutions were added to it

➢ The test tube was heated for 1-2 min

2.2.4 Synthesis of fluorescent PNA oligomers on solid support by click reaction

The resin bound azide functional group was reacted with the alkyne derivative of 5(6)-carboxyfluorescein to get the 1,2,3-triazole unit as fluorescent PNA oligomer. To the solid supported azide (10 mg, 0.35 mmol/g), a mixture of alkyne (8.66 mg, 6 eq), CuI (12 mg, 18 eq), ascorbic acid (3.1 mg, 5 eq) and DIPEA (15 µL, 24 eq) in DMF:pyridine (1:2) was added. The reaction was continued for 24 h at room temperature to get the desired click product. The click reaction gave around 60-70 % yield in all cases which were confirmed by RP-HPLC.

2.2.5 Cleavage of the PNA oligomers from solid support

The MBHA resin (10 mg) with oligomers attached to it was stirred with thioanisole (20 µL) and 1, 2-ethanedithiol (8 µL) in an ice bath for 10 min. TFA (200 µL) was added to it in cooled condition and kept in ice bath. TFMSA (16 µL) was added slowly with stirring to dissipate the heat generated. The reaction mixture was stirred for 1.5 to 2 h at room temperature. The resin was removed by filtration under reduced pressure and washed twice with TFA. The filtrate was combined evaporated on rota evaporator at ambient temperature. The remaining amount of TFA was transferred to eppendorf tube and the peptide was precipitated with cold dry ether. The peptide was isolated by centrifugation and the precipitate was dissolved in de-ionized water.

2.2.6 Purification of the PNA oligomers by RP-HPLC

PNA purification was carried out on Dionex ICS 3000 HPLC system. For the purification of peptides, semi-preparative BEH130 C18 (10X250 mm) column was used. Purification of PNA oligomers was performed with gradient elution method: A to 100% B in 20 min; A= 0.1% TFA in $CH_3CN:H_2O$ (5:95); B= 0.1% TFA in $CH_3CN:H_2O$ (1:1) with flow rate of 3 mL/min. All the HPLC profiles were monitored at 254 nm wavelength. Fluorescent PNA oligomers were monitored on both 254 and 490 nm wavelengths.

2.3 References

1. Nielsen, P. E.; *Acc. Chem. Res*. **1999**, *32*, 624-630.

2. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.

3. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc*. **1992**, *114*, 1895-1897.

4. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566.

5. Demidov, V. V.; Potaman, V N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sonnichsen, S.H.; Nielsen, P.E. *Biochern. Phormucol*. **1994**, *48,* 1310-1313.

6. a) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed.* **1998,** *37*, 2797-2823. b) Koppelhus, U.; Nielsen, P. E. *Adv. Drug Delivery Rev.* **2003**, *55,* 267-280.

7. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38,* 404–412.

8. a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **1996**, *35*, 1939-1942. b) Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett*. **1998**, *39*, 4707-4710. c) Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 7,591-598.

9. Sugiyama, T.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 7317-7320.

10. Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. *Tetrahedron Lett.* **2005**, *46*, 8395-8399.

11. Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem*. **2000**, *4*, 931-943.

12. Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 591-598.

13. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem.Soc. Chem. Commun.* **1993**, 518-519.

14. Hyrup, B.; Egholm, M.; Nielsen, P.E.; Wittung, P.; Nordén, B.; Buchardt, O. *J. Am. Chem. Soc*. **1994,** *116*, 7964-7968.

15. (a) Nielsen, P. E. *Curr. Opin. Struct. Biol.* **1999**, *9*, 353-357. (b) Ganesh, K. N.; Nielsen. P. E. *Curr. Org. Chem.* **2000,** *4*, 931-943.

16. (a) Kumar, V. A. *Eur. J. Org. Chem*. **2002**, 2021-2032. (b) Govindaraju, T., Kumar, V. A.; Ganesh, K. N. *J. Org. Chem*. **2004**, *69*, 5725–5734.

17. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed*. **1996**, *35*, 1939-1942.

18. (a) Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli. R. *Eur. J. Org. Chem.* **2000**, 2905-2913. (b) Menchise, V.; Simone G. D.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, r.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Nat. Acad. Sci*. USA, **2003**, *100*, 12021-12026.

19. Kleiner, R.; Brudno, Y. M.; Birnbaum, E.; Liu, D. *J. Am. Chem. Soc*. **2008**, *130*, 14, 4646.

20. Dose, C.; Seitz, O.; *Org. Lett.* **2005**, *7*, 4365.

21. Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. *Tetrahedron Lett*. **2005**, *46*, 8395.

22. Dragulescue-Andrasi, A.; Rapireddy, S.; Freeza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc*. **2006**, *128*, 10258.

23. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6878.

24. Dragulescue-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. Commun*. **2005**, 244.

25. Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas, S. M.; Zon, G.; Livak, K. J.; Ly, D. H. *J. Org. Chem.* **2009**, *74*, 1509.

26. Deglane, G.; Abes, S.; Michel, T.; Privot, P.; Vives, E.; Debart, F.; Barvik, I.; Lebleu, B. Vasseur, J. *Chembiochem*. **2006**, *7*, 684.

27. Mitra, R.; Ganesh, K. N. *Chem. Commun*. **2011**, *47*, 1198-1200.

28. Mitra, R.; Ganesh, K. N. *J. Org*. *Chem*. **2012**, *77*, 5696-5707.

29. Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas, S. M.; Zon, G.; Livak, K. J.; Ly, D. H. *J. Org. Chem.* **2009**, *74*, 1509.

30. Manicardi, A.; Accetta, A.; Tedeschi, T.; Sforza, S.; Marchelli, R.; Corradini, R. *Artificial DNA:PNA & XNA* **2012**, *3(2)*, 53-62.

31. Devi, G.; Ganesh, K. N. *Artificial DNA:PNA & XNA* **2010**, *1(2)*, 68-75.

32. Amant, A. H. St.; Engvers, C.; Hudson, R. H. E. *Artificial DNA:PNA & XNA* **2013**, *4(1)*, 1-7.

33. Zhang, L. H.; Kauffman, G. S.; Pesti, J. A.; Yin, J. *J. Org*. *Chem.* **1997**, *62*, 6918-6920.

34. Liu, Q.; Tor, Y.; *Org*. *Lett*. **2003**, *5(14)*, 2571-2572.

35. (a) Bodansky, M.; Bodansky, A. *The Practice of Peptide Synthesis*, Springer-Verlog, Berlin, **1984**. (b) Stewart, J. M.; Young, J. D. *Solid Phase Peptide Synthesis*, W. H. Freeman & Co, New York, **1969**.

36. Merrifield, R. B. *J. Am. Chem. Soc.***1963**, *85*, 2149-2154.

37. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K.; Hansen, H. F.; Koch, C.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Peptide Sci.* **1995**, *3*, 175-183.

38. (a) Erickson, B. W.; Merrifield, R. B. *Solid Phase Peptide Synthesis. In the Proteins,* Vol. II, 3rd ed.; Neurath, H.; Hill, R. L. eds.; Academic Press, New York, **1976**, 255. (b) Merrifield, R. B.; Stewart, J. M.; Jernberg, N. *Anal. Chem.* **1966**, *38*, 1905-1914.

39. (a) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595-598 (b) Kaiser, E.; Bossinger, C. D.; Cplescott, R. L.; Olsen, D. B. *Anal. Chim. Acta*. **1980**, *118*, 149-151 (c) Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem.* **1981**, *117*, 147.

40. Devi, G.; Ganesh, K. N. *Artificial DNA:PNA & XNA* **2010**, *1(2)*, 68-75.

41. 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*, 175-183.

42. *Organic Syntheses* **1990**, *7*, 70; **1985**, *63*,160.

43. Zhang, L. H.; Kauffman, G. S.; Pesti, J. A.; Yin, J. *J. Org. Chem.* **1997**, *62*, 6918-6920.

44. Punna, S.; Kaltgrad, E.; Finn, M. G. *Bioconjugate Chem.* **2005**, *16*, 1536-1541.

2.4 Appendix I: Characterization data of synthesized compounds/PNA

Compound	Description of spectral characterization	Page No.
Compound **4**	^{1}H NMR and ^{13}C NMR of compound **4**	98
Compound **5**	^{1}H NMR and ^{13}C NMR of compound **5**	99
Compound **6**	^{1}H NMR and ^{13}C NMR of compound **6**	100
Compound **8**	^{1}H NMR and ^{13}C NMR of compound **8**	101
Compound **9**	^{1}H NMR and ^{13}C NMR of compound **9**	102
Compound **10**	^{1}H NMR and ^{13}C NMR of compound **10**	103
Compound **11**	^{1}H NMR and ^{13}C NMR of compound **11**	104
Compound **12**	^{1}H NMR and ^{13}C NMR of compound **12**	105
Compound **13**	^{1}H NMR and ^{13}C NMR of compound **13**	106
Compound **14**	^{1}H NMR of compound **14**	107
Compound **19**	^{1}H NMR and ^{13}C NMR of compound **19**	108
Compound **20**	^{1}H NMR and ^{13}C NMR of compound **20**	109
Compound **22**	^{1}H NMR and ^{13}C NMR of compound **22**	110
Compound **23**	^{1}H NMR and ^{13}C NMR of compound **23**	111
Compound **24**	^{1}H NMR and ^{13}C NMR of compound **24**	112
Compound **25**	^{1}H NMR and ^{13}C NMR of compound **25**	113
Compound **26**	^{1}H NMR and ^{13}C NMR of compound **26**	114
Compound **29**	^{1}H NMR and ^{13}C NMR of compound **29**	115
Compound **30**	^{1}H NMR and ^{13}C NMR of compound **30**	116
Compound **32**	^{1}H NMR and ^{13}C NMR of compound **32**	117
Compound **33**	^{1}H NMR and ^{13}C NMR of compound **33**	118
Compound **34**	^{1}H NMR and ^{13}C NMR of compound **34**	119
Compound **35**	^{1}H NMR and ^{13}C NMR of compound **35**	120
Compound **36**	^{1}H NMR and ^{13}C NMR of compound **36**	121
Compound **2** & **3**	MALDI TOF Mass of compound **2** and **3**	122
Compound **4** & **5**	MALDI TOF Mass of compound **4** and **5**	123
Compound **6** & **8**	MALDI TOF Mass of compound **6** and **8**	124
Compound **9** & **10**	MALDI TOF Mass of compound **9** and **10**	125
Compound **11** & **12**	MALDI TOF Mass of compound **11** and **12**	126
Compound **13** & **14**	MALDI TOF Mass of compound **13** and **14**	127

^{1}H NMR of Compound 4

^{13}C NMR of Compound 4
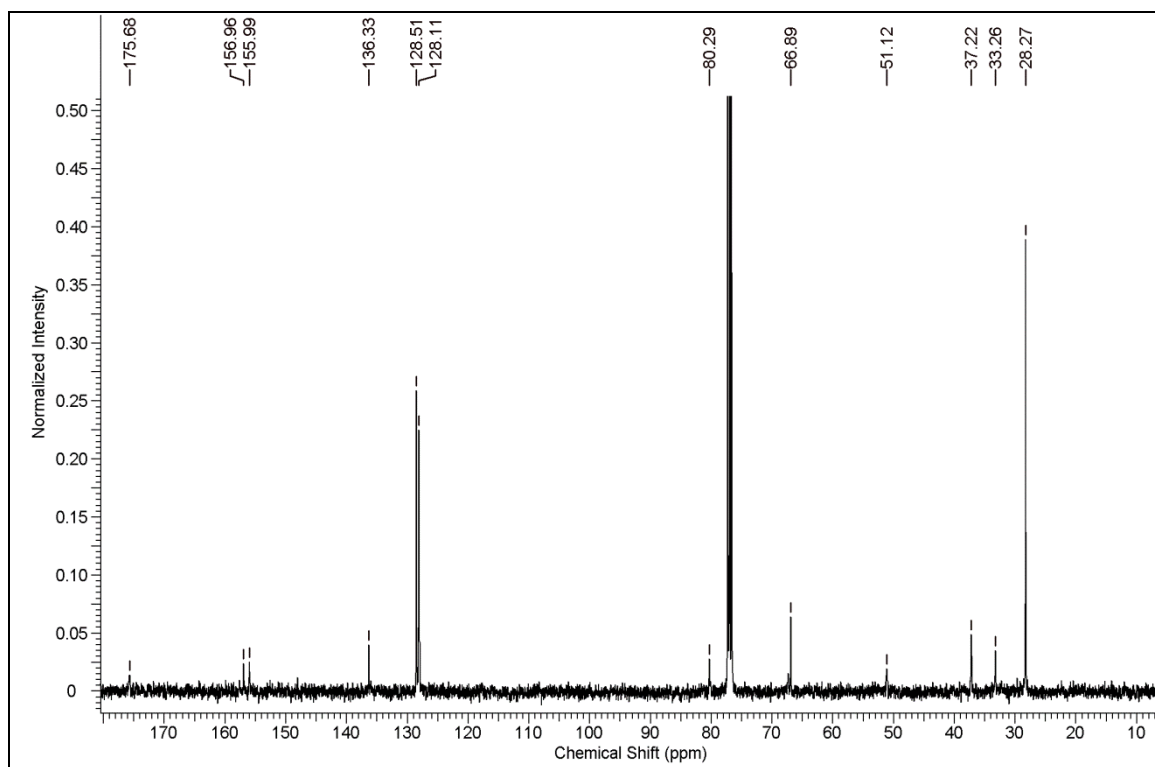
¹H NMR of Compound 5
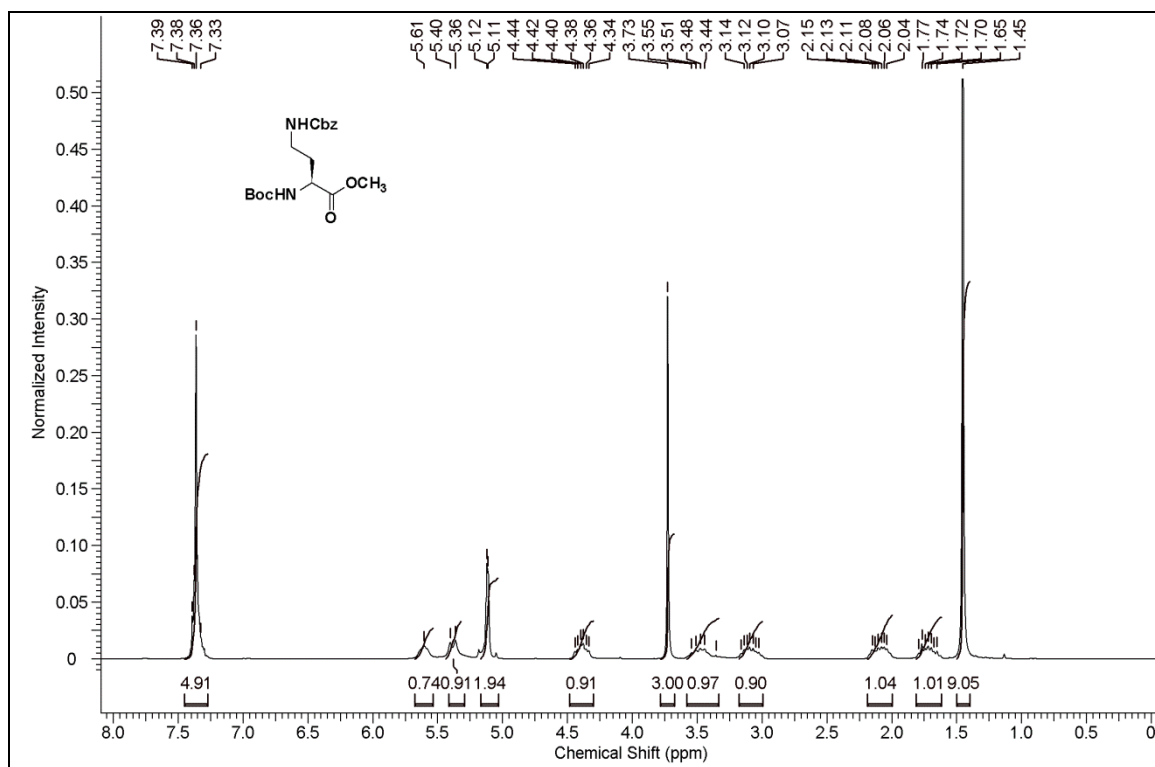
¹³C NMR of Compound 5
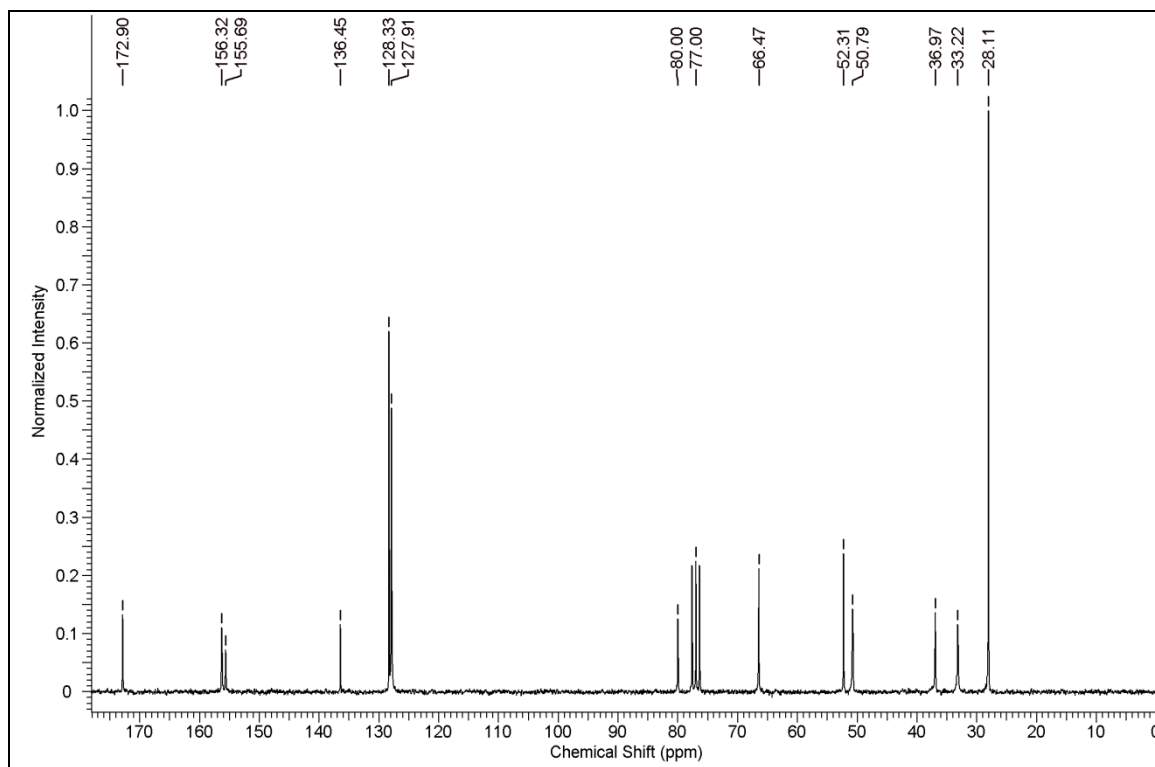
^{1}H NMR of Compound 6
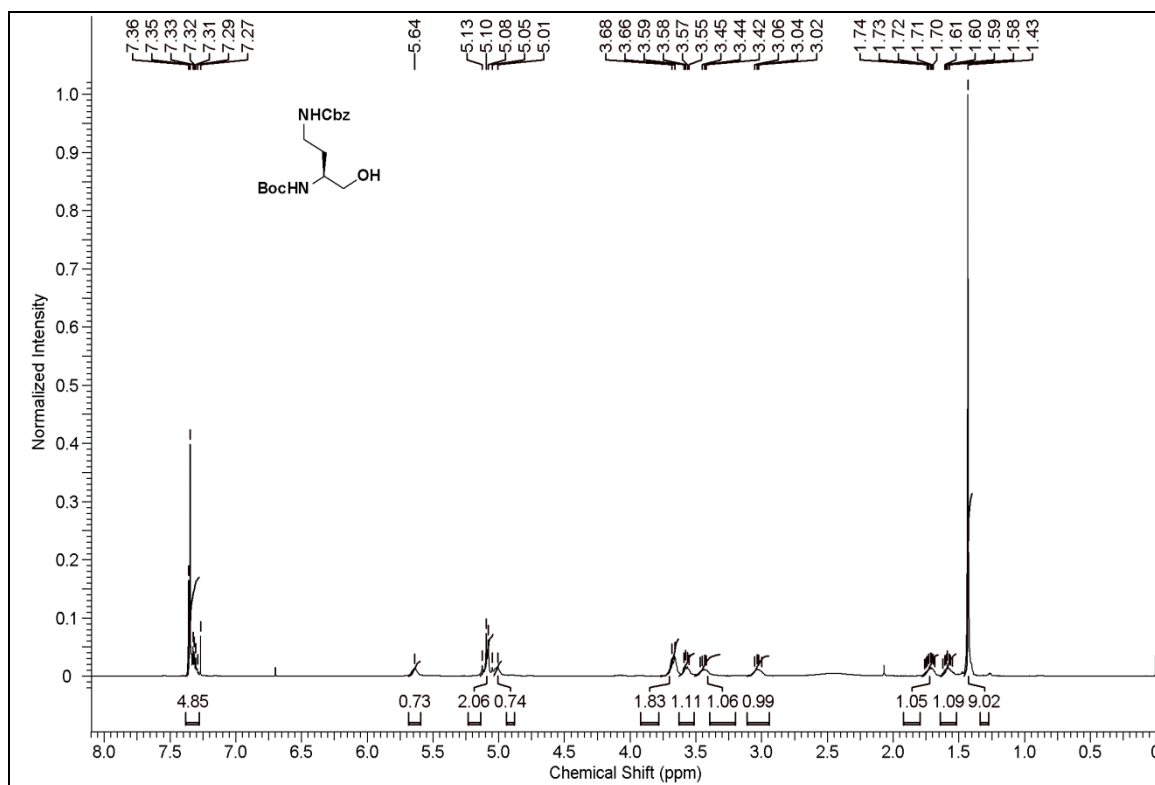
^{13}C NMR of Compound 6
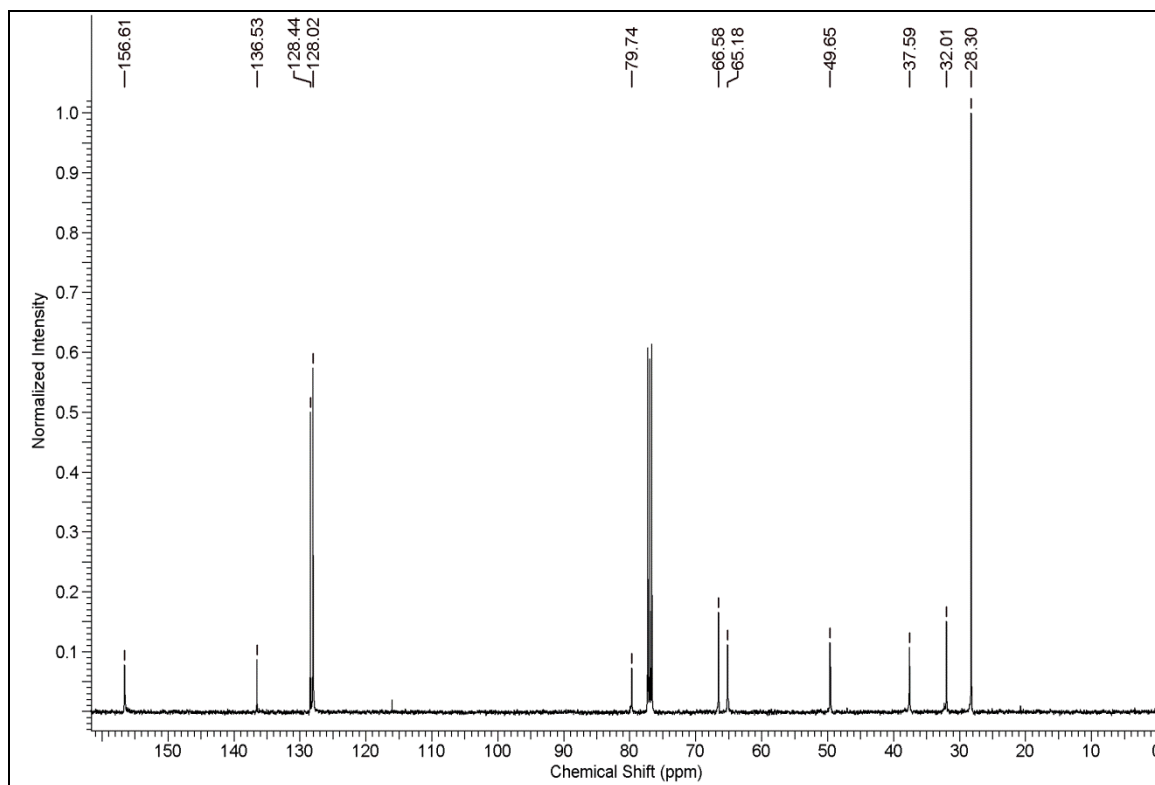
^{1}H NMR of Compound 8
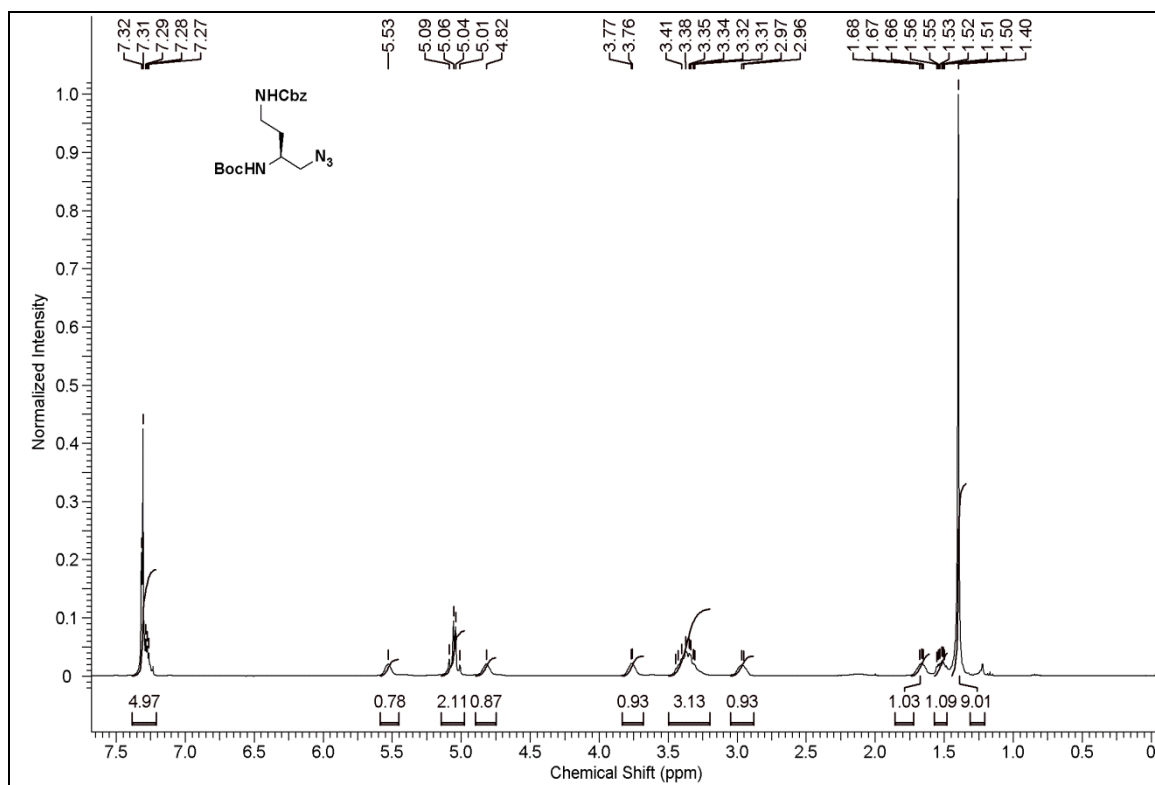
^{13}C NMR of Compound 8
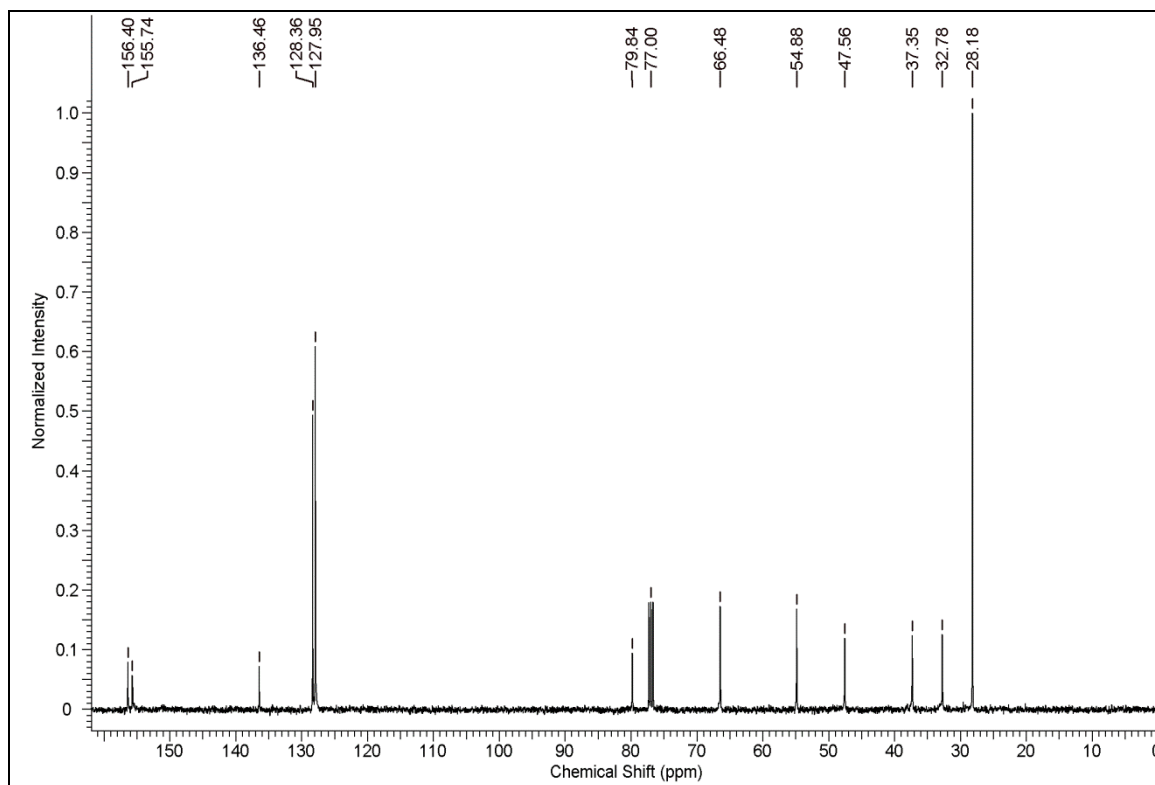
^{1}H NMR of Compound 9
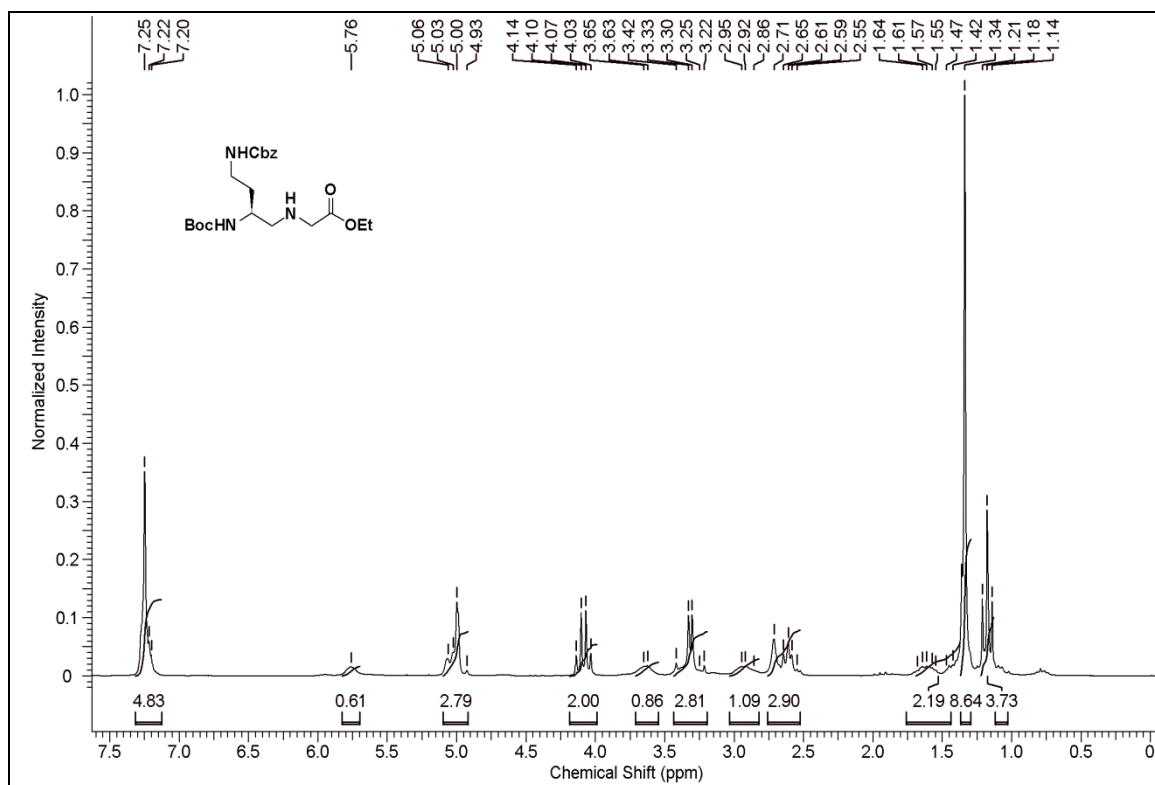
^{13}C NMR of Compound 9
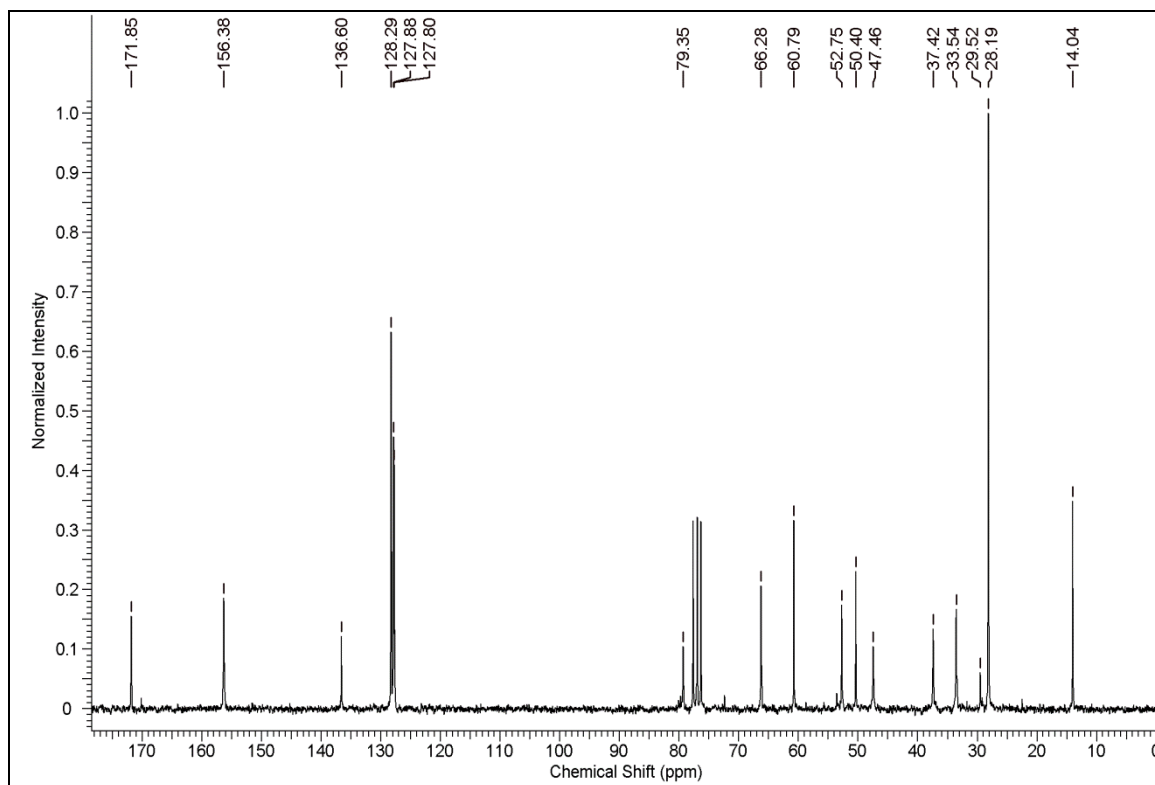
^{1}H NMR of Compound 10
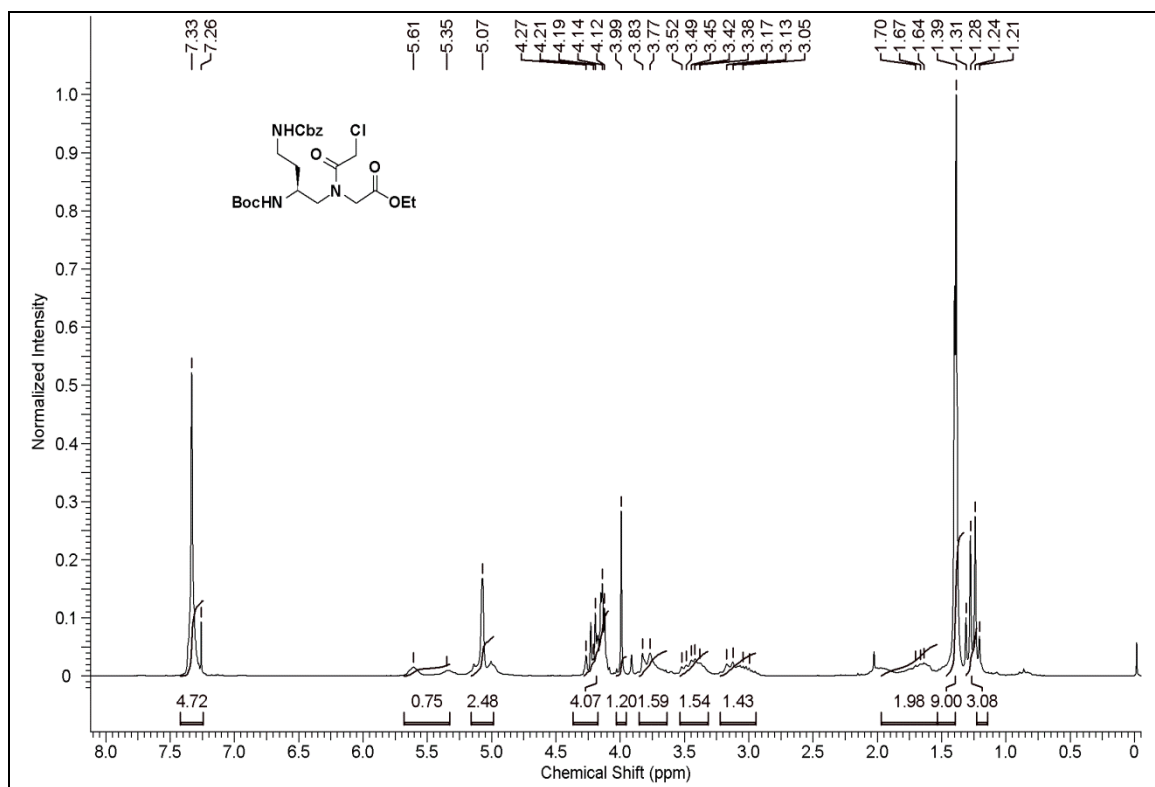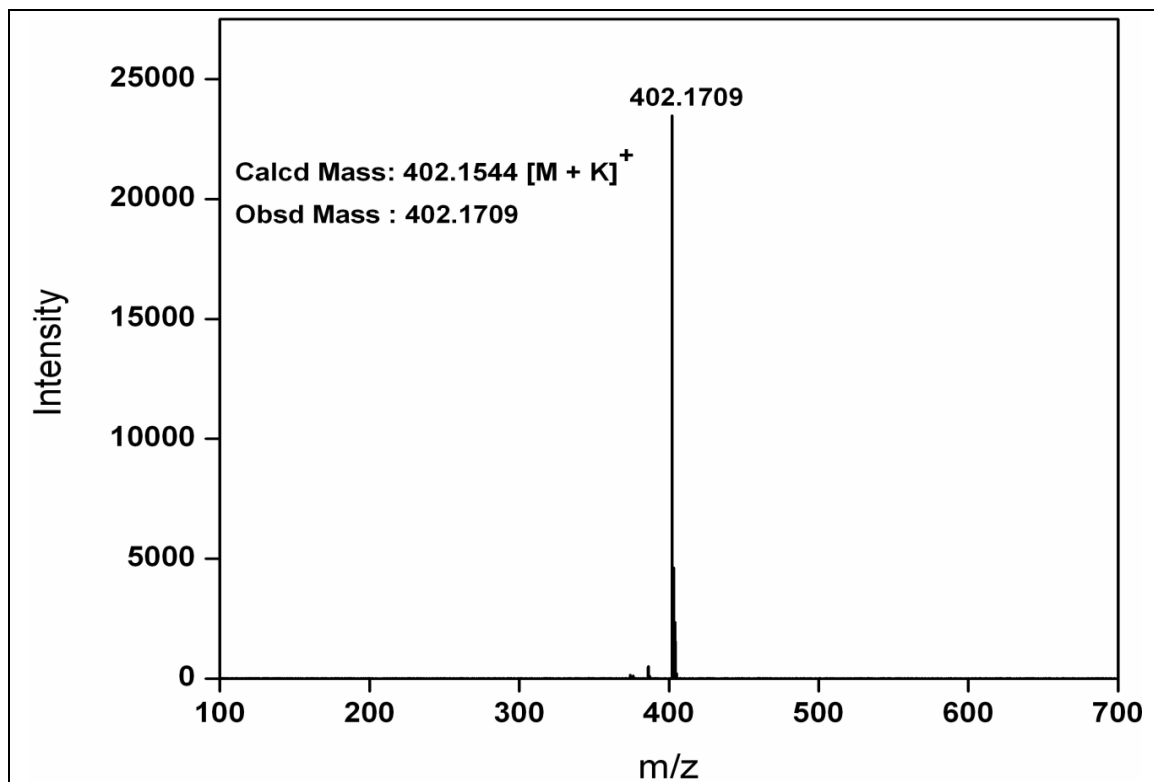

^{13}C NMR of Compound 10
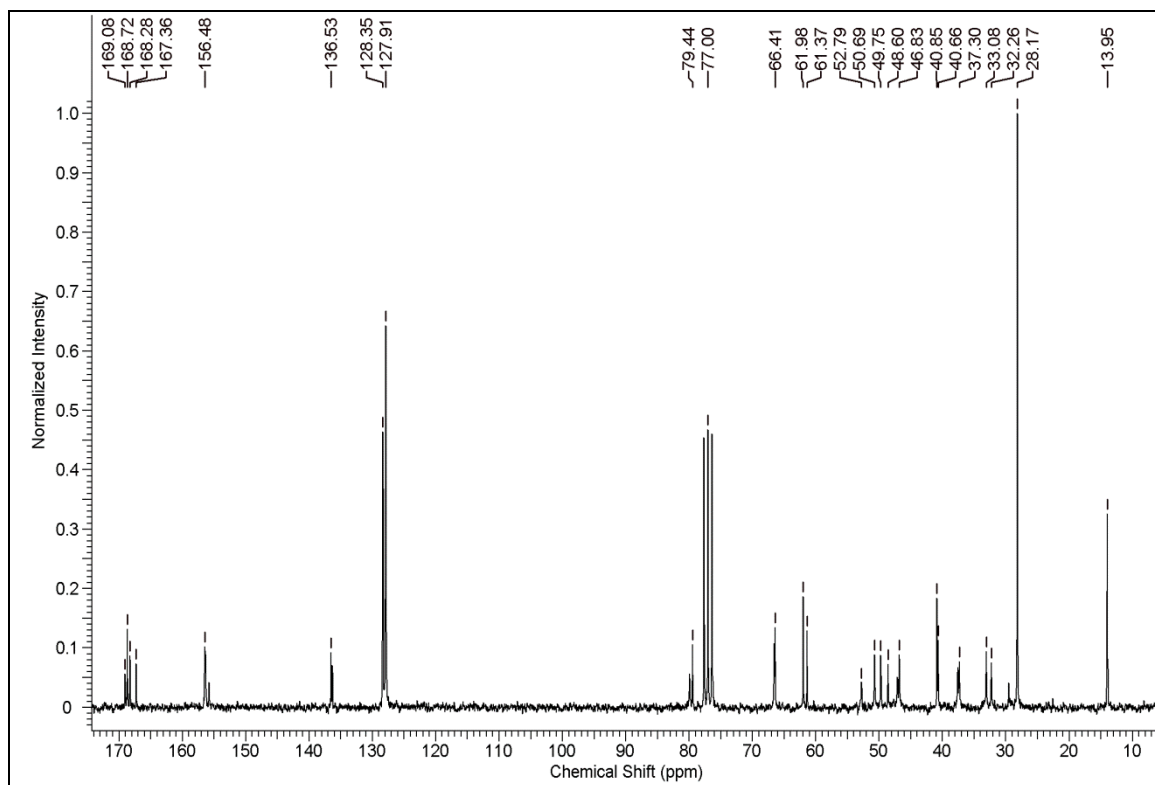
¹H NMR of Compound 11
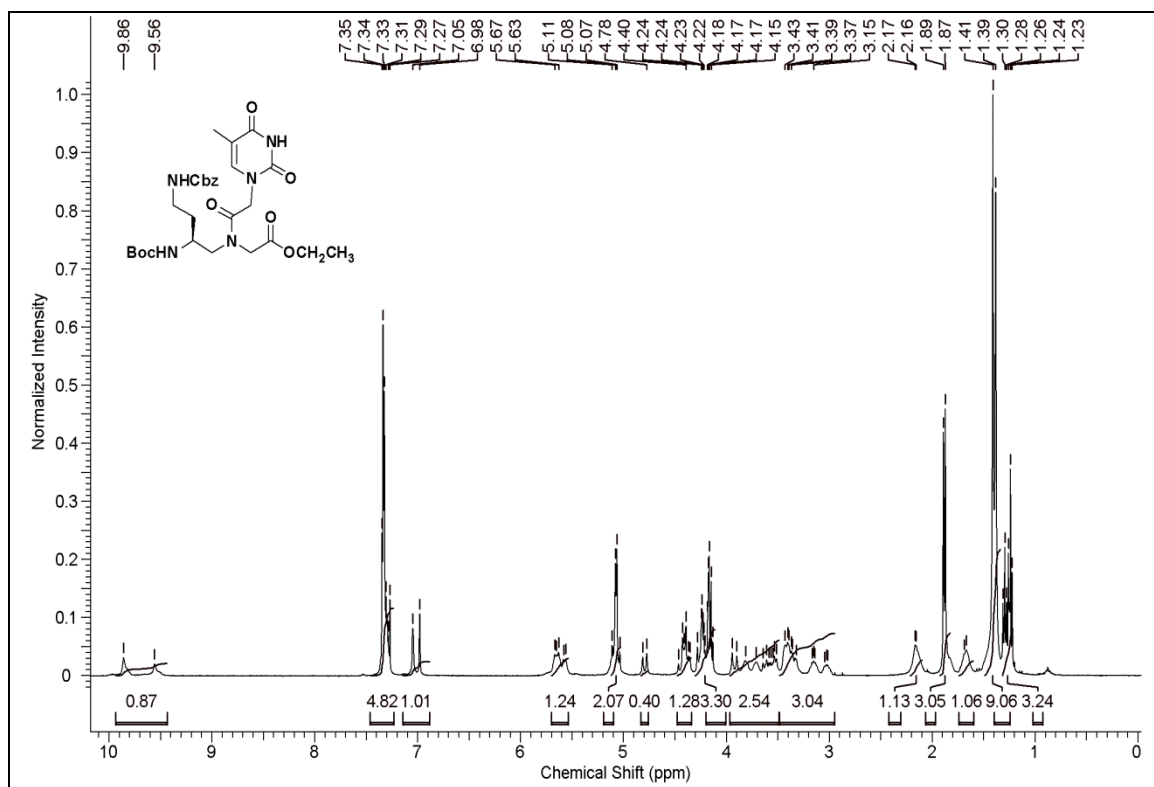
¹³C NMR of Compound 11
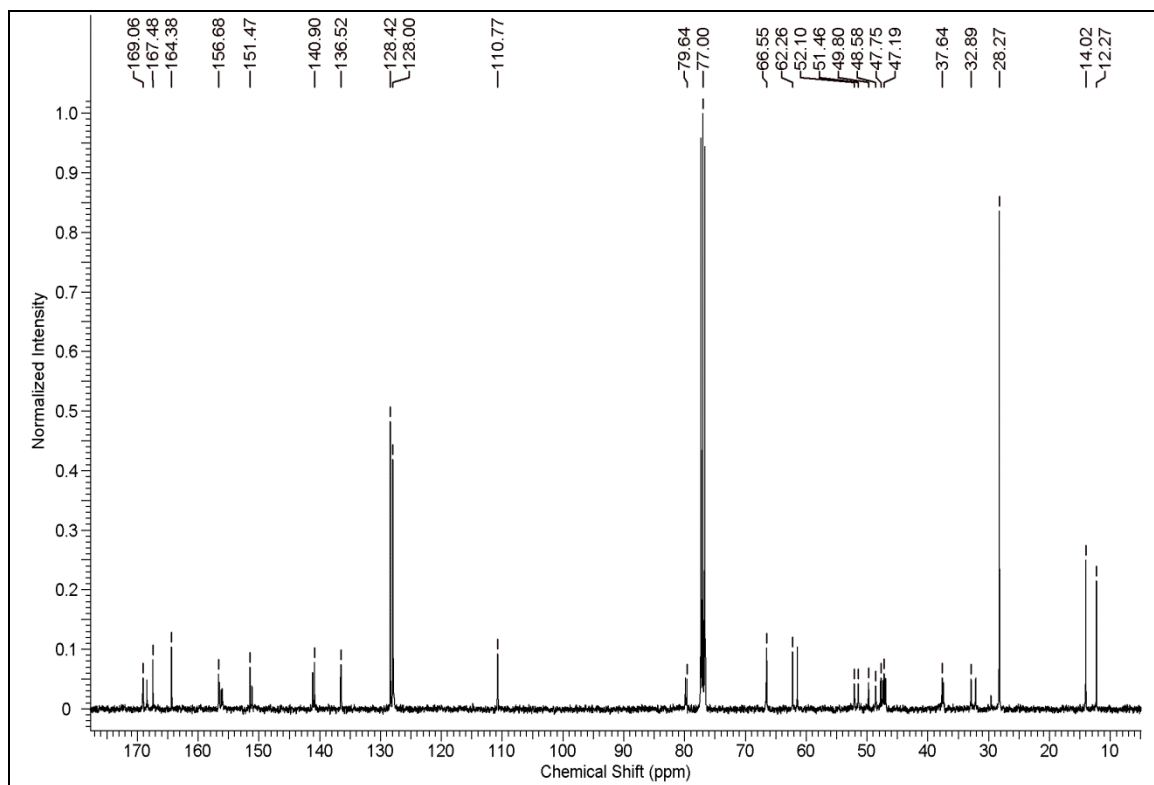
¹H NMR of Compound 12
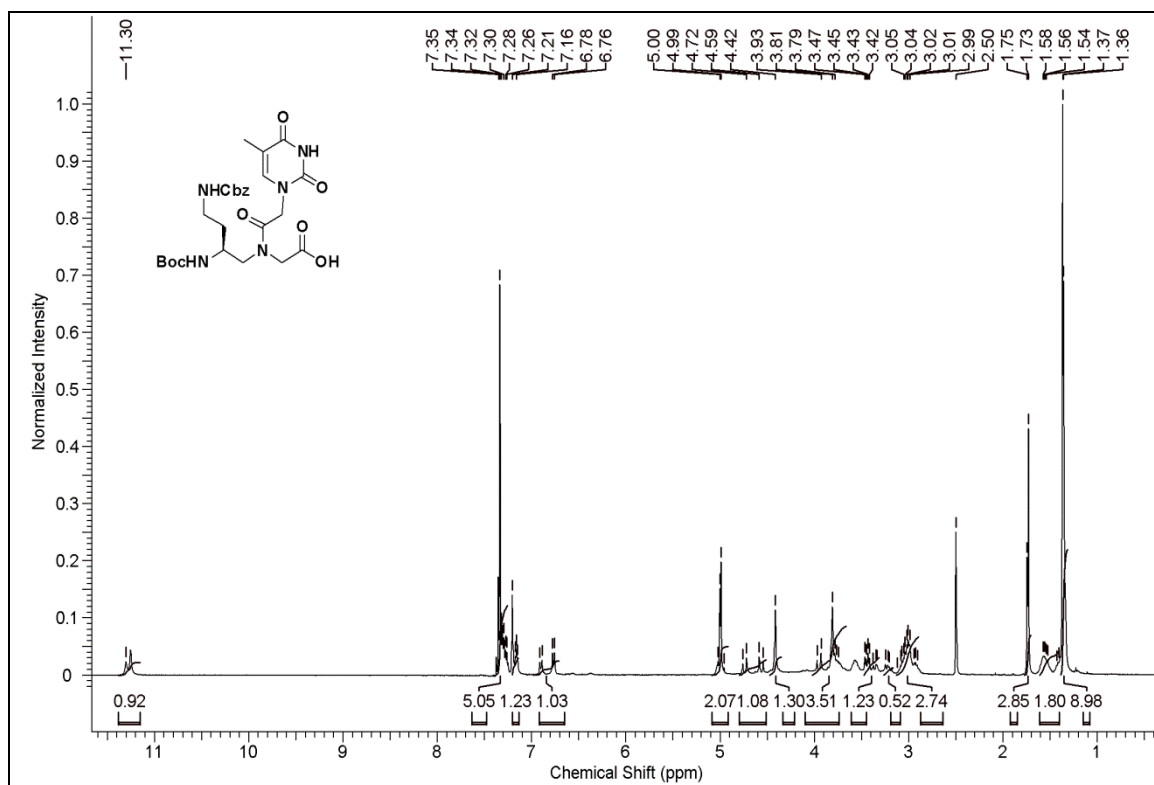
¹³C NMR of Compound 12
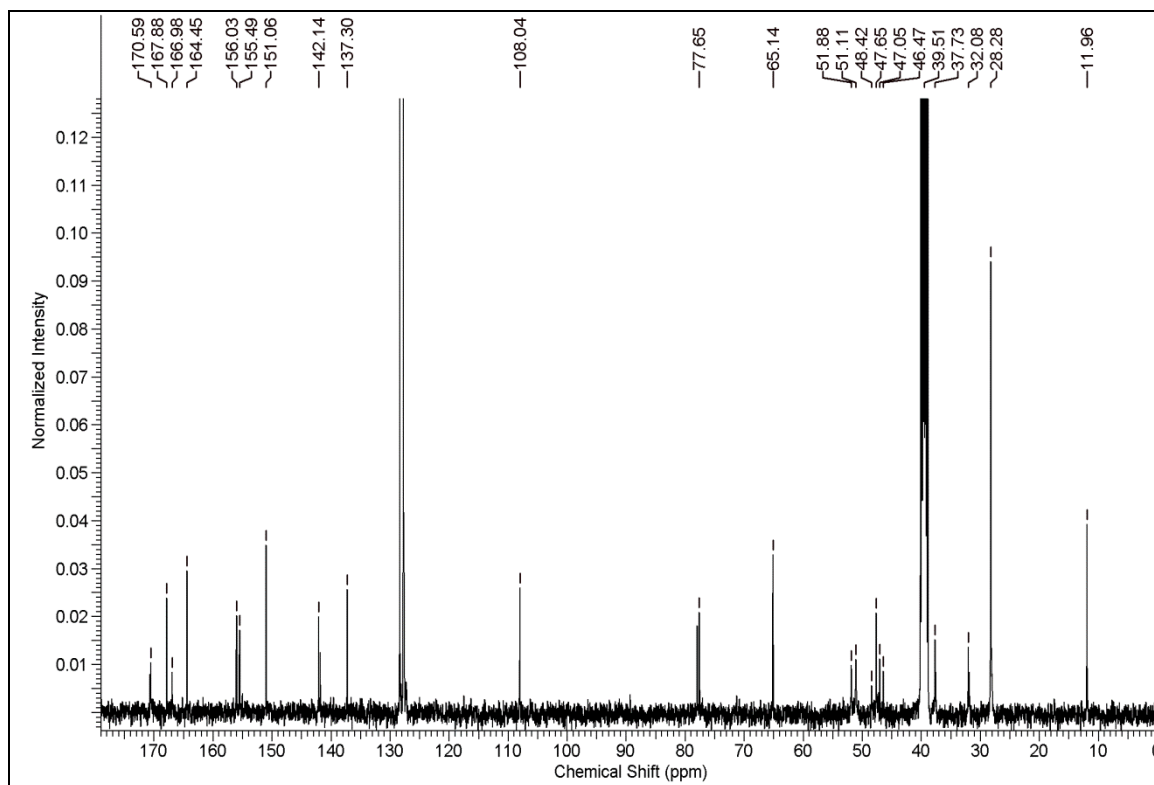
¹H NMR of Compound 13
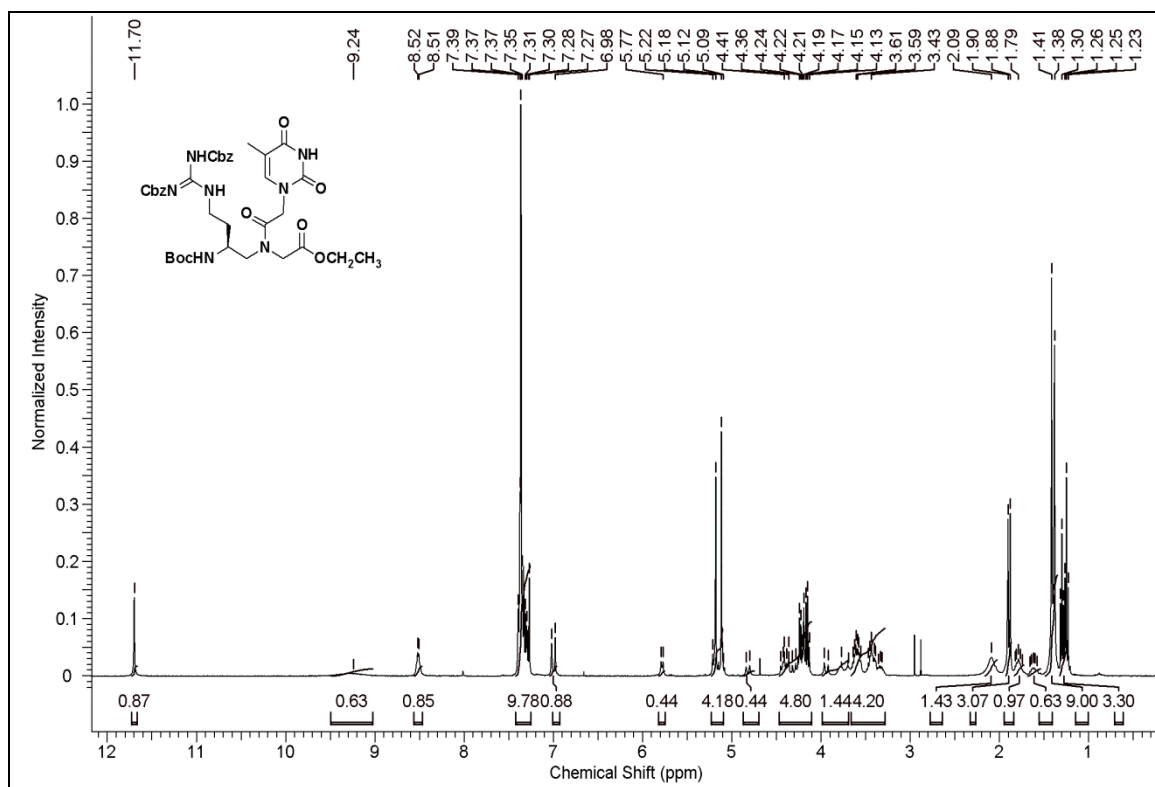
¹³C NMR of Compound 13
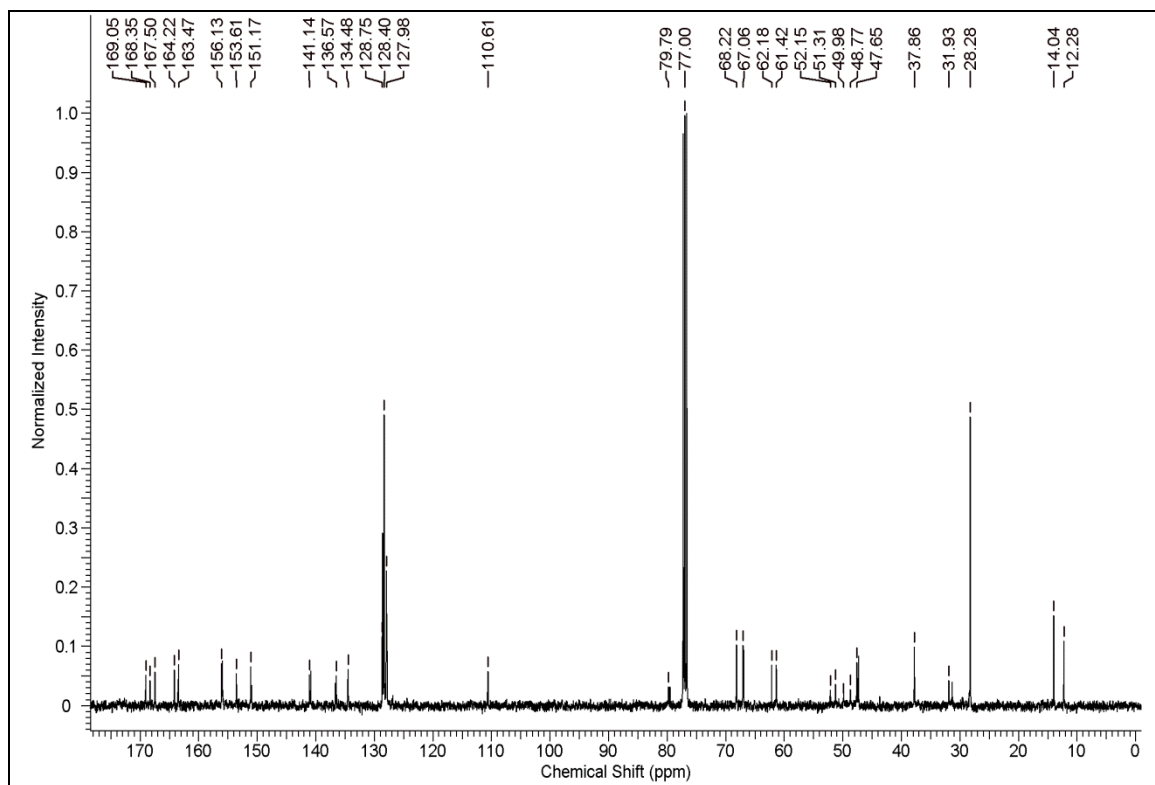
¹H NMR of Compound 14
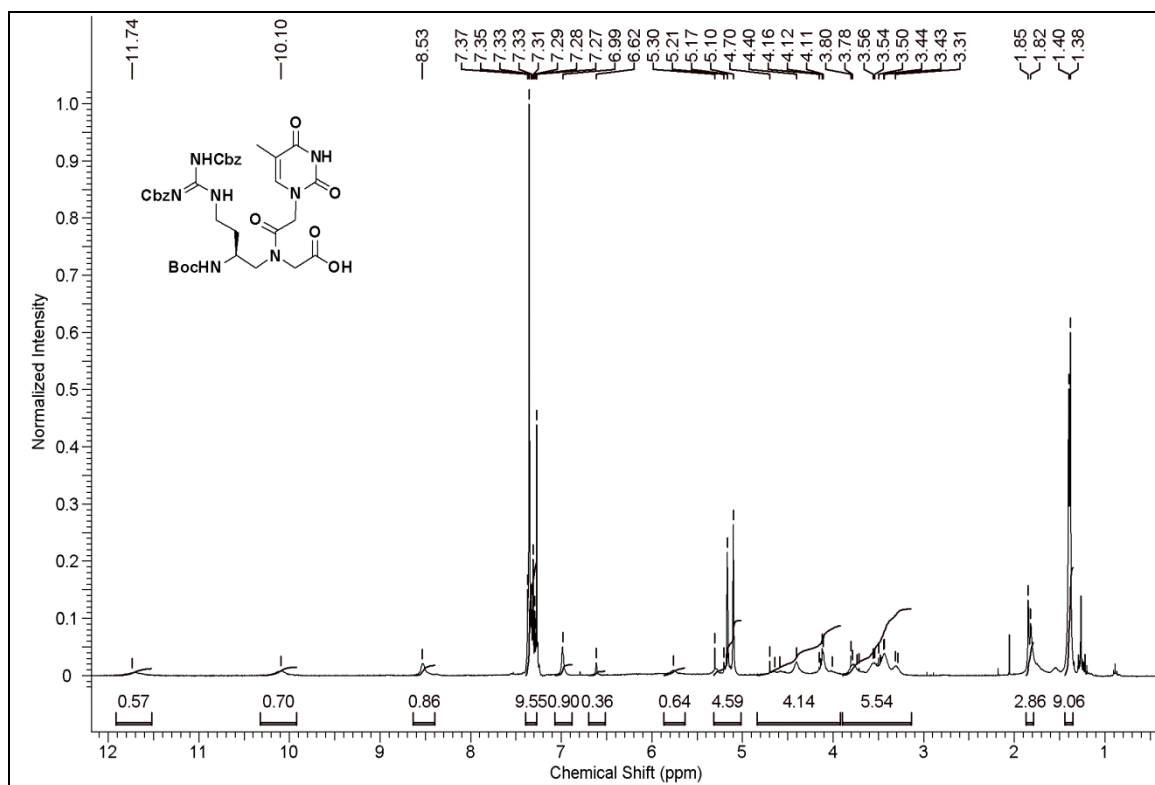
^{1}H NMR of Compound 19
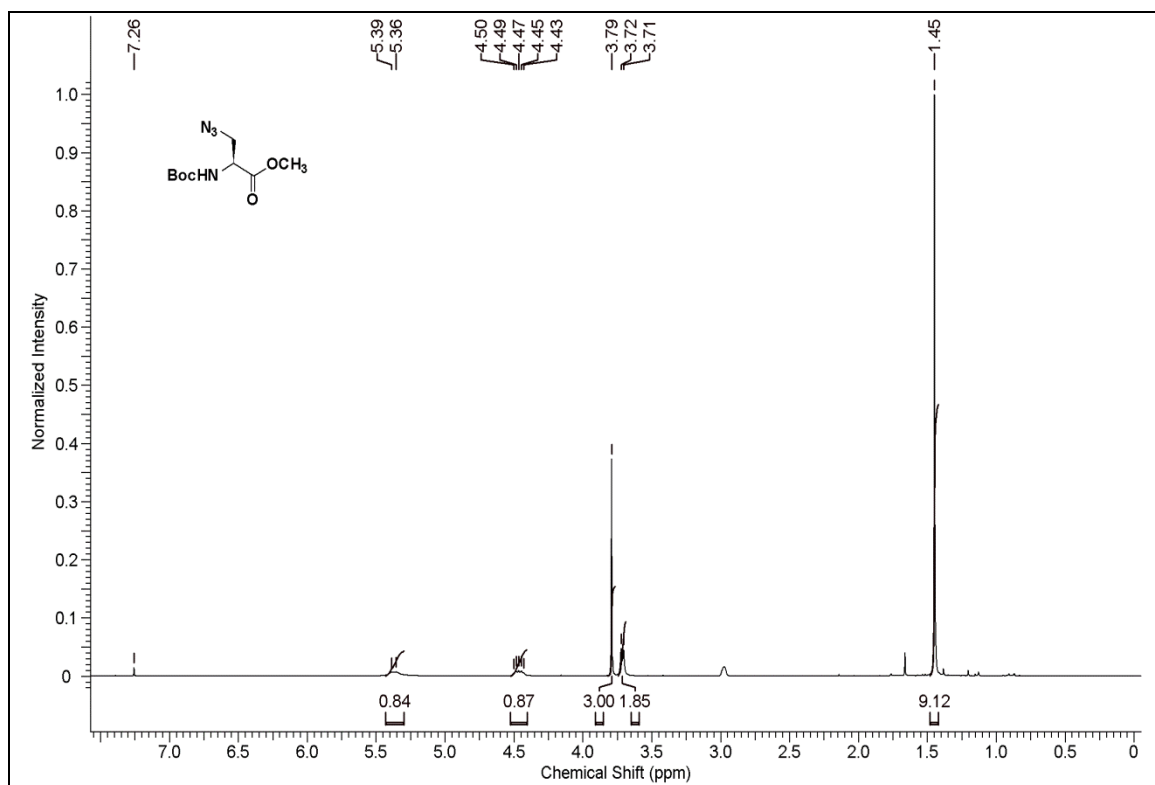
^{13}C NMR of Compound 19
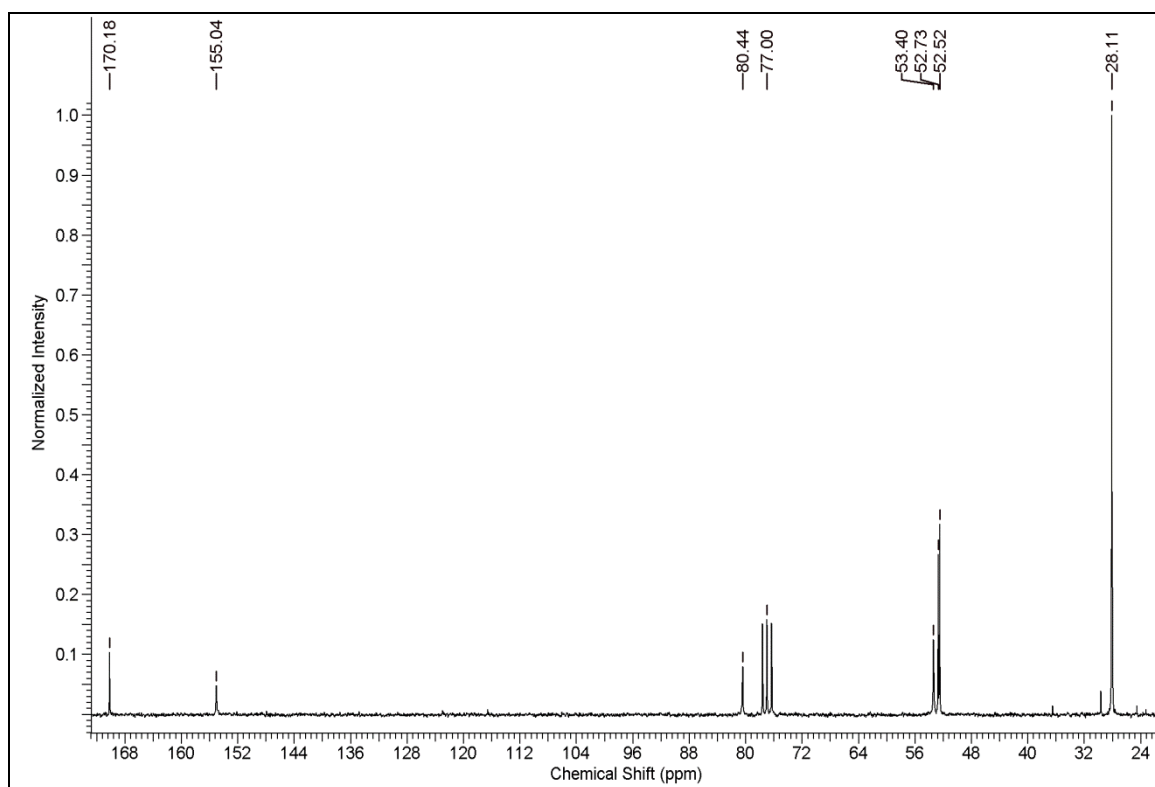
^{1}H NMR of Compound 20
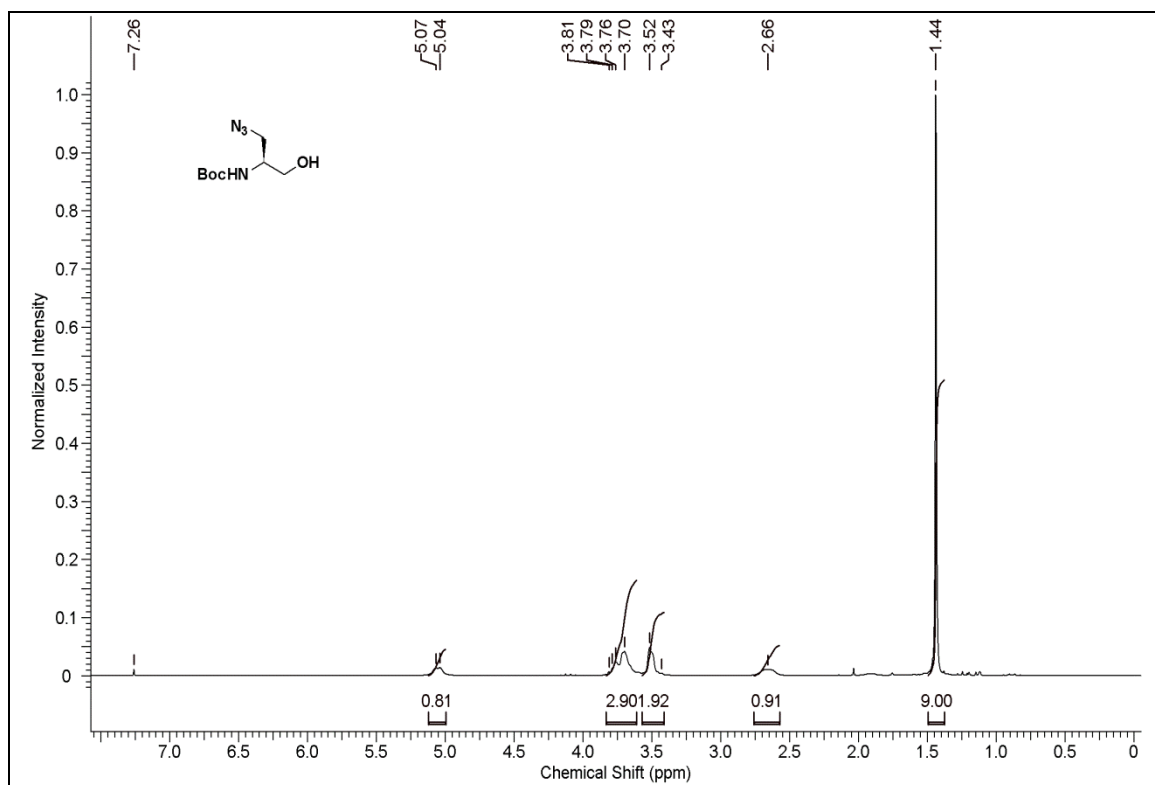
^{13}C NMR of Compound 20
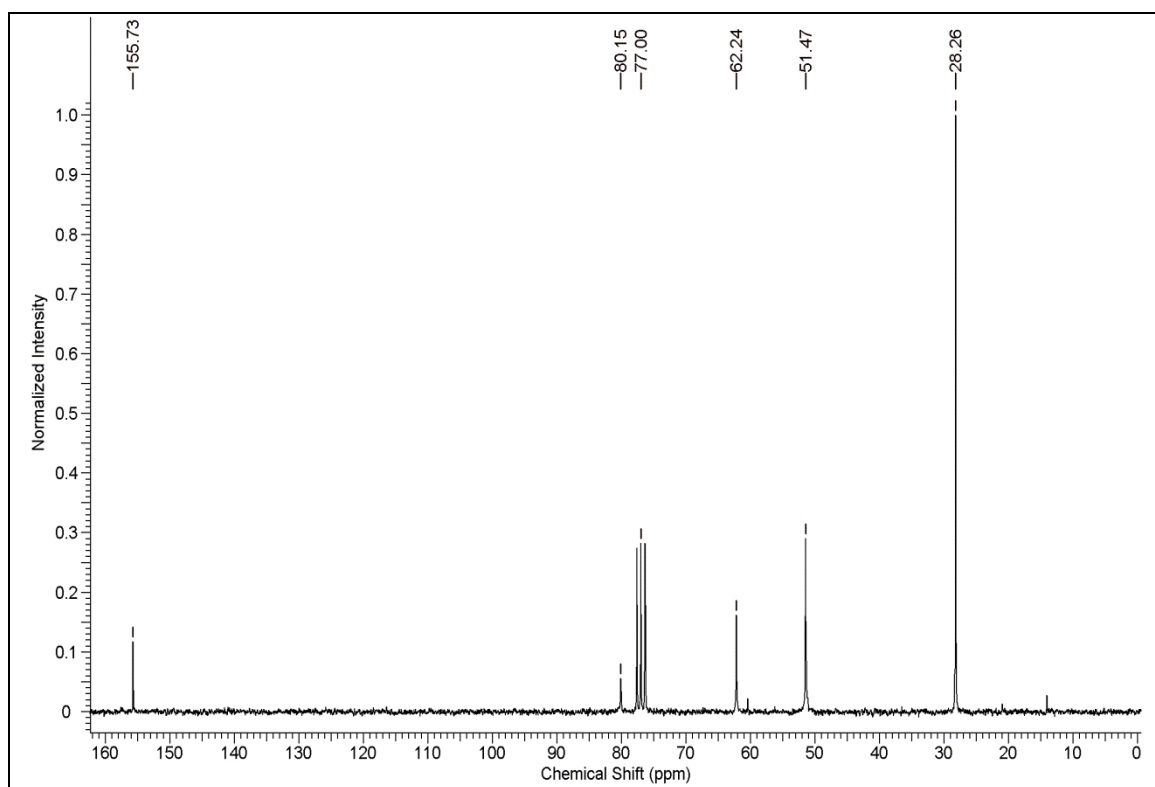
^{1}H NMR of Compound 22
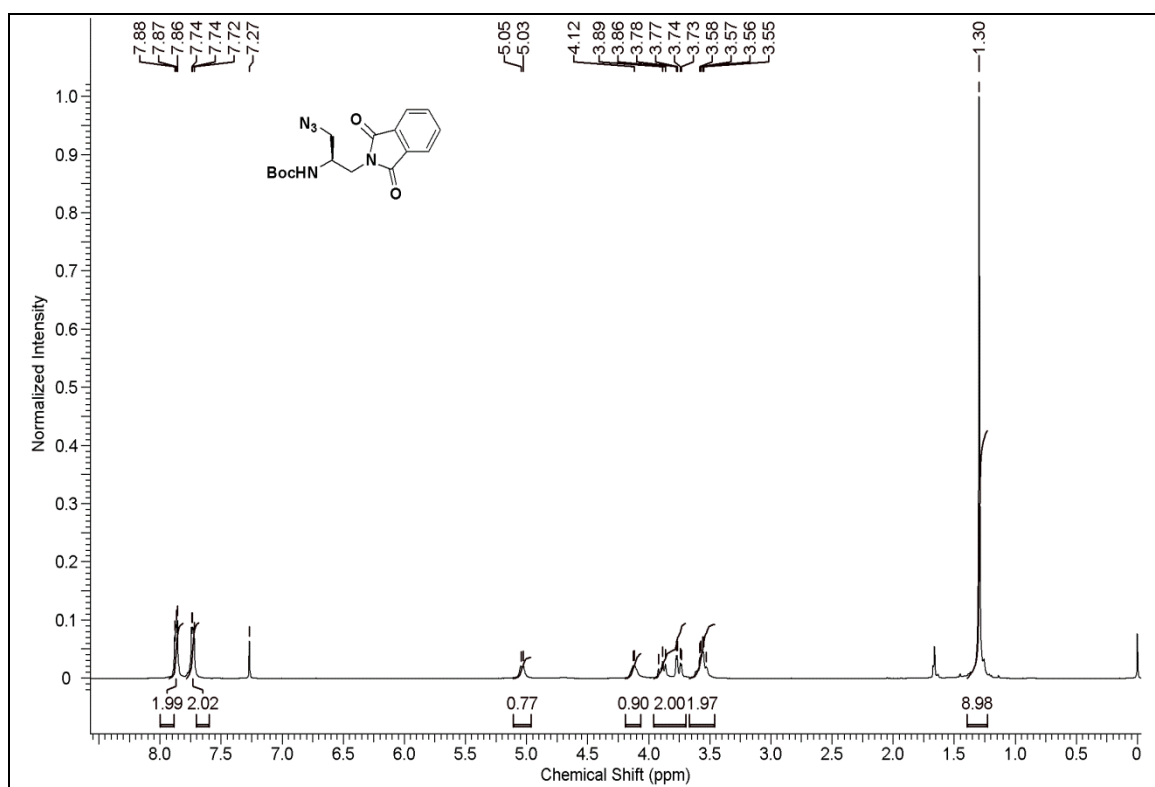
^{13}C NMR of Compound 22
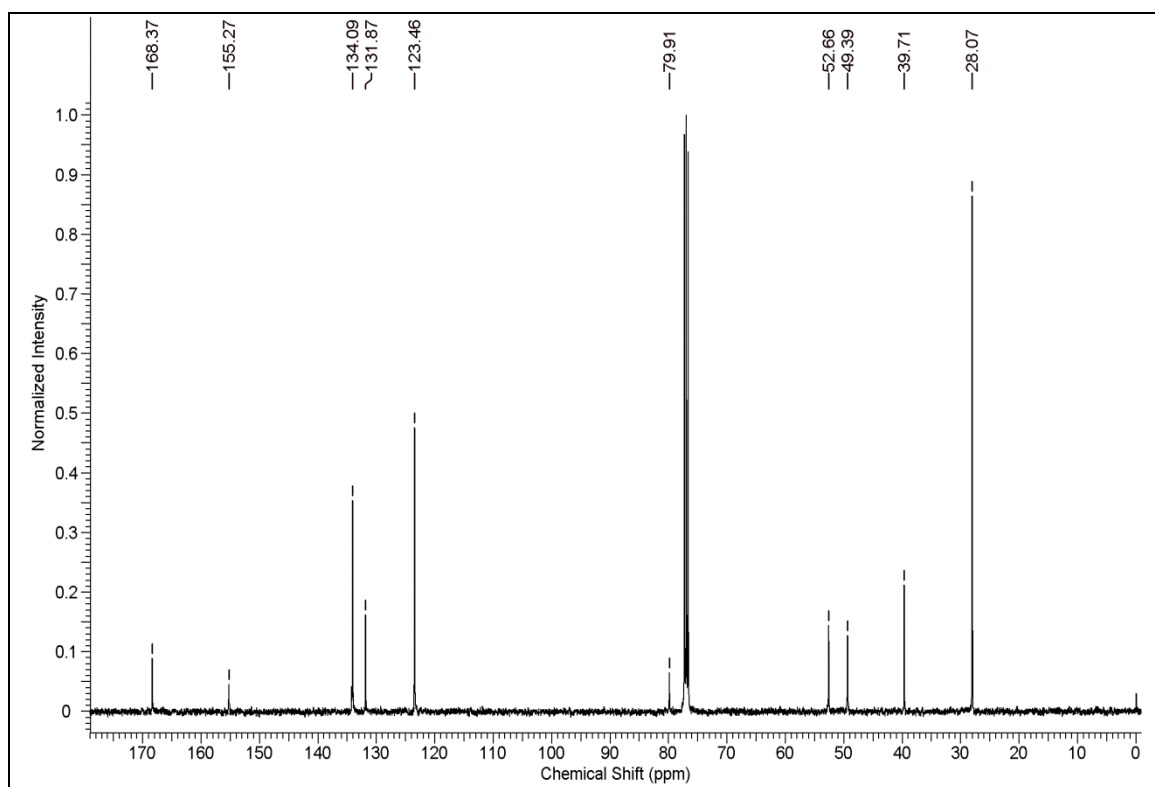
^{1}H NMR of Compound 23
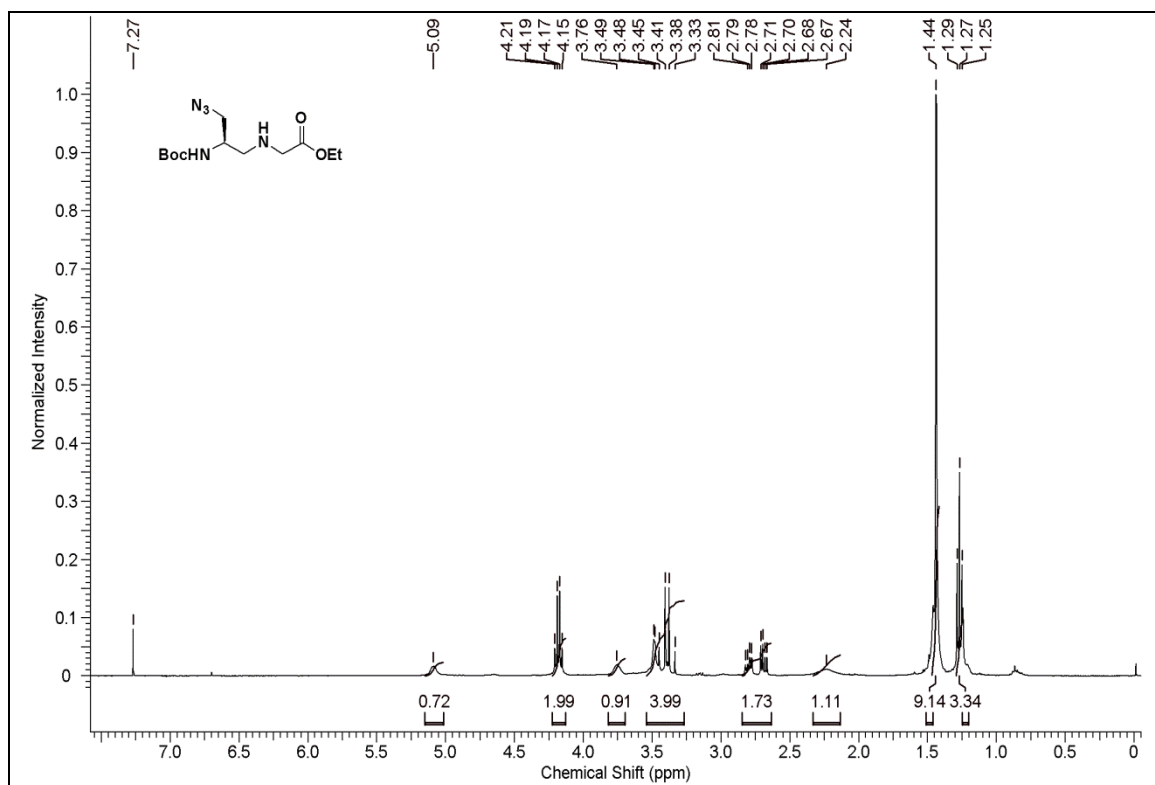
^{13}C NMR of Compound 23
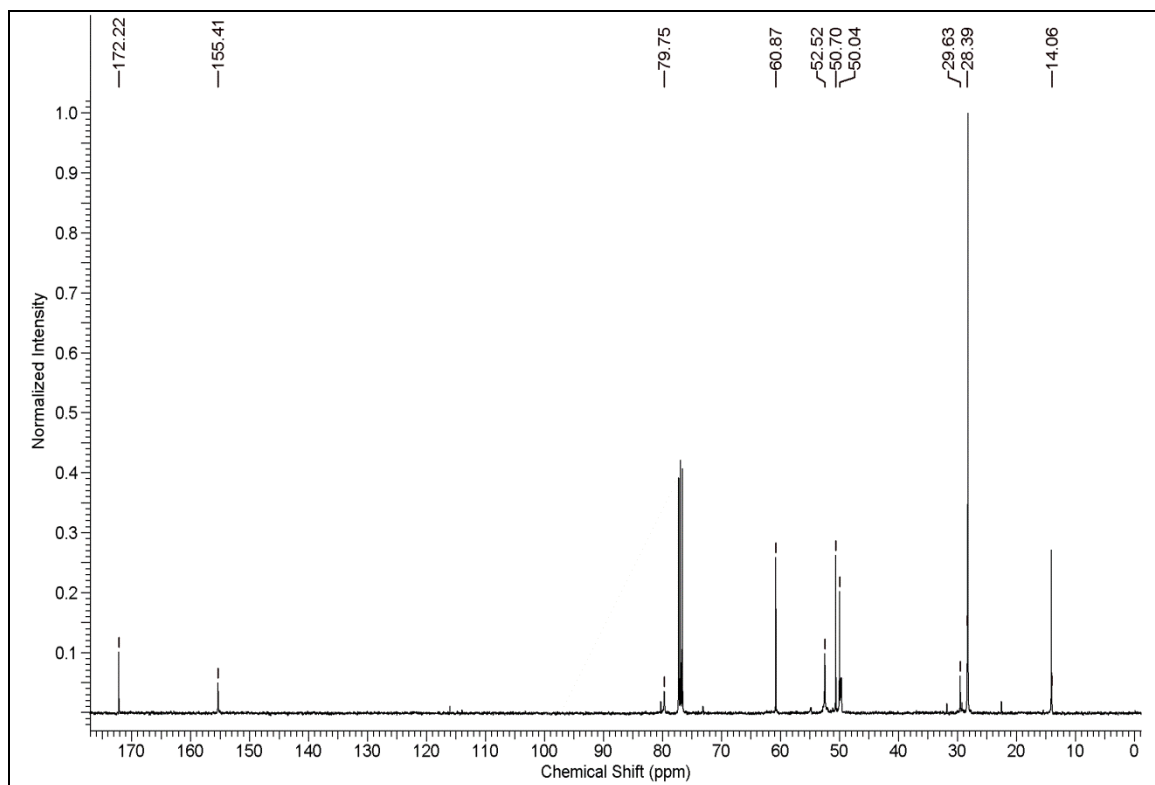
^{1}H NMR of Compound 24
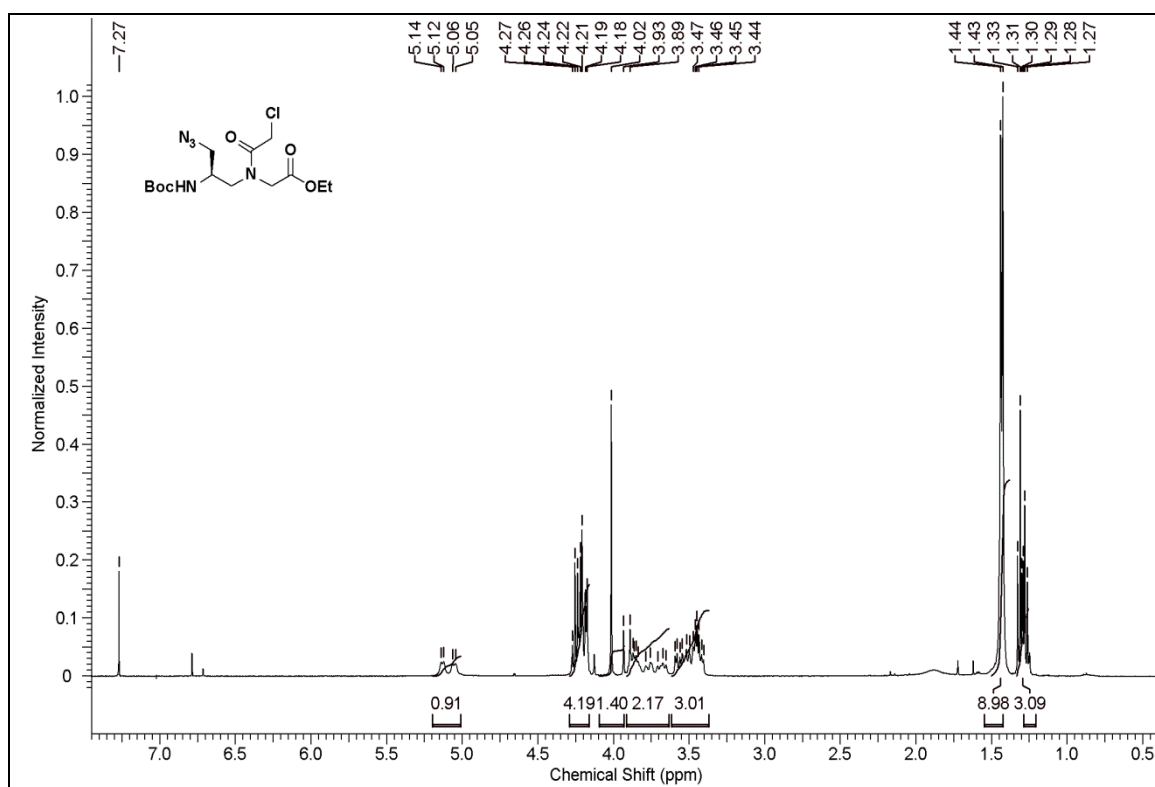
^{13}C NMR of Compound 24
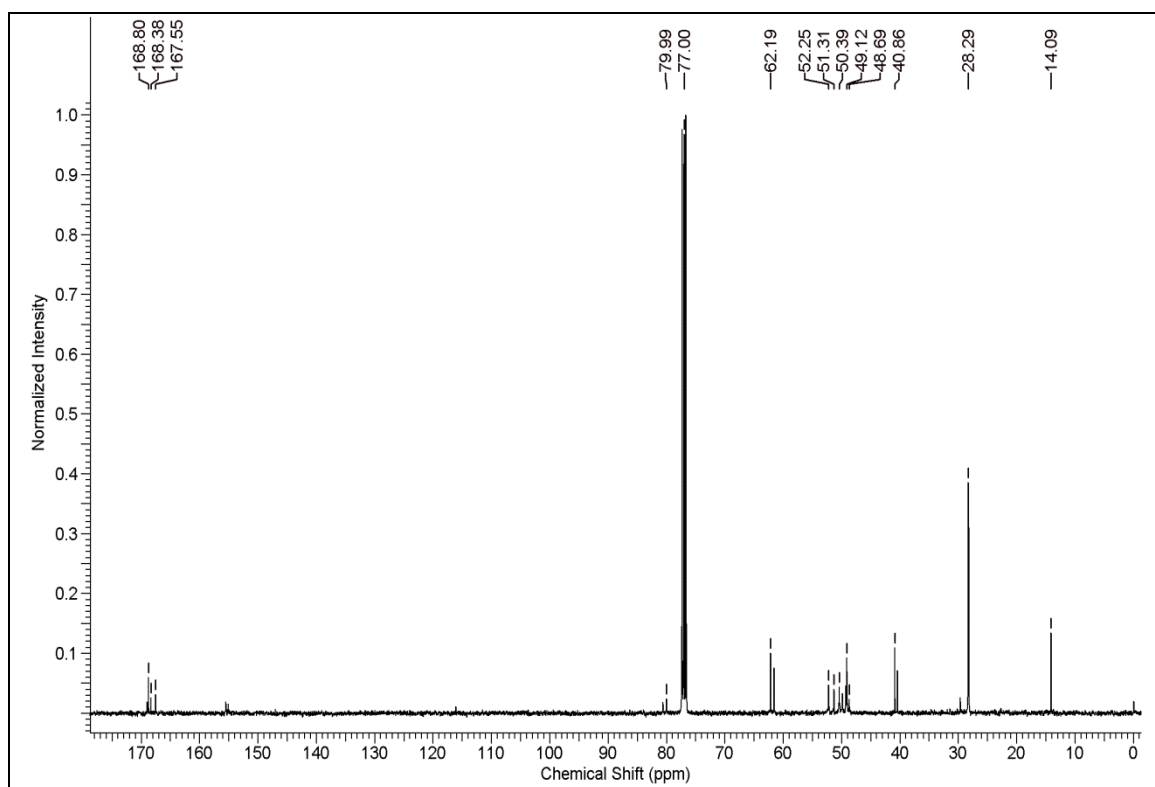
^{1}H NMR of Compound 25
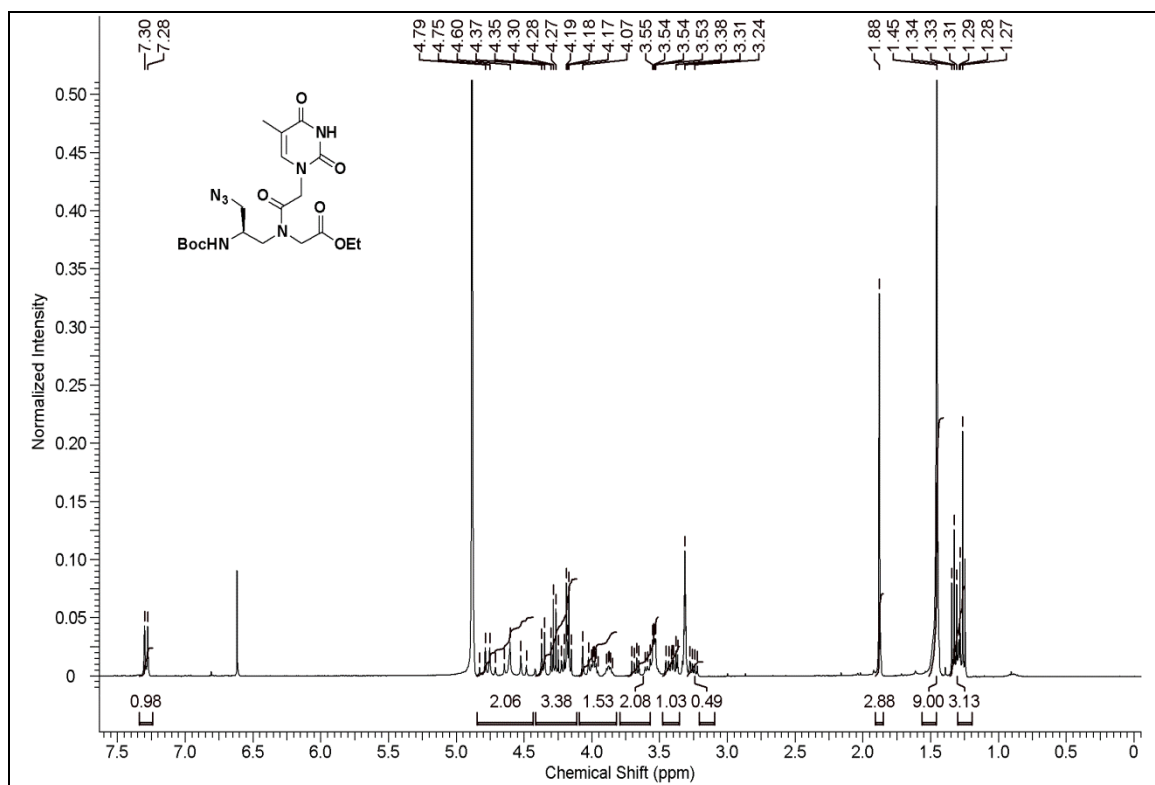
^{13}C NMR of Compound 25
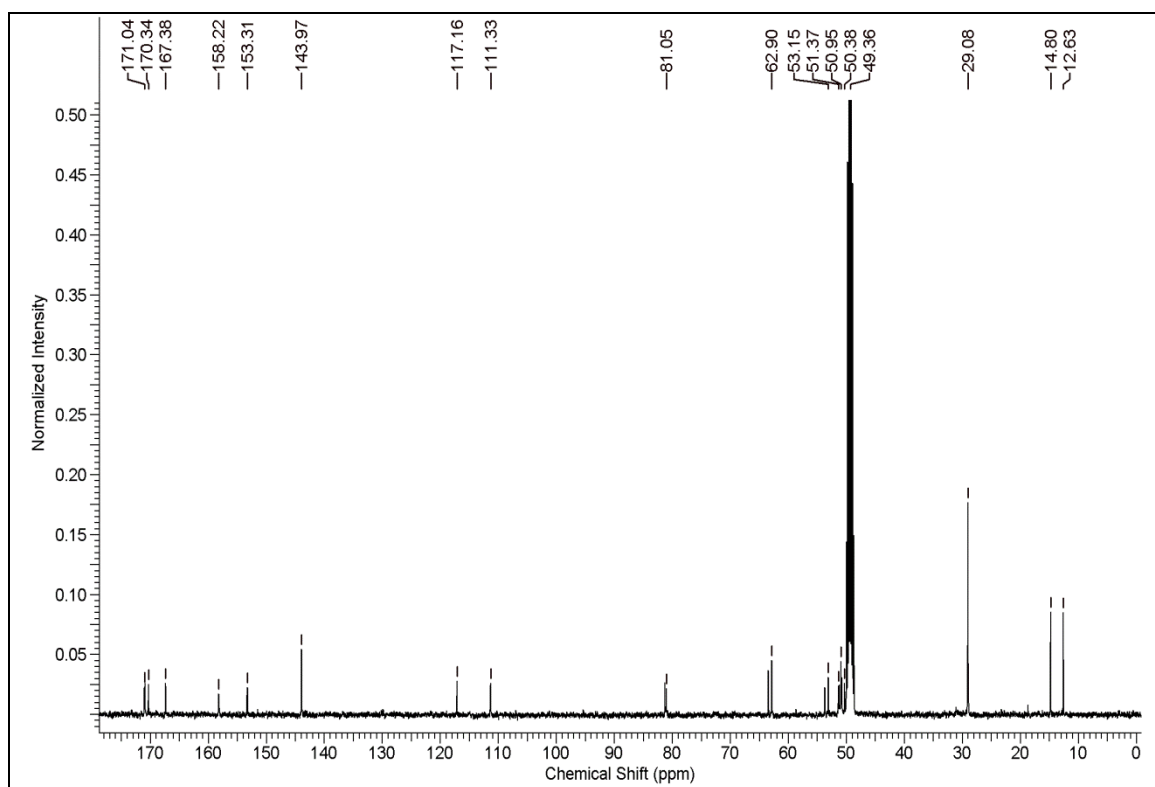
^{1}H NMR of Compound 26
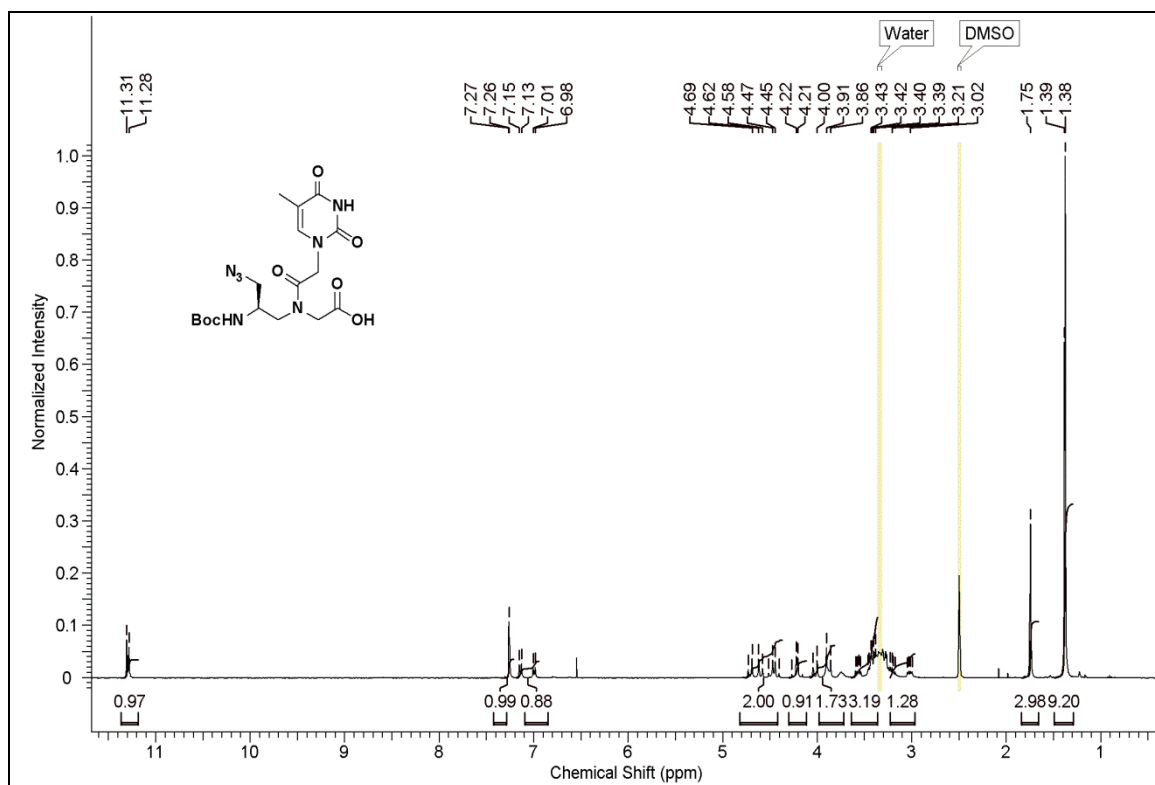
^{13}C NMR of Compound 26
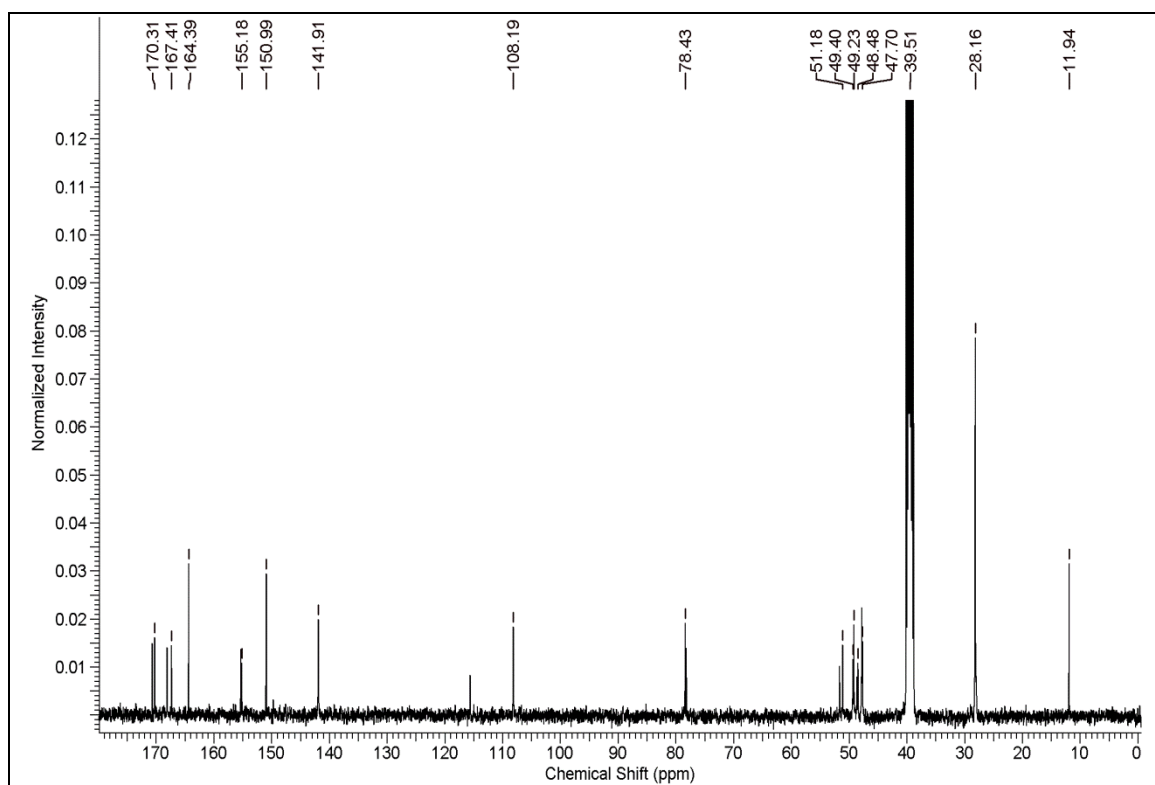
^{1}H NMR of Compound 29
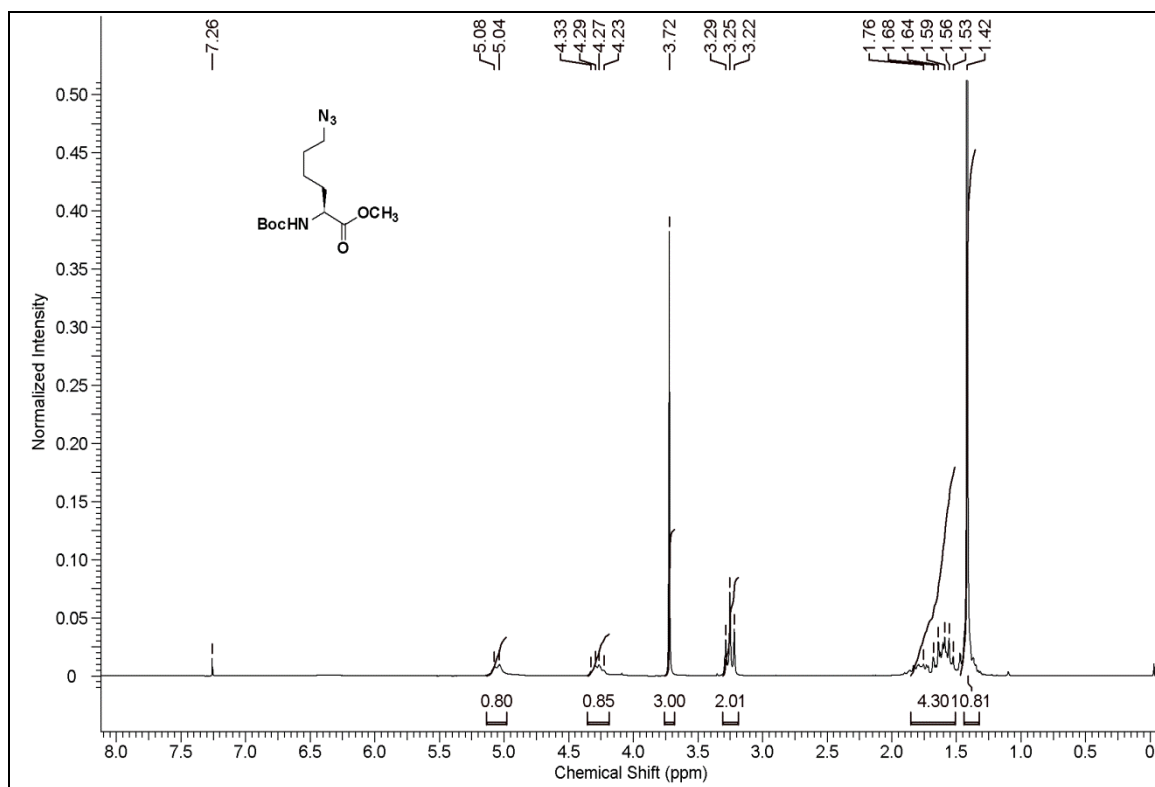
^{13}C NMR of Compound 29
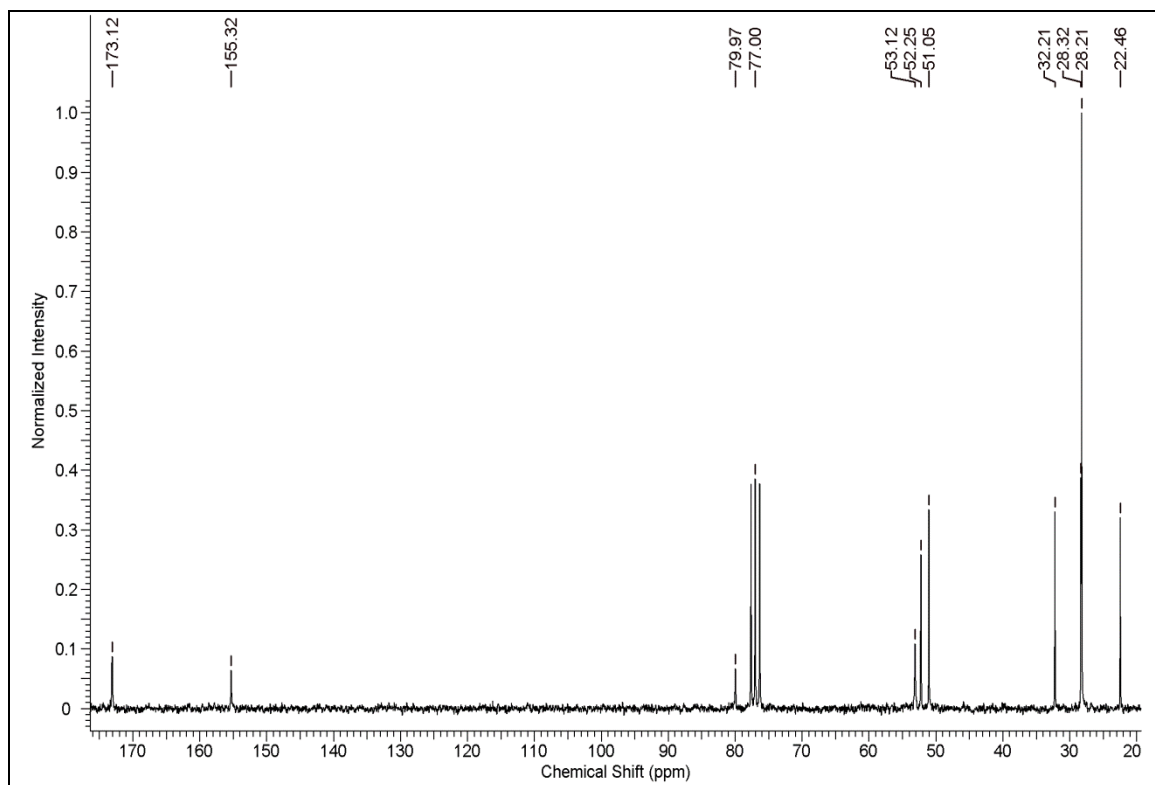
^{1}H NMR of Compound 30
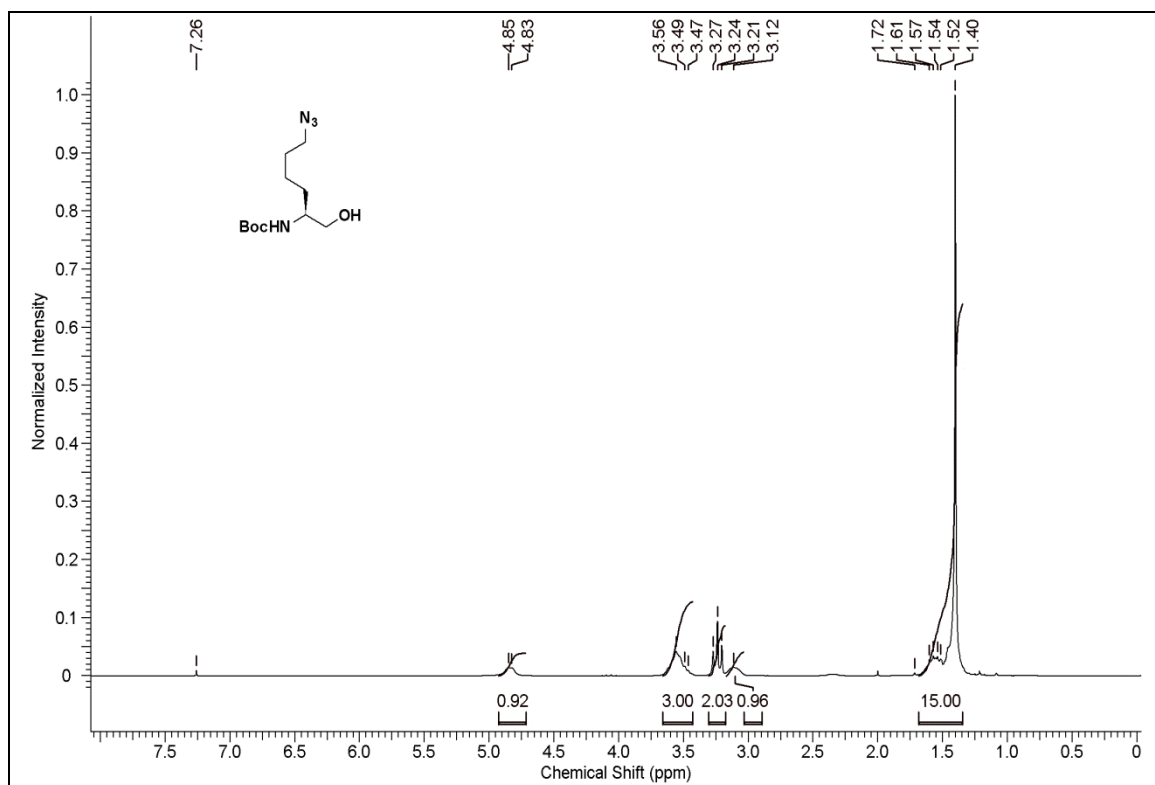
^{13}C NMR of Compound 30
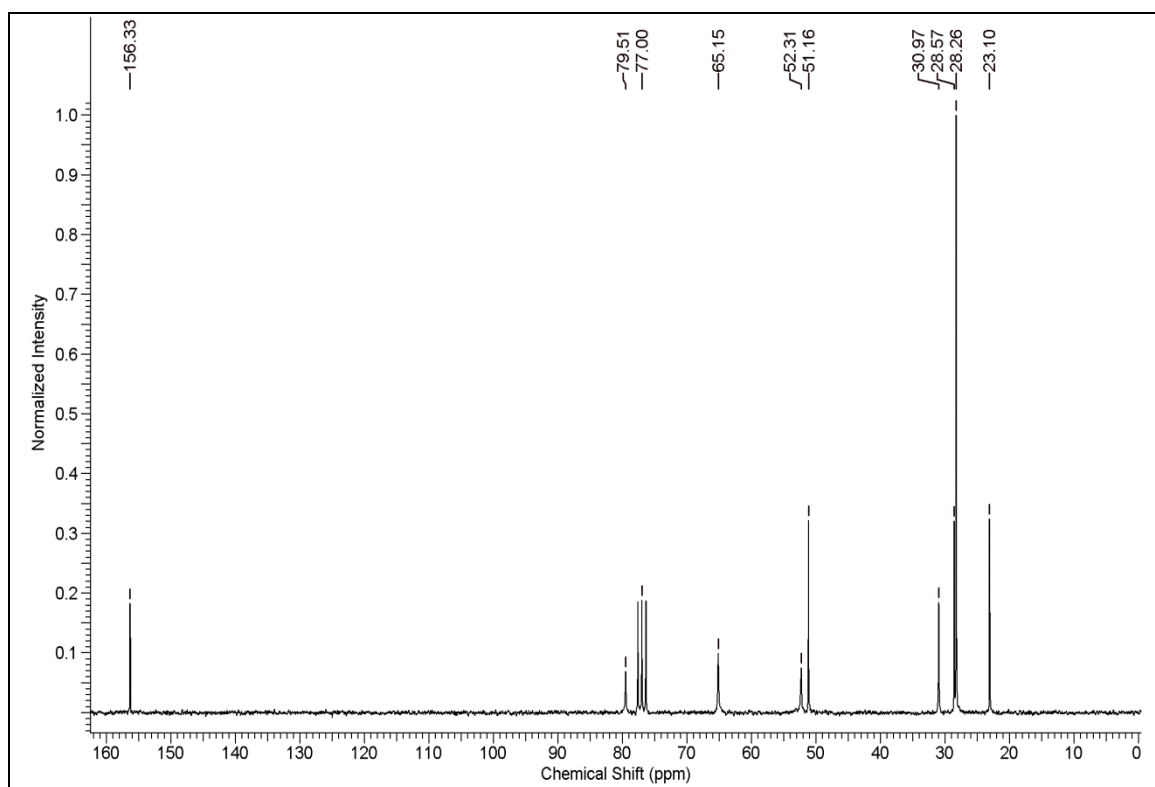
¹H NMR of Compound 32
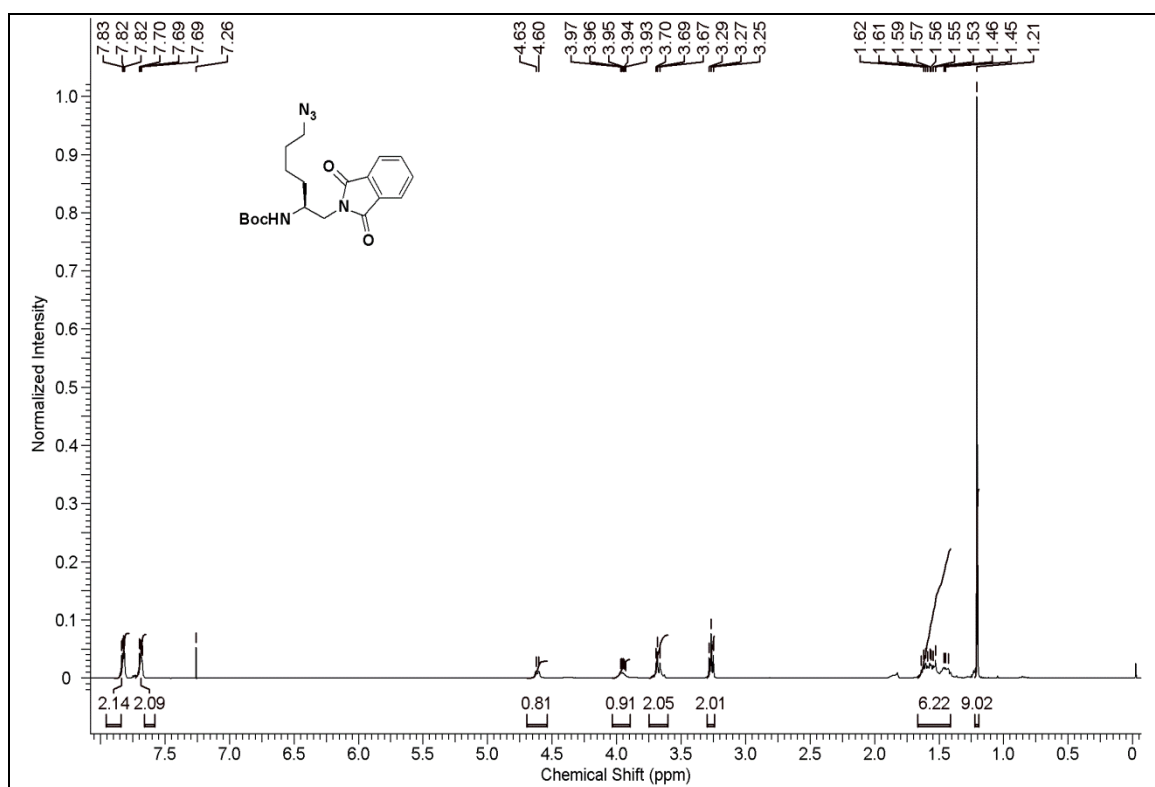
¹³C NMR of Compound 32
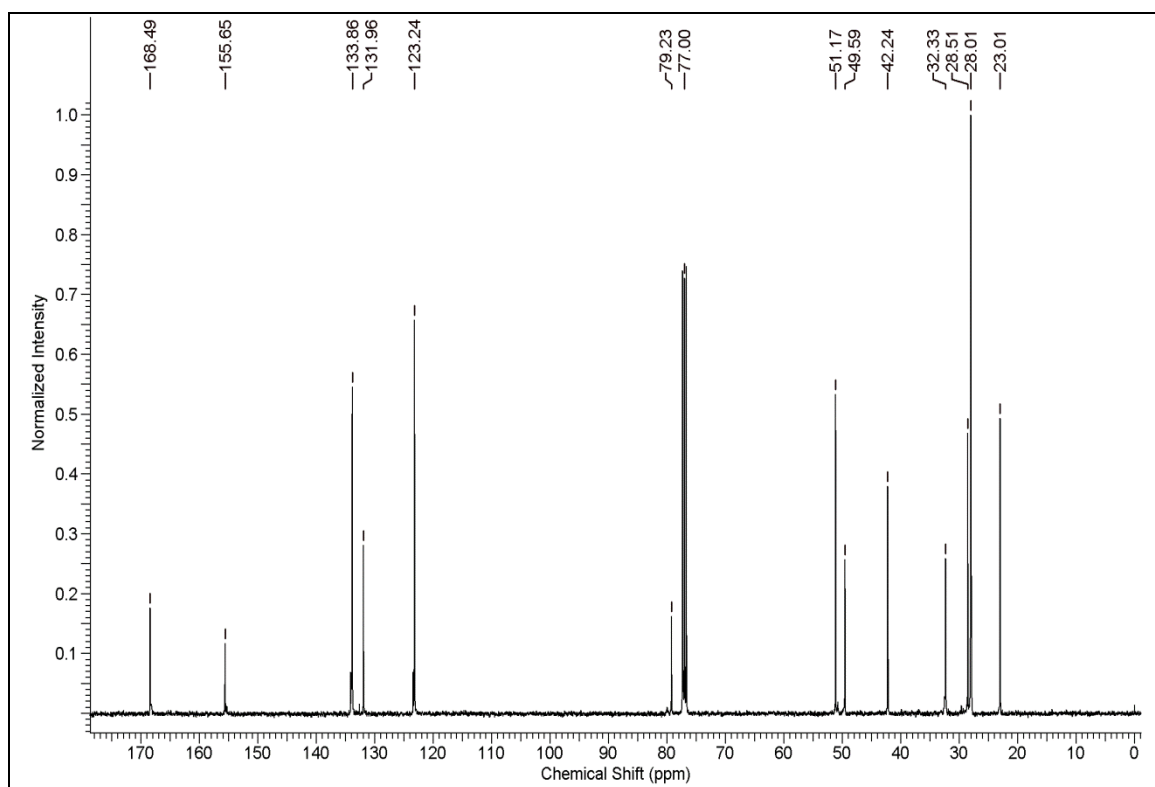
¹H NMR of Compound 33
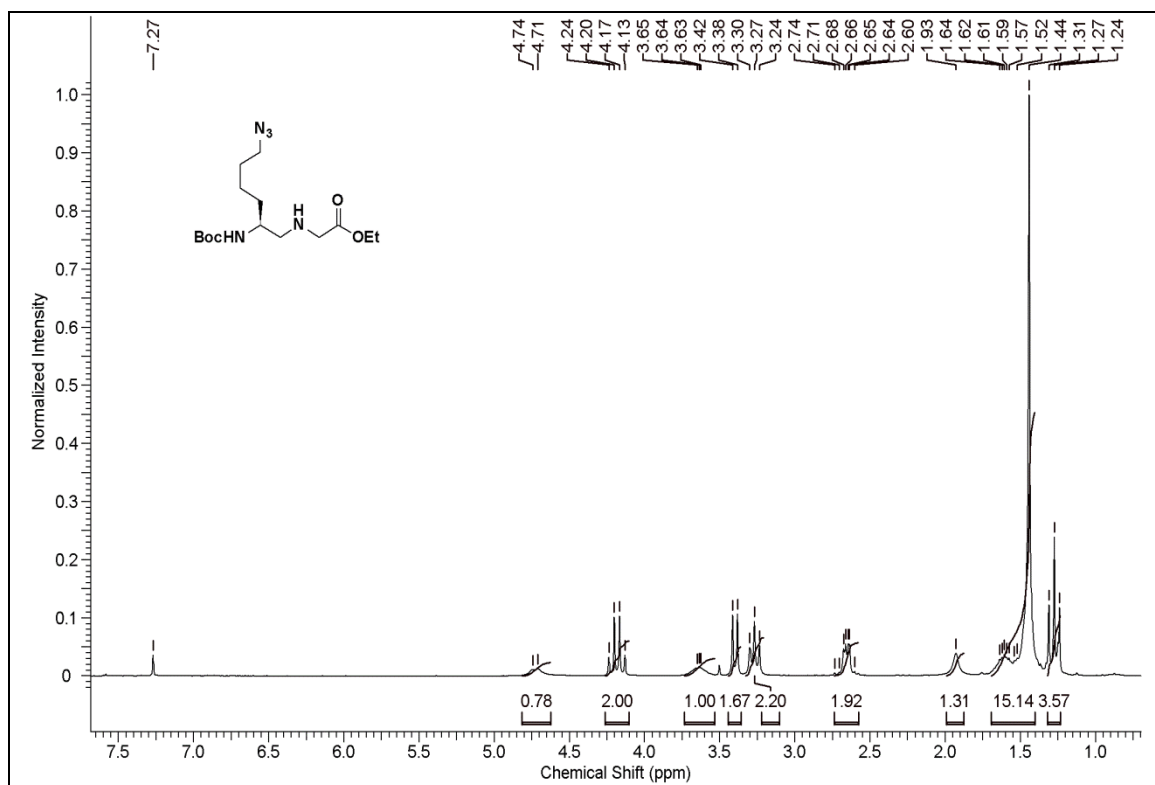
¹³C NMR of Compound 33
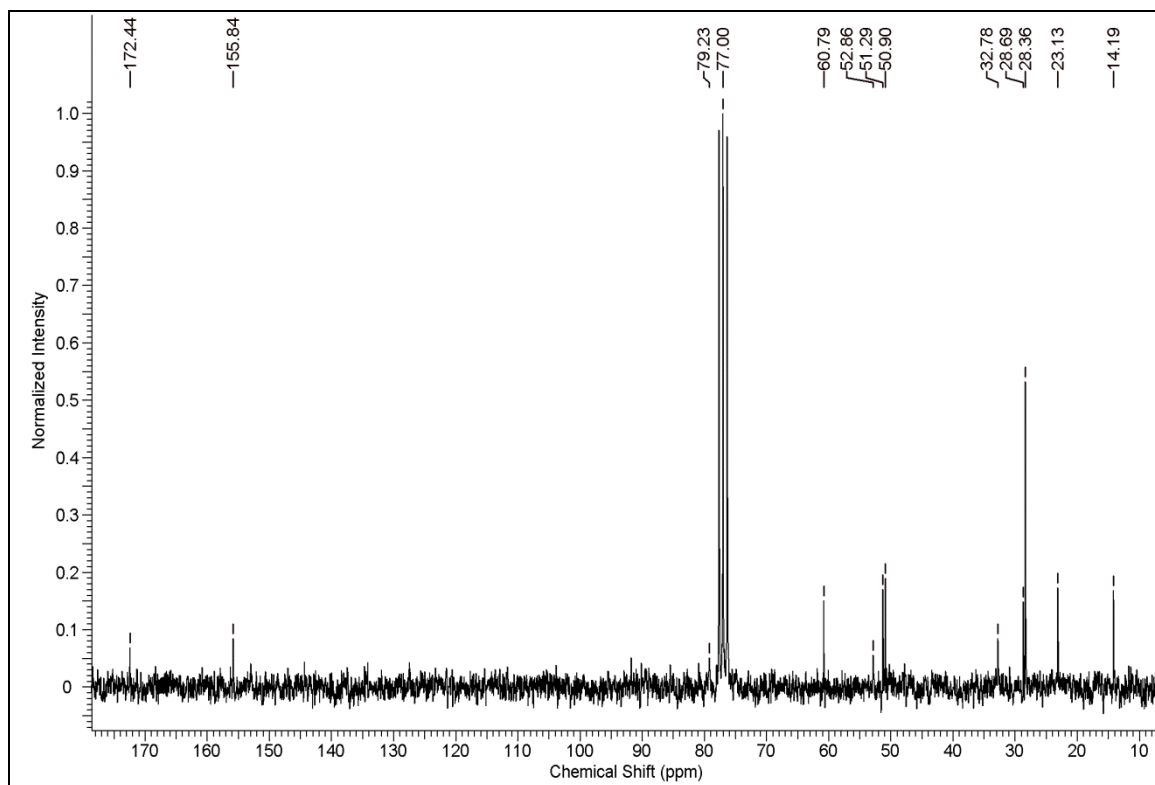
¹H NMR of Compound 34
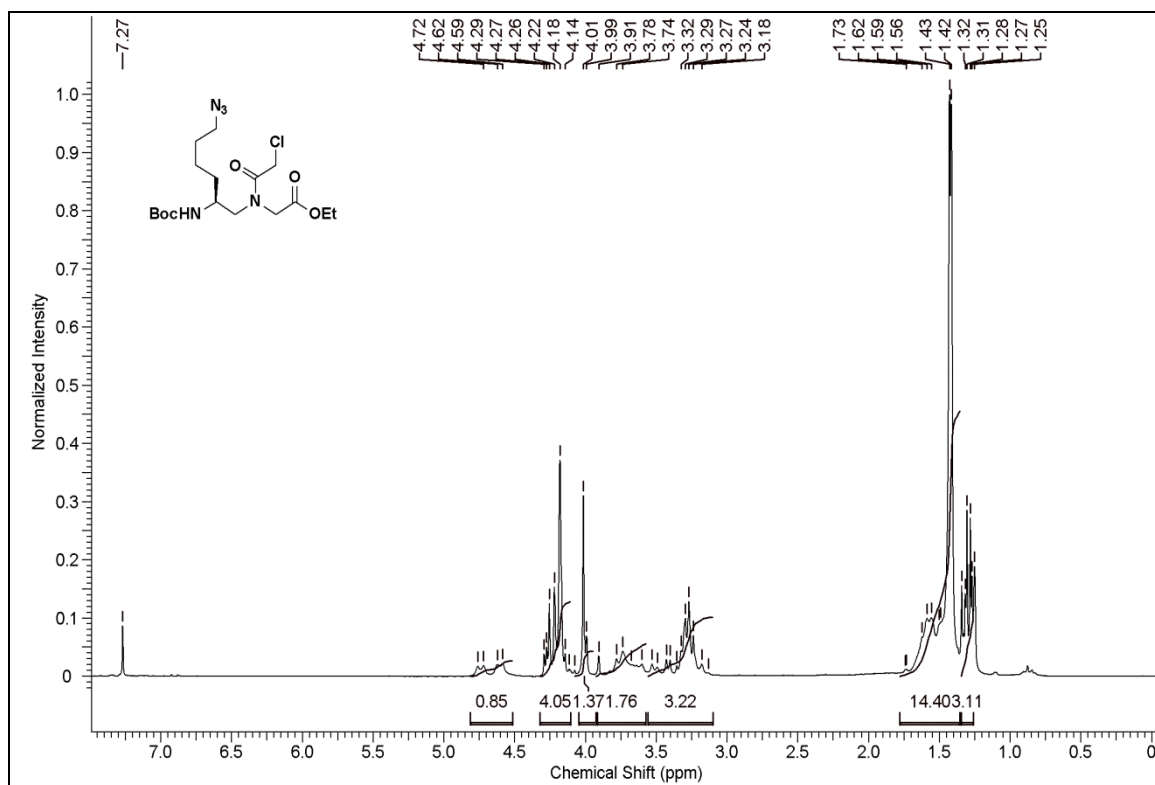
¹³C NMR of Compound 34
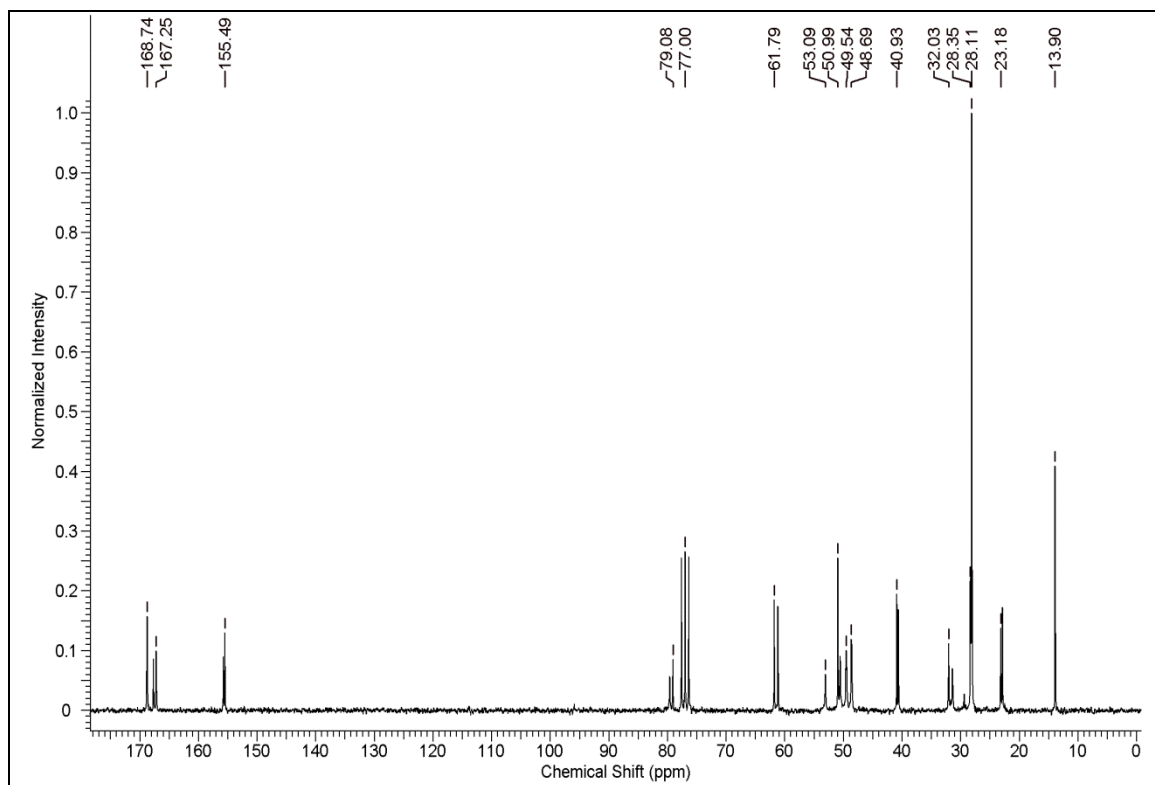
^{1}H NMR of Compound 35
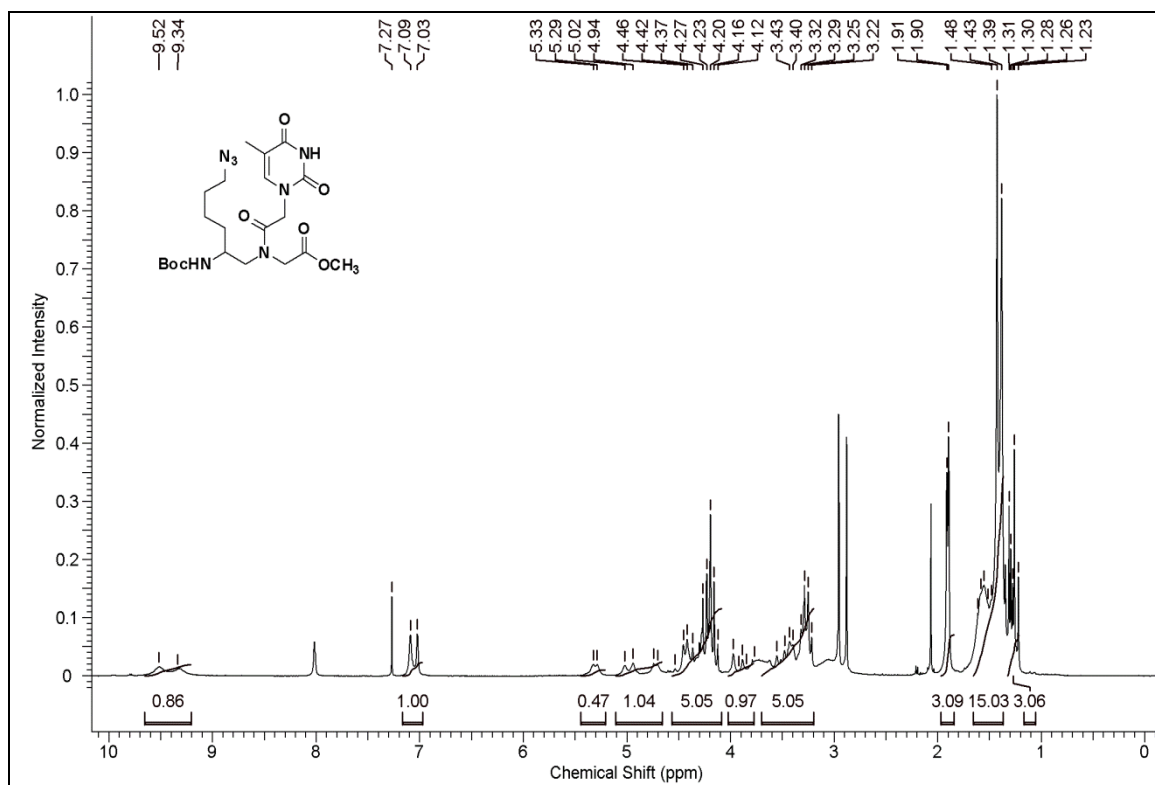
^{13}C NMR of Compound 35
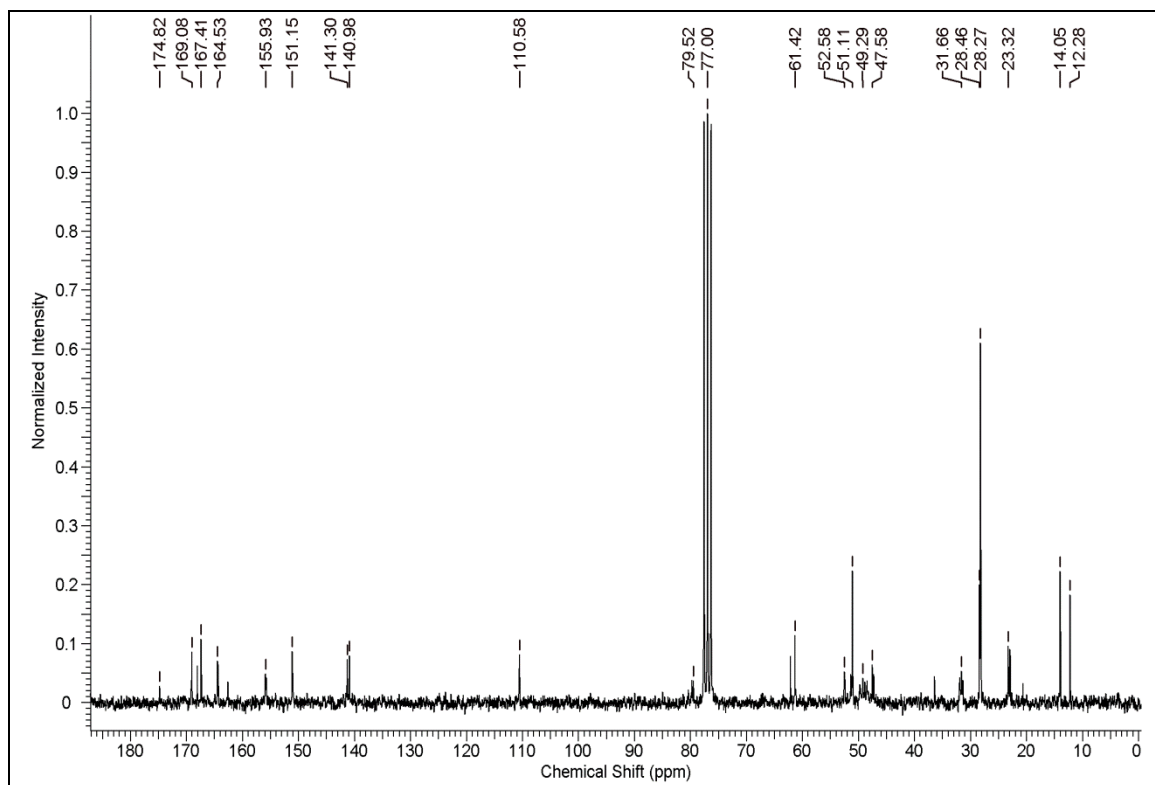
¹H NMR of Compound 36
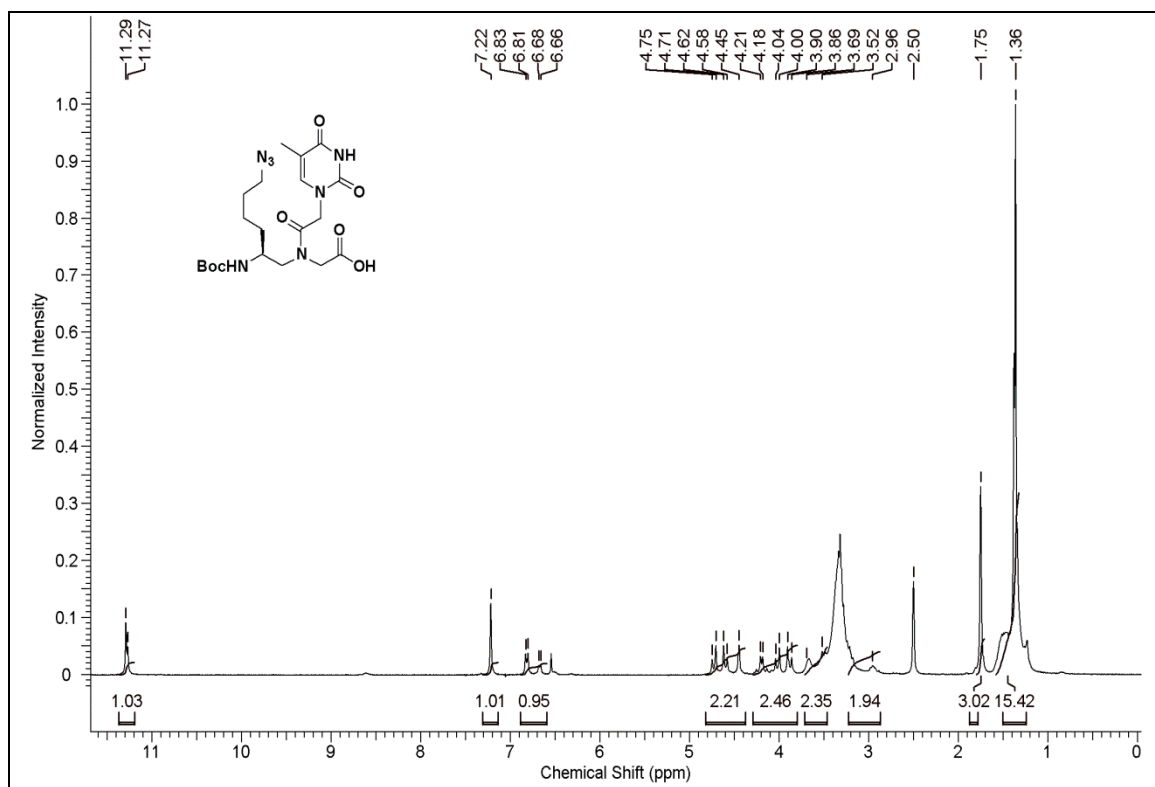
¹³C NMR of Compound 36
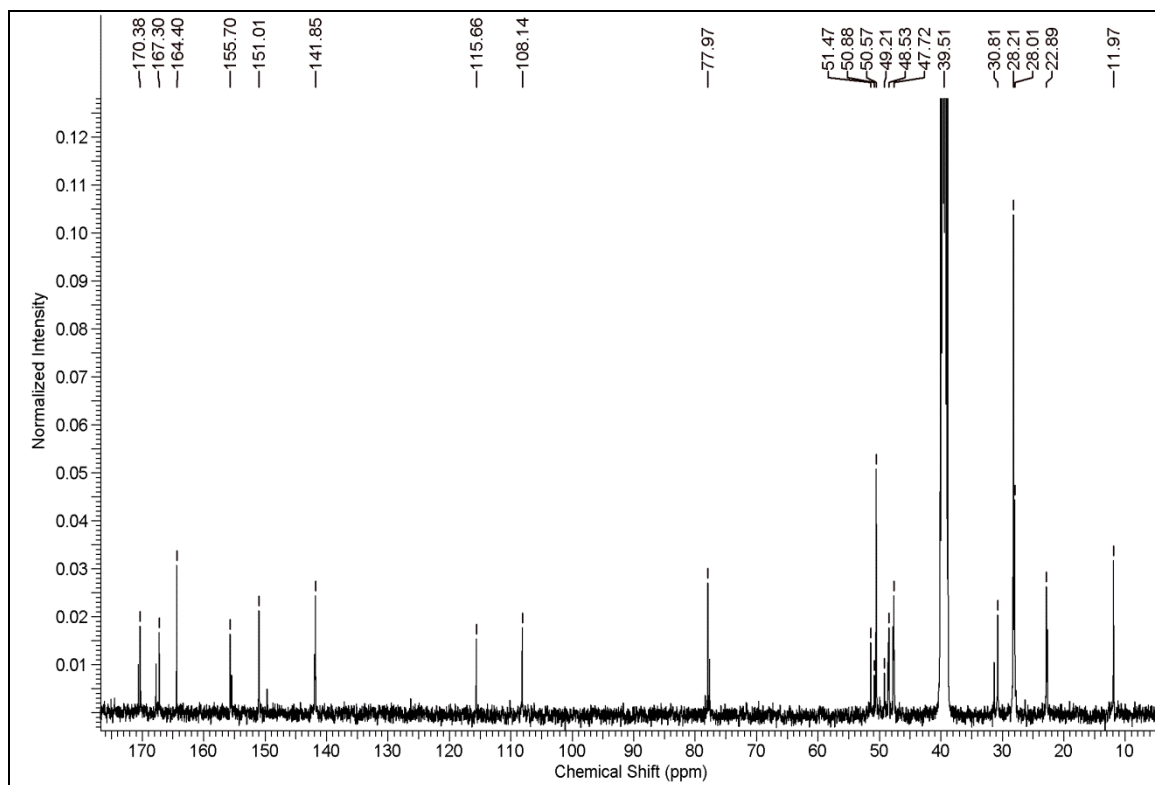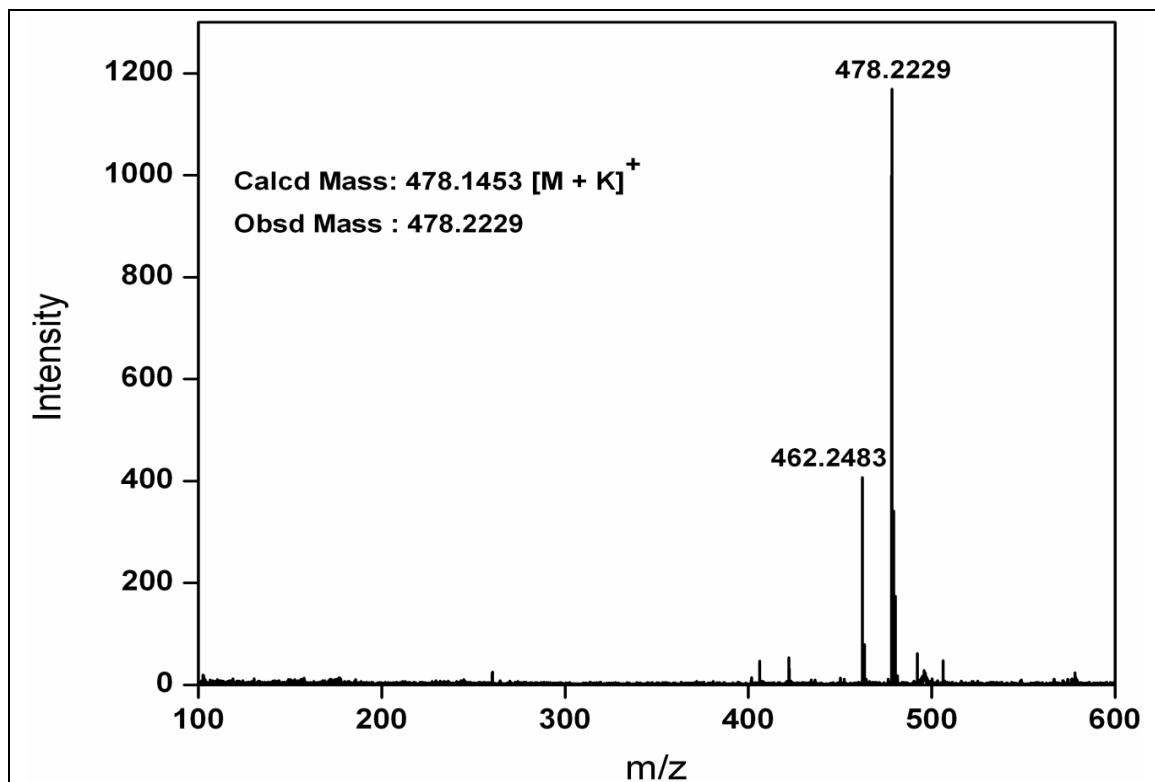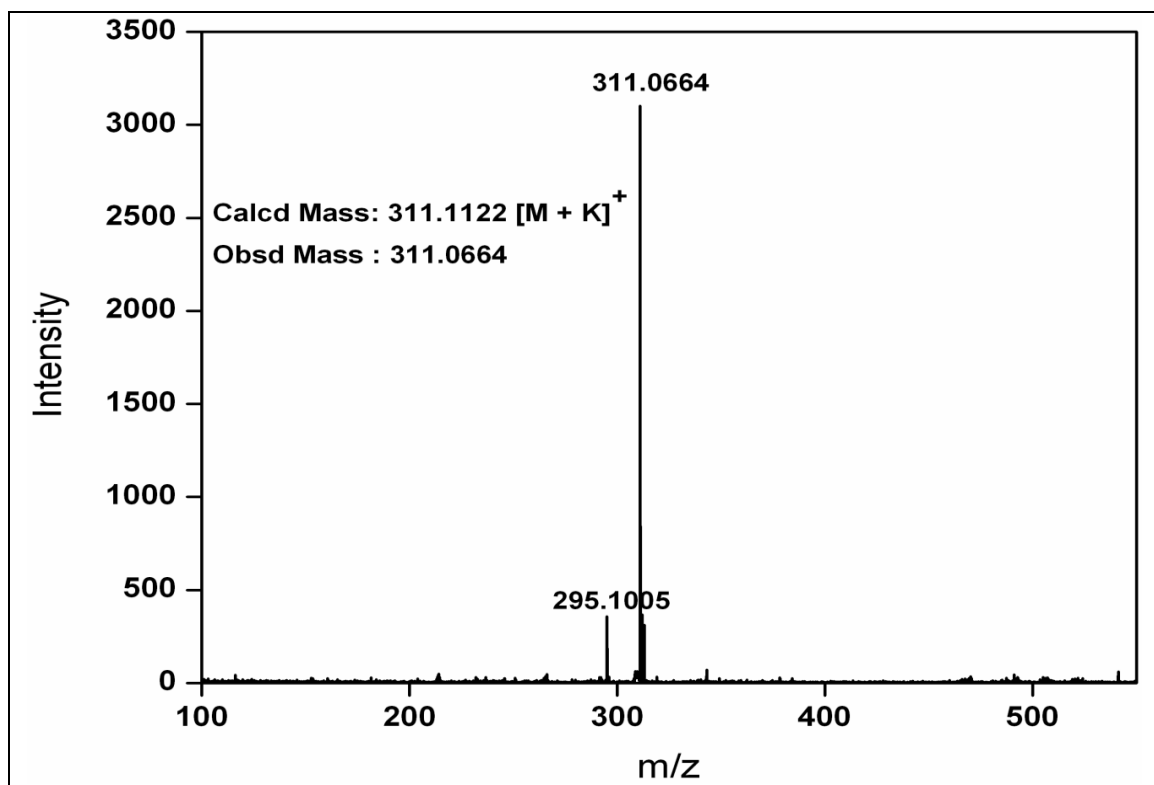
MALDI-TOF Mass of compound 2
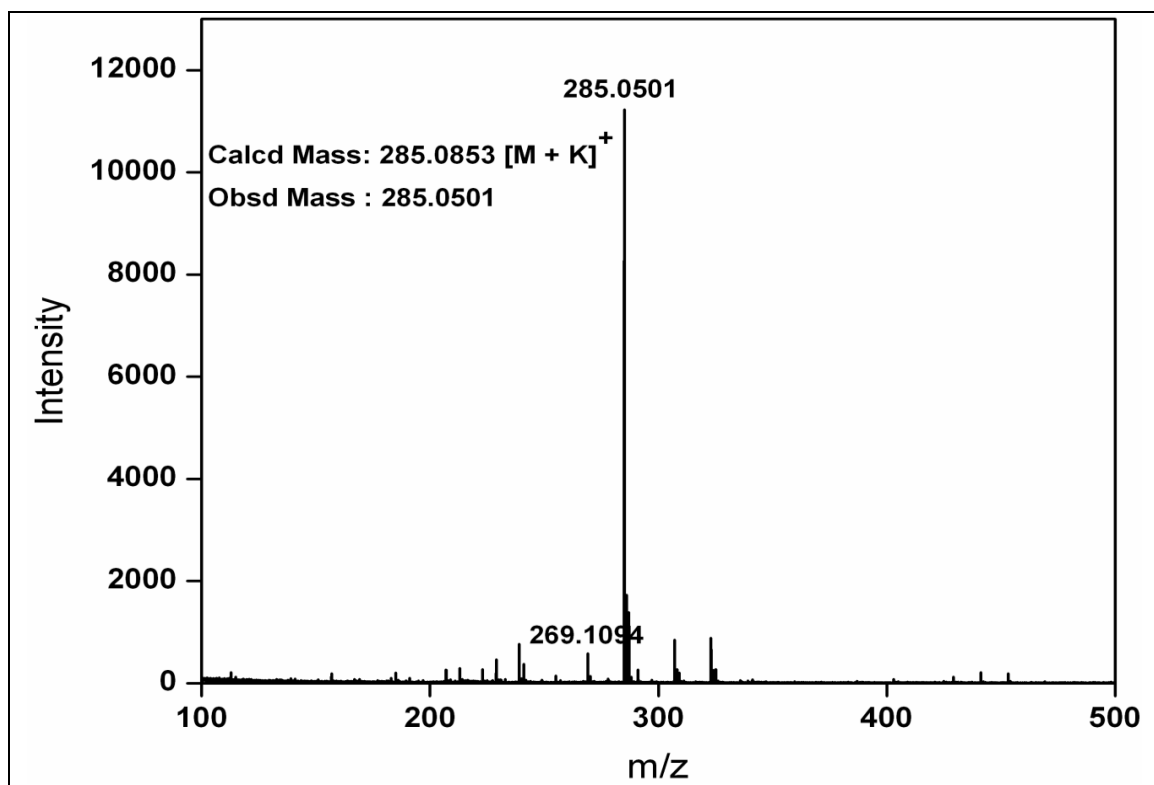
MALDI-TOF Mass of compound 3

MALDI-TOF Mass of compound 4

Calcd Mass: 391.1271 [M + K]$^+$
Obsd Mass : 391.1075

391.1075

MALDI-TOF Mass of compound 5

Calcd Mass: 405.1428 [M + K]$^+$
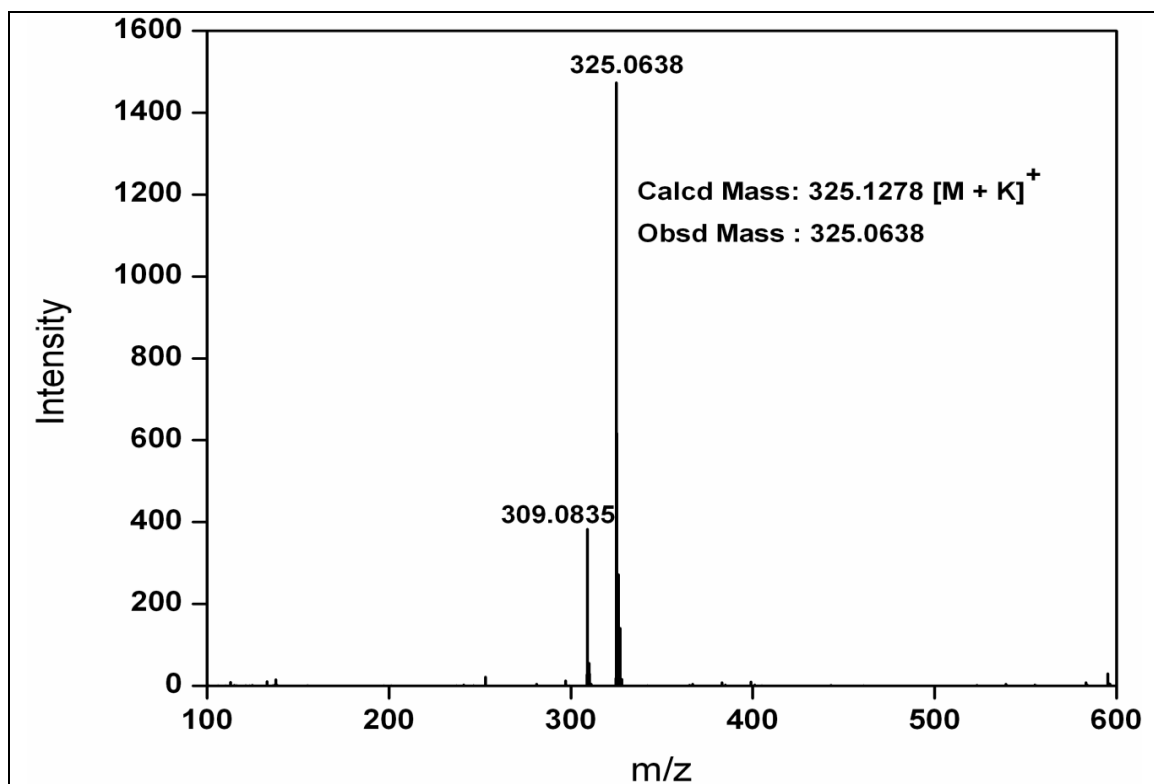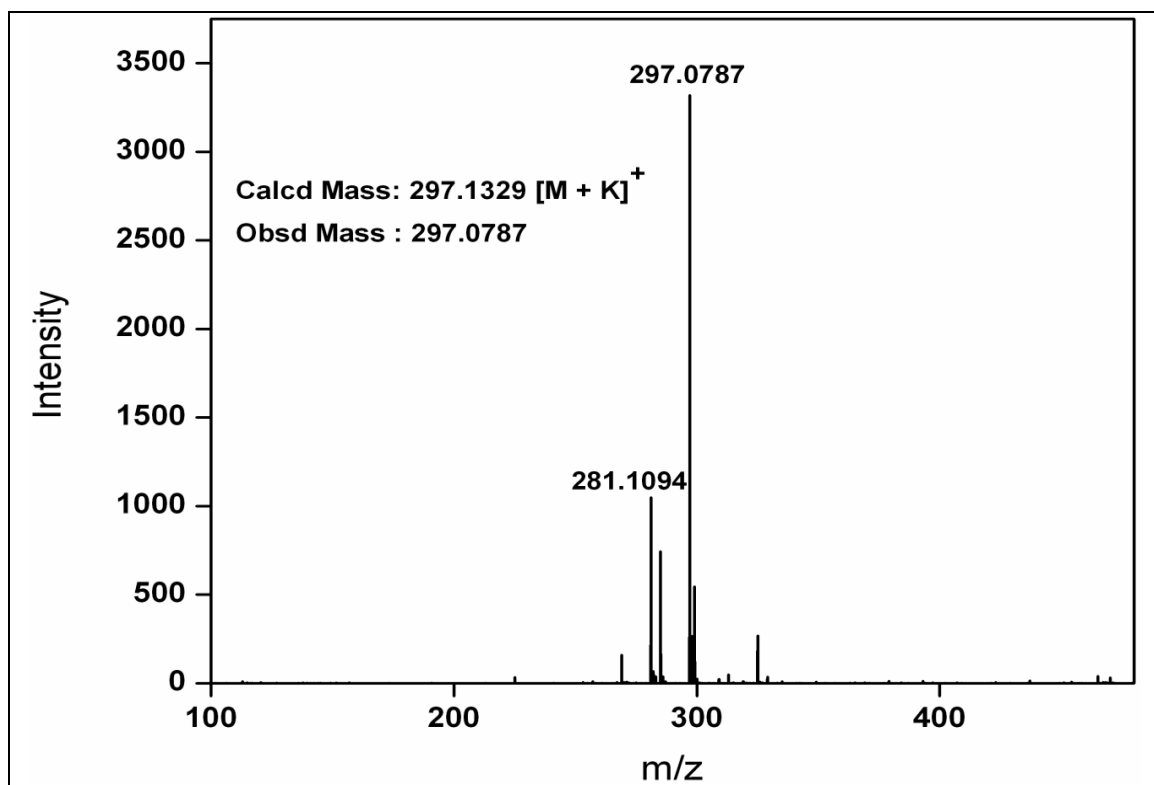Obsd Mass : 405.1199

405.1199

389.1465
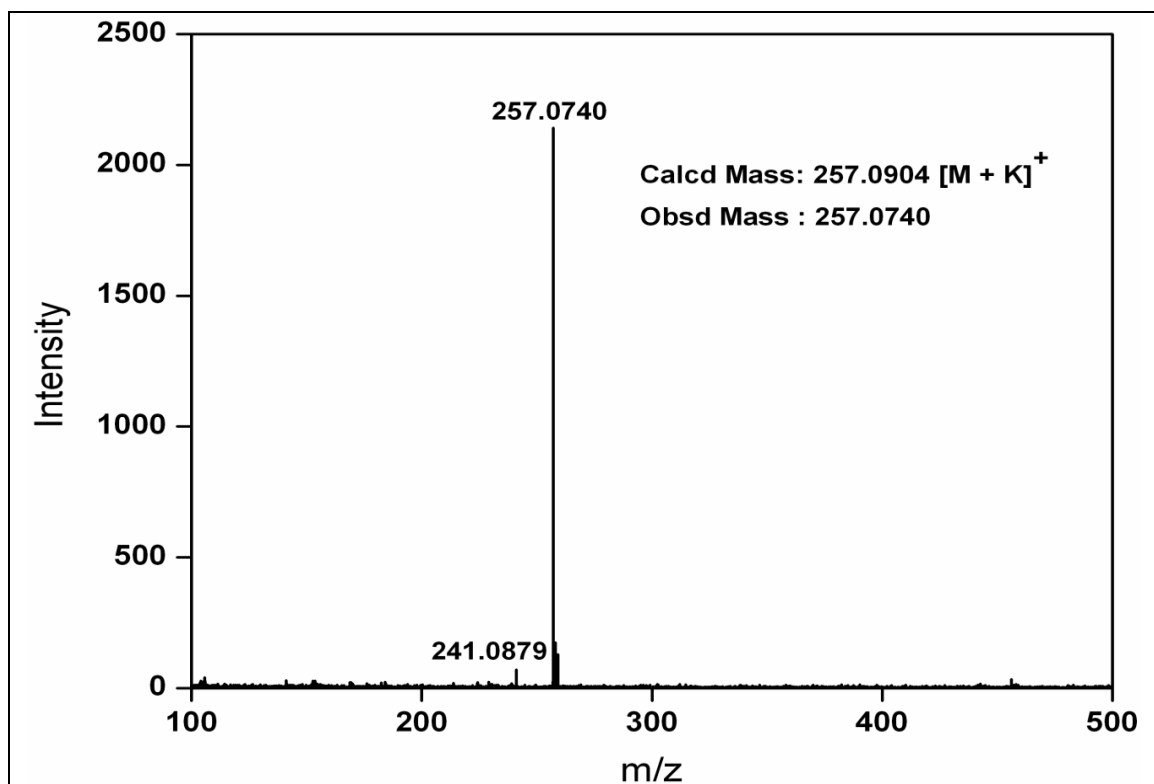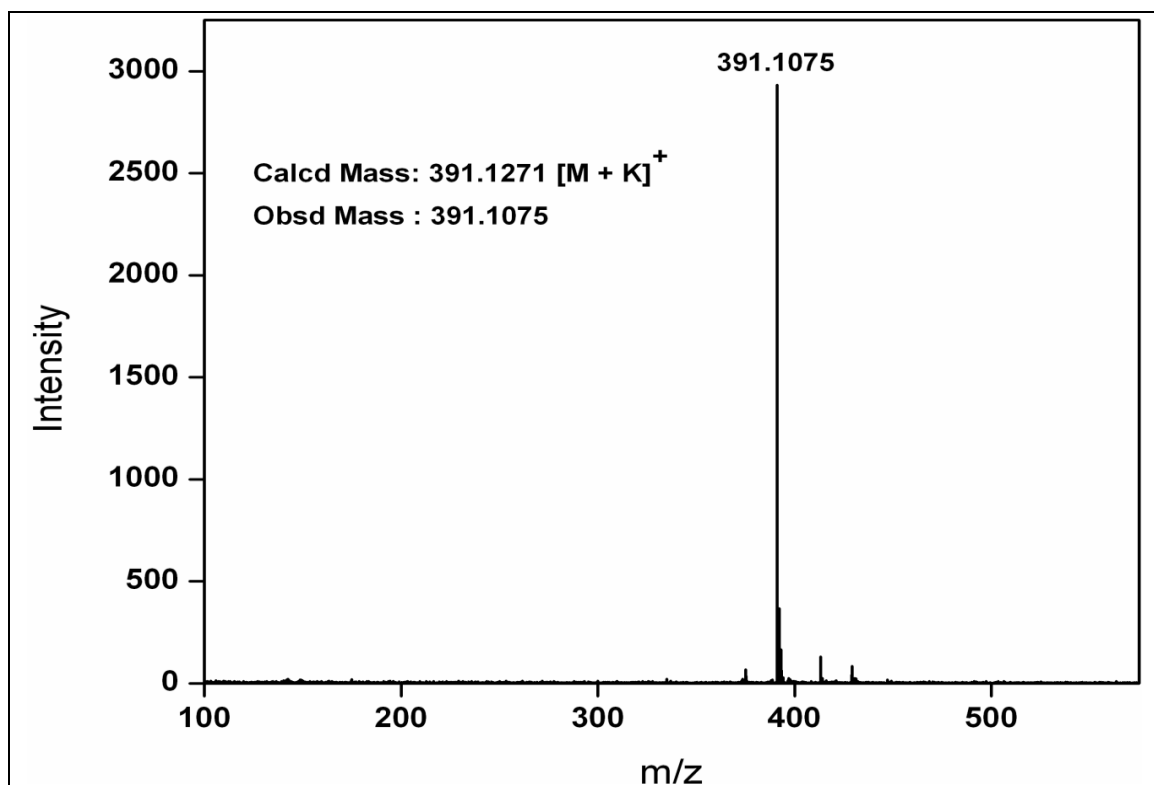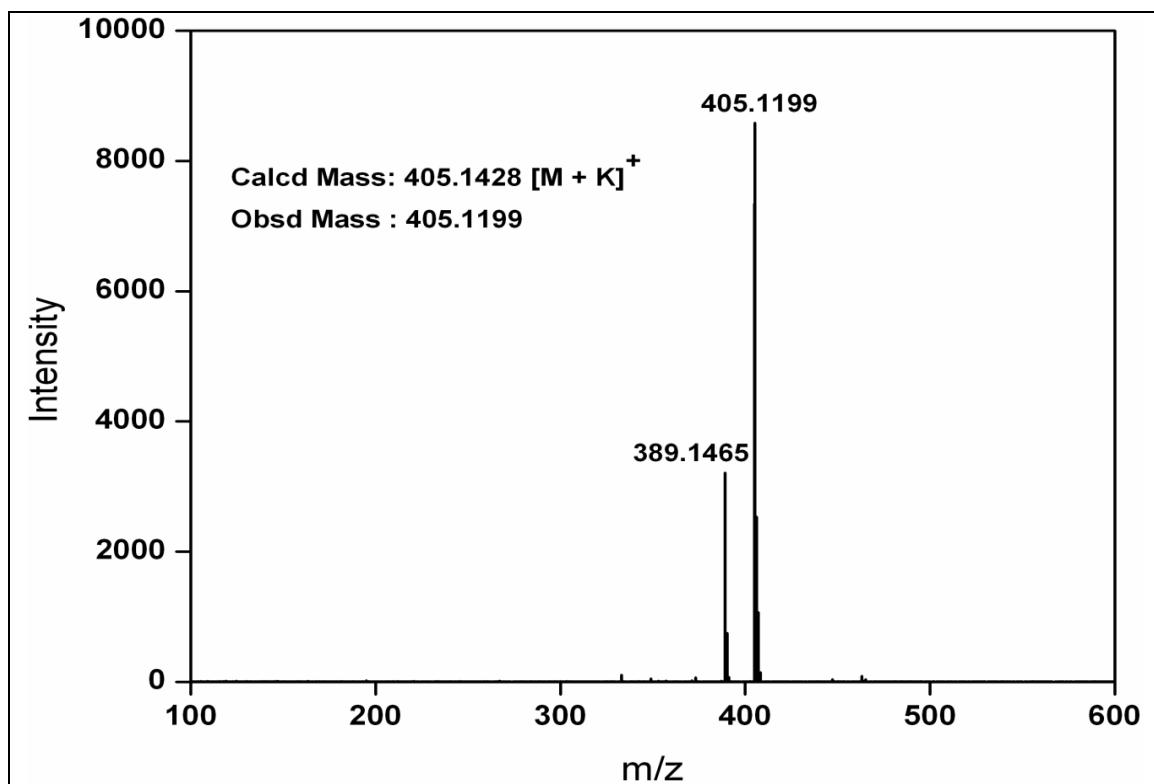
MALDI-TOF Mass of compound 6
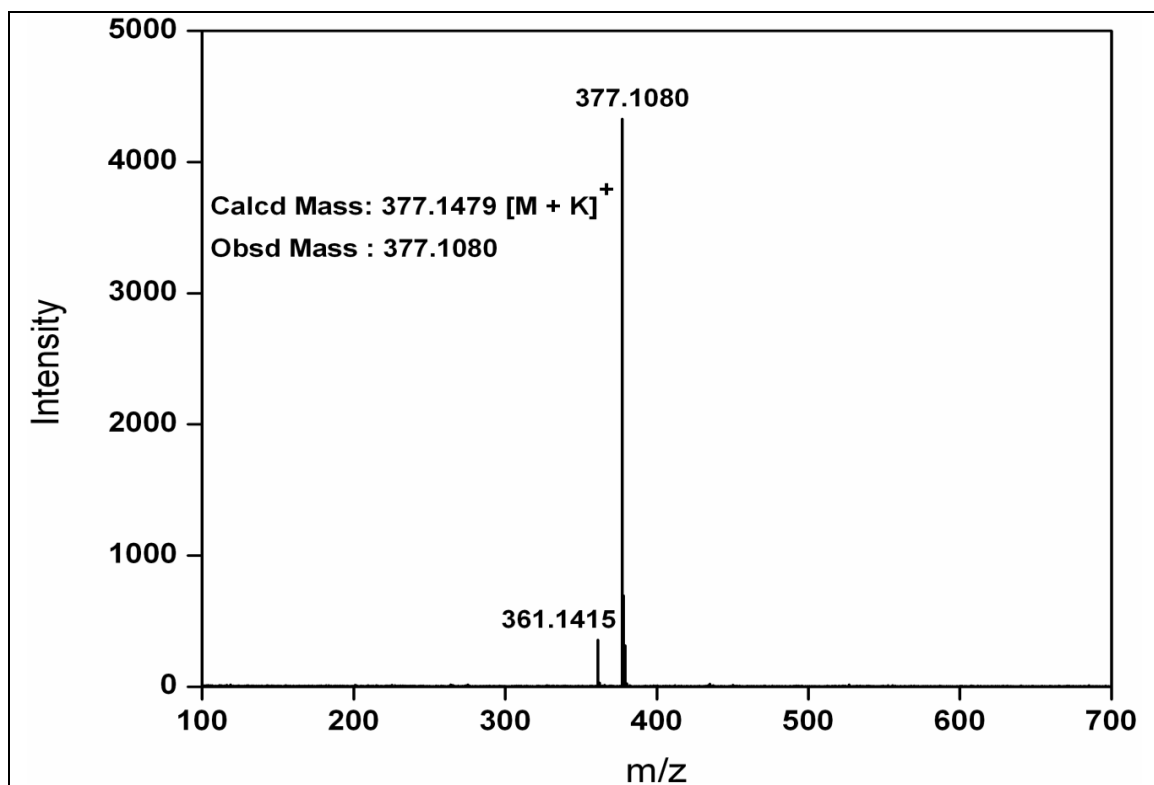
MALDI-TOF Mass of compound 8

MALDI-TOF Mass of compound 9
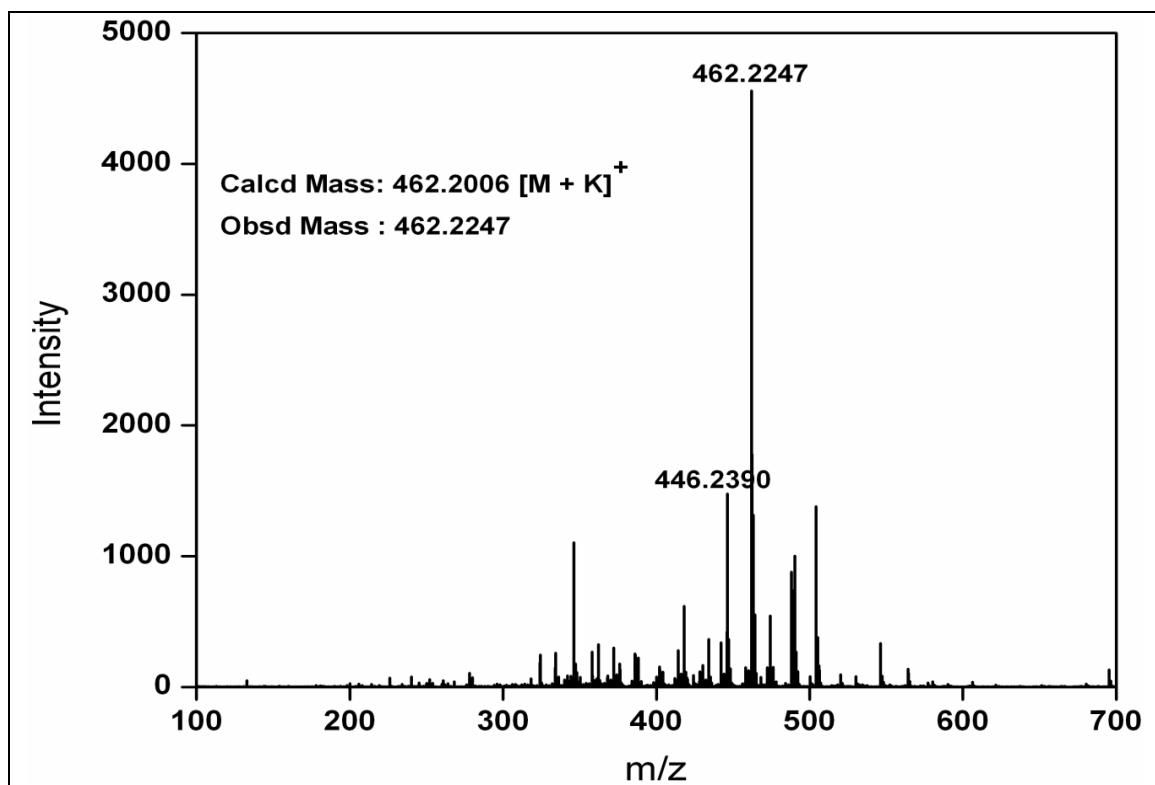
Calcd Mass: 462.2006 [M + K]⁺
Obsd Mass : 462.2247

MALDI-TOF Mass of compound 10

Calcd Mass: 538.1722 [M + K]⁺
Obsd Mass : 538.1591
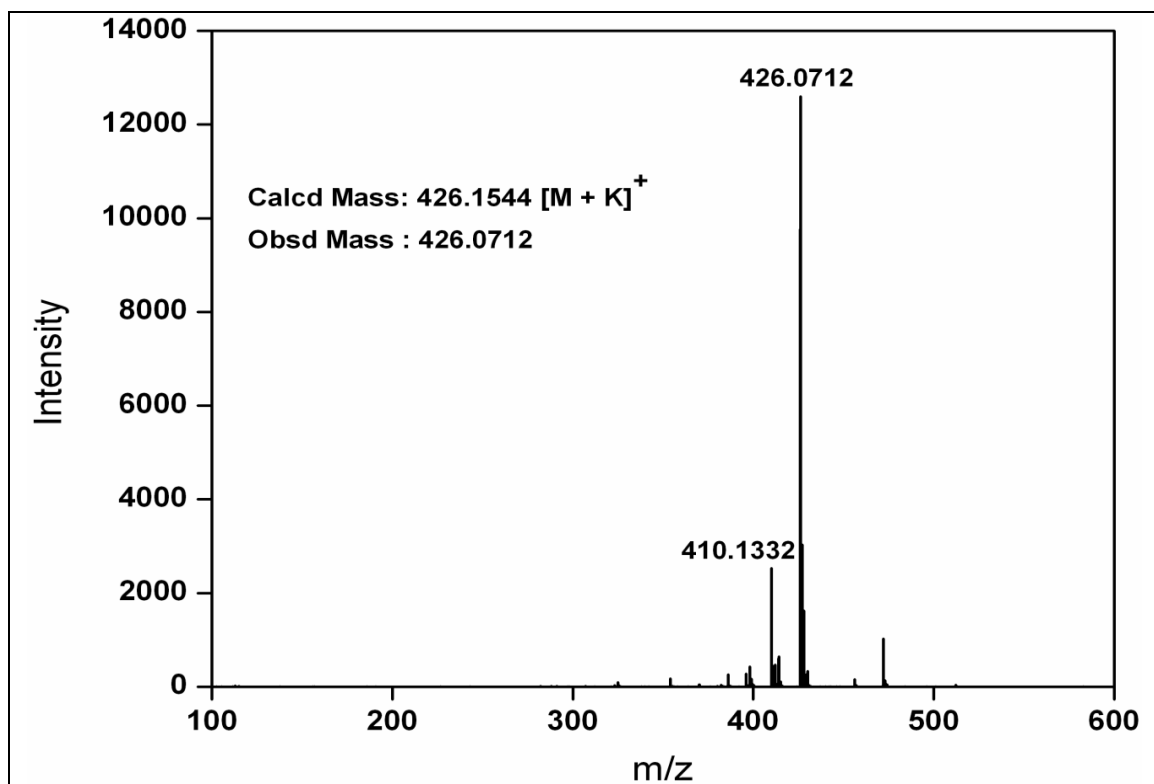
MALDI-TOF Mass of compound 11

Calcd Mass: 628.2385 [M + K]$^+$

Obsd Mass : 628.2361

MALDI-TOF Mass of compound 12

Calcd Mass: 600.2072 [M + K]$^+$

Obsd Mass : 600.1787
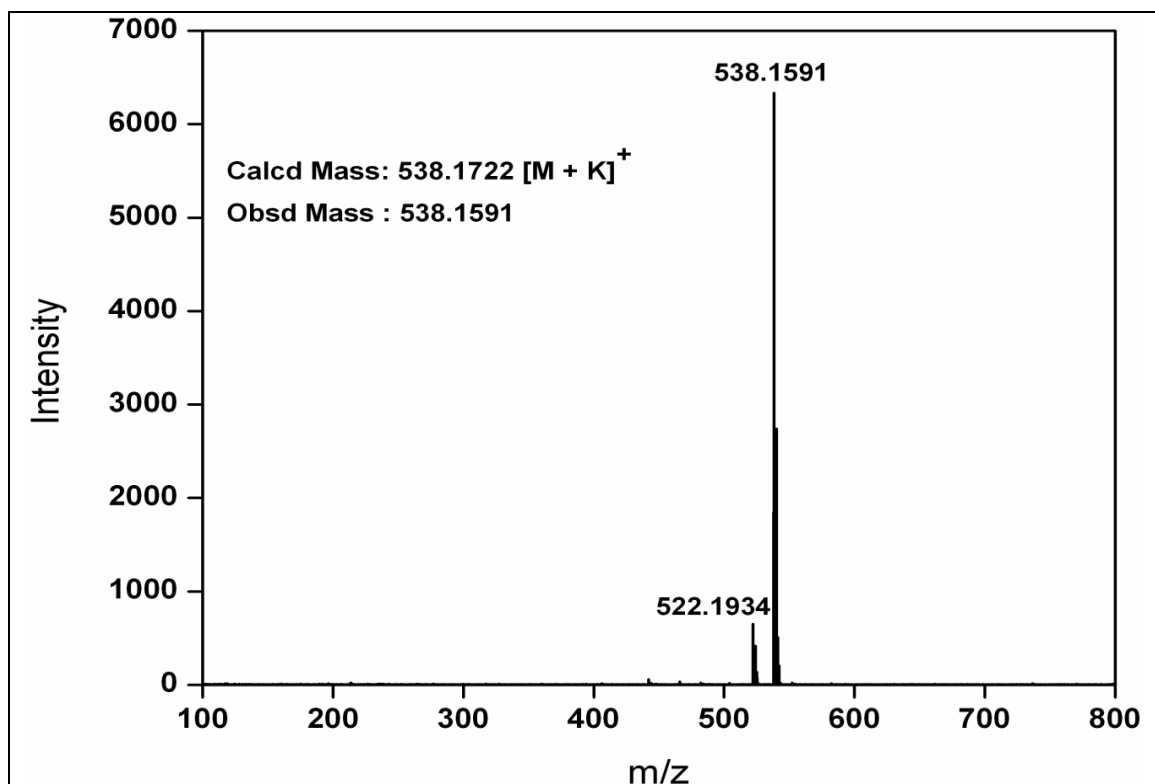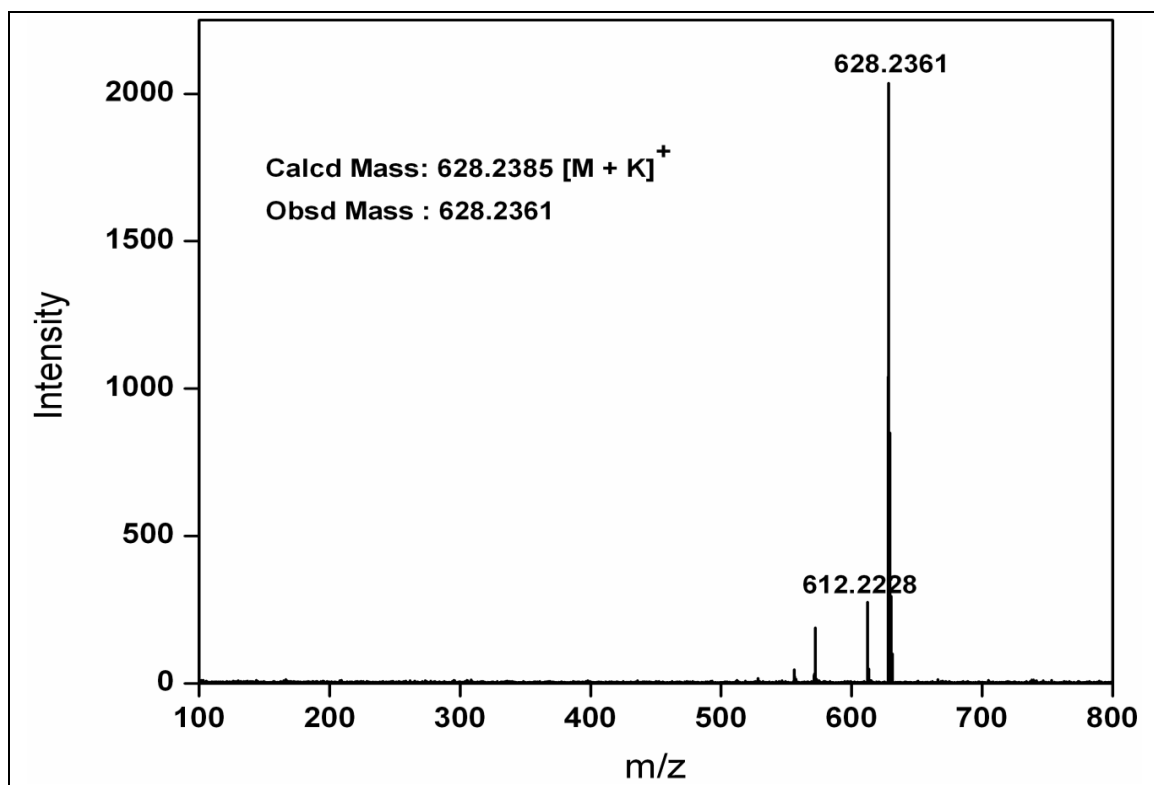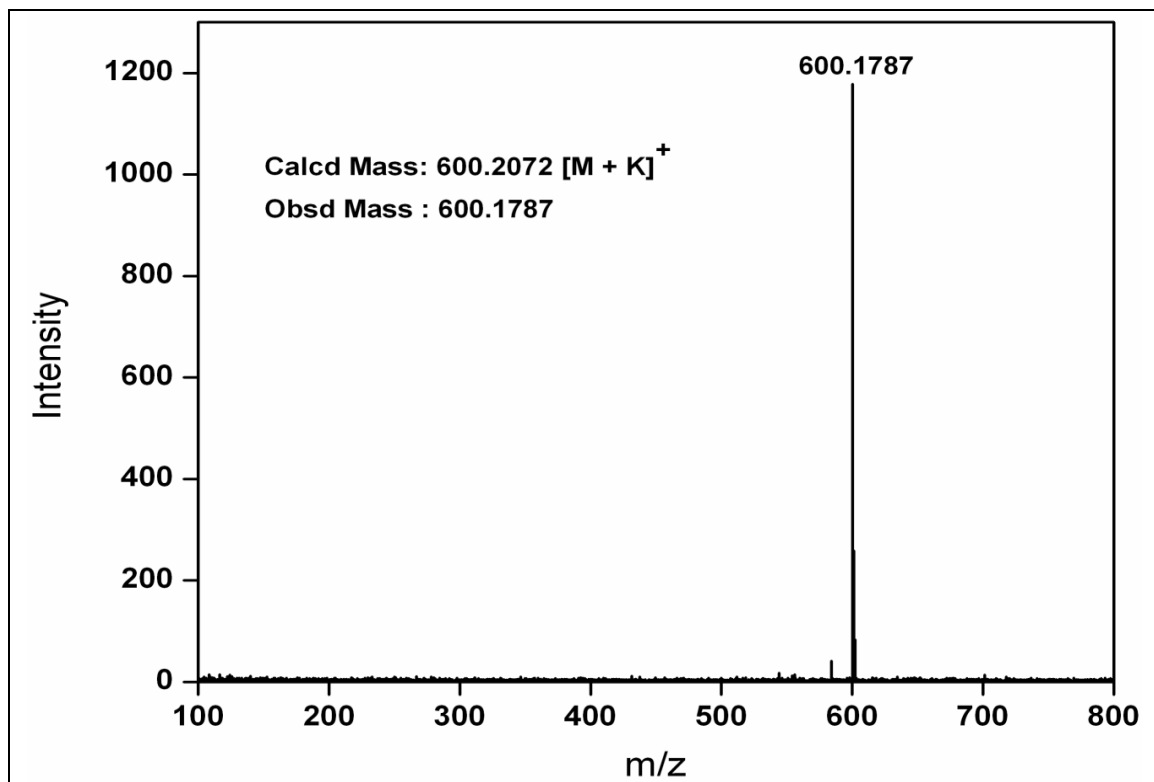
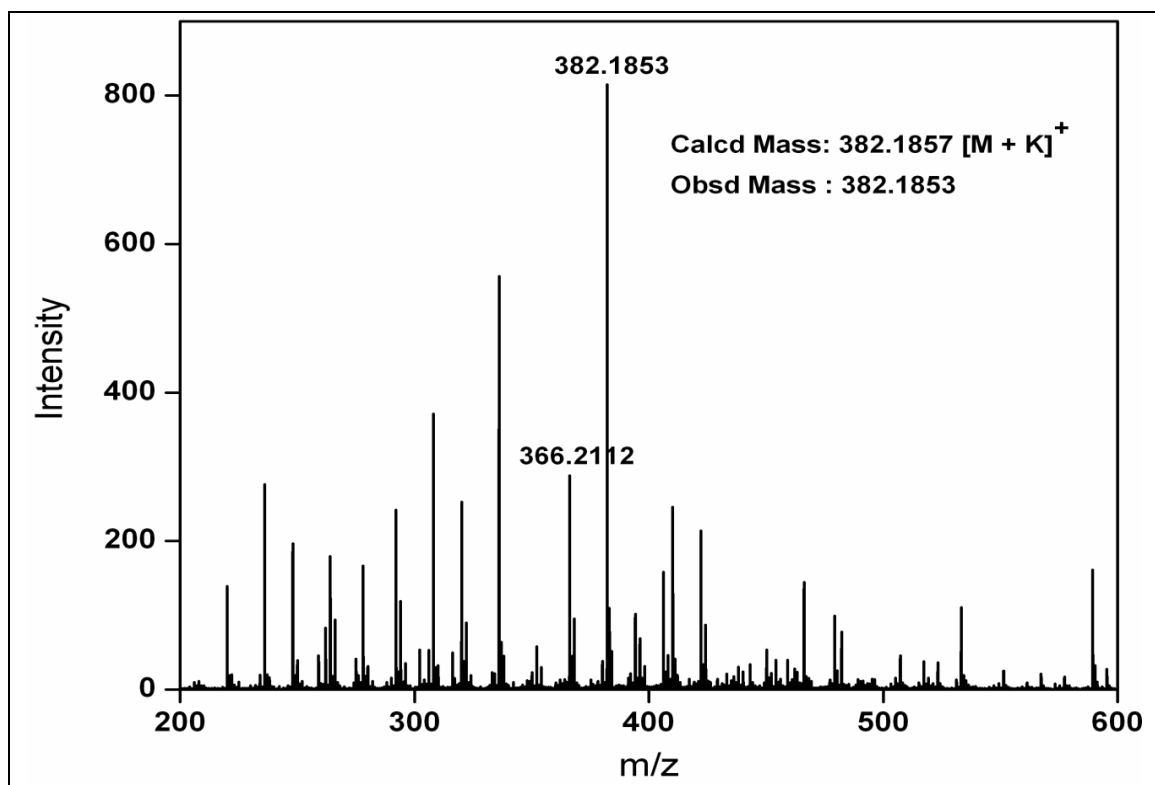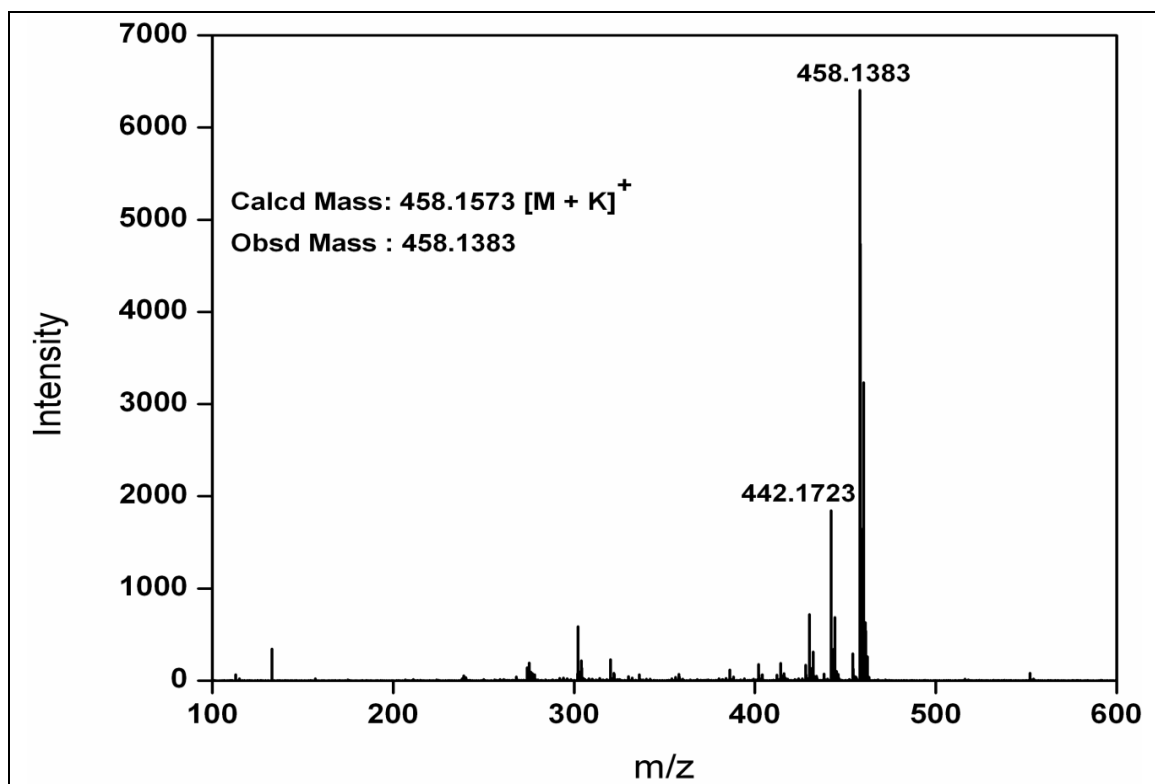**MALDI-TOF Mass of compound 13**

MALDI-TOF Mass of compound 14

127

MALDI-TOF Mass of compound 16

Calcd Mass: 271.0696 [M + K]$^+$

Obsd Mass : 271.0638

MALDI-TOF Mass of compound 17

Calcd Mass: 243.0747 [M + K]$^+$

Obsd Mass : 243.0456
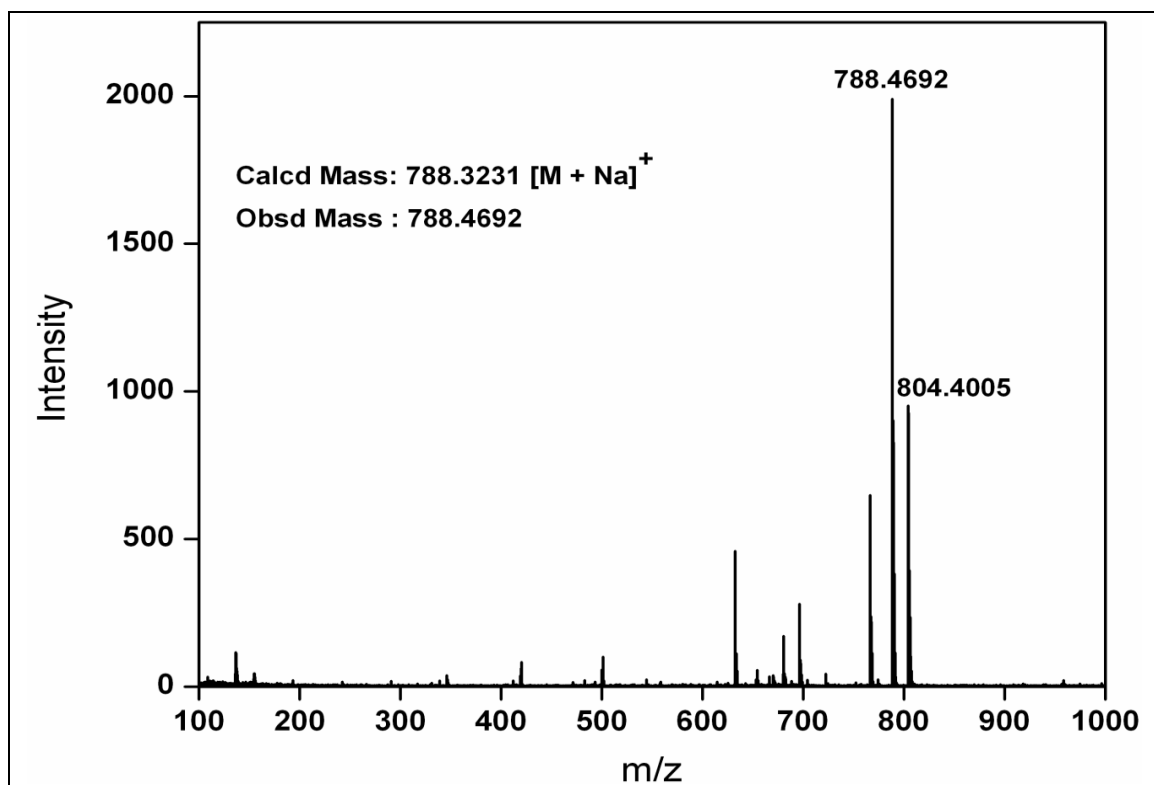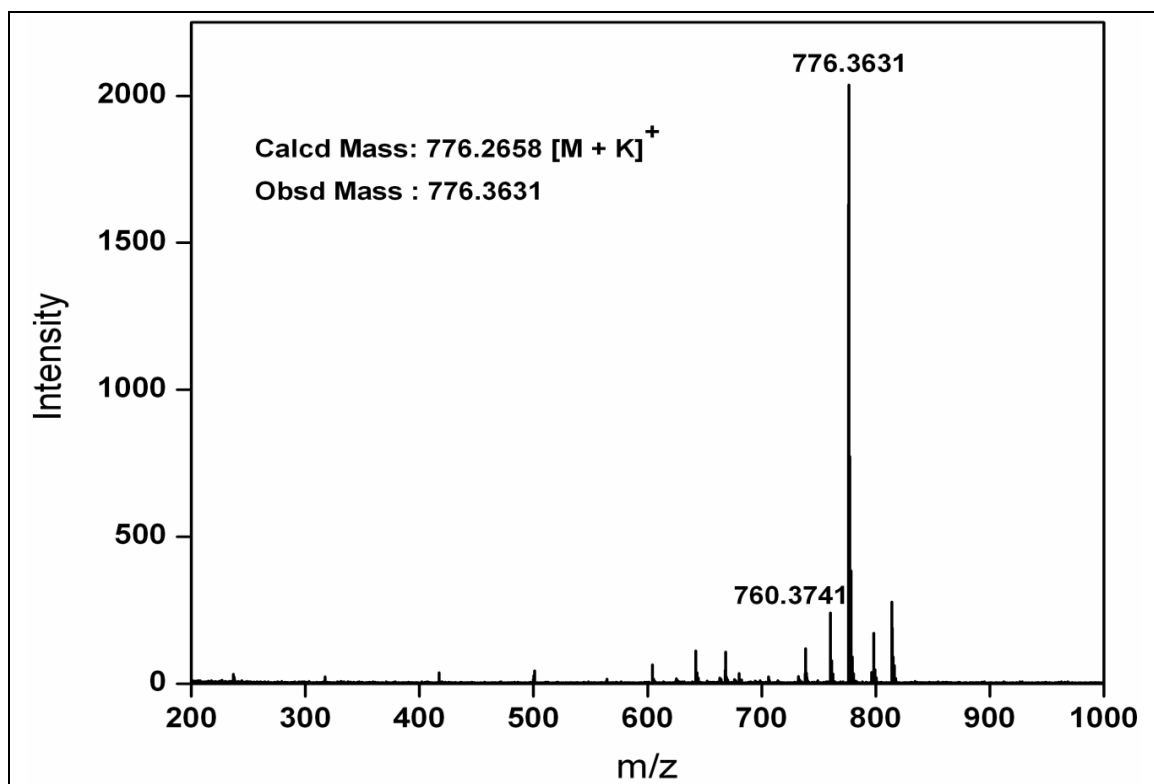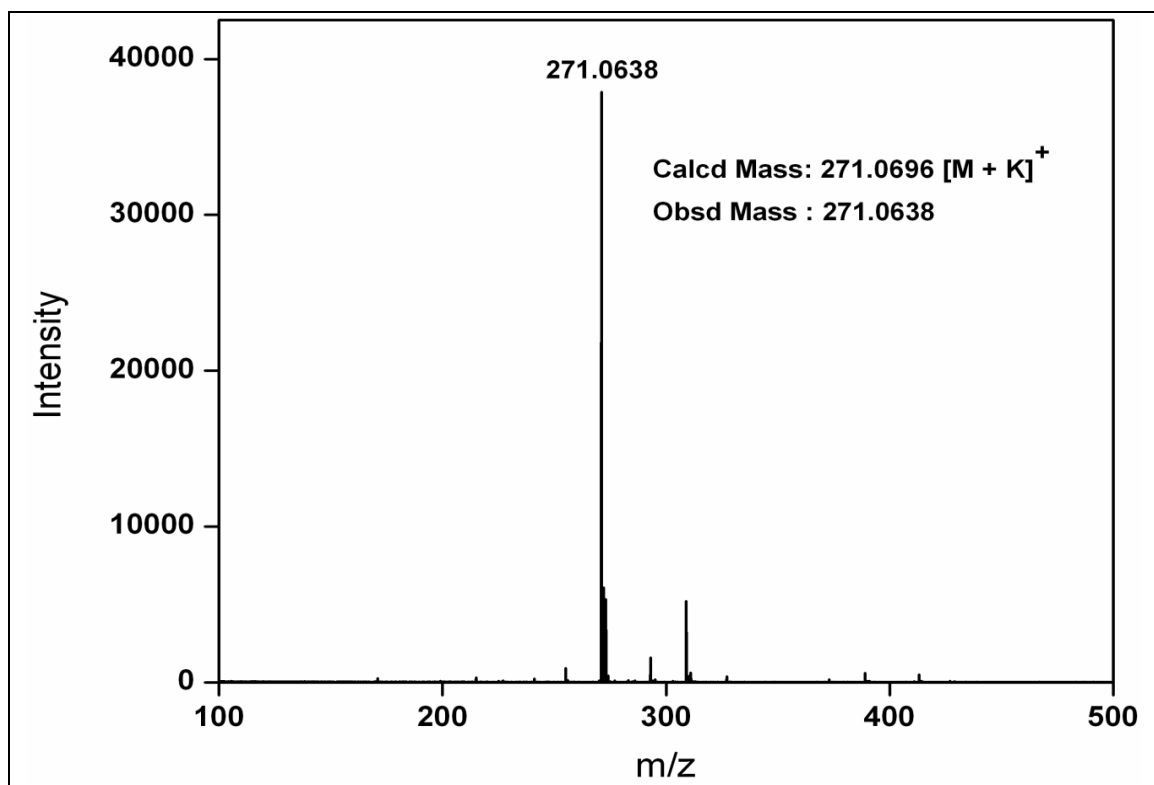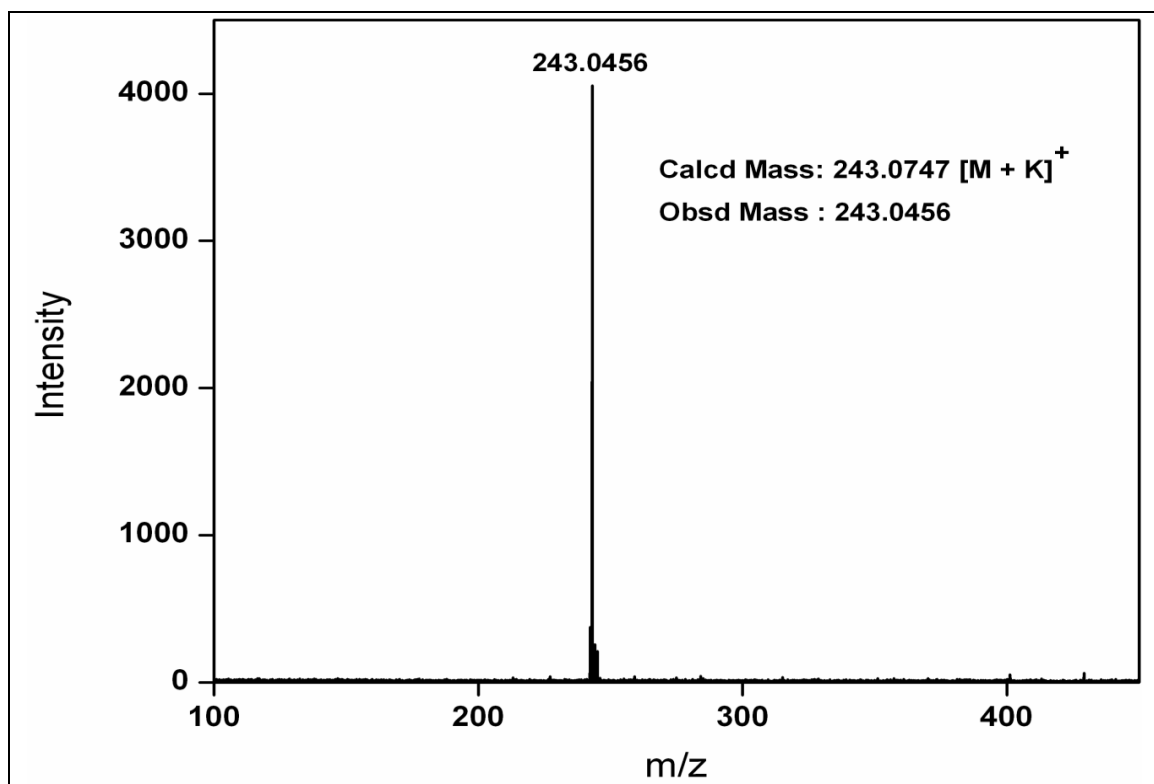
MALDI-TOF Mass of compound 18
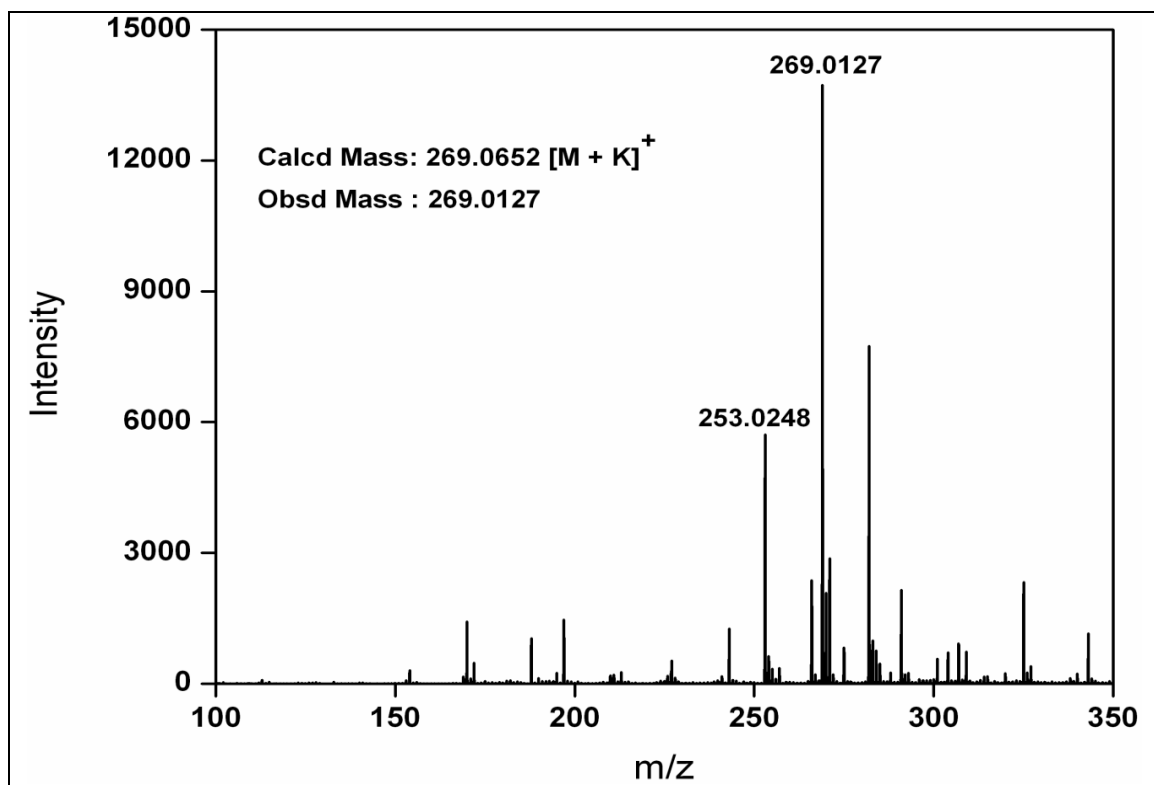
MALDI-TOF Mass of compound 19
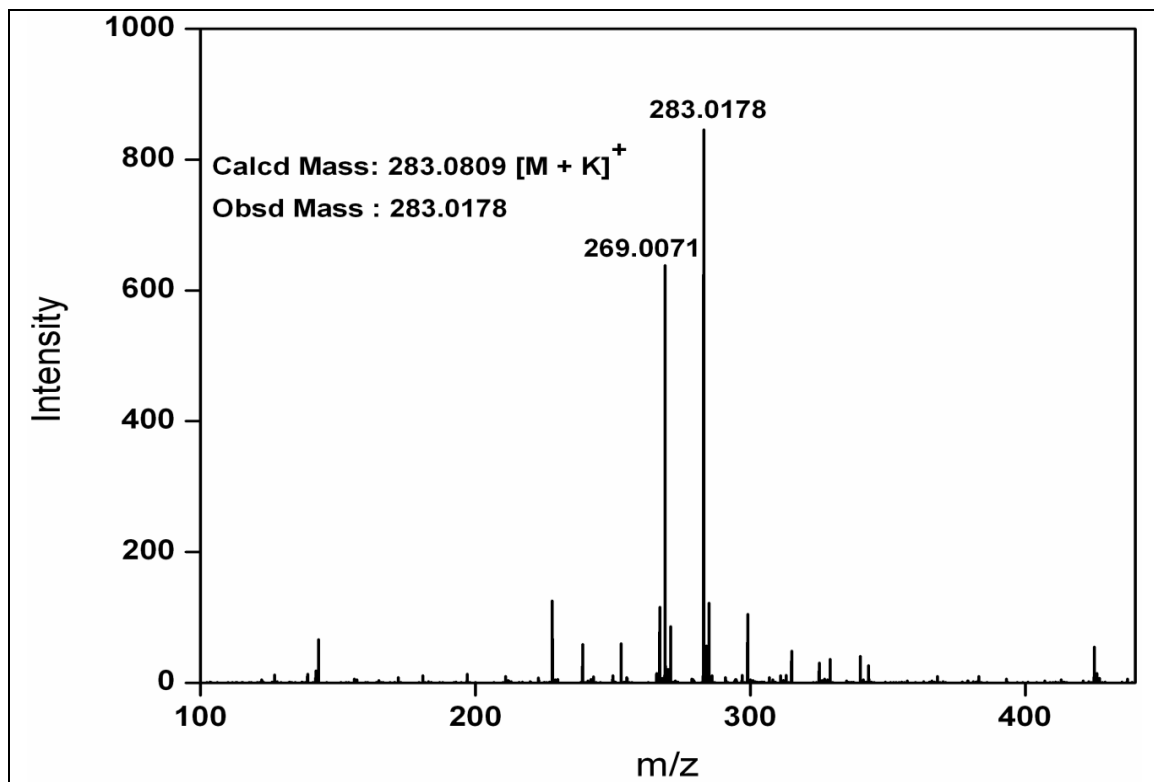
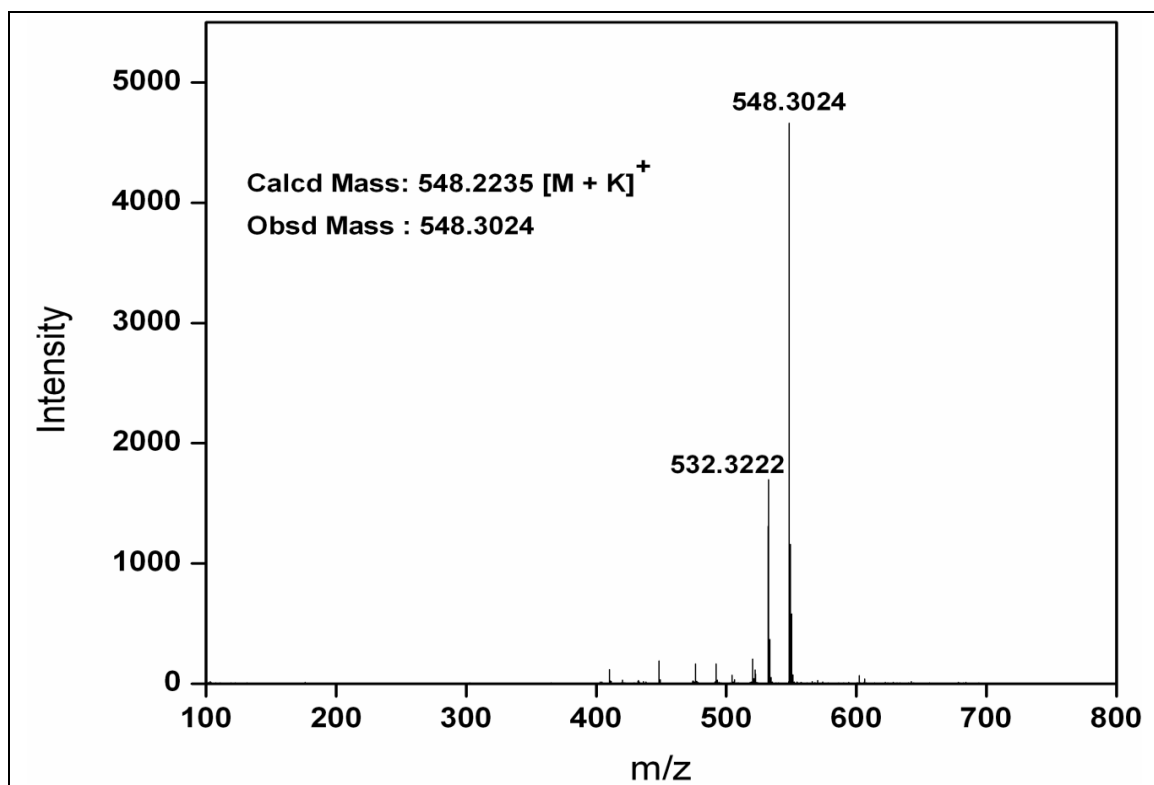**MALDI-TOF Mass of compound 20**
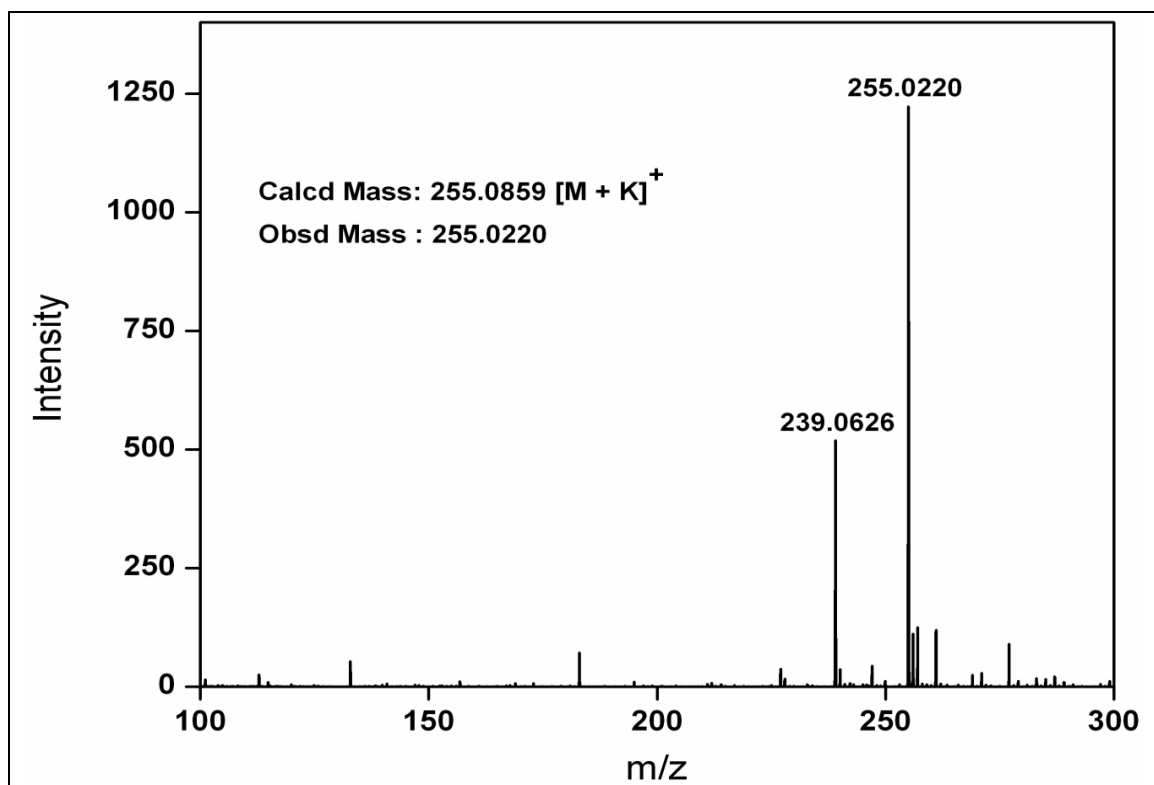
Calcd Mass: 255.0859 [M + K]$^{+}$

Obsd Mass : 255.0220

(Peaks labeled: 255.0220, 239.0626)

MALDI-TOF Mass of compound 22

Calcd Mass: 384.1074 [M + K]$^{+}$

Obsd Mass : 384.0857

(Peak labeled: 384.0857)

MALDI-TOF Mass of compound 23

Calcd Mass: 340.1387 [M + K]$^+$
Obsd Mass : 340.0984

MALDI-TOF Mass of compound 24

Calcd Mass: 416.1103 [M + K]$^+$
Obsd Mass : 416.1190
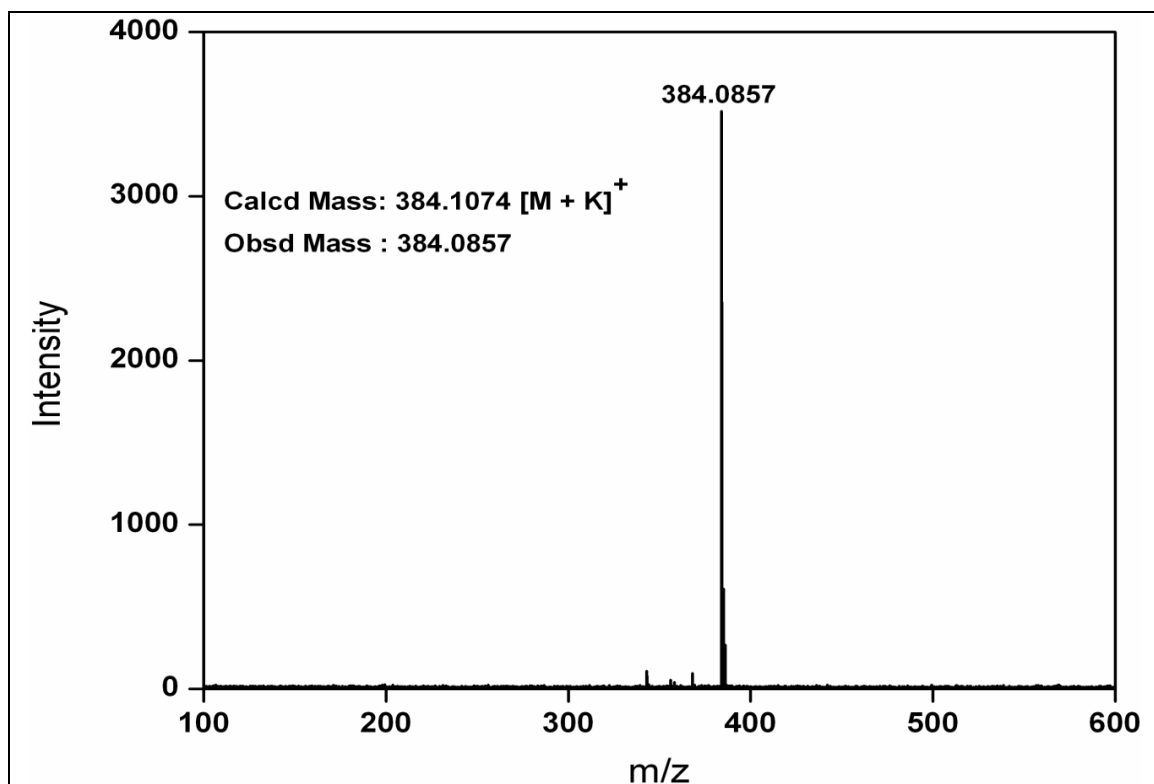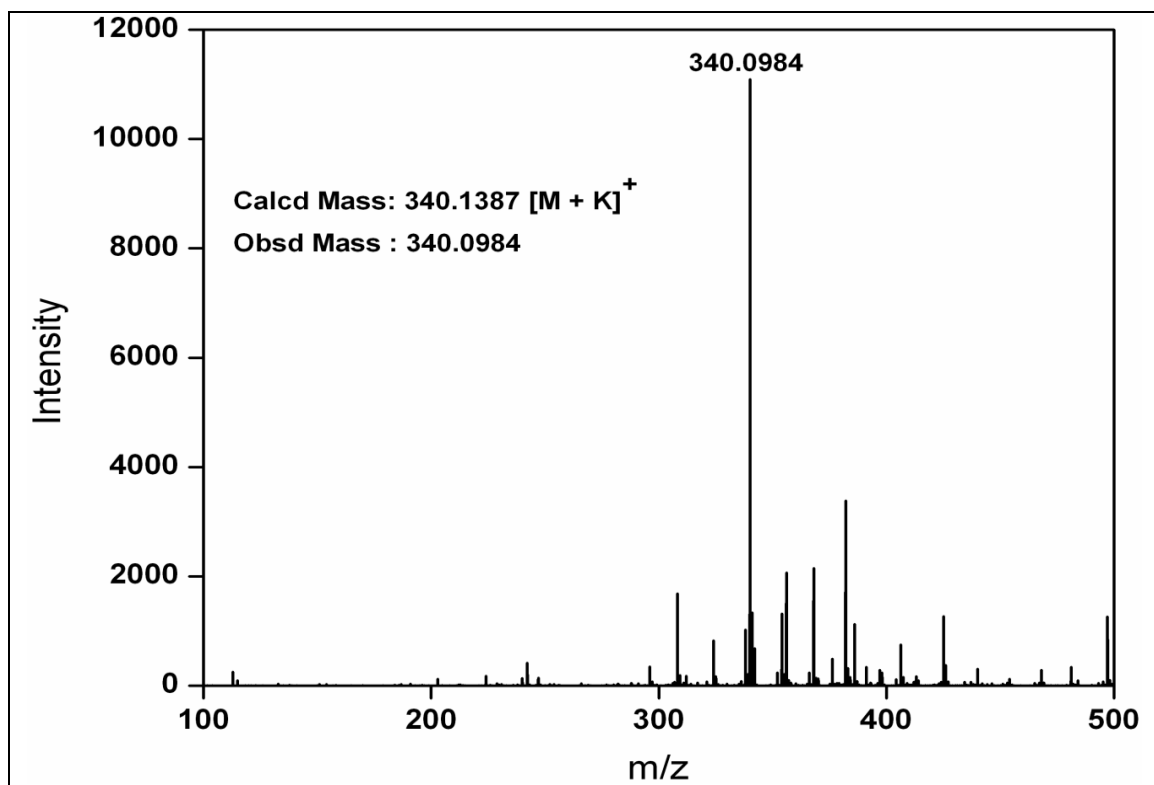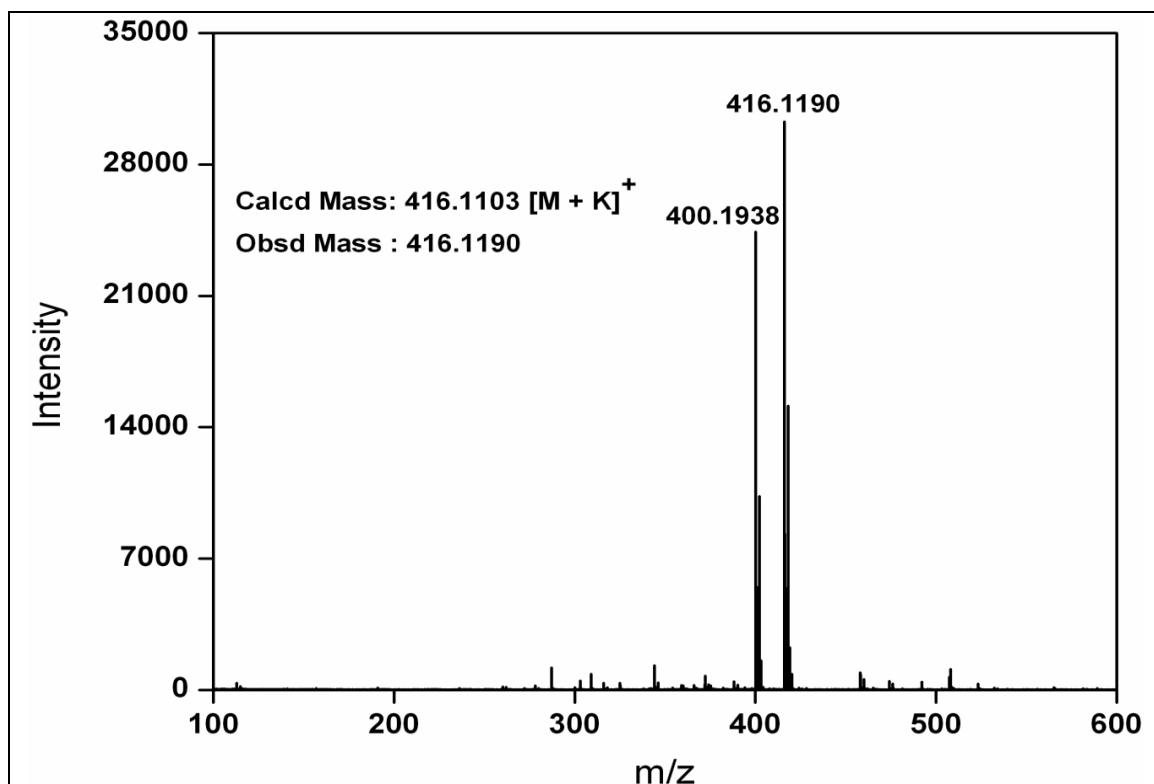
**MALDI-TOF Mass of compound 25**
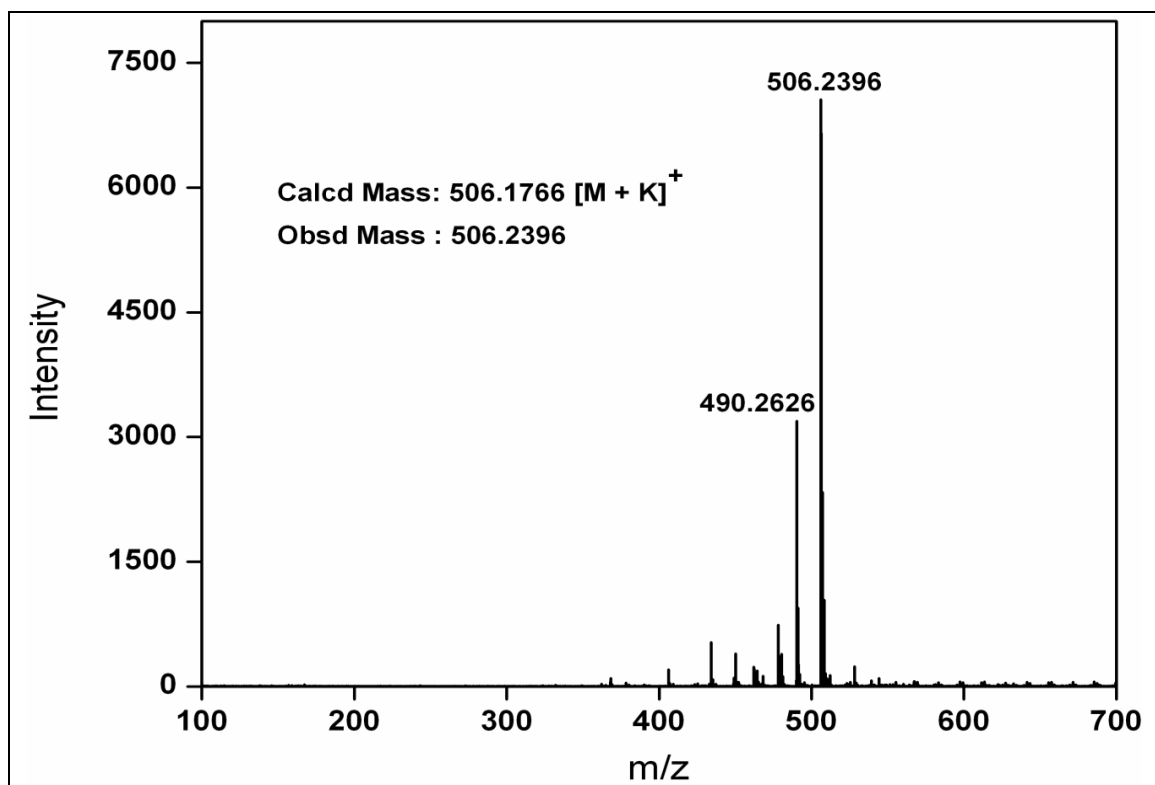

MALDI-TOF Mass of compound 26

MALDI-TOF Mass of compound 28

MALDI-TOF Mass of compound 29

MALDI-TOF Mass of compound 30

Calcd Mass: 297.1329 [M + K]$^+$
Obsd Mass : 297.0787

MALDI-TOF Mass of compound 32

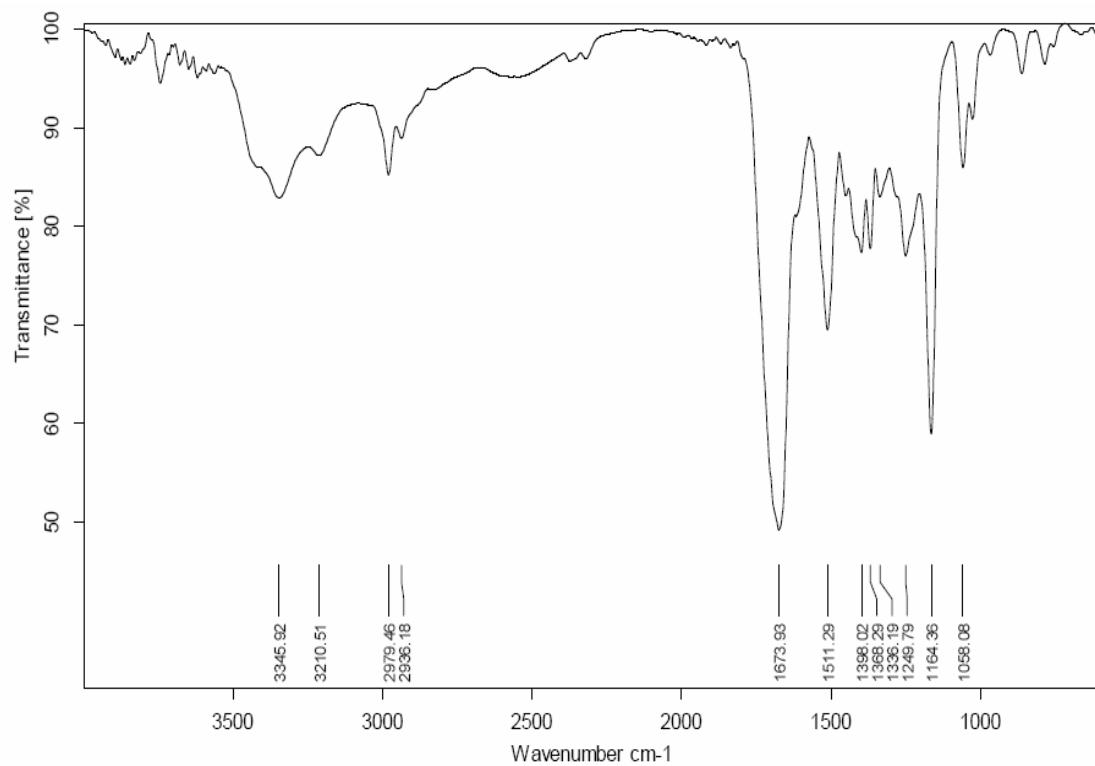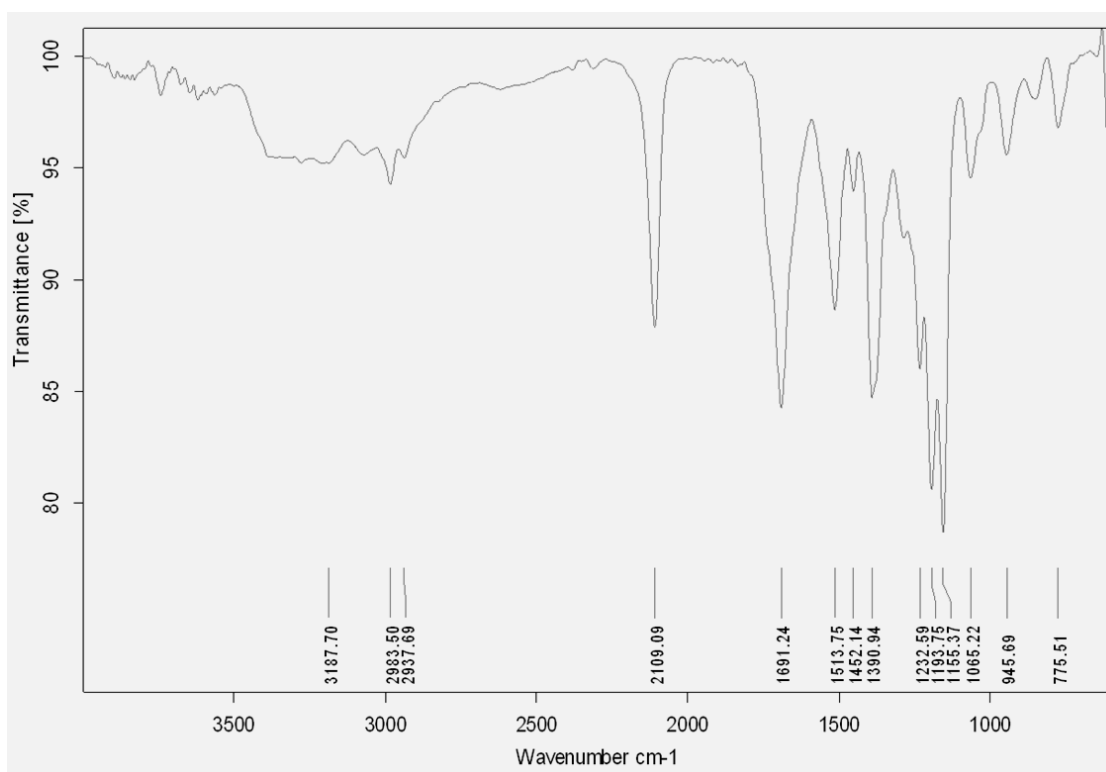Calcd Mass: 426.1544 [M + K]$^+$
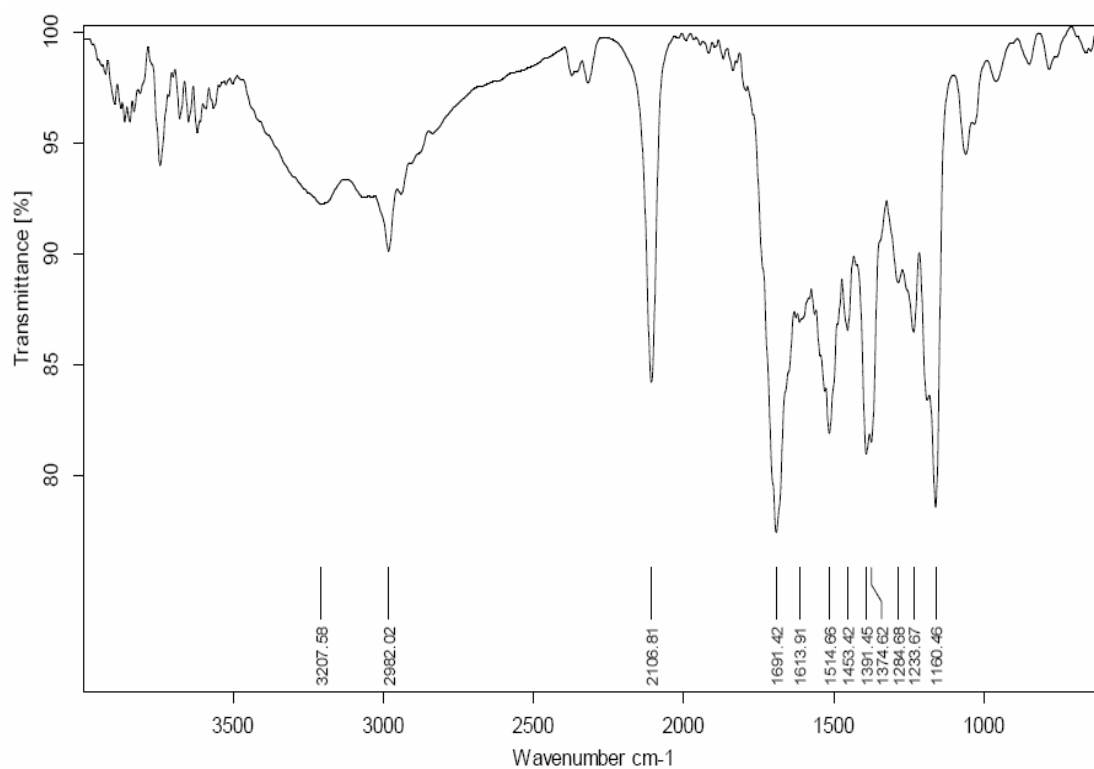Obsd Mass : 426.0712

MALDI-TOF Mass of compound 33

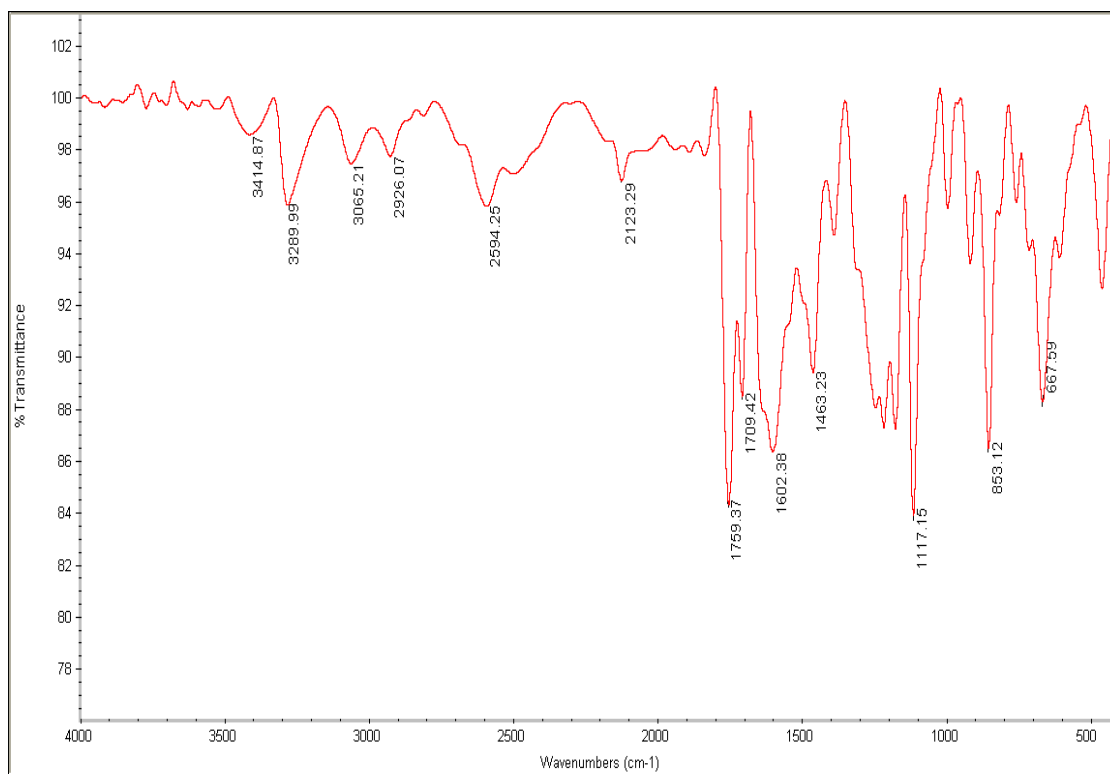**MALDI-TOF Mass of compound 34**

MALDI-TOF Mass of compound 35

MALDI-TOF Mass of compound 36
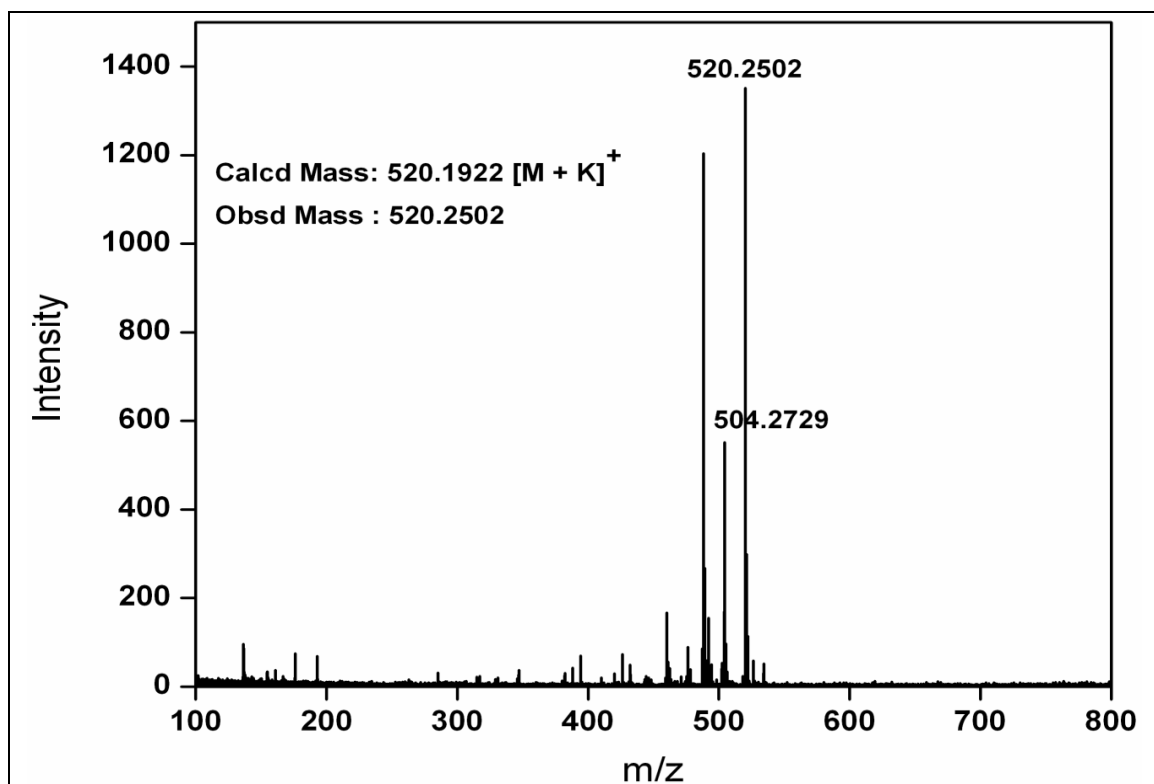
IR of Compound 2
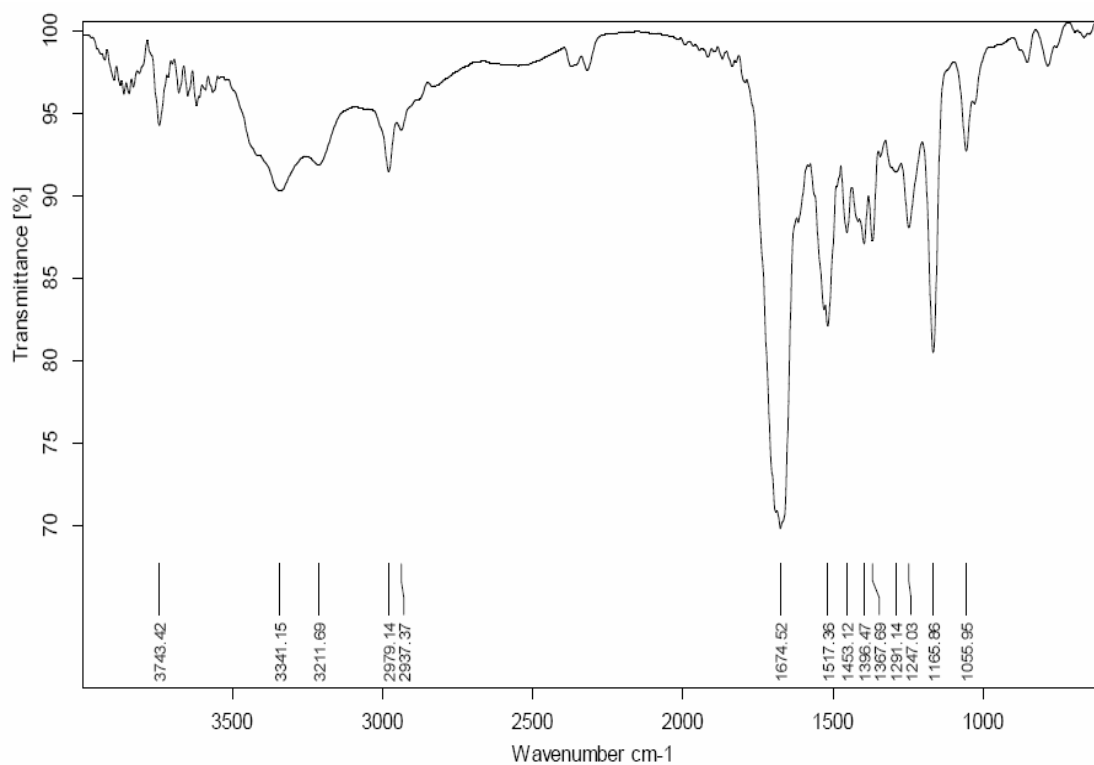
IR of Compound 3
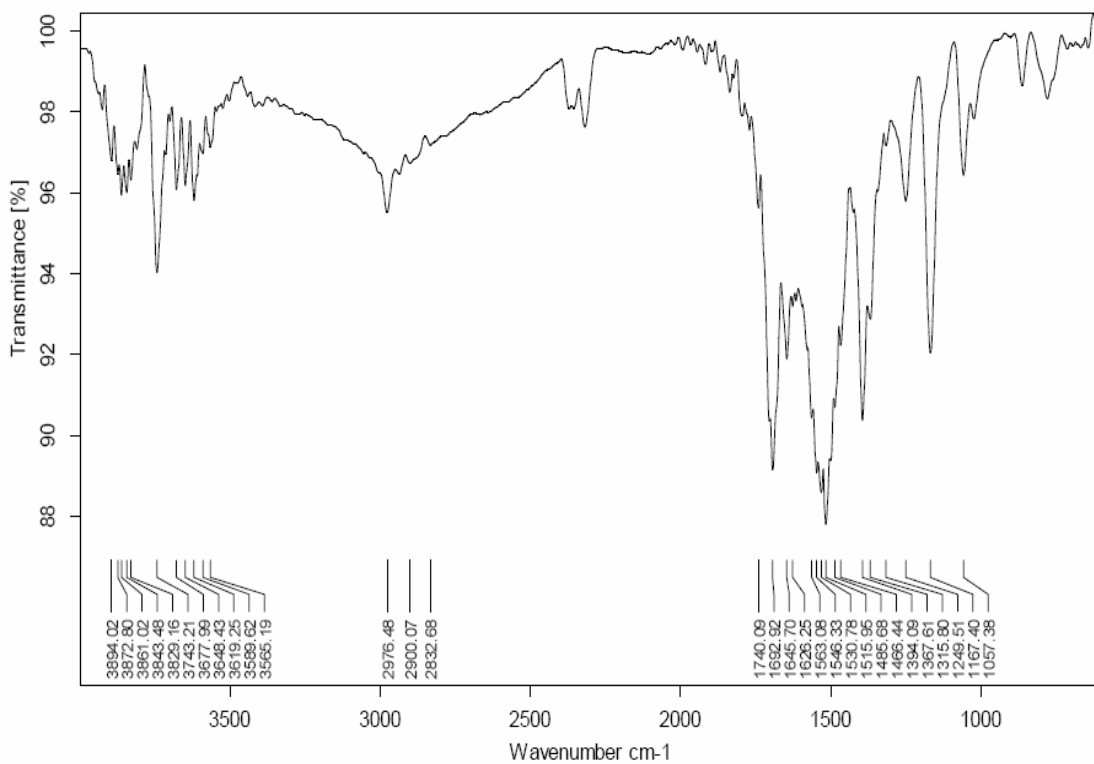
IR of Compound 8
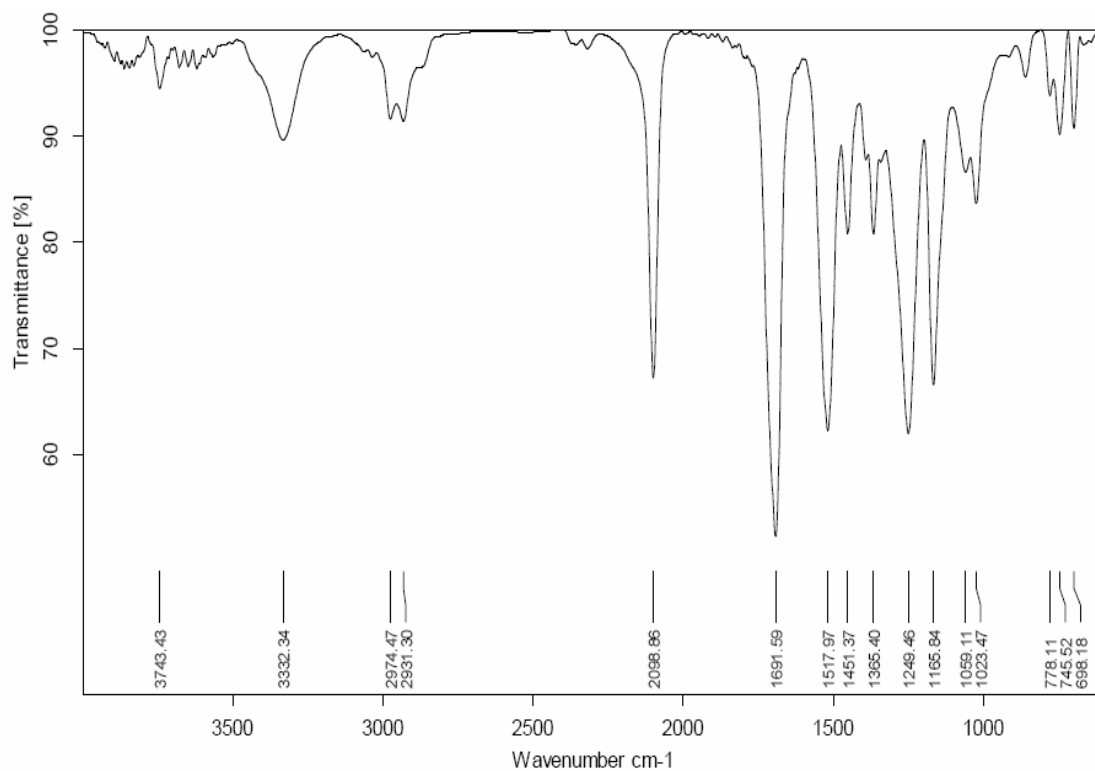
IR of Compound 16

IR of Compound 18

IR of Compound 28

IR of compound 38

HPLC Trace of PNA 1

aeg PNA 1: H-TTACCTCAGT-LysNH$_2$

HPLC Trace of PNA 2

eam-t$_2$ PNA 2: H-TtACCTCAGT-LysNH$_2$

HPLC Trace of PNA 3

eam-t_6 PNA 3: H-TTACC**t**CAGT-LysNH$_2$

HPLC Trace of PNA 4

eam-t_{10} PNA 4: H-TTACCTCAG**t**-LysNH$_2$

HPLC Trace of PNA 5

eam-t$_{2,6}$ PNA 5: H-Tt̲ACCt̲CAGT-LysNH$_2$

9.80

mAU

Time (min)

HPLC Trace of PNA 6

egd-t$^{\gamma}_2$ PNA 6: H-Tt̲$^{\gamma}$ACCTCAGT-LysNH$_2$

10.40

mAU

Time (min)

HPLC Trace of PNA 7

egd-t$^{\gamma}_{6}$ PNA 7: H-TTACCt$^{\gamma}$CAGT-LysNH$_2$

10.38

HPLC Trace of PNA 8

egd-t$^{\gamma}_{10}$ PNA 8: H-TTACCTCAGt$^{\gamma}$-LysNH$_2$

9.82

144

HPLC Trace of PNA 9

egd-t$^\gamma_{2,6}$ PNA 9: H-T$\underline{t}^\gamma$ACC$\underline{t}^\gamma$CAGT-LysNH$_2$

10.07

HPLC Trace of PNA 10

egd-t$^\gamma_{2,6,10}$ PNA 10: H-T$\underline{t}^\gamma$ACC$\underline{t}^\gamma$CAG$\underline{t}^\gamma$-LysNH$_2$

10.21

HPLC Trace of PNA 11

azm-T$^\gamma_2$ PNA 11: H-TT$^\gamma$ACCTCAGT-LysNH$_2$

9.74

HPLC Trace of PNA 12

azm-T$^\gamma_6$ PNA 12: H-TTACCT$^\gamma$CAGT-LysNH$_2$

10.08

HPLC Trace of PNA 13

azm-T$^\gamma_{10}$ PNA 13: H-TTACCTCAG$\underline{T}^\gamma$-LysNH$_2$

9.63

HPLC Trace of PNA 14

azm-T$^\gamma_{2,6}$ PNA 14: H-T$\underline{T}^\gamma$ACC$\underline{T}^\gamma$CAGT-LysNH$_2$
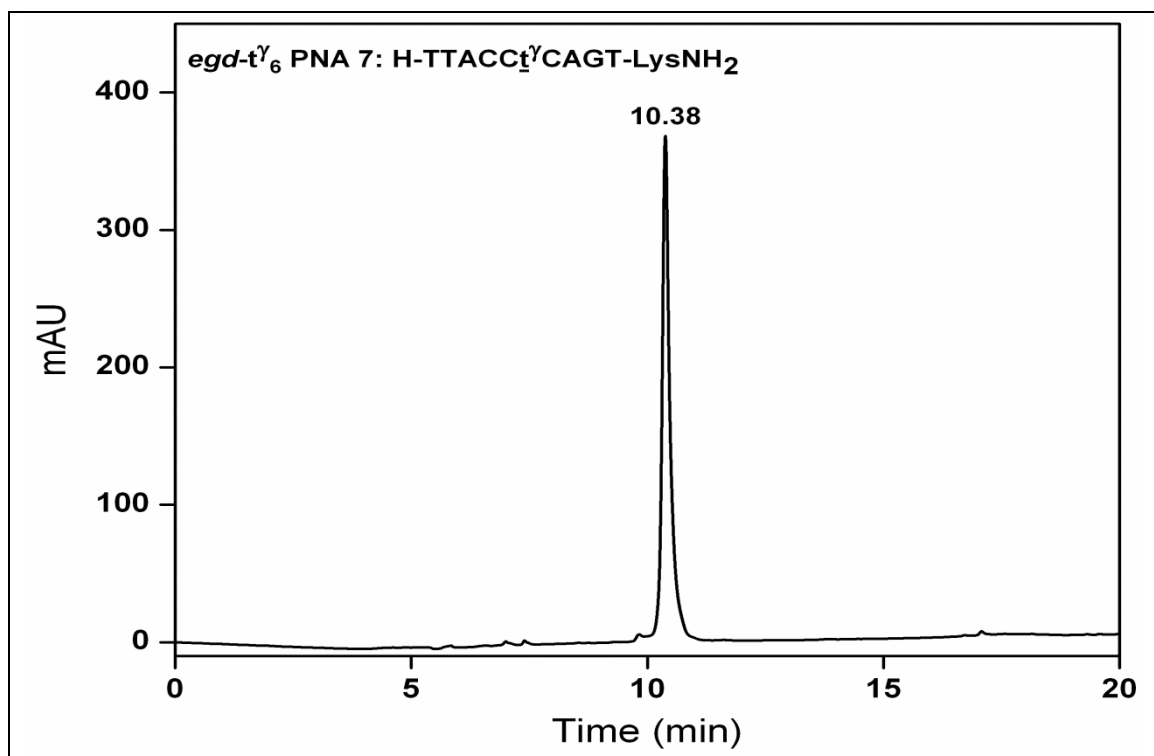
10.12
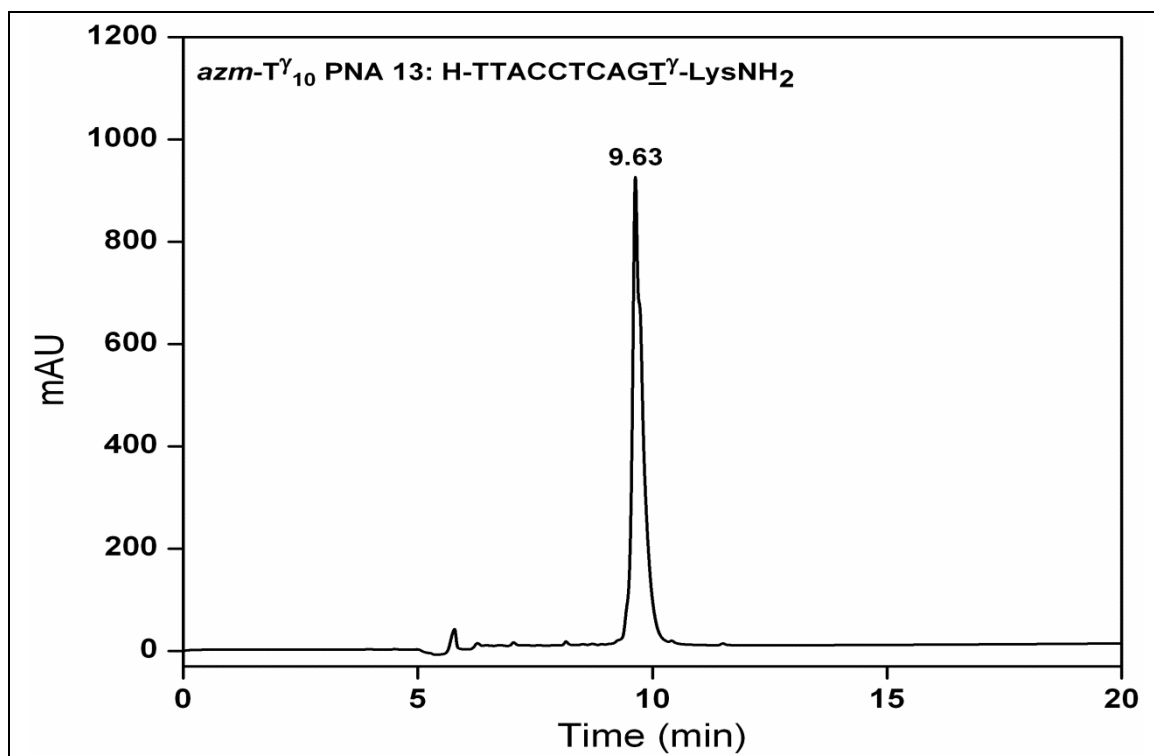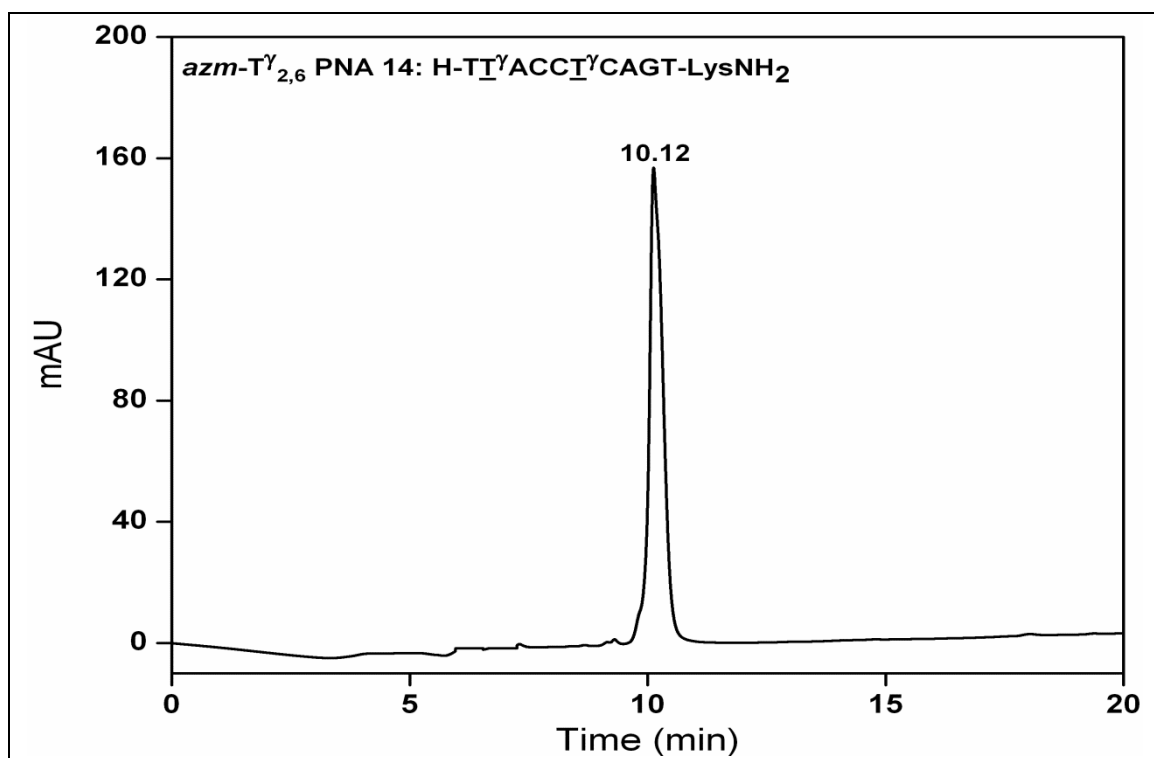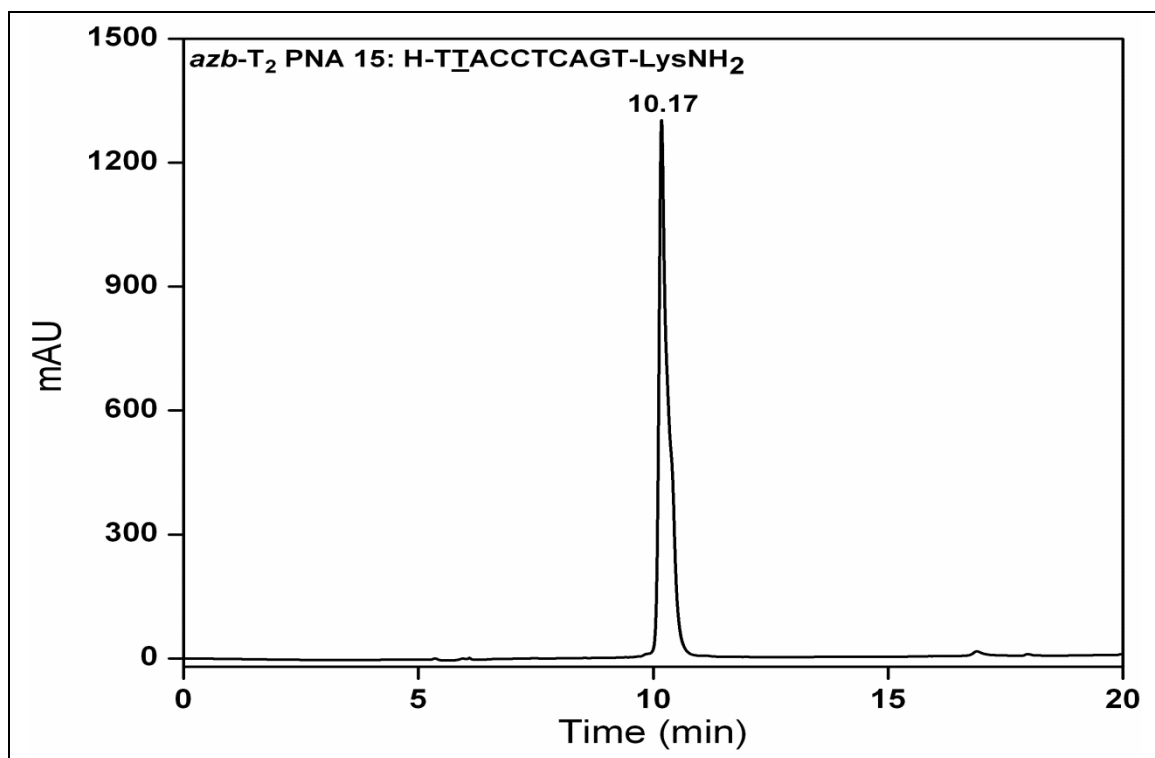
HPLC Trace of PNA 15

HPLC Trace of PNA 16
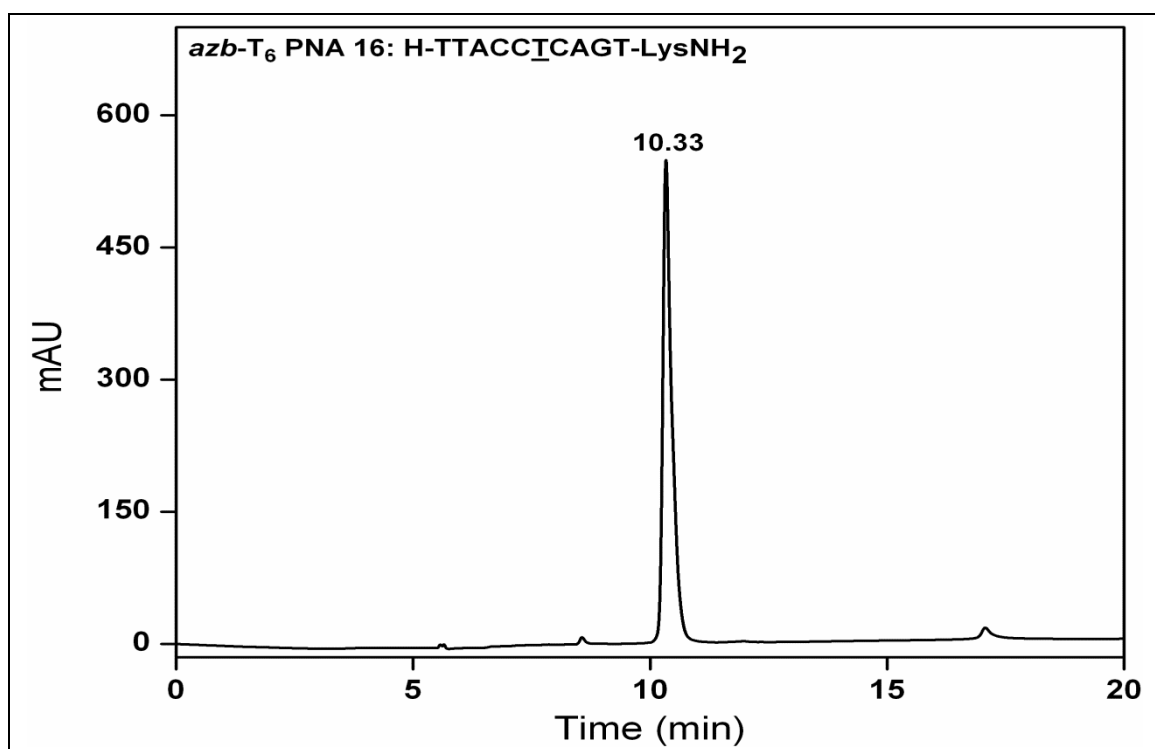
HPLC Trace of PNA 17

azb-T$_{10}$ PNA 17: H-TTACCTCAG<u>T</u>-LysNH$_2$

10.30

HPLC Trace of PNA 18

azb-t$_{2,6}$ PNA18: H-T<u>T</u>ACC<u>T</u>CAGT-LysNH$_2$

10.26

HPLC Trace of PNA 19

azm-t$^f_2$ PNA 19: H-Tt$\underline{t}^f$ACCTCAGT-LysNH$_2$
13.67

HPLC Trace of PNA 20

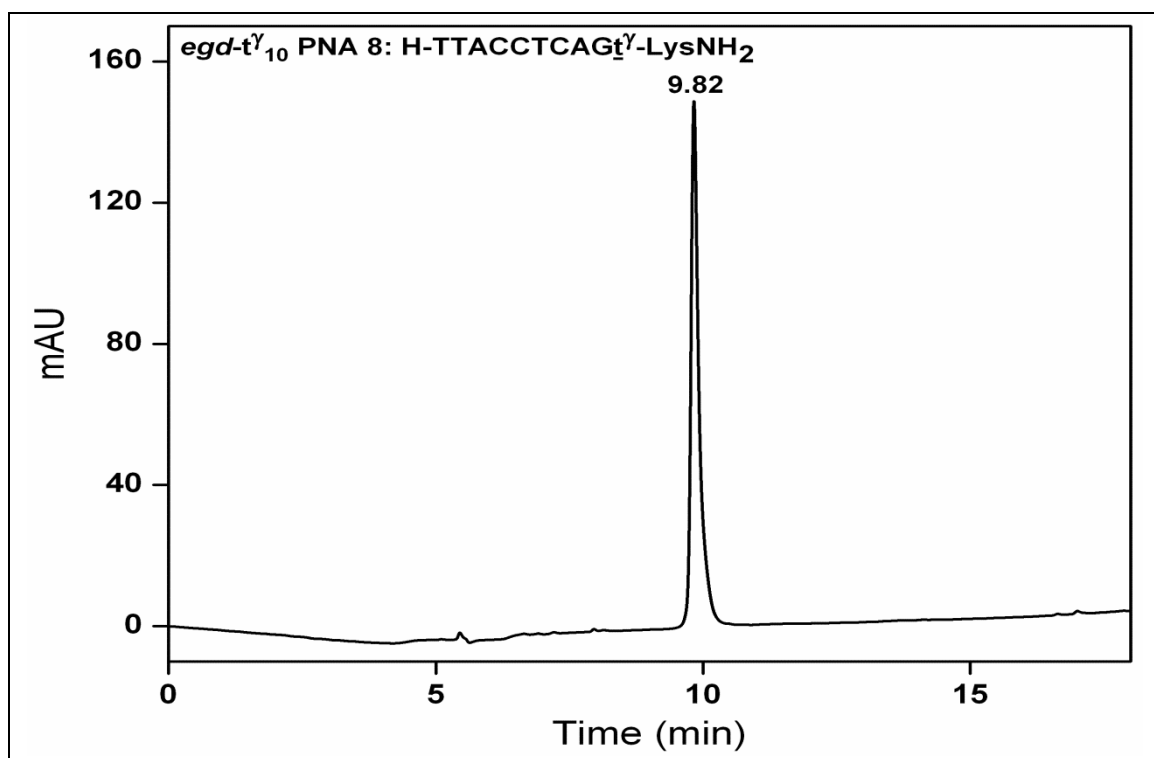*azm*-t$^f_6$ PNA 20: H-TTACC$\underline{t}^f$CAGT-LysNH$_2$
12.95
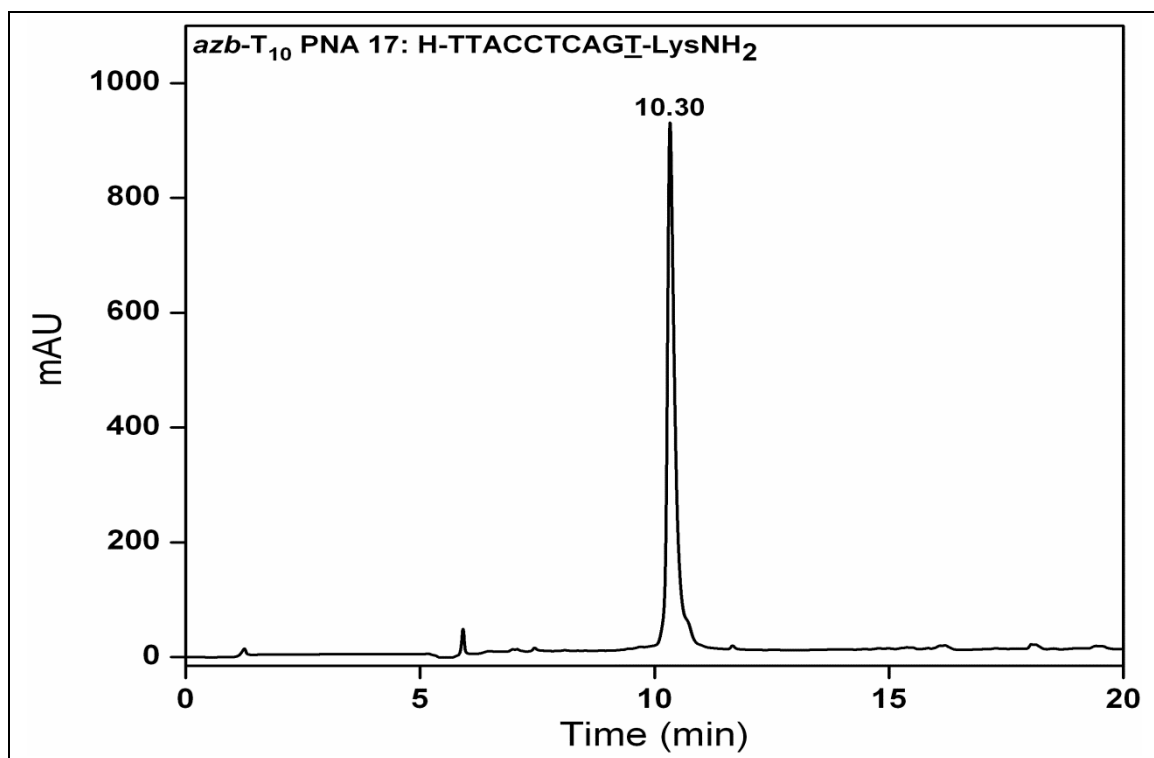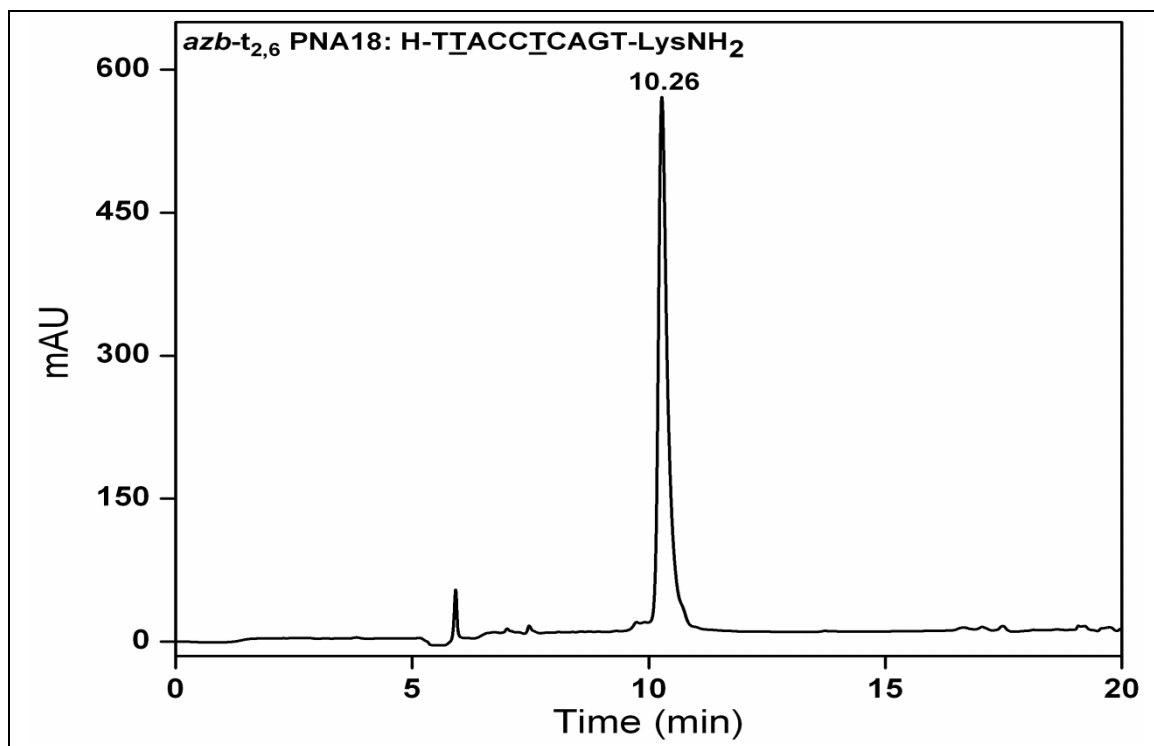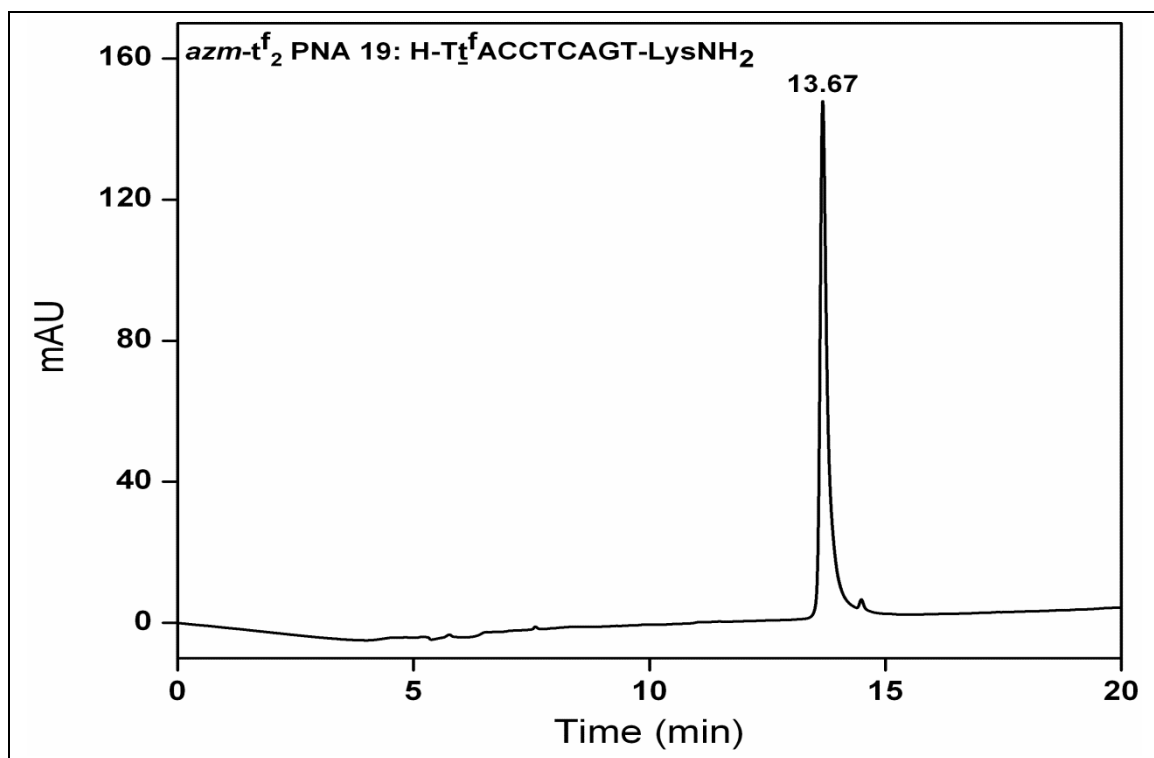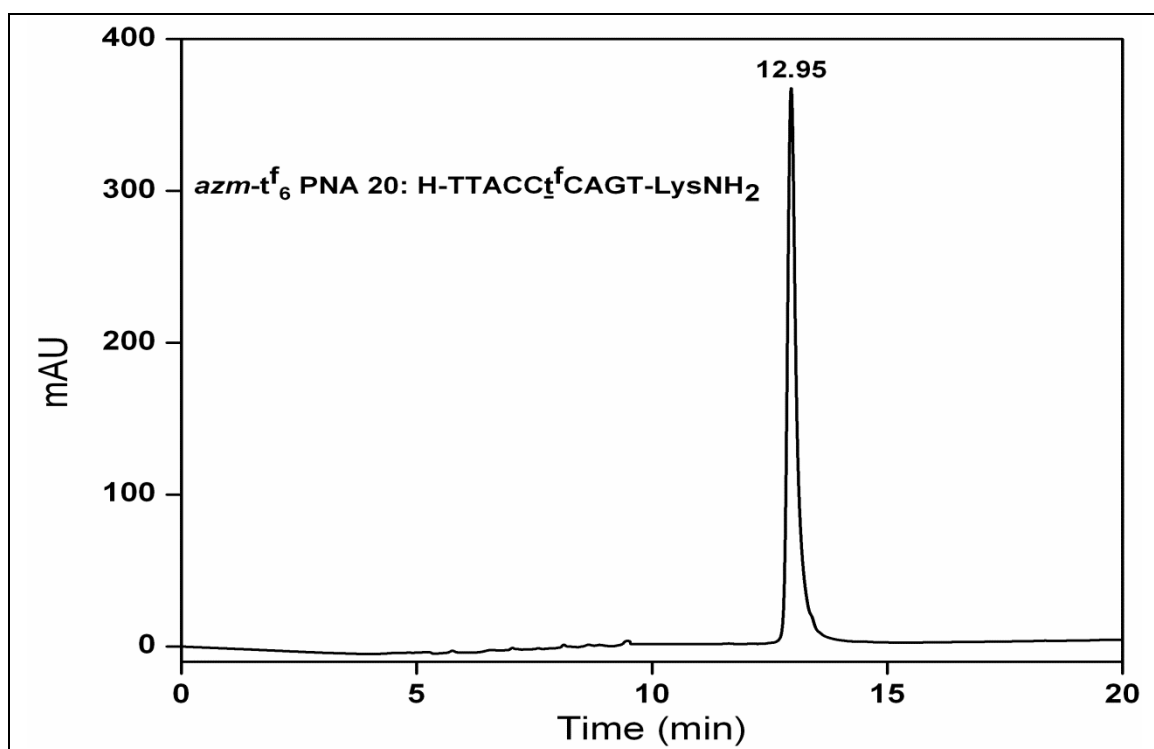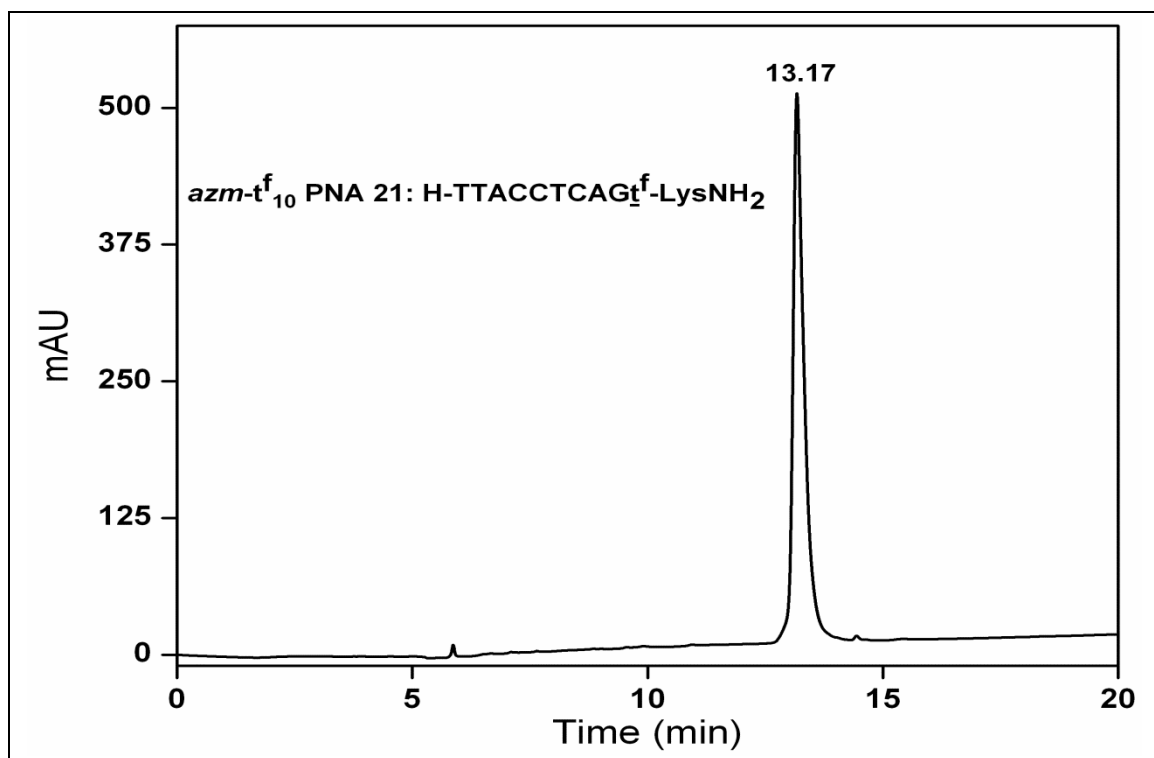
HPLC Trace of PNA 21

HPLC Trace of PNA 22
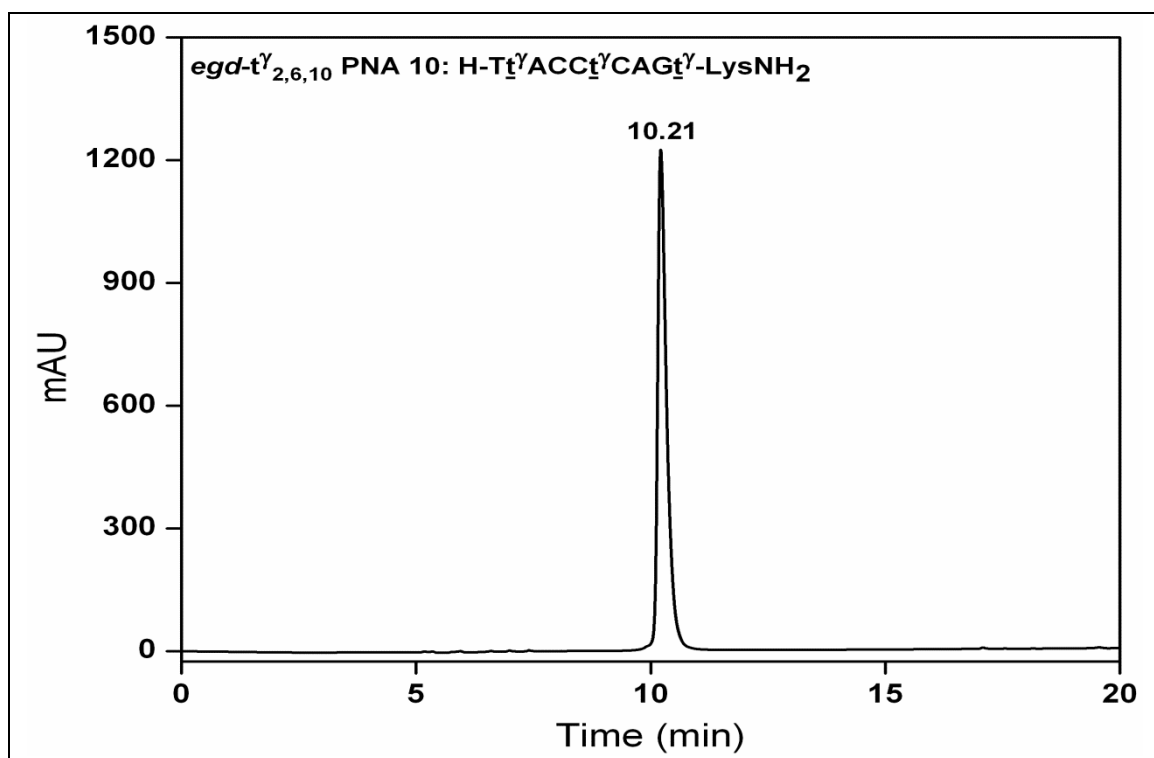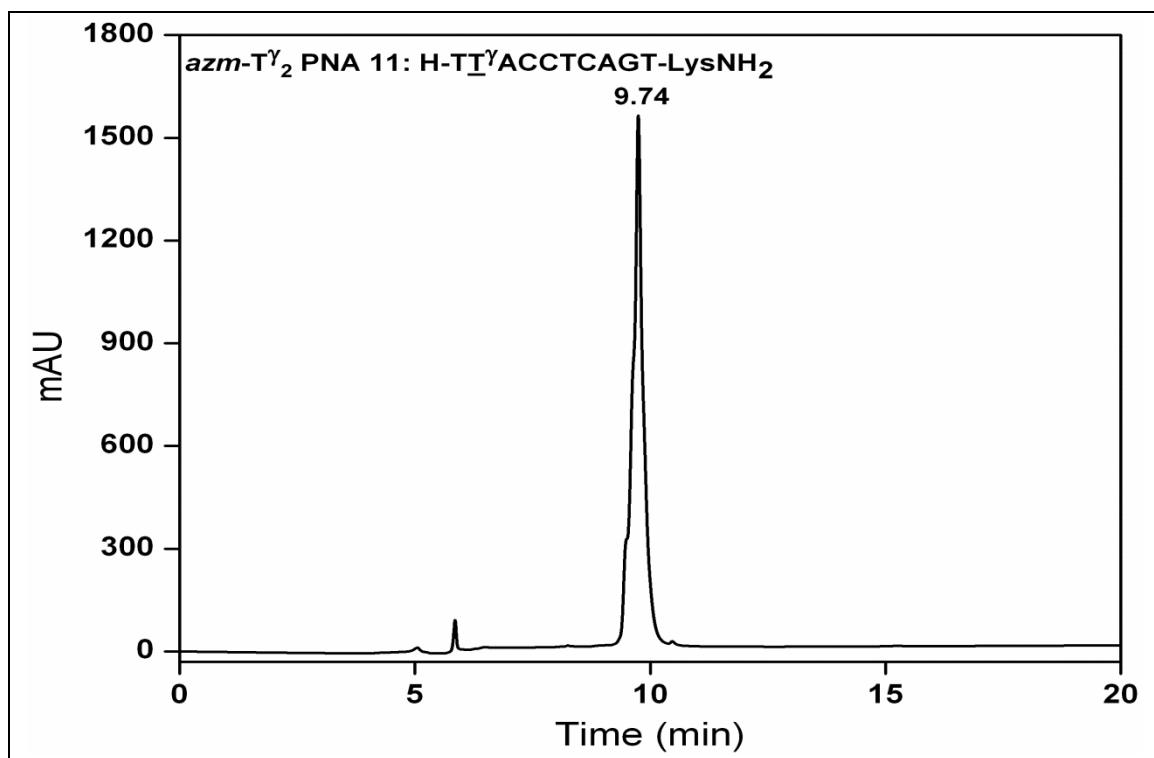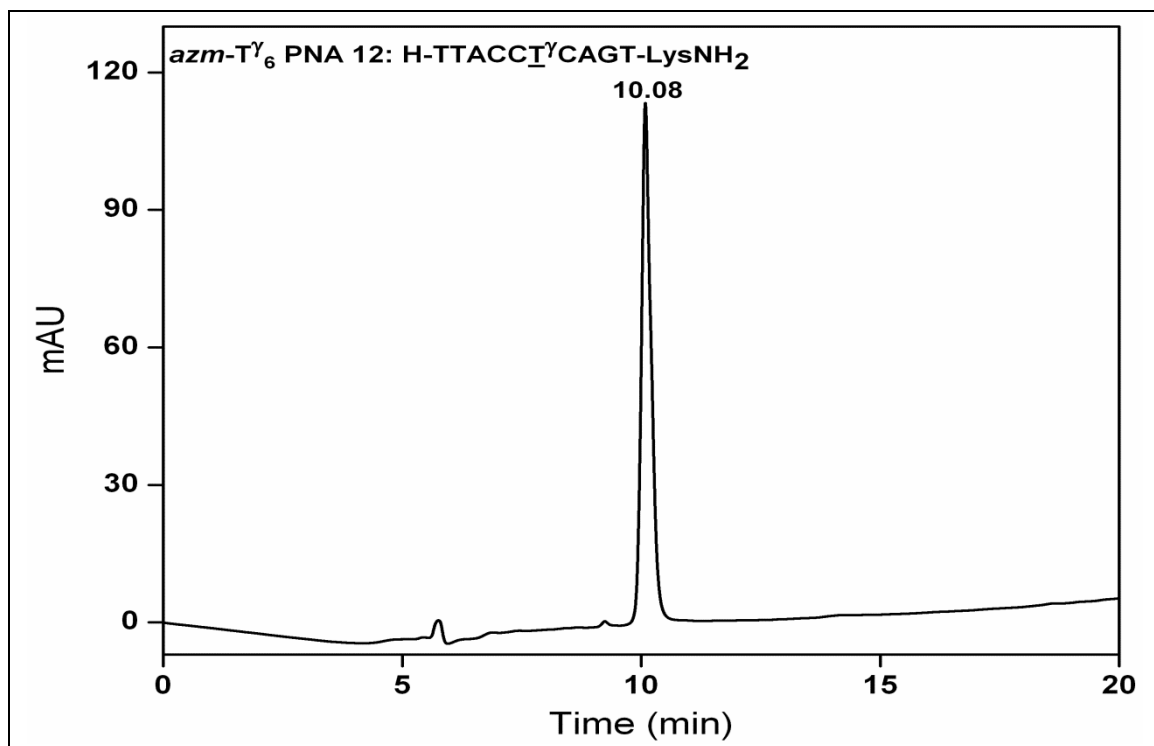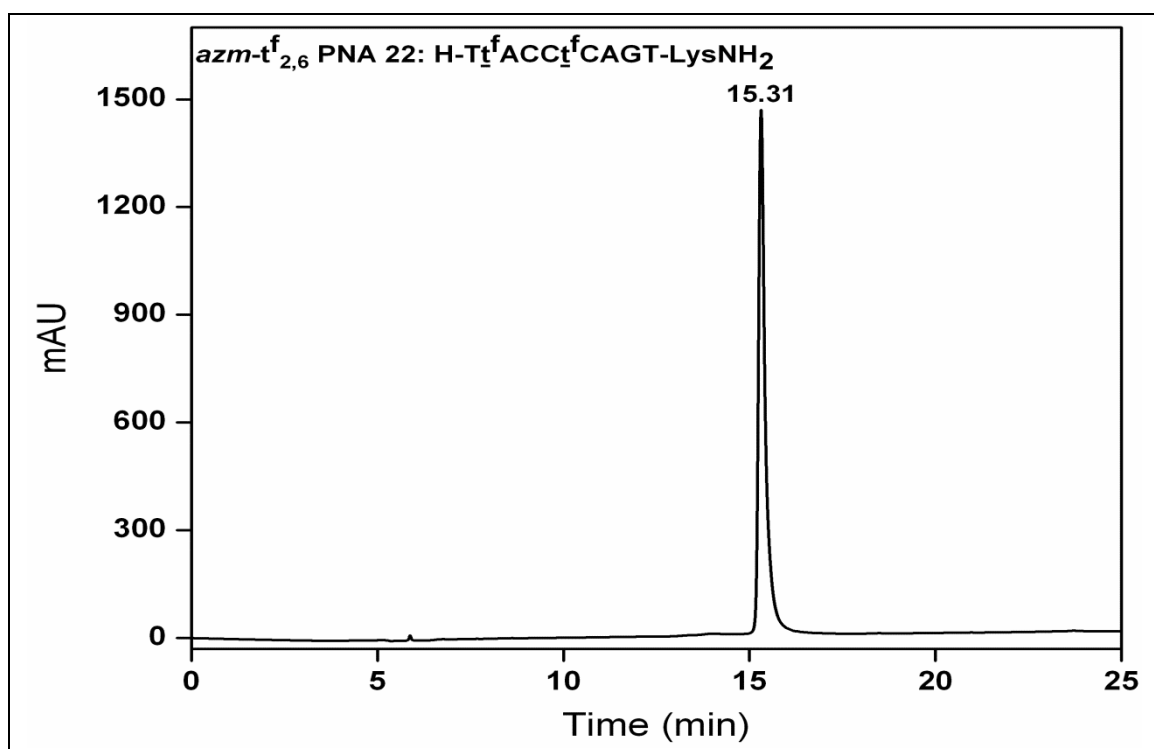
HPLC Trace of PNA 23
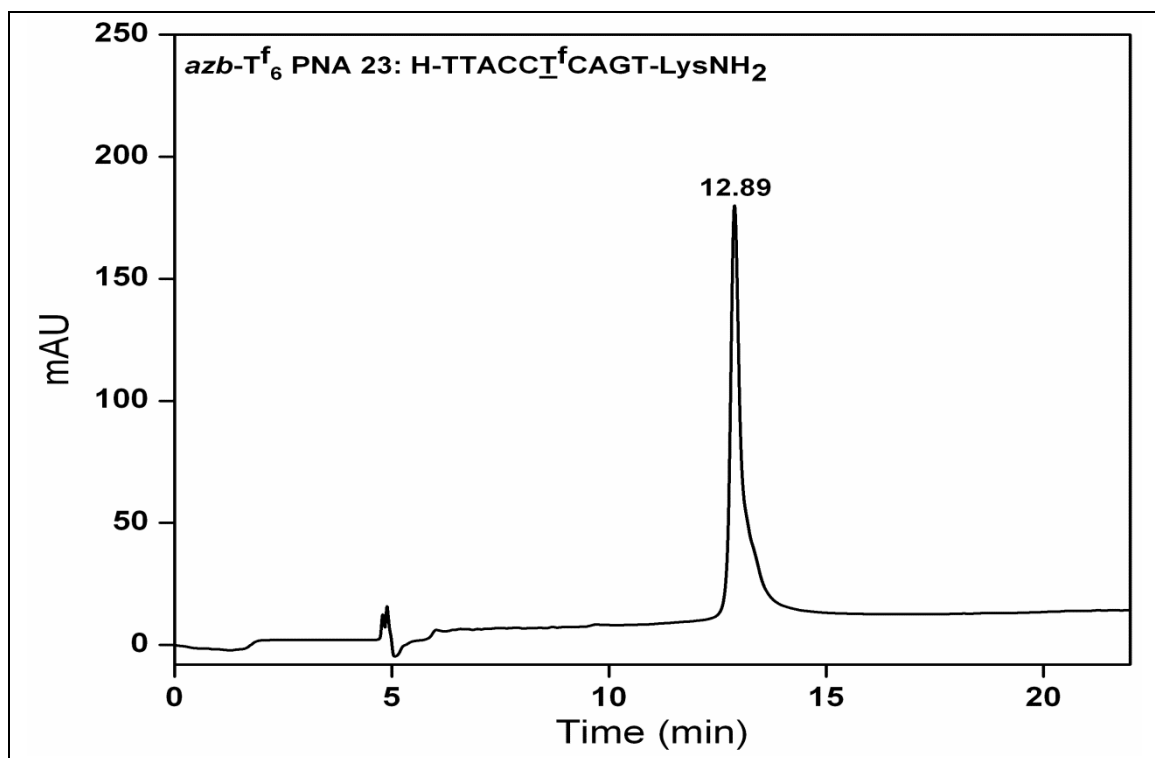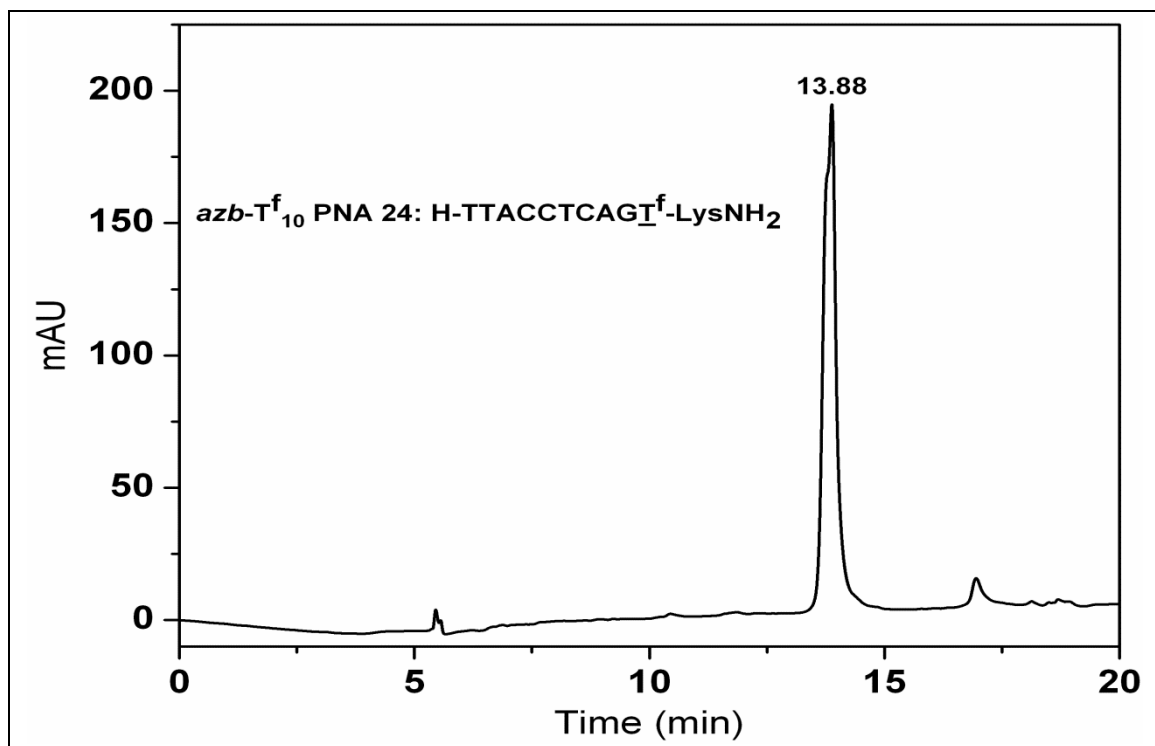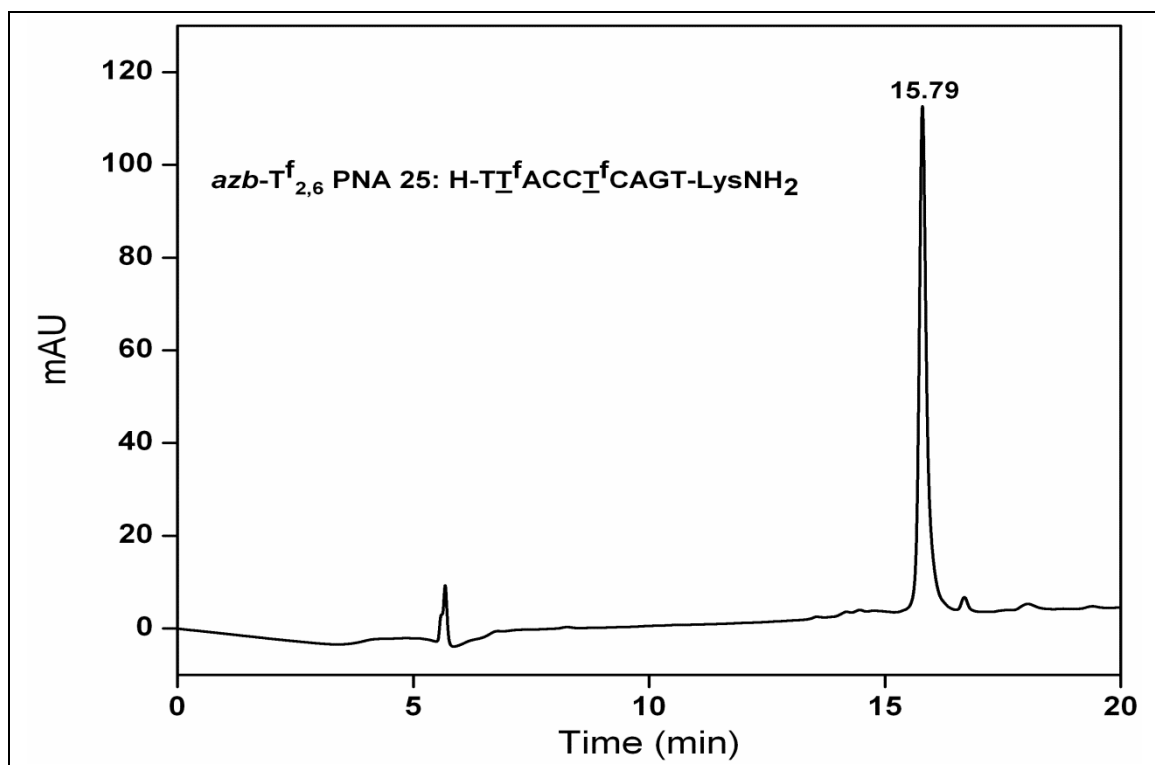
azb-T$^f_6$ PNA 23: H-TTACC$\underline{T}^f$CAGT-LysNH$_2$

HPLC Trace of PNA 24

azb-T$^f_{10}$ PNA 24: H-TTACCTCAG$\underline{T}^f$-LysNH$_2$

HPLC Trace of PNA 25

azb-T$^f_{2,6}$ PNA 25: H-TT̲^{f}ACCT̲^{f}CAGT-LysNH$_2$

15.79

MALDI-TOF Mass of PNA 1

Calcd Mass: 2805.1750 [M + H]$^+$
Obsd Mass: 2805.0652
2805.0652

MALDI-TOF Mass of PNA 2

Calcd Mass: 2848.2172 [M + H]$^+$
Obsd Mass: 2848.3081
2848.3081

MALDI-TOF Mass of PNA 3

Calcd Mass: 2848.2172 [M + H]$^{+}$
Obsd Mass: 2848.2515

2848.2515

MALDI-TOF Mass of PNA 4

Calcd Mass: 2848.2172 [M + H]$^{+}$
Obsd Mass: 2848.8708

2848.8708
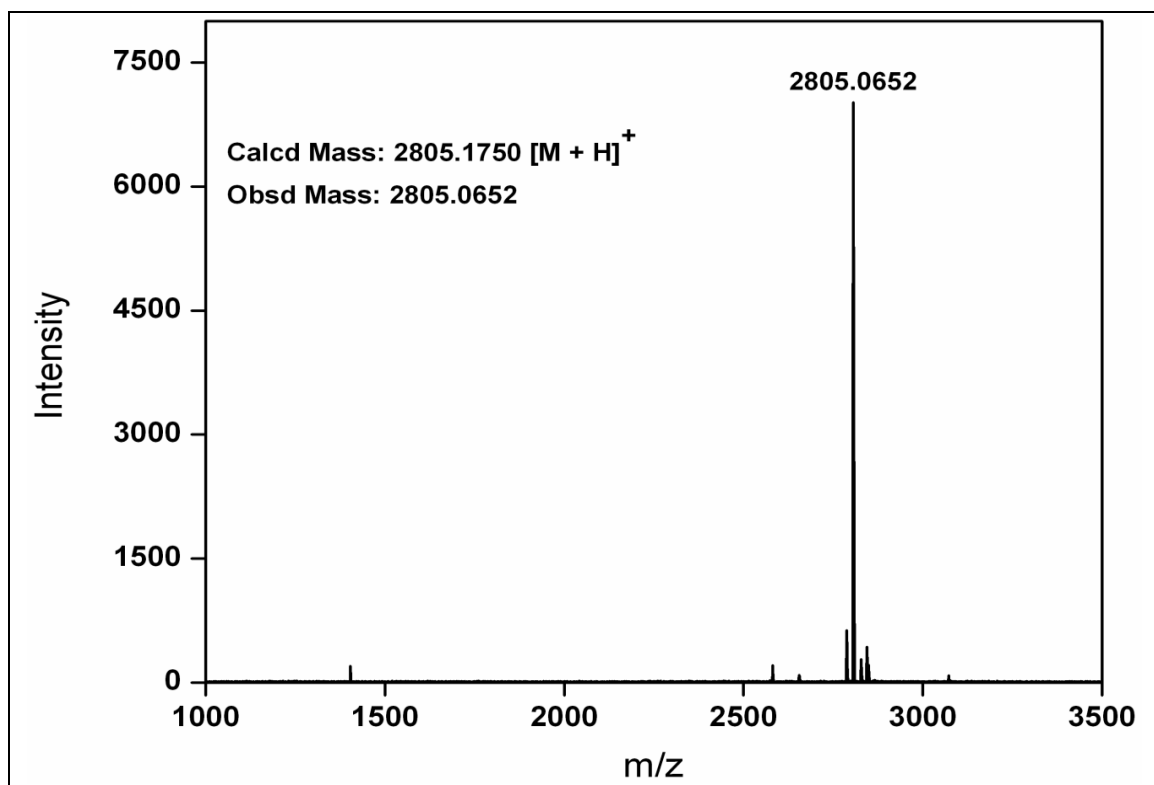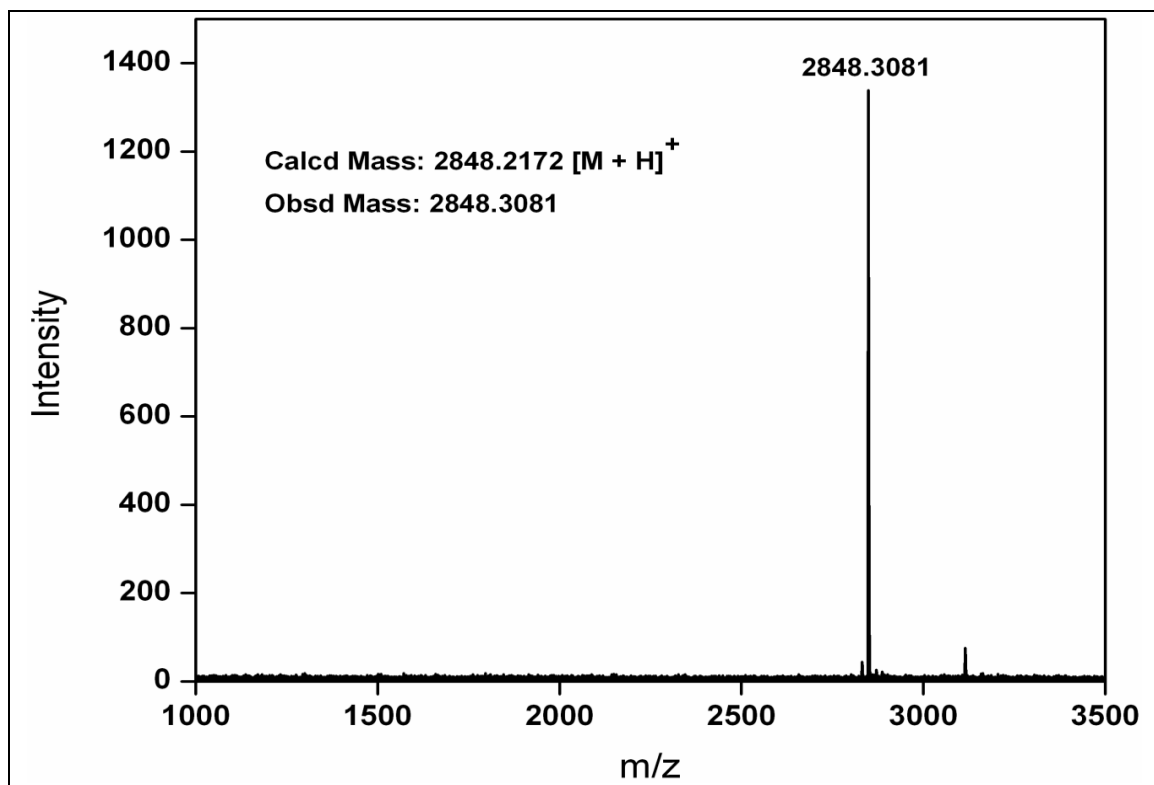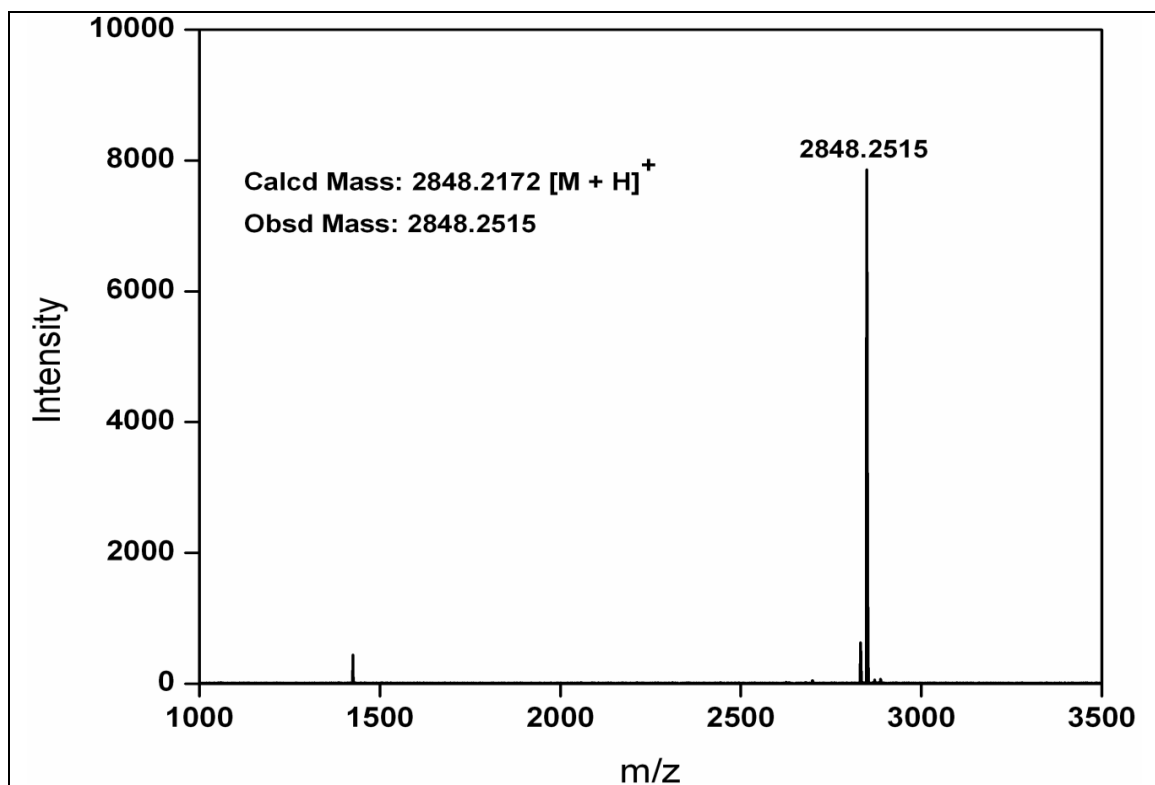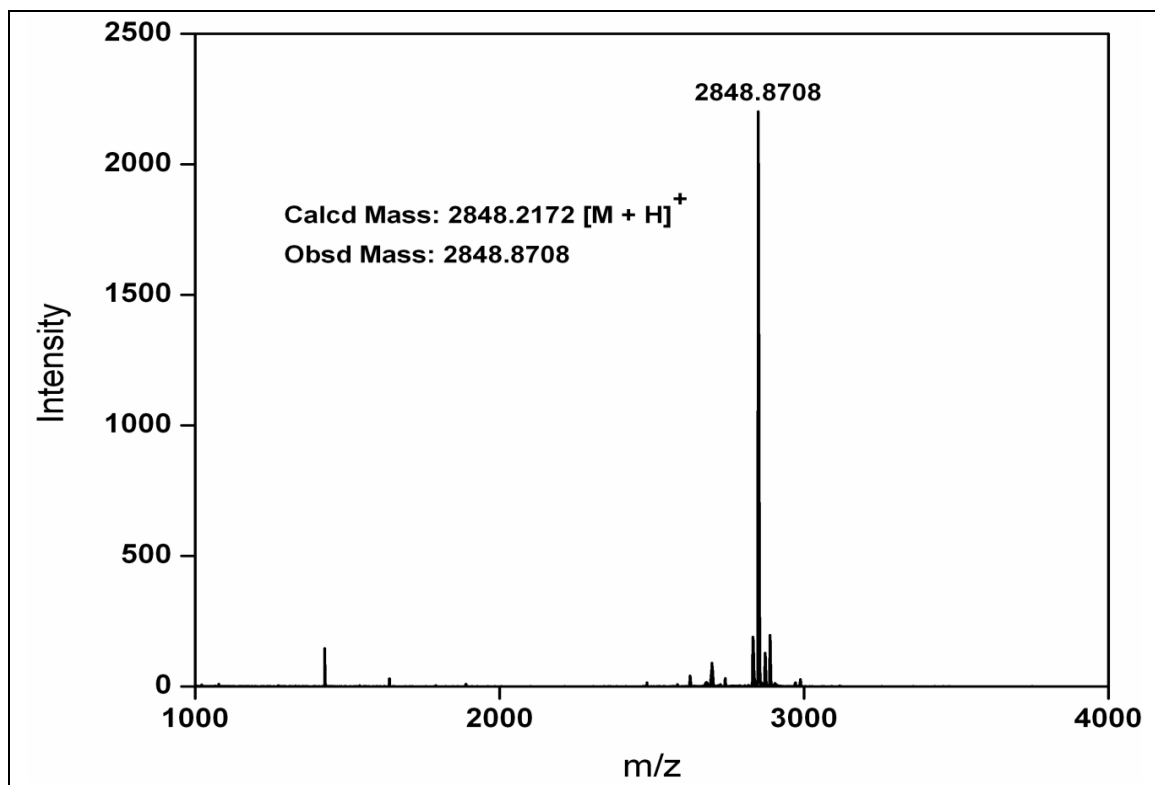
MALDI-TOF Mass of PNA 5
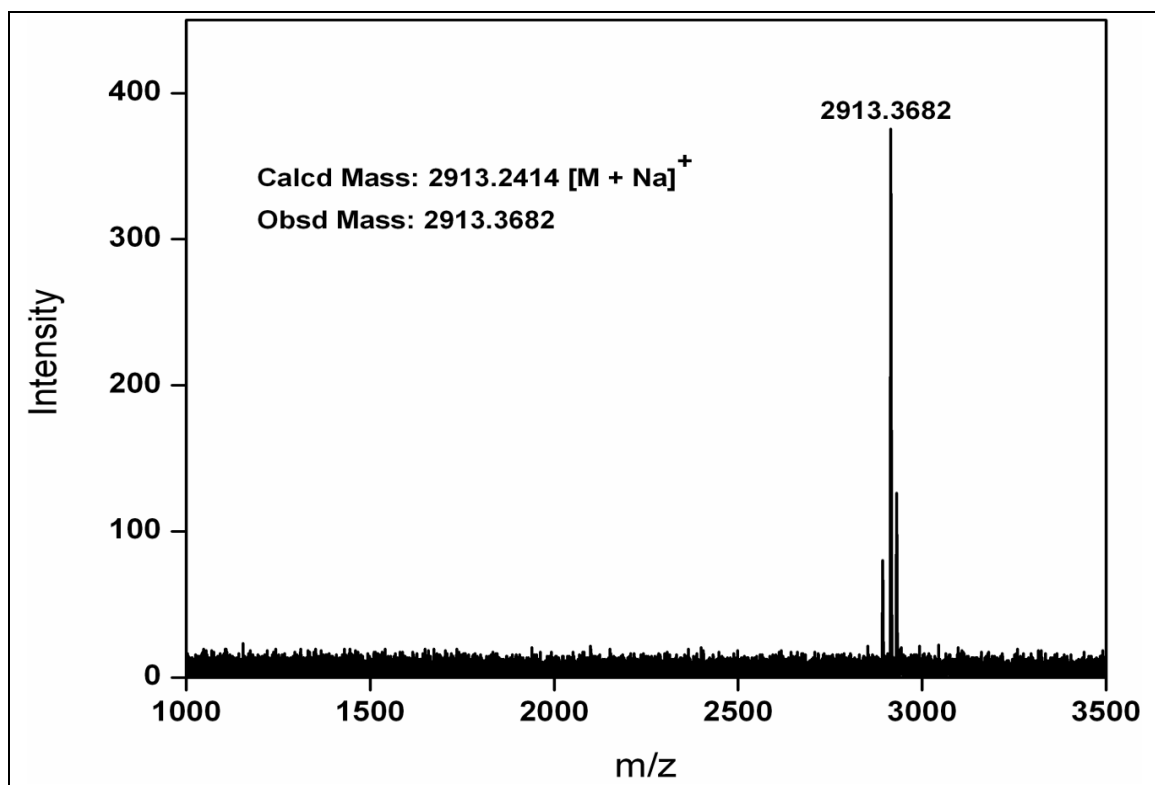
MALDI-TOF Mass of PNA 6
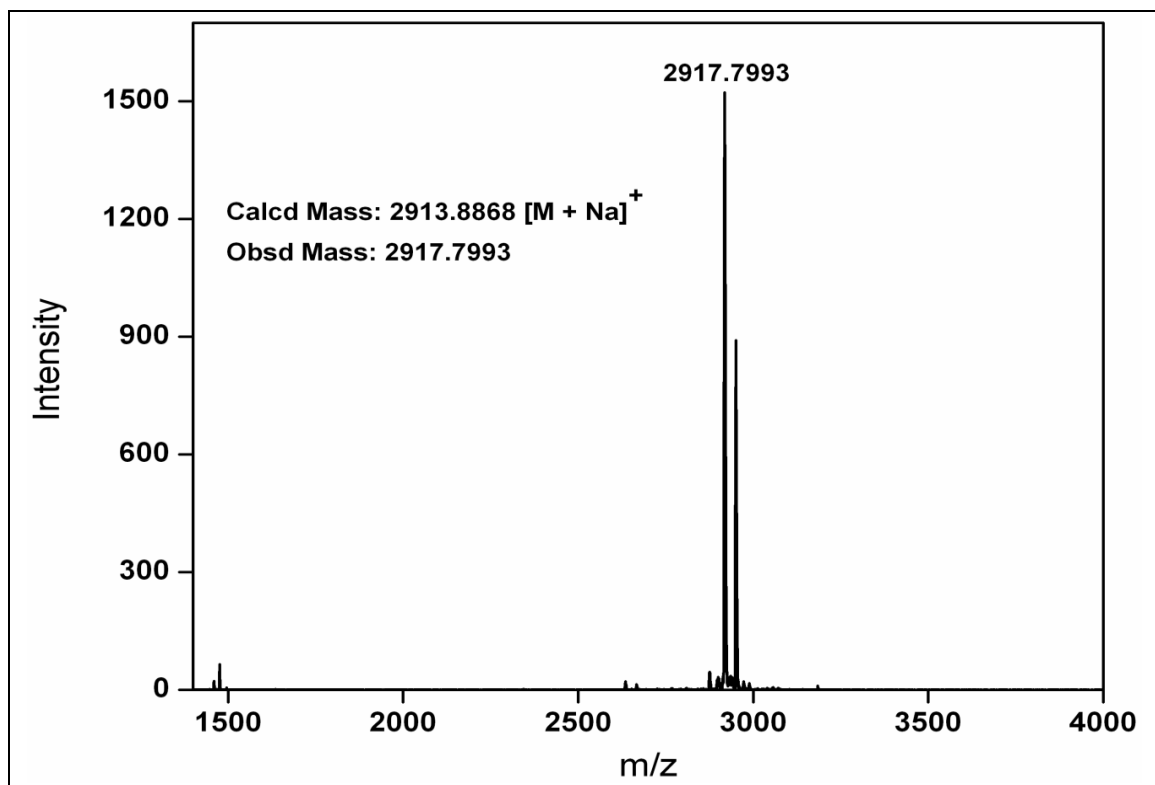
MALDI-TOF Mass of PNA 7
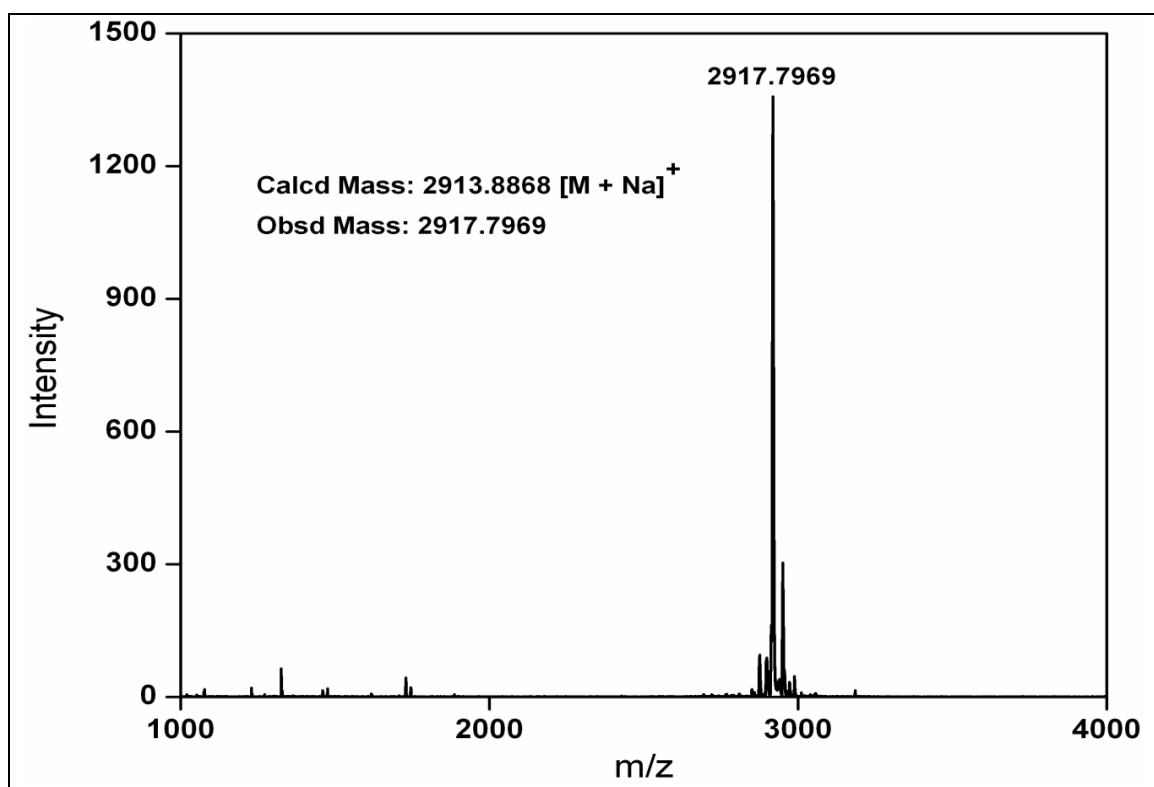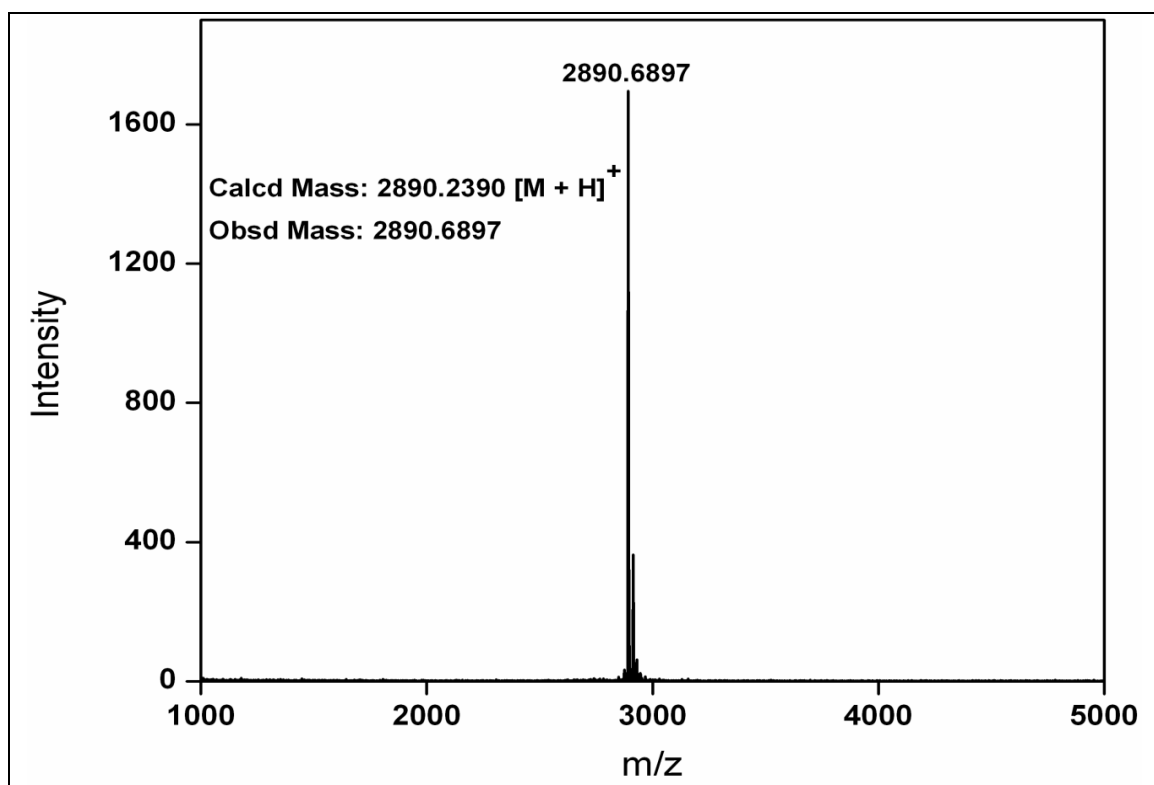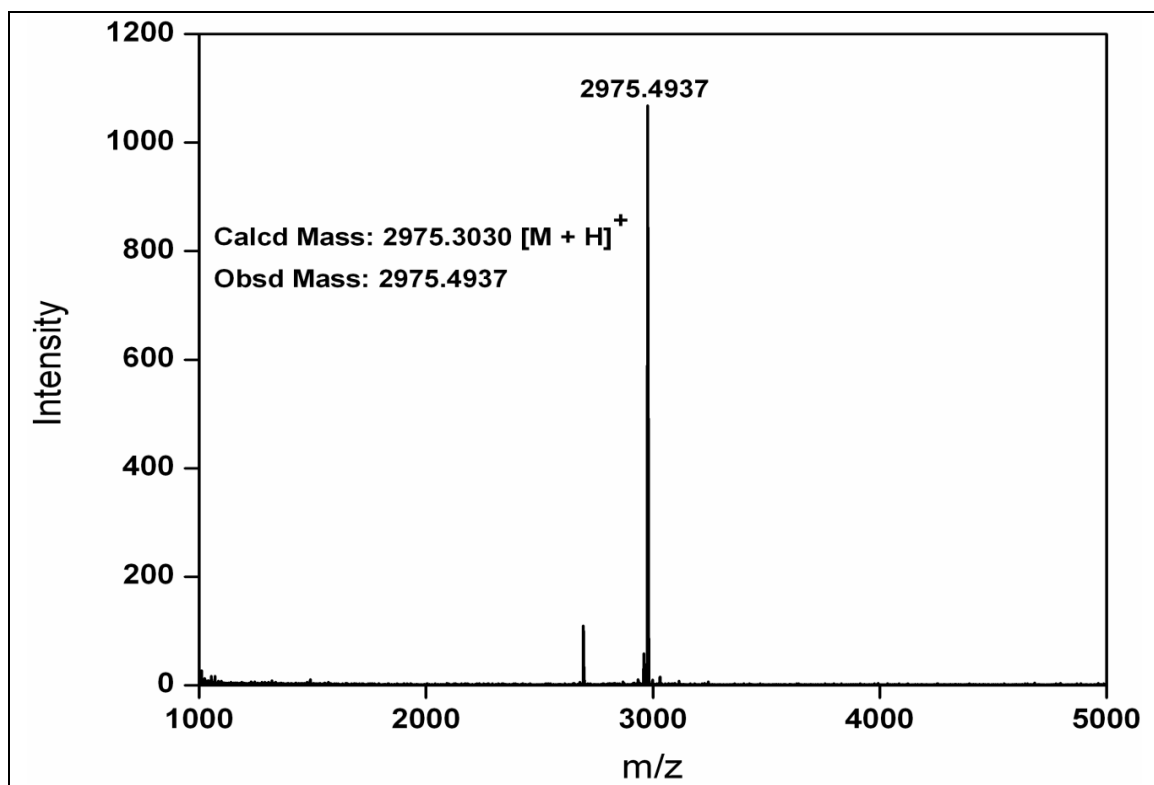
Calcd Mass: 2913.8868 [M + Na]$^+$
Obsd Mass: 2917.7969

Peak: 2917.7969

MALDI-TOF Mass of PNA 8

Calcd Mass: 2890.2390 [M + H]$^+$
Obsd Mass: 2890.6897

Peak: 2890.6897

MALDI-TOF Mass of PNA 9

Calcd Mass: 2975.3030 [M + H]$^{+}$
Obsd Mass: 2975.4937

2975.4937

MALDI-TOF Mass of PNA 10

Calcd Mass: 3060.3670 [M + H]$^{+}$
Obsd Mass: 3060.1923

3060.1923

MALDI-TOF Mass of PNA 11

Calcd Mass: 2901.2312 [M]$^+$
Obsd Mass: 2900.5166

2900.5166

MALDI-TOF Mass of PNA 12

Calcd Mass: 2901.2312 [M]$^+$
Obsd Mass: 2899.7097

2899.7097
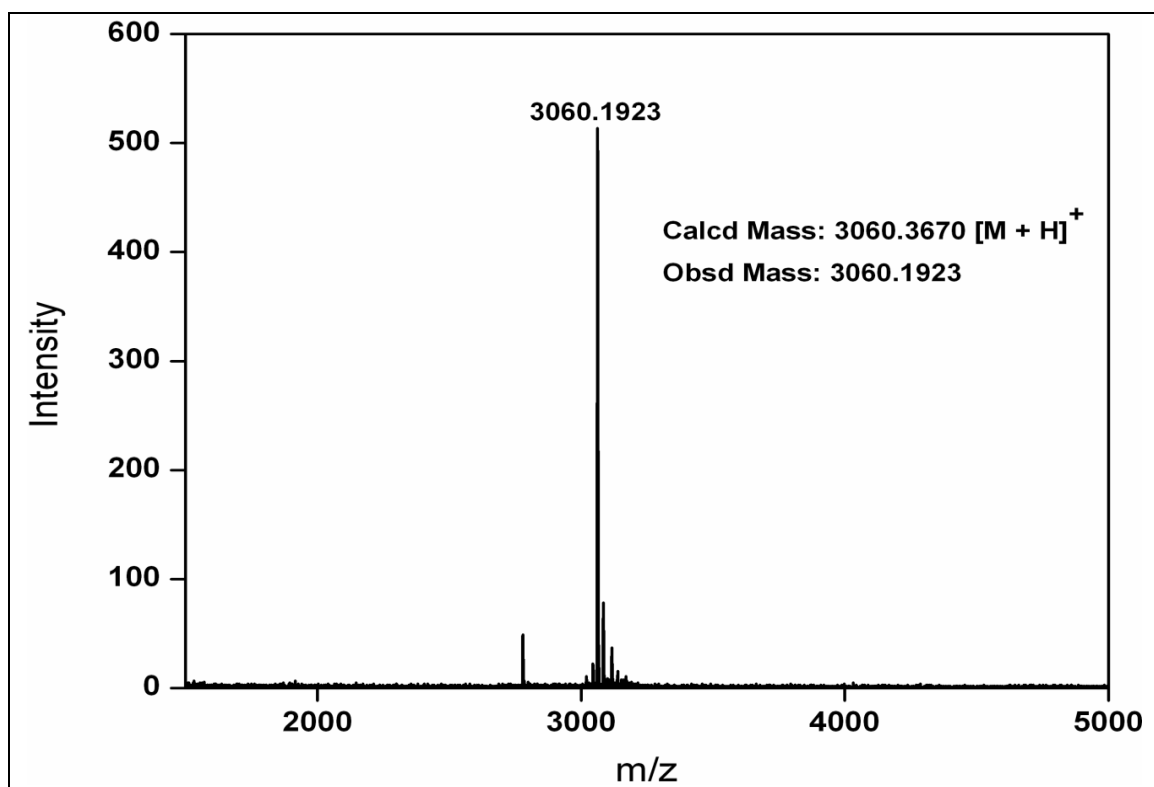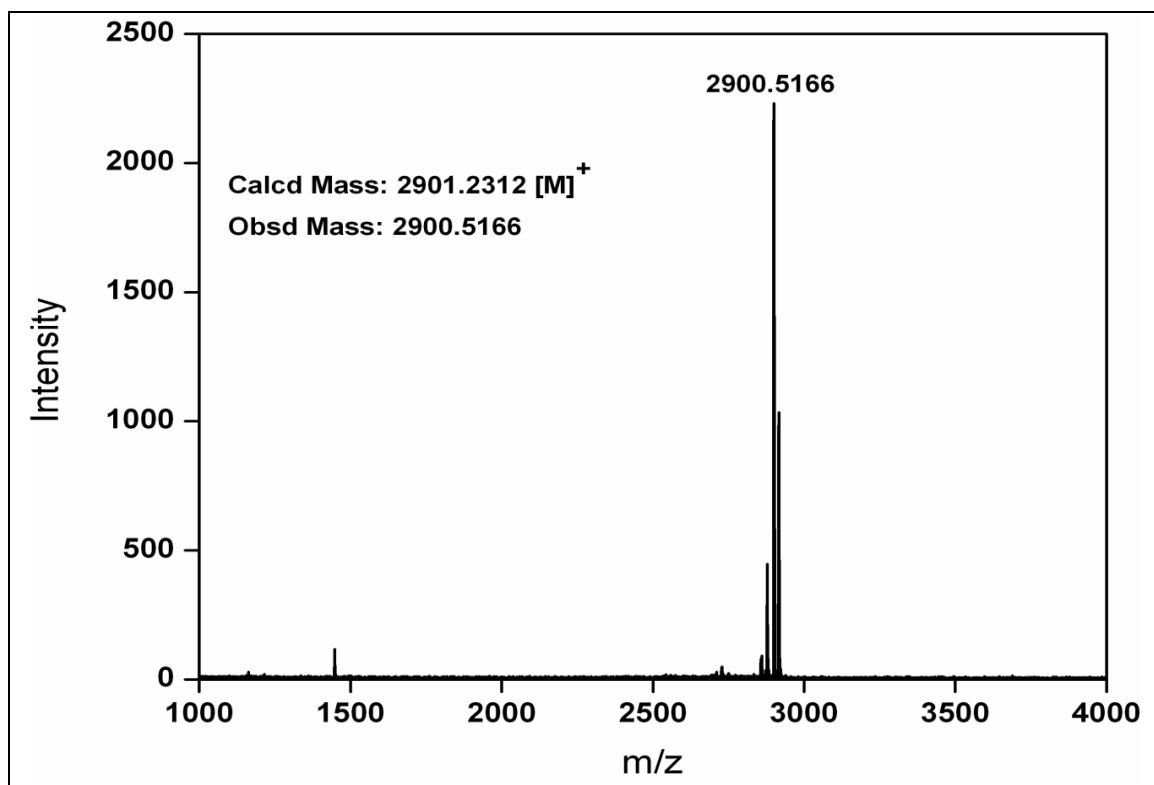
MALDI-TOF Mass of PNA 13

MALDI-TOF Mass of PNA 14
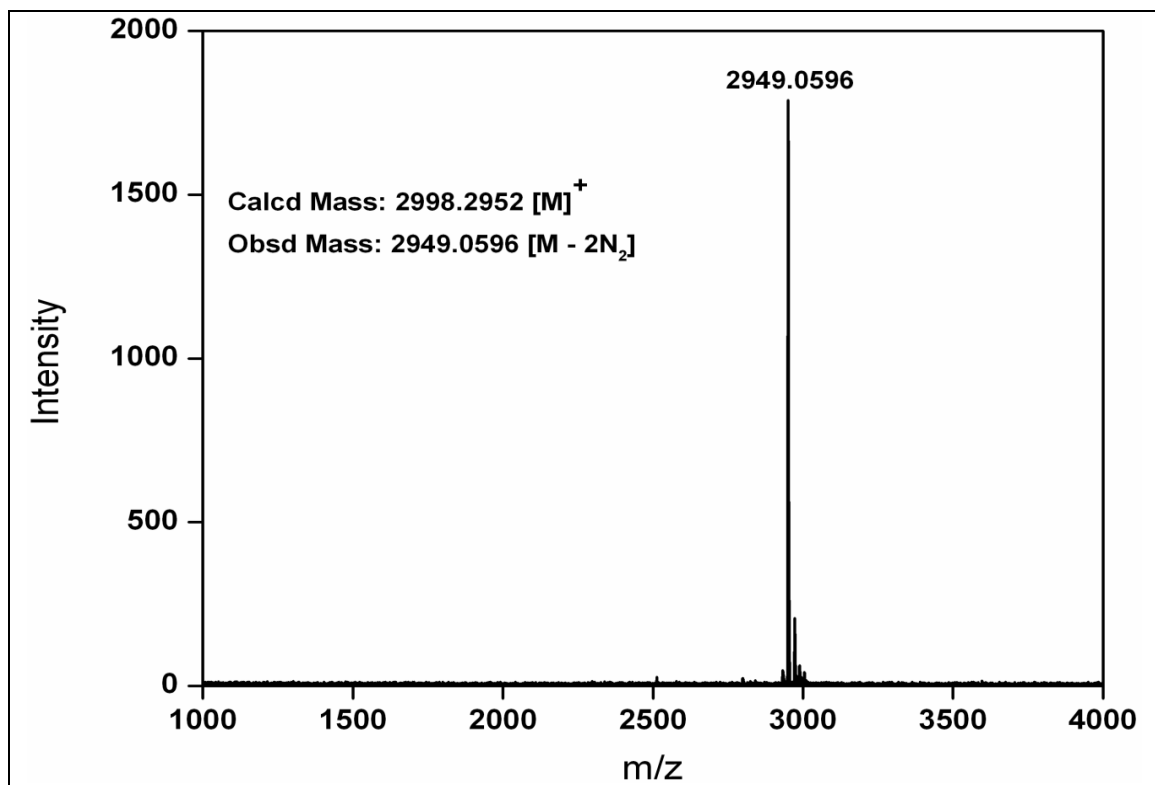
MALDI-TOF Mass of PNA 15
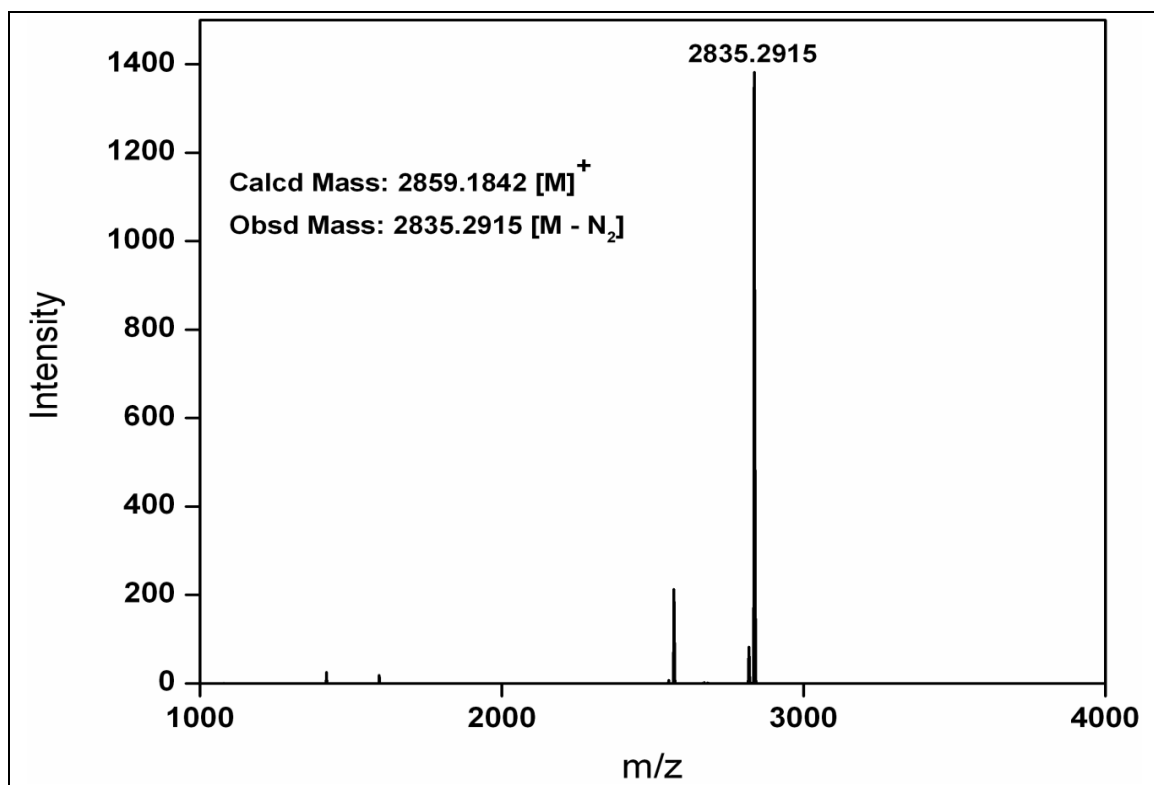
Calcd Mass: 2859.1842 [M]$^+$
Obsd Mass: 2835.2915 [M - N$_2$]

MALDI-TOF Mass of PNA 16

Calcd Mass: 2859.2312 [M]$^+$
Obsd Mass: 2835.3149 [M - N$_2$]

MALDI-TOF Mass of PNA 17

Calcd Mass: 2859.2312 [M]$^+$

Obsd Mass: 2835.2590 [M - N$_2$]

MALDI-TOF Mass of PNA 18

Calcd Mass: 2914.2952 [M]$^+$

Obsd Mass: 2865.3467 [M - 2N$_2$]

MALDI-TOF Mass of PNA 19

Calcd Mass: 3315.3289 [M + H]$^+$
Obsd Mass: 3318.2996

MALDI-TOF Mass of PNA 20

Calcd Mass: 3315.3289 [M + H]$^+$
Obsd Mass: 3315.9607

MALDI-TOF Mass of PNA 21

Calcd Mass: 3825.4829 [M + H]$^+$
Obsd Mass: 3828.1565

MALDI-TOF Mass of PNA 22

Calcd Mass: 3273.2820 [M + H]$^+$
Obsd Mass: 3275.2666
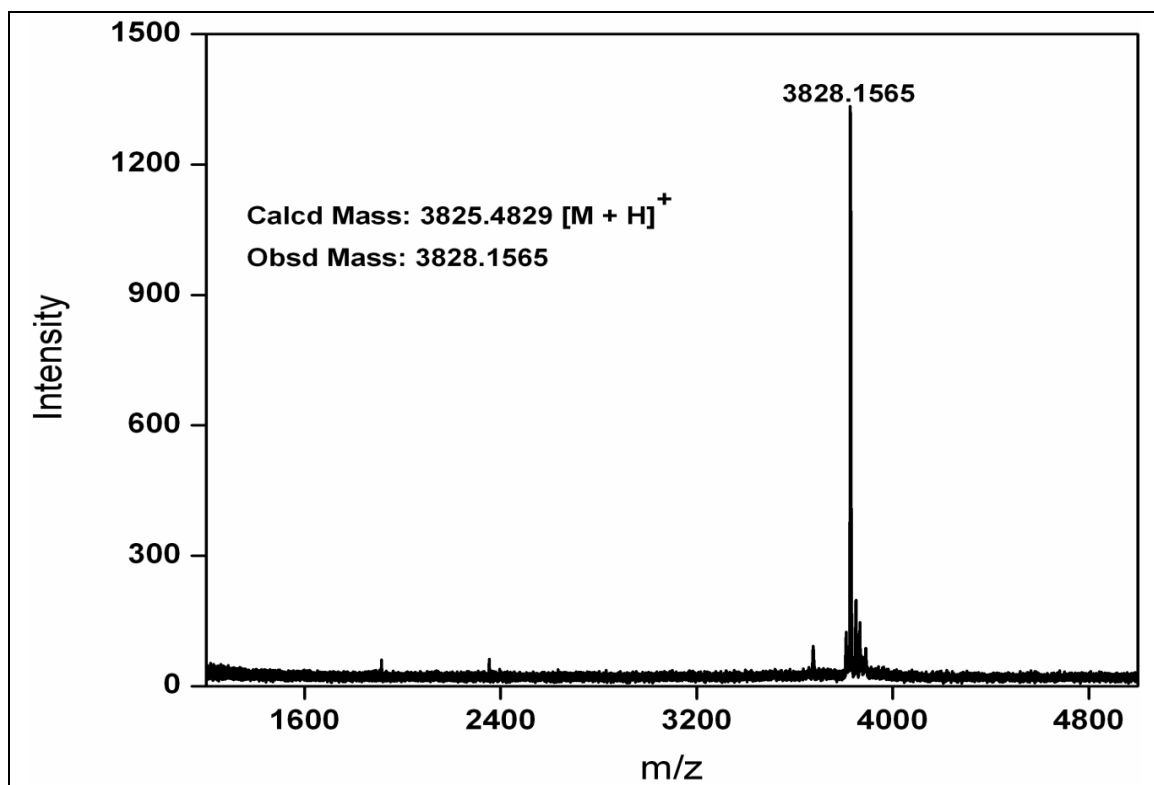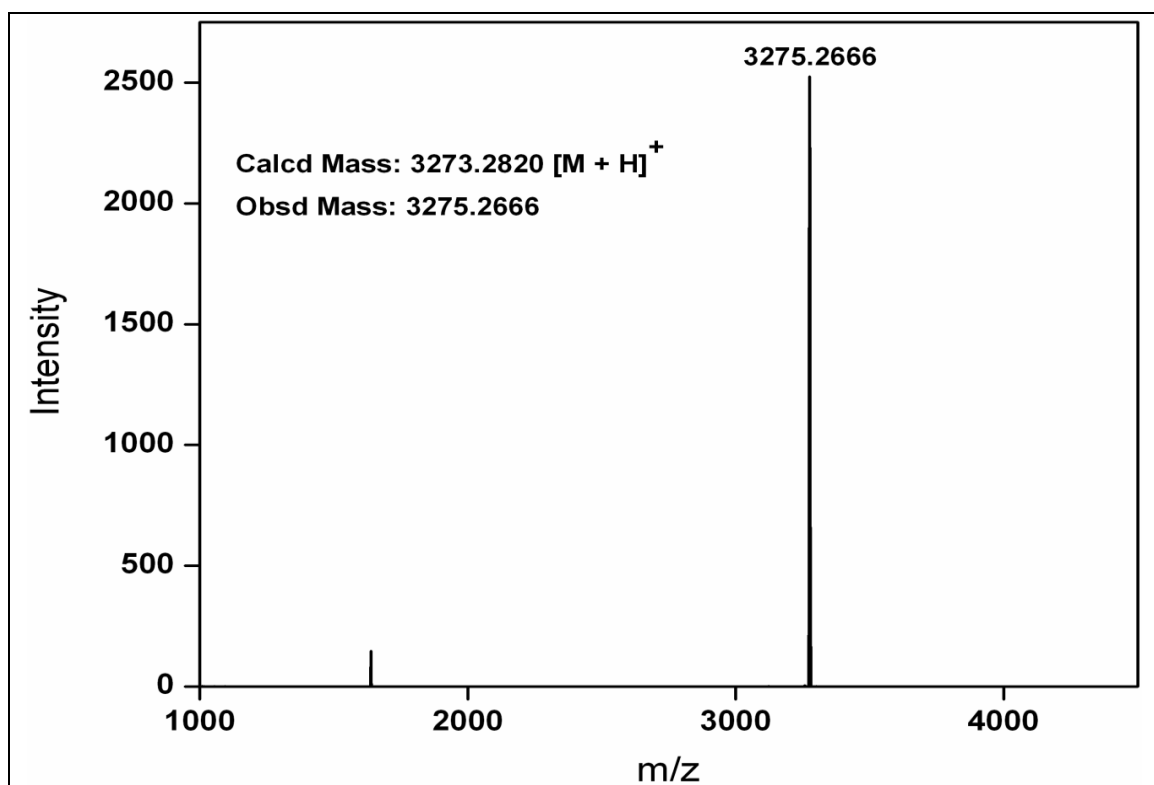
MALDI-TOF Mass of PNA 23
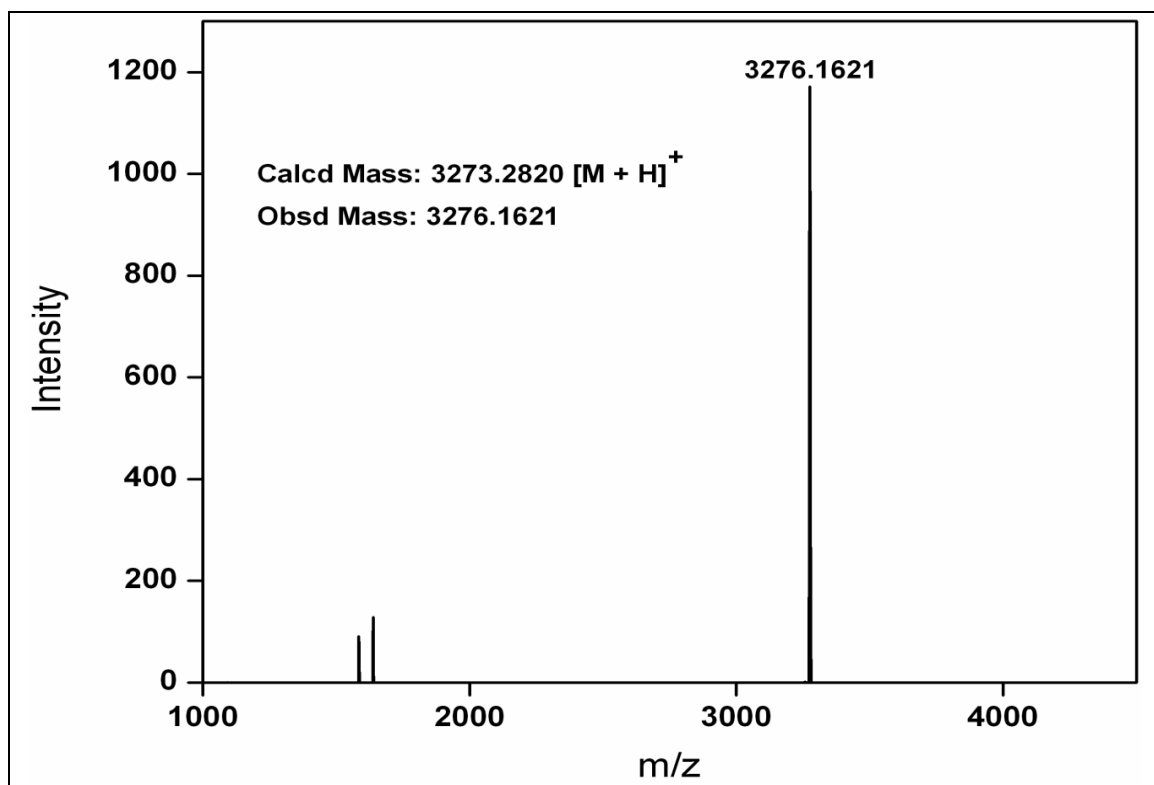
MALDI-TOF Mass of PNA 24
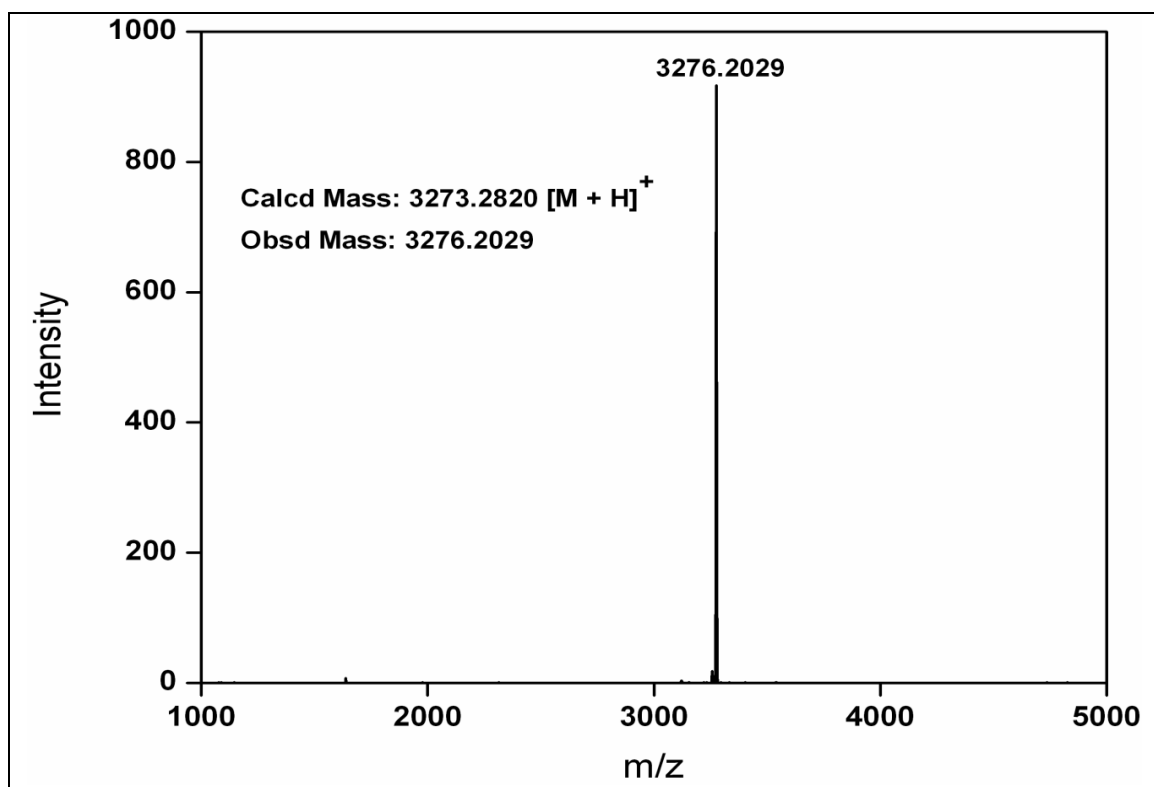
MALDI-TOF Mass of PNA 25

Calcd Mass: 3741.3890 [M + H]$^+$
Obsd Mass: 3744.2063

3744.2063
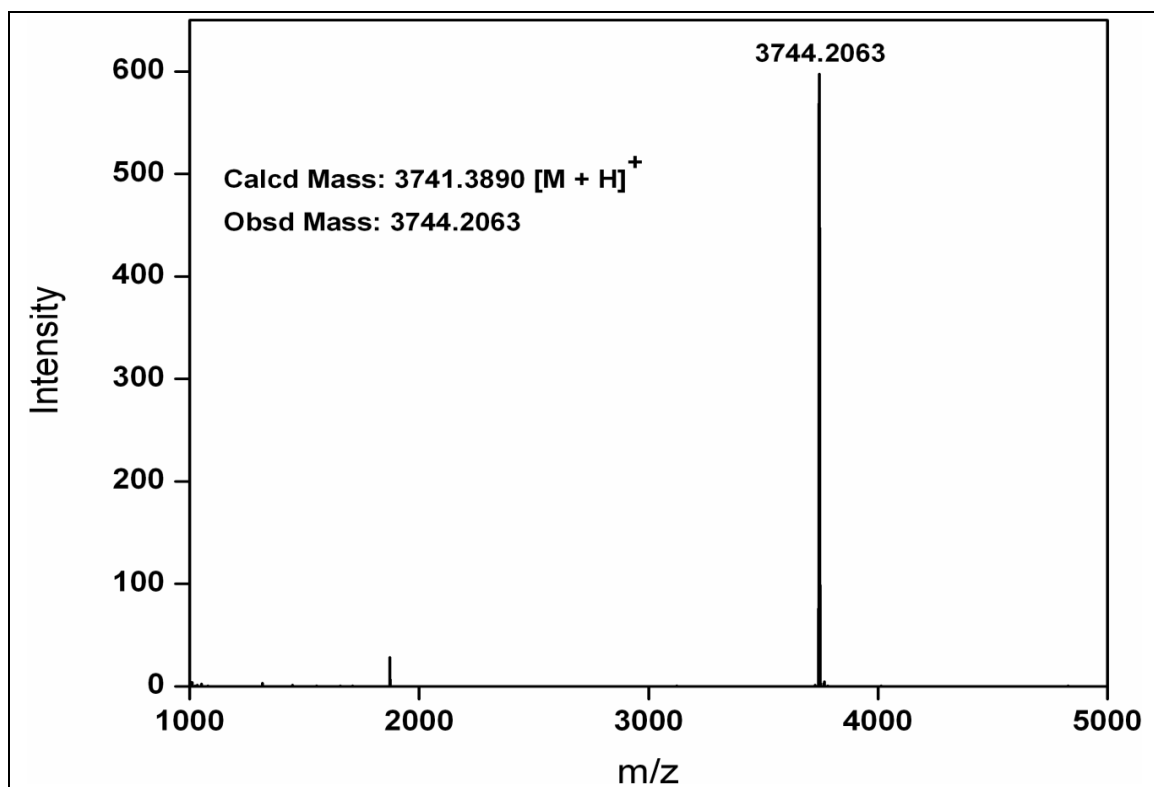
Chapter **3**

Biophysical Evaluation of Modified PNA Oligomers

The effect of chiral, cationic and neutral functionalities on the affinity of PNA towards the complementary DNA and their thermal stabilities has been studied using various biophysical techniques. The comparative binding studies of various modified PNAs have been discussed.

3.1 Introduction

Biophysics is an interdisciplinary science which uses the techniques from the physical sciences in order to understand the biological structure and their functions. These techniques are useful in studying the structure and properties of nucleic acids, proteins, peptides and their analogs.[1] The ability of antisense/antigene oligonucleotides to bind *in vitro* to the target DNA/RNA can be investigated by using various biophysical techniques like thermal UV-melting (UV-T_m), Circular Dichroism (CD), Fluorescence Spectroscopy and Gel Electrophoresis.

The preceding chapter discusses about the synthesis of rationally designed γ-C-substituted multifunctional PNA analogs. The modified PNA monomers were introduced into achiral *aeg* PNA at various positions by solid phase synthesis in order to study the effect of chirality and side chain functional groups in influencing the binding properties to target nucleic acids.

3.2 Rationale of the present work

This chapter addresses the biophysical studies of various modified PNAs and their hybrids with complementary/mismatch nucleic acids using temperature dependent UV spectroscopy, CD spectroscopy, fluorescence spectroscopy and Gel mobility shift assay. The biophysical studies has been carried out on the designed PNA analogs in order to investigate the effective conjugation of side chain functional groups on the binding affinity to target nucleic acids. These studies give information about the binding selectivity, binding specificity, structural organization and the duplex stability which are important parameters to evaluate the antisense/antigene agents.

3.3 Biophysical techniques used to study the hybridization properties

This section describes the applications of various biophysical techniques used to study the properties of single stranded PNAs and PNA:DNA duplexes.

3.3.1 UV-melting

The two strands of complementary nucleic acids are held together by hydrogen bonding and the stacking interactions between the adjacent nucleobases. The stability

of duplexes depends on the conditions like temperature, pH and the ionic strength. If the conditions that disrupts the hydrogen bonding and stacking interactions are applied, the strands are no longer held together and the double helix is denatured. The duplex is said to melt if temperature is used as denaturing agent. The thermal stability of various nucleic acid complexes including PNA-DNA/RNA hybrids has been studied by monitoring the UV absorption at 260 nm as a function of temperature.[2] The hetero-aromatic bases interact via their π electron clouds when stacked together in the duplex in water. Because the UV absorbance of the bases is a consequence of π electron transitions, the magnitude of these transitions is affected when bases are stacked. The denaturation or dissociation of two strands results in loss of stacking and disruption of hydrogen bonding which causes the increase in absorbance at 260 nm called hyperchromicity. The process is co-operative and the plot of absorbance at 260 nm vs temperature is sigmoidal (Figure **3.1**). The midpoint of the sigmoidal transition is termed as the melting temperature, T_m.
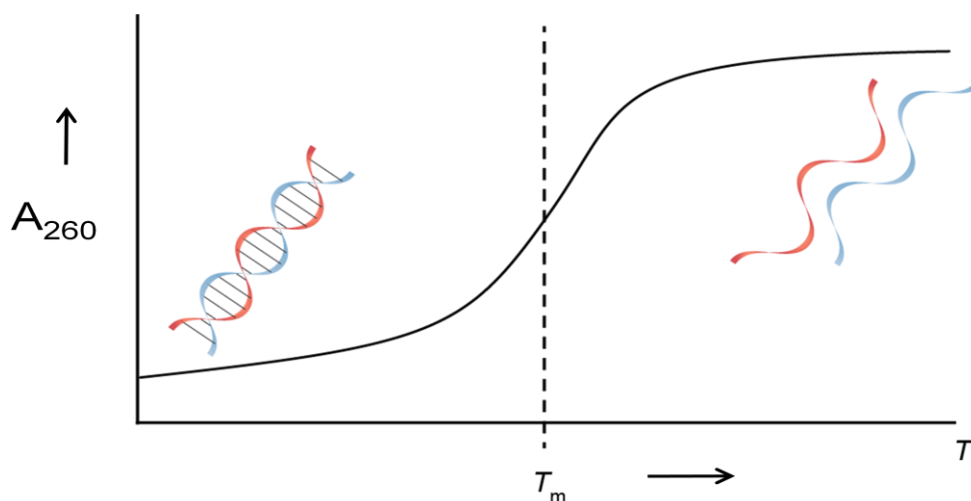
FIGURE 3.1 Sigmoidal graph of UV-T_m

The sigmoidal nature of transition suggests that the nucleic acids exist in two states, either as duplexes or as single strands and at varying temperatures, the relative proportions of two states change. A nonsigmoidal (e.g. sloping linear) transition with low hyperchromicity is a consequence of non-complementation. In many cases, the transitions are broad and the exact T_m values are obtained from the peak in first derivative plots (Figure **3.2**). This technique has provided valuable information regarding the strength of base-dependent complementary interactions in nucleic acid hybrids involving DNA, RNA and PNA.

The binding stoichiometry of nucleic acids can be determined from the UV-titration and the CD-mixing using Job's plot. The combination of UV absorption and the CD spectra provides the correct determination of complex formation and the strand stoichiometry. In literature, it has been shown that polypyrimidine PNA binds to complementary DNA in a 2:1 ratio forming PNA_2:DNA triplex[3] whereas the mixed purine-pyrimidine PNA binds in 1:1 ratio forming PNA:DNA duplex.[4] In this chapter all PNAs contain mixed purine-pyrimidine bases so the PNA and DNA are mixed in 1:1 ratio.
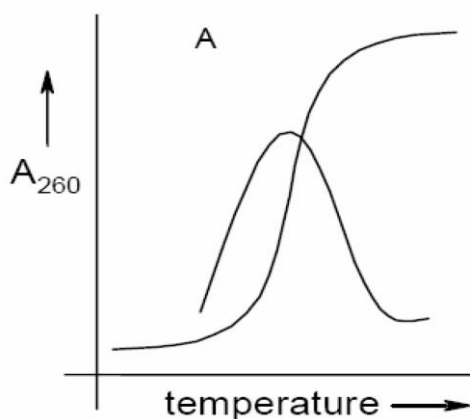
FIGURE 3.2 Derivative plot of PNA:DNA duplex

3.3.2 Circular dichroism

Circular dichroism (CD) is a spectroscopic technique used to measure the difference between the absorption of left handed circularly polarized light (L-CPL) and the right handed circularly polarized light (R-CPL) and results from molecules that contain one or more chromophores adjacent to chiral centre. The chiral molecules exist as isomers of non-superimposable mirror images called enantiomers. The physical and chemical properties of the pair of enantiomers are identical except the way they interact with polarized light and the way they interact with other chiral molecules.

CD spectra are particularly valuable in determining the following aspects:

➢ Whether the individual strands themselves are self-coiled

➢ Whether the individual strands change their conformation upon binding to complementary oligonucleotide strand

➢ How the conformations of hybrid PNA:DNA complexes are related with hybrid DNA:DNA complexes

➢ Nature of helices: α, π (proteins) A, B (DNA)

CD spectra provide a reliable estimation of the overall conformational state of biopolymers and structural changes induced by modification as compared with the reference compound. In case of nucleic acids, the sugar units of the backbone possess chirality and the bases attached to sugars are the chromophores. CD spectroscopy monitors the structural changes of nucleic acids in solution and helps to diagnose whether new or unusual structures are formed by particular polynucleotide sequences.

3.3.3 Fluorescence spectroscopy

Fluorescence is generated when a substance absorbs light energy in uv-visible region and then emits radiation at a longer wavelength. Fluorescence corresponds to the emission of radiation as a molecule returns to its ground state from an excited electronic state.[5] The duration of time interval between absorption and emission is usually in the order of 10^{-9} to 10^{-8} seconds. The fundamentals of this phenomenon can be explained by Jablonski diagram (Figure **3.3 A**). The difference (in wavelength or frequency units) between the positions of band maxima of absorption and emission spectra is called Stokes shift (Figure **3.3 B**).
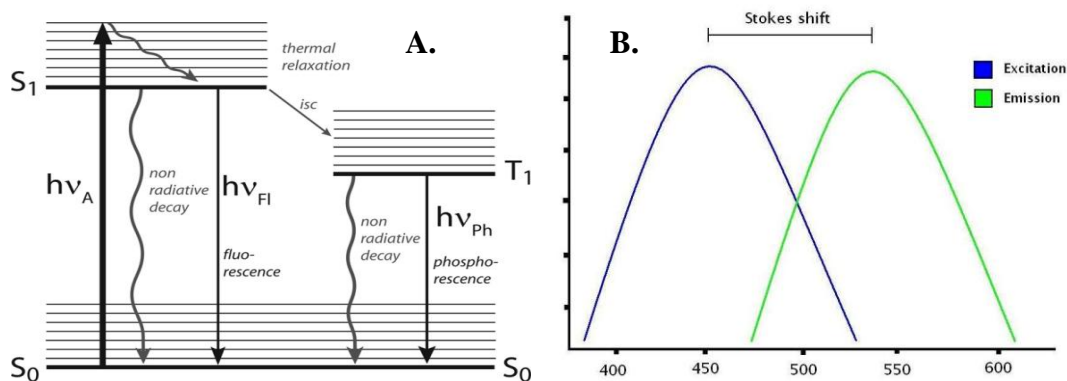
FIGURE 3.3 (A) A simplified Jablonski diagram of a fluorescence event. The fluorescent molecule begins in its ground energy state, S_0, and is converted to an excited singlet state, S_1', by absorbing energy ($h\nu_A$) in a specific wavelength. The molecule has a transition to the relaxed singlet state, S_1 or T_1, by releasing some of the absorbed energy. Finally, the molecule returns to its ground energy state by releasing the remaining energy ($h\nu_{Fl}$ - fluorescence) or ($h\nu_{Ph}$ - phosphorescence). The duration of a single fluorescence event is a few nanoseconds. (B) Generalized representation of stokes shift

Fluorescence spectroscopy has applications in examining the interaction between macromolecules and a ligand. Many DNA binding ligands are not fluorescent or have very little fluorescence in aqueous solutions. However, upon binding to a DNA duplex, the ligand-nucleic acid complex shows changes in fluorescence because the ligand is in a hydrophobic environment where solvent can no longer quench the intrinsic ligand fluorescence. Therefore, fluorescence emission at particular wavelength can be used as a direct probe of ligand-nucleic acid interactions.

Ethidium bromide (EtBr) interacts with duplex DNA by intercalation between the base pairs. It is a weakly fluorescent molecule, but when intercalated to DNA duplex, exhibits a strong fluorescence with emission maxima at 595 nm upon excitation at 475 nm. A competent ligand binding to DNA can displace the intercalated ethidium bromide thus leading to a decrease in the fluorescence intensity of the complex.[6] An attractive feature of the ethidium binding is that its fluorescence can increase upto 20 fold upon intercalation with DNA making it a very useful probe for monitoring the interactions of DNA duplex with oligonucleotides (to form triplex) or small molecules.[7] In ethidium bromide displacement assay, when DNA-ethidium bromide complexes are challenged by other DNA binding agents (like PNA) where ethidium bromide is displaced from the complex leading to a fall in fluorescence intensity (Figure **3.4**). In such experiments, a measure of the decrease in fluorescence intensity of ethidium bromide is useful to compute the relative binding strengths of the added PNA to DNA.
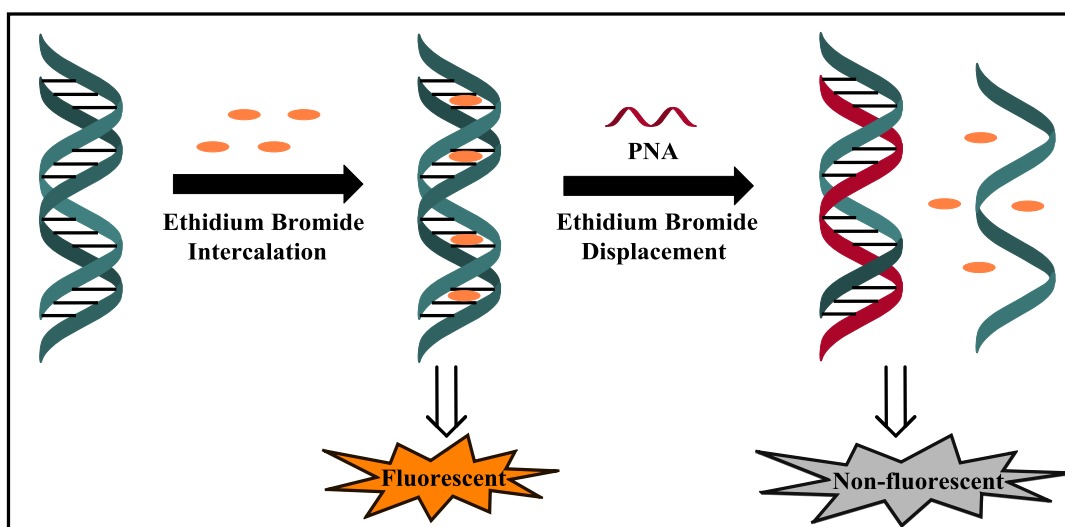
FIGURE 3.4 Ethidium bromide displacement assay

3.3.4 Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assay (EMSA) or Gel retardation assay, is a common technique used to study protein-nucleic acid interactions. EMSA can be used to identify and estimate the binding of oligonucleotides to complementary DNA/RNA. The advantage of electrophoretic assay is its ability to resolve complexes of different stoichiometries and conformations.[8] In free solution, the movement of DNA in an electric field is dependent on its charge and independent of shape and molecular mass. However when DNA is exposed to an electric field in a gel matrix, the movement is dependent on size, shape as well as charge. Gels commonly used for nucleic acid electrophoresis are made of agarose or polyacrylamide. Both types of materials consist of 3D networks of cross-linked polymer strands, with pores whose size varies according to the concentration of polymer used. The mobility of DNA in such gels is dependent mainly on size and shape, since the charge per unit length of DNA is effectively constant. Shorter molecules move faster and migrate farther than longer ones because shorter molecules migrate more easily through the pores of the gel. This phenomenon is called sieving.[9]

Agarose gel electrophoresis is generally useful for resolving nucleic acid fragments in the size range of 100 nucleotides to around 10-15 kb. Below this range, fragments are difficult to separate and hard to visualize because of rapid diffusion within the gel matrix. Polyacrylamide gel electrophoresis (PAGE) can separate fragments as small as 10 bp and upto 1 kb with a resolution as little as 1 bp. It can also accommodate much larger quantities of DNA (up to 10 μg) and allow recovery of very pure DNA. Figure **3.5** shows the principle of PAGE technique.
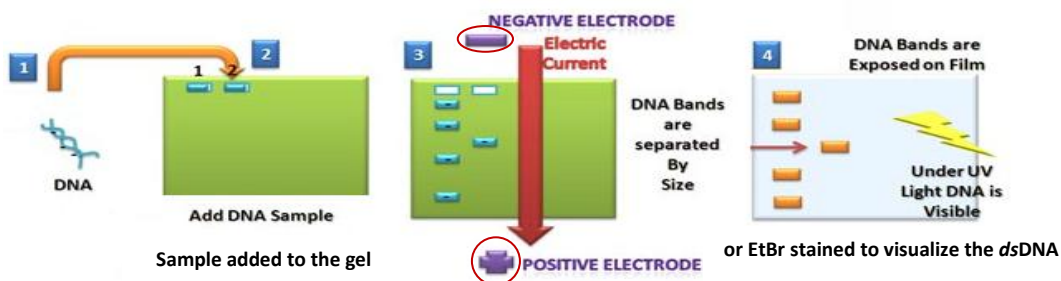
FIGURE 3.5 Protocol for polyacrylamide gel electrophoresis

Gel Electrophoresis

Electrophoresis can be carried out under two conditions: Denaturing or nondenaturing (native) polyacrylamide gel.

Denaturing gels are run under the conditions that disrupt the higher order structures in proteins causing them to unfold into a linear chain, random coil or disrupting the nucleic acid structures into single stranded fragments. This involves polymerization in the presence of denaturing agents like urea, sodium dodecyl sulphate (SDS), glyoxal etc. which disrupts the hydrogen bondings in nucleic acid bases. The mobility of macromolecule under these conditions depends on its length/size and its mass to charge ratio.

When gels are run under the non-denaturing conditions (absence of urea, SDS etc), there is no disruption of nucleic acid structures or complexation. As a general rule, *ds*DNAs migrate through nondenaturing polyacrylamide gels at rates that are inversely proportional to the $\log_{10}$ of their size but are affected by their base composition and sequence, so that duplex DNAs of exactly the same size can differ in mobility by upto 10 %.

3.4 Objectives of the present work

This section presents biophysical studies of the modified PNAs in terms of following aspects:

- ➢ Thermal stability, binding selectivity, specificity and discrimination of modified PNA towards complementary DNA using temperature-dependent UV spectroscopy
- ➢ Characterization of the PNA:DNA duplex structures by CD spectroscopy
- ➢ Study of fluorescence properties of single strand fluorescent PNA and derived duplexes with complementary DNA
- ➢ Determination of binding of PNA to DNA by electrophoretic mobility shift and ethidium bromide competition displacement assays
- ➢ Analysis and comparison of the results of various biophysical studies

The structures of the modified PNA units incorporated in the *aeg* PNA sequence have been shown in the Figure **3.6**.
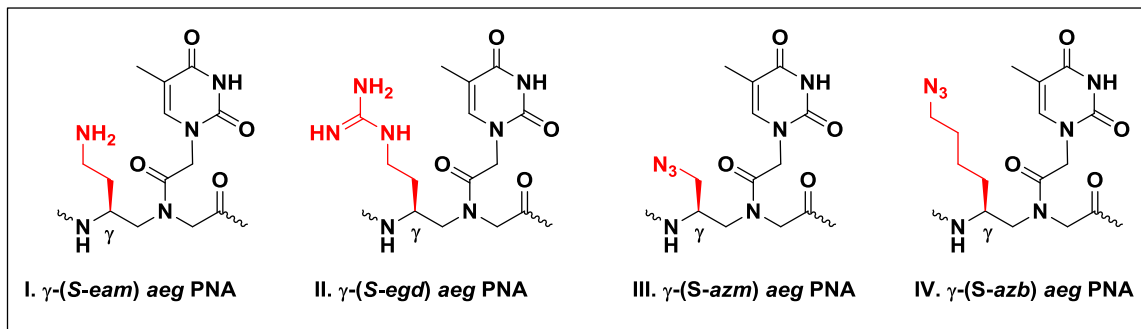
FIGURE 3.6 Modified PNA units

3.5 Results

The interaction of modified PNAs with DNA can be investigated by changes in UV, CD and fluorescence spectroscopy upon complexation. The site-specific effects of introduction of modified PNA units at various positions in a 10 mer *aeg* PNA sequence on their duplex formation are studied by these techniques. Thermal stabilities of PNA:DNA hybrids can be examined by temperature dependent UV absorption at 260 nm. The knowledge of binding selectivity, specificity and discrimination of chiral modified PNAs towards target complementary/mismatch DNA sequences is important for designing new PNAs. The effect of chiral modified PNAs on the structures of single stranded PNAs and PNA:DNA hybrids has been studied by CD spectroscopy. The binding affinity of modified PNAs with DNA was investigated by gel retardation and ethidium bromide displacement assays.

3.5.1 UV melting studies of PNA:DNA hybrids

The hybridization studies of modified PNA oligomers with their complementary **DNA 1** and mismatch **DNA 2** were carried out by temperature dependent changes in UV-absorbance. The thermal stabilities of PNA(**I-IV**):DNA**1** duplexes were determined for different PNA modifications from UV-temperature plots. The effect of mismatch on stability of PNA(**I-IV**):DNA**2** duplexes were studied with **DNA 2** having single base mismatch in the middle. The effect of different γ-PNA side chains on the thermal stabilities of PNA:DNA complexes was studied using PNAs with amino (*eam*), guanidine (*egd*) and azido (*azm*, *azb*) functional groups in the side chain (Figure **3.6**). The T_m values were obtained from mid-points of the thermal stabilities of various modified PNAs.

3.5.1a Thermal stability of *aeg* PNA and γ-ethyleneamino [γ- (*S-eam*)] *aeg* PNA:DNA duplexes

In order to investigate the effect of modified PNA monomers on their corresponding duplexes, PNAs (**PNA 1-PNA 5**) were hybridized with complementary **DNA 1** in antiparallel orientation (Figure **3.7**). The PNA:DNA duplex derived from unmodified *aeg* **PNA 1** showed a melting (T_m) of 43.4 °C (Table **3.1**, entry **1**), while the T_m of *eam* modified PNA:DNA duplex showed variation as a function of the number and site of modification.

Single modifications: The thermal stability of different ethyleneamino (*eam*) modified PNAs is generally enhanced irrespective of the position of modification in *aeg* PNA sequence (**PNA 1**). Among these, the C-terminal modification (**PNA 4**) showed relatively higher stabilization ($\Delta T_m = + 5.8$ °C, Table **3.1**, entry **4**) compared to the N-terminal and middle modifications. The N-terminal (**PNA 2**) and middle (**PNA 3**) modified PNAs showed T_m stabilization of + 3.9 °C and + 3.7 °C (Table **3.1**, entry **2** & **3**) respectively, over the unmodified control PNA**1**:DNA**1** duplex.

FIGURE 3.7 Temperature dependent UV absorbance curves for complementary γ-(*S-eam*) *aeg* PNA:DNA 1 duplexes (DNA 1 = 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

Bi-modification: Incorporation of two modified units of γ-(*S-eam*) *aeg* PNA (**PNA 5**) into the *aeg* PNA sequence (one towards N-terminus and the other in middle) caused considerable increase in T_m values ($\Delta T_m = + 6.3$ °C, Table **3.1**, entry **5**) for its duplex formed with complementary **DNA 1**. These results indicate that the each γ-(*S-eam*) modification of *aeg* PNA in the 10 mer *aeg* PNA sequence stabilizes duplex by ~ 3 °C.

γ-(*S-eam*) **aeg** PNA

TABLE 3.1 UV-T_m values of complementary PNA:DNA duplexes with γ-(*S-eam*) *aeg* PNA units

Entry	PNA Code	PNA Sequence	DNA 1: 3' AATGGAGTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*aeg* PNA **1**	H-TTACCTCAGT-Lys NH$_2$	43.4	_
2	*eam*-t$_2$ PNA **2**	H-T**t**ACCTCAGT-LysNH$_2$	47.3	+ 3.9
3	*eam*-t$_6$ PNA **3**	H-TTACC**t**AGT-LysNH$_2$	47.1	+ 3.7
4	*eam*-t$_{10}$ PNA **4**	H-TTACCTCAG**t**-LysNH$_2$	49.2	+ 5.8
5	*eam*-t$_{2,6}$ PNA **5**	H-T**t**ACC**t**CAGT-LysNH$_2$	49.7	+ 6.3

ΔT_m indicates the difference in T_m with control *aeg* PNA, ΔT_m values are accurate to ±0.5°C

3.5.1b Thermal stability of γ-ethyleneguanidino [γ-(*S-egd*)] *aeg* PNA:DNA duplexes

Thermal stability results of γ-*S*-ethyleneguanidino (*egd*) modified PNAs (**PNA 6-10**) with complementary **DNA 1** are shown in Figure **3.8** and Table **3.1**.

Single modifications: Incorporation of single unit of γ-(*S-egd*) into *aeg* PNA **1** at the C-terminus (**PNA 8**) showed an increased stabilization of (ΔT_m = + 8.6 °C) compared to N-terminal and middle modifications (Table **3.2**, entry **3**). Among the middle (**PNA 7**) and N-terminal (**PNA 6**) modifications, the middle modifiied (**PNA 7**) showed a higher stability (ΔT_m = + 5.4 °C) compared to the N-terminal modification (**PNA 6**) that stabilized the corresponding duplex by + 2.5 °C (Table **3.2**, entry **1** & **2**). The observed results are similar to that seen with γ-(*S-eam*) *aeg* PNA that showed higher stabilization for C-terminal modification.

Bi-modification: The stabilization seen by incorporating two modifications of γ-(*S-egd*) units in *aeg* PNA **1** was higher than that with single modification with the exception of C-terminal single modification which had stabilization equal to bi modification. Two modified units of γ-(*S-egd*) *aeg* PNA (**PNA 9**) in the 10 mer *aeg* PNA 1 stabilized the duplex by + 8.5 °C (Table **3.2**, entry **4**). The results suggested a

ΔT_m stabilization of 4-4.5 °C/modification, with exception of **PNA 8** which had maximum stabilization of 8.6 °C/modification.
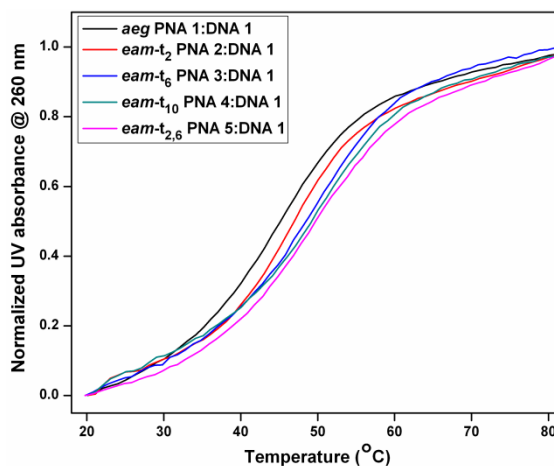

FIGURE 3.8 Temperature dependent UV absorbance curves for complementary γ-(*S-egd*) *aeg* PNA:DNA 1 duplexes (DNA 1 = 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

γ-(*S-egd*) *aeg* PNA

TABLE 3.2 UV-T_m values of complementary PNA:DNA duplexes with γ-(*S-egd*) *aeg* PNA units

Entry	PNA Code	PNA Sequence	DNA 1: 3' AATGGAGTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*egd*-t$^{\gamma}_2$ PNA **6**	H-Tt$^{\gamma}$ACCTCAGT-LysNH$_2$	45.9	+ 2.5
2	*egd*-t$^{\gamma}_6$ PNA **7**	H-TTACCt$^{\gamma}$CAGT-LysNH$_2$	48.8	+ 5.4
3	*egd*-t$^{\gamma}_{10}$ PNA **8**	H-TTACCTCAGt$^{\gamma}$-LysNH$_2$	52.0	+ 8.6
4	*egd*-t$^{\gamma}_{2,6}$ PNA **9**	H-Tt$^{\gamma}$ACCt$^{\gamma}$CAGT-LysNH$_2$	51.9	+ 8.5
5	*egd*-t$^{\gamma}_{2,6,10}$ PNA **10**	H-Tt$^{\gamma}$ACCt$^{\gamma}$CAGt$^{\gamma}$-LysNH$_2$	57.4	+ 14.0

ΔT_m indicates the difference in T_m with control *aeg* PNA, ΔT_m values are accurate to ±0.5°C

Tri-modification: Introduction of three modified units of γ-(*S-egd*) *aeg* PNA as in **PNA 10** remarkably increased the thermal stability of PNA:DNA duplex. The tri modified **PNA 10** stabilized the duplex by + 14 °C (Figure **3.8**, Table **3.2**, entry **5**)

which was the highest stabilization among all modifications. The degree of stabilization seen with this corresponds to about ~ 5 °C per modification.

3.5.1c Thermal stability of γ-azidomethylene [γ-(*S-azm*)] *aeg* PNA:DNA duplexes

Azido functional group was incorporated at γ-position through side chains having 1 or 4-carbons to enable linking a fluorophore by click reaction and for comparison of chiral and neutral modifications with unmodified *aeg* PNA in terms of thermal stability (Figure **3.9**).

Single modifications: Incorporation of azidomethylene (*azm*) side chain at the γ-position in *aeg* PNA backbone showed no effect on its DNA binding ability. N-terminal (**PNA 11**) and C-terminal (**PNA 13**) modifications stabilized the DNA duplex by + 0.5 °C and + 2.4 °C respectively (Table **3.3**, entry **1** & **3**). It is significant to note that the middle modification (**PNA 12**) showed a higher stabilization with ΔT_m = + 6.3 °C (Table **3.3**, entry **2**), compared to terminal modifications.

Bi-modification: Introduction of two modified units of γ-(*S-azm*) *aeg* PNA into 10 mer *aeg* PNA sequence did not destabilize the duplex. The doubly (N-terminus and the middle) modified **PNA 14** formed hybrid with **DNA 1** (T_m = 43.6 °C) which was as much stable as that with *aeg* PNA (Table **3.3**, entry **4**).

FIGURE 3.9 Temperature dependent UV absorbance curves for complementary γ-(*S-azm*) *aeg* PNA:DNA 1 duplexes (DNA 1 = 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)
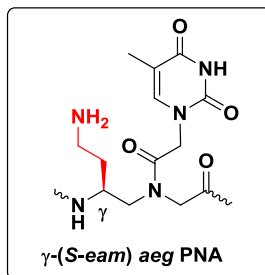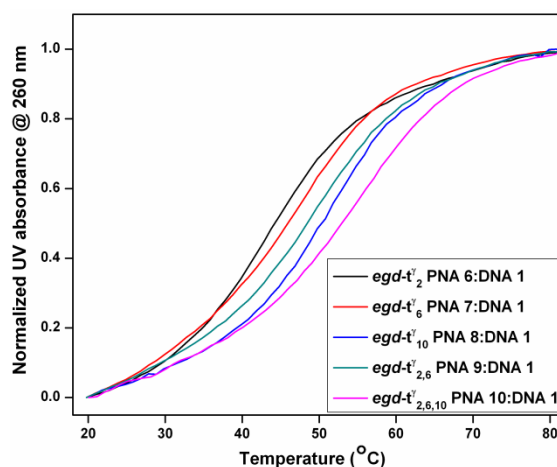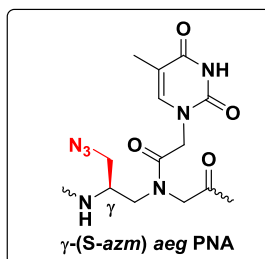
γ-(S-*azm*) *aeg* PNA

TABLE 3.3 UV-T_m values of complementary PNA:DNA duplexes with γ-(*S-azm*) *aeg* PNA units

Entry	PNA Code	PNA Sequence	DNA 1: 3' AATGGAGTCA 5'	
			UV-T_m (°C)	ΔT_m
1	azm-T$^γ_2$ PNA **11**	H-T**T**^{γ}ACCTCAGT-LysNH$_2$	43.9	+ 0.5
2	azm-T$^γ_6$ PNA **12**	H-TTACC**T**^{γ}CAGT-LysNH$_2$	49.7	+ 6.3
3	azm-T$^γ_{10}$ PNA **13**	H-TTACCTCAG**T**^{γ}-LysNH$_2$	45.8	+ 2.4
4	azm-T$^γ_{2,6}$ PNA **14**	H-T**T**^{γ}ACC**T**^{γ}CAGT-LysNH$_2$	43.6	+ 0.2

ΔT_m indicates the difference in T_m with control *aeg* PNA, ΔT_m values are accurate to ±0.5°C

3.5.1d Thermal stability of γ-azidobutylene [γ -(*S-azb*)] *aeg* PNA:DNA duplexes

Azidobutylene substitution having 4-carbon side chain was incorporated at γ-position to investigate the effect of a longer side chain on derived PNA:DNA hybrids (Figure **3.10**).

Single modifications: Incorporation of azidobutylene (*azb*) side chain at γ-position enhances the duplex stability better than the shorter azidomethylene (*azm*) side chain. The PNA with C-terminal modification (**PNA 17**) showed relatively higher duplex stabilization among all singly modified γ-(*S-azb*) *aeg* PNA oligomers (ΔT_m = + 7.5 °C) whereas the N-terminal (**PNA 15**) and middle (**PNA 16**) modifications stabilized the duplex by only 1.5 °C and 5.8 °C respectively (Table **3.4**, entry **1**, **2**, **3**).

Bi-modification: Incorporation of two modified units of γ-(*S-azb*) into *aeg* PNA **1** stabilized the derived PNA:DNA duplex. The doubly (N-terminus and middle) modified **PNA 18** showed a stabilization of + 7.2 °C upon hybridization with complementary **DNA 1** (Table **3.4,** entry **4**).
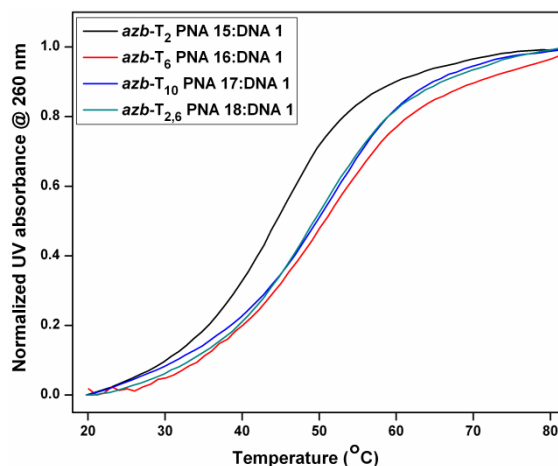
FIGURE 3.10 Temperature dependent UV absorbance curves for complementary γ-(*S-azb*) *aeg* PNA:DNA 1 duplexes (DNA 1 = 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)
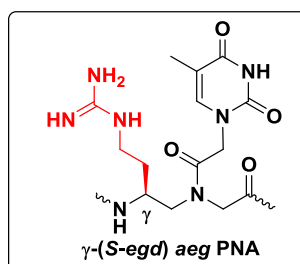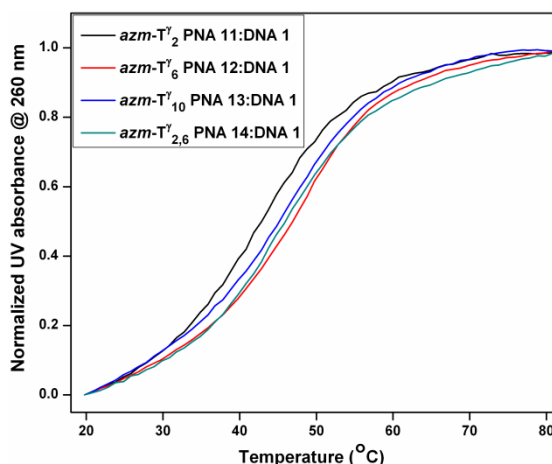

TABLE 3.4 UV-T_m values of complementary PNA:DNA duplexes with γ-(*S-azb*) *aeg* PNA units

Entry	PNA Code	PNA Sequence	DNA 1: 3' AATGGAGTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*azb*-T$_2$ PNA **15**	H-T**T**ACCTCAGT-LysNH$_2$	44.9	+ 1.5
2	*azb*-T$_6$ PNA **16**	H-TTACC**T**CAGT-LysNH$_2$	49.2	+ 5.8
3	*azb*-T$_{10}$ PNA **17**	H-TTACCTCAG**T**-LysNH$_2$	50.9	+ 7.5
4	*azb*-T$_{2,6}$ PNA **18**	H-T**T**ACC**T**CAGT-LysNH$_2$	50.6	+ 7.2

ΔT_m indicates the difference in T_m with control *aeg* PNA, ΔT_m values are accurate to ±0.5°C

3.5.2 UV-T_m mismatch studies of PNA:DNA hybrids

Sequence specific binding is an important criterion for antisense therapeutics to avoid off-target binding of antisense oligonucleotides. Peptide nucleic acids have attracted attention because of its high sequence specificity which can be determined by base mismatch studies. Therefore, sequence specificity of modified PNA:DNA duplexes was determined by studying the hybridization properties of PNA with complementary **DNA 2** (5' ACTG**C**GGTAA 3') carrying single base mismatch in the middle of the sequence (Figure **3.11**).

T T A C C T C A G T   PNA

A A T G G C G T C A   Mismatch DNA 2

FIGURE 3.11 Mismatch PNA:DNA complex

The PNA:DNA complexes derived from unmodified *aeg* **PNA 1** and various other modified PNAs were subjected to UV-melting experiments to examine their ability to discriminate mismatch base in the target DNA.

3.5.2a UV-T_m studies with *aeg* PNA, γ-(*S-eam*) *aeg* PNA and γ-(*S-egd*) *aeg* PNA

The results of UV-T_m studies for duplexes of *aeg* PNA, γ-(*S-eam*) *aeg* PNA and γ-(*S-egd*) *aeg* PNA modifications with mismatch **DNA 2** are shown in Figure **3.12** and Table **3.5**. The *aeg* **PNA 1** showed a destabilization of 6 °C when hybridized with complementary **DNA 2** having single base mismatch in the middle of the sequence, compared to full complementary **PNA1:DNA1** duplex. Incorporation of γ-(*S-eam*) *aeg* PNA unit at N-terminus (**PNA 2**) and in centre (**PNA 3**) destabilized the corresponding **PNA2:DNA2** and **PNA3:DNA2** duplexes by 10.6 °C and 12.7 °C respectively (Table **3.5**, entry **2**, **3**). The duplexes from C-terminus (**PNA 4**) and double modification (**PNA 5**) showed destabilization with $\Delta T_m = 13$ °C and 12.5 °C respectively (Table **3.5**, entry **4**, **5**). Among all single modifications, C-terminus modification (**PNA 8**) showed a higher destabilization ($\Delta T_m = 15$ °C) than centre (**PNA 7**) and N-terminus (**PNA 6**) modifications which destabilize the duplex by 12.5 °C and 10.5 °C respectively (Table **3.5**, entry **6**, **7**, **8**). In PNAs with two and three modified units of γ-(*S-egd*) *aeg* PNA (**PNA 9** & **PNA 10**) higher mismatch discrimination was observed ($\Delta T_m = 13.6$ °C & 18.5 °C) (Table **3.5**, entry **9**, **10**).

FIGURE 3.12 Temperature dependent UV absorbance curves for mismatch (A) γ-(*S-eam*) *aeg* PNA:DNA 2 and (B) γ-(*S-egd*) *aeg* PNA:DNA 2 duplexes (DNA 2 = 5' ACTG**C**GGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

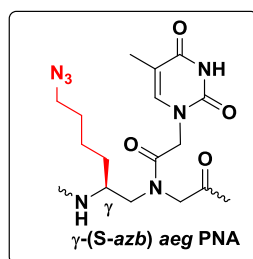**TABLE 3.5 UV-T_m values of mismatch PNA:DNA duplexes with γ-(*S-eam*) *aeg* PNA and γ-(*S-egd*) *aeg* PNA units**

Entry	PNA Code	PNA composition	DNA 2: 3' AATGG**C**GTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*aeg* PNA **1**	H-TTACCTCAGT-Lys NH$_2$	37.4	- 6.0
2	*eam*-t$_2$ PNA **2**	H-T**t**ACCTCAGT-LysNH$_2$	36.7	- 10.6
3	*eam*-t$_6$ PNA **3**	H-TTACC**t**AGT-LysNH$_2$	35.0	- 12.7
4	*eam*-t$_{10}$ PNA **4**	H-TTACCTCAG**t**-LysNH$_2$	36.2	- 13.0
5	*eam*-t$_{2,6}$ PNA **5**	H-T**t**ACC**t**CAGT-LysNH$_2$	37.2	- 12.5
6	*egd*-t$^\gamma_2$ PNA **6**	H-T**t**$^\gamma$ACCTCAGT-LysNH$_2$	35.4	- 10.5
7	*egd*-t$^\gamma_6$ PNA **7**	H-TTACC**t**$^\gamma$CAGT-LysNH$_2$	36.3	- 12.5
8	*egd*-t$^\gamma_{10}$ PNA **8**	H-TTACCTCAG**t**$^\gamma$-LysNH$_2$	37.0	- 15.0
9	*egd*-t$^\gamma_{2,6}$ PNA **9**	H-T**t**$^\gamma$ACC**t**$^\gamma$CAGT-LysNH$_2$	38.4	- 13.6
10	*egd*-t$^\gamma_{2,6,10}$ PNA **10**	H-T**t**$^\gamma$ACC**t**$^\gamma$CAG**t**$^\gamma$-LysNH$_2$	38.9	- 18.5

ΔT_m indicates the difference in T_m with complementary DNA **1** of respective PNA, ΔT_m values are accurate to ±0.5°C

3.5.2b UV-T_m studies with γ-(*S-azm*) *aeg* PNA and γ-(*S-azb*) *aeg* PNA

The results of UV-T_m studies for duplexes of γ-(*S-azm*) *aeg* PNA and γ-(*S-azb*) *aeg* PNA with mismatch **DNA 2** are shown in Figure **3.13** and tabulated in Table **3.6**. In case of γ-(*S-azm*) *aeg* PNA, when the modification was located in the middle of the sequence (**PNA 12**), the destabilization was higher (ΔT_m = 14.4 °C) (Table **3.6**, entry **2**). However, not much difference was observed when γ-(*S-azm*) unit was incorporated into *aeg* PNA at N-terminus (**PNA 11**) or C-terminus (**PNA 13**). Doubly modified **PNA 14** destabilized the corresponding duplex by 10.4 °C (Table **3.6**, entry **1**, **3**, **4**). The T_m results of incorporating γ-(*S-azb*) *aeg* PNA at various positions was similar but slightly higher mismatch discrimination was observed when the γ-(*S-azb*) modification was located at C-terminus as in **PNA 17** (ΔT_m = 14.6 °C) (Table **3.6**, entry **7**). PNAs with modifications at the N-terminus (**PNA 15**) and in the centre of the sequence (**PNA 16**) showed a decrease in T_m by 7.7 °C and 13.4 °C respectively. The doubly modified **PNA 18** (N-terminus and middle) destabilized the duplex by 12.8 °C (Table **3.6**, entry **5**, **6**, **8**).
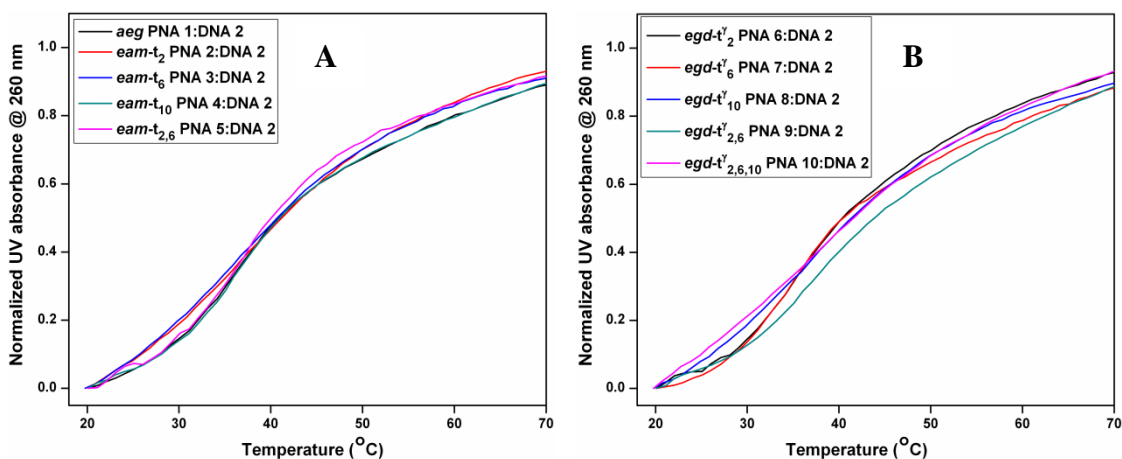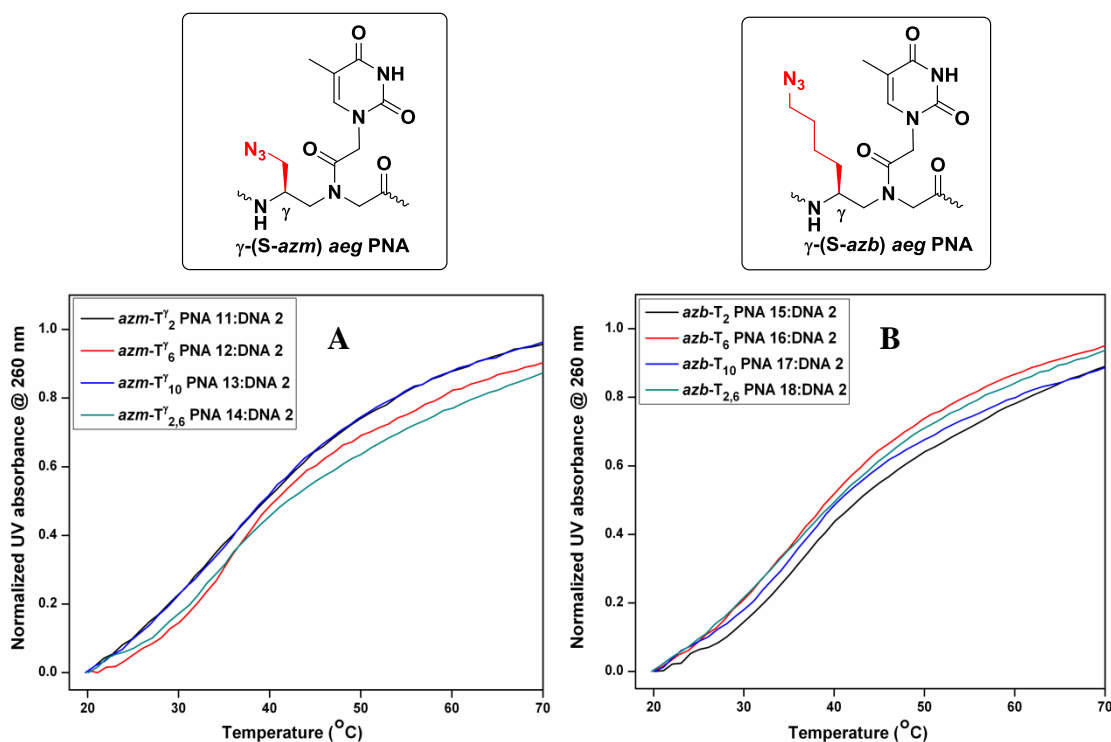

FIGURE 3.13 Temperature dependent UV absorbance curves for mismatch (A) γ-(*S-azm*) *aeg* PNA:DNA 2 and (B) γ-(*S-azb*) *aeg* PNA:DNA 2 duplexes (DNA 2 = 5' ACTG**C**GGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)
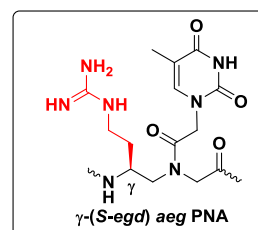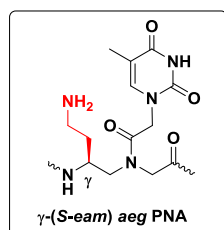
TABLE 3.6 UV-T_m values of mismatch PNA:DNA duplexes with γ-(*S-azm*) *aeg* PNA and γ-(*S-azb*) *aeg* PNA units

Entry	PNA Code	PNA composition	DNA 2: 3' AATGG**C**GTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*azm*-T$^{γ}_2$ PNA **11**	H-T**T**^{γ}ACCTCAGT-LysNH$_2$	33.8	- 10.2
2	*azm*-T$^{γ}_6$ PNA **12**	H-TTACC**T**^{γ}CAGT-LysNH$_2$	35.3	- 14.4
3	*azm*-T$^{γ}_{10}$ PNA **13**	H-TTACCTCAG**T**^{γ}-LysNH$_2$	35.8	- 10.0
4	*azm*-T$^{γ}_{2,6}$ PNA **14**	H-T**T**^{γ}ACC**T**^{γ}CAGT-LysNH$_2$	33.2	- 10.4
5	*azb*-T$_2$ PNA **15**	H-T**T**ACCTCAGT-LysNH$_2$	37.2	- 7.7
6	*azb*-T$_6$ PNA **16**	H-TTACC**T**CAGT-LysNH$_2$	35.9	- 13.4
7	*azb*-T$_{10}$ PNA **17**	H-TTACCTCAG**T**-LysNH$_2$	36.3	- 14.6
8	*azb*-T$_{2,6}$ PNA **18**	H-T**T**ACC**T**CAGT-LysNH$_2$	37.9	- 12.8

ΔT_m indicates the difference in T_m with complementary DNA **1** of respective PNA, ΔT_m values are accurate to ±0.5°C

3.5.3 CD spectroscopic studies of *ss* PNAs and PNA:DNA duplexes

The effect of γ-side chain substitution on the conformation of PNA:DNA duplexes was studied by CD spectroscopy. The CD spectrum of single stranded DNA shows a positive band of moderate intensity between 200 to 220 nm alongwith a lower intensity positive band in the region of 270 to 290 nm (Figure **3.14**). The single stranded PNA being achiral does not show any significant CD signatures. However, being a polyamide, it can form either left or right handed coil structures with equal facility. Even though *ss*PNA shows low profile CD signals, PNA:DNA complexes exhibit characteristic CD signatures because of the helicity induced by the chiral DNA in PNA:DNA complex (Figure **3.14**). The PNA:DNA duplex shows two positive bands, one with moderate intensity between 260 to 270 nm and another of lower intensity between 218 to 225 nm. It also showed a less intensive negative band near 240 nm.
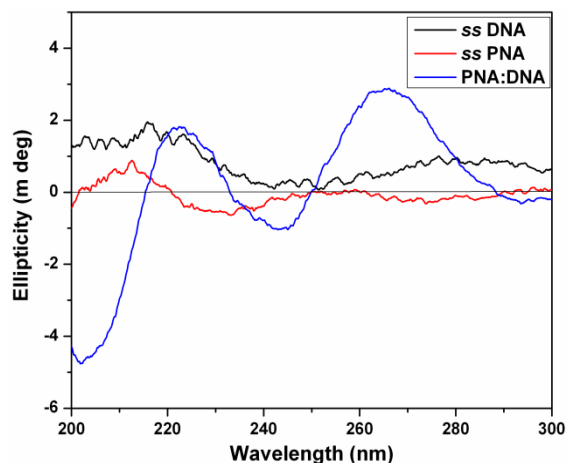
FIGURE 3.14 CD signatures of representative *ss* DNA 1, *ss* PNA and PNA:DNA duplex

The CD spectra for mixed purine pyrimidine single stranded *aeg* **PNA 1**, PNA oligomers containing chiral, cationic *γ-(S-eam) aeg* **PNA 2-5**, *γ-(S-egd) aeg* **PNA 6-10** and their complexes formed with complementary **DNA 1** are shown in Figure **3.15 A**, **B** and **3.16 A**, **B**.
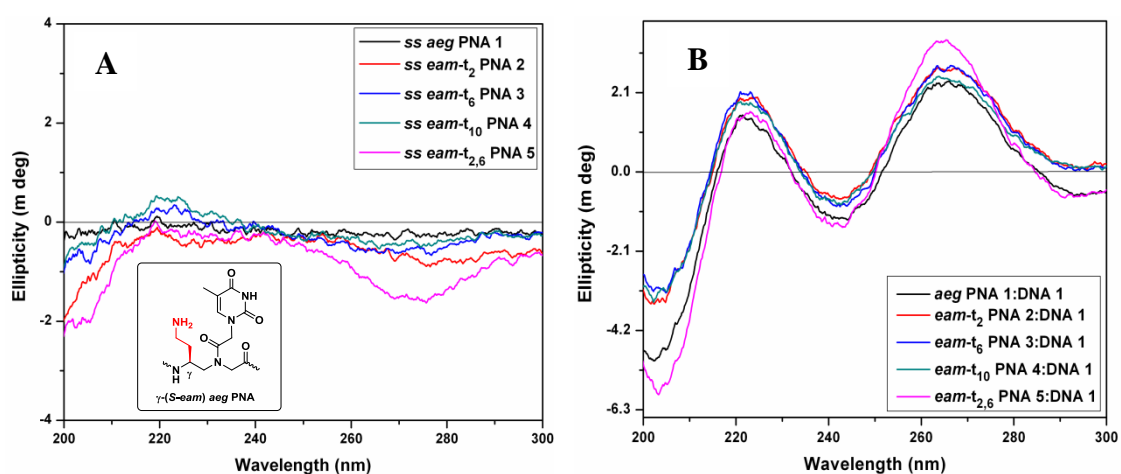
FIGURE 3.15 CD spectra of (A) *ss aeg* PNA 1, *ss* **γ*-(S-eam) aeg* PNA 2-5 (B) Duplexes of PNA 1-5 with complementary DNA 1** (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)
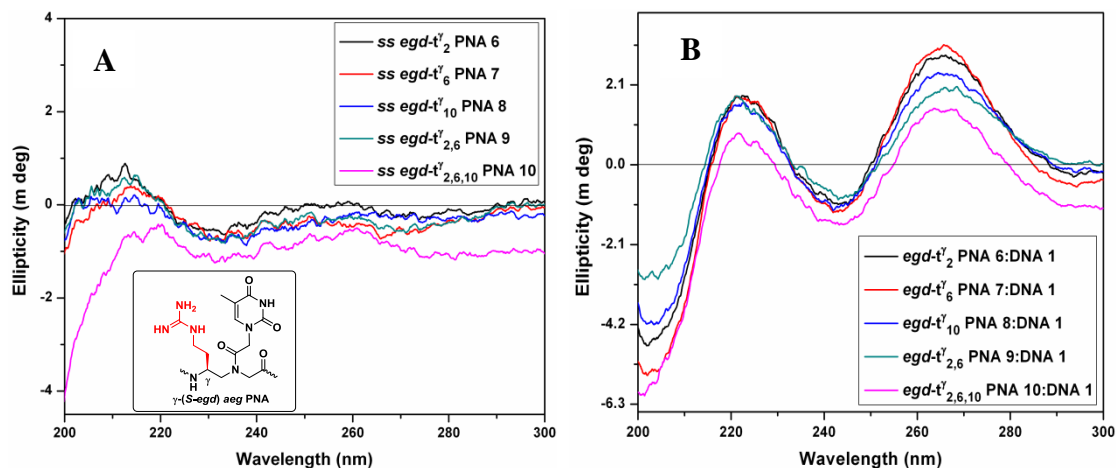
FIGURE 3.16 CD spectra of (A) *ss* **γ-(S-egd) aeg PNA 6-10 (B) Duplexes of PNA 6-10 with complementary DNA 1** (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

The CD spectra of single stranded chiral, azide modified *γ-(S-azm) aeg* **PNA 11-14**, *γ-(S-azb) aeg* **PNA 15-18** and their complexes formed with complementary **DNA 1** are shown in Figure **3.17 A, B** and **3.18 A**, **B**.
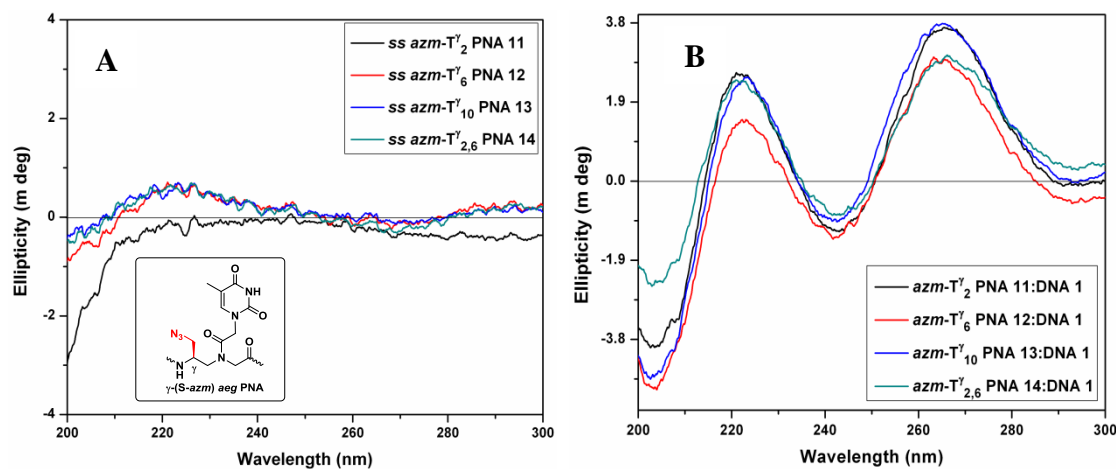
FIGURE 3.17 CD spectra of (A) *ss* **γ-(S-azm) aeg PNA 11-14 (B) Duplexes of PNA 11-14 with complementary DNA 1** (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)
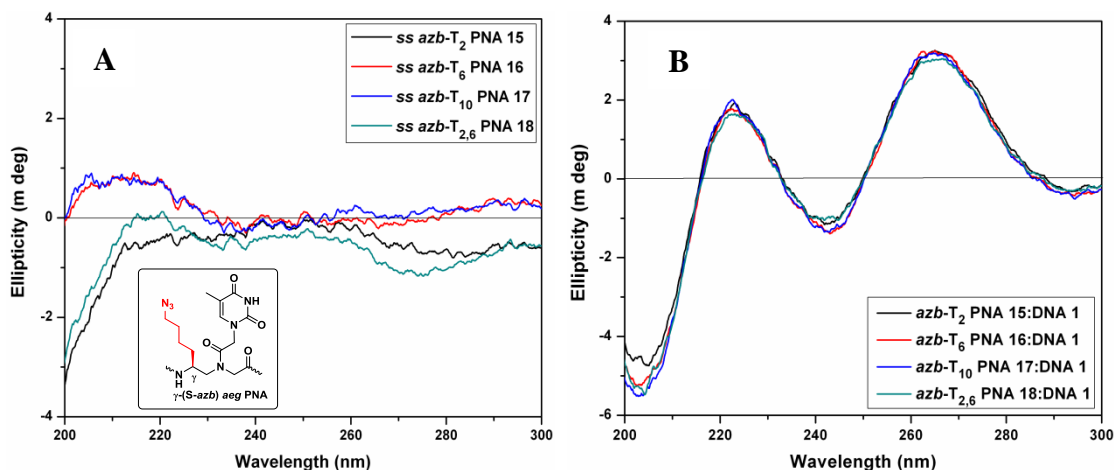
FIGURE 3.18 CD spectra of (A) *ss* **γ-(S-azb) aeg PNA 15-18 (B) Duplexes of PNA 15-18 with complementary DNA 1** (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

From the overall results it is seen that single stranded *aeg* **PNA 1** as well as backbone modified *ss*PNAs do not show significant CD spectral signature. However, the formation of corresponding PNA:DNA duplexes result in two positive bands with maxima one in the region of 262 to 270 nm (higher intensity) and another of lower intensity in between 218 to 225 nm. PNA:DNA duplexes also showed a negative maxima at 240 to 250 nm with cross-over points at approximately 235 and 250 nm.

3.5.4 Ethidium bromide displacement assay from *ds*DNA

Ethidium bromide (EtBr) is a weakly fluorescent and strong DNA intercalating agent. The intercalation into DNA is a result of unwinding of DNA (around 26°) by ethidium cation which causes base pairs to separate or rise, creating an opening of about 3.4 Å. Also, charge repulsion between the two strands causes continuous breathing in the DNA which helps ethidium bromide to get inside the two strands. Intercalation of EtBr into the duplex DNA results in the 20 fold increase in its fluorescence. However, in case of PNA:DNA duplex, the unwinding of duplex is minimal and the breathing rate is decreased as a result of reduced repulsion between the backbone units which does not allow EtBr to intercalate between the PNA containing hybrid duplexes[15] (Figure **3.19**).
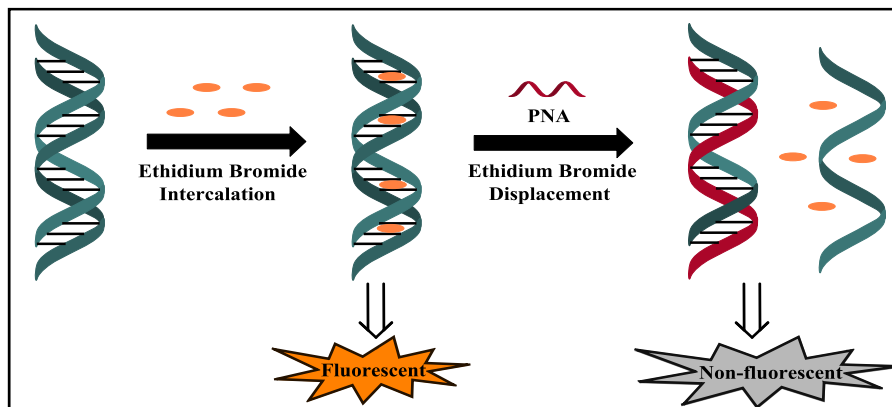
FIGURE 3.19 Ethidium bromide displacement assay

The differential binding ability of various modified PNAs towards the complementary DNA was examined by the displacement of EtBr intercalated into the DNA duplex and the displacement was followed by fluorescence decay at 610 nm. Initially, the EtBr (250 µM) was titrated against the DNA duplex (2 µM) and the increase in fluorescence intensity was found to be saturated at 5.58 µM EtBr. The resulting EtBr-*ds*DNA complex was individually titrated with aqueous solutions of various PNA oligomers (Figure **3.20** and **3.21**).

In case of control *aeg* **PNA 1**, around 84 % displacement of EtBr was observed at equimolar concentration (2 µM) of DNA duplex and *aeg* PNA 1 (Figure **3.20 A**). However, the guanidino modified cationic PNA oligomers were less efficient at displacing EtBr from the DNA duplex though, these cationic PNAs have high thermal stabilities. Further, increase in guanidino units in the PNA sequence does not alter the EtBr displacement ability. Incorporating one guanidino modified unit (**PNA 8**) in the sequence displaces the EtBr by 56 % whereas PNA oligomers having two (**PNA 9**) and three (**PNA 10**) guanidino modified units displaces the EtBr by 56 % and 57 % respectively at equimolar concentration (Figure **3.20 B**).
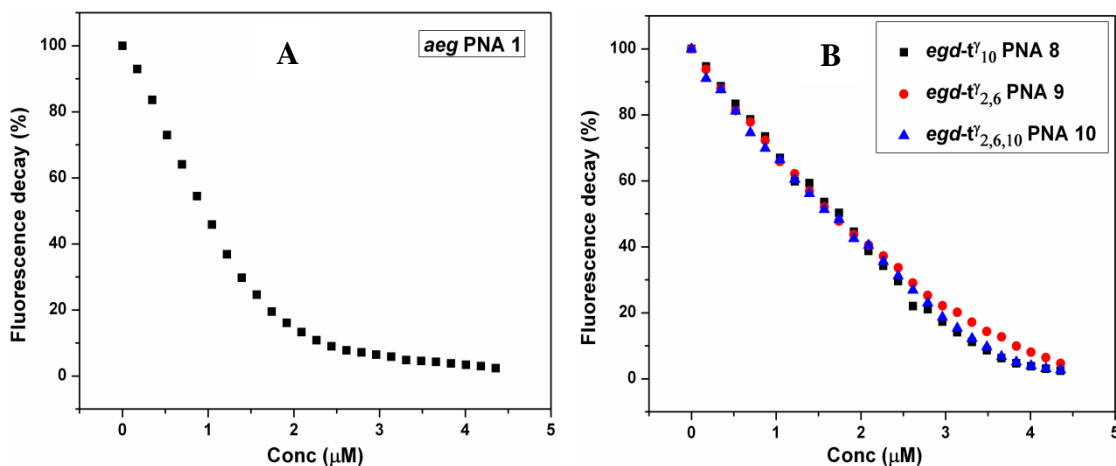
FIGURE 3.20 Ethidium bromide displacement assay of (A) *aeg* **PNA 1 (B) Guanidino modified PNA 8, PNA 9 and PNA 10 from** *ds***DNA** ($\lambda_{excitation}$ = 520 nm; $\lambda_{emission}$ = 610 nm)

The EtBr displacement ability of amino modified cationic PNAs was found to be better than guanidino modified PNAs but less efficient than control *aeg* PNA 1. Incorporating one amino modified unit (**PNA 4**) in the sequence displaces the EtBr by 76 % whereas EtBr displacement was 74 % when two amino modified units (**PNA 5**) were introduced in the sequence at equimolar concentration of DNA duplex and PNA (Figure **3.21 A**). Ethidium bromide displacement from the DNA duplex was also carried out with neutral azidomethylene (shorter) and azidobutylene (longer) side chain modified PNAs. In case of azidobutylene modified PNA (**PNA 17**), the ethidium bromide displacement was 89 % which is highest among all unmodified and modified PNAs whereas azidomethylene modified PNA (**PNA 13**) displaces EtBr by 63 % from the DNA duplex at equimolar concentration (Figure **3.21 B**).
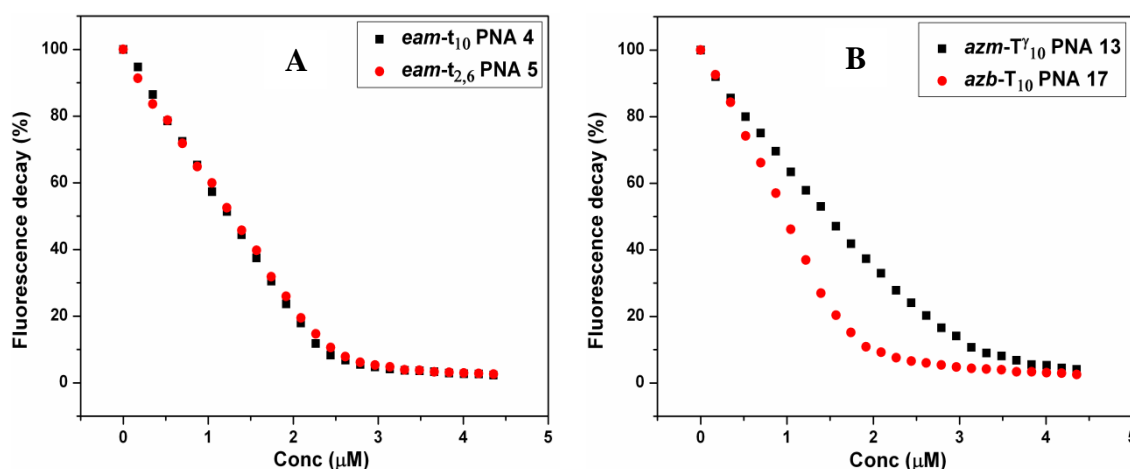
FIGURE 3.21 Ethidium bromide displacement assay of (A) Amino modified PNA 4 and PNA 5 (B) Azidobutylene modified PNA 17 and azidomethylene modified PNA 13 from

The comparative ethidium bromide displacement studies from duplex DNA showed that neutral PNA oligomers are more efficient at displacing EtBr than cationic PNA oligomers. The significantly less efficiency of cationic PNAs at displacing EtBr from the DNA duplex was perhaps due to the charge repulsion between the cationic groups of PNA and the positively charged EtBr (Figure **3.22**).
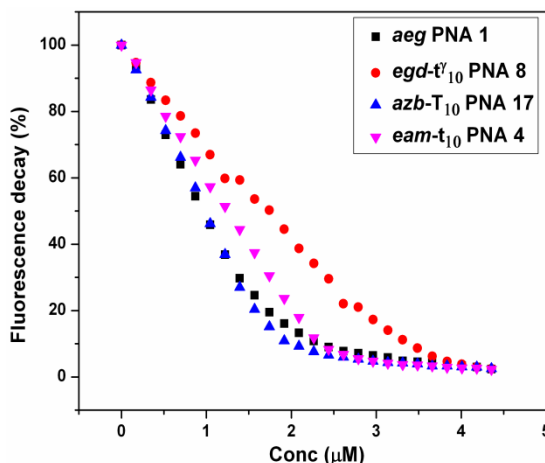
FIGURE 3.22 Comparative ethidium bromide displacement studies from *ds*DNA

3.5.5 Electrophoretic mobility shift assay and competition binding experiment

Nucleic acids have differential mobility on polyacrylamide gel on the basis of their mass to charge ratio. Therefore, single stranded species and their duplexes run differently under electrophoretic conditions. Electrophoretic gel shift assay was used to establish the binding of different PNAs to the complementary DNA. Control *aeg* **PNA 1** and **PNA 10** with three guanidino modified units were individually treated with complementary **DNA 1** and the complex formations were monitored by gel electrophoresis at ~ 18 °C. The spots were visualized through UV shadowing by illuminating the gel placed on a fluorescent silica gel plate, F254 using UV light (Figure **3.23**). Single stranded oligonucleotides (DNA 1, DNA 3) DNA**1**:DNA**3** duplex and PNA**10**:DNA**1** duplex showed differential gel mobility under electrophoretic conditions depending on their charge and mass. Being negatively charged in nature, single stranded DNAs moved faster on gel whereas neutral *aeg* **PNA 1** and cationic **PNA 10** did not move much out of the well (Figure **3.23** lane **2**, **3**, **4** and **8**).

DNA**1**:DNA**3** duplex showed retarded mobility more than single stranded DNAs because of its increased mass.
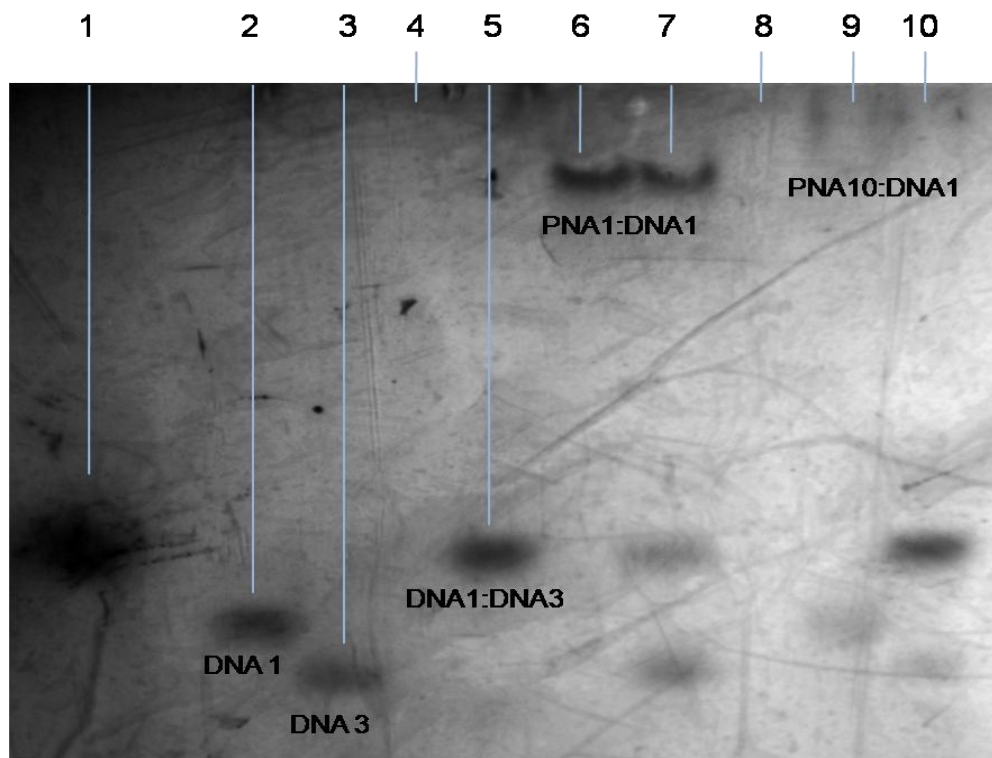
FIGURE 3.23 Electrophoretic mobility shift assay & competition binding experiment. lane 1, Bromophenol blue; lane 2, DNA 1; lane 3, DNA 3; lane 4, PNA 1; lane 5, DNA 1:DNA 3; lane 6, PNA 1: DNA 1; lane 7, DNA 1: DNA 3 + PNA 1; lane 8, PNA 10; lane 9, PNA 10:DNA1; lane 10, DNA 1: DNA 3 + PNA 14 (DNA 1 = 5' ACTGAGGTAA 3'; DNA 3 = 5' TTACCTCAGT 3')

However, the mobility of *aeg* PNA**1**:DNA**1** duplex was highly retarded in comparison with DNA duplex due to overall reduced negative charge on the complex of PNA with DNA (Figure **3.23**, lane **5**, **6**). The retardation was even higher for cationic PNA**10**:DNA**1** duplex where guanidine group neutralizes the charge on DNA, decreasing the overall negative charge present on the complex (Figure **3.23**, lane **9**).

The competition binding experiment was also carried out by adding PNAs to the DNA duplex followed by annealing and gel electrophoresis (see the equation).

DNA 1:DNA 3 + PNA = DNA 1: PNA + DNA 3
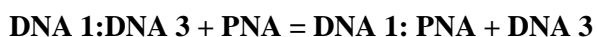
The PNA:DNA complex formation was confirmed from the release of single stranded **DNA 3** and the appearance of PNA:DNA complex similar to positive control (lane **6** & **9**). From the observed results, the *aeg* **PNA 1** as well as the cationic **PNA 10** compete with **DNA 3** to bind to complementary **DNA 1** (Figure **3.23**, lane **7**, **10**).

3.6 Biophysical evaluation of fluorescently labeled PNA:DNA duplexes

This section describes the stability of fluorescently labeled PNA:DNA duplexes by temperature dependent UV spectroscopy, CD studies of single stranded PNAs and their duplexes with complementary DNA. The UV-Visible and fluorescence spectroscopic characterization has also been discussed.

3.6.1 Thermal stability of *azm*-t(f) and *azb*-T(f) PNA:DNA duplexes

The click reaction of PNAs having azide side chains with alkyne derivative of 5(6)-carboxyfluorescein yielded fluorescent PNA oligomers which may be used as probes of hybridization. Hence the thermal stability of these fluorescent PNA oligomers upon hybridization with DNA was measured (Figure **3.24**).
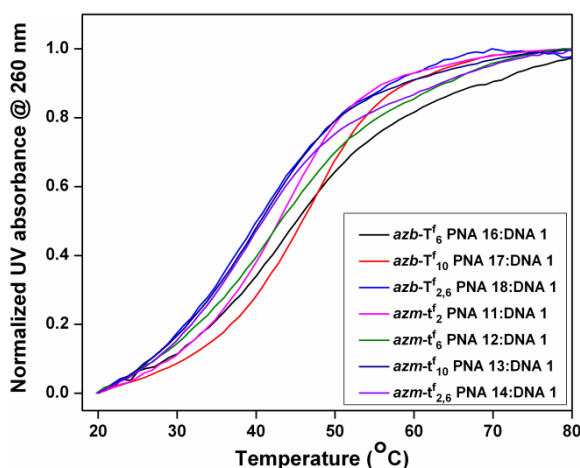
FIGURE 3.24 Temperature dependent UV absorbance curves for complementary *azm*-t(f) PNA and *azb*-T(f) PNA:DNA 1 duplexes (DNA 1 = 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

Single modifications: The tolerance of bulky fluorophore incorporated via click reaction is better when the distance is more (*azb*-T(f) PNA) from the backbone of PNA in terms of its duplex stability. In case of azidobutylene (longer) side chain, incorporation of fluorophore in the middle of PNA sequence [*azb*-**T(f)$_6$ PNA 16**] stabilized the duplex by 3 °C whereas the stabilization was slightly higher (ΔT_m = + 4.4 °C) when fluorophore is at C-terminal end [*azb*-**T(f)$_{10}$ PNA 17**] (Table **3.7**, entry **1**, **2**). In case of azidomethylene with shorter side chain, introduction of fluorophore at N-terminus [*azm*-**t(f)$_2$ PNA 11**] does not affect the DNA binding (ΔT_m = 0.4 °C). However, introduction of fluorophore in middle [*azm*-**t(f)$_6$ PNA 12**] and C-terminal end

[*azm*-t(f)$_{10}$ PNA **13**] led to destabilization of the duplex by 1.4 °C (Table **3.7**, entry **4**, **5**, **6**).

Bi-modifications: Incorporation of two fluorophoric PNA units in a 10 mer PNA sequence destabilized PNA:DNA duplex irrespective of the length of side chain. In both [*azb*-T(f)$_{2,6}$ PNA **18** and *azm*-t(f)$_{2,6}$ PNA **14**], thermal stability (T_m = 40 °C) was lower than the control *aeg* PNA **1** (ΔT_m = - 3.4 °C) (Table **3.5**, entry **3**, **7**).
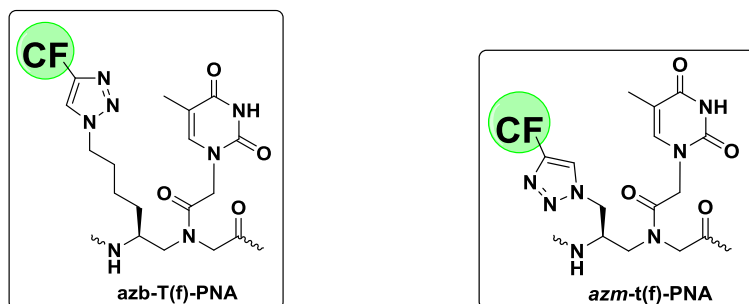
TABLE 3.7 UV-T_m values of complementary PNA:DNA duplexes with fluorescent *azm*-t(f) and *azb*-T(f) PNA units

Entry	PNA Code	PNA Sequence	DNA 1: 3' AATGGAGTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*azb*-T$^f_6$ PNA **16**	H-TTACC**T^f**CAGT-LysNH$_2$	46.4	+ 3.0
2	*azb*-T$^f_{10}$ PNA **17**	H-TTACCTAG**T^f**-LysNH$_2$	47.8	+ 4.4
3	*azb*-T$^f_{2,6}$ PNA **18**	H-T**T^f**ACC**T^f**CAGT-LysNH$_2$	40.0	- 3.4
4	*azm*-t$^f_2$ PNA **11**	H-T**t^f**ACCTCAGT-LysNH$_2$	43.8	+ 0.4
5	*azm*-t$^f_6$ PNA **12**	H-TTACC**t^f**CAGT-LysNH$_2$	42.0	- 1.4
6	*azm*-t$^f_{10}$ PNA **13**	H-TTACCTCAG**t^f**-LysNH$_2$	42.0	- 1.4
7	*azm*-t$^f_{2,6}$ PNA **14**	H-T**t^f**ACC**t^f**CAGT-LysNH$_2$	40.0	- 3.4

ΔT_m indicates the difference in T_m with control *aeg* PNA, ΔT_m values are accurate to ±0.5°C

3.6.2 UV-T_m mismatch studies of fluorescently labeled *azm*-t(f) and *azb*-T(f) PNA

The T_m results for fluorescent PNA oligomers are shown in Figure **3.25** and Table **3.8**. Incorporation of fluorophore in the middle of sequence [*azb*-T(f)$_6$ PNA **16**] through click reaction destabilized the corresponding **PNA:DNA 2** duplex by 13.6 °C whereas destabilization was higher when fluorophore was present at C-terminus [*azb*-T(f)$_{10}$ PNA **17**] with ΔT_m = 15.1 °C (Table **3.8**, entry **1**, **2**). Introducing two fluorophore units as in *azb*-T(f)$_{2,6}$ PNA **18** does not show any binding with mismatch **DNA 2** (Table **3.8**, entry **3**). In case of shorter chain carrying fluorophore at N-terminus [*azm*-t(f)$_2$ PNA **11**], destabilizaion of the duplex was by 12 °C (Table **3.8**, entry **4**). The

193

destabilization of duplex was similar when fluorophore was introduced at C-terminus [*azm*-t(f)$_{10}$ **PNA 13**], middle of the sequence [*azm*-t(f)$_6$ **PNA 12**] or doubly modified [*azm*-t(f)$_{2,6}$ **PNA 14**] with $\Delta T_m \sim 9$ °C (Table **3.8**, entry **5, 6, 7**).
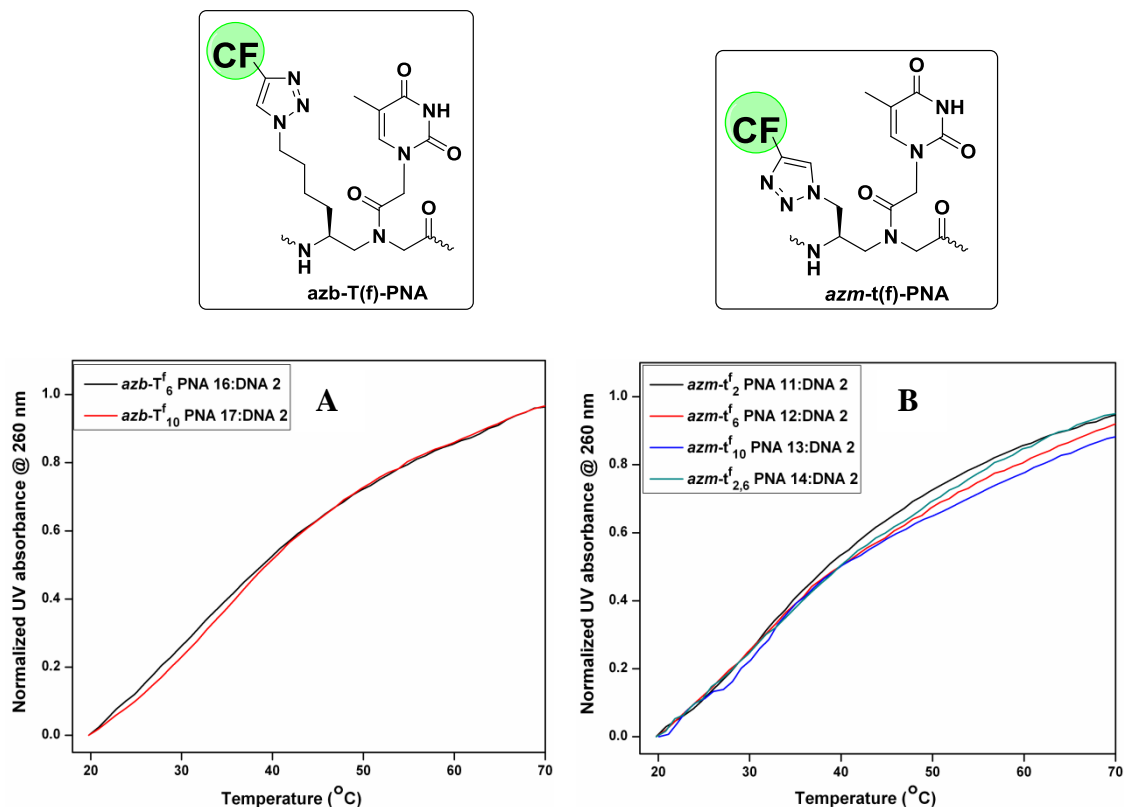
FIGURE 3.25 Temperature dependent UV absorbance curves for mismatch (A) *azb*-**T(f) PNA:DNA 2 and (B)** *azm*-**t(f) PNA:DNA 2 duplexes** (DNA 2 = 5' ACTG**C**GGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

TABLE 3.8 UV-T_m values of mismatch PNA:DNA duplexes with fluorescent *azm*-t(f) and *azb*-T(f) PNA units

Entry	PNA Code	PNA composition	DNA 2: 3' AATGG**C**GTCA 5'	
			UV-T_m (°C)	ΔT_m
1	*azb*-T$^f_6$ PNA **16**	H-TTACC**T^f**CAGT-LysNH$_2$	32.8	- 13.6
2	*azb*-T$^f_{10}$ PNA **17**	H-TTACCTAG**T^f**-LysNH$_2$	32.9	- 15.1
3	*azb*-T$^f_{2,6}$ PNA **18**	H-T**T^f**ACC**T^f**CAGT-LysNH$_2$	nb	-
4	*azm*-t$^f_2$ PNA **11**	H-T**t^f**ACCTCAGT-LysNH$_2$	31.7	- 12.0
5	*azm*-t$^f_6$ PNA **12**	H-TTACC**t^f**CAGT-LysNH$_2$	32.8	- 9.2
6	*azm*-t$^f_{10}$ PNA **13**	H-TTACCTCAG**t^f**-LysNH$_2$	33.2	- 8.8
7	*azm*-t$^f_{2,6}$ PNA **14**	H-T**t^f**ACC**t^f**CAGT-LysNH$_2$	31.0	- 9.0

ΔT_m indicates the difference in T_m with complementary DNA **1** of respective PNA, nb = not binding,

3.6.3 UV-visible and fluorescence studies of *azm*-t(f) and *azb*-T(f) PNA

As discussed in chapter 2 (Figure **2.15**), alkyne derivative of 5(6)-carboxyfluorescein was reacted with azide modified PNAs (**PNA 11-14** and **PNA 16-18**) under click reaction conditions to get fluorescent PNAs. The absorption spectra were recorded for single stranded ***azm*-t(f) PNA 11-14** and ***azb*-T(f) PNA 16-18**. The absorption spectra showed absorbance maxima at 260 nm due to nucleobases whereas absorbance maxima at 458 and 484 nm arise from the excitation of 5(6)-carboxyfluorescein residue in the PNA oligomer (Figure **3.26 A** & **B**).
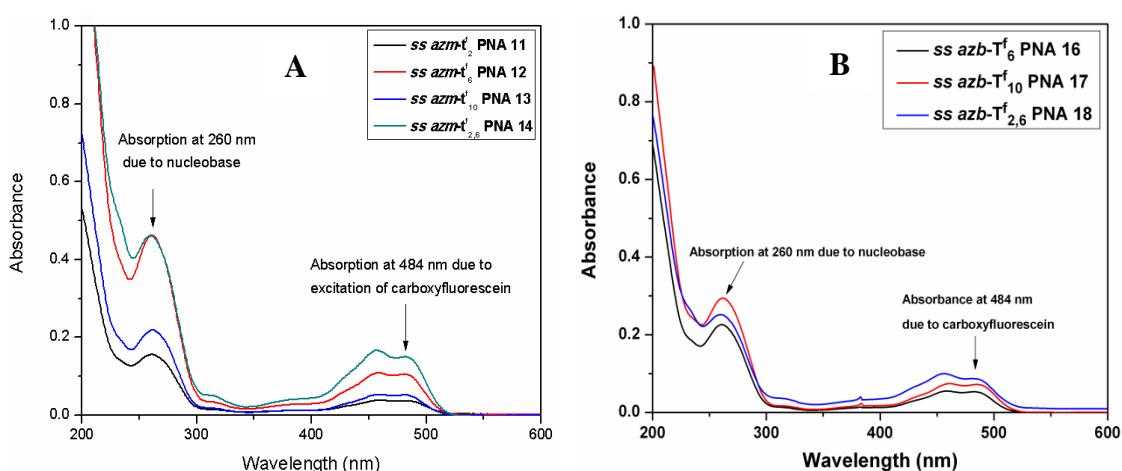
FIGURE 3.26 UV-Visible absorption spectra of (A) *azm*-t(f) **PNA 11-14 and (B)** *azb*-T(f) **PNA 16-18**

The fluorescence spectra were recorded for single stranded ***azm*-t(f) PNA 11-14**, ***azb*-T(f) PNA 16-18** and their complexes with complementary **DNA 1.** These fluorescent PNAs upon excitation at 458 nm and 484 nm showed the emission spectra with maximum at 520 nm. The fluorescence intensity of single stranded fluorescent PNAs as well as their complexes with complementary **DNA 1** at 484 nm was approximately double than that of 458 nm excitation wavelength. We observed slightly decreased fluorescent intensities in case of duplexes with complementary **DNA 1** than that of single stranded fluorescent PNAs when excited at either of the excitation wavelengths at 458 nm or 484 nm. However, the change observed in the fluorescent intensities of single stranded fluorescent PNAs and their duplexes was not significant. The fluorescence spectra for single stranded ***azm*-t(f) PNA 11-14**, ***azb*-T(f) PNA 15-18** and their complexes with complementary **DNA 1** are shown in Figure **3.27 A**, **B** and **3.28 A**, **B** respectively.

FIGURE 3.27 Fluorescence spectra of (A) *ss azm*-t(f) PNA 11-14 and (B) Duplexes of *azm*-t(f) PNA **11-14 with complementary DNA 1** (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM; $\lambda_{excitation}$ = 484 nm, emission λ_{max} = 520 nm)
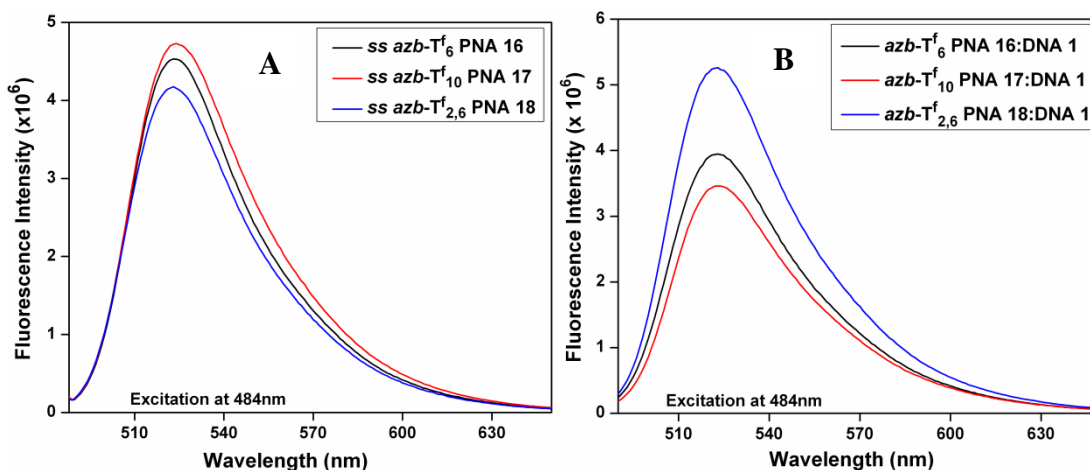
FIGURE 3.28 Fluorescence spectra of (A) *ss azb*-T(f) PNA 16-18 and (B) Duplexes of *azb*-T(f) PNA **16-18 with complementary DNA 1** (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM; $\lambda_{excitation}$ = 484 nm, $\lambda_{emission}$ = 520 nm)

3.6.4 CD spectroscopic studies of fluorescently labeled *ss* PNAs and PNA:DNA duplexes

The effect of bulky fluorophore introduced through a click reaction between the azido component of azide modified PNAs and the alkyne derivative of 5(6)-carboxyfluorescein on the conformation of derived PNA:DNA duplexes was studied by CD spectroscopy. To see whether the bulky fluorophore on PNA affects the overall secondary structure, CD spectra were recorded for single stranded *azm*-t(f) PNA 11-14, *azb*-T(f) PNA 16-18 and their complexes formed with complementary DNA 1 (Figure **3.29 A, B** and **3.30 A**, **B**).
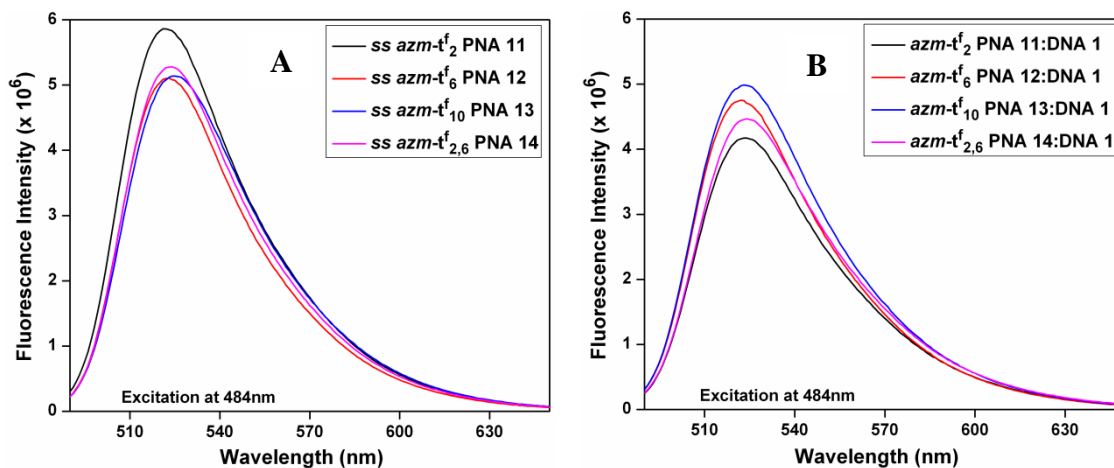
196
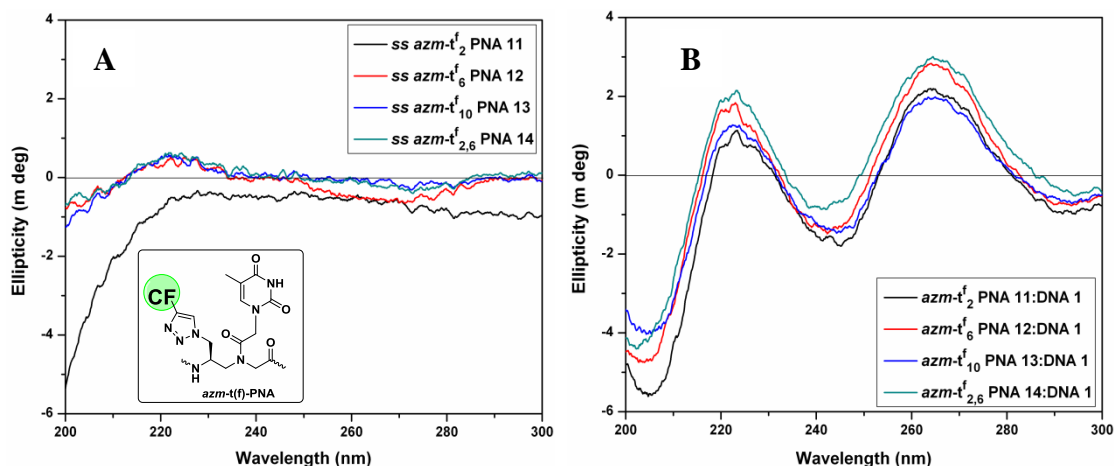
FIGURE 3.29 CD spectra of (A) *ss azm*-t(f) PNA 11-14 (B) Duplexes of *azm*-t(f) PNA 11-14 with complementary DNA 1 (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)
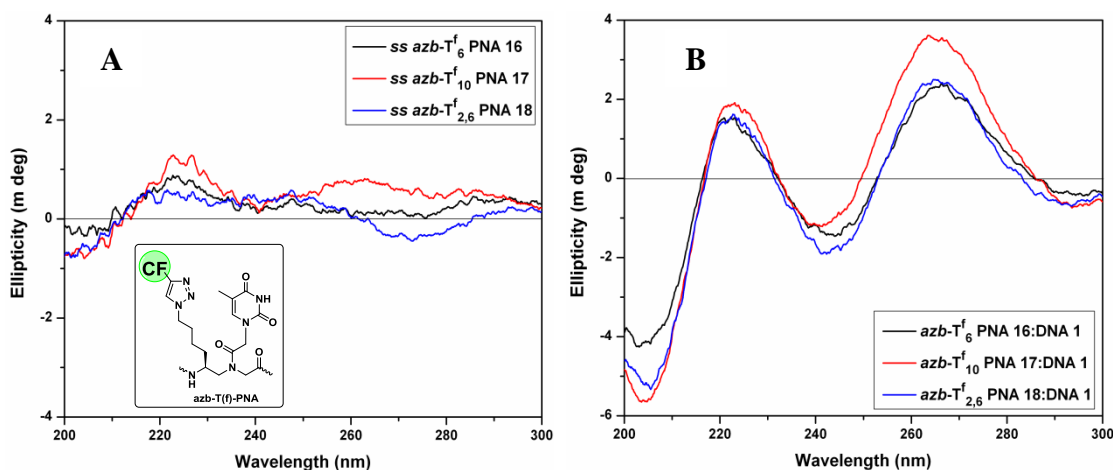
FIGURE 3.30 CD spectra of (A) *ss azb*-T(f) PNA 16-18 (B) Duplexes of *azb*-T(f) PNA 16-18 with complementary DNA 1 (DNA 1: 5' ACTGAGGTAA 3'; Buffer: 10 mM sodium phosphate, pH 7.2, NaCl 10 mM, EDTA 0.1 mM)

From the CD spectroscopic studies, it is seen that single stranded fluorescently labeled PNAs do not show significant CD spectral signature. However, the formation of corresponding PNA:DNA duplexes result in two positive bands with maxima one in the region of 262 to 270 nm (higher intensity) and another of lower intensity in between 218 to 225 nm. PNA:DNA duplexes also showed a negative maxima at 240 to 250 nm with cross-over points at approximately 235 and 250 nm.

3.7 Discussion

This section describes the comparative studies of thermal stabilities of various modified PNA oligomers. The effect of carbon chain length carrying cationinc (amine/guanidine) functional groups on the PNA:DNA duplex stability and the abilities of various modified PNAs to displace ethidium bromide from DNA duplex have also been discussed.

3.7.1 Comparative study of UV-T_m values of modified PNAs

The temperature dependent UV-absorbance data shows variation in the stabilities of modified PNA:DNA duplexes depending on the type of modification incorporated (amine, guanidine and azide), position of modification (Figure **3.31**) and the number of modified units incorporated into *aeg* PNA sequence. Almost in all the cases, the incorporation of modified unit enhanced the binding stability and selectivity of the complex irrespective of the position of modification. Also an increase in the number of modified units in the *aeg* PNA backbone amplified the induced thermal stability.

N   M   C

N-terminus   C-terminus

H-TTACCTCAGT-LysNH$_2$ PNA
3' AATGGAGTCA 5'   DNA 1
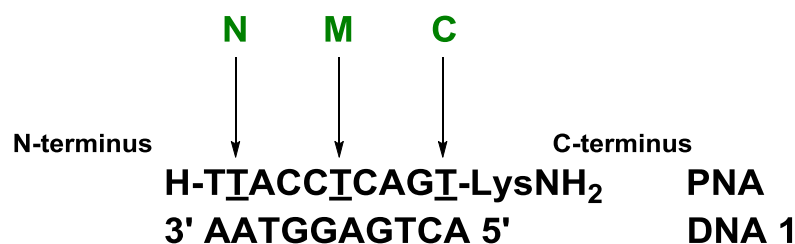
FIGURE 3.31 Schematic representation showing the position of modification incorporated

3.7.2 Comparative study of γ-(*S-eam*) *aeg* PNA and γ-(*S-egd*) *aeg* PNA

For both γ-(*S-eam*) *aeg* PNA and γ-(*S-egd*) *aeg* PNA, the UV-T_m results indicate that among the single modifications (C-terminus, N-terminus and middle), the stability of PNA:DNA duplex was more pronounced when modification was incorporated at C-terminus into the sequence (Figure **3.32**). The order of induced stability for ethyleneamino (*eam*) modified PNA was C-terminus > N-terminus > Middle while that for ethyleneguanidino (*egd*) modified PNA was C-terminus > Middle > N-terminus. The possible explanation for observed higher stabilization at C-terminus modification is the interaction of positively charged (amine/guanidine) side chain with

198

the PNA dipole. Moreover, the stability of PNA:DNA duplex was found to increase with increase in number of modified units (amino/guanidino) into the system.

FIGURE 3.32 Comparative ΔT_m studies for PNA:DNA duplex formed by γ-(*S-eam*) *aeg* PNA and γ-(*S-egd*) *aeg* PNA modifications on *aeg* PNA backbone at various positions on mixed decamer sequence with complementary DNA. (ΔT_m indicates the difference in T_m with control *aeg* PNA)

3.7.3 Comparative study of γ-(*S-azm*) *aeg* PNA and γ-(*S-azb*) *aeg* PNA

Incorporation of neutral azide functionality in the *aeg* PNA backbone at γ-position gives better stabilizing effect when connected with a longer chain (4 X CH$_2$) to the backbone as compared to azide at shorter distance (1 X CH$_2$) (Figure **3.33**). In case of azidomethylene (*azm*) PNAs, the PNA:DNA duplex stability is better when modification is placed in the middle of the sequence than that of C-terminus. The N-terminus and doubly (one towards N-terminus and the other at middle) modified PNAs showed stability similar to that of unmodified *aeg* PNA. In case of longer side chain azidobutylene (*azb*) PNAs, single modification at C-terminus gave higher stability than the N-terminus and middle modifications. N-terminus modification was found to be least stabilizing among all *azb* modifications whereas incorporating two *azb* modified units (one at N-terminus and the other in the middle) magnifies the PNA:DNA duplex stability.
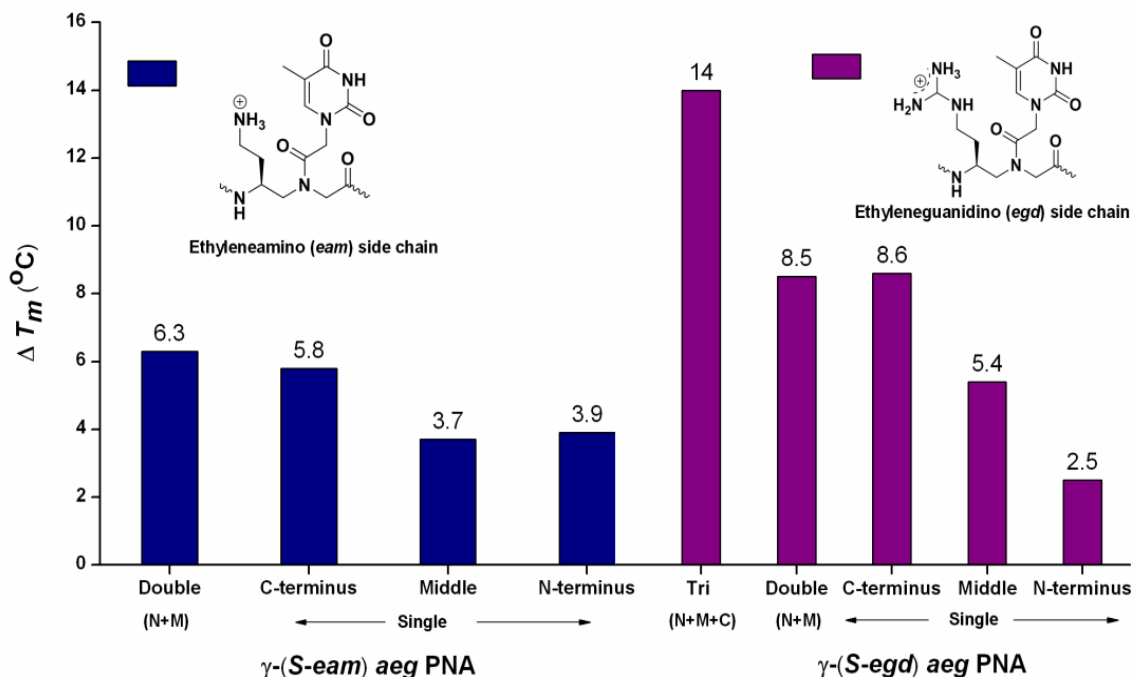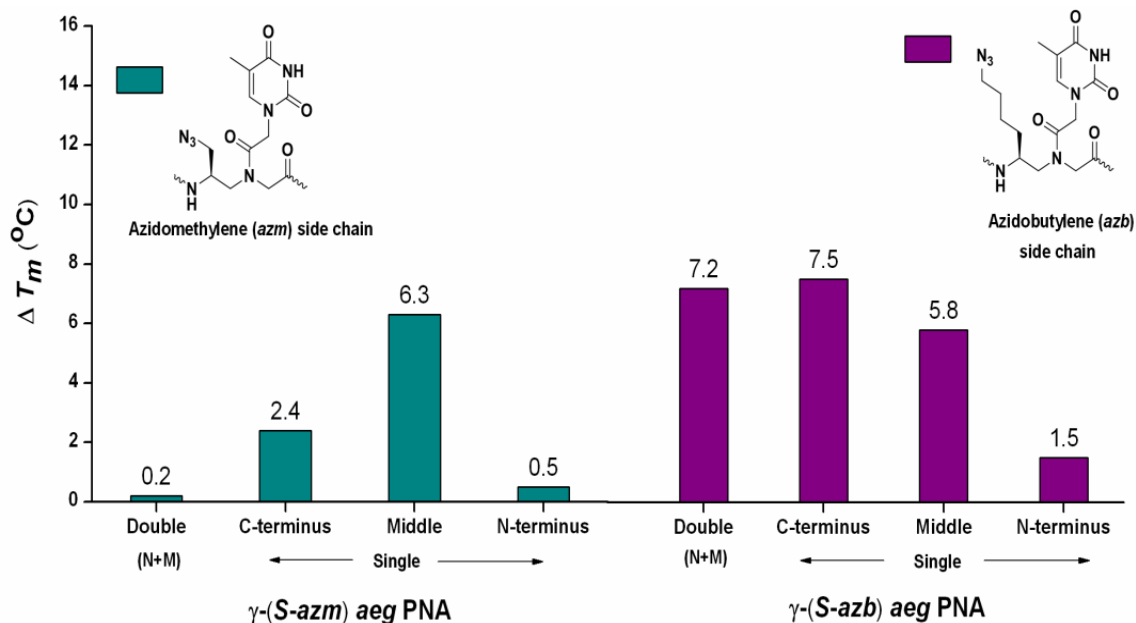
FIGURE 3.33 Comparative ΔT_m studies for PNA:DNA duplex formed by γ-(*S-azm*) *aeg* PNA and γ-(*S-azb*) *aeg* PNA modifications on *aeg* PNA backbone at various positions on mixed decamer sequence with complementary DNA. (ΔT_m indicates the difference in T_m with control *aeg* PNA)

3.7.4 Comparative study of *azm*-t(f) PNA and *azb*-T(f) PNA

Strategic placement of fluorophores in the PNA backbone is useful for diagnostic purpose. In the present study, 5(6)-Carboxyfluorescein was attached at γ-position via copper-mediated azide-alkyne cycloaddition (CuAAC) reaction and the fluorescent analogs were investigated for binding ability to the complementary DNA target. The observed UV-T_m results (Table **3.7**) indicated that the bulky carboxyfluorescein is tolerated when placed away from the PNA backbone as in butylenes side chain analogs (Figure **3.34**). The incorporation of fluorophore at C-terminus stabilized the PNA:DNA duplex better than that in the centre whereas incorporation of two fluorophore units at N-terminus and in the middle destabilized the PNA:DNA duplex by 3.4 °C.
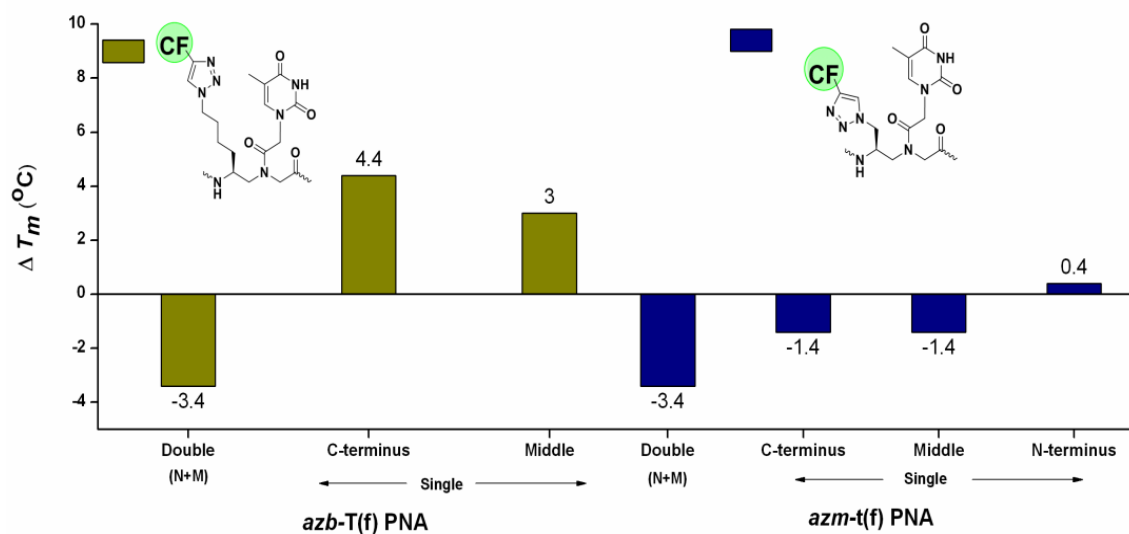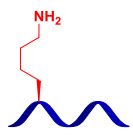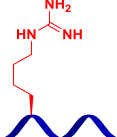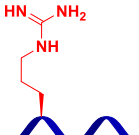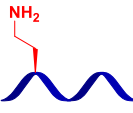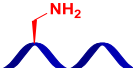
FIGURE 3.34 Comparative ΔT_m studies for PNA:DNA duplex formed by *azm*-t(f) PNA and *azb*-T(f) PNA modifications on *aeg* PNA backbone at various positions on mixed decamer sequence with complementary DNA. (ΔT_m indicates the difference in T_m with control *aeg* PNA)

To investigate the specificity of these modified PNA oligomers towards their target, PNAs were challenged to bind to DNA with a single base mismatch in the middle of the sequence. UV-T_m mismatch studies showed that the modified PNA oligomers are highly specific compared to unmodified *aeg* PNA. The mismatch discrimination observed for modified PNA oligomers was remarkably higher than the control *aeg* PNA.

3.7.5 Effect of carbon chain length on thermal stability of cationic PNA:DNA duplexes

The PNA:DNA duplex stabilization was found to be dependent on the carbon chain length used as a linker between the PNA backbone and the cationic (amino/guanidino) functional groups at γ-position. The observed results are consistent with literature studies, wherein the PNA:DNA duplex stability increases on reducing the carbon chain length from 4-carbons to 1-carbon. The relative ΔT_m values showing the effect of carbon chain length on the thermal stability of PNA:DNA duplexes have been shown in the Table **3.9**.

TABLE 3.9 Comparative thermal stabilities of cationic γ-C-Substituted PNAs

Carbon Chain Length	Modification Incorporated		ΔT_m (°C) per modification		Reference
	Amino	Guanidino	Amino	Guanidino	
4			1.7	2	10, 11
3	NR		NR	3	12
2			3	5	**Present Work**
1		NR	6.5	NR	13, 14

ΔT_m indicates the difference in T_m with control *aeg* PNA; NR: Not reported

CD spectroscopic studies showed that the incorporation of cationic amino/guanidine or charge-neutral azide functionalities through various carbon chain spacers does not disturb the overall PNA:DNA duplex structures. The CD signatures observed for all modified PNA:DNA duplexes were similar to unmodified *aeg* PNA:DNA duplex which forms a right handed helix like in B-form DNA. Also, the incorporation of bulky fluorophore (via click reaction) does not alter the overall conformation of PNA:DNA duplex.

γ-C-substituted PNA oligomers were tested to displace the intercalated ethidium bromide from the duplex DNA. Surprisingly, the guanidino modified PNAs were least efficient at displacing ethidium bromide from the *ds*DNA followed by amino modified PNAs. Also, increase in number of cationic functional groups in the 10 mer *aeg* PNA does not enhance the ethidium bromide displacement ability. However, charge neutral azido modified PNAs and control *aeg* PNA which is also neutral in charge were highly

efficient at displacing ethidium bromide from the duplex DNA. Therefore, the order of ethidium bromide displacement by various PNA is:

Azido PNA ≥ *aeg* PNA > amino PNA > Guanidino PNA

The formation of PNA:DNA duplexes was confirmed by the differential mobility of single stranded nucleic acid fragments and their duplex on polyacrylamide gel. The negatively charged single stranded DNAs move faster under electrophoretic conditions compared to charge-neutral *aeg* PNA and cationic guanidino modified PNA. In case of duplexes, cationic guanidine modified PNA**10**:DNA**1** duplex was highly retarded followed by *aeg* PNA**1**:DNA**1** which was followed by DNA**1**:DNA**3** duplex. Competitive binding experiment of PNA and DNA towards their complementary DNA strand was also carried out using gel electrophoresis which showed that both control *aeg* PNA and cationic guanidine modified PNA compete to bind to target DNA.

3.8 Conclusions

Following important inferences can be drawn from the biophysical studies on the cationic (amine/guanidine) and charge-neutral azide modified PNA oligomers:

➢ All modified (amine, guanidine or azide) PNA oligomers showed higher binding affinity for complementary DNA when placed at any position in the sequence.

➢ Modified PNA oligomers bind to mismatch DNA with lower stability than for complementary DNA suggesting a higher sequence specificity and remarkable mismatch discrimination compared to control *aeg* PNA.

➢ γ-(*S-egd*) *aeg* PNAs showed an increased thermal stability among all unmodified and modified PNAs with ~ 5 °C increase in T_m per modification whereas γ-(*S-eam*) *aeg* PNAs showed ~ 3 °C increase in T_m per modification. These findings suggest that the site of modification of cationic substitutions in the *aeg* PNA backbone is important.

➢ Shorter the distance of cationic functional group from the backbone at γ-position, better the thermal stability with complementary DNA.

- In case of azido PNAs, the azidobutylene (longer side chain) modified PNAs were better in terms of binding affinity and sequence specificity compared to azidomethylene (shorter side chain) modifications.

- The bulky fluorophore when placed away (4 x CH_2) from the PNA backbone gave higher thermal stability of the duplex suggesting the role of steric interactions. Fluorescence studies for single stranded fluorescently labeled PNA as well as their duplexes with complementary DNA showed that the fluorescence intensity for PNA:DNA duplex was slightly less than that of single stranded PNA.

- CD studies showed that incorporation of various modified units in the *aeg* PNA does not alter the structure of PNA:DNA duplex which forms right handed helix when *S*-stereochemistry is used at γ-position.

- Although the thermal stability was better for cationic PNAs, they were less efficient at displacing ethidium bromide (EtBr) from the DNA duplex perhaps due to charge repulsion between positively charged EtBr and cationic PNAs. However, neutral PNAs were highly efficient at displacing ethidium bromide.

- Electrophoretic mobility shift assay showed more retardation in case of PNA:DNA duplex than DNA:DNA duplex. Also both control *aeg* PNA and guanidino modified PNA showed competition towards target DNA.

3.9 Summary

To summarize, this chapter deals with the biophysical investigation of modified PNA oligomers using temperature dependent UV-visible spectroscopy, CD spectroscopy and fluorescence spectroscopy. The differential binding ability of modified PNA oligomers to the complementary DNA was also investigated by ethidium bromide displacement and gel retardation assays (Figure **3.35**). The next chapter deals with the cell penetration abilities of these modified PNA oligomers.
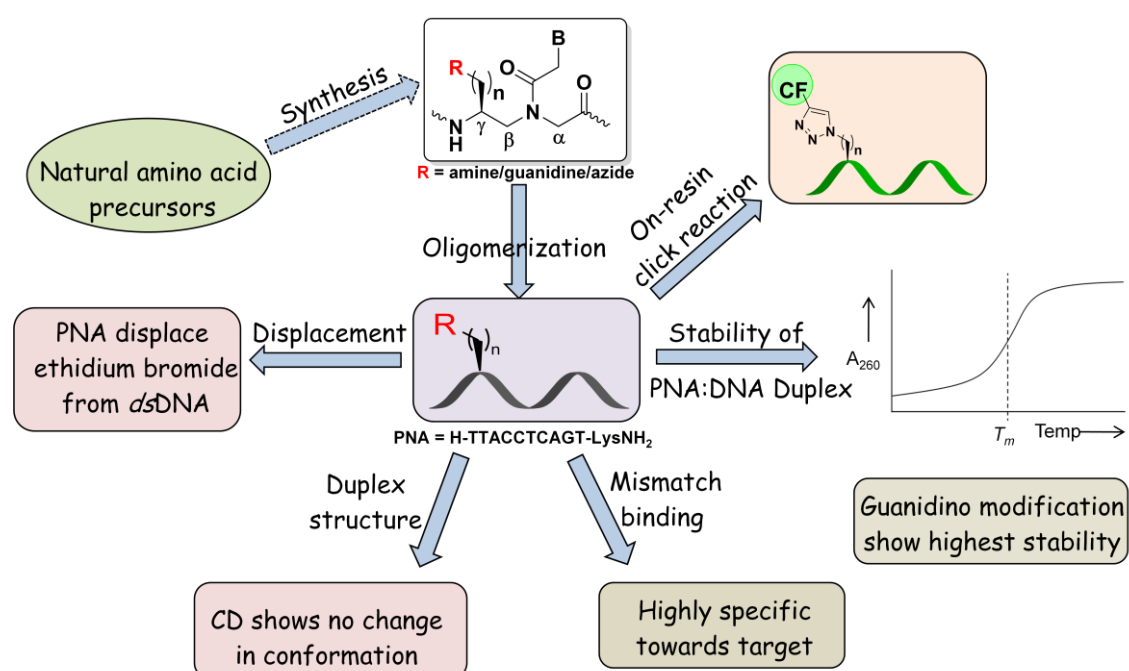

FIGURE 3.35 Summary of biophysical studies of modified PNA oligomers

3.10 Experimental procedures

Chemicals: The unmodified and modified PNA oligomers were synthesized manually by general solid phase PNA synthesis protocol using *Boc*-strategy. Complementary and mismatch DNA oligonucleotides were obtained commercially from Integrated DNA Technologies (IDT). Salts and reagents used in buffer preparation such as NaCl, Na_2BO_4, NaH_2PO_4, Na_2HPO_4, and Tris base etc. were obtained from Sigma-Aldrich. Reagents like acrylamide, *N,N'*-methylenebisacrylamide, ammonium persulphate (APS), tetramethylethylenediamine (TEMED) sucrose, glycerol etc. were procured from Sigma Aldrich to use for gel electrophoresis. The pH of the buffer solutions was adjusted using NaOH or HCl. Ethidium bromide for displacement assay was also obtained from Sigma Aldrich.

3.10.1 UV-T_m measurement

UV-melting experiments were carried out on Varian Cary 300 UV-spectrophotometer equipped with a peltier. The samples for T_m measurement were prepared by mixing the calculated amounts of respective oligonucleotides in stoichiometric ratio (1:1, for duplex) in sodium phosphate buffer (10 mM) containing EDTA (0.1 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 2 μM for each strand. The samples were annealed by heating at 90 °C for 3 min. followed by slow cooling to room temperature for at least 6-8 h and then refrigerated for at least 4 to 5 h. The samples (500 μL) were transferred to quartz cell and equilibrated at the starting temperature for 2 min. The OD at 260 nm was recorded in steps from 20-85 °C with temperature increment of 1 °C/min. Each melting experiment was repeated at least twice. The normalized absorbance at 260 nm was plotted as a function of the temperature. The T_m was determined from the first derivative of normalized absorbance with respect to temperature and is accurate to ± 0.5 °C. The data were processed using Microcal Origin 8.0/8.5. [The concentration of all oligonucleotides were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases i.e. T = 8.8 cm^2/μmol; C = 7.3 cm^2/μmol; G = 11.7 cm^2/μmol and A = 15.4 cm^2/μmol]

3.10.2 Circular dichroism

CD spectra were recorded on JASCO J-815 spectropolarimeter connected with a peltier. The calculated amounts for PNA oligomers and the complementary DNA 1 were mixed together in stoichiometric ratio (1:1, for duplex) in sodium phosphate buffer (10 mM) containing EDTA (0.1 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 5 μM for each strand. The samples were annealed by heating at 90 °C for 3 min. followed by slow cooling to room temperature for at least 6-8 h. The cooled samples were transferred to refrigerator for at least 4 to 5 h. To record the CD spectra of PNA:DNA duplexes and single stranded PNAs, the temperature of circulating water was kept at 10 °C. The CD spectra were recorded as an accumulation of 5 scans from 300 to 190 nm using 2 mm cell, a resolution of 0.1 nm, band-width of 1 nm, sensitivity of 2 m deg, response of 2 sec and a scan speed of 50 nm/min.

3.10.3 UV-visible and fluorescence studies

UV-Visible spectra for all PNA oligomers and complementary oligonucleotides were recorded on Perkin Elmer Lambda-45 UV-Visible Spectrophotometer and for fluorescence studies, experiments were performed on Horiba Jobin Yvon Fluorolog 3 spectrophotometer. The samples for fluorescence spectra were prepared by mixing calculated amounts of PNA and DNA in stoichiometric ratio (1:1, for duplex) in sodium phosphate buffer (10 mM) containing EDTA (0.1 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 1 μM for each strand. The annealed samples were used to record fluorescence spectra in a rectangular quartz cell at ambient temperature with λ_{exc} 458 and 484 nm; λ_{em} 520 nm; excitation slit width of 1 nm and emission slit width of 6 nm.

3.10.4 Ethidium bromide displacement assay from *ds*DNA

In case of ethidium bromide displacement, calculated amounts of DNA strands were mixed together to get a DNA duplex (2 μM, 400 μL) in the above said buffer system. The annealed samples were titrated against EtBr (250 μM) in the increment of 1 μL and the fluorescence intensity at 610 nm was recorded to achieve saturation point which was found to be 8 μL, 5.58 μM EtBr. The resulting EtBr-*ds*DNA complex was

titrated individually against various PNA oligomers in the aliquots of 0.4 μL and change in fluorescence intensity at 610 nm was recorded. The fluorescence decay observed upon displacement of EtBr was plotted against the concentration of PNA oligomer using Microcal Origin 8.5.

3.10.5 Electrophoretic mobility shift assay and competition binding experiment

The calculated amounts of PNA oligomers (3 nmoles) were individually mixed with DNA strands (3 nmoles) in 1:1 ratio in sodium phosphate buffer (10 mM, NaCl 100 mM, pH 7.2; 10 μL). The samples were annealed by heating to 85 °C for 3 min followed by slow cooling to RT and refrigeration at 4 °C overnight. To this, 5 μL loading buffer (10 mM Tris-HCl, pH 8; 30 % sucrose and 10 % glycerol) was added and samples were loaded on the gel. Bromophenol blue (BPB) was used as the tracer dye and separately loaded in an adjacent well. Gel electrophoresis was performed on 20 % nondenaturing polyacrylamide gel (acrylamide:*N*,*N'*-methylenebisacrylamide, 19:1) with 1X-TBE containing 500 mM NaCl as tank buffer at constant power supply of 14 W and 120 V for 30 h. During electrophoresis, the temperature was maintained at ~18 °C. The spots were visualized through UV shadowing by illuminating the gel placed on a fluorescent silica gel plate, F254 using UV light.

3.11 References

1. Cantor, C. R.; Warshaw, M. W.; Shapiro, H. *Biopolymers* **1970**, *9*, 1059-1070.

2. Puglisi, J. D.; Tinoco, I. Jr. *Methods Enzymol*. **1989**, *180*, 304-325.

3. (a) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897. (b) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.

4. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Frier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Neilsen, P. E. *Nature*, 1993, *365*, 566-568.

5. Brand, L.; Johnson, M. L. *Methods Enzymol*. **1997**, *278*, 1-628.

6. Cain, B. F.; Baguley, B. C.; Denny, W. A. *J. Med. Chem*. **1978**, *21*, 658-668.

7. Gershan, H.; Chiralando, R.; Guttman, S. B.; Minsky, A. *Biochemistry* **1993**, *32*, 7143.

8. (a) Hendrickson, W. *BioTechniques* **1985**, *3*, 346-354. (b) Revzin, A. *BioTechniques* **1989**, *7*, 346-354.

9. Sambrook, J.; Russel, D. W. *Molecular Cloning: A Laboratory Manual* 3rd Ed. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York, **2001**.

10. Enguld, E. A.; Appella, D. H. *Angew. Chem. Int. Ed*. **2007**, *46*, 1414-1418.

11. Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas S. M.; Zon, G.; Livak, K. J.; Ly, D. H. *J. Org. Chem.* **2009**, *74*, 1509-1516.

12. Manicardi, A.; Calabretta, A.; Bencivenni, M.; Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. *Chirality* **2010**, *22*, 161-172.

13. Mitra, R.; Ganesh, K. N. *Chem. Commun*. **2011**, *47*, 1198-1200.

14. Mitra, R.; Ganesh, K. N. *J. Org. Chem*. **2012**, *77*, 5696-5704.

15. Wttung, P.; Kim, S. K.; Buchardt, O.; Nielsen, P.; Norden, B. *Nucleic Acids Research* **1994**, *22*, 5371-5377.

Chapter 4

Cell Permeation Studies of PNA Oligomers

The cell penetration ability of carboxyfluorescein tagged PNA oligomers was investigated by live cell imaging in NIH 3T3 and MCF-7 cell lines using confocal microscopy. The quantification of cell permeation has been carried out using Fluorescence Activated Cell Sorter (FACS) analysis. The results show those cationic functional groups incorporated in the side chain at γ-position remarkably increase the efficiency of cell permeation.

4.1 Introduction

Synthetic oligonucleotides have been proposed as a new class of potential therapeutic molecules that can interact with the messenger RNA (mRNA) of disease-related protein which specifically inhibits the protein synthesis.[1,2] The fascinating approach of antisense strategy represents one of the most powerful systems for molecular recognition designed by nature.[3] Naturally occuring oligonucleotides (DNA/RNA) suffer from the various hurdles to act as potential drug candidates. Several classes of modified oligonucleotides (ONs) have been developed that can recognize and bind to target DNA/RNA with high affinity and sequence specificity. The modified oligonucleotides are resistant to degradation by cellular enzymes like proteases and nucleases, however none of them can efficiently traverse the cell membrane on their own.[4] Various ways have been attempted to improve the cell permeation properties of the antisense oligonucleotides.[5] The most important ones include the use of liposome as a carrier of antisense ONs where the therapeutic ONs are encapsulated in the liposome and specifically delivered to the target cells.[6] The use of liposomes is limited due to their short half life in the serum. The attachment of poly-L-lysine[7] and additional chemical, mechanical and electrical transduction means have been employed to improve the transport of these ONs to the cells.[8] Although these ways are useful for small-scale experimental setups, the use of such exogenous transduction reagents may lead to off-target and cytotoxic effects.[9] It is therefore desirable to modify the structures and chemical functionalities of these ONs to improve the cell permeability as well as pharmacokinetic and pharmacodynamic properties.

As discussed earlier, the attractive features in terms of binding affinity, strand invasion, resistance to degrading enzymes and high stability in cellular extracts[10] make peptide nucleic acid (PNA) a promising molecule for the development of gene-specific drugs. However, the low cellular entry of PNA limits their usefulness as an antigene/antisense drug.[11] Therefore, several different ways have been employed to increase the cellular delivery of unmodified PNAs.[12] Some of the transfection protocols for PNA are listed below:

a) Microinjection
b) Electroporation

c) Co-transfection with DNA

d) Permeabilised cells

e) Direct delivery at high concentrations

However, the faster growth of PNA development as an antisense agent depends on the finding of more efficient and general delivery protocols for PNA. To meet these demands, several modifications of PNA have been introduced and different cellular delivery protocols have been attempted. One such modification is the conjugation of PNA to lipophilic adamantyl acetic acid. But this approach was partially successful giving the cell permeability depending on the cell type and the PNA sequence used.[13] Similarly another lipophilic moieties like triphenylphosphonium, cholesterol, bile acid etc. have been conjugated to improve the cell permeability of PNA.[14] Considerable effort has been invested in exploring the effect of 'Trojan peptides' and several other 'cell penetrating peptides' as carriers for cellular delivery of PNA. The most thoroughly studied peptides are penetratin, a 16 residue peptide derived from Drosophila homeodomain transcription factor, antenapedia and 35-amino acid basic sequence (TAT) from the HIV tat protein.[15] PNA when conjugated to penetratin or its retroinverso derivative has shown diffusion in cytoplasm or localization in the nucleus. Transportan, a 27 amino acid long peptide and 7-residue basic nuclear localization signal (NLS) peptide have also been investigated as conjugates of PNA to improve the antisense activity.[16] The risk of adverse reactions in non-targeted cells have been tried to avoid by conjugation of PNA to cell-specific receptor ligands. Wickstrom et al. have conjugated PNA to a four residue peptide with specific binding capability to the cell surface receptor for insulin like growth factor 1 (IGF-1R). These studies showed that only IGF-1R expressing cells internalized conjugated PNA enabling specific entry to the cells. Unfortunately, the PNA was entrapped in the cytosolic compartments limiting the availability of antisense PNA to the target.[17] To explore this strategy, Corey's group has conjugated anti-telomerase PNA to lactose which can be recognized by the hepatic asialoglycoprotein receptor (ASGP-R) where it was found that among all cell types used, only the liver derived cell type internalized the PNA conjugate.[18] The general routes taken by these PNA conjugates to get inside the cells are either direct translocation or endocytosis as shown in Figure **4.1**[19].
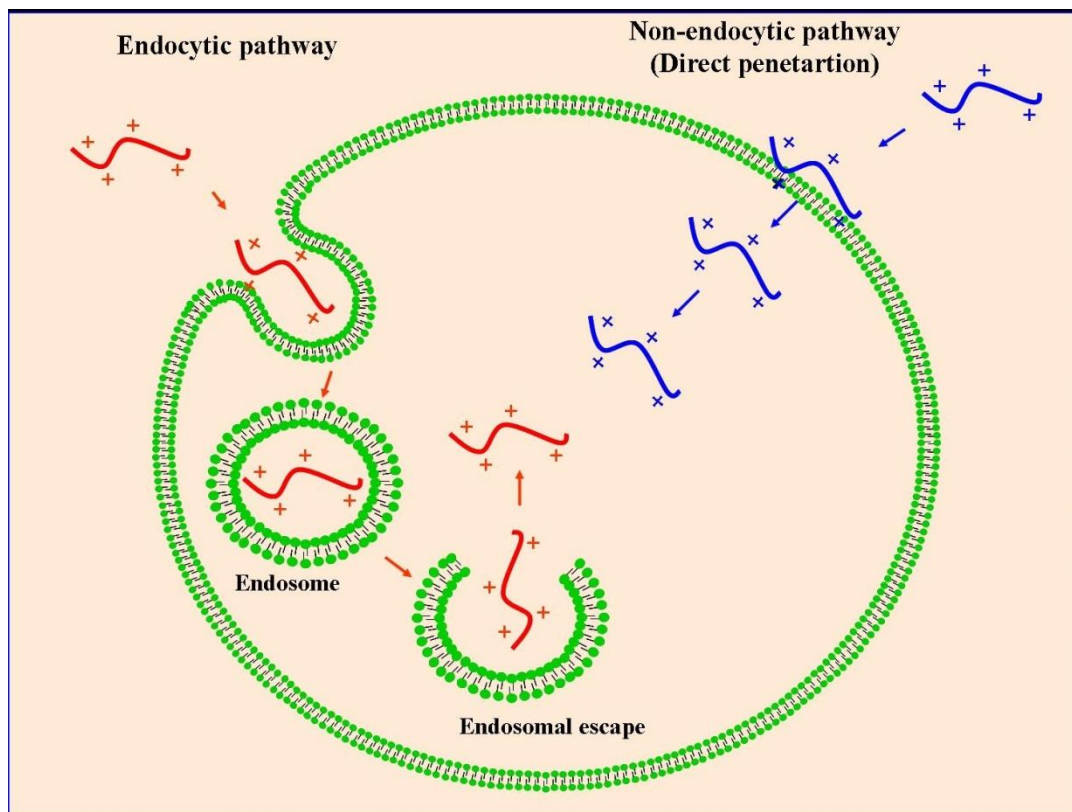
FIGURE 4.1 General routes for cellular entry of PNA or its conjugates[19]

The various protocols mentioned to improve the cellular uptake of PNA have suffered from the obstacles like endosomal entrapment, cytotoxicity induced by conjugating highly cationic cell penetrating peptides or polylysine and polyarginine, low plasma half life of liposomes etc. Therefore, increasing the cell permeation ability of PNA and improvement in its use as effective antisense agent is still in progress that can help in producing more nucleic acid-based therapeutic approaches and diagnostics.

4.2 Rationale of the work

Compounds within a narrow range of molecular size, polarity and net charge are able to diffuse through cell membranes, which remains as an impermeable barrier to the rest of macromolecules. Cell membrane impermeability is the major limiting factor for PNA for use as antisense agents. In order to overcome this barrier for effective delivery of membrane-impermeable PNA, cationic cell penetrating peptides have been designed, but resulted in increased cytotoxicity. Attachment of polylysine (cationic amines) and polyarginine (cationic guanidines) residues have improved the cell penetration but suffer from cytotoxicity[20]. Therefore, the synthesis of inherently cationic peptide

nucleic acids was expected to improve the binding affinity to the target and enhance the intrinsic cell permeability. The cationic guanidine functional group greatly contributes to the enhancement of cellular uptake by forming stronger bidentate hydrogen bond network with the negatively charged carboxylates, phosphates and sulfates[21] (Figure **4.2**). The cationic amine functionality (with single hydrogen bonding site) has also been shown to improve the cell entry but less efficiently than guanidinium groups. To explore the strategy of making inherently cationic PNA, ethyleneamino (*eam*) and ethyleneguanidino (*egd*) substitutions have been introduced at γ-position in the PNA backbone. The rationally designed novel amino and guanidino modified peptide nucleic acids have been shown to bind to target DNA with higher thermal stabilities as compared to control *aeg* PNA (Chapter **3**).
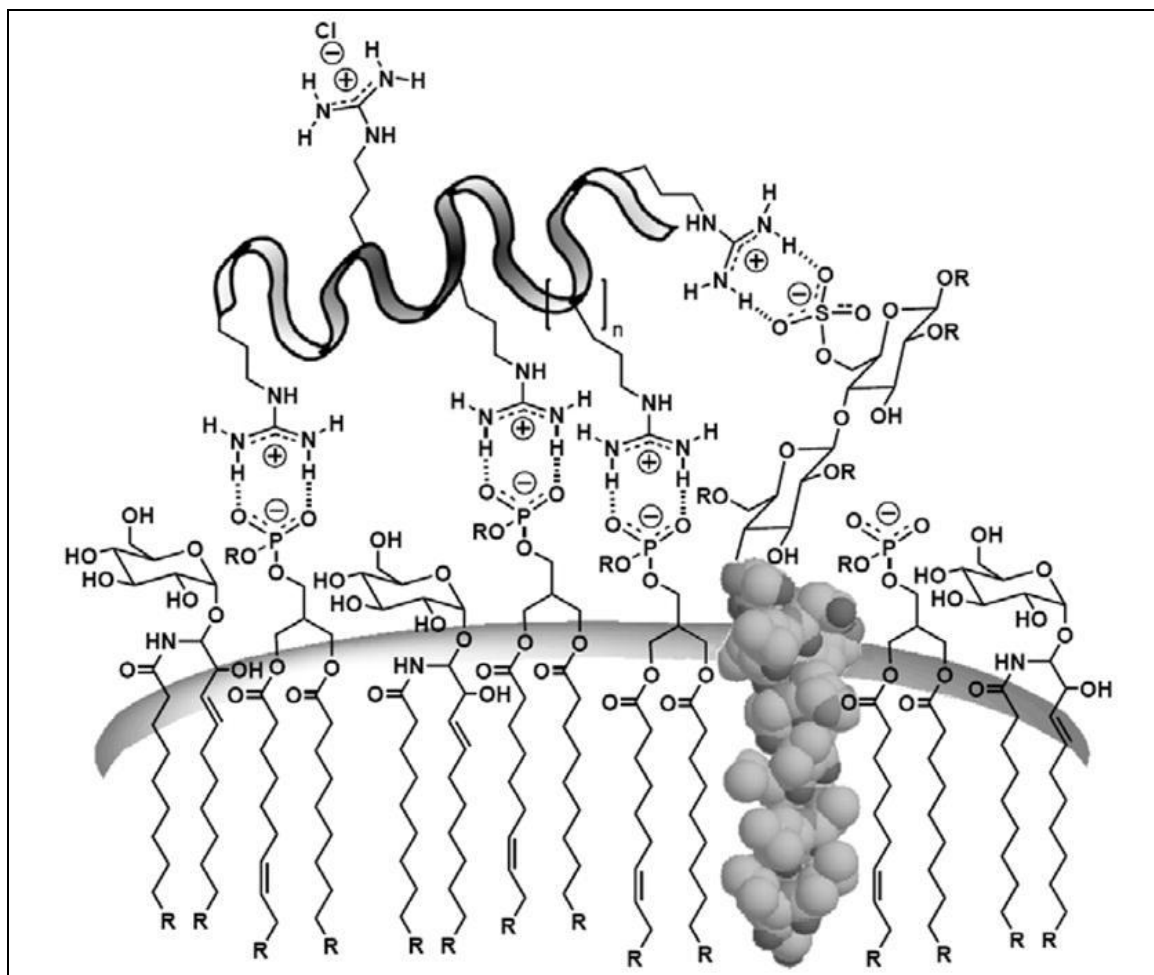

FIGURE 4.2 Association of cationic guanidinium groups with anionic cell membrane constituents[21]

In order to produce more information about these inherently cationic peptide nucleic acids at cellular level, the cell uptake properties have been targeted to be

determined on various cell lines. This chapter elucidates the cell permeation studies based on the number of amino and guanidino modified PNA units incorporated in the *aeg* PNA sequence. The quantification of the cell entry of these modified PNAs has been investigated by Fluorescence Activated Cell Sorter (FACS) studies. The chiral and charge-neutral azide modified PNAs were tested for cell penetration to compare with charge effects of cationic amino and guanidino modified PNAs and also with control achiral, neutral *aeg* PNA.

4.3 Aim of the present work

The specific objectives of this section are:

➢ Synthesis of 5(6)-carboxyfluorescein N-terminal-tagged PNA oligomers
➢ UV-visible and fluorescence studies of tagged fluorescent PNA oligomers
➢ Investigation of the uptake efficiency of modified and control *aeg* PNA oligomers into NIH 3T3 and MCF-7 cell lines by confocal microscopy
➢ Quantification of cell permeation abilities of PNA oligomers by FACS studies

4.4 Results and discussion

The synthesis, purification and characterization of fluorescently tagged PNA oligomers have been discussed in this section.

4.4.1 Tagging of PNA oligomers with 5(6)-carboxyfluorescein

The modified PNA and control *aeg* PNA sequences were synthesized for studying their cell permeation properties. Using solid phase peptide synthesis protocol, each of these selected sequences were tagged with 5(6)-carboxyfluorescein [5(6)-CF] at the final step in presence of HOBt and DIC, just before their cleavage from the resin (Figure **4.3**).
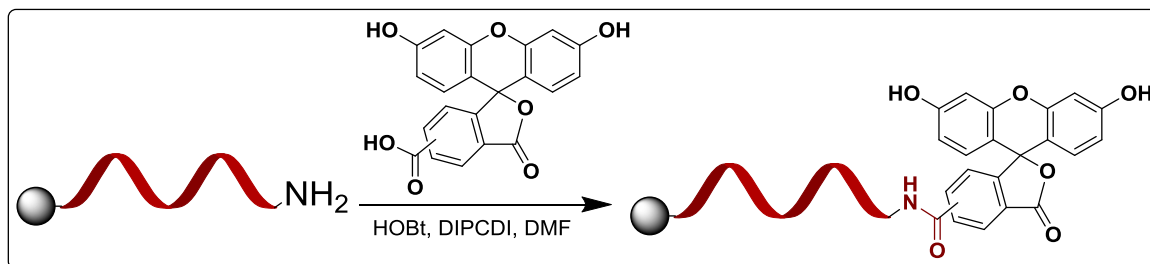

FIGURE 4.3 Tagging of PNA oligomers with fluorescent 5(6)-carboxyfluorescein

4.4.2 Cleavage of the fluorescent PNA oligomers from the solid support

After coupling time was over, the resin was washed thoroughly with methanol to remove excess 5(6)-CF. The peptides were cleaved from the solid support using tifluoromethanesulfonic acid (TFMSA) in the presense of trifluoroacetic acid (TFA) (Low, High TFMSA-TFA method)[20] and thioanisole-ethanedithiol as scavengers which yielded PNA oligomers having L-lysine amide at C-terminus and carboxyfluorescein at N-terminus. After cleavage reaction was over, the PNA oligomers obtained in solution were precipitated by addition of cold diethyl ether and the PNA oligomers were dissolved in de-ionized water. The 5(6)-carboxyfluorescein tagged PNA oligomers are listed in Table **4.1** that have been further used for cellular uptake studies.

TABLE 4.1 5(6)-carboxyfluorescein tagged PNA oligomers

Entry	Sequence code	5(6)-CF tagged PNA oligomers
1	*aeg* PNA **1-CF**	**CF**-T T A C C T C A G T-LysNH$_2$
2	*eam*-t$_2$ PNA **2-CF**	**CF**-T **t** A C C TC A G T-LysNH$_2$
3	*eam*-t$_{2,6}$ PNA **5-CF**	**CF**-T **t** A C C **t** C A G T-LysNH$_2$
4	*egd*-t$^\gamma_2$ PNA **6-CF**	**CF**-T **t$^\gamma$**A C C T C A G T-LysNH$_2$
5	*egd*-t$^\gamma_{2,6}$ PNA **9-CF**	**CF**-T **t$^\gamma$**A C C **t$^\gamma$**C A G T-LysNH$_2$
6	*egd*-t$^\gamma_{2,6,10}$ PNA **10-CF**	**CF**-T **t$^\gamma$**A C C **t$^\gamma$**C A G **t$^\gamma$**-LysNH$_2$
7	*azb*-T$_{10}$ PNA **17-CF**	**CF**-T T A C C T C A G **T**-LysNH$_2$
8	*azm*-T$^\gamma_{10}$ PNA **13-CF**	**CF**-T T A C C T C A G **T$^\gamma$**-LysNH$_2$

aeg = aminoethylglycine, ***eam*** = ethyleneamino, ***egd*** = ethyleneguanidino, ***azb*** = azidobutylene, ***azm*** = azidomethylene and **CF** = 5(6)-carboxyfluorescein

4.4.3 Purification and characterization of synthesized PNA oligomers

After cleavage from the solid support, fluorescently tagged PNA oligomers were purified by reverse phase high performance liquid chromatography (RP-HPLC). The purification of PNA oligomers was carried out on a semi-preparative C18 column using a gradient system of acetonitrile and water. The purity of PNA oligomers was checked by reinjecting the sample on the same C18 semi-preparative column. All HPLC chromatograms are shown in Appendix II.

The integrity of these synthesized PNA oligomers was confirmed by MALDI-TOF mass spectrometry. 2,5-dihydroxybenzoic acid (DHB) was used as a matrix to record MALDI-TOF spectra for all synthesized PNA oligomers. The calculated and observed molecular weights for all PNAs with their molecular formulae and HPLC retention time in minutes are mentioned in Table **4.2**. The MALDI-TOF data for the confirmation of synthesis of tagged PNA oligomers are shown in Appendix II.

TABLE 4.2 MALDI-TOF spectral analyses and HPLC retention times of tagged PNA oligomers

Sr. No.	PNA sequence code	Molecular Formula	Calculated Mass	Observed Mass	Retention Time (min)
1	*aeg* PNA **1-CF**	$C_{134}H_{160}N_{55}O_{39}$	3163.2227 $[M + H]^+$	3163.5088	13.85
2	*eam*-t_2 PNA **2-CF**	$C_{136}H_{165}N_{56}O_{39}$	3206.2649 $[M + H]^+$	3208.2456	13.31
3	*eam*-$t_{2,6}$ PNA **5-CF**	$C_{138}H_{170}N_{57}O_{39}$	3249.3071 $[M + H]^+$	3253.3552	13.33
4	*egd*-t^γ_2 PNA **6-CF**	$C_{137}H_{166}N_{58}NaO_{39}$	3272.1389 $[M + Na]^+$	3276.4036	14.12
5	*egd*-$t^\gamma_{2,6}$ PNA **9-CF**	$C_{140}H_{174}N_{61}O_{39}$	3335.3200 $[M + H]^+$	3337.3982	13.58
6	*egd*-$t^\gamma_{2,6,10}$ PNA **10-CF**	$C_{143}H_{181}N_{64}O_{39}$	3418.4147 $[M + H]^+$	3420.3511	13.25
7	*azb*-T_{10} PNA **17-CF**	$C_{138}H_{166}N_{58}O_{39}$	3261.2130 $[M]^+$	3236.4409*	14.16
8	*azm*-T^γ_{10} PNA **13-CF**	$C_{135}H_{160}N_{58}O_{39}$	3219.1320 $[M]^+$	3193.6467*	14.22

* represents the mass after subtracting N_3 from the actual mass

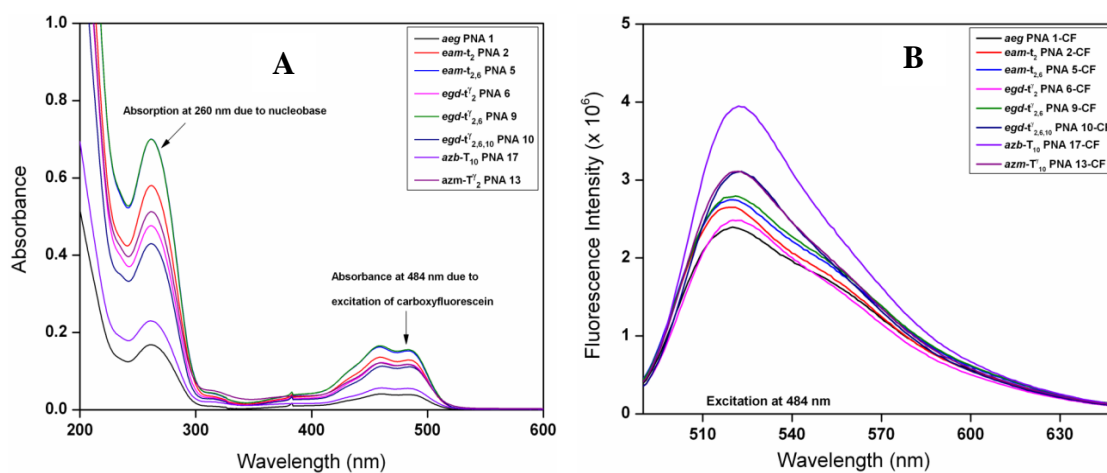## 4.4.4 UV-Visible and fluorescence studies of tagged PNA oligomers

FIGURE 4.4 A) UV-Visible absorption spectra of tagged PNA oligomers and B) Fluorescence spectra of tagged PNA oligomers

216

The absorption spectra were recorded for tagged PNA oligomers which showed absorbance maxima at 260 nm due to nucleobases whereas absorbance maxima at 458 and 484 nm were because of the excitation of 5(6)-carboxyfluorescein unit in the PNA oligomer (Figure **4.4 A**). The fluorescence spectra were also recorded for these tagged PNA oligomers which showed emission maxima at 520 nm when excited at 484 nm wavelength (Figure **4.4 B**).

4.5 Cellular uptake studies

To gain an insight into the cell permeation abilities of these γ-C-substituted multifunctional PNA oligomers, the intracellular distribution of selected PNAs (Table **4.1**) were investigated in two different cell lines (MCF-7 and NIH 3T3). The confocal microscopic technique was used to observe the cellular uptake of 5(6)-carboxyfluorescein tagged PNAs.

4.5.1 Cell lines used in study

a) MCF-7: MCF-7 is a human breast adenocarcinoma cell line and it was first isolated in 1970 from the breast tissue of 69 year old Caucasian woman. MCF-7 cells are useful for in vitro breast cancer studies because the cell line has retained several ideal characteristics particular to the mammary epithelium. These are estrogen receptor (ER) positive control cell lines due to the ability of these cells to process estrogen in the form of estradiol.

b) NIH 3T3: These are the mouse embryonic fibroblast cells isolated and initiated in 1962 at New York university. 3T3 refers to the cell transfer and inoculation protocol for the line, and means '3 day transfer, inoculation 3×10^5 cells.' The 3T3 cells can be inhibited by tenazepam and other benzodiazepines. These cells are also contact inhibited and sensitive to sarcoma virus and leukemia virus focus formation.

4.5.2 Cellular uptake experiment using confocal microscopy

MCF-7 and NIH 3T3 cells were plated in 8-well chambered coverglass in 200 µL Dulbecco's Modified Eagle Medium (DMEM) containing 10 % Fetal Bovine Serum (FBS) at the concentration of 8000 cells per well. The cells were grown by maintaining

at 37 °C in a humidified atmosphere containing 5 % CO_2 for 16 h. The required amounts of 5(6)-carboxyfluorescein tagged PNA stock solutions were added to the corresponding wells to achieve the desired final concentration of 1 μM. The cells incubated with tagged PNA oligomers were maintained at 37 °C in a humidified atmosphere containing 5 % CO_2 for 24 h. After the incubation period was over, the medium was aspirated and cells were rinsed or washed thrice with 1 medium-volume equivalent of ice-cold PBS. The cells were then replenished with 400 μL of OPTIMEM medium containing Hoechst 33342 at a dilution of 1/1000 of the supplied stock solution and left for 30 minutes at 37 °C. After incubation for 30 min, the medium containing nuclear stain Hoechst 33342 was removed and fresh OPTIMEM medium was added to the cells. The cells were then immediately visualized using 40X objective of Zeiss LSM 710 laser scanning confocal microscope. The well of cells without incubation of any PNA oligomer was also visualized as control experiment. The images were acquired with same camera settings for all PNA oligomers including unmodified *aeg* PNA and various modified PNA oligomers. The confocal microscopy imaging has been repeated at least twice for each PNA.

4.5.3 Observations of confocal microscopy images

Cellular permeability of tagged PNA oligomers has been monitored by confocal microscopy images using above said protocol for cell internalization. The results showed that all PNA oligomers including modified and unmodified PNAs can permeate the cell membrane. The green fluorescent signal for 5(6)-CF tagged PNA oligomers was found to be localized within the vicinity of the nucleus which was stained with Hoechst 33342 giving blue color. The punctates observed for PNA oligomers in confocal images suggest that PNAs probably take the endocytotic route to get inside the cells. The pattern of all PNA oligomers getting inside the cells was almost similar irrespective of the cell line used and the modification incorporated in the PNA. Following are the results obtained after the incubation of PNA oligomers for 24 h in MCF-7 and NIH 3T3 cell lines.

A) Cell permeation in MCF-7 cells

The relative uptake efficiency of the PNA oligomers by MCF-7 cell lines is shown in figures **4.5-4.13**. The untreated control cells showed only the blue colored staining in the nucleus by nuclear Hoechst 33342 stain and no green fluorescent signal for PNA internalization (**Figure 4.5**).
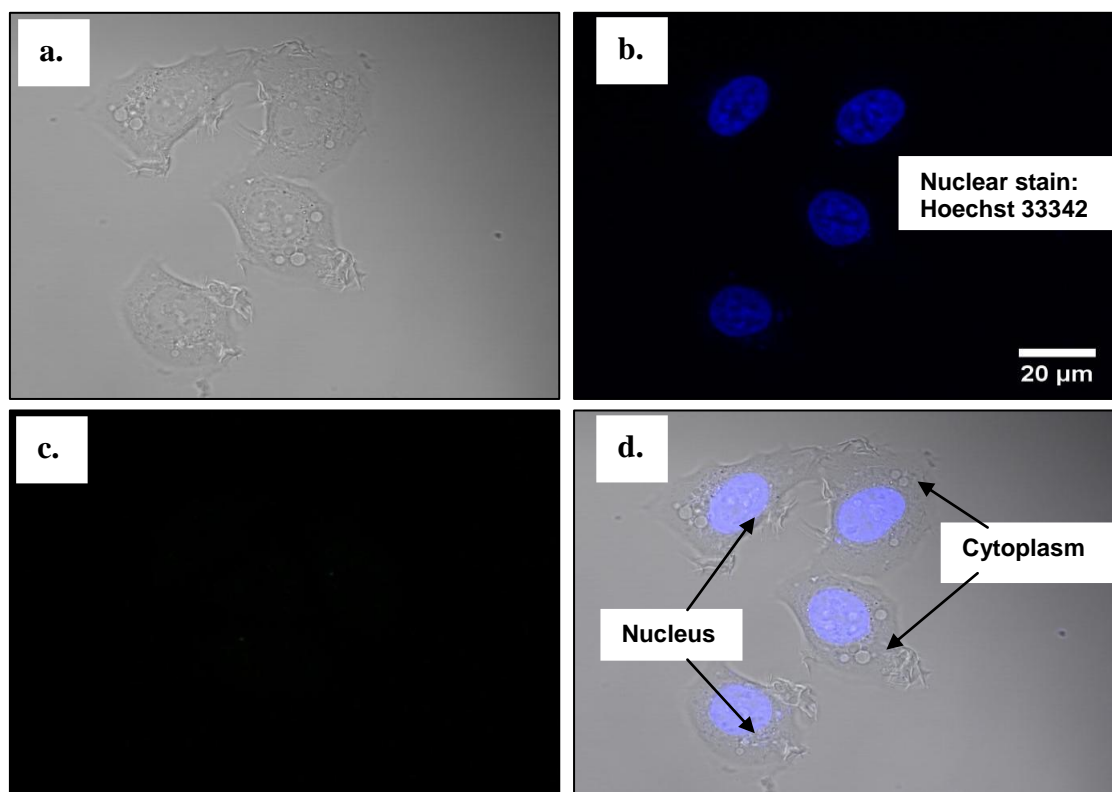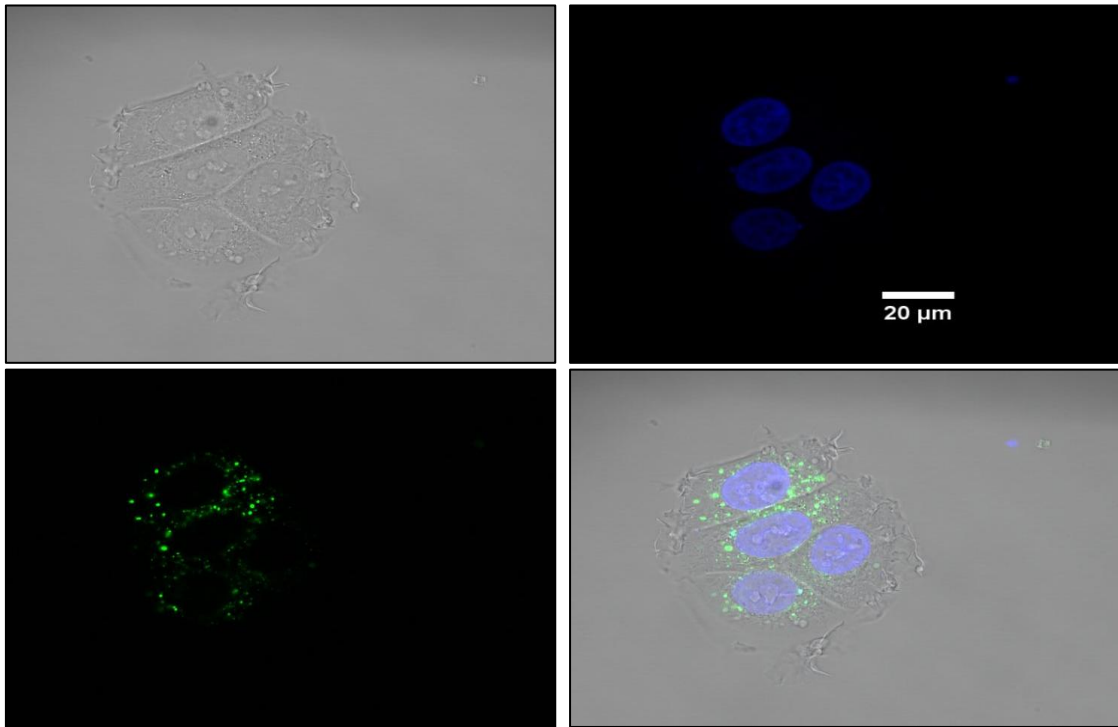

FIGURE 4.5 Confocal microscopy images for control untreated cells a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

The cell permeation of PNA oligomers can be confirmed by visualization of green fluorescence in the cell compartments. A more convincing evidence of the PNA uptake by cells comes from the superimposition of bright field image, fluorescent image and the Hoechst 33342 stained image. The cellular uptake of unmodified *aeg* PNA 1 and various amino, guanidino and azido modified PNA oligomers have been shown below (**Figures 4.6-4.13**). All images are organized in a similar way as shown in Figure **4.5 a**, **b**, **c** and **d**.

FIGURE 4.6 Confocal microscopy images for unmodified *aeg* PNA 1-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
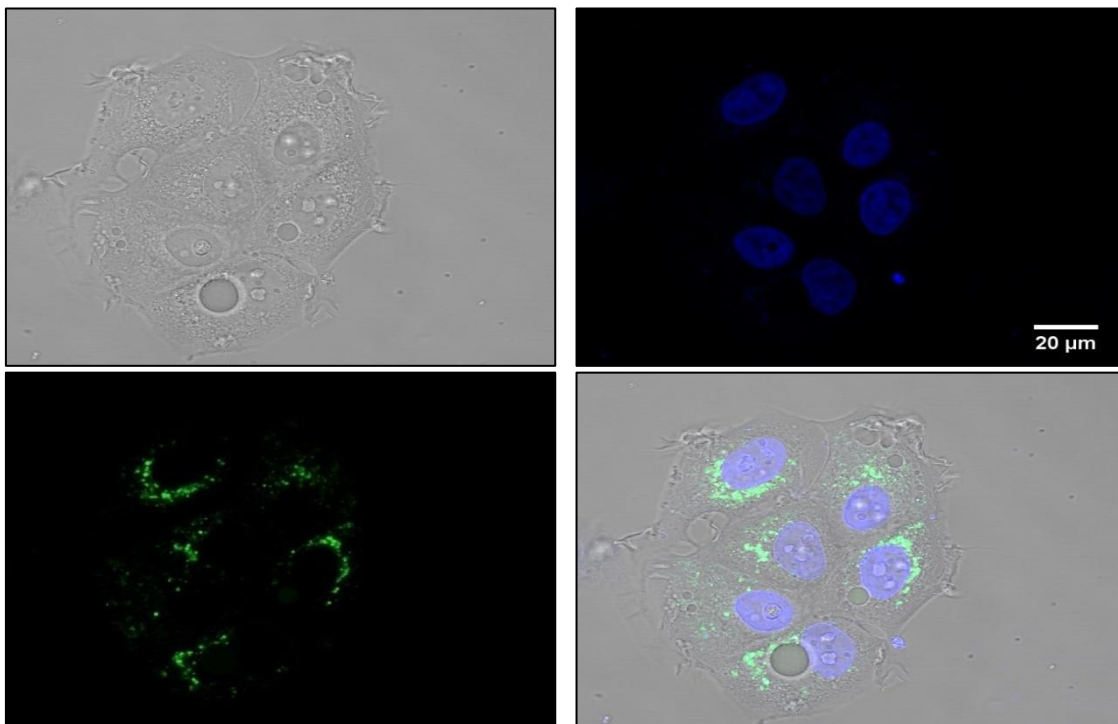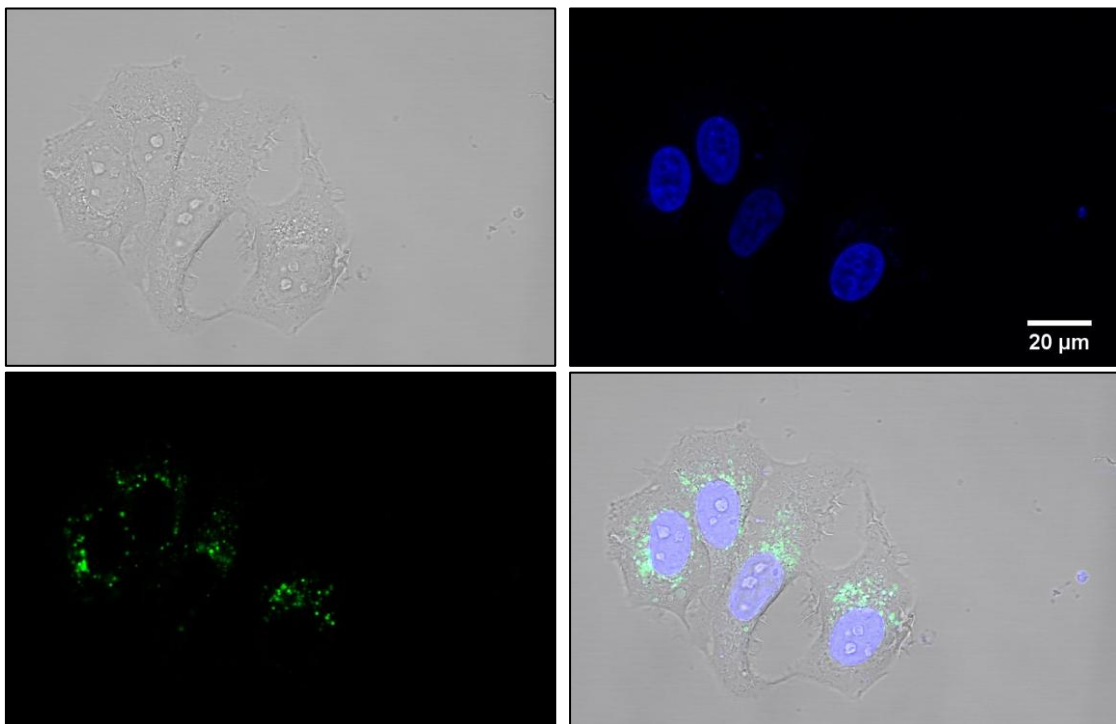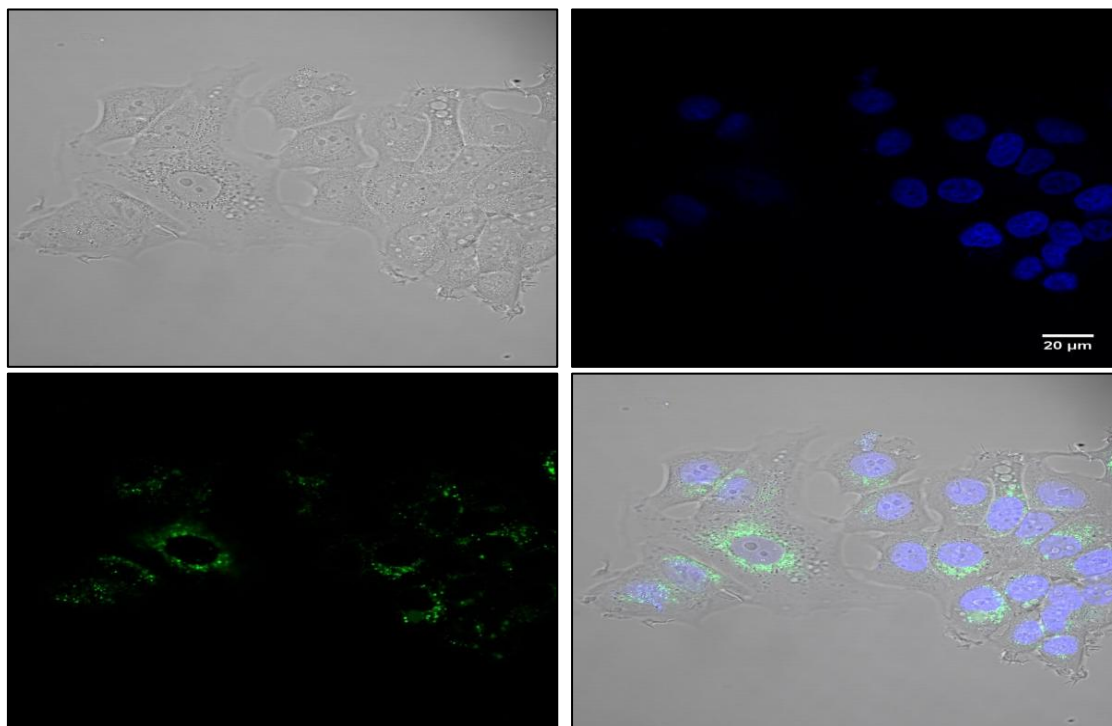

FIGURE 4.7 Confocal microscopy images for 1-amino modified *eam*-t$_2$ PNA 2-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
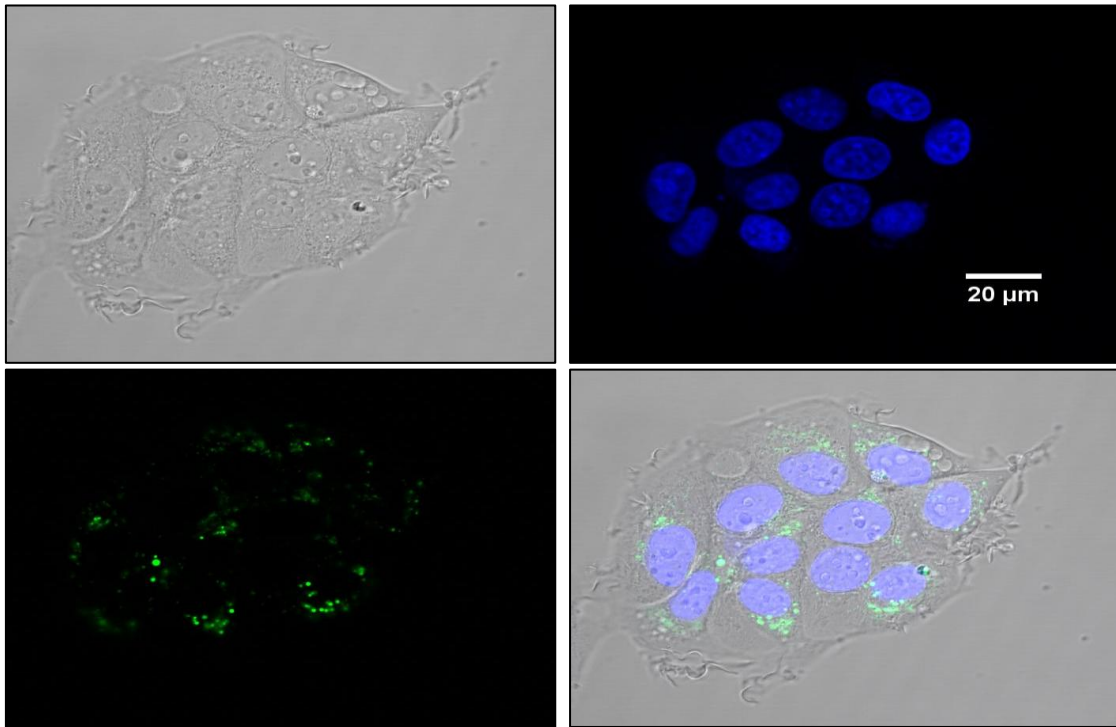
FIGURE 4.8 Confocal microscopy images for 2-amino modified *eam*-t$_{2,6}$ PNA 5-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
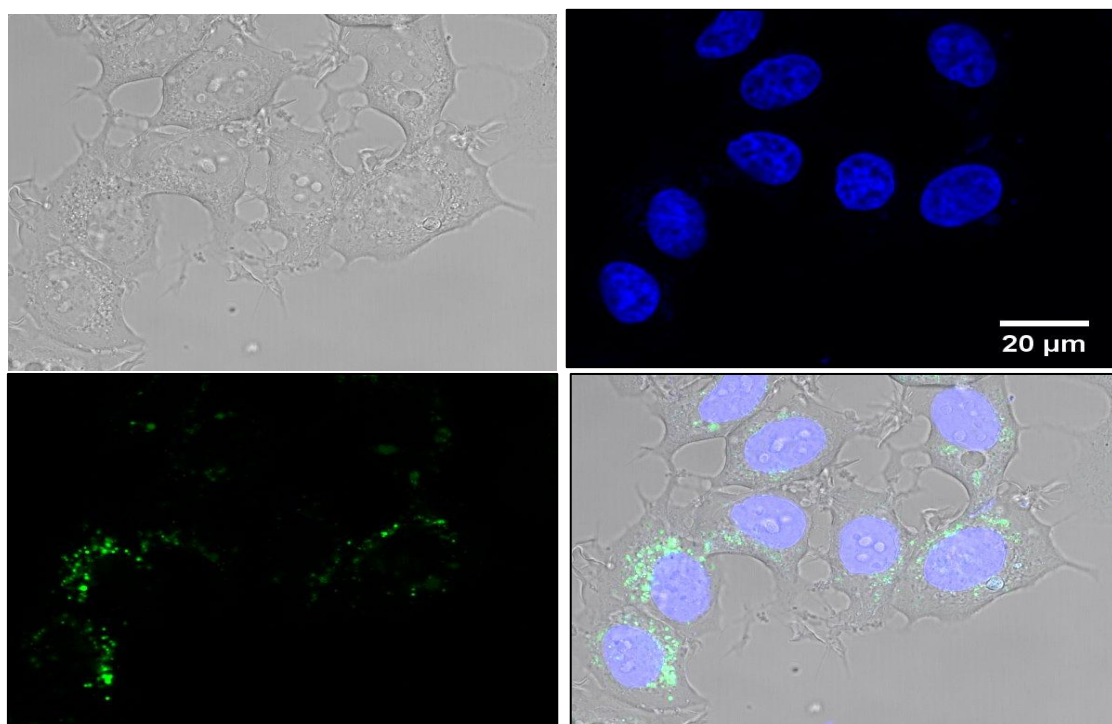

FIGURE 4.9 Confocal microscopy images for 1-guanidino modified *egd*-t$^{\gamma}_{2}$ PNA 6-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
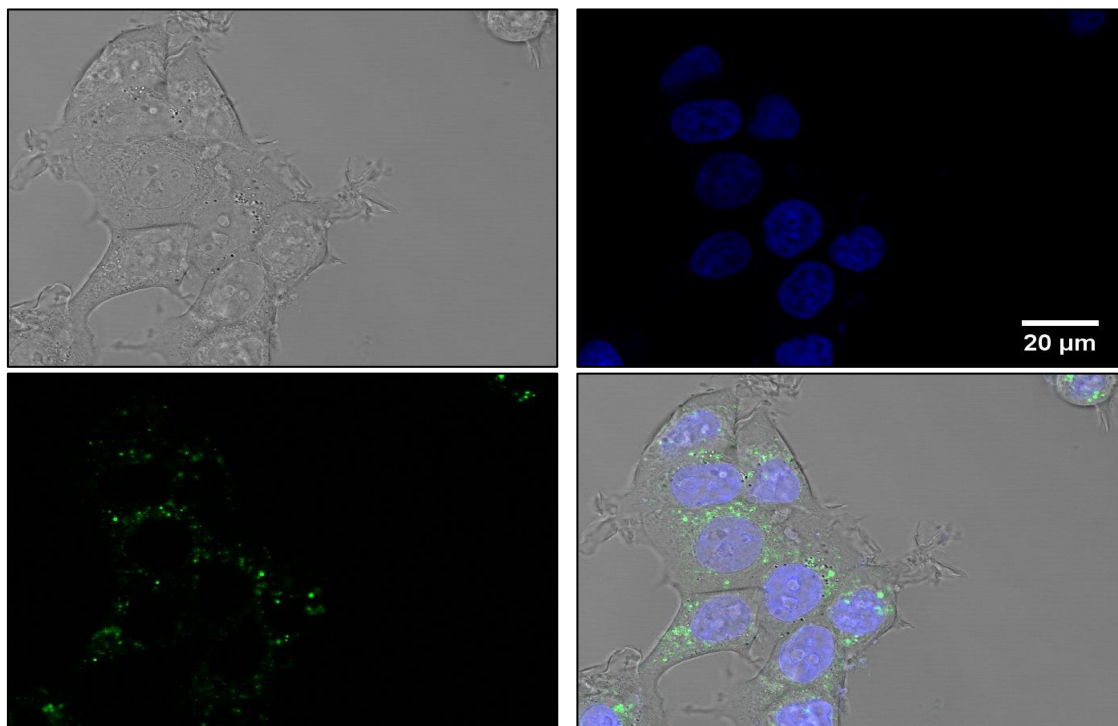
FIGURE 4.10 Confocal microscopy images for 2-guanidino modified *egd*-t$^{\gamma}_{2,6}$ PNA 9-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

FIGURE 4.11 Confocal microscopy images for 3-guanidino modified *egd*-t$^{\gamma}_{2,6,10}$ PNA 10-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
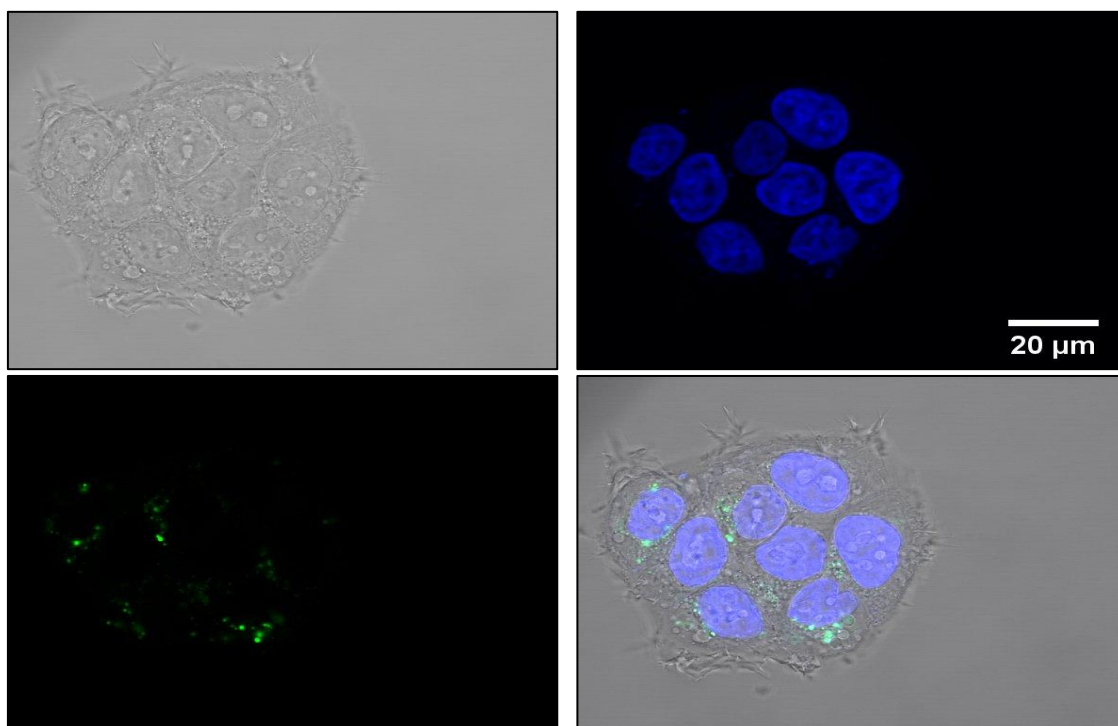
222

FIGURE 4.12 Confocal microscopy images for longer chain azide modified *azb*-T$_{10}$ PNA 17-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

FIGURE 4.13 Confocal microscopy images for shorter chain modified *azm*-T$^{\gamma}_{10}$ PNA 13-CF a) Bright field image of MCF-7 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

B) Cell Permeation in NIH 3T3 cells

Cell permeability of PNA oligomers was also investigated in normal mouse embryonic fibroblast cells to evaluate the ability of PNA oligomers getting inside the normal versus cancerous cell lines. The similar protocol mentioned in the section **4.5.2** was followed to study the cell permeability of various PNA oligomers in NIH 3T3 cells. The control untreated cells showed the blue staining of nucleus due to nuclear stain Hoechst 33342 and no signals for PNA internalization which gives green fluorescence (**Figure 4.14**).
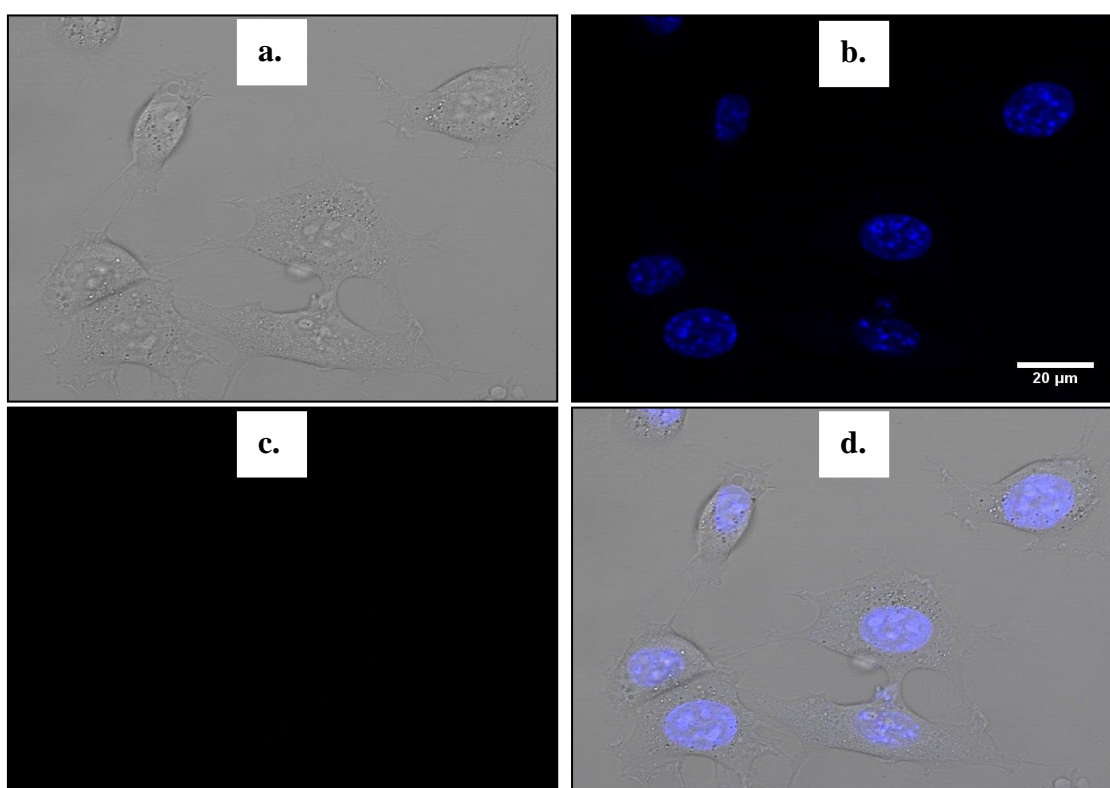

FIGURE 4.14 Confocal microscopy images for control untreated cells a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

A more convincing evidence of the PNA uptake by cells comes from the superimposition of bright field image, fluorescent image and the Hoechst 33342 stained image. The cellular uptake of unmodified *aeg* PNA 1 and various amino, guanidino and azido modified PNA oligomers have been shown below (**Figures 4.15-4.22**). All images are organized in a similar way as shown in Figure **4.14 a**, **b**, **c** and **d**.
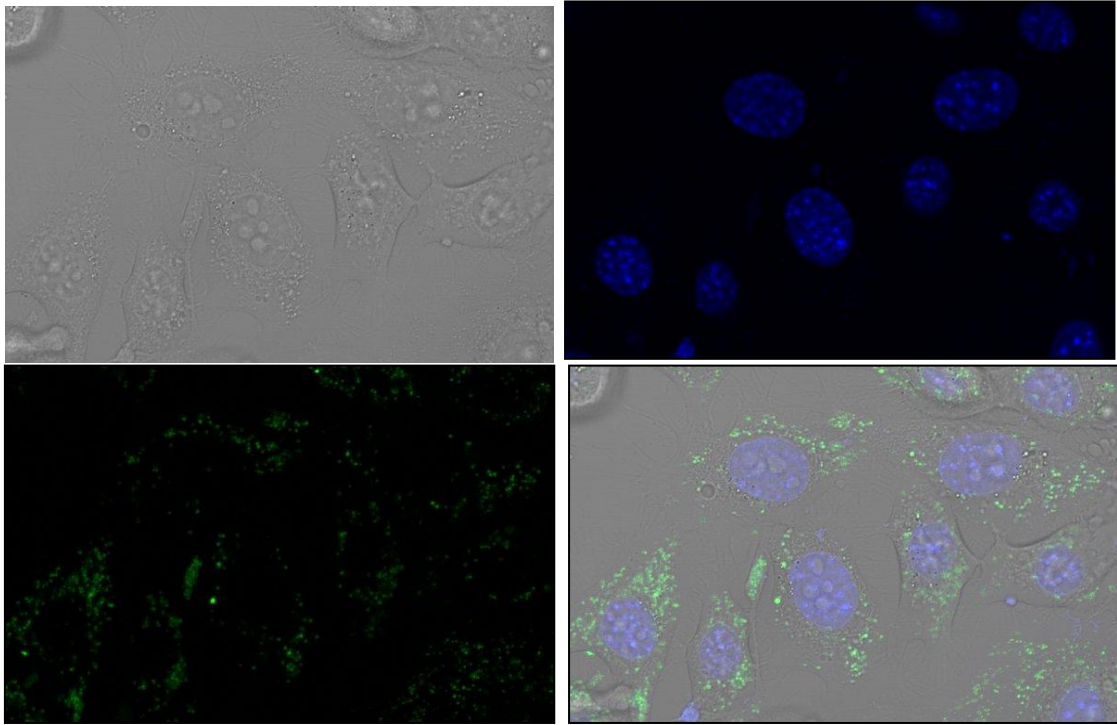
FIGURE 4.15 Confocal microscopy images for unmodified *aeg* PNA 1-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
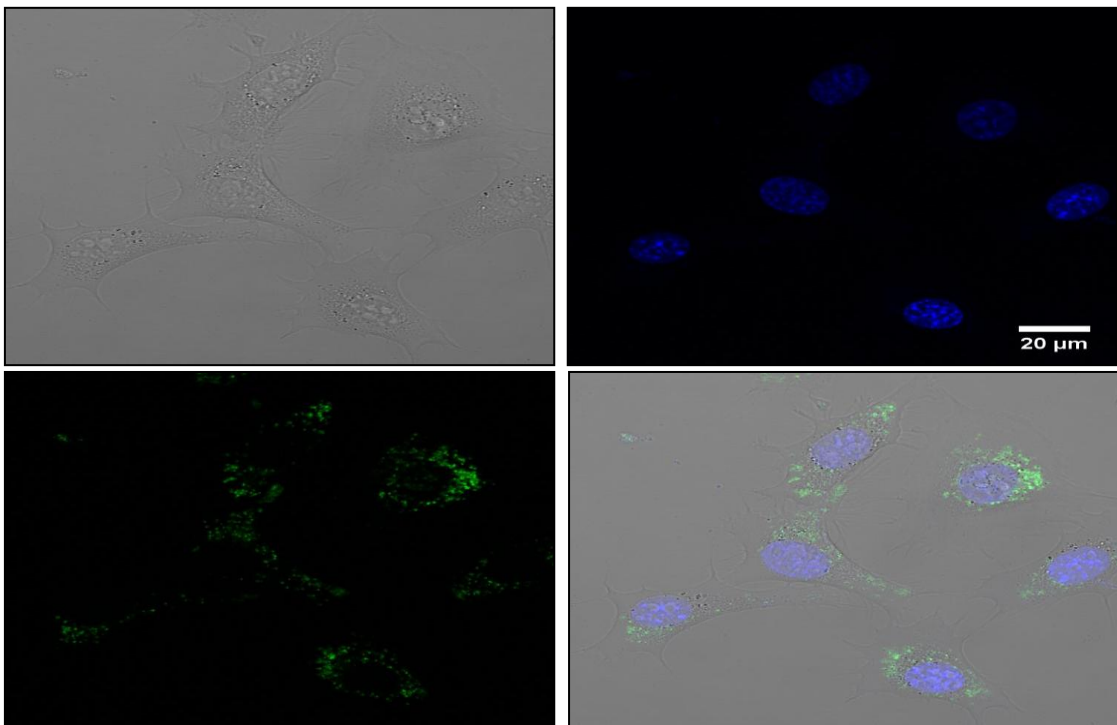

FIGURE 4.16 Confocal microscopy images for 1-amino modified *eam*-t$_2$ PNA 2-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
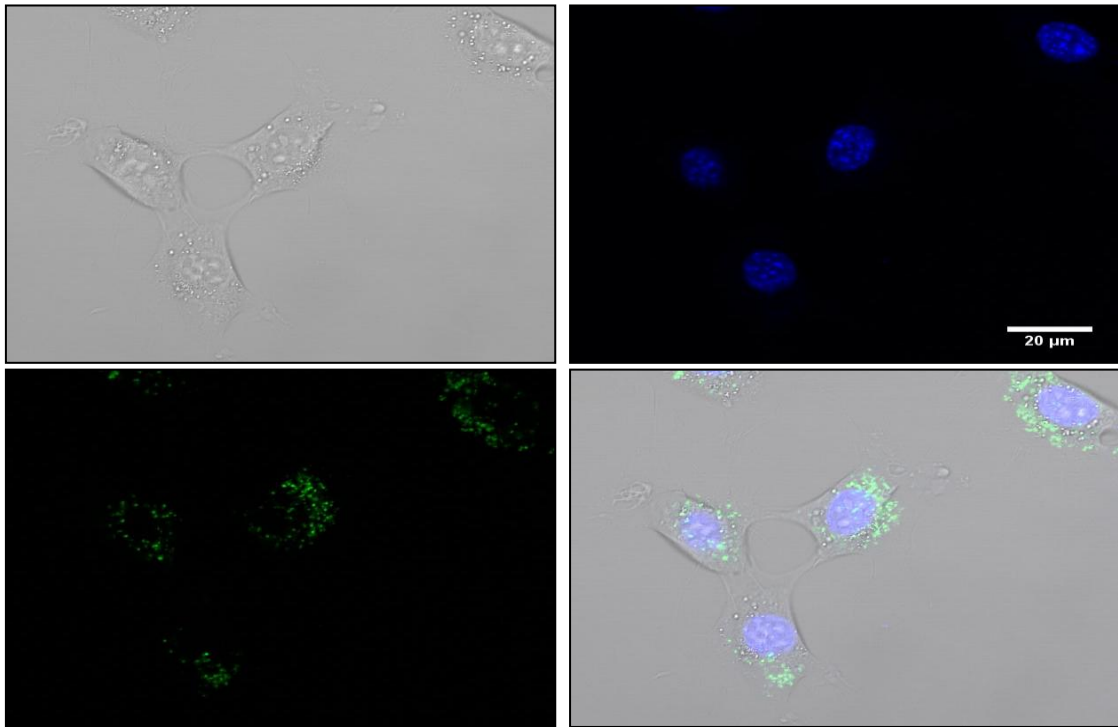
225

FIGURE 4.17 Confocal microscopy images for 2-amino modified *eam*-t$_{2,6}$ PNA 5-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

FIGURE 4.18 Confocal microscopy images for 1-guanidino modified *egd*-t$^{\gamma}_2$ PNA 6-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
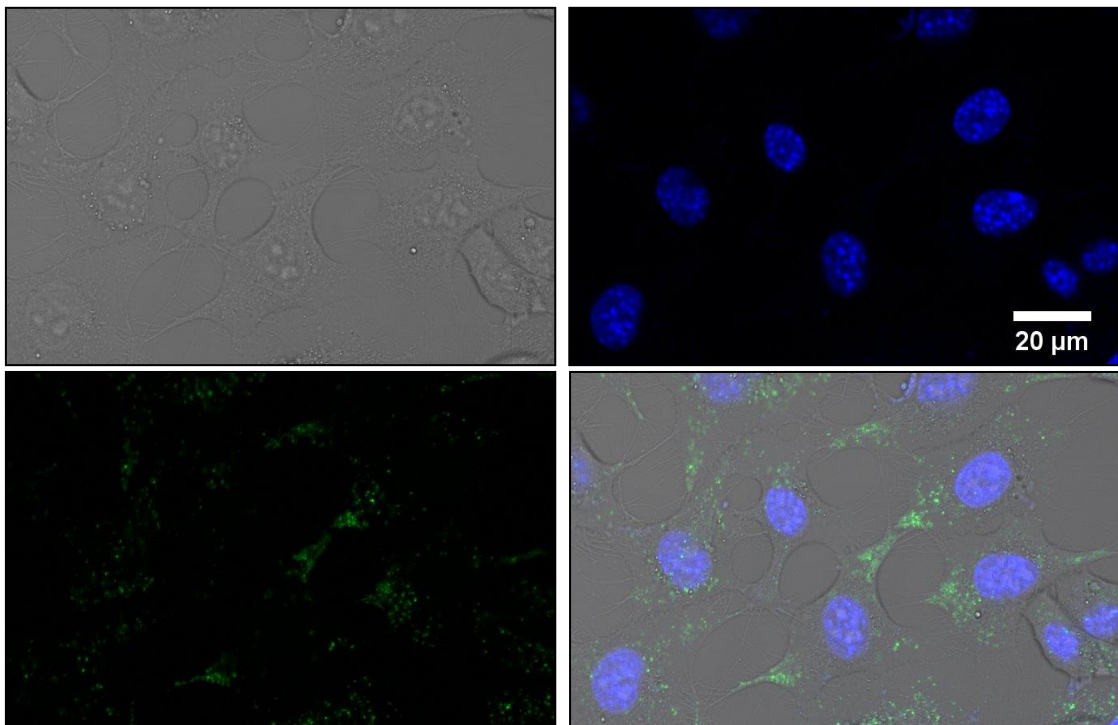
FIGURE 4.19 Confocal microscopy images for 2-guanidino modified _egd_-t$^\gamma_{2,6}$ PNA 9-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
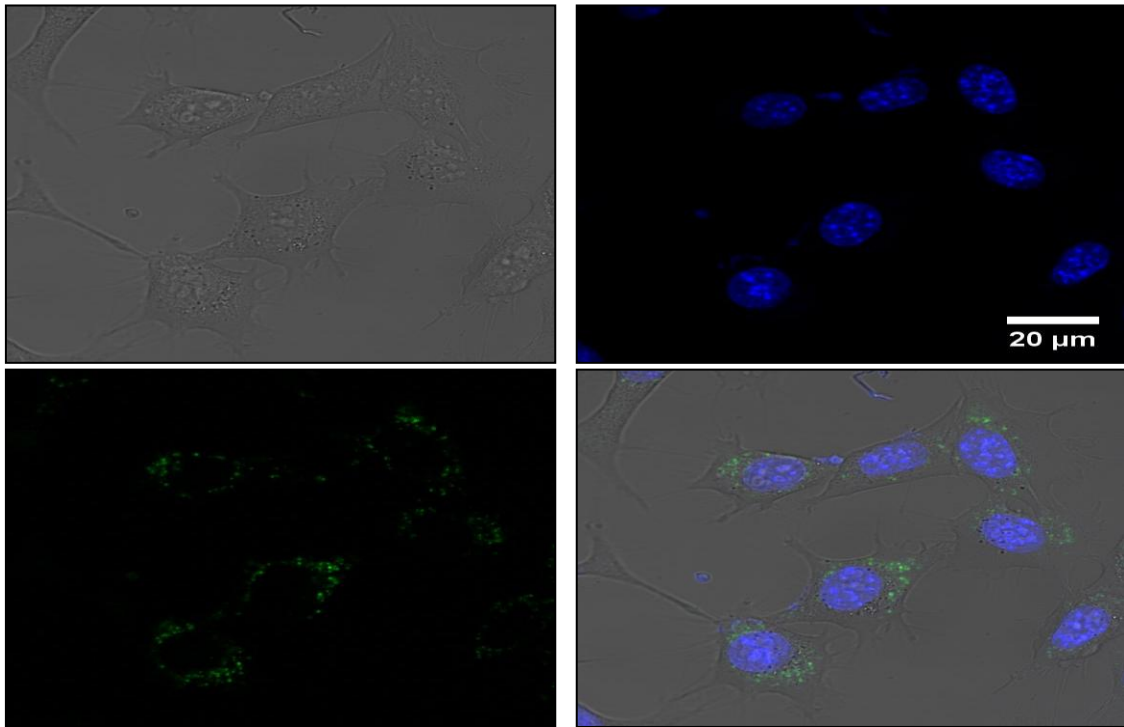

FIGURE 4.20 Confocal microscopy images for 3-guanidino modified _egd_-t$^\gamma_{2,6,10}$ PNA 10-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)

FIGURE 4.21 Confocal microscopy images for azidobutylene modified *azb*-T$_{10}$ PNA 17-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
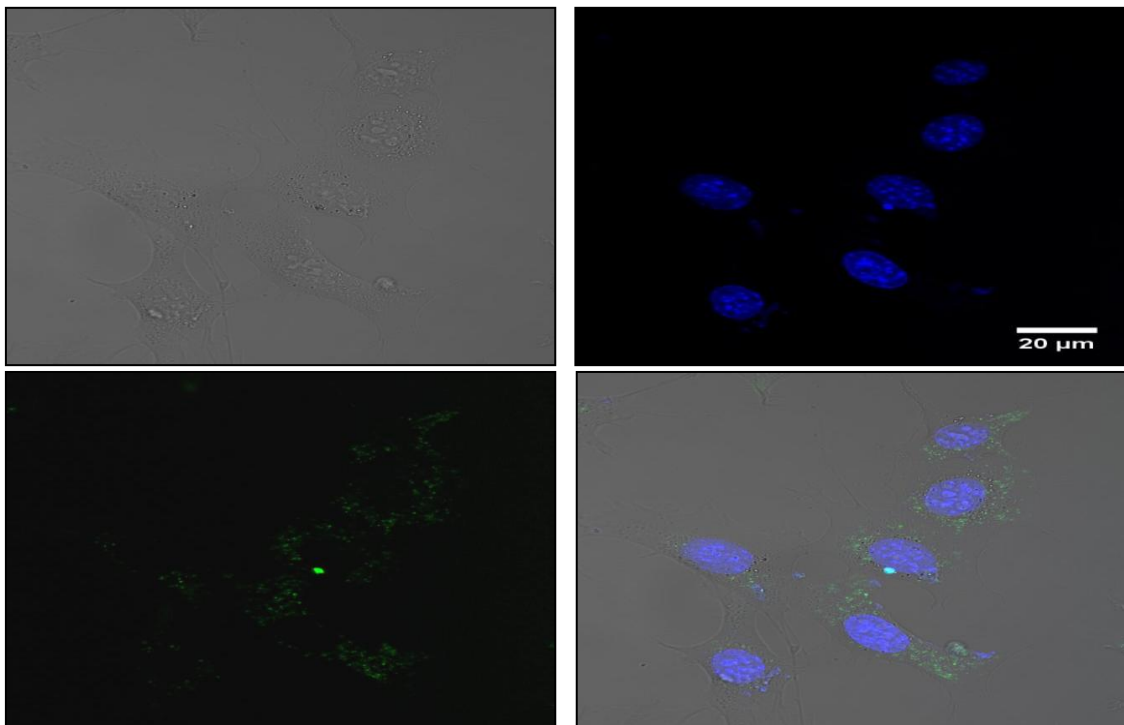

FIGURE 4.22 Confocal microscopy images for azidomethylene modified *azm*-T$^{\gamma}_{10}$ PNA 13-CF a) Bright field image of NIH 3T3 cells b) Hoechst 33342 stained image c) Green fluorescent image and d) Superimposed image of images a), b) and c)
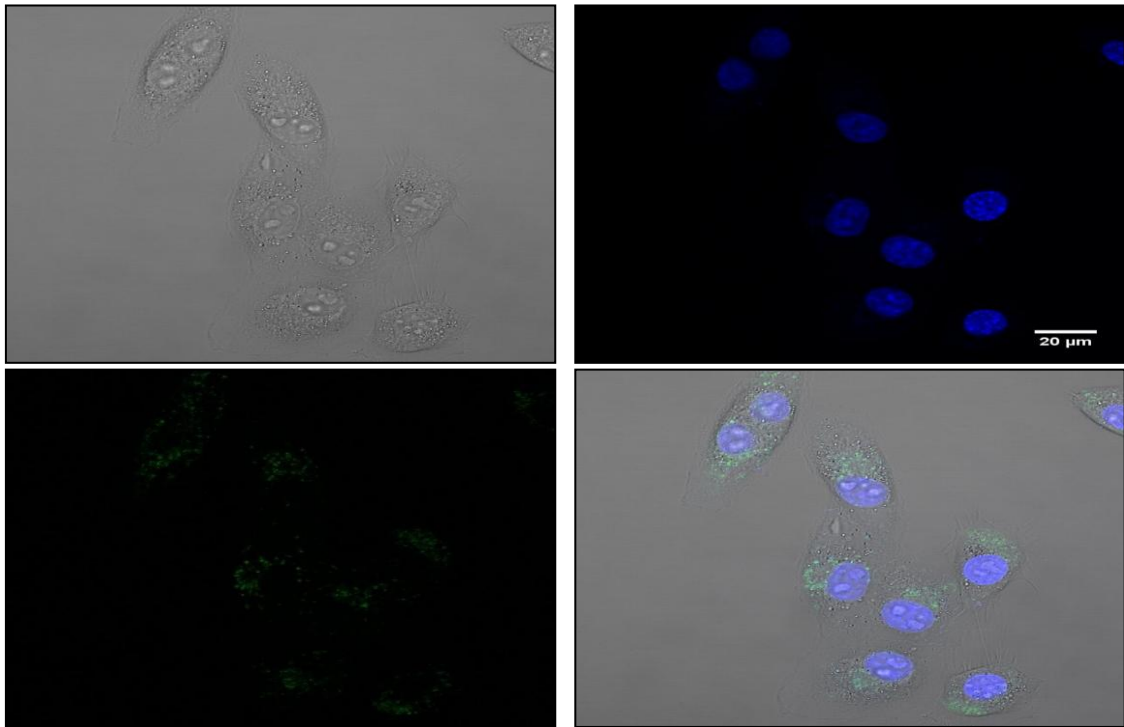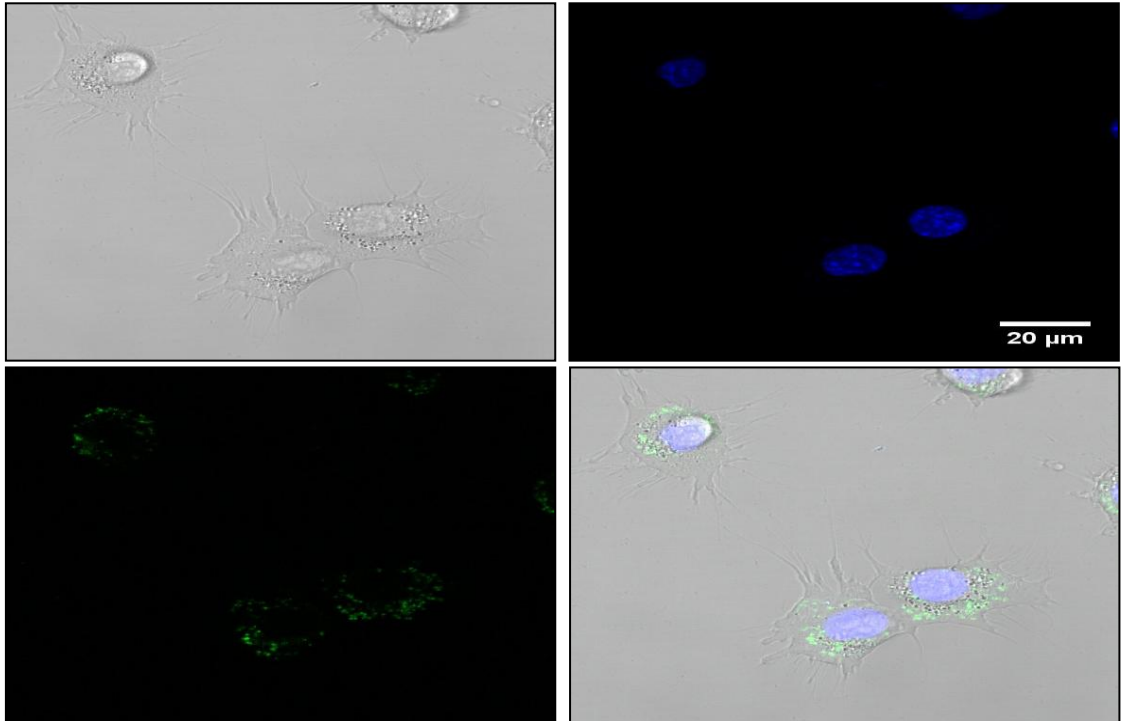
228

4.6 Quantitative estimation of cellular uptake

The confocal microscopy images showed that all PNA oligomers including modified and unmodified PNAs can penetrate the cell membrane and give fluorescent spots in the cytoplasm. It is important to quantify the cellular uptake of the different modified PNA oligomers to differentiate their abilities of cell permeation. Quantification of cell permeation has been carried out using a Fluorescence Activated Cell Sorter (FACS) technique. FACS is a specialized type of Flow cytometry which provides a method for sorting homogeneous as well as heterogeneous mixtures of cells based on the specific light scattering and fluorescent characteristics of each cell. FACS gives the quantitative recording of fluorescent signals from each individual cell as well as it provides important information about the positive and negative cells in terms of their fluorescent signal. The cells which give fluorescent signal are the positive cells whereas negative cells do not give any fluorescent signal. Thus, FACS data would provide quantitative estimation of cell permeability of all PNA oligomers which can be differentiated from their percent positive cells and mean fluorescence intensity. Quantitative estimation of PNA oligomers has been carried out for both MCF-7 and NIH 3T3 cells.

4.6.1 Quantification of cell permeability using FACS experiment

MCF-7 and NIH 3T3 cells were plated in 9 X 60 mm dishes in 2 mL DMEM medium at a concentration of 1 x 10^6 cells per dish. The cells were maintained at 37 °C in a humidified atmosphere containing 5 % CO_2 for 16 to 18 h. The required amounts of 5(6)-carboxyfluorescein tagged PNA stock solutions were added to the corresponding wells to achieve the desired final concentration of 1 μM. The cells incubated with tagged PNA oligomers were maintained at 37 °C in a humidified atmosphere containing 5 % CO_2 for 24 h. After the incubation period was over, the medium was aspirated and cells were rinsed or washed thrice with 1 medium-volume equivalent of ice-cold PBS. Cells were collected by trypsinization using 0.5 mL of 0.05 % trypsin-EDTA for each dish. The trypsinization processing was stopped using 3 mL DMEM medium and the cells were transferred to 15 mL tubes. These tubes containing cells were centrifuged at 500 g for 5 min at 4 °C, the supernatant was aspirated and the cells were washed thrice with ice-cold 1X PBS. The cell suspension in PBS was then

transferred to FACS tube and the samples were immediately analysed on BD Biosciences FACS Calibur flow cytometer. The data obtained from FACS was processed using CellQuest Pro acquisition software. The cell permeation of PNA oligomers has been confirmed by histograms obtained from the experiment. The percent positive cells for each individual PNA were plotted on Microcal origin 8.

4.6.2 Results and discussion

4.6.2a Quantitative estimation of cell permeation in NIH 3T3 cells

The FACS analysis data for the quantification of cell permeation of PNA oligomers in NIH 3T3 cells are shown in Figure **4.23**.
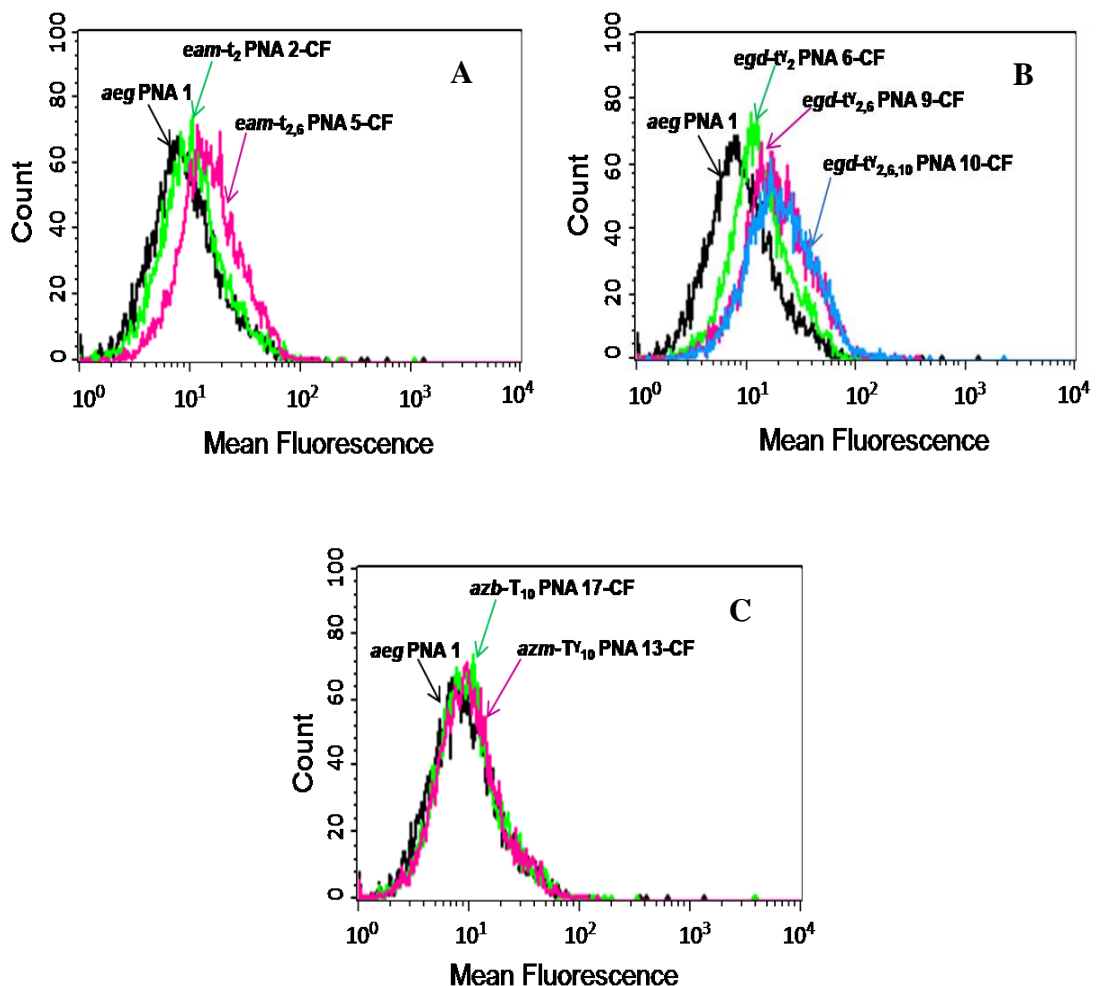

FIGURE 4.23 FACS analysis showing mean fluorescence for quantification of PNA cell permeation in NIH 3T3 cells

In case of amino modified PNAs, two amino modified PNA (*eam*-t$_{2,6}$ PNA **5**-CF) units incorporated into a PNA sequence showed slightly increased mean fluorescence compared to unmodified *aeg* PNA (PNA **1**-CF), while one amino modified PNA (*eam*-t$_2$ PNA **2**-CF) unit gave similar mean fluorescence like control *aeg* PNA (Figure **4.23 A**). Incorporation of guanidino modified PNA units into a PNA sequence increased the mean fluorescence intensity compared to control *aeg* PNA. Two and three guanidino modified PNA oligomers (*egd*-t$^\gamma_{2,6}$ PNA **9**-CF and *egd*-t$^\gamma_{2,6,10}$ PNA **10**-CF) showed almost similar mean fluorescence intensities but slightly increased than that of one guanidino modified PNA (*egd*-t$^\gamma_2$ PNA **6**-CF) (Figure **4.23 B**). Azide modified *azm*-T$^\gamma_{10}$ PNA **13**-CF (shorter chain) and *azb*-T$_{10}$ PNA **17**-CF (longer chain) oligomers showed almost similar mean fluorescence intensities like control *aeg* PNA oligomer (Figure **4.23 C**). Thus, mean fluorescence intensities were found to be minimally changed for control *aeg* PNA and modified PNA oligomers.

FACS experiment gives the sorting of individual cells depending on the fluorescence signal observed from the fluorescently tagged PNA. When an individual cell shows green fluorescence signal of PNA, it is called 'positive cell', while 'negative cell' does not give fluorescence signal for PNA. Therefore, % positive cell count was used to quantify the cellular uptake of different modified PNAs (Figure **4.24**). The count of % positive cells actually gives the number of cells that have taken up the PNA.

Control *aeg* PNA (PNA **1**-CF) can penetrate into ~ 18 % of cells, whereas incorporation of one amino modified unit (*eam*-t$_2$ PNA **2**-CF) into a sequence increase the % positive cells to ~ 30 %. Incorporation of two amino modified PNA units (*eam*-t$_{2,6}$ PNA **5**-CF) into a sequence resulted in penetration of 50 % of cells with 3-fold increase in the cell permeation compared to control *aeg* PNA. Incorporation of one guanidino modified unit (*egd*-t$^\gamma_2$ PNA **6**-CF) in the PNA sequence gave ~ 42 % positive cells, while two guanidino modified units (*egd*-t$^\gamma_{2,6}$ PNA **9**-CF) in the sequence increase the % positive cells to ~ 62 %. Incorporation of three guanidine modified units (*egd*-t$^\gamma_{2,6,10}$ PNA **10**-CF) in the PNA sequence gives almost 4-fold increase in the cell penetration giving ~ 70 % of cells positive compared to control *aeg* PNA. Azide modified PNA oligomers (*azm*-T$^\gamma_{10}$ PNA **13**-CF and *azb*-T$_{10}$ PNA **17**-CF) can penetrate into ~ 25 % of cells. Thus, the efficiency of cell penetration was found to be increased when cationic functional groups were incorporated into PNA backbone.
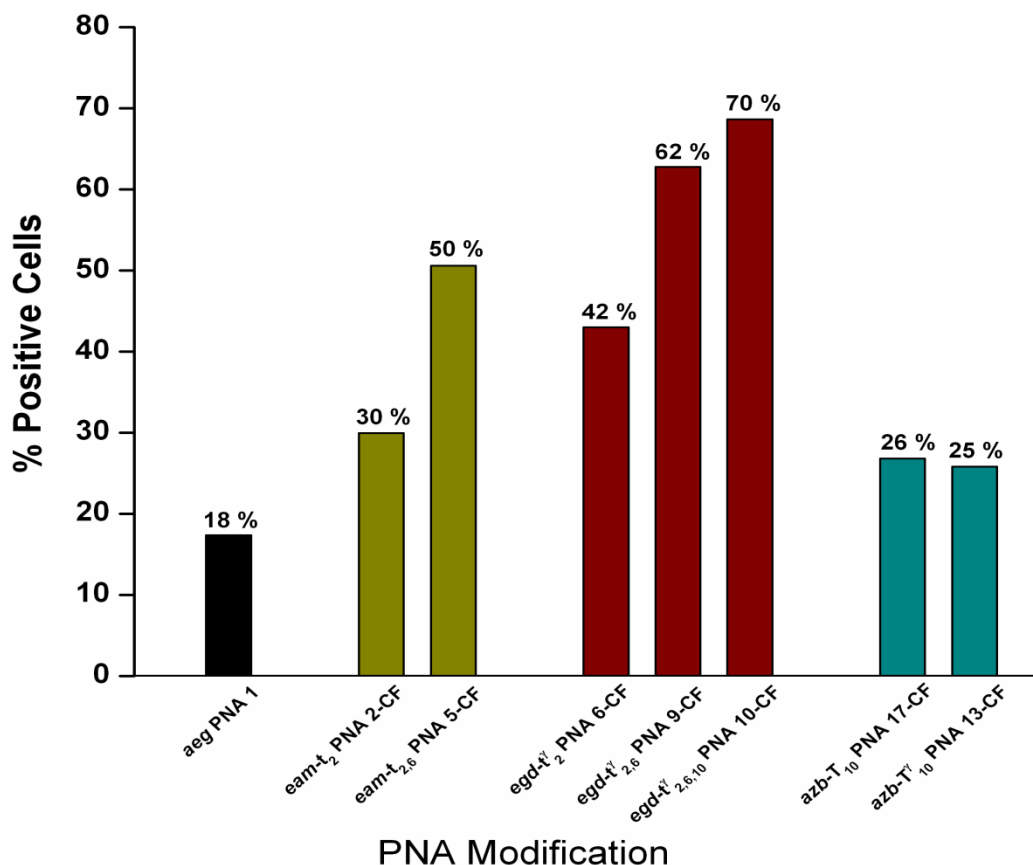
FIGURE 4.24 FACS analysis showing the % positive cells for quantification of cell permeation in NIH 3T3 cells for various PNA oligomers (represented % positive cell values are mean of three independent results, standard deviation: $\pm$ 6 %)

4.6.2b Quantitative estimation of cell permeation in MCF-7 cells

The quantification of cell penetration in MCF-7 cells for control *aeg* and modified PNA oligomers is shown in Figure **4.25**.

The mean fluorescence intensities were found to be similar for unmodified *aeg* PNA (PNA **1**-CF) and modified PNA oligomers. The shift observed in mean fluorescence intensity was very less after increasing the number of amino as well as guanidino modified PNA units in the PNA sequence (Figure **4.25 A**, **B**). Azide modified (through longer and shorter carbon chain) PNA oligomers showed mean fluorescence intensities similar to control *aeg* PNA (Figure **4.25 C**).
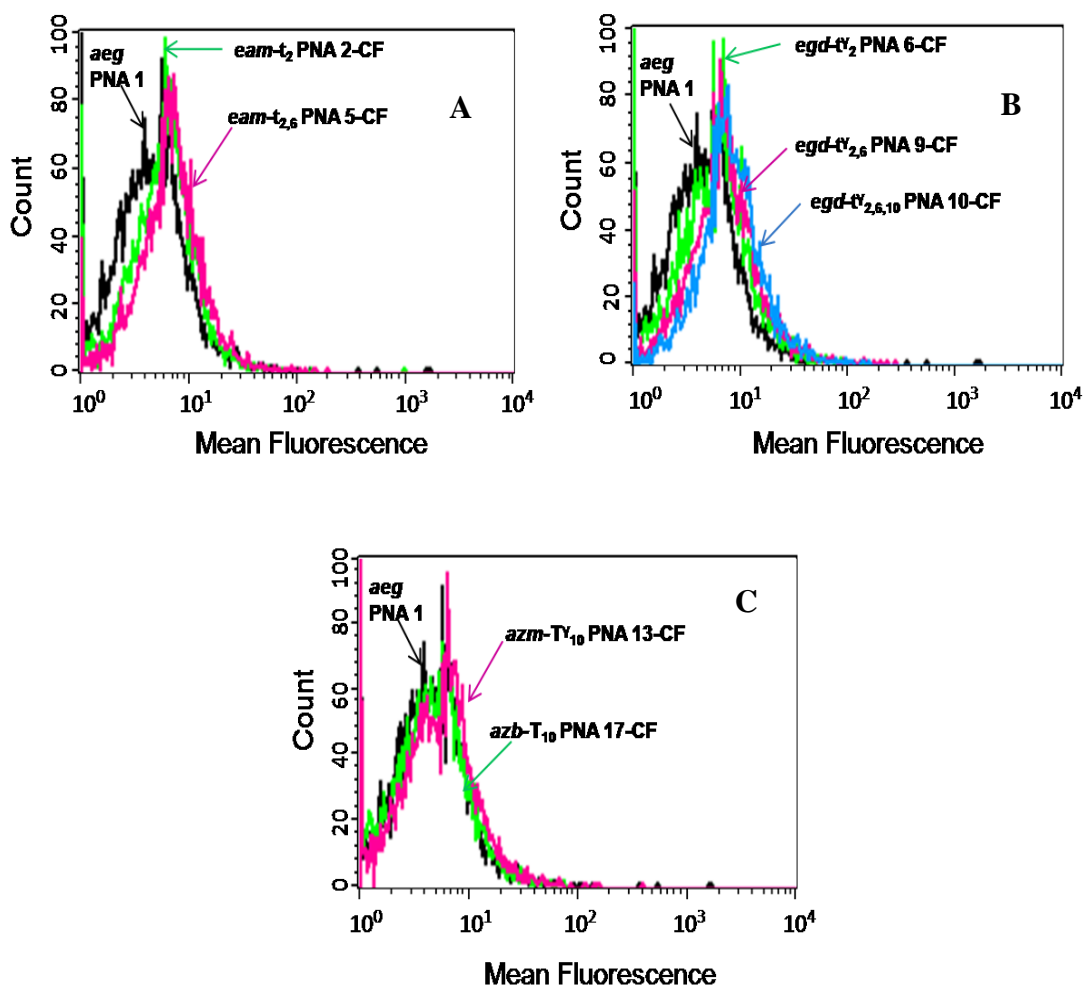
FIGURE 4.25 FACS analysis showing mean fluorescence for quantification of PNA cell permeation in MCF-7 cells

The FACS analysis data of % positive cells representing the number of cells that have taken up the unmodified and modified PNA oligomers is shown in Figure **4.26**. The unmodified *aeg* PNA (PNA **1**-CF) can penetrate ~ 11 % of MCF-7 cells; whereas one amino modified PNA unit (*eam*-t_2 PNA **2**-CF) in the sequence gives ~ 15 % positive cells. Incorporation of two amino modified PNA units (*eam*-$t_{2,6}$ PNA **5**-CF) into a sequence increased the % positive cells to ~ 22 % with 2-fold increase in the cell penetration compared to control *aeg* PNA. When one guanidino modified unit (*egd*-t^{γ}_2 PNA **6**-CF) was incorporated in the PNA sequence it gave ~ 19 % of cells positive, while incorporation of two guanidino modified units (*egd*-$t^{\gamma}_{2,6}$ PNA **9**-CF) in the sequence gives ~ 25 % positive cells. % positive cells increased to ~ 31 % (3-fold compared to control *aeg* PNA) when three guanidino modified PNA units (*egd*-$t^{\gamma}_{2,6,10}$ PNA **10**-CF) were incorporated into a sequence. The azidomethylene (shorter carbon

233

chain) modified PNA (*azm*-T$^\gamma_{10}$ PNA **13**-CF) can penetrate into ~ 17 % of cells, while azidobutylene (longer carbon chain) modified PNA (*azb*-T$_{10}$ PNA **17**-CF) can penetrate ~ 12 % of cells which is similar to control *aeg* PNA. The quantification of cell permeation in MCF-7 cells also showed that cationic functional groups help to increase the efficiency of cell penetration.
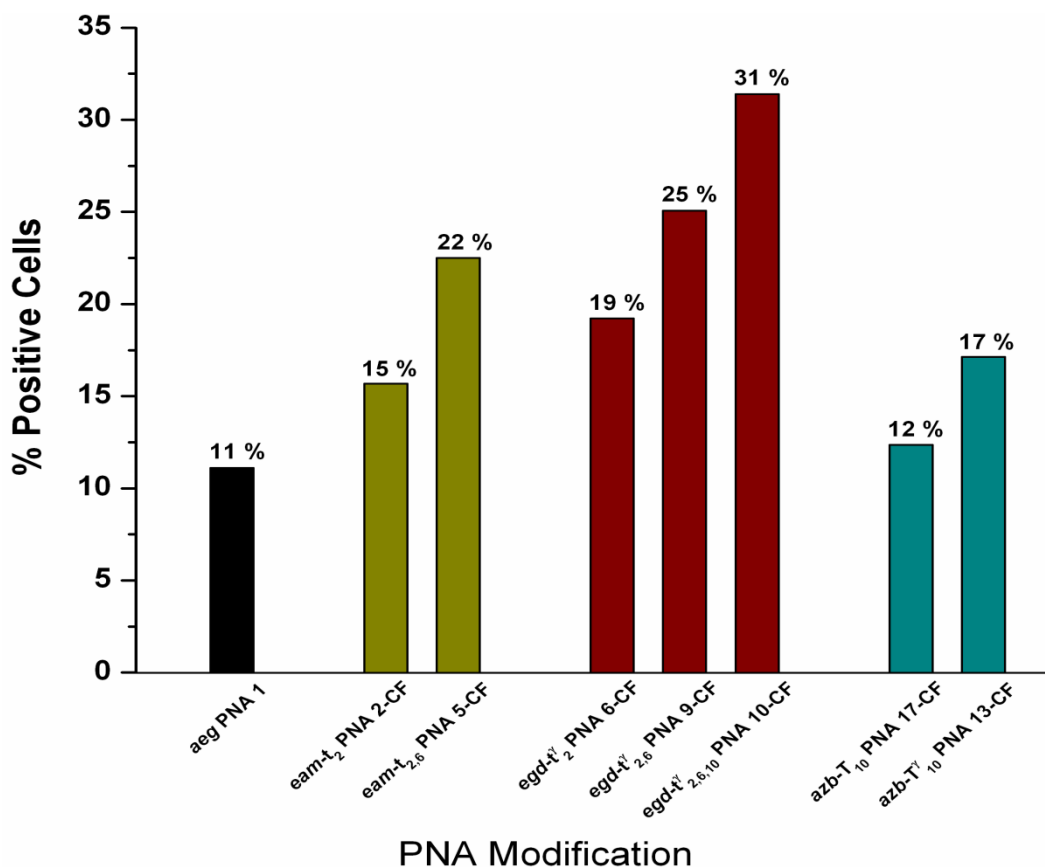

FIGURE 4.26 FACS analysis showing the % positive cells for quantification of cell permeation in MCF-7 cells for various PNA oligomers (represented % positive cell values are mean of three independent results, standard deviation: ± 5 %)

To summarize, cationic amino/guanidino and charge-neutral azido modified and control *aeg* PNA oligomers have been tagged with 5(6)-carboxyfluorescein to visualize the PNA inside the cells. The fluorescently tagged PNA oligomers were studied for cell penetration using confocal microscopy. The live cell imaging for cell permeation was done for all unmodified and modified PNA oligomers into NIH 3T3 and MCF-7 cell lines. The quantification for differential cellular uptake of all PNA oligomers was achieved by FACS analysis.

4.7 Conclusions

➢ The confocal microsopy images clearly indicate that unmodified as well as modified PNA oligomers can penetrate into both cell lines used for the study.

➢ The PNA oligomers were observed to be localized in the vicinity of nucleus in the cytoplasm.

➢ The quantification of cell permeation by FACS analysis showed that cationic functional groups (amine and guanidine) incorporated in the PNA backbone at γ-position remarkably increases the efficiency of cell penetration.

➢ The guanidino modified PNA was found to be better among all modifications to enhance the cell penetration in both cell lines (NIH 3T3 and MCF-7) used for this study.

➢ Differential cellular uptake was observed in two different cell lines where in NIH 3T3 cells, three guanidino modified PNA (*egd*-t$^{\gamma}_{2,6,10}$ PNA **10**-CF) gave 5-fold increase in the cell penetration while 3-fold increase in cell penetration was observed in MCF-7 cells for the same PNA.

➢ The amino modified PNA oligomers were also found to be better than control *aeg* PNA in both cell lines but these are less efficient than guanidino modified PNAs.

➢ Incorporation of charge-neutral azide modified (through longer and shorter carbon chain) unit in the PNA sequence showed only marginal increase in the cell penetration compared to control *aeg* PNA.

4.8 Experimental procedures

4.8.1 Tagging of PNA oligomers with 5(6)-carboxyfluorescein

5(6)-carboxyfluorescein, a fluorescent dye was attached to the N-terminal end of PNA oligomers. Using the solid phase peptide synthesis protocol, *Boc* group of final N-terminal amine was deprotected by 50 % TFA in DCM. The TFA salt generated after Boc-deprotection was neutralized using DIPEA as base to get free amine. The amine was coupled with 5(6)-carboxyfluorescein (13.16 mg, 10 eq) using HOBt (4.72 mg, 10 eq) and DIC (5.51 μL, 10 eq) in DMF as coupling reagents. The coupling reaction was completed within 12 h which was confirmed by Kaiser Test.

4.8.2 Cleavage of the PNA oligomers from the solid support

The resin beads were washed with methanol several times to remove excess 5(6)-carboxyfluorescein. The peptides were cleaved from the solid support using trifluoromethane sulfonic acid (TFMSA) in the presence of trifluoroacetic acid (TFA); (Low, High TFMSA-TFA method which is discussed in chapter 2) (section **2.2.5**). Thioanisole and ethanedithiol were used as scavengers in the cleavage protocol. This yielded PNA oligomers having L-lysine amide at C-terminus and carboxyfluorescein attached at N-terminus. After cleavage reaction was over, the peptide was precipitated with cold dry ether. The peptide was isolated by centrifugation and the precipitate was dissolved in de-ionized water.

4.8.3 Purification of the fluorescently tagged PNA oligomers

PNA purification was carried out on Dionex ICS 3000 HPLC system. For the purification of peptides, semi-preparative BEH130 C18 (10X250 mm) column was used. Purification of PNA oligomers was performed with gradient elution method: A to 100% B in 20 min; A= 0.1% TFA in $CH_3CN:H_2O$ (5:95); B= 0.1% TFA in $CH_3CN:H_2O$ (1:1) with flow rate of 3 mL/min. All the HPLC profiles were monitored at 254 and 490 nm wavelength.
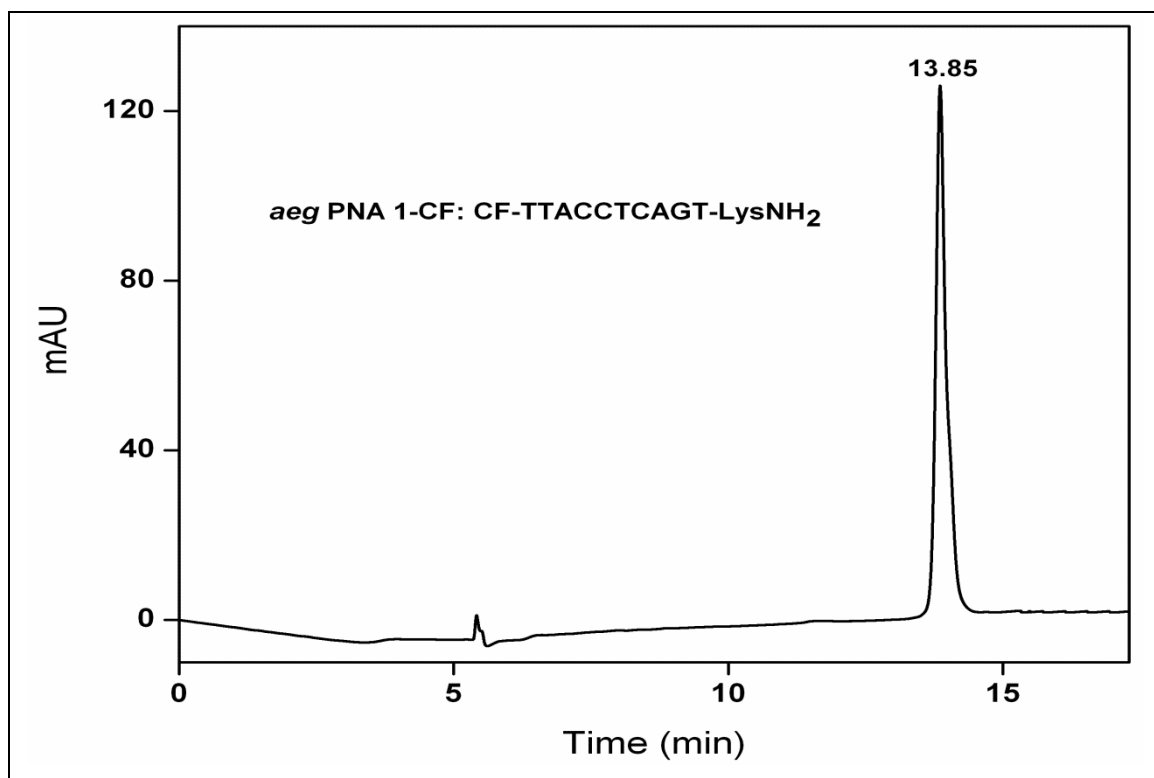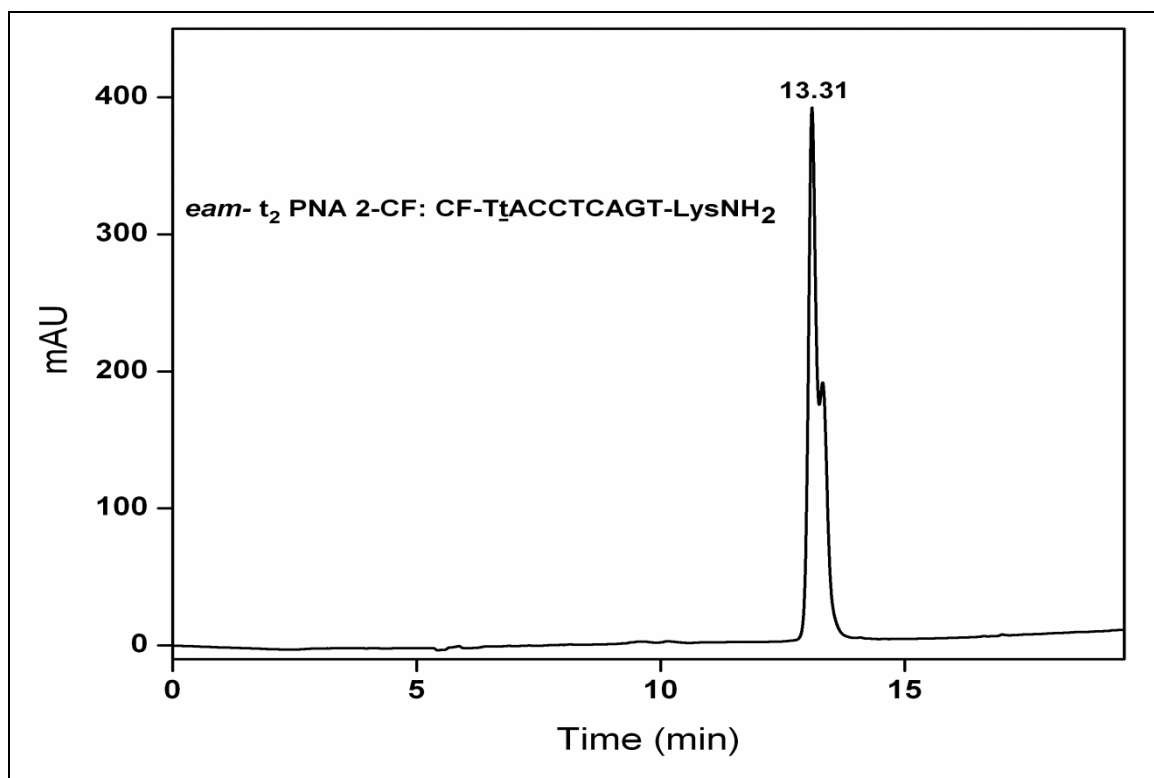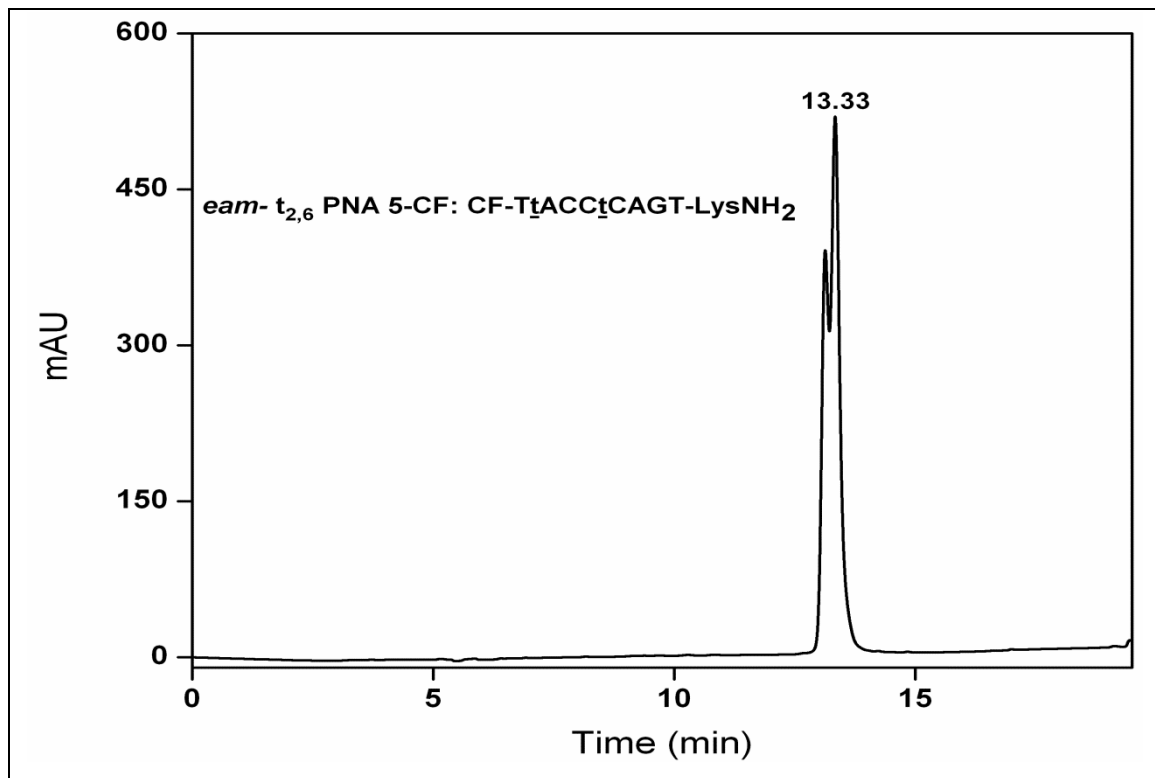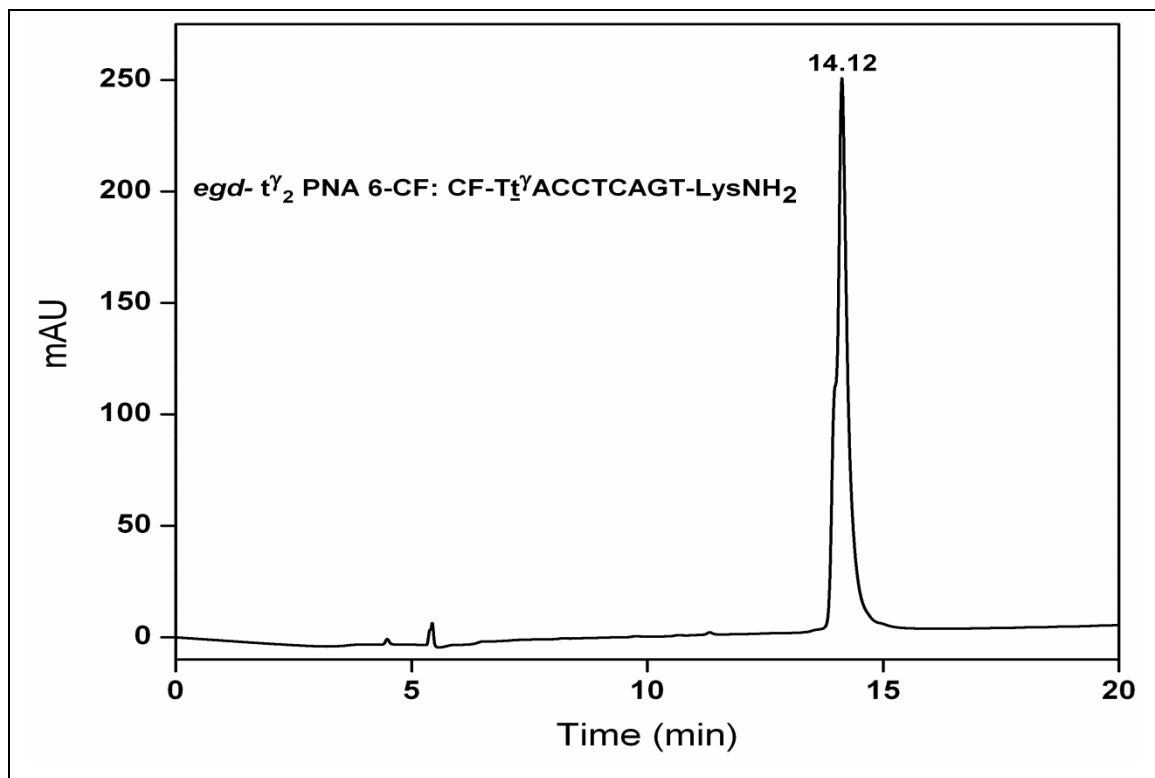
4.9 References

1. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 1, 280-284.

2. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 1, 285-288.

3. Mesmaeker, A. D.; Haner, R.; Martin, P.; Moser, h. E. *Acc. Chem. Res.* **1995**, *28*, 366-374.

4. (a) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543-584. (b) Wengel, J. *Acc. Chem. Res.* **1999**, *32*, 301–310. (c) Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624-630.

5. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543-584.

6. Bayard, B.; Leserman, L. D.; Bisbal, C.; Lebleu, B. *Eur. J. Biochem.* **1985**, *151*, 319-325.

7. Bisbal, C.; Silhol, M.; Lemaitre, M.; Bayard, B.; Salehzada, T.; lebleu, B. *Biochemistry* **1987**, *26*, 5172-5178.

8. (a) Juliano, R. L.; Yoon, H. *Curr. Opin. Mol. Therap.* **2000**, *2*, 297-303. (b) Lebedeva, I.; Benimetskaya, L.; Stein, C. A.; Vilenchik, M. *Eur. J. Pharm. Biopharm* **2000**, *50*, 101-119.

9. Stein, C. A. *Nat. Biotechnol.* **1999**, *17*, 209.

10. Demidov, V.; Potaman, V. N.; Franck-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sonnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1309-1313.

11. Bonham, M. A.; Brown, S.; Boyd, A. L.; Brown, P. H.; Bruckenstein, D. A.; Hanvey, J. C.; Thomson, S. A.; Pipe, A.; Hassman, F.; Bisi, J. E.; Froehler, B. C.; Matteucci, M. D.; Wagner, R. W.; Noble, S. A.; Babiss, L. E. *Nucleic Acids Res.* **1995**, *23*, 1197-1203.

12. Koppelhus, U.; Nielsen, P. E. *Adv. Drug Delivery Rev.* **2003**, *55*, 267-280.

13. Ljungstrøm, T.; Knudsen, H.; Nielsen, P. E. *Bioconjugate Chem.* **1999**, *10*, 965-972.

14. (a) Muratovska, A.; Lightowlers, R. N.; Taylor, R. W.; Turnbull, D. M.; Smith, R. A. J.; Wilce, J. A.; Martin, S. W.; Murphy, M. P. *Nucleic Acids Res.* **2001**, *29*, 6, 1852-1863. (b) Geall, A.J.; Taylor, R. J.; Earll, M. E.; Eaton, M. A. W.; Blagbrough, I. S. *Bioconjugate Chem.* **2000**, *11*, 314-326. (c) Shiraishi, T.; Nielsen, P. E.; *Bioconjugate Chem.* **2012**, *23*, 196-202. (d) Zhilina, Z. V.; Ziemba, A. J.; Ebbinghaus, S. W. *Curr. Top. Med. Chem.* **2005**, *5*, 1119-1131.

15. Fisher, A. A.; Sergueev, D.; Fisher, M.; shaw, B. R.; Juliano, R. L. *Pharmaceutical Res.* **2002**, *19*, 6, 744-754.

16. (a) Pooga, M.; Soomets, U.; Hallbrink, M.; Valkna, A.; Saar, K.; Rezaei, K.; Kahl, U.; Hao, j. X.; Xu, X. J.; Hallin, Z. W.; Hokfelt, T.; Bartfai, T.; Langel, U. *Nat. Biotechnol.* **1998**, *16*, 857-861. (b) Cutrona, G.; Carpaneto, E. M.; Ulivi, M.; roncella, S.; Landt, O.; Ferrarini, M.; Boffa, L. C. *Nat. Biotechnol.* **2000**, *18*, 300-303.

17. Basu, S.; Wickstrom, E. *Bioconjugate Chem.* **1997**, *8*, 481-488.
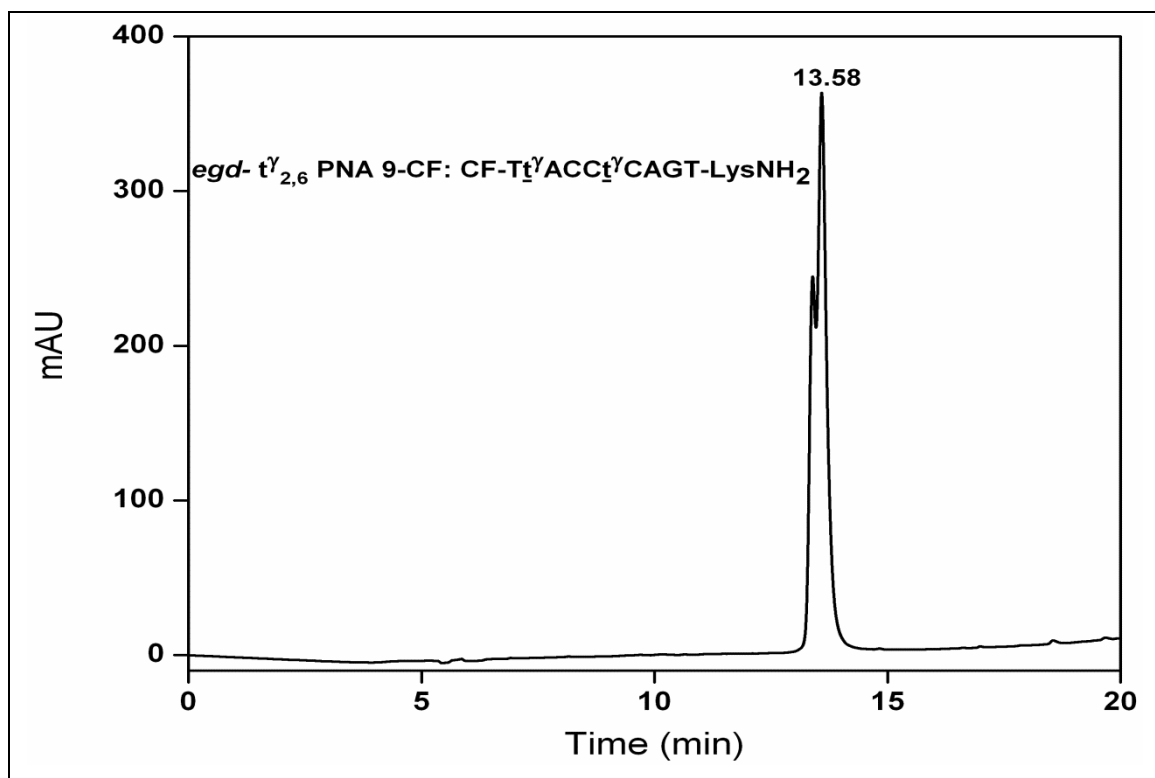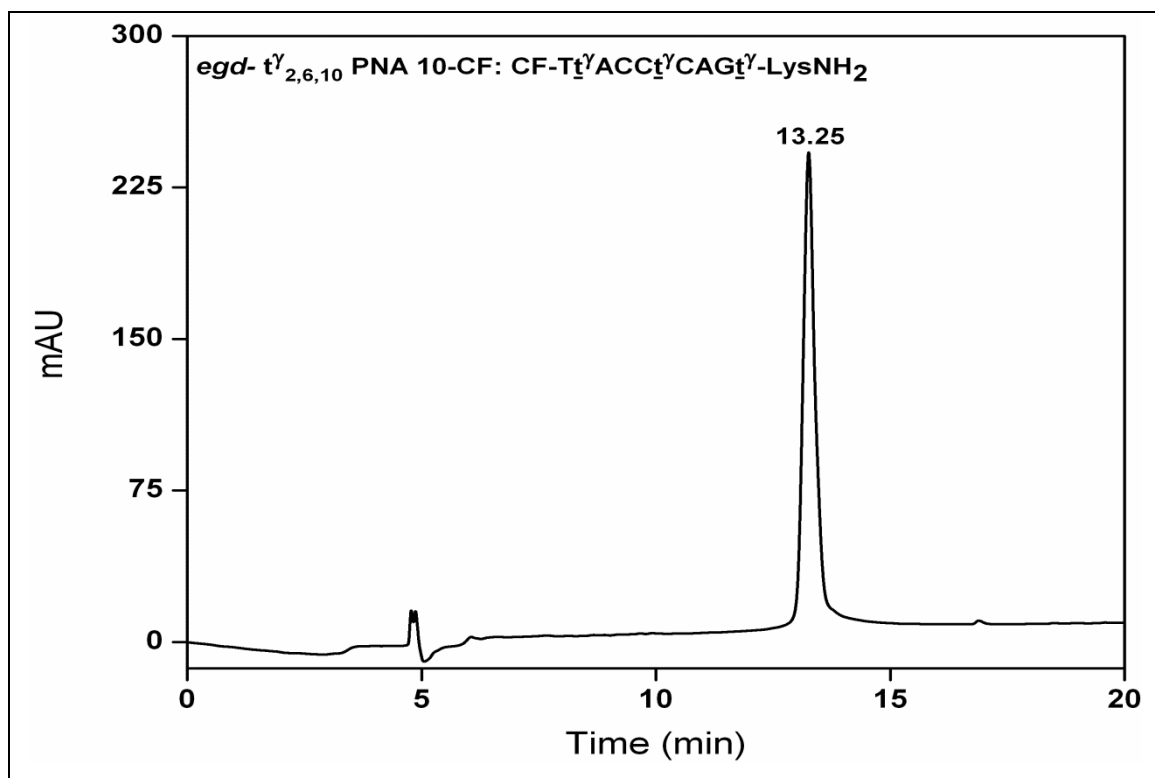
18. Zhang, X.; Simmons, C. G.; Corey, D. R. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1269-1272.

19. CPPsite: a website for cell penetrating peptides. crdd.osdd.net/raghava/cppsite/help.php

20. Morgan, D. M. L.; Larvin, V. L.; Pearson, J. D. *J. Cell Sci.* **1989**, *94*, 553-559.

21. Wender, P.A.; Galliher, W. C.; Goun, E. A.; Jones, L.R.; Pillow, T. H. *Adv. Drug Delivery Rev*. **2008**, *60*, 452-472.
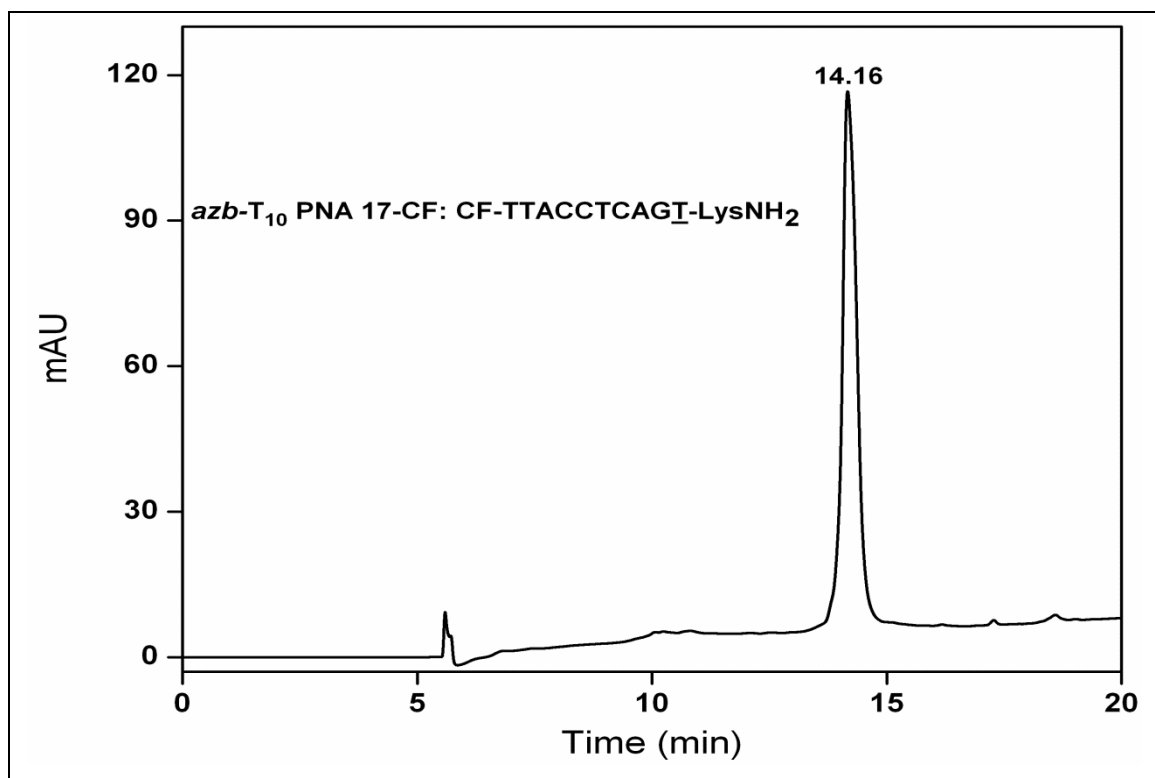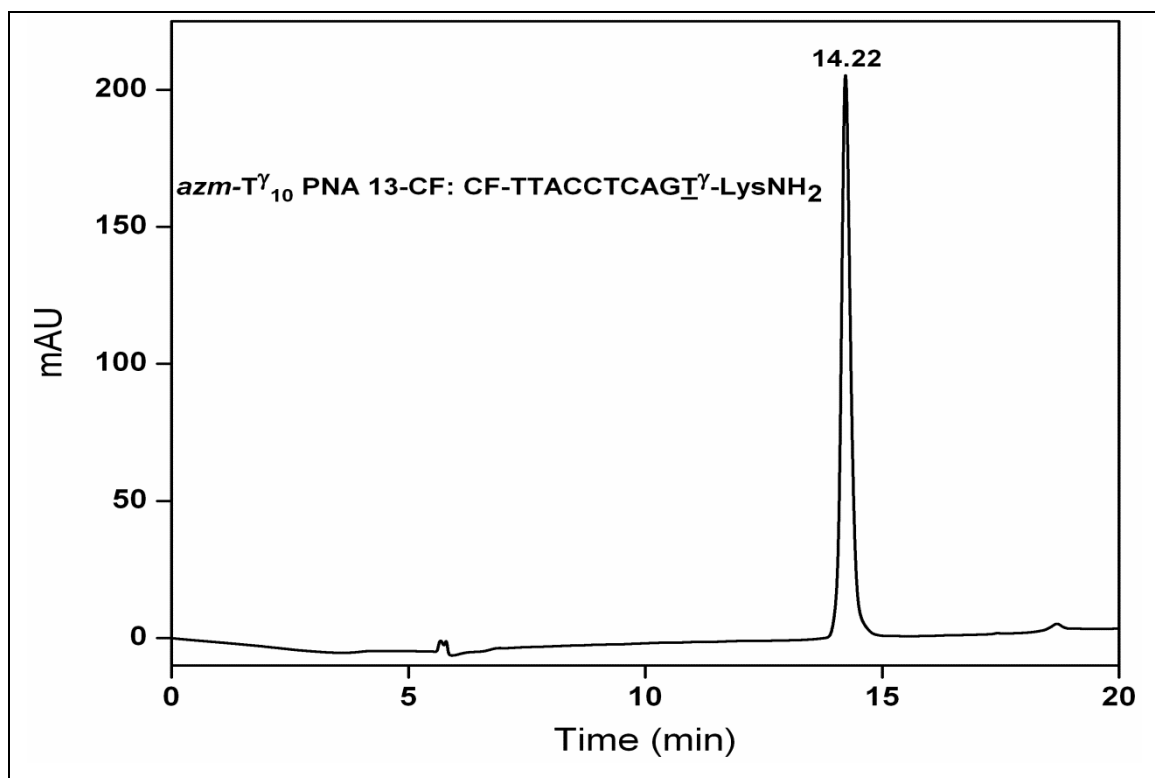
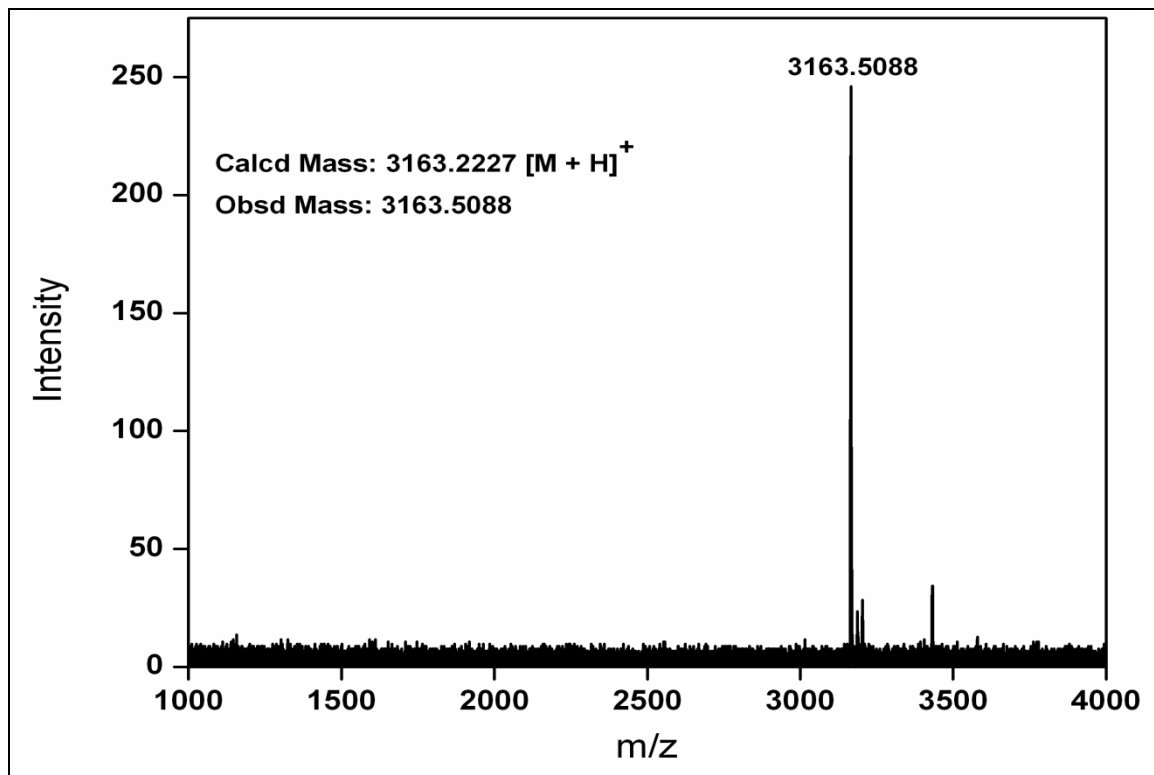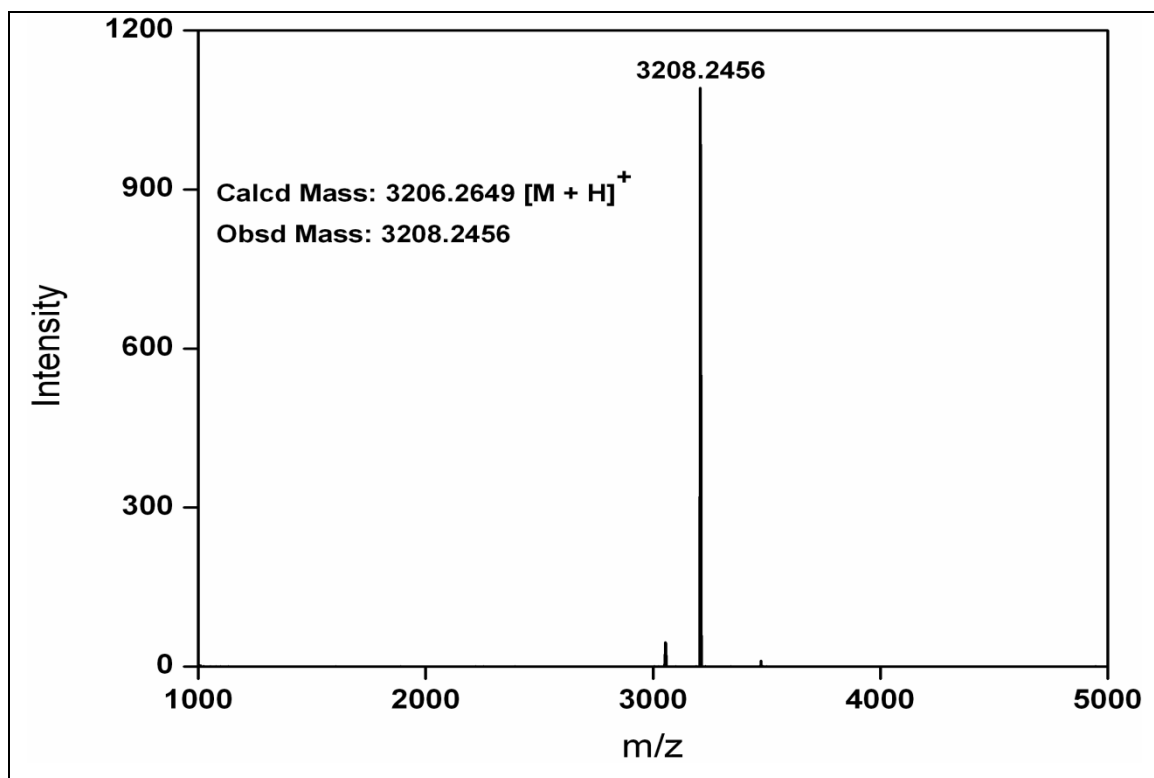## 4.10 Appendix II Characterization data of synthesized compounds

Compound	Description of spectral characterization	Page No.
PNA **1-CF** & **2-CF**	HPLC spectra of PNA **1-CF** & PNA **2-CF**	240
PNA **5-CF** & **6-CF**	HPLC spectra of PNA **5-CF** & PNA **6-CF**	241
PNA **9-CF** & **10-CF**	HPLC spectra of PNA **9-CF** & PNA **10-CF**	242
PNA **13-CF** & **17-CF**	HPLC spectra of PNA **13-CF** & PNA **17-CF**	243
PNA **1-CF** & **2-CF**	MALDI-TOF spectra of PNA **1-CF** & PNA **2-CF**	244
PNA **5-CF** & **6-CF**	MALDI-TOF spectra of PNA **5-CF** & PNA **6-CF**	245
PNA **9-CF** & **10-CF**	MALDI-TOF spectra of PNA **9-CF** & PNA **10-CF**	246
PNA **13-CF** & **17-CF**	MALDI-TOF spectra of PNA **13-CF** & PNA **17-CF**	247

HPLC Trace of PNA 1-CF

aeg PNA 1-CF: CF-TTACCTCAGT-LysNH$_2$

HPLC trace of PNA 2-CF

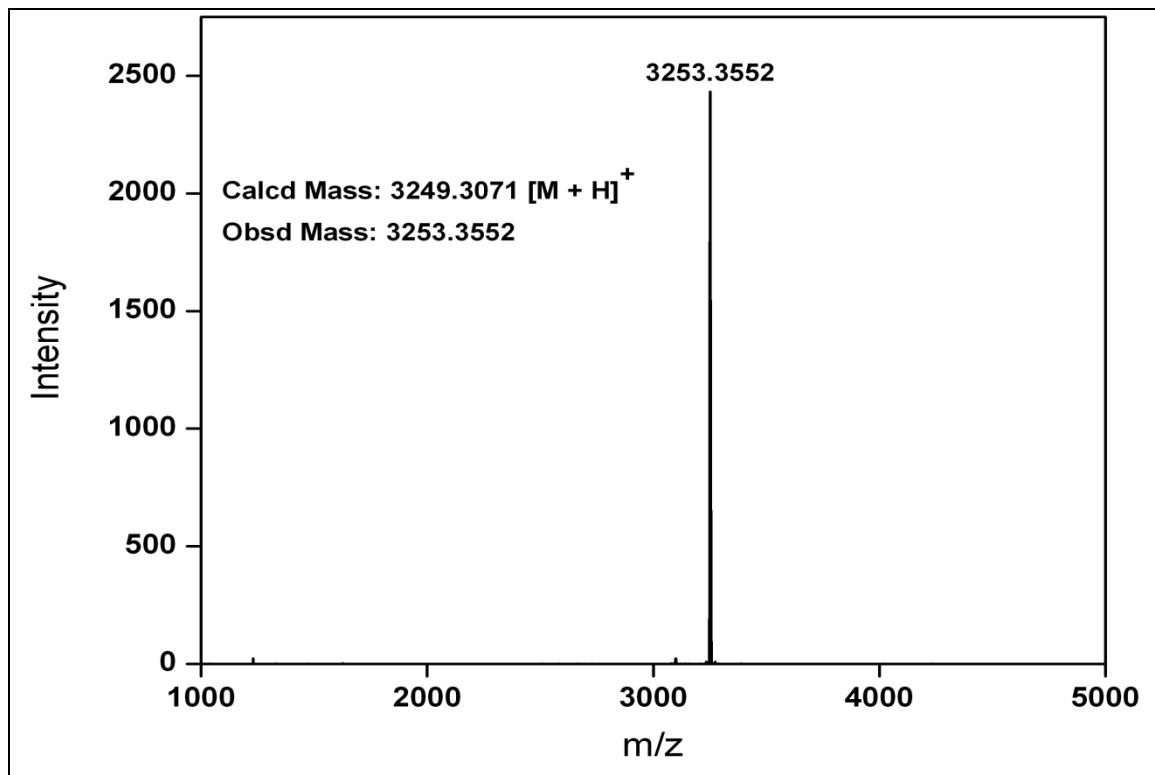*eam-* t$_2$ PNA 2-CF: CF-TtACCTCAGT-LysNH$_2$

HPLC Trace of PNA 5-CF

HPLC Trace of PNA 6-CF

HPLC Trace of PNA 9-CF

egd- $t^{\gamma}_{2,6}$ PNA 9-CF: CF-T$\underline{t}^{\gamma}$ACC$\underline{t}^{\gamma}$CAGT-LysNH$_2$

HPLC Trace of PNA 10-CF

egd- $t^{\gamma}_{2,6,10}$ PNA 10-CF: CF-T$\underline{t}^{\gamma}$ACC$\underline{t}^{\gamma}$CAG$\underline{t}^{\gamma}$-LysNH$_2$

HPLC Trace of PNA 17-CF

azb-T$_{10}$ PNA 17-CF: CF-TTACCTCAG$\underline{T}$-LysNH$_2$

HPLC Trace of PNA 13-CF

azm-T$^{\gamma}_{10}$ PNA 13-CF: CF-TTACCTCAG$\underline{T}^{\gamma}$-LysNH$_2$

MALDI-TOF Mass of PNA 1-CF

Calcd Mass: 3163.2227 [M + H]$^+$
Obsd Mass: 3163.5088
3163.5088

MALDI-TOF Mass of PNA 2-CF

Calcd Mass: 3206.2649 [M + H]$^+$
Obsd Mass: 3208.2456
3208.2456

MALDI-TOF Mass of PNA 5-CF

Calcd Mass: 3249.3071 [M + H]$^+$
Obsd Mass: 3253.3552

MALDI-TOF Mass of PNA 6-CF

Calcd Mass: 3272.1389 [M + Na]$^+$
Obsd Mass: 3276.4036
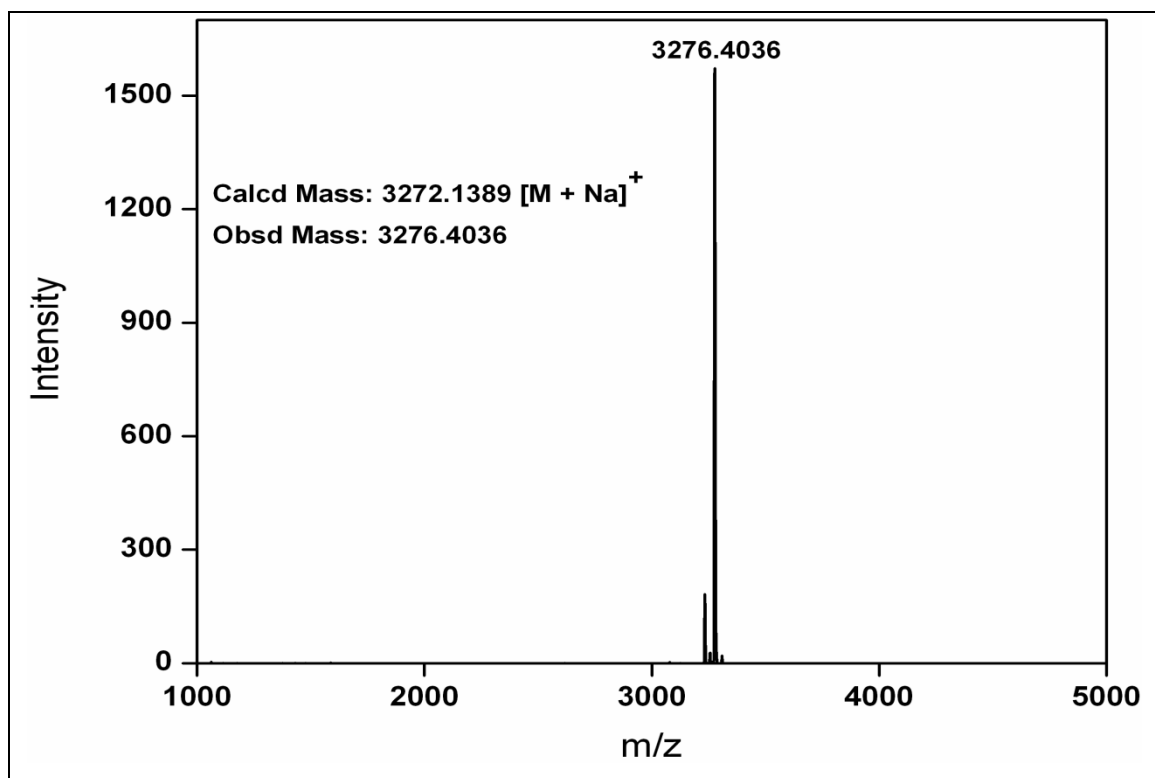
MALDI-TOF Mass of PNA 9-CF
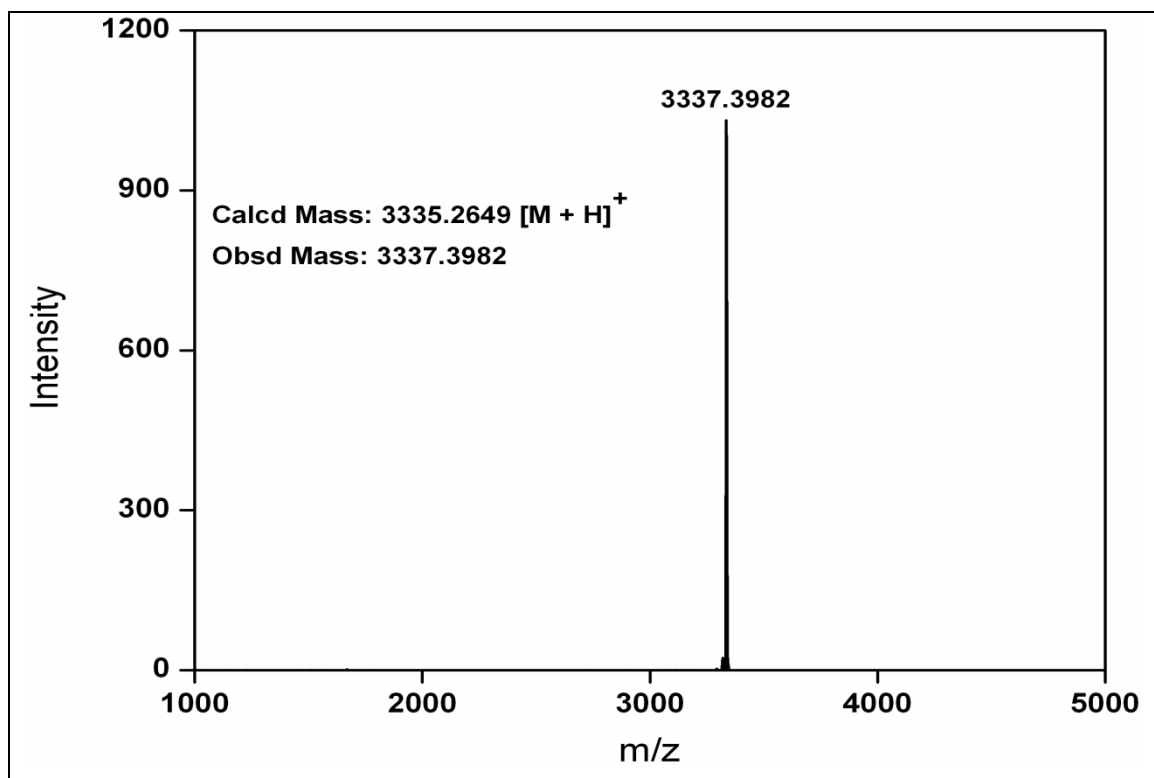
MALDI-TOF Mass of PNA 10-CF
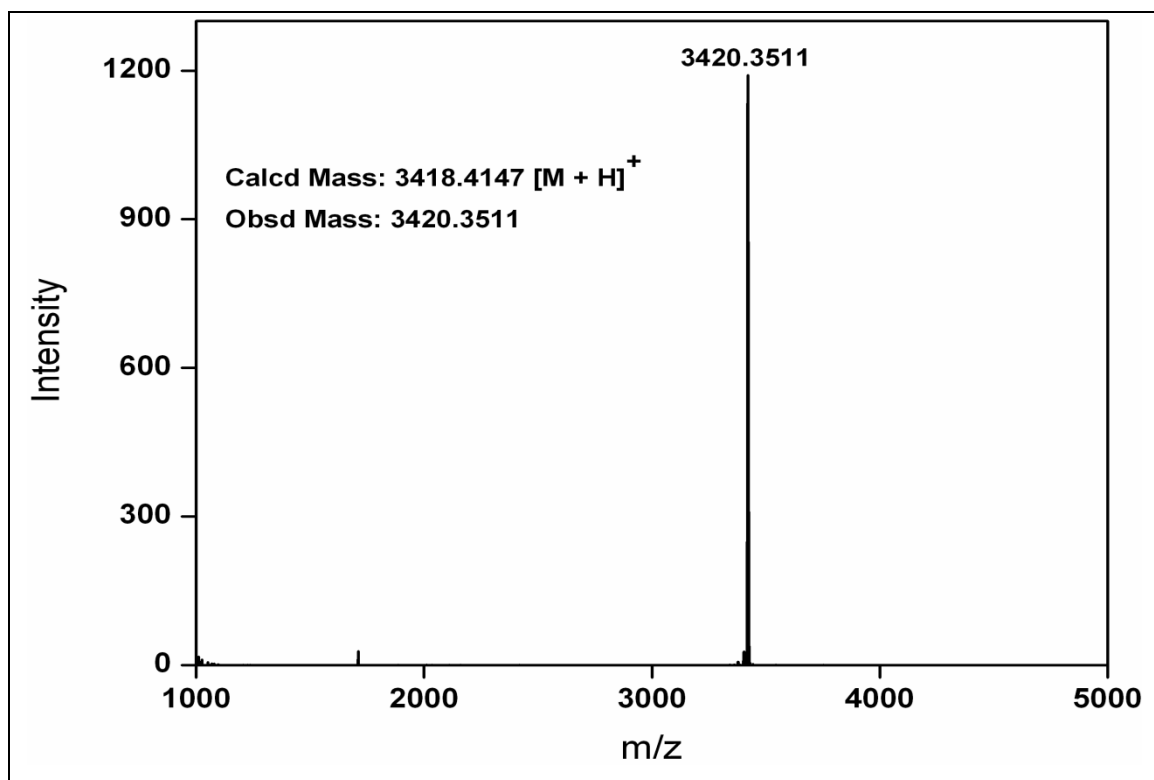
MALDI-TOF Mass of PNA 17-CF

MALDI-TOF Mass of PNA 13-CF