

# **Design and Synthesis of Imidazole Conjugated PNA and Polypeptides as Nuclease Mimics**

**A thesis**

**submitted in partial fulfillment of the requirements of the degree  
of**

**Doctor of Philosophy**

**By**

**Prabhakar Uttamrao Pawar**

**ID: 20123202**

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



**INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE**

**March 2019**

## **CERTIFICATE**

I certify that the work incorporated in the thesis entitled “**Design and Synthesis of Imidazole Conjugated PNA and Polypeptides as Nuclease Mimics**” submitted by **Mr. Prabhakar Uttamrao Pawar** 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:**

**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:

Prabhakar Uttamrao Pawar

## Acknowledgements

I thank my thesis supervisor Prof. Krishna N. Ganesh for his guidance and support through all these years. His words of support and encouragement have been the most important source of strength for me at all times. I am grateful for the lessons I have learnt from him, by listening to his lectures, by discussing with him, and by simply observing him. For that I will be grateful to him for the rest of my life.

I express my sincere thanks to Prof. Srinivas Hotha, Dr. Vaijayanti Kumar and Dr. S. G. Srivatsan for their suggestions and advice. Their insights and questions in RAC meetings have helped me a great deal to understand my project. I am also thankful to IISEE Pune chemistry faculty. Those interactions were short, but very useful as it helped me find fresh perspectives in the projects.

The technical staff in chemistry department (Swati, Mahesh, Megha, Sandip, Yatish and Nitin) has been very kind to me all these years. I am greatly indebted to all of them for their help.

I thank my labmates and my seniors for helping me all these years. Especially, Madan, Nitin, Vijay, Deepak, Satheesh, Tanpreet, Mahesh, Shahaji and Pradnya had been kind and supportive. I am greatly indebted to Pramod, Manoj, Iranna, Dhruv, Shiraj, Om Shanker, Pradeep, Rajat (chhote), Isha and Gaurav for helping me at numerous occasions. I had a wonderful company with my friends Trimbak (Pappu), Anil<sup>2</sup>, Kishor, Sager, Mubarak, Raosaheb, Nandu, Nitin, Pavan, Ajay, Pralhad Aniket, and all IISER, NCI friends. I thank my uncle Dattatray Pawar and cousin brother Sunil Pawar for education awareness. I am grateful to all of them for teaching me so many things in the lab.

Finally, I am forever indebted to my parents, brother Ravindra, sisters in law Swati and sweet nieces Naveli. Their affection and love are the pillars of strength I depend on. I cannot thank them enough, for they keep me sane and they keep me alive.

Prabhakar Uttamrao Pawar



# Contents

Abbreviations	i
Abstract	iv
<b>Chapter 1: Introduction</b>	
1.1 Introduction to nucleic acids	1
1.2 Base Pairing through Hydrogen bonding	2
1.3 Stability of DNA and RNA	4
1.4 Peptide Nucleic Acids (PNAs)	12
1.5 PNA as Artificial nucleases	24
1.6 Scope of present work	25
1.7 References	28
<b>Chapter 2: Synthesis of <i>aeg</i>-Imidazole Monomers and derived PNA Oligomers</b>	
2.1 Introduction	38
2.2 Rationale behind the work	40
2.3 Objectives	43
2.4 Results and Discussion	44
2.4.1 Synthesis of modified PNA monomers	44
2.4.2 Solid Phase Peptide Synthesis	50
2.4.3 Cleavage of the PNA oligomers from the solid support	57
2.4.4 Purification and characterization of the PNA oligomers	57
2.5 Summary	59
2.6 Experimental section	60
2.6.1 General	60
2.6.2 Synthesis of compounds	61
2.6.3 Synthesis of PNA oligomers by solid phase PNA synthesis	81
2.6.4 Cleavage of the PNA oligomers from solid support	82
2.6.5 Purification of the PNA oligomers by RP-HPLC	82
2.7 References	82
2.8 Appendix I	85
<b>Chapter 3: Biophysical Study of Imidazolyl PNA</b>	
3.1 Introduction	130
3.2 Rationale	131

3.3 Objectives	133
3.3.1 Studies with cDNA	134
3.3.2 Studies with cPNA	134
3.4 Results and Discussion	135
3.4.1 Thermal melting studies of antiparallel PNA:DNA hybrids	136
3.4.2 <i>Thermal melting studies of antiparallel PNA:PNA hybrids</i>	143
3.4.3 <i>Thermal melting studies of parallel PNA: PNA hybrids</i>	149
3.4.4 <i>CD studies of Anti Parallel duplexes of PNA 2, PNA 3, PNA 4, PNA 5 with cDNA</i>	154
3.5 Summary	155
3.6 Experimental procedures	156
3.6.2 <i>UV- Tm measurements</i>	156
3.6.3 <i>Circular Dichroism</i>	156
3.7 References	157
<b>Chapter 4: Metal Complexation, Biophysical and RNA Cleavage Studies by Imidazole-PNA conjugates</b>	
4.1 Introduction	160
4.2 Rational and objective of present work	161
4.3 Results and Discussion	162
4.3.1 UV studies of imidazole <i>aeg</i> monomers	162
4.3.2 UV-Vis spectroscopic studies of metal complexation of <i>NH</i> -Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate (II)	163
4.3.3 UV-Vis spectroscopic studies of metal complexation of <i>NH</i> -Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate (II)	164
4.3.4 <i>aeg</i> -Imidazolyl PNA oligomers: Thermal stability stabilities of duplexes and triplexes	166
4.3.5 Thermal melting studies of triplexes	170
4.3.6 RNA cleavage studies	171
4.4 Cleavage of RNA by Zn <sup>2+</sup> ion	181
4.5 Summary	182
4.6 Reference	183
<b>Chapter 5: Design and Synthesis of Imidazole Polyproline Peptides Nuclease Mimics</b>	
5.1 Introduction	187
5.2 Rational and Objectives	189

5.3 Results and Discussion	191
5.3.1 Synthesis of imidazole conjugated 4-aminoproline monomer	191
5.3.1a <i>Synthesis of (2S,4S)-4-aminoproline monomer 6</i>	191
5.3.1b <i>Synthesis of imidazole polyprolyl peptide</i>	192
5.3.1c <i>Purification and characterization of peptides</i>	194
5.3.1d <i>Effect of different solvent and length on imidazole polyprolyl peptide</i>	194
5.3.1e <i>Conformational Studies of 4-N-acetyl (C4-Imidazole) prolyl polypeptide P2 in presence of metal ions</i>	196
5.3.1f <i>Conformational Studies of 4S-amido-(Imidazole-C4-acetyl) Prolyl polypeptide P2 with NiCl<sub>2</sub></i>	197
5.3.1g <i>RNA cleavage study with imidazole polyprolyl peptide and with metal.</i>	198
5.4 Summary	200
2.5 Experimental section	200
2.5.1 Synthesis of compounds	200
5.6 References	204
5.7 Appendix I	207

## Abbreviations

A	Adenine
Abs	Absolute
Ac <sub>2</sub> O	Acetic anhydride
CAN	Acetonitrile
<i>Aeg</i>	Aminoethylglycine
<i>Ap<sub>g</sub></i>	N-(3-aminopropyl)glycine
<i>Aep</i>	Aminoethylpropyl
Ap	Antiparallel
<i>aq.</i>	Aqueous
(Boc) <sub>2</sub> O	Boc anhydride
Bn	Benzyl
5FBn	Pentafluorophenylmethylene
BIAB	(Diacetoxyiodo)benzene
Bt	Benzotriazole
C	Cytosine
Calcd	Calculated
Obsvd	Observed
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
CF/5(6)-CF	5(6)-Carboxyfluorescein
Ch	Cyclohexyl
CHCA	$\alpha$ -cyano-4-hydroxycinnamic acid
Cp	Cyclopentyl
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DHB	2,5-dihydroxybenzoic acid
DIC	N,N'-diisopropylcarbodiimide
DIPEA	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
DNase	Deoxyribonuclease
Ds	Double stranded
EBA	Ethylbromo acetate
EDTA	Ethylene diamine tetraacetic acid
Et	Ethyl
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- hexafluorophosphate
HOBt	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HRMS	High resolution mass spectrometry
in situ	In the reaction mixture
in vivo	Within the living
IR	Infra-red
LC-MS	Liquid Chromatography-Mass Spectrometry
Lys	Lysine
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
PCR	Polymerase chain reaction
Pd	Palladium
ppm	Parts per million
PPh <sub>3</sub>	Triphenylphosphine
PNA	Peptide Nucleic Acid
PS-oligo	Phosphorothioate-oligo
Rt	Retention time
RNA	Ribonucleic Acid
RNase	Ribonuclease
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 <sub>3</sub> N	Triethylamine
TEMPO	Tetramethylpiperidine 1-oxyl
TFA	Trifluoroacetic acid
Tf <sub>2</sub> O	Triflicanhydride
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TMSCl	Trimethylsilyl chloride
TBDMSCl	<i>tert</i> -butyldimethylsilyl chloride
<i>T<sub>m</sub></i>	Melting temperature
UV-Vis	Ultraviolet-Visible

## Abstract

The thesis entitled “**Design and Synthesis of Imidazole Conjugated PNA and Polypeptides as Nucleases Mimics**” comprises of a systematic investigation into the use of peptide nucleic acid (PNA) and polyproline scaffolds in the rational design of imidazole based artificial nucleases.

Peptide nucleic acids are a class of DNA analogues that are stable to cellular enzymes, and allow sequence specific targeting of DNA and RNA. Thus, PNAs are suitable scaffolds for rational design of artificial nucleases. In the present work, PNAs have been modified with imidazole, which plays a key role in the active sites of ribonucleases, at various positions to produce potential PNA based artificial nucleases. The different sequences with one or more imidazolyl PNA monomeric units incorporated in them were synthesized using solid phase peptide synthesis protocol. The RNA hydrolysing activity of the various rationally constructed PNAs were evaluated in the presence and absence of divalent metal ions using high performance liquid chromatography. Additionally, the ability of imidazole to act as an unnatural surrogate for the usual nucleobases was assessed by examining the effect of the modification on thermal melting properties of duplexes with complementary DNA and PNA. In order to examine whether a protein-like environment aids the catalytic activity, imidazole modified polyprolines of different lengths were also constructed and their hydrolytic activity was studied as well.

This thesis is presented in five chapters:

**Chapter 1:** Introduction

**Chapter 2:** Synthesis of *aeg*-Imidazole Monomers and Derived PNA Oligomers

**Chapter 3:** Biophysical Study of Imidazolyl PNA

**Chapter 4:** Metal Complexation, Biophysical and RNA Cleavage Studies by Imidazole-PNA Conjugates

**Chapter 5:** Design and Synthesis of Imidazole Polyproline Peptide Nuclease Mimics

---

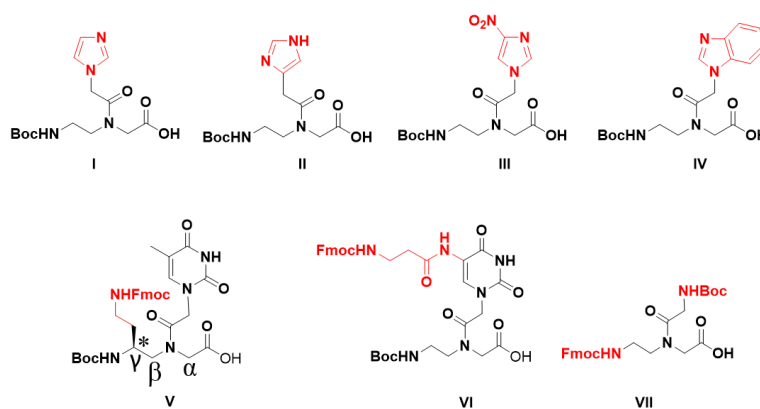
**Chapter 1: Introduction**

This chapter gives an overview of the background literature for the undertaking research work, emphasizing the recent advancements in the field of peptide nucleic acids and various artificial nuclease systems. It focuses on the attempts made in the last few years to optimize the properties of oligonucleotides and their analogues through chemical modifications for targeted RNA cleavage.

Peptide Nucleic Acid (PNA), first introduced by Nielsen *et al* in 1991, is a synthetic analogue of DNA/RNA. It has higher binding affinity and better specificity towards complementary sequences than the corresponding DNA/RNA. PNAs have a neutral N-(2-aminoethyl)-glycine backbone with nucleobases attached via a methylene linker. This renders them resistant to hydrolytic enzymes like proteases and nucleases. In addition to regular duplexes, PNAs can also form stable PNA<sub>2</sub>:DNA/RNA triplexes with complementary DNA/RNA. These exceptional properties of PNAs are favourable for their applications as RNA inhibition agents and as artificial nucleases.

## Chapter 2: Synthesis of *aeg*-Imidazole Monomers and derived PNA Oligomers

This chapter deals with the introduction of imidazole in the peptide nucleic acid structure at various positions. It describes the synthesis and characterization of rationally designed PNA monomers (Figure 1), where the imidazole moieties are incorporated at various sites: (a) in the backbone, (b) on the nucleobase and (c) by replacing the base with imidazole. Bis-imidazole PNA was also synthesized which contained the modifications both on the backbone and on the sidechain. All the intermediates were characterized by <sup>1</sup>H and <sup>13</sup>C NMR spectroscopy, mass spectral analysis and other appropriate analytical data.

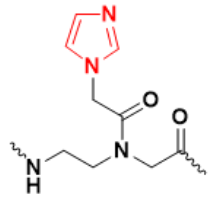
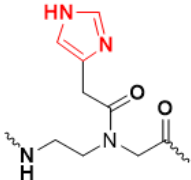
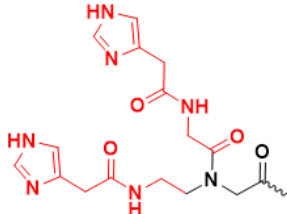


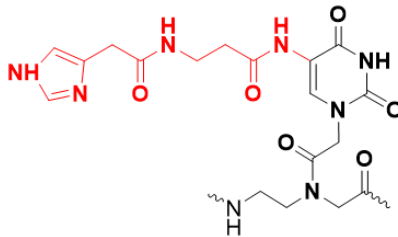
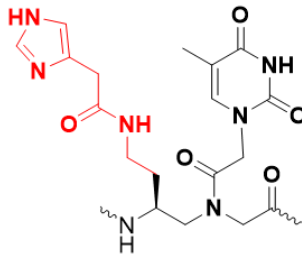
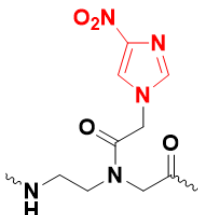
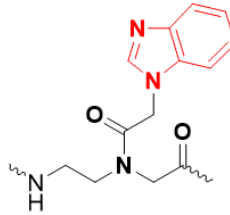
**Figure 1** Structure of synthesized modified PNA monomers



The rationally designed modified PNA monomers have been incorporated into 10-mer pyrimidine and, 9-mer and 12-mer purine-pyrimidine PNA sequences. This imidazole modified monomers were inserted into *aeg* PNA sequences at specific positions using HBTU as coupling reagent by solid phase peptide synthesis protocol. All the imidazole modified and unmodified PNA oligomers (Table 1) were synthesized by following Boc-protocol of solid phase synthesis. After cleavage from the solid support, the PNA oligomers were purified by RP-HPLC and characterized by MALDI-TOF spectrometry

**Table 1** PNA oligomers with modified/unmodified monomers at various positions

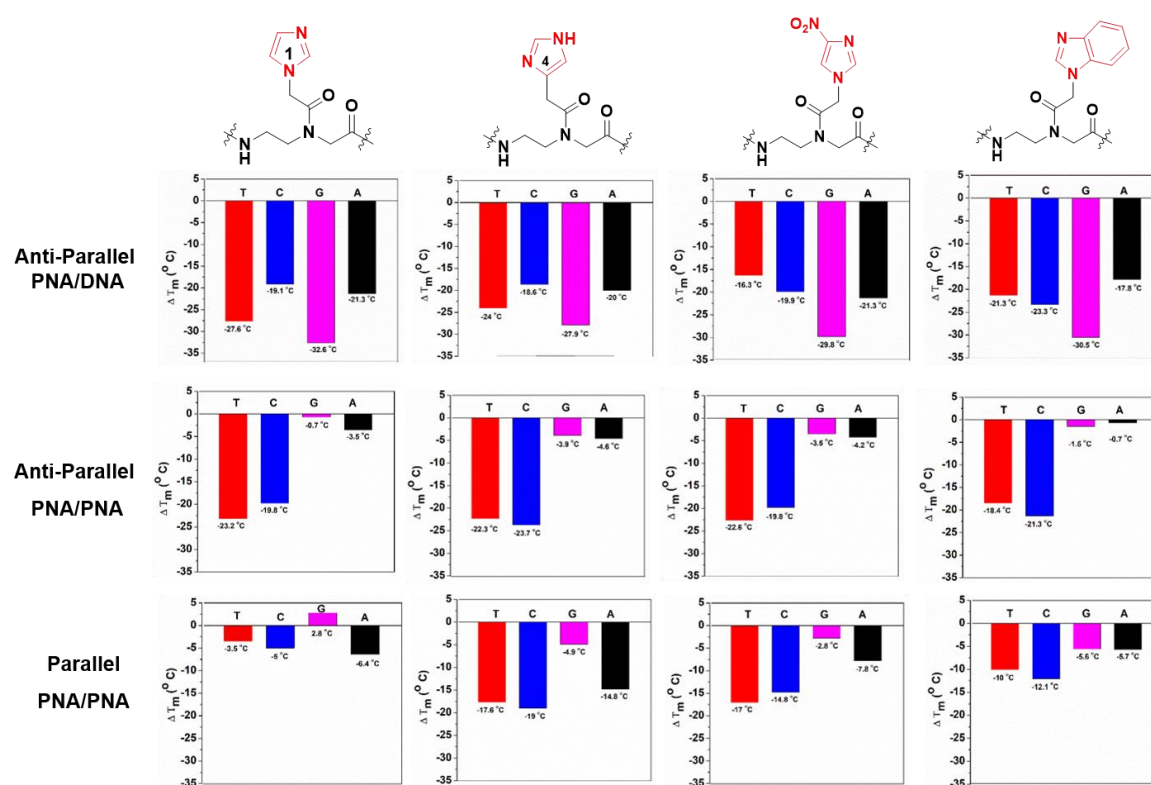
Entry	Sequence Code	PNA sequences	Imidazole modifications
1	<i>aeg</i> PNA 1	H-G G C A T G C C LysNH <sub>2</sub>	A/G/C/T = <i>aeg</i> PNA
2	<i>Im(N1)</i> -PNA 2 <i>Im(N1)</i> -PNA 9	H-GGCA <i>Im(N1)</i> TGCCLysNH <sub>2</sub> H-GTAG <i>Im(N1)</i> TACTTLysNH <sub>2</sub>	 <i>Im(N1)</i> =aminoethyl(imidazole N1-acetamido) glycyl
3	<i>Im(C4)</i> -PNA 3 <i>Im(C4)</i> -PNA 10	H-GGCA <i>Im(C4)</i> TGCCLysNH <sub>2</sub> H-GTAG <i>Im(C4)</i> TACTTLysNH <sub>2</sub>	 <i>Im(C4)</i> =aminoethyl(imidazole C4-acetamido) glycyl
4	<i>Bis-Im</i> -PNA 4	<i>Bis-Im</i> -TTTTTGCLysNH <sub>2</sub>	 <i>Bis-Im</i> = <i>aeg</i> -Bis-imidazole

5	$U^{Im}$ T <sub>10</sub> -PNA 5	H-TTTT $U^{Im}U^{Im}$ TTTTLysNH <sub>2</sub>	
6	$U^{Im}$ -PNA 6	H-TCTCAAG $U^{Im}$ TGGGLysNH <sub>2</sub>	$U^{Im}$ = aeg-C5-(β-Ala-Im-U)
7	$\gamma C^{Im}$ T <sub>10</sub> -PNA 7	H-TTTT $\gamma C^{Im}\gamma C^{Im}$ TTTTLysNH <sub>2</sub>	
8	$\gamma C^{Im}$ -PNA 8	H-TCTCAAG $\gamma C^{Im}$ TGGGLysNH <sub>2</sub>	$\gamma C^{Im}$ = aminoethyl-γC-(S-eam-Im)glycyl
9	$ImNO_2$ -PNA 11	H-GTAG $ImNO_2$ TACTTLysNH	
			$ImNO_2$ = ae-(4-NO <sub>2</sub> -imidazole - N1-acetamido) glycyl
10	$Bzim$ -PNA 12	H-GTAG $Bzim$ TACTTLysNH <sub>2</sub>	
			$Bzim$ = ae-(benzimidazole N1-acetamido) glycyl
11	PNA 13	H-AAGTAACTACLysNH <sub>2</sub>	A/G/C/T = aeg PNA
12	PNA 14	H-AAGTATCTACLysNH <sub>2</sub>	
13	PNA 15	H-AAGTACCTACLysNH <sub>2</sub>	
14	PNA 16	H-AAGTAGCTACLysNH <sub>2</sub>	
15	PNA 17	H-CATCAATGAALysNH <sub>2</sub>	
16	PNA 18	H-CATCTATGAALysNH <sub>2</sub>	
17	PNA 19	H-CATCCATGAALysNH <sub>2</sub>	

18	PNA <b>20</b>	H-CATC <b>G</b> ATGAALysNH <sub>2</sub>
19	PNA <b>21</b>	H-GTAG <b>C</b> TACTTlysNH <sub>2</sub>
20	PNA <b>22</b>	H-GTAG <b>T</b> TACTTlysNH <sub>2</sub>

### Chapter 3: Biophysical Study of Imidazolyl PNA

This chapter deals with studies examining the role of imidazole as an unnatural nucleobase. Among the unnatural nucleobases, 3-nitropyrrole is interesting since it can complement with any of the four standard nucleobases. This universal base shares its structural properties with 3-nitroimidazole. Imidazole and 3-nitroimidazole, like 3-nitropyrrole, have five membered rings and are both aromatic in nature. Therefore, hybridization properties of N1-linked imidazolyl PNA modifications and C4-linked imidazolyl PNA modifications were studied.



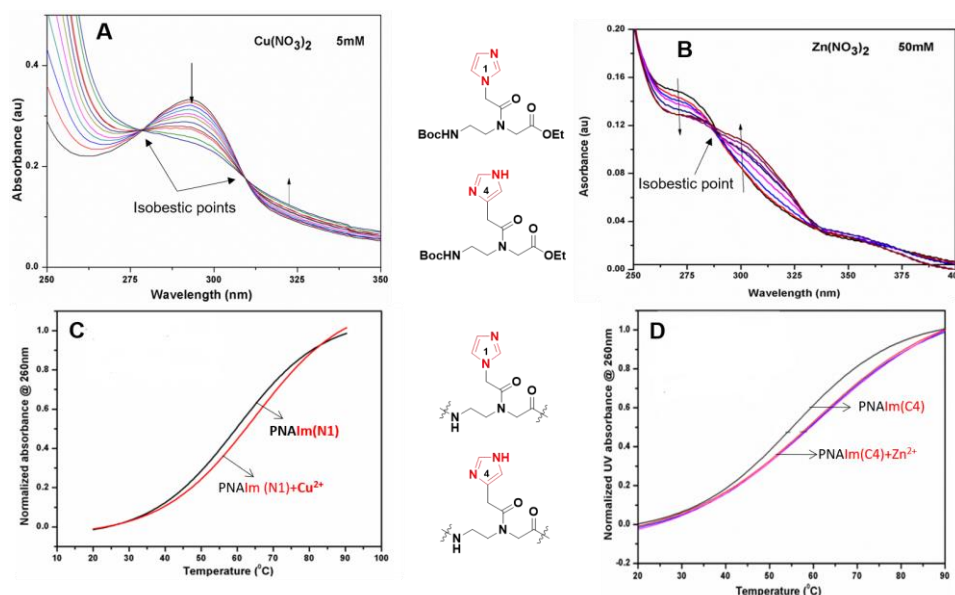
**Figure 2** Relative thermal Stability of PNA derived duplexes

Investigation of duplexes from all imidazolyl PNAs with complementary antiparallel DNA showed significant destabilisation with all natural nucleobases in

cDNA. In case of antiparallel PNA:PNA duplexes, imidazolyl PNA showed discrimination between purines (adenine (A) and guanine (G)) and pyrimidines (cytosine (C) and thymine (T)). These imidazolyl PNAs showed more stability with complementary G and A PNAs whereas, T and C PNAs showed significantly less stable duplexes. In case of parallel PNA:PNA duplexes, all imidazolyl PNA showed G to be more compatible against imidazole modifications and stabilise corresponding duplex better than other three nucleobases A, C and T.

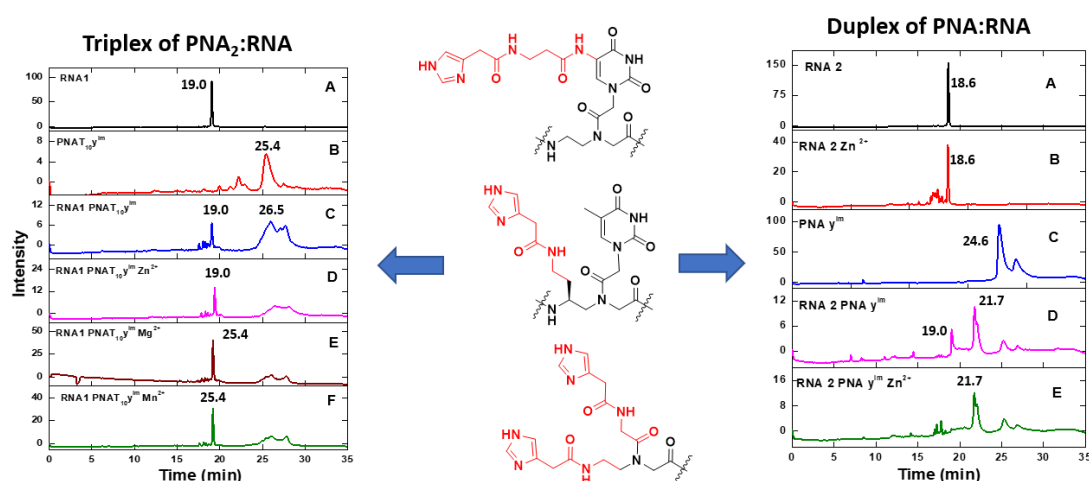
#### Chapter 4: Metal Complexation, Biophysical and RNA Cleavage Studies by Imidazole-PNA conjugates

In this chapter the *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate and *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate monomers showed selective metal complex with  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  respectively (Figure 3 A and B). These monomers are incorporated into the self-complementary PNA sequences. The thermal stability ( $T_m$ ) of PNA:PNA duplexes of *Im(N1)*-PNA **1** and *Im(C4)*-PNA **2** are significantly increased 5 °C specifically in presence of  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  respectively (Figure 3 C and D).



**Figure 3** A) UV absorption spectra of Metal binding study with *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** and B) *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate **II**; C) Melting curves of the Imidazole N1-linked self-complementary duplex [*Im(N1)*PNA **1**]<sub>2</sub> in the presence of  $\text{Cu}^{2+}$  ion and D) Imidazole-C4- conjugated self-complementary [*Im(C4)*PNA **2**]<sub>2</sub> duplex in the presence of  $\text{Zn}^{2+}$  ion.

The base modified imidazole PNAs of  $U^{Im}T_{10}$ -PNA **4**, backbone modified  $\gamma C^{Im}T_{10}$ -PNA **5** and end modified *Bis-Im*-PNA **6** were studied for their potential to cleave complementary RNA **1** in the presence and absence of metal ions. In triplex and half duplex ( $U^{Im}T_{10}$ -PNA **4**)<sub>2</sub>:RNA **1**, ( $\gamma C^{Im}T_{10}$ -PNA **5**)<sub>2</sub>:RNA **1** and *Bis-Im*-PNA **6**:RNA **1**, the imidazole modified PNAs cleaved 10% RNA **1** in the absence of metal ions. In the presence of metal ions, cleavage increased up to 20%. However, no sequence directed selectivity was seen in any of the cases. With duplexes,  $U^{Im}$ -PNA **7**:RNA **1**/RNA **2** and  $\gamma C^{Im}$ -PNA **8**:RNA **2**/RNA **3**, the imidazole modified PNAs did not cleave RNA **2**/ RNA **3** in the presence or absence of  $Zn^{2+}$  metal ions, but partial degradation of RNA **2**/ RNA **3** was observed only in the presence of  $Zn^{2+}$  ion.



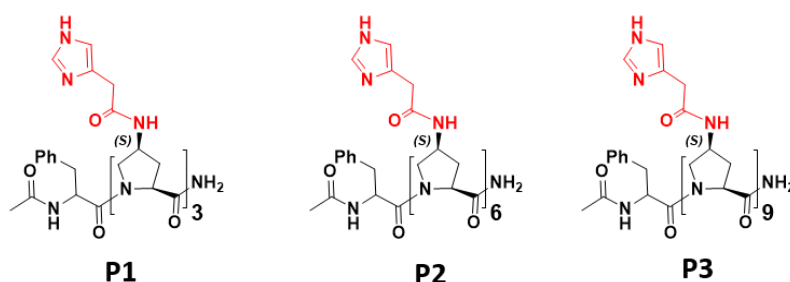
**Figure 4** Representative HPLC graph for RNA cleavage study

## Chapter 5: Design and Synthesis of Imidazole Polyproline Peptides Nuclease Mimics

Small peptides mimic the protein environment. In this context, polyproline offers a simple, conformationally controllable, and structurally amenable scaffold for the design of peptide based RNase mimics. In continuation with the goal of previous chapters to create artificial ribonucleases, wherein imidazole units were placed on nucleic acid templates, herein, the rationale was to functionalise the C4 position of prolines in polyprolines with imidazole for imparting nucleolytic properties for hydrolysing the bound RNA strands.

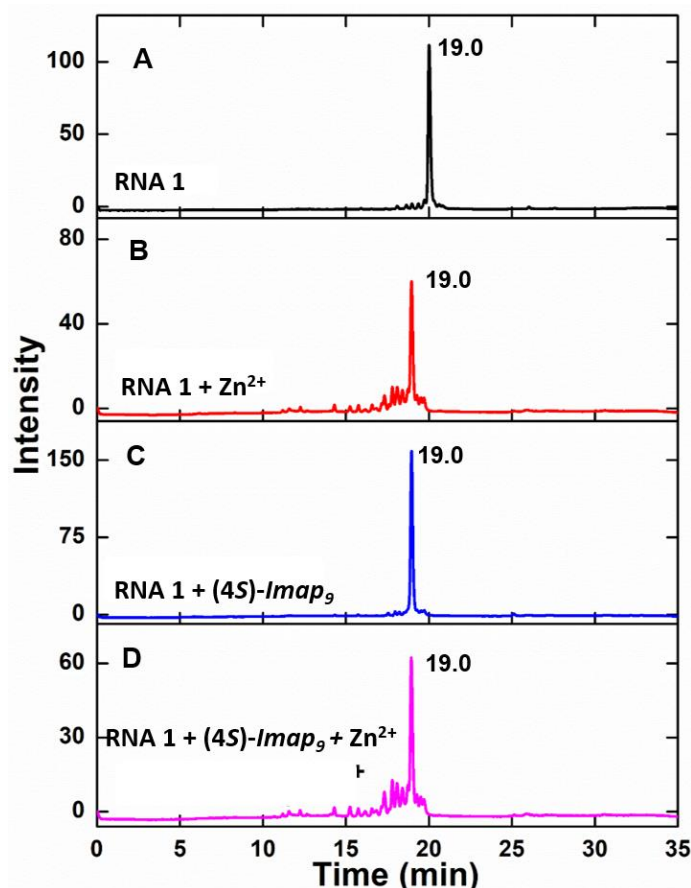
The synthesis of (2*S*,4*S*) aminoproline monomer and its incorporation into polyproline peptide by solid phase peptide synthesis protocol were accomplished

successfully. Further, the 4*S*-aminopolyprolyl peptides were coupled to imidazole-4-acetic acid using solid phase synthesis (Figure 5), cleaved from resin, purified by HPLC and characterised by MALDI TOF.



**Figure 5** Imidazolyl-4-aminoproline peptides

The conformation of the peptides derived from monomers in which imidazole was conjugated to 4-amino prolines were studied by CD. This suggested that they retained PP-II conformation in both buffer and TFE. The **P1**, **P2**, and **P3** peptides (Figure 5) of increasing lengths exhibited slight increase in the magnitude of positive CD bands in buffer and TFE. The titration of imidazolyl polyproline peptide with  $\text{ZnCl}_2$ ,  $\text{MgCl}_2$ ,  $\text{CoCl}_2$ , and  $\text{CuCl}_2$  solutions showed no changes in PP-II conformation. When the imidazolyl polyproline peptide was titrated with  $\text{NiCl}_2$ , the CD signature was found to be drastically altered. This was attributed to the formation of Ni-complex with the imidazolyl moiety, which is well known with polyhistidines and did not match with any of the known secondary structures.



**Figure 6** RNA hydrolysis by **P3** peptide

Further, the imidazolyl-4-aminoproline peptides were studied for the RNA hydrolysis activity in the presence and absence of metal ions (Figure 6). We found that they did not cause any **RNA 1** degradation in the absence of metal ions; however, they caused hydrolysis of **RNA 1** by less than 10% in the presence of Zn ion.

### Summary of thesis

The imidazole moiety is incorporated on PNA monomers by replacing the base with imidazole. All monomers were characterized by  $^1\text{H}$  &  $^{13}\text{C}$  NMR spectroscopy, mass spectral analysis and other appropriate analytical data. The rationally designed modified PNA monomers have been incorporated into 10-mer pyrimidine and, 9-mer and 12-mer mixed purine- pyrimidine PNA sequence. The PNA oligomers after cleavage were purified by RP-HPLC and characterized by MALDI-TOF spectrometry. Duplexes from all imidazolyl PNAs with complementary antiparallel DNA showed significant destabilisation with all natural nucleobases in cDNA. In case of antiparallel PNA:PNA duplexes, these imidazolyl PNAs showed more stability with complementary G and A

PNAs whereas, T and C PNAs showed significantly less stable duplexes. In case of parallel PNA:PNA duplexes, all imidazolyl PNAs showed G to be more compatible against imidazole modifications and stabilise corresponding duplex better than other three nucleobases A, C and T. The *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate and *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate monomers showed selective metal complex with  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  respectively and the thermal stability ( $T_m$ ) of PNA:PNA duplexes of *Im(N1)*-PNA **1** and *Im(C4)*-PNA **2** are significantly increased 5 °C specifically in presence of  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  respectively. In triplex and half duplex (*U<sup>Im</sup>*T<sub>10</sub>-PNA **4**)<sub>2</sub>:RNA **1**, (*γC<sup>Im</sup>*T<sub>10</sub>-PNA **5**)<sub>2</sub>:RNA **1** and *Bis-Im*-PNA **6**:RNA **1**, the imidazole modified PNAs cleaved 10% RNA **1** in the absence of metal salts and in presence of metal salt cleavage increased upto the 20%. With duplexes *U<sup>Im</sup>*-PNA **7**:RNA **1**/RNA **2** and *γC<sup>Im</sup>*-PNA **8**:RNA **2**/RNA **3**, the imidazole modified PNAs did not cleave RNA **2**/ RNA **3** in presence or absence of  $\text{Zn}^{2+}$  metal ions, but RNA **2**/ RNA **3** partial degradation was observed only in presnsce of only  $\text{Zn}^{2+}$  ion. Further, the imidazolyl-4-aminoproline peptides were studied for the RNA hydrolysis in presence and absence of metal; however, they did not cause any significant degradation (observed degradation less than 10%) of RNA **1** in the absence or in presence of Zn salt.

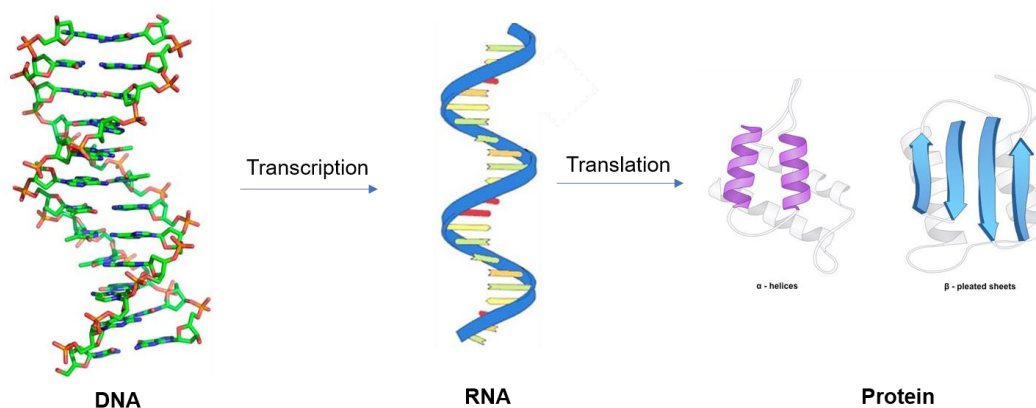


# **Chapter 1**

## **Introduction**

## 1.1 Introduction to nucleic acids

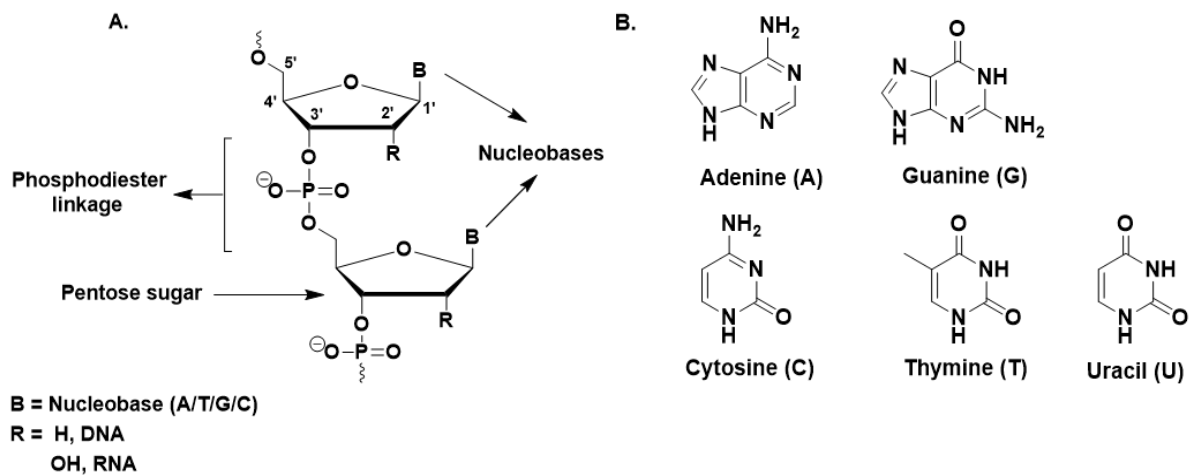
Nucleic acids are the essential biological macromolecules present in all known forms of life. The most supreme biological macromolecules are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Their important functions include storage, transmission, and expression of the genetic information within the biological systems (Figure 1.1).



**Figure 1.1** Biomacromolecules involved in the flow of genetic information

The DNA and RNA molecules are made up of repeating units of nucleotides. Each nucleotide unit consists of a nitrogenous base (purine or pyrimidine), a pentose sugar<sup>1</sup> and a phosphate group. Both DNA and RNA comprise of four nucleobases: adenine (A) and guanine (G) which are purines, and cytosine (C) and thymine (T, in DNA) or uracil (U, in RNA) that are pyrimidines (Figure 1.2B). The structures of DNA and RNA differ from each other in their sugar: DNA contains a deoxyribose unit at 2' position, while RNA has a ribose sugar at the same 2' position which has a tremendous effect on stability and conformational preferences of RNA. DNA and RNA also differ in their nucleobase composition. The three heterocyclic bases (A, G and C) are common in both DNA and RNA; U is present only in RNA and T is found in DNA. In 1953, Watson and Crick proposed that the molecular structure of DNA consists of two helical chains, each coiled around the same axis with a right handed twist.<sup>2</sup> In these linear co-polymers, 3'-5' phosphodiester bonds link successive  $\beta$ -D-deoxyribofuranose of the nucleotides to form the negatively charged backbone. The two DNA strands are held together by specific hydrogen bonds between complementary base pairs (A:T & G:C, commonly known as Watson-Crick base pairs) to form antiparallel double helical structure. The backbone of the double helical strands is positioned outside of the main helix with the H-bonded

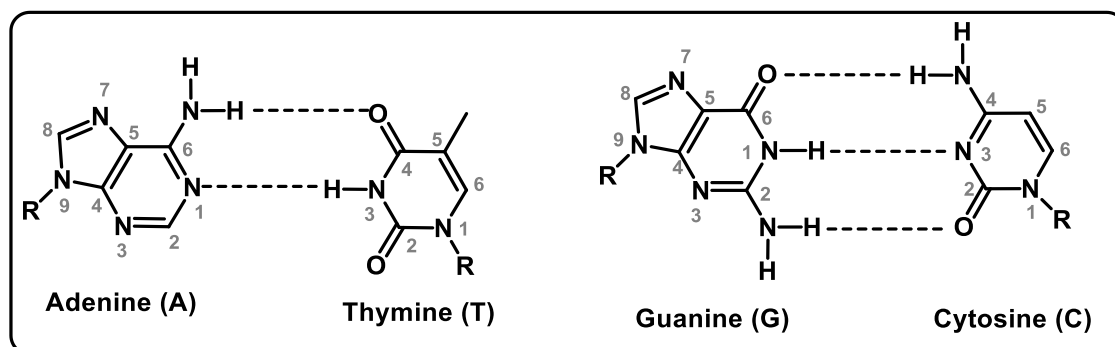
hydrophobic nucleobases stacked perpendicular to the helix axis located in the interior (Figure 1.1A).



**Figure 1.2** Chemical structures of A) DNA and RNA B) Nucleobases

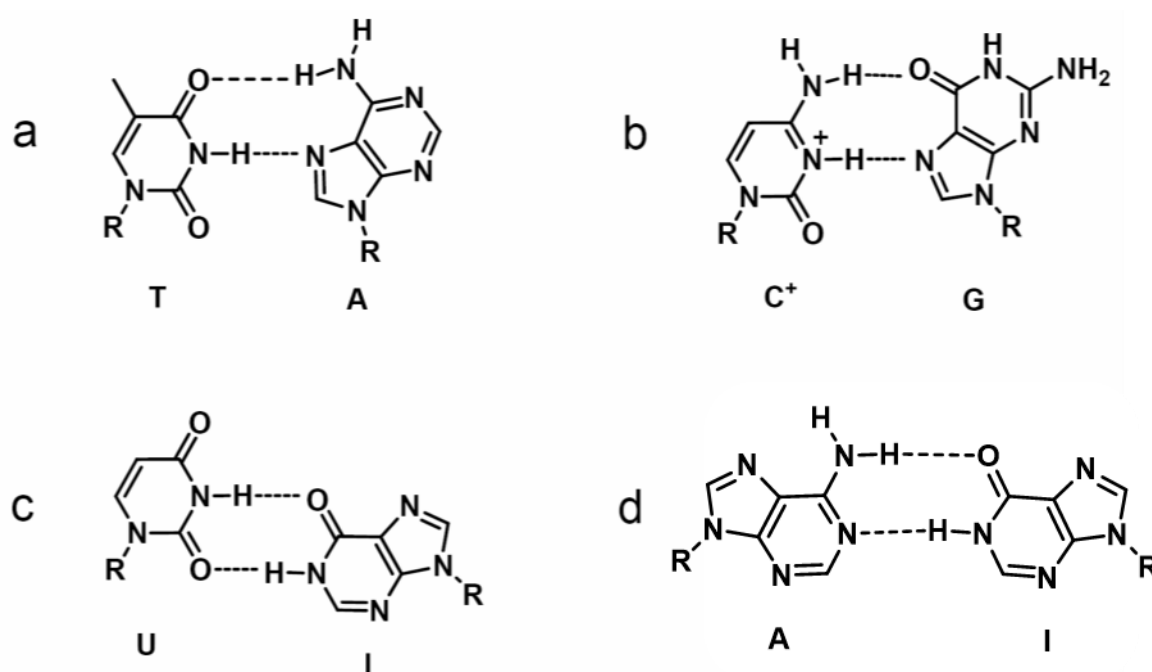
## 1.2 Base Pairing through Hydrogen bonding

Sequence dependent molecular recognition between strands in nucleic acids through complementary hydrogen bonding is one of the most important principles of molecular self-assembly that governs information processing in the complex biological systems. The hydrogen bonds are formed specifically between the *amino-keto* tautomers of the bases that establishes the high fidelity in DNA transcription and translation processes. The N-H groups of the bases are potent hydrogen bond donors, while the  $sp^2$  hybridized electron pairs on oxygen of the carbonyl (C=O) groups and nitrogens present in the aromatic ring are good hydrogen bond acceptors. This leads to Watson-Crick hydrogen bonding with 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

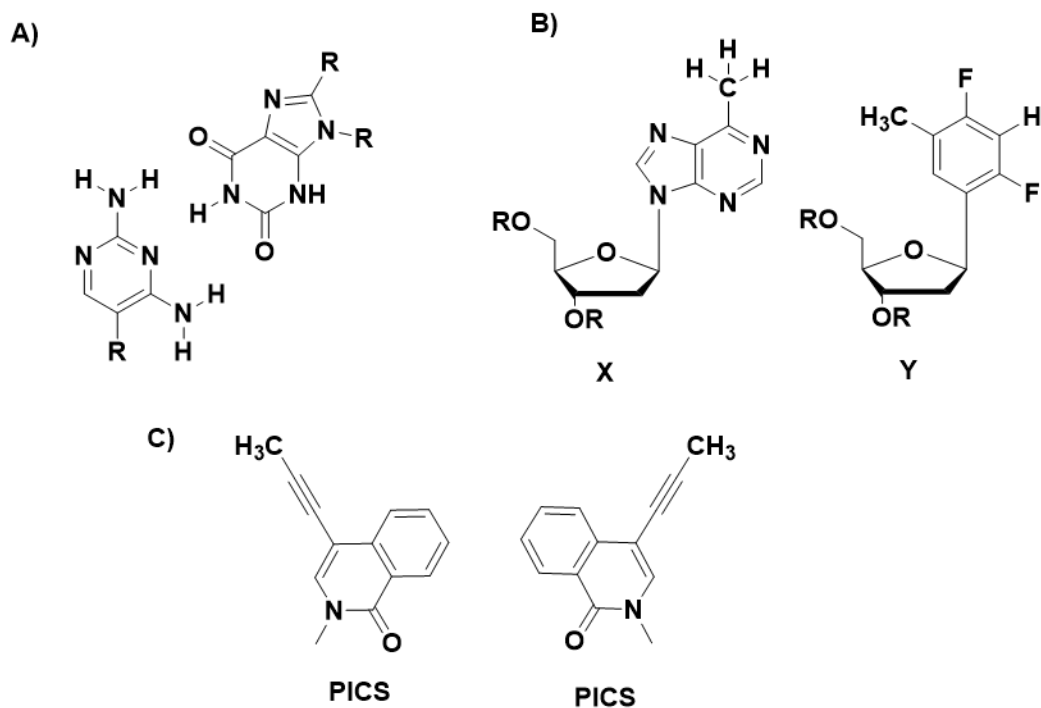
Other possible hydrogen bonding schemes of biological relevance are the Hoogsteen<sup>3</sup> (HG) and Wobble<sup>4</sup> base pairs. Hoogsteen base pairing involves simultaneous H-bonding both from the 5- and 6-membered purine rings (N7 / <sup>6</sup>NH<sub>2</sub>, <sup>6</sup>CO). It is not isomorphous with Watson-Crick base pairing because they have 80° angle between the glycosidic bonds and 8.6 Å separation between the anomeric carbons (Figure 1.4 a, b). In addition to Hoogsteen base pairing, the C6-substituents (NH<sub>2</sub>/CO) of 6-membered purine ring of A and G can also be involved in simultaneous WC H-bonding with T and C and this feature leads to formation of triple helix (Figure 1.3) by DNA/RNA. Wobble base pairing (Figure 1.4 c and d, UI/AI where I=Inosine) observed in RNA strands are a fundamental feature of RNA secondary structures that play a crucial role in translation of genetic code.



**Figure 1.4** (a & b) Hoogsteen base pairing and (c & d) Wobble base pairing

### 1.2.1 Hydrogen bonding of unnatural nucleobases

Unnatural nucleobases have attracted considerable attention. For the last few decades, researchers have been developing unnatural base pairs and exploring ways to expand the genetic alphabets.<sup>5-9</sup> Several types of unnatural base pairs that are compatible with the replication machinery with high fidelity and efficiency in combination with natural A-T and G-C pairs have been reported. Some of these are shown in Figure 1.5.

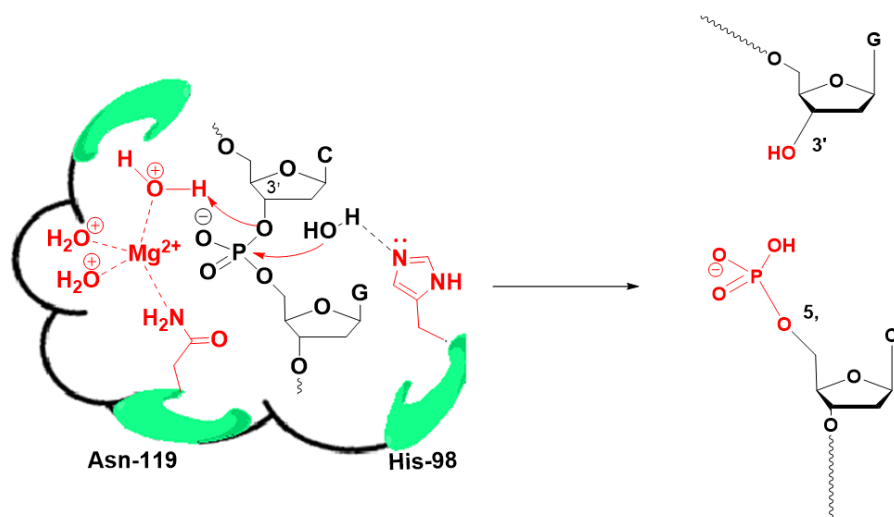


**Figure 1.5** (a & b) Unnatural nucleobase base pairing

Benner's group<sup>10</sup> synthesized four types of unnatural base pairs with nonstandard hydrogen bonding patterns (Figure 1.5A). Kool *et.al.*<sup>11</sup> designed non-hydrogen-bonded base pairs between shape complementarity of the natural base pairs (Figure 1.5B). Romesberg *et. al.*,<sup>12</sup> synthesized hydrophobic self-pair of propynyl isocarbostyryl (PICS), which is the first successful demonstration of a synthetic base pair (Figure 1.5 C). A plethora of nonstandard base pairs have been developed through improvements in rational design of synthetic analogues of nucleobases employing principles of various hydrogen bonding patterns, shape complementarity, hydrophobicity, and electrostatic interactions.<sup>13-17</sup>

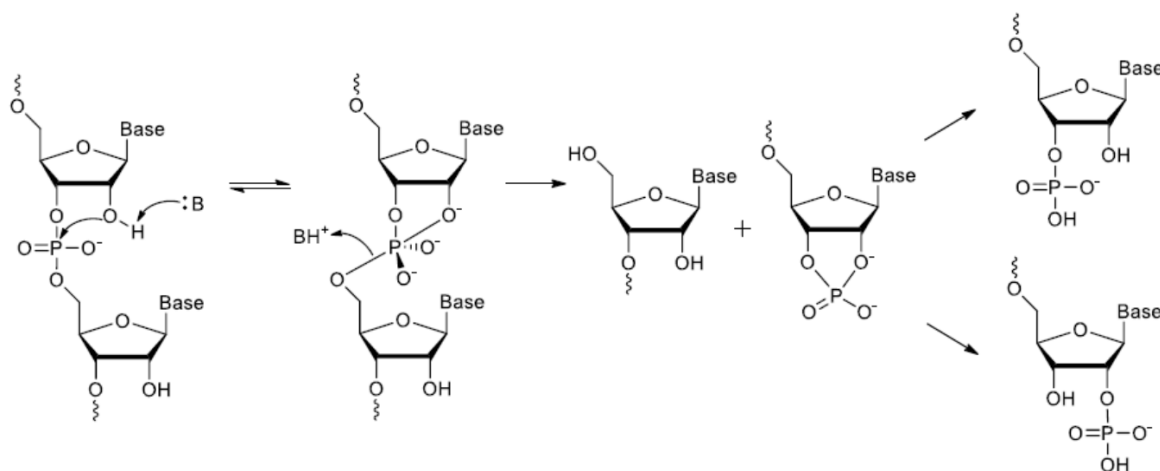
### 1.3 Stability of DNA and RNA

Phosphodiester bond is quite stable in DNA than that in RNA since RNA has 2'-hydroxyl group wherein it acts like internal nucleophile cleaving the phosphodiester bond. This stability of DNA makes it a genetic material in all living organism. It has been known that DNA is susceptible to enzymatic cleavage. Amongst various DNA cleaving enzymes, endonuclease has His-98, Asn-119 and  $Mg^{2+}$  at its active site (Figure 1.6) which is responsible for the DNA cleavage.



**Figure 1.6** Endonucleases cleave the phosphodiester bond from 3' hydroxy group

The cleavage of the phosphodiester bond in RNA is usually initiated by the 2'-hydroxyl group, because it can act as a nucleophile attacking the phosphate linkage.<sup>18</sup> The transesterification reaction proceeds through the formation of a phosphorane intermediate (or phosphorane-like transition state), characterized by a pentacoordinated phosphorus (Figure 1.7). The cyclic phosphate formed is then hydrolyzed. In studies of ribozymes and deoxyribozymes, it has been observed that both the deprotonation of the attacking nucleophile (2'-OH) and protonation of departing nucleophile (5'-OR) can accelerate the reaction by a factor of up to  $10^6$ .<sup>19</sup> Protein enzymes and ribozymes that cleave RNA catalyse via the transesterification reaction.



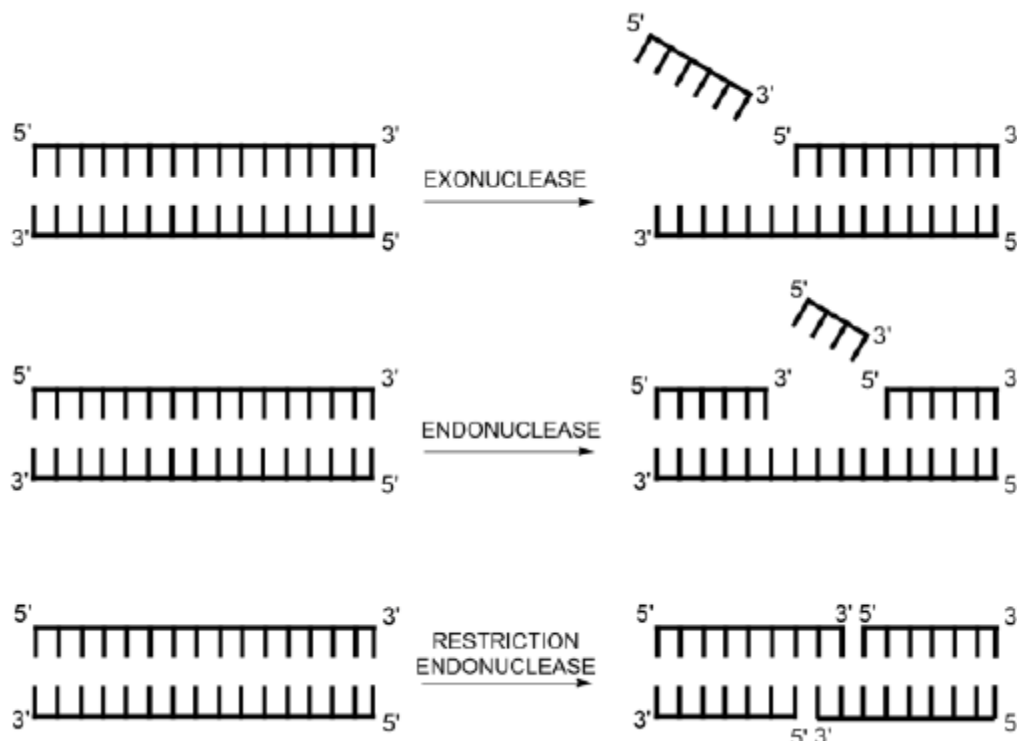
**Figure 1.7** Cleavage of phosphodiester bond in RNA.<sup>20b</sup>

This reaction proceeds through a  $SN_2$ -like mechanism, accommodating in-line conformation at the site of cleavage.<sup>20</sup> It has been established that this in-line orientation

provides a minimum of 10- to 20-fold rate enhancement for the cleavage of RNA by an intramolecular transesterification mechanism.<sup>21</sup>

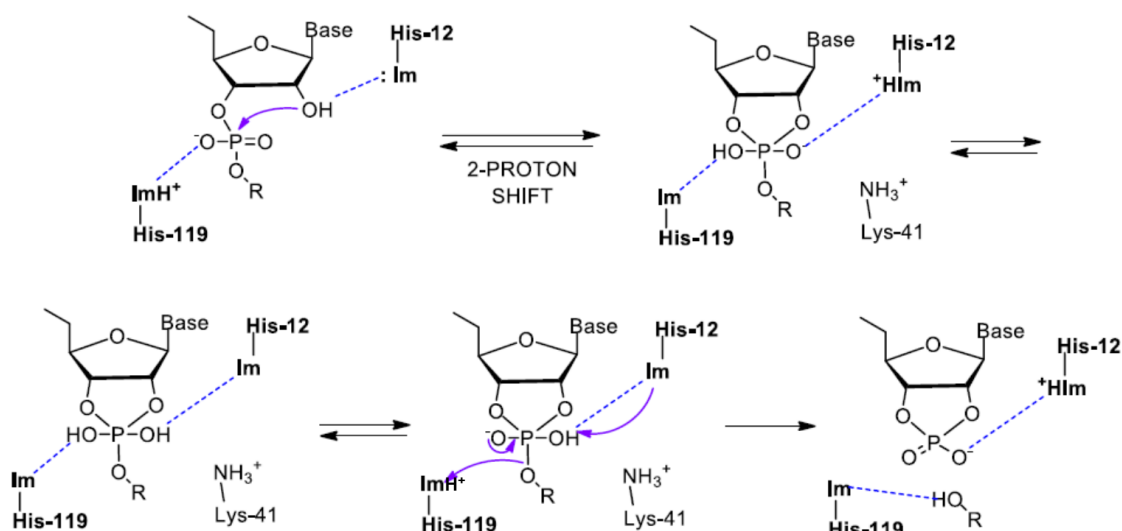
### 1.3.1 Enzymatic cleavage of RNA by nucleases

Phosphodiesterases (PDEs) are a class of enzyme that catalyze the reaction of cleavage of the phosphodiester linkage in a RNA strand, and are the major cause of the lability of the oligonucleotides *in vivo*. PDEs are divided into subclasses, I, II, and III.<sup>22</sup> Phosphodiesterases can be classified also as exonucleases or endonucleases, and in both cases the product of the cleavage is the 3'-phosphomonoester or the 5'-phosphomonoester (Figure 1.8).



**Figure 1.8** *Exo* and *endo* nucleases action on double stranded RNA.

There are three major classes of intracellular RNA-degrading enzymes (ribonucleases or RNases): *endonucleases* that cut RNA internally, *5'-exonucleases* that hydrolyze RNA from the 5'-end, and *3'-exonucleases* that degrade RNA from the 3'-end. *Endo* and *3'-exo* nucleases have been characterized in all domains of life, whereas *5'-exonucleases* until recently were believed to be absent in bacteria.<sup>23</sup>



**Figure 1.9** Mechanism of ribonuclease <sup>20b</sup>

Breslow et.al.<sup>24</sup> proposed chemical mechanism of RNA cleavage which involves a general acid-base catalysis and is promoted by two histidine residues (Figure 1.8). In RNA cleavage, the enzyme ribonuclease A acts as a bifunctional catalyst employing the imidazole groups of the residues, His-12 and His-119 as well as Lys-41, which stabilizes the negative charge on oxygen of the phosphate group. His-12 abstracts proton from 2'-OH group, while His-119 protonates 5' oxygen (Figure 1.9). Then, the stable cyclic intermediate is cleaved by hydroxyl anion and His-12 was protonated to 2'-OH group.

### 1.3.2 Artificial ribonucleases

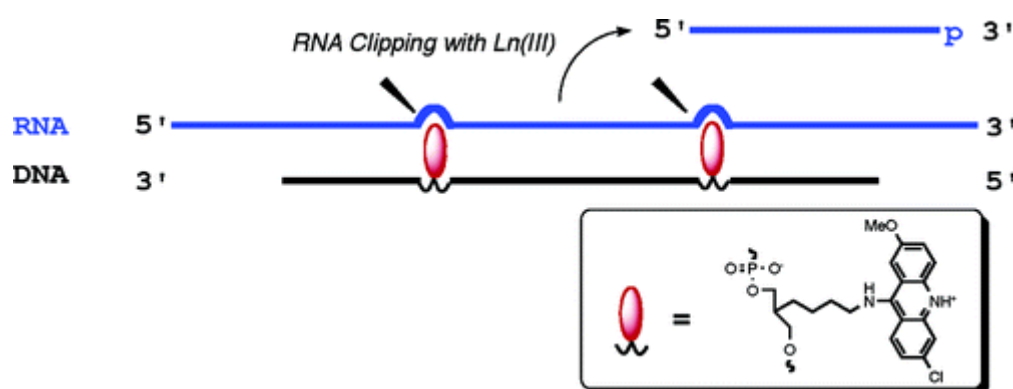
Synthetic mimics of ribonucleases have attracted considerable interest in the past few decades. These act as molecular scissors, which can cleave the ribonucleic acids at the phosphodiester linkage. In these, an oligonucleotide strand is used for sequence specific recognition of RNA substrate, and it helps to bring the molecular scissors close to the target phosphodiester bond. The oligonucleotide based artificial nucleases (OBAN) may carry different kinds of catalytic moieties. These OBAN can be divided into three different groups based on the nature of their catalytic moiety: i) lanthanide ion chelates, ii)  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  chelates, and iii) metal ion-independent artificial ribonucleases.<sup>24-28</sup>



### 1.3.2a Metal ion based artificial ribonucleases

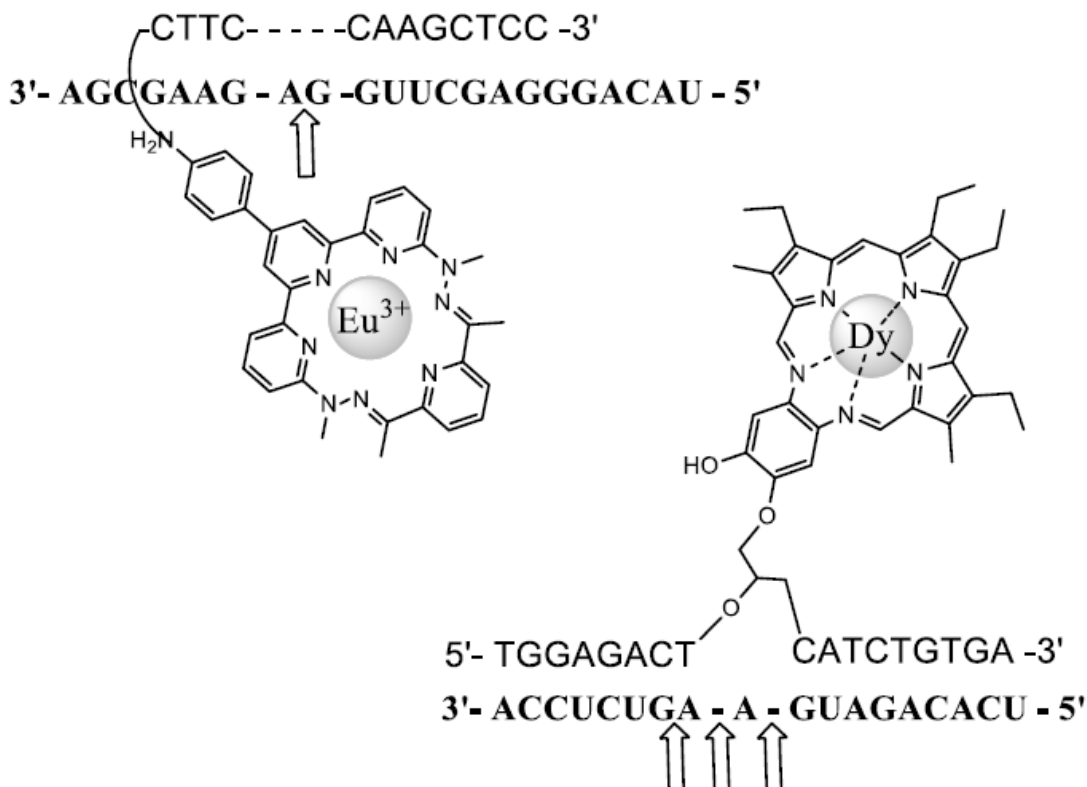
The metals present in the natural metalloenzymes are known to be involved in the catalytic mechanisms with various functions.<sup>29,30</sup> In the last few decades, many research groups have developed synthetic mimics of metalloenzymes; most of them used lanthanides or  $\text{Cu}^{2+}$  and  $\text{Zn}^{2+}$  as catalytic metal.<sup>27, 31-34</sup>

Komiyama group<sup>33-37</sup> developed systems based on lanthanide ion as RNA cleavers (Figure 1.10). The molecular scissors were usually placed at the termini and this cleaved the RNA outside the recognition sequence, that is complementary to the oligonucleotide. However, it did not give significant substrate turnover. When the acridine group was attached to the central part of the recognizing strand, it makes the target RNA more susceptible to cleavage by lanthanide ions in solution, which are close the acridine moieties, making it a successful design with high selectivity and better turnover. In the proposed mechanism, acridine pushes the unpaired ribonucleotide out of the heteroduplex changing the conformation of the RNA at the target specific for the sequence-selective activation.<sup>35</sup>



**Figure 1.10** Representation of the site-selective cleavage of RNA by acridine-DNA conjugates using  $\text{Ln}^{3+}$ .<sup>36</sup>

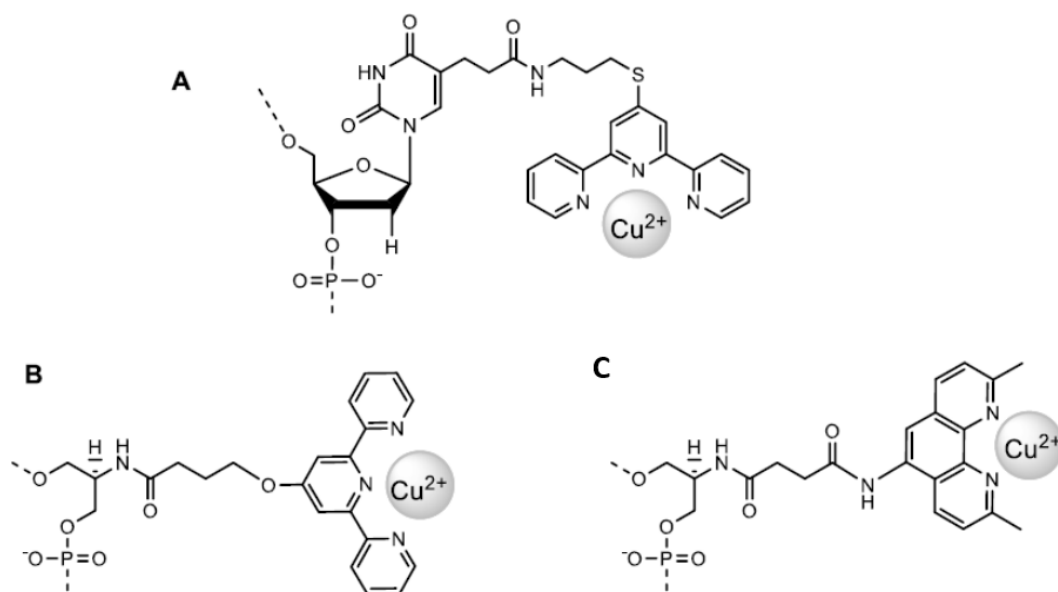
Magda *et. al.*<sup>31,32</sup> synthesized artificial nucleases with europium (EuTx) and dysprosium thexaphyrin (DyTx), which are quite active (Figure 1.11). DyTx derivative attached at the middle of the oligomer using a glyceryl linker gave good turnover of the substrate RNA. In case of EuTx, this modification was attached at the end of oligonucleotide so, it did not show turnover in RNA cleavage, presumably because the cleaved substrate was not released.<sup>38</sup>



**Figure 1.11** Lanthanide metal ion complexes and oligonucleotide based artificial nucleases<sup>20b</sup>

The cleavage of double helical DNA/RNA was not efficient using metal ion chelators; it was characterized by a strong base-stacking interaction, which is known to hamper the 5' nucleoside from taking an apical orientation. Haner *et.al.*<sup>39</sup> showed that RNA cleavage oligonucleotide lanthanide chelate based cleavers is increased at bulged-out sites. For this reason, the sequence of the oligonucleotide based artificial nuclease must be rationally designed so that a bulge or loop is formed in the target RNA strand upon hybridization with the oligonucleotide part of the OBAN.

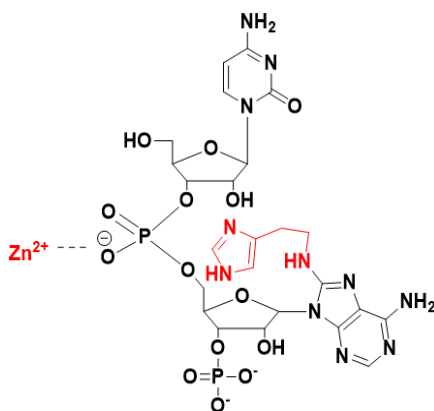
In 1994, Bashkin's group<sup>26,40,41</sup> reported for the first time a  $\text{Cu}^{2+}$  based artificial nuclease, which is based on a 17-mer oligonucleotide with a C5 terpyridine conjugated uracil base. After optimization, including the formation of a trinucleotide internal loop at the cleavage site, turnover of the target increased in excess and the half-life reduced to 40 h in pH 7.5 at 37 °C.



**Figure 1.12** Copper ion chelates used in artificial nucleases: A and B) terpyridine and C) 5-amino-2,9-dimethyl-1,10-phenanthroline.<sup>20b</sup>

Another ligand that is shown to be useful in artificial nucleases is 2,9-dimethyl-5-aminophenanthroline, which chelates with  $\text{Cu}^{2+}$  to cleave oligonucleotides. If the 2,9-dimethyl-5-aminophenanthroline is attached to serinol linker, the cleavage reaction is 5-times faster than the one catalyzed by the corresponding terpyridine conjugate (Figure 1.12 C).<sup>42</sup>

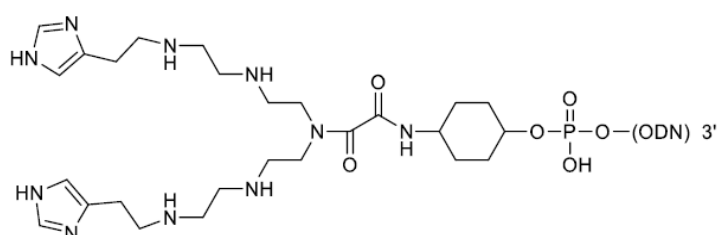
Ganesh *et.al.*<sup>43</sup> modified the ribodinucleotide CA at C8-position of adenine using histamine. This modification showed the self-cleavage properties in the dinucleotide specifically in presence of  $\text{Zn}^{2+}$  (Figure 1.13) and they studied kinetics and mechanistic of self-hydrolysis. It followed a mechanistic pathway, similar to that followed by RNAses.



**Figure 1.13** Dinucleotide self cleavage<sup>43b</sup>

### 1.3.2b Metal-free artificial ribonucleases

The major advantage of using non-metal ion-based systems for *in vivo* applications is that concerns of metal ion toxicity are precluded. Moreover, since the stability of the metal complexes and their catalytic activity are dependent on the concentration of metal ions at the target site, these systems are likely to be unreliable. In contrast, metal free approach avoids such complications and usually does not require the presence of high concentrations of any specific metal ions. Vlassov *et.al.*<sup>44,45</sup> showed a number of studies with the imidazole as cleaving agent. They synthesized different deoxyoligonucleotide derivatives attached with two histamine moieties, which can mimic histidines of the active site of RNase A (His-12 and His-119). In these systems, one imidazole acts like His-12 and the other acts like His-119. The best system was a construct where the 5'-terminus was linked to bis-imidazole, which was ultimately used to cleave a short sequence in the yeast tRNA<sup>Phe</sup> in a site-selective way, (Figure 1.14).



**Figure 1.14** Bis-imidazole cleaving system used by Vlassov *et. al.*<sup>44,45</sup>

The bis-imidazole moiety was conjugated to the end of the phosphates of the oligonucleotides (Figure 1.14) so that in the complexes they could reach two CA sequences in the tRNA, known to be susceptible to RNase A. It is important to note that the results obtained are controversial due to the natural predisposition of the target to enzymatic cleavage. Maximum cleavage (60%) of the target was obtained after 8 h at 37 °C and pH 7.0. In another example from Zenkovas group, oligoarginine-leucine peptide conjugates [G(RL)<sub>4</sub>] was used as a non-metal cleaver of oligonucleotide, which showed RNA cleavage in a catalytic, but not in a sequence-specific way.<sup>46</sup>

### 1.3.3c Free radicals based artificial nucleases

Fedorova *et.al.*<sup>47</sup> found that synthetic 2-dimethylamino-3chloro naphthoquinone and the natural anticancer quinone daunomycin could be employed in artificial nucleases. This was added to a system containing NADPH-cytochrome P-450 reductase, NADPH, ferric ions and oxygen, It generated radicals and caused single strand cleavage of supercoiled DNA. Tullius *et. al.*<sup>48</sup> showed rate of the hydroxyl radical cleavage of double-stranded DNA by the hydroxyl radical (Figure 1.15). To investigate this question, they synthesised a set of double-stranded DNA molecules in which deuterium incorporated specifically at every position in the deoxyribose of one of the four nucleotides.

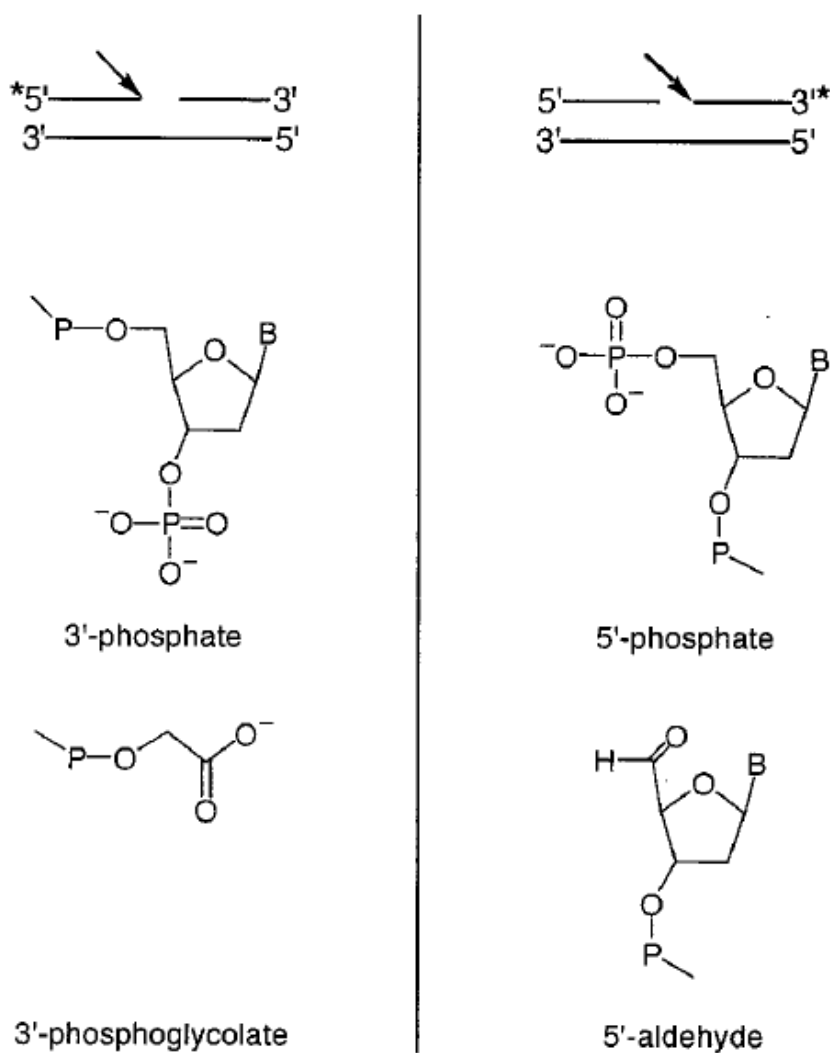
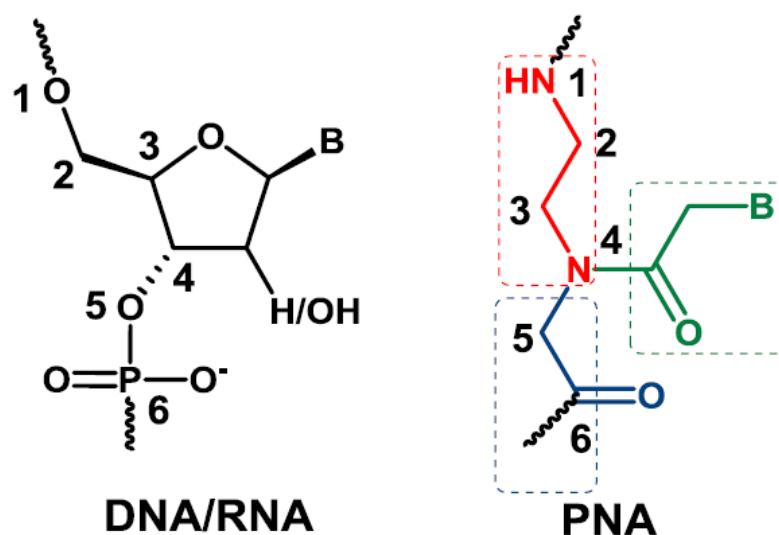


Figure 1.15 Hydroxyl radical mediated DNA cleavage.<sup>48</sup>

#### 1.4 Peptide Nucleic Acids (PNAs)

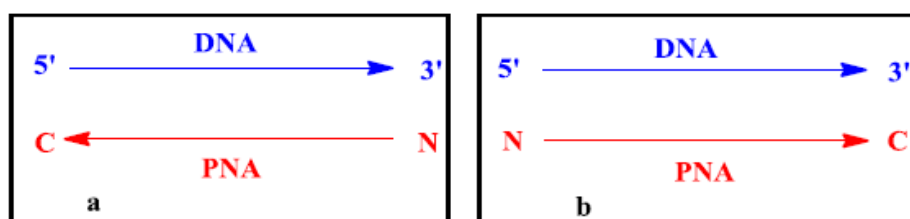
Peptide nucleic acids (PNAs) are one of the prominent classes of nucleic acid analogs pioneered by Buchardt and Nielsen in 1991.<sup>49</sup> In these DNA/RNA analogs the negatively charged sugar phosphate backbone is replaced with a pseudopeptide backbone

(2-aminoethyl-glycine unit). The nucleobase is attached to the backbone through a methylene carbonyl linker (Figure 1.16). The PNA backbone is constituted by six atoms distance for each repeating unit, which is conserved with backbone distances in DNA.<sup>50</sup> The inter nucleobase distance in PNA are the same as in DNA which allows its binding to the target DNA and RNA sequences by adhering to the Watson-Crick base pairing rules.<sup>51,52</sup> Since the PNA backbone is neutral in nature, the duplex stability of PNA:DNA is stronger than that of DNA:DNA duplexes, due to absence of interstrand electrostatic repulsion. PNAs are highly specific to their complementary sequences of DNA/RNA and it is very sensitive to mismatch in base pairing, with single mismatch resulting in a reduction of  $T_m$  by 8-20 °C.<sup>53</sup>



**Figure 1.16** Chemical structures of DNA/RNA and PNA

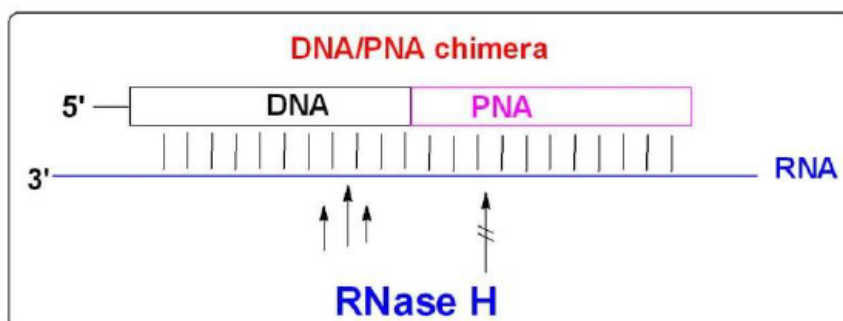
In DNA:DNA duplexes, the two strands are in an antiparallel orientation, whereas a PNA:DNA duplex can form either in antiparallel or parallel orientations (Figure 1.17), with the antiparallel orientation showing higher stability than parallel orientation.<sup>54</sup>



**Figure 1.17** Antiparallel and parallel modes of PNA:DNA duplex

PNAs being neither “true” peptides nor “true” nucleic acids, are resistant to degradation by both nucleases and proteases. Hence, they are suitable for translation

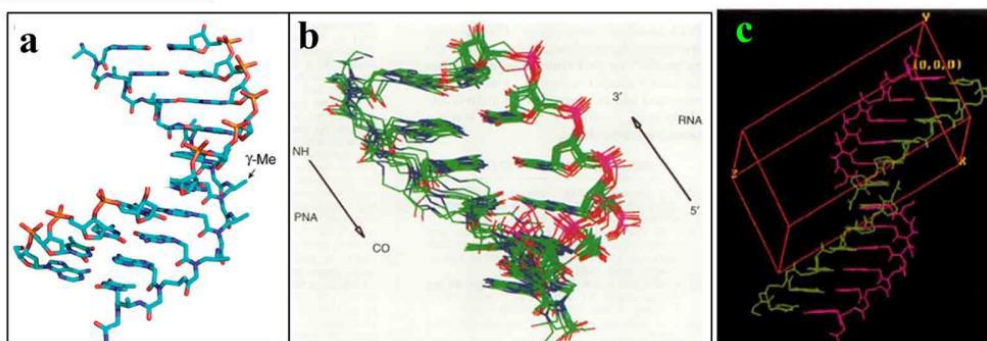
inhibition<sup>55</sup> and splicing modulation by antisense mechanisms.<sup>56-58</sup> However, the DNA/PNA chimeras are capable of stimulating RNA cleavage by RNase H *via* formation of chimeric-RNA duplexes. RNA cleavage occurs at the ribonucleotides which base pair with the DNA part of the chimera but not in the PNA part of the chimera (Figure 1.18).<sup>59</sup>



**Figure 1.18** PNA/DNA chimera

#### 1.4.1 Duplex formation with complementary DNA/RNA/PNA

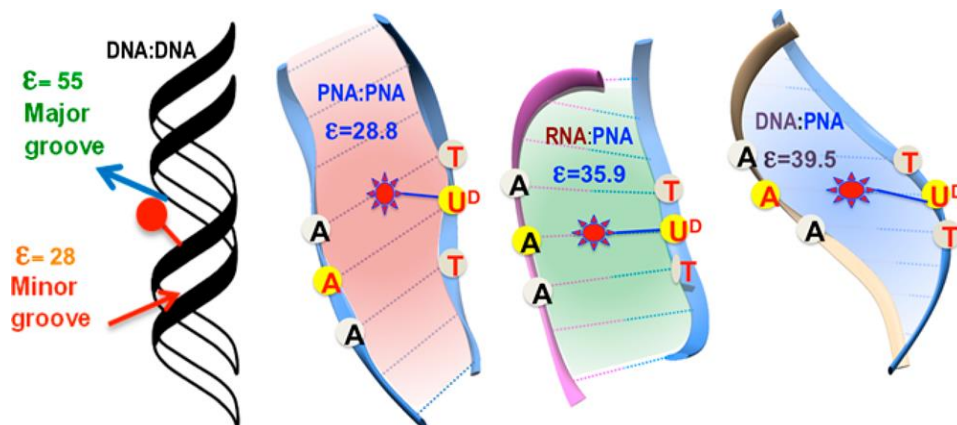
Brown *et al.*<sup>60</sup> showed the structure of a PNA:RNA duplex with nuclear magnetic resonance. A hexameric PNA (H-GAACTC-NH<sub>2</sub>) formed a 1:1 duplex with a complementary RNA [5'-r(GAGUUC)-3'] in an antiparallel orientation. This duplex formed right-handed double helix with Watson-crick base pairing (Figure 19b). Recently, Ly *et al.*<sup>61</sup> reported the crystal structure of chiral  $\gamma$ -PNA with complementary DNA strand. The crystal structure is a uniquely preorganized  $\gamma$ -PNA:DNA duplex (Figure 1.19a).



**Figure 1.19** (a) Crystal structure of  $\gamma$ -PNA:DNA duplex,<sup>61</sup> (b) NMR model for PNA:RNA duplex,<sup>60</sup> (c) Crystal structure of PNA:PNA duplex.<sup>62</sup>

The crystal structure of PNA:PNA duplex is known as P-form helix.<sup>62</sup> The P-form helix constitutes a wide PNA duplex (28 Å diameter) with a large base pair helical displacement and a pitch covering 18 base pairs. Also, in the P-form, the base pairs are placed away from the helix axis leaving a central tunnel in the helix (Figure 1.19c).

Ganesh et. al.<sup>63-65</sup> used the fluorescent probe strategy to measure the polarity of the major groove of PNA:DNA/RNA/PNA duplexes (Figure 1.20). They have revealed that in PNA:DNA/RNA/PNA duplexes, major grooves are more nonpolar compared to those DNA:DNA duplexes.



**Figure 1.20** Representation of the DNA:DNA, PNA:PNA, RNA:PNA and DNA:PNA duplexes, where one strand is modified with fluorophore.<sup>65</sup>

### 1.4.2 Triple helix formation of PNA

Polypyrimidine PNAs are able to form stable triplexes with complementary polypurine DNA, through the formation of PNA<sub>2</sub>:DNA triplexes.<sup>66</sup> The base pairing in these triplexes occurs through Watson-Crick and Hoogsteen hydrogen bonds. When only one PNA strand is used to form a PNA<sub>2</sub>:DNA triplex, both strands are necessarily either antiparallel or parallel to DNA strand. When two different homopyrimidine PNA sequences are used, the Watson-Crick PNA strand binds to the homopurine strand of DNA in antiparallel orientation, but the Hoogsteen strand binds to the DNA strand in parallel orientation to form a stable triplex. The sequence specificity of triplexes 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,<sup>66,67</sup> a property which is unique to PNAs.

### 1.4.3 Applications of PNA in Biological and material science

#### 1.4.3a Biological applications of PNA

PNAs can be used as a gene therapeutic agent because of their unique strand invasion properties, binding affinity towards the complementary DNA/RNA and their chemical as well as biological stability. Therefore, two strategies involved in using PNAs



as therapeutic drugs antigene and antisense methods. Moreover, no sign of any general toxicity of PNA has been observed so far.

### 1.4.3b PNA as antisense and antigene agents

#### A) Inhibition of transcription

PNAs are capable of stopping transcriptional processes by virtue of their ability to form a stable triplex structures, by strand invasion or by forming strand displacement complexes with DNA. If the template DNA were to be the structurally obstructed, the polymerase could be stopped, halting transcription (Figure 1.21).<sup>68</sup> PNA targeted against the promoter region can form a stable PNA/DNA duplex that stop the DNA access of the corresponding polymerase. Even an 8-mer length of PNA ( $T_8$ ) is ability to blocking phage  $T_3$  polymerase activity.<sup>69</sup>

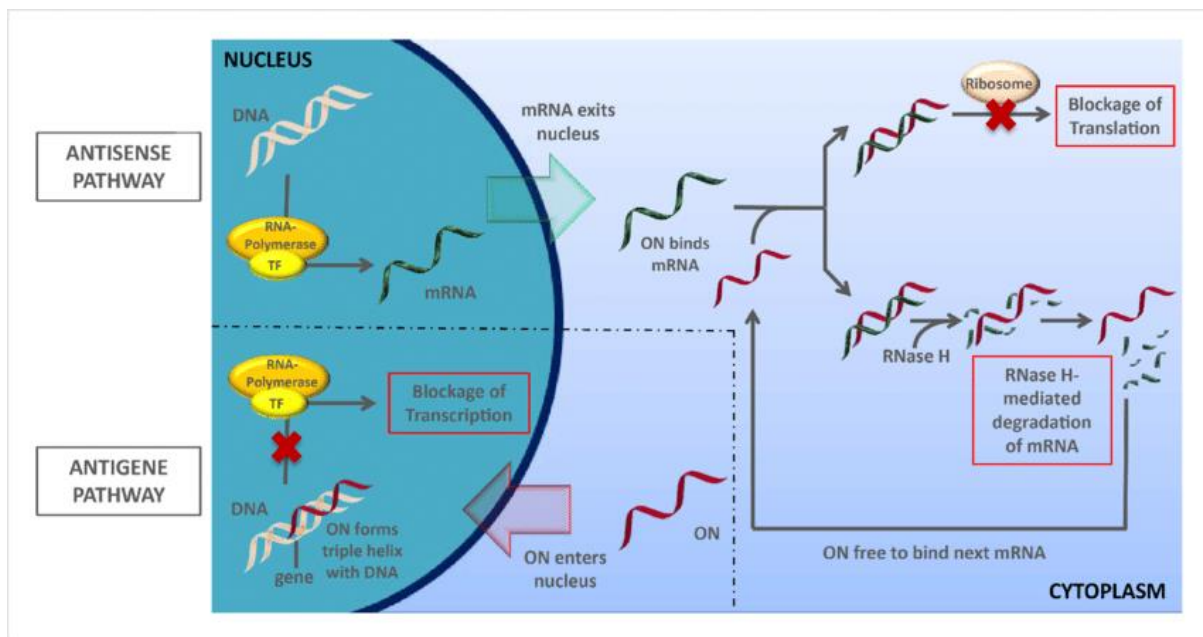


Figure 1.21 Antigene and antisense inhibition strategy<sup>70</sup>

#### B) Inhibition of translation

Antisense strategy, the nucleic acid analogs can be designed to recognize and hybridize to complementary sequences in mRNA and so, its inhibit translation (Figure 1.21). Normally, the PNA antisense is depend on the steric blocking of mRNA during translation. It has been showed *in vitro* translation experiments involving rabbit

reticulocyte lysates that both duplex-forming and triplex-forming PNAs are able to inhibiting translation at targets overlapping the AUG start codon.<sup>71</sup>

### **C) *Inhibition of replication***

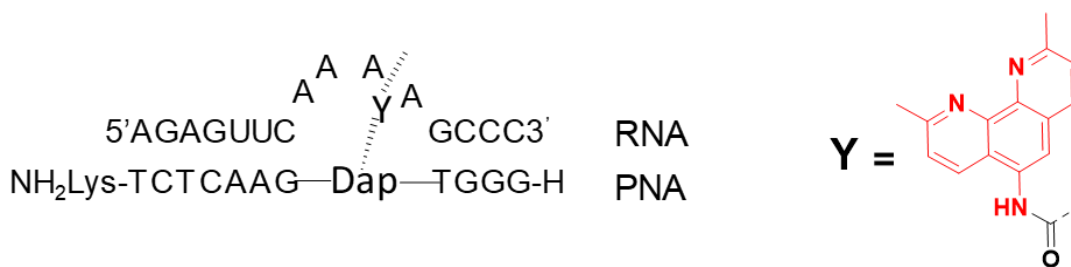
PNA can also inhibit the elongation of DNA primers by DNA polymerase. Further, the inhibition of DNA replication should be possible if the DNA duplex is subjected to strand invasion by PNA under physiological conditions, or if the DNA is single stranded during the replication process. Taylor *et al.*<sup>72</sup> has reported efficient inhibition of extra-chromosomal mitochondrial DNA. this is largely single-stranded during replication.

#### **1.4.3c *Tools for molecular biology***

Molecular genetics applications such as amplification of VNTR (variable number tandem repeat) loci for genetic typing make wide 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. Enhanced amplification of a specific VNTR product is possible by PNA. This has been done in the case of VNTR locus D1S80.<sup>73</sup> For PCR amplification, the template is stopped by a small PNA and becomes unavailable for intra- and inter-strand interactions during the re-association step. Although PNA blocks re-association, 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 has the potential of PNA application for PCR amplification where fragments of various sizes are required to be more accurately and evenly amplified.

### **A) *Artificial restriction enzyme***

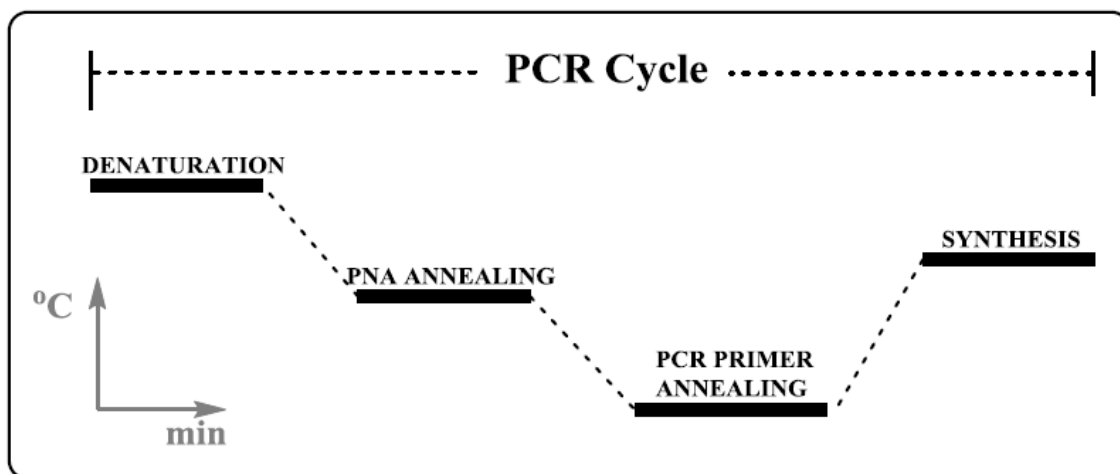
Strömberg *et al.*<sup>74</sup> have synthesized chelating ligand 5-amino-2,9-dimethyl-1,10-phenantroline attached to a diamino propionic acid unit and incorporated it in the middle of a PNA oligomer. This PNA was used for complementary RNA cleavage in the presence of Zn<sup>2+</sup> and Cu<sup>2+</sup> (Figure 1.22). A remarkable improvement in the catalytic activity was seen in the case of Cu<sup>2+</sup>; it showed site-specific cleavage and in less time. In this PNA modification, rate of RNA cleavage was the highest among the artificial nucleases till date.



**Figure 1.22** PNA as Artificial restriction enzyme

**B) Single base pair mutation analysis using PNA-directed PC clamping**

Single-base-pair mutation or single-nucleotide polymorphism (SNP) analysis is possible by the PCR technique, if PNA sequence targets the primer binding site (Figure 1.23).<sup>75</sup> 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 more than that for normal PCR primer annealing. In this condition the PNA is selectively bound to the DNA molecule. The PNA binds to the primer binding site instead of the primer. This PCR product formation is then stopped by PNA effectively. PNA is able to discriminate between single mismatch and fully complementary targets (mutations) in a mixed target PCR. Hence, out-competing PNA annealing will be favoured over binding of primer. Consequently, mutated sequences will be favourably amplified.



**Figure 1.23** representation of cyclic profile used in PNA-directing clamping<sup>75</sup>

**1.4.3d PNA based Self-Assembly**

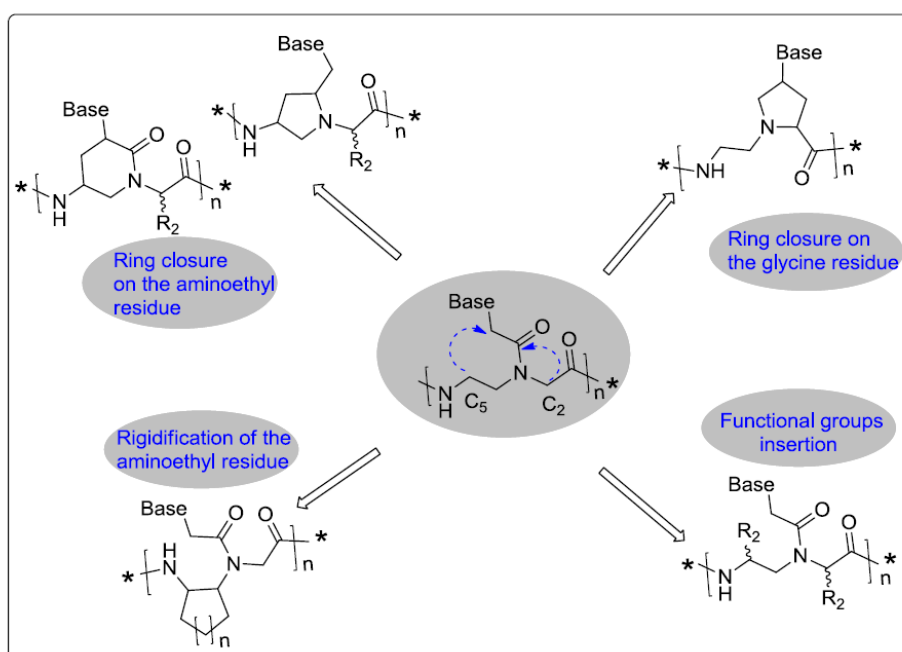
The control of supramolecular assembly of a greater number of molecules is possible with high degree of precision and predictability, when the oligonucleotides are

programmed using simple rules of Watson-Crick basepairing and duplex geometry. As PNA shares its informational blocks, the nucleobases with DNA, PNA can also be used in the same manner. With an amide backbone, the PNA offers versatility, flexibility and structural complexity.

Similar to peptides, PNA were also attached to alkyl chains to induce self-assembly in aqueous solvents. Schneider et. al.<sup>76</sup> prepared amphiphiles of 10-mer PNA and reported that it still retained its DNA hybridizing capacity. Stupp et. al.<sup>77</sup> showed the self-assembly of peptide PNA conjugate into nanofibers. Ganesh et. al.,<sup>78</sup> reported that fluorinated chain in the PNA structure thus caused compaction of the assembled particles. PNA strands attached to a perfluoroalkyl formed particles that were smaller by 2-3 fold than the PNA amphiphiles attached to hydrocarbon tail.

#### 1.4.4 Chemical modifications of PNA

The binding affinity of PNA towards DNA and RNA can be tuned via suitable conformational preorganization of PNA backbone. Being non-chiral, PNAs adapt both righthanded / lefthanded conformations with equal ease. An approach to improve DNA binding affinity is the design and synthesis of preorganized PNAs that prefer a righthanded helical conformation, which can be achieved by adding substituents to the backbone or by cyclization of the PNA backbone (Figure 1.24).<sup>79,80</sup>

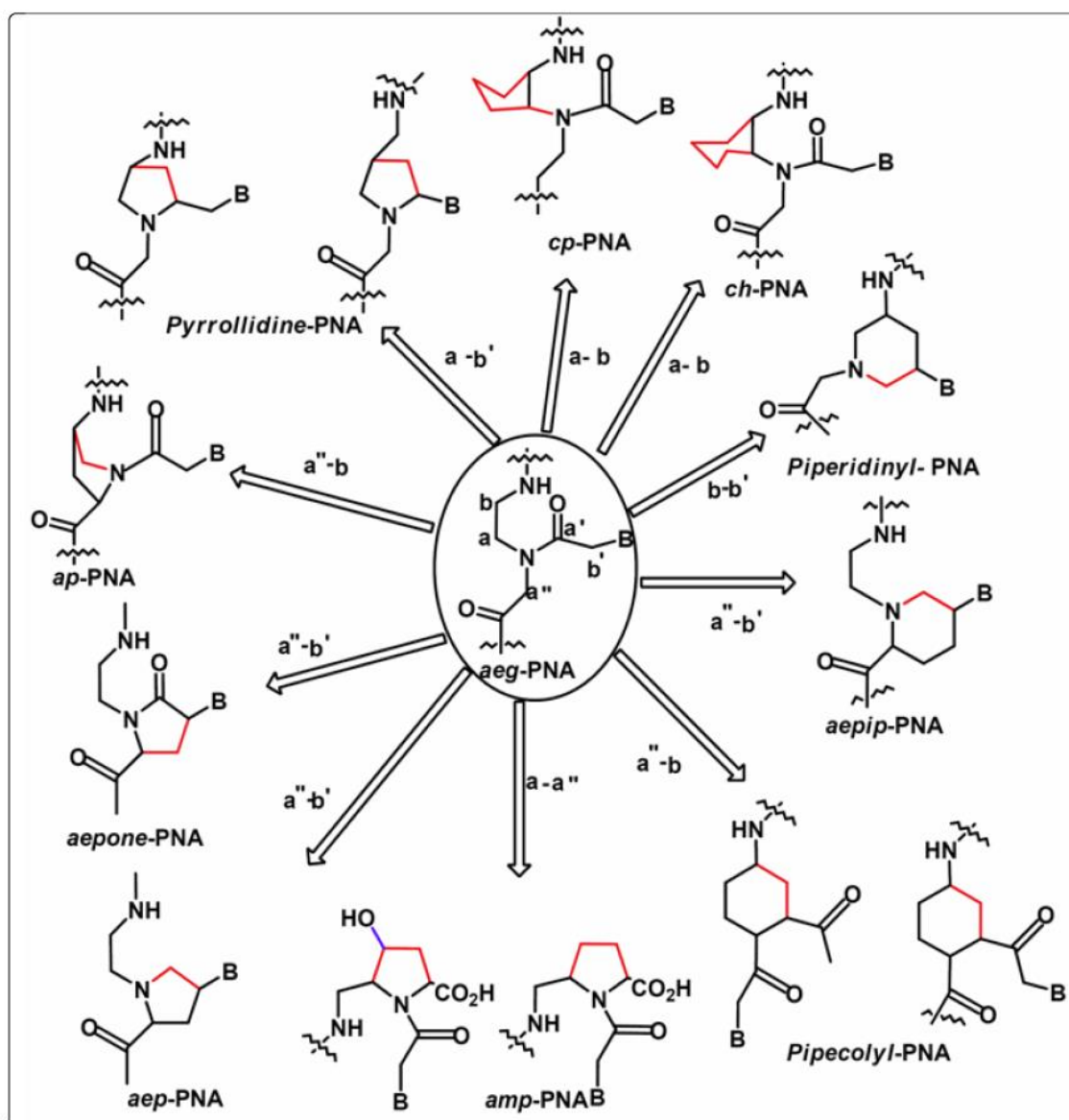


**Figure 1.24** Strategies for inducing preorganization in the PNA structure

#### 1.4.4a Preorganization through cyclic PNAs

Systematic efforts have been carried out in rational design and synthesis of conformationally constrained PNA analogues towards evolving PNAs for sequence specific recognition of DNA/RNA. Conformational preorganization has been addressed which are depend on insertion of methylene or ethylene groups to bridge the aminoethyl glycylic backbone and methylene carbonyl side chain to generate various five or the six membered nitrogen heterocyclic analogues (Figure 1.25).<sup>79</sup> The nucleobases are directly conjugated to the ring have defined nucleobase orientation in cyclic analogs, this is solving the rotamer problem. It also introduces chiral centres, which may information of directional sequence specific of PNA with chiral DNA/RNA.

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  $\alpha$ -carbon atom of the glycine unit and the  $\beta$ '-carbon atom of the nucleobase linker were attached through a methylene bridge.<sup>81</sup> The flexibility in the aminoethyl segment of *aeg* PNA was retained. The nucleobase conjugated to the pyrrolidine ring was fixed by virtue of the chirality of C-4 therefore, possibility of any rotameric populations was removed. The tertiary amine group in the backbone was found to be at less partially protonated at physiological pH ( $pK_a \sim 6.8$ ). So, 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 more favourable binding properties towards the target sequences of PNA/DNA without compromising the specificity.

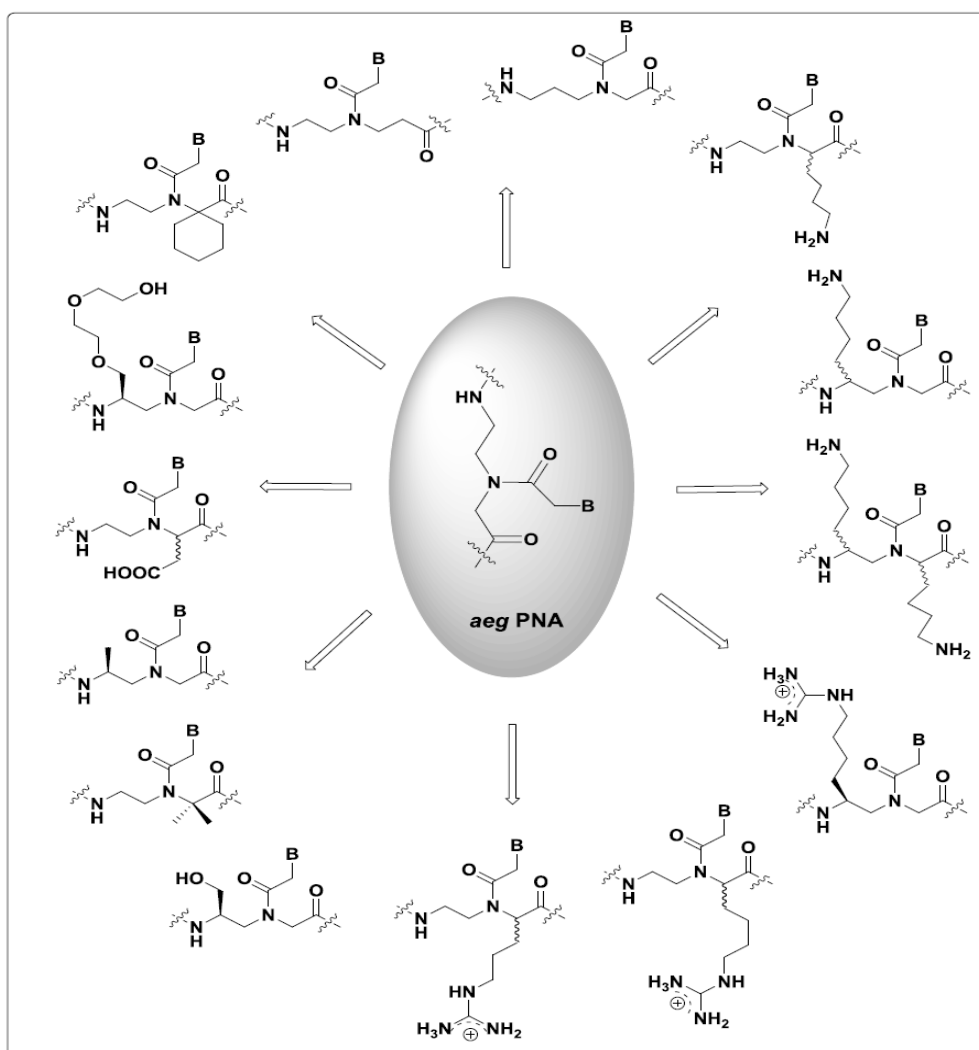


**Figure 1.25** Cyclic and conformationally constrained PNA analogs<sup>79</sup>

Insertion of a methylene bridge between the  $\beta$ -carbon atom of the aminoethyl part and the  $\alpha''$ -carbon of the glycine part of *aeg*-PNA resulted in 4-aminopropyl (*ap*) PNA having two chiral centers.<sup>82</sup> None of the homochiral aminopropyl thyminyi PNAs corresponding to any of the diastereomers form the complex with sequence specific DNA sequences,<sup>83</sup> which due to high rigidity in the backbone showed structural incompatibility. However, insertion of single chiral D-trans or L-trans propyl PNA monomer into *aeg*-PNA at the N-terminus or within the PNA sequence showed higher binding affinity towards DNA with specific parallel or an antiparallel mode unlike the unmodified PNA.<sup>84</sup>

### 1.4.4b Preorganization through acyclic PNAs

L/D-amino acid synthons have been used to synthesize C2-substituted and C5-substituted chiral PNAs with desired stereochemistry (Figure 1.26). Chiral C2-substituted PNAs derived from alanine,<sup>85</sup> arginine and lysine<sup>86</sup> side chains showed increased PNA:DNA stability due to the small steric hindrance and the electrostatic interaction with negatively charged phosphates in DNA. The introduction of stereogenic centres in the PNA strand showed a predominant helix handedness.



**Figure 1.26** Various modified acyclic PNAs

As revealed from CD spectroscopy, PNAs containing D-amino acid derived monomers with the stereogenic center at C2-position (*R* stereochemistry) forms right-handed conformation in PNA-PNA complex, whereas PNAs containing L-amino acid derived monomers with the stereogenic center in the same position (*S* stereochemistry) forms 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 and it has a greater DNA binding affinity than their mirror images.<sup>87,88</sup>

To address the issue of poor cell permeability of *aeg* PNA, Ly *et al.*<sup>89</sup> synthesized guanidinium group at C2-position of PNA backbone ( $\alpha$ -GPNA) (Figure 1.27 a). Thermal stability studies showed that although the guanidinium group of  $\alpha$ -GPNA is positively charged at physiological pH, incorporation of one modified unit into a decamer PNA sequence destabilized the duplex by 2-4 °C. In addition, incorporation of multiple  $\alpha$ -D-GPNA units at alternate positions in the sequence increased the stability of PNA:DNA duplex in antiparallel orientation.<sup>90</sup> Cellular uptake of fully modified  $\alpha$ -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.<sup>89</sup> Cell penetration of  $\alpha$ -GPNA especially in ES cells was very important because these cells are extremely difficult to transduce, even with the best transfecting agents.<sup>91</sup>

Ly *et al.*<sup>92</sup> have synthesized the second generation guanidino modified PNA analogs ( $\gamma$ -L-GPNA) based on the homo-arginine side chain at  $\gamma$ -position of PNA backbone (Figure 1.27 b). The PNA:DNA duplex stability was found to be superior for  $\gamma$ -L-GPNA compared to  $\alpha$ -GPNA and Cell permeation studies showed that  $\gamma$ -L-GPNAs are easily taken up by HeLa cells and the uptake efficiency was comparable to that of the TAT transduction domain.

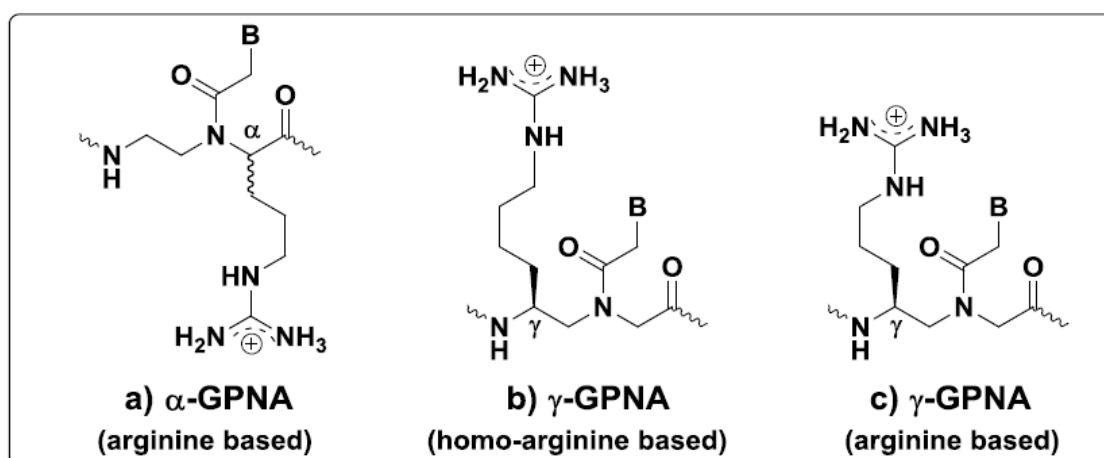


Figure 1.27 guanidino modified  $\alpha$ - and  $\gamma$ -GPNA analogs<sup>92</sup>



Manicardi *et al.*<sup>93</sup> have reported the inhibition of micro-RNA by GPNA based on arginine side chain (Figure 1.27 c). Anti-miR-210 activity of several 18-mer PNAs (*aeg* PNA,  $\alpha$ -GPNA and  $\gamma$ -GPNA) in leukemic K562 cells was tested. The incorporation of GPNA was either alternating in the sequence or consecutive at the 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  $\gamma$ -GPNA with consecutive placement.

Many backbone modified PNAs have been reported, some of these modifications were designed, synthesized and characterized in this laboratory. Ganesh *et al.*<sup>94,95</sup> reported the design and synthesis of backbone modified chiral C $\alpha$ - and C $\gamma$ -aminoethyl PNAs (*am*-PNAs) with substitutions in the PNA backbone (Figure 1.28 A and B). The *am*-PNAs formed stable PNA:DNA complex and the order of stabilization was,  $\gamma$ -(*S*)-*am*-PNA >  $\alpha$ -(*R*)-*am*-PNA >  $\alpha$ -(*S*)-*am*-PNA. The *am*-PNAs are taken up by HeLa cells, with the decreasing order of uptake efficiency as  $\gamma$ -(*S*)-*am*-PNA >  $\alpha$ -(*R*)-*am*-PNA >  $\alpha$ -(*S*)-*am*-PNA.

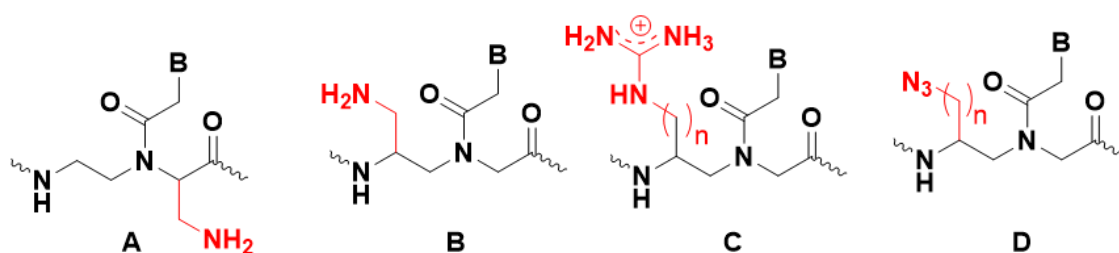


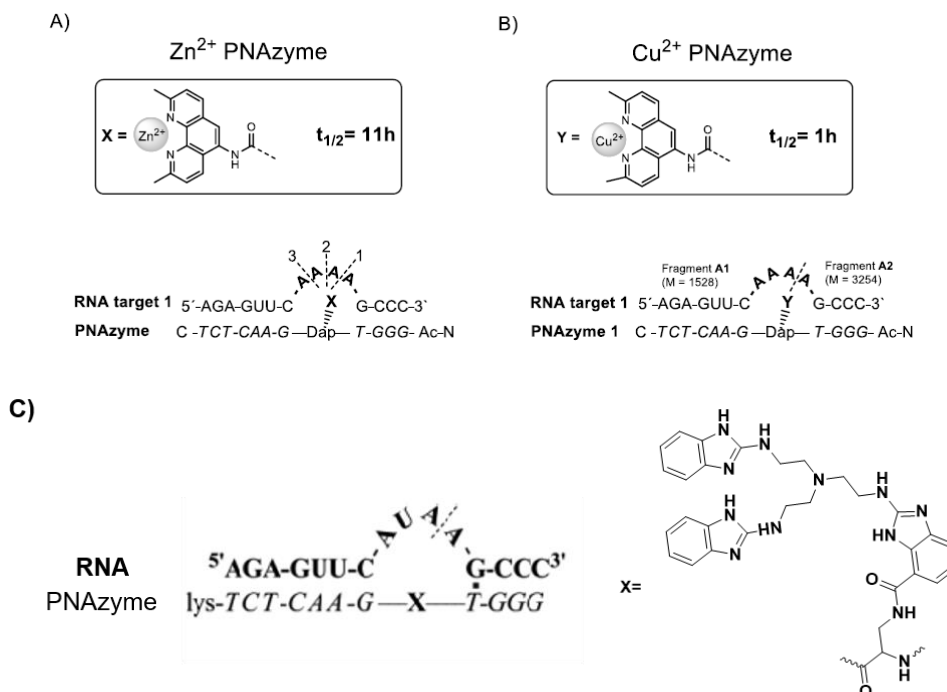
Figure 1.28 Modified acyclic PNAs<sup>95b</sup>

## 1.5 PNA as Artificial nucleases

As mentioned in the introduction, PNA has a neutral N-(2-aminoethyl)-glycine backbone, and nucleobases are attached through a methylene linker. It forms duplexes with complementary RNA/DNA more strongly and form stable PNA<sub>2</sub>:DNA/RNA triplexes with complementary RNA/DNA. All these features make PNA a suitable candidate for the further development of the artificial nucleases.

Strömberg *et al.*<sup>74</sup> used chelating ligand 5-amino-2,9-dimethyl-1,10-phenantroline attached to a diamino propionic acid unit and incorporated it in the middle of a PNA oligomer. This PNA was used for complementary RNA cleavage in the presence of Zn<sup>2+</sup> (Figure 1.29 A) and Cu<sup>2+</sup> (Figure 1.29 B). A remarkable improvement in the catalytic

activity was seen in the case of  $\text{Cu}^{2+}$ ; it showed site-specific cleavage and in less time. In this PNA modification, rate of RNA cleavage was the highest among the artificial nucleases till date.



**Figure 1.29** Artificial PNAzymes<sup>20b</sup>

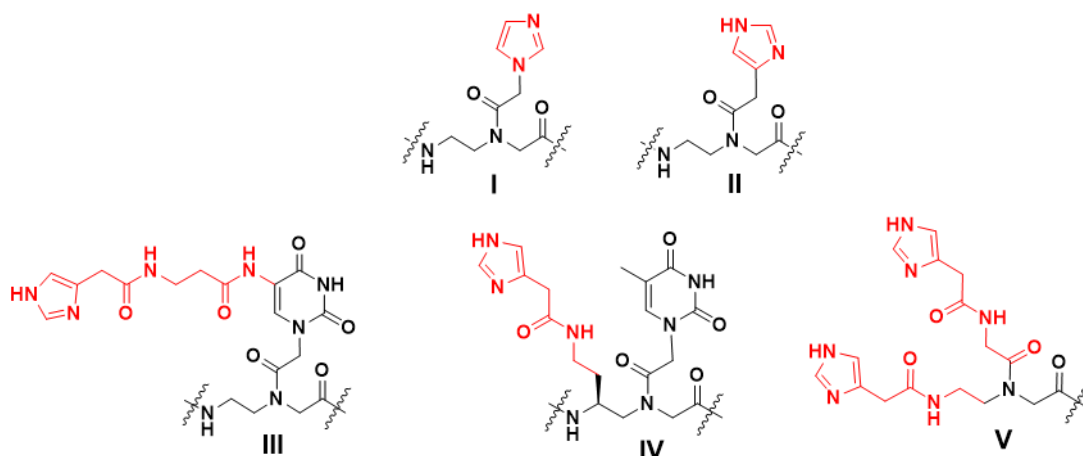
Göbel *et. al.*<sup>96</sup> showed RNA cleavage by tris(2-aminobenzimidazole)-PNA (Figure 1.29 C), by non-metal-based artificial nucleases which hydrolyzed RNA at specific manner at the bulge site in the middle of the oligomer. This PNAzyme showed slightly lower rate of RNA hydrolysis.

## 1.6 Scope of present work

This thesis deals with studies on PNAs modified with catalytic imidazole moiety towards creating nucleases mimics. The work involves design, synthesis and characterization of PNAs conjugated with imidazoles at different positions and imidazole substituted polyproline peptides with a focus on their ability to catalyze hydrolysis of RNA.

## Chapter 2: Synthesis of *aeg*-Imidazole Monomers and Derived PNA Oligomers

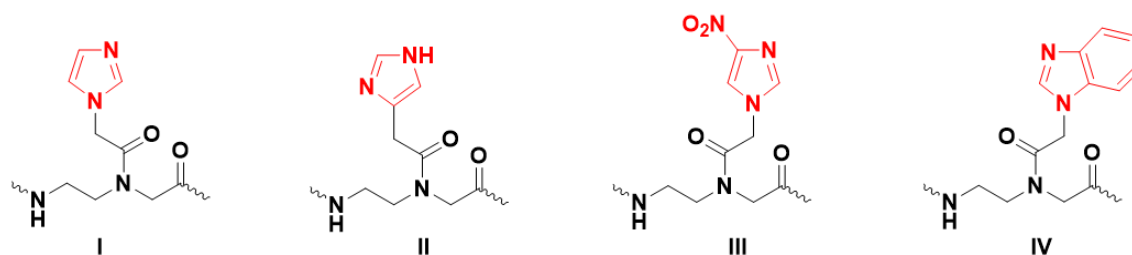
This chapter describes rational design, and synthesis of novel imidazole modified PNA monomers and their incorporation into PNA oligomers (Figure 1.30) using solid phase synthesis to obtain prospective imidazole based artificial nucleases.



**Figure 1.30** Designed Imidazole modified PNA oligomers

## Chapter 3: Biophysical Study of Imidazolyl PNA

This chapter involves studies on probing the hybridization properties of imidazole substituted PNAs with complementary DNA or PNA using UV- $T_m$  analysis and circular dichroism techniques. The PNAs, where one of the nucleobases was replaced by N1- or C4- imidazole groups (Figure 1.31) were studied. It examines whether imidazole can behave as a universal base like 3-nitropyrrole, in addition to exploring base-pairing preference of imidazole towards the natural nucleobases. Although imidazoles did not show universal base pairing properties, the results of thermal melting studies with PNA:DNA and PNA:PNA duplexes using temperature dependent UV absorption spectroscopy revealed the role of backbone in determining the ability of imidazole to distinguish between purine and pyrimidine nucleobases in the complementary strand.



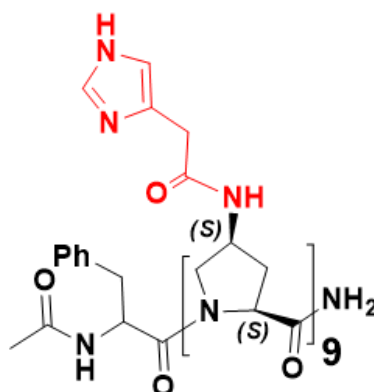
**Figure 1.31** Imidazole modified PNA oligomers

## Chapter: 4 Metal Complexes, Biophysical and RNA Cleavage Studies

This chapter investigates the ability of the imidazole modified PNAs to hydrolyse complementary RNA site-selectively, using HPLC. PNA oligomers presenting imidazoles at different sites on the structure were tested against complementary RNA in three different architectures: PNA/RNA/RNA triplexes, PNA/RNA duplexes, and PNA/RNA duplexes with a bulge. Their preference for binding to specific metal ions were also studied to further examine whether complexation with metal ions confers nuclease like activity on the PNAs. The results showed that these rationally designed imidazolyl PNAs were unable to exhibit any notable hydrolysis activity against RNA strands indicating the need for further optimization of the structure of imidazolyl PNA analogs in order to employ them as artificial nucleases.

## Chapter 5: Design and Synthesis of Imidazole Polyproline Peptides Nuclease Mimics

In continuation with the goal of previous chapters to create artificial ribonucleases, wherein imidazole units were placed on nucleic acid templates, herein, the rationale was to functionalise the C4 position of prolines in a polyprolines with imidazole (Figure 1.32) for imparting nucleolytic properties for hydrolysing RNA strands.



**Figure 1.32** Imidazole modified PNA oligomers

In summary, this thesis attempts the synthesis of different rationally designed imidazole modified PNA analogues and examines the effect of the imidazole substitutions on the base-pairing properties. The preference for specific metal ions to bind different imidazole moieties and their effects on possible catalytic properties to drive target-specific and site-specific RNA hydrolysis is explored. The ability of imidazoles anchored on polyproline peptides for possible nuclease activity is also explored.

## 1.7 References

1. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Wlaser, P. *Molecular Biology of the Cell* (4<sup>th</sup> Ed.). Garland Science. **2002**, pp. 120-121.
2. Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, **1953**, *171*, 737-738.
3. Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta. Crystal.* **1963**, *16*, 907-916.
4. (a) Crick, F. H. C. *J. Mol. Biol.* Codon—anticodon pairing: The wobble hypothesis. **1966**, *19*, 548-555. (b) Soll, D.; Cherayil, J. D.; Bock, R. M. Studies on polynucleotides: LXXV. Specificity of tRNA for codon recognition as studied by the ribosomal binding technique. *J. Mol. Biol.* **1967**, *29*, 97-112.
5. Benner, S. A. Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res.* **2004**, *37*, 784–797.
6. Henry, A. A.; Romesberg, F. E. Beyond A, C, G and T: augmenting nature's alphabet. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727–733.
7. Krueger, A. T.; Kool, E. T. Redesigning the architecture of the base pair: toward biochemical and biological function of new genetic sets. *Chem. Biol.* **2009**, *16*, 242–248.
8. Hirao, I. Unnatural base pair systems for DNA/RNA-based biotechnology. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622–627.
9. Kimoto, M.; Cox, R. S., 3rd; Hirao, I. Unnatural base pair systems for sensing and diagnostic applications. *Expert Rev. Mol. Diagn.* **2011**, *11*, 321–331.
10. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **1990**, *343*, 33–37.

11. Morales, J. C.; Kool, E. T. Efficient replication between non-hydrogen-bonded nucleoside shape analogs. *Nat. Struct. Biol.* **1998**, *5*, 950–954.
12. McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly
13. Ohtsuki, T.; Kimoto, M.; Ishikawa, M.; Mitsui, T.; Hirao, I.; Yokoyama, S. Unnatural base pairs for specific transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4922–4925.
14. Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Nishikawa, K.; Yokoyama, S. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **2002**, *20*, 177–182.
15. Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. An unnatural hydrophobic base pair with shape complementarity between pyrrole-2-carbaldehyde and 9-methylimidazo[(4,5)-b]pyridine. *J. Am. Chem. Soc.* **2003**, *125*, 5298–5307.
16. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **2006**, *3*, 729–735.
17. Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. An efficient unnatural base pair for PCR amplification. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555.
18. Brown, D. M.; Todd, A. R. Nucleic acids. *Annu. Rev. Biochem.* **1955**, *24*, 311-38.
19. Emilsson, G. M.; Nakamura, S.; Roth, A.; Breaker, R. R. Ribozyme speed limits. *RNA* **2003**, *9*, 907-918.
20. (a) Kuimelis, R. G.; McLaughlin, L. W. Mechanisms of ribozyme-mediated RNA cleavage. *Chem. Rev.* **1998**, *98*, 1027-1044. (b) Ghidini, A. Ph.D thesis, Karolinska Institutet, **2015**.
21. Soukup, G. A.; Breaker, R. R. Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* **1999**, *5*, 1308-1325.
22. Omori, K.; Kotera, J. Overview of PDEs and their regulation. *Circ. Res.* **2007**, *100*, 309-327.
23. Houseley, J.; Tollervey, D. The many pathways of RNA degradation. *Cell*, **2009**, *136*, 763-776.

24. Breslow, R.; Chapman, W. H., Jr., On the mechanism of action of ribonuclease A: relevance of enzymatic studies with a p-nitrophenylphosphate ester and a thiophosphate ester. *Proc Natl Acad Sci U S A* **1996**, *93*, 10018-21.
25. Niittymäki, T.; Lonnberg, H. Artificial ribonucleases. *Organic & biomolecular chemistry* **2006**, *4*, 15-25.
26. Trawick, B. N.; Daniher, A. T.; Bashkin, J. K. Inorganic Mimics of Ribonucleases and Ribozymes: From Random Cleavage to Sequence-Specific Chemistry to Catalytic Antisense Drugs. *Chem. Rev.* **1998**, *98*, 939-960.
27. Komiyama, M.; Sumaoka, J.; Kuzuya, A.; Yamamoto, Y. Sequence-selective artificial ribonucleases. *Methods Enzymol.* **2001**, *341*, (Ribonucleases, Part A), 455-468.
28. Kuzuya, A.; Komiyama, M. Site-selective artificial ribonucleases and their applications. *Curr. Org. Chem.* **2007**, *11*, 1450-1459.
29. Liu, C. T.; Neverov, A. A.; Maxwell, C. I.; Brown, R. S. Demonstration of Prominent Cu(II)-Promoted Leaving Group Stabilization of the Cleavage of a Homologous Set of Phosphate Mono-, Di-, and Triesters in Methanol. *J. Am. Chem. Soc.* **2010**, *132*, 3561-3573.
30. Williams, N. H.; Takasaki, B.; Wall, M.; Chin, J. Structure and Nuclease Activity of Simple Dinuclear Metal Complexes: Quantitative Dissection of the Role of Metal Ions. *Acc. Chem. Res.* **1999**, *32*, 85-493. 57. Magda, D.; Crofts, S.; Lin, A.; Miles, D.; Wright, M.; Sessler, J. L., Synthesis and Kinetic Properties of Ribozyme Analogs Prepared Using Phosphoramidite Derivatives of Dysprosium(III) Texaphyrin. *J. Am. Chem. Soc.* **1997**, *119*, 2293-2294.
31. Magda, D.; Crofts, S.; Lin, A.; Miles, D.; Wright, M.; Sessler, J. L. Synthesis and Kinetic Properties of Ribozyme Analogs Prepared Using Phosphoramidite Derivatives of Dysprosium(III) Texaphyrin. *J. Am. Chem. Soc.* **1997**, *119*, 2293-2294.
32. Magda, D.; Miller, R. A.; Sessler, J. L.; Iverson, B. L. Site-Specific Hydrolysis of RNA by Europium(III) Texaphyrin Conjugated to a Synthetic Oligodeoxyribonucleotide. *J. Am. Chem. Soc.* **1994**, *116*, 7439-40.
33. Komiyama, M. Sequence-selective scission of DNA and RNA by lanthanide ions and their complexes. *Met. Ions Biol. Syst.* **2003**, *40*, (Lanthanides and Their Interrelations with Biosystems), 463-475.

34. Kuzuya, A.; Machida, K.; Mizoguchi, R.; Komiyama, M. Conjugation of Various Acridines to DNA for Site-Selective RNA Scission by Lanthanide Ion. *Bioconjugate Chem.* **2002**, *13*, 365-369.
35. Kuzuya, A.; Mizoguchi, R.; Morisawa, F.; Machida, K.; Komiyama, M. Metal Ion-Induced Site-Selective RNA Hydrolysis by Use of Acridine-Bearing Oligonucleotide as Cofactor. *J. Am. Chem. Soc.* **2002**, *124*, 6887-6894.
36. Kuzuya, A.; Mizoguchi, R.; Sasayama, T.; Zhou, J.-M.; Komiyama, M. Selective Activation of Two Sites in RNA by Acridine-Bearing Oligonucleotides for Clipping of Designated RNA Fragments. *J. Am. Chem. Soc.* **2004**, *126*, 1430-1436.
37. Kuzuya, A.; Shi, Y.; Tanaka, K.; Machida, K.; Komiyama, M. Efficient Site-selective RNA Activation and Scission Achieved by Geometry Control of Acridine Intercalation in RNA/DNA Heteroduplex. *Chemistry Letters* **2009**, *38*, 432-433.
38. Magda, D.; Wright, M.; Crofts, S.; Lin, A.; Sessler, J. L. Metal Complex Conjugate of Antisense DNA Which Displays Ribozyme-Like Activity. *J. Am. Chem. Soc.* **1997**, *119*, 6947-6948.
39. Hall, J.; Huesken, D.; Haener, R., Towards artificial ribonucleases: the sequence-specific cleavage of RNA in a duplex. *Nucleic Acids Res.* **1996**, *24*, 3522-3526. 68
40. Bashkin, J. K.; Frolova, E. I.; Sampath, U. Sequence-Specific Cleavage of HIV mRNA by a Ribozyme Mimic. *J. Am. Chem. Soc.* **1994**, *116*, 5981-2.
41. Trawick, B. N.; Osiek, T. A.; Bashkin, J. K., Enhancing sequence-specific cleavage of RNA within a duplex region: Incorporation of 1,3-propanediol linkers into oligonucleotide conjugates of serinol-terpyridine. *Bioconjugate Chem.* **2001**, *12*, 900-905.
42. Putnam, W. C.; Daniher, A. T.; Trawick, B. N.; Bashkin, J. K., Efficient new ribozyme mimics: direct mapping of molecular design principles from small molecules to macromolecular, biomimetic catalysts. *Nucleic Acids Res.* **2001**, *29*, 2199-2204.
43. Prakash, T. P.; Kunte, S. S.; Ganesh, K. N. Self-cleavage of C8-histamino-r(UpA) promoted by ZnCl<sub>2</sub>: mechanistic studies on a designed ribonuclease mimic. *Tetrahedron* **1994**, *50*, 11699–11708. (b) Prakash, T. P.; Ganesh, K. N. Ribonuclease mimic: Zn<sup>2+</sup> promoted cleavage of C8-histamino-r(UpA) proceeds through 2',3'-cUMP as intermediate. *J. Chem. Soc., Chem. Commun.* **1994**, 1357–1358.



44. Vlassov, V.; Abramova, T.; Godovikova, T.; Giege, R.; Silnikov, V. Sequence-specific cleavage of yeast tRNAP<sup>He</sup> with oligonucleotides conjugated to a diimidazole construct. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 39-42.
45. Vlassov, V. V.; Vlassov, A. V., Cleavage of RNA by imidazole. *Nucleic Acids Mol. Biol.* **2004**, *13*, 49-60.
46. Mironova, N. L.; Pyshnyi, D. V.; Shtadler, D. V.; Fedorova, A. A.; Vlassov, V. V.; Zenkova, M. A. RNase T1 mimicking artificial ribonuclease. *Nucleic Acids Res.* **2007**, *35*, 2356-2367.
47. Rumyantseva, G. V.; Weiner, L. M.; Frolova E. I.; Fedorova, O.S.; Hydroxyl radical generation and DNA strand scission mediated by natural anticancer and synthetic quinones. **1989**, *242*, 397-400,
48. Balasubramanian, B.; Pogozelski, W. K.; Tullius, D.T. DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone *Proc. Natl. Acad. Sci. USA.* **1998**, *95*, 9738–9743.
49. Nielsen, P. E.; Egholm, M.; Berg, R.H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*, 1497-1500.
50. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. Peptide nucleic acids (PNA). Oligonucleotide analogs with an achiral peptide backbone. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.
51. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund,  
A. Ionic Effects on the Stability and Conformation of Peptide Nucleic Acid Complexes. *J. Am. Chem. Soc.* **1996**, *118*, 5544.
52. Knudsen, H.; Nielsen, P. E. Antisense properties of duplex- and triplex-forming PNAs. *Nucleic Acid Res.* **1996**, *24*, 494.
53. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordon, B.; Nielsen, P. E. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993**, *365*, 566.
54. Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Rytte, A. Synthesis and Properties of PNA/DNA Chimeras *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2632- 2635.

55. Nulf, C. J.; Corey, D. Intracellular inhibition of hepatitis C virus (HCV) internal ribosomal entry site (IRES)-dependent translation by peptide nucleic acids (PNAs) and locked nucleic acids (LNAs). *Nucleic Acids Res.* **2004**, *32*, 3792-3798.
56. Sazani, P.; Gemignani, F.; Kang, S. H.; Maier, M. A.; Manoharan, M. Systemically delivered antisense oligomers upregulate gene expression in mouse tissues. *Nat. Biotechnol.* **2002**, *20*, 1228-1233.
57. Siwkowski, A. M.; Malik, L.; Esau, C.C.; Maier, M. A.; Wancewicz, E.V. Identification and functional validation of PNAs that inhibit murine CD40 expression by redirection of splicing. *Nucleic Acids Res.* **2004**, *32*, 2695-2706.
58. Abes, S.; Turner, J. J.; Ivanova, G. D.; Owen, D.; Williams, D. Efficient splicing correction by PNA conjugation to an R6-Penetratin delivery peptide. *Nucleic Acids Res.* **2007**, *35*, 4495-4502.
59. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. PNA: Synthetic Polyamide Nucleic Acids with Unusual Binding Properties. *Angew. Chem. Int. Ed.* **1998**, *37*, 2796-2823.
60. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. NMR solution structure of a peptide nucleic acid complexed with RNA. *Science* **1994**, *265*, 777-780.
61. Yeh, J.I.; Boris Shivachev, B.; Rapireddy, S.; Crawford, M.J.; Gil, R.R.; Du, S.; Madrid, M.; Ly, D. H. Crystal Structure of Chiral  $\gamma$  PNA with Complementary DNA Strand—Insights into the Stability and Specificity of Recognition and Conformational Preorganization. *J. Am. Chem. Soc.* **2010**, *132*, 10717-10727.
62. (a) Rasmussen, H.; Kastrop, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 Å resolution. *Nat. Struct. Biol.* **1997**, *4*, 98-101. (b) Eldrup, A. B.; Nielsen, B. B.; Haaima, G.; Rasmussen, H.; Kastrop, J. S.; Christensen, C.; Nielsen, P. E. 1,8-Naphthyridin-2(1H)-ones. Novel Bicyclic and Tricyclic Analogues of Thymine in Peptide Nucleic Acids (PNAs). *Eur. J. Org. Chem.* **2001**, *9*, 1781-1790. (c) Haaima, G.; Rasmussen, H.; Schmidt, G.; Jensen, D. K.; Kastrop, J. S.; Stafshede, P. W.; Norden, B.; Buchardt, O.; Nielsen, P. E. Peptide nucleic acids (PNA) derived from N-(N-methylaminoethyl)glycine. Synthesis, hybridization and structural properties. *New J. Chem.* **1999**, *23*, 833-840. (d) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. NMR solution structure of a peptide nucleic acid complexed with RNA. *Science* **1994**, *265*, 777-780. (e) He, W.; Hatcher, E.; Balae, A.; Beratan, D. N.; Gil, R. R.; Madrid, M.; Achim, C. Solution

- Structure of a Peptide Nucleic Acid Duplex from NMR Data: Features and Limitations. *J. Am. Chem. Soc.* **2008**, *130*, 13264–13273. (f) He, W.; Crawford, M. J.; Rapireddy, S.; Madrid, M.; Gil, R. R.; Ly, D. H.; Achim, C. The structure of a c-modified peptide nucleic acid duplex. *Mol. Biol. Syst.* **2010**, *6*, 1619–1629
63. Barawkar, D. A.; Ganesh, K. N. Fluorescent d- (CGCGAATTCGCG): Characterization of Major Groove Polarity and Study of Minor Groove Interactions through a Major groove Semantophore Conjugate. *Nucleic Acids Res.* 1995, *23*, 159–164.
64. Jadhav, V. R.; Barawkar, D. A.; Ganesh, K. N. Polarity Sensing by Fluorescent Oligonucleotides: First Demonstration of Sequence- Dependent Microenvironmental Changes in the DNA Major Groove. *J. Phys. Chem. B* 1999, *103*, 7383–7385. (17) Kimura, T.; Kawai, K.; Majima,
65. Vijay N. Kadam, Kayarat Saikrishnan, and Krishna N. Ganesh. 5-Amidodansyl-U (UD) Peptide Nucleic Acid (PNA) as a Fluorescent Sensor of the Local Dielectric Constant ( $\epsilon$ ) in PNA Duplexes: Major Grooves in PNA Duplexes Are More Hydrophobic Than Major Grooves in DNA–DNA Duplexes. *J. Phys. Chem. C* 2018, *122*, 14004–14013.
66. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **1991**, *254*, 1497-1501.
67. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **1991**, *254*, 1497-1501. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. Evidence for (PNA)<sub>2</sub>/DNA triplex structure upon binding of PNA to dsDNA by strand displacement. *J. Mol. Recogn.* **1994**, *7*, 165- 170.
68. Ray, A.; Norden, B. Peptide nucleic acid (PNA): its medical and biotechnical applications and promise for the future. *FASEB J.* **2000**, *14*, 1041-1060.
69. Nielsen, P. E.; Egholm, M.; Buchardt, O. Sequence-specific transcription arrest by peptide nucleic acid bound to the DNA template strand. *Gene* **1994**, *149*, 139-145.
70. Meng, M.; Ducho, C. Oligonucleotide analogues with cationic backbone linkages. *Beilstein J. Org. Chem.* **2018**, *14*, 1293–1308.
71. Knudsen, H.; Nielsen, P. E. Antisense properties of duplex- and triplex-forming PNAs. *Nucleic Acids Res.* **1996**, *24*, 494-500.

72. Taylor, R. W.; Chinnery, P. F.; Turnbull, D. M.; Lightowlers, R. N. Selective inhibition of mutant human mitochondrial DNA replication in vitro by peptide nucleic acids. *Nature Genet.* **1997**, *15*, 212-215.
73. Demers, D.B.; Curry, E.T.; Egholm, M.; Sozer, A. C. Enhanced PCR amplification of VNTR locus D1S80 using peptide nucleic acid (PNA). *Nucleic Acids Res.* **1995**, *23*, 3050-3055.
74. Murtola, M.; Wenska, M.; Strömberg, R. PNAzymes That Are Artificial RNA Restriction Enzymes. *J. Am. Chem. Soc.* **2010**, *132*, 8984-8990.
75. Orum, H.; Nielsen, P.E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res.* **1993**, *21*, 5332-5336.
76. Vernille, J. P.; Kove, L. C.; Schneider, J. W. Peptide Nucleic Acid (PNA) Amphiphiles: Synthesis, Self-Assembly, and Duplex Stability. *Bioconjugate Chem.* **2004**, *15*, 1314-1321.
77. Guler, M. O.; Pokorski, J. K.; Appella, D. H.; Stupp, S. I. Enhanced oligonucleotide binding to self-assembled nanofibers. *Bioconjugate Chem.* **2005**, *16*, 501-503.
78. Ellipilli, S.; Vasudeva M. R.; Ganesh, K. N. Perfluoroalkylchain conjugation as a new tactic for enhancing cell permeability of peptide nucleic acids (PNAs) via reducing the nanoparticle size. *Chem. Commun.* **2016**, *52*(3), 521-524.
79. Kumar, V. A.; Ganesh, K. N. Conformationally Constrained PNA Analogues: Structural Evolution toward DNA/RNA Binding Selectivity. *Acc. Chem. Res.* **2005**, *38*, 404-412.
80. Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. Peptide nucleic acids with a structurally biased backbone. Updated review and emerging challenges. *Curr. Top. Med. Chem.* **2011**, *11*, 1535-1554.
81. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. Aminoethylpropyl Peptide Nucleic Acids (aepPNA): Chiral PNA Analogues That Form Highly Stable DNA:aepPNA<sub>2</sub> Triplexes. *Org. Lett.* **1999**, *1*, 1513-1516.
82. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. Synthesis of N $\alpha$ -(purinyl/pyrimidinyl acetyl)-4-aminoproline diastereomers with potential use in PNA synthesis. *Tetrahedron* **1996**, *52*, 15017-15030.
83. Gangamani, B. P.; D'Costa, M.; Kumar, V. A.; Ganesh, K. N. Conformationally restrained chiral PNA conjugates: synthesis and DNA complementation studies. *Nucleosides Nucleotides.* **1999**, *18*, 1409-1011.

84. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. Chiral analogs of peptide nucleic acids: synthesis of 4-aminopropyl nucleic acids and DNA complementation studies using UV/CD spectroscopy. *Tetrahedron* **1999**, *55*, 177-192.
85. Dueholm, K. L.; Pettersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg.* Peptide nucleic acid (PNA) with a chiral backbone based on alanine. *Med. Chem. Lett.* **1994**, *4*, 1077-1080.
86. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. Peptide Nucleic Acids (PNAs) Containing Thymine Monomers Derived from Chiral Amino Acids : Hybridization and Solubility Properties of D-Lysine PNA. *Angew. Chem. Int. Ed.* **1996**, *35*, 1939-1942
87. Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. Chiral Peptide Nucleic Acids (PNAs): Helix Handedness and DNA Recognition. *Eur. J. Org. Chem.* **1999**, 197-204.
88. Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. Induction of Helical Handedness and DNA Binding Properties of Peptide Nucleic Acids (PNAs) with Two Stereogenic Centres. *Eur. J. Org. Chem.* **2007**, 5879-5885.
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. Antisense and antigene properties of peptide nucleic acids. *Science* **1992**, *258*, 1481-1485.
90. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Peptide nucleic acids (PNAs): Potential anti-sense and anti-gene agents. *Anti-Cancer Drug Design*, **1993**, *8*, 53-63.
91. Mologni, L.; leCoutre, P.; Nielsen, P. E.; Gambacorti-Passerini, C. Additive antisense effects of different PNAs on the in vitro translation of the PML/RAR $\alpha$  gene. *Nucleic Acids Res.* **1998**, *26*, 1934-1938.
92. Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas, S. M.; Zon, G.; Livak, K. J.; Ly, D. H. Synthesis of Conformationally Preorganized and Cell-Permeable Guanidine-Based  $\gamma$ -Peptide Nucleic Acids ( $\gamma$ GPNAs). *J. Org. Chem.* **2009**, *74*, 1509-1516.
93. Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Bianchi, N.; Brognara, E.; Gambari, R.; Marcgelli, R.; Corradini, R. Cellular Uptakes, Biostabilities and Anti-miR-210 Activities of Chiral Arginine-PNAs in Leukaemic K562 Cells. *ChemBioChem* **2012**, *13*, 1327-1337.

94. Mitra, R.; Ganesh, K. N. PNAs grafted with ( $\alpha/\gamma$ , R/S)-aminomethylene pendants: Regio and stereo specific effects on DNA binding and improved cell uptake. *Chem. Commun.* **2011**, *47*, 1198-1200.
95. (a) Mitra, R.; Ganesh, K. N. Aminomethylene Peptide Nucleic Acid (am-PNA): Synthesis, Regio-/Stereospecific DNA Binding, And Differential Cell Uptake of ( $\alpha/\gamma$ ,R/S)am-PNA Analogues. *J. Org. Chem.* **2012**, *77*, 5696-5704. (b) Jain D. R. Ph.D thesis, Indian Institute Of Science Education And Research, Pune, **2013**
96. Dogandziyski, P.; Ghidini, A.; Danneberg, F. Strömberg, R.; Göbel, M. W. Studies on Tris(2-aminobenzimidazole)-PNA Based Artificial Nucleases: A Comparison of Two Analytical Techniques. *Bioconjugate Chem.* **2015**, *26*, 2514-2519.

## **Chapter 2**

# **Synthesis of *aeg*-Imidazole Monomers and Derived PNA Oligomers**

## 2.1 Introduction

In the biological systems, RNA is crucial for the sustenance of life. The pivotal role it plays in protein synthesis makes it ubiquitous in all living organisms. As messenger RNA it relays genetic information from the DNA in the nucleus to the cytoplasm. As transfer RNA, it activates the amino acids and primes them for incorporation into protein chains, and delivers them to the ribosomal sites for protein synthesis. In the ribosomal machinery, it catalyses the formation of peptide bonds. RNA also plays key regulatory roles as miRNA, siRNA etc., through various mechanisms that determine the extent to which a specific gene should be expressed in the cells.

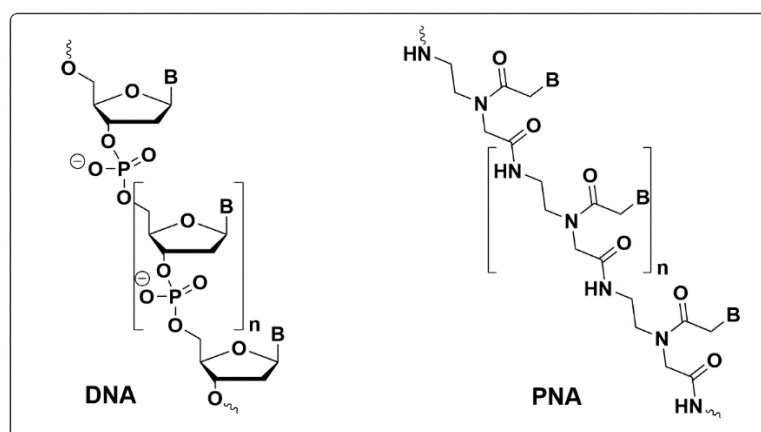
Given the critical role it plays in various forms in the universal life processes, it is not surprising to note that there are many sophisticated enzymes that act on RNA and manipulate them. RNA is synthesized by DNA dependent RNA polymerase such as RNA polymerase I,<sup>1</sup> RNA polymerase II, RNA polymerase III etc. RNA is degraded by ribonucleases such as RNase A<sup>2</sup>, RNase C<sup>3</sup>, RNase P<sup>4</sup> etc. Bovine pancreatic RNase A (EC 3.1.27.5) is one of the well-studied ribonucleases. As a model enzyme it has helped to advance our understanding of protein folding, stability, enzyme chemistry and molecular evolution.<sup>5</sup> The significance of this enzyme is underlined by the fact that Christian Anfinsen, Stanford Moore and William Stein were awarded the Nobel prize in chemistry for their seminal work on RNase A.

Developing artificial mimics of RNase A has been an active area of research for many decades. Such an enzyme mimic would be useful for sequence selective manipulations as artificial RNA restriction enzymes.<sup>6</sup> They also have potential application in cancer therapy<sup>7</sup> and nucleic acid therapeutics.<sup>8</sup> The promising candidates as artificial nucleases are those that have in their design, a functional catalytic moiety tethered to a sequence directing structural oligonucleotide unit. Using the principle of base complementarity elucidated by Watson and Crick<sup>9</sup> it is possible to design the oligonucleotide in such a way that it binds selectively to a specific target RNA, these systems provide the unique advantage of tuneable substrate specificity through sequence design. It also allows us to control and place the catalytic moiety near the backbone with high precision for cleavage at specific predetermined site in the sequence. However, the



use of DNA/RNA as structural scaffolds in these nucleic acid based systems suffer from the limitations of low stability *in vivo* as they are susceptible to nucleases. Analogues of the natural nucleic acids with a modified backbone that exhibit resistance to enzymatic degradation are therefore useful alternatives.

Over the past several decades peptide nucleic acid (PNA)<sup>10</sup> has emerged as one of the most promising oligonucleotide analogues. In PNAs, the sugar-phosphate backbone of DNA is replaced with a pseudo-peptide backbone with repeating (2-aminoethyl) glycine units. The nucleobases are attached to this achiral neutral backbone through a methylene carbonyl linker.



**Figure 2.1** Structural comparison of DNA and PNA (B = nucleobase)

The advantage of using PNA over DNA/RNA as scaffold are the following<sup>11</sup>: i) higher binding affinity, ii) lower tolerance to mismatch, which leads to better target specificity, iii) resistance to endogenous nucleases and proteases, which improves its stability *in vivo*, iv) as opposed to RNA, the polyamide backbone eliminates the possibility of self-cleavage.<sup>12</sup> With various PNA analogues reported in the recent literature<sup>13</sup> that address its limitations such as poor water solubility and poor cellular uptake properties, PNA has emerged as a suitable candidate for use as structural scaffold in sequence specific artificial nucleases.

Breslow, *et. al.*,<sup>14</sup> showed that imidazole buffer can catalyse RNA cleavage. Furthermore, functionalising a DNA/RNA binding units with catalytic imidazole moieties help understand the details of the catalytic mechanisms. Lonnberg *et. al.*,<sup>15</sup> have shown that a multi-nucleating azacrown conjugate selectively recognizes short oligoribonucleotides

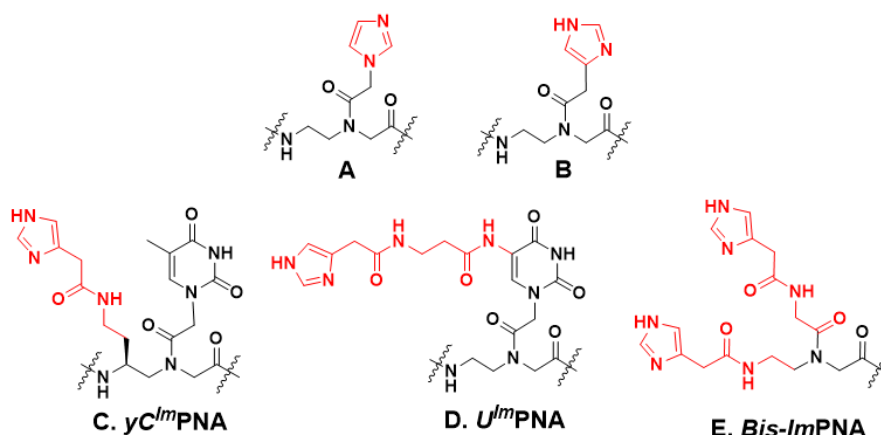
and cleave them in presence of  $Zn^{2+}$  metal ion. Komiyama, *et. al.*,<sup>16,17</sup> showed that PNA bearing catalytic Ce(IV)/EDTA species at the terminal position is efficient in cleavage of dsDNA. Also, they have shown site specific RNA cleavage using acridine conjugated DNA. Efforts towards the development of PNA based RNase mimic (PNAzymes) have focused mainly on the use of metal complexes as catalytic moieties. Strömberg, *et. al.*,<sup>18</sup> used Cu(II)-2,9 dimethyl phenanthroline to target non-base paired, bulge region in the RNA to demonstrate site and sequence specific hydrolysis of RNA. DNA/RNA cleavage can be affected either through hydrolysis of phosphodiester backbone or bi hydroxyl radical mediated cleavage of sugar units. There are also a few reports of non-metal based artificial PNAzymes. For instance, Göbel, *et. al.*,<sup>19</sup> used tris (2-amino benzimidazole) to cleave phosphodiester bond at specific sites on the backbone of target RNA analogues for DNA/RNA recognition.

## 2.2 Rationale behind the work

Imidazole moieties are indispensable in the active site of ribonucleases. In RNase A, His119 and His12 play catalytic role to hydrolyze the phosphodiester bond of the substrate RNA. The imidazole units are crucial here, because at physiological pH, they can act either as acid catalyst or as base catalyst, based on the necessity. While the imidazole moieties perform the catalytic function, the rest of the protein provides structural scaffold, acting as a support to hold the active site residues in place for optimal activity.

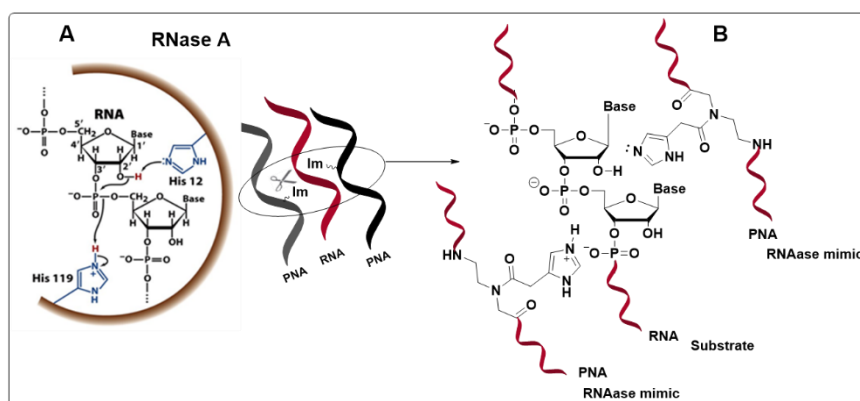
The present work deals with rationally designed PNA based hydrolytic nuclease mimics where a PNA strand is used as a structural scaffold and imidazole group is used as the functional catalytic unit. Earlier investigations into the catalytic potential of imidazole modified PNAs as RNase mimics involved a construct where the nucleobase was replaced with an imidazole at specific positions as shown in Figure 2.2A and B. This design showed no significant catalytic performance against complementary RNA.<sup>20</sup> Here, it is envisaged that the ability of PNA to hydrolyze the phosphodiester bonds on the complementary RNA backbone could be imparted by incorporating imidazole groups at various sites on the PNA. To this end, the following modified PNA were synthesized where the imidazole group is added on to: (i) the PNA backbone through amino ethyl

linkage at  $\gamma$ -C (Figure 2.2C), (ii) the nucleobase U at C5 through  $\beta$ -alanyl acetyl linkage (Figure 2.2D), and (iii) at both on the backbone and side chain (Figure 2.2E).



**Figure 2.2** PNA oligomers designed for RNA hydrolysis

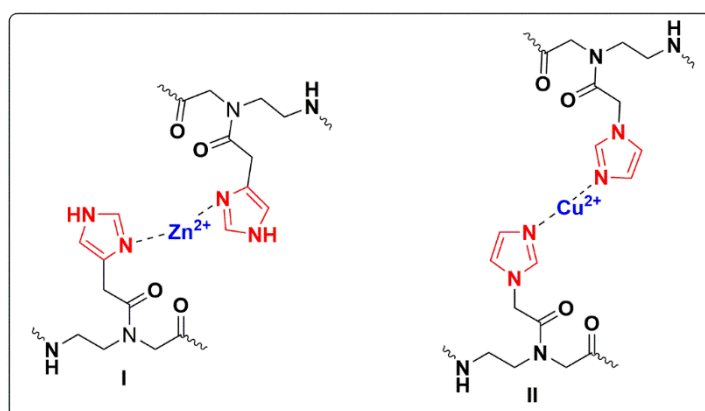
A PNA oligomer of a defined sequence modified by conjugation with a cleaving agent at a specific position (Figure 2.3) is likely to be a useful tool in RNA biochemistry and molecular biology, complementing the existing repertoire of enzymes such as restriction endonucleases and CRISPR/Cas9 etc. Such a PNA construct can also be a promising antisense agent since it would not rely on the cellular nuclease machineries for target mRNA hydrolysis.



**Figure 2.3** (A) Mechanism of RNA cleavage of RNase A at its active site and (B) Proposed scheme for PNA based artificial nuclease.

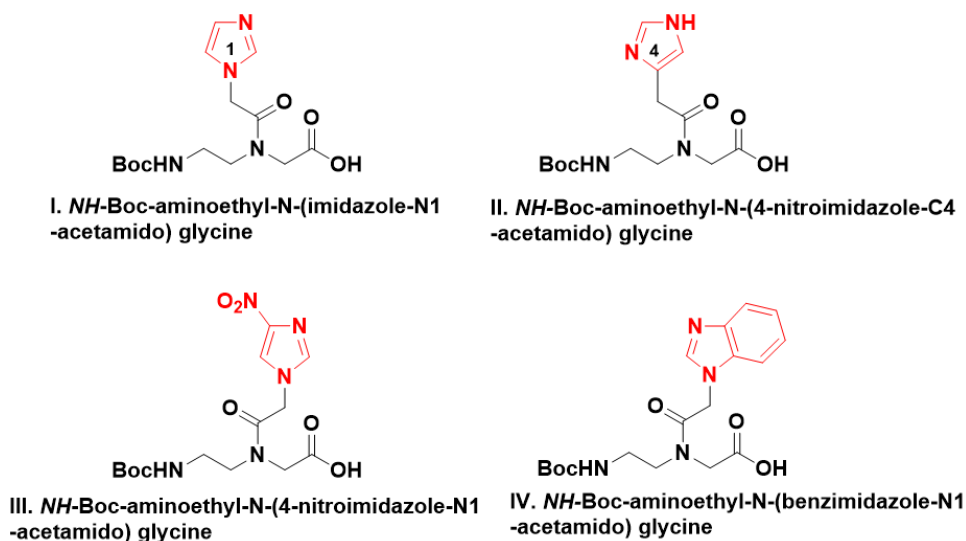
Since there are evidences of significant role played by divalent metal ions such as  $Zn^{2+}$  in some of the ribonucleases, we also examined metal complexation properties of the

designed monomers and oligomers. The PNAs, where the nucleobases are replaced with imidazole groups are likely to exhibit metallo-basepairing with metal ions like  $Zn^{2+}$  and  $Cu^{2+}$ . The metallo-basepairing in PNA was expected to result in the formation of stable PNA:PNA duplexes and discrete metal assembly inside the duplex structure. Since imidazole could be linked to PNA backbone either *via* N-alkylation or through C-alkylation, the positional effect on the selectivity towards co-ordinating metal ion was examined (Figure 2.4).



**Figure 2.4** Metal complexes with designed ligands

Since the primary role of the PNA strand in the artificial nucleases is to recognize and to bind to the complementary RNA, it is necessary to understand the effect of the modifications on the stability of the duplex and its base-pairing properties. This is especially relevant in those designs where an imidazole replaces the canonical nucleobases. Therefore, we examined the affinity of N1-linked and C4-linked imidazole for the four natural nucleobases in the duplex. In literature, 3-nitro pyrrole is known as a universal base.<sup>21</sup> The structural similarities between pyrrole and imidazole prompted us to study the imidazolyl PNA to ascertain whether they exhibit similar molecular recognition properties. For this, PNAs with nitroimidazole and benzimidazole modifications were also synthesized as control and their base-pair specificity was compared with those of the N1-linked and C4-linked imidazole PNA (Figure 2.5).

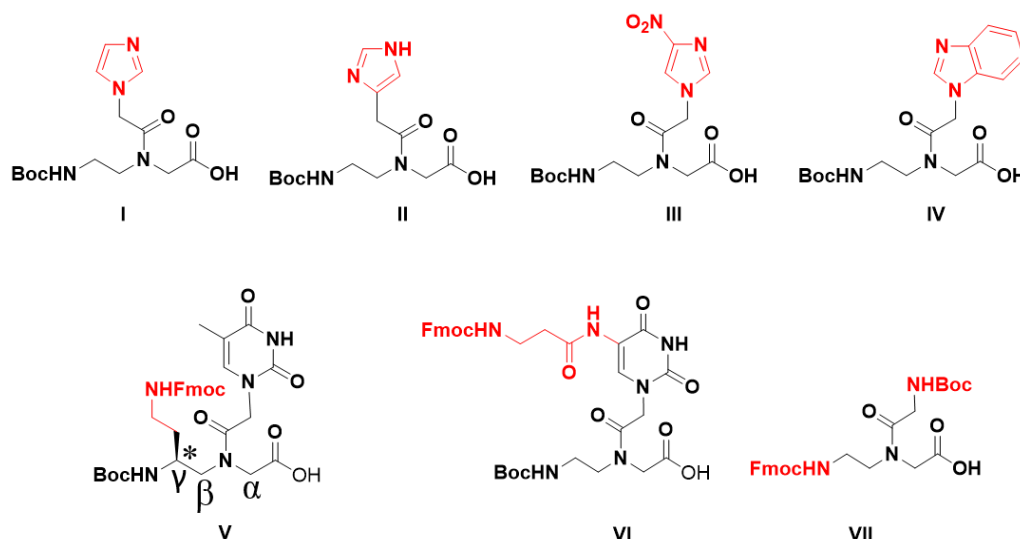


**Figure 2.5** Proposed PNA monomers for metal complexation and molecular recognition studies

## 2.3 Objectives

The specific objectives of this chapter are

- Synthesis of *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) glycine (I)
- Synthesis of *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) glycine (II)
- Synthesis of *NH*-Boc-aminoethyl-N-(4-nitroimidazole-N1-acetamido) glycine (III)
- Synthesis of *NH*-Boc-aminoethyl-N-(benzimidazole-N1-acetamido) glycine (IV)
- Synthesis of *NH*-Boc-aminoethyl- $\gamma$ C-(*S*)-(NH-Fmoc-aminoethyl)-N-(thymine-N1-acetamido) glycine (V)
- Synthesis of *NH*-Boc-aminoethyl-N-[5-amino(NH-Fmoc- $\beta$ -alanyl) uracil N1-acetamido] glycine (VI)
- Synthesis of *NH*-Fmoc-aminoethyl-N1-(NH-Boc-glycinamido) glycine (VII)
- Characterization of glycine derivatives (I-VII) and their intermediates by various spectroscopic techniques
- Synthesis of PNA oligomers incorporating monomers (I-VII)
- Purification and characterization of PNA oligomers



**Figure 2.6** Structure of target PNA monomers (I-VII)

## 2.4 Results and Discussion

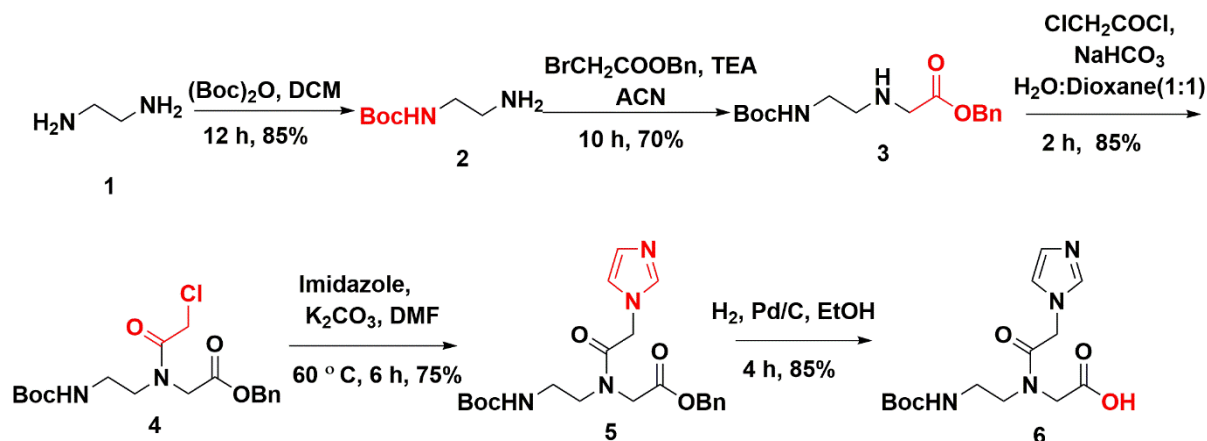
All imidazole modified PNA monomers (Figure 2.6 I-VII) were synthesised and characterised by various spectroscopic techniques. The modified PNA monomers were incorporated into PNA oligomers using solid phase synthesis.

### 2.4.1 Synthesis of modified PNA monomers

This section describes the synthesis of rationally designed imidazole modified PNA monomers.

#### 2.4.1a Synthesis of *NH-Boc-aminoethyl-N-(imidazole-N1-acetamido) glycine 6*

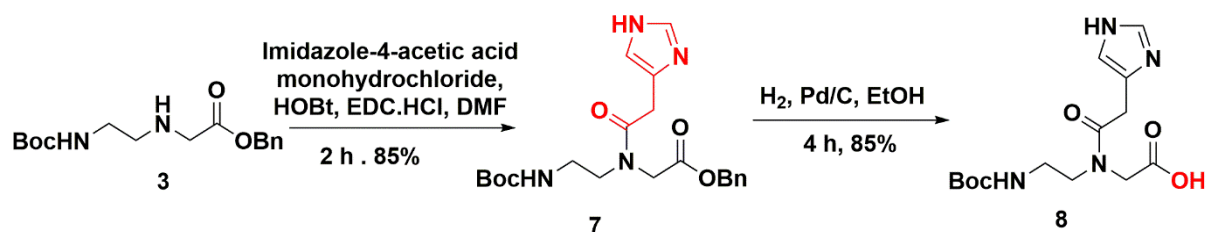
The commercially available ethylenediamine **1** was treated with Boc-anhydride to get mono *NH*-Boc-protected ethylenediamine **2**. This was N-alkylated by the reaction with benzylbromoacetate, triethylamine in acetonitrile solvent to form compound **3**, that was N-acylated by treating it with chloroacetyl chloride to get N-chloroacetyl compound **4**. This compound was reacted with imidazole to get *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) benzyl glycinate **5**, which upon hydrogenolysis using Pd/C in ethanol gave the *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) glycine **6**.



**Scheme 2.1** Synthesis of *NH*-Boc-aminoethyl-*N*-(imidazole-*N*1-acetamido) glycine **6**

#### 2.4.1b Synthesis of *NH*-Boc-aminoethyl-*N*-(imidazole-*C*4-acetamido) glycine **8**

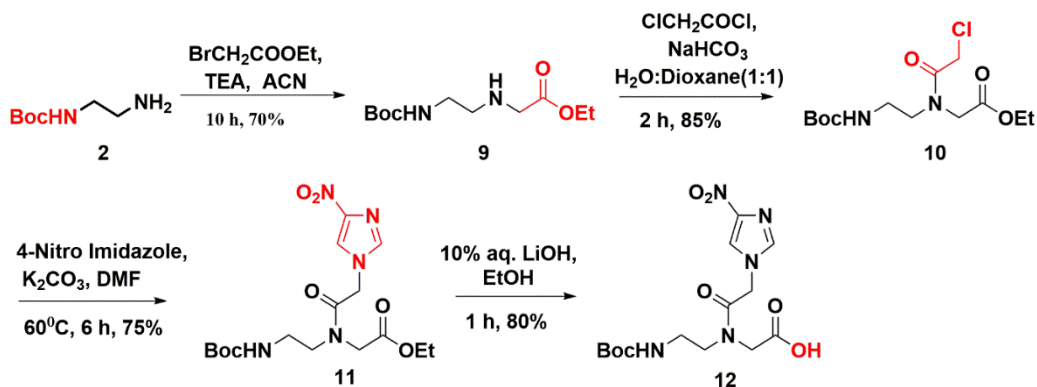
Compound **3** was coupled with imidazole-4-acetic acid in presence of HOBt, EDC and DIPEA to obtain *NH*-Boc-aminoethyl-*N*-(imidazole-*C*4-acetamido) benzyl glycinate **7**, which upon hydrogenolysis with Pd/C in ethanol gave the *NH*-Boc-aminoethyl-*N*-(imidazole-*C*4-acetamido) glycine **8**.



**Scheme 2.2** Synthesis of *NH*-Boc-aminoethyl-*N*-(imidazole-*C*4-acetamido) glycine **8**

#### 2.4.1c Synthesis of *NH*-Boc-(4-nitro-imidazole-*N*1-acetamido) aminoethyl glycine **12**

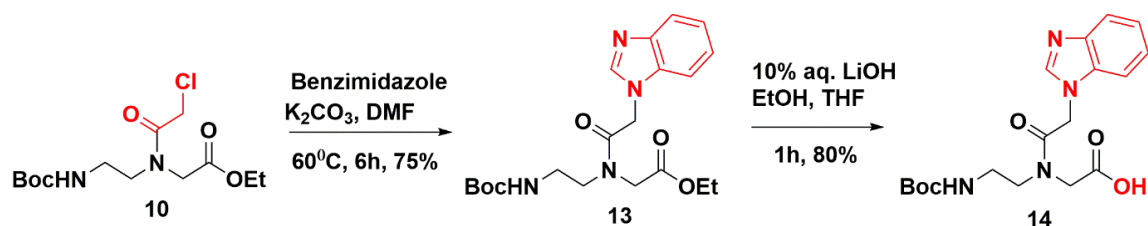
The mono *NH*-Boc-protected ethylenediamine compound **2** was *N*-alkylated by the reaction with ethylbromoacetate to form compound **9**, which was *N*-acylated by treating it with chloroacetyl chloride to obtain compound **10**. This was reacted with 4-nitroimidazole to obtain *NH*-Boc-aminoethyl-*N*-(4-nitroimidazole-*N*1-acetamido) ethyl glycinate **11**, which upon hydrolysis using aq. LiOH in ethanol gave the *NH*-Boc-aminoethyl-*N*-(4-nitroimidazole *N*1-acetamido) glycine **12**.



**Scheme 2.3** Synthesis of *NH*-Boc-aminoethyl-*N*-(4-nitroimidazole-*N*1-acetamido) glycine **12**

#### 2.4.1d Synthesis of *NH*-Boc-aminoethyl-*N*-(benzimidazole *N*1-acetamido) glycine **14**

Synthesis of *aeg*-*N*-acetyl-benzimidazole monomer (**14**) started from the previously synthesized compound **10** which was reacted with benzimidazole to obtain *NH*-Boc-ethylamino-*N*-(benzimidazole *N*1-acetamido) ethyl glycinate **13**, which upon hydrolysis using aq. LiOH in ethanol gave the *NH*-Boc-aminoethyl-*N*-(benzimidazole-*N*1-acetamido) glycine **14**.



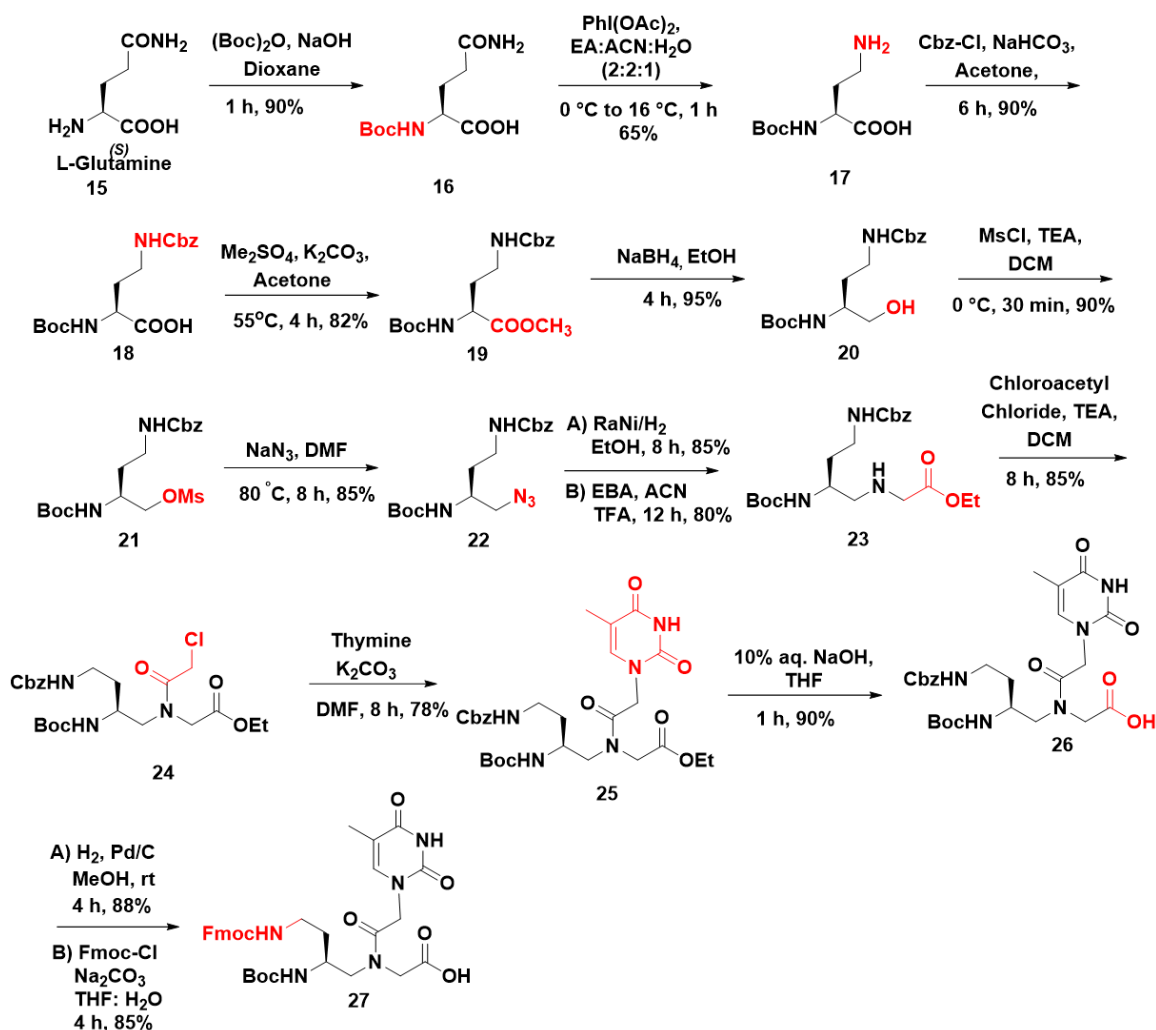
**Scheme 2.4** Synthesis of *NH*-Boc-aminoethyl-*N*-(benzimidazole-*N*1-acetamido) glycine **14**

#### 2.4.1e Synthesis *NH*-Boc-aminoethyl- $\gamma$ C-(*S*)-(NH-Fmoc-aminoethyl)-*N*-(thyminy-*N*1-acetamido) glycine **27**

The commercially available L-glutamine **15** was treated with Boc-anhydride in aq. NaOH-dioxane to obtain *NH*-Boc-protected L-glutamine **16** (Scheme 2.5). This was followed by the reaction of *N*- $\alpha$ -Boc-L-glutamine **16** with iodobenzene diacetate (PIDA)<sup>22</sup> which led to the formation of 4-amino-2-(*N*-Boc-amino)butanoic acid **17**. This was treated with benzylchloroformate in toluene in the presence of NaHCO<sub>3</sub> to obtain orthogonally protected 4-(Cbz-amino)-2-(Boc-amino) butanoic acid **18** in quantitative yield. Compound **18** was converted to methyl ester derivative **19** using dimethyl sulfate



and activated  $K_2CO_3$  and the conversion was confirmed by the appearance of  $^1H$  NMR peak at  $\delta$  3.73 ppm for  $-CH_3$  group of methyl ester in the product. The methyl ester was reduced using sodium borohydride in absolute ethanol to give the alcohol derivative **20**. Mesylation of the primary hydroxyl group by controlled addition of mesyl chloride in DCM triethyl amine gave the mesylate derivative **21** which was immediately treated with sodium azide in dry DMF to obtain the azido compound **22**. The reduction of azide derivative using Raney Ni under hydrogenation conditions yielded the free amine yield which was in-situ alkylated with ethyl bromoacetate to the alkylated compound **23**. It was then acylated with chloroacetyl chloride to yield the chloro compound **24**. This upon condensation with thymine afforded the  $\gamma$ C-(*S*-ethyleneamino) aminoethylglycyl ethyl glycinate **25** in quantitative amount. The appearance of peaks at  $\delta$  7.05 ppm and 1.87 ppm in  $^1H$  NMR shows the presence of thymine in the desired product **25**. Compound **25** was hydrolysed using aq. NaOH in THF to obtain compound **26** which upon hydrogenation with Pd/C in methanol yielded the free amine which was protected using Fmoc-Cl and aq. 10%  $Na_2CO_3$  in THF to yield the desired Fmoc derivative **27**. All the intermediates were purified by column chromatography and characterized by  $^1H$ ,  $^{13}C$  NMR and mass spectral analysis.

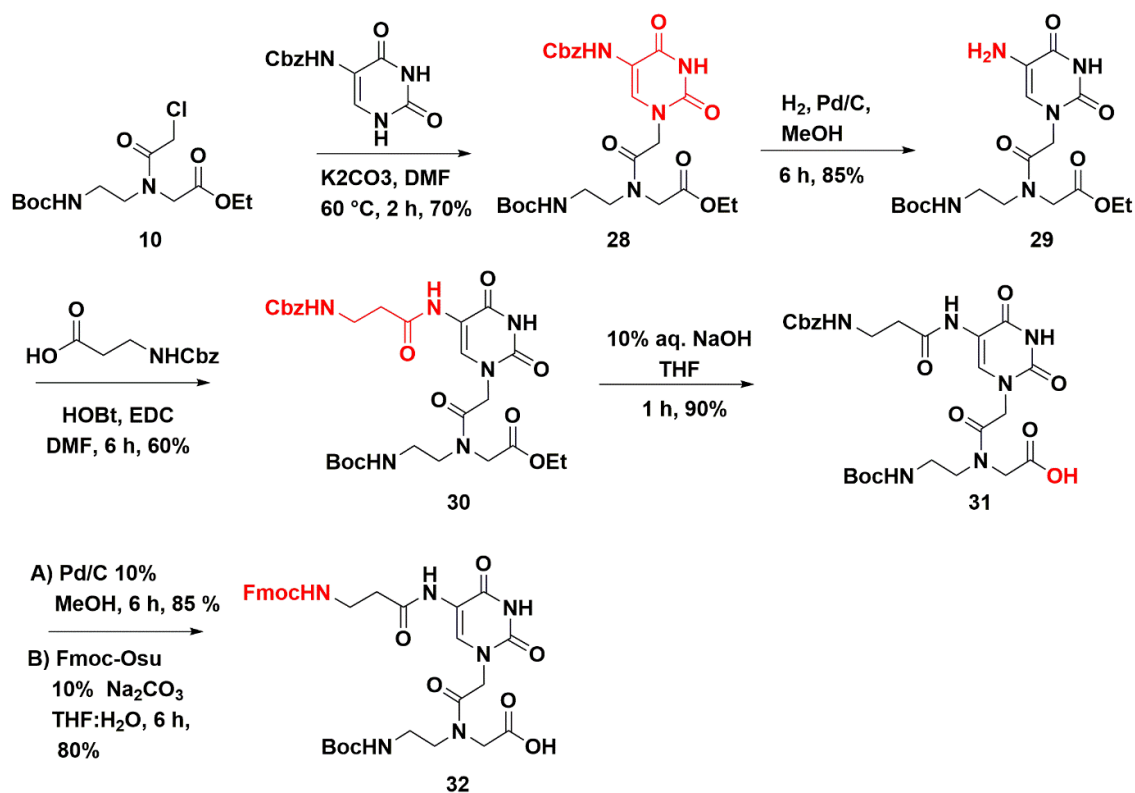


**Scheme 2.5** Synthesis of *NH*-Boc-aminoethyl- $\gamma$ C-(*S*)-(NH-Fmoc-aminoethyl)-*N*-(thyminyln1-acetamido) glycine **27**

#### 2.4.1f Synthesis of *NH*-Boc-aminoethyl-*N*-[5-amino(NH-Fmoc- $\beta$ -alanyl) uracil *N*1-acetamido] glycine **32**

Compound **10** was treated with *NH*-Cbz protected 5-amino uracil in DMF using activated  $\text{K}_2\text{CO}_3$  as base to obtain *NH*-Cbz-5-amino uracil ethyl glycinate **28** which was subjected for hydrogenation to get *NH*-Boc-aminoethyl-*N*-(5-aminouracil-*N*1-acetamido) ethyl glycinate **29** (Scheme 2.6). The free amine was coupled with *NH*-Cbz-protected  $\beta$ -alanine using EDC, HOBt and DIPEA to obtain *NH*-Boc-aminoethyl-*N*-(*NH*-Cbz- $\beta$ -alanyl-5-aminouracil-*N*1-acetamido) ethyl glycinate **30**. It was hydrolysed using aq. NaOH in THF to the *NH*Boc-aminoethyl-*N*-[5-(*NH*Cbz- $\beta$ -alanyl) uracil *N*1-acetamido] ethyl glycinate **31** followed by removal of Cbz-group by hydrogenation using Pd/C to get the free amine

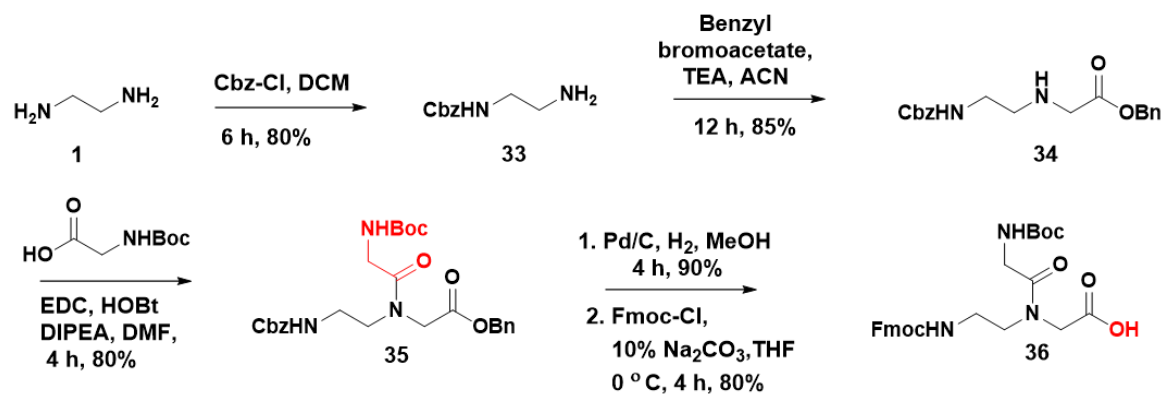
which was in-situ reacted with Fmoc-Cl to obtain *NH*-Boc-aminoethyl-N-[5-amino(*NH*-Fmoc- $\beta$ -alanyl) uracil-N1-acetamido] glycine **32**.



**Scheme 2.6** Synthesis of *NH*-Boc-aminoethyl-N-[5-amino(*NH*-Fmoc- $\beta$ -alanyl) uracil N1-acetamido] glycine **32**

#### 2.4.1g Synthesis of *NH*-Fmoc-aminoethyl-N1-(*NH*-Boc-glycinamido) glycine

The commercially available ethylenediamine **1** was protected with Cbz-Cl to get mono *NH*-Cbz-protected ethylenediamine **33**. This was N-alkylated by reaction with benzylbromoacetate and triethylamine in acetonitrile solvent to form *NH*-Cbz-aminoethylglycyl benzyl glycinate **34**, which was coupled with Boc-protected glycine in presence of EDC, HOBt and DIPEA to get *NH*-Cbz-aminoethyl-N1-(*NH*-Boc-glycinamido) aminoethylglycine benzyl glycinate **35**. Compound **35** was subjected with hydrogenation using Pd/C in methanol to obtain the free amine acid. That was reacted with Fmoc-Cl to get *NH*-Fmoc-ethylamino-N-(*NH*-Boc glycinamido) glycine **36**.

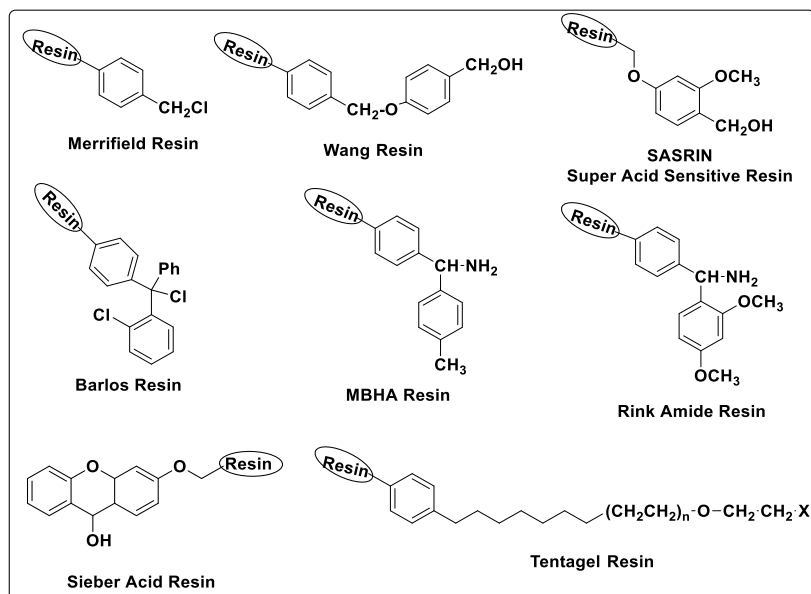


**Scheme 2.7** Synthesis of *NH*-Fmoc-aminoethyl-*N*-(*NH*-Boc-glycinamido) glycine **36**

### 2.4.2 Solid Phase Peptide Synthesis

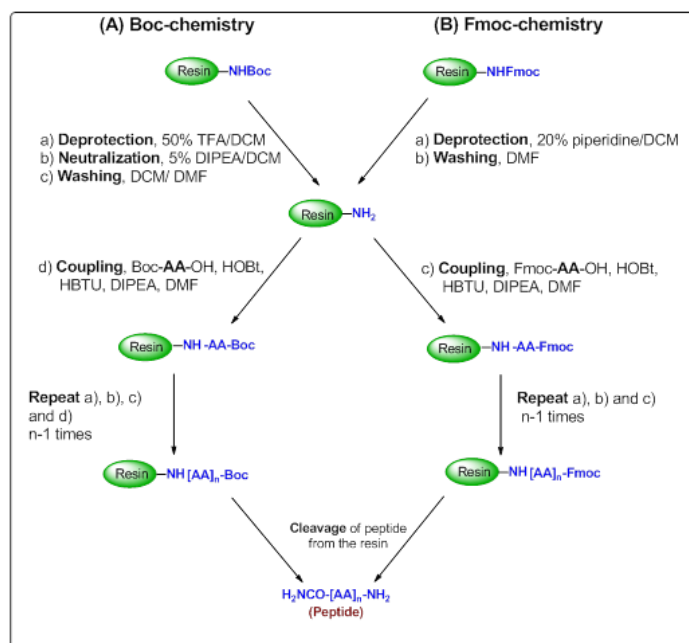
Peptides can be synthesized either in solution phase or by solid phase synthesis protocol.<sup>23</sup> Synthesis of short chain peptides can be more efficient by solution phase strategy. Solution phase synthesis requires tedious separation and purification steps after each coupling reaction. On the other hand, solid phase peptide synthesis can be efficiently used in the synthesis of several short and long chain peptides as well as in the synthesis of PNA oligomers.

Solid phase peptide synthesis, first invented by Merrifield,<sup>24</sup> utilizes polymeric the bead which have functional groups located on their surfaces and in their pores<sup>25</sup> (Figure 2.7) Small solid beads are insoluble in organic solvents. The peptide chains are grown over them by sequentially linking the aminoacid monomers to the functional groups or linkers on the solid support. In each step, the resin beads are immersed in the appropriate solvent containing the reagents for the reaction to proceed. Following the reaction, the solvent and excess reagents and the by products are washed away by filtration. The next *N*<sup>α</sup>-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. Because the *C*-terminal amino acid is linked to the insoluble solid support, it also acts as a protection for the carboxylic acid during the synthesis (Figure 2.7). The excess amino acid is washed out and the deprotection and coupling reactions are repeated until the desired peptide sequence is achieved. Finally, the resin bound peptide and the side chain protecting groups are cleaved in a global deprotection step.



**Figure 2.7** Representative structures of resin used in SPPS

There are two routinely followed protocols for solid phase peptide synthesis-*Fmoc* strategy and *Boc* strategy, which use base labile and acid labile protecting groups respectively (Figure 2.8). The solid-phase peptide synthesis proceeds in a C-terminal to N-terminal fashion. The N-termini of amino acid monomers are protected by these two groups and added onto a deprotected amino acid chain. First protocol uses the *t*-butoxycarbonyl (*t*-Boc) group as  $N^\alpha$ -protection that is removed by 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 dimethylsulfide or TFMSA in TFA. In the alternative protocol fluorenylmethyloxycarbonyl (Fmoc) group is used as  $N^\alpha$ -protection which is extremely stable to acidic conditions but can be cleaved off efficiently with a base such as piperidine. The final peptide and side chain protecting groups can be cleaved with acid (50% TFA in DCM).



**Figure 2.8** General protocols for SPPS via (A) Boc-chemistry (B) Fmoc-chemistry

### 2.4.2a Synthesis of PNA oligomers

Synthesis of PNA oligomers were carried out using solid phase synthesis protocol using Boc strategy and Fmoc strategy. The modified PNA monomers were incorporated at various positions in the unmodified *aeg* PNA sequence. PNA oligomerization was carried out from C-terminus to the N-terminus end using monomeric units with protected amino and carboxylic acid functions maintaining the orthogonality.

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/Ci-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 reduced from 0.6 mmol/g to 0.35 mmol/g by partial acylation of amine content using calculated amount of acetic anhydride.<sup>26</sup> The free amine groups on the resin available for coupling was confirmed before starting synthesis by Kaiser's test.

The deprotection of the Boc protecting group and the completion of coupling reaction were monitored by Kaiser's test.<sup>27</sup> The Boc deprotection leads to a positive Kaiser's test,

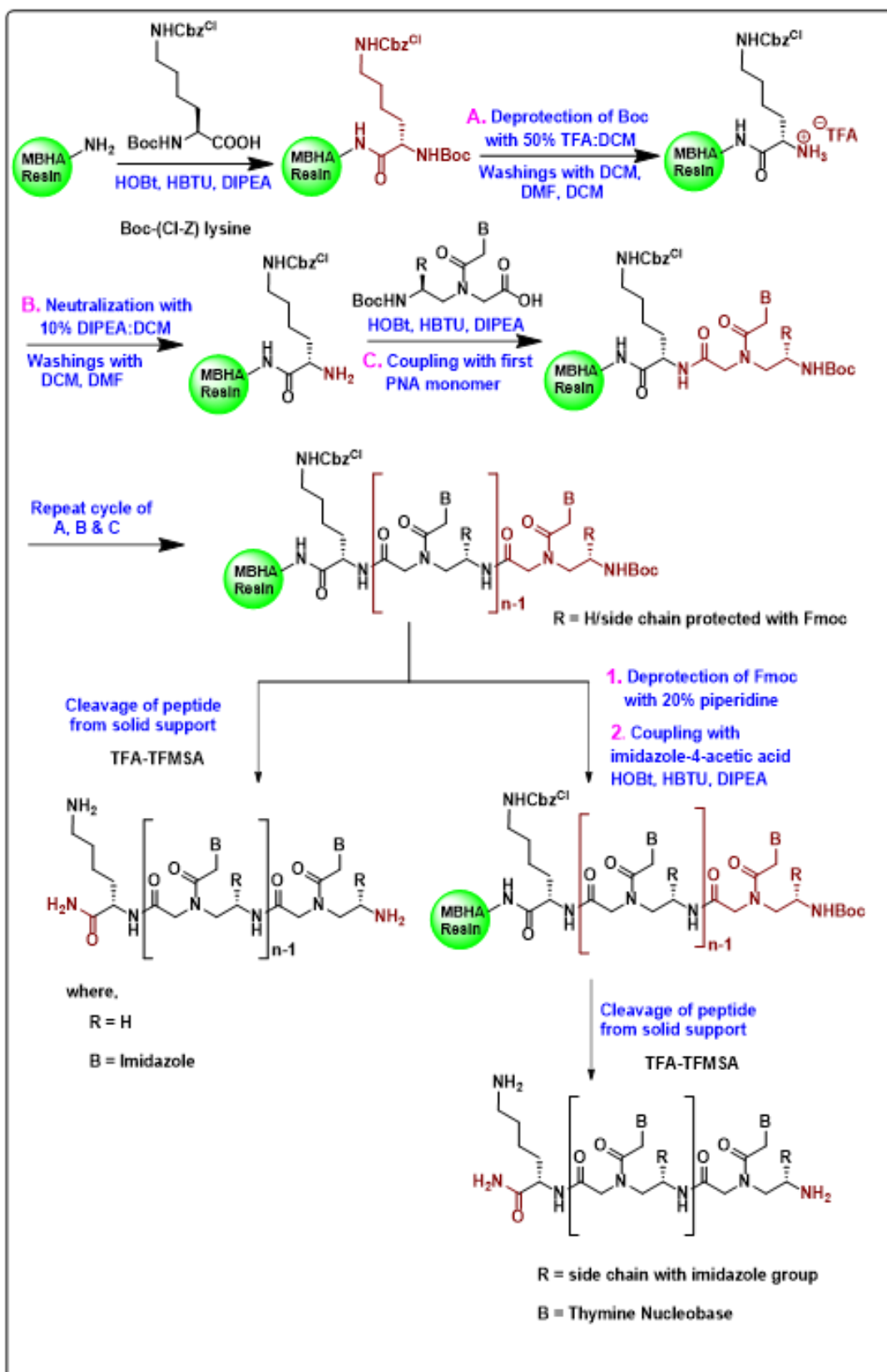


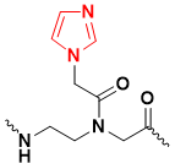
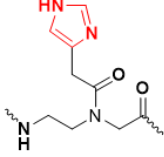
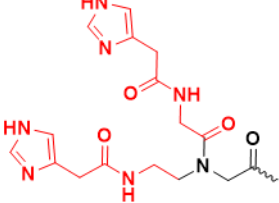
Figure 2.9 Solid phase PNA synthesis protocol by *Boc* strategy

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 2.9), the PNA oligomers of desired length incorporating modified as well as unmodified PNA monomers at desired positions were synthesized.

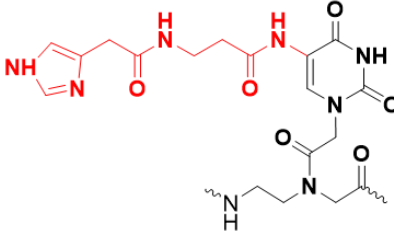
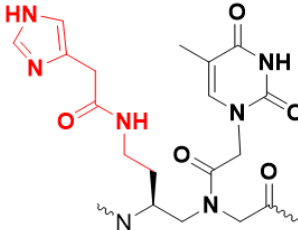
#### 2.4.2b Synthesis of imidazolyl PNA oligomers as nuclease mimics

For the purpose of the study of RNA hydrolysis and metal complexes various PNA oligomers were synthesized following the solid phase protocol using Boc-strategy (Table 2.1).

**Table 2.1** PNA oligomers with modified/unmodified monomers at various positions

Entry	Sequence Code	PNA sequences	Monomers used
1	<i>aeg</i> PNA 1	H-G G C A T G C C LysNH <sub>2</sub>	A/G/C/T = <i>aeg</i> PNA
2	<i>Im</i> (N1) PNA 2	H-GGCA <i>Im</i> (N1) TGCCLysNH <sub>2</sub>	 <i>Im</i> (N1)=aminoethyl(imidazole N1-acetamido) glycol
3	<i>Im</i> (C4) PNA 3	H-GGCA <i>Im</i> (C4) TGCCLysNH <sub>2</sub>	 <i>Im</i> (C4)=aminoethyl(imidazole C4-acetamido) glycol
4	<i>Bis-Im</i> PNA 4	<i>Bis-Im</i> -TTTTTGCLysNH <sub>2</sub>	 <i>Bis-Im</i> = <i>aeg</i> -Bis-imidazole

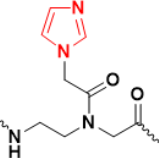
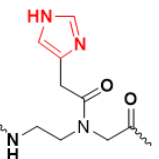


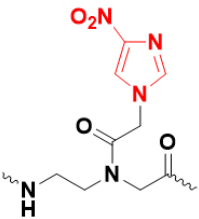
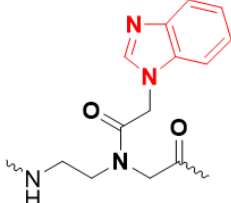
5	$U^{Im}$ T <sub>10</sub> PNA <b>5</b>	H-TTTT $U^{Im}U^{Im}$ TTTTLysNH <sub>2</sub>	
6	$U^{Im}$ PNA <b>6</b>	H-TCTCAAG $U^{Im}$ TGGGLysNH <sub>2</sub>	$U^{Im} = aeg-C5-(\beta-Ala-Im-U)$
7	$\gamma C^{Im}$ T <sub>10</sub> PNA <b>7</b>	H-TTTT $\gamma^{Im}\gamma^{Im}$ TTTTLysNH <sub>2</sub>	
8	$\gamma C^{Im}$ PNA <b>8</b>	H-TCTCAAG $\gamma^{Im}$ TGGGLysNH <sub>2</sub>	$\gamma^{Im} = aminoethyl-\gamma C-(S-eam-Im)glycyl$

### 2.4.2c Synthesis of imidazolyl PNA oligomers for nucleases

To study the imidazole binding with natural bases various PNA oligomers were synthesized following the solid phase protocol using Boc-strategy (Table 2.2)

**Table 2.2** PNA oligomers with modified/unmodified monomers for base recognition

Entry	Sequence Code	PNA sequences	Monomers used
9	$Im(N1)$ PNA <b>9</b>	H-GTAG $Im(N1)$ TACTTLysNH <sub>2</sub>	 $Im(N1) = aminoethyl (imidazole N1-acetamido) glycyl$
10	$Im(C4)$ PNA <b>10</b>	H-GTAG $Im(C4)$ TACTTLysNH <sub>2</sub>	 $Im(C4) = aminoethyl (imidazole C4-acetamido) glycyl$

11	<i>Im NO<sub>2</sub></i> PNA <b>11</b>	H-GTAG <b><i>ImNO<sub>2</sub></i></b> TACTTlysNH	 <p><i>ImNO<sub>2</sub></i> = <i>ae</i>-(4-NO<sub>2</sub>-imidazole - N1-acetamido) glycyL</p>
12	<i>Bzim</i> PNA <b>12</b>	H-GTAG <b><i>Bzim</i></b> TACTTlysNH <sub>2</sub>	 <p><i>Bzim</i> = <i>ae</i>-(benzimidazole N1-acetamido) glycyL</p>
13	PNA <b>13</b>	H-AAGTA <b>A</b> CTAClysNH <sub>2</sub>	A/G/C/T = <i>ae</i> g PNA
14	PNA <b>14</b>	H-AAGTA <b>T</b> CTAClysNH <sub>2</sub>	
15	PNA <b>15</b>	H-AAGTA <b>C</b> CTAClysNH <sub>2</sub>	
16	PNA <b>16</b>	H-AAGTA <b>G</b> CTAClysNH <sub>2</sub>	
17	PNA <b>17</b>	H-CATC <b>A</b> ATGAALysNH <sub>2</sub>	
18	PNA <b>18</b>	H-CATC <b>T</b> ATGAALysNH <sub>2</sub>	
19	PNA <b>19</b>	H-CATC <b>C</b> ATGAALysNH <sub>2</sub>	
20	PNA <b>20</b>	H-CATC <b>G</b> ATGAALysNH <sub>2</sub>	
21	PNA <b>21</b>	H-GTAG <b>C</b> TACTTlysNH <sub>2</sub>	
22	PNA <b>22</b>	H-GTAG <b>T</b> TACTTlysNH <sub>2</sub>	

### 2.4.3 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 amides at their C-termini.<sup>28</sup> In this cleavage condition, side chain protecting group was removed accompanied by deprotection of nucleobases. After the 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.

### 2.4.4 Purification and characterization of the PNA oligomers

After cleavage from the solid support, reverse phase high performance liquid chromatography (RP-HPLC) was used to purify PNA oligomers. 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.

MALDI-TOF mass spectrometry was used to confirm the integrity of these synthesized PNA oligomers. In literature, various matrices like sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), picolinic acid (PA), 2,5-dihydroxybenzoic acid (DHB),  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) etc. have been reported to record MALDI-TOF spectrum. Of these, DHB and CHCA were 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.

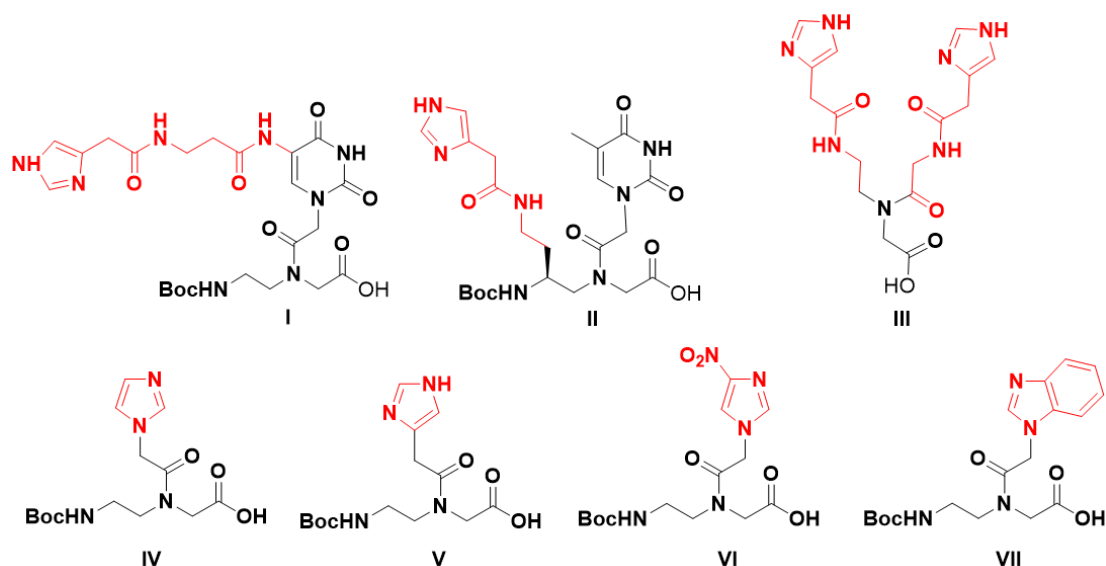
**Table 2.3** MALDI-TOF spectral analysis of the synthesized PNA oligomers

<b>Sr. No</b>	<b>PNA sequence code</b>	<b>Molecular Formula</b>	<b>Calculated Mass</b>	<b>Observed Mass</b>	<b>Retention Time (min)</b>
1	<i>aeg</i> PNA <b>1</b>	C <sub>91</sub> H <sub>120</sub> N <sub>50</sub> O <sub>25</sub>	2314.26	2314.49	12.7
2	<i>Im</i> (N1) PNA <b>2</b>	C <sub>100</sub> H <sub>132</sub> N <sub>54</sub> O <sub>27</sub> Na	2522.48	2522.00	16.6
3	<i>Im</i> (C4) PNA <b>3</b>	C <sub>100</sub> H <sub>132</sub> N <sub>54</sub> O <sub>27</sub> Na	2522.49	2522.32	15.1
4	<i>Bis-Im</i> PNA <b>4</b>	C <sub>98</sub> H <sub>130</sub> N <sub>42</sub> O <sub>31</sub>	2415.37	2414.32	11.2
5	<i>U<sup>Im</sup></i> T <sub>10</sub> PNA <b>5</b>	C <sub>130</sub> H <sub>171</sub> N <sub>51</sub> O <sub>45</sub>	3168.11	3167.58	13.6
6	<i>U<sup>Im</sup></i> PNA <b>6</b>	C <sub>143</sub> H <sub>183</sub> N <sub>75</sub> O <sub>40</sub>	3608.52	3608.55	14.2
7	<i>γ<sup>Im</sup></i> T <sub>10</sub> PNA <b>7</b>	C <sub>130</sub> H <sub>173</sub> N <sub>49</sub> O <sub>43</sub> Na	3132.08	3133.88	13.8
8	<i>γ<sup>Im</sup></i> PNA <b>8</b>	C <sub>143</sub> H <sub>184</sub> N <sub>74</sub> O <sub>40</sub>	3579.52	3579.58	13.7
9	<i>Im</i> (N1) PNA <b>9</b>	C <sub>113</sub> H <sub>148</sub> N <sub>56</sub> O <sub>32</sub>	2802.78	2802.39	12.7
10	<i>Im</i> (C4) PNA <b>10</b>	C <sub>113</sub> H <sub>148</sub> N <sub>56</sub> O <sub>32</sub>	2802.78	2802.32	14.4
11	<i>Im</i> NO <sub>2</sub> PNA <b>11</b>	C <sub>113</sub> H <sub>147</sub> N <sub>57</sub> O <sub>34</sub>	2870.76	2870.10	13.5
12	<i>Bzim</i> PNA <b>12</b>	C <sub>117</sub> H <sub>150</sub> N <sub>56</sub> O <sub>32</sub>	2875.17	2874.82	13.3
13	PNA <b>13</b>	C <sub>114</sub> H <sub>147</sub> N <sub>63</sub> O <sub>28</sub> Na	2870.82	2870.22	13.5
14	PNA <b>14</b>	C <sub>114</sub> H <sub>148</sub> N <sub>60</sub> O <sub>30</sub> Na	2861.80	2861.30	13.2
15	PNA <b>15</b>	C <sub>113</sub> H <sub>147</sub> N <sub>61</sub> O <sub>29</sub> Na	2846.79	2846.11	13.5
16	PNA <b>16</b>	C <sub>114</sub> H <sub>147</sub> N <sub>63</sub> O <sub>29</sub> Na	2886.82	2886.43	13.0

17	PNA <b>17</b>	C <sub>114</sub> H <sub>147</sub> N <sub>63</sub> O <sub>28</sub>	2847.84	2847.41	13.3
18	PNA <b>18</b>	C <sub>114</sub> H <sub>148</sub> N <sub>60</sub> O <sub>30</sub> Na	2861.80	2861.59	13.1
19	PNA <b>19</b>	C <sub>113</sub> H <sub>147</sub> N <sub>61</sub> O <sub>29</sub> Na	2846.79	2846.17	13.2
20	PNA <b>20</b>	C <sub>114</sub> H <sub>147</sub> N <sub>63</sub> O <sub>29</sub>	2863.84	2864.18	13.4
21	PNA <b>21</b>	C <sub>113</sub> H <sub>147</sub> N <sub>61</sub> O <sub>29</sub> Na	2868.80	2867.92	13.0
22	PNA <b>22</b>	C <sub>114</sub> H <sub>148</sub> N <sub>60</sub> O <sub>30</sub> Na	2861.80	2860.89	13.4

## 2.5 Summary

To summarize, this section describes the synthesis and characterization of rationally designed PNA monomers (Figure 2.10) where the imidazole moiety is incorporated on the backbone, on the nucleobase and by replacing the base with imidazole. Bis-imidazole PNA was also synthesized which contained the modifications both on the backbone and on the sidechain. All the intermediates have been characterized by <sup>1</sup>H & <sup>13</sup>C NMR spectroscopy, mass spectral analysis and other appropriate analytical data.



**Figure 2.10** Structure of synthesized PNA monomers

The rationally designed modified PNA monomers have been incorporated into 10-mer pyrimidine and 9-mer and 12-mer mixed purine pyrimidine PNA sequences. These imidazole modified monomers were inserted into *aeg* PNA sequences at specific positions using HOBt, HBTU and DIPEA as coupling reagent by solid phase peptide synthesis protocol. All the imidazole modified and unmodified PNA oligomers obtained by standard Boc-protocol of 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. The next chapter deals with the investigation of biophysical properties of PNA oligomers.

## 2.6 Experimental section

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

### 2.6.1 General

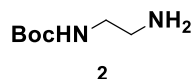
The chemicals used were of laboratory or analytical grade. All the solvents used were distilled or dried to carry out different reactions. Reactions were monitored by thin layer chromatography (TLC). Usual workup involved sequential washing of the organic extract with water and brine followed by drying the organic layer over anhydrous sodium sulphate and evaporation of solvent under vacuum. TLCs were carried out on pre-coated silica gel GF<sub>254</sub> 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).

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded using Bruker AC-200 (200 MHz) or JEOL 400 MHz NMR spectrometers. The delta ( $\delta$ ) 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<sub>2</sub> or 2,5-dihydroxybenzoic acid (DHB) and the integrity of PNA oligomer was checked on the same instrument using DHB or CHCA as matrix. 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.

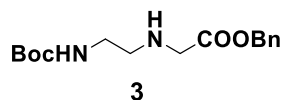
### 2.6.2 Synthesis of compounds

#### ***N*1- (Boc)-1, 2-diaminoethane (2)**

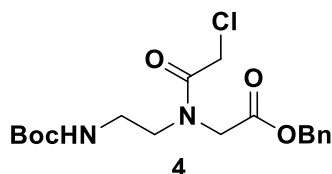


To an ice-cold stirred solution of 1,2-diaminoethane **1** (20g, 0.33 mol) in DCM (300 ml) was added solution of Boc-anhydride ((5 g, 35 mmol) in DCM (50 ml) solution. The mixture was stirred for 12 h and the resulting solution was concentrated to 100 mL. The *N*1, *N*2-di-Boc derivative not being soluble in water, precipitated out and was removed by filtration. The corresponding *N*-mono-Boc derivative was obtained by repeated extraction from the filtrate in dichloromethane. Removal of solvents yielded the mono-Boc-diaminoethane (**5**) (3.40 g, 60%) which was used for further reaction without any purification.

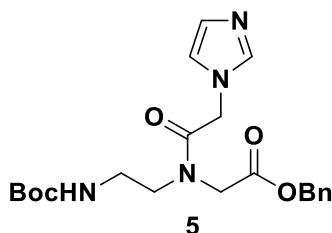
#### **Ethyl *N*-(2-Boc-aminoethyl) glycinate (3)**



To a stirred solution of *N*-(Boc)-1, 2-diaminoethane **2** (3.2 g, 20 mmol) in acetonitrile (100 ml) was added ethyl bromoacetate (2.2 mL, 20 mmol) slowly. Et<sub>3</sub>N (5.5 mL, 40 mmol) was added in the RT. The RM was stirred for 10 h. TLC analysis showed completion of the reaction. After completion of reaction, solvent was removed on rota evaporator. Reaction mixture was then diluted with water (50 mL) and extracted with ethyl acetate (3 × 50 mL), followed by washing with brine. The collected organic layer was dried over anhydrous sodium sulphate and filtered and then concentrated to get crude product. The crude product was purified by silica gel column chromatography using petroleum ether: ethyl acetate (3.5:1.5) as eluent to obtain ethyl *N*-(2-Boc-aminoethyl) glycinate **6** (3.4 g, 70%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: H 7.32-7.29 (m, 5H), 5.12 (s, 2H), 3.43 (s, 2H), 3.20-3.16 (m, 2H), 2.73-2.70 (t, 2H, J=12 Hz), 1.40 (s, 9H) ppm. <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 172.14, 156.11, 135.5, 128.70, 128.70, 128.43, 79.26, 66.77, 50.37, 48.80, 40.07, 28.48 ppm.

**Synthesis of ethyl *N*-(2-Boc-aminoethyl)chloroacetyl glycinate (4)**

Compound **3** (3 g, 11.5 mmol) at 0 °C was dissolved in dioxane (50 mL) and NaHCO<sub>3</sub> (1.9 g, 23 mmol) in water (50 mL) was added to reaction mixture at 0 °C and stirred for 5 min. To this, chloroacetyl chloride (1.0 mL, 12.6mmol) was added dropwise and reaction was continued for stirring at 0 °C for another 5 min. After 5 min, the reaction mixture was brought to room temperature and stirred for 1.5 h. After completion of reaction (TLC), dioxane was removed under reduced pressure. Reaction mixture was then diluted with water (25 mL) and extracted with ethyl acetate (3 × 75 mL). The combined organic layer was washed with saturated solution of NaHCO<sub>3</sub> followed by washing with brine. Organic layer was dried over anhydrous sodium sulphate, collected by filtration and concentrated on rota evaporator. The residual material was purified on column using silica gel in petroleum ether: ethyl acetate (4:1) to obtained **4** as gummy (3.2 g, 85 %). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.37-7.26 (m, 5H), 5.21 (maj) 5.19 (min) (s, 2H), 4.21 (maj) 4.15 (min) (s, 2H), 4.0 (min) 3.98 (maj) (s, 2H), 3.53 (t, 2H, J= 8 Hz), 3.28-3.24 (m, 2H), 1.44 (min) 1.42 (maj) (s, 9H) ppm; <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 169.72, 167.79, 156.14, 135.13, 128.83, 128.72, 128.47, 80.09, 67.53, 49.93, 49.37, 41.17, 40.75, 38.82, 28.47ppm.

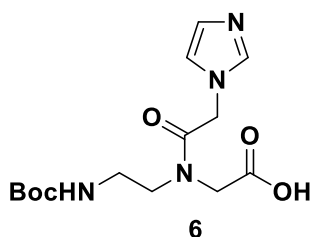
***NH*-Boc-aminoethyl-*N*-(imidazole-N1-acetamido) benzyl glycinate (5)**

To a stirred solution of Compound **4** (3 g, 9.3mmol) in dry DMF (20ml) under inert atmosphere was added dry K<sub>2</sub>CO<sub>3</sub> at 0 °C for 5 min. The imidazole (0.58 g 11.2mmol) was added at RM. The reaction was stirred for 6 h at 60 0 °C. TLC analysis showed completion of the reaction. The hot reaction mixture was cooled at room temperature and then reaction mixture was diluted with water (40 mL) and extracted with ethyl acetate (3



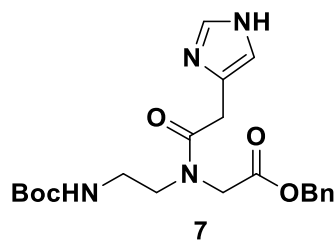
× 75 mL). The combined organic layer was washed with saturated solution of NaHCO<sub>3</sub>(20mL) and then again washed with water (3 × 40 mL) followed by washing with brine (25mL). The collected organic layer was dried over anhydrous sodium sulphate, collected by filtration and concentrated on a rotary evaporator. The crude product was purified on a silica gel column in petroleum ether: ethyl acetate (1.5:3.5) (2.5 g, 75%) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ H 7.38-7.31 (m, 6H), 7.13 (s, 1H), 7.10 (s, 2H), 5.23 (maj) 5.16 (min) (s, 2H), 4.12 (s, 2H), 3.70 (bs, 2H), 3.53 (t, 2H, J= 12 Hz), 3.33-3.29 (m, 2H) 1.40 (s, 9H) ppm. <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 169.39, 166.87, 156.27, 137.58, 135.04, 129, 128.91, 128.78, 128.79, 128.41, 121.32, .88, 67.49, 49.03, 48.96, 48.78, 48.20, 38.60, 28.47. HRMS (ESI-TOF) m/z calcd for C<sub>21</sub>H<sub>28</sub>N<sub>4</sub>O<sub>5</sub> [M + H]<sup>+</sup> 417.2178, found 417.1552 ppm.

### Synthesis of *NH*-Boc-aminoethyl-*N*-(imidazole-*N*1-acetamido) glycine (6)



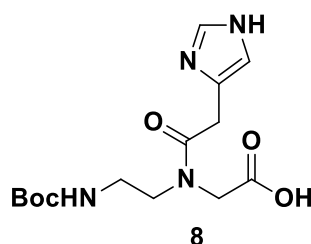
Compound **5** (1 g, 2.8mmol) was dissolved in ethanol and stirred at 0 °C. To this, 10% LiOH (1.5 mL) was dropwise added at the same temperature for 10 min. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on rotary evaporator. Ethyl acetate was added to the residue and it was acidified by addition of saturated aqueous solution of potassium bisulfate till the pH comes down to 3-4. Organic layer was collected and washed with brine solution. The organic layer was concentrated on a rotary evaporator to get solid <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.52 (maj) and 7.49 (min), (s, 1H), 7.02 (maj) 7.00 (min) (s, 1H), 6.87 (maj) 6.84 (min) (s, 1H), 4.99 (maj) 4.80 (min) (s, 2H), 3.87 (maj) 3.85 (min) (s, 2H), 3.41-3.28 (m, 2H), 3.04-2.99 (m, 2H), 1.35 (s, 9H) ppm. <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 156.21 ,1.56.12, 138.73, 135.62, 122.09, 121.38, 78.59, 78.22, 47.42, 38.46, 38.11, 28.66 ppm. HRMS (ESI-TOF) m/z calcd for C<sub>14</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub> [M + H]<sup>+</sup> 327.1668, found 327.1671.

### *NH*-Boc-aminoethyl-*N*-(imidazole-*C*4-acetamido) benzyl glycinate (7)



Imidazole-4-acetic acid monohydrochloride (1.5g 9.1mmol) was activated by coupling reagent EDC (1.6g 8.6mmol) and DIPEA (1.5ml8.1mmol) in DMF under argon atmosphere. Compound **3** (2g 8.1 mmol) dissolved in DMF was then added to the reaction was stirred at room temperature for 6 hrs. DMF was evaporated under reduced pressure and the left residue was extracted with ethyl acetate (30 ml  $\times$  3). Combined organic layer was washed with saturated  $\text{Na}_2\text{HCO}_3$  solution followed by brine. Organic layer was evaporated in a rotary evaporator and the compound was purified by column chromatography. (yield 1.8g, 62%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.49 (s, 1H), 7.36-7.33 (m, 5H), 6.89 (s, 1H), 5.16 (maj) 5.15 (min) (s, 2H), 4.26 (maj) 4.07 (min) (s, 2H), 3.75 (s, 2H), 3.62-3.50 (m, 2H), 3.27-3.22 (m, 2H), 1.40 (min) 1.39 (maj) (s, 9H) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  169.84, 156.4, 135.16, 134.95, 128.74, 128.67, 128.28, 79.68, 67.21, 49.58, 48.81, 38.90, 29.70, 28.40 ppm. HRMS (ESI-TOF)  $m/z$  calcd for  $\text{C}_{21}\text{H}_{28}\text{N}_4\text{O}_5$   $[\text{M} + \text{H}]^+$  417.2138, found 417.2145

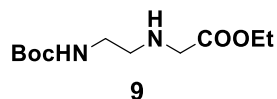
#### ***NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) glycine (**8**)**



Compound **7** (1 g, 5.6 mmol) was dissolved in ethanol and stirred at 0 °C. To this, 10% LiOH (1.5 mL) was drop wise added at the same temperature for 10 min. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on rotary evaporator. Ethyl acetate was added to the residue and it was acidified by addition of saturated aqueous solution of potassium bisulfate till the pH comes down to 3-4. Organic layer was collected and washed with brine solution. The organic layer was concentrated on a rotary evaporator to get solid  $^1\text{H}$  NMR (400 MHz,  $\text{DMSO-d}_6$ )  $\delta$  7.6 (s, 2H), 6.91 (s, 1H), 4.18 (min) 3.92 (maj) (s, 2H), 3.61 (maj) 3.49 (min) (s, 2H), 3.44-3.27

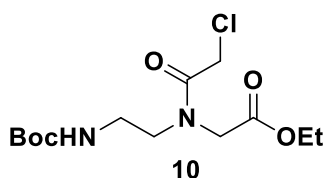
(m, 2H), 3.11-3.00 (m, 2H), 1.36 (s, 9H) ppm.  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta$  171.59, 170.97, 170.67, 156.23, 156.08, 135.22, 78.40, 48.53, 47.86, 38.79, 38.30, 28.71 ppm. HRMS (ESI-TOF)  $m/z$  calcd for  $\text{C}_{14}\text{H}_{22}\text{N}_4\text{O}_5$   $[\text{M} + \text{H}]^+$  327.1668, found 327.1679.

### Ethyl *N*-(2-Boc-aminoethyl) glycinate (**9**)



The *N*-(Boc)-1, 2-diaminoethane **2** (3.2 g, 20 mmol) was treated with ethyl bromoacetate (2.2 mL, 20 mmol) in acetonitrile (100 mL) in the presence of triethylamine (5.5 mL, 40 mmol) and the mixture was stirred at ambient temperature for 10 h. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on a rotary evaporator. Reaction mixture was then diluted with water (50 mL) and extracted with ethyl acetate (3  $\times$  50 mL), followed by washing with brine. The collected organic layer was dried over anhydrous sodium sulphate and filtered and then concentrated to get crude product. The crude product was used further for next reaction without purification.

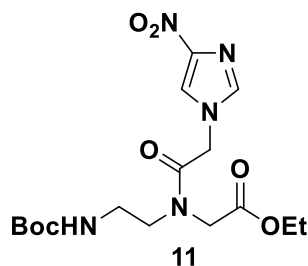
### Ethyl *N*-(2-Boc-aminoethyl) chloroacetyl glycinate (**10**)



Compound **9** (3 g, 11.5 mmol) was dissolved in dioxane (50 mL) at 0 °C and  $\text{NaHCO}_3$  (1.9 g, 23 mmol) in water (50 mL) was added to reaction mixture at 0 °C and stirred for 5 min. To this, chloroacetyl chloride (1.0 mL, 12.6 mmol) was added dropwise and reaction was continued for stirring at 0 °C for another 5 min. After 5 min, the reaction mixture was brought to room temperature and stirred for 1.5 h. After completion of reaction (TLC), dioxane was removed under reduced pressure. Reaction mixture was then diluted with water (25 mL) and extracted with ethyl acetate (3  $\times$  75 mL). The combined organic layer was washed with saturated solution of  $\text{NaHCO}_3$  followed by washing with brine. Organic layer was dried over anhydrous sodium sulphate, collected by filtration and concentrated on a rotary evaporator. The residual material was purified on column using silica gel in petroleum ether: ethyl acetate (4:1) (3.2 g, 85%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : H 5.02

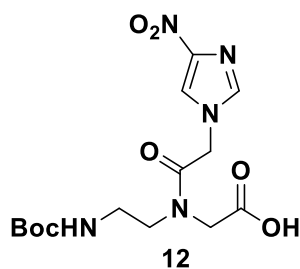
(bs, 1H), 4.5 (s, 2H), 4.22 (q, 2H J=12), 3.35 (s, 2H J=16), 3.20 (t, 2H J=16,) 2.76 (t, 2H), 1.46 (s, 9H), 1.28 (t, 3H J=12);  $^{13}\text{C}$ -NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ :14.1, 28.3, 38.7, 49.7, 61.7, 80.0, 156.1, 167.7, 169.8 ppm.

***NH*-Boc-aminoethyl-N-(4-nitroimidazole-N1-acetamido) ethyl glycinate (11)**



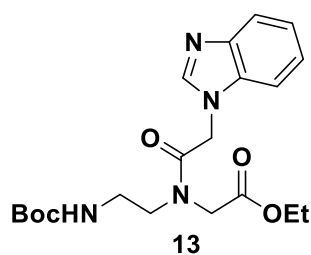
Compound **4** (3 g, 9.3mmol) was dissolved in dry DMF under inert atmosphere (25 mL) and stirred at 0 °C temperature for 5 min. After 5 min, dry  $\text{K}_2\text{CO}_3$  was added and stirring was continued at same temp for another 5 min. Then Imidazole (0.58 g 11.2mmol) was added and stirring continued at same temperature. After 5 min, reaction mixture was heated at 60°C for 6h. Completion of reaction was monitored by TLC. The hot reaction mixture was cooled at room temperature and then reaction mixture was diluted with water (40 mL) and extracted with ethyl acetate (3 × 75 mL). The combined organic layer was washed with saturated solution of  $\text{NaHCO}_3$ (20mL) and then again washed with water (3 × 40 mL) followed by washing with brine (25mL). The collected organic layer was dried over anhydrous sodium sulphate, collected by filtration and concentrated on rota evaporator. The crude product was purified on column using silica gel in petroleum ether: ethyl acetate (1.5:3.5) (Yield 2.5 g, 75%).  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  8.07 (s, 1H), 7.69 (s, 1H), 5.24 (min) 5.04 (maj) (s, 2H), 4.33-4.14 (m, 4H), 3.52-3.31 (m, 4H), 1.42 (min) 1.41 (maj) (s, 9H), 1.27-1.23 (m,3H) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  169.41, 165.92, 162.65, 162.42, 137.30, 121.08, 80.62, 61.99, 48.31, 38.84, 36.58, 31.51, 28.47, 14.19 ppm. HRMS (ESI-TOF) m/z calcd for  $\text{C}_{16}\text{H}_{25}\text{N}_5\text{O}_9$   $[\text{M} + \text{Na}]^+$  422.1654, found 422.1651.

***NH*-Boc-(4-nitro-imidazole-N1-acetamido) aminoethyl glycine (12)**



Compound **11** (1 g, 5.6 mmol) was dissolved in ethanol and stirred at 0 °C. To this, 10% LiOH (1.5 mL) was drop wise added at the same temperature for 10 min. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on rotary evaporator. Ethyl acetate was added to the residue and it was acidified by addition of saturated aqueous solution of potassium bisulfate till the pH comes down to 3-4. Organic layer was collected and washed with brine solution. The organic layer was concentrated on a rotary evaporator to get solid. <sup>1</sup>H NMR (400 MHz, DMSO-d<sup>6</sup>) δ 8.24 (min) 8.23 (maj) (s, 1H), 7.75 (s, 1H), 5.23 (min) 5.02 (maj) (s, 2H), 4.22 (min) 4.00 (maj) (s, 2H), 3.43-3.32 (m, 2H), 3.23-3.05 (m, 2H), 1.37 (s, 9H) ppm. <sup>13</sup>C NMR (100 MHz, DMSO-d<sup>6</sup>) δ 170.76, 166.70, 156.24, 147.05, 138.91, 123.28, 78.60, 48.47, 48.22, 38.60, 38.38, 28.59 ppm. HRMS (ESI-TOF) m/z calcd for C<sub>14</sub>H<sub>21</sub>N<sub>5</sub>O<sub>7</sub> [M - H]<sup>-</sup> 370.1363, found 370.1363.

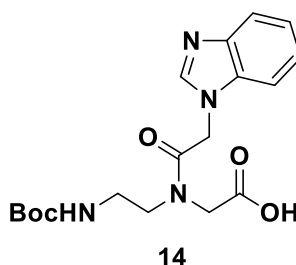
### ***NH*-Boc-ethylamino-N-(benzimidazole N1-acetamido) ethyl glycinate (13)**



Compound **4** (3 g, 9.29 mmol) was dissolved in dry DMF under inert atmosphere (25 mL) and stirred at 0 °C temperature for 5 min. After 5 min, dry K<sub>2</sub>CO<sub>3</sub> was added and stirring was continued at same temp for another 5 min. Then Benzimidazole (1.32 g 11.15 mmol) was added and stirring continued at same temperature. After 5 min, reaction mixture was heated at 60°C for 6h. Completion of reaction was monitored by TLC. The hot reaction mixture was cool at room temperature and then reaction mixture was diluted with water (40 mL) and extracted with ethyl acetate (3 × 75 mL). The combined organic layer was washed with saturated solution of NaHCO<sub>3</sub>(20mL) and then again washed with water (3 ×

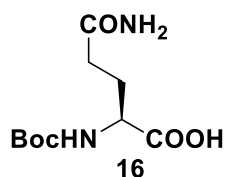
40 mL) followed by washing with brine (25mL). The collected organic layer was dried over anhydrous sodium sulphate, collected by filtration and concentrated on rota evaporator. The crude product was purified on column using silica gel in petroleum ether: ethyl acetate (1.5:3.5) (Yield 2.25 g, 60%)  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  7.92 (min) 7.39 (maj) (s, 1H), 7.79- 7.77 (m, 1H), 7.32- 7.24 (m, 3H), 5.70 (bs, 1H), 5.01 (min) 4.85 (maj) (s, 2H), 4.30-4.14 (m, 2H), 4.01 (s, 2H), 3.56-3.31 (m, 4H), 1.42 (s, 9H), 1.25-1.22 (t, 3H,  $J=12$  Hz) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  169.34, 167.00, 156.25, 143.33, 143.25, 134.24, 123.43, 122.47, 120.34, 109.77, 80.46, 61.91, 49.18, 49.03, 45.03, 45.69, 38.80, 28.47, 14.13 ppm. MS (MALDI-TOF)  $m/z$  calcd for  $\text{C}_{10}\text{H}_{18}\text{N}_2\text{O}_5$   $[\text{M} + \text{K}]^+$  443.16, found 443.18.

***NH*-Boc-aminoethyl-N-(benzimidazole N1-acetamido) glycine (14)**



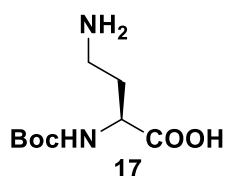
Compound **13** (1 g, 2.47 mmol) was dissolved in ethanol and stirred at 0 °C. To this, 10% LiOH (1.5 mL) was drop wise added at the same temperature for 10 min. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on rotary evaporator. Ethyl acetate was added to the residue and it was acidified by addition of saturated aqueous solution of potassium bisulfate till the pH comes down to 3-4. Organic layer was collected and washed with brine solution. The organic layer was concentrated on a rotary evaporator to get solid (Yield 0.74 g, 80%).  $^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  8.05 -8.01 (s, 1H), 7.39-7.29 (m, 4H), 5.22-5.01 (s, 2H), 4.03 (min) 3.86 (maj) (s, 2H), 3.53 (t, 2H,  $J=12\text{Hz}$ ), 3.42 (t, 2H,  $J=12\text{Hz}$ ), 3.16-3.14 (m, 2H), 1.28 (s, 9H) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{D}_2\text{O}$ )  $\delta$  181.4, 175.7, 178.5, 171.0, 169.4, 158.0, 123.7, 122.8, 122.8, 118.8, 110.6, 48.85, 45.8, 37.4, 36.4, 27.6, 27.6, 23.2 ppm. HRMS (ESI-TOF)  $m/z$  calcd for  $\text{C}_{18}\text{H}_{24}\text{N}_4\text{O}_5$   $[\text{M} + \text{H}]^+$ , found HRMS (ESI-TOF)  $m/z$  calcd for  $\text{C}_{18}\text{H}_{24}\text{N}_4\text{O}_5$   $[\text{M} + \text{H}]^+$  377.1825, found 377.1826

**amino-2-((tert-butoxycarbonyl)amino)-5-oxopentanoic acid (16)**



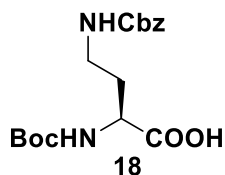
A solution of di-*tert*-butyl dicarbonate [(Boc)<sub>2</sub>O] (16.4 g, 17.26 mL, 75.3 mmol) in dioxane (100 mL) was added in portion to an ice-cold solution of L-glutamine (10 g, 68.5 mmol) in 1 N NaOH (100 mL). The reaction mixture was further stirred at 0 °C for 1 h. After completion of reaction dioxane was removed completely under vacuum from the reaction mixture. The aqueous layer was washed with diethyl ether to remove excess [(Boc)<sub>2</sub>O]. The aqueous layer was cooled in ice-water bath, acidified to pH 2-3 by slow addition of saturated KHSO<sub>4</sub> solution and then extracted with ethyl acetate (3 × 150 mL). The combined organic extracts were dried over an. Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated to give compound **16** as a white powder which was used without further purification (14.5 g, 86% yield). mp = 119-121 °C; R<sub>f</sub> = 0.39 EtOAc/MeOH (50:50); [α]<sup>25</sup><sub>D</sub> - 2.960 (c 0.5, Methanol); MS (MALDI-TOF) *m/z* calcd for C<sub>10</sub>H<sub>18</sub>N<sub>2</sub>O<sub>5</sub> [M + K]<sup>+</sup> 285.0853, found 285.0501.

#### 4-amino-2-((*tert*-butoxycarbonyl)amino)butanoic acid (**17**)



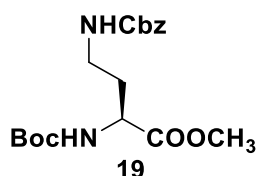
A slurry of compound **16** (5 g, 20.3 mmol), ethyl acetate (24 mL), acetonitrile (24 mL), water (12 mL) and iodobenzene diacetate (7.87 g, 24.4 mmol) was cooled and stirred at 16 °C for 30 min. The temperature was allowed to reach 20 °C and the reaction mixture was stirred until completion (approximately 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** (2.65 g, 65% yield). mp = 200-201 °C; R<sub>f</sub> = 0.2 EtOAc/MeOH (50:50); [α]<sup>25</sup><sub>D</sub> + 13.6 (c 0.5, Methanol); MS (MALDI-TOF) *m/z* calcd for C<sub>9</sub>H<sub>18</sub>N<sub>4</sub>O<sub>4</sub> [M + K]<sup>+</sup> 257.0904, found 257.0740.

#### 4-(((benzyloxy)carbonyl)amino)-2-((*tert*-butoxycarbonyl)amino) butanoic acid (**18**)



The solution of  $\text{NaHCO}_3$  (1.15 g, 15 mL, 13.7 mmol) in water was added to an ice-cold solution of compound **17** (1 g, 4.58 mmol) in acetone (25 mL) and stirred for 10 min at 0 °C. To this, benzylchloroformate (1.87 g 1.85 mL, 5.5 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 ( $2 \times 30$  mL). The aqueous layer was acidified to pH 2-3 with sat.  $\text{KHSO}_4$  solution and extracted with ethyl acetate ( $3 \times 60$  mL). The combined organic layer was dried over an.  $\text{Na}_2\text{SO}_4$ , filtered and concentrated to give compound **18** as sticky oil (1.45 g, 90% yield).  $R_f = 0.67$  EtOAc/MeOH (50:50).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.40-7.31 (m, 5H), 5.65-5.63 (m, 1H), 5.47-5.42 (m, 1H), 5.15-5.03 (m, 2H), 4.34-4.33 (m, 1H), 3.50-3.06 (m, 2H), 2.07-1.76 (m, 2H), 1.43 (s, 9H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  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  $\text{C}_{17}\text{H}_{24}\text{N}_2\text{O}_6$   $[\text{M} + \text{K}]^+$  391.1271, found 391.1075.

**Methyl-4-(((benzyloxy)carbonyl)amino)-2-(((tert-butoxycarbonyl)amino) butanoate (19)**

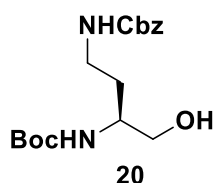


To a stirred solution of compound **18** (5.13 g, 14.5 mmol),  $\text{K}_2\text{CO}_3$  (5.02 g, 36 mmol) in acetone (70 mL) was added dimethyl sulfate (1.7 mL, 17.4 mmol) and reaction mixture was heated to 55 °C for 5 h under reflux condenser. Acetone was evaporated completely and water (90 mL) was added to the concentrate, which was then extracted with ethyl acetate ( $3 \times 50$  mL). The combined organic layer was washed with brine, dried over an.  $\text{Na}_2\text{SO}_4$ , filtered and concentrated. The compound **19** was purified on column chromatography to give as white solid (2.5 g, 93% yield). mp = 65-68 °C;  $R_f = 0.5$  petroleum ether/EtOAc (70:30);  $[\alpha]^{25}_{\text{D}} - 18.8$  ( $c$  0.5, Methanol);  $^1\text{H}$  NMR (200 MHz,  $\text{CDCl}_3$ )  $\delta$  7.38-7.33 (m, 5H), 5.62 (br, 1H), 5.38 (app d,  $J = 8$  Hz, 1H), 5.2-5.12 (m, 2H),



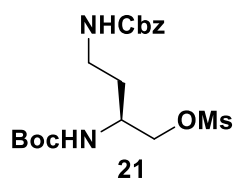
4.43-4.35 (m, 1H), 3.73 (s, 3H), 3.55-3.07 (m, 2H), 2.14-1.66 (m, 2H), 1.45 (s, 9H) ppm.  $^{13}\text{C}$  NMR (50 MHz,  $\text{CDCl}_3$ )  $\delta$  172.9, 156.3, 155.7, 136.5, 128.3, 127.9, 80.0, 66.5, 52.3, 50.8, 37.0, 33.2, 28.1 ppm. MS (MALDI-TOF)  $m/z$  calcd for  $\text{C}_{18}\text{H}_{26}\text{N}_2\text{O}_6$   $[\text{M} + \text{K}]^+$  405.1428, found 405.1199.

### Benzyl tert-butyl (4-hydroxybutane-1,3-diyl)dicarbamate (**20**)



To a stirred solution of compound **19** (2.5 g, 12.9 mmol) in absolute ethanol (30 mL) was added sodium borohydride (0.6 gm, 20.5 mmol) and reaction mixture was stirred for 6 h under nitrogen atmosphere at RT. Ethanol was evaporated completely and water (60 mL) was added to the concentrate which was extracted with ethyl acetate ( $3 \times 30$  mL). The combined organic layer was washed with brine, dried over an.  $\text{Na}_2\text{SO}_4$ , filtered and concentrated. The residue was then purified on column chromatography to give compound **20** as white solid (2.05 g, 89 % yield). mp = 80-82 °C;  $R_f$  = 0.4 petroleum ether/EtOAc (50:50);  $[\alpha]_D^{25}$  - 24.4 ( $c$  0.5, Methanol);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.36-7.29 (m, 5H), 5.65 (br, 1H), 5.13-5.05 (m, 2H), 5.01 (br, 1H), 3.67-3.67 (app d,  $J$  = 8 Hz, 2H), 3.59-3.42 (m, 2H), 3.05-3.0 (m, 1H), 1.77-1.56 (m, 2H), 1.43 (s, 9H) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  155.6, 136.5, 128.4, 128.0, 79.7, 66.6, 65.2, 49.7, 37.6, 32.0, 28.3 ppm. MS (MALDI-TOF)  $m/z$  calcd for  $\text{C}_{17}\text{H}_{26}\text{N}_2\text{O}_5$   $[\text{M} + \text{K}]^+$  377.1479, found 377.1080.

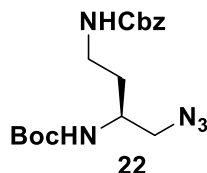
### 4-(benzyloxy)carbonylamino)-2-((tert-butoxycarbonyl)amino)butyl methane sulfonate (**21**)



To an ice-cold solution of compound **20** (2 g, 6 mmol), triethyl amine (2.1 mL, 15 mmol) in dry DCM (30 mL) was added mesyl chloride (0.6 mL, 7.6 mmol) and reaction mixture was stirred for 30 min at 0 °C under nitrogen atmosphere. To the reaction mixture DCM (30 mL) was added which was washed with water (30 mL) and brine (20 mL). The

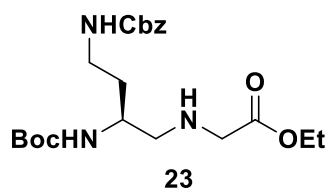
organic layer was dried over an.  $\text{Na}_2\text{SO}_4$ , filtered and concentrated on a rotary evaporator to give compound **21** (2.25 g, 92% crude yield).  $R_f = 0.53$  petroleum ether/EtOAc (50:50). This compound was used for next step without further purification.

**Benzyl tert-butyl (4-azidobutane-1,3-diyl)dicarbamate (22)**



The solution of compound **21** (2.25 g, 5.5 mmol) and sodium azide (5.43 g, 82.5 mmol) in dry DMF (55 mL) was heated for 6 h at 80 °C. To the reaction mixture water (60 mL) was added which was extracted with ethyl acetate (3 × 30 mL). The ethyl acetate layer was washed with water (25 mL) and brine (25 mL). The combined organic layer was dried over an.  $\text{Na}_2\text{SO}_4$ , filtered and concentrated. The residue obtained was then purified on column chromatography to give compound **22** as sticky yellowish oil (1.76 g, 87% yield).  $R_f = 0.73$  petroleum ether/EtOAc (50:50);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.32-7.28 (m, 5H), 5.54 (br, 1H), 5.08-5.02 (m, 2H), 4.83 (br, 1H), 3.81-3.75 (m, 1H), 3.45-3.38 (m, 2H), 3.36-2.98 (m, 2H), 1.69-1.50 (m, 2H), 1.40 (s, 9H) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  156.4, 155.7, 136.5, 128.4, 128.0, 19.8, 66.5, 54.9, 47.6, 37.4, 32.8, 28.1 ppm. MS (MALDI-TOF)  $m/z$  calcd for  $\text{C}_{17}\text{H}_{25}\text{N}_5\text{O}_4$   $[\text{M} + \text{K}]^+$  402.1544, found 402.1709.

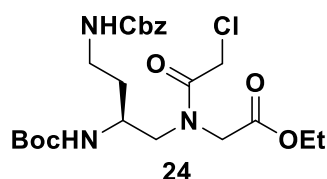
**Ethyl 2-((4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino)butyl)amino)acetate (23)**



To a solution of compound **22** (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  $\text{H}_2$  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

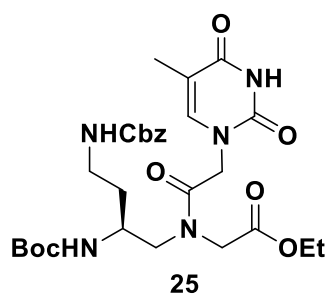
acetonitrile (20 mL) using triethyl amine (0.77 mL, 5.5 mmol) and the reaction mixture was stirred at room temperature for 12 h. Acetonitrile was evaporated completely under vacuum and water (50 mL) was added to the concentrate. The aqueous layer was extracted with ethyl acetate (3 × 40 mL). The combined organic layer was washed with sat. NaHCO<sub>3</sub>, brine, dried over an. Na<sub>2</sub>SO<sub>4</sub>, 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 oil (700 mg, 80%).  $R_f = 0.48$  petroleum ether/EtOAc (20:80); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 7.24-7.21 (m, 5H), 5.77 (br, 1H), 5.05-4.94 (m, 3H), 4.13-4.03 (q,  $J = 8$  Hz, 2H), 3.66-3.62 (m, 1H), 3.42-3.22 (m, 3H), 2.97-2.87 (m, 1H), 2.72-2.56 (m, 3H), 1.64-1.42 (m, 2H), 1.34 (s, 9H), 1.21-1.14 (t,  $J = 7$  Hz, 3H) ppm. <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>) 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 ppm MS (MALDI-TOF)  $m/z$  calcd for C<sub>21</sub>H<sub>33</sub>N<sub>3</sub>O<sub>6</sub> [M + K]<sup>+</sup> 462.2006, found 426.2247.

**Ethyl 2-(N-(4-(((benzyloxy)carbonyl)amino)-2-((tert-butoxycarbonyl)amino) butyl)-2-chloroacetamido)acetate (**24**)**



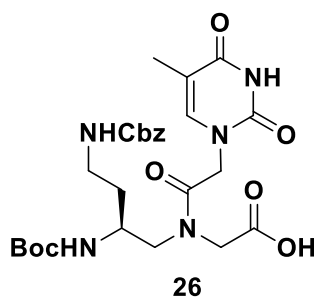
To an ice-cold solution of compound **23** (3.1 g, 7.3 mmol) and triethyl amine (2.96 g; 4 mL, 29.2 mmol) in dry DCM (50 mL) was added chloroacetyl chloride (0.82 g; 0.58 mL, 7.3 mmol) and reaction mixture was stirred for 8 h. To the reaction mixture DCM (20 mL) was added and washed with water (50 mL) and brine (50 mL). The organic layer was dried over an. Na<sub>2</sub>SO<sub>4</sub>, 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 colourless sticky oil (2.63 g, 72 %).  $R_f = 0.59$  petroleum ether/EtOAc (40:60); <sup>1</sup>H NMR (200 MHz, CDCl<sub>3</sub>) δ 7.337.26 (m, 5H), 5.62, 5.34 (br, 1H), 5.15-5.1 (m, 2H), 4.26-4.11 (m, 4H), 3.99 (s, 2H), 3.87-3.66 (m, 2H), 3.53-3.37 (m, 2H), 3.24-2.94 (m, 2H), 1.71-1.63 (m, 2H), 1.41 (maj) 7 1.40 (man) (s, 9H), 1.31-1.21 (m, 3H) ppm. <sup>13</sup>C NMR (50 MHz, CDCl<sub>3</sub>) δ 169.1, 168.8, 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 ppm MS (MALDI-TOF)  $m/z$  calcd for C<sub>23</sub>H<sub>34</sub>ClN<sub>3</sub>O<sub>6</sub> [M + K]<sup>+</sup> 538.1722, found 538.1591.

***NH*-Boc-aminoethyl- $\gamma$ C-(*S*)-(NH-Cbz-aminoethyl)-N-(thyminy-N1-acetamido) ethyl glycinate (25)**



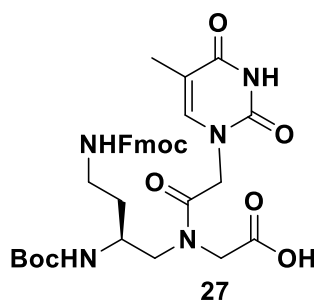
The solution of compound **24** (1 g, 2 mmol),  $K_2CO_3$  (0.33 g, 2.4 mmol) and thymine (0.3 g, 2.4 mmol) in dry DMF (20 mL) was stirred at room temperature for 12 h. To the reaction mixture water (50 mL) was added and extracted with ethyl acetate ( $3 \times 40$  mL). The ethyl acetate layer was washed with water (40 mL) and brine (20 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 **25** as white solid (0.95 g, 81%). mp = 92-94 °C;  $R_f = 0.47$  EtOAc (100);  $[\alpha]_D^{25} - 8.1$  ( $c$  1, Methanol);  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  9.86 (maj) 7 9.56 (min) (br, 1H), 7.37-7.27 (m, 5H), 7.06 (min) 6.99 (maj) (s, 1H), 5.68-5.64 (maj) & 5.58-5.56 (min.) (comp, 1H), 5.11-5.04 (m, 2H), 4.82, 4.77 (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), 2.16 (br, 1H), 1.89 (min) & 1.87 (maj) (s, 3H), 1.69-1.67 (comp, 1H), 1.42 (maj) & 1.39 (min) (s, 9H), 1.31-1.23 (m, 3H) ppm.  $^{13}C$  NMR (100 MHz,  $CDCl_3$ )  $\delta$  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 ppm. MS (MALDI-TOF)  $m/z$  calcd for  $C_{28}H_{39}N_5O_9$   $[M + K]^+$  628.2385, found 628.2361.

***NH*-Boc-aminoethyl- $\gamma$ C-(*S*)-(NH-Cbz-aminoethyl)-N-(thyminy-N1-acetamido) glycine (26)**



To a stirred solution of compound **25** (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<sup>+</sup> 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 **26** as white solid (0.42 g, 88%). mp = 241-245 °C; R<sub>f</sub> = 0.5 EtOAc/MeOH (50:50); [α]<sup>25</sup><sub>D</sub> - 3.0 (c 0.5, Methanol); <sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>) 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) ppm <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) 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 ppm. MS (MALDI-TOF) m/z calcd for C<sub>26</sub>H<sub>35</sub>N<sub>5</sub>O<sub>9</sub> [M + K]<sup>+</sup> 600.2072, found 600.1787.

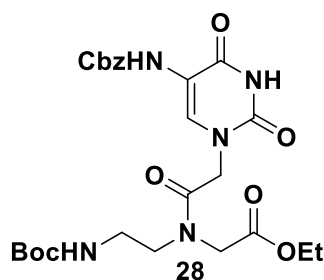
***NH-Boc-aminoethyl-γC-(S)-(NH-Fmoc-aminoethyl)-N-(thyminy-N1-acetamido) glycine (27)***



To the Compound **26** (1 g, 1.78 mmol) dissolved in ethanol, 10% Pd/C on charcoal was added under H<sub>2</sub> atmosphere and stirred at room temperature for 6 h. The completion of reaction was monitored by TLC. The reaction mixture was filtered on celite-545 pad and filtrate was collected, evaporated to get solid product (crude weight 0.8 g). The free amine compound (0.8 g, 1.87 mmol) was dissolved in 10 ml THF:H<sub>2</sub>O (1:1) taken in a clean RBF and stirred at 0 °C. Na<sub>2</sub>CO<sub>3</sub> (0.5 g, 4.67 mmol) was added to the reaction mixture followed by slow dropwise addition of Fmoc-Cl (0.58 g, 2.24 mmol). The reaction mixture was stirred at 0 °C for 1 h and then at 25 °C for 3 h. TLC analysis showed completion of reaction after 4 h. The solvent was evaporated on a rotary

evaporator. The residue was dissolved in 10 ml of water and washed with 20 ml of diethyl ether. The reaction mixture was neutralised with 10% HCl and extracted with ethyl acetate (25 ml×3). The organic layer was concentrated and the residue obtained was purified by column chromatography to yield compound **27** (1.02 g, 85%). <sup>1</sup>H NMR (400 MHz, DMSO-d<sup>6</sup>) δ 11.28 (s, 1H), 7.89-7.87 (d, 2H, J = 8 Hz), 7.68- 7.67 (d, 2H, J = 4 Hz), 7.40 (t, 2H, J = 16 Hz), 7.32 (t, 2H, J = 16 Hz), 7.21 (s, 1H), 7.17 (s, 1H), 4.80-4.46 (m, 2H), 4.31-4.20 (m, 4H), 4.03-3.60 (m, 3H), 3.23-2.97 (m, 4H), 4.03-360 (m, 3H), 3.23-2.97 (m, 4H), 1.74 (maj) 1.72 (min) (S, 2H), 1.37 (maj) 1.35 (min) (s, 9H) ppm. <sup>13</sup>C NMR (100 MHz, DMSO-d<sup>6</sup>) δ 171, 168.31, 167.70, 156.10, 151.29, 144.43, 144.31, 141.19, 128.06, 127.53, 125.56, 120.57, 106.61 ppm. HRMS (ESI-TOF) m/z calcd for C<sub>28</sub>H<sub>39</sub>N<sub>5</sub>O<sub>9</sub> [M + Na]<sup>+</sup> 672.2645, found 672.2612.

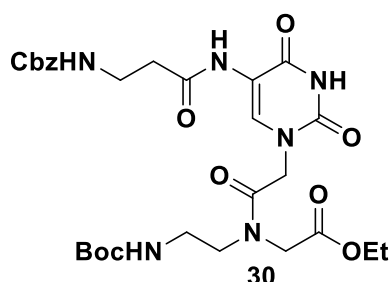
### ***NH*-Cbz-5-amino uracil ethyl glycinate (**28**)**



Compound **4** (3 g, 9.3 mmol) was dissolved in dry DMF under inert atmosphere (25 mL) and stirred at 0 °C temperature for 5 min. After 5 min, dry K<sub>2</sub>CO<sub>3</sub> and 5-benzyloxyaminouracil (2 g, 9.3 mmol) was added and stirring continued at same temperature. After 5 min, reaction mixture was moved to oil bath and heated at 60 °C for 2 h. Completion of reaction was monitored by TLC. The hot reaction mixture was allowed to cool at room temperature and then reaction mixture was diluted with water (40 mL) and extracted with ethyl acetate (3 × 75 mL). The combined organic layer was washed with saturated solution of NaHCO<sub>3</sub> (20 mL) and then again washed with water (3 × 40 mL) followed by washing with brine (25 mL). The collected organic layer was dried over anhydrous sodium sulphate, collected by filtration and concentrated on a rotary evaporator. The crude product was purified on column using silica gel in petroleum ether: ethyl acetate (1.5:3.5) (3.5 g, 75%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.82 (bs, 1H), 8.03 (bs, 1H), 7.44-7.33 (m, 5H), 7.17 (maj) 7.13 (min) (s, 1H), 5.64(bs, 1H), 5.16 (maj) 5.15 (min), (s, 2H), 4.61 (min) 4.47 (maj) (s, 2H,) 4.28 (maj) 4.20 (min) (q, 2H, J = 8 Hz), 4.05

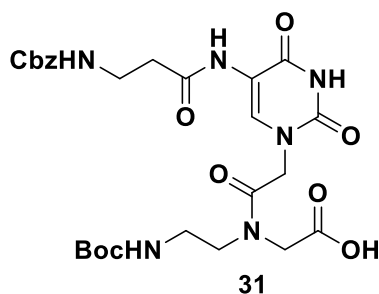
(s, 2H), 3.53 (t, 2H,  $J = 6$  Hz), 3.33 (t, 2H,  $J = 8$  Hz), 1.45 (min) 1.43 (maj) (s, 9H), 1.32 (maj) 1.27 (min) (t, 3H,  $J = 6$  Hz) ppm.  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$ : 169.4, 169.0, 167.03, 159.7, 156.1, 152.9, 149.1, 135.7, 129.6, 128.5, 128.1, 114.8, 80.0, 67.3, 62.3, 61.6, 48.8, 38.8, 28.3, 14.0 ppm. MS (HRMS-ESI):  $m/z$  calcd for  $\text{C}_{25}\text{H}_{33}\text{N}_5\text{O}_9\text{Na}$  547.2278, found 570.2175 (M+Na), 448.1895 (M-Boc).

***NH*-Boc-aminoethyl-*N*-(*NH*-Cbz- $\beta$ -alanyl-5-aminouracil-*N*1-acetamido) ethyl glycinate (**30**)**



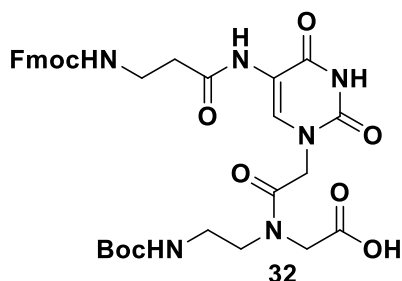
To the Compound **29** (1.7 g, 2.7 mmol) dissolved in ethanol, 10% Pd/C on charcoal was added under  $\text{H}_2$  atmosphere and stirred at room temperature for 6 h. The completion of reaction was monitored by TLC. The reaction mixture was filtered on celite-545 pad and filtrate was collected, evaporated to get solid product. Free amine intermediate (1.5 g, 3.6 mmol) was coupled with protected  $\beta$ -alanine (0.9 g 4.32 mmol) was activated by coupling reagent EDC (0.8 g 4.32 mmol), HOBT (0.6 g, 4.32 mmol) and DIPEA (0.7 ml, 4.32 mmol) in DMF under argon atmosphere. The reaction mixture was stirred at room temperature for 6 hrs. DMF was evaporated under reduced pressure and the left residue was extracted with ethyl acetate (30 ml  $\times$  3). Combined organic layer was washed with saturated  $\text{Na}_2\text{HCO}_3$  solution followed by brine. Organic layer was evaporated on a rotary evaporator and the compound was purified by column chromatography to get **30** as white solid (1.5 g, 60%). HRMS (ESI-TOF)  $m/z$  calcd for  $\text{C}_{28}\text{H}_{38}\text{N}_6\text{O}_{10}$   $[\text{M} + \text{Na}]^+$  641.2546, found 641.2548.

**(*N*-Boc-aminoethylglycyl)-5-(*N*-Cbz- $\beta$ -alanyl) uracil acetamide (**31**)**



To a stirred solution of compound **30** (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 10 % of HCl till pH of the solution turned ~ 5-6 and extracted with ethyl acetate (25 ml×3). The organic layer was concentrated and the residue obtained was purified on column chromatography to yield compound **31** (0.43 g, 90%). HRMS (ESI-TOF)  $m/z$  calcd for  $C_{33}H_{38}N_6O_{10}$   $[M + Na]^+$  701.2547, found 701.2546.

***NH*-Boc-aminoethyl-*N*-[5-amino(*NH*-Fmoc- $\beta$ -alanyl) uracil *N*1-acetamido] glycine (32)**



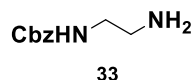
To the Compound **31** (1 g, 1.69 mmol) dissolved in ethanol, 10 % Pd/C on charcoal was added under  $H_2$  atmosphere and stirred at room temperature for 6 h. The completion of reaction was monitored by TLC. The reaction mixture was filtered on celite-545 pad and filtrate was collected, evaporated to get solid product (crude weight 0.8 g). The free amine compound (0.8 g, 1.75 mmol) was dissolved in 10 ml THF:H<sub>2</sub>O (1:1) taken in a clean RBF and stirred at 0 °C.  $Na_2CO_3$  (0.45 g, 4.3 mmol) was added to the reaction mixture followed by slow dropwise addition of Fmoc-Cl (0.54 g, 2.1 mmol). The reaction mixture was stirred at 0 °C for 1 h and then at 25 0 °C for 3 h. TLC analysis showed completion of reaction after 4 h. The solvent was evaporated on a rotary evaporator. The residue was dissolved in 10 ml of water and washed with 20 ml of diethyl ether. The



reaction mixture was neutralised with 10 % HCl and extracted with ethyl acetate (25 ml×3). The organic layer was concentrated and the residue obtained was purified on column chromatography to yield compound **32** (1.12 g, 80%).

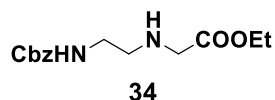
$^1\text{H}$  NMR (400 MHz, DMSO- $d_6$ )  $\delta$  9.25 (s, 1H), 8.23 (maj.) 8.17 (man.), (s, 1H), 7.92-8.91 (d, 2H,  $J = 4$  Hz) 7.72-7.70 (d, 2H,  $J = 8$ Hz), 7.46-7.42 (t, 2H,  $J = 16$  Hz), 7.37-7.34 (t, 2H,  $J = 12$  Hz), 4.70 (maj) 4.53 (min) (s, 2H), 4.31-4.23 (m, 2H), 3.87-3.80 (s, 2H), 3.40-3.25 (m, 6H), 3.06-3.05 (s, 2H), 1.41 (maj) 1.38 (min) (s, 9H) ppm.  $^{13}\text{C}$  NMR (100 MHz, DMSO- $d_6$ )  $\delta$  170.2, 160.7, 156.4, 149.8, 144.3, 141.2, 139.9, 137.9, 129.4, 128.1, 127.7, 127.5, 125.6, 121.8, 120.5, 120.5, 113.8, 110.2, 78.0, 65.8, 49.1, 47.1, 36.2, 28.7, 28.6 ppm. HRMS (ESI-TOF)  $m/z$  calcd for  $\text{C}_{33}\text{H}_{38}\text{N}_6\text{O}_{10}$   $[\text{M} + \text{Na}]^+$  701.2546, found 701.2524.

### ***N*1-Cbz-1, 2-diaminoethane (33)**



1, 2-diaminoethane **1** (20 g, 330 mmol) was taken in DCM (300 mL) and cooled in an ice-bath. Cbz-Cl (4.71 ml, 33 mmol) in DCM (50 mL) was slowly added with stirring. The mixture was stirred for 12 h and the resulting solution was concentrated to 100 mL. The *N*1, *N*2-di-Cbz derivative not being soluble in water, precipitated out and was removed by filtration. The corresponding *N*-mono-Cbz derivative was obtained by repeated extraction from the filtrate in dichloromethane. Removal of solvents yielded the mono-Cbz-diaminoethane (**5**) (4.0 g, 60%) which was used for further reaction without any purification.

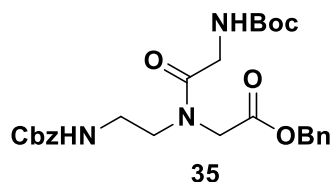
### **Ethyl *N*-(2-Cbz-aminoethyl) glycinate (34)**



The *N*-(Cbz)-1, 2-diaminoethane **33** (3.2 g, 16.5 mmol) was reacted with ethyl bromoacetate (3.13 mL, 19.8 mmol) in acetonitrile (100 mL) in the presence of triethylamine (5.74 mL, 41.2 mmol) and the mixture was stirred at ambient temperature for 10 h. The completion of reaction was monitored by TLC. After completion of reaction, solvent was removed on rota evaporator. Reaction mixture was then diluted with water (50 mL) and extracted with ethyl acetate (3 × 50 mL), followed by washing with

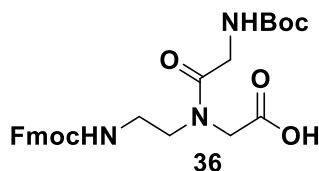
brine. The collected organic layer was dried over anhydrous sodium sulphate and filtered and then concentrated to get crude product. The crude product was used further for next reaction without purification.

***N*-(Cbz-aminoethylglycyl)-*N*-Boc-acetyl benzyl ester (**35**)**



Boc-protected glycine (1.5g 7.01 mmol) was activated by coupling reagent EDC (1.33 g 7.01 mmol), HOBt (0.95 g, 7.01 mmol) and DIPEA (1.22 ml, 7.01 mmol) in DMF under argon atmosphere. Compound **34** (2 g, 5.84 mmol) dissolved in DMF was then added to the reaction was stirred at room temperature for 6 h. DMF was evaporated under reduced pressure and the left residue was extracted with ethyl acetate (30 ml × 3). Combined organic layer was washed with saturated Na<sub>2</sub>HCO<sub>3</sub> solution followed by brine. Organic layer was evaporated in rotavapour and the compound was purified by column chromatography to obtained **35** as white solid. (1.17 g, 80%)

***NH*-Fmoc-aminoethyl-*N*1-(*NH*-Boc-glycinamido) glycine (**36**)**



To the stirred solution of **35** (1 g, 2.22 mmol) in ethanol (10ml) was added 10% Pd/C at room temperature. The R. M. was stirred for 6h under H<sub>2</sub> atm. The completion of reaction was monitored by TLC. The reaction mixture was filtered on celite-545 pad and filtrate was collected, evaporated to get solid product (crude weight 0.8 g). The free amine compound (0.8 g, 2.9 mmol) was dissolved in 10 ml THF:H<sub>2</sub>O (1:1) taken in a clean RBF and stirred at 0 °C. Na<sub>2</sub>CO<sub>3</sub> (0.73 g, 7.25 mmol) was added to the reaction mixture followed by slow dropwise addition of Fmoc-Cl (0.9 g, 3.48 mmol). The reaction mixture was stirred at 0 °C for 1 h and then at 25 0 °C for 3 h. TLC analysis showed completion of reaction after 4 h. The solvent was evaporated on a rotary evaporator. The residue was dissolved in 10 ml of water and washed with 20 ml of diethyl ether. The reaction mixture

was neutralised with 10 % HCl and extracted with ethyl acetate (25 ml×3). The organic layer was concentrated and the residue obtained was purified on column chromatography to yield compound **36** as white solid (1.45 g, 80 %). <sup>1</sup>H NMR (400 MHz, DMSO-d<sup>6</sup>) δ 7.77-7.75 (d, 2H, J=8Hz), 7.60-7.58 (d, 2H, J=8Hz), 7.42-7.38 (t, 2H, J=16 Hz), 7.34-7.31 (m, 2H), 4.41-4.35 (m, 2H), 4.23-4.06 (m, 4H), 3.96-3.81 (m, 2H), 3.56-3.38 (m, 4H), 1.45 (min) 1.42 (maj) (s, 9H) ppm. <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 143.90, 141.28, 127.90, 127.69, 127.33, 127.08, 126.13, 125.15, 128.08, 119.97, 61.11, 67.04, 66.94, 47.16, 47.13, 38.90, 31.18, 28.33 ppm. HRMS (ESI-TOF) m/z calcd for C<sub>26</sub>H<sub>31</sub>N<sub>3</sub>O<sub>7</sub> [M + H]<sup>+</sup> 498.2240, found 498.2232

### 2.6.3 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 having 0.35 mmol/g loading value. 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 × 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 × 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

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

The MBHA resin (10 mg) with oligomers attached to it was stirred with thioanisole (20  $\mu\text{L}$ ) and 1, 2-ethanedithiol (8  $\mu\text{L}$ ) in an ice bath for 10 min. TFA (200  $\mu\text{L}$ ) was added to it in cooled condition and kept in ice bath. TFMSA (16  $\mu\text{L}$ ) 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 a rotary 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.

#### 2.6.5 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  $\text{CH}_3\text{CN}:\text{H}_2\text{O}$  (5:95); B= 0.1% TFA in  $\text{CH}_3\text{CN}:\text{H}_2\text{O}$  (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.

### 2.7 References

1. Burgess, R. R. RNA Polymerase. *Annu. Rev. Biochem.* **1971**, *40*, 770.
2. Raines, R. T. Ribonuclease A. *Chem. Rev.* **1998**, *98*, 1045-1065.
3. Court, D. L.; Gan, J.; Liang, Y.; Shaw, G. X.; Tropea, J. E.; Costantino, N.; Waugh, D. S.; Ji, X. RNase III: Genetics and Function; Structure and Mechanism. *Annu. Rev. Genet.* **1998**, *67*, 153–180.
4. Frank, D. N.; Pace, N. R. Ribonuclease P: Unity and Diversity in a tRNA Processing Ribozyme. *Annu. Rev. Biochem.* **1998**, *67*, 153–80.

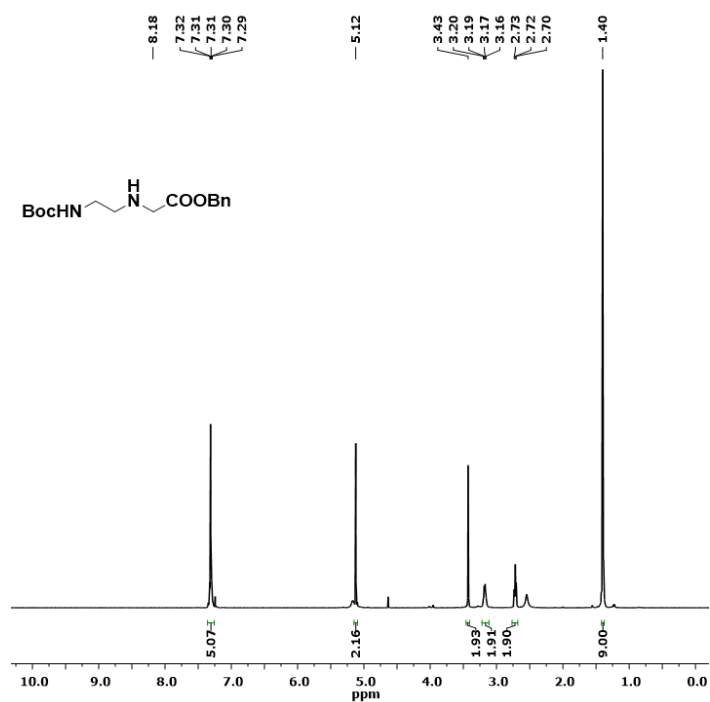
5. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 4096.
6. Choudhury, R.; Tsai, Y. S.; Dominguez, D.; Wang, Y.; Wang, Z. Engineering RNA endonucleases with customized sequence specificities. *Nat. Commun.* **2012**, *3*, 1-8.
7. Waldmann T. A. Effective cancer therapy through immunomodulation. *Annu. Rev. Med.* **2006**, *57*, 65–81.
8. Lachelf, U.; Wagner, E. Nucleic Acid Therapeutics Using Polyplexes: A Journey of 50 Years (and Beyond). *Chem. Rev.* **2015**, *115*, 11043–11078.
9. Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737-738.
10. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide *Science* **1991**, *254*, 1497-1500.
11. Ganesh, K. N.; Krishnan Y. Nucleic Acids – Chemistry and Applications *J. Org. Chem.* **2013**, *78*, 12283-12287.
12. (a) Prakash, T. P.; Kunte, S. S.; Ganesh, K. N. Self-cleavage of C8-histamino-r(UpA) promoted by ZnCl<sub>2</sub>: mechanistic studies on a designed ribonuclease mimic. *Tetrahedron* **1994**, *50*, 11699–11708. (b) Prakash, T. P.; Ganesh, K. N. Ribonuclease mimic: Zn<sup>2+</sup> promoted cleavage of C8-histamino-r(UpA) proceeds through 2',3'-cUMP as intermediate. *J. Chem. Soc., Chem. Commun.* **1994**, 1357–1358.
13. Nielsen, P. E. Addressing the challenges of cellular delivery and bioavailability of peptide nucleic acids (PNA). *Q. Rev. Biophys.* **2005**, *38*, 345–350
14. Breslow, R. Bifunctional acid-base catalysis by imidazole groups in enzyme mimics. *J. Mol. Catal.* **1994**, *91*, 161-174.
15. Laine, M.; Lönnberg, T.; Helkearo, M.; Lönnberg, H. Cleavage of short oligoribonucleotides by a Zn<sup>2+</sup> binding multi-nucleating azacrown conjugate. *Inorg. Chim. Acta.* **2016**, *452*, 111-117.
16. Kameshima, W.; Ishizuka, T.; Minoshima, M.; Yamamoto, M.; Sugiyama, H.; Xu, Y.; Komiyama, M. Conjugation of Peptide Nucleic Acid with a Pyrrole/Imidazole Polyamide to Specifically Recognize and Cleave DNA. *Angew. Chem. Int. Ed.* **2013**, *52*, 13681-13684.

17. Kuzuya, A.; Machida, K.; Shi, Y.; Tanaka, K.; Komiyama, M. Site-Selective RNA Activation by Acridine-Modified Oligodeoxynucleotides in Metal-Ion Catalyzed Hydrolysis: A Comprehensive Study. *ACS Omega*. **2017**, *2*, 5370–5377.
18. Murtola, M.; Wenska M.; Stromberg, R. PNAzymes That Are Artificial RNA Restriction Enzymes. *J. Am. Chem. Soc.* **2010**, *132*, 8984-8990.
19. Dogandziyski, P.; Ghidini, A.; Danneberg, F.; Strömberg, R.; Göbel, M. W. Studies on Tris(2-aminobenzimidazole)-PNA Based Artificial Nucleases: A Comparison of Two Analytical Techniques. *Bioconjugate Chem.* **2015**, *26*, 2514-2519.
20. Devi, G. Ph.D thesis, Savitribai Phule Pune University, **2009**
21. Challa, H.; Styers, M. L.; Woski, S. A. Nitroazole Universal Bases in Peptide Nucleic Acids. *Org. Lett.* **1999**, *1*, 1639-1641.
22. Huang, X.; Seid, M.; Keillor, J. W.; L. H.; Kauffman, G. S.; Pesti, J. A.; Yin, J. A Mild and Efficient Modified Hofmann Rearrangement. *J. Org. Chem.* **1997**, *62*, 7495-7496.
23. (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**.
24. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide1. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
25. 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.
26. (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.
27. (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.
28. 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. Solid-phase synthesis of peptide nucleic acids. *J. Peptide Sci.* **1995**, *3*, 175-183.

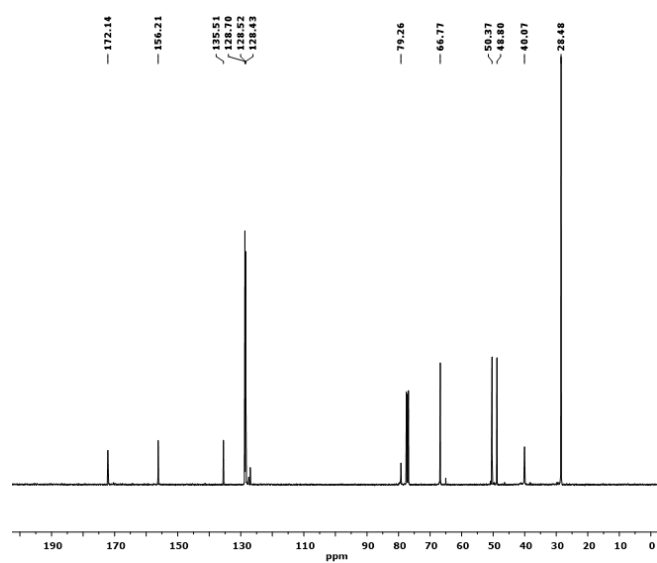
## 2.8 Appendix I: Characterisation data of synthesized compounds and PNA

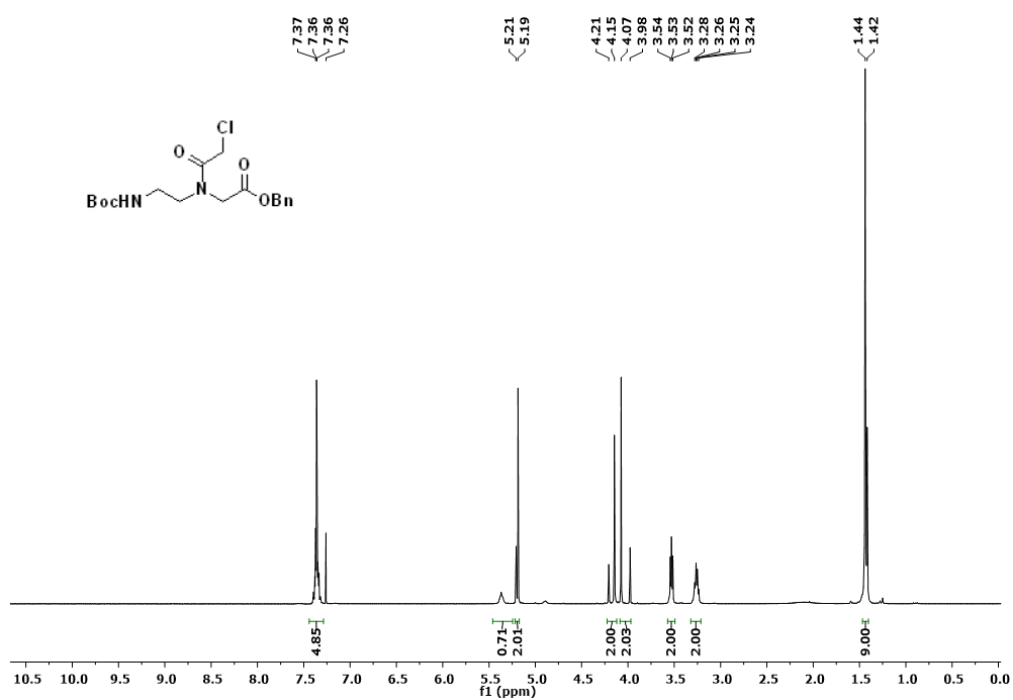
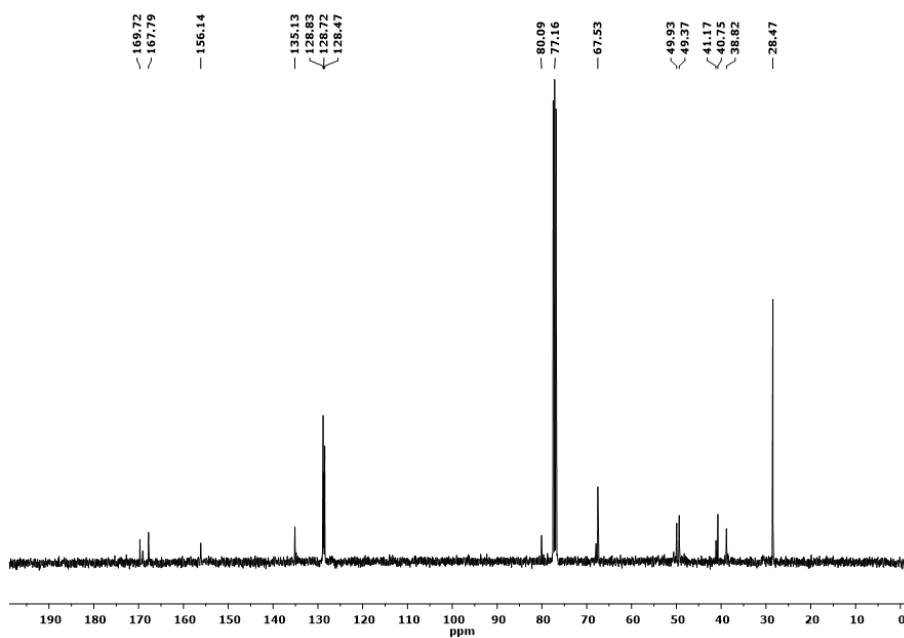
### $^1\text{H}$ and $^{13}\text{C}$ NMR of Compounds

#### $^1\text{H}$ NMR of Compound 3

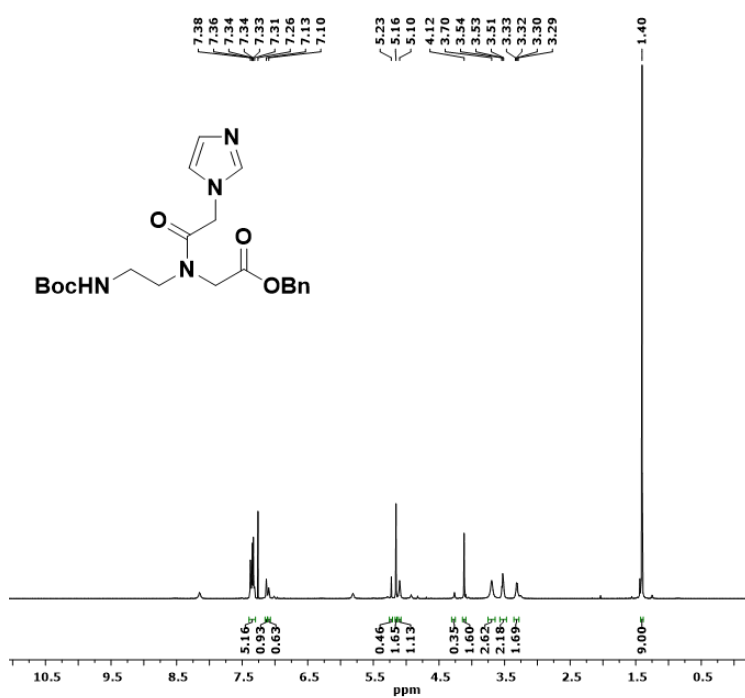
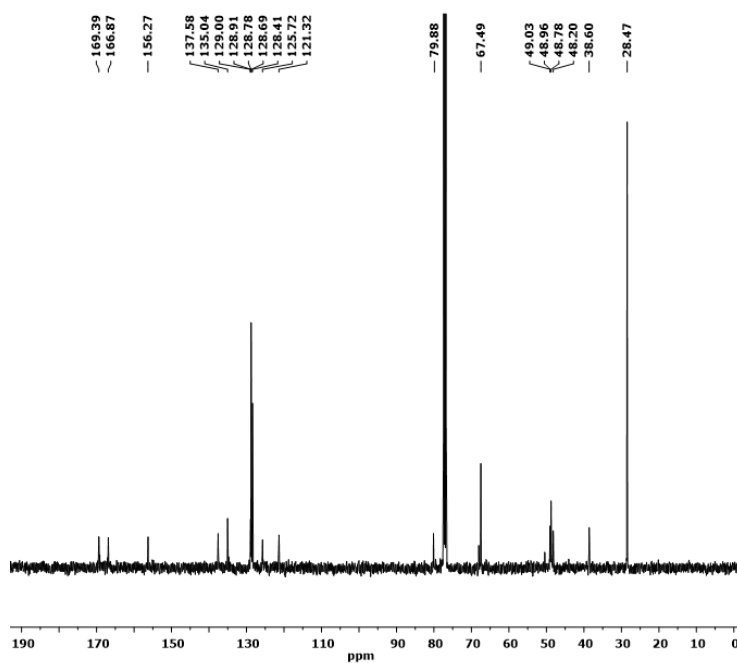


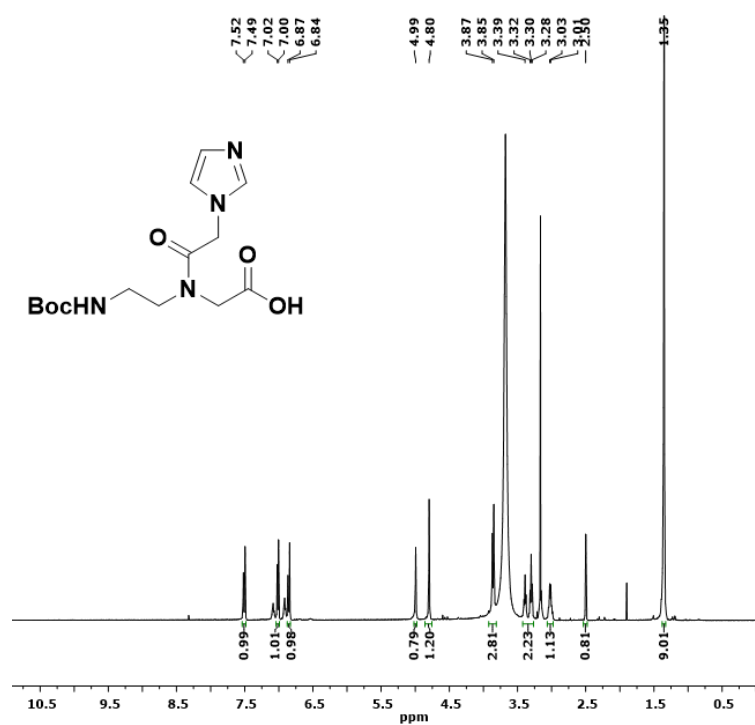
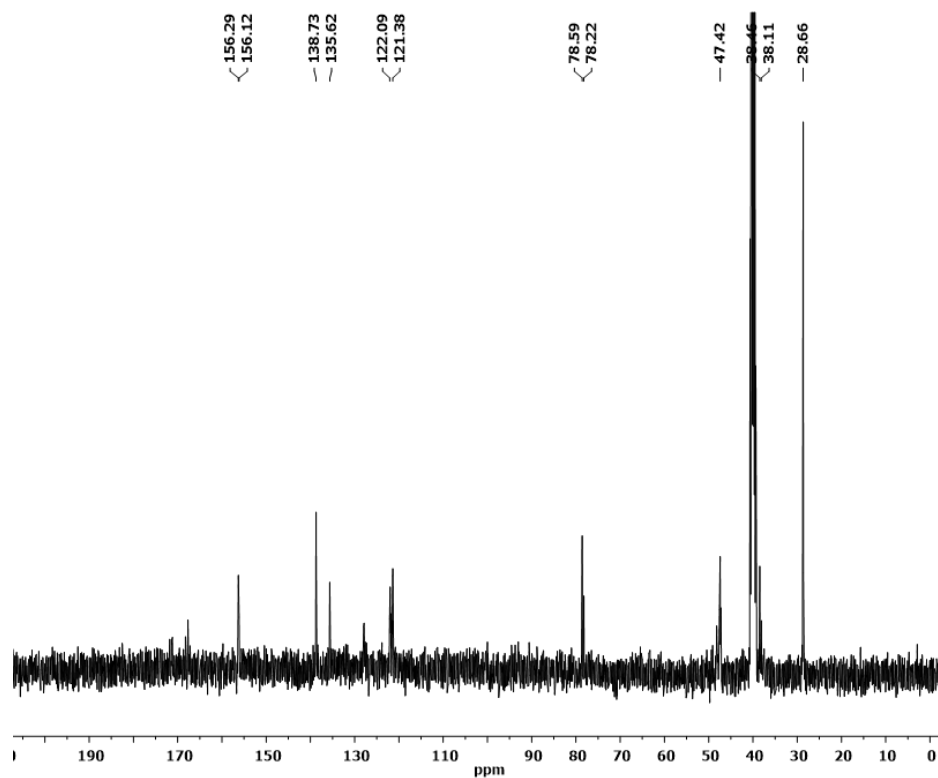
#### $^{13}\text{C}$ NMR of Compound 3

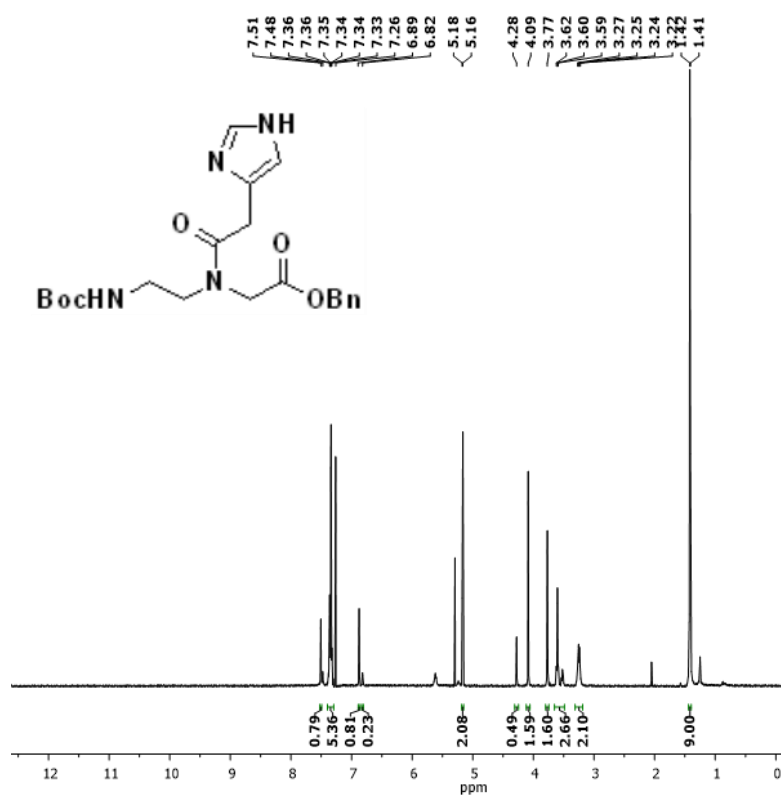
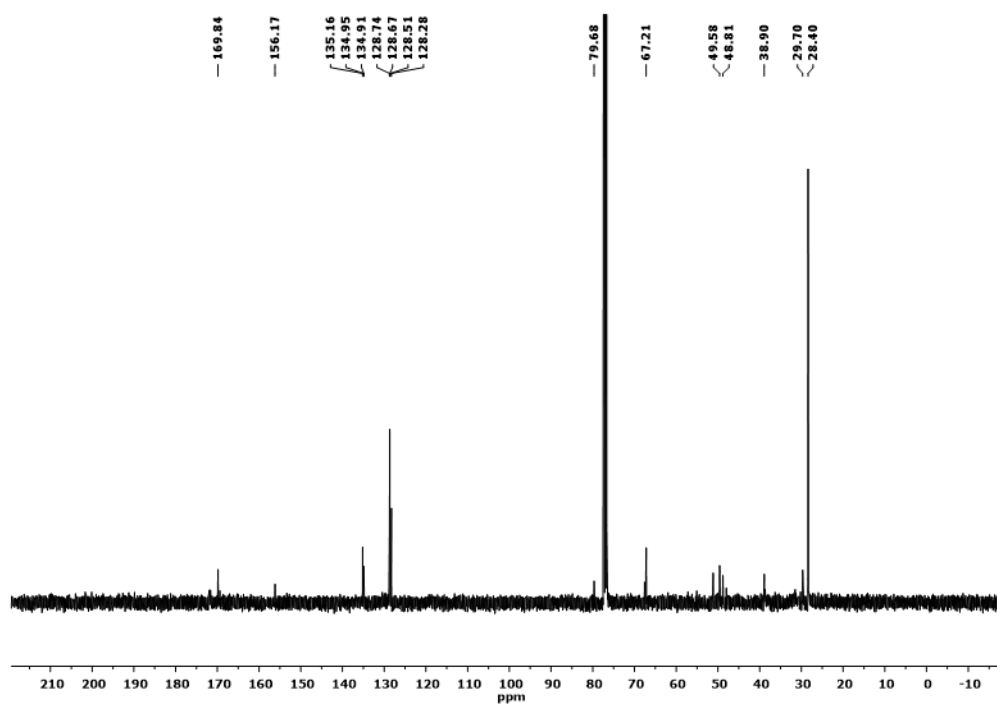


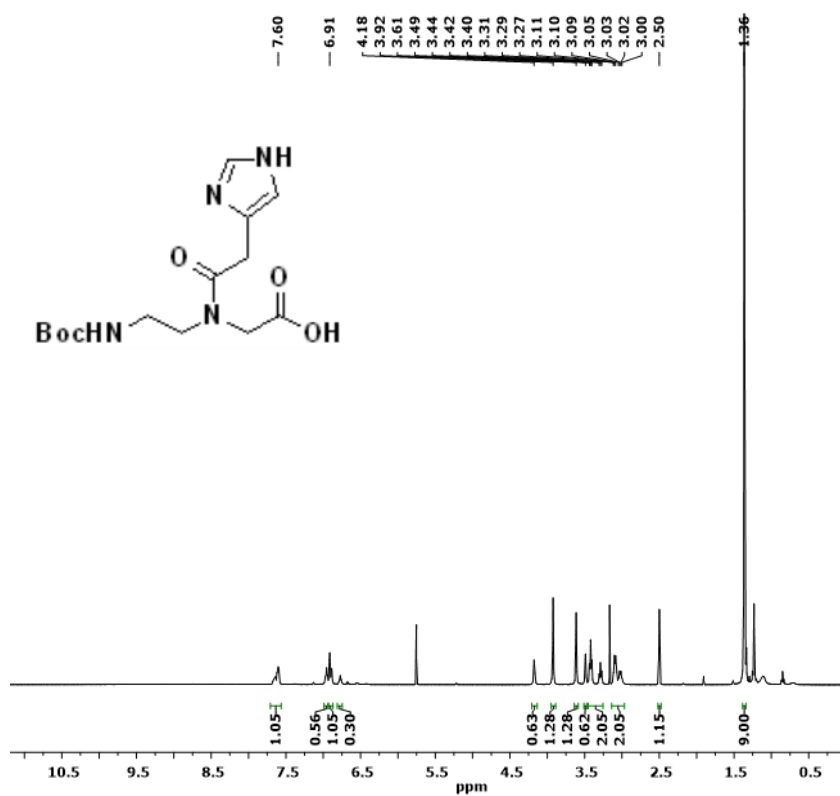
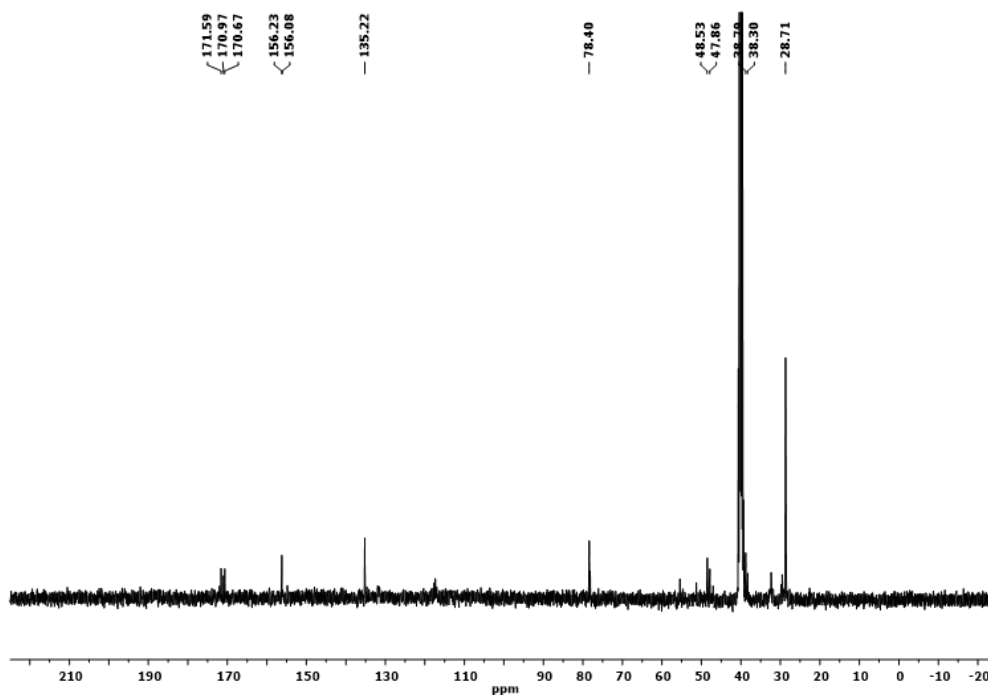
$^1\text{H}$  NMR of Compound 4 $^{13}\text{C}$  NMR of Compound 4

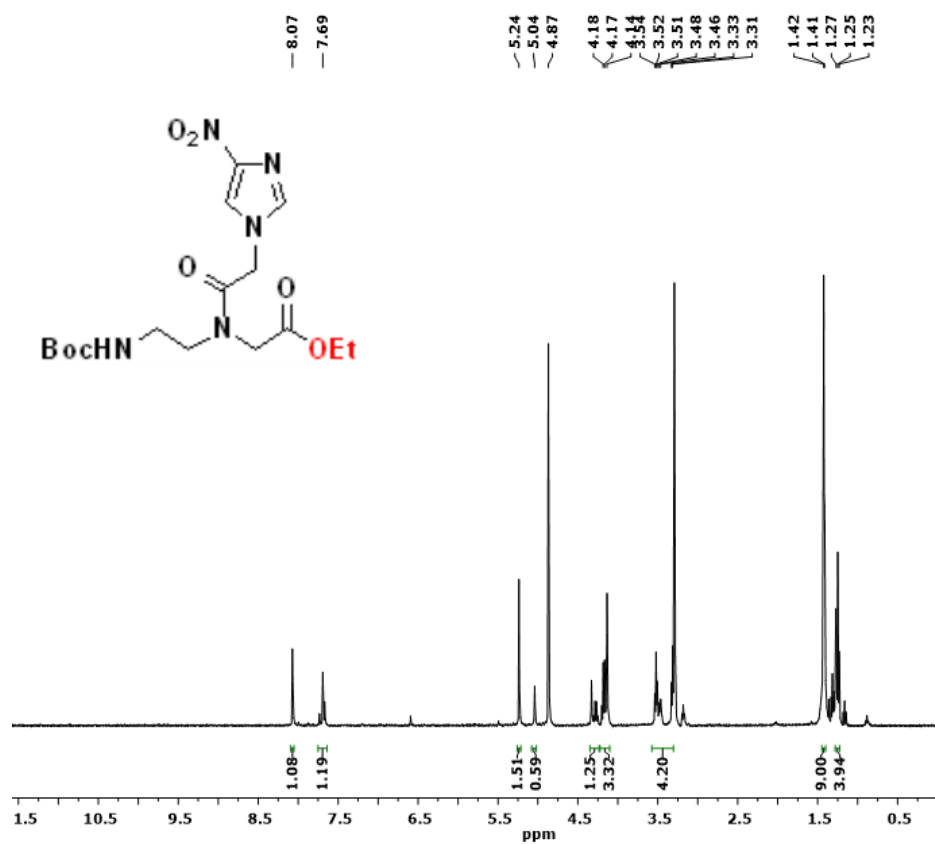
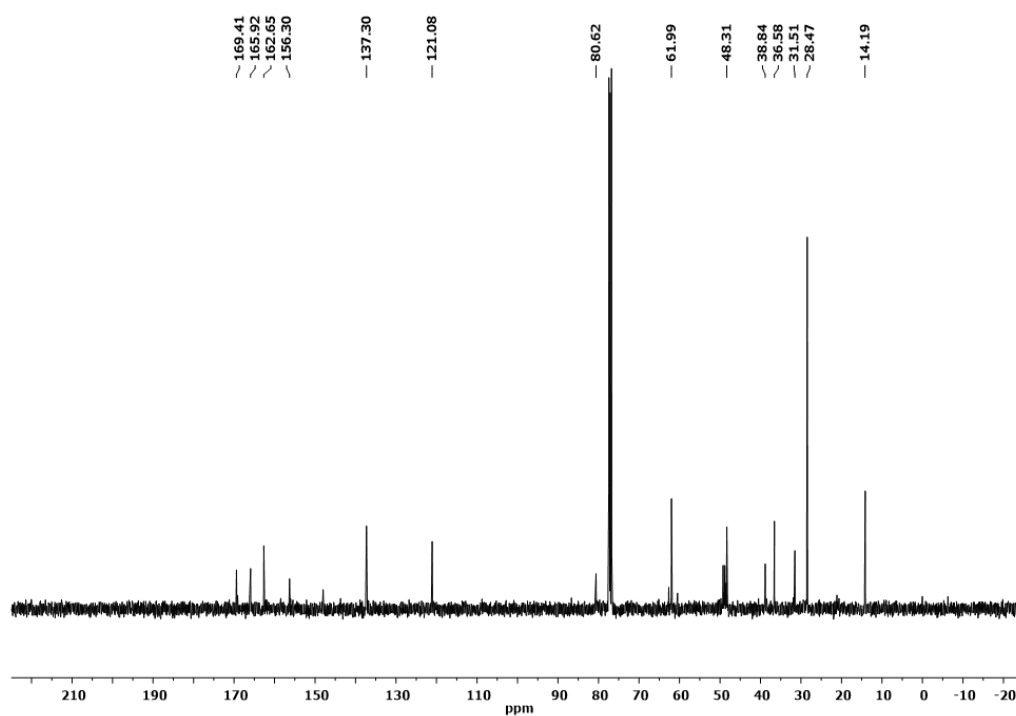


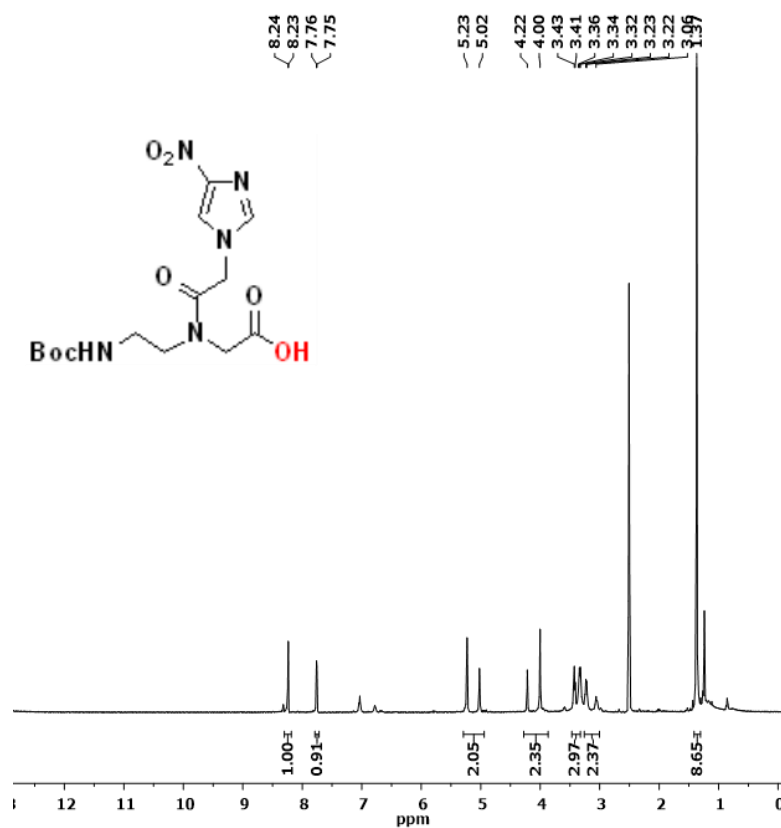
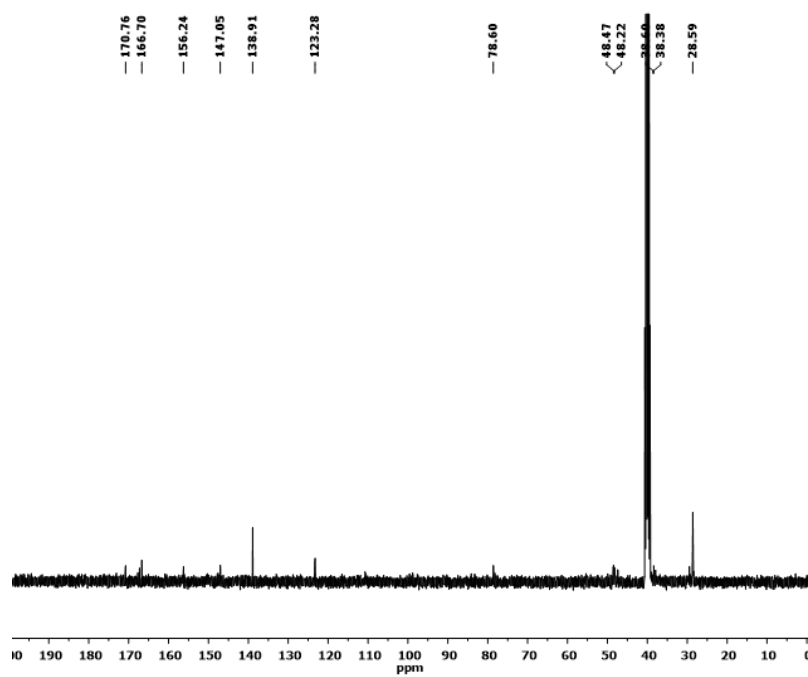
$^1\text{H}$  NMR of Compound **5** $^{13}\text{C}$  NMR of Compound **5**

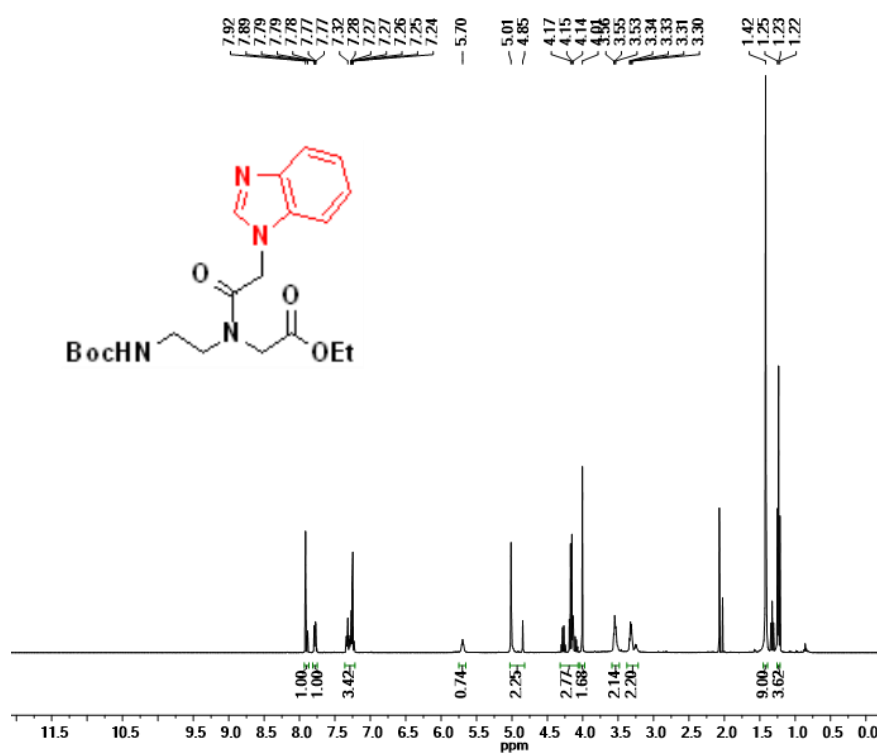
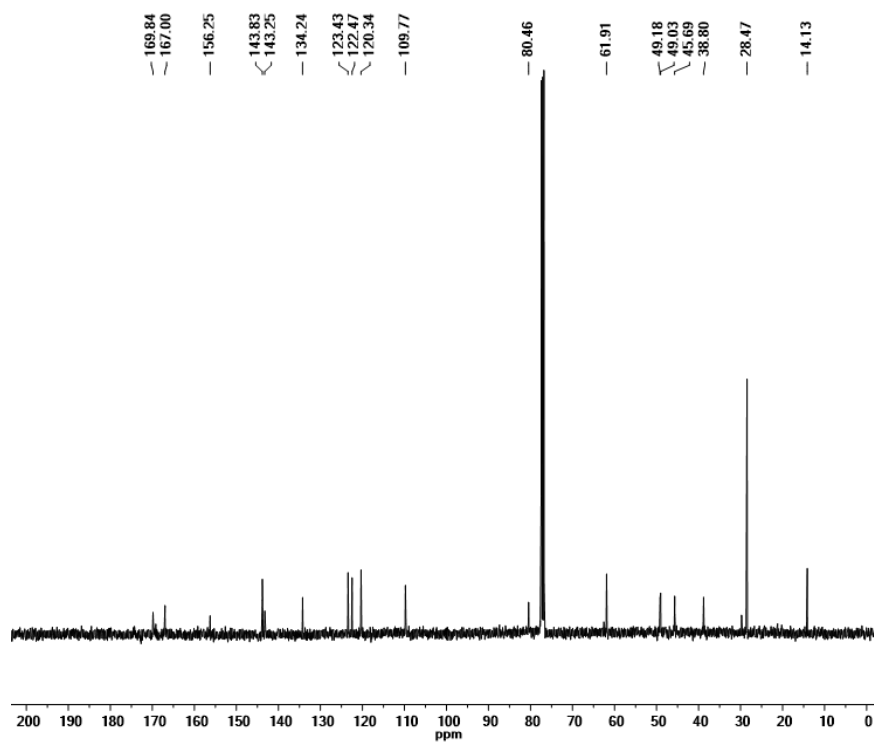
$^1\text{H}$  NMR of Compound **6** $^{13}\text{C}$  NMR of Compound **6**

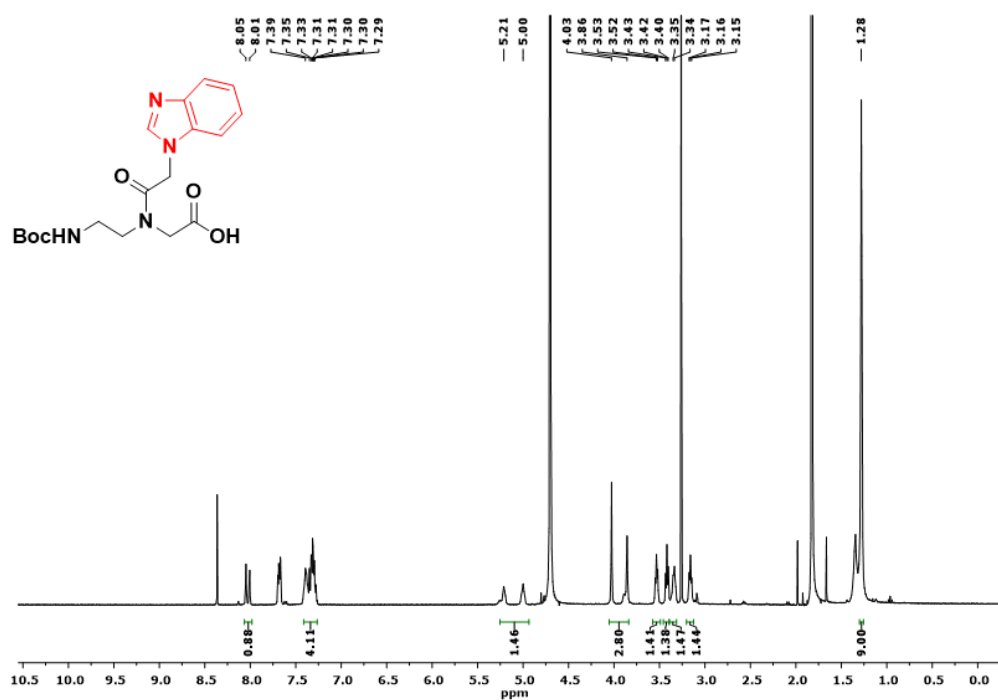
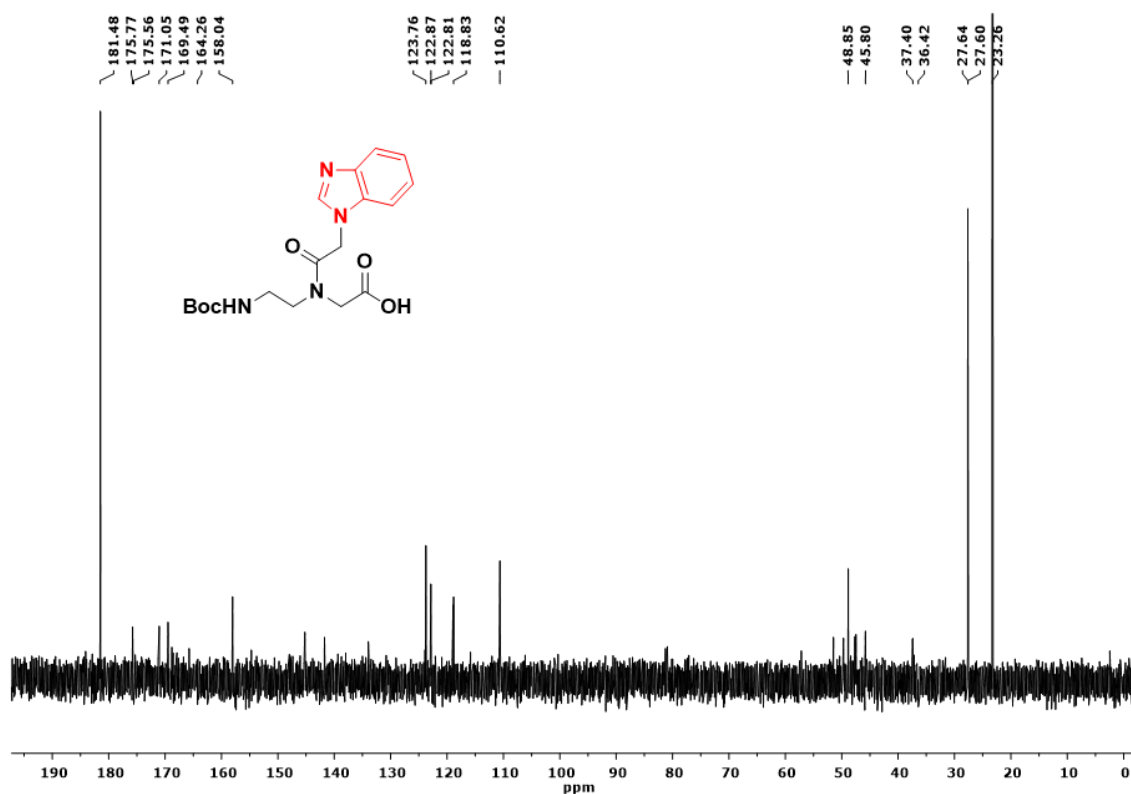
$^1\text{H}$  NMR of Compound 7 $^{13}\text{C}$  NMR of Compound 7

$^1\text{H}$  NMR of Compound **8** $^{13}\text{C}$  NMR of Compound **8**

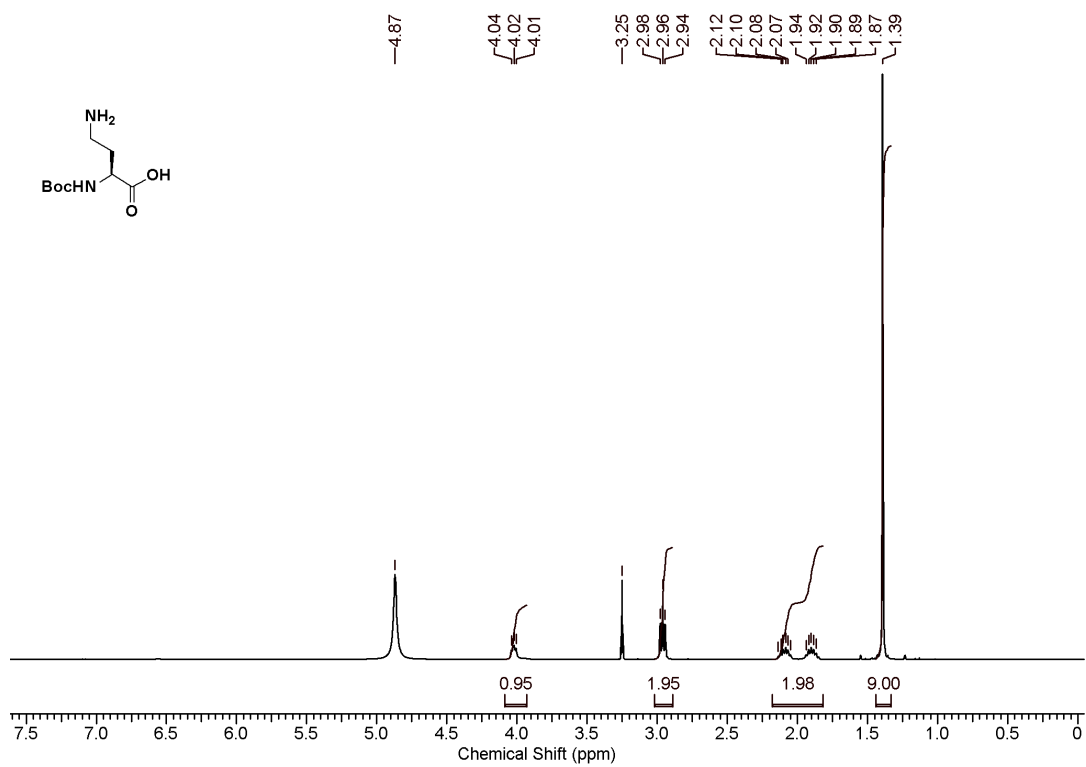
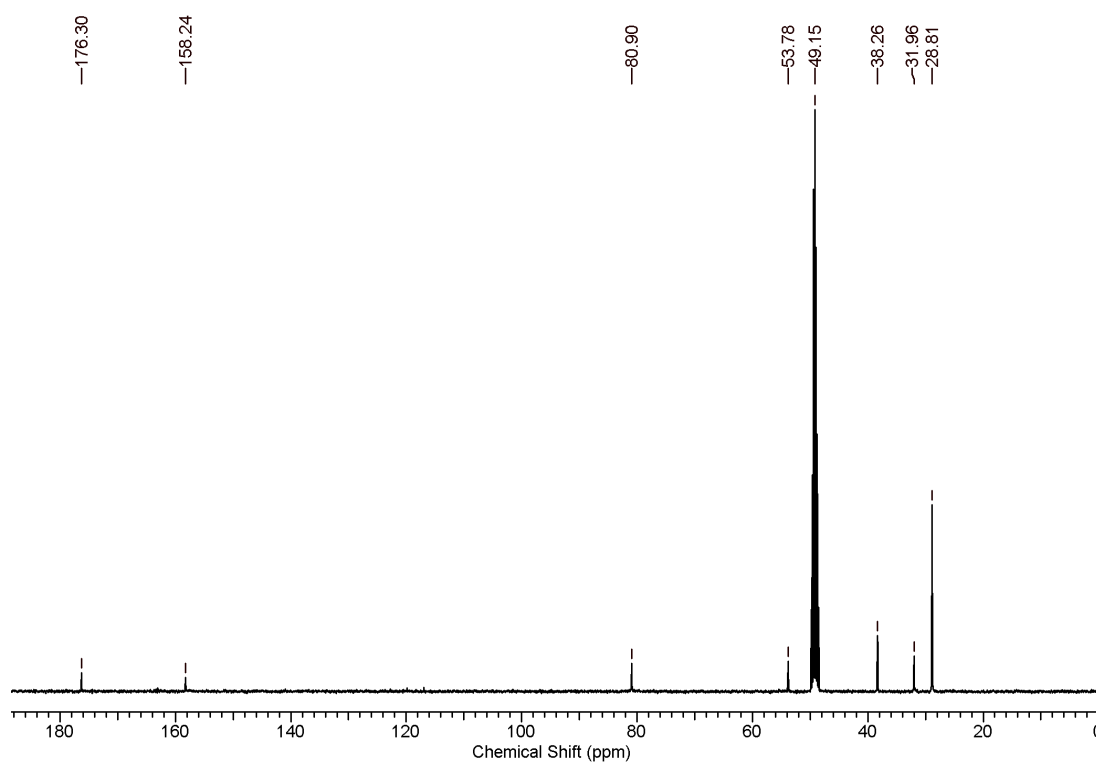
$^1\text{H}$  NMR of Compound **11** $^{13}\text{C}$  NMR of Compound **11**

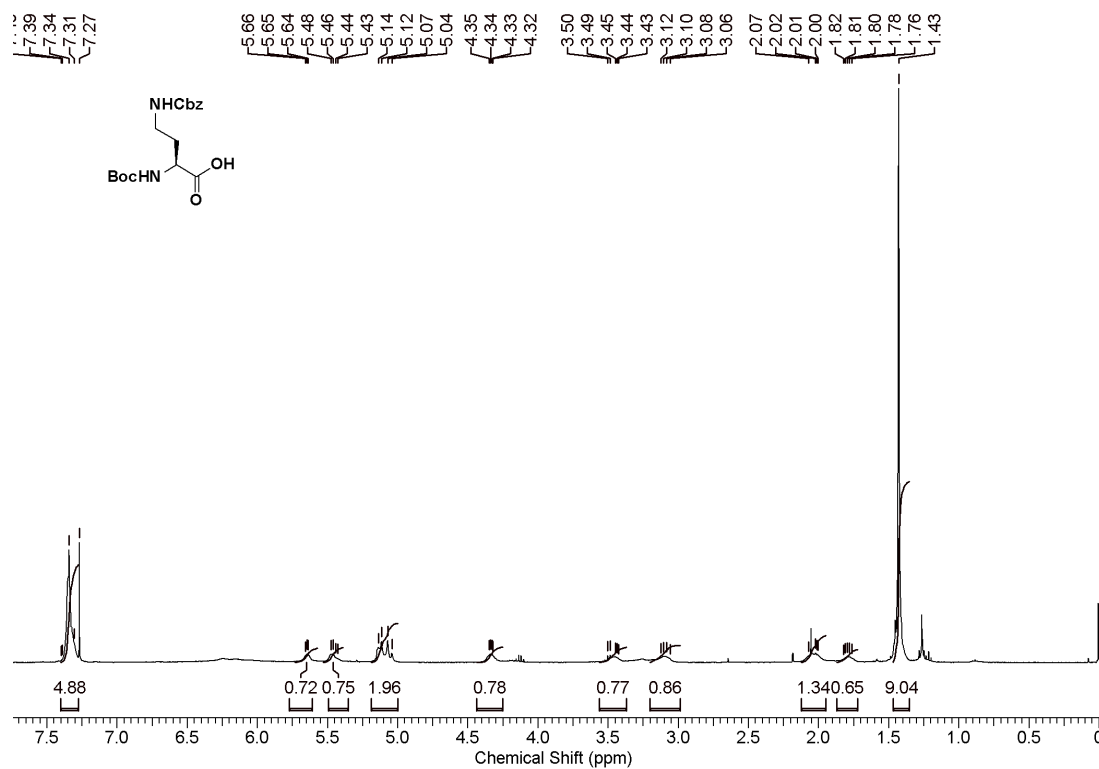
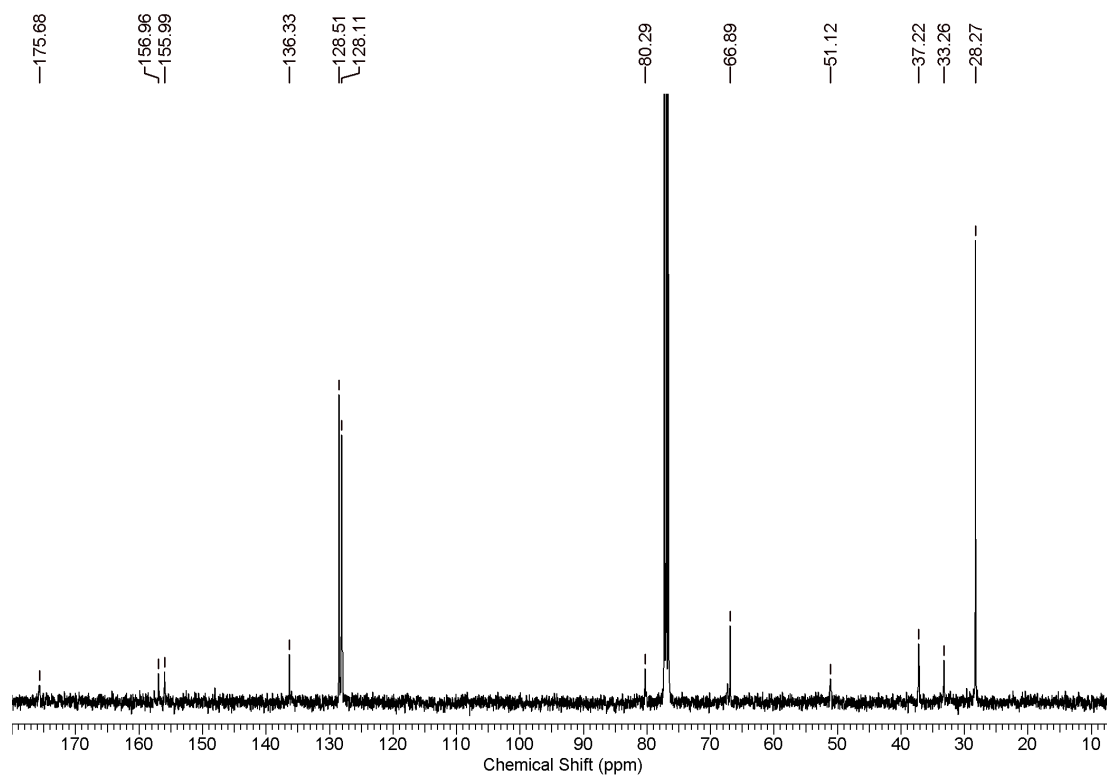
$^1\text{H}$  NMR of Compound **12** $^{13}\text{C}$  NMR of Compound **12**

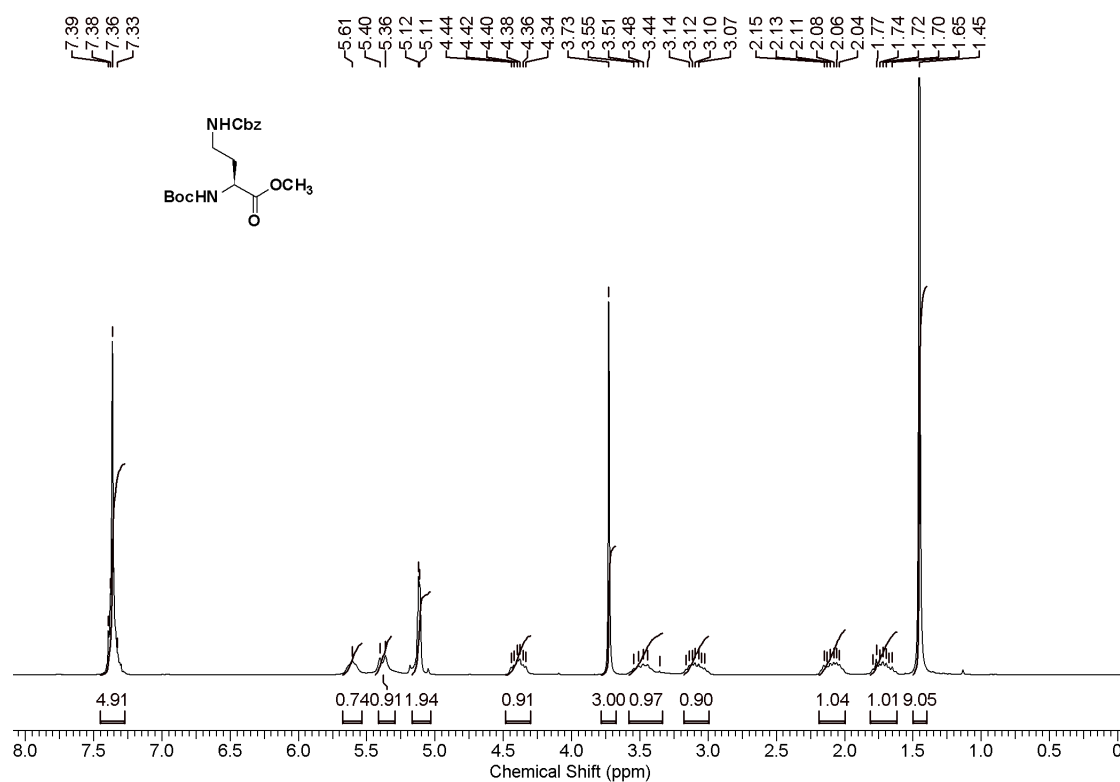
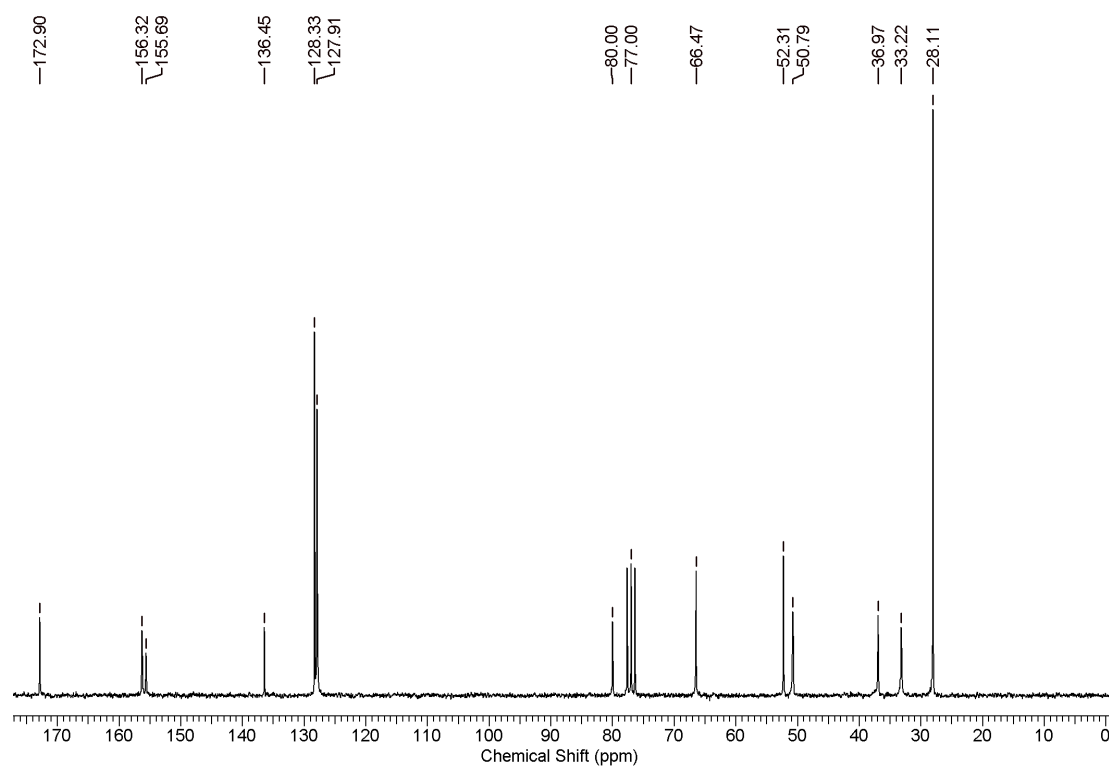
$^1\text{H}$  NMR of Compound **13** $^{13}\text{C}$  NMR spectrum of Compound **13**

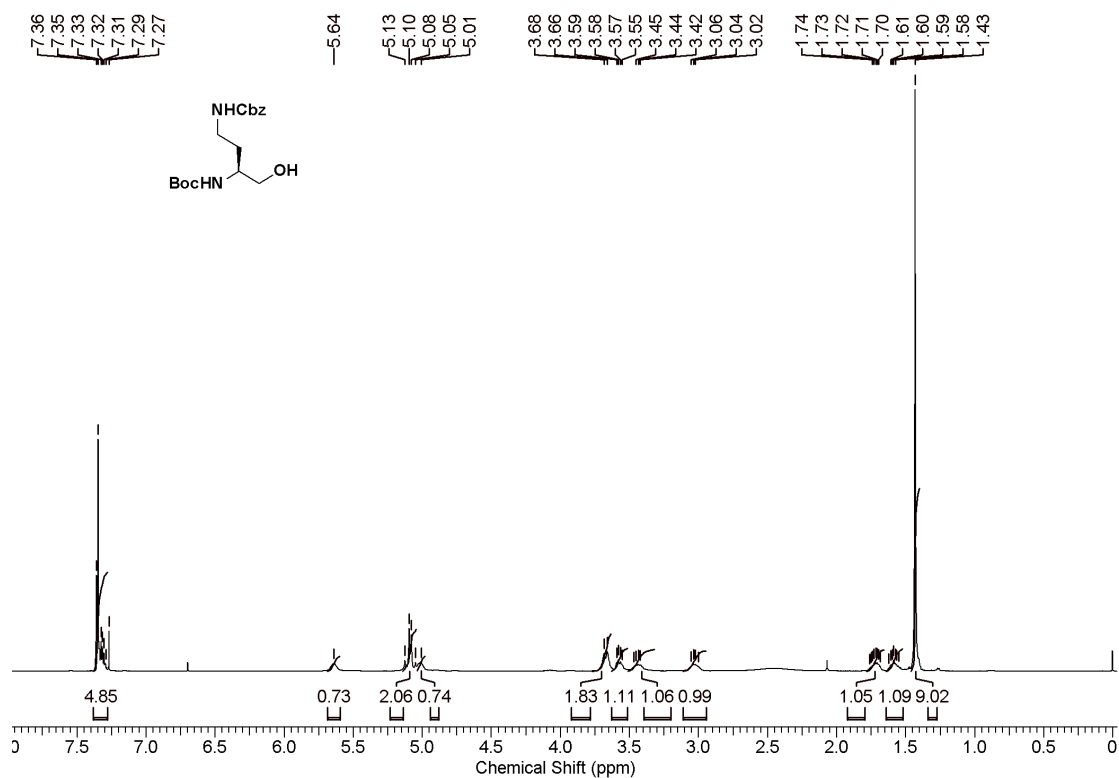
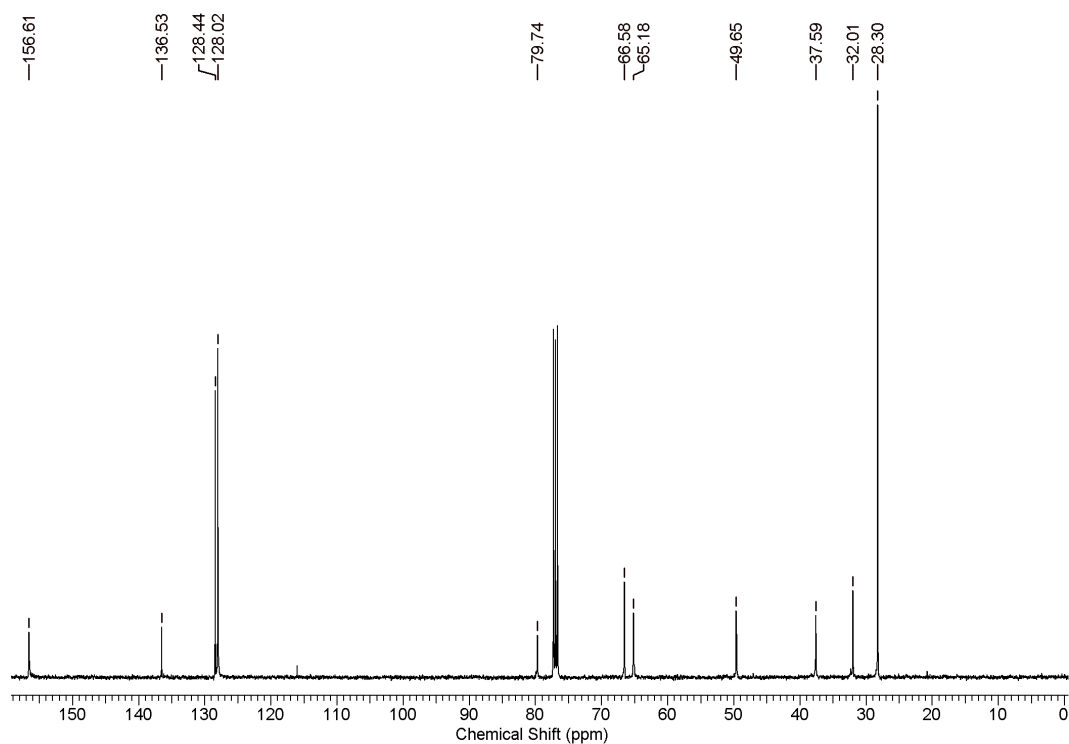
$^1\text{H}$  NMR of Compound **14** $^{13}\text{C}$  NMR of Compound **14**

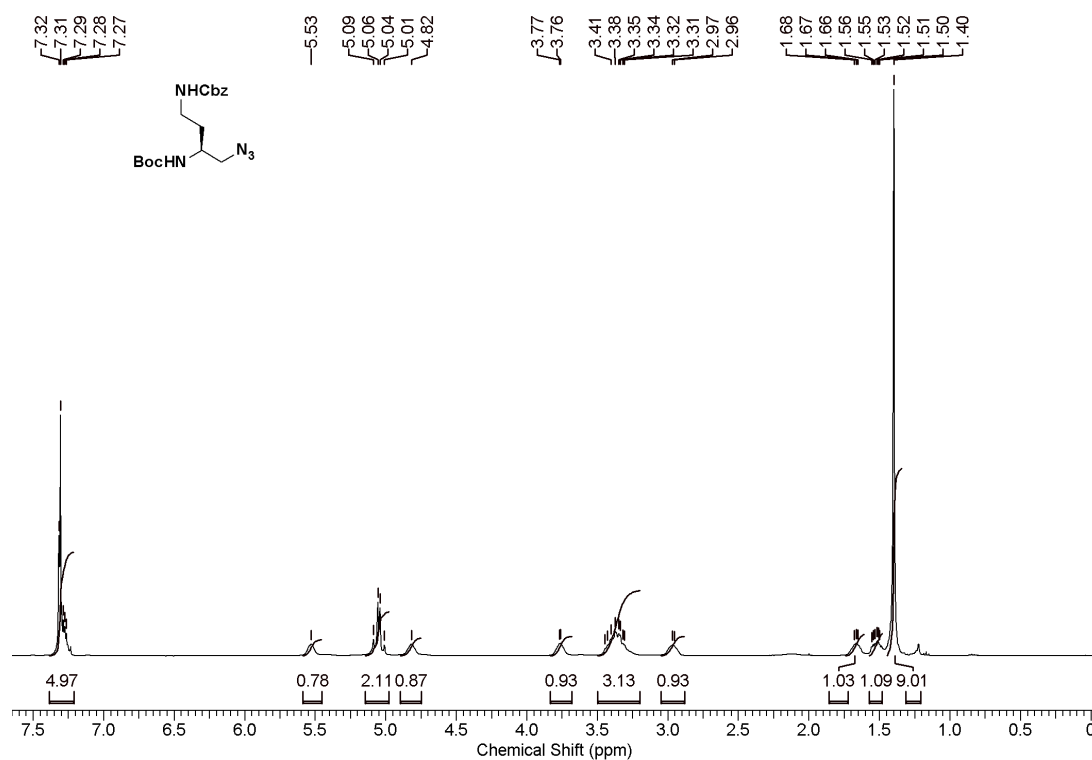
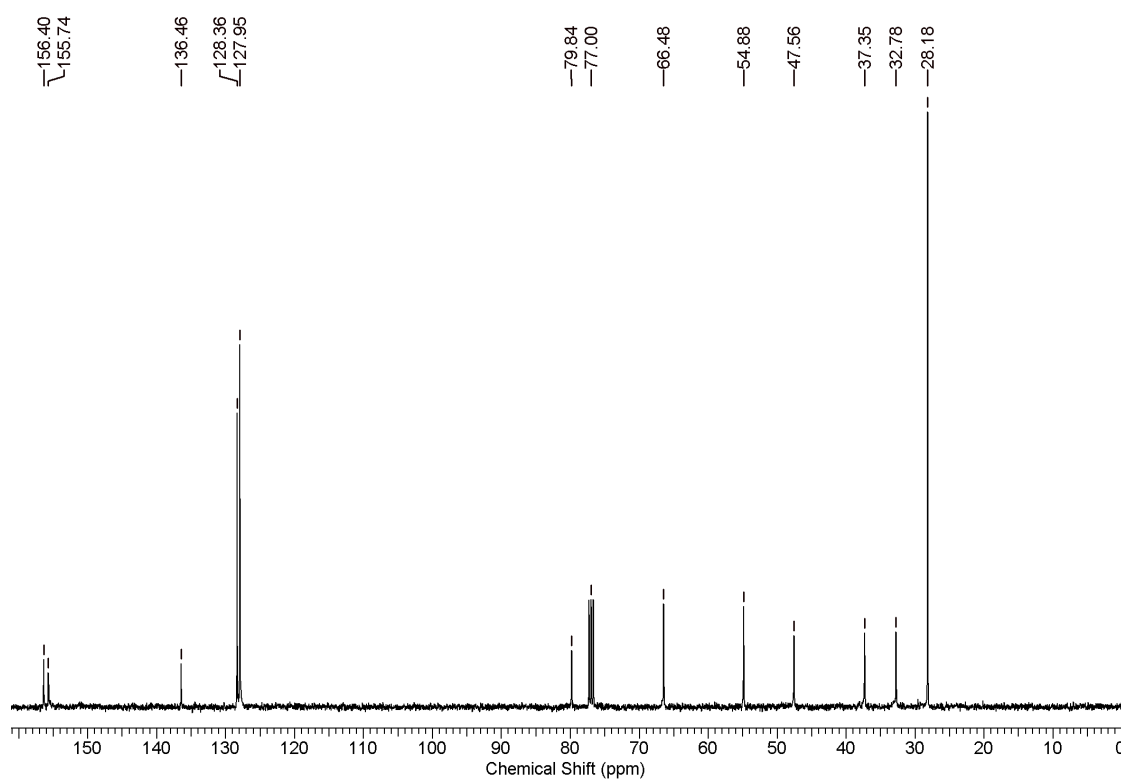


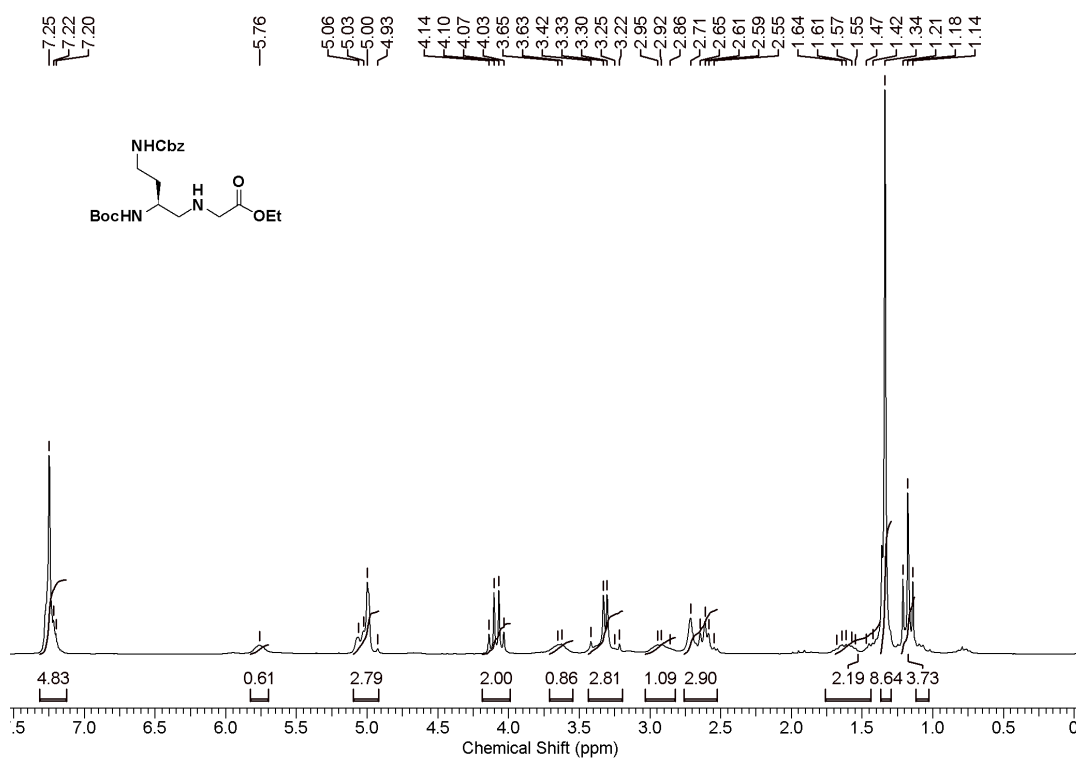
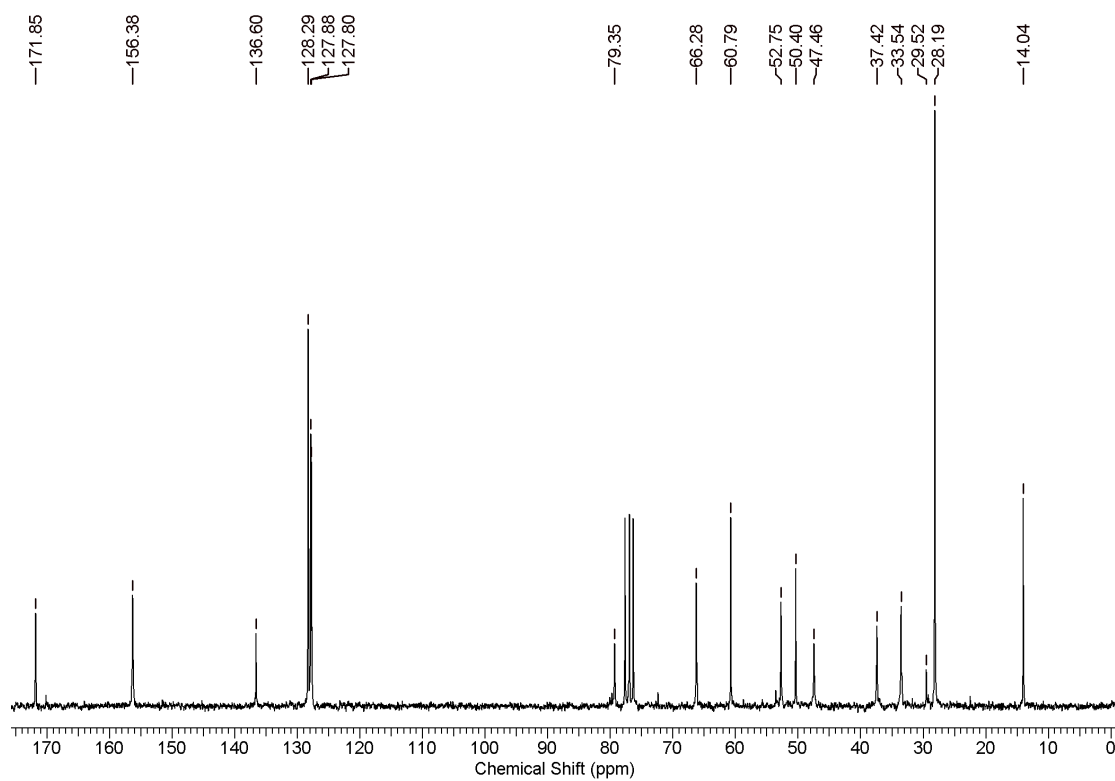
$^1\text{H}$  NMR of Compound **17** $^{13}\text{C}$  NMR of Compound **17**

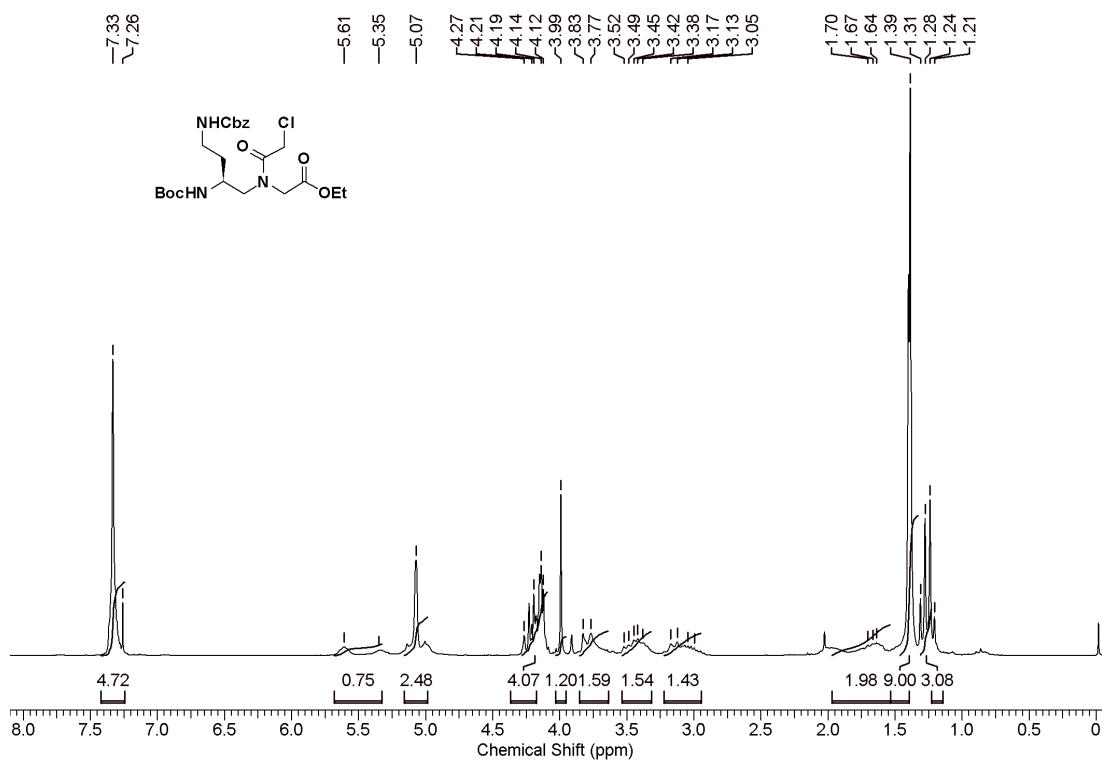
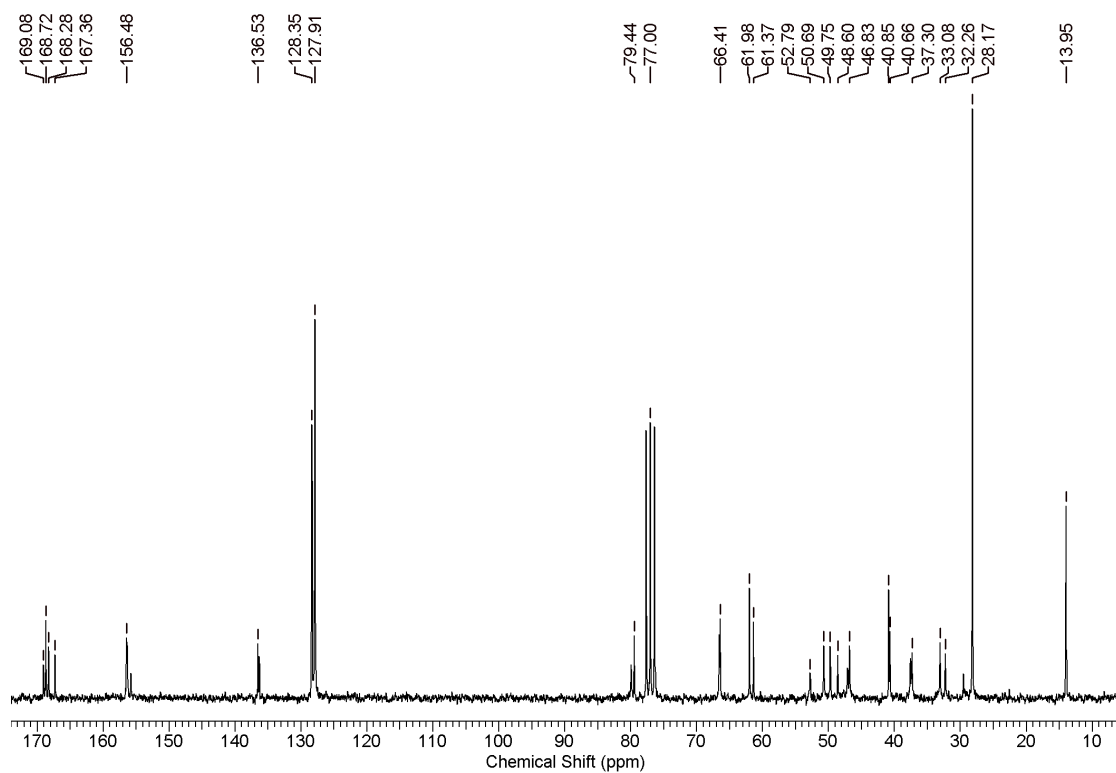
$^1\text{H}$  NMR of Compound **18** $^{13}\text{C}$  NMR spectrum of Compound **18**

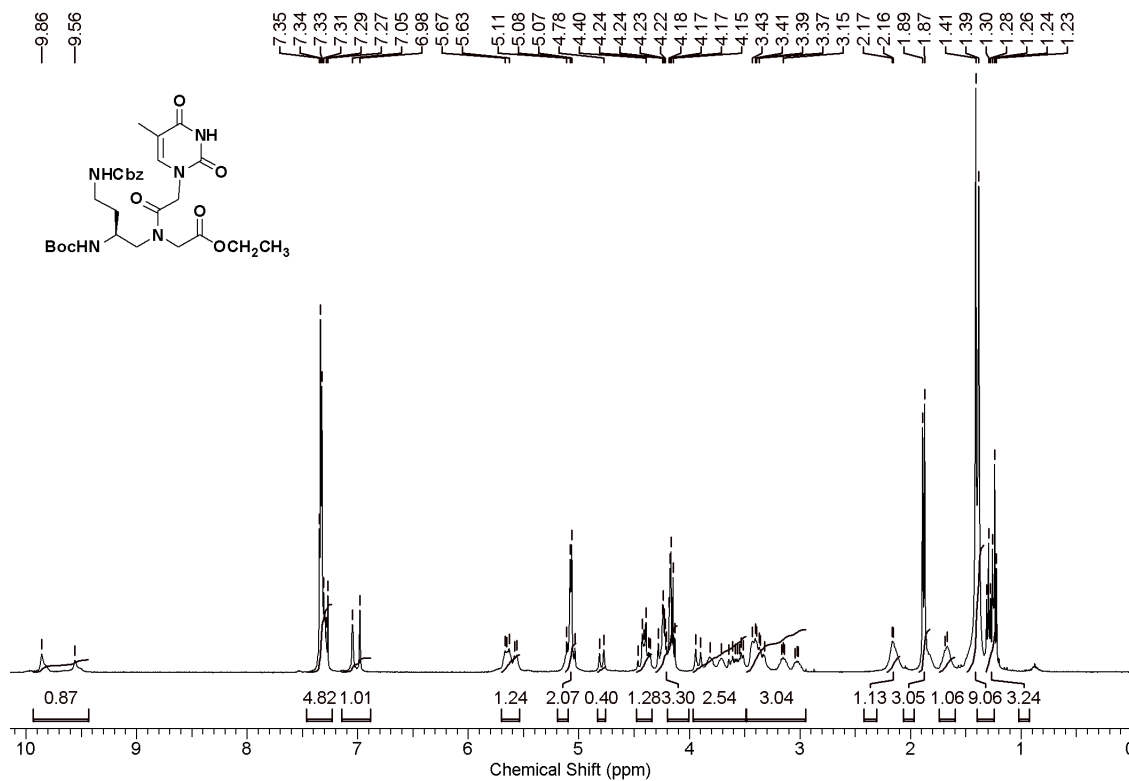
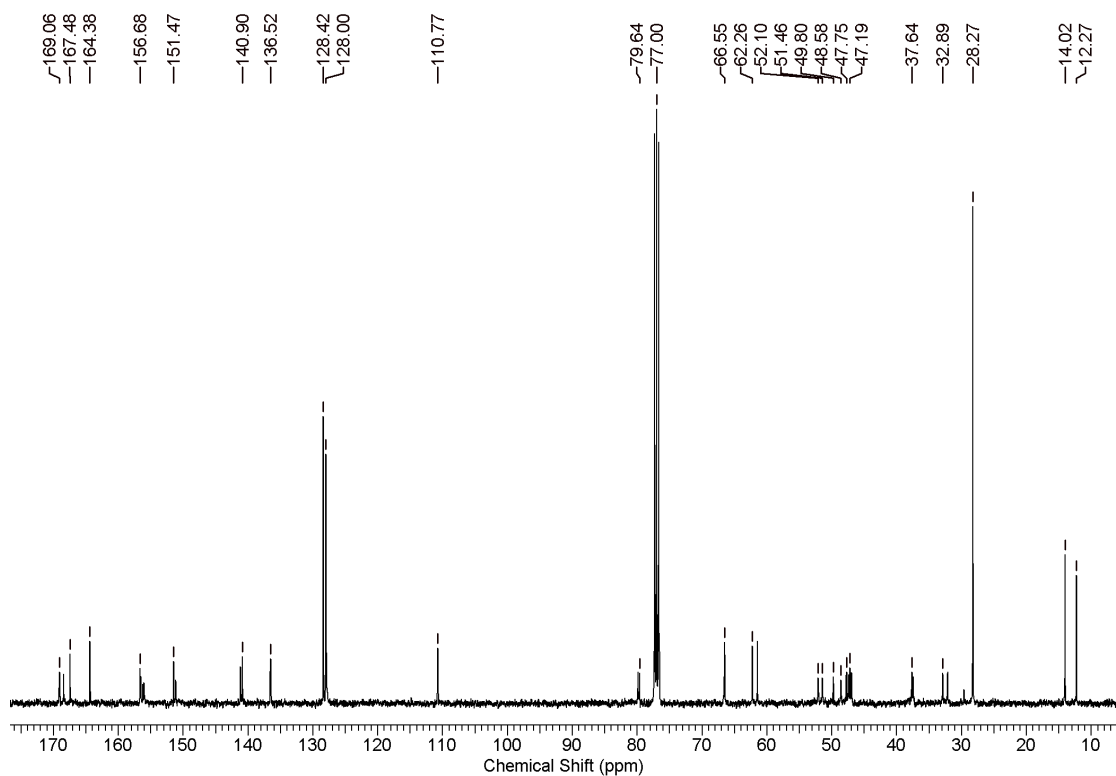
$^1\text{H}$  NMR of Compound **19** $^{13}\text{C}$  NMR of Compound **19**

$^1\text{H}$  NMR of Compound **20** $^{13}\text{C}$  NMR of Compound **20**

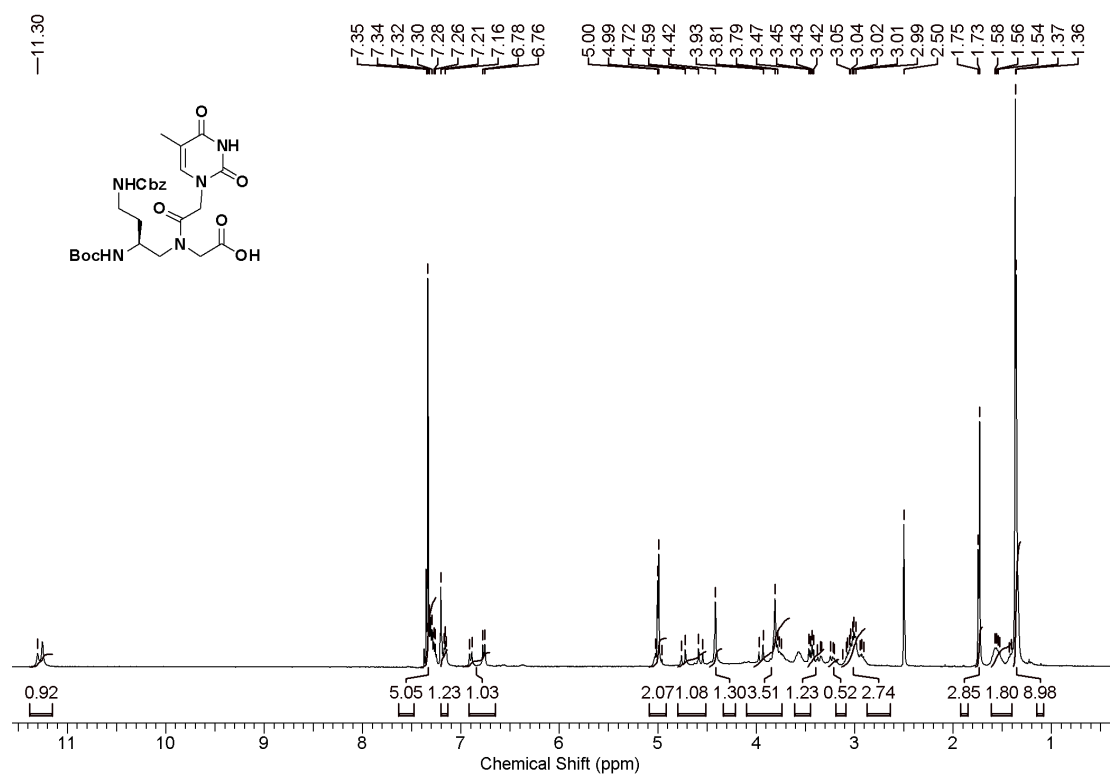
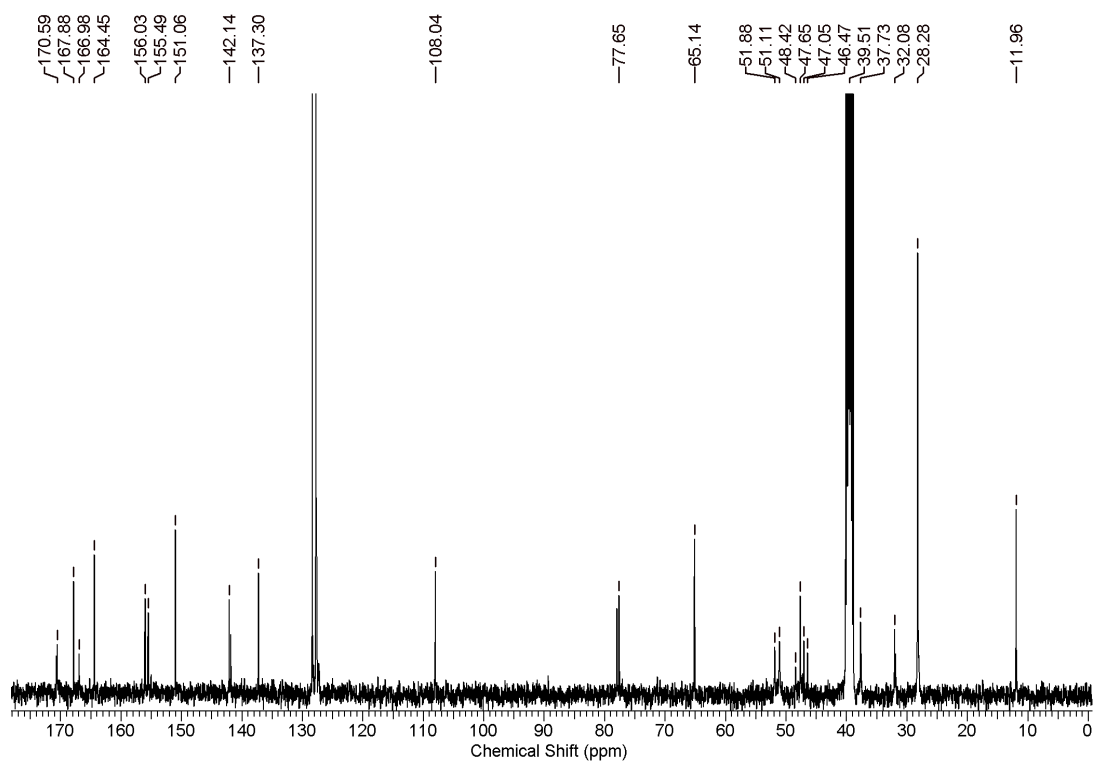
$^1\text{H}$  NMR of Compound **22** $^{13}\text{C}$  NMR of Compound **22**

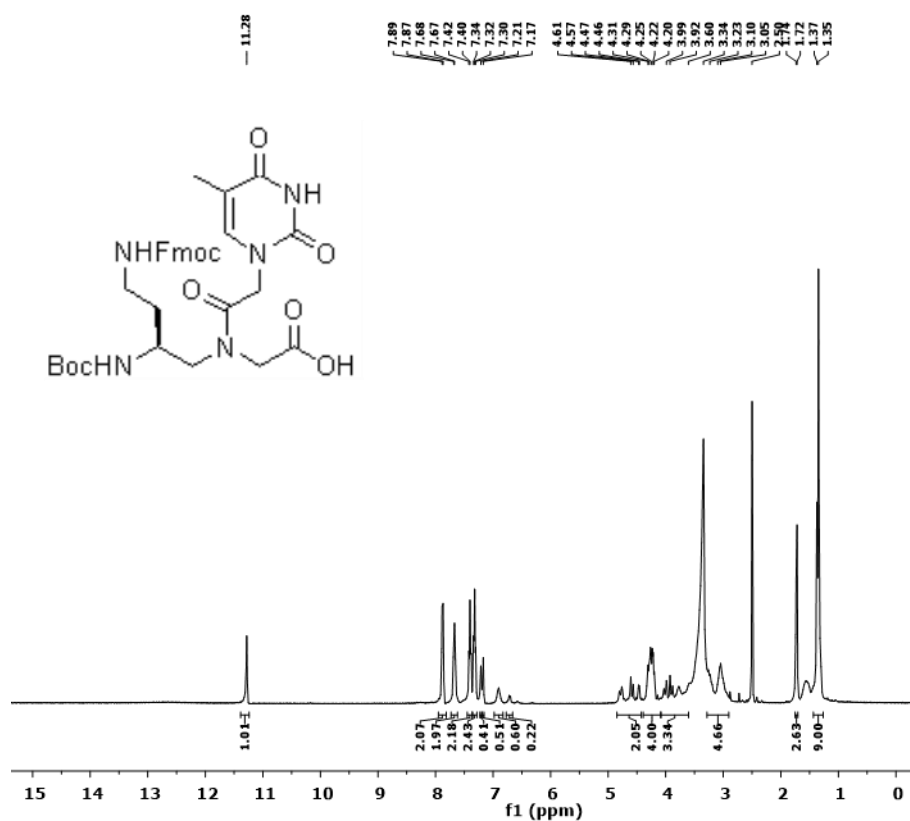
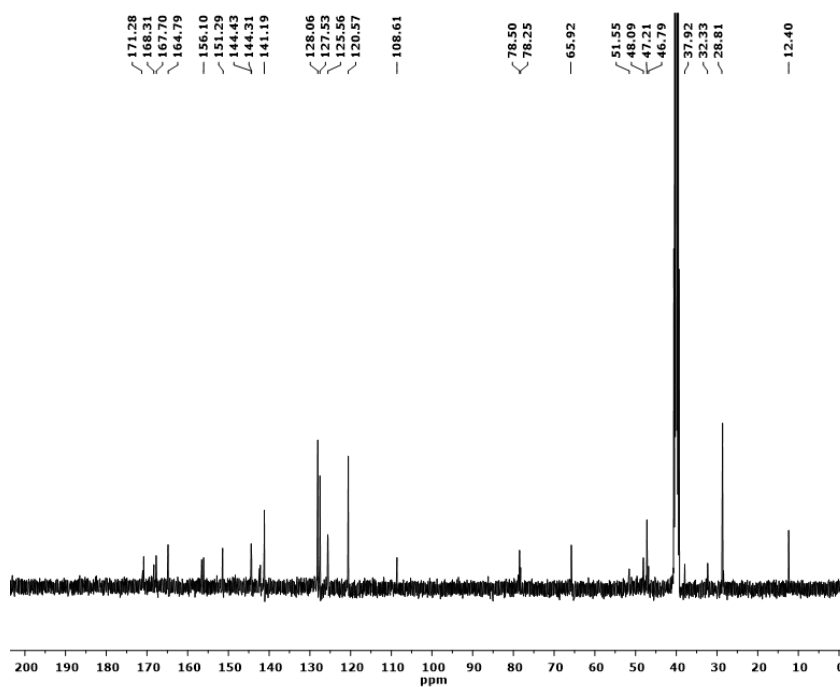
$^1\text{H}$  NMR of Compound **23** $^{13}\text{C}$  NMR of Compound **23**

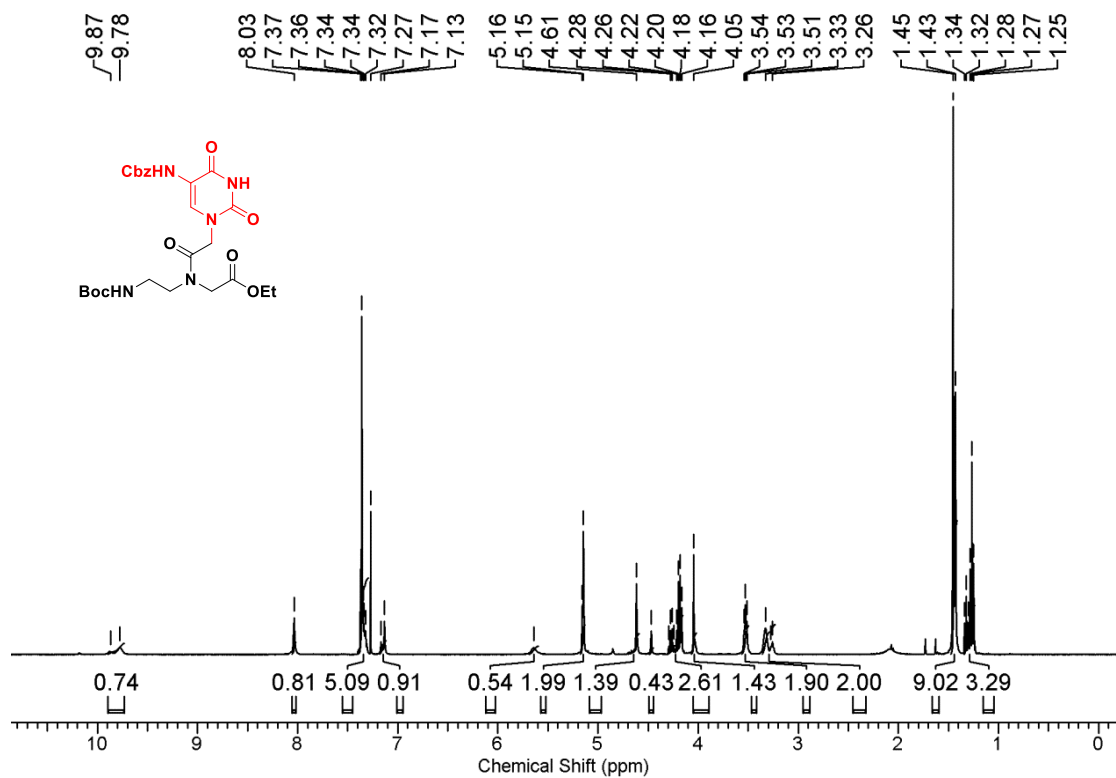
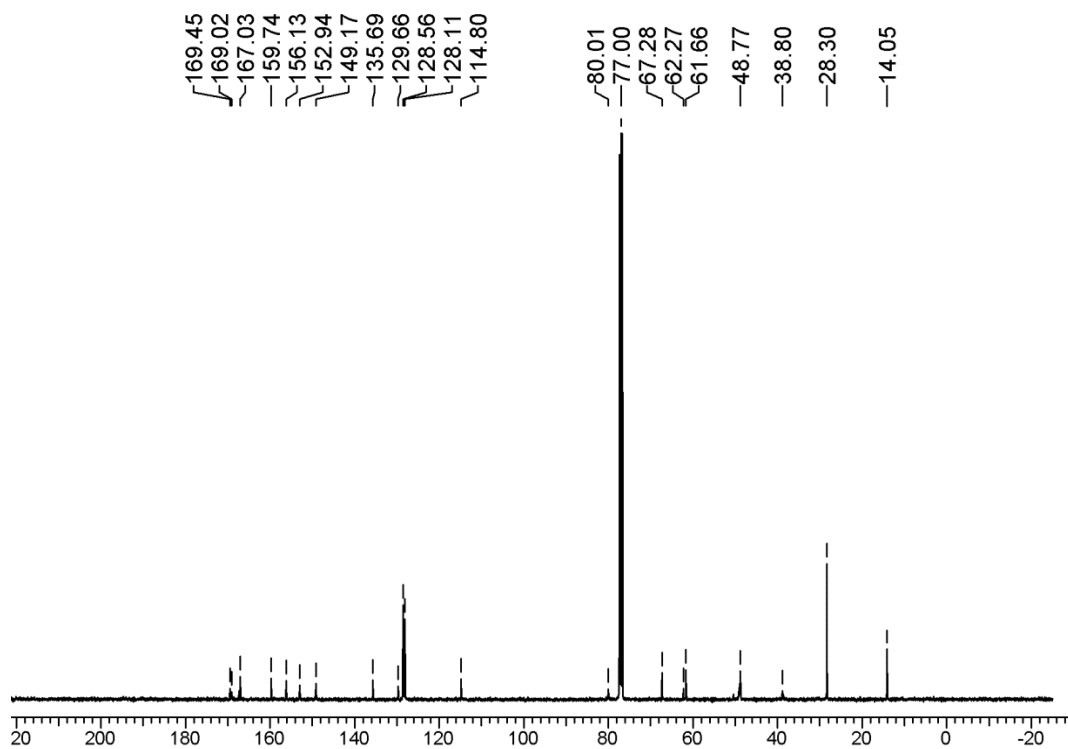
$^1\text{H}$  NMR of Compound **24** $^{13}\text{C}$  NMR of Compound **24**

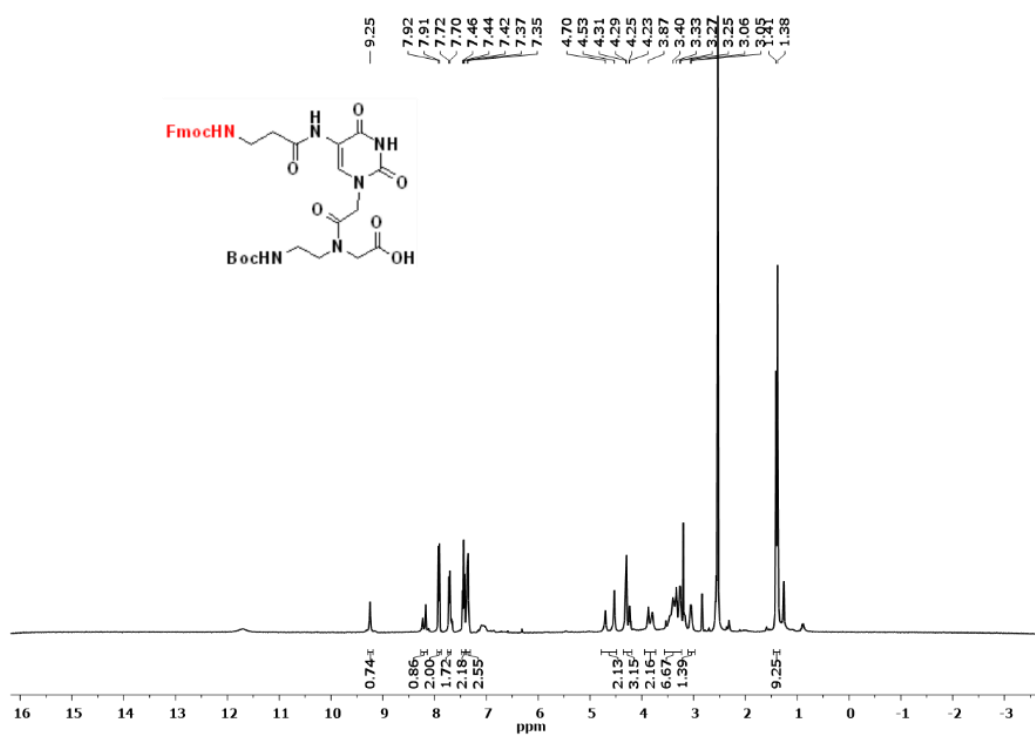
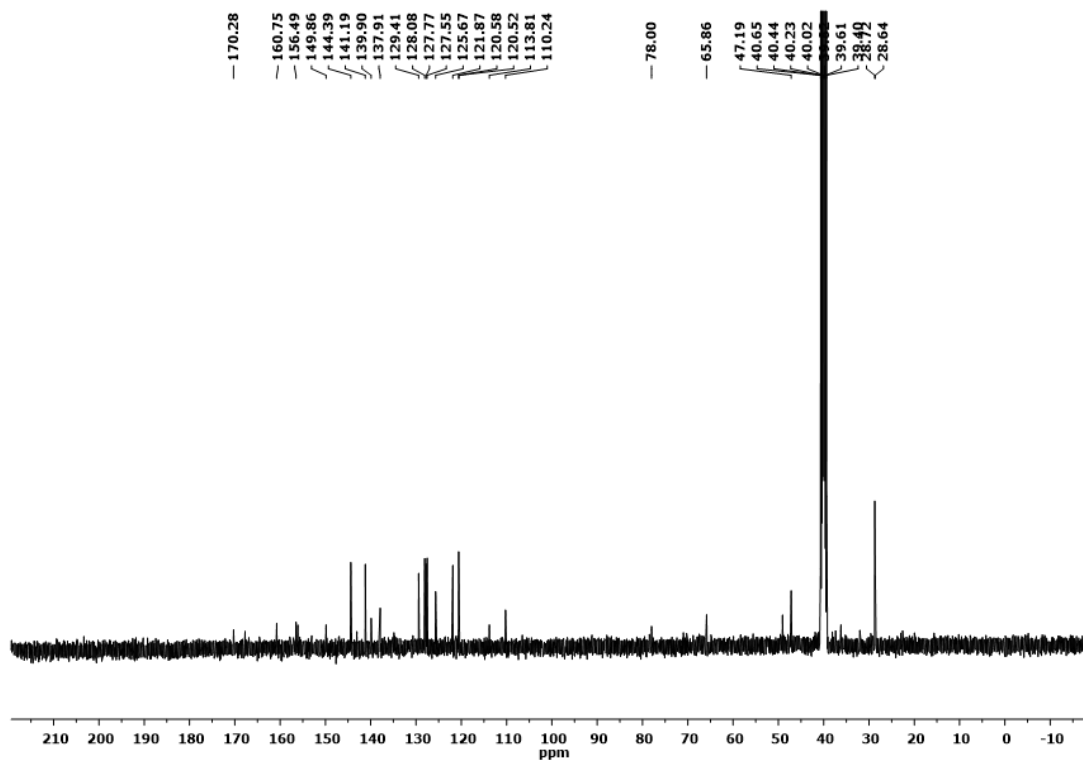
$^1\text{H}$  NMR of Compound **25** $^{13}\text{C}$  NMR of Compound **25**

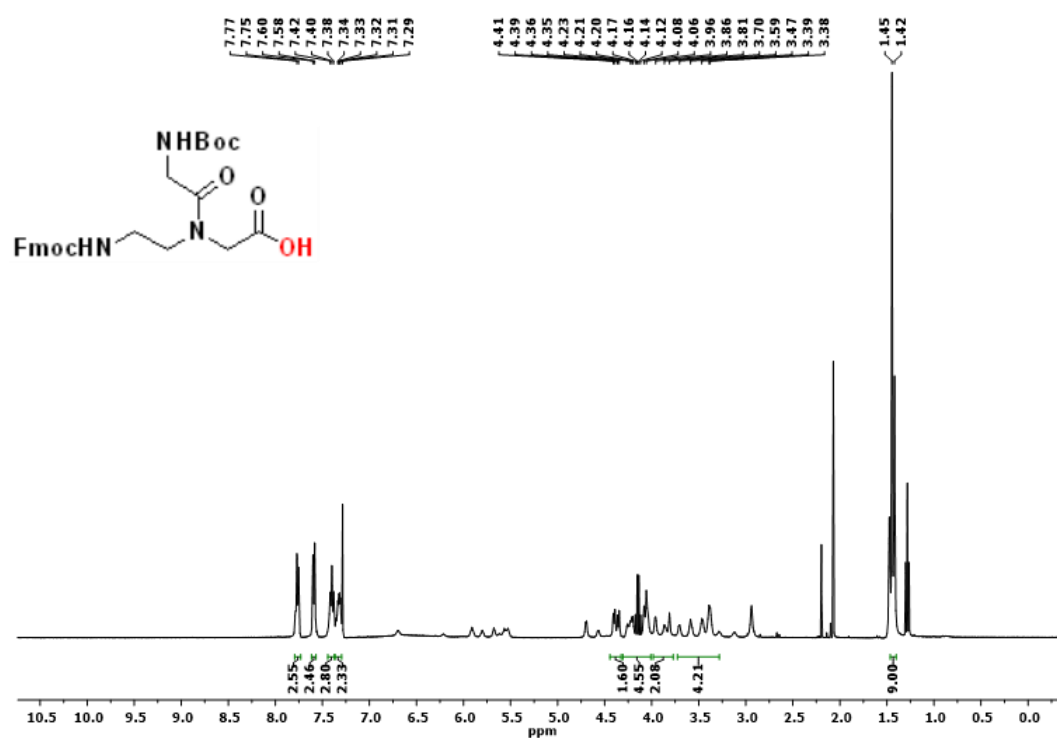
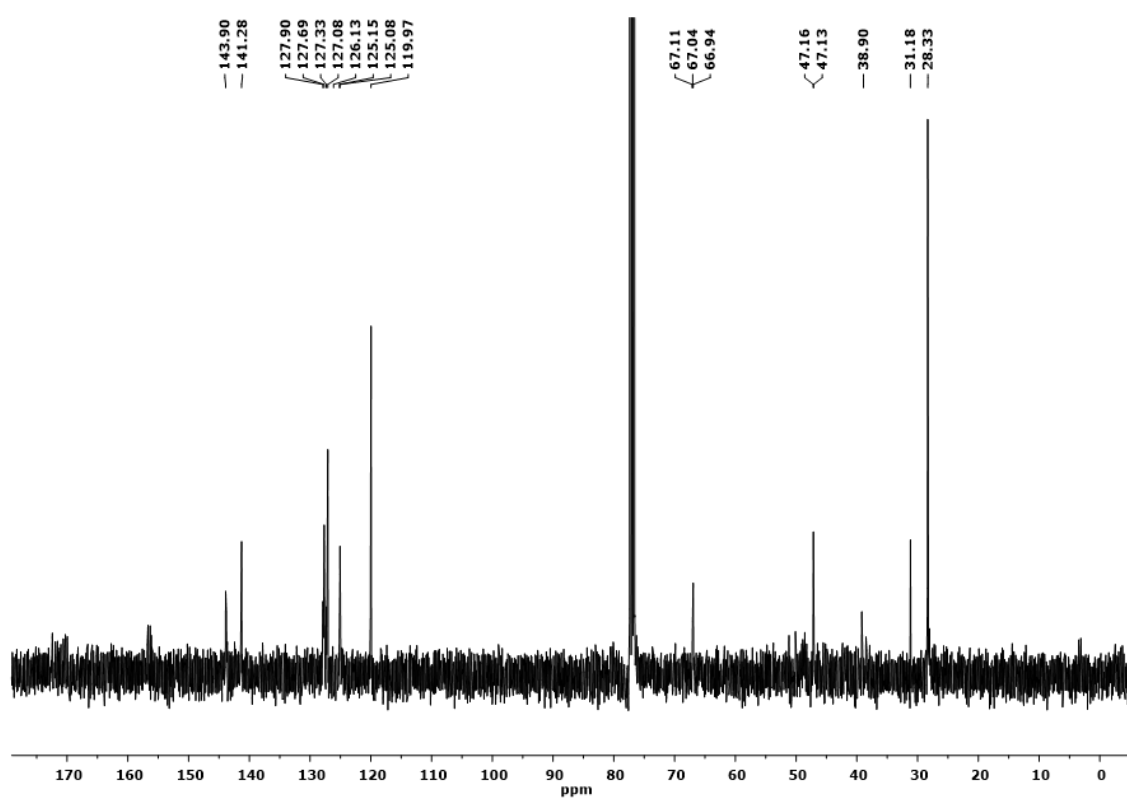


<sup>1</sup>H NMR of Compound 26<sup>13</sup>C NMR of Compound 26

$^1\text{H}$  NMR of Compound **27** $^{13}\text{C}$  NMR of Compound **27**

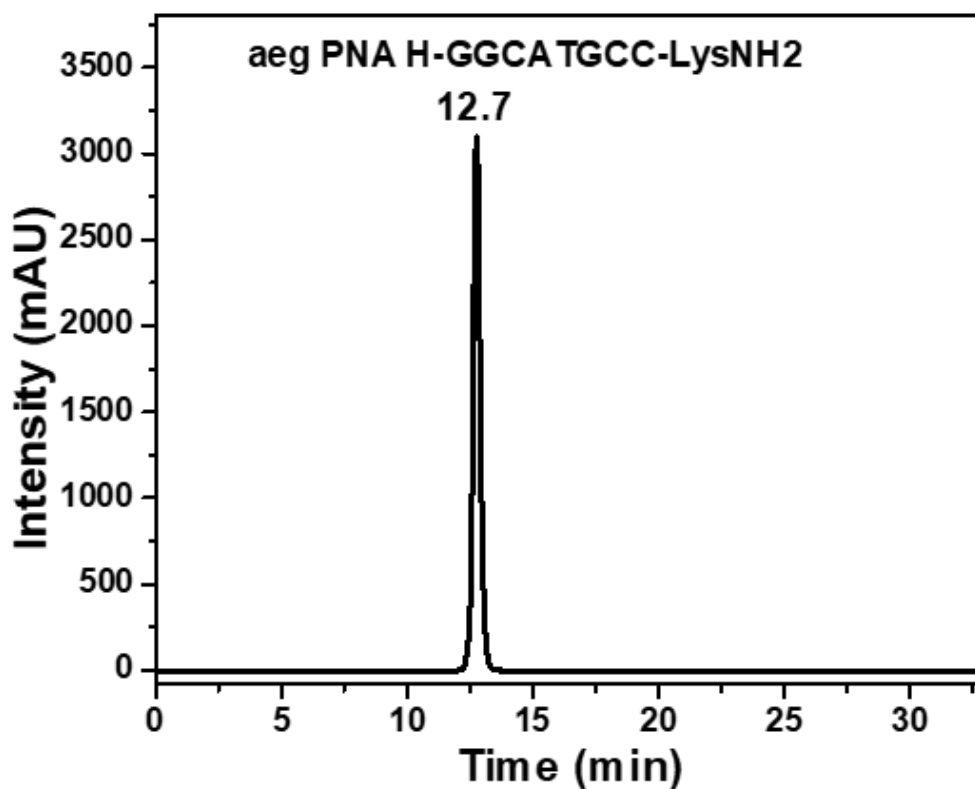
$^1\text{H}$  NMR of Compound **28** $^{13}\text{C}$  NMR of Compound **28**

$^1\text{H}$  NMR of Compound **32** $^{13}\text{C}$  NMR of Compound **32**

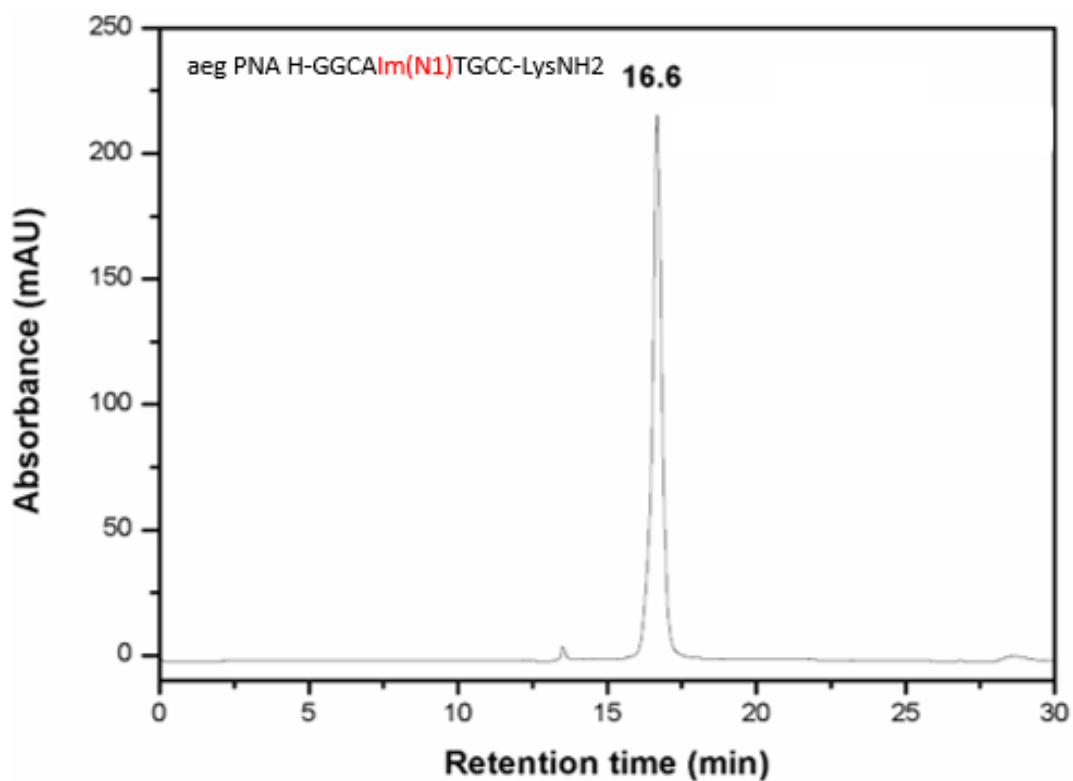
$^1\text{H}$  NMR of Compound **36** $^{13}\text{C}$  NMR of Compound **36**

## HPLC chromatograms of peptides

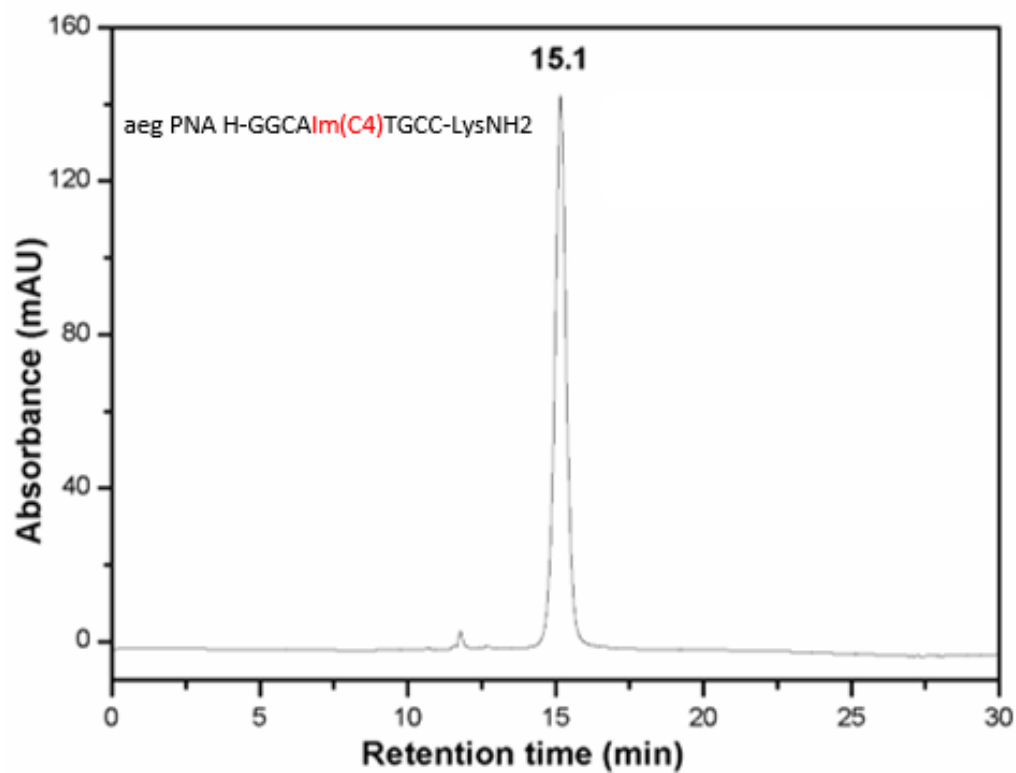
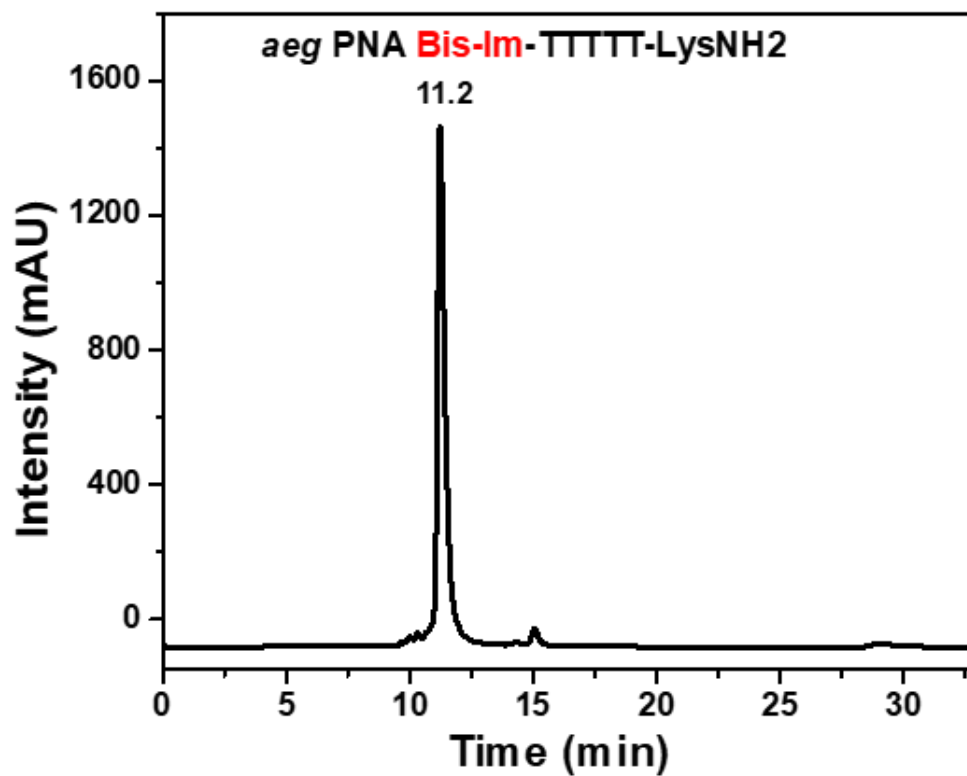
HPLC chromatogram of aeg PNA 1

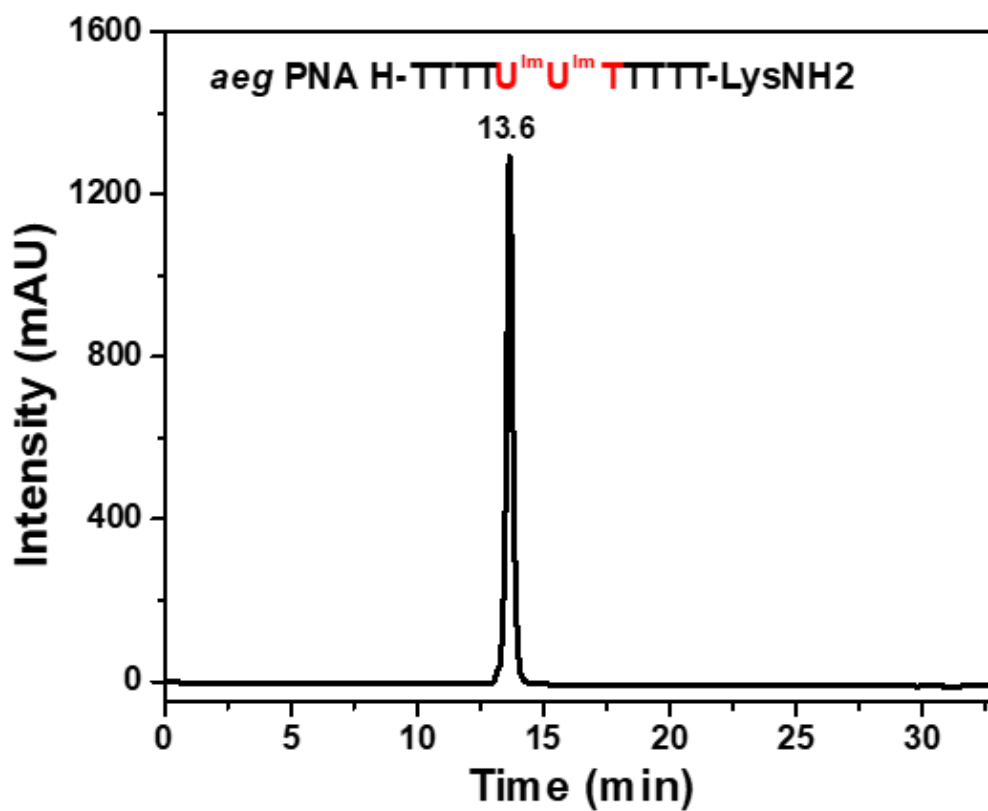
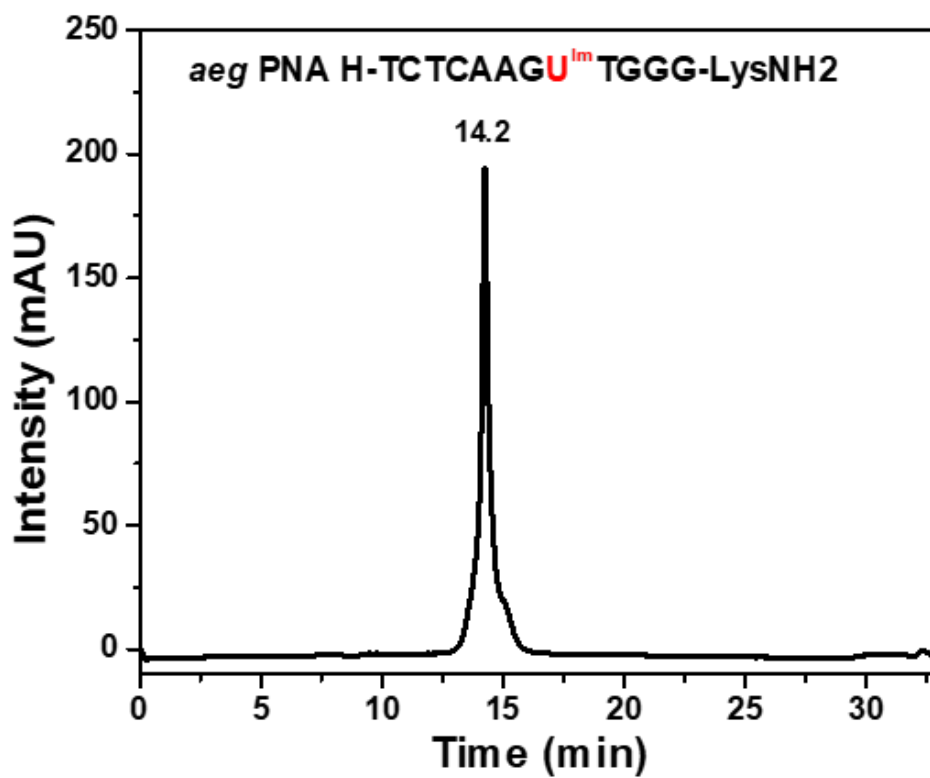


HPLC chromatogram of aeg PNA 2

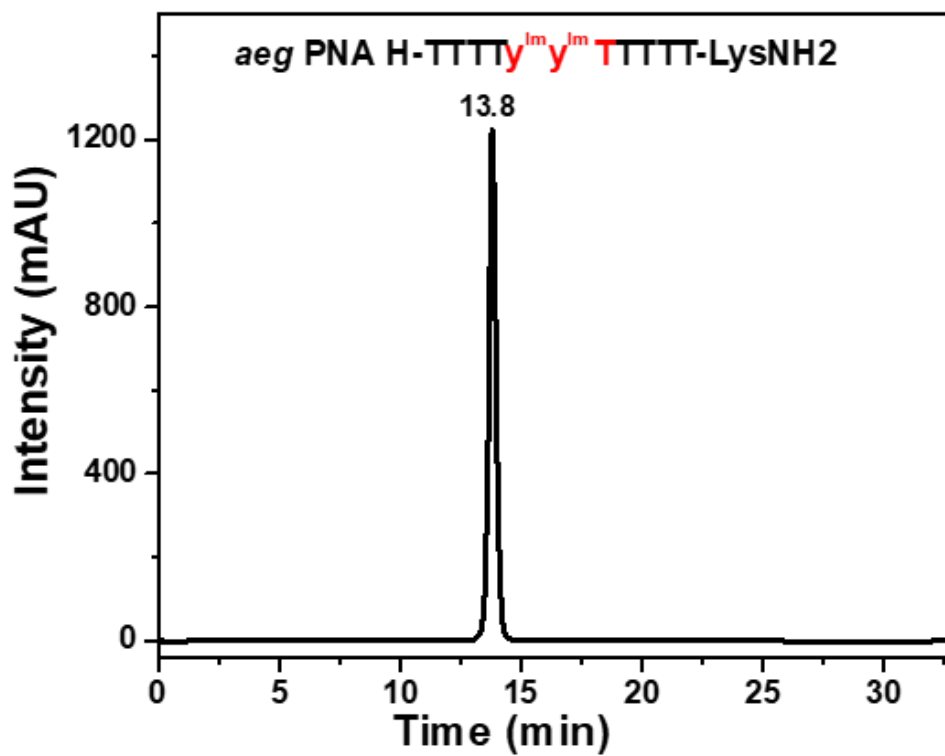
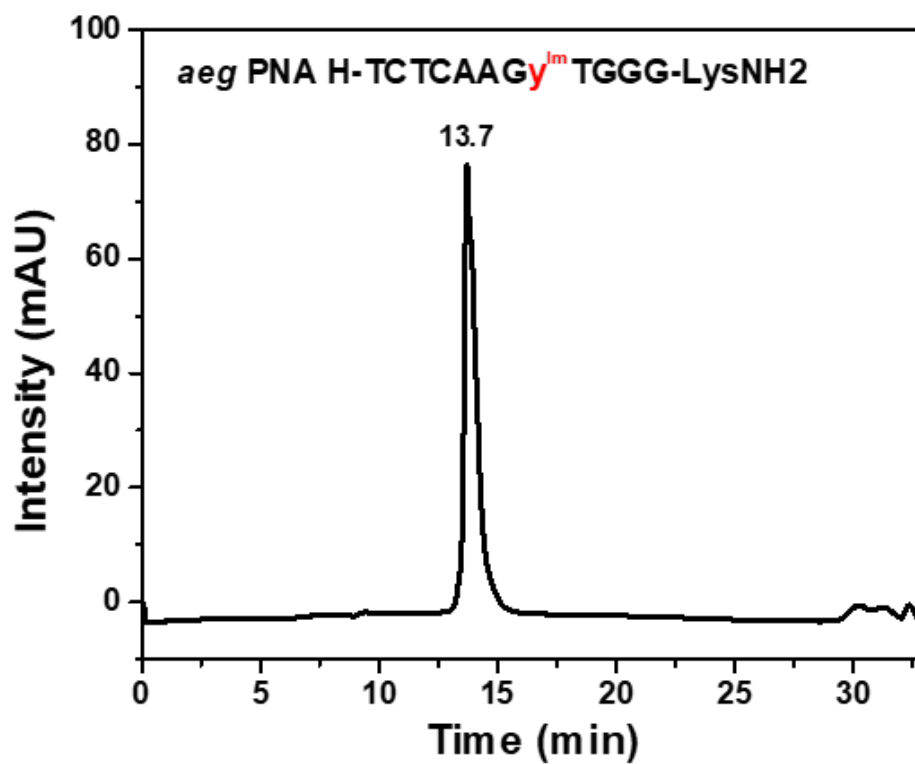


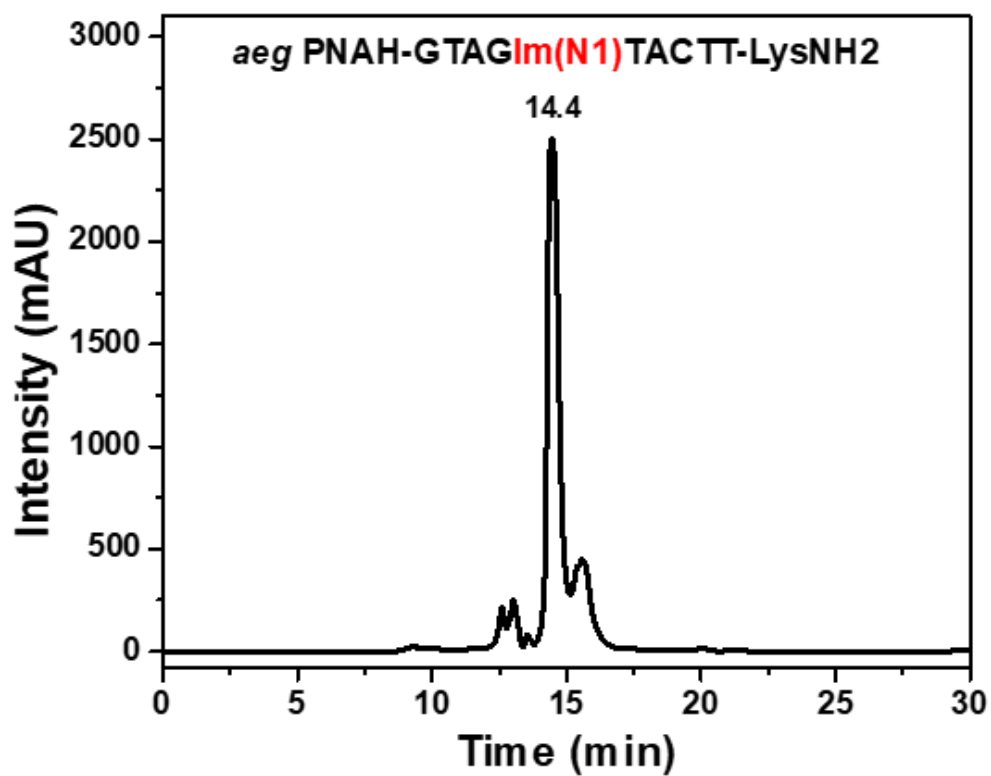
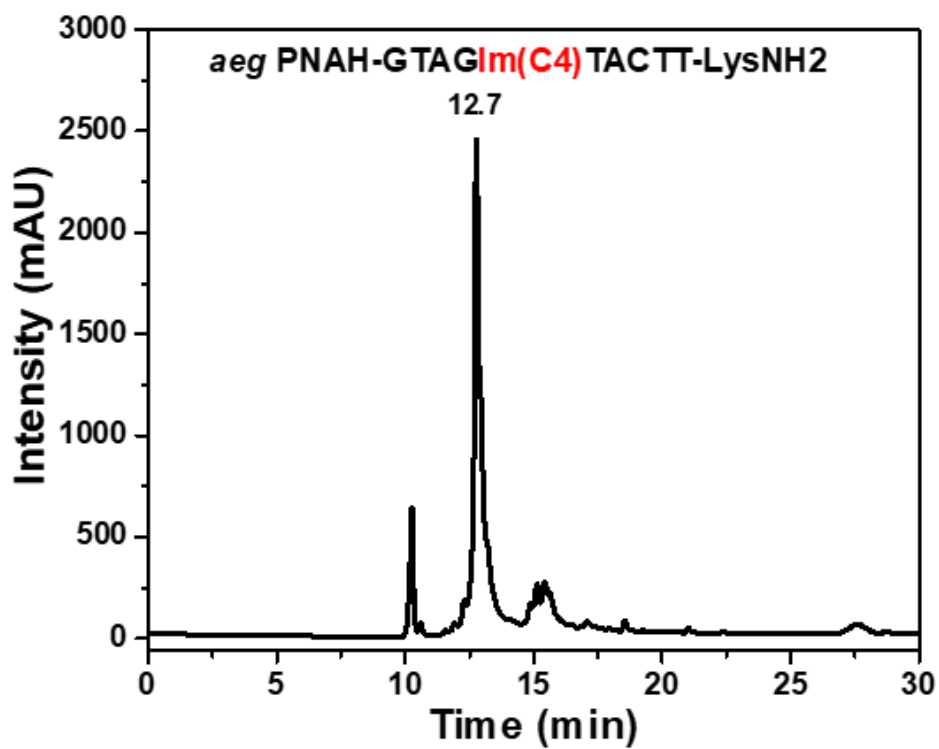
HPLC chromatogram of aeg PNA 3

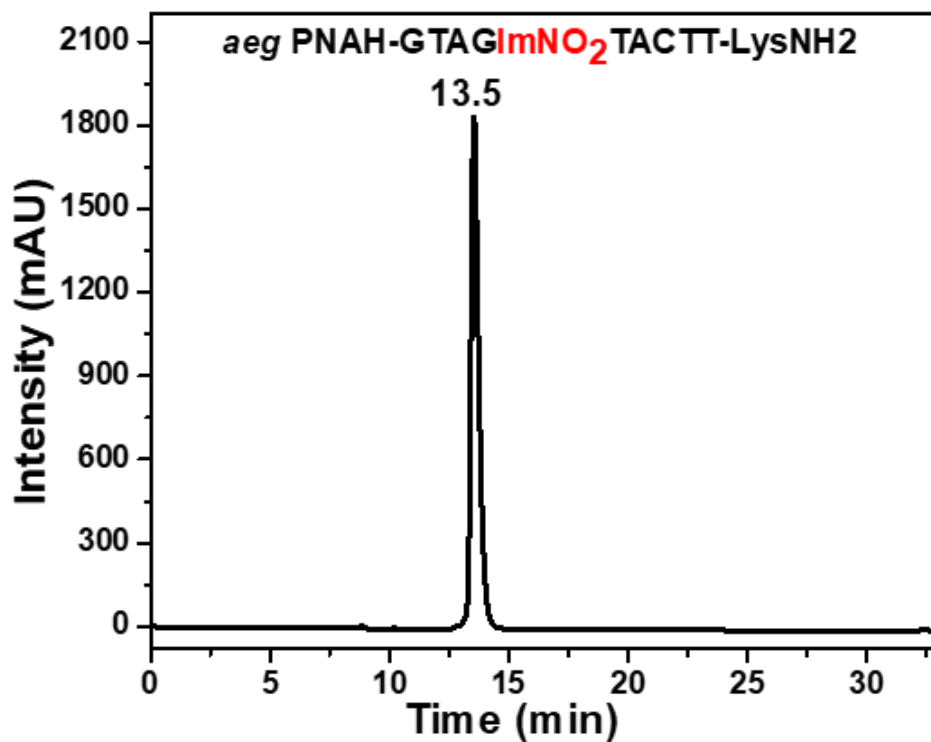
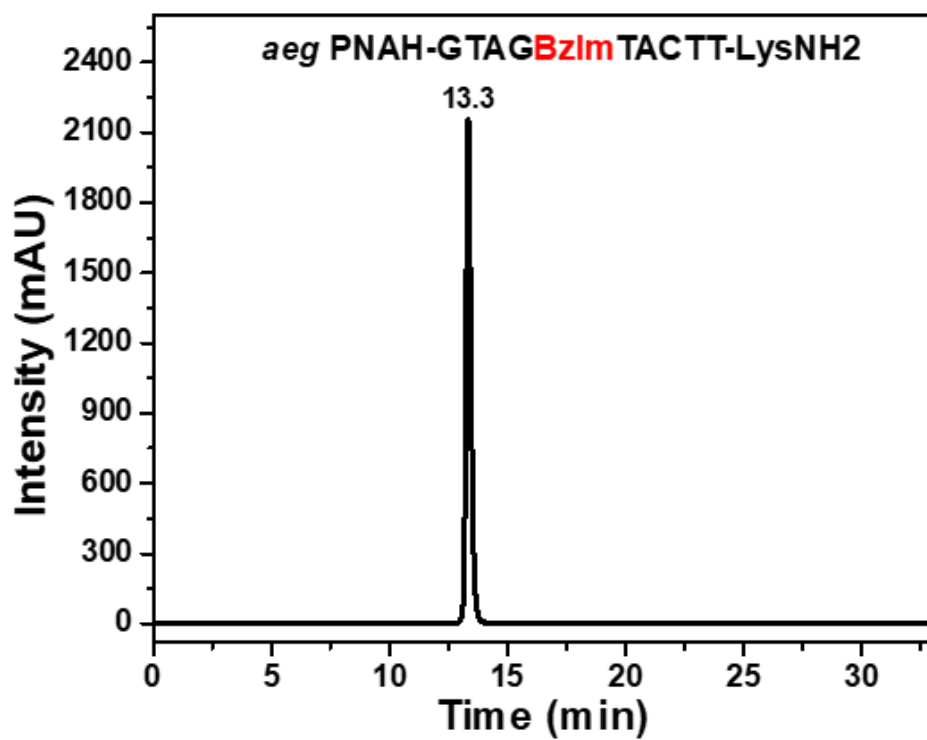
HPLC chromatogram of *Bis-Im* PNA 4

HPLC chromatogram of  $U^{Im}T_{10}$ -PNA 5HPLC chromatogram of  $U^{Im}$ -PNA 6

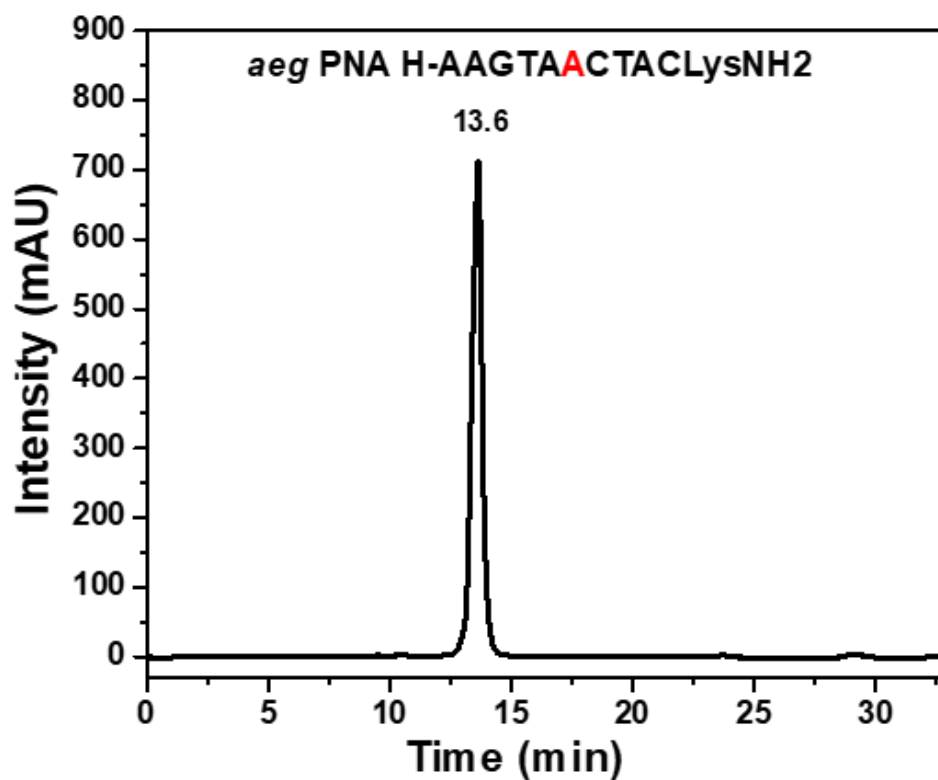


HPLC chromatogram of  $\gamma^{Imr}T_{10}$ -PNA 7HPLC chromatogram of  $\gamma^{Im}$ -PNA 8

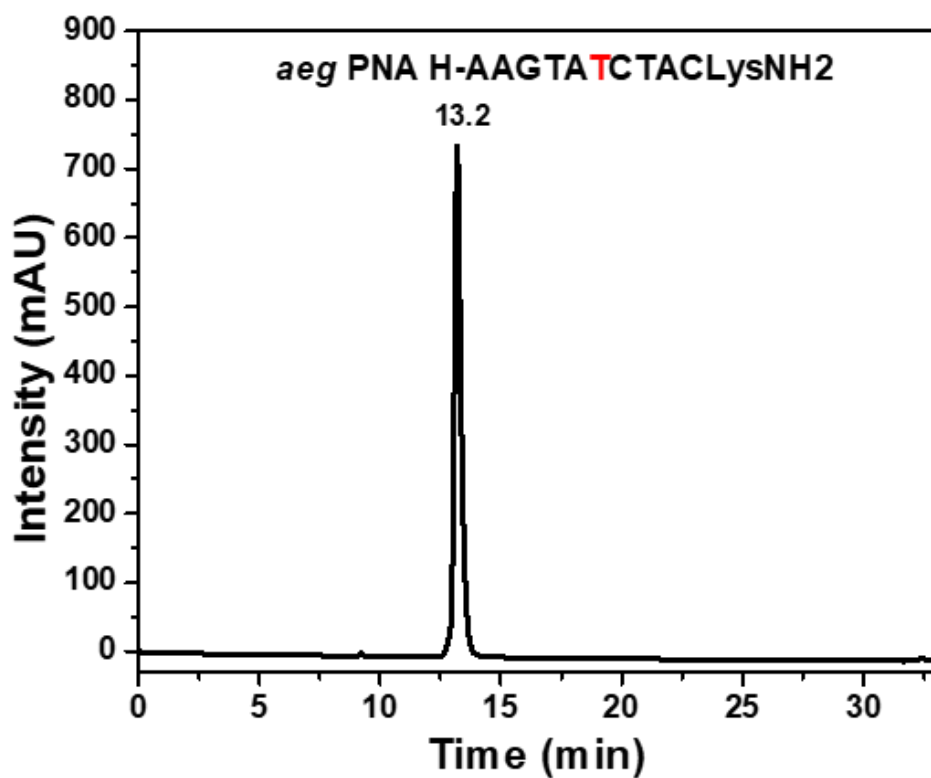
HPLC chromatogram of *Im(N1)*-PNA 9HPLC chromatogram of *Im(C4)*-PNA 10

HPLC chromatogram of *Im NO2* PNA 11HPLC chromatogram of *Bzim* PNA 12

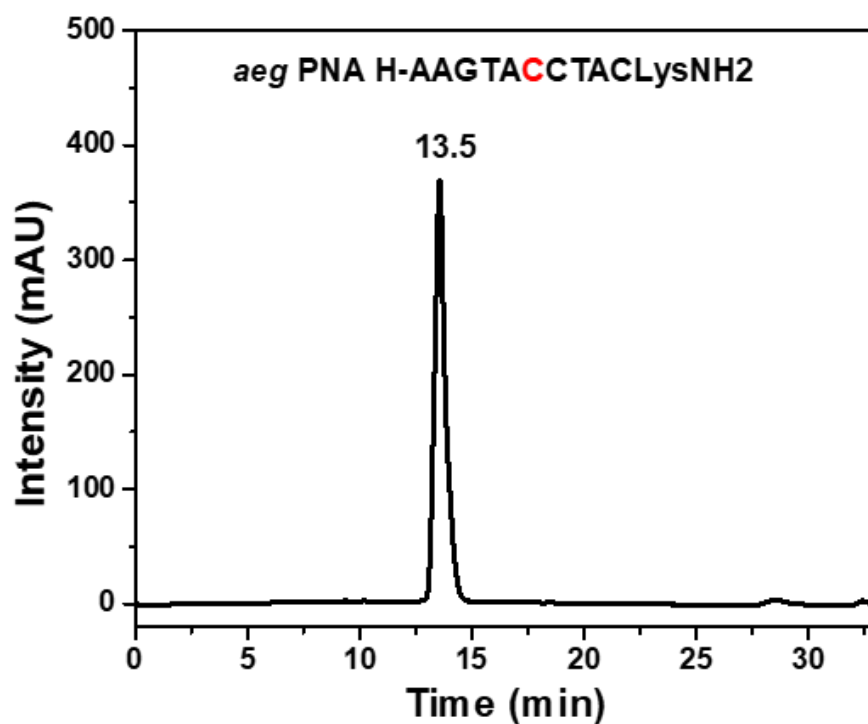
HPLC chromatogram of PNA 13



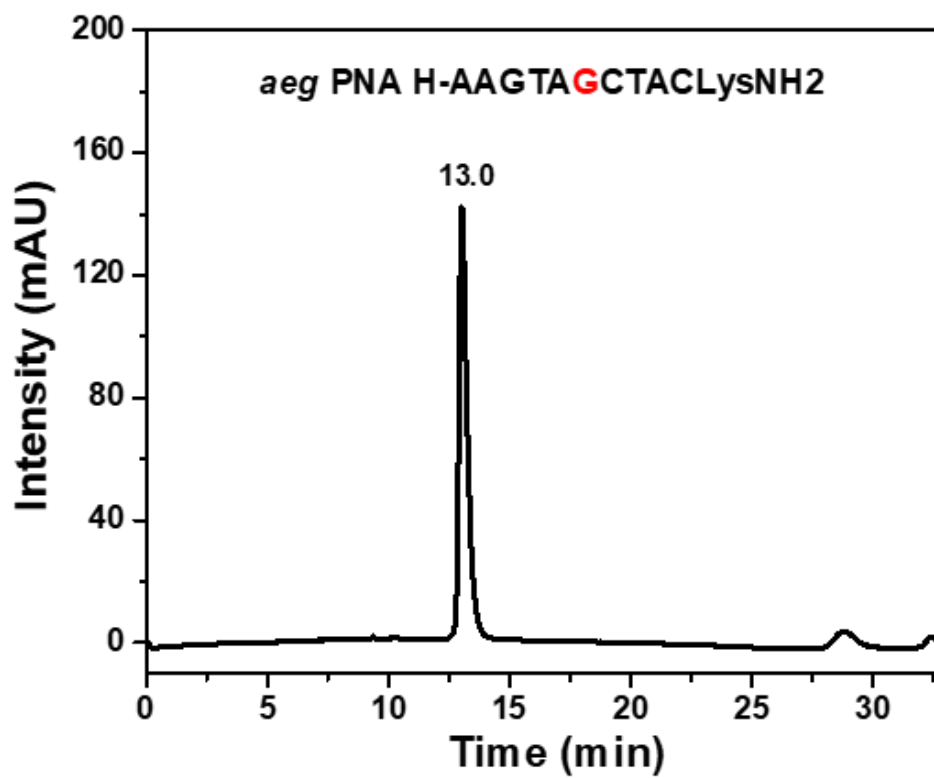
HPLC chromatogram of PNA 14



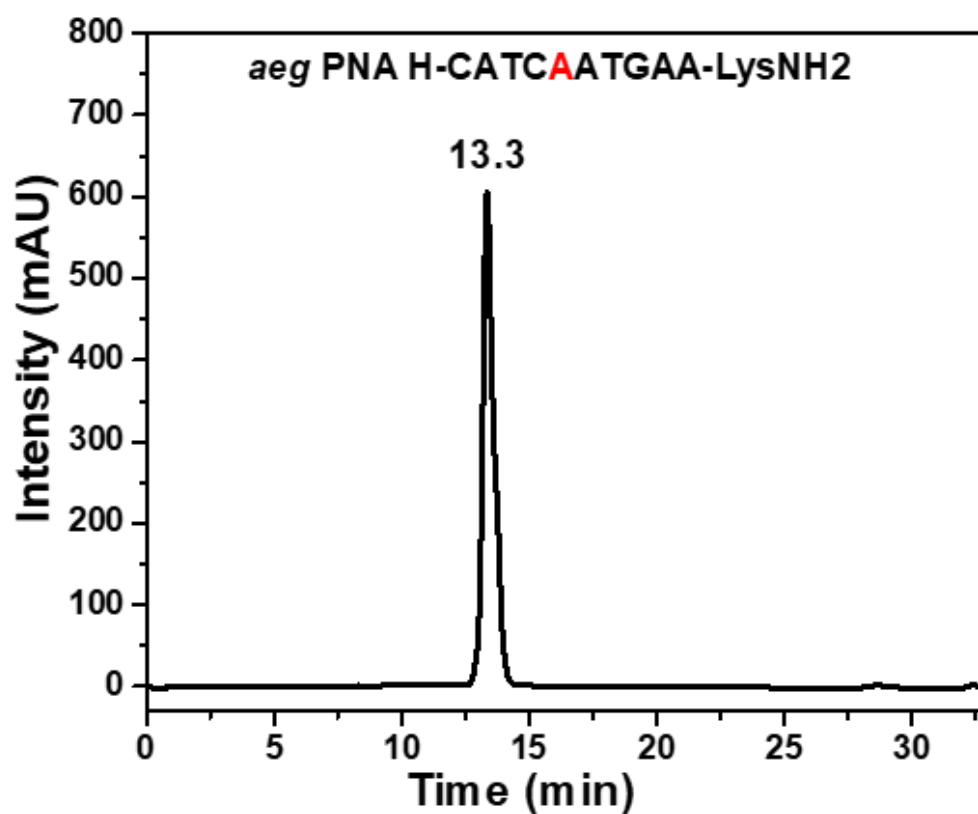
HPLC chromatogram of PNA 15



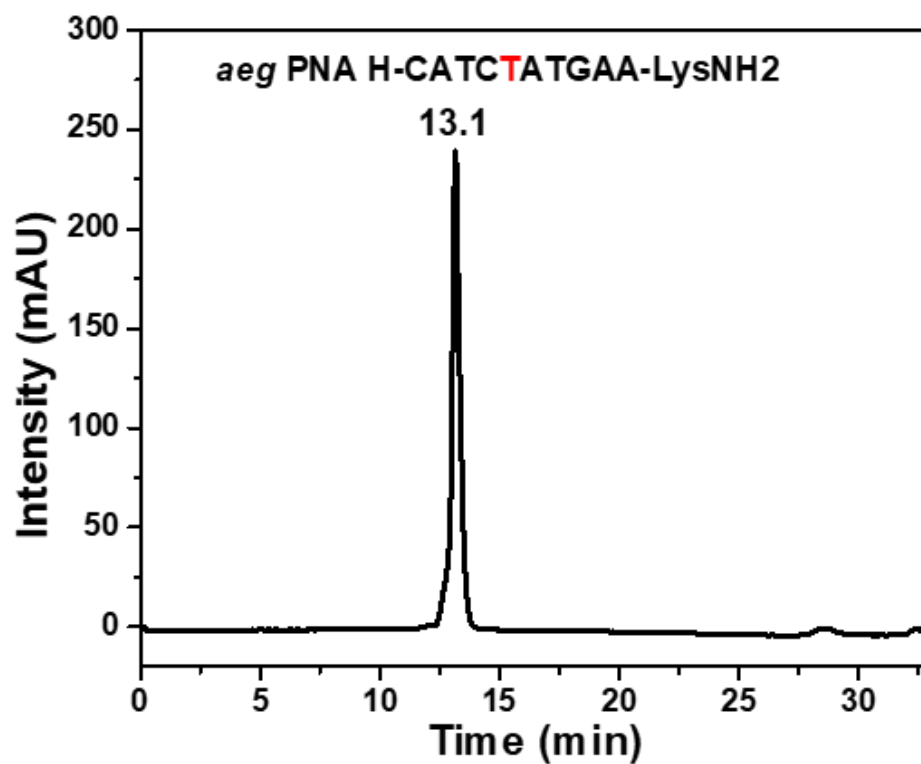
HPLC chromatogram of PNA 16



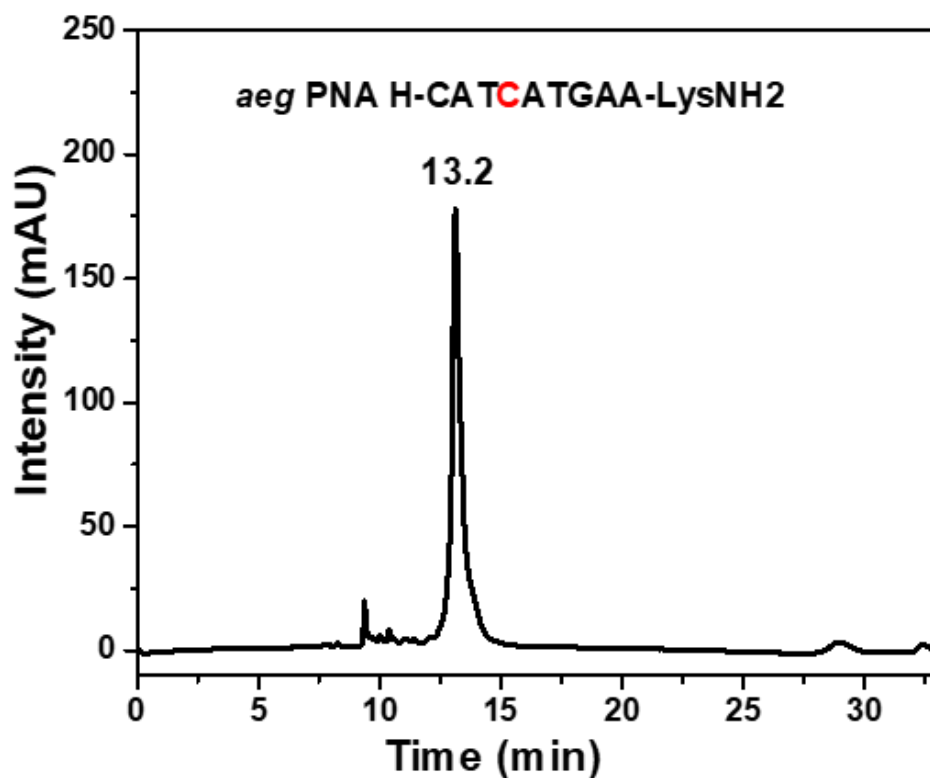
HPLC chromatogram of PNA 17



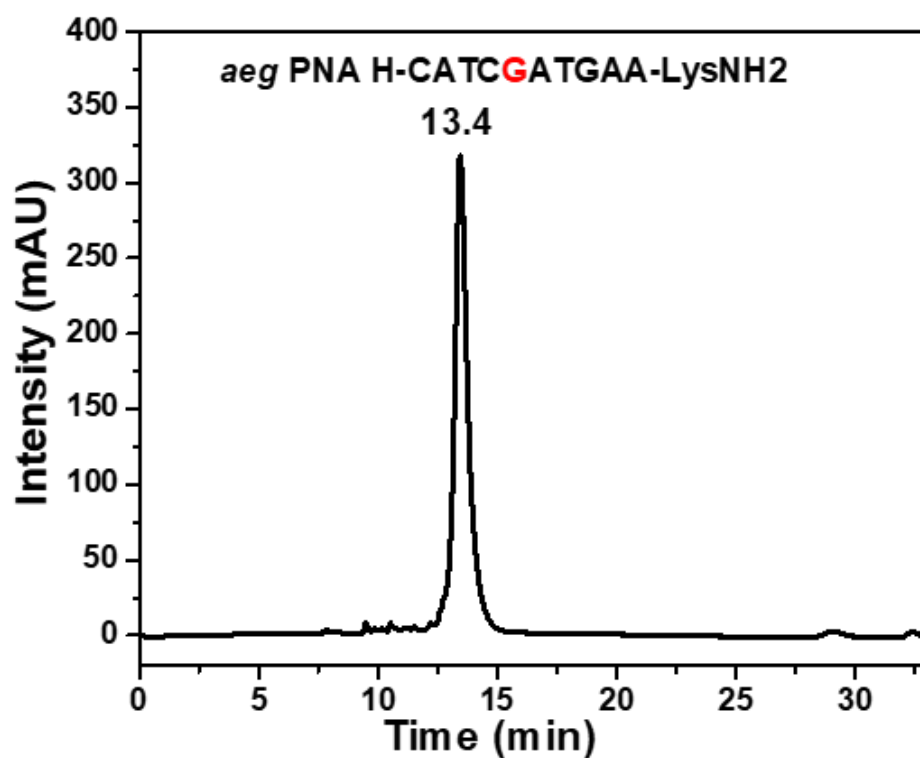
HPLC chromatogram of PNA 18



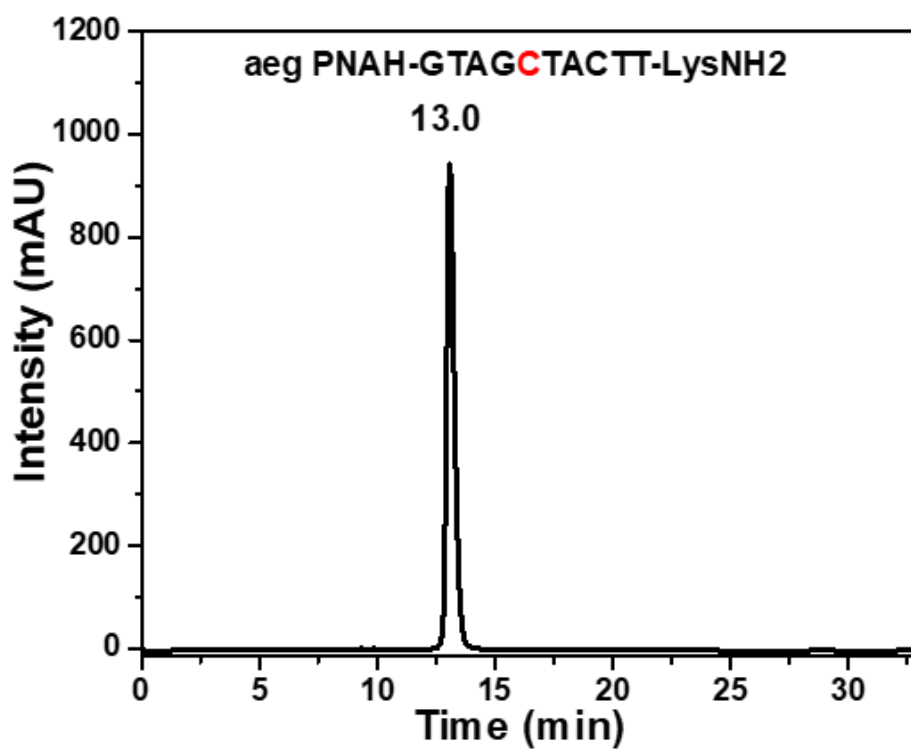
HPLC chromatogram of PNA 19



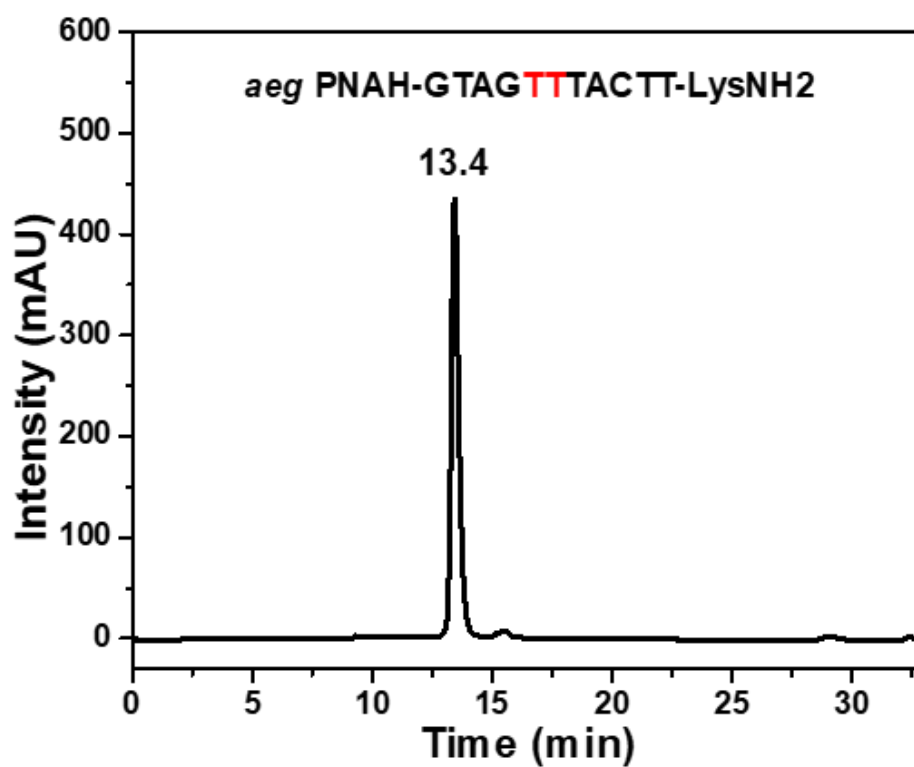
HPLC chromatogram of PNA 20



HPLC chromatogram of PNA 21



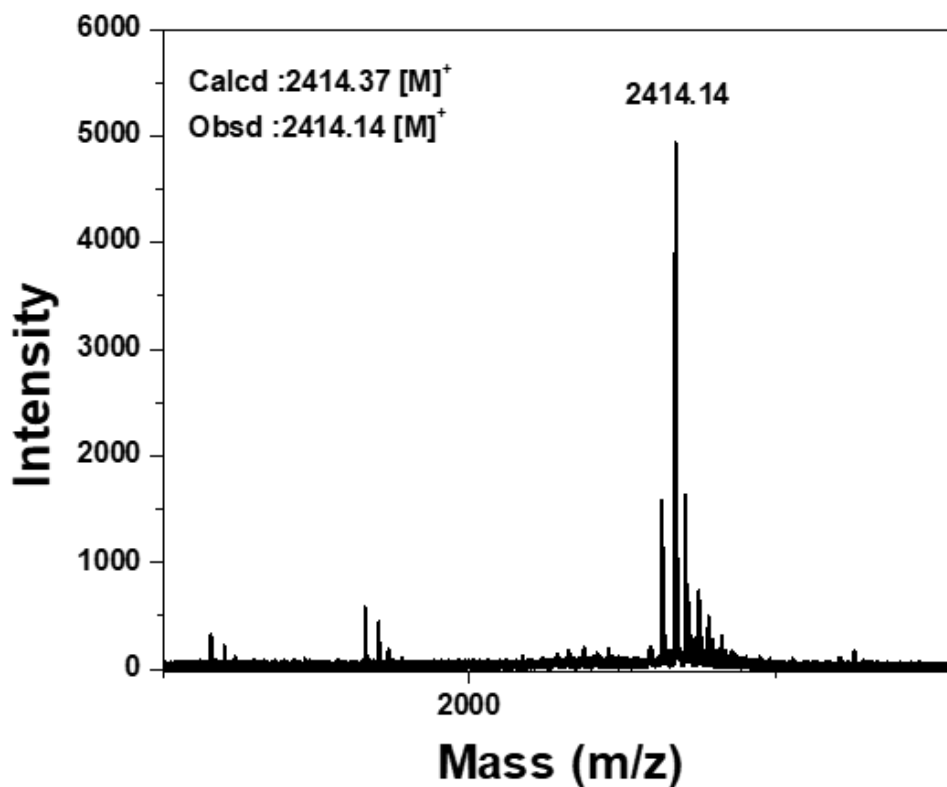
HPLC chromatogram of PNA 22



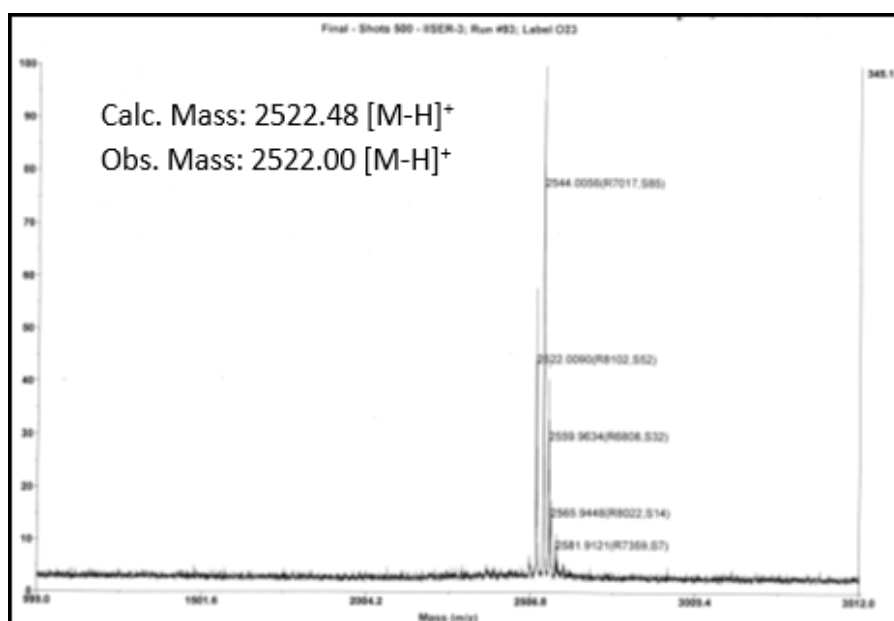


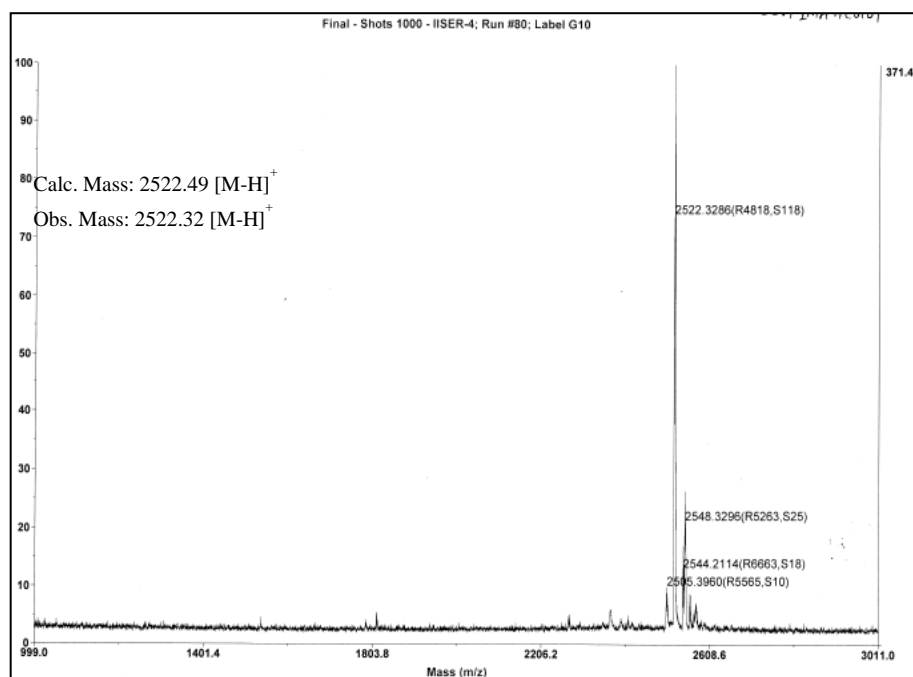
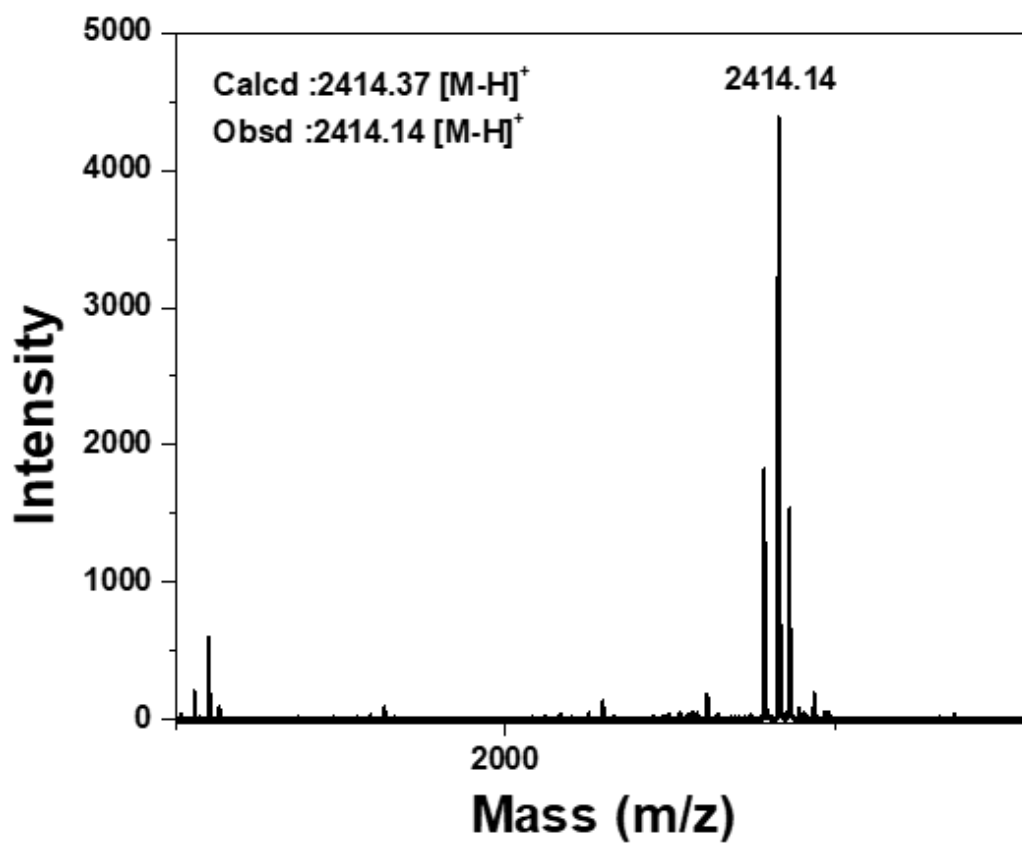
## MALDI-TOF Spectra of Peptides

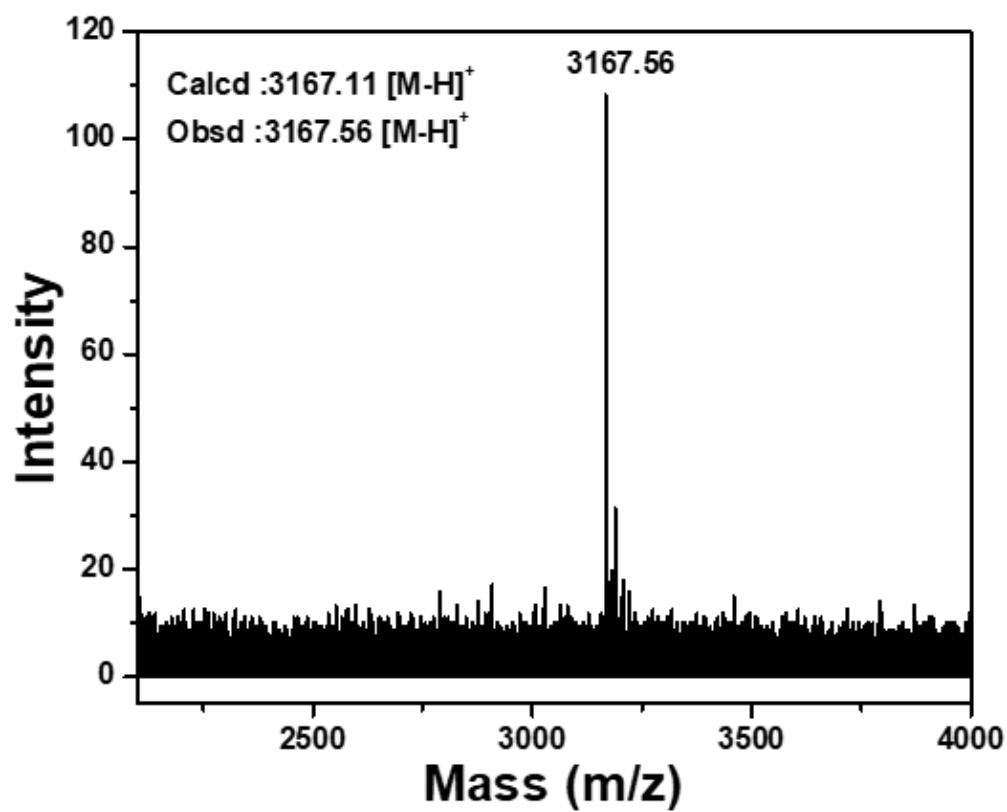
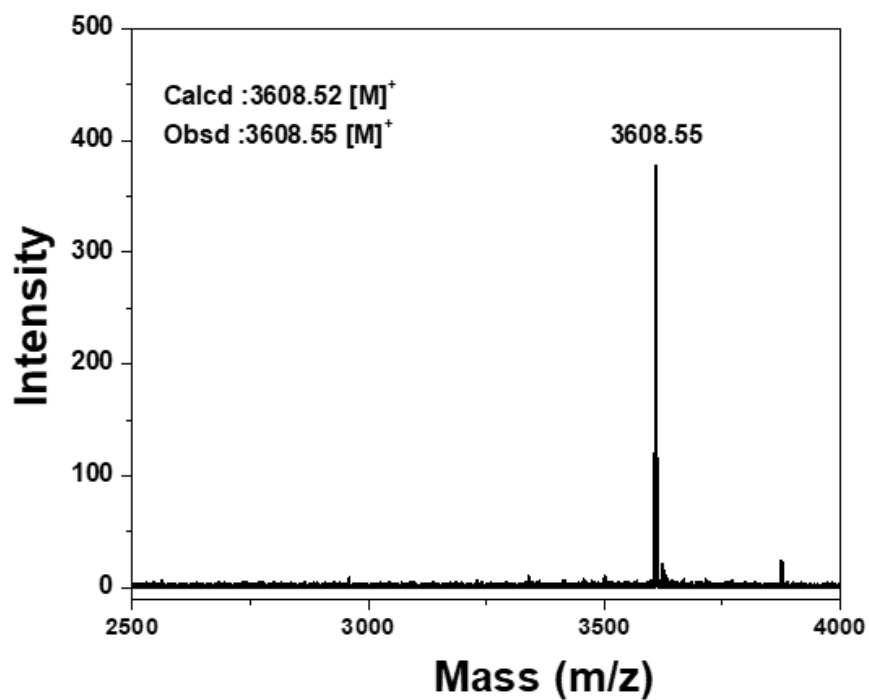
### MALDI-TOF Spectra of *aeg* PNA 1

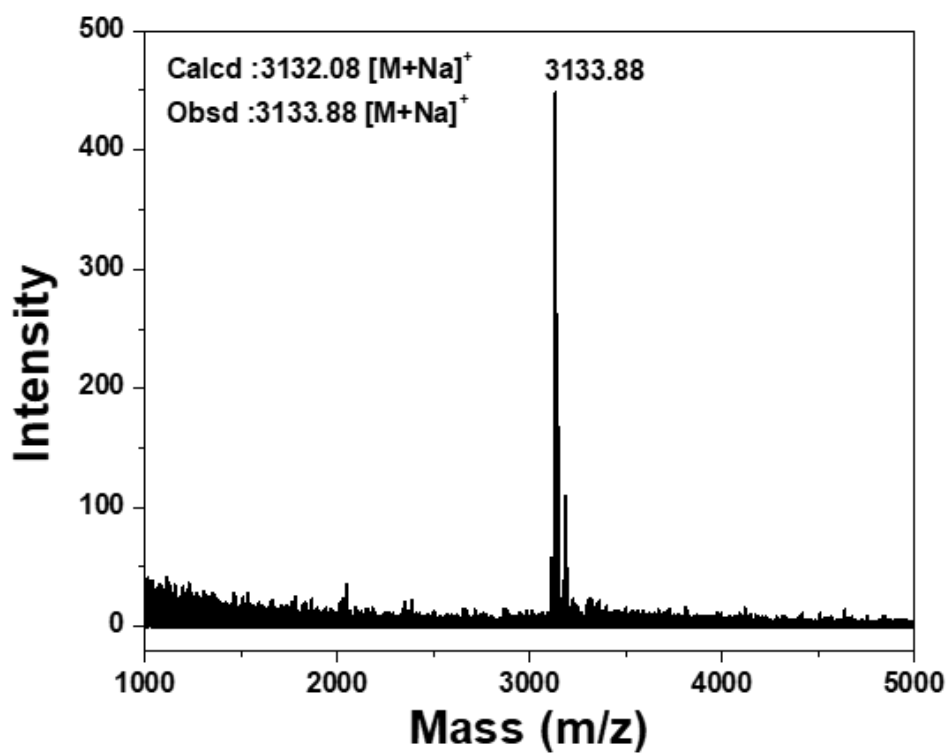
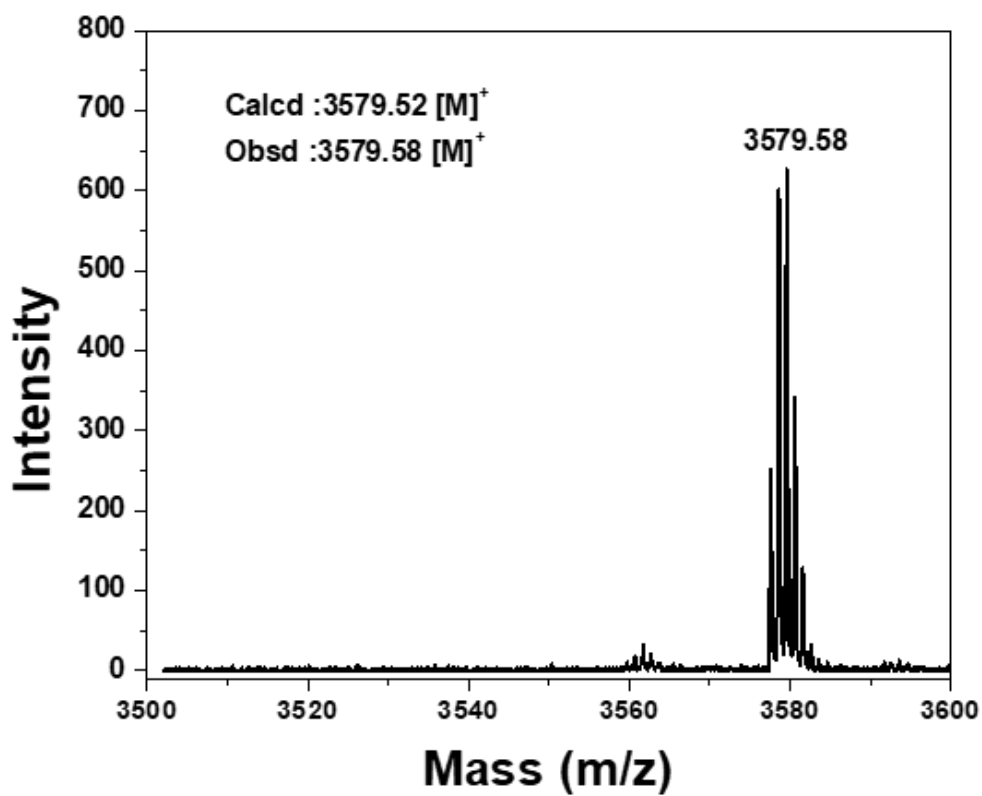


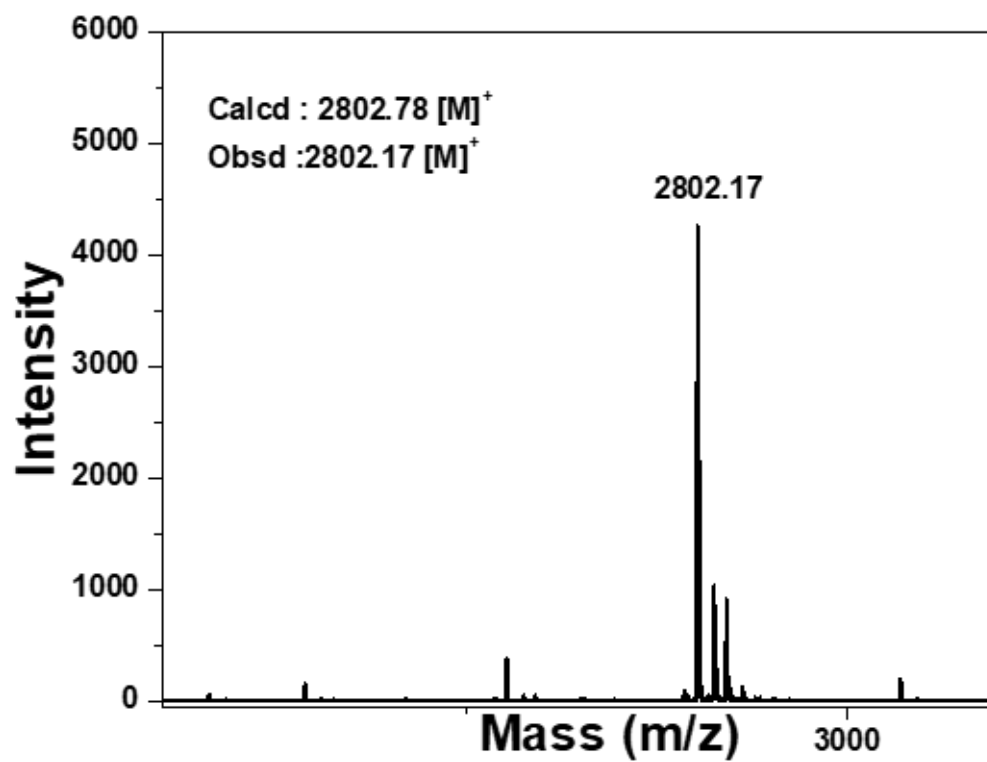
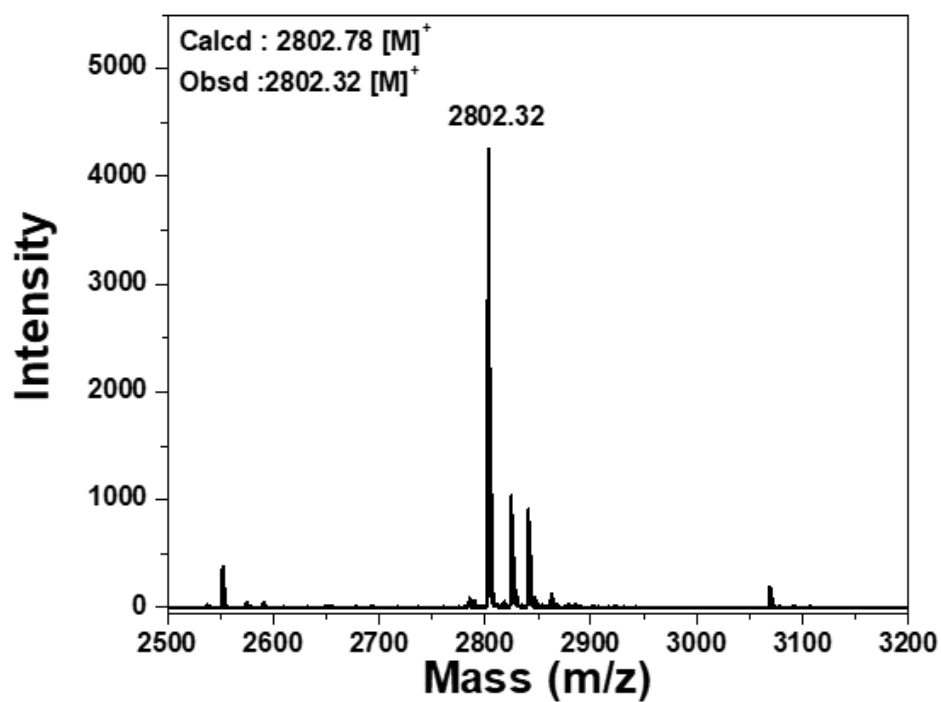
### MALDI-TOF Spectrum of *Im*(N1) PNA 2

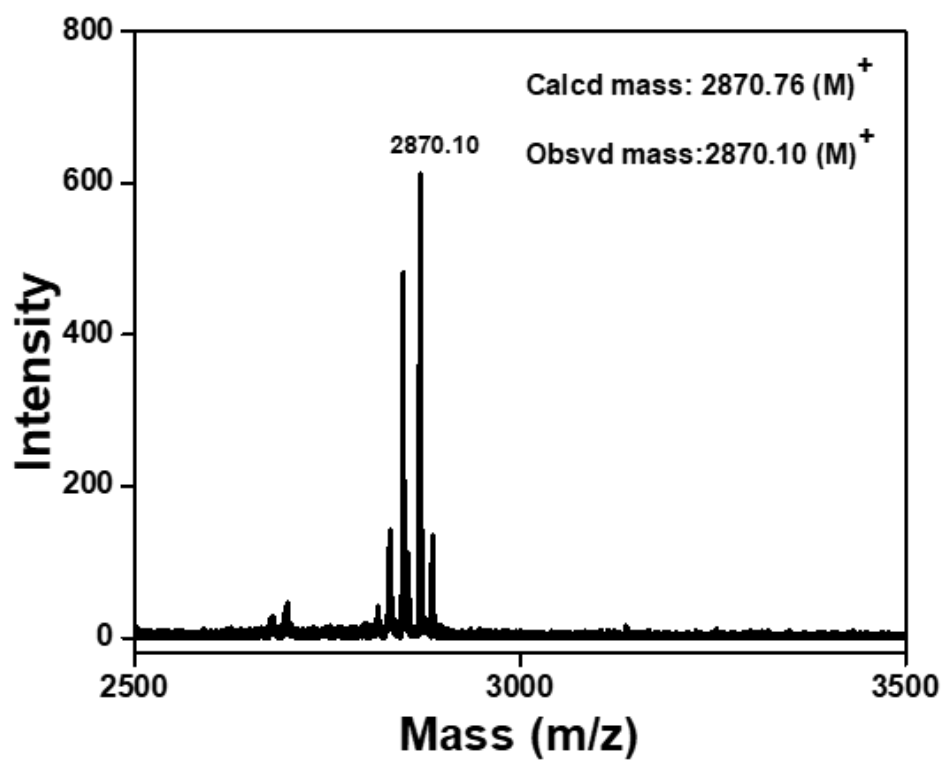
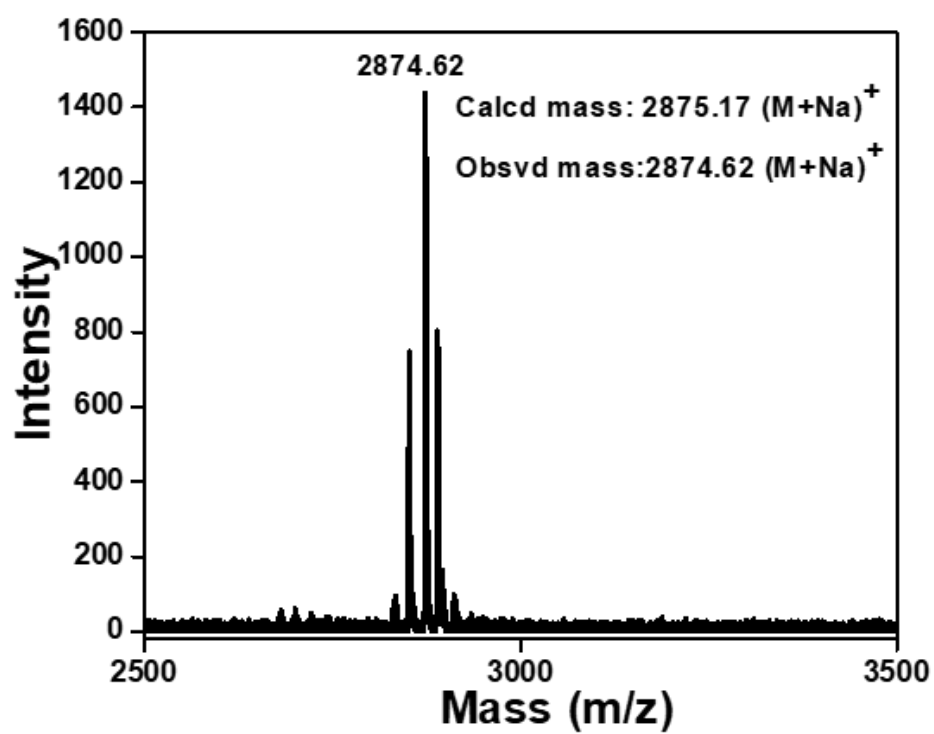


MALDI-TOF Spectrum of *Im*(C4) PNA 3MALDI-TOF Spectrum of *Bis-Im* PNA 4

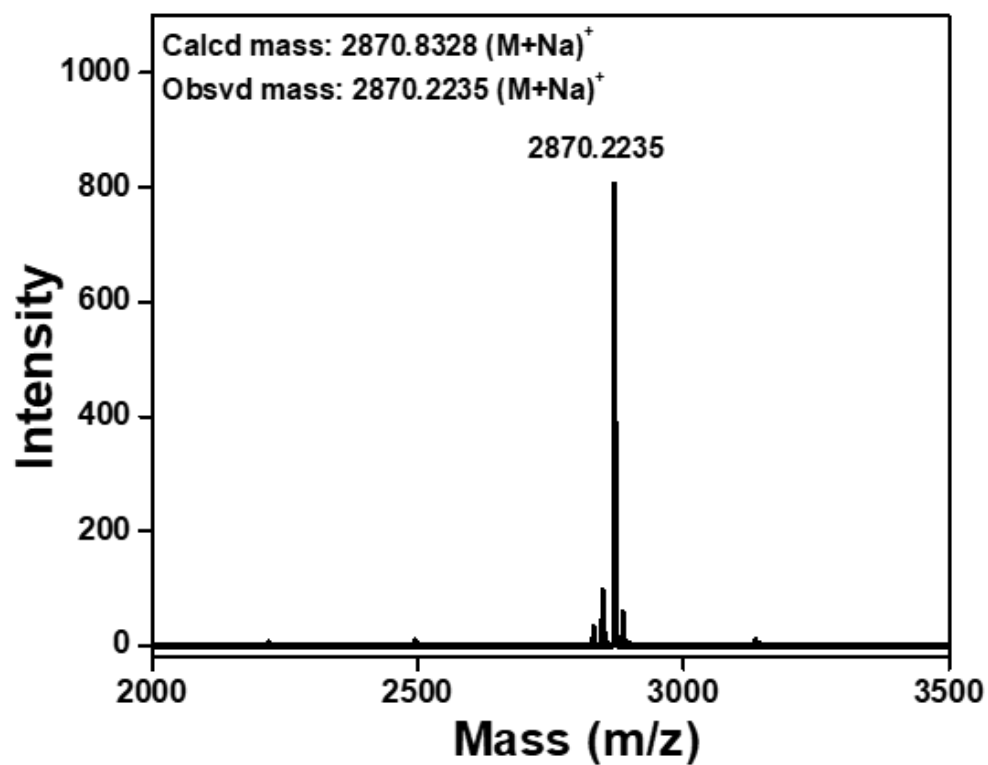
MALDI-TOF Spectrum of  $U^{lm}T_{10}$ -PNA 5MALDI-TOF Spectrum of  $U^{lm}$ -PNA 6

MALDI-TOF Spectrum of  $\gamma^{Im}$  T<sub>10</sub>-PNA 7MALDI-TOF Spectrum of  $\gamma^{Im}$ -PNA 8

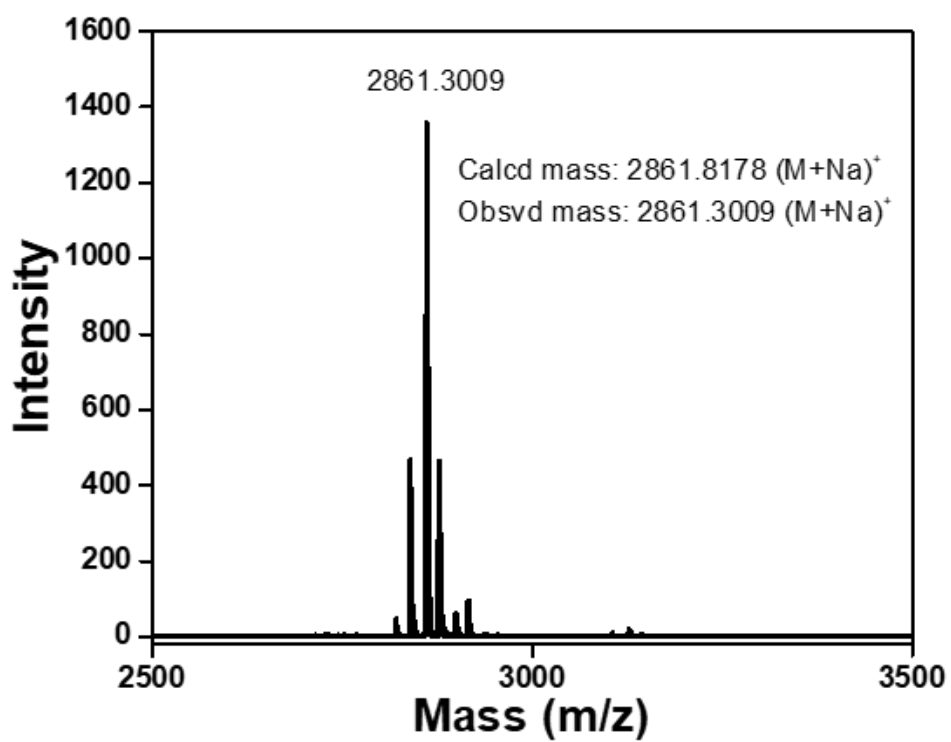
MALDI-TOF Spectrum of *Im(NI)*-PNA 9MALDI-TOF Spectrum of *Im(C4)*-PNA 10

MALDI-TOF Spectrum of *Im* NO<sub>2</sub> PNA **11**MALDI-TOF Spectrum of *BzIm*PNA **12**

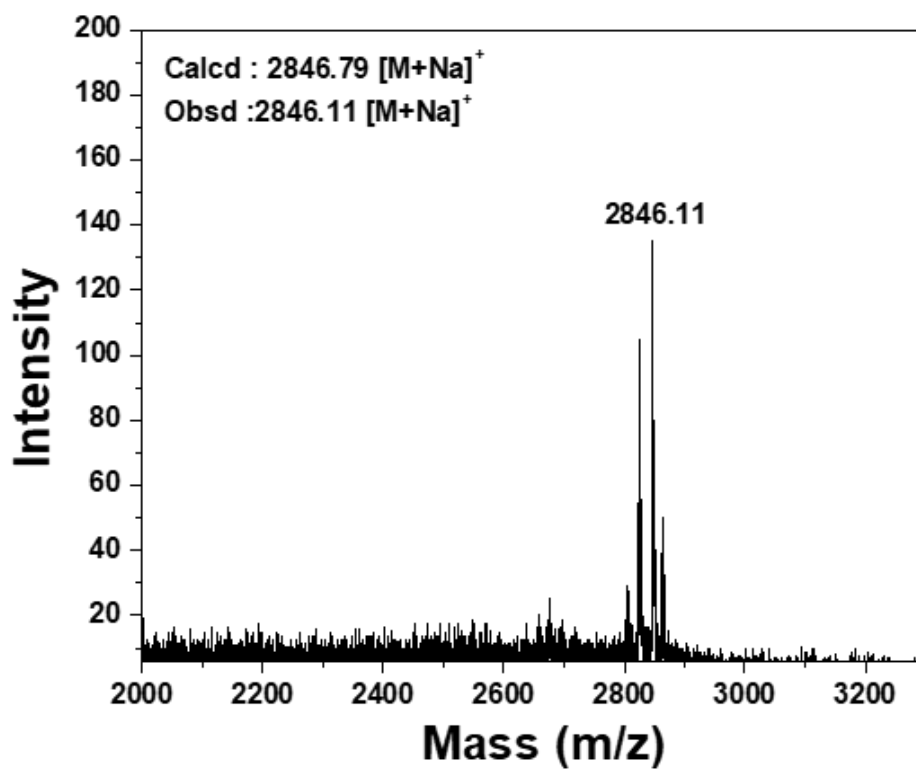
MALDI-TOF Spectrum of PNA 13



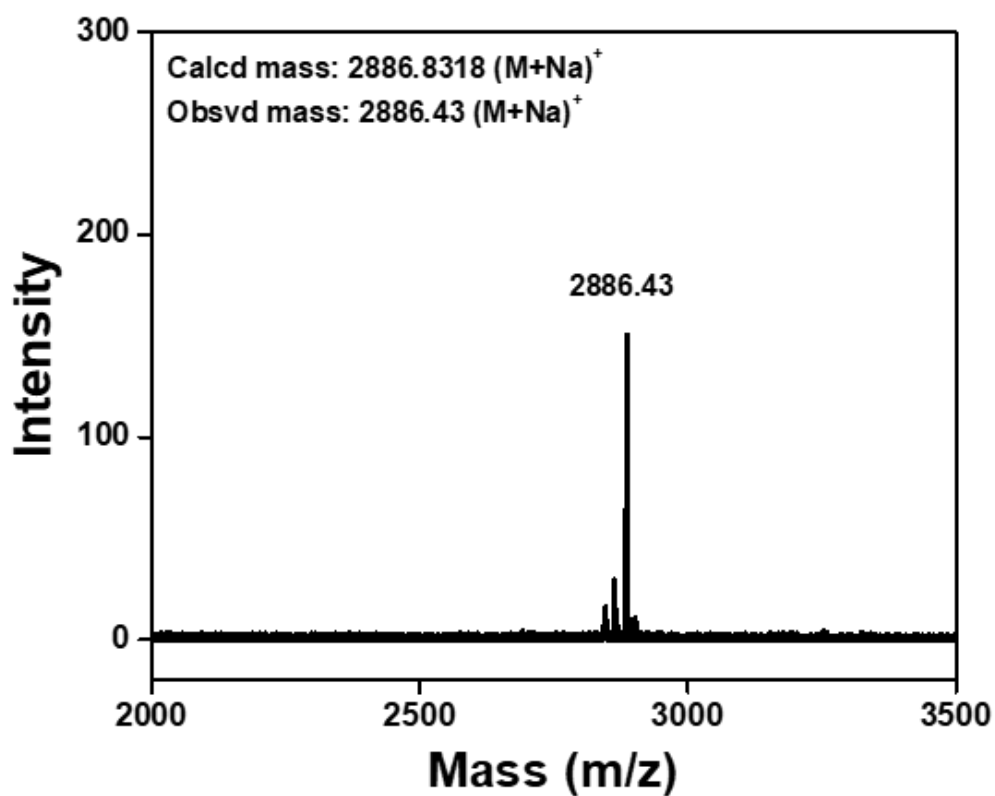
MALDI-TOF Spectrum of PNA 14



MALDI-TOF Spectrum of PNA 15

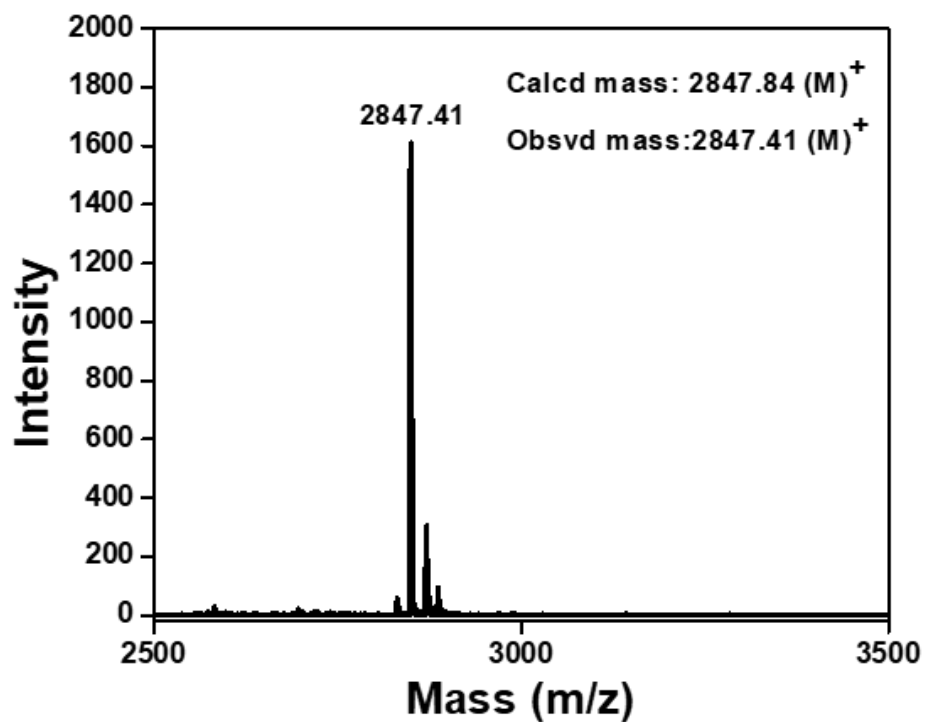


MALDI-TOF Spectrum of PNA 16

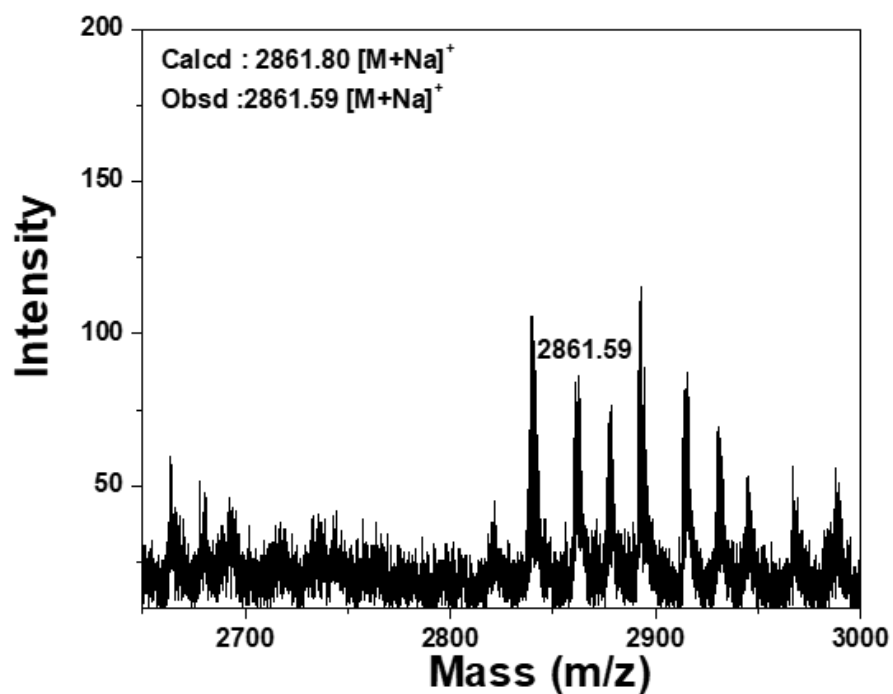




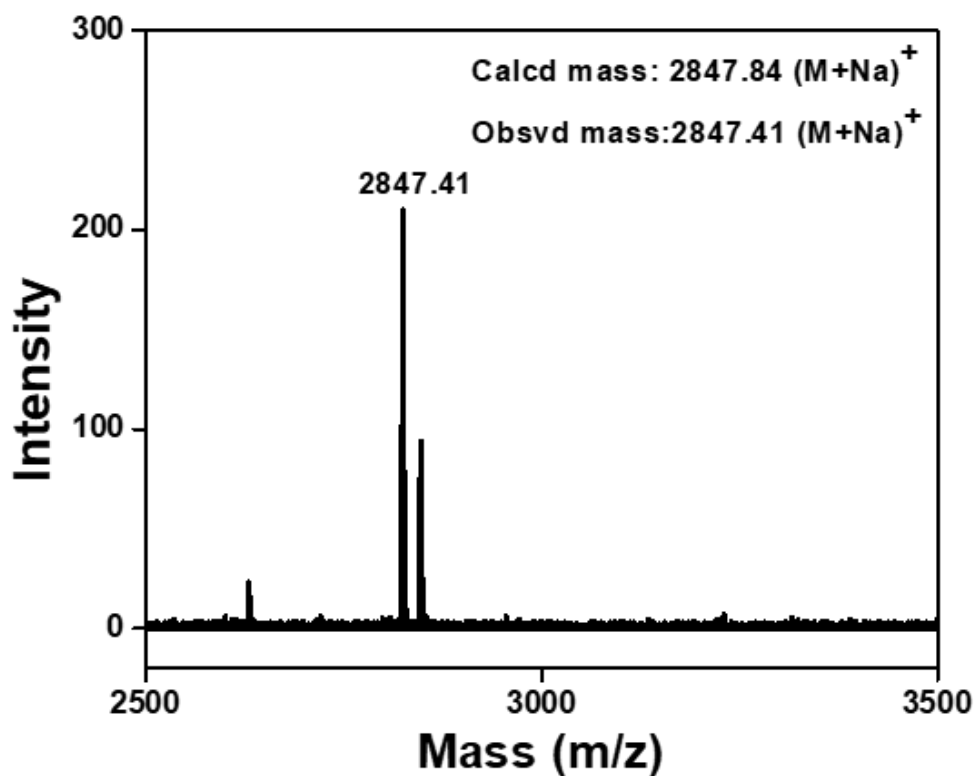
MALDI-TOF Spectrum of PNA 17



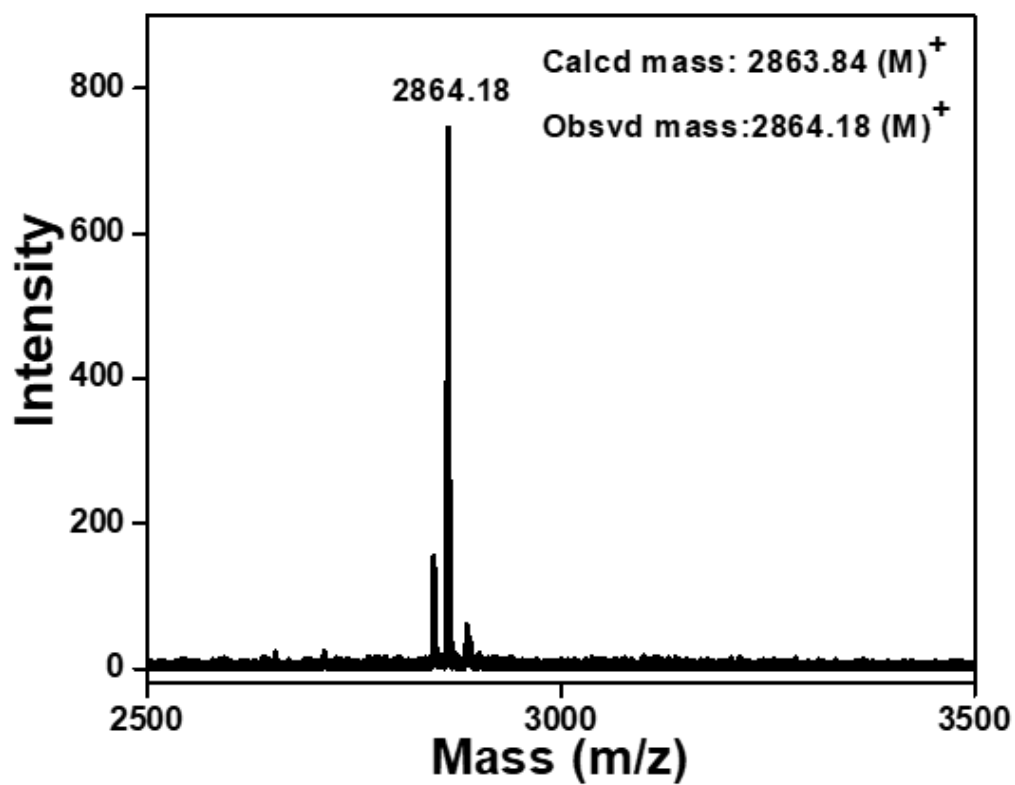
MALDI-TOF Spectrum of PNA 18



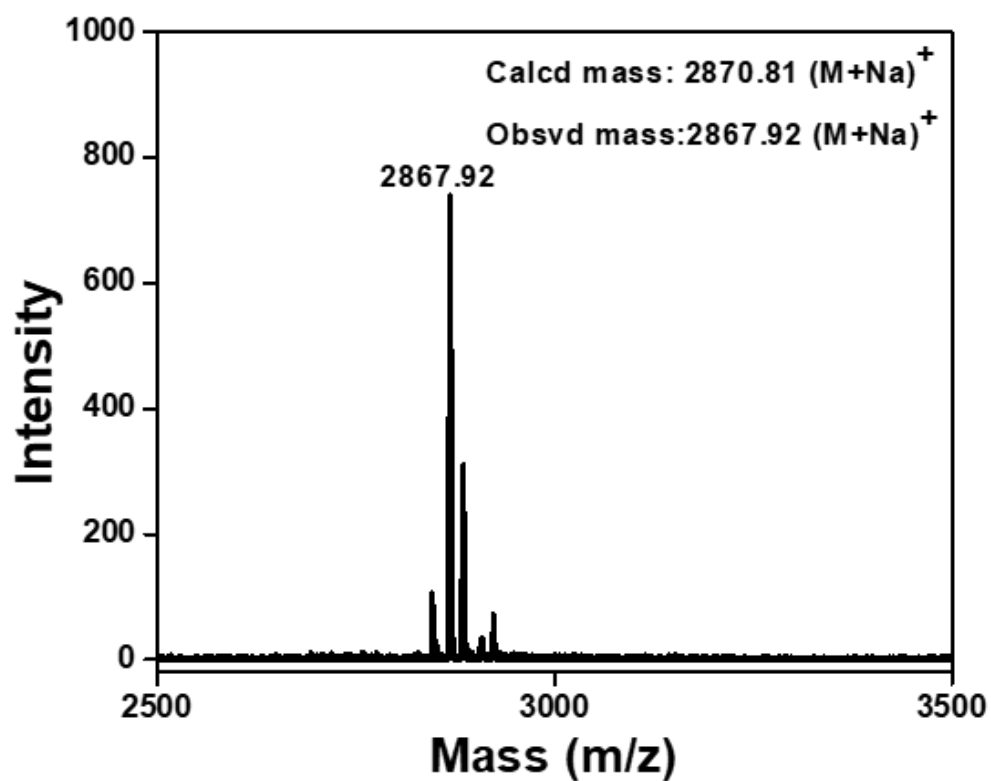
MALDI-TOF Spectrum of PNA 19



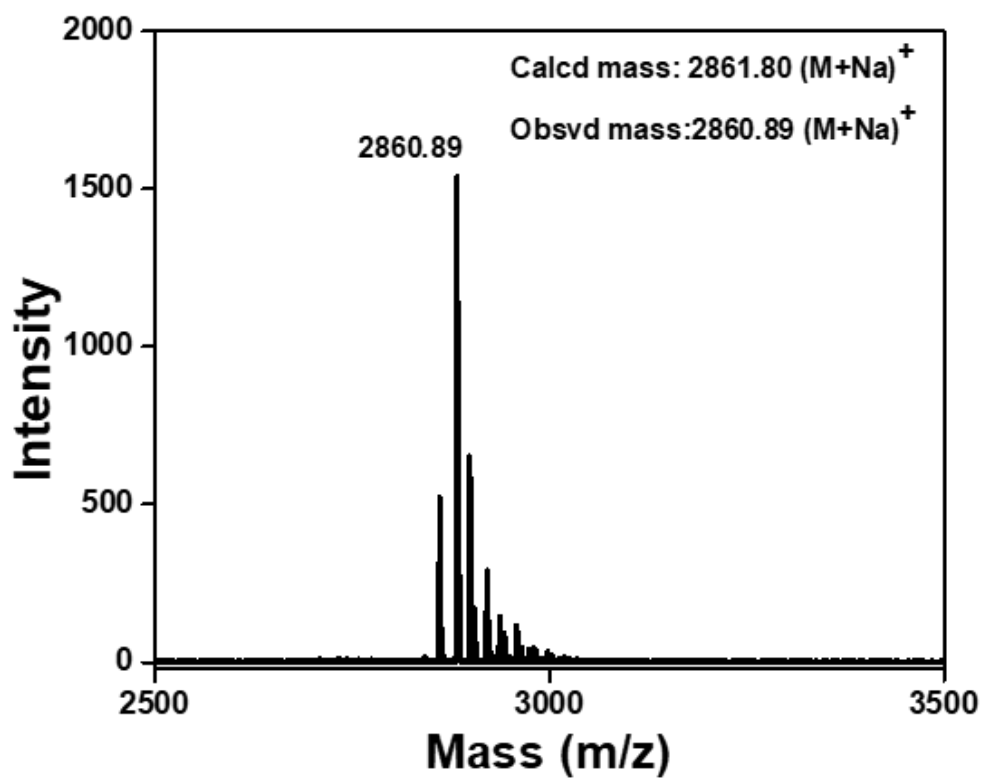
MALDI-TOF Spectrum of PNA 20



MALDI-TOF Spectrum of PNA 21



MALDI-TOF Spectrum of PNA 22



**Chapter 3**

**Biophysical Study of Imidazolyl**

**PNA**

### 3.1 Introduction

DNA double helix consists of two linear strands of oligodeoxyribonucleotides held together by specific hydrogen bonds between the complementary pairs of nucleobases on the two strands. James Watson and Francis Crick recognized the significant implications of the base-pairing of adenine (A) with thymine (T) and guanine (G) with cytosine (C) and suggested in their seminal paper on the double helical structure of DNA<sup>1</sup> that this elegant scheme could position DNA as the fulcrum of all the biological processes concerning storage and transmission of genetic information. It is because of this feature that all the genetic information stored in nucleic acids are passed on faithfully from one generation to the next.

While the two purines (A and G) and the three pyrimidines (U, T and C) were selected by billions of years of evolution, similar double helical constructions can be made with non-natural nucleobases as well<sup>2</sup> Innumerable variants of synthetic nucleobases have been reported in the literature and many have been shown to be useful as surrogates for the five naturally occurring nucleobases; for instance, 5-amino uracil is a substitute for thymine and may be a purine mimic.<sup>3</sup> These artificial nucleobases have enriched the genetic alphabet and have also been used to expand the genetic code. In addition to ensuring that they are compatible with the complementarity scheme, they also need to be compatible with various biological machinery and techniques. For example, amplification of oligonucleotides containing one or more of the nucleobases requires that the modification does not hinder the DNA polymerase enzyme used in polymerase chain reaction (PCR). Such design considerations are necessary to employ the unnatural nucleobases for various functions.

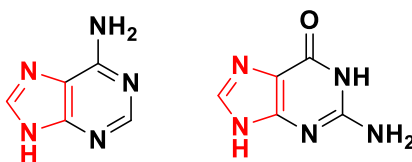
Over the years, libraries of base pairs have also been developed. In 1962, Alexander Rich first proposed a third base pair to supplement the natural A-T and G-C pairs.<sup>4</sup> In the late 1980s, four types of unnatural base pairs were shown by Benner *et. al.*,<sup>5</sup> with nonstandard hydrogen bonding patterns, which differed from the natural base pairs. Meanwhile, Kool *et al.*,<sup>6</sup> synthesized nonhydrogen-bonded base pairs between 4-methylbenzimidazole and 2,4-difluoro-1-methylbenzene, which are analogous in shape to the natural nucleobases. This work demonstrated the importance of shape complementarity, in the natural base-pairing. Many hydrophobic base analogues were synthesized by Romesberg's group<sup>7</sup> and they have found optimized unnatural base pairs for Polymerase

chain reaction (PCR).<sup>8,9</sup> Yamshige *et.al*<sup>10</sup> has also developed highly specific hydrophobic base pairs. Constant improvements by combining the concepts of nonstandard hydrogen bonding patterns, shape-complementarity, hydrophobicity, and electrostatic repulsion are going on.<sup>11-15</sup>

Among the unusual nucleobases, universal nucleobases are those that can base-pair with all the four natural nucleobases without any significant discrimination. Several modified nucleobases have been tested for potential universal base-pairing property. These include naturally occurring hypoxanthine,<sup>16</sup> and synthetic nucleobases such as 3-nitropyrrole,<sup>17-19</sup> 5-nitropyrrole,<sup>20-23</sup> andazole-4-carboxamides.<sup>24-26</sup> Universal bases can be used for construction of nucleic acid probes and primers, where the identity of the base in the target nucleic acid is unknown. Therefore, these can be useful in molecular biology and biotechnolog

### 3.2 Rationale

In nature, the base-pairing is typically between a purine and a pyrimidine. As shown in Figure 3.1, imidazole moiety is an integral part of the natural purine nucleobases, adenine and guanine, since purine is a fusion of the pyrimidine and imidazole rings.

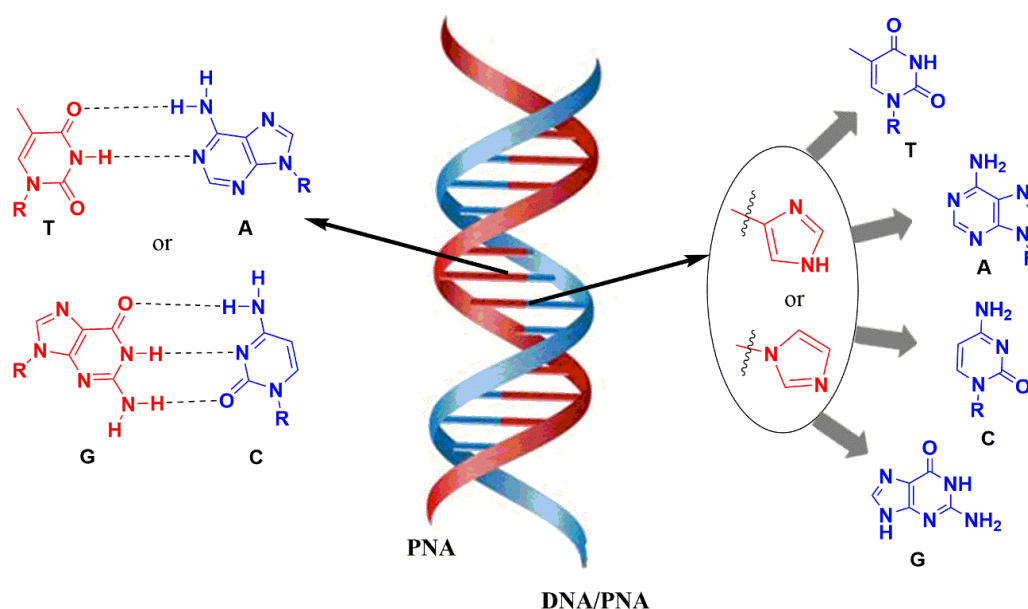


**Figure 3.1** Chemical structure of purines A and G showing the constituent imidazole ring.

It is worth noting that in purines, only the N1 and the substituents at C2 and C6 (of the constituent pyrimidine ring) are involved in the hydrogen bonding interactions with the complementary bases in the Watson-Crick system, which determines sequence selectivity among nucleic acids. Even though the 5-membered imidazole ring participates in Hoogsteen base pairing in triplexes, and also involved in quadruplexes formed by G, it is not involved in the canonical Watson-Crick base pairing scheme.<sup>27</sup>

While replacement of the usual nucleobases with an imidazole has been well-documented in DNA oligonucleotides,<sup>28</sup> the base-pairing properties of imidazoles in PNA and its hybrids (Figure 3.2) is not well studied. Since the strength of the base-pairing interactions between nucleobases depend on whether the nucleobases are conjugated to the

sugar phosphate backbone of DNA or the pseudopeptide backbone of PNA, this investigation is non-trivial. It also assumes importance by the fact that results of such investigations could augment our understanding of base-pair engineering specifically for PNA based systems and help in the rational design of self-assembling motifs that employ PNA.

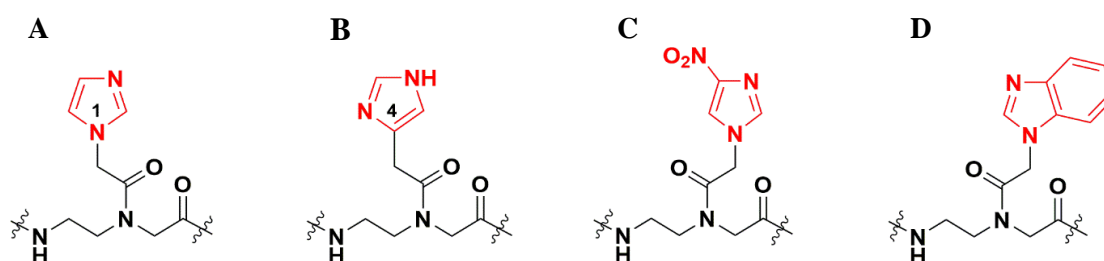


**Figure 3.2** Schematic representation of duplexes of PNA:DNA and PNA:PNA showing the canonical base-pairs and the imidazolyl replacements used in this study

Imidazole moiety can be linked to the PNA through N1, as in the purine nucleobases, or through C4. The way the imidazole moiety is linked to the backbone is likely to alter the nature and the strength of its interactions with nucleobases in the complementary strand. Therefore, hybridization properties of N1-linked imidazolyl PNA modifications and C4-linked imidazolyl PNA modifications would be interesting to study (Figure 3.2).

Among the unnatural nucleobases, 3-nitropyrrole is interesting since it can complement with any of the four standard nucleobases. This universal base shares its structural properties with 3-nitroimidazole. Imidazole and 3-nitroimidazole, like 3-nitropyrrole, have five membered rings and are both aromatic in nature. However, they differ from 3-nitropyrrole in the number of nitrogen atoms in the ring; while the imidazole has two, 3-nitropyrrole has only one nitrogen atom in the ring. Examining the influence of this structural difference in the molecular recognition property of PNA oligomers could help in designing base-pairing moieties for PNA based molecular recognition schemes and

devices that exploit broad specificity afforded by such interactions. One of the distinct features of PNAs is their low tolerance to mismatch. Because this property minimizes off-target interactions, it is an advantage in various applications; however, it may also restrict its application. The low tolerance to mismatch allows targeting a very specific sequence. The ability to broaden the specificity at particular sites in a sequence can be useful for targeting regions in the genome with consensus sequences where the degree of variability in the sequences can be expected to be slightly more. Since biological targets, especially the nucleic acid targets are replete with such variations, concessions in the sequence selectivity without hampering stability of the duplex can help in the design of better PNA probes and therapeutic agents for a wide variety of applications. To this end, the potential role of both N-alkylated and C-alkylated imidazole as universal base in the modified PNA will be explored in this chapter (Figure 3.3).



**Figure 3.3** PNA modifications used in the present investigation: A) N1-linked imidazolyl PNA, C4-linked imidazolyl PNA, C) N1-linked nitroimidazolyl PNA, and D) N1-linked benzimidazolyl PNA

Since modifications at the middle of sequence are likely to produce more drastic changes, the PNA oligomers in this study were designed to have a single modification at position 5. To further analyse the nature of interactions, oligomers with 4-nitroimidazole, and with benzimidazole will be synthesized and their properties compared with the imidazole modified PNA oligomers. The molecular recognition properties and base-pairing preferences of these oligomers towards complementary DNA/PNA strands are compared with those of oligomers synthesized with 4-nitroimidazole and benzimidazole modifications at the corresponding position.

### 3.3 Objectives

This study aims to investigate how replacement of the canonical nucleobases with an imidazole moiety in a PNA strand influences base-pairing specificity with its complementary PNA and DNA. For this purpose, PNA oligomers with an aminoethylglycyl

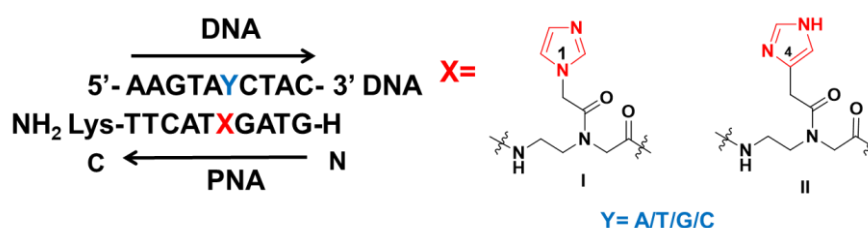


(*aeg*) backbone bearing a single imidazole modified monomeric unit in the middle of the sequence were used, as shown in Chapter 2.

The objectives of this chapter include the following.

### 3.3.1 Studies with cDNA

- Hybridization of the modified PNA with complementary *antiparallel DNA* strands containing the four natural nucleobases opposite the site of modification (Figure 3.4).

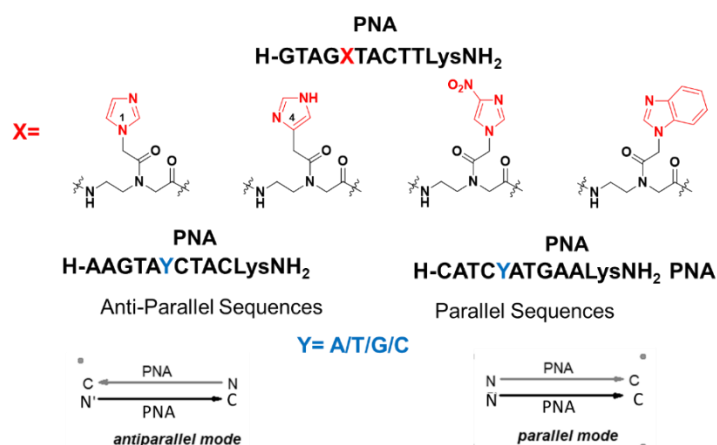


**Figure 3.4** Antiparallel PNA:DNA hybrids used in this study

- Evaluation of thermal stability of the hybrid duplexes and their comparative analysis to understand whether the imidazolyl modifications can distinguish nucleobases in the complementary DNA strand.
- Circular dichroic studies to establish the effect of modification on the helical structure of the hybrid duplexes.

### 3.3.2 Studies with cPNA

- Hybridization of the modified PNA with complementary *parallel and antiparallel PNA* strands containing the four natural nucleobases opposite the site of modification (Figure 3.5).



**Figure 3.5** Antiparallel and parallel PNA:PNA hybrids used in this study

- Evaluation of thermal stability of the PNA duplexes and their comparative analysis to examine whether the imidazolyl modifications can distinguish nucleobases in the complementary PNA strand; whether the orientation of the complementary strand (parallel vs antiparallel) affects the base-pairing properties of the nucleobases; whether the nature of the complementary strand (PNA vs DNA) affects the molecular recognition properties.
- Circular dichroic studies to establish the effect of modification on the helical structure of the PNA:PNA duplexes.

### 3.4 Results and Discussion

Introduction of imidazole modifications on a 10-mer *aeg* PNA oligomer was expected to affect the stability of duplexes with its complementary PNA or DNA. The PNAs with and without modifications that are used in this investigation is listed in Table 3.1. The effect of the modifications on the stability of the duplex and conformational integrity is studied in this chapter by using temperature dependent UV absorption spectroscopy (UV-*T<sub>m</sub>*) and CD spectroscopy.

**Table 3.1 PNA oligomers used in the present study**

Entry	Sequence Code	PNA sequences	Remarks
1	<i>aeg</i> PNA 1	H-GTAGCTACTTlysNH <sub>2</sub>	Control
2	<i>Im</i> (N1) PNA 2	H-GTAG <i>Im</i> (N1)TACTTlysNH <sub>2</sub>	N1-Linked Imidazole PNA
3	<i>Im</i> (C4) PNA 3	H-GTAG <i>Im</i> (C4)TACTTlysNH <sub>2</sub>	C4-Linked Imidazole PNA
4	<i>Im</i> NO <sub>2</sub> PNA 4	H-GTAG <i>Im</i> NO <sub>2</sub> TACTTlysNH	3-Nitroimidazole PNA
5	<i>Bzim</i> PNA 5	H-GTAG <i>Bzim</i> TACTTlysNH <sub>2</sub>	Benzimidazole PNA
6	PNA 6	H-AAGTATCTAClysNH <sub>2</sub>	
7	PNA 7	H-AAGTACCTAClysNH <sub>2</sub>	
8	PNA 8	H-AAGTAGCTAClysNH <sub>2</sub>	Complementary antiparallel PNAs
9	PNA 9	H-AAGTAACTAClysNH <sub>2</sub>	

10	<b>PNA 10</b>	H-CATCTATGAALysNH <sub>2</sub>	
11	<b>PNA 11</b>	H-CATCCATGAALysNH <sub>2</sub>	Complementary parallel PNAs
12	<b>PNA 12</b>	H-CATCGATGAALysNH <sub>2</sub>	
13	<b>PNA 13</b>	H-CATCAATGAALysNH <sub>2</sub>	

### 3.4.1 Thermal melting studies of antiparallel PNA:DNA hybrids

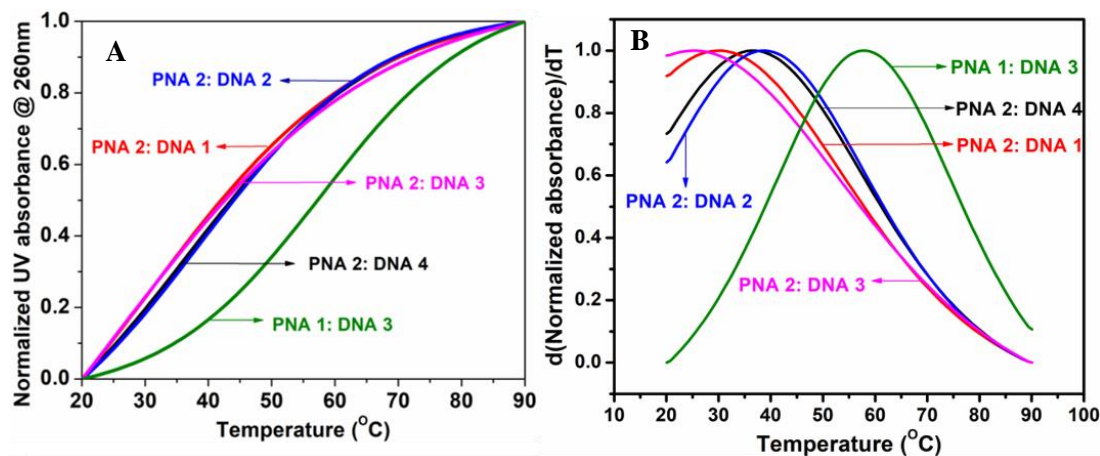
As a DNA mimic, PNA is known to hybridize with complementary DNA forming duplexes similar to the standard DNA:DNA double helix. With a neutral backbone, the PNA strands form duplexes that are thermodynamically more stable than the typical DNA:DNA duplexes. Similar to DNA duplexes, PNA:DNA hybrids also exhibit hyperchromicity on denaturation. Therefore, the thermal stability of the hybrid duplexes can be studied by monitoring the changes in the absorbance at 260 nm during a heat-denaturation process. The midpoint of the transition where half the molecules exist in duplex form and the other half in single stranded form is defined as melting temperature ( $T_m$ ). It is typically used as a measure of thermal stability of the PNA:DNA duplexes. Comparison of the measured  $T_m$  values between the duplexes formed by modified PNAs and the unmodified ones provides a way to evaluate the effect of the modifications on the strength of base-pairs at the site of modification.

In the present study, the **PNAs 2-5** were individually hybridized with respective complementary DNAs **1-4**. This produced a combination of duplexes where the unnatural nucleobase (N1-linked imidazole, C4-linked imidazole, 3-nitroimidazole, and benzimidazole) in the PNA is held opposite each of the natural nucleobases (A, G, C, and T). The thermal stability of each of these duplexes was studied in 10 mM phosphate buffer at pH 7.2 in the presence of 10 mM NaCl. The  $T_m$  measured for the control pair of PNA:DNA hybrid duplex formed between unmodified **PNA 1** and **DNA 3** under the same conditions was compared to evaluate the stabilizing or destabilizing effect of the modification.

#### 3.4.1a Duplexes of N1-Imidazolyl PNA 2 with cDNAs 1-4

In order to investigate the effect of PNAs with a single N1-imidazole modification in their corresponding duplexes, **PNA 2** was hybridized with antiparallel complementary DNAs (**DNA 1 - DNA 4**). The UV-melting transition of these duplexes are shown in (Figure 3.6A).

The first derivative of the sigmoidal plots was generated for each of the duplexes to identify the respective melting temperature ( $T_m$ ) (Figure 3.6B).



**Figure 3.6** UV-Melting curves of PNA:DNA duplexes (A) and their first derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.2** UV- $T_m$  values of duplexes of N1-imidazole modified PNA 2 with complementary DNA

Sr. No.	PNA : DNA Duplex Sequence		UV- $T_m$ ( $^{\circ}\text{C}$ )	
			$T_m$	$\Delta T_m$
1	PNA 1 DNA 3	H-GTAGCTACTT-Lys NH <sub>2</sub> 3'-CATCGATGAA-5'	57.6	-
2	PNA 2 DNA 1	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> 3'-CATC T ATGAA-5'	30.0	-27.6
3	PNA 2 DNA 2	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> 3'-CATC C ATGAA-5'	38.5	-19.1
4	PNA 2 DNA 3	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> 3'-CATC G ATGAA-5'	25.0	-32.6
5	PNA 2 DNA 4	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> 3'-CATC A ATGAA-5'	36.3	-21.3

The  $T_m$  values are accurate to  $\pm 0.5$   $^{\circ}\text{C}$ .

The PNA:DNA duplex derived from unmodified *aeg* PNA 1 showed a melting ( $T_m$ ) of 57.7  $^{\circ}\text{C}$  (Table 3.2, entry 1). In case of duplexes formed by N1-imidazole modified PNA 2 with its complementary DNA strands DNA 1–4, the  $T_m$  values were found to be significantly lesser than that of the control *aeg* PNA. The differential melting temperature ( $\Delta T_m$ ), which is the difference between the  $T_m$  observed with the modified PNA and that obtained from the unmodified PNA with a G instead of the imidazole at the site of modification, gives the extent of stabilization or destabilization caused by the modification.

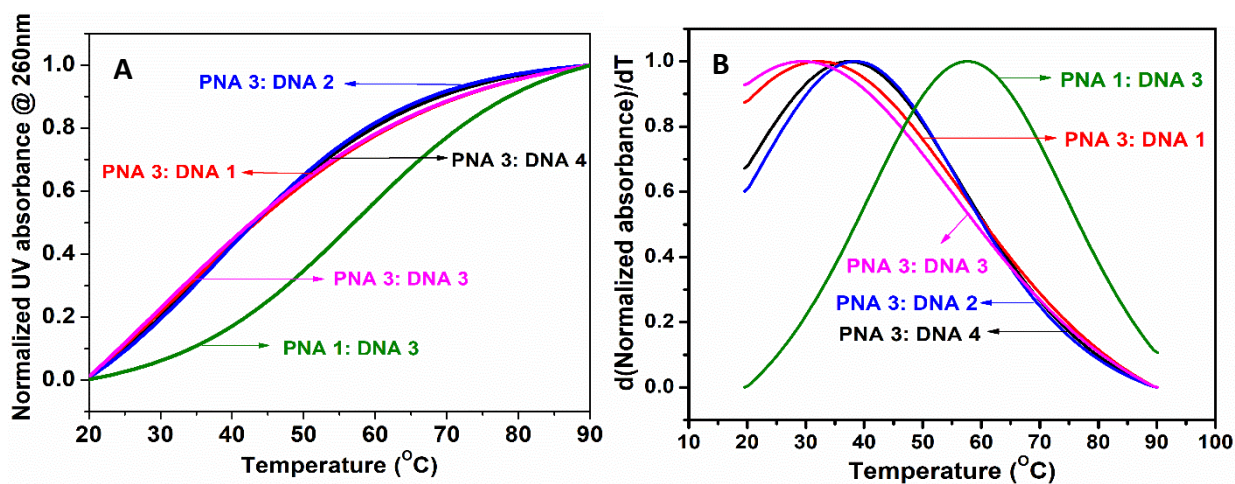
It is also a measure of the strength of the base-pair formed by the imidazole with the usual nucleobases (T, C, G and A) in the complementary strand (**DNA 1–4**) respectively.

Although all the four nucleobases in DNA, opposite to imidazole site on PNA seem to produce significant destabilization of the duplex, the degree of destabilization was largest with G ( $\Delta T_m = - 32.6$  °C) and the least destabilization was observed with C ( $\Delta T_m = - 19.1$  °C) at the complementary site. The degree of destabilization seemed to be affected by the nucleobase that was present in the complementary DNA, opposite the N1-imidazole modification site. The order of stability of the duplexes with respect to the complementary nucleobase was found to be  $C > A > T > G$ .

The observed decrease in the thermal stability shows that in a PNA:DNA duplex, the N1-imidazole is not tolerated on PNA and is a mismatch. This observation could be attributed to the inability of N1-imidazole to establish any strong interactions either through H-bonding or by stacking with any of the nucleobases on the other complementary strand. This is expected due to the presence of only one potential H-bond acceptor in the N1-imidazole, which also fails to render significant contributions through stacking between bases in the PNA:DNA duplex. The order of destabilization showed no apparent distinction between purines and pyrimidines.

#### **3.4.1b Duplexes of C4-Imidazolyl PNA 3 with cDNAs 1-4**

To examine the effect of C4-imidazole on the stability of corresponding PNA:DNA duplexes, **PNA 3** containing C4-imidazolyl modification at the middle of the 10 mer sequence was hybridized with antiparallel complementary DNAs (**DNA 1 - DNA 4**). The sigmoidal UV-melting transition was monitored (Figure 3.7A) and the corresponding  $T_m$  of the duplexes were measured from the respective first derivative curves (Figure 3.7B).



**Figure 3.7** UV-Melting curves of PNA:DNA duplexes (A) and their first derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.3** UV- $T_m$  values of duplexes of C4-imidazole modified PNA 3 with complementary DNA

Sr. No.	PNA : DNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 DNA 3	H-GTAGCTACTT-Lys NH <sub>2</sub> 3'-CATCGATGAA-5'	57.6	-
2	PNA 3 DNA 1	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> 3'-CATC <b>T</b> ATGAA-5'	33.4	-24.2
3	PNA 3 DNA 2	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> 3'-CATC <b>C</b> ATGAA-5'	38.8	-18.8
4	PNA 3 DNA 3	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> 3'-CATC <b>G</b> ATGAA-5'	29.5	-28.1
5	PNA 3 DNA 4	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> 3'-CATC <b>A</b> ATGAA-5'	37.4	-20.2

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

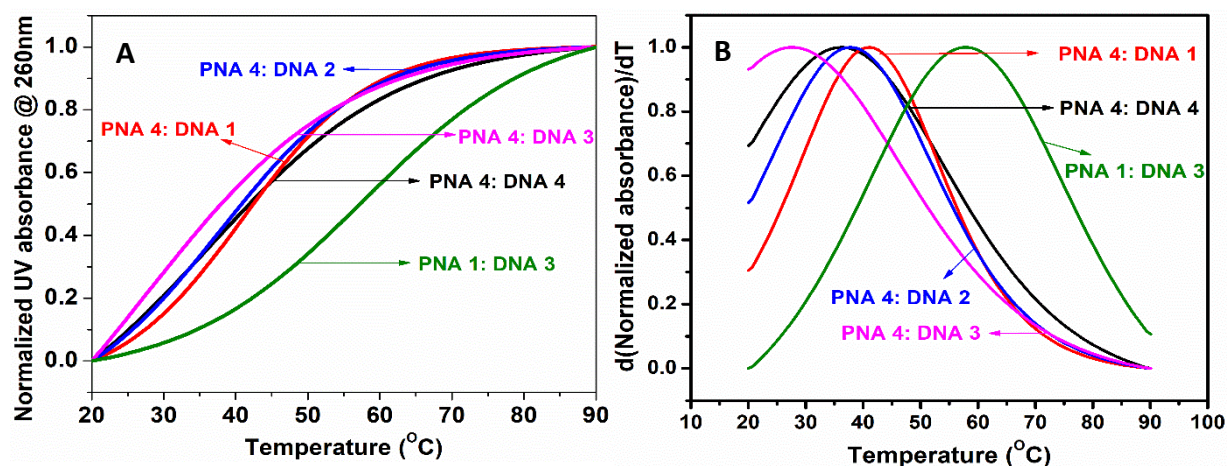
Similar to the observations with PNA oligomers containing N1-imidazole modification, the PNA oligomer with a single C4 imidazolyl modification also formed duplexes with DNA with significantly lower thermal stability than the unmodified PNA. With a  $T_m$  of 29.5 °C, the extent of destabilization was the highest with G at the complementary DNA site ( $\Delta T_m = -27.9$  °C). It was the least with cytosine, which showed a  $T_m$  of 38.8 °C ( $\Delta T_m = -18.6$  °C). Thus, an average  $T_m$  of 35.0 °C ( $\Delta T_m = -22.4$  °C) was observed with PNA:DNA duplexes formed by C4-imidazole modified PNA 3.

In comparison with N1-imidazolyl modified **PNA 2**, which showed an average  $T_m$  of 32.5 °C ( $\Delta T_m = - 25.2$  °C), the duplexes formed by C4-imidazolyl **PNA 3** were slightly more stable (relatively less destabilised). However, the order of duplex stability, C > A > T > G was found to be the same for both the N1-imidazole and C4-imidazolyl modifications.

Compared to N1-imidazolyl modification, the C4-imidazolyl modification produces only a marginal stabilization of the PNA:DNA duplex. However, when compared to the control **PNA 1**, the imidazolyl modifications were found to produce significant destabilization. Even though C4 imidazole moiety presents both H-bond donor and H-bond acceptor, it is incompatible to establish strong base-pairing via H-bonding with the natural nucleobases, due to lack of structural (geometrical) complementarity between the C4 imidazole moiety and any of the four nucleobases studied.

### 3.4.1c Duplexes of 4-NO<sub>2</sub>-Imidazolyl PNA 4 with cDNAs 1-4

In order to examine the base-pairing properties of 4-NO<sub>2</sub>-imidazole, **PNA 4** was hybridized with antiparallel complementary **DNA 1–4** and the corresponding duplexes were studied using the temperature dependent UV-absorption spectroscopy as was done for N1-imidazole and C4-imidazole modification (Figure 3.8). The absorbance at 260 nm showed the typical sigmoidal transition as the PNA:DNA underwent controlled and gradual temperature dependent denaturation to single strands.



**Figure 3.8** UV-Melting curves of PNA:DNA duplexes (A) and their first derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).



**Table 3.4** UV- $T_m$  values of duplexes of 4-NO<sub>2</sub>-imidazole modified PNA **4** with complementary DNA

Sr. No.	PNA: DNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA <b>1</b> DNA <b>3</b>	H-GTAGCTACTT-Lys NH <sub>2</sub> 3'-CATCGATGAA-5'	57.6	-
2	PNA <b>4</b> DNA <b>1</b>	H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> 3'-CATC <b>T</b> ATGAA-5'	41.3	-16.3
3	PNA <b>4</b> DNA <b>2</b>	H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> 3'-CATC <b>C</b> ATGAA-5'	37.7	-21.3
4	PNA <b>4</b> DNA <b>3</b>	H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> 3'-CATC <b>G</b> ATGAA-5'	27.8	-29.8
5	PNA <b>4</b> DNA <b>4</b>	H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> 3'-CATC <b>A</b> ATGAA-5'	36.3	-19.9

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

4-NO<sub>2</sub>-imidazole modified PNA **4** formed duplexes with complementary antiparallel DNA strands DNA **1–4** with an average  $T_m$  of 35.8 °C ( $\Delta T_m = -21.8$  °C). The magnitude of  $T_m$  change was similar to that observed with N1-imidazole (PNA **2**; average  $T_m = 32.8$ ) and C4-imidazole (PNA **3**) modifications. However, the order of stability (T > C > A > G) was found to be different. The lack of overlap between their preference for nucleobase on the opposite strand indicated that the modes of interactions between the NO<sub>2</sub>-imidazole and the nucleobases differ from those between the N1-imidazole or C4-imidazole. Notably, guanine still produced the greatest destabilization in this case as well.

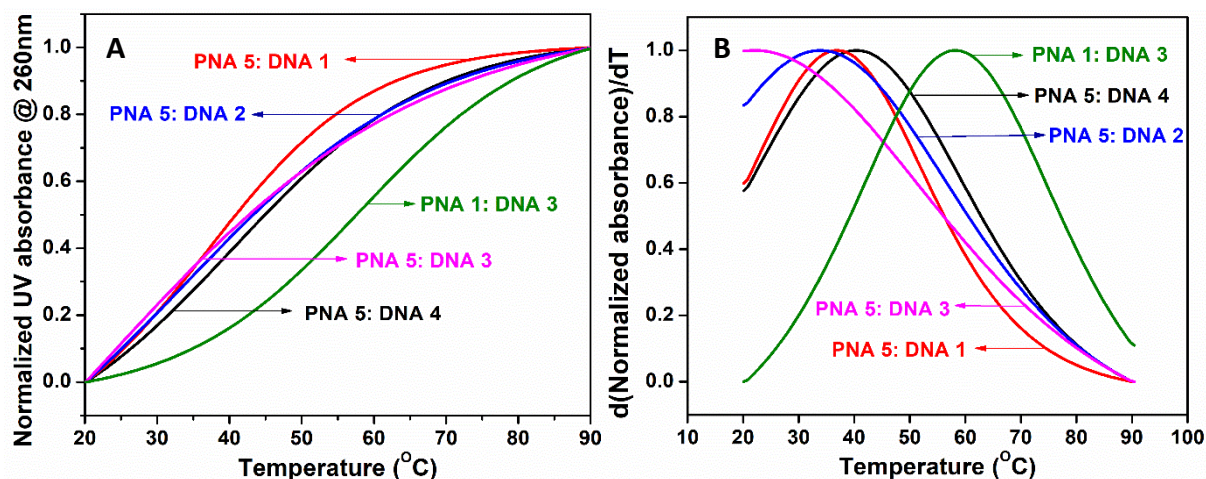
While 3-nitropyrrole is known to be a universal base in PNA: DNA duplexes, its analogue 4-nitroimidazole does not show such properties. On the contrary, it causes destabilization of the duplex.

#### 3.4.1d Duplexes of BzIm-Imidazolyl PNA **5** with cDNAs **1-4**

Benzimidazole, an analogue of purine with an imidazole ring fused to benzene ring was used to elucidate the role of stacking interactions in the base-pairing with complementary DNA strands in the present context. For this, PNA **5**, which contains



benzimidazole group as a substitute for nucleobase in the middle of the oligomer, was hybridized with antiparallel complementary DNA 1–4. The corresponding sigmoidal melting transitions monitored by absorbance at 260 nm and the respective first derivatives are shown below (Figure 3.9).



**Figure 3.9** UV-Melting curves of PNA : DNA duplexes (A) and their first derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.5** UV- $T_m$  values of duplexes of BzIm-imidazole modified PNA 5 with complementary DNA

Sr. No.	PNA : DNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 DNA 3	H-GTAGCTACTT-Lys NH <sub>2</sub> 3'-CATCGATGAA-5'	57.6	-
2	PNA 5 DNA 1	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> 3'-CATC <b>T</b> ATGAA-5'	37.2	-16.3
3	PNA 5 DNA 2	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> 3'-CATC <b>C</b> ATGAA-5'	33.6	-21.3
4	PNA 5 DNA 3	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> 3'-CATC <b>G</b> ATGAA-5'	22.2	-29.8
5	PNA 5 DNA 4	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> 3'-CATC <b>A</b> ATGAA-5'	40.7	-19.9

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

The control PNA: DNA duplex showed a UV-melting  $T_m$  of 57.6 °C (Table 3.5, entry 1). Similar to the other modifications studied in this chapter, benzimidazole

substitution also produced significant destabilization. The average  $T_m$  was found to be 33.4 °C ( $\Delta T_m = - 24.2$  °C). Compared to N1-imidazole, which has an average  $T_m$  of 32.8 °C, the benzimidazole substitution is not able to produce any significant increase in the thermal stability. This indicates that the factors that contributed to the destabilization of the duplex could not be countered by improved stacking between the bases in PNA:DNA double helix. This is also indicated by changes in the order of duplex stability. In this case, it was found to be  $A > T > C > G$ , which was different from the ones observed with other modifications. No distinct pattern was discernible from the trend and consistent with observations with other modifications, G in the opposite produced the highest destabilization of the duplex. Interestingly, in spite of the potential steric hindrance, the purine A, has the least destabilizing effect on the duplex.

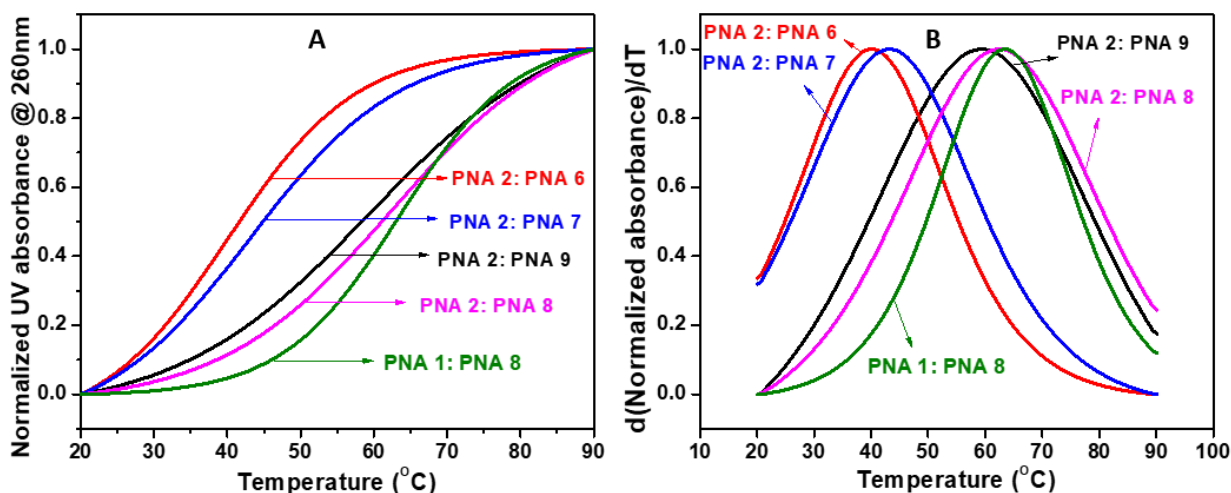
### **3.4.2 Thermal melting studies of antiparallel PNA:PNA hybrids**

It is known PNA:PNA duplexes are thermodynamically more stable than DNA:DNA and DNA:PNA duplexes. Therefore, the thermal stability of the hybrid duplexes can be studied by monitoring the changes in the absorbance at 260 nm during a heat-denaturation process. Comparison of the measured  $T_m$  values between the duplexes formed by modified PNAs and the unmodified ones provides a way to evaluate the effect of the modifications on the strength of base-pairs at the site of modification.

Herein, the modified PNAs **2-5** were hybridised separately with corresponding complementary PNAs **6-9**. These PNAs were modified with the unnatural nucleobases (N1-linked imidazole, C4-linked imidazole, 3-nitroimidazole, and benzimidazole) against these modifications to base pair with all natural nucleobases (A, G, C, and T). The thermal stability of each of these duplexes were studied in 10 mM phosphate buffer at pH 7.2 in the presence of a 10 mM NaCl. The  $T_m$  measured for the control pair of PNA:PNA hybrid duplex formed between unmodified **PNA 1** and **PNA 8** under the same conditions was compared to evaluate the stabilizing or destabilizing effect of the modification.

### 3.4.2a Duplexes of N1-Imidazolyl PNA 2 with cPNAs 6-9

In order to investigate the effect of N1-imidazole modified PNA monomer, from their corresponding duplexes, **PNA 2** was hybridized with antiparallel complementary PNAs (**PNA 6 - PNA 9**). The UV-melting transition of these duplexes are shown in (Figure 3.10A) and its derivative of the sigmoidal plots was generated for each of the duplexes to identify the respective melting temperature (Figure 3.10 B).



**Figure 3.10** UV-Melting curves of PNA:PNA duplexes (A) and their first derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.6** UV- $T_m$  values of duplexes of N1-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence	UV- $T_m$ (°C)	
		$T_m$	$\Delta T_m$
1	PNA 1 H-GTAGCTACTT-Lys NH <sub>2</sub> PNA 8 Lys NH <sub>2</sub> -CATCGATGAA-H	63.2	-
2	PNA 2 H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> PNA 6 Lys NH <sub>2</sub> -CATC T ATGAA-H	40.0	-23.2
3	PNA 2 H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> PNA 7 Lys NH <sub>2</sub> -CATC C ATGAA-H	43.4	-19.8
4	PNA 2 H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> PNA 8 Lys NH <sub>2</sub> -CATC G ATGAA-H	62.5	-0.7
5	PNA 2 H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> PNA 9 Lys NH <sub>2</sub> -CATC A ATGAA-H	59.7	-3.5

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

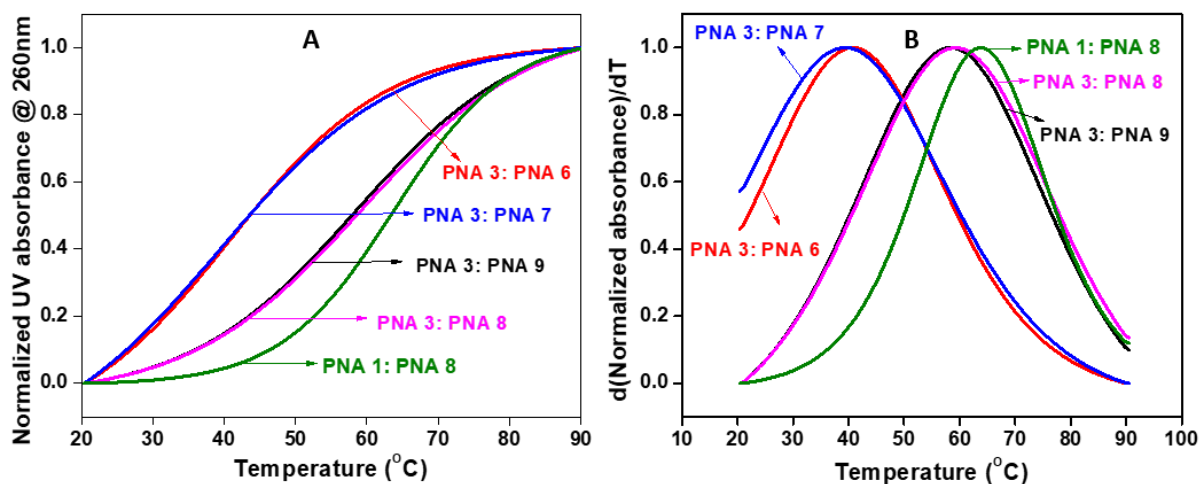
The unmodified control PNA:PNA duplexes showed a melting ( $T_m$ ) of 63.2 °C (Table 3.6, entry 1). In case of modified PNA:PNA duplexes, N1-imidazole modified **PNA 2** formed complex with antiparallel PNA strands **PNA 6-9**. The  $T_m$  values for N1-imidazole

(PNA 2) against G and A (PNA 8 and PNA 9) duplexes showed slightly less stabilization ( $\Delta T_m = -0.7$  °C and  $-3.5$  °C, Table 3.6, entry 4 and 5) than control PNA:PNA duplex whereas, duplexes with complementary T and C PNAs were found to have significantly lesser  $T_m$  than control duplex.

The duplexes derived from N1-imidazole modified PNA 2 with antiparallel PNA:PNA duplexes showed lower destabilisation against PNAs with G and A than PNAs with T and A. This suggests that N1- imidazole acts like a pyrimidine, stabilizing better against G and A than C and T in PNA:PNA duplexes.

### 3.4.2b Duplexes of C4-Imidazolyl PNA 3 with cPNAs 6-9

To examine the effect of PNA with C4-imidazole modification that has two hydrogen bond donors, PNA 3 was hybridized with antiparallel complementary PNAs (PNA 6 - PNA 9). The UV-melting transition of these duplexes are shown in (Figure 3.11A). The derivative of the sigmoidal plots was generated for each of the duplexes to identify the respective melting temperature (Figure 3.11B).



**Figure 3.11** UV-Melting curves of PNA: PNA duplexes (A) and their first derivative plots (B). (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.7** UV- $T_m$  values of duplexes of C4-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 PNA 8	H-GTAGCTACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATCGATGAA-H	63.2	-
2	PNA 3 PNA 6	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC T ATGAA-H	40.9	-22.3
3	PNA 3 PNA 7	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC C ATGAA-H	39.5	-23.7
4	PNA 3 PNA 8	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC G ATGAA-H	59.3	-3.9
5	PNA 3 PNA 9	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC A ATGAA-H	58.6	-4.6

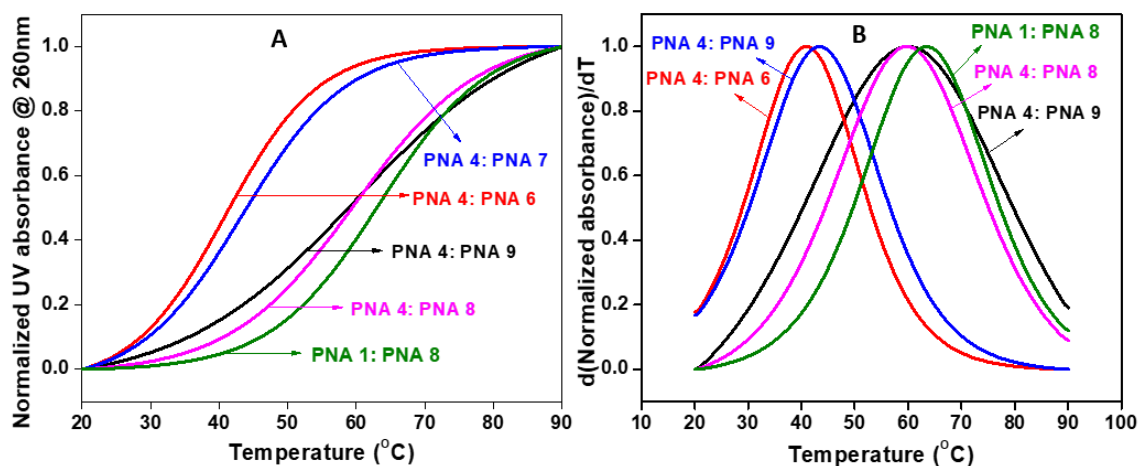
The  $T_m$  values are accurate to  $\pm 0.5$  °C.

The control PNA:PNA duplex was showed UV-melting 63.2 °C (Table 3.7, entry 1). The duplexes from C4-imidazole modified **PNA 3** with antiparallel complementary G and A PNAs (**PNA 8** and **9**) showed relatively less destabilisation ( $\Delta T_m = -3.9$  °C and  $-4.6$  °C Table 3.7, entry 4 and 5), compared to duplexes with T and C PNAs (**PNA 6** and **PNA 7**;  $\Delta T_m = -22.3$  °C and  $-23.7$  °C, Table 3.7, entry 2 and 3)

C4-imidazole modified PNA was stabilising only those PNAs with opposite G and A indicating that it also acts like pyrimidine nucleobases, stabilizing far better with opposite purines than pyrimidines.

### 3.4.2c Duplexes of 4-NO<sub>2</sub>-Imidazolyl PNA 4 with cPNAs 6-9

To examine the effect 4-NO<sub>2</sub>-imidazole modified **PNA 4** was individually hybridised with PNAs (**PNA 6** – **PNA 9**) in 10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl and the UV- $T_m$  data is shown in Figure 3.12 A along with  $T_m$  values found from derivative curve in Figure 3.12 B.



**Figure 3.12** UV-Melting curves of PNA: PNA duplexes (A) and their derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.8** UV- $T_m$  values of duplexes of 4-NO<sub>2</sub>-imidazole modified PNA with complementary PNA

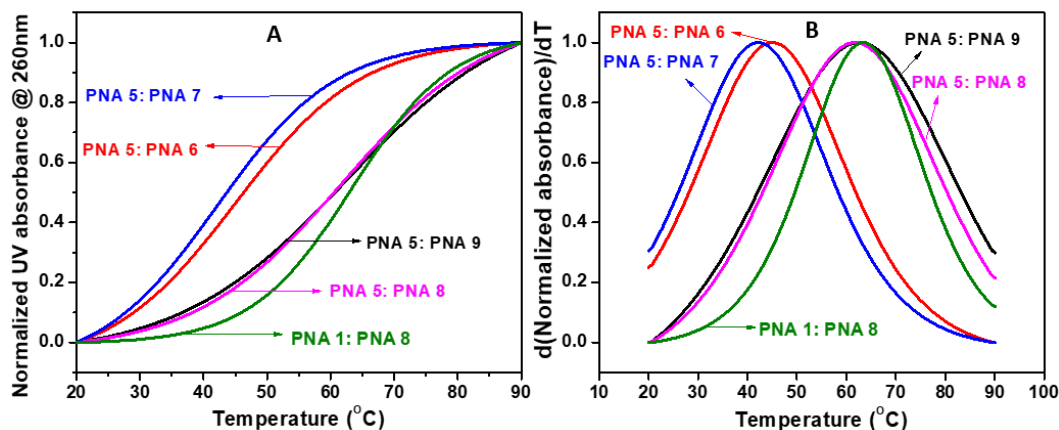
Sr. No.	PNA: PNA Duplex Sequence	UV- $T_m$ (°C)	
		$T_m$	$\Delta T_m$
1	PNA 1 H-GTAGCTACTT-Lys NH <sub>2</sub> PNA 8 Lys NH <sub>2</sub> -CATCGATGAA-H	63.2	-
2	PNA 4 H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> PNA 6 Lys NH <sub>2</sub> -CATC T ATGAA-H	40.6	-22.6
3	PNA 4 H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> PNA 7 Lys NH <sub>2</sub> -CATC C ATGAA-H	43.4	-19.8
4	PNA 4 H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> PNA 8 Lys NH <sub>2</sub> -CATC G ATGAA-H	59.7	-3.5
5	PNA 4 H-GTAGNO <sub>2</sub> ImTACTT-Lys NH <sub>2</sub> PNA 9 Lys NH <sub>2</sub> -CATC A ATGAA-H	59.0	-4.2

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

The PNA:PNA duplex derived from unmodified *aeg* PNA 1 showed a melting ( $T_m$ ) of 63.2 °C (Table 3.8, entry 1). The duplexes derived from 4-NO<sub>2</sub>-imidazole modified PNA 4 with complementary PNA 8 and PNA 9 showed lower destabilisation ( $\Delta T_m = -3.5$  °C and  $-4.2$  °C Table 3.8, entry 4 and 5) than duplexes with PNA 6 and PNA 7. Thus, 4-NO<sub>2</sub>-imidazole also behaves like a pyrimidine, stabilizing against purines G and A more than pyrimidines T and C in corresponding antiparallel PNA:PNA duplexes.

### 3.4.2d Duplexes of BzIm-Imidazolyl PNA 5 with cPNAs 6-9

To examine the effect on duplex stability of benzimidazole modified **PNA 5** with complementary **PNA 6-PNA 9**, the UV-melting curves for individual duplexes were recorded (Figure 3.13 A) and  $T_m$  extracted from derivative plots (Figure 3.13 B).



**FIGURE 3.13** UV-Melting curves of PNA: PNA duplexes (A) and their derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.9** UV- $T_m$  values of duplexes of BzIm-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 PNA 8	H-GTAGCTACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATCGATGAA-H	63.2	-
2	PNA 5 PNA 6	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC <b>T</b> ATGAA-H	44.8	-18.4
3	PNA 5 PNA 7	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC <b>C</b> ATGAA-H	41.9	-21.3
4	PNA 5 PNA 8	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC <b>G</b> ATGAA-H	61.7	-1.5
5	PNA 5 PNA 9	H-GTAG <i>BzIm</i> TACTT-Lys NH <sub>2</sub> Lys NH <sub>2</sub> -CATC <b>A</b> ATGAA-H	62.5	-0.7

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

The PNA:PNA duplex derived from unmodified *aeg* **PNA 1** showed a melting ( $T_m$ ) of 63.2 °C (Table 3.9, entry 1). The  $T_m$  of duplexes from benzimidazole modified **PNA 5** with antiparallel complementary PNAs (**PNA 6 - PNA 9**), showed negligible destabilisation ( $\Delta T_m = -1.5$  °C and  $-0.7$  °C Table 3.9, entry 4 and 5) compared with control

**PNA 1:PNA 8** duplex In comparison, duplexes of **PNA 5** with **PNA 6** and **PNA 7** were found to be highly destabilised.

In summary, antiparallel duplexes of all imidazole modified PNAs with complementary PNAs with purines G and A are more stable than duplexes of PNAs with pyrimidines C and T. To further check the compatibility, benzimidazole that mimics the purine ring was used in PNA modification. But it also exhibited similar preference pattern for PNAs with G and A than for PNAs with C and T. Benzimidazole may have additional stacking interaction with G and A than with T and C leading to lower destabilisation of their PNA:PNA duplexes.

### **3.4.3 Thermal melting studies of parallel PNA: PNA hybrids**

It is known antiparallel PNA:PNA duplexes are thermodynamically more stable than corresponding DNA:DNA and DNA:PNA and parallel PNA:PNA duplexes are less stable than antiparallel PNA:PNA duplexes. Therefore, to check the relative thermal stabilities of various duplexes, UV-melting studies at 260 nm were carried out with parallel duplexes.

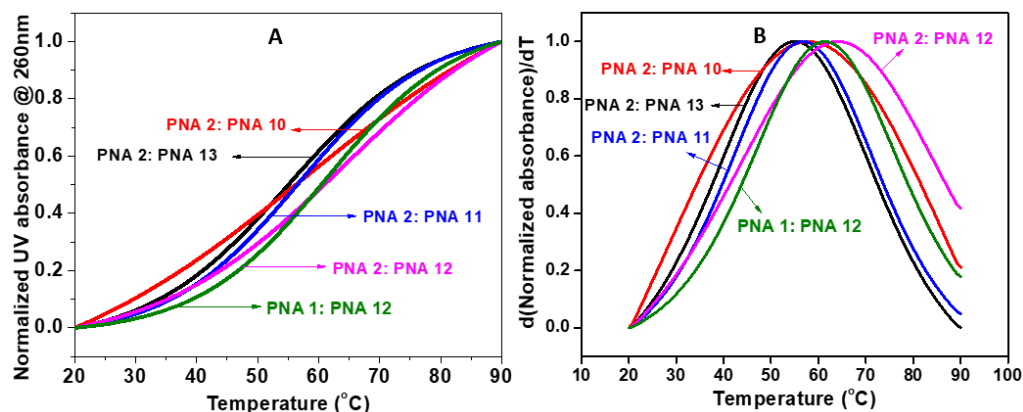
The modified PNAs **2-5** were hybridised individually with the corresponding complementary PNAs **10-13** designed to bind in antiparallel manner. The PNAs were modified with the unnatural nucleobases (N1-linked imidazole, C4-linked imidazole, 3-nitroimidazole, and benzimidazole) held opposite each of the natural nucleobases (A, G, C, and T). As before, the thermal stability of each of these duplexes were studied in 10 mM phosphate buffer at pH 7.2 in the presence of 10 mM NaCl. The  $T_m$  measured for the control parallel PNA:PNA duplex formed between unmodified **PNA 1** and **PNA 12** under the same conditions was compared to evaluate the stabilizing or destabilizing effect of the modification.

#### **3.4.3a Parallel duplexes of N1-Imidazolyl PNA 2**

In order to investigate the effect of N1-imidazole modified PNA on its corresponding parallel duplexes, **PNA 2** was hybridized with parallel complementary PNAs (**PNA 10 - PNA 13**). The UV-melting transition of these duplexes are shown in



(Figure 3.14A). The derivative of the sigmoidal plots was generated for each of the duplexes to identify the respective melting temperature (Figure 3.14B).



**Figure 3.14** UV-Melting curves of PNA: PNA duplexes (A) and their derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.10** UV- $T_m$  values of duplexes of N1-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 PNA 12	H-GTAGCTACTT-Lys NH <sub>2</sub> H-CATCGATGAA-Lys NH <sub>2</sub>	61.7	-
2	PNA 2 PNA 10	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> H-CATC T ATGAA-Lys NH <sub>2</sub>	58.2	-3.5
3	PNA 2 PNA 11	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> H-CATC C ATGAA-Lys NH <sub>2</sub>	56.7	-5.0
4	PNA 2 PNA 12	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> H-CATC G ATGAA-Lys NH <sub>2</sub>	63.8	+2.1
5	PNA 2 PNA 13	H-GTAGIm(N1)TACTT-Lys NH <sub>2</sub> H-CATC A ATGAA-Lys NH <sub>2</sub>	55.3	-6.4

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

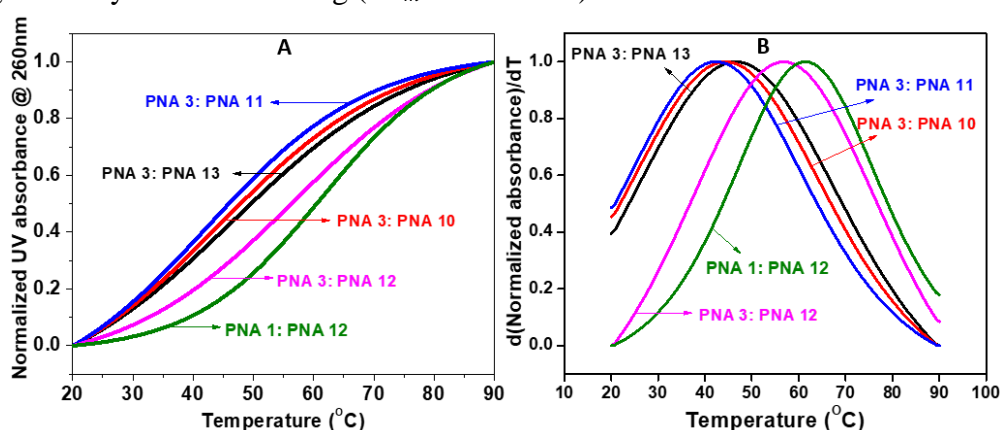
The control PNA:PNA parallel duplex showed melting ( $T_m$ ) of 61.7 °C (Table 3.10, entry 1). The  $T_m$  of duplexes derived from benzimidazole modified PNA 2 with parallel complementary PNAs (PNA 10 - PNA 13) showed marginal stabilisation ( $\Delta T_m = + 2.1$  °C Table 3.9, entry 4) in case of PNA 12 (G), while duplexes with PNA 10 (T), PNA 11 (C) and PNA 13 (A) were found to be destabilised by 3.5 – 6.5 °C

Thus N1-imidazole modified parallel PNA:PNA duplexes showed marginal stabilisation of duplex with parallel PNA (G) more than that with other natural

nucleobases. Since N1-imidazole cannot form H-bonding, it is possible that this arises from a different stacking interaction of N1-imidazole with G in parallel duplexes, better than that with other natural nucleobases (A, C and T).

### 3.4.3b Parallel duplexes of C4-Imidazolyl PNA 5

To examine the effect of C4-imidazole modified **PNA 5** on its duplexes with parallel complementary PNAs (**PNA 6-PNA 9**), UV-melting experiments (Figure 3.15 A) were done on corresponding duplexes and  $T_m$  extracted from first derivative plots. (Figure 3.15 B). The control PNA:PNA duplexes showed a melting ( $T_m$ ) of 61.7 °C (Table 3.11, entry 1). The C4-imidazole modified PNA:PNA parallel duplexes, showed lower destabilisation ( $\Delta T_m = -4.9$  °C Table 3.11, entry 4) in case of **PNA 12** (G) compared to control PNA:PNA duplex, while parallel duplexes **PNA 10** (T), **PNA 11** (C) and **PNA 13** (A) were found to be significantly more destabilising ( $\Delta T_m = > 14.0$  °C).



**Figure 3.15** UV-Melting curves of PNA:PNA duplexes (A) and their derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.11** UV- $T_m$  values of duplexes of C4-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 PNA 12	H-GTAGCTACTT-Lys NH <sub>2</sub> H-CATCGATGAA-Lys NH <sub>2</sub>	61.7	-
2	PNA 3 PNA 10	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> H-CATC T ATGAA-Lys NH <sub>2</sub>	44.1	-17.6
3	PNA 3 PNA 11	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> H-CATC C ATGAA-Lys NH <sub>2</sub>	42.7	-19
4	PNA 3 PNA 12	H-GTAGIm(C4)TACTT-Lys NH <sub>2</sub> H-CATC G ATGAA-Lys NH <sub>2</sub>	56.8	-4.9

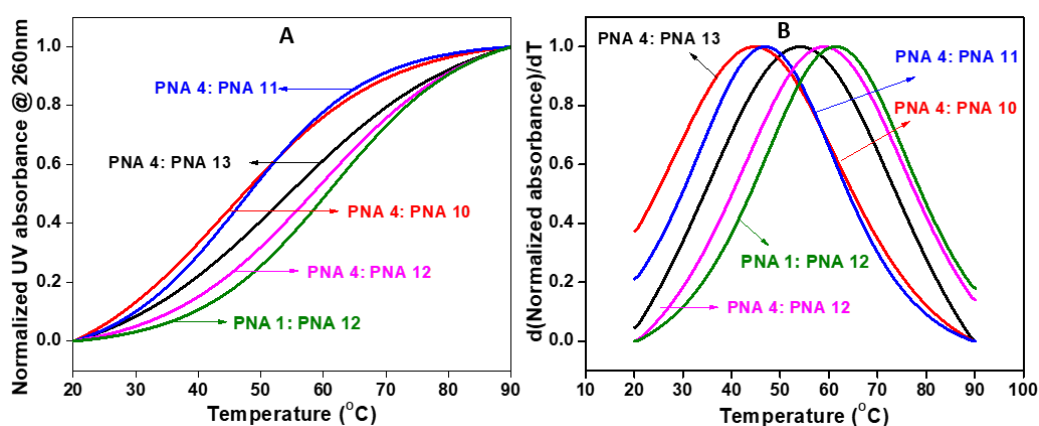
5	PNA 3 PNA 13	H-GTAG <i>Im(C4)</i> TACTT-Lys NH <sub>2</sub> H-CATC <b>A</b> ATGAA-Lys NH <sub>2</sub>	46.9	-14.8
---	-----------------	---	------	-------

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

Thus parallel PNA:PNA duplexes derived from C4-imidazole PNA 3 showed slightly more stabilisation with complementary PNA 12 (G) than that with other nucleobases (T, C and A), similar to the pattern observed for other modified PNAs.

### 3.4.3c Parallel duplexes of 4-NO<sub>2</sub>-Imidazoly PNA 4

To investigate the effect of 4-NO<sub>2</sub>-imidazole PNA 4 on its parallel duplexes with complementary PNAs 10-13, UV-melting studies of these duplexes were done and the results are shown in Figure 3.16A. The derivative of the sigmoidal plots generated for each of the duplexes to identify the respective melting temperature are depicted in Figure 3.16B. The control PNA: PNA duplex  $T_m$  was 61.7 °C and NO<sub>2</sub> modified PNA 4 was hybridised with complementary parallel PNAs (PNA 10- PNA 13). Among all combination PNA 12 showed very less destabilisation ( $\Delta T_m = - 2.9$  °C Table 3.12, entry 4) with PNA 4 whereas, all PNA 10, PNA 11 and PNA 13 showed higher destabilisation.



**Figure 3.16** UV-Melting curves of PNA: PNA duplexes (A) and their derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

**Table 3.12** UV- $T_m$  values of duplexes of 4-NO<sub>2</sub>-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence		UV- $T_m$ (°C)	
			$T_m$	$\Delta T_m$
1	PNA 1 PNA 12	H-GTAGCTACTT-Lys NH <sub>2</sub> H-CATCCGATGAA-Lys NH <sub>2</sub>	61.7	-

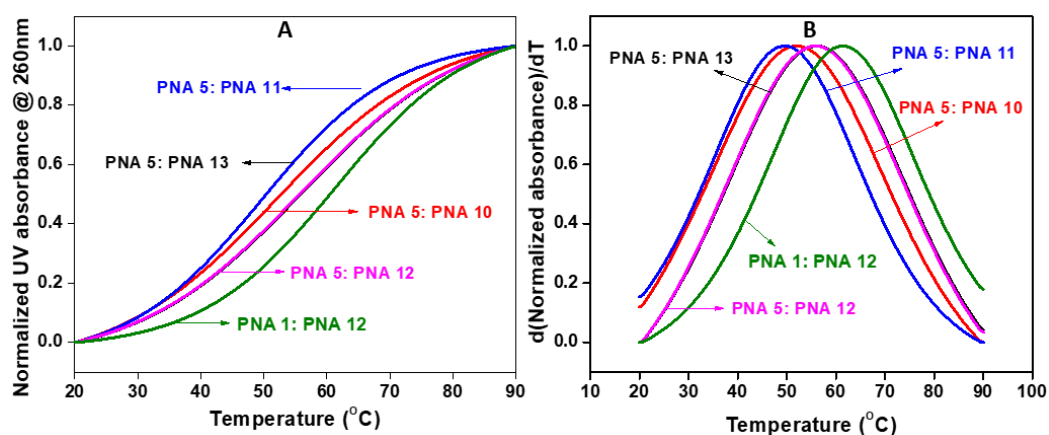
2	PNA 4 PNA 10	H-GTAG $NO_2Im$ TACTT-Lys NH <sub>2</sub> H-CATC <b>T</b> ATGAA-Lys NH <sub>2</sub>	44.8	-16.9
3	PNA 4 PNA 11	H-GTAG $NO_2Im$ TACTT-Lys NH <sub>2</sub> H-CATC <b>C</b> ATGAA-Lys NH <sub>2</sub>	47.0	-14.7
4	PNA 4 PNA 12	H-GTAG $NO_2Im$ TACTT-Lys NH <sub>2</sub> H-CATC <b>G</b> ATGAA-Lys NH <sub>2</sub>	59.0	-2.7
5	PNA 4 PNA 13	H-GTAG $NO_2Im$ TACTT-Lys NH <sub>2</sub> H-CATC <b>A</b> ATGAA-Lys NH <sub>2</sub>	54.0	-7.7

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

Thus 4- $NO_2$ -imidazole also stabilises PNA (G) in parallel PNA:PNA duplexes compared to other duplexes perhaps through a better stacking interaction with G.

### 3.4.3 d Parallel duplexes of Bzim-Imidazolyl PNA 5

To examine the effect of Bzim-imidazole modified PNA 5 on derived parallel duplexes, PNA 5 was hybridised with parallel PNA strands PNA 10-13. The UV-melting transition of these duplexes are shown in (Figure 3.17A). The derivative of the sigmoidal plots was generated for each of the duplexes to identify the respective melting temperature (Figure 3.17B). The  $T_m$  of control PNA 1:PNA 12 parallel duplex was 61.7 °C and Bzim modified PNA 5 was hybridised with complementary parallel PNAs (PNA 10- PNA 13). Among all duplex combinations purine PNA 12 (G) and PNA 13 (A) showed very less destabilisation ( $\Delta T_m = - 5.6$  °C and  $- 5.7$  °C Table 3.13, entry 4 and 5) whereas, the pyrimidine PNA 10 (T) and PNA 11 (C) showed higher destabilisation ( $\Delta T_m > 10$  °C). Thus Bzim-imidazole stabilises G and A PNA:PNA parallel duplexes better than C and T PNA:PNA parallel duplexes



**Figure 3.17.** UV-Melting curves of PNA: PNA duplexes (A) and their derivative plots (B) (10 mM sodium phosphate buffer, pH = 7.4, 10 mM NaCl).

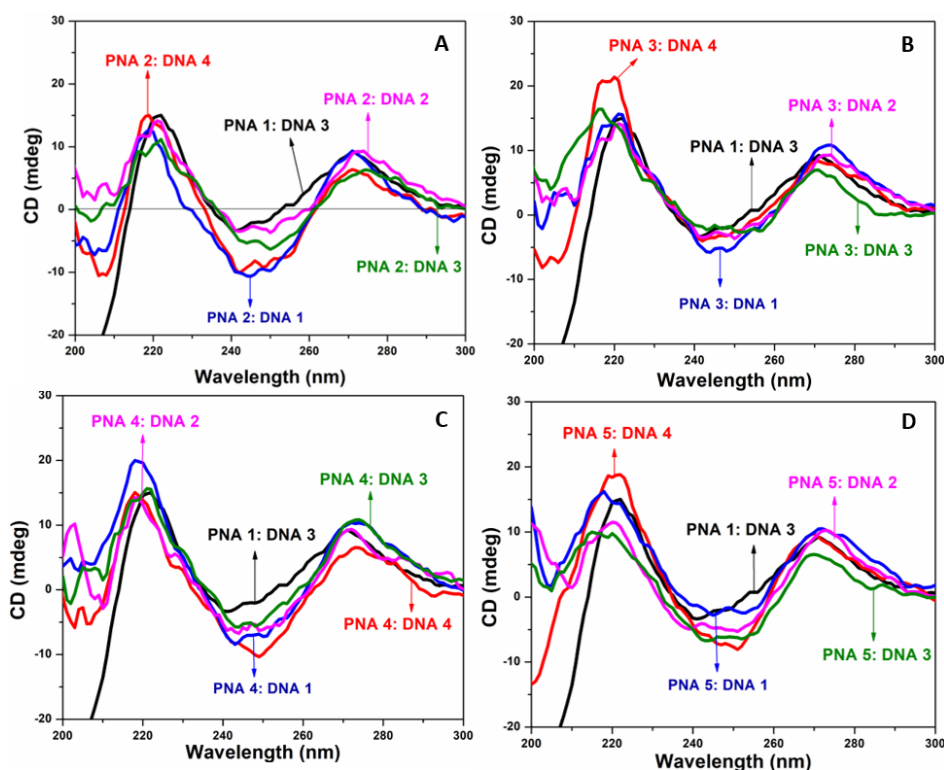
**Table 3.13** UV- $T_m$  values of duplexes of BzIm-imidazole modified PNA with complementary PNA

Sr. No.	PNA: PNA Duplex Sequence	UV- $T_m$ (°C)	
		$T_m$	$\Delta T_m$
1	PNA 1 PNA 12 H-GTAGCTACTT-Lys NH <sub>2</sub> H-CATCGATGAA-Lys NH <sub>2</sub>	61.7	-
2	PNA 3 PNA 10 H-GTAGBzImTACTT-Lys NH <sub>2</sub> H-CATC T ATGAA-Lys NH <sub>2</sub>	51.8	-10
3	PNA 3 PNA 11 H-GTAGBzImTACTT-Lys NH <sub>2</sub> H-CATC C ATGAA-Lys NH <sub>2</sub>	49.7	-12.1
4	PNA 3 PNA 12 H-GTAGBzImTACTT-Lys NH <sub>2</sub> H-CATC G ATGAA-Lys NH <sub>2</sub>	56.2	-5.6
5	PNA 3 PNA 13 H-GTAGBzImTACTT-Lys NH <sub>2</sub> H-CATC A ATGAA-Lys NH <sub>2</sub>	56.1	-5.7

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

#### 3.4.4 CD studies of Anti Parallel duplexes of PNA 2, PNA 3, PNA 4, PNA 5 with cDNA

The effect of various imidazole modified substitutions on the conformation of PNA:DNA duplexes was studied by CD spectroscopy. All modified PNA:DNA duplexes show



**Figure 3.18** CD spectra of PNAs 1-5 duplexed with cDNAs 1-4 (Phosphate buffer: 10 mM sodium phosphate, pH=7.2, 10 mM NaCl).

Two positive bands at 220 nm, 270 nm and negative band intensity between 240 nm to 250 nm. These are characteristic bands of standard PNA:DNA duplexes. N-1 imidazole, C-4 imidazole, 4-nitroimidazole and benzimidazole modified PNAs (**PNA 2 - PNA 5**) formed duplexes with complementary DNAs (**DNA 1- DNA 4**) as shown in (Figure 3.18 A, B, C and D). All modified and unmodified PNAs formed duplexes with cDNA, without much conformational distortion as they all have the same CD profile

### 3.5 Summary

Duplexes from all imidazolyl PNAs with complementary antiparallel DNA showed significant destabilisation with all natural nucleobases in cDNA. These modifications were not fitting for base pairing with the four natural nucleobases. In case of antiparallel PNA:PNA duplexes, imidazolyl PNA showed discrimination with purines (A,G) and pyrimidines (C,T). These imidazolyl PNAs showed more stability with complementary G and A PNAs whereas, T and C PNAs showed significantly less stable duplexes. In case of parallel PNA:PNA duplexes, all imidazolyl PNA showed G to be more compatible against

imidazolemodifications and stabilise corresponding duplexes better than other three nucleobases A, C and T.

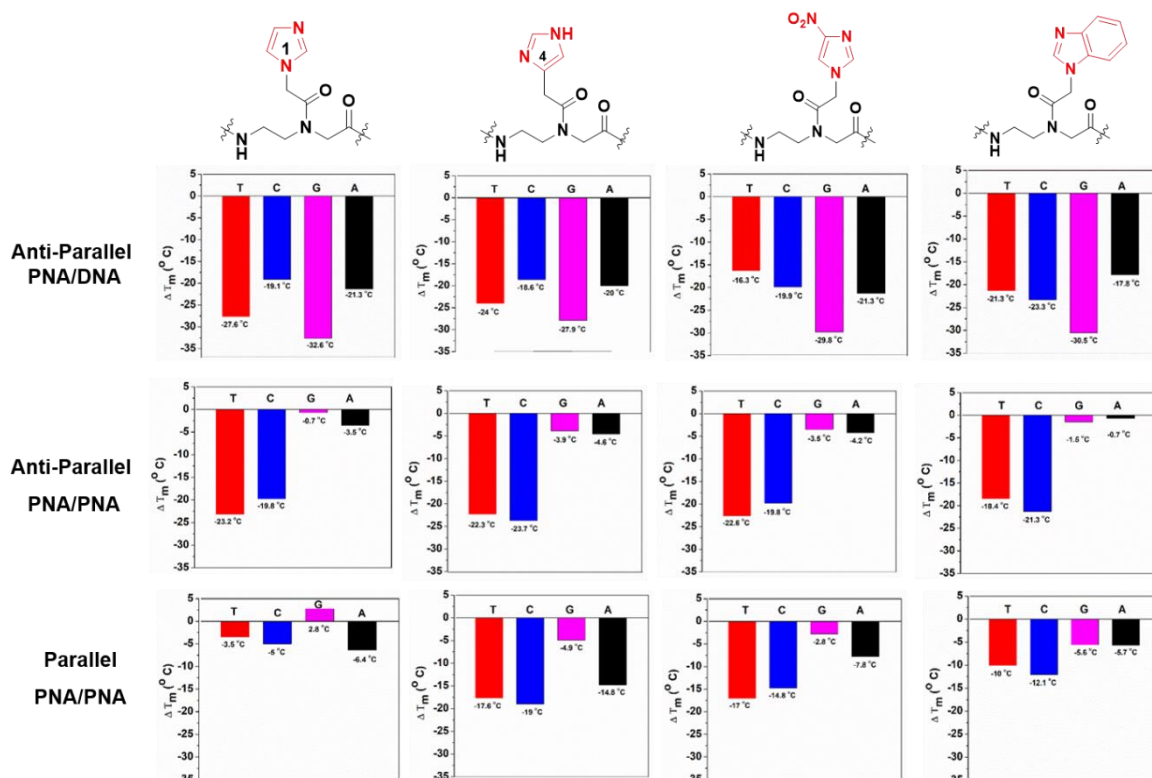


Figure 3.19 Summary of Chapter

## 3.6 Experimental procedures

### 3.6.1 Chemicals

The PNA oligomers were synthesized by solid phase protocol manually and under microwave conditions using Boc-Strategy as described in Chapter 2. The DNA and RNA oligonucleotides were obtained commercially from IDT (Integrated DNA Technologies, USA). All other chemicals were of analytical grade.

### 3.6.2 UV- $T_m$ measurements

UV-melting experiments were carried out on Varian Cary 300 UV spectrophotometer equipped with a Peltier temperature programmer and Julabo water circulator. The samples for  $T_m$  measurements were prepared at 2.5  $\mu\text{M}$  concentration of each PNA and DNA (500  $\mu\text{L}$  of 10 mM phosphate buffer, pH 7.2 and 10 mM NaCl). The PNA and cDNA or cPNA were mixed together in stoichiometric amounts of 1:1 and

annealed at 90 °C for 5 min and cooled to room temperature slowly. The samples were transferred to quartz cell, sealed with Teflon stopper and the optical density (OD) was recorded at 260 nm with a rate of 1.0 °C/min temperature increment from 20 °C to 90 °C. The normalized absorbance at 260 nm was plotted as a function of the temperature and the  $T_m$  was determined from the first derivative plots with respect to temperature and is accurate to  $\pm 0.5$  °C. The data were processed using Microcal Origin 8.5. The concentration of DNA and PNA were calculated with the help of extinction coefficients of nucleobases ( $A = 15.4 \times 10^3$ ,  $T = 8.8 \times 10^3$ ,  $C = 7.3 \times 10^3$  and  $G = 11.7 \times 10^3$ )  $\text{m}^{-1} \text{cm}^{-1}$ .

### 3.6.3 Circular Dichroism

CD spectra were recorded on JASCO J-715 spectro-polarimeter. The CD spectra of the PNA:DNA complexes and the relevant single strands were recorded using 10  $\mu\text{M}$  duplexes in 10 mM sodium phosphate buffer, 10 mM NaCl at pH 7.2. The temperature of the circulating water was kept below the melting temperature of the PNA:DNA complexes, i.e., at 20 °C. The CD spectra of PNA:DNA duplexes were recorded by addition of 3 scans from 300 to 190 nm, with a resolution of 0.1 nm, bandwidth of 1.0 nm, sensitivity of 2 mdeg, response of 2 sec and a scan speed of 50 nm/min.

## 3.7 References

1. Rich, A. Problems of evolution and biochemical information transfer. In *Horizons Biochemistry*. Kasha, M. P. B., Ed.; Academic Press: 1962; pp 103\_126.
2. Wu, Y.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. Efforts toward expansion of the genetic alphabet: Optimization of interbase hydrophobic interactions. *J. Am. Chem. Soc.* **2000**, *122*, 7621–7632.
3. Rana, V. S.; Ganesh, K. N. Recognition of 5-aminouracil (U#) in the central strand of a DNA triplex: orientation selective binding of different third strand bases. *Nucleic Acids Res.* **2000**, *28*, 1162–1169.
4. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **1990**, *343*, 33–37.



5. Morales, J. C.; Kool, E. T. Efficient replication between non-hydrogen-bonded nucleoside shape analogs. *Nat. Struct. Biol.* **1998**, *5*, 950–954.
6. Marcus, B.; Yiqin, W.; Anthony, K. O.; Distin, L. M.; Peter, G. S.; Floyd, E. R. Universal bases for hybridization replication and chain termination. *Nucleic Acids Research* **2000**, *28*, 2911–2914
7. McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly stable, selfpairing hydrophobic base. *J. Am. Chem. Soc.* **1999**, *121*, 11585–11586.
8. Malyshev, D. A.; Seo, Y. J.; Ordoukhanian, P.; Romesberg, F. E. PCR with an expanded genetic alphabet. *J. Am. Chem. Soc.* **2009**, *131*, 14620–14621.
9. Malyshev, D. A.; Pfaff, D. A.; Ippoliti, S. I.; Hwang, G. T.; Dwyer, T. J.; Romesberg, F. E. Solution structure, mechanism of replication, and optimization of an unnatural base pair. *Chemistry* **2010**, *16*, 12650–12659.
10. Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I. An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. *Nucleic Acids Res.* **2009**, *37*, e14.
11. Ohtsuki, T.; Kimoto, M.; Ishikawa, M.; Mitsui, T.; Hirao, I.; Yokoyama, S. Unnatural base pairs for specific transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4922–4925.
12. Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokoyama, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Nishikawa, K.; Yokoyama, S. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **2002**, *20*, 177–182.
13. Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. An unnatural hydrophobic base pair with shape complementarity between pyrrole-2-carbaldehyde and 9-methylimidazo[4,5-b]pyridine. *J. Am. Chem. Soc.* **2003**, *125*, 5298–5307.
14. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods* **2006**, *3*, 729–735.
15. Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. An efficient unnatural base pair for PCR amplification. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555.
16. Ohtsuka, E.; Matsuki, S.; Ikehara, M.; Takahashi, Y.; Matsubara, K. An Alternative Approach to Deoxyoligonucleotides as Hybridization Probes by Insertion of Deoxyinosine at Ambiguous Codon Positions. *J. Biol. Chem.* **1985**, *260*, 2605–2608.

17. Amosova, O.; George, J.; Fresco, J. R. Effect of the 1-(2'-deoxy-β-D-ribofuranosyl)-3-nitropyrrole residue on the stability of DNA duplexes and triplexes. *Nucleic Acids Res.* **1997**, *25*, 1930-1934.
18. Nichols, R.; Andrews, P. C.; Zhang, P.; Bergstrom, D. E. A universal nucleoside for use at ambiguous sites in DNA primers. *Nature* **1994**, *369*, 492-493.
19. Bergstrom, D. E.; Zhang, P.; Toma, P. H.; Andrews, P. C.; Nichols, R. Synthesis, Structure, and Deoxyribonucleic Acid Sequencing with a Universal Nucleoside: 1-(2'-Deoxy-β-D-ribofuranosyl)-3-nitropyrrole. *J. Am. Chem. Soc.* **1995**, *117*, 1201-1209.
20. Loakes, D.; Brown, D. M. 5-Nitroindole as an universal base analogue. *Nucleic Acids Res.* **1994**, *22*, 4039-4043.
21. Loakes, D.; Hill, F.; Linde, S.; Brown, D. M. Nitroindoles as Universal Bases *Nucleosides Nucleotides* **1995**, *14*, 1001-1003.
22. Loakes, D.; Brown, D. M.; Linde, S.; Hill, F. 3-Nitropyrrole and 5-nitroindole as universal bases in primers for DNA sequencing and PCR *Nucleic Acids Res.* **1995**, *23*, 2361-2366.
23. Loakes, D.; Hill, F.; Brown, D. M.; Salisbury, S. A. Stability and Structure of DNA Oligonucleotides Containing Non-specific Base Analogues. *J. Mol. Biol.* **1997**, *270*, 426-435.
24. Johnson, W. T.; Zhang, P.; Bergstrom, D. E. The synthesis and stability of oligodeoxyribonucleotides containing the deoxyadenosine mimic 1-(2'- deoxy-β-D-ribofuranosyl)imidazole-4-carboxamide. *Nucleic Acids Res.* **1997**, *25*, 559-567.
25. Zhang, P.; Johnson, W. T.; Klewer, D.; Paul, N.; Hoops, G.; Davisson, V. J.; Bergstrom, D. E. Exploratory studies onazole carboxamides as nucleobase analogs: thermal denaturation studies on oligodeoxyribonucleotide duplexes containing pyrrole-3-carboxamide. *Nucleic Acids Res.* **1998**, *26*, 2208-2215.
26. Pochet, S.; Dugue, L. Imidazole-4-Carboxamide and 1,2,4-Triazole-3-Carboxamide Deoxynucleotides as Simplified DNA Building Blocks with Ambiguous Pairing Capacity. *Nucleosides Nucleotides* **1998**, *17*, 2003-2009
27. Bissler, J. J. Triplex DNA and human disease. *Front. Biosci.* **2007**, *12*, 4536-4546.
28. Hensel, S.; Megger, N.; Schweizer, K.; Müller, J. Second generation silver(I)-mediated imidazole base pairs. *Beilstein J. Org. Chem.* **2014**, *10*, 2139-2144.

## **Chapter 4**

# **Metal Complexation, Biophysical and RNA Cleavage Studies by Imidazole-PNA conjugates**

## 4.1 Introduction

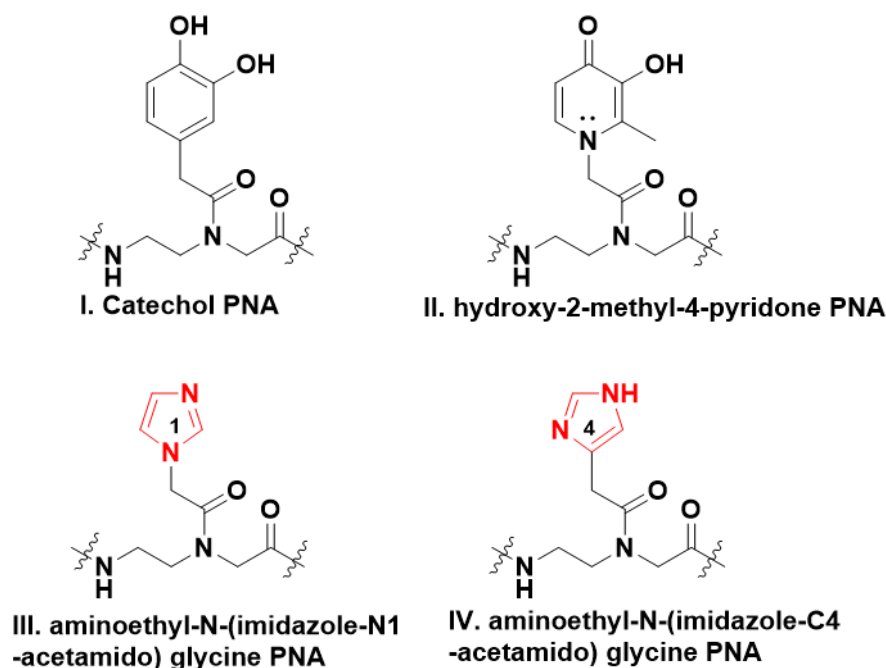
During the past decades there is a great interest in trying to mimic ribonucleases. Towards this end, researchers have used several sequence recognizing natural / synthetic chemical moieties functionalized with RNA cleaving moieties such as metal complexes. Most of these studies employ lanthanides,  $\text{Cu}^{2+}$ , or  $\text{Zn}^{2+}$  as metal ion in its catalytic complex.<sup>1-5</sup> Komiyama *et.al.*<sup>4-7</sup> developed RNA cleaving systems based on lanthanide ions.

The PNA is a designed analogue of DNA which structurally permits the incorporation of distinct, high affinity metal binding sites either on the backbone or at N- / C- termini. In this context, the metal binding moieties can also be incorporated at the base pairing site. Upon co-ordination with metal, the overall charge of the otherwise neutral PNA molecule is determined by the metal ion charge of the conjugate. When DNA is employed for constructing such artificial nucleases and metal-based supramolecular nanostructures, the negative charge of the phosphate backbone neutralises the metal positive charge, decreasing the net charge on the backbone. The various types of metal binding ligands, such as pyridine, ferrocene, 8-hydroxyquinoline, ter-pyridine, tetrazole and bipyridine have been attached to *aeg*-PNA.<sup>8-12</sup> The ligand-modified PNAs have been used as ribonucleases mimics; for instance, dimethyl phenanthroline ligand attached PNA has shown site specific RNA cleavage in the presence of  $\text{Cu}^{2+}$  ion.<sup>13</sup> Furthermore, tris-2-benzimidazole modified PNA showed site-specific RNA cleavage without metal ion.<sup>14</sup> Other applications include PNA-derived biosensors with capability of distinguishing closely related sequences as compared to DNA under less restrictive reaction conditions.<sup>15,16</sup> The PNA recognition layer was used for hosting free diffusing redox mediators,<sup>17</sup> redox-active intercalators or minor-groove binders<sup>18-21</sup> or as covalently bound redox labels.<sup>22,23</sup>

The preceding chapters discussed the synthesis of rationally designed imidazolyl PNA monomers, and imidazole modified PNA oligomers by solid phase synthesis, the thermal stability of derived hybrids with complementary DNA, RNA and PNA. In this chapter, hydrolytic activity of RNA induced by the modified PNA oligomers containing imidazole moieties at specific site in the sequence and at rationally designed position on the backbone or on nucleobase has been investigated using HPLC in the presence and in the absence of different metal ions.

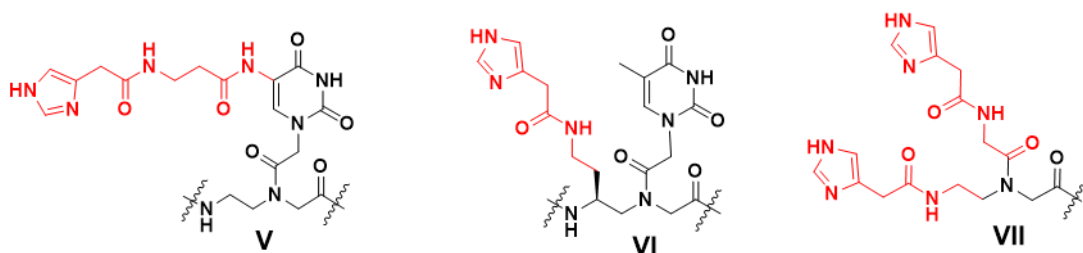
## 4.2 Rational and objective of present work

The present chapter describes the biophysical properties of imidazole modified PNAs, their metal complexes and RNA cleavage induced by them. Previous work from the laboratory involved the use of catechol and 3-hydroxy 2-methyl 4-pyridone on PNA backbone as metal binding ligands.<sup>24</sup> The catechol linked PNA monomer and oligomer (I) showed poor binding affinity to  $\text{Cu}^{2+}$  and  $\text{Ni}^{2+}$  and PNA monomer with 3-hydroxy 2-methyl 4-pyridone (II) exhibited strong affinity with rare earth metal ions ( $\text{Eu}^{3+}$ ,  $\text{Tb}^{3+}$ ,  $\text{Ho}^{3+}$  and  $\text{La}^{3+}$ ). In view of the known metal complexation of imidazoles, PNA conjugated to imidazole on backbone through acetyl linker at N1 (III) and C4- (IV) sites were chosen for metal complexation studies.



**Figure 4.1** Structure of catechol PNA (I), 3-Hydroxy 2-methyl 4-pyridone PNA (II), aminoethyl-N-(imidazole-N1-acetamido) glycine PNA and (III) aminoethyl-N-(imidazole-C4-acetamido) glycine PNA (IV)

The biophysical studies have been carried out on the designed imidazole modified PNAs (Figure 4.1) in order to investigate the effect of incorporated imidazole groups on the hydrolysis of target cRNA. HPLC has been used extensively for RNA hydrolysis study.



**Figure 4.2** Structure of imidazole modified PNAs for RNA hydrolysis

The specific objectives of this chapter are

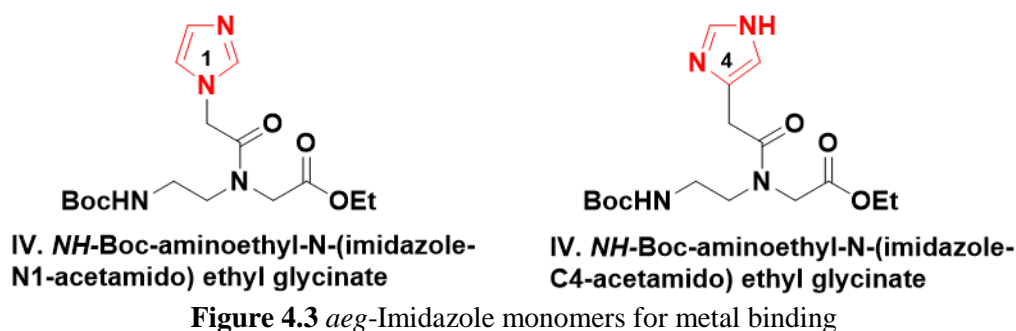
- UV spectroscopic study of various metal ions binding with *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate (III) and *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate (IV)
- UV spectroscopic study of various metal ions binding with PNA oligomers linked with *imidazole* at N1 and C4-sites and measuring UV- $T_m$  of derived PNA-PNA duplexes
- UV spectroscopic and UV-  $T_m$  studies to examine the triplex formation of PNA oligomers conjugated with imidazole via N1 / C4, *aeg*-C5-( $\beta$ -Ala-Im)-U (V) and *aeg*- $\gamma$ C-(*S*-eam-Im) (VI) with RNA
- Attempted RNA hydrolysis using imidazole conjugated PNA oligomers carrying *aeg* -C5-( $\beta$ -Ala-Im-U) PNA (V) and *aeg* - $\gamma$ C-(*S*-eam-Im) PNA (VI) and bis-imidazole (*aeg*-Bis-Im) moieties and product analysis using HPLC.

### 4.3 Results and Discussion

The imidazole modified *aeg*-PNA monomers *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) glycine and *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) glycine were synthesized and incorporated into PNA oligomers at defined sites. The metal complexation of imidazole units in monomers and the corresponding PNA oligomers were studied by UV-Vis and temperature dependent UV spectroscopic techniques. The RNA hydrolysis catalysed by imidazole units in different PNA oligomers (Figure 4.2 V, VI and VII) were studied by HPLC.

#### 4.3.1 UV studies of imidazole *aeg* monomers

This section describes the studies of metal complexation of imidazole *aeg* monomers (Figure 4.3) and metals ion by UV spectroscopy.



### 4.3.2 UV-Vis spectroscopic studies of metal complexation by the *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate (I)

The complexation studies of *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** with different metal ions were carried out using copper nitrate, nickel nitrate, cobalt nitrate, zinc nitrate, and silver nitrate. A change in the ultraviolet (UV) absorption upon complexation with individual metal ions was used to quantify the strength of complex formation. The titration with solutions of different metal salts with *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** was followed by changes in their UV-Vis spectra.

**4.3.2a Titration of I with  $\text{Cu}(\text{NO}_3)_2$ :** Upon addition of aliquots of  $\text{Cu}(\text{NO}_3)_2$  (5mM) to *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** in methanol, a progressive reduction in the intensity at 292 nm was observed along with two isobestic points at 278 nm and 310 nm. This indicates the formation of complex between *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** and  $\text{Cu}^{2+}$  metal ion (Figure 4.4).

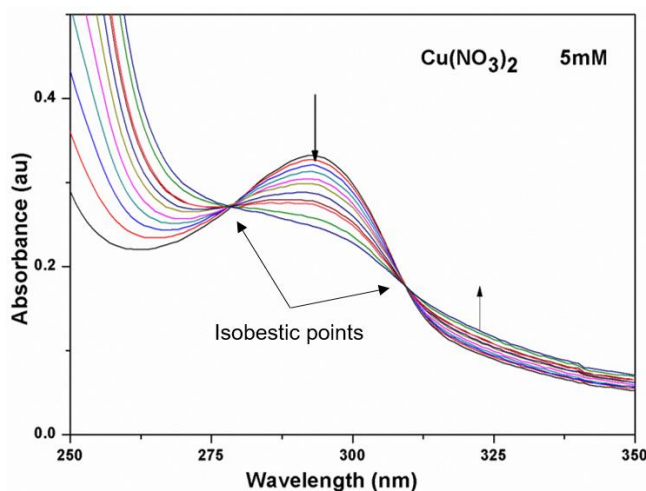
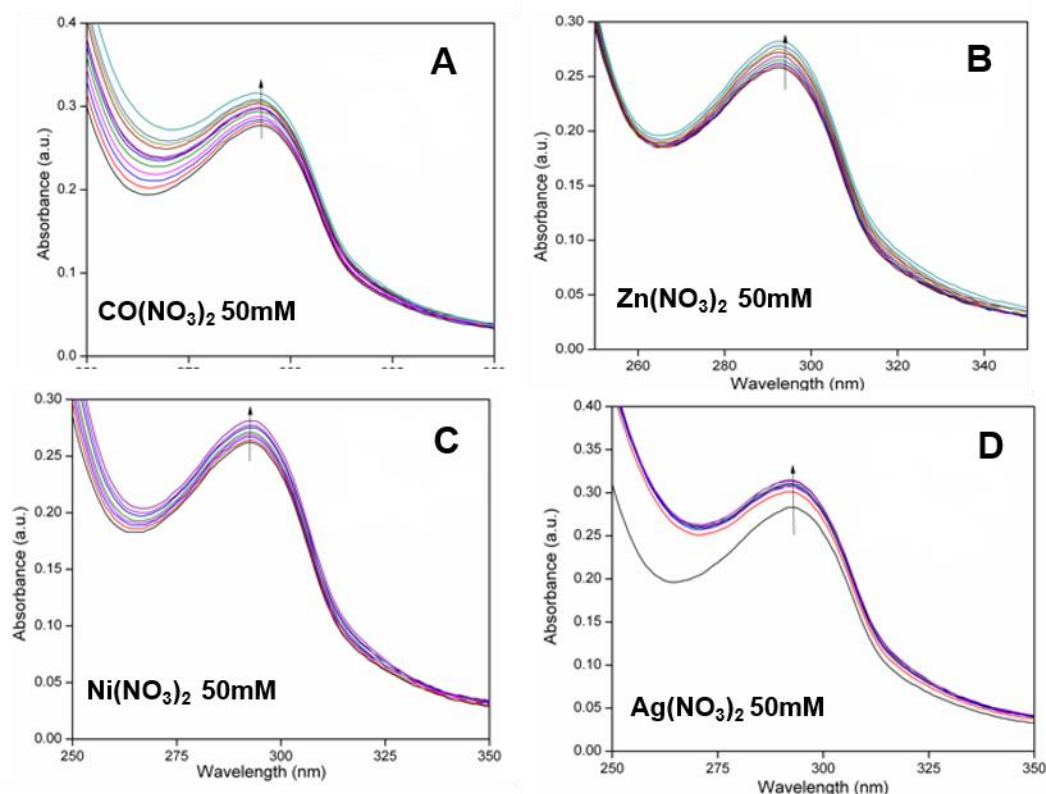


Figure 4.4 Change in the absorption spectra of the *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** (5mM) in methanol upon the addition of  $\text{Cu}^{2+}$  metal ions (5mM).

**4.3.2b Titration of I with  $\text{Co}(\text{NO}_3)_2$ ,  $\text{Ni}(\text{NO}_3)_2$ ,  $\text{Zn}(\text{NO}_3)_2$  and  $\text{Ag}(\text{NO}_3)_2$ :** Compound **I** was titrated individually with metal salts cobalt nitrate, zinc nitrate, nickel nitrate and silver nitrate in methanol and the results revealed no appreciable changes or non-appearance of new absorption bands (Figure 4.5 A-D) in the UV-Vis spectra. This suggested that the *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** did not form any complex with these metal salts, in contrast to that seen in titration with  $\text{Cu}(\text{NO}_3)_2$ .



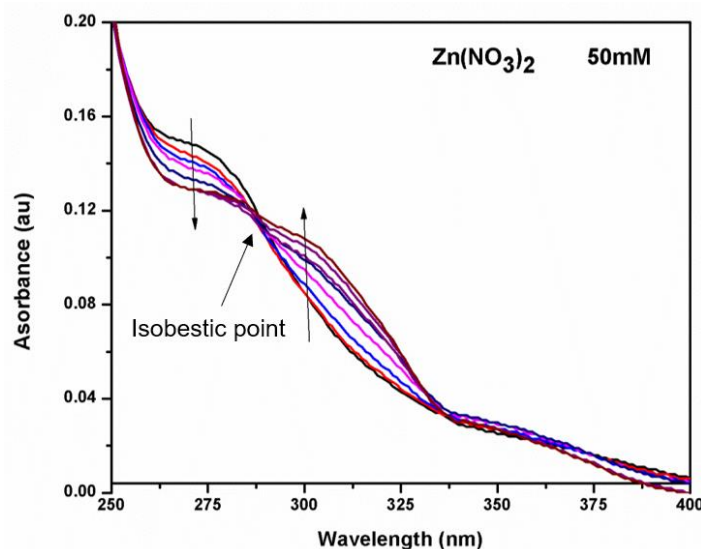
**Figure 4.5** UV absorption spectra of the *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate **I** (5mM) in methanol upon the addition of metal salt; A)  $\text{Co}(\text{NO}_3)_2$  (50 mM); B)  $\text{Ni}(\text{NO}_3)_2$  (50 mM); C)  $\text{Zn}(\text{NO}_3)_2$  (50 mM) and D)  $\text{Ag}(\text{NO}_3)_2$ .

### 4.3.3 UV-Vis spectroscopic studies of metal complexation of *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate (**II**)

The complexation studies of *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate (**II**) with different metal ions were carried out using copper nitrate, nickel nitrate, cobalt nitrate, zinc nitrate, silver nitrate. A change in the ultraviolet (UV) absorption upon complexation with individual metal ions was used to quantify the strength of complex formation. The titration of **II** with solutions of different metal salts *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate **II** followed by changes in their UV-Vis spectra.

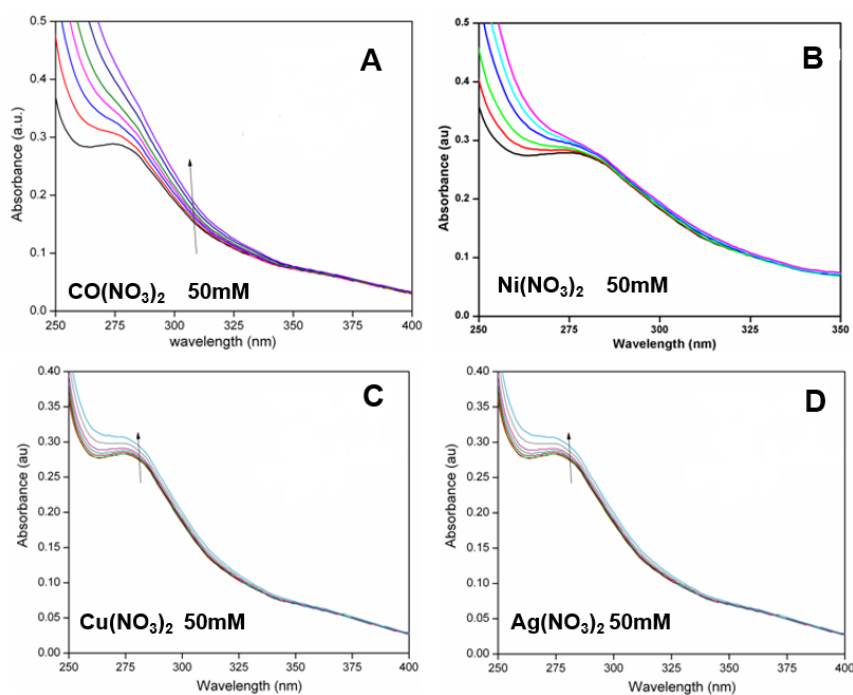


**4.3.3a Titration of II with  $Zn(NO_3)_2$ .** Upon titration of *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate **II** with  $Zn(NO_3)_2$  (5mM), a reduction in the intensity at 274 nm was observed along with two isosbestic points at 290 nm and 337 nm (Figure 4.6). This indicated the formation of complex between *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate **II** and  $Zn^{2+}$  ion.



**Figure 4.6** Change in the absorption spectra of the *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate **II** (5 mM) in methanol upon the addition of metal salt  $Zn(NO_3)_2$  (50 mM).

**4.3.3b Titration of II with  $Co(NO_3)_2$ ,  $Ni(NO_3)_2$ ,  $Cu(NO_3)_2$  and  $Ag(NO_3)$ ,** Compound **II** was titrated individually with solutions of cobalt nitrate, zinc nitrate, nickel nitrate and silver nitrate in methanol. The results showed the absence of any changes or appearance of new absorption bands (Figure 4.7 A-D) in the spectra. This indicated that *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate **II** does not form complexes with these metal salts.



**Figure 4.7** UV absorption spectra of *NH*-Boc-aminoethyl-*N*-(imidazole-C4-acetamido) ethyl glycinate **II** (5mM) in methanol upon the addition of metal salt; A)  $\text{Co}(\text{NO}_3)_2$  (50 mM); B)  $\text{Ni}(\text{NO}_3)_2$  (50 mM); C)  $\text{Cu}(\text{NO}_3)_2$  (50 mM) and D)  $\text{Ag}(\text{NO}_3)_2$ .

#### 4.3.4 *aeg*-Imidazolyl PNA oligomers: Thermal stability stabilities of duplexes and triplexes

The self complementary PNA oligomers derived from incorporation of N1-imidazole (**PNA 1**) and C4-imidazole (**PNA 2**) containing monomers (I and II) in self-complementary PNA sequence were expected to affect the stability of derived PNA-PNA duplexes and their complexes with metal ions (Table 4.1 entry 1 and 2). In addition, the imidazole containing PNA oligomer ( $U^{Im}T_{10}$ -**PNA 4**), the backbone conjugated imidazole PNA ( $\gamma C^{Im}T_{10}$ -**PNA 5**), the bis-imidazole PNA (*Bis-Im* **PNA 6**) and the corresponding control  $T_{10}$ -**PNA 3** (Table 4.1 entry 4, 5, 6 and 3 respectively) were synthesised for complexation study. Mixed sequence PNA oligomers with imidazole conjugation ( $U^{Im}$ -**PNA 7**) and backbone conjugation ( $\gamma C^{Im}$ -**PNA 8**) in were also synthesised to examine thermal stability of their duplexes. These PNAs were also examined for their ability for potential cleavage of cRNAs (**RNA 1-3**) through **PNA<sub>2</sub>:RNA 1** triplex (with  $T_{10}$ -PNAs) and **PNA:RNA 2-3** duplex formation.

**Table 4.1** Sequences of oligomers PNAs 1-8 and RNAs 1-3

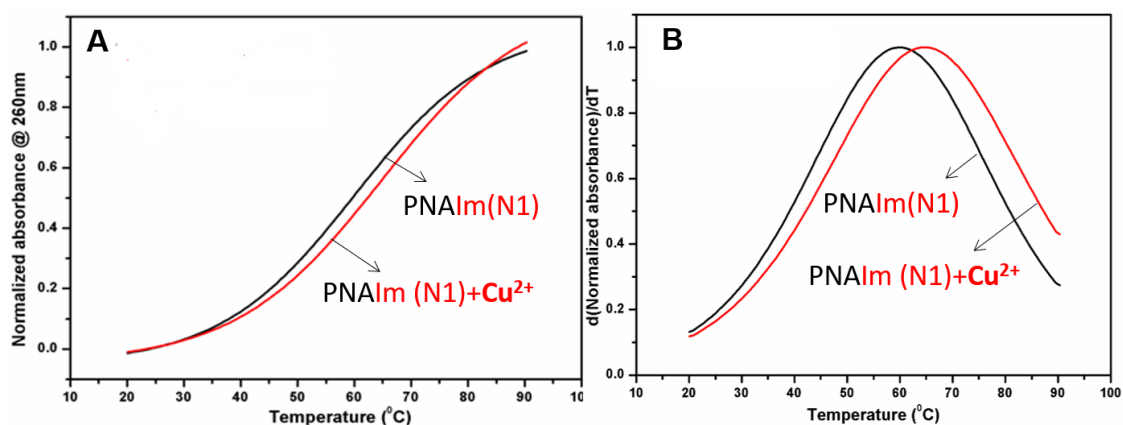
Entry	Sequence Code	PNA sequences	Remarks
1	<i>Im(N1)</i> -PNA 1	H-GGCA <i>Im(N1)</i> TGCCLysNH <sub>2</sub>	N1 and C4-Linked Imidazole modified for metal complexes
2	<i>Im(C4)</i> -PNA 2	H-GGCA <i>Im(C4)</i> TGCCLysNH <sub>2</sub>	
3	T <sub>10</sub> -PNA 3	H-TTTTTTTTTTTLysNH <sub>2</sub>	Control
4	<i>U<sup>Im</sup></i> T <sub>10</sub> -PNA 4	H-TTTT <i>U<sup>Im</sup></i> <i>U<sup>Im</sup></i> TTTTLysNH <sub>2</sub>	Base imidazole modified
5	<i>γC<sup>Im</sup></i> T <sub>10</sub> -PNA 5	H-TTTT <i>γ<sup>Im</sup></i> <i>γ<sup>Im</sup></i> TTTTLysNH <sub>2</sub>	Backbone imidazole modified
6	<i>Bis-Im</i> -PNA 6	<i>Bis-Im</i> -TTTTTGCLysNH <sub>2</sub>	Bis-imidazole modified
7	<i>U<sup>Im</sup></i> -PNA 7	H-TCTCAAG <i>U<sup>Im</sup></i> TGGGLysNH <sub>2</sub>	Base imidazole modified
8	<i>γ<sup>Im</sup></i> -PNA 8	H-TCTCAAG <i>γ<sup>Im</sup></i> TGGGLysNH <sub>2</sub>	Backbone imidazole modified
	<b>RNA Code</b>	<b>RNA Sequence</b>	
10	RNA 1	5' GCAAAAAAAAAACG 3'	Complementary RNAs
11	RNA 2	5' AGAGUUCAUAAGCCC 3'	
11	RNA 3	5' AGAGUUCAAAAGCCC 3'	

Formation of stable PNA:RNA duplexes / triplexes is critical to the eventual step of sequence directed hydrolysis. In order to investigate the effect of imidazole modified PNA on the thermal stability of various PNA:PNA/RNA duplexes, temperature dependent UV absorbance were measured in the absence and presence of metal salts. The PNA:PNA duplexes were obtained by annealing corresponding PNA oligomers in 10 mM phosphate buffer (pH 7.2) containing 10 mM NaCl in the absence and presence of metal salts. NaCl was added to decrease the electrostatic repulsion between two strands. UV- $T_m$  was employed to understand the effect of modifications on the PNA and to determine whether metal complexation affects duplex stability.

The thermal stability ( $T_m$ ) of the PNA:PNA duplexes was evaluated from melting curves monitored at 260 nm. The temperature-absorbance curves were sigmoidal which suggested the formation of 1:1 PNA:PNA duplexes. The  $T_m$  derived from midpoint of first derivative plots.

#### 4.3.4a UV melting studies of self-complementary N1-imidazole linked PNA:PNA duplex.

The thermal stability of self-complementary  $[Im(N1)PNA \mathbf{1}]_2$  duplex ( $T_m$ ) was 59.7 °C (Figure 4.8). Upon addition of 1 eq. of  $Cu^{2+}$  ion, the duplex melting ( $T_m$ ) increased to 64.7 °C (Figure 4.8) ( $\Delta T_m = +5$  °C Table 4.2, entry 1) as confirmed by first derivative spectra. It may be pointed out that the N1-imidazole monomer formed complex selectively with  $Cu^{2+}$  ions and the PNA self-complementary duplex held together by two imidazoles mediated by  $Cu^{2+}$  ion (Figure 4.9). The addition of  $Cu^{2+}$  ions thus significantly enhanced the stability of ligand modified PNA duplexes. Figure 4.9 shows the possible structure of the  $Cu^{2+}$  complex of  $[Im(N1)PNA \mathbf{1}]_2$  duplex.

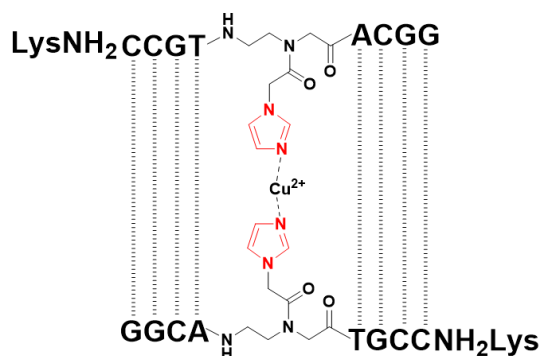


**Figure 4.8** A) Melting curves of the Imidazole N1-linked self-complementary duplex  $[Im(N1)PNA \mathbf{1}]_2$  in the presence of  $Cu^{2+}$  ion.  $[Im(N1)PNA \mathbf{1}] = 10 \mu M$ ,  $[Cu^{2+}] = 5 \mu M$  in 10 mM PBS buffer (pH 7.4), 10 mM NaCl.

**Table 4.2:**  $T_m$  (°C) of PNA-1 in the absence and presence of metal salt.

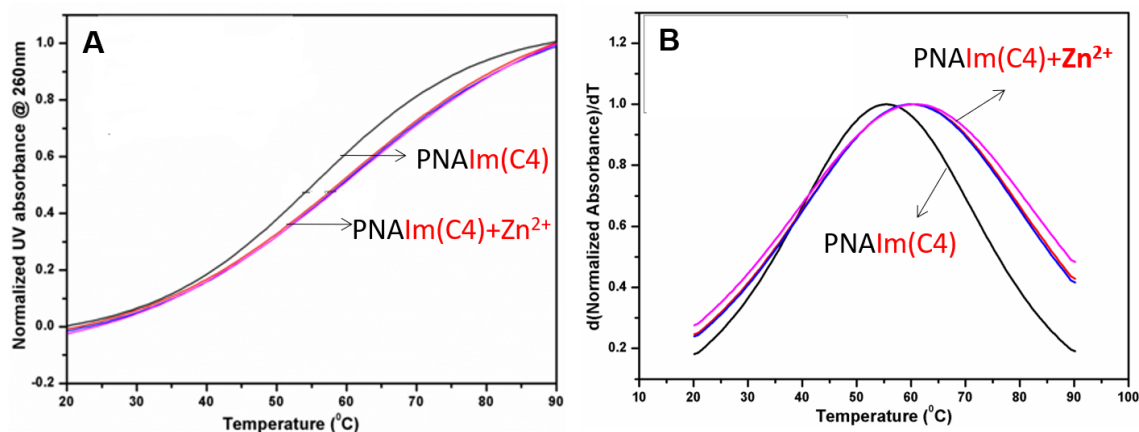
Sr No.	PNA <b>1</b>	No Metal ( $T_m$ )	$Cu^{2+}$ (1eq) ( $T_m$ )	( $\Delta T_m$ )
1	H-GGCA $Im(N1)$ TGCCLysNH <sub>2</sub>	59.7 °C	64.7 °C	+5

The  $T_m$  values are accurate to  $\pm 0.5$  °C.



**Figure 4.9.** Structure of N1-imidazole incorporated PNA:PNA duplex in antiparallel orientation.

**4.3.4b UV melting studies of self-complementary C4-imidazole linked PNA:PNA duplex.** The thermal stability of self-complementary  $[Im(C4)PNA\ 2]_2$  duplex ( $T_m$ ) was 55.4 °C (Figure 4.10). Upon addition of with one equivalent  $Zn^{2+}$  ion to  $Im(C4)PNA\ 2$ , the duplex melting ( $T_m$ ) increased to 60.4 °C ( $\Delta T_m = +5.0$  °C Table 4.3). UV melting studies of self complementary  $[Im(C4)PNA\ 2]_2$  duplex done with increasing equivalents of  $Zn^{2+}$  ion (Figure 4.10) did not show any further increase in the  $T_m$ . Thus, the addition of  $Zn^{2+}$  ions enhanced the stability of Im-C4 linked PNA duplex This is similar to the C4-imidazole monomer forming selective complexation with  $Zn^{2+}$  ions. Figure 4.11 shows the possible structure of the  $Zn^{2+}$  complex of  $[Im(C4)PNA\ 2]_2$  duplex.

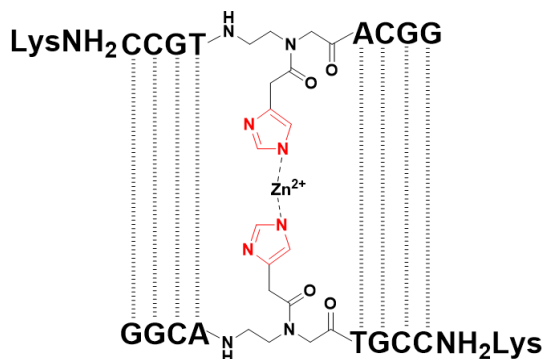


**Figure 4.10** A) Melting curves of the Imidazole-C4-conjugated self-complementary  $[Im(C4)PNA\ 2]_2$  duplex in the presence of  $Zn^{2+}$  ion  $[Im(C4)PNA\ 2]= 10\ \mu M$ ,  $[Zn^{2+}]= 5\ \mu M$  in 10 mM PBS buffer (pH 7.4), 10 mM NaCl

**Table 4.3:**  $T_m$  (°C) of PNA-2 in the absence and presence of metal salt.

Sr No.	PNA 2	No Metal ( $T_m$ )	$Zn^{2+}$ (1eq) ( $T_m$ )	( $\Delta T_m$ )
1	H-GGCA $Im(C4)$ TGCC $LysNH_2$	55.4 °C	60.4 °C	+5

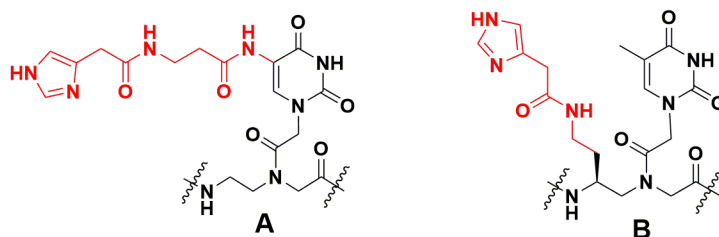
The  $T_m$  values are accurate to  $\pm 0.5$  °C.



**Figure 4.11.** Structure of imidazole C4-linked antiparallel PNA:PNA duplex.

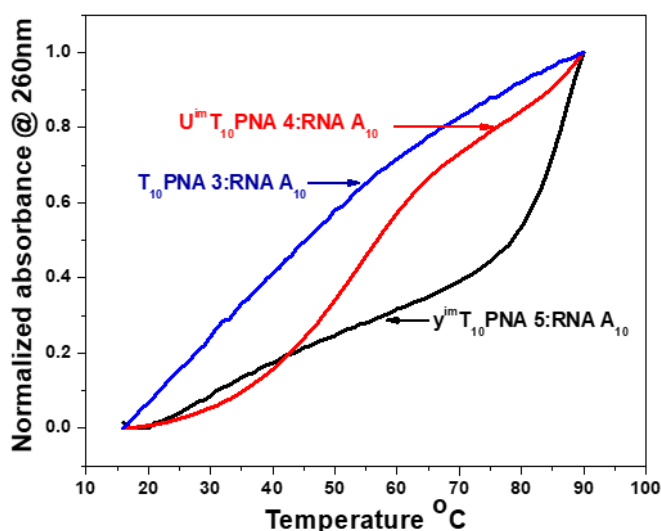
### 4.3.5 Thermal melting studies of triplexes

Biophysical studies of PNA<sub>2</sub>:RNA triplexes derived from *U*<sup>Im</sup>T<sub>10</sub>-PNA **4** and *γC*<sup>Im</sup>T<sub>10</sub>-PNA **5** (Figure 4.12 A and B) with cRNA **1** were investigated for the effect of imidazole modification on formation of triplexes. These are polypyrimidine sequences and hence form triplex with complementary polypurine RNA **1**. The thermal stability of triplexes was evaluated by temperature dependent UV absorbance change monitored at 260 nm and showed sigmoidal melting curves and the *T*<sub>m</sub> derived from midpoint of first derivative plots.



**Figure 4.12** Imidazole modified PNA **3** and PNA **4**

Unlike DNA and RNA triplexes, which show two distinct transitions corresponding to lower triplex and a higher duplex melting, and PNA triplexes with RNA and DNA (PNA<sub>2</sub>:DNA/RNA) normally show single transition with simultaneous dissociation of both strands. In comparison, it was seen that imidazole modified PNA<sub>2</sub>:RNA triplexes show double transitions similar to that observed in DNA/RNA triplexes.



**Figure 4.13**  $T_m$  curves of triplexes from  $U^{Im}T_{10}$ -PNA 4 and  $\gamma C^{Im}T_{10}$ -PNA 5 with RNA 1. Buffer: 10mM sodium phosphate, pH 7.2, NaCl 10 mM.

Two equivalents of the homopyrimidine  $T_{10}$ -PNA 3 was hybridised with one equivalent of complementary RNA 1 to form  $(PNA\ 3)_2:RNA\ 1$  triplex. As expected, this showed single transition in the melting with  $T_m$  of 45.5 °C (Figure 4.13). However, the two imidazole modified PNAs  $U^{Im}T_{10}$ -PNA 4 and  $\gamma C^{Im}T_{10}$ -PNA 5 (Figure 4.13) showed double sigmoidal melting curves, with  $T_{ms}$  56.5 °C and 54.5 °C respectively, corresponding to stabilization ( $\Delta T_m = +11.2$  °C and +9.2 °C Table 4.4, entry 2 and 3) over control  $(PNA\ 3)_2:RNA\ 1$  triplex.

**Table 4.4:** Melting studies of PNA: RNA: PNA triplexes

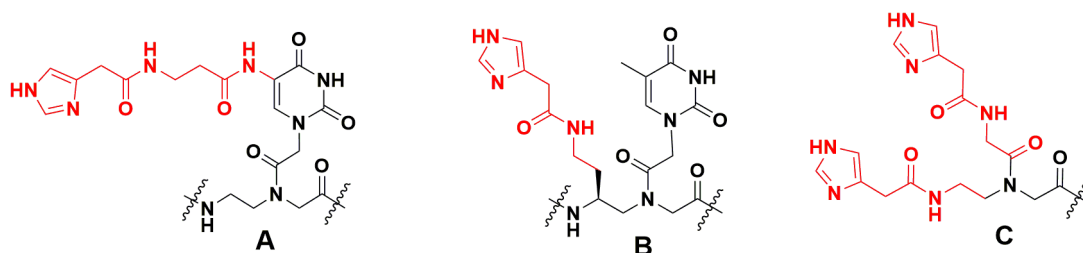
Sr. No.	PNA Code	PNA Sequence	RNA 5' GCAAAAAAAAAACG 3'	
			UV- $T_m$ (°C)	( $\Delta T_m$ ) (°C)
1	$T_{10}$ PNA 3	H-TTTTTTTTTT $LysNH_2$	45.3	Nd
2	$U^{Im}T_{10}$ PNA 4	H-TTTT $U^{Im} U^{Im}$ TTTT $LysNH_2$	56.5	+11.2
3	$\gamma C^{Im}T_{10}$ PNA 5	H-TTTT $\gamma^{Im} \gamma^{Im}$ TTTT $LysNH_2$	54.5	+9.2

The  $T_m$  values are accurate to  $\pm 0.5$  °C.

#### 4.3.6 RNA cleavage studies

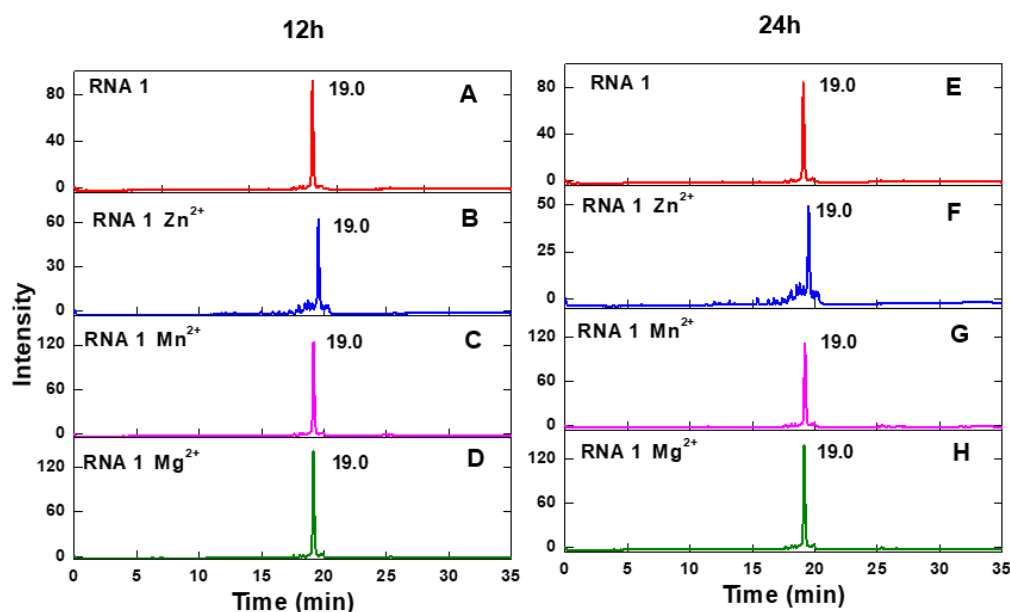
The base modified imidazole PNAs viz PNAs  $U^{Im}T_{10}$ -PNA 4, backbone modified  $\gamma C^{Im}T_{10}$ -PNA 5 and end-modified *Bis-Im*-PNA 6 were used for sequence directed cRNA cleavage under a variety of conditions (Table 4.1) and the reactions were followed

by high performance liquid chromatography (HPLC) for analysis of RNA cleavage products.



**Figure 4.14** Imidazole modified PNAs for RNA cleavage A)  $U^{Im}T_{10}$ -PNA **4**, B)  $\gamma C^{Im}T_{10}$ -PNA **5** and C) *Bis-Im*-PNA **6**

**4.3.6a Cleavage studies of RNA 1 with metals  $ZnCl_2$ ,  $MgCl_2$ ,  $MnCl_2$ .** To examine the metal ion catalysed cleavage of **RNA 1**, it was incubated with  $ZnCl_2$ ,  $MgCl_2$  and  $MnCl_2$  (300  $\mu M$ ) for 12 h (Figure 4.15 A-D) at pH 7.2 and the reaction was followed by HPLC. These studies showed that **RNA 1** was stable for 12h-24 h in presence of  $MnCl_2$  and  $MgCl_2$ . (Figure 4.15 C,D, G and H), while in presence of  $ZnCl_2$  a partial cleavage was noticed with appearance of slightly faster eluting minor peak (Figure 4.15 B and F)

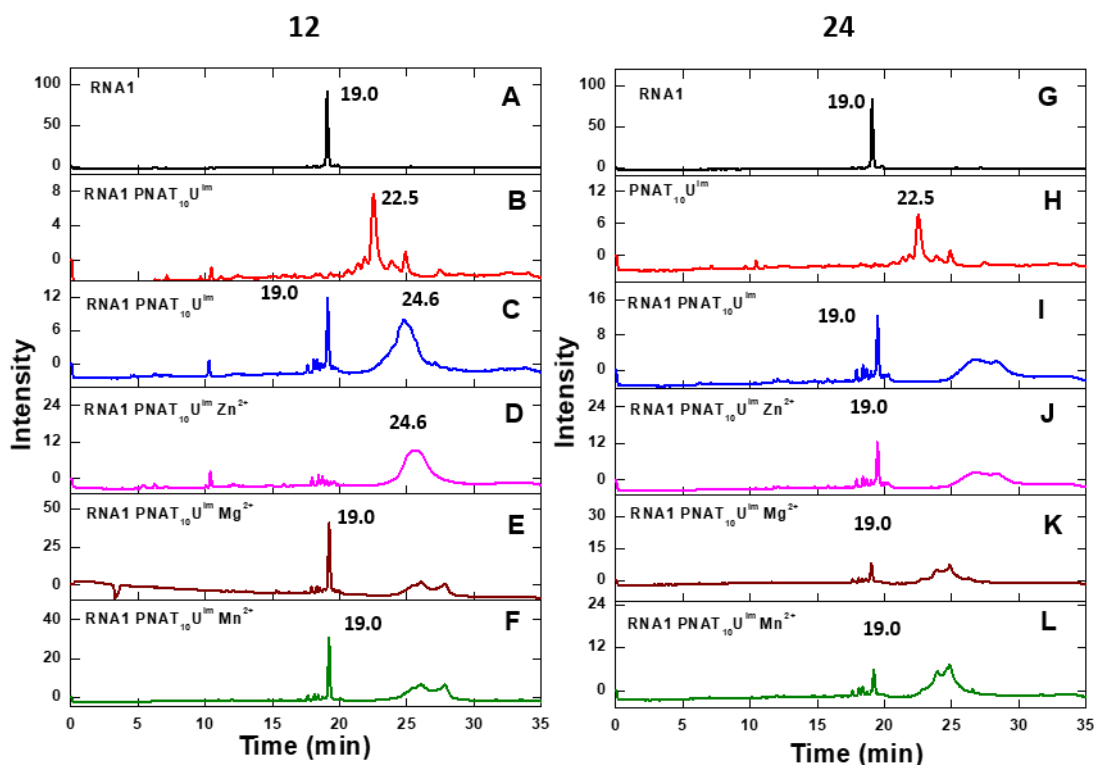


**Figure 4.15** (A-D) 12 h and E-H 24 h) Cleavage studies of **RNA 1** with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C,  $M^{2+}$  300  $\mu M$ ); A and E) **RNA 1** (10  $\mu M$ ); B and F) **RNA 1** +  $Zn^{2+}$ ; C and G) **RNA 1** +  $Mg^{2+}$  and D and H) **RNA 1** +  $Mn^{2+}$ .

**4.3.6b Cleavage studies of RNA 1 directed by  $U^{Im}T_{10}$ -PNA **4** and metals  $ZnCl_2$ ,  $MgCl_2$ ,  $MnCl_2$  via triplex formation.** To study of cleavage of **RNA 1** directed by imidazole PNA, two equivalents of imidazole modified  $U^{Im}T_{10}$ -PNA **4** was hybridised with one equivalent of **RNA 1** to form the  $(PNA\ 4)_2:RNA\ 1$  triplex. Two imidazole units from each of  $U^{Im}T_{10}$ -PNA **4** strand would come together to increase the local concentration of putative



catalytic imidazole moiety for **RNA 1** cleavage. All samples prepared in 10 mM HEPES buffer (pH 7.2) containing 10 mM NaCl and separately each with 300  $\mu\text{M}$  of  $\text{ZnCl}_2$ ,  $\text{MgCl}_2$  or  $\text{MnCl}_2$ , and the control sample without any metal salt at 37  $^\circ\text{C}$ . At regular time intervals, small fractions of aliquots were removed and subjected for RP-HPLC analysis using triethyl amine acetate buffer (pH 7.2) as the mobile phase on C18 column.



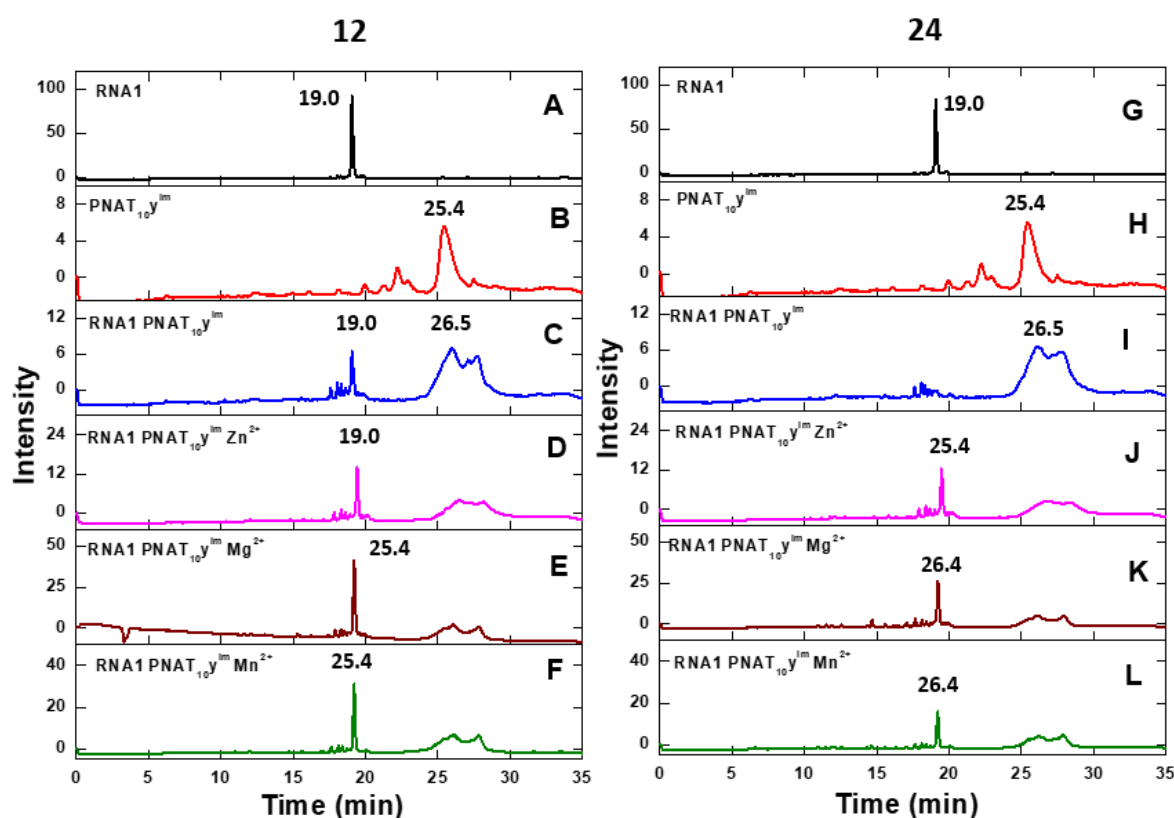
**Figure 4.16** (A-F 12 h and G-L 24 h) Cleavage study of **RNA 1** by  $U^{Im}T_{10}$ -**PNA 4** with metals and without metal (10mM HEPES buffer, pH 7.4, 10 mM NaCl, 37  $^\circ\text{C}$ ,  $M^{2+}$  300  $\mu\text{M}$ ); A) and G) **RNA 1**; B and H)  $U^{Im}T_{10}$ -**PNA 4** (10  $\mu\text{M}$ ); C and I) **RNA 1**+  $U^{Im}T_{10}$ -**PNA 4** (10  $\mu\text{M}$ : 20  $\mu\text{M}$ ); D) and J) **RNA 1**+  $U^{Im}T_{10}$ -**PNA 4** +  $\text{Zn}^{2+}$ ; E and K) **RNA 1** +  $U^{Im}T_{10}$ -**PNA 4** +  $\text{Mg}^{2+}$  and F) and L) **RNA 1**+  $U^{Im}T_{10}$ -**PNA 4**+  $\text{Mn}^{2+}$ .

$U^{Im}T_{10}$ -**PNA 4** bearing two imidazole at middle eluted at 22.5 min in HPLC (Figure 4.16 B and H) and was reacted with with **cRNA 1** at 37  $^\circ\text{C}$  for 12 h and 24h in presence and absence of metal (Figure 4.16 A-F and G-L). The reaction mixture analysed by RP-HPLC showed a major peak at 19.0 min corresponding to **RNA 1** (Figure 4.16 C-F and I-L) with some minor satellite peaks in presence of metal salts. This suggested insignificant cleavage of **RNA 1**. Additionally, a broad second peak was observed at 24.6 min which could be the complex of  $U^{Im}T_{10}$ -**PNA 4**:**RNA 1** (Figure 4.15 C-F and I-L).

**4.3.6c Cleavage studies of RNA 1 directed by  $\gamma C^{Im}T_{10}$ -PNA 5 and metals  $\text{ZnCl}_2$ ,  $\text{MgCl}_2$ ,  $\text{MnCl}_2$  via triplex formation.** The cleavage of **RNA 1** was attempted by  $\gamma C^{Im}T_{10}$ -**PNA 5** containing imidazole linked at  $\gamma C$  via ethylaminoacetyl sidechain  $\gamma C^{Im}T_{10}$ -**PNA 5** was

hybridised with **RNA 1** in 2:1 ratio to form (**PNA 5**)<sub>2</sub>:**RNA 1** triplex, wherein the two imidazole moieties from each strand of  $\gamma C^{Im}T_{10}$ -**PNA 5** would converge to increase the local concentration of imidazole, making the site better for catalytic **RNA 1** cleavage. All samples were prepared in 10 mM HEPES buffer, (pH 7.2) containing 10 mM NaCl and individual reactions carried out at 37 °C in presence of 300  $\mu$ M ZnCl<sub>2</sub>, MgCl<sub>2</sub> and MnCl<sub>2</sub>. The control reactions did not have any metal salts.

Imidazole modified  $\gamma C^{Im}T_{10}$ -**PNA 5** bearing the two imidazole moieties eluted at 25.4 min (Figure 4.17 B and H) in RP HPLC. **RNA 1** was subjected to cleavage reaction at 37 °C for 12 h and 24 h with the complementary  $\gamma C^{Im}T_{10}$ -**PNA 5** to with and without metal ions (Figure 4.17 C-F and 4.17 I-L). The reaction mixture was analysed by RP-HPLC at 12 h and 24 h and the minor peaks observed before the major RNA peak at 19 min in presence and (Figure 4.17 C-F and I-L)) indicate partial cleavage of **RNA 1**, although not significant. The other broad peak observed at 24.6 min is perhaps due to the



**Figure 4.17** (A-F 12 h and G-L 24 h) Cleavage study of **RNA 1** by  $\gamma C^{Im}T_{10}$ -**PNA 5** with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C,  $M^{2+}$  300  $\mu$ M); A) and G) **RNA 1**; B) and H)  $\gamma C^{Im}T_{10}$ -**PNA 5**; C) and I) **RNA 1**+20  $\mu$ M  $\gamma C^{Im}T_{10}$ -**PNA 5** (10  $\mu$ M:20  $\mu$ M); D)and J) **RNA 1**  $\gamma C^{Im}T_{10}$ -**PNA 5**+ Zn<sup>2+</sup>; E) and K) **RNA 1** +  $\gamma C^{Im}T_{10}$ -**PNA 5**+ Mg<sup>2+</sup> and F) and L) **RNA 1** +  $\gamma C^{Im}T_{10}$ -**PNA 5**+ Mn<sup>2+</sup>.

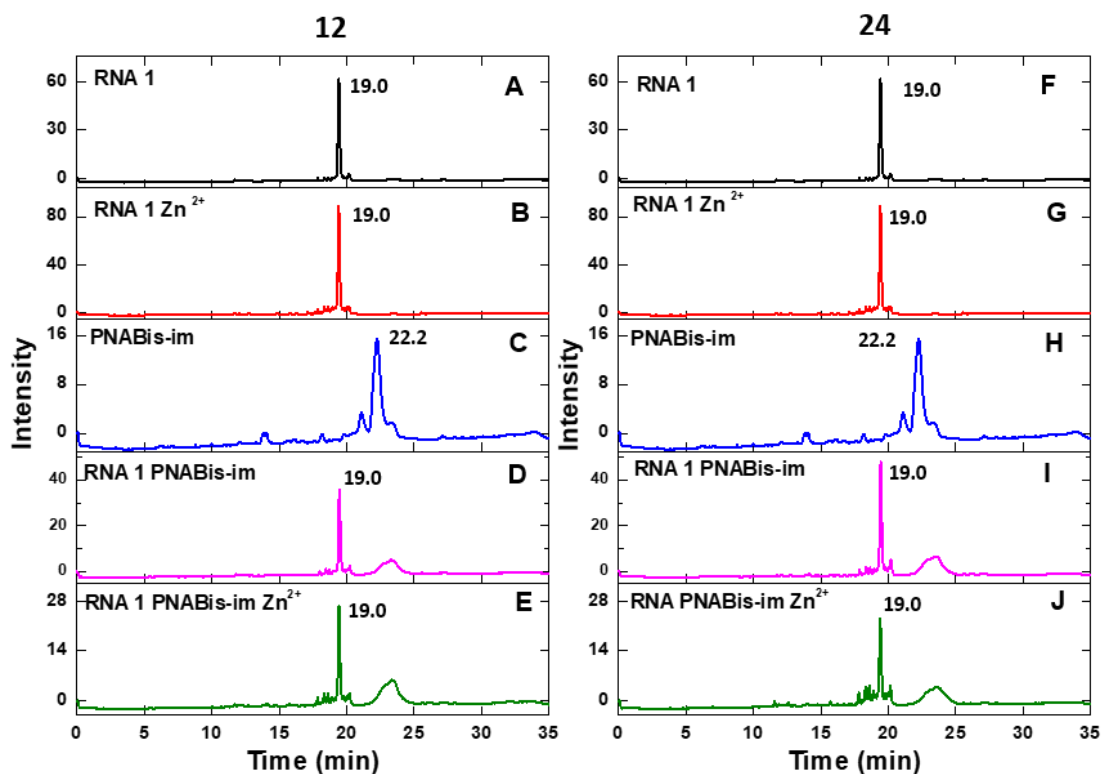
complex of  $\gamma C^{Im}T_{10}$ -PNA 5: RNA 1 (Figure 4.17 C-F and I-L). It is seen from these results that in case of  $Zn^{2+}$  reaction the amount of cleavage is slightly higher than that with other metals, suggesting that Zn-imidazole-PNA complex is slightly effective in inducing cleavage.

**4.3.6d Cleavage studies of RNA 1 directed by Bis-Im-PNA 6 and metals  $ZnCl_2$ , via duplex formation.** To study RNA 1 cleavage by Bis-Im-PNA 6, both components were hybridized in 1:2 equivalents to form PNA<sub>2</sub>:RNA triplex. One strand of Bis-Im-PNA 6 forms half duplex with RNA 1 and the other strand of Bis-Im-PNA 6 forms one half of a duplex with RNA 1 (Figure 4.18). This results in bringing four imidazole together to perform the RNA cleavage. All samples were prepared in 10 mM HEPES buffer, (pH 7.2) containing 10 mM NaCl with 300  $\mu$ M  $ZnCl_2$  and without metal ion. All fractions of reaction mixture were incubated separately at 37 °C.

Imidazole modified Bis-Im-PNA 6 eluted in RP HPLC at 22.2 min. (Figure 4.19 C and H). RNA 1 was subjected to cleavage reaction Bis-Im-PNA 6 at 37 °C for 12 h and 24 h to determine its catalytic activity without metal. The reaction mixture was injected into the RP-HPLC at 12 h and 24 h exhibited peaks around 19 min corresponding to RNA 1 (Figure 4.19 D and I) with almost no cleavage without metal. In presence of  $Zn^{2+}$  metal ion showed small satellite peaks before the major peak at 19 min (RNA 1) indicating 10-15% of cleavage of RNA 1 (Figure 4.19 E and J). The second broad peak seen at 22.3 min corresponded to Bis-Im-PNA 6:RNA 1 complex (Figure 4.19 D,E and I,J). Thus RNA 1 showed 10-15% cleavage specifically in presence of  $Zn^{2+}$  by Bis-Im-PNA 6.

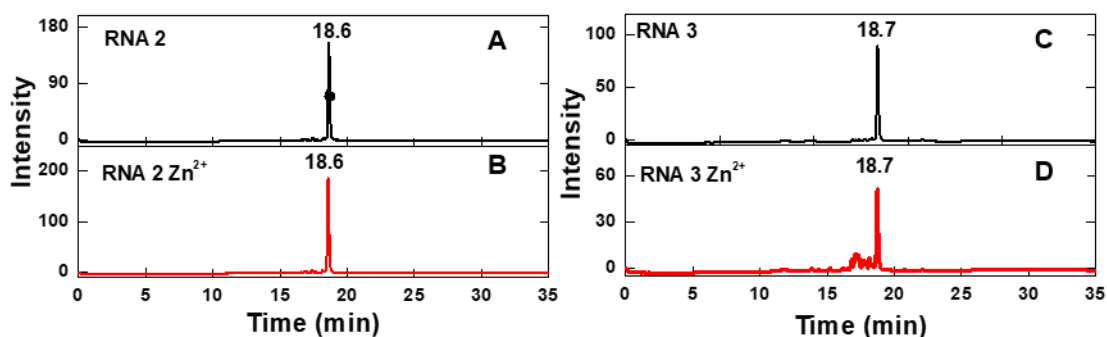


**Figure 4.18** Half duplex of Bis-Im-PNA 6:RNA 1



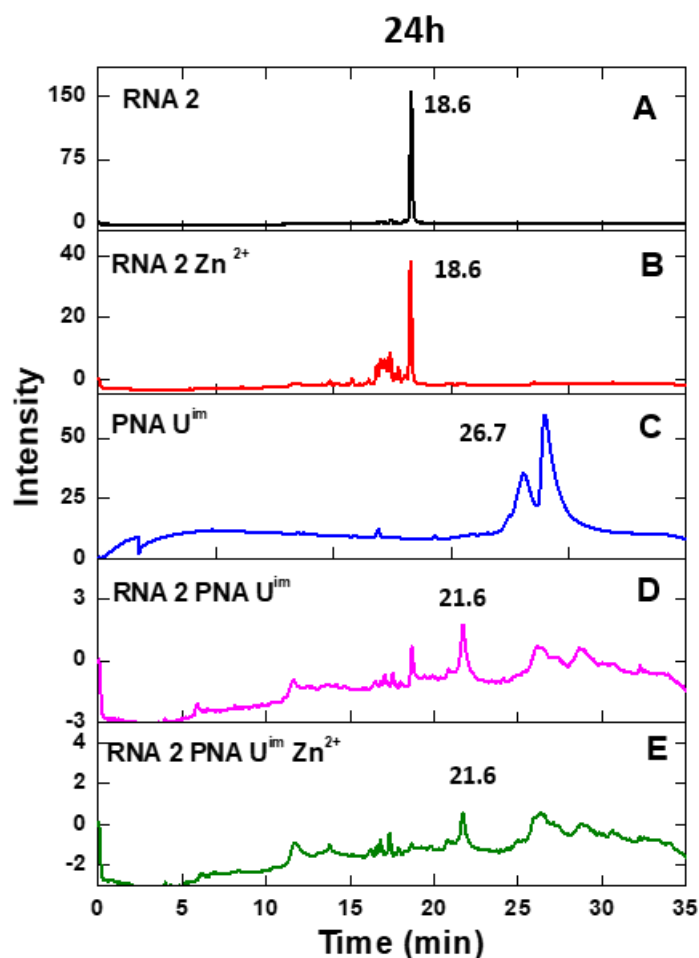
**Figure 4.19** (A-E 12 h and F-J 24 h) Cleavage study of **RNA 1** by *Bis-Im-PNA 6* with and without  $\text{Zn}^{2+}$  (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h,  $\text{Zn}^{2+}$  300  $\mu\text{M}$ ); A) and F) **RNA 1**; B and G) **RNA 1**  $\text{Zn}^{2+}$ ; C) and H) *Bis-Im-PNA 6*; D) and I) **RNA 1** + *Bis-Im-PNA 6* (10  $\mu\text{M}$ :20  $\mu\text{M}$ ); E and J) **RNA 1** + *Bis-Im-PNA 6* +  $\text{Zn}^{2+}$ .

**4.3.6e Cleavage studies of RNA 2 and RNA 3 in presence of  $\text{ZnCl}_2$ .** To examine the cleavage of non-self complementary **RNA 2** and **RNA 3** with and without  $\text{ZnCl}_2$  (300  $\mu\text{M}$ ), the reaction was initially carried out for 24 h at pH 7.2, 10mM NaCl. These studies indicated **RNA 2** and **RNA 3** were stable to reaction conditions without metal for 24 h (Figure 4.20 A and C). The reactions done in presence of  $\text{ZnCl}_2$  for 24 h showed **RNA 3** was stable (Figure 4.20 B and D).



**Figure 4.20** (A-D) Cleavage study of **RNA 2** with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h,  $\text{Zn}^{2+}$  300  $\mu\text{M}$ ); A) **RNA 2** (10  $\mu\text{M}$ ); B) **RNA 2** +  $\text{Zn}^{2+}$ ; C) **RNA 3** (10  $\mu\text{M}$ ) and D) **RNA 3** +  $\text{Zn}^{2+}$ .



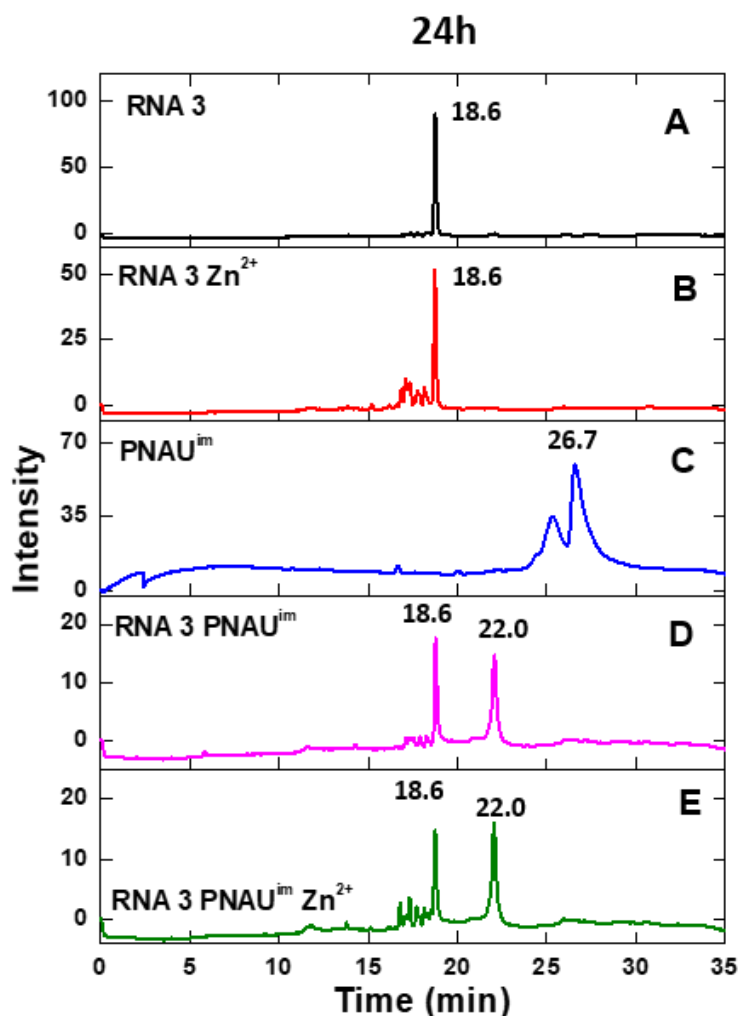


**Figure 4.22** (A-E) Cleavage study of **RNA 2** by  $U^{Im}$ -**PNA 7** with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h,  $Zn^{2+}$  300  $\mu$ M) A) **RNA 2**; B) **RNA 2**  $Zn^{2+}$ ; C)  $U^{Im}$ -**PNA 7**; D)  $U^{Im}$ -**PNA 7** (20  $\mu$ M); E) **RNA 2** +  $U^{Im}$ -**PNA 7** (10  $\mu$ M:20  $\mu$ M) and C) **RNA 2** +  $U^{Im}$ -**PNA 7** +  $Zn^{2+}$ .

**4.3.6g Cleavage of RNA 3 via duplex formation with  $U^{Im}$ -PNA 7.** To examine the comparative cleavage of **RNA 3** which forms duplex with complementary  $U^{Im}$ -**PNA 7** with a RNA bulge at AAAA, both were mixed to form duplex (Figure 4.23). This was subjected to cleavage reactions at 37°C for 24 h without metal and analysed by RP HPLC. The peak observed around 19 min (Figure 4.24 D and E) suggested no significant cleavage of **RNA 3**. The second broad peak observed at 22.3 min was due to the complex of  $U^{Im}$ -**PNA 7**:**RNA 3** (Figure 4.24 D and E).

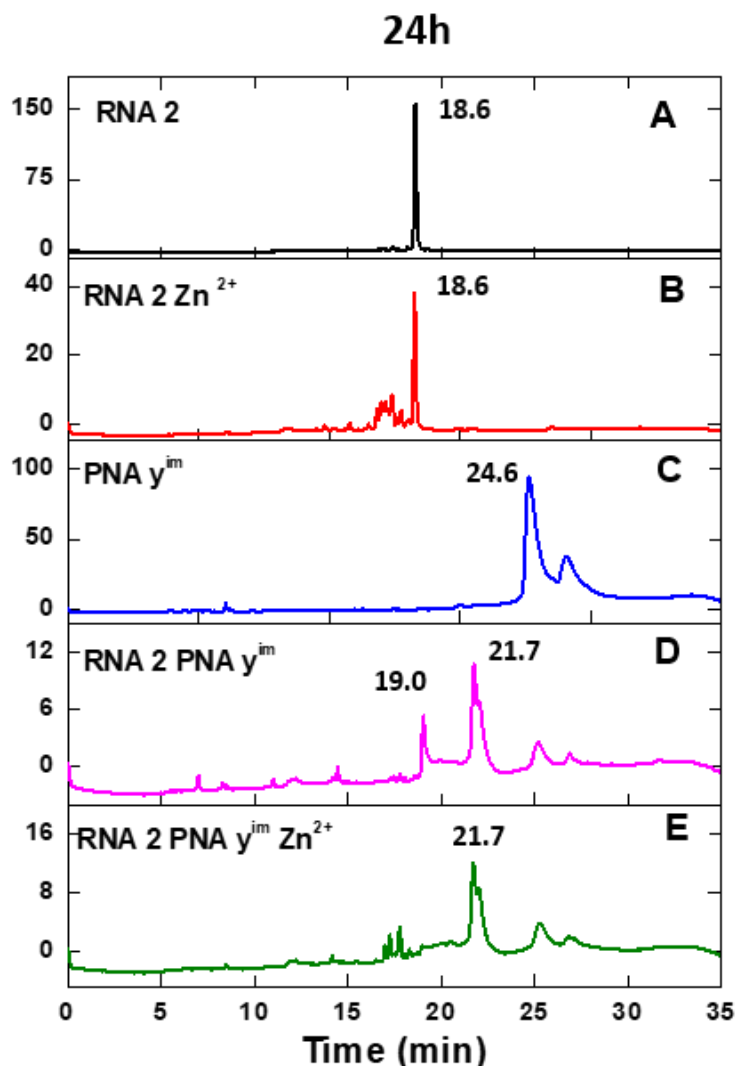


**Figure 4.23**  $U^{Im}$ -**PNA 7**: **RNA 3** duplex structure



**Figure 4.24** (A-E) Cleavage study of **RNA 3** by  $U^{Im}$ -**PNA 7** with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h,  $Zn^{2+}$  300  $\mu$ M); A) **RNA 3**; B) **RNA 3**,  $Zn^{2+}$ ; C)  $U^{Im}$ -**PNA 7**; D) **RNA 3** +  $U^{Im}$ -**PNA 7** (10  $\mu$ M:20  $\mu$ M) and E) **RNA 3** +  $U^{Im}$ -**PNA 7** +  $Zn^{2+}$ .

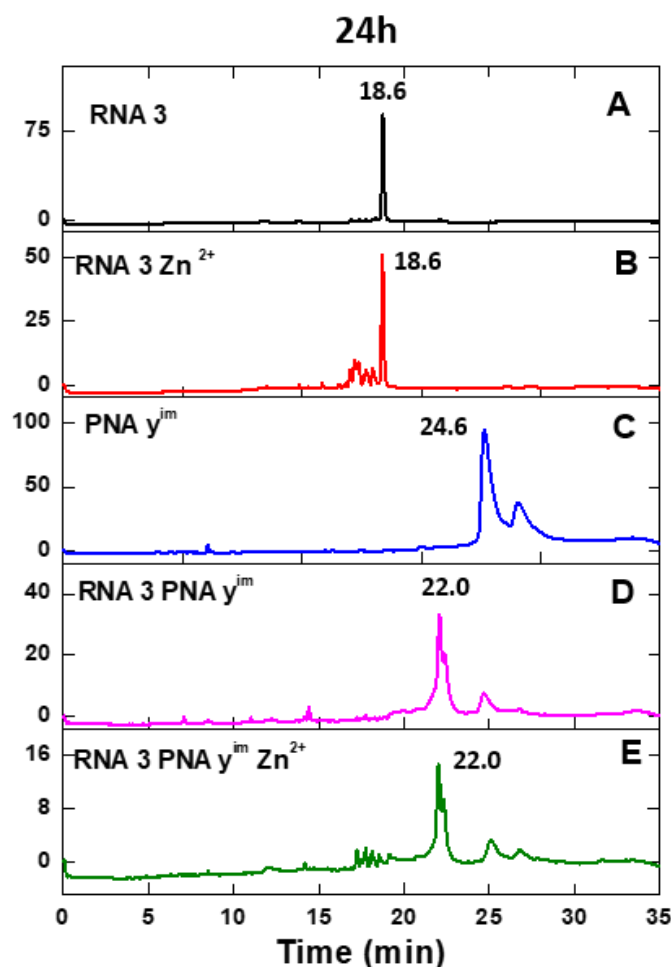
**4.3.6h Cleavage of RNA 2 via Duplex formation with  $\gamma C^{Im}$ -PNA 8.** Backbone modified  $\gamma C^{Im}$ -**PNA 8** elutes at 24.6 min. (Figure 4.25). It was employed for cleavage of **cRNA 2** by reaction at 37°C for 24 h to determine its catalytic activity with and without metal. The  $\gamma C^{Im}$ -**PNA 8** eluted at 24.6 min in RP HPLC (Figure 4.25 B) and **RNA 2** eluted at 18.6 min in presence and absence of metal (Figure 4.25 A and B). The reaction mixture was injected into the RP-HPLC at 24 h. The peak was broad and it was observed at 21.7 min corresponded to a complex of  $\gamma C^{Im}$ -**PNA 8**:**RNA 2** (Figure 4.25 D and E ). No PNA directed cleavage of **RNA 2** was observed and the minor peaks observed around 18 min were from simple cleavage of **RNA 2** by  $Zn^{2+}$  (Figure 4.25 B and E) of **RNA 3** mediated



**Figure 4.25** (A-E) Cleavage study of **RNA 2** by  $\gamma C^{Im}$ -**PNA 8** with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h,  $Zn^{2+}$  300  $\mu$ M). A) **RNA 2**; B) **RNA 2** +  $Zn^{2+}$ ; C)  $\gamma C^{Im}$ -**PNA 8**; D) **RNA 2** +  $\gamma C^{Im}$ -**PNA 8** (10  $\mu$ M:20  $\mu$ M); E) **RNA 2** +  $\gamma C^{Im}$ -**PNA 8**+  $Zn^{2+}$ .

**4.3.3i Cleavage of RNA 3 by duplex formation with  $\gamma C^{Im}$ -PNA 8.** The cleavage reaction by  $\gamma C^{Im}$ -**PNA 8** in presence and absence of metal ions was similarly carried out at 37 °C for 24 h and the products analysed by HPLC. The reaction mixture was injected into the RP-HPLC at 24 h. The broad peak observed at 22.0 min in presence and absence of metal due to  $\gamma C^{Im}$ -**PNA 8**:**RNA 3** complex (Figure 4.26 D) and E). The PNA directed cleavage of **RNA 3** was negligible.



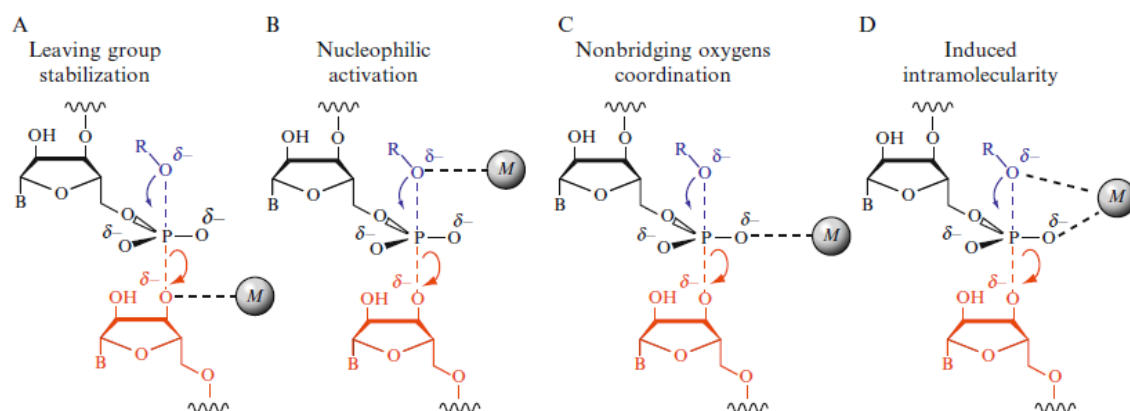


**Figure 4.26 (A-E)** Cleavage study of **RNA 3** by  $\gamma C^{Im}$ -**PNA 8** with and without  $Zn^{2+}$  (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h,  $Zn^{2+}$  300  $\mu$ M). A) **RNA 3**; B) **RNA 3** +  $Zn^{2+}$ ; C)  $\gamma C^{Im}$ -**PNA 8**; D) **RNA 3** +  $\gamma C^{Im}$ -**PNA 8** (10  $\mu$ M:20  $\mu$ M) and E) **RNA 3** +  $\gamma C^{Im}$ -**PNA 8** +  $Zn^{2+}$ .

#### 4.4 Cleavage of RNA by $Zn^{2+}$ ion

Sigurdsson *et.al.*<sup>30,31</sup> reported RNA cleavage motif (Figure 4.27) using  $Zn^{2+}$  salt which was found in the hammerhead ribozyme. It was observed that  $Zn^{2+}$  cleaves the nucleotide at G8 and A9 site specifically that yields a free 5'-hydroxyl group and a 2',3'-cyclic phosphate. The cleavage was dependent on divalent metal ions and was the first evidence wherein a metalloribozyme was found to show preference for  $Zn^{2+}$  ion. Further studies have shown that cleavage between C3 and U4 sites was unusually dependent on concentration of  $Zn^{2+}$  ion as well as on the pH of buffer system. Additionally at higher concentrations, the yield of U4 cleavage product was found to be decreased due to non-specific cleavage of the ribozyme. This phenomenon was also observed in case of A9

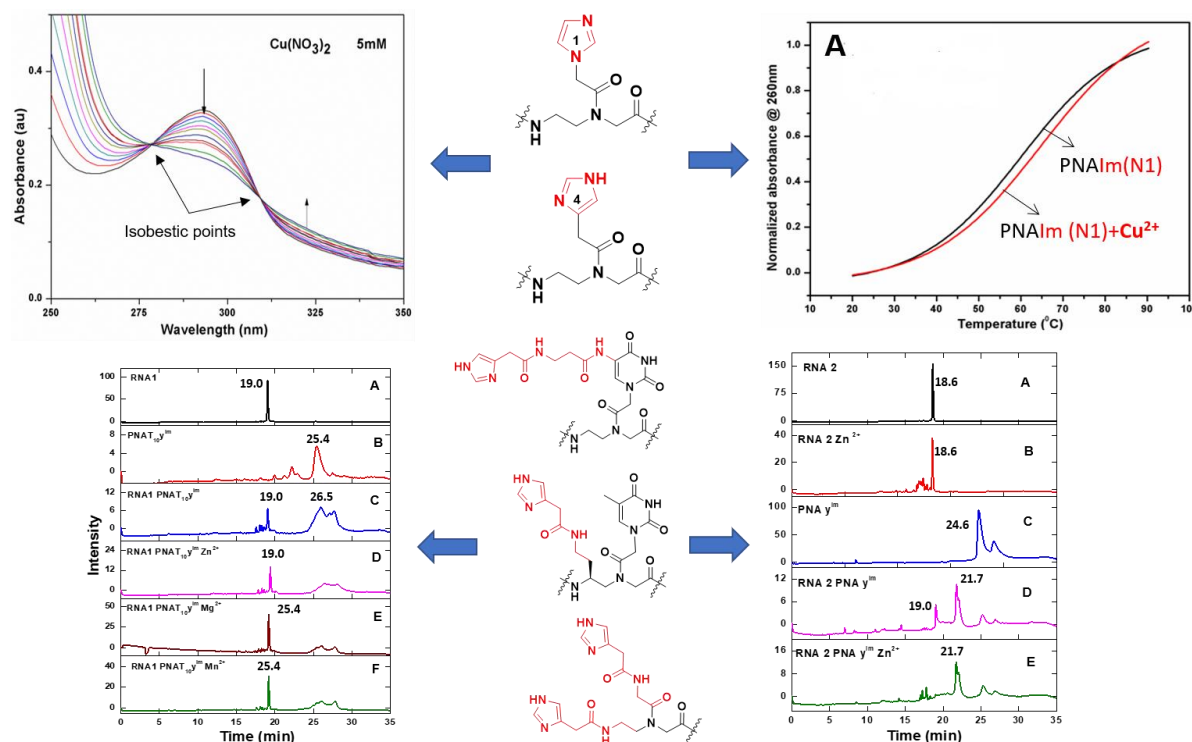
cleavage. Husket *et.al.*<sup>32</sup> showed the cleavage of RNA by  $Mg^{2+}$  and  $Ca^{2+}$  metal ion. Zagórska *et.al.*<sup>33</sup> reported RNA hydrolysis using  $Mg^{2+}$ ,  $Zn^{2+}$  and  $Pd^{2+}$  metal ion.



**Figure 4.27** Mechanism of RNA cleavage by metal ions (A) 3'-oxygen stabilization by metal (B) Nucleophile generated by metal (C) Oxygen atoms coordinated by metal (D) RNA cleavage by 3'-oxygen by divalent metal<sup>34,35</sup>

## 4.5 Summary

The *NH*-Boc-aminoethyl-N-(imidazole-N1-acetamido) ethyl glycinate and *NH*-Boc-aminoethyl-N-(imidazole-C4-acetamido) ethyl glycinate monomers showed selective metal complex with  $Cu^{2+}$  and  $Zn^{2+}$  respectively. These monomers are incorporated into the self-complementary PNA sequences. The thermal stability ( $T_m$ ) of PNA:PNA duplexes of *Im(N1)*-PNA **1** and *Im(C4)*-PNA **2** are significantly increased 5 °C specifically in presence of  $Cu^{2+}$  and  $Zn^{2+}$  respectively.



**Figure 4.27** Summary of chapter

The base modified imidazole PNAs of  $U^{Im}T_{10}$ -PNA **4**, backbone modified  $\gamma C^{Im}T_{10}$ -PNA **5** and end modified Bis-Im-PNA **6** were studied for their potential to cleave RNA **1** in the presence and absence of metal ions. In triplex and half duplex ( $U^{Im}T_{10}$ -PNA **4**)<sub>2</sub>:RNA **1**, ( $\gamma C^{Im}T_{10}$ -PNA **5**)<sub>2</sub>:RNA **1** and Bis-Im-PNA **6**:RNA **1**, the imidazole modified PNAs cleaved 10% RNA **1** in the absence of metal salts and in presence of metal salt cleavage increased upto the 20%. However no sequence directed selectivity was seen in any of the cases. With duplexes  $U^{Im}$ -PNA **7**:RNA **1**/RNA **2** and  $\gamma C^{Im}$ -PNA **8**:RNA **2**/RNA **3**, the imidazole modified PNAs did not cleave RNA **2**/ RNA **3** in presence or absence of Cu, and Mn metal ions, but RNA **2**/ RNA **3** partial degradation was observed only in presnsce of only Zn<sup>2+</sup> ion.

## 4.6 References

1. Komiyama, M.; Sumaoka, J.; Kuzuya, A.; Yamamoto, Y., Sequence-selective artificial ribonucleases. *Methods Enzymol.* **2001**, *341*, (Ribonucleases, Part A), 455-468.
2. Magda, D.; Crofts, S.; Lin, A.; Miles, D.; Wright, M.; Sessler, J. L., Synthesis and Kinetic Properties of Ribozyme Analogs Prepared Using Phosphoramidite Derivatives of Dysprosium(III) Texaphyrin. *J. Am. Chem. Soc.* **1997**, *119*, 2293-2294.

3. Magda, D.; Miller, R. A.; Sessler, J. L.; Iverson, B. L., Site-Specific Hydrolysis of RNA by Europium(III) Texaphyrin Conjugated to a Synthetic Oligodeoxyribonucleotide. *J. Am. Chem. Soc.* **1994**, *116*, 7439-40.
4. Komiyama, M., Sequence-selective scission of DNA and RNA by lanthanide ions and their complexes. *Met. Ions Biol. Syst.* **2003**, *40*, 463-475.
5. Kuzuya, A.; Machida, K.; Mizoguchi, R.; Komiyama, M., Conjugation of Various Acridines to DNA for Site-Selective RNA Scission by Lanthanide Ion. *Bioconjugate Chem.* **2002**, *13*, 365-369.
6. Komiyama, M., Sequence-selective scission of DNA and RNA by lanthanide ions and their complexes. *Met. Ions Biol. Syst.* **2003**, *40*, 463-475.
7. Kuzuya, A.; Machida, K.; Mizoguchi, R.; Komiyama, M., Conjugation of Various Acridines to DNA for Site-Selective RNA Scission by Lanthanide Ion. *Bioconjugate Chem.* **2002**, *13*, 365-369.
8. Watson, R. M.; Skorik, Y. A.; Patra, G. K.; Achim, C. Influence of Metal Coordination on the Mismatch Tolerance of Ligand-Modified PNA Duplexes. *J. Am. Chem. Soc.* **2005**, *127*, 14628-14639.
9. Popescu, D.; Parolin, T. J.; Achim, C. Metal Incorporation in Modified PNA Duplexes. *J. Am. Chem. Soc.* **2003**, *125*, 6354-6355.
10. De Leon, A. R.; Olatunde, A. O., Morrow, J. R.; Achim, C. Binding of EuIII to 1,2-Hydroxypyridinone-Modified Peptide Nucleic Acids *Inorg. Chem.* **2012**, *51*, 12597-12599.
11. a) Gasser, G.; Hüsken, N.; Köster, S. D.; Metzler-Nolte, N. Synthesis of organometallic PNA oligomers by click chemistry. *Chem. Commun.* **2008**, *45*, 3675-3677; b) Gasser, G.; Sosniak, A. M.; Metzler-Nolte, N. Metal-containing peptide nucleic acid conjugates *Dalton Trans.*; **2011**, *40*, 7061-7076.
12. Bazer, S.; Rapireddy, S.; Skorik, Y. A.; Ly Danith, H.; Achim, C. Coordination-Driven Inversion of Handedness in Ligand-Modified PNA. *Inorg. Chem.* **2011**, *50*, 11929-11937.
13. Murtola, M.; Wenska, M.; Stromberg, R. PNAszymes That Are Artificial RNA Restriction Enzymes. *J. Am. Chem. Soc.* **2010**, *132*, 8984.
14. Dogandzhiyski, P.; Ghidini, A.; Danneberg, F.; Strömberg, R.; Göbel, M. W. Studies on Tris(2-aminobenzimidazole)-PNA Based Artificial Nucleases: A Comparison of Two Analytical Techniques. *Bioconjugate Chem.* **2015**, *26*, 2514.

15. Wang, J. DNA biosensors based on Peptide Nucleic Acid (PNA) recognition layers. A review<sup>1</sup>. *Biosens. Bioelectron.* **1998**, *13*, 757-762.
16. Wang, PNA Biosensors for Nucleic Acid Detection. *J. Curr. Issues Mol. Biol.* **1999**, *1*, 117-122.
17. Kolasa, K. A., Morrow, J. R., and Sharma, A. P. Trivalent lanthanide ions do not cleave RNA in DNA-RNA hybrids. *Inorg. Chem.* **1993**, *32*, 3983-4.
18. Ozkan, D.; Erdem, A.; Kara, P.; Kerman, K.; Gooding, J. J.; Nielsen, P. E.; Ozsoz, M. Electrochemical detection of hybridization using peptide nucleic acids and methylene blue on self-assembled alkanethiol monolayer modified gold electrodes. *M. Electrochem. Commun.* **2002**, *4*, 796-802.
19. Hashimoto, K.; Ishimori, Preliminary evaluation of electrochemical PNA array for detection of single base mismatch mutations. *Y. Lab Chip*, **2001**, *1*, 61-63.
20. Hejazi, M. S.; Pournaghi-Azar, M. H.; Ahour, Electrochemical detection of short sequences of hepatitis C 3a virus using a peptide nucleic acid-assembled gold electrode. *F. Anal. Biochem.* **2010**, *399*, 118-124.
21. Ozkan, D.; Kara, P.; Kerman, K.; Meric, B.; Erdem, A.; Jelen, F.; Nielsen, P. E.; Ozsoz, M. DNA and PNA sensing on mercury and carbon electrodes by using methylene blue as an electrochemical label. *Bioelectrochemistry*, **2002**, *58*, 119-126.
22. Aoki, H.; Tao, Signal Enhancement for Gene Detection Based on a Redox Reaction of [Fe(CN)<sub>6</sub>]<sup>4-</sup> Mediated by Ferrocene at the Terminal of a Peptide Nucleic Acid as a Probe with Hybridization-amenable Conformational Flexibility. *H. Anal. Sci.* **2008**, *24*, 929-933.
23. Aoki, H.; Tao, Label- and marker-free gene detection based on hybridization-induced conformational flexibility changes in a ferrocene-PNA conjugate probe. *H. Analyst*, **2007**, *132*, 784-791.
24. Kadam, V. Ph.D thesis, Indian Institute Of Science Education and Reserch, Pune. **2016**.
25. Usher, D. A., and McHale, A. H. Hydrolytic stability of helical RNA: A selective advantage for the natural 3',5'-bond. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 1149-1153.
26. Kolasa, K. A., Morrow, J. R., and Sharma, A. P. Trivalent lanthanide ions do not cleave RNA in DNA-RNA hybrids. *Inorg. Chem.* **1993**, *32*, 3983-4.

27. Hüsken, D., Goodall, G., Blommers, M. J. J., Jahnke, W., Hall, J., Häner, R., and Moser, H. E. (1996) Creating RNA Bulges: Cleavage of RNA in RNA/DNA Duplexes by Metal Ion Catalysis. *Biochemistry*, **1996**, 35, 16591–16600.
28. Portmann, S., Grimm, S., Workman, C., Usman, N., and Egli, M. Crystal structures of an A-form duplex with single-adenosine bulges and a conformational basis for site-specific RNA self-cleavage. *Chem. Biol.* **1996**, 3, 173–184.
29. Kaukinen, U.; Bielecki, L., Mikkola, S., Adamiak, R. W., and Lönnberg, H. The cleavage of phosphodiester bonds within small RNA bulges in the presence and absence of metal ion catalysts. *J. Chem. Soc., Perkin Trans.* **2001**, 2, 1024–1031.
30. Markley, J. C.; Godde, F.; Sigurdsson, S. T. Identification and Characterization of a Divalent Metal Ion-Dependent Cleavage Site in the Hammerhead Ribozyme. *Biochemistry*, **2001**, 40, 13849-13856.
31. Borda, E. J.; Markley, J. C.; Sigurdsson, S. T.; Zinc-dependent cleavage in the catalytic core of the hammerhead ribozyme: evidence for a pH-dependent conformational change. *Nucleic Acids Res.*, **2003**, 31, 102595-2600.
32. Husken, D.; Goodall, G.; Blommers, M. J. J.; Jahnke, W.; Hall, J.; Haner, R.; Moser, H. E. Creating RNA Bulges: Cleavage of RNA in RNA/DNA Duplexes by Metal Ion Catalysis. *Biochemistry*, **1996**, 35, 16591-16600.
33. Zagórowska, I.; Kuusela, S.; Lönnberg, H. Metal ion-dependent hydrolysis of RNA phosphodiester bonds within hairpin loops. A comparative kinetic study on chimeric ribo/2'-O-methylribo oligonucleotides. *Nucleic Acids Res.*, **1998**, 26, 3392–3396.
34. Frederiksen, J. K., Fong, R., and Piccirilli, J. A. Metal ions in RNA catalysis. In "Nucleic Acid-Metal Ion Interactions," (N. V. Hud, ed.), Royal Society of Chemistry, Cambridge, UK, **2009**.
35. Forconi, M.; Herschlag, D.; Chapter five. Metal Ion-Based RNA Cleavage as a Structural Probe.

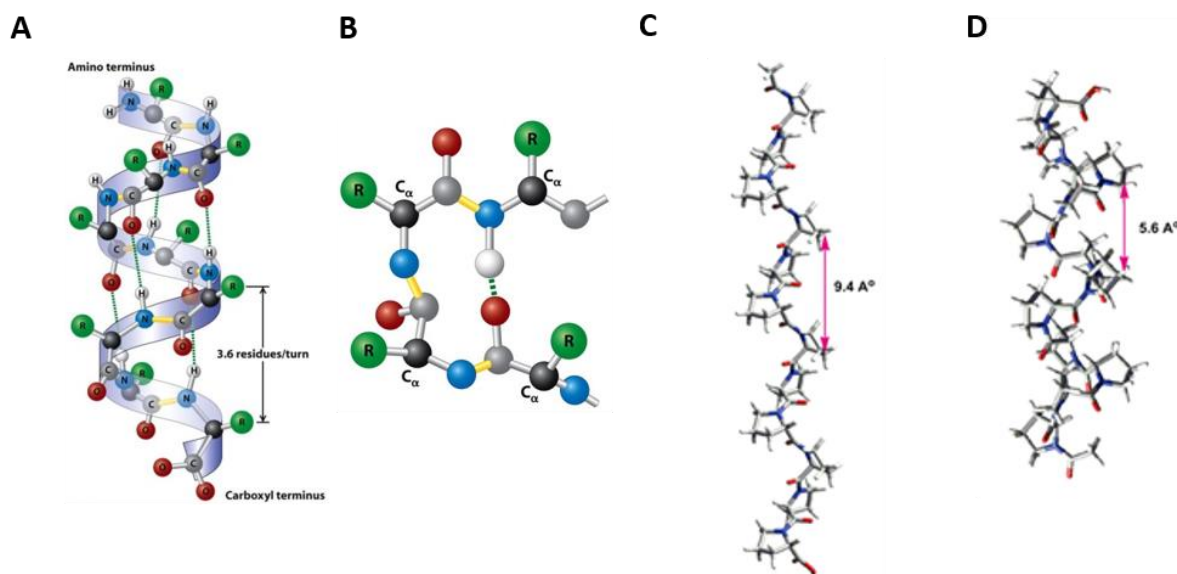
## **Chapter 5**

# **Design and Synthesis of Imidazole Polyproline Peptides Nuclease Mimics**

## 5.1 Introduction

Proteins are important biopolymers made up of repeating units of amino acids. They have various biological functions in all living organism. The type of amino acids present in the protein sequence determines its structure. The overall structure of the protein backbone, and the sidechain of the specific amino acid residues, held in the appropriate positions are responsible for the function of the protein. All proteins are composed of different secondary structures: alpha helix, beta sheet, and polyproline helix etc. The alpha helix (Figure 5.1 A), beta sheet and turns (Figure 5.1 B) arise due to intra and interchain H-bonding from backbone amide groups. In comparison, polypeptides formed from the proline residues lack amide H-bond donor and hence cannot form H-bonds. The observed polyproline peptide helical structures arise entirely due to conformational constrains of Ramachandran dihedral angles ( $\phi$  and  $\Psi$ ). Polyproline helices are divided into two classes, PP-I and PP-II (Figure 5.1 C and D). Out of which, homo-oligomeric polyproline II (PP-II) helix is a common secondary structure that is present in both folded and unfolded proteins.<sup>1</sup> It is semi-extended left-handed helix and has three residues per turn with no intrastrand hydrogen bonds stabilizing the helix (Figure 5.1 C).<sup>2</sup> It is usually seen in peptides containing repeating proline residues that lack H-bond donor in its backbone, but it is also adopted by polypeptides and proteins with hereo amino acid residues as well. In PP-II helix, the peptide bonds are in *trans* conformation, which is the favoured form in aqueous solvents for polyproline peptides. In non-polar alcoholic solvents however, the polyproline peptides adopt an alternate conformation polyproline I (PP-I) form, which is a right handed helix in which the amide bonds are in *cis* conformation.<sup>3</sup> PP-I is more compact than the PP-II helix with 3.3 residues per turn.





**Figure 5.1** Schematic representation of secondary structures: A)  $\alpha$ -helix<sup>4</sup> B)  $\beta$ -turn<sup>4</sup> C) PP-II<sup>5</sup> and D) PP-I<sup>5</sup>

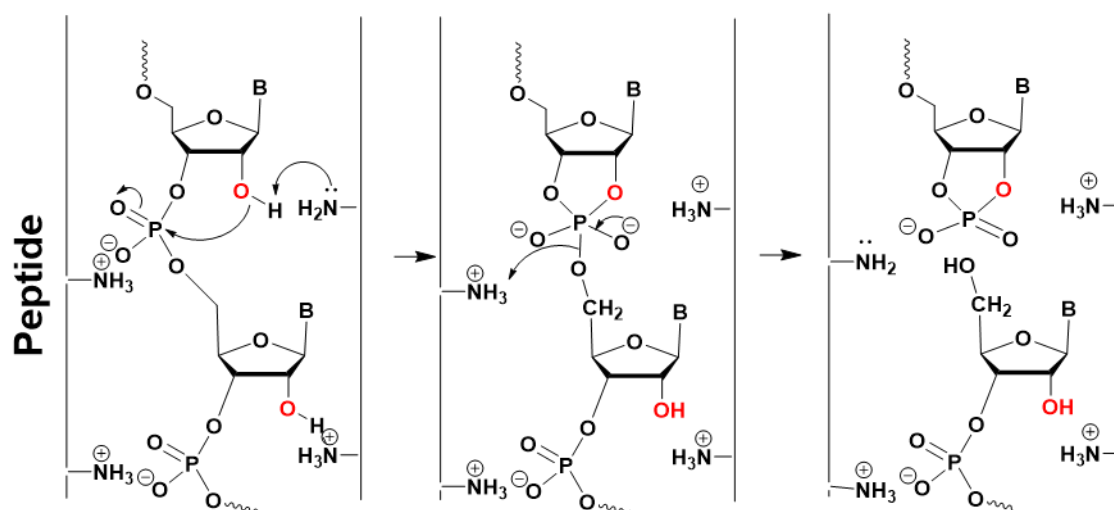
Short synthetic polyproline peptides adopt PP-II and PP-I structures, depending on the nature of solvent. Since polyproline peptides lack NH hydrogen bond donors, they cannot form regular secondary structures like alpha helices or beta structure arising from intra and interchain H-bonding from the peptide backbone. H-bonding substituents are placed at C4 position of proline (C4-NH<sub>2</sub>/OH/SH) may change the nature of secondary structures PPII and PPI. Previous works from this laboratory demonstrated that *cis*-(2*S*,4*S*)-aminopolyproline<sup>3</sup> and *cis*-(2*S*,4*S*)-hydroxypolyproline<sup>4</sup> (4*S*-*amp*<sub>9</sub>/*hyp*<sub>9</sub>) switch to unusual, non-classical  $\beta$ -structures in TFE, while both show PPII form in water.<sup>6-8</sup> Employing such conformational dynamics of 4(OH/NH<sub>2</sub>)-substituted polyprolines for specific functions could lead to multiple applications.

Ribozymes catalyse either the hydrolysis of their own phosphodiester bonds, or the hydrolysis/ligation of phosphodiester of other RNA molecules.<sup>9-11</sup> In different RNA cleaving enzymes have revealed diversity in kinetic behaviour as well as in catalytic mechanism. Many cases these enzyme uses one or two divalent metal ions (e.g. Zn<sup>2+</sup>, Mg<sup>2+</sup>) in the active sites.<sup>12</sup>

To inhibit the expression of any gene at mRNA level by antisense technique, sequence specificity for binding to that gene is important. Researchers have been developing various approaches towards sequence specificity and non-random RNA cleavage.<sup>13-16</sup> Artificial ribonucleases typically consist of two parts, one for sequence specific recognition and another for cleavage of the phosphodiester bonds. Many strategies

use of metal ion chelates to cleave RNA phosphodiester bond *via* transesterification reaction followed by hydrolytic cleavage of produced cyclic-monophosphate.<sup>14,15,17</sup> Another interesting approach is metal ion independent cleavage using purely organic structure for hybridisation and hydrolysis of target RNA.<sup>13,14,18-21</sup>

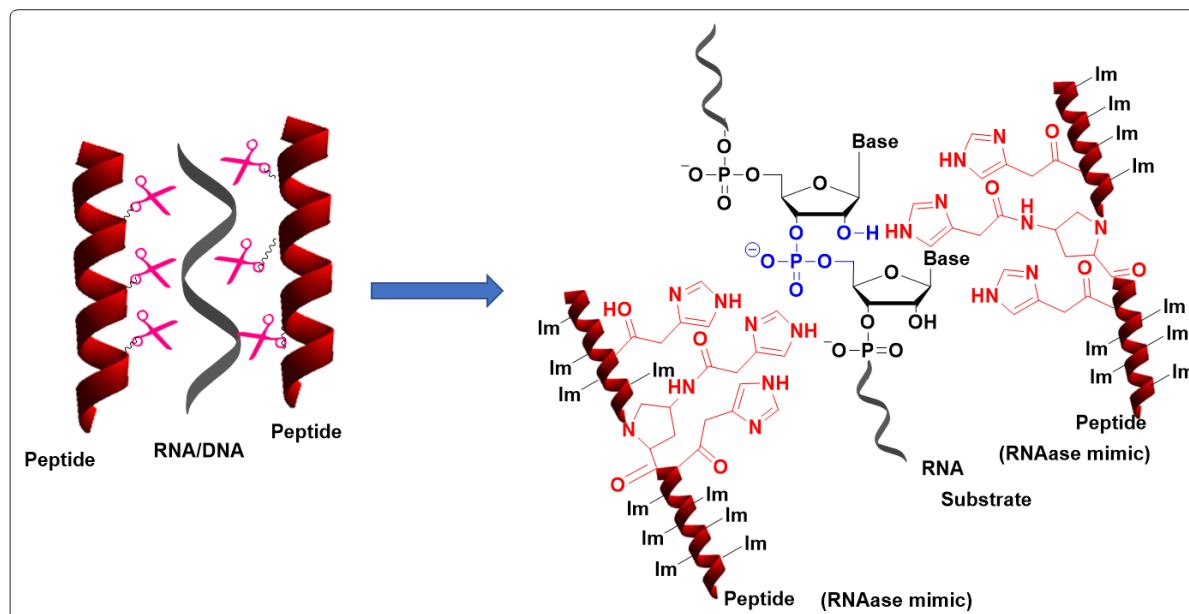
Among the most studied enzymes are the ubiquitous ribonucleases. Synthetic peptides are being designed to mimic enzymes. For instance, polypeptides with alternating hydrophilic and hydrophobic residues (Leu-Lys) have the highest activity RNA cleavage (Figure 5.2).<sup>22-24</sup> Breslow et. al.<sup>25</sup> showed RNA hydrolysis by imidazole buffer to mimic the mechanistic pathway of RNA cleavage using imidazole catalyst. Furthermore, there are reports where polyimidazoles were used to bind to poly(U) RNA and to cleave them.<sup>26</sup> Therefore, it was premised that imidazole conjugated peptides may induce RNA hydrolysis.



**Figure 5.2** Tentative model for the mechanism of hydrolysis of a polyribonucleotide by poly(Lys).

## 5.2 Rational and Objectives

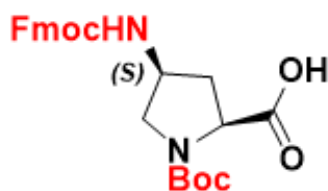
Ribonucleases are involved in hydrolysis of the phosphodiester bonds in RNA. In the enzyme RNase, imidazoles from histidines present at the active site are crucial for hydrolysis. Bringing imidazole units together in a protein like environment may be a useful strategy to mimic the natural RNase environment to design artificial nucleases (Figure 5.3).



**Figure 5.3** Proposed scheme for imidazole polyproline based artificial nuclease.

Small aggregated peptides may mimic the protein environment. In this context, polyproline offers a simple, conformationally tunable and structurally amenable scaffold for the design of peptide based RNase mimics. As demonstrated by previous works in this laboratory and others, functionalization at C4 position of proline can impart solvent dependent conformational properties to the peptides. In continuation with the goal of previous chapters to create artificial ribonucleases, wherein imidazole units were placed on nucleic acid templates, herein, the rationale was to functionalise the C4 position of prolines in polyprolines with imidazole for imparting nucleolytic properties for hydrolysing the bound RNA strands. Such polyfunctionalization would increase the local concentration of imidazole. Based on previous results of conformational studies, polyprolines were functionalized with imidazole at C4 in 4S configuration. Since imidazole is known to complex with metals, the ability of such templated metallo-imidazoles on a polyproline scaffold and studying consequential effect of ribonucleolytic activity would be interesting. The specific objectives of this chapter are

- Synthesis of (2*S*,4*S*)-4-aminoproline monomer for polypeptide synthesis
- Incorporation of imidazole conjugated prolines into polyproline peptides, purification and characterization of imidazole conjugated polyproline peptides
- Metal complexation and biophysical studies of synthesized peptides
- RNA hydrolysis study with imidazole polyproline peptides by HPLC



**Figure 5.4** N-protected (2*S*,4*S*)-4-aminoproline monomer

## 5.3 Results and Discussion

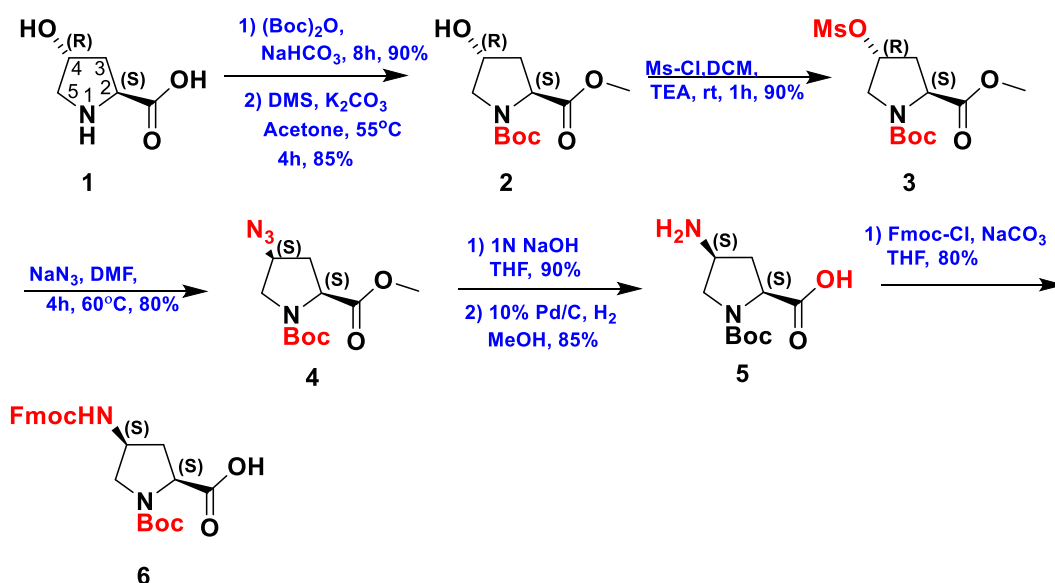
All 4*S*-aminoproline monomers were synthesised by previously reported routes and characterised by various analytical techniques. The 4*S*-aminoproline monomer was incorporated into polyproline peptides using solid phase synthesis and imidazole 4-acetic acid was coupled with 4-amino group of prolines, on resin to get polyproline peptides with imidazole conjugated at C4 of prolyl residues. CD studies were done on imidazole linked amino polyproline peptides. The imidazole polyproline peptide was used for studying RNA hydrolysis.

### 5.3.1 Synthesis of imidazole conjugated 4-aminoproline monomer

This section describes the synthesis of rationally designed 4*S*-aminoproline monomer for imidazole polyproline peptides.

#### 5.3.1a Synthesis of (2*S*,4*S*)-4-aminoproline monomer **6**

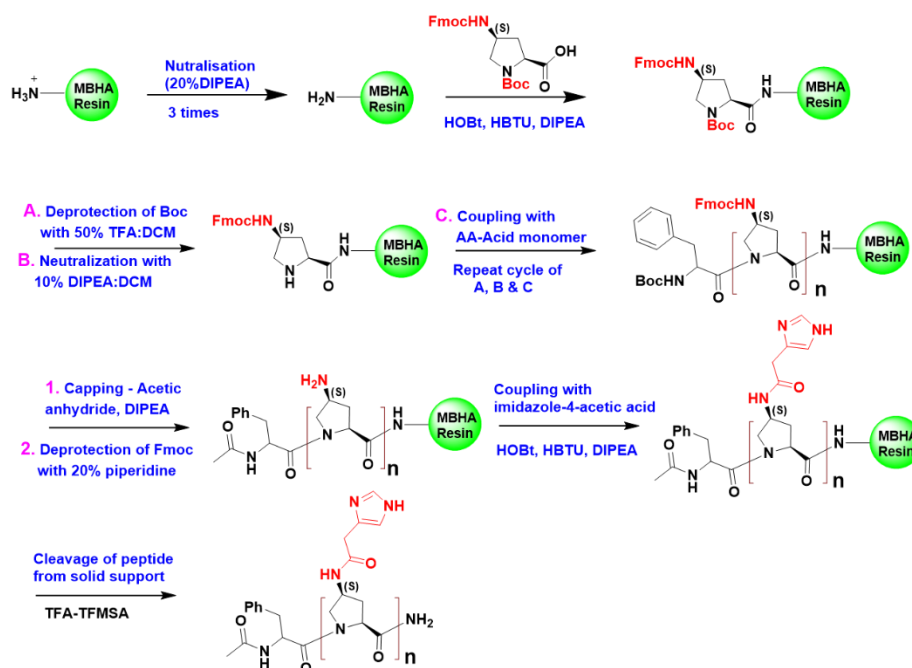
For the synthesis of (2*S*,4*S*)-4-amino proline monomer **6**, 4(*R*)-hydroxyproline **1** was treated with *bis-t*-butoxycarbonyl anhydride in presence of NaOH in water-dioxane mixture to get N<sup>1</sup>-Boc protected 4(*R*)-hydroxyproline. The carboxyl group was *in-situ* protected as methyl ester **2** using dimethyl sulfate and potassium carbonate in anhydrous acetone. The 4*S*-OH group of the methyl ester **2** was reacted with methanesulfonylchloride in DCM in the presence of Et<sub>3</sub>N to obtain the corresponding 4*R*-mesyl derivative **3** (Scheme 2.2). Treatment of **3** with NaN<sub>3</sub> at 55 °C in DMF resulted in an S<sub>N</sub>2 displacement of 4(*S*)-O-mesyl group to yield the 4(*S*)-azide **4**. The methyl ester of compound **4** was hydrolysed with aq. NaOH to yield (2*S*,4*S*)-azido acid and subsequent catalytic hydrogenation of the 4(*S*)-azide yielded the corresponding 4(*S*)-amine **5**. This was treated with 9-fluorenylmethyl chloroformate in THF:water in the presence of Na<sub>2</sub>CO<sub>3</sub> to yield the fully protected (2*S*,4*S*)-4-aminoproline monomer **6**.



**Scheme 5.1** Synthesis of (2*S*,4*S*) aminoproline monomer **6**

### 5.3.1b Synthesis of imidazole polyprolyl peptide

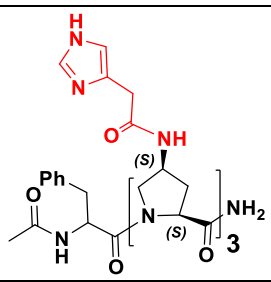
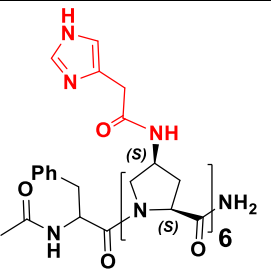
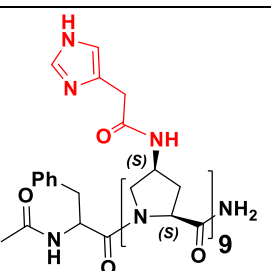
4(*S*)-Imidazolyl polyprolyl peptides [4(*S*)-*Imap<sub>n</sub>*] **P1**, **P2** and **P3** corresponding to different lengths were assembled by solid phase peptide synthesis protocol as shown in Figure 5.5. For this orthogonally protected monomer (2*S*,4*S*)-*N*<sup>1</sup>-(Boc)-*N*<sup>4</sup>-(Fmoc)-aminoproline monomer **6** was synthesized as described in Scheme 5.1. MBHA resin was used for the synthesis of the peptides and a modified *t*-Boc strategy was employed.



**Figure 5.5** Solid phase polyproline peptide synthesis protocol by *t*-Boc strategy

The protocol followed for the synthesis of imidazole polyproline peptides **P1-P3** was similar to the ones used for the linear PNA as described in Chapter 2, except for a few steps for the conjugation of imidazole on polyproline peptide. The commercially available MBHA resin was allowed to swell and was treated with 20% DIPEA (in DCM). A lower loading of the resin enhances the efficacy of coupling. Hence, the loading capacity of the resin was reduced to 0.3 mmol/g by partial coupling of the first proline residue and capping all the unreacted free amines on the resin with acetic anhydride. After incorporation of proline residues and repeating the cycle of coupling, deprotection and neutralization steps for the desired number of times with 4-substituted prolyl polypeptides of different lengths with their sidechains protected with Fmoc were obtained. As the last residue, the amino acid phenylalanine was coupled and the peptides were capped with acetic anhydride. This amino acid is useful for determining the peptide concentration, Finally, all N<sup>4</sup>-Fmoc protecting groups were deprotected with 20% of piperidine and the liberated 4-amino groups were coupled on resin with imidazole-4-acetic acid. The peptides were cleaved from the resin using TFMSA/TFA in the presence of scavengers, ethanedithol and thioanisole. The N-capped peptides were thus obtained as C-terminal amides.

**Table 5.1:** HPLC retention time and MALDI-TOF mass spectral analysis of the peptides

Sequence	Molecular Formula	Ret. time	Cal. Mass	Obs. Mass
<b>P1</b> 	C <sub>41</sub> H <sub>50</sub> N <sub>14</sub> O <sub>8</sub>	11.8	866.39	867.24 [M+H] <sup>+</sup>
<b>P2</b> 	C <sub>71</sub> H <sub>86</sub> N <sub>26</sub> O <sub>14</sub>	11.9	1527.60	1549.48 [M+Na] <sup>+</sup>
<b>P3</b> 	C <sub>101</sub> H <sub>122</sub> N <sub>38</sub> O <sub>20</sub>	13.7	2188.33	2188.49 [M] <sup>+</sup>

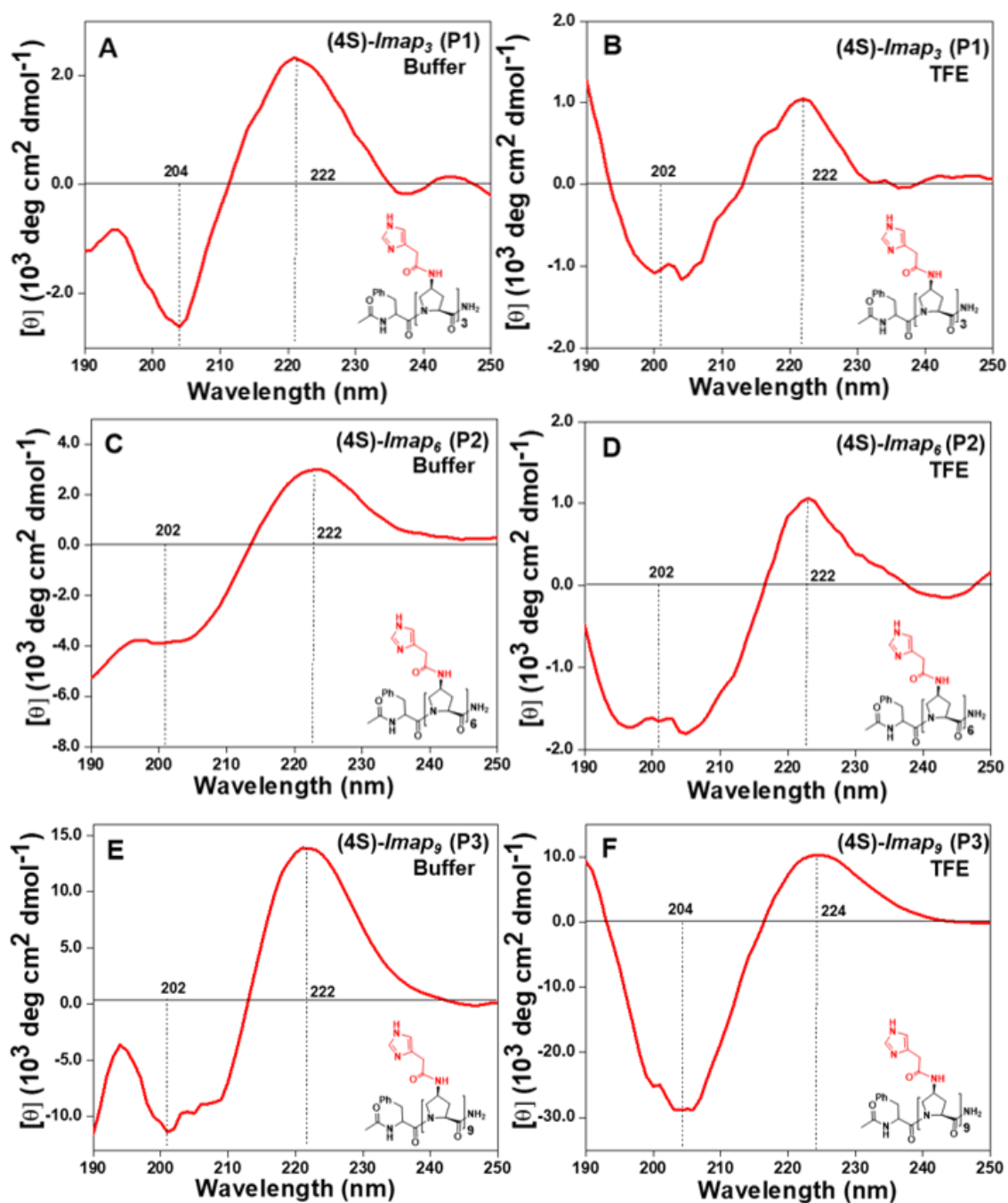
### 5.3.1c *Purification and characterization of peptides*

The peptides were purified using the standard reverse phase HPLC using a semipreparative C18 column. A linear gradient system between water and acetonitrile was used. The purity of the peptides were finally checked on C18 analytical column. The integrity of the purified peptides was confirmed by MALDI-TOF spectrometry using DHB as matrix. The HPLC retention time, calculated and observed mass of the peptides are listed in the Table 5.1. The HPLC chromatograms and MALDI-TOF mass spectra are shown in Appendix.

### 5.3.1d *Effect of different solvent and length on imidazole polyprolyl peptide*

The 4-NH<sub>2</sub>/OH prolyl polypeptides have been shown to have interesting solvent dependent conformational properties dictated by the *R/S*-configuration at C4.<sup>6,7</sup> It is known that the functionalization of 4(*S*)-NH<sub>2</sub>/OH with acetyl group failed to show structural conversion.<sup>8</sup> In this present study conformational properties of synthesised analogues of 4-N-acetyl (C4-Imidazole) prolyl polypeptide **P1**, **P2** and **P3** were studied by CD spectroscopy in different solvents and length of peptides.

The imidazole conjugated polyprolyl peptides **P1**, **P2** and **P3** showed a typical PP-II CD spectra, in buffer and TFE. All the imidazole polyprolyl peptides **P1**, **P2** and **P3** showed a positive band in the 220-225 nm wavelength while a negative band in the 200-205 nm wavelength (Figure 5.6 A-F). The negative and positive band appeared in the same region in all the peptides in both buffer and TFE.



**Figure 5.6** Solvent and length dependent CD spectra of peptides: A) (4S)-*Imap*<sub>3</sub> (P1) in 10 mM phosphate buffer B) in TFE C) (4S)-*Imap*<sub>6</sub> (P2) in phosphate buffer D) in TFE and E) (4S)-*Imap*<sub>9</sub> (P3) in 10 mM phosphate buffer F) in TFE.

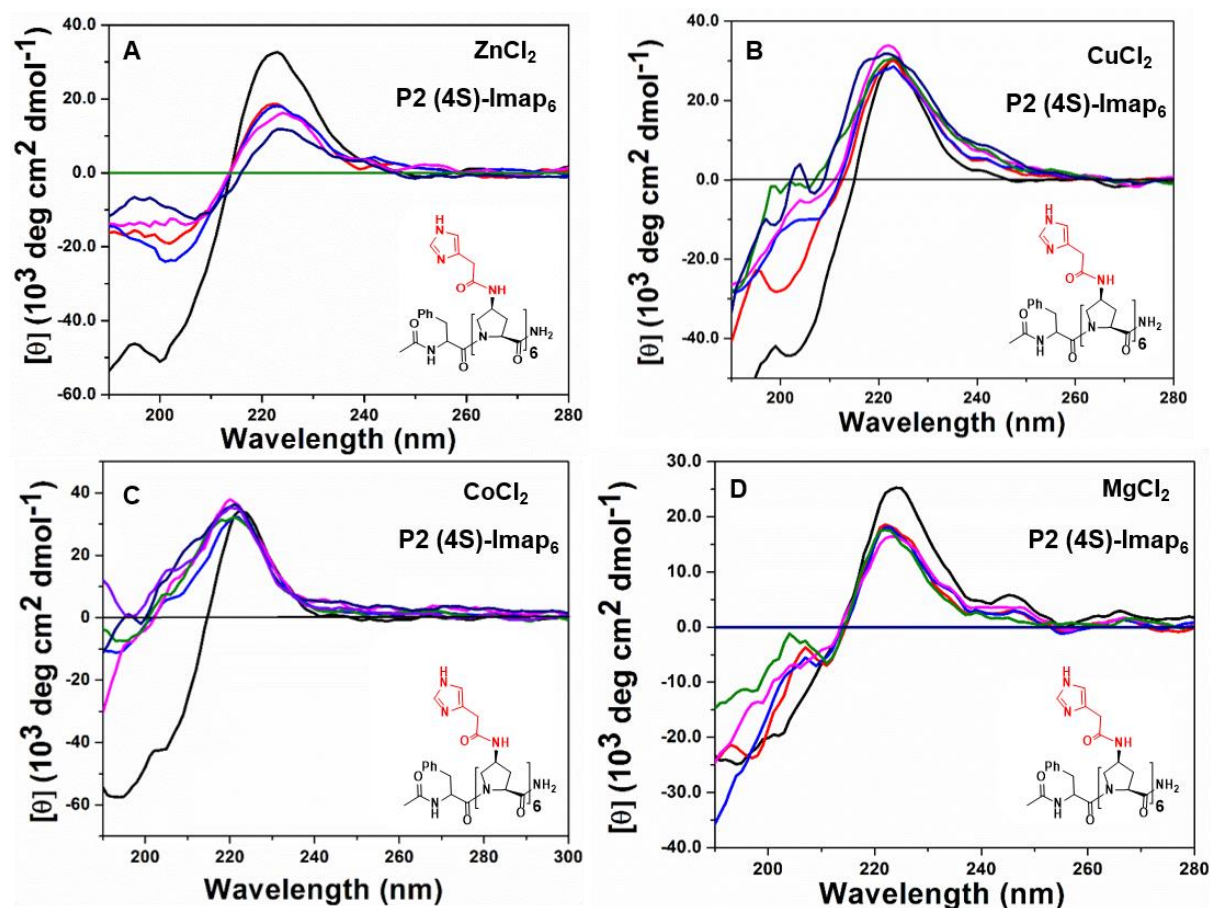
It was observed that magnitudes of positive and negative bands depend upon the length of the peptides. In case of (4S)-*Imap*<sub>3</sub> (P1) and (4S)-*Imap*<sub>6</sub> (P2) though length increases, the magnitude of positive band was almost same (Figure 5.6 A, B, C and D) whereas for (4S)-*Imap*<sub>9</sub> (P3) peptide, intensity of positive band increased a significantly (Figure 5.6 E and F) in buffer and TFE. The intensity of negative band decreased in buffer and TFE.



These results suggested that even after incorporation of imidazole side chain at C4 on (4*S*)-*amp*<sub>n</sub> peptides, it retained the PP-II structure and no solvent dependent conformational change was observed. The higher magnitudes of positive band at 225 nm, both in buffer and in TFE, in peptides **P1**, **P2** and **P3** indicated that PP-II conformation is more ordered and stable even in **P1** peptide. The effect of length of imidazole conjugated proline peptides **P1-P3** lead to increase in the content of PP-II in buffer and TFE.

### 5.3.1e Conformational Studies of 4-*N*-acetyl (C4-Imidazole) prolyl polypeptide **P2** in presence of metal ions

The 4*S*-amidoacetyl (C4-imidazole) polypeptide **P2** was titrated individually with metal salts of ZnCl<sub>2</sub>, CuCl<sub>2</sub>, CoCl<sub>2</sub> and MgCl<sub>2</sub>. and accompanying changes in conformation were followed by CD spectroscopy (Figure 5.5).



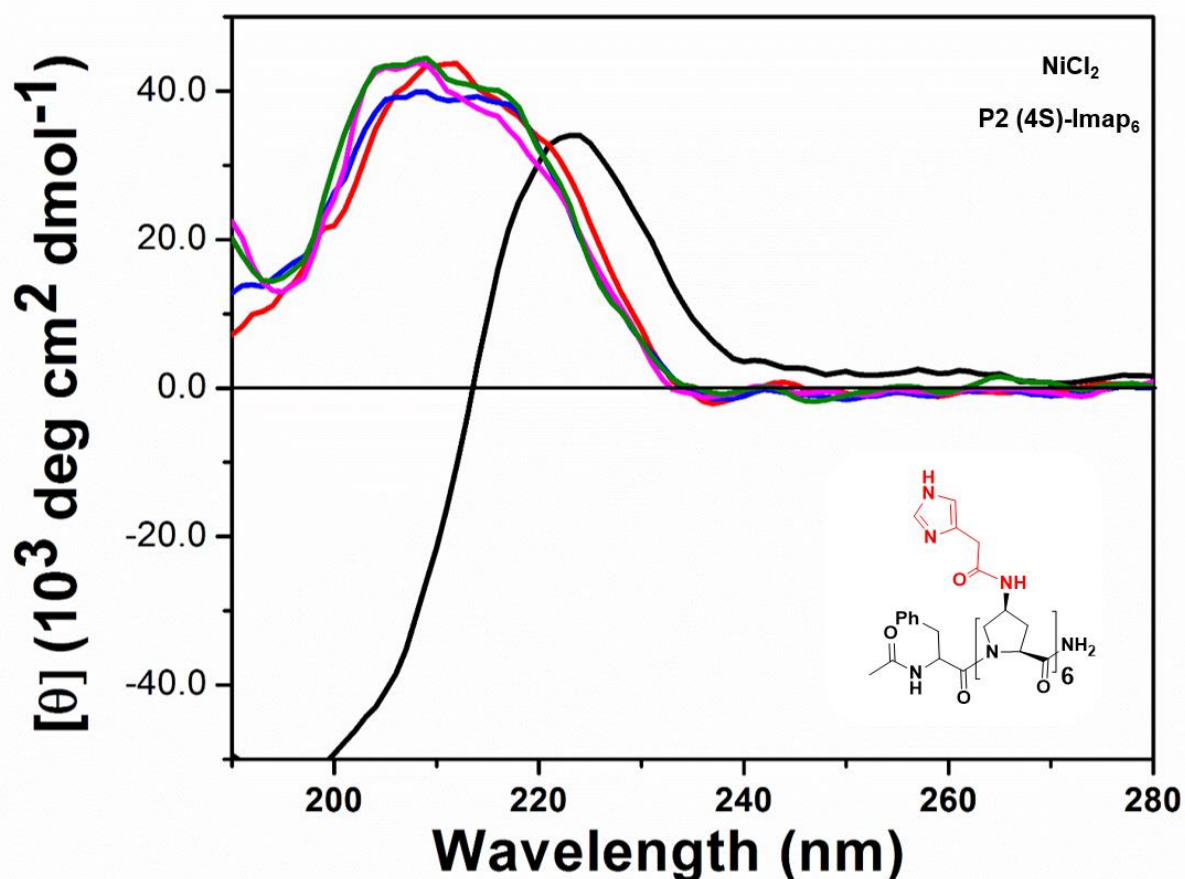
**Figure 5.5** Effect of metals on conformational of imidazole peptide (**P2**) in water. A) CD titration of ZnCl<sub>2</sub> in water of peptide **P2** (4*S*)-*Imap*<sub>6</sub>; B) CuCl<sub>2</sub>; C) CoCl<sub>2</sub>; D) MgCl<sub>2</sub>.

In water, the peptide **P2** showed typical PP-II like CD spectra in presence of ZnCl<sub>2</sub>, (Figure 5.5 A) while with other metal ions (CuCl<sub>2</sub>, CoCl<sub>2</sub> and MgCl<sub>2</sub>), the negative band in

200-205 nm region started collapsing, without much change in the positive band at 220 nm (Figure 5.5 B, C and D, either in its intensity or position. These results grossly suggested that the metal ions studied here did not change the conformational preference of (4*S*)-*Imap*<sub>6</sub> **P2** from PP-II to other conformation at least in presence of ZnCl<sub>2</sub>. However, with other metal salts, the loss of signals at the 205 nm made it difficult to identify whether the peptide conformation is still intact.

### 5.3.1f Conformational Studies of 4*S*-amido-(Imidazole-C4-acetyl) prolyl polypeptide **P2** with NiCl<sub>2</sub>

The imidazole polypeptide **P2** with repeating units of 4(*S*)-amido-(imidazole-C4-acetyl) prolyl polypeptide was titrated with metal NiCl<sub>2</sub>. The C4(*S*)-imidazole conjugated amino prolyl peptide **P2** showed PP-II conformation.



**Figure 5.6** Effect of metals on conformational of imidazole peptide (**P2**) in water. CD titration of NiCl<sub>2</sub> in water of peptide **P2** (4*S*)-*Imap*<sub>6</sub> Inset: CD spectra changes conformation.

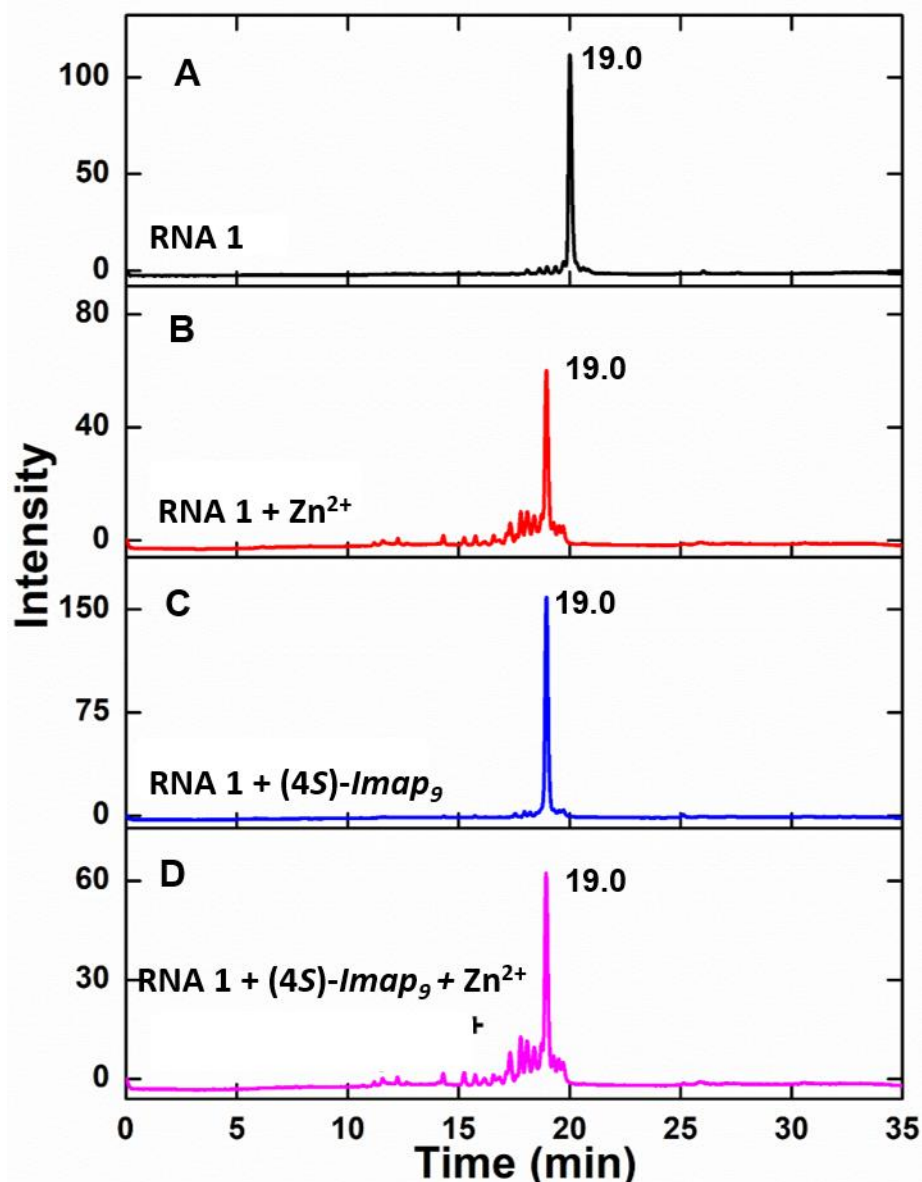
Addition of NiCl<sub>2</sub> caused a loss of typical peptide bands, positive  $n \rightarrow \pi^*$  band in the 220-225 nm region and negative  $\pi \rightarrow \pi^*$  band at 205 nm were disappeared (Figure 5.6).

These results suggested that peptide **P2** changed the conformation from PP-II to a different unidentified state in the presence of  $\text{NiCl}_2$ .

### **5.3.1g RNA cleavage study with imidazole polyprolyl peptide and with metal.**

To study cleavage of **RNA 1**, one equimolar imidazole modified (4S)-*Imap<sub>9</sub>* (**P3**) was hybridised with one equimolar **RNA1** (1:1), which will enormously increase the local concentration of imidazole to act almost like imidazole buffer to cleave the **RNA 1**. All samples prepared in 10 mM HEPES buffer, pH 7.2 containing 10 mM NaCl with 300  $\mu\text{M}$   $\text{ZnCl}_2$  without metal ion. All fractions of reaction mixture were incubated separately at 37 °C. At regular time intervals, small fractions of aliquots were removed and subjected for RP-HPLC analysis using triethyl amine acetate buffer (pH 7.2) as the mobile phase on C18 column.

The peptide **P3** bearing nine imidazoles could not be eluted out with the solvent system (TEAA buffer) as it strongly bound to the HPLC column; however, **RNA 1** was eluted at 19.0 min. The stability of **RNA 1** with  $\text{ZnCl}_2$  for 24 h was checked and it did not show any degradation (Figure 5.7 B). The peptide **P3** was used for cleavage of **RNA 1** by reaction at 37 °C for 24 h to determine its catalytic activity without metal. The reaction mixture analysed by RP-HPLC at 24 h showed only peak at 19.0 min (Figure 5.7 C) corresponding to **RNA 1** indicating that **P3** also failed to hydrolyse the **RNA 1** in the absence of metal.



**Figure 5.7** (A-D) Cleavage study of **RNA 1** by *(4S)-Imap<sub>9</sub>* with metals and without metal (10 mM HEPES buffer, pH 7.4, 10 mM NaCl, 37 °C, 24 h, Zn<sup>2+</sup> 300 μM). A) **RNA 1** (10 μM); B) **RNA 1** + Zn<sup>2+</sup>; C) **RNA 1** + *(4S)-Imap<sub>9</sub>*; D) **RNA 1** + *(4S)-Imap<sub>9</sub>* + Zn<sup>2+</sup>.

In presence of ZnCl<sub>2</sub>, **RNA 1** was incubated with peptide **P3** at 37 °C for 24 h. The reaction mixture analysed by RP-HPLC at 24 h showed additional minor peaks at 16-18 min (Figure 5.7 D) suggesting small amount of cleaved RNA by peptide **P3**, followed by a major peak at 19.0 min corresponding to intact **RNA 1**. Thus *(4S)-Imap<sub>9</sub>* peptide **P3** cleaved **RNA 1** by <10% in presence of metal.

*4(S)* substituted polyprolyl peptide backbone is rigid than the unsubstituted polyprolyl peptide, so it can't reach out to phosphodiester target for effective cleave of RNA.

## 5.4 Summary

In summary, the synthesis of (2*S*,4*S*) aminoproline monomer and its incorporation into polyproline peptide by solid phase peptide synthesis protocol were accomplished successfully. Further, the 4*S*-aminopropylproline peptides were coupled to imidazole-4-acetic acid using solid phase synthesis, cleaved from resin, purified by HPLC and characterised by MALDI TOF.

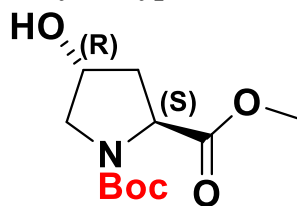
The conformation of polypeptides derived from monomers in which imidazole was conjugated to 4-amino prolines were studied by CD. This suggested that they retained PP-II conformation in both buffer and TFE. The **P1**, **P2** and **P3** peptides of increasing lengths exhibited slight increases in magnitude of positive CD bands in buffer and TFE. The titration of imidazolyl polyproline peptide with ZnCl<sub>2</sub>, MgCl<sub>2</sub>, CoCl<sub>2</sub> and CuCl<sub>2</sub> solutions showed no changes in PP-II conformation. When the imidazolyl polyproline peptide was titrated with NiCl<sub>2</sub>, the CD signature was found to be drastically altered. This was attributed to the formation of Ni-complex with the imidazolyl moiety which is well known with polyhistidines and did not match with any of the known secondary structures.

Further, the imidazolyl-4-aminoproline peptides were studied for the RNA hydrolysis in presence and absence of metal; however, they did not cause any significant degradation (observed degradation less than 10%) of **RNA 1** in the absence or in presence of Zn salt.

## 2.5 Experimental section

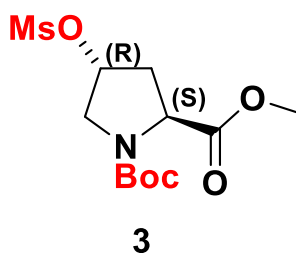
General methods of peptide synthesis, purification of peptides by HPLC, characterization using MALDI-TOF, conformational analysis by CD spectroscopy and preparation of samples for AFM and FE0SEM are described in Chapter 2.

### 2.5.1 Synthesis of compounds 2-6

**(2*S*, 4*R*)-N1-(*t*-butoxycarbonyl)-4-hydroxyproline methylester (2)****2**

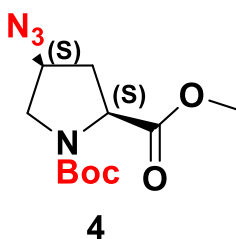
A solution of *trans*-4*R*-hydroxyproline (1 g, 7.6 mmol) in a solution of aq. NaOH (2*N*, 10 mL) and dioxane (10 mL) was cooled to 0° C and Boc-anhydride (2 ml, 8.4 mmol) was added dropwise using addition funnel. The reaction mixture was stirred at 0 C for 1 h, and dioxane was removed under reduced pressure. The residue was extracted with diethyl ether (5 x 25 mL) to remove unreacted (Boc)<sub>2</sub>O. The aqueous layer was then vigorously stirred by ethyl acetate under ice-cold condition and acidified with saturated KHSO<sub>4</sub> solution to pH 3. The ethyl acetate layer was separated and the aqueous layers were further extracted into ethyl acetate (3 x 20 mL). The organic layer were pooled and washed with water followed by brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. Removal of ethyl acetate under reduced pressure, gave a white solid which was recrystallized with EtOAc/Hexane. The white solid compound (1 g, 4.3 mmol) obtained in the above step was mixed with anhydrous K<sub>2</sub>CO<sub>3</sub> (1.8 g, 13 mmol) and dissolved in anhydrous acetone (25 mL). The reaction mixture was stirred for 30 min at room temperature and dimethylsulphate (4.9 ml, 5.18 mmol) was added. Stirring was continued under reflux conditions for 5 h, when TLC confirmed maximum product formation. The solvent was evaporated under reduced pressure and resulting solid was dissolved in ethyl acetate (50 ml), washed with water (2 X 30 ml), followed by saturated brine solution and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. The ethyl acetate extract was concentrated under reduced pressure to a pale yellow solid which was then purified by silica gel chromatography, eluting with ethylacetate:hexane (1:1) to yield compound **2** as white crystalline solid (1.54 g, 82%) over the two steps. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: H 4.48 – 4.31 (m, 2H), 3.69 (s, 3H), 3.54 (m, 2H), 2.25 (m, 1H), 2.07 – 1.96 (m, 1H), 1.39 (s, 9H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 173.8, 154.1, 80.5, 80.4, 70.1, 69.4, 58.0, 57.6, 54.7, 52.4, 52.2, 39.1, 38.5, 28.4, 28.3 ppm. HRMS (ESI-TOF) *m/z* calcd for C<sub>11</sub>H<sub>19</sub>NO<sub>5</sub> [M+Na]<sup>+</sup> 268.1160, found 268.1166 [M+Na]<sup>+</sup>

**2*S*,4*R*)-N-(*t*-butyloxycabonyl)-4-*O*-mesyl proline methylester (3)**

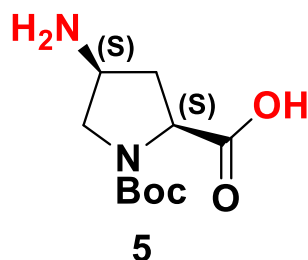


The solution of methylester **2** (1 g, 4.1 mmol) and triethylamine 1.7 mL (12.2 mmol) in 20 mL of dry dichloromethane was cooled to 0 ° C on ice bath under Argon. While stirring (4.7 mL, 6.1 mmol) methanesulfonyl chloride was added in one shot. Stirring was continued over a period of 3 hrs at 0 ° C. The reaction mixture was then washed with water, followed by saturated brine solution. The organic layer was dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude material was purified by silica gel chromatography (30% ethyl acetate/hexane) which afforded mesylated compound **3** as a white solid (1.26 g 95%). HRMS (ESI-TOF) m/z calcd for C<sub>12</sub>H<sub>21</sub>NO<sub>7</sub>S [M+Na]<sup>+</sup> 346.0936, found 346.0939 [M+Na]<sup>+</sup>

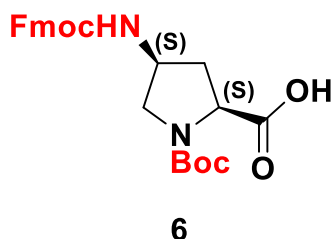
**(2S,4S)-N1-(t-butyloxycarbonyl)-4-azidoproline methyl ester (4)**



A solution of compound **3**, (1 g, 2.5 mmol) and NaN<sub>3</sub>, (1.3 g, 20 mmol) in dry DMF (10 mL) was stirred at 55-60 ° C for 12 hrs under nitrogen. When TLC indicated complete conversion of the starting material, DMF was removed under vacuum and the residue was dissolved in water. The aqueous layer was extracted with ethyl acetate (3 x 25 mL). The combined organic layer was washed with water followed by brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude product obtained was purified by silica gel chromatography (40% ethyl acetate/hexane elute) to afford compound **4** as colorless thick oil. (0.65 g 90%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: H 4.34 (m, 1H), 4.17 – 4.05 (m, 1H), 3.79 – 3.60 (m, 4H), 3.44 (m, 1H), 2.54 – 2.35 (m, 1H), 2.13 (m, 1H), 1.44 (min) 1.38 (maj) (s, 9H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 172.3, 172.0, 154.0, 153.5, 80.6, 59.3, 58.3, 57.8, 57.4, 52.5, 52.4, 51.3, 50.9, 36.1, 35.2, 28.4, 28.3 ppm. HRMS (ESI-TOF) m/z calcd for C<sub>11</sub>H<sub>18</sub>N<sub>4</sub>O<sub>4</sub> [M+Na]<sup>+</sup> 293.1225, found 293.1227 [M+Na]<sup>+</sup>

**(2*S*,4*S*)N1-(*t*-butyloxycabonyl)-4-aminoproline (5)**

The methyl ester **4** (1 g, 3.7 mmol) was hydrolysed using aqueous NaOH (2N) in water MeOH (1:1, 10 ml) mixture. The reaction mixture was stirred in RT for 2 hrs and methanol was removed under reduced pressure. The aqueous layer was acidified with saturated KHSO<sub>4</sub> solution, extracted with ethyl acetate and dried over Na<sub>2</sub>SO<sub>4</sub>. The free acid intermediate was obtained as a dark brown liquid upon concentration. The free acid intermediate (1 g, 3.9 mmol) was taken in methanol (10 mL) and to this was added 10% Pd-C (0.15 g). The mixture was subjected to hydrogenation for 4 hrs. Pd-C suspension was removed by filtration through Celite and the filtrate was evaporated under reduced pressure to give reddish brown liquid. The (2*S*,4*R*)N1-(*t*-butoxycarbonyl)-4-aminoproline **5** thus obtained (0.88 g, 98% yield) was used without further purification.

**(2*S*,4*S*)N1-(*t*-butoxycarbonyl)-4-(fluorenylmethyloxycarbonylamino) proline (6)**

An aqueous solution (100 mL) of compound **5** (1 g, 4.3 mmol) was cooled to 0 °C and 10% aqueous Na<sub>2</sub>CO<sub>3</sub> (20 mL) and dioxane (100 mL) were added. Fmoc-Cl (1.7 g, 6.5 mmol) dissolved in dioxane (10 mL) was added dropwise during 1 hr. The resulting mixture was stirred for 15 hrs at room temperature maintaining the pH of solution around 8.0. The reaction mixture was concentrated on a rotary evaporator to remove dioxane and to the resulting slurry 50 mL of water was added. The mixture was extracted with diethyl ether (5 x 50 mL) to remove unreacted Fmoc-Cl. The aqueous layer was then vigorously stirred by ethyl acetate under ice-cold condition, and acidified with saturated KHSO<sub>4</sub> solution to pH 3. The ethyl acetate layer was separated and the aqueous layer was further extracted into



fresh ethyl acetate (3 x 50 mL), the organic layers were pooled and washed with water followed by brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude material was purified by silica gel chromatography (80% ethyl acetate/hexane elute) afford compound **6** as white solid. Yield 1.28 g (55%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ: H 7.74 (d, 2H, J= 8 Hz), 7.58 (d, 2H, J= 8 Hz), 7.34 (t, 2H, J= 16 Hz), 7.26 (t, 2H, J= 16 Hz), 4.29 (d, 2H, J= 8 Hz), 4.15 (m, 3H), 3.73 (m, 1H), 3.19 (m, 1H), 2.64 – 2.45 (m, 1H), 1.93 – 1.79 (m, 1H), 1.42 (min) 1.39 (maj) (2s, 9H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 174.0, 157.2, 156.0, 143.9, 141.3, 127.8, 127.2, 125.3, 120.1, 82.9, 67.2, 66.0, 58.7, 54.0, 50.7, 47.2, 33.2, 28.4, 15.3 ppm. HRMS (ESI-TOF) m/z calcd for C<sub>25</sub>H<sub>28</sub>N<sub>2</sub>O<sub>6</sub> [M+H]<sup>+</sup> 453.2025, found 453.2032 [M+H]<sup>+</sup>

## 5.6 References

1. Shi, Z.; Woody, R. W.; Kallenbach, N. R. Is polyproline II a major backbone conformation in unfolded proteins? *Adv. Protein Chem.* **2002**, *62*, 163-240. b) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. Conformation of the Backbone in Unfolded Proteins. *Chem. Rev.* **2006**, *106*, 1877–1897.
2. Cubellis, M. V.; Caillez, F.; Blundell, T. L.; Lovell, S. C. Properties of Polyproline II, a Secondary Structure Element Implicated in Protein–Protein Interactions. *Proteins: Struct., Funct., Bioinf.* **2005**, *58*, 880–892.
3. (a) Zhong, H.; Carlson, H. Conformational Studies of Polyprolines. *J. Chem. Theory Comput.* **2006**, *2*, 342-353.4. (b) Lodish, A.; Berk, A.; Kaiser, C. A.; Krieger, M.; Scott, M. P.; Bretscher, A.; Ploegh, H.; Matsudaira, P. Protein Structure and Function. *Molecular Cell Biology*, 5<sup>th</sup> Ed; W. H. Freeman: New York, 2008; pp 63-110.
5. Kuemin, M.; Schweizer, S.; Ochsenfeld, C.; Wennemers, H. Effects of Terminal Functional Groups on the Stability of the Polyproline II Structure: A Combined Experimental and Theoretical Study. *J. Am. Chem. Soc.* **2009**, *131*, 15474–15482
6. Mahesh V. S.; Ganesh K. N. Water-Induced Switching of β -Structure to Polyproline II Conformation in the 4S-Aminoproline Polypeptide via H-Bond Rearrangement. *Org. Lett.*, **2010**, *12*, 5390-5393.
7. Nitin D. B.; Mahesh V. S.; Ganesh K. N. A nanofiber assembly directed by the non-classical antiparallel β-structure from 4S-(OH) proline polypeptide. *Chem. Commun.*, **2016**, *52*, 4884-4887

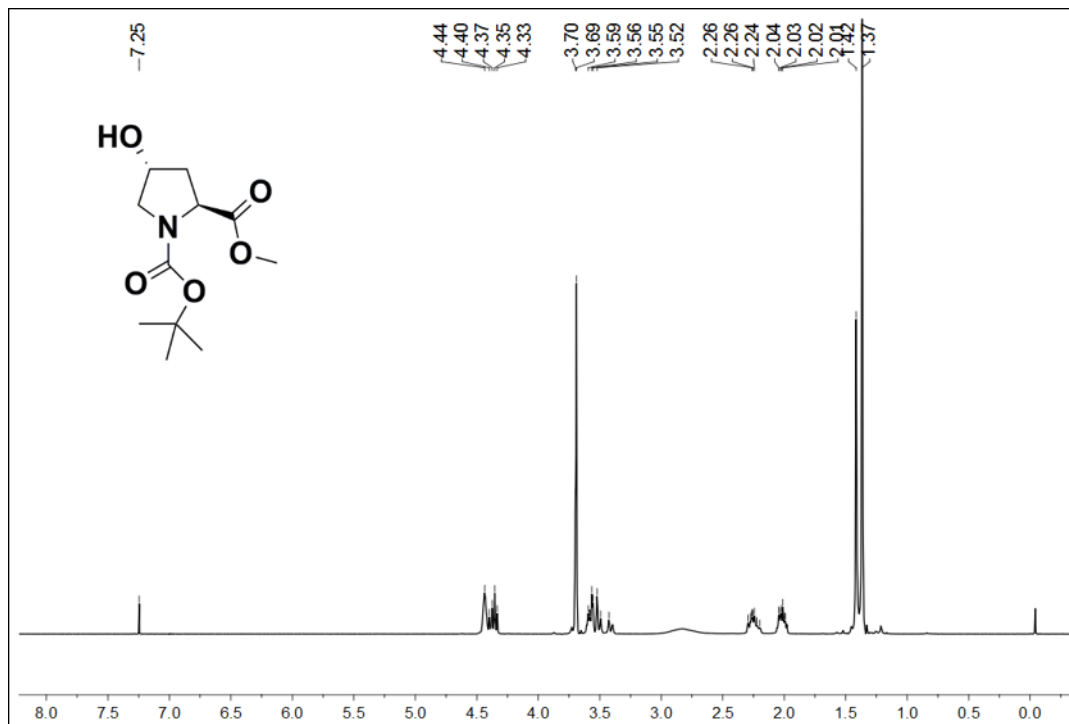
8. Bansode N.D.; Madhanagopal B.; Sonar M.V.; Ganesh K.N. Stereodependent and solvent-specific formation of unusual  $\beta$ -structure through side chain-backbone H-bonding in C4(S)-(NH<sub>2</sub>/OH/NHCHO)-L-prolyl polypeptides. **2017**, *108*, 1-11
9. Cech, T. R., Zaug, A. J., and Grabowski, P. J. In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence, *Cell*, **1981**, *27*, 487-496.
10. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell*, *35*, 849-857.
11. Takagi, Y., Ikeda, Y. and Taira, K. Ribozyme mechanisms, *Top. Curr. Chem.* **2004**, *232*, 213-251
12. Koroleva, L. S., Serpokrylova, I. Y., Vlassov, V. V., and Silnikov, V. N. (2007) Design and synthesis of metal-free artificial ribonucleases, *Protein and peptide letters* *14*, 151-163.
13. Weston, J. Mode of action of bi- and trinuclear zinc hydrolases and their synthetic Analogues. *Chem. Rev.* **2005**, *105*, 2151-2174.
14. Niittymaki, T., and Lonnberg, H. Artificial ribonucleases, *Org. biomol. Chem.* **2006**, *4*, 15-25.
15. Komiyama, M., Sumaoka, J., Kuzuya, A., and Yamamoto, Y. Sequence-selective artificial ribonucleases, *Methods in enzymol.* **2001**, *341*, 455-468.
16. Kuzuya, A., and Komiyama, M. Site-Selective Artificial Ribonucleases and their Applications. *Curr. Org. Chem.* **2007**, *11*, 1450-1459.
17. Mancin, F., Scrimin, P., Tecilla, P., and Tonellato, U. Artificial metallonucleases, *Chemcomm.* **2005**, 2540-2548.
18. Gnaccarini, C., Peter, S., Scheffer, U., Vonhoff, S., Klussmann, S., and Gobel, M. W. 2006 Site-specific cleavage of RNA by a metal-free artificial nuclease attached to antisense oligonucleotides, *J. Am. Chem. Soc.* **2006**, *128*, 8063-8067.
19. Riguet, E., Tripathi, S., Chaubey, B., Desire, J., Pandey, V. N., and Decout, J. L. A peptide nucleic acid-neamine conjugate that targets and cleaves HIV-1 TAR RNA inhibits viral replication, *J. Med. Chem.*, **2004**, *47*, 4806-4809.
20. Komiyama, M., and Inokawa, T. Selective hydrolysis of tRNA by ethylenediamine bound to a DNA oligomer, *J. Biochem.*, **1994**, *116*, 719-720.
21. Komiyama, M., and Yoshinari, K. Kinetic Analysis of Diamine-Catalyzed RNA Hydrolysis, *J. Org. Chem.*, **1997**, *62*, 2155-2160.

22. Antonini, E.; Brunori, M.; Colosimo, A.; Kuiper, H. A.; Zolla, L. Kinetic and thermodynamic parameters for oxygen binding to the allosteric states of *panulirus interruptus* hemocyanin. *Biophys. Chem.*, **1983**, *18*, 117-124.
23. Klarman, A.; Daniel, E. Oxygen binding properties of stripped (calcium ion and magnesium ion free) hemocyanin from the scorpion *Leirus quinquestriatus*. *Biochemistry* **1980**, *19*, 5176-5180.
24. Er-el, Z.; Shaklai, N.; Daniel, E. Oxygen binding properties of haemocyanin from Levantina hierosolima. *J. Mol. Biol.*, **1972**, *64*, 341-352.
25. Ronald B. Bifunctional acid-base catalysis by imidazole groups in enzyme mimics. *J. Mol. Catal.*, **1994**, *91*, 161-174.
26. Liang C.; Abhilash K.G.; Breslow, R. Binding and biomimetic cleavage of the RNA poly(U) by synthetic polyimidazoles. *Proc Natl Acad Sci U S A.* **2012**, *109*, 12884–12887

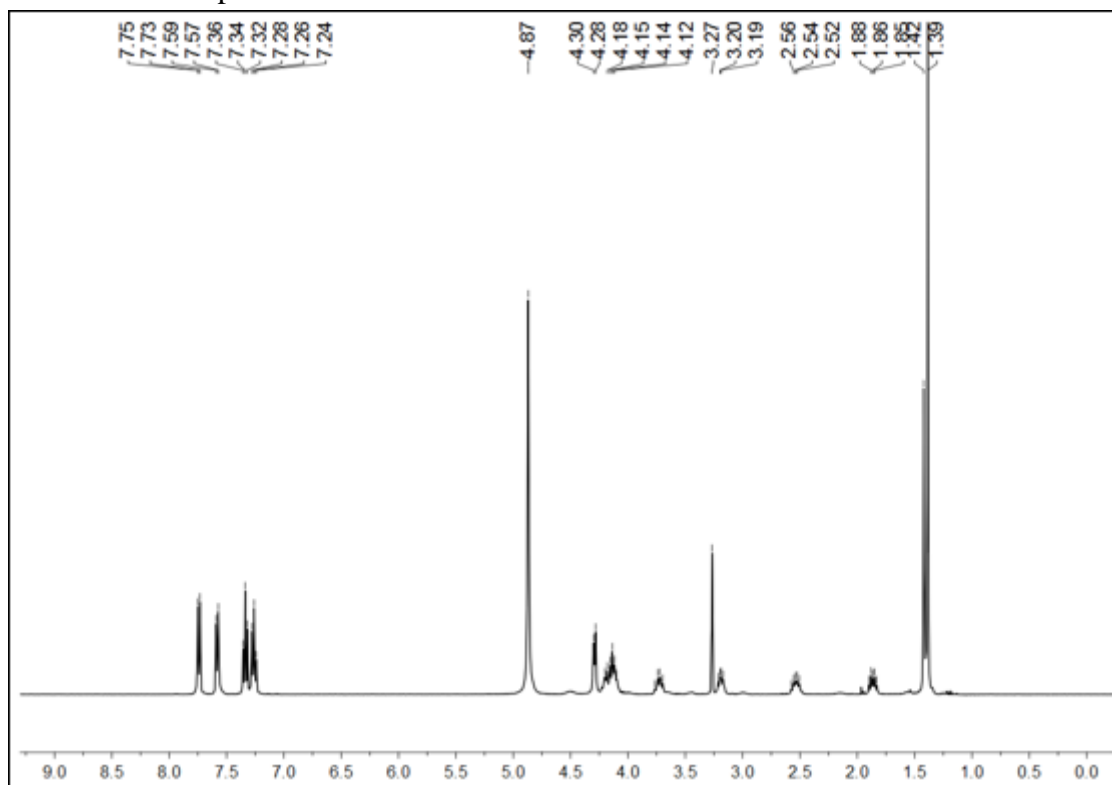
## 5.7 Appendix I: Characterisation data of synthesized compounds and peptides

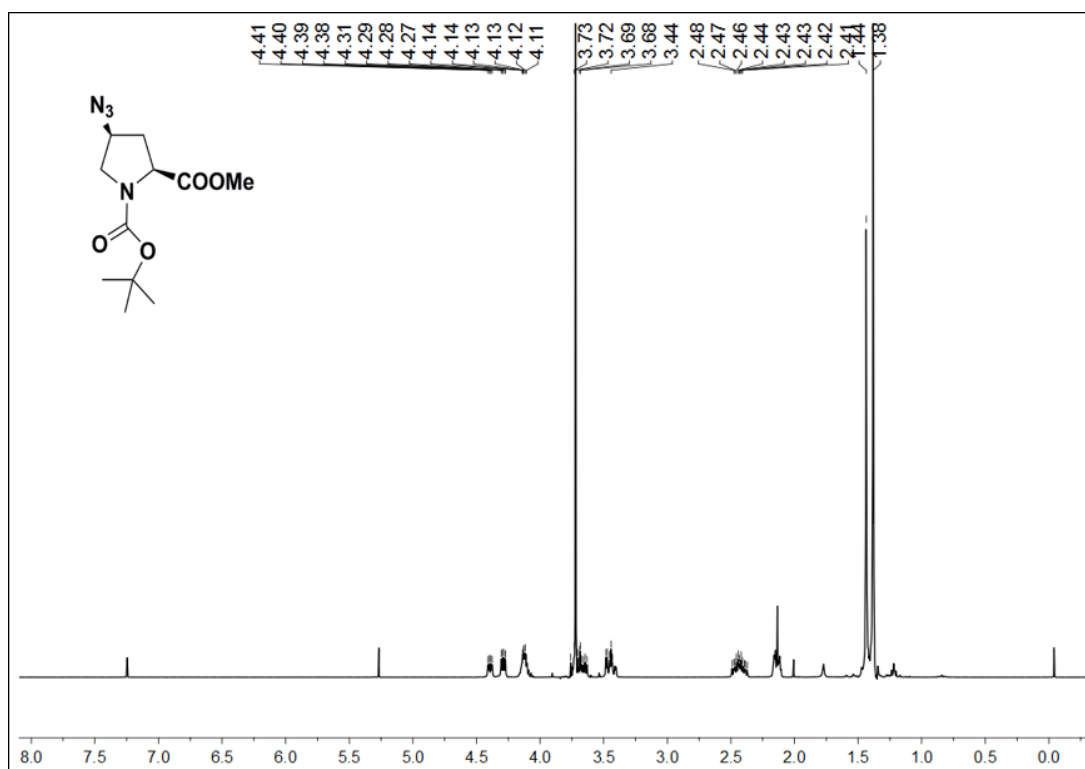
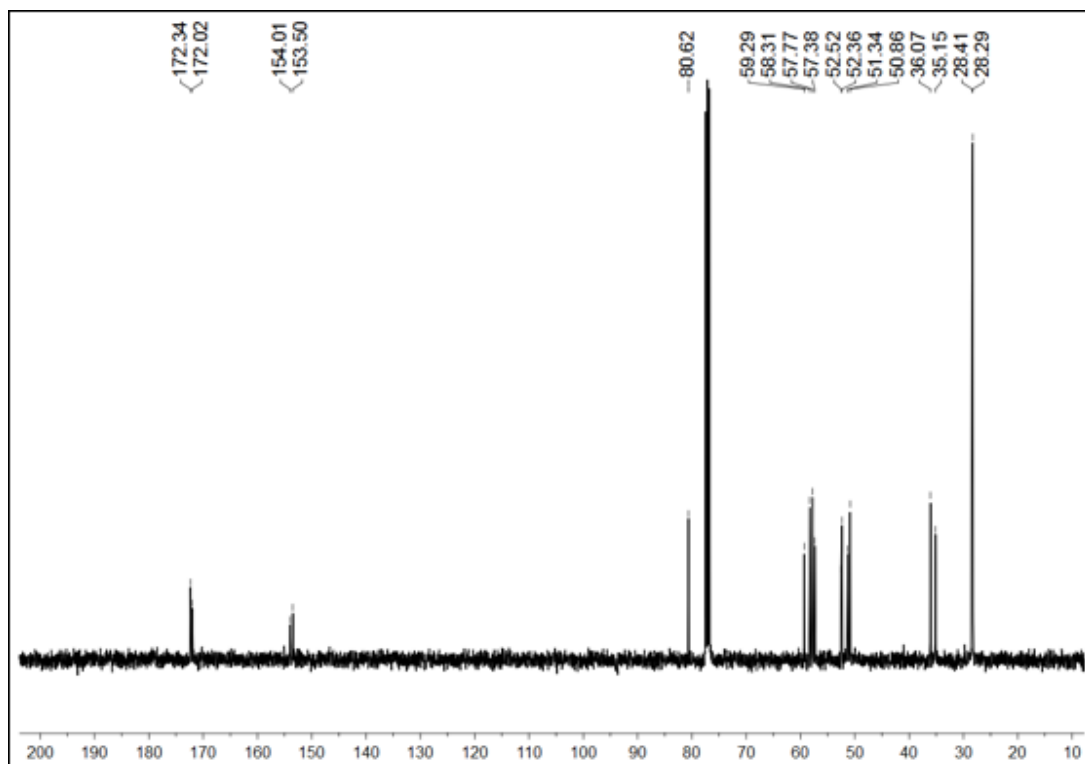
### $^1\text{H}$ and $^{13}\text{C}$ NMR of Compounds

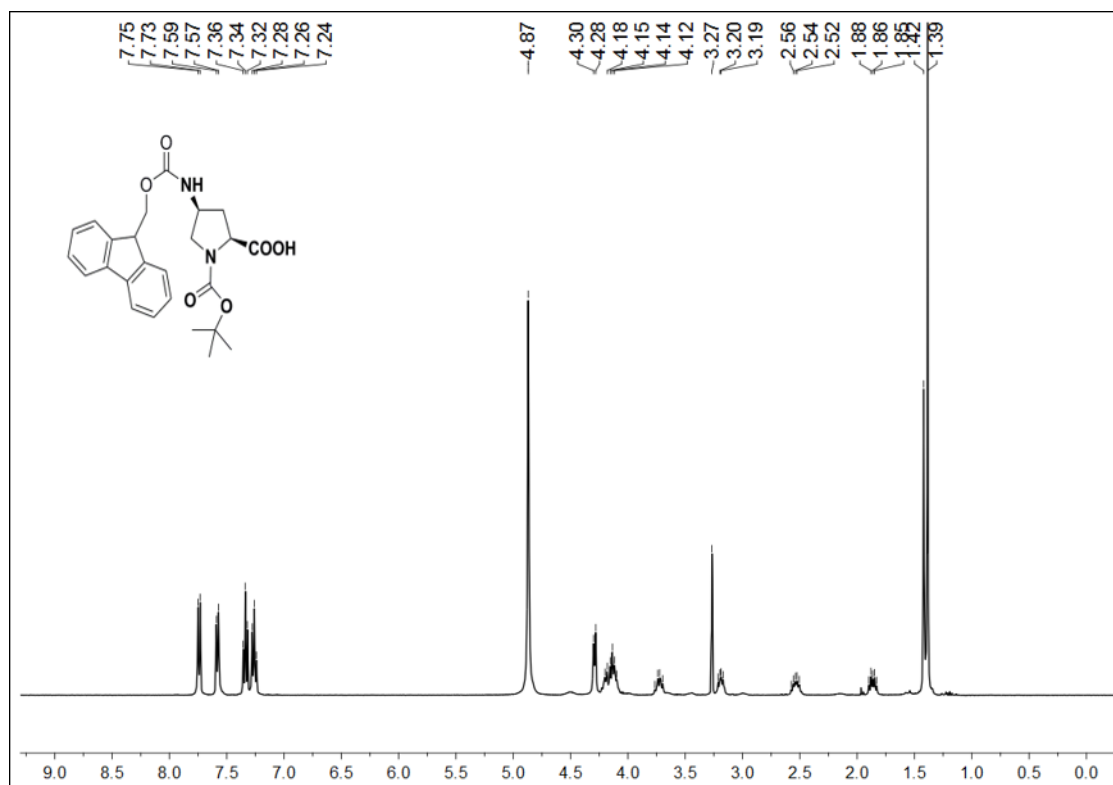
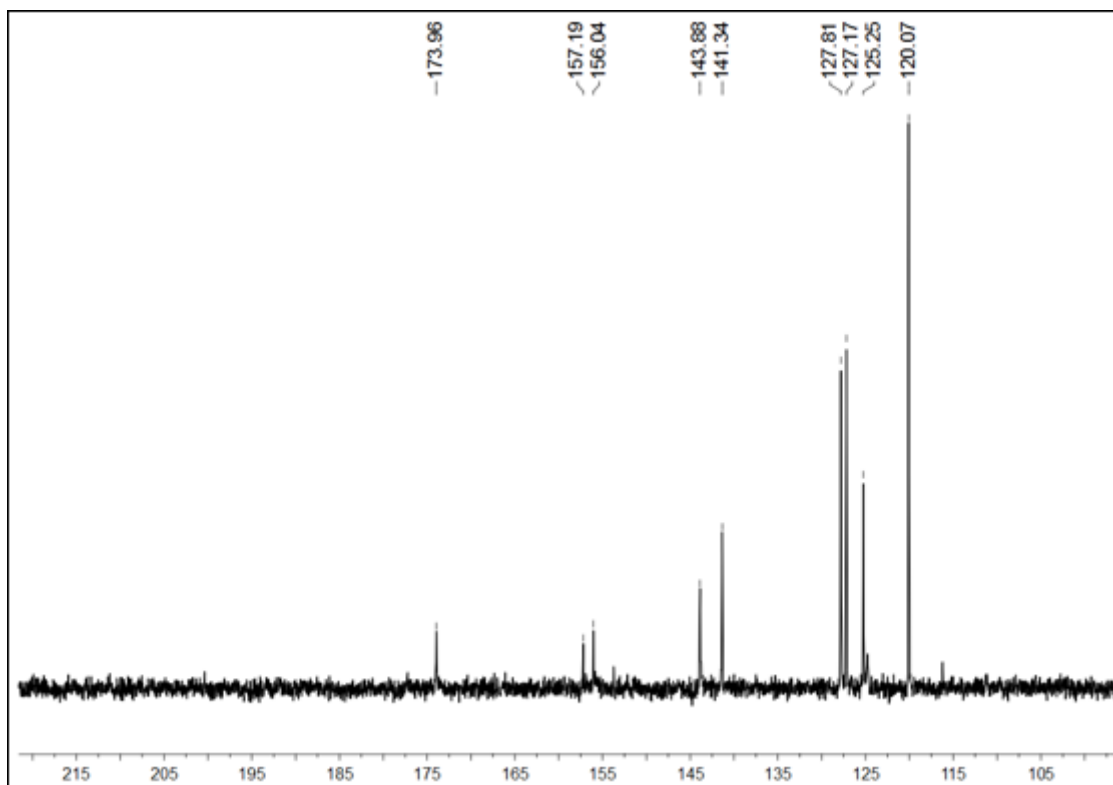
#### $^1\text{H}$ NMR of Compound 2



#### $^{13}\text{C}$ NMR of Compound 2

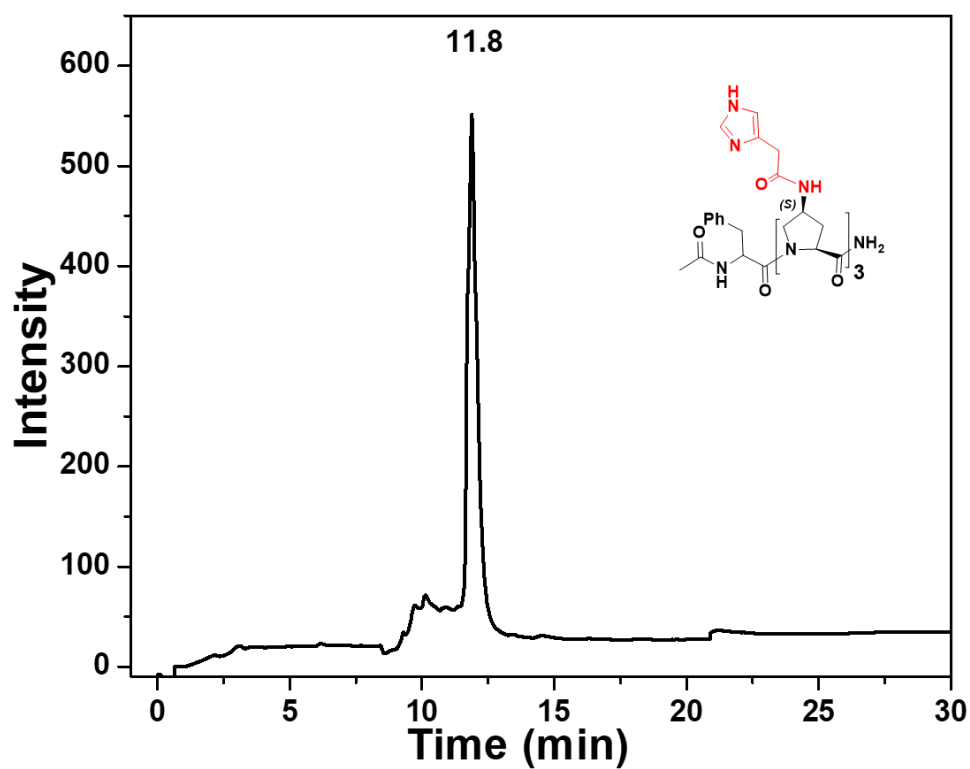


$^1\text{H}$  NMR of Compound 4 $^{13}\text{C}$  NMR of Compound 4

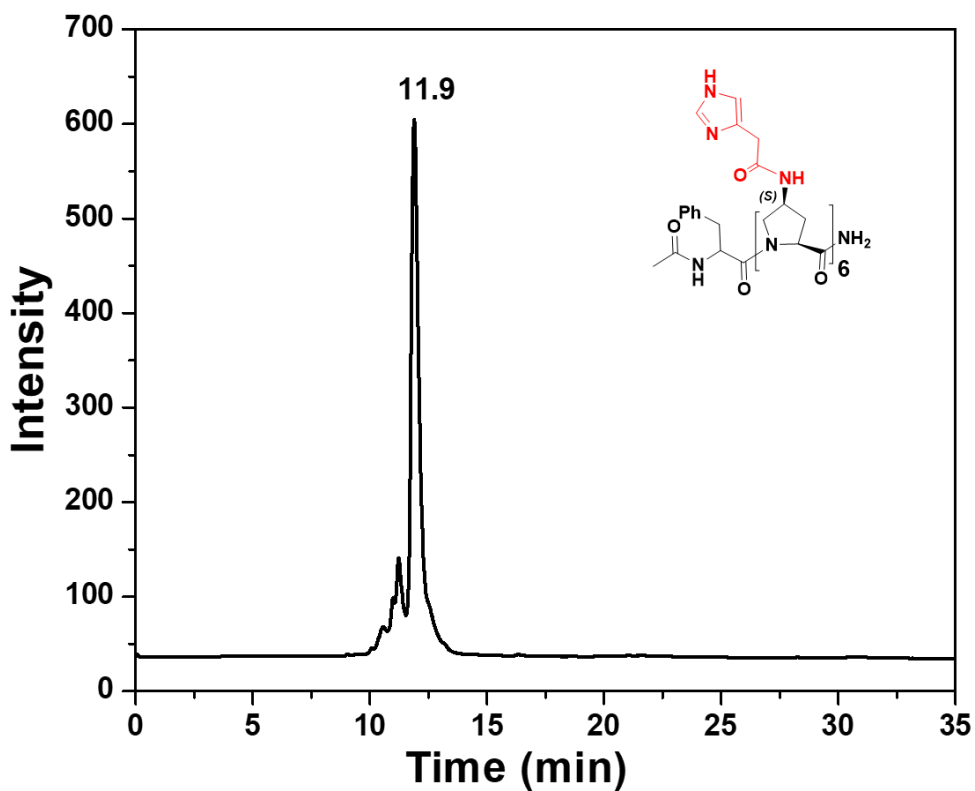
$^1\text{H}$  NMR of Compound **6** $^{13}\text{C}$  NMR of Compound **6**

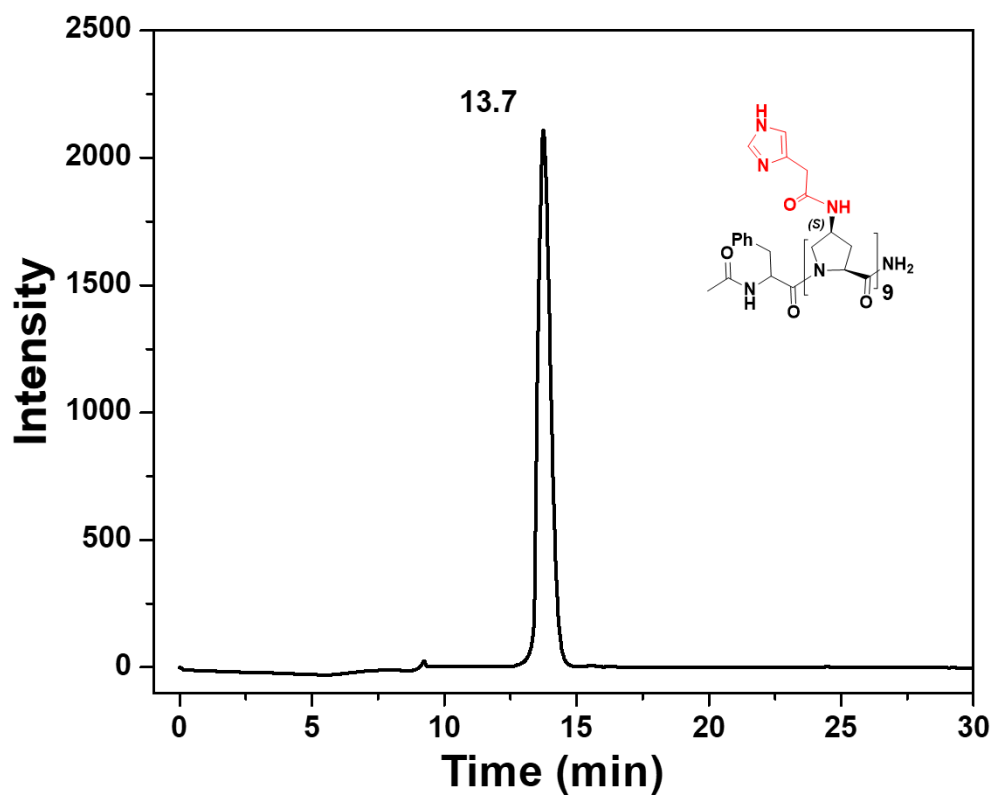
## HPLC chromatograms of peptides

## HPLC chromatograms of peptide P1

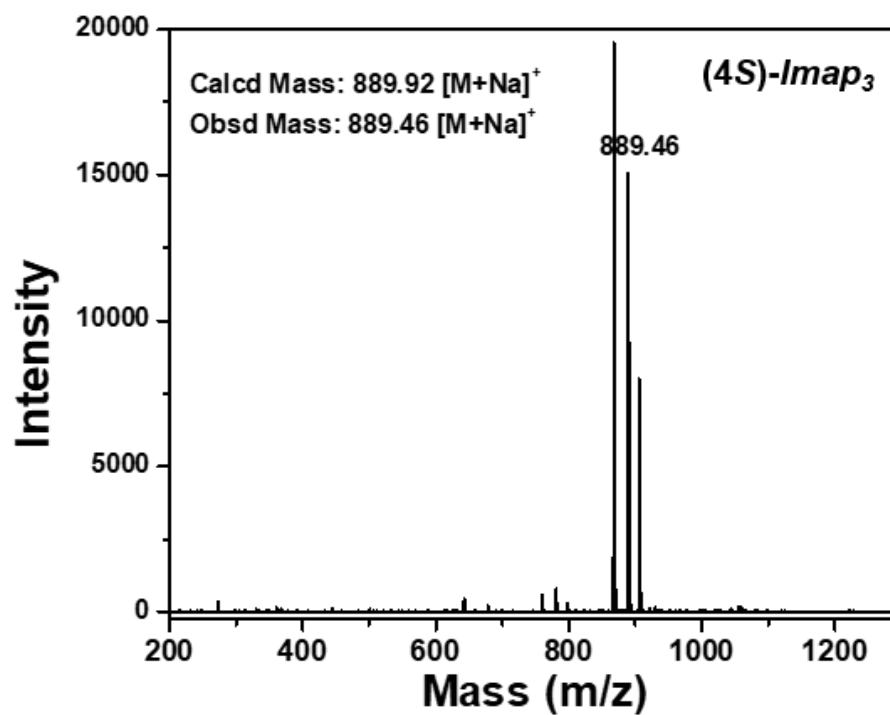


## HPLC chromatograms of peptide P2

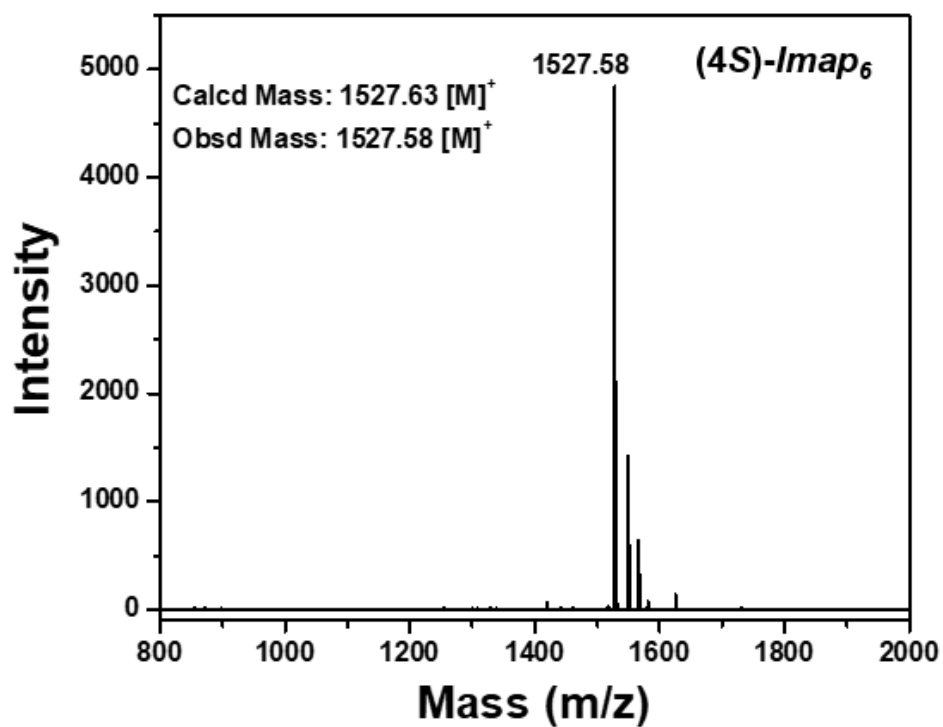


HPLC chromatograms of peptide **P3**

## MALDI -TOF Spectra of Peptides

MALDI -TOF Spectra of Peptide **P1**



MALDI -TOF Spectra of Peptide **P2**MALDI -TOF Spectra of Peptide **P3**