

Peptide Nucleic Acid (PNA)-GalNAc Conjugates for Targeted Cellular Delivery and C^γ(S/R)-Janus PNAs for Double Duplex Formation with cDNA

A thesis

**Submitted in partial fulfilment of the requirements of the degree
of**

Doctor of Philosophy

By

Pramod Prabhakar Bhingardeve

ID: 20143338

Research Supervisor

Prof. Krishna N. Ganesh



INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE

January 2020

This thesis is dedicated to...

My family

and

Teacher

CERTIFICATE

I certify that the work incorporated in the thesis entitled **“Peptide Nucleic Acid (PNA)-GalNAc Conjugates for Targeted cellular Delivery and Cy(S/R)-Janus PNAs for Double Duplex Formation with cDNA”** submitted by **Mr. Pramod Prabhakar Bhingardeve** 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: 1 January 2020



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:



Pramod Prabhakar Bhingardeve

(20143338)

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 Dr. S. G. Srivatsan, Dr. Vaijayanti Kumar and Dr. H. V Thulasiram 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 IISER Pune Chemistry faculty. Those interactions were short, but very useful as it helped me find fresh perspectives in the projects.

The technical staff in the chemistry department (Swati, Mahesh, Dnyaneshwar, Suresh, Sandip, Yatish, Nilesh) 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, Dr. Madhan, Satheesh, Vijay, Deepak, Nitin, Mahesh, Shahaji, Prabhakar and Pradnya had been kind and supportive. I am greatly indebted to Om Shanker, Manoj, Iranna, Shiraj, Pradeep, Rajat, Isha and Gaurav for helping me at numerous occasions. I had a wonderful company with my friends Ganesh, Sohan, Rupal, Sachin, Sanjeet. My special thanks to Anjana who is always besides me in all difficulties and happy moments.

Last but not the least; I shall always remain indebted to my parents, sisters and teachers, for their unconditional love, blessings, patience, support and encouragement.

Pramod Prabhakar Bhingardeve

Table of Contents

Chapter 1: Introduction to Peptide Nucleic Acids		
1.1	Introduction to nucleic acids	2
1.2	Base Pairing through Hydrogen bonding	3
1.3	Secondary structures of nucleic acids	4
1.4	Applications of nucleic acids	5
1.5	Chemical Modifications of DNA	10
1.6	Peptide nucleic acids (PNA)	14
1.7	Methodologies: Synthesis and Biophysical Techniques	33
1.8	Present work	40
1.9	References	44
Chapter 2: Synthesis and Characterization of Triantennary Galactosamine (GalNAc) and PNA Oligomers		
2.1	Introduction	53
2.2	Rationale for the present study	55
2.3	Objectives of the present work	56
2.4	Result and Discussion	56
2.5	Summary	64
2.6	Experimental Methods	64
2.7	Characterization data of synthesized compounds and PNA	83
2.8	Reference	111
Chapter 3: Biophysical evaluation and Cell Permeation Studies of GalNAc₃-PNA oligomers		
3.1	Introduction:	113
3.2	Rational for the Present Work	113
3.3	Objective of the Present work	114
3.3.1	GalNAc ₃ -PNA oligomers used for biophysical studies	114
3.3.2	DNA/RNA oligonucleotides used for biophysical studies	115

3.4	Result and discussion	116
3.4.1	UV melting studies of GalNAc ₃ -PNA:DNA/RNA hybrids	116
3.4.1a	<i>Thermal stability of PNA, Triantennary GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA:DNA duplexes</i>	117
3.4.1b	<i>Thermal stability of PNA, Triantennary GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA:RNA duplexes</i>	118
3.4.2	CD Studies of PNA:DNA and PNA:RNA Duplexes	119
3.4.3	UV-visible and fluorescence studies of <i>Cf</i> -PNA, <i>Cf</i> -GAINAc ₃ -PNA and <i>Cf</i> -[C ^γ (S)-GalNAc-T] ₃ -PNA and their duplexes	120
3.5	Cellular Uptake Studies	121
3.5.1	Cellular uptake experiment using confocal microscopy	122
3.5.1a	<i>Cellular uptake studies with HepG2 cells</i>	122
3.5.1b	<i>Cellular uptake studies with HEK-293 cells.</i>	123
3.5.1c	<i>Flow cytometry studies</i>	125
3.5.1d	<i>Cytotoxicity studies</i>	125
3.6	Conclusion	126
3.7	Experimental procedures	127
3.8	References	131
Chapter 4: C^γ-Janus PNAs: Design, Synthesis and Characterization of C^γ-substituted Janus PNA monomers		
4.1	Introduction	133
4.2	Rational of present Work	134
4.3	Objective of the present work	135
4.4	Synthesis of modified PNA monomers	135
4.5	Summary	143
4.6	Experimental Section	144
4.7	Characterization data of synthesized compounds	167
4.8	Reference	204
Chapter 5: Design, Synthesis and Characterization of C^γ Janus PNA Oligomers		
5.1	Aim of Present Work	206

5.2	Rational for choice of sequence	207
5.3	Results and discussion	209
5.3.1	Solid phase Synthesis of PNA oligomers	209
5.3.2	Synthesis of mixed purine-pyrimidine <i>Janus</i> PNA Oligomers	210
5.3.2a	<i>Synthesis of homo Janus PNA oligomer by Boc chemistry on solid phase</i>	210
5.3.2b	<i>Synthesis of hetero Janus PNA with mixed sequence on both amide and Cγ face</i>	212
5.3.2c	<i>Synthesis of PNA analogues with nucleobase on only aminoethyl glycylic Cγ face</i>	213
5.3.2d	<i>Synthesis of PNA analogues with nucleobase on only tertiary amide face aeg PNA 7- 10</i>	214
5.3.2e	Summary of all PNA sequences	215
5.3.3	Cleavage of PNA oligomers from the solid support	218
5.3.4	Purification and characterization of the PNA oligomers	218
5.4	Summary	220
5.5	Experimental Methods	220
5.6	References	221
5.7	Appendix I	223
Chapter 6: Biophysical Studies of <i>Janus</i> PNA Oligomers		
6.1	Introduction	243
6.2	Objectives of the present work	243
6.3	<i>Janus</i> PNA (JP) oligomers used for biophysical studies	244
6.4	DNA oligonucleotides used for biophysical studies	245
6.5	<i>Janus</i> PNA complexes	246
6.6	Results	249
6.6.1	Thermal stability of <i>aeg</i> - γ R/S-PNA with complementary DNA duplexes	249
6.6.1a	<i>Thermal stability of aeg-γ PNAs p8-γS-C₆ and p8-γR-C₆ with mismatch DNA duplexes</i>	250
6.6.1b	<i>CD spectra of DNA duplexes of Cγ-PNA:DNA duplexes</i>	251
6.6.1c	<i>Thermal stability of aeg-γ(S/R)-G₆ PNA with complementary DNA duplexes</i>	252
6.6.1d	<i>Thermal stability of p8-γS-G₆ and p8-γR-G₆ PNAs with mismatch DNA duplexes</i>	253

6.6.1e	<i>CD spectra of DNA duplexes of Cγ(S/R) PNAs p8-γS-G₆ (PNA 13S) and p8-γR-G₆ (PNA 14R)</i>	254
6.6.2	Binding studies of homo Janus C γ -PNAs T _{8jp} - γ S-C ₆ and T _{8jp} - γ R-C ₆ with complementary DNA	254
6.6.2a	<i>Thermal stability of complex of Cγ(S/R) homo Janus PNA T_{8jp}-γS-C₆:DNA duplexes</i>	255
6.6.2b	<i>Thermal stability of duplex of triplex of homo Janus-PNA T_{8jp}-γR-C₆:DNA duplexes</i>	257
6.6.2c	<i>CD spectra of Cγ(S/R) homo Janus triplexes and duplexes</i>	258
6.6.2d	<i>Thermal stability of duplexes and double duplex from Cγ-Janus PNA duplexes with mismatch DNA</i>	260
6.6.2e	<i>CD spectral study of order of duplex formation</i>	264
6.6.3	Thermal stability of duplexes and double duplex from Janus PNA JP 3S (T _{8jp} - γ S-G ₆) and JP 4R (T _{8jp} - γ R-G ₆) duplexes with complementary DNA	266
6.6.3a	<i>Thermal stability complexes from Janus PNA JP 3S and JP 4R with mismatch DNA</i>	269
6.6.3b	<i>CD spectra of Janus PNA JP 3S (T_{8jp}-γS-G₆) and JP 4R (T_{8jp}-γR-G₆) with cDNA</i>	273
6.6.4	Thermal stability of duplexes and double duplex from hetero Janus PNA JP 5S (mjp- γ S-PNA) and JP 6R (mjp- γ R-PNA) with cDNA	274
6.6.4a	<i>Thermal stability of duplexes and double duplex of Cγ(S/R) Janus PNAs with mismatch DNA</i>	278
6.6.4b	<i>CD spectra of duplexes of Cγ(S/R) hetero Janus PNAs, with cDNA 4/cDNA 5</i>	281
6.7	Thermal stability of non-Janus C γ -PNA-C _n with complementary DNA	282
6.7.1	i-Motif studies by pH-dependent UV-T _m	281
6.7.2	Thermal stability of quadruplex formation by non-Janus PNA-G _n -oligomers	284
6.7.3	G-quadruplex studies by Variable-Temperature Circular Dichroism	285
6.8	Ethidium bromide displacement assay from dsDNA	286
6.9	Isothermal Titration Calorimetry	288

6.9.1	Binding of the JP 1S ($T_8jp-\gamma S-C_6$) and PNA 7S ($T_8jp-\gamma S-NH^{Ac}$) double duplex with of complementary DNA.	289
6.9.2	Binding of the JP 2R ($T_8jp-\gamma R-C_6$) and PNA 8R ($T_8jp-\gamma R-NH^{Ac}$) with cDNA (dA ₈)	291
6.10	Conclusion	293
6.11	Summary	295
6.12	Experimental procedures	295
6.13	References	299
6.14	Appendix II	300
	Summary of thesis and outlook	302

Abbreviations

A	Adenine
Abs.	Absolute
AcOH	Acetic acid (glacial)
Ac ₂ O	Acetic anhydride
ACN	Acetonitrile
<i>aeg</i>	Aminoethylglycine
<i>eam</i>	Ethylamino
<i>eaz</i>	Ethylazide
ASGPR	Asialoglycoprotein receptor
<i>tz</i>	Triazolyl
<i>etz</i>	Ethyltriazolyl
ap	Antiparallel
<i>aq.</i>	Aqueous
(Boc) ₂ O	Boc anhydride
Bn	Benzyl
BIAB	(Diacetoxyiodo)benzene
Bt	Benzotriazole
C	Cytosine
Calcd	Calculated
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
CF	5(6)-Carboxyfluorescein
CHCA	α -cyano-4-hydroxycinnamic acid
DCC	Dicyclohexylcarbodiimide
DCE	Dichloroethane
DCM	Dichloromethane
DHB	2,5-dihydroxybenzoic acid
DHBtOH	3-Hydroxy-1,2,3-benzotriazin-4(3H)-one
DIPEA	N,N-Diisopropylethylamine

DIC	Differential Interference Contrast
DMAP	N,N-Dimethyl-4-aminopyridine
DMF	N,N-dimethylformamide
DMSO	N,N-Dimethyl sulfoxide
DNA	2'-deoxyribonucleic acid
ds	Double stranded
EBA	Ethylbromo acetate
EDTA	Ethylene diamine tetraacetic acid
Et	Ethyl
EtOAc	Ethyl acetate
FESEM	Field Emission Scanning Electron Microscope
Fmoc	9-Fluorenylmethoxycarbonyl
g	gram
G	Guanine
GalNAc	N-Acetylgalactosamine
gly	Glycine
h	Hours
HBTU	2-(1H-Benzotriazole-1-yl)- 1,1,3,3 tetramethyl-uronum-hexafluoro-phosphate
HEK	Human Embryonic Kidney Cell
HOBt	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
in situ	In the reaction mixture
in vivo	Within the living
ITC	Isothermal Titration Calorimetry
IR	Infra-red
L-	Levo-
Lys	Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight

MBHA	4-Methyl benzhydryl amine
MEOH	Methanol
mg	milligram
MHz	Megahertz
MsCl	Mesyl Chloride
μL	Microliter
μM	Micromolar
min	minutes
mL	milliliter
mM	millimolar
mmol	millimoles
mp	melting point
MW	Molecular weight
N	Normal
nm	Nanometer
NMR	Nuclear Magnetic Resonance
Obsvd.	Observed
ONs	Oligonucleotides
p	Parallel
PCR	Polymerase chain reaction
Pd	Palladium
PIDA	iodosobenzene diacetate (phenyliodonium diacetate)
ppm	Parts per million
PNA	Peptide Nucleic Acid
PS-oligo	Phosphorothioate-oligo
R_f	Retention factor
RNA	Ribonucleic Acid
RP	Reverse Phase (-HPLC)
rt	Room temperature
RT	Retention time

S	Sinister
SPPS	Solid Phase Peptide Synthesis
<i>ss</i>	Single strand/single stranded
T	Thymine
TBAc	tert-Butyl bromoacetate
TEA/Et ₃ N	Et ₃ N/Triethylamine
TEMPO	Tetramethylpiperidine 1-oxyl
TFA	Trifluoroacetic acid
Tf ₂ O	Triflicanhydride
TfN ₃	Trifluoromethanesulfonyl azide
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
TMSOTf	Trifluoromethanesulfonate
TIPS	Triisopropyl Silane
TLC	Thin layer chromatography
<i>T_m</i>	Melting temperature
UV-Vis	Ultraviolet-Visible

Abstract

Peptide Nucleic Acid (PNA)-GalNAc Conjugates for Targeted Cellular Delivery and C^γ(S/R)-Janus PNAs for Double Duplex Formation with cDNA

The thesis is comprised of two components. The first part deals with studies on peptide nucleic acids (PNA) which are conjugated to triantennary N- acetyl galactosamine (GalNAc) for targeted delivery to hepatocytes via asialoglycoprotein receptor (ASGPR)¹. This receptor is abundantly expressed exclusively on hepatocytes and clears glycoproteins and desialylated platelets². Inspired from recent reports of siRNAs conjugated with GalNAc efficiently permeating hepatocyte cells, the work presented in this thesis aims to conjugate triantennary GalNAc to PNA and study their entry into hepatocyte derived cells. The second part of thesis is concerned with the design, synthesis and evaluation of an entirely new class of PNA analogues -“Janus PNA” in which a single strand of PNA is designed to carry two sets of nucleobase sequences, one set same as the standard t-amide linked bases as in standard PNA and the other set linked at C^γ on backbone through a secondary amide linker. Such Janus PNAs can simultaneously bind to two cDNA sequences and opens up a new paradigm for PNA recognition by cDNA for biological and material applications. The thesis is presented as six chapters:

Chapter 1: Introduction to PNA

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 (PNA). It focuses on attempts made in the last few years to optimize the properties of natural oligonucleotides through chemical modifications at phosphate, ribose or the nucleobase. PNA introduced by Nielsen et al.³ is a novel DNA mimics in which the sugar-phosphate backbone is replaced with a pseudo peptide backbone in the form of achiral 2-(aminoethyl)-glycine (*aeg*) linkage (Figure 1). Nucleobases (A/T/G/C) are attached through methylene carbonyl via t-amide link to the backbone. PNA hybridizes with complementary DNA/RNA via Watson–Crick base-pairing with higher avidity compared to DNA:DNA or DNA:RNA duplexes. PNAs are resistant to both proteases and nucleases and consequently have higher stability in the cellular environment. The simplicity of chemical structure, ease of synthesis,

superior binding properties and lack of toxicity have made PNA an attractive gene regulatory agent. However, the major drawbacks of PNA are poor water solubility, inefficient cellular uptake and self-aggregation that restrict applications for PNA-based gene-targeting agents. This problem has been addressed by a number of chemically modification of PNA. The introduction chapter reviews this area and the effects of different modifications on biophysical and biochemical properties of PNA to rationalize the aims of the present work.

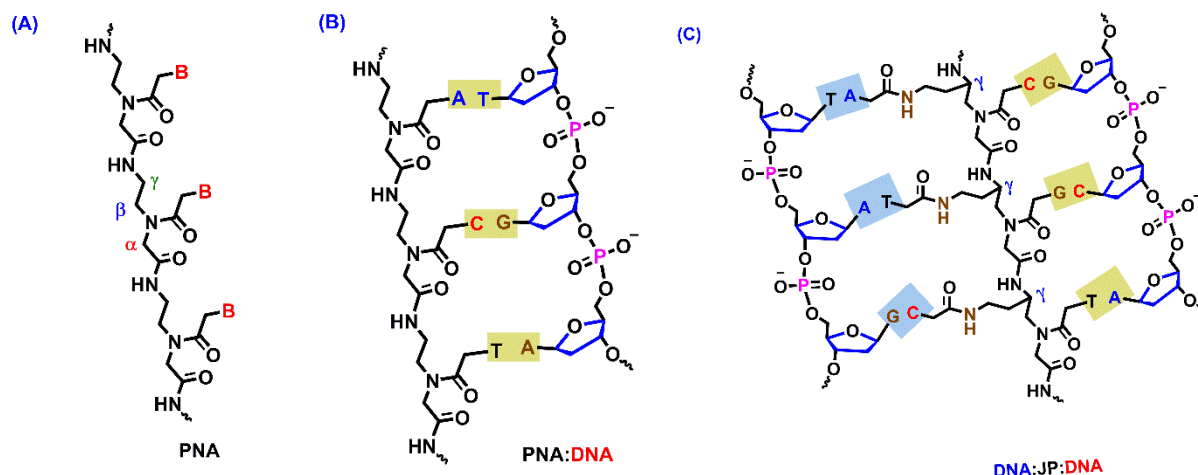


Figure 1. Structure of (A) peptide nucleic acid (B = nucleobase), (B) Complex of PNA with complementary DNA, (C) Janus PNA with complementary DNA.

Chapter 2: Synthesis and characterization of triantennary N-acetylgalactosamine (GalNAc) conjugated PNA oligomers. The siRNAs conjugated to triantennary N-acetylgalactosamine (GalNAc₃) induce robust RNAi-mediated gene silencing in the liver owing to the asialoglycoprotein receptor (ASGPR) mediated uptake.⁴⁻⁶ This chapter aims to extend this idea to PNA for its conjugation with triantennary galactosamine. Starting from GalNAc pentanoic acid **1**, the triantennary galactosamine **2** was synthesized by known protocols for conjugation with desired PNA sequence at N-terminus (Figure 2), assembled on solid phase, followed by cleavage from resin. The resulting conjugates were purified by HPLC and characterized by mass spectra (Figure 3). Another structural variation of GalNAc-PNA consisted of linking GalNAc pentanoic acid to C^γ-sidechain of PNA monomer **3** and the modified PNA monomer was used to couple to N-terminus of PNA to generate a linear trimer [C^γ(S)-GalNAc-T]₃ rather than triantennary GalNAc (Figure 3). All intermediates and final PNAs were characterized by standard protocols.

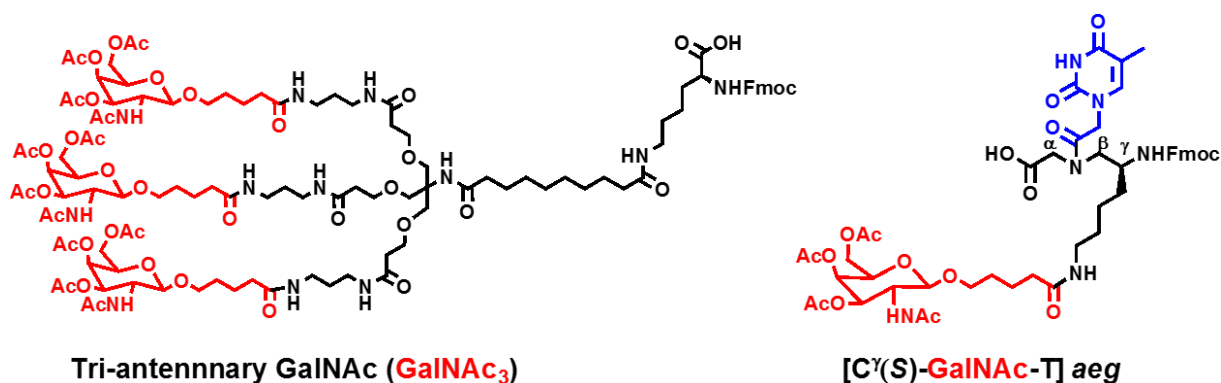
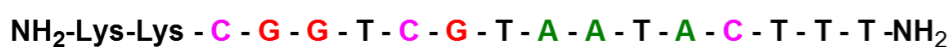
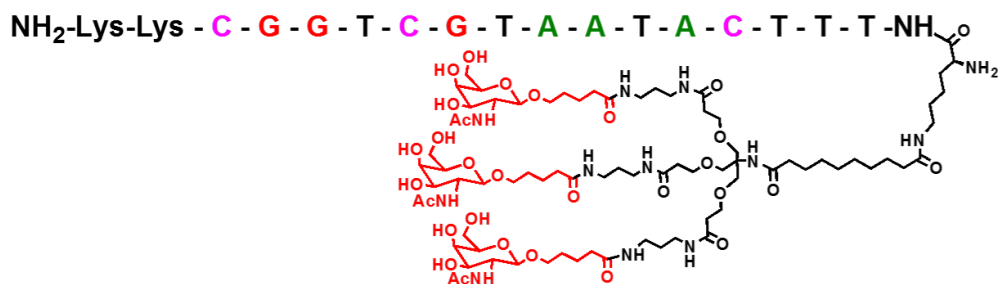


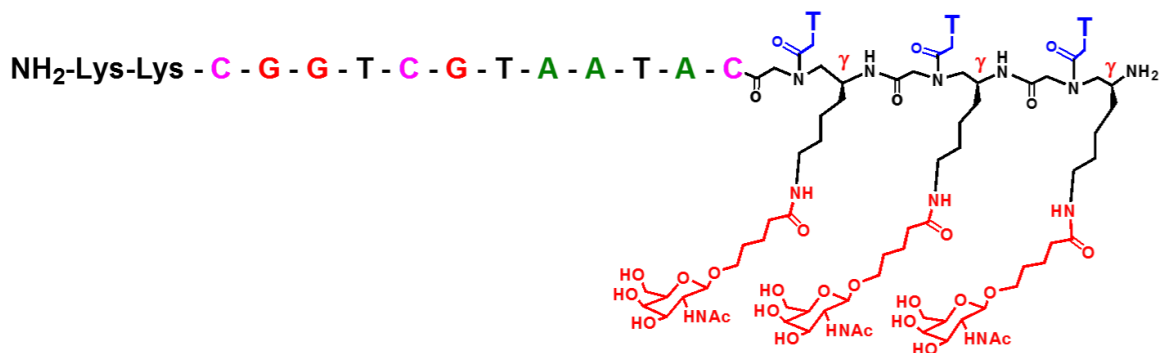
Figure 2 Chemical structures of target GalNAc moieties



A) PNA



B) GalNAc₃-PNA



C) [C^γ(S)-GalNAc-T]₃-PNA

Figure 3 Structures of control PNA (A) triantennary GalNAc₃-PNA and (B) linear trimer [C^γ(S)-GalNAc-T]₃ oligomers (C).

Chapter 3: Biophysical evaluation and cell permeation studies of GalNAc₃-PNA oligomers

The temperature dependent UV absorbance studies of PNA:DNA duplexes were used to determine the effects of GalNAc conjugation on thermal stabilities of derived

Abstract

PNA:DNA/RNA duplexes with complementary DNA/RNA. GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA with complementary DNA showed stabilization of corresponding duplexes by 3.6 °C and 2.4 °C respectively (Table 1). In contrast, the triantennary GalNAc₃-PNA destabilized the derived RNA duplex by -3.7 °C, while the linear trivalent [C^γ(S)-GalNAc-T]₃ PNA destabilized the RNA duplex by -5.7 °C. The relative order of stability of the PNA:RNA duplexes were in the order PNA > GalNAc₃-PNA > [C^γ(S)-GalNAc-T]₃-PNA (Table 2).

Table 1 UV-*T_m* values of complementary PNA:DNA duplexes with triantennary-GalNAc₃ PNA and trivalent- [C^γ(S)-GalNAc-T]₃ -PNA units.

Entry	PNA Sequence	5'GCCAGCATTATGAAA3'	
		UV- <i>T_m</i> (°C)	Δ <i>T_m</i> (°C)
1	H-TTTCATAATGCTGGC-Lys-Lys-NH ₂	48.9	-
2	GalNAc ₃ -TTTCATAATGCTGGC-Lys-Lys-NH ₂	52.5	+ 3.4
3	[C ^γ (S)-GalNAc-T] ₃ -TTTCATAATGCTGGC-Lys-Lys-NH ₂	51.3	+ 2.4

Table 2 UV-*T_m* values GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃ PNA:RNA duplexes

Entry	PNA Sequence	5'GCCAGCAUUAUGAAA3'	
		UV- <i>T_m</i> (°C)	Δ <i>T_m</i>
1	H-TTTCATAATGCTGGC-Lys-Lys NH ₂	59.5	-
2	GalNAc ₃ -TTTCATAATGCTGGC-Lys-Lys NH ₂	55.8	-3.7
3	[C ^γ (S)-GalNAc-T] ₃ -CATAATGCTGGC-Lys-Lys NH ₂	53.8	-5.7

The effect of GalNAc conjugation on the conformation of PNA:DNA/RNA duplexes was investigated by circular dichroism (CD) spectra. The spectra indicated that the GalNAc conjugated PNA:DNA/RNA retained similar conformation as unmodified PNA:DNA/RNA duplexes. In order to gain insights into the cellular uptake properties of GalNAc -PNA oligomers, the PNAs were tagged with 5(6)-carboxyfluorescein at the N-terminus of the PNA for visualization in the cells (Figure 4). The fluorescently labeled GalNAc-PNA analogs were tested for cell permeation in human embryonic Kidney (Hek 293) cell line and human liver cancer (Hep G2) cell line. The GalNAc-conjugated PNAs were selectively uptaken in Hep G2 cells and very interestingly, the linear trimeric *Cf*-[C^γ(S)-GalNAc-T]₃ PNA was found to be better than tri-antennary *Cf*-GalNAc₃-PNA (Figure 5). Moreover, both the PNAs showed much enhanced cell penetration in Hep G2 cells than the unmodified PNA and the negative control HEK 293 cells, clearly providing the proof of concept for hypothesis.

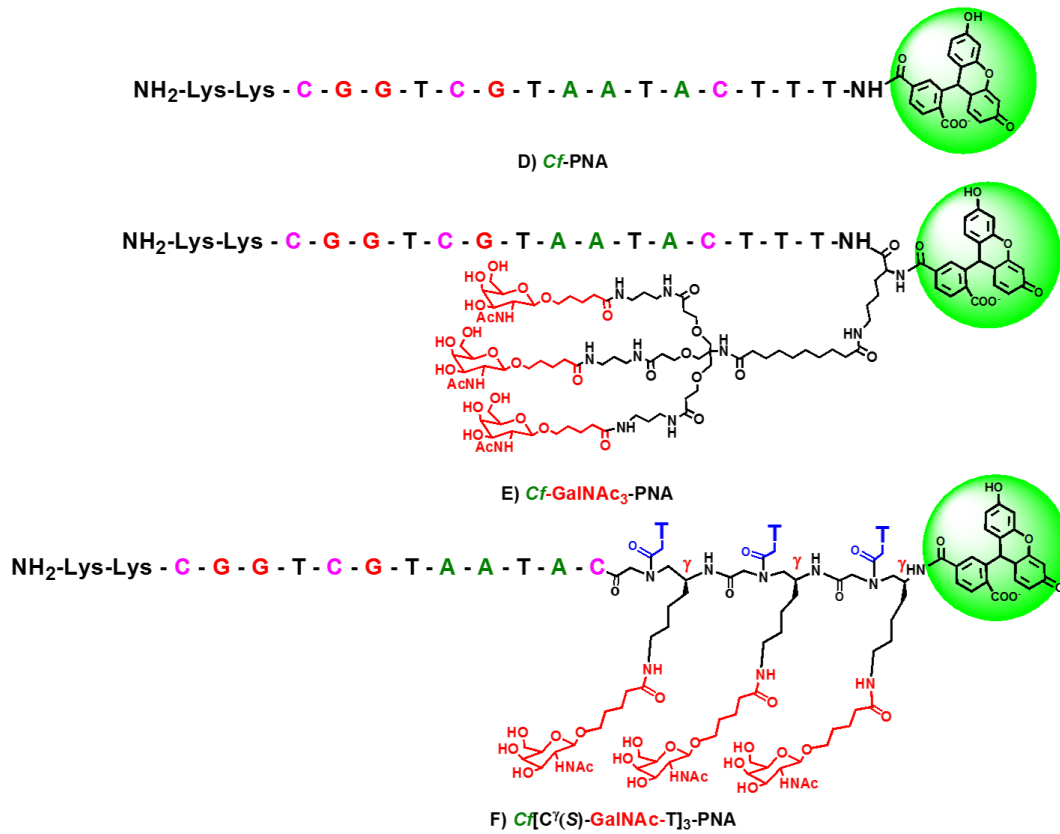


Figure 4 Carboxyfluorescein tagged aeg-PNAs and GalNAc-PNAs.

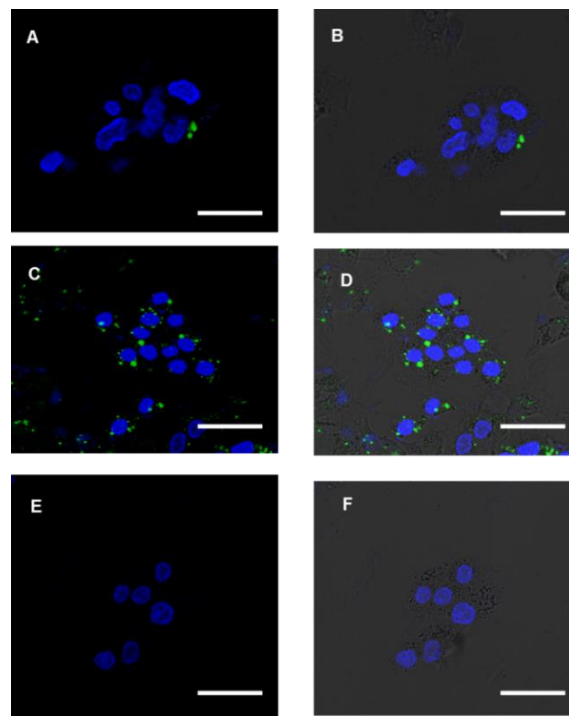


Figure 5 Confocal microscopy images of HepG2 cells treated with *Cf*-GalNAc₃-PNA (A-B), *Cf*-[C^γ(S)-GalNAc-T]₃ PNA (C-D) and *Cf*-PNA (E-F). A, C, and E are merged images of DAPI and carboxyfluorescein channels. B, D, and F shows the overlay of the merged images of DAPI and

carboxyfluorescein channels with differential interference contrast image showing the intact cells (gray), nuclei (blue), and internalized PNAs (green). Scale bars represent 25 μm .

Chapter 4: C^γ -Janus PNAs: Design, Synthesis and Characterization of C^γ -substituted *Janus* PNA Monomers

This chapter aims to address the design, synthesis of C^γ -ethylamino substituted *aeg* PNA monomers (1-4) as rationally designed PNA analogs having side chains carrying amino function for conjugating different nucleobases (eg. 6-9, Figure 6). The presence of nucleobases on both faces provides opportunity for C^γ -Janus PNAs to bind simultaneously to two cDNAs, as compared to just one cDNA by standard PNAs. The modified PNA monomers 1 and 2 were used to construct Janus PNA by post synthetic modification through global amide conjugation using base linked acetic acids 6-9 for making homo Janus PNAs or stepwise synthesis on solid phase for making hetero Janus PNAs, with mixed sequences on both sides.

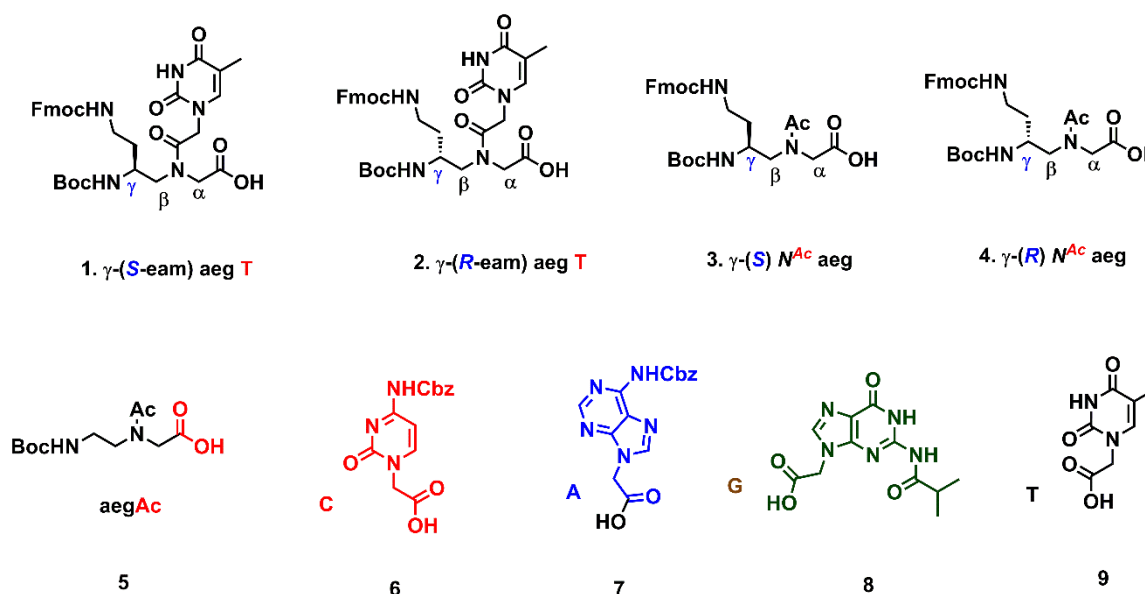


Figure 6 C^γ substituted PNA monomers (1-4)

Chapter 5: Design, Synthesis and Characterization of C^γ -*Janus* PNA Oligomers

This chapter deals with incorporation of C^γ -modified PNA units into Janus PNA oligomers by solid phase followed by global coupling reaction to synthesize various kinds of *Janus* peptide nucleic acids oligomers (Figure 7), *homo Janus* PNA (Figure 8) and *Mix*

Janus PNA (Scheme 2, Figure 9), depending on the type of C^γ -stereochemistry (*R/S*) of the monomers, bases(A/T/G/C) and linkage connecting them to the backbone for eg. amide. The synthesis of PNA oligomers were followed by cleavage from resin, purification by RP-HPLC and characterization by MALDI-TOF spectrometry.

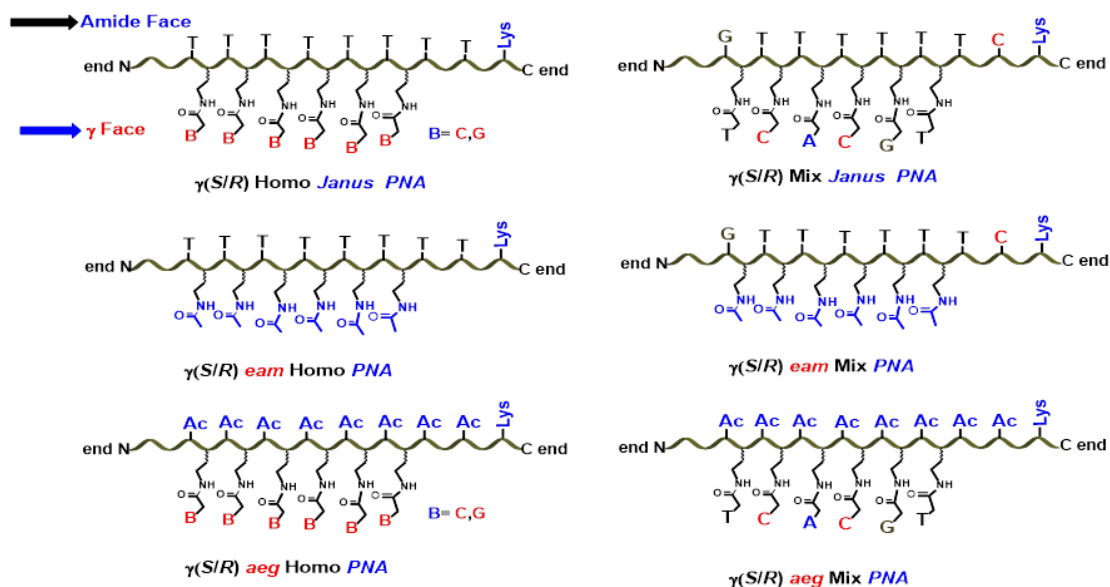


Figure 7 Types of *Janus* PNA oligomers

Scheme 1 Synthetic strategy of *homoJanus* PNA on solid support

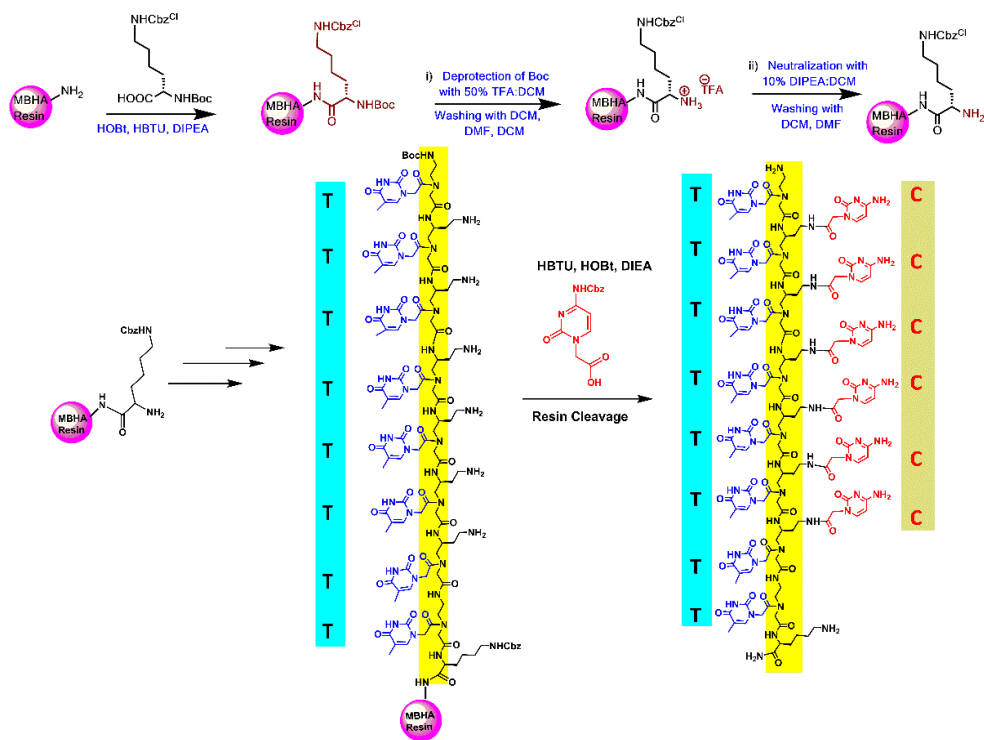


Figure 8 Synthetic strategy of *homoJanus* PNA oligomer (**JP 1**) on solid support

Scheme 2 Synthetic strategy of hetero Janus PNA oligomer.

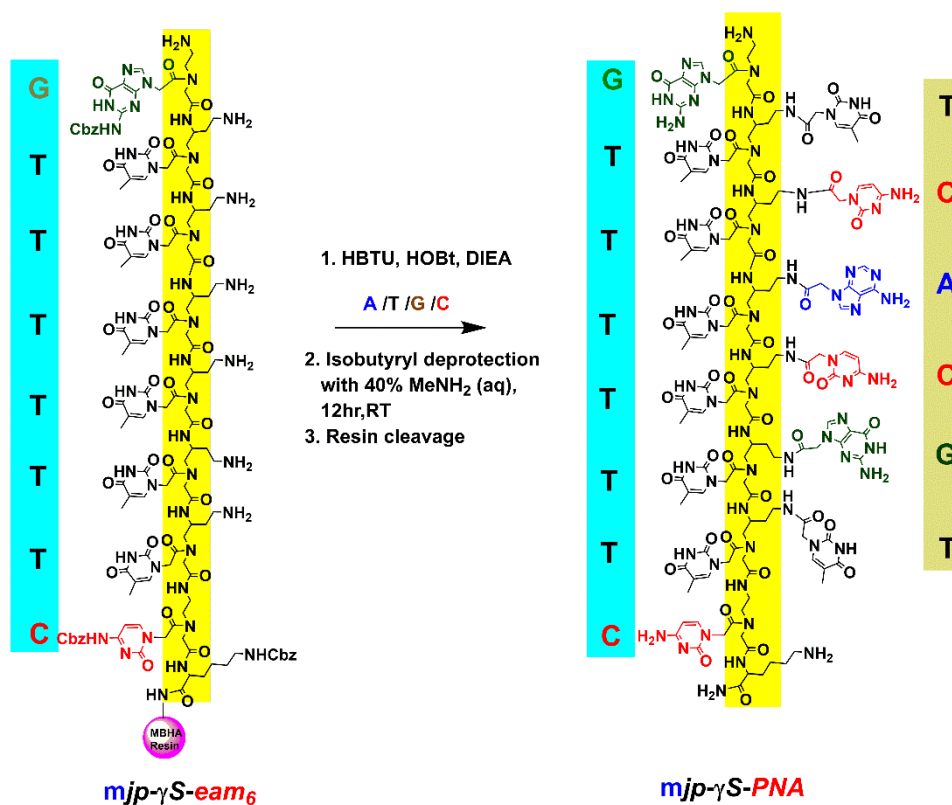
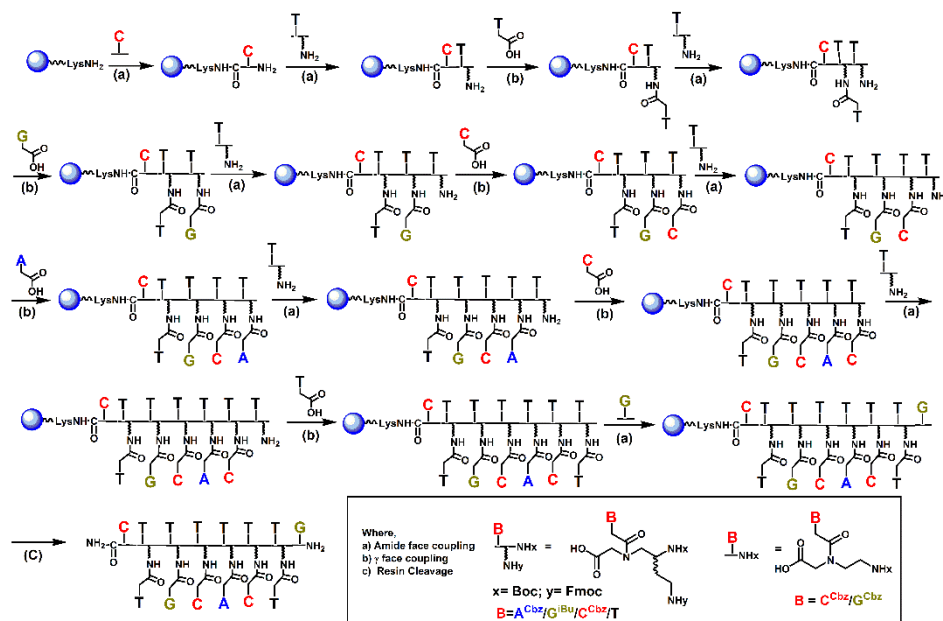


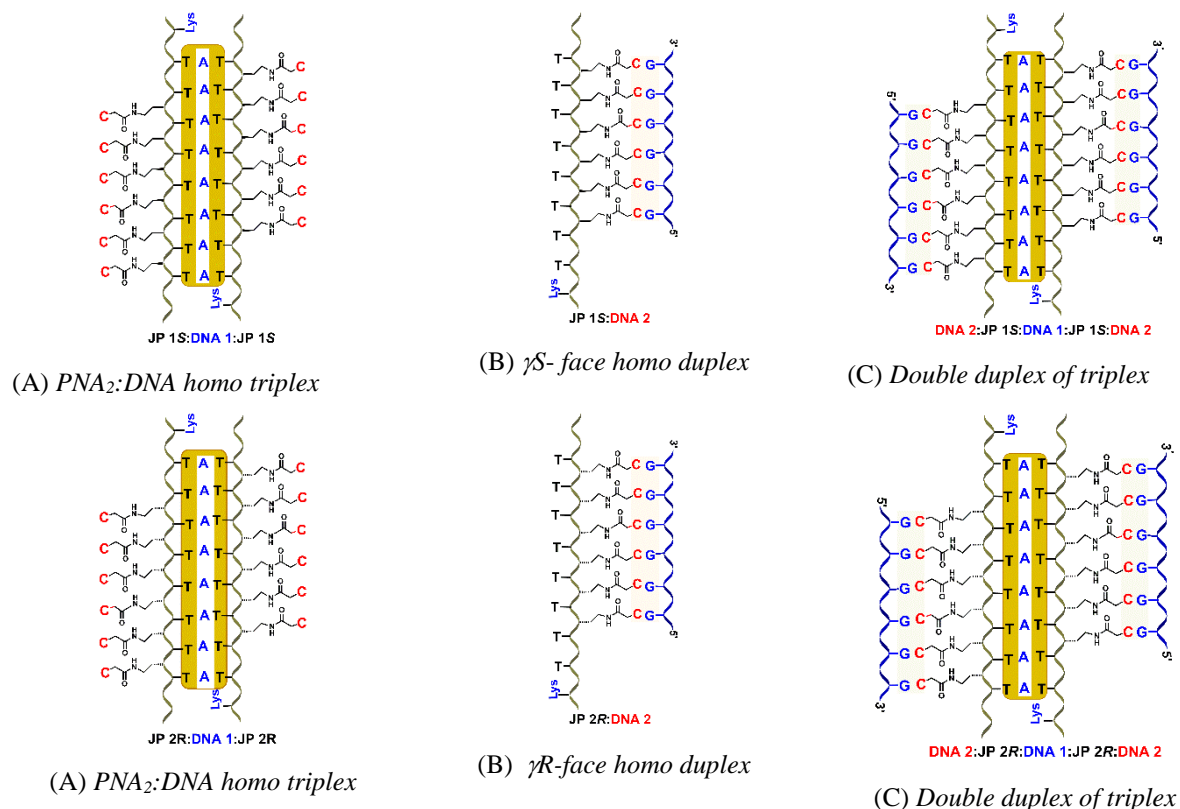
Figure 9 Synthetic strategy of mix Janus PNA (JP 5) on solid support

Control PNA analogues with nucleobase only on aminoethyl glyceryl C γ face were synthesized to study hybridization exclusively from C γ face and amide faces. The different Janus PNAs, *aeg*- γ PNA, and *aeg*-PNA oligomers synthesized as above by solid phase

synthesis were cleaved from the resin and purified by reverse phase HPLC followed by mass characterization using MALDI-TOF spectrometry.

Chapter 6: Biophysical Studies of *Janus* PNA Oligomers

This chapter reports on the biophysical evaluation of various kinds of *Janus* PNA oligomers. Temperature-dependent UV absorbance studies lead to determination of thermal stabilities T_m of *Janus* PNA:DNA duplexes, triplexes, double duplexes and duplex of triplex with complementary antiparallel DNA. The effect of various substitutions at C^γ -position on the conformation of *Janus* PNA:DNA duplexes, triplexes, double duplexes and duplex of triplex (Figure 10) is investigated by circular dichroism (CD) studies. The thermodynamic studies of *Janus* PNA:DNA duplexes and triplexes are investigated by isothermal titration calorimetry (ITC) experiments. Formation of *Janus* PNA derived *i*-motif and G-quadruplex are studied by UV- T_m and CD. It is demonstrated that *Janus* PNAs indeed bind to two cDNAs simultaneously to form double duplexes, which have different T_m s. The interesting observation was that the T_m s of duplexes in double duplexes are higher than the corresponding single duplexes and the duplex composed from C^γ -sidechain nucleobases were higher than that formed from the normal t-amide sidechain. There is a synergistic effect on stabilization of both duplexes in *Janus* PNAs.



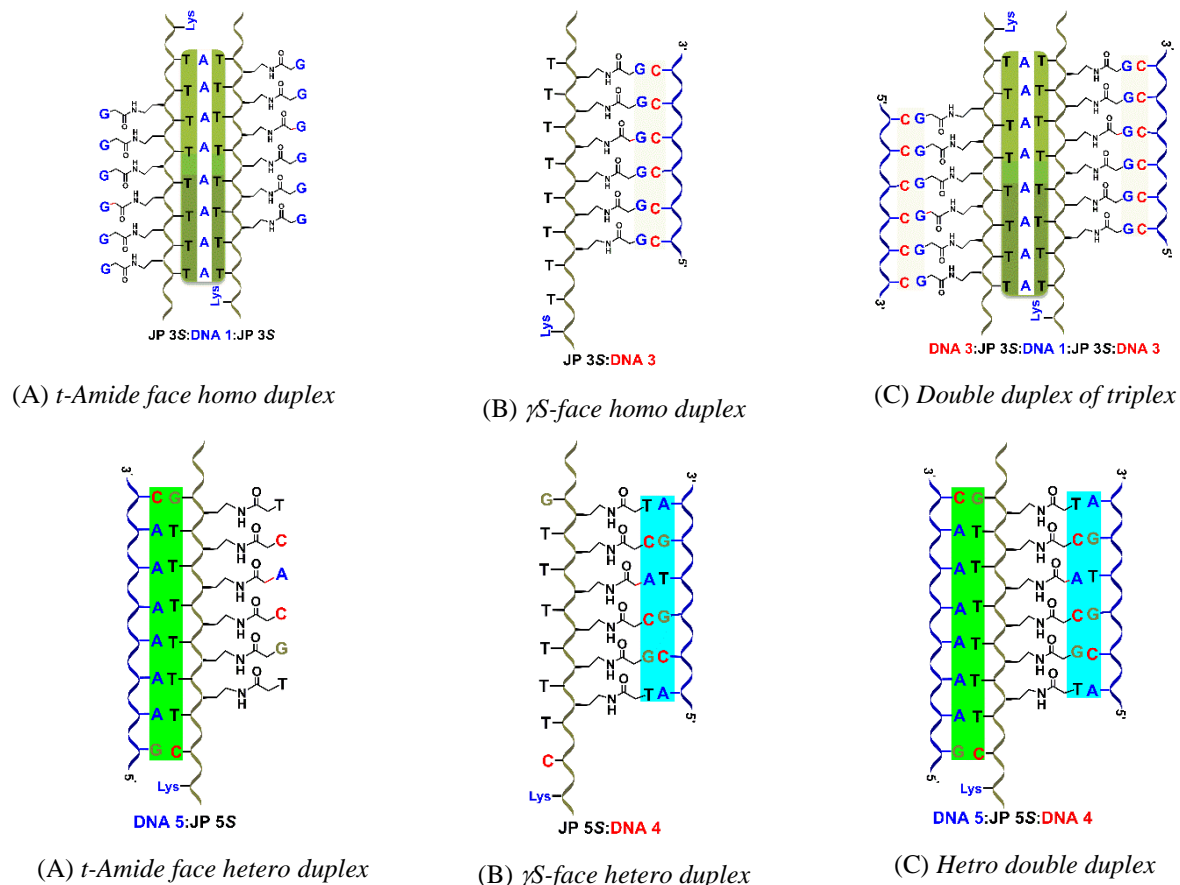


Figure 10 The various types of complexes with cDNA formed by different Janus PNAs

Janus PNAs having homo oligomeric C sequences on one face can form corresponding *i*-motif at acidic pH, with stability similar to that from *aeg*-PNA or DNA sequences (Figure 11). Similarly *Janus* PNA-G₆ also forms conventional G-tetraplexes with the thermal stability that are slightly higher than that of tetraplexes from *aeg*-PNA-G₆.

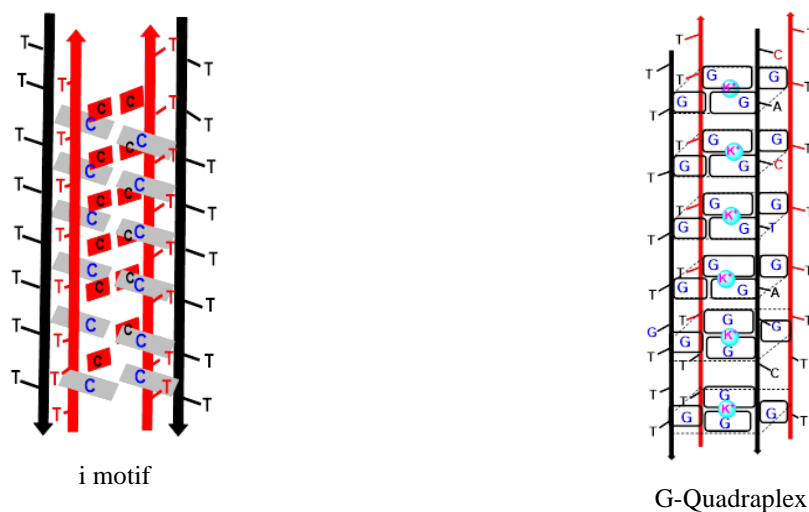


Figure 11 Schematic representation of *i*-motif and G-quadruplex.

Summary of Thesis

The triantennary (GalNAc)₃ and trivalent C^γS-substituted GalNAc [C^γ(S)-GalNAc-T] moieties were synthesized from commercially available D-(+)-galactosamine. These GalNAc moieties were incorporated at N-terminus of *aeg* PNA oligomers by solid-phase synthesis. By coupling with carboxyfluorescein to obtain carboxyfluorescein tagged PNA oligomers, purified by RP-HPLC and characterized by MALDI-TOF spectrometry.

- The GalNAc₃ modified PNAs, tri-antennary GalNAc₃ and trivalent [C^γ(S)-GalNAc-T]₃-PNA oligomers formed stable duplexes with complementary DNA and RNA.
- Tri-antennary GalNAc₃-PNA showed an increased thermal stability of DNA duplexes with a ΔT_m of about +5 °C, whereas [C^γ(S)-GalNAc-T]₃-PNA showed a ΔT_m of ~ +3 °C compared to unmodified PNA duplex with DNA. These findings suggest that even though the GalNAc₃ modifications are large and bulky at N-terminus and at C^γ-positions, they do not adversely affect the stability of PNA:cDNA duplexes.
- The stability of tri-antennary GalNAc₃-PNA as well as trivalent [C^γ(S)-GalNAc-T]₃-PNA with complementary RNA duplexes showed slightly lesser stability than that of unmodified PNA:RNA duplexes. However, both DNA and RNA duplexes with GalNAc₃-PNA were more stable than that from unmodified PNA:DNA duplexes. With melting temperatures above 50 °C for all three duplexes, the effect of destabilization induced by the GalNAc₃ modifications on the PNA:RNA duplexes are not significantly large.
- CD studies showed that conjugation of bulky tri-antennary GalNAc₃ and trivalent [C^γ(S)-GalNAc-T]₃-PNA at N-terminus of PNA did not affect the conformation of the duplexes they formed with complementary DNA or RNA. The CD signatures were similar to those of the unmodified PNA:DNA and PNA:RNA duplexes.
- Fluorescence studies for single stranded fluorescently labeled PNA as well as their duplexes with complementary DNA showed that the fluorescence intensity for PNA:DNA duplexes were slightly lesser than that of corresponding single stranded PNA.
- The trivalent *Cf*-[C^γ(S)-GalNAc-T]₃-PNA was found to be better than triantennary *Cf*-GalNAc₃-PNA showing enhanced cell penetration in HepG2 cell line than the unmodified PNA.

- The selectivity of the GalNAc₃-PNA to enter specifically hepatocytes HepG2 cells expressing ASGR receptor is demonstrated by its inability to get internalized by HEK293 cell line.

Janus PNAs that have nucleobases (A/G/C/T) linked to C^γ(S/R)-side chains via amide linkage were designed to create new and novel PNA backbones capable of binding to two cDNAs simultaneously on its both faces (amide/C^γ). These C^γ(S/R)-substituted *aeg*-PNA monomers were assembled to various *Janus* PNA oligomers by solid phase synthesis, purified by RP-HPLC and characterized by MALDI-TOF spectrometry.

- The thermal stability of duplexes of *Janus* PNA oligomers with each face complementary DNA and double duplexes generated by simultaneous binding of complementary DNAs on both faces were investigated by temperature dependent UV-visible spectroscopy. The T_m of C^γ face duplex was always higher than the amide face duplex both in individual duplexes as well as in double duplexes. The homo *Janus* PNA oligomers exhibit double duplex or triplex.
- The duplex from each face (amide/C^γ) show characteristic CD bands. Double duplexes of *Janus* PNA can be generated by sequential addition of each face complementary DNA and the order of addition does not matter. The final double duplex formed has the same conformation as seen by CD.
- The *Janus* PNAs with oligo C's form tetrameric i-motif similar to that of *aeg*-PNA-C_n and DNA C_n oligomers at acidic pH while become duplexes. The PNA/*Janus* PNA-G_n oligomers also form G-quadruplexes.
- The implication of present work for novel biological and material applications are discussed in the thesis.

Janus PNAs are a new class of nucleic acid analogues that can simultaneously bind complementary DNA/RNA/PNA strands from two sides by canonical Watson-Crick base pairing and have great potential for new material and biological applications.

Reference

1. Wittrup, A.; Ai, A.; Liu, X.; Hamar, P.; Trifonova, R.; Charisse, K.; Manoharan, M.; Kirchhausen, T.; Lieberman, J. *Nat. Biotechnol.* **2015**, *33*, 870–876.
2. Dalton, S. R.; Wiegert, R. L.; Casey, C.A. *Liver Int.* **2003**, *23*, 484-491.
3. Nielsen, P. E.; Christensen, L. *J. Am. Chem. Soc.* **1996**, *118*, 2287-2288.

Abstract

4. Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F. *Nucleic Acids Res.* **2014**, *42*, 8796–8807.
5. Prakash, T. P.; Yu, J.; Migawa, M. T.; Kinberger, G. A.; Wan, W. B.; Ostergaard, M.E.; Carty, R. K.; Vasquez, G.; Low, A.; Chappell, A. *et al. J. Med. Chem.* **2016**, *59*, 2718–2733.
6. Nair, J. K.; Willoughby, J. L.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Kandasamy, P.; Kel'in, A.V.; Milstein, S. *J. Am. Chem. Soc.*, **2014**, *136*, 16958–16961.

Chapter 1

Introduction to Peptide Nucleic Acids

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).

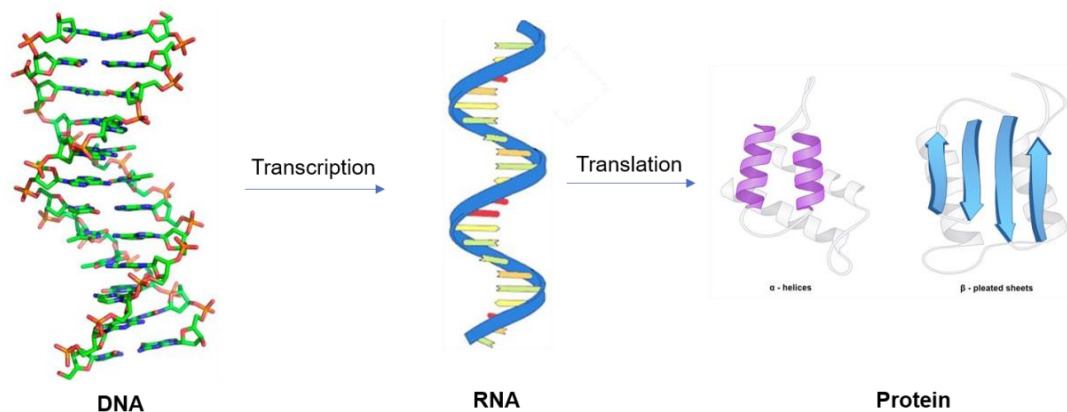


Figure 1.1 Biomacromolecules involved in the flow of genetic information

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 sugar¹ and a phosphate group. Both DNA and RNA comprise of four nucleobases: adenine (A) and guanine (G), which are purines, and cytosine (C), thymine (T, in DNA), and uracil (U, in RNA), which are pyrimidines (Figure 1.2 B). The structures of DNA and RNA differ from each other in their sugar: DNA contains a deoxyribose unit, while RNA has a ribose sugar. This has a tremendous effect on the 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, but 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.² In these linear copolymers, 3'-5' phosphodiester bonds link successive β -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 and G:C) commonly known as Watson-Crick base pairs to form an 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 1.2).

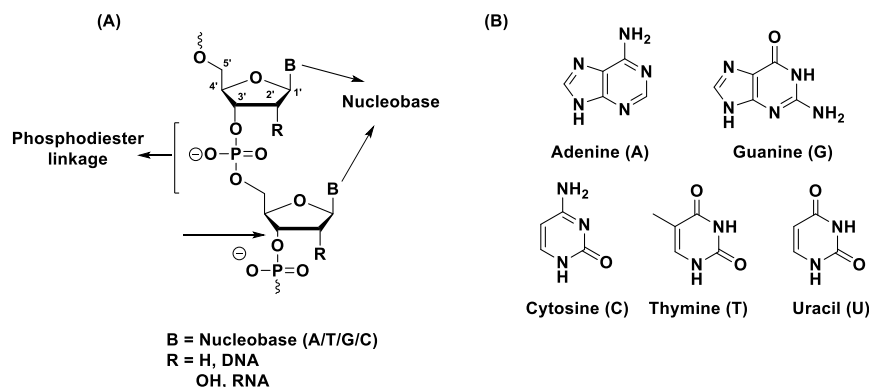


Figure 1.2. Chemical structures of DNA and RNA (A)¹ and nucleobases (B)¹

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* tautomer of the bases that establish 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 the 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).

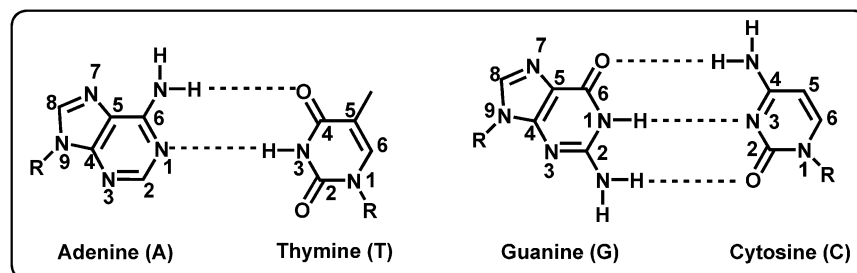


Figure 1.3. Watson-Crick hydrogen bonding for A:T and G:C base pairs³

Other important hydrogen bonding pairs are the Hoogsteen⁴ (HG) and Wobble⁵ base pairs. Hoogsteen base pairing is not isomorphous with Watson-Crick base pairing because they have an 80° angle between the glycosidic bonds and 8.6 \AA separation of the anomeric carbons (Figure 1.4 a, b). Hoogsteen hydrogen bonding has importance in triple helix formation and in protein-DNA complexes. Wobble base pairing involves non-Watson-Crick pairing between two nucleotides in RNA molecules. Wobble base pairs are fundamental in RNA secondary structure and are important for the proper translation of genetic code (Figure 1.4c, d).

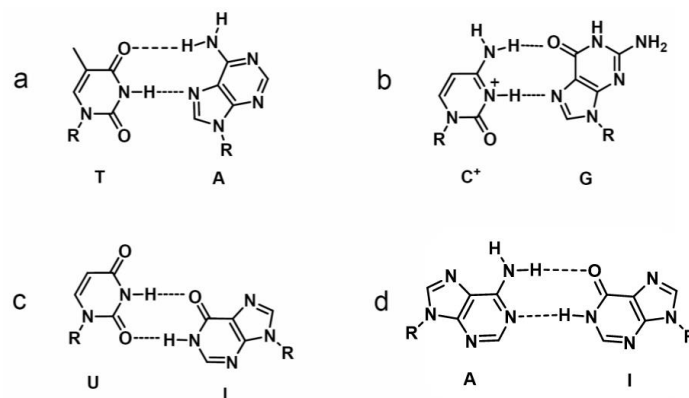


Figure 1.4. Hoogsteen base pairing⁴ (a & b) and Wobble base pairing⁵ (c & d)

1.3 Secondary structures of nucleic acids

DNA exists in various possible conformations like A-DNA, B-DNA, and Z-DNA, although, the most common is the B-DNA which is a right-handed double helix.⁶ It has a wide and deep major-groove with a narrow and shallow minor-groove wherein the bases lie perpendicular to the helical axis. A-DNA also forms a right-handed helix where the major groove is deep and narrow while the minor groove is broad and shallow. In both A and B forms of DNA, the Watson-Crick base pairing is maintained along with *anti*-glycosidic conformation. Z-DNA is a left-handed double helix and the most favored conformation in alternating G-C sequences. The left-handed helix for Z-DNA is a result of a switch in the glycoside bond with a *syn* conformation (Figure 1.5). Moreover, the sugar conformations are different in the two forms, with the B-DNA form showing C2'-endo pucker and the A-DNA form exhibiting C3'-endo sugar-pucker (Figure 1.6).

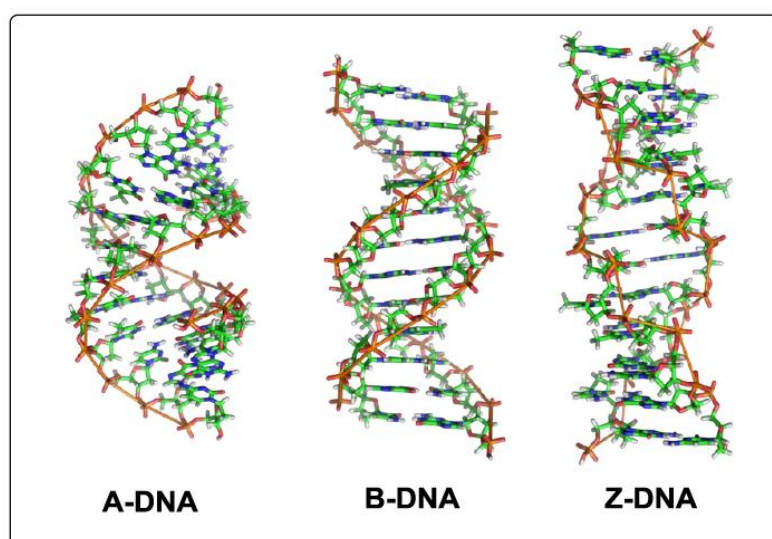


Figure 1.5. Molecular models of A-DNA, B-DNA and Z-DNA⁶

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

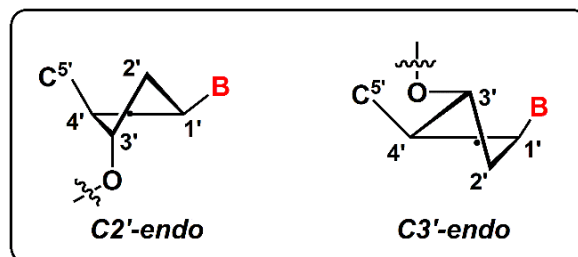


Figure 1.6. Structures of *C2'-endo* and *C3'-endo* sugar pucker⁶

1.4 Applications of nucleic acids

Although the double helical structure of DNA was elucidated by Watson and Crick in 1953, only recently the molecular and supramolecular properties of DNA has been used to design nucleic acid based drugs, diagnostics, and new materials⁷ through explorations of chemical space in both modified DNA and RNA structures accompanied by expanded functionality. The recent decades have seen DNA evolve from just “genetic” material to a “generic” material^{8,9} for several applications from the intention of antisense activity of nucleic acid analogues, covalent conjugation of functional ligands and the explosion of DNA as a nanoscale construction material.

1.4.1 Disease diagnosis

Nucleic acids are natural biomarkers that can be used as probes for disease diagnosis based on the fact that they recognize and bind to specific complementary sequences of nucleic acids (i.e. DNA or RNA). Nucleic acid-based diagnostics detect the presence of a pathogen either by directly detecting the presence of DNA or RNA nucleic acids in the host or by first amplifying the pathogen DNA or RNA. In the case of infectious diseases, nucleic acid-based diagnostics detect DNA or RNA from the infecting organism. For non-infectious diseases, nucleic acid-based diagnostics may be used to detect a specific gene or the expression of a gene associated with disease. Nucleic acid-based diagnostics are used to diagnose a wide range of conditions, including cancer, genetic markers associated with a high

risk of cancer, and genetic diseases (e.g. cystic fibrosis). Common nucleic-acid based diagnostic techniques used to diagnose infectious diseases are PCR/RT-PCR, isothermal amplification, hybridization, and sequencing. Nucleic acid-based diagnostics are also used to detect infectious diseases, such as anthrax, *Clostridium difficile* (a common hospital acquired bacterial infection), chlamydia, and gonorrhea and are useful for diagnosis of leishmaniasis, tuberculosis, and HIV.

1.4.2 Therapeutics

Nucleic acid therapeutics include antisense oligonucleotides, aptamers and small interfering RNAs, and are typically considered in cases where specific inhibition of the function of a particular gene involved in disease is thought to be therapeutically desirable. Several steps in the process of gene expression may be modulated, including the transcription, RNA splicing and translation by various mechanisms. In the last two decades, therapeutic nucleic acid technology has developed as a promising tool to fight human diseases without secondary effects often observed with nucleoside and nucleotide analogues. This novel approach to drug design was more promising than rational drug design in the sense that it can be used for specific control of gene expression at the level of nucleic acid with increased specificity against malignant cells or viral entities and reduced side effects. Traditional small molecule therapeutic agents, such as nucleoside and nucleotide analogues work by interfering or modulating the function of enzymes, receptors, transport or structural proteins, whereas the therapeutic oligonucleotides alter the flow of genetic information at the DNA, mRNA or even protein level, and in many cases stimulate immune responses resulting in the suppression of disease-associated gene products.¹⁰

1.4.2a Antisense and antigene oligonucleotides

An antisense oligonucleotide (ASO) is a short strand of oligonucleotide analogue that hybridizes with the complementary mRNA in a sequence-specific manner via Watson-Crick base pairing and downregulates the target protein expression. Antisense technology was first reported by Zamecnik and Stephenson in 1978 in Rous sarcoma virus (RSV).¹¹ On the basis of mechanism of action, two classes of antisense oligonucleotide can be discerned: (a) the RNase H-dependent oligonucleotides, which induce the degradation of mRNA; and (b) the steric-blocker oligonucleotides, which physically prevent or inhibit the progression of splicing or the translational machinery. RNase H is a ubiquitous enzyme that hydrolyzes the RNA strand of an RNA/DNA duplex. Oligonucleotide-assisted RNase H-dependent reduction

of targeted RNA expression can be quite efficient, reaching 80–95% down-regulation of protein and mRNA expression. RNase H-dependent oligonucleotides can inhibit protein expression when targeted to virtually any region of the mRNA. An antigene drug is a triplex-forming oligonucleotide that recognizes and attaches directly to a specific DNA sequence and prevents transcription of the blocked DNA sequence into mRNA.¹² The strategy of blocking transcription and inducing specific mutations, both *in vitro* and *in vivo* has been successfully demonstrated.¹³ The pre-requisites for an oligonucleotide to be an effective and successful therapeutic candidate are the following: (i) easy synthesis in ample amounts, (ii) *in vivo* stability towards cellular degrading enzymes, (iii) ability to penetrate into cells, (iv) retention by the target cell, (v) ability to interact with their cellular targets (DNA/RNA), (vi) no non-specific interaction with other macromolecules. Naturally occurring DNA or RNA can act as therapeutic agents but they lack binding affinity and get degraded by nucleases. Numerous chemical modifications have been developed to overcome limitations and to make these oligonucleotides successful therapeutic agents.

1.4.2b DNAzymes

DNAzymes are a catalytically active class of antisense reagents discovered in the 1990s. These are single stranded DNA molecules that bind to their target mRNA by Watson–Crick base pairing and cleave the link between an unpaired purine (A or G) and a paired pyrimidine (C or U) in the RNA by the cation-dependent domain of the DNAzyme.¹⁴ DNAzymes are relatively small and inexpensive to synthesize and are resistant to nuclease degradation. DNAzymes have been used as inhibitory agents in a variety of experimental disease settings, suggesting their possible clinical utility. The first demonstration of DNAzyme activity in animals was in 1999 when a DNAzyme against the zinc finger transcription factor Egr-1 inhibited restenosis in balloon-injured rat carotid arteries.¹⁵ DNAzyme directed against vascular endothelial growth factor receptor 2 was confirmed to be capable of tumor suppression by blocking angiogenesis upon intratumoral injections in mice.¹⁶ Thus, DNAzymes may be useful potentially as interventional tools in common disorders.

1.4.2c Gene therapy

Gene therapy is an experimental technique that uses genes to treat or prevent diseases and allow doctors to treat a disorder by inserting a gene into a patient's cells instead of using drugs or surgery. Researchers are testing several approaches to gene therapy, including:

- Replacing a mutated gene that causes disease with a healthy copy of the gene.
- Inactivating, or “knocking out,” a mutated gene that is functioning improperly.
- Introducing a new gene into the body to help fight a disease. Although gene therapy is a promising treatment option for a number of diseases, the technique remains risky and is under development to make sure that it is safe and effective.

1.4.2d RNA interference (RNAi)

RNA interference (RNAi) is an evolutionarily conserved mechanism by which small non-coding double-stranded RNA (dsRNA) silences gene expression, either by inducing the sequence specific degradation of complementary mRNA or by inhibiting translation via endogenous cellular machinery. The concept of RNAi was introduced in 1998 and was employed to inhibit the expression of homologue genes in *C. elegans*.¹⁷

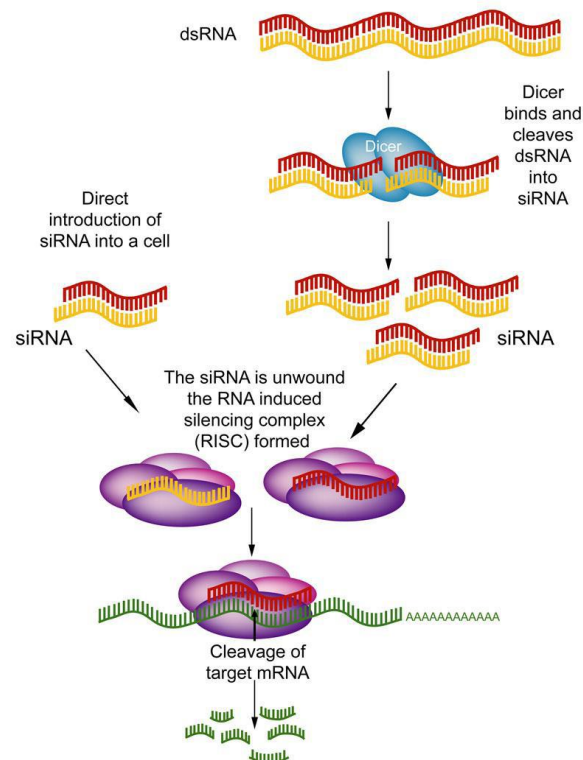


Figure 1.7. Gene silencing by RNA interference¹⁷

RNAi effect is triggered by long pieces of dsRNA, which are cleaved into fragments known as siRNA (21–23 nucleotides long) by the enzyme Dicer.¹⁸ In the cytoplasm of the cell, siRNA is incorporated into a protein complex called the RNA induced silencing complex (RISC), which unwinds the siRNA, after which the sense strand (or passenger strand) of the siRNA is cleaved.¹⁹ The activated RISC containing the antisense strand (or

guide strand) of the siRNA then selectively degrades mRNA that is complementary to the antisense strand.²⁰ The cleavage of mRNA occurs at a position between nucleotides 10 and 11 on the complementary antisense strand, relative to the 5'-end.²¹ The activated RISC complex can further move on to destroy additional mRNA targets (Figure 1.6).

1.4.2e MicroRNAs

MicroRNAs (miRNAs) are another class of small non-coding RNAs that have been found to regulate gene expression at the post-transcriptional level. Mature miRNAs originate from longer transcripts, called primary miRNAs (pri-miRNAs) that are transcribed in the cell nucleus by polymerase II.¹⁶ In the nucleus, pri-miRNAs are processed into pre-miRNAs by the microprocessor complex, which consists of the RNase III enzyme Drosha.¹⁷ The pre-miRNAs are then transported from the nucleus to the cytoplasm via Exportin 5 complex.¹⁸ These pre-miRNAs are additionally processed by the RNase III enzyme Dicer to generate miRNA.¹⁹ These miRNAs are incorporated into a RISC-like complex which can either lead to mRNA cleavage or translational repression.²⁰ Hence, a single miRNA can regulate multiple genes due to its successful gene silencing ability even with partial complementarity between the miRNA and its target.

1.4.2f Catalytically active RNA molecules (ribozymes)

Another strategy for drug development is by targeting the transcriptional process using ribozymes, a unique class of RNA molecules that not only store information but also possess catalytic activity.²¹ In the early 1980s, Cech and coworkers discovered that RNA molecules are capable of catalysing reactions even in the absence of any protein component²² and were named ribozymes. The hammer-head and hairpin ribozymes, among various ribozymes, have been extensively studied due to their small size and high cleavage efficiency.²³ Chemical modification of the RNA molecules is required to withstand nucleolytic degradation. Various clinical trials are in progress to evaluate the potential of ribozymes to fight cancer, e.g. angiozyme, is in phase II trial for treatment of metastatic colorectal cancer,²⁴ while herzyme (Zinzyme) is in phase I clinical trial for treatment of breast and ovarian cancer.²⁵

1.4.3 Nucleic acid nanotechnology

Inspired by nature, researchers over the past four decades have explored nucleic acids as convenient building blocks to assemble novel nanodevices. Nucleic acid nanotechnology has come a long way since its inception a quarter century ago. In 1982, Seeman first proposed

using branched DNA building blocks to construct ordered arrays, which dramatically accelerated progress in nucleic acid nanotechnology by increasing the simplicity, precision, and fidelity of the design principles available for generating spatially addressable nanoscale structures.²⁶ Nucleic acid scaffolds are nanostructures that may be functionalized for ordering and arraying materials with nanometer precision. Scaffolds were originally formed by combining different short double-stranded DNA domains joined in a programmable fashion using single-stranded DNA overhangs known as “sticky ends”.²⁷ These sticky ends can be used as molecular walkers or as an adhesive to bring other macromolecules like proteins or lipids to form nucleic acid hybrids. Reduction in errors in self-assembly, extension of two-dimensional self-assembly to three dimensional, and scale up in self-assembly will help in the construction of nucleic acid based molecular devices and patterned superstructures on nanometer-scale.

1.5 Chemical modifications of DNA

To improve the affinity and specificity of oligonucleotide recognition, enhance the cell membrane permeability and resistance to nuclease degradation, several chemical modifications of DNA have been attempted²⁸(Figure 1.8). In general, three types of chemical modifications of DNA can be distinguished as

- i) Modified phosphate linkages
- ii) Modified sugars
- iii) Altered sugar-phosphate backbone

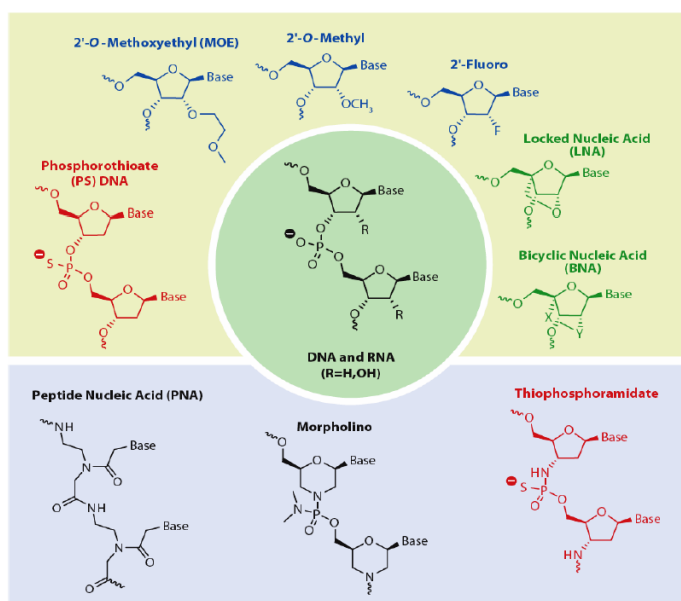


Figure 1.8. Examples of chemical modifications of oligonucleotides²⁸

1.5.1 Phosphate backbone analogues

The modifications of phosphate moiety resulted in the development of phosphorothioates²⁹ (1), methyl phosphonates³⁰ (2), phosphoroamidates³¹ (3), phosphotriesters³² (4) and borano phosphonates³³ (5) (Figure 1.9). Phosphorothioate (PS)-containing oligonucleotides differ from natural nucleic acids in that one of the nonbridging phosphate oxygen atoms is replaced with a sulfur atom. Phosphorothioate oligonucleotides are one of the earliest and most widely used backbone modifications for antisense drugs. The substitution of sulfur for oxygen in PS-oligos greatly increases stability to nucleolytic degradation. PS-oligos are able to efficiently elicit the RNase H cleavage of target mRNA, which is critical in the mechanism of action of antisense oligonucleotides.³⁴ Their binding to plasma proteins protects them from rapid renal excretion and is responsible for increased serum half-life. Vitravene (Fomiversen) is the first FDA approved antisense drug, which is based on PS-oligos. Some of the antisense drugs which are under various stages of clinical trials incorporate PS-modifications. The binding of PS-oligos to certain proteins, particularly those which interact with polyanions such as hairpin-binding proteins, prove to be their major drawback.^{35,36,37}

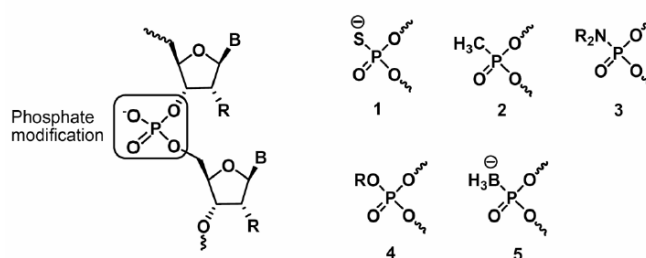


Figure 1.9. Structure of alternative phosphate linkages

Other backbone modifications (methylphosphonates, phosphoroamidates etc.) of oligonucleotides have been less successful at improving the oligonucleotides properties. In methylphosphonates, one of the nonbridging oxygen atoms is replaced with a methyl group and is neutral in charge. Although it provides high nuclease resistance, it does not induce RNase H activity. Additionally, increasing the number of methylphosphonate units in an oligomer leads to loss of affinity towards the target mRNA and to poor water solubility.

1.5.3 Sugar modifications

This class of ONs includes those containing nucleotides with alkyl substitutions at the 2'-position of ribose sugar (Figure 1.10). Organization of the sugar into RNA-like *C3'*-endo pucker increases the binding affinity of these ONs towards the complementary RNA.

Furthermore, the 2'-substituent in an oligonucleotide increases the steric bulk near the 3'-phosphate and makes it difficult for degrading enzymes to cleave the phosphodiester bond. The increase in binding affinity of 2'-modified ONs is energetically driven by the electronegative substituent at the 2'-position. Among all 2'-modified ONs, 2'-fluoro modification (Figure 1.10a) imparts the highest binding affinity towards the target RNA. The 2'-O-methyl (Figure 1.10b) group enhances the binding affinity to a lesser extent than the 2'-fluoro modification, but it imparts a substantial degree of nuclease resistance to the corresponding oligonucleotide. The 2'-O-Methoxyethyl (MOE) (Figure 1.10c) modification is currently the most advanced in the 2'-modified series and many antisense drugs having this modification have entered clinical trials. Unfortunately, the lack of RNase H activity restricts the use of 2'-modified oligonucleotides for antisense purpose. To overcome this drawback, the gapmer strategy has been used where regions of 2'-modified residues flank a central DNA region of the oligonucleotide.

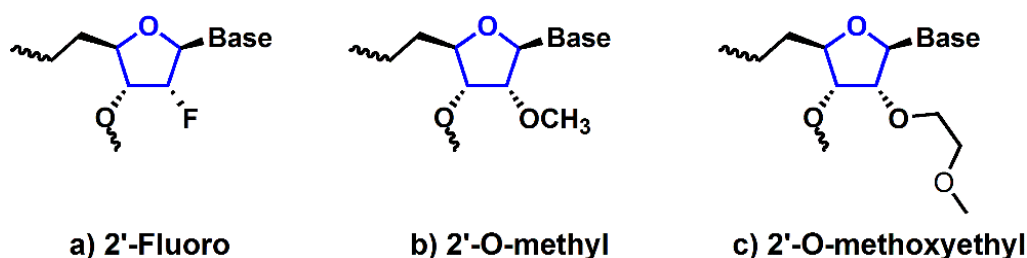


Figure 1.10 2'-modified oligonucleotides

1.5.4 Sugar-phosphate backbone modifications

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

1.5.4a Locked nucleic acid (LNA). These oligonucleotides developed by Jesper Wengel *et al.*³⁸ in 1998 are one of the most promising class of chemically modified ONs. These analogs are bicyclic systems that contain a methylene bridge that connects the 2'-O- of the ribose with the 4'-C, locking the ribose in *C3'-endo* conformation (Figure 1.11). Introduction of LNA into a DNA sequence induces a conformational change of the DNA:RNA duplex towards the A-type helix³⁹ but prevents the RNase H cleavage of the target RNA. LNAs and LNA:DNA chimeras have been shown to efficiently inhibit gene expression when targeted to

a variety of regions within the luciferase mRNA.⁴⁰ LNA shows remarkably increased hybridization properties relative to a DNA:RNA duplex and improves nuclease resistance. Few analogs of LNA have been reported, which have improved activity and/or toxicity profiles in animals.^{41,42} Thus, LNA offers attractive properties, such as nuclease stability, high target affinity, potent biological activity, and lack of acute toxicity.

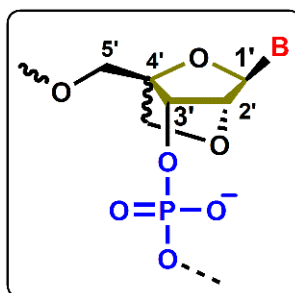


Figure 1.11. Locked nucleic acid (LNA)

1.5.4b Morpholino oligonucleotides (MF). Phosphorodiamidate morpholino oligonucleotide has morpholine ring replacing the furanose ring in DNA/RNA. It has the phosphorodiamidate linkage which connects the morpholine nitrogen atom with the hydroxyl group of the 3'-side residue (Figure 1.12).⁴³ Morpholino ONs are non-ionic, and therefore unlikely to have unwanted electrostatic interactions with nucleic acid binding proteins. These ONs are stable to nucleases and have a similar affinity as in DNA:DNA duplexes. However, morpholino ONs do not activate the RNase H and are primarily used in translation arrest or in other steric blocking mechanisms, such as alteration of splicing.^{44,45} Eteplirsen (EXONDYS 51) is the first morpholino oligo-based antisense drug approved by FDA for treatment some forms of Duchenne muscular dystrophy (DMD).

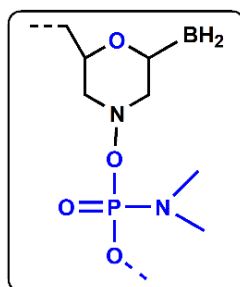


Figure 1.12 Morpholino oligonucleotide (MF)

1.5.4c N3'-P5' phosphoroamidates (NPs). N3'-P5' phosphoroamidate (NPs) is a modified phosphate backbone, in which the 3'-oxygen of the deoxyribose ring is substituted with an amino group (Figure 1.13). NPs are resistant to nucleases and have high affinity towards a complementary RNA strand but do not activate the RNase H.⁴⁶ The sequence specificity of

phosphoramidate-mediated antisense effects by steric blocking of translation initiation has been demonstrated in cell culture and *in vivo*.⁵⁵

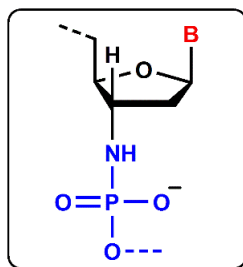


Figure 1.13. $N3'$ - $P5'$ phosphoramidates (NPs)

1.6 Peptide nucleic acids (PNA)

Peptide nucleic acids (PNAs) are a radically different class of oligonucleotide modifications that contain a peptide link instead of phosphate link. PNA was first introduced by Peter Nielsen and coworkers in 1991.⁴⁸ PNAs are highly resistant to degrading enzymes like proteases and nucleases, exhibit high binding affinity towards target DNA/RNA, but do not activate RNase H and, as such, have been used primarily in translation inhibition and splicing modulation antisense mechanisms. PNAs are known to be non-toxic, as they are uncharged molecules with low affinity for proteins that normally bind nucleic acids.

PNAs are DNA analogs where the sugar-phosphate backbone is replaced with a pseudopeptide backbone in the form of 2-aminoethyl-glycine linkage. Nucleobases are attached through a methylene carbonyl linker to this backbone at the amino nitrogen. The PNA backbone is constituted by six atoms for each repeating unit and a two-atom spacer between the backbone and the nucleobase, similar to the natural DNA (Figure 1.14).⁴⁹ PNA was originally designed and developed as a mimic of DNA recognizing, major groove binding, triplex-forming oligonucleotide.^{50,51} However, the polyamide backbone of PNA has proven to be a surprisingly good structural mimic of the sugar-phosphate backbone of nucleic acids.

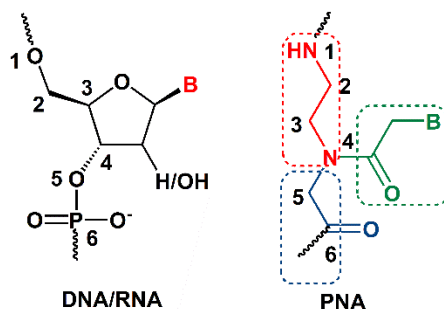


Figure 1.14 Chemical structures of DNA/RNA and PNA

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

1.6.1 PNA structure

PNA binds to a complementary DNA/RNA through classical Watson-crick base-pairing mechanism. The PNA bases form a helical π -stack similar to DNA but the smaller twist of the PNA double helix and the larger π -overlap between the neighboring bases makes it different from the DNA. NMR methods and X-ray crystallography have been used to derive the three-dimensional structures of the major families of PNA complexes including PNA:RNA,⁵² PNA:PNA,⁵³ PNA:DNA duplexes⁵⁴ and PNA₂:DNA triplex⁵⁵ (Figure 1.15). Although PNA oligomer, to some extent, is able to structurally adapt to the oligonucleotide complement, it also has a preferred structure of its own, termed the '*P-form*' helix.

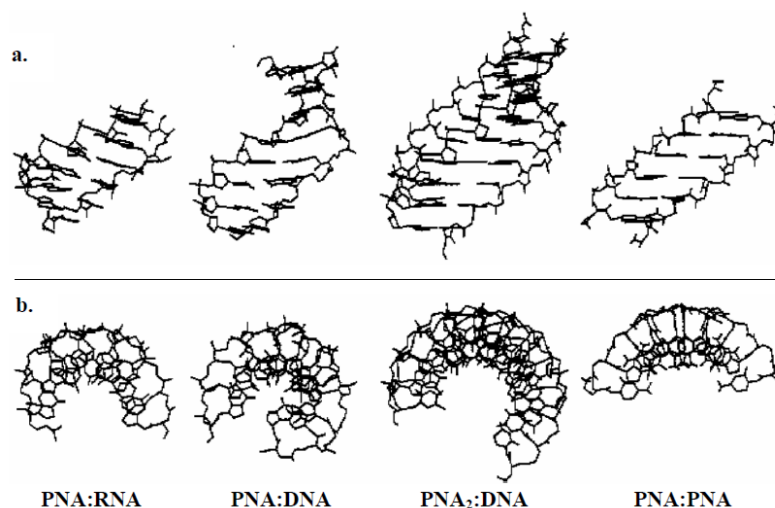


Figure 1.15. Structures of various PNA complexes shown in a) side view and b) top view⁵⁶

The *P-form* helix constitutes a very wide PNA duplex (28 Å diameter) with an accordingly large base pair helical displacement and a very large pitch (18 bp). A canonical *B-form* helix which is typical for DNA duplexes has a diameter of 20 Å and a pitch of 10 bp

per turn. A canonical *A-form* helix, typical of RNA duplexes, also has a diameter of 20 Å but a pitch of 11 bp per turn, and the base pairs are tilted 20° relative to the helix axis. Also, the base pairs are displaced away from the helix leaving a central 'tunnel' in the helix, analogous to that seen in the *P-form*. These structures suggest that PNA can adapt well to its nucleic acid partner, as the RNA strand in the PNA:RNA duplex is essentially in *A-form* conformation, whereas PNA:DNA duplex adopts a *B-form* conformation.

1.6.2 Physico-chemical properties of PNA

PNA has proved itself to be a promising antisense or antigene agent on the basis of its superior properties, such as highly sequence-specific binding to the complementary DNA/RNA targets, high biological and chemical stability, and high mismatch discrimination.

1.6.2a Duplex formation with complementary oligonucleotides. PNA was originally designed for sequence-specific targeting of double-stranded DNA via major groove recognition. PNA hybridizes to complementary oligonucleotides obeying Watson-Crick base pairing rule. In DNA:DNA duplexes, the two strands are always in antiparallel orientation (with the 5'-end of one strand opposed to the 3'-end of the other). However PNA:DNA adducts can be formed in two different orientations, arbitrarily termed *parallel* and *antiparallel* (Figure 1.16). Both adducts are formed at room temperature, with the antiparallel orientation showing higher stability.⁵⁷ This creates the possibility for PNAs to bind two DNA tracts of opposite sequence.

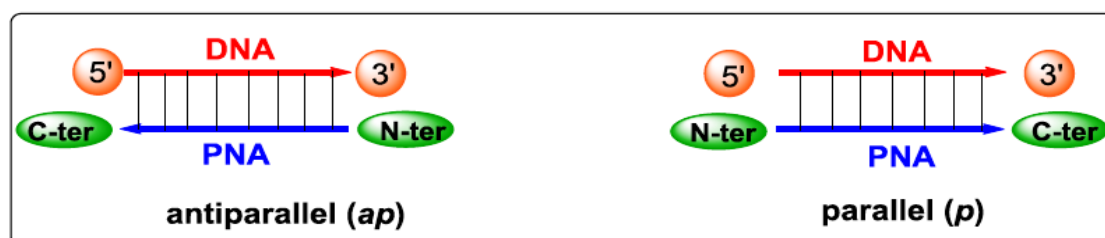


Figure 1.16. Antiparallel and parallel modes of PNA:DNA duplex formation

1.6.2b Triple helix formation of PNA. Polypyrimidine PNAs are able to form stable adducts with complementary polypurine DNA, through the formation of PNA₂:DNA triplexes.⁵⁸ The base pairing in these complexes occurs via Watson-Crick and Hoogsteen hydrogen bonds. When only one PNA strand is used to form a PNA₂:DNA triplex, both strands are necessarily either antiparallel or parallel to DNA strand. When two different homopyrimidine PNA sequences are used, Watson-Crick PNA strand is oriented in antiparallel and the Hoogsteen strand is in parallel orientation to form a stable triplex with homopurine strand of DNA. The

sequence specificity of triple helix formation is based on the selectivity of formation of the intermediate PNA:DNA duplex, whereas binding of the third strand contributes only slightly to selectivity. The stability of structures enables PNA to perform strand invasion,^{59,60} a property uniquely shown by PNAs.

1.6.2c Strand invasion by PNA. The unique property of PNAs to displace one strand of DNA double helix to form strand invasion complexes,⁶¹ is a favorable attribute for their application as antigene agents. Four modes of binding for sequence-specific targeting of double-stranded DNA by PNA have been identified (Figure 1.17), three of which involve invasion of the DNA duplex by PNA strands.

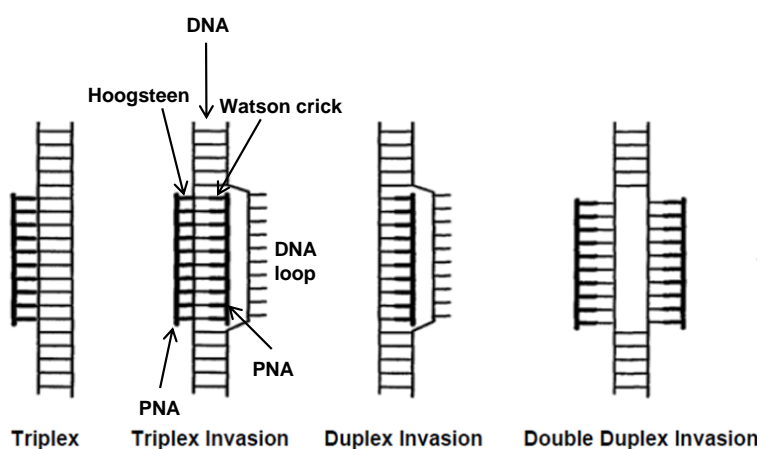


Figure 1.17. Various modes for binding of PNA to sequence specific targets in *dsDNA*⁶¹

It is possible for PNA (homopurine) single strand to either invade (duplex invasion) via Watson-Crick base pairing, or alternatively, invasion may be accomplished by two pseudo-complementary PNA strands, each of which binds to one of the DNA strands of the target (double duplex invasion). These pseudo complementary PNAs contain modified adenine and thymine nucleobases⁶¹ (Figure 1.18) that do not allow stable hybridization between the two complementary sequence PNAs, but does permit good binding to the DNA. The 'triplex invasion' requires a homopurine DNA target and complementary homopyrimidine PNAs that bind the purine DNA strand through combined Watson-Crick and Hoogsteen base pairing *via* formation of a very stable PNA₂:DNA triplex. For many applications, the two PNA strands are connected in a bis-PNA designed such that the one strand is antiparallel (W-C strand) and the other strand is parallel (H-strand) to the DNA target.

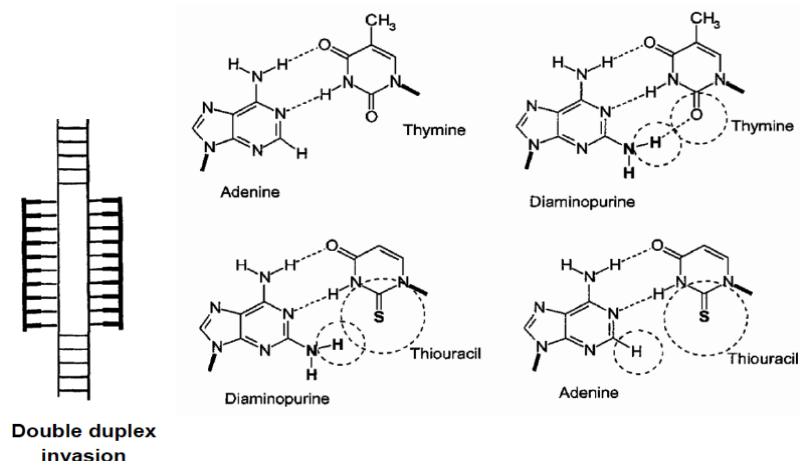


Figure 1.18 Double duplex invasion of pseudocomplementary PNAs⁶¹

1.6.2d G-Quadruplex and i-motif formation by PNA. DNA and RNA oligomers that contain multiple stretches of consecutive guanine (G) nucleotides are able to fold into a stable secondary structure known as G-quadruplex^{62,63} that is gaining increasing attention due to its implication in regulation of gene expression.⁶⁴ Balasubramanian *et al.*⁶⁵ have shown the formation of quadruplex composed entirely of PNA (Q-PNA) (Figure 1.19b). A homologous PNA (i.e. a PNA having the same sequence as the target) forms a stable PNA₂:DNA₂/RNA₂ hybrid quadruplex by disrupting a bimolecular DNA/RNA G-quadruplex (Figure 1.19a).⁶⁶ Ganesh *et al.*⁶⁷ demonstrated the formation of C-C⁺ tetraplex (i-motif) with unmodified PNAs TC₄ and TC₈ in acidic pH. The tetraplex formation was monitored using UV-thermal melting at 295 nm, show a reverse sigmoidal pattern, which is characteristic of C-C⁺ tetraplex formation (Figure 1.19c).

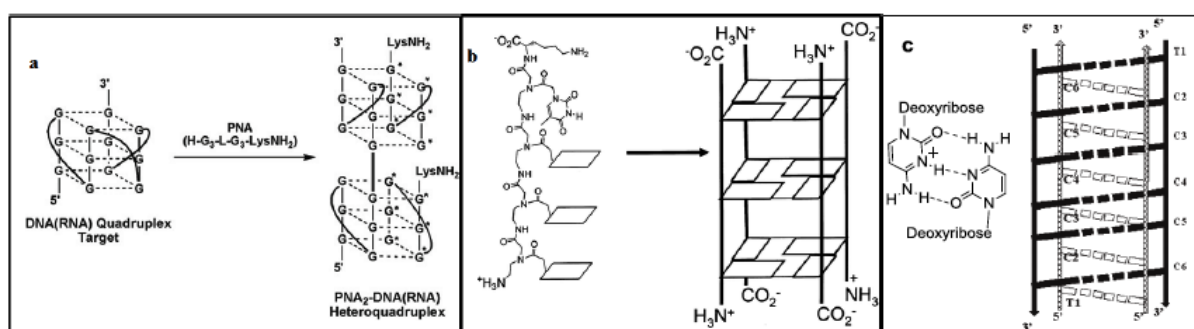


Figure 1.19 (a) Schematic drawing for PNA:DNA/RNA homologous quadruplex⁶⁶ (b) Schematic diagram for PNA-G4 quadruplex⁶⁵ (c) Schematic diagram of i-motif formation⁶⁷

1.6.2e Higher specificity towards target nucleic acid. PNA shows greater specificity in binding to complementary DNA. A PNA/DNA mismatch is more destabilizing than a mismatch in a DNA/DNA duplex. A single mismatch in mixed PNA:DNA (15-mer) sequence, destabilizes the duplex by 8-20 °C (15 °C on average), while in the corresponding

DNA:DNA duplex, a single mismatch destabilizes the duplex by 4-16 °C (11 °C on average).⁶⁸

1.6.2f Stronger binding independent of salt concentration. Another important consequence of the neutral backbone is that the T_m values of PNA:DNA duplexes are practically independent of salt concentration. In contrast, the T_m values of DNA:DNA duplexes are highly dependent on ionic strength.⁶⁹

1.6.2g Solubility of PNA. PNAs are charge-neutral compounds and hence have poor water solubility compared with DNA. Neutral PNA molecules have a tendency to aggregate to a degree that is dependent on the sequence of the oligomer. PNA solubility is also related to the length of the oligomer and to the purine/pyrimidine ratio.^{69,70} Some of the recent modifications, including the incorporation of positively charged lysine residues (carboxy-terminal or backbone modification in place of glycine), have shown improvements in solubility of PNA.

1.6.2h Cellular uptake of PNA. Although PNA binds to complementary DNA/RNA with high affinity, specificity and stability in biological fluids, the progress in the exploration of PNA as antisense/antigene agents and gene expression regulation has been hampered by their poor cellular uptake. Thus, efficient cellular delivery systems for PNAs are required if these are to be developed into antisense and antigene agents. However, a number of transfection protocols for PNA have been established like microinjection, electroporation, co-transfection with DNA, conjugation to lipophilic moieties, conjugation to cell penetrating peptides etc.⁷¹ To address the issues like poor cell penetration, solubility and ambiguity in binding orientation, various chemical modifications of PNA as described below.

1.6.3 Chemical modifications of PNA

The antisense and antigene potential of PNAs can be improved by enhancing the binding affinity for DNA and RNA by suitable conformational preorganization. One approach for improving DNA binding affinity is the design and synthesis of preorganized PNAs that prefer a right-handed helical conformation which can be achieved by adding of substituents to the backbone or by cyclization of the PNA backbone (Figure 1.20).^{72,73}

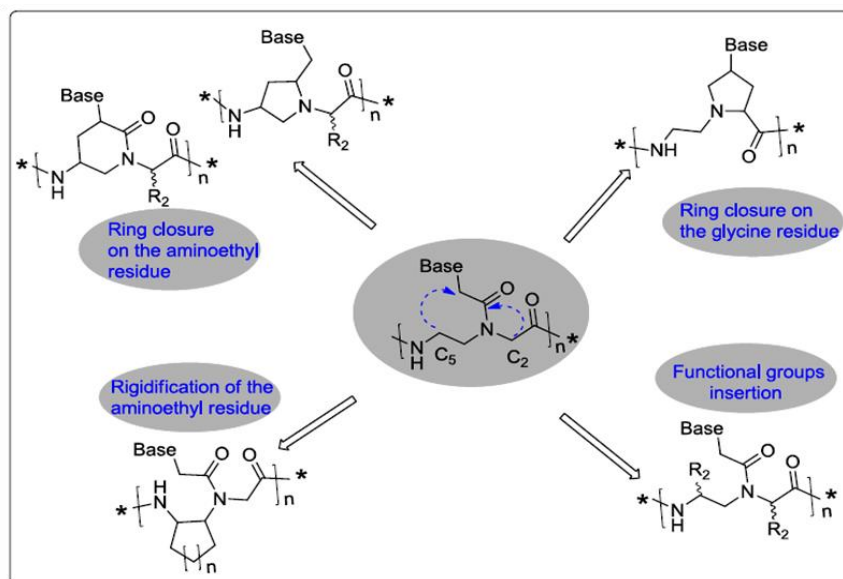


Figure 1.20 Strategies for inducing preorganization in the PNA structure

1.6.3a Preorganization of acyclic PNAs. The N-(2-aminoethyl) glycine unit in PNA backbone is a versatile motif for the modification of the PNA. Replacement of glycine by other α -substituted amino acids and substitution in ethylenediamine results in chiral PNAs that bear substituents at different positions (α , β , γ) (Figure 1.21)

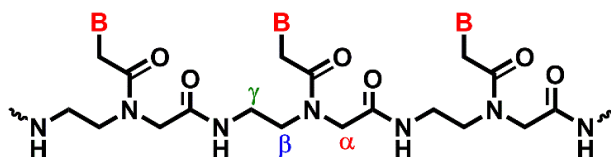


Figure 1.21. Structures of achiral PNA, B = A/T/G/C. (A = Adenine, T = Thymine, G = Guanine and C = Cytosine)

C $^{\alpha}$ -PNAs The first C $^{\alpha}$ -PNAs were reported by Nielsen *et al.*⁷⁴ in 1994, derived from both L and D alanine (Figure 1.22a). PNA:DNA duplex stability derived from D-alanine in PNA is almost equal to unmodified PNA whereas C $^{\alpha}$ -substituted PNA derived from L-alanine drastically destabilised the duplex stability. PNAs derived from D-lysine (Figure 1.22b) in place of glycine exhibited higher T_m than that from the corresponding PNA derived from L-lysine, with T_m more than that with unmodified PNAs.⁷⁵ C $^{\alpha}$ -substituted PNAs derived from D-amino acids prefer P-helix formation and mismatch discrimination ability is more than those of unmodified PNAs.⁷⁶ PNA oligomer containing three units of Lys-D-PNA and its peptide conjugate tested for cellular uptake ability, were found to be taken in most of the tested cells.⁷⁷

D-Arginine in place of glycine leads to C^α -guanidinium PNAs (GPNAs, Figure 1.22c) that showed destabilization of the derived PNA:DNA duplex. However, incorporation of multiple C^α -guanidinium PNA units (α -GPNA) in a PNA decamer improved the binding affinity, partly from electrostatic interactions between guanidinium group and phosphates.⁷⁸ C_α -PNAs with negatively charged side chains (derived from glutamic acid and aspartic acid) destabilized the PNA:DNA duplex due to electrostatic repulsions between negatively charged PNA backbone and polyanionic DNA backbone.⁷⁹

C^α -substituted PNAs derived from α -amino acids having bulky side chains such as tyrosine, histidine, tryptophan, phenylalanine, and valine showed lower T_m values due to steric hindrance.⁸⁰ 1-(2,2-dimethylcyclobutyl) ethanone containing α -PNA (Figure 1.22e) showed almost equal binding affinity with DNA and RNA. Gem-dimethyl substitution at C^α increases the T_m of derived PNA:DNA duplexes, and the homologous aminopropyl-(α,α -dimethyl)glycine (*apdmg*) PNAs improved the binding affinity to DNA. These gem-dimethyl substituted PNAs have showed preferential binding to DNA than to RNA.

C_β -PNAs In 2011, Sugiyama *et al.*⁸¹ reported the first C_β -substituted PNA bearing a methyl group at β -position using D and L enantiomers of alanine (Figure 1.22f). C_β -PNA oligomer derived from L-alanine showed similar binding affinity as that of unmodified *aeg* PNAs, whereas the corresponding C_β -PNA derived from D-alanine did not bind to DNA. Circular dichroism revealed that C_β -S-PNA exhibited a right-handed helical structure whereas C_β -R-PNA exhibited left-handed helical structure.⁸¹

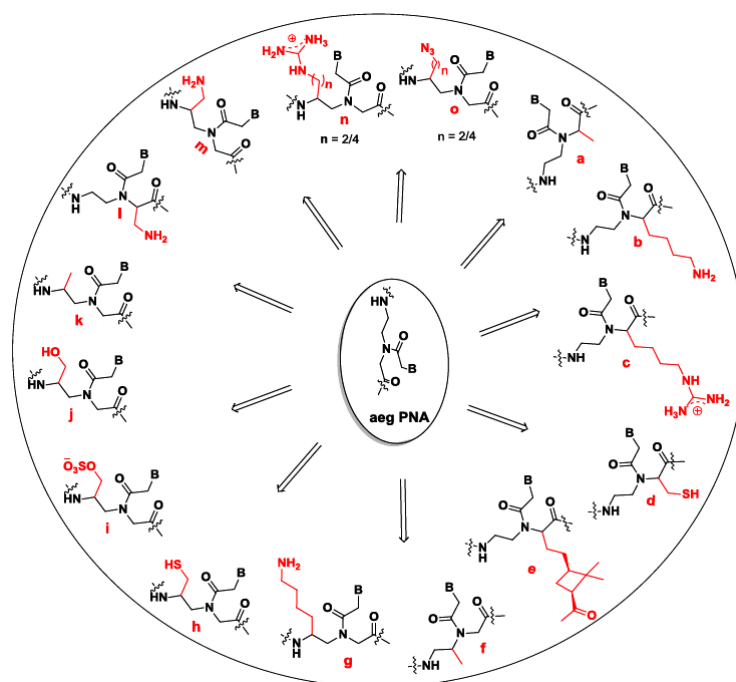


Figure 1.22 Modified acyclic PNAs.

C^γ-PNAs Although the first chiral C^γ-PNA was reported in 1994, C^γ-PNA oligomers did not appear in literature until 2005.⁸²⁻⁸⁶ C^γ-(L)-Lys-modified PNAs (Figure 1.22g) stabilized PNA:DNA duplexes and showed good mismatch discrimination relative to unmodified PNAs, whereas C^γ-D-Lys-PNAs markedly destabilized the duplex.⁸⁷ Cysteine based PNA substitution (Figure 1.22h) at the N-terminus of a PNA oligomer can be used for the native chemical ligation with PNAs carrying thioesters to yield long chain PNAs. The effect of stereochemistry on DNA binding was more effective in C^γ-PNAs than that of C^α-PNAs.⁸⁸ Homopyrimidines incorporated with C^γ-PNA bearing a sulfate group (Figure 1.22i) formed a triplex with lower stability which arises from the repulsions between negatively charged phosphate DNA and sulfate of modified PNAs. The modified PNA could be lipofected into human breast cancer (SKB3) cells and exhibited antigene activity against ErbBr gene.⁸⁷

Serine or alanine based C^γ-PNAs (Figure 1.22j and 1.22k) revealed that single-stranded γ - backbone modifications preorganized to right-handed helix, similar to that in PNA:DNA duplex.^{89,90} Ly *et al.*⁹¹ have reported a variety of C^γ-PNAs and demonstrated that C^γ-position can accommodate hindered side chains unlike C^α-PNAs without affecting the hybridization properties considerably. The cellular uptake of a fully modified alternate C^γ-GPNA decamer by HeLa cells is comparable to that of TAT transduction domain. Recently, Manicardi *et al.*⁹² reported the inhibition of micro-RNA by GPNA carrying an arginine side chain. Anti-miR-210 activity of PNAs in leukemic K562 cells was examined using a series of modified 18-mer PNAs (α or γ , PNAs conjugated with arginine octamer and unmodified PNAs). The best anti-miR-210 activity was observed with C^γ-GPNA with consecutive placement.

Ganesh *et al.*⁹³ have reported the design and synthesis of chiral C _{α} - and C _{γ} -aminoethyl PNAs (*am*-PNAs) with substitutions in the PNA backbone (Figure 1.19l and 1.19m). The *am*-PNAs formed stable PNA:DNA duplexes and the order of stabilization was, γ -(S)- *am*-PNA > α -(R)- *am*-PNA > α -(S)-*am*-PNA. The *am*-PNAs are taken up by HeLa cells, with decreasing order of uptake efficiency as γ -(S)-*am*-PNA > α -(R)-*am*-PNA > α -(S)-*am*-PNA. Recently,⁹⁴ it was shown in γ -GPNA (Figure 1.19n, n = 2) and γ -azido PNAs (Figure 1.19o) that side chain lengths is important for hybridization with DNA. The guanidinium PNAs with shorter spacer chain increased the PNA:DNA duplex stability. These PNAs taken up in 3T3 and HeLa cells were visualized by confocal microscopy and quantified using fluorescence assisted cell sorter (FACS).⁹⁵ The azido functionality of C^γ-substituted methylene/butylene azido PNAs enabled the attachment of multisite labeling and introduction

of fluorophores in a single step through click reaction without any protection/deprotection steps. The azido fluorescent PNA oligomers have been shown to accumulate around the nuclear membrane in 3T3 cells.⁹⁶

1.6.3b Conformational preorganization through cyclic PNAs. The PNA being acyclic, is conformationally flexible and formation of ordered PNA:DNA/RNA complexes is accompanied by enthalpic advantage through hydrogen bonding and base stacking interactions but has an undesirable loss in entropy of the PNA strand.⁹⁷ A high rotation barrier is encountered in the interconversion of *cis* and *trans* rotamers around the tertiary amide linkage in PNA and leads to different PNA:DNA/RNA hybridization kinetics in parallel and antiparallel hybrids.^{98,99} Ganesh *et al.*¹⁰⁰ have synthesized many modified cyclic PNAs that can potentially preorganize PNA to form PNA:DNA/RNA complexes and achieved hybridization preferences (decreased loss in entropy) for either DNA or RNA complementation and parallel or antiparallel orientation.

Many strategies have been developed to enrich populations of single-stranded PNAs that have the favorable, pre-organised conformation for binding to complementary DNA/RNA. The preorganized conformers could trigger a shift in equilibrium toward the desired complexation because of the net reduced entropy loss upon complexation. The strategies to conformational preorganization are based on the introduction of methylene/ethylene groups to bridge the aminoethyl-glycyl backbone and methylene carbonyl side chain to generate diverse five- or six-membered nitrogen heterocyclic analogs. Some of these strategies are described below in detail.

Aminoprolyl PNA (ap-PNA) 4(*R/S*)-Aminoprolyl PNA synthesized by the introduction of a methylene bridge between β -carbon atom of the aminoethyl segment and the α' -carbon of the glycine segment on *aeg*-PNA backbone (Figure 1.23).¹⁰¹ None of the homochiral aminoprolyl thymine PNAs corresponding to any of the diastereomers bound to target DNA sequences¹⁰² and this might be attributed to high rigidity in the backbone resulting in structural incompatibility.¹⁰³

Aminoethylprolyl PNA (aep-PNA) Chimeric aminoethylprolyl PNA (*aep*-PNA) analogues arise from connecting the glycine on backbone with side chain to give 4(*S*)-2(*S/R*)-*aep*-T oligomers (Figure 1.23) that showed strong and specific binding properties toward target DNA sequences.¹⁰⁴ The homooligomer *aep*-PNA-T₈ with (2*R*,4*R*) stereochemistry displayed significant stabilization of the complexes with poly rA,¹⁰⁵ while the adenine A₈ homooligomer with (2*S*,4*S*) stereochemistry showed improved binding to the target DNA.¹⁰⁶ Nucleobase dependent binding and orientation selectivity were observed with these constructs.

Aminoethylpyrrolidinone PNA (aepone-PNA) The aminoethylprolyl-5-one thymine (*aepone*) monomers contain cyclic amide and were incorporated into *aeg*-PNA- T_8 backbone at different positions.¹⁰⁷ The *aepone*-PNAs (Figure 1.23) showed remarkable stabilization of derived PNA₂:DNA triplexes compared to *aeg*-PNA.¹⁰⁸ The homooligomers of *aepone*-PNA- T_8 binds to the complementary 12-mer DNA more strongly to cRNA. This suggested that structurally preorganized PNAs can discriminate their binding selectivity to DNA or RNA.

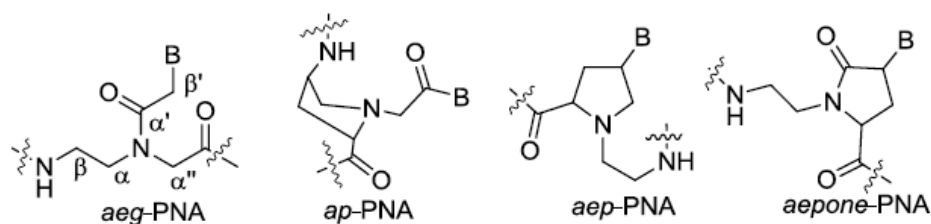


Figure 1.23 Cyclic PNAs: *ap*-PNA, *aep*-PNA and *aepone*-PNA

Pyrrolidine PNAs Deletion of the *endocyclic* carbonyl group in pyrrolidinone PNA gave the pyrrolidine PNA¹⁰⁹ (Figure 1.24a). The derived (2*R*,4*S*) stereomeric homo adenylate chimeric *aeg*-PNA oligomer formed a stable complex with both DNA and RNA.¹⁰⁹ The (2*R*,4*R*) version of this PNA analogue was shown to bind target DNA and RNA with high affinity and kinetic selectivity toward RNA.¹¹⁰ (2*S*,4*S*) pyrrolidine PNA analogue, and its thyminyl oligomer resulted in a decreased binding efficiency with target DNA and RNA sequences.¹¹¹ PNA:DNA dimer prepared from (2*R*,4*R*) pyrrolidine PNA-*T*, which when placed in a PNA-DNA chimera led to decreased DNA triplex stability.¹¹¹

Introduction of the α' - β -methylene bridge led to another pyrrolidine-PNA (Figure 1.24b)¹¹² having the nucleobase away from the pyrrolidine ring by one carbon. The (2*R*,4*S*) pyrrolidine-*T* monomer when introduced into the middle of the *aeg*-PNA- T_8 , bound to the target DNA better than the diastereomeric (2*S*,4*S*) PNA. This is an example of stereochemical discrimination effects in PNA:DNA recognition. The 4*R* pyrrolidine-*T* PNA sequences stabilized the PNA:DNA duplex with a significant difference in parallel/antiparallel binding compared to *aeg*-PNA:DNA duplexes,¹¹³ while the PNAs with 4*S* modifications destabilized the duplexes without much parallel/antiparallel binding differences as compared to the unmodified hybrids. The enantiomeric pairs (2*S*,4*S*) and (2*R*,4*R*) formed antiparallel complexes with RNA much stronger than that of *aeg*-PNA or other diastereomers.

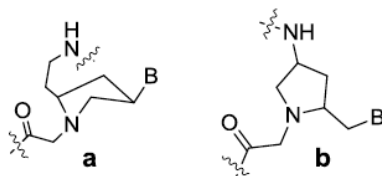


Figure 1.24. Pyrrolidine PNAs⁹⁶

Pipecolic and Piperidinyl PNAs The pipecolyl PNA oligomers were designed by introducing a methylene, or an ethylene bridge between the C^γ- or C^β-carbon of the aminopropyl segment and the α''-carbon of the glycyloxy segment into an aminopropyl-glycyl PNA analogue (Figure 1.25).¹¹⁴ The modified PNA-T₁₀ oligomer (*pip*-PNA-1) carrying the *trans* (2*S*,4*S*) monomer destabilized the PNA₂:DNA triplex formation. However, the homothyminyly mixed *aeg*-PNA sequences incorporating the 5-aminopipecolyl unit (*pip*-PNA-2) formed stable complexes with target DNA oligomers.

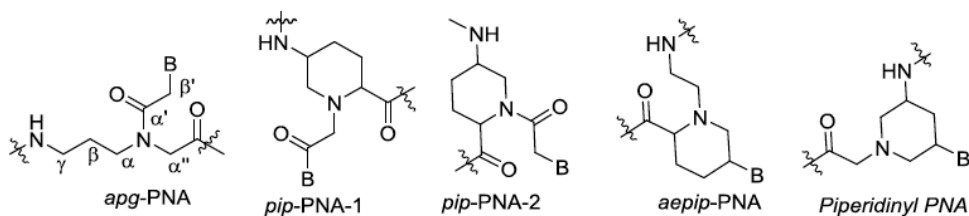


Figure 1.25 Pipecolic and Piperidinyl PNAs¹⁰²

The chiral six-membered analogues (2*S*,5*R*) *aepip*-PNA and the *trans* (3*S*,5*S*)-piperidinyl PNA-T upon incorporation into the *aeg*-PNA-T₈ homooligomer at different positions stabilized the corresponding PNA₂:DNA triplexes.¹¹⁵

Cyclohexyl PNA One of the earliest PNA modifications was to constrain the flexibility in the aminoethyl segment around a single C-C bond by introducing a cyclohexyl ring (Figure 1.26).¹¹⁶ The derived *trans*-(*S,S*)-cyclohexyl PNA oligomer hybridized with complementary DNA as good as the unmodified *aeg*-PNA, while the enantiomeric *trans*-(*R,R*)-cyclohexyl PNA oligomers lacked such a property. Significant stereo differentiation was observed with *SR*- and *RS*-*ch*-PNAs, the *SR* isomer was more destabilizing than the *RS* isomer in *ch*-PNA:DNA complexes, while in case of RNA complexes, *RS* was more destabilizing than *SR*. From *T_m* values it was observed that (*R,S*)-*ch*-PNAs bound to RNA with higher affinity than to DNA. This is attributed to the inherent rigidity of the *cis*-*ch*-PNAs that forbids structural readjustments to bind to DNA (PNA:DNA β ~ 140°) and prefers binding to RNA (PNA:RNA). This discrimination was achieved *via* the concept of preorganization.

Cyclopentyl PNAs Ganesh *et al.*¹¹⁷ also synthesized *cis*-(1*S*,2*R*/1*R*,2*S*)-cyclopentyl PNA-T oligomers in which the characteristic *endo-exo* puckering that dictates the pseudoaxial/pseudoequatorial dispositions of substituents may allow better torsional adjustments to attain the necessary hybridization-competent conformations. The homothymine *aeg*-PNA-T₁₀ oligomers with modifications at defined positions exhibited binding affinities with DNA and RNA. It was observed that the *RS*-*cp*-PNA enantiomer showed higher affinity for DNA as compared to *SR*-*cp*-PNA isomer and in case of RNA, this was reversed.

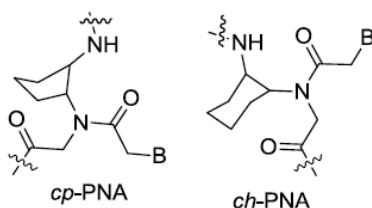


Figure 1.26. Cyclohexyl and cyclopentyl PNAs

The all-modified homooligomers of both enantiomers exhibited significant stabilization of their triplexes with DNA and poly rA without sacrificing the base specificity. *ch*-PNA and *cp*-PNA-T monomers of both (*S,R*) and (*R,S*) enantiomers were introduced into mixed sequences, in all the cases PNA:RNA hybrids were more stable than the corresponding PNA:DNA hybrids, and significantly, *cp*-PNA oligomers showed much higher T_m s compared to *ch*-PNA.

1.6.4 Synthetic heterocycles as *Janus* nucleobases

The term Janus heterocycle, was first coined by Nobel laureate J. M. Lehn, after the two-faced Roman god.^{118,119} The term Janus heterocycle is employed to describe heterocycles that are capable of forming hydrogen bonds on both faces of the heterocycle, in contrast to natural nucleobases that typically present only one face for functional hydrogen bonding (not counting the potential for simultaneous formation of Hoogsteen pairs) (Figure 1.27).

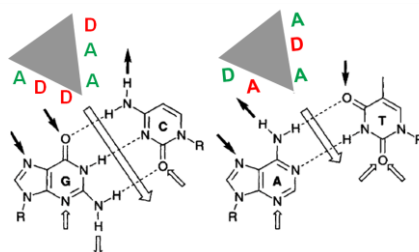


Figure 1.27. Janus-wedge recognizing regular DNA base pairs (A: hydrogen bond acceptors, D: hydrogen bond donors, R=deoxyribose)

A Janus heterocycle therefore represents one type of artificial molecule that can be rationally designed with predefined hydrogen bonding donors and acceptors relative to regular nucleobases. For example, if one side is complementary to A (which has the character of T) the other side can complement C (which has the character of G base), this Janus heterocycle is named Janus-TG. Typically, this Janus-TG heterocycle can form Watson-Crick base pair with A on one strand and C on another strand, as AC cannot form direct associate through base pairing. Taking this complementarity into account, there are two kinds of Janus heterocycles, one of which can form base pair with itself whereas the other cannot. So far, several Janus heterocycles have been reported, including the reported Janus-GA (Lehn *et al.*¹²⁰) and subsequently, the Janus-GC (Mascal *et al.*¹²¹) and Janus-AT (Perrin *et al.*¹²²) (Figure 1.28).

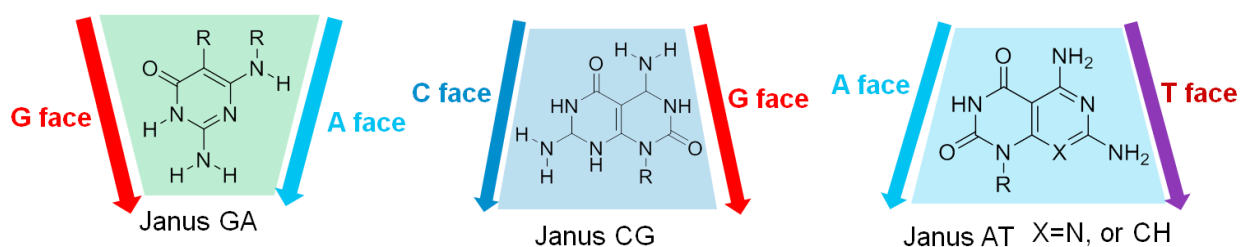


Figure 1.28 Janus heterocycles (R=alkyl groups in all three kinds of Janus heterocycles)

Due to self-complementary hydrogen-bonding donor and acceptor groups on two faces, Janus-GC and Janus-AT can aggregate to form interesting non-covalent supramolecules, the driving force being the formation of hydrogen bonding as well as stacking interactions that favor creation of higher order structures. For example, Fenniri *et al.*¹²³ reported the occurrence of helical rosette nanotubes in water by employing Janus-GC heterocycle (figure 1.29b).

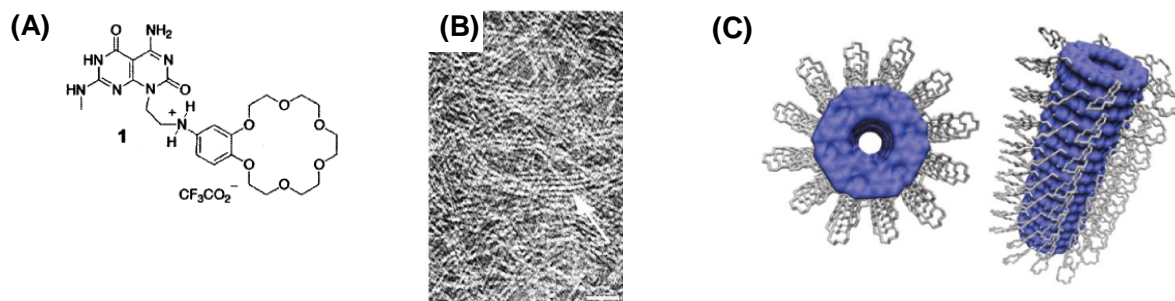


Figure 1.29. Self-assembly rosette nanotubes of Janus GC conjugate (A) GC conjugate 4-aminobenzo-18-crown-6-ether (B) TEM image of A (C) hierarchical self-assembly of A.¹²³

This group also reported that the property of helical rosette nanotubes could be adjusted by changing the side chain attached to the Janus-GC heterocycle. It was shown that chirality of rosette nanotubes could be induced with the addition of chiral amino acids when the side chain of Janus-GC was a crown ether. In addition, when Janus-GC heterocycle was combined with the ribose forming Janus-GC nucleosides, Janus-GC nucleosides also have the capacity to self-aggregate forming the aforementioned rosette nanotubes but with chiral properties owing to the presence of the ribose moiety.¹²⁴ When the Janus-AT nucleoside¹²⁴ was examined under similar conditions, instead of rosette cyclic structure, a linear sheet-like structure was observed. Although a given Janus-GC heterocycle or Janus-GC nucleoside can form the rosette nanotube superstructure, one significant limitation to the application of these kinds of rosette nanotubes is the lack of reproducibility of the identical nanotubes.¹²⁵ In order to achieve pre-defined nanotubes, they hypothesized that a certain number of Janus heterocycles should be incorporated into oligonucleotides to form a Janus heterocycle oligonucleotide, where sequence specificity can be modulated through suitable length as well as the incorporation of standard nucleosides. This is now being pursued by several laboratories. Besides its utility in the construction of nanotubes, another potential application of a Janus heterocycle would be to act as a device by which one could recognize specific DNA sequences based on Watson-Crick base pairings.

However bifacial *Janus* bases (that bind one face by Watson-Crick and second face by non-natural H-bonding) have been used for invasion of double helical DNA in specific and stable manner Ly *et al.*¹²⁶ By utilization of bivalent nucleic acid ligands for recognition of RNA repeated expansion associated with Huntington's disease by Ly *et al.*¹²⁶

1.6.4a Peptide Nucleic Acids: incorporation of Janus wedge motifs. PNA recognizes single-stranded and double-stranded nucleic acids through hydrogen bonding via Watson-Crick and Hoogsteen faces of the nucleobases. This approach works well for triplexes T-A-T, C⁺-G-C, A-A-T, and G-G-C but does not cover the full scope of possibilities, such as mismatched nucleobases, due to difficulty in triplex formation with pyrimidine bases. McLaughlin *et al.*¹²⁷ successfully used an oligo-PNA with Janus wedge motifs (Figure 1.30) to target T-C mismatches in DNA. Using a heterocycle called W which presents hydrogen bonding motifs that would complement the Watson-Crick faces of thymine and cytosine, they were able to form stable triplexes with a duplex DNA having 8 consecutive T-C mismatches flanked by eleven canonical bps. In later work, heterocycles targeting A-T and G-C bps were able to successfully form triplexes when canonical bps were flanked by multiple mismatched bases

and discriminate against non-complementary nucleobases.¹²⁸ Ganesh *et al*¹²⁹ utilized cyanuric acid and 8-aminoadenine moiety displayed on a PNA backbone to strengthen triplex formation with oligo-A DNA. They found that a single cyanuric acid moiety in the middle of an oligo-T displaying PNA stabilized the triplex by a 12 °C in T_m measurements by providing a Watson-Crick competent binding face, irrespective of the rotamer of the cyanuric acid presented to the DNA nucleobase.

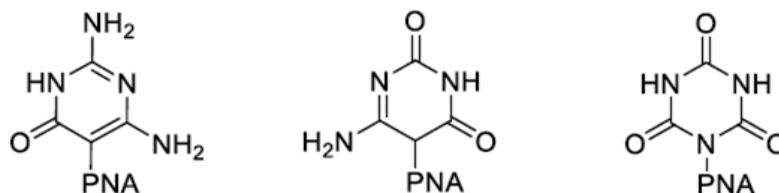


Figure 1.30. Janus wedge motifs used in PNA

1.6.4b Peptide Nucleic Acids: native peptide backbone. After Nielsen discovered PNA as a competent nucleic acid mimic, other groups have looked at native peptide backbones as a way of displaying genetic information. Lenzi *et al*¹³⁰ first attempted a native peptide backbone nucleic acid with a self-complementary sequence incorporating thymine and adenine nucleobases, but the hairpin had a modest T_m of 19 °C, possibly due to a lack of flexibility in the backbone. Diederichsen *et al*¹³¹ synthesized alanyl peptide nucleic acids with alternating stereo centers to create nucleobase displaying beta sheets, again with modest success compared to PNA. Huang *et al*¹³² introduced α -PNA in which the nucleobases are displayed on an alpha-helical peptide. Using a positively-charged helical structure, cooperative binding to the DNA sequence GGAGG was demonstrated as well as high melting temperatures in a possible triplex structure.

Eschenmoser and Krishnamurthy,¹³³ interested in the primordial evolution of genetic material, studied α -peptides and peptoids displaying triazine heterocycles as potential precursors to modern nucleic acids (Figure 1.31). In their systems 2,4-diamino-1,3,5-triazine and 2,4-dioxo-1,3,5-triazine rings were displayed on a glutamyl, aspartyl iminodiacetic acid, and ethylenediamine monomers with alternating aspartic acid or dialkyl-ammonium groups for solubility. Their findings showed that the binding of these small oligos, displaying as few as six nucleobase mimics, were robust in binding to RNA and DNA when containing the 2,4-diamino-1,3,5-triazine heterocycle displayed on the longest linker. They found that the 2,4-dioxo-1,3,5-triazine nucleobase mimic did not bind to nucleic acid targets due to deprotonation in the assay buffer. While they ruled out that the triazine heterocycles as

possible nucleobase precursors, they did show that α -peptides are competent as nucleic acid analogs even with negatively charged backbones. Furthermore, while not a major finding of the paper, they found that the 2,4-diamino-1,3,5-triazine is effective as a nucleobase mimic capable of recognizing uracil and thymine. In a follow up communication using α -peptides, the researchers showed that pK_a of the nucleobase is important.¹³⁴ Their findings yielded a relationship between pK_a of the base and pH of the solution, in that if the base becomes deprotonated or tautomerize in the buffer (at pK_a near or equal to the pH of the solution), then molecular recognition is weak.

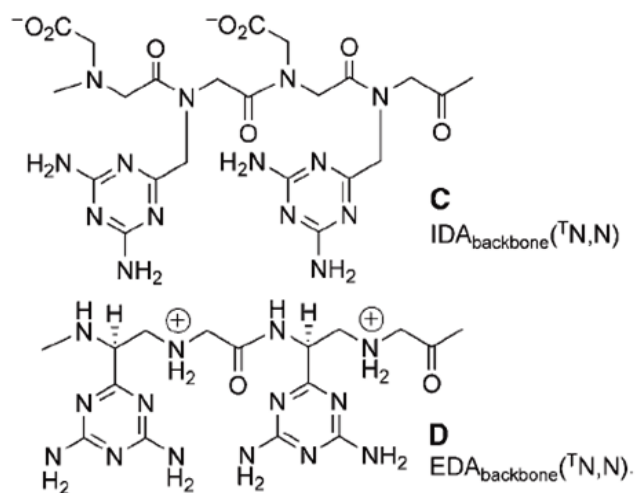


Figure 1.31. Structures of triazine-tagged oligomers (IDA=iminodiacetic acid; EDA=ethylenediamine; ^(T,N,N)= (2,4-diamino)triazin-6-yl)¹⁴²

The purines (A and G) are capable of forming base pairing from both faces, through Watson-Crick / Hoogsteen H-bonding schemes, while pyrimidines (C and T) cannot achieve the same. In a way, A and G are “Janus” bases with two non-identical faces and are important in the formation of triplexes (A/T*A:T and C⁺/G*G:C) and G-quadruplexes. Mutual recognition of non-complementary bases has been achieved by insertion of wedge-like heterocyclic “Janus bases” having two faces designed with H-bonding sites, matching that of the target nucleobases. The versatility of synthetic “Janus” bases to simultaneously recognize two natural nucleobases and the geometric separation of adjacent bases in *aeg*-PNA attuned to that in natural DNA / RNA duplexes motivated us to design “Janus” PNAs.

1.6.5 Biological applications of PNAs

This section describes applications PNA and modified PNAs in gene therapeutics, molecular biology, and programmable supramolecular assemblies.

1.6.5a PNA can use as antisense and antigene agents. PNA can be used as an effective gene therapeutic agent by unique strand invasion property and their chemical as well as biological stability. There are basically two strategies involved in using PNAs as therapeutic drugs, namely antigene and antisense methods. Moreover, no sign of any general toxicity of PNA has so far been observed.

Inhibition of transcription. PNAs can inhibit the transcriptional processes by virtue of their ability to form a stable triplex, duplex or duplex structure or a strand invasion or strand displacement complex with DNA. Such complexes can create a structural hindrance to block the stable functioning of RNA polymerase and thus are capable of working as antigene agents.¹³³Nielsen *et al.* have demonstrated that even an 8-mer PNA (T₈) is capable of blocking phage T₃ polymerase activity.¹³⁴

Inhibition of translation In the case of antisense strategy, the nucleic acid analogs can be designed to recognize and hybridize to complementary sequences in *mRNA* and thereby inhibit its translation.¹³⁵It has been established from the results of *in vitro* translation experiments involving rabbit reticulocyte lysates that both duplex-forming (mixed sequence) and triplex-forming (pyrimidine rich), PNAs can be useful for inhibiting translation at targets overlapping the AUG start codon.

Inhibition of replication PNA can also inhibit the elongation of DNA primers by DNA polymerase. Further, the inhibition of DNA replication should be possible if the DNA duplex is subjected to strand invasion by PNA under physiological conditions or if the DNA is single-stranded during the replication process. Efficient inhibition of extra-chromosomal mitochondrial DNA, which is largely single-stranded during replication by PNA, has been demonstrated by Taylor *et al.*¹³⁴

Although PNA has been shown to interfere in several biological functions, till now there are no PNA based molecules which have reached clinical trials. The first drug approved by FDA is Vitravene¹³⁵ It is an antisense drug developed by Ionis pharmaceuticals in 1998 for the treatment of cytomegalovirus (CMV) retinitis in AIDS patients. Mipomersen¹³⁶ is the second FDA approved antisense drug which is a cholesterol reducing drug candidate that targets the mRNA for Apo lipoprotein B. There are many antisense drugs which are recently approved by FDA the examples are tabulated in the Table 1.1, but none of them are based on PNA scaffold.

Table 1.1 Examples of antisense drugs

DISEASE	Drug	Chemistry	Target	Clinical Phase	Ref.
Retinitis	Fomiversen	PS DNA	CMV IE2	Approved	135
Cardiovascular	KYNAMRO Isis G	2'-MOE,PS DNA	ApoB-100	Approved	136
DMD	Exondyse 51	PMO	Exon 51	Approved	137
DMD	Drisapersen	2'-OMePS	Exon 51	Approved	137
SMA	SPINRAZA	2'-MOE	Exon 7	Approved	138
Cancer	Custirsen	2'-MOE, PS DNA	Clusterin	Approved	139
AMD	Mucugen	PS	VEGF165	Approved	140

DMD: Duchenne muscular dystrophy; **SMA:** Spinal muscular atrophy; **AMD:** Age-related muscular degeneration.

Recently bifacial *Janus* bases (that bind one face by Watson-Crick and second face by non-natural H-bonding) have been used for invasion of double helical DNA in specific and stable manner by Ly *et al.*¹⁴¹ Bivalent nucleic acid ligands have been recently used for recognition of RNA repeated expansion associated with Huntington's Disease by Ly *et al.*¹⁴¹ In a recent work, Bong *et al.*¹⁴² described a versatile non-covalent strategy based on triplex hybridization of oligo-uridylyate RNA with bifacial polymer nucleic acid (bPoNA). bPoNA was prepared and side chain-functionalized with N-acetylgalactosamine (GalNAc), which is known to enable delivery to hepatocytes and liver via binding to the asialoglycoprotein receptor (ASGPR) cell. Alnylum Pharmaceutical (USA) has used tri-antennary N-acetylgalactosamine (GalNAc) moiety for targeted delivery of siRNA to hepatocytes. It has been observed that using GalNAc as a vehicle to delivery of siRNA improves potency 10-fold in mice.¹⁴³

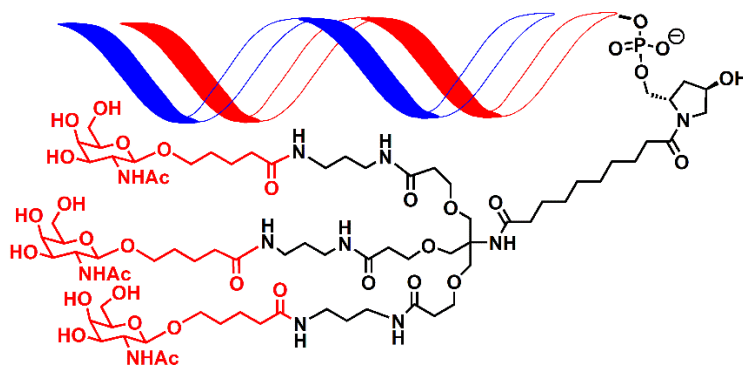


Figure 1.32. Structure of the triantennary GalNAc₃-siRNA conjugate¹⁴⁴

1.7 Methodologies: Synthesis and Biophysical Techniques

This section describes the method of solid phase synthesis of PNA oligomers, resin cleavage, purification, characterization and various methods for studying hybrid formation with cDNA, such as temperature dependent absorbance, circular dichroism spectroscopy, and isothermal titration calorimetry.

1.7.1 Solid phase synthesis of PNA oligomers

Peptides can be synthesized either in solution or by using solid phase synthesis protocols.¹⁴⁵ Synthesis of short peptides can be more efficient by solution phase strategy; however it 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,¹⁴⁶ utilizes polymeric beads, which has functional groups located on their surfaces and in their pores¹⁴⁷ (Figure 1.32). Small solid beads are insoluble in organic solvents and act as solid support. The peptide chains are grown over them by sequentially linking the amino acid monomers to the functional groups or linkers on the bead. 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, excess reagents and the byproducts are washed away by filtration. The next *N*^α-protected amino acid is coupled to the resin-bound amino acid either by using 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. 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.

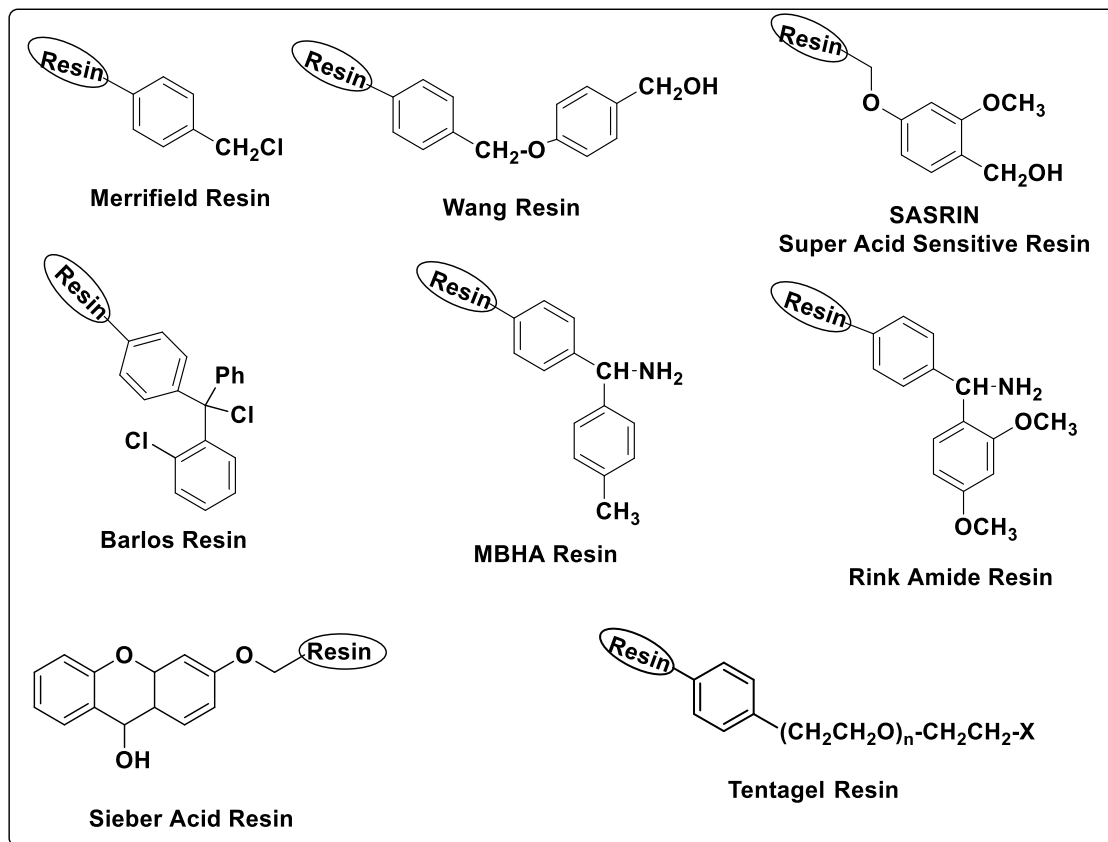


Figure 1.32 Representative structures of resin used in SPPS

There are two routinely followed protocols for solid phase peptide synthesis: *Fmoc* and *Boc* strategies, which use base labile and acid labile protecting groups respectively (Figure 1.33). The solid-phase peptide synthesis proceeds in C-terminal to N-terminal fashion. The N-terminus of amino acid monomers are protected by these two groups and added onto a deprotected amino acid chain. For N^α -protection, the first protocol uses the *t*-butoxy carbonyl (*t*-Boc) group, which can be removed by acidic conditions, such as 50% TFA in DCM. The reactive side chains are protected with groups that are stable to *t*-Boc deprotecting conditions and can be removed under strongly acidic conditions using HF^{148} in dimethylsulfide or TFMSA^{149} in TFA. In the alternative protocol, fluorenylmethyloxycarbonyl (Fmoc)¹⁵⁰ group is used for N^α -protection, it is stable to acidic conditions but can be cleaved off efficiently with a base such as piperidine.¹⁵¹ The final peptide and side chain protecting groups can be cleaved with acid (50% TFA in DCM).

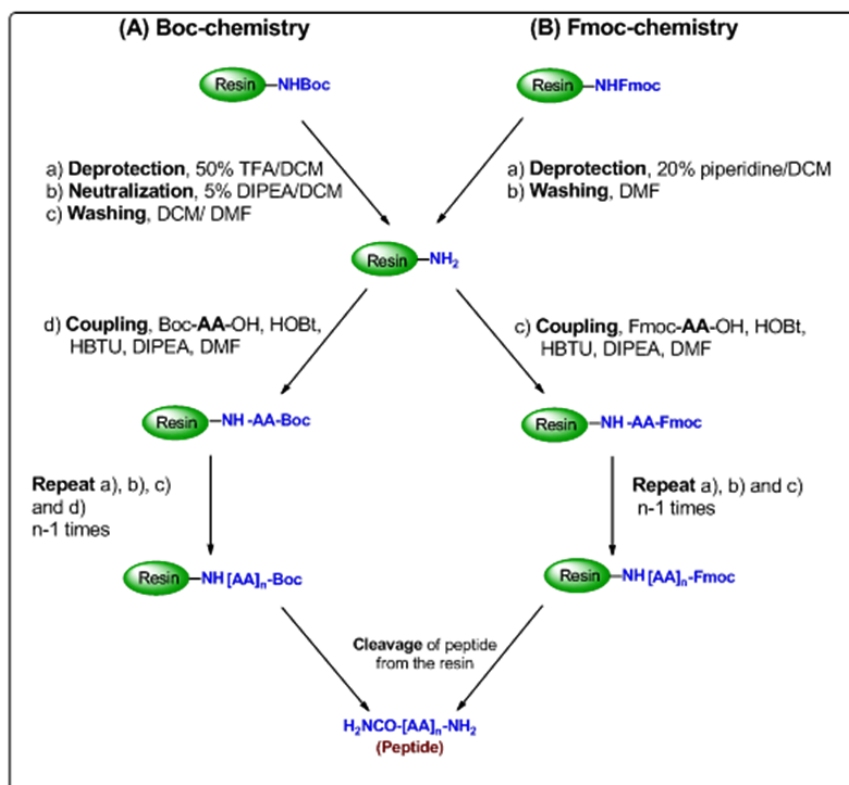


Figure 1.33. General protocols for SPPS via (A) Boc-chemistry (B) Fmoc-chemistry

After cleavage from the solid support, completely deprotected crude PNA oligomer was dissolved in MilliQ water and then purified by preparative HPLC and desired product peak is identified and characterized by mass spectroscopy. The purified PNA used for biophysical study.

1.7.2 Biophysical techniques used to study the hybridization properties

This section describes the application of various biophysical techniques to study the properties of both single stranded PNAs and double stranded PNA:DNA and PNA:RNA hybrids.

1.7.2a Temperature dependent UV absorbance (UV-melting). In nucleic acids, the two strands are held together by hydrogen bonds between nucleobases across the two complementary strands and stacking interactions between adjacent basepairs along the helix axis. The stability of the duplexes depends on factors such as temperature, pH, and ionic strength. Under conditions that disrupt the hydrogen bonding and stacking interactions, the double helix denatures. The duplex is said to “melt” if temperature is used as denaturing agent. The thermal stability of various nucleic acid complexes, including PNA:DNA and PNA:RNA hybrids, has been studied by monitoring the UV absorption at 260 nm as a

function of temperature.¹⁵² In water, the hetero-aromatic bases interact via their π electron clouds when stacked together in the duplex. Because the UV absorbance of the nucleobases is a consequence of π electron transitions, the magnitude of these transitions is dependent on whether the bases are stacked or not. The denaturation of the duplexes results in the loss of stacking interactions through disruption of hydrogen bonds, which leads to increase in the absorbance at 260 nm called hyperchromicity. The process is co-operative and the plot of absorbance at 260 nm vs temperature is sigmoidal (Figure 1.34). The midpoint of the sigmoidal transition is termed as the melting temperature, T_m .

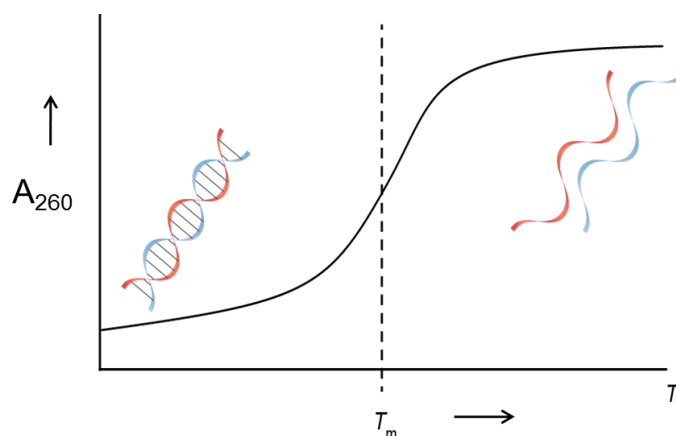


Figure 1.34. A typical sigmoidal plot showing melting temperature (T_m) of a double stranded DNA.

The sigmoidal nature of the transition suggests that the nucleic acids exist in two states, either as duplexes or as single strands. Further, at varying temperatures, the relative proportions of the two states change. A nonsigmoidal (e.g. sloping linear) transition with low hyperchromicity could indicate the lack of complementarity among the sequences in the participating strands. In many cases, the transitions are broad and the exact T_m values are obtained from the peak in first derivative plots (Figure 1.35). This technique provides valuable information about the strength of sequence-dependent complementary interactions in nucleic acid hybrids involving DNA, RNA and PNA.

The binding stoichiometry of nucleic acids can be determined from the UV-Job's plot. The combination of UV absorption and the CD spectra provides the correct determination of complex formation and the strand stoichiometry. In literature, it has been shown that polypyrimidine PNA binds to complementary DNA in a 2:1 ratio forming PNA₂:DNA triplex¹⁵³ whereas the mixed purine-pyrimidine PNA binds in 1:1 ratio forming PNA:DNA

duplex.¹⁵⁴ In this chapter, all PNAs contain mixed purine-pyrimidine bases, therefore the PNA and DNA were mixed in 1:1 ratio.

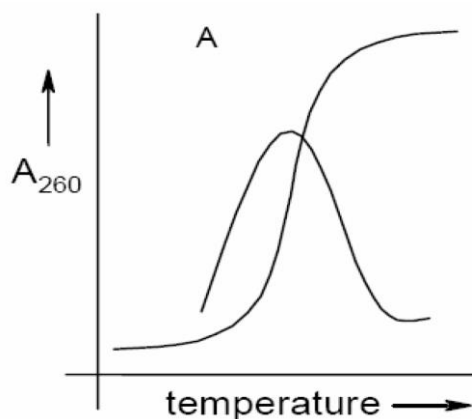


Figure 1.35 First derivative plot of UV- T_m curve of a PNA:DNA duplex

1.7.2b Circular dichroism. Circular dichroism (CD) is a spectroscopic technique used to measure the difference between the absorption of left-handed circularly polarized light (L-CPL) and the right-handed circularly polarized light (R-CPL).¹⁵⁵ CD is exhibited by molecules that contain one or more chromophores adjacent to a chiral centre. The chiral molecules exist as isomers of non-superimposable mirror images called enantiomers. The physical and chemical properties of the pair of enantiomers are identical except the way they interact with polarized light and the way they interact with other chiral molecules. CD spectra are particularly valuable in determining the following aspects:

- Whether the individual strands themselves are self-coiled
- Whether the individual strands change their conformation upon binding to complementary oligonucleotide strand
- How the conformations of hybrid PNA:DNA complexes are related with hybrid DNA:DNA complexes
- Nature of the secondary structures: A-, B-, Z-form, etc., in case of nucleic acids.

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

1.7.2c Fluorescence spectroscopy. Fluorescence is generated when a compound absorbs light energy in UV-visible region and emits radiations at a longer wavelength. Fluorescence emission the consequence of the molecule returning to its ground state from an excited electronic state.¹⁵⁵ The duration of time interval between absorption and emission is usually in the order of 10^{-9} to 10^{-8} seconds. The fundamentals of this phenomenon can be explained by Jablonski diagram (Figure 1.36 A). The difference (in wavelength or frequency units) between the positions of band maxima of absorption and emission spectra is called Stoke's shift (Figure 3.3B). Fluorescence spectroscopy has applications in examining the interaction between macromolecules and a ligand.

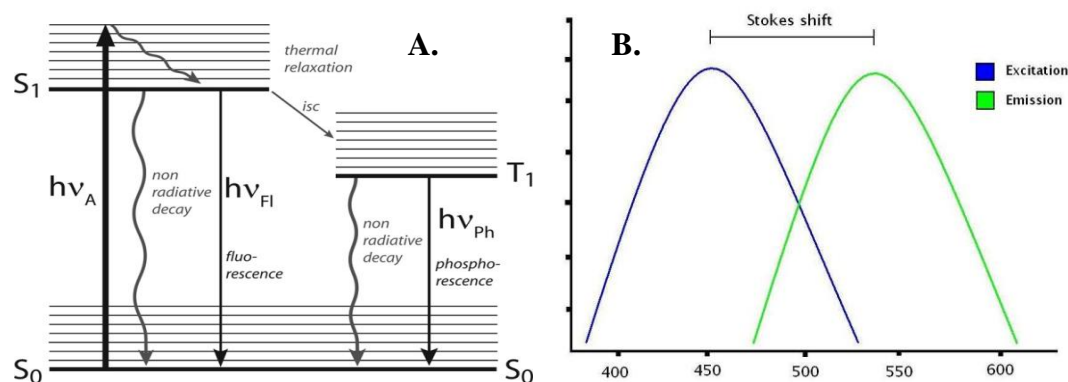


Figure 1.36 (A) A simplified Jablonski diagram of a fluorescence event. The fluorescent molecule begins in its ground energy state, S_0 , and is converted to an excited singlet state, S_1 , by absorbing energy ($h\nu_A$) in a specific wavelength. The molecule undergoes a transition to the relaxed singlet state, S_1 or T_1 , by releasing some of the absorbed energy. Finally, the molecule returns to its ground energy state by releasing the remaining energy ($h\nu_{Fl}$ - fluorescence) or ($h\nu_{Ph}$ - phosphorescence). The duration of a single fluorescence event is a few nanoseconds. (B) Generalized representation of Stoke's shift.

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

Ethidium bromide (EtBr) interacts with duplex DNA by intercalation between the base pairs. It is a weakly fluorescent molecule, but when intercalated to DNA duplex, exhibits a strong fluorescence with emission maxima at 595 nm upon excitation at 475 nm. A

competent ligand binding to DNA can displace the intercalated ethidium bromide thus leading to a decrease in the fluorescence intensity of the complex.¹⁵⁶ An attractive feature of the ethidium binding is that its fluorescence can increase upto 20 fold upon intercalation with DNA making it a very useful probe for monitoring the interactions of DNA duplex with oligonucleotides (to form triplex) or small molecules.¹⁵⁷ In ethidium bromide displacement assay, when DNA-ethidium bromide complexes are challenged by other DNA binding agents (like PNA) where ethidium bromide is displaced from the complex leading to a fall in fluorescence intensity (Figure 1.37). In such experiments, a measure of the decrease in fluorescence intensity of ethidium bromide is useful to compute the relative binding strengths of the added PNA to DNA.

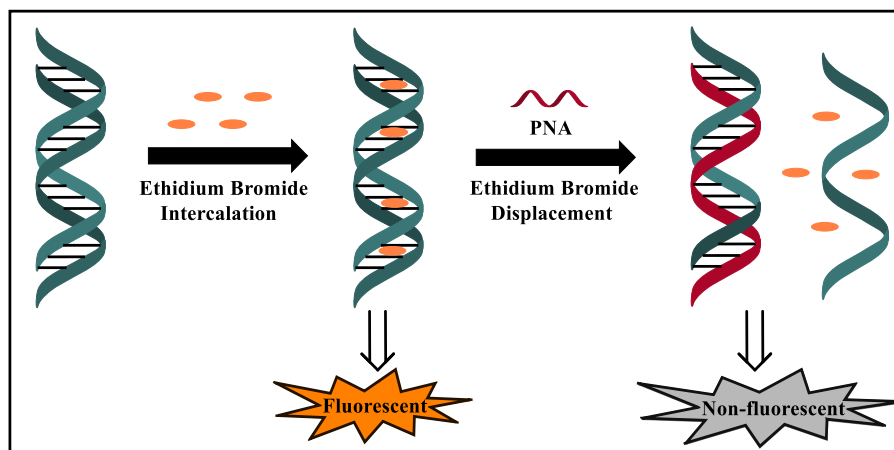


Figure 1.37. Ethidium bromide displacement assay

1.7.2d Confocal microscopy. The principle of confocal imaging was patented in 1957 by Marvin Minsky.¹⁵⁸ In a conventional fluorescence microscope (wide-field microscope), all parts of the specimen in the optical path are excited at the same time and the resulting fluorescence detected by the photodetector/ camera, including a large unfocused background. In contrast, a confocal microscope uses point illumination and a pinhole in an optically conjugate plane in front of the detector to eliminate out-of-focus signal. Because fluorescence light emitted only by the molecules in the focal plane can be detected, the optical resolution of the image, particularly along the sample depth direction, is much better than that of wide-field microscopes. However, as most of the light from sample fluorescence is blocked at the pinhole, this increased resolution is achieved at the cost of decreased signal intensity. The schematic illustration of the principle of confocal microscopy is shown in Figure 1.38.

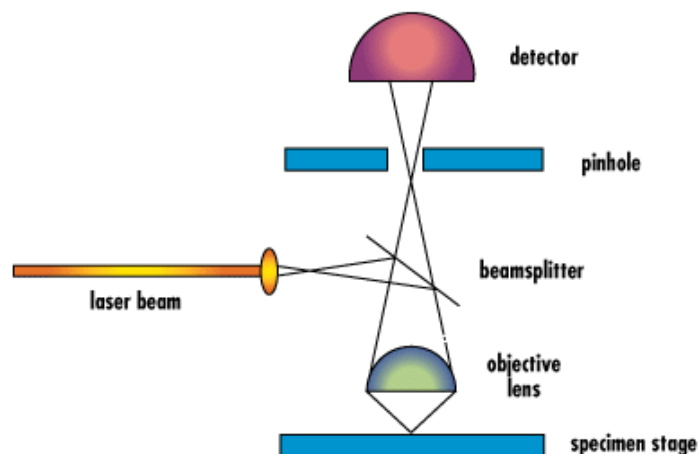


Figure 1.38 Schematic diagram of principle of confocal scanning light microscope¹⁵⁹

1.7.2e Isothermal Titration Calorimetry (ITC). The strength of interactions between the two species in a complex is represented by binding or association constant (K_a) or its inverse dissociation constant (K_d).¹⁶⁰ These can be obtained by titrating one component to another and following the extent of complexation by different analytical techniques such as UV, CD, NMR, etc. The thermodynamic details (enthalpy/entropy) of binding reaction can be assessed by isothermal titration calorimetry (ITC),¹⁶¹ a technique in which the heat of a reaction produced during the bimolecular binding is measured at constant temperature during titration of one component into another. In a typical ITC experiment, small aliquots of a titrant (eg. DNA, ligand etc) solution is added to an analyte (eg. proteins, DNA, PNA etc) solution at specific temperatures (below T_m duplex when the analyte PNA being titrated with complementary DNA) and the released heat is monitored. Judicious selection of nucleic acid concentrations and buffer conditions results in a titration curve that can be analyzed to yield the stoichiometry of the association reaction (N), the enthalpy of association (ΔH), the equilibrium association constant (K), and thus the free energy of association (ΔG). Once ΔH and ΔG are known, the entropy of association (ΔS) can also be obtained. Thus, a single ITC experiment yields a wealth of thermodynamic information about the association reaction.

1.8 Present work

Based on literature survey so far, it emerges that PNA structures need to be developed further to emerge as potential antisense therapeutic agents. In view of their strong DNA / RNA binding affinity they have a good scope towards development as probes and for material applications. The present work illustrates our approach towards these aims.

1.8.1 Rationale for present work

Triantennary galactosamine has emerged as a potential ligand for targeting hepatocytes as nicely demonstrated by their siRNA conjugates selectively entering the hepatocyte cells. In this context, the first part of the present work aims to synthesise triantennary galactosamine conjugates of PNAs to examine their potential to enter hepatocyte cells. As a variation of triantennary concept, each galactosamine moiety was also linked individually to adjacent PNA monomeric units at C γ on backbone, so that they are now in a linear format, but with same type of separation distance among them. These two types of PNA conjugates were examined for their effectiveness in permeating hepatocytes cells. In the second part of the work, a new type of PNA analogs termed “Janus PNA” have been designed in which, the C γ of sidechain carries nucleobases in addition to the classical t-amide position, to yield single stranded PNAs that can simultaneously bind two cDNA sequences using one set of sequences from t-amide site and another set from C γ -sidechain.

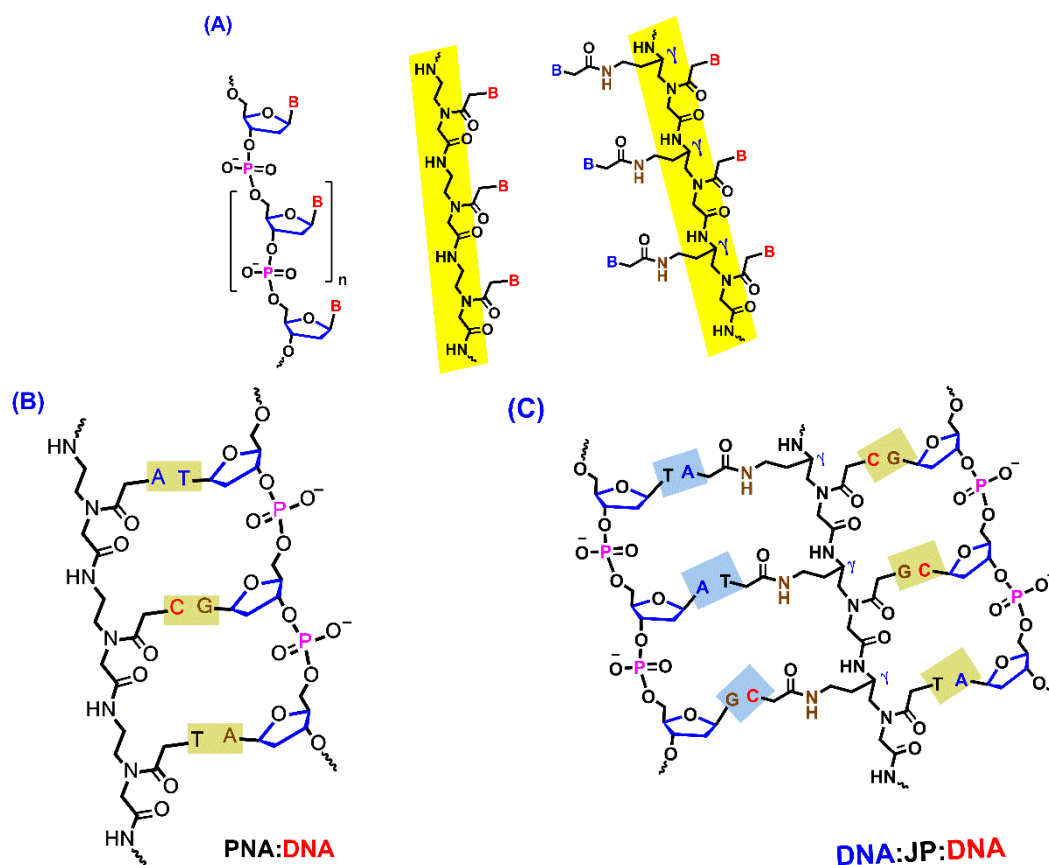


Figure 1.39. (A) Structural comparison of DNA, PNA and Janus PNA (B = nucleobase) (B) Complex of PNA with complementary DNA (C) Complex of PNA and Janus PNA with complementary DNA

1.8.2 Tri-antennary N-acetyl Galactosamine (GalNAc) Conjugated PNA

Chapter 2 describes the synthesis of tri-antennary N-acetyl galactosamine (GalNAc) and its conjugation with a 15-mer PNA (Fig.1.32.A) with the sequence in the report on siRNA-GalNAc conjugate.¹⁶² Further, three moieties of GalNAc units were consecutively conjugated at C γ on PNA backbone at N-terminus (Fig. 1.32.B) to examine site of conjugation would have any effect on their biological action. The two types of GalNAc conjugates were synthesized on solid phase, purified and characterized.

Chapter 3 consists of their biophysical studies on duplex formation with cDNA and experiments to investigate their comparative targeted delivery to hepatocytes, Hep G2 cells.¹⁵⁰

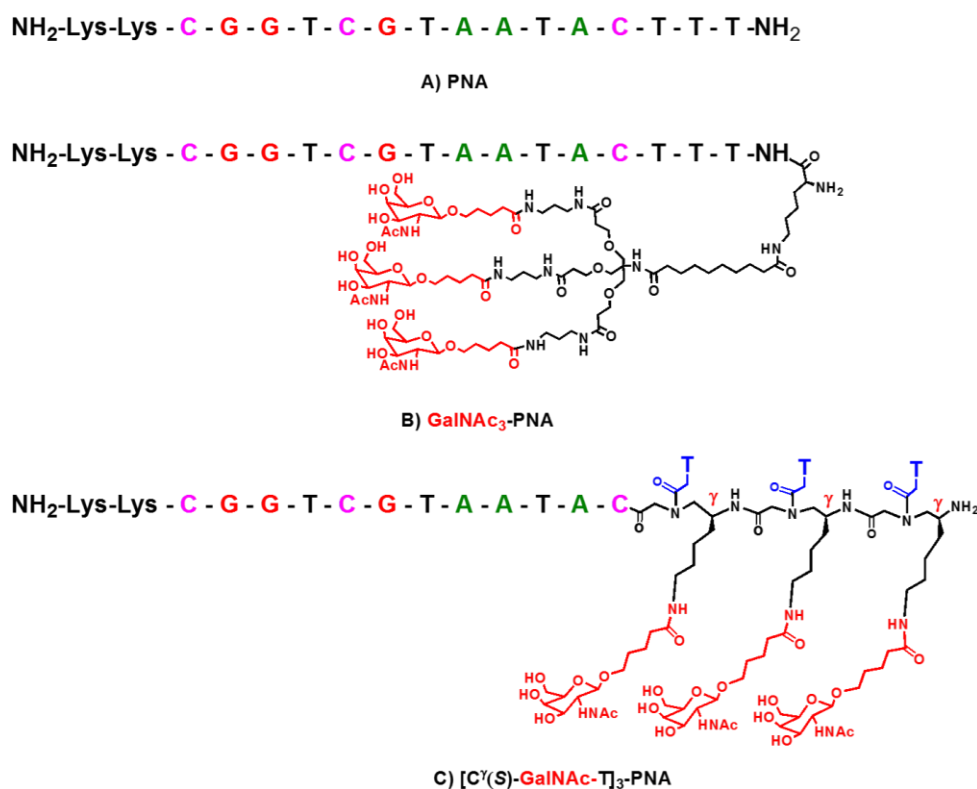


Figure 1.40. Structures of target PNA and GalNAc PNAs

1.8.3 Janus Peptide Nucleic Acid (Janus PNA)

Chapter 4 deals with design of *Janus* PNAs that bear additional natural nucleobase (A, G, C and T) on sidechain at C γ in each unit, in addition to the t-amide linked nucleobases as in standard PNA. In such PNAs, the backbone itself acquires “Janus” character and a single strand *Janus* PNA is equipped with the capability to simultaneously recognize and base pair with two complementary DNA / RNA / PNA strands from both faces of the

backbone. This would lead to the formation of double duplexes perhaps through canonical Watson-Crick H-bonding on both faces. In this chapter the required monomers in both $C^{\gamma}(S)$ and $C^{\gamma}(R)$ stereomers corresponding to different bases (A/G/C/T) were synthesized and characterized.

Chapter 5 describes how the synthesized monomers were assembled to obtain various types of Janus PNAs in which (i) both t-amide and C^{γ} -sites have homo oligomeric sequences (ii) t-amide side has mixed sequence and C^{γ} -site is a homo oligomer (G/C) (iii) t-amide and C^{γ} - side both have mixed sequences. Janus PNAs corresponding to both $C^{\gamma}(S)$ and $C^{\gamma}(R)$ stereooligomers were synthesized.

Chapter 6 mainly investigates the complexing properties of various Janus PNAs with cDNA. Depending on the type of sequences, Janus PNAs can form duplexes from either side, triplexes, double duplexes and double duplex of triplex. The different types of complexes generated are shown in Figure 1.41. The thermal stability of these various complexes are studied by temperature dependent UV absorbance, conformational features by CD spectroscopy and thermodynamic parameters by isothermal titration calorimetry. It is importantly seen that the C^{γ} -amide face duplexes are inherently more stable than the t-amide face duplex and in double duplex complexes, the duplexes are more stable than those in isolated duplexes. The high fidelity of base pairing is also established using mismatched DNA sequences.

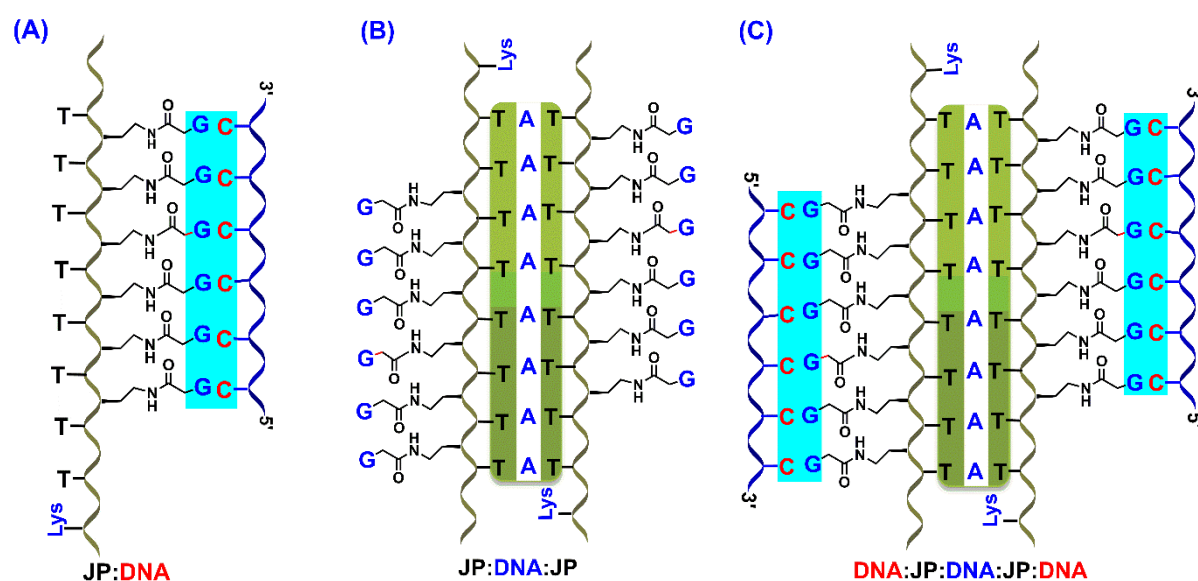


Figure 1.41. Janus PNA Can Form (A) Duplex (B) Double duplex (C) Double duplex of triplex with complementary DNA. (JP=Janus PNA)

1.9 References

1. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Wlater, P. *Molecular Biology of the Cell* (4th Ed.). *Garland Science* **2002**, pp. 120-121.
2. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.
3. Srinivasan, A. R.; Sauers, R. R.; Fenley, M. O.; Boschitsch, A. H.; Matsumoto, A.; Colasanti, A. V. Olson, W. K. *Biophysical Reviews* **2009**, *1*, 13.
4. Hoogsteen, K. *Acta. Crystal.* **1963**, *16*, 907-916.
5. (a) Crick, F. H. C. *J. Mol. Biol.* **1966**, *19*, 548-555. (b) Soll, D.; Cherayil, J. D.; Bock, R. M.; *J. Mol. Biol.* **1967**, *29*, 97-112.
6. (a) Dickerson, R. E. *Methods in Enzymol.* **1992**, *211*, 67-111. (d) Saenger, W. Principles of Nucleic Acids structure. Springer-Verlag, New York, **1984**. (b) Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Van der M. G.; Van Boom, J. H.; Rich, *Science* **1981**, *211*, 171-176.
7. Bennett, C.F. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 215-24.
8. Ganesh, K. N. and Krishnan, Y. Nucleic Acids – Chemistry and Applications. *J. Org. Chem.* **2013**, *78*, 12283–12287.
9. (a) Seeman, N. C. *Nature* **2003**, *421*, 427-431. (b) Seeman, N. C. *Annu. Rev. Biochem.* **2010**, *79*, 65-87.
10. Uherek, C.; Wels, W. *Adv. Drug Deliv.* **2000**, *Rev. 44*:153–166
11. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 280.
12. (a) Armitage, B. A. *Nat. Chem. Biol.* **2005**, *1*, 185. (b) Quigley, H. A. *Lancet*, **2011**, *377*, 1367.
13. Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature*, **1998**, *391*, 806.
14. <http://www.landesbioscience.com>
15. Bernstein, E.; Caudy, A. A. Hammond, S. M. *et al. Nature* **2001**, *409*, 363–366
16. (a) Rand, T. A.; Ginalski, K.; Grishin, N. V.; Wang, X. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14385. (b) Matranga, C.; Tomari, Y.; Shin, C.; Bartel, D. P.; Zamore, P. D. *Cell* **2005**, *123*, 607.
17. Bernstein, E.; Caudy, A. A.; Hammond, S. M.; Hannon, G. J. *Nature*, **2001**, *409*, 363.
18. Lee, Y.; Kim, M.; Han, J.; Yeom, K.H.; Lee, S.; Baek, S. H.; Kim, V. N. *EMBO J.* **2004**, *23*, 4051.
19. Han, J.; Lee, Y.; Yeom, K. H. *et. al. Genes. Dev.* **2004**, *18*, 3016–3027
20. Lund, E.; Guttinger, S.; Calado, A.; Dahlberg J. E.; Kutay, U. *Science* **2004**, *303*, 95.
21. He, L.; Hannon, G. J. *Nat. Rev. Genet.* **2004**, *5*, 522.
22. Kiehntopf, M.; Esquivel, E. L.; Brach, M. A.; Herrmann, F. *Lancet* **1995**, *345*, 1027
23. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling D. E.; Cech, T. R.

- Cell*, **1982**, 31, 147
24. (a) Robinson, R. *PLoS Biol.* **2004**, 2, e28. (b) Mueller, D.; Stahl, U.; Meyer, V. J. *Microbiol. Methods.* **2006**, 65, 585; (c) Jason, T. L. H.; Koropatnick, J.; Berg, R. W. *Toxicol. Appl. Pharmacol.* **2004**, 201, 66.
25. (a) Usman, N.; Blatt, L. M.; Clin, J. *Invest.* **2000**, 106, 1197. (b) Peracchi, A. *Rev. Med. Virol.* **2004**, 14, 47.
26. Auser, J. M.; Barker, K. S.; Liu, T.T.; BlaB-Warmuth, J.; Homayouni, R.; Rogers, P. D. *PLoS Pathog.* **2007**, 3, e164.
27. Bielinska, A.; Shivdasani, R. A.; Zhang L. Q.; Nabel, G.J. *Science* **1990**, 250, 997.
28. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010**, 50, 259-293.
29. Stein, C. A.; Cohen, J. S. London: Macmillan Press. **1989**, p. 97.
30. Millar, P. S. London: Macmillan Press, **1989**, p. 79.
31. Froehler, B.; Ng, P.; Matteucci, M. *Nucleic Acids Res.* **1988**, 16, 4831-4839.
32. Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G. *Nucleic Acids Res.* **1986**, 14, 7421-7437.
33. (a) Sood, S.; Shaw, B. R.; Spielvogel, B. F.; *J. Am. Chem. Soc.* **1990**, 112, 9000-9001. (b) Shaw, B. R.; Madison, J.; sood, S.; Spielvogel, B. F. Humana Press, Inc. **1993**, 225-243. (c) Sergueev, D. S.; Shaw, B. R. *J. Am. Chem. Soc.* **1998**, 120, 9417-9427.
34. Uhlmann, E.; Peyman, A. *Chem Rev.* **1990**, 90, 543-584.
35. Brown, D. A.; Kang, S. H.; Gryaznov, S. M.; De Dionisio, L.; Heidenreich, O.; Sullivan, S.; Xu, X.; Neerenberg, M. I. *J. Biol. Chem.* **1994**, 43, 26801-26805.
36. Guvakova, M. A.; Yakubov, L. A.; Vlodaysky, I.; Tonkinson, J. L.; Stein, C. A. *J. Biol. Chem.* **1995**, 270, 2620-2627.
37. Rockwell, P.; O'Connor, W.; King, K.; Goldstein, N. I.; Zhang, L. M.; Stein, C. A. *Proc. Natl Acad. Sci. USA.* **1998**, 94, 6523-6528.
38. (a) Braasch, D. A.; Corey, D. R. *Chem. Biol.* **2001**, 8, 1-7. (b) Orum, H.; Wengel, J. *Curr. Opin. Mol. Ther.* **2001**, 3, 239-243.
39. Altona, C.; Sundaralingam, M.; *J. Am. Chem. Soc.* **1972**, 94, 8205-8212.
40. Teplova, M.; Minasov, G.; Tereshka, V.; Inamati, G. B.; Cook, P. D. *et al. Nat. Struct. Biol.* **1999**, 6, 535-539.
41. Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel, J.; Jacobsen, J. P. *Chem. Eur. J.* **2000**, 6, 2687-2695.
42. Braasch, D. A.; Liu, Y.; Corey, D. R. *Nucleic Acids Res.* **2002**, 30, 5160-5167.
43. Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S. *et al. J. Med. Chem.* **2009**, 52, 10-13.
44. Koizumi, M. *Curr. Opin. Mol. Ther.* **2006**, 8, 144-149.

45. Sazani, P.; Gemignani, F.; Kang, S. H.; Maier, M. A.; Manoharan, M. et al. *Nature Biotechnol.* **2002**, *20*, 1228-1233.
46. Alter, J.; Lou, F.; Rabinowitz, A.; Yin, H.; Rosenfeld, J. *Nat. Med.* **2006**, *12*, 175-177.
47. Wang, J.; Verbeure, B.; Luyten, I.; Lescrinier, E.; Froeyen, M.; Hendrix, C.; Rosemeyer, H.; Seela, F.; van Aerschot, A.; Herdewijn, P. *J. Am. Chem. Soc.* **2000**, *122*, 8595-8602.
48. Verbeure, B.; Lescrinier, E.; Wang, J.; Herdewijn, P. *Nucleic Acids Res.* **2001**, *29*, 4941-4947.
49. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.
50. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.
51. Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.
52. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science*, **1994**, *265*, 777-780.
53. Rasmussen, H.; Kastrop, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98-101.
54. (a) Leijon, M.; Graeslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry*, **1994**, *22*, 9820-9825. (b) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol.* **1996**, *3*, 410-413.
55. Bets, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838-1841.
- 56.(a) Eriksson, M.; Nielsen, P. E.; *Q. Rev. Biophys.* **1996**, *29*, 369-394. (b) Nielsen, P. E.; Egholm, M. *Curr. Iss. Molec. Biol.* **1999**, *1*, 89-104.
57. Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Rytte, A. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2632-2635.
58. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1501.
59. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1501. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Mol. Recogn.* **1994**, *7*, 165-170.
60. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Anti-Cancer Drug Design.* **1993**, *8*, 53-63. (b) Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481-1485.
61. Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624-630.
62. Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668-698.
63. Simonsson, T. *Biol. Chem.* **2001**, *382*, 621-628.
64. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *265*, 566-568.
65. Ghosh, K. Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944-5945.
66. Datta, B.; Schmitt, C.; Armitage, B. A. *J. Am. Chem. Soc.* **2003**, *125*, 4111-4118. (b) Marin, V. L.; Armitage, B. A. *J. Am. Chem. Soc.* **2005**, *127*, 8032-8033.
67. Sharma, N. K.; Ganesh, K. N. *Chem. Commun.* **2005**, 4330-4332.
68. Demidov, V. V.; Frank-Kamenetskii, M. D.; *Trends Biochem. Sci.* **2004**, *29*, 62-71.

69. Nielsen, P. E. *Pure Appl. Chem.* **1998**, *70*, 105-110.
70. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5-23.
71. Koppelhus, U.; Nielsen, P. E. *Adv. Drug Delivery Rev.* **2003**, *55*, 267-280.
72. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404-412.
73. Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. *Curr. Top. Med. Chem.* **2011**, *11*, 1535-1554.
74. Dueholm, K. L.; Petersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1077-1080.
75. Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. *Eur. J. Org. Chem.* **2000**, *2000*, 2905-2913.
76. Menchise, V.; de Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 12021-12026.
77. Koppelhus, U.; Awasthi, S. K.; Zachar, V.; Holst, H. U.; Ebbesen, P.; Nielsen, P. E. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 51-63.
78. Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; O'Keefe, H.; Vatta, P.; Hyldig-Nielsen, J. J.; Ly, D. H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4931-4935.
79. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939-1942.
80. Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998**, *39*, 4707-4710.
81. Sugiyama, T.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 7317-7320.
82. Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett.* **1994**, *35*, 5173-5176.
83. Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. *Tetrahedron Lett.* **2005**, *46*, 8395-8399.
84. Englund, E. A.; Appella, D. H. *Org. Lett.* **2005**, *7*, 3465-3467.
85. Dose, C.; Seitz, O. *Org. Lett.* **2005**, *7*, 4365-4368.
86. Ficht, S.; Dose, C.; Seitz, O. *ChemBioChem.* **2005**, *6*, 2098-2103.
87. Englund, E. A.; Wang, D.; Fujigaki, H.; Sakai, H.; Micklitsch, C. M.; Ghirlando, R.; Martin-Manso, G. *Nat. Commun.* **2012**, *3*, 614.
88. Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. *Eur. J. Org. Chem.* **2007**, *2007*, 5879-5885.
89. Avitabile, C.; Moggio, L.; Malgieri, G.; Capasso, D.; Gaetano, S. D.; Saviano, M.; Pedone, C.; Romanelli, A. *PLoS One* **2012**, *7*, e35774.
90. Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258-10267.
91. Yeh, J. I.; Boris Shivachev, B.; Rapireddy, S.; Crawford, M. J.; Gil, R. R.; Du, S.; Madrid, M.; Ly, D. H. *J. Am. Chem. Soc.* **2010**, *132*, 10717-10727.

92. Crawford, M. J.; Rapireddy, S.; Bahal, R.; Sacui, I.; Ly, D. H. J. *Nucleic Acids* **2011**, 2011, doi:10.4061/2011/652702.
93. Mitra, R.; Ganesh, K. N. *J. Org. Chem.* **2012**, 77, 5696-5704.
94. Jain, D. R.; Ganesh, K. N. *J. Org. Chem.* **2014**, 79, 6708-6714.
95. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc.* **1996**, 118, 5544-5549.
96. Hollenstein, M.; Leumann, C. *J. Org. Lett.* **2003**, 5, 1987-1990.
97. Rose, D. J. *Anal. Chem.* **1993**, 65, 3545-3549.
98. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010**, 50, 259-293
99. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1996**, 52, 15017-15030.
100. Gangamani, B. P.; D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides* **1999**, 18, 1409-1011.
101. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, 55, 177-192.
102. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **1999**, 1, 1513-1516.
103. Vilaivan, T.; Khongdeesameor, C.; Harnyuttanokam, P.; Westwell, M. S.; Lowe, G. *Bioorg. Med. Chem. Lett.* **2000**, 10, 2541-2545.
104. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett.* **2001**, 3, 1281-1284.
105. Sharma, N.; Ganesh, K. N. *Tetrahedron Lett.* **2004**, 45, 1403-1406.
106. Sharma, N.; Ganesh, K. N. *Chem. Commun.* **2003**, 0, 2484-2485.
107. Puschl, A.; Boesen, T.; Zuccarello, G.; Dahl, O.; Pitsch, S.; Nielsen, P. E. *J. Org. Chem.* **2001**, 66, 707-712.
108. Hickman, D. T.; King, P. M.; Cooper, M. A.; Slater, J. M.; Mickelfield, J. *Chem. Commun.* **2000**, 2251-2252.
109. Kumar, V. A.; Pallan, P. S.; Meena,; Ganesh, K. N. *Org. Lett.* **2001**, 3, 1269-1272.
110. Kumar, V. A.; Meena. *Nucleos Nucleot Nucleic acids* **2003**, 22, 1101-1104.
111. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron. Lett.* **2002**, 43, 883-886.
112. Lonkar, P.; Ganesh, K. N.; Kumar, V. A. *Org. Biomol. Chem.* **2004**, 2, 2604-2611
113. Lonkar, P. S.; Kumar, V. A.; Ganesh, K. N. *Nucleos. Nucleot Nucleic Acids* **2001**, 20, 1197-1200.
114. Lescrinier, E.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res.* **2003**, 31, 2975-2989.
115. Lagrioule, P.; Wittung, P.; Eriksson, M.; Jensen, K. K.; Norden, B.; Buchardt, O.; Nielsen, P. E. *Chem. Eur. J.* **1997**, 3, 912-919.
116. a) Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2004**, 69, 1858-1865. b) Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2004**, 69, 5725-5734.
117. Branda, N.; Kurz, G.; Lehn, J.-M. *Chem. Commun.* **1996**, 2, 2443-2444.
118. Marsh, A.; Nolen, E. G.; Gardinier, K. M.; Lehn, J.-M. *Tetrahedron Letters* **1994**, 35, 397-400.

119. Fenniri, H.; Mathivanan, P.; Vidale, K. L.; Sherman, D. M.; Hallenga, K.; Wood, K. V.; Stowell, J. G. *J. Am. Chem. Soc.* **2001**, *123*, 3854-3855.
120. (a) Mascal, M.; Hext, N. M.; Warmuth, R.; Arnall-Culliford, J. R.; Moore, M. H.; Turkenburg, J. P. *J. Org. Chem.* **1999**, *64*, 8479-8484. (b) Mark, M.; Hext, N. M.; Warmuth, R.; Moore, M. H.; Turkenburg, J. P. *Angew. Chem. Int. Ed.* **1996**, *35*(19), 2204-2206
121. Asadi, A.; Patrick, B. O.; Perrin, D. M. *J. Org. Chem.* **2007**, *72*, 466- 475.
122. (a) Morales, J. G.; Ruez, J.; Yamazaki, T.; Motkuri, R. K.; Kovalenko, A.; Fenniri, H. *J. Am. Chem. Soc.* **2005**, *127*, 8307-8309. (b) Fenniri, H.; Deng, B. L.; Ribbe, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11064-11072. (c) Fenniri, H.; Deng, B. L.; Ribbe, A. E.; Hallenga, K.; Jacob, J.; Thiagarajan, P. *Proc. Natl. Acad. Sci. U.S.A* **2002**, *99*, 6487-6492.
123. (a) Zhao, H.; Huang, W.; Wu, X.; Xing, Z.; He, Y.; Chen, Q. *Chem Commun.* **2012**, *48*, 6097-6099. (b) Yang, H. Z.; Pan, M. Y.; Jiang, D. W.; He, Y. *Org. Biomol. Chem.* **2011**, *9*, 1516-1522. (c) Pan, M.Y.; Wu, X. H.; Luo, D. B.; Huang, W.; He, Y. *Acta Crystallographica Section C.* **2011**, *67*, 0175-0178. (d) Pan, M. Y.; Hang, W.; Zhao, X. J.; Zhao, H.; Deng, P. C.; Xing, Z. H.; Qing, Y.; He, Y. *Org. Biomol. Chem.* **2011**, *9*, 5692-5702.
124. Asadi, A.; Patrick, B. O.; Perrin, D. M. *J. Org. Chem.* **2006**, *72*, 466-475.
125. Beingessner, R. L.; Diaz, J. A.; Hemraz, U. D.; Fenniri, H. *Tetrahedron Letters* **2011**, *52*, 661-664.
126. Chen, D.; Meena; Sharma, S. K.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2004**, *126*, 70-71.
127. Chen, H.; Meena; McLaughlin, L. W. *J. Am. Chem. Soc.* **2008**, *130*, 13190-13191.
128. Vysabbattar, R.; Ganesh, K. N. *Tetrahedron Letters* **2008**, *49*, 1314-1318.
129. Lenzi, A.; Reginato, G.; Taddei, M.; Trifilieff, E. *Tetrahedron Letters* **1995**, *36*, 1717-1718.
130. Diederichsen, U. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 445-448.
131. Huang, Y.; Dey, S.; Zhang, X.; Sonnichsen, F.; Garner, P. *J. Am. Chem. Soc.* **2004**, *126*, 4626-4640.
132. Mittapalli, G. K.; Reddy, K. R.; Xiong, H.; Munoz, O.; Han, B.; De Riccardis, F.; Krishnamurthy, R.; Eschenmoser, A. *Angew. Chem. Int. Ed.* **2007**, *46*, 2470-2477.
133. Taylor, R. W.; Chinnery, P. F.; Turnbull, D. M.; Lightowers, R. N. *Nature Genet.* **1997**, *15*, 212-215.
134. Colton, M. M.; Tanowitz, M.; Donner, A. J.; Prakash, T. P.; Swayze, E.E.; Harris, E. N.; Seth, P.P. *Nucleic Acid Therapeutics.* **2018**, *00*, 1-9.
135. Kanasty, R.; Dorkin, J.; Vegas, A. *et al. Nature Mater* **2013**, *12*, 967-977
136. Kurreck, J. *Eur. J. Biochem.* **2003**, *270*, 1628-1644.
137. Isis Pharmaceuticals / Genzyme
138. Zhang, et al. *Medicine* **2019**, *98*:6

139. Stein, C. A.; Castanotto D. *Molecular Therapy Rev.* **2017**, *25*, 1069-1075.
140. Thadke, S. A.; Hridya, V. M.; Perera, J. D. R.; Gil, R. R.; Mukherjee, A.; Ly, D. H. *Commun. Chem.* **2018**, *1*, 79.
141. Hsieh, W. C.; Bahal, R.; Thadke, S. A.; Bhatt, K.; Sobczak, K.; Thornton, C.; Ly, D. H. *Biochemistry* **2018**, *57*, 907-911.
142. Xia, X.; Zhou, Z.; Desantis, C.; Rossi, J. J.; Bong, D. *ACS Chemical Biology* **2019**, *14*, 1310-1318.
143. Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.* **2014**, *42*, 8796-8807.
144. (a) Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G. *Mol. Biotechnol.* **2006**, *33*, 239-254. (b) Bodansky, M.; Bodansky, A. *Practice of Peptide Synthesis*, Springer-Verlog, Berlin, **1984**. (c) Stewart, J. M.; Young, J. D. *Solid Phase Peptide Synthesis*, W. H. Freeman & Co, New York, **1969**.
- 145.(a) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154. (b) Connah, L.; Joshi, R.; Vibhute, S.; Gambino, G.; Correia, J. D. G.; Angelovski, G.; *Org. Lett.* **2019**, *21*, 5378-5382. (c) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoet, M.; Monbaliu, J. C. M.; Melnyk, O. *Chem. Rev.* **2019**, *119*, 7328-7443. (d) Isidro-Llobet, A.; Kenworthy, M. N.; Mukherjee, S.; Kopach, M. E.; Wegner, K.; Gallou, F.; Smith, A. G.; Roschangar, F. *J. Org. Chem.* **2019**, *84*, 4615-4628. (e) Kulkarni, S. S.; Wang, C.C.; Sabbavarapu, N. M.; Podilapu, A. R.; Liao, P. H.; Hung, S. C. *Chem. Rev.* **2018**, *118*, 8025-8104.
146. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K.; Hansen, H. F.; Koch, C.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Peptide Sci.* **1995**, *3*, 175-183.
147. Muttenthaler, M.; Albericio, F.; Dawson, P. E. *Nat. Protoc.* **2015**, *10*, 1067.
148. Orain, D.; Ellard, J.; Bradley, M. *J. Comb. Chem.* **2002**, *4*, 1-16.
149. Hansen, P. R.; Oddo, A. in *Peptide Antibodies: Methods and Protocols* (Ed.: G. Houen), Springer New York, New York, NY, **2015**, pp. 33-50.
150. Brieke, C.; Cryle, M. J. *Org. Lett.* **2014**, *16*, 2454-2457.
151. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Frier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Neilsen, P. E. *Nature*, **1993**, *365*, 566-568.
152. (a) Kempen E. C.; Brodbelt J. S. *Anal. Chem.* **2000**, *72*, 5411-5416. (b) Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.
153. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. *Eur. J. Mass Spectrum.* **2001**, *36*, 849-865.
154. Brand, L.; Johnson, M. L. *Methods Enzymol.* **1997**, *278*, 1-628.
155. Cain, B. F.; Baguley, B. C.; Denny, W. A. *J. Med. Chem.* **1978**, *21*, 658-668.
156. Gershan, H.; Chiralando, R.; Guttman, S. B.; Minsky, A. *Biochemistry* **1993**, *32*, 7143.
157. Minsky, M. *Scanning* **1988**, *10*, 128-138.

158. Minsky, M. US patent 3013467, 12 Dec 1961.
159. <http://emu.uct.ac.za/training/intro-to-electron-microscopy-for-biologists/confocal-laserscanning-microscopy/>
160. Gourishankar, A.; Shukla, S.; Ganesh, K. N.; Sastry, M. *J. Am. Chem. Soc.* **2004**, *126*, 13186-13187.
161. Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *19*, 213.
162. Tanowitz M., Hettrick L., Revenko A., Kinberger G.A., Prakash T.P., Seth P.P. *Nucleic Acids Res.* **2017**; *45*:12388–12400.

Chapter 2

Synthesis and Characterization of Tri antennary N-Acetyl Galactosamine (GalNAc) and PNA Oligomers

2.1 Introduction

Antisense oligonucleotide therapeutics originate from the specific molecular recognition event between the mRNA of the gene to be inhibited and the synthetic oligonucleotide drug. In addition to this key molecular recognition process, pharmacokinetic and pharmacodynamics requirements must be satisfied in order to reach the desired therapeutic end point. The success of antisense drug development technology relies on efficient delivery of therapeutic agent into target cells in the patient to inhibit the translation of the desired mRNA or its degradation.

In comparison to antisense inhibition, RNAi is an endogenous pathway for post-transcriptional silencing of gene expression that is triggered by double-stranded RNA (dsRNA), including endogenous microRNA (miRNA) and synthetic short interfering RNA (siRNA). By activating this pathway, siRNAs can silence the expression of virtually any gene with high efficiency and specificity, including targets traditionally considered as 'undruggable'. The therapeutic potential of this method is far-reaching, and siRNA-based therapeutics are under development for the treatment of diseases ranging from viral infections^{1,2} to cancer and other gene related diseases. siRNA molecules must be delivered to the interior of target cells for incorporation into the RNAi machinery. As siRNA molecules are too large and too hydrophilic to diffuse across cell membranes alone, delivery mechanisms through chemical modification is necessary to assist their uptake by target cells³. When administered systemically, siRNA faces several physiological barriers in reaching its site of action. The delivery systems must be engineered to provide (i) stability against serum nucleases, (ii) avoidance of non-specific interactions with serum proteins and non-target cells, (iii) evasion of the immune system (iv) prevention of renal clearance, (v) exit from blood vessels to reach target tissues, (vi) cell entry and (vii) incorporation into the RNAi machinery.^{3,4-6} A key challenge to realizing the broad potential of siRNA-based therapeutics is the need for safe and effective delivery methods. The conjugation of trivalent N-acetyl galactosamine (GalNAc₃) to siRNA duplexes or single stranded ASOs enhances potency by 10-60 fold for inhibiting gene targets expressed in hepatocytes.^{7,8} The improved potency is a result of targeted delivery of the GalNAc₃ conjugated oligonucleotides to hepatocytes via the asialoglycoprotein receptor (ASGR).⁹ This receptor is abundantly expressed almost exclusively on hepatocytes and clears glycoproteins and desialylated platelets from circulation.¹⁰

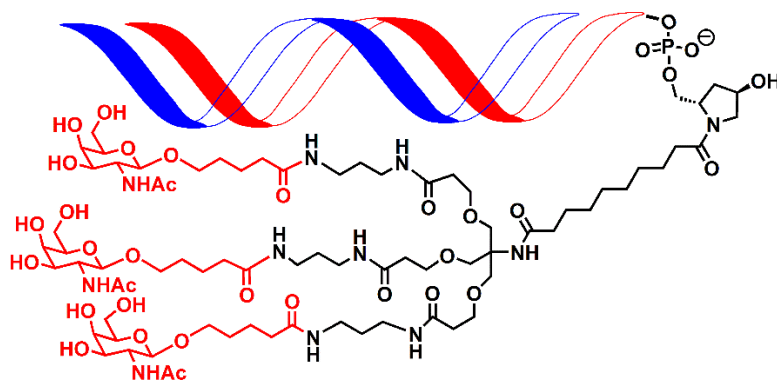


Figure 2.1 Structure of the triantennary GalNAc₃-siRNA conjugate⁷

Over the past several decades peptide nucleic acid (PNA)¹¹ has emerged as one of the most potential 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 2.2).

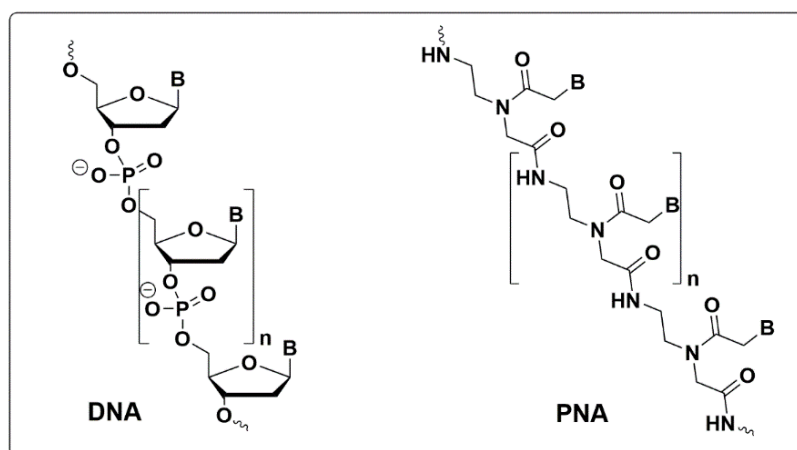


Figure 2.2 Structural comparison of DNA and PNA (B = nucleobase)

The advantage of using PNA over DNA/RNA as antisense scaffold are the following: 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* and iv) as opposed to RNA, the polyamide backbone eliminates the possibility of self-cleavage.¹² However PNA suffers from limitations of poor water solubility and low cellular uptake. Various chemically edited PNA analogues have been reported in the recent literature¹⁴ to address these drawbacks.

2.2 Rationale of the present work

Conjugation of drugs to ligands that bind to cell-surface receptors abundantly expressed by certain cell types on specific tissues in certain diseased conditions is a promising approach for targeted drug delivery. Recently Manoharan *et al.*⁸ have shown that efficient antisense activity of siRNA conjugated to N-acetylgalactosamine (GalNAc) that is a specific ligand for the asialoglycoprotein cell surface receptor (ASGPR) expressed by hepatocytes. Upon subcutaneous (SC) administration, these GalNAc₃-siRNA conjugates robustly suppress gene expression of the targeted mRNA in liver. They found that conjugation of optimized, chemically modified siRNAs to an engineered ASGPR ligand resulted in improved systemic stability against nucleases and favourable pharmacokinetics relative to the unconjugated siRNAs. These conjugates also mediated durable silencing of the targeted gene in the liver following single or multiple low-volume SC administrations. As PNA is a potential antisense oligonucleotide, it was surmised that the conjugation of triantennary GalNAc₃ moiety with a spacer (1) to PNA would impart receptor mediated cell penetration of conjugated PNA (2).

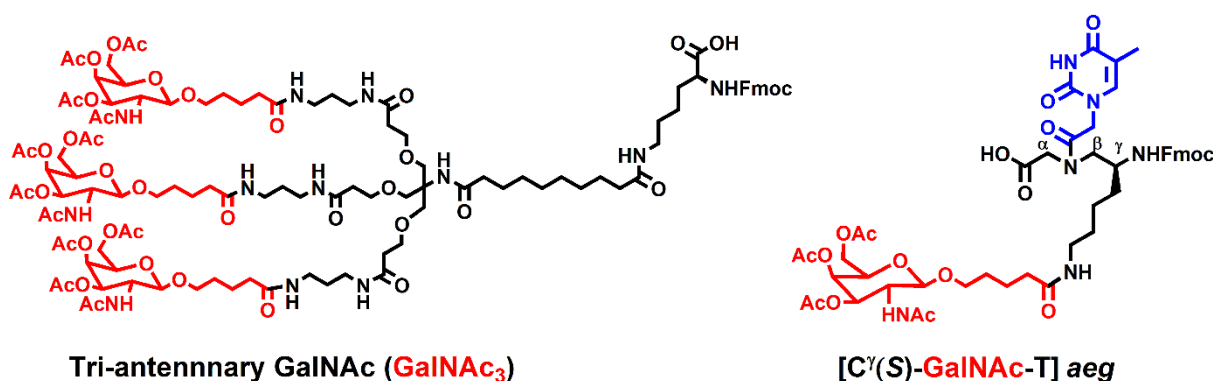


Figure 2.3 Chemical structures of target compounds

This PNA is analogous to GalNAc₃ conjugated siRNA targeting transthyretin (TTR) mRNAs that exhibited good activity for reducing TTR protein in blood in early stage human trials. The present work deals with the synthesis of triantennary GalNAc from precursors and its conjugation with PNA at the N-terminus (Figure 2.4 A). In addition, PNA analogues conjugated with mono GalNAc at C^γ position of PNA backbone on three successive units at the N-terminus (Figure 2.4B) was also synthesized to compare its properties with the triantennary PNA conjugate.

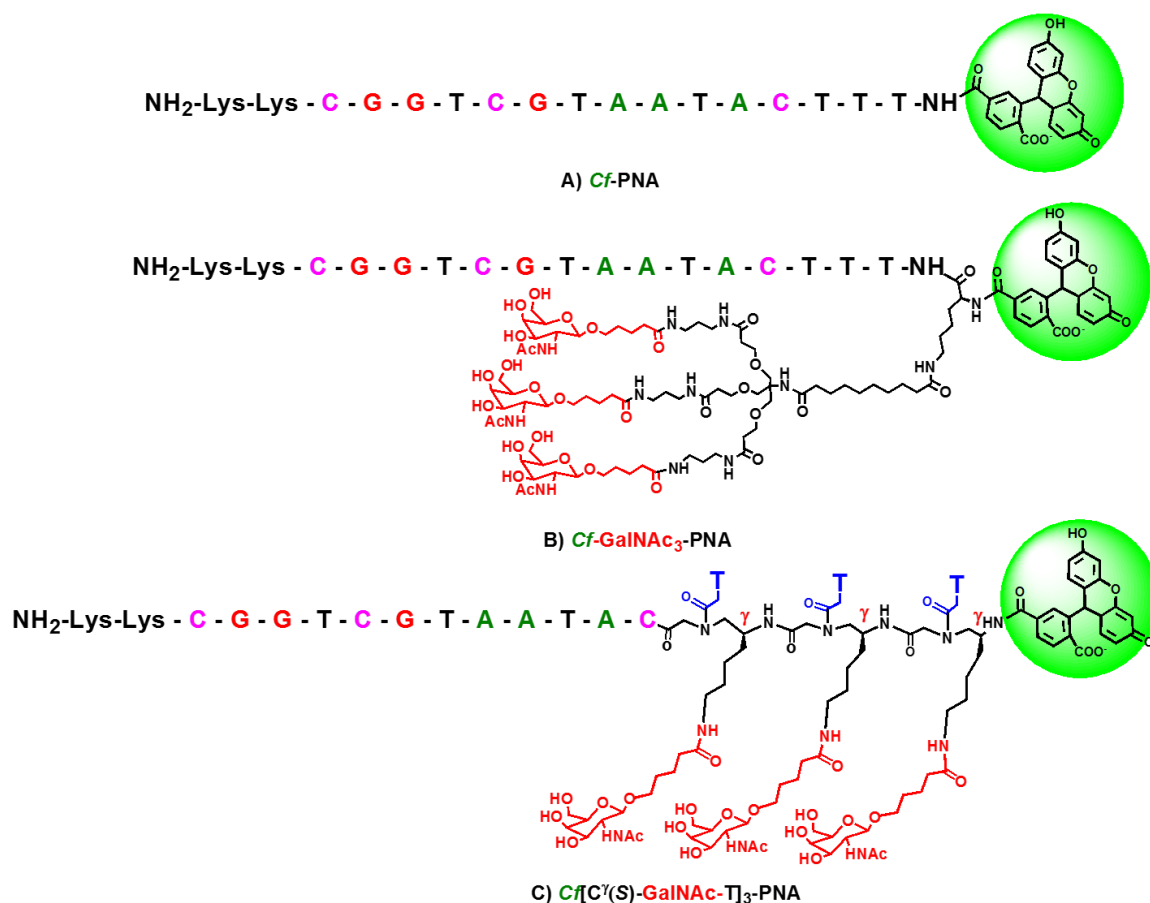


Figure 2.4 Structure of Carboxyfluorescein tagged control PNA and GalNAc PNA

2.3. Objective of present Work

- Synthesis and characterization of tri-GalNAc Fmoc Lysine moiety (1)
- Synthesis and characterization of [C^γ(S)-GalNAc-T] aeg monomer (2)
- Synthesis of triantennary GalNAc₃ conjugated PNA oligomer (A)
- Synthesis of [C^γ(S)-GalNAc-T]₃ conjugated PNA oligomer (B)
- Purification by HPLC and characterization by MALDI-TOF spectroscopy of GalNAc₃ conjugated PNA oligomers.

2.4 Results and Discussion

This section describes the synthesis of triantennary GalNAc₃ (1) and [C^γ(S)-GalNAc-T] PNA monomers (2) and their characterization by NMR and mass spectral data. Conjugation of triantennary GalNAc₃ and [C^γ(S)-GalNAc-T] PNA monomers to target PNA oligomers at N-terminus was done by solid phase synthesis. The conjugates were cleaved from solid support, purified by HPLC and characterized by mass spectral data.

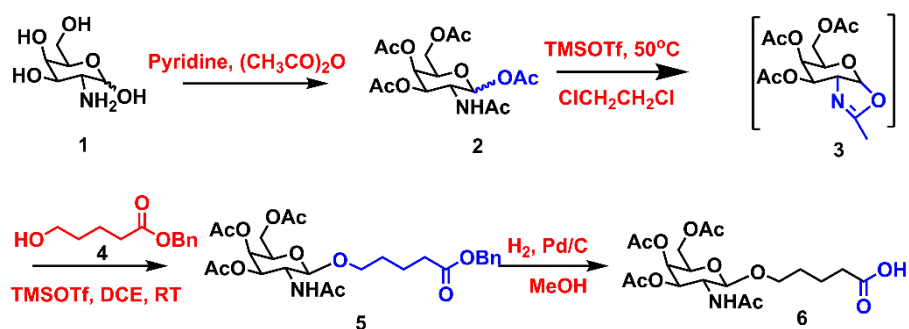
2.4.1 Synthesis of triantennary GalNAc (1) and [C^γ(S)-GalNAc-T]PNA monomer (2).

This section describes the synthesis of triantennary N-Acetyl Galactosamine (Scheme 2.1).

2.4.1a Synthesis of 5-[[3,4,6-tri-O-acetyl-2-(acetylamino)-2-deoxy-β-D-galactopyranosyl]oxy]-pentanoic acid¹ (6)

The commercially available D-(+)-galactosamine was acetylated to obtain the pentacetate **1**, which was transformed into the oxazoline¹⁴ derivative **2** by treatment with TMA-triflate and in situ glycosylated with benzyl 5-hydroxy pentanoate **4** to obtain the galactosyl pentanoate ester **5**. The glycosylation afforded exclusively β-isomer due to the anchimeric assistance from the C₂ equatorial N-acetyl at the transition stage of the reaction.¹⁵ This was deprotected to the corresponding acid 5-[[3,4,6-tri-O-acetyl-2-(acetylamino)-2-deoxy-β-D-galactopyranosyl]oxy]-pentanoic acid⁷ **6** (Scheme 2.1).

Scheme 2.1 Synthesis Of 5-[[3,4,6-tri-O-acetyl-2-(acetylamino)-2-deoxy-β-D-galactopyranosyl]oxy]-pentanoic acid¹ (**6**).

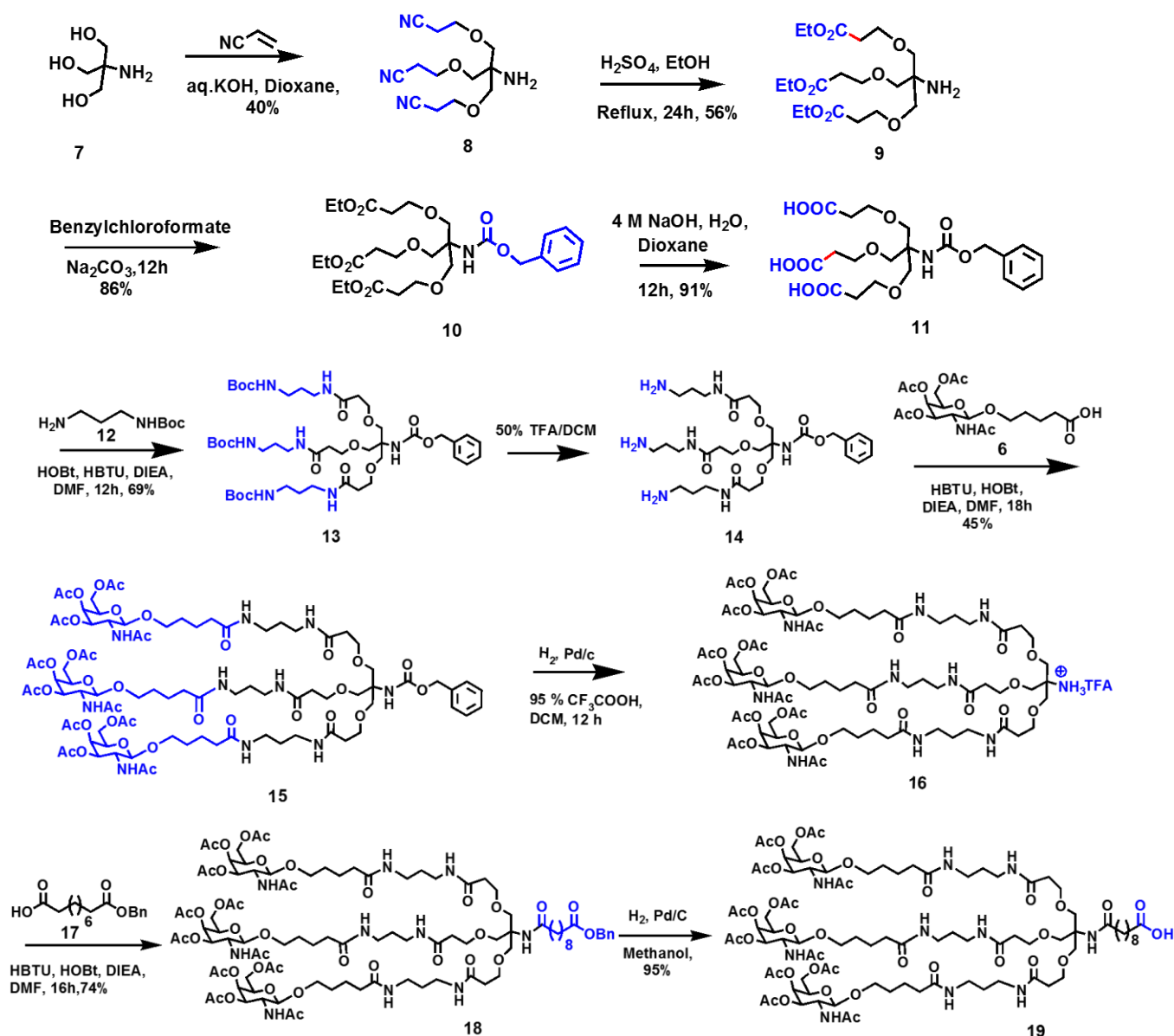


2.4.1b Synthesis of triantennary N-acetyl galactosamine (Tri-GalNAc₃) (19)

The alkylation of the three hydroxyl functions of TRIS (**7**) via a Michael addition to acrylonitrile¹⁶ yielded aminotrinitrile **8** (Scheme 2.2). By refluxing compound **8** with a solution of hydrochloric acid in ethanol, triethylester **9** was obtained¹⁷. Subsequently, the amino function of **9** was protected with benzyl chloroformate to Cbz protected tri-ester **10**. Saponification of the ethyl esters of **10**, followed by HBTU/ HOBt-mediated condensation¹⁶ of the resulting trivalent carboxylic acid **11** with N-Boc-1,3-diaminopropane¹⁸ **12**, furnished the dendrimeric synthon¹⁷ **13**. The three aminopropyl groups of **13** were selectively deprotected with TFA in dichloromethane, leading to **14**.

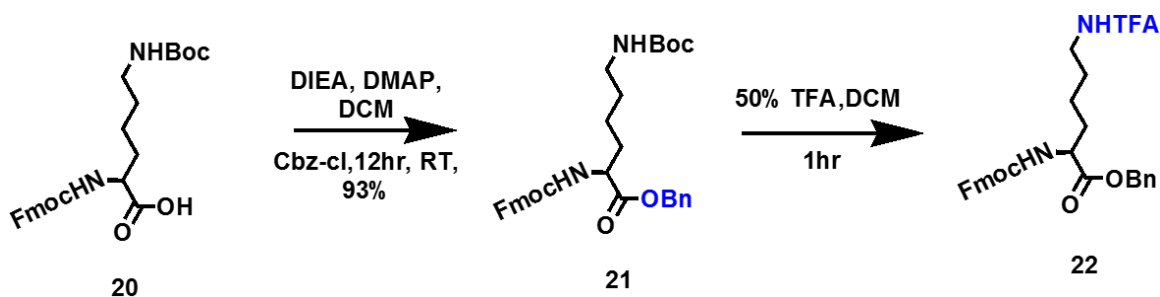
The activation of the carboxylic acid **6** with HBTU in the presence of DIEA in DMF followed by addition of the triamine¹⁹ **14** afforded the triantennary GalNAc scaffold **15** in

58% yield. Catalytic hydrogenation of **15** at atmospheric pressure over 10% Pd-C in the presence of acetic acid in methanol for 18 h afforded the amine⁷ **16**. After completion of the reaction, the compound **16** was treated with trifluoroacetic acid to isolate the amine in the form of ammonium salt to avoid possible *O*→*N* acetyl migration from the sugar moieties. Compound **16** was treated with HBTU-activated monobenzyl 1,10-decanoic acid **17** in the presence of DIEA in DMF to afford the benzyl ester **18** in 74% yield. Catalytic hydrogenation of the benzyl ester **18** over 10% Pd-C at atmospheric pressure for 18 h afforded carboxylic acid⁸ **19** in 95% yield.

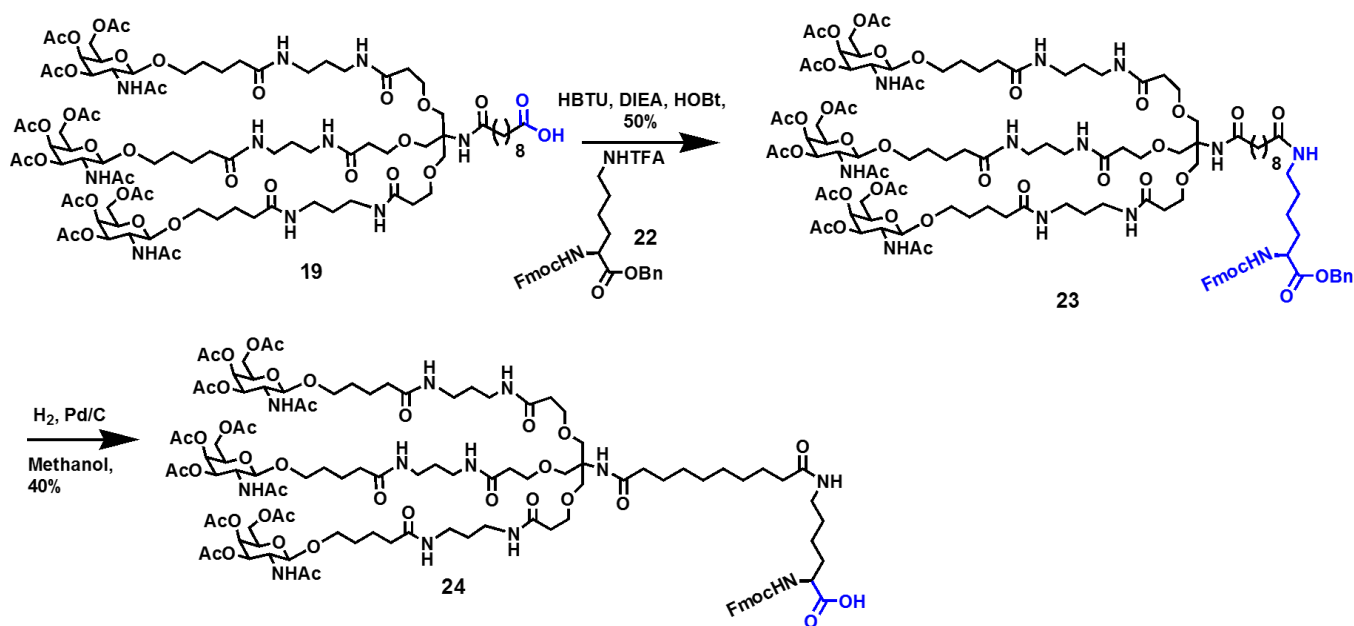
Scheme 2.2 Synthesis of triantennary GalNAc₃ (**19**)

2.4.1c. Synthesis of Fmoc lysine NH^{TFA} benzyl ester (**22**)

Benylation of $C_2(NHFmoc)-\omega(NHBoc)$ -LysOH with benzyl chloroformate gave the benzyl ester **21** as shown in scheme 2.3²⁰. Deprotection of $\omega(NHBoc)$ by 50% TFA yielded ω -amine **21** in the form of TFA salt of Fmoc Lysine benzyl ester **22**.

Scheme 2.3 Synthesis of Fmoc Lysine NH^{TFA} benzyl ester.(**22**)2.4.1d Synthesis of Fmoc Lys triantennary *N*-acetyl galactosamine (**24**)

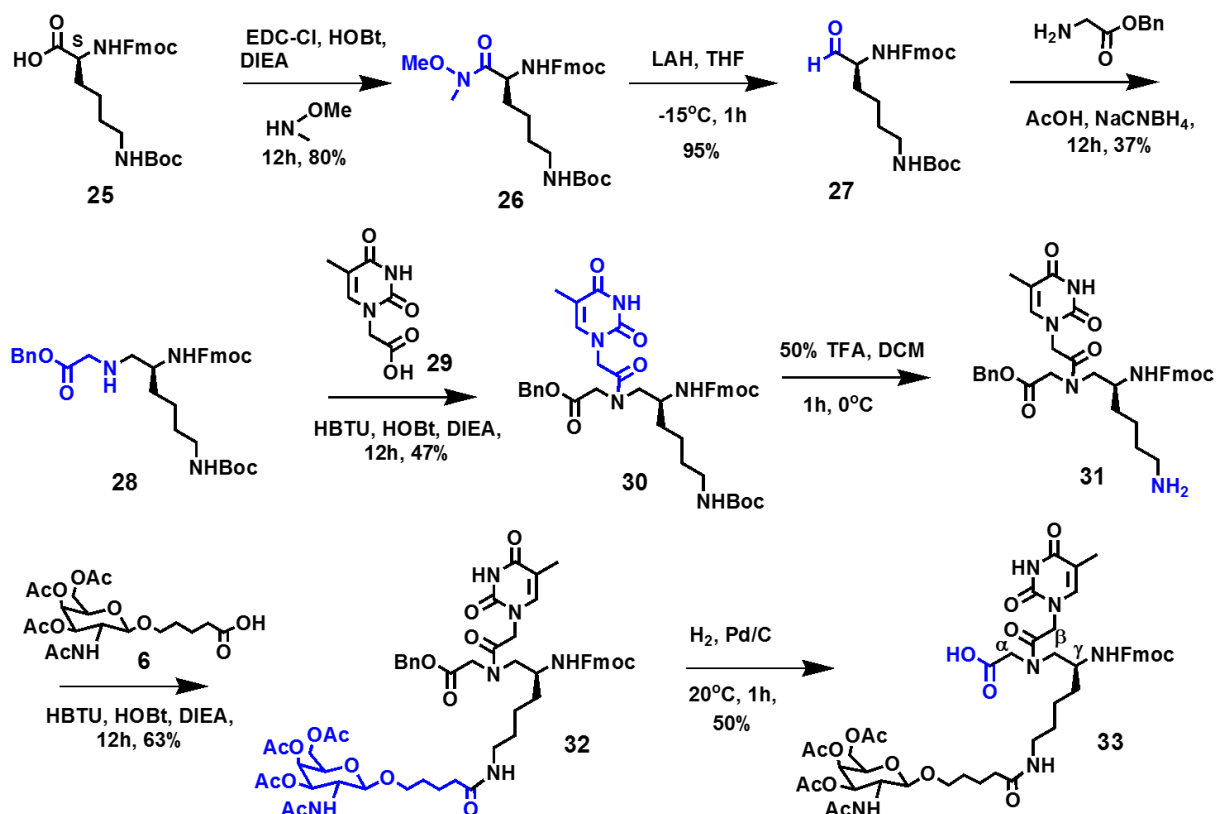
Activation of the carboxylic acid **19** with HBTU in the presence of DIEA in DMF followed by addition of the Fmoc Lysine TFA salt **22** afforded Fmoc lysine triantennary GalNAc benzyl ester **23** (Scheme 2.4). Further catalytic hydrogenation of the benzyl ester **23** over 10% Pd-C at atmospheric pressure and by maintaining temperature at 20-25 °C for 1hr to **24** in 96% yield.

Scheme 2.4 Synthesis of Fmoc Lys- NH -GalNAc₃ (**24**)

2.4.1e Synthesis of (*C*^γ(*S*)-GalNAc-*T*) PNA Monomer (**33**)

The activation of the acid function of Fmoc Lys (Boc)-OH **25** by *N,N*-methoxymethyl amine yielded compound **26**. The activated ester was then subjected to oxidation in presence of lithium aluminium hydride in THF at -15 °C to get aldehyde **27**, which was coupled with glycine benzyl ester under reductive amination conditions to obtain (*S*)-NHFmoc (ω -NHBoc) Lys benzyl glycinate **28**. This was reacted with thymine acetic acid **29** in the presence HBTU, HOBt, DIEA to result in the benzyl ester *C*^γ-substituted PNA monomer **30**. The deprotection of the Boc group to get free amine **31**, which on coupling with 5-[[3,4,6-tri-*O*-acetyl-2-(acetylamino)-2-deoxy- β -D-galactopyranosyl]oxy]-pentanoic acid **6** yielded *C*^γ(*S*)-GalNAc substituted compound **32**. Finally catalytic hydrogenation of **32** at atmospheric pressure over 10% Pd-C in the presence of acetic acid in methanol for 1h at 20 °C- 25 °C afforded *C*^γ-GalNAc substituted PNA monomer [*C*^γ(*S*)-GalNAc-*T*]-OH **33**.

Scheme 2.4. Synthesis of (*S*) γ (butylamido)- GalNAc PNA monomer (**33**)



2.4.2. Solid Phase Peptide Synthesis

Synthesis of PNA oligomers were carried out using solid phase synthesis protocol using Fmoc strategy. PNA oligomerization was carried out from C-terminus to the N-

terminus end using PNA monomers with protected amino and carboxylic acid functions maintaining the orthogonality.

2.4.2a Synthesis of PNA oligomers

Rink amide 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 (Fmoc/Boc) 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 0.22 mmol/g by partial acylation of amine content using calculated amount of acetic anhydride. The free amine groups on the resin available for coupling was confirmed before starting synthesis by performing Kaiser's test.

The deprotection of the Fmoc protecting group and the completion of the coupling reaction were monitored by Kaiser's test.²¹ The Fmoc deprotection leads to a positive Kaiser's test, where the resin beads show blue color (Rheumann's purple). On the other hand, after completion of coupling reaction the resin beads were colorless which means a negative Kaiser's test. It is the most widely used qualitative test for the presence or absence of free amino group (deprotection/coupling). Using the standard solid phase synthesis protocol (Figure.2.5), the PNA oligomers of desired length incorporating modified as well as unmodified PNA monomers at desired positions were synthesized.

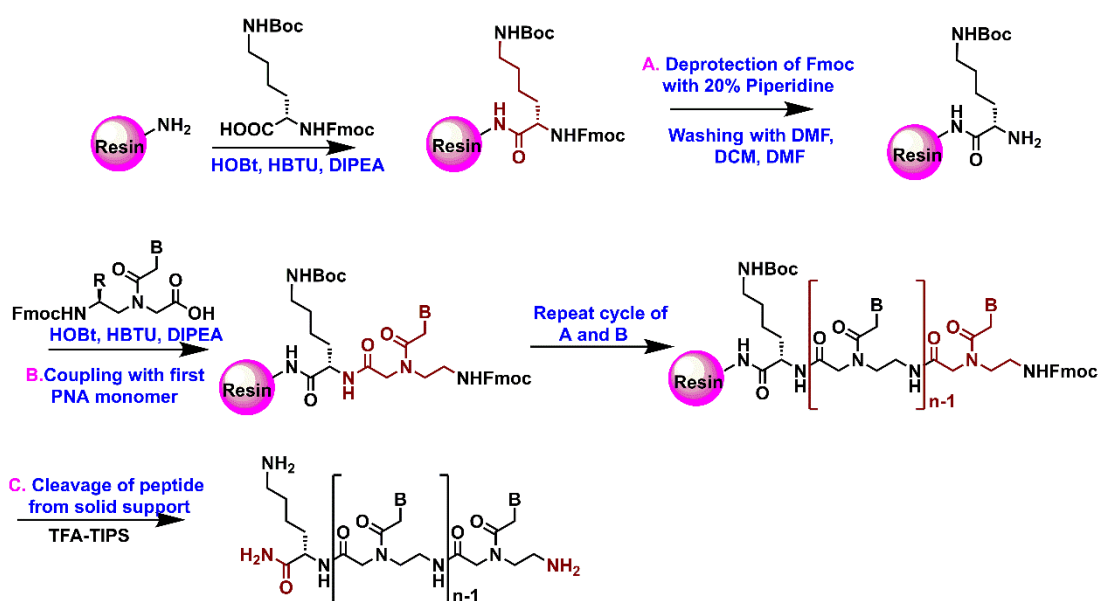
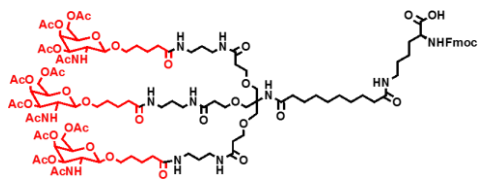
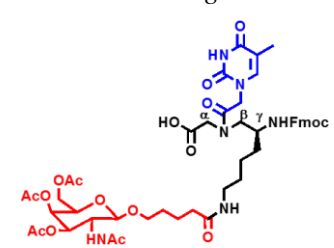


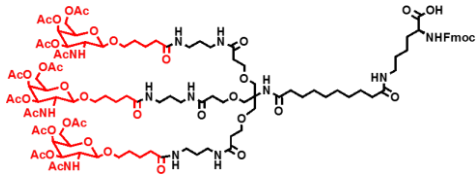
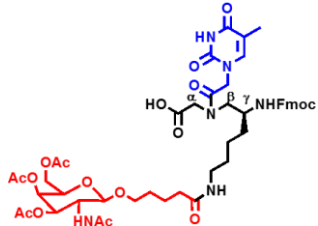
Figure 2.5 Solid phase PNA synthesis protocol by *Fmoc* strategy; B = Nucleobase such as Adenine (A), Guanine (G), Cytosine (C) and Thymine (T)

2.4.2b Synthesis of GalNAc conjugated PNA oligomers

For the purpose of the cell permeability and DNA/RNA binding studies, various PNA oligomers were synthesized following the solid phase protocol using Fmoc-strategy (Table 2.1). After the final step of regular PNA monomer addition, coupling with carboxy fluorescein gave the fluorescent PNA oligomer *Cf*-PNA **4**. (Table 2.1, Entry 4). The triantennary GalNAc acid **19** was coupled using pybop, HOBt and DIEA to obtain the GalNAc₃-PNA **2** (Table 2.1, Entry 2). Conjugation of triantennary GalNAc₃ Fmoc lysine **24** with PNA **1** followed by Fmoc deprotection and coupling with carboxy fluorescein to obtain carboxy fluorescein tagged *Cf*-GalNAc₃-PNA **5** (Table 2.1, Entry 5). The 12 mer PNA synthesized by on solid phase and three successive coupling with (S) C^γ(butylamido)-GalNAc PNA monomer **33** (Table 2.1, Entry 3) by using same coupling reagents to obtain 15 mer [C^γ(S)GalNAc-T]₃ PNA **3** followed by deprotection of Fmoc group and coupling with carboxy fluorescein gave *Cf*-[C^γ(S)GalNAc-T]₃ PNA **6** (Table 2.1, Entry 6).

Table 2.1 PNA oligomers and GalNAc conjugated PNA oligomers

Entry	PNA sequences	Monomers used
1	H-TTTCATAATGCTGGC- <i>Lys-Lys</i> -NH ₂ PNA 1	A/G/C/T = <i>aeg</i> PNA
2	GalNAc ₃ -TTTCATAATGCTGGC- <i>Lys-Lys</i> -NH ₂ PNA 2	A/G/C/T = <i>aeg</i> PNA 
3	[C ^γ (S)GalNAc-T] ₃ -TTTCATAATGCTGGC- <i>Lys-Lys</i> NH ₂ PNA 3	A/G/C/T = <i>aeg</i> PNA 
4	<i>Cf</i> -TTTCATAATGCTGGC- <i>Lys-Lys</i> -NH ₂ PNA 4	A/G/C/T = <i>aeg</i> PNA

5	<p>Cf-GalNAc₃-TTTCATAATGCTGGC-Lys-Lys-NH₂</p> <p style="text-align: center;">PNA 5</p>	<p style="text-align: center;">A/G/C/T=ae^g PNA</p> 
6	<p>Cf-[C'(S)GalNAc-T]₃-TTTCATAATGCTGGC-Lys-lys-NH₂</p> <p style="text-align: center;">PNA 6</p>	<p style="text-align: center;">A/G/C/T=ae^g PNA</p> 

2.4.2c Cleavage of the PNA oligomers from the solid support

The oligomers were cleaved from the solid support (L-lysine derivatized Rink amide resin), using 50% TFA in DCM in the presence of triisopropyl silane (TIPS), which yielded PNA oligomers having L-lysine amides at their C-termini. Under this cleavage condition, the nucleobase side chain protecting groups were also removed. After the cleavage reaction, the PNA oligomers in solution were precipitated by addition of cold diethyl ether and the PNA oligomers were dissolved in de-ionized water.

2.4.2d 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 the 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 re injecting the sample on the same C18 analytical 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 such as sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid, SA), Picolinic acid (PA), 2,5-dihydroxybenzoic acid (DHB) and α -cyano-4-hydroxycinnamic acid (CHCA) have been reported to record MALDI-TOF spectrum. Out of these, DHB, CHCA or SA 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 2.2. The MALDI-TOF data for confirmation of mixed purine-pyrimidine PNA oligomers are shown in Appendix II

TABLE 2.2 MALDI-TOF spectral analysis of the synthesized PNA oligomers

Sr. No.	PNA sequence Code	Molecular Formula	Calc Mass	Obs Mass	Rt (min)
1	PNA 1	C ₁₇₄ H ₂₂₈ N ₈₆ O ₅₀	4321.77	4344.26[M+Na] ⁺	14.2
2	GalNAc ₃ -PNA 2	C ₂₄₅ H ₃₅₁ N ₉₅ O ₈₀	5907.08	5907.44[M] ⁺	10.9
3	[C ^γ (S)-GalNAc-T] ₃ PNA 3	C ₂₂₅ H ₃₁₇ N ₉₁ O ₇₂	5448.57	5487.18[M +K] ⁺	11.5
4	Cf-PNA 4	C ₁₉₅ H ₂₃₈ N ₈₆ O ₅₆	4680.80	4681.60[M +H] ⁺	14.9
5	Cf-GalNAc ₃ -PNA 5	C ₂₇₂ H ₃₇₄ N ₉₈ O ₈₆	6392.58	6392.86[M] ⁺	11.2
6	Cf-[C ^γ (S)-GalNAc-T] ₃ PNA 6	C ₂₄₆ H ₃₂₇ N ₉₁ O ₇₈	5806.88	5806.19[M] ⁺	12.4

2.5 Summary

To summarize, this chapter describes the synthesis and characterization of rationally designed triantennry GalNAc moiety and [C^γ(S)-GalNAc-T] monomers (**1** and **2**). The GalNAc₃ moiety and [C^γ(S)-GalNAc-T] monomer were incorporated at the N terminus of PNA. All the intermediates have been characterized by ¹H & ¹³C NMR spectroscopy, mass spectral analysis and other appropriate analytical data. The detailed experimental procedure (Section 2.6) and spectral data of all intermediates discussed in (Section 2.8). The next chapter deals with the investigation of biophysical properties of PNA oligomers.

2.6 Experimental Methods

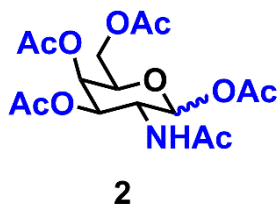
This section describes the detailed synthetic procedures and spectral characterization of the GalNAc₃ moieties.

2.6.1 General procedures

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 GF₂₅₄ 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).

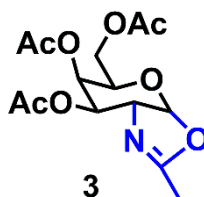
^1H and ^{13}C NMR spectra were recorded using Bruker AC-200 (200 MHz) or JEOL 400 MHz NMR spectrometers. The delta (δ) values for chemical shifts are reported in ppm and are referred to internal standard TMS or deuterated NMR solvents. The optical rotation values were obtained on Rudolph Research Analytical Autopol V polarimeter. Mass spectra for reaction intermediates were obtained by Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using TiO_2 or 2,5-dihydroxybenzoic acid (DHB) and the integrity of PNA oligomer was checked on the same instrument using SA, 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 Tetra(*O*-acetyl)-*N*-(acetyl) galactosamine (2)



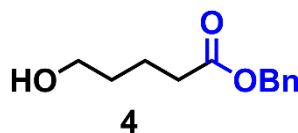
D-(+)-Galactosamine hydrochloride **1** (2.0 g, 10 mmol) was dissolved in anhydrous pyridine (20 mL), and acetic anhydride (10.5 mL, 0.1 mol, 10 eq) was added. The reaction mixture was stirred at room temperature until disappearance of the starting material, and was poured into a beaker with ice-cold water (200 mL). A white solid precipitated was collected by vacuum filtration, washed with ice-cold water and co-evaporated with toluene (3×20 mL) to remove residual water to yield **2** (2.9 g, 80%) as a powdery solid: ^1H NMR (400 MHz, CDCl_3) δ 5.73 (d, $J = 8.8$ Hz, 1 H), 5.46 (d, $J = 9.5$ Hz, 1 H), 5.40 (d, $J = 2.8$ Hz, 1 H), 5.12 (dd, $J = 3.3, 11.3$ Hz, 1 H), 4.52-4.43 (m, 1 H), 4.24-4.11 (m, 2 H), 4.07-4.02 (m, 1 H), 2.20 (s, 3 H), 2.16 (s, 3 H), 2.08 (s, 3 H), 2.05 (s, 3 H), 1.97 (s, 3 H). ^{13}C NMR (101 MHz, CDCl_3) δ 171.7, 170.7, 169.9, 169.0, 168.6, 100.0, 90.7, 70.7, 69.7, 67.5, 61.6, 51.1, 23.0, 20.9 (MALDI-TOF) m/z calcd. for $\text{C}_{16}\text{H}_{23}\text{NO}_{10}$: 389.1322; found: 412.2691 [$\text{M} + \text{Na}$] $^+$.

2.6.3 GalNAc oxazolin (3)



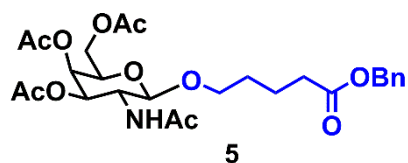
Galactosamine pentaacetate **2** (1 g, 2.5 mmol) was dissolved in DCM (10 mL) at room temperature and TMSOTf (0.7 ml, 3.1 mmol) was added to the reaction. The reaction was stirred at 50 °C for 90 minutes and allowed to come to room temperature and stirred for another 12 hours. The reaction mixture was poured into an ice cold sodium bicarbonate solution, extracted with DCM, washed with water and dried (Na₂SO₄) and filtered. Solvents were removed the residue dried under high vacuum overnight to provide **3** as dark gum (0.8 g, 90%) which was used for the next reaction without any further purification. (MALDI-TOF) *m/z* calcd. for C₁₄H₂₉NO₈; 329.11; found: 352.0817 [M + Na]⁺.

2.6.4 Benzyl 5-hydroxypentanoate (**4**)



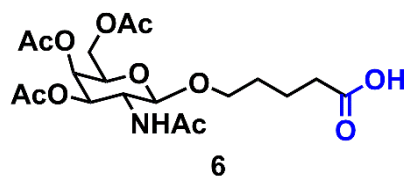
A mixture of δ -valerolactone (9.2 ml, 92.0 mmol) and NaOH (3.6 g, 92 mmol) in 20 mL of distilled water was refluxed for 6 h. After cooling in an ice-water bath, the water was evaporated and the mixture was dissolved in hot EtOH. After filtration, the insoluble residue in hot ethanol was removed a white solid was obtained. Benzyl bromide (3.4 g, 20 mmol) was added dropwise to a solution of sodium-5-hydroxypentanoate (5.0 g, 40 mmol) containing Na₂CO₃ (600 mg) in of distilled water (20 mL). After stirring at room temperature for 12 h the reaction was quenched with water (400 mL) and extracted with EtOAc (2 × 300 mL). The EtOAc layer was dried over Na₂SO₄, filtered, and concentrated under reduced pressure. The residue obtained was then purified on silica gel (60–120 mesh) flash column chromatography using DCM and MeOH (Methanol:DCM 10:90). to give compound **4** as sticky yellowish oil (3.5 g, 85% yield). ¹H NMR (400 MHz, CDCl₃) δ 7.39 – 7.27 (m, 5H), 5.11 (s, 2H), 4.69 (s, 1H), 3.63 (t, *J* = 6.3 Hz, 2H), 3.47 (s, 2H), 2.40 (t, *J* = 7.3 Hz, 2H), 1.78 – 1.69 (m, 2H), 1.58 (dt, *J* = 13.4, 6.7 Hz, 2H). ¹³C NMR (400 MHz, CDCl₃) δ 173.7, 136.1, 128.7, 128.4, 127.7, 127.1, 66.4, 62.3, 34.0, 32.1, 21.2. HRMS (ESI-TOF) *m/z* calcd. for C₁₂H₁₆O₃ 208.1099 found 231.0996 [M + Na]⁺.

2.6.5 6-Benzyloxy [tri(*O*-acetyl)-*N*-(acetyl)galactosamine]-pentanoate (**5**):



Compound **3** (4 g, 12.1 mmol) and benzyl ester **4** (3.5 g, 16.92 mmol) were dissolved in DCM (50 mL) and molecular sieves (4 Å, 5.7 g) were added to the reaction and stirred for 30 minutes. TMSOTf (1.2 mL, 5.4 mmol) was added to the reaction mixture and was stirred at room temperature for 12 hours after which it was poured in to an ice cold solution of Na₂CO₃ and extracted with DCM, washed with water and dried over anhydrous Na₂SO₄ and filtered. Solvents were removed and the residue purified by column chromatography (silica gel, eluting with 20 to 100% ethyl acetate in hexanes) to yield **5** (4.3 g, 65%). ¹H NMR (400 MHz, CDCl₃) δ: 7.66 (d, *J*=9.2 Hz, 1H), 7.32–7.06 (m, 5H), 5.08 (d, *J*=3.3 Hz, 1H), 4.95 (s, 2H), 4.83 (dd, *J*=3.5, 11.3 Hz, 1H), 4.35 (d, *J*=8.4 Hz, 1H), 4.00–3.84 (m, 3H), 3.83–3.66 (m, 1H), 3.64–3.51 (m, 1H), 3.37–3.22 (m, 1H), 2.23 (t, *J*=7.2 Hz, 2H), 2.10 (s, 3H), 1.99 (s, 3H), 1.89 (s, 3H), 1.76 (s, 3H), 1.51–1.27 (m, 4H); ¹³C NMR (400 MHz, CDCl₃) δ 173.7, 170.6, 169.64, 137.0, 128.6, 127.8, 100.7, 99.8, 77.2, 70.1, 66.8, 66.4, 61.6, 51.7, 33.9, 28.6, 23.5, 21.3, 20.8 ppm. HRMS (ESI- TOF) *m/z* calculated for C₂₆H₃₆NO₁₁ 560.21; Found 560.21[M+H]⁺.

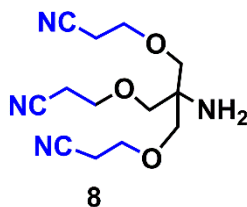
2.6.6 [Tri(*O*-acetyl)-*N*-(acetyl)galactosamine] pentanoic acid (**6**)



A mixture of Pd/C (1.3 g, 10 wt% wet Degussa type) and compound **5** (1 g, 1.8 mmol) were suspended in MeOH (5 mL) and ethyl acetate (60 mL). The reaction vessel was flushed with hydrogen gas and hydrogenated under balloon pressure for 16 hours. The reaction was filtered through celite and the filter bed was washed with MeOH. The combined filtrates were concentrated under reduced pressure to yield **6** (0.7 g, 95%). ¹H NMR (400 Hz, DMSO-*d*₆) δ: 7.81 (d, *J*=9.2 Hz, 1H), 5.22 (d, *J*=3.3 Hz, 1H), 4.97 (dd, *J*=3.3, 11.3 Hz, 1H), 4.49 (d, *J*=8.6 Hz, 1H), 4.14–3.95 (m, 3H), 3.88–3.80 (m, 1H), 3.81–3.63 (m, 1H), 3.54–3.16 (m, 2H), 2.40–2.20 (t, *J*=6.9 Hz, 2H), 2.11 (s, 3H), 2.00 (s, 3H), 1.90 (s, 3H), 1.77 (s, 3H), 1.59–1.44 (m, 4H); ¹³C NMR (400 MHz, CDCl₃) δ 171.3, 170.8, 100.9, 77.1, 70.7, 70.2, 69.1,

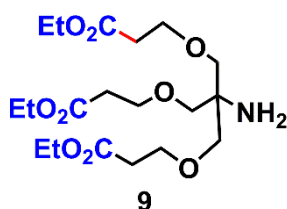
66.9, 61.7, 51.6, 33.7, 28.6, 23.4, 21.5, 20.8.. HRMS (ESI-TOF) m/z calculated for $C_{19}H_{30}NO_{11}$ 448.2; Found 448.2 $[M+H]^+$.

2.6.7 Tris(cyanoethoxymethyl)aminomethane (8)

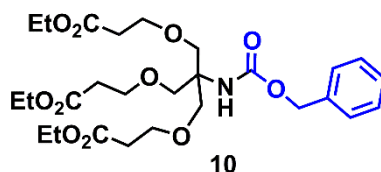


Compound **7** (1.4 g, 5.0 mmol) was dissolved in dry ethanol (25 mL). The mixture was refluxed for 3 h and subsequently left stirring overnight at room temperature. Next, the mixture was filtered, and the filtrate was concentrated under reduced pressure. Water (40 mL) was added, and the mixture was neutralized with aqueous NH_3 (25%). The mixture was extracted with DCM (2 x 50 mL), and the combined organic layers were dried over $MgSO_4$. The crude product was purified over a silica gel column using Methanol/DCM (1/0 f 94/6, v/v) as eluent to Yield **8** (1.2 g, 56%). 1H NMR (400 MHz, $CDCl_3$) δ 4.15 (q, 6 H, $J = 7.1$ Hz), 3.69 (t, 6 H, $J = 6.3$ Hz), 3.32 (s, 6 H), 2.55 (t, 6 H), 1.27 (t, 9 H), ^{13}C NMR (400 MHz, $CDCl_3$) δ 117.6 , 71.9 , 65.1, 55.4, 18.1 ppm. HRMS (ESI-TOF) calculated for $C_{20}H_{13}N_4O_3$ 280.15 Found 281.1614 $[M+H]^+$.

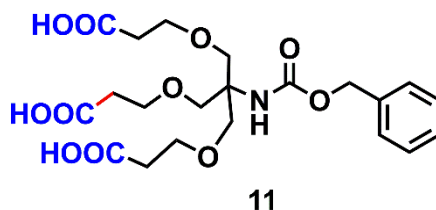
2.6.8 Tris(carboxyethoxyethyl)aminomethane triethyl ester (9)



The trinitrile **8** (1.4 g, 5.0 mmol) was added to a H_2SO_4 (10 mL) in dry ethanol (25 mL). The mixture was refluxed for 24 h. After completion of reaction, the mixture was neutralized by aq. saturated $NaHCO_3$ solution, filtered, and the filtrate was extracted with DCM (2 x 50 mL). The combined organic layers were dried over anhydrous Na_2SO_4 . The crude product obtained after concentration was purified over a silica gel column using DCM/Methanol (94/6, v/v) as eluent. Yield **9** (1.2 g, 56%). 1H NMR ($CDCl_3$): δ 4.15 (q, 6 H, $J = 7.1$ Hz), 3.69 (t, 6 H, $J = 6.3$ Hz), 3.32 (s, 6 H), 2.55 (t, 6 H), 1.27 (t, 9 H). ^{13}C NMR (400 MHz, $CDCl_3$) δ 171.2, 72.5, 66.6, 60.1, 55.7, 34.8, 14.0 ppm. MS (MALDI-TOF) m/z calculated for $C_{19}H_{35}NO_9$ 421.23 found 460.19 $[M+K]^+$.

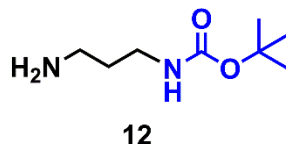
2.6.9 *N*H-Cbz-tris(carboxyethoxymethyl) aminomethane triethyl ester (**10**)

To stirred solution of H-TRIS[OEt]₃ **9** (3.8 g, 9.1 mmol) in dioxane was added solution of Na₂CO₃ (1.0 g, 10.0 mmol) in H₂O (7.8 mL). To this solution was added benzyl chloroformate (1.7 g, 10.0 mmol) and the reaction was allowed to continue for 12 hours. The solvent was removed under reduced pressure and the resulting oil was purified by silica gel column chromatography (20-30% EtOAc/Hexane) to yield **10** (4.5 g, 90%). $R_F = 0.23$ (20% EtOAc/Hexane). ¹H NMR (400 MHz, CDCl₃) δ 7.32 (m, 5H, Phenyl), 5.23 (s, 1H, NH), 5.01 (s, 2H benzyl), 4.1 (q, 6H, CH₂CH₃), 3.65 (m, 12H, CH₂O and OCH₂), 2.5 (t, 6H, CH₂C=O), 1.21 (q, 9H, CH₃). ¹³C NMR (400 MHz, CDCl₃) δ 171.6, 136.9, 128.6, 128.1, 77.2, 69.6, 66.9, 66.3, 60.6, 58.9, 35.2, 14.4 ppm. HRMS (ESI-TOF) m/z calculated for C₂₇H₄₁NO₁₁ 578.26 found 578.25 [M+Na]⁺.

2.6.10 *N*-(Benzyloxycarbonyl)-tris-(carboxyethoxymethyl) aminomethane (**11**)

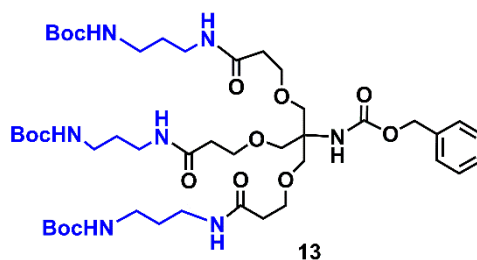
The triester **10** (0.3 g, 0.50 mmol) was dissolved in a mixture of 1,4-dioxane (30 mL) and water (9 mL). To this solution was added aqueous NaOH (4 M, 1 mL). After stirring for 3 h, TLC analysis (DCM/methanol/acetic acid 18/2/1, v/v/v) showed complete conversion of the starting material into one product. The mixture was acidified by the addition of aqueous HCl (36-38%; 1 mL). Water (50 mL) was added, and the resulting mixture was extracted with DCM (3 x 50 mL). The combined organic layers were dried over anhydrous MgSO₄ and concentrated to an oil. The crude **11** thus obtained was immediately used for the synthesis of **13**. MS (MALDI-TOF) m/z calculated for C₂₁H₂₉NO₁₁ 471.17 found 494.16 [M+Na]⁺.

2.6.11 *tert*-butyl (3-aminopropyl)carbamate (**12**)



1,3-Diaminopropane (2.9 g, 40.0 mmol, 4 eq) was dissolved in 100 mL of anhydrous DCM and stirred under an ice bath. A solution of 2.2 g (10.0 mmol, 1 eq) of Boc₂O in DCM (20 mL) was slowly added dropwise into the flask. The adding procedure was completed within 30 min. With the addition of Boc₂O, a white precipitate began to form. After the completion of the addition, the reaction mixture was stirred at room temperature for an additional 16 h. The mixture was then filtered through a Buchner funnel to remove the white precipitate. The collected organic phase was washed three times with saturated saline, and the aqueous phase was then extracted with DCM once more. The combined organic phase was dried against anhydrous Na₂SO₄. The product (sticky transparent liquid) was obtained after vacuum evaporation (1.5 g, 87.0%). Compound **12** [tert-butyl (3-aminopropyl) carbamate]: ¹H NMR (400 MHz, CDCl₃) δ 1.31–1.62 (11H, m, C(CH₃)₃ and C-CH₂-C), 2.73 (2H, t, NH₂CH₂), 3.15 (2H, t, CONH-CH₂), 5.05 (2H, s, NH₂); ¹³C NMR (400 MHz, CDCl₃) δ 156.4, 79.3, 77.2, 39.5, 38.4, 33.0, 28.6 ppm. HRMS(ESI-TOF) *m/z* calculated for C₈H₁₈N₂O₂ 175.2 [M + H]⁺.

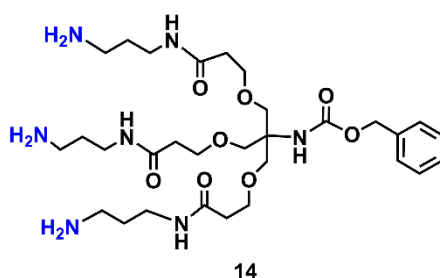
2.6.12 *N*-(Benzyloxycarbonyl)-tris-[tertbutoxy(carboxyethoxymethyl)aminopropyl] aminomethane carbamate (**13**)



Compound **11** (3 g, 6.3 mmol) and mono-Boc protected 1,3- diaminopropane **12** (4.4 g, 25.4 mmol) were dissolved in DMF (5 mL). HBTU (8.4 g, 22.3 mmol) and *N,N*-diisopropylethylamine (6.6 mL, 51.3 mmol) were added and the reaction was stirred overnight. The reaction mixture was poured on to ice-water and extracted with DCM. The organic layer was washed with saturated NaHCO₃ solution, brine, dried with anhydrous Na₂SO₄ and filtered. The organic phase was concentrated under reduced pressure and the residue was purified by column chromatography (silica gel, eluting with EtOAc, followed by 2 to 10 % methanol in DCM to yield **13** (4.1 g, 69%). ¹H NMR (300 MHz, DMSO-*d*₆) δ: 7.82

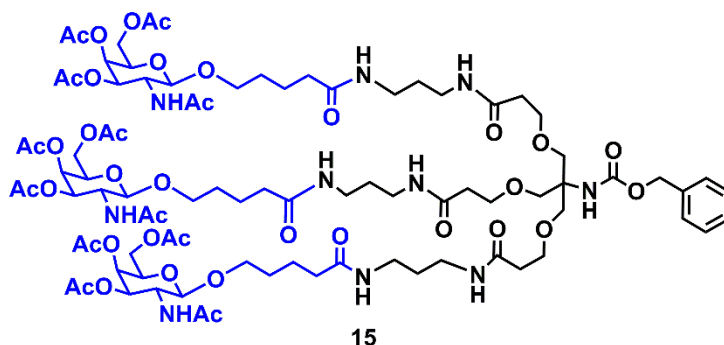
(m, 3H), 7.34 (m, 5H), 6.76 (m, 3H), 6.61–6.44 (m, 1H), 4.97 (s, 2H), 3.54 (t, $J=6.3$ Hz, 6H), 3.47 (br.s., 6H), 3.08–2.95 (m, 6H), 2.96–2.80 (m, 6H), 2.26 (t, $J=6.3$ Hz, 6H), 1.48 (t, $J=6.9$ Hz, 6H), 1.42–1.28 (m, 27H); ^{13}C NMR (101 MHz, CDCl_3) δ 171.9, 156.5, 155.3, 136.4, 128.5, 79.2, 77.1, 69.6, 67.4, 66.9, 66.3, 58.7, 37.1, 36.8, 36.1, 30.0, 28.4 ppm. HRMS (ESI-TOF) m/z calculated for $\text{C}_{45}\text{H}_{76}\text{N}_7\text{O}_{14}$ 939.5529; Found 940.5567 $[\text{M}+\text{H}]^+$.

2.6.13 *N*-(Benzyloxycarbonyl)-tris-[(carboxyethoxymethyl)aminopropyl]aminomethane carbamate (14):



Compound **13** (3.6 g, 3.8 mmol) was dissolved in DCM (75 mL) and trifluoroacetic acid (15 mL) and the reaction was stirred at room temperature for 30 minutes. The reaction mixture was diluted with toluene (100 mL) and concentrated under reduced pressure. The residue was co-evaporated with toluene (2 x 100 mL) and dried under reduced pressure using a high vacuum pump to yield **14** as the TFA salt (2.4 g, 95%) which was used for the next reaction without any further purification. MS [MALDI-TOF]: m/z calculated for $\text{C}_{30}\text{H}_{54}\text{N}_7\text{O}_8$. 939.3956; Found 940.4013 $[\text{M}+\text{H}]^+$.

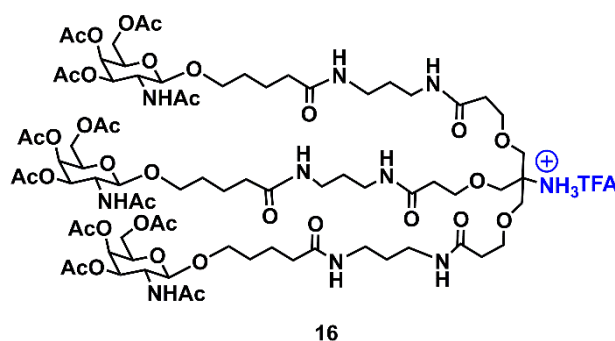
2.6.14 *N*-(Benzyloxycarbonyl)-tris-[(carboxyethoxymethyl)aminopropyl-tri-(*O*-acetyl) *N*-acetyl galactosamine]aminomethane carbamate (15)



Compound **6** (0.15 g, 0.23 mmol) was dissolved in DMF (0.5 mL). To this HBTU (0.13 g, 0.34 mmol), HOBt (52 mg, 0.34 mmol) and *N,N*-diisopropylethylamine (0.15 mL, 1.0 mmol) were added and stirred for 5 min. A solution of compound **14** (60 mg, 9.5 mmol) in DMF (0.5 mL) was added and the reaction was stirred for 18 h at room temperature.

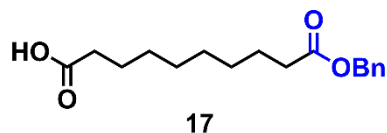
Solvent was removed under reduced pressure and the residue was dissolved in DCM (10 mL), washed with saturated aqueous sodium bicarbonate solution (2 mL), water (2 mL) and brine (2 mL). The organic layer was dried over anhydrous Na₂SO₄ and filtered. The solvent was removed and under reduced pressure and the residue was purified by column chromatography over silica gel, eluting first with EtOAc, followed by gradient elution 5 to 25% methanol in DCM to yield **15** (90 mg, 45%). ¹H NMR (400 MHz, DMSO-d₆) δ 7.96–7.78 (m, 6H), 7.74 (t, *J*=5.5 Hz, 3H), 7.46–7.08 (m, 5H), 5.20 (d, *J*=3.3 Hz, 3H), 5.03–4.83 (m, 5H), 4.47 (d, *J*=8.4 Hz, 3H), 4.01 (s, 9H), 3.93–3.78 (m, 3H), 3.76–3.62 (m, 3H), 3.59–3.35 (m, 15H), 3.02 (m, 12H), 2.26 (t, *J*=6.3 Hz, 6H), 2.09 (s, 9H), 2.07–2.00 (m, 6H), 1.99 (s, 9H), 1.92–1.83 (m, 9H), 1.76 (s, 9H), 1.60–1.34 (m, 18H); ¹³C NMR (400 MHz, DMSO-d₆) δ 171.9, 170.0, 169.9, 137.2, 128.3, 127.7, 101.0, 79.2, 70.5, 69.8, 68.7, 68.2, 67.3, 66.7, 61.4, 58.8, 49.4, 36.4, 35.0, 29.3, 28.6, 22.8, 21.8, 20.5 ppm. HRMS [ESI-TOF] *m/z* calculated for C₈₇H₁₃₃N₁₀O₃₈ 1926.8860; Found 1950.22 [M+Na]⁺

2.6.15 *NH*^{TFA}-tris[(carboxyethoxymethyl)aminopropyl-tri-(O-acetyl) N-acetyl galactosamine]aminomethane carbamate (**16**)



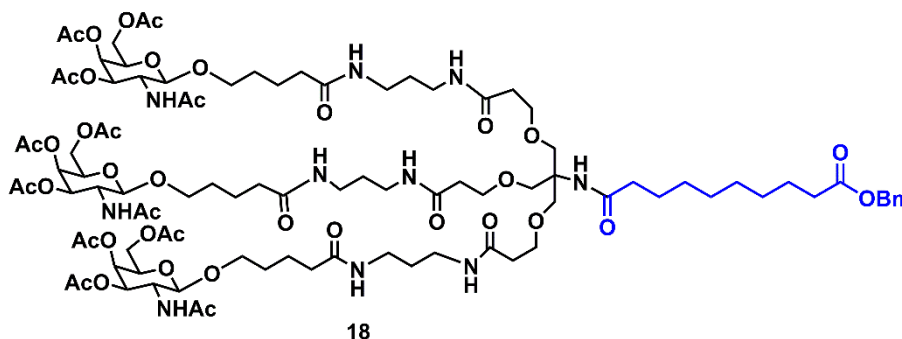
Compound **15** (1.5 g, 2.1 mmol) was dissolved in methanol (30 mL) and Pd/C (0.45 g, 10 wt% Degussa type wet) was added. The reaction vessel was flushed with hydrogen gas and the reaction mixture was hydrogenated (balloon pressure) for 12 hours. The reaction mixture was filtered through a small pad of celite and washed with methanol. Trifluoroacetic acid (0.25 mL, 3.1 mmol) was added and the solvent was removed under reduced pressure. The residue was co-evaporated with toluene (2 times) and dried under high vacuum for 16 hours to yield the TFA salt of **16** (3.9 g, 95%) as a pale yellow solid. MS (MALDI-TOF) *m/z* calculated for C₇₉H₁₂₉N₁₀O₃₆ 1793.8565 found 1794.9445 [M+H]⁺.

2.6.16 10-(Benzyloxy)-10-oxodecanoic acid (**17**)



Decanoic acid (2 g, 9.8 mmol) was dissolved in 1:4 dioxane (10 ml) and water (10 ml) to which solid NaHCO_3 (0.85 g, 9.9 mmol) was added. After five minutes benzyl bromide (2.6 ml, 8.1 mmol) was added dropwise at 0 °C. After completion of addition of benzyl bromide, the reaction was stirred at room temperature overnight. After completion of reaction, dioxane was removed on rotary evaporator, water added and extracted with DCM (20 ml). The organic layer was washed with aq. NaHCO_3 (5 ml) and brine (5 ml). The crude product was subjected to flash column chromatography (Hexane/EtOAc 90:10 v/v) to obtain **17**, (1.4 g, 45%). ^1H NMR (400MHz, CDCl_3) δ (ppm): 10.88 (s, 1H, COOH), 7.34 (m, 5H, Ar-H), 5.11 (s, 2H, $\text{CH}_2\text{-Ar}$), 2.34 (m, 4H, COCH_2), 1.61 (m, 4H, COCH_2CH_2), 1.29 (s, 8H, $(\text{CH}_2)_4$). ^{13}C NMR (101, CDCl_3) δ ppm: 179.8, 173.7, 136.1, 128.5, 66.1, 34.3, 29.0, 8.9, 24.9 ppm. HRMS (ESI-TOF) m/z calcd. For $\text{C}_{17}\text{H}_{24}\text{O}_4$ 315.1572; found 315.1572 $[\text{M} + \text{Na}]^+$.

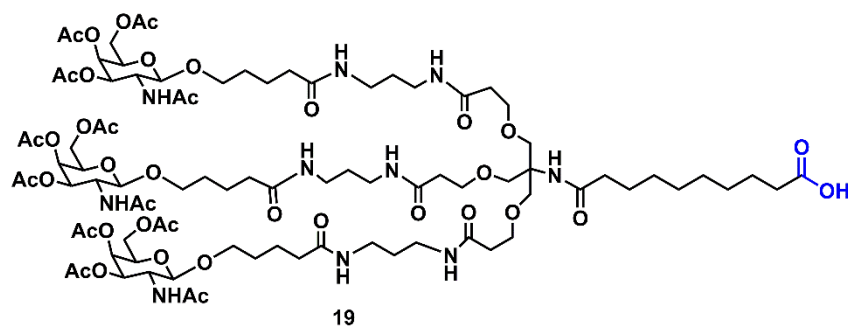
2.6.11 *N*-(Benzyloxy decanoate)-tris-[(carboxyethoxymethyl)aminopropyl- tri-(*O*-acetyl) *N*-acetyl galactosamine]aminomethane carbamate (**18**)



To a solution of decanoic acid monobenzyl ester **17** (0.2 g, 0.6 mmol) in DMF (2 mL) were added HBTU (34 mg, 0.9 mmol) and DIEA (0.6 mL, 4.1 mmol), and the resulting mixture was stirred for few minutes. A solution of compound **16** (1 g, 0.56 mmol) in DMF (2 mL) was added, and stirring was continued at room temperature overnight. Solvents and volatiles were removed under reduced pressure, and the residue was dissolved in DCM (20 mL), washed with saturated NaHCO_3 (5 mL) and water (5 mL). After drying over anhydrous Na_2SO_4 the solvent was evaporated under reduced pressure. The residue was purified by silica gel chromatography (eluent: 3-15% MeOH in DCM) to obtain compound **18** as off-white solid (0.810 g, 74%). ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ 7.85–7.76 (m, 6H); 7.71 (t, $J = 5.7$ Hz, 3H); 7.39–7.29 (m, 5H); 6.97 (s, 1H); 5.20 (d, $J = 3.4$ Hz, 3H); 5.07 (s, 2H); 4.94 (dd,

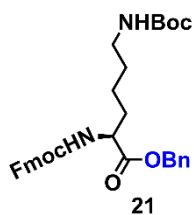
$J = 3.4, 11.1$ Hz, 3H); 4.47 (d, $J = 8.5$ Hz, 3H); 4.06–3.97 (m, 9H); 3.86 (dt, $J = 8.9, 11.2$ Hz, 3H); 3.69 (dt, $J = 5.9, 9.9$ Hz, 3H); 3.61–3.44 (m, 12H); 3.40 (dt, $J = 6.3, 9.9$ Hz, 3H); 3.08–2.96 (m, 12H); 2.32 (t, $J = 7.4$ Hz, 2H); 2.25 (t, $J = 6.3$ Hz, 6H); 2.09 (s, 9H), 2.02 (t, $J = 7.0$ Hz, 8H); 1.99 (s, 9H); 1.88 (s, 9H), 1.77 (s, 9H); 1.56–1.35 (m, 22H); 1.28–1.14 (m, 12H). ^{13}C NMR (100 MHz, DMSO- d_6): δ 172.8, 171.9, 170.1, 169.9, 136.3, 128.4, 128.0, 127.9, 101.0, 70.5, 69.8, 68.7, 67.3, 66.7, 65.3, 61.4, 59.5, 49.4, 36.3, 35.9, 35.1, 33.5, 29.5, 28.9, 25.3, 24.5, 22.8, 21.9, 20.5 ppm. (MALDI-TOF) m/z calcd. for $\text{C}_{97}\text{H}_{154}\text{N}_{10}\text{O}_{39}$: 2067.29; found: 2090.14 $[\text{M} + \text{Na}]^+$.

2.6.12 *N*-(decanoic acid)- tris-[(carboxyethoxymethyl)aminopropyl- tri-(*O*-acetyl) *N*-acetyl galactosamine]aminomethane carbamate (19)



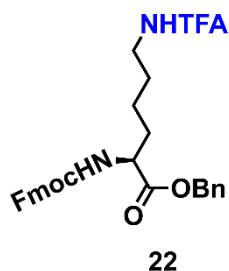
Compound **18** (1 g, 0.48 mmol) was dissolved in methanol, evaporated under reduced pressure to remove traces of chlorinated solvent from previous step, re-dissolved in MeOH. 10 wt% Pd-C (1.0 g, wet Degussa type) was added, and the reaction was hydrogenated under normal pressure overnight. The reaction mixture was filtered through celite and washed with MeOH; combined filtrate was evaporated under reduced pressure to afford **19** as a white solid (0.9 g, 95%). ^1H NMR (400 MHz, DMSO- d_6): δ 11.92 (s, 1H); 7.92–7.80 (m, 6H); 7.73 (t, $J = 5.6$ Hz, 3H); 6.96 (s, 1H); 5.21 (d, $J = 3.4$, 3H); 4.96 (dd, $J = 3.4, 11.2$ Hz, 3H); 4.47 (d, $J = 8.4$, 3H); 4.06–3.96 (m, 9H); 3.85 (dt, $J = 8.8, 11.0$ Hz, 3H); 3.69 (dt, $J = 6.0, 9.9$ Hz, 3H); 3.56–3.46 (m, 12H); 3.41 (dt, $J = 6.3, 9.9$ Hz, 3H); 3.08–2.98 (m, 12H); 2.28 (t, $J = 6.4$ Hz, 6H); 2.18 (t, $J = 7.4$ Hz, 2H); 2.10 (s, 9H); 2.03 (t, $J = 7.2$ Hz, 8H); 1.99 (s, 9H); 1.88 (s, 9H); 1.77 (s, 9H); 1.55–1.36 (m, 22H); 1.26–1.16 (m, 12H). ^{13}C NMR (101 MHz, DMSO- d_6): δ 174.8, 172.6, 172.0, 170.0, 169.6, 101.1, 70.5, 69.9, 68.7, 68.3, 67.4, 66.6, 61.5, 59.5, 49.4, 36.3, 35.9, 35.2, 34.0, 29.3, 28.5, 25.3, 24.7, 22.8, 21.9, ppm. (MALDI- TOF) m/z calcd. for $\text{C}_{89}\text{H}_{144}\text{N}_{10}\text{O}_{39}$: 1976.95; found: 2000.70 $[\text{M} + \text{Na}]^+$.

2.6.13 *N*-(Boc)- *N*-[(Fmoc)methoxy]carbonyl]-benzyl lysinate (21)



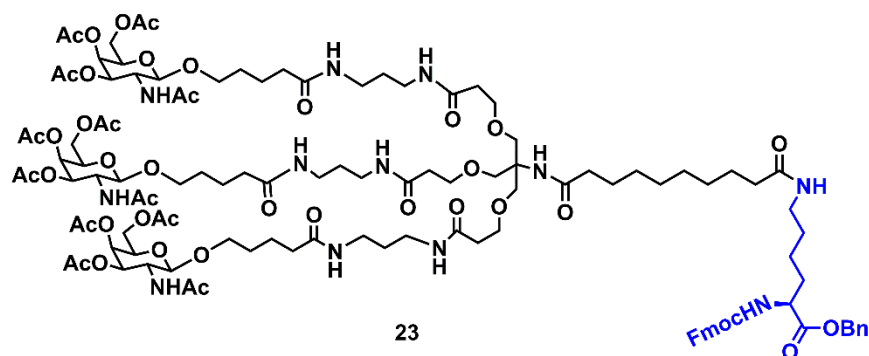
To a stirred solution of NHFmoc Lysine Boc OH **20** (5 g, 6.6 mmol), DIEA (1.2 g, 9.6 mmol), and DMAP (80 mg, 0.66 mmol) in CH₂Cl₂ (20 mL), was added dropwise at 0 °C a solution of Cbz-Cl (1.3 g, 1.00 mL, 7.7 mmol) in DCM (25 mL). The mixture was stirred at 0 °C for 3 h, then washed with aqueous KHSO₄ (1 N) solution (3 × 100 mL), dried and concentrated under reduced pressure. The residue was taken up in EtOAc (3 mL) and purified by column chromatography [hexane / EtOAc 2 : 1] to give the benzyl ester of Fmoc Lysine **21** (5.5, 93%). ¹H NMR (400 MHz, CDCl₃) δ 7.76 (d, *J* = 7.5 Hz, 2H), 7.60 (d, *J* = 7.4 Hz, 2H), 7.43 – 7.28 (m, 9H), 5.43 (d, *J* = 7.3 Hz, 1H), 5.24 – 5.12 (m, 2H), 4.56 – 4.31 (m, 4H), 4.21 (t, *J* = 7.0 Hz, 1H), 3.07 (d, *J* = 5.8 Hz, 2H), 1.91 – 1.81 (m, 1H), 1.69 (s, 2H), 1.43 (s, 11H), 1.26 (s, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 172.1, 155.8, 143.7, 143.5, 141.1, 135.1, 128.3, 127.5, 126.9, 125.0, 119.6, 79.01, 77.1, 66.8, 53.7, 46.9, 39.8, 32.1, 29.4, 27.9, 22.5, 22.0, 13.9 ppm. HRMS (ESI-MS) *m/z* calcd. for C₃₃H₃₈N₂O₆, 558.2730; found: 559.2806 [M+H]⁺.

2.6.14 *N*-(TFA)-*N*-[(Fmoc)methoxy]carbonyl]-benzyl lysinate (**22**):



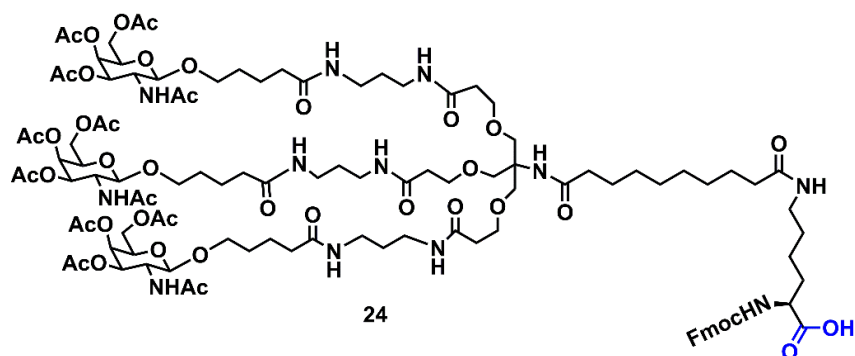
Fmoc Lysine benzyl ester **21** (3.6 g, 3.30 mmol) was dissolved in DCM (75 mL) containing trifluoroacetic acid (15 mL) and the reactants were stirred at room temperature for 30 minutes. The reaction mixture was diluted with toluene (100 mL) and concentrated under reduced pressure. The residue was co-evaporated with toluene (2 × 100 mL) and dried under reduced pressure using a high vacuum pump to obtain compound **22** as the TFA salt (2.4 g, 95%) which was used for the next reaction without any further purification. HRMS (ESI-MS) *m/z* calcd. for C₂₈H₂₉N₂O₄: 458.2127; found: 459.2281 [M+H]⁺.

2.6.15 *N*-(Decanoic)-tris-[(carboxyethoxymethyl)aminopropyl]-tri(O-acetyl) N(acetyl galactosamine)aminomethane [(Fmoc)methoxy]carbonyl]-benzyl lysinate (**23**)



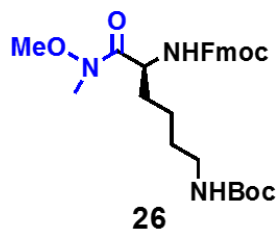
To a solution of **19** (0.5 g, 0.25 mmol) in DMF (2 mL) were added HBTU (0.3 g, 0.76 mmol) and DIEA (0.14 mL, 0.92 mmol), and the resulting mixture was stirred for few minutes. A solution of compound **22** (0.42 g, 0.76 mmol) in DMF (2 mL) was added, and stirring was continued at room temperature overnight. Solvents and volatiles were removed under reduced pressure, and the residue was dissolved in DCM (20 mL), washed with saturated NaHCO₃ (5 mL) and water (5 mL). After drying over anhydrous Na₂SO₄ the solvent was evaporated under reduced pressure. The residue was purified by silica gel chromatography (eluent: 3-15% MeOH in DCM) to obtain compound **23** (0.6 g, 50%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.91 – 7.69 (m, 14H), 7.44 – 7.28 (m, 8H), 6.99 (s, 1H), 5.75 (s, 1H), 5.21 (d, *J* = 3.4 Hz, 2H), 5.12 (s, 1H), 4.97 (dd, *J* = 11.2, 3.4 Hz, 2H), 4.49 (d, *J* = 8.5 Hz, 2H), 4.32 – 4.18 (m, 2H), 4.07 – 3.96 (m, 9H), 3.88 (dd, *J* = 18.9, 10.0 Hz, 3H), 3.70 (dt, *J* = 10.8, 5.6 Hz, 3H), 3.55 (m, *J* = 12.0, 5.7 Hz, 12H), 3.44 – 3.36 (m, 4H), 3.07 – 2.96 (m, 13H), 2.28 (t, *J* = 6.3 Hz, 6H), 2.10 (s, 8H), 2.04 (dd, *J* = 13.1, 6.2 Hz, 9H), 1.99 (d, *J* = 2.5 Hz, 9H), 1.89 (s, 8H), 1.77 (s, 8H), 1.55 – 1.40 (m, 22H), 1.33 (s, 1H), 1.24 – 1.16 (m, 9H), 1.10 (s, 2H). ¹³C NMR (150 MHz, DMSO-*d*₆) δ 172.5, 172.0, 170.1, 169.6, 156.2, 143.7, 140.7, 136.0, 128.4, 128.0, 127.6, 127.0, 125.2, 120.1, 115.6, 101.0, 70.5, 69.8, 68.3, 67.3, 66.7, 65.7, 64.9, 61.4, 59.5, 54.9, 54.0, 49.4, 46.6, 39.5, 36.0, 35.1, 29.4, 28.6, 21.8, 20.5 ppm. MS (MALDI-TOF) *m/z* calcd. for C₁₁₇H₁₇₂N₁₂O₄₂ [M+H]⁺, 2418.70 ; found: 2441.92 [M+Na]⁺.

2.6.16 *N*-(decanoic) -tris-[(carboxyethoxymethyl)aminopropyl]-tri-(*O*-acetyl) *N*-acetyl galactosamine]aminomethane [(Fmoc)methoxy)carbonyl]-Lysine (24**):**



Compound **23** (0.5 g, 0.22 mmol) was dissolved in methanol, evaporated under reduced pressure to remove traces of chlorinated solvent from previous step, re-dissolved in MeOH, 10 wt% Pd-C (1.0 g, wet Degussa type) was added, and the reaction was hydrogenated under normal pressure overnight. The reaction mixture was filtered through celite and washed with MeOH; combined filtrates were evaporated under reduced pressure to afford **24** (0.16g, 40%) as a white solid. ^1H NMR (400 MHz, DMSO- d_6) δ 7.90 (t, $J = 7.9$ Hz, 8H), 7.81 (t, $J = 5.5$ Hz, 3H), 7.77 – 7.67 (m, 3H), 7.41 (t, $J = 7.4$ Hz, 2H), 7.32 (t, $J = 7.4$ Hz, 2H), 7.00 (d, $J = 6.0$ Hz, 1H), 5.21 (d, $J = 3.4$ Hz, 3H), 5.00 – 4.94 (m, 3H), 4.50 (d, $J = 8.5$ Hz, 3H), 4.28 – 4.19 (m, 3H), 4.05 – 3.96 (m, 9H), 3.87 (dt, $J = 11.0, 8.9$ Hz, 3H), 3.70 (dt, $J = 10.8, 5.6$ Hz, 3H), 3.53 (dd, $J = 13.6, 7.1$ Hz, 14H), 3.43 – 3.37 (m, 5H), 3.07 – 2.96 (m, 14H), 2.28 (t, $J = 6.3$ Hz, 6H), 2.10 (s, 9H), 2.03 (dd, $J = 13.1, 6.2$ Hz, 10H), 1.99 (s, 10H), 1.88 (s, 9H), 1.77 (s, 9H), 1.49 (dq, $J = 19.4, 6.5$ Hz, 24H), 1.33 (s, 2H), 1.25 – 1.18 (m, 18H). ^{13}C NMR (600 MHz, DMSO- d_6) δ 174.0, 172.5, 172.0, 171.9, 170.0, 169.6, 167.0, 155.6, 153.4, 144.0, 143.9, 142.6, 140.7, 140.2, 139.4, 138.9, 137.4, 134.1, 131.7, 128.9, 127.6, 127.3, 127.1, 125.2, 123.2, 122.7, 121.4, 120.0, 115.6, 114.6, 109.8, 101.0, 70.5, 69.8, 68.6, 68.2, 67.3, 66.7, 65.4, 61.4, 59.5, 39.5, 36.0, 35.0, 31.5, 31.3, 29.0, 28.6, 25.30, 22.7, 22.1, 21.8, 20.5, 13.9, 10.8 ppm. MS (MALDI-TOF) m/z calcd. for $\text{C}_{110}\text{H}_{166}\text{N}_{12}\text{O}_{42}$ $[\text{M}+\text{H}]^+$, 2328.58 ; found: 2351.57 $[\text{M}+\text{Na}]^+$.

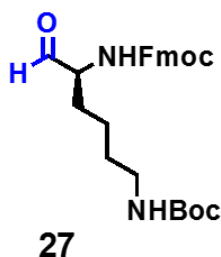
2.6.17 *N*-(Boc)-*N*-[(Fmoc)methoxycarbonyl]-*N,N* methoxy methyl lysinate (**26**)



To a stirring solution of Fmoc-protected Boc Lysine **25** in (5 g, 10.7 mmol) CH_2Cl_2 (50 ml) was added EDC (1.8 g, 11.8 mmol) and HOBt (2 g, 13 mmol). After 10 min stirring

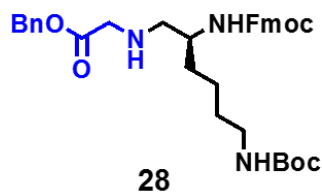
at 23 °C, N, O-Dimethylhydroxylamine hydrochloride (1.2 g, 12 mmol) and DIEA (3 ml, 24 mmol) was added and the reaction mixture was stirred 12 h. The reaction mixture was diluted with EtOAc and washed 0.1 N HCl, 10% K₂CO₃, brine, dried over MgSO₄ and concentrated. The residue obtained was purified on silica gel (100–200 mesh) using Hexane and EtOAc to give compound **26** as a white powder (4.5 g, 80%). ¹H NMR (400 MHz, CDCl₃) δ 7.75 (d, *J* = 8.1 Hz, 2H), 7.60 (t, *J* = 6.6 Hz, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.33 – 7.28 (m, 2H), 5.61 (d, *J* = 8.9 Hz, 1H), 4.68 (d, *J* = 49.4 Hz, 2H), 4.36 (d, *J* = 7.2 Hz, 2H), 4.21 (t, *J* = 7.0 Hz, 2H), 3.76 (s, 3H), 3.21 (s, 3H), 3.11 (d, *J* = 4.5 Hz, 2H), 1.92 (s, 1H), 1.76 (td, *J* = 13.4, 7.6 Hz, 1H), 1.61 (dd, *J* = 13.9, 5.5 Hz, 2H), 1.51 (dd, *J* = 14.7, 7.5 Hz, 2H), 1.42 (s, 9H), 1.28 – 1.18 (m, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 172.5, 156.0, 155.8, 143.7, 143.5, 141.0, 127.4, 126.8, 124.9, 119.5, 78.9, 77.1, 67.0, 61.4, 50.5, 46.9, 40.0, 32.2, 31.8, 29.3, 28.2, 22.2, 20.8 ppm. HRMS (ESI-TOF) *m/z* calcd. For C₂₈H₃₇N₃O₆: 511.2682 and found 534.2571[M+ Na]⁺.

2.6.18 *N*-(*tert*-butoxycarbonyl)-*N*-[(Fmoc)methoxy]carbonyl]-*N,N* methoxy methyl lysinaldehyde (**27**)



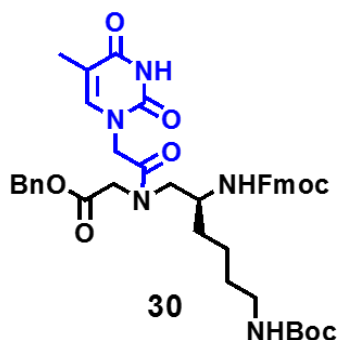
Compound **26** (3 g, 6 mmol) was dissolved in THF (15ml), cooled to -20°C, and LiAlH₄ (0.2 gm, 5.2 mmol) was added slowly. After 1 h, the reaction was quenched with EtOAc and the organic layer washed with 0.1N HCl, dried over MgSO₄ and concentrated. The residue obtained was purified on silica gel (100–200 mesh) using petroleum ether and EtOAc to give compound **27** as a pale yellow oil (2.6 g, 95 %). ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.39 – 7.27 (m, 5H), 7.24 (m, *J* = 5.2 Hz, 1H), 7.19 (m, *J* = 5.9 Hz, 1H), 6.70 (m, *J* = 8.7 Hz, 1H), 5.00 (d, *J* = 5.2 Hz, 2H), 4.75 (d, *J* = 16.8 Hz, 2H), 4.4 (d, *J* = 16.7 Hz, 2H), 4.20 (dd, *J* = 33.7, 18.8 Hz, 1H), 3.95 (dd, *J* = 44.8, 17.4 Hz, 1H), 3.81 – 3.70 (m, 1H), 3.49 – 3.18 (m, 4H), 3.05 (dd, *J* = 19.4, 7.5 Hz, 2H), 1.75 (d, *J* = 1.8 Hz, 2H), 1.64–1.47 (m, 2H), 1.36 (s, 9H). ¹³C NMR (101 MHz, CDCl₃) δ 199.5, 156.3, 143.9, 141.5, 127.9, 127.2, 125.2, 120.1, 79.4, 77.1, 67.1, 60.2, 47.3, 39.9, 29.9, 28.5, 22.2 ppm. HRMS (ESI-TOF) *m/z* calcd. For C₂₈H₃₇N₃O₆ 452.2311 and found 452.5510[M+H]⁺.

2.6.19 *N*-Ethylamino-NH(Fmoc)-C γ (S)-[(tert-butoxycarbonyl)aminobutyl] benzyl glycinate (**28**)



To a solution of the aldehyde **27** in ethanol, benzyl glycine ester hydrochloride (3 g, 66 mmol) was added and then neutralized with DIPEA (0.48 mL, 38 mmol) at ice cold condition and stirred for 5-10 min, and then sodium cyanoborohydride (1.7 g, 26 mmol) was added in small lots at the same temperature and stirred for 4 h. After completion of reaction (monitored by TLC), solvent ethanol was evaporated and the product was extracted with EtOAc. The organic layer was concentrated and subjected for column chromatography, and the product was eluted with 35% EtOAc: Pet-ether to obtain compound **28**. Yield (1.5 g, 37%). ¹H NMR (400 MHz, CDCl₃) δ 7.75 (d, J = 7.5 Hz, 2H), 7.61 (d, J = 7.1 Hz, 6H), 7.41 – 7.27 (m, 2H), 5.15 (t, J = 6.8 Hz, 2H), 4.60 (s, 1H), 4.41 (d, J = 6.4 Hz, 1H), 4.21 (t, J = 6.7 Hz, 2H), 3.69 (s, 1H), 3.45 (q, J = 17.5 Hz, 2H), 3.10 (d, J = 4.9 Hz, 1H), 2.67 (dd, J = 12.5, 4.5 Hz, 1H), 1.54 – 1.29 (m, 13H), 1.28 – 1.18 (m, 2H). ¹³C NMR (101 MHz, CDCl₃) δ 173.6, 156.5, 144.2, 141.5, 128.9, 127.7, 127.1, 125.2, 121.1, 120.1, 108.1, 77.2, 66.8, 53.5, 53.0, 52.0, 50.9, 49.7, 34.2, 28.6, 23.0 ppm. HRMS (ESI-TOF) m/z calcd. For C₃₅H₄₃N₃O₆ 601.3152 and found 601.7440 [M+H]⁺.

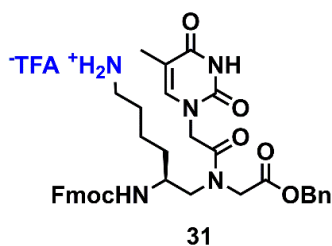
2.6.20 *N*-Ethylamino-NH(Fmoc)-*N*-(acetamido-N1-thyminy)-C γ (S)-[(tert-butoxycarbonyl)aminobutyl] benzyl glycinate (**30**)



To thymine acetic acid **29** (1 g, 55 mmol) in anhydrous DMF (4 mL), secondary amine **28** (5 g, 81 mmol) was added and cooled to 0 °C. After 5 min stirring, HBTU (3 g, 81.5 mmol) and DIPEA (1.5 mL, 12 mmol) were added to the reaction mixture and stirred for 12 h at RT. After completion of reaction, reaction mixture was poured into water (25 ml),

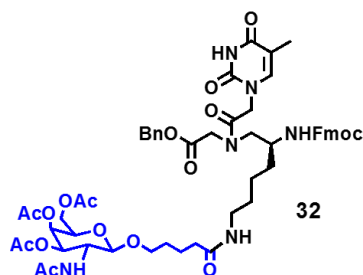
neutralized with *sat.aq.* KHSO₄ (10 mL), and extracted with ethyl acetate (3 x 20 mL). The organic layer was dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure and product was purified by column chromatography, eluting with petroleum ether/EtOAc (1:1) to obtain the compound **30**. (Yield 2.8 g, 74%). ¹H NMR (400 MHz, CDCl₃) δ 9.06 (d, *J* = 51.6 Hz, 1H), 7.73 – 7.25 (m, 13H), 6.78 (t, *J* = 54.4 Hz, 1H), 5.52 (d, *J* = 8.0 Hz, 1H), 5.21 – 4.97 (m, 2H), 4.74 – 4.45 (m, 2H), 4.41 – 3.96 (m, 4H), 3.86 – 3.66 (m, 2H), 3.43 (dd, *J* = 25.6, 5.1 Hz, 2H), 3.14 – 2.75 (m, 3H), 1.80 (t, *J* = 7.4 Hz, 3H), 1.42 (d, *J* = 18.8 Hz, 12H), 1.24 (s, 2H). ¹³C NMR (151 MHz, CDCl₃) δ 169.3, 168.4, 168.0, 164.4, 156.5, 152.1, 144.0, 143.8, 141.5, 141.1, 135.2, 129.0, 128.9, 128.4, 127.8, 127.2, 125.3, 120.1, 110.9, 79.3, 77.2, 68.0, 67.4, 66.8, 50.5, 50.0, 49.4, 48.2, 47.2, 40.1, 39.9, 31.9, 31.5, 29.7, 28.5, 23.0, 22.9, 12.3. HRMS (ESI-TOF) *m/z* calcd. For C₄₂H₄₉N₅O₉ 767.3530 and found 768.3619 [M+H]⁺.

2.6.21 *N*-Ethylamino-NH(Fmoc)-*N*-(acetamido-N1-thyminy)-C γ (*S*)-[(TFA)aminobutyl]benzyl glycinate (**31**)



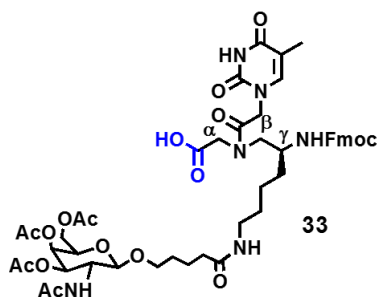
Compound **30** (2 g, 2.8 mmol) was dissolved in DCM (75 mL) and trifluoroacetic acid (15 mL) and the reaction was stirred at room temperature for 30 minutes. The reaction mixture was diluted with toluene (100 mL) and concentrated under reduced pressure. The residue was co-evaporated with toluene (2 x 100 mL) and dried under reduced pressure using a high vacuum pump to yield **31** as the TFA salt (1.7 g) which was used for the next reaction without any further purification. HRMS [ESI-TOF] *m/z* calculated for C₃₇H₄₁N₅O₇. 667.3006; Found 690.1689 [M+Na]⁺.

2.6.22 *N*-Ethylamino-NH(Fmoc)-*N*-(acetamido-N1-thyminy)-C γ (*S*)-[tri (*O*-acetyl) *N*-acetyl galactosamine]aminobutyl]benzyl glycinate (**32**)



To a solution of tri-*O*-acetyl-*N*(acetyl)-galactosamine pentaanoic acid **6** (1.6 g, 3.5 mmol) in DMF (2 mL) were added HBTU (2 g, 5 mmol) and DIEA (1 mL, 7.3 mmol), and the resulting mixture was stirred for few minutes. A solution of compound **31** (3.5 g, 5.1 mmol) in DMF (2 mL) was added, and stirring was continued at room temperature overnight. Solvents and volatiles were removed under reduced pressure, and the residue was dissolved in DCM (20 mL), washed with saturated NaHCO₃ (5 mL) and water (5 mL). After drying over anhydrous Na₂SO₄ the solvent was evaporated under reduced pressure. The residue was purified by silica gel chromatography (eluent: 3-15% MeOH in DCM) to obtain compound **32** (1.6 g, 63 %). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.29 (s, 1H), 7.89 – 7.84 (m, 2H), 7.81 (d, *J* = 9.2 Hz, 1H), 7.73 – 7.64 (m, 3H), 7.42 – 7.24 (m, 10H), 7.18 (s, 1H), 7.11 (d, *J* = 18.3 Hz, 1H), 5.19 (d, *J* = 9.2 Hz, 2H), 5.10 (s, 1H), 4.96 (d, *J* = 14.6 Hz, 1H), 4.69 (dd, *J* = 44.0, 16.7 Hz, 1H), 4.47 (d, *J* = 8.5 Hz, 2H), 4.39 – 4.17 (m, 4H), 4.15 – 3.97 (m, 4H), 3.87 (dd, *J* = 19.7, 9.1 Hz, 1H), 3.69 (dd, *J* = 8.2, 6.9 Hz, 2H), 3.45 – 3.35 (m, 2H), 2.99 (dd, *J* = 13.4, 6.9 Hz, 2H), 2.11 – 1.85 (m, 12H), 1.73 (dd, *J* = 26.6, 6.5 Hz, 6H), 1.51 – 1.21 (m, 10H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 171.8, 170.0, 169.3, 168.7, 167.8, 164.3, 156.1, 150.9, 143.8, 141.8, 140.7, 135.5, 128.4, 127.8, 127.0, 125.1, 120.0, 108.1, 101.0, 70.5, 69.8, 68.6, 66.7, 65.9, 65.3, 61.4, 54.9, 49.8, 49.3, 48.3, 47.7, 46.8, 39.5, 38.4, 35.0, 31.2, 29.5, 28.6, 23.0, 22.7, 21.8, 20.5, 13.6, 11.90. HRMS (ESI-TOF) *m/z* calcd. For C₅₆H₆₈N₆O₁₇; 1097.1814 and found 1097.4727 [M]⁺.

2.6.23 *N*-Ethylamino-NH(Fmoc)-*N*-(acetamido-*N*1-thyminy)-C^γ(*S*)-[tri(*O*-acetyl)-*N*-acetyl galactosamine)aminobutyl] glycine (**33**)



Compound **31** (0.6 g, 0.6 mmol) was dissolved in MeOH, evaporated under reduced pressure to remove traces of chlorinated solvent from previous step, re-dissolved in MeOH. Pd-C (1.0 g, 10 wt% wet, Degussa type) was added at 0° C, and the reaction was hydrogenated under normal pressure by maintaining temp 20-25 °C for 1 h. After completion of reaction (monitored by TLC). The reaction mixture was filtered through celite and washed with MeOH, the combined filtrate was evaporated under reduced pressure to afford **33** as a white solid (0.5 g, 96 %); ¹H NMR (400 MHz, DMSO-d₆) δ 11.26 (s, 1H), 8.71 (s, 2H), 7.85 (d, *J* = 7.6 Hz, 2H), 7.74 (t, *J* = 5.2 Hz, 1H), 7.58 (d, *J* = 7.2 Hz, 1H), 7.40 – 7.28 (m, 3H), 5.21 (d, 1H), 4.97 (d, *J* = 11.2 Hz, 1H), 4.49 (d, *J* = 8.4 Hz, 1H), 4.43 (d, *J* = 8.6 Hz, 2H), 4.02 (s, 3H), 3.97 – 3.82 (m, 25H), 3.72 (dd, *J* = 21.6, 11.6 Hz, 2H), 3.54 (d, *J* = 13.2 Hz, 1H), 3.38 (dd, *J* = 16.4, 10.1 Hz, 5H), 3.26 (s, 2H), 3.20 – 3.12 (m, 1H), 3.01 (d, *J* = 5.5 Hz, 2H), 2.10 (s, 3H), 2.01 (m, *J* = 15.7 Hz, 6H), 1.89 (s, 4H), 1.75 (d, *J* = 15.1 Hz, 6H), 1.46 (m, *J* = 7.4 Hz, 8H), 1.23 (m, 2H). ¹³C NMR (101 MHz, DMSO-d₆) δ 171.9, 170.0, 169.7, 168.5, 168.0, 167.3, 164.4, 151.0, 148.6, 142.2, 139.9, 127.0, 124.2, 120.0, 108.0, 101.0, 70.5, 69.8, 68.7, 66.7, 61.4, 57.5, 51.3, 50.8, 50.0, 49.4, 48.5, 47.9, 46.6, 41.9, 39.5, 38.2, 35.0, 31.5, 30.4, 28.6, 23.1, 22.8, 22.1, 21.9, 20.6, 19.2, 18.1, 13.5, 11.9. HRMS (ESI-TOF) *m/z* calcd. For C₄₉H₆₂N₆O₁₇; 1006. 4171 and found 1029.4069 [M+Na]⁺.

2.6.24 Synthesis of GalNAc₃ conjugated PNA on solid support

The PNA oligomers were synthesized by solid phase coupling reaction by using Fmoc strategy on rink amide resin. The resin-bound 15 mer PNA oligomers (20 mg, 0.22 mmol/g) in DMF (1:1) were coupled with corresponding GalNAc moieties (30 mg, 3eq) in the presence of Pybop (6 mg, 3 eq), HOBT (2 mg, 3eq) and DIPEA (2.5 μL, 3eq). The reaction was done for 10 min in microwave at 75 °C and 40 W and then 24 h at room temperature. Excess reagents were removed by filtration and the resin was washed with DMF, DCM, MeOH.

The rink amide resins (10 mg) after assembly of PNA oligomers were stirred with 90% TFA in DCM (200 μL) and triisopropyl silane (10 μL) in an ice bath for 10 min. 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 and the filtrate was evaporated on a rotary evaporator at ambient temperature. The filtrate was transferred to microfuge tube and the peptide was precipitated with cold dry ether. Precipitate were dissolved in Milli Q water (1 mL) and purified by RP-HPLC.

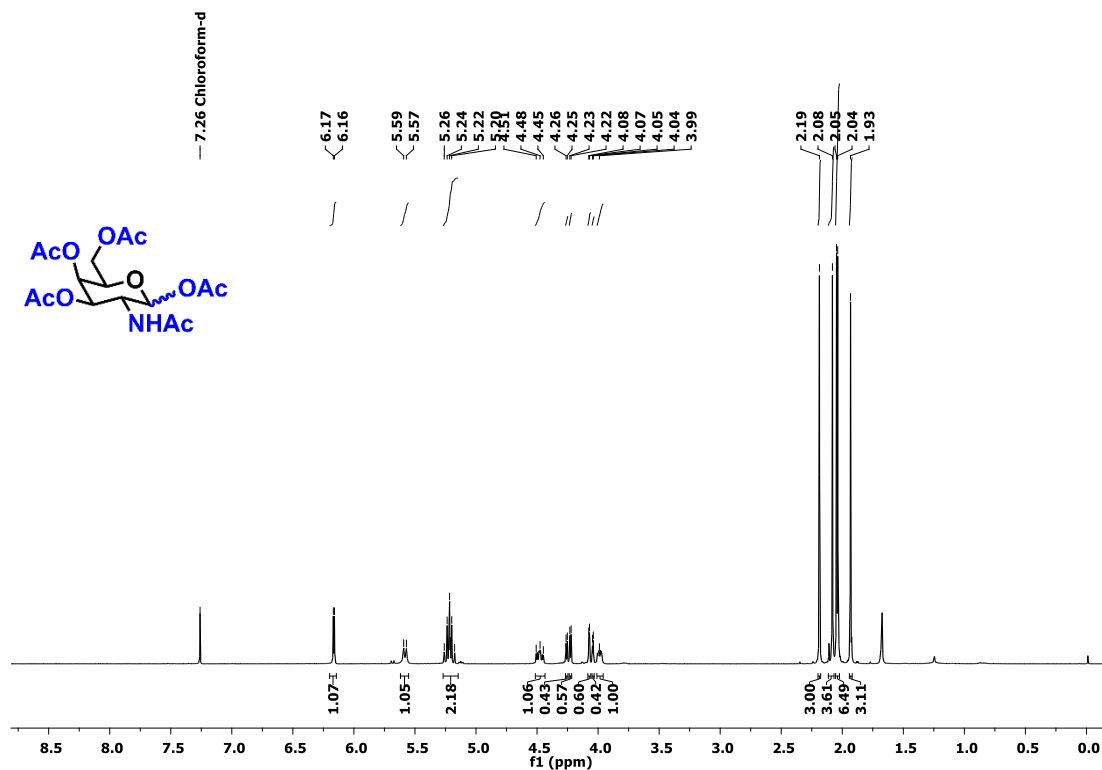
2.6.25 Purification of the GalNAc PNA oligomers by RP-HPLC

The purification of PNAs was carried out on Agilent 1200 HPLC system with semi-preparative phenomenex C18 (10×250 mm) column using solvents water and acetonitrile with composition A: 0.1% TFA in CH₃CN:H₂O (5:95) and B= 0.1% TFA in CH₃CN:H₂O (95:5). The gradient for elution was 100% A to 100% B in 20 min, with flow rate of 2 mL/min. The HPLC elutions were monitored at 220 and 260 nm wavelength.

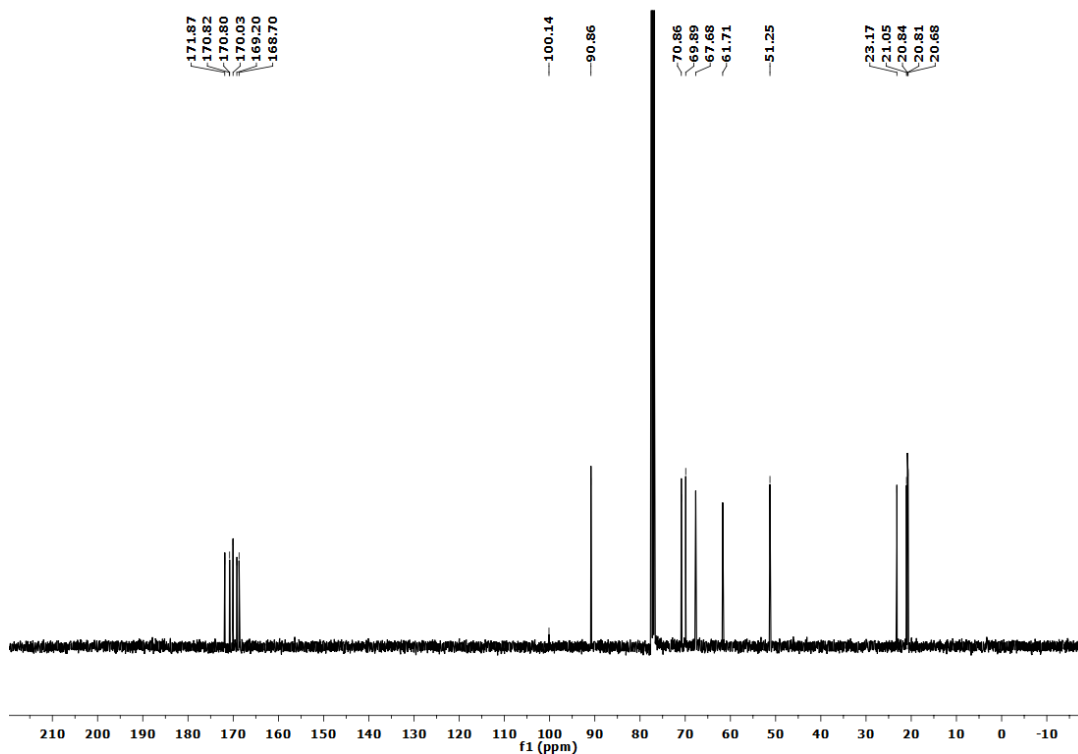
2.7 Characterization data of synthesized compounds and PNA

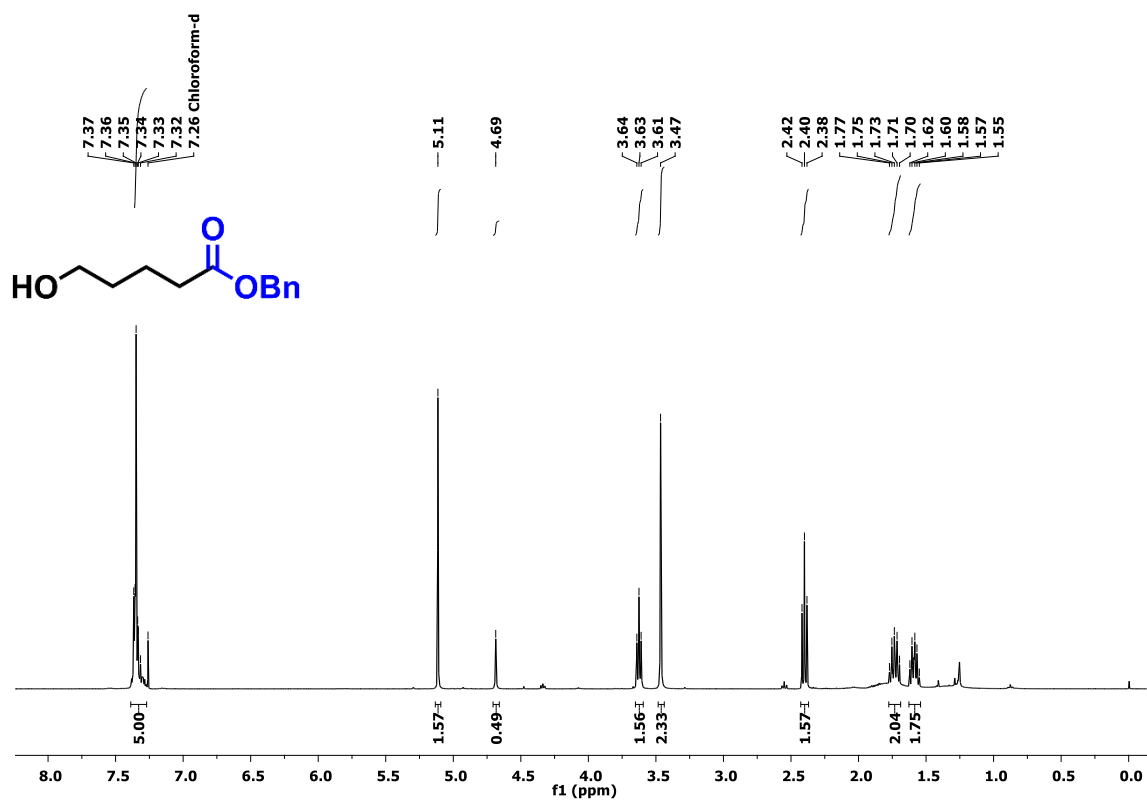
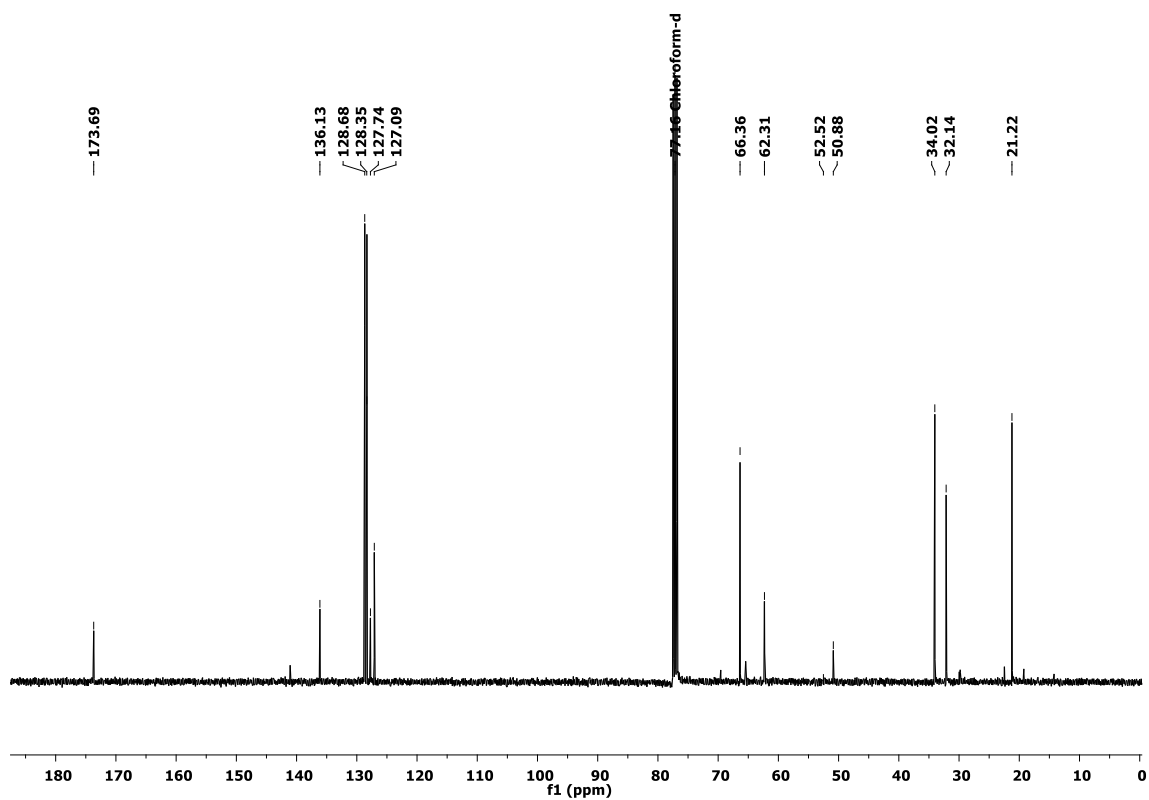
^1H and ^{13}C NMR of Compounds.

^1H NMR of Compound 2

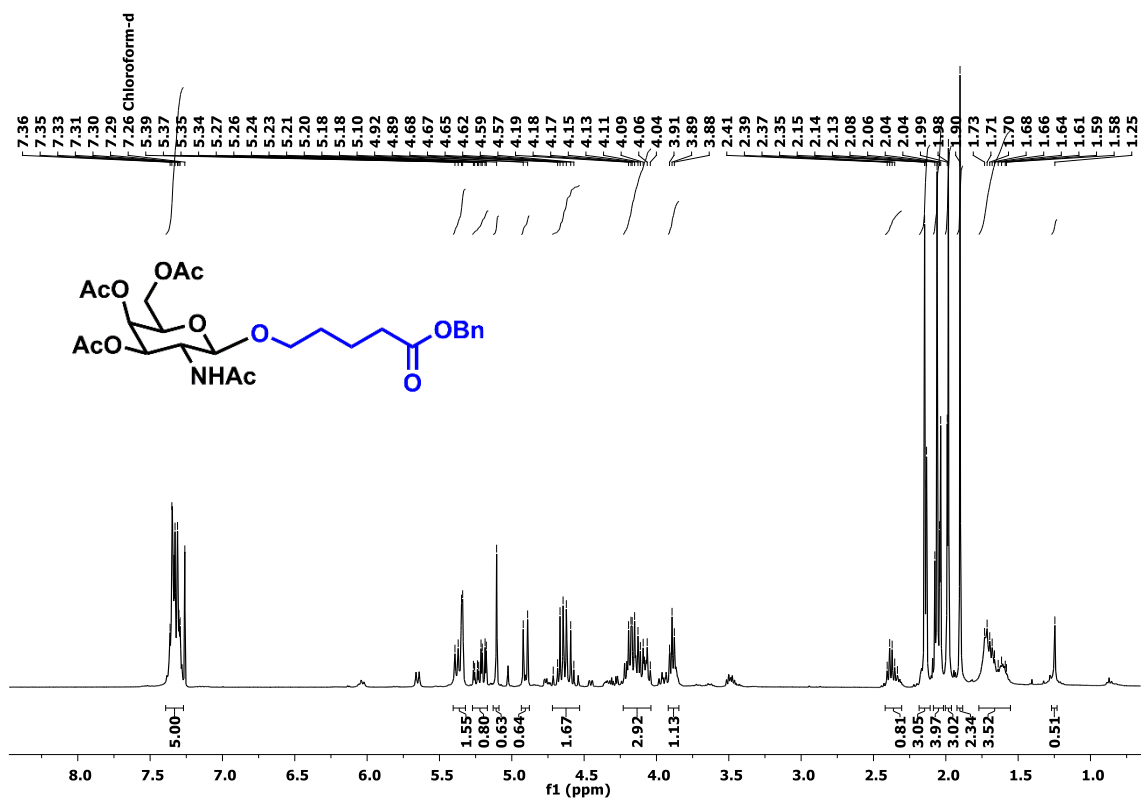


^{13}C NMR of Compound 2

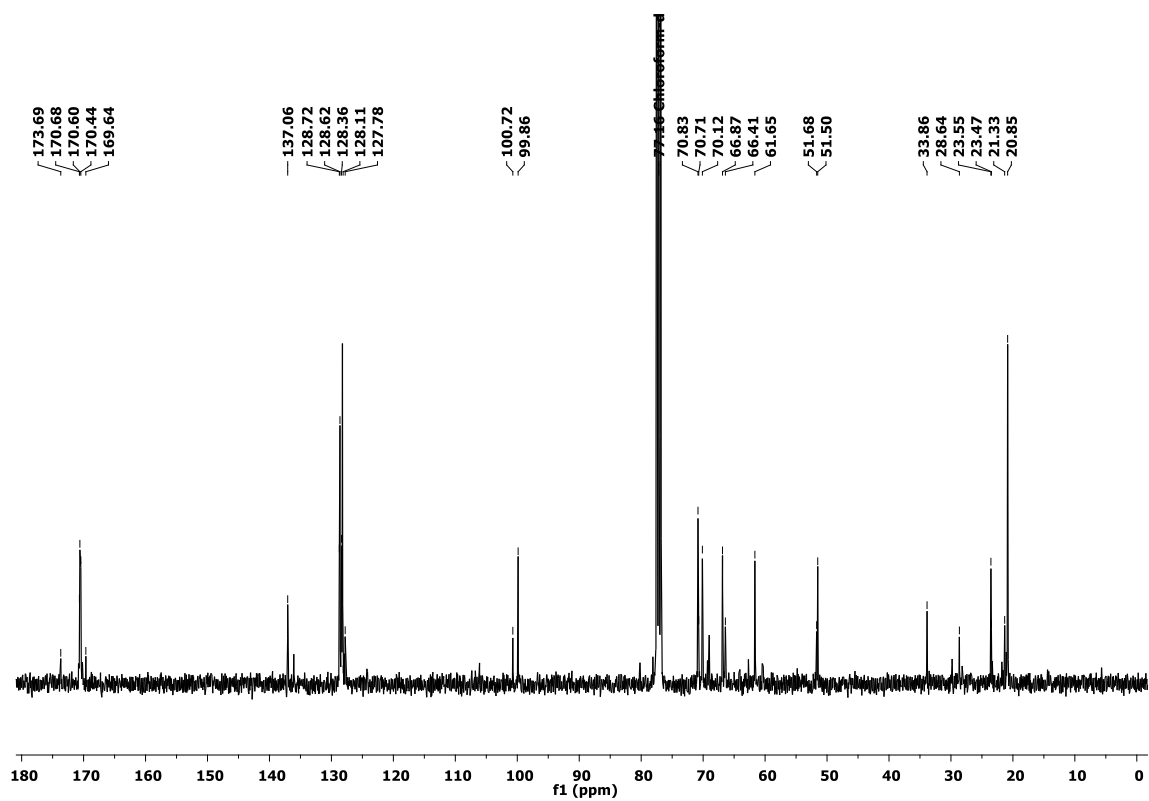


^1H NMR of Compound 4 ^{13}C NMR of Compound 4

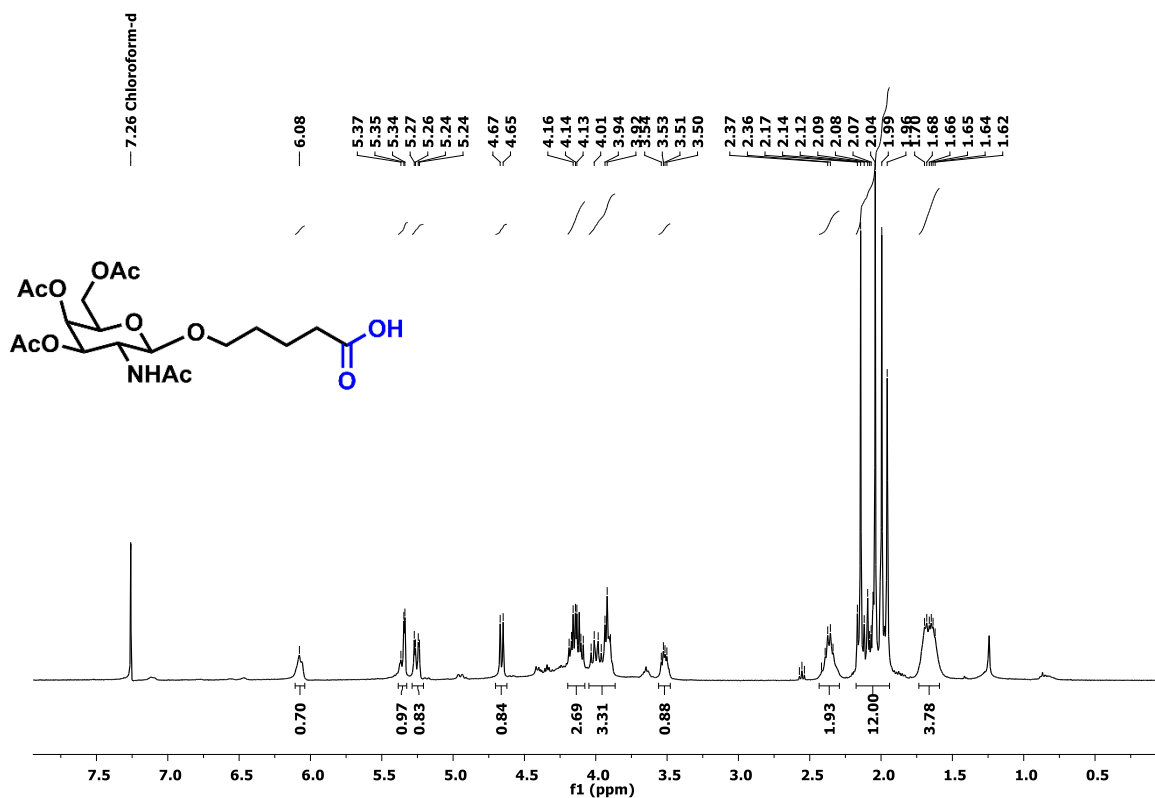
¹³C NMR of Compound 5



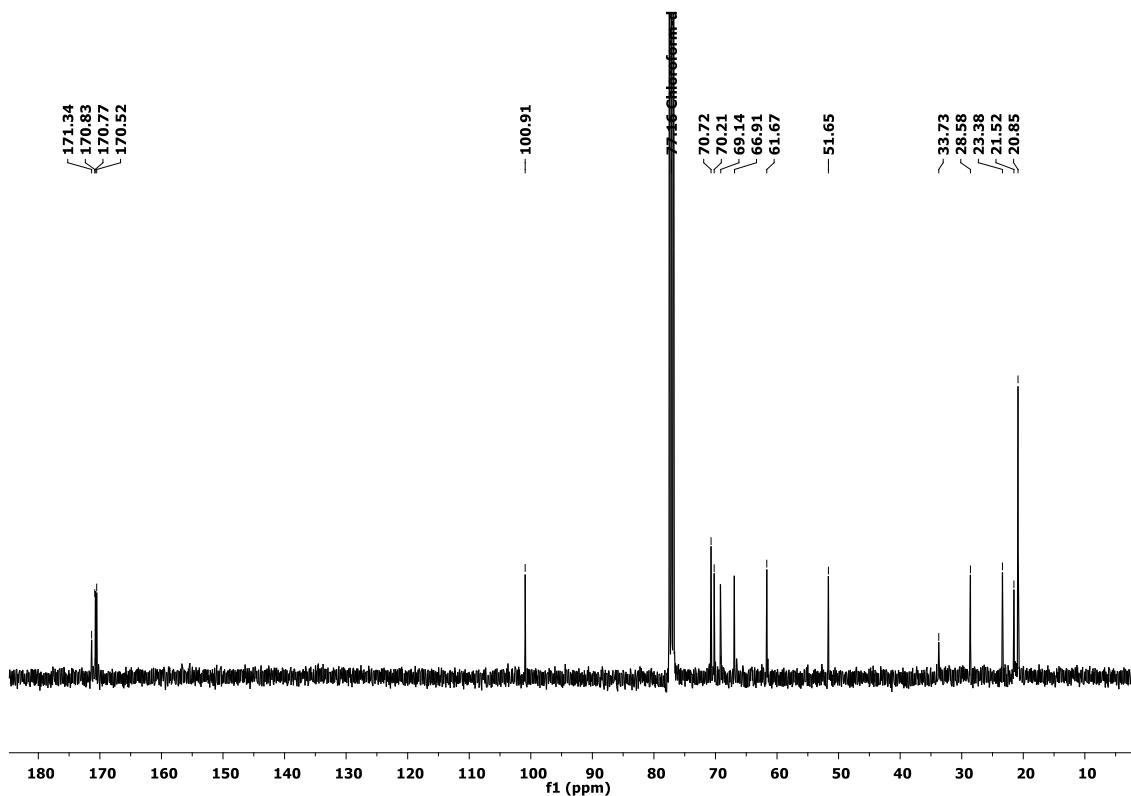
¹³C NMR of Compound 5

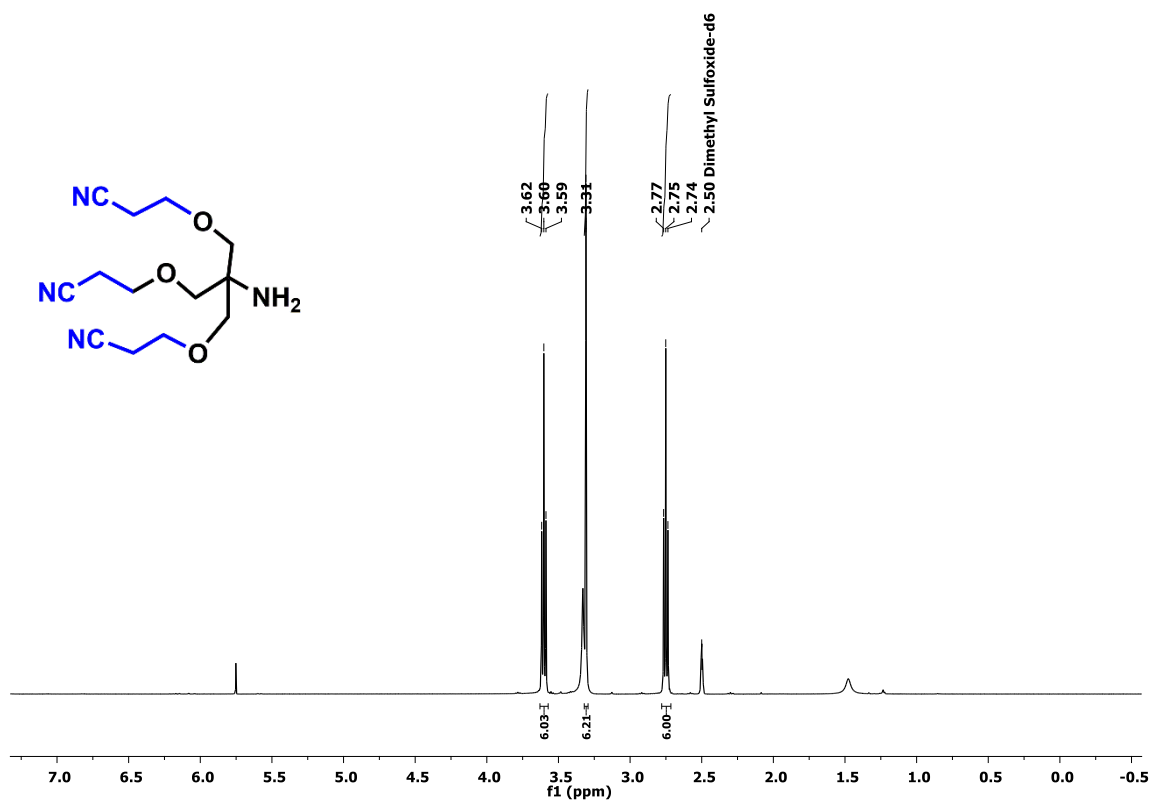
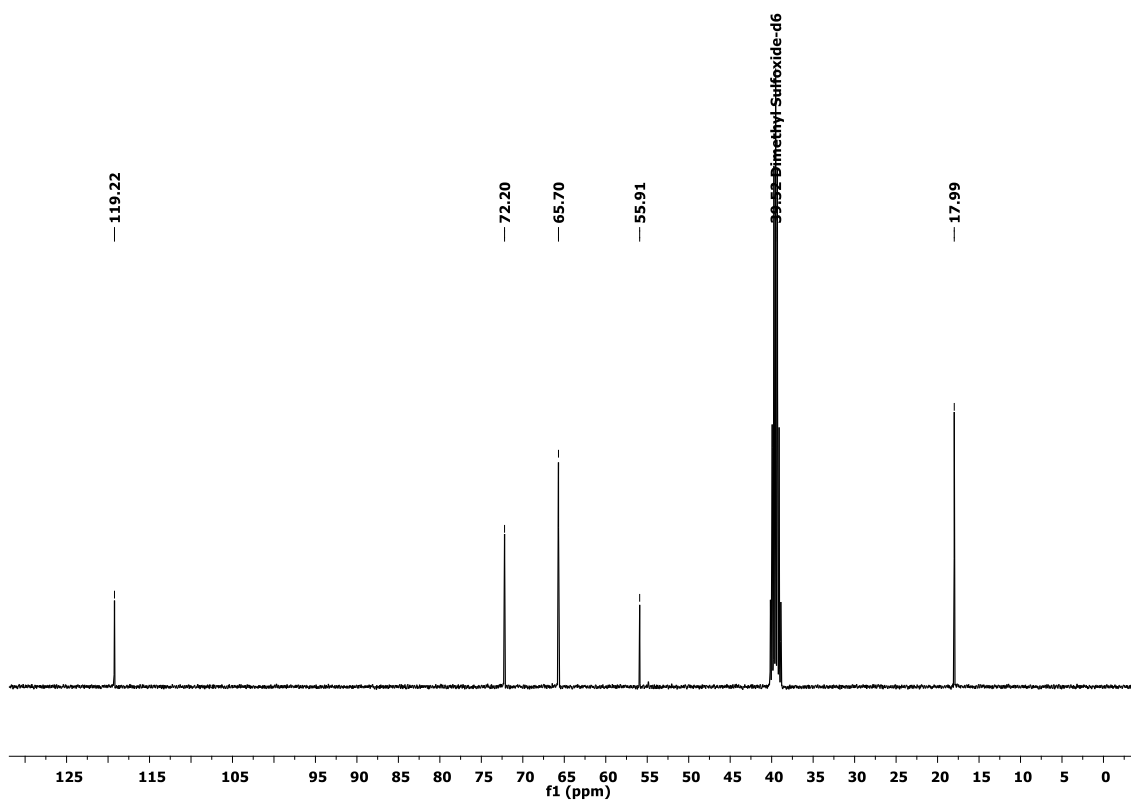


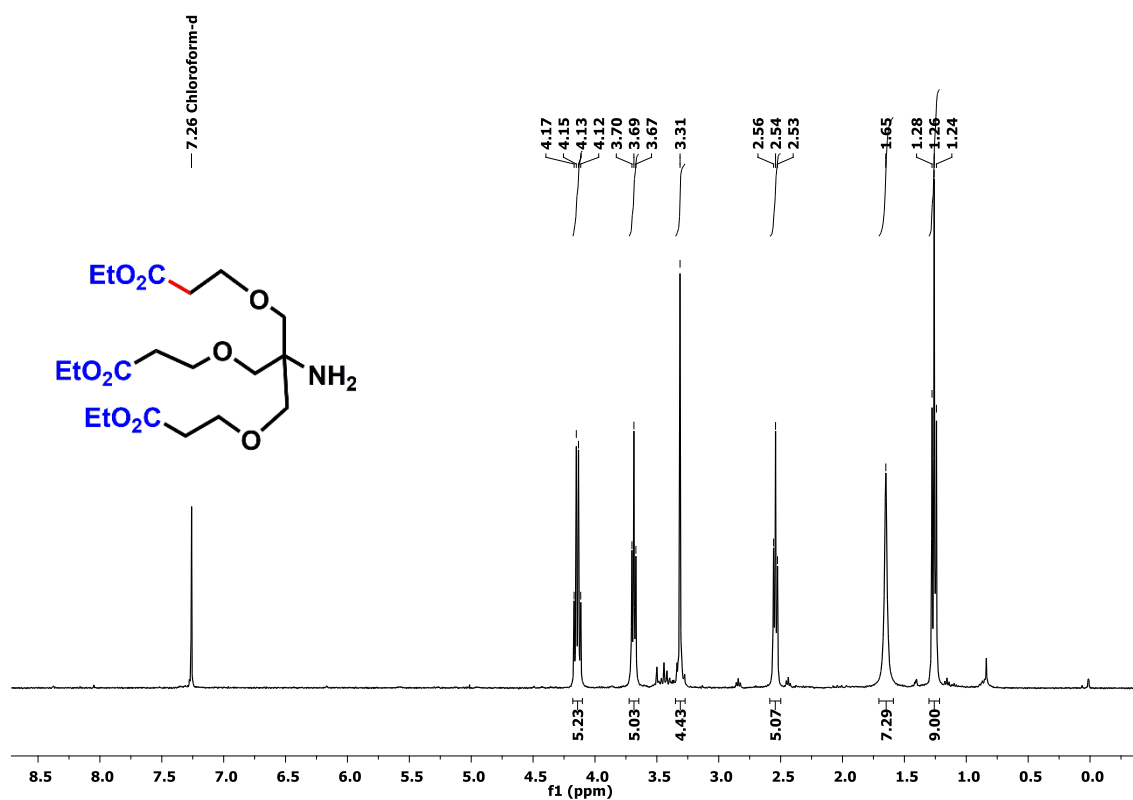
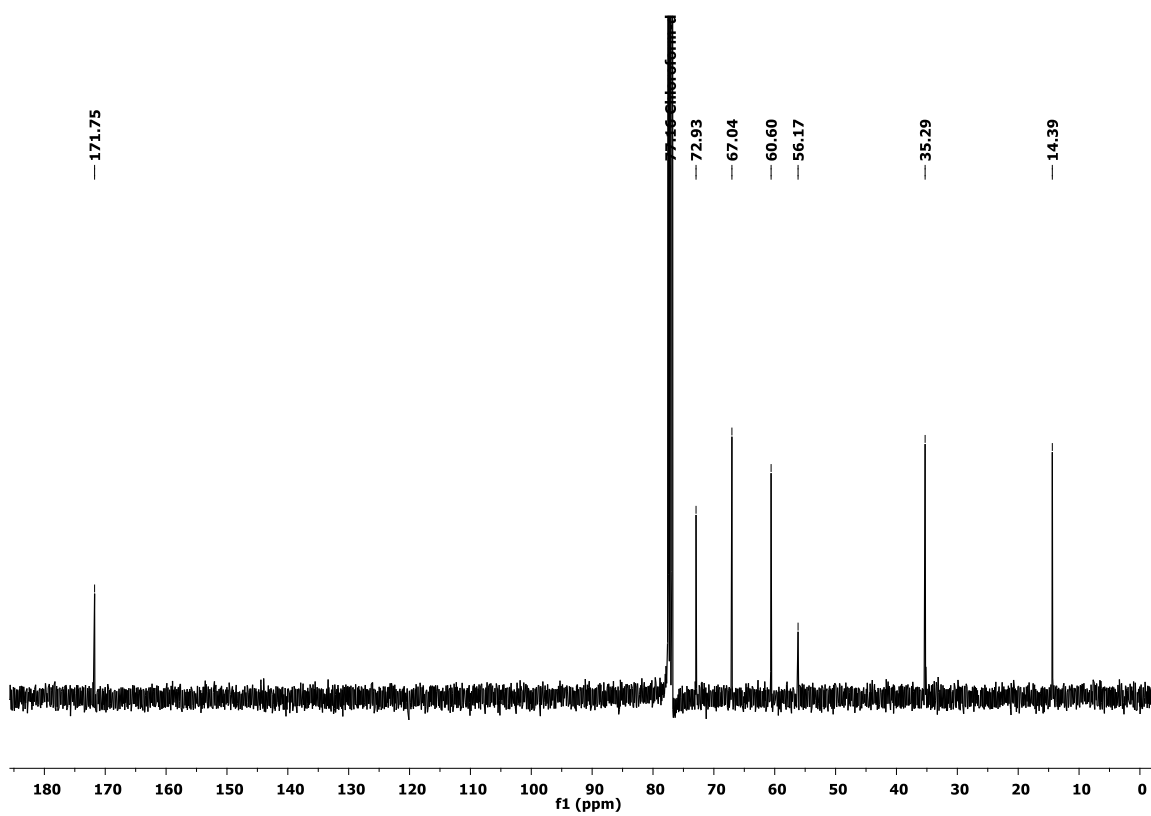
¹H NMR of Compound 6

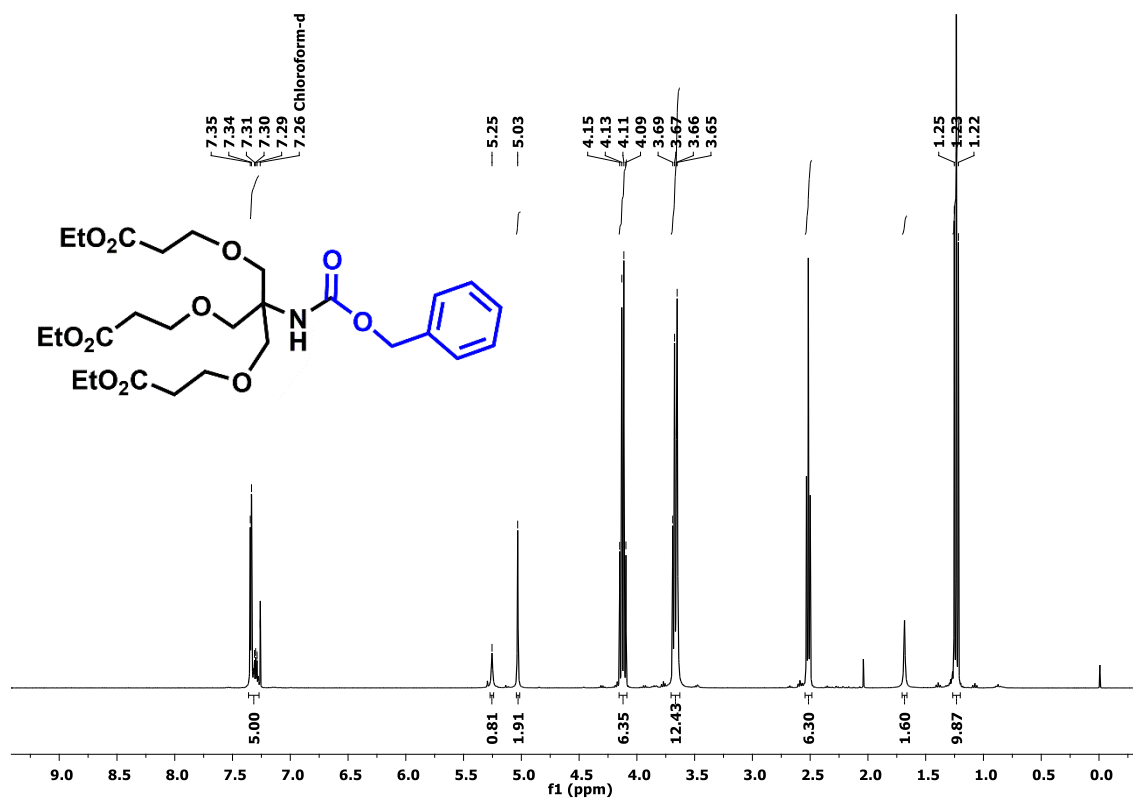
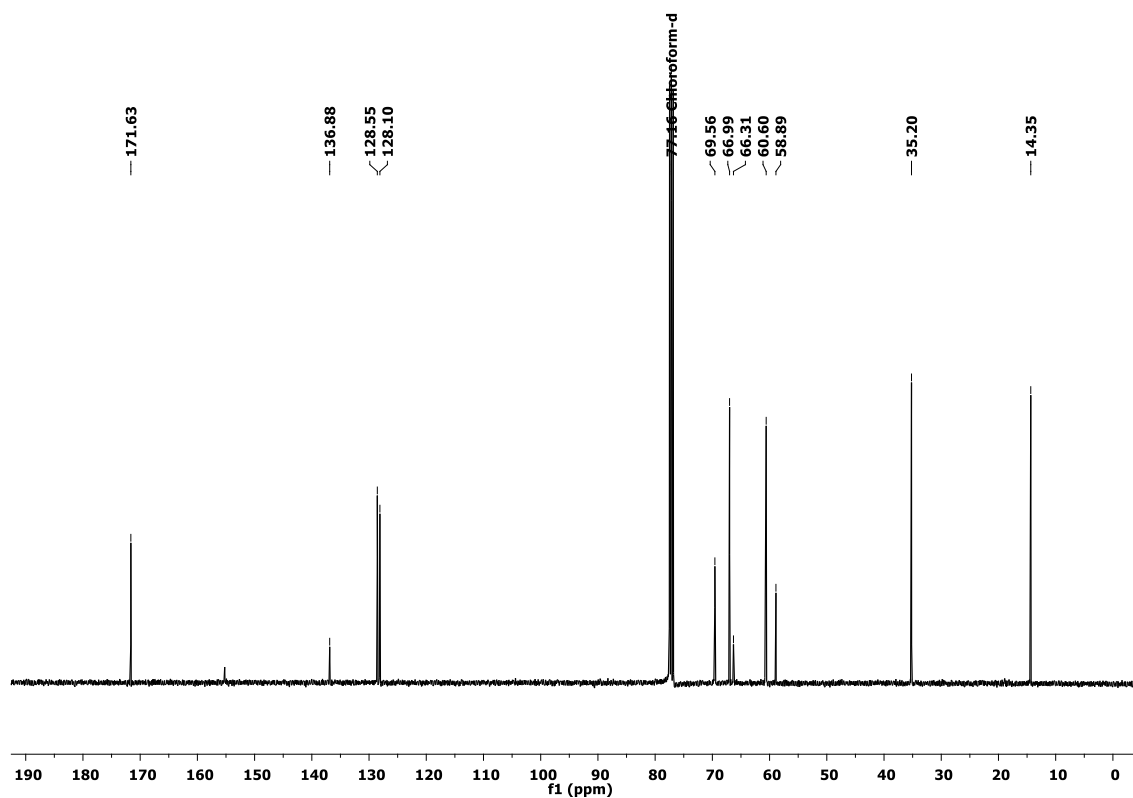


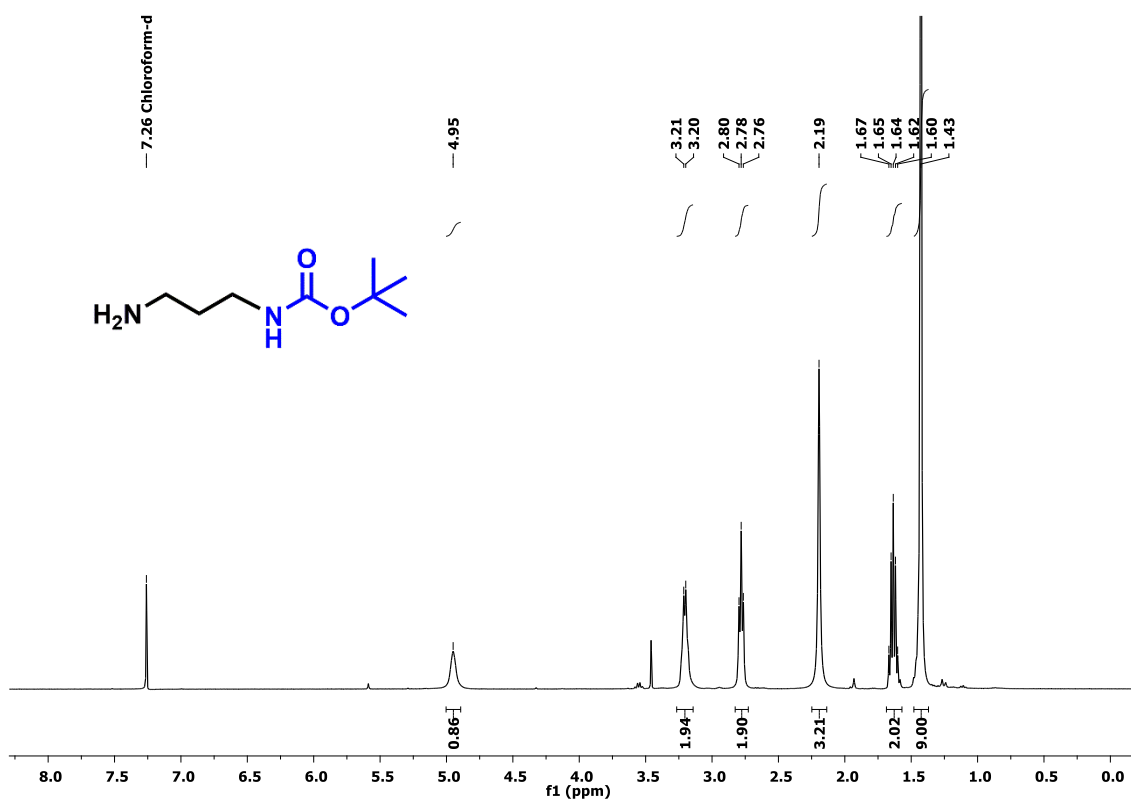
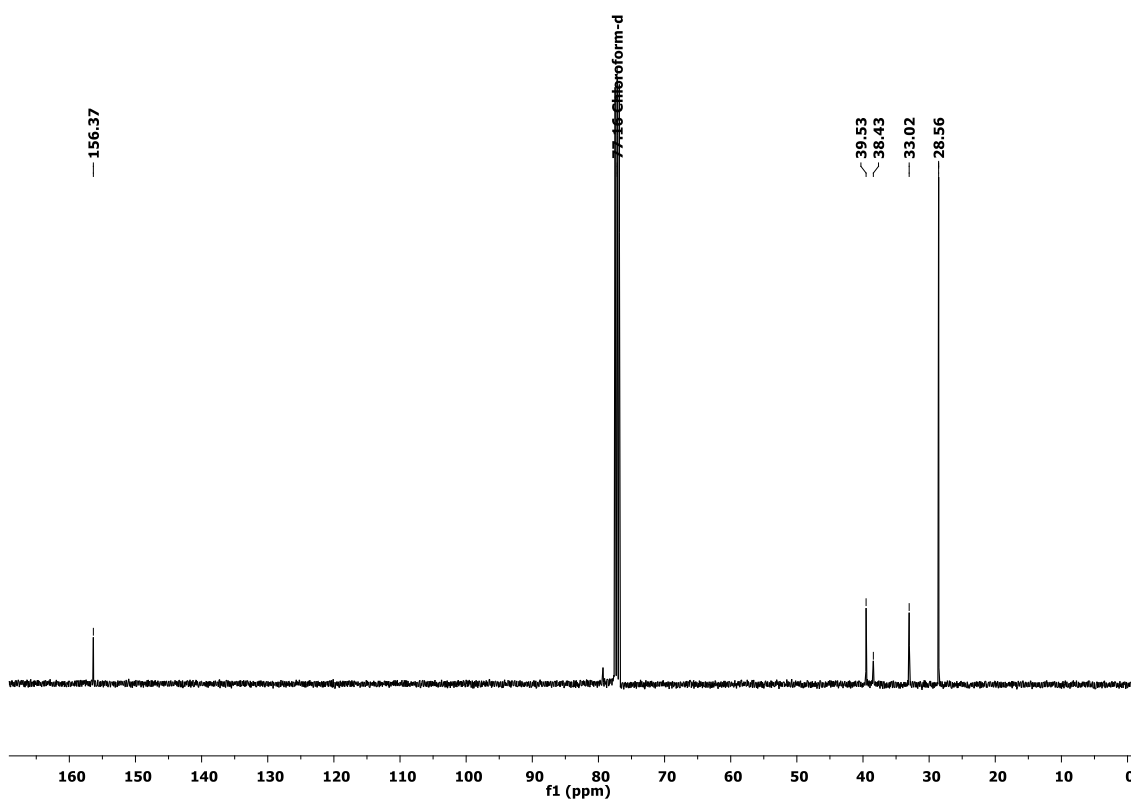
¹³C NMR of Compound 6



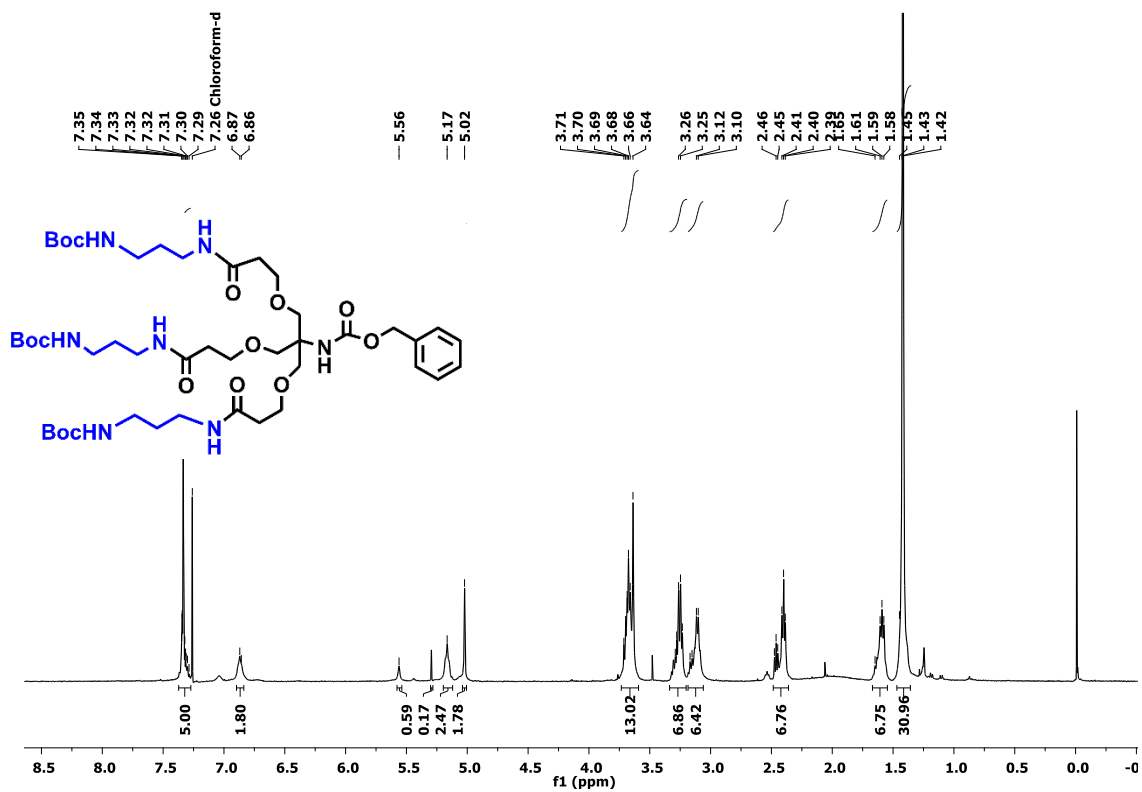
^1H NMR of Compound 8 ^{13}C NMR of Compound 8

^1H NMR of Compound 9 ^{13}C NMR of Compound 9

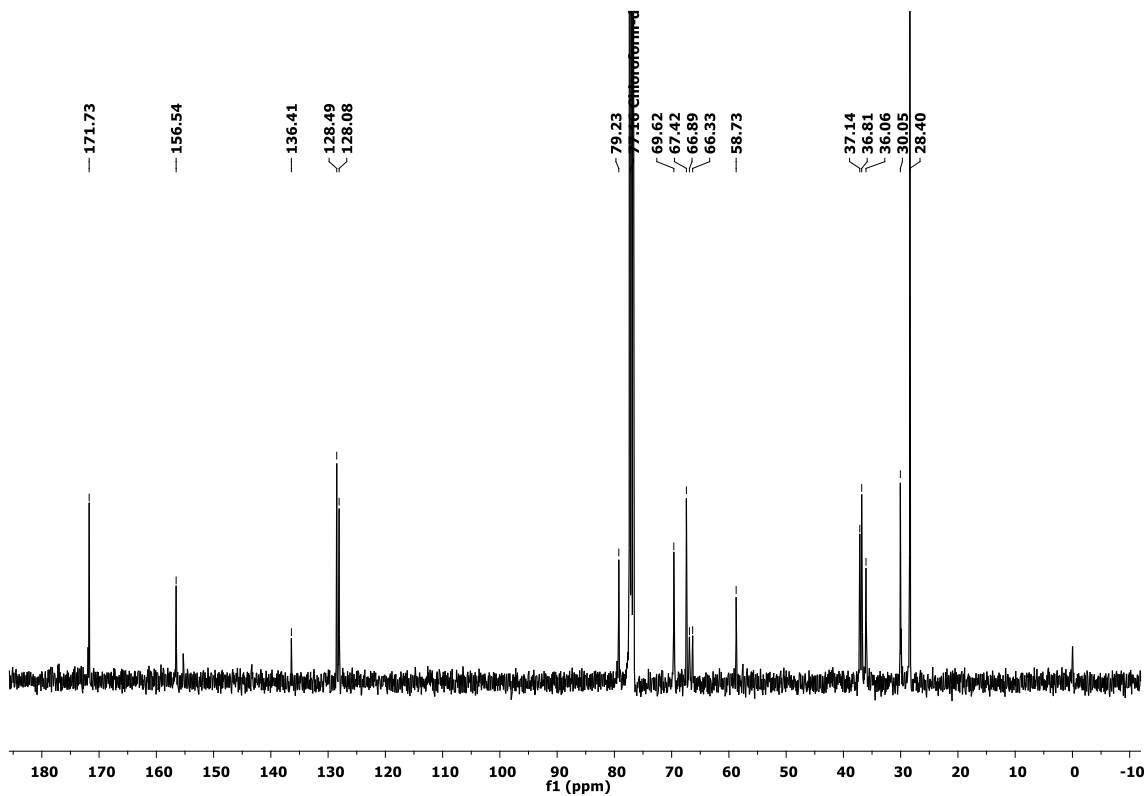
^1H NMR of Compound 10 ^{13}C NMR of Compound 10

^1H NMR of Compound 12 ^{13}C NMR of Compound 12

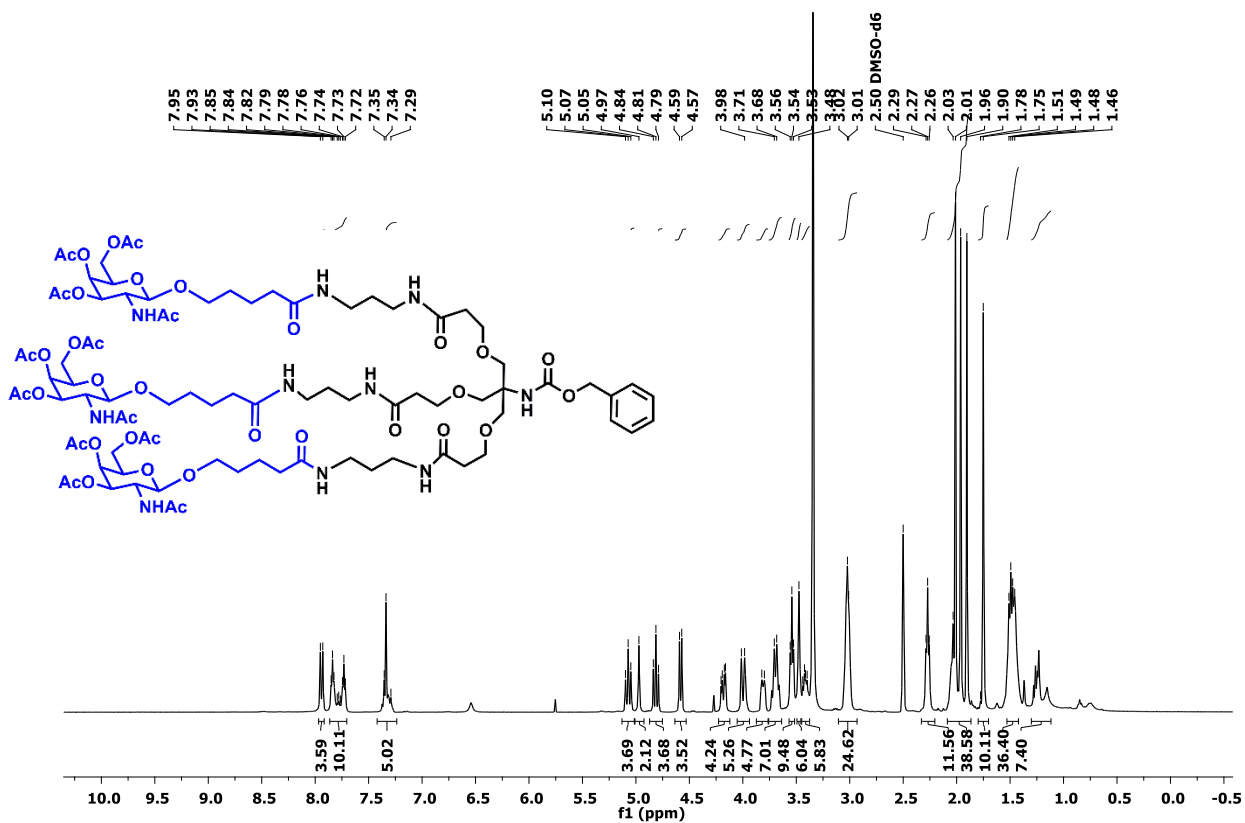
¹H NMR of Compound 13



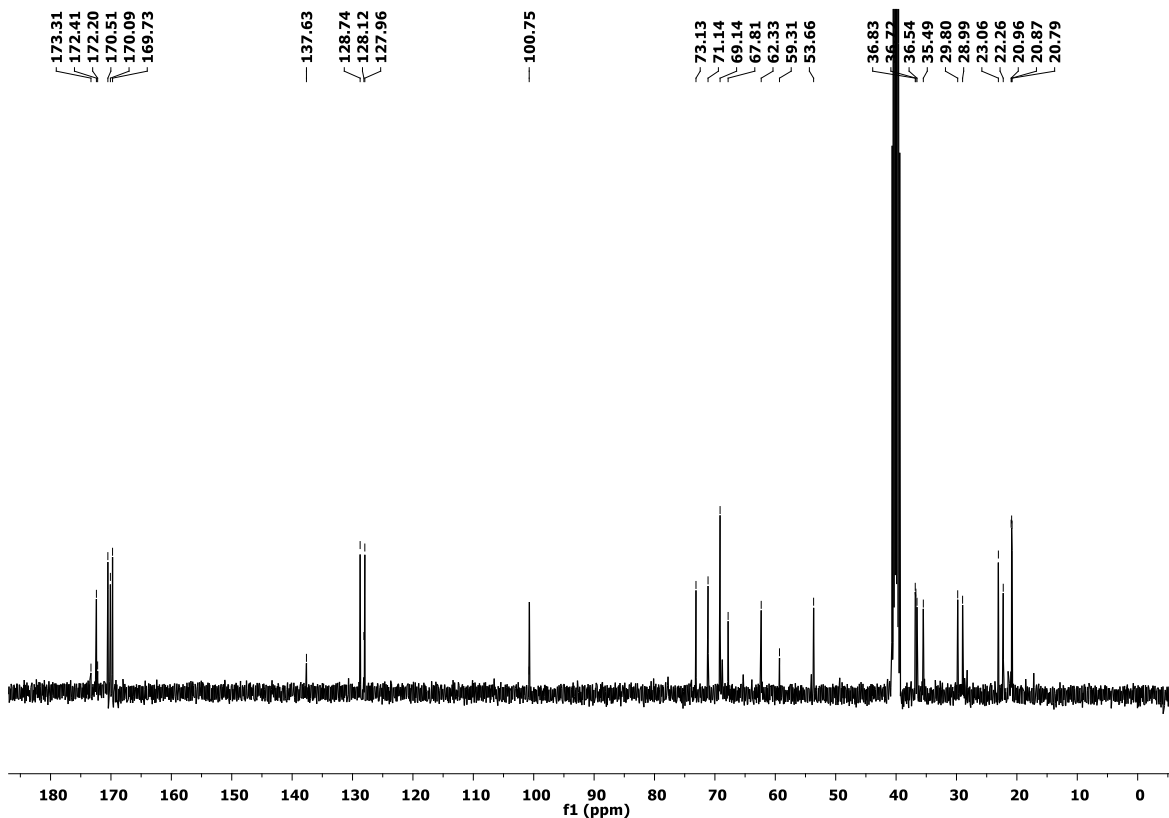
¹³C NMR of Compound 13

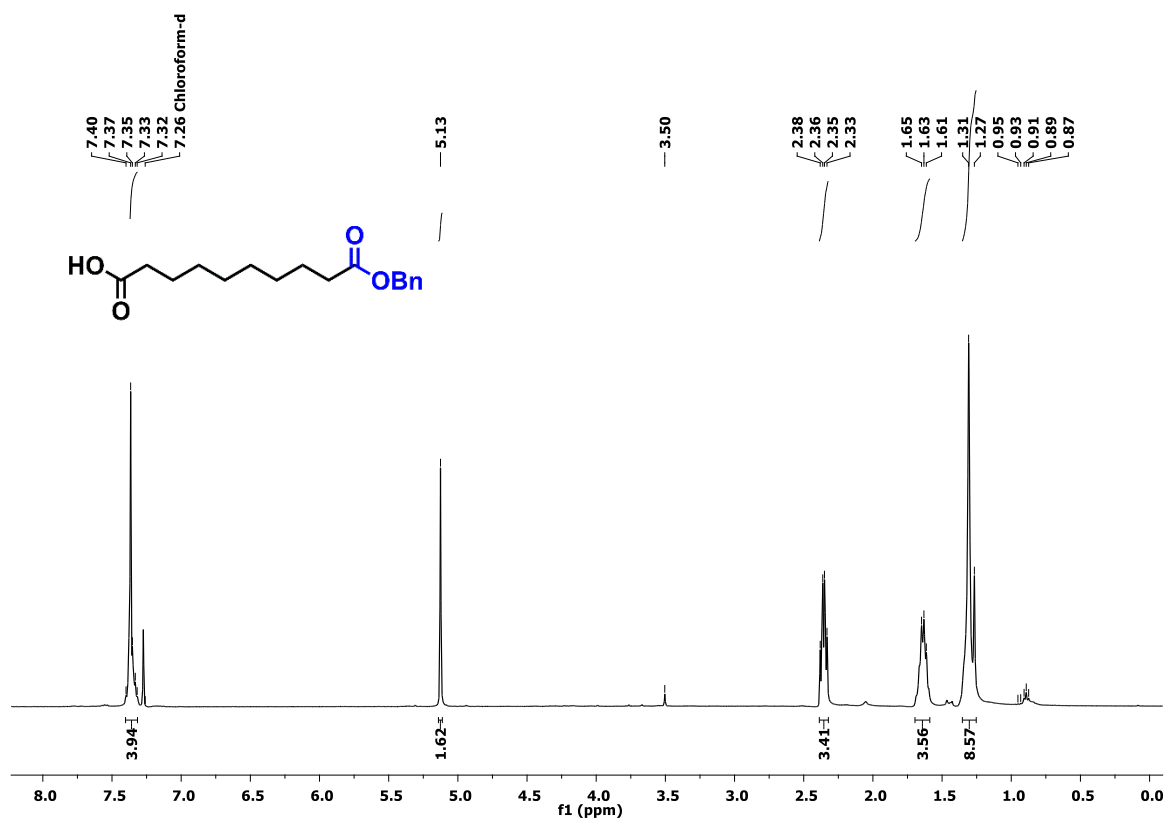
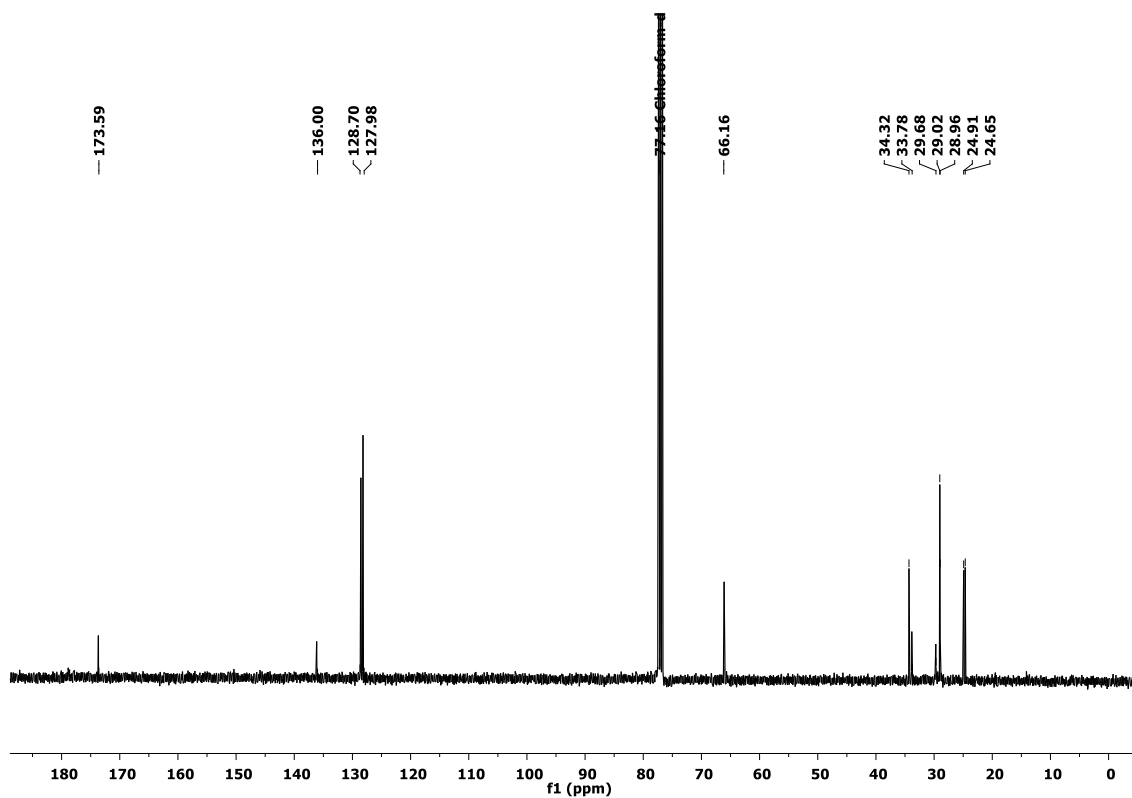


¹H NMR of Compound 15

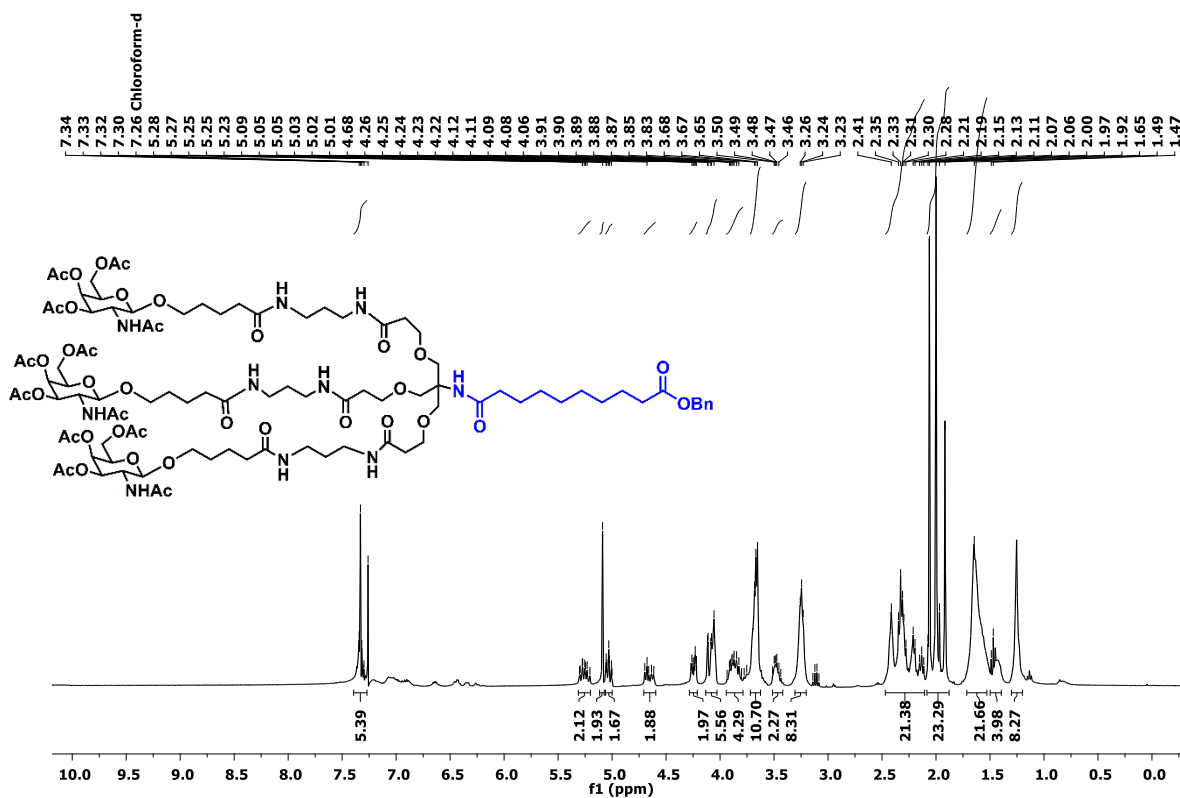


¹³C NMR of Compound 15

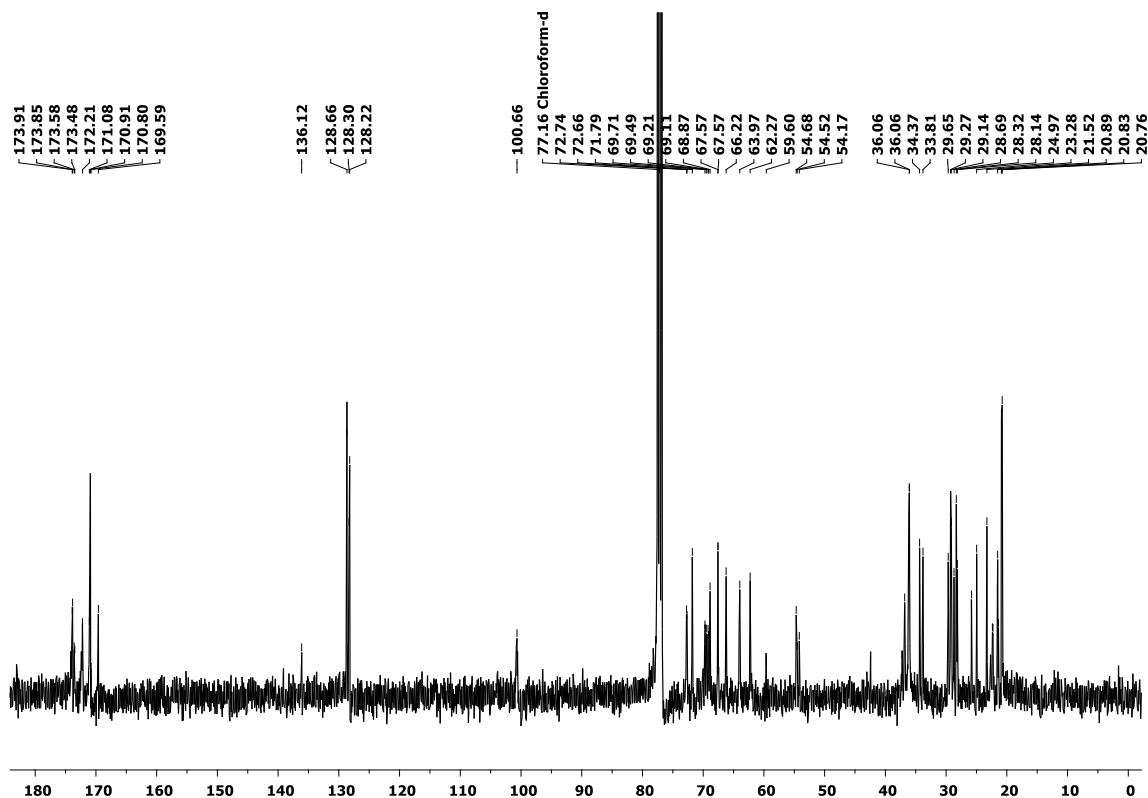


^1H NMR of Compound 17 ^{13}C NMR of Compound 17

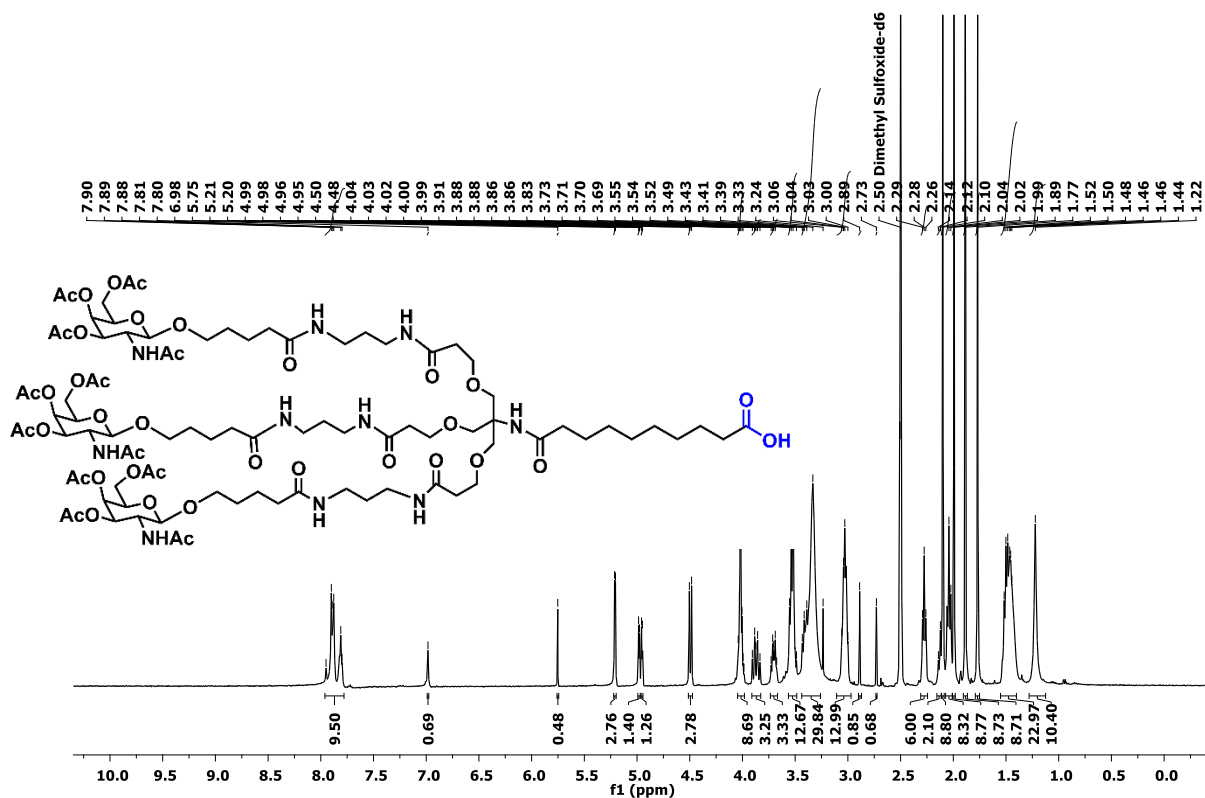
¹H NMR of Compound 18



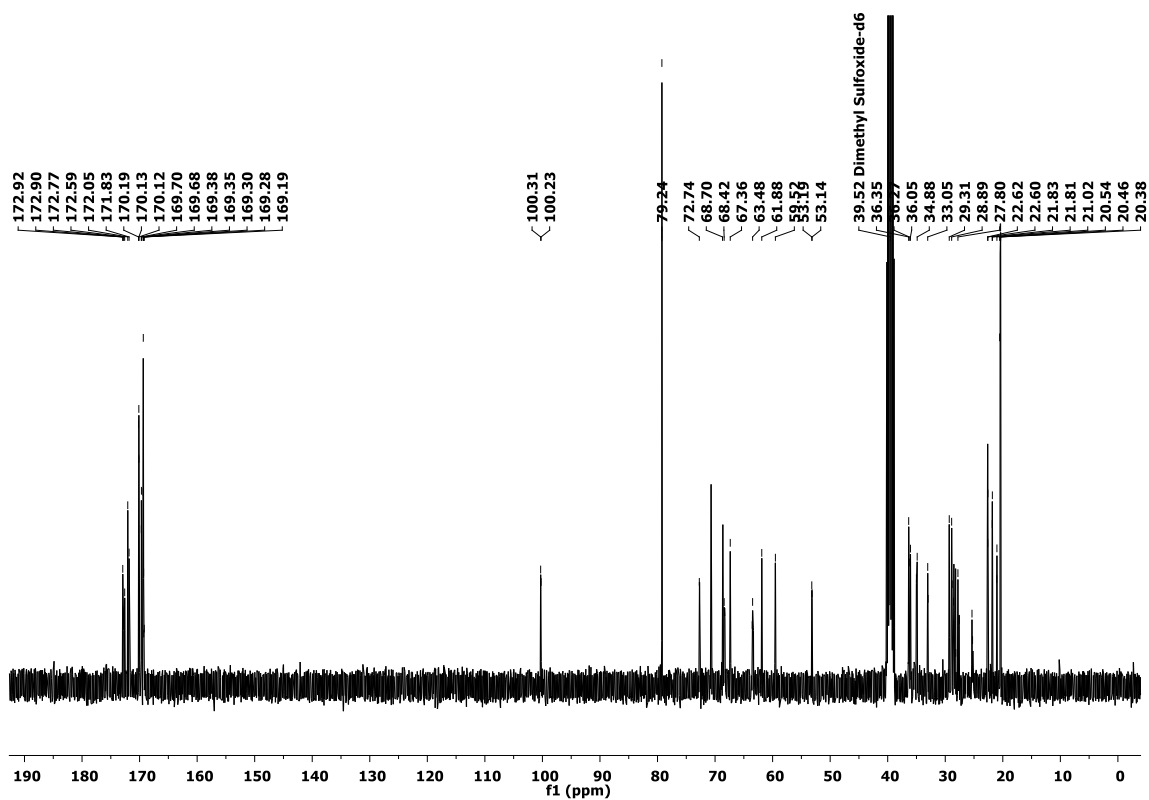
¹³C NMR of Compound 18

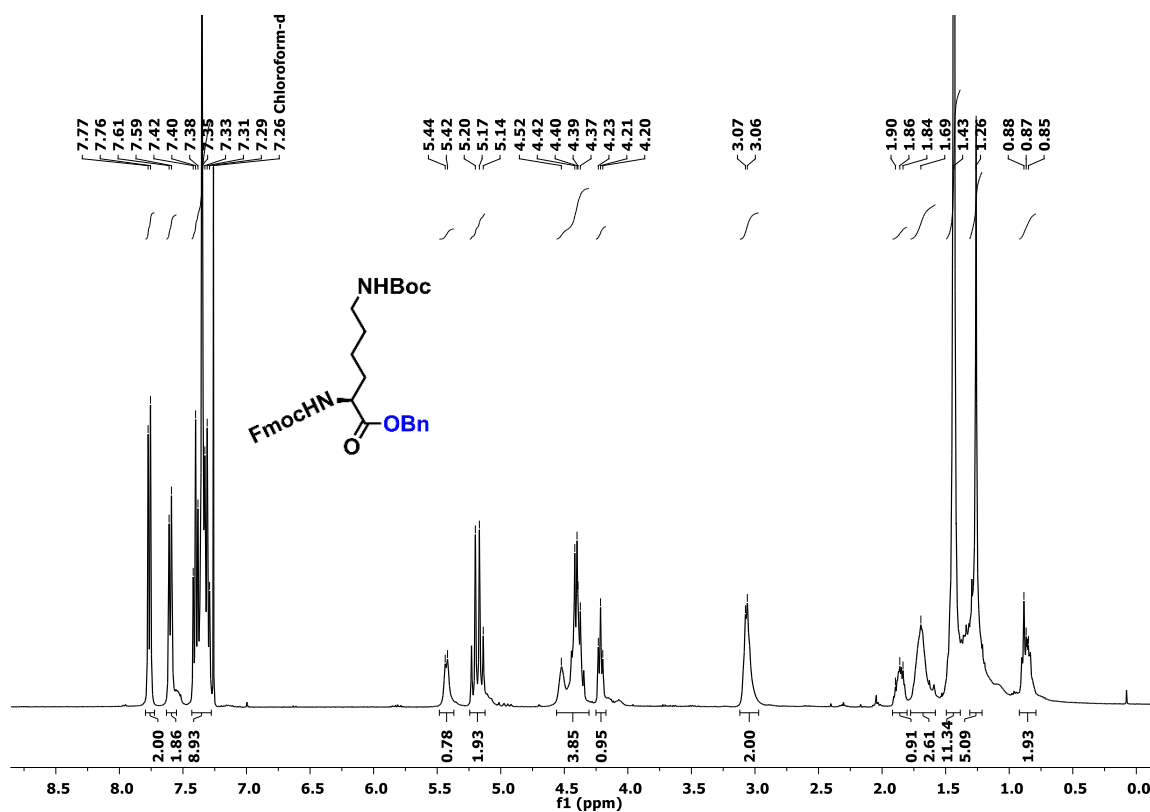
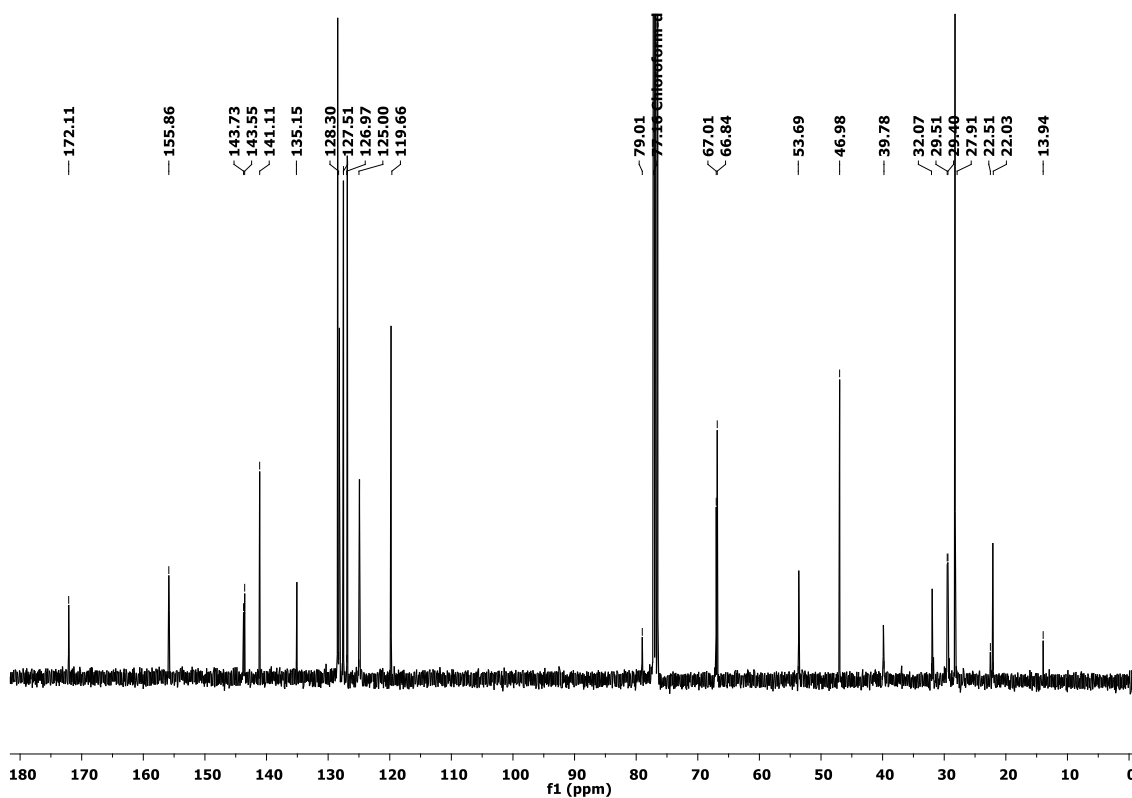


¹H NMR of Compound 19

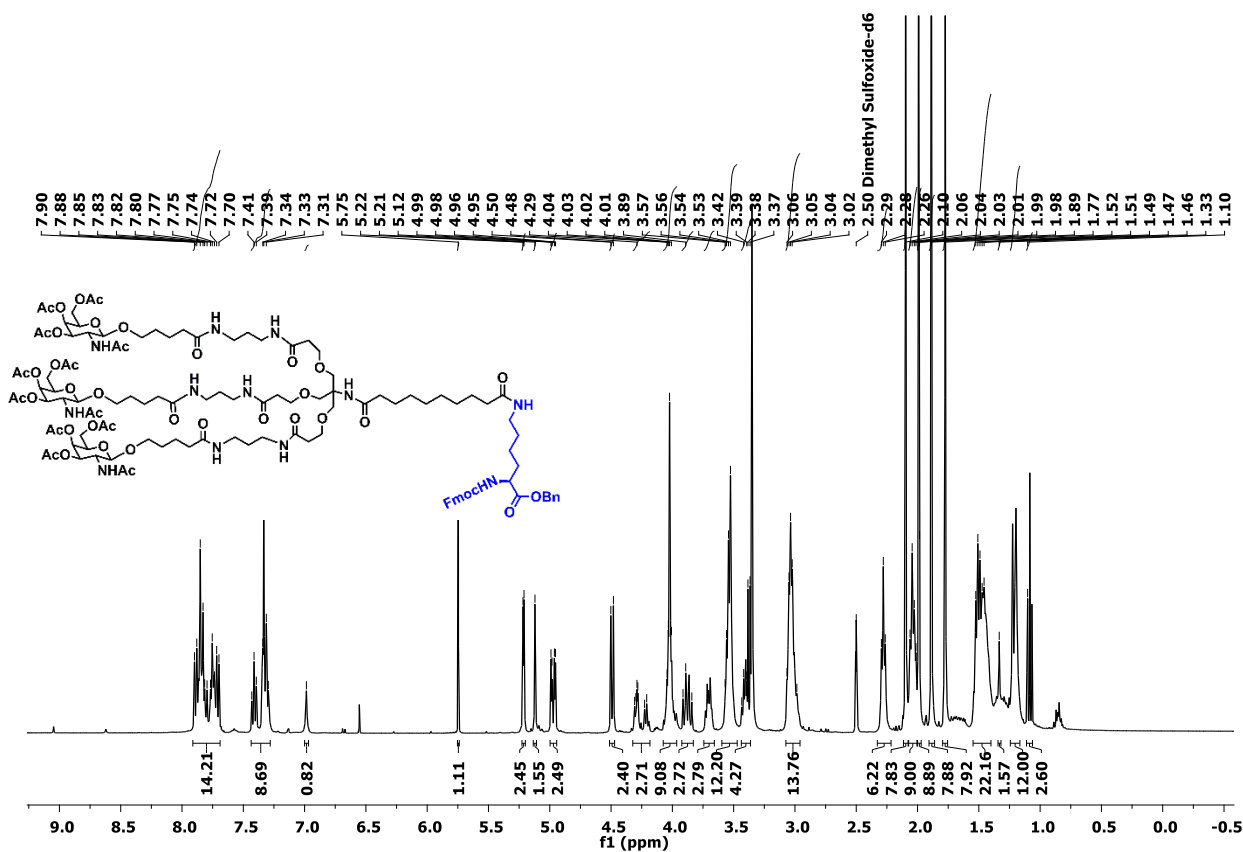


¹³C NMR of Compound 19

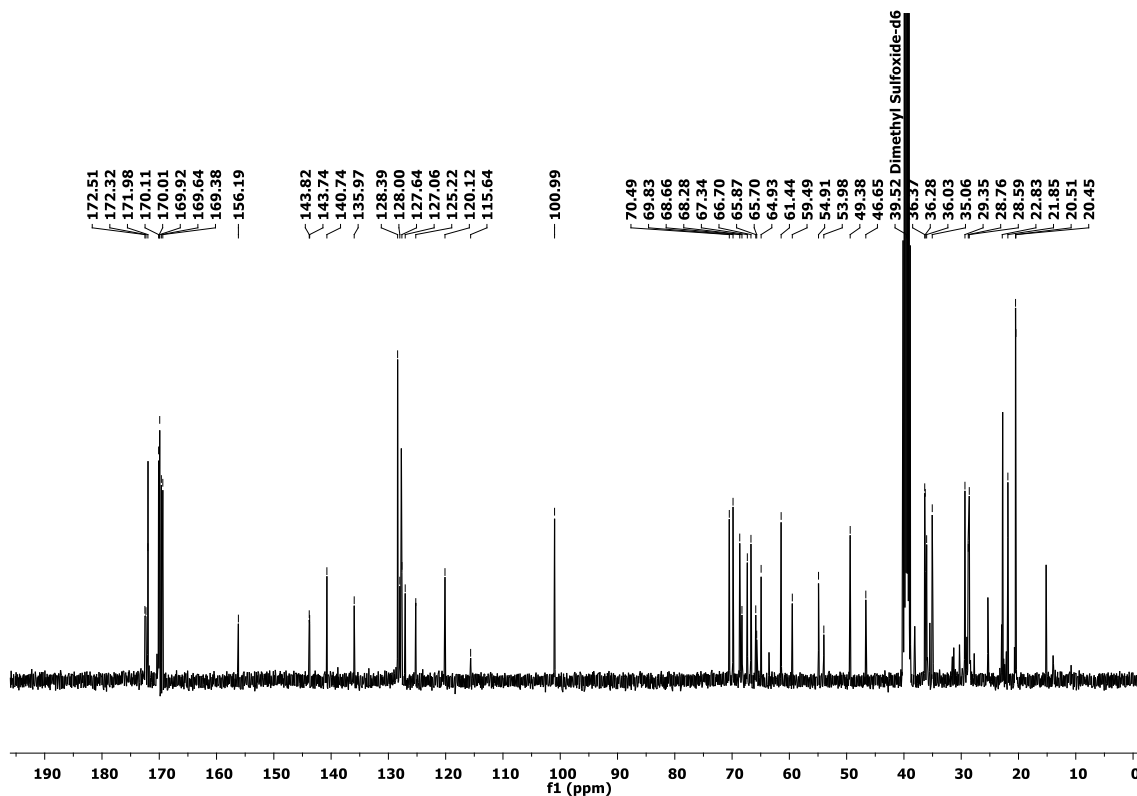


^1H NMR of Compound 21 ^{13}C NMR spectrum of Compound 21

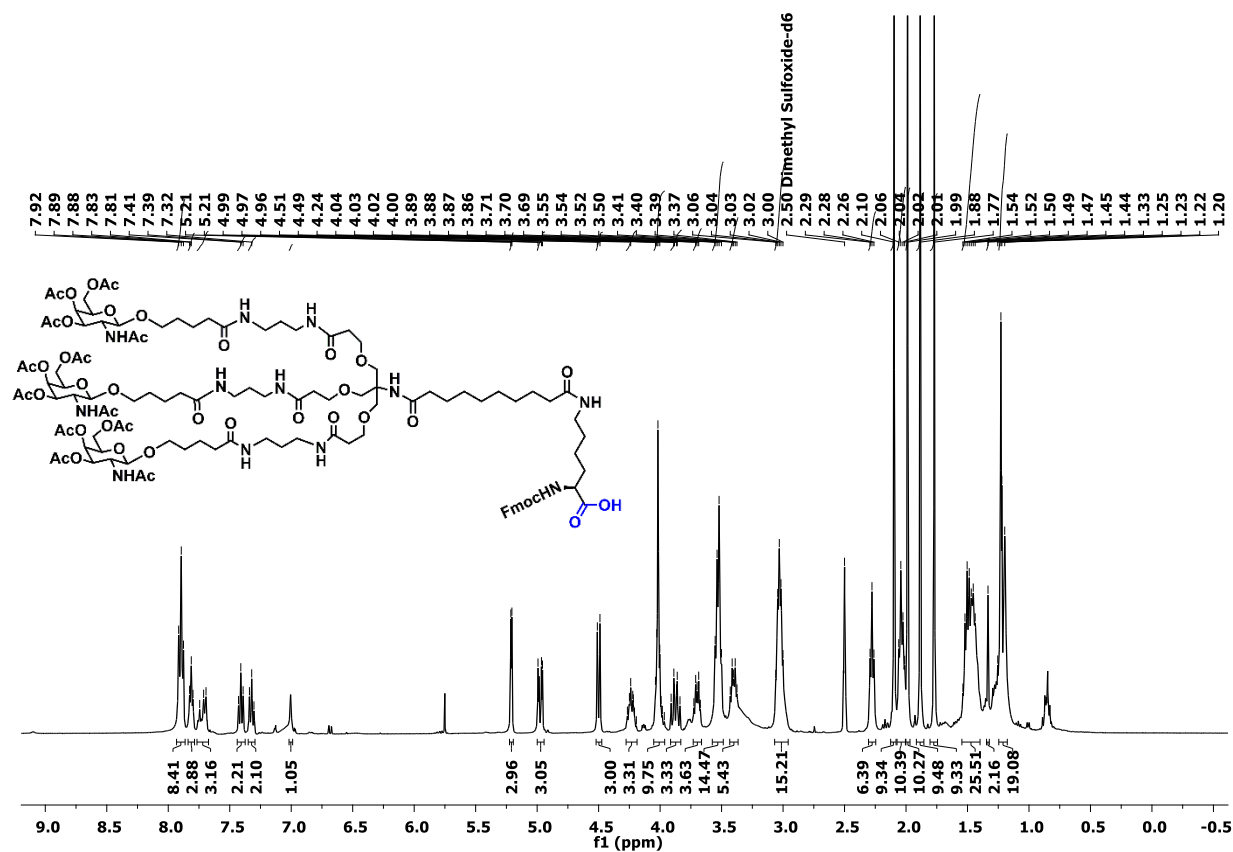
¹H NMR of Compound 23



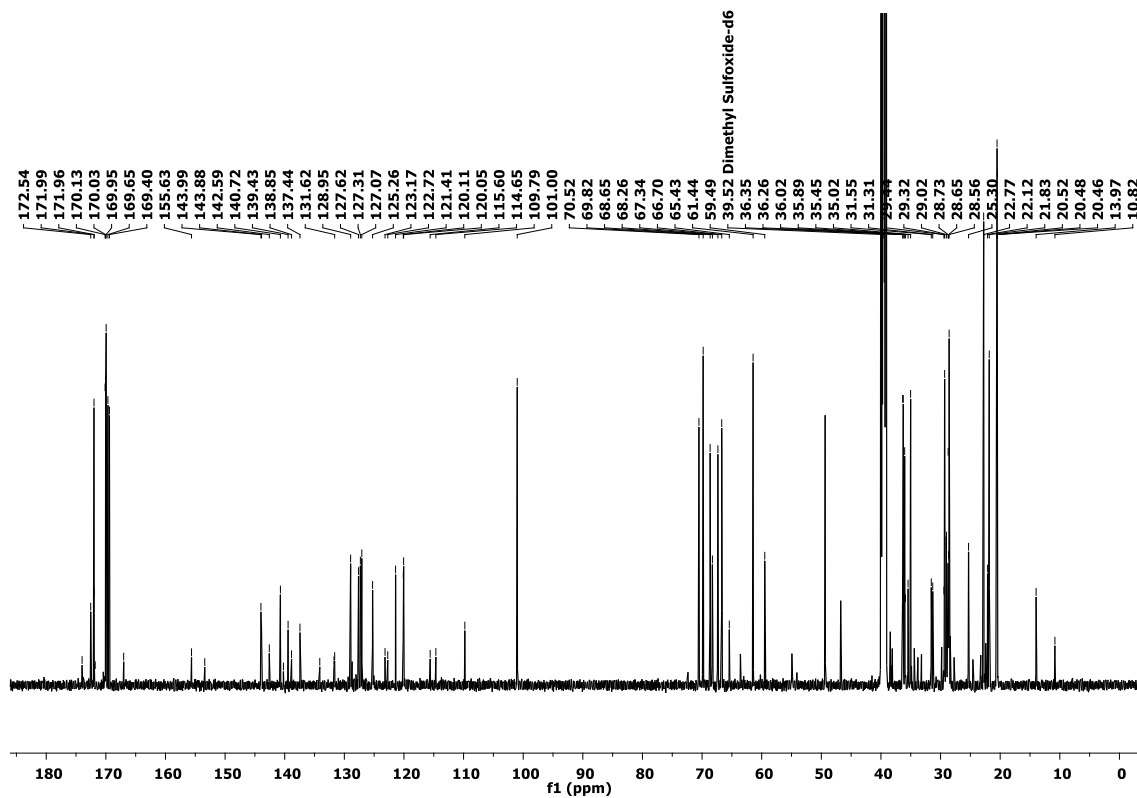
¹³C NMR of Compound 23

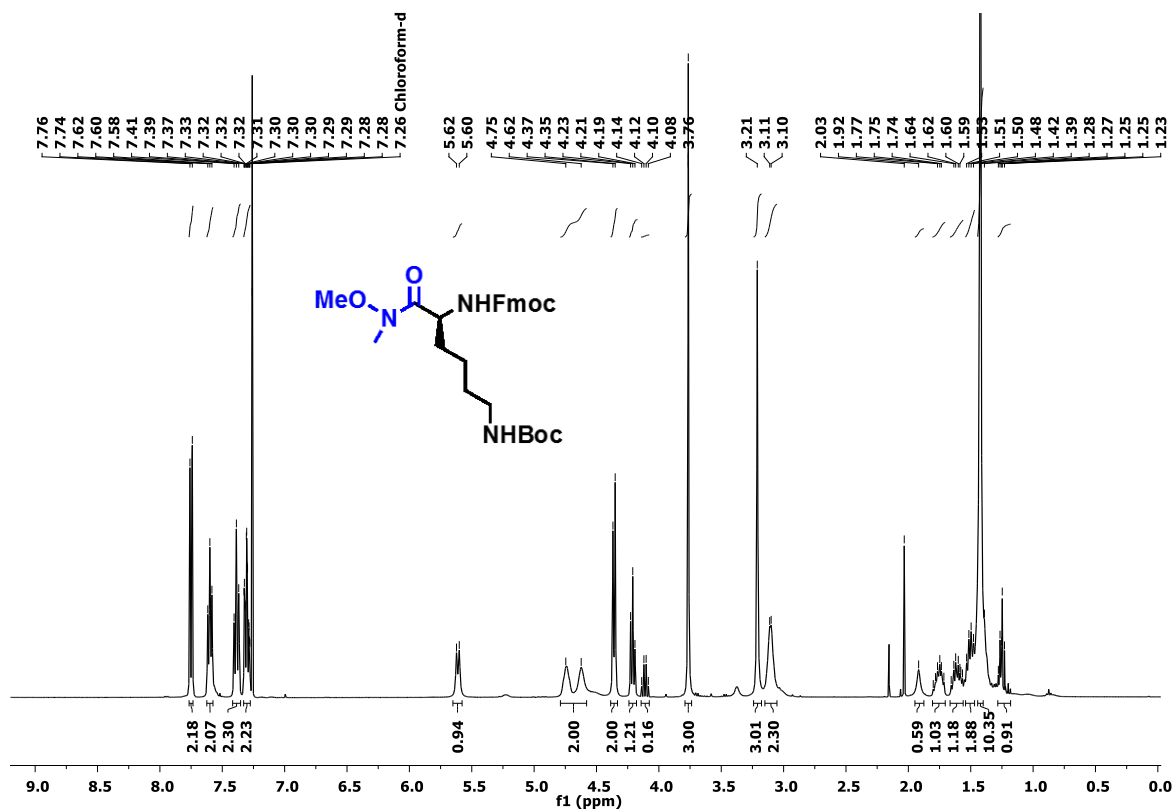
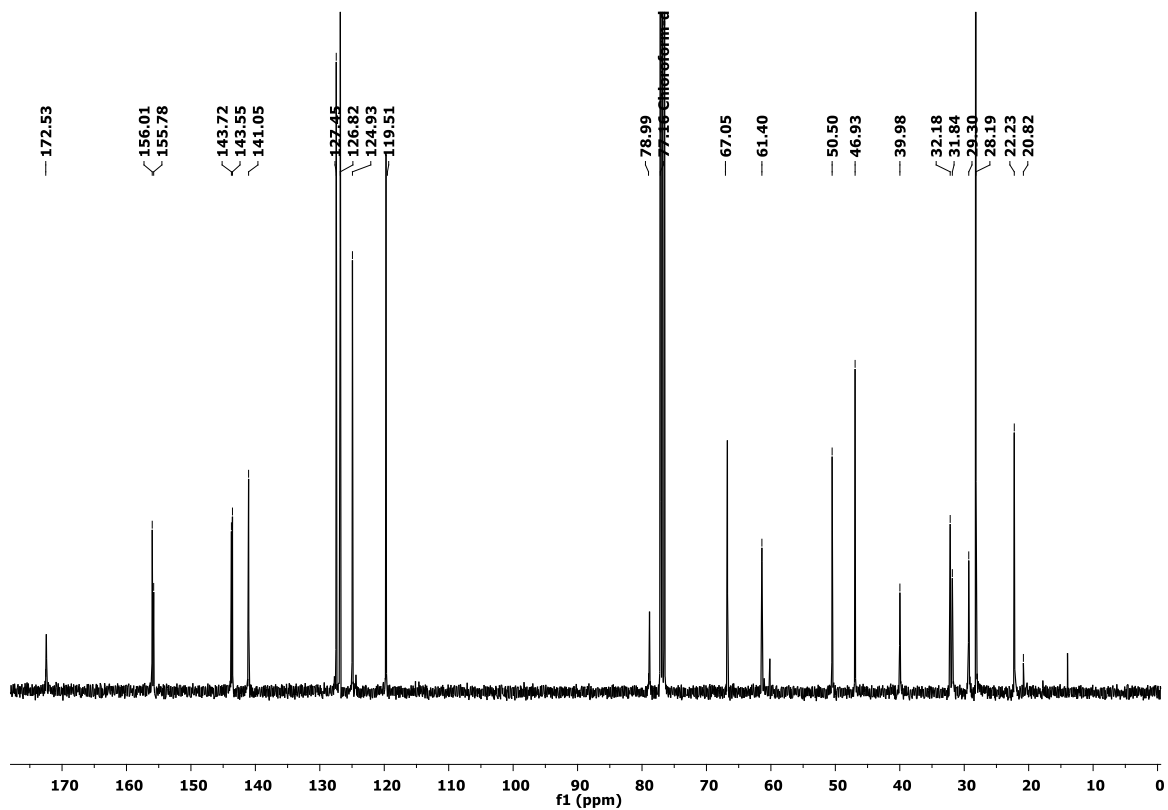


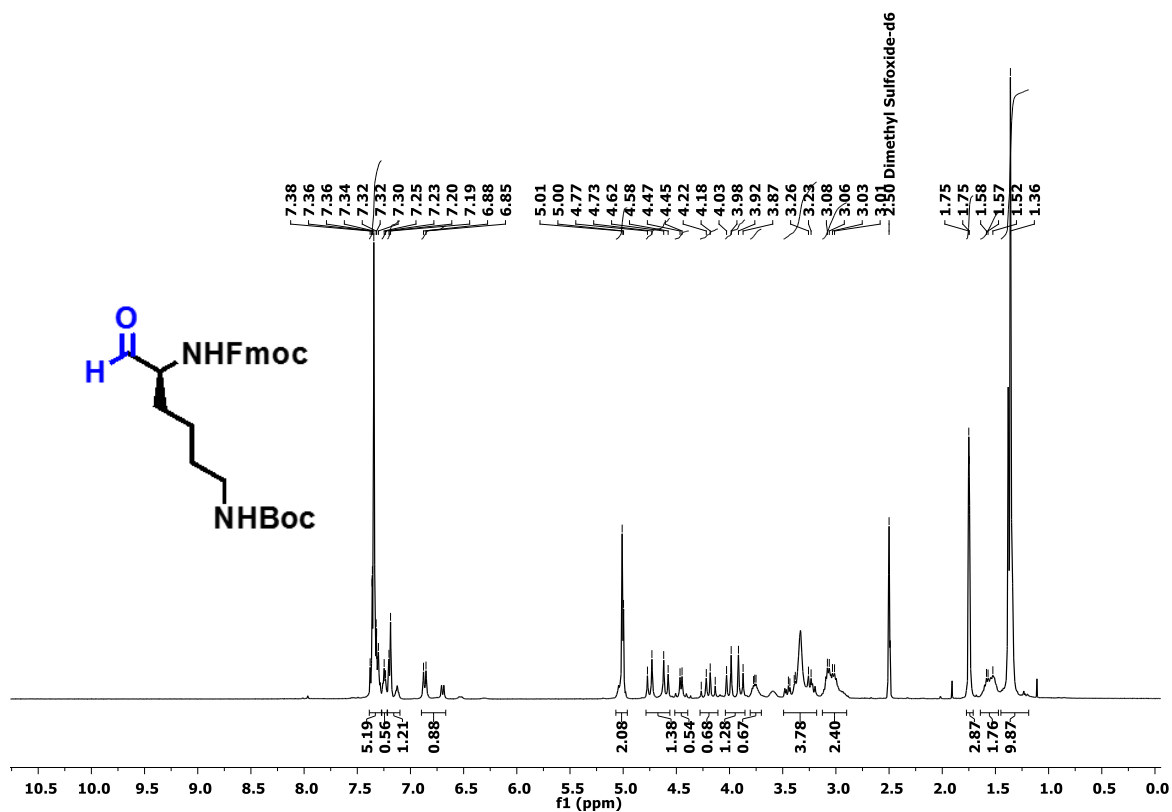
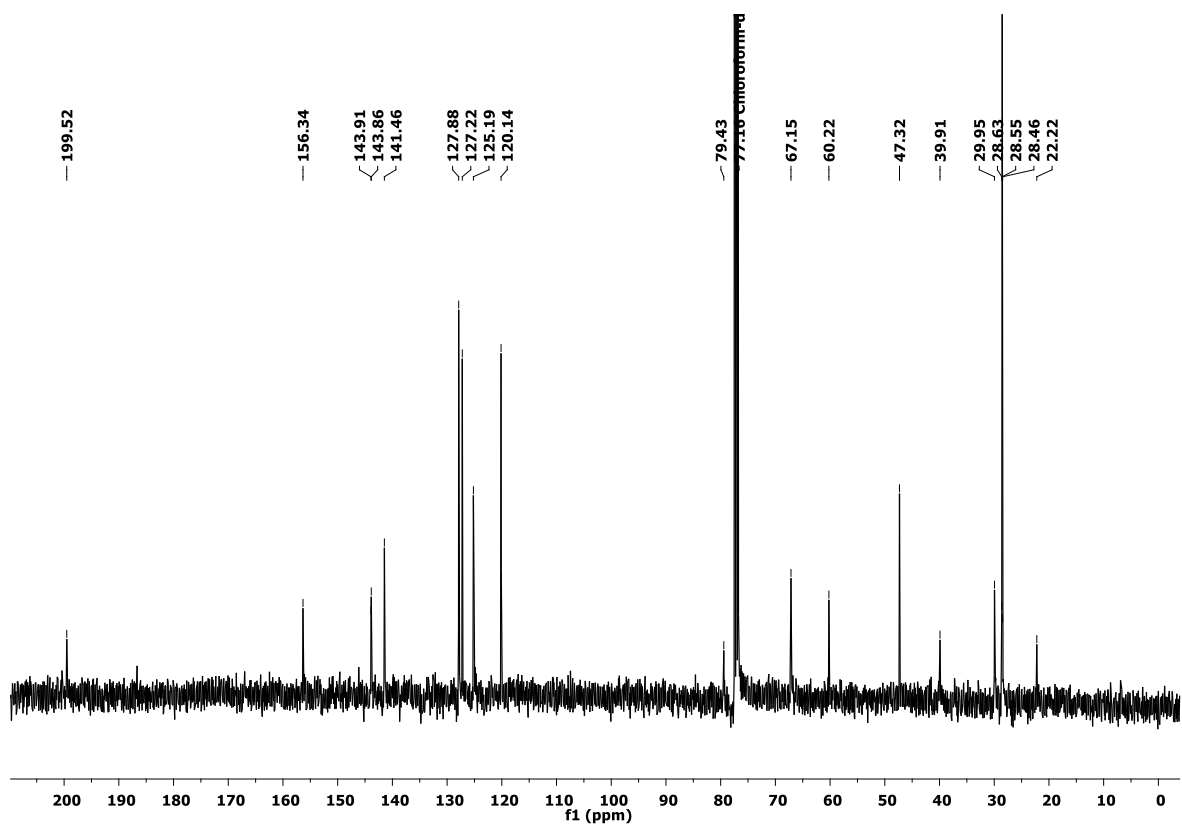
¹H NMR of Compound 24

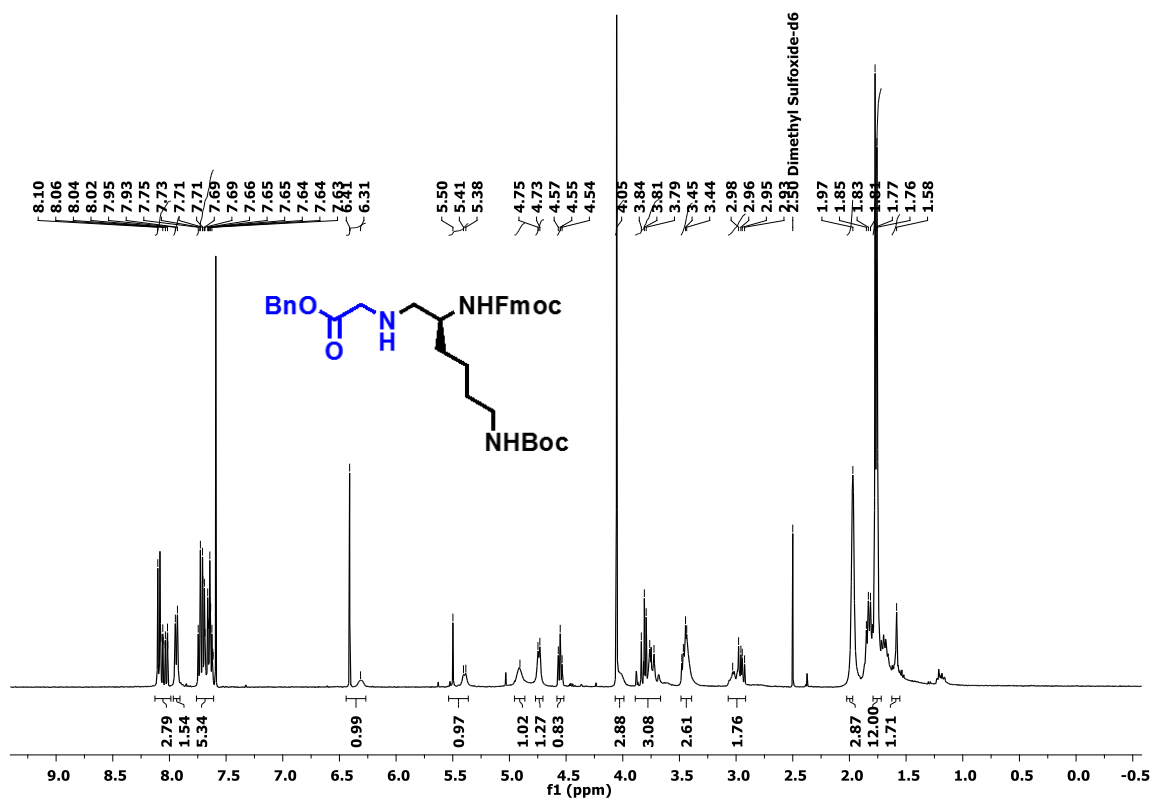
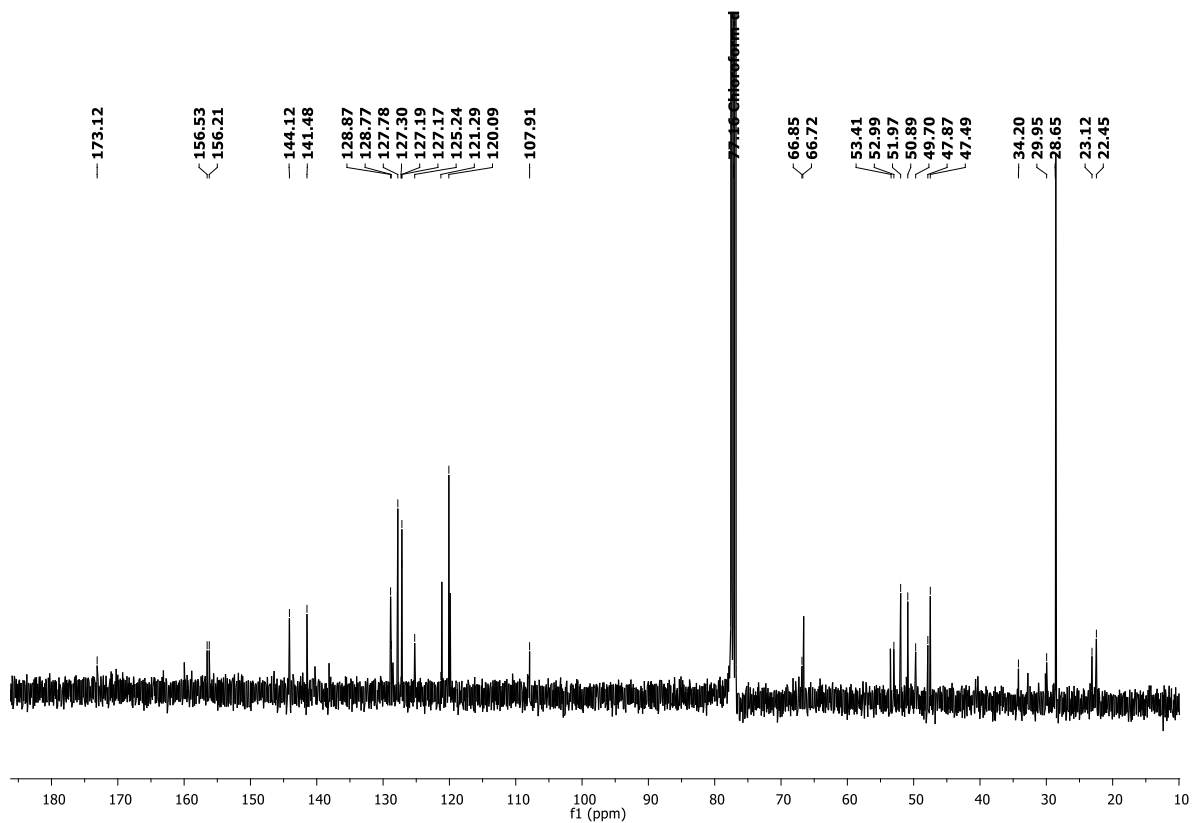


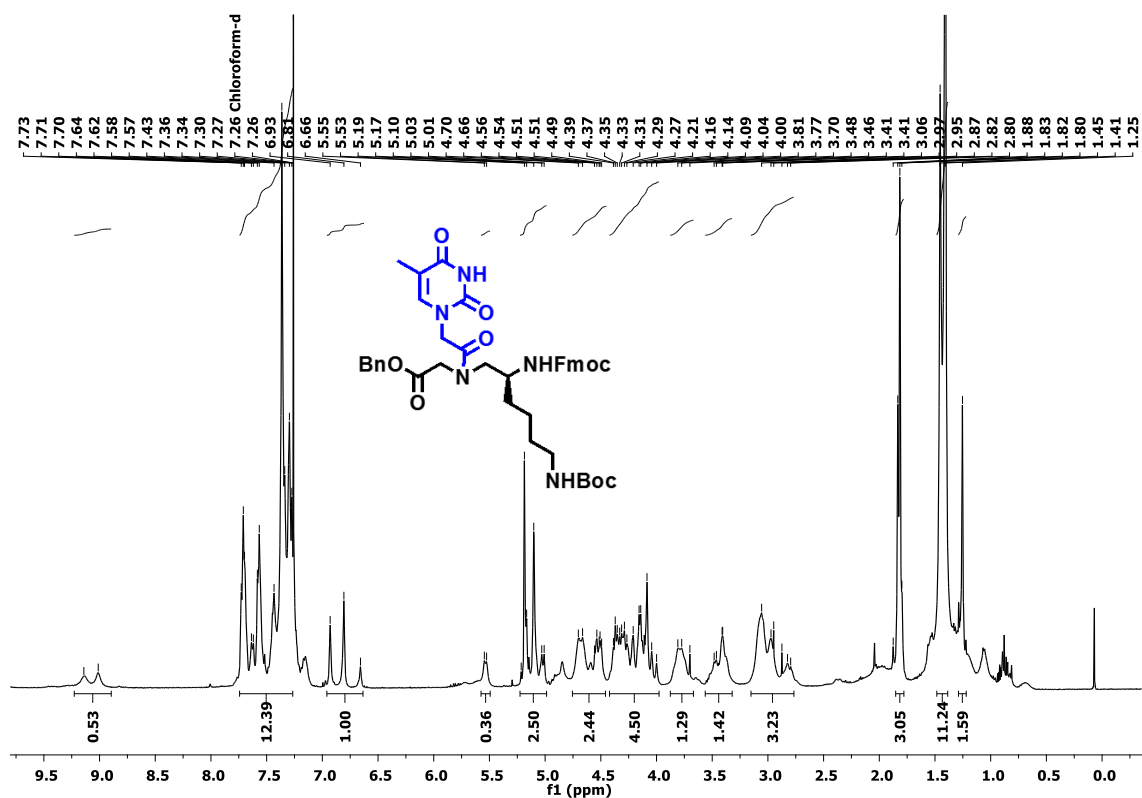
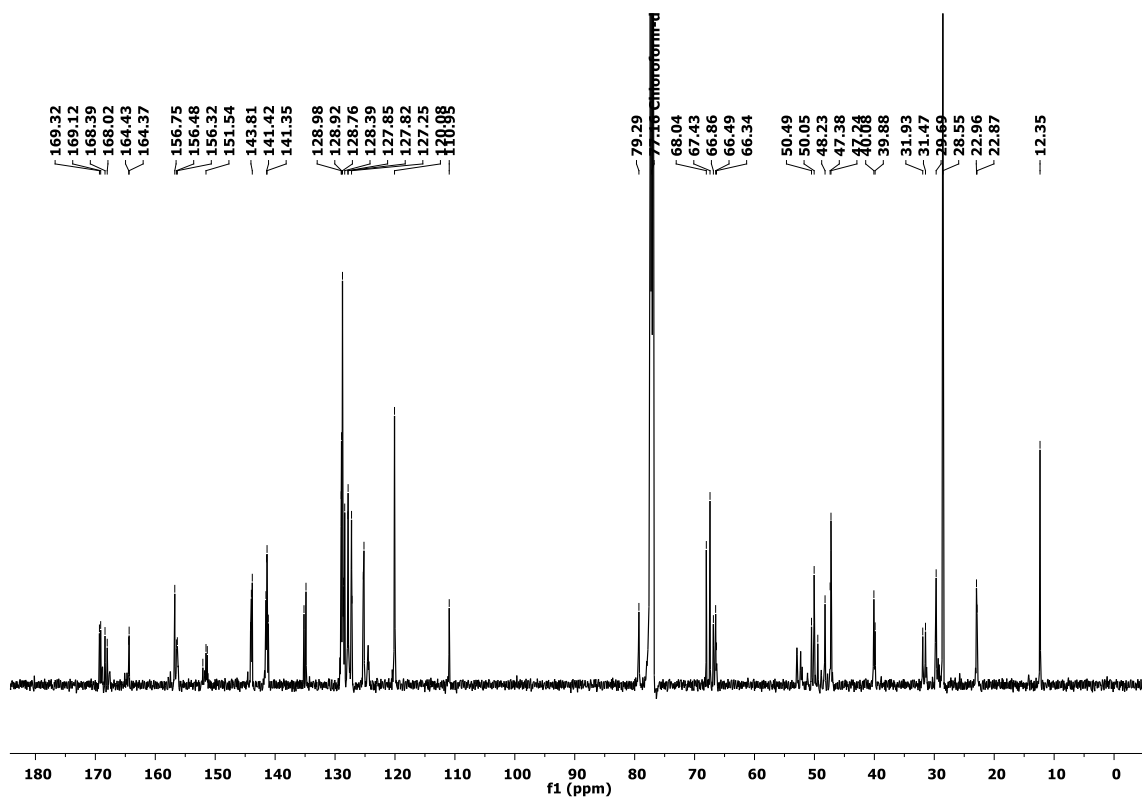
¹³C NMR of Compound 24

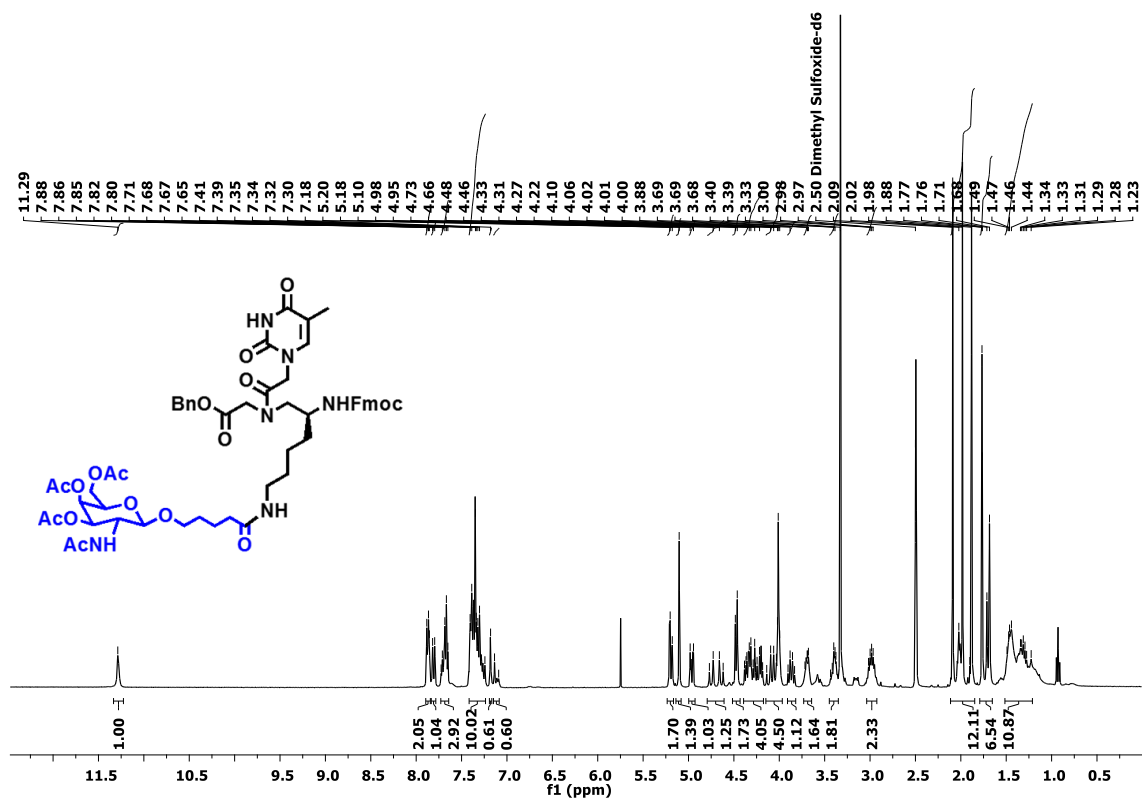
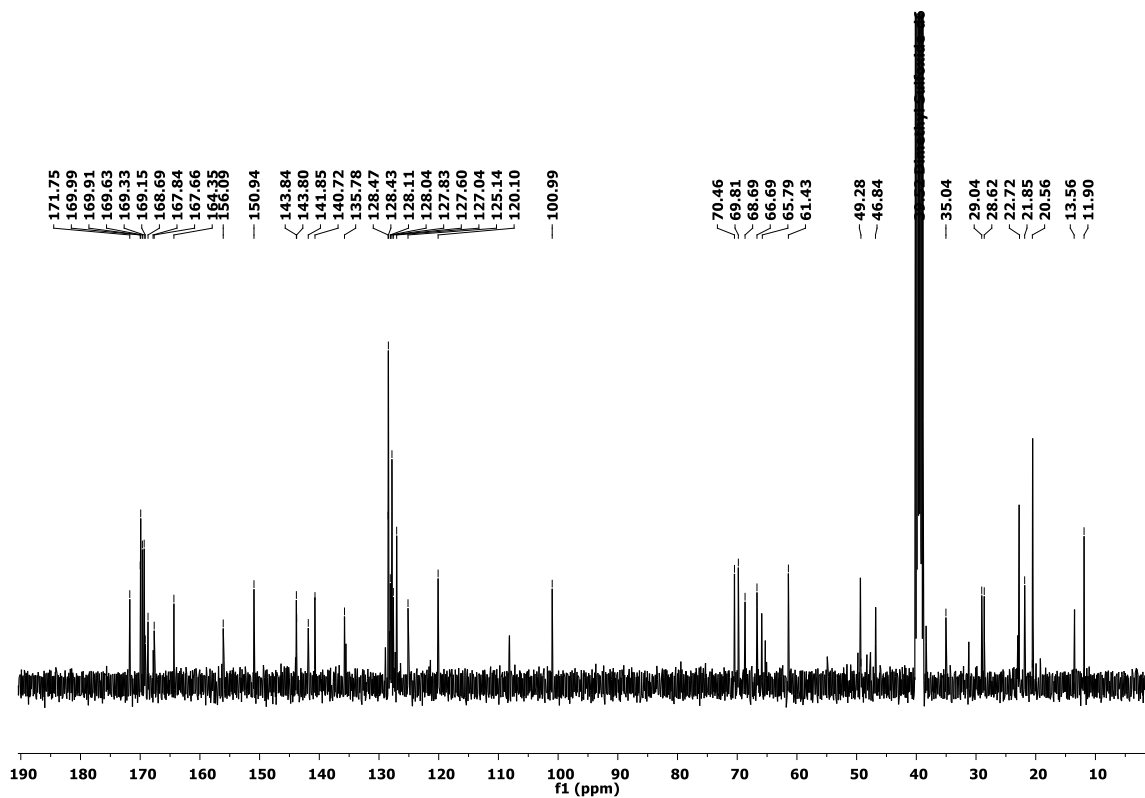


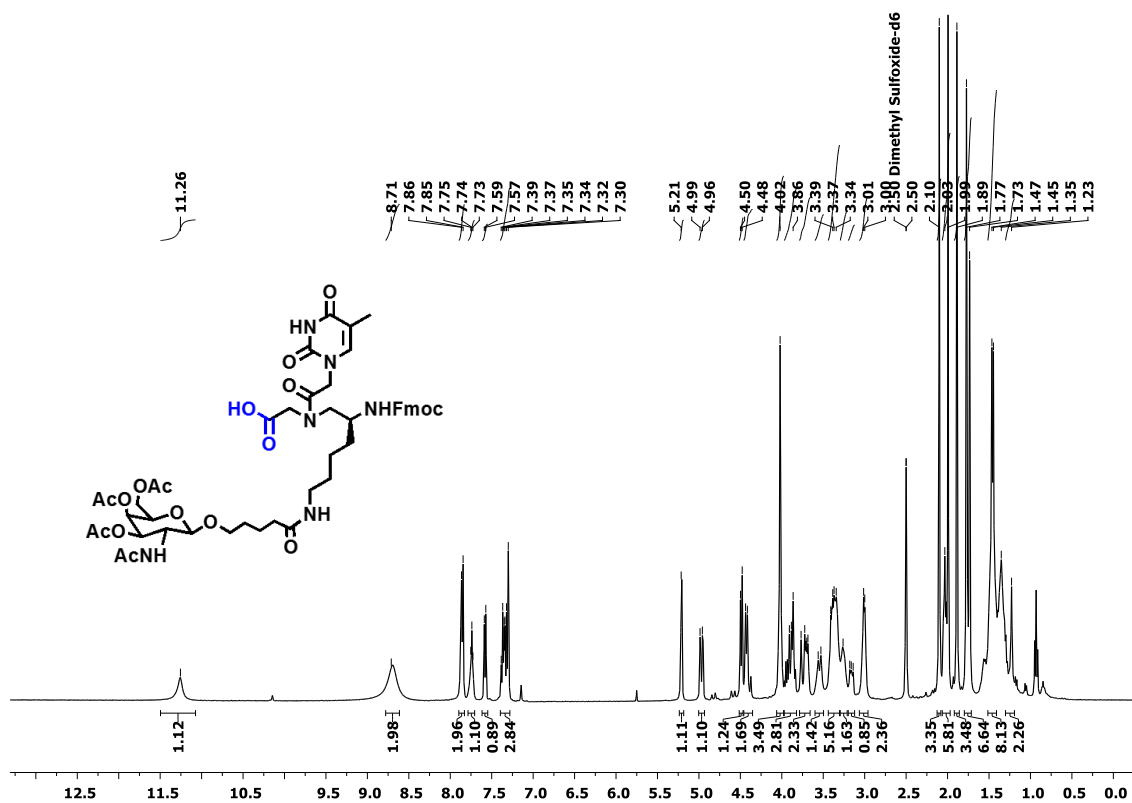
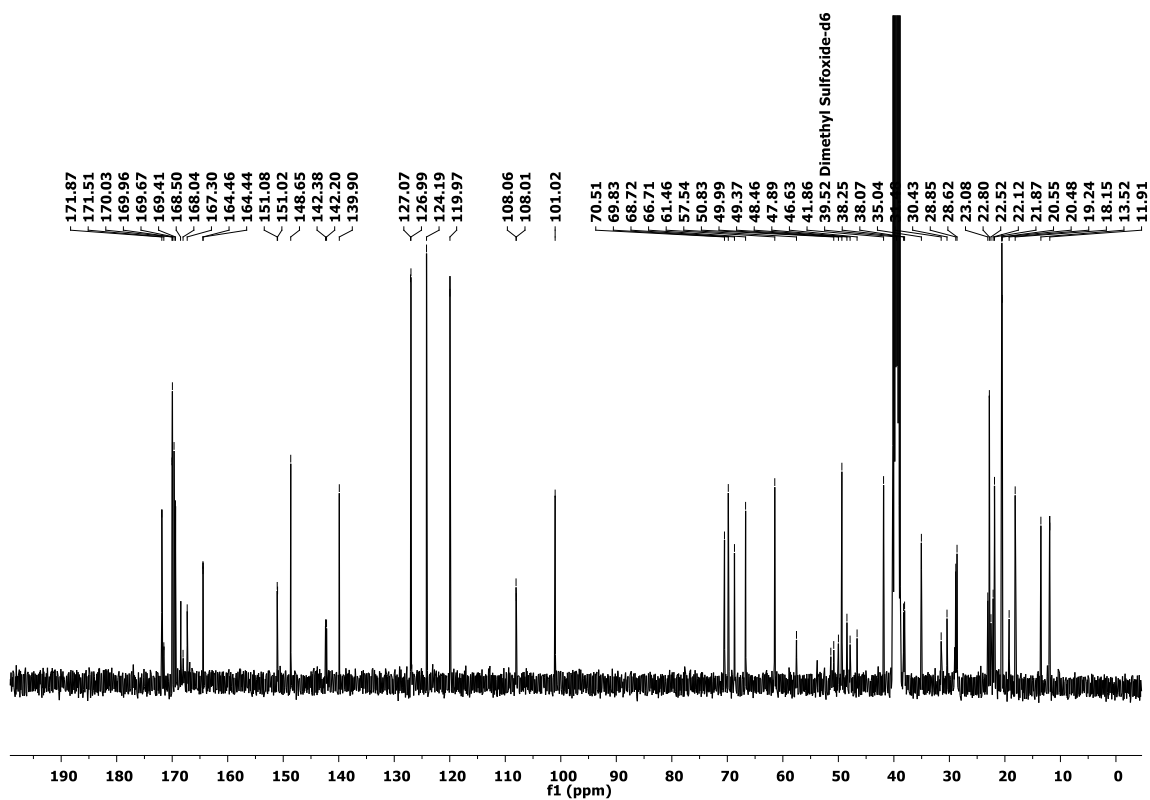
^1H NMR of Compound 26 ^{13}C NMR of Compound 26

^1H NMR of Compound 27 ^{13}C NMR of Compound 27

^1H NMR of Compound 28 ^{13}C NMR spectrum of Compound 28

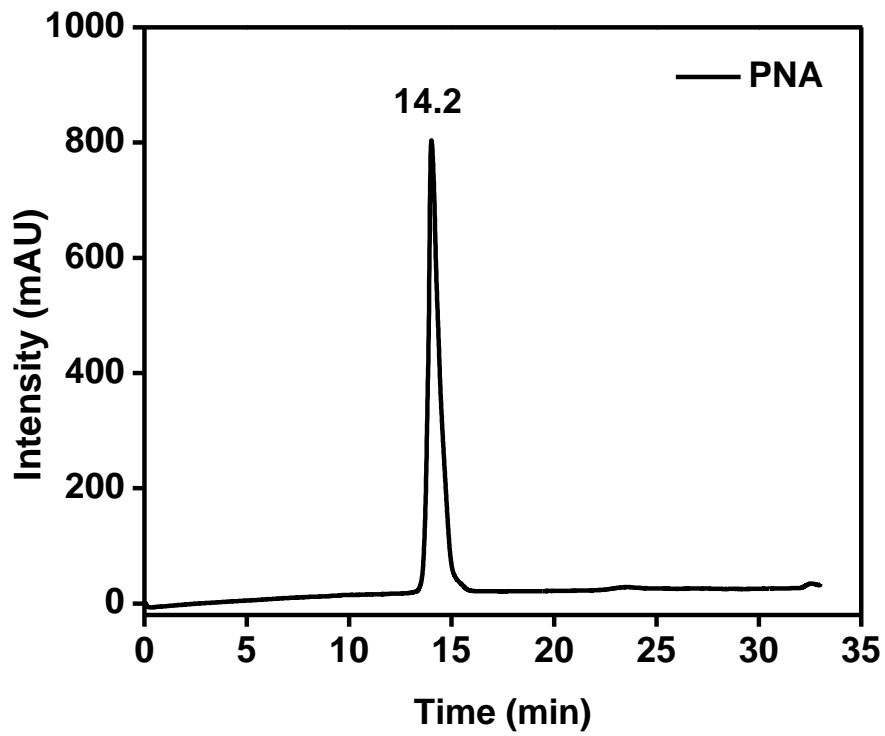
¹H NMR of Compound 30¹³C NMR of Compound 30

^1H NMR of Compound 32 ^{13}C NMR of Compound 32

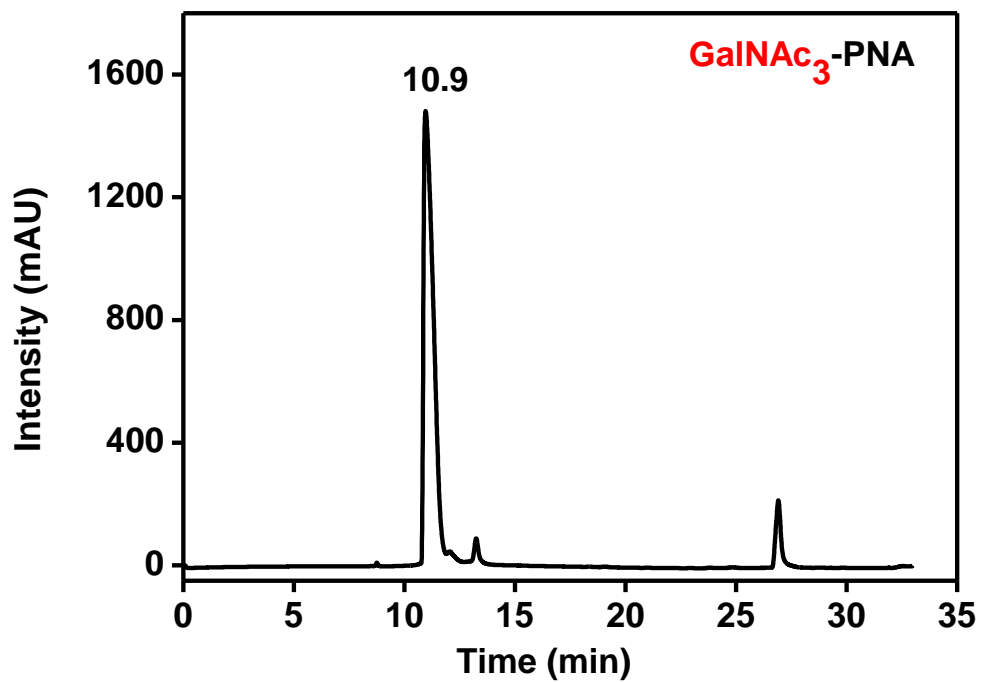
^1H NMR of Compound 33 ^{13}C NMR of Compound 33

HPLC chromatograms of peptides:

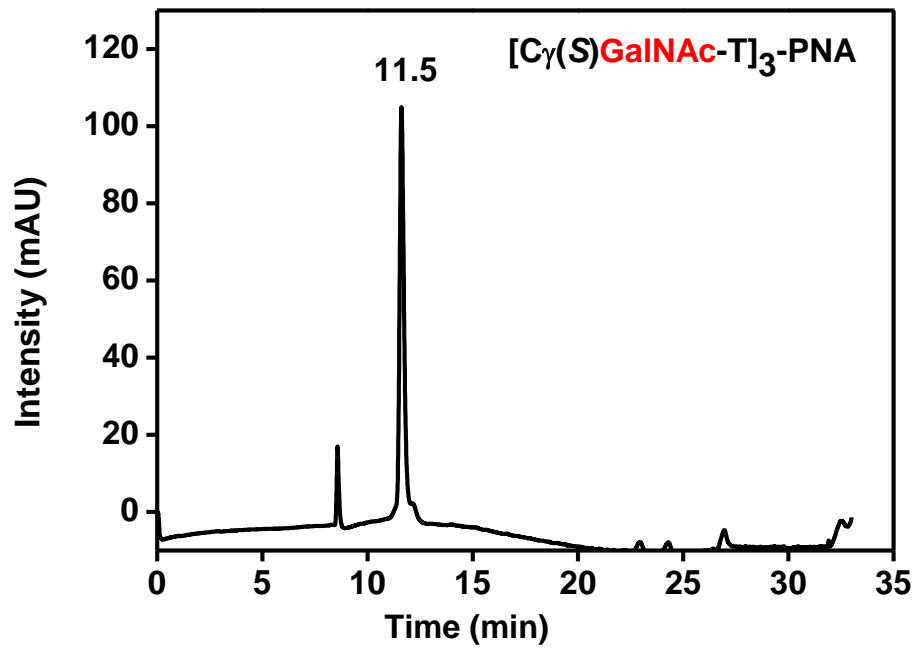
HPLC chromatogram of PNA 1



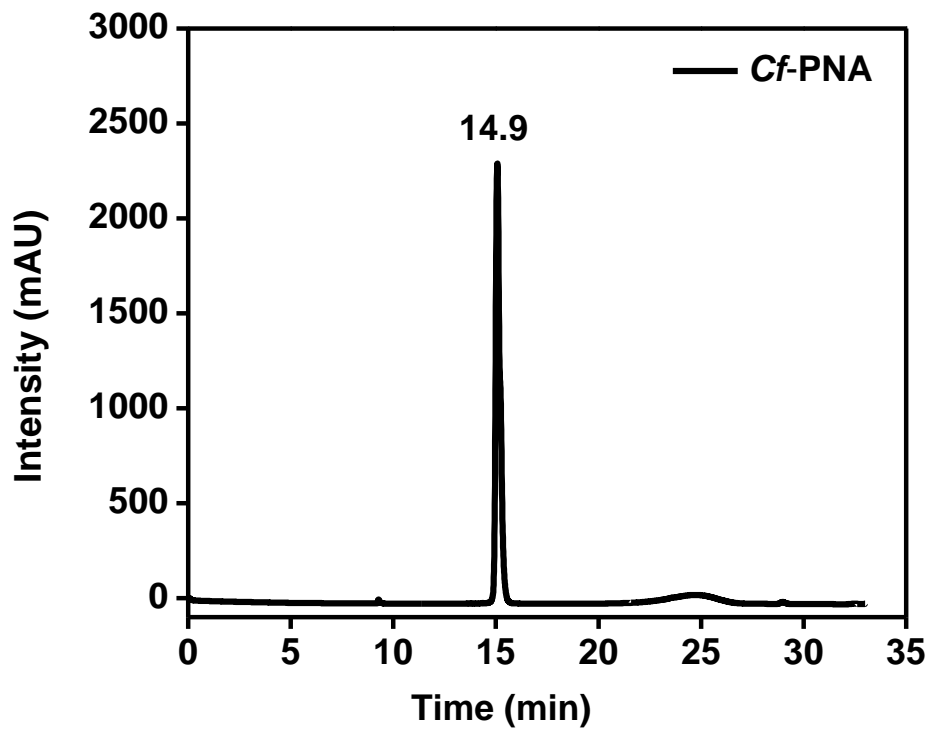
HPLC chromatogram of PNA 2



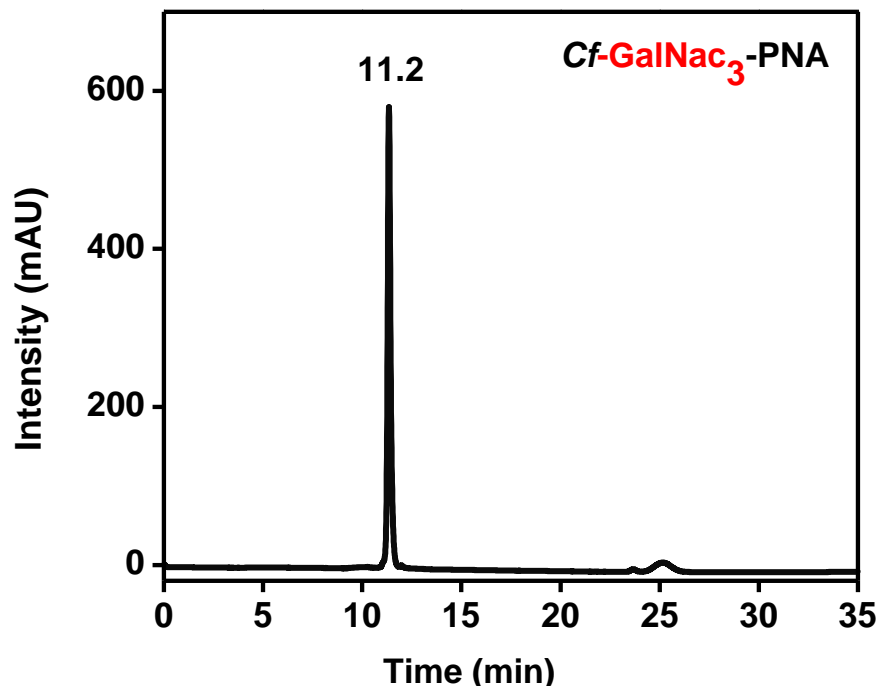
HPLC chromatogram of PNA 3



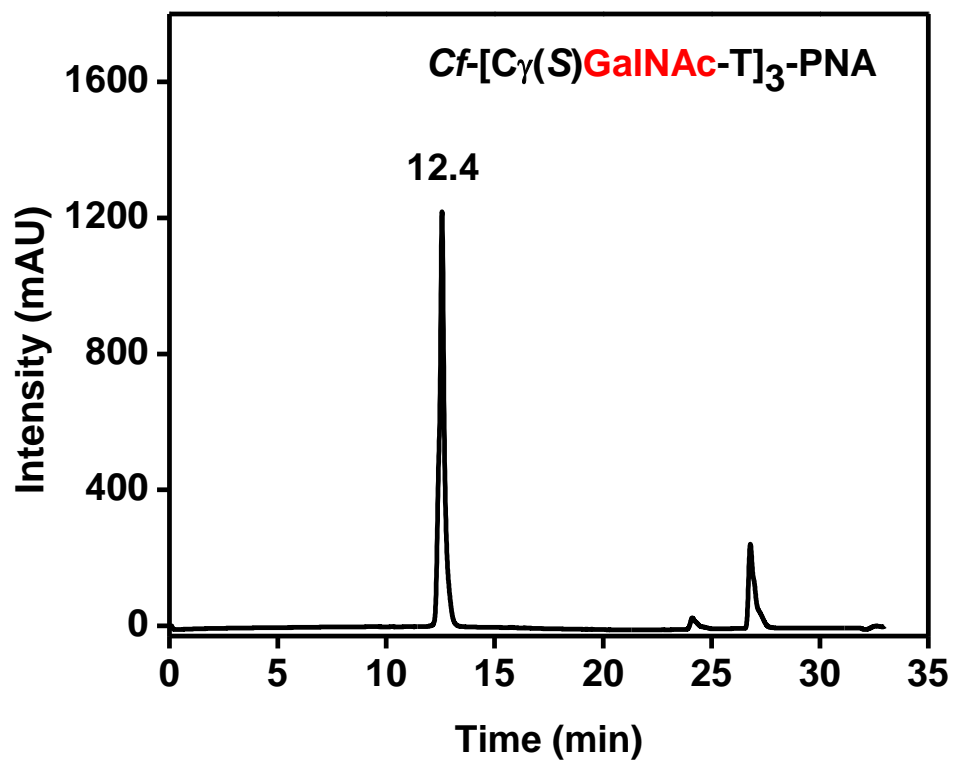
HPLC chromatogram of PNA 4



HPLC chromatogram of PNA 5

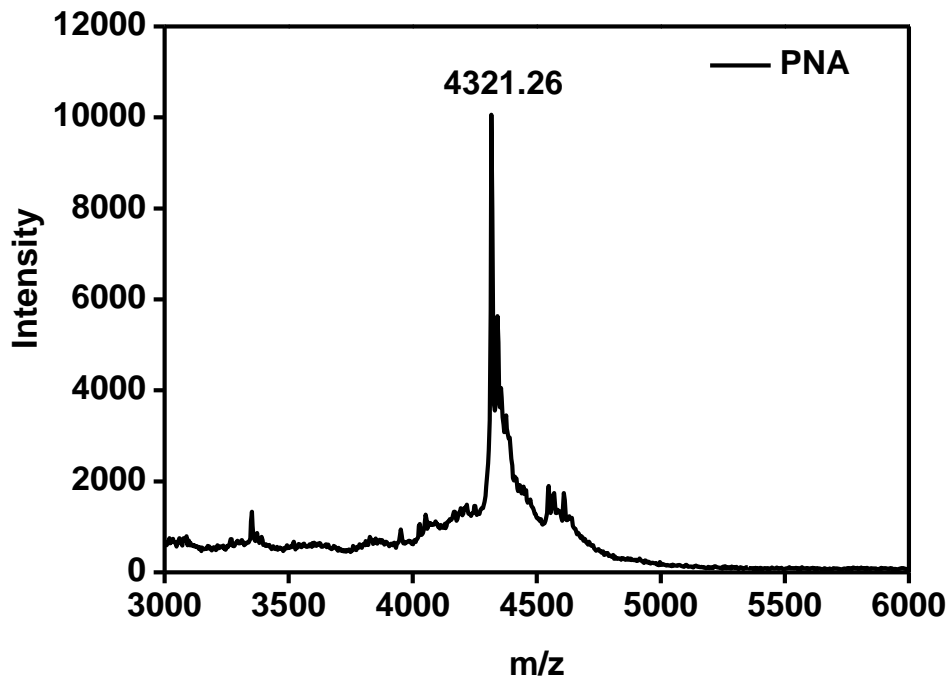


HPLC chromatogram of PNA 6

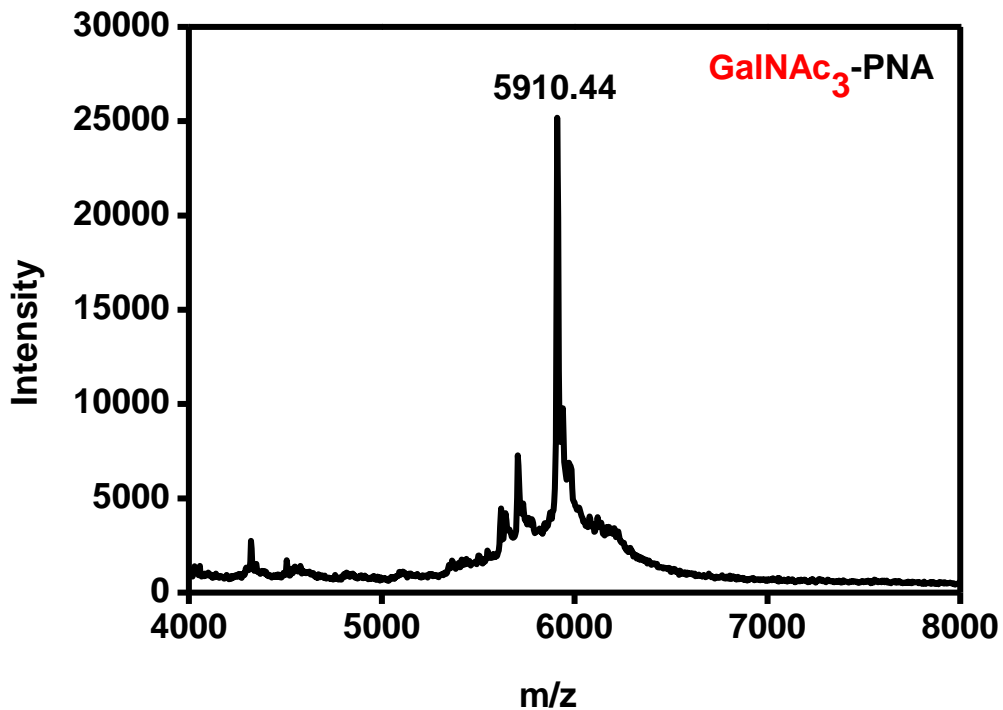


MALDI-TOF Spectra of Peptides:

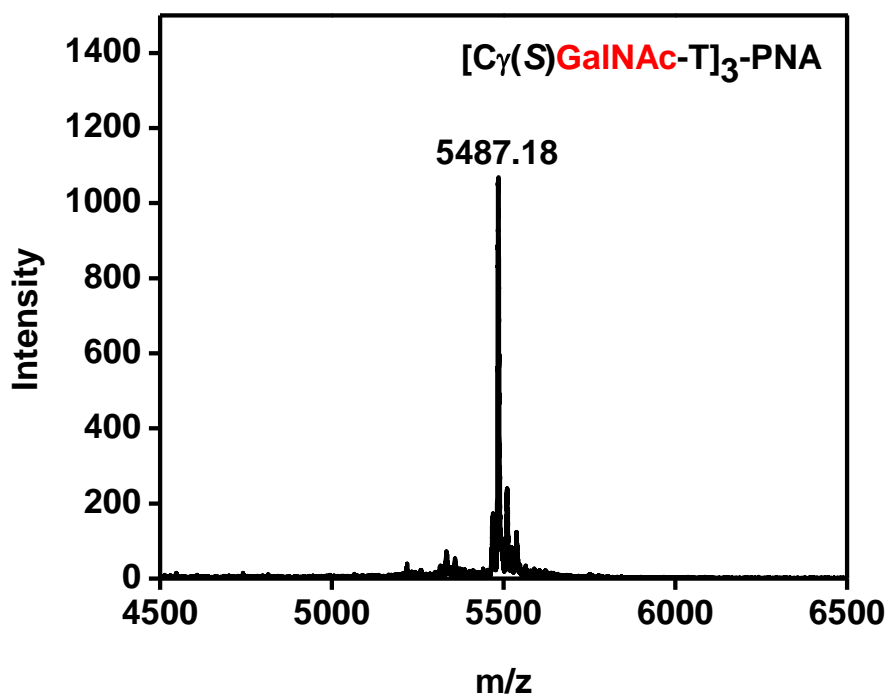
MALDI-TOF Spectra of PNA 1



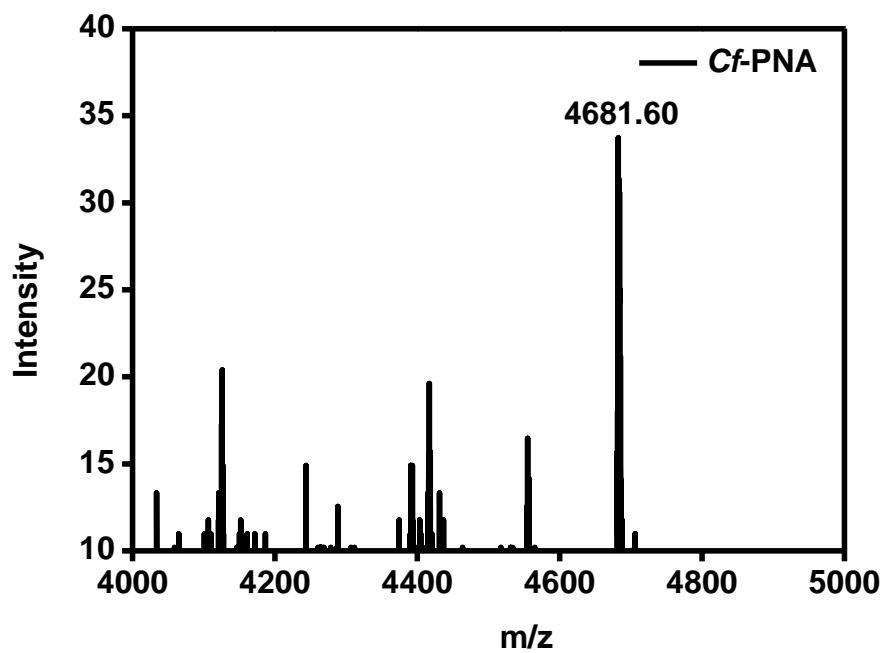
MALDI-TOF Spectra of PNA 2



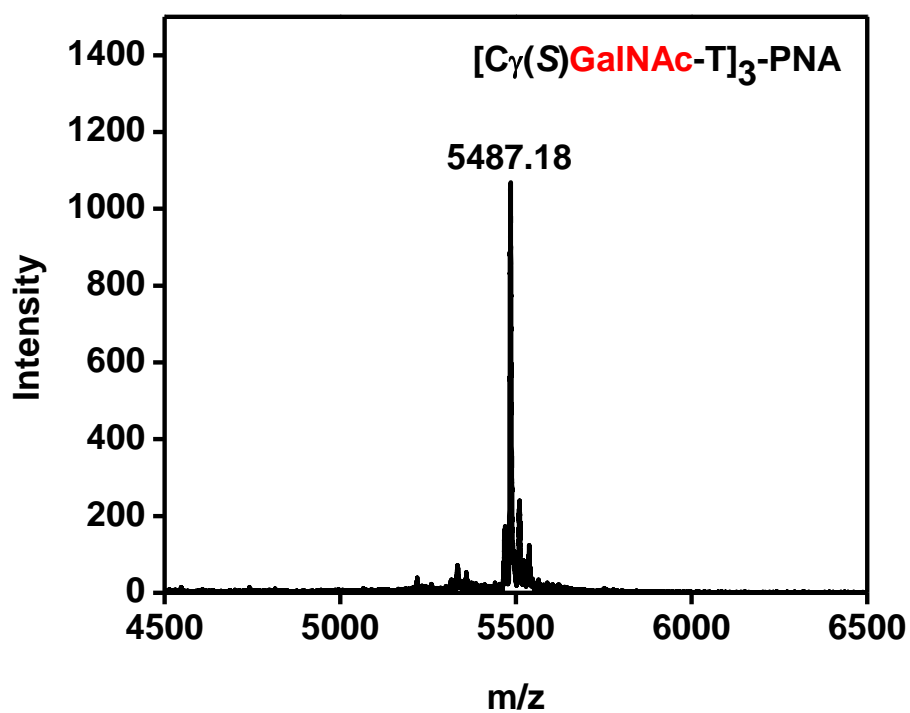
MALDI-TOF Spectra of PNA 3



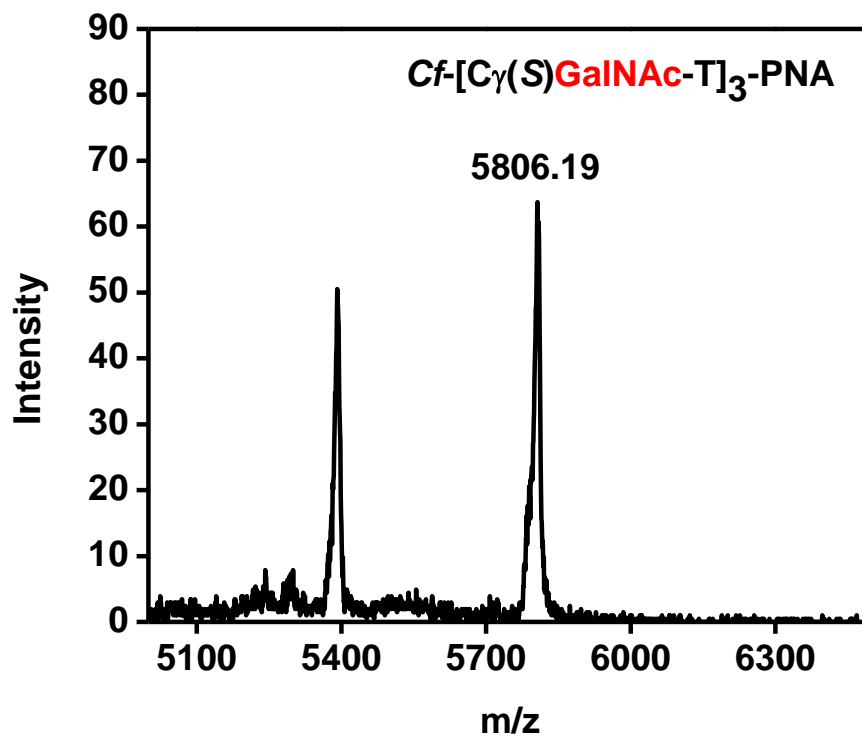
MALDI-TOF Spectra of PNA 4



MALDI-TOF Spectra of PNA 5



MALDI-TOF Spectra of PNA 6



2.8 References

1. Morrissey, D. V. *et al. Nature Biotechnol.* **2005**, *23*, 1002–1007.
2. Okumura, A.; Pitha, P. M.; Harty, R. N. *Proc. Natl Acad. Sci. USA.* **2008**, *105*, 3974–3979.
3. Whitehead, K. A.; Langer, R.; Anderson, D. G. *Nature Rev. Drug Discov.* **2009**, *8*, 129–138.
4. Alexis, F.; Pridgen, E.; Molnar, L. K.; Farokhzad, O. C. *Mol. Pharm.* **2008**, *5*, 505–515.
5. Petros, R. A.; Desimone, J. M. *Nature Rev. Drug Discovery*, **2010**, *9*, 615–627.
6. Kanasty, R. L.; Whitehead, K. A.; Vegas, A. J.; Anderson, D. G. *Mol. Ther.* **2012**, *20*, 513–524.
7. Nair, J. K.; Willoughby, J. L.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Kandasamy, P.; Kel'in, A. V.; Milstein, S.; Taneja, N.; O'Shea, J.; Shaikh, S.; Zhang, L.; van der Sluis, R. J.; Jung, M. E.; Akinc, A.; Hutabarat, R.; Kuchimanchi, S.; Fitzgerald, K.; Zimmermann, T.; van Berkel, T. J.; Maier, M. A.; Rajeev, K. G.; Manoharan, M. *J. Am. Chem. Soc.* **2014**, *136*, 16958-16961.
8. Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.* **2014**, *42*, 8796-8807.
9. Stockert, R. J. *Physiol. Rev.* **1995**, *75*, 591-609.
10. Steirer, L. M.; Park, E. I.; Townsend, R. R.; Baenziger, J. U. *J. Biol. Chem.* **2009**, *284*, 3777-3783.
11. Grewal, P. K.; Uchiyama, S.; Ditto, D.; Varki, N.; Le, D. T.; Nizet, V.; Marth, J. D. **2008**, *14*, 648-655.
12. Nielsen, P. E. *Peptide Nucleic Acids Protocols and Applications*, 2nd edition. **2004**, 1-5.
13. Jacquinet, J. C.; Zurabyan, S. E.; Khorlin, A. Y. *Carbohydr. Res.* **1974**, *32*, 137.
14. Drouillard, S.; Armand, S.; Davies, G. J.; Vorgias, C. E.; Henrissat, B. *Biochem. J.* **1997**, *328*, 945.
15. Guo, J.; Ye, X.-S. *Molecules.* **2010**, *15*, 7235.
16. Leo, A.J. *et.al. J. Med. Chem.* **1999**, *42*, 609-618.
17. Garth, et. al. *J. Am. Chem. Soc.*, **2002**, *124*, 1562-15163.
18. Krapcho, A. P.; Kuell, C. S. *Synth. Commun.* **1990**, *20*, 2559-2564.
19. Jones, S. P.; Pavan, G. M.; Danani, A.; Pricl, S.; Smith, D. K. *Chem. - Eur. J.* **2010**, *16*, 4519.
20. Larionov, O. V.; Meijere, A. *Org.Lett.* **2004**, *13*, 2153-2156.
21. (a) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595-598 (b) Kaiser, E.; Bossinger, C. D.; Cpllescott, 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.

Chapter 3

Biophysical Evaluation and Cell Permeation Studies of GalNAc₃-PNA Oligomers

3.1 Introduction

In order to use PNAs as antisense and antigene oligonucleotides, their biophysical properties and ability to penetrate into target cells need to be studied. Several biophysical techniques are used to understand the properties of the molecules under physiological conditions. This allows us to fine tune their chemical structure for specific functions.¹ The ability of antisense/antigene oligonucleotides to bind *in vitro* to the target DNA/RNA can be investigated using techniques such as thermal UV-melting (UV- T_m), circular dichroism (CD), fluorescence spectroscopy and cell permeation properties of the oligomers can be studied using confocal microscopy.

The preceding chapter dealt with the synthesis of rationally designed triantennary GalNAc₃ and trivalent [C^γ(S)-GalNAc-T]₃ substituted multifunctional PNA analogs. The triantennary GalNAc₃ and trivalent [C^γ(S)-GalNAc-T]₃ moieties were incorporated at the N-terminus of the PNAs by solid phase synthesis. The triantennary GalNAc₃ sequence is known to bind to asialoglycoprotein receptor (ASGPR), which is specific to hepatocytes, enabling its internalization. The GalNAc₃-PNA tagged with carboxyfluorescein have been studied for their ability to get internalized selectively by hepatocytes. Experiments are also reported with cells that do not express asialoglycoprotein as negative controls.

3.2 Rationale for the present work

Incorporation of three GalNAc moieties in PNA was intended to aid the selective uptake of PNAs into hepatocytes that exhibit ASGPR on their surface. Since the utility of GalNAc₃ modified PNAs depends on whether or not they form stable PNA:RNA or PNA:DNA hybrids under physiological conditions, the thermal stability of the hybrids formed by the modified PNAs with complementary nucleic acids were determined using temperature dependent UV spectroscopy. To examine perturbations in the conformation of the duplexes, the CD spectroscopy studies were carried out. For the carboxyfluorescein tagged PNAs, their fluorescence emission characteristics were studied prior to their use in the cell internalization studies using confocal laser scanning microscopy (CLSM).

The biophysical studies have been carried out on the designed PNA analogs in order to investigate the effective conjugation of side chain functional groups on the binding affinity to target nucleic acids. These studies provide information about the binding selectivity, binding specificity, structural organization and the duplex stability, which are important parameters to evaluate their potential as antisense/antigene agents.

3.3 Objectives of the present work

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

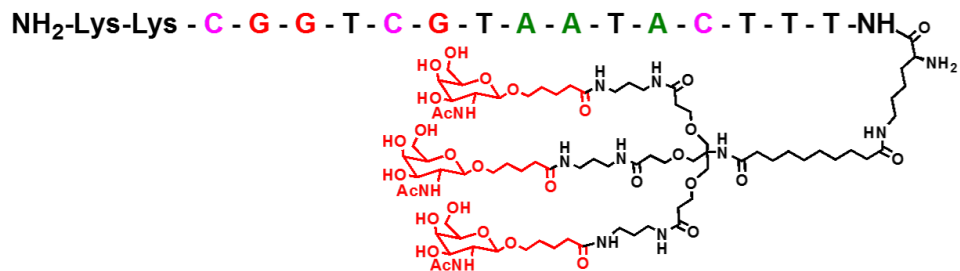
- Thermal stability, binding selectivity, specificity and discrimination of control *aeg* PNA as well as modified PNA towards complementary DNA using temperature-dependent UV spectroscopy
- Characterization of the PNA:DNA duplex structures by CD spectroscopy
- Study of fluorescence properties of single stranded fluorescent PNA and derived duplexes with complementary DNA
- UV-visible and fluorescence studies of tagged fluorescent PNA oligomers
- Investigation of the uptake efficiency of modified and control *aeg* PNA oligomers into HEK 293 and Hep-G2 cell lines by confocal microscopy

3.3.1 GalNAc₃-PNA oligomers used for biophysical studies

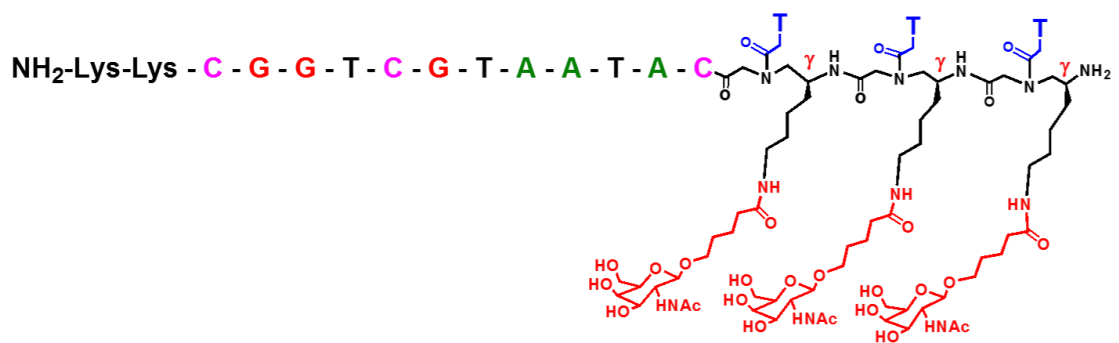
The triantennary (GalNAc)₃ and trivalent C^γS-substituted GalNAc [C^γ(S)-GalNAc-T] moieties were incorporated at the N-terminus of *aeg* PNA oligomers by solid-phase synthesis (Figure 3.5 A, B, C) followed by coupling with carboxyfluorescein reaction to obtain carboxyfluorescein tagged PNA oligomers (Figure 3.1 D, E, F) and used for the biophysical studies with complementary antiparallel DNAs (Table 3.1).



A) PNA



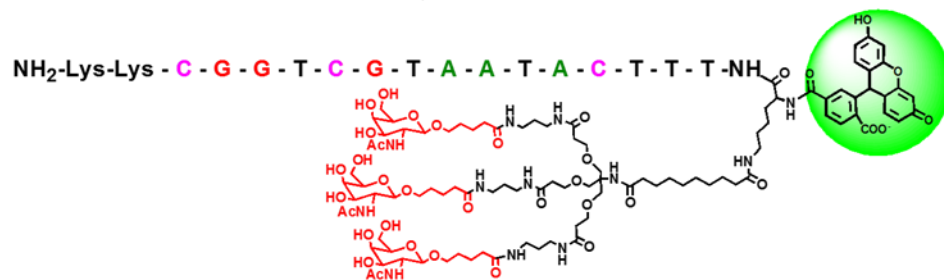
B) GalNAc₃-PNA



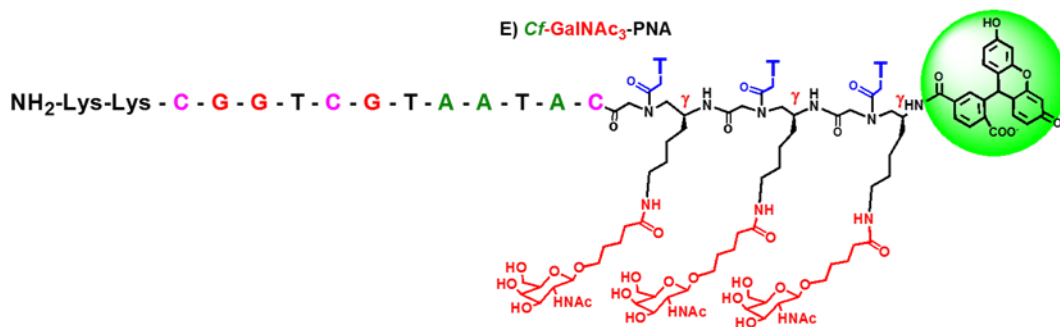
C) [C^γ(S)-GalNAc-T]₃-PNA



D) Cf-PNA



E) Cf-GalNAc₃-PNA



F) Cf[C^γ(S)-GalNAc-T]₃-PNA

Figure 3.1 Different types of GalNAc₃-PNA and control PNA oligomers

3.3.2 DNA/RNA oligonucleotides used for biophysical studies

The DNA and RNA oligonucleotides were obtained from commercial sources and used without further purification in the biophysical studies of PNAs.



Figure 3.2 DNA and RNA oligonucleotide

Control PNA, GalNAc₃-PNA and [C^γ(S)-GalNAc-T] (Figure 3.1) sequences were hybridized with appropriate complementary DNA/RNA (Figure 3.2) in antiparallel orientation to generate PNA:DNA/RNA duplexes (Figure 3.3).

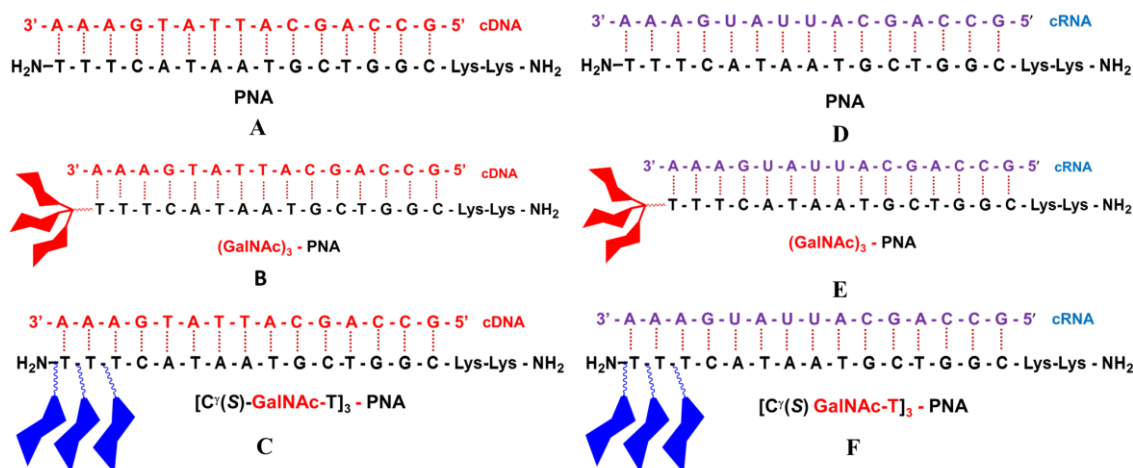


Figure 3.3 (A) Control PNA:cDNA, (B) GalNAc₃-PNA:cDNA, (C) [C^γ(S)-GalNAc-T]₃-PNA:cDNA and (D) Control PNA:cRNA, (E) GalNAc₃-PNA:cRNA, (F) [C^γ(S)-GalNAc-T]₃-PNA:cRNA

3.4. Result and Discussion

The interaction of modified and unmodified PNAs with DNA or RNA can be investigated by changes in UV, CD and fluorescence spectra upon complexation. The site-specific effects of introduction of modified PNA units at various positions in a 15-mer *aeg* PNA sequence on their duplex formation are studied by these techniques. Thermal stabilities of PNA:DNA or PNA:RNA hybrids can be examined by temperature dependent UV absorption spectroscopy. The effect of GalNAc modification on the conformation of the duplexes formed by GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA with cDNA or cRNA has been studied by CD spectroscopy.

3.4.1 UV melting studies of GalNAc₃-PNA:DNA/RNA hybrids

The hybridization studies of modified and control PNA oligomers with their complementary DNA and complementary RNA were carried out by temperature dependent changes in UV-absorbance at 260 nm. The T_m values were obtained from mid-points of the thermal stabilities of various modified PNAs.

3.4.1a Thermal stability of PNA, Triantennary GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃ PNA:DNA duplexes

In order to investigate the effect of GalNAc₃ modifications on the thermal stability of PNA:DNA hybrids, the modified PNA oligomers were hybridized with complementary DNA in antiparallel orientation (Figure 3.4). The PNA:DNA duplex derived from unmodified *aeg*-PNA showed a melting (T_m) of 48.9 °C, while the T_m of triantennary GalNAc₃-PNA:DNA duplex showed enhanced stabilization 52.5 °C and [C^γ(S)-GalNAc-T]₃-PNA:DNA showed T_m 51.3 °C (Figure 3.4). GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA duplexes thus showed stabilization of corresponding duplexes by 3.6 °C and 2.4 °C respectively (Table 3.1).

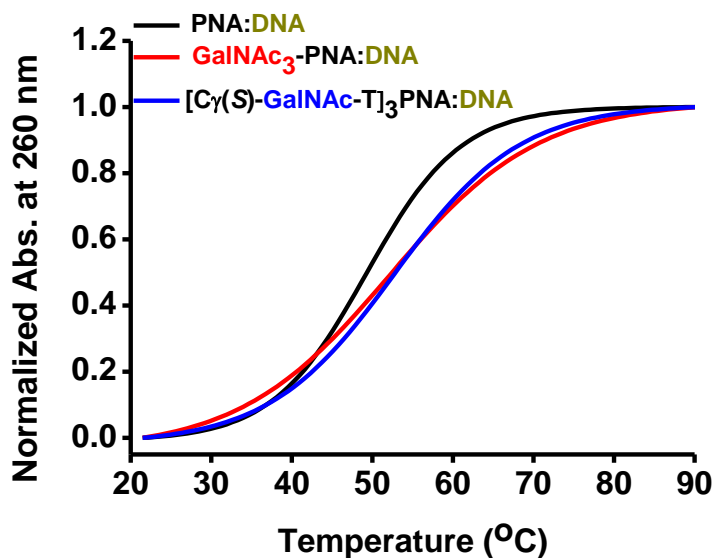


Figure 3.4 Melting curves for *aeg* PNA, triant-GalNAc₃-PNA:DNA and [C^γ(S)-GalNAc-T]₃-PNA:DNA duplexes; (DNA=5'GCCAGCATTATGAAA 3'; Buffer: 10mM sodium cacodylate, pH 7.2, NaCl 10 mM).

Table 3.1. UV- T_m values of complementary PNA:DNA duplexes with triantennary-GalNAc₃ PNA and trivalent- [C^γ(S)-GalNAc-T]₃-PNA units

Entry	PNA Sequence	5'GCCAGCATTATGAAA3	
		UV- T_m (°C)	ΔT_m (°C)
1	H-TTTCATAATGCTGGC-Lys-Lys-NH ₂	48.9	-
2	GalNAc ₃ -TTTCATAATGCTGGC-Lys-Lys-NH ₂	52.5	+ 3.4
3	[C γ (S)GalNAc-T] ₃ -TTTCATAATGCTGGC-Lys-Lys-NH ₂	51.3	+ 2.4

3.4.1b. Thermal stability of PNA, triantennary GalNAc₃-PNA and [C γ (S)-GalNAc-T]₃-PNA:RNA duplexes

To study the effect of GalNAc₃ modifications on the stability of PNA:RNA hybrids, the PNA oligomers were hybridized with complementary RNA in antiparallel orientation (Figure 3.5). The PNA:RNA duplex derived from unmodified *aeg*-PNA showed a melting (T_m) of 59.5 °C. The T_m of triantennary GalNAc₃-PNA:RNA duplex was of 55.8 °C, whereas the T_m of [C γ (S)-GalNAc-T]₃-PNA:RNA duplex was 53.8 °C (Figure 3.5). In contrast to the results obtained from the PNA:DNA hybrids where both GalNAc₃ conjugated PNAs stabilized duplexes over unmodified duplexes, the corresponding PNA:RNA hybrids were destabilized. The triantennary GalNAc₃-PNA destabilized the derived RNA duplex by -3.7 °C, while the [C γ (S)-GalNAc-T]₃-PNA destabilized the RNA duplex by -5.7 °C. The stability of the PNA:RNA duplexes were in the order PNA > GalNAc₃-PNA > [C γ (S)-GalNAc-T]₃-PNA (Table 3.2). The destabilization of PNA:RNA hybrids induced by GalNAc₃ modification was only marginal and the melting temperatures were still higher than the corresponding PNA:DNA hybrids.

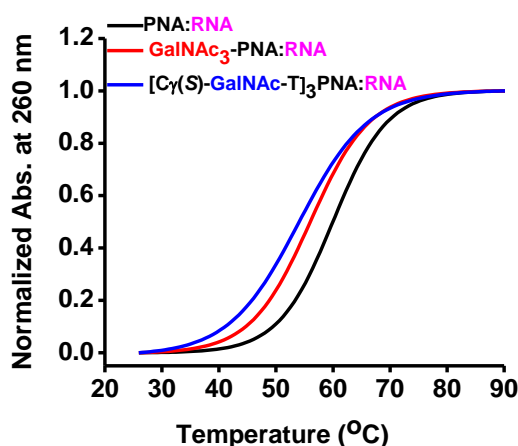


Figure 3.5 Melting curves for PNA, GalNAc₃-PNA:RNA and [C γ (S)-GalNAc-T]₃-PNA:RNA duplexes; (RNA=5'GCCAGCAUUAUGAAA3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM).

Table 3.2. UV- T_m values GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA:RNA duplexes

Entry	PNA Sequence	5'GCCAGCAUUAUGAAA3'	
		UV- T_m (°C)	ΔT_m
1	H-TTTCATAATGCTGGC-Lys-Lys NH ₂	59.5	-
2	GalNAc ₃ -TTTTCATAATGCTGGC-Lys-Lys NH ₂	55.8	-3.7
3	[C ^γ (S)-GalNAc-T] ₃ -CATAATGCTGGC-Lys-Lys NH ₂	53.8	-5.7

3.4.2. CD Studies of PNA:DNA and PNA:RNA duplexes

The effect of conjugation of GalNAc₃ to PNA and GalNAc₃ as side chain substitution at C^γ of thymine on PNA on the conformation of PNA:DNA/RNA duplexes was studied by CD spectroscopy. The CD-spectra for *ss*PNA (A, B and C), *ss*DNA, and *ss*RNA were recorded. The CD spectra of single stranded DNA and RNA show a positive band between 200 to 220 nm along with a positive band in the region of 270 to 290 nm. The single stranded aeg PNA being achiral does not show any characteristic CD signatures. However, being a polyamide, it can form either left or right handed coil structures with equal facility (Figure 3.6). While the triantennary GalNAc₃ modified showed no characteristic CD signature as the control aeg-PNA, [C^γ(S)-GalNAc-T]₃-PNA showed a negative band at 275 nm. The appearance of the band indicates that the chiral induction from GalNAc₃ moiety at C^γ has an effect on the nucleobase stacking arrangement.

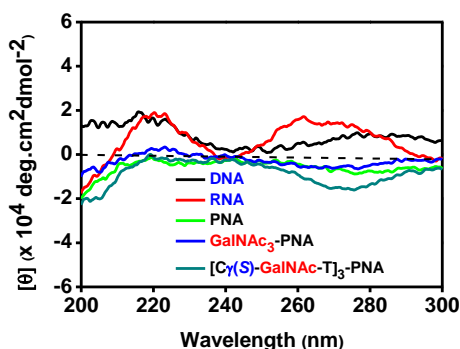


Figure 3.6 CD signatures of *ss*DNA, *ss*RNA, *ss*PNAs (Buffer: 10 mM sodium cacodylate, pH=7.2, NaCl 10 mM).

The CD-spectra of control PNA:DNA, GalNAc₃-PNA:DNA and [C^γ(S)-GalNAc-T]₃-PNA:DNA duplexes showed weak positive bands at 222 nm and strong positive band at 275 nm, and a weak negative band at 250 nm (Figure 3.7A). In case of RNA corresponding duplexes, positive band at 220 nm was relatively strong, accompanied by strong positive band at 265 nm and weaker negative band at 245 nm (Figure 3.7 B).

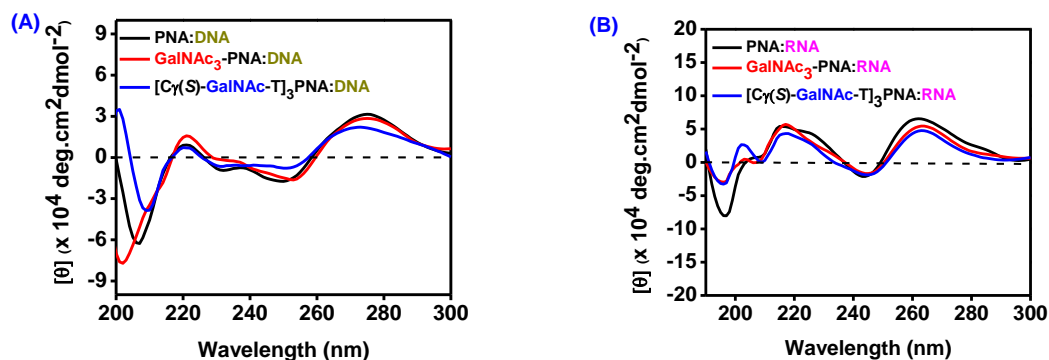


Figure 3.7 CD signatures of PNA:DNA duplexes (A) and PNA:RNA duplexes (B) formed by *aeg*-PNA, GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA. (Buffer: 10 mM sodium cacodylate, pH=7.2, NaCl 10 mM).

All the PNA:DNA and PNA:RNA duplexes showed similar CD spectra suggesting that GalNAc₃ modifications does not perturb the PNA:DNA/RNA conformation.

3.4.3 UV-visible and fluorescence studies of *Cf*-PNA, *Cf*-GalNAc₃-PNA and *Cf*[C^γ(S)-GalNAc-T]₃-PNA and their duplexes

As discussed in Chapter 2 (Figure 2.4), 5(6)-carboxyfluorescein was coupled at the N-terminus of control PNA, GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA under acid-amine coupling reaction conditions to obtain corresponding fluorescent PNAs (Figure 3.1 D, E and f). The UV absorption spectra were recorded for single stranded PNAs. The spectra showed absorbance maxima at 260 nm due to nucleobases whereas lower intensity broad absorbance with maxima at 458 nm and 484 nm are indicative of presence of 5(6)-carboxyfluorescein in the PNA oligomers (Figure 3.8).

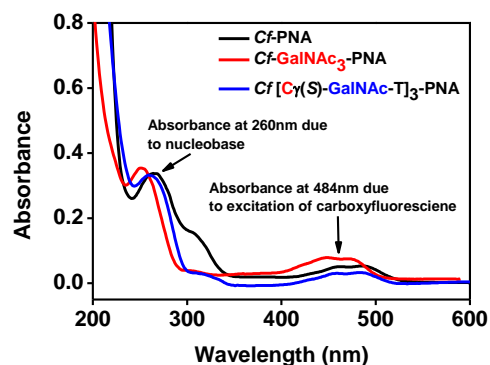


Figure 3.8 UV-Visible absorption spectra of control PNA, GalNAc₃-PNA and [C^γ(S)-GalNAc-T]₃-PNA

The fluorescence spectra of single stranded *Cf*-PNA, *Cf*-GalNAc₃-PNA and *Cf*-[C^γ(S)-GalNAc-T]₃-PNA and their corresponding duplexes with cDNA were recorded (Figure 3.9 A and B). The fluorescent PNAs upon excitation at 458 nm and 484 nm showed the emission spectra with maximum at 520 nm. The fluorescence intensity of single stranded fluorescent PNAs as well as their complexes with complementary DNA excited at 484 nm was approximately double than that of excitation at 458 nm. The fluorescent intensity of the duplexes were slightly lower than that of single stranded fluorescent PNAs when excited at either at 458 nm or 484 nm. However, the changes observed in the fluorescent intensities among the different single stranded fluorescent PNAs or their duplexes were not significant.

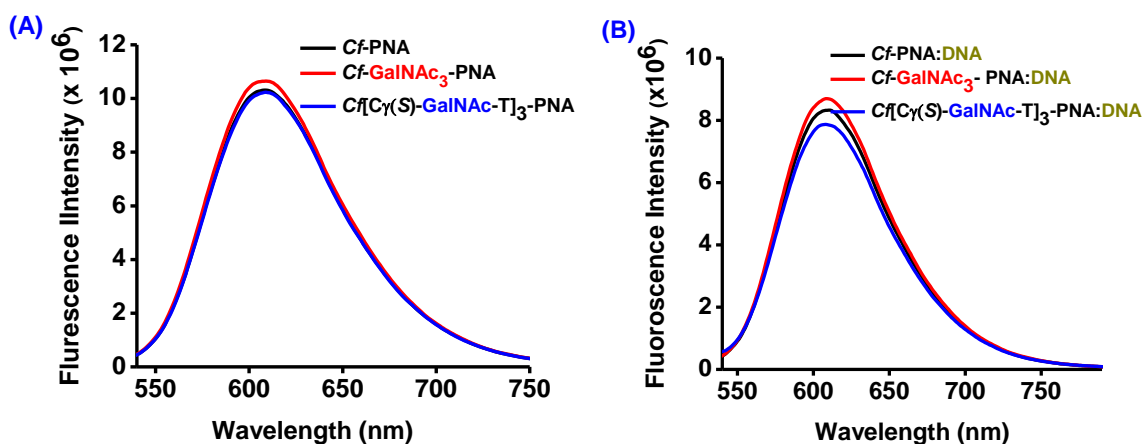


Figure 3.9 Fluorescence spectra of (A) *Cf*-PNA, *Cf*-GalNAc₃-PNA and *Cf*-[C^γ(S)-GalNAc-T]₃-PNA and (B) Duplexes of *Cf*-PNA, *Cf*-GalNAc₃-PNA and *Cf*-[C^γ(S)-GalNAc-T]₃-PNA with complementary DNA (Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM; $\lambda_{\text{excitation}} = 484$ nm, emission $\lambda_{\text{max}} = 520$ nm).

3.5 Cellular uptake studies

The GalNAc₃ conjugated PNAs were designed for selective uptake by hepatocytes. To examine their internalization in hepatocytes, HepG2 cells which have asialoglycoprotein on their cell surfaces, were treated with 5(6)-carboxyfluorescein tagged *Cf*-GalNAc₃-PNA and *Cf*-[C^γ(S)-GalNAc-T]₃-PNA and their uptake was studied by confocal microscopy. In these experiments, 5(6)-carboxyfluorescein tagged *aeg*-PNA of the same sequence was used as control. To check the selectivity of uptake, the experiments were conducted with HEK-293 as negative control.

HEK293 cell line is derived from human embryonic kidney cells that were grown in tissue culture. They are widely used for transfection experiments because they are easy

to culture and exhibit high propensity for transfection. In this study, HEK293 cells were used as a negative control because they lack asialoglycoprotein receptors, that are required for internalization of GalNAc₃ containing PNAs.

HepG2 cell line is derived from liver cells of a patient diagnosed with hepatocellular carcinoma. They are used as models for investigating various intracellular trafficking processes and for studying liver metabolism and xenobiotics toxicity. They express ASGPRs in large amounts. The presence of ASGPRs on the surface of HepG2 in large numbers make them suitable models for the cellular uptake studies to evaluate selective uptake of GalNAc₃ modified PNAs into liver.

3.5.1 Cellular uptake experiment using confocal microscopy

3.5.1a Cellular uptake studies with HepG2 cells. To examine whether GalNAc₃ modification enables PNAs to be internalized into hepatocytes, HepG2 cells were grown in MEM medium at 37°C under 5% CO₂ to reach a confluency of 80-90%. They were then treated with 5(6)-carboxyfluorescein functionalized PNAs *Cf*-GalNAc₃-PNA, *Cf*-[C^γ(S)-GalNAc-T]₃-PNA and control unmodified *Cf*-PNA each in 4 μM concentrations.

After incubation for 24 h, the cells were washed to remove excess PNAs in the medium and the internalized PNAs were visualized by imaging the cells with confocal laser scanning microscopy. Prior to imaging, 4',6-diamidino-2-phenylindole (DAPI) was added to stain the nuclei. The images of HepG2 cells following treatment with PNAs were obtained with various channels. The panels, A, C, and E of Figure 3.10 correspond to the DAPI emission signals showing the nuclei of the cells. The panels, B, D, and F of the Figure 3.10 show the merged image of the DAPI and carboxyfluorescein channels along with differential interference contrast (DIC) image.

In case of *Cf*-GalNAc₃-PNA treated cells, the presence of green fluorescence indicated that the PNAs were taken up by HepG2 cells. This is evident from Figure 3.10 A-B, which show the nuclei and the cell boundary along with green punctas that appear within the cells. Similar results were obtained with *Cf*-[C^γ(S)-GalNAc-T]₃-PNA. However, the extent of internalization of *Cf*-[C^γ(S)-GalNAc-T]₃-PNA was higher than that exhibited by *Cf*-GalNAc₃-PNA. The localization of green fluorescence within the

cell was clearly observed in Figure 3.10 D. The appearance of discrete punctas in the vicinity of the nucleus and distributed throughout the cell suggested that the internalization could be due to endocytosis, as expected for ASGPR mediated cellular uptake. Occasionally, there were also a large concentration of the punctas. In contrast to the GalNAc₃ modified PNAs, the unmodified *Cf*-PNA did not show any green fluorescence in the cells, indicating no internalization. These results showed that the cell penetrating ability of the GalNAc₃ modified PNAs was indeed conferred by the GalNAc₃ functionalization.

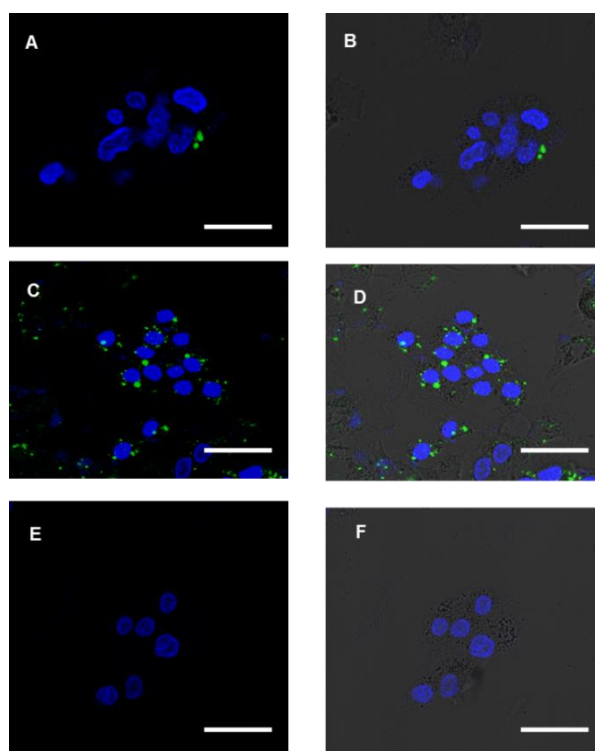


Figure 3.10. Confocal microscopy images of HepG2 cells treated with *Cf*-GalNAc₃-PNA (A-B), *Cf*-[C^γ(S)-GalNAc-T]₃-PNA (C-D) and *Cf*-PNA (E-F). A, C, and E are merged images of DAPI and carboxyfluorescein channels. B, D, and F shows the overlay of the merged images of DAPI and carboxyfluorescein channels with differential interference contrast image showing the intact cells (gray), nuclei (blue), and internalized PNAs (green). Scale bars represent 25 μm.

3.5.1b Cellular uptake studies with HEK-293 cells. The HEK-293 cells in which asialoglycoprotein receptors (ASGPR) are not expressed were used as negative controls to confirm that observed internalization of GalNAc₃ modified PNAs was indeed mediated by specific interactions with the cell surface receptor ASGR. The cells grown in MEM medium at 37 °C under 5% CO₂ to reach 80-90% confluency. They were then treated with *Cf*-GalNAc₃-PNA, *Cf*-[C^γ(S)-GalNAc-T]₃-PNA or control unmodified *Cf*-

PNA (Figure 3.11). The concentration of the PNAs were maintained at 4 μM , as in the experiment with HepG2 cells.

After incubating the HEK-293 cells with the PNAs at 37 °C for 24 h, the cells were imaged under confocal laser scanning microscopy. Prior to imaging, the nuclei of the cells were stained with DAPI. Figure 3.11 shows the confocal images of HEK-293 cells following treatment with PNAs obtained from various channels. The panels, A, D, and G correspond to the DAPI signals indicating the nuclei of the cells. The panels, B, E, and H shows the image obtained from the detector capturing carboxyfluorescein emission signals. The panels C, F, and I show the merged image of the DAPI and carboxyfluorescein channels along with differential interference contrast image.

In case of *Cf*-GalNAc₃-PNA treated cells, no green fluorescence emission was observed. This is evident from Figure 3.11 C, which shows the nuclei and the cell boundary but no PNA within the cells. Similar results were obtained with *Cf*-[*C*'(*S*)-GalNAc-T]₃-PNA and unmodified *Cf*-PNA as well. These results indicate that these PNAs, with or without *Cf*-PNA modifications, are incapable of cell penetration.

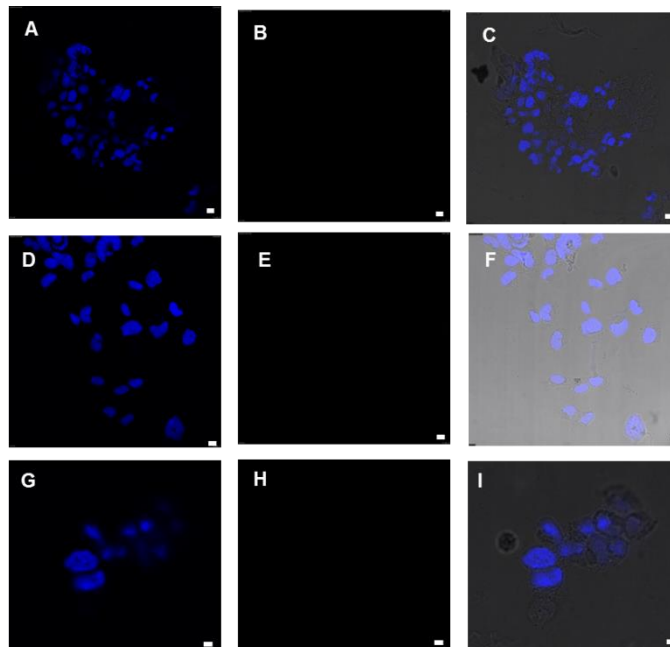


Figure 3.11 Confocal microscopy images of HEK-293 cells treated with *Cf*-GalNAc₃-PNA (A-B), *Cf*-[*C*'(*S*)-GalNAc-T]₃-PNA (C-D) and *Cf*-PNA (E-F). A, D, and G are merged images of DAPI. B, E, and H show signals from carboxyfluorescein channel. C, F, and I show the merged image of DAPI and carboxyfluorescein channels along with differential interference contrast image showing the intact cells (grey), nuclei (blue), and internalized PNAs (green). Scale bar represents 8 μm .

3.5.1c Flow cytometry studies. For quantitative comparison of the uptake of the GalNAc modified PNAs by HepG2 cells flow cytometry studies were carried out. HepG2 cells were cultured as mentioned above. The cells were treated with 5(6)-carboxyfluorescein functionalized PNAs *Cf*-GalNAc₃-**PNA 5**, *Cf*[C^γ(S)-GalNAc-T]₃-**PNA 6** or unmodified control *Cf*-**PNA 4** each at 4 μM concentrations for 12 hours. The results show that compared to untreated cells and those that were incubated with unmodified control *Cf*-**PNA 4**, the cells that were treated with GalNAc modified PNAs showed significant uptake (Figure 3.12). Similar to the results of the confocal microscopy studies, the cells were found to internalize *Cf*[C^γ(S)-GalNAc-T]₃-**PNA 6** more than *Cf*-GalNAc₃-**PNA 5**. While the measured uptake was only 3.1% for triantennary *Cf*-GalNAc₃-**PNA 5**, it was 39.5% for the *Cf*[C^γ(S)-GalNAc-T]₃-**PNA 6**. Thus, conjugation of three GalNAc moieties to the last three PNA monomeric residues at the N terminus was found to effect a 12.7 fold increase in the uptake of the PNA by HepG2 cells compared to the triantennary GalNAc modification.

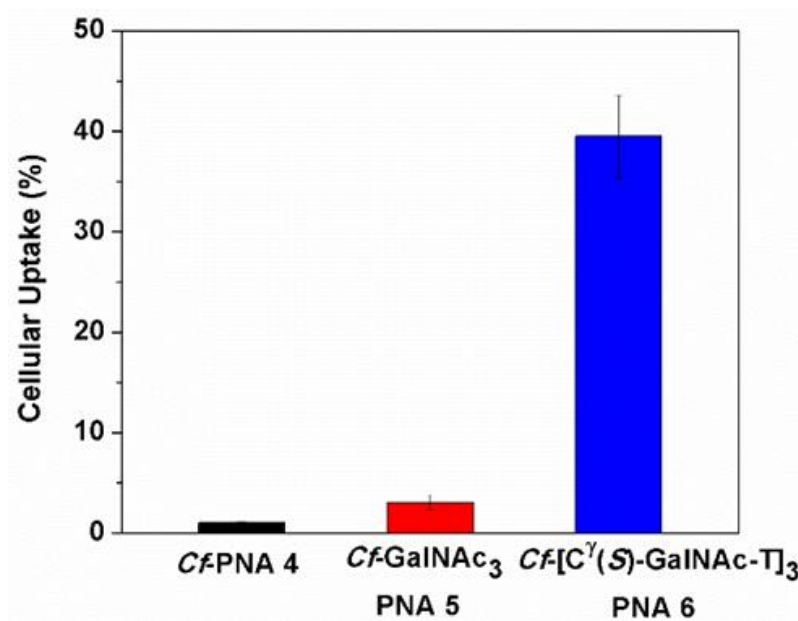


Figure 3.12 Uptake of unmodified control *Cf*-**PNA 4**, *Cf*-GalNAc₃-**PNA 5**, and *Cf*[C^γ(S)-GalNAc-T]₃-**PNA 6** by HepG2 cells as measured by flow cytometry. The results shown are average of two independent measurements. The error bars represent standard deviations.

3.5.1d Cytotoxicity studies. For ultimate therapeutic application of the GalNAc modified PNAs, it is essential to confirm whether the modified PNAs elicit any adverse effects in the target cells. We assessed the cytotoxicity of the GalNAc modified PNAs

GalNAc₃-PNA **2** and [C^γ(S)-GalNAc-T]₃-PNA **3** towards HepG2 cells and compared them with the unmodified control PNA **1** using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Figure 3.13). The cells were incubated with PNAs at the concentration of 4 μM for 12 hours, as in the cellular uptake studies, prior to treatment with MTT. The results show that the viability of the cells were not significantly altered by the treatment of the cells with the PNAs. The viability of the cells were marginally reduced when they were treated with GalNAc₃-PNA **2**, while treatment with [C^γ(S)-GalNAc-T]₃-PNA **3** did not lead to any reduction in the viability compared to the control PNA. On the contrary, treatment of the cells with [C^γ(S)-GalNAc-T]₃-PNA **3** slightly enhanced the cell viability compared to the control and untreated PNAs; however, this effect was not significant.

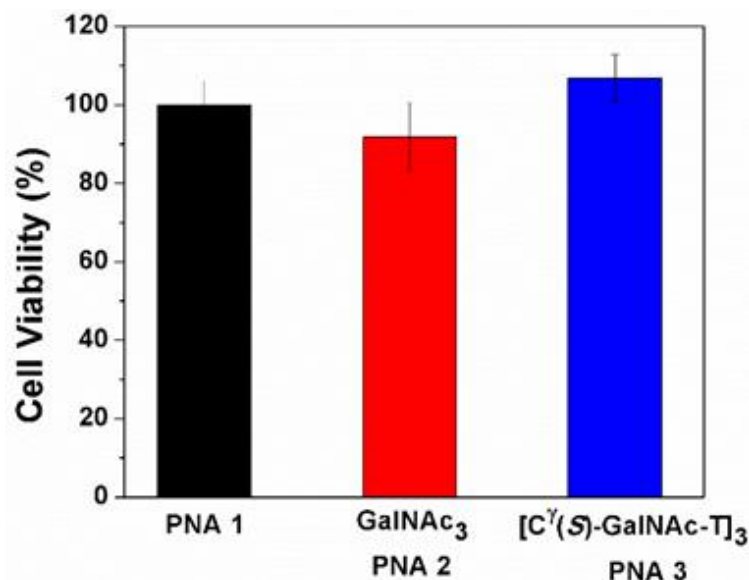


Figure 3.13 Cell viability measured using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay for HepG2 cells after treatment with 4 μM of PNA **1**, GalNAc₃-PNA **2** or [C^γ(S)-GalNAc-T]₃-PNA **3** for 12 hours. The data shown are average of three measurements. The error bars represent standard deviations.

3.6 Conclusions

- The GalNAc₃ modified PNAs, tri-antennary GalNAc₃ and trivalent [C^γ(S)-GalNAc-T]₃-PNA oligomers formed stable duplexes with complementary DNA and RNA.
- Tri-antennary GalNAc₃-PNA showed an increased thermal stability of DNA duplexes with a ΔT_m of about +5 °C, whereas [C^γ(S)-GalNAc-T]₃-PNA showed

a ΔT_m of $\sim +3$ °C compared to unmodified PNA duplex with DNA. These findings suggest that eventhough the GalNAc₃ modifications are large and bulky at N-terminus and at C^γ-positions, they do not adversely affect the stability of PNA:cDNA duplexes.

- The stability of tri-antennary GalNAc₃-PNA as well as tri-valent [C^γ(S)-GalNAc-T]₃-PNA with complementary RNA duplexes showed slightly lesser stability than that of unmodified PNA:RNA duplexes. However, both DNA and RNA duplexes with GalNAc₃-PNA were more stable than that from unmodified PNA:DNA duplexes. With melting temperatures above 50 °C for all three duplexes, the effect of destabilization induced by the GalNAc₃ modifications on the PNA:RNA duplexes are not significantly large.
- CD studies showed that conjugation of bulky triantennary GalNAc₃ and trivalent [C^γ(S)-GalNAc-T]₃-PNA at N-terminus of PNA did not affect the conformation of the duplexes they formed with complementary DNA or RNA. The CD signatures were similar to those of the unmodified PNA:DNA and PNA:RNA duplexes.
- Fluorescence studies for single stranded fluorescently labeled PNA as well as their duplexes with complementary DNA showed that the fluorescence intensity for PNA:DNA duplexes were slightly lesser than that of corresponding single stranded PNA.
- The trivalent Cf-[C^γ(S)-GalNAc-T]₃-PNA was found to be better than triantennary Cf-GalNAc₃-PNA showing enhanced cell penetration in HepG2 cell line than the unmodified PNA.
- The selectivity of the GalNAc₃-PNA to enter specifically hepatocytes HepG2 cells expressing ASGPR receptor is demonstrated by its inability to get internalized by HEK293 cell line.

3.7 Experimental procedures

3.7.1 Biophysical Studies

3.7.1a Chemicals. The unmodified and modified PNA oligomers were synthesized manually by general solid-phase PNA synthesis protocol using fmoc-strategy as described in Chapter 2. Complementary DNA and RNA oligonucleotides were obtained commercially from Integrated DNA Technologies (IDT). The unmodified PNA

monomer were obtained from ASM technologies Ltd. Germany. Salts and reagents used in buffer preparation such as NaCl, sodium cacodylate etc. were obtained from Sigma-Aldrich. The pH of the buffer solutions was adjusted using NaOH or HCl, from Sigma Aldrich.

3.7.1b. UV- T_m measurement. UV-melting experiments were carried out on Varian Cary 300 UV spectrophotometer equipped with a peltier. The samples for T_m measurements were prepared by mixing the calculated amounts of respective oligonucleotides in the stoichiometric ratio (1:1, duplex) in sodium cacodylate buffer (10 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 2 μM / 3 μM for each strand. The samples were annealed by heating at 90 °C for 10 min. followed by slow cooling to room temperature for at least 8-10 h and then refrigerated for at least 12 to 24 h. The samples (500 μL) were transferred to quartz cell and equilibrated at the starting temperature for 5 min. The OD at 260 nm was recorded in steps from 20-92 °C with temperature increment of 0.5 °C or 1 °C/min. Each melting experiment was repeated at least twice. The normalized absorbance at 260 nm was plotted as a function of the temperature. The T_m was determined from the first derivative of normalized absorbance with respect to temperature and is accurate to ± 1.0 °C. The data were processed using Origin 8.5. The concentration of all oligonucleotides were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases i.e. T = 8.8 $\text{cm}^2/\mu\text{mol}$; C = 6.6 $\text{cm}^2/\mu\text{mol}$; G = 11.7 $\text{cm}^2/\mu\text{mol}$ and A = 13.7 $\text{cm}^2/\mu\text{mol}$.^{9,10}

3.7.1c Circular dichroism. CD spectra were recorded on JASCO J-815 spectropolarimeter connected with a peltier. The calculated amounts of PNA oligomers and the complementary DNA were mixed together in stoichiometric ratio (1:1 for duplex) in sodium cacodylate buffer (10 mM) containing and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 5 or 10 μM for each strand. The samples were annealed by heating at 90 °C for 10 min. followed by slow cooling to room temperature for at least 8-10 h. The cooled samples were transferred to refrigerator for at least 8 to 12 h. To record the CD spectra of PNA:DNA duplexes and single stranded PNAs, the temperature of circulating water was kept at 10 °C. The CD spectra were recorded as an accumulation of 3 scans from 300 to 190 nm using 1 mm cell, a resolution of 0.1 nm, band-width of 1 nm, response of 1 sec and a scan speed of 100 nm/min.

3.7.1d. UV-visible and fluorescence studies. UV-Visible spectra for all PNA oligomers and complementary oligonucleotides were recorded on Perkin Elmer Lambda-45 UV-Visible Spectrophotometer and for fluorescence studies, experiments were performed on Horiba Jobin Yvon Fluorolog 3 spectrophotometer. The samples for fluorescence spectra were prepared by mixing calculated amounts of PNA and DNA in stoichiometric ratio (1:1, for duplex) in sodium cacodylate buffer (10 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 5 μ M for each strand. The annealed samples were used to record fluorescence spectra in a rectangular quartz cell at ambient temperature with λ_{exc} 458 and 484 nm; λ_{em} 520 nm; excitation slit width of 1 nm and emission slit width of 6 nm.

3.7.1e Cell permeation studies. HEK293 cells were obtained from American Type Culture Collection (ATCC: CRL-1573) and HepG2 cells (ATCC: HB-8065) were procured from National Centre for Cell Science (NCCS), Cell Repository, Pune. The cells were cultured using Eagle's minimum essential medium (MEM), which was supplemented with 10% fetal bovine serum (FBS, Sigma Aldrich). The cells were incubated at 37°C with 5% CO₂ and maintained at the exponential growth phase. About 1-2 million cells were transferred to the new flask and the cells were split. The medium was discarded, and the cells were gently washed with 2 mL of sterile phosphate buffered saline (PBS). Following PBS wash, the adhered cells were treated with 3 mL of 0.25% trypsin and 0.2% EDTA. After trypsin treatment for about 3 minutes at 37°C, 8 mL cell medium was added to stop trypsinization. The concentration of live cells was estimated using a cell counter. The cells suspension was centrifuged and the pellet was resuspended in MEM. Finally, 1-2 million cells were passaged to a new flask.

3.7.1f Confocal laser scanning microscopy analysis

Preparation of Cells. The cells were seeded in 35 x 15 mm cell imaging dishes containing 500 μ L MEM with 10% FBS and allowed to grow at 37 °C in a humidified atmosphere containing 5% CO₂ for at least 24 h. After the cells have grown to reach 50% confluency they were treated with PNAs.

Treatment of cells with CF conjugated PNAs. The media from all the imaging dishes were discarded carefully and 500 μ L fresh media containing respective 5(6)-carboxyfluorescein tagged PNAs (4 μ M) was added to the corresponding dishes. The

adhered cells were incubated at 37° C under 5% CO₂ for another 12 h. After treatment with PNAs, the medium was carefully aspirated out and the adhered cells were washed with PBS three times. The cells were then replenished with 500 µL of the media. Prior to imaging, 20 µL of DAPI was added to stain the nuclei of the cells. The confocal microscopy imaging has been repeated at least twice for each PNA.

Cell imaging. For imaging cells treated with carboxyfluorescein tagged PNAs, Leica TCS SP8 confocal microscope was used. The pin hole was set to 1 airy unit. For visualization of internalized PNAs, the carboxyfluorescein dye tagged to the PNAs were excited with 514 nm laser and the fluorescence emission signals from the molecules were detected from 550 nm to 660 nm using the HyD detector. For visualization of nuclei, DAPI was excited at 405 nm with UV laser and the signals over the range of 420 nm to 500 nm were detected. Bright field images were captured using UV laser illumination and PMT detector.

Flow cytometry. HepG2 cells were seeded in a 12-well plate (1 x 10⁵ cells/well) and cultured for 24 hours in DMEM, which was supplemented with 10% fetal bovine serum (FBS, Sigma Aldrich) at 37°C with 5% CO₂. Then, media were removed and replaced with fresh media containing either no PNA or 4 µM PNA. After an incubation period of 12 hours, the media from all the wells were discarded and the cells were washed with cold PBS containing 2 mM EDTA and harvested by trypsinization. The detached cells were diluted and suspended in PBS. The cells were then analyzed using BD FACS celesta (BD Biosciences, San Jose, CA, USA).

MTT assay. HepG2 cells were seeded in a 96-well plate with a density of 1 x 10⁴ cells/well, each containing 100 µL of DMEM. After culturing them overnight at 37 °C in a 5% CO₂ incubator, they were treated the PNAs (4 µM). Untreated cells were used as controls. After an incubation of 12 hours, the media was discarded. Fresh media (96 µL) was added to all the wells. Then, 4 µL of MTT solution (1 mg/mL) was added to each well and the cells were incubated for 4 h at 37 °C in a 5% CO₂ incubator. Finally, the media was discarded and 100 µL of the DMSO was added to each well to solubilize the formazan crystals. After 1 hour, absorbance at 570 nm was measured using EnSight multimode plate reader (Perkin Elmer, USA).

3.8 References

1. Kandasamy, P.; Kell'in, A. V.; Milstein, S.; Taneja, N.; O'Shea, J.; Shaikh, S.; Zhang, L.; van der Sluis, R. J.; Jung, M. E.; Akinc, A.; Hutabarat, R.; Kuchimanchi, S.; Fitzgerald, K.; Zimmermann, T.; van Berkel, T. J.; Maier, M. A.; Rajeev, K. G.; Manoharan, M. *J. Am. Chem. Soc.* **2014**, *136*, 16958-16961 Cantor, C. R.; Warshaw, M. W.; Shapiro, H. *Biopolymers*, **1970**, *9*, 1059-1070.
2. Nair, J. K.; Willoughby, J. L.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Puglisi, J. D.; Tinoco, I. Jr. *Methods Enzymol.* **1989**, *180*, 304-325.
3. (a) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897. (b) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.
4. Miller, C. M.; Tanowitz, M.; Donner, A. J.; Prakash, T. P.; Swayze, E. E.; Harris, E. N.; Seth, P. P. *Nucl Acid Ther.* **2018**, *28*, 119-127.
5. Kelly, S. M.; Price, N. C. *Curr Protein Pept Sci*, **2000**, *1*, 349- 384.
6. Ivanov, V. I.; Minchenkova, L. E.; Minyat, E. E.; Frank-Kamenetskii, M. D.; Schyolkina, A. K.; *J. Mol. Biol.* **1974**, *87*, 817-833.
7. Jain, R. D.; Libi, V. A.; Lahiri, M.; Ganesh, K. N. *J. Org. Chem.* **2014**, *79*, 9567-9577.
8. Cavaluzzi, M. J.; Borer, P. N. *Nucleic Acids Res.* **2004**, *32*, 13e
9. Kallansrud, G.; Ward, B. *Anal. Biochem.* **1996**, *236*, 134-138.

Chapter 4

C^γ-Janus PNAs: Design, Synthesis and Characterization of C^γ-substituted *Janus* PNA monomers

4.1 Introduction

Functionalization of the PNA backbone may dramatically change the physicochemical properties of PNA and influence the bio-distribution affecting the pharmacokinetic profile. A number of efforts have been made to synthesize conformationally constrained cyclic PNA analogs using the concept of structural preorganization to address interesting attributes of PNA:DNA/RNA hybridization.¹ Further, modifications in the acyclic PNA backbone to make them chiral² with stereocentres at C^α or C^γ have led to improvement in the properties of derived PNAs (Figure 4.1). Introduction of gem-dimethyl substitution at C^α led to PNA analogs that show preferential binding to cDNA rather than cRNA, reverse to that observed earlier with cyclohexyl PNAs,³ that were more specific for binding cRNA. Recently, Sugiyama *et al.*⁴ introduced methyl substitution at the least studied C^β position. The functionalization at these positions of the monomer may preorganize the PNA backbone to effect selectivity in PNA:DNA/RNA recognition. Specifically, configuration at the three stereogenic centres (C^α, C^β and C^γ) has been reported to influence the preference of the PNA to form left-handed or right handed helices. This, in turn, impacts the stability of the PNA:DNA duplex through the control of helix handedness.⁵

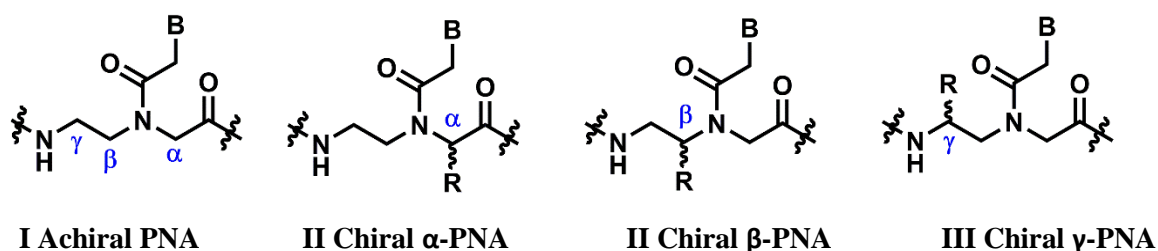


Figure 4.1 Structures of achiral PNA, chiral show α -, β - and γ -modified PNA (R = amino acid side chain or other modifications, B = nucleobase)

In order to be a good efficient DNA binding molecule, the modified PNA backbone should be neither too flexible nor too rigid. A highly flexible backbone would require a very large loss of entropy for efficient binding whereas a very rigid backbone would prevent DNA binding due to a difficult fit to adopt to target structure.⁵ The ability of unmodified aminoethyl glycyl (*aeg*) PNA to bind strongly with DNA has been interpreted as a result of the ‘constrained flexibility’ in its structure⁶ imposed by the tertiary amide link. The earliest and simplest modification involved in extension of PNA backbone by linking with a methylene group individually in each of the structural subunits (aminoethyl⁷ and glycine^{8,9}) of the PNA monomer. However, such modifications resulted in a significant decrease in the melting temperatures of the derived PNA:DNA duplexes. These studies suggested that the

inter-residue distance between the nucleobases is important and when compatible with that present in DNA and RNA, it becomes critical for competent binding with complementary oligonucleotides.

4.2 Rationale for the present study

Several PNA derivatives have been obtained by insertion of side chains at either C α - or C γ -carbons on *aeg* PNA backbone.¹⁰ These include hydroxymethyl (serine), aminobutyl (lysine), guanidinopropyl (arginine) and many other functional groups from amino acid side chains.⁹ Such acyclic chiral PNA analogues, retained the constrained flexibility of original PNA backbone, without much preorganization observed earlier with cyclic PNA analogues. However, the stereochemistry of the side chain substituents at C α and C γ seems to be important for the derived PNA oligomers in exhibiting efficient hybridization properties: *R* configuration at C α and *S* configuration at C γ .¹¹ The side chains of the D/L amino acids are versatile for easy synthesis of such chiral substituted PNA oligomers. This chapter describes synthesis and characterization of various C γ -ethylazido substituted PNA monomers (Figure 4.2) that can become precursors for assembling a new class of PNA analogues - *Janus* PNAs. Such PNA molecules will have two nucleobase sidechains per monomer and PNA oligomers derived from such monomers can recognise two molecules of complementary DNA or RNA from both faces. In such Janus PNAs this is favored because interbase separation distance on both tertiary amide (t-amide) face and the C γ face is identical to that in DNA/RNA, enabling simultaneous hybridization with cDNA/RNA from both faces. Recently Ly *et.al*¹² has successfully demonstrated such a concept by introducing side chain bearing nucleobases at C α of glycine unit in addition to the nucleobases at tertiary amide inherent in PNA. Such C α -Janus PNAs formed double duplexes with DNA sequences complementary to each face. These duplexes had different individual thermal stabilities and in double duplexes exhibited synergistic stabilization effects, with thermal stabilities higher than the corresponding individual duplexes. Inspired from such observations, the work in this chapter concerns on design of analogous C γ -Janus PNAs, where a similar internucleobase distance logic holds good. Further, the sidechain bearing backbone site C γ becomes chiral and the effect of chirality (*R* and *S*) on differential hybridization is also studied.

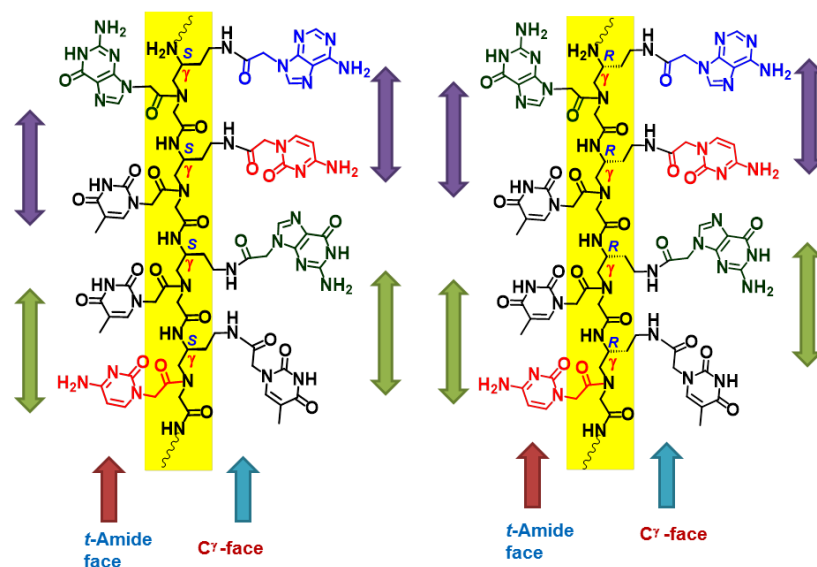


Figure 4.2 Representation of a *R* and *S* Janus oligomers having a *t*-amide and a γ face.

The precursors for *C γ -Janus* PNA oligomers are *aeg*-PNA (A/G/C/T) monomers (Figure 4.3) having ethylamino side chain at *C γ* in both *R* and *S* configurations that can be conjugated with A/G/C/T acid bases via coupling reaction to get amide group at side chain. The resulting amide linker being non-ionic is less prone for nonspecific secondary interactions with nucleotide phosphate groups in the cDNA sequence. The chirality of PNAs may have some stereospecific effects in complementation. The *C γ* -ethylamino *aeg* (A/C/G/T) PNA monomers (**1** and **2**) can be incorporated in PNA sequences on solid phase followed by single global amide coupling reaction with any nucleobase acetic acid (Figure 4.3) to yield the corresponding *Janus* PNAs that are homo-oligomeric on both faces (Figure 4.4) or have mixed sequence on the tertiary-amide face but homo-oligomer on the *C γ* face (Figure 4.5). Such *Janus* PNA oligomers possess nucleobases attached by tertiary amide bond through a carbon linker on one face as in standard PNA, nucleobases attached at *C γ* of glycine in the same PNA residue through amide linker (*C γ* face), Thus both faces have amide linker as in standard PNA. The monomer precursor of such *Janus* PNA will have nucleobases at amide face wherein *C γ* side chain will have *NHFmoc* (orthogonal protection) to enable synthesis (Figure 4.3) by using both Boc strategy for PNA assembly and Fmoc strategy to install nucleobases at *C γ* by deprotection of *NHFmoc* after the assembly PNA and couple with required nucleobase derivative at *C γ* (Figure 4.4). In such *Janus* PNAs, the inter- nucleobase geometric distance in neighboring units on either amide or *C γ* face is same and correspond to that in DNA and RNA. This facilitates simultaneous complementary binding with DNA or RNA from both faces and such *Janus* PNAs also become chiral.

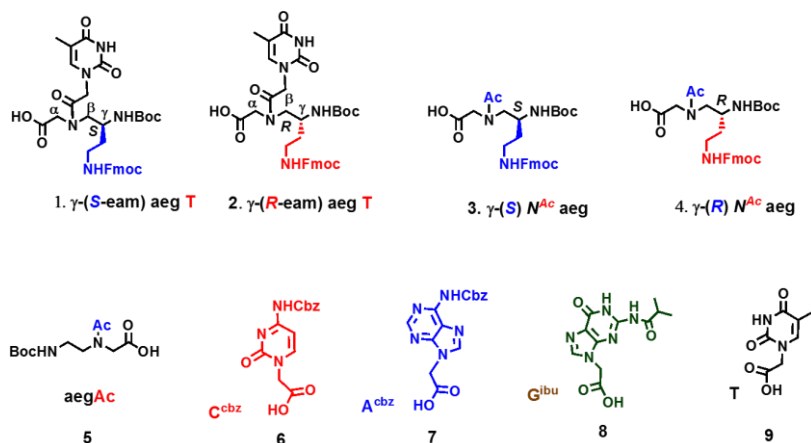


Figure 4.3 Target *Janus* PNA monomers modified at C^γ -position (1,2,3,4 and 5) and modified nucleobase acids (6,7,8 and 9).

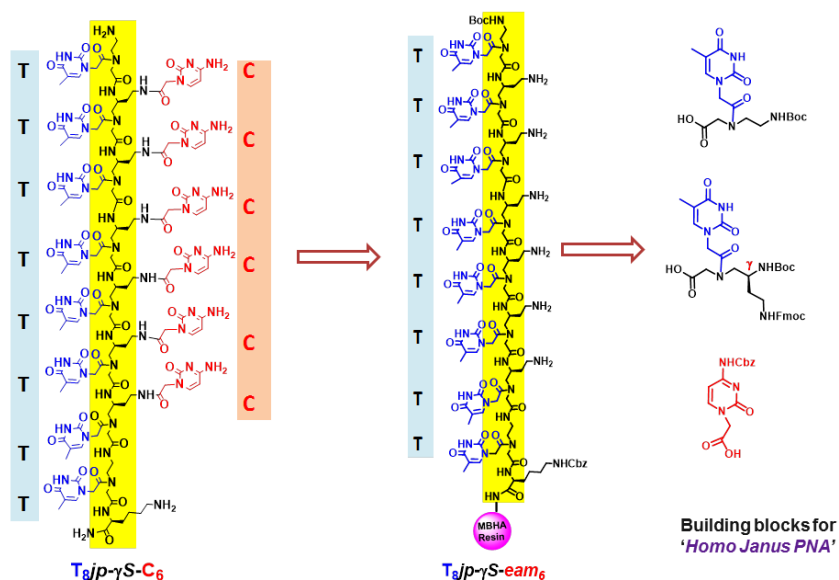


Figure 4.4 Retro synthetic strategy of **JP 1S** *Homo Janus PNA*. (eam_6 =ethylamino; aeg = aminoethyl glycol)

As far as the regio and stereospecificity of the chiral monomers are concerned, PNAs having monomers derived from L-amino acids (C^γ -S) in the glycolyl fragment of *aeg* backbone bind to the complementary DNA with greater thermal stability than those derived from the corresponding D-amino acids (C^γ -R) in the ethylene diamine (*eda*) fragment. In C^γ -substituted *eda* fragment of *aeg* backbone derived from D-amino acids translate into C^γ -R PNA oligomers and are known to preorganize into a right-handed helix, binding to complementary DNA/RNA with high affinity.¹² The situation turns out to be the reverse for substitutions of chiral amino acids at C^γ in glycine segment in *aeg* backbone; the L-amino

acids with C^γ - S at glycine hybridize with cDNA better than that with D -amino acids with C^γ - R at glycine site.¹³

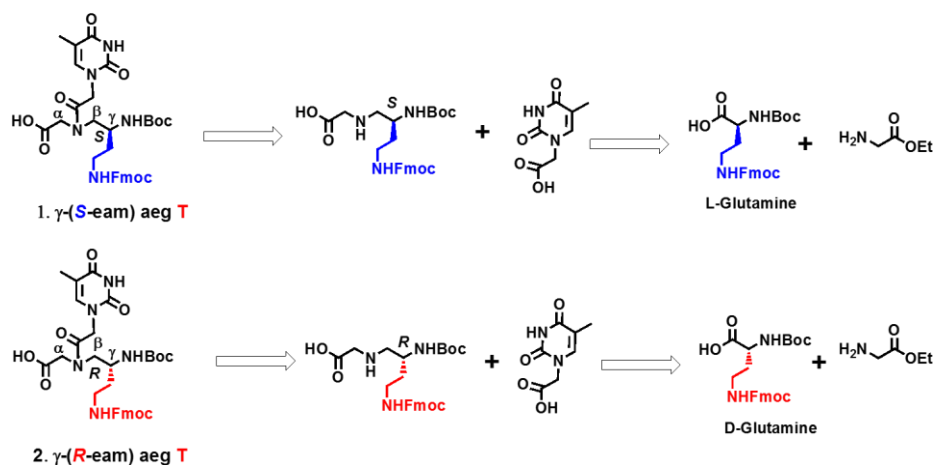


Figure 4.5 Retro synthetic strategy of S and R Janus PNA Monomers.

These results provided rationale for the design and synthesis of C^γ -substituted *Janus* PNA oligomers. Incorporation of S -ethylamino substitution at C^γ in the *aeg* PNA backbone to result C^γ -(S -ethylamino) aminoethylglycyl A/T/C/G PNA monomers such as C^γ -(S -eam)-T-*aeg* **1**, C^γ -(R -eam)-T-*aeg* **2**, C^γ -(S -eam)- N^{Ac} -*aeg* **3**, C^γ -(R -eam)- N^{Ac} -*aeg* **4** required for solid phase synthesis of *Janus* PNA homooligomers by post solid phase global coupling reaction. Such kind of monomers can be used to mix sequences by using both *Boc* and *Fmoc* strategy to synthesize *Janus* PNAs with mixed nucleobase sequence on both the faces.

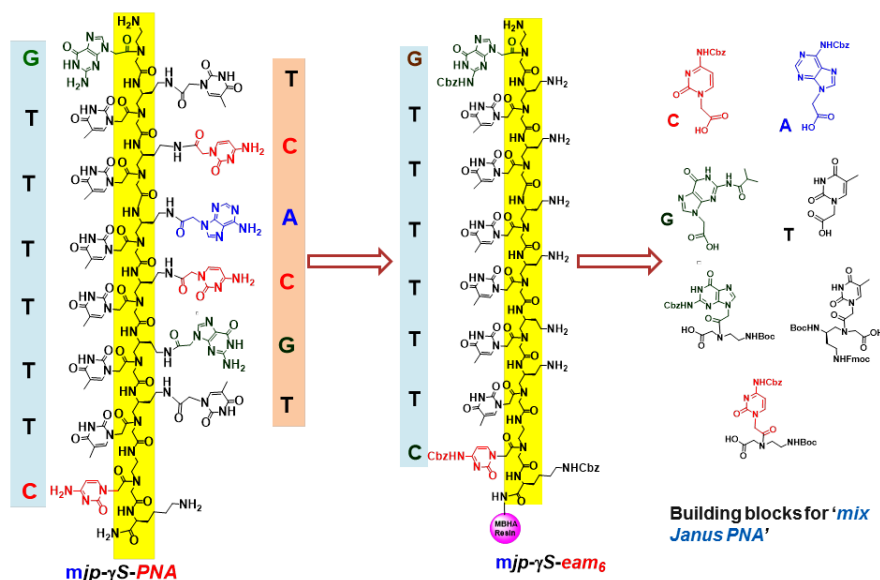


Figure 4.6 Retro synthetic strategy of **JP 3S Mix Janus PNA**. (*eam*₆=ethylamino; *aeg*= aminoethyl glyceryl)

4.3 Objectives of the present work

The specific objectives of this section are

- Synthesis of $C^\gamma(S\text{-ethylamino})\text{-aeg}$ and $C^\gamma(R\text{-ethylamino})\text{-aeg}$ PNA A/T/C/G monomers (**1, 2**)
- Synthesis of $C^\gamma(S\text{-ethylamino})\text{-N}^{\text{Ac}}\text{-aeg}$ and $C^\gamma(R\text{-ethylamino})\text{-N}^{\text{Ac}}\text{-aeg}$ monomers (**3, 4**)
- Synthesis of linker $\text{NHBoc-N}^{\text{Ac}}\text{ aeg monomer}$ (**5**)
- Synthesis of nucleobase-acetic acids (**6-9**)
- Characterization of rationally synthesized PNA monomers and their intermediates by various spectroscopic techniques.

4.4 Synthesis of modified PNA monomers

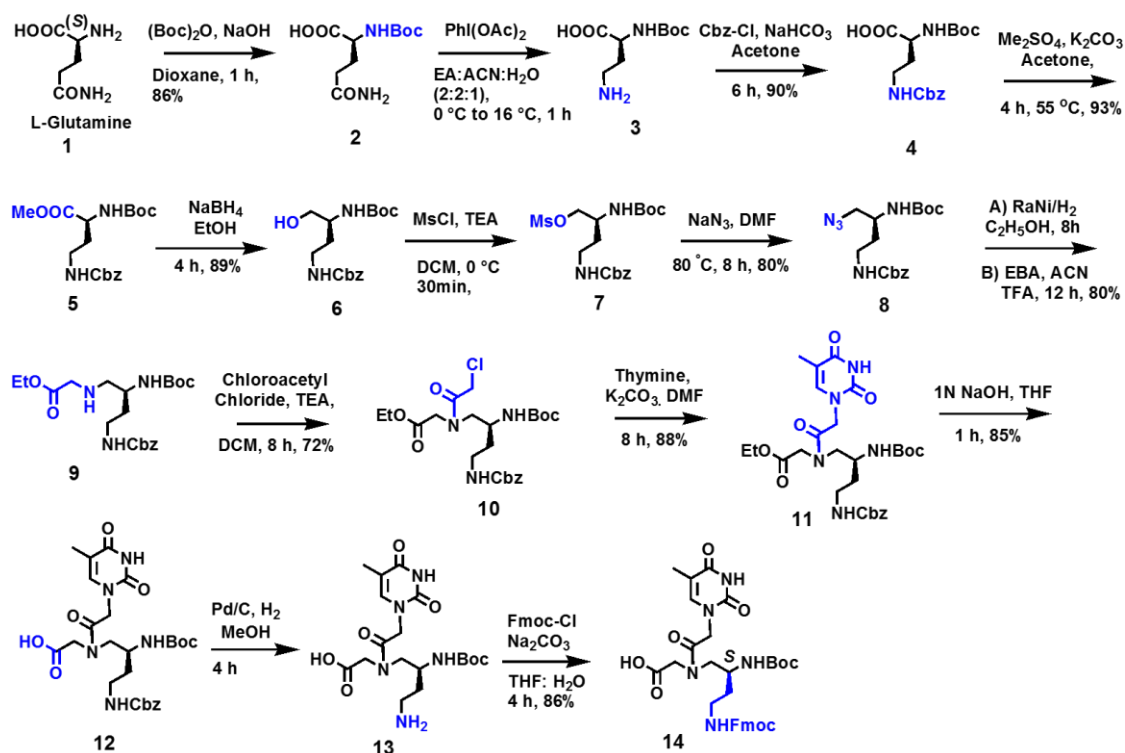
This section describes the synthesis of rationally designed $C^\gamma(S\text{-eam})\text{-aeg}$ PNA A/T/G/C monomers **1-5**.

4.4.1 Synthesis of N-(acetamido-N1-thyminy)-(NHBoc-aminoethyl)-[$C^\gamma(S)$ -(NHFmoc) aminoethyl] glycine (**14**)

The commercially available L-glutamine **1** was treated with Boc-anhydride to obtain 2(*S*)-NHBoc-L-glutamine **2** (Scheme 4.1). This was followed by its reaction with iodobenzene diacetate (PIDA)¹⁴ which led to the formation of 2(*S*)-NHBoc-4-aminobutanoic acid **3**. This was treated with benzylchloroformate in toluene in the presence of NaHCO_3 to obtain orthogonally protected 2(*S*)-NHBoc-4-(NHCbz)-aminobutanoic acid **4** in quantitative yield. Compound **4** was converted to its methyl ester **5** using dimethyl sulfate and the conversion was confirmed by the appearance of ^1H NMR peak at δ 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 **6**. Mesylation of the primary hydroxyl group by controlled addition of mesyl chloride in DCM/triethyl amine gave the mesylate derivative **7** which was immediately treated with sodium azide in dry DMF to obtain the azido compound **8**. The reduction of azide derivative using Raney Ni under hydrogenation conditions yielded the free amine yield that was in-situ alkylated with ethyl bromoacetate to the alkylated compound **9**. The secondary amine was then acylated with chloroacetyl chloride to yield the chloroacetyl compound **10**. This upon condensation with thymine afforded the $C^\gamma(S)$ -(NHFmoc)-aminoethyl)-(NHBoc-aminoethyl) ethyl glycinate **11** in good yield. The

appearance of peaks at δ 7.05 ppm and 1.87 ppm in ^1H NMR shows the presence of thymine in the desired product **11**. This ester was hydrolysed using aq. NaOH in THF to obtain the acid **12** which upon hydrogenation with Pd/C in methanol yielded the free amine **13**. The amino moiety in **13** was protected using Fmoc-Cl and 10% aq. Na_2CO_3 in THF to yield the desired Fmoc derivative **14**. All the intermediates were purified by column chromatography and characterized by ^1H , ^{13}C NMR and mass spectral analysis.

Scheme 4.1 Synthesis of C^γ -(S)-(NHFmoc)-NHBoc-T aeg monomer **14**

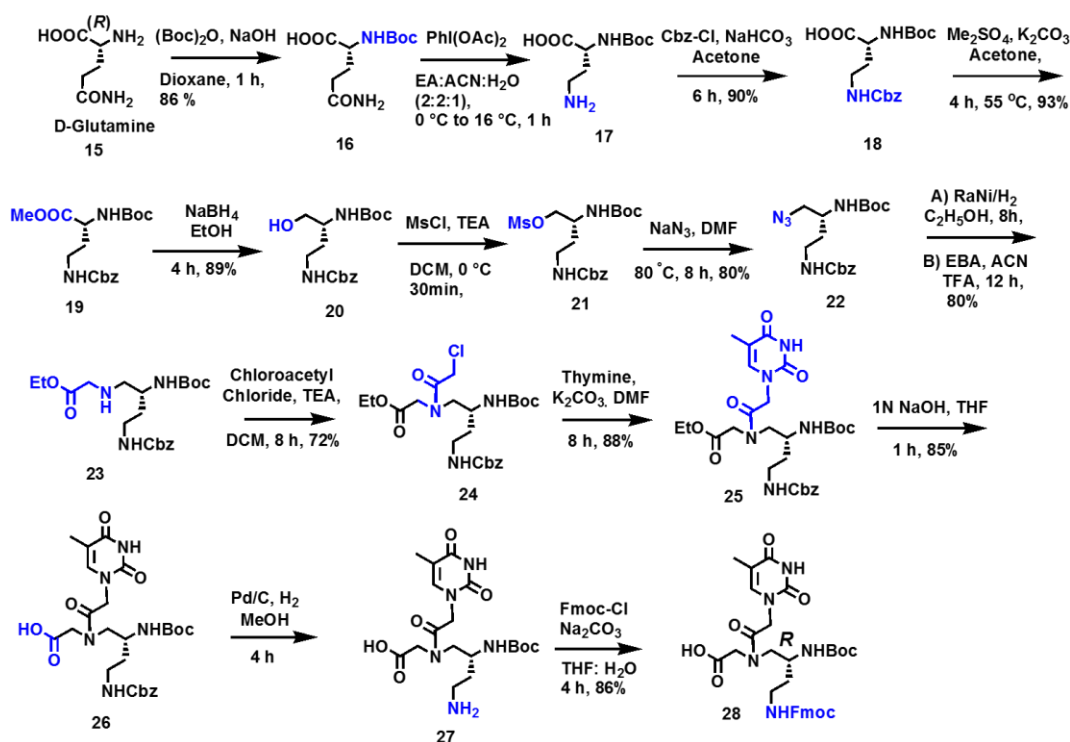


4.4.2 Synthesis of N-(acetamido-N1-thyminy)-(NHBoc-aminoethyl)-[C^γ (R)(NHFmoc) aminoethyl glycine] (**28**)

The commercially available D-glutamine **15** was treated with Boc anhydride to obtain *NH*-Boc protected D-glutamine **16** (Scheme 4.3.2). This was followed by the reaction of 2(*R*)-NHBoc-D-glutamine **16** with iodobenzene diacetate (PIDA)²⁴ which led to the formation of 2(*R*) NHBoc-4-aminobutanoic acid **17**. This was treated with benzylchloroformate in toluene in the presence of NaHCO_3 to obtain the orthogonally protected 2(*R*)-NHBoc-4-(NHCbz)-aminobutanoic acid **18** in quantitative yield. This was converted to methyl ester **19** using dimethyl sulfate and the conversion was confirmed by the appearance of ^1H NMR peak at δ 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 Boc compound **23**. It was then acylated with chloroacetyl chloride to yield the N-chloro compound **24**. This upon condensation with thymine afforded the C^γ(*R*)-(NH-Fmoc- aminoethyl)-(NH-Boc-aminoethyl) ethyl glycinate **25** in good yield. The appearance of peaks at δ 7.05 ppm and 1.87 ppm in ¹H NMR shows the presence of thymine in the desired product **25**¹⁹. The ester was hydrolysed using aq. NaOH in THF to obtain the acid **26** which upon hydrogenation with Pd/C in methanol yielded the free amine **27**. The amino moiety in **27** was protected using Fmoc-Cl and 10% aq. Na₂CO₃ in THF to yield the desired Fmoc derivative **28**. All the intermediates were purified by column chromatography and characterized by ¹H, ¹³C NMR and mass spectral analysis.

Scheme 4.2 Synthesis of C^γ-(*R*)-(NH-Fmoc)-NH-Boc –T aeg monomer (**28**)

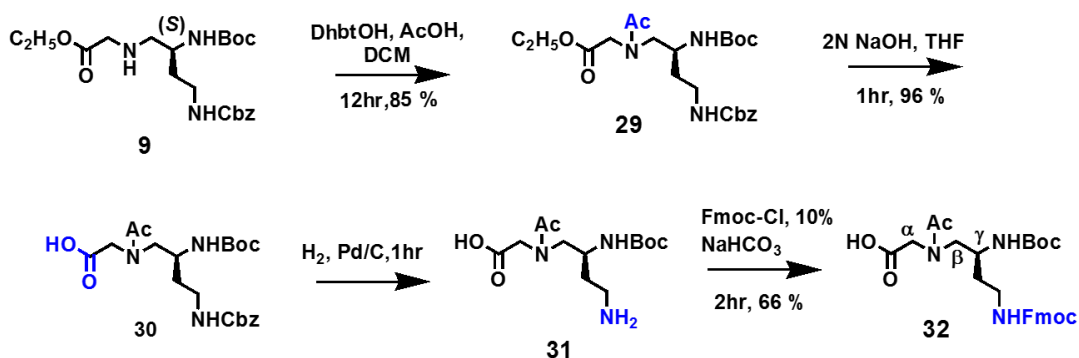


4.4.3 Synthesis of N(acetyl)-(NH-Boc)aminoethyl [C^γ(*S*)(NH-Fmoc)ethylamino] glycine (**32**)

The N-(NH-Boc)aminoethyl [C^γ(*S*)(NH-Cbz) ethylamino] glycinate **9** was protected with acetyl group in presence of coupling reagents DhbtOH and DCC in dry DCM to obtain

compound **29** (Scheme 4.3). This ester was hydrolyzed under basic conditions in presence of NaOH and THF to afford the acid **30** which upon hydrogenation with Pd/C in methanol yielded the C^γ(S)-ethylamine **31**. Reaction of the amine with Fmoc-Cl and 10% aq. Na₂CO₃ in THF to yielded the desired Fmoc derivative **32** in good yield and it was used for synthesis of aeg oligomer. All the intermediates were purified by column chromatography and characterized by ¹H, ¹³C NMR and mass spectral analysis.

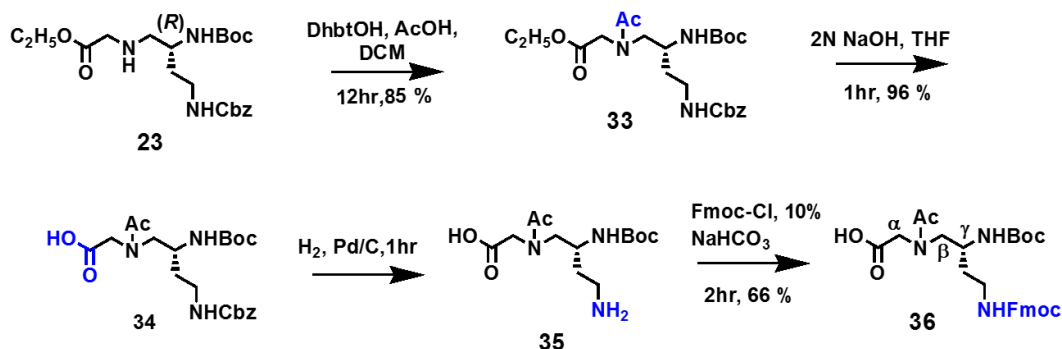
Scheme 4.3. Synthesis of NHBoc- (S) C^γ- NHFmoc- N^{Ac} aeg monomer (32)



4.4.4 Synthesis of N(acetyl)-(NHBoc)aminoethyl [C^γ(R)(NHFmoc)ethylamino] glycine (36)

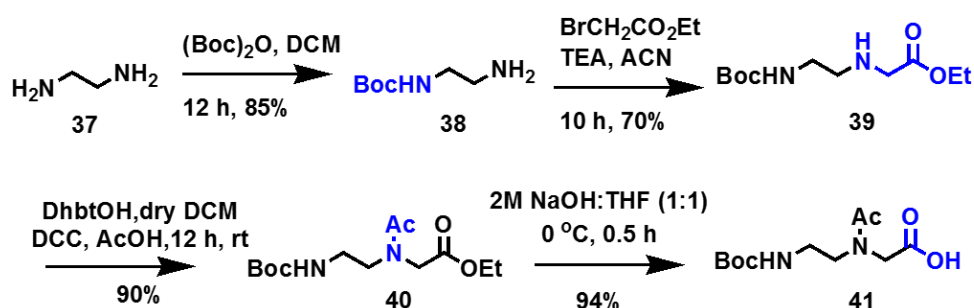
The N-(NHBoc)aminoethyl [C^γ(S)(NHCbz) ethylamino] glycinate **23** was protected with acetyl group in presence of coupling reagents DhbtOH and DCC in dry DCM to obtain compound **33**, this ester was hydrolyzed under basic conditions in presence of NaOH and THF to afford the acid **34** which upon hydrogenation with Pd/C in methanol yielded the C^γ(S)-ethylamine **35**. Reaction of the amine with Fmoc-Cl and 10% aq. Na₂CO₃ in THF to yielded the desired Fmoc derivative **36** in good yield and it was used for synthesis of aeg oligomer. All the intermediates were purified by column chromatography and characterized by ¹H, ¹³C NMR and mass spectral analysis.

Scheme 4.4. Synthesis of NHBoc-aminoethyl C^γ(R)-NHFmoc-aminoethyl-N-acetyl glycine (36)



4.4.5 Synthesis of *NHBoc-N-acetyl aeg* (41)

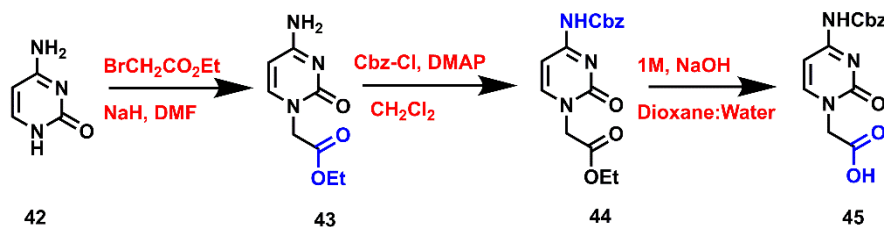
The commercially available ethylenediamine **37** was treated with di-*t*-butyl dicarbonate [(Boc)₂O] to get mono *NH*-Boc-protected ethylenediamine **38** (Scheme 4.5). This was *N*-alkylated by the reaction with ethylbromoacetate, triethylamine in acetonitrile solvent to form compound **39**. Compound **39** was then subjected to acylation with glacial acetic acid in presence of DhbtOH in dry DCM to obtain Compound **40**, which was further hydrolyzed under basic conditions to afford the desired BocNH protected monomer **41** in very good yield for synthesis of control *aeg* oligomer.

Scheme 4.5 Synthesis of *NHBoc N-acetyl aeg* monomer (41)

4.4.6 Synthesis of nucleobase acetic acid

This section describes the synthesis of nucleobase (A/G/C/T)-acetic acid from the corresponding nucleobases.

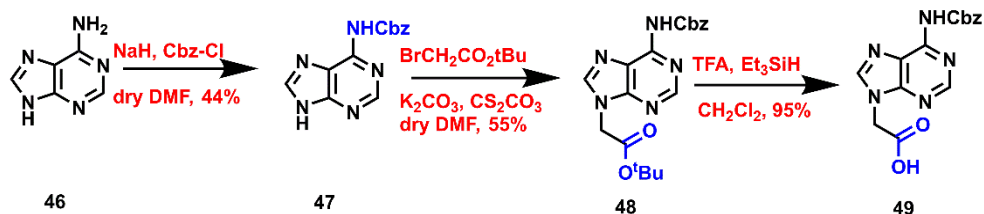
4.4.6a Synthesis of (*N*⁴-Benzyloxycarbonyl)cytidinyl-acetic acid (45). The commercial available cytosine **42** was alkylated with ethyl bromoacetate in the presence of NaH to obtain compound **43** followed by protection of 4-NH₂ function as Cbz derivative in presence of DMAP to obtain the ester **44**. The ester was hydrolyzed in presence of basic medium to afford cytosine acetic acid **45** in 43% overall yield.¹⁵

Scheme 4.6 Synthesis of (*N*⁴-Benzyloxycarbonyl)-cytidinyl-*N*1-acetic acid (45)

4.4.6b Synthesis of *N*₆(Benzyloxycarbonyl)-adeninyl-*N*₉-acetic acid (49). The synthesis was started from compound **46** in which 6-NH₂ group of adenine was protected with Cbz-Cl to get compound **47** followed by *N*₉-alkylation by *t*-butylbromoacetate to get compound **48** in the next

step tertiary butyl group was undergo cleavage to afford the desired N9 derivative acid **49**. (Scheme 3.7)¹⁶

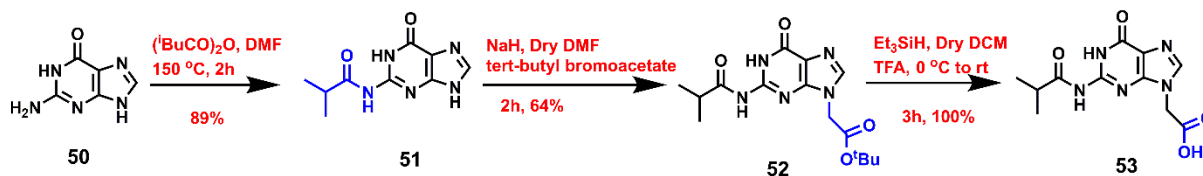
Scheme 4.7 Synthesis of *N6*(Benzyloxycarbonyl)-adeninyl-*N9*-acetic acid (49**)**



4.4.6c Synthesis of *N2*-(Isobutanoyl)-guaninyl-*N9*-acetic acid (**53**)

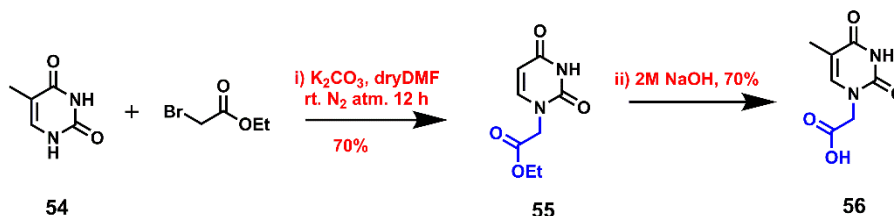
The 2-NH₂ group of guanine **50** was protected in presence of isobutyric anhydride and dry DMF to get compound **51**, followed by selective alkylation of N9 of guanine in presence of t-butylbromoacetate to get c **52**. The tertiary butyl ester group was hydrolysed using triethylsilane and DCM to yield the desired *N2*-(Isobutanoyl)-guaninyl-*N9*-acetic acid **53** (Scheme 4.8)¹⁷⁻²¹

Scheme 4.8 Synthesis of *N2*-(Isobutanoyl)-guaninyl-*N9*-acetic acid (53**)**



3.4.6d Synthesis of thymine-1-yl acetic acid (56**)**. The alkylation of thymine **54** in presence of ethylbromoacetate to **54**, followed by in situ hydrolysis under basic conditions gave the desired thymine-1-yl acetic acid **56**. (Scheme 4.9)²²

Scheme 4.9 Synthesis of thymine-1-ylacetic acid (56**)**



4.5 Summary

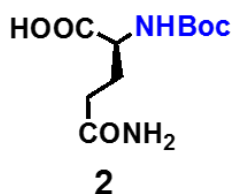
In summary, this section described the synthesis and characterization of orthogonally

protected precursors for synthesis of designed C γ -substituted Janus PNA monomers and synthesis of nucleobase acetic acids. The precursor for Janus PNA monomer has nucleobases A/T/C/G at one face (t-amide) and other face (C γ -face) has the ethylamino chain with the terminal amine protected as NHFmoc. Such monomers can be used in synthesis of the homo as well as mixed sequence Janus PNA oligomers on solid phase. Deprotection of C γ -sidechain NHFmoc group on solid phase and coupling with appropriate nucleobase acetic acid would ultimately yield the desired Janus PNA oligomers as described in next chapter. All the intermediates have been characterized by ^1H & ^{13}C NMR spectroscopy and mass spectral analysis as shown in Section 4.6. The next chapter deals with the incorporation of these modified monomers into PNA oligomers using solid phase peptide synthesis and generation of Janus oligomers.

4.6 Experimental section

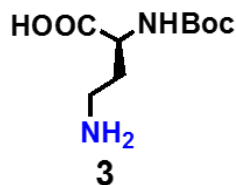
This section details the synthetic procedures and spectral characterization of various rationally designed PNA monomers. The chemicals used were of laboratory or analytical grade. All the solvents used were distilled or dried to carry out different reactions. Reactions were monitored by thin layer chromatography (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 GF₂₅₄ 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). ^1H and ^{13}C NMR spectra were recorded using Bruker AC-400 (400 MHz) or JEOL 400 MHz NMR spectrometers. The delta (δ) values for chemical shifts are reported in ppm and 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, Synapt G2 High Definition Mass Spectrometry

4.6.1. 5-amino-2-(S)-[(tert-butoxycarbonyl)amino]-5-oxopentanoic acid (2)



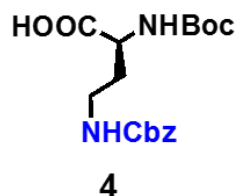
A solution of di-*tert*-butyl dicarbonate [(Boc)₂O] (16.5 g, 17 mL, 75 mmol) in dioxane (100 mL) was added in portion to an ice-cold solution of L-glutamine (10 g, 68 mmol) in aq. NaOH (1 N, 100 mL). The reaction mixture was further stirred at 0 °C for 1 h. After completion of reaction dioxane was removed completely under vacuum from the reaction mixture. The aqueous layer was washed with diethyl ether to remove excess [(Boc)₂O] and cooled in ice-water bath, acidified to pH 2-3 by slow addition of saturated aq. KHSO₄ solution and then extracted with ethyl acetate (3 × 150 mL). The combined organic extracts were dried over an. Na₂SO₄, filtered and concentrated to give compound **2** as a white powder which was used without further purification (14.5 g, 86% yield). M.p = 119-121 °C; R_f = 0.39 EtOAc/MeOH (50:50); [α]²⁵_D - 2.960 (c 0.5, MeOH); MS (MALDI-TOF) *m/z* calcd for C₁₀H₁₈N₂O₅ [M + K]⁺ 285.0853, found 285.0501.

4.6.2. 4-amino-2-(S)-[(*tert*-butoxycarbonyl)amino] butanoic acid (**3**)



A slurry of compound **2** (5 g, 20 mmol), ethyl acetate (25 mL), acetonitrile (25 mL), water (12 mL) and iodobenzene diacetate (8 g, 24 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) and cooled to 0 °C and filtered under vacuum. The filter cake was washed with EtOAc and dried in vacuum to obtain compound **3** (2.6 g, 65% yield). M.p = 200-201 °C; R_f = 0.2 EtOAc/MeOH (50:50); [α]²⁵_D + 13.6 (c 0.5, MeOH); MS (MALDI-TOF) *m/z* calcd for C₉H₁₈N₂O₄ [M + K]⁺ 257.0904, found 257.0740.

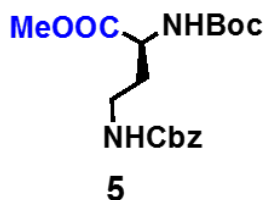
4.6.3. 4-amino (benzyloxy)carbonyl)-2-(S)-[(*tert*-butoxycarbonyl)amino]butanoic acid (**4**)



The solution of NaHCO₃ (1 g, 15 mL, 14 mmol) in water was added to an ice-cold solution of compound **3** (1 g, 4.6 mmol) in acetone (25 mL) and stirred for 10 min at 0 °C. To this, benzylchloroformate (2 g, 2 mL, 5.5 mmol) in toluene was added and the reaction mixture was stirred overnight at room temperature. Acetone was removed completely under vacuum and the aqueous layer was washed with diethyl ether (2 x 30 mL). The aqueous layer

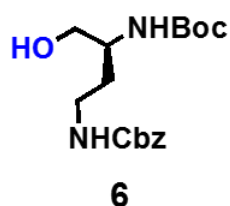
was acidified to pH 2-3 with saturated aq. KHSO_4 solution and extracted with EtOAc (3 x 60 mL). The combined organic layer was dried over anhyd. Na_2SO_4 , filtered and concentrated to give compound **4** as sticky oil (1.45 g, 90% yield). $R_f = 0.67$ EtOAc/MeOH (50:50). ^1H NMR (400 MHz, CDCl_3) δ 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}C NMR (100 MHz, CDCl_3) δ 175.7, 157.0, 156.0, 136.3, 128.5, 128.1, 80.3, 66.9, 51.1, 37.2, 33.3, 28.3; MS (MALDI-TOF) m/z calcd for $\text{C}_{17}\text{H}_{24}\text{N}_2\text{O}_6$ $[\text{M} + \text{K}]^+$ 391.1271, found 391.1075.

4.6.4 4-amino (benzyloxy)carbonyl)-2-(S)-[(tert-butoxycarbonyl)amino] methyl butanoate (**5**)



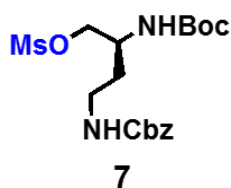
To a stirred solution of compound **4** (5 g, 15 mmol), K_2CO_3 (5.0 g, 36 mmol) in acetone (70 mL) was added dimethyl sulfate (2 mL, 17.5 mmol) and reaction mixture was heated to 55°C for 5 h under reflux condenser. Acetone was evaporated completely and water (90 mL) was added to the concentrate, which was then extracted with ethyl acetate (3 x 50 mL). The combined organic layers was washed with brine, dried over an. Na_2SO_4 , filtered and concentrated. The compound **5** was purified on column chromatography to give as white solid (2.5 g, 93% yield). mp = $65-68^\circ\text{C}$; $R_f = 0.5$ petroleum ether/EtOAc (70:30); $[\alpha]_D^{25} -18.8$ (*c* 0.5, Methanol); ^1H NMR (200 MHz, CDCl_3) δ 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}C NMR (50 MHz, CDCl_3) δ 172.9, 156.3, 155.7, 136.5, 128.3, 127.9, 80.0, 66.5, 52.3, 50.8, 37.0, 33.2, 28.1 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.

4.6.5 4-Amino(benzyloxycarbonyl)-2 (S)-[(tert-butoxycarbonyl)amino]-butanol (**6**)



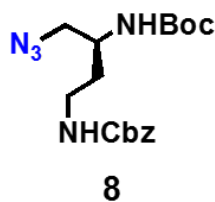
To a stirred solution of compound **5** (2.5 g, 13.0 mmol) in absolute ethanol (30 mL) was added sodium borohydride (0.6 gm, 20.5 mmol) and reaction mixture was stirred for 6 h under nitrogen atmosphere at RT. Ethanol was evaporated completely and water (60 mL) was added to the concentrate which was extracted with ethyl acetate (3 x 30 mL). The combined organic layer was washed with brine, dried over an. Na₂SO₄, filtered and concentrated. The residue was then purified on column chromatography to give compound **6** as white solid (2.0 g, 89 % yield). M.p = 80-82 °C; R_f = 0.4 petroleum ether/EtOAc (50:50); [α]²⁵_D - 24.4 (c 0.5, Methanol); ¹H NMR (400 MHz, CDCl₃) δ 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. ¹³C NMR (100 MHz, CDCl₃) δ 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 C₁₇H₂₆N₂O₅ [M + K]⁺ 377.1479, found 377.1080.

4.6.6 4-Amino(benzyloxycarbonyl)-2(S)-[(tert-butoxycarbonyl)amino]butyl methane sulfonate (**7**)



To an ice-cold solution of compound **6** (2 g, 6 mmol), triethyl amine (2.0 mL, 15 mmol) in dry DCM (30 mL) was added mesyl chloride (0.6 mL, 8 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 anhyd. Na₂SO₄, filtered and concentrated on a rotary evaporator to give compound **7** (2.3 g, 92% crude yield). R_f = 0.53 petroleum ether/EtOAc (50:50). This compound was used for next step without further purification.

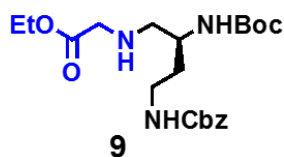
4.6.7 4-Amino(benzyloxy)carbonyl)-2(S)-[(tert-butoxycarbonyl)amino]-butylazide (**8**)



The solution of compound **7** (2.0 g, 5.5 mmol) and sodium azide (5.5 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

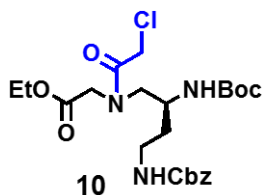
added which was extracted with ethyl acetate (3 x 30 mL). The ethyl acetate layer was washed with water (25 mL) and brine (25 mL). The combined organic layer was dried over an. Na_2SO_4 , filtered and concentrated. The residue obtained was then purified on column chromatography to give compound **8** as sticky yellowish oil (1.76 g, 87% yield). $R_f = 0.73$ petroleum ether/EtOAc (50:50); ^1H NMR (400 MHz, CDCl_3) δ 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}C NMR (100 MHz, CDCl_3) δ 156.4, 155.7, 136.5, 128.4, 128.0, 19.8, 66.5, 54.9, 47.6, 37.4, 32.8, 28.1 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.

4.6.8 N-ethylamino-(tert-butoxycarbonyl)amino)-2-(S)-[(benzyloxycarbonyl)aminoethyl] ethyl glycinate (**9**)



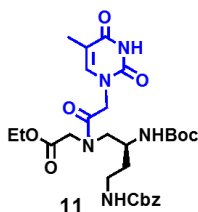
To a solution of compound **8** (750 mg, 2.0 mmol) in absolute ethanol (15 mL) taken in hydrogenation flask was added Raney Nickel (2 mL). The reaction mixture was hydrogenated in a Parr apparatus for 6 h at room temperature and H_2 pressure of 50-55 psi. The catalyst was filtered off from reaction mixture and the filtrate was concentrated under reduced pressure to yield a residual yellowish oil of amine (630 mg, 1.8 mmol) was treated with ethylbromo acetate (0.2mL, 1.7 mmol) in acetonitrile (20 mL) using triethyl amine (0.8 mL) and the reaction mixture was stirred at room temperature for 12 h. Acetonitrile was evaporated completely under vacuum and water (50 mL) was added to the concentrate. The aqueous layer was extracted with ethyl acetate (3 x 40 mL). The combined organic layer was washed with sat. NaHCO_3 , brine, dried over an. Na_2SO_4 , filtered and concentrated on rotary evaporator. The residue obtained was purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **9** as yellowish oil (700 mg, 80%). $R_f = 0.48$ petroleum ether/EtOAc (20:80); ^1H NMR (200 MHz, CDCl_3) 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. ^{13}C NMR (50 MHz, CDCl_3) δ 171.9, 156.4, 136.6, 128.3, 127.9, 127.8, 79.4, 66.3, 60.8, 52.8, 50.4, 47.5, 37.4, 33.5, 29.5, 28.2, 14.0 ppm MS (MALDI-TOF) m/z calcd for $\text{C}_{21}\text{H}_{33}\text{N}_3\text{O}_6$ [$\text{M} + \text{K}$] $^+$ 423.2369, found 462.2006.

4.6.9 N-ethylamino(tertbutoxycarbonyl)-N(chloroacetamido)-2(S)-[(benzyloxycarbonyl) aminoethyl] ethyl glycinate (**10**)



To an ice-cold solution of compound **9** (3.0 g, 7.3 mmol) and triethyl amine (3 g; 4 mL, 30 mmol) in dry DCM (50 mL), chloroacetyl chloride (0.6 mL, 7.3 mmol) was added and the 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 anhyd. Na₂SO₄, filtered and concentrated. The residue was then purified on silica gel (100-200 mesh) using petroleum ether and EtOAc to give compound **10** as colourless sticky oil (2.6 g, 72 %). *R_f* = 0.59 petroleum ether/EtOAc (40:60); ¹H NMR (200 MHz, CDCl₃) 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. ¹³C NMR (50 MHz, CDCl₃) 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 C₂₃H₃₄ClN₃O₆ [M + K]⁺ 538.1722, found 538.1591.

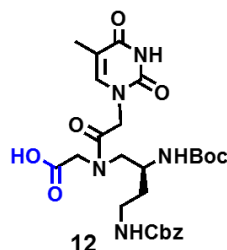
4.6.10 N-ethylamino(tertbutoxycarbonyl)-N(acetamido-N1-thyminylyl)-2(S)-[(benzyloxy carbonyl) aminoethyl] ethyl glycinate (**11**)



A solution of compound **10** (1 g, 2 mmol), K₂CO₃ (0.3 g, 2.5 mmol) and thymine (0.3 g, 2.5 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 x 40 mL). The ethyl acetate layer was washed with water (40 mL) and brine (20 mL). The combined organic layer was dried over an. Na₂SO₄, filtered and concentrated. The residue obtained was then purified on silica gel (100-200 mesh) using petroleum ether and ethyl acetate to give compound **11** as white solid (0.95 g, 81%). mp = 92-94 °C; *R_f* = 0.47 EtOAc (100); [α]_D²⁵ - 8.1 (c 1,

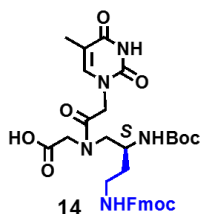
Methanol); ^1H NMR (400 MHz, CDCl_3) δ 9.86 (maj, 7H) 9.56 (min) (br, 1H), 7.37-7.27 (m, 5H), 7.06 (min) 6.99 (maj) (s, 1H), 5.68-5.64 (maj) & 5.58-5.56 (min.) (comp, 1H), 5.11-5.04 (m, 2H), 4.82, 4.77 (br, 1H), 4.47-4.37 (m, 1H), 4.28-4.13 (m, 3H), 3.94-3.51 (m, 3H), 3.43-3.02 (m, 3H), 2.16 (br, 1H), 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) δ 169.1, 167.5, 164.4, 156.7, 151.5, 140.9, 136.5, 128.4, 128.0, 110.8, 79.6, 66.6, 62.3, 52.1, 51.5, 49.8, 48.6, 47.8, 47.2, 37.6, 32.9, 28.3, 14.0, 12.3 ppm. MS (MALDI-TOF) m/z calcd for $\text{C}_{28}\text{H}_{39}\text{N}_5\text{O}_9$ $[\text{M} + \text{K}]^+$ 628.2385, found 628.2361.

4.6.11 N-ethylamino(tertbutoxycarbonyl)-N(acetamido-N1-thyminyl)-2(S)-[(benzyloxy carbonyl) aminoethyl] glycine (12)



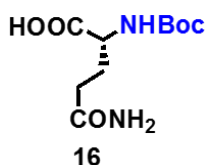
To a stirred solution of compound **11** (0.5g, 1 mmol) in methanol was added 10% aq. LiOH and reaction mixture was stirred at room temperature for 3-4 h. Methanol was removed under vacuum and the aqueous layer was washed with diethyl ether. The aqueous layer was then neutralized with activated Dowex H^+ resin till pH of the solution turned 5-6. The resin was removed by filtration and the filtrate was concentrated to obtain the resulting compound **12** as white solid (0.4 g, 88%). M.p = 241-245 $^\circ\text{C}$; R_f = 0.5 EtOAc/MeOH (50:50); $[\alpha]_D^{25}$ - 3.0 (c 0.5, Methanol); ^1H NMR (400 MHz, DMSO-d_6) δ 11.30 (min) & 11.25 (maj) (br, 1H), 7.35-7.26 (m, 5H), 7.21-7.15 (m, 1H), 6.91-6.89 (min) & 6.78-6.76 (maj) (d, J = 8 Hz, 1H), 5.02-4.96 (m, 2H), 4.76-4.72 (maj) & 4.59-4.55 (min.) (d, J = 16 Hz, 1H), 4.42 (s, 2H), 3.97-3.74 (m, 4H), 3.47-3.21 (m, 2H), 3.12-2.91 (m, 3H), 1.75 (min) & 1.73 (maj) (s, 3H), 1.58-1.41 (m, 2H), 1.37 (maj) & 1.36 (min) (s, 9H) ppm ^{13}C NMR (100 MHz, CDCl_3) δ 170.6, 167.9, 167.0, 164.5, 156.0, 155.5, 151.1, 142.1, 137.3, 108.0, 77.7, 65.1, 51.9, 51.1, 48.4, 47.7, 47.1, 46.5, 37.7, 32.1, 28.3, 12.0 ppm. MS (MALDI-TOF) m/z calcd for $\text{C}_{26}\text{H}_{35}\text{N}_5\text{O}_9$ $[\text{M} + \text{K}]^+$ 600.2072, found 600.1787.

4.6.12 N-ethylamino(tertbutoxycarbonyl)-N(acetamido-N1-thyminyl)-2S[(Fmoc) aminoethyl] glycine (14)



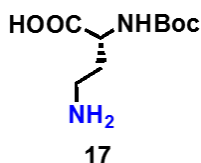
To the Compound **12** (1 g, 2 mmol) dissolved in ethanol, 10% Pd/C on charcoal was added under H₂ atmosphere and stirred at room temperature for 6 h. The completion of reaction was monitored by TLC. The reaction mixture was filtered on celite-545 pad and filtrate was collected, evaporated to get solid product (crude weight 0.8 g). The free amine compound (1 g, 2 mmol) was dissolved in 10 ml THF:H₂O (1:1) taken in a clean RBF and stirred at 0 °C. Na₂CO₃ (0.5 g, 5 mmol) was added to the reaction mixture followed by slow dropwise addition of Fmoc-Cl (0.6 g, 2 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 **14** (1.0 g, 85%). ¹H NMR (400 MHz, DMSO-d₆) 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), 4.03-3.60 (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. ¹³C NMR (100 MHz, DMSO-d₆) δ 171, 168.3, 167.7, 156.1, 151.3, 144.4, 144.3, 141.2, 128.1, 127.5, 125.6, 120.6, 106.6 ppm. HRMS (ESI-TOF) m/z calcd for C₂₈H₃₉N₅O₉ [M + Na]⁺ 672.2645, found 672.2612.

4.6.13 5-Amino-2-(R)-[(tert-butoxycarbonyl)amino]-5-oxopentanoic acid (**16**)



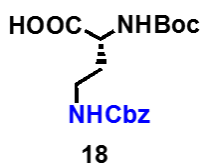
The compound **16** was synthesized starting from D-Glutamine by following procedure same as described for **2**. MS (MALDI-TOF) m/z calcd for C₁₀H₁₈N₂O₅ [M + K]⁺ 285.0853, found 285.0205.

4.6.14 4-amino-2-(R)-[(tert-butoxycarbonyl)amino] butanoic acid (**17**)



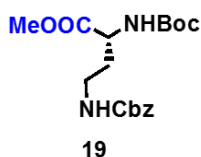
The compound **17** was synthesized starting from 5-amino-2-(*R*)-[(*tert*-butoxycarbonyl)amino]-5-oxopentanoic acid **16** by following procedure same as described for **3**. MS (MALDI-TOF) m/z calcd for $C_9H_{18}N_4O_4$ [$M + K$] $^+$ 257.0904, found 257.0740

4.6.15 4-amino (benzyloxy)carbonyl)-2-(*R*)-[(*tert*-butoxycarbonyl)amino]butanoic acid (**18**)

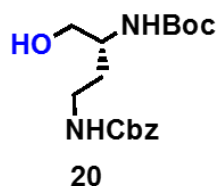


The compound **18** was synthesized starting from 4-amino-2-(*R*)-[(*tert*-butoxycarbonyl)amino]butanoic acid **17** by following procedure same as described for **4**. 1H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.27 (m, 5H), 5.58 (d, $J = 49.7$ Hz, 2H), 5.35 (m, $J = 40.1$ Hz, 1H), 5.14 – 5.02 (m, 2H), 4.40 – 4.28 (m, 1H), 3.71 (s, 2H), 3.54 – 3.41 (m, 1H), 3.14 – 3.00 (m, 1H), 2.07 (d, $J = 3.9$ Hz, 1H), 1.69 (dd, $J = 9.0, 4.8$ Hz, 1H), 1.46 – 1.38 (m, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 175.0, 173.2, 156.6, 156.1, 156.0, 156.9, 136.6, 128.6, 128.2, 128.2, 80.4, 77.2, 66.9, 66.8, 52.6, 51.0, 37.4, 37.2, 33.6, 30.1, 28.4. MS (MALDI-TOF) m/z calcd for $C_{17}H_{24}N_2O_6$ [$M + K$] $^+$ 391.1271, found 391.1075.

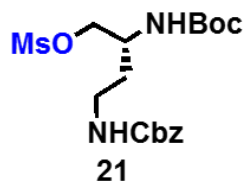
4.6.16 4-amino (benzyloxy)carbonyl)-2-(*R*)-[(*tert*-butoxycarbonyl)amino] methyl butanoate (**19**)



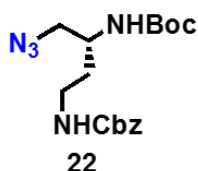
The compound **19** synthesized from 4-amino (benzyloxy)carbonyl)-2-(*R*)-[(*tert*-butoxy carbonyl)amino] butanoic acid **18** by following same procedure described for **5**. 1H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.28 (m, 5H), 5.50 (s, 1H), 5.26 (s, 1H), 5.10 (q, $J = 12.3$ Hz, 2H), 4.37 (t, $J = 8.5$ Hz, 1H), 3.72 (s, 3H), 3.54 – 3.43 (m, 1H), 3.11 – 3.00 (m, 1H), 2.07 (dd, $J = 8.4, 4.4$ Hz, 1H), 1.89 (s, 1H), 1.73 – 1.62 (m, 1H), 1.43 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.15, 156.5, 156.0, 136.7, 128.6, 128.2, 128.2, 80.4, 77.2, 66.8, 52.6, 51.0, 37.2, 33.7, 28.4. MS (MALDI-TOF) m/z calcd for $C_{18}H_{26}N_2O_6$ [$M + K$] $^+$ 405.1428, found 405.1199.

4.6.17 4-Amino(benzyloxycarbonyl)-2-(*R*)-[(*tert*-butoxycarbonyl)amino]-butanol (**20**)

The compound **20** synthesized from 4-amino (benzyloxy)carbonyl)-2-(*R*)-[(*tert*-butoxy carbonyl)amino] methyl butanoate **19** by following same procedure described for **6**. ^1H NMR (400 MHz, CDCl_3) δ 7.82 – 7.72 (m, 5H), 6.27 (s, 1H), 5.65 (d, $J = 8.4$ Hz, 1H), 5.53 (q, $J = 12.3$ Hz, 2H), 4.16 – 3.94 (m, 4H), 3.87 (dd, $J = 12.8, 6.1$ Hz, 1H), 3.47 (d, $J = 5.3$ Hz, 1H), 3.04 (s, 8H), 2.22 – 2.10 (m, 1H), 2.08 – 1.96 (m, 1H), 1.93 – 1.83 (m, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 156.7, 136.8, 128.6, 128.2, 80.0, 77.1, 66.7, 65.6, 49.9, 37.8, 32.3, 28.5. 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. 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.

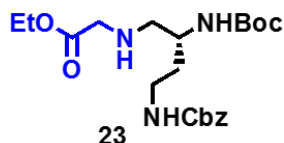
4.6.18 4-Amino(benzyloxycarbonyl)-2(*S*)-[(*tert*-butoxycarbonyl)amino]butyl methane sulfonate (**21**)

The compound **21** synthesized from **20** by following same procedure described for **7** and used for next step without any further purification.

4.6.19 4-Amino(benzyloxy)carbonyl)-2(*R*)-[(*tert*-butoxycarbonyl)amino]-butylazide (**22**)

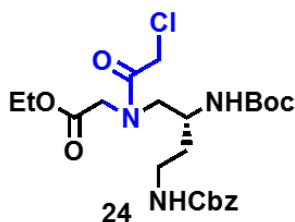
The compound **22** synthesized from **21** by following same procedure described for **8**. ^1H NMR (400 MHz, CDCl_3) δ 7.44 – 7.30 (m, 31H), 5.47 (s, 1H), 5.11 (q, $J = 12.3$ Hz, 2H), 4.73 (s, 1H), 3.84 (s, 1H), 3.45 (d, $J = 16.0$ Hz, 3H), 3.08 – 2.96 (m, 1H), 1.74 (dd, $J = 8.3, 4.3$ Hz, 1H), 1.59 (dt, $J = 13.7, 7.2$ Hz, 1H), 1.46 (s, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 156.5, 155.9, 136.6, 128.5, 128.1, 80.1, 66.7, 55.2, 47.7, 37.5, 33.1, 28.3. 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.

4.6.20 N-ethylamino-(tert-butoxycarbonyl)amino)-2-(R)-[(benzyloxycarbonyl)amino]ethyl glycinate (23).



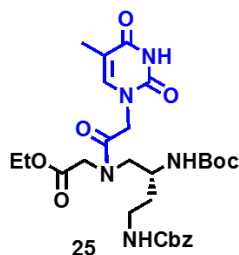
The compound **23** synthesized from **22** by following same procedure described for **9**. ^1H NMR (400 MHz, CDCl_3) δ 7.40 – 7.31 (m, 5H), 5.61 (s, 1H), 5.05 (ddd, $J = 26.6, 19.7, 9.8$ Hz, 4H), 4.24 – 4.14 (m, 3H), 4.03 (s, 1H), 3.98 – 3.67 (m, 2H), 3.50 (ddd, $J = 21.0, 18.4, 6.3$ Hz, 2H), 3.17 (d, $J = 10.4$ Hz, 1H), 3.10 – 2.99 (m, 1H), 1.90 – 1.65 (m, 2H), 1.42 (s, 9H), 1.32 – 1.27 (m, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 168.9, 168.6, 156.7, 136.8, 136.6, 128.7, 128.6, 128.3, 128.2, 128.2, 79.8, 77.2, 66.9, 66.7, 61.7, 53.1, 51.0, 50.0, 49.0, 47.5, 47.2, 41.0, 40.8, 37.8, 37.6, 33.6, 32.8, 28.5, 14.3, 14.2, 14.2. MS (MALDI-TOF) m/z calcd for $\text{C}_{21}\text{H}_{33}\text{N}_3\text{O}_6$ $[\text{M} + \text{K}]^+$ 462.2006, found 426.2247.

4.6.21 N-ethylamino(tertbutoxycarbonyl)-N(chloroacetamido)-2(R)-[(benzyloxycarbonyl) aminoethyl] ethyl glycinate (24)



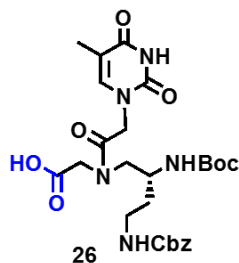
The compound **24** synthesized from **23** by following same procedure described for **10**. ^1H NMR (400 MHz, CDCl_3) δ 7.33 (dt, $J = 8.3, 4.0$ Hz, 5H), 7.02 (d, $J = 29.6$ Hz, 2H), 5.08 (q, $J = 12.8$ Hz, 2H), 4.41 (dd, $J = 39.3, 16.3$ Hz, 2H), 4.30 – 4.08 (m, 4H), 3.86 – 3.61 (m, 2H), 3.37 (dd, $J = 46.5, 9.9$ Hz, 2H), 3.20 – 2.85 (m, 4H), 2.05 (d, $J = 9.5$ Hz, 2H), 1.91 – 1.88 (m, 2H), 1.40 (d, $J = 8.1$ Hz, 9H), 1.25 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 169.2, 168.6, 167.6, 164.4, 156.9, 156.7, 151.5, 151.2, 141.3, 141.0, 136.8, 136.7, 128.6, 128.2, 110.96, 79.88, 77.2, 77.2, 62.5, 61.7, 60.5, 52.4, 51.7, 50.0, 47.9, 37.9, 33.2, 32.5, 28.5, 14.3, 14.2, 12.4. MS (MALDI-TOF) m/z calcd for $\text{C}_{23}\text{H}_{34}\text{ClN}_3\text{O}_6$ $[\text{M} + \text{K}]^+$ 538.1722, found 538.1591.

4.6.22 N-ethylamino(tertbutoxycarbonyl)-N(acetamido-N1-thyminy)-2(S)-[(benzyloxy carbonyl) aminoethyl] ethyl glycinate (25)



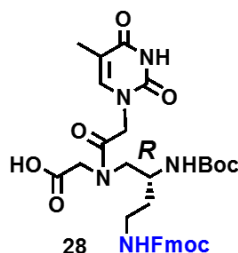
The compound **25** synthesized from **24** by following same procedure described for **11**. ^1H NMR (400 MHz, CDCl_3) δ 9.70 (d, $J = 120.7$ Hz, 1H), 7.35 – 7.27 (m, 5H), 7.00 (d, $J = 26.8$ Hz, 1H), 5.70 – 5.51 (m, 1H), 5.06 (q, $J = 12.5$ Hz, 2H), 4.79 (d, $J = 16.2$ Hz, 1H), 4.47 – 4.32 (m, 1H), 4.30 – 4.08 (m, 4H), 3.91 (d, $J = 17.2$ Hz, 1H), 3.76 (d, $J = 41.0$ Hz, 1H), 3.65 – 3.48 (m, 2H), 3.47 – 3.26 (m, 1H), 3.06 (ddd, $J = 19.7, 14.2, 6.1$ Hz, 3H), 2.15 (s, 1H), 1.87 (d, $J = 7.2$ Hz, 3H), 1.73 – 1.58 (m, 1H), 1.39 (d, $J = 9.8$ Hz, 9H), 1.26 (dt, $J = 21.1, 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 169.3, 168.6, 167.6, 164.5, 156.8, 156.7, 156.4, 156.2, 151.6, 151.3, 141.3, 141.0, 136.7, 136.7, 128.5, 128.1, 111.0, 80.0, 79.8, 66.7, 62.4, 61.6, 52.2, 51.59, 49.93, 48.71, 47.9, 47.6, 47.3, 47.1, 37.9, 37.6, 33.0, 32.3, 28.4, 14.2, 12.4 ppm. MS (MALDI-TOF) m/z calcd for $\text{C}_{28}\text{H}_{39}\text{N}_5\text{O}_9$ [$\text{M} + \text{K}$] $^+$ 628.2385, found 628.2361.

4.6.23 N-ethylamino(tertbutoxycarbonyl)-N(acetamido-N1-thyminy)-2(S)-[(benzyloxy carbonyl) aminoethyl] glycine (**26**):



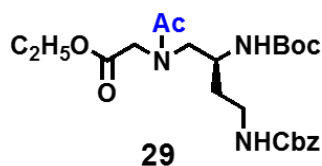
The compound **26** synthesized from **25** by following same procedure described for **12**. ^1H NMR (400 MHz, CDCl_3) δ 7.78 (d, $J = 8.4$ Hz, 2H), 7.66 – 7.50 (m, 2H), 7.35 (dd, $J = 28.0, 8.6$ Hz, 4H), 5.55 – 5.10 (m, 1H), 4.65 – 4.32 (m, 2H), 4.22 (s, 1H), 3.96 (d, $J = 14.0$ Hz, 1H), 3.64 (d, $J = 72.1$ Hz, 1H), 3.11 (s, 1H), 2.78 (d, $J = 46.7$ Hz, 1H), 2.08 – 1.86 (m, 1H), 1.84 – 1.61 (m, 2H), 1.46 (dd, $J = 16.2, 4.8$ Hz, 9H). ^{13}C NMR (101 MHz, DMSO) δ 172.3, 170.8, 170.5, 168.1, 167.50, 164.57, 156.20, 155.83, 151.19, 151.16, 142.08, 141.92, 137.42, 128.57, 127.96, 108.5, 108.4, 78.3, 78.0, 65.4, 51.3, 48.8, 47.9, 46.6, 39.5, 37.9, 32.0, 28.4, 21.3, 12.2 ppm. MS (MALDI-TOF) m/z calcd for $\text{C}_{26}\text{H}_{35}\text{N}_5\text{O}_9$ [$\text{M} + \text{K}$] $^+$ 600.2072, found 600.1787.

4.6.24 N-ethylamino(tert-butoxycarbonyl)-N(acetamido-N1-thyminy)-2S[(Fmoc) aminoethyl] glycine (**28**)



The compound **28** synthesized from **26** by following same procedure described for **14**. ^1H NMR (400 MHz, DMSO- d_6) δ 12.58 (s, 1H), 11.28 (d, $J = 3.0$ Hz, 1H), 7.88 (d, $J = 7.5$ Hz, 2H), 7.68 (d, $J = 6.5$ Hz, 2H), 7.41 (t, $J = 7.4$ Hz, 2H), 7.36 – 7.27 (m, 2H), 7.23 – 7.14 (m, 1H), 6.79 (dd, $J = 66.5, 9.0$ Hz, 1H), 4.58 (ddd, $J = 46.1, 28.4, 12.1$ Hz, 2H), 4.37 – 4.12 (m, 3H), 3.98 (dt, $J = 39.4, 12.3$ Hz, 1H), 3.51 – 3.15 (m, 4H), 3.05 (ddd, $J = 18.4, 13.5, 8.0$ Hz, 2H), 1.95 (d, $J = 30.2$ Hz, 1H), 1.77 – 1.69 (m, 2H), 1.37 (d, $J = 10.7$ Hz, 9H), 1.17 (t, $J = 7.1$ Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 172.0, 172.0, 170.6, 170.6, 170.4, 167.9, 167.3, 164.3, 156.1, 155.1, 155.6, 150.1, 143.1, 143.9, 143.9, 141.9, 141.7, 140.7, 127.6, 127.1, 125.1, 120.1, 108.2, 78.0, 77.8, 66.4, 65.3, 59.8, 47.6, 46.8, 46.3, 39.5, 37.5, 31.9, 28.3, 28.2, 21.1, 20.8, 14.1, 12 ppm. HRMS (ESI-TOF) m/z calcd. for $\text{C}_{28}\text{H}_{39}\text{N}_5\text{O}_9$ $[\text{M} + \text{Na}]^+$ 672.2645, found 672.2612.

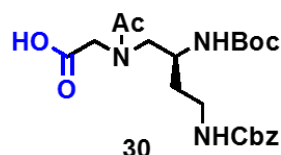
4.6.25 N(Acetyl)-N-ethylamino(tert-butoxycarbonyl)-2(S)-[(benzyloxycarbonyl) aminoethyl] ethyl glycinate (**29**)



To a stirred solution of **23** (1 g, 2.5 mmol) in dried DCM (10 mL) under an inert atmosphere were added DCC (0.7 g, 4 mmol) and DhbtOH (0.6 g, 4 mmol). The resulting mixture was stirred at room temperature for 10 min after that glacial acetic acid (0.2 mL, 4 mmol) was added to the above reaction mixture, and stirred overnight. The solvent was removed under reduced pressure, and the remaining residue was dissolved in EtOAc (300 mL) and washed with a saturated aq. NaHCO_3 (300 mL) followed by 10% aq. KHSO_4 (300 mL). The organic layer was washed with brine (3×150 mL) and dried over Na_2SO_4 . The solvent was removed under reduced pressure. The crude product was purified by column chromatography to afford a **29** (1.5 gm, 85%) as a sticky oil. ^1H NMR (400 MHz, CDCl_3) δ

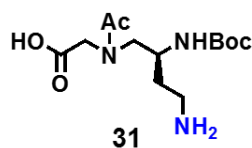
8.00 (s, 1H), 7.37 – 7.28 (m, 5H), 5.62 (s, 1H), 5.08 (q, $J = 12.4$ Hz, 2H), 4.19 (ddd, $J = 20.4$, 13.7, 7.0 Hz, 2H), 4.02 (d, $J = 3.0$ Hz, 1H), 3.91 – 3.64 (m, 2H), 3.43 (ddd, $J = 20.6$, 14.4, 6.6 Hz, 2H), 3.16 – 2.97 (m, 2H), 2.95 – 2.86 (m, 1H), 2.12– 2.03 (d, $J = 7.5$ Hz, 3H), 1.87 – 1.54 (m, 2H), 1.40 (d, $J = 5.1$ Hz, 9H), 1.26 (dt, $J = 11.2$, 7.1 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 172.8, 169.4, 157.0, 156.8, 137.2, 136.7, 128.8, 128.7, 128.4, 128.3, 128.3, 79.8, 79.7, 77.2, 67.0, 66.8, 62.1, 61.62, 54.14, 51.2, 50.7, 49.2, 48.0, 47.7, 38.0, 37.8, 33.7, 32.7, 28.6, 21.7, 21.5, 14.4, 14.4 ppm. HRMS (ESI-TOF) m/z calcd. for $\text{C}_{23}\text{H}_{35}\text{N}_3\text{O}_7$ $[\text{M} + \text{H}]^+$ 465.2475, found 465.5470.

4.6.26 N(Acetyl)-N-ethylamino(tertbutoxycarbonyl)-2(S)-[(benzyloxycarbonyl)aminoethyl] glycine (30)



To a stirred solution of compound **29** (0.70 g, 1.5 mmol) in THF (15 mL) was added aq. NaOH (2 M, 15 mL) at 0 °C, and the reaction mixture was stirred for 1 hr at 0 °C. Upon completion of the reaction as confirmed by TLC, H_2O (50 mL) was added, and resulting mixture was extracted with EtOAc (2 \times 35 mL). The aqueous layer was neutralized with saturated aq. KHSO_4 at 0 °C to pH 3–4 and then extracted with EtOAc (4 \times 35 mL). The combined organic extracts were dried over anhydrous Na_2SO_4 . The solvent was evaporated in vacuo, and crude product was purified by column chromatography to afford **30** as a white solid (0.63 g, 96%). ^1H NMR (400 MHz, CDCl_3) δ 7.38 – 7.28 (m, 5H), 5.50 (d, $J = 89.1$ Hz, 1H), 5.13 – 5.03 (m, 2H), 4.18 – 3.92 (m, 3H), 3.75 (t, $J = 23.4$ Hz, 2H), 3.46 (ddd, $J = 34.8$, 10.5, 7.1 Hz, 2H), 3.09 (dd, $J = 28.3$, 9.8 Hz, 1H), 2.17 – 2.02 (m, 3H), 1.83 – 1.60 (m, 2H), 1.38 (dd, $J = 17.1$, 6.5 Hz, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 173.0, 171.3, 157.0, 136.3, 128.3, 77.2, 66.7, 50.3, 47.7, 37.5, 29.7, 28.2, 21.4, 20.8 ppm. HRMS (ESI-TOF) m/z calcd. for $\text{C}_{21}\text{H}_{31}\text{N}_3\text{O}_7$ $[\text{M} + \text{Na}]^+$ 437.2162, found 460.2060.

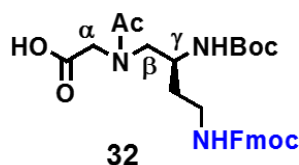
4.6.27 N(Acetyl)-N-ethylamino(tertbutoxycarbonyl)-2- (S)- [(Fmoc)aminoethyl] glycine (31)



The compound **31** synthesized from **30** by following same procedure described for **14** (Yield, 90 %). ^1H NMR (400 MHz, CDCl_3) δ 7.75 (d, $J = 7.5$ Hz, 2H), 7.59 (dd, $J = 11.1$, 7.9

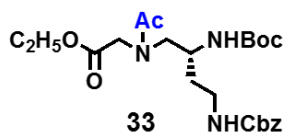
Hz, 2H), 7.38 (t, $J = 7.4$ Hz, 2H), 7.29 (t, $J = 7.4$ Hz, 2H), 4.48 – 4.28 (m, 2H), 4.24 – 4.15 (m, 1H), 4.14 – 3.95 (m, 3H), 3.84 – 3.73 (m, 1H), 3.46 – 3.33 (m, 1H), 3.18 – 3.02 (m, 1H), 2.16 (d, $J = 7.4$ Hz, 3H), 2.06 (m, $J = 12.6$ Hz, 2H), 1.41 (d, 9H). ^{13}C NMR (101 MHz, CDCl_3) δ 173.2, 171.4, 144.1, 141.5, 127.9, 127.1, 125.3, 125.2, 120.1, 80.0, 77.2, 67.0, 50.8, 47.3, 37.79, 33.41, 29.8, 28.4, 21.5, 21.5 ppm. HRMS (ESI-TOF) m/z calcd. for $\text{C}_{28}\text{H}_{35}\text{N}_3\text{O}_7$ $[\text{M} + \text{Na}]^+$ 525.2475, found 548.2373.

4.6.28N(Acetyl)-N-ethylamino(tertbutoxycarbonyl)-2(S)-[(benzyloxycarbonyl)aminoethyl] ethyl glycinate (32):



The compound **32** synthesized from **31** by following same procedure described for **14**. (Yield, 85%). ^1H NMR (400 MHz, CDCl_3) δ 7.37 – 7.27 (m, 5H), 5.63 (s, 1H), 5.13 – 5.03 (m, 2H), 4.27 – 4.10 (m, 3H), 4.08 – 3.96 (m, 2H), 3.88 (d, $J = 17.1$ Hz, 3H), 3.82 – 3.62 (m, H), 3.51 – 3.35 (m, 2H), 3.17 – 2.83 (m, 2H), 2.07 (d, $J = 41.0$ Hz, 3H), 1.91 – 1.55 (m, 3H), 1.40 (d, $J = 4.9$ Hz, 9H), 1.26 (dt, $J = 11.2, 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 172.6, 171.5, 170.1, 169.2, 156.8, 156.0, 136.8, 136.6, 128.6, 128.6, 128.2, 128.9, 128.1, 80.1, 79.59, 77.16, 66.8, 66.6, 61.9, 61.5, 51.0, 50.5, 47.4, 37.9, 37.6, 33.5, 32.6, 28.5, 21.5, 21.30, 14.3, 14.2 ppm. HRMS (ESI-TOF) m/z calcd. for $\text{C}_{23}\text{H}_{35}\text{N}_3\text{O}_7$ $[\text{M} + \text{Na}]^+$ 465.2475, found 488.2373.

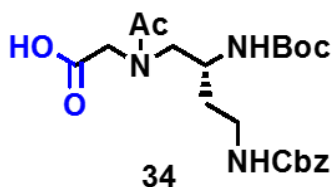
4.6.29N(Acetyl)-N-ethylamino(tertbutoxycarbonyl)-2(R)-[(benzyloxycarbonyl)aminoethyl] ethyl glycinate (33)



The compound **33** synthesized from **23** by following same procedure described for **29**. (1.5 gm, 85%) as a sticky oil. ^1H NMR (400 MHz, CDCl_3) δ 7.37 – 7.27 (m, 5H), 5.63 (s, 1H), 5.13 – 5.03 (m, 2H), 4.27 – 4.10 (m, 3H), 4.08 – 3.96 (m, 2H), 3.88 (d, $J = 17.1$ Hz, 3H), 3.82 – 3.62 (m, H), 3.51 – 3.35 (m, 2H), 3.17 – 2.83 (m, 2H), 2.07 (d, $J = 41.0$ Hz, 3H), 1.91 – 1.55 (m, 3H), 1.40 (d, $J = 4.9$ Hz, 9H), 1.26 (dt, $J = 11.2, 7.1$ Hz, 3H). $^{13}\text{C}\{^1\text{H}\}$ NMR (101 MHz, CDCl_3) δ 172.6, 171.5, 170.1, 169.2, 156.8, 136.8, 128.6, 128.1, 80.1, 79.6, 77.2,

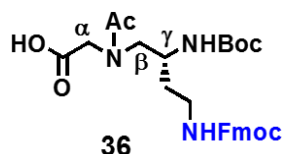
66.6, 61.9, 61.5, 51.0, 50.5, 47.4, 37.9, 37.6, 33.5, 32.6, 28.5, 21.5, 14.2 ppm. HRMS (ESI-TOF) m/z calcd. for $C_{23}H_{35}N_3O_7$ 466.2475, found 466.3684 $[M + H]^+$.

4.6.30 N(Acetyl)-N-ethylamino(tertbutoxycarbonyl)-2(R)-[(benzyloxycarbonyl)aminoethyl] glycine (34)



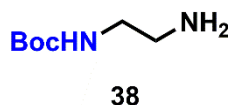
The compound **34** synthesized from **33** by following same procedure described for **30**. (yield, 96%). 1H NMR (400 MHz, $CDCl_3$) δ 7.31 (dt, $J = 12.0, 4.1$ Hz, 5H), 5.09 (t, $J = 9.6$ Hz, 2H), 4.21 – 3.86 (m, 2H), 3.74 (t, $J = 18.5$ Hz, 5H), 3.44 (d, $J = 32.9$ Hz, 1H), 3.17 – 3.00 (m, 8H), 2.15 – 2.02 (m, 3H), 1.84 – 1.54 (m, 2H), 1.48 – 1.34 (m, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.0, 171.3, 157.0, 136.3, 128.3, 77.2, 66.7, 50.3, 47.7, 37.5, 29.8, 28.2, 21.4, 20.8 ppm. HRMS (ESI-TOF) m/z calcd. for $C_{20}H_{29}N_3O_7$ $[M + Na]^+$ 437.2162, found 460.2060.

4.6.31 N(Acetyl)-N-ethylamino(tertbutoxycarbonyl)-2- (R)- [(Fmoc)aminoethyl] glycine (36)



The compound **36** synthesized from **34** by following same procedure described for **14**. (yield, 96%). 1H NMR (400 MHz, $CDCl_3$) δ 7.74 (d, $J = 7.5$ Hz, 2H), 7.64 – 7.50 (m, 2H), 7.38 (t, $J = 7.4$ Hz, 2H), 4.81 – 4.28 (m, 7H), 4.24 – 3.89 (m, 1H), 3.84 – 3.63 (m, 2H), 3.55 – 3.28 (m, 5H), 3.27 – 2.87 (m, 6H), 2.19 – 1.98 (m, 3H), 1.82 – 1.47 (m, 2H), 1.36 (t, $J = 32.1$ Hz, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 175.4, 141.1, 127.5, 126.9, 119.7, 77.2, 47.1, 28.1, 20.4 ppm. HRMS (ESI-TOF) m/z calcd. for $C_{28}H_{35}N_3O_7$ $[M + Na]^+$ 525.2475, found 548.2373.

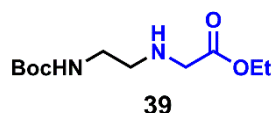
4.6.32 N-(tertbutoxycarbonyl)-1, 2-diaminoethane (38)



To an ice-cold stirred solution of 1,2-diaminoethane **37** (2.0 g, 33 mmol) in DCM (300 ml) was added solution of Boc-anhydride (5 g, 35 mmol) in DCM (50 ml) solution. The mixture

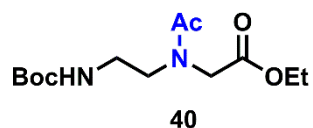
was stirred for 12 h and the resulting solution was concentrated to 100 ml. The *N*₁,*N*₂-di-Boc derivative not being soluble in water, precipitated out and was removed by filtration. The corresponding *N*-mono-Boc derivative was obtained by repeated extraction from the filtrate in dichloromethane. Removal of solvents yielded the mono-Boc-diaminoethane **38** (3.5 g, 60%) which was used for further reaction without any purification.

4.6.33 N-(tert-butoxycarbonyl)aminoethyl ethyl glycinate (**39**)



The *N*-(Boc)-1,2-diaminoethane **38** (3 g, 20 mmol) was treated with ethyl bromoacetate (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 × 50 ml), followed by washing with brine. The collected organic layer was dried over anhydrous sodium sulphate and filtered and purified with flash column to afford (3.5 g, 14 mmol) compound **39** (yield 70 %), IR (neat) 3887, 3735, 3388, 2973, 2100, 1697, 1510, 1456, 1362, 1246, 1162, 1084, 1046, 1002, 864, 735, 699, 607 cm⁻¹; ¹H NMR (400 MHz, CDCl₃) δ 5.1 (s, 1H), 4.2 (m, 2H), 3.4 (s, 2H), 3.21 (d, *J* = 4 Hz, 2H), 2.74 (t, 2H) 1.44 (d, *J* = 31.0 Hz, 9H), 1.28 (t, 3H); ¹³C NMR (101 MHz, CDCl₃) δ 173, 156, 79.4, 61.0, 50.6, 48.9, 40.3, 28.6, 14.4 ppm; HRMS (ESI-TOF) *m/z* calcd for C₁₁H₂₂N₂O₄ [M + H]⁺ 246.3119, found 246.3070.

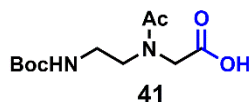
4.6.34 N-(acetyl)-N-(tert-butoxycarbonyl)aminoethyl ethyl glycinate (**40**)



To a stirred solution of **39** (0.7 g, 3 mmol) in dried DCM (10 mL) under an inert atmosphere were added DCC (0.87 g, 4.2 mmol) and DhbtOH (0.7 g, 4 mmol). The resulting mixture was stirred at room temperature for 10 min after that glacial acetic acid (0.2 ml, 4 mmol) was added to the above reaction mixture, and stirred overnight. The solvent was removed under reduced pressure, and the remaining residue was dissolved in EtOAc (300 mL) and washed with a saturated aq. NaHCO₃ (300 mL) followed by 10% aq. KHSO₄ (300 mL). The organic layer was washed with brine (3×150 mL) and dried over Na₂SO₄. The

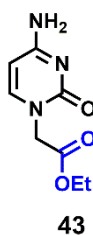
solvent was removed under reduced pressure. The crude product was purified by column chromatography to afford a **40** (0.6 gm, 75%). TLC in EtOAc 4:1, $R_f = 0.26$; HRMS (ESI-TOF) m/z calcd for $C_{13}H_{24}N_2O_5$ $[M + H]^+$ 288.1685, found 288.2908.

4.6.35 N-(acetyl)-N-(tertbutoxy carbonyl)aminoethyl) glycine (**41**)



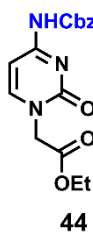
Compound **41** (1 g, 3.0 mmol) was prepared, purified, and characterized the same way as describe for compound **32**, starting from compound **40** (1 g, 3 mmol), to afford yield (99%). 1H NMR (400 MHz, $CDCl_3$) δ 5.61 (s, 1H), 4.02 (d, $J = 23.6$ Hz, 4H), 3.49 (s, 4H), 3.24 (s, 4H), 2.14 (d, $J = 12.0$ Hz, 4H), 2.04 (s, 2H), 1.43 (d, $J = 17.2$ Hz, 19H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.9, 172.6, 156.3, 79.9, 77.2, 50.4, 49.0, 38.8, 28.5, 20.9 ppm; HRMS (ESI-TOF) m/z calcd for $C_{11}H_{20}N_2O_5$ $[M + K]^+$ 260.1372, found 299.1009.

4.6.36 Ethyl cytosine-1-yl acetate (**43**)



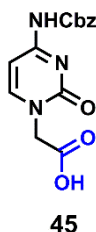
To a suspension of cytosine **42** (5.0 g, 45.0 mmol) in dry DMF (100 ml) was added NaH (1.0 g, 45.0 mmol) under N_2 at 0 °C. The mix was stirred for 2 h at rt then ethyl bromoacetate (4.3 ml, 45.0 mmol) was added. The mix was stirred for 48 h at rt. The solvent was then evaporated in vacuo. The crude residue was triturated with water (100 ml) and the precipitate was filtered off. Recrystallization from MeOH/ H_2O gave of compound **43** (5.3 g, 27 mmol 60%) as white crystals (mp: 225 – 227 °C); HRMS (ESI-TOF) m/z calcd for $C_8H_{12}N_3O_3$ $[M + H]^+$ 198.0878, found 198.0887.

4.6.37 Ethyl (N^4 -benzyloxycarbonyl) cytidin-1-yl acetate (**44**)



Cbz-Cl (3 ml, 22 mmol) and DMAP (2.7 g, 22 mmol) were dissolved at -15 °C in DCM (20 ml). The mixture was stirred for 15 min and compound **43** (2.0 g, 10 mmol) was gradually added. After stirring for 15 min at -15 °C then 5 h at rt, the mixture was evaporated in vacuo and the crude residue was taken up in CHCl₃. The organic layer was washed with 1 M HCl, with water then dried over Na₂SO₄ and lastly evaporated in reduced pressure. Trituration of the crude residue in ether gave a white precipitate which was then filtered off, affording **44** (2.6 g, 7.9 mmol 77%) as an amorphous powder. TLC (EtOAc/MeOH 1:1) R_f = 0.79. ¹H NMR (400 MHz, DMSO-d₆) δ 10.82 (s, 1H), 8.04 (d, *J* = 7.3 Hz, 1H), 7.45 – 7.33 (m, 5H), 7.04 (d, *J* = 7.2 Hz, 1H), 5.19 (s, 2H), 4.60 (s, 2H), 4.14 (q, *J* = 7.1 Hz, 2H), 1.20 (t, *J* = 7.1 Hz, 3H); ¹³C NMR (101 MHz, DMSO-D₆) δ 168.0, 163.5, 155.1, 153.2, 150.5, 136.0, 128.6, 128.3, 128.0, 94.2, 66.6, 61.2, 50.7, 14.1 ppm.

4.6.38 N⁴(Benzyloxycarbonyl) cytidin-1-yl acetic acid (**45**)



Compound **44** (2.0 g, 6.0 mmol) was dissolved in dioxane (50 ml) and aqueous 1 M NaOH (9 ml) was added. The mix was stirred 5 h at rt and then concentrated under reduced pressure. The residue was taken up in a aqueous KHSO₄ (1 M). The resultant white precipitate was isolated by filtration, washed with water and dried in a vacuum desiccator affording **45** (1.6 g, 5 mmol, 87%). TLC (EtOAc/MeOH 1:1) R_f = 0.10. ¹H NMR (400 MHz, DMSO-d₆) δ 10.84 (s, 1H), 8.04 (d, *J* = 7.3 Hz, 1H), 7.46 – 7.31 (m, 5H), 7.02 (d, *J* = 7.3 Hz, 1H), 5.20 (s, 2H), 4.53 (s, 2H); ¹³C NMR (101 MHz, DMSO-D₆) δ 169.4, 163.4, 155.1, 153.2, 150.5, 136.0, 128.5, 128.2, 128.0, 94.1, 66.6, 50.6 ppm; HRMS (ESI-TOF) *m/z* calcd for C₁₄H₁₄N₃O₅ [M + H]⁺ 304.0933, found 304.0934.

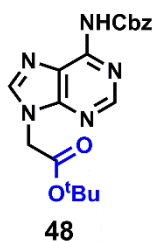
4.6.39 N⁶(Benzyloxycarbonyl)adenine (**47**)



Sodium hydride (6 g, 0.25 mmol, 60% dispersion in oil) was washed with petroleum ether (3 x). After cooling in an ice bath, dry DMF (150 ml) was added, followed by adenine

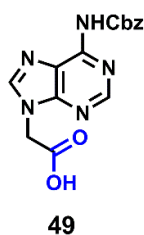
46 (5.0 g, 37 mmol) in small portions. The suspension was stirred vigorously for 3 min and then benzyl chloroformate (11.5 mL, 81.5 mmol) was added dropwise. After stirring for 4 h the reaction mixture was poured into ice water (300 mL) and the pH adjusted to 7 with aq. HCl (1 N). The light yellow precipitate was collected by filtration and washed with water, and ether to afford crude product (8.5 g). Recrystallization from methanol/chloroform afforded two crops of **47** (4 g, 16.0 mmol 44%) as a white solid: m.p = 222 °C; ¹H NMR (400 MHz, DMSO-d⁶) δ 12.25 (s, 1H), 11.05 (s, 1H), 8.59 (s, 1H), 8.44 (s, 1H), 7.61 – 7.31 (m, 5H), 5.29 (s, 2H); ¹³C NMR (101 MHz, DMSO-d⁶) δ 162.3, 154.2, 151.7, 146.5, 145.0, 136.4, 129.0, 128.8, 128.6, 113.3, 67.4 ppm; HRMS (ESI-TOF) m/z calcd for C₁₃H₁₂N₅O₂ [M + H]⁺ 270.0991, found 270.0981.

4.6.40 N⁶(Benzyloxycarbonyl)-N⁹-(methyl-*t*-butylcarboxy) adenine (**48**)



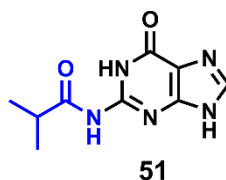
To a suspension of **47** (0.5 g, 2 mmol) in dry DMF (4 ml) was added anhydrous Cs₂CO₃ (0.06 g, 0.2 mmol) and anhydrous K₂CO₃ (0.3g, 2 mmol). After stirring for 5 min, tert-butyl bromoacetate (0.3 ml, 2.0 mmol) was added dropwise. After stirring for ca.23 h, the mixture was evaporated to dryness in vacuo and the residue partitioned between ethyl acetate (35 ml) and water (10 ml). The organic phase was washed with water and brine, dried (Na₂SO₄) and concentrated in vacuo. The residue was crystallized from ethyl acetate/hexane to afford **48** (0.4 g, 1.0 mmol 55%) as a white solid: m.p = 141-142 °C; ¹H NMR (400 MHz, CDCl₃) δ 8.77 (s, 1H), 8.58 (s, 1H), 7.98 (s, 1H), 7.47 – 7.33 (m, 5H), 5.30 (s, 2H), 4.86 (s, 2H), 1.47 (s, 9H); ¹³C NMR (101 MHz, CDCl₃) δ 165.9, 153.1, 151.7, 151.0, 149.5, 143.4, 135.6, 128.8, 128.8, 128.7, 121.5, 84.0, 67.9, 45.0, 28.1 ppm; HRMS (ESI-TOF) m/z calcd for C₁₉H₂₂N₅O₄ [M + H]⁺ 384.1672, found 384.1670.

4.6.41 N⁶(Benzyloxycarbonyl)-N⁹-(carboxymethyl) adenine (**49**):



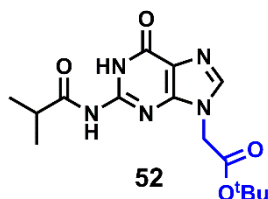
A solution of **48** (0.15 g, 0.4 mmol) in dry DCM (3.0 ml) was treated with triethylsilane (0.6 ml), cooled to 0 °C and TFA (1.5 ml) was added slowly. After 5 min the reaction was allowed to warm to rt. After 7.5 h, the mixture was evaporated to dryness in vacuo and any remaining volatiles removed by azeotroping with chloroform (3 x) to give **49** (0.12 g, 0.4 mmol, 97%) as a white foam. An analytically pure sample of **49** was obtained by recrystallization from acetone: m.p = 139-143 °C; ¹H NMR (400 MHz, DMSO-d⁶) δ 13.39 (s, 1H), 10.72 (s, 1H), 8.62 (s, 1H), 8.44 (s, 1H), 7.49 – 7.29 (m, 5H), 5.22 (s, 2H), 5.09 (s, 2H); ¹³C NMR (101 MHz, DMSO-d⁶) δ 169.0, 152.2, 151.6, 149.4, 144.8, 136.3, 128.4, 128.2, 128.0, 127.8, 122.9, 66.3, 44.3 ppm; HRMS (ESI-TOF) m/z calcd for C₁₅H₁₄N₅O₄ [M + H]⁺ 328.1046, found 328.1046.

4.6.42 N²(Isobutyryl) guanine (**51**)



Isobutyric acid anhydride (21 g, 0.13 mol) was added at room temperature to a suspension of guanine **50** (7.5 g, 50.3 mmol) in N, N-dimethyl acetamide (100 ml). The reaction mixture was heated at 150 °C for 2 hours, cooled to room temperature and evaporated under reduced pressure to 1/10 of its volume. The precipitated crude product was collected and crystallised from boiling ethanol:water (1:1, 150 ml) to yield **51** (10. g, 45 mmol, 89%) as colorless crystals, and used for next step, without any purification.

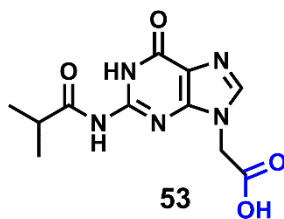
4.6.43 N²(Isobutyryl)-N⁹-(methyl-t-butylcarboxy) guanine (**52**):



To a suspension of N²-isobutyrylguanine **51** (6.5 g, 30.0 mmol) in dry DMF (250 ml) was added NaH (0.7 g, 30.0 mmol) at 0 °C. The mixture was stirred for 1 h and then tert-butyl bromoacetate (6.5 g, 33.0 mmol) was added dropwise over a period of 30 min at 0 °C. The reaction was stopped after 2 h by the addition of a small amount of solid CO₂ and methanol (5 ml). The reaction mixture was evaporated in vacuo and the residue was treated with dichloromethane and water. The water phase was re-extracted with dichloromethane. The dichloromethane fractions were combined and concentrated in vacuo to give the crude

product. The resulting crude product was recrystallized from ethyl acetate to give the desired N9-isomer product **52** with a yield (6.5 g 19.40 mmol 64%) mp = 204–205 °C (decomp.); R_f = 0.13 (dichloromethane/methanol, 9.5:0.5); $^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 12.10 (s, 1H), 11.65 (s, 1H), 7.95 (s, 1H), 4.88 (s, 2H), 2.86 – 2.69 (m, 1H), 1.42 (s, 9H), 1.11 (d, J = 6.8 Hz, 6H); $^{13}\text{C NMR}$ (101 MHz, DMSO- d_6) δ 180.2(s), 166.7, 154.8, 148.9, 148.1, 140.3, 119.6, 82.3, 44.8, 34.7), 27.7, 18.9 ppm; HRMS (ESI-TOF) m/z calcd for $\text{C}_{15}\text{H}_{22}\text{N}_5\text{O}_4$ [$\text{M} + \text{H}$] $^+$ 336.1672, found 336.1671.

4.6.44 N²-(isobutyryl)-N9-(carboxymethyl) guanine (**53**):



To a suspension of **52** (2.0 g, 6.0 mmol) dry dichloromethane (10 ml) was added triethylsilane (5 ml, 30 mmol). The mixture was cooled to 0 °C and TFA (15 ml) was then added over a period of 5 min. After 30 min the reaction was allowed to warm to room temperature. Upon completion of the reaction (traced by TLC), the mixture was concentrated in vacuo to give a foam which was stirred with diethyl ether, filtered, washed with diethyl ether, and dried in vacuo to afford the desired product **53** with a yield (1.7 g, 6.0 mmol, 100%). R_f = 0.44 (n-butanol/acetic acid/ H_2O , 3:1:1); $^1\text{H NMR}$ (400 MHz, DMSO) δ 12.08 (s, 1H), 11.67 (s, 1H), 7.95 (s, 1H), 4.90 (s, 2H), 2.86 - 2.69 (m, 1H), 1.11 (d, J = 6.8 Hz, 6H); $^{13}\text{C NMR}$ (101 MHz, DMSO- d_6) δ 180.2, 169.1, 154.8, 149.0, 148.1, 140.3, 119.7, 44.4, 34.7, 18.9 ppm; HRMS (ESI-TOF) m/z calcd for $\text{C}_{11}\text{H}_{14}\text{N}_5\text{O}_4$ [$\text{M} + \text{H}$] $^+$ 280.1046, found 280.1049.

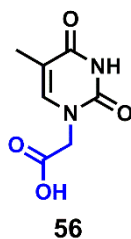
4.6.45 Ethyl 2-(thyminy)acetate (**55**).



Thymine **54** (1 g, 8 mmol) and potassium carbonate (1 g, 8 mmol) were dissolved in anhydrous DMF (6 mL) under N_2 . To this mixture, ethyl bromoacetate (1 mL, 8 mmol) was added dropwise. The reaction mixture was stirred overnight at room temperature and dissolved in water. The product was extracted by ethyl acetate (15 mL \times 3). The combined

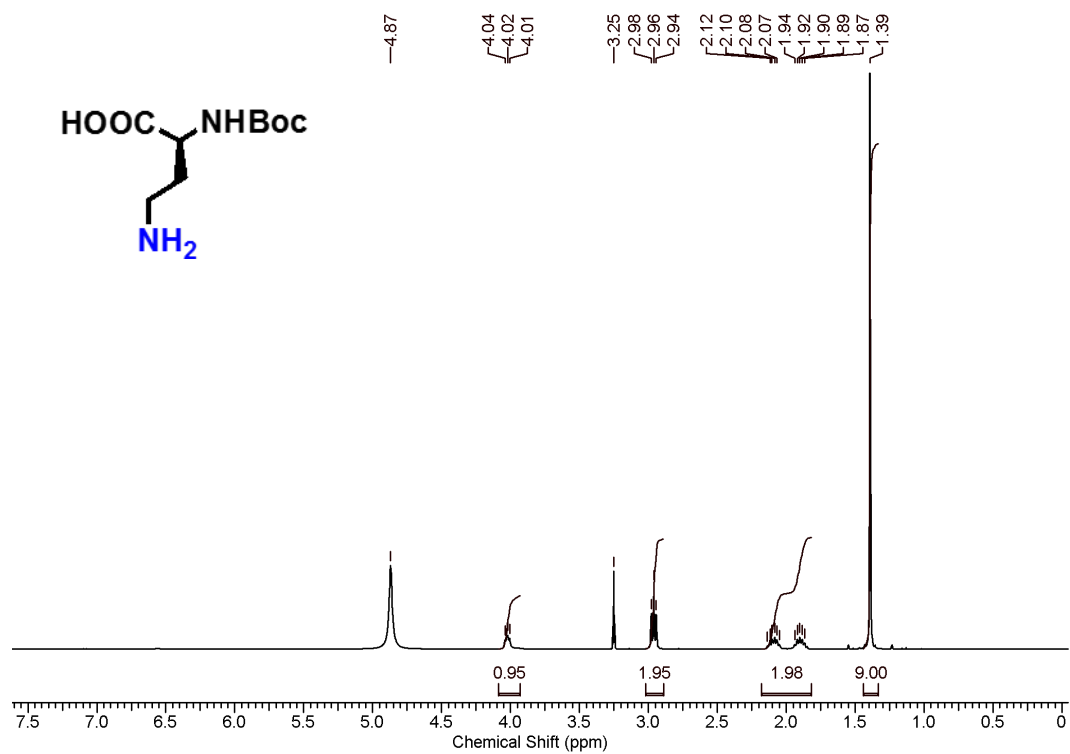
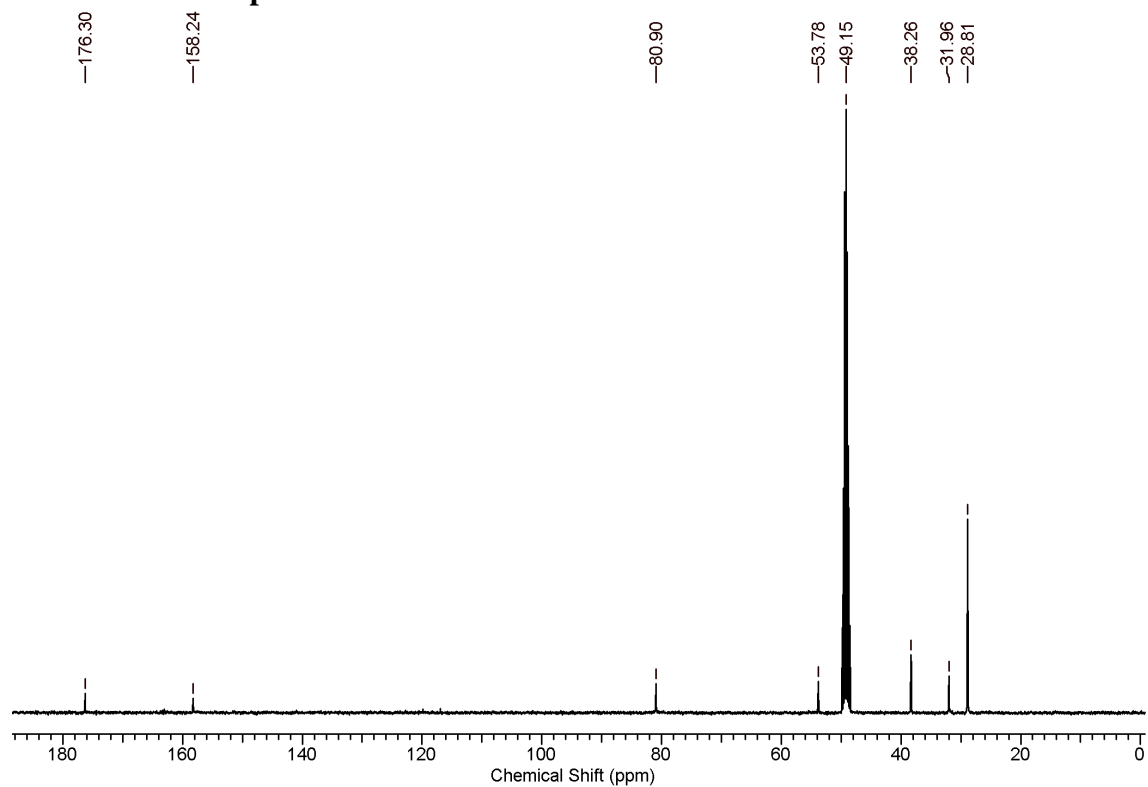
organic phase was washed with saturated aqueous KHSO₄ solution (10 mL × 2), brine (5 mL × 2) and dried over MgSO₄. Solvent was removed under reduced pressure to get compound **55** (1.30 g) as a white solid with 78% isolated yield. ¹H NMR (300 MHz, CDCl₃): δH 8.92 (s, 1H), 6.91 (s, 1H), 4.41 (s, 2H), 4.26-4.19 (q, 2H), 1.91 (s, 3H), 1.30-1.25 (t, 3H); ¹³C NMR (100 MHz, CDCl₃): δC 167.8, 164.7, 151.1, 140.3, 110.6, 61.8, 48.5, 14.0, 12.9 ; ESI-MS: *m/z*: calcd. for C₉H₁₂N₂O₄ [M+H]⁺; 213.08, found: 213.10.

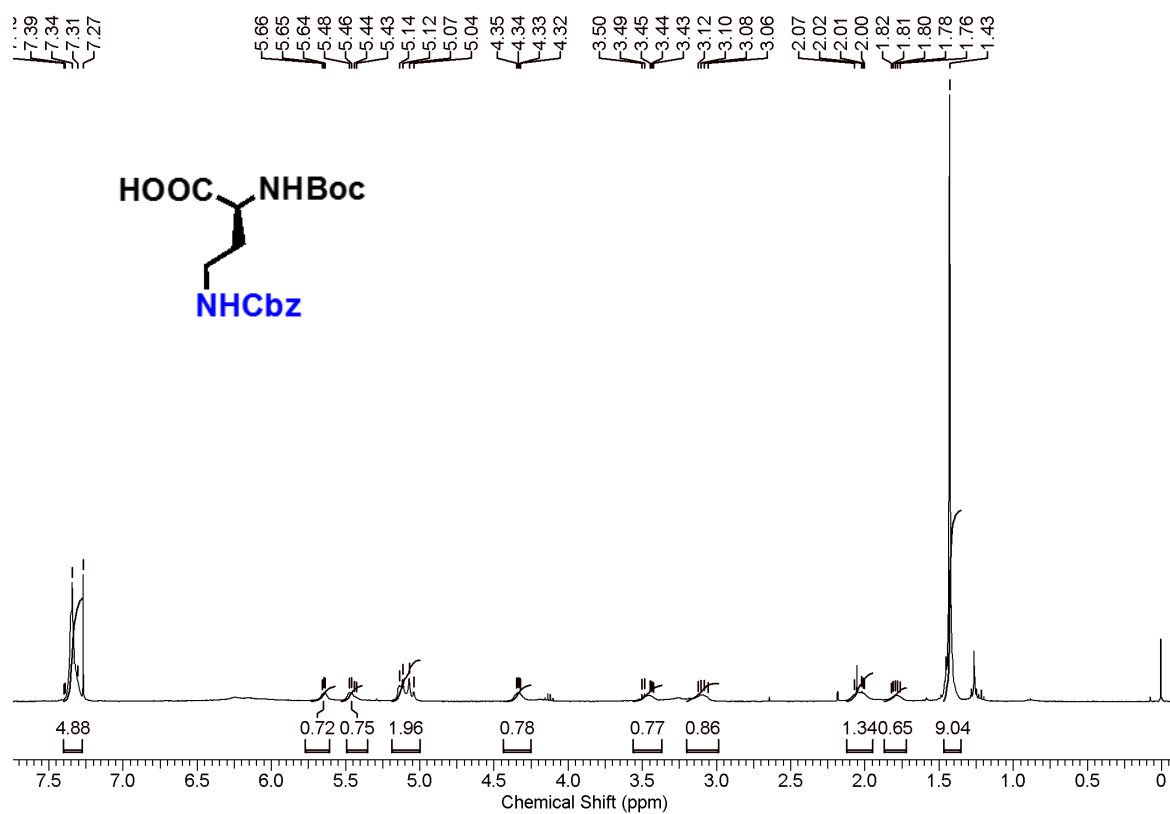
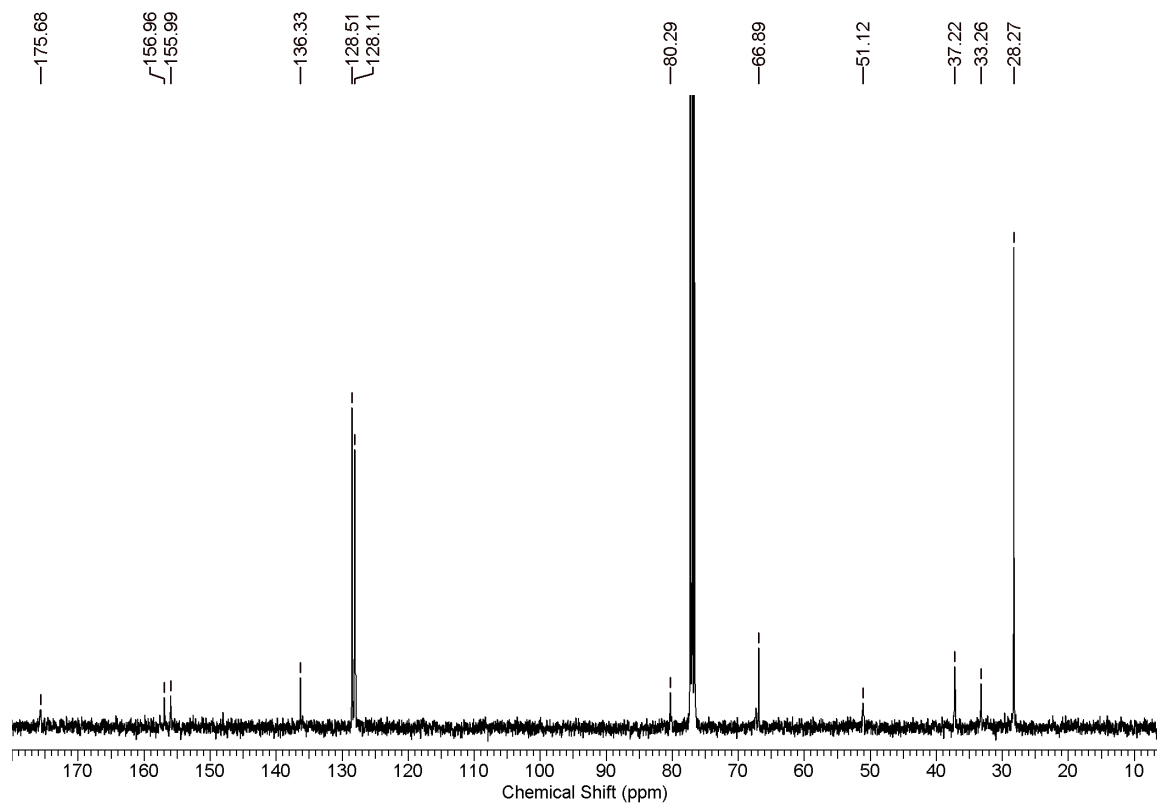
4.6.46 Thymin-1-yl acetic Acid (**56**):

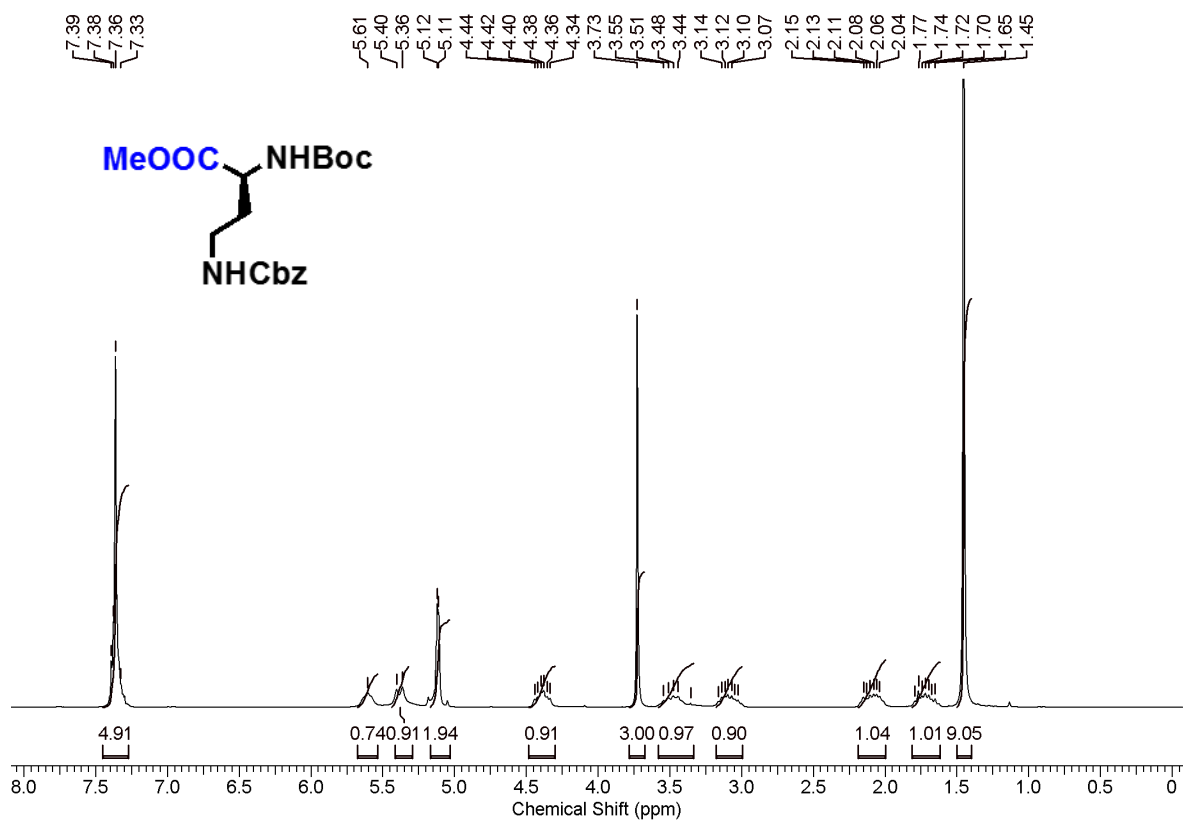
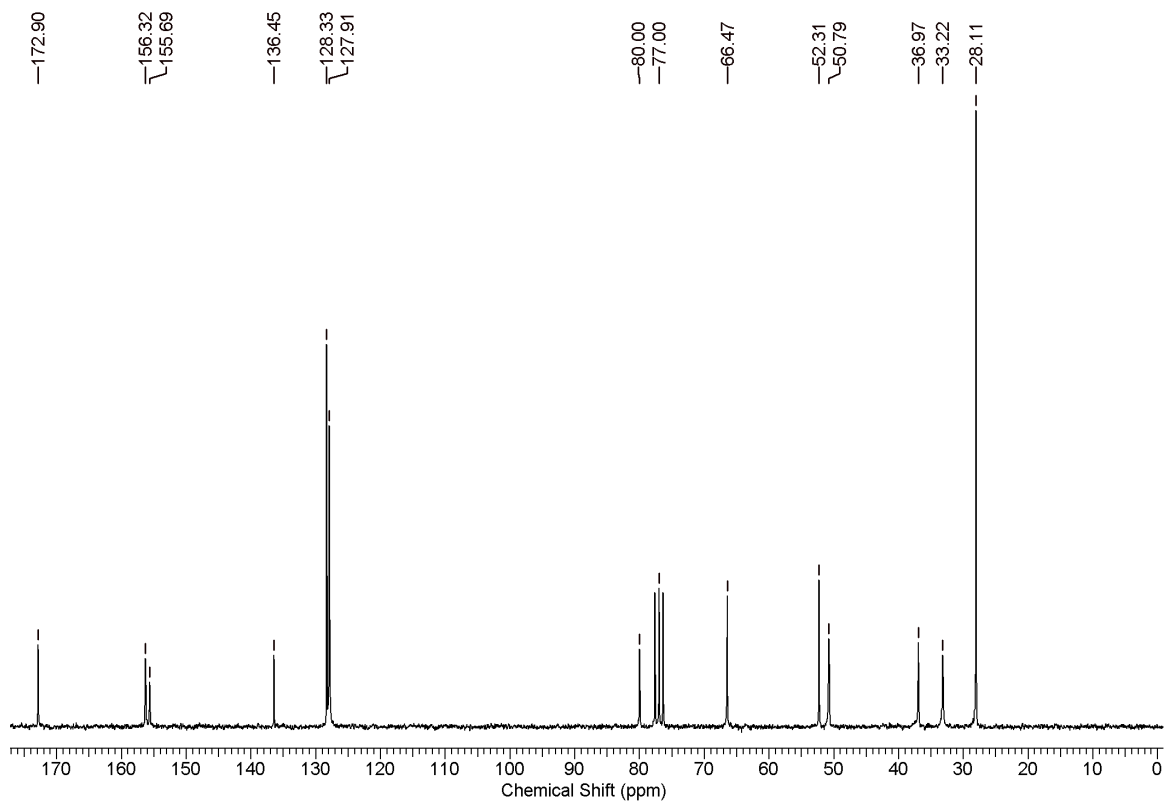


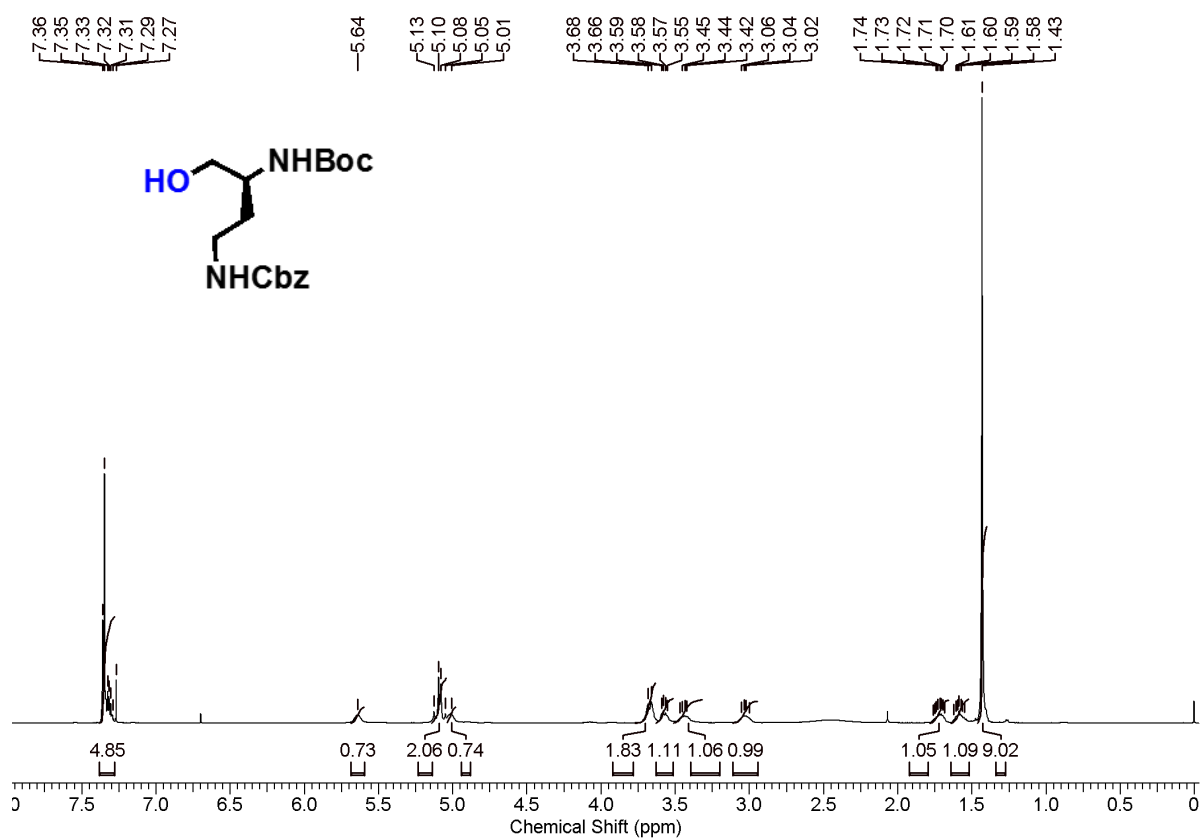
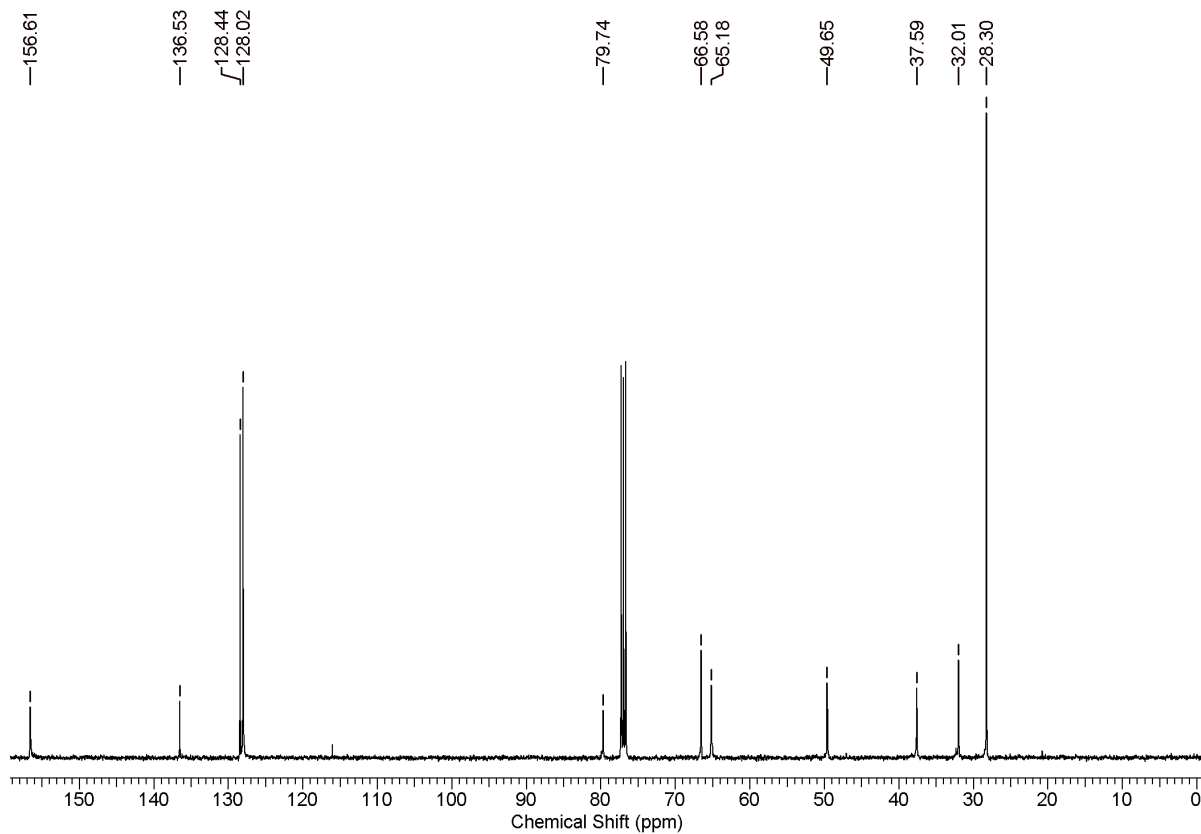
Compound **55** (2.0 g, 9.4 mmol) was dissolved in THF (30 mL) and stirred in ice bath. Aqueous 1 M LiOH was added dropwise. The reaction was left stirred for 0.5 h at room temperature and the completion of the reaction was checked by TLC. The reaction mixture was then neutralized by 2 M aqueous HCl. Product was obtained as white precipitate after filtration by using a Büchner funnel. The product was washed with water and dried in a vacuum desiccator to get **56** (2.6 g) as white solid, yield (1.6 g, 70%); ¹H NMR (400 MHz, DMSO-*d*⁶) δ 13.1 (s, 1H), 11.3 (s, 1H), 7.5 (d, *J* = 1.1 Hz, 1H), 4.4 (s, 2H), 1.8 (d, *J* = 0.9 Hz, 3H). HRMS (ESI-TOF) *m/z* calcd for C₇H₉N₂O₄ [M + H]⁺ 185.0562, found 185.0566.

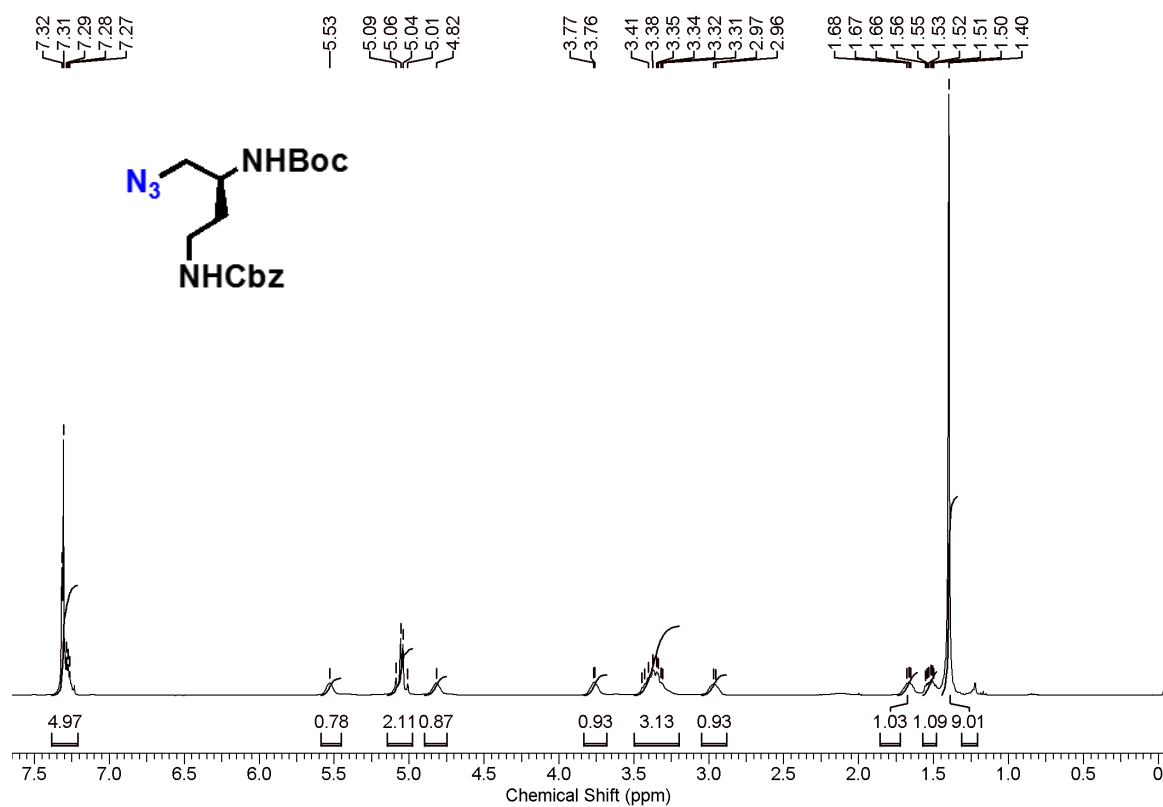
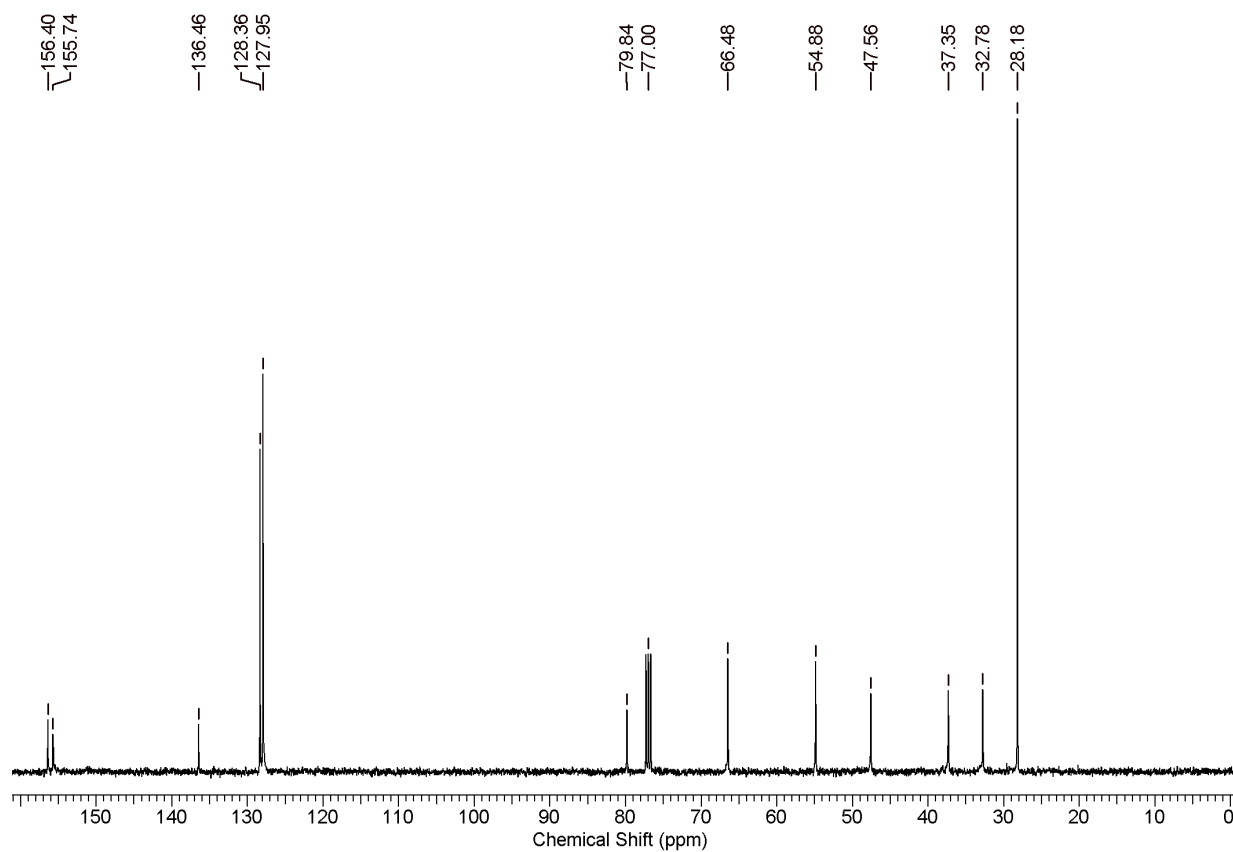
4.7 Characterization data of synthesized compounds

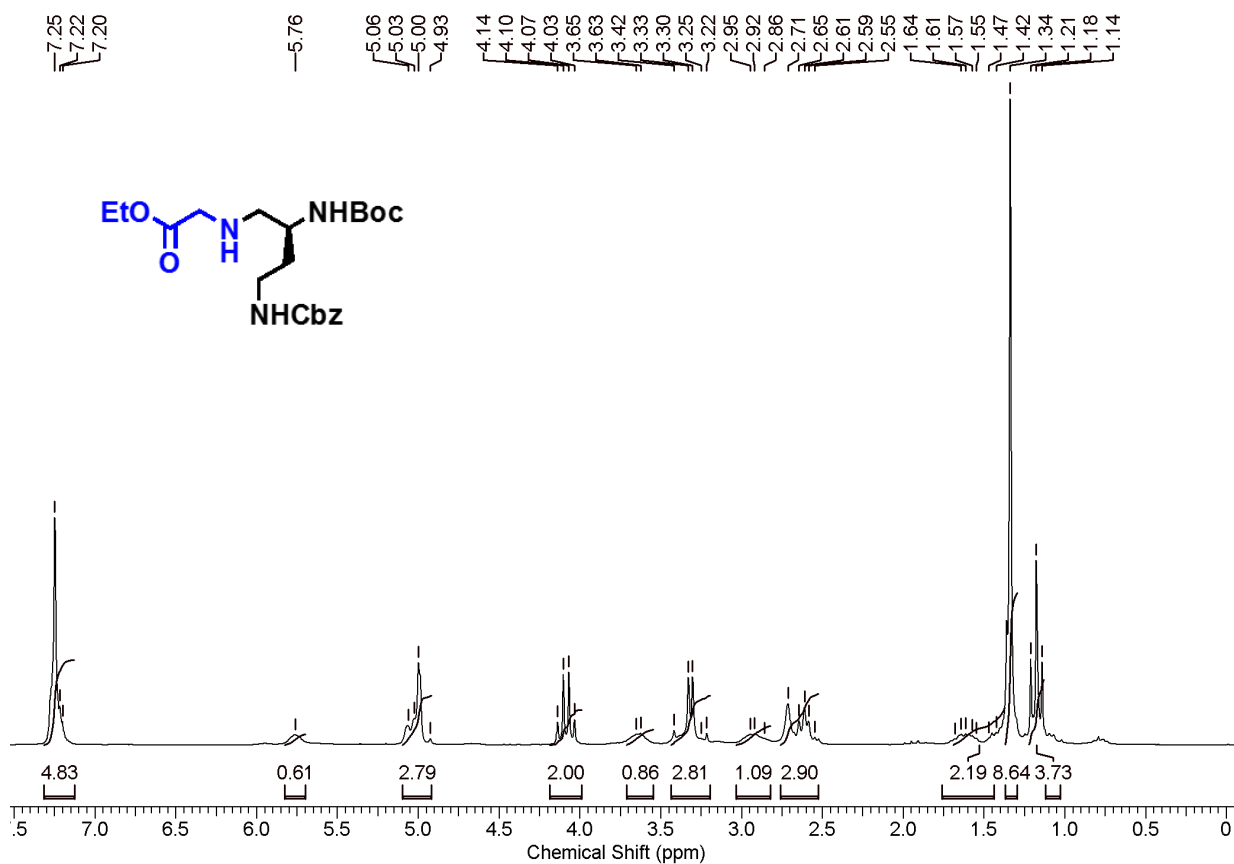
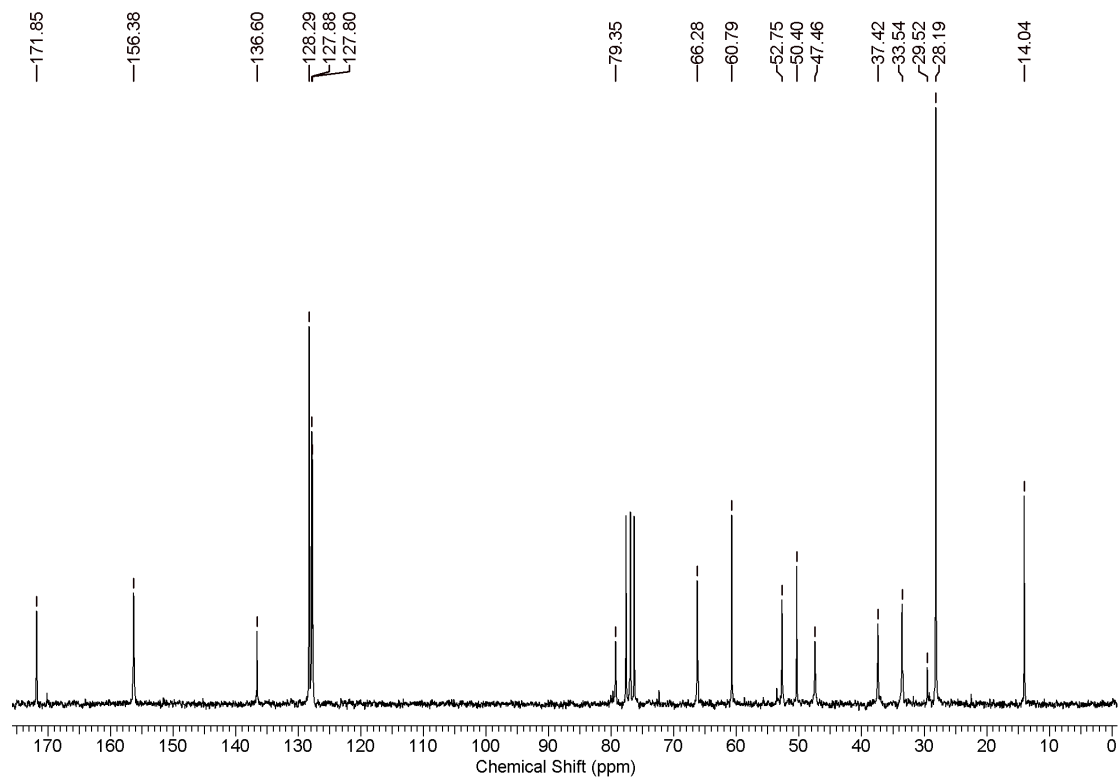
 ^1H and ^{13}C NMR of Compounds. ^1H NMR of Compound 3 ^{13}C NMR of Compound 3

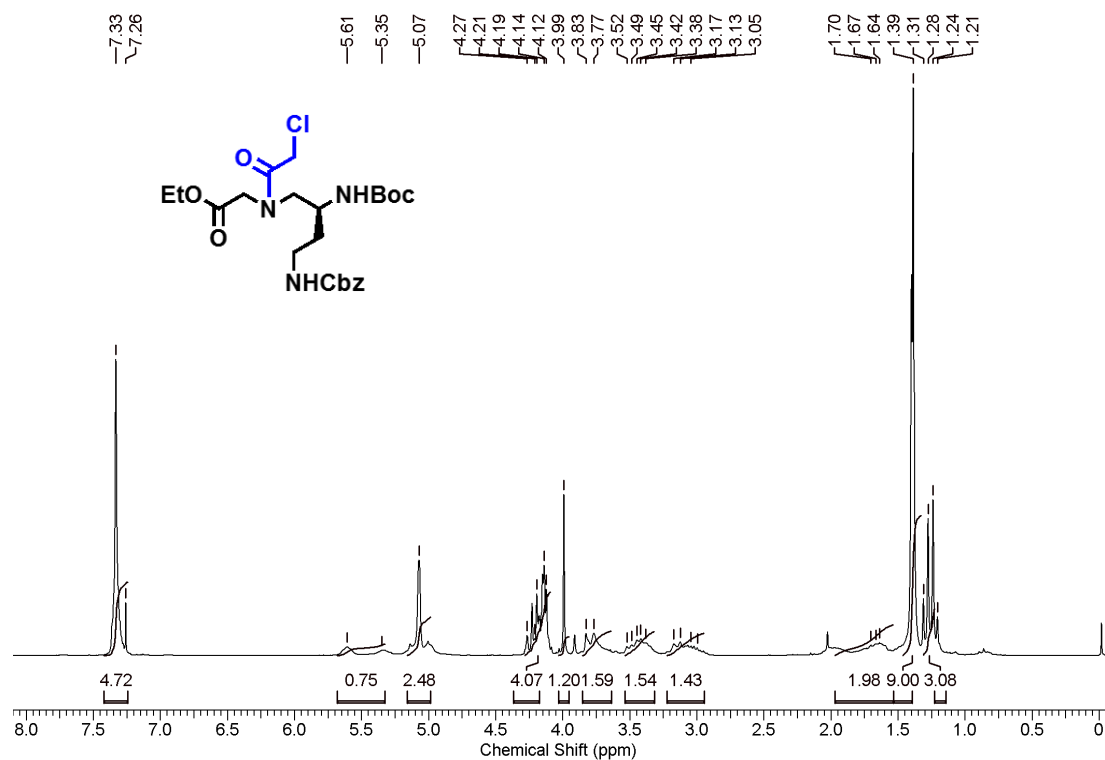
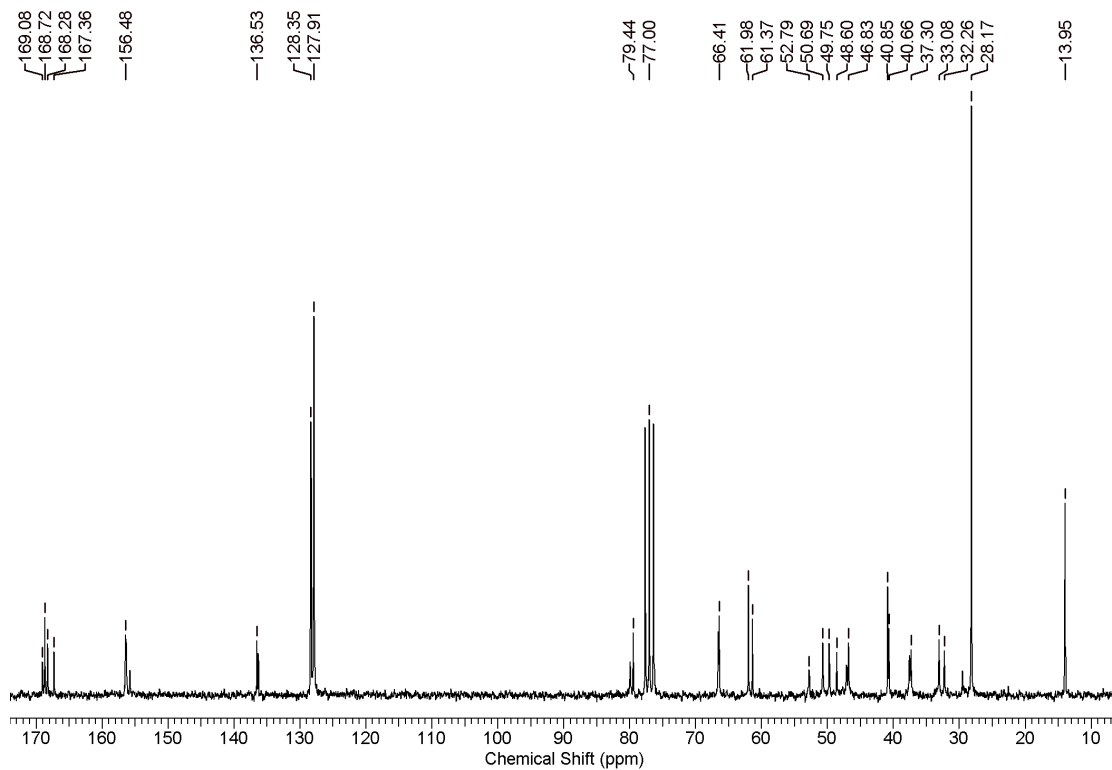
^1H NMR of Compound 4 ^{13}C NMR of Compound 4

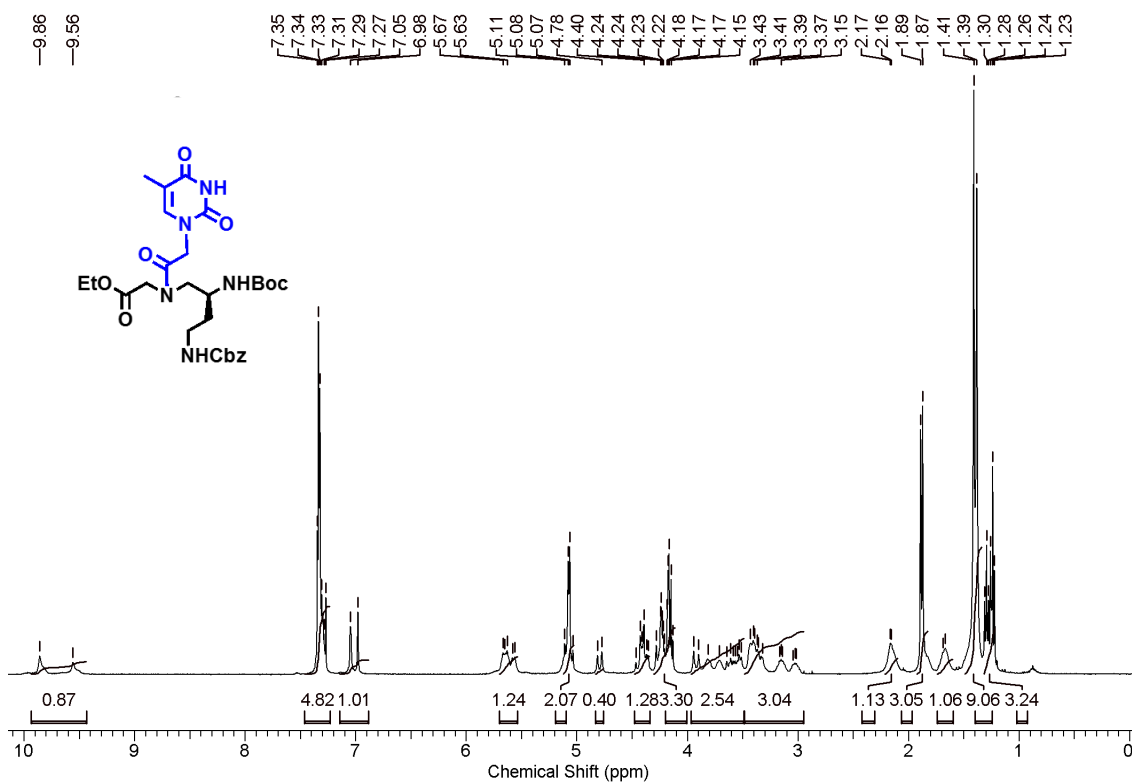
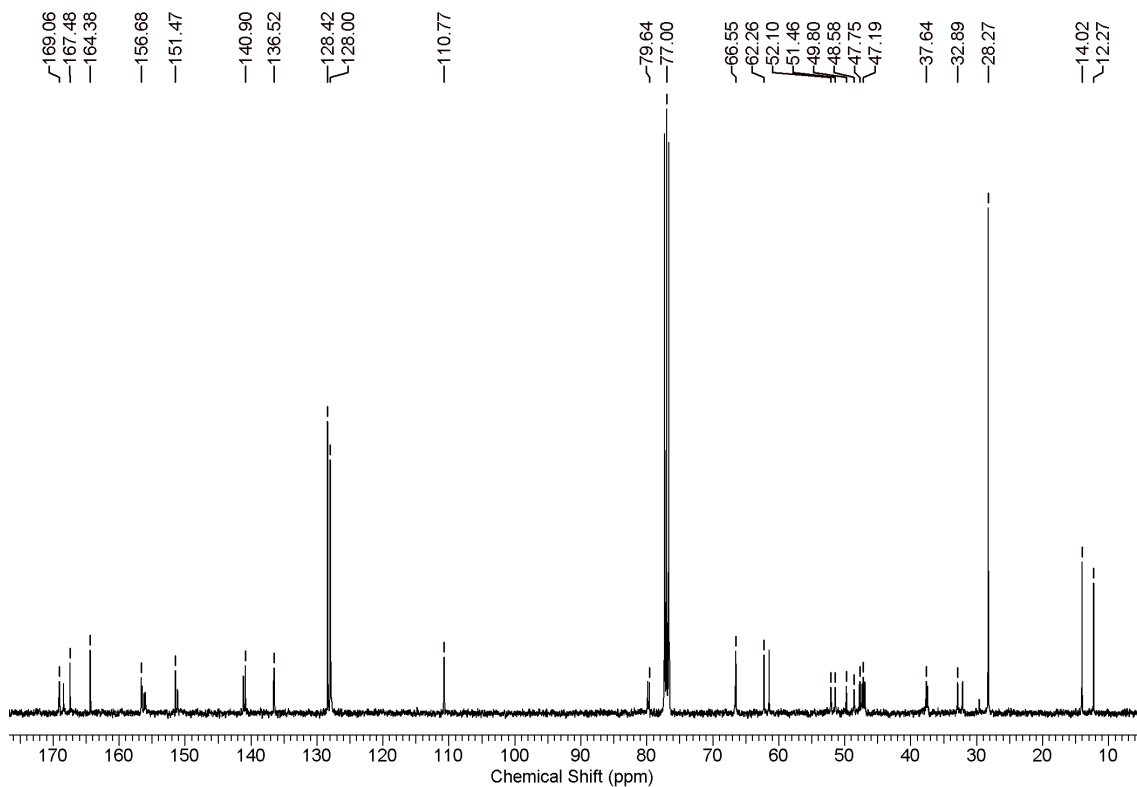
^1H NMR of Compound 5 ^{13}C NMR of Compound 5

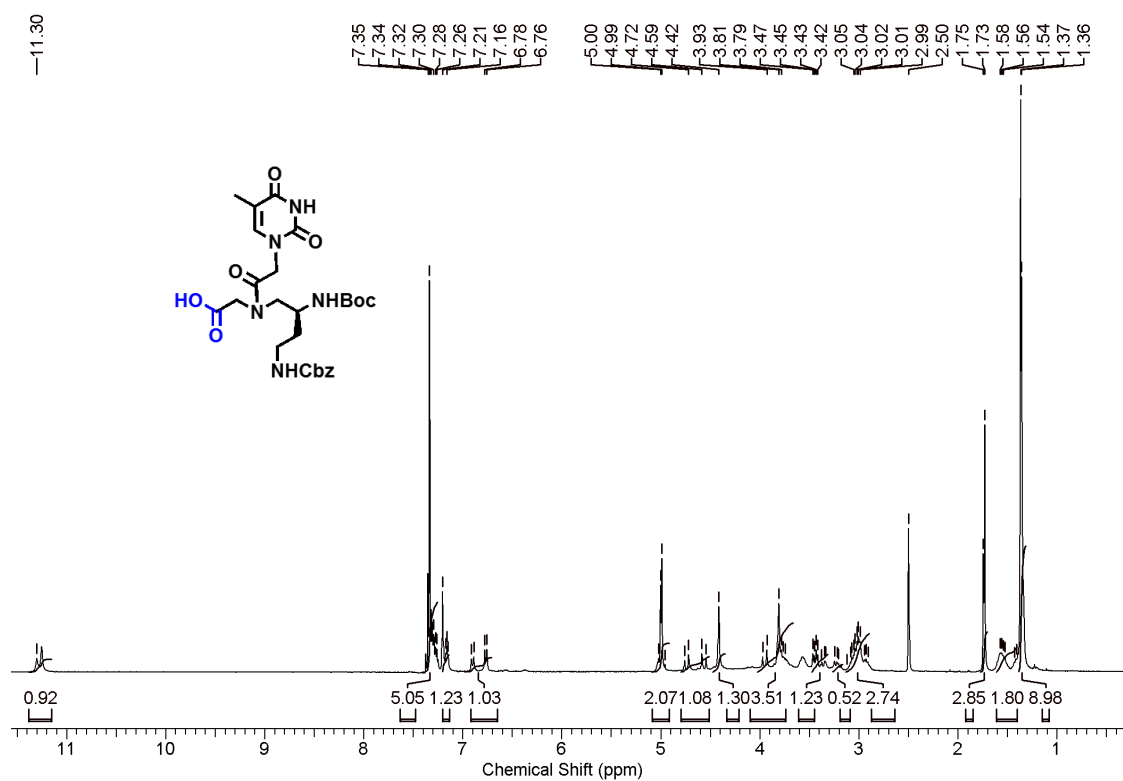
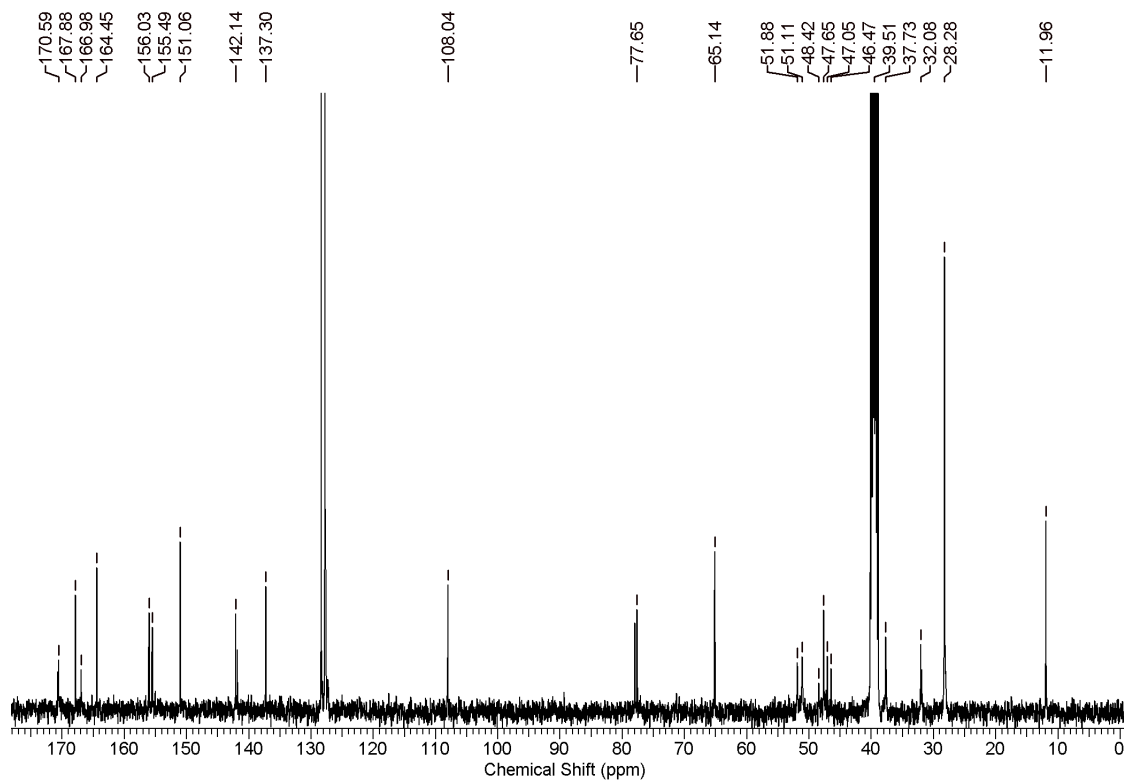
¹H NMR of Compound 6**¹³C NMR of Compound 6**

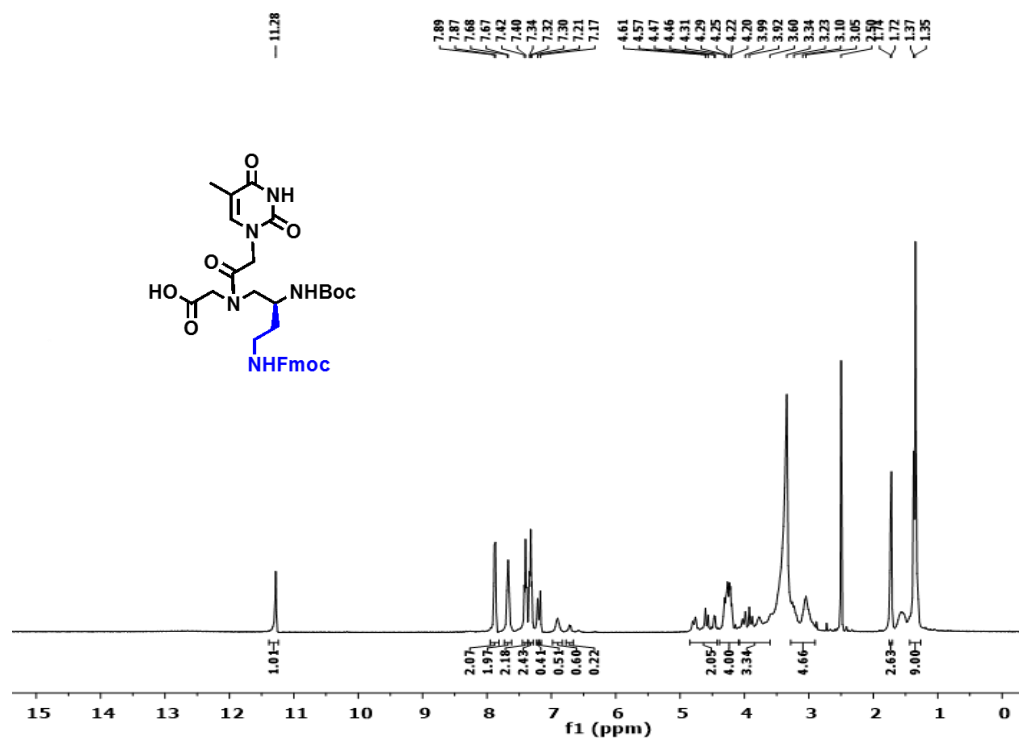
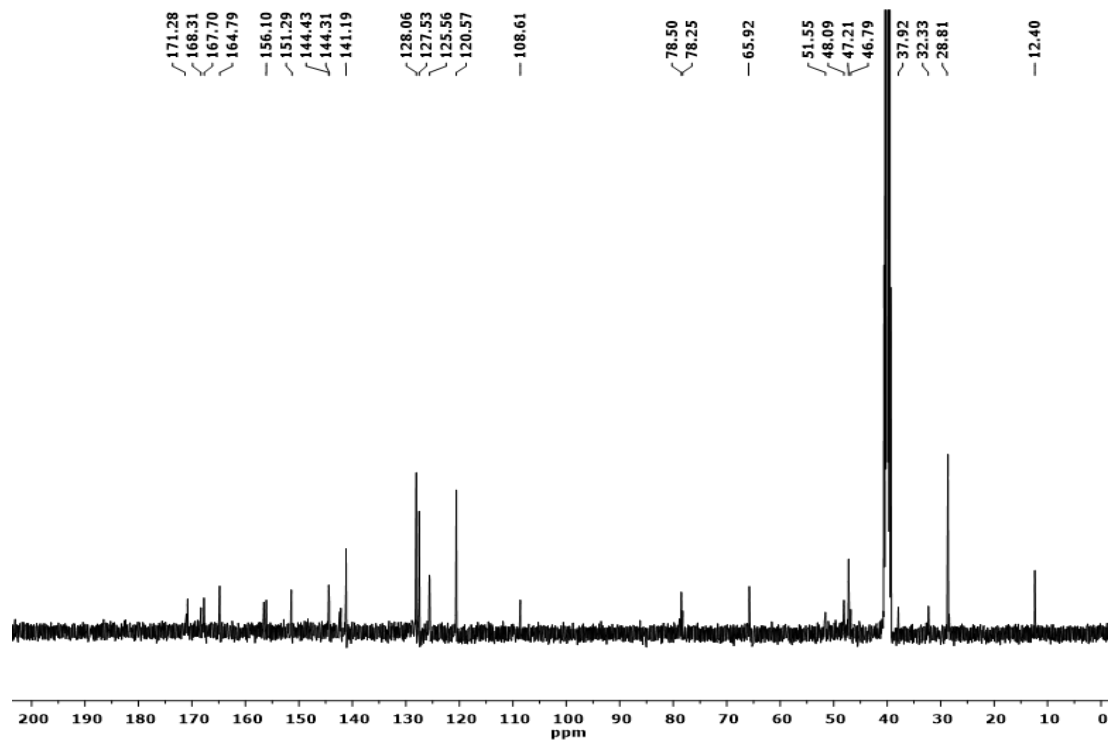
^1H NMR of Compound 8 ^{13}C NMR of Compound 8

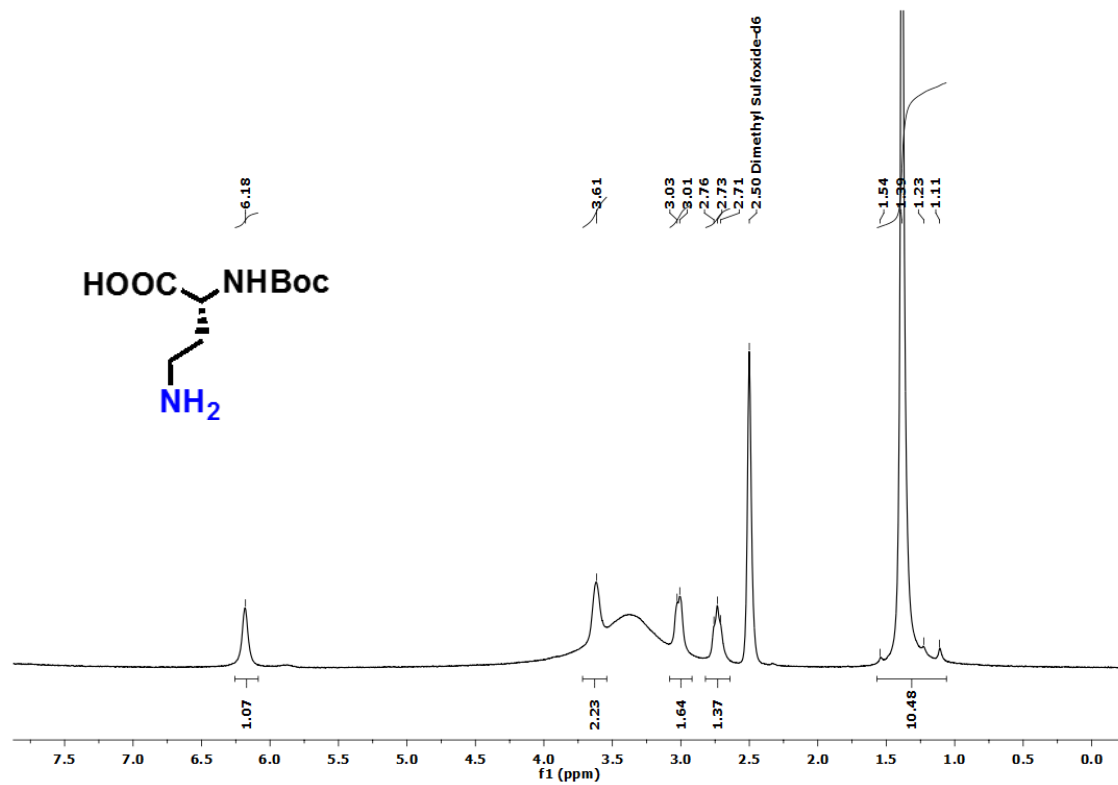
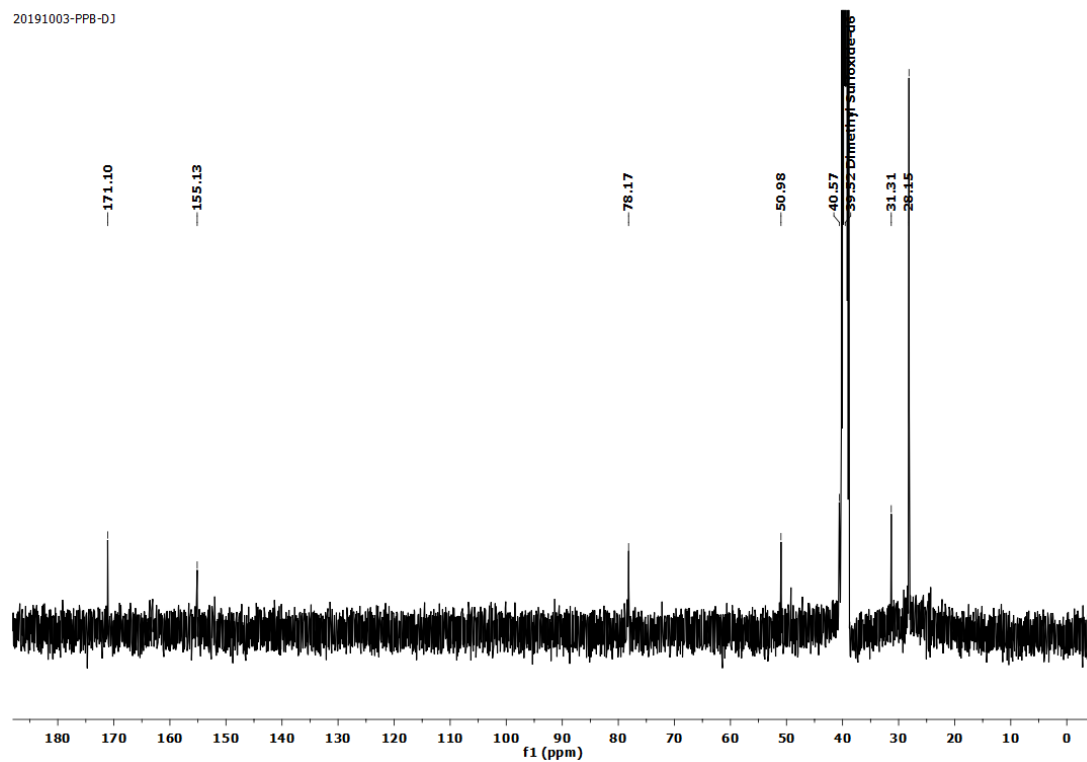
^1H NMR of Compound 9 **^{13}C NMR of Compound 9**

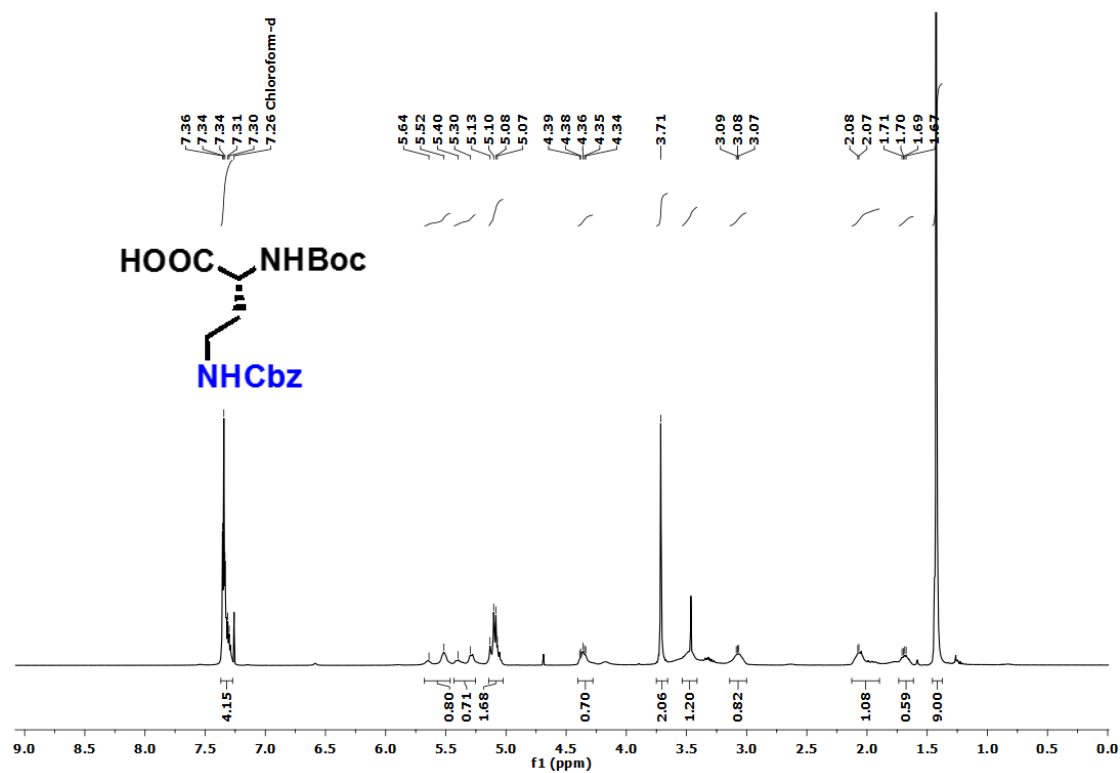
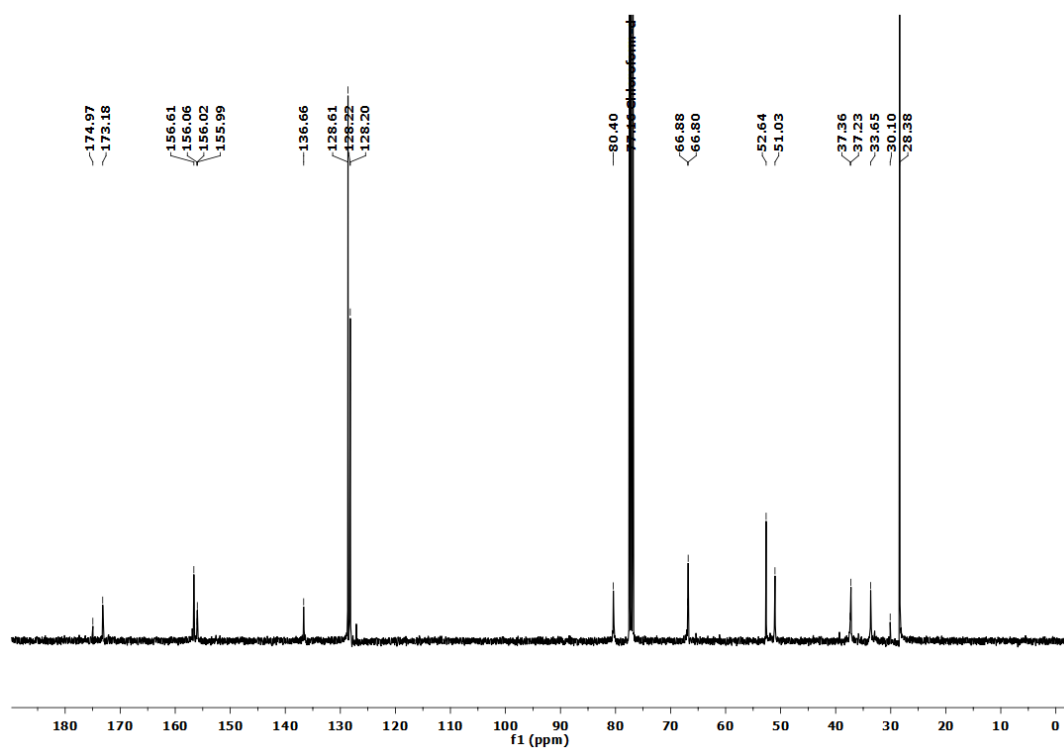
^1H NMR of Compound 10 **^{13}C NMR of Compound 10**

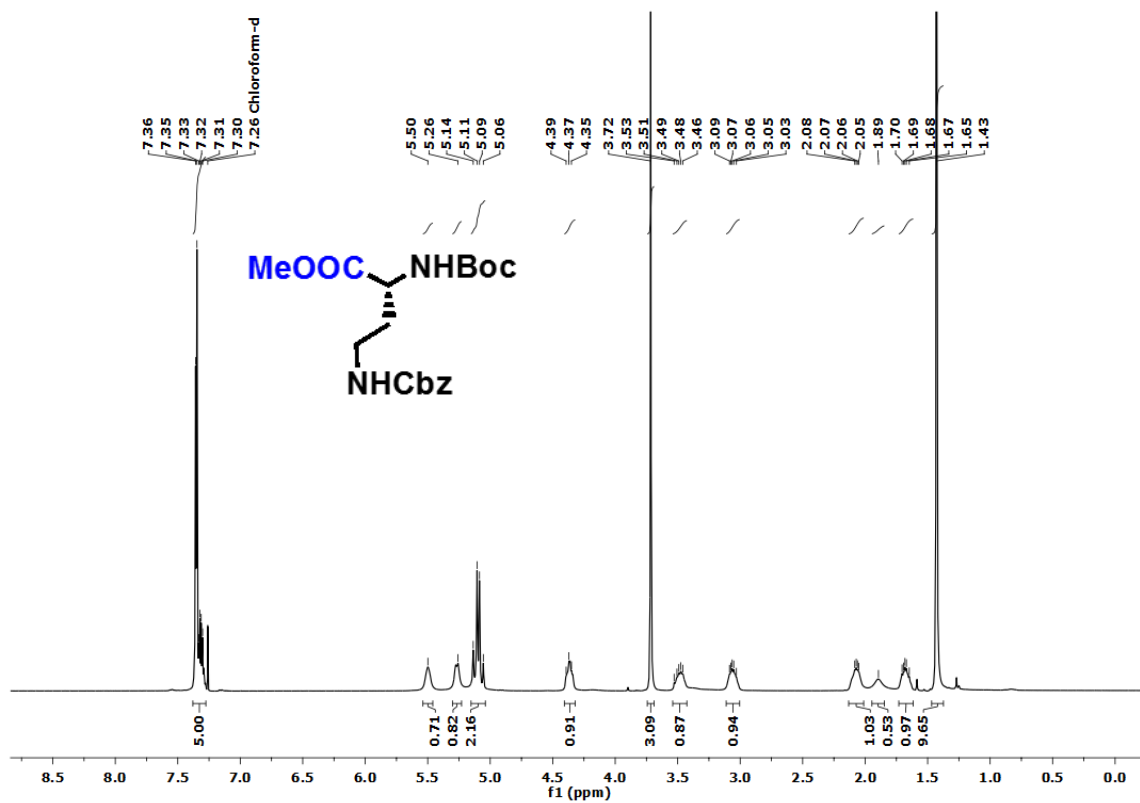
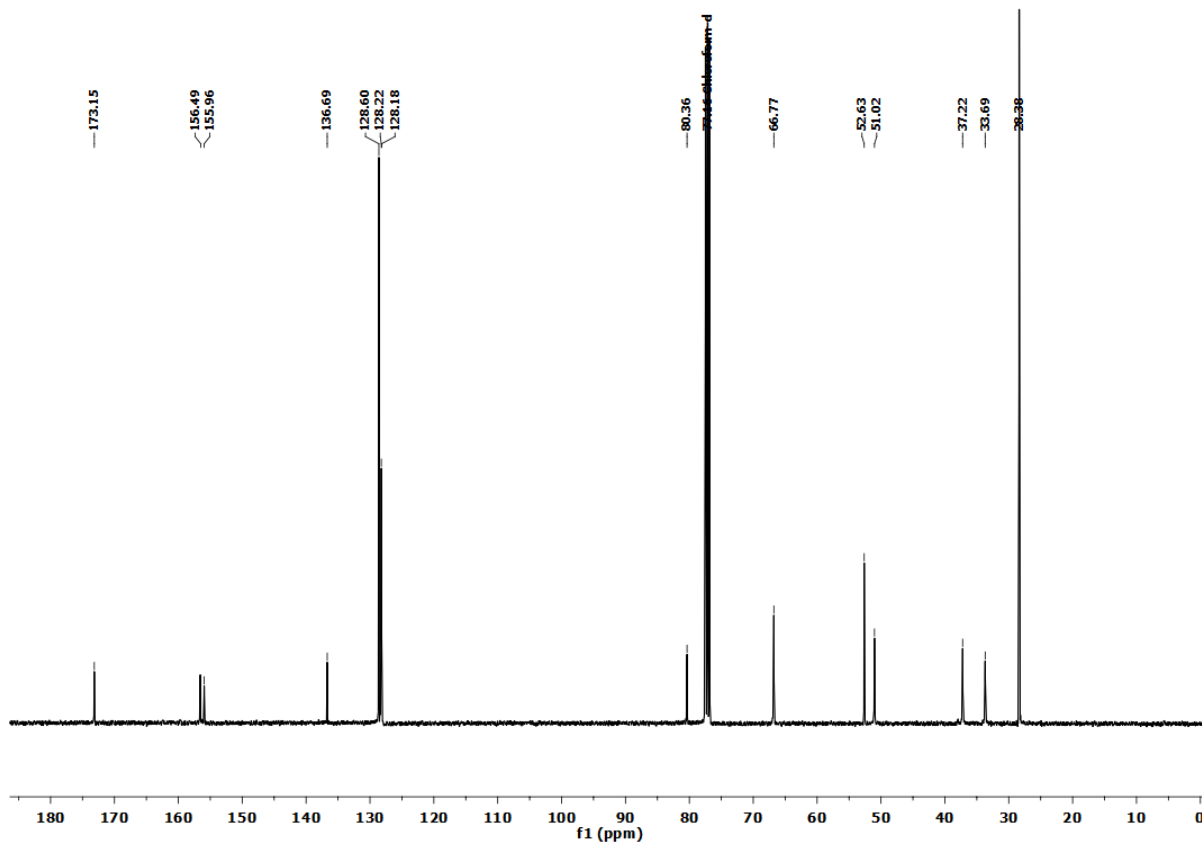
¹H NMR of Compound 11**¹³C NMR of Compound 11**

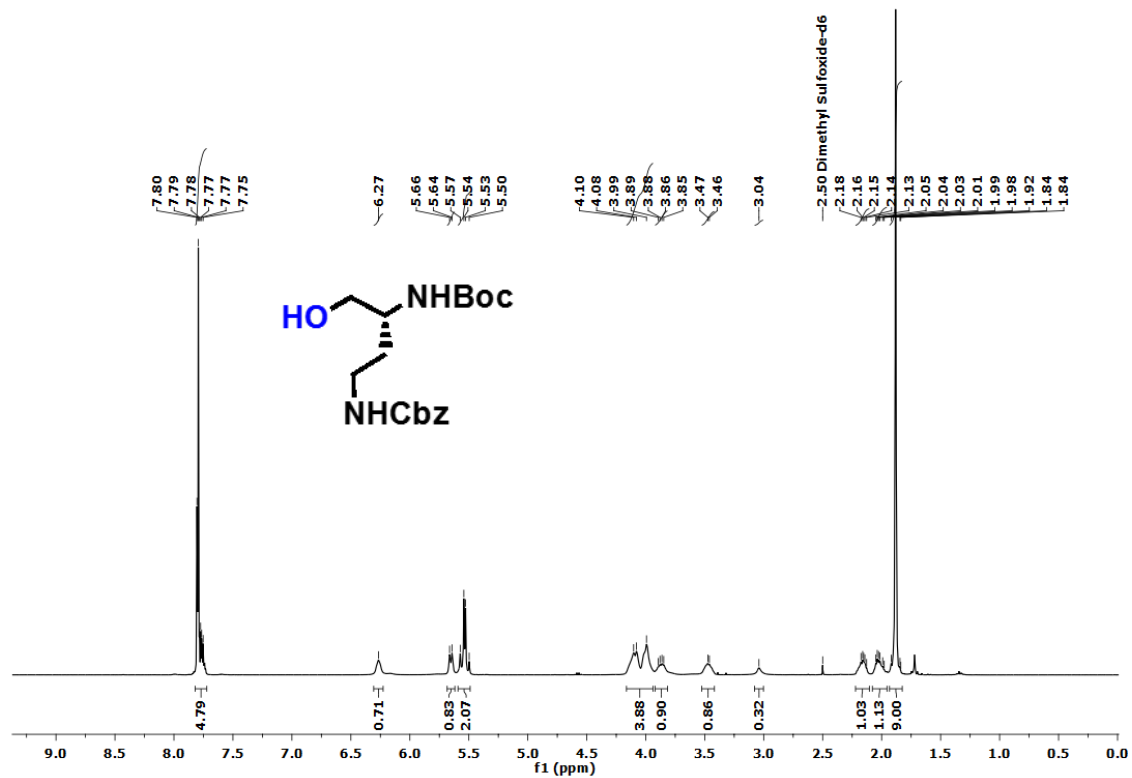
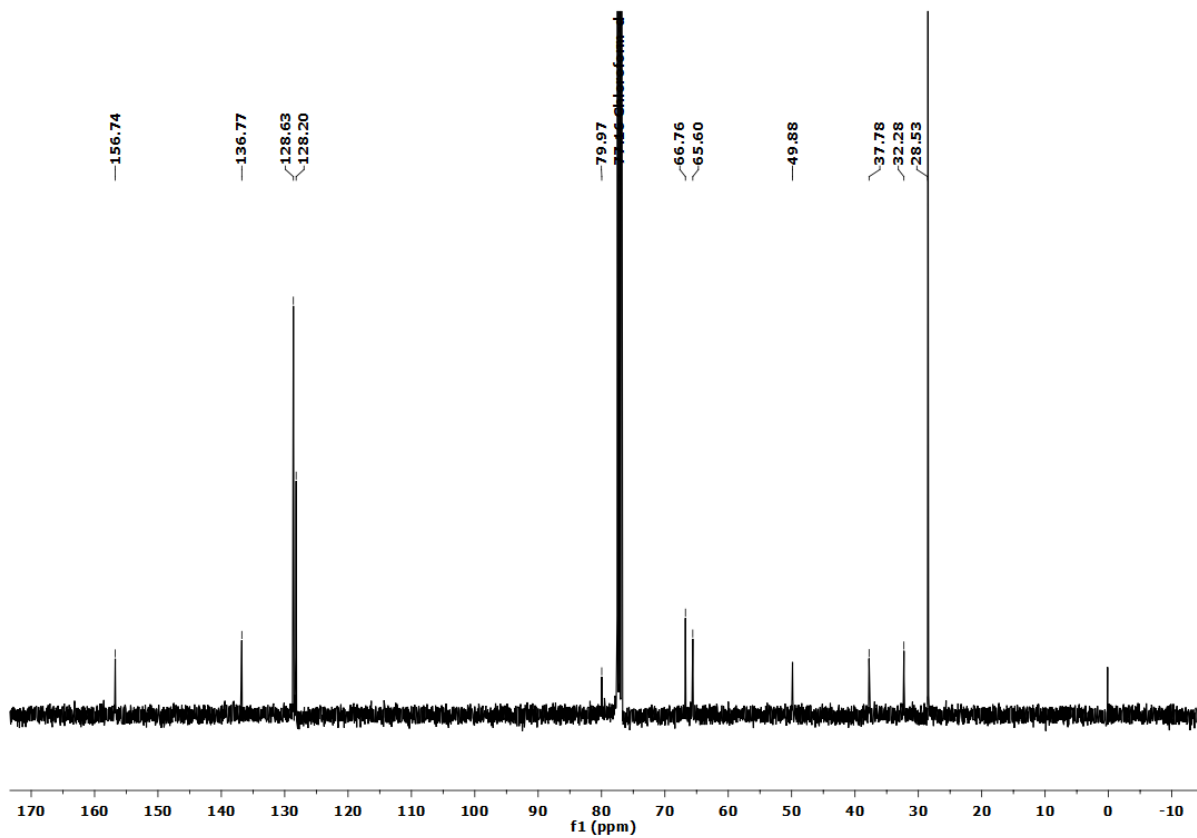
^1H NMR of Compound 12 **^{13}C NMR of Compound 12**

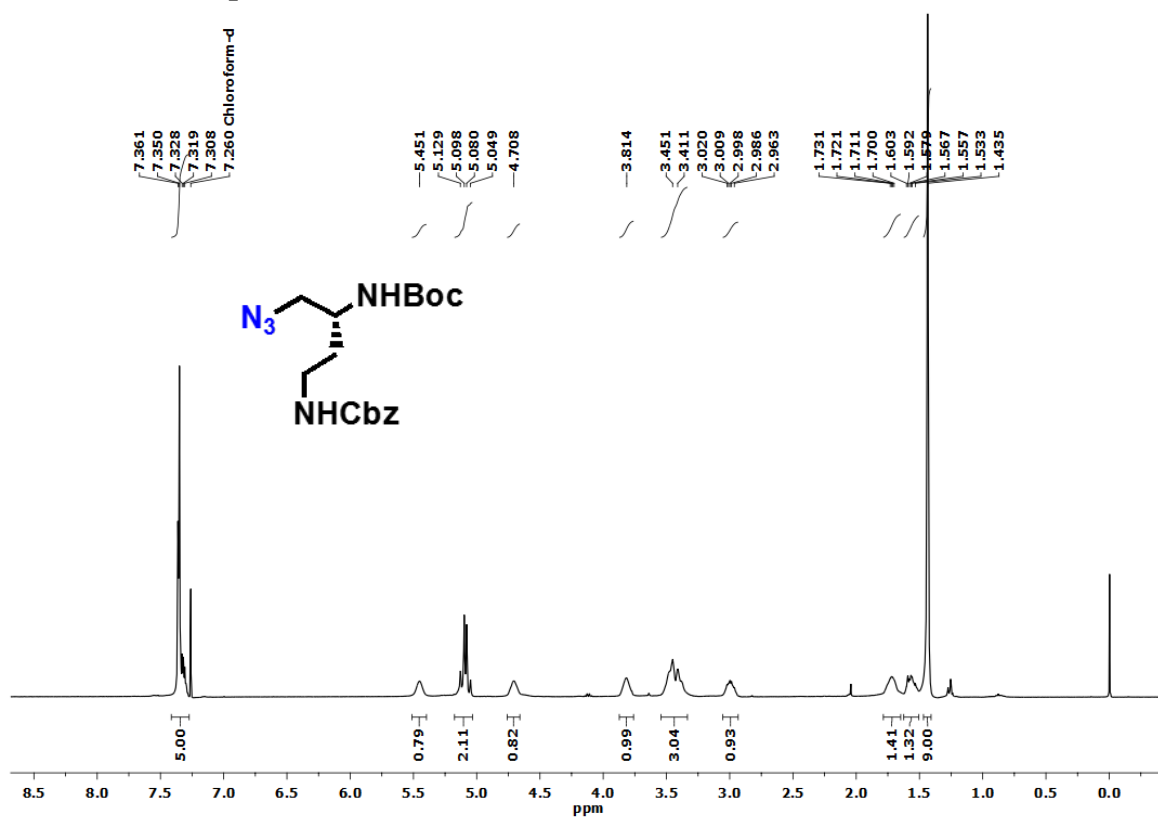
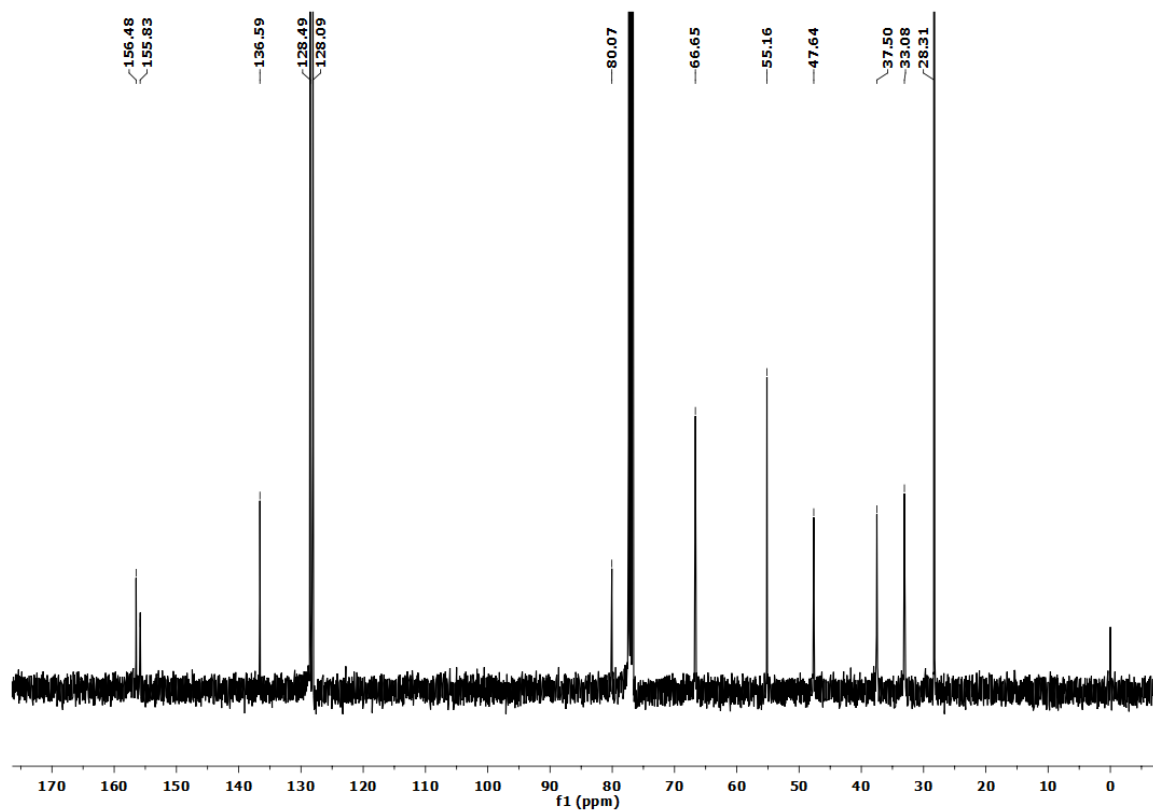
^1H NMR of Compound 14 ^{13}C NMR of Compound 14

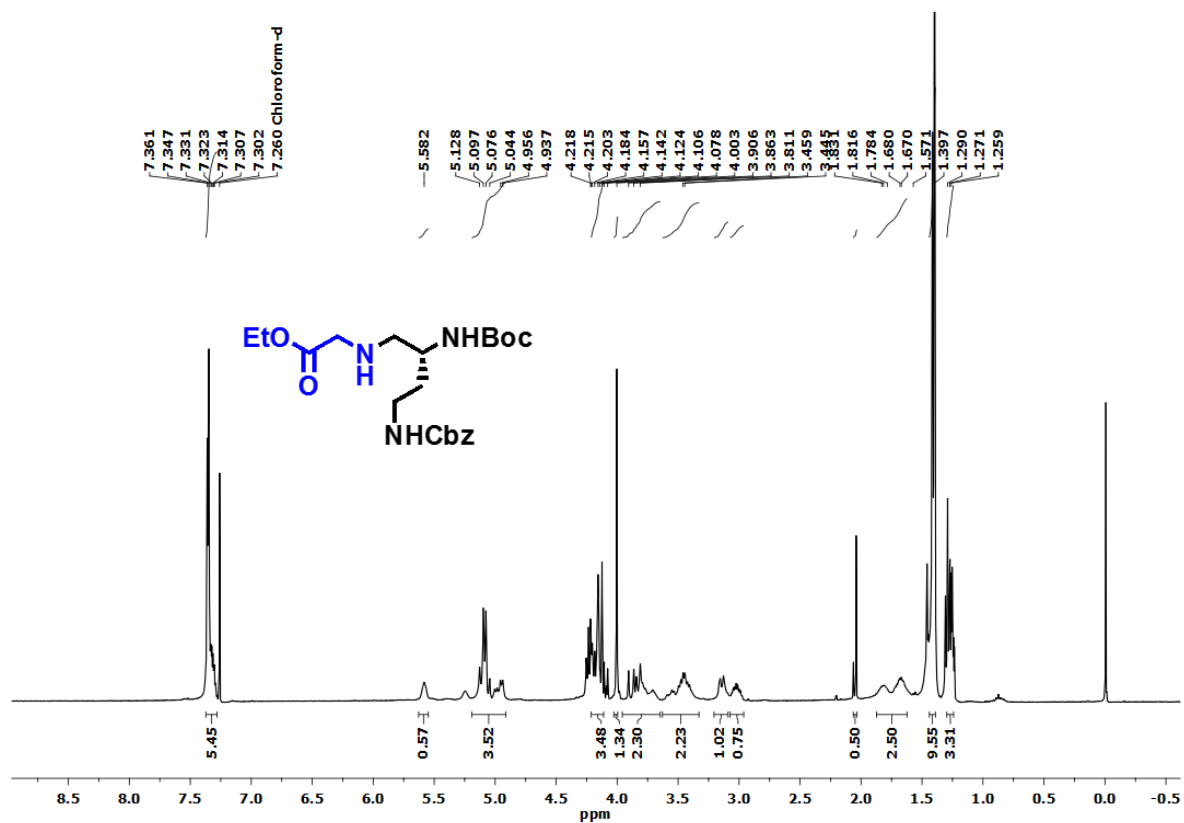
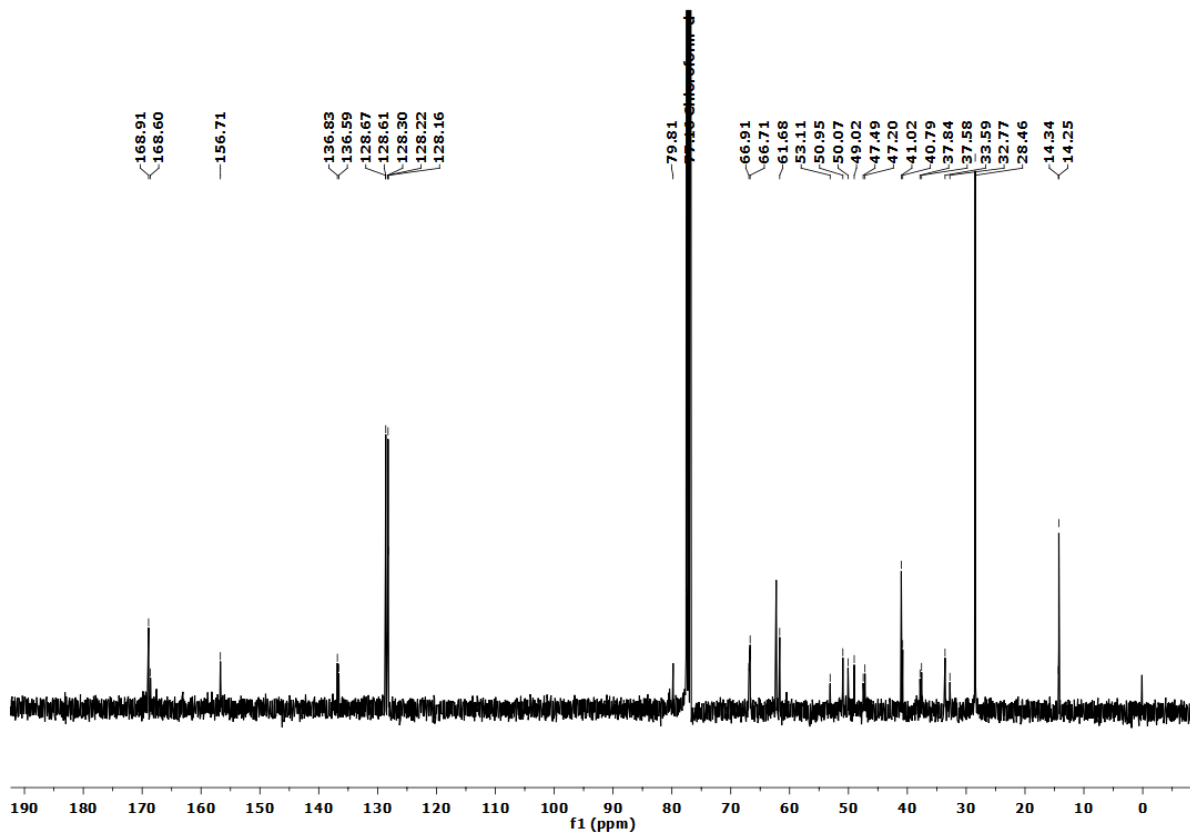
^1H NMR of Compound 17 ^{13}C NMR of Compound 17

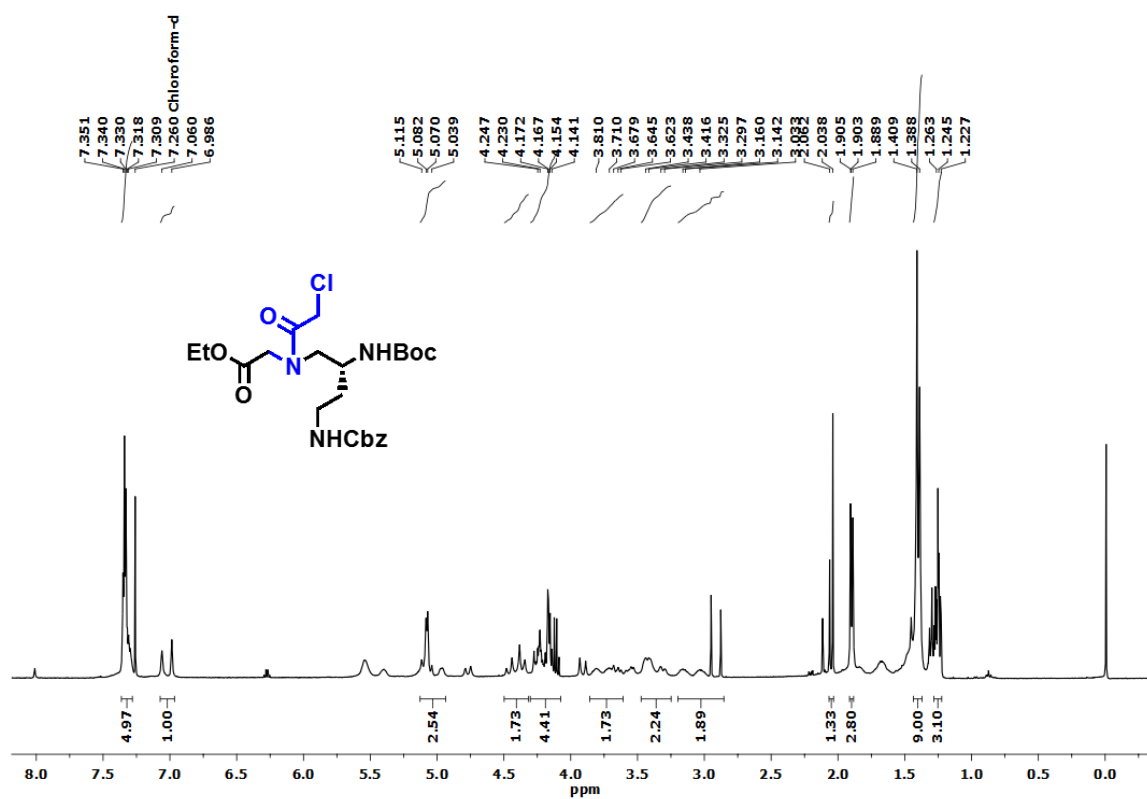
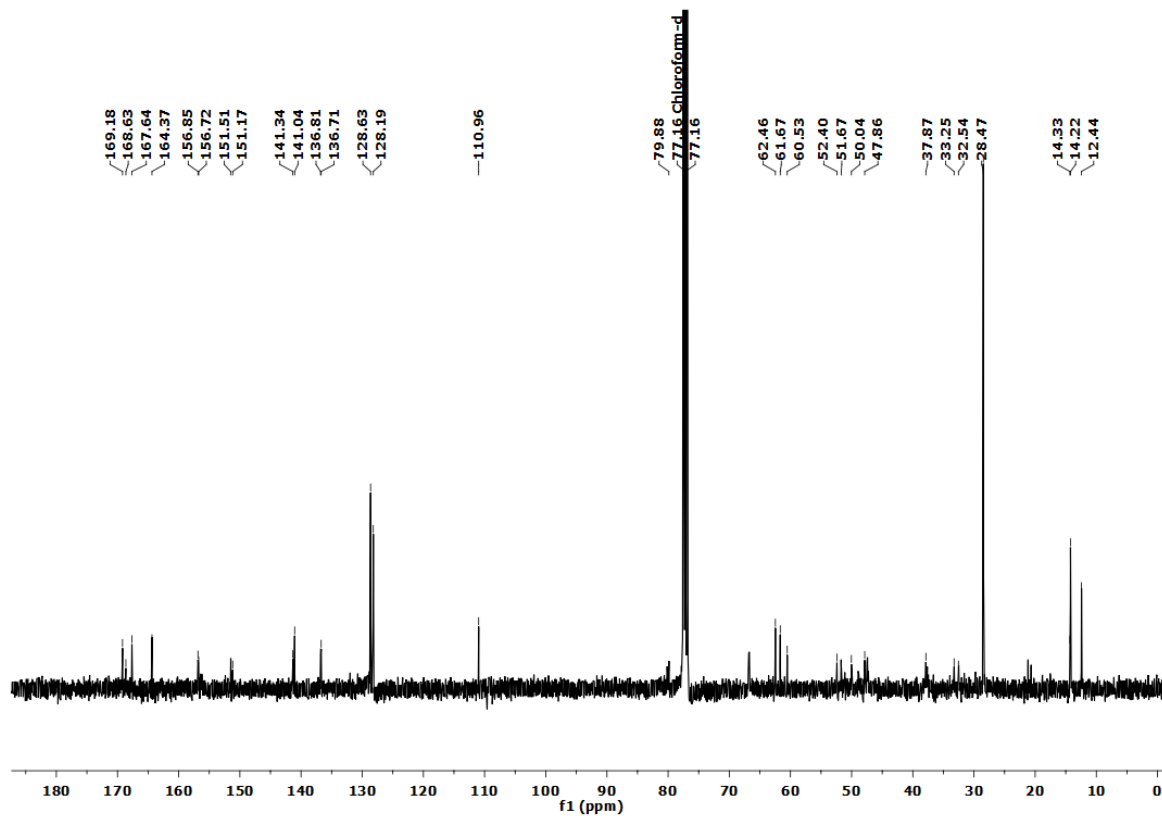
^1H NMR of Compound 18 ^{13}C NMR of Compound 18

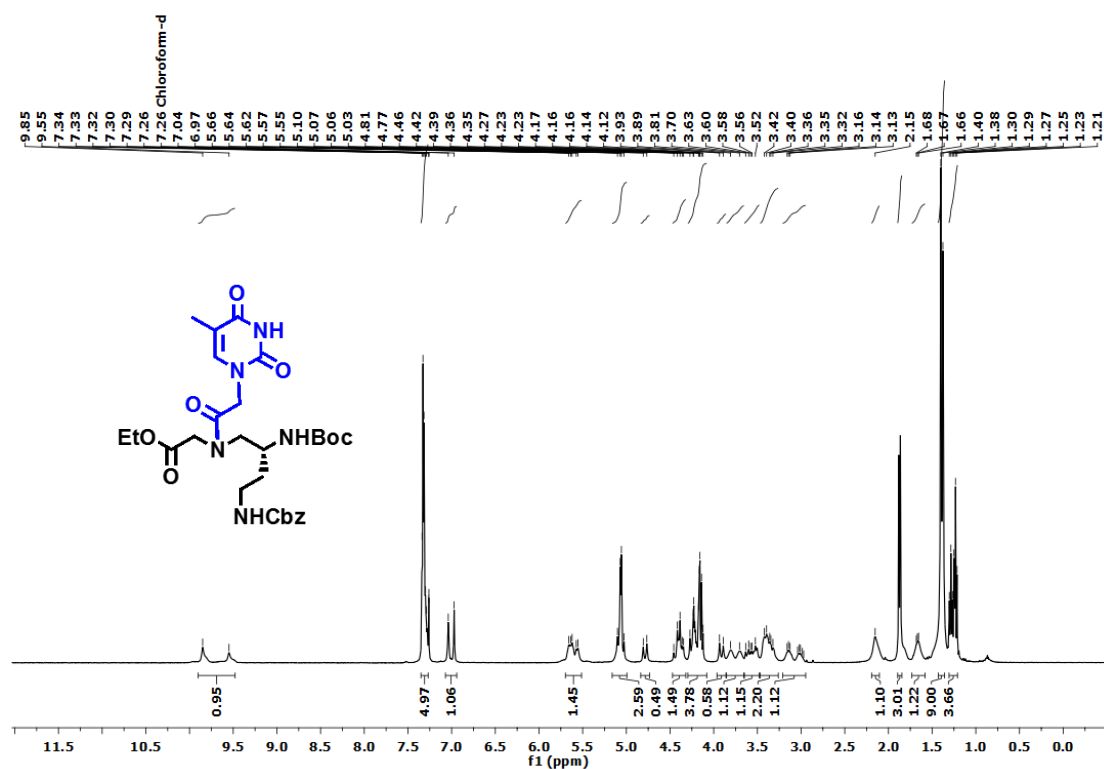
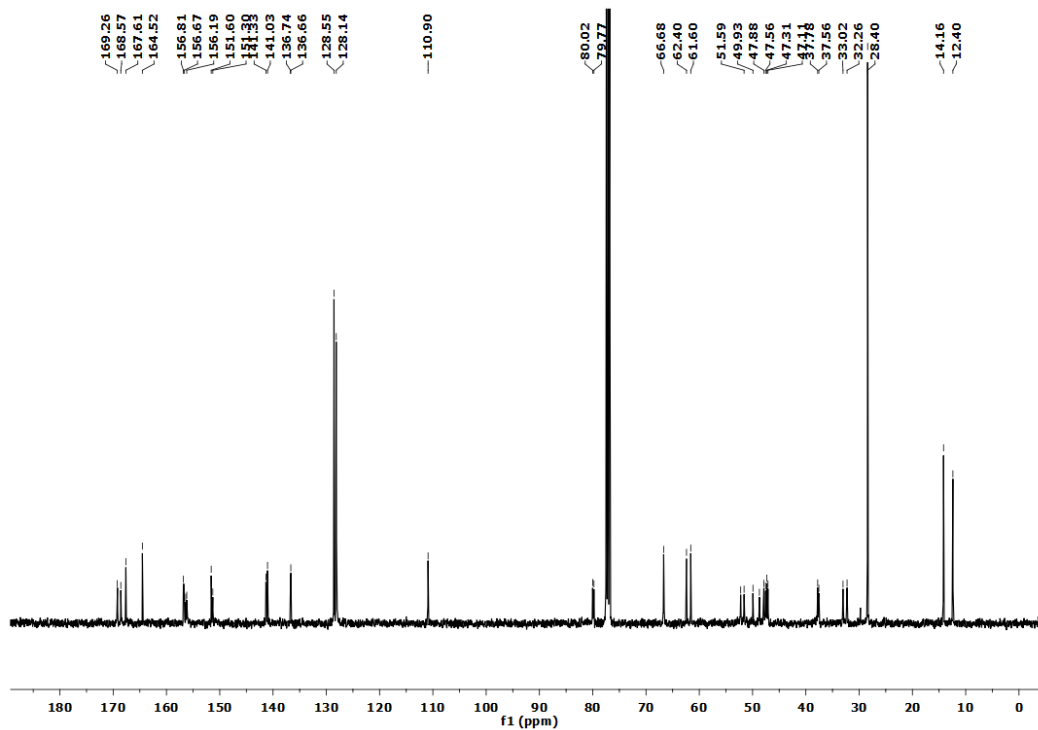
^1H NMR of Compound 19 ^{13}C NMR of Compound 19

^1H NMR of Compound 20 ^{13}C NMR spectrum of Compound 20

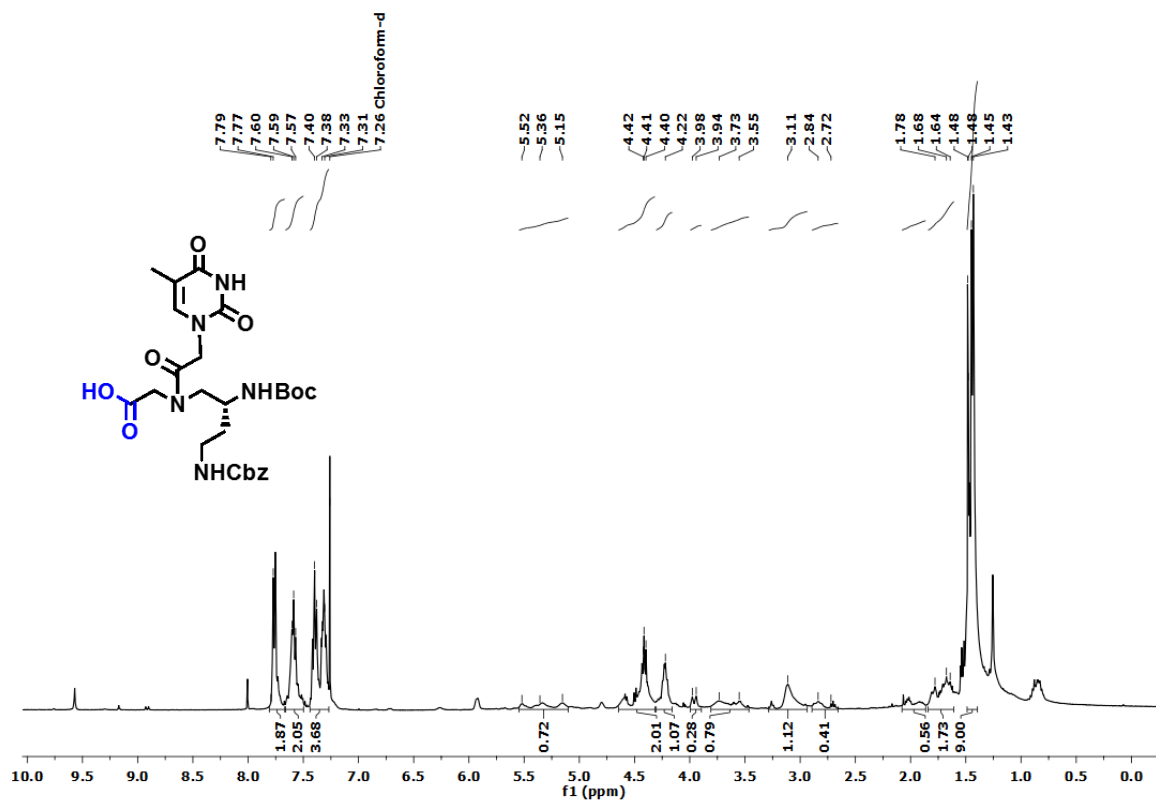
^1H NMR of Compound 22 ^{13}C NMR of Compound 22

^1H NMR of Compound 23 ^{13}C NMR of Compound 23

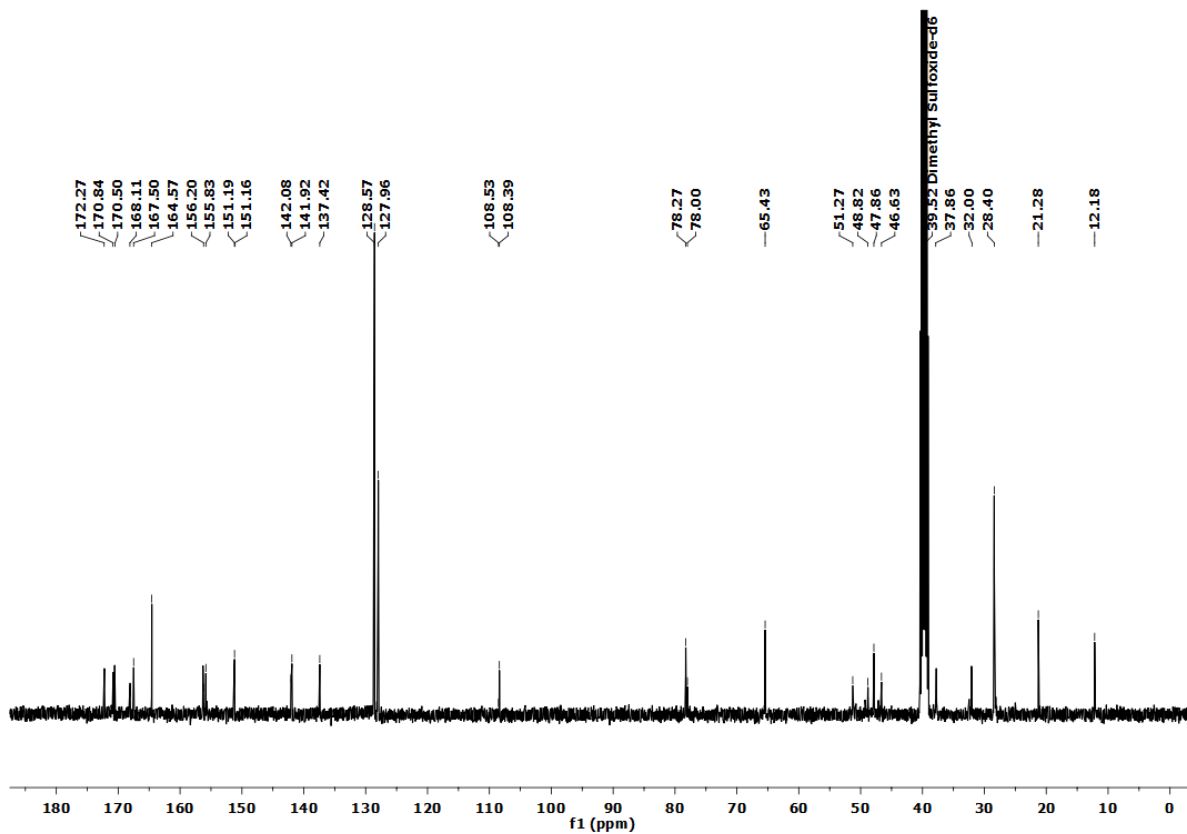
^1H NMR of Compound 24 ^{13}C NMR of Compound 24

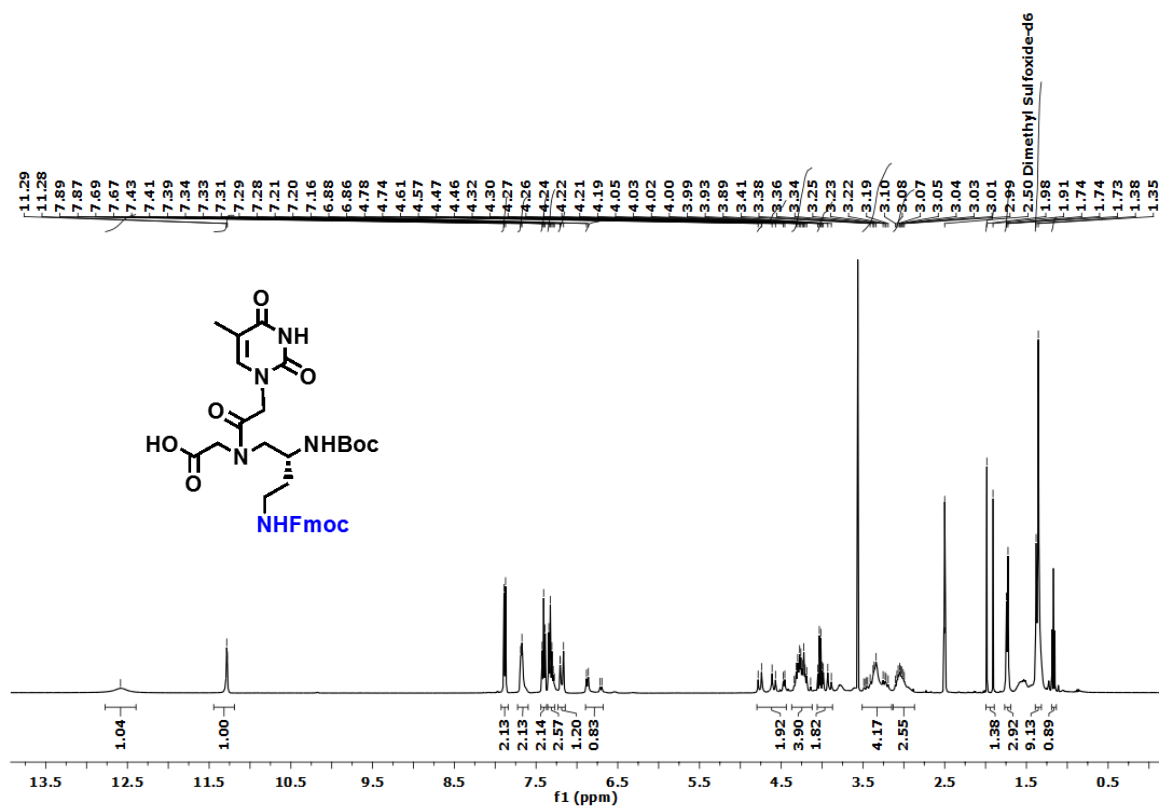
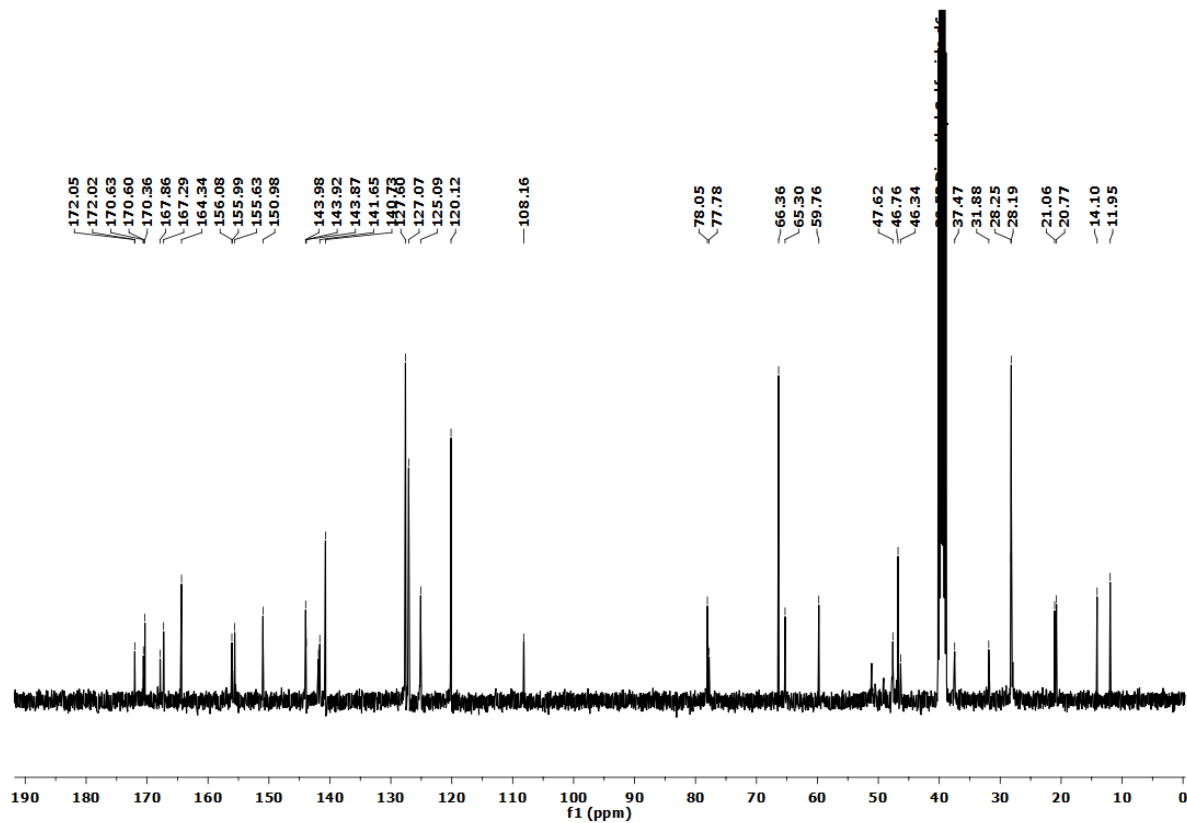
¹H NMR of Compound 25¹³C NMR of Compound 25

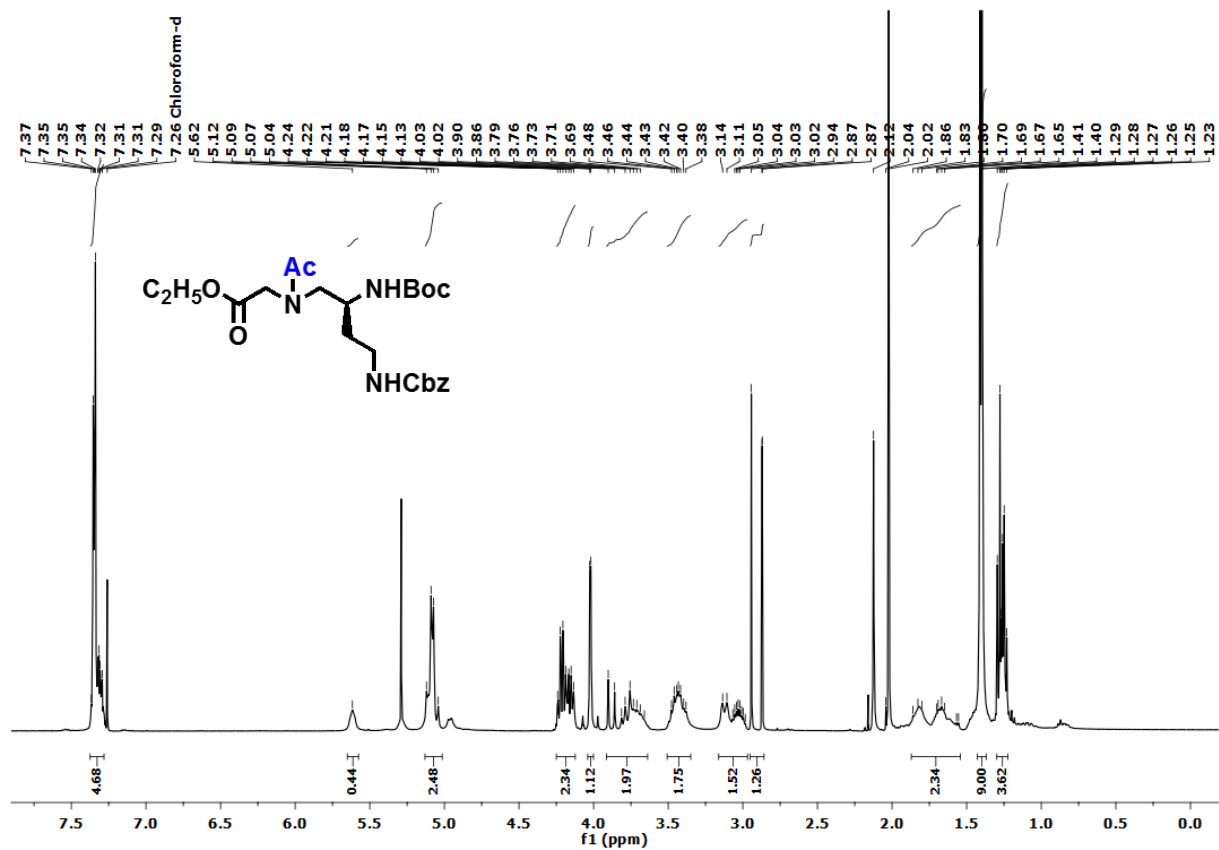
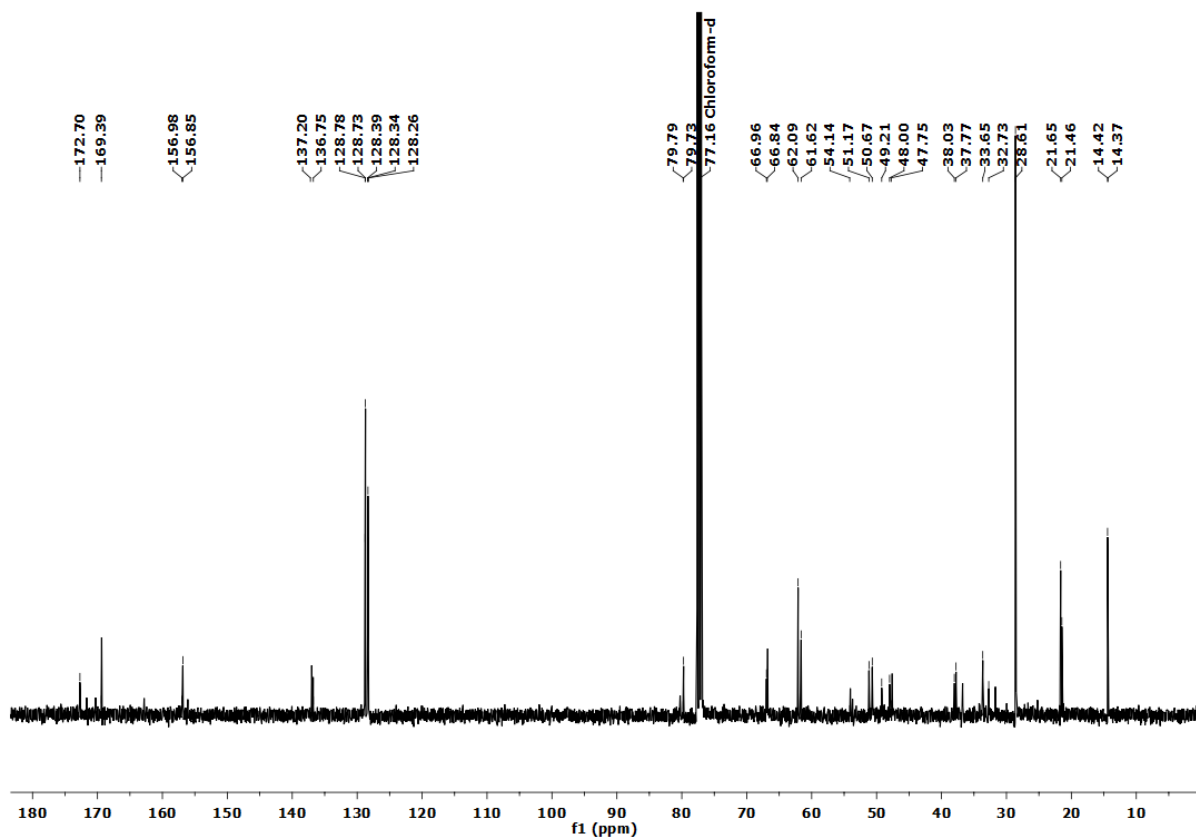
¹H NMR of Compound 26

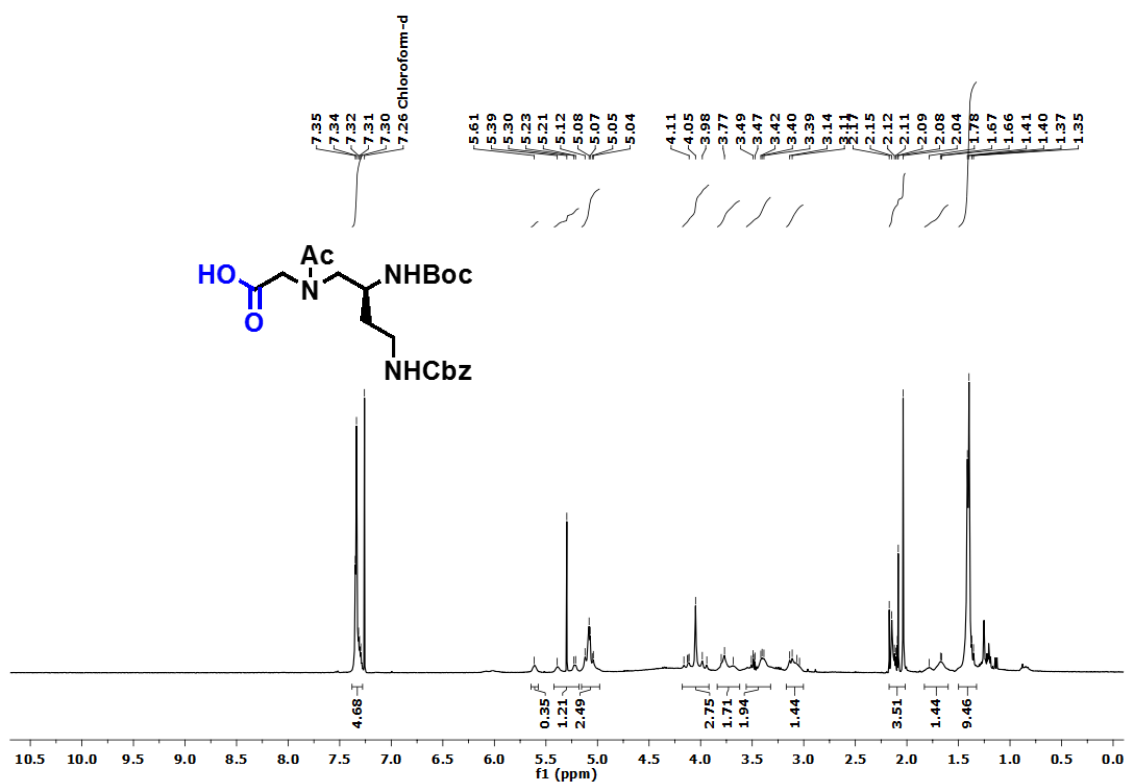
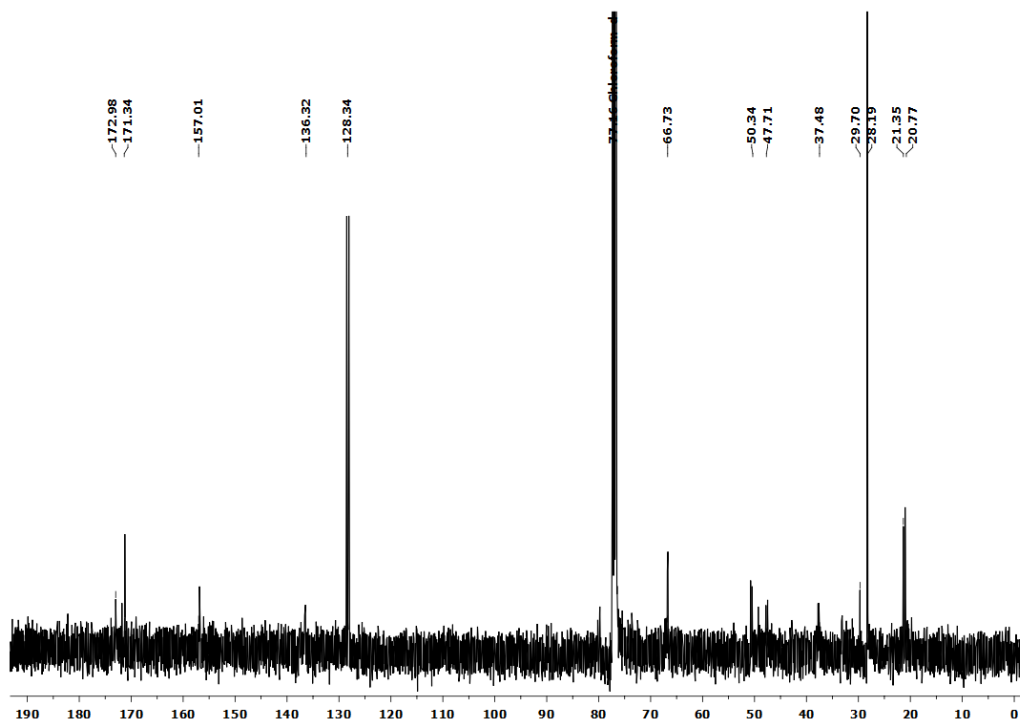


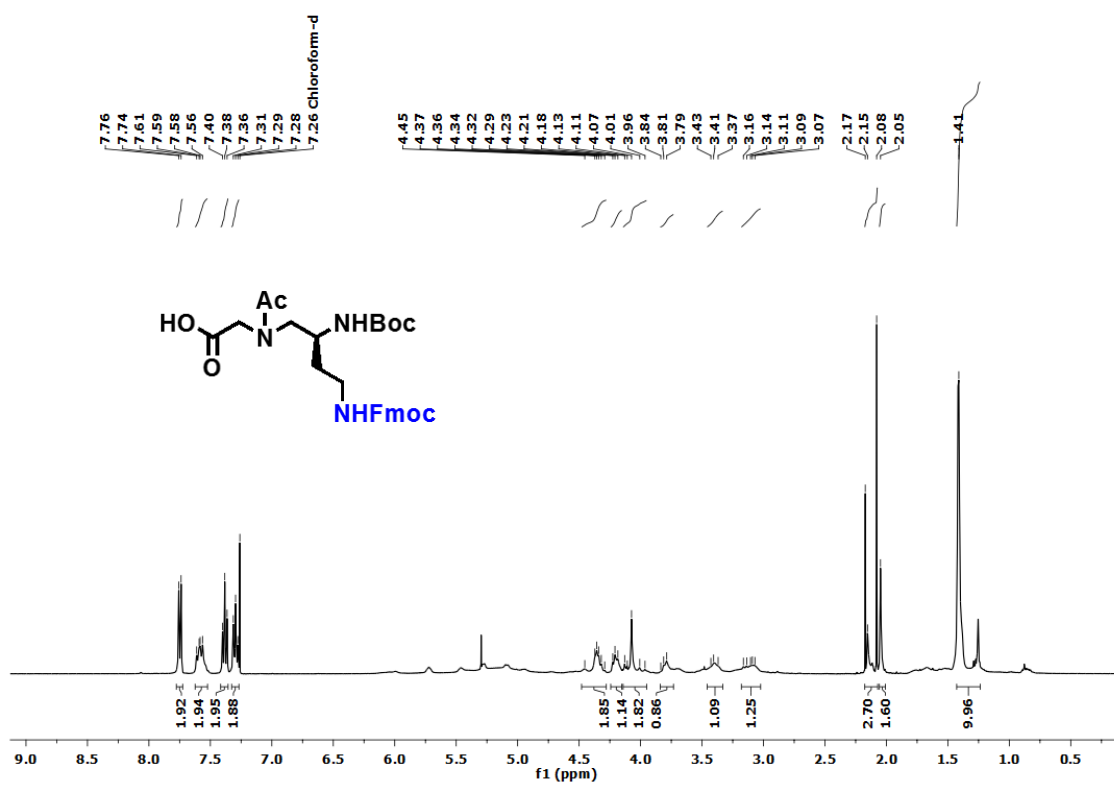
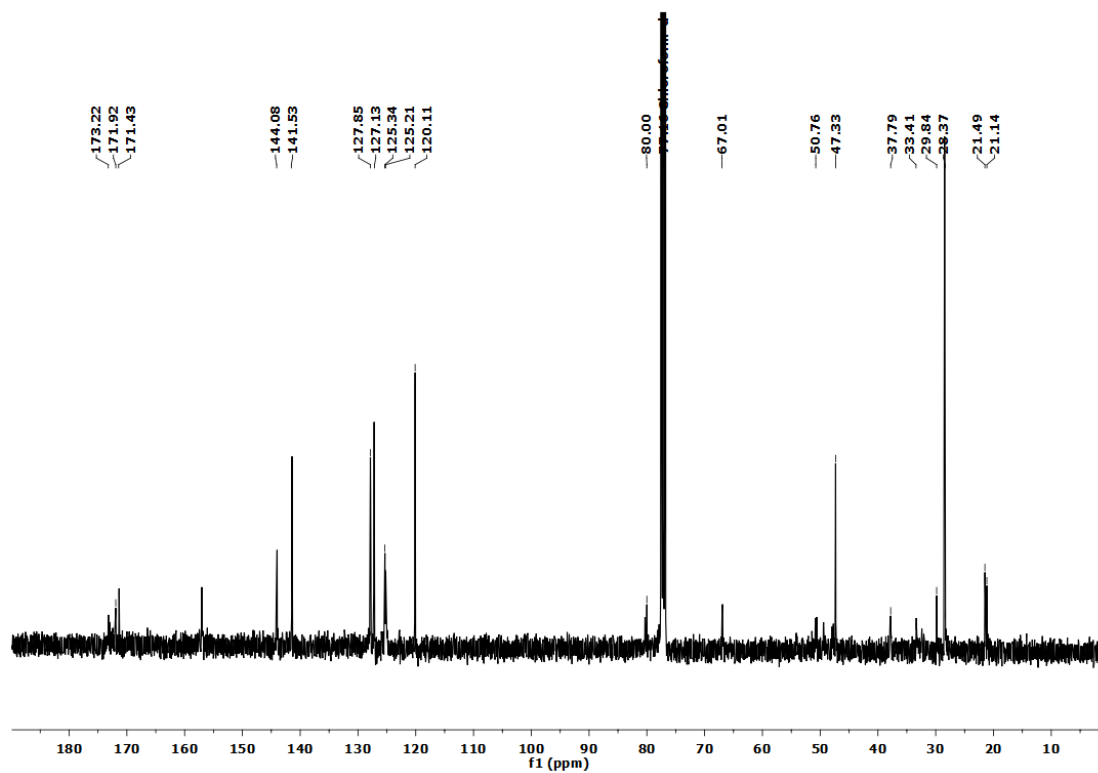
¹³C NMR of Compound 26

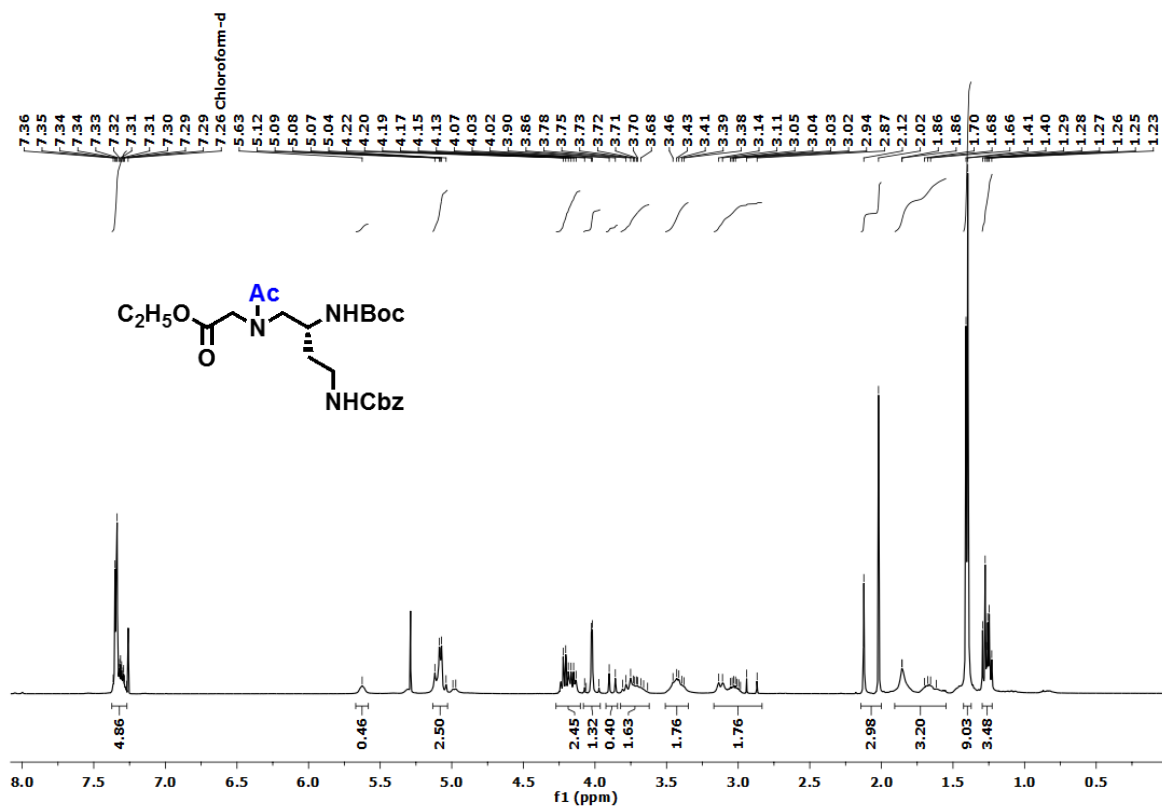
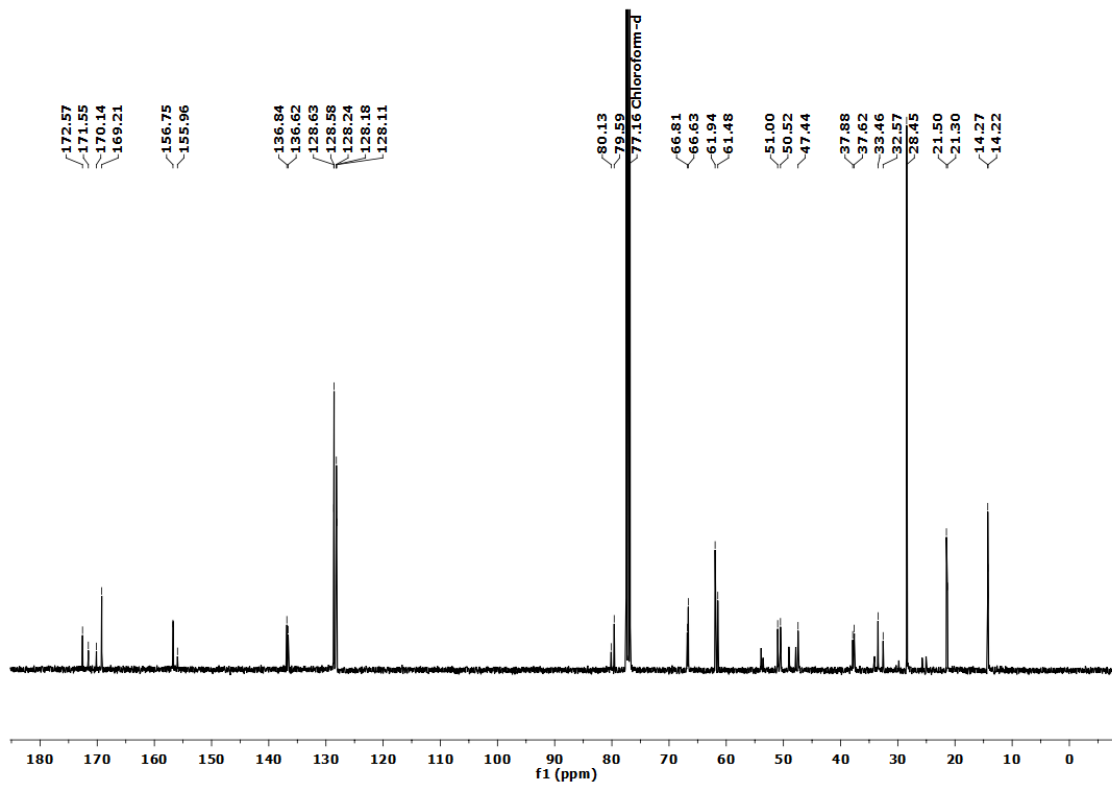


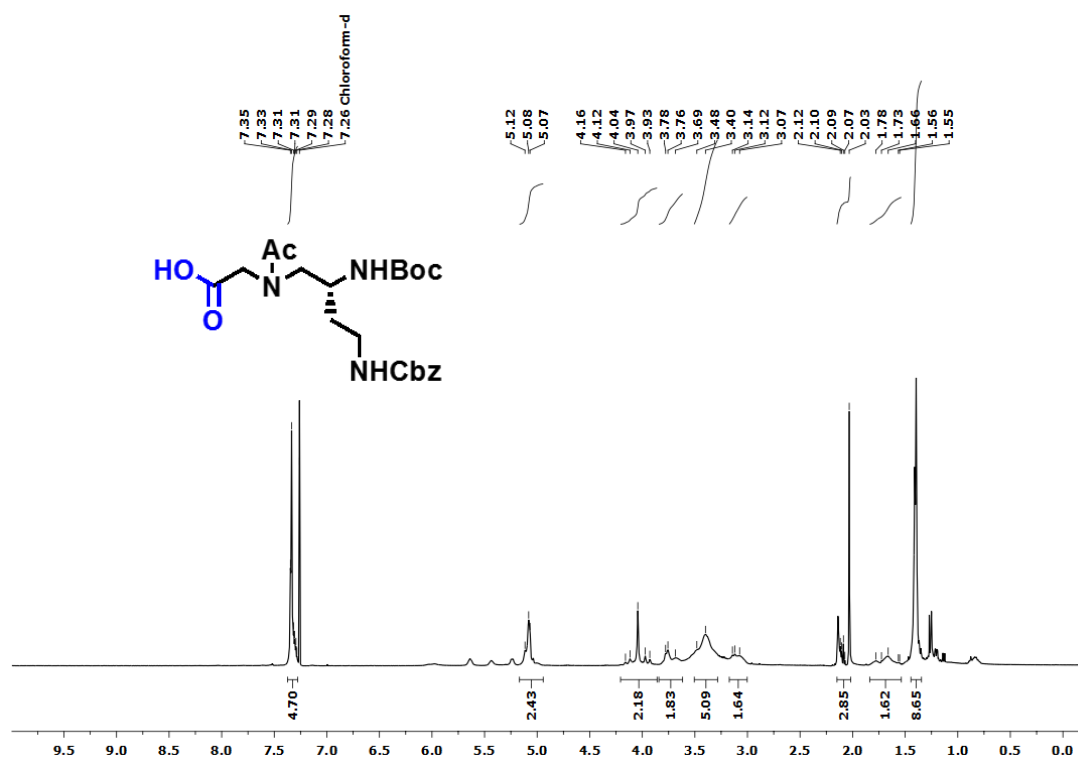
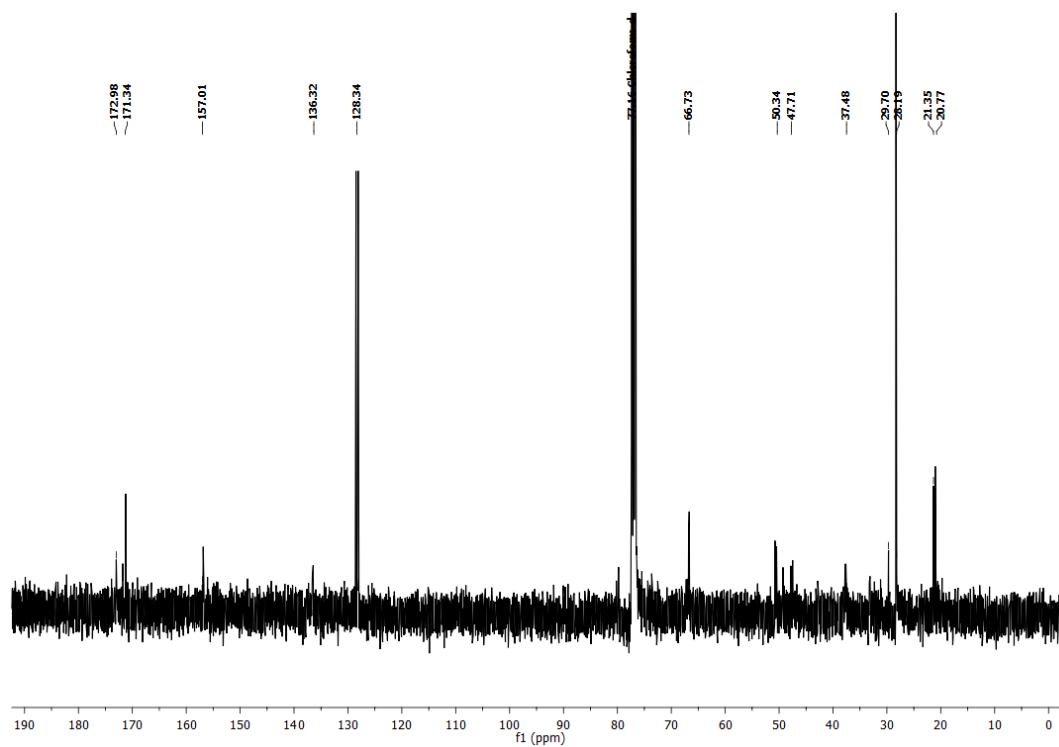
^1H NMR of Compound 28 ^{13}C NMR of Compound 28

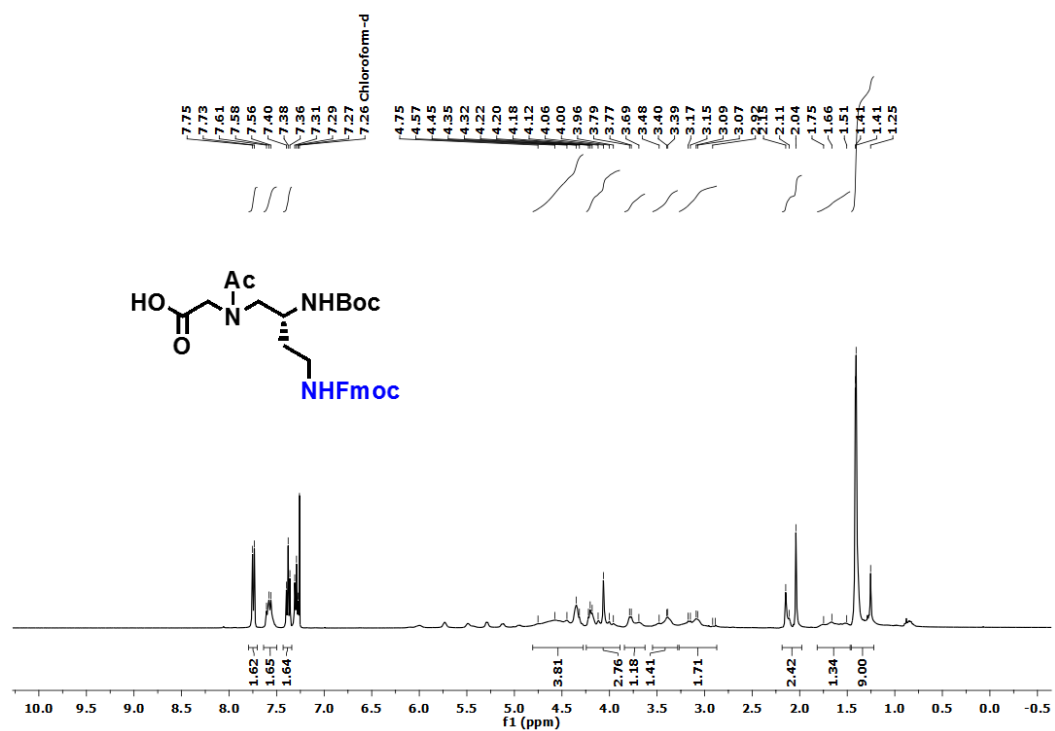
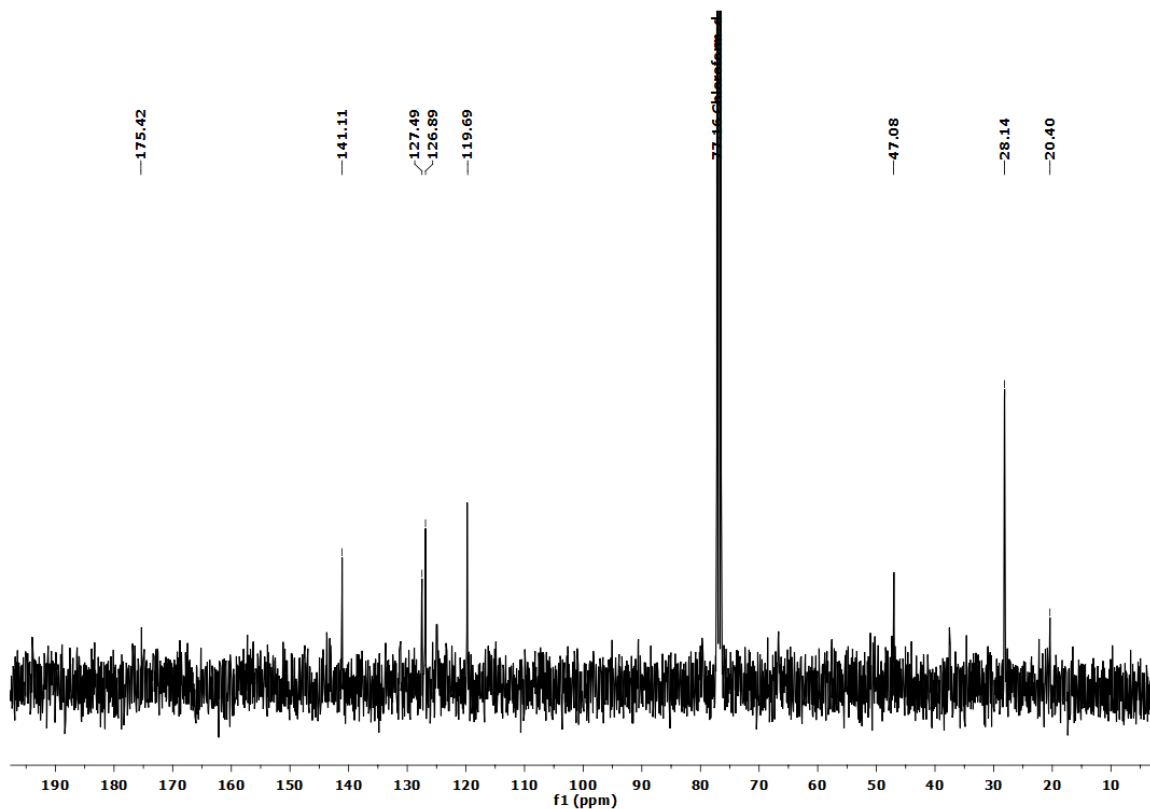
^1H NMR of Compound 29 ^{13}C NMR of Compound 29

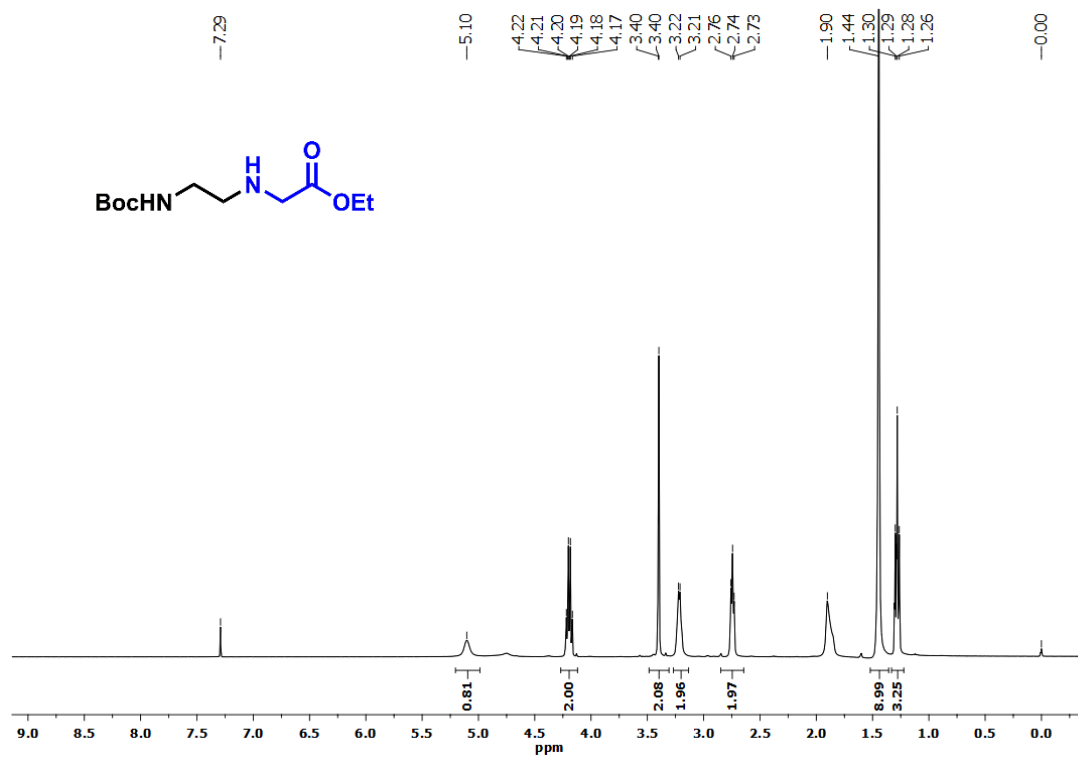
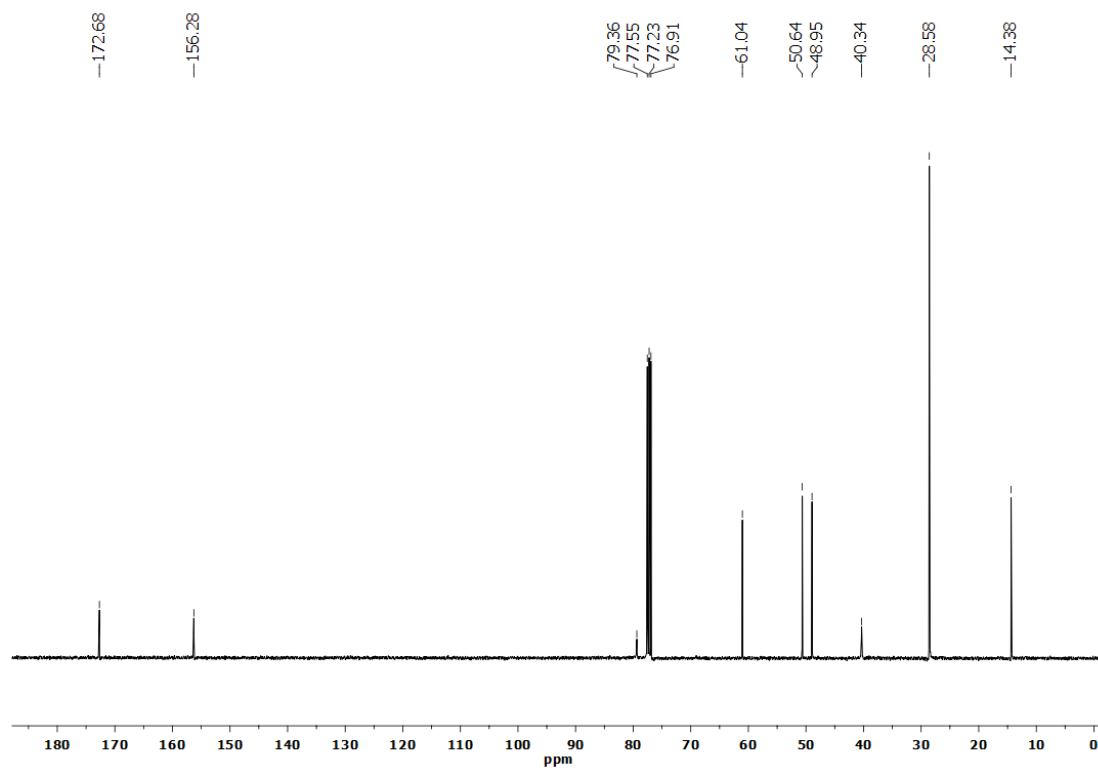
^1H NMR of Compound 30 ^{13}C NMR of Compound 30

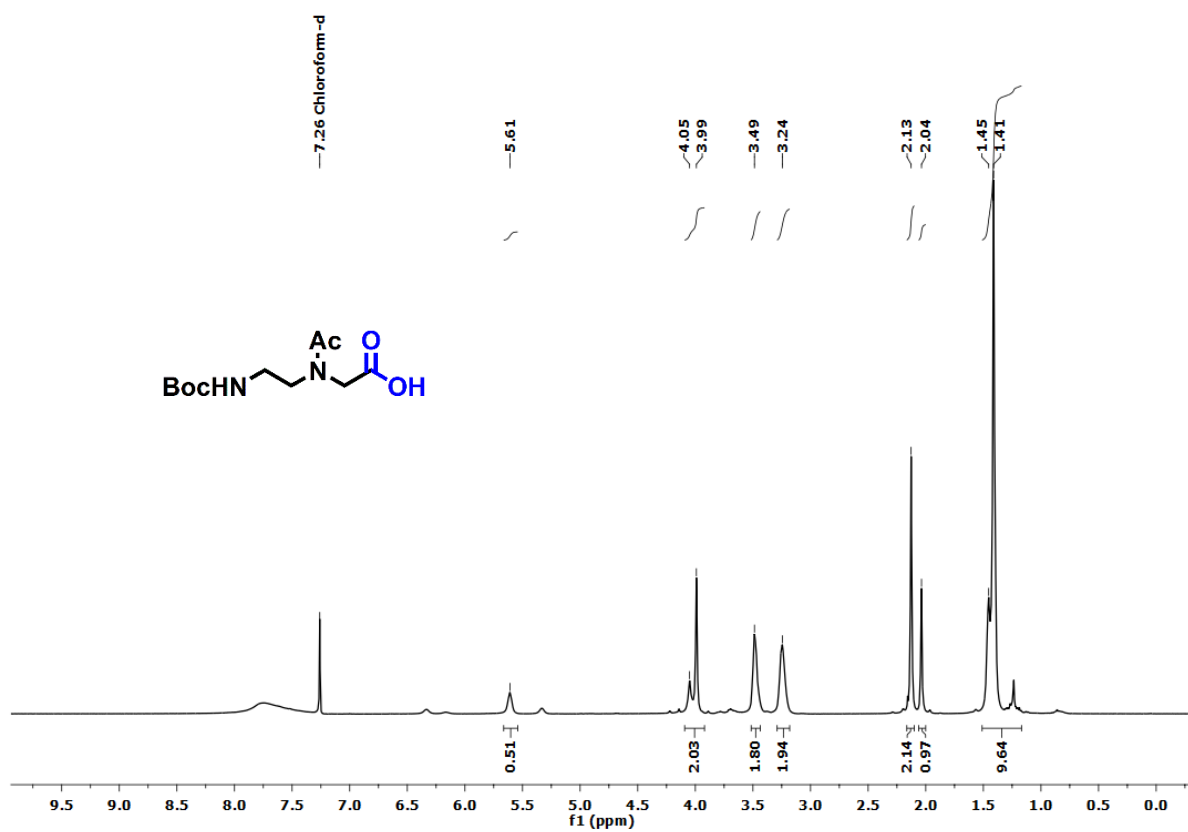
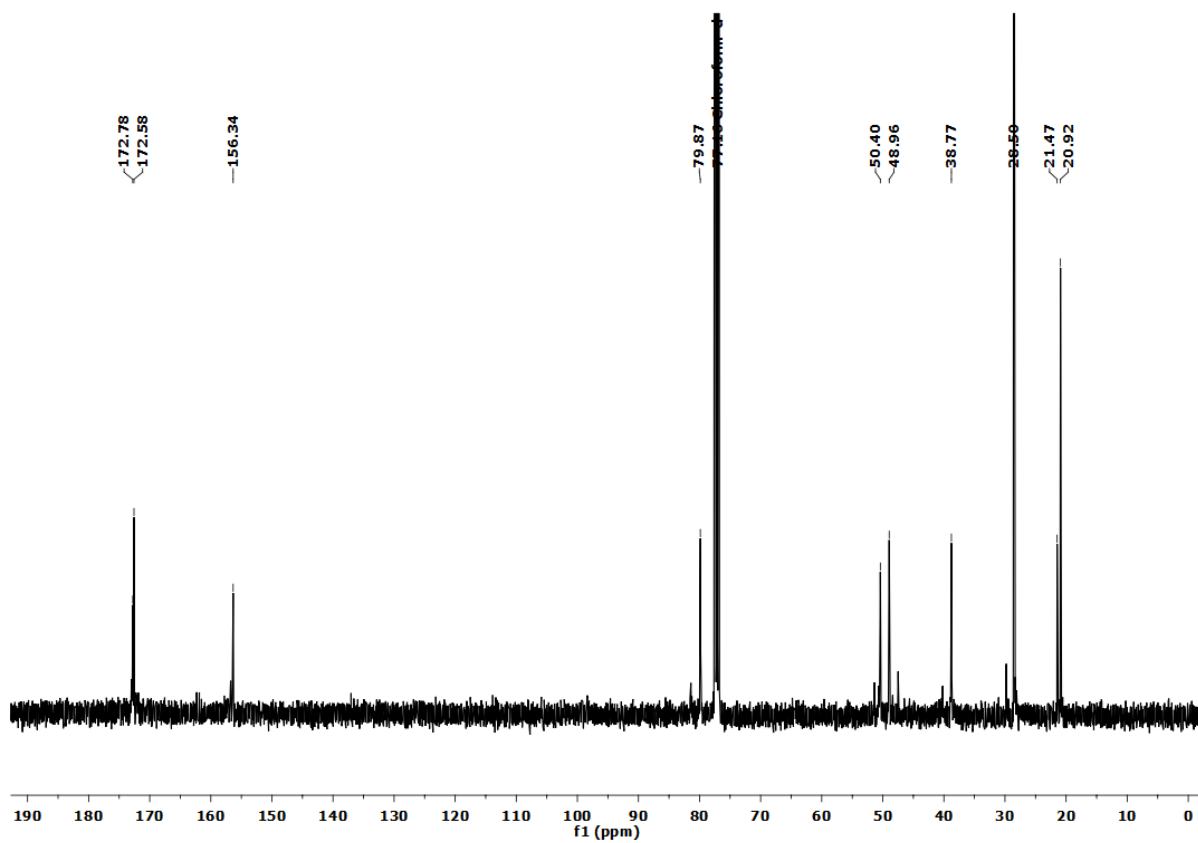
^1H NMR of Compound 32 ^{13}C NMR of Compound 32

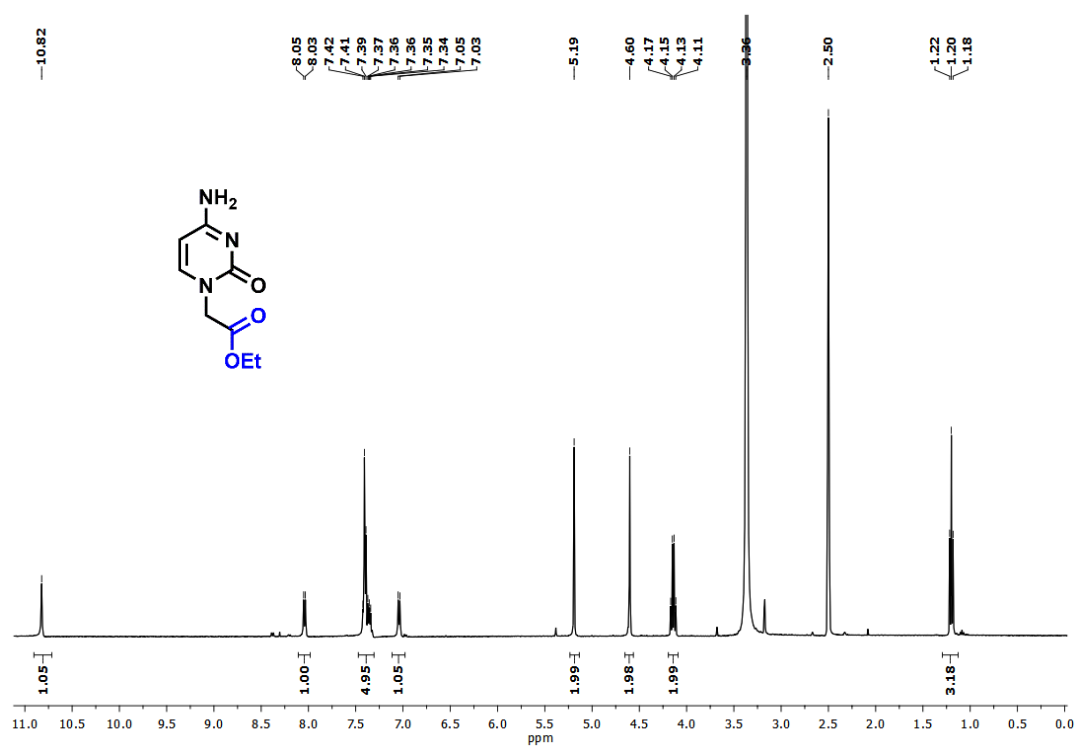
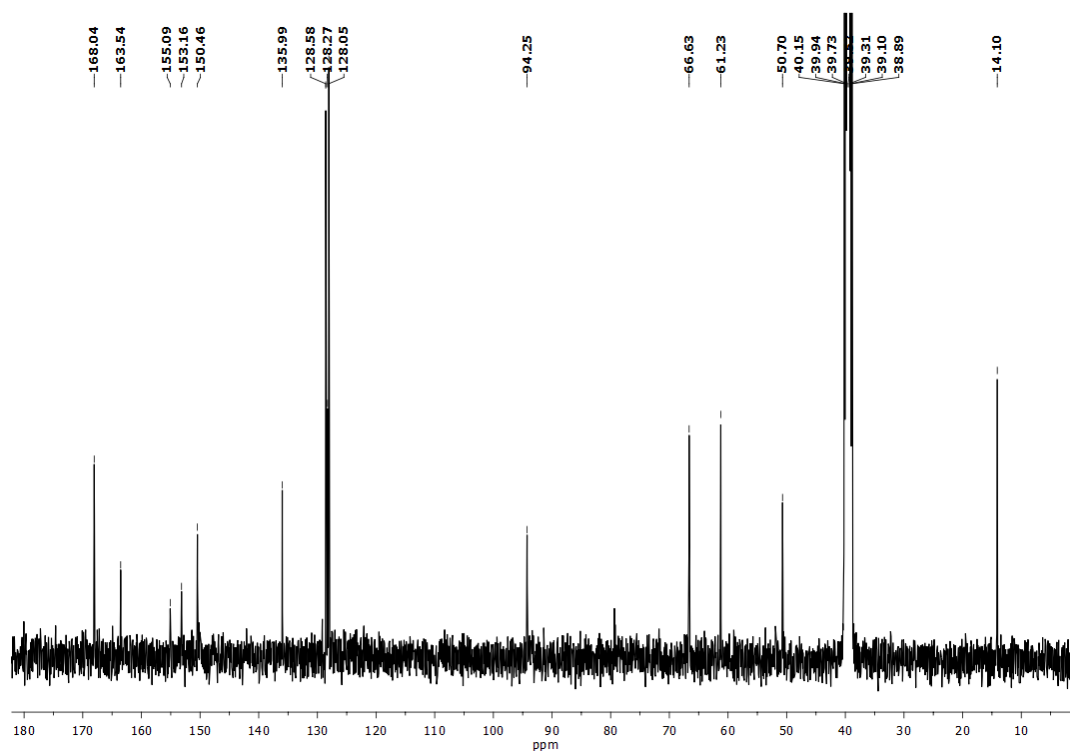
^1H NMR of Compound 33 ^{13}C NMR of Compound 33

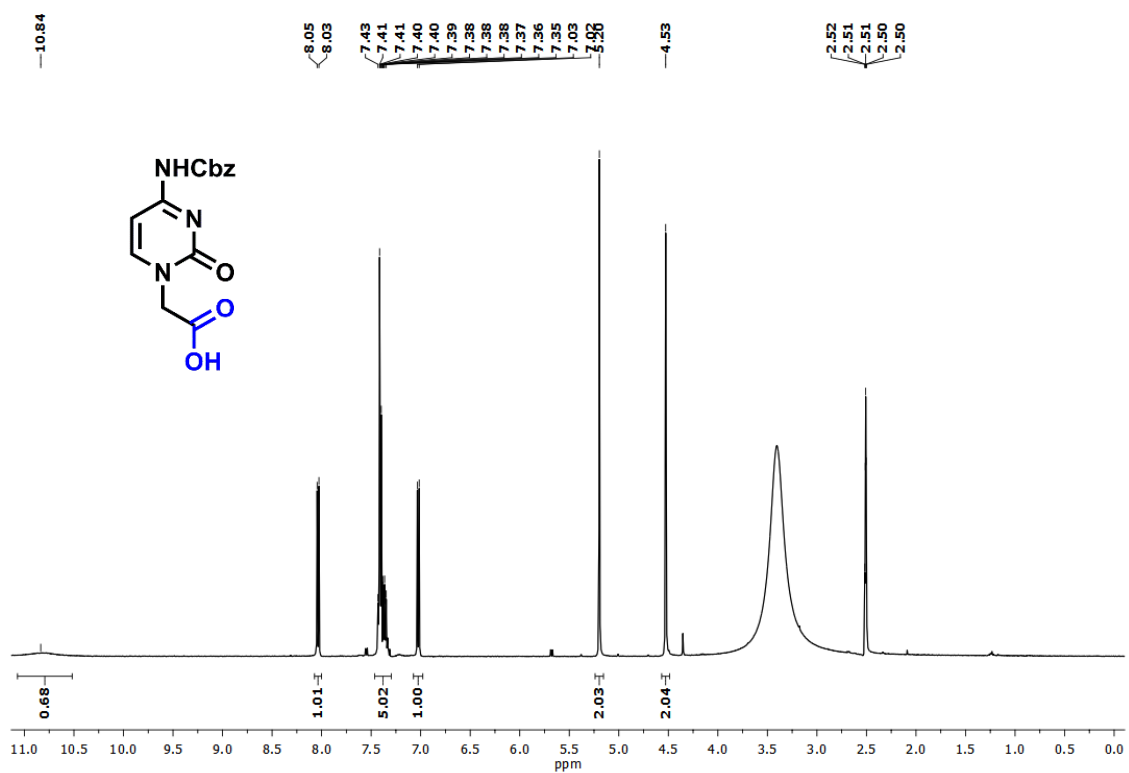
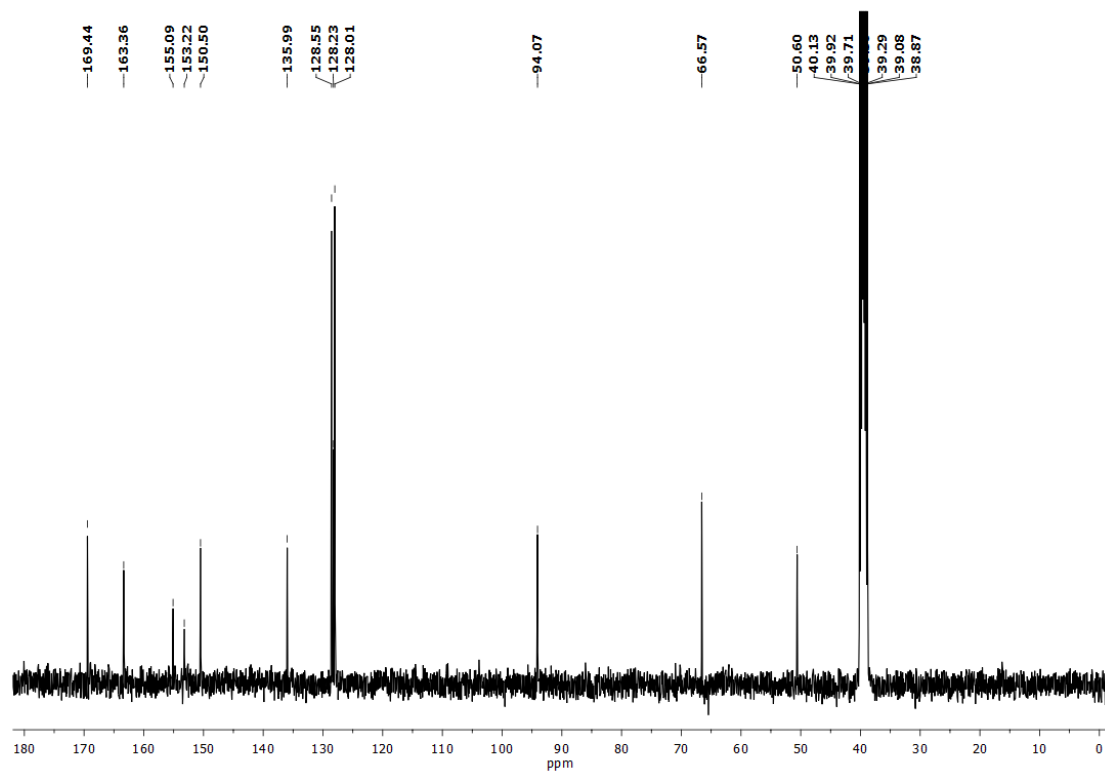
^1H NMR of Compound 34 ^{13}C NMR of Compound 34

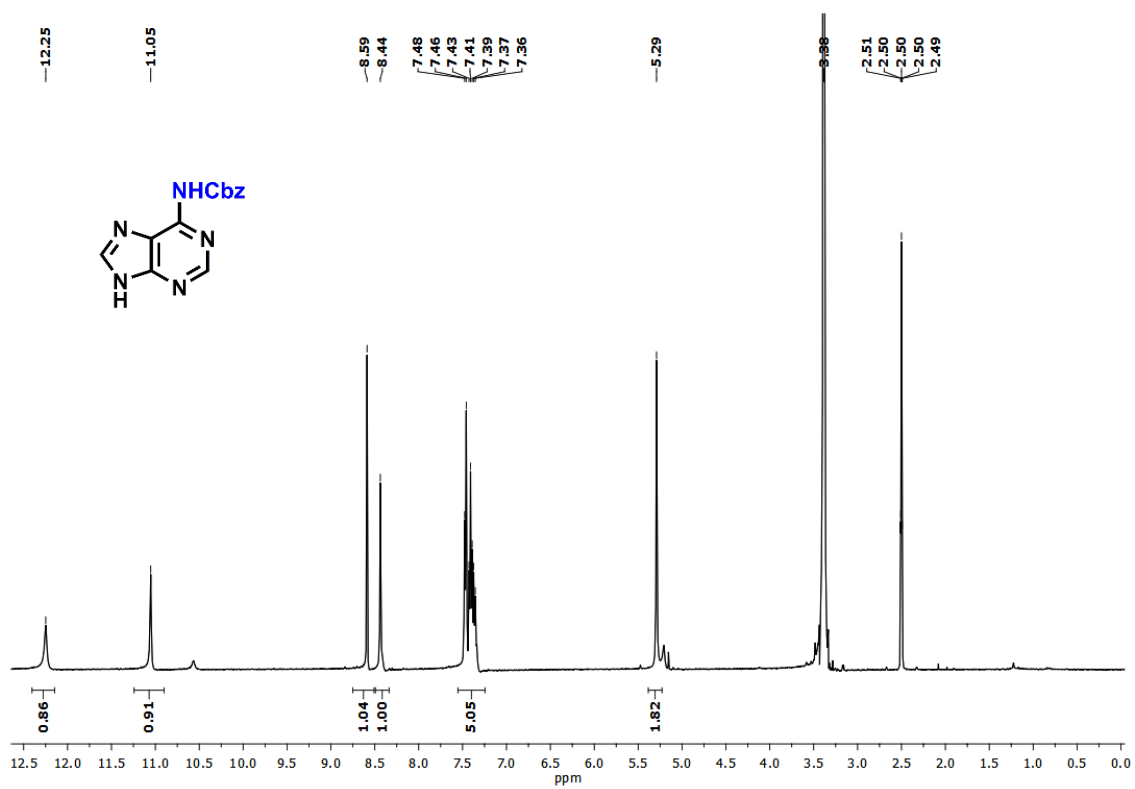
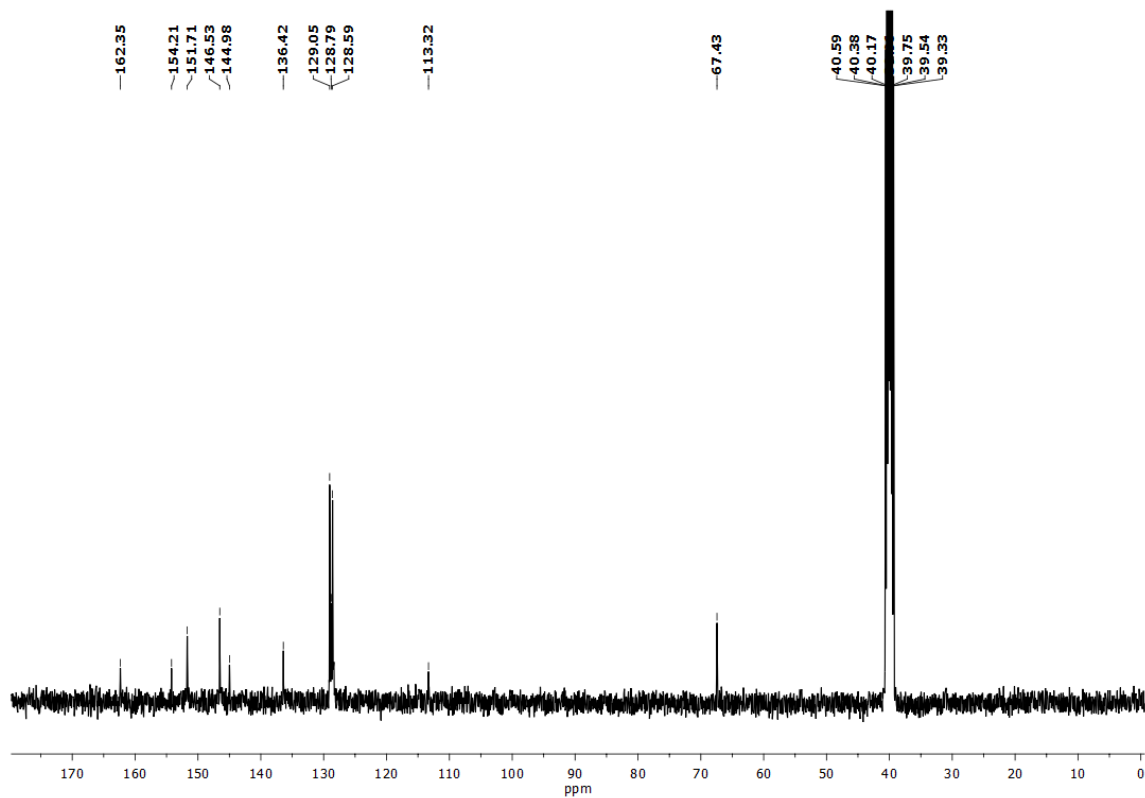
^1H NMR of Compound 36 ^{13}C NMR of Compound 36

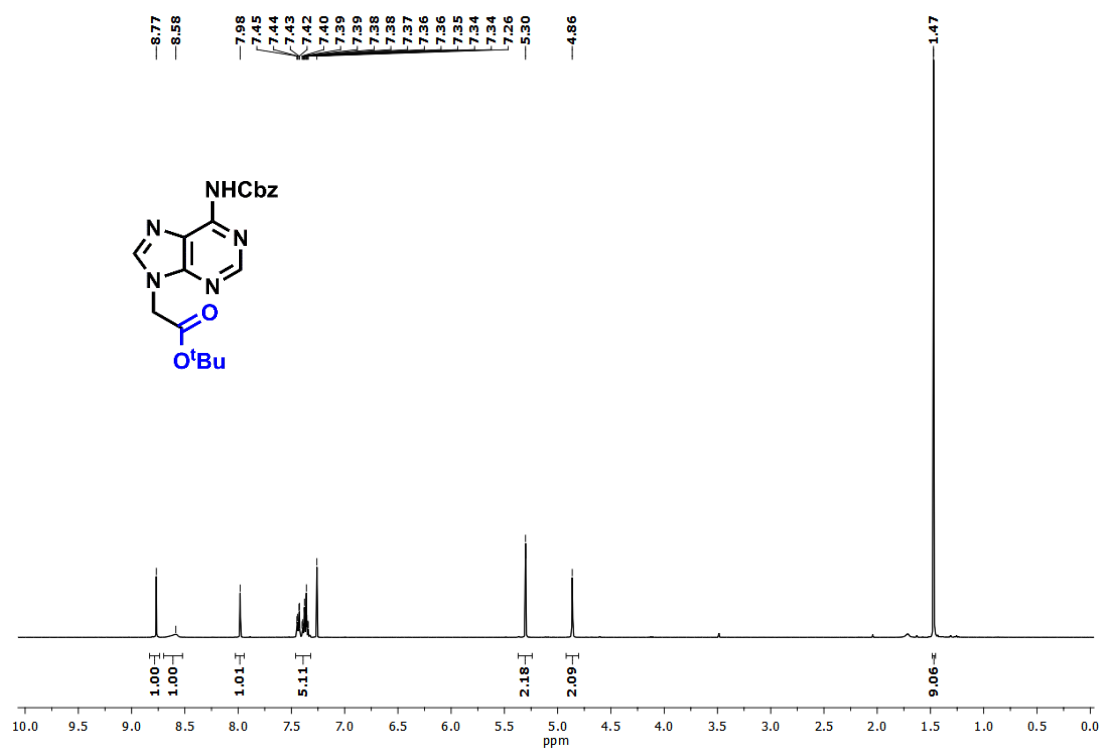
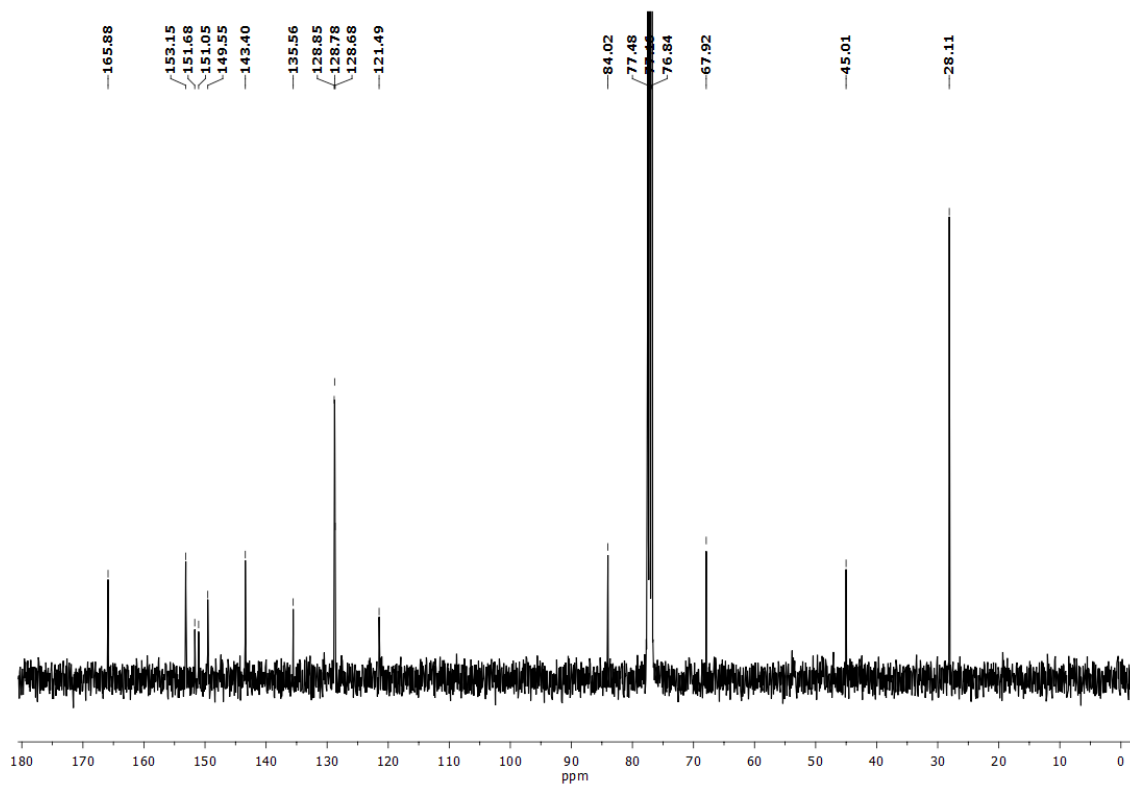
^1H NMR of Compound 39 ^{13}C NMR of Compound 39

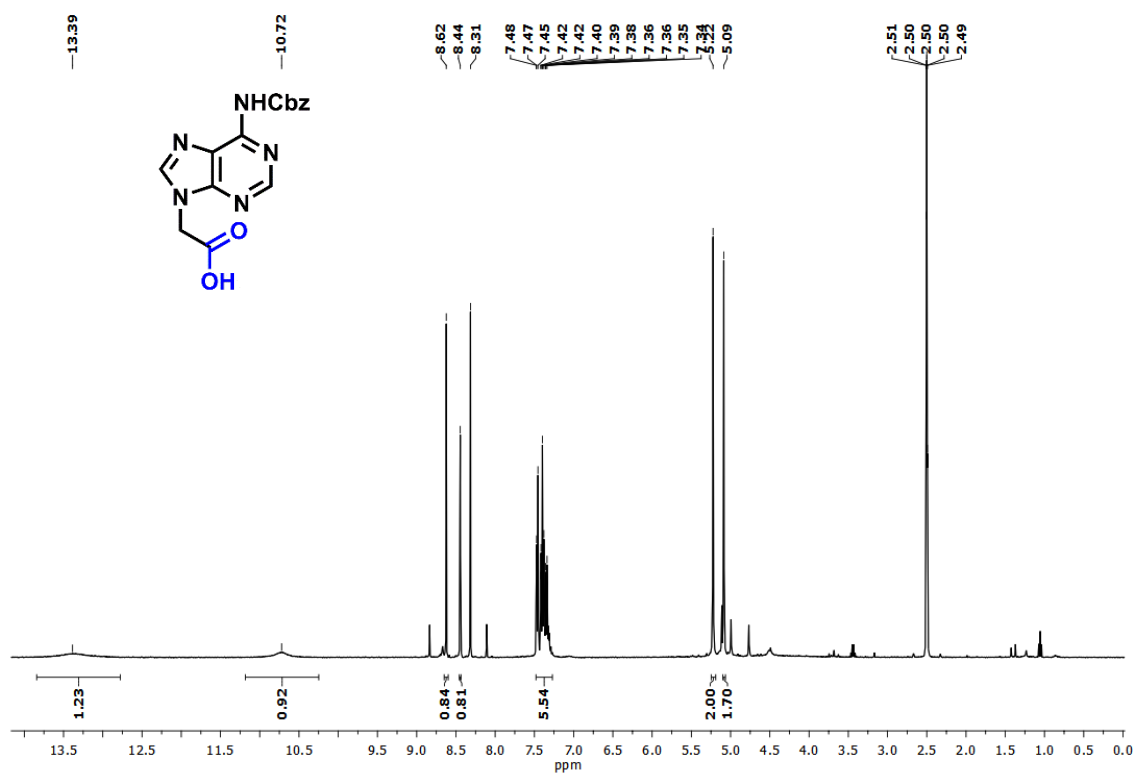
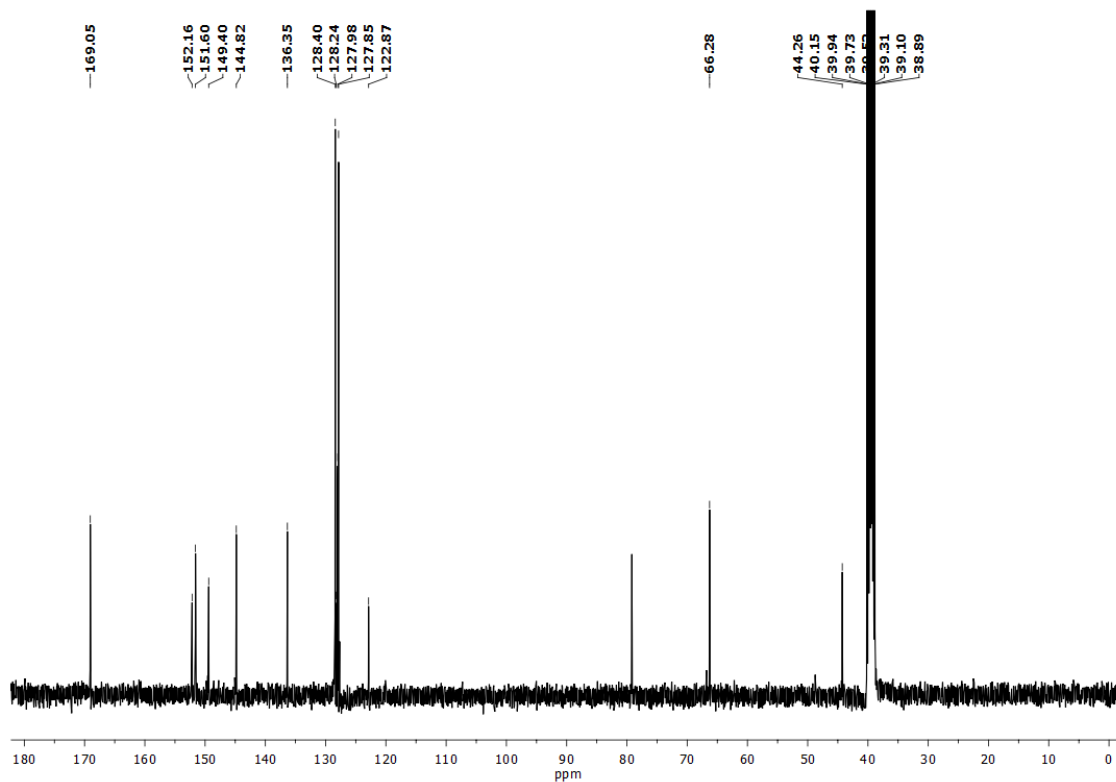
^1H NMR of Compound 41 ^{13}C NMR of Compound 41

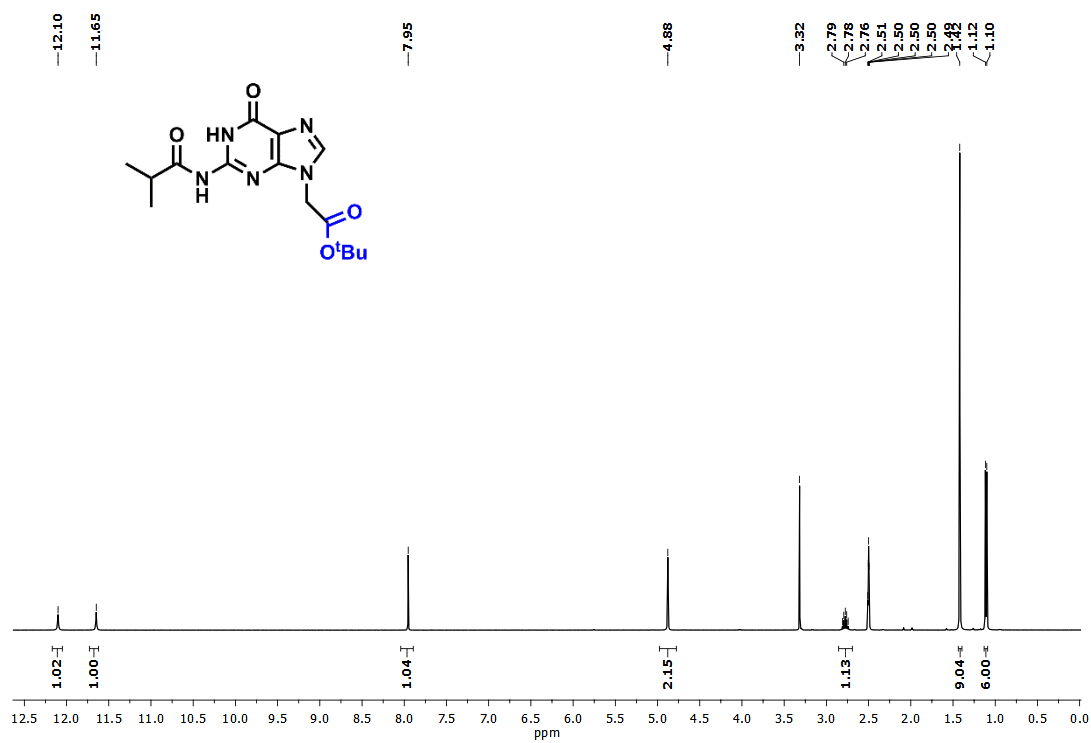
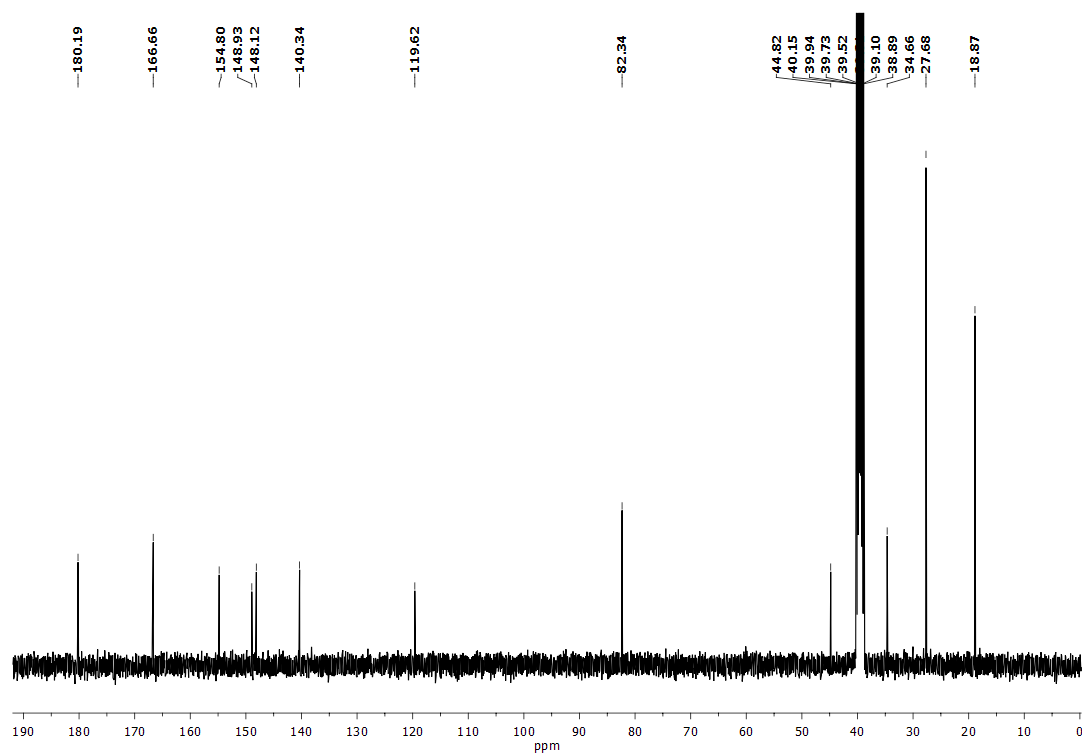
^1H NMR of Compound 44 ^{13}C NMR of Compound 44

^1H NMR of Compound 45 ^{13}C NMR spectrum of Compound 45

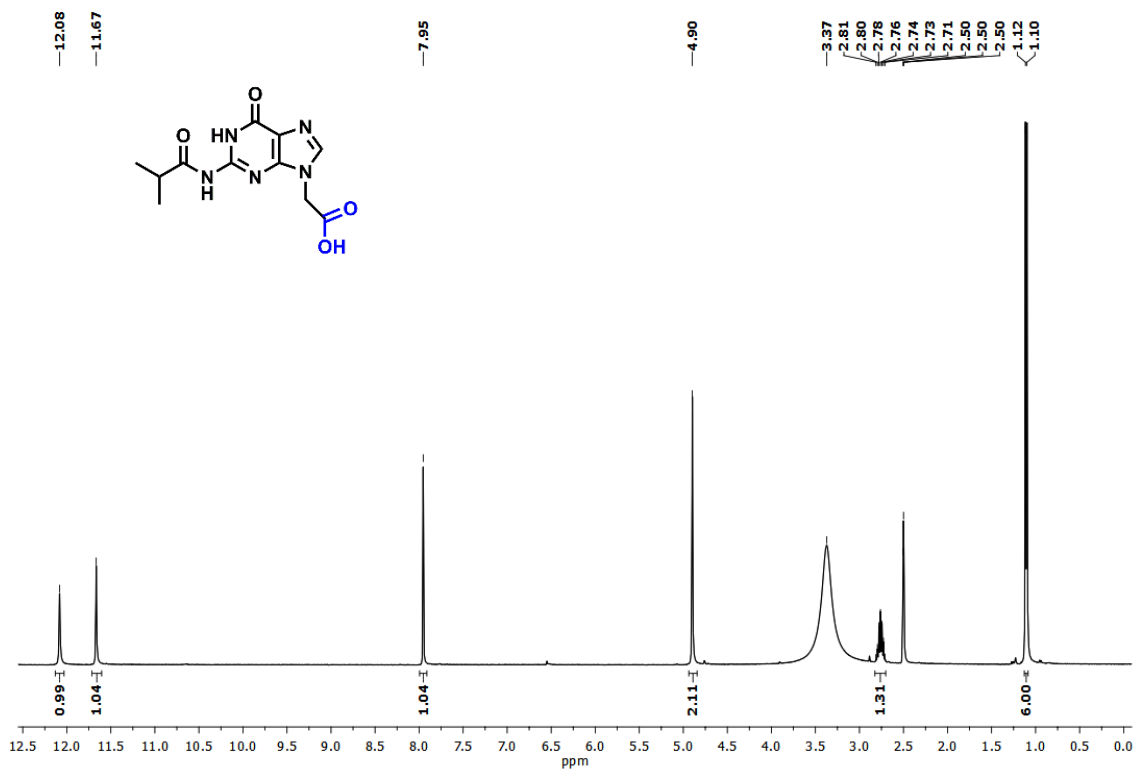
^1H NMR of Compound 47 ^{13}C NMR of Compound 47

^1H NMR of Compound 48 ^{13}C NMR of Compound 48

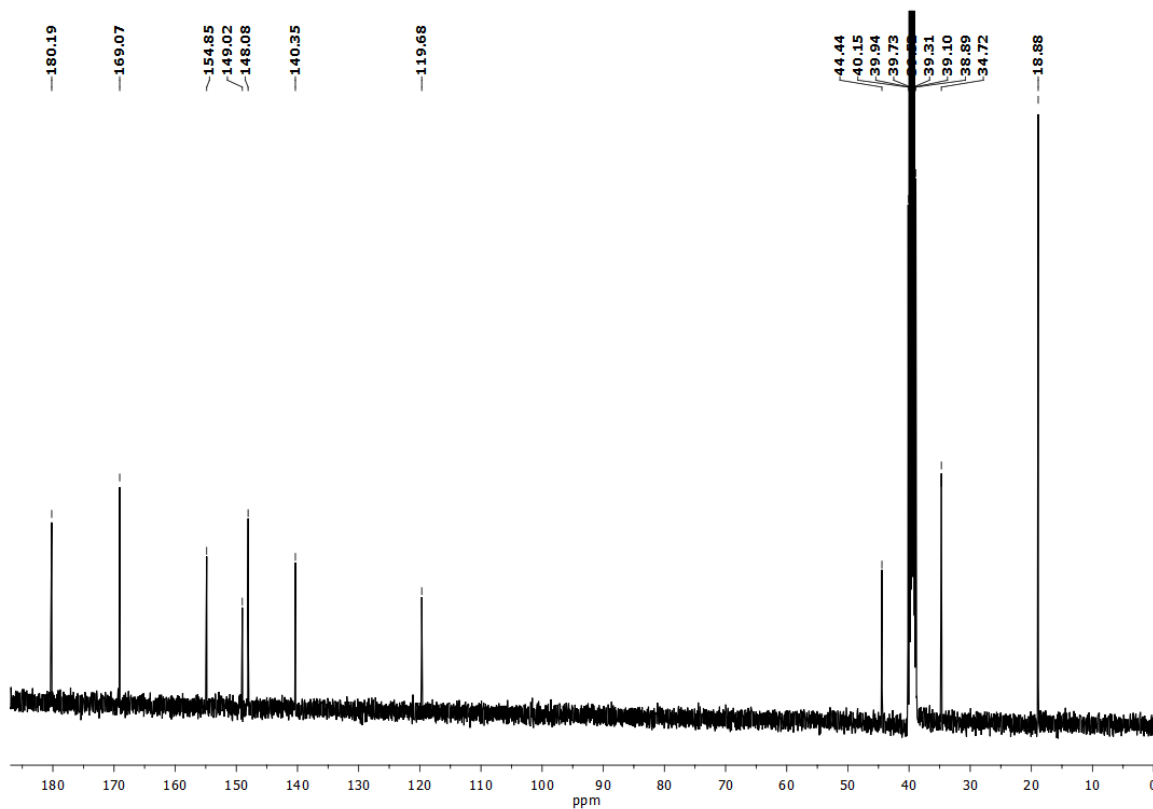
^1H NMR of Compound 49 ^{13}C NMR of Compound 49

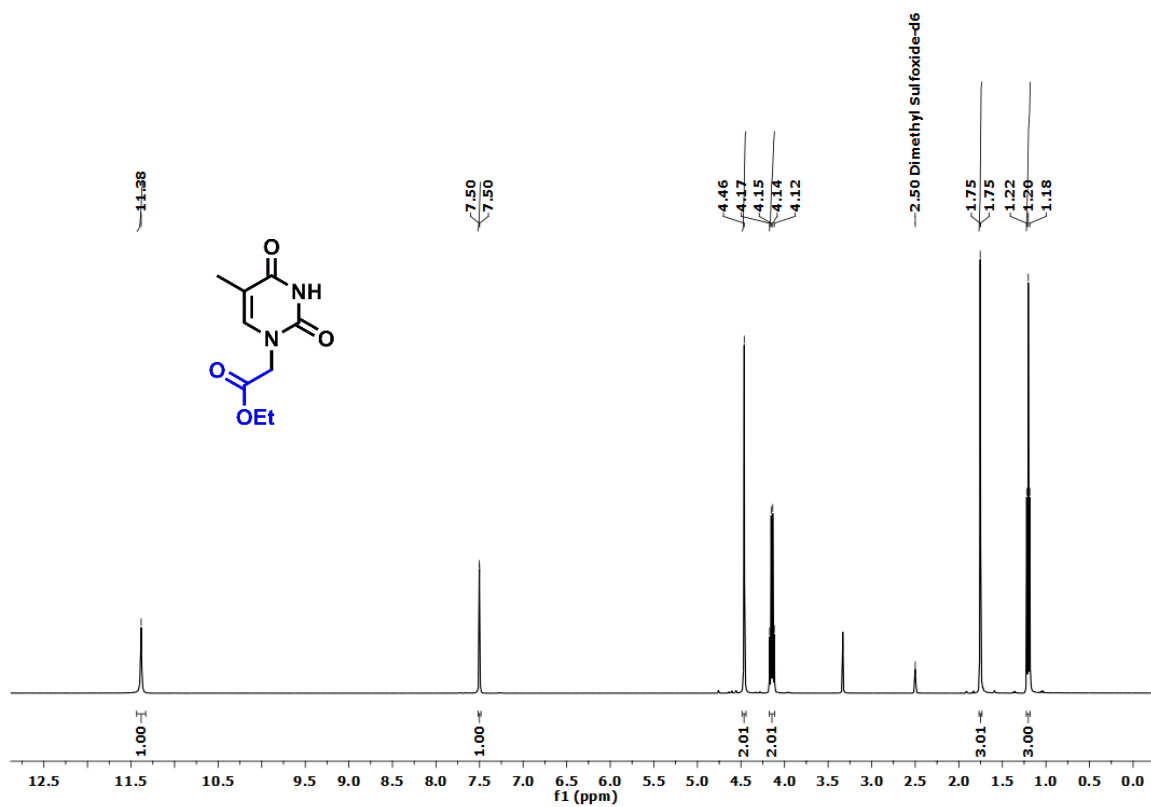
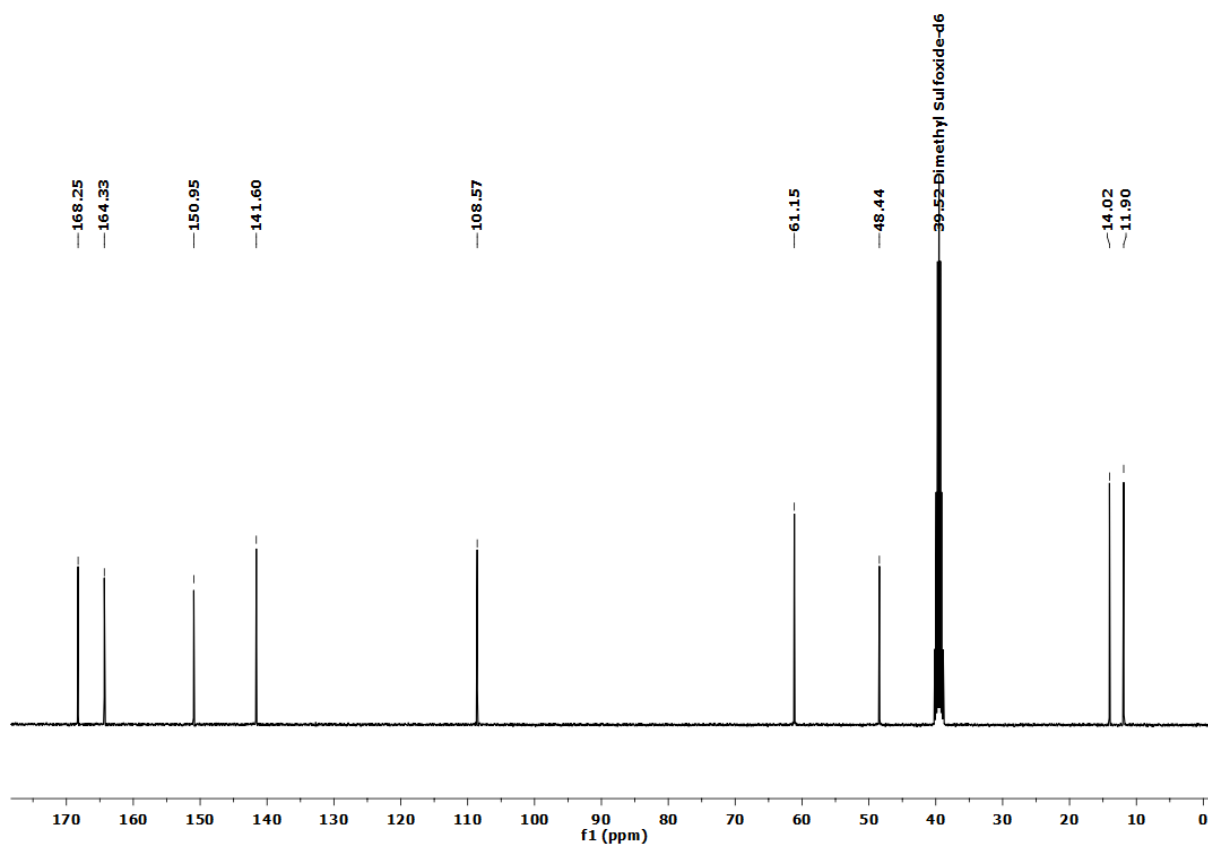
^1H NMR of Compound 52 ^{13}C NMR of Compound 52

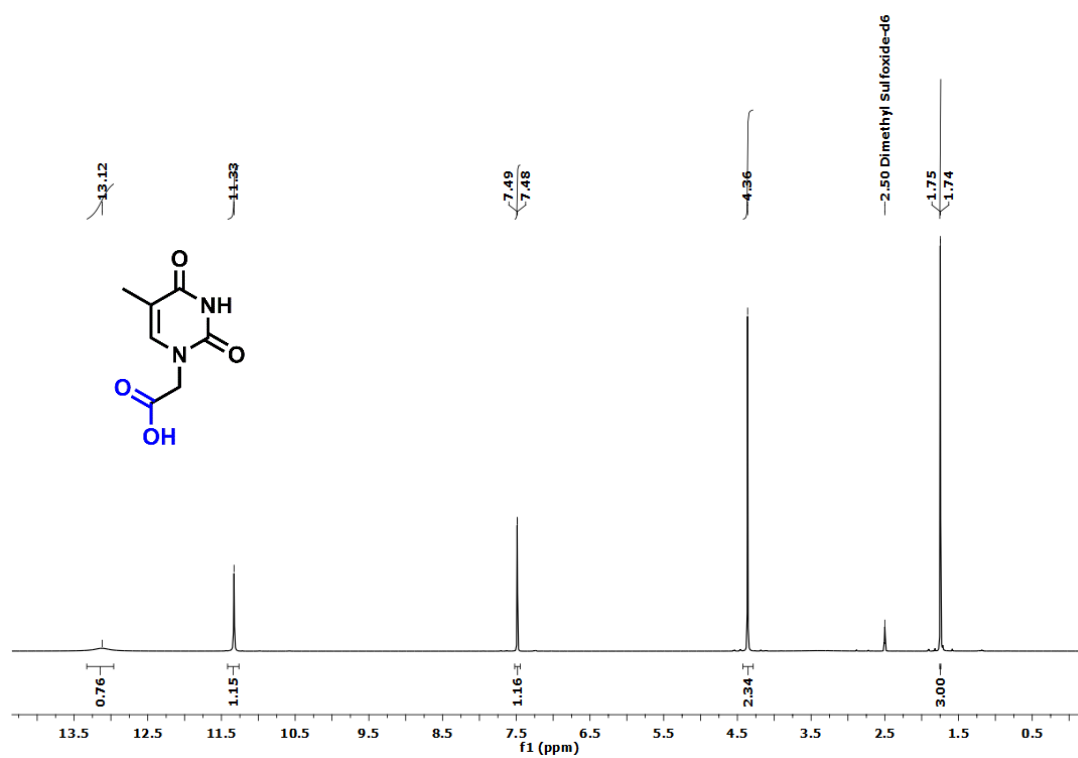
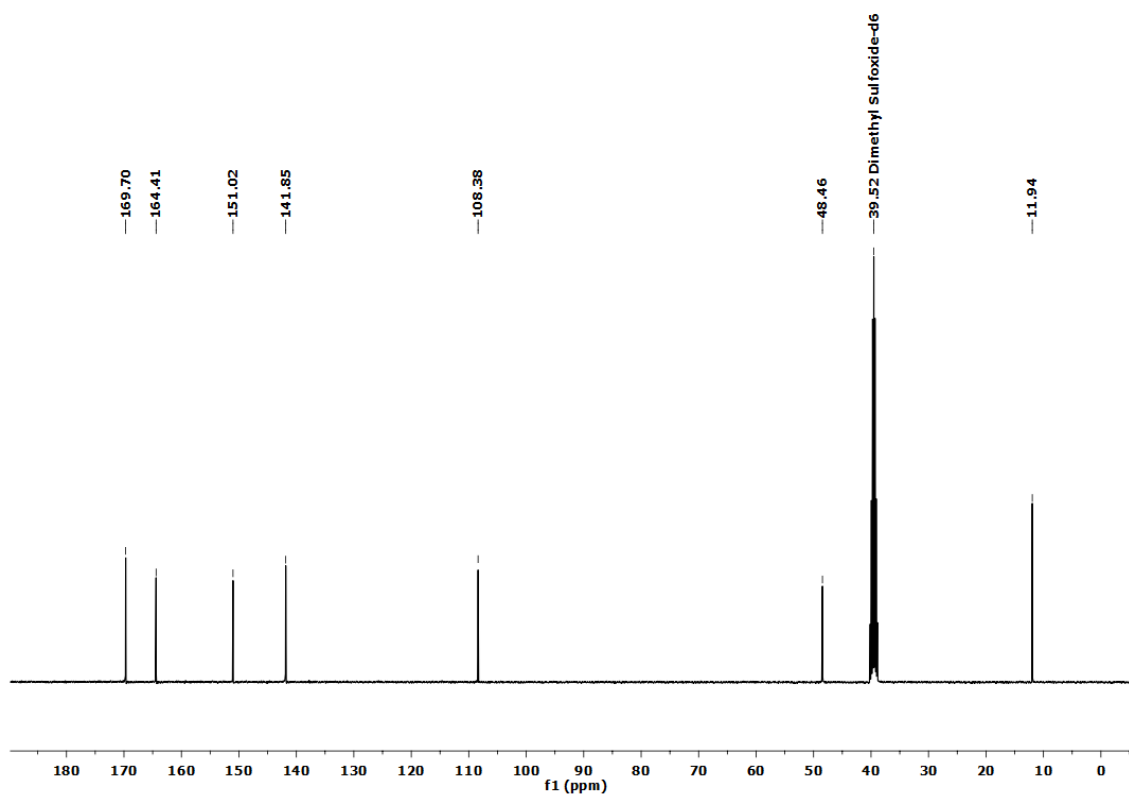
¹H NMR of Compound 53



¹³C NMR of Compound 53



^1H NMR of Compound 55 ^{13}C NMR of Compound 55

^1H NMR of Compound 56 ^{13}C NMR of Compound 56

4.8 References

1. Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404–412.
2. a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **1996**, *35*, 1939–1942.
b) Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998**, *39*, 4707–4710. c) Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 591–598.
3. Gourishankar, A.; Ganesh, K. N. *Artificial DNA: PNA & XNA* **2012**, *3*, 5–13.
4. Sugiyama, T.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 7317–7320.
5. Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. *Tetrahedron Lett.* **2005**, *46*, 8395–8399.
6. Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem.* **2000**, *4*, 931–943.
7. Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 591–598.
8. Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc. Chem. Commun.* **1993**, 518–519.
9. (a) Hyrup, B.; Egholm, M.; Nielsen, P. E.; Wittung, P.; Norden, B.; Buchardt, O. *J. Am. Chem. Soc.* **1994**, *116*, 7964–7968. (b) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed.* **1996**, *35*, 1939–1942.
10. Wojciechowski, F.; Robert H. E. Hudson *J. Org. Chem.*, **2008**, *10*, 3807–3816
11. (a) Mitra, R.; Ganesh, K. N. *Chem Commun.* **2011**, *47*, 1198–1200. (b) Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. *Eur. J. Org. Chem.* **2007**, 5879–5885.
12. Thadke, S. A.; Hridya, V. M.; Perera, J. D. R.; Gil, R. R.; Mukherjee, A.; Ly, D. H.; *Commun Chem.* **2018**, *1*, 79
13. Mitra, R.; Ganesh, K. N. *J. Org. Chem.* **2012**, *77*, 5696–5707
14. Huang, X.; Seid, M.; Keillor, J. W.; L. H.; Kauffman, G. S.; Pesti, J. A.; Yin, J. *J. Org. Chem.* **1997**, *62*, 7495–7496
15. Narindoshvili, A. R. T. *Org. Biomol. Chem.* **2008**, *6*, 3171–3176.
16. Thomson, S. A.; Josey, J. A.; Cadilla, R.; Gaul, M. D.; Fred Hassman, C.; Luzzio, M. J.; Pipe, A. J.; Reed, K. L.; Ricca, D. J.; Wiethe, R. W.; Noble, S. A. *Tetrahedron* **1995**, *51*, 6179–6194.
- 17 (a) Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 16456–16457. (b) Hudson, R. H. E.; Goncharenko, M.; Wallman, A. P.; Wojciechowski, F. *Synlett* **2005**, *2005*, 1442–1446.
18. Hudson, R. H. E.; Goncharenko, M.; Wallman, A. P.; Wojciechowski, F. *Synlett* **2005**, *2005*, 1442–1446.
19. N. Nagapradeep, S. Verma, *Chem. Commun.* **2011**, *47*, 1755–1757;
20. R. H. E. Hudson, M. Goncharenko, A. P. Wallman, F. Wojciechowski, *Synlett* **2005**, *2005*, 1442–1446.

21. Will, D. W.; Breipohl, G.; Langner, D.; Knolle, J.; Uhlmann, E. *Tetrahedron* **1995**, *51*, 12069-12082
22. Jenny, T. F.; Schneider, K. C.; Benner, S. A. *Nucleosides and Nucleotides* **1992**, *11*, 1257-1261.
23. Bernardes, G. J. L.; Linderoth, L.; Doores, K. J.; Boutureira, O.; Davis, B. G. *ChemBioChem* **2011**, *12*, 1383–1386.

Chapter 5

Design, Synthesis and Characterization of *C^γ-Janus* PNA Oligomers

5.1 Aim of the present work

The specific objectives of this chapter are

- Incorporation of $C^\gamma(S\text{-}eam)\text{-}aeg$ PNA (A/T/G/C) monomers and $aeg\text{-PNA}$ (T/C) monomers into *Janus* PNA sequences by solid phase synthesis
- Coupling of C^γ -amine with nucleobase acetic acids on solid phase to get *Janus* PNA oligomers.
- Cleavage of oligomers from the solid support, purification by RP-HPLC and characterization by MALDI-TOF spectrometry.

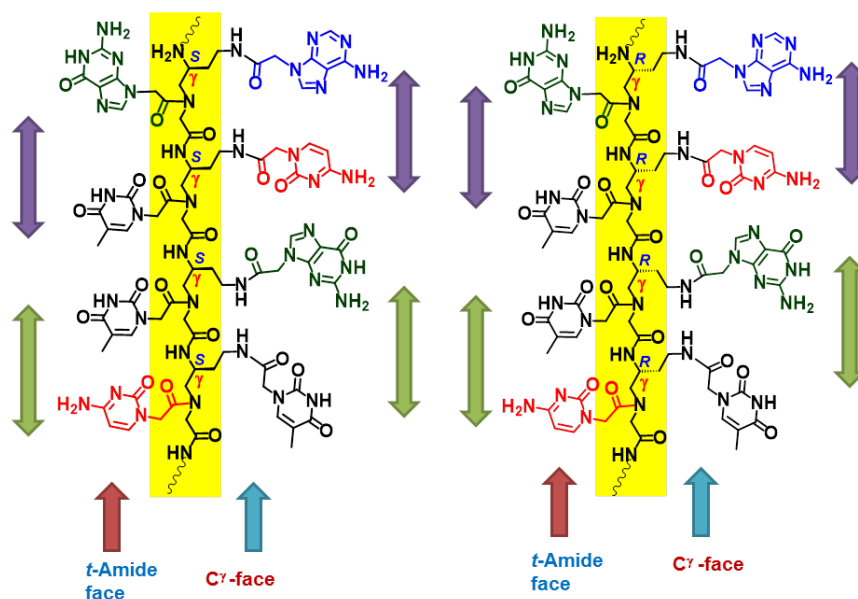


Figure 5.3. Representation of a *S* and *R* *Janus* oligomer having an *t*-amide and a C^γ face.

Generally, PNAs form stable duplexes of antiparallel orientation with their complementary sequences. By convention, in the antiparallel orientation of PNA:DNA duplexes, the *N*-terminus of the PNA faces the 3'-end of the DNA and *C*-terminus faces the 5'-end of DNA. In parallel orientation¹, the *C*-terminus of PNA faces the 3'-end of the DNA and the *N*-terminus faces the 5'-end of the DNA (Figure 4.4).

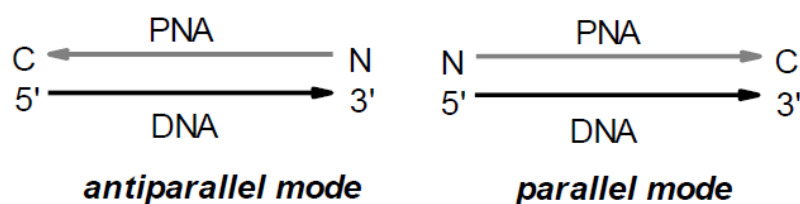


Figure 3.5 Antiparallel and parallel binding of PNA to DNA

In order to study the potential of duplex formation and the discrimination in DNA binding of the C γ -substituted PNAs, various kinds of *Janus* PNA oligomers were synthesized. These include -

- *Homo Janus PNA of S/R* (pyrimidine nucleobases on both faces, T on amide face and C/G on C γ face)
- *Hetero Janus PNA of S/R* (mixed purine-pyrimidine sequences on both amide and C γ face)

The appropriate *aeg* PNA sequences bearing A/T/C/G bases only on the amide face and PNAs having bases only on the C γ face (amino ethylglycyl C γ PNA oligomers) as controls to examine the hybridization effectiveness through double duplex formation by *Janus* PNAs through their supramolecular self-assembly.

5.2 Rationale for Choice of sequence

The conformational rigidity of PNAs is enhanced by the presence of the chiral centers, that may also enhance ability of PNA strands and enrich hybridization competent conformations and, ultimately, increase the selectivity in molecular recognition. A simple C γ -backbone modification can transform a randomly folded peptide nucleic acid (PNA) into a right-handed helix². These conformationally preorganized helical PNAs bind to DNA and RNA with exceptionally high affinity and sequence selectivity³. The two molecular entities are prepared from the same monomeric chemical scaffold, with the *S* and *R* stereochemistry at the γ -backbone that determines if the corresponding oligo adopts a right-handed or left-handed helix. These conformers hybridize to each other with exquisite affinity, sequence selectivity, and level of orthogonality. This has been demonstrated by biophysical, crystal structure and NMR studies of a L-serine based chiral PNA-DNA duplex³. Taking this into consideration, PNA **9** (Figure 4.6) that contains (A/T/G/C) units on the amide face and S-ethylamino substitution at C γ position of glycine was used as a primary motif for design of all *Janus* PNAs. The *homo S/R*, and *hetero S/R Janus* PNAs with mixed sequences were synthesized incorporating various kinds of modified PNA monomer units as (A/T/G/C).

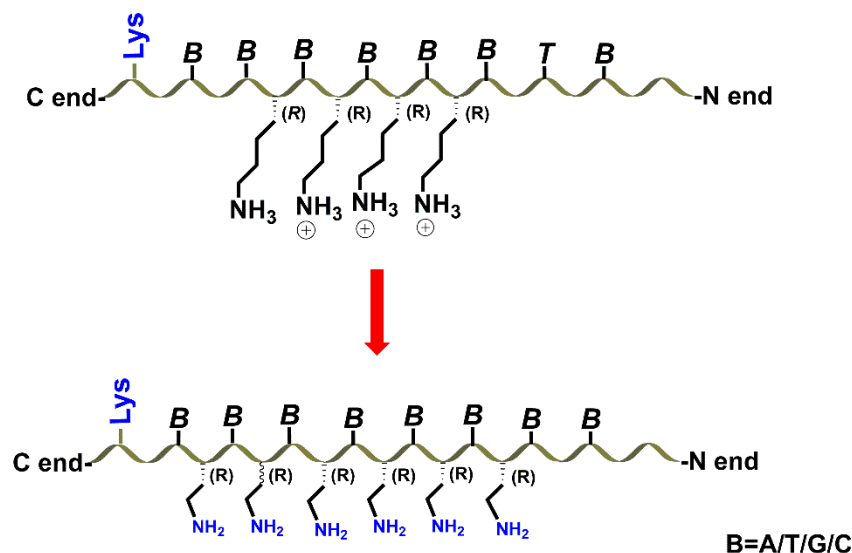


Figure 5.4 Sequence of the 8-mer PNA (PNA 9) chosen for synthesis of Janus PNA

Based on the modifications on the two faces (t-amide and C γ face), various kinds of *Janus*-PNA oligomers were synthesized as shown in (Figure 4.7).

- i) *S Homo Janus* PNA oligomers
- ii) *R Homo Janus* PNA oligomers
- iii) *S Hetero Janus* PNA oligomers
- iv) *R hetero Janus* PNA oligomer
- iv) Aminoethyl glyceryl C γ PNA oligomers (*aeg- γ* PNA)

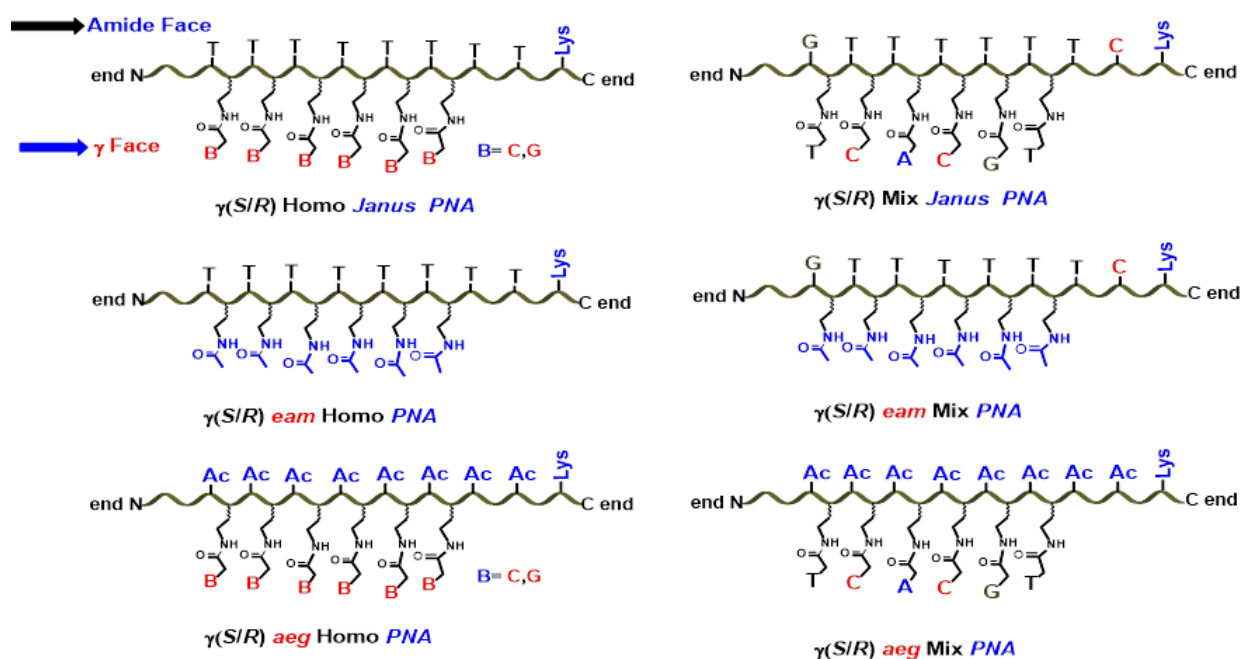


Figure 5.6 Different Types of *Janus* PNA oligomers

5.3 Results and discussion

This section describes the synthesis, purification and characterization of *Janus PNA* oligomers incorporating modified (C^γ -ethylamine/ N^{Ac}) monomers as well as unmodified *aeg* PNA monomers (Figure 5.7) into PNA sequences at desired positions. After the assembly of *Janus PNA* oligomers and deprotection, purification was done by HPLC followed by characterization using mass spectral data.

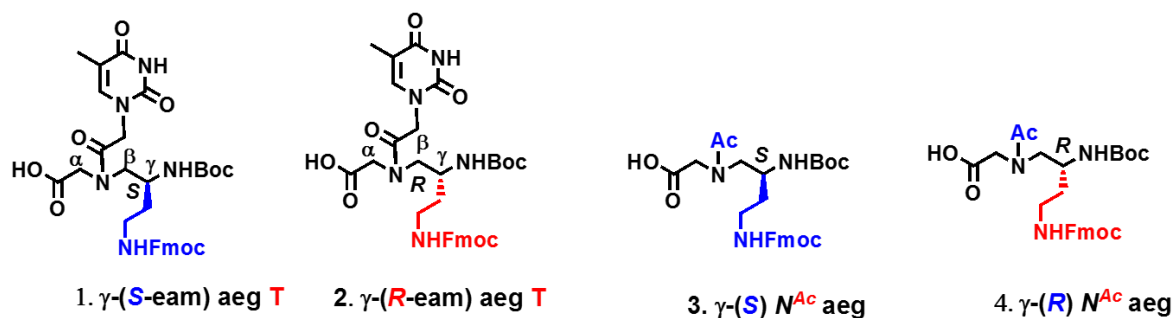


Figure 5.7 Chemical Structures of *Janus PNA* monomers.

5.3.1 Solid phase Synthesis of PNA oligomers

The synthesis of PNA oligomers was carried out using solid phase synthesis protocol using the Boc strategy on MBHA (4-methyl-benzhydryl amine) resin² and Fmoc strategy⁴ on rink amide resin. The modified PNA monomers were incorporated at specific positions in *aeg*-PNA sequences. PNA oligomer synthesis was carried out from the C-terminus to N-terminus end using modified monomeric units with protected amino and carboxylic acid functions maintaining the orthogonality.

MBHA resin (4-methyl-benzhydryl amine resin) was used as the solid support on which the oligomers were built. The monomers were coupled by *in situ* activation with HBTU/HOBt. The synthesis of all oligomers was initiated by first coupling the orthogonally protected (Boc/Cbz/Bhoc/Fmoc) L-lysine (*S*) as the C-terminal spacer-amino acid, linked to MBHA resin through an amide bond. The amine loading on the resin was reduced from 0.6 mmol/g to 0.2 mmol/g by partial acylation of amine groups using the calculated amount of acetic anhydride.⁵ This was done to increase the efficiency of coupling by lessening steric crowding on the resin. The free amine groups on the resin available for coupling was confirmed before starting the synthesis by Kaiser's test.

The deprotection of the Boc protecting group and the completion of coupling reactions were monitored by Kaiser's test,^{6,7} which is the most widely used qualitative test for the presence or absence of free amino group (deprotection/coupling). The *t*-Boc

deprotection leads to a positive Kaiser's test in which the resin beads show blue color (Rheumann's purple). After completion of the coupling reaction, the resin beads will become colorless leading to a negative Kaiser's test. Using the standard solid phase synthesis protocol (Figure 5.4)⁸, the PNA oligomers of the desired length incorporating PNA monomers were synthesized.

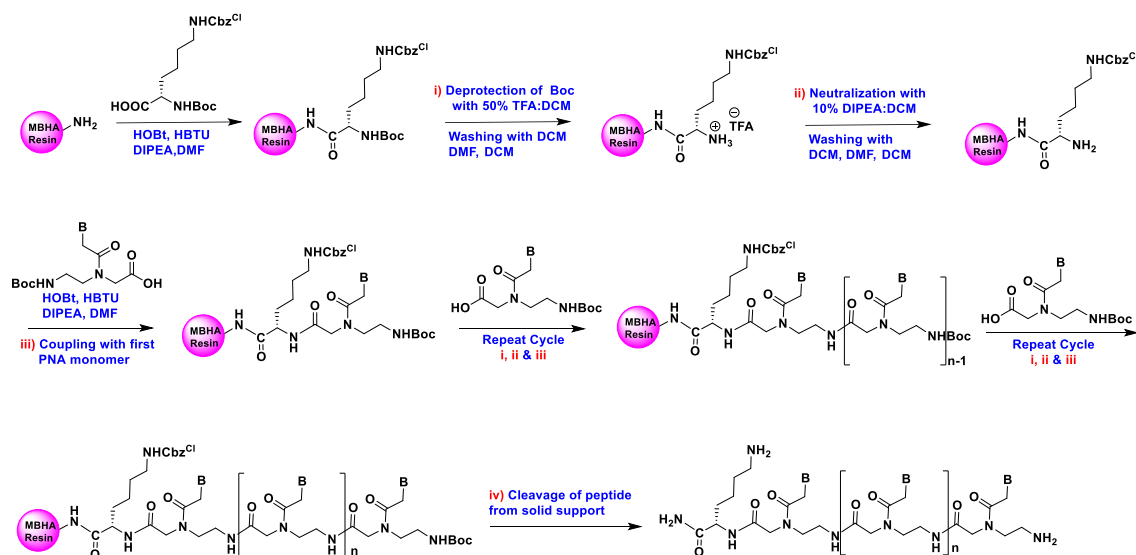


Figure 5.8 Solid phase PNA synthesis protocol by *Boc* strategy; B = Nucleobase such as Adenine (A), Guanine (G), Cytosine (C) and Thymine (T).

5.3.2 Synthesis of mixed purine-pyrimidine *Janus* PNA Oligomers

The strategy employed for the synthesis of various kinds of *Janus* PNA oligomers is described in this section.

5.3.2a Synthesis of homo *Janus* PNA oligomer by *Boc* chemistry on solid phase.

PNA monomers were incorporated into 8-mer PNA sequence by solid phase synthesis on L-lysine derivatized MBHA functionalized resin having 0.20 mmol/g loading value⁹. Deprotection and neutralization followed by coupling of monomers (Figure 5.8) and repetition of the same steps in every cycle gave the homo-oligomeric PNAs T₈-*aeg-γS-eam*₆ (PNA) and T₈-*aeg-γR-eam*₆ (PNA) (Figure 5.9) having six ethylamine side chains at C^γ with *S* and *R* stereochemistry respectively. In these PNAs, both C and N termini used unmodified nucleobases to avoid the steric hindrance in synthesis and DNA hybridization reactions. After the assembly of PNA sequences, and single step deprotection of all sidechain Fmoc groups, global coupling reaction of C^γ-ethyl amino groups on solid phase with cytosinyl-N1-acetic

acid (**11**) using the coupling agents HBTU, HoBt and DIEA gave resin bound Janus PNAs $T_{8jp-\gamma S-C_6}$ and $T_{8jp-\gamma R-C_6}$. These were cleaved from the resin by reaction with TFA/TFMSA to yield *homo Janus* PNA oligomers $T_{8jp-\gamma S-C_6}$ (**JP 1S**) and $T_{8jp-\gamma R-C_6}$ (**JP 2R**) (Figure 5.8). Global coupling with guaninyl-N9-acetic acid gave *homo Janus* PNA oligomers $T_{8jp-\gamma S-G_6}$ (**JP 3S**) and $T_{8jp-\gamma R-G_6}$ (**JP 4R**).

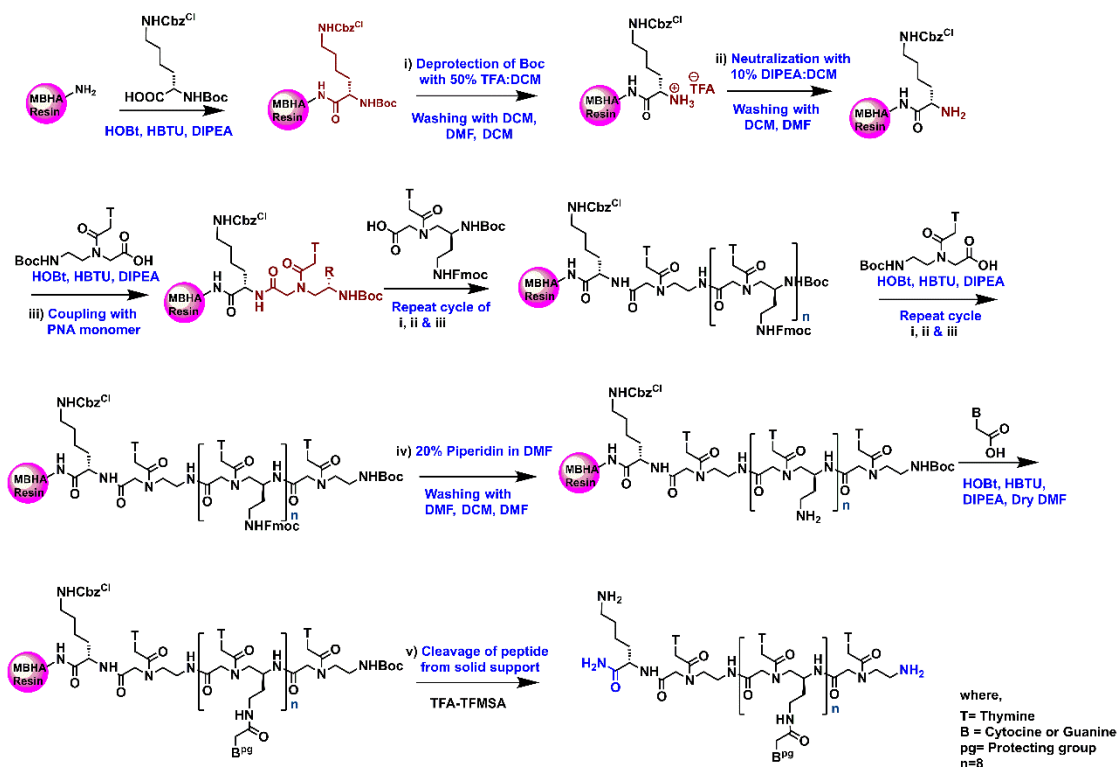


Figure 5.9 Solid phase *homo Janus* PNAs (**JP 1S**) and (**JP 1R**) synthesis protocol by *Boc* strategy

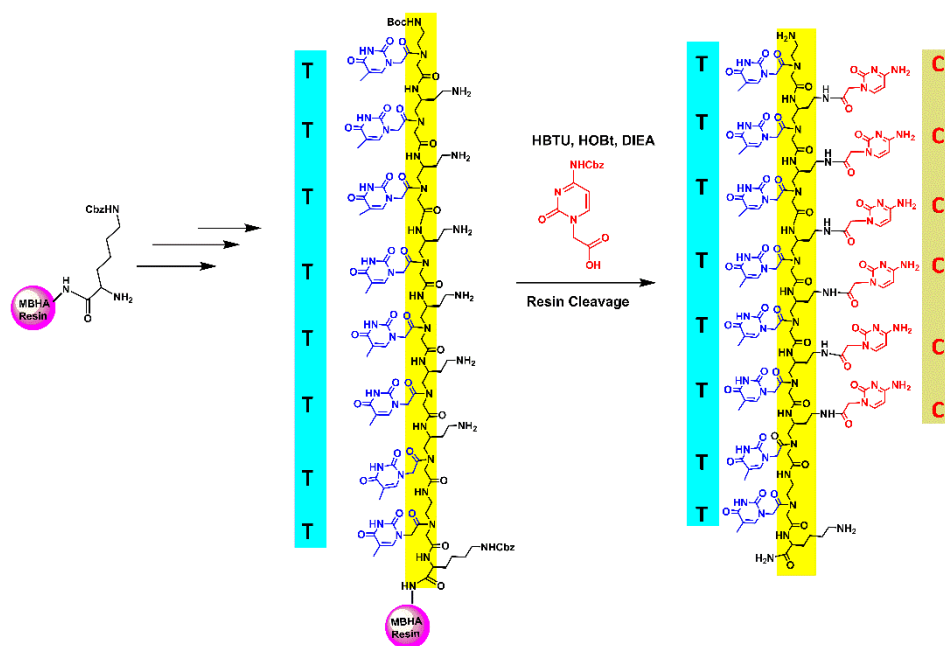


Figure 5.10 Synthetic strategy for *homo Janus* PNA oligomer (**JP 1** to **JP 4**)

5.3.2b Synthesis of hetero Janus PNA with mixed sequence on both amide and C γ face. The hetero Janus PNAs that possess mixed purine/pyrimidine bases on both the amide and γ faces were synthesized by using combination of boc and Fmoc protocols on MBHA resin. In order to synthesise hetero Janus PNAs having T sequence on t-amide face and mixed sequence on C γ -face aeg monomer with pre-installed base T on amide face and C γ side chain with NHFmoc (Figure 5.7) was synthesized. Each coupling of (S/R)-C γ -ethylamine PNA-T monomer was followed by deprotection of Fmoc on side chain and reaction of the C γ -side chain NH $_2$ with desired (A/C/G/T)-acetic acid to generate the Janus PNA fragment at every step. Such stepwise alternating boc/Fmoc synthetic protocols⁹ lead to building sequences on both sides simultaneously to final hetero Janus PNAs with Ts on t-amide face and mixed sequence on C γ -face. The synthesized hetero Janus PNAs were cleaved from resin purified by HPLC to yield **JP 5S** and **JP 6R**. They were characterized by mass spectral data. To ensure successful and efficient coupling at all steps, a small amount of the resin was cleaved at intermediate stages of synthesis at difficult steps of coupling and the intermediate products were characterized by mass spectral data.

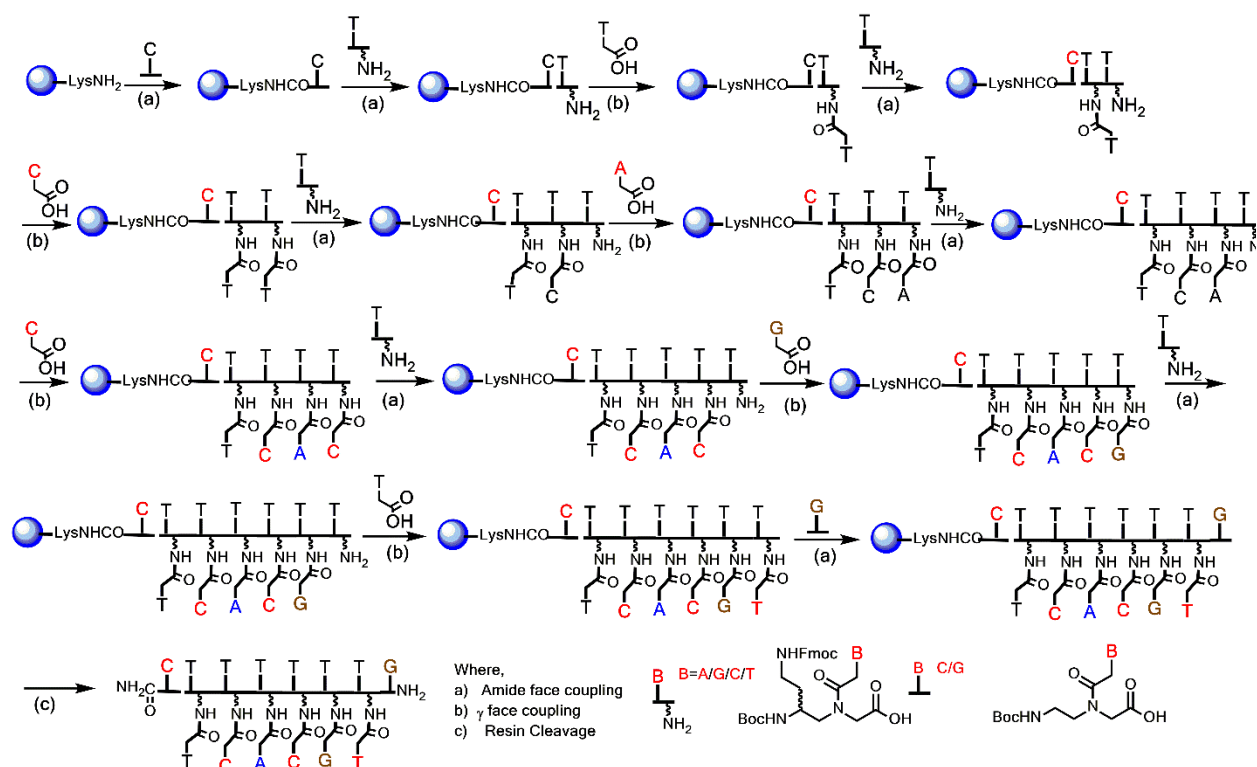


Figure 5.11 Synthesis strategy of *mjp*- γ -PNA oligomer (**JP 5S** and **JP 6R**) on solid support.

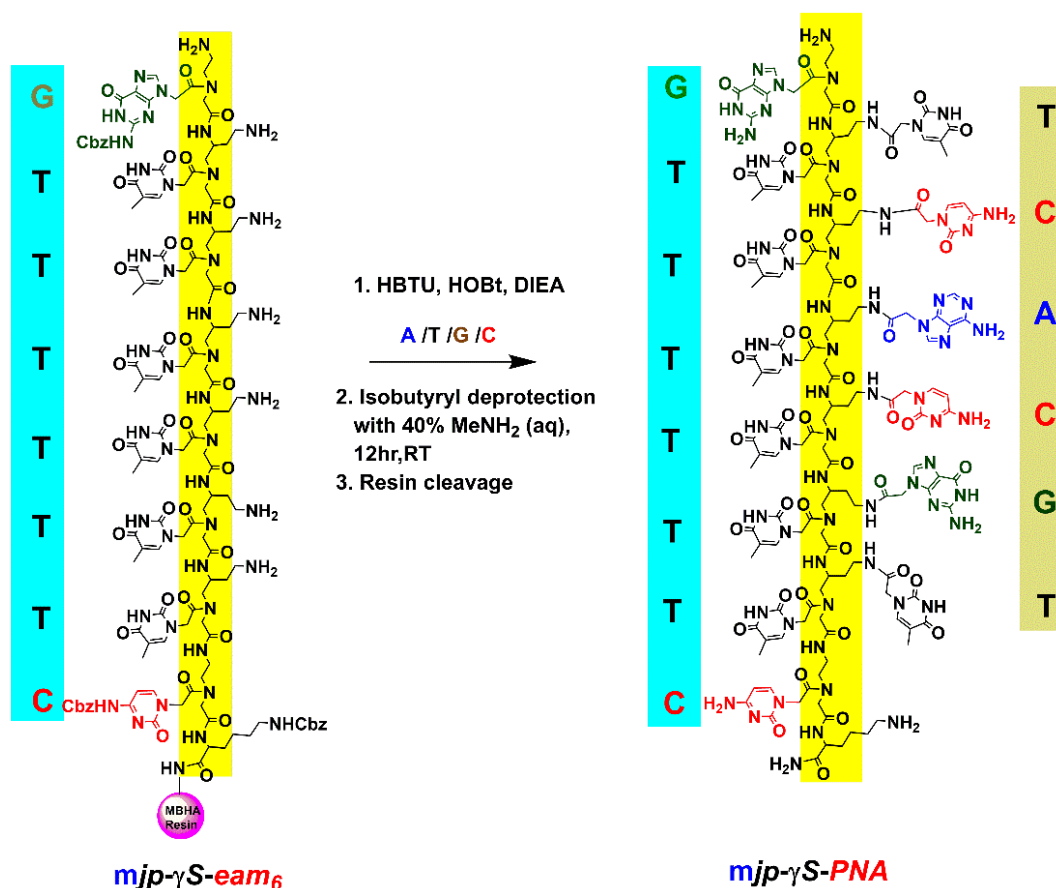


Figure 5.12 Synthetic strategy of *mjp- γ S-PNA* oligomer (JP 6)

5.3.2c Synthesis of PNA analogues with nucleobase on only aminoethyl glycol C γ face.

For studying hybridization exclusively from C γ -face, synthesis of aminoethylglycine backbone oligomers having bases only on C γ -substituted face (*aeg- γ* PNA), and without bases on t-amide face and the backbone amino groups protected with acetyl group were synthesized (Figure 5.14). This involved the use of modified monomers γ S-(*eam*)-N^{Ac}-*aeg* (**8S**) and γ R-(*eam*)-N^{Ac}-*aeg* (**8R**) and individual coupling on resin followed by global deprotection and reaction with nucleobase acetic acids (A/T/G/C) (**11**) to get the corresponding C γ derived homo PNAs p8- γ -C₆ (PNA **13S** and **13R**) and p8- γ -G₆ (PNA **14S** and **14R**)¹⁰ and mixed sequence p8- γ -*mjp* (PNA **15S** and **16R**) with secondary NH^{Ac} groups present on the backbone. After solid phase assembly and deprotection, the oligomers were purified by HPLC and characterized by mass spectral data.

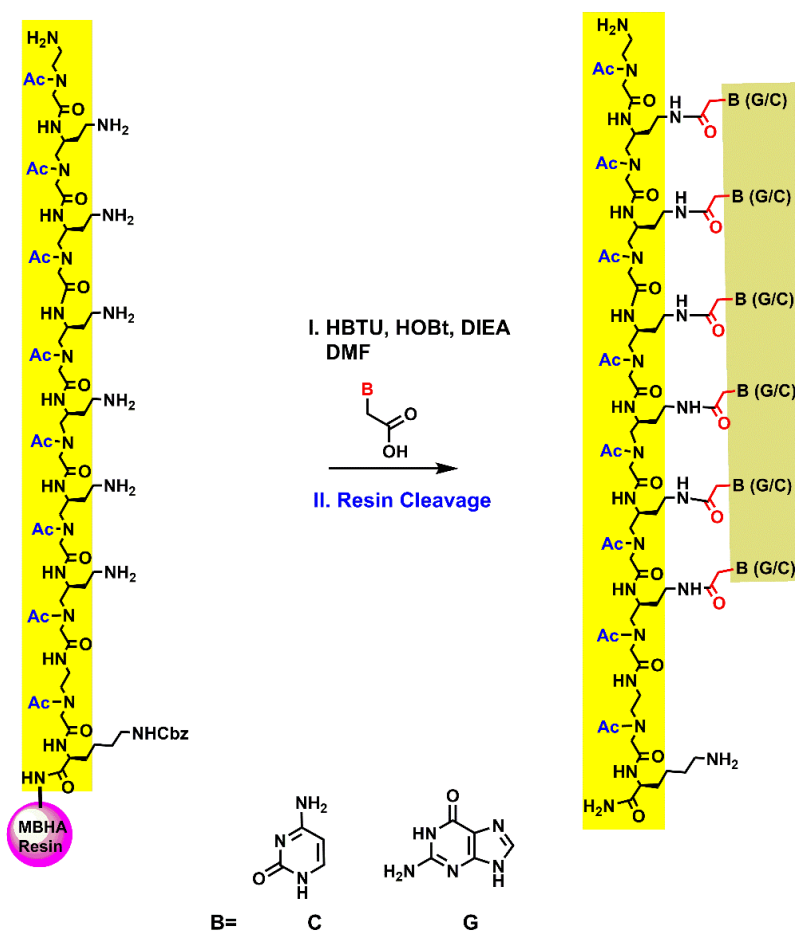


Figure 5.13 Strategy for solid phase synthesis of aminoethyl glyceryl C^{γ} ethylamino PNA (*aeg- γ -eam₆* PNA) oligomers (PNA 12 to PNA 14)

5.3.2d Synthesis of PNA analogues with nucleobase on only tertiary amide face *aeg* PNA 7-10

Similarly as C^{γ} face, the nucleobase on only t-amide face synthesized for studying hybridization exclusively. Synthesis of aminoethylglycine backbone oligomers having bases only on t-amide face as normal PNA and amine group of C^{γ} face protected with acetyl group. (Figure 5.16). This involved the use of modified monomers $\gamma(S\text{-eam})aeg\text{-}T$ and $\gamma(R\text{-eam})aeg\text{-}T$ and individual coupling on resin followed by global deprotection and reaction with Acetic anhydride to get the corresponding homo PNAs $T_{8jp}\text{-}\gamma(S)\text{-}NH^{Ac}$ (PNA-7) and $T_{8jp}\text{-}\gamma(R)\text{-}NH^{Ac}$ (PNA-8) and mixed sequence $mjp\text{-}\gamma(S)\text{-}NH^{Ac}$ (PNA-9) and $mjp\text{-}\gamma(R)\text{-}NH^{Ac}$ (PNA-10). After solid phase assembly and deprotection, the oligomers were purified by HPLC and characterized by mass spectral data.

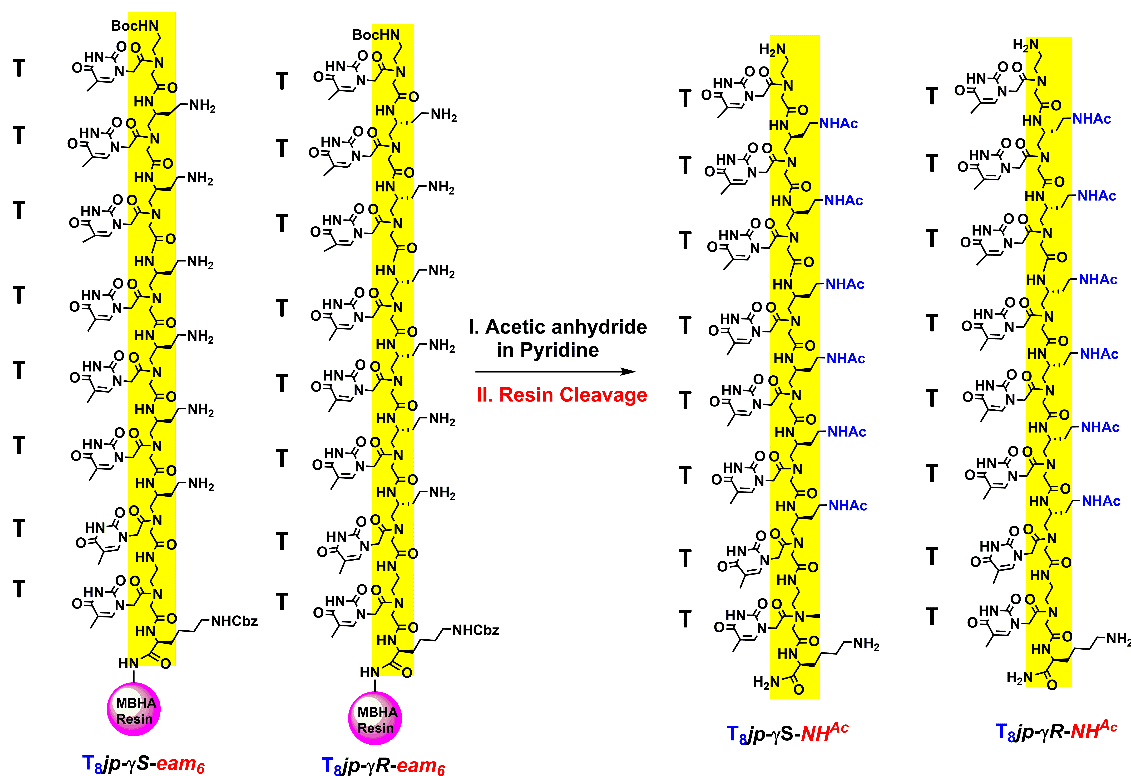


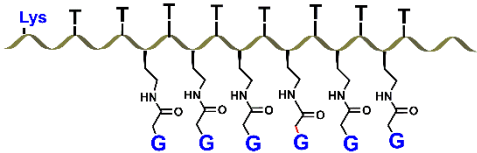
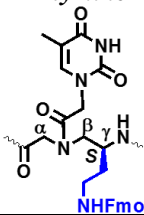
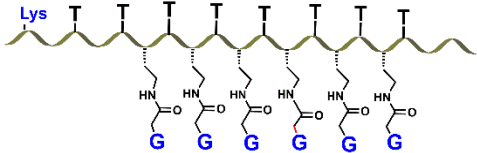
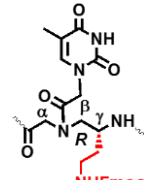
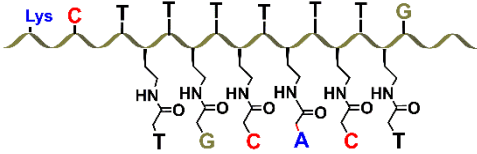
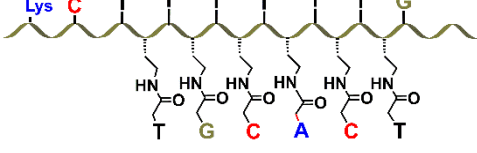
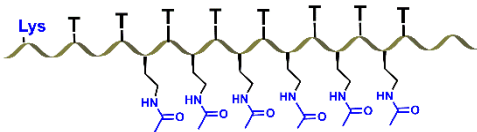
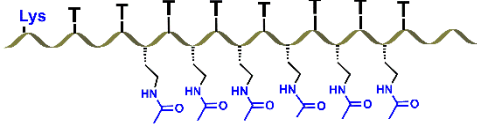
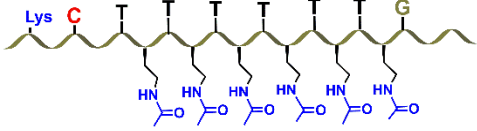
Figure 5.14 Synthetic strategy for $T_{8jp-\gamma S-NHAc}$ and $T_{8jp-\gamma R-NHAc}$ oligomer (JP 7 to JP 8)

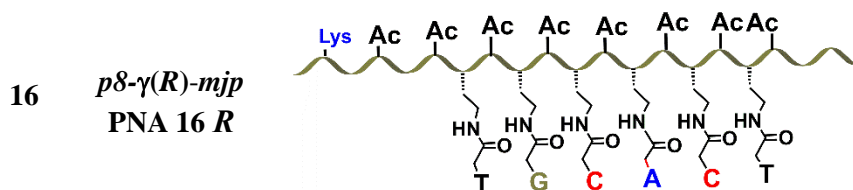
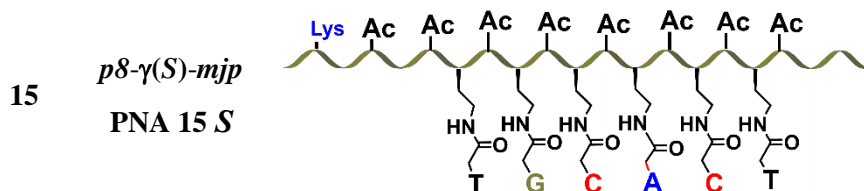
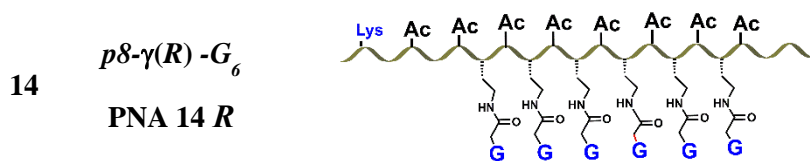
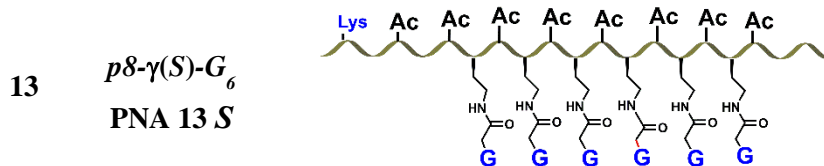
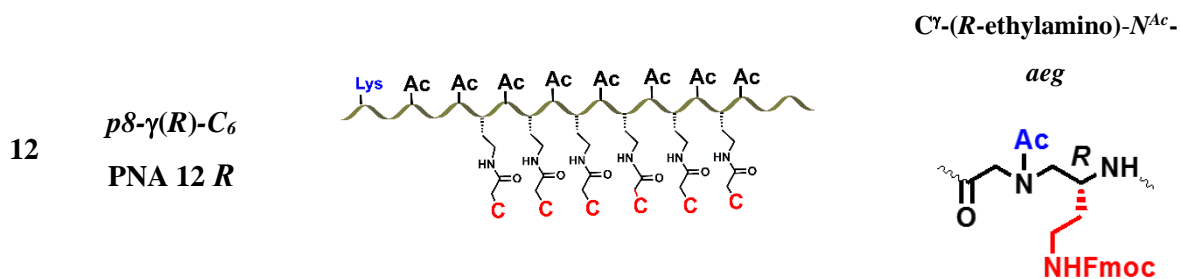
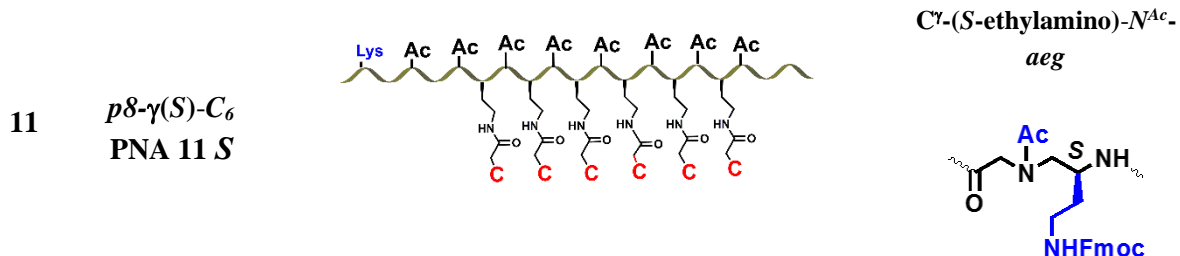
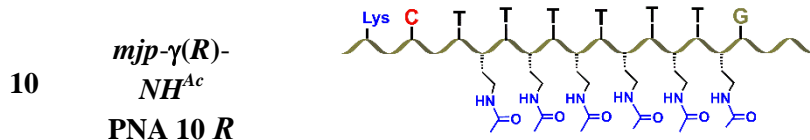
5.3.2e Summary of all PNA sequences.


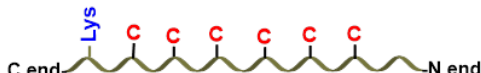

The sequences and structures of all PNA oligomers synthesized by solid phase in this chapter and the corresponding monomers used are summarized in Table 5.1.

Table 5.1 Janus-PNA oligomers with modified/unmodified monomers for base recognition

Entry	Sequence Code	PNA Sequences	Monomers used
1	$T_{8jp-\gamma(S)-C_6}$ JP 1S		C^γ -(S-ethylamino) aeg Thymine
2	$T_{8jp-\gamma(R)-C_6}$ JP 2R		C^γ -(R-ethylamino) aeg Thymine

3	$T_8 jp-\gamma(S)-G_6$ JP 3S		$C^\gamma-(S)\text{-ethylaamino) aeg}$ Thymine 
4	$T_8 jp-\gamma(R)-G_6$ JP 4R		$C^\gamma-(R)\text{-ethylaamino) aeg}$ Thymine 
5	$mjp-\gamma(S)-PNA$ JP 5S		
6	$mjp-\gamma(R)-PNA$ JP 6R		
7	$T_8 jp-\gamma(S)-NH^Ac$ PNA 7S		
8	$T_8 jp-\gamma(R)-NH^Ac$ PNA 8R		
9	$mjp-\gamma(S)-NH^Ac$ PNA 9S		



17	PNA-T ₈ PNA 17	
18	PNA-C ₆ PNA 18	
19	PNA-G ₆ PNA 19	

5.3.3 Cleavage of PNA oligomers from the solid support

The synthesized *Janus* PNA oligomers were cleaved from the solid support using trifluoromethane sulphonic acid (TFMSA) in the presence of trifluoroacetic acid (TFA). This yielded PNA oligomers having L-lysine amides at their C-termini.¹¹ In this cleavage condition, side chain protecting group was also 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.

5.3.4 Purification and characterization of the PNA oligomers

After the cleavage from the solid support, all PNA oligomers were purified by reverse phase high performance liquid chromatography (RP-HPLC) on semi-preparative C18 column using a gradient system of acetonitrile-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 the synthesized PNA oligomers. In literature, various matrices like sinapinic acid (3, 5-dimethoxy-4-hydroxycinnamic acid), 2, 5-dihydroxybenzoic acid (DHB), α -cyano-4-hydroxycinnamic acid (CHCA) etc. have been reported to record the MALDI-TOF spectra. Among these, DHB and CHCA were used here as a matrix to record MALDI-TOF spectra for all the synthesized PNAs. The calculated as well as observed molecular weights for all the PNAs with their molecular formulas, the MALDI-TOF data for confirmation of mixed purine-pyrimidine PNA oligomers are shown in Table 5.2.

Table 5.2 MALDI-TOF spectral analysis of the synthesized Janus PNA/ γ -aeg PNA/aeg PNA oligomers

Sr. No.	PNA sequenceCode	Molecular Formula	Calculated Mass	Observed Mass	Retention Time (min)
1	$T_8jp-\gamma(S)-C_6$ JP 1 S	$C_{142}H_{187}N_{59}O_{45}$	3440.42 [M+K] ⁺	3479.72	13.4
2	$T_8jp-\gamma(R)-C_6$ JP 2 R	$C_{142}H_{187}N_{59}O_{45}$	3440.42[M +K] ⁺	3479.52	13.8
3	$T_8jp-\gamma(S)-G_6$ JP 3 S	$C_{148}H_{187}N_{71}O_{45}$	3680.57 [M+ K] ⁺	3719.42	14.2
4	$T_8jp-\gamma(R)-G_6$ JP 4 R	$C_{148}H_{187}N_{71}O_{45}$	3680.57 [M+H] ⁺	3681.23	14
5	$mjp-\gamma(S)-PNA$ JP 5 S	$C_{145}H_{187}N_{65}O_{44}$	3542.43 [M] ⁺	3542.50	13.3
6	$mjp-\gamma(R)-PNA$ JP 6 R	$C_{145}H_{187}N_{65}O_{44}$	3542.43 [M +H] ⁺	3543.07	13.5
7	$T_8jp-\gamma(S)-NH^{Ac}$ PNA 7 S	$C_{118}H_{169}N_{41}O_{39}$	2784.25 [M+K] ⁺	2808.31	14.6
8	$T_8jp-\gamma(R)-NH^{Ac}$ PNA 8 R	$C_{118}H_{169}N_{41}KO_{39}$	2784.25 [M + K] ⁺	2808.99	15.5
9	$mjp-\gamma(S)-NH^{Ac}$ PNA 9 S	$C_{117}H_{167}N_{45}O_{37}$	2795.90 [M+Na]	2834.07	13.2
10	$mjp-\gamma(R)-NH^{Ac}$ PNA 10 R	$C_{117}H_{167}N_{45}O_{37}$	2795.90 [M+H]	2796.83	13.8
11	$p8-\gamma(S)-C_6$ PNA 11 S	$C_{102}H_{155}N_{43}O_{29}$	2446.19 [M+] ⁺	2446.79	13.6
12	$p8-\gamma(R)-C_6$ PNA 12 R	$C_{102}H_{155}N_{43}O_{29}$	2446.19 [M+H] ⁺	2446.74	13.8
13	$p8-\gamma(S)-G_6$ PNA 13 S	$C_{107}H_{153}N_{55}O_{29}$	2673.75 [M+K] ⁺	2712.37	12.1
14	$p8-\gamma(R)-G_6$ PNA 14R	$C_{107}H_{153}N_{55}O_{29}$	2673.75 [M+Na] ⁺	2696.71	12.2
15	$p8-\gamma(S)-mjp$ PNA 15 S	$C_{106}H_{157}N_{45}O_{30}$	2541.70 [M+K] ⁺	2580.69	13.1
16	$p8-\gamma(R)-mjp$ PNA 16 R	$C_{106}H_{157}N_{45}O_{30}$	2541.70 [M+K] ⁺	2580.66	13.5
17	PNA- T_8 PNA 17	$C_{94}H_{127}N_{35}O_{33}$	2273.93 [M+Na] ⁺	2296.39	13.9
18	PNA- C_6 PNA 18	$C_{66}H_{93}N_{33}O_{19}$	1651.73 [M+H] ⁺	1615.15	11.6
19	PNA- G_6 PNA 19	$C_{72}H_{93}N_{45}O_{19}$	1891.76 [M + H] ⁺	1892.13	11.2

5.4. Summary

The rationally designed modified PNA and *Janus* PNA monomers have been incorporated into 8-mer of *homo* and *mix Janus* PNA sequences. Modified monomers (**1** to **7**) incorporated at solid phase by using HOBt, HBTU and DIPEA as coupling reagent by solid phase peptide synthesis protocol. All the *Janus* PNA and unmodified *aeg* PNA oligomers obtained by solid phase synthesis were cleaved from solid support using appropriate protocol. The *Janus* PNA, γ -*aeg* PNA and *aeg* PNA oligomers after cleavage were purified by RP-HPLC and characterized by MALDI-TOF spectrometry.¹² Detailed experimental procedures (sections 5.5.1, 5.5.2 and 5.5.3) and spectral data of all intermediates are discussed in sections 4.5.5 and 4.5.6). The next chapter (Chapter 6) deals with the investigation of biophysical properties of PNA oligomers.

5.5 Experimental Methods

General: The MBHA resin (0.6 mmol/g) used for the Solid Phase Peptide Synthesis were bought from Novabiochem and were used without further purification. The loading value of resin reduced to 0.22 mmol/g. The Boc and Fmoc-protected amino acids were bought from ASM Technology Germany. The chemicals were used of laboratory or analytical grade obtained from Sigma Aldrich, Spectrochem and TCI. The solvents used for HPLC were obtained from Rankem. PNA oligomers were purified on Agilent HPLC system using semi-preparative BEH130 C18 (10X250 mm) column. Mass spectra for all PNAs were obtained by Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using 2, 5-dihydroxybenzoic acid (DHB) or α -Cyano-4-hydroxycinnamic acid (CHCA) as matrix.

5.5.1 Synthesis of Janus PNA oligomers on solid support.

The *Janus*-PNA oligomers were synthesized on solid phase, by using Boc and Fmoc strategy^{13,14}. The resin-bound amine PNA oligomers (10 mg, 0.20 mmol/g) in DMF were reacted with corresponding nucleobase A/T/G/C (12 mg, 3 eq) in the presence of HBTU (6 mg, 3 eq), HOBt (6 mg, 3 eq), and DIPEA (10 μ L, 6 eq). The reaction was done for 5 min in microwave at 65 °C and 25 W and then 6h at room temperature. Excess reagents were removed by filtration and the resin was washed with DMF, DCM, MeOH.

5.5.2 Cleavage of the Janus PNA oligomers from solid support

The MBHA resin (10 mg) after assembly of *Janus* PNA oligomers were treated with 40% methylamine and kept for another 8 h at room temperature for removal of isobutyryl group from guanine. After that the resin was treated with thioanisole (20 μ L) and 1,2-ethanedithiol (8 μ L) in an ice bath for 10 min. TFA (200 μ L) was added and cooled in an ice bath. TFMSA (16 μ L) was added slowly with stirring and 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 and the filtrate was evaporated on a rotary evaporator at ambient temperature. The filtrate 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 milli-Q water. Then, the peptides were filtered and purified by HPLC.

5.5.3 Purification of the Janus PNA oligomers by RP-HPLC

The purification of PNAs was carried out on Agilent HPLC system with semi-preparative BEH130 C18 (10x250 mm) column using solvents water and acetonitrile with composition A: 0.1% TFA in CH₃CN:H₂O (5:95) and B= 0.1% TFA in CH₃CN:H₂O (1:1). The gradient for elution was 100% A to 100% B in 20 min, with flow rate of 2 mL/min. The HPLC elutions were monitored at 220 and 260 nm wavelength.

5.6 References

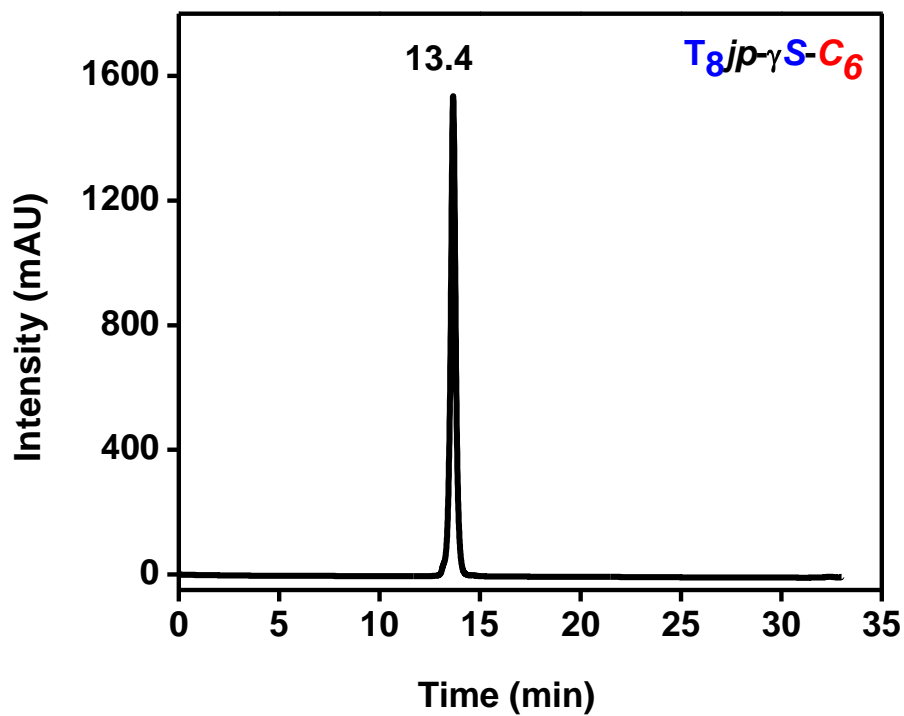
1. Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12021-12026.
2. Dragulescu Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc.* **2006**, *128*, 10258-10267.
3. He, G.; Rapireddy, S.; Bahal, R.; Sahu, B.; Ly, D. H. *J. Am. Chem. Soc.* **2009**, *131*, 12088–12090.
4. Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180-193.
5. Bayermann, M.; Bienert, M. *Tetrahedron Letters*. **1992**, *33*, 3745–3748
6. Nielsen, P. E.; Haaima, G.; Lohse, A.; Buchardt, O. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939-1942.
7. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595–598. 18. Sarin, V. K.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem.* **1981**, *117*, 147-157.
8. Coin, I.; Bayermann M.; Bienert, M. *Nat Protoc.* **2007**, *2*, 3247–3256.
9. Pipkorn, R.; Wiessler, M.; Waldeck, W.; Hennrich, U.; Nokihara, K.; Beining, M.; Braun, K. *Int. J. Med. Sci.* **2012**, *9*, 1-10.

10. 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. Pept. Sci.* **1995**, *3*, 175–183.
11. Kulkarni, S. S.; Wang, C. C.; Sabbavarapu, N. M.; Podilapu, A. R.; Liao, P.-H.; Hung, S. C. *Chem. Rev.* **2018**, *118*, 8025-8104.
12. Cui, Z.; Theruvathu, J. A.; Farrel, A.; Burdzy, A.; Sowers, L. C. *Anal. Biochem.* **2008**, *379*, 196–207.
13. Sheppard, R. J. *Pept. Sci.* **2003**, *9*, 545.
14. Behrendt, R.; White, P.; Offer, J. J. *Pept. Sci.* **2016**, *22*, 4.

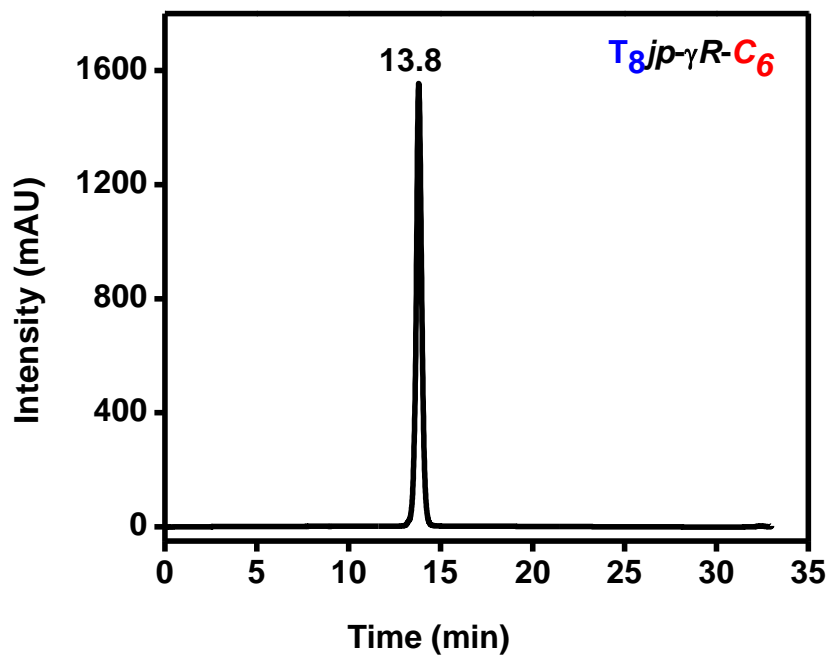
5.8 Appendix I

5.8.1 HPLC chromatograms of peptides

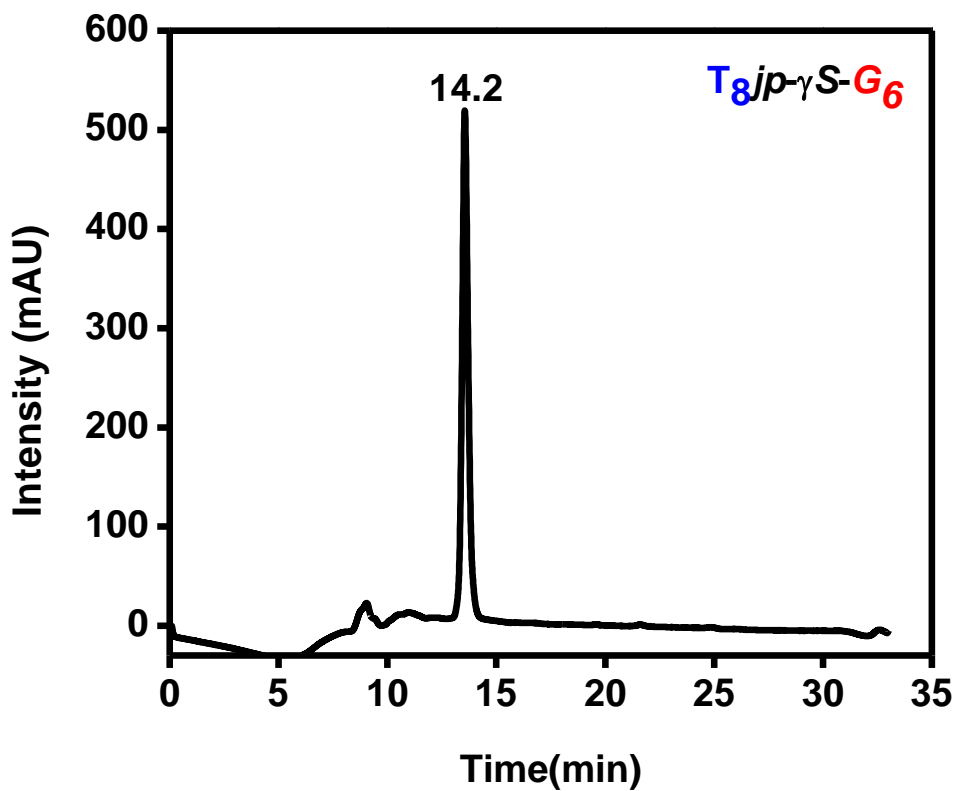
HPLC chromatogram of (JP 1)



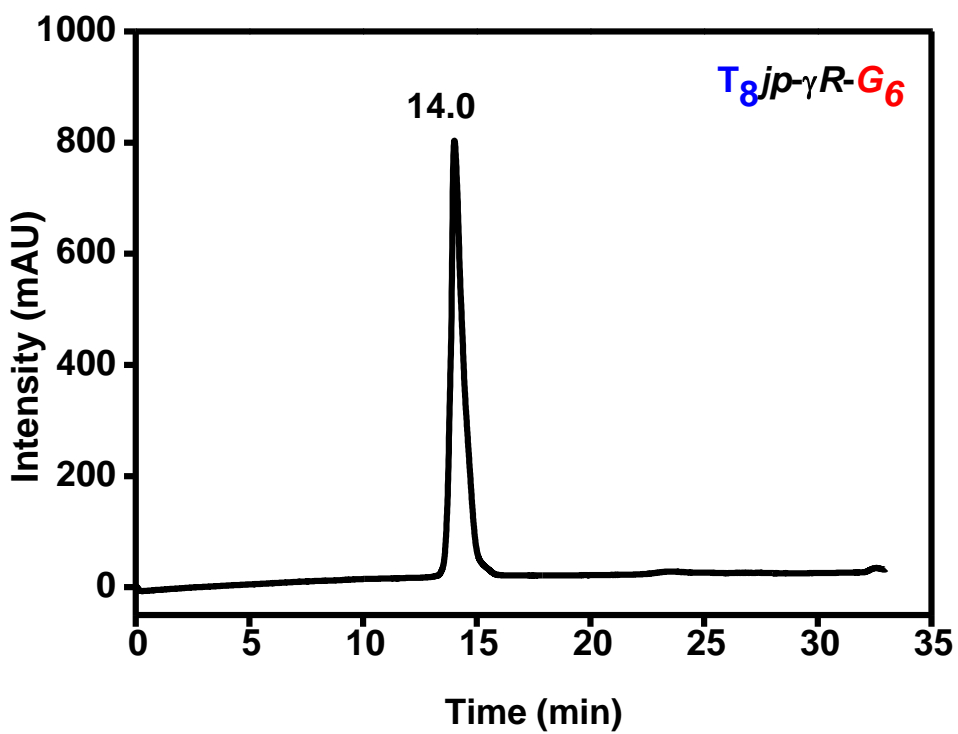
HPLC chromatogram of (JP 2)



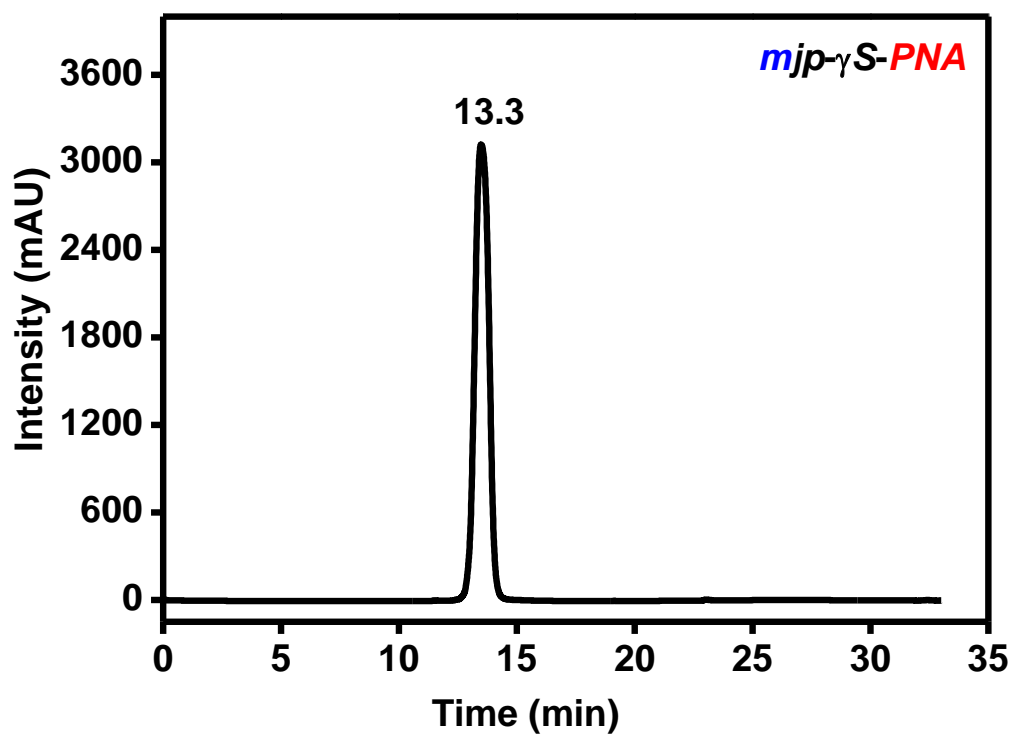
HPLC chromatogram of (JP 3)



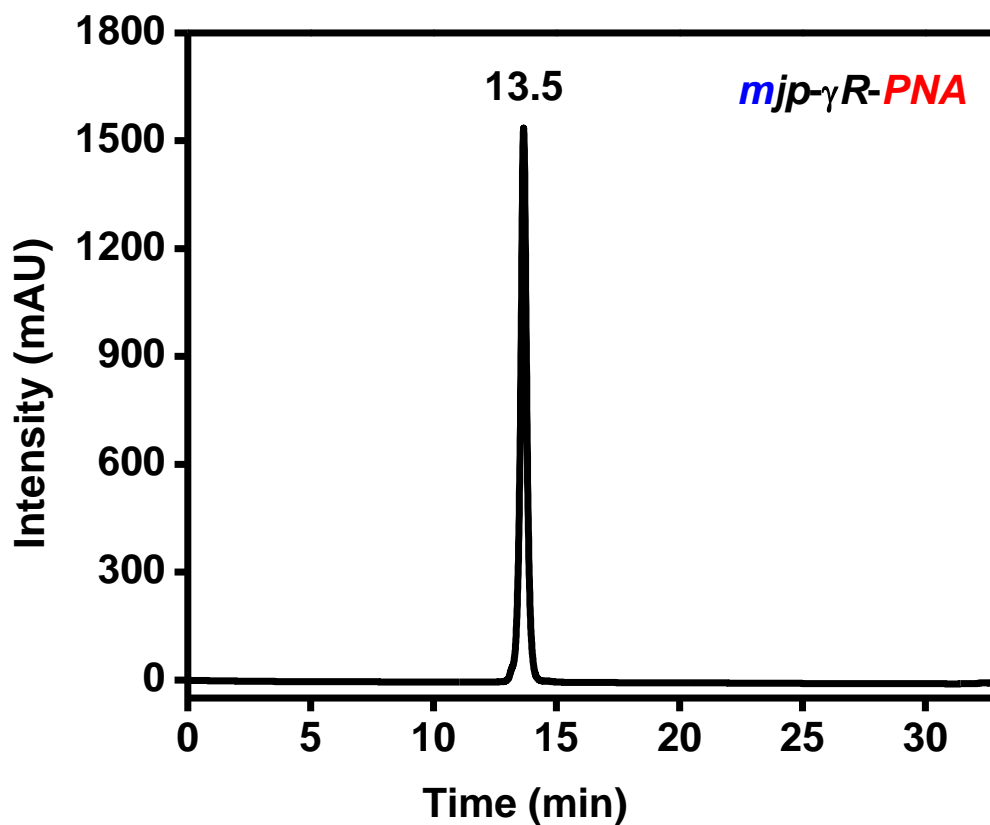
HPLC chromatogram of (JP 4)



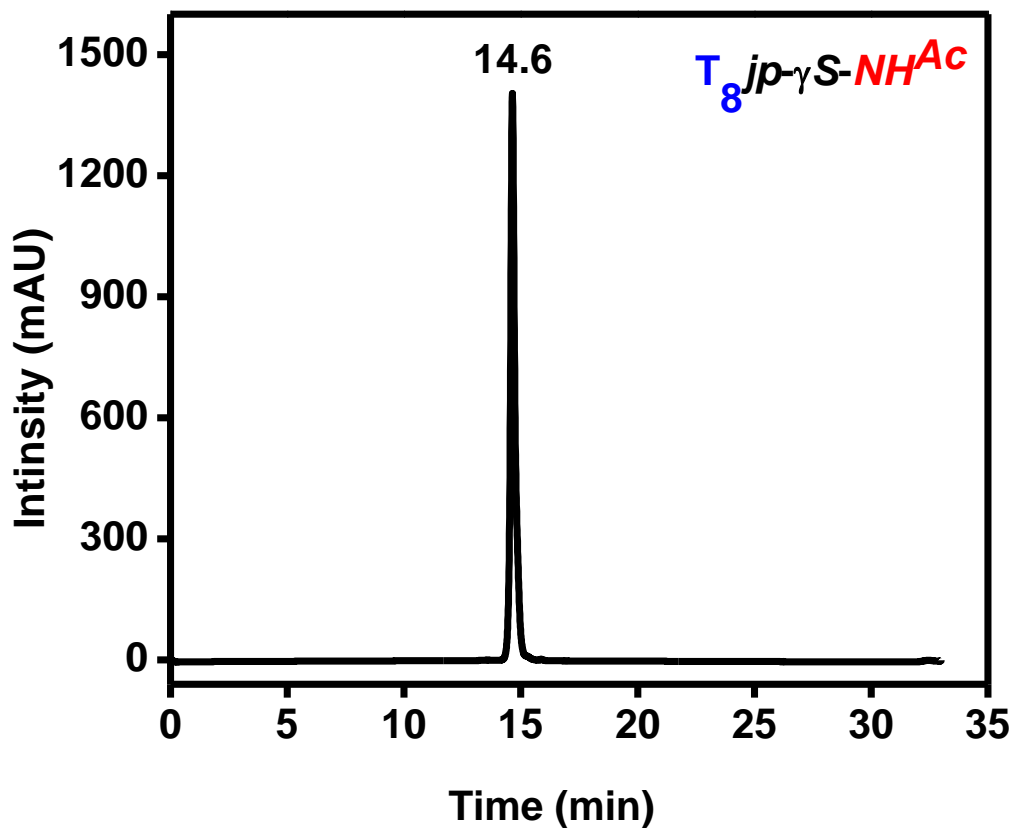
HPLC chromatogram of (JP 5)



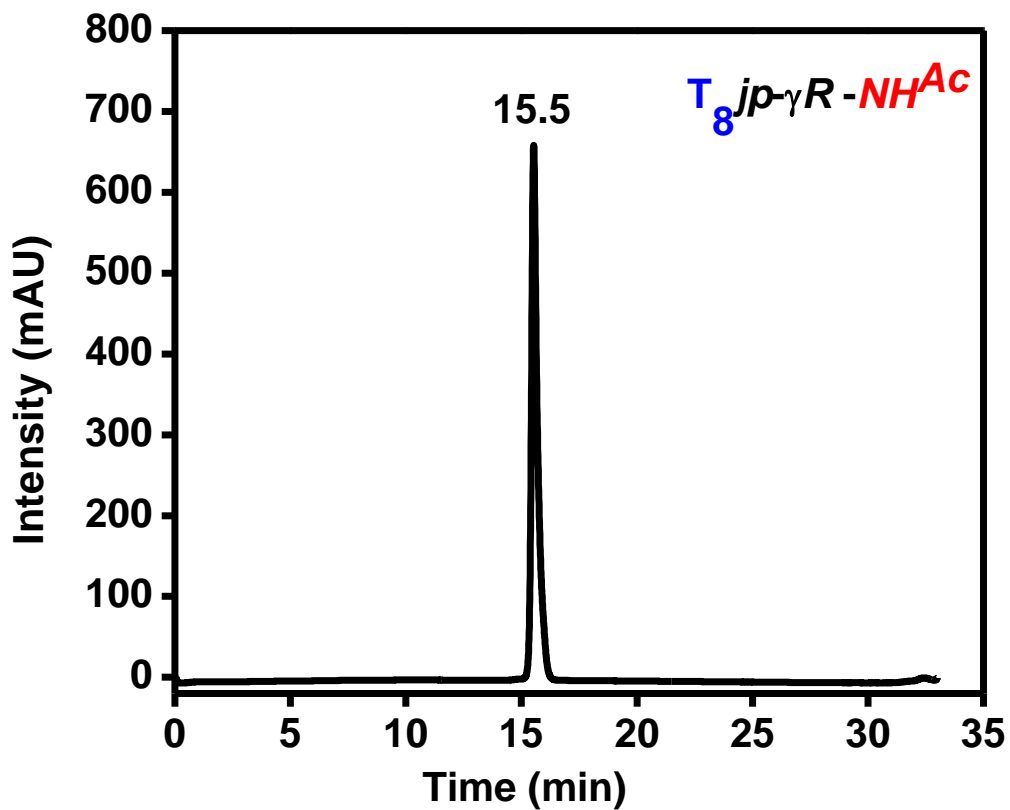
HPLC chromatogram of (JP 6)



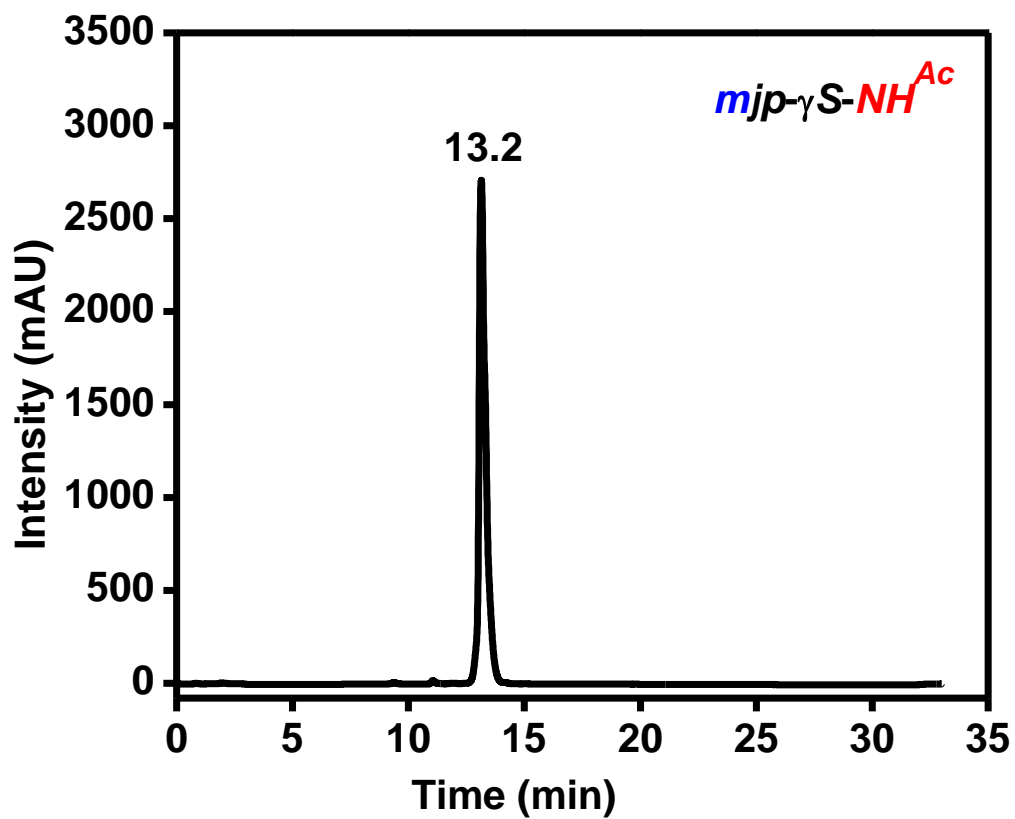
HPLC chromatogram of (JP 7)



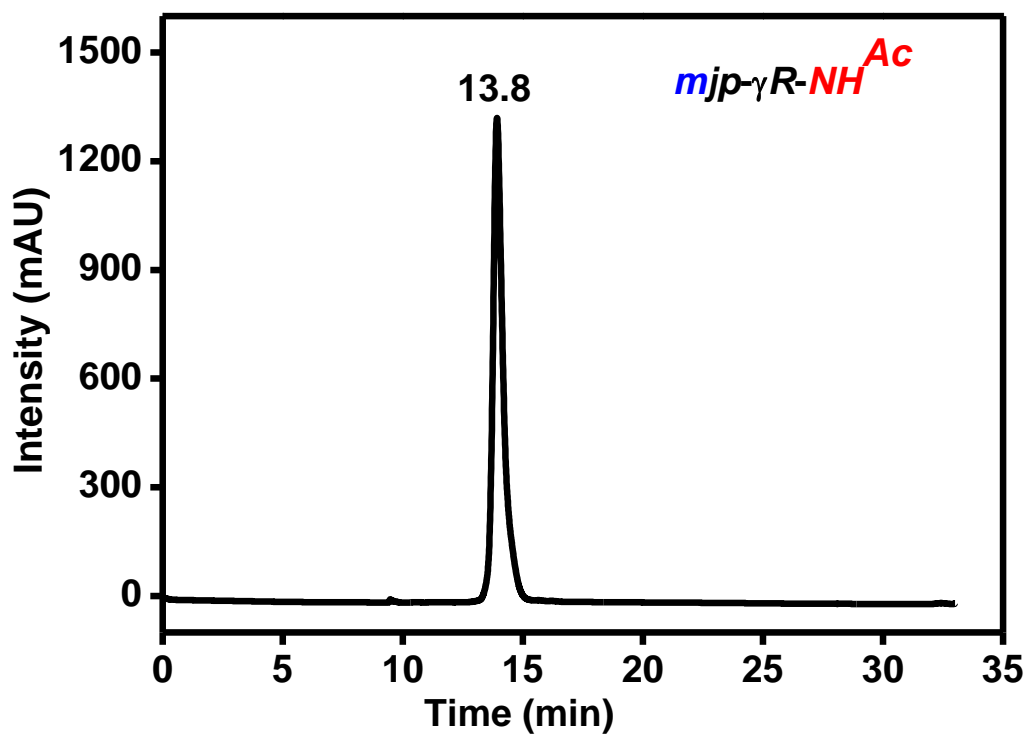
HPLC chromatogram of (JP 8)



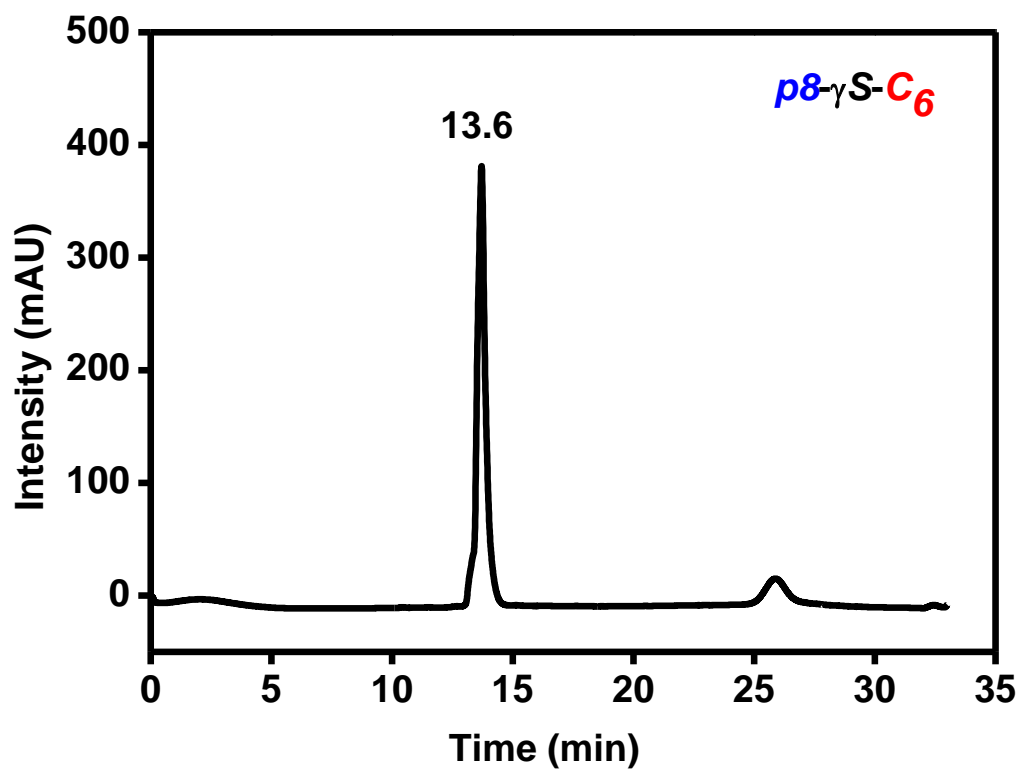
HPLC chromatogram of (JP 9)



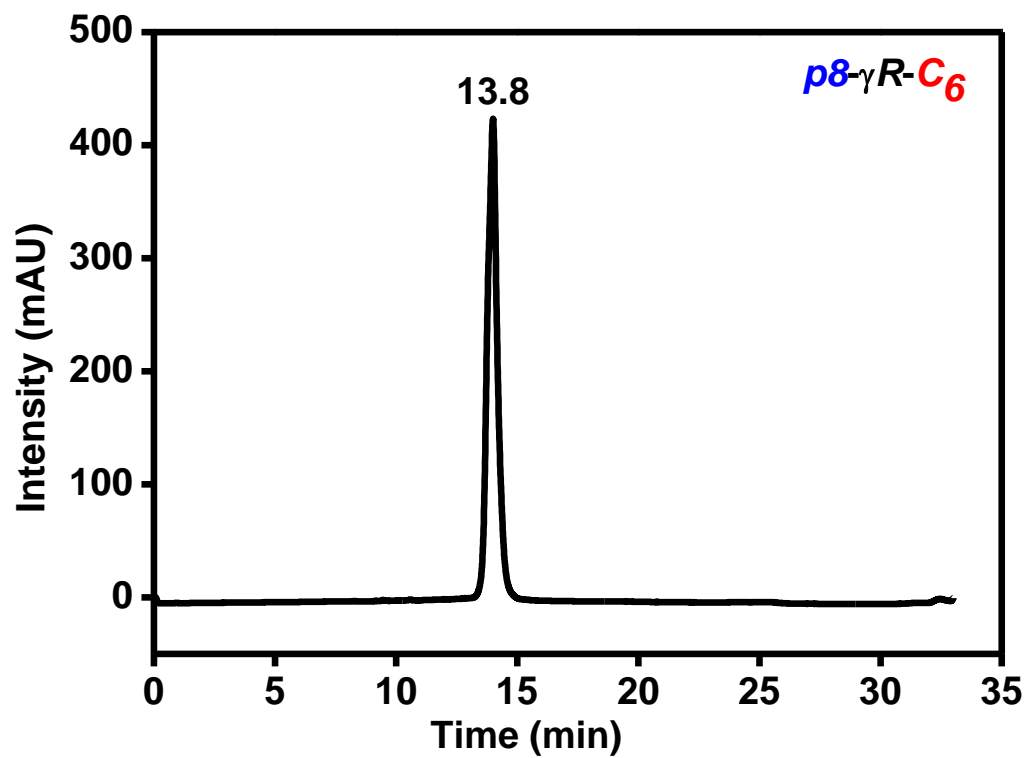
HPLC chromatogram of (JP 10)



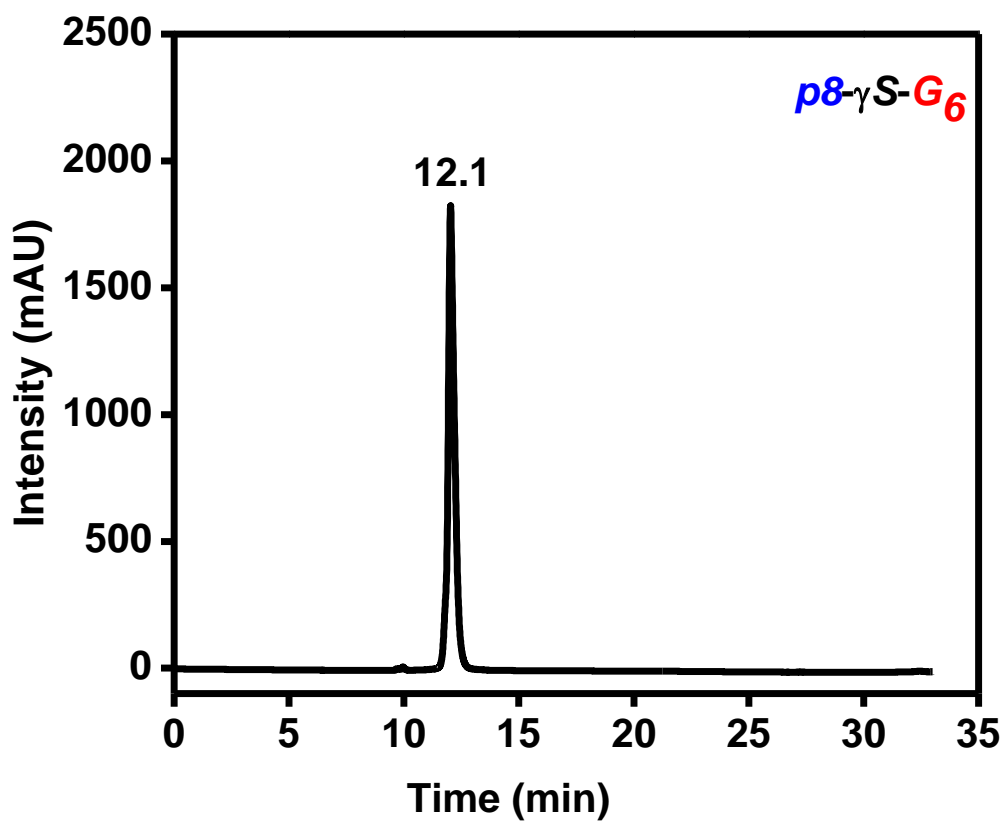
HPLC chromatogram of (PNA 11)



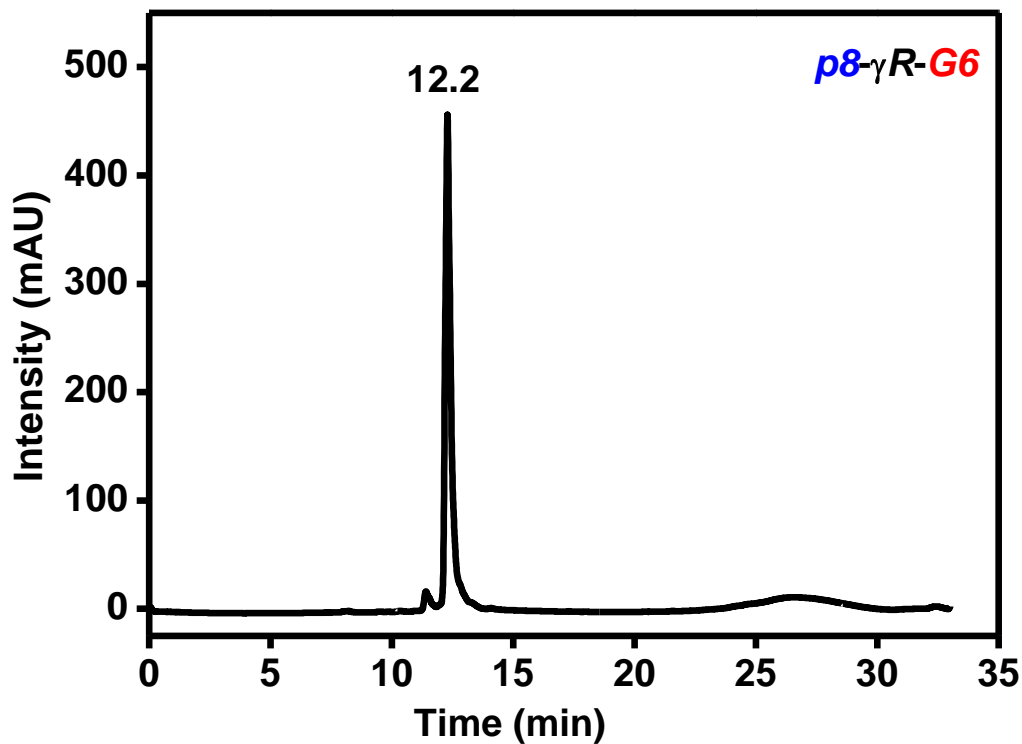
HPLC chromatogram of (PNA 12)



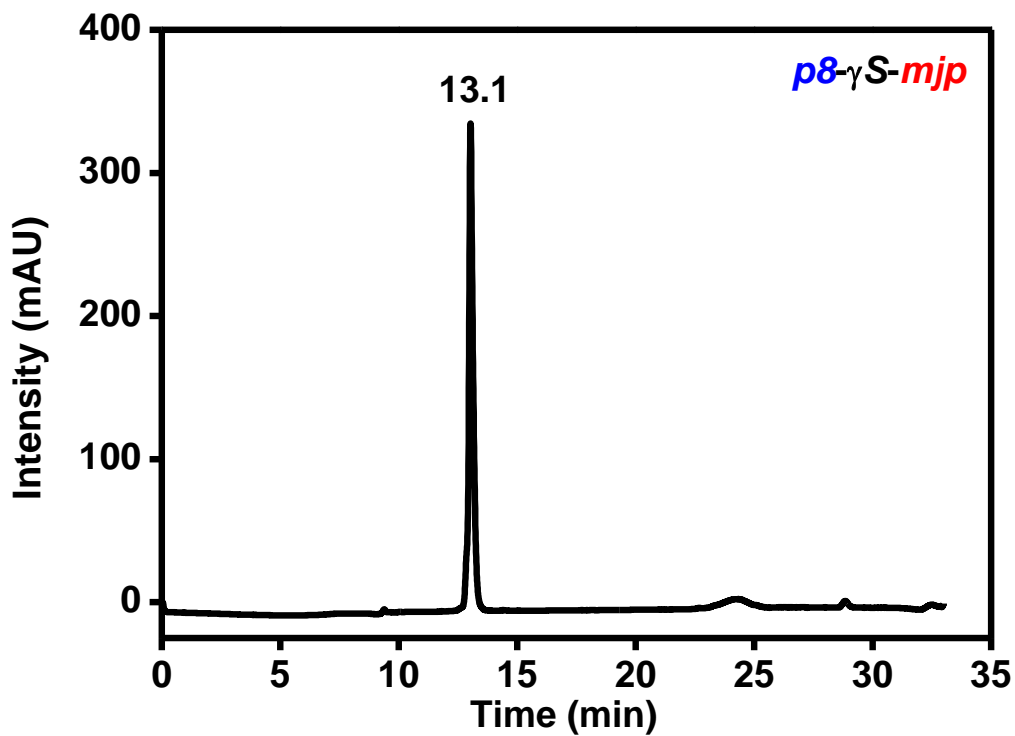
HPLC chromatogram of (PNA 13)



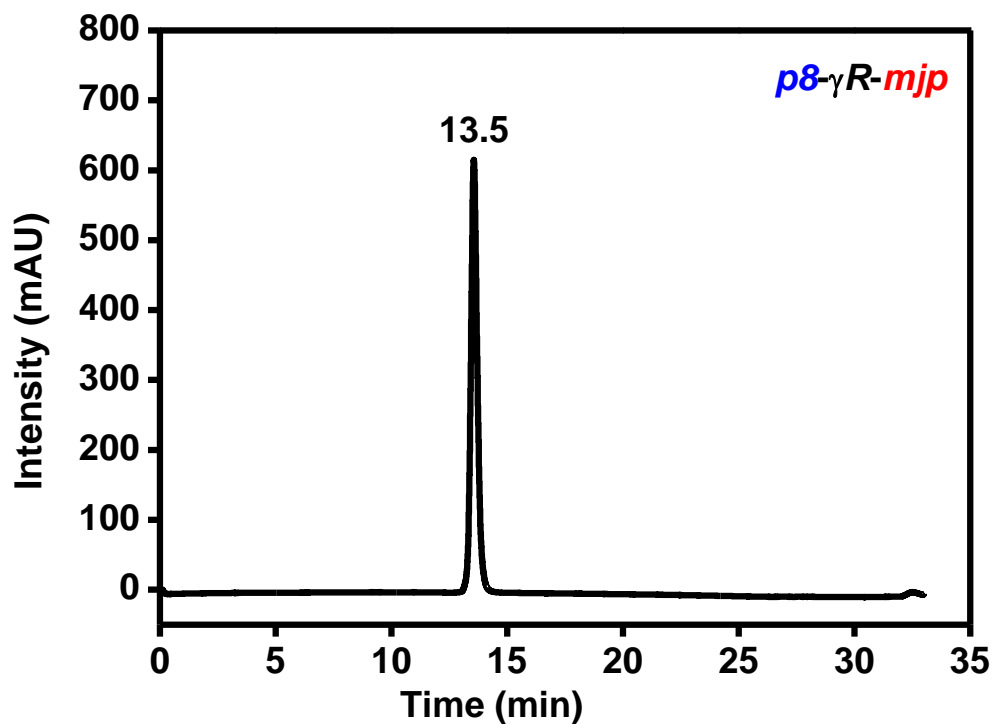
HPLC chromatogram of (PNA 14)



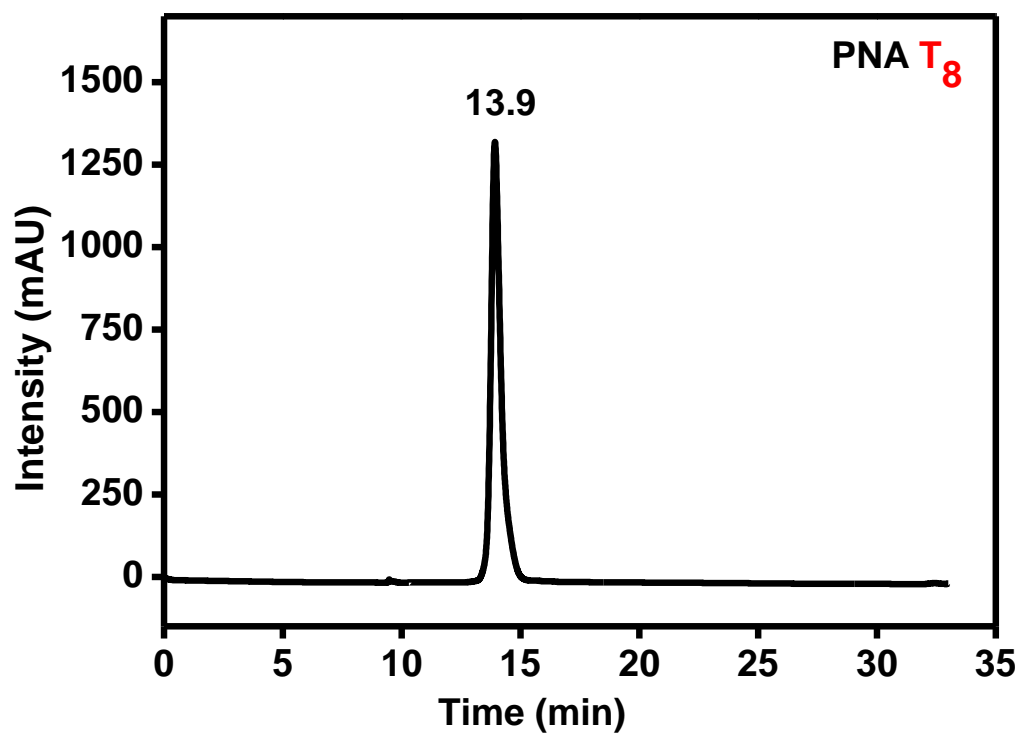
HPLC chromatogram of (PNA 15)



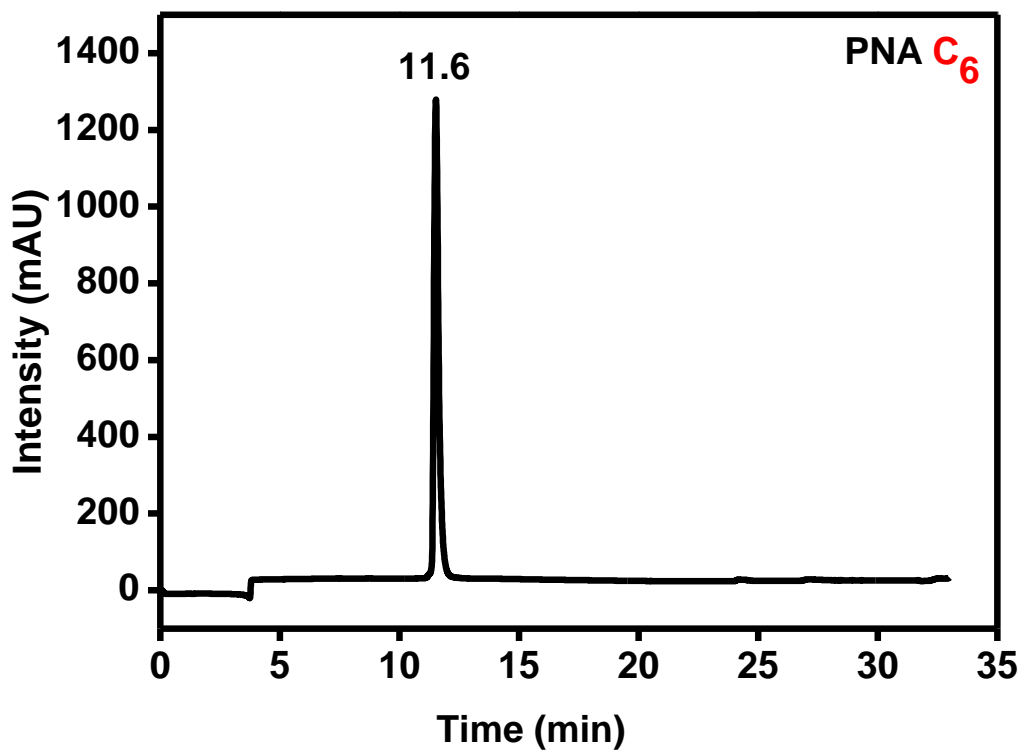
HPLC chromatogram of (PNA 16)



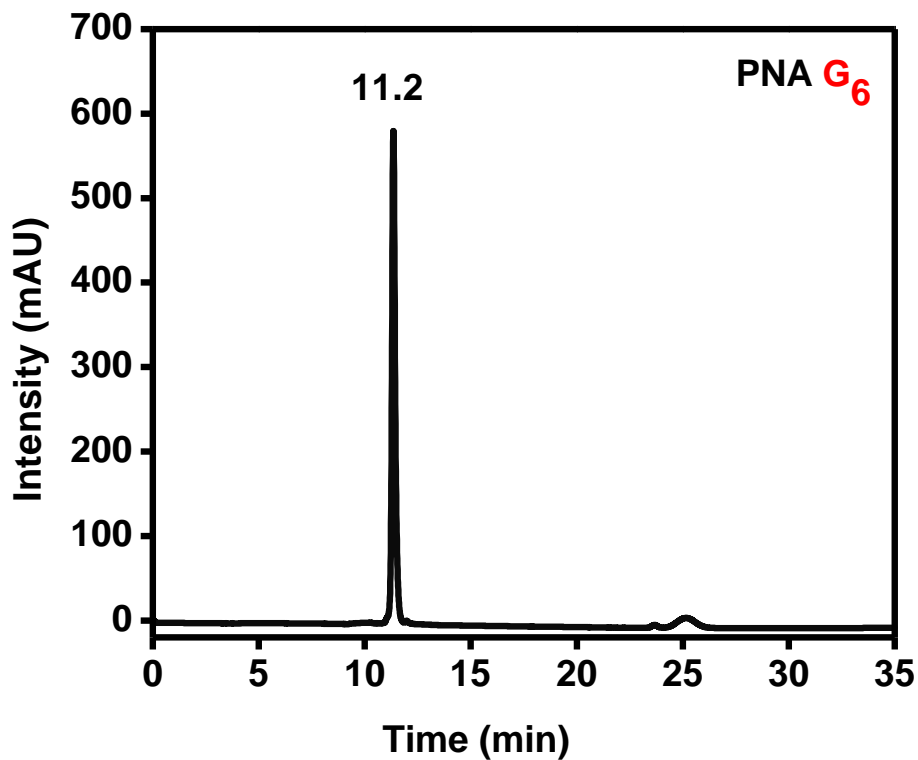
HPLC chromatogram of (PNA 17)



HPLC chromatogram of (PNA 18)

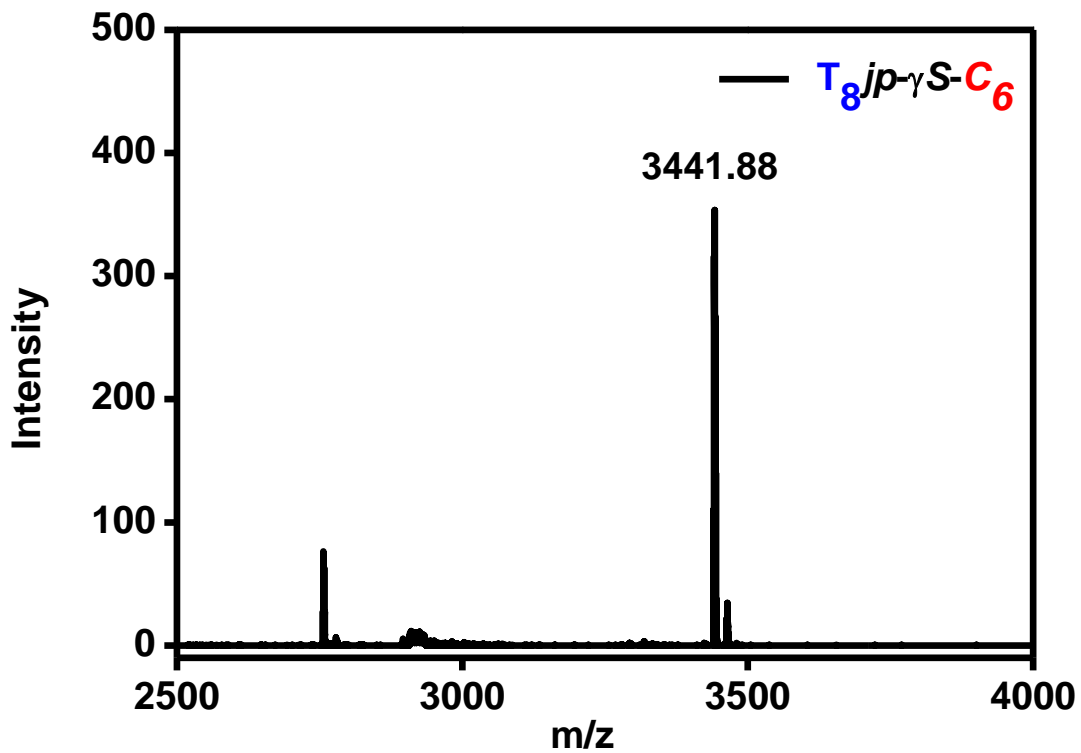


HPLC chromatogram of (PNA 19)

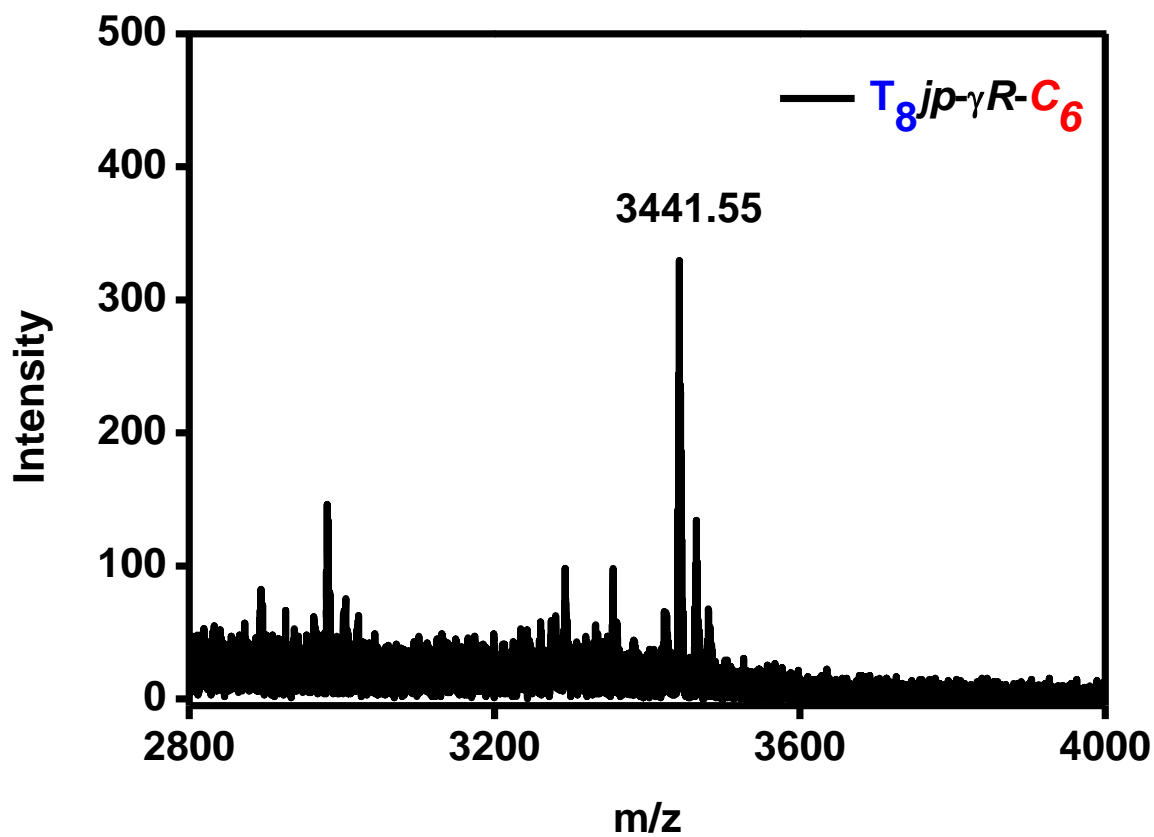


5.8.2 MALDI TOF spectra of peptides

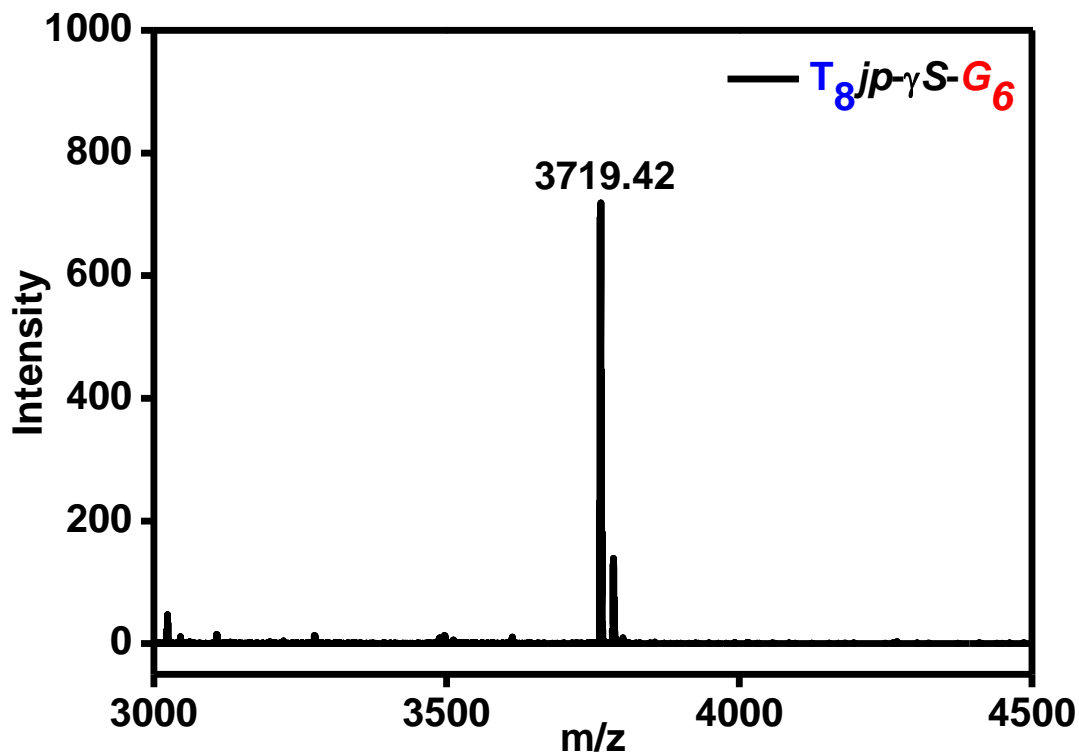
MALDI-TOF Spectrum of (JP 1)



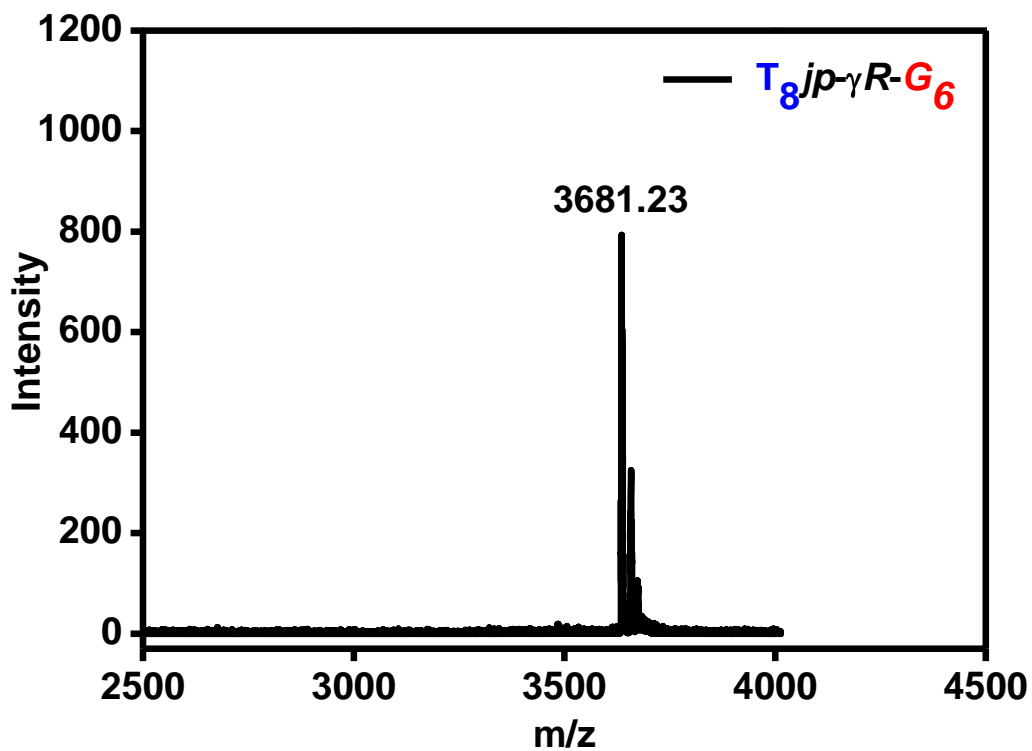
MALDI-TOF Spectrum of (JP 2)



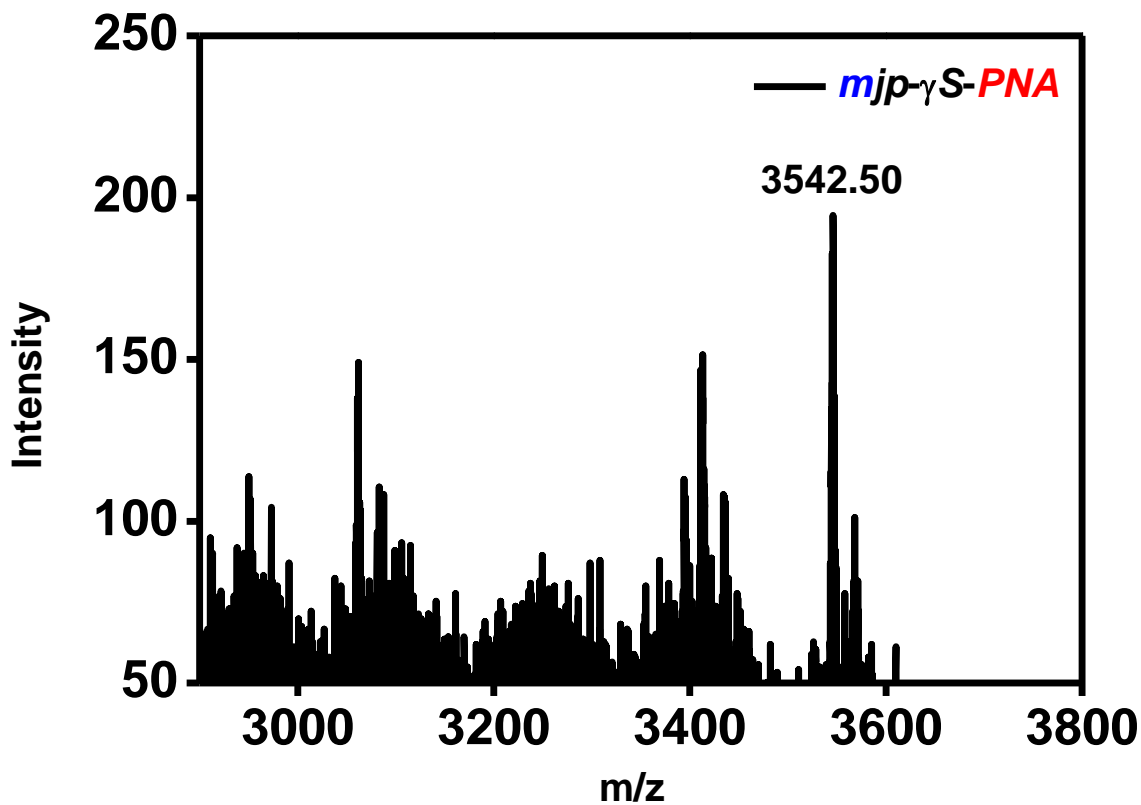
MALDI-TOF Spectrum of (JP 3)



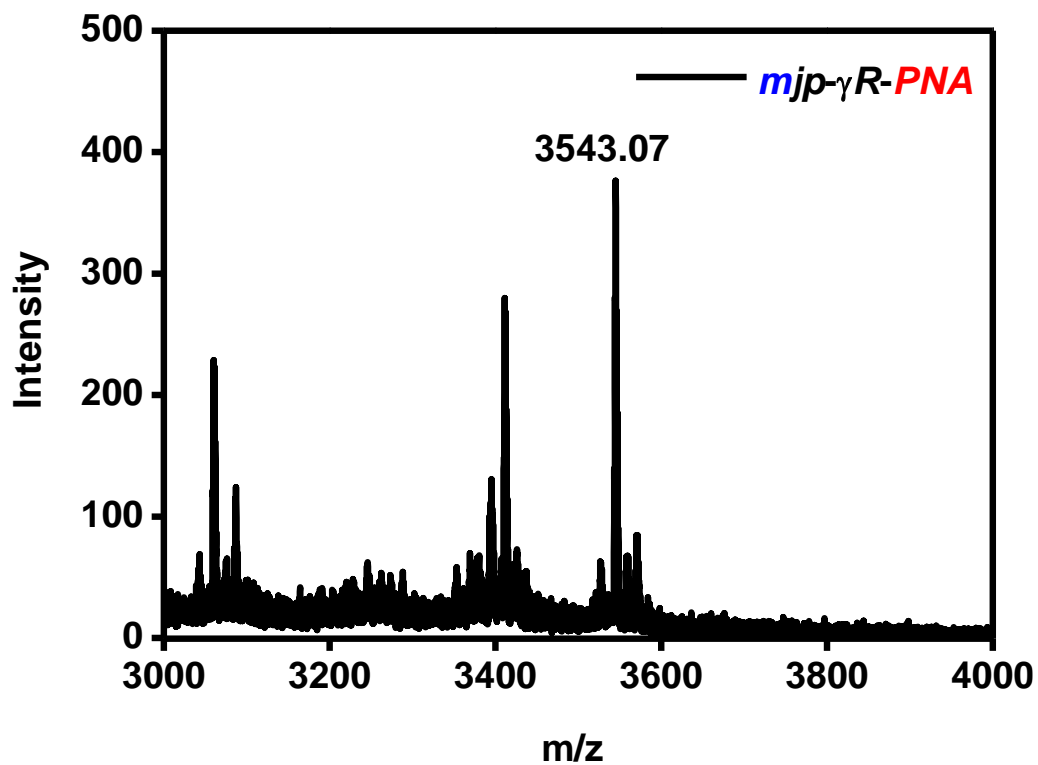
MALDI-TOF Spectrum of (JP 4)



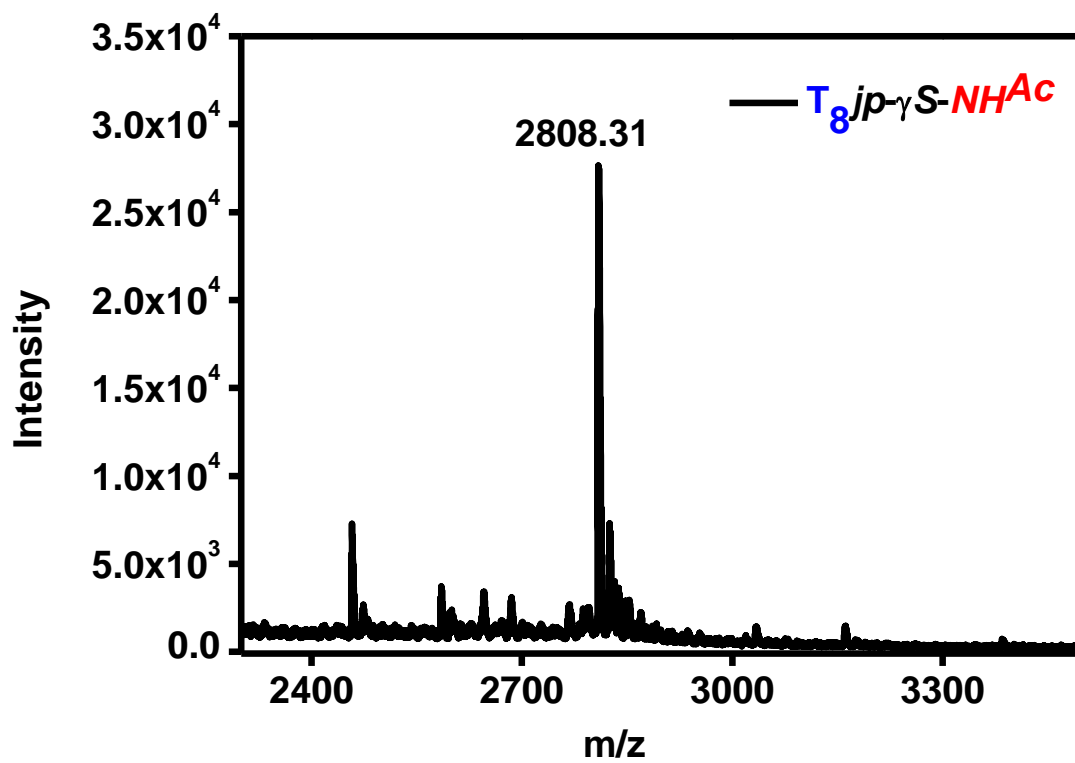
MALDI-OF Spectrum of (JP 5)



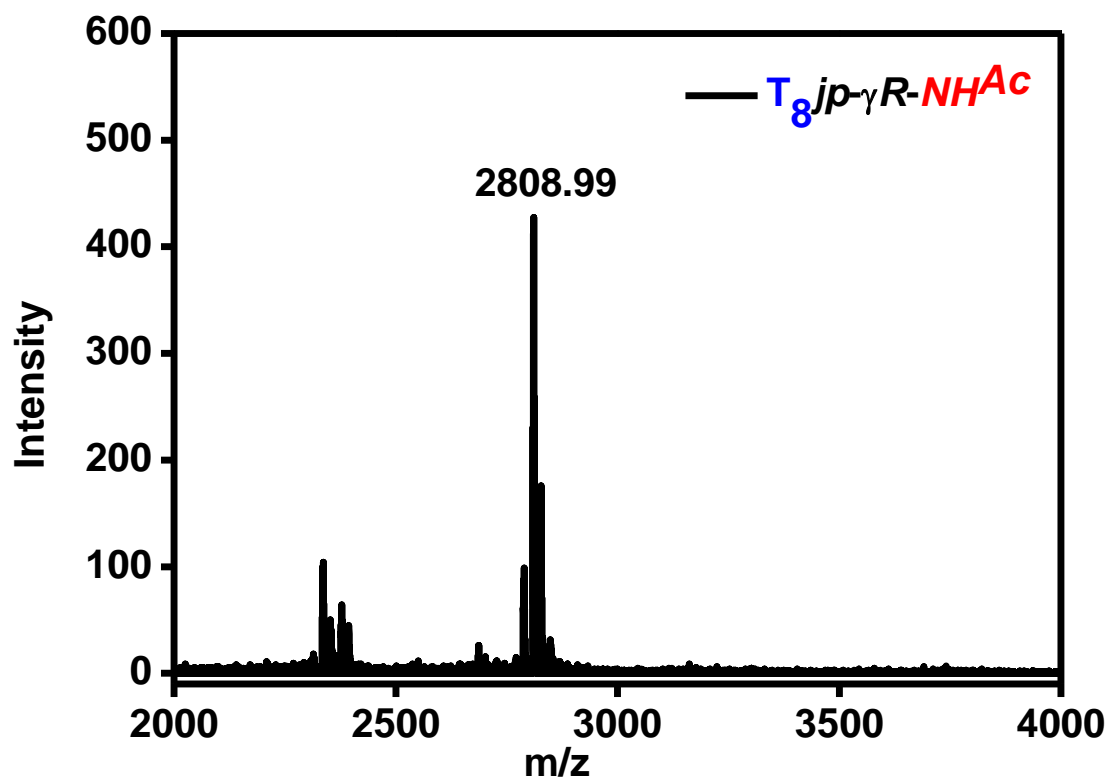
MALDI-TOF Spectrum of (JP 6)



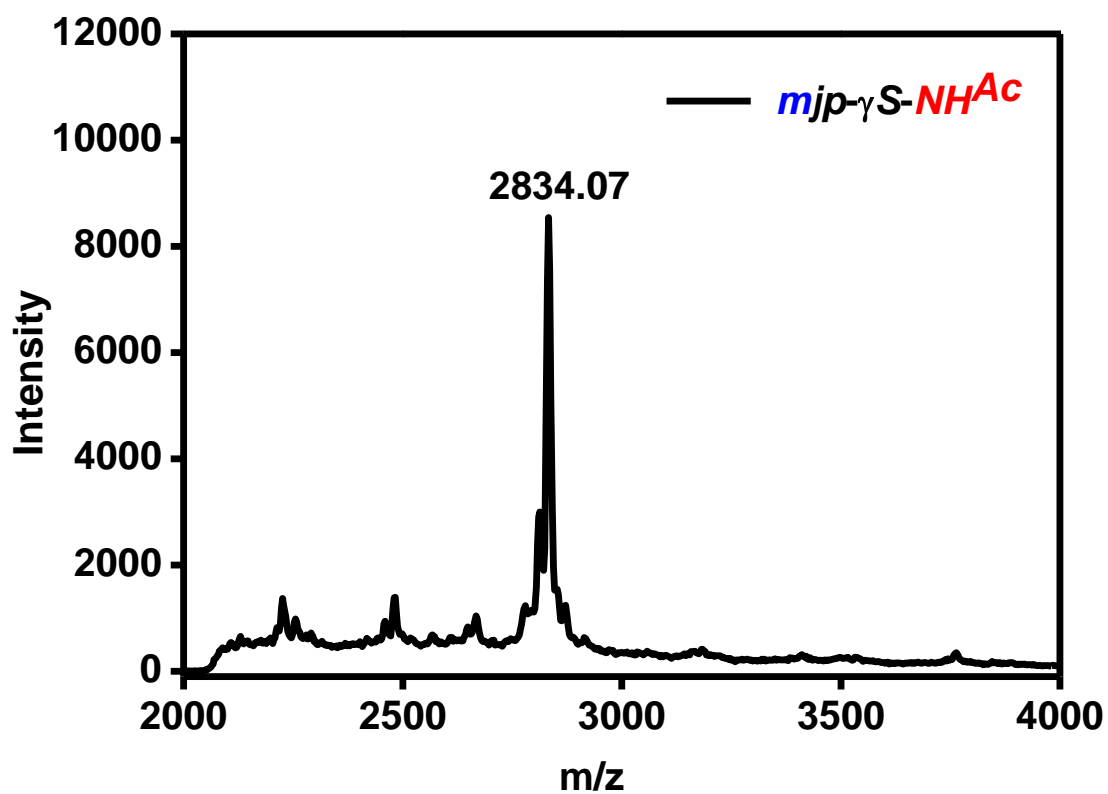
MALDI-TOF Spectrum of (JP 7)



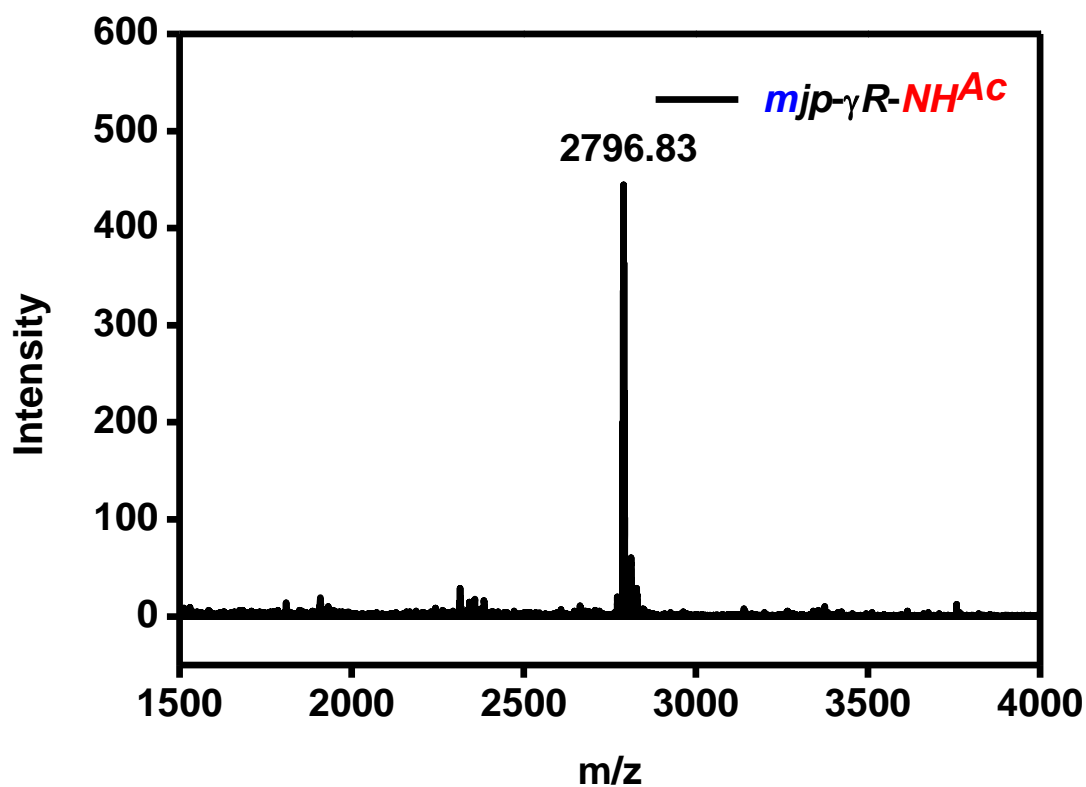
MALDI-TOF Spectrum of (PNA 8)



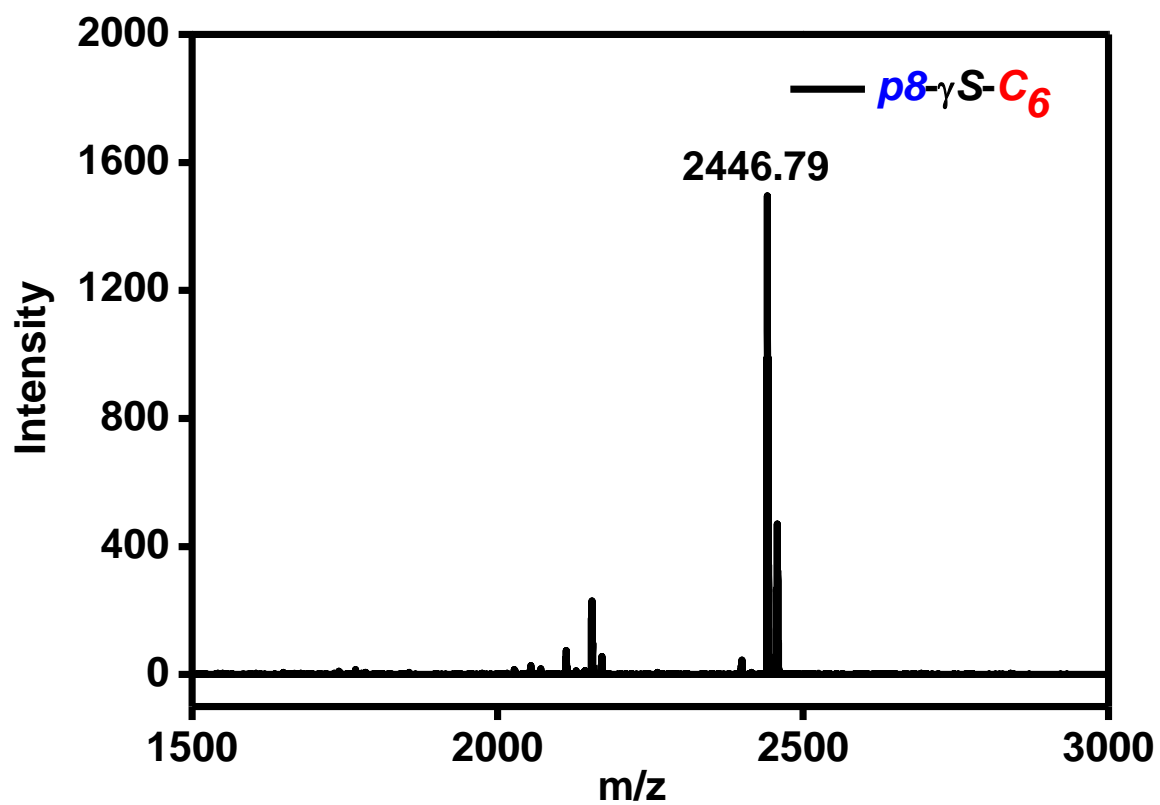
MALDI-TOF Spectrum of (PNA 9)



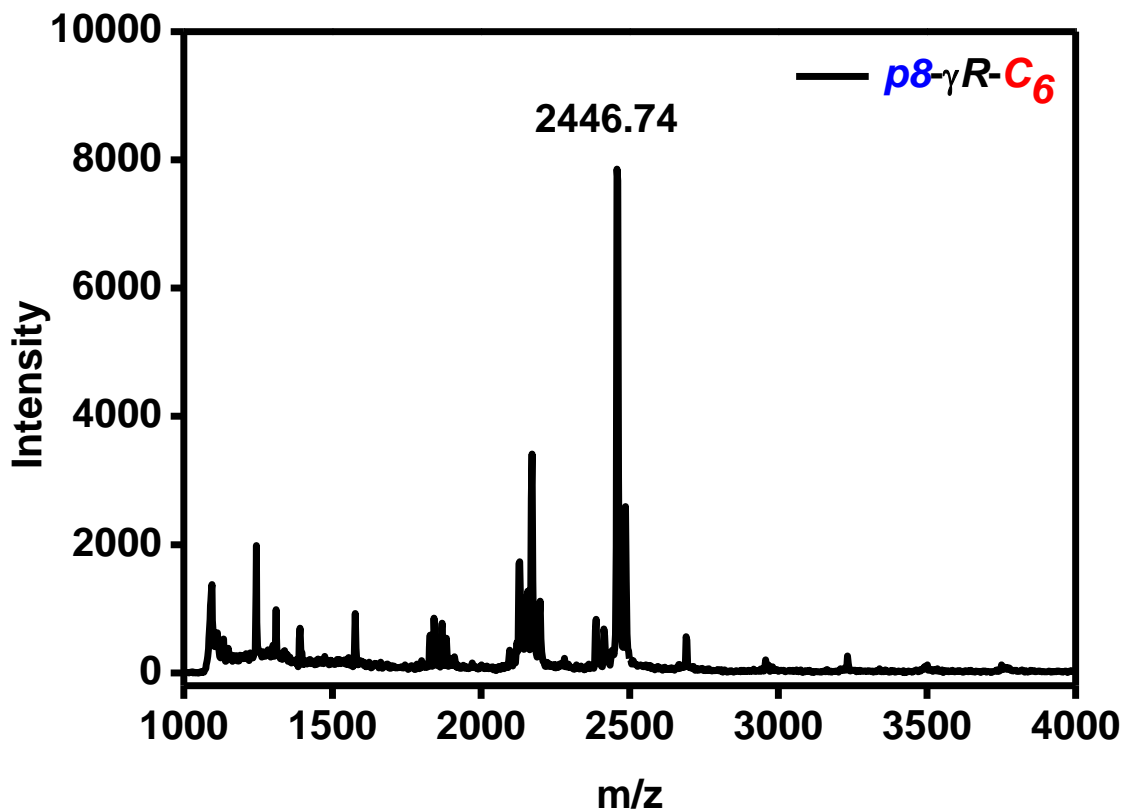
MALDI-TOF Spectrum of (JP 10)



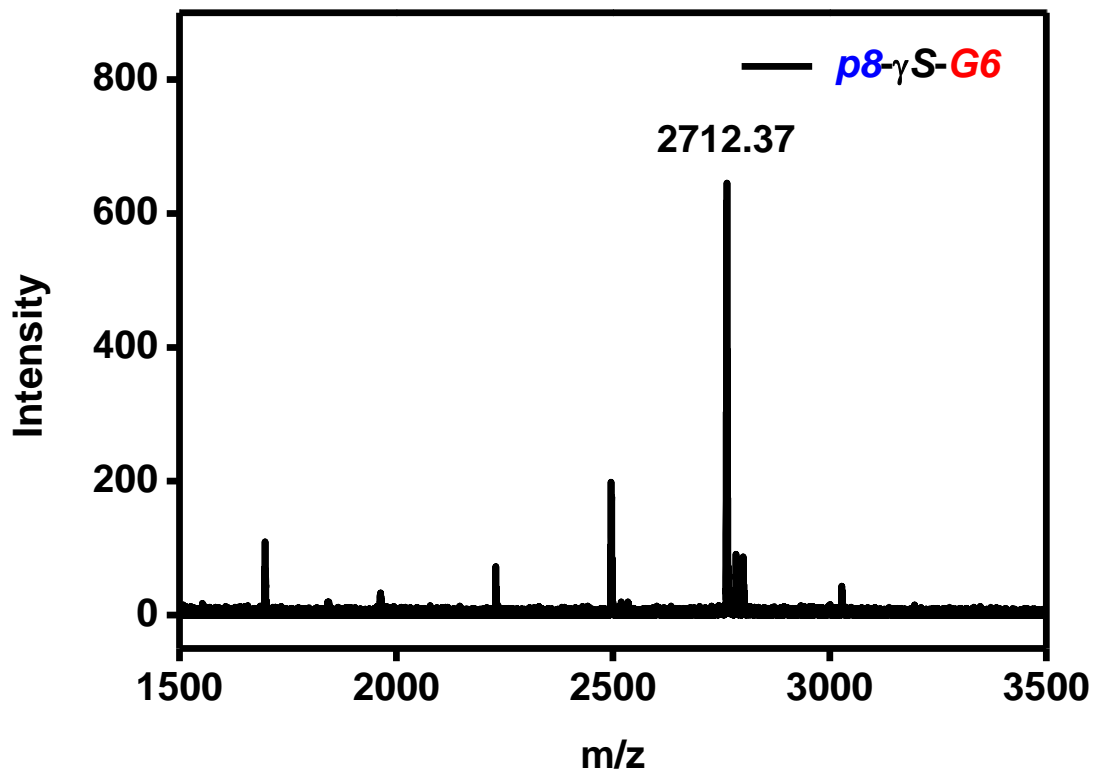
MALDI-TOF Spectrum of (JP 11)



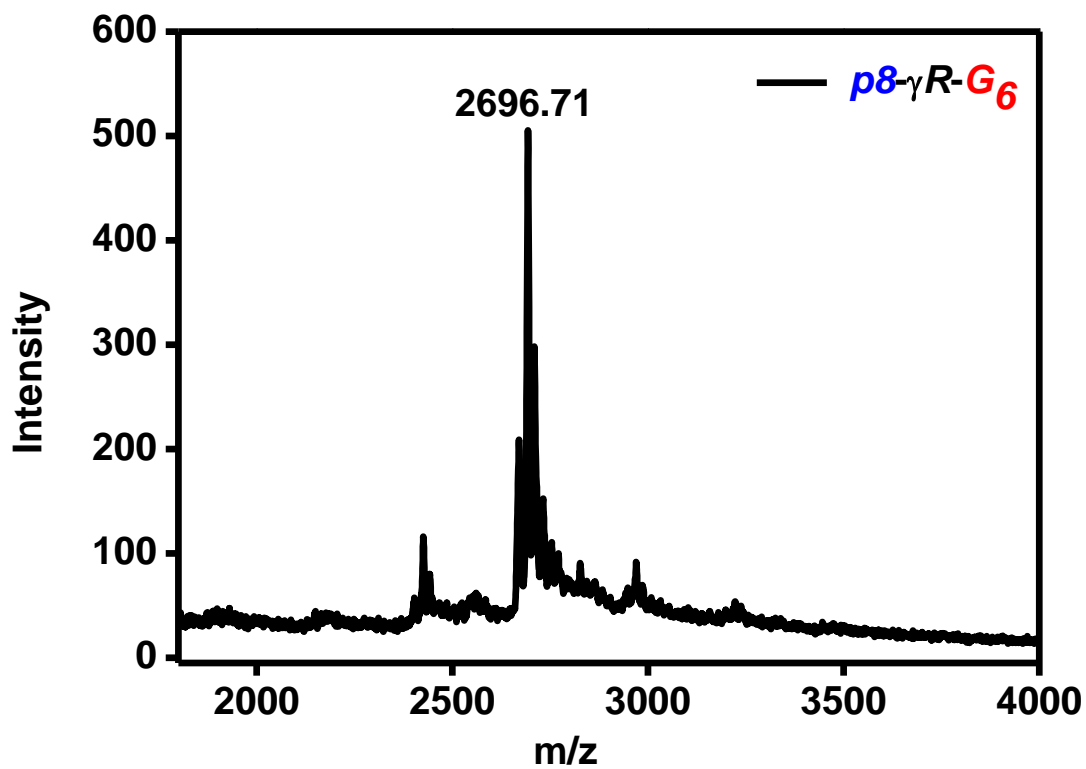
MALDI-TOF Spectrum of (JP 12)



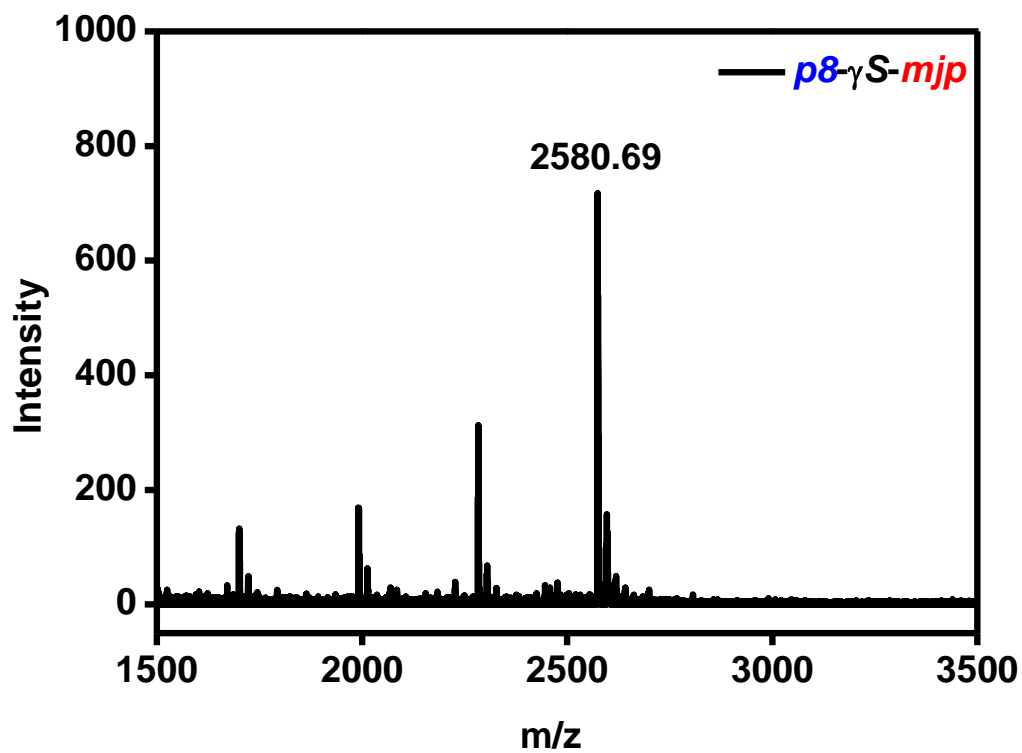
MALDI-TOF Spectrum of (PNA 13)



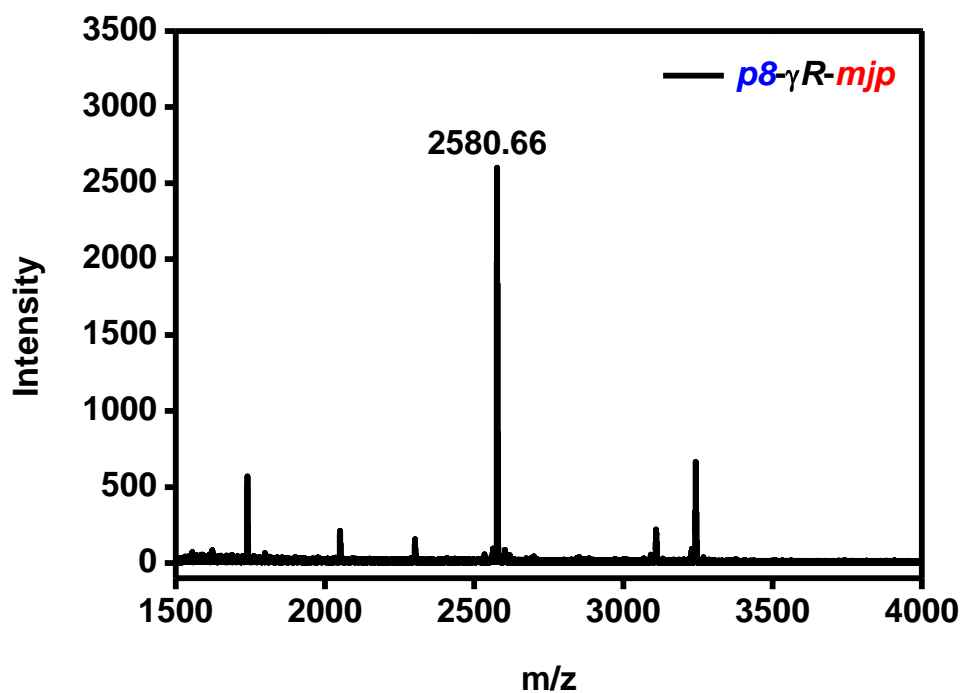
MALDI-TOF Spectrum of (PNA 14)



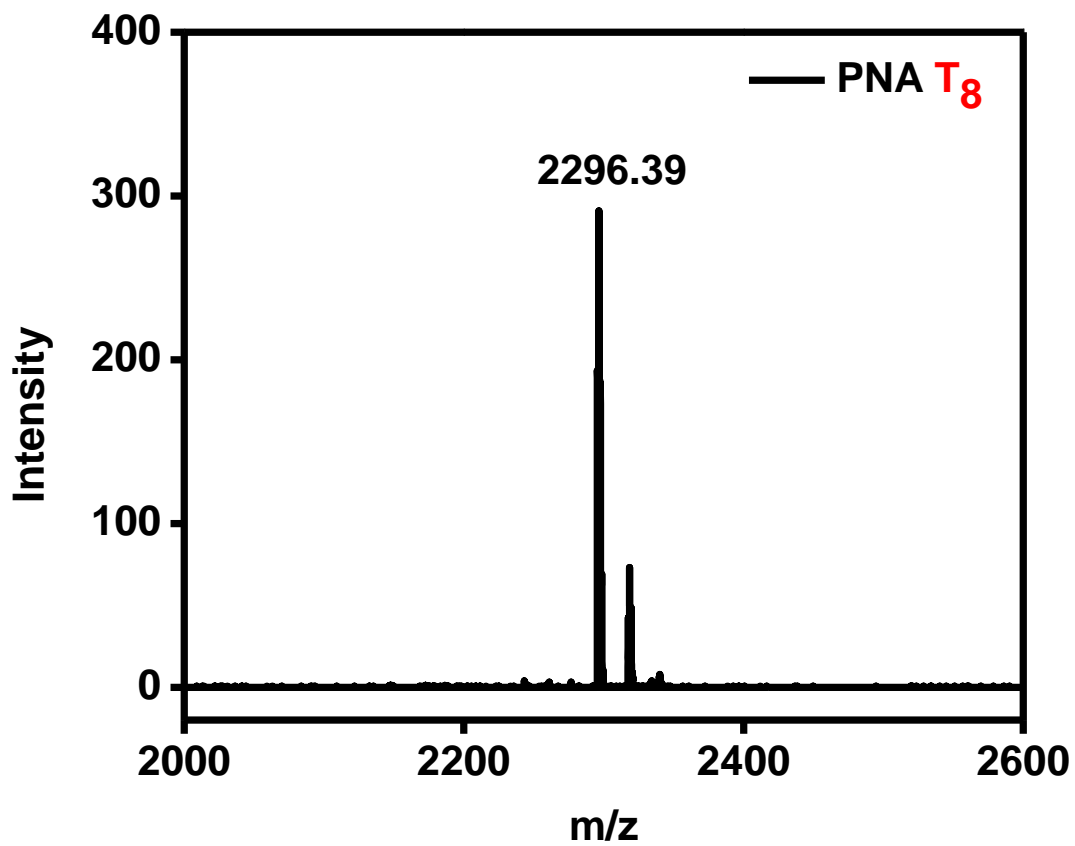
MALDI-TOF Spectrum of (PNA 15)



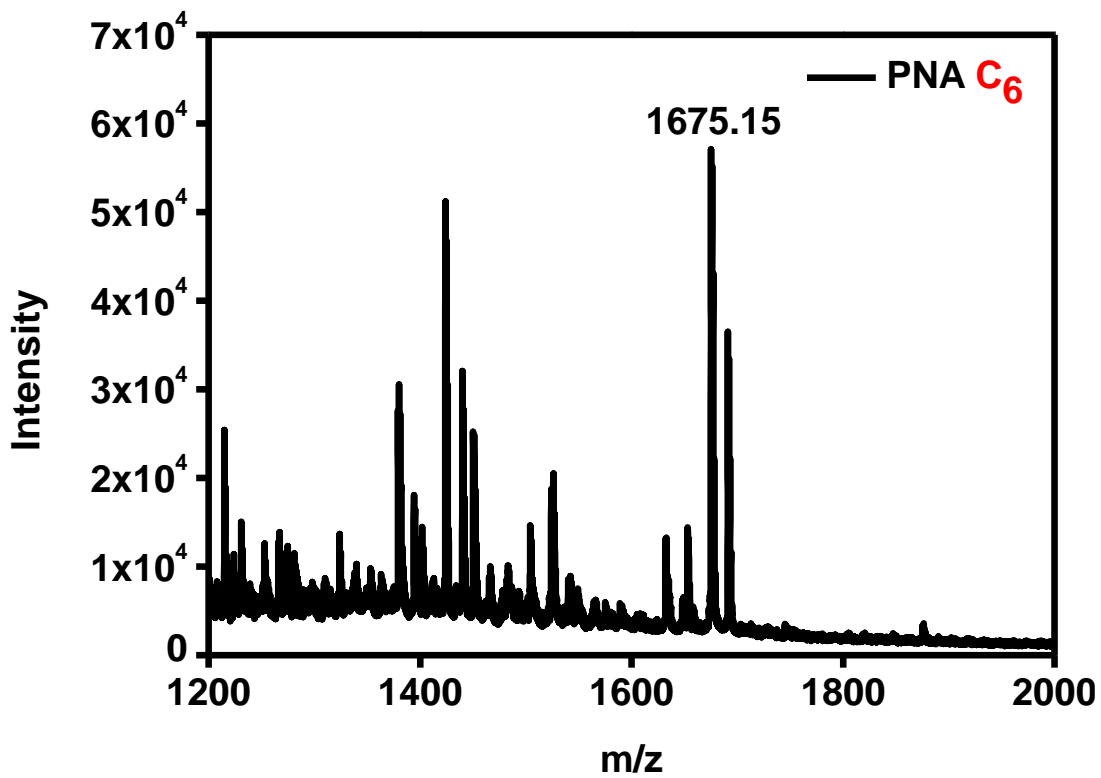
MALDI-TOF Spectrum of (PNA 16)



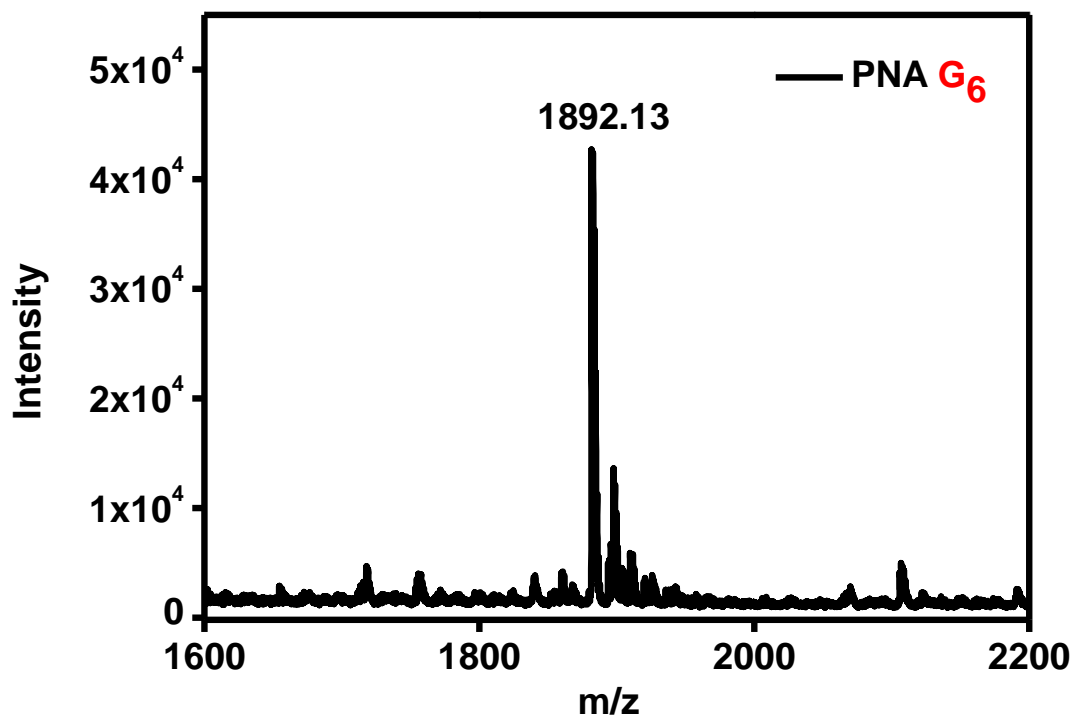
MALDI-TOF Spectrum of (PNA 17)



MALDI-TOF Spectrum of (PNA 18)



MALDI-TOF Spectrum of (PNA 19)



Chapter 6
**Biophysical Studies of *Janus* PNA
Oligomers**

6.1 Introduction

The ability of oligonucleotides to bind to the target complementary DNA/RNA can be investigated *in vitro* by employing various biophysical techniques, such as temperature-dependent UV absorbance, thermal UV-melting (UV- T_m), Circular Dichroism (CD)¹ for conformational features adopted by duplexes², Isothermal Titration Calorimetry (ITC)³ for evaluating thermodynamic features of binding.

The preceding chapter discusses the synthesis of rationally designed C γ -substituted PNA analogs in which nucleobases are attached to C γ -sidechain. The C γ -modified PNA monomers were introduced by solid phase synthesis to generate various modified PNA and *Janus* PNA oligomers to study the effect of chirality and side chain functional groups in influencing the binding properties to complementary oligonucleotides. This chapter reports comparative biophysical studies of modified PNA and *Janus* PNA oligomers and their hybrids with complementary nucleic acids using temperature dependent UV spectroscopy, CD spectroscopy⁴, Fluorescence spectroscopy⁵ and ITC^{6,7}. The objective is to understand the relative binding efficiency and selectivities of nucleobase sequences on each face (*t*-amide / C γ) of the *Janus* PNA, individually and simultaneously with complementary DNA sequences. It would provide us details of the binding strength, face selectivity and the synergetic effects of binding of cDNA on one face (*eg.* *t*-amide face) with the binding of cDNA on other face (C γ -face) and *vice versa*. The homo and mix *Janus* PNAs were evaluated for formation of single duplexes from each face and double duplexes from both face, in terms of thermal stability of each of the duplexes using temperature dependent UV absorbance,¹ conformation and effects on sequential formation of double duplexes from each initial duplex by CD spectroscopy, and further ethidium bromide assay by fluorescence study to investigate compute the binding strengths of the added PNA to DNA and thermodynamics of duplex formation by ITC.⁸⁻¹⁰ The formation of various types of complexes include triplexes (*Janus* PNA₂:DNA), duplexes (*Janus* PNA:DNA), double duplex (DNA:*Janus* PNA:DNA) and double duplex of triplex [DNA:(*Janus* PNA:DNA:*Janus* PNA):DNA] from *homo mixed Janus* PNAs. This chapter reports also studies on their ability to form i-motif¹¹ and G-quadruplex¹² stability in comparison with control *aeg* PNA/*aeg*-C γ PNA/DNA.

6.2 Objectives of the present work

The work presented in this chapter involves biophysical studies of the *Janus* PNAs in evaluating the following aspects:

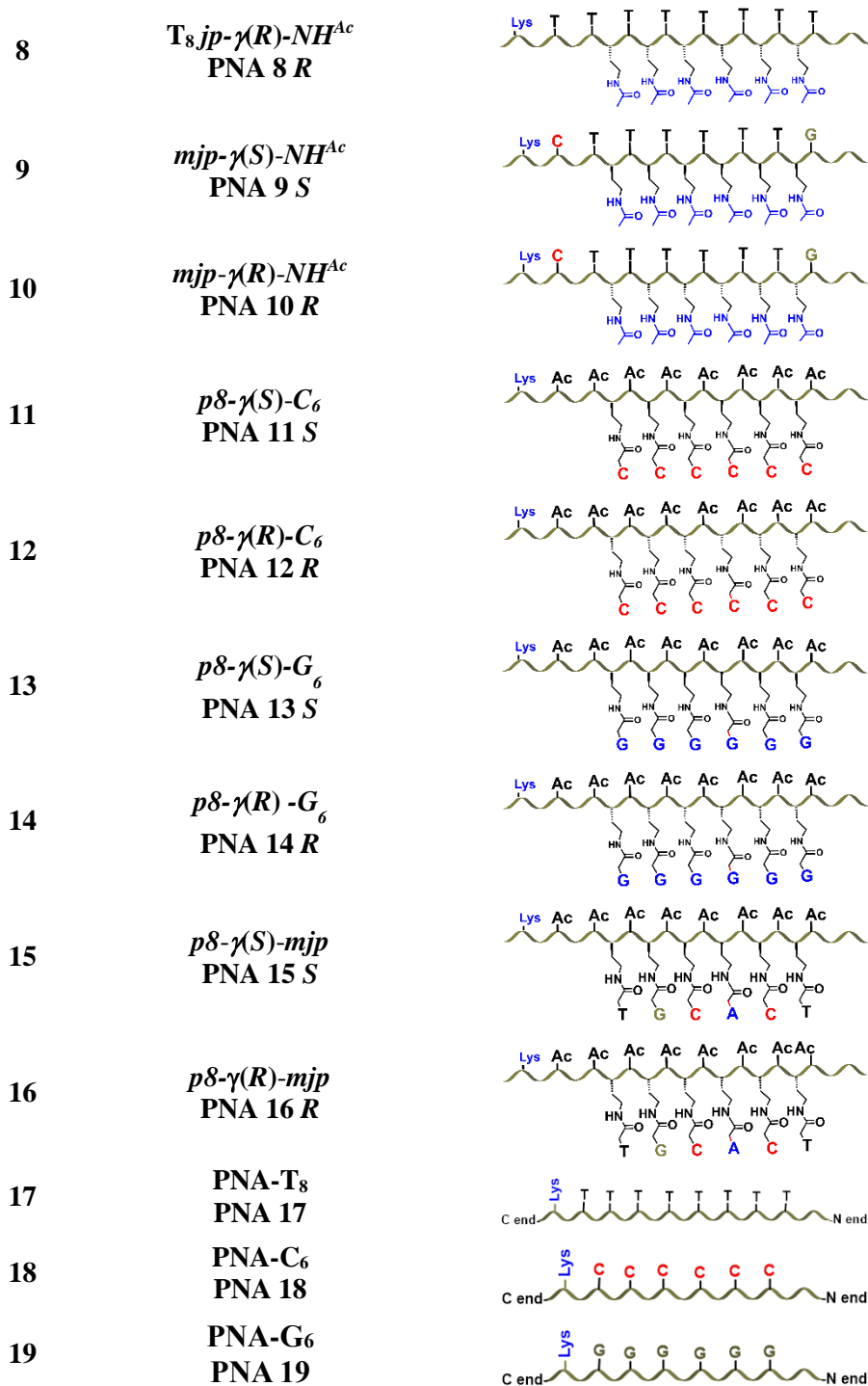
- Thermal stability, binding specificity and selectivity of *Janus* PNA with cDNA using temperature-dependent UV spectroscopy
- Conformational studies of *Janus* PNA:DNA duplex structures by CD spectroscopy
- Studying the thermal stability and conformation of *i*-motif of *Janus* PNA-C_n, *Janus* PNA-G₆ by UV-*T*_m and CD
- Measurement of thermodynamic parameters of *Janus* PNA:DNA interaction by ITC

6.3 *Janus* PNA (JP) oligomers used for biophysical studies

The modified PNA monomers with C^γ-substituted side chains were incorporated into PNA oligomers by solid-phase synthesis, followed by nucleobase coupling to obtain various kinds of *Janus* PNA oligomers (FigTable 6.1) for biophysical studies with complementary antiparallel DNAs (Table 6.2).

Table 6.1 Sequences of *Janus* PNA and aeg-PNA oligomers

Entry	Sequence Code	PNA Sequences
1	$T_8jp-\gamma(S)-C_6$ JP 1S	
2	$T_8jp-\gamma(R)-C_6$ JP 2R	
3	$T_8jp-\gamma(S)-G_6$ JP 3S	
4	$T_8jp-\gamma(R)-G_6$ JP 4R	
5	$mjp-\gamma(S)-PNA$ JP 5S	
6	$mjp-\gamma(R)-PNA$ JP 6R	
7	$T_8jp-\gamma(S)-NH^{Ac}$ PNA 7S	



6.4 DNA oligonucleotides used for biophysical studies

The DNA oligonucleotides **1-10** (Table 6.1) were obtained from commercial sources and used without further purification in the biophysical studies of PNAs or *Janus* PNAs.

Table 6.2 Complementary DNA used for biophysical studies of *Janus* PNA

Entry	DNA	Sequence (5' to 3')
1	DNA 1	AAAAAAAA or dA ₈
2	DNA 2	GGGGGG or dG ₆
3	DNA 3	CCCCCC or dC ₆
4	DNA 4	ACGTGA
5	DNA 5	GAAAA AC
6	DNA 6m	AAAACAAA
7	DNA 7m	GGGTGG
8	DNA 8m	CCCTCC
9	DNA 9m	GAAACAAC
10	DNA 10m	ACGCCA

6.5 Janus PNA complexes

The different *Janus* PNA (Table 6.1) sequences were hybridized with appropriate complementary DNAs (Table 6.2) in antiparallel orientation to generate PNA:DNA complexes that involved triplex, duplex, double duplex and duplex of triplex are shown in Figures 6.1 to 6.7.

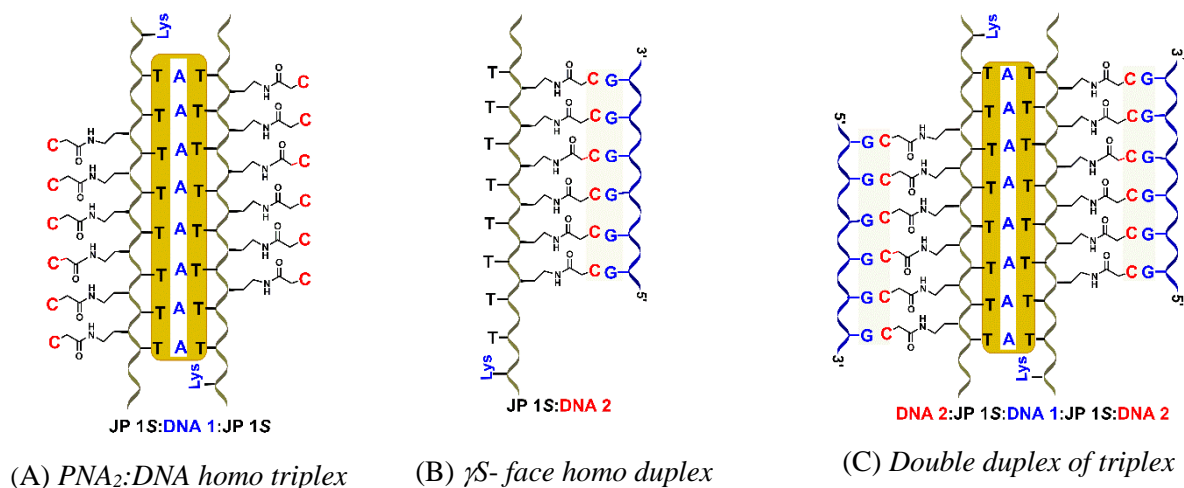


Figure 6.1 Homo Janus γ S PNA:DNA complex (A) triplex C_6 - γ S-*jpT*₈:dA₈:*T*₈*jp*- γ S- C_6 (**JP 1S:DNA 1:JP 1S**) (B) γ - face duplex *T*₈*jp*- γ S- C_6 :dG₆ (**JP 1S:DNA 2**) (C) double duplex of triplex dG₆: C_6 - γ S-*jpT*₈:dA₈:*T*₈*jp*- γ S- C_6 :dG₆ (**DNA 2:JP 1S:DNA 1:JP 1S:DNA 2**)

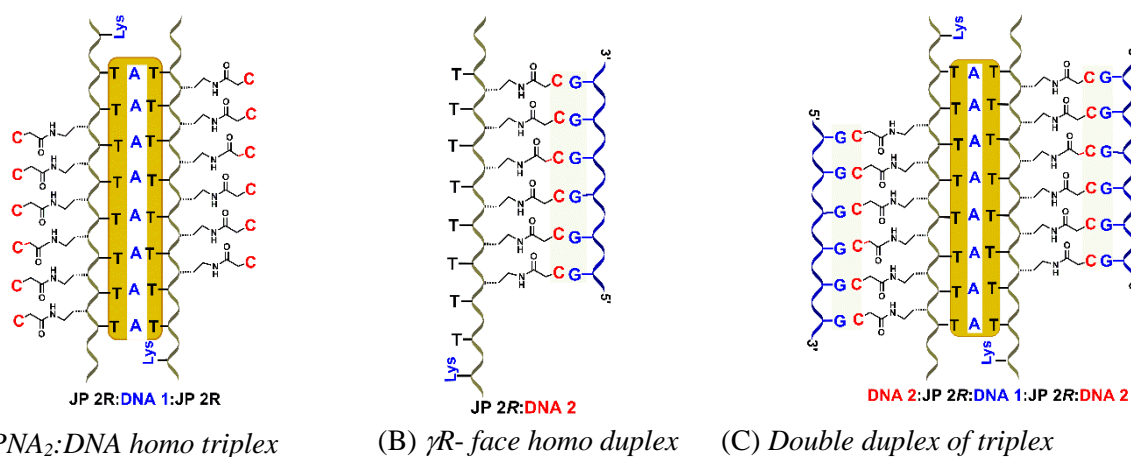


Figure 6.2 Homo Janus γ R PNA:DNA complexes (A) triplex $C_6\text{-}\gamma$ R- jpT_8 :dA₈: $T_8jp\text{-}\gamma$ R- C_6 (JP 2R:DNA 1:JP 2R) (B) γ R-face duplex $T_8jp\text{-}\gamma$ R- C_6 :dG₆ (JP 2R:DNA 2) (C) double duplex of triplex dG₆: $C_6\text{-}\gamma$ R- jpT_8 :dA₈: $T_8jp\text{-}\gamma$ R- C_6 :dG₆ (DNA 2:JP 2R:DNA 1:JP 2R:DNA 2)

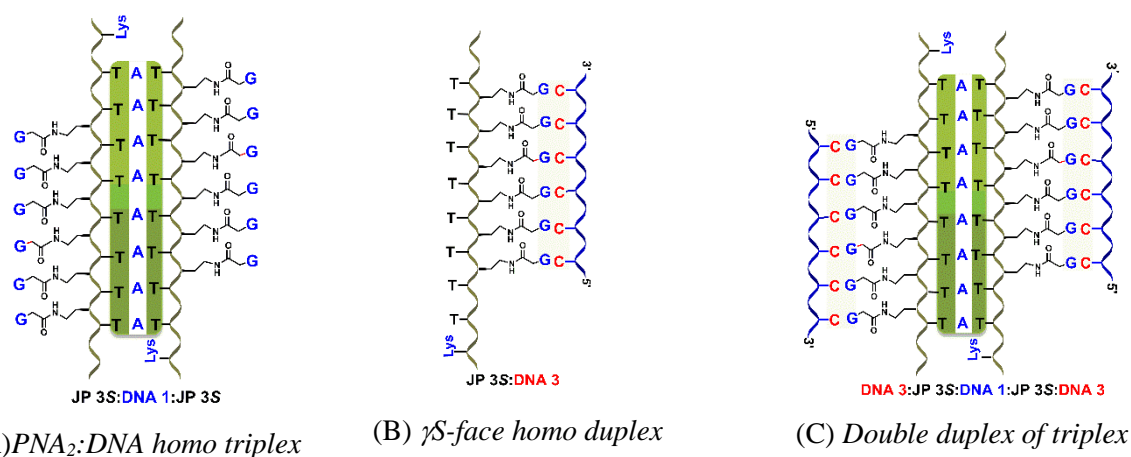


Figure 6.3 Homo Janus γ S PNA:DNA complex (A) triplex $G_6\text{-}\gamma$ S- jpT_8 :dA₈: $T_8jp\text{-}\gamma$ S- G_6 (JP 3S:DNA 1:JP 3S) (B) γ -face duplex $T_8jp\text{-}\gamma$ S- G_6 :dC₆ (JP 3S:DNA 3) (C) double duplex of triplex dC₆: $G_6\text{-}\gamma$ S- jpT_8 :dA₈: $T_8jp\text{-}\gamma$ S- G_6 :dC₆ (DNA 3:JP 3S:DNA 1:JP 3S:DNA 3)

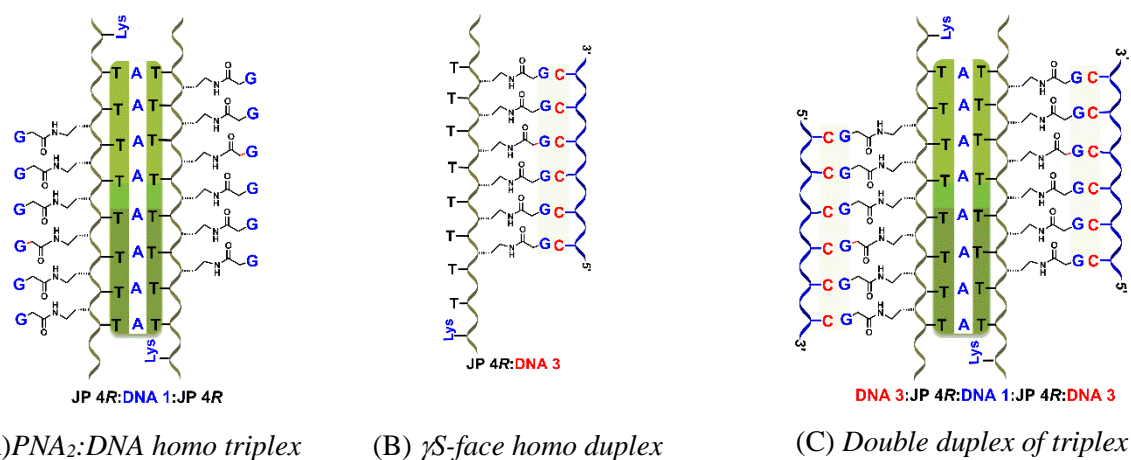


Figure 6.4 Homo Janus γ R PNA:DNA complex (A) triplex $G_6\text{-}\gamma$ R- jpT_8 :dA₈: $T_8jp\text{-}\gamma$ R- G_6 (JP 3R:DNA 1:JP 3R) (B) γ -face duplex $T_8jp\text{-}\gamma$ R- G_6 :dC₆ (JP 3R:DNA 3) (C) double duplex of triplex dC₆: $G_6\text{-}\gamma$ R- jpT_8 :dA₈: $T_8jp\text{-}\gamma$ R- G_6 :dC₆ (DNA 3:JP 3R:DNA 1:JP 3R:DNA 3)

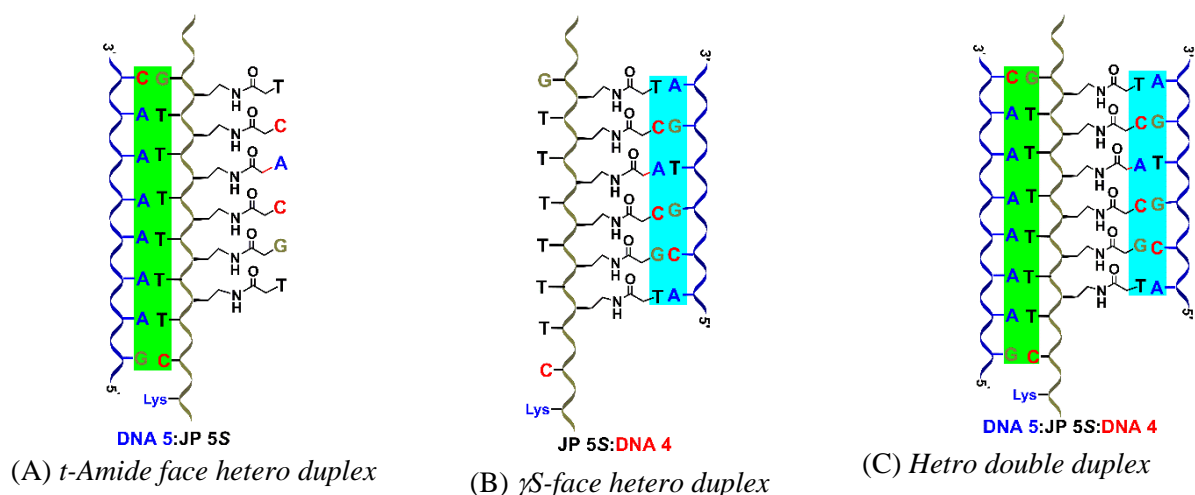


Figure 6.5 Hetero Janus PNA:DNA complexes (A) duplex *mjp*- γ S-PNA (DNA 5:JP 5S) (B) duplex *mjp*- γ S-PNA:DNA 6 (JP 5S:DNA 6) (C) double duplex DNA 5:*mjp*- γ S-PNA:DNA 6 (DNA 5:JP 5S:DNA 6)

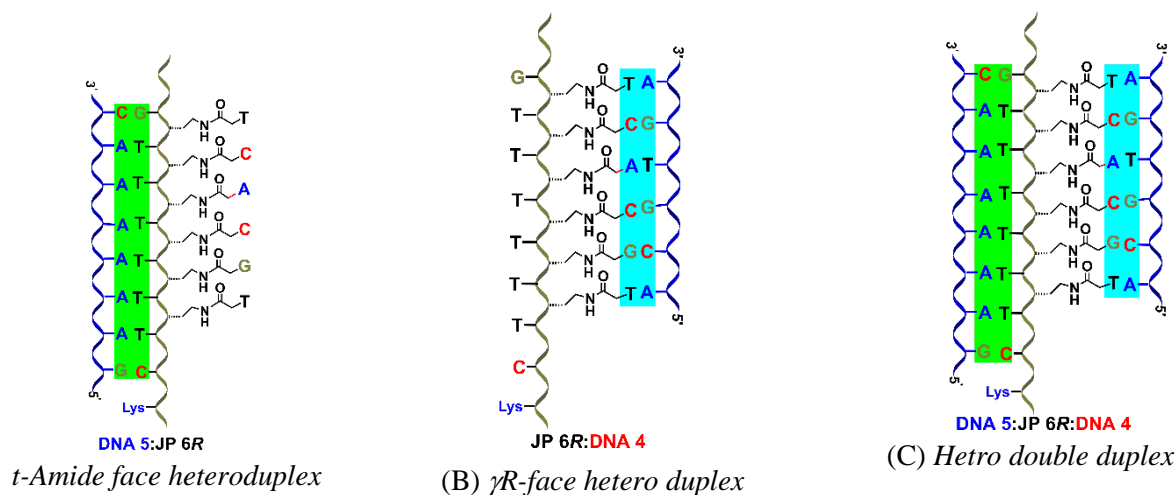


Figure 6.6 Hetero Janus PNA:DNA complexes (A) duplex *mjp*- γ R-PNA (DNA 5:JP 6R) (B) duplex *mjp*- γ R-PNA:DNA 6 (JP 6R:DNA 6) (C) double duplex DNA 5:*mjp*- γ R-PNA:DNA 6 (DNA 5:JP 6R:DNA 6)

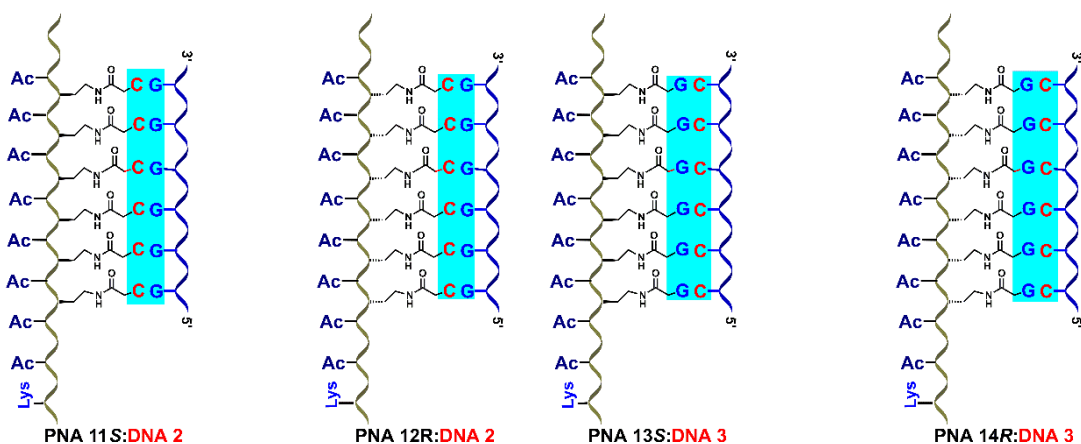


Figure 6.7 Homo duplex for control studies (A) duplex *p8*- γ S-*C*₆:d*G*₆ (PNA 11S:DNA 2) (B) duplex *p8*- γ R-*C*₆:d*G*₆ (PNA 12R:DNA 10) (C) duplex *p8*- γ S-*G*₆:d*C*₆ (PNA 13S:DNA 3) (D) duplex *p8*- γ R-*G*₆:d*C*₆ (PNA 14R:DNA)

6.6 Results

Thermal stabilities of PNA:DNA hybrids were examined by temperature-dependent UV absorption at 260 nm. The effect of chiral *Janus* PNAs on the conformation of derived *Janus* PNA:DNA hybrids were studied by CD spectroscopy. The binding affinity of modified PNAs with DNA was investigated by ethidium bromide displacement assays. ITC was used for thermodynamic studies for *Janus*-PNA:DNA interactions. The results obtained with various *Janus* PNAs with complementary DNA studies are described in the following sections. The first experiments done were to examine the base pairing abilities of nucleobases that are linked at C^γ(*R/S*) of PNA backbone through an amide sidechain. As this is a new PNA backbone without any tertiary amide linked nucleobases as in standard *aeg*-PNA, it is important to establish the base pairing properties of C^γ(*R/S*) sidechain linked nucleobases.

6.6.1 Thermal stability of *aeg*- γ *R/S*-PNA with complementary DNA duplexes

Hybridization studies of γ -*Janus* PNA surrogates *p8*- γ *S*-C₆ (PNA **11S**) and *p8*- γ *R*-C₆ (PNA **12R**) that have only secondary amide linked nucleobases in side chain and devoid of N-tertiary amide in backbone were carried by complexing with cDNA **2** by measuring UV absorbance as a function of temperature. The 8-mer *S*-PNA *p8*- γ *S*-C₆ (PNA **11S**) formed duplex *p8*- γ *S*-C₆:dG₆ with DNA **2** and in UV-T plot showed a nice sigmoidal transition characteristic of PNA:DNA duplexes and a single *T*_m corresponding to 44.6 °C (Figure 6.8). In comparison, the corresponding 8-mer *R*-PNA (PNA **12R**) also exhibited a sigmoidal transition with a single *T*_m of 38.9 °C. This clearly suggested the ability of C^γ-*S/R*-PNAs to form duplexes with cDNA and also the effect of stereochemistry at C^γ on duplex stability: the C^γ(*S*)-PNA:DNA duplex being more stable than C^γ(*R*)-PNA:DNA duplex by 5.7 °C. The standard *aeg*-PNA-C₆ (PNA **18**) formed duplex PNA-C₆:dG₆ (PNA **18**:DNA **2**) with *T*_m of 42.1 °C, while the corresponding DNA **2**:DNA **3** duplex melted at *T*_m of 39.9 °C. Thus the duplex from C^γ(*S*)-PNA *p8*- γ *S*-C₆ (PNA **11**) was more than *p8*- γ *R*-C₆ (PNA **12**), PNA **18**:DNA and DNA **2**:DNA **3** duplexes (Table 6.3). These results suggested that the nucleobases attached to *aeg* backbone through sidechain at C^γ are capable of forming stand alone duplexes with the stereochemistry of sidechain in both in *S* and *R* stereochemistry. This is a significant observation as the C^γ-PNAs PNA **11** and PNA **12** are new hitherto unknown PNA structures, devoid of tertiary amide linked nucleobases that can bind cDNA sequences.

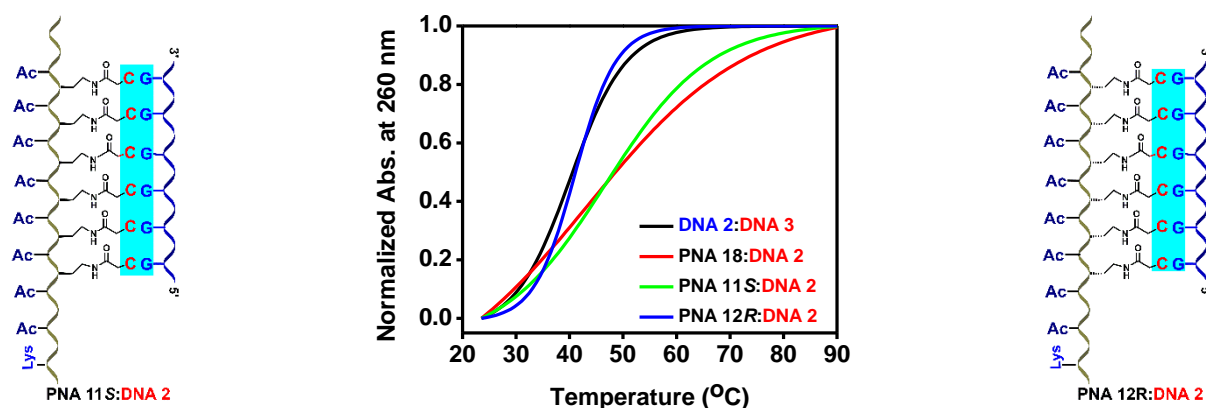


Figure 6.8 Temperature dependent UV absorbance curves for DNA duplexes of *aeg* PNA/*aeg-γ* PNA/DNA. PNA **11** = *pδ-γS-C₆*, PNA **12**= *pδ-γR-C₆*, PNA **18** = PNA-*C₆*, DNA **2** = 5'-GGGGGG-3', DNA **3** = 5'-CCCCC-3' Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

Table 6.3 UV- T_m (°C) values of DNA duplexes with *aeg/aeg-γ* PNA

Entry	PNA:DNA complexes	T_m (°C)		ΔT_m (°C)- duplex (PNA:DNA) –(DNA:DNA)
		t-Amide	C^γ	
1	PNA 11S :DNA 2 (<i>pδ-γS-C₆</i> :dG ₆)	-	44.6	+5.7
2	PNA 12R :DNA 2 (<i>pδ-γR-C₆</i> :dG ₆)	-	38.9	-1
3	PNA 18 :DNA 2 (PNA <i>C₆</i> :dG ₆)	42.1		+2.2
4	DNA 2 :DNA 3 (dG ₆ :dC ₆)	39.9		-

T_m values are accurate to ± 1.0 °C

6.6.1a Thermal stability of *aeg-γ*PNAs *pδ-γS-C₆* and *pδ-γR-C₆* with mismatch DNA duplexes

To examine the sequence specific binding of such new C^γ PNA structures, the stability of duplexes containing mismatched base pairs were investigated. The results of UV- T_m studies for PNA duplexes of *pδ-γS-C₆* (PNA **11S**) and *pδ-γR-C₆* (PNA **12R**) and *aeg* PNA **18** with DNA **7** (5'-GGGTGG-3) that has a T which introduces a T-C mismatch in the derived duplexes are shown in Figure (6.9). The C^γ -PNA *pδ-γS-C₆* (PNA **11S**) shows a destabilization of 5.8 °C, and *pδ-γR-C₆* (PNA **12R**) was destabilized by 4.5 °C compared to corresponding perfect matched duplexes (Table 6.4). These ΔT_m destabilization were similar to that of *aeg*-PNA **18** which exhibited a destabilization of 4.8 °C with its duplex with mismatch DNA **7**. These results clearly affirmed that the C^γ (S/R) PNAs PNA **11S** and PNA **12R** form sequence base pairing with cDNA in the same way as the standard *aeg*-PNA and so are appropriate for for constructing designed C^γ -Janus PNAs for simultaneous binding of two cDNAs, which is the objective of this work.

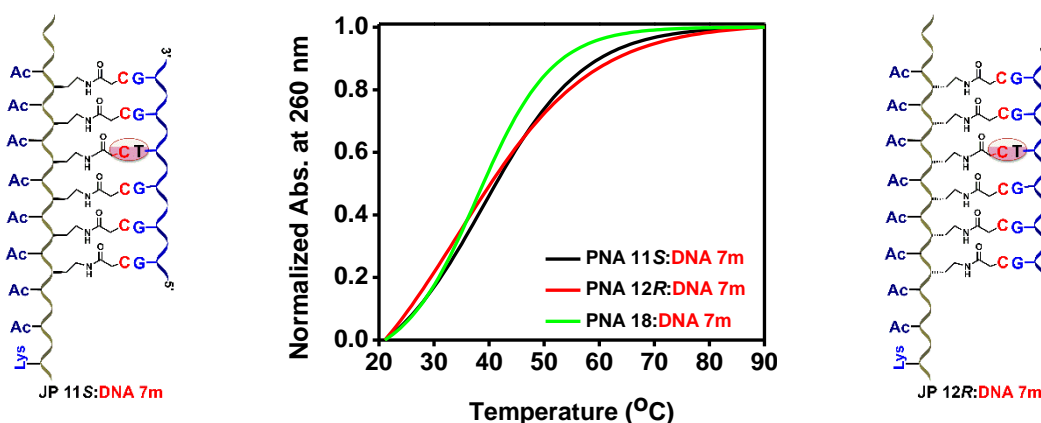


Figure 6.9 Temperature dependent UV absorbance curves for mismatch DNA/ *aeg* PNA/ *aeg* - C^γ PNA:DNA duplexes, PNA **11** = $p8\text{-}\gamma S\text{-}C_6$, PNA **12** = $p8\text{-}\gamma R\text{-}C_6$, PNA **18** = PNA- C_6 , and DNA **7** = 5'-GGGTGG-3; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

Table 6.4 UV- T_m ($^\circ\text{C}$) values of mismatch DNA duplexes with *aeg/aeg*- γ PNA

Entry	PNA:DNA mismatch duplexes	T_m ($^\circ\text{C}$)		ΔT_m^1 ($^\circ\text{C}$) PNA:DNA(mismatch) - PNA:DNA (complementary)
		t-Amide	C^γ	
1	PNA 11S :DNA 7m	-	38.8	- 5.8
2	PNA 12R :DNA 7m		34.4	- 4.5
3	PNA 18 :DNA 7m	37.3		- 4.8

T_m values are accurate to ± 1.0 $^\circ\text{C}$, ΔT_m^1 indicates the difference in T_m of PNA with mismatch DNA and cDNA

6.6.1b CD spectra of DNA duplexes of C^γ -PNA:DNA duplexes

For studying relative conformations of the different PNA:DNA duplexes and the effect of C^γ -stereochemistry, CD spectra of duplexes of PNA $p8\text{-}\gamma S\text{-}C_6$ (PNA **11S**), $p8\text{-}\gamma R\text{-}C_6$ (PNA **12**) and *aeg*-PNA **18** with cDNA **2** were recorded (Figure 6.10). The CD spectra of $C^\gamma(S)$ duplex $p8\text{-}\gamma S\text{-}C_6$:DNA **2** (PNA **11S**:DNA **2**) shows a positive CD band at 260 nm a negative maxima at 239 nm, whereas the $C^\gamma(R)$ duplex $p8\text{-}\gamma R\text{-}C_6$:DNA **2** (PNA **12R**:DNA **2**) shows a slightly weak positive band at 270 nm and a negative band at 248 nm with cross-over point at 255 nm. The CD spectra of duplex *aeg*-PNA **18**:DNA **2** shows strong positive bands at 261 nm and 208 nm (Figure 6.10). The CD profile suggests that the $C^\gamma(S)$ PNA:DNA duplex is more akin to *aeg*-PNA duplex, while the $C^\gamma(R)$ PNA:DNA duplex is very closely resembles PNA:DNA duplex with a slight shift in positive band. It was previously observed that the $C^\gamma(S)$ PNA:DNA has better duplex stability compared to $C^\gamma(R)$ PNA:DNA. Overall, both UV-T and CD spectra suggests that the C^γ -PNAs form duplexes with DNA comparable to that from *aeg*-PNA.

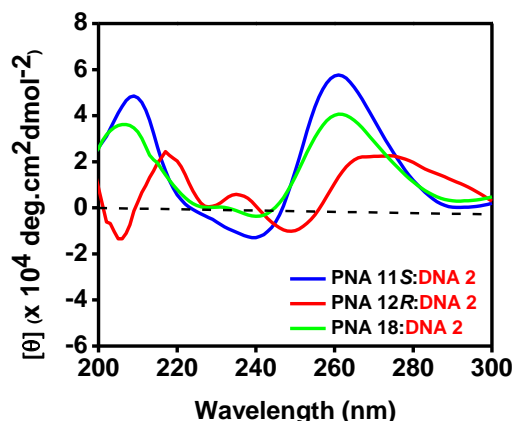


Figure 6.10 CD spectra of duplexes (A and B) **PNA 11S**= $p\delta$ - γ S- C_6), **PNA 12R**= $p\delta$ - γ R- C_6 and **PNA 18**= (PNA- C_6), DNA 2 (dC₆= 5' CCCCCC 3'), Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

6.6.1c Thermal stability of aeg - γ (S/R)- G_6 PNA with complementary DNA duplexes.

In order to examine the sequence effects on the ability of C^γ (S/R) PNAs to form duplex, $p\delta$ - γ S- G_6 (PNA **13S**) and $p\delta$ - γ R- G_6 (PNA **14R**) in which G_6 is linked to C^γ sidechain were duplexed with dC₆ (DNA **3**). The duplexes of PNA **13S** and PNA **14R** with DNA **3** gave nice sigmoidal transitions with T_m of 53.6 °C and 52.4 °C respectively. In comparison, the aeg PNA- G_6 (PNA **19**:DNA **3**) duplex showed a slightly higher T_m of 59.8 °C (Figure 6.11). All three PNA- G_6 duplexes gave T_m s considerably higher than the T_m of 40 °C for dG₆:dC₆ (Table 6.5). These results clearly confirmed that all 4 PNAs of the series C^γ (S/R)- C_6 C^γ (S/R)- G_6 formed duplexes successfully with cDNAs.

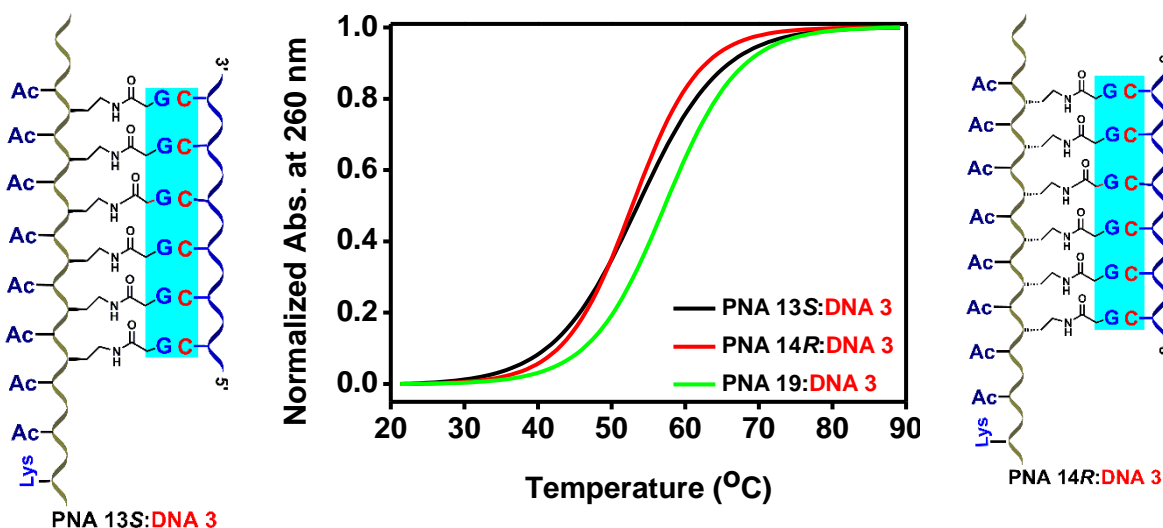


Figure 6.11 Temperature dependent UV absorbance curves for complementary DNA/ aeg PNA/ aeg - γ PNA:DNA duplexes, PNA **13** = S $p\delta$ - γ - G_6 , PNA **14** (R $p\delta$ - γ - G_6), PNA **19**=PNA G_6 , DNA **3** (dC₆= 5' CCCCCC 3') Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM).

Table 6.5 UV- T_m ($^{\circ}\text{C}$) values of DNA duplexes with *aeg* PNA/*aeg*- C^{γ} PNA

Entry	PNA:DNA complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)- duplex (PNA:DNA) - (DNA:DNA)
		t-Amide	C^{γ}	
1	PNA 13S :DNA 3 (<i>p8-γS</i> - G_6 :dC $_6$)	-	53.6	+13.7
2	PNA 14R :DNA 3 (<i>p8-γR</i> - G_6 :dC $_6$)		52.4	+12.5
3	PNA 19 :DNA 3 (PNA G_6 :C $_6$)	59.8		+19.9
4	DNA 2 :DNA 3 (dG $_6$:dC $_6$)	39.9		-

T_m values are accurate to ± 1.0 $^{\circ}\text{C}$

6.6.1d Thermal stability of *p8- γ S*- G_6 and *p8- γ R*- G_6 PNAs with mismatch DNA duplexes

The duplexes from *p8- γ S*- G_6 (PNA **13S**), *p8- γ R*- G_6 (PNA **14R**) and *aeg*-PNA- G_6 (PNA **19**) with DNA **8** (5'-CCCTCC-3') that inserts a T-G mismatch showed destabilization with $\Delta T_m = 5.8$ $^{\circ}\text{C}$, 4.5 $^{\circ}\text{C}$ and 4.8 $^{\circ}\text{C}$ respectively compared with their perfect duplexes (Figure 6.12 Table 6.16, entry 1, 2 and 3). These results further reinforce the sequence specific base pairing of C^{γ} (S/R)-PNAs with cDNA.

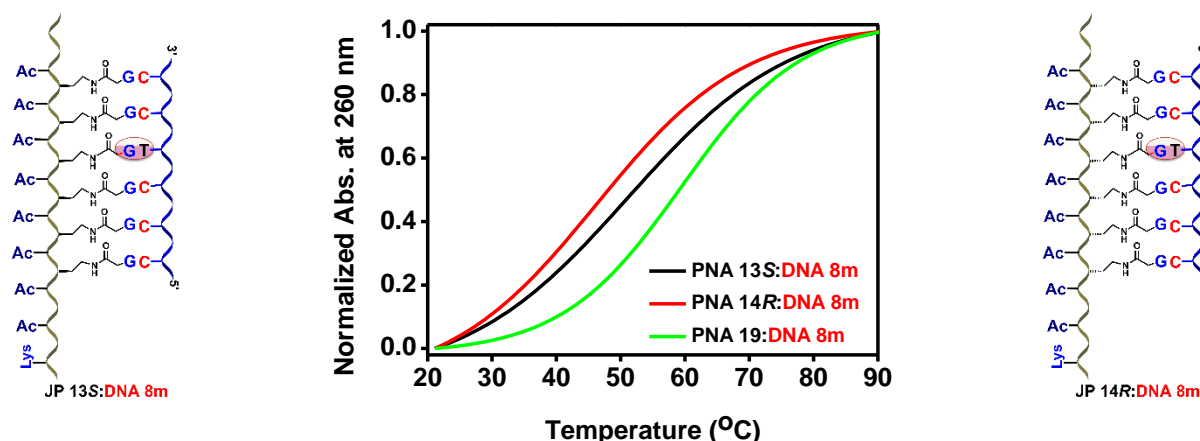


Figure 6.12 Temperature dependent UV absorbance curves for mismatch DNA/ *aeg* PNA/ *aeg*- C^{γ} PNA:DNA duplexes, PNA **13S**= *p8- γ S*- G_6 , PNA **14R** = *p8- γ R*- G_6 , PNA **19** = PNA- G_6 , and DNA **8m** = 5'-CCCTCC-3; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

Table 6.6 UV- T_m ($^{\circ}\text{C}$) values of mismatch DNA duplexes with *aeg/aeg- γ* PNA

Entry	PNA:DNA mismatch duplexes	T_m ($^{\circ}\text{C}$)		ΔT_m^1 ($^{\circ}\text{C}$) PNA:DNA(mismatch) - PNA:DNA (complementary)
		t-Amide	C^{γ}	
1	PNA 13S :DNA 8m	-	50.9	- 5.8
2	PNA 14R :DNA 8m		45.7	- 4.5
3	PNA 19 :DNA 8m	58.6		- 4.8

T_m values are accurate to ± 1.0 $^{\circ}\text{C}$, ΔT_m^1 indicates the difference in T_m of PNA with mismatch DNA and cDNA.

6.6.1e CD spectra of DNA duplexes of $C^{\gamma}(S/R)$ PNAs $p8-\gamma S-G_6$ (PNA 13S) and $p8-\gamma R-G_6$ (PNA 14R)

The CD spectra of duplexes of $p8-\gamma S-G_6$:DNA 3 (PNA 13S) shows a positive CD band at 273 nm, a negative band at 249 nm with cross-over points at 232 nm and 258 nm (Figure 6.13). The $C^{\gamma}(R)$ duplex $p8-\gamma R-G_6$:DNA 3 (PNA 14R:DNA 3) shows a positive band at 274 nm, a negative band at 250 nm with crossover at 260 nm (Figure 6.13). These CD profiles match very well with that of the duplex of *aeg*-PNA 19:DNA 3 that shows strong positive band at 275 nm and negative band at 249 nm. These CD results for $C^{\gamma}(S/R)-G_6$ PNAs along with CD profiles of $C^{\gamma}(S/R)-C_6$ (section 6.8.1b) endorse that the $C^{\gamma}(R/S)$ PNAs form duplexes with cDNA with duplex conformation same as in standard *aeg*-PNA.

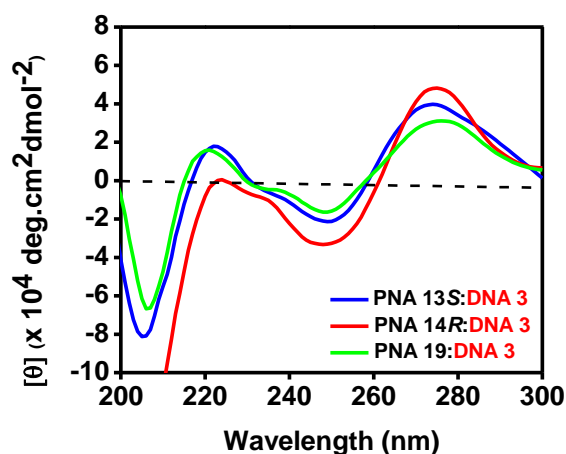


Figure 6.13 CD spectra of duplexes PNA 13S= $p8-\gamma S-G_6$, PNA 12R= $p8-\gamma R-G_6$ and PNA 19=PNA- G_6 , DNA 3(dC₆= 5' CCCCCC 3'), Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

6.6.2 Binding studies of homo Janus C^{γ} -PNAs $T_{8jp}-\gamma S-C_6$ and $T_{8jp}-\gamma R-C_6$ with complementary DNA

After obtaining overwhelming confirmation from results of previous sections that new PNA backbone with nucleobases linked to a C^{γ} -sidechain at in both *S* and *R* configurations and without the tertiary amide linked nucleobases as in standard *aeg*-PNA form sequence specific duplexes with cDNA, this section describes results with Janus PNAs, which are composites carrying nucleobases attached in both manner on same PNA strand. These were designed to bind two cDNA strands simultaneously, one to sequences located on N-tertiary amide linkage and the other to sequences carried by C^{γ} -sidechain. The duplex generated from binding to N-tertiary amide linked bases are designated as “t-amide face duplex” and duplex from binding to C^{γ} -sidechain linked nucleobases are termed ‘ C^{γ} -face duplex’

6.6.2a Thermal stability of complex of C^γ(S/R) homo Janus PNA T₈jp-γS-C₆:DNA duplexes

Homo Janus PNA T₈jp-γS-C₆ (**JP 1S**) was hybridized individually with t-amide face cDNA **1** (dA₈) and γ-face cDNA **2** (dG₆) and the temperature dependent absorbance plots are shown in Figure 6.14. As per literature precedence¹³ complexation of dA_n:dT_n always lead to dT:dA:dT triplexes, which can be seen from characteristic CD profiles (see later). In case of dC:dG complexes, duplexes prevail at physiological pH, while triplexes that need CH⁺ are formed only at acidic pH. The γ-face duplex T₈jp-γS-C₆:dG₆ (**JP 1S**:DNA **2**) gave a sigmoidal transition corresponding to T_m of 68.2 °C (Figure 6.14A). The t-amide face triplex C₆-Sγ-jpT₈:dA₈:T₈jp-γS-C₆ (**JP 1S**:DNA **1**:**JP 1S**) showed a single transition corresponding T_m of 58.8 °C (Table 6.7, entry 2). Thus the γ-face duplex T₈jp-Sγ-C₆:dG₆ (**JP 1S**:DNA **2**) shows a higher T_m (+9.4°C) compared to t-amide face triplex in spite of having less number of base pairs. Further hybridization of triplex C₆-γS-jpT₈:dA₈:T₈jp-γS-C₆ with γ-face cDNA dG₆ gave a complex that now exhibited double sigmoidal transition with two different melting temperatures T_{m1} (57.4°C) and T_{m2} (77.9°C). This should correspond to double duplex of triplex dG₆:C₆-γS-jpT₈:dA₈:T₈jp-γS-C₆:dG₆ (DNA **2**:**JP 1S**:DNA **1**:**JP 1S**:DNA **2**) originating from duplex formation on both γ-faces on the initially formed triplex (Figure 6.14C) (Table 6.7, entry 3). Significantly the two T_m s observed in double duplex of triplex are higher than those observed with individual duplexes, suggesting that both triplex and duplex are stabilized by mutual presence of each other in a synergistic manner. Extending the observation that γ-face duplex is more stable than the t-amide face triplex (as seen in individual complexes), it is seen that in Janus PNA complex, the γ-face duplex showed stabilization of +19.1 °C and the t-amide duplex was stabilized by +10.8 °C over that in individual duplex and triplex respectively (Table 6.2, entry 3). The differential T_m between γ-face duplex and t-amide face duplex in duplex of triplex was also similar (+8.4 °C) compared to +9.1 °C in isolated duplexes. These results indicated that the designed C^γ-Janus PNA **JP1S** can indeed bind simultaneously to two cDNAs giving rise to double duplex of a triplex seen by sequential melting in UV-T plots. This gave the first proof of concept results.

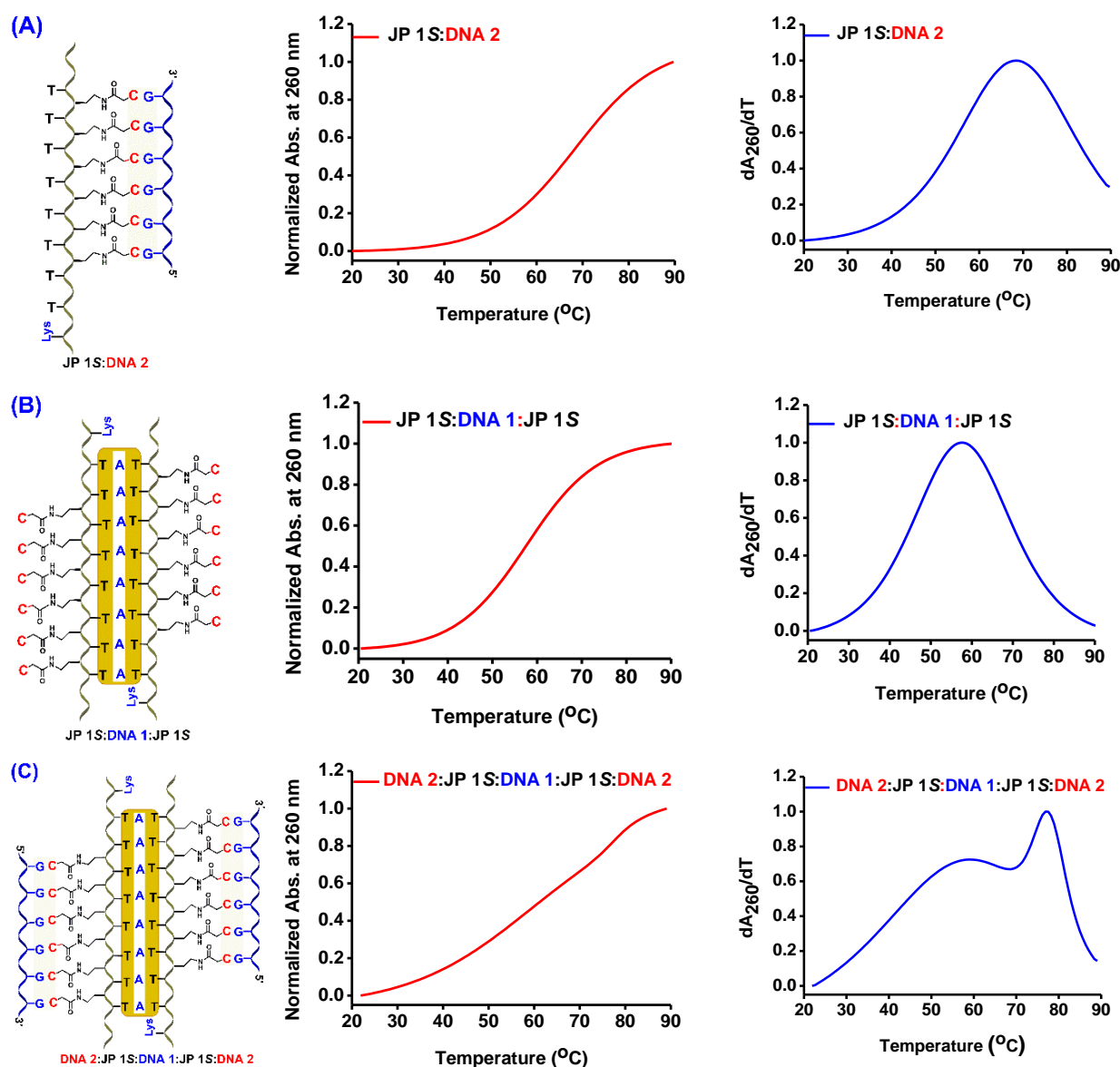


Figure 6.14 Temperature dependent UV absorbance curves for DNA complexes of **JP 1S** ($T_8jp-\gamma S-C_6$) with DNA 1 ($dA_8 = 5' AAAAAAAAA 3'$) and DNA 2 ($dG_6 = 5' GGGGGG 3'$); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (A) **JP 1S:DNA 2** ($T_8jp-\gamma S-C_6:dG_6$) (B) **JP 1S:DNA 1:JP 1S** ($C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6$) (C) **DNA 2:JP 1S:DNA 1:JP 1S:DNA 2** ($dG_6:C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6:dG_6$).

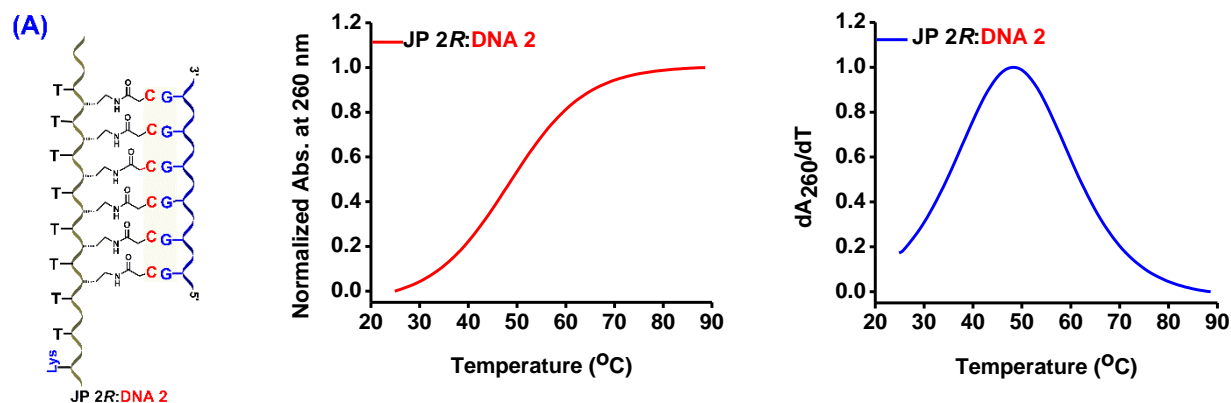
Table 6.7 UV- T_m ($^{\circ}C$) of complexes of **JP 1S** with complementary DNA 1 and DNA 2

Entry	<i>Homo Janus</i> PNA:DNA complexes	T_m ($^{\circ}C$)		ΔT_m ($^{\circ}C$)		
		<i>t</i> -Amide	C^{γ}	<i>t</i> -Amide	C^{γ}	C^{γ} - <i>t</i> -Amide
1	$T_8jp-\gamma S-C_6:dG_6$ (JP 1S:DNA 2)	-	68.2	-	-	+9.4
2	$C_6-\gamma S-jpT_8:dA_8:ST_8jp-\gamma S-C_6$ (JP 1S:DNA 1:JP 1S)	58.8	-	-	-	
3	$dG_6:C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6:dG_6$ (DNA 2:JP 1S:DNA 1:JP 1S:DNA 2)	57.4	77.9	+10.8	+19.1	+8.3

DNA **1** ($\mathbf{dA}_8 = 5' \text{ AAAAAAAAAA } 3'$); DNA **2** ($\mathbf{dG}_6 = 5' \text{ GGGGGG } 3'$); ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C

6.6.2b Thermal stability of duplex of triplex of homo Janus-PNA $T_{8jp-\gamma R-C_6}$:DNA duplexes

In order to examine the effect of stereochemistry at C^γ - on complexation of homo Janus PNA, the $C^\gamma(R)$ Janus PNA $T_{8jp-\gamma R-C_6}$ (**JP 2R**) was individually hybridized with t-amide face cDNA **1** (\mathbf{dA}_8) and γ -face cDNA **2** (\mathbf{dG}_6) and the UV- T_m plots are shown in Figure 6.15. The C^γ -face duplex $T_{8jp-\gamma R-C_6}:\mathbf{dG}_6$ (**JP 2R**:DNA **2**) gave a T_m 48.34 °C (Table 6.8, entry 1) and the t-amide face triplex $C_6-\gamma R-jpT_8:\mathbf{dA}_8:T_{8jp-\gamma R-C_6}$ (DNA **1**:**JP 2R**) showed a single transition corresponding T_m of 43.6 °C (Table 6.8, entry 2). Thus γ -face duplex **JP 2R**:DNA **2** ($T_{8jp-\gamma R-C_6}:\mathbf{dG}_6$) had a higher T_m +4.7°C compared to t-amide face triplex as seen for the corresponding $C^\gamma(S)$ analogues. The hybridization of triplex with γ -face cDNA (\mathbf{dG}_6) lead to formation of double duplex of triplex $\mathbf{dG}_6:C_6-\gamma R-jpT_8:\mathbf{dA}_8:T_{8jp-\gamma R-C_6}:\mathbf{dG}_6$ (DNA **2**:**JP 2**:DNA **1**:**JP2**:DNA **2**) as seen by two different melting temperatures T_{m1} (41.0°C) and T_{m2} (65.8°C) (Table 6.8, entry 3). Among these two T_m s the γ -face duplex T_m is higher than duplex from t-amide face. The γ -face duplex showed stabilization of +17.5 °C and the t-amide face duplex was stabilized by +2.6 °C over that in individual duplex and triplex (Table 6.8,entry 3). This huge difference lead a very nice separation of the transitions. The differential T_m between γ -face duplex and t-amide face duplex in double duplex of triplex (+14.9 °C) was higher compared to +4.7 °C seen in isolated duplexes. These clearly pointed out to synergistic effect on duplex stabilization in Janus PNA duplex of triplex. Thus the C^γ -R-stereomeric PNA **JP 2R** has lower stability of duplexes/triplexes compared the C^γ -S-PNA.



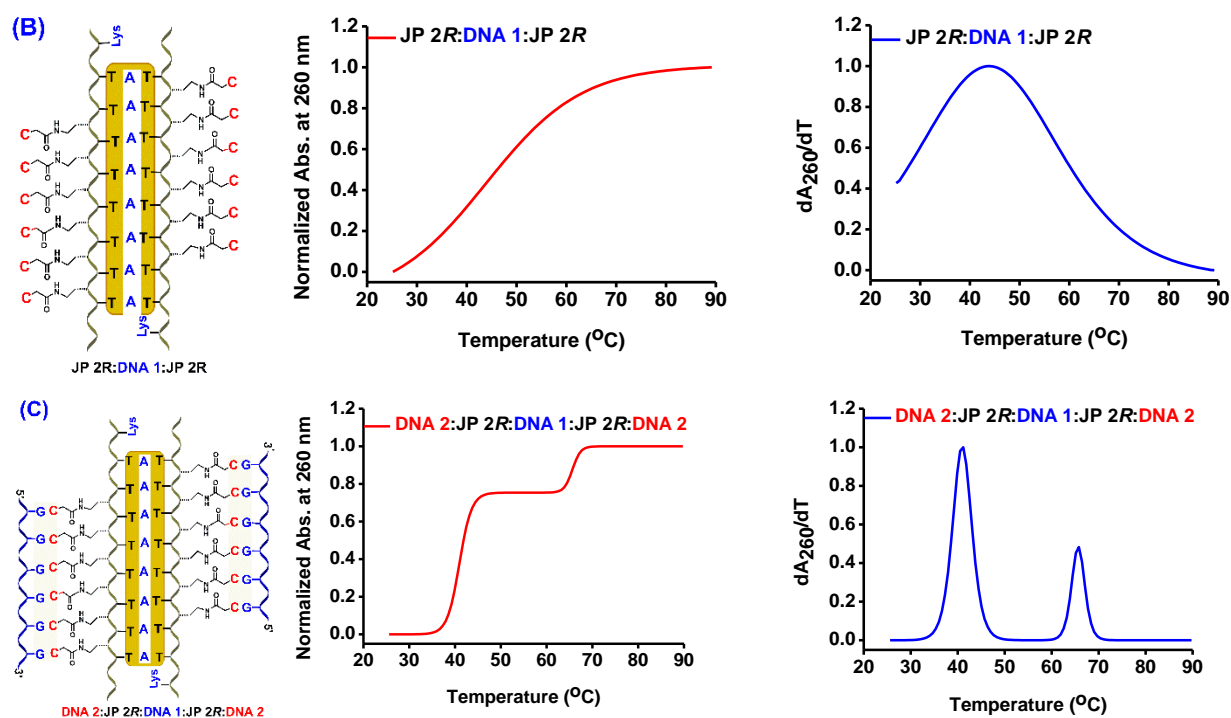


Figure 6.15 Temperature dependent UV absorbance curves for DNA complexes of **JP 2R** ($T_{8jp-\gamma R-C_6}$) with **DNA 1** ($dA_8 = 5' \text{ AAAAAAAAAA } 3'$) and **DNA 2** ($dG_6 = 5' \text{ GGGGGG } 3'$); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (A) **JP 2R:DNA 2** ($T_{8jp-\gamma R-C_6}:dG_6$) (B) **JP 2R:DNA 1:JP 2R** ($C_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-C_6}$) (C) **DNA 2:JP 2R:DNA 1:JP 2R:DNA 2** ($dG_6:C_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-C_6}:dG_6$).

Table 6.8. UV- T_m ($^{\circ}\text{C}$) of complexes of **JP 2R** with complementary **DNA 1** and **DNA 2**

Entry	<i>Homo Janus</i> PNA:DNA complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)		
		<i>t</i> -Amide	C^{γ}	<i>t</i> -Amide	C^{γ}	C^{γ} - <i>t</i> -Amide
1	$T_{8jp-\gamma R-C_6}:dG_6$ (JP 2R:DNA 2)	-	48.3	-	-	+4.7
2	$C_6-\gamma jpT_8R:dA_8:RT_{8jp-\gamma C_6}$ (JP 2R:DNA 1:JP 2R)	43.6	-	-	-	
3	$(dG_6:C_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-C_6}:dG_6)$ DNA 2:JP 2R:DNA 1:JP 2R:DNA 2	41.0	65.8	+2.6	+17.5	+14.9

DNA 1 ($dA_8 = 5' \text{ AAAAAAAAAA } 3'$); **DNA 2** ($dG_6 = 5' \text{ GGGGGG } 3'$); ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 $^{\circ}\text{C}$

6.2.2c CD spectra of $C^{\gamma}(S/R)$ homo Janus triplexes and duplexes

The formation of triplexes and duplexes were characterized by CD spectroscopy. The CD spectra of *Janus* PNA complexes $C_6-\gamma S-jpT_8:dA_8:T_{8jp-\gamma S-C_6}$ and $C_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-C_6}$ showed a double hump pattern with positive bands at 265 nm and 282 nm (Figure 6.16) and major positive band at 220 nm with a shoulder. The positive bands are accompanied by

a negative minimum at 249 nm with cross-over points at 240 nm and 259 nm. This CD profile is characteristic of the (PNA-T₈)₂:poly dA triplex composed from right-handed helices.¹³ The very close overlapping CD profiles of C^γ-Janus PNAs with *aeg*-PNA profile suggests the formation of triplex dA₈:(T₈*jp-γS-C*₆)₂ DNA **1**:(**JP 1S**)₂ (1:2) and dA₈:(T₈*jp-γR-C*₆)₂ (DNA **1**:**JP 2R**)₂ with DNA:PNA (1:2) stoichiometry.

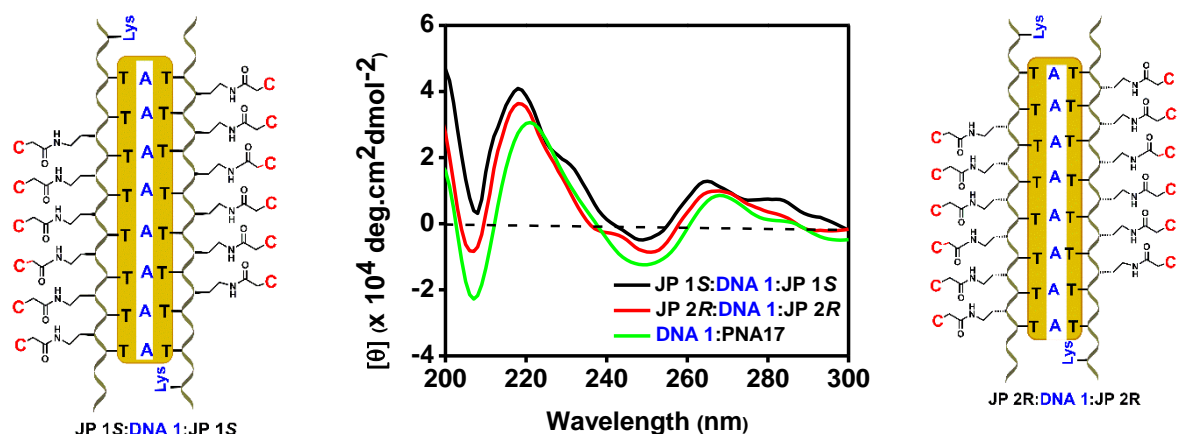
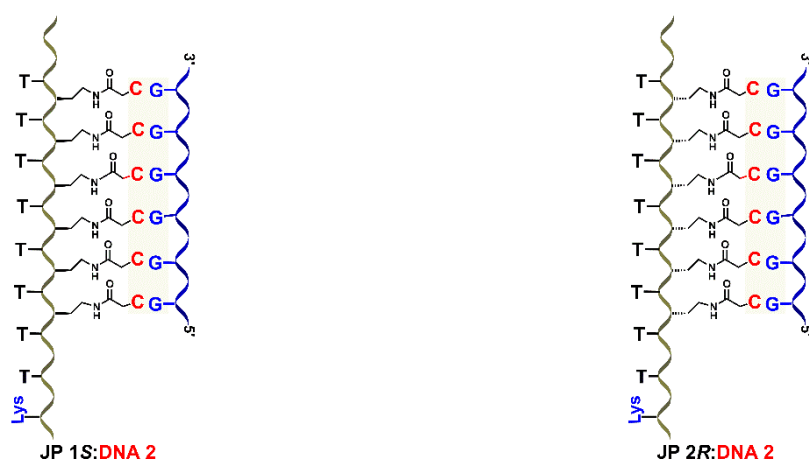


Figure 6.16 CD spectra of Amide face DNA duplexes from *aeg* PNA, Janus PNA T₈*jp-γS-C*₆ (**JP 1S**) and T₈*jp-γR-C*₆ (**JP 2R**) Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

The CD spectra of C^γ Janus PNA duplexes T₈*jp-γS-C*₆:dG₆ (**JP 1S**:DNA **2**), T₈*jp-γR-C*₆:dG₆ (**JP 2R**:DNA **2**) and *aeg*-PNA C₆:dG₆ (**PNA 18**:DNA **2**) are shown in Figure 6.17 along with CD spectra of DNA duplex dG₆:dC₆ (DNA **2**:DNA **3**). The CD spectra are very different from that of triplexes seen in Figure 6.17. All duplexes showed a positive band in the region 260 nm - 263 nm and a very mild negative band around 240 nm with cross-over points at 245 nm. The near overlapping pattern of all duplexes indicated their similar duplex conformation.



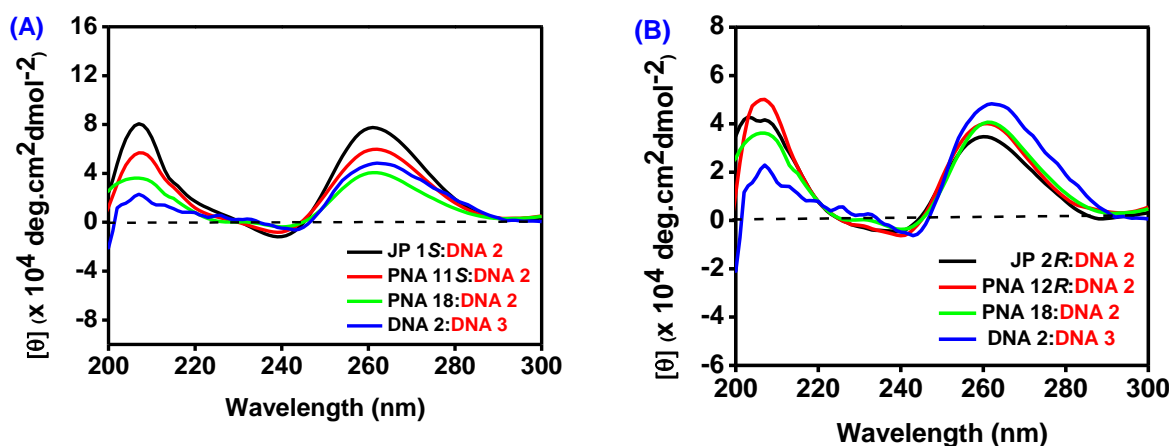


Figure 6.17 CD spectra of C^γ face of DNA duplexes from, Janus PNA (**JP 1S** or $T_{8jp-\gamma S-C_6}$), (**JP 2R** or $T_{8jp-\gamma R-C_6}$), aeg PNA, aeg γ PNA, Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 Mm.

In comparison with the CD spectra of isolated triplexes and duplexes as shown above, the double duplex of Janus triplex constituted from $T_{8jp-\gamma S-C_6}$ (**JP 1S**) and $T_{8jp-\gamma R-C_6}$ (**JP 2R**) with DNA 1 (dA_8) and DNA 2 (dG_6) (Figure 6.18) shows two distinct positive bands at 260 nm to 218 nm, with a minor negative band at 242 nm with cross-over points at 250 nm and 238 nm. The CD spectra is a composite of the t-amide triplex and γ -duplex and the fact that this complex showed two T_m s with enhanced T_m of both the triplex and duplex components suggested that the CD to be characteristic of double duplex of triplex. It is not a result of simple additive spectra of the isolated triplex and duplex.

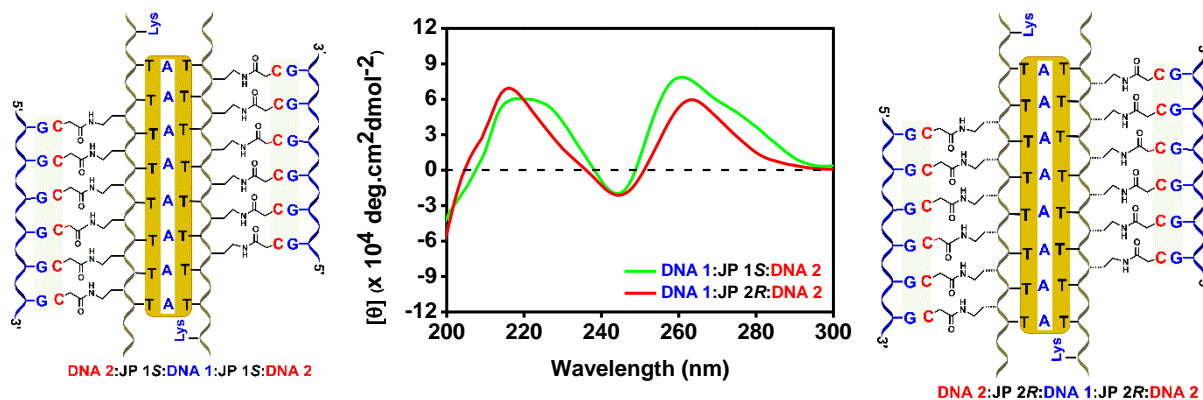


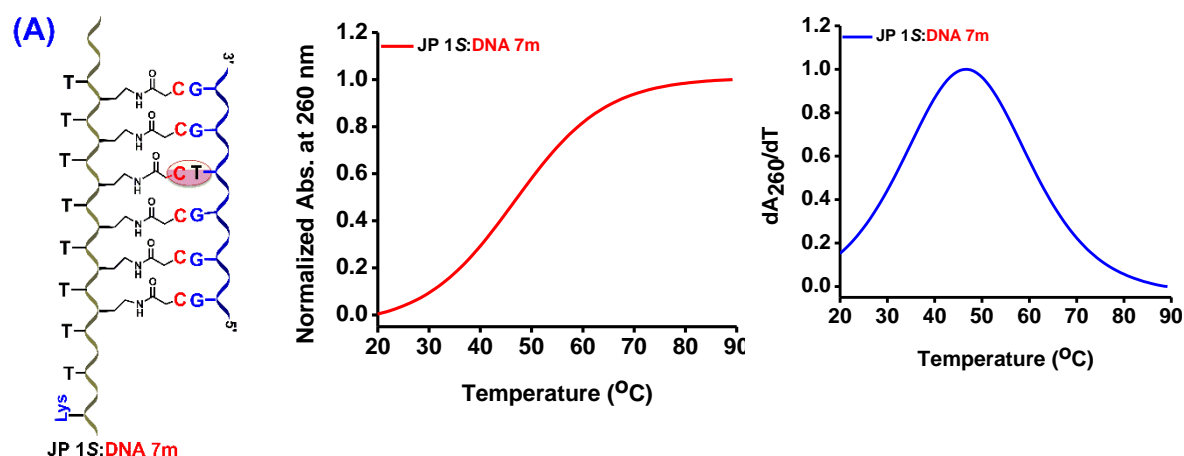
Figure 6.18 CD spectra of duplex of triplex DNA 1:**JP 1S**:DNA 2 ($dG_6:C_6-\gamma S-jpT_8:dA_8:T_{8jp-\gamma S-C_6}:dG_6$) and DNA 1:**JP 2R**:DNA 2 ($dG_6:C_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-C_6}:dG_6$) complex. Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

6.6.2d Thermal stability of duplexes and double duplex from C^γ -Janus PNA duplexes with mismatch DNA

In order to understand the enhanced stability of each face duplex and sequence specificity of interactions, UV absorbance studies were done using DNA 6 (5' GGGTGG 3')

and DNA 7 (5'AAAACAAA3') that have one base mismatch with the Janus PNA sequence. The γ -face duplex **JP 1S:DNA 7** (T_{8jp} - γS - C_6 :DNA 7) has one C-A mismatch and showed a single transition corresponding $T_m = 46.1$ °C (Figure 6.19A), which is lower by 19.7 °C than the T_m observed with matched complementary duplex **JP 1S:DNA 2** (T_{8jp} - γS - C_6 :dG₆, $T_{m1} = 65.8$ °C). The mismatched triplex from t-amide face **JP 1S:DNA 6:JP 1S** (C_6 - γS - jpT_8 :DNA 6: T_{8jp} - γS - C_6) also showed a single transition with T_m of 39.2 °C (Figure 6.19B) with T_m lower by 14.2 °C than the perfect individual duplex.

When both t-amide and γ -face mismatches were present in double duplex of triplex of *Janus* PNA **JP 1S** with **DNA 6** and **DNA 7** (**DNA 7:JP 1S:DNA 6:JP 1S:DNA 7**), only one melting transition with $T_m = 58.5$ °C (Figure 6.19C) was observed, unlike the two transitions seen in perfect double duplex. Compared with their individual isolated duplexes, in double duplex, T_{ms} of both t-amide and γ -face duplexes were enhanced by +19.2 °C and +12.4 °C respectively (Table 6.4). Upon such stabilization, both duplex and triplex they have similar T_{ms} and due to low difference transitions are not resolved and only one T_m is observed. Enhancement of T_{ms} of duplex and triplex similar to that seen in perfect duplexes, substantiated formation of double duplexes in C^γ -*Janus* PNA. Thus destabilization of individual duplexes in presence of mismatches and subsequent enhancement on simultaneous binding of cDNAs on both faces, supports sequence specific complementary base pairing in these C^γ -*Janus* PNAs.



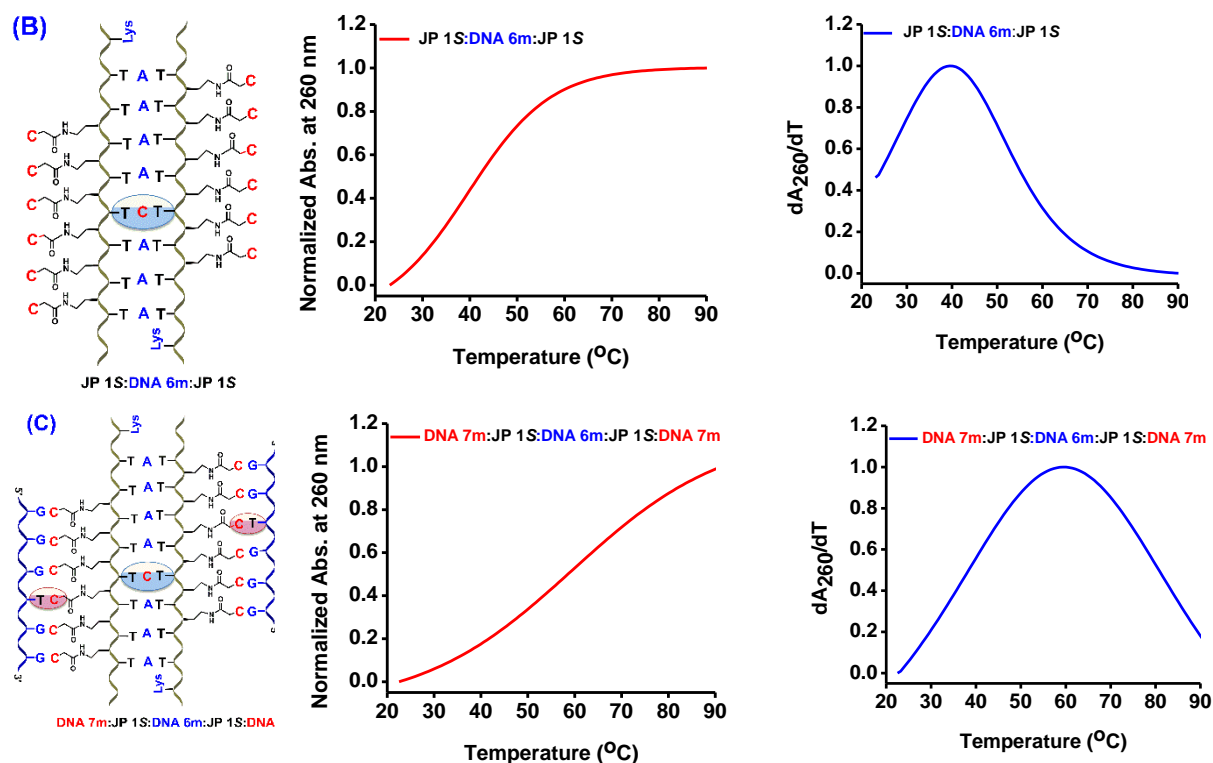


Figure 6.19 Temperature dependent UV absorbance curves for **JP 1S** ($T_{8jp-\gamma S-C_6}$) with mismatch DNA duplex and double duplexes, **DNA 6** (5' GGGTGG 3'), **DNA 7** (5' AAAACAAA 3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) (A) **JP 1S:DNA 7** (B) **DNA 6:JP 1S** (C) **DNA 6:JP 1S:DNA 7**.

Table 6.9 UV - T_m (°C) of **JP 1S** ($T_{8jp-\gamma S-C_6}$) with mismatch **DNA 6** and **DNA 7**

Entry	Janus PNA:mismatchDNA complexes	T_m (°C)		ΔT_m (°C)		
		t-Amide	C^γ	t-Amide	C^γ	C^γ -t-Amide
1	JP 1S:DNA 7m	-	46.1	-	-	+6.9
2	DNA 6m:JP 1S	39.2	-	-	-	
3	DNA 6m:JP 1S:DNA 7m	58.5		+19.2	+12.4	

DNA 6 (5' GGGTGG 3') and **DNA 13** (5' AAAACAAA 3'), **JP 1S** = $T_{8jp-\gamma S-C_6}$, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C.

The $C^\gamma(R)$ homo Janus PNA **JP 2R** ($T_{8jp-\gamma R-C_6}$) was individually hybridized with mismatch **DNA 6** at t-amide face and mismatch **DNA 7** on γ -face. The UV-Temp plots of the complexes are shown in Figure 6.20. The single mismatch duplex **JP 2R:DNA 7** ($T_{8jp-\gamma R-C_6}$: **DNA 7**) gave a T_m of 40.2 °C (Figure 6.20A), and the single mismatch triplex **JP 2R:DNA 6:JP 2** ($T_{8jp-\gamma R-C_6}$:**DNA 6**: $T_{8jp-\gamma R-C_6}$) showed a corresponding of T_m 39.1 °C (Figure 6.20B). The T_m of γ -face duplex **JP 2R:DNA 7** ($T_{8jp-\gamma R-C_6}$:**DNA 7**) was lower by 8.1 °C and

T_m of t-amide face triplex lower T_m by 12.1°C compared to T_{ms} observed for perfect duplexes (Table 6.10). The double duplex of triplex **DNA 6:JP 2R:DNA 7:JP 2R:DNA 6** exhibited single melting transition with of T_m 38.4 °C (Figure 6.20C), which is slightly destabilized by -0.7 °C on t-amide face triplex and -1.8 °C for C γ -face duplex. The decrease in T_m arises from the non-complementary DNAs on both the side destabilization leads to overall lowering of T_m .

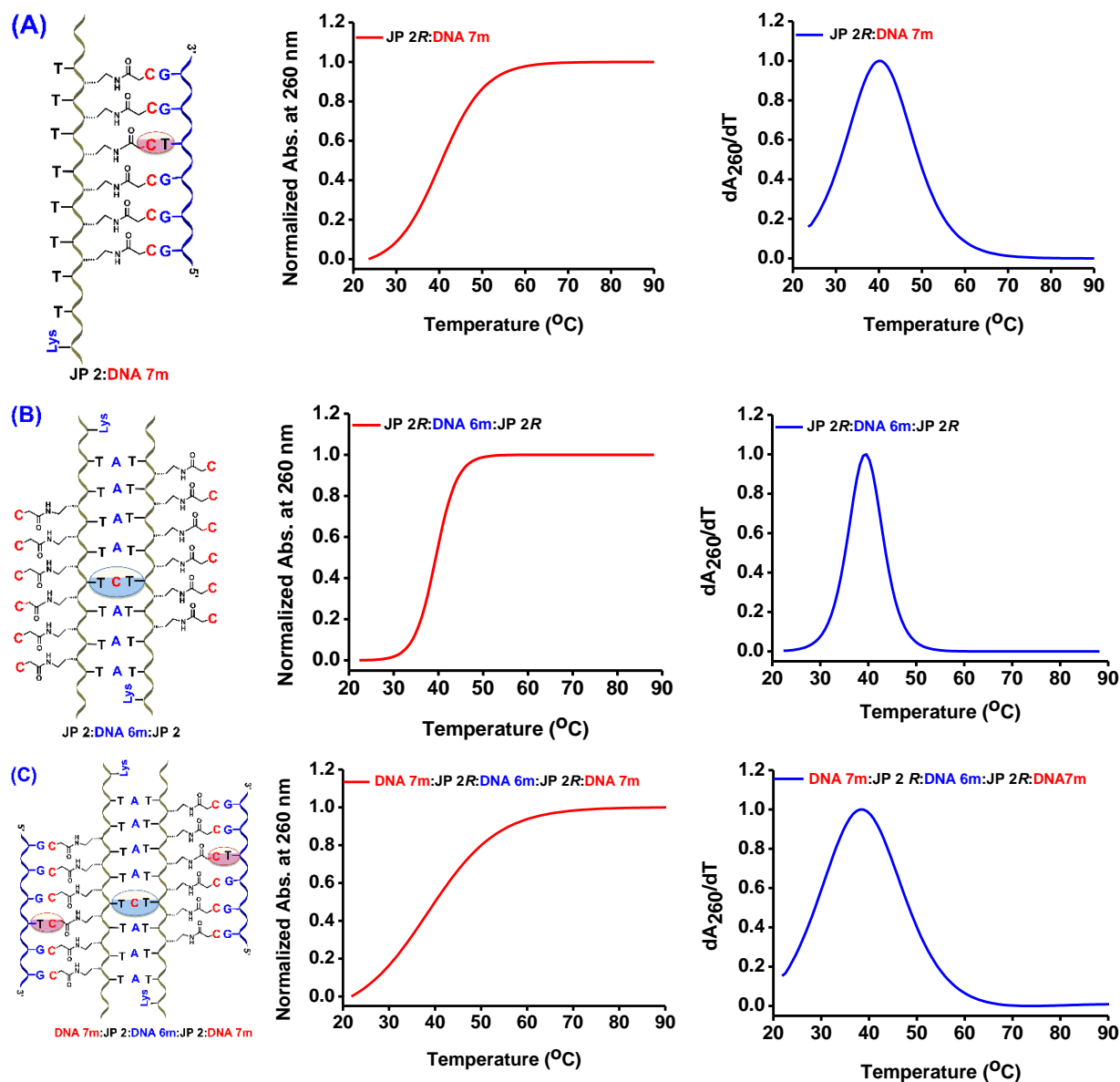


Figure 6.20 Temperature dependent UV absorbance curves for **JP 1R** or *Tsjp- γ R-C₆* with antiparallel mismatch DNA duplex and double duplexes, DNA 6 (5' GGGTGG 3'), DNA 7 (5' AAAACAAA3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) (A) **JP 2R:DNA 7** (B) **DNA 6:JP 2R** (C) **DNA 6:JP 2R:DNA 7**

Table 6.10 UV - T_m (°C) of **JP 2R** ($T_{8jp-\gamma R-C_6}$) with mismatch DNA **6** and DNA **7**

Entry	Janus PNA:mismatchDNA complexes	T_m (°C)		ΔT_m (°C)		
		t-Amide	C $^{\gamma}$	t-Amide	C $^{\gamma}$	C $^{\gamma}$ -t-Amide
1	JP 2R:DNA 7m	-	40.2	-	-	+1.1
2	DNA 6m:JP2R	39.1	-	-	-	
3	DNA 6m:JP 2R:DNA 7m	38.4		-0.7	-1.8	

DNA **6** (5' GGGTGG 3') and DNA **7** (5' AAAACAAA 3'), **JP 2R** = $T_{8jp-\gamma R-C_6}$, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C.

6.6.2e CD spectral study of order of duplex formation

Whether the order of triplex / duplex formation matters to achieve the final observed CD profile of double duplex of triplex was examined by studying CD patterns in a sequential manner:

- (i) *Homo Janus* PNA, **JP 1S** ($T_{8jp-\gamma S-C_6}$, 10 μ M) and **JP 2R** ($T_{8jp-\gamma R-C_6}$, 10 μ M) taken separately in sodium cacodylate buffer and NaCl at pH 7.1 showed weak positive band in the region 240 nm – 260 nm and negative band centered around 270 and 285 nm.
- (ii) They were mixed with stoichiometric amounts of the complementary DNA **2** (dG₆) and kept for equilibration for 10 mins and the CD spectra of duplexes **JP 1S:DNA 2** and **JP 2R** were recorded (Figure 6.21A and 6.21B)
- (iii) This was followed by similar addition of stoichiometric amounts of DNA **1** (dA₈) to (ii), equilibration and recording of CD spectrum of double duplex of triplex DNA **2:JP 1S:DNA 1:JP 1S:DNA 2** (dG₆:C₆- γS -jpT₈:dA₈: $T_{8jp-\gamma S-C_6}$:dG₆) (Figure 6.21 A) and DNA **2:JP 2R:DNA 1:JP 2R:DNA 2** (dG₆:C₆- γR -jpT₈:dA₈: $T_{8jp-\gamma R-C_6}$:dG₆) (Figure 6.21B)
- (iv) The sequential CD experiment was repeated by reversing the order of DNA additions to **JP 1S** and **JP 2R**. First addition of DNA **1** (dA₈) (Figure 6.22), followed by addition of DNA **2** (dG₆) to obtain DNA **2:JP 1S:DNA 1:JP 1S:DNA 2** (dG₆:C₆- γS -jpT₈:dA₈: $T_{8jp-\gamma S-C_6}$:dG₆) (Figure 6.22A) and DNA **2:JP 2R:DNA 1:JP 2R:DNA 2** (dG₆:C₆- γR -jpT₈:dA₈: $T_{8jp-\gamma R-C_6}$:dG₆) (Figure 6.22 B).

It is seen that in both experiments the CD spectra of final products DNA **2:JP 1S:DNA 1:JP 1S:DNA 2** (dG₆:C₆- γS -jpT₈:dA₈: $T_{8jp-\gamma S-C_6}$:dG₆) (Figure 6.23A) and DNA **2:JP**

2R:DNA 1:JP 2R:DNA 2 ($dG_6:C_6-\gamma R-jpT_8:dA_8:T_8jp-\gamma R-C_6:dG_6$) (Figure 6.23B) were identical in terms of spectral bands and intensities. This, experiment indicates that the order in which the duplexes are formed from both faces does not affect the conformational state of final double duplex of triplex.

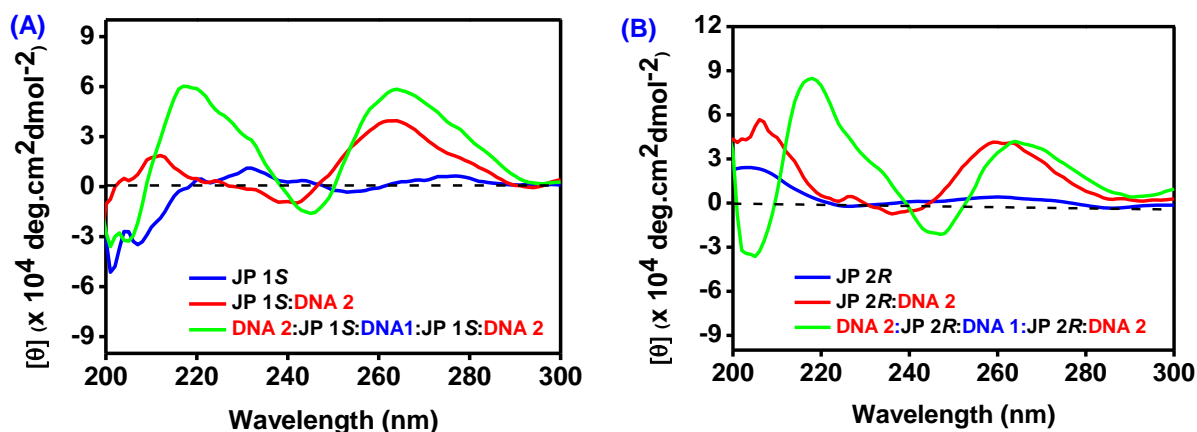


Figure 6.21 CD spectra of (A) *ssJP 1S* ($T_8jp-\gamma S-C_6$), duplex **JP 1S:DNA 2** ($T_8jp-\gamma S-C_6:dG_6$) and duplex of triplex ($dG_6:C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6:dG_6$) (**DNA 2:JP 1S DNA 1:JP 1S:DNA 2**) (B) *ssJP 2R* ($T_8jp-\gamma R-C_6$), triplex **JP 2R:DNA 1** ($T_8jp-\gamma R-C_6:dG_6$) and duplex of triplex **DNA 2:JP 2R:DNA 1:JP 2R:DNA 2** ($dG_6:C_6-\gamma R-jpT_8:dA_8:T_8jp-\gamma R-C_6:dG_6$).

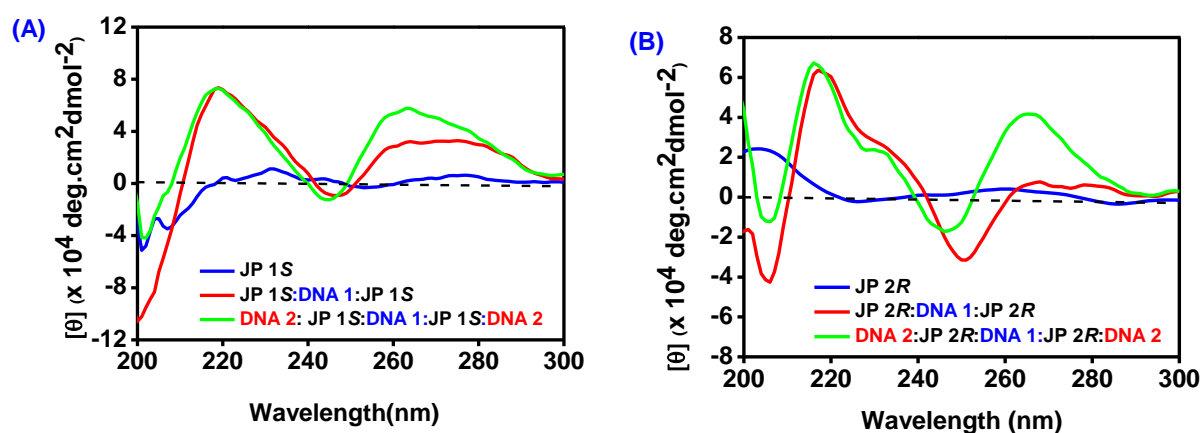


Figure 6.22 CD spectra of (A) *ssJP 1S* ($T_8jp-\gamma S-C_6$), triplex **JP 1S:DNA 1:JP 1S** ($C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6$) and duplex of triplex ($dG_6:C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6:dG_6$) (**DNA 2:JP 1S DNA 1:JP 1S:DNA 2**, 1:1:1) (B) *ssJP 2R* ($T_8jp-\gamma R-C_6$), triplex **JP 2R:DNA 1:JP 2R** ($C_6-\gamma R-jpT_8:dA_8:T_8jp-\gamma R-C_6$) and duplex of triplex **DNA 2:JP 2R:DNA 1:JP 2R:DNA 2** ($dG_6:C_6-\gamma R-jpT_8:dA_8:T_8jp-\gamma R-C_6:dG_6$).

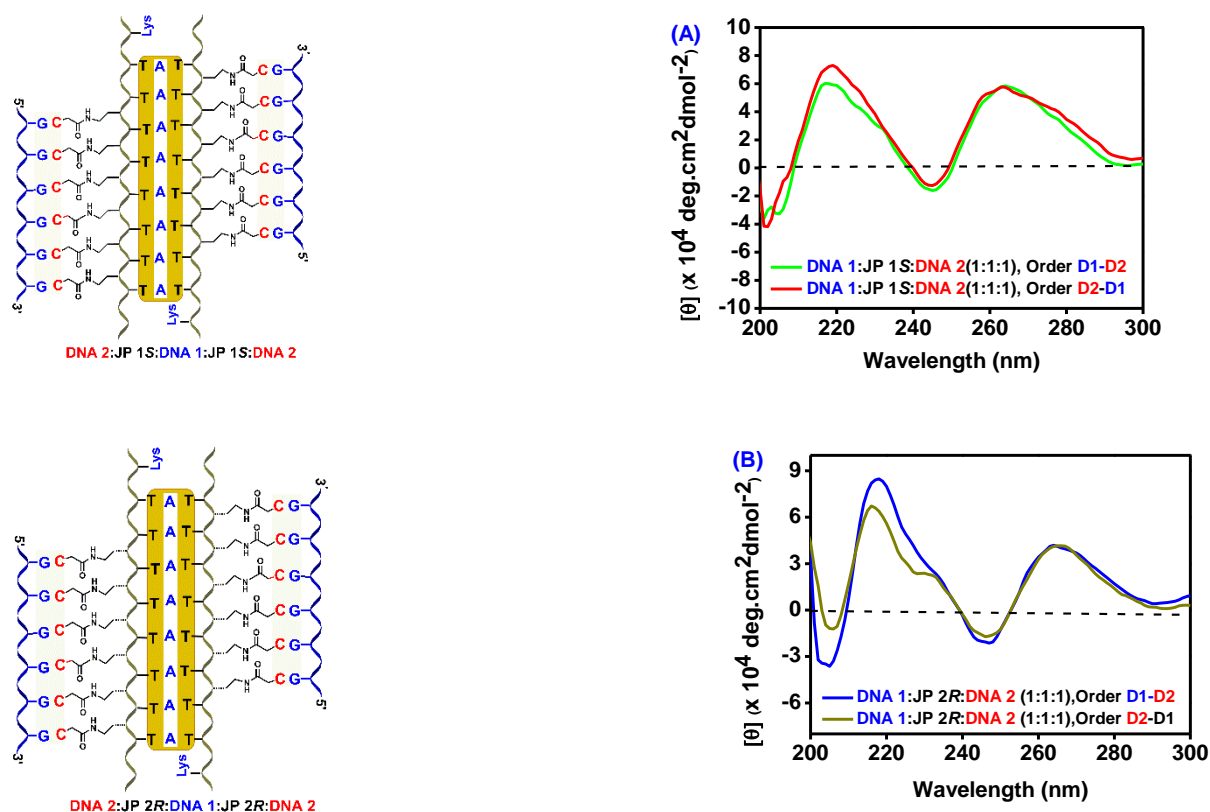


Figure 6.23 (A) Final CD spectra of DNA 1:JP 1S:DNA 2 ($dG_6:C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6:dG_6$) (1:1:1) from Figure 6.21A and 6.22A (B) Final CD spectra of DNA 1:JP 2R:DNA 2 ($dG_6:C_6-\gamma R-jpT_8:dA_8:T_8jp-\gamma R-C_6:dG_6$) (1:1:1) ratio, from figure 6.21B and 6.22 B.

6.6.3 Thermal stability of duplexes and double duplex from Janus PNA JP 3S $T_8jp-\gamma S-G_6$ and JP 4R ($T_8jp-\gamma-G_6$) duplexes with complementary DNA

To further examine the effect of sequence on double duplex formation, homo Janus PNA having a G_6 on γ -face (instead of the C_6 in previous cases) and retaining T_8 on t-amide face were studied. ((JP 3S, $T_8jp-\gamma S-G_6$ and JP 4R, $T_8jp-\gamma R-G_6$). The *homo Janus* PNAs (JP 3S) $T_8jp-\gamma S-G_6$) was hybridized individually with complementary DNAs: γ -face cDNA 3 (dC_6) and amide face with cDNA 1 (dA_8). The γ -face duplex JP 3S:DNA 3 ($T_8jp-\gamma S-G_6:dC_6$) showed a single transition corresponding $T_m = 43.2$ °C (Figure 6.24A). The triplex from t-amide face DNA1:JP 3S:DNA 1 ($G_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-G_6$) also showed a single transition with T_m of 45.3 °C (Figure 6.24B). In presence of both face complementary DNAs, DNA 2 and DNA 3, the *homo Janus* PNA formed double duplex of triplex DNA 3:JP 3S:DNA 1:JP 3S:DNA 3 ($dC_6:G_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-G_6:dC_6$) that exhibited two melting transitions with $T_{m1} = 40.1$ °C and $T_{m2} = 73.3$ °C (Figure 6.24C) similar to that observed previously with DNA 2:JP 1S DNA 1:JP 1S:DNA 2 ($dG_6:C_6-\gamma S-jpT_8:dA_8:T_8jp-\gamma S-C_6:dG_6$). The T_m of t-amide face duplex was increased by +3.1°C while that from γ -face was elevated by +28 °C in

double duplexes compared with their individual isolated duplexes with **JP 3S** ($T_8jp-\gamma S-G_6$) (Table 6.11, entry 3). This again supported the synergistic stabilizing effects mutually exerted by each face duplex on the other duplex. It is again noticed that the stability of the γ -face duplex was higher than that of the t -amide face duplex.

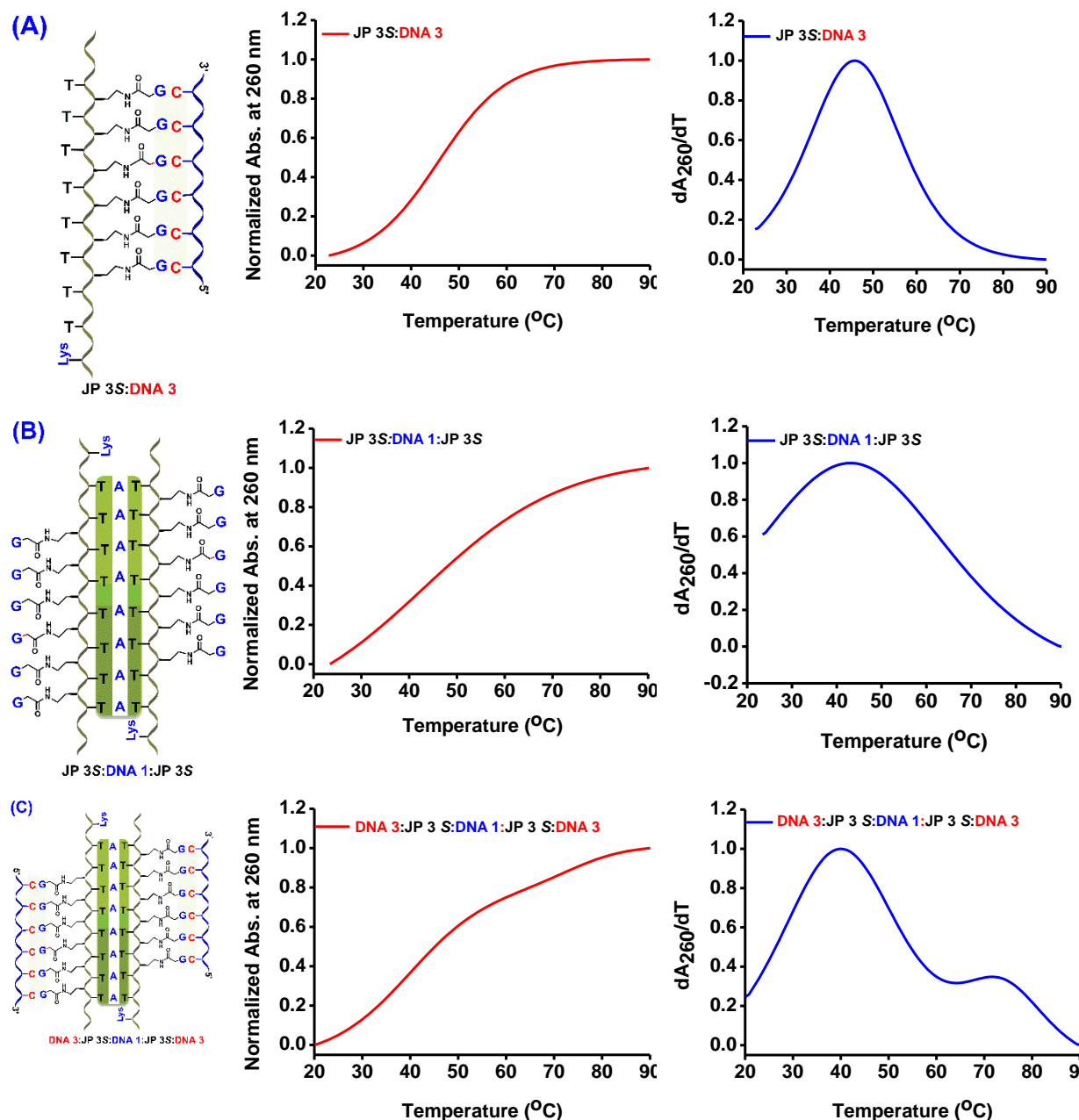


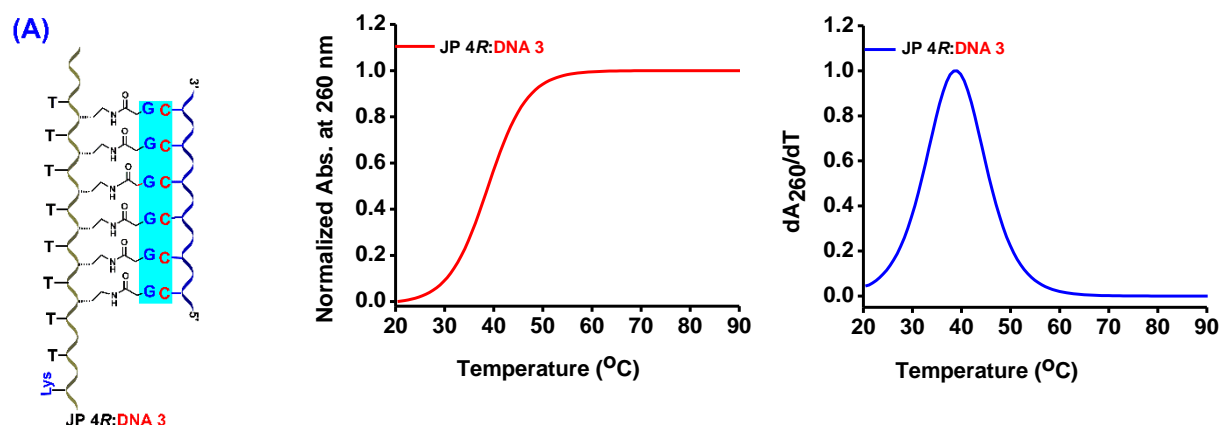
Figure 6.24 Temperature dependent UV absorbance curves for DNA complexes of **JP 3S** ($T_8jp-\gamma S-G_6$) with DNA 1 ($dA_8 = 5' \text{ AAAAAAAAAA } 3'$) and DNA 3 ($dC_6 = 5' \text{ CCCCCC } 3'$); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (A) **JP 3S:DNA 3** ($T_8jp-\gamma S-G_6:dC_6$) (B) **JP 3S:DNA 1:JP 3S** ($G_6-\gamma S.jpT_8:dA_8:T_8jp-\gamma S-G_6$) (C) **DNA 3:JP 3S:DNA 1:JP 3S:DNA 3** ($dC_6:G_6-\gamma S.jpT_8:dA_8:T_8jp-\gamma S-G_6:dC_6$)

Table 6.11 UV - T_m ($^{\circ}\text{C}$) of **JP 3S** ($\text{T}_8\text{jp}-\gamma\text{S}-\text{G}_6$) with complementary DNA **1** and DNA **3**.

Entry	<i>HomoJanus</i> PNA:DNA complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)	
		t-Amide	C^{γ}	t-Amide	C^{γ}
1	JP 3S:DNA 3 ($\text{T}_8\text{jp}-\gamma\text{S}-\text{G}_6:\text{dC}_6$)	-	45.3	-	-
2	JP 3S:DNA 1:JP 3S ($\text{G}_6-\gamma\text{S}-\text{jpT}_8:\text{dA}_8:\text{T}_8\text{jp}-\gamma-\text{G}_6$)	43.2	-	-	-
3	DNA 3:JP 3S:DNA 1:JP 3S:DNA 3 ($\text{dC}_6:\text{G}_6-\gamma\text{S}-\text{jpT}_8:\text{dA}_8:\text{T}_8\text{jp}-\gamma\text{S}-\text{G}_6:\text{dC}_6$)	40.1	73.3	+3.1	+28

DNA **1** ($\text{dA}_8 = 5'-\text{AAAAAAAA}-3'$) and DNA **3** ($\text{dC}_6 = 5'-\text{CCCCCC}-3'$), **JP 3S** = $\text{T}_8\text{jp}-\gamma\text{S}-\text{G}_6$, ΔT_m indicates the difference in T_m s of single duplex with each face *with that corresponding in double duplex*. T_m values are accurate to ± 1.0 $^{\circ}\text{C}$.

The stereoisomeric *homo Janus* PNA **JP 4R** ($\text{T}_8\text{jp}-\gamma\text{R}-\text{G}_6$) was also studied similarly for its complexation with corresponding cDNAs and the UV- T_m hybridization data is shown in Figure 6.25 and T_m s in Table 6.12. The γ -face duplex **JP 4R:DNA 3** and the t-amide face triplex **JP 4R:DNA 1:JP 4R** both showed single transitions corresponding to T_m s of 39.5 $^{\circ}\text{C}$ and $T_m = 45.5$ $^{\circ}\text{C}$, respectively, with the γ -face duplex having less stability than the t-amide face triplex. The double duplex of triplex DNA **3:JP 4R:DNA 1:JP 4R:DNA 3** displayed two transitions with $T_{m1} = 37.8$ $^{\circ}\text{C}$ for t-amide face duplex and $T_{m2} = 54.8$ $^{\circ}\text{C}$ for C^{γ} face duplex. The thermal stability of t-amide face duplex slightly decreased by -9.7°C and the stability of γ -face duplex was enhanced by $+15^{\circ}\text{C}$ compared to T_m s of individual duplexes (Table 6.31, entry 3). As in previous cases, the γ -face showed higher stabilization than individual duplexes and demonstrating that the relative stabilization of two stereoisomeric *Janus* PNAs are different, with *S*-Janus PNA **JP 3S** being more stable than *R*-Janus PNA **JP 4R**.



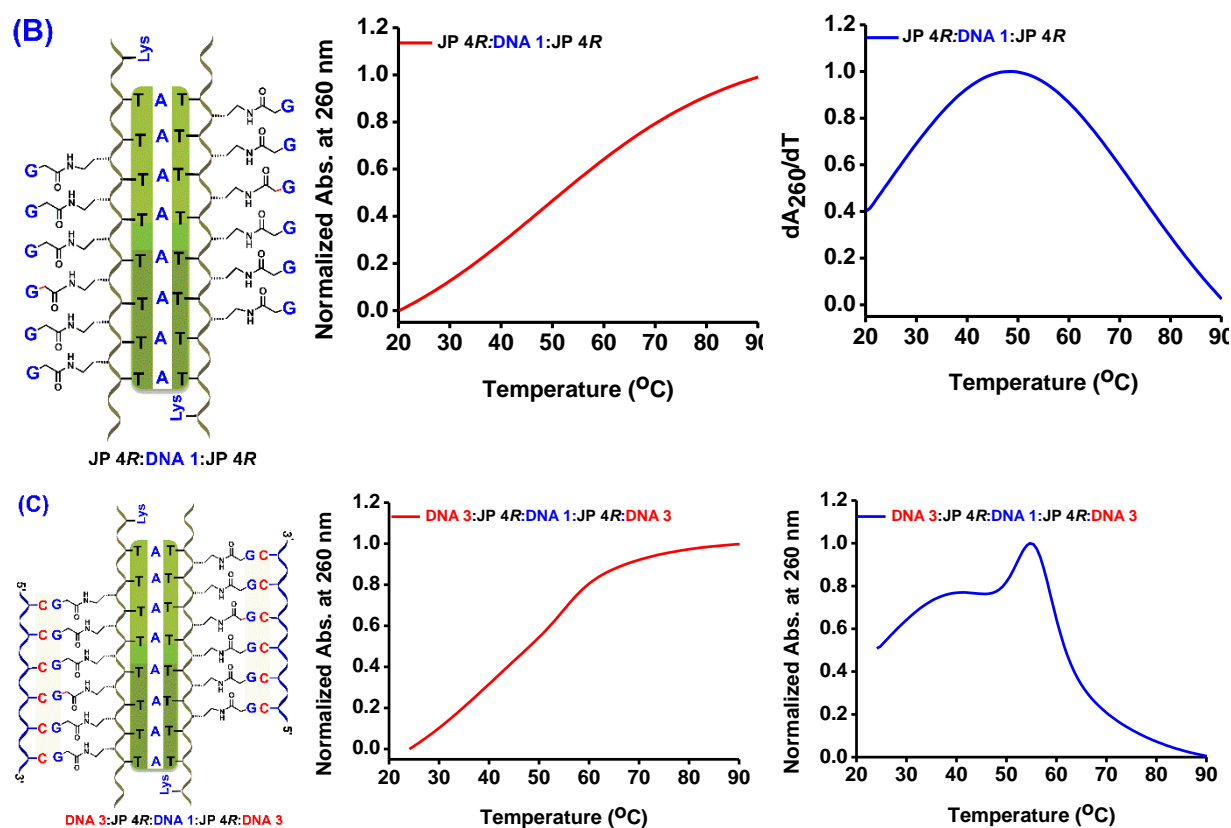


Figure 6.25 Temperature dependent UV absorbance curves for DNA complexes of **JP 4R** ($T_{8jp-\gamma R-G_6}$) with DNA **1** ($dA_8 = 5' \text{ AAAAAAAAAA } 3'$) and DNA **3** ($dC_6 = 5' \text{ CCCCCC } 3'$); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (A) DNA **1:JP 4R** ($dA_8:T_{8jp-\gamma R-G_6}$) (B) **JP 4R:DNA 1:JP 4R** ($G_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-G_6}$) (C) DNA **3:JP 4R:DNA 1:JP 4R:DNA 3** ($dC_6:G_6-\gamma S.jpT_8:dA_8:T_{8jp-\gamma S-G_6}:dC_6$)

Table 6.12 UV - T_m ($^{\circ}\text{C}$) of **JP 4R** ($T_{8jp-\gamma R-G_6}$) with complementary DNA **1** and DNA **3**

Entry	<i>HomoJanus</i> PNA:DNA complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)	
		t-Amide	C^{γ}	t-Amide	C^{γ}
1	JP 4R:DNA 3 ($T_{8jp-\gamma R-G_6}:dC_6$)	-	39.5	-	-
2	JP 4R:DNA 1:JP 4R ($G_6-\gamma R-jpT_8:dA_8:T_{8jp-\gamma R-G_6}$)	47.5	-	-	-
3	DNA 3:JP 4R:DNA 1:JP 4R:DNA 3 ($dA_8:T_{8jp-\gamma R-G_6}:dC_6$)	37.8	54.8	-9.7	+15

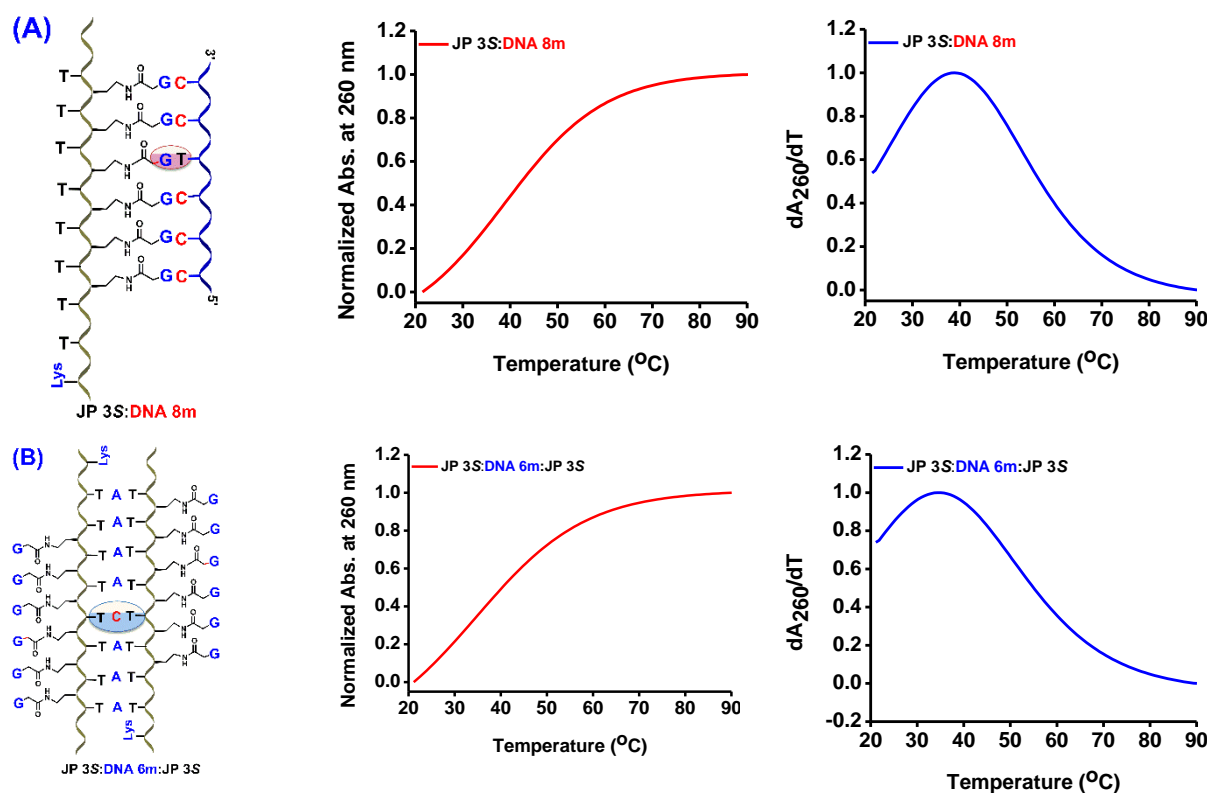
DNA **1** ($dA_8 = 5' \text{ AAAAAAAAAA } 3'$) and DNA **3** ($dC_6 = 5' \text{ CCCCCC } 3'$), **JP 4R** = $T_{8jp-\gamma R-G_6}$, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 $^{\circ}\text{C}$.

6.6.3a Thermal stability complexes from Janus PNA **JP 3S** and **JP 4R** with mismatch DNA

In order to establish the sequence specificity of binding, the thermal stability of duplexes and double duplexes from homo C^{γ} -Janus PNAs **JP 3S** ($T_{8jp-\gamma S-G_6}$) and **JP 4R**

($T_{8jp-\gamma R-G_6}$) with cDNA having single mismatch were studied. The melting studies were done using DNA **8m** (5' CCCTCC 3') that has a single site mismatched T for γ -face duplex **JP 3S:DNA 6m** ($T_{8jp-\gamma S-G_6}$:DNA **8m**). This duplex with a T-G mismatch showed a single transition corresponding $T_m = 38.3$ °C (Figure 6.26A), which is lower by 7 °C than the T_m observed with perfect complementary duplex **JP 3S:DNA 3** ($T_{8jp-\gamma S-G_6}$:dC₆, $T_{m1} = 45.3$ °C). The mismatched triplex from t-amide face **JP 3S:DNA 6m:JP 3S** ($T_{8jp-\gamma S-G_6}$:DNA **6m**: $T_{8jp-\gamma S-G_6}$) also showed a single sigmoidal transition with T_m of 34.3 °C (Figure 6.26B), which is lower by 8.9 °C compared to corresponding perfect triplex **JP 3S:DNA 3:JP 3S**.

In presence of both face mismatch DNAs (DNA **6m** and DNA **8m**), the C^γ -Janus PNA **JP 3S** formed double duplex of triplex DNA **8m:JP 3S:DNA 6m:JP 3S:DNA 8m** and exhibited a single melting transition with $T_m = 55.5$ °C (Figure 6.26C). Compared with their individual isolated duplexes, the T_m s of both t-amide and C^γ duplexes in the double duplex were enhanced by +21.3 °C and +17.2 °C respectively and due to overlapping of T_m s, a single transition is observed. The initial destabilization of individual duplex and triplex and their stability enhancement, upon forming double duplex of triplex substantiated the formation of complexes in C^γ -Janus PNA by sequence specific complementary base pairing.



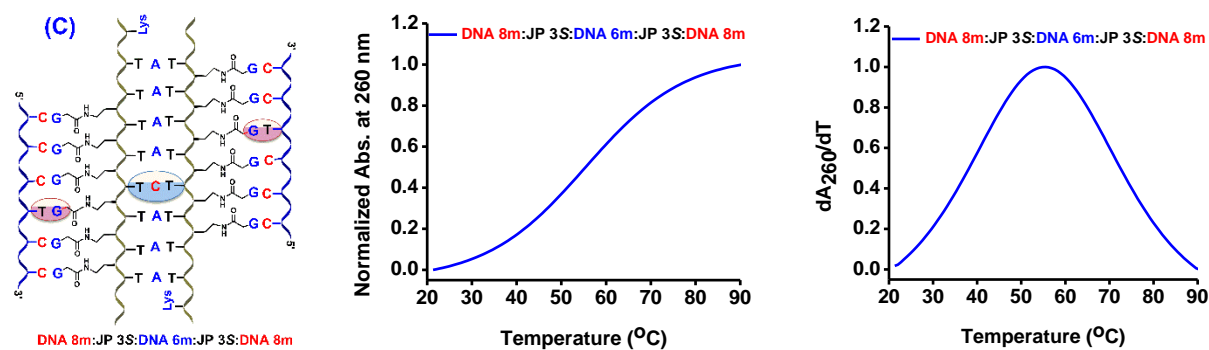


Figure 6.26 Temperature dependent UV absorbance profiles for **JP 3S** ($T_{8jp-\gamma S-C_6}$) with mismatch DNA duplex, triplex and double duplex of triplex. DNA **6m** (5' CCCTCC 3'), DNA **8m** (5' AAAACAAA 3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) (A) duplex **JP 3S:DNA 8m** (B) triplex **JP 3S:DNA 6m:JP 3S** (C) duplex of triplex DNA **8m:JP 3S:DNA 6m:JP 3S:DNA 8m**.

Table 6.13 UV - T_m (°C) of **JP 3S** ($T_{8jp-\gamma S-G_6}$) with complementary DNA **6** and DNA **8**.

Entry	Janus PNA:mismatchDNA complexes	T_m (°C)		ΔT_m (°C)	
		t-Amide	C'	t-Amide	C'
1	JP 3S:DNA 8m	-	38.3	-	-
2	JP 3S:DNA 6m:JP 3S	34.3	-	-	-
3	DNA 8m:JP 3S:DNA 6m:JP 3S:DNA 8m	55.5		+21.3	+17.2

DNA **6m** (5' CCCTCC 3') and DNA **8m** (5' AAAACAAA 3'), **JP 3S** = $T_{8jp-\gamma S-G_6}$, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C.

Similar experiments were done with stereomeric *homo Janus* PNA **JP 4R** ($T_{8jp-\gamma R-G_6}$) by hybridization with mismatch DNA **6m** at t-amide face and mismatch DNA **8m** with γ -face and the UV- T plots are shown in Figure 6.33. The duplex **JP 4R:DNA 8** ($T_{8jp-\gamma R-G_6}$:DNA 8m) gave T_m 30.8 °C (Figure 6.26A) and the t-amide face triplex **JP 4R:DNA 6m:JP 4R** (DNA 6m: $T_{8jp-\gamma R-G_6}$) showed T_m of 37.4 °C (Figure 6.26B). The T_m of γ -face duplex **JP 4R:DNA 8m** ($T_{8jp-\gamma R-G_6}$:DNA 8m) was lower by 8.7 °C and the T_m of t-amide face duplex was lower by 8 °C compared to the T_m s seen for corresponding perfect complexes. The double duplex of triplex DNA **8m:JP 4R:DNA 6m:JP 4R:DNA 8m** exhibited single melting transition with T_m of 45.9 °C (Figure 6.27C), which is enhanced by +15.1 °C for t-amide face triplex and +8.5 °C for γ -face duplex. The T_m s of both complexes in composite double duplex of triplex are higher, similar to that seen in perfect duplexes. The initial destabilization of individual duplex and triplex and their stability enhancement, upon forming double duplex of

triplex substantiated the formation of complexes in C^{γ} -Janus PNA by sequence specific complementary base pairing.

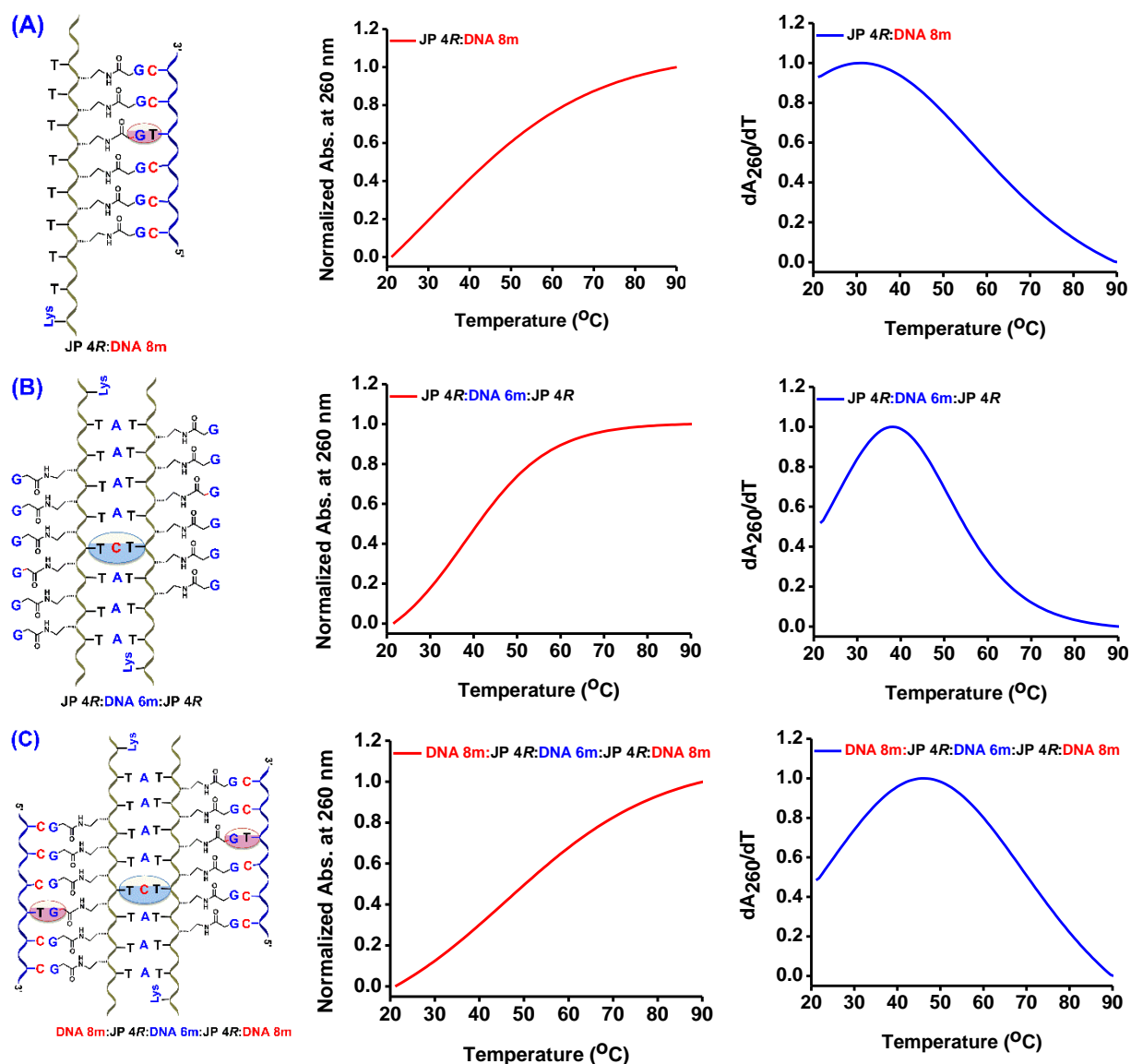


Figure 6.27 Temperature dependent UV absorbance curves for **JP 4** or $R Tsjp-\gamma C_6$ with antiparallel mismatch DNA duplex and double duplexes, DNA **6m** (5' CCCTCC 3'), DNA **8m** (5' AAAACAAA3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) (A) **JP 4R:DNA 8m** (B) **JP 4R:DNA 6m:JP 4R** (C) **DNA 8m:JP 4R:DNA 6m:JP 4R:DNA 8m**

Table 6.14 UV - T_m ($^{\circ}\text{C}$) of **JP 4R** ($Tsjp-\gamma R-G_6$) duplexes with mismatch DNA **6m** and DNA **8m**

Entry	Janus PNA:mismatchDNA complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)	
		t-Amide	C^{γ}	t-Amide	C^{γ}
1	JP 4R:DNA 8m	-	30.8	-	-
2	JP 4R:DNA 6m:JP 4R	37.5	-	-	-
3	DNA 8m:JP 4R:DNA 6m:JP 4R:DNA 8m	45.9		+8.5	+15.1

DNA 6 (5' CCCTCC 3') and DNA 8 (5' AAAACAAA 3'), **JP 4R**= $T_{8jp-\gamma R-G_6}$, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C.

6.6.3b CD spectra of Janus PNA **JP 3S** ($T_{8jp-\gamma S-G_6}$) and **JP 4R** ($T_{8jp-\gamma R-G_6}$) with cDNA

The CD spectra of the single stranded C^γ -Janus PNA, DNA and the derived t-amide and C^γ face complexes of Janus PNA, **JP 3S** ($T_{8jp-\gamma S-C_6}$) and **JP 4R** ($T_{8jp-\gamma R-G_6}$) are shown in (Figure 6.28). The t-amide face triplexes **JP 3S:DNA 1:JP 3S** ($T_{8jp-\gamma S-G_6}$:DNA 1: $T_{8jp-\gamma S-G_6}$) and **JP 4R: DNA 1:JP 4R** ($T_{8jp-\gamma R-G_6}$:DNA 1: $T_{8jp-\gamma R-G_6}$) a positive maxima at 270 nm-275 nm and a negative maxima at 245 nm- 250 nm with cross-over points at 234 and 259 nm and the C^γ face duplex **JP 3S:DNA 2** ($T_{8jp-\gamma S-G_6}$:DNA 2) and **JP 4R** ($T_{8jp-\gamma R-G_6}$) shows a positive maxima at 272nm -276 nm similar to C^γ -face of **JP 1S** and **JP 2R** (Figure 6.28). The **JP 3S** ($T_{8jp-\gamma S-G_6}$:dC₆) shows a negative maxima at 248 nm with cross-over points at 258 nm and **JP 4R** ($T_{8jp-\gamma R-G_6}$:dC₆) shows a negative maxima at 226 nm with cross over point 233 nm.

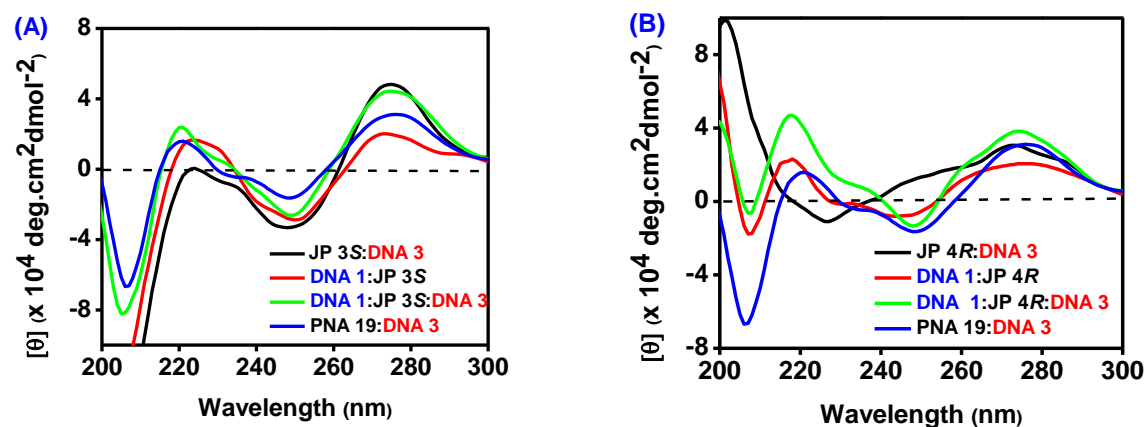


Figure 6.28 CD spectra of (A) γ -face duplex **JP 3S:DNA 3** ($T_{8jp-\gamma S-G_6}$:dC₆), t-amide face triplex DNA 1:**JP 3S** (:dA₈: $T_{8jp-\gamma S-G_6}$) duplex of duplex DNA 1:**JP 3S:DNA 3** (dA₈: $T_{8jp-\gamma S-G_6}$:dC₆) (B) γ -face duplex **JP 4R:DNA 3** ($T_{8jp-\gamma R-G_6}$:dC₆), t-amide face triplex DNA 1:**JP 4R** (dA₈: $T_{8jp-\gamma R-G_6}$) duplex of duplex DNA 1:**JP 4R:DNA 3** (dA₈: $T_{8jp-\gamma R-G_6}$:dC₆), and *aeg*PNA DNA duplex, Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

The *homo Janus* double duplex DNA 1:**JP 3S:DNA 3** (DNA 1: $T_{8jp-\gamma S-G_6}$:dC₆) shows a positive maxima at 275 nm and a negative band with maxima at 249 nm with cross-over points at 236 and 258 nm where as DNA 1:**JP 4R:DNA 3** (DNA 1: $T_{8jp-\gamma S-G_6}$:dC₆) shows positive maxima at 273 nm and a negative band with maxima at 247 nm with cross-over points at 239 and 254 nm (6.29A and 6.29B). In comparison with the CD spectrum of DNA complex of Janus PNA **JP 1S** ($T_{8jp-\gamma S-C_6}$) and **JP 2R** ($T_{8jp-\gamma R-C_6}$) with both amide and C^γ

faces using appropriate complementary DNAs, DNA 1 (dA₈) and DNA 2 (dG₆). As shown in Figure 6.29, the duplex of triplex DNA 1:JP 1S:DNA 2 (dG₆:C₆-S γ -jpT₈:dA₈:T₈jp- γ S-C₆:dG₆) and DNA 1:JP 2R:DNA 3 (dG₆:C₆-R γ -jpT₈:dA₈:T₈jp- γ R-C₆:dG₆) shows a positive maxima in between the 253 nm to 285 nm region and a negative maxima in 238 nm – 251 nm region.

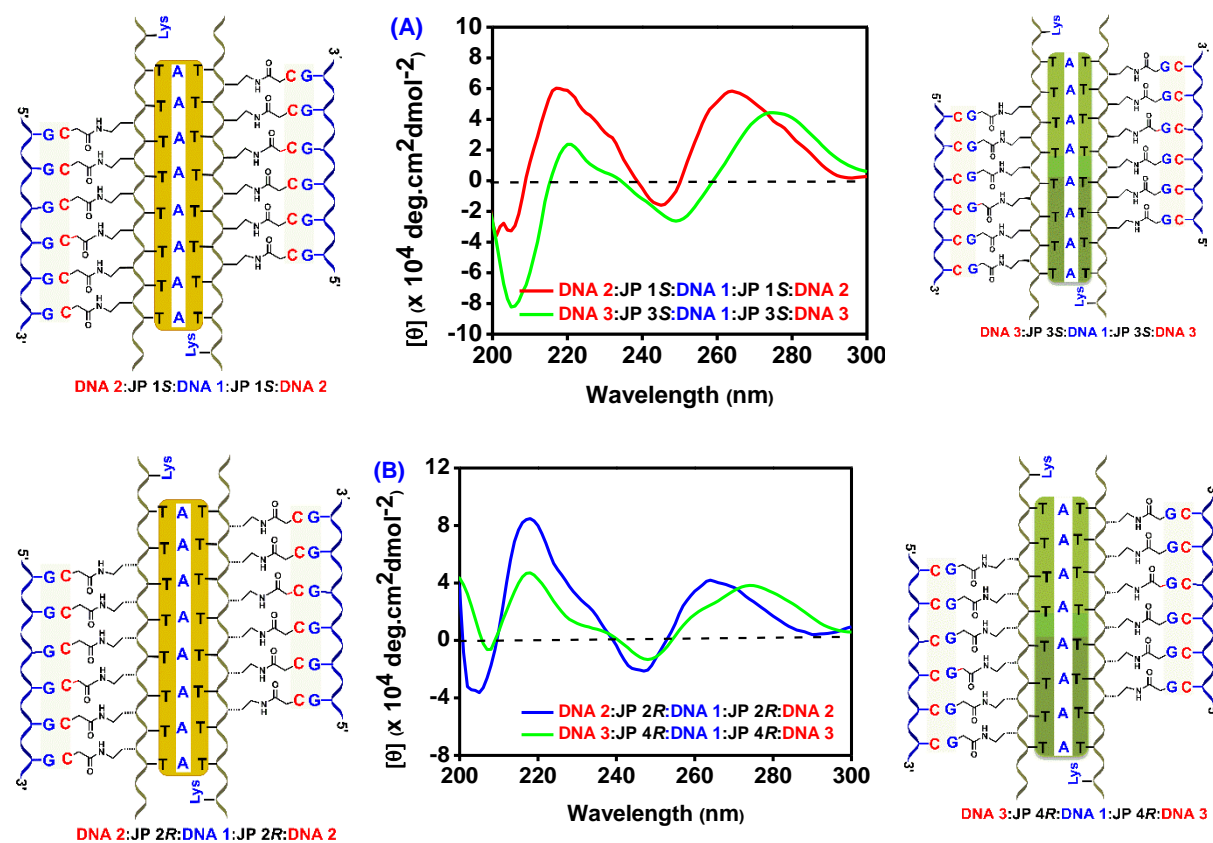


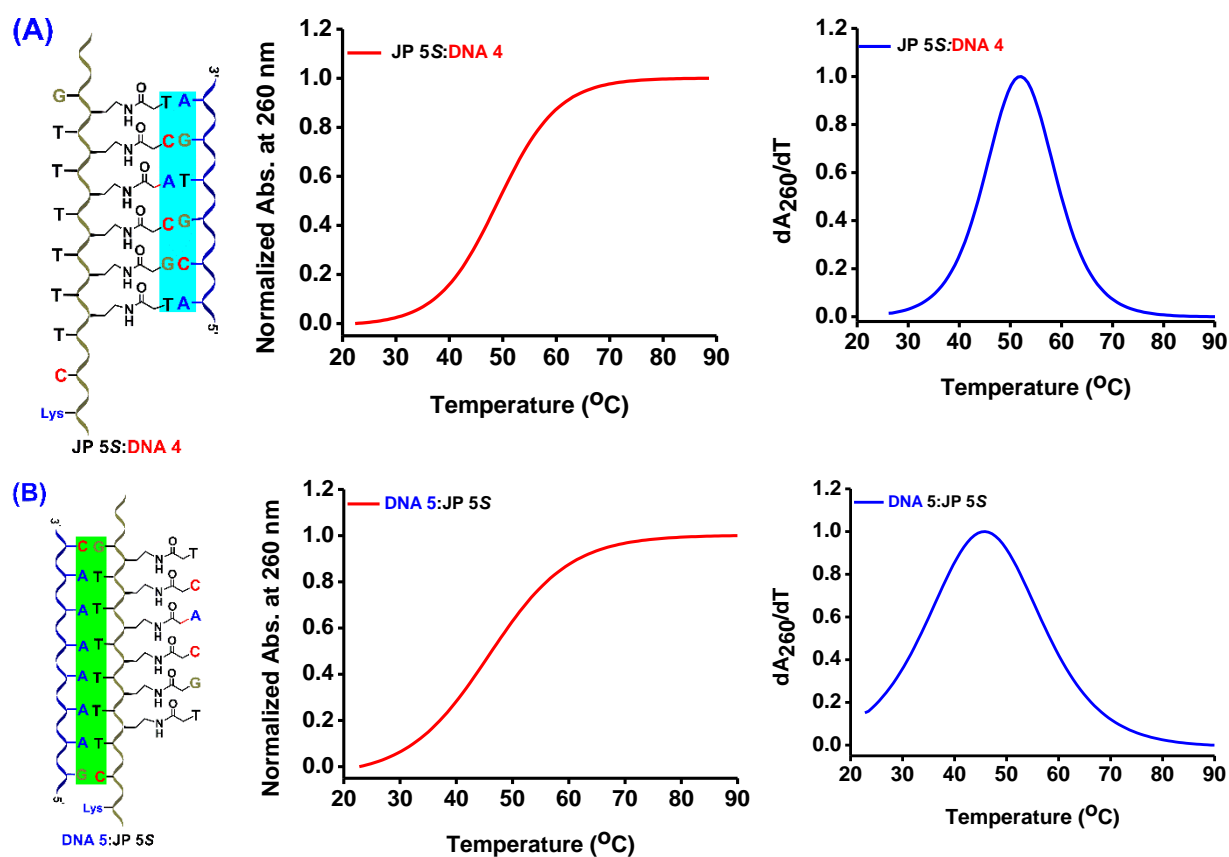
Figure 6.29 CD spectra of (A) duplex of triplex DNA 1:JP 1S:DNA 2 (dG₆:C₆- γ S-jpT₈:dA₈:T₈jp- γ S-C₆:dG₆) vs DNA 1:JP 3S:DNA 3 (dC₆:G₆- γ S-jpT₈:dA₈:T₈jp- γ S-G₆:dC₆) complex. (B) duplex of triplex DNA 1:JP 2R:DNA 2 (dG₆:C₆- γ R-jpT₈:dA₈:T₈jp- γ R-C₆:dG₆) vs DNA 1:JP 4R:DNA 3 (dC₆:G₆- γ R-jpT₈:dA₈:T₈jp- γ R-G₆:dC₆) complex. (Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 Mm).

6.6.4 Thermal stability of duplexes and double duplex from hetero Janus PNA JP 5S (*mjp*- γ S-PNA) and JP 6R (*mjp*- γ R-PNA) with cDNA

The sequences of *Janus* PNAs on t-amide and C^γ-faces described so far are homooligomeric sequences of T and C/G and they form triplexes and sometimes higher order structures. To generate only double duplexes, mixed sequences on both sides are ideal and formation of such structures were examined from C^γ-*Janus* PNA sequences JP 5S (*mjp*- γ S-PNA) and *Janus* PNA JP 6R (*mjp*- γ R-PNA) which have mixed base sequence on t-amide

face. These stereomeric PNAs are derived from 8-mer mixed base sequence on t-amide face and 6-mer mixed base sequence on C γ -face and were separately hybridised with individual cDNA sequences from both sides to generate respective duplexes and simultaneously hybridized with cDNAs to generate double duplexes. (Figure 6.30).

The UV-*T* hybridization data of the *mix Janus* PNA **JP 5S** (*mjp- γ S-PNA*) with cDNA **4** is shown in Figure 6.30 and Table 6.15. The C γ -face duplex **JP 5S:DNA 4** and the t-amide face duplex **DNA 5:JP 5S** showed well defined single transitions corresponding to T_{ms} of 51.7 °C and $T_m = 45.5$ °C, respectively. The results again indicated that the C γ -face duplex has higher stability (+6.2 °C) compared to t-amide duplex. The double duplex from mixed sequence *Janus* PNA **DNA 5:JP 5S:DNA 4** displayed two transitions with $T_{m1} = 52.7$ °C for t-amide face duplex and $T_{m2} = 73.9$ °C for C γ -face duplex. The thermal stability of both duplexes is enhanced in double duplex complex, compared to individual duplex alone by +7.2 °C for t-amide face and +22.2 °C for C γ -face duplex. (Tabl 6.15, entry 3). As in previous cases, the C γ -face showed higher stabilization than amide face by +6.2 °C in mono duplex form and +21.2 °C in double duplex state.



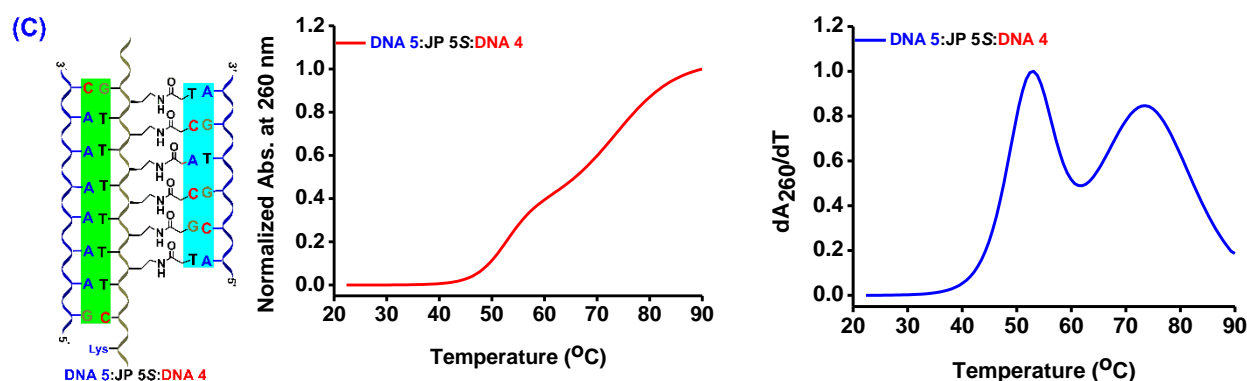


Figure 6.30 Temperature dependent UV absorbance curves for DNA complexes of **JP 5S** (*mjp-γS-PNA*) with DNA **5** (5'-CAAAAAAG-3') and DNA **4** (5'-ACGTGA-3'); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (A) duplex **JP 5S:DNA 4** (*mjp-γS-PNA:DNA 4*) (B) duplex DNA **5:JP 5S** (DNA **5**:*mjp-γS-PNA*) (C) double duplex DNA **5:JP 5S:DNA 4** (DNA:*mjp-γS-PNA:DNA*).

Table 6.15 UV - T_m (°C) of **JP 5** (*S mjp-γ-PNA*) with complimentary DNA **4** and DNA **5**

Entry	Mix Janus PNA:cDNA complexes	T_m (°C)		ΔT_m (°C)		
		<i>t</i> -Amide	C^γ	<i>t</i> -Amide	C^γ	C^γ - <i>t</i> -Amide
1	JP 5S:DNA 4 (<i>mjp-γS-PNA:DNA 4</i>)	-	51.7	-	-	+6.2
2	DNA 5:JP 5S (DNA 5 : <i>mjp-γS-PNA</i>)	45.5	-	-	-	
3	DNA 5:JP 5S:DNA 4 (cDNA 5 : <i>mjp-γS-PNA:DNA 4)</i>	52.7	73.9	+7.2	+22.2	+21.2

DNA **5** (5'-CAAAAAAG-3') and DNA **4** (cDNA = 5'- GTGATCT-3'), **JP 5S** = *mjp-γS-PNA*, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C

The stereomeric C^γ -mix Janus PNA **JP 6R** (*mjp-γR-PNA*) hybridized with cDNA **4** from C^γ -face **JP 6R:DNA 4** (*mjp-γR-PNA:DNA 4*) showed duplex melting with single transition with $T_m=38.1$ °C (Figure 6.31A) and duplex from the *t*-amide face DNA **5:JP 6R** (DNA **5**:*mjp-γR-PNA*) melting transition at $T_m =33.4$ °C, (Figure 6.31B) with a difference in T_m of + 4.7 °C between the two faces. However, the double duplex DNA **5:JP 6R:DNA 4** (DNA **5**:*mjp-γR-PNA:DNA 4*) displayed two transitions (Figure 6.31C) with a $T_{m1} = 38.7$ °C for *t*-amide face duplex and $T_{m2} = 69.8$ °C for C^γ -face duplex (Table 6.16, entry 3). These T_m s which is higher than the T_m of each of the individual duplexes. Thus duplexes from both *t*-amide and C^γ -face are stabilized, with the *t*-amide face duplex by +5.3 °C and the C^γ -face duplex to a higher extent by +31.7 °C.

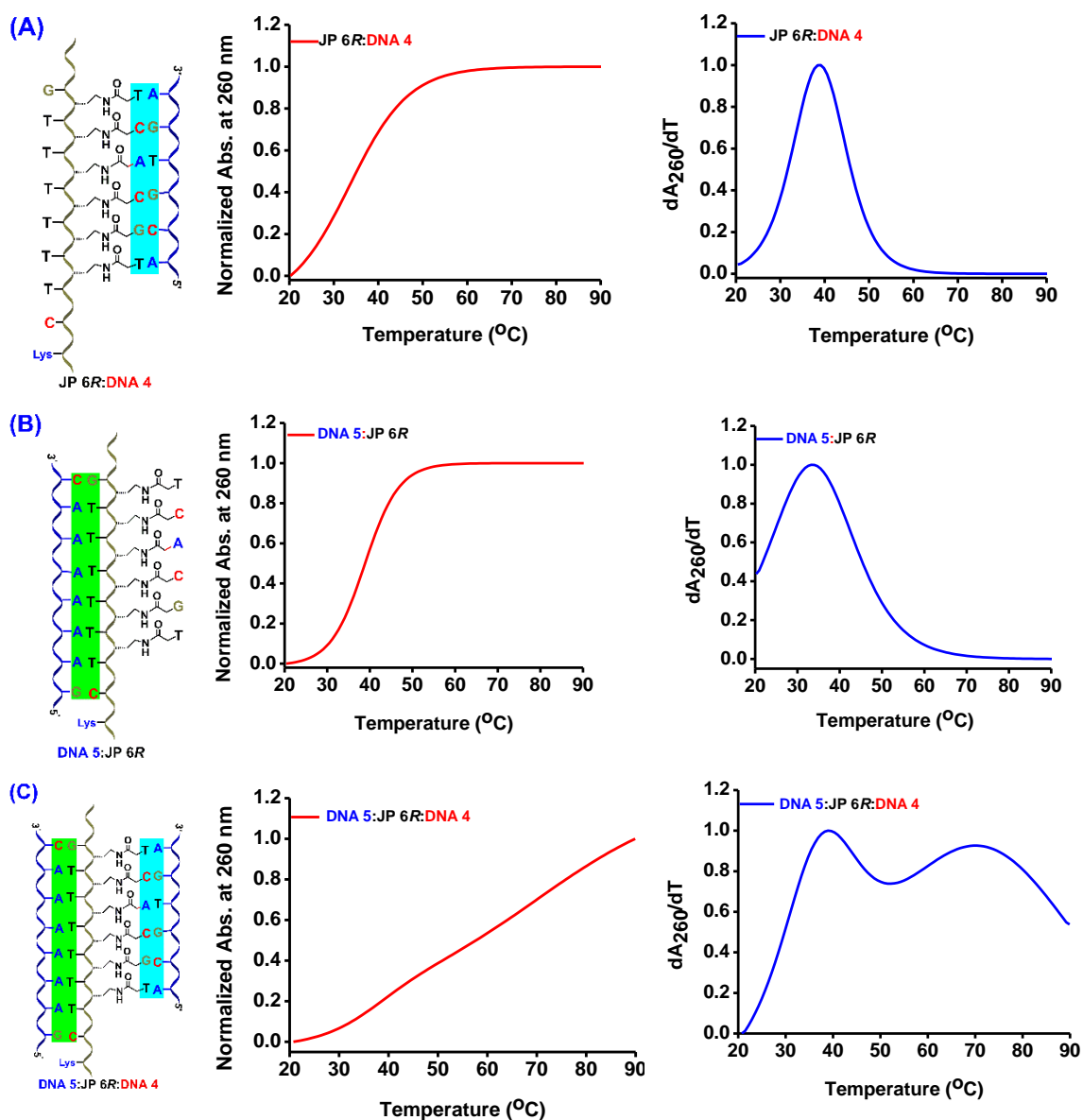


Figure 6.31 Temperature dependent UV absorbance curves for DNA complexes of **JP 5R** (*mjp-γR-PNA*) with DNA 5 (DNA = 5'-CAAAAAG-3') and DNA 4 (5'-ACGTGA-3'); Buffer 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM (A) **JP 6R:DNA 4** (*mjp-γR-PNA*:DNA 4) (B) **DNA 5:JP 6R** (DNA 5:*mjp-γR-PNA*) (C) **DNA 5:JP 6R:DNA 4** (DNA 5:*mjp-γR-PNA*:DNA 4).

Table 6.16 UV - T_m (°C) of **JP 6R** (*mjp-γR-PNA*) with complementary DNA 4 and DNA 5

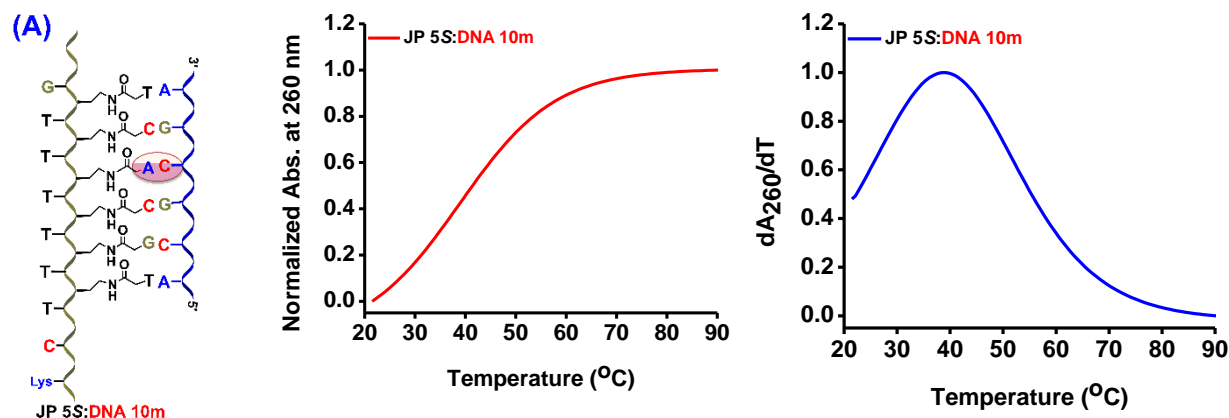
Entry	Mix Janus PNA:cDNA complexes	T_m (°C)		ΔT_m (°C)		
		<i>t</i> -Amide	C γ	<i>t</i> -Amide	C γ	C γ - <i>t</i> -Amide
1	JP 6R:DNA 4 (<i>mjp-γR-PNA</i> :DNA 4)	-	38.1	-	-	+4.7
2	DNA 5:JP 6R (DNA 5: <i>mjp-γR-PNA</i>)	33.4	-	-	-	
3	DNA 5:JP 6R:DNA 4 (DNA 5: <i>mjp-γR-PNA</i> :DNA 4)	38.7	69.8	+5.3	+31.7	+31.1

DNA **5** (5'-CAAAAAAG-3') and DNA **4** (5'-ACGTGA-3'), **JP 6R** = *mjp- γ R-PNA*, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C

6.6.4a Thermal stability of duplexes and double duplex of C γ (S/R) Janus PNAs with mismatch DNA

The sequence specificity of binding of the Janus PNA **JP 5S** with cDNAs were examined by studying their hybridization with cDNA **9** and cDNA **10** containing single-site mismatch for t-amide and C γ -face duplexes respectively (Figure 6.32). The C γ -face mismatch duplex **JP 5S**:DNA **10** (*mjp- γ S-PNA*:DNA **10**) showed a single transition corresponding T_m = 38.7 °C which is lower by 13 °C than the T_m seen with perfect complementary duplex **JP 5S**:DNA **4** (*mjp- γ S-PNA*:cDNA, **4**, T_{m1} = 51.7 °C). The mismatched duplex from t-amide face DNA **9**:**JP 5S** (DNA **9**:*mjp- γ S-PNA*) also showed a single transition with T_m of 32.9 °C (Figure 6.32) with T_m lower by 12.6 °C than the perfect duplex.

In presence of both face mismatch DNAs (DNA **9** and DNA **10**), the Janus PNA **JP 5S** formed double duplex DNA **9**:**JP 5S**:DNA **10** (DNA **9**:**JP 5S**:DNA **10**), but exhibited a single melting transition with T_m = 61.5 °C (Figure 6.32C). The observance of single transition in this double duplex is similar to situations observed before in mismatched complexes, due to the similar T_m s of both duplexes. Compared with their individual isolated duplexes, the T_m s of both t-amide and C γ -duplexes in the double duplex were enhanced by +28.9 °C and +22.8 °C respectively (Table 6.17). The initial destabilization of individual duplexes and enhancement in T_m s similar to that seen in perfect duplexes substantiated the formation of double duplexes in C γ (S/R)-Janus PNAs by sequence specific complementary base pairing.



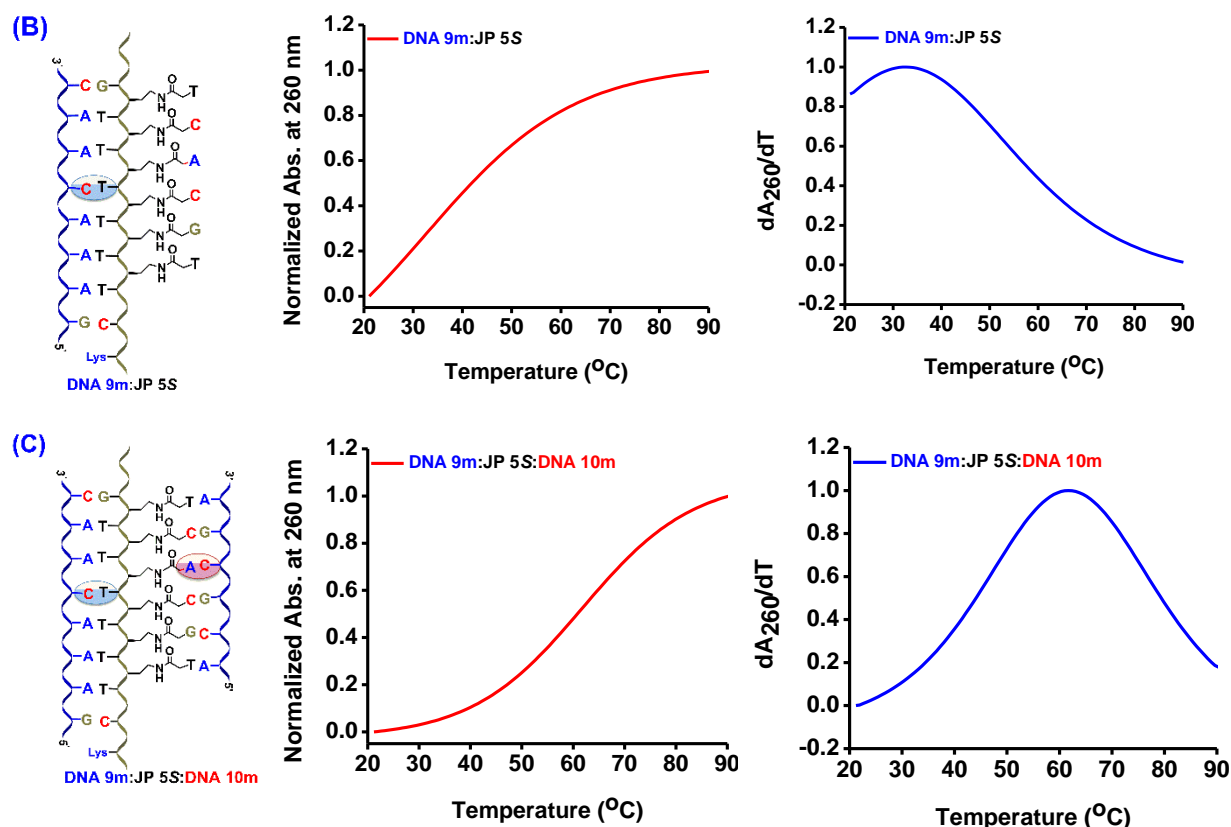


Figure 6.32 Temperature dependent UV absorbance curves for **JP 5S** (*mjp- γ -PNA*) with mismatch DNA duplex and double duplexes, DNA **9** (5'-GAAACAAC-3'); DNA **10** (5'-ACGCGA-3'); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) (A) **JP 5S:DNA 10** (B) DNA **9:JP 5S** (C) DNA **9:JP 5S:DNA 10**

Table 6.17 UV - T_m ($^{\circ}\text{C}$) of **JP 2** (*S mjp- γ -PNA*) with mismatch DNA **9** and DNA **10**

Entry	Janus PNA:mismatchDNA complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)	
		t-Amide	C^{γ}	t-Amide	C^{γ}
1	JP 5S:DNA 10m	-	38.7	-	-
2	DNA 9m:JP 5S	32.9	-	-	-
3	DNA 9m:JP 5S:DNA 10m	61.5		+28.9	+22.8

DNA **9** (5'-GAAACAAC-3') and DNA **10** (5'-ACGCGA-3'), **JP 5S** = *mjp- γ S-PNA*, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 $^{\circ}\text{C}$.

The stereomeric **JP 6R** (*mjp- γ R-PNA*) was also hybridized with cDNAs having single base mismatch at the middle of the sequence. The Janus PNA **JP 6R** (*mjp- γ R-PNA*) hybridized with mismatch DNA **10** from C^{γ} -face to form duplex **JP 6R:DNA 10** (*mjp- γ R-PNA:DNA 10*) with single transition (Figure 6.33A) and $T_m=34.2$ $^{\circ}\text{C}$, which is lower by 3.9 $^{\circ}\text{C}$ than the T_m of complementary duplex **JP 6R:DNA 4** (*mjp- γ R-PNA:cDNA*, $T_{m1} = 38.1^{\circ}\text{C}$).

The *t*-amide face DNA **9:JP 6R** (DNA **9:mjp- γ R-PNA**) exhibited single transition (Figure 6.33B) at $T_m = 29.6$ °C, with destabilization of 5.8 °C compared to perfect duplex. The double duplex DNA **9:JP 6R:DNA 10** (DNA **9:mjp- γ R-PNA:DNA 10**) displayed a single transition (Figure 6.33C) with a $T_m = 52.3$ °C (Table 6.18, entry 3), which is higher than the T_m of each of the individual duplexes. The stability of mismatched complex of *S*-stereomer **JP 5S** DNA **9:mjp- γ PNA:DNA 10** was higher than the that of R-stereomer **JP 6R** DNA **9:mjp- γ R-PNA:DNA 10** and both are less stable than their perfect duplexes **JP 5S** (DNA **5:mjp- γ S-PNA:DNA 4**) and **JP 6R** (DNA **5:mjp- γ R-PNA:DNA 4**) respectively.

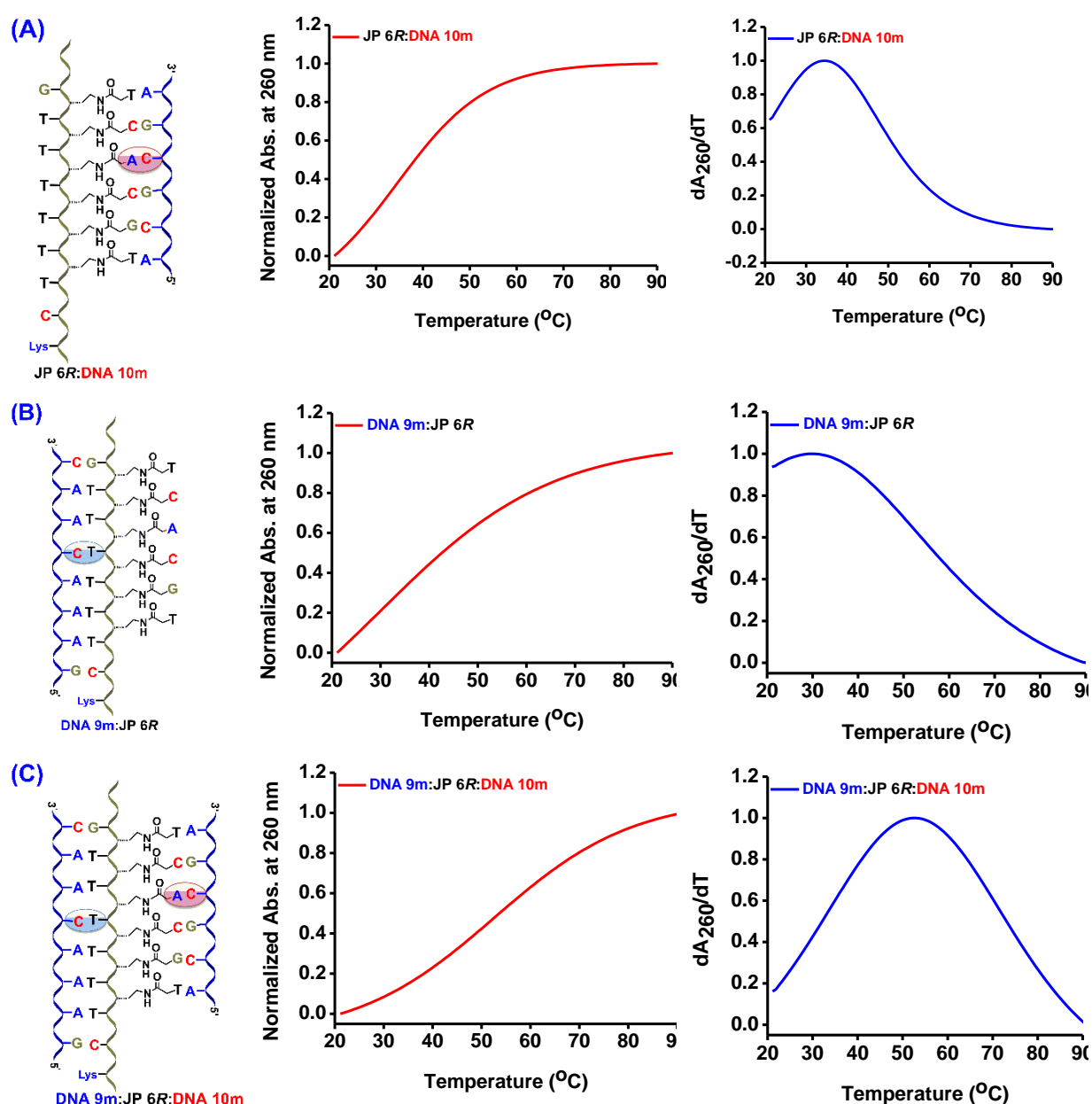


Figure 6.33 Temperature dependent UV absorbance curves for **JP 6R** (*mjp- γ R-PNA*) with mismatch DNA duplex and double duplexes, DNA **9** (5'-GAAACAAC-3'); DNA **10** (5'-

ACGCGA-3'; Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM) (A) **JP 6R**:DNA **10**
(B) DNA **9**:**JP 6R** (C) DNA **9**:**JP 6R**:DNA **10**

Table 6.18 UV - T_m (°C) of **JP 6R** (*mjp- γ R-PNA*) with mismatch DNA **9m** and DNA **10m**

Entry	Janus PNA:mismatchDNA complexes	T_m (°C)		ΔT_m (°C)	
		t-Amide	C^γ	t-Amide	C^γ
1	JP 6R :DNA 10m	-	34.2	-	-
2	DNA 9m : JP 6R	29.6	-	-	-
3	DNA 9m : JP 6R :DNA 10m	52.3		+22.7	+18.1

DNA **10** (5'-GAAACAAC-3') and DNA **9** (5'-ACGCGA-3'), **JP 6R** = *mjp- γ R-PNA*, ΔT_m indicates the difference in T_m s of single duplex with each face with that corresponding in double duplex. T_m values are accurate to ± 1.0 °C

6.6.4b CD spectra of duplexes of C^γ (S/R) hetero Janus PNAs, with cDNA 4/cDNA 5

The CD spectra of t-amide duplex DNA **5**:**JP 5S** (*mjp- γ S-PNA*) shows a positive maximum at 263 nm and a negative maximum at 244 nm with cross-over points at 239 nm. The CD spectra of C^γ -duplex **JP 5S** (*mjp- γ S-PNA*):DNA **4** shows positive maxima at 261 nm and a negative maxima at 240 nm with cross-over points at 239 nm (Figure 6.34). In comparison, the double duplex DNA **5**:**JP 5** (*S mjp- γ PNA*):DNA **6** shows positive maxima at 263 nm and a negative maximum at 244 nm (Figure 6.42A). The Cd spectra of all duplexes showed similar pattern with slight intensity variations and resembled the CD spectra characteristic of PNA:DNA duplexes

The CD spectra of stereometric C^γ (R) Janus PNA (**JP 6R**) with t-amide face duplex DNA **5**:**JP 6R** (*mjp- γ R-PNA*) shows a positive band at 259 nm and a negative band at 237 nm with cross-over points at 240 nm. The C^γ -face duplex of **JP 6R** (*R mjp- γ PNA*):DNA **4** shows slightly weaker intensity band with positive maxima 260 nm and a negative band at 239 nm with cross-over points at 232 nm and 240 nm. The corresponding double duplex DNA **5**:**JP 6R**:DNA **6** (*mjp- γ PNA*):DNA **4** shows a positive maximum at 264 nm and a negative maximum at 244 nm with cross over point 239 nm (Figure 6.42 B). The CD spectra of duplexes of C^γ (R) Janus PNA suggest that the double duplex has a slightly altered conformation compared to individual duplexes, but not much differences were seen between the two stereometric Janus double duplexes.

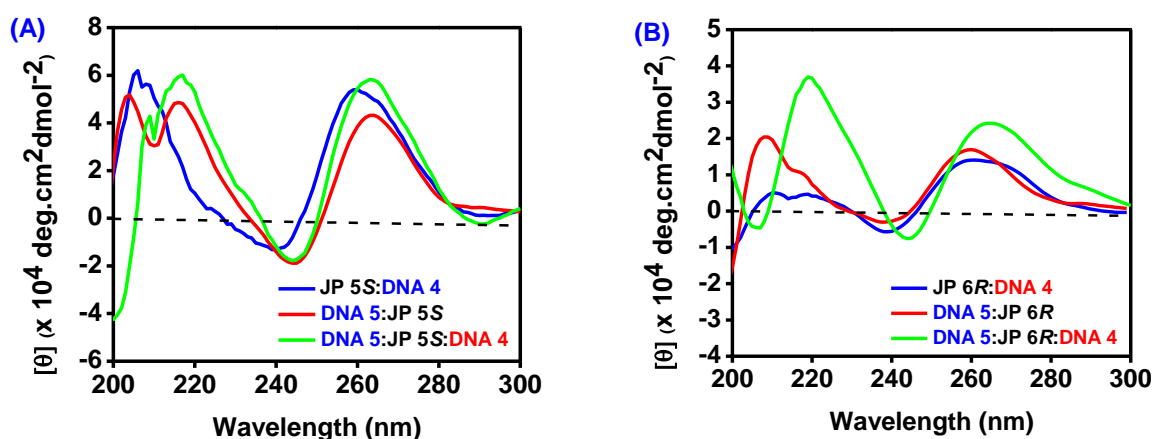


Figure 6.34 CD spectra of (A) *Hetero Janus* PNA:DNA duplex, DNA 5:JP 5S (DNA 5:mjp- γ S-PNA), JP 5S:DNA 4 (mjp- γ S-PNA:DNA 4) and *Hetero Janus* PNA:DNA double duplex DNA 5:JP 5S:DNA 4 (DNA 5:mjp- γ S-PNA:DNA 4) (B) *Hetero Janus* PNA:DNA duplex, DNA 5:JP 6R (DNA 6:mjp- γ R-PNA), JP 6R:DNA 4 (mjp- γ R-PNA:DNA 4) and *Hetero Janus* PNA:DNA double duplex DNA 5:JP 6R:DNA 4 (DNA 5:mjp- γ R-PNA:DNA 4) Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

6.7 Thermal stability of non Janus C ^{γ} -PNA-C_n with complementary DNA

In this section, the UV- T_m experiments of DNA duplexes from non Janus γ PNA (*aeg- γ* PNA) and *aeg*-PNA with the complementary DNA to are comparatively examined to see the roles of amide-nucleobases and C ^{γ} -nucleobases in stabilizing the PNA:DNA duplexes.

6.7.1 i-Motif studies by pH-dependent UV- T_m

The UV- T_m studies of the *i*-motif formation by *homo Janus* PNA JP 1S (T₈j_p- γ S-C₆), JP 2R (T₈j_p- γ S-C₆), PNA 11S (p₈- γ S-C₆), PNA 12R (p₈- γ R-C₆), PNA 18 (PNA-C₆, *aeg* PNA) and DNA 3 (dC₆), were done at acidic pH [sodium acetate (100 mM) buffer (pH 3.2)] by monitoring absorbance at 295 nm which is characteristic of C⁺ (N3-H⁺).¹²

The UV-temperature plot of PNA 18, DNA 3, *Janus* PNA JP 1S and *Janus* PNA JP 2R monitored at 295 nm at acidic pH 3.2 and the corresponding T_m of the *i*-motif structures were evaluated from first derivative. DNA 3 and PNA 18 showed broad transitions and negative inverse first derivative curves indicating the formation of *i*-motif tetraplexes with T_m 55.1 °C and 75.5 °C respectively, with T_m of PNA 18 > T_m of DNA 3 (Figure 6.35A and B). The *Janus* PNA JP 1S and amino-C ^{γ} PNA 11S show transition at 58.2 °C and 56.6 °C respectively with T_m of PNA JP 1S > T_m of PNA 11S (Figure 6.36A and B). However *Janus* PNA 2 and amino-C ^{γ} PNA 12 showed negative transition at 55.5 °C and 54.9 °C respectively with T_m of JP 2R > T_m of PNA 12 (Figure 6.37A and B). From these experiments, it is

clearly seen that *Janus* PNA **JP 1S** and *aeg* PNA **18** form i-motif with better stability compared to DNA **10**, **JP 2R** and PNA **12R**.

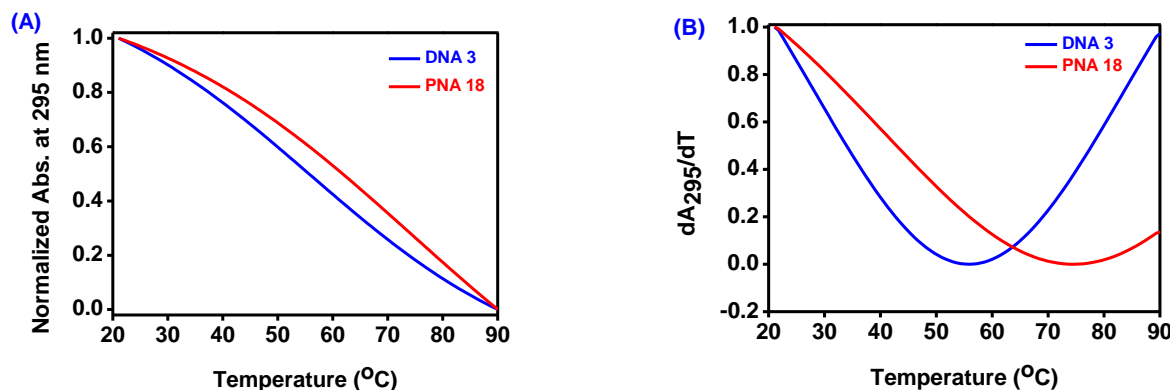


Figure 6.35 UV- T_m of C_n -tetraplexes of DNA/PNA/ amino PNA **18** at pH 3.27 (A) normalised absorbance at 295 nm (B) Ist derivative curve (dA_{295}/dT).

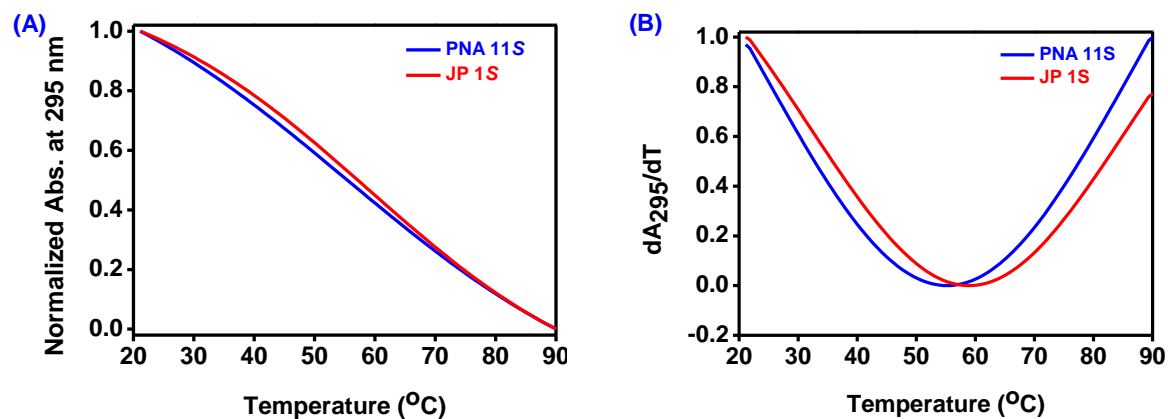


Figure 6.36 UV- T_m of C_n -tetraplexes of DNA/PNA/ amino PNA **11S** and *Janus* PNA **JP 1S** at pH 3.27 (A) normalised absorbance at 295 nm (B) Ist derivative curve (dA_{295}/dT).

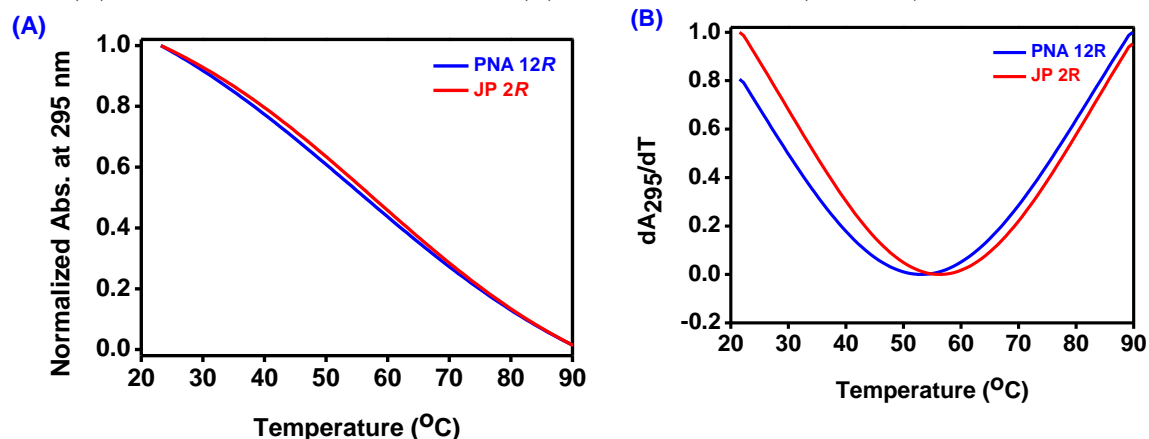


Figure 6.37 UV- T_m of C_n -tetraplexes of DNA/PNA/ amino PNA **12R** and *Janus* PNA **JP 2R** at pH 3.27 (A) normalised absorbance at 295 nm (B) Ist derivative curve (dA_{295}/dT)

Table 6.19. Summary of pH-dependent UV- T_m ($^{\circ}\text{C}$) of *homo Janus* PNA **JP 1S**, **JP 2R**, *aeg* PNA **11S** $p8\text{-}\gamma\text{-C}_6$, *aeg*-PNA **12 S** $p8\text{-}\gamma\text{-C}_6$ (amino PNA), PNA **18** PNA- C_6 (*aeg* PNA) and DNA **3** or dC_6

Entry	PNA:DNA Complexes	T_m ($^{\circ}\text{C}$)		ΔT_m ($^{\circ}\text{C}$)- duplex PNA- DNA
		t-amide	C^{γ}	
1	JP 1S ($T_{8jp}\text{-}\gamma\text{-C}_6$)		58.2	+3.1
2	JP 2R ($T_{8jp}\text{-}\gamma\text{-R-C}_6$)		55.5	+0.4
3	PNA 11S ($p8\text{-}\gamma\text{-C}_6$)		56.6	+1.5
4	PNA 12R ($p8\text{-}\gamma\text{-R-C}_6$)		54.9	-0.2
5	PNA 18 (PNA C_6)	75.5		+20.4
6	DNA 3 (dC_6)	55.1		-

T_m values are accurate to ± 1.0 $^{\circ}\text{C}$, ΔT_m^1 indicates the difference in T_m of PNA with DNA

6.7.2 Thermal stability of quadruplex formation by Janus PNA- G_n -oligomers

As pointed out earlier, various G_4 -tetraplexes are favoured in presence of metal ions Na^+ or K^+ , which co-ordinatively bind the O^6 of carbonyls of tetrameric G residues¹⁴. Hence formation of G_4 -tetraplexes from different G-oligomers **PNA 13S**, **PNA 14R** and *aeg*-PNA **19** (PNA- G_6) were examined by UV- T_m , with 100 mM salt concentration KCL and each oligomer strand concentration of 5 μM at pH 7.1. The transitions were monitored at 295 nm which is characteristic band for G-tetraplex in 20 mM sodium cacodylate buffer having 100 mM KCl concentration. The UV- T_m experiments monitored at 295 nm¹⁵ showed formation of G-tetraplexes from *aeg* PNA**19** characteristic inverse sigmoidal transition at 295 nm (Figure 6.38A) and the T_m of transition was 36.5 $^{\circ}\text{C}$ (Figure 6.38B).

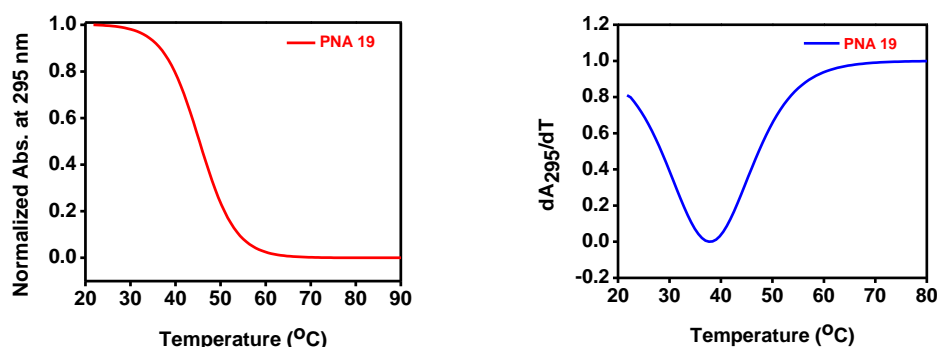


Figure 6.38 UV- T_m of PNA **19** at 20 mM sodium cacodylate buffer having 100 mM KCl, pH 7.1 (A) Normalised abs. at 295 nm (B) 1st derivative curve of $\text{d}A_{295}/\text{d}T$.

The Janus PNA **JP 3S** and **JP 4R** show an inverse sigmoidal transition with T_m 45.6 $^{\circ}\text{C}$ and 44.2 $^{\circ}\text{C}$ respectively. PNA G-oligomers *aeg* PNA **13** (S $p8\text{-}\gamma\text{-G}_6$) and PNA **14** (R $p8\text{-}\gamma\text{-G}_6$) also exhibit inverse sigmoidal transition with T_m 52.1 $^{\circ}\text{C}$ and 50.3 $^{\circ}\text{C}$ respectively (Figure

6.39A and B). The decreasing order of T_m at 100 mM KCl concentration is *aeg* PNA **19** **JP 4** ($T_{8jp-\gamma S-G_6}$) < **JP 3** ($T_{8jp-\gamma R-G_6}$) < PNA **14** ($p8-\gamma R-G_6$) < PNA **13** ($p8-\gamma S-G_6$).

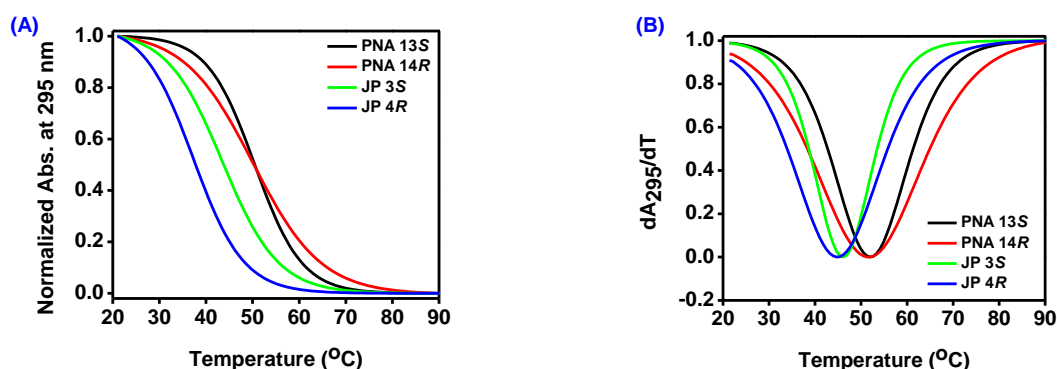


Figure 6.39 UV- T_m of PNA 13S, PNA 14R, JP 3S and JP 4R at 20 mM sodium cacodylate buffer having 100 mM KCl, pH 7.1 (A) Normalised abs. at 295 nm (B) 1st derivative curve of dA_{295}/dT .

Table 6.20. UV- T_m (°C) of PNA G-quadruplexes (5 μ M) in sodium cacodylate (20 mM, pH 7.1).

Entry	PNA:DNA complexes	T_m (°C)	ΔT_m (°C) C γ -PNA – <i>aeg</i> PNA
1	PNA 13S ($p8-\gamma S-G_6$)	52.1	+15.6
2	PNA 14R ($p8-\gamma R-G_6$)	50.3	+ 13.8
4	JP 3S ($T_{8jp-\gamma S-G_6}$)	45.6	+9.1
5	JP 4R ($T_{8jp-\gamma R-G_6}$)	44.2	+7.9
3	PNA 19 (PNA G_6)	36.5	-

T_m values are accurate to ± 1.0 °C, ΔT_m^1 indicates the difference in T_m of PNA with DNA

6.7.3 G-quadruplex studies by Variable-Temperature Circular Dichroism

For further proof of G-quadruplex formation and its stability, temperature-dependent CD experiments of *Janus PNA* JP **3** and JP **4**, amino C γ PNAs PNA **13** ($p8-\gamma S-G_6$) and PNA **14** ($p8-\gamma R-G_6$) were done under identical conditions as that used for UV- T_m experiments. The JP **3**, JP **4**, PNA **13** and PNA **14** 100 μ M strand concentration was prepared individually in 20 mM sodium cacodylate buffer containing 100 mM KCl at pH 7.1. The sample was heated to 90 °C for 15 min, cooled slowly to room temperature over 12 h and equilibrated at 4 °C for 24 h. The CD spectra were recorded at temperature 10 °C between 200 nm -320 nm range as average of 3 successive runs and baseline corresponding to buffer alone was subtracted from all spectra. The comparative CD spectra of PNA **13** and PNA **14** are shown in Figure 6.54 A. The both PNA **13** and PNA **14** shows characteristic G-quadruplex positive band at 298 nm and 297 nm and negative band at 283 nm and 286 nm respectively.

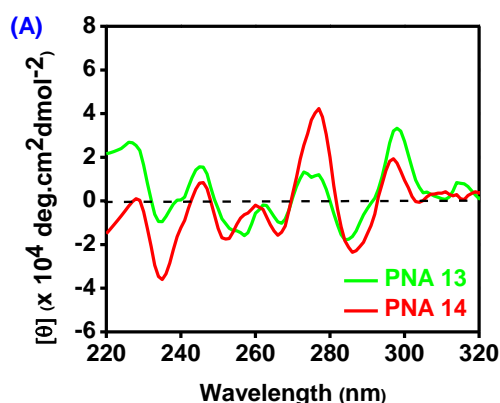


Figure 6.40 CD spectra of PNA 13 and PNA 14 (100 μ M), Sodium cacodylate 20 mM, KCl 100 mM.

The CD spectra of *Janus PNA JP 3* (*S T₈jp- γ -G₆*) and *JP 4* (*R T₈jp- γ -G₆*) were recorded at 10 $^{\circ}$ C- 70 $^{\circ}$ C with heating rate of 5 $^{\circ}$ C/min. The CD spectra of *JP 3* at 10 $^{\circ}$ C present as quadruplex showed a negative band centering at 264 nm and 284 nm and a weak positive band at 272 nm and strong positive band at 300 nm (Figure 6.41 A). In case of *JP 4* negative band at 274 nm and 290 nm where as positive band at 275nm and 301nm observed. At 70 $^{\circ}$ C these characteristic peak disappear in both cases (Figure 6.41 B). The observed CD at 10 $^{\circ}$ C arises from quadruplex of PNA 13 and PNA 14 and resembles that observed for antiparallel DNA quadruplexes such as d(GGTTTTGGTTTTGGTTTTGG).²⁹

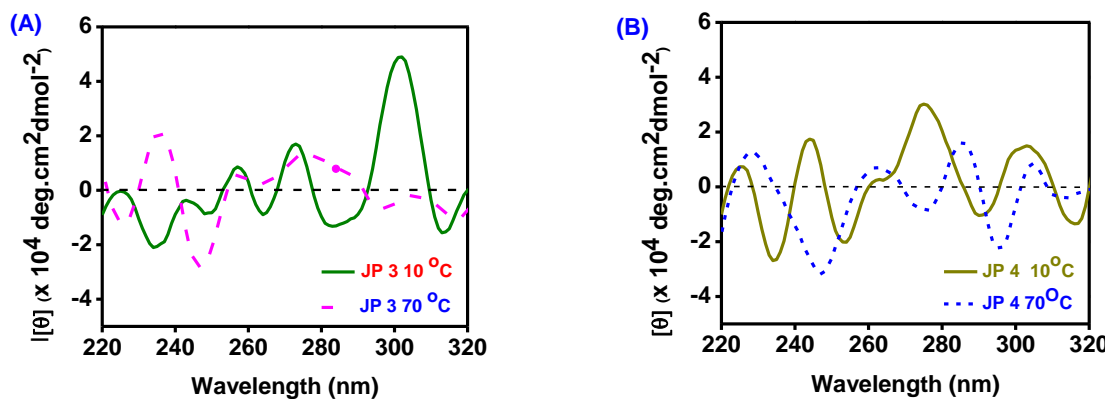


Figure 6.41 Temperature dependant CD spectra of (A) JP 3 at 10 $^{\circ}$ C and 70 $^{\circ}$ C (100 μ M) (B) JP 4 at 10 $^{\circ}$ C and 70 $^{\circ}$ C (100 μ M), Sodium cacodylate 20 mM, KCl 100 Mm.

6.8 Ethidium bromide displacement assay from *ds*DNA

Ethidium bromide (EtBr) is a weakly fluorescent and strong DNA intercalating agent¹⁶. The intercalation into DNA is a result of unwinding of DNA (around 26 $^{\circ}$) by ethidium cation which causes base pairs to separate or rise, creating an opening of about 3.4

Å. Also, charge repulsion between the two strands causes continuous breathing in the DNA which helps ethidium bromide to get inside the two strands. Intercalation of EtBr into the duplex DNA results in the 20 fold increase in its fluorescence. However, in case of PNA:DNA duplex, the unwinding of duplex is minimal and the breathing rate is decreased as a result of reduced repulsion between the backbone units which does not allow EtBr to intercalate between the PNA containing hybrid duplexes¹⁶. The differential binding ability of Janus PNAs **JP 1S** ($T_{8jp-\gamma S-C_6}$), **JP 2R** ($T_{8jp-\gamma R-C_6}$), **JP 3S** ($T_{8jp-\gamma S-G_6}$) and **JP 4R** ($T_{8jp-\gamma R-G_6}$) towards the complementary DNA **2** (dG₆) and DNA **3** (dC₆) was examined by the displacement of EtBr intercalated into the DNA duplex DNA **2** (dG₆):DNA **3** (dC₆). EtBr (6 μM) was treated with the DNA duplex dG₆:dC₆ (5 μM) and the resulting EtBr-dsDNA complex was individually titrated with aqueous solutions of various Janus PNA oligomers (Figure 6.42).

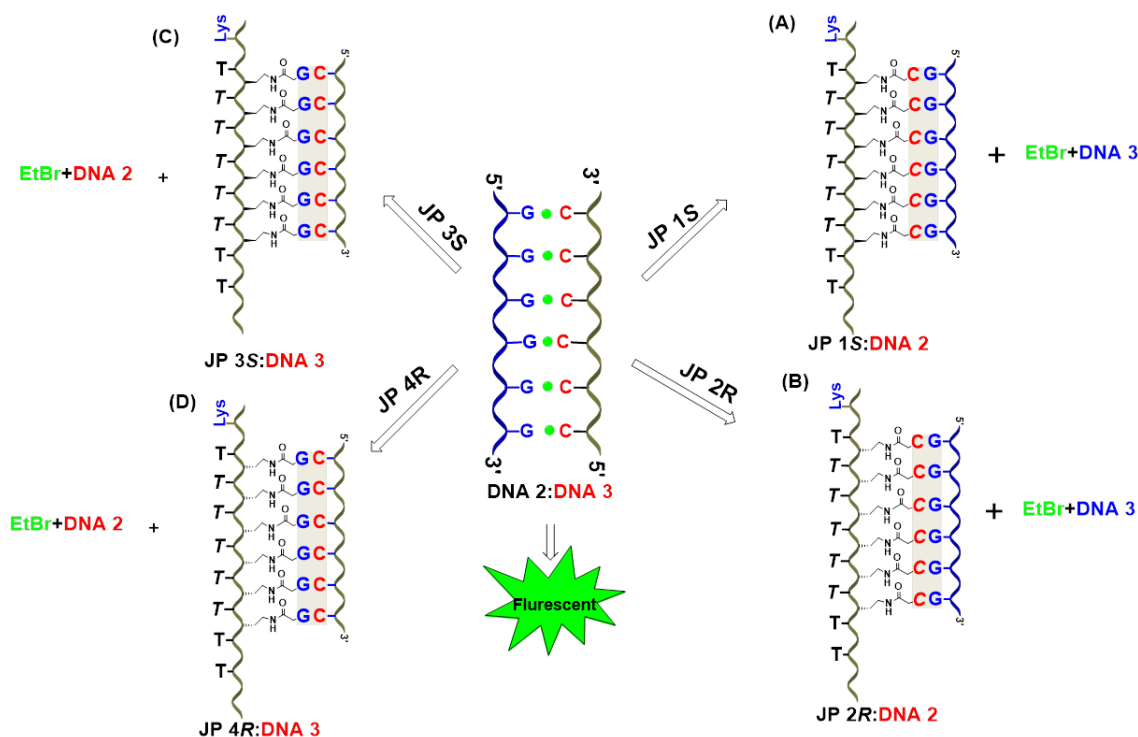


Figure 6.42 Schematic ethidium bromide displacement assay of **JP 1S**, **JP 2R**, **JP 3S** and **JP 4R** with **DNA 2:DNA 3** duplex.

The ethidium bromide intercalates between the base-pairs of duplex **DNA 2** (dG₆):**DNA 3** (dC₆) and the fluorescence spectra were recorded shows an enhancement of fluorescence emission. On binding of the *Janus PNA* **JP 1S** ($T_{8jp-\gamma S-C_6}$) and **JP 2R** ($T_{8jp-\gamma R-C_6}$) to the ds DNA causes a displacement of the intercalator resulting in a reduction in the fluorescence emission to form duplex of **JP 1S:DNA 2** ($T_{8jp-\gamma S-C_6}$:dG₆) and **JP 2R:DNA 3**

($T_{8jp-\gamma S-C_6}:\text{dG}_6$). It revealed distinct fluorescence quenching profiles for binding of **JP 1S** and **JP 2R** to (dG_6):(dC_6) DNA duplex. With **JP 2R: DNA 2** duplex, the fall in fluorescence was steeper than with **JP 1S:DNA 3** duplex. This revealed a similar affinity of Janus PNA **JP 1S** ($T_{8jp-\gamma S-C_6}$) to (dG_6):(dC_6) duplex and **JP 2R** ($T_{8jp-\gamma R-C_6}$) (Figure 6.43A).

The Janus PNAs **JP 3S** ($T_{8jp-\gamma S-G_6}$) and **JP 4R** ($T_{8jp-\gamma R-G_6}$) were also able to displace ethidium bromide from $\text{dG}_6:\text{dC}_6$ duplex to form **JP 3S:DNA 3** ($T_{8jp-\gamma S-G_6}:\text{dC}_6$) and **JP 4R:DNA 3** ($T_{8jp-\gamma R-G_6}:\text{dC}_6$). The fall in intensity of fluorescence with **JP 4R:DNA 3** is steeper than **JP 3S:DNA 3** (Figure 6.43B) similar to that observed in case of **JP 2R:DNA 2** duplexes. However, a comparison of the quenching fluorescence profiles of ($T_{8jp-\gamma S-C_6}$), ($T_{8jp-\gamma R-C_6}$), ($T_{8jp-\gamma S-G_6}$) and ($T_{8jp-\gamma R-G_6}$) showed that the nature of interaction between the Janus PNA and the double helical DNA was dependent on the stereochemistry of the nucleobase substitutions at C' position of the Janus PNA.

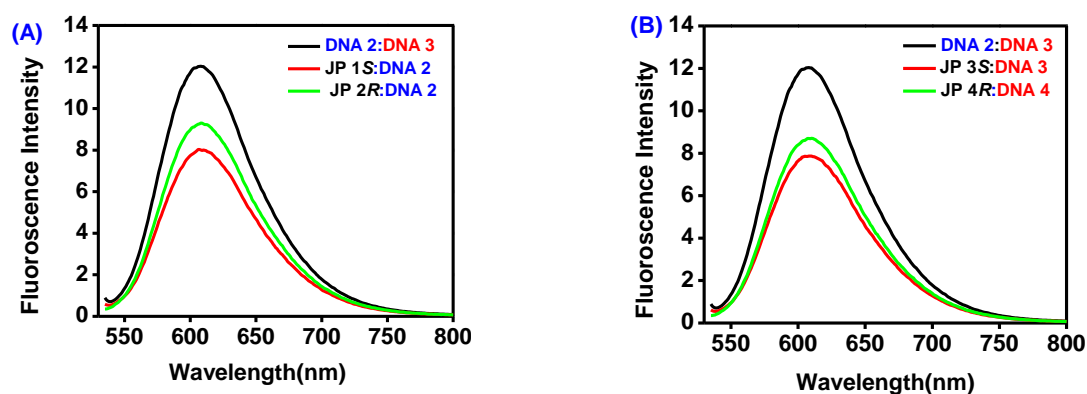


Figure 6.43 (A) Fluorescence emission decay by Janus PNA **JP 1S** ($T_{8jp-\gamma S-C_6}$) and **JP 2R** ($T_{8jp-\gamma R-C_6}$) (B) Fluorescence emission decay by Janus PNA **JP 3S** ($T_{8jp-\gamma S-G_6}$) and **JP 4R** ($T_{8jp-\gamma R-G_6}$) on duplexes **DNA 2** (dG_6):**DNA 3** (dC_6) duplex. DNA 2 ($\text{dG}_6=5'$ GGGGGG $3'$), DNA 3 ($\text{dC}_6=5'$ CCCCCC $3'$), Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM.

6.9 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) was used to study the thermodynamics of hybridization of the *homo* Janus PNA **JP 1S** ($T_{8jp-\gamma S-C_6}$) and **JP 2R** ($T_{8jp-\gamma R-C_6}$) with DNA **1** (dA_8), which is complementary to the T_8 on the t-amide face of the Janus PNA. The results were compared with those obtained from hybridization of PNA **7S** ($T_{8jp-\gamma S-NH^{Ac}}$) and PNA **8R** ($T_{8jp-\gamma R-NH^{Ac}}$) with no nucleobase at t-amide side to reveal the effect of nucleobase (cytosine) on the **JP 1S** ($T_{8jp-\gamma S-C_6}$) and **JP 2R** ($T_{8jp-\gamma R-C_6}$) on thermodynamic parameters of formation of $T_8:\text{A}_8:T_8$ triplex on the t-amide face of the Janus PNA. The comparison of

the studies between the **JP 1S** ($T_8jp-\gamma S-C_6$) and **PNA 7S** ($T_8jp-\gamma S-NH^{Ac}$), **JP 2R** ($T_8jp-\gamma R-C_6$) and **PNA 8R** ($T_8jp-\gamma R-NH^{Ac}$) also was drawn to understand the effect of configuration on the stability of the duplexes formed on the t-amide face.

Aliquots (2 μ L) of DNA dA_8 solution were dispensed from the syringe into the sample cell containing 200 μ L of PNA at a rate of 5.5 μ L per second. A total of 19 injections were made with an interval of 2.5 min between injections. The stirring speed was maintained at 700 rpm during the titrations. The reference cell and the sample cell were supplied with a constant current to regulate the differential power at 5 μ cal/sec, while maintaining the temperature at 8 $^{\circ}$ C. The buffered solutions of PNA and DNA contained identical concentrations of sodium cacodylate (10 mM) and sodium chloride (10 mM). Reference experiments (DNA vs Buffer) were performed by titrating DNA into buffer under the same conditions. The measured heat of dilution was subtracted from the sample experiments to obtain the heat changes only due to the PNA:DNA binding interactions. The binding isotherm was fitted to model of one set of binding sites using MicroCal data analysis software and the thermodynamic parameters were obtained. In each experiment, the concentration was expressed in terms of nucleobases for comparison of data across different samples. The *Janus* PNA concentration was derived from this titration by fixing the number of sites to 0.5, expected for $T_8:A_8:T_8$ base pair in a triplex. Then, the thermodynamic parameters, such as K_D , ΔG , ΔH , and ΔS were obtained for all the binding experiments.

6.9.1 Binding of the **JP 1S** ($T_8jp-\gamma S-C_6$) and **PNA 7S** ($T_8jp-\gamma S-NH^{Ac}$) double duplex with of complementary DNA.

The homooligomeric sequence of T_8 on the t-amide face of **JP 1S** ($T_8jp-\gamma S-C_6$, 42 μ M) was taken in cell and titrated with its complementary DNA **1** dA_8 (400 μ M) at 8 $^{\circ}$ C and Figure 6.42A shows the binding isotherm which showed a saturation at molar ratio of 0.9. The K_D was found to be 0.74 μ M, ΔG -7.9 kcal/mol, ΔH -11.4 kcal/mol and entropy change ΔS -12.4 cal/mol. The stoichiometry of binding N which is the ratio of number of bases involved in base pairing from DNA to number of bases from PNA was found to be 0.5 clearly suggesting the formation of (**JP 1S**)₂:DNA **1** triplex. This supports the results of UV-T and CD experiments that showed features characteristic of formation of triplex.

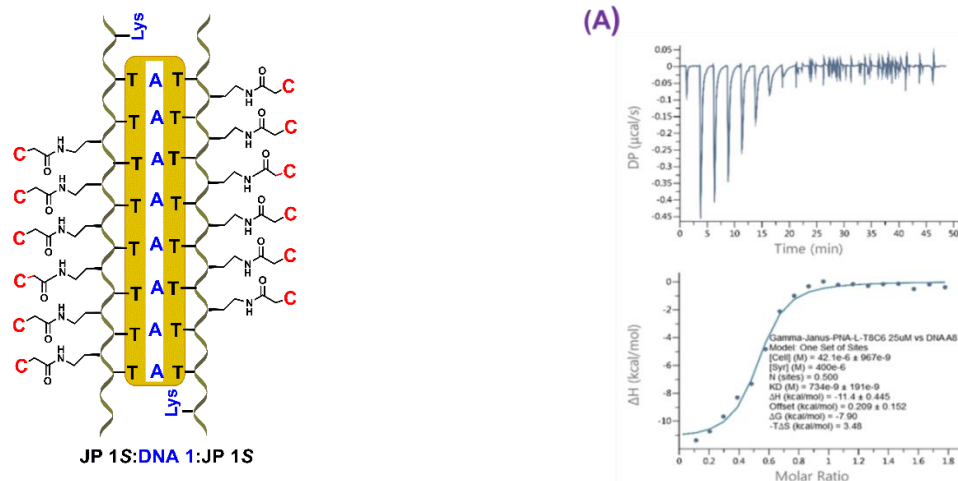


Figure 6.44 ITC binding isotherms for complex formation of **JP 1S** ($T_{8jp-\gamma S-C_6}$) with DNA **1** (dA_8)

To investigate the role played by nucleobases at the C^γ -face, the hybridization of control PNA **7S** ($T_{8jp-\gamma S-NH^{Ac}}$, 91 μM) titrated with DNA **1** (dA_8 , 800 μM) at 8°C was studied (Figure 6.45B). The isotherm was similar to that of **JP 1S** ($T_{8jp-\gamma S-C_6}$) with PNA **7S** ($T_{8jp-\gamma S-NH^{Ac}}$). The formation of PNA **7S** ($T_{8jp-\gamma S-NH^{Ac}}$):DNA **1** (dA_8) complex reached saturation at molar ratio of 0.9. The K_D was found to be 3.12 μM . The ΔG of -7.1 kcal/mol accompanied by enthalpy change ΔH of -12.2 kcal/mol and entropy change ΔS of -18.3 cal/mol were obtained from the binding isotherm. The stoichiometry of binding N which is the ratio of number of bases involved in base pairing from DNA to number of bases from PNA was found to be 0.5 clearly suggesting the formation of (PNA **7S**)₂:DNA **1** triplex. This supports the results of UV-T and CD experiments that showed features characteristic of formation of triplex.

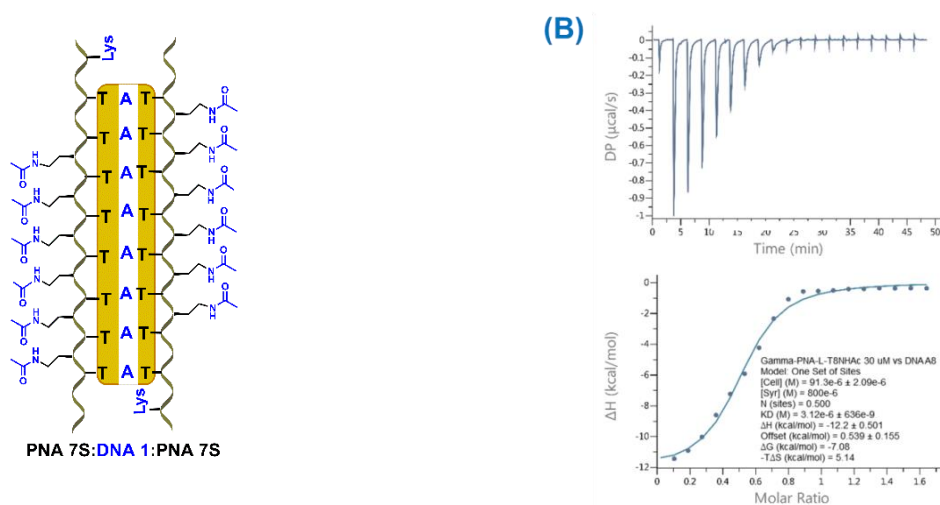


Figure 6.45 ITC binding isotherms for PNA **7S** ($T_{8jp-\gamma S-NH^{Ac}}$):DNA **1** (dA_8) hybrid formation

6.9.2 Binding of the JP 2R ($T_8jp-\gamma R-C_6$) and PNA 8R ($T_8jp-\gamma R-NH^{Ac}$) with cDNA (dA_8)

The stereomeric **JP 2R** ($T_8jp-\gamma R-C_6$) was titrated with the complementary DNA **1** (dA_8) and the ITC results are shown in Figure 6.46A. **JP 2R** (59.6 μM) was taken in ITC cell and DNA **1** (800 μM) was titrated at 8°C. The binding reaction reached a saturation at molar ratio of 1.25. The K_D was found to be 2.54 μM , free energy change ΔG -7.2 kcal/mol, enthalpy ΔH and entropy ΔS changes were -8.2 kcal/mol and -3.3 cal/mol respectively.

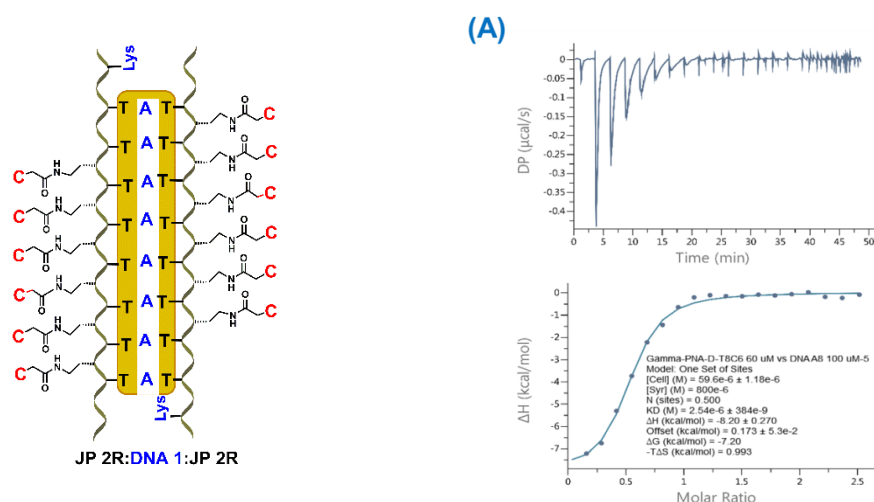


Figure 6.46 ITC binding isotherms for duplex **JP 2R** ($T_8jp-\gamma S-C_6$):DNA **1**(dA_8) formation. The hybridization of DNA **1**(dA_8) with control PNA **8R** ($T_8jp-\gamma R-NH^{Ac}$) was studied to probe the effect of the nucleobase containing sidechain at the C^γ site on the **JP 2R** ($T_8jp-\gamma R-C_6$). The PNA **8R** ($T_8jp-\gamma R-NH^{Ac}$, 140 μM) taken in ITC cell was titrated against DNA **1** dA_8 (800 μM) taken in the syringe at 8°C. Figure 6.46B shows the binding isotherm with saturation achieved at at 1.0 molar ratio. The K_D was found to be 14.9 μM , ΔG was -6.21 kcal/mol, enthalpy ΔH and entropy ΔS changes to be -10.3 kcal/mol and -14.4 cal/mol respectively. The stoichiometry of binding N which is the ratio of number of bases involved in base pairing from DNA to number of bases from PNA was found to be 0.5 clearly suggesting the formation of (**JP 2R**)₂:DNA **1** triplex. This also supports the results of UV-T and CD experiments that showed features characteristic of formation of triplex.

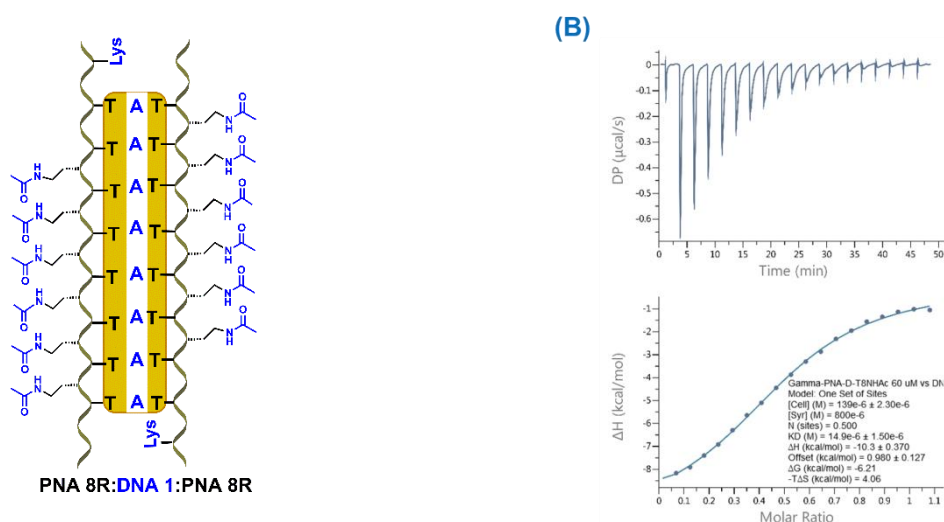


Figure 6.47 ITC binding isotherms of PNA **8R** ($T_{8jp}\text{-}\gamma R\text{-NH}^{Ac}$):DNA **1** (dA_8) hybrid formation.

The thermodynamic parameters derived from the ITC studies performed on the four different PNAs is tabulated in Table 1. It shows the comparative K_D , ΔG , ΔH and ΔS values of complexes formed by L- and D-Janus PNAs with and without nucleobases at the C^γ position.

Table 6.20 Thermodynamic parameters of PNA:DNA hybridization studies with ITC

PNA	K_D ($\times 10^{-6} M$)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (cal/mol.K)	Wiseman Constant
$T_{8jp}\text{-}\gamma S\text{-}C_6$ (JP 1S) :DNA 1	0.73 ± 0.19	-7.90 ± 0.21	-11.4 ± 0.45	-12.38 ± 2.5	57.36
$T_{8jp}\text{-}\gamma S\text{-NH}^{Ac}$ (PNA 7S) :DNA 1	3.12 ± 0.64	-7.08 ± 0.16	-12.2 ± 0.50	-18.28 ± 2.6	29.26
$T_{8jp}\text{-}\gamma R\text{-}C_6$ (JP 2R) :DNA 1	2.54 ± 0.38	-7.20 ± 0.12	-8.2 ± 0.27	-3.53 ± 1.5	23.46
$T_{8jp}\text{-}\gamma R\text{-NH}^{Ac}$ (PNA 8R) :DNA 1	14.9 ± 1.50	-6.21 ± 0.08	-10.3 ± 0.37	-14.44 ± 1.9	9.33

In summary, the ITC results described above suggest that

- (i) K_D : (a) The Janus PNA *S*-oligomers bind to cDNA **1** better than the corresponding *R*- oligomers as indicated by lower K_D values: **JP 1S**:DNA **1** < **JP 2R**:DNA **1**; PNA **7S**:DNA **1** < **JP 2R**:DNA **1**
- (b) K_D : Janus PNAs bind better than control *aeg*-PNAs: **JP 1S**:DNA **1** < PNA **7S**:DNA **1**; **JP 2R**:DNA **1** < **JP 2R**:DNA **1**

These support the conclusions from order of T_m values from UV-T data.

- (ii) **Stoichiometry N:** Expressed in terms of ratio (DNA/PNA) of number of bases involved in base pair formation, in all cases $N = 0.5$ which corresponds to (PNA)₂:DNA triplex as supported by UV and CD experiments in terms of triplex characteristic features.
- (iii) **Free energy changes ΔG :** The negative values suggest favourable binding reactions and the range -6.0 to -8.0 kcal/mol is similar to that observed with standard PNA:DNA triplexes.²⁸
- (iv) **Enthalpy changes ΔH :** The negative values observed suggest favourable reaction, with more negative values for Janus PNA *S*-oligomers than Janus PNA *R*-oligomers. This again suggests better binding by *S*-stereo oligomers compared to *R* stereomer.
- (v) **Entropy changes ΔS :** The negative values of entropy changes indicate that the binding reactions are not favoured entropically. The most less negative values for **JP 2R:DNA 1** indicates it is relatively better favoured entropically than other complexes. The fact that all duplexes have net negative free energy suggests a good entropy-enthalpy compensation and the binding reactions are driven essentially by enthalpic factors. The large number of base paired H-bonds in triplex are responsible for high enthalpy driven binding reactions. The full understanding of relative thermodynamic parameters favouring Janus PNA complex formation requires ITC experiments with other duplexes and double duplexes.

6.10 Conclusion

This chapter illustrates effective complex formation by Janus PNA $C^\gamma(S/R)$ -oligomers with cDNA. The ability of nucleobases linked to C^γ -side chain in PNAs that do not have any nucleobases linked through t-amide linkage (as in standard *aeg*-PNAs), to base pair with cDNA to form duplex / triplex were first demonstrated by their single sigmoidal transition in UV-T absorbance plot. These showed lower T_m with mismatched cDNA sequence confirming sequence dependence of binding. The *aeg*- $C^\gamma(S/R)$ -PNA designed are a new class of PNA analogues, which are devoid of t-amide linked nucleobases as in standard *aeg*-PNA and nucleobases attached /anchored to backbone through a flexible linker still form sequence specific base pairing similar to that in classical PNA.

The homo Janus $C^\gamma(S/R)$ $T_{7jp-\gamma-C_5}$ PNAs form (**JP 1S/2R**)₂:DNA **1**(dA₈) triplexes as expected for (poly dT)₂:dA system. In the presence of cDNA (dG₆), the triplex formed double duplex of triplex $dG_6:C_6$ -**JP1S/2R**-T₈:dA₈:T₈-**JP 1S/2R**-C₆:dG₆. This was indicated by two

transitions in UV-T plot, one for the triplex and the other for the identical duplexes. Both T_m s were higher than the T_m s of individual duplex T_8 -**JP 1S/2R**-C₆:dG₆ and triplex **JP 1S/2R**₂:DNA **1**(dA₈). The mixed Janus PNA sequences (Figure 6.5 and 6.6) also exhibited duplex formation with cDNA from both faces, with T_m in double duplexes higher than individual duplexes. It was also found that the *S*-stereomeric Janus PNAs always gave complexes with stability higher than the *R*-stereomeric Janus PNAs. The sequence specificity of base pairing was established by observance of lower T_m in mismatched cDNA sequences. The strand invasion through competitive binding was illustrated by displacement of intercalated ethidium bromide from DNA duplex by Janus PNA. The CD profiles of Janus duplexes and triplexes were characteristic of typical PNA:DNA duplexes and triplexes showing conformational similarity of Janus duplexes and triplexes. It was demonstrated that during formation of double duplex or triplex, the order of formation of complex does not matter; first duplex followed by triplex or vice versa; the conformation of the final complex is always same. The formation of triplex was supported by ITC experiments, which confirmed the stoichiometry (N) of binding as ratio of DNA/PNA bases and the relative strength of complexes (K_D) which correlated well with their relative T_m seen in UV-T plots. The ΔG for the binding reaction were all negative clearly signifying favourable association and the ΔH and ΔS values implied that they were driven mainly by enthalpic factors, since formation of triplexes are not entropically favourable.

It should be noted that through out this chapter, the phrase “base pairing” was used without stating the expected “Watson-Crick (WC)/Hoogsteen (HG) pairing”. HG mode of binding is weaker than WC and in classical triplexes they are always induced in presence of stronger WC binding. Given the strength of binding in terms of T_m and K_D values and the vast literature precedence on base pairing in duplexes including that in modified nucleobases, it is imperative that the base pairings observed in Janus PNA complexes are indeed WC in nature. Further, like *aeg*-PNA, strand invasion is possible with Janus PNA as demonstrated by ethidium bromide displacement assay. Since there is no direct experimental proof (X-ray / NMR) in this work, it has been refrained from using the term “WC base pairing” although all evidences and lowering of T_m in single mismatched duplex strongly supports it. In such a case the novelty of Janus PNAs described here gets even more amplified as two WC duplexes are simultaneously formed from a single strand of Janus PNA in double duplexes. *There are no precedence of such complexes in literature*; significantly, formation of one duplex increases the stability of other duplex leading to new phenomenon of synergistic stabilization

of duplexes hosted on a single anchor Janus PNA strand. The importance of this property and possible application are outlined in the section “Summary of this and Outlook.”

6.11 Summary

To summarize, this chapter deals with the biophysical investigation of Janus PNA monomers and various types of Janus PNA oligomers using temperature dependent UV-visible spectroscopy, CD spectroscopy, Fluorescence spectroscopy and ITC. The *S* and *R* Janus PNAs having homo oligomeric C or G sequences on one face can form corresponding i- or G-tetraplexes with stability similar to that from *aeg*-PNA or DNA sequences. The differential binding ability of Janus PNAs oligomers to the complementary DNA was also investigated by ethidium bromide displacement and metal complexation study. The formation of duplex of tetraplex (i-motif and G4) are interesting and have scope for further development to functionally useful architectures.

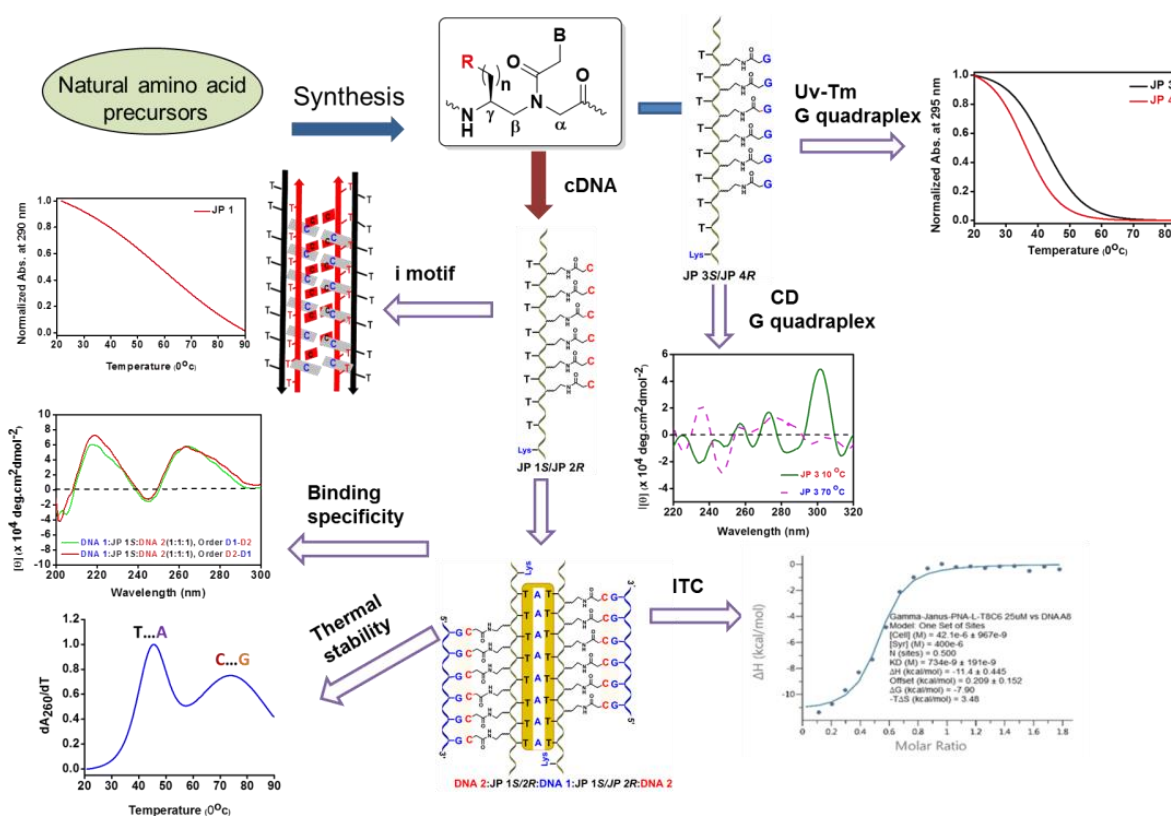


Figure 6.48 Summary of biophysical studies of Janus PNA oligomers

6.12 Experimental procedures

Chemicals: The unmodified and modified PNA oligomers were synthesized manually by general solid-phase PNA synthesis protocol using *Boc*-strategy as described in preceding

chapter. Complementary DNA oligonucleotides were obtained commercially from Integrated DNA Technologies (IDT). Salts and reagents used in buffer preparation such as NaCl, Sodium cacodylate etc. were obtained from Sigma-Aldrich. The pH of the buffer solutions was adjusted using NaOH or HCl, from Sigma Aldrich.

6.12.1 UV- T_m measurement

UV-melting experiments were carried out on Varian Cary 300 UV spectrophotometer equipped with a peltier. The samples for T_m measurement were prepared by mixing the calculated amounts of respective oligonucleotides in the stoichiometric ratio (1:1, duplex) in sodium Cacodylate buffer (10 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 2 μ M / 3 μ M for each strand. The samples were annealed by heating at 90 °C for 15 min. followed by slow cooling to room temperature for at least 10-12 h and then refrigerated for at least 24 h. The samples (500 μ L) were transferred to quartz cell and equilibrated at the starting temperature for 5 min. The OD at 260 nm was recorded in steps from 20-90 °C with temperature increment of 0.5 °C or 1 °C/min. Each melting experiment was repeated at least twice. The normalized absorbance at 260 nm or 295 nm was plotted as a function of the temperature. The T_m was determined from the first derivative of normalized absorbance with respect to temperature and is accurate to ± 1.0 °C. The data were processed using Microcal Origin 8.5. [The concentration of all oligonucleotides were calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases i.e. T = 8.8 cm²/ μ mol; C = 6.6 cm²/ μ mol; G = 11.7 cm²/ μ mol and A = 13.7 cm²/ μ mol]⁴⁰

6.12.2 Circular Dichroism

CD spectra were recorded on JASCO J-815 spectropolarimeter connected with a peltier. The calculated amounts of PNA/*Janus* PNA oligomers and the complementary DNA were mixed together in stoichiometric ratio (1:1 for duplex) in sodium cacodylate buffer (10 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 5 or 10 μ M for each strand. For Gquadruplex study, the samples were prepared in sodium cacodylate buffer containing 100 mM KCl, at pH 7.2. *Janus PNA* and aegPNA (150 μ M)/strand. The samples were annealed by heating at 90 °C for 15 min. followed by slow cooling to room temperature for at least 12 h. The cooled samples were transferred to refrigerator for at least 10 to 12 h. To record the CD spectra of PNA:DNA duplexes and single stranded PNAs, the temperature of circulating water was kept at 10°C. The CD spectra were recorded as an accumulation of 3

scans from 300 to 190 nm using 2 mm cell, a resolution of 0.1 nm, band-width of 1 nm, sensitivity of 2 m deg, response of 2 sec and a scan speed of 100 nm/min.

6.12.3 UV-visible and fluorescence studies

UV-Visible spectra for all PNA oligomers and complementary oligonucleotides were recorded on Perkin Elmer Lambda-45 UV-Visible Spectrophotometer and for fluorescence studies, experiments were performed on Horiba Jobin Yvon Fluorolog 3 spectrophotometer. The samples for fluorescence spectra were prepared by mixing calculated amounts of PNA and DNA in stoichiometric ratio (1:1, for duplex) in sodium cacodylate buffer (10 mM) and NaCl (10 mM); pH 7.2 to achieve a final strand concentration of 5 μ M for each strand. The annealed samples were used to record fluorescence spectra in a rectangular quartz cell at ambient temperature with λ_{exc} 458 and 484 nm; λ_{em} 520 nm; excitation slit width of 1 nm and emission slit width of 6 nm.

6.12.4 Ethidium bromide displacement assay from dsDNA

In case of ethidium bromide displacement, calculated amounts of DNA strands were mixed together to get a DNA duplex (5 μ M, 200 μ L) in the above said buffer system. The annealed samples were titrated against EtBr (250 μ M) in the increment of 1 μ L and the fluorescence intensity at 610 nm was recorded to achieve saturation point which was found to be 8 μ L, 5.58 μ M EtBr. The resulting EtBr-dsDNA complex was titrated individually against various PNA oligomers in the aliquots of 0.4 μ L and change in fluorescence intensity at 610 nm was recorded. The fluorescence decay observed upon displacement of EtBr was plotted against the wavelength of PNA oligomer using Microcal Origin 8.5.

Mass Spectrometry: Model Agilent 6540 UHD QTOF MS; Source ESI; Gas Temp ($^{\circ}$ C) 325; Gas Flow (l/min) 8; Nebulizer (psig) 30; Sheath Gas Temp ($^{\circ}$ C) 300; VCap (V) 2500; Nozzle Voltage (V) 0; Fragmentor (V) 200; Skimmer1: 45; Mass range (m/z) 100-3200; Scan speed (Hz) 1; Data type Profile; Software Masshunter workstation software v. B.05.01

Chromatography: Model Agilent Binary LC 1260; Column No column –Directly injected through UPLC; Mobile phase: A) Water (0.1% Formic Acid); B) Acetonitrile (LC-MS grade, J.T. Baker); Program type Isocratic 50% B, 0.2 mL/min; Injection volume 20 μ L; Injection wash with methanol.

6.12.5 Isothermal Titration Calorimetry (ITC)

Thermodynamic properties such as enthalpy, entropy and binding constants of *Janus* PNA complexation to cDNA were determined using isothermal titration calorimetry (ITC). The binding interactions between *Janus* PNAs and their complementary DNA oligonucleotides were carried out on Malvern MicroCal PEAQ ITC instrument. All titration experiments were performed at 15 °C in 10 mM sodium cacodylate buffer (pH 7.2) containing NaCl (10 mM) using same buffer solution to prepare all the solutions used in the experiment. The sample cell was loaded with *Janus* PNA solution and the reference cell contained only the buffer.

For determination the binding stoichiometry ratio for *homo Janus* PNA, each experiment performed at constant concentration (81 μM), the syringe was loaded with DNA solution (40 μL) 220 μM to 1.5 mM depend upon experiment (concentration is given in terms of nucleobase for comparison) depending upon experiments. The instrument was equilibrated at 15 °C until the baseline was flat and stable. *Janus* PNA was taken in the cell for binding experiment with cDNA to each face and total of 19 injections (2 μL /injection) were performed into a solution of *Janus* PNA in the cell.

N_b (nucleobase pair ratio) obtained from the ratio of number of nucleobases involved in such a complex formation.

6.13 References

1. Ranjbar, B.; Gill, P. *Chem. Biol. Drug Des.* **2009**, *74*, 101-120.
2. Nakano, S. I.; Kanzaki, T.; Sugimoto, N. *J. Am. Chem. Soc.* **2004**, *126*, 1088-1095.
3. Leavitt, S.; Freire, E. *Curr. Opin. Struct. Biol.* **2001**, *11*, 560-566.
4. Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y.H.; Goto, Y.; Refregiers, M.; Kardos, J. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E3095-E3103.
5. Cain, B. F.; Baguley, B. C.; Denny, W. A. *J. Med. Chem.* **1978**, *21*, 658-668.
6. Kantonen, S. A.; Henriksen, N. M.; Gilson, M. K. *Biochim. Biophys. Acta, Gen. Subj.* **2017**, *1861*, 485-498.
7. Sturtevant, J. M.; *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 2236-2240.
8. Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *19*, 213.
9. Barbieri, C. M.; Li, T. K.; Guo, S.; Wang, G.; Shalloo, A. J.; Pan, W.; Yang, G.; Gaffney, B. L.; Jones, R. A.; Pilch, D. S. *J. Am. Chem. Soc.* **2003**, *125*, 6469.
10. Kunne, A.; Sieber, M.; Meierhans, D.; Allemann, R. K. *Biochemistry* **1998**, *37*, 4217.
11. Dhakal, S.; Yu, Z.; Konik, Cui, R. Y.; Koirala, D.; Mao, H. *Biophys J* **2012**, *102*, 2575-2584.
12. Laisne, A.; Pompon, D.; Leroy, J.-L. *Nucleic Acids Res.* **2010**, *38*, 3817-3826.
13. Kim, S. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477-6481.
14. (a) Krishnan-Ghosh, Y.; Stephens, E.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 5944-5945. (b) Dapic, V.; Abdomerović, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic acids Res.* **2003**, *31*, 2097-2107.
15. (a) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, Gabelica, E. V. *J. Am. Chem. Soc.* **2008**, *130*, 10208-10216. (b) Huppert, J. L. *The FEBS Journal* **2010**, *277*, 3452-3458. (c) Risitano, A.; Fox, K. R. *Nucleic Acids Res.* **2004**, *32*, 2598-2606. (d) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2004**, *126*, 16405-16415. (e) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. *J. Am. Chem. Soc.* **2010**, *132*, 17105-17107. (f) B. Heddi, A. T. Phan, *J. Am. Chem. Soc.* **2011**, *133*, 9824-9833. (g) Hansel, R.; Lohr, F.; Foldynova-Trantirkova, S.; Bamberg, E.; Trantirek, L.; Dotsch, V. *Nucleic Acids Res.* **2011**, *39*, 5768-5775. (h) Guedin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. *Nucleic acids Res.* **2010**, *38*, 7858-7868.
16. Scaria P.V.; Shafer R. H. *J. Biol. Chem.* **1990**, *266*, 5417-5423.

6.14 Appendix II

CD spectroscopic studies of Janus PNA:DNA duplexes

The conformation of PNA:DNA, PNA:RNA and PNA:PNA duplexes have been well studied using CD spectroscopy. While the classical A and B forms of DNA/RNA duplexes contain distinct major and minor grooves the conformations of PNA-derived duplexes with DNA²⁸ and RNA are slightly different. PNA:PNA duplexes correspond to the P-helical form in which the distinction of major and minor grooves is not that apparent. Each of these conformations have distinct CD spectra. The effect of chiral C^γ-side chain substitution on the conformation of *Janus* PNA:DNA duplexes was studied by CD spectroscopy. The *homo Janus* PNAs, *Hetero Janus* PNAs, *aeg-C^γ* oligomers and control PNAs exhibited very low induced CD signals. (Figure 6.76 to 6.79)

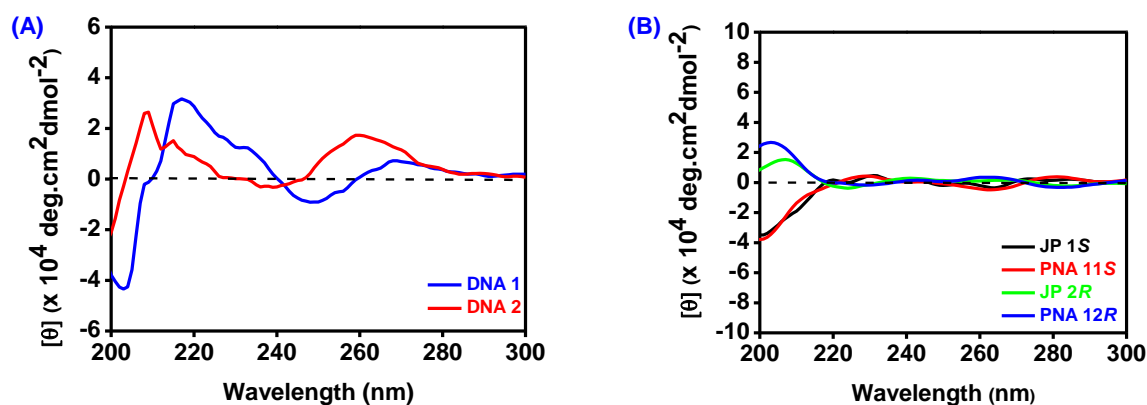


Figure 6.49 CD spectra of (A) *ss* DNAs: DNA 1 (dA₈), DNA 2 (dG₆) (B) *ss*PNAs: JP 1 (*S* T₈jp-γ-C₆), JP 2 (*R* T₈jp-γ-C₆), PNA 11 (*S* p₈-γ-C₆), PNA 12 (*R* p₈-γ-C₆); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM)

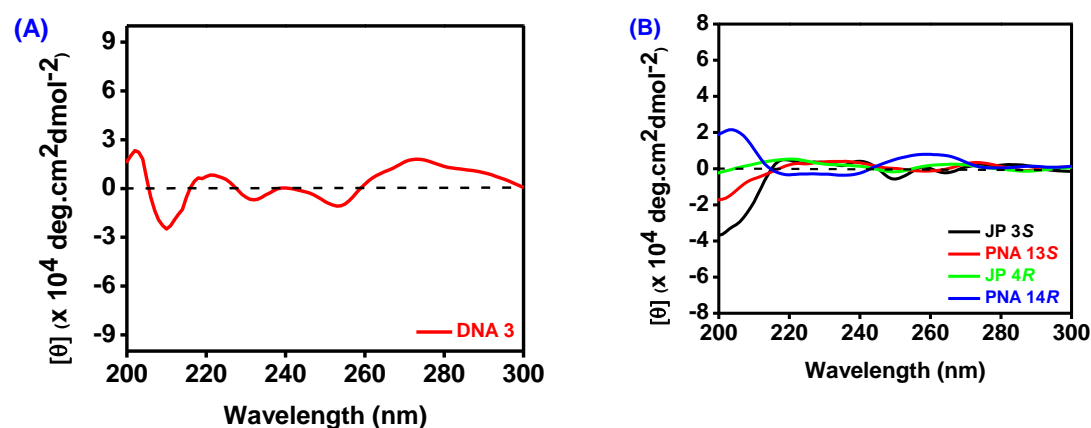


Figure 6.50 CD spectra of (A) *ss* DNAs: DNA 3 (dC₆) (B) *ss*PNAs: JP 3 (*S* T₈jp-γ-G₆), JP 4 (*R* T₈jp-γ-G₆), PNA 13 (*S* p₈-γ-G₆), PNA 14 (*R* p₈-γ-G₆); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM)

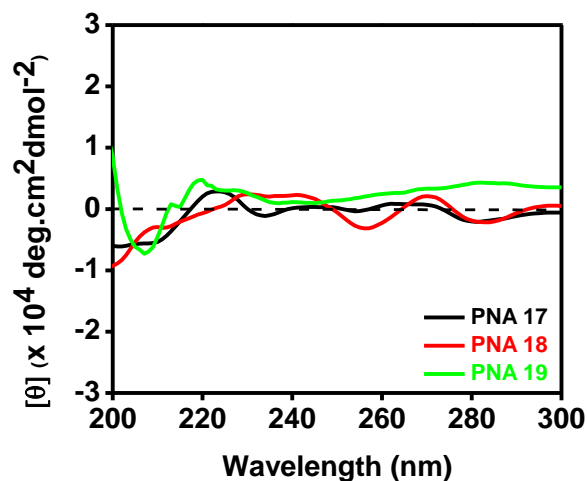


Figure 6.51 CD spectra of ssPNAs: PNA 17 (*PNA-T₈*), PNA 18 (*PNA C₆*), PNA 19 (*PNA G₆*); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM)

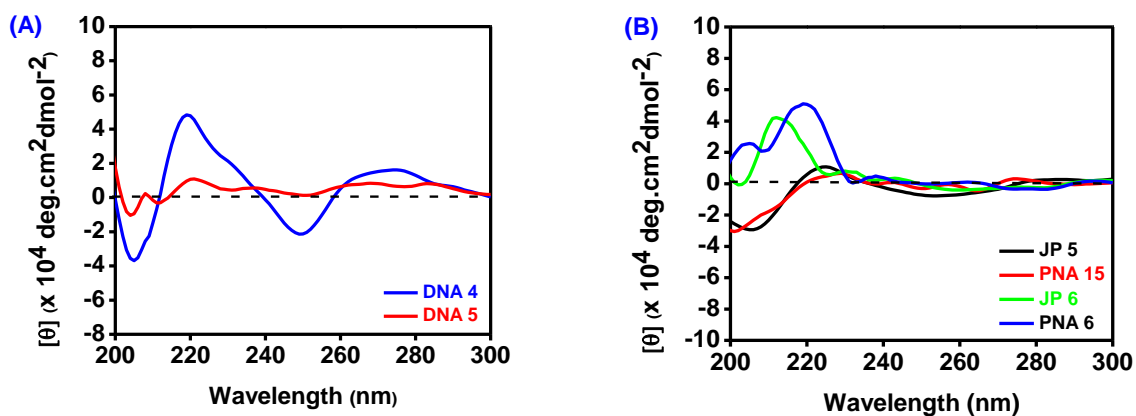


Figure 6.52 CD spectra of (A) ss DNAs: (A) DNA 4, DNA 5 (B) ssPNAs: JP 5 (*S mjp-γ-PNA*), JP 6 (*R mjp-γ-PNA*), PNA 15 (*S p8-γ-mjp*), PNA 16 (*R p8-γ-mjp*); Buffer: 10 mM sodium cacodylate, pH 7.2, NaCl 10 mM)

Summary of thesis and outlook

This thesis presents a new and novel type of PNA structure termed “Janus PNA” in which each PNA unit consists of an additional side chain at C^γ position which carries a nucleobase through a spacer having amide link. Thus the PNA backbone of Janus PNA is endowed with two nucleobase sequences on two sides of the backbone termed “t-amide face” and “ C^γ -face”. Since C^γ becomes chiral with a side chain substitution, both *S* and *R* stereomeric PNA oligomers were synthesized to examine the effect of chiral centre. Thus, the C^γ -Janus PNA can hybridize with two cDNAs, one from the amide face and the other from the C^γ -face leading to formation of double duplex. Based on the nature of sequences on each face they have been classified into different types: homo Janus PNAs ($T_{7jp-\gamma(S/R)-C_5}$, homo oligomeric sequences on both face), mixed Janus PNA ($mjp-\gamma(S/R)-C_5$, mixed sequences on amide face and homo oligomeric sequence on C^γ -face) and hetero Janus PNA (mixed sequence on both face) These were synthesized on solid phase manually using appropriate protected monomers, cleaved from resin, purified and characterized by mass spectra.

The complex formation of different Janus PNAs was studied by UV-T plots and observation of single sigmoidal transitions suggested successful duplex/triplex formation. The $C^\gamma(S/R)$ PNA sequences having only nucleobases conjugated via C^γ -side chain via amide linker and devoid of tertiary amide linked triazole side chain (as in standard *aeg*-PNA) formed duplex with cDNA, indicating individual capability of these new C^γ -PNAs to base pair with cDNA. The different Janus PNAs having base sequences on both amide and C^γ -faces were hybridized with appropriate cDNAs to generate duplexes (amide face/ C^γ - faces), triplexes and double duplex of triplex (with homo Janus PNAs), and double duplexes (hetero Janus PNAs). The T_m s of each of these duplexes were measured by temperature dependent UV absorbance. In isolated individual duplexes with one cDNA binding, it was found that the C^γ -side duplex always had a higher T_m than t-amide face duplex. In complexes composed in presence of both cDNAs, double duplexes were observed with simultaneous binding on both sides to duplexes on both faces. These gave two transitions and hence two T_m s corresponding to melting of each duplex. Three interesting observations: (i) The T_m of both amide and C^γ -duplex in the double duplex were much higher than analogous T_m s in individual isolated duplexes and (ii) the C^γ - duplexes always had higher T_m than amide face duplex and (iii) the *S*-stereomeric oligomer always had higher thermal stability than *R*-stereomeric oligomer. To

establish the sequence dependence of base pair formation, all individual and Janus double duplexes were constituted from mismatched DNA sequences. These exhibited lower T_{ms} than the perfect duplexes indicating the sequence dependence of duplex formation.

The conformational aspects of duplex and triplex formation from Janus PNA was studied by CD spectroscopy. The derived duplexes and triplexes generated by stoichiometric additions of Janus PNA and DNA components exhibited CD profiles characteristic of PNA:DNA duplexes and triplexes confirming similar conformation. The double duplex of Janus PNA can be generated in two ways depending on the order of duplex formation. CD studies of sequential formation of duplexes indicated that irrespective of the order, the final duplex always has same conformation.

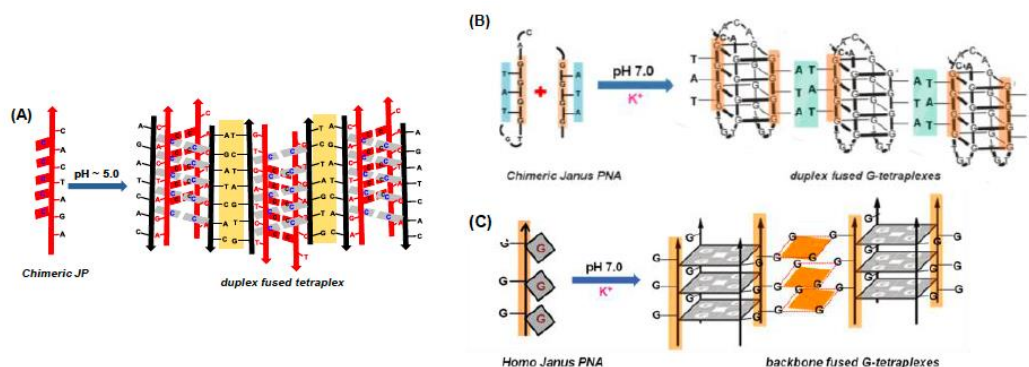
The thermodynamic parameters for binding reactions of triplexes from Janus PNAs with cDNA were obtained from Isothermal Titration Calorimetry. The stoichiometric and cooperative nature of binding reactions were indicated by sigmoidal nature of binding isotherm. The enthalpy and free energy of various binding reaction were negative and quantitatively similar to that seen in standard *aeg*-PNA:DNA complexes. The stoichiometry N of complexation in different experiments computed in terms of ratio of number of PNA and DNA bases involved in base pair formation and corresponding to the proposed triplex matched with that obtained by the experiment. This further supported the formation of triplexes through specific base pairing compositions to yield various complexes. The combined biophysical results established the proof of concept of design of Janus PNAs and demonstrated their ability to simultaneously bind to cDNAs from both sides to form double duplexes.

Nature of base pairing: The overall work in the thesis clearly confirms interaction of $C^{\gamma}(S/R)$ Janus PNAs from both sides to form sequence specific base pairing through complementary H-bonding. Based on all literature of H-bonding among natural nucleobases, it is tempting to suggest that all base pairing observed in this work are Watson-Crick base pairing. The other not so prevalent form is Hoogsteen base pairing. There are not many cases where isolated Hoogsteen base pairing occurs in solution though it is well characterized in crystals. They are much weaker than canonical Watson-Crick base pairing and in DNA and PNA triplexes, they occur only when the bases are paired by Watson-Crick H-bonding on one face and not in isolation. These facts clearly point out that in $C^{\gamma}(S/R)$ -Janus PNAs, all base pairing seen among various duplexes are Watson-Crick in motif. It can only be established by

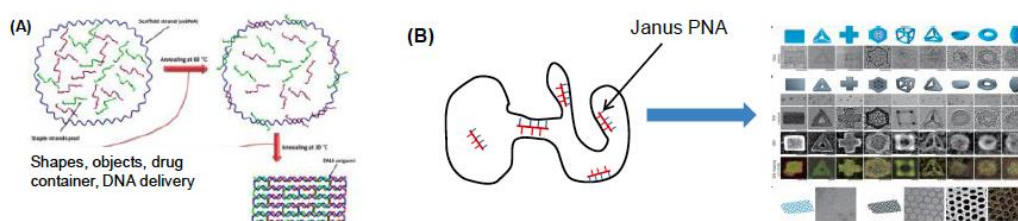
NMR spectroscopic studies or X-ray crystal structures, which need enough materials. Hopefully in future this can be achieved. Hence in this thesis it has been refrained from the use of “Watson-Crick H-bonding.”

Future outlook: If base pairing in Janus PNA is by Watson-Crick H-bonding, which is most possibly true, the double duplex formation by C^γ-Janus PNAs by simultaneous recognition of a single strand Janus PNA by two complementary DNA strands, both by sequence specific Watson-Crick H-bonding is the most novel aspect of the work. This would certainly open up new biological applications for Janus PNAs.

- (i) When Janus PNAs form i-motif or G₄-tetraplexes, on the backbone another set of sequences are available for further growth of self-assembly for higher order structures as shown in the following figures. One can generate tetraplex of tetraplexes and contiguous i-motif structures. If the sequence is a mixed sequence, one can generate duplexes of i-motif or tetraplexes. Given the high T_m and robustness of these structures, these may have interesting material properties.

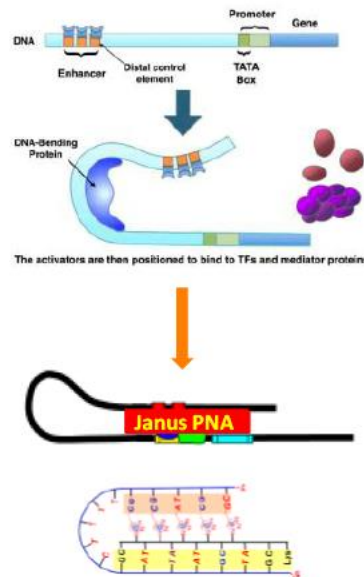


- (ii) Janus PNAs can also replace DNA staples in DNA origami. Engineering ordered staples to complement specific regions in plasmid DNA, they can systematically generate various shaped structures by DNA origami.

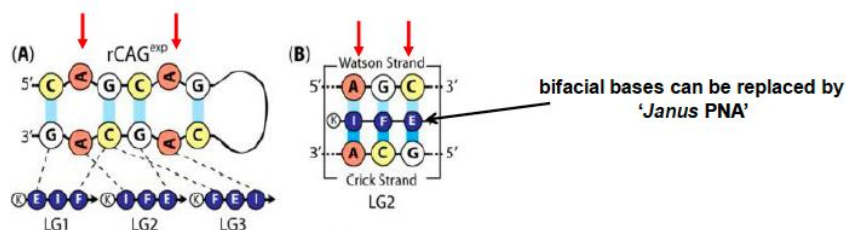


- (iii) There could be specific biological applications where a single probe may be necessary to simultaneously target two different genes and Janus PNAs would be

very versatile. Availability of such molecules may generate new applications as well. One potential application would be in Janus PNA or bundles or self-aggregates simplifying the enzyme machinery in eukaryotic transcription, with reference to enhancer elements. Janus PNAs can “in principle” bring together the relevant parts of distantly located DNA regions together to modulate gene expression. This could be a far-fetched application, availability of Janus PNA type of molecules may generate simpler applications.



- (iv) Recently Danith Ly et al have published a nice application of control of disease related RNA hairpins using Janus bases. Janus PNAs could also be employed in such application.



Thus the designed Janus PNAs and results presented in this thesis offers potential biological and material applications in future.