Structural Control of Peptide Nucleic Acid Rotamers: Aza-PNA and Optical Properties of Gemdimethyl PNA, Bimodal PNA Monomers and Oligomers

A Thesis

Submitted in partial fulfilment of the requirements

of degree of

Doctor of Philosophy

By

Abdul Shiraj

ID: 20163469

Indian Institute of Science Education and Research, Pune

June 2022

This thesis dedicated to….

My Supervisor

And

My Spouse

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान, पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH (IISER), PUNE
(An Autonomous Institution, Ministry of Human Resource Development, Govt. of India)
900 NCL Innovation Park, Dr. Homi Bhabha Road, Pune 411008

Prof. Krishna N. Ganesh
FNA, FASc, FTWAS
Director, IISER-Tirupati
Co-Editor: ACS Omega
Honorary faculty, IISER Pune

CERTIFICATE

Certified that the work incorporated in the thesis entitled **Structural control of Peptide Nucleic Acid Rotamers: Aza-PNA, Bimodal PNA and Optical Properties of Gemdimethyl PNA monomers and oligomers** submitted by Mr. Abdul Shiraj 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: 29th June 2022, Pune

Prof. Krishna N. Ganesh
(Research Supervisor)

Kn.ganesh@iiserpune.ac.in • Ph: +91 (0877) 2500 202 • Fax: +91 (0877) 2500 250

DECLARATION

I declare that this written submission represents my ideas in my own words and whether other's ideas have been included, I have adequately cited and referenced the original source. I also declare that I have adhere 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 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: 29th June 2022, Pune

Mr. Abdul Shiraj

ID: 20163469

Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor Prof. Krishna N. Ganesh for his constant guidance and immense support during my PhD tenure in IISER-Pune, without which it would have been impossible for me to complete the PhD work in 5-year tenure. Encouragement and support during the conversations with Prof. K. N. Ganesh in the tough time of my PhD always given me the strength to tackle the obstacles and challenges of scientific research. It is always a pleasure for me to learn presentation skills, communication skill, public speaking ability etc from Prof. Ganesh during his talks in group meeting and in numerous occasions. Apart from that it is always delighted to learn research project problem solving approach, scientific writing, scientific ethics, academic honesty, and research project designing approach from him, for which I will be grateful to him for the rest of my scientific career and life as well.

I would like to express my sincere thanks to my RAC members Prof. H.N. Gopi and Dr. Moneesha Fernandes for their suggestions and important advice during several RAC meetings. The questions asked during RAC meeting by Prof. H.N. Gopi and Dr. Moneesha Fernandes and the comments after the presentation helped me a lot to improve my scientific skills and to decide the future progress of my research project. Apart from that I would like to express my sincere thanks to Dr. S. G. Srivatson for his suggestions during my departmental seminar and pre-synopsis seminar. I am also thankful to all the faculty members of Department of Chemistry IISER-Pune specially Dr. Pinaki Talukdar, Prof. M. Jayakannan, Dr. Pramod Pillai and Dr. Partha Hazra for giving access to their instrument facilities.

The technical staffs of Department of Chemistry, IISER Pune has always been very helpful and kind during my PhD tenure in IISER. Nitin Dalvi, Dr. Sandeep Kumar Mishra, Dr. Sandeep Kanade, Suresh Prajapati, Ravinder Malothu was very help helpful to obtain good quality data during the PhD tenure. Specially Mahesh Jadav and Dnayaneshwar Ekshinge were very helpful for technical support for HPLC and other instruments.

I am very much thankful to my lab seniors Dr. Shahaji H More, Dr. Prabhakar Pawar, Dr. Manoj Gupta, Dr, Pramod Bhingardeve, Dr. Madhanagopal, Dr. Iranna Todkari for their help and suggestions during these years. Apart from that I am very much thankful to Dr. Dhrubjyoti Dutta for his constant help and suggestions during the early years of my PhD. I am also thankful to Dr. Gade Chandrasekhar (IISER-Tirupati), Dr. Om Shanker Tiwari, Rajat

Halemane S (IISER-Tirupati), Gaurav Borkar (IISER-Tirupati), Pradeep, Ashwini (IISER-Tirupati) for helping me on several occasions. During my PhD tenure I have enjoyed wonderful company with my friends T. Anand Kumar, Muzaffar Hussain, Prakash Warghure, Ankit, Subhajit, Korak Biswas, Dipayan, Puneeth, Ruma Ghosh, Padmini Sahoo, Neetu etc.

I am very much thankful to CSIR-New Delhi for providing fellowship during my 5-year PhD tenure and IISER-Pune for providing Half fellowship for the extended 6-month period.

Finally, I would like to express my deepest gratitude to my Parents, younger sisters, and my spouse. Their affection and unconditional love gave me enough strength which helped me to reach my scientific goals.

Abdul Shiraj

Publications

1. Accepted manuscript:

Abdul Shiraj, Raghunath O Ramabhadran* and Krishna N Ganesh*, Aza-PNA: Engineering E-Rotamer Selectivity through Extended γ-Turn Formation, *Organic Letters*

Manuscript ID: ol-2022-02993e.R2

Date of acceptance: 30.09.2022

2. Manuscript under preparation:

1**. Abdul Shiraj**, Krishna N Ganesh, Rotamer dependent fluorescence of α/β-*gemdimethyl*-*aeg*-Guanine monomers in solution and solid state.

2. **Abdul Shiraj**, Krishna N Ganesh, Effect of C_α-cyclisation on the rotamer preferences of *aeg*-PNA.

Symposia attended/Poster presentation

1. Attended the **A**sian **C**hemical **B**iology **I**nitiative-**2022** (**ACBI-2022**) conference on chemical biology (*14th September, 2022, IISER Pune*).

2. Presented poster in the **A**sian **C**hemical **B**iology **I**nitiative-**2022** (**ACBI-2022**) conference on chemical biology (*14th September, 2022, IISER Pune).*

Table of contents

Chapter 1: Introduction to Peptide Nucleic Acids

1.1. **Introduction to Nucleic Acids** 1

1.1.1 Base pairing and duplex structure of DNA 2

1.1.2 Secondary structure of DNA 4

1.1.3 Secondary structure of RNA 6

1.1.4 DNA triplexes and quadruplex structure 7

1.1.5 Application of Nucleic Acids 8

1.1.6 Chemical modification of DNA 11

1.2 **Peptide Nucleic Acid** 14

1.2.1 PNA duplex structure 16

1.2.2 Triple helix formation by PNA 17

1.2.3 Strand invasion by PNA 17

1.2.4 Sequence discrimination and other properties of PNA 20

1.2.5 Cell permeation of PNA 20

1.3 **Chemical modification of PNA** 21

1.3.1 Backbone modification of PNA 21

1.3.1a C_{α} modified PNA 22

1.3.1b C_{β} modified PNA 23

1.3.1c $C\gamma$ modified PNA 24

1.4 **Conformational preorganization of PNA (Cyclic PNA)** 26

1.5 **Application of PNA** 29

1.6 **Synthesis of PNA and bio-physical techniques** 30

1.6.1 Solid Phase PNA synthesis 31

1.6.2 Biophysical techniques 33

1.6.2a Temperature dependent UV and Jobs Plot 33

1.6.3a Circular Dichroism (CD) 36

1.6.4 Fluorescence and Isothermal Calorimetry (ITC) 37

1.7 **Present work** 38

1.8 References 44

Chapter 2: E-Rotamer Selectivity in N_{α}-Aza-Peptide Nucleic Acid Monomers via 8-membered γ-Type Turn Formation

2.1 **Introduction** 55

2.1.1 PNA rotamers: Crystal structure of PNA monomer 59

2.2 **Rationale and objective** 63

2.2.1 Specific objectives 65

2.3. **Result and Discussion** 65

2.3.1 Synthesis of α-*aza*-PNA-T 65

2.3.2 Synthesis of β-*ala*-*aza*-PNA-T 66

2.3.3 Analysis of ^{1}H NMR spectra of monomers 67

2.3.4 Assignment of backbone protons and rotamers of *aza*-PNA 69

2.3.5 Effect of protecting group for the stabilization of rotamer 77

2.3.6 Solvent dependent NMR and NMR experiments 79

2.3.7 Temperature dependent NMR shift and NMR experiments 82

2.3.8 Crystal structure of α-*aza* chloro intermediate **7** 86

2.3.9 Infrared spectroscopy data of *aza*-PNA monomers 89

2.3.10 Computational studies of *aza*-PNA rotamers 92

2.3.11 Example of *aza*-modification in biomolecules 100

2.4 **Conclusion** 101

2.5 **Summary** 103

2.6 **Experimental section** 103

2.6.1 General information 103

2.6.2 Experimental procedure 104

2.7 **References** 117

2.8 **Appendix-I** 121

Chapter 3: α/β-*Gemdimethyl(gdm)* Guanine PNA monomers and dimers-Rotamer dependent fluorescence

3.1 **Introduction** 154

3.1.1 Introduction to luminescence 154

3.1.2 Fluorescence and Phosphorescence 155

3.1.3 Fluorescent nucleobases 156
3.1.4 Canonical fluorescent purine nucleobases 157
3.1.5 Canonical fluorescent pyrimidine nucleobases 158
3.1.6 Fluorescent compounds lacking conventional fluorophores 159
3.1.7 Origin of fluorescence in *clusteroluminogens* 161
3.2 **Rationale and objective** 163
3.3 **Result and discussion** 166
3.3.1 Synthesis of α-*gemdiemthyl*-*aeg*-G/C PNA monomer 166
3.3.2 Synthesis of β-*gemdiemthyl*-*aeg*-G PNA monomer 167
3.3.3 Synthesis of fatty acid conjugated and free amine β-*gemdimethyl* monomer 167
3.3.4 Synthesis of fatty acid conjugated and free amine α-*gemdimethyl* monomer 168
3.3.5 Synthesis of dimers 169
3.3.6 Conformational study of the *gemdiemthyl* monomers 170
3.3.7 Fluorescence spectroscopy of the *gemdimethyl* Guanine monomers 174
3.3.8 Fluorescence study of *gemdimethyl* Pyrimidine monomers 179
3.3.9 Effect of fatty acid conjugation on fluorescence 182
3.3.10 Langmuir-Blodgett monolayer formation 185
3.3.11 Red Edge Excitation Shift effect 188
3.3.12 Fluorescence emission of dimers 192
3.3.13 Crystal structure of the Chloro intermediate of α-*gemdimethyl* **11** 194
3.4 **Discussion** 197
3.5 **Summary** 199
3.6 **Experimental procedure** 201
3.6.1 General information 201
3.6.2 Experimental procedure and spectral data 202
3.7 **References** 212
3.8 **Appendix-II** 218

Chapter 4: Optical properties of Bimodal PNA monomers and oligomers

4.1 **Introduction** 246

4.2 **Rationale and objective** 250

4.3 **Result and discussion** 251

4.3.1 Synthesis of *bimodal* PNA monomer 251

4.3.2 Solid phase PNA synthesis 252

4.3.3 Rotamer preferences of *bimodal* PNA monomer 253

4.3.4 UV-Vis and fluorescence spectra of *bimodal* PNA monomer 255

4.3.5 UV and fluorescence of oligomer 256

4.4 **Discussion** 258

4.5 **Summary** 259

4.6 **Experimental section** 259

4.6.1 General information 259

4.6.2 Experimental procedure and spectra 261

4.7 **References** 269

4.8 **Appendix-III** 272

Chapter 5: Synthesis and characterization of succinimide analogues of Peptide Nucleic Acid

5.1 **Introduction** 285

5.2 **Biologically active succinimide derivatives** 286

5.3 **Rationale and objective** 287

5.4 **Result and discussion** 288

5.4.1 Synthesis of succinimide modified PNA monomer 288

5.4.2 NMR studies of the succinimide modified monomer 290

5.4.3 Crystal structure of succinimide modified PNA monomer 292

5.5 **Discussion** 296

5.6 **Summary** 297

5.7 **Experimental section** 297

5.7.1 General information 297

5.7.2 Experimental procedures and spectra 299

5.8 **References** 302

5.9 **Appendix-IV** 303

Summary and outlook of the thesis 309

Abbreviations

A	Adenine
AcOH	Acetic acid (glacial)
ACN	Acetonitrile
aeg	amino ethyl glycine
aq.	Aqueous
$(Boc)_2O$	Boc anhydride
Bn	Benzyl
C	Cytosine
Calcd	Calculated
Cbz	Benzyloxy carbonyl
CD	Circular Dichroism
CHCA	α-Cyano-4-hydroxy cinnamic acid
COSY	Homonuclear corelation spectroscopy
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DHB	2,5-dihydroxybenzoic acid
DhbtOH	3-Hydroxy-1,2,3-benzotriazin-4(3H)-one
DIPEA	N, N-Diisopropylethylamine
DMF	N, N- Dimethylformamide
DMSO	N, N-Dimethyl sulphoxide
DNA	2'-deoxyribonucleic acid
ds	Double stranded
Et	Ethyl
EtOAc	Ethyl acetate
g	gram
G	Guanine
gly	Glycine
h	hours

HBTU	2-(1-H-Benzotriazole-1-yl)-1,1,3,3-tetramethyl-uronum-hexafluoro-phosphate
HOBt	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
IR	Infra-red
Lys	Lysine
MALDI-TOF	Matrix Assisted laser Desorption Ionization- Time of Flight
MBHA	4-Methyl benzhydryl amine
mg	milligram
MHz	Megahertz
µL	Microliter
µM	Micromolar
min	minutes
ml	milliliter
mM	millimolar
mmol	millimoles
N	Normal
nm	nanometer
NOSEY	Nuclear Overhauser effect spectroscopy
NMR	Nuclear Magnetic Resonance
ODNs	oligonucleotides
ppm	Parts per million
PS-oligo	Phosphorothioate-oligo

PNA	Peptide Nucleic Acid
PCR	Polymerisation Chain Reaction
R_f	Retardation factor
RNA	Ribonucleic acid
RP	Reverse Phase
rt	room temperature
RT	Retention time
SPPS	Solid Phase Peptide Synthesis
TEA/Et_3N	Triethyl amine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethyl sulphonic acid
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TOCSY	Total Corelation Spectroscopy
T_m	Melting temperature
UV-Vis	Ultraviolet-visible

Abstract

Structural Control of Peptide Nucleic Acid Rotamers: Aza-PNA and Optical Properties of Gemdimethyl PNA, Bimodal PNA Monomers and Oligomers

The thesis entitled **"**Structural control of Peptide Nucleic Acid (PNA) rotamers and their fluorescence properties: *aza*-PNA, *gemdimethyl* G-PNA and Cα(*S*)-*bm* PNA monomers and oligomers" describes studies done on the synthesis, characterisation and rotamer study of novel *aza*-PNA monomers, α/β-*gemdimethyl* guanine PNA, Cα(S)-(acetamido)-*bm* PNA monomers and oligomers. The effect of intramolecular H-bonding in Cα-NH (*aza*-PNA) and *gem-dimethyl* substitution induced conformational constraint on rotamer selectivity and the fluorescence properties are explored in this work. The thesis is presented in five chapters

Chapter 1: Introduction to PNA

This chapter gives an overview of DNA/RNA duplex, triplex and tetraplex structure, chemically modified DNA and Peptide nucleic acid. Synthetically modified oligonucleotide (ODNs) of DNA and RNA are capable of forming Watson-Crick hydrogen bonded duplex between the two complementary strands as natural DNA. Polypyrimidine sequences of DNA/RNA oligomers are capable of forming triplexes by combination of Watson-crick and Hoogsteen hydrogen bonding among the three composed strands. The specificity in complementary base pairing to forming stable duplexes and triplexes make synthetic DNA an ideal candidate to suppress undesirable genes and is the basis of new medicinal chemistry concept of antigene and antisense therapy. Natural oligonucleotides (DNA/RNA) lack essential requirements of nuclease stability and cell permeability to be successful antisense agent and

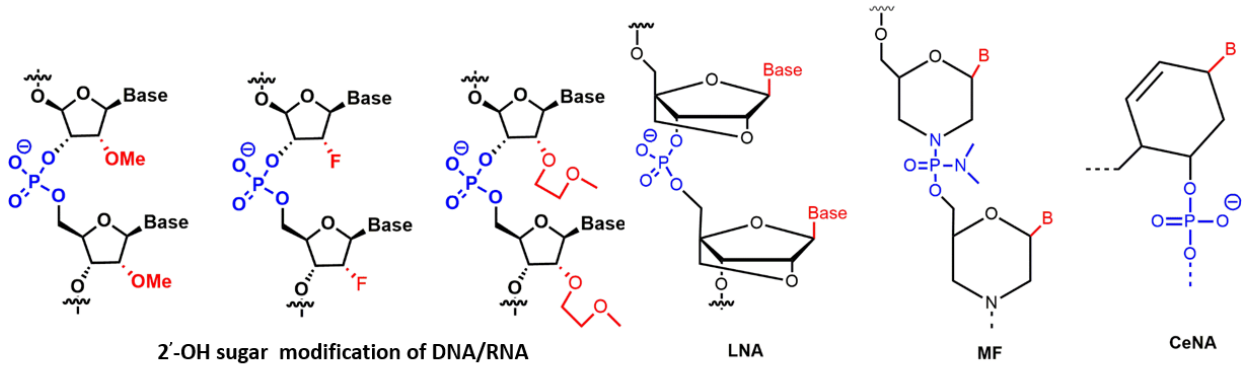


Figure 1. Backbone modified analogues of DNA/RNA

hence chemical modification is necessary for successful and rational design of antisense oligonucleotides. This chapter summarises common chemical modifications (Figure 1), leading to a versatile type of chemical modification – Peptide Nucleic acid, which is the focus of this thesis.

Peptide Nuclei Acid (PNA) acids are achiral synthetic mimic of DNA where the negatively charged sugar phosphate backbone of DNA is replaced by neutral amino ethyl glycine (*aeg*) unit (Figure 2). The nucleobases (A/T/G/C) are separated by six atom spacers in PNA similar as in DNA and the backbone is attached to the nucleobase by a two-atom methylene carbonyl spacer. PNA binds with complementary DNA/RNA and PNA to form stable duplexes with higher binding affinity compared to that of DNA/RNA duplexes. Polypyrimidine PNAs forms PNA_2:DNA triplexes. This interesting property makes PNA a perfect candidate for its application as antisense oligonucleotide. Unsubstituted *aeg*-PNA suffers from major drawbacks including low aqueous solubility, poor cell permeability and ambiguity of binding towards parallel and antiparallel DNA/RNA etc, which restrict its use as an antisense oligonucleotide and hence modification of the PNA backbone is necessary to improve the properties of PNA compared to the parent *aeg*-PNA oligonucleotide. This chapter gives a detailed overview of the advancement in the field of Peptide Nucleic Acid including chemical modifications and explains the rationale and aim of the works reported in this thesis.

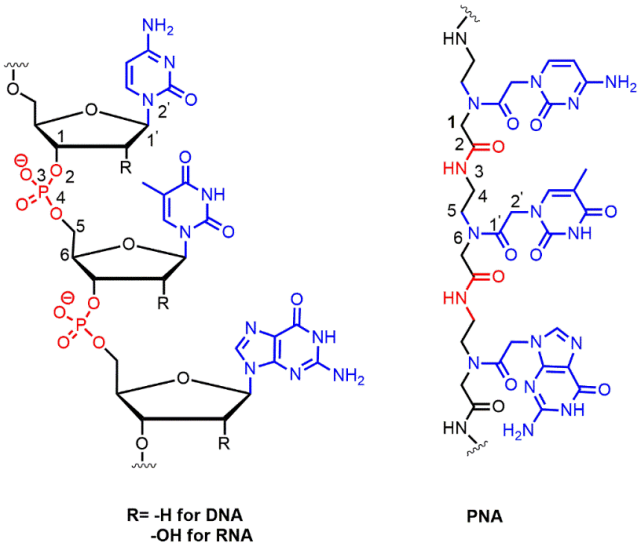


Figure 2. Basic chemical structure of DNA and PNA

Most chemical modifications are centered around the backbone of PNA and the role of central tertiary amide carbonyl group of PNA in the PNA: DNA binding has not been explored. Due to restricted rotation around N-CO bond, PNA monomers exist in two rotameric forms –

Z and E. The work presented in this thesis is about a new PNA analogue "*aza*-PNA" in which the α-methylene is substituted by NH and the reported *gemdimethyl* PNA, where α-methylene is replaced by *gemdimethyl* group (Figure 3). It is demonstrated that both stabilize the unusual E-rotamer which affects the conformation of PNA backbone and induce intramolecular folding in *aza*-PNA and fluorescence in *gemdimethyl* PNA.

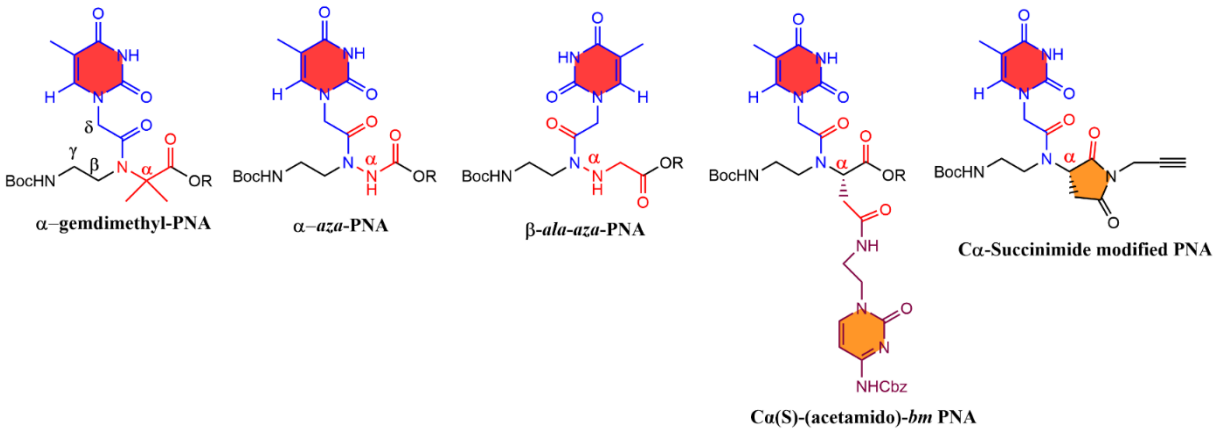


Figure 3. Modified PNA monomers reported in this thesis to explain the role of central tertiary amide carbonyl group

Chapter 2: *aza*-PNA monomers and study of rotamers

Analysis of the single crystals of PNA monomers and oligomers reveal preferential stabilization of Z-rotamer in solid state but a mixture of Z and E rotamer in solution state. Recently it was reported that incorporation of *gemdimethyl* group at the β and α position of the *aeg*-PNA monomer selectively stabilize E and Z rotamer in solution by avoiding steric clash of the *gemdimethyl* group with δ-methylene of the PNA backbone and the favorable C-H…O interaction in E-rotamer. Replacing the α-methylene by NH group at the α-position on the PNA backbone gives *aza*-PNA which can induce folding of the backbone via intramolecular hydrogen bonding with the N-terminal carbonyl (Figure 4). This leads to extended γ-turn formation in which the tertiary amide carbonyl group is locked up in E-rotamer form. This chapter reports synthesis of *aza*-PNA monomer (Schemes 1), β-*ala*-*aza*-PNA monomer (Scheme 2) and study of rotamer selectivity by NMR and validation of the results by theoretical calculations (Figure 5), comparative ^{13}C NMR and X-Ray crystallography. It is seen from ^{1}H-^{1}H NOESY NMR and solvent dependent studies that the preferred rotamer in *aza* PNA monomers is E-rotamer, stabilized by intramolecular H-bonding and supported by DFT computation and X-ray structure of a precursor intermediate.

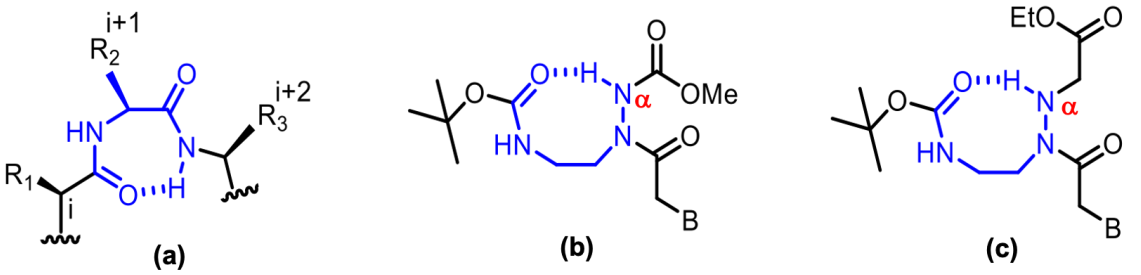
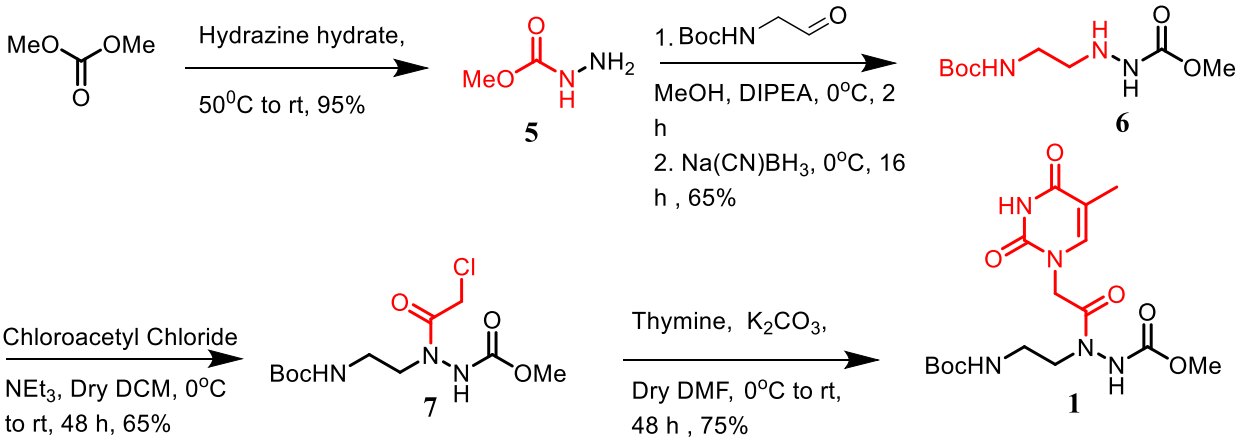
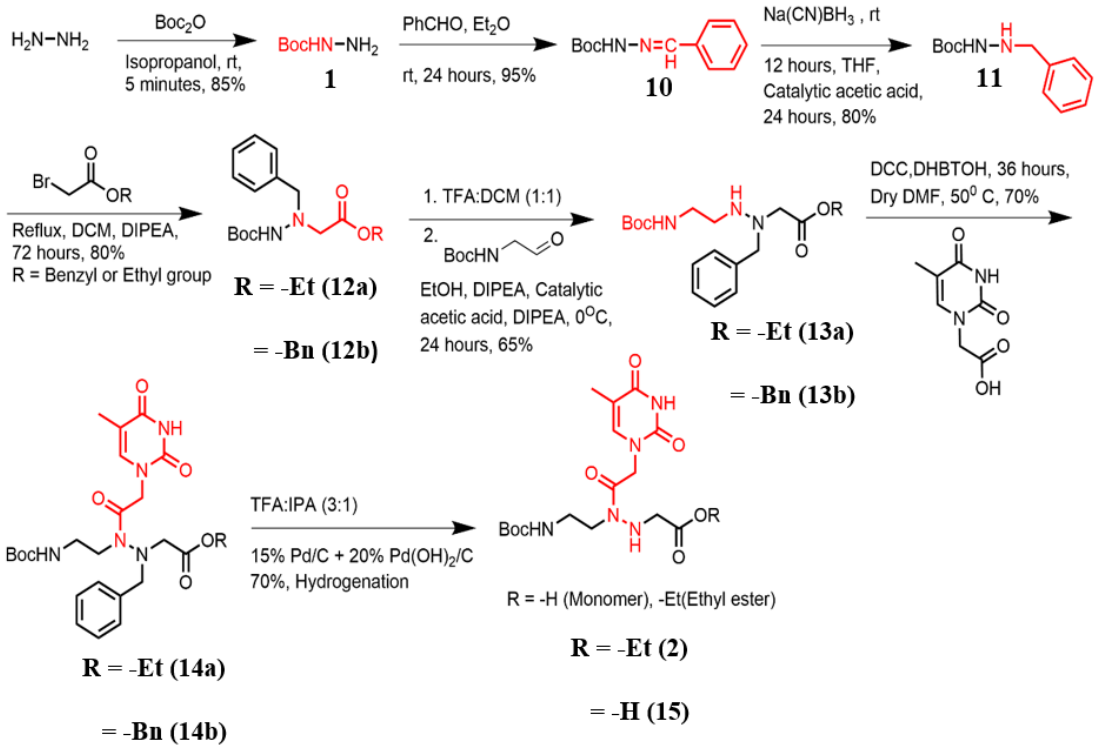


Figure 4. (a) γ-turn peptides (b) *aza*-PNA monomer (c) b-*ala*-*aza*-PNA monomer



Scheme 1: Synthesis of α-*aza*-PNA-T monomer



Scheme 2: Synthesis of β-ala-*aza*-PNA-T monomer

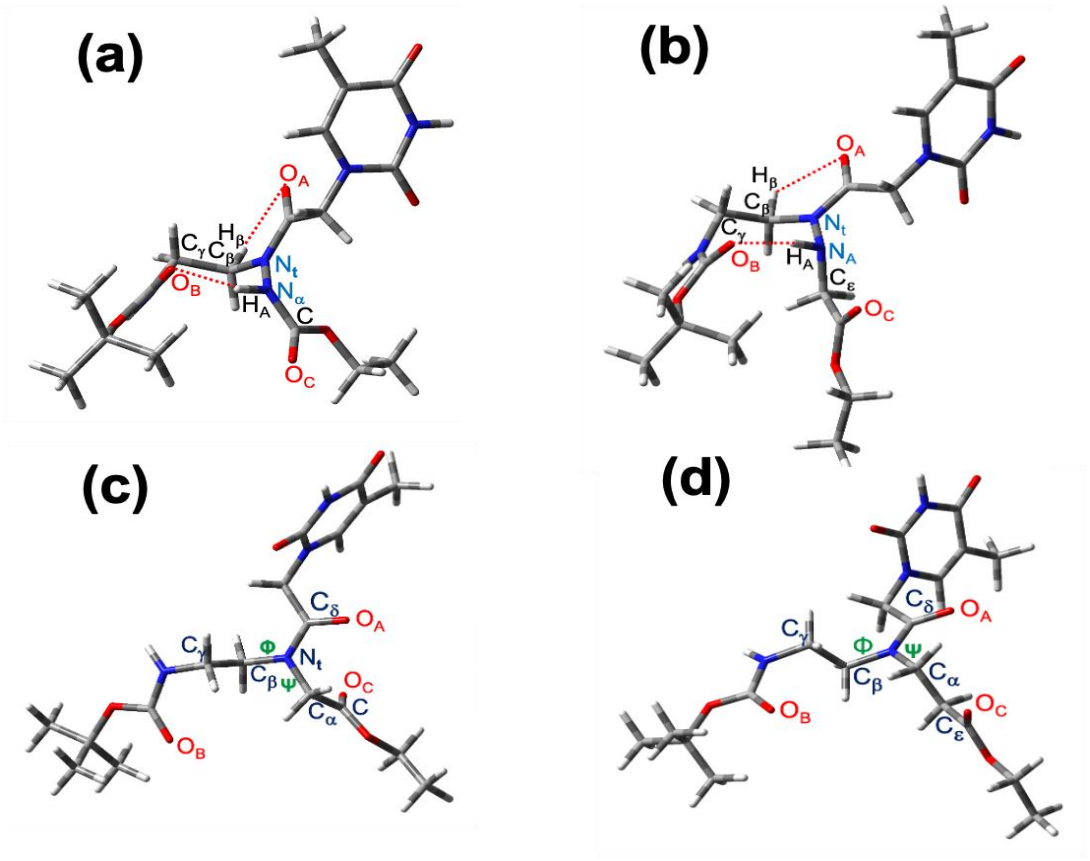


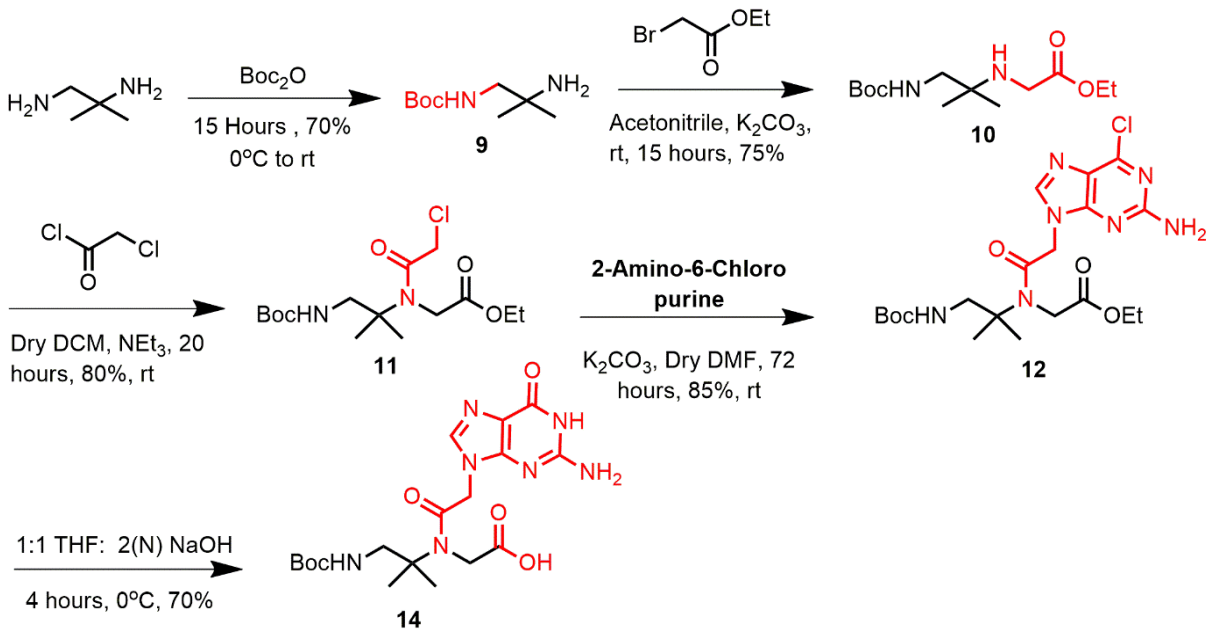
Figure 5. DFT optimized structures of (a) α-*aza*-PNA **1** (b) *aza*-β-*ala* PNA **2** (c) *aeg*-PNA **3** (d) β-*ala*-PNA **4**, with H-bond lengths and selected dihedral angles

Chapter 3- α/β-*gemdimethyl(gdm)*-Guanine PNA monomers and dimers: Rotamer dependent fluorescence

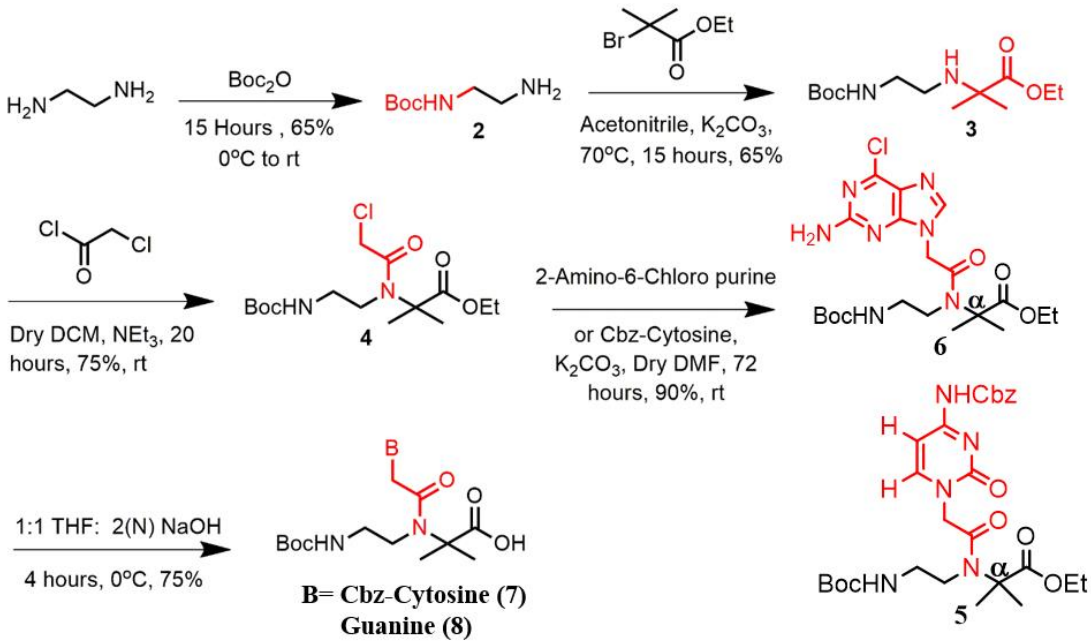
aeg-PNA GC and *aeg*-PNA GG dimer had been shown to exhibit fluorescence emission despite the fact that they lack any conventional fluorophore in the structure. π-π stacking interaction between the two base paired strands of GC/GG dimers in Z-rotamer form was hypothesized to be the origin of fluorescence in this non-conjugated system. Since incorporation of *gemdimethyl* group in the β/α-position of *aeg*-PNA-G monomer would freeze the rotamer in Z or E -form, the relative stacking interactions in these may differ and influence their fluorescence emissions. With this rationale, this chapter reports on the differential effects of E/Z-rotamers on the fluorescence emission of β/α-*gemdimethyl*-*aeg*-PNA Guanine monomers, dimers and the effect of fatty acid conjugation on the fluorescence emission and physical properties of β/α-*gemdimethyl*-(6-chloro-2-amino)-*aeg* purine monomers.

The syntheses of α,α-*gemdimethyl*-*aeg*-G/C-PNA monomers were done by same strategy as reported before, but by using 2-amino-6-chloroprurine as surrogate for Guanine to

avoid N7-alkylation, followed by its conversion to Guanine (Scheme 3). The *gem-dimethyl* N^4-Cbz-cytosine -PNA monomer was also similarly synthesized (Scheme 4).



Scheme 3. Synthetic scheme of β-*gemdimethyl-aeg*-G-PNA monomers



Scheme 4: Synthetic scheme of α,α-*gemdimethyl-aeg*-G/C-PNA monomers

In order to examine the role of backbone NH and effect of aggregation on fluorescence properties of α/β-*gemdimethyl*-G-PNA monomers, the corresponding dodecanoyl conjugated derivatives were also synthesized (Figure 6). To examine any amplification of fluorescence

effects in dimers and for comparison with previous work on dimers *aeg*-PNA GC/CG, the dimers β-*gemdimethyl*-*aeg*-GG-PNA and *aeg*-GG-PNA were also synthesized using standard solid phase peptide synthesis using Boc strategy (Scheme 5). All compounds were characterized by NMR and mass spectral data.

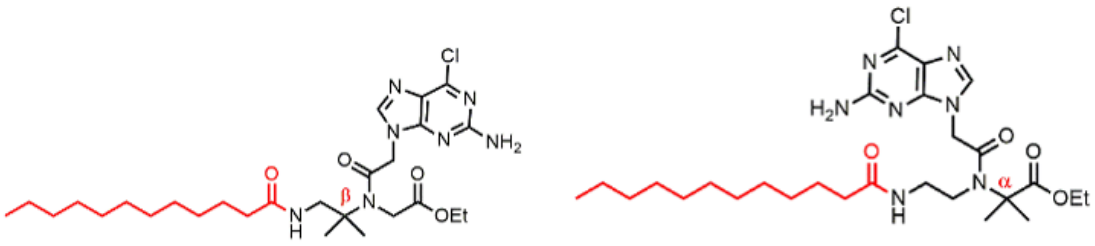
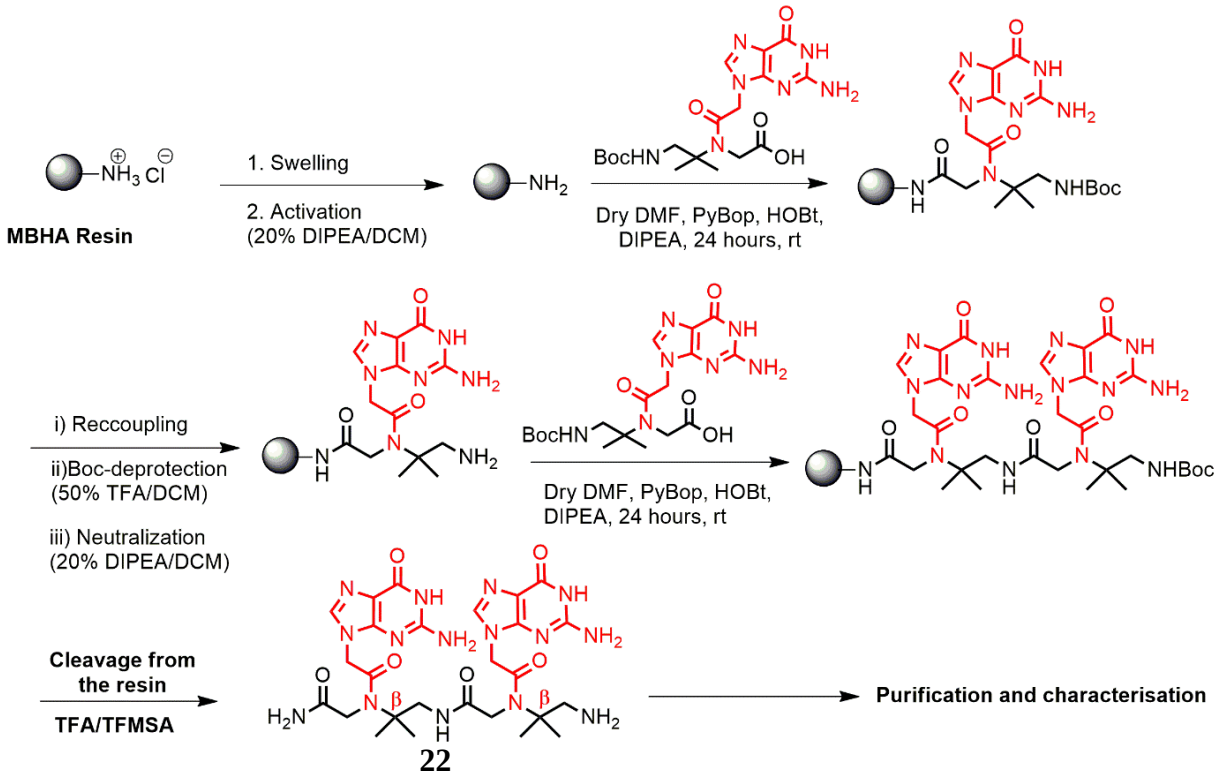


Figure 6. Structures of dodecanoyl conjugated *gemdimethyl*-G-PNA monomers



Scheme 5. Synthesis of dimers using SPPS

^{1}H NMR analysis of *gemdimethyl*-*aeg*-G-PNA monomers revealed a single Z-rotamer for α-*gemdimethyl*-*aeg*-PNA monomers and E-rotamer for α-*gemdimethyl*-*aeg*-PNA. Interestingly, the gemdimethyl-G monomers exhibited distinct UV absorption band at 309 nm, absent in *aeg*-PNA-G monomer. Further β-*gemdimethyl*-*aeg*-PNA monomer with E-rotamer displayed bright green fluorescence emission with maxima at 475 nm. α-*gemdimethyl*-*aeg*-

PNA monomer having Z-rotamer showed blue fluorescence emission with maxima at 402 nm and 428 nm (Figure 7). The monomers also exhibited fluorescence in solid powder forms as well.

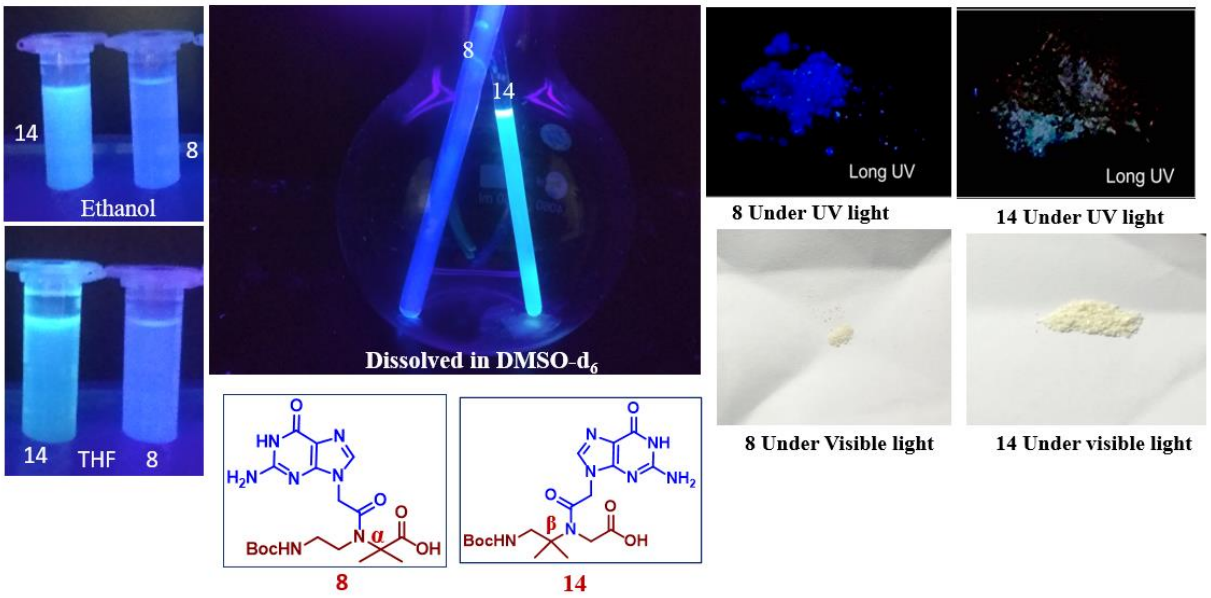


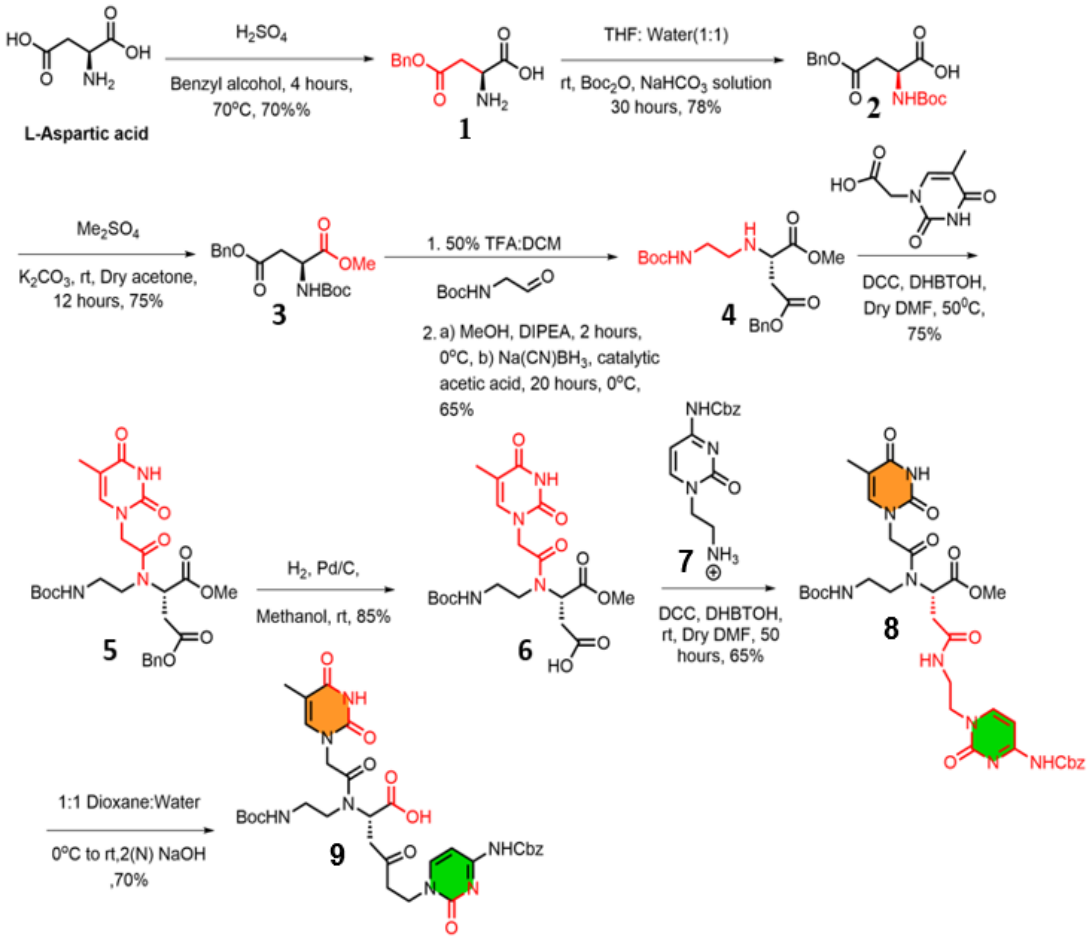
Figure 6. Solution and solid-state fluorescence of β-*gdm-aeg*-PNA-G monomer (**14**) and α-*gdm-aeg*-PNA-G monomer (**8**)

Both monomers exhibited the phenomenon of excitation dependent emission spectra in solution with slightly different spectral patterns suggesting that E and Z rotamers have different electronic and optical properties. Recent reports on origin of fluorescence in molecules lacking conventional fluorophores suggest that through space conjugation (TSC) involving n-π* interactions enhanced in aggregated systems are responsible for fluorescence in such syatems.[1] The fluorescence properties change in dodecanoyl conjugated α/β-*gemdimethyl*-G-PNA monomers solution, which caused aggregation induced quenching. The amphilic nature of dodecanoyl conjugated α/β-*gemdimethyl*-G-PNA monomers enable them form Langmuir monolayers at the air water interface which are characterized with respect to various aspects. The β-(*gdm-aeg*-G)$_2$-PNA and (*aeg*-G)$_2$-PNA dimers lacked UV-absorption band at 309 nm accompanied by weak emission in organic solvents buffer and solid powder form.

Chapter 4: Optical properties of bimodal PNA monomers and oligomers

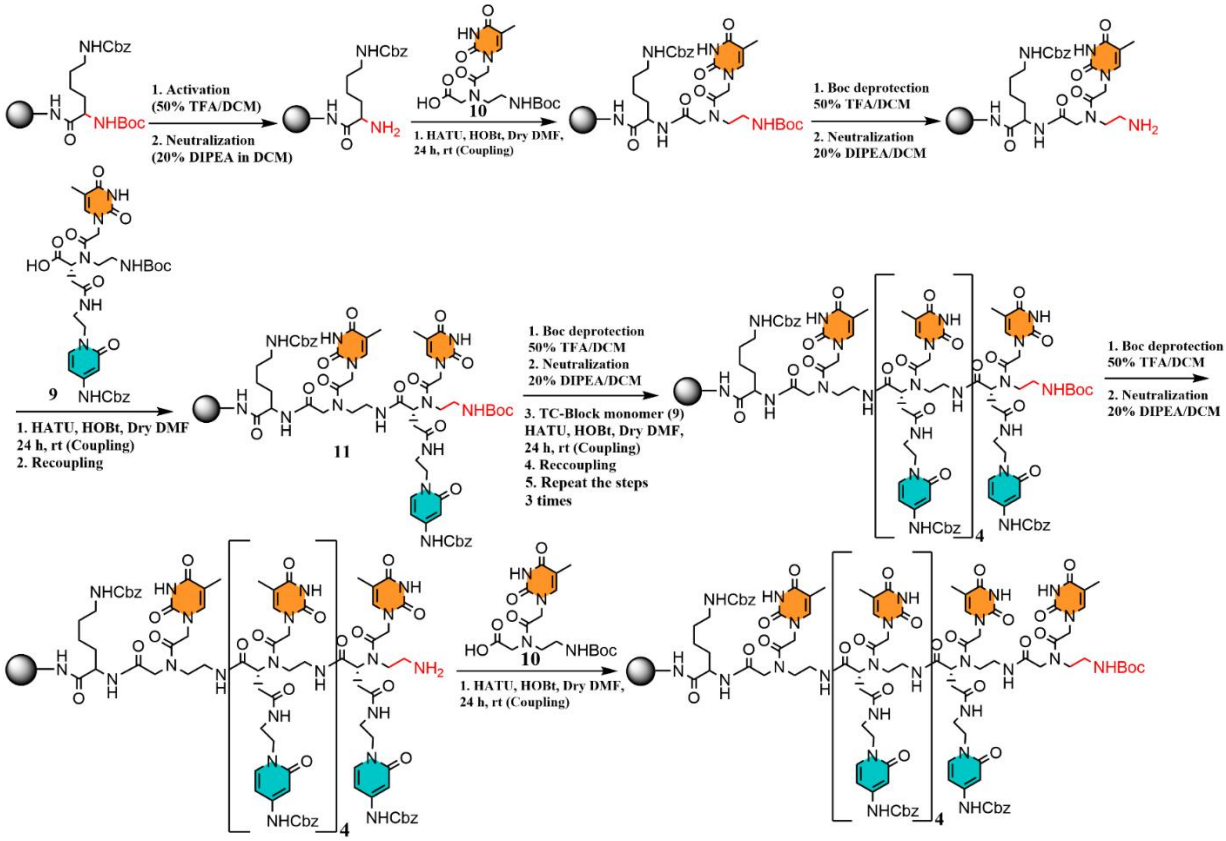
The interesting and unexpected fluorescence properties of *gemdimethyl* substituted PNA-G monomers prompted extension of studies to the monosubstituted novel bimodal PNA monomers and oligomers. The solid phase synthesis of bimodal PNAs with mixed base

sequence on both sides involve complex synthetic schemes[2] involving orthogonal protection to enable selective deprotection / coupling steps to extend the main chain and simultaneously enable side chain nucleobase derivatization.[2] An alternative method is to synthesize *bm*-PNA units in which both nucleobases are already present in the monomer stage and then use it for sequential coupling. This needs synthesis of several possible nucleobase combinations in the block monomer. To explore this strategy, synthesis of such bi-monomer blocks for use in bimodal PNA synthesis was attempted. Scheme 6 shows synthesis of Cα(*S*) -(acetamido)-*bm*-TC-PNA monomer, which was used for solid phase synthesis of the bm-(PNA-TC)$_n$ oligomer (Scheme 7).



Scheme 6. Synthesis of bm-TC-PNA block monomer

Optical properties of TC-block PNA monomer and oligomer: ^{1}H NMR analysis of Cα(S)-(acetamido)-*bm*-PNA monomer (**9**) revealed it to be a mixture of rotamers in solution with a Z:E ratio (55:45) compared to *aeg*-PNA (70:30). In the NOSEY, cross peak was noticed for α-CH with the β/γ-methylene suggesting a somewhat folded structure in solution. UV absorption of the Cα(S)-(acetamido)-*bm*-PNA monomer (**9**) revealed bands at 300 nm, 280 nm and 246



Scheme 7. Solid phase synthesis of *bm*-(PNA-TC)$_n$ oligomer

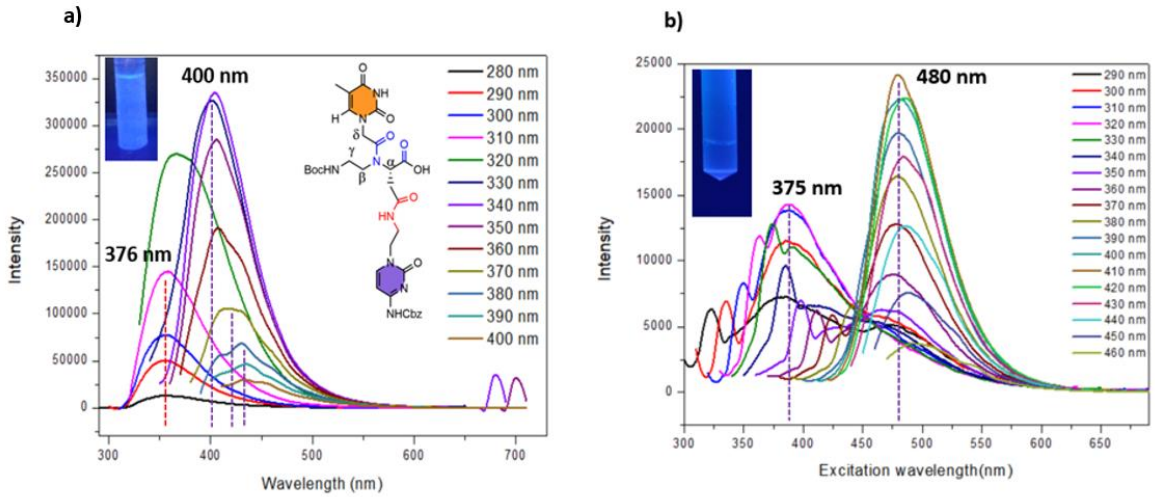


Figure 7. a) Fluorescence emission of Cα(S)-(acetamido)-*bm*-PNA monomer (**9**) in ethanol (concn, 250 μM). b) Fluorescence emission of the Cα(S)-(acetamido)-*bm*-PNA derived oligomer (**12**) in phosphate buffer (10 mM, pH 12.75)

nm. Notably, the TC-block monomer (**9**) shows bright blue fluorescence emission in solution and solid state with maxima at 400 nm at excitation wavelength of 300 nm, accompanied by excitation dependent emission spectra (REES effect) (Figure 7a). Attachment of two different

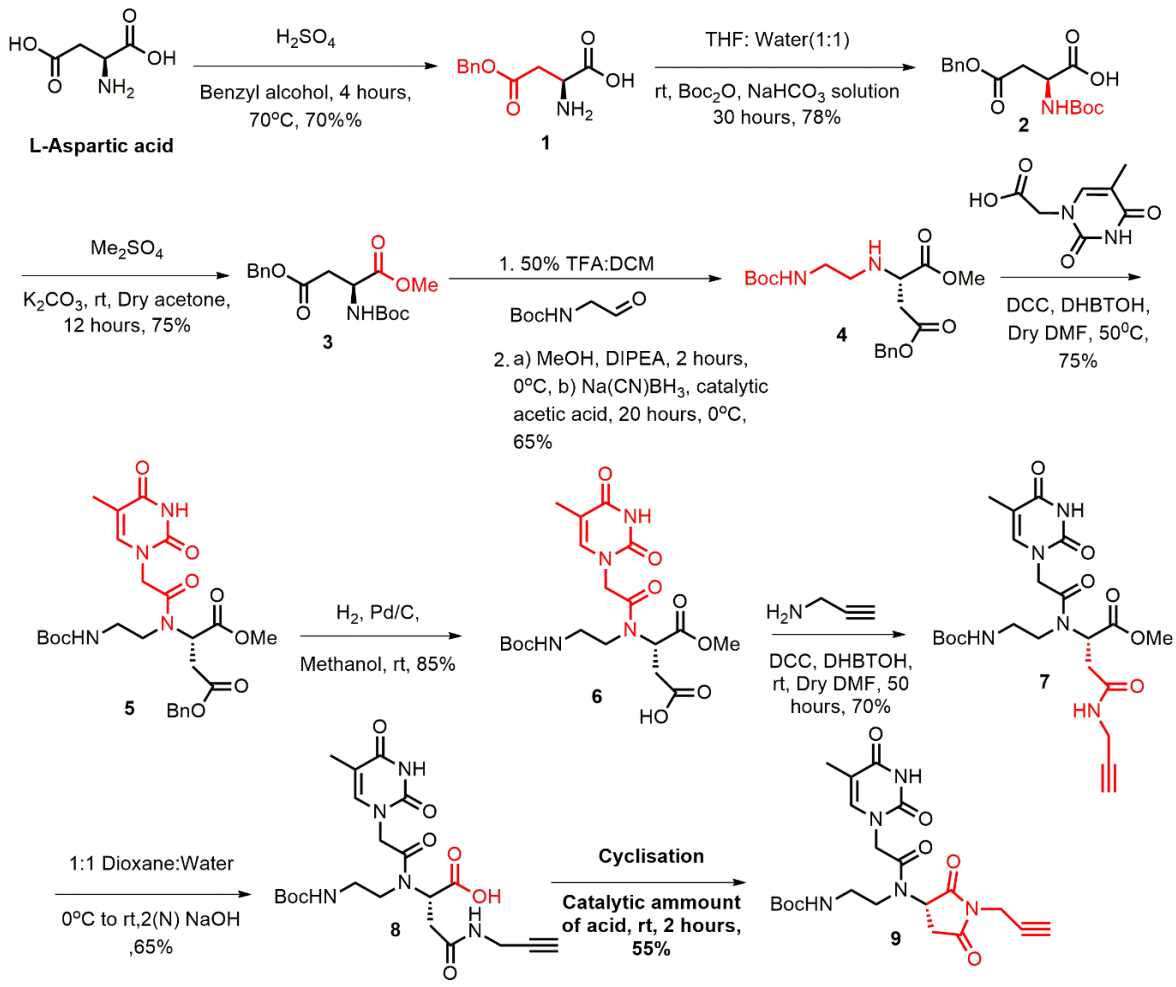
nucleobases in the same PNA backbone seem to effect efficient TSC interaction reducing the S_1 to S_0 band gap, with multiple emission bands at lower wavelength 375 nm and a single emission band at 425 nm. The presence of mixture two rotamers in Cα-monosubstituted derivative **9** may induce differences in excited state solvation. The UV-absorption of oligomer **12** shows a single absorption band at 270 nm. Study of fluorescence spectra of oligomer **12** in water and alkaline phosphate buffer (pH 12.75) revealed emission and excitation dependent fluorescence emission in the range 290 nm-460 nm (Figure 7b). The fluorescence emission in phosphate buffer was seen only at alkaline pH (12.75), but not at lower pH (7.0-10.0) suggested suggesting deprotonated T and C at pH 12.75 may be crucial for the fluorescence phenomenon of the TC block monomer **9**.

Chapter 5: Synthesis, characterization and rotamer properties of Succinimide analogues of Peptide Nucleic Acid

Most of the cyclic modification reported for PNA involves either cyclisation of the β/γ-methylene and δ-methylene or cyclisation of α-methylene with the δ-methylene group. But cyclisation involving the glycyl methylene (α-methylene) and C-terminal amide group has not been reported yet. Cyclisation of suitable side chain attached at C_α-methylene of PNA backbone with amide group at C-terminus can lead to formation of lactam or succinimide ring at the C-terminus of the PNA monomer. Many natural and synthetic succinimide derivatives including *tecovirimat, lurasidone, perospirone* etc. are FDA approved antitumor, antimicrobial, antiviral, anti-inflammatory agents. Succinimide based PNA monomers would add new repertoire for PNA based small molecule drug design. This chapter reports synthesis (Scheme 8), rotamer study and the crystal structure of C_α-succinimide modified PNA monomer.

^{1}H NMR analysis of cyclic Cα-succinimide PNA monomer **9** reveals a mixture of rotamers similar as in acyclic *aeg*-PNA, but with rotamer ratio of 1:1 compared to 3:1 of acyclic *aeg*-PNA. This reflects the effect of cyclisation around the C-terminal on the rotamer energies. Further analysis by ^{1}H-^{1}H NOSEY revealed the presence of Z and E-rotamers. assigned from expected NOSEY cross peak from α-CH with δ-methylenes and the crosspeaks to backbone β/γ-methylenes suggested a folded E and Z rotamer. The single crystal X-ray structure of Cα-succinimide PNA monomer **9** exhibits Z-rotamer in the crystal and shows folded turn structure for the monomer supporting the NOSEY observation. Apart from that the crystal exhibit T: T duplex structure and intermolecular hydrogen bonding between the N-terminal BocNH and

thymine C=O group, which is rare in the crystal structure of the PNA monomer and oligomer (Figure 8).



Scheme 8. Synthesis of Cα-succinimide PNA monomer **9**

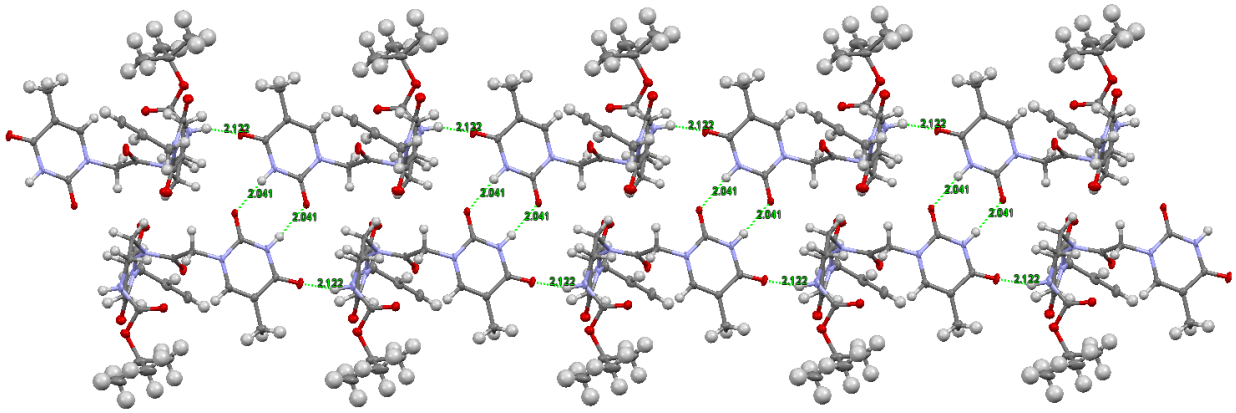


Figure 8. Crystal structure of Cα-succinimide PNA monomer **9**

Summary of the thesis

The objective of this thesis is to establish a relation between the structure of PNA rotamers with specific properties such as fluorescence and induced secondary structures.

- Novel *aza*-modified PNA monomer α-*aza*-PNA-T and β-ala-*aza*-PNA-T exhibited extended γ-turn assisted E-rotamer stabilization which is unique in the field of peptide nucleic Acid (PNA). α-*aza*-modification of *aeg*-PNA and β-ala-PNA to give *aza*-PNAs resulted in interesting folded structures on backbone mediated by H-bonding of α-NH group in *aza*-PNA, supported by NMR studies and crystal structure. Theoretical DFT calculations supports the experimental observation of extended γ-turn formation in *aza*-PNAs.

- β/α-*gemdimethyl*-*aeg*-G-PNA monomers despite lacking fluorochromophores show rotamer dependent fluorescence emission in solution and solid state. The Z-rotamer (α-*gemdimethyl*-G-PNA) gives blue fluorescence while the E-rotamer (β-*gemdimethyl*-G-PNA) shows green fluorescence. Fatty acid conjugated *gemdimethyl* monomers exhibit fluorescence emission and monolayer formation property. Fluorescence emission of all the *gemdimethyl* Guanine monomers display excitation dependent emission spectra or REES effect.

- Synthesis of modified di-block PNA monomers for sequential coupling / synthesis of *bimodal* PNA oligomers on solid phase by orthogonal strategy is demonstrated. TC-block PNA monomer and the derived homo oligomer exhibit fluorescence and excitation dependent emission property (REES) in solution.

- Cyclisation of C_α-methylene and C-terminal carbonyl group results C_α-succinimide modified PNA having reduced rotamer ratio (1:1) and folded Z and E rotamer in solution which explains the higher binding affinity of cyclic PNA derivatives. Single crystal structure of the monomer shows a folded Z-rotamer stabilization and interesting hydrogen bonding properties.

In conclusion, the work presented in this thesis makes original contribution on hitherto unaddressed role of PNA Z/E-rotamers on physical and optical properties of monomers and derived oligomers, including the interesting *bimodal*-PNA oligomers. The results will have implications for developing practical applications of the new PNA analogues – *aza*-PNA and *bimodal, gemdimethyl*-G-PNA monomers and oligomers.

References

1. a) Qian, H.; Cousins, M.E.; Horak, E.H.; Wakefield, A.; Liptak, M.D.; and Aprahamian, I. *Nat. Chem*., **2017**, *9*, 83–87. b) Guo, J.; Fan, J.; Lin, L.; Zeng, J.; Liu, H.; Wang, C.K.; Zhao, Z.; and Tang, B.Z. *Adv. Sci.*, **2019**, *6*, 1801629. c) He, Z.; Zhao, W.; Lam, J.W.Y.; Peng, Q.; Ma, H.; Liang, G.; Shuai, Z.; and Tang, B.Z. *Nat Commun*., **2017,** *8*, 416. d) halpern, A.M.; Roebber, J.; and Weiss, K. *J. Chem. Phys*, **1968**, *49*, 1348–1357. e) Ohrn, A. and Karlstro, G. *J. Phys. Chem. A*., **2006**, *110*, 1934–1942.

2. a) Gupta, M. K..; Madhanagopal, B. R.; Datta, D. and Ganesh, K. N. *Org. Lett*., **2020**, *22*, 5255−5260. b) Gupta, M. K.; Madhanagopal, B. R.; and Ganesh, K. N. *J. Org. Chem*., **2021**, *86*, 414−428. c) Bhingardeve, P.; Madhanagopal, B. R.; and Ganesh, K. N. *J. Org. Chem.*, **2020**, *85*, 13680−13693. d) Bhingardeve, P.; Jain, P.; and. Ganesh, K. N. *ACS Omega*., **2021**, *6*, 19757−19770. e) Todkari, I.; Gupta, M. K.; and Ganesh, K. N. *Chem. Commun*., **2022**, *58*, 4083–4086.

Chapter 1

Introduction

Chapter-1

Introduction

1.1. Introduction to Nucleic acids

Nucleic acids are the most fundamental biopolymer that define the functions of a cell and thereby the organism. The double helix structure of deoxyribonucleic acid (DNA) was discovered by Watson and Crick in 1953[1]. The major roles of DNA[1] as a genetic material include storage of genetic information within the cell and its transmission to progeny through replication producing identical copies of original DNA. The central dogma of the molecular biology involves synthesis of RNA from DNA (transcription) and synthesis of proteins from RNA (translation). DNA and RNA are long chain biopolymers containing a linear arrangement

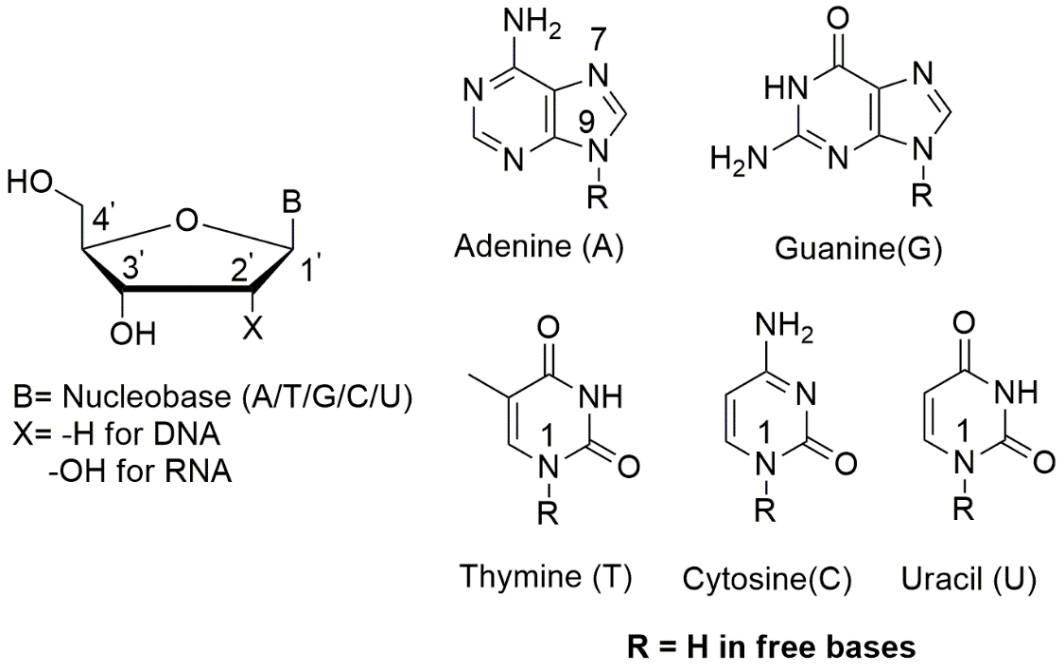


Figure 1.1. Structure of Nucleoside (Sugar ring and Bases)

of basic unit of nucleotides which are the phosphate esters of nucleoside. A nucleoside is made up of a five membered pentose sugar ring and a heterocyclic nitrogenous base that are pyrimidines T (thymine), U (uracil) and C (cytosine) and purines A (adenine) and G (Guanine). DNA and RNA differ in the nature of the five membered sugar ring: 2'-deoxy ribose pentose in DNA and ribose in RNA. Further, the heterocyclic pyrimidine base U is present in RNA instead of T in DNA. The pyrimidines (T, C and U) are the six membered aromatic heterocyclic bases, while the purines (A and G) are bicyclic aromatic heterocyclic with a six membered ring

fused to a five membered ring. The heterocyclic bases are connected to the five membered sugar ring in DNA/RNA via a glycosidic bond at C1' of sugar. The purines are linked through N9 nitrogen while pyrimidines are linked via N1 nitrogen (Figure 1.1).

1.1.1 Base pairing and duplex structure of DNA

Watson and Crick proposed a novel duplex structure of DNA in 1953.[1] The two helical chains of DNA are coiled around each other on the same axis and each polymer chain consists of nucleotides connected via phosphate link between 3' and 5' hydroxyls of sugar moieties. The heterocyclic bases are attached to sugars at C1' via β-glycosidic linkages. The two chains run in opposite directions, one in 3'-5' direction and another in 5'-3' direction (antiparallel) leading to a right-handed double helix in DNA. The distance between the two bases (and the in the duplex structure of DNA is 3.4 Å, and the angle between the two adjacent base planes in the same chain is 36°. Hence after 10 residues in each chain, the helical unit repeats leading to 34 Å length per each turn of the helix. The water content of the duplex structure of DNA determines its structural profile: at high water content it assumes so called B-DNA and low water content it adapts the A-form and several structural parameters such as base tilts in relation to axis, pitch of the helix etc also differ between the two forms. Notable feature of the duplex structure of DNA is the way two strands of DNA held together to form helix.

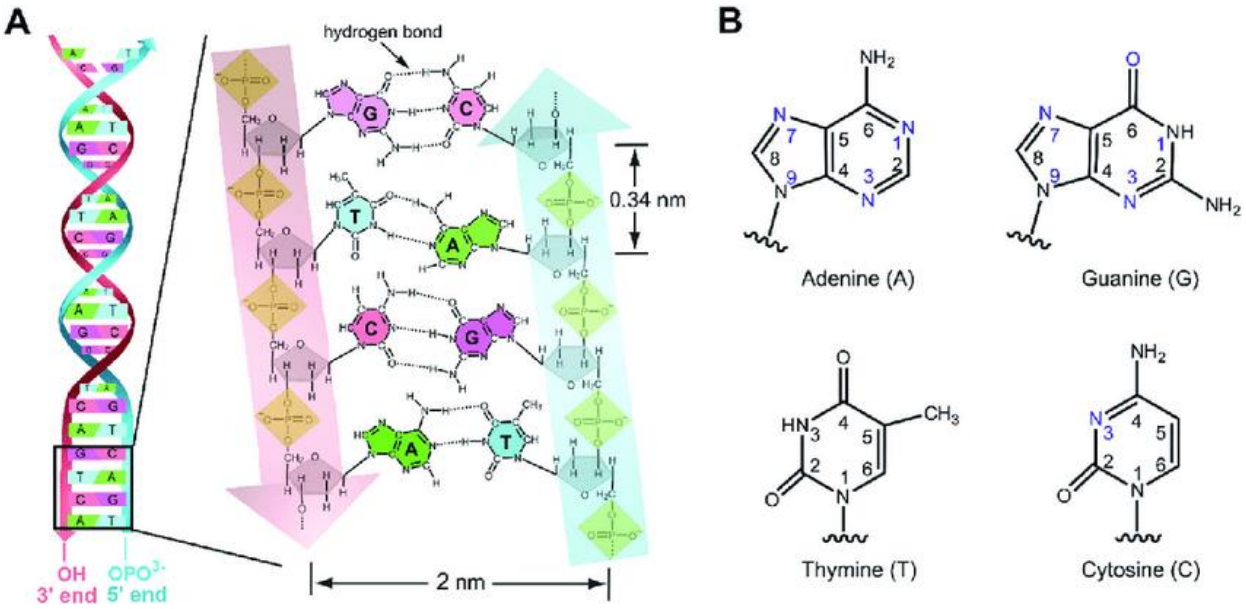


Figure 1.2. Duplex structure of DNA and the hydrogen bonding interaction

The two chains of DNA are held through complementary base pairing between the purine and pyrimidine bases. The sp^2 hybridised carbonyl group in bases is a potential hydrogen

bond acceptor while the NH is a potential hydrogen bond donor. The purine bases from one strand undergoes hydrogen bonding with pyrimidine base on the opposite strand of DNA. Adenine (A) forms complementary hydrogen bonding with thymine (T) while guanine (G) complements hydrogen bonding with cytosine (C). The complementary base pairs originate from pairing of the six-membered rings of purines and pyrimidines. The specificity of base pairing is achieved with the nucleobases in keto form rather than enol which also fixes NH as H-bond donor. The specific base pairing of adenine with thymine and guanine with cytosine is known as Watson Crick hydrogen bonding[1] (Figure 1.3).

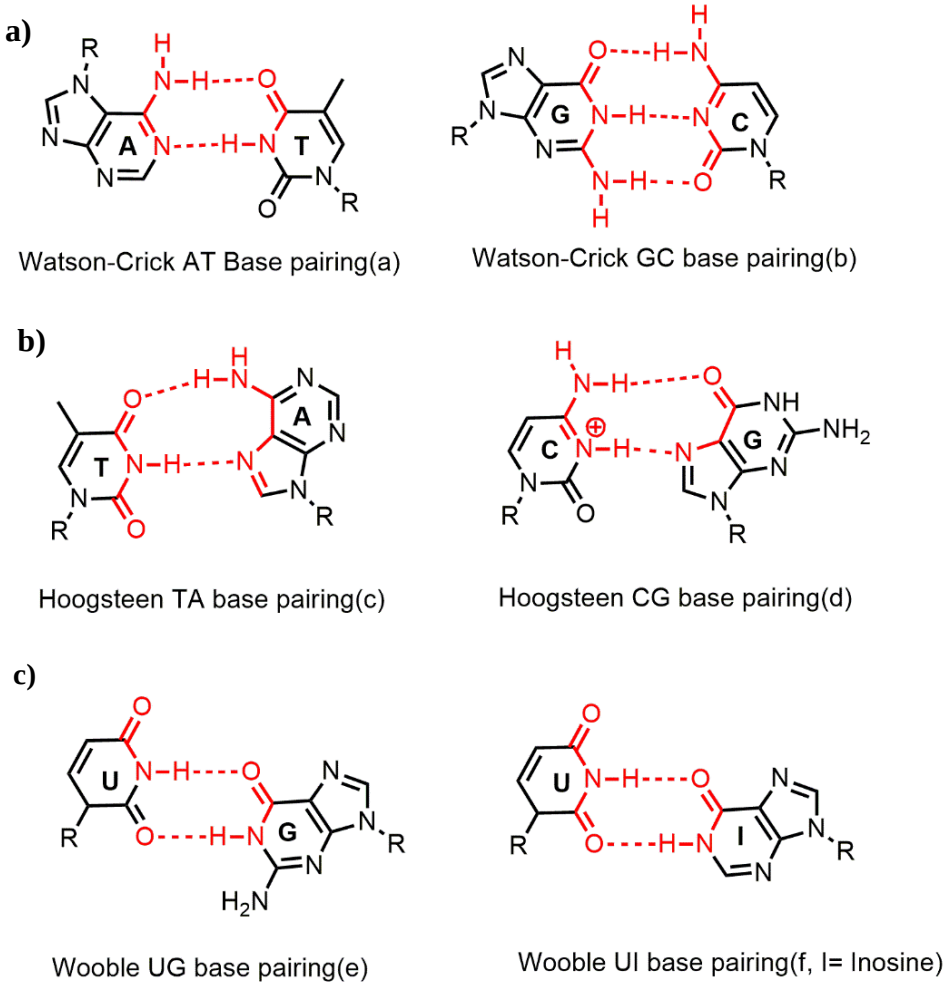


Figure 1.3. Watson-Crick base pairing[1] (a), Hoogsteen base pairing[2] (b) and Wobble base pairing[3] (c)

Apart from the Watson and Crick base pairing between complementary nucleobases, Hoogsteen[2] and Wooble[3] base pairing are the two other significant base pairing modes for nucleobases in DNA/RNA. Hoogsteen base pairing involves hydrogen bonding between T and A (and between N3 protonated C and G) and in both cases the fused 5-membered purine rings

are also involved in base pairing (Figure 1.3b). The orientation of the nucleobases along with the hydrogen bond donors and acceptors changes drastically in Hoogsteen base pairs compared to the Watson Crick base pairs (Figure 1.3b). The biological significance of Hoogsteen base pairs arises from their participation in triplex formation. Wobble base pairs[3] involve non complementary base pairs such as G:U (Figure 1.3c) and observed during the protein synthesis in the ribosome, where the messenger RNA (m-RNA) interacts with transfer RNA (t-RNA) (codon-anticodon interactions). X-Ray studies of oligodeoxynucleotide have revealed several mismatch bases pairs and anomalous hydrogen bonding patterns between the nucleobases.[4]

1.1.2 Secondary structures of DNA

Crystallographic studies of the heterogenous DNA fibres suggest two distinct conformations for DNA double helix.[5] Low humidity and high salt concentration favour the highly crystalline A-DNA and in contrary the high humidity and low salt concentration favour the B-DNA structure. A-DNA is, right-handed, anti-parallel double helix, following the Watson-Crick base pairing. The phosphate groups of the backbone stay outside of the cylindrical helix, with 11 bases in each turn of the A-DNA with 28 Å per helical turn corresponding to 2.56 Å per base pair raise. The furanose ring of A-DNA prefers C3ʾ endo puckering, and the glycoside is in the anti-conformation. In the A-DNA duplex major groove the backbones are far apart compared to the minor groove, wherein the backbones are close and the major groove is deep compared to minor groove that is quite shallow. The B-DNA structure is the most prominent form of DNA is solution. It is also a right-handed double helix, with a wide major groove and narrow minor groove. The sugar ring has C2ʾ endo pucker and all the glycosidic bonds have anti-conformations. The adjacent phosphates are slightly farer in B-DNA compared to the A-DNA (Figure 1.4).

Z-DNA is left-handed antiparallel DNA,[6] seen in case o alternating purine/pyrimidine sequences (CG/GC) that arises from the preferred unusual syn-conformation for purine bases in dimeric unit. This results an alternate anti-syn conformation in each DNA strand with the phosphates a zig-zag orientation. Sugar puckering changes from C2ʾ endo / anti pyrimidine residues to C3ʾ endo / syn in purine residues. Such conformational changes affect the base stacking and the groove topologies, with the minor groove of Z-DNA very deep and encompasses the helix axis. Another notable feature of the Z-DNA is the flat convex surface of the major groove on which guanine N7/ C8 and cytosine C5 and are exposed. Z-DNA formation is also favoured in high salt ($MgCl_2$, NaCl) concentrations. B-DNA with alternating

repeat CG sequences changes to Z-DNA at higher salt concentration. The salient structural features of the A-DNA, B-DNA and Z-DNA are summarised in the Table 1.

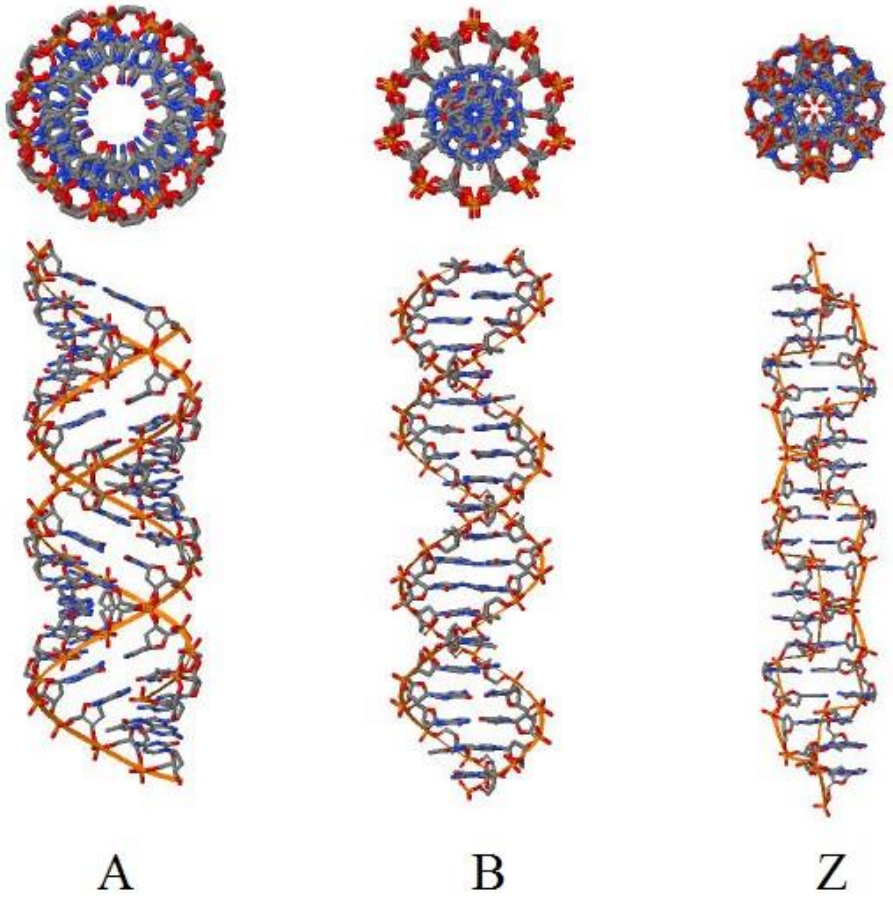


Figure 1.4. Molecular models of different forms of DNA[5,6]

Parameters	**A-form[5]**	**B-form[5]**	**Z-form[6]**
Handedness	Right-handed helix	Right-handed helix	left-handed helix
Conditions	Low humidity and high salt	Dilute aqueous solution	High salt and GC rich sequences
Residues per turn	11	10	12 base pairs
Base tilt relative To helix axis(degree)	20	6	7
Rise per base pair	2.3 Å	3.4 Å	3.8 Å
Major groove	Narrow and deep	Wide and deep	Flat
Width of major major groove	11 Å	5.7 Å	2.0 Å

Parameters	**A-form[5]**	**B-form[5]**	**Z-form[6]**
Minor groove	Shallow	Narrow and deep	Narrow and deep
Width of minor groove	2.7 Å	11.7 Å	8.8 Å
Orientation of N-glycosidic bond	*Anti*	*Anti*	*Syn* for Purines and *anti* for pyrimidine

Table 1.1. Structural features of A, B and Z-DNA

1.1.3 Secondary structures of RNA

In general RNA is a single stranded molecule but has potential to form complicated and unusual structures due to chain folding which results because of intramolecular base pairing. t-RNA forms base paired stems, bulges, and open loops and overall, the structure the shape of completely folded t-RNA is L-shaped. B-type helix formation is completely ruled out for RNA as the steric hindrance caused by 2'-OH group completely destabilizes the B-helix. But the 2'-OH can be accommodated in A-helix. Thus, RNA has more structural diversity than DNA. At low ionic strength RNA adopts a right-handed, anti-parallel double helix with 11 base pairs per turn. The sugar prefers a C3'-endo pucker conformation, and the other parameters of the helix matches with the A-DNA. If the concentration of the salts increased above 20%, A-RNA changes to A'-RNA form where there are 12 base pairs per turn of the duplex. A-RNA duplexes have very deep major groove but rather shallow minor groove. Both the A-RNA and A'-RNA shows typical properties of a Watson-Crick duplexes[1]. However, the presence of 2'-OH in the RNA nucleotide makes it more labile, due to 2'-OH acting as a potential nucleophile for intermolecular attack on nearest phosphate group. In alkaline medium RNA degrades much faster than DNA as the 2'-OH deprotonates in the alkaline medium.

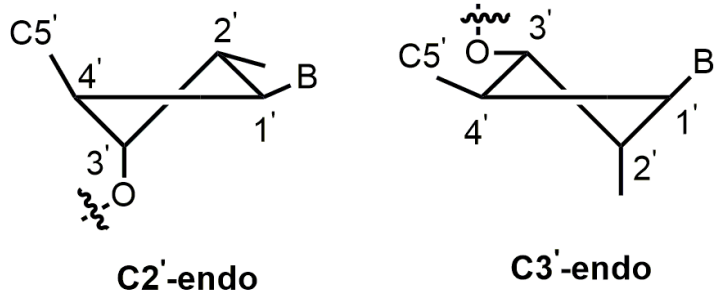


Figure 1.5. Structure of *C2'-endo* and *C3'-endo* sugar puckering

1.1.4. DNA triplexes and quadruplex structures

Pyrimidine rich DNA or RNA sequences forms multistranded triple helical structure.[6b] Poly thymine or protonated (C^+) poly cytosines forms DNA_2: DNA triplexes in which one of strand is held together by Watson-Crick[1] base pairing (Figure 1.14 and 1.5A), while the triplex strand is held together by Hoogsteen[2] base pairing between the complementary nucleobases (Figure 1.14 and 1.5A). Triplex forming oligonucleotide could serve as a potential gene-drugs and could be useful in anti-gene strategy[9]. DNA triplexes can be generated either by combining two pyrimidine rich sequence and one purine rich sequences (Y R^* Y) or by combining the two purine rich sequences and one pyrimidine rich DNA sequences[6b] (Y R^* R, Y= Purine and R= pyrimidine, R^* = Central strand). Interestingly presence of bivalent metal cations affects the stability of Y R^* R triplexes but Y R^* Y triplexes remains unaffected in presence of bivalent metal cations.[6c] Irrespective of the differences both the combination of triplexes have similar characteristics as a) One of duplex of triple helical DNA must have a homopurine sequence b) The two chemically homologous strands are antiparallel to each other.[6b]

Apart from that Guanine (G) rich DNA sequences forms folded intramolecular G-quadruplex (Figure 1.5A) structures where four guanines are held together by Hoogsteen hydrogen bonding[6d] to form G-quartret. The stability of the quadruplex structure affected by the presence of monovalent cation and the decreasing order of stability of mono cations are ($K^+ > Na^+ > NH_4^+ > Li^+$).[6d] Depending on the orientation of DNA strands G-quadruplex are divided into following categories a) Unimolecular parallel b) Tetramolecular parallel c) Unimolecular antiparallel and d) Unimolecular bulged.[6d] Cytosine rich DNA sequences having protonated (C^+) Cytosine at acidic pH (<5 to 6) folds into i-tetraplex structure where the cytosines are held together by base stacking interaction.[6e]

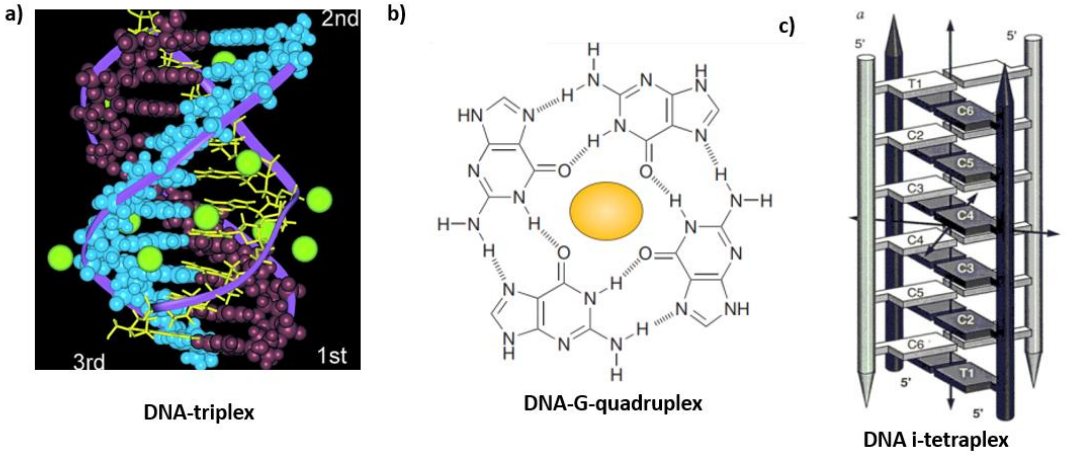


Figure 1.5A. a) DNA-Triplex b) DNA-G quadruplex c) DNA i-tetraplex

1.1.5 Application of Nucleic Acids

The unique properties of DNA such as duplex formation by sequence selective base pairing with complementary DNA, triplex and tetraplex formation leads to supramolecular structures that make DNA a perfect scaffold for designing future drugs. Chemically modified analogues of DNA and RNA can serve as better candidates for designing antisense drugs as they show nuclease resistant properties compared to the original DNA. Electrostatic interactions between two complementary DNA strands via Watson-Crick hydrogen bonding, Hoogsteen hydrogen bonding etc inspires scientist to use DNA as a building block to form novel DNA based nanostructures. Applications of nucleic acids include antisense therapeutics, diseases diagnostics and DNA nanotechnology.[7]

1.1.5a. DNA Nanotechnology[7]

Nucleic acids can be used as a novel building block for the construction of DNA-based Nanostructures[7]. The pioneering vision of Ned Seeman to use DNA as a building block of nano structures in the early1980s became the foundation of the field DNA nanotechnology[7]. Invention of DNA origami by Poul Rothemund in 2006 added another milestone in the field, which since then has expanded in an impressive pace. DNA properties such as chemical and thermal stability, easy synthesis of the short oligonucleotides, sequence dependent predictable nature of the DNA double helix and the dynamic nature of the DNA hybridisation had inspired chemists and biologists to choose DNA as building block for novel nanostructure formation. It is possible to obtain and design 2D and 3D DNA nanostructures by DNA nanotechnology. DNA nanostructures are made up of short double-stranded DNA domains joined in a programmable fashion to create different shapes and sizes by using single stranded DNA overhangs known as 'sticky ends.[7] Modification of nucleobases may bring more structural diversity in the field of DNA nanotechnology as unnatural nucleobases like iso-guanine, cyanuric acid prefers interesting pentameric and hexameric rosette like assemblies. Metal mediated base pairing of DNA may add another layer of structural diversity of DNA nanostructure as introduction of metal adds interesting properties like redox activities, photochemical, catalytic, and magnetic properties. First metal-DNA branched structure was reported in 2001, where two Ru-bipyridine complexes containing complementary DNA strands was brought together and they formed a cyclic metal DNA-nanostructures. Application of DNA nanotechnology includes designing DNA based electronic devices by assembling inorganic nanostructures with DNA, design of novel bio-physical and bio-medical techniques etc.

Covalent attachment of specific protein binding ligands (ex biotin etc) can bring corresponding proteins, whose placement on nanostructure surface can be controlled by ligand disposition.[7]

1.1.5b. Disease diagnostics

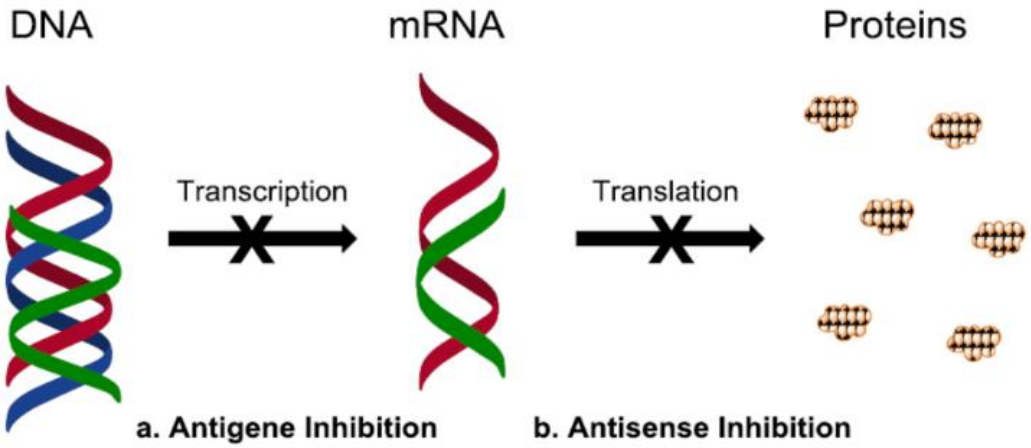
Inherent properties of nucleic acids like recognition and binding with specific complementary DNA sequences makes it a natural biomarker in molecular biology. The main principle of nucleic acid-based diagnostics is the direct detection of DNA genes or associated RNA sequence of the host or by first amplifying the pathogen DNA or RNA. Nucleic acid-based diagnostics detect the DNA or RNA from the disease infected organism by detecting the specific gene or level of gene expression associated with the discase. Cancer, genetic discase (e.g cystic fibrosis) etc are conditions where such gene level diagnostic techniques are very useful. For viruses, nucleic acid based diagnostic techniques involve use of are RTPCR, isothermal amplification, hybridization and sequencing.[8]. Infectious disease like anthrax, *Clostridium difficile* (a common hospital acquired bacterial infection), chlamydia, gonorrhea, tuberculosis, and HIV is detected using nucleic acid based discase diagnostics.[8] An efficient sensitive way is to conjugate DNA with a fluorescent probe, ligands, or enzymes for detection by techniques such as fluorescence, ELISA and chemiluminescence.[8]

1.1.5c. Oligonucleotide as therapeutic agents

Conventional approach of designing small molecule drugs requires complete understanding of the secondary and tertiary structure of the target proteins, binding site, dissociation constant etc. Due to the unpredictability of the protein folding process structure-based drug design for targets whose crystal structure is not known has some serious limitation. In contrary, where the structure of nucleotide sequence DNA and RNA of the associated / disease related proteins are known, nucleic acid-based drug approach is very useful. It is possible to design DNA / RNA sequence complementary to gene (DNA) or m-RNA sequence corresponding to the disease related protein and achieve sequence specific and selective silencing of target gene through antigene, antisense and siRNA therapeutics. In other words, a rationally designed short oligonucleotide (DNA / RNA) can act as drug.

The antigene or triplex approach[9] involves selective inhibition of the transcription process of the corresponding gene unwanted protein to stop the discase to spread further. The mechanism involves use of a third strand of DNA to bind in the major groove of DNA duplex via Hoogsteen pairing and form triplex,[9] and thereby inhibit DNA replication. The *antisense strategy*,[10] involves prevention of protein synthesis by selective inhibition of the translation

process. The rationally designed DNA / RNA sequence for antisense strategy must is a complementary oligonucleotide of target sense strand as the mechanism of action involves Watson crick base pairing between the target mRNA and the antisense oligonucleotide to form hybrids,[10] that is not recognized by protein synthesis machinery and may even be susceptible for degradation of m-RNA by RNAse H,[10] to stop protein production.



Antisense oligonucleotides. Grineva *et. al*[11] first proposed the idea of the use of antisense oligonucleotide in 1960. Later Ts'o et. *al.*[12] showed that chemical modification of the parent DNA/RNA would stabilise them from degradation by *endo*/*exo* nucleases. Zamecnik and Stephenson[13] first illustrated the idea of antisense therapeutics in 1978 and demonstrated that replication of *Rous sarcoma* virus in a cellular system can be inhibited by antisense oligonucleotide. The limitation to explore the idea at that time was unavailability of synthetic methodology for scaling up the oligonucleotide in large quantity. An effective and successful therapeutic oligonucleotide candidate should meet the following criteria[15]

(i) Ability to selectively interact with the target sequence in the cell

(ii) Must be stable towards degrading enzymes like nucleases and proteases

(iii) Should be able to penetrate the cell membrane

(iv) Should be retained by target cells

(v) No non-specific interaction with the other macromolecules presents in the cells

(vi) Should not be toxic to cell

(vii) Ease of scale up

A systematic surmounting of all these barriers,[14] has lead to success of antisense technology, which is based on selective inhibition of protein synthesis during translation process. Currently some antisense oligonucleotides have been approved and many other oligonucleotides are under investigation for antisense therapy in viral infections, cancers, inflammatory disease and genetic disorders.[14,15] Crucial to success is the rational chemical modification of oligonucleotides to meet many of the requirements and intense biological research on their effectiveness and toxicity.

1.1.6 Chemical modifications of DNA

To fulfil the criteria of an ideal antisense oligonucleotide several chemical modifications of the original DNA molecule have been performed.[16] These can be classified into three groups (Figure 1.7): (a) the first generation oligonucleotide modification involved modification of the backbone phosphate group to obtain nuclease resistance properties (b) the second generation oligonucleotide modification involved derivatisation of the 2ˆ-OH group of the ribonucleotides, such as alkylation, replacement by fluorine, bridging with C4' etc to achieve both nuclease resistance and pre-organise the sugar conformation to RNA like C3'-

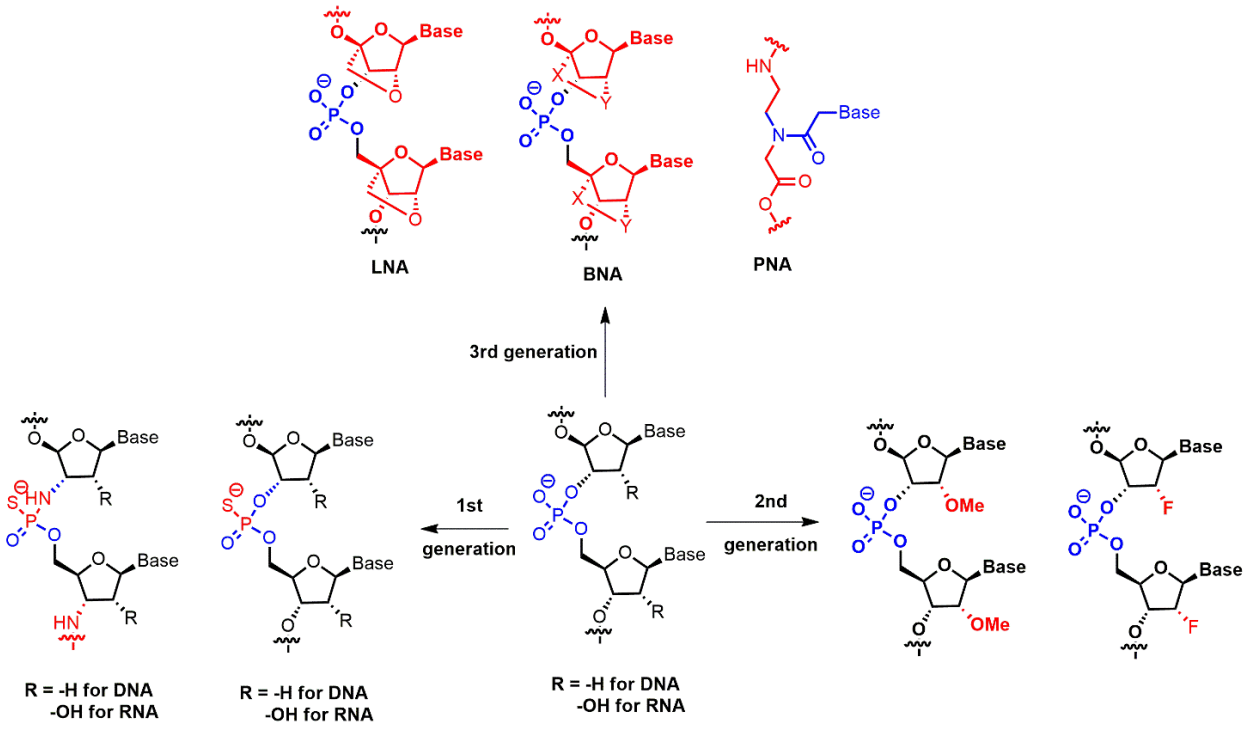


Figure 1.7. First, second and third generation modified oligonucleotide

endo form (c) the third-generation oligonucleotide modifications involved alteration of the sugar ring and inter nucleotide phosphate group. Examples includes morpholino

phosphoroamidate (MF),[24] peptide nucleic acids (PNA),[27] threose nucleic acids (TNA)[21b, 21c] etc. Apart from the above mentioned three classes of modifications, the nucleobase modifications have also been explored in the literature.[26c-26i]

1.1.6a First generation modification[17]

The first-generation chemical modification of oligonucleotides[17] (Figure 1.8) involve the backbone phosphate group is replaced by several linkers like 5'-N-carbamate, methylene-methylimine (MMI), amide, triazole. phosphorothioate (PS), phoshorodithioate, thioether, methyl phosphonate, boranophosphate, N3''-phosphoramidate (NP) S-methylthiourea and guanidinium linkage. Depending on the the linkers, these can be grouped as anionic, neutral, and cationic in nature. The most important modifications among the first-generation modification is phosphorothioates (PS-ODNs). These were successfully used as a gene silencing agent as they exhibit higher stability towards nucleases and proteases, improved cell permeability, but suffer from non-specific interactions with some proteins and toxicity.[18] It also makes P of phosphate group chiral leading to a mixture of diastereomers, but FDA approved the first antisense PS drug Vitravene in 1998 which provided the proof of concept for antisense principle. Other first generation modified oligonucleotide suffer from drawbacks like poor water solubility and less successful to improve the properties of original DNA Oligonucleotide.

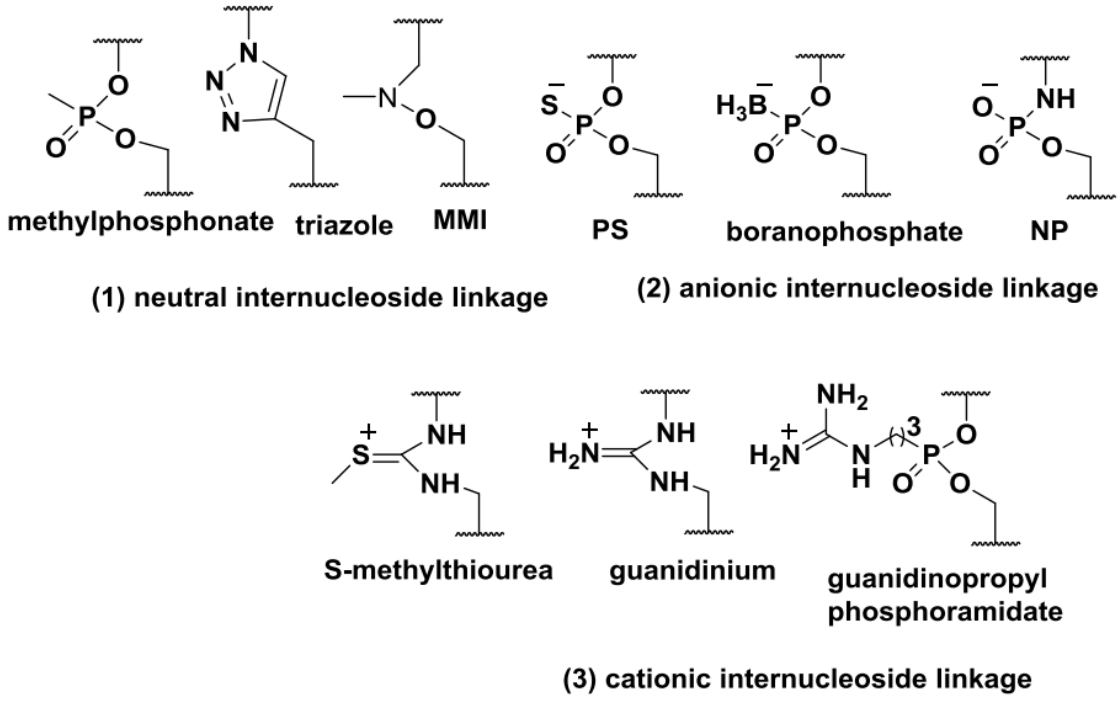


Figure 1.8. Chemically modified linkers of first generation ODN modification[17,18]

1.1.6b Second generation modifications

Second generation modifications of oligonucleotides involve modification at the 2'-OH group of the oligonucleotide (Figure 1.9). These include alkylation of 2'-OH group of ribose, replacing the 2'-OH by electronegative elements like fluorine and inverting the stereochemistry at C2' to yield arabino sugars. These modifications induce *C3'-endo* puckering for the sugar as in RNA and hence exhibit higher binding affinity towards cRNA than cDNA.[19] C2'-substitution increases the steric bulk around the phosphate group which consequently makes difficult for nucleases to cleave the phosphodiester bond. Introduction of electronegative fluorine atom at the 2' position enhances the binding affinity towards target cRNA. Although 2'-OMe shows lesser binding affinity towards cRNA compared to the 2'-F modification[19] it exhibited better resistance towards proteases and nucleases. 2'-O-methoxyethyl (MOE)[20] is the most advanced modification in the second-generation oligonucleotides, with many antisense drugs currently under clinical trial. Although second generation of oligonucleotides show properties improved over the first generation, the unsatisfactory RNase H activity was overcome by "Gapmer" strategy.[21] FDA has approved RNA aptamer drug Macugen used for the treatment of macular degeneration is based on the 2'-F and 2'-OMe substituted sugar moieties of oligonucleotides.

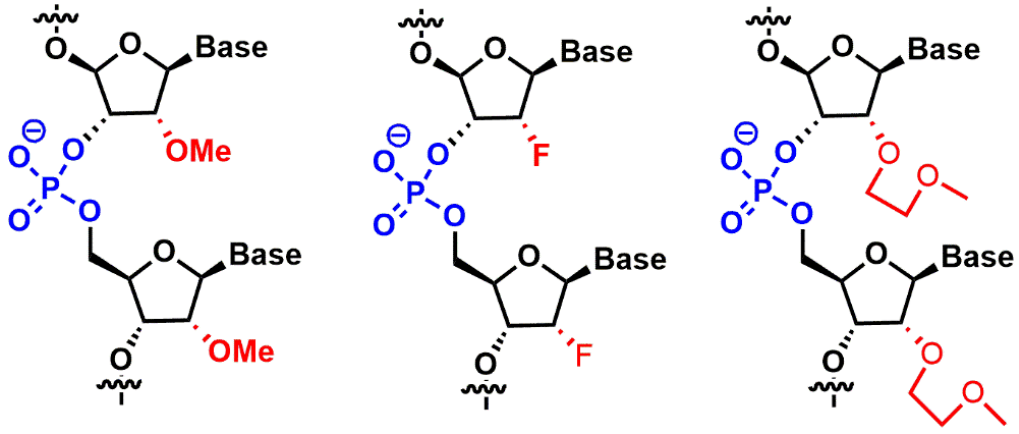


Figure 1.9. Second generation modification of the oligonucleotide[19,20,21]

1.1.6c Third generation modifications

Third generation modification of the oligonucleotides involving replacement of sugar ring and phosphodiester backbone and seems to be most effective. Some the promising and potential candidates from third generation modification are locked nucleic acid (LNA), Cyclohexene nucleic acids (CeNA), morpholino oligonucleotides (MF) and peptide Nucleic acids (PNA, Figure 1.10). Imanishi and Wengel[22] introduced Locked Nucleic Acid (LNA) in

1998, which is a novel analogue with the 2'-O and 4'-C positions of the five membered sugar ring joined via O-methylene, S-methylene, or N-methylene group. Due to the bicyclic nature, the furanose conformation of the sugar ring is locked in C3'-*endo* conformation. LNAs shows high binding affinity towards complementary DNA/RNA, extraordinary antisense activity, nuclease resistance, good aqueous solubility, and almost negligible toxicity,[23a-23b] but LNAs are unable to activate the RNase H. To overcome this drawback, LNA-DNA-LNA gapmer is designed where the central DNA consisting of 7-10 PS oligonucleotides are connected with the 3 or 4 LNA units on each side.[23c-23c] Morpholino oligonucleotides (MF) are the neutral analogues of DNA where the sugar ring is replaced by a six membered morpholino ring and the phosphodiester bond is replaced by an isosteric phosphorodiamidate linkage.[24] MF ODNs are stable against the enzyme nucleases with the target affinity same as the unmodified ODNs. The neutral character of MF ODNs affects their cell uptake properties and do not activate the RNase H. They can be used as steric blocker to inhibit the gene expression.[25] In the cyclohexene nucleic acid (CeNA), the five membered sugar ring is replaced by as cyclohexene ring keeping the phosphodiester linkage intact.[26] The presence of cyclohexene ring in CeNA makes it rigid, resistant to nucleases and activate the RNase H. They form stable duplexes with complementary DNA.[26]

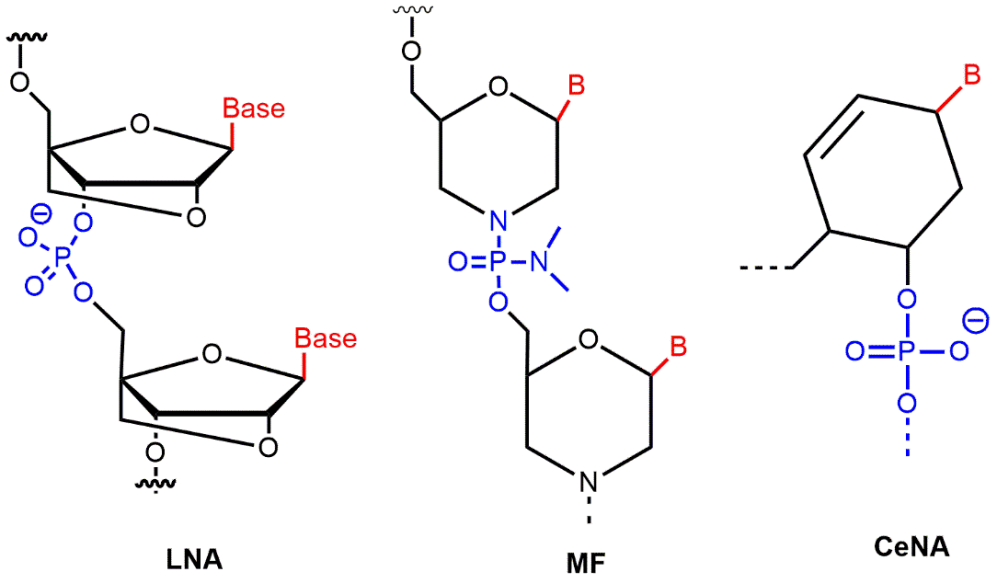


Figure 1.10. Third generation modification of oligonucleotides[22-26]

1.2 Peptide Nucleic Acid (PNA)

Peptide Nucleic Acids (PNAs) are synthetic mimics of DNA where the 2'-deoxyribose sugar ring of the DNA is replaced by a 2-aminoethyl-glycine backbone (Figure 1.11).[27] The

phosphodiester bond of DNA is replaced by an amide linkage in PNA. The nucleobases in PNA are connected to backbone through a methylene carbonyl linker. One of the notable features of the PNA molecule is its charge neutrality, which affects the properties of PNA. Although originally developed Peter Nielsen in 1991as a triplex forming oligonucleotide for recognition of duplex DNA as third strand, the extraordinary features and properties of PNA has made it a versatile mimic of DNA.[27] The structural features of PNA with six-atom spacing between the adjacent nucleobases in the PNA backbone similar to natural DNA and two atom spacing

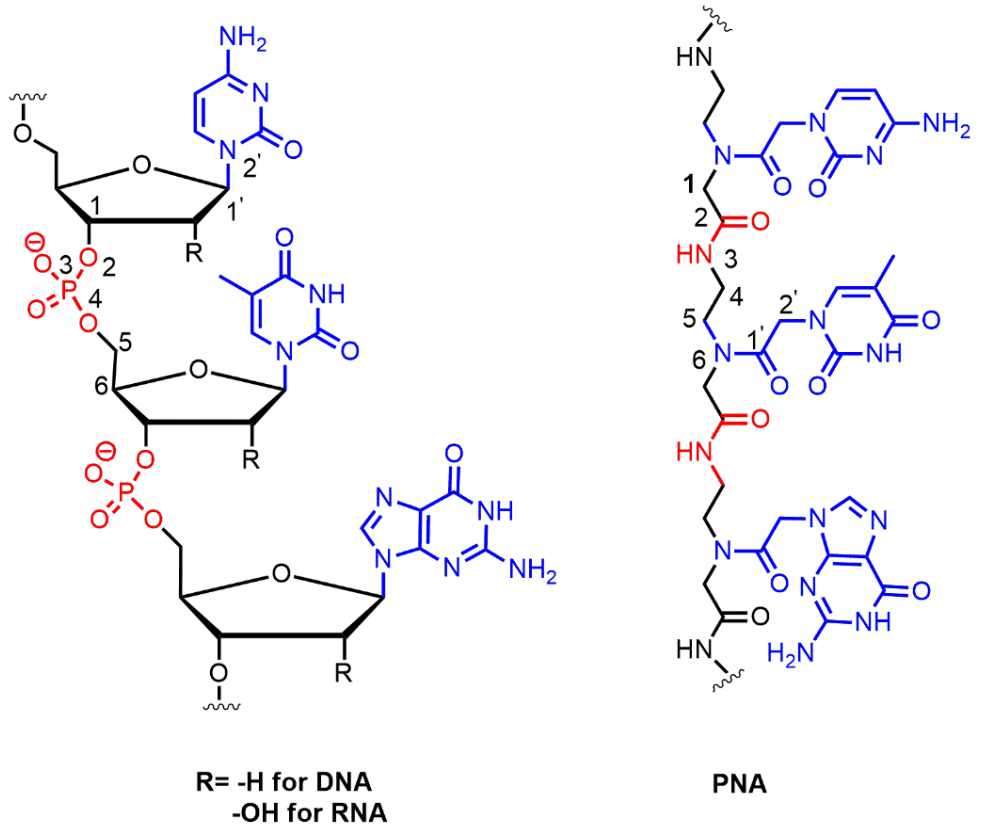


Figure 1.11. Structures of DNA and PNA

between the backbone and nucleobases makes it bind to complementary DNA / RNA. As it is neither a peptide nor a nucleic acid, PNA is stable to both proteases and nucleases, negligibly toxic but fail to trigger the RNase H activity. PNA exactly mimics the properties of DNA like Watson Crick base pairing with complementary DNA / RNA and triplex formation with the DNA duplexes. PNA:DNA and PNA:RNA hybrids show higher stability (Tm, melting temperature) compared to DNA:DNA/DNA:RNA hybrids and a complementary strand can invade a DNA duplex and displace the appropriate DNA strand (strand invasion). The higher stability of PNA:PNA/DNA duplexes is due to the neutral character of PNA backbone, which obviates the electrostatic repulsions of phosphate anionic charges present in DNA:DNA/RNA

duplexes. The effect of a mismatch in sequence on the thermal stability of PNA:PNA/DNA duplexes is more pronounced compared to the DNA:DNA duplexes. Because of these superior properties, PNA has attracted the attention of medicinal chemists and molecular biologist for the development of gene therapeutics, diagnostics and for material science.[27c-27h]

1.2.1 PNA duplex structure

The X-ray structure of PNA:PNA, PNA:DNA, PNA:RNA duplexes and NMR spectroscopy analysis have revealed a number of interesting features of PNA complexes.[29] The structural differences between PNA and DNA is reflected the π-π stacking interaction between the nucleobases, the profile and dimensions of the composed helix, the nature of major and minor grooves (Figure 1.12).

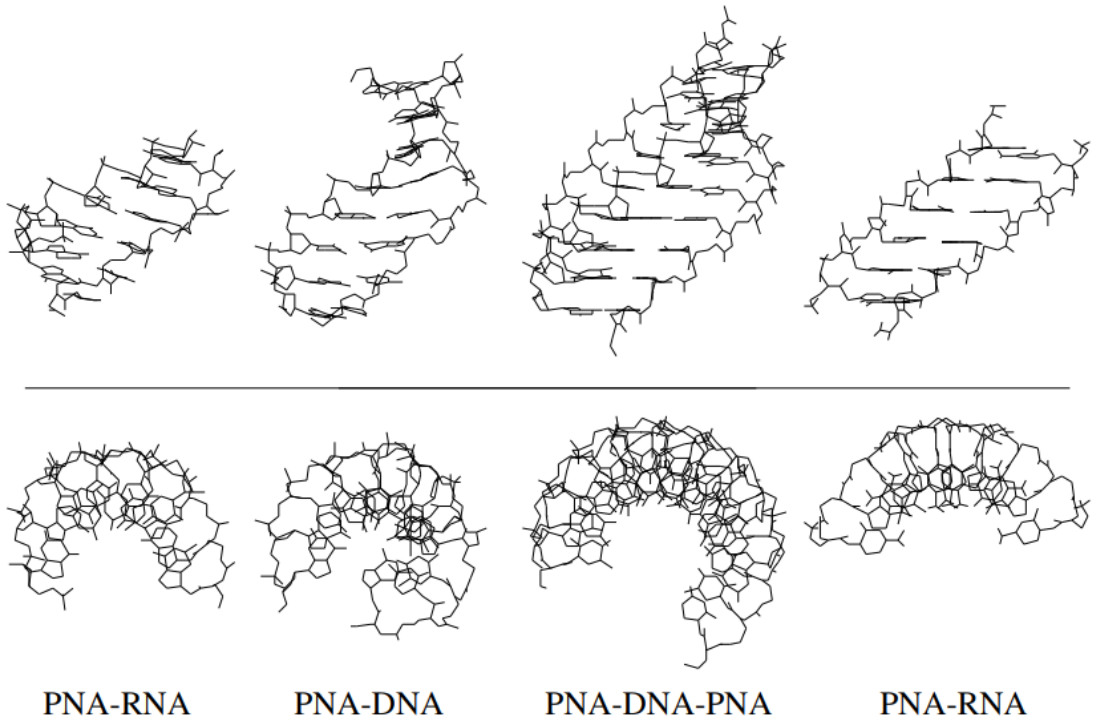


Figure 1.12. Structure of various PNA complexes (Upper- side view, below-top view)[29]

Typical B-helix seen in DNA duplex has a diameter of 20 Å and pitch of 10 bp / turn. A-form of helix typically observed in RNA duplexes is characterised by a diameter of 20 Å and a pitch of 11 bp/ turn, with the bases slightly tilted relative to the helix axis (20^{o}). In contrast PNA duplexes adopt the P- helix, which is characterised by a wide PNA duplex (28 Å), larger helical displacement of base pairs and a large helical pitch (18 bp). This suggests that PNA duplexes (PNA:cDNA/cRNA) adapt a structure which is combination of of DNA and RNA duplexes. PNA can form duplexes with complementary DNA/RNA in both parallel and

antiparallel orientations (Figure 1.13) obeying Watson-Crick base pairing rules. In parallel duplex, the N-terminus of PNA is priented towards 5'-terminus of DNA/RNA, while in antiparallel duplexes, the C-terminus of PNA is oriented towards the 5'-terminus of DNA (Figure 1.13). Antiparallel duplexes generally exhibit higher thermal stability compared to the corresponding parallel duplex.

Figure 1.13. Antiparallel and Parallel mode of PNA:DNA duplex formation[30]

1.2.2 Triple helix formation by PNA

Polypyrimidine PNA hybridises with complementary polypurine DNA to form PNA_2: DNA triplexes[31] (Figure 1.14). One of the PNA strand in the triple helix forms Watson-Crick hydrogen bonding while the opther PNA strand pairs with DNA via the Hoogsteen hydrogen bond formation. For a single poly pyrimidine DNA strand in PNA_2 : DNA triplex, the orientation of both PNA strand can be either parallel or antiparallel to the DNA strand (Figure 1.14). For two different polypyrimidine PNA strand to form triple helix the Watson-crick strand should be antiparallel orientation while the hoogsteen strand should be in parallel orientation with respect to the DNA. The stability of the triplexes depends on the selectivity of the PNA: DNA duplex while the third strand contributes moderately to selectivity. Triplex formation of polypyrimidines facilitates the strand invasion process, a unique property of Peptide Nucleic Acid.

1.2.3 Strand invasion by PNA

Strand invasion is a unique property exhibited by PNA oligomers. PNAs displace one strand of double helical DNA to form strand invasion complexes.[32] Strand invasion property of PNA makes it a potential candidate for antigene therapeutics. Depending on the mode of binding, strand invasion process can be classified into four types as: strand invasion via (i) triplex formation (ii) triplex invasion (iii) invasion of duplexes and (iv) double duplex invasion. The strand invasion by triplex involves formation of triplex strand with the double stranded DNA to form DNA_2:PNA complex. Triplex invasion involves formation of a triplex with single strand of DNA via Watson-Crick and Hoogsteen hydrogen bondings, while the second DNA

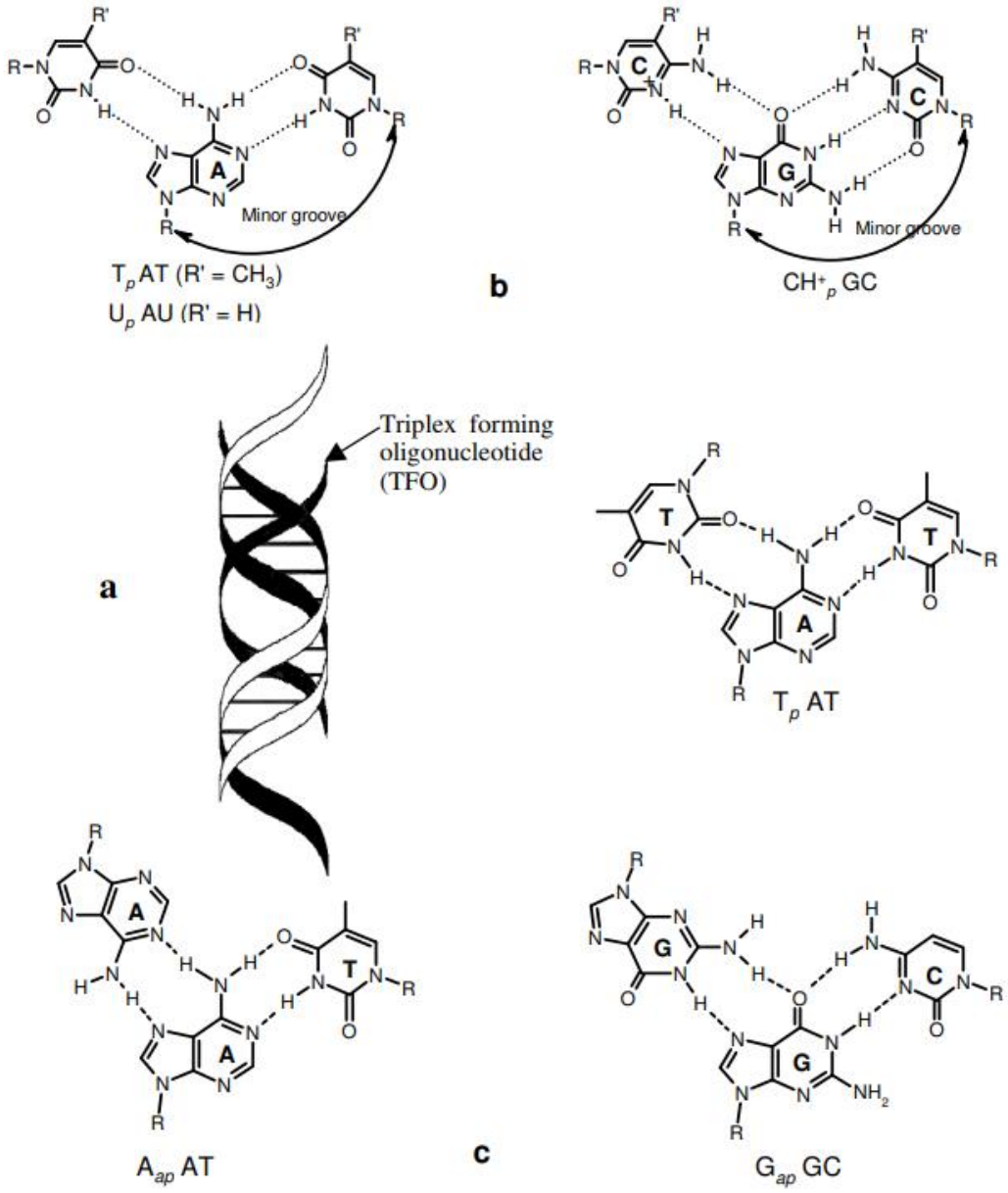


Figure 1.14. (a) A ribbon model demonstrating the relative position of a triplex forming oligonucleotide (TFO) in the major groove of DNA, (b) pyrimidine binding motif with the binding of TFO in a parallel (p) orientation to the polypurine strand of DNA involving protonated cytosine in the CH^+_p G:C triplet and T_p A:T. (c) The purine motif-the binding of TFO in antiparallel (ap) orientation. The three canonical base triplets in this motif: A_{ap}AT, G_{ap}GC and T_p AT.

strand of the original duplex remains as a loop (PNA_2:DNA_2). It is possible for a single complementary PNA strand to bind with one of the DNA via Watson Crick hydrogen bondingto invade the duplex (DNA_2:PNA) . Two complementary PNA strands can invade duplex DNA by forming duplex with each DNA strand (DNA_2:PNA_2, Figure 1.15).[33] Selective

double duplex invasion can be achieved by modifying the purines and pyrimidines nucleobases of the PNA sequence. Replacing adenine with 2,6-diaminopurine provides one extra hydrogen bond interaction compared to the original adenine nucleobase and favours duplex formation with DNA over triplexes (Figure 1. 16). To avoid the self-complementarity between the two PNA strands, uracil of one of the PNA strands is replaced with thiouracil which lacks one carbonyl group (C=O is a better hydrogen bond acceptor than C=S) and hence avoids duplex formation with diaminopurine substituted PNA strand but selectively forms duplex with adenine containing DNA strand. Two homopyrimidine PNA strands bind with one complementary DNA strand in 2:1 ratio to form PNA_2:DNA triplexes and hence invades the DNA duplex.

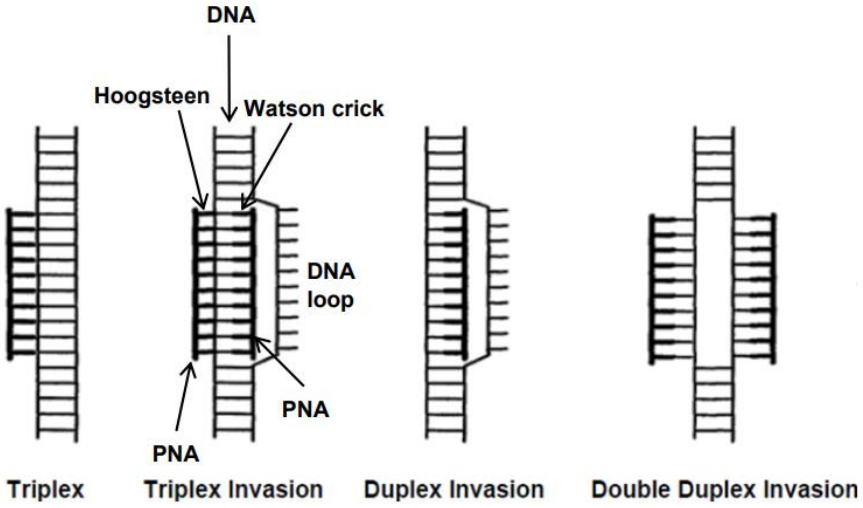


Figure 1.15. Various modes of strand invasion of ds DNA by PNA[32]

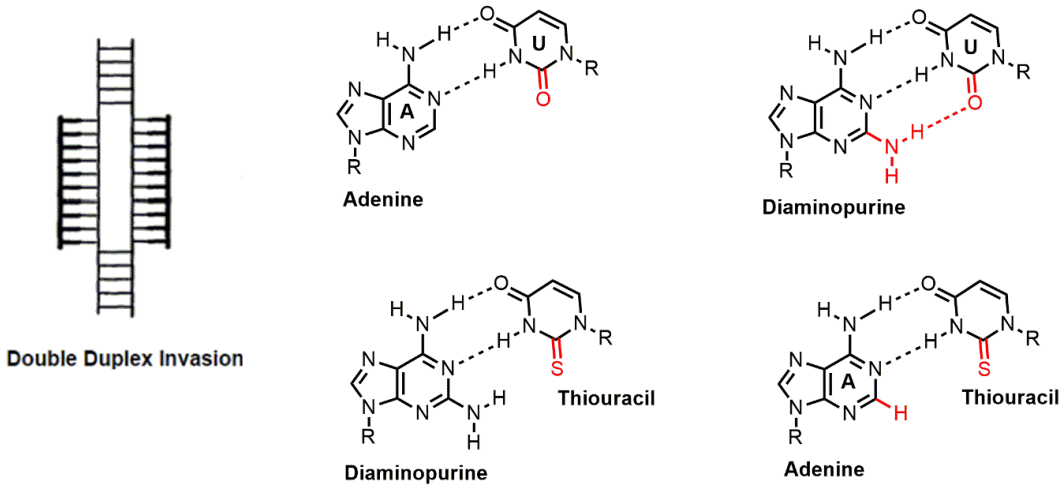


Figure 1.16. Double duplex invasion of pseudo complementary[33] PNAs

1.2.4 Higher sequence discrimination by PNA, salt effects and improved solubility

Compared with the original DNA, PNA shows higher specificity in binding towards the PNA:DNA duplex formation. PNA/DNA duplexes exhibit more destabilization per mismatch compared to DNA/DNA duplex.[34] Average reduction of melting temperature (T_m) due to a single mismatch in PNA/DNA duplex is 15^o (Range 8-20^o) while the destabilization due to a single mismatch in DNA sequence of DNA duplex is 11^o (Range 4-16^o). Thus, PNAs are more sequence specific in terms of hybridizing with complementary DNA and show higher discrimination for mismatch sequences and are more useful as hybridization probes.

Further, the T_m of the PNA:DNA duplexes are independent of salt concentration since the PNA backbone is neutral and hence the duplexes experience less electrostatic repulsion copared to DNA:DNA duplexes with both strands are polyanionic. The melting temperature (T_m) of the DNA:DNA duplexes are hence highly dependent on the ionic strength of the medium,[35] requiring neutralisation of the electrostatic repulsions by Na^+ cations.

Generally, PNAs are insoluble in water as the neutral backbone of PNA is hydrophobic and PNA oligomers have tendency to aggregate up to a degree depending on the length and base composition (pyrimidine/purine ratio) of the sequence.[36] Introduction of amino acids with positive charge NH_3^+ group (lysine) or negative charge COO^- (aspartic and glutamic acid) group at the α/β or γ position of the PNA backbone enhances the solubility of the PNA oligomers in water. This is achieved by coupling unmodified PNA oligomers with one to three lysines at the N-terminal of the PNA backbone at the end of PNA synthesis on solid phase.

1.2.5 Cell permeability of PNA

The superior and unique properties of PNA like strand invasion, triplex formation, higher stability of the PNA:DNA duplexes compared to the DNA:DNA duplexes, hybridization with DNA and resistance towards enzymes like nucleases and proteases make PNA a potential candidate for its application in antisense therapy. However, the biological applications are limited by poor cellular uptake of the PNA oligomers. Several modifications like conjugation of PNA with cell penetrating peptides,[37] sugar moieties (*GalNAc*), co-transfection with DNA, electroporation, microinjection, conjugation with lipophilic moieties etc has been employed to enhance the cell permeability of PNA. In addition, chemical modification of the pseudo peptide backbone, including introduction of chirality have been done to improve properties of PNA and selective binding with chiral targets (DNA/RNA).

1.3 Chemical modifications of PNA

PNA single strand with backbone composed of all single bonds (C-C and C-N) assumes random conformation and formation of ordered duplexes with DNA / RNA involves overcoming the entropy barrier. Higher and selective binding of PNA with DNA or RNA can be achieved by preorganization of the random conformation of PNA backbone by introduction of a variety of conformational constrains, varying from simple substitutions on α/β/γ sites to cyclised units on backbone.[38] Such chiral preorganization of backbone has led to improved affinity and selectivity of PNA towards binding the complementary DNA (Figure 1.17).

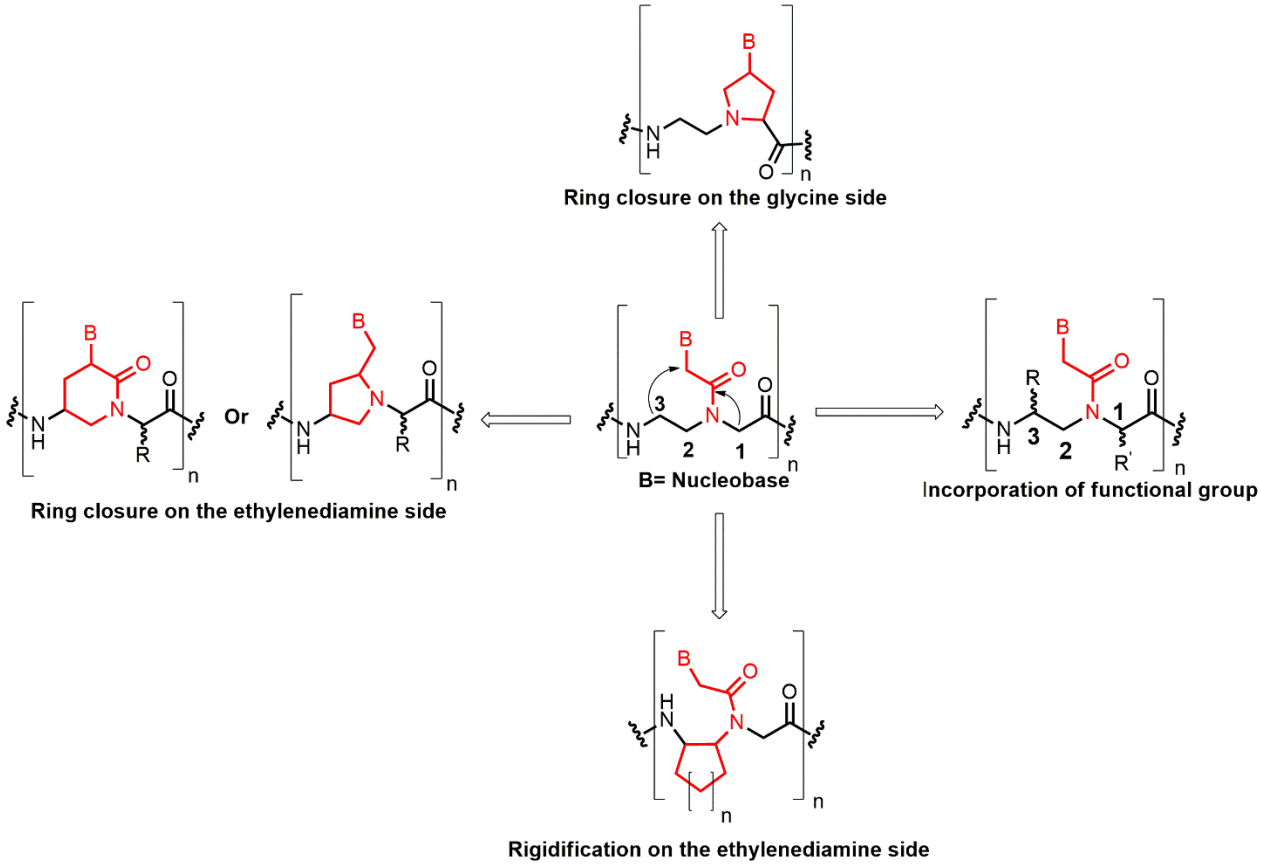


Figure 1.17. Strategies for inducing preorganization in the PNA backbone

1.3.1 Backbone substitution in acyclic PNAs

The N-(2-aminoethyl) glycine unit of PNA consisting of β/γ and α-methylene is the source of modification to get a chiral preorganized PNA. β/γ and α-methylene group of PNA backbone can be substituted with sidechains of α-amino acids to result in a chiral modified PNA (Figure 1.18). Properties of the derived PNA oligomers depends on the type of substitution in PNA monomers.

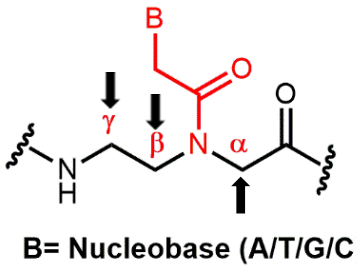


Figure 1.18. Structure of achiral unmodified PNA

1.3.1a *C_α-modified PNAs*: Nielsen et. al.,[39] reported the first C_α-modified PNA in 1994, where the Cα carried a methyl group (Figure 1.19). The Cα-methyl PNAs were derived from L/D alanine and the C_α-L-alanine substituted (*S*)- PNA destabilized the PNA:DNA duplex while C_α-D-alanine substituted (*R*)-PNA was similar in binding as unmodified PNA. D-lysine derived (*R*)-PNA carrying alkylamino sidechain at Cα (Figure 1.19) showed higher binding affinity for complementary DNA compared to the L-lysine derived (*S*)-PNA oligomers. This indicated the role of positive charge in the enhancement of melting temperature of PNA:DNA duplex.[39] Like standard PNA, the C_α modified (*R*)-PNA derived from D-amino acids prefer P-helix formation in the PNA:DNA duplexes.[39] The effect of mismatch sequences on the duplex stability of D-amino acids derived (*R*)-PNA:DNA duplexes is more pronounced compared to the unmodified PNA:DNA duplexes.[39] PNA-peptide conjugates incorporated with three or four D-lysine derived (*R*)-PNA were tested for cell penetrating ability shows significant cell permeability in most of the tested cells[40].

D-Arginine derived (*R*)-PNA oligomers consisting of guanidium group at Cα-position (Figure 1.19) exhibit destabilization of the PNA:DNA duplexes. However, PNA decamer containing multiple Cα-guanidium PNA unit (α-GPNA) shows improved binding affinity suggesting that electrostatic interactions between positive charge of guanidium and negative charge of the anionic phosphate group of DNA,[41] dominate the binding. C_α L(*S*)/D(*R*)- aspartic acid and glutamic acid derived PNA oligomers showed significant destabilization of the PNA:DNA duplex due to the destabilizing electrostatic repulsions between the negative charges of PNA backbone and phosphate group of DNA.[42] Steric hindrance induced by the bulky sidechain of tyrosine, histidine, tryptophan, phenylalanine and valine destabilize the PNA:DNA duplexes with lower T_m compared to the unmodified PNA.[43] α-PNA containing 1-(2,2-dimethylcyclobutyl) ethenone substitution exhibits equal affinity for complementary DNA and RNA.[44] The D-amino acid substitutions correspond to Cα(*R*) configuration and L-aminoacids lead to Cα(*S*) configuration and thus the Cα(*R*) chiral substituents tend to stabilize

the PNA oligomers better than Cα(*S*) substituents. One of the interesting Cα-modification is introduction of *gemdimethyl* (*gdm*) modification.[45] Unlike other monosubstituted modifications, *gemdimethyl* (*gdm*) substitution produces achiral PNA and the oligomers derived from C_{α}-*gemdimethyl* substitution exhibits improved binding affinity for DNA/RNA, with preferential binding to DNA than RNA. C_{α}-*gemdimethyl* substitution also results in selective stabilization of Z-rotamer around the tertiary amide bond.[45,47]

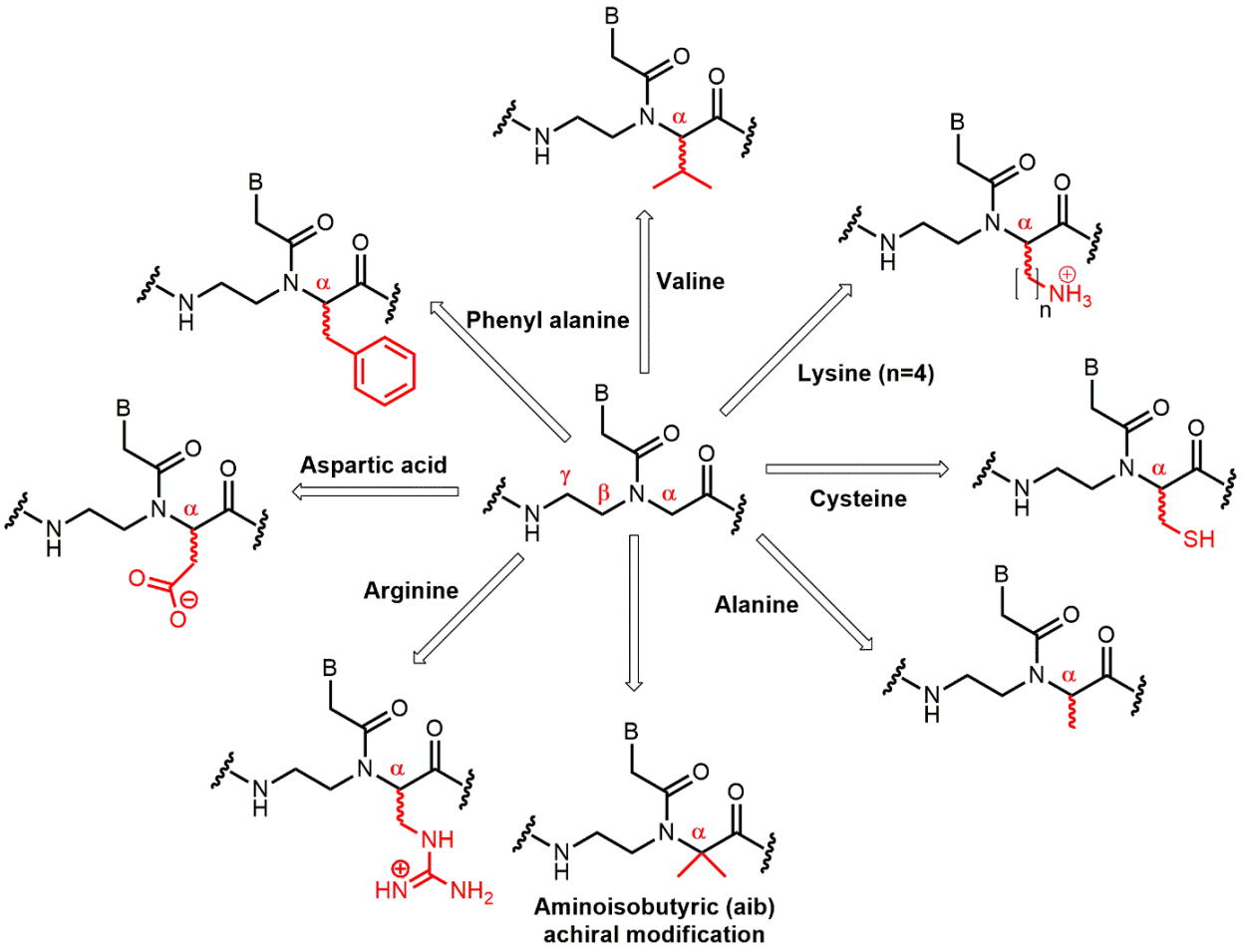


Figure 1.19. Cα-modified analogues of PNA

1.3.1b C_{β}-modified PNAs: Sugiyama *et. al.*[46] first reported the β-substituted PNA oligomer derived from L and D alanine, where the Cβ carries a methyl group (Figure 1.20). These PNAs exhibit similar binding affinity with complementary DNA as unmodified *aeg*-PNA.[46] C_{β}-PNA oligomer derived from Cβ-(*R*) methyl substituted ethylene diamine shows poor binding with complementary DNA . Circular dichroism (CD) study of C_{β}-*S*-PNA and C_{β}-*R*-PNA reveals that C_{β}-*S*-PNA prefers right-handed helical structure, while C_{β}-*R*-PNA prefers left-handed helical structure.[46] Another notable modification is the *β-gemdimethyl*(*gdm*) modification, which leads to achiral PNAs.[47] This substitution preferentially stabilized the E-rotamer and exhibits

differential binding affinity for cDNA and cRNA. The plausible reason for the rotamer selectivity is the avoidance of steric clash between bulky *gemdimethyl* group and the δ-methylene of the PNA backbone. C*β-gemdimethyl*(*gdm*)-PNA-T stabilised the E-rotamer while *α-gemdimethyl*(*gdm*)-PNA-T, stabilised the Z-rotamer.[47]

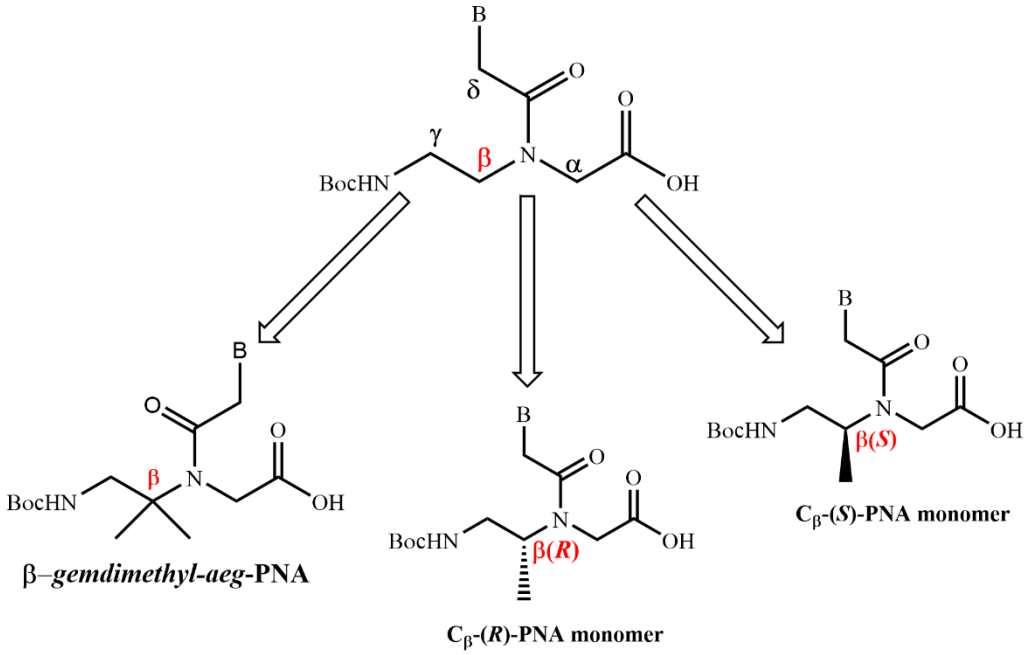


Figure 1.20. Cβ-modified analogues of PNA

1.3.1c Cγ-modified PNAs: In contrast to Cβ-substituted PNAs, Cγ-substituted PNAs have been widely studied (Figure 1.21). The first Cγ-substituted PNA gain attention in 2005 due to its interesting properties.[48] The Cγ-substituted PNAs can be synthesised starting from D/L-amino acids by synthetic routes for them to finally appear as substituted moieties in the ethylenediamine fragment of *aeg*-PNA backbone. Cγ(L/*R*)-lysine substituted PNA shows higher binding affinity for complementary DNA compared to the Cγ(D/*S*)-substituted PNAs. The effect of mismatch sequences on the binding affinity of Cγ(*R*)-lysine-PNA:DNA duplexes is more pronounced than unmodified PNA. Cγ-(*R*/*S*)-cysteine-PNA (Figure 1.21) can be used for the application of native chemical ligation (NCL).[49] The Cγ-(*R*/*S*)-amino acid modified PNAs exhibit hybrid stabilities dependent on Cγ-stereochemistry.[50] Cγ-polypyrimidine PNA carrying Cγ-sulphate group destabilized the PNA_2:DNA triplexes due to repulsion of negative charges of sulphate on PNA and phosphates of DNA. However, the γ- sulphate modified PNAs showed antigene activity against ErbBr gene.[51] Incorporation of serine and alanine in the γ-PNA showed preference for right-handed helix in PNA: DNA duplexes similar to unmodified PNA.[52] Danith Ly *et al.*[53] substituted Cγ-position of the PNA with hindered side chains and

proved that the binding affinity of derived γ-PNAs are generally independent of the bulk of the side chain. Fully modified C_γ-guanidine -PNA oligomer carrying an arginine side chain shows better cell permeability compared to other PNA analogues. Manicardi *et al.*[54] showed that C_γ-guanidine PNAs also inhibit the activity of micro-RNA. γ-PNAs carrying consecutive γ-guanidine-PNA units exhibit significant anti-miR-210 activity compared to the similar alpha analogue.[54] Recently Danith Ly *et al.*[55] showed that C_γ(*R*)-ethylene glycol PNA derived from C_γ(*R*)-serine-PNA preorganized into a right-handed helix exhibiting better binding affinity and higher sequence selectivity for DNA than RNA. The solubility of C_γ(*R*)-ethylene glycol PNAs in water are considerably higher than the parent *aeg*-PNA.

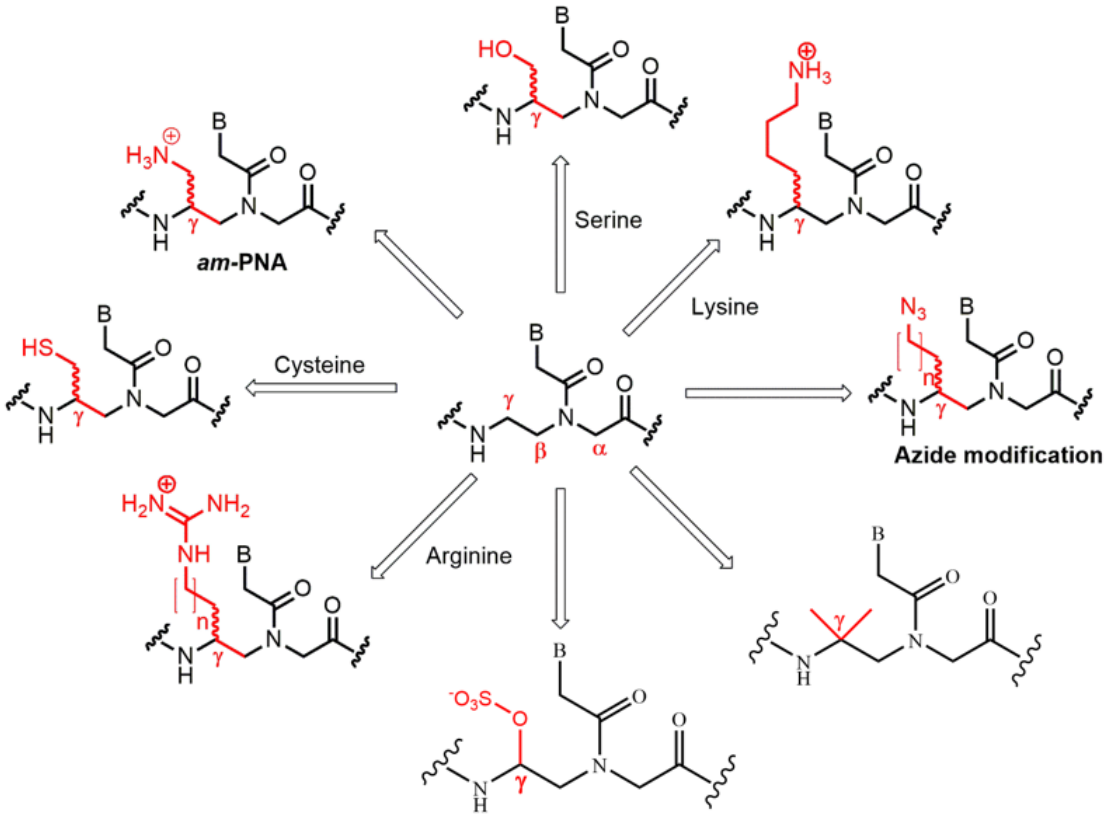


Figure 1.21. Important γ modified PNA analogues

Ganesh *et al.* [56] reported the synthesis and biophysical studies of C_α and C_γ aminoethyl PNA (*am*-PNAs) and showed that *am*-PNAs form stable duplexes with complementary DNA with the order of stabilization γ-(S)-*am*-PNA > α-(R)-*am*-PNA> α-(S)-*am*-PNA. These PNAs exhibit better cell penetrating ability with the decreasing order of uptake efficiency as γ(S)-*am*-PNA > α(R)-*am*-PNA> α(S)-*am*-PNA. The length of the side chain is one of the determining factor for the better duplex affinity of γ-guanidino and γ-azido PNAs.[57] A shorter spacer in the side chain increases the binding affinity for γ-guanidino PNA and stabilize the PNA:DNA

duplexes. The azido functionality at the Cγ-position has the applicability of the the azide for easy conversion to triazole by click reaction to link fluorophores. The achiral C*γ*-*gemdimethyl*(*gdm*) PNA T monomers show a rotamer mixture in solution unlike *α/β*-*gemdimethyl*(*gdm*) which prefer single rotamers in solution.[55]

1.4 Conformational preorganization through cyclisation of backbone

Acyclic PNA oligomers are conformationally flexible and the formation of acyclic PNA:DNA/RNA duplexes suffer from the undesirable loss on entropy of the PNA strand.[58] Loss of entropy during duplex formation of acyclic PNA is compensated by enthalpic advantages through Watson Crick hydrogen bonding and base stacking interactions.[58] Rotational barrier between cis (Z) and trans (E) interconversion around tertiary carbonyl group is moderately high for these PNAs and consequently hybridisation kinetics are different for parallel and anti-parallel hybrids.[59] Ganesh *et al.* showed that cyclisation of the PNA backbone leads to minimal loss of entropy during duplex formation and hence guided preferences for hybridisation with complementary DNA and RNA.[60] Many strategies have been reported to minimise the entropy loss during duplex formation and to have the favourable preorganized conformation for binding with complementary DNA. A structurally preorganized and favourable conformation of the PNA oligomer shifts the equilibrium towards the duplex formation with minimum loss of entropy. Common strategies of preorganization include cyclisation of the *aeg*-PNA backbone in the glycine side, cyclisation in the ethylene diamine side, locking the conformation of PNA monomers by introducing cyclic moieties in the glycine/ethylene diamine side of the backbone etc (Figure 1.22).

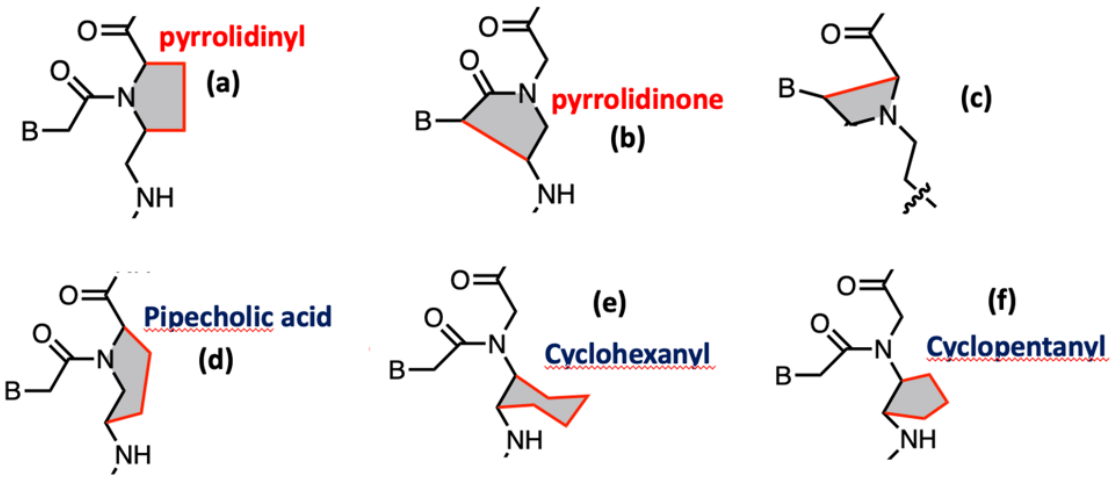


Figure 1.22. Conformationally constrained PNA monomer fragments

Pyrrolidinyl PNA (*py*-PNA): These arise from connecting the Cα-methylene of the backbone with the β/γ-methylene through a five membered ring[61] (Figure 1.22a). High rigidity and structural incompatibility of the *py*-PNA backbone restricts the *py*-PNA:DNA duplex formation and consequently no significant binding of *ap*-PNA with the DNA was observed.

Aminoethylprolyl PNA (*aep*-PNA): Aminoethylprolyl PNA (*aep*-PNA) were designed by connecting the α-methylene with the Cδ in side chain of the PNA backbone via a five membered ring[63] (Figure 1.22c). 4(*S*)-2(*S*/*R*)-*aep*-T PNA oligomers show strong binding affinity for complementary DNA sequences. (2*R*, 4*R*)-*aep*-PNA-T_8 exhibits strong binding with the complementary Poly rA to form duplexes.[64]

Aminoethylpyrrolidinone PNA (*aepone*-PNA): Aminoethylpyrrolidinone PNA were obtained by connecting the Cβ with the side chain Cδ of the PNA backbone (Figure 1.23) giving rise to a five membered lactum ring.[65] PNAs derived from *aepone*-PNA monomers exhibit exceptional stability of the PNA_2:DNA triplexes compared to the *aeg*-PNA.[66] *aepone*-PNA-T_8 oligomers preferentially bind with complementary DNA over complementary RNA showing structural discrimination.

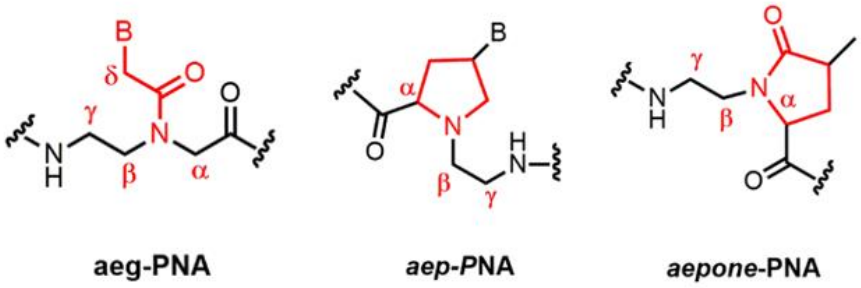


Figure 1.23. Molecular structures of *aeg*-PNA, *aep*-PNA, *aepone*-PNA

***trans*-Cyclopentane PNA (*tcyp*-PNA):** Appella *et al* [67] reported the (*S*,*S*)trans-cyclopentyl PNA (Figure 1.24) which is derived from fusing a cyclopentyl ring to Cβ-Cγ bond and it showed poor binding with complementary DNA.

Pyrrolidine PNAs: Pyrrolidine PNAs were derived by connecting Cβ with Cδ [68] (Figure 1.23). or Cγ with sidechain carbonyl C of t-amide group. (2*R*,4*S*)-Pyrrolidine-A_8 PNA shows stable duplex formation with the DNA/ RNA.[69] Similarly, the (2*R*,4*R*)- pyrrolidine-A_8 PNA shows strong binding affinity towards the DNA/RNA.[70] (2*R*,4*R*)-pyrrolidine-T PNA oligomer exhibits destabilization of the PNA:DNA/RNA duplexes.[71] Introduction of a bridging

methylene between N-terminal amine and tertiary amide nitrogen gives another derivative of pyrrolidine PNA where the nucleobase is away from the ring[72,73] (Figure 1.24). Enantiomeric pairs (2*S*,4*S*) and (2*R*, 4*R*) are found to prefer the formation of antiparallel complexes with RNA with higher binding affinity than *aeg*-PNA or other diastereomers.[73]

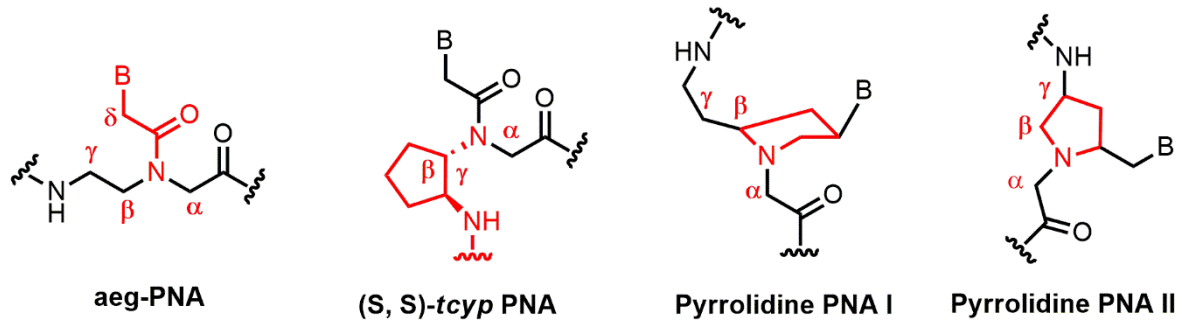


Figure 1.24. Structure of cyclopentane *aeg*-PNA, *trans*-cp PNA, pyrrolidine PNA (I and II)

Pipecolic PNA and piperidinyl PNAs: The pipecolyl PNAs are derived from aminopropyl glycyl-PNA (*apg*-PNA, Figure 1.25). Cyclisation of the backbone glycine methylene with the N-terminal methylene by introducing ethylene bridges gives the pipecolic PNA[74] (Figure 1.25). Incorporation of trans (2S,4S) monomer in fully modified PNA-T_{10} oligomer destabilized the PNA_2:DNA triplexes. *aeg*-PNA sequence containing a 5-aminopipecolyl unit exhibits stronger binding affinity for complementary DNA. Isomeric *aepip*-PNA and *piperidinyl*-PNA units substituted at the various positions of *aeg*-PNA-T_8 sequence shows a stronger tendency for PNA_2:DNA triplex formation.[75]

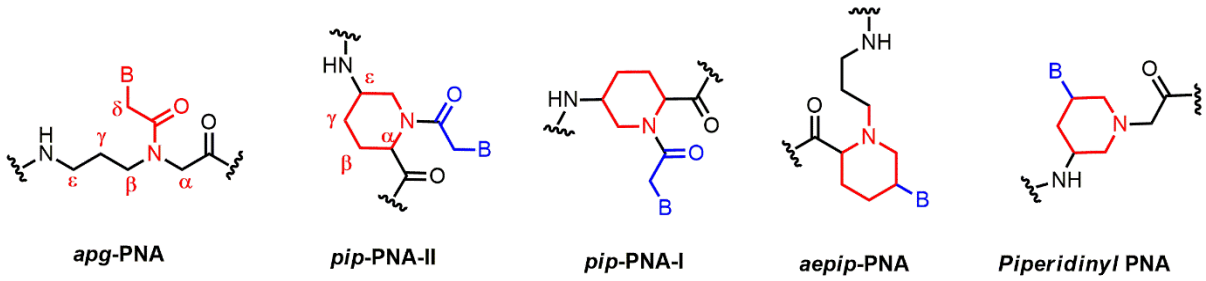


Figure 1.25. Pipecolic and piperidinyl PNAs

Cyclohexyl (*ch*-PNA) and cyclo-pentyl PNAs (*cp*-PNA): Ganesh *et. al.* first reported the cyclohexyl PNA where cyclohexane ring was introduced in the PNA backbone for preorganization of the oligomer[76] (Figure 1.26). Cyclohexane ring was introduced at the C_β-C_γ bond to restrict the rotation around only one C_β-C_γ bond (Figure 1.26). *trans*-(*S*,*S*)-cyclohexyl-PNA derived oligomer shows similar binding affinity for complementary DNA as standard *aeg*-PNA but the enantiomer *trans-(R,R)*-cyclohexyl-PNA destabilizes the PNA:DNA duplex.

Interestingly the properties of *cis*-(*R,S*) and *(S,R)*-cyclohexyl PNA (*ch*-PNA) differ significantly in binding complementary DNA and RNA. *cis*-(*R,S*) and *(S,R)*-*ch*-PNAs show preferential binding of complementary RNA compared to DNA due to the structural preorganization with dihedral angle around C_β-C_γ bond similar to that found in PNA:RNA X-ray structure. Ganesh *et al.* also reported the synthesis and biophysical study of cyclopentyl[77] PNAs (Figure 1.26). Unlike cyclohexane, cyclopentane shows characteristic *exo/endo* puckering dictating the pseudo axial/pseudo equatorial dispositions allowing tortional adjustment to attain the required hybridisation competent conformations. *aeg*-PNA-T_{10} modified with *cis-(1S,2R)-cyclopentyl*-PNA and *cis-(1R,2S)-cyclopentyl*-PNA at defined positions show that *cis(1R,2S)- and cis(1S,2R)-cyclopentyl*-PNA modified oligomers show higher binding affinity for both DNA and RNA. Homo-oligomers of the *cis-(1R,2S)-cyclopentyl*-PNA-T and *cis-(1S,2R)-cyclopentyl*-PNA-T stabilized the PNA_2:DNA triplexes maintaining the base specificity. *(1S,2R)* and *(1R,2S)* substituted *ch*-PNA and *cp*-PNA incorporated into mix sequence show preferential RNA binding (higher T_m for PNA:RNA duplex) over DNA (lower T_m for PNA:DNA duplexes) irrespective of the ring size. Interestingly the melting temperature of cyclopentyl PNA (*cp*-PNA):DNA/RNA duplexes are always higher than corresponding cyclohexyl PNA (*ch*-PNA):cDNA/cRNA duplexes suggesting entropy loss due to preorganization is less for *cp*-PNA:DNA complexes compared to *ch*-PNA:DNA complexes.

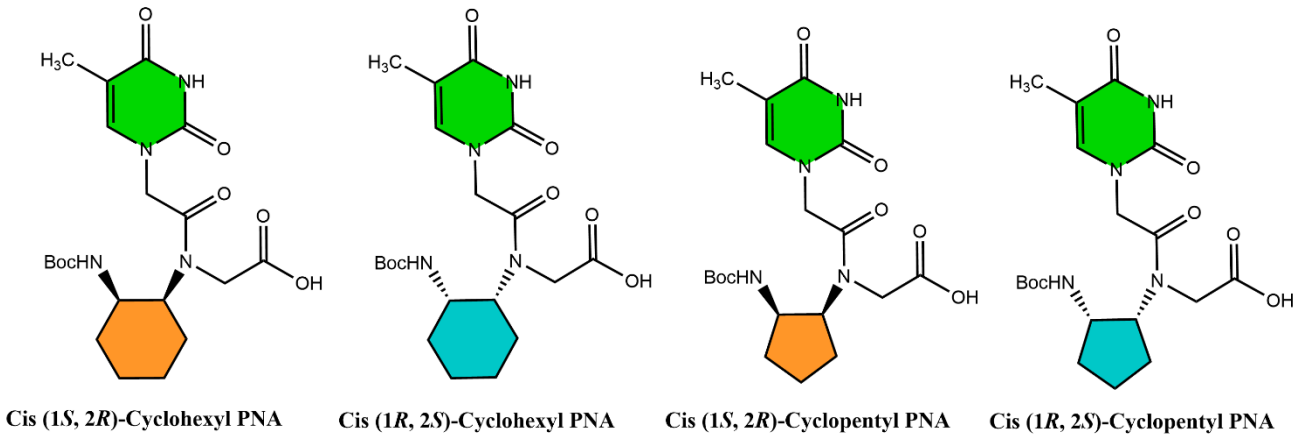


Figure 1.26. Structure of *cp*-PNA and *ch*-PNA

1.5 Applications of PNA

The versatile and unique properties of unmodified and modified PNAs make them perfect candidates for applications in gene therapy, molecular biology, medicinal chemistry,

and antisense therapy and to design supramolecular assembly. This section describes the applications of PNA in molecular biology and in therapeutics.

PNA as a potential antisense agent: With nearly undetectable toxicity, unique strand invasion property, stability towards hydrolytic enzymes like proteases and nucleases etc are the novel characteristic properties of PNA which makes PNA a superior gene therapeutic agent compared to other synthetic nucleic acid analogue. Two strategies are employed in the molecular biology for the therapeutic application of PNA include antigene and antisense methods.[9-15]

Inhibition of transcription: PNAs have potential to inhibit the transcription process through various mechanism including via triplex formation, strand invasion, strand displacement and by a duplex of duplex structure.[78] Peter E Nielson *et al.* have successfully displayed the inhibition properties of simple 8-mer PNA-T_8 which blocks the phase T_3 polymerase activity.[79]

Inhibition of translation: Antisense PNAs bind with m-RNA produced from DNA by transcription process to form the PNA:*m*-RNA duplexes, which affect the participation in translation process and hence stop the protein synthesis process.[80] In-vitro translation experiment in rabbit reticulocyte lysates proved that both duplex forming mix-PNA sequences and pyrimidine rich triplex forming sequences are useful for the inhibition of translation process.[81] Inhibition of protein translation process mediated by AtGCN1 is essential criteria for the survival of *Arabidopsis thaliana* in cold condition.[81b]

Inhibition of replication: Synthesis of daughter DNA from parent DNA molecules via DNA replication process can also be inhibited by PNA. Technically inhibition of replication process is possible if PNA binds with the complementary DNA and invades the stands of duplex DNA under physiological conditions. Another approach to inhibit replication process involves formation of duplexes with the single stranded DNA to form PNA:DNA duplexes. Taylor *et al.* [82] showed that DNA replication process of extra chromosomal mitochondrial DNA which largely stays single stranded during replication process, can be efficiently inhibited by PNA. Although several experiments prove the concept antigene and antisense therapy by PNAs, so far not a single PNA drug is in the clinical trials for applications of antisense therapy.

1.6 Synthesis of PNA and biophysical techniques

PNAs are synthesised on solid phase by using standard methods developed solid phase peptide synthesis. Hybridisation with complementary DNA/RNA is studied by biophysical

techniques like UV-visible spectroscopy, Temperature dependent absorbance measurement, circular dichroism and fluorescence spectroscopy.

1.6.1 Solid phase PNA synthesis

Peptides are synthesized either in solution or by solid phase synthesis (SPPS).[83] Peptide synthesis in solution involves coupling, separation of unconverted starting materials from desired product and purification of the peptides after each coupling step. The latter process is time consuming and tedious. In solid phase peptide synthesis developed by Merrifield,[84] long chains of peptides (10-mer -20-mers) can easily be synthesized and also capable of automation. The method involves porous polymeric beds functionalized with functional groups through linkers at the surface (Figure 1.27). The C-terminal aminoacidic with protected amino group is connected to the solid support through a cleavable linker. After the attachment of first amino acid, the N-terminus is deprotected to liberate free amine group. The reactive amino groups on the solid support is coupled with the desired N-protected amino acid monomer with free carboxylic acid, in presence of activating and coupling agents. Successive steps of deprotection and coupling leads to growth of the peptide chain with desired sequence of amino acid. The coupling reactions are normally performed using coupling agents like HBTU, HATU, DIC etc in presence of HOBt, to suppress possible racemization. Since the polymeric beds are insoluble in organic solvents, at every step, the product on the solid support can be easily separated from unreacted monomer and coupling agents by simple washing by filtration. After the completion of synthesis, the peptides are cleaved from the solid resin using appropriate reagents dependent on the nature of linker.

Currently the two most widely used strategies for SPPS are based on the nature of amino protection (Figure 1.28). Generally, the SPPS proceeds from C-terminus to N-terminus. In Fmoc strategy, the first step involves reaction of the primary amine group present on MBHA resin followed by coupling with the desired N(Fmoc) protected amino acid or PNA monomer with free carboxylic acid. The peptide product obtained remains linked to the resin through carboxy terminus and protected at N-terminus with Fmoc group. The base labile Fmoc group is then deprotected with piperidine and the liberated free amine is coupled with the next N(Fmoc) protected PNA monomer using standard coupling agents. The cycle of base deprotection, washing and coupling step is continued until the required length of the PNA oligomer is reached. In the alternative Boc-strategy, the primary amine on Rink-amide resin is coupled with N(Boc)-protected PNA monomer with free carboxyl group to get peptide on the

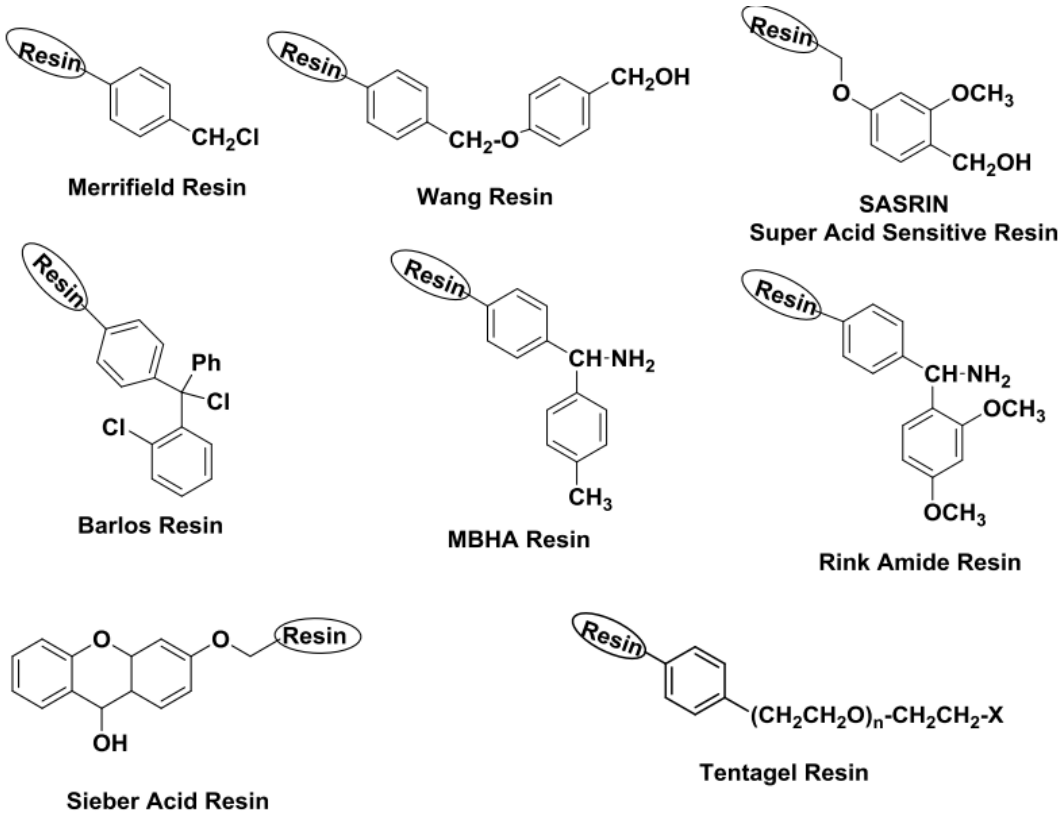


Figure 1.27. Chemical structures of the commonly used resins for SPPS

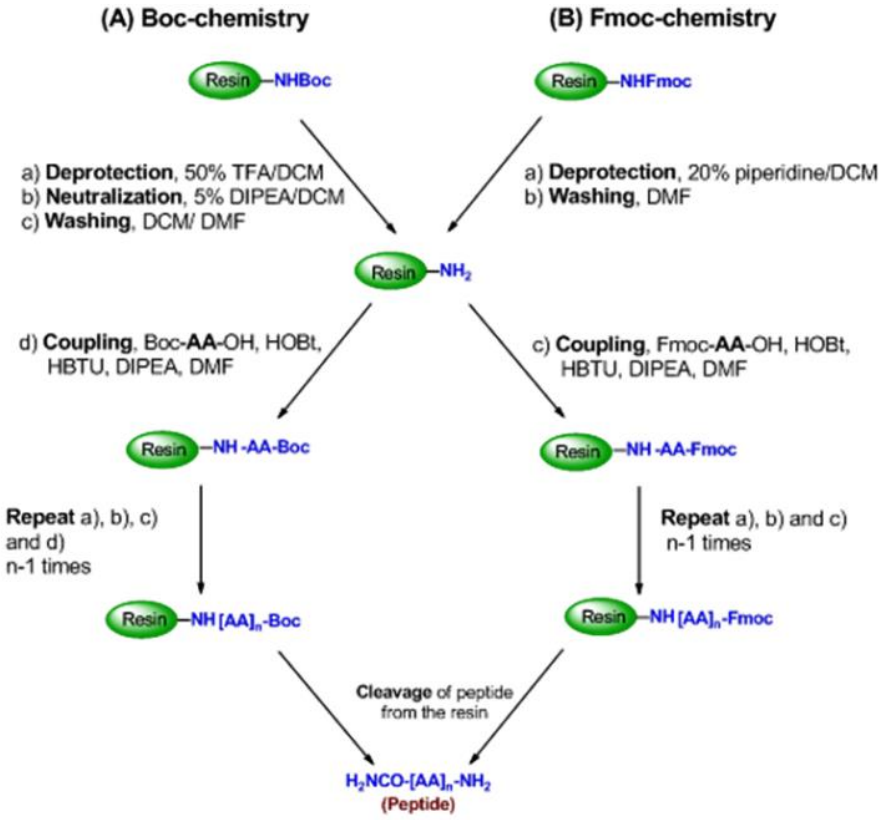


Figure 1.28. Schematic diagram showing steps of Boc and Fmoc strategy of SPPS

resin that carries Boc protected N-terminus. The acid labile Boc group is deprotected with TFA to liberate free amino group that is then subjected coupling with next Boc protected amino acid and the cycle of acid deprotection, washing and coupling continued until the required length of the PNA oligomer is reached. The resin bound PNA oligomer is cleaved using TFMSA[85] in TFA to liberate the product into solution.[86] The completely deprotected PNA oligomers obtained in solution are purified by reverse phase HPLC and characterised by MALDI mass spectrometry.

1.6.2 Biophysical techniques

Several biophysical techniques are used to study the properties of PNA:DNA duplexes and higher order structures. The most common techniques include temperature dependent UV absorbance measurement (UV-T_m) for assessing the strength of association of complexes, circular dichroism (CD) to examine the conformation of complexes, isothermal calorimetry (ITC) for evaluating the thermodynamics of complex formation and fluorescence spectroscopy for florescent labelled oligomers, microscopy to examine their nanoparticles and examine their cell permeation properties.

1.6.2a Temperature dependent UV-absorbance (UV-T_m) and Jobs plot

The duplex structures of DNA with DNA/RNA/PNA are held by a combination of base stacking and complementary base pairing through H-bonding. Upon heating, these interactions break down resulting in disassociation of the double helix into single DNA strands known as melting of DNA. This process can be monitored by following the change of absorbance of DNA bases at 260 nm as a function of temperature.[87] UV absorbance of nucleobases arise from the π-electron density of the nucleobase chromophores. The stacking of nucleobases in duplexes leads to interaction of π-electrons of the adjacent nucleobases, leading to lower UV absorbance known as hyperchromicity. During the melting of DNA, the breaking of H-bonds is accompanied by unstacking of nucleobases and consequently the UV absorbance of the nucleobases increases. The process of DNA melting is a co-operative process and can be monitored by following UV absorbance at 260 nm as a function of temperature. The plot of absorbance at 260 nm vs temperature shows a sigmoidal transition (Figure 1.29a) which suggests presence of two states of nucleic acid complexes in solution. At lower temperature duplex form predominates and shows low absorbance. Upon increasing the temperature, the duplex breaks into single strands resulting in increased absorbance with increase in temperature. The midpoint of the sigmoidal curve is known as melting temperature (T_m) at

which 50% DNA exists as duplex while 50% exist as single strands. Depending in the role of association and dissociation the transition can be sharp or slightly broadened. The accurate T_m can be obtained from the maxima of the first derivative plot of the absorbance (Figure 1.29b). Non-sigmoidal transition suggests single strand or random conformation of the DNA. The melting temperature represent the thermal stability of the corresponding Nucleic Acid Complexes, higher the melting temperature higher will be the stability of the nucleic acid complex.

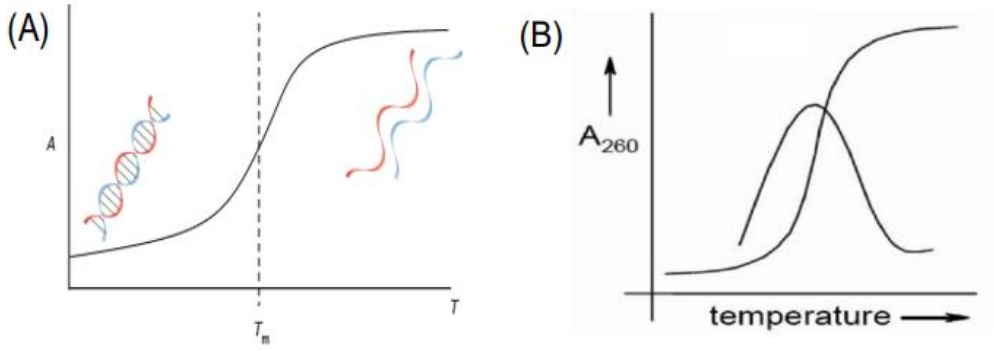


Figure 1.29. Temperature dependent UV-absorbance curve of nucleic acid duplexes at 260 nm. (A) Sigmoidal curve for DNA duplex (B) derivative plot of duplex.

In DNA triplexes, two strands are held by Watson-Crick hydrogen bonding and the third strand is held by Hoogsteen Hydrogen-bonding to the duplex. Upon heating, the triplex dissociates in a sequential way, with first transition arising from Hoogsteen strand melting followed by duplex melting. This is reflected as two transitions in the UV-melting curve (Figure 1.30a) for triplex formation. G-rich DNA sequences can also form tetraplexes, which show characteristic UV absorption at 295 nm. As tetraplex melts, the intensity of band at 295 absorption decreases and the UV melting curve followed at 295 nm shows negative sigmoidal profile. A similar situation exists with C-C^+ i-Motif formed by C-rich DNA sequences. This motif requires protonation N3 of C and this also shows UV band at 295 nm at acidic pH (<5.0). The melting transitions show a negative sigmoidal transition (Figure 1.30b). Thus, temperature dependent UV absorbance is an important technique to characterize DNA duplex, triplex and tetraplexes. Polypyrimidine PNA binds with complementary DNA in 2:1 ratio to form PNA_2: DNA triplexes[90] while the mix sequence PNA sequences form duplexes with the complementary DNA in 1:1 ratio. PNA complexes also exhibit similar UV profiles reflecting formation of duplexes, triplexes and tetraplexes. Figure 1.30 shows the melting profile of triplexes, G-quadruplexes and i-tetraplexes.

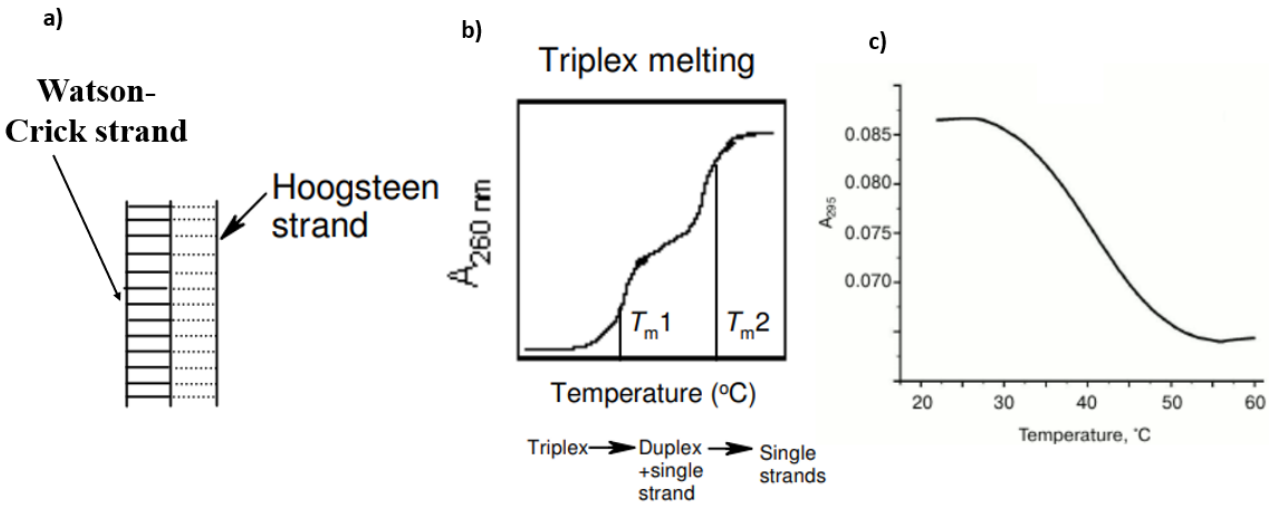


Figure 1.30. a) PNA_2: DNA triplex formation b) Dual melting UV-T_m profile of PNA_2: DNA complexes c) Negative sigmoidal melting profile of G-quadruplex and tetraplex motif.

The determination of the binding stoichiometry in the complex is important to assess the nature of the complex such as duplex, triplex etc., which is done by UV-Jobs plot.[88] In continuous variation method the relative concentrations of the two strands in the binary mixture is varied keeping the total concentration constant. UV absorbance at each composition of such generated binary mixture are recorded and plotted against the mole fraction of the two components (Figure 1.31). The total concentrations chosen must be within the range of the Lambert-Beer law applicability (absorbance range 0.1-1.0) and the sum concentration of the two components should be same at all compositions. In such a Jobs plot the point of inflection gives the stoichiometry of complex, for ex, for duplexes, it is 1:1, while for triplexes, it is 2:1.

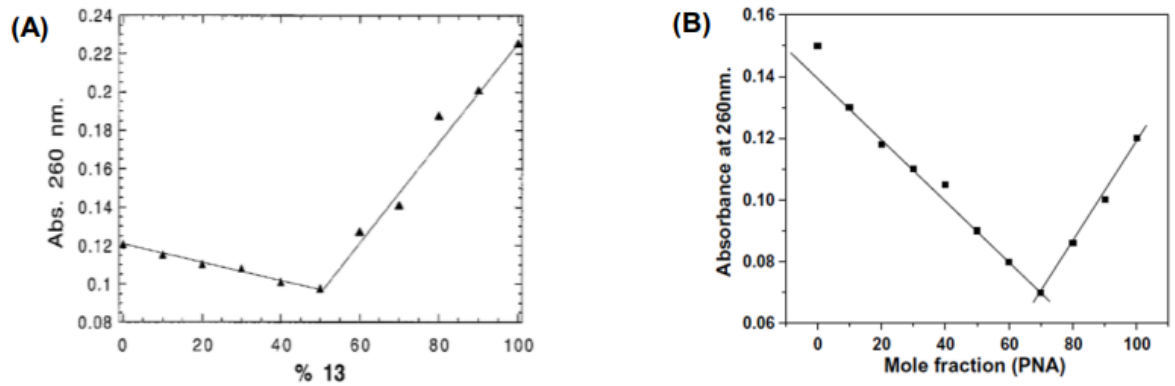


Figure 1.31. Job plot of UV-Continuation method A) Duplex B) Triplex

1.6.3a Circular Dichroism (CD)

The difference between the left-handed circularly polarised light and right-handed circularly polarized light upon interacting with a chiral medium is measured in the CD spectra[91]. Non-superimposable mirror images are known as enantiomers. Generally, the physical and chemical properties of enantiomer are same but the patten of interaction of an enantiomeric pair with left-handed and right-handed circularly polarised light are entirely different. CD spectral features give valuable information on formation of coiled structure or helical structure and sensitive to change in conformation of oligonucleotide or peptides upon binding with a ligand, formation of hybrid PNA with DNA / RNA , nature of different types of helices (α-helix and β-sheet, g-turn etc) formed by peptides and proteins, different helical pattern of A-DNA, B-DNA and Z-DNA.

Nucleobases in DNA/RNA are attached with the chiral sugar molecule and hence shows characteristic CD signals in UV region. The CD spectra of duplexes A-DNA, B-DNA and Z-DNA[92] with different helical structures exhibit unique and signature CD bands (Figure 1.31). The standard PNAs are achiral and may be composed of equal amounts of right-handed and left-handed helices and hence lack or show very weak CD bands. Upon binding with chiral DNA to form duplexes or triplexes the PNA: DNA or PNA_2: DNA complexes acquire chirality and hence shows signature CD bands[93] (Figure 1.31b). α/β/γ-modified PNAs are chiral in nature and show moderate CD bands without complexation with DNA and with complexation with DNA.[93b]

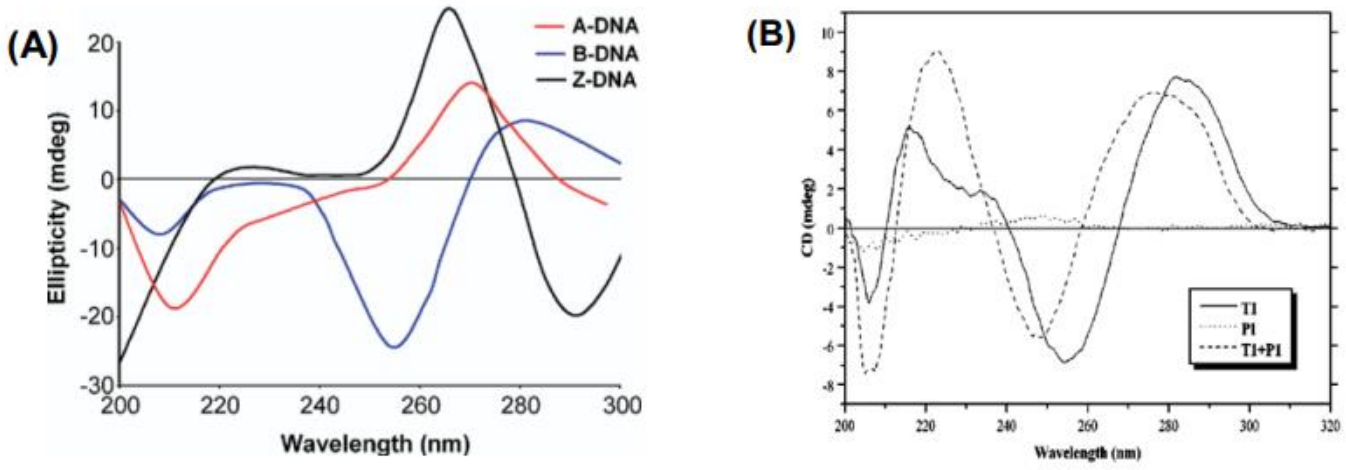


Figure 1.31. Typical CD spectra of (A) A-DNA, B-DNA and Z-DNA[92](B) PNA:DNA[93] hybrid

1.6.4 Fluorescence and ITC

Nucleobases (A, C, G, T, U) are inherently non-fluorescent, but upon substitution, or modification may become fluorescent (see Chapter 3). DNA and PNA bearing such chromophores are useful as hybridisation probes. The application of fluorescent nucleic acids is discussed in later chapter (Chapter 3).

Dissociation constant (K_a) represents the strength of intramolecular interaction between two species whereas the inverse of dissociation constant (K_a) is known as inverse dissociation constant (K_d)[94]. Determination of dissociation constant (K_a or K_d) can be performed by titrating one component to another and following the progress of complexation reaction using analytical techniques like UV, CD, NMR etc. The thermodynamic details of the complexation reaction are obtained by Isothermal Calorimetry (ITC) which measures the heat of reaction produced during the complexation at a constant temperature.

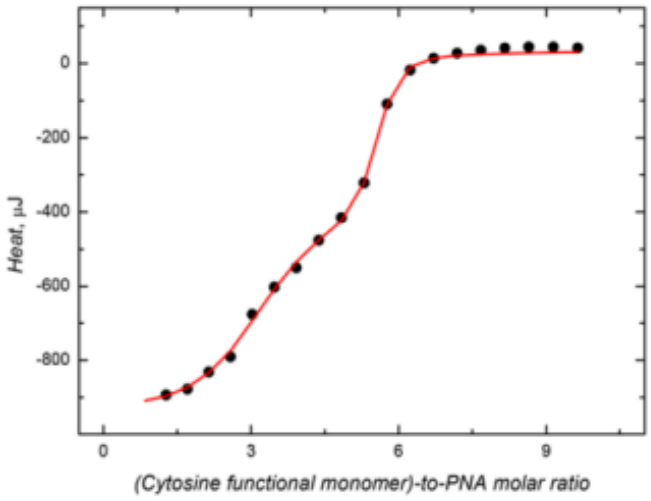


Figure 1.32. Typical ITC isotherm for 2-step binding complex

A small portion of the solution of titrant (eg. DNA, ligand etc.)[95] was added to the analyte (eg. DNA, RNA, Protein, PNA etc)[96] sequentially at a constant temperature (Below the melting temperature of the duplex for titration of PNA with complementary DNA) and the released heat was monitored during a typical ITC experiment. Experimentally standardized concentration of the oligonucleotide and buffer of the medium results titration curve which provides information regarding 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). Having the value of the enthalpy of association (ΔH) and the free energy of

association (ΔG), entropy of association (ΔS) can be calculated using the standard thermodynamic equations. Most of the cases the analyte DNA or PNA consist of unequal number of nucleobases and binds with the titrant with cooperative manner to result a double transition in the titration curve.[97] Hence values of all the major thermodynamic binding parameters can be obtained from a single ITC experiment. Figure 1.32 shows the standard titration curve obtained during an ITC experiment.

1.7 Present work: Rationale and overview

Classical *aeg*-PNA consisting of C-C, C-N and C-O bonds have very flexible in conformation. The slight conformational rigidity arises from of the tertiary amide bond which has restricted rotation around N-CO bond, resulting in two distinct rotamers.[98] The orientation where the tertiary amide carbonyl group is pointed towards the C-terminus glycine side is known as Z-rotamer, and when the carbonyl group is pointed towards the N-terminus ethylene diamine side the rotamer is E[98] (Figure 1.33). The rotational energy barrier calculated around the tertiary carbonyl group using variable temperature NMR was found to be around 19±2 kcal/mol and the Z:E rotamer ratio is 70:30.[98] It was recently shown that incorporation of *gemdimethyl* (*gdm*) group at the α and β position of the PNA backbone preferentially stabilizes Z-rotamer for α-*gemdimethyl*-PNA, E-rotamer for β-*gemdimethyl*-PNA and γ-*gemdimethyl* shows a mixture of rotamers as in *aeg*-PNA[99] (Figure 1.34). The possible reason for the selective stabilization of rotamers is the avoidance of steric clash between side chain δ-methylene and the gemdimethyl group.[99] For β/α-*gemdimethyl*-PNA the side chain δ-CH_2 and *gemdimethyl* groups are adjacent to each other and hence shows single rotamer to avoid the steric clash. In γ-*gemdimethyl*-PNA the side chain δ-CH_2 and gemdimethyl group is far away from each hence shows a mixture of rotamers[99] (Figure 1.34). In addition to the steric factors, the attractive C-H…O bonding between the δCO and β/γ-CH_2 also contributes to selective stabilization of E-rotamer.

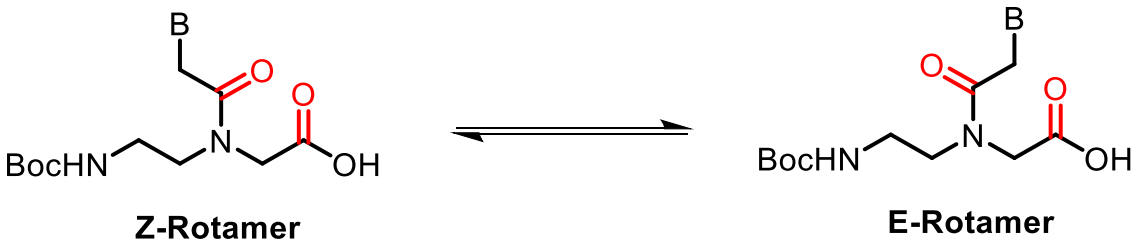


Figure 1.33. E and Z rotamers of *aeg*-PNA[98]

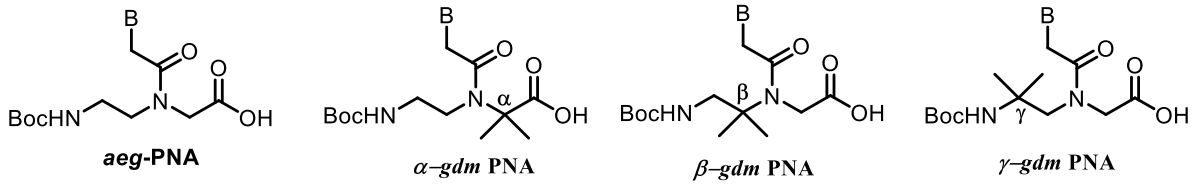


Figure 1.34. Chemical structures of substituted *gemdimethyl(gdm)* PNA[99] monomers

Chapter 2: Based on the above literature. it was thought to delete the repulsive steric interactions by eliminating *gemdimethyl* substitutions at Cα and introduce attractive H-bonding incorporating NH group at α position of the PNA backbone (Figure 1.35). Cα substitution by NH can manifest in potential hydrogen bonding opportunities in PNA backbone. In several peptides and biomolecules, such H-bonds redefine the local conformational constrains and the adaptive chirality of the nitrogen atom may impart interesting chiral the properties tothe molecules.[100] *aza*-PNAs with α-NH may participate in intramolecular hydrogen bonding with the N-terminal amide carbonyl group to turn the PNA backbone in a γ-turn like orientation as in peptides[101] (Figure 1.36). Formation of an extended γ-turn like motif in *aza*-PNAs may influence the rotamer selectivity and DNA/RNA binding properties. The homologue β-ala-*aza*-PNA extends the sixe of such intramolecular H-bond rings.

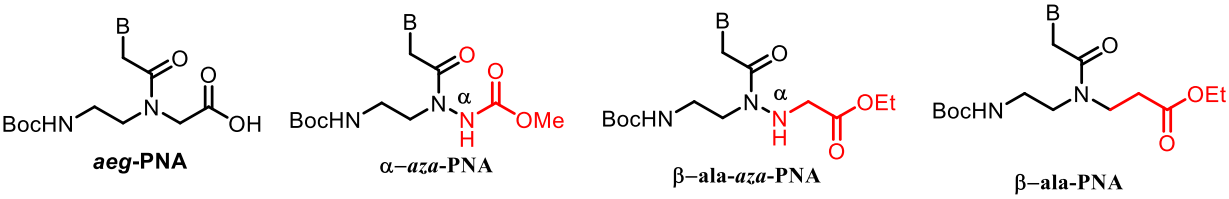


Figure 1.35. Chemical structure of the designed *aza*-PNA monomers

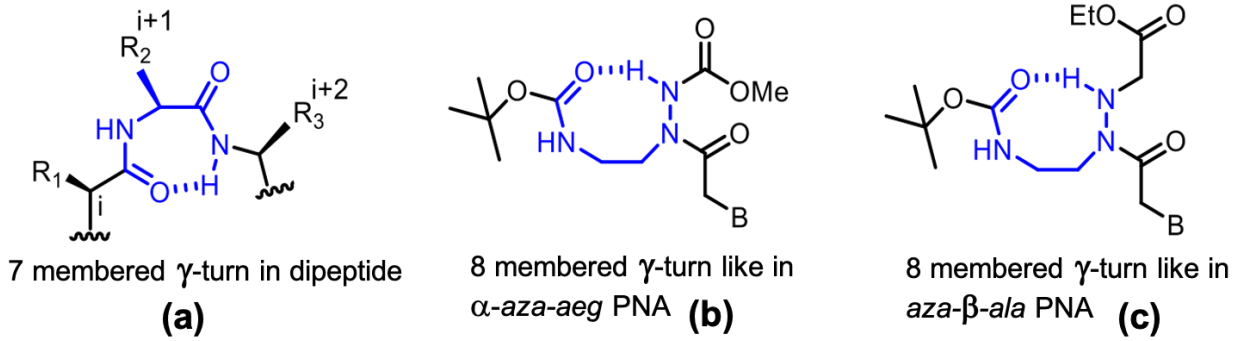


Figure 1.36. Strcutures of (a) γ-turn in dipeptides (b) γ-turn-like *aza*-PNA and (c) γ-turn-like *aza*-β-*ala*-PNA

The PNA oligomer of α-*aza*-PNA would result in a repeating planar urea linkage (Figure 1.37) could rigidify the oligomer and consequently impact the hybridisation efficiency of the α-*aza*-PNA, making it inappropriate for hybridization with complementary DNA/RNA. To overcome such problems in α-*aza*-PNA, the homologue with an additional methylene group (ε-CH_2, β- ala-*aza*-PNA, Figure 1.35) between α-NH and was designed. The oligomer derived from β-ala-*aza*-PNA would result in a polyamide backbone (Figure 1.37) like in *aeg*-PNA, which may be more compatible for hybridisation with DNA/RNA. To understand the structural role of α-NH in β-ala-*aza*-PNA, control monomer of β-ala-*aza*-PNA was synthesized where the α-NH is replaced by isosteric methylene group (Figure 1.35) termed as β-*ala*-PNA. It is shown by ^{1}H NMR and computational studies that the *aza*-PNA monomers selectively form E-rotamers driven by intramolecular H-bonding in backbone.

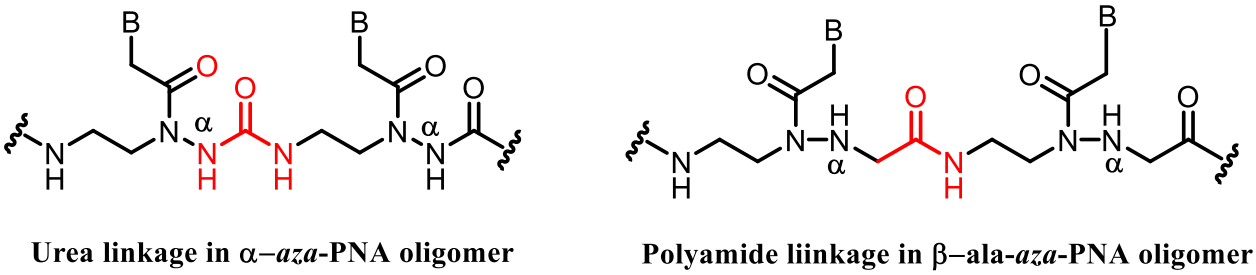


Figure 1.37. Urea linkage and polyamide linkage in α-*aza*-PNA and β-ala-*aza*-PNA oligomers

Chapter 3: Gazit *et.al.*[102] reported the single crystal structure of GC di-PNA. The crystal structure exhibits Watson-Crick hydrogen bonding between the two complementary strands of GC di-PNA and effective π-π stacking interaction between the nucleobases.[102] The carbonyl group of the tertiary amide is oriented towards the glycine side of the backbone in both the G and C monomers with a Z-rotamer stabilization in the solid state. The GC-Dimer dissolved in buffer solution shows excitation dependent emission spectra although it lacks the presence of any conventional fluorophore in the molecule. Apart from the fluorescence it exhibits electroluminescence properties. The fluorescence of *aeg*-PNA-GC dimer could arise from structural packing of the molecule leading to stacking interactions between the nucleobases, directed by Z-rotamers.[102] With this literature background, it was thought to examine the effect of E-rotamers on stacking of bases, crystal packing and fluorescent properties. Exclusive stabilization of E and Z rotamers is possible by incorporating *gemdimethyl* group in the α and β position of the PNA backbone (Figure 1.38) as reported earlier. [99]

Hence the objective was to synthesise and examine the α/β-*gemdimethyl* guanine monomers and dimers (Figure 1.38) and study their fluorescence properties, extending to fatty

acid conjugated monomers and dimers to induce hydrophobicity based aggregation. The studies revealed interesting rotamer dependent fluorescent behaviour in α/β-*gemdimethyl* guanine monomers in solid as well as solution state. Apart from the *guanine* monomers, the corresponding *cytosine* monomers are also studied.

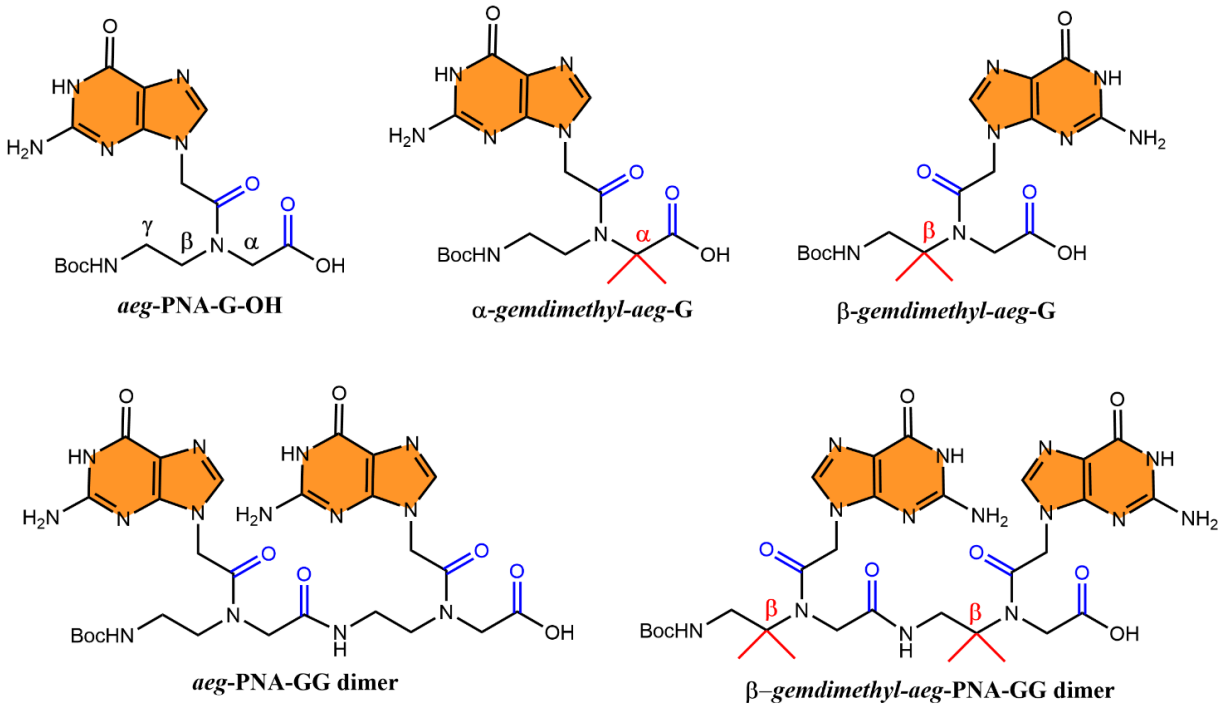


Figure 1.38. β/α-Gemdimethyl-PNA-Guanine monomers and dimers

Chapter 4: The solid phase synthesis of *bimodal* PNAs with mixed base sequence on both sides involve complex synthetic schemes.[103] These employ orthogonal protecting groups to enable selective deprotection of Boc at N-terminus for coupling to extend the main chain and selective deprotection of Fmoc-group on sidechain to enable coupling of second nucleobase on side chain.[103] An alternative method is to synthesize *bm*-PNA monomers in which both nucleobases are already present (Figure 1.39) and then use it for sequential coupling. This needs synthesis of several possible nucleobase combinations in the block monomer. This chapter reports the alternative synthesis of *bi-modal* PNA monomers from block PNA monomers.

Based on the results of previous chapter where Cα / Cβ *gemdimethyl* disubstituted PNA monomers showed fluorescence, it would be interesting to study if Cα-monosubstituted monomers would be fluorescent and if so the nature of bimodal oligomers assembled from such monomers. Fluorescence of pyrimidine-based block bimodal PNA monomer were explored in this chapter to observe the effect of incorporating two different pyrimidine bases in the same

aeg-PNA based monomer (Figure 1.39). Hence the objective of this chapter is to study the nature of the *bi-modal* PNA rotamers in and their effect on the fluorescence emission of designed block bimodal PNA monomers. The study reveals interesting excitation dependent fluorescence emission of bi-modal PNA monomer and the homo oligomer in buffer.

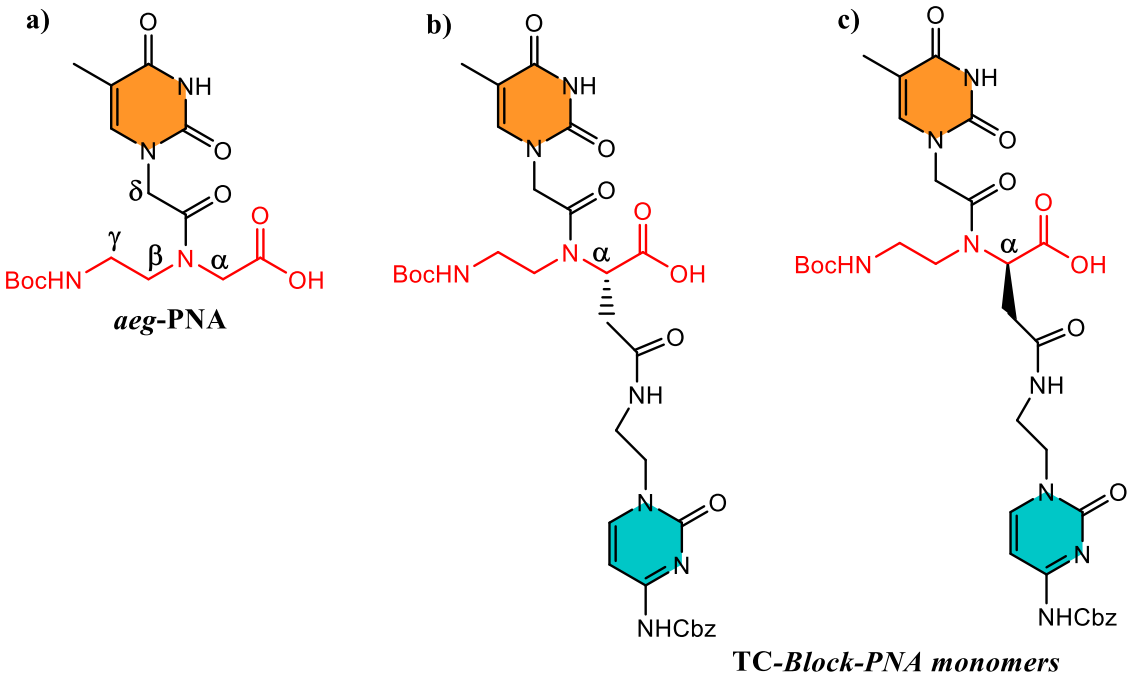


Figure 1.39. a) *aeg*-PNA monomer and (b) and (c) target block bimodal PNA monomers

Chapter 5: Cyclisation involving the glycyl methylene and C-terminal amide group of the PNA backbone to afford a succinimide ring in the C-terminal of the PNA backbone remains unexplored in PNA literature due to complexity of the synthesis. Formation of ring structure in the C-terminus could influence the rotamer distribution of the derived *aeg*-PNA and may resulst interesting structural features of the derived PNA monomer. Many succinimido derivatives like *tecovirimat*, *lurasidone*, *perospirone* etc are FDA approved drugs consisting of succinimide pharmacophore.[104] Diverse nature of substitution around succinimide pharmacophore gives new opportunities for succinimide based drug discovery as synthesis of succinimide PNA derivatives could open path for PNA based small molecule drug design. Compared to the conventional succinimide based drug molecule, PNA based drugs will have extra advantage as the nucleobase could potentially and selectively interact with the other nucleobases of DNA via hydrogen bonding interaction, which the conventional drugs lacks.

Hence the objective of this chapter is to synthesize Cα-Succinimide modified PNA monomer and study the rotmaer preferences of the monomer along with the crystal structure of

the PNA monomer which may potentially exhibit interesting structural features. Indeed, the study reveals Z-rotamer stabilization similar as *aeg*-PNA but folded conformation of the monomer in solution. The crystal structure reveals intesrsting turn type folded strucure of the succinimide modified PNA monomer and a Z-rotamer stabilization in the crystal structure.

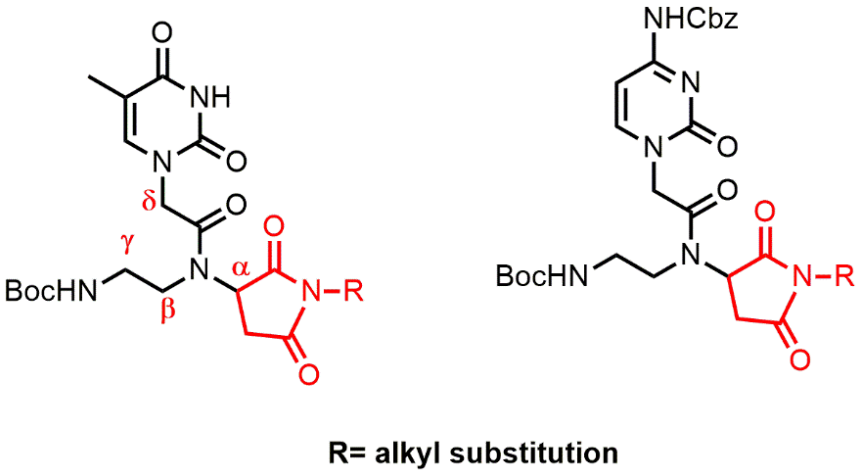


Figure 1.40. Structure of rationally designed succinimide analogues of PNA monomers

1.8 References

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

2. Hoogsteen, K. *Acta. Cryst*.,**1963**, *65*, 907.

3. Crick, F. H. C. *J. Mol. Biol*., **1966**, 19,548-555.

4. Seeman, N. C.; Rosenberg, J. M.; Rich, A. *Proc. Natl. Acad. Sci. USA*., **1976**, *73*, 804-807.

5. (a) Brahms, J.; Mommaerts, W. F. H. M. *J. Mol. Biol.*,**1964**, *10*, 73-85. (b) Fuller, W.; Wilkins, M. H. F.; Wilson, H. R.; Hamilton, L.; D. *J. Mol. Biol*., **1965,** *12*, 60-80. (c) Dickerson, R. E: DNA structure from A-Z. *Methods in Enzymology*., **1992,** *211*, 67-111. (d) Saenger, W: (Ed) *Principles of Nucleic Acids structure*. Springer-Verlag, New York, **1984**.

6. a) Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Van der, M. G.; Van Boom, J. H.; Rich, A. *Science*., **1981**, *211*, 171-176. b) Kamenetskii, M.D.F.; Mirkin, S.M. *Annu. Rev. Biochem.*, **1995**, *64*, 65-95. c) Potaman, V. N.; Soyfer, V.N. *J. Biomol. Struc. Dyn*., **1994**, *11*, 1035-40. (DOI: 10.1080/07391102.1994.10508050) d) . Hänsel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. *Nature Reviews Molecular Cell Biology*., **2017**, *18*, 279-284. e) Gehring, K.; Leroy, J.-L.; Guéron, M. *Nature*.,**1993**, *363*, 561-565.

7. a) Seeman N.C and Sleiman H.F. *Nature Reviews Materials*., **2018**, *3*, 17068. (DOI: https://doi.org/10.1038/natrevmats.2017.68) b) Madsen and Gothelf K.V. *Chem. Rev*., **2019**, *119*, 6384−6458. c) Seeman, N. C. *Nature*., **2003**, *421*, 427-431. (d) Seeman, N. C. *Annu Rev Biochem*., **2010**, *79*, 65-87. e) Petrillo M. L, Newton C. J.; Cunningham R. P.; Ma RI, Kallenbach N. R.; Seeman N. C. *Biopolymers*., **1998**, *27*, 1337-1352. f) Bujold, K.E, Lacroix A, Sleiman, H.F , *Chem 4*., **2018**, 495–521. g) Walsh, A.S.; Yin, Haifang.; Erben, C.M.; Wood, M.J.A.; and Turberfield, *ACS Nano*., **2011**, *5*, 5427-5432.

8. a) Weile, J.; and C, Knabbe. *Anal Bioanal Chem*., **2009**, *394*, 731-742 (b) Muldrew, K L. *Curr OpinPediatr*., **2009**, *21*, 102-111. b) Green, C.; Huggett, J.F.; Talbot, E.; Mwaba, P.; Reither, K.; Zumla, A., *Lancet Infect Dis.*, **2009**, **9**, 505–11. c) oehlenschlager, F.; schwille, P.; and eigen, M. *Proc. Natl. Acad. Sci. USA*., **1996**, *93*, 12811-12816. d) Whiley, D.M.; Tapsall, J.W.; and Sloots, T.P., *Journal of Molecular Diagnostics.*, **2006**, *8*, 3-15 (DOI: 10.2353/jmoldx.2006.050045). e) Katzmeier, F.; Aufinger, L.; Dupin, A.; Quintero, J.; Lenz, M.; Bauer, L.; Klumpe, S.; Sherpa, D.; Du¨rr, B.; Honemann, M.; Styazhkin. I.;

Simmel, F.C.; Heymann, M. (DOI: https://doi.org/10.1371/journal.pone.0220091). f) Mahler, M.; Bentow, C.; Serra, J. & Fritzler, M.J., *Immunopharmacol Immunotoxicol.*, **2016**, *38*, 14–20.

9. Chan, P. P.; Glazer, P. M. *J. Mol. Med*., **1997**, *75*, 267-282.

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

11. Belikova, A. M.; Zarytova, V. F.; Grineva, N. *Tetrahedron Lett.*, **1967**, *37*, 3557-3562.

12. Miller, P. S.; Braiterman, L. T.; Ts'o, P. O. *Biochemistry*., **1977**, *13*, 4897.

13. a) Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A*., **1978,** *75*, 280- 284. b) Stephenson, M. L.; Zamecnik, P. C. *Proc. Natl. Acad. Sci. U.S.A*., **1978**, *75*, 285-289.

14. (a) Akhtar, S.; Agrawal, S. *Trends in Pharmacol Sci*., **1997**, *18*, 12-18. (b) Bennet, C. F. *Biochem Pharmacol*., **1998**, *55*, 9-19. (c) Wickstrom, E: Clinical Trials of Genetic Therapy with Antisense DNA and DNA vectors. Eds. Marcel Dekker; New York, **1998**.

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

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

17. a) Stein, C. A.; Cohen, J. S. Phoshorothioate oligodeoxynucleotide analogues. In Cohen J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 97. b) Millar, P. S. Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): Oligodeoxynucleotides-antisense inhibitors of gene expression. London: Macmillan Press, **1989**, p. 79. c) Froehler, B.; Ng, P.; Matteucci, M. *Nucleic Acids Res*., **1988**, *16*, 4831-4839. d). Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G. *Nucleic Acids Res.*, **1986**, *14*, 7421-7437. e) Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol*., **2010**, *50*, 259-293. f) Sood, S.; Shaw, B. R.; Spielvogel, B. *J. Am. Chem. Soc*., **1990**, *112*, 9000. g) Sergueev, D. S.; Shaw, B. R. *J. Am. Chem. Soc*., **1998**, *120*, 9417-9427.

18. (a) Bennett, C.F and Swayze, E.E, *Annu. Rev. Pharmacol. Toxicol*., **2010**, *50*, 259– 293 (b) Rahman, S.M, Baba, T, Kodama, T, Islam, M.A and Obika, S, *Bioorg. Med. Chem*., **2012**, *20*, 4098–4102; (c) DeClerq, E, Eckstein, F and Merigan, T.C, *Science.*, **1969**, *165*, 1137–1139. d) 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. (e) Guvakova, M. A.; Yakubov, L. A.; Vlodavsky, I.; Tonkinson, J. L.; Stein, C. A. *J. Biol. Chem*., **1995**, *270*, 2620-2627. (iii) 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.

19. a) Altona, C.; Sundaralingam, M.; *J. Am. Chem. Soc*., **1972**, *94*, 8205-8212. b) Guschlbauer, W.; and Jankowski, K. , *Nucleic Acids Res.,* **1980**, *8*, 1421–1433.

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

21. a) Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol*., **2010**, *50*, 259-293. b) Scho¨ning, K.U.; Scholz, P.; Guntha, S.; Wu X.; Krishnamurthy, R.; Eschenmoser, A.;*Science*., **2000**, *290*, 1347-1351. c) Pitsch, S.; Wendeborn, S.; Eschenmoser, *A. Helv. Chim. Acta*., **1993**, *76*, 2161-2183.

22. (a) Obka, S.; Nanbu, D.; Hari, Y.; Andoh, J. –I.; Mori, K. –I.; Doi, T.; Imanishi, T. *Tetrahedron Lett*., **1998**, *39*, 5401-5404. (b) Petersen, M.; Wengel, J. *Trends Biotechnol*., **2003**, *21*, 74-81.

23. a) Seth, P. P.; Siwkowski, A.; Allerson, C. R.; Vasquez, G.; Lee, S. *et al. J. Med. Chem.*, **2009**, *52*, 10-13. b) Koizumi, M. *Curr. Opin. Mol. Ther*., **2006**, *8*, 144-149. c) Kurreck, J.; Wyszko, E.; Gillen, C.; Erdmann, V.A., *Nucleic Acid Research.*, **2002**, *30*, 1911-1918. d) GruÈnweller, A.; Wyszko, E., Bieber, B.; Jahnel, R.; Erdman, V.A.; and Kurreck, J., *Nucleic Acids Research.*, **2003**, *31*, 3185-3193.

24. Schnell, F.J.; Crumley, S.L.; Mourich, D.V.; and Iversen, P.L., BioRes. *Open Access*, **2013**, *2*, 61.

25. (a) Sazani, P.; Gemignani, F.; Kang, S. H.; Maier, M. A.; Manoharan, M. *et al*. *Nature Biotechnol*., **2002**, *20*, 1228-1233. (ii) Alter, J.; Lou, F.; Rabinowitz, A.; Yin, H.; Rosenfeld, J. *et al. Nat. Med*., **2006**, *12*, 175-177.

26. a) 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. b) Verbeure, B.; Lescrinier, E.; Wang, J.; Herdewijn, P. *Nucleic Acids Res*., **2001**, *29*, 4941-4947. c) Aboul-Fadl, T., *Curr. Med. Chem*., **2005**, *12*, 2193–2214. 58. d) Seitz,O.,*Chem., Int. Ed*., **1999**, *38*, 3466–3469. e) Lin, K.Y.; and Matteucci, M.D., *J. Am.*

Chem. Soc., **1998**, *120*, 8531–8532. f) Lin, K.Y.; Jones, R.J.; and M. Matteucci, *J. Am. Chem. Soc*., **1995**, *117*, 3873–3874. 59. g) Herdewijn, P., *Antisense Nucleic Acid Drug Dev.*, **2000**, *10*, 297–310. h) Peacock, H.; Kannan, A.; Beal P.A.; and Burrows, C.J., *J. Org. Chem*., **2011**, *76*, 7295– 7300. i) Verma, S.; and Eckstein, F., *Annu. Rev. Biochem*., **1998**, *67*, 99–134.

27. a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science.*, **1991**, *254*, 1497-1500. b) Qrum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic acid Res*. **1993,** *21*, 5332- 5336. c) Thiede, C.; Bayerdoerffer, E.; Blasczyk, R.; Wittig, B.; Neubauer, A. *Nucleic Acid Res*. **1996,** *24*, 983-984. d) Demers, D. B.; Curry, E. T.; Egholm, M.; Sozer, A. C. *Nucleic Acid Res*.**1995,** *23*, 3050-3055. e) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Nucleic Acid Res*. **1993,** *21*, 197-200. f) Good, L.; Nielsen, P. E. *Antisense Nucleic Acid Drug Dev*. **1997**, *7*, 431. g) De Mesmaeker, A. et al: *Curr. Open. Struct. Biol.* **1995**, *5*, 343. h) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W: PNA. *Angew. Chem. Int. Ed*. **1998**, *37*, 2796-2823.

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

29. a) Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science*., **1994**, *265*, 777-780. b) Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol*., **1997**, *4*, 98-101. c) Leijon, M.; Graeslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, S. M.; Eriksson, M. *Biochemistry*., **1994,** *22*, 9820-9825. (d) Eriksson, M.; Nielsen, P. E. *Nat. Struct. Biol*., **1996**, *3*, 410-413. f) Bets, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science*., **1995**, *270*, 1838-1841.

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

31. a) Felsenfeld, G.; Davies, R. D.; Rich, A. *J. Am. Chem. Soc*., **1957**, *79*, 2023-2024. (b) Le Deon, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucleic Acid Res*., **1987**, *15*, 7749-7760. (c) Helene, C.; Toulene, J. J: *Oligonucleotides: Antisense Inhibitors of Gene Expression*, McMillan Press, London, **1989**. (d) Moser, H. E.; Dervan, P. B. *Science.*, **1987**, *238*, 645-650. (e) Felsenfeld, G.;

Rich, A. *Biochem. Biophys. Acta*., **1957**, *26*, 457- 468. f) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*., **1991**, *254*, 1497-1501. (g) Nielsen, P. E.; Egholm, M.; Buchardt, O. J. *Mol. Recogn*., **1994**, *7*, 165-170.

32. (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.

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

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

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

36. a) Nielsen, P. E, *Pure Appl. Chem*., **1998**, *70*, 105-110. b) Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem*., **1996**, *4*, 5-23.

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

38. a) Corradini, R.; Sforza, S.; Tedeschi, T.; Totsingan, F.; Manicardi, A.; Marchelli, R. *Curr. Top. Med. Chem*., **2011**, *11*, 1535-1554. b). Dueholm, K. L.; Petersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.*, **1994,** *4*, 1077-1080.

39. a) Dueholm, K. L.; Petersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett*., **1994**, *4*, 1077-1080. b) Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. Eur. *J. Org. Chem*., **2000**, *2000*, 2905-2913. c) 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.

40. Koppelhus, U.; Awasthi, S. K.; Zachar, V.; Holst, H. U.; Ebbesen, P.; Nielsen, P. E. *Antisense Nucleic Acid Drug Dev*., **2002**, *12*, 51-63.

41. 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.

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

43. Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett*., **1998**, *39*, 4707-4710.

44. Aguado, G. P.; Rua, F.; Branchadell, V.; Nielsen, P. E.; Ortuno, R. M. *Tetrahedron Asymmetry*., **2006**, *17*, 2499-2503.

45. a) Gourishankar, A.; Ganesh, K. N. *Artificial DNA PNA XNA*., **2012**, *3*, 5-13. b) Kulkarni, P.; Datta, D.; Ramabhadran, R.O.; and Ganesh, K.N., *Org. Biomol. Chem*., **2021**, *19*, 6534.

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

47. Kulkarni, P.; Datta, D.; Ramabhadran, R.O.; and Ganesh, K.N., *Org. Biomol. Chem*., **2021**, *19*, 6534.

48.a) Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett*., **1994**, *35*, 5173-5176. b) Tedeschi, T.; Sforza, S.; Corradini, R.; Marchelli, R. *Tetrahedron Lett*., **2005**, *46*, 8395-8399. c) Englund, E. A.; Appella, D. H. *Org. Lett*., **2005**, *7*, 3465-3467. d) Dose, C.; Seitz, O. *Org. Lett*., **2005**, *7*, 4365-4368. e) Ficht, S.; Dose, C.; Seitz, O. *ChemBioChem*., **2005,** *6*, 2098-2103.

49. Englund, E. A.; Wang, D.; Fujigaki, H.; Sakai, H.; Micklitsch, C. M.; Ghirlando, R.; Martin-Manso, G. *Nat. Commun*., **2012**, *3*, 614.

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

51. Avitabile, C.; Moggio, L.; Malgieri, G.; Capasso, D.; Gaetano, S. D.; Saviano, M.; Pedone, C.; Romanelli, A. *PLoS One*., **2012**, *7*, e35774.

52. a) Dragulescu-Andrasi, A.; Rapireddy, S.; Frezza, B. M.; Gayathri, C.; Gil, R. R.; Ly, D. H. *J. Am. Chem. Soc*., **2006**, *128*, 10258-10267. b) 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.

53. Crawford, M. J.; Rapireddy, S.; Bahal, R.; Sacui, I.; Ly, D. H. *J. Nucleic Acids* **2011**, 2011, (doi:10.4061/2011/652702).

54. Manicardi, A.; Fabbri, E.; Tedeschi, T.; Sforza, S.; Nicoletta Bianchi, N.; Brognara, E.; Gambari, R.; Marchelli, R.; Corradini, R. *Chem. Bio. Chem*., **2012**, *13*, 1327-1337.

55. a) Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, B. A.; Ly, D. H. *J. Org. Chem*., **2011**, *76*, 5614-5627 b) Canady, T. D.; Telmer, C. A.; Oyaghire, S. N.; Armitage, B. A.; Bruchez, M. P. *J. Am. Chem. Soc*., **2015**, *137*, 10268-10275.

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

57. a) Canady, T. D.; Telmer, C. A.; Oyaghire, S. N.; Armitage, B. A.; Bruchez, M. P. *J. Am. Chem. Soc*. **2015**, *137*, 10268-10275. b) Shiraishi, T.; Pankratova, S.; Nielsen, P. E. *Chem. Biol*., **2005**, *12*, 923-929.

58. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc*., **1996**, *118*, 5544-5549.

59. a) Hollenstein, M.; Leumann, C. J. *Org. Lett*., **2003**, *5*, 1987-1990. b) Rose, D. *J. Anal. Chem*., **1993**, *65*, 3545-3549.

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

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

62. a) Gangamani, B. P.; D' Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides*., **1999**, *18*, 1409-1011. b) Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron*., **1999**, *55*, 177-192.

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

64. a) Vilaivan, T.; Khongdeesameor, C.; Harnyuttanokam, P.; Westwell, M. S.; Lowe, G. *Bioorg. Med. Chem. Lett*., **2000**, *10*, 2541-2545. c) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **2001**, *3*, 1281-1284.

65. Sharma, N.; Ganesh, K. N. *Tetrahedron Lett*., **2004**, *45*, 1403-1406.

66. Sharma, N.; Ganesh, K. N. *Chem. Commun*., **2003**, *0*, 2484-2485.

67. Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc*., **2006**, *128*, 16456 – 16457.

68. Puschl, A.; Boesen, T.; Zuccarello, G.; Dahl, O.; Pitsch, S.; Nielsen, P. E. *J. Org. Chem*., **2001**, *66*, 707-712.

69. Hickman, D. T.; King, P. M.; Cooper, M. A.; Slater, J. M.; Mickelfield, *J. Chem. Commun*., **2000**, 2251-2252.

70. Kumar, V. A.; Pallan, P. S.; Meena, Ganesh, K. N. *Org. Lett*., **2001**, *3*, 1269-1272.

71. Puschl, A.; Boesen, T.; Zuccarello, G.; Dahl, O.; Pitsch, S.; Nielsen, P. E. *J. Org. Chem*., **2001**, *66*, 707-712.

72. D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron. Lett*., **2002**, *43*, 883-886.

73. Lonkar, P.; Ganesh, K. N.; Kumar, V. A. *Org. Biomol. Chem*., **2004**, *2*, 2604-2611.

74. Lonkar, P. S.; Kumar, V. A.; Ganesh, K. N. *Nucleos Nucleot Nucleic Acids*. **2001**, *20*, 1197-1200.

75. Lescrinier, E.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res.*, **2003**, *31*, 2975-2989.

76. a) Govindaraju, T.; Gonnade, G.R.; Bhadbhade, M.M.; Kumar, V.A.; and Ganesh, K.N. *Org. Lett*., **2003**, *5*, 3013-3016. b)Govindaraju, T.; Kumar, V.A.; and Ganesh. K.N. *J. Org. Chem*., **2004**, *69*, 1858-1865. c) Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Am. Chem. Soc.* **2005**, *127*, 4144–4145.

77. a) Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Org. Chem.* **2004**, *69*, 5725–5734. b) Govindaraju, T.; Kumar, V. A. *Chem. Commun.*, **2004**, 860–861.

78. Huang, Y.; Dey, S.; Zhang, X.; Sonnichsen, F.; Garner, P. *J. Am. Chem. Soc*., **2004**, *126*, 4626-4640.

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

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

81. a) Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res*.,**1996**, 24, 494-500. b) Wang, L.; Li1, H.; Zhao, C.; Li1, S.; Kong, L.; Wu, W.; Kong, W.; Liu, Y.; Wei, Y.; Zhu, J.K.; and Zhang, H. *Plant Cell Environ.*, **2017** , *40*, 56–68.

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

83. 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.**

84.(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. f) 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.

85. Orain, D.; Ellard, J.; Bradley, M. *J. Comb. Chem*., **2002**, *4*, 1-16.

86. Brieke, C.; Cryle, M. J. *Org. Lett*., **2014**, *16*, 2454-2457.

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

88. (a) Barawkar D. A.; Bruice T. C. *Proc. Natl. Acad. Sci. U.S.A*., **1998**, *95*, 11047-11052. (b) Yadav D. K.; Lokhande R. S.; Pitale S. M.; Janwadkar S. P.; Navarkar P. S.; Rana P. K. *J. Anal. Chem*., **2014**, *2*, 10-14.

89.(a) Huang, C. Y. *Methods in Enzymology*., (Ed.: D. L. Purich), Academic Press, **1982**, *87*, 509-525. (b) Nielsen, P. E.; Christensen, L. *J. Am. Chem. Soc*., **1996**, *118*, 2287-2288.

90. Egholm, M.; Buchardt, 0.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc*., **1992**, *114*, 1895-1897.

91. 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.

92. Baker, E. S.; Bowers, M. T. *J. Am. Soc. Mass Spectrom*., **2007**, *18*, 1188-1195.

93. a) Kushon, S. A.; Jordan, J. P.; Seifert, J. L.; Nielsen, H.; Nielsen, P. E.; Armitage, B. A. *J. Am. Chem. Soc*., **2001**, *123*, 10805-10813. b) Andrasi, A.D.; Rapireddy, S.; Frezza, B.M.; Gayathri, C.; Gil, R.R.; and Ly, D.H. *J. Am. Chem. Soc*., **2006**, *128*, 10258-10267.

94. Gourishankar, A.; Shukla, S.; Ganesh, K. N.; Sastry, M. *J. Am. Chem. Soc*., **2004**, *126*, 13186-13187.

95. Pierce, M. M.; Raman, C. S.; Nall, B. T. **Methods**.,**1999**, *19*, 213.

96. (a) Barbieri, C. M.; Li, T. K.; Guo, S.; Wang, G.; Shallop, A. J.; Pan, W.; Yang, G.; Gaffney, B. L.; Jones, R. A.; Pilch, D. S. *J. Am. Chem. Soc*., **2003**, *125*, 6469. (b) Kunne, A.; Sieber, M.; Meierhans, D.; Allemann, R. K. *Biochemistry*., **1998,** *37*, 4217. (c) Egholm, M.; Buchardt, 0.; Christensen, L.; Behrens, C.; Frier, S. M.; Driver, D. A.;Berg, R. H.; Kim, S. K.; Norden, B.; Neilsen, P. E. *Nature*, **1993**, *365*, 566-56.

97. Bartold, K.; Pietrzyk-Le, A.; Golebiewska, K.; Lisowski, W.; Cauteruccio, S.; Licandro, E.; D Souza, F.; Kutner, W. *ACS Appl. Mater. Interfaces*., **2018**, *10*, 27562-27569.

98. a) Brown, S.C.; Thomson, S.A.; Veal, J.M.; Davis, D.G., *Science* **1994**, *265*, 777 - 780. b) Oleszczuk, M.; Motowidlo, S.R.; and Falkiewicz, B., *nucleosides, nucleotides & nucleic acids*, **2001**, *20*, 1399–1402. c) Meltzer, P.C.; Liang, A.Y.; and Matsudaira. P., *J. Org. Chem*. **1995**, *60*, 4305-4308.

99. Kulkarni, P.; Datta, D.; Ramabhadran, R.O.; and Ganesh, K.N., *Org. Biomol. Chem*., **2021**, *19*, 6534.

100. Zhang, Y.; Malamakal, R.M.; and Chenoweth, D.M., *Angew.Chem.Int.Ed.*, **2015**, *54*, 10826–10832. b) McMechen, M.A.; Willis, E.L.; Gourville, P.C.; and Caroline Proulx, *Molecules* **2019**, *24*, 1919.

101. Vass, E.R.; Hollo´si, M.; Besson, F.; and Buchet, R., *Chem. Rev*. **2003**, *103*, 1917−1954.

102. Berger, O.; Abramovich, L.A.; Levy-Sakin, M.; Grunwald, A.; Liebes-Peer, Y.; Bachar, M.; Buzhansky, L.; Mossou, E.; Forsyth, V.T.; Schwartz, T.; Ebenstein, Y.; Frolow,F.; Shimon, L.J.W.; Patolsky, F.; and Gazit, E., *Nature Nanotechnology*, **2015**, *10*, 353–360.

103. a) Gupta, M.K.; Madhanagopal, B.R.; Datta, D.; and Ganesh, K.N. *Org. Lett*., **2020**, *22*, 5255−5260. b) Gupta, M.J.; Madhanagopal, B.R.; and Ganesh, K.N. *J. Org. Chem*., **2021**, *86*, 414−428. c) Bhingardeve, P.; Madhanagopal, B.R.; and Ganesh, K.N. *J. Org. Chem.*, **2020**, *85*, 13680−13693.d) Bhingardeve, P.; Jain, P.; and Ganesh, K.N. *ACS Omega*., **2021**, *6*, 19757−19770. e) Todkari, I.; Gupta, M.J.; and Ganesh, K.N., *Chem. Commun*., **2022**, *58*, 4083–4086.

104. a) Zhao, Z.; Yue , J.; Ji , X.; Nian , M.; Kang, K.; Qiao , H.; Zheng, X., *Bioorganic Chemistry*., **2021**, *108*, 104557. b) Needham, J.; Kelly, M.T.; Ishige, M.J.; Andersen, R.J.; Andrimid and Moiramides. *J. Org. Chem*., **1994**, *59*, 2058–2063. c) Pohlmann, J.; Lampe, T.; Shimada, M.; Nell, P.G.; Pernerstorfer, J.; Svenstrup, N.; Brunner, N.A.;

Schiffer, G.; Freiberg,C., *Bioorg. Med. Chem. Lett*., **2005**, *15*, 1189–1192. d) Silvers, M.A.; Robertson, G.T.; Taylor, C.M.; Waldrop. G.M., *J. Med. Chem*., **2014**, 57, 89478959. e) Yang, A.; Si, L.; Shi, Z.; Tian, L.; Liu, D.; Zhou, D.; Proksch, P.; Lin, W.., *Org. Lett*., **2013**, 15, 5366–5369. f) Philkhana, G.R.; Jachak, V.B.; Gunjal, N.M.; Dhage, A.H.; Bansode, D.S. Reddy., *Tetrahedron Lett*., **2015**, *56* ,1252–1254.g) Zhao, Y.; Nie, M.; Xu, P.; Li, H.; Xu, L.; Cheng, X.Q.; Lei., Y.Q., *Biochem. Biophys. Res. Commun*., **2019**, *515* ,474–480. d) Philkhana, S.C.; Verma, A.K.; Jachak, G.R.; Hazra, B.; Basu, A.; Reddy, D.S, *Eur. J. Med. Chem*., **2017**, 135 ,89–109.

Chapter 2

E-Rotamer Selectivity in N_{α}-Aza-Peptide Nucleic Acid Monomers via 8-membered γ-Type Turn Formation

Chapter 2

E-Rotamer Selectivity in N_{α}-Aza-Peptide Nucleic Acid Monomers via 8-membered γ-Type Turn Formation

2.1. Introduction

Much after the discovery of duplex structure of DNA[1] by Watson and Crick in 1953, several chemical modifications of DNA molecule were reported[2] in the literature for the application of antisense therapeutics.[3] Among several chemical modifications reported in the field of modified nucleic acids and oligonucleotide analogues, Peptide Nucleic Acid (PNA)[4] stands out as one of the simplest synthetic mimics of DNA. In PNA, the sugar-phosphate backbone of DNA is replaced by achiral and neutral N-(2-aminoethyl) glycine (*aeg*) unit to which the nucleobase is connected through a methylene carbonyl linker[4] (Figure 2.1). The nucleobase arrangement in PNA efficiently mimics the specific base pairing properties as seen in DNA.

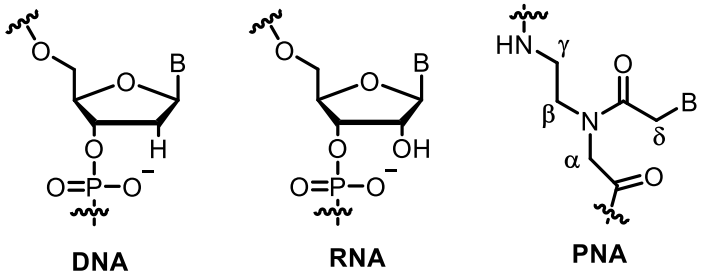


Figure 2.1. Basic chemical structure of DNA, RNA and PNA

PNA binds to complementary DNA/RNA via Watson-Crick base pairing to yield PNA:DNA duplexes. Like DNA triplexes, PNA can form PNA_2/DNA triplexes, in which the DNA forms Watson-Crick base pairing with one of the PNA strands and gets involved in Hoogsteen hydrogen bonding[5] with the second PNA strand. PNA:DNA/RNA complexes have higher thermal stability compared to the native DNA:cDNA/cRNA complexes.[5] Although originally developed as a complementary ligand, the extraordinary binding and the attractive properties of PNA like stability towards proteases and nucleases and low toxicity makes PNA a potential tool in molecular biology and diagnostics,[6] with great interest for the medicinal chemists.[7] In analogy with DNA, numerous chemical modifications of PNA have been performed at the α, β and γ positions[8] (Figure 2.2) to understand the structure-activity relationship (SAR).[9] Inspite

of having extraordinary biophysical properties, PNA suffers from drawbacks like poor aqueous solubility, insufficient cell permeability and orientational (parallel/antiparallel) ambiguity in binding DNA/RNA. Attempts towards addressing these shortcomings involved numerous backbone modifications like introduction of ionizable groups (amine, hydroxy, guanidium cation, Figure 2.2) at the α, γ and β sites on PNA backbone.[8,10] Modifications that involved increasing either the flexibility or rigidity of backbone resulted in reduced binding affinity with complementary DNA or RNA due to the accompanying entropic disadvantages compared to the standard aminoethyl glycine (*aeg*) backbone.[10] A balanced preorganization of PNA backbone to attain hybridization competent conformation is necessary to yield unambiguous PNA:DNA/RNA complexes.[11] PNA being non-chiral, can assume both right handed and left handed helical coil structures and introduction of particular chirality in the PNA backbone may shift equilibrium towards the correct handedness and provide subtle rigidification of the backbone to entropically favour duplex formation.[11] Incorporation of 5/6-membered ring structures through folding of backbone or its cyclization with sidechain also gave chiral structures. Based on the nature of chirality and ring size, some modifications preorganized the PNA backbone, to entropically favour it to attain hybridisation competent conformation.[11] Inclusion of spiro cyclohexyl ring in the glycine part of the *aeg* PNA backbone (Figure 2.3) favours the PNA:DNA duplex formation.[11] Introducing ring structure in the PNA backbone also brought about preferential binding of PNA towards complementary DNA or RNA.[11] Substitution of the α position of PNA backbone (glycine side) with D/L amino acids exhibited interesting properties.[12] Incorporation of L/D alanine in PNA introduced preferences of antiparallel binding over parallel binding, with D-alanine more inclined for anti-parallel binding compared to L-alanine.[12] PNA derived from D-lysine improved the hybridisation property of unmodified *aeg*-PNA,[13] due to the positively charged side chain amino group enhancing the affinity to negatively charged phosphates of DNA. On the other hand, in L/D-aspartic acid and glutamic acid derived PNAs, the negative charges of the sidechain carboxyl group significantly reduced the hybridisation strength due to repulsion with the anionic phosphates on DNA backbone.[14] However, the presence of charged sidechains in PNA increased the solubility of PNA in water. Incorporation of the chiral PNA monomers strengthen the sequence selectivity in hybridisation. The extension of *aeg*-PNA backbone via introduction of a methylene group on glycine/ethylene diamine sides or on the methylene carbonyl linker (Figure 2.4), lead to reduced hybridization efficiency with DNA, compared to the unmodified *aeg*-PNA:cDNA duplexes.[15]

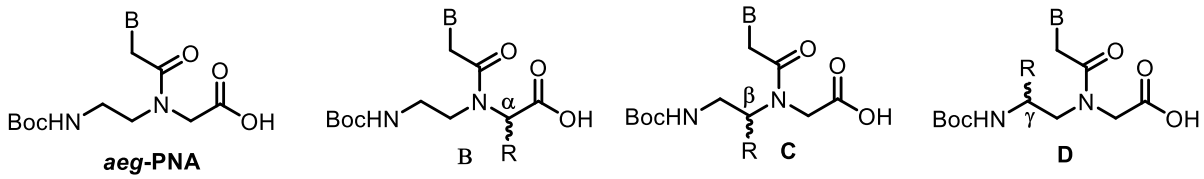


Figure 2.2. Structures of *aeg*-PNA and Chiral PNA modifications (B-D, R= amino acid side chain or other modifications, B= Nucleobase)

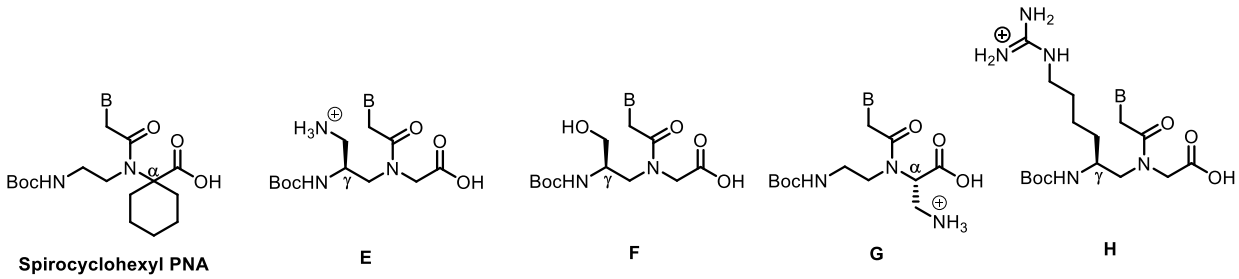


Figure 2.3. Spiro cyclohexyl PNA and modified PNA to address the shortcomings of PNA(E-H)

The central N-tertiary amide bond of *aeg*-PNA is a key structural element for PNA. The partial double bond character of tertiary amide restricts the rotation around N-CO bond, leading to two distinct orientations for the carbonyls of the tertiary amide bond resulting in rotamers[16] When the t-amide bond carbonyl is oriented towards the glycine side, the rotamer is assigned as the Z-rotamer, while in E-rotamer, the tertiary amide bond carbonyl is pointed towards the ethylene diamine side (Figure 2.5). The energy barrier of ~19±2 kcal/mol for rotation around

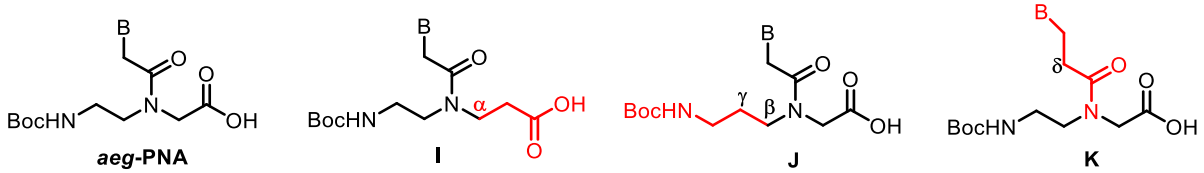


Figure 2.4. Modified *aeg*-PNA extended by one methylene group (I-K)

the N-CO bond allows the simultaneous existence of Z:E rotamer mixture at ambient temperature and can be identified by ^{1}H NMR, where the backbone CH_2 protons (α, β and γ) and the sidechain δ-CH_2 show doubling of signals corresponding to each of the rotamers,[17,16] with the ratio Z:E rotamer to be about 60-70:40-30, in favor of Z in monomers. However, the crystal structures of PNA:cDNA, PNA:cRNA and PNA:cPNA duplexes show exclusively the Z-rotamer,[18] the preferred rotamer form in monomer.

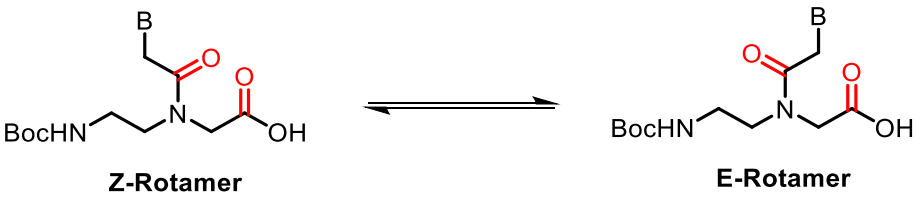


Figure 2.5. E and Z rotamer of *aeg*-PNA

It would be interesting to study PNA duplexes derived from the analogues in which the rotamer is exclusively E in monomers. In a first attempt to validate the same, the central amide bond of PNA was replaced by an isostructural olefinic double bond and the nucleobase, backbone linker was fixed into E and Z isomeric form (Figure 2.6). Thermal stability of the fully modified E-OPA:DNA duplexes decreased significantly and ability of PNA to form triplexes completely diminishes in the fully modified OPA-T sequences (E and Z) with complementary DNA,[19] emphasising the central role of tertiary amide bond in structural features that impart the hybridisation affinity and the preferred strand orientation of PNA in the PNA:DNA complexes.

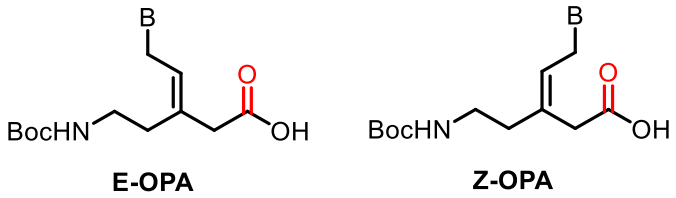


Figure 2.6. Structure of Olefinic E and Z OPA monomers (B= Nucleobase)

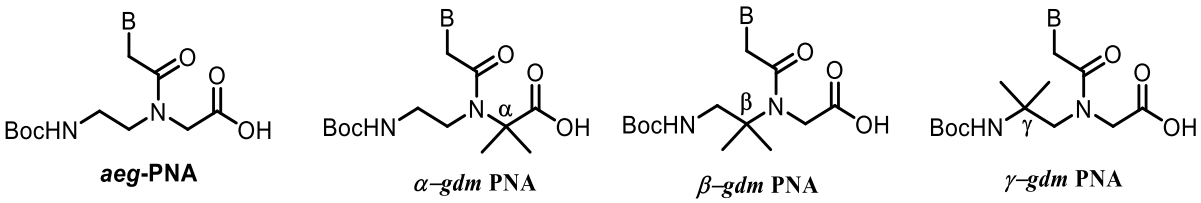


Figure 2.7. Structures of *aeg*-PNA and *gemdimethyl* PNA

Interestingly, the introduction of achiral *gemdimethyl(gdm)* group (Figure 2.7) in the α-position of *aeg*-PNA monomer exclusively stabilized the Z-rotamer, while in β-position only the E-rotamer was detected.[20] The substitution of *gdm* group in the γ-position gave a mixture of rotamers. These results indicated that structural factors with steric clash of *gemdimethyl* group with δ-CH_2 disfavours E-rotamer in α-*gdm* monomer while the possibility of C-H···O

bonding from δ-CO to β-*gemdimethyl* group favours E-rotamer. These results pointed out that the incorporation of *gemdimethyl* group on either of the adjacent carbons of central tertiary amide bond of PNA leads to single rotamer in contrast to unsubstituted *aeg*-PNA that exhibits a rotamer mixture.[20] In summary, parameters like length of the backbone, nature of the linker attached with nucleobase, chirality of the carbon atoms, steric hindrance, spatial orientation of the sidechain within the backbone, nature of functional groups present in the side chain of the backbone etc are important for rational design of new PNA monomers.

2.1.1 PNA rotamers: Crystal structures of Peptide Nucleic Acid (PNA)

In general, *aeg*-PNA and modified β/γ, α-substituted PNA monomers in solution exhibit mixture of rotamers in solution. In contrary, PNA monomers and oligomers in solid state crystalize into a single rotamer. Ganesh *et.al*[21] reported the single crystal structure of cyanuric acid PNA monomer (Figure 2.8) which showed it to be Z-Rotamer, with the tertiary amide

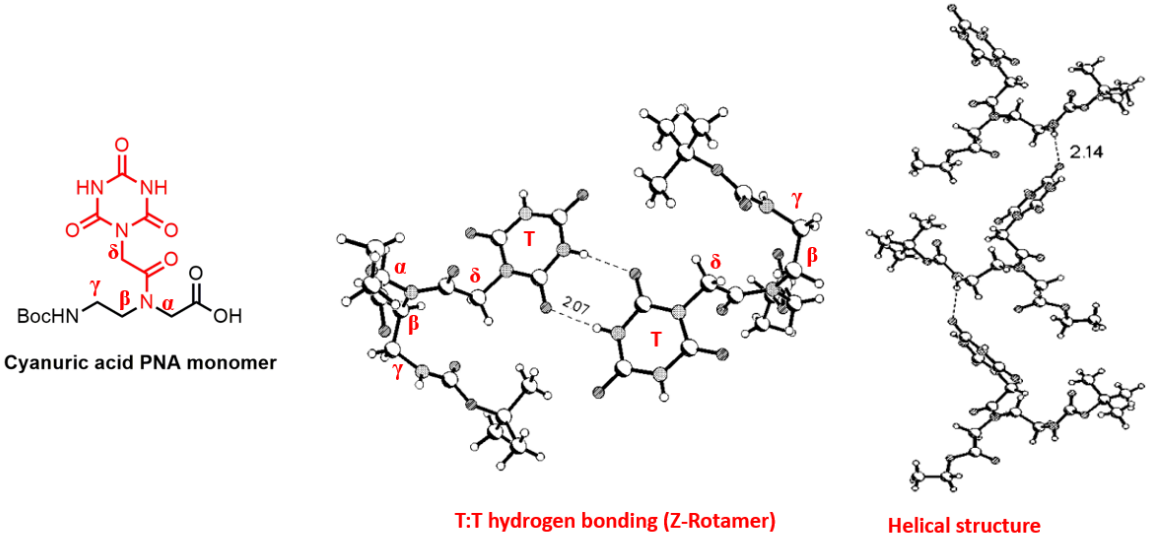


Figure 2.8. Molecular and Crystal structure of Cyanuric acid PNA monomer[21]

carbonyl oriented towards the glycine side of the PNA backbone. In the crystal packing, extensive self-hydrogen bonding between the thymine nucleobases of two PNA monomer units facilitated a helical arrangement of the chains in solid state. The solution NMR of *aeg*-PNA suggested a mixture of rotamers with the major one being the Z form, in solution. The crystal structure of γ-*gdm*-T PNA monomer[20] (Figure 2.9) showed it to be the E-rotamer with δ-CO oriented towards ethylene diamine side. In solution, the ^{1}H NMR of γ-*gdm*-T PNA monomer exhibited a mixture of rotamers, with the major rotamer being the E-Rotamer. However, the

α-*gdm*-T PNA monomer exhibited Z-rotamer in both solution and solid states. Thus, the exclusive existence of E-rotamers in PNAs was established for the first time in β-*gdm*-T PNA monomer and Z-rotamer in α-*gdm*-T PNA monomer.

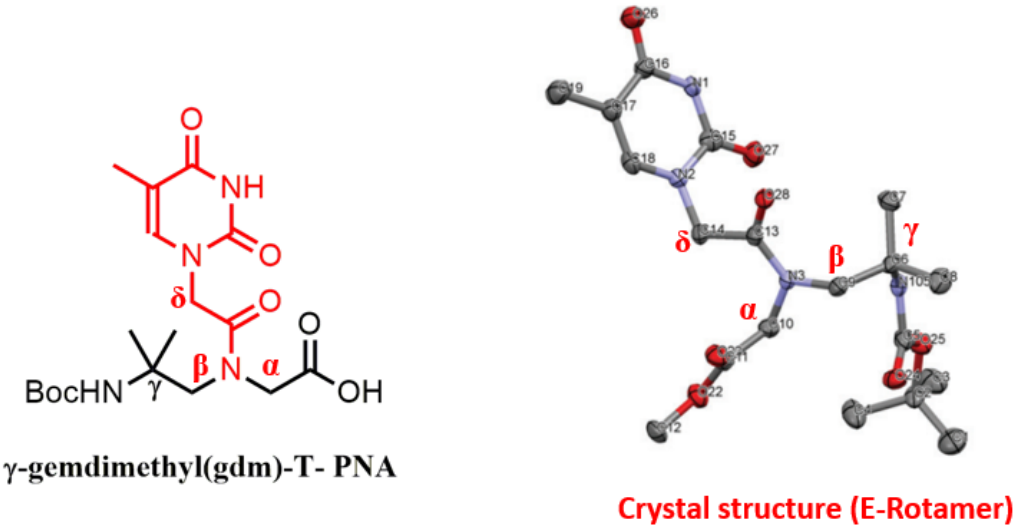


Figure 2.9. Molecular and crystal structure of γ-*gemdimethyl*-T PNA monomer[20]

Howarth *et al.*[22] reported the crystal structure of the backbone modified PNA where the ethylene diamine side is replaced by propargyl group. The solution ^{1}H NMR analysis of the modified monomer revealed a mixture of E and Z rotamer in solution, while the crystal

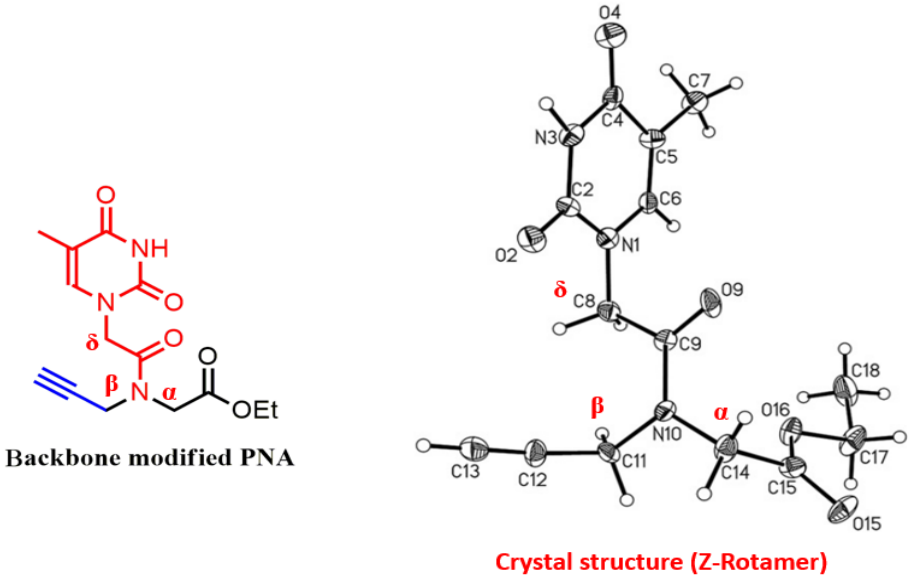


Figure 2.10. Molecular structure and crystal structure of the backbone modified PNA

structure showed Z-rotamer in crystal[22] (Figure 2.10). Analysis of packing of the crystal reveals a centrosymmetric hydrogen bonded dimeric structure of the modified PNA monomer and stacking interaction between the nucleobases. The alkenyl CH (H-6) of the thymine nucleobase faced the tertiary amide carbonyl of the backbone.[22]

Gazit *et al.*[18] reported the crystal structure of GC di-PNA which showed the complementary Watson-Crick intermolecular base pairing between the guanine (G) and cytosine (C) nucleobases to form the PNA:PNA duplex structure in solid state. The π-π stacking interaction between the G and C nucleobases stabilizes the dimeric structure in solid state.[18] The tertiary carbonyl group is oriented towards the glycine side of the PNA backbone as in Z-rotamer in the GC dimer as observed for individual monomers in solution state [18] (Figure 2.11).

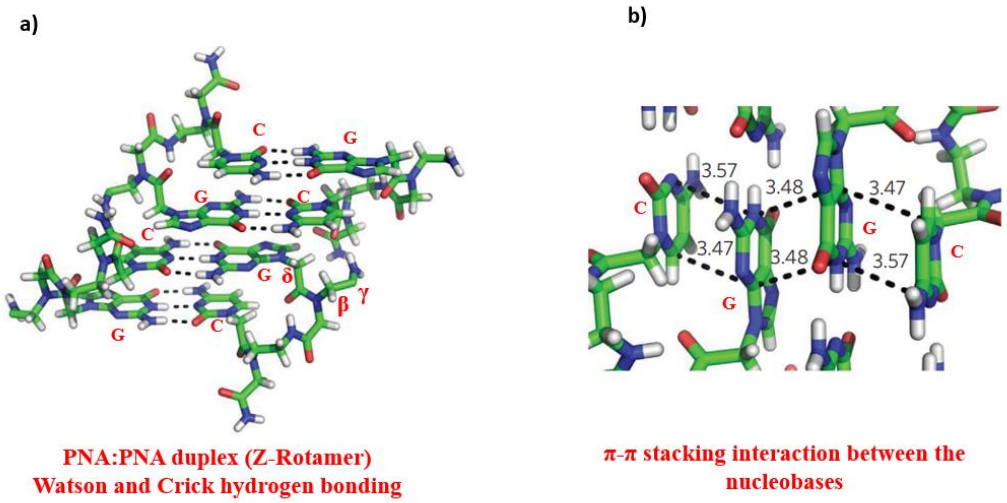


Figure 2.11. a) Watson crick base pairing in GC Dimer and Z-rotamer stabilization b) π-π stacking interaction between the Guanine and Cytosine in GC Dimer

In the crystal structures of *cis*-(1*S*,2*R*)- and *cis*-(1*R*,2*S*)- cyclohexyl PNAs,[23] wherein the ethylene diamine of *aeg*-PNA is replaced by a cyclohexyl group inducing rigidity in the PNA backbone. The tertiary carbonyl group is oriented towards the ethylene diamine segment of the PNA backbone, indicating it to be Z-Rotamer (Figure 2.12).[23] The backbone C-C torsion angle β for (1*S*,2*R*)-cyclohexyl PNA and (1*R*,2*S*)-cyclohexyl PNA are -63º and 66º respectively. Thus Z-rotamer has higher compatibility and adaptability of for both acyclic and cyclic PNA monomers. Comparative structural studies (X-ray and NMR) of PNA:DNA/RNA duplexes indicated specific differences in C-C dihedral angle β, which was ~60º in PNA:RNA duplexes, while ~120 º in PNA:DNA complexes. This means that *cis*-cyclohexyl PNA is pre-organised for preferential binding with RNA than DNA. Indeed the *cis*-cyclohexyl PNA

incorporated oligomers selectively stabilised PNA:RNA duplexes compared to PNA:DNA duplexes.[23a] The trans-Cyclohexyl PNA in which dihedral angle is 120^{o} neither stabilized the PNA: DNA nor PNA: RNA duplexes.[23b]

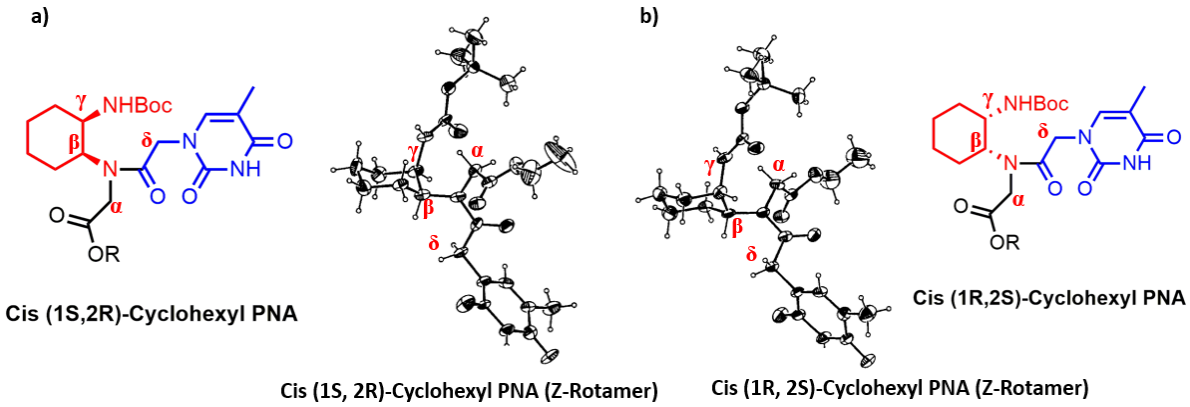


Figure 2.12. a) Molecular and crystal structure of cis-disubstituted (1S,2R)-cyclohexyl PNA b) Molecular and crystal structure of cis-disubstituted (1R,2S)-cyclohexyl PNA

In the crystal structure of *cis*-(1*S*,2*R*)- and *cis*-(1*R*,2*S*)- cyclopentyl PNAs (*cp*-PNA),[23c] wherein the ethylene diamine of *aeg*-PNA is replaced by a cyclopentyl group inducing rigidity in the PNA backbone. The tertiary carbonyl group is oriented towards the ethylene diamine segment of the PNA backbone, indicating it to be Z-Rotamer[23c] in cyclopentyl PNA (Figure 2.13), similar as cyclohexyl PNA.[23a] Thus Z-rotamer has higher compatibility and adaptability of for both acyclic and cyclic PNA monomers compared to E-rotamer as the cis cyclohexyl and cyclopentyl PNAs exhibit Z-rotamer is crystal. *cis*-(1*S*,2*R*) and *cis*-(1*R*,2*S*)-cyclopentyl PNAs forms stronger duplexes (PNA:DNA) and triplexes (PNA_2:DNA) with complementary DNA and RNA without any specific discrimination between DNA and RNA.[23d] The higher stability of the *cp*-PNA : cDNA duplexes compared to acyclic *aeg*-PNA: cDNA could be attributed due to the higher flexibility and ring puckering of the cyclopentyl ring which preorganized the *cp*-PNA : cDNA duplexes compared to *aeg*-PNA: cDNA duplexes.[23d] cyclopentyl PNAs shows very high sequence discrimination ability with mismatch DNA.[23d] Interestingly unlike *trans*-(1*S*,2*S*) and *trans*-(1*R*,2*R*)-cyclohexyl PNA which destabilizes the PNA: DNA duplexes,[23b] the *trans*-(1*S*,2*S*) and *trans*-(1*R*,2*R*)-cyclopentyl PNA exhibit improved PNA:DNA duplex stability and higher binding affinity compared to *aeg*-PNA.[23e] Reenetly Daniel H. apella *et al.*[23f] showed that melting temperature of fully modified heterooligomer of trans cyclohexyl PNA duplexes (*trans cp*-PNA: cDNA) could reach up to 86^{o}C.[23f]

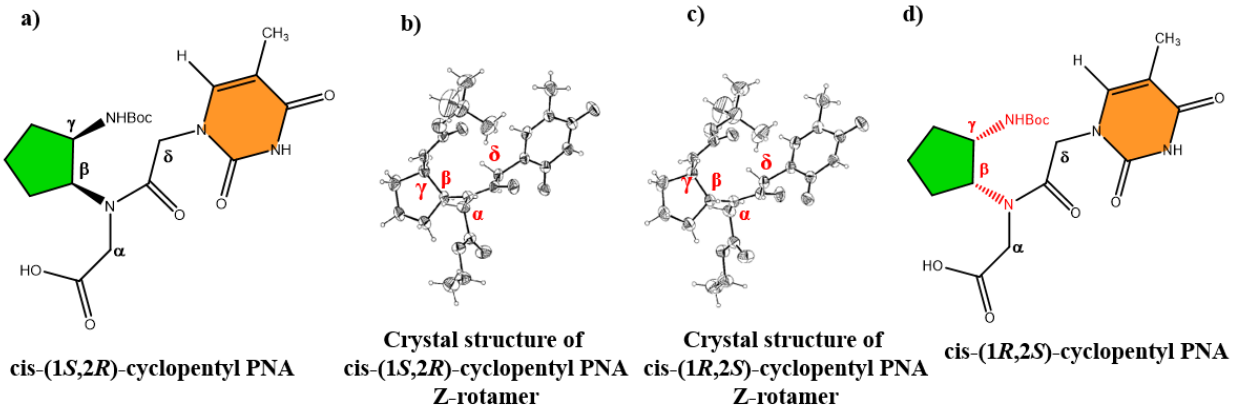


Figure 2.13. a) Molecular structure of cis-disubstituted (1*S*,2*R*)-cyclopentyl PNA b) Crystal structure of cis-disubstituted (1*S*,2*R*)-cyclopentyl PNA c) Crystal structure of cis-disubstituted (1*R*,2*S*)-cyclopentyl PNA c) Molecular structure of cis-disubstituted (1*R*,2*S*)-cyclopentyl PNA.

2.2 Rationale and objective of the present work

Classical *aeg*-PNA consist of C-C, C-N and C-O single bonds that are flexible to adopt hybridisation competent conformation for binding to complementary DNA/RNA. Recent report from Ganesh *et al*.[20] suggested that incorporation of *gemdimethyl* group in the backbone of *aeg*-PNA preferentially stabilized one rotamer over another[20] (Figure 2.7). The rotamer selectivity arises from avoiding the steric clash between δ-CH_2 and α or β *gemdimethyl* group and the theoretical study also suggested favorable CH•••O interaction in the stabilized rotamer.[20] The rationale of the present work is to design PNA analogues where the E or Z-rotamers can be stabilized by intra-residue H-bonding in backbone rather than steric constrains. Such conformationally constrained, but backbone folded *aza*-PNA analogues may have better helical structures in PNA single strands and thereby may show preference for stabilising PNA:DNA/RNA duplexes. The steric factors at α or β-sites can be deleted by incorporating potential H-bonding NH group in the α position of the *aeg*-PNA backbone (Figure 2.14). Cα substitution by NH in *aza* aminoacids alters the hydrogen bonding properties of derived peptides by redefining the local conformational features, with the adaptive chirality of the nitrogen atom affecting the properties of the molecules.[24] This interesting potential of NH was the motivation for this work to engineer *aza*-PNAs, in which the C_α of *aeg*-PNA backbone is replaced by isosteric NH group. α-NH in *aza*-PNA (Figure 2.14) has potential to get involved in intramolecular hydrogen bonding with the (i) N-terminal amide CO group (ii) N-tertiary amide CO group or (iii) C-terminal CO and fold the *aeg* PNA backbone into a β or γ-turn[29] like

conformation as observed in peptides (Figure 2.15). Formation of such turn like motifs in *aza*-PNAs may influence their binding propensity with complementary DNA/RNA by the rotation of central tertiary amide (N-CO) bond of PNA backbone and thus locking it into a single E or Z rotamer. Unlike CH_2 group, NH group carries lone pair of electrons and may rotate around the molecular axis by 180^0 to invert the molecule known as umbrella inversion,[30] thus even imparting partial chirality. α-*aza*-PNA (Figure 2.14b) is analogous to *aza*-peptides and the repeating units generates urea link (Figure 2.16) which is also known to stabilizes the β-sheet structure. However, the planar urea linkage is not conducive to fold the derived oligomer into a helical form and hence the homologue of α-*aza*-PNA - the *aza*-β-*ala*-PNA (Figure 2.14c) which may lead to a regular helical polyamide oligomer (Figure 2.16) was also designed for study. This is the α-*aza* analogue of β-*ala*-PNA monomer (Figure 2.14d). With this rationale the target α-*aza*-PNA-T/C, β-ala-*aza*-PNA-T and β-*ala*-PNA-T monomers (Figure 2.14) were synthesized, and the rotamer structures of the monomers were studied by 1D and 2D NMR techniques and by IR spectroscopy. To validate and further explain the obtained results theoretical calculations were performed.

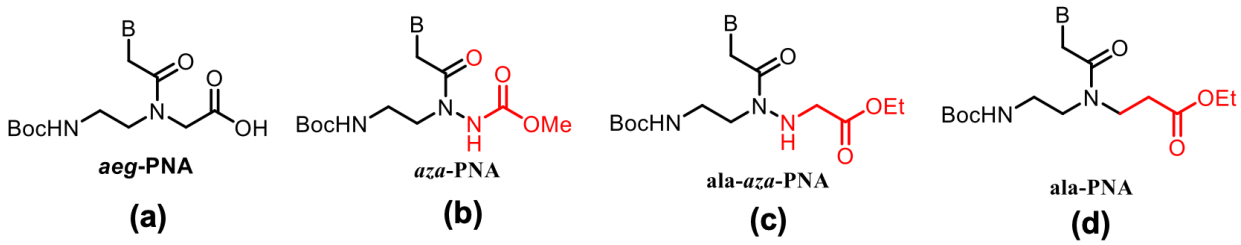


Figure 2.14. Chemical structure of rationally designed *aza*-monomers

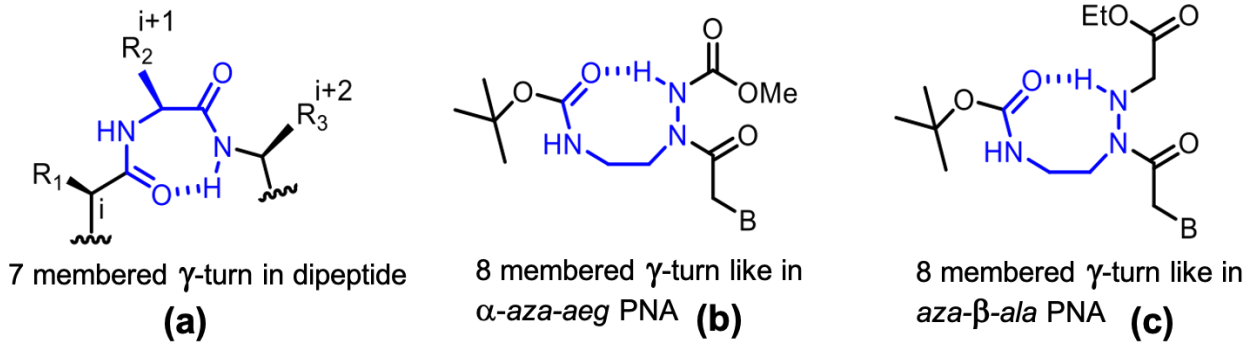


Figure 2.15. Strcutures of (a) γ-turn[29] in dipeptides (b) γ-turn-like *aza*-PNA and (c) γ-turn-like *aza*-β-*ala*-PNA

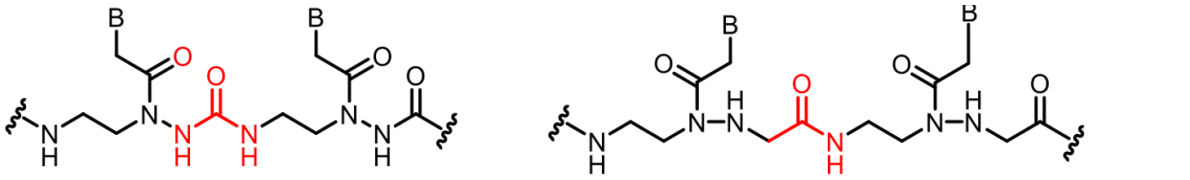


Figure 2.16. (a) Urea linkage in α-*aza*-PNA and (b) polyamide linkage from β-ala-*aza*-PNA oligomers

2.2.1 Specific objectives

The main objectives of this chapter are

(i) Synthesis and characterisation of α-*aza*-PNA-T/C (Figure 2.14b), β-ala-*aza*-PNA-T (Figure 2.14c) and β-ala-PNA-T (Figure 2.14d) monomers

(ii) Study and identification of Z and E rotamers of α-*aza*-PNA-T, β-ala-*aza*-PNA-T and β-ala-PNA T monomers by ^{1}H NMR, 2D-NOSEY and 2D-TOCSY in $CDCl_3$ and DMSO-d_6.

(iii) Study of hydrogen bonding in *aza*-monomers and in β-ala-PNA-T using IR spectroscopy, ^{1}H NMR, variable temperature and solvent dependent ^{1}H NMR.

(iv) Computational studies of rotamer equilibrium in α-*aza*-PNA-T, β-ala-*aza*-PNA-T and β-ala-PNA-T monomers to validate the experimental results.

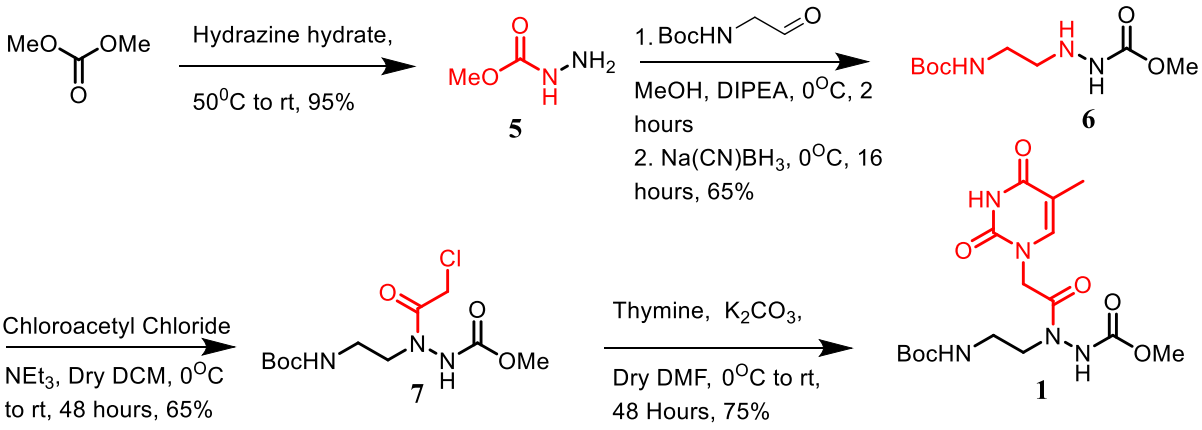
(v) X-Ray structure analysis of the N-chloroacetyl-α-*aza*-*aeg* ester.

2.3 Results: Synthesis, characterization and NMR, Theoretical studies

2.3.1. Synthesis of α-*aza*-PNA-T ester 1

The commercially available dimethyl carbonate was treated with hydrazine hydrate[31] to afford the monoprotected hydrazine derivative **5** (Scheme 1). This was reacted with tert-butyl(2-oxoethyl) carbamate under the condition of reductive amination to generate N-Boc-aminoethyl hydrazido methyl carboxylate **6**. This was reacted with chloroacetyl chloride in dichloromethane (DCM) under basic conditions to yield the N-Boc-aminoethyl (N-chloroacetyl) hydrazido methyl carboxylate **7** having N-chloroacetyl group attached to a-*aza*-aminoethylglycine backbone. The chloro compound was reacted with thymine(T)/N^4-Cbz-protected cytosine (C) in dry DMF under basic condition to obtain the methyl ester of the

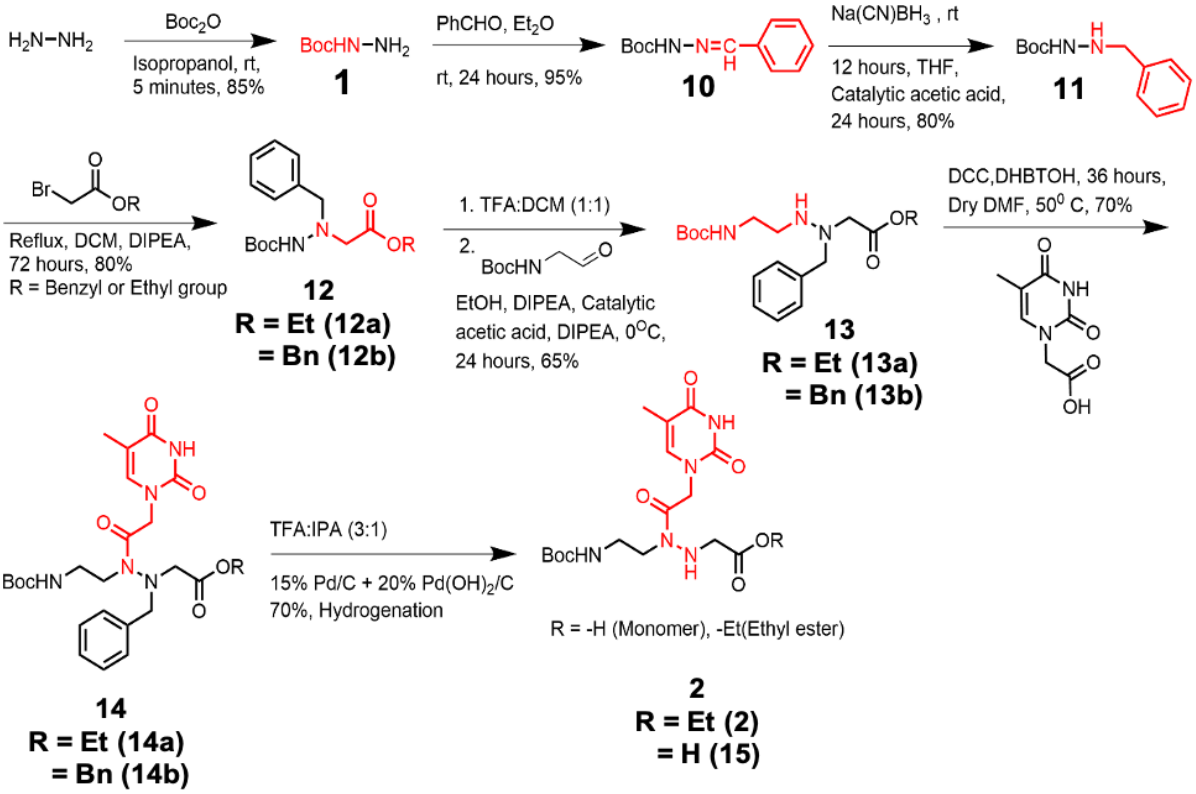
desired α-*aza*-PNA-T/C monomers **1** as a white hygroscopic solid in 75% yield. The synthesized monomers and intermediates were characterised by NMR and Mass spectrometry.



Scheme 1: Synthesis of α-*aza*-PNA-T monomer ester **1**

2.3.2 Synthesis of β-ala-*aza*-PNA-T ester

The commercially available hydrazine hydrate was treated with Boc_2O to afford the mono N-Boc protected hydrazine[32] (**1**, Scheme 2). The obtained N-monoprotected hydrazine derivative was reacted with benzaldehyde to afford the Schiff base **10**, which was reduced by



Scheme 2: Synthesis of β-ala-*aza*-PNA-T monomer and ethel ester

sodium cyanoborohydride ($NaCNBH_3$) to afford the amine derivative **11**. N-alkylation of **11** with benzyl or ethyl bromoacetate under basic conditions afforded the alkylated derivative of hydrazine (Compound **12a** ethyl ester and **12b** benzyl ester). These were individually reacted with tert-butyl(2-oxoethyl) carbamate under the condition of reductive amination to yield the corresponding β-ala-*aza*-PNA esters (Compound **13a** and **13b**, Scheme 2). DCC/DhbtOH mediated coupling of thymine acetic acid with β-ala-*aza*-PNA esters **13a** and **13b**, afforded the β-ala-*aza*-PNA -T esters (**14a** and **14b**). Debenzylation of **14** under the hydrogenolysis condition using mixture of Pd/C and Pd $(OH)_2$/C gives the desired β-ala-*aza*-PNA-T ester monomer (**2**, Scheme 2) and β-ala-*aza*-PNA-T acid (**15**, Figure 2.18) as white hygroscopic solid in 70% yield (Scheme 2). The obtained monomers and intermediate compounds were characterised by NMR and mass spectrometry. *aeg*-PNA monomers and the β-*ala*-PNA-T monomers were synthesized according to reported literature procedure.[33]

2.3.3 Analysis of ^{1}H NMR spectra of monomers

aeg-PNA monomer presents two rotamers around the N-tertiary amide bond with rotational energy barrier (around tertiary N-CO bond) found to be 19±2 kcal/mol using variable temperature NMR.[17,16] The presence of two rotamers of *aeg*-PNA (E and Z) can be identified by ^{1}H NMR spectra, in which each backbone methylene (α, β and γ) and side chain δ methylene protons appear as double peaks with a ratio of 3:1. Isosteric replacement of α-methylene group by NH in *aeg*-PNA and its homologue results α-*aza*-PNA-T (Figure 2.17) and β-ala-*aza*-PNA-T (Figure 2.18). The full and partial ^{1}H NMR spectra from δ 2.5-6.0 ppm of the a-*aza*-PNA-T ester are shown in the Figures 2.17. α-*aza*-PNA-T exhibited two sets of peaks for backbone β-CH_2 and side chain δ-CH_2 (Figure 2.17b). The difference of chemical shift between the two rotamers is 0.1 ppm and the relative ratio of rotamers calculated from the area of integration of the double peaks was 60:40. The hydrazamido α-NH (a_{NH}) and the N-terminal amide NH (b_{NH}) appear around δ 5.1 ppm with the latter (b_{NH}) downfield shifted by 0.2 ppm compared to the hydrazamido α-NH (a_{NH}). Figure 2.17 shows the structure (Figure 2.17b) and full ^{1}H NMR spectra (Figure 2.17a) and enlarged ^{1}H NMR (Figure 2.17c) spectra of α-*aza*-PNA-T in $CDCl_3$.

In comparison, the homologue β-ala-*aza*-PNA-T ester and the free monomer exhibited single peak for each backbone CH_2 (α, β and γ) and the sidechain δ-CH_2 protons indicating presence of a single rotamer for β-ala-*aza*-PNA-T (Figure 2.18). Compared to the α-*aza*-PNA-T, the δ, γ and β-CH_2 protons shows downfield shift in the β-ala-*aza*-PNA-T ester. The downfield shift is maximum for δ-CH_2 (0.7 ppm) compared to β (0.1 ppm) and γ -CH_2 (0.4 ppm) protons in β-ala-*aza*-PNA-T respectively (Figure 2.18). The hydrazamido α-NH (a_{NH})

and the N-terminal amide NH (b_{NH}) protons for β-ala-*aza*-PNA-T appears downfield compared to the shifts of similar protons in α-*aza*-PNA-T (Figure 2.17). Similar to α-*aza*-PNA-T the N-terminal amide NH (b_{NH}) protons (δ 5.3 ppm) appear in the down fielded region compared to the hydrazamido α-NH ($a_{NH,}$, δ 5.1 ppm) proton.

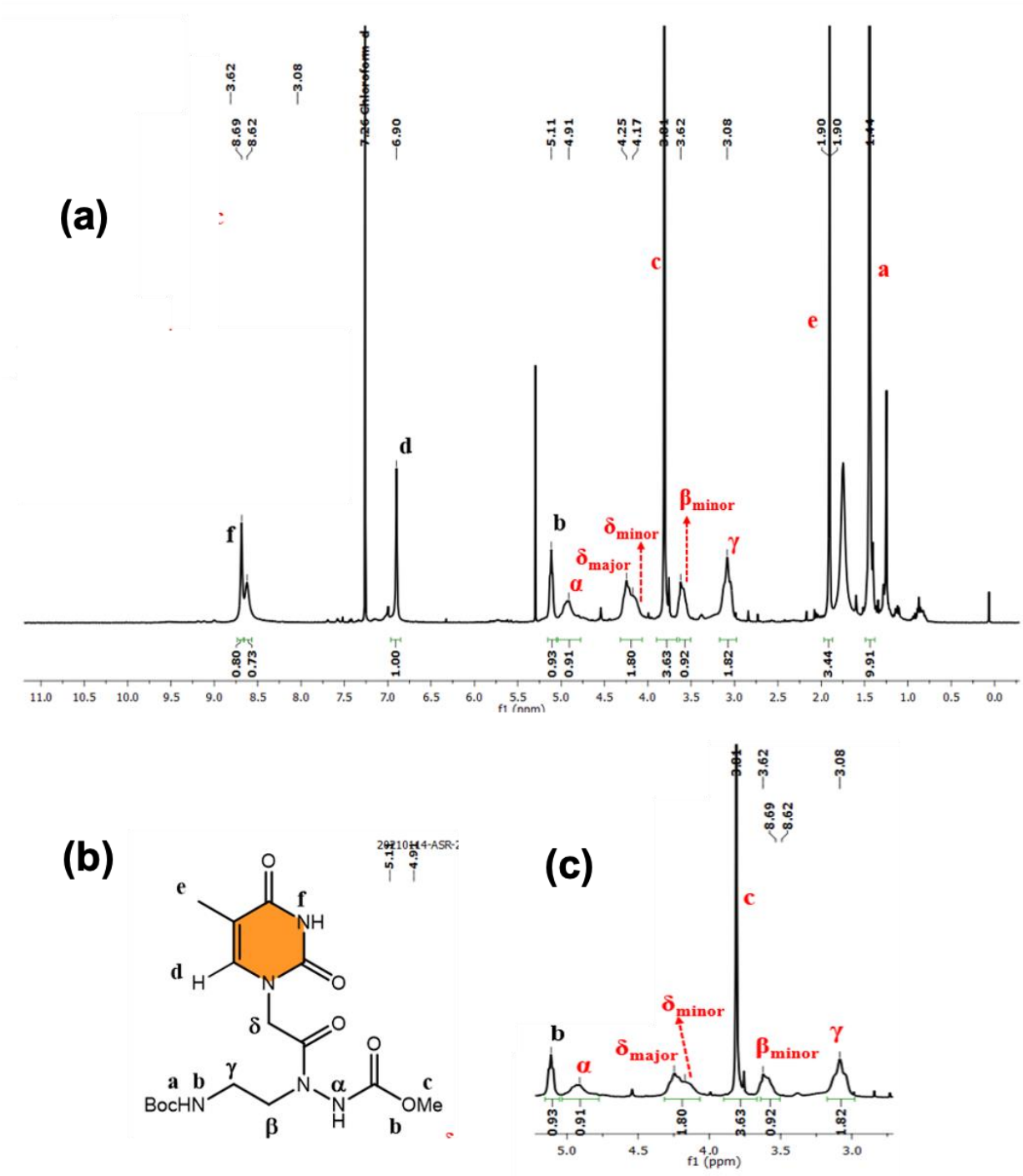


Figure 2.17. (a) Full ^{1}H NMR of α-*aza*-PNA-T (b) Structure of α-*aza*-PNA-T ester and (c) expanded region of 3.0 to 5.0 ppm spectra. Solvent, $CDCl_3$

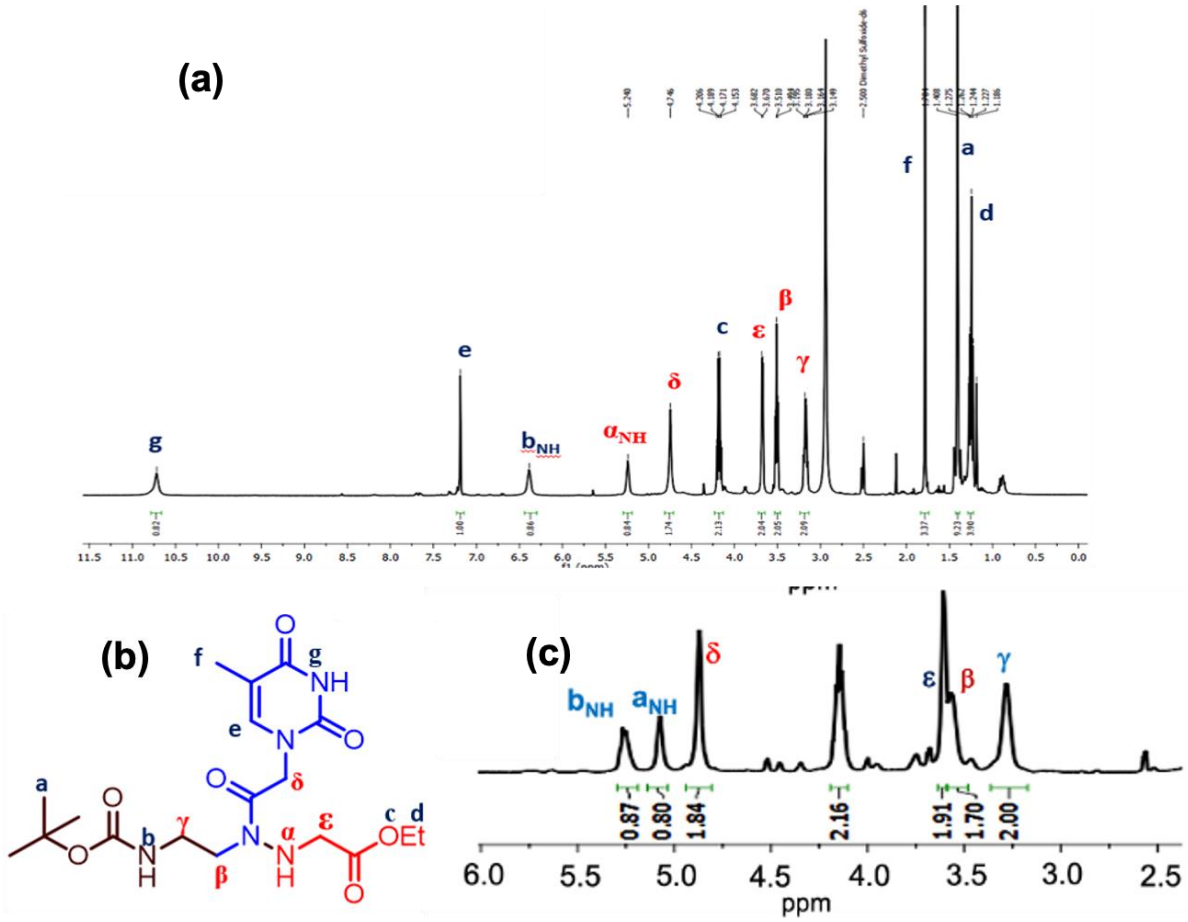


Figure 2.18. (a) Full ^{1}H NMR of *aza*-β-ala PNA-T (b) *aza*-β-ala PNA-T ester and (c) expanded region of 3.1 to 3.8 ppm spectra. Solvent DMSO-d_6.

2.3.4 Assignments of backbone protons in *aza* -PNA-T esters

The assignment of individual protons and identification of Z/E rotamers in *aza*-PNAs were done by combination of 2D TOCSY and 2D NOE connectivity. In Z-rotamer the δ-methylene group is spatially closer to β/γ protons of the backbone, with the expected NOE cross peak between side chain δ-CH_2 and β/γ protons of the backbone (Figure 2.19a). In contrast, in the E-rotamer the δ-methylene group of is closer to α-CH_2 protons of the backbone and so this rotamer is characterised by NOE cross peak between side chain δ-CH_2 and α-CH_2 protons of the backbone (Figure 2.19b and 2.19c). This strategy has been used earlier in assignment of backbone protons in *aeg*-PNA and α/β / γ *gdm*-PNAs.[16,18,20]

The chemical shift assignments of β, γ and α-CH_2 protons were performed by TOCSY and gradient COSY NMR. The β and γ methylene protons are connected through covalent bond and hence show cross peak in gradient COSY (Figure 2.20). The α methylene proton is not J-coupled to any other protons and does not show cross peaks with any of the other protons (Figure 2.20). The total correlated spectroscopy (TOCSY) was employed for the assignment of

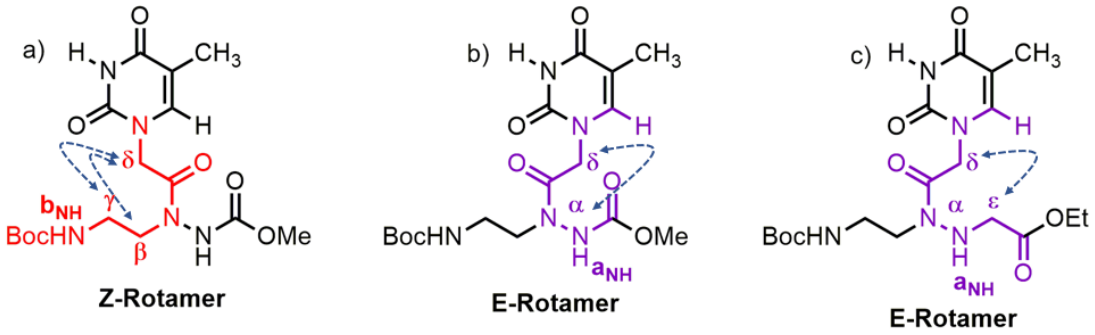


Figure 2.19. Expected NOE connectivity. (a) Z rotamer of α-*aza*-PNA-T (b) E rotamer of α-*aza*-PNA-T and (c) E rotamer of β-ala-*aza*-PNA-T

terminal BocNH (b_{NH}) and backbone NH(α). The b_{NH} is covalently connected with γ methylene protons of the backbone showing cross peak with γ methylene protons (Figure 2.19a) that in turn shows cross peak with β methylenes. In β-ala-*aza*-PNA-T (Figure 2.19c), α-NH is J-connected with ε-CH_2 in and hence shows cross peak with ε-CH_2 (Figure 2.19c). After the assignment of methylene protons, NH protons were assigned by TOCSY and COSY experiments. The peak at δ 4.25 ppm for δ-CH_2 corresponding to major rotamer exhibits cross peak with β/γ -CH_2 protons suggesting the major rotamer to be Z (Figure 2.19a and 2.21, red dotted lines). The peak at δ 4.91 ppm for α-NH proton shows NOE cross peak with δ 4.17 ppm (δ-CH_2, minor rotamer) suggesting the minor rotamer to be E (Figure 2.19b and 2.21, blue dotted lines). The observed NOSEY cross peaks of α-NH with δ-CH_2 further validates the assignment of 4.9 ppm to α-NH and the peak at 5.1 ppm to the carbamate -NH (BocNH, b_{NH}). The weak cross peak of TH-6 at 4.17 ppm to δ_{minor} and α-NH suggests that the H-6 face of the thymine of minor E-rotamer is oriented towards the carboxy end of the backbone (Figure 2.21, purple dotted line).

In contrast, the simplicity of 1D ^{1}H NMR of β-ala-*aza*-PNA-T with only one peak for each of the methylene protons, suggests a single rotamer for the β-ala-*aza*-PNA-T ester (Figure 2.18). The primary assignment of the backbone methylene protons at β and γ protons was done at δ 3.5 ppm and 3.2 ppm respectively, while the downfield signal at δ 3.7 ppm corresponded to ε methylene The backbone a_{NH} was assigned to signal at δ 5.4 ppm, upfield compared to the terminal BocNH (b_{NH}, Figure 2.18b). ^{1}H-^{1}H gradient COSY spectra of β-ala-*aza*-PNA-T in the DMSO-d_6 (Figure 2.22) exhibited cross peak between methylene protons at δ 3.5 and 3.2 ppm confirming the J-connectivity of β and γ backbone protons (Figure 2.22). The methylene proton at δ 3.7 ppm lacked any cross peak in the ^{1}H-^{1}H gradient COSY spectra of β-ala-*aza*-PNA-T

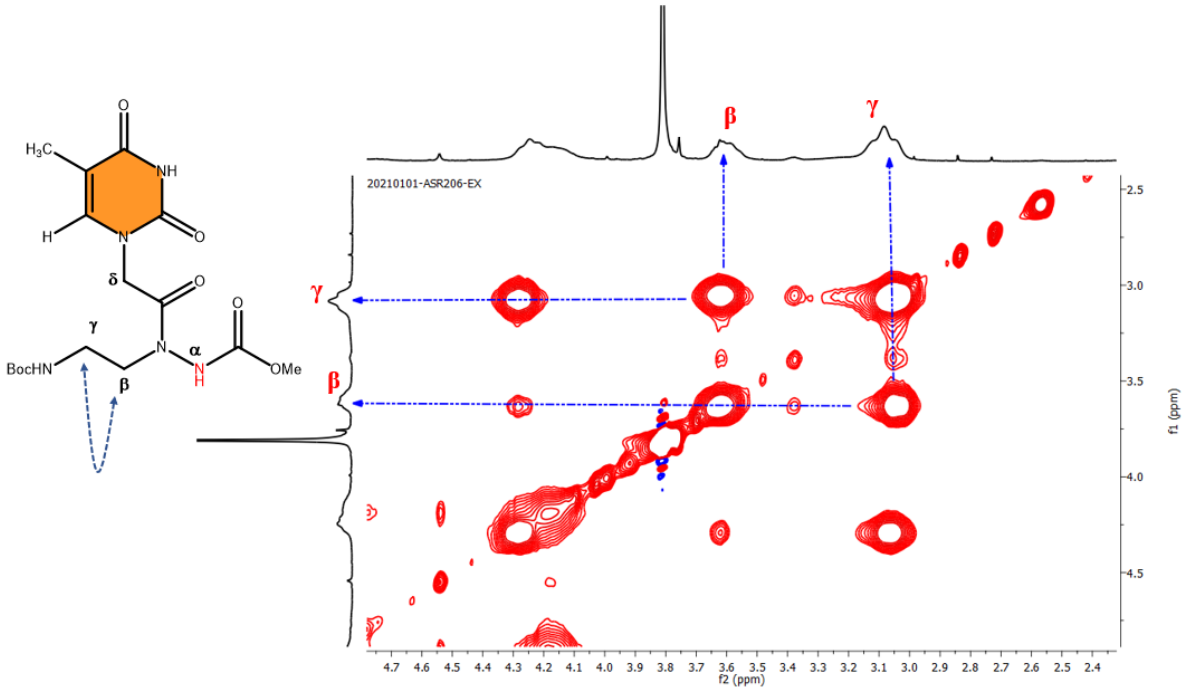


Figure 2.20. Partial Gradient COSY spectra of α-*aza*-PNA-T in $CDCl_3$ showing connectivity's of β and γ-methylene protons

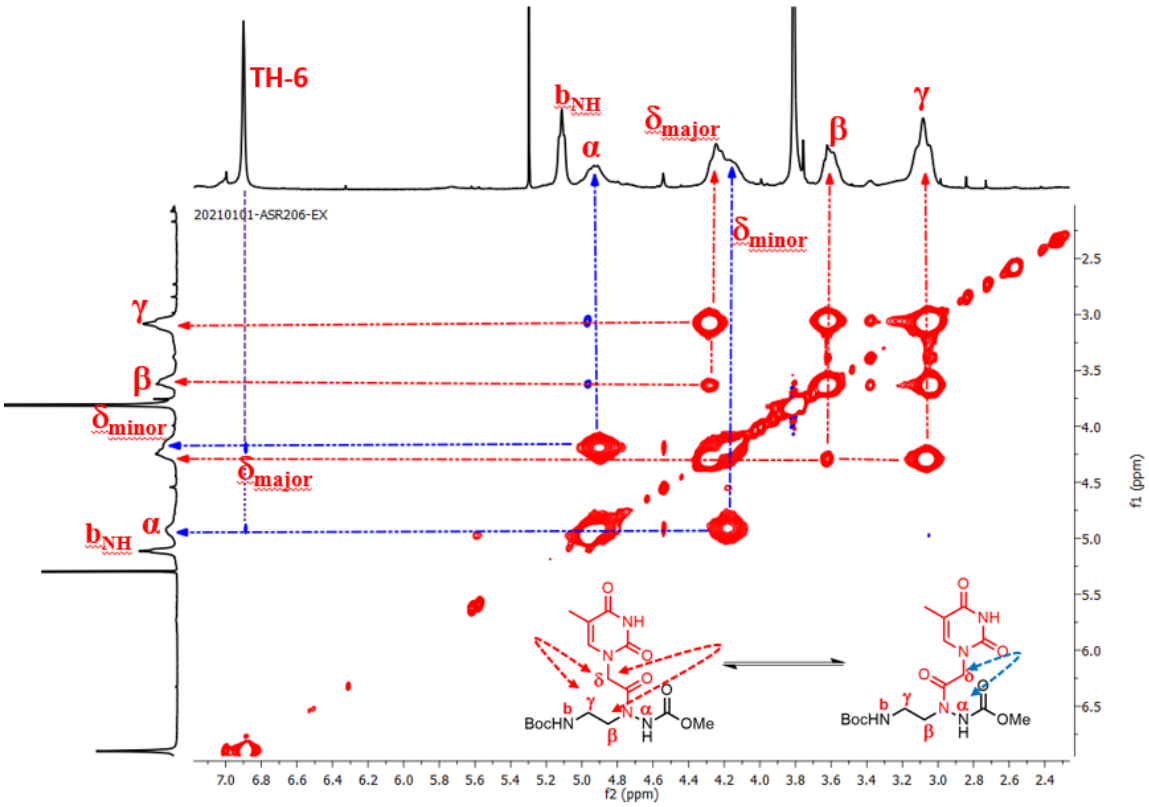


Figure 2.21. Partial 2D-NOSEY spectra of α-*aza*-PNA-T in $CDCl_3$ showing connectivity's of major and minor rotamer of α-*aza*-PNA-T

confirming its assignment to ε-CH_2, of methylene protons (Figure 2.22).

The ^{1}H-^{1}H TOCSY spectra (Figure 2.23) of β-ala-*aza*-PNA-T in the DMSO-d_6 showed the NH proton at δ 5.4 ppm having cross peak with ε-CH_2 suggesting it to be hydrazamido α-NH (a_{NH}) The NH peak at δ 6.45 ppm exhibited J-cross peak with γ backbone protons in the ^{1}H-^{1}H COSY spectra suggesting it to be terminal BocNH (b_{NH}). These observations collectively confirmed all primary assignments of protons (Figure 2.18).

Figure 2.24 shows the ^{1}H-^{1}H NOSEY spectra of the β-ala-*aza*-PNA-T in $CDCl_3$. The simplicity of the NOSEY spectra reflects existence of single rotamer for β-ala-*aza*-PNA-T in solution. The side chain δ-CH_2 shows cross peak (cross peak 1, blue dotted line, Figure 2.24)

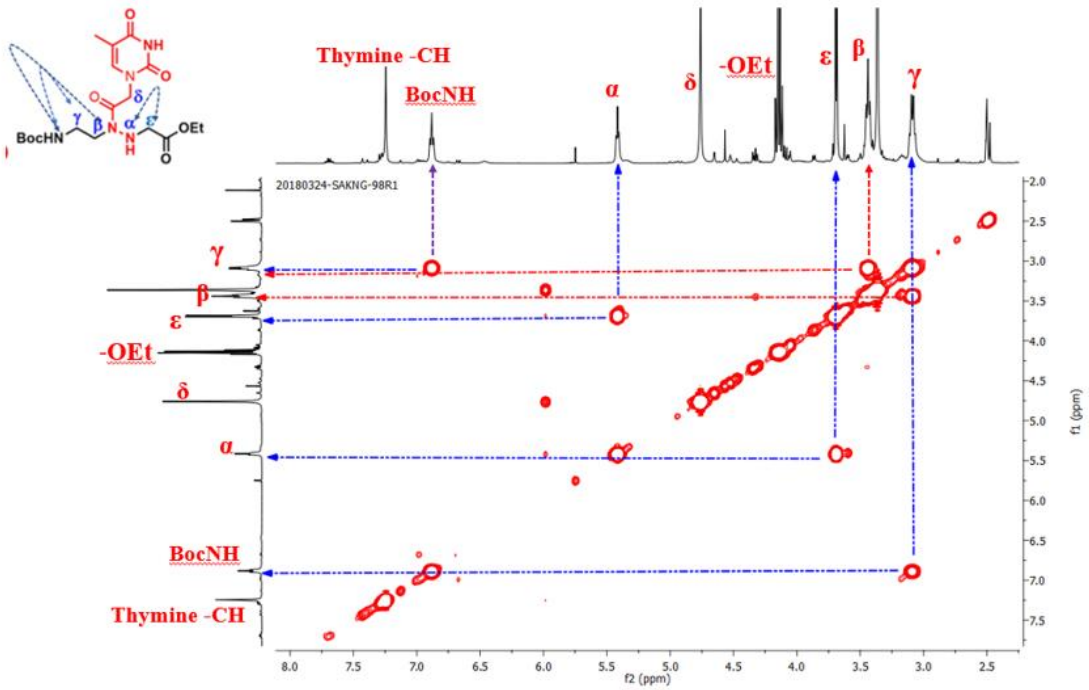


Figure 2.22. Partial gradient COSY spectra of β-ala-*aza*-PNA-T in DMSO-d_6

with the ε-CH_2 suggesting the single rotamer to be E-Rotamer. The NOSEY Cross peak between the δ methylene of the sidechain with α-NH (a_{NH}) of the backbone (purple dotted line, cross peak 3, Figure 2.24) and absence of cross peak between side chain δ-CH_2 and β/γ-CH_2 further confirms the assignment of E-rotamer in $CDCl_3$. The cross peak between α-NH (a_{NH}) and γ-CH_2 is interesting (cross peak 2) and hints at possible folding of backbone of the β-ala-*aza*-PNA-T monomer in $CDCl_3$. The COSY cross peak of backbone α-NH (a_{NH}) and ε-methylene (cross peak 4, Figure 2.24) reconfirms the assignment of δ 5.1 ppm peak as backbone α-NH (a_{NH}) and δ 5.35 ppm as N-terminal BocNH (b_{NH}).

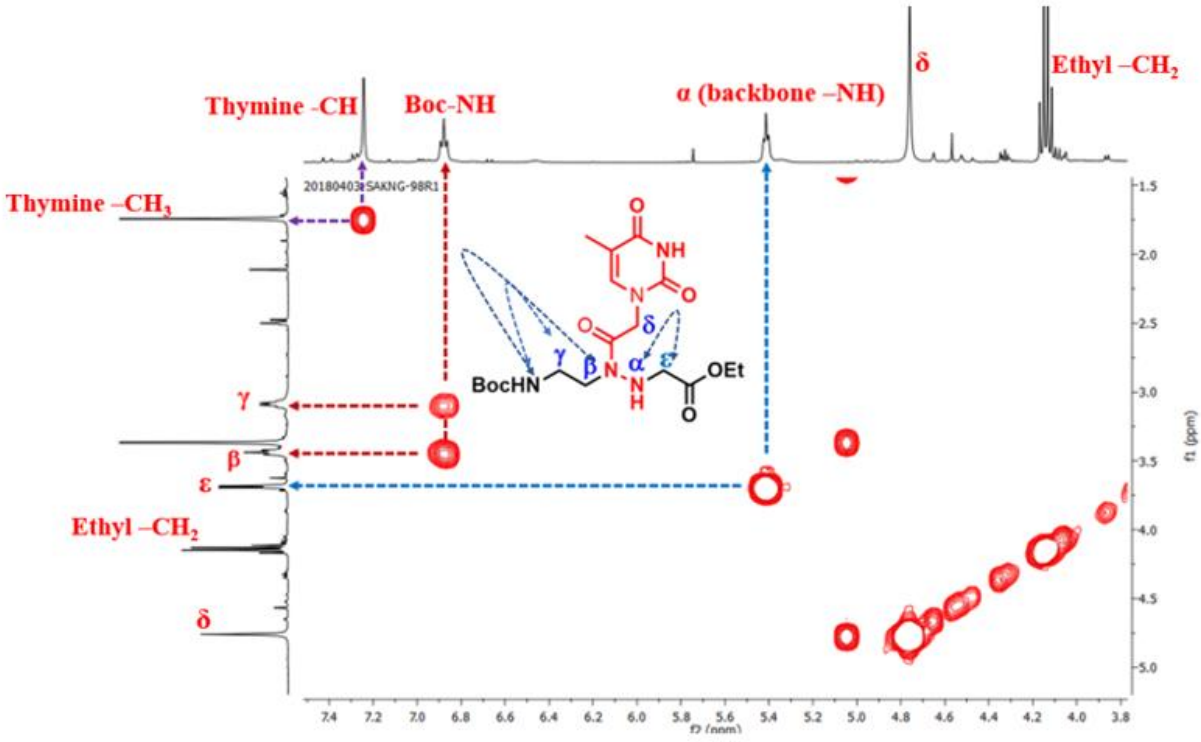


Figure 2.23. Partial ^{1}H-^{1}H TOCSY spectra of β-ala-*aza*-PNA-T in DMSO-d_6

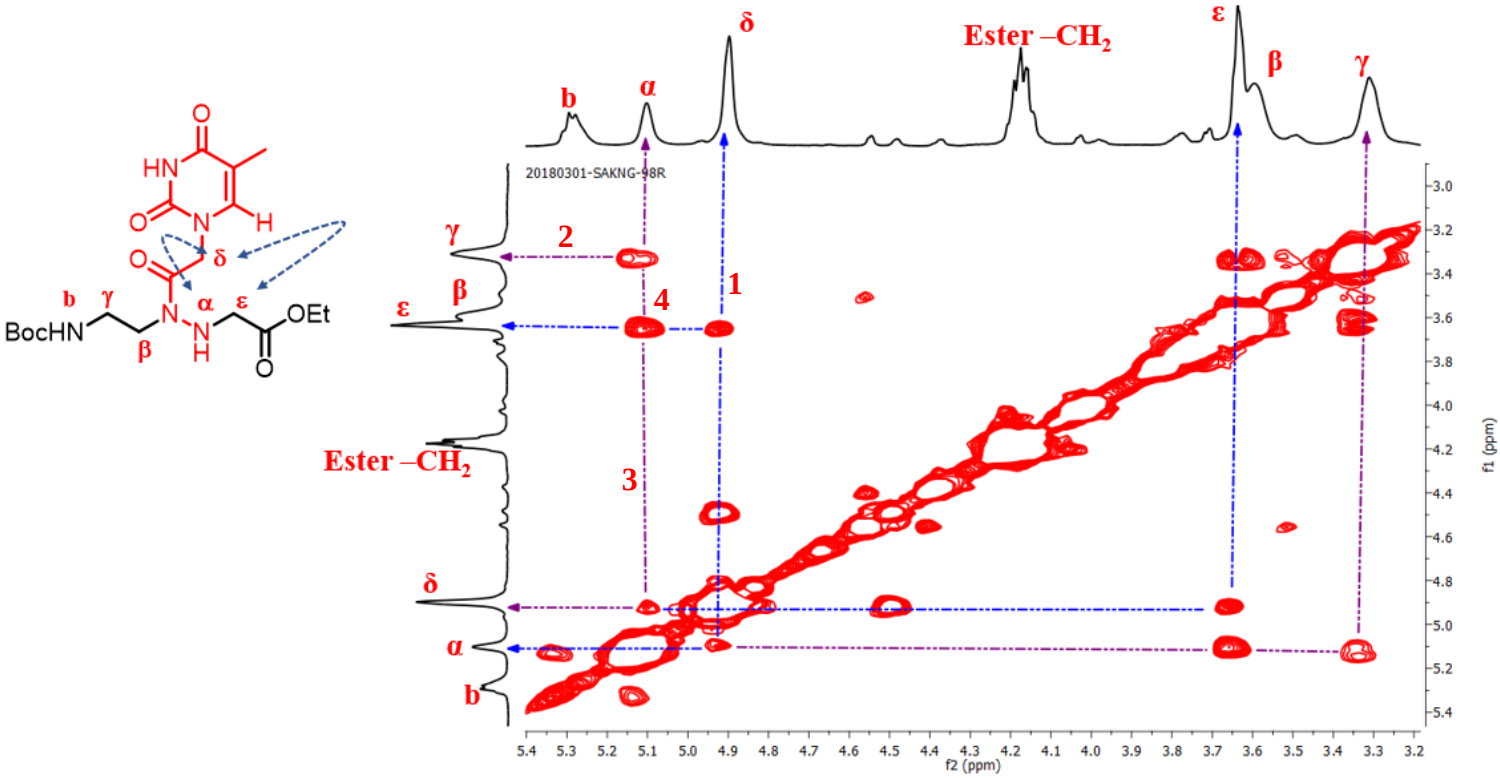


Figure 2.24. Partial 2D-NOSEY spectra of β-ala-*aza*-PNA-T in $CDCl_3$

The observance of single E-rotamer in β-ala-*aza*-PNA-T has parallel to similar exclusive E-Rotamer seen in β-*gemdimethyl*-PNA-T (β-*gdm*) monomer.[20] The existence of single E-rotamer in β-*gemdimethyl*-PNA-T (β-*gdm*) monomer is dictated by the steric effect of *gemdimethyl* group and stabilization through C-H…O bonding.[20] The presently studied β-ala-*aza*-PNA-T and α-*aza*-PNA-T ester monomers lack such steric effects and instead offer hydrogen bonding possibilities of α-NH (a_{NH}, Figure 2.15) with several carbonyls (δ-CO, ester-CO, and carbamate) which may lead to folding of the backbone with concomitant selective stabilization of E-rotamer. To understand the effect H-bonding and solvation on the rotamer preferences, NOSEY experiment was performed in DMSO-d_6. DMSO is known for its solvation ability and the highly polar nature of DMSO also disrupt the hydrogen bonding and may potentially influence the rotamer preferences of the PNA monomers and oligomers.[35] The unmodified *aeg*-PNA prefers Z-Rotamers over E-Rotamer in $CDCl_3$ but in DMSO-d_6, the E-Rotamer is preferred over the Z-rotamer.[16] Hence the effect of DMSO on the rotamer preferences of *aza*-PNA monomers were studied.

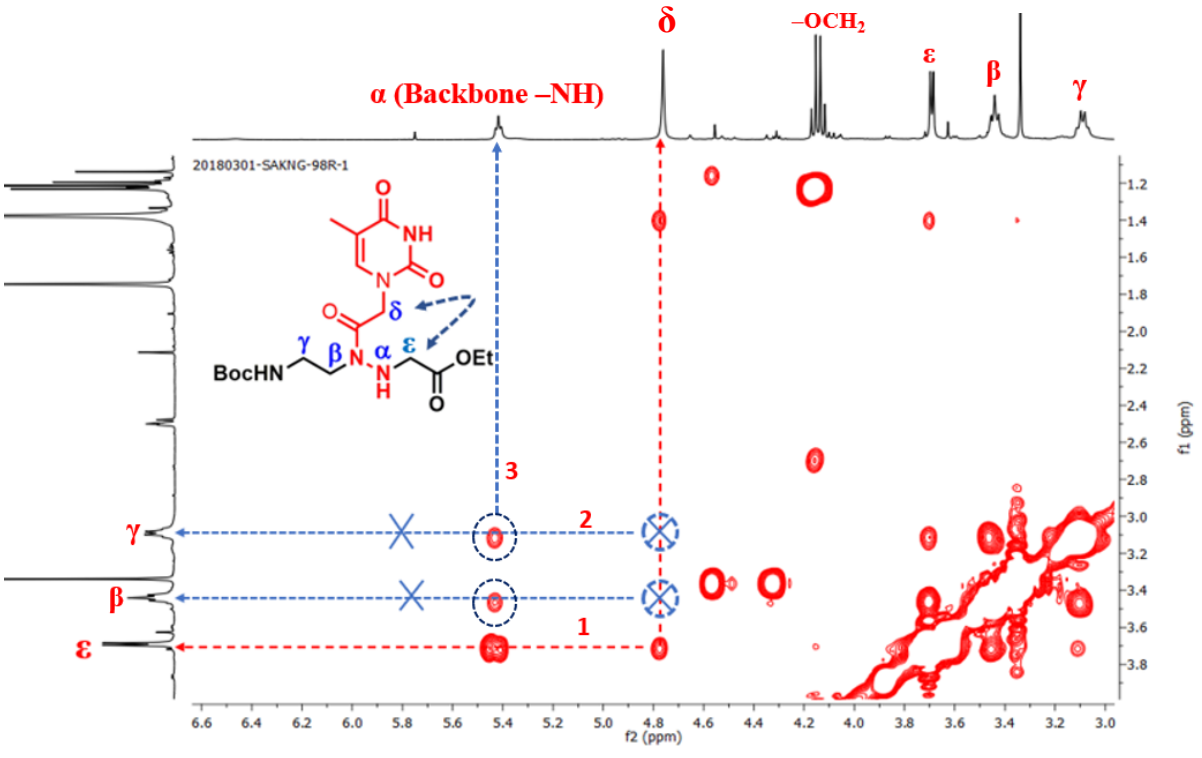


Figure 2.25. Partial 2D-NOSEY spectra of β-ala-*aza*-PNA-T in DMSO-d_6

Figure 2.25 shows the ^1H-^1H NOSEY spectra of β-ala-*aza*-PNA-T in DMSO-d_6, the simplicity of NOSEY spectra reflects the single rotamer nature of the β-ala-*aza*-PNA-T in DMSO-d_6, similar as in $CDCl_3$. The observed cross peak between the side chain δ-CH_2 and the

ε-CH_2 suggests that the single rotamer to be E-Rotamer (cross peak 1, red dotted line, Figure 2.25). The non-observance of cross peak of δ-CH_2 with the β/γ-CH_2 protons further confirms the observation and indirectly supports the rotamer to be E and not Z (Cross peak 2, blue dotted line, Figure 2.25). Similarly, as the NOSEY spectra in $CDCl_3$, the α-NH (a_{NH}) exhibited cross peaks with β/γ-CH_2 (cross peak 3, blue dotted line, Figure 2.25) in DMSO-d_6 suggesting a possible backbone folded structure for the β-ala-*aza*-PNA-T in DMSO-d_6. Thus, the E-rotamer preference of β-ala-*aza*-PNA-T remains unaffected in DMSO-d_6 due to solvation. However, the downfield chemical shift of NH protons by 0.2 ppm suggests an intramolecular hydrogen bonding of b_{NH} in the β-ala-*aza*-PNA-T monomer. This and the effect of solvation switches the observed rotamer preference for α-*aza*-PNA-T in $CDCl_3$ to reverse preference for rotamer stabilization in DMSO-d_6.

In both β-ala-*aza*-PNA-T and α-*aza*-PNA-T, the chemical nature of α-NH (a_{NH}) the a_{NH} being hydrazido secondary amine in β-ala-*aza*-PNA-T, while in α-*aza*-PNA-T that lacks ε-CH_2 makes it carbamate like that is more rigid compared to a secondary amine. Carbamate NH is more susceptible towards solvation than a hydrogen bonded secondary amine.[35] These factors not only affect their relative chemical shifts, but also lead to a reversal of rotamer preferences for α-*aza*-PNA-T in DMSO-d_6. Very weak cross peaks observed in the NOSEY spectra of α-*aza*-PNA-T in DMSO-d_6 (See appendix-I), indicate loss of folded structure in DMSO-d_6. The side chain δ-CH_2 (δ_{major}) shows weak cross peak with the α-NH (a_{NH}) suggesting the major rotamer in DMSO to be E. Thus, solvation alters the rotamer preferences for α-*aza*-PNA-T in solution perhaps with an extended structure of the monomer in solution.

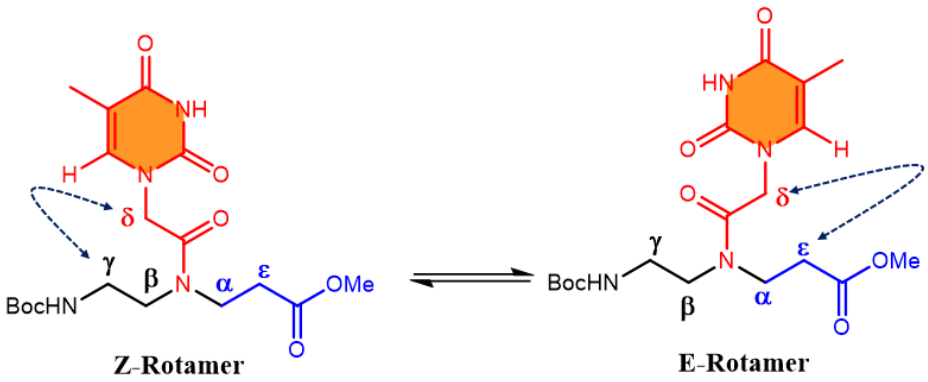


Figure 2.26. E and Z rotamer of β-*ala*-PNA-T

To determine the role of α-NH (a_{NH}) in stabilization of E-rotamer, β-*ala*-PNA-T ester was synthesized where the a_{NH} was replaced by isosteric methylene group. The primary ^{1}H NMR analysis suggests a mixture of rotamer for β-*ala*-PNA-T with a ratio of 53:47 in $CDCl_3$. Figure 2.27 shows the partial NOSEY spectra of the β-*ala*-PNA-T in $CDCl_3$, whose complexity reflects a rotamer mixture composition in solution. The major rotamer (δ_{major}) shows cross peak with α/ε-methylene protons of the backbone suggesting a major E rotamer in $CDCl_3$ (Figure 2.27, blue dotted line). The minor rotamer shows (δ_{minor}) cross peak with β/γ-methylene protons of the PNA backbone suggesting a minor Z-Rotamer in $CDCl_3$ (Figure 2.26 and 2.27, red dotted line). Thus, replacing α-NH with isosteric methylene group results a rotamer mixture with major E and minor Z in $CDCl_3$ (Figure 2.27). These results suggest an active role of a_{NH} towards stabilization of E rotamer over Z rotamer for β-ala-*aza*-PNA-T in $CDCl_3$ and DMSO-d_6.

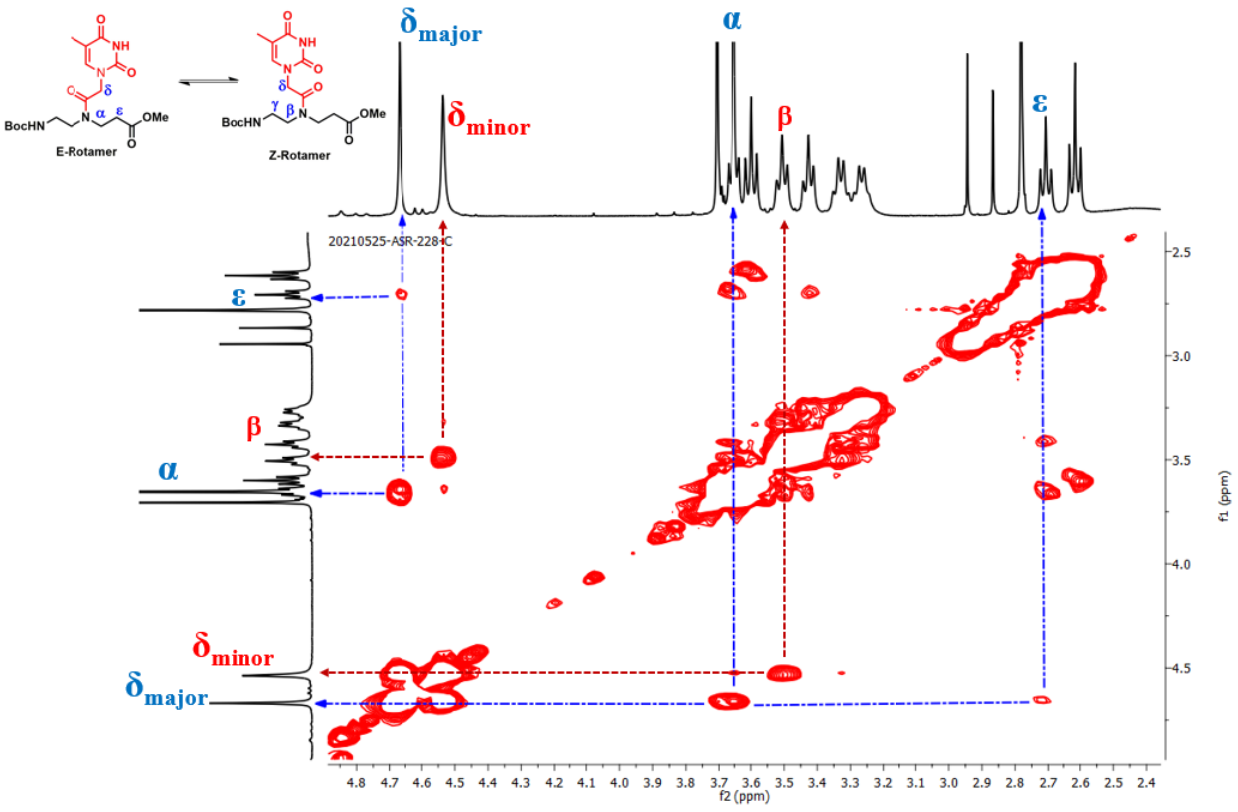


Figure 2.27. Partial ^{1}H-^{1}H 2D NOSEY spectra of β-ala-PNA-T (Major E and minor Z, Solvent $CDCl_3$)

A reverse preference of rotamer stabilization (Z > E) was observed for β-*ala*-PNA-T in DMSO-d_6, following a opposite trend as seen in *aeg*-PNA. Figure 2.28 shows the partial NOSEY spectra of β-*ala*-PNA-T in DMSO-d_6. The major rotamer (δ_{major}) shows cross peak with the β/γ-CH_2 protons of the backbone suggesting a major Z-rotamer in DMSO-d_6 (Figure

2.28, red dotted line). The minor rotamer (δ_{minor}) shows cross peak with the α/ε -CH_2 protons, suggesting it to be E rotamer in DMSO-d_6 (Figure 2.28, Blue dotted line). Hence *deaza*-analogue, the β-*ala*-PNA-T shows mixture of rotamers with preferences changing drastically upon solvation. The results suggest intra-residue H-bonding in backbone that stabilizes the E-Rotamer over Z-rotamer irrespective of solvation.

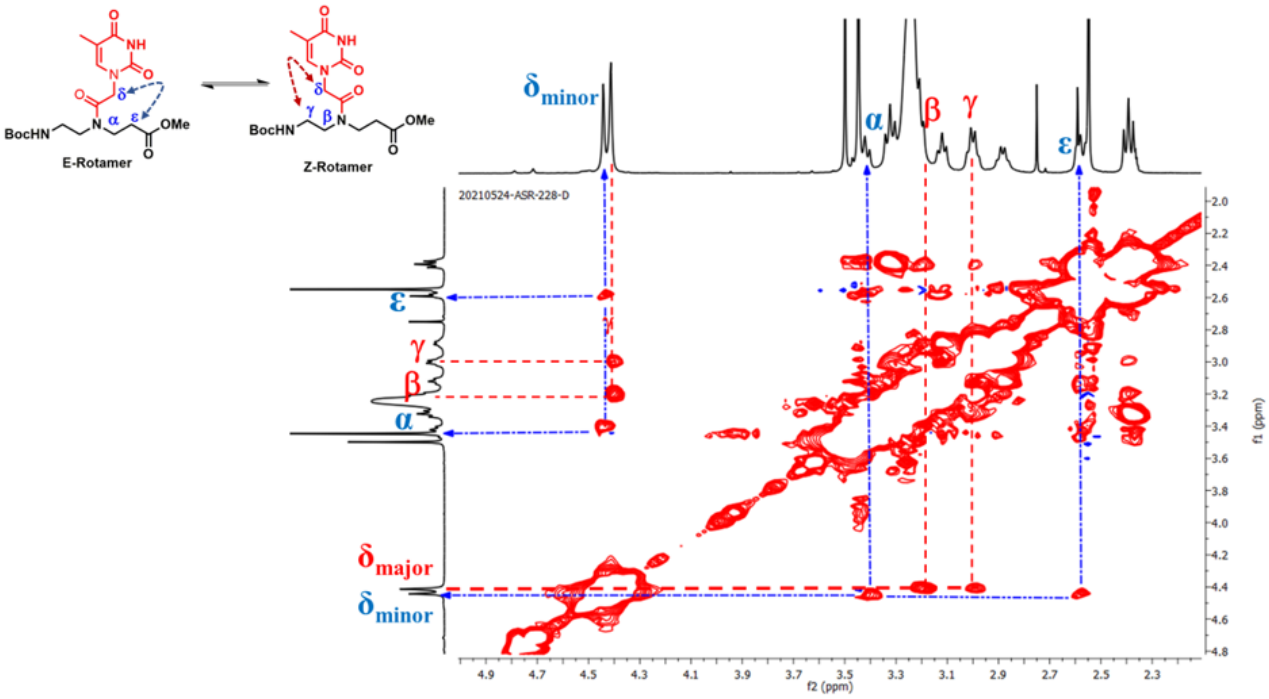


Figure 2.28. Partial ^{1}H-^{1}H 2D NOSEY spectra of β-ala-PNA-T (Major E and minor Z, Solvent DMSO-d_6)

2.3.5 Role of C-terminal and N-terminal protecting groups in the rotamer stabilization of β-ala-*aza*-PNA-T

To further understand the role of N-terminal amine and C-terminal acid group of the PNA backbone towards the stabilization of the preferred rotamers, analysis of the free acid monomers and the free amine monomer were also performed. Free acid monomers of the β-ala-*aza*-PNA-T and β-*ala*-PNA-T exhibit similar trends of rotamer preferences as observed in corresponding ester analogues. Free acid β-ala-*aza*-PNA-T shows single E rotamer in DMSO-d_6 and the free acid β-*ala*-PNA-T shows mixture of rotamer as observed for the methyl/ethyl ester (major E and minor Z in $CDCl_3$). This ruled out the possibility of C-terminal acid group affecting the rotamer preferences of the β-ala-*aza*-PNA-T and the β-*ala*-PNA-T. The deprotection of the N-terminal Boc group afforded the free amine monomers and the primary ^{1}H NMR analysis of the free amine (Figure 2.29) shows that β-ala-*aza*-PNA-T retains its single

rotamer property suggesting that the N-terminal protecting group does not have much effect on the rotamer preferences.

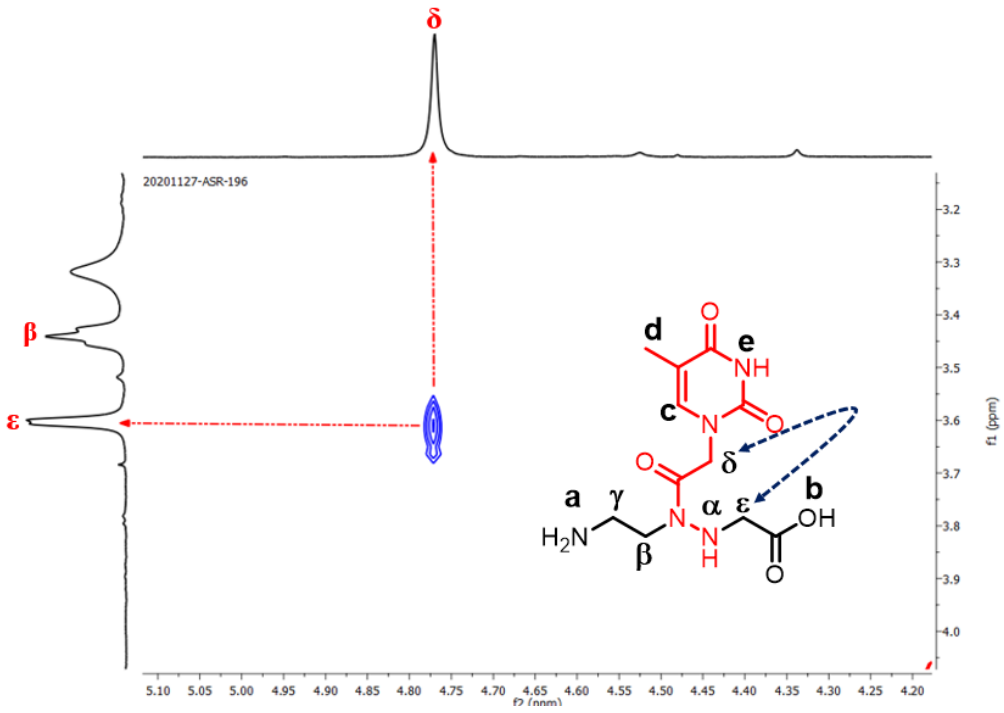


Figure 2.29. Partial ^{1}H-^{1}H NOSEY spectra of the free amine β-ala-*aza*-PNA-T monomer

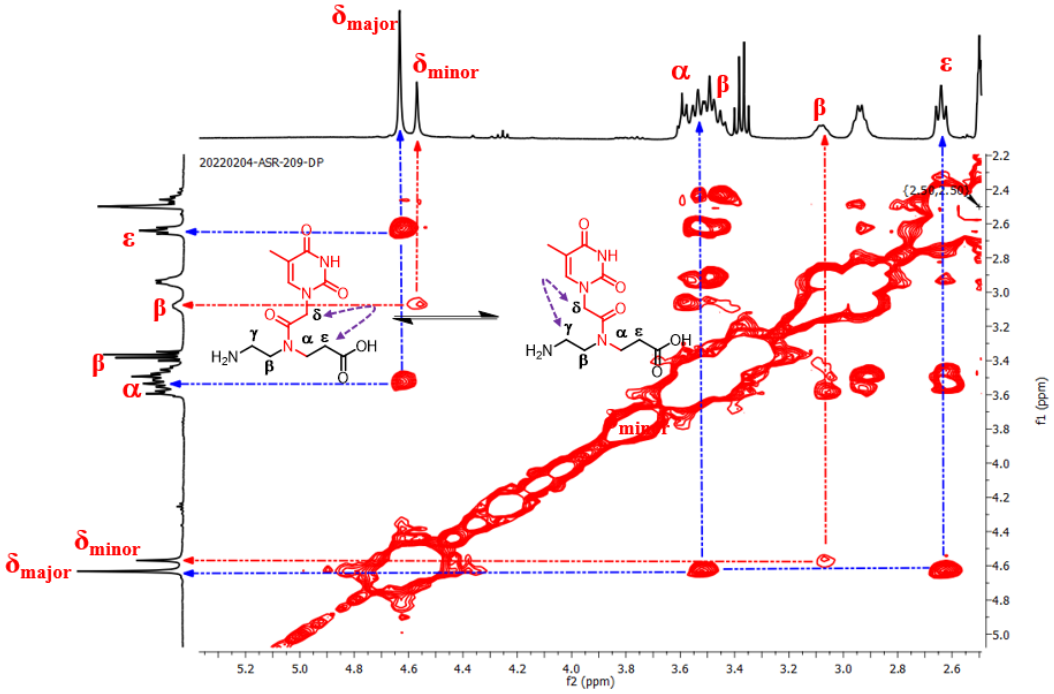


Figure 2.30. Partial ^{1}H-^{1}H NOSEY spectra of the free amine β-*ala*-PNA-T monomer (major Z minor E, Solvent DMSO-d_6)

The free amine β-*ala*-PNA-T shows a mixture of rotamer in $CDCl_3$ and DMSO-d_6. Figure 2.30 shows the partial NOSEY spectra of the free amine β-*ala*-PNA-T in DMSO-d_6. The major rotamer (δ_{major}) shows cross peak with the α/ε-CH_2 protons of the backbone suggesting the major rotamer to be E in DMSO-d_6 (Figure 2.30, blue dotted line). The minor rotamer (δ_{minor}) shows weak cross peak with the β-CH_2 of the backbone suggesting a minor rotamer to be Z (Figure 2.30, red dotted line). These results are reverse of the rotamer preferences of fully protected β-*ala*-PNA-T in DMSO-d_6 (major Z, minor E). The differential selectivity of rotamers in the free amine β-*ala*-PNA-T and fully protected/free acid β-*ala*-PNA-T reveals the importance N-terminal tertiary amide carbonyl group in preferential stabilization of rotamers in fully protected/free acid β-*ala*-PNA-T monomer.

^{1}H NMR of β-ala-*aza*-PNA-T precursor (compounds **14a** and **14b**) that has Nα-benzyl group shows a complex mixture of rotamers in solution (see appendix-I). The deprotection of N-benzyl group affords the β-ala-*aza*-PNA-T monomer, which shows a single E-rotamer in solution. This result reveals the importance of a_{NH} (α-NH) in the stabilization of rotamers in *aza*-PNA systems. Hence solvent and temperature dependent NMR studies were performed to understand the importance of a_{NH} for the stabilization of rotamers in *aza*-PNA monomers.

2.3.6 Solvent dependent NMR shifts and NMR experiments

Comparative ^{1}H NMR spectra of α-*aza*-PNA-T in $CDCl_3$ and DMSO-d_6 are shown in Figure 2.31. In the ^{1}H NMR spectra of α-*aza*-PNA-T in DMSO-d_6, the β-CH_2 shows slight upfield shift compared to that in $CDCl_3$, while the side chain δ-CH_2 and the backbone γ-CH_2 show downfield shifts in DMSO-d_6. A notable feature is the large downfield shift of 2.0 ppm and 4.6 ppm for a_{NH} and b_{NH} protons in α-*aza*-PNA-T, suggesting possible involvement in internal hydrogen bonding in $CDCl_3$. Figure 2.31b shows the comparative ^{1}H NMR spectra of β-ala-*aza*-PNA-T in DMSO-d_6 and $CDCl_3$. The β-ala-*aza*-PNA-T exhibits single rotamer in both $CDCl_3$ and DMSO-d_6. β/γ-CH_2 of backbone and side chain δ-CH_2 show an upfield shift of 0.2 ppm-0.3 ppm in DMSO-d_6, but slight downfield shift was observed for ε-CH_2 protons. The downfield shift of a_{NH} (0.2 ppm) and b_{NH} (1.3 ppm) for β-ala-*aza*-PNA-T in DMSO-d_6 is much less compared to α-*aza*-PNA-T, but still suggest their involvement in intramolecular hydrogen bonding. Considering a very small downfield shift of a_{NH} (α-NH, 0.2 ppm) in β-ala-*aza*-PNA-T in DMSO-d_6 suggests the hydrogen binding to be intramolecular. To further investigate the role of a_{NH} in intramolecular hydrogen bonding, the titration of $CDCl_3$ solution with small aliquots of DMSO-d_6 or D_2O and the temperature dependence of NMR were studied.

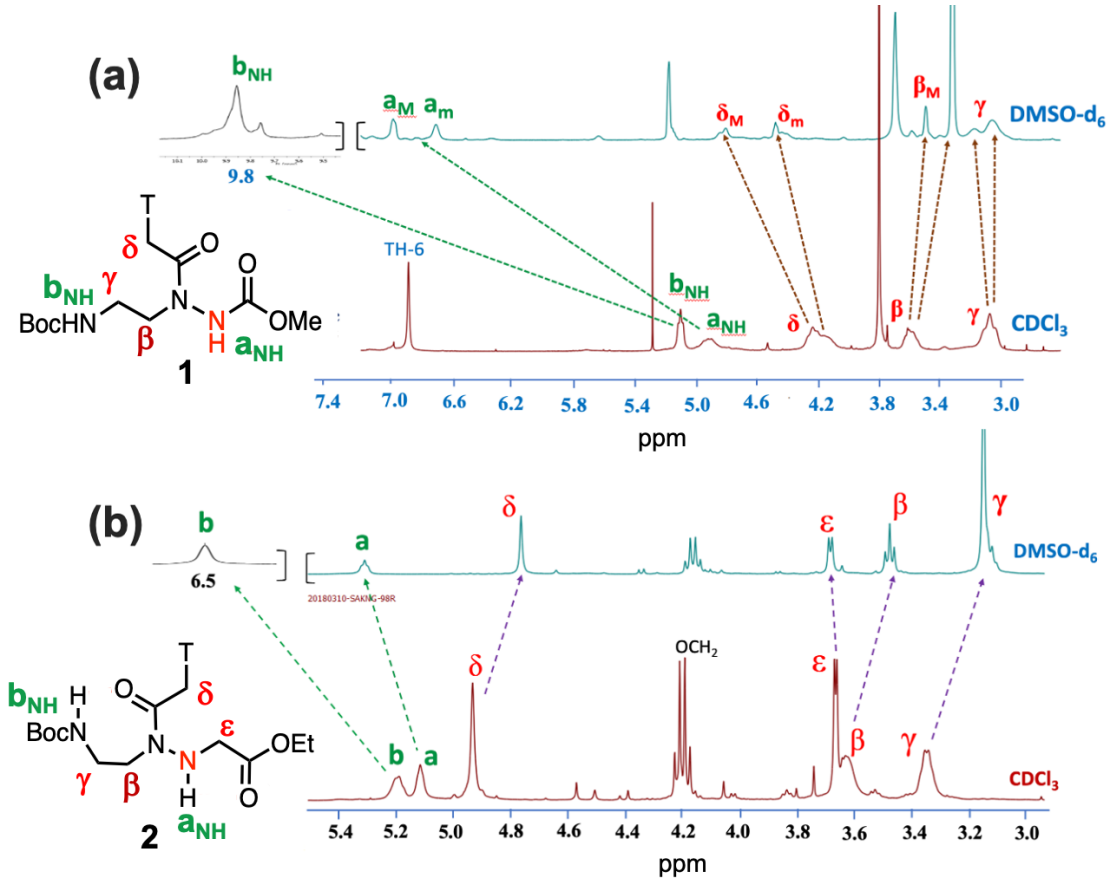


Figure 2.31. a) Comparative ^{1}H NMR spectra of a-*aza*-PNA T and (b) *aza*-b-*ala*-PNA-T in $CDCl_3$ and DMSO-d_6 at 24 ^{o}C.

Dimethyl sulfoxide (DMSO) is known for its ability to break the intermolecular interactions and weak intramolecular interactions through solvation of the polar groups in the molecules.[35] Sequential addition of DMSO-d_6 in a $CDCl_3$ solution of β-ala-*aza*-PNA-T should therefore incrementally break the intramolecular hydrogen bonding, and consequent solvation of NH groups leads to downfield chemical shift in ^{1}H NMR spectra. β-ala-*aza*-PNA-T consist of three NH groups in the molecule, backbone α-NH (a_{NH}), BocNH (b_{NH}) and the nucleobase -amide NH. Variation of NH chemical shift of backbone NH (a_{NH}) during DMSO-titration will be of primary interest since this being adjacent to the t-amide group may influence the stabilization of rotamers. DMSO titration study (Figure 2.32) revealed that backbone NH (a_{NH}) chemical shift shifted downfield by just 0.125 ppm, whereas the BocNH shifted by 1.03 ppm. (8 times that of backbone a-NH). This suggested that in β-ala-*aza*-PNA-T, the backbone α-NH is perhaps participating in a strong intramolecular hydrogen bonding interaction. In comparison, with β-*ala*-PNA-T that lacks the a_{NH} shows two rotamers and hence two signals for t-boc b_{NH}, separated by 0.3 ppm, with the major isomer being the downfied shifted signal (Figure 2.33).

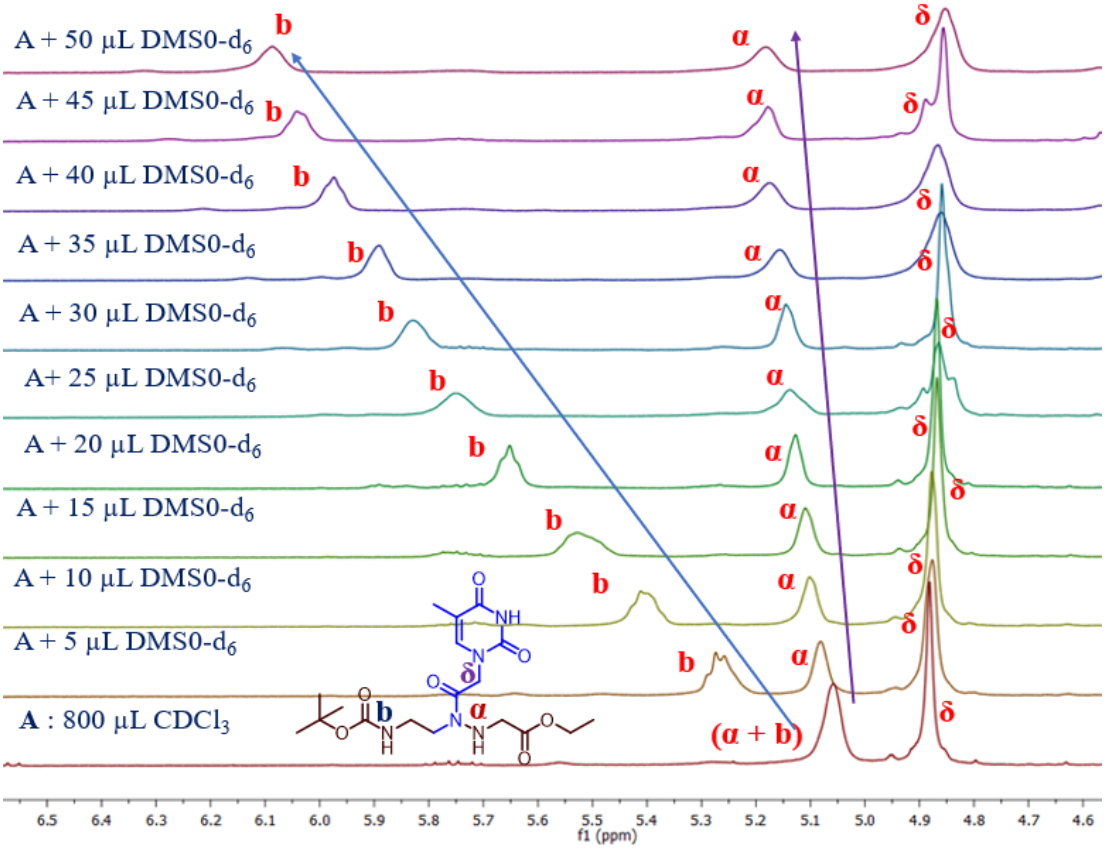


Figure 2.32. Variation of chemical shift of backbone NH (α, a_{NH}) and BocNH (b_{NH}) in DMSO titration experiment of *aza*-β-*ala*-PNA-T.

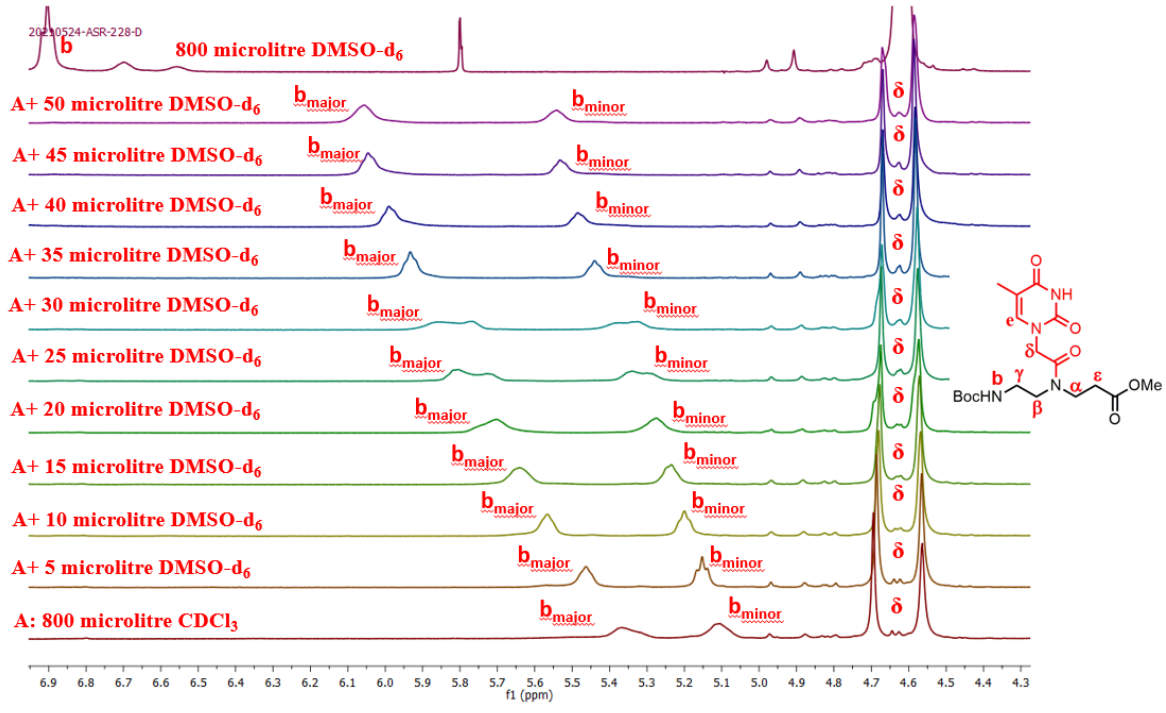


Figure 2.33. Variation of chemical shift of BocNH (b_{NH}) in DMSO titration experiment of β-ala-PNA-T.

The sequential addition of DMSO-d_6 to a $CDCl_3$ solution of β-*ala*-PNA-T shifted both the b_{NH} signals to downfield. The downfield major and upfield minor rotamer both exhibited downfield shifts of 0.8 ppm and 0.43 ppm respectively in the DMSO titration experiment (Figure 2.33). The shift of both major and minor b_{NH} protons (0.8 ppm and 0.43 ppm) are much higher compared to the shift of a_{NH} in β-ala-*aza*-PNA-T (0.125 ppm) suggesting stronger solvation effects, due to a weaker hydrogen bonding effect. The shift of major b_{NH} rotamer (0.8 ppm) is comparable to the shift of b_{NH} in β-ala-*aza*-PNA-T (1.03 ppm) suggesting similar intramolecular interaction if any in both cases.

2.3.7 Temperature dependent NMR shifts and NMR experiments

To further investigate the strength of the intramolecular interaction in β-ala-*aza*-PNA-T, temperature dependent NMR (heating in DMSO-d_6) was performed in the temperature range 298K-378K (Figure 2.34). Heating provides thermal energy and disrupts the H-bonding and consequently shifts the chemical shift of NH upfield in the ^{1}H NMR spectra. β-ala-*aza*-PNA-T shows single E-rotamer and upfield shift of both b_{NH} and a_{NH} protons were observed in the

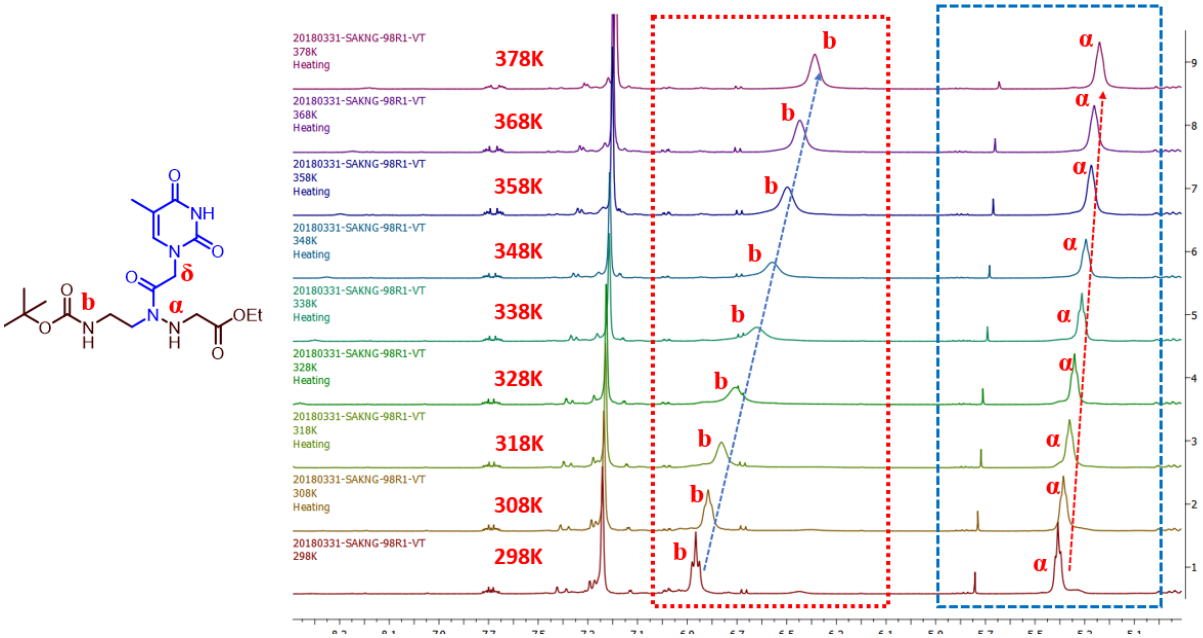


Figure 2.34. Variation of chemical shift of backbone -NH (α, a_{NH}) and BocNH (b_{NH}) in the VT-NMR experiment of *aza*-β-*ala*-PNA-T in DMSO-d_6 (Heating).

temperature range 298K-378K (Figure 2.34) except for small peak broadening seen at higher temperature. A small upfield chemical shift (0.15 ppm) of backbone α-NH (a_{NH}) along with a broadening of NMR signal was observed for β-ala-*aza*-PNA-T at higher temperature (Figure 2.34, blue dotted rectangle). In comparison, a much larger upfield shift (0.45 ppm) was seen for of b_{NH} (Figure 2.34, red dotted rectangle) suggesting that b_{NH} is more free than in a_{NH}.. The

temperature coefficient of -0.0056 ppm/^{o}C calculated for b_{NH} is three times of that for a_{NH}- 0.0018 ppm/^{o}C, which clearly indicated a much stronger intramolecular hydrogen bonding for a_{NH} compared to b_{NH} in DMSO-d_6.

The comparative variable temperature NMR of β-*ala*-PNA-T in DMSO-d_6 is shown in Figure 2.35. The upfield shift of b_{NH} was observed with increasing temperature along with peak broadening at higher temperature. The change of chemical shift and the calculated temperature coefficient for b_{NH} are 0.46 ppm and -0.0067 ppm/^{o}C respectively, which is similar to that in β-ala-*aza*-PNA-T suggesting similar temperature effects for b_{NH} in both monomers. Thus, a lower shift and lower temperature coefficient for a_{NH} in β-ala-*aza*-PNA-T hints at its involvement in intramolecular hydrogen bonding, while a higher temperature coefficient for b_{NH} in both β-ala-*aza*-PNA-T and β-*ala*-PNA-T suggests a very weak or absence of such intramolecular effects for b_{NH} in β-ala-*aza*-PNA-T and β-*ala*-PNA-T.

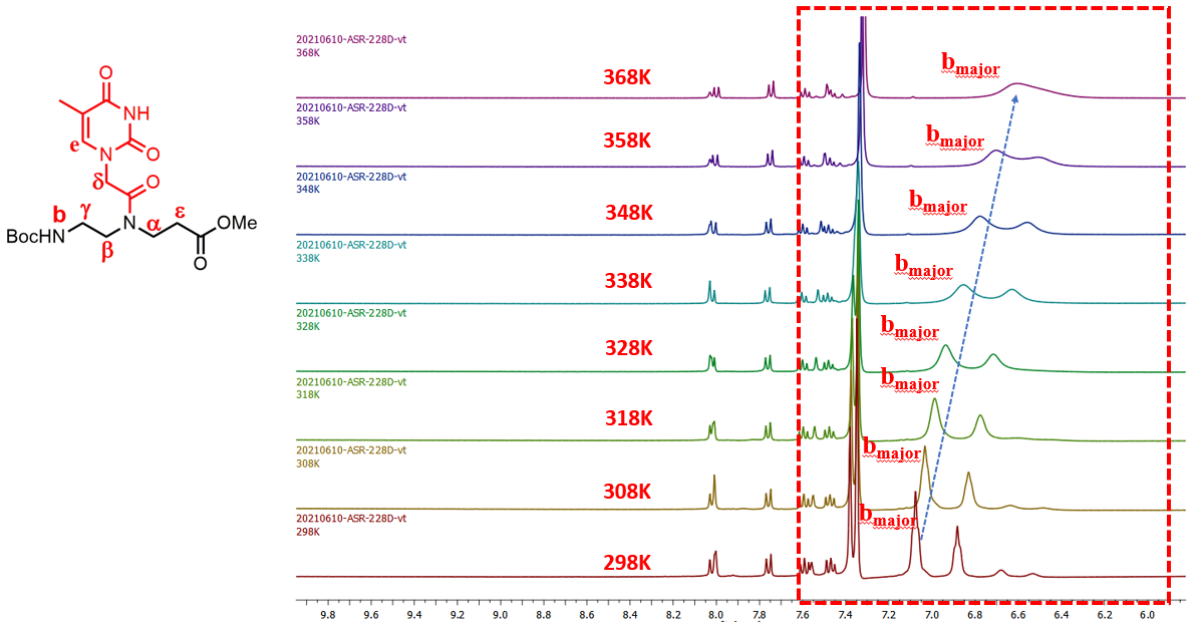


Figure 2.35. Variation of Chemical Shift of BocNH (b_{NH}) in the VT-NMR experiment of β-ala-PNA-T in DMSO-d_6 (Heating)

Figure 2.36 shows the variable temperature NMR of α-*aza*-PNA-T in DMSO-d_6 upon heating. Upfield shifts of ~ 0.46 ppm for both a_{NH} and b_{NH} along with peak broadening was seen at higher temperature. The temperature coefficient values for a_{NH} and b_{NH} are also similar ~ -0.0057 ppm/^{o}C and equal to that seen in b_{NH} of β-ala-*aza*-PNA-T in DMSO-d_6, suggesting weak intramolecular effects. It should be pointed out that different chemical nature of a_{NH} in β-ala-*aza*-PNA-T (secondary amine) and α-*aza*-PNA-T (carbamate -NH) along with the

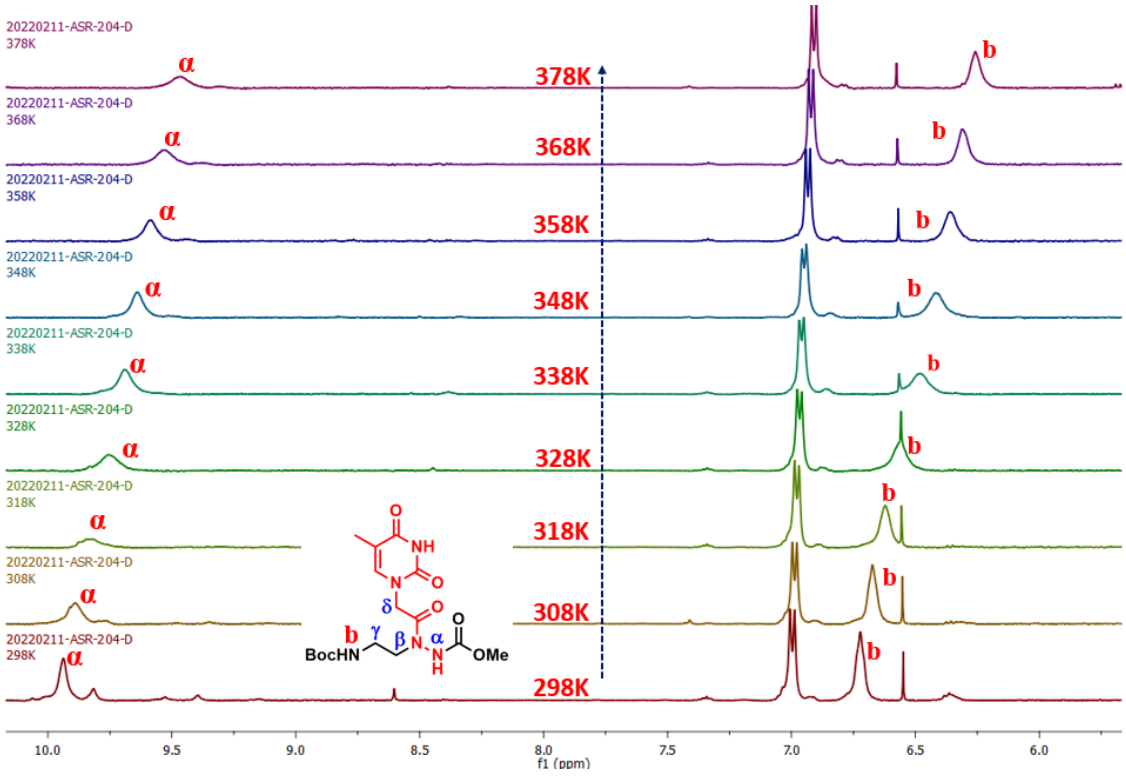


Figure 2.36. Variation of chemical chift of cackbone -NH (α, a_{NH}) and BocNH (b_{NH}) in the VT-NMR experiment of α-*aza*-PNA-T in DMSO-d_6 (heating).

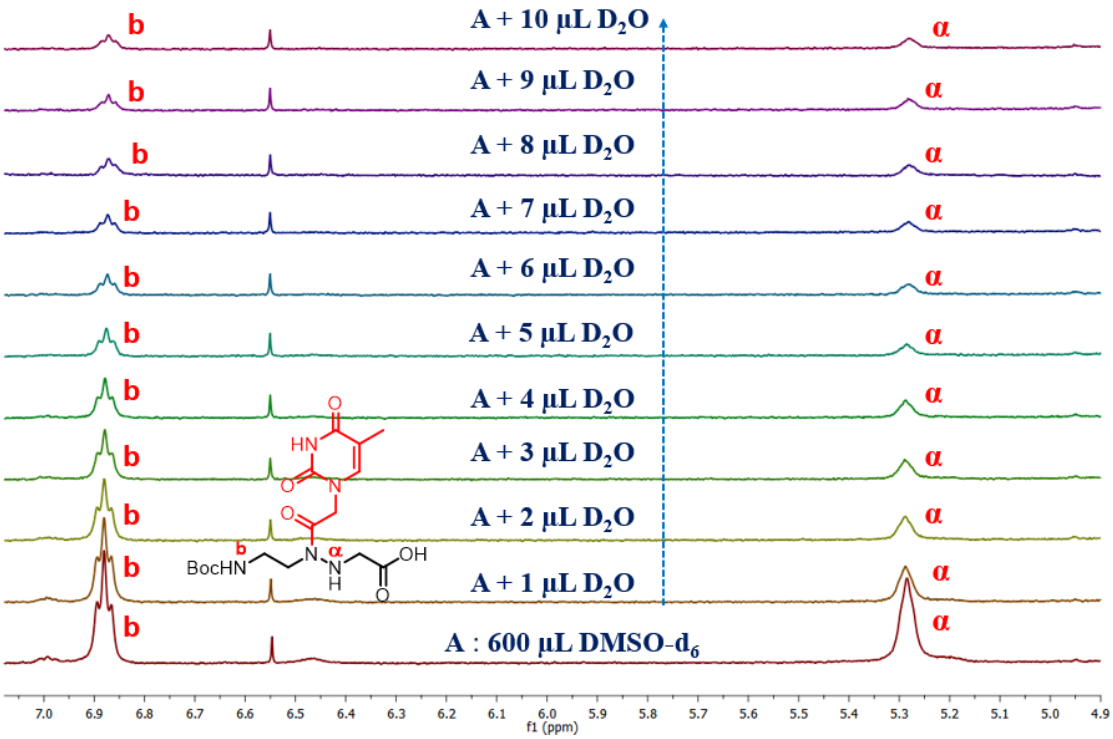


Figure 2.37. Variation of intensity of BocNH (b_{NH}) and backbone -NH (α, a_{NH}) in D_2O titration experiment of β-ala-*aza*-PNA-T (40 mM).

distinct rotameric distribution of the monomers in solution also may cause different behavior for both the monomers in VT-NMR.

H/D exchange reactions are also useful to determine the availability of NH protons for exchange with deuterium. Free NH protons undergo exchange with deuterium faster than hydrogen bonded NH which are less available for exchange with deuterium. Figure 2.37 shows the change in intensity of a_{NH} and b_{NH} protons of β-ala-*aza*-PNA-T with increasing amounts of added D_2O in the DMSO-d_6. The intensity of peaks of a_{NH} and b_{NH} slowly decreases upon addition of D_2O, with slight broadening and saturation of peaks after addition of 7 μL of D_2O and no significant changes in chemical shifts occurred with increasing proportion of D_2O. (Figure 2.37). Since secondary amines are better hydrogen bond acceptors than deuterated (ND) amine, only 0.01 ppm upfield shift of a_{NH} was seen after the addition of D_2O, with no changes in chemical shift of b_{NH}. The combined results of DMSO-titration, VT-NMR and D_2O exchange experiments suggested the participation of a_{NH} of β-ala-*aza*-PNA-T in intramolecular hydrogen bond.

The difference in chemical shifts of NH protons in $CDCl_3$ and DMSO can empirically be used to calculate the acidity of hydrogen bonded protons using the equation[35] $A = 0.0065+0.133\Delta\delta$, where $\Delta\delta$=difference of chemical shift of OH/NH_2 protons in $CDCl_3$ and DMSO-d_6. A value of A below or equal to 0.1, indicates involvement of OH in an intramolecular hydrogen bond and when it is equal or crosses 0.5 it suggests absence of intramolecular hydrogen bond. For NH protons, A value less than or equal to 0.05 indicates an intramolecular hydrogen bonded and if the A value greater than 0.16 suggests absence of intramolecular hydrogen bond. Table 1 shows the A value calculated for different NH protons of *aza*-monomers and control PNA monomers. From the table it is evident that for all b_{NH} protons in *deaza*-PNA monomers, the 'A' value is more than 0.16, suggesting their non-involvement in intramolecular hydrogen bond. Among the *aza*-monomers only a_{NH} of β-ala-*aza*-PNA-T shows an 'A' value of 0.04 strongly indicating possibility of intramolecular hydrogen bonding. The a_{NH} of α-*aza*-PNA T exhibited an 'A' value of 0.67 suggesting no intramolecular hydrogen bonding, however it shows rotamers and hence this is inconclusive. The b_{NH} of α-*aza*-PNA T and β-ala-*aza*-PNA-T shows 'A' value of 0.22 and 0.23 respectively, indicating its non-participation in intramolecular hydrogen bond. The literature precedence for hydrogen bond acidity (A) is more for possibility of 5 and 6-membered intramolecular hydrogen bond formation by -NH and -OH protons but not certain if it holds for larger (7/8-

membered) rings, and hence the correlation only provides a rough idea about the possibility of intramolecular hydrogen bond formation.

Calculation of Hydrogen bond acidity (A) value

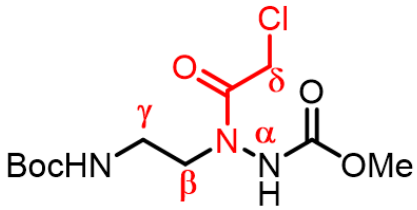
Compound	δ(ppm) $CDCl_3$	δ(ppm) DMSO-d_6	Δδ(ppm)	A Value	Intramolecular Hydrogen Bonding
aeg-PNA-T	5.61	6.94 (major rotamer)	1.33	0.18	No intramolecular Hydrogen bond
		6.74 (minor rotamer)	1.13	0.15	weak Intramolecular Hydrogen bond
α-*aza*-PNA-T	4.930(α-NH)	9.85 (α-NH)	5.02	0.67	No intramolecular Hydrogen bond
	4.930(BocNH)	6.72 (BocNH)	1.60	0.22	No intramolecular Hydrogen bond
β-ala-*aza*-PNA-T	5.12(α-NH)	5.42(α-NH)	0.3	0.04	Strong intramolecular Hydrogen bond
	5.14(BocNH)	6.88(BocNH)	1.74	0.23	No intramolecular Hydrogen bond
β-ala-PNA-T	5.37(b_{major})	7.05(b_{major})	1.68	0.22	No intramolecular Hydrogen bond
	5.10(b_{minor})	6.85(b_{minor})	1.75	0.23	No intramolecular Hydrogen bond

Table 1: Hydrogen bond acidity (A) values of -NH protons of *aza* and *deaza* monomer

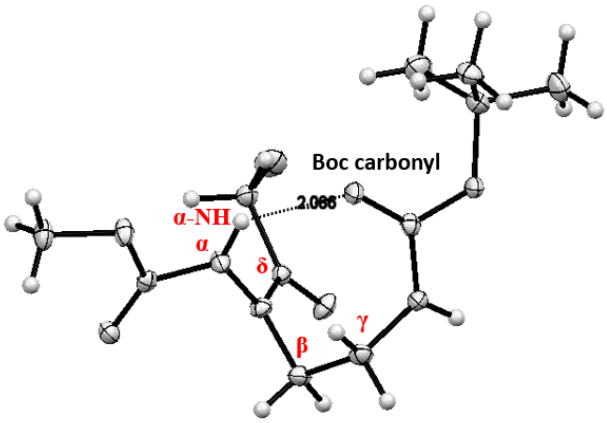
2.3.8 Crystal structure of the N-acetylchloro aminoethyl-α-*aza* glycine (7)

The N-acetylchloro aminoethyl-α-*aza* glycine (**7**) which is the precursor for synthesis of α-*aza*-PNA T ester crystallised from its methanol solution kept at 4°C. The homogeneous solution was filtrated two times by cotton filter and then transferred into glass vials, which were sealed by para film with a few small holes. Then the vials were placed into a ventilated fridge to evaporate the solvents slowly. The crystals were collected after 2 ~ 3 days and stored in dark. The single crystal sample was mounted on top of a cryoloop and transferred into the cold nitrogen stream (100 K) of a Bruker-AXS D8 Venture diffractometer. Data collection and reduction was performed using the Bruker software suite APEX.[36] The final unit cell was obtained from the xyz centroids of 9824 reflections after integration. A multiscan absorption correction was applied, based on the intensities of symmetry-related reflections measured at different angular settings (SADABS). The structures were solved by direct methods using SHELXT[37] and refinement of the structure was performed using SHLELXL.[38] The hydrogen atoms were generated by geometrical considerations.

The crystal structure exhibited a folded backbone with δ-CO in E-rotamer conformation (Figure 2.38). In the backbone folded structure, the N-terminal carbamate carbonyl and α-NH came in close proximity with 2.098 angstrom distance between them. This suggests the



Compound 7



Crystal structure of Compound 7
E-Rotamer
CCDC Number- 2150552

Figure 2.38. Molecular structure and crystal structure of N-acetylchloro aminoethyl-α-*aza*-glycine (**7**)

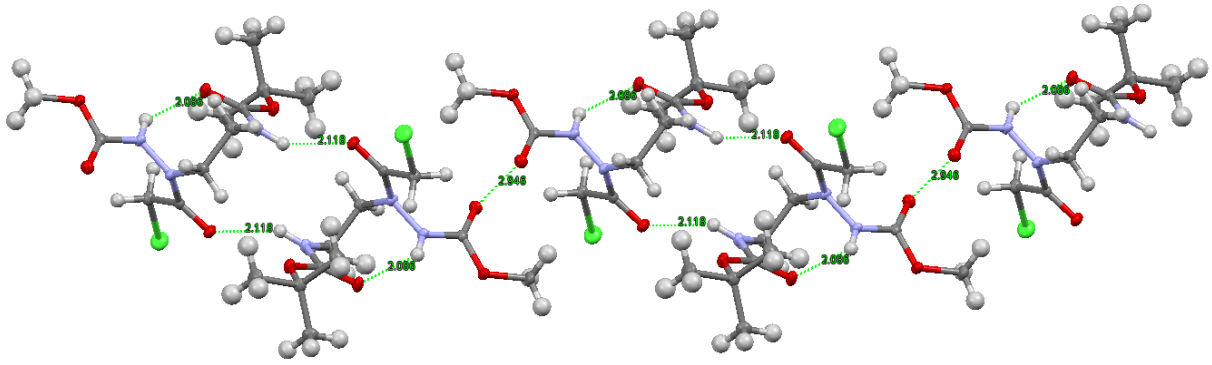


Figure 2.39. Packing diagram of the N-acetylchloro aminoethyl-α-*aza* glycine (**7**) in crystal

possibility of intramolecular H-bond through formation of an eight membered ring intramolecular hydrogen bonding between a_{NH} and N-terminal amide carbonyl group on PNA backbone. The 7-memebered intramolecular γ-turns are well known in literature[29] but 8-memebered rings are of lower propensity. The crystal packing structure (Figure 2.39) reveals dimeric units held through H- bonding (2.118 A) between the δ-CO and the carbamate NH in solid state which is again possible only if δ-CO assumes E-rotameric configuration, with the side chain δ-CH_2 oriented towards the *aza*-glycine side of the PNA backbone. The crystal structure thus indicates that α-NH in *aza*-PNAs is involved in intramolecular H-bonding with carbamate CO, while b_{NH} participates in intermolecular bonding with δ-CO of paired residue. The Cl atom does not seem to have any effect on the structure and perhaps the presence of T

instead of Cl also does not affect the structure. The H-bonding pattern differences between a_{NH} and b_{NH} seen in the crystal structure of N-(acetylchloro) aminoethyl-α-aza glycine **7** are also retained in *aza*-PNA T structures in solution as revealed by solvent and temperature dependent NMR experimental results. The single crystal structure data is deposited in the Cambridge Structural Database with CCDC number of the crystal 2150552. And the data is summarised in Table 2.2.

Crystal data and structure refinement of compound 7

Identification code	trictyl 1_a
Empirical formula	C_{11} H_{20} N_3 O_5 Cl
Formula weight	309.75
Crystal system, space group	triclinic, P-1
Refines used	9389
Measurement temperature	296 K
Unit cell dimension	a = 8.726 angstrom, α = 81.69 degree
	b = 9.362 angstrom, β = 68.40 degree
	c =10.33 angstrom, γ = 78.54 degree
Cell measurement theta max	28.34 degree
Cell measurement theta min	2.68 degree
Crystal description	Block crystal
Crystal color	Colorless
Crystal density	1.32
Crystal size	0.13
Absorption coefficient	0.271
Absorpt correction type	Multi scans
Absorpt correction T-min	0.7126
Absorpt correction T-max	0.7457
Absorpt process details	SADBAS
Radiation wavelength and type	0.7107 angstrom and MoK\a
Diffraction source	Fine focused seal tube
Diffraction measurement device	Bruker APEX-II CCD

Diffraction method	Omega and Phi Scan
Diffraction reflection number	12911
Diffraction reflection h min	-11
Diffraction reflection h max	11
Diffraction reflection l min	-10
Diffraction reflection l max	13
Diffraction reflection k min	-12
Diffraction reflection k min	12
Diffraction reflection theta min	2.226
Diffraction reflection theta min	28.35
Total reflection number	3797
Computing data collection	Bruker Apex-II
Computing cell refinement	Bruker SAINT
Computing data reduction	Bruker SAINT
Computing structure solution	SHELXT 2014/5 (sheldrick, 2014)
Computing data refinement	SHELX-L 2018/3 (sheldrick, 2018)
Computing molecular graphics	Bruker SHELXTL
Computing publication material	Bruker SHELXTL
Refine ls extinction method	SHELX-L 2018/3 (sheldrick, 2018)
Refine ls extinction coefficient	0.11 (2)
Refine ls R factor all	0.0398
Refine ls R factor gt	0.0354

2.3.9 Infrared spectroscopy data of monomers

Infrared spectroscopy is a very important tool for organic chemists to study the secondary interactions in molecules, especially intramolecular and intermolecular hydrogen bonding.[39] Typical IR stretching frequency[39] for an intramolecular hydrogen bonded amine is in the range 3200 cm^{-1}- 3400 cm^{-1}. Non hydrogen bonded amines[39] exhibit IR stretching frequency in the range 3400 cm^{1} – 3600 cm^{-1}. Figure 2.40 shows the enlarged IR spectra of β-ala-*aza*-PNA-T and β-*ala*-PNA-T in $CDCl_3$ and DMSO-d_6. β-ala-PNA-T shows very weak IR stretching at 3450 cm^{-1} for N-terminal amide NH (b_{NH}) suggesting no intramolecular hydrogen bond formation for β-*ala*-PNA-T in organic solvents. In comparison, IR spectra of β-ala-*aza*-

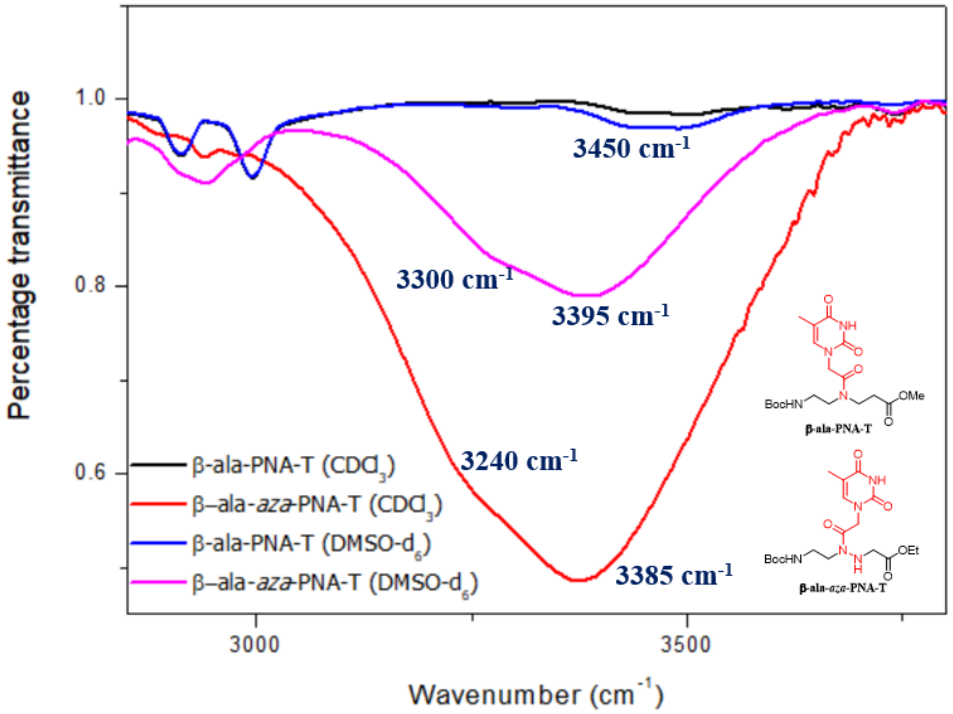


Figure 2.40. IR spectra of β-ala-*aza*-PNA-T and β-ala-PNA-T in $CDCl_3$ and DMSO-d_6

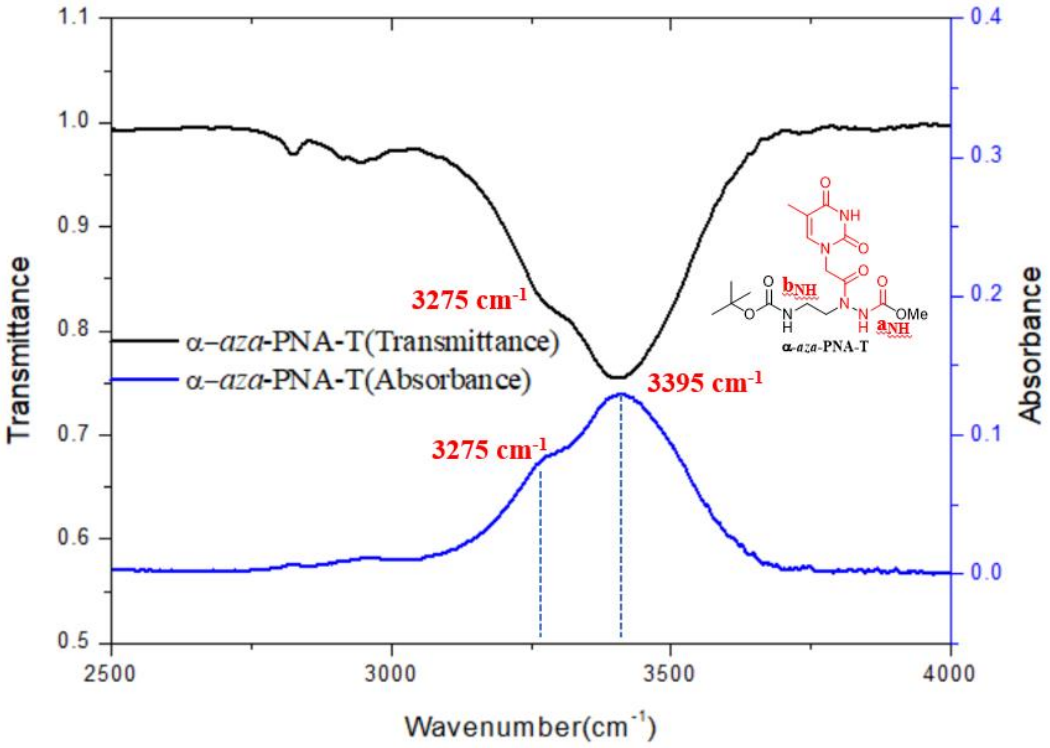


Figure 2.41. IR spectra of α-*aza*-PNA-T in $CDCl_3$

PNA-T in $CDCl_3$ and DMSO-d_6 exhibits a broader IR peak in $CDCl_3$ compared to DMSO-d_6. The broad IR spectra of β-ala-*aza*-PNA-T arises due to the presence of both hydrogen bonded (a_{NH}) and non-hydrogen bonded amines (b_{NH}). a_{NH} (α-NH) of β-ala-*aza*-PNA-T participates in eight membered intramolecular hydrogen bonding with amide carbonyl of N-terminal amide resulting in IR stretching frequency at 3300 cm^{-1} and 3395 cm^{-1} in $CDCl_3$. In DMSO-d_6 the 3395 cm^{-1} band shifted to 3385 cm^{-1}. IR stretching frequency after 3400 cm^{-1} suggests the presence of non-hydrogen bonded N-terminal amide NH (b_{NH}) group in β-ala-*aza*-PNA-T monomer. Figure 2.41 shows the enlarged IR spectra of the α-*aza*-PNA-T in $CDCl_3$ in both Absorbance and transmittance unit. Presence of a strong peak at 3395 cm^{-1} and 3275 cm^{-1} suggests eight membered intramolecular hydrogen bonding of a_{NH} with amide carbonyl of N-terminal amide group.

Secondary amides show characteristics Amide-I, Amide-II and Amide-III bands respectively.[40] Amide-I bands are combination of C=O stretching, C-N out of phase stretching and C_α-CN deformation bands ranging from 1620 cm^{-1} to 1710 cm^{-1}. Amide-II bands are combination of N-H in plane bending, C-N stretching and C=O in plane bending ranging from 1500 cm^{-1} to 1600 cm^{-1}. Amide-III bands are combination of N-H in plane bending and C-N stretching ranging from 1200 cm^{-1} to 1400 cm^{-1}. Amide-III bands above 1290 cm^{-1} are considered as characteristic band for turn motif (β/γ-turn) in peptides.[40] α-helix and β-sheet structures lacks IR Amide-III bands[41] above 1290 cm^{-1}. Hydrogen bonding in amides affects the Amide-I band of amides and shifts the band to lower wavenumber, non-hydrogen bonded amides or in the dilute samples where intermolecular hydrogen bonding is minimal shows C=O stretching at higher wavenumber. Intramolecular hydrogen bonding reduces the amide-I stretching frequency of the carbonyl groups and shifts in the lower wavenumber region. β-*ala*-PNA-T lacks intramolecular hydrogen bonding and hence shows amide-I bands at 1737 cm^{-1} (See appendix). β-ala-*aza*-PNA-T shows amide-I band at 1672 cm^{-1} suggesting an intramolecular hydrogen bonding, Presence of N-H stretching band at 3296 cm^{-1} further confirms the observation as hydrogen bonded -NH stretching frequency appears in the range 3200-3400 cm^{-1}. Amide-II band of β-ala-*aza*-PNA-T appears at 1516 cm^{-1}. Presence of amide-III band at 1370 cm^{-1} and 1245 cm^{-1} confirms γ-turn type folded structure in β-*ala*-*aza*-PNA-T.[40] a_{NH} protected (Benzyl protected) monomer of β-ala-*aza*-PNA-T (Compound **14a** and **14b**) shows C=O stretching at 1676 cm^{-1} but lacks -NH stretching band at 3300 cm^{-1} suggesting a non-intramolecular hydrogen bonded structure of the monomer (See appendix) in solution. IR spectra of α-*aza*-PNA-T in $CDCl_3$ shows C=O stretching frequency at 1650 cm^{-1} suggesting

hydrogen bonded carbonyl group in α-*aza*-PNA-T. Presence of -NH stretching peak at 3395 cm^{-1} and 3275 cm^{-1} further support the observation and suggests an intra/intermolecular hydrogen bonding in α-*aza*-PNA-T monomer (See appendix). Presence of amide-II band at 1250 cm^{-1} and 1380 cm^{-1} hints a folded γ-turn structure for α-*aza*-PNA-T in $CDCl_3$. Hence IR spectra suggest a hydrogen bonded, folded γ-turn structure for *aza*-PNA monomers in solution.

2.3.10 Computational studies of *aza*-PNA rotamer structure (in Collaboration with Dr Raghunath Ramabhadran, IISER Tirupati)

Structures of E and Z rotamers of *aza*-PNA monomers were optimized by using B3-LYP density function in combination with 6-31+G(d,p) basis-sets[41] by using the SMD implicit solvation method (DMSO method).[42] During the optimization D3-BJ dispersion corrections were introduced at each point. Much larger 6-311++G(3df,2p) basis sets were used to perform the single point calculations and to confirm that the energetics would remain same as observed with 6-31+G(d,p) basis-sets. Intramolecular weak interactions were studied by established NBO method.[43] To verify the calculations M06-2x and w97X-D methods were used. The optimized structures of *aza*-PNA monomers **1** and **2** along with the *deaza* analogue of **1** (*aeg*-PNA, Compound **3**) and **2** (β-ala-PNA-T, Compound **4**) are shown in Figure 2.42 along with the expanded region in Figure 2.43. The computationally optimized structures of different *aeg*/*aza*-PNA monomers are shown in Figure 2.42: α-*aza*-PNA-T (Figure 2.42a), β-ala-*aza*-PNA-T (Figure 2.42b) and the corresponding *deaza* analogues *aeg*-PNA-T (Figure 2.42c) and β-*ala*-PNA-T (Figure 2.42d). Notable features of the optimized structures of α-*aza*-PNA-T and β-ala-*aza*-PNA-T are the presence of eight membered hydrogen bonding between the $N_\alpha H$ and (H_A) and backbone amide carbonyl O_B (N_α-H…O_B) along with the bonding between C_β-H (H_β) with t-amide carbonyl δ-CO (C_β-H_β…O_A). The selected expanded regions in the figure 2.43 (2.43a and 2.43b) shows the interactions. The calculated distance between the H_A and O_B is 1.95 angstrom with bond angle 148-163^o suggesting a very strong intramolecular hydrogen bonding between the $N_{\boldsymbol{\alpha}}H$ (H_A) and the backbone amide carbonyl O_B ($N_{\boldsymbol{\alpha}}$-H_A⋯O_B).

Interestingly no intramolecular or intermolecular hydrogen bonding features were observed for $N_{\boldsymbol{\alpha}}H$ with either δ-CO (O_A) or ester CO (O_C). The formation of extended 8 membered $N_{\boldsymbol{\alpha}}$-H_A⋯O_B γ-turn type intramolecular hydrogen bond in *aza*-PNA monomers (Figure 2.42a and 2.42b) compacts the *aza*-PNA backbone compared to the *deaza* analogues *aeg*-PNA-T (Figure 2.42c) and β-ala-PNA-T (Figure 2.42d). Interestingly, *deaza* PNA monomers *aeg*-PNA-T (Figure 2.42c) and β-*ala*-PNA-T (Figure 2.42d) lacks the $N_{\boldsymbol{\alpha}}H$ group and hence no intra/intermolecular hydrogen bonding were observed for the *deaza* PNA

monomers (Figure 2.42c and 2.42d). These results suggest that a_{NH} (N_{α}-H_A) in *aza*-PNA monomers are primarily responsible for the preferential stabilization of rotamers in solution. Significant reduction of dihedral angles Ψ (N_t-N_{α}) and Φ (C_{β}-N_t) was observed in α-*aza*-PNA-T (17^o and 7^o) and β-ala-*aza*-PNA-T (27^o and 25^o) respectively, compared to the corresponding *deaza* analogues *aeg*-PNA-T (Figure 2.42c) and β-*ala*-PNA-T (Figure 2.42d). Table 2.3

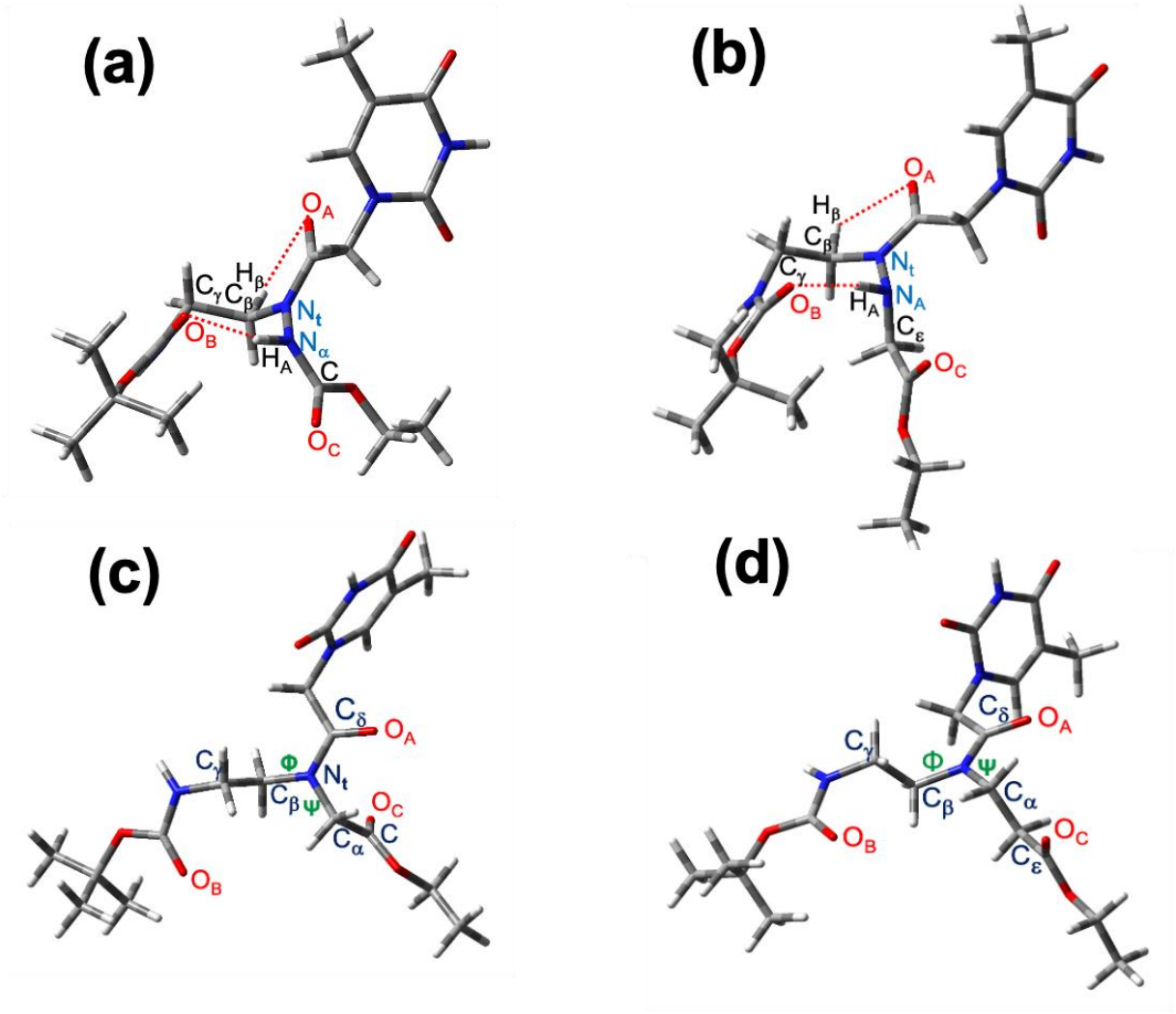


Figure 2.42. DFT optimised structures of (a) α-*aza*-PNA **1** (b) *aza*-β-*ala* PNA **2** (c) *aeg*-PNA **3** (d) β-*ala*-PNA **4**, with H-bond lengths and selected dihedral angles

summarises the values of calculated Ψ (N_t-N_{α}) and Φ (C_{β}-N_t) dihedral angles, bond angles and values of bond lengths. anticlockwise rotations around N_t-C_{α} and N_t-C_{β} bonds driven by $N_{\alpha}H$ triggers the folding of the *aza*-PNA backbone in a γ-turn fashion so that the N-terminal carbamate CO (O_B) comes closer to the $N_{\alpha}H$ for H-bond formation. This γ-turn induced folding of the *aza*-PNA backbone reorients the rotamers to E-Form (Figure 2.43a and 2.43b). Another notable feature of the optimized structure is the C_{β}-H···O_A interaction in *aza*-PNA monomers.

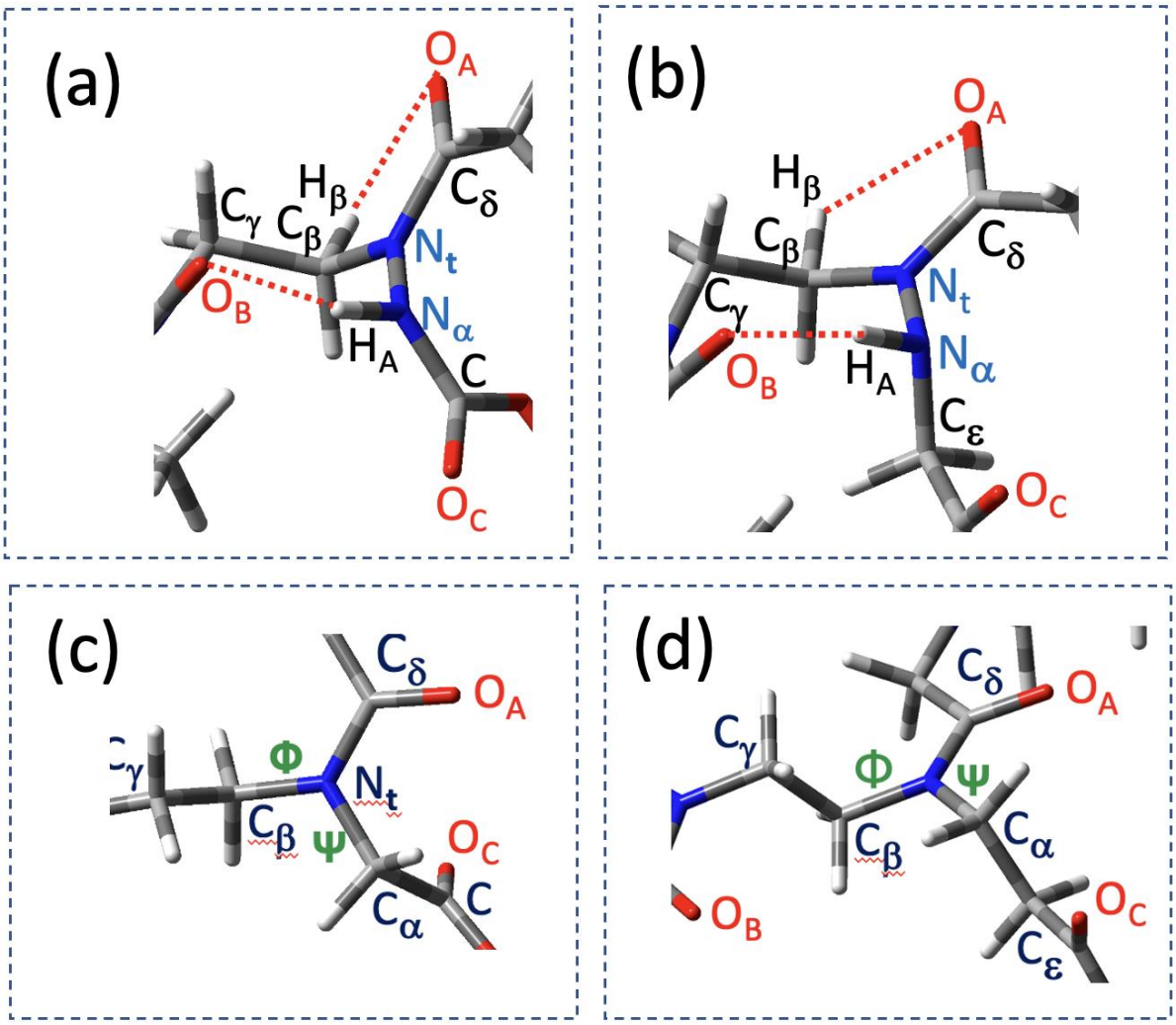


Figure 2.43. Expanded regions of computed structures of *aza*-PNA (a,b) and *deaza*-PNA monomers (c,d) showing H-bonded connections

Dihedral and bond angles (°) in *aza/aeg*-PNAs*

PNA	Ψ	Φ	Bond angles		Bond lengths	
			H-bond	CHO	$N_\alpha H_A \cdots O_B$	C_β-H$\cdots O_A$
1 (E)	79.4	79.0	148	100	1.95Å	2.04 Å
2 (E)	62.4	60.8	163	106	1.91 Å	1.95 Å

* Ψ (C_β-N_t-N_α/C_α-C); Φ (N_α-N_t-C_β-C_γ); **1**(Z), Ψ (96.5°), Φ (86.4°); **2**(Z), Ψ (89.6°), Φ (86.3°).

Table 2.3. The values of calculated Ψ (N_t-N_α) and Φ (C_β-N_t) dihedral angles, Bond angles and values of bond lengths

Folding of the backbone in a γ-turn like orientation in *aza*-PNA brings the C_β-H closer to the O_A, which is confirmed by the observed NOSEY cross peak of N_α-H_A with the β/γ-backbone methylene protons. The C_β-H and O_A distances in *aza*-PNAs are 2.04 angstrom in α-*aza*-PNA-T and 1.95 angstrom in β-ala-*aza*-PNA-T respectively, which favors the C_β-H···O_A interaction in *aza*-PNA monomers. The values of the computed NBO interaction energies are significant for the preferential stabilization of E-Rotamer in *aza*-PNAs.

In peptide structures, the five membered hydrogen bonding between amide NH and carbonyls are abundantly known in literature. Raines *et. al.* in 2016 reported the intraresidual five membered hydrogen bonding between the amide -NH and carbonyl group that stabilizes the primary and secondary structures of proteins.[39] Formation of five and four membered hydrogen bonding is possible in β-ala-*aza*-PNA-T and α-*aza*-PNA-T. Figure 2.44 shows the possible intra-residual five and four membered hydrogen bonding in *aza*-PNA monomers. To prove the importance and uniqueness of gamma-turn structure of *aza*-PNA it is necessary to rule out the alternative hydrogen bonded structures of the monomers through DFT calculation. Formation of 4 membered hydrogen bonding in the Z-Rotamer of α-*aza*-PNA-T between α-NH (a_{NH}) and ester carbonyl group (Figure 2.44b), 5- membered hydrogen bonding between α-NH (a_{NH}) and tertiary amide carbonyl (Figure 2.44c) is possible in solution. These and the

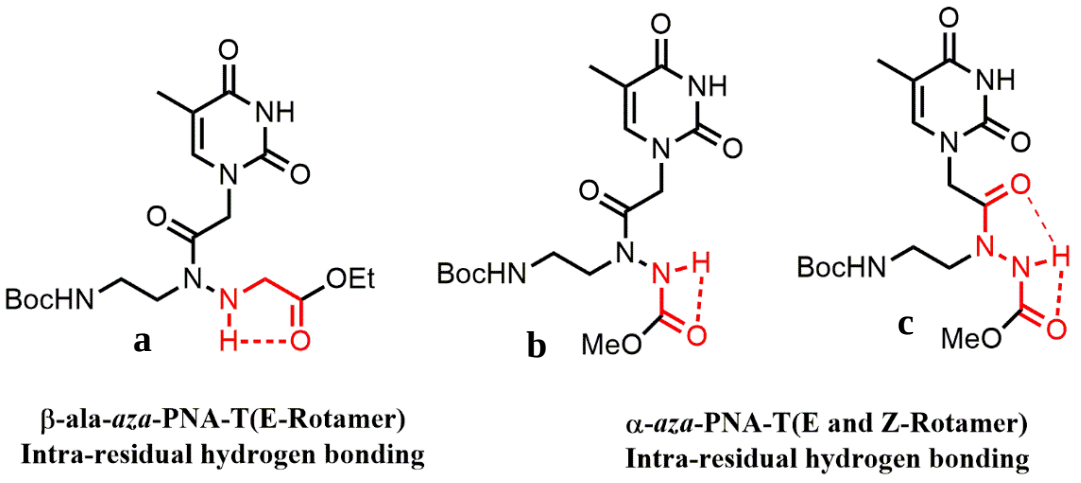


Figure 2.44. Possible intra-residual hydrogen bonding pattern in *aza*-PNA monomers

Fully optimized zig-zag conformer of the α-*aza* PNA monomer were scanned along the NH•••O distance by performing a relaxed scan calculation (Gaussian Keyword –opt= modreduandant).Figure 2.45 shows the complete scan profiles. The scanning shows lowering of energy at NH•••O distance 2.32 Å. Starting with this minimum point fully optimised structure of the α-*aza*-PNA-T provides a novel structure where the carbonyl oxygen from the

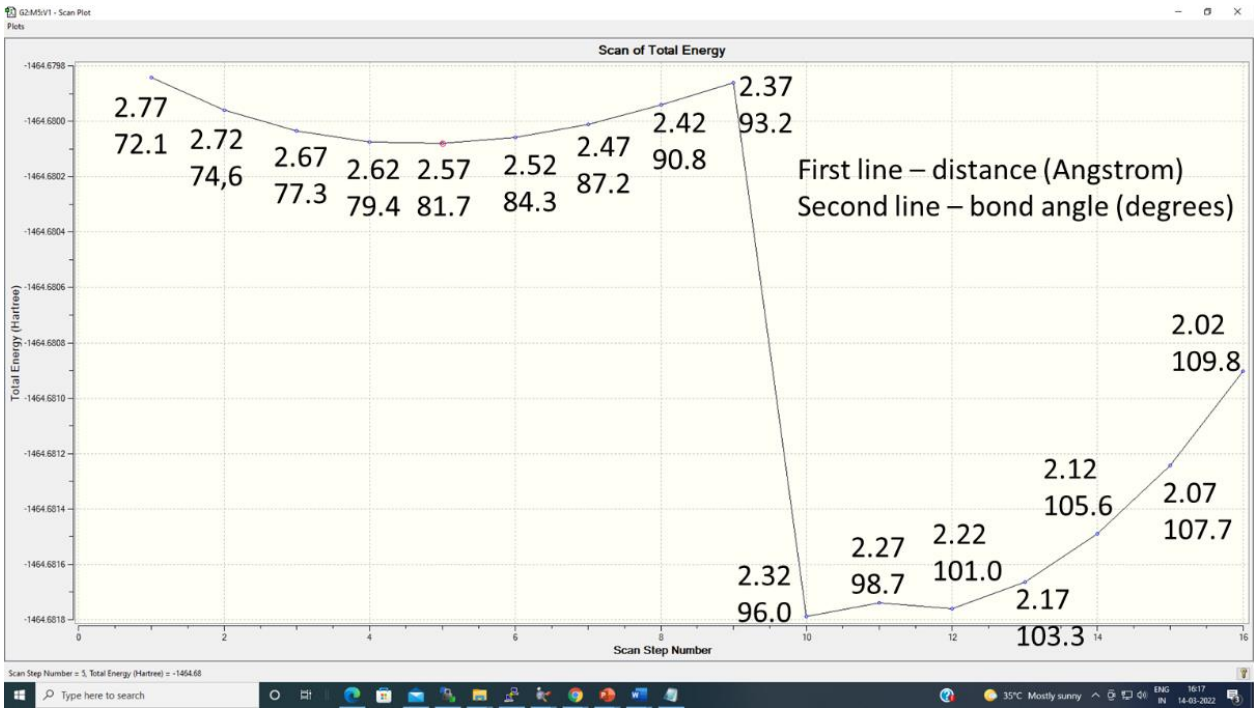


Figure 2.45. Scanning along α-NH•••O (from t-amide CO) distance starting with the Z-zigzag conformer of α-*aza* PNA-T monomer. The bond angle reported in the scan is the α-NH•••O (from t-amide CO) angle

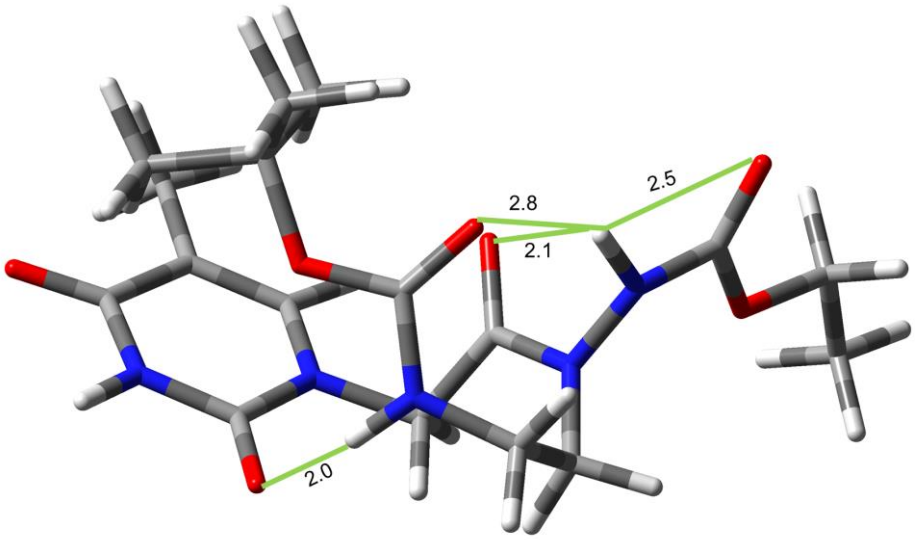


Figure 2.46. Computationally obtained alternate hydrogen bonded structure of α-*aza*-PNA-T

base hydrogen bonds with the NH of N-terminal NHBoc. This kind of experimentally established structure or computationally established structure was never reported in the PNA literature. The obtained structure completely opposed the experimental solution state structure

of α-*aza*-PNA-T as the computationally obtained structure predicts preferential Z-rotamer stabilization for α-*aza*-PNA-T, although α-*aza*-PNA-T shows 55:45 ratio of Z and E rotamer in $CDCl_3$. Hence the alternative patterns of hydrogen bonding are ruled out for α-*aza*-PNA-T monomer. Figure 2.45 and 2.46 shows the scan profile and the computationally obtained alternative hydrogen bonded structure.

Alternative hydrogen bonded structures of β-*ala*-*aza*-PNA-T is shown in the Figure 2.44a. In the most stable conformer of β-*ala*-*aza*-PNA-T the α-NH (a_{NH}) lacks hydrogen bonding with the COOEt ester carbonyl. To computationally validate the obtained lower energy structure and possibility of alternate hydrogen bonded structure a relaxed scan based on the distance between α-NH and carbonyl O of COOEt was performed. Figure 2.47 shows the scanning profile of the β-*ala*-*aza*-PNA-T monomer (Gaussian Keyword – Opt = modreduandant).

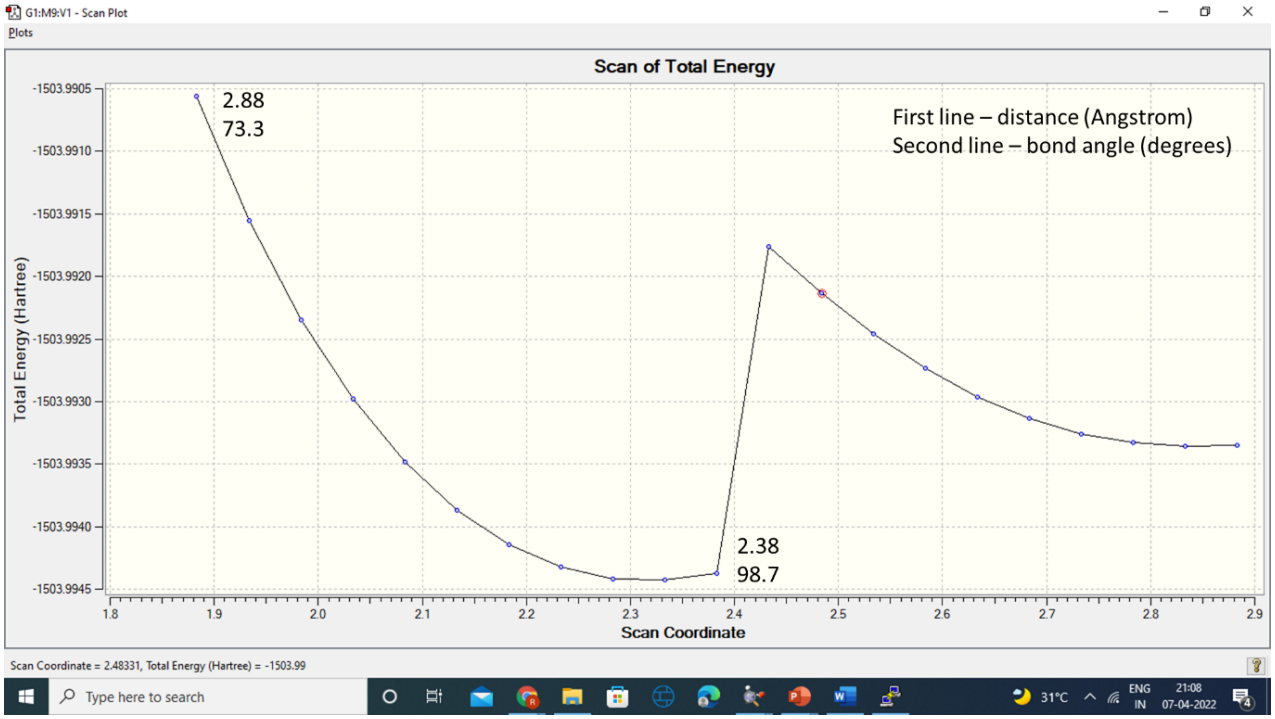


Figure 2.47. Scanning along α-NH•••O (carbonyl oxygen from COOEt) distance starting with the β-*ala*-*aza* conformer of β-ala-*aza* PNA monomer. Only selected distances and bond angles are mentioned. The bond angle reported in the scan is the α-NH•••O (carbonyl oxygen from COOEt) angle.

Starting with the lowest scan point having lowest distance between the α-NH and carbonyl of COOEt (2.38 Å) results an unusual structure of PNA which was never reported in the PNA literature. In the computationally obtained structure CH group from the base

(Thymine) hydrogen bonds with the O from CO of N-terminal NHBoc group. This structure has no literature reference, and the obtained structure opposed the experimental findings regarding the solution state structure of β-*ala*-*aza*-PNA-T in $CDCl_3$ and DMSO-d_6. Hence such structure with alternative hydrogen bonding between α-NH and carbonyl of COOEt group is ruled out as it contradicts the experimental findings regarding the solution state structure of β-*ala*-*aza*-PNA-T and shows an unusual hydrogen bonding between the nucleobase -CH (Thymine) and carbonyl of N-terminal BocNH group. Figure 2.48 shows the alternative hydrogen bonded structure. Notable feature of the optimized folded (γ-turn like) structure of *aza*-PNA monomers (Figure 2.42 and 2.43) is the favourable bonding between $C_{\beta}H$ (H_{β}) with t-amide δ-CO (C_{β}-H_{β}•••O_A). Interaction between the $C_{\beta}H$ (H_{β}) with t-amide δ-CO (C_{β}-H_{β}•••O_A) results more electrophilic carbonyl carbon of t-amide δ-CO compared to the *deaza* analogue.

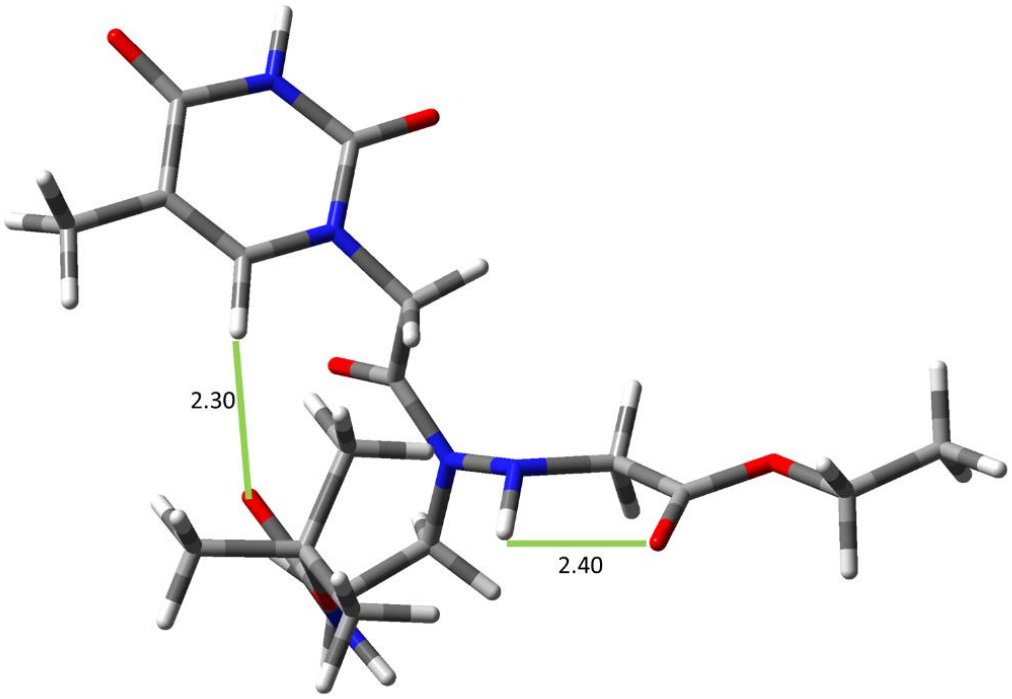


Figure 2.48. The structure obtained by starting with the lowest energy point in the scan (Figure 2.43). It features a carbonyl oxygen from NHBoc hydrogen bonding with a CH group from the base which disagree with the NMR results.

A comparison of ^{13}C NMR spectra of the *aza-monomers* with *deaza-monomers* is expected to exhibit more deshielded t-amide δ-CO for *aza-monomers* compared to the *deaza-monomers.* Figure 2.49 shows the comparison of carbonyl ^{13}CNMR chemical shift of β-*ala*-*aza*-PNA-T with *deaza* analogue β-*ala*-PNA-T. Comparison of spectra reveals more

deshielded t-amide δ-CO for β-*ala*-*aza*-PNA-T compared to the *deaza* analogue β-*ala*-PNA-T in DMSO-d_6. Except t-amide δ-CO all the carbonyl shows upfield shift compared to the carbonyls of β-*ala*-PNA-T. The carbonyls of β-*ala*-PNA-T exhibit double peaks for individual carbonyls suggesting rotamer mixture but the carbonyls of β-*ala*-*aza*-PNA-T shows single peak for individual carbonyls suggesting single rotamer stabilization in solution. This results directly support the Interaction between the $C_\beta H$ (H_β) with t-amide δ-CO (C_β-H_β•••O_A) in folded structure and supports γ-turn mediated rotamer stabilization in β-*ala*-*aza*-PNA-T.

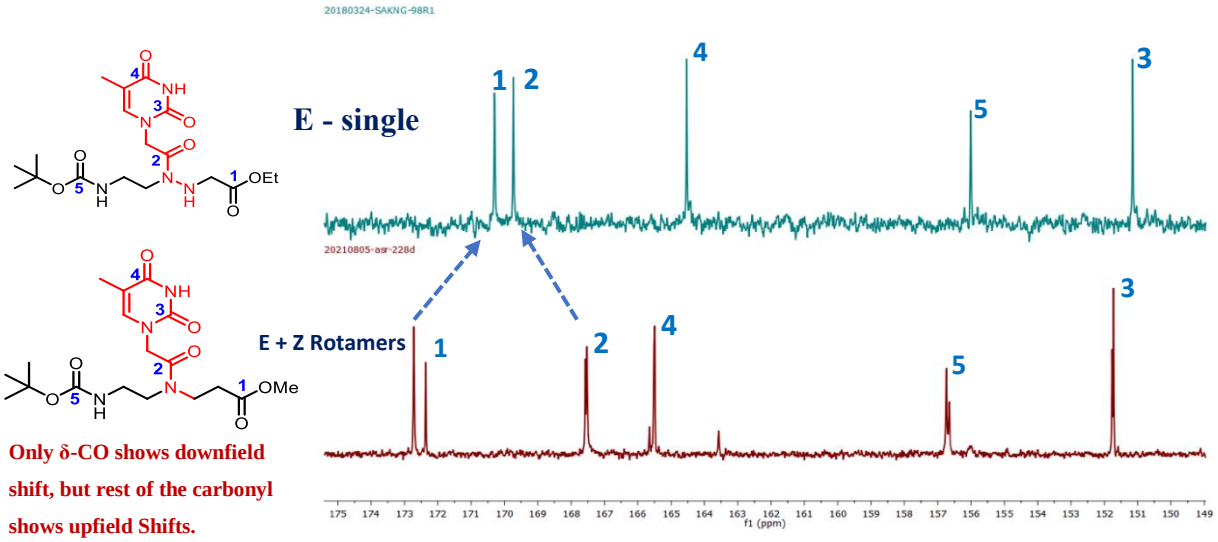


Figure 2.49. Comparison of carbonyl chemical shifts of β-*ala*-PNA-T and β-*ala*-*aza*-PNA-T in DMSO-d_6

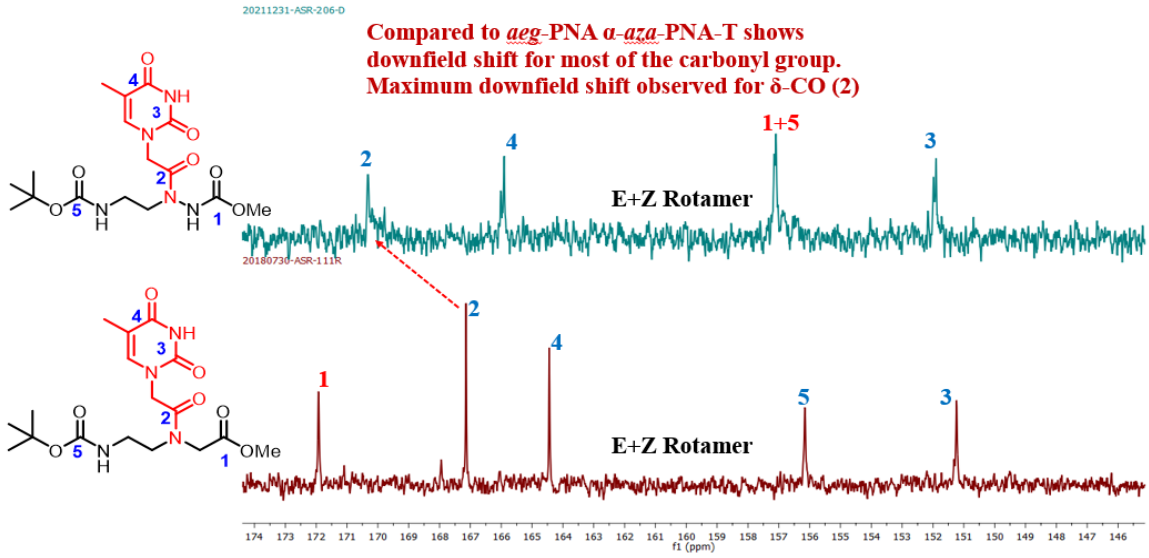


Figure 2.50. Comparison of carbonyl chemical shift of α-*aza*-PNA-T and *aeg*-PNA-T in $CDCl_3$

Figure 2.50 shows the comparison of carbonyl ^{13}CNMR chemical shift of α-*aza*-PNA-T and *aeg*-PNA-T. The comparison was performed in $CDCl_3$ as the α-*aza*-PNA-T monomers loss the folded structure in DMSO-d_6. Comparison of carbonyl ^{13}CNMR chemical shift reveals more deshielded δ-CO for α-*aza*-PNA-T compared to the *deaza* monomer *aeg*-PNA-T.

Interestingly, each carbonyl peak of α-*aza*-PNA-T shows double peak due to presence of E and Z rotamer is solution. In contrary all the carbonyl peaks of *aeg*-PNA-T shows single peak suggesting accidental match of NMR chemical shift for E and Z rotamer for *aeg*-PNA-T. All the carbonyls of α-*aza*-PNA-T except C-terminal hydrazamido carbonyl exhibit down field shift compared to the *aeg*-PNA-T. These results directly support the Interaction between the $C_{\beta}H$ (H_{β}) with t-amide δ-CO (C_{β}-H_{β}•••O_A) in folded structure and supports γ-turn mediated rotamer stabilization in the α-*aza*-PNA-T monomer. Hence from this comparison it is concluded that in both α-*aza*-PNA-T and β-*ala*-*aza*-PNA-T monomer the γ-turn formation stabilizes the preferred rotamers.

2.3.11 Examples of *aza*-modification in biomolecules

In *aza*-peptides and peptoids the α-NH is known to participate in the hydrogen bonding with the carbonyl group of peptide bond to induce the β-sheet structures.[25] Incorporation of *aza*-Amino acids in the β-turn forming peptide sequence significantly stabilizes the β-turn secondary structures.[25] Single *aza*-amino acid substitution in the β-hairpin forming peptide shows destabilization of the β-hairpin compared to the original peptide.[24] Side chain *aza* modifications of *aza*-peptoids dimer shows intramolecular hydrogen bonding with peptoid carbonyl group and drives preferential stabilization of Trans amide isomer (100%) confirmed by NMR and computational studies.[26] Literature suggests that replacement of glycine CH_2 in collagen peptide by isosteric NH group (termed as *aza*-Collagen) retains all the properties of original collagen peptide sequence *pro-pro(hydroxyproline)-gly* and induce hyperstability for the collagen triple helical structure.[27] Improved thermal stability, synergic stabilization of the collagen triple helix and self-assembly of the fully modified 12-mer collagen mimetic peptides (CMP) into a defined protein tertiary structure is the keynote properties of the *aza*-glycine modified CMPs.[27] Replacing the middle proline of collagen mimetic peptide sequence (*pro-pro-gly*) by *aza-proline* shows extreme effect on the self-assembly properties of CMPs.[28]

γ-turn: γ-turn[29] is a classical turn motif in the peptides where the CO residue of i and NH residue of (i+2) undergoes intramolecular hydrogen bonding to fold the peptide backbone. phi

and psi dihedral angles of the (i+1) residue in a classical γ turn are (70, -60) but for an inverse γ-turn[29] are (-70, 60) respectively.

Effect of urea linkage and polyamide linkage on the hybridisation of PNA with complementary DNA: Derived PNA oligomer of α-*aza*-PNA would results a repeating urea linkage in the oligomer (Figure 2.16) which could rigidify the oligomer and consequently the hybridisation efficiency of the α-*aza*-PNA will reduce significantly, making it inappropriate for hybridization with complementary DNA/RNA. To overcome the difficulty of hybridisation in α-*aza*-PNA with cDNA/cRNA to form complexes, homologue of α-*aza*-PNA was synthesized where *aza*-glycine side of α-*aza*-PNA is extended by one methylene group (ε-CH_2, β-ala-*aza*-PNA, Figure 2.14) between α-NH and CO and the resulting monomer named as β-*ala*-*aza*-PNA (Figure 2.14). Oligomer derived from β-ala-*aza*-PNA would results a polyamide backbone (Figure 2.16) like *aeg*-PNA, which makes it an appropriate candidate for hybridisation with cDNA/cRNA. To understand the role of α-NH in β-ala-*aza*-PNA in an effective way, control monomer of β-ala-*aza*-PNA was synthesized where the α-NH is replaced by isosteric methylene group (β-ala-PNA, Figure 2.14) named as β-*ala*-PNA (Figure 2.14)

2.4. Conclusion

Rationally designed β-ala-*aza*-PNA-T and α-*aza*-PNA-T monomers were synthesized where the α-CH_2 group is replaced by isosteric -NH group. To avoid the repeating urea linkage in the derived oligomer of α-*aza*-PNA-T, *aza*-glycine side of α-*aza*-PNA-T is extended by one methylene (ε-CH_2) to obtain the β-ala-*aza*-PNA-T monomer. The primary ^{1}H NMR spectral analysis of the synthesized monomers reveals a mixture of rotamers for α-*aza*-PNA-T but interestingly preferential stabilization of a single rotamer was observed for β-ala-*aza*-PNA-T. Further investigation by ^{1}H-^{1}H NOSEY, TOCSY and Gradient COSY shows a solvent independent E- rotamer stabilization of β-ala-*aza*-PNA-T in $CDCl_3$ (100% E-Rotamer) and DMSO-d_6 (100% E-Rotamer). In contrary, α-*aza*-PNA-T exhibits a solvent dependent rotamer stabilization in $CDCl_3$ (major Z, minor E) and DMSO-d_6 (major E, minor Z). The different effect of solvation on the *aza*-PNA monomers is attributed due to the different chemical nature of a_{NH} in the *aza*-PNA monomers. In α-*aza*-PNA-T a_{NH} is an active carbamate NH with higher acidity compared to a secondary amine but in the β-ala-*aza*-PNA-T the a_{NH} is a secondary amine with less acidity than an active carbamate NH. The observation of a mixture of rotamers in β-ala-PNA-T and β-ala-*aza*(N^{Bz})-PNA-T (a_{NH} is protected by benzyl group) indicates a major role of α-NH (a_{NH}) for the stabilization of single E-rotamer in β-ala-*aza*-PNA-T. NMR

experiments (DMSO titration, variable temperature NMR and D_2O titration) and solvent dependent NMR analysis of β-ala-*aza*-PNA-T suggests a strong intramolecular hydrogen bonding in β-ala-*aza*-PNA-T involving a_{NH} and one of the backbone carbonyl groups. Strong NOSEY cross peak of a_{NH} with the backbone β/γ-methylene protons hint a folded structure of the *aza*-PNA backbone in $CDCl_3$ and DMSO-d_6. N_α-benzylated derivative of β-ala-*aza*-PNA-T shows a mixture of rotamer which further support the NMR observation of β-ala-*aza*-PNA-T. Quantification of intramolecular hydrogen bonding by calculating the acidity of hydrogen bond (A value), NMR experiments of *deaza* monomer β-ala-PNA-T (variable temperature NMR, DMSO titration) and Infrared (IR) spectral studies of *aza*-monomers and *deaza*-monomers further support the NMR results of β-ala-*aza*-PNA-T. NMR experiments of α-*aza*-PNA-T (DMSO titration, variable temperature NMR, NOSEY) and IR spectral studies of α-*aza*-PNA-T proves the existence intramolecular hydrogen bonding in a folded structure of the *aza*-PNA backbone in $CDCl_3$. The folded structure of α-*aza*-PNA-T opens up in DMSO-d_6. The crystal structure of the N-acetylchloro-α-*aza*-aeg (**7**) shows an eight membered intramolecular hydrogen bonding between N-terminal amide carbonyl and the backbone a_{NH} suggesting an extended γ-turn type folded structure of the *aza*-PNA backbone. Favourable bonding distance between the a_{NH} and N-terminal amide carbonyl of 2.006 Å facilitates the intramolecular hydrogen bonding interaction. Apart from the intramolecular hydrogen bonding the crystal structure shows dimeric units held by intermolecular hydrogen bond between the δ CO group and the carbamate NH of the PNA backbone. The side chain δ-CH_2 is oriented towards the *aza*-glycine side suggesting a E-rotamer in the crystal structure. Theoretical calculations support the experimental observation of *aza*-PNA and show strong evidence for a γ-turn type folded structure of the *aza*-PNA backbone. The E-rotamer formation in *aza*-PNA is stabilized by favourable C_β-H$\cdots O_A$ interaction, with favorable bond lengths of 2.04 Å in α-*aza*-PNA-T and 1.95 Å in β-ala-*aza*-PNA-T. Calculated distance between a_{NH} and N-terminal carbonyl group in *aza*-PNA (1.95 Å in α-*aza*-PNA-T and 1.91 Å in β-ala-*aza*-PNA-T) are in close agreement with the same observed in crystal structure (2.006 Å). The results of theoretical calculations of the *deaza* monomers further supports the γ-turn motif in the *aza*-PNA monomers. Although formation of intraresidue five membered hydrogen bonding is possible in β-ala-*aza*-PNA-T, the overall experimental results and theoretical study of *aza*-PNAs ruled out such possibility. In conclusion favourable intramolecular hydrogen bonding interaction of a_{NH} with the N-terminal amide carbonyl of *aza*-PNA monomers (α-*aza*-PNA-T and β-ala-*aza*-PNA-T) and favourable C_β-H$\cdots O_A$ interaction drives the rotamer selectivity by slightly compacting the *aza*-PNA backbone through the formation of extended γ-turn type motif.

2.5. Summary

- *aza*-PNA monomers were designed, synthesized and the purpose of the design explained clearly.
- β-ala-*aza*-PNA-T exhibits solvent independent single E-rotamer stabilization in $CDCl_3$ and DMSO-d_6.
- α-*aza*-PNA-T exhibits solvent dependent preference of rotamer mixture in $CDCl_3$ (major Z and minor E) and DMSO-d_6 (major E and minor Z).
- Crystal structure of the synthetic precursor of α-*aza*-PNA-T (compound **7**) shows extensive intramolecular/ intermolecular hydrogen bonding and C=O...O=C interaction driven E-rotamer stabilization in solid state.
- Eight membered Intramolecular hydrogen bonding between a_{NH} and N-terminal amide carbonyl in the crystal structure of compound 7 confirms the formation of an extended γ-turn like motif in the *aza*-PNA monomers.
- Results of NMR experiments, IR spectral studies and Crystal structure of compound 7 supports γ-turn (eight membered intramolecular hydrogen bond between a_{NH} and N-terminal amide carbonyl) driven E-rotamer stabilization in β-ala-*aza*-PNA-T.
- Theoretical calculations and computational studies performed on the *aza*-PNA rotamers strongly support the experimental results.

2.6. Experimental section

This section describes the detailed synthesis of monomers and the spectral characterisation of the rationally designed *aza*-PNA and *deaza*-PNA monomers.

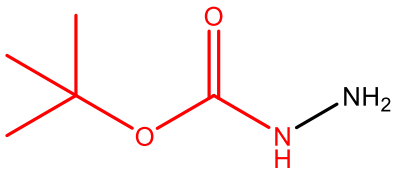
2.6.1 General information

Laboratory grade chemicals were used for setting up standard chemical reactions. Chemicals were purchased from Sigma Aldrich and Spectrochem. For moisture sensitive reactions the reagents were redistilled to ensure the maximum purity of the reagents. Solvents used for performing chemical reactions were dried using standard procedures available in the literature. Reactions were monitored using Thin Layer Chromatography (TLC). Standard UV-Lamp, Iodine spray and spraying with ninhydrin solution followed by heating the TLC plate were used to analyse the TLC. Standard procedure of work up involves extracting the desired

organic compound by washing the aqueous layer with organic solvent (DCM, EtOAc) followed by washing of the combined organic layer with brine. The combined organic layer was dried over Na_2SO_4 or Calcium chloride ($CaCl_2$ For DCM) and evaporated under reduced pressure to afford the desired organic compound. Precoated silica gel GF_{254} (Merck 5554) TLC plates were used for TLC analysis. Standard chemical reactions were performed in room temperature, wherever required temperature was mentioned. Silica gel (60-120 or 100-120 mesh) were used to perform the column chromatography. Double distilled EtOAc, hexane, MeOH and DCM were used as a mobile phase to perform the column chromatography. ^{1}H NMR, ^{13}C NMR and 2D-NMR experiments were performed in Bruker-400 MHz, JEOL-400 MHz, Bruker-500 MHz, and Bruker-600 MHz NMR spectrometers. NMR solvents were purchased from Sigma Aldrich. Deuterated chloroform ($CDCl_3$) and DMSO-d_6 were used as an NMR solvent. Chemical shifts were reported in δ ppm with respect to the internal standard TMS or deuterated NMR solvent. Mass spectra of the reaction intermediates were recorded in an Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using TiO_2 or 2,5-dihydroxybencoic acid (DHB) as matrix. To further confirm the mass spectra of small molecules (Mass 100-1200) obtained in MALDI, High Resolution Mass Spectrometry (HRMS) was performed for each reaction intermediates. IR spectra of the samples were recorded in a Bruker IR spectrometer.

2.8.2. Experimental procedures and spectral data

t-butyl carbazate (1)[32]

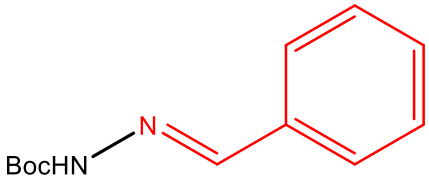


1

The synthesis of mono Boc-protected hydrazine was performed using the known literature procedure.[32] Hydrazine monohydrate (32 g, 0.5 mol) was mixed with isopropanol (400 ml) at 0^{o}C, then a solution of Boc_2O (50 g, 0.2 mol) in isopropanol (200 ml) was added dropwise at 0^{o}C. The reaction mixture turned cloudy upon addition and stirring was continued at room temperature for additional 5 min after the addition (*Caution-Hydrazine hydrate is a very reactive amine, keeping the reaction at room temperature more than 5 min produces di Boc substituted hydrazine, which is crystalline solid and difficult to separate from mono Boc*

substituted hydrazine by column chromatography). The solvent was evaporated under reduced pressure, the residue was dissolved in water and the mono Boc-protected hydrazine was extracted using EtOAc (3 x 250 ml). Then EtOAc was evaporated under reduced pressure to give **1** as yellowish liquid. The crude yellow liquid was purified by flash chromatography using 35% EtOAc in hexane to afford **1** as a colourless liquid which solidify at low temperature (22.8 g, 75%). R_f 0.20 (50% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ 1.42 (s, 9H), 3.68 (s, 2H), 6.13 (s, 1H) ppm. ^{13}C NMR (100 MHz, CDCl3, TMS): δ 28.5, 77.2, 158.3 ppm. The obtained white crystal of mono Boc-hydrazine was used for the next step without further analysis.

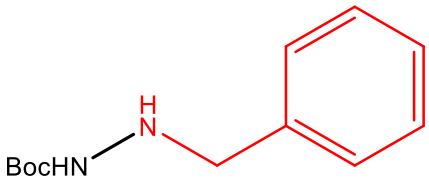
t-Butyl 2-benzylhydrazinecarboxylate (10)[32]



10

t-Butyl 2-benzylhydrazinecarboxylate (**10**) was prepared as per as the literature procedure and used without further purification and characterisation.[32]

tert-Butyl 2-benzylhydrazinecarboxylate (11)[32]

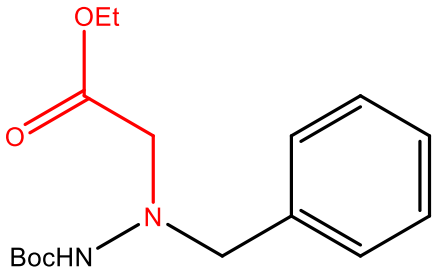


11

Orthogonally protected hydrazine derivative **11** was synthesized from the Schiff base precursor **10** using literature reported procedure.[32] To a suspension of Schiff base precursor **10** in THF (5 ml) at 0^{o}C was added acetic acid (3 ml) followed by sodium cyanoborohydride (0.3 g, 4.3 mmol). After 5 min, the reaction cleared up and then stirred at room temperature for additional 16 h. EtOAc (15 ml) and saturated $NaHCO_3$ (15 ml) were added to the reaction mixture. After the organic layer was collected, the aqueous layer was neutralized with solid $NaHCO_3$, and extracted two times with EtOAc (2 x 100 ml). The organic fractions were pooled, washed with saturated sodium bicarbonate and brine (3 x 100 ml), dried over Na_2SO_4, and

concentrated in vacuo. The titled orthogonally protected hydrazine derivative was obtained as a white hygroscopic solid. The crude was purified by flash chromatography to afford **11** as colourless liquid which solidify at low temperature (0.3 g, 82% yield), R_f = 0.44 (2 % MeOH in CH_2Cl_2), ^{1}H NMR (400 MHz, CHLOROFORM-D) δ 7.36 – 7.16 (m, 5H), 6.81 (s, 1H), 4.21 (s, 1H), 3.93 (s, 2H), 1.48 (s, 9H), ^{13}C NMR (100 MHz, CHLOROFORM-D) δ 156.7 , 137.4 , 129 , 128.2, 127.1 , 80.1, 55.4 , 28.2, Calculated mass $(C_{12}H_{18}N_2O_2 + Na)^+$ 245.1265 and observed mass 245.1270.

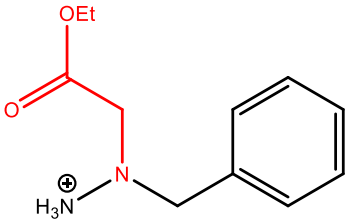
tert-butyl 2-benzyl-2-(2-ethoxy-2-oxoethyl)hydrazinecarboxylate (12a)



12a

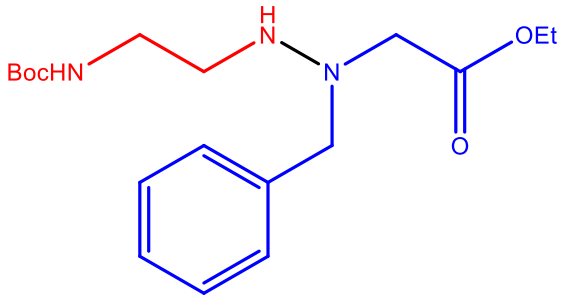
Orthogonally protected hydrazine amine (**11**, 200 mg, 0.9 mmol) was dissolved in CH_2Cl_2 (5 ml) and to it was added DIPEA (0.3 ml, 1.5 mmol). After 5 min ethyl bromoacetate (0.2 ml, 1.4 mmol) was added to the reaction mixture and the reaction mixture was refluxed at 45^{o}C for 36 h. After the completion of the reaction brine was added to the reaction mixture and the aqueous layer was extracted with DCM (3 x 100 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over $MgSO_4$, concentrated in vacuo to give the orthogonally protected N-alkylated hydrazine derivative **12a** as yellow oil. The obtained crude was purified by flash chromatography using 25% EtOAc in hexane as eluent to afford **12a** as yellow oil (0.2 g, 63% yield), R_f = 0.65 (20% EtOAC in hexane), ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.00 (m, 5H), 6.72 (s, 1H), 4.19 (q, *J* = 7.1 Hz, 2H), 4.11 (s, 2H), 3.66 (s, 2H), 1.39 (s, 9H), 1.28 (t, *J* = 7.2 Hz, 3H), ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.2, 154.9, 136.9, 129.4, 128.4, 127.7, 80.12, 60.8, 60.4, 55.4, 28.3, 14.2, Calculated mass ($C_{16}H_{24}N_2O_4$ +Na), 331.1634; observed mass, 331.1641.

Ethyl 2-(1-benzyl-2-(2-((terbutoxycarbonyl)amino)ethyl)hydrazinyl)acetate (13a)



12c

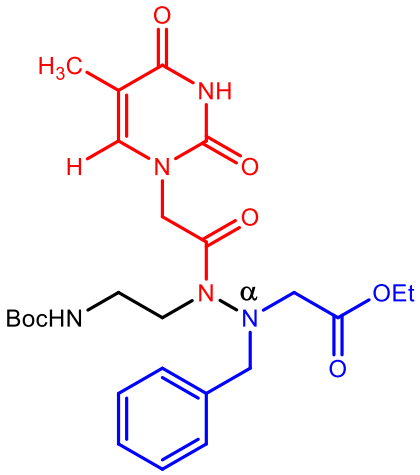
N-alkylated hydrazine amine **12a** was dissolved in 50% TFA in DCM and the reaction mixture was stirred at room temperature for 4 h. Azeotropic mixture of TFA was prepared by adding toluene in the reaction mixture and evaporation of the solvent under reduced pressure followed by drying in vacuo gives the free amine derivative of N-alkylated hydrazine ester **12c**. The obtained free amine of **12a** (**12c**) was used for the next step without further purification and analysis.



13a

Free amine derivative of N-alkylated hydrazine ester **12c** (966 mg, 3.1 mmol) was dissolved in ethanol (5 ml) and tert-butyl (2-oxoethyl) carbamate (250 mg, 1.5 mmol) dissolved in ethanol (3 ml) was added to the reaction mixture and stirred at 0^{o}C for 5 min. DIPEA (0.9 ml, 4.9 mmol) was added to the reaction mixture. After that the mixture was stirred for 60 min at 0^{o}C and followed by $Na(CN)BH_3$ (150 mg, 2.3 mmol) and catalytic amount of acetic acid was added to it. The reaction mixture was stirred at 0^{o}C for additional 16 h. After the completion of the reaction ethanol was evaporated under reduced pressure and $NaHCO_3$ solution was added to it. The organic layer was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the β-ala-*aza*-PNA backbone **13a** as yellow oil. The crude was purified by flash chromatography using 40% EtOAc in hexane as eluent to afford the *aza*-PNA backbone **13a** as yellow oil (170 mg, 55% yield), R_f = 0.50 (25% EtOAc in hexane).^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.26 (m, 5H), 5.27 (s, 1H), 4.18 (q, *J* = 7.2 Hz, 2H), 3.91 (s, 2H), 3.38 (s, 2H), 3.25 (dd, *J* = 10.4, 5.8 Hz, 2H), 2.85 – 2.81 (m, 2H), 1.44 (s, 9H), 1.27 (t, *J* = 7.1 Hz, 3H), ^{13}C NMR (101 MHz, $CDCl_3$) δ 171, 156.1, 137.4, 129.3, 128.5, 127.5, 78.6, 60.5, 59.2, 55.9, 47.3, 39.4, 37.1, 28.4, 14.2, Calculated mass ($C_{18}H_{29}N_3O_4$+H^+), 352.2236; observed mass, 352.2235.

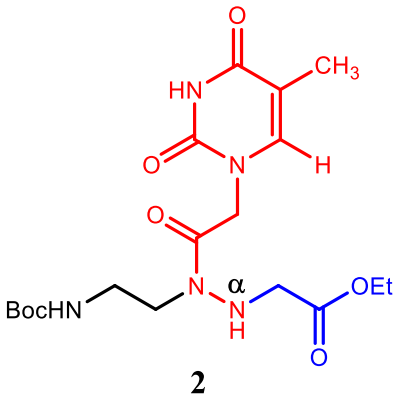
Ethyl 2-(1-benzyl-2-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)hydrazinyl)acetate (14a)



14a

Thymine acetic acid (480 mg, 2.6 mmol) was dissolved in dry DMF (4 ml), and the reaction mixture was cooled at -73°C. N-methyl morpholine (0.7 ml, 6.5 mmol) and isobutyl chloroformate (0.9 ml, 6.8 mmol) was added to it and the reaction mixture was stirred at -73° C for 30 min. Solution of free amine *aza*-PNA backbone **13a** (0.5 g, 1.4 mmol) in dry DMF (2 ml) and DIPEA (0.7 ml, 3.9 mmol) was added to it. The reaction mixture was kept at 0°C for 30 min and then kept for stirring at room temperature for additional 24 h. Brine was added to the reaction mixture and the desired compound was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the nucleobase coupled N_α-benzylated *aza*-PNA monomer **14a** as yellow oil. The crude was purified by flash chromatography using 75% EtOAc in hexane as eluent to afford **14a** as white hygroscopic solid (200 mg, 45% yield), R_f = 0.40 (75% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.42 (d, *J* = 17.7 Hz, 1H), 7.43 – 7.26 (m, 5H), 6.58 (s, 1H), 5.82 (s, 1H), 5.07 (dd, *J* = 21.3, 12.4 Hz, 1H), 4.50 (d, *J* = 17.1 Hz, 1H), 4.19 – 4.10 (m, 2H), 4.05 – 3.88 (m, 2H), 3.69 (d, *J* = 17.0 Hz, 2H), 3.56 (dd, *J* = 11.8, 5.1 Hz, 3H), 3.40 – 3.29 (m, 2H), 1.84 (d, *J* = 4.1 Hz, 3H), 1.43 (d, *J* = 16.8 Hz, 9H), 1.26 (dd, *J* = 9.5, 4.7 Hz, 3H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 170.1, 168.5, 164.5, 156, 151, 141, 137, 135, 129.2, 128.4, 128, 109.3, 78.6, 60.6, 60, 58.7, 54, 48.6, 41, 38.2, 28, 13.6, 12, Calculated mass $(C_{25}H_{35}N_5O_7 + H)^+$ 518.2614; Observed mass, 518.2613.

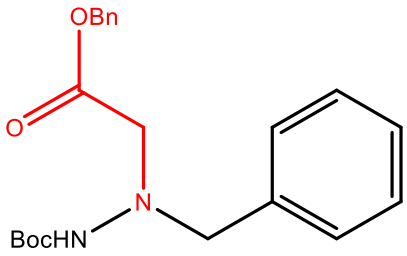
Ethyl2-(2-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)hydrazinyl)acetate (2)



2

N_α-benzylated thymine-*aza*-PNA monomer **14a** was dissolved in a 3:1 mixture of THF and IPA and to that catalytic amount of 20% $Pd(OH)_2/C$ and catalytic amount of 15% Pd/C was added. The reaction mixture was subjected for hydrogenation for 72 h at 1 atm pressure. After completion of the reaction the reaction mixture was filtered using celite and washed with MeOH. Mixture of 20% $Pd(OH)_2/C$ and 15% Pd/C was quenched by 2(N) HCl solution. Evaporation of MeOH under reduced pressure gives the β-ala-*aza*-PNA-T ethyl ester **2** as white hygroscopic solid. Purification of the crude by flash chromatography using 10% MeOH in DCM as eluent gives the β-ala-*aza*-PNA-T ethyl ester **2** as white solid (70%), R_f 0.4 (80% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.78 (s, 1H), 6.97 (d, *J* = 20.5 Hz, 1H), 5.11 (s, 2H), 4.93 (s, 2H), 4.24 – 4.17 (m, 2H), 3.70 – 3.59 (m, 4H), 3.35 (d, *J* = 5.1 Hz, 2H), 1.90 (s, *J* = 0.8 Hz, 3H), 1.44 (s, 9H), 1.32 – 1.25 (t, 3H), ^{13}C NMR (101 MHz, $CDCl_3$) δ 171, 170.48 , 164.5, 157.1, 141.5, 110.4, 81.3, 80.2, 61.5, 49.7, 49, 42.4, 36.8, 28.4, 14.2, 12.4, Calculated mass $(C_{18}H_{29}N_5O_7 + Na)^+$ 450.1964 and Observed mass 450.1964.

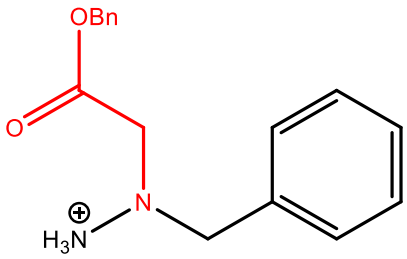
tert-butyl 2-benzyl-2-(2-(benzyloxy)-2-oxoethyl)hydrazinecarboxylate (12b)



12b

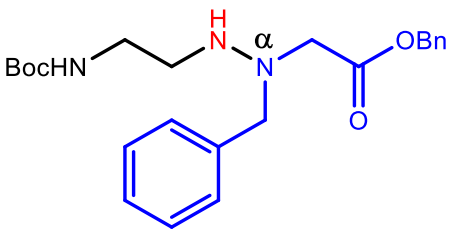
Orthogonally protected hydrazine amine (**11**, 5 g, 22.4 mmol) was dissolved in CH_2Cl_2 (80 ml) and to it was added DIPEA (7.9 ml, 45 mmol). After 5 min benzyl bromoacetate (5.4 ml, 34 mmol) was added to the reaction mixture and the reaction mixture was refluxed at 45^oC for 48 h. After the completion of the reaction brine was added to the reaction mixture and the aqueous layer was extracted with DCM (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over $MgSO_4$, concentrated in vacuo to give the orthogonally protected N-alkylated hydrazine derivative **12b** as yellow oil. The obtained crude was purified by flash chromatography using 20% EtOAc in hexane as eluent to afford **12b** as yellow oil (4.2 g, 50% yield), R_f = 0.80 (20% EtOAC in hexane), ^{1}H NMR (400 MHz, CHLOROFORM-D) δ 7.25 – 7.08 (m, 10H), 5.00 (s, 2H), 4.51 (s, 1H), 3.97 (s, 1H), 3.83 (s, 1H), 3.77 (s, 1H), 3.57 (s, 1H), 1.28 (d, *J* = 25.0 Hz, 9H), ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.5, 156.5, 154.7, 141.3, 137.1, 136.6, 135.3, 129.2, 129, 128.5, 128.3, 128.2, 127.5, 127.2, 126.8, 67.2, 64.7, 60.3, 55.5, 28.2, Calculated mass $(C_{21}H_{26}N_2O_4+Na)^+$ 393.1790 and Observed mass 393.1791.

Benzyl2-(1-benzyl-2-(((tert-butoxycarbonyl)amino)ethyl)hydrazinyl)acetate (13b)



12d

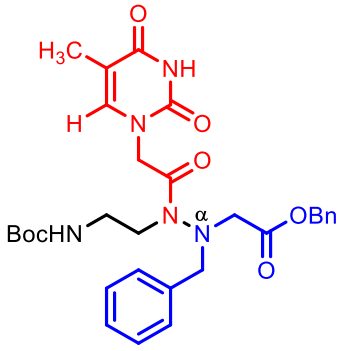
N-alkylated hydrazine amine **12b** was dissolved in 50% TFA in DCM and the reaction mixture was stirred at room temperature for 5 h. Azeotropic mixture of TFA was prepared by adding toluene in the reaction mixture and evaporation of the solvent under reduced pressure followed by drying in vacuo gives the free amine derivative of N-alkylated hydrazine ester **12d**. The obtained free amine of **12b** (**12d**) was used for the next step without further purification and analysis.



13b

Free amine derivative of N-alkylated hydrazine ester **12d** (1.2 g, 3.1 mmol) was dissolved in ethanol (10 ml) and tert-butyl (2-oxoethyl) carbamate (250 mg, 1.5 mmol) dissolved in ethanol (3 ml) was added to the reaction mixture and stirred at 0°C for 5 min. DIPEA (0.8 ml, 4.7 mmol) was added to the reaction mixture. After that the mixture was stirred for 60 min at 0°C and followed by $Na(CN)BH_3$ (147 mg, 2.3 mmol) and catalytic amount of acetic acid was added to it. The reaction mixture was stirred at 0°C for additional 20 h. After the completion of the reaction ethanol was evaporated under reduced pressure and $NaHCO_3$ solution was added to it. The organic layer was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the β-ala-*aza*-PNA backbone **13b** as yellow oil. The crude was purified by flash chromatography using 35% EtOAc in hexane as eluent to afford the *aza*-PNA backbone **13b** as yellow oil (175 mg, 62% yield), R_f = 0.65 (25% EtOAc in hexane). ^{1}H NMR (400 MHz, CHLOROFORM-D) δ 7.36 – 7.22 (m, 10H), 5.25 (t, *J* = 4.5 Hz, 1H), 5.21 (s, 2H), 5.14 (d, *J* = 7.0 Hz, 2H), 3.88 (s, 2H), 3.42 (s, 2H), 3.16 (dd, *J* = 44.9, 4.8 Hz, 2H), 2.78 (dd, *J* = 19.5, 14.6 Hz, 2H), 1.47 – 1.40 (m, 9H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 170.8, 156, 137.2, 135.6, 129.3, 128.6, 128.5, 128.4, 127.6, 66.82, 59.1, 55.8, 47.3, 39.4, 37.1, 28.5, Calculated mass $(C_{23}H_{31}N_3O_4+H)^+$ 414.2393 and Observed mass 414.2397.

Benzyl2-(1-benzyl-2-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)hydrazinyl)acetate (14b)



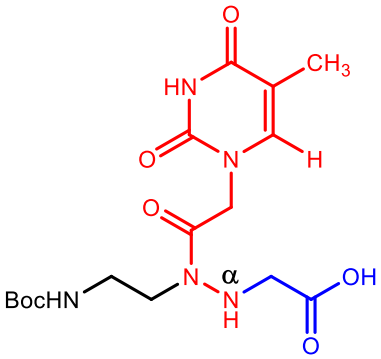
14b

Thymine acetic acid (480 mg, 2.6 mmol) was dissolved in dry DMF (4 ml), and the reaction mixture was cooled at -73°C. N-methyl morpholine (0.7 ml, 6.5 mmol) and isobutyl chloroformate (0.9 ml, 6.8 mmol) was added to it and the reaction mixture was stirred at -73° C for 30 min. Solution of free amine *aza*-PNA backbone **13b** (592 mg, 1.4 mmol) in dry DMF (2 ml) and DIPEA (0.7 ml, 3.9 mmol) was added to it. The reaction mixture was kept at 0°C

for 30 min and then kept for stirring at room temperature for additional 24 h. Brine was added to the reaction mixture and the desired compound was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the nucleobase coupled N_α-benzylated *aza*-PNA monomer **14b** as yellow oil. The crude was purified by flash chromatography using 70% EtOAc in hexane as eluent to afford **14b** as white hygroscopic solid (300 mg, 44% yield), R_f = 0.50 (75% EtOAc in hexane). ^{1}H NMR (400 MHz, CHLOROFORM-D) δ 9.20 (s, 1H), 7.37 – 7.31 (m, 10H), 6.64 (s, 1H), 6.46 (s, 1H), 5.21 (t, *J* = 5.7 Hz, 1H), 5.12 (dd, *J* = 19.5, 7.4 Hz, 2H), 5.01 (d, *J* = 17.2 Hz, 1H), 4.45 (dd, *J* = 17.8, 6.9 Hz, 1H), 3.99 (d, *J* = 12.6 Hz, 1H), 3.86 (d, *J* = 12.6 Hz, 1H), 3.72 (d, *J* = 17.1 Hz, 1H), 3.62 – 3.52 (m, 3H), 3.36 – 3.24 (m, 2H), 1.84 (s, 3H), 1.42 (s, 9H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 170.7, 169, 164.5, 156.5, 151.4, 141.4, 135.1, 129.8, 129, 128.7, 128.6, 128.5, 116.1, 110.3, 67.1, 59.4, 54.6, 50.6, 41.5, 38.9, 28.45, 12.4. Calculated mass $(C_{30}H_{37}N_5O_7 + H)^+$ 580.2771 and Observed mass 580.2779.

Alternative Procedure: To a stirred solution of Thymine acetic acid (0.5 mmol, 100 mg) in dry DMF (5 ml) was added DCC (0.5 mmol, 112 mg), DHbtOH (0.5 mmol, 89 mg) and stirred for 1 h at room temperature. After 1 h, solution of *aza*-PNA backbone **13b** in dry DMF (0.3 mmol, 135 mg) was added to the reaction mixture containing activated thymine acetic acid. The reaction mixture was heated at 50^{o}C for 40 hours. After the completion of the reaction, the reaction was quenched by addition of saturated $NaHCO_3$ solution, and the desired organic compound was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine solution (3 x 150 ml) and dried over Na_2SO_4. Evaporation of EtOAc under reduced pressure gives the N_α-benzylated *aza*-PNA monomer **14b** as a white hygroscopic solid. Purification of the crude by flash chromatography using 70% EtOAc in hexane gives the N_α-benzylated *aza*-PNA monomer **14b** as a hygroscopic white solid (477 mg, 70% yield).

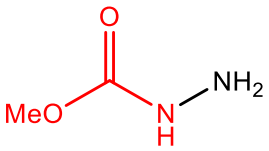
2-(2-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)hydrazinyl)acetic acid (15)



15

N_{α}-benzylated thymine-*aza*-PNA monomer **14b** was dissolved in a 3:1 mixture of THF and IPA and to that catalytic amount of 20% $Pd(OH)_2/C$ and catalytic amount of 15% Pd/C was added. The reaction mixture was subjected for hydrogenation for 72 h at 1 atm pressure. After completion of the reaction the reaction mixture was filtered using celite and washed with MeOH. Mixture of 20% $Pd(OH)_2/C$ and 15% Pd/C was quenched by 2(N) HCl solution. Evaporation of MeOH under reduced pressure gives the β-ala-*aza*-PNA-T free acid monomer **15** as white solid. Purification of the crude by flash chromatography using 20% MeOH in DCM as eluent gives the β-ala-*aza*-PNA-T free acid monomer **15** as white solid (75%), R_f 0.35 (10% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO) δ 7.24 (s, 1H), 6.88 (t, *J* = 5.5 Hz, 1H), 5.30 (s, 1H), 4.78 (s, 2H), 3.45 (t, *J* = 6.0 Hz, 2H), 3.10 (d, *J* = 5.9 Hz, 2H), 1.74 (s, 3H), 1.40 (d, *J* = 19.0 Hz, 9H), ^{13}C NMR (101 MHz, DMSO-d_6) δ 171.8, 169.7, 164.5, 156.0, 151.2, 142.2, 107.9, 78.0, 58.0, 48.8, 48.6, 43.0, 37.0, 28.2, 12.0, Calculated mass $(C_{16}H_{25}N_5O_7 + Na)^+$ 422.1651 and Observed mass 422.1657.

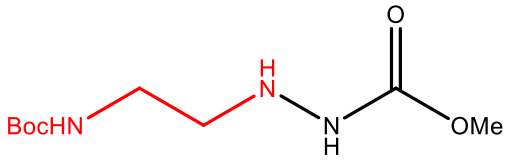
Methyl hydrazinecarboxylate(5)[31]



5

Methyl hydrazinecarboxylate **5** was synthesized using the standard procedure reported in the literature.[31] Hydrazine monohydrate (7.8 ml) and dimethyl carbonate (15 ml) were mixed and stirred at 50°C for 20 minutes. The reaction mixture was allowed to cool at rt and stirred overnight at room temperature. The unreacted dimethyl carbonate was removed from the reaction mixture by evaporating the reaction mixture under reduced pressure. The crude product was dried in vacuo to obtain Methyl hydrazinecarboxylate **5** as a white solid. The obtained white solid was use for the next step without further purification.

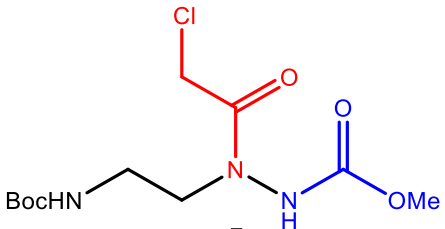
Methyl 2-(2-((tert-butoxycarbonyl)amino)ethyl)hydrazine-1-carboxylate(6)



6

Methyl hydrazinecarboxylate **5** (2 gm, 22.2 mmol) was dissolved MeOH and tert-butyl(2-oxoethyl) carbamate (3.5 gm, 22.2 mmol) followed by DIPEA (7.8 ml, 44.4 mmol) was added to the reaction mixture at 0^{o}C. The reaction mixture was allowed to stirred at 0^{o}C for 2 h. $Na(CN)BH_3$ (2 gm, 33.3 mmol) and catalytic amount of glacial acetic acid was added to the reaction mixture. The reaction was kept for stirring at 0^{o}C for additional 15 h. After the completion of the reaction MeOH was evaporated under reduced pressure and the crude was dissolved in water. The desired α-*aza*-PNA backbone **6** was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (2 x 100 ml) and $NaHCO_3$ solution (2 x 100 ml). Combined organic layer was dried over Na_2SO_4 and evaporated under reduced pressure to give the α-*aza*-PNA backbone **6** as yellow oil. The crude was purified by flash chromatography using 45% EtOAc in hexane as eluent to afford **6** as yellow oil (3.4 gm, 65% yield), R_f = 0.35 (75% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.02 (s, 1H), 5.31 (t, *J* = 42.9 Hz, 2H), 3.88 (t, *J* = 4.6 Hz, 2H), 3.75 (d, *J* = 8.7 Hz, 2H), 1.39 (s, 9H), ^{13}C NMR (101 MHz, $CDCl_3$) δ 156, 154.6, 143.8, 78.0, 52.8, 41.8, 28.36. Calculated Mass ($C_9H_{19}N_3O_4$ +Na)$^{+}$ 256.1273, Observed Mass 256.1277.

Methyl2-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(2-chloroacetyl)hydrazine-1-carboxylate (7)



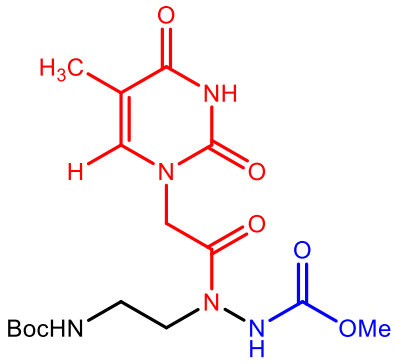
7

α-*aza*-PNA backbone **6** (1 g, 3.2 mmol) was dissolved in dry DCM (15 ml) and Triethyl amine (1.2 ml, 8.1 mmol) was added in to the reaction mixture at 0^{o}C. Chloroacetyl chloride (0.4 µL, 4.8 mmol) was added to the reaction mixture drop by drop (*Caution-Addition of Chloroacetyl chloride produces heat, add the reagent slowly and drop by drop*) at 0^{o}C. The reaction mixture was stirred at room temperature for additional 30 h. After the completion of the reaction saturated $NaHCO_3$ solution was added to the reaction mixture and the N-acetylchloro aminoethyl-α-*aza* glycine **7** was extracted by DCM (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and dried over $MgSO_4$. The organic layer was evaporated under reduced pressure to give the N-acetylchloro aminoethyl-α-*aza* glycine **7** as yellow oil. The obtained crude was purified by flash chromatography using 40% EtOAc in hexane as eluent to afford **7** as yellow oil (0.8 g, 65% yield), R_f = 0.65 (75% EtOAc in hexane).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (s, 1H), 4.97 (s, 1H), 4.36 (s, 1H), 4.19 (d, *J* = 21.1 Hz, 2H), 3.78 (s, *J* = 9.1 Hz, 4H), 3.63 (s, 1H), 3.01 (s, 2H), 1.40 (s, 9H), ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.6, 157.8, 155.8, 80.6, 53.4, 46.5, 42.2, 36.5, 28.3, Calculated Mass $(C_{11}H_{20}N_3O_5Cl + Na)^+$ 332.0989, Observed Mass 332.0991.

Alternative procedure: α-*aza*-PNA backbone **6** (500 mg, 2.1 mmol) was dissolved in 1:1 Dioxane (20 ml) and water (20 ml). The reaction mixture was cooled to 0^{o}C. Saturated $NaHCO_3$ solution (40 ml) was added to it. Chloroacetyl chloride (340 µL, 4.2 mmol) was added in the cold reaction mixture and the mixture was stirred at room temperature for additional 30 h. After the completion of the reaction saturated $NaHCO_3$ solution was added to the reaction mixture and the desired compound was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml). The organic layer was dried over Na_2SO_4 and evaporation of EtOAc under reduced pressure to give the N-acetylchloro aminoethyl-α-*aza* glycine **7** as yellow oil. The crude was purified by flash chromatography using 40% EtOAc in hexane as eluent to afford **7** as yellow oil.

Methyl2-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)hydrazine-1-carboxylate (1)

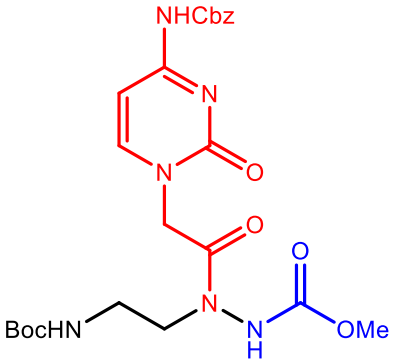


1

To a suspension of thymine (112 mg, 0.9 mmol) in dry DMF (4 ml) was added anhydrous K_2CO_3 (114 mg, 0.9 mmol) and stirred at room temperature for 5 min. N-acetylchloro aminoethyl-α-*aza* glycine (**7,** 200 mg, 0.6 mmol) was added to the reaction mixture. Resulting mixture was stirred at room temperature for additional 48 h. The reaction mixture was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml) . The combined organic layer was washed with brine (3 x 150 ml) and the same was dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the methyl ester of α-*aza*-PNA-T monomer **1** as a white hygroscopic solid. The crude was

purified by flash chromatography using 10% MeOH in DCM as eluent to afford the **1** as white solid (195 mg, 75% yield), R_f = 0.5 (100% EtOAc).^{1}H NMR (400 MHz, CHLOROFORM-D) δ 8.69 (s, 1H), 8.62 (s, 1H), 6.90 (s, 1H), 5.11 (s, 1H), 4.25 (s, 2H), 3.81 (s, 4H), 3.62 (s, 1H), 3.08 (s, 2H), 1.90 (d, *J* = 0.9 Hz, 3H), 1.44 (s, 9H), ^{13}C NMR (101 MHz, DMSO-d_6) δ 170.3, 165.9, 164.1, 157, 151.9, 143.2, 109.7, 79.8, 54, 49, 28.9, 12.6, Calculated Mass ($C_{16}H_{25}N_5O_7$ +Na)$^+$ 422.1652,Observed mass 422. 1656.

Methyl2-(2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)acetyl)-2-(2-((tert-butoxycarbonyl)amino)ethyl)hydrazine-1-carboxylate (1C)



1C

To a suspension of (N^4-Cbz)-cytosine (296 mg, 1.2 mmol) in dry DMF (5 ml) was added anhydrous K_2CO_3 (168 mg, 1.2 mmol) and stirred at room temperature for 5 min. N-acetylchloro aminoethyl-α-*aza* glycine (**7,** 250 mg, 0.8 mmol) was added to the reaction mixture. Resulting mixture was stirred at room temperature for additional 48 h. The reaction mixture was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml) . The combined organic layer was washed with brine (3 x 150 ml) and the same was dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the methyl ester of α-*aza*-PNA-C monomer **1C** as a white hygroscopic solid. The crude was purified by flash chromatography using 10% MeOH in DCM as eluent to afford the **1C** as white solid (318 mg, 75% yield), R_f = 0.65 (100% EtOAc), ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.94 (s, 1H), 7.92 (s, 1H), 7.47 – 7.30 (m, 5H), 7.00 (s, 1H), 6.72 (s, 1H), 5.19 (s, 2H), 4.84 (d, *J* = 16.4 Hz, 1H), 4.47 (d, *J* = 13.7 Hz, 1H), 3.70 (s, 3H), 3.50 (s, 1H), 3.34 (s, 1H), 3.06 (s, 2H), 1.36 (s, 9H), ^{13}C NMR (101 MHz, DMSO-d_6) δ 170.3, 165.9, 164.1, 157, 151.9, 143.2, 109.7, 79.8, 54, 49, 28.9, 12.6, Calculated Mass ($C_{23}H_{30}N_6O_8$+H)$^+$ 519.2203, Observed Mass 519.2180.

2.7. References

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

2. a) Stein, C. A.; Cohen, J. S: Phoshorothioate oligodeoxynucleotide analogues. In Cohen, J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 97 b) Millar, P. S: Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. London: Macmillan Press, **1989**, p. 79. c) Froehler, B.; Ng, P.; Matteucci, M. *Nucleic Acid Res.* **1988**, *16*, 4831-4839.

3. Toule, J-J: in Antisense oligonucleotides and antisense RNA: Novel Pharmacological and Therapeutic Agents. (Weiss, B., ed), CRC press, 1997, pp. 1-16.

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

5. a) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordon, B.; Nielsen, P. E. *Nature,* **1993**, *365*, 566-568. b) Jensen, K. K.; Qrum, H.; Nielsen, P. E.; Norden, B. *Biochemistry* **1997**, *36*, 5072-5077.

6. a) Qrum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic acid Res*. **1993,** *21*, 5332- 5336. b) Thiede, C.; Bayerdoerffer, E.; Blasczyk, R.; Wittig, B.; Neubauer, A. *Nucleic Acid Res*. **1996,** *24*, 983-984. c) Demers, D. B.; Curry, E. T.; Egholm, M.; Sozer, A. C. *Nucleic Acid Res*.**1995,** *23*, 3050-3055. d) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Nucleic Acid Res*. **1993,** *21*, 197-200.

7. (a) Good, L.; Nielsen, P. E. *Antisense Nucleic Acid Drug Dev*. **1997**, *7*, 431. (b) De Mesmaeker, A. et al: *Curr. Open. Struct. Biol.* **1995**, *5*, 343. (c) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W: PNA. *Angew. Chem. Int. Ed*. **1998**, *37*, 2796-2823.

8. (a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed*. **1996,** *35*, 1939-1942. (b) Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett*. **1998**, *39*, 4707-4710. (c) Sforza, S.; Galaverna, G.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality* **2002**, *14*, 7,591-598. d) Sugiyama, T.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Bioorg. Med. Chem. Lett*. **2011**, *21*, 7317-7320.

9. Drauz, K.; Waldmann, H; (eds): Enzyme Catalysis in Organic Synthesis. Vol. II, Wiley-VCH, 2002. (b) Bornscheuer, U. T.; Kazlauskas, R. J; (eds): Hydrolases in Organic Synthesis: Regio- and stereoselective *Biotransformations*. Wiley-VCH, **1999.**

10. (a) Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem*. **2000**, *4*, 1931-1943. (c) Kumar, V. A. *Eur. J. Org. Chem*. **2002**, 2021-2032.

11. a) Maison, W.; Schlemminger, I.; Westerhoff, O.; Martens, J. *Bioorg. Med. Chem. Lett*. **1999**, *9*, 581. b) Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc*., **1996**, *118*, 5544-5549. c) Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol*., **2010**, *50*, 259-293. d) Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron*., **1996**, *52*, 15017-15030. e) Gangamani, B. P.; D' Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides*., **1999**, *18*, 1409-1011. f) Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron*., **1999**, *55*, 177-192. g) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **1999**, *1*, 1513-1516. h) Vilaivan, T.; Khongdeesameor, C.; Harnyuttanokam, P.; Westwell, M. S.; Lowe, G. *Bioorg. Med. Chem. Lett*., **2000**, *10*, 2541-2545. i) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **2001**, *3*, 1281-1284.

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

13. Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. *Eur. J. Org. Chem*. **2000**, *2000*, 2905-2913.

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

15. (a) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed.* **1998,** *37*, 2796-2823. (b) Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc. Chem. Commun*. **1993**, 519.

16. a) Brown, S.C.; Thomson, S.A.; Veal, J.M.; Davis, D.G. *Science* **1994**, *265*, 777 - 780. b) Oleszczuk, M.; Motowidlo, S.R.; Falkiewicz, B., *nucleosides, nucleotides & nucleic acids*, **2001**, *20*, 1399–1402. c) Meltzer, P.C.; Liang, A.Y.; and Matsudaira, P. *J. Org. Chem*. **1995**, *60*, 4305-4308.

17. (a) Ovadia, R.; Lebrunb, A-L.; Barvikd, I.; Vasseurc, J.-J; Baragueyc, C.; Alvarez, K. *Org. Biomol. Chem.*, **2015**, *13*, 11052-11071.

18. Berger, O.; Abramovich, L.A.; Sakin, M.L.; Grunwald, A.; Peer, Y.L.; Bachar, M.; Buzhansky, L.; Mossou, E.; Forsyth, V.T.; Schwartz, T.; Ebenstein, Y.; Frolow, F.; Shimon, L.J.W.; Patolsky, F.; and Gazit. E. *Nature Nanotechnology*, **2015**, *10*, 353–360.

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

20. Kulkarni, P.; Datta, D.; Ramabhadran, R.O.; Ganesh, K.N. *Org. Biomol. Chem*., **2021**, *19*, 6534.

21. Sanjayan, G.J..; Pedireddi, V.R.; Ganesh, K.N. *organic letters.,* **2000**, *18*, 2825-2828.

22. Howarth, N.M.; Ricci, J. *Tetrahedron.,* **2011**, *67*, 9588-9594.

23. a) Govindaraju, T.; Gonnade, R.G.; Bhadbhade, M.M.; Kumar, V.A.; Ganesh, K.N. *organic letters.,* **2003**, *17*, 3013-3016. b) Lagriffoule, P.; Wittung, P.;Eriksson, M.; Jensen, K.K.; Nordkn, B.; Buchardt, O.; Nielsen, P.E. *Chem. Eur. J*., **1997**, *3*, 912–919. c) Govindaraju, T.; Kumar, V.A.; and Ganesh, K.N. *Chem. Commun*., **2004**, 860–861. d) Govindaraju, T.; Kumar, V.A.; and Ganesh, K.N. *J. Org. Chem*., **2004**, *69*, 5725-5734. e) Pokorski, J.K.; Witschi, M.A.; Purnell, B.A.; and Appella, D.H. *J. Am. Chem. Soc*., **2004**, *126*, 15067-15073. f) Zheng, H.; Botos, I.; Clausse, V.; Nikolayevskiy, H.; Rastede, E.E.; Fouz1, M.F.; Mazur, S.J.; and Appella, D.H. *Nucleic Acids Research.,* **2021**, *49*, 713–725.

24. Zhang, Y.; Malamakal, R.M.; and Chenoweth, D.M. *Angew.Chem.Int.Ed.,* **2015**, *54*,10826–10832. b) McMechen, M.A.; Willis, E.J.; Gourville, P.C.; and Proulx, C. *Molecules* **2019**, *24*, 1919.

25. (a) Chingle, R.; Proulx, C.; Lubell, W. D. *Acc. Chem. Res*. **2017**, *50*, 1541-1556 (b) Kasznel, A. J.; Zhang, Y.; Hai, Y. *J. Am. Chem. Soc*. **2017**, *139*, 9427−9430.

26. Davern, C.M.; Lowe, B.D.; Rosfi, A.; Ison, E.A.; and Proulx C. *Chem. Sci*., **2021**, *12*, 8401.

27. Zhang, Y.; Malamakal, R.M.; and Chenoweth, D.M. *J. Am. Chem. Soc*. **2015,** *137*, 12422−12425. b) Zhang, Y.; Herling, M.; and Chenoweth, D.M. *J. Am. Chem. Soc*. **2016**, *138*, 9751−9754.

28. Zhang, Y.; Malamakal, R.M.; and Chenoweth, D.M. *Angew.Chem.Int.Ed.,* **2015**, *54*,10826–10832.

29. Vass, E.; Hollo´si, M.; Besson, F.; and Buchet, R. *Chem. Rev*. **2003**, *103*, 1917−1954.

30. Bouakline. F. *Phys. Chem. Chem. Phys*., **2021**, *23*, 20509–20523.

31. Wagner, B.; Hiller, W.; Ohnob, H.; and Krause, N. *Org. Biomol. Chem*., **2016**, *14*, 1579–1583.

32. a) Urban, M.; Frank, M.; Hofmanova, M.; Cisarova, I.; and Vesely, J. *Org. Biomol. Chem*., **2017**, *15*, 9071–9076. b) Nguyen, T.; Coover, R.A.; Verghese, J.; Moran, R.G.; and Ellis, K.C. *ACS Med. Chem. Lett*. **2014,** *5*, 462−467.

33. Egholm, M.; Buchardt, O.; Nielsen, P. E. *J. Am. Chem. Soc*. **1992**, *114*, 1895–1897. (b) Egholm, M.; Nielsen, P. E.; Buchardt, O. *J. Am. Chem. Soc*. **1992**, *114*, 9677–9678. (c) Egholm, M.; Buchardt, O; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P.E. *Nature* **1993**, *365*, 566–568. (d) Nielsen, P. E. *Acc Chem Res 1999*, *32*, 624-630. e) Kumar, V.; Brodyagin, N.; and Rozners, E. *ChemBioChem*., **2020**, *21*, 3410–3416.

34. Gasparro, F.P.; Kolodny, Nancy. *J. Chem. Educ*. **1977,** *54*, 258. b) Akhmedov, N.G.; Myshakin, E.M.; and Hall, C.D. *Magn Reason Chem.* **2004,** *42*, 39-48.

35. Abraham, M. J.; Abraham, R. J.; Acree, Jr. W. E., Aliev, A. E.; J. Leo, A, J. and Whaley, L. *J. Org. Chem.* **2014**, *79*, 11075−11083.

36. Bruker, (2016). *APEX3* (v2016.1-0), *SAINT* (Version 8.18C) and *SADABS* (Version 2012/1). Bruker AXS Inc., Madison, Wisconsin, USA.

37. Sheldrick, G. M. *Acta Cryst*. **2015**, *A71*, 3-8.

38. Sheldrick, G. M. *Acta Cryst*. **2008**, *A64*, 112-122.

39. Newberry, R. W.; and Raines, R.T. *Nature Chemical Biology*., **2016,** *12*, 1084–1088.

40. a) Vass, E.; Hollo´si, M.; Besson, F.; and Buchet, R. *Chem. Rev*. **2003**, *103*, 1917−1954. b) Drewes, J.A.; and Rowlen, K.L. *Biophysical Journal*., **1993**, *65*, 985-991. c) lianc, C.; Rito, C.J.; and Cellman, S.C. *Biopolymers*., **1992**, *32*, 293-301.

41. (a) Becke, A. D. *J. Chem. Phys*. **1992**, *98*, 1372–1377. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988,** *37*, 785-790.

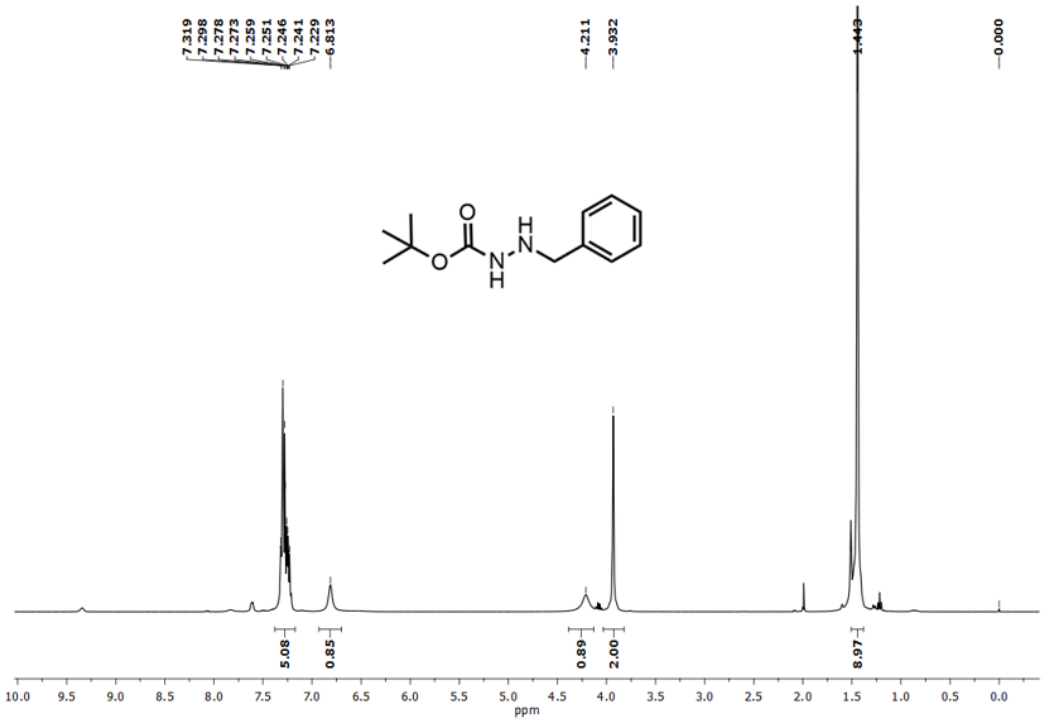
42. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.

43. Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO 3.0. Madison, WI, **1990**, as implemented in *Reference 10.*

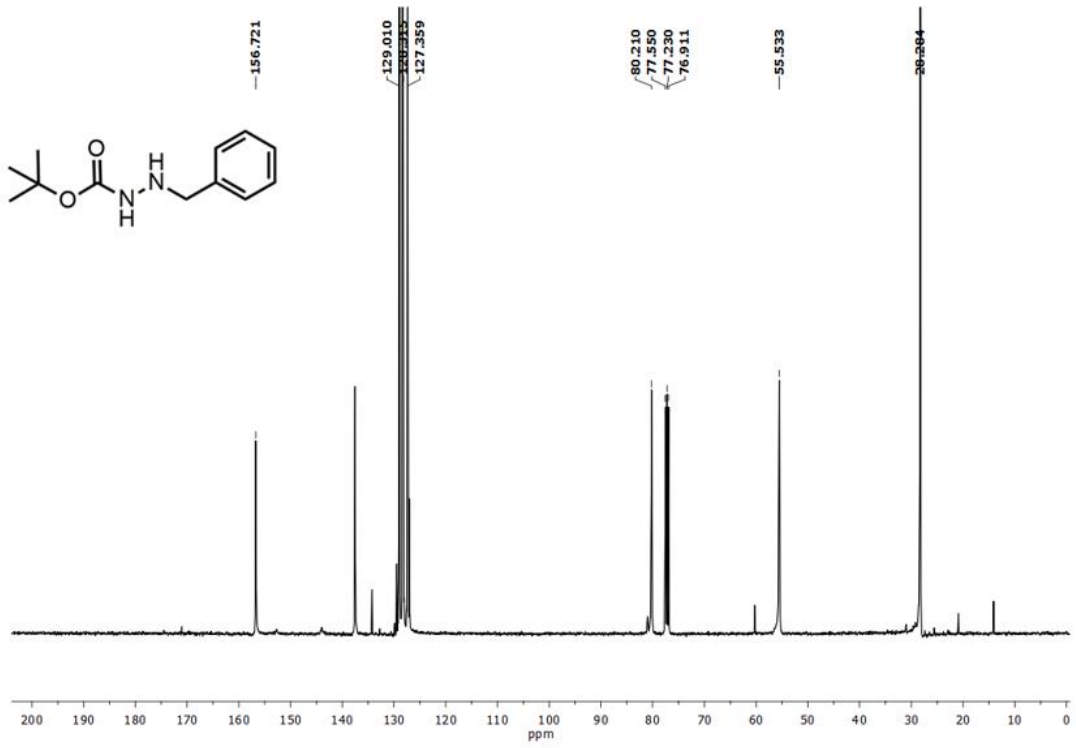
2.8. Characterisation data of the synthesized compounds and monomers (Appendix-I)

Entry	Table of contents	Page no
1	^{1}H NMR, ^{13}CNMR, IR and HRMS spectra of the monomers	122-145
2	NOSEY, HSQC and 2D-NMR of monomers	146-153

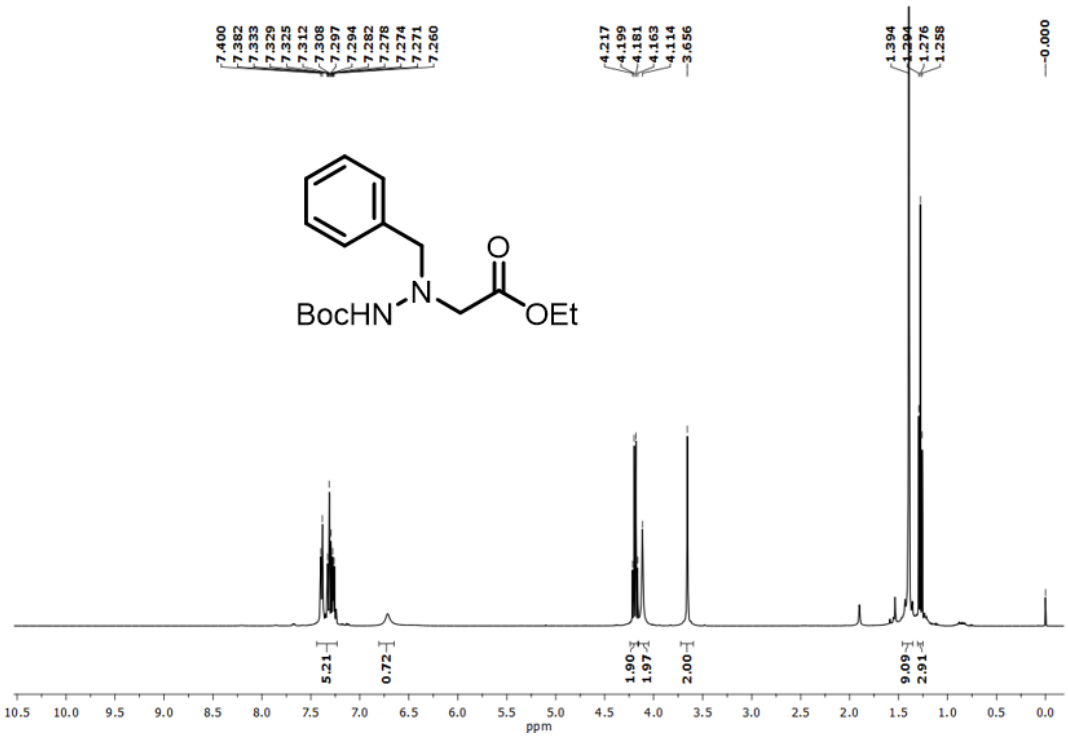
^{1}H NMR of compound 11 in $CDCl_3$ (400 MHz)



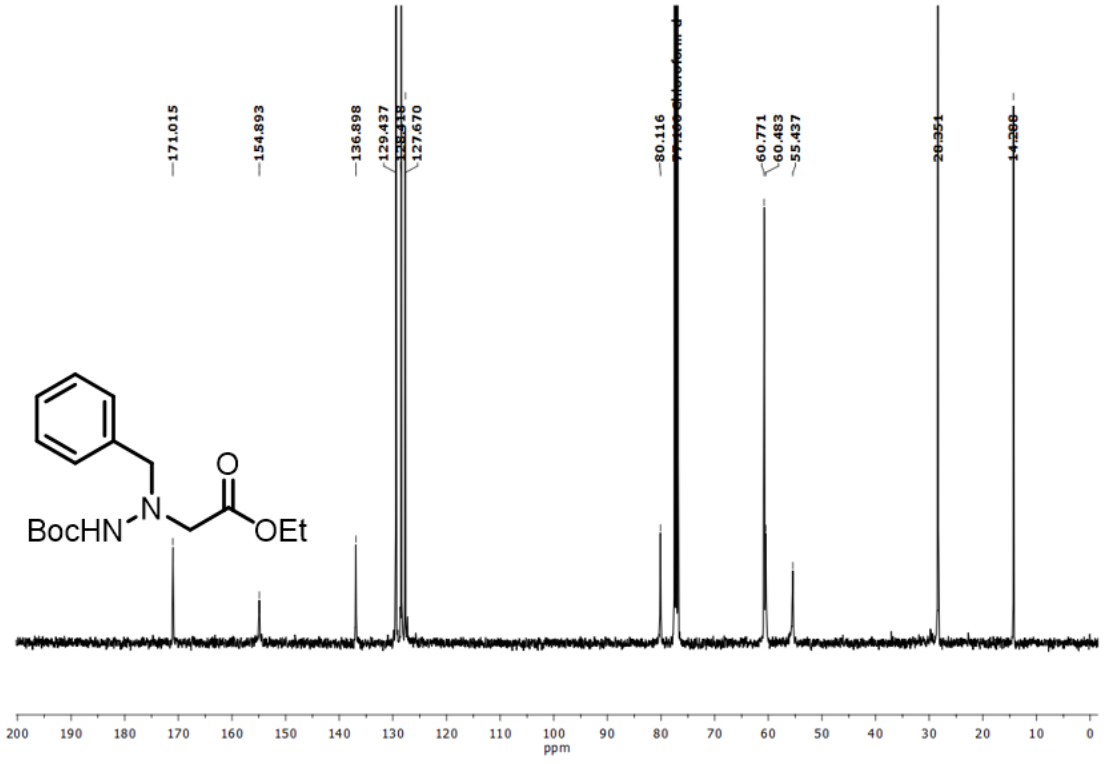
^{13}C {1H} NMR of compound 11 in $CDCl_3$ (101 MHz)



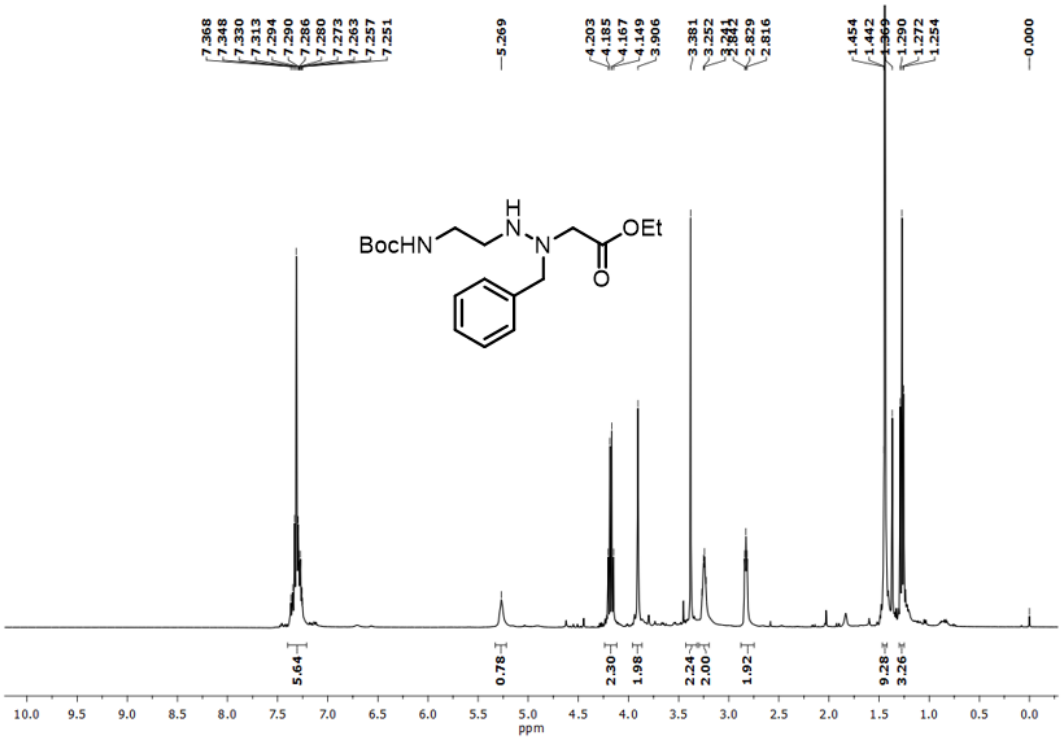
^{1}H NMR of compound 12a in $CDCl_3$ (400 MHz)



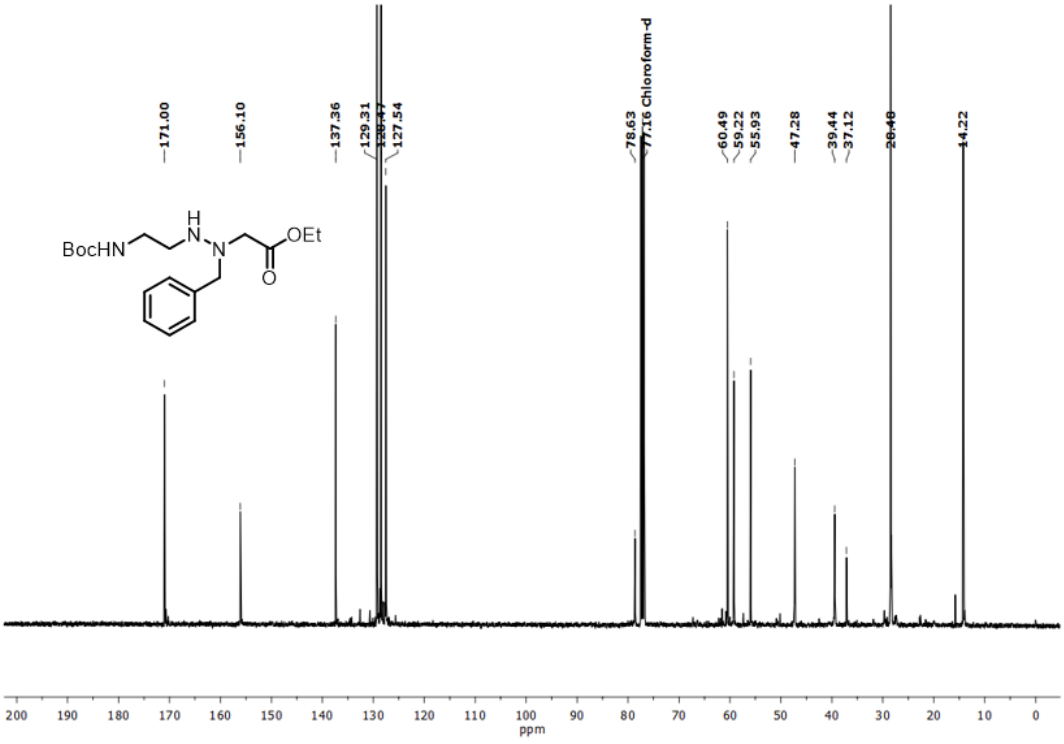
^{13}C {1H} NMR of compound 12a in $CDCl_3$ (101 MHz)



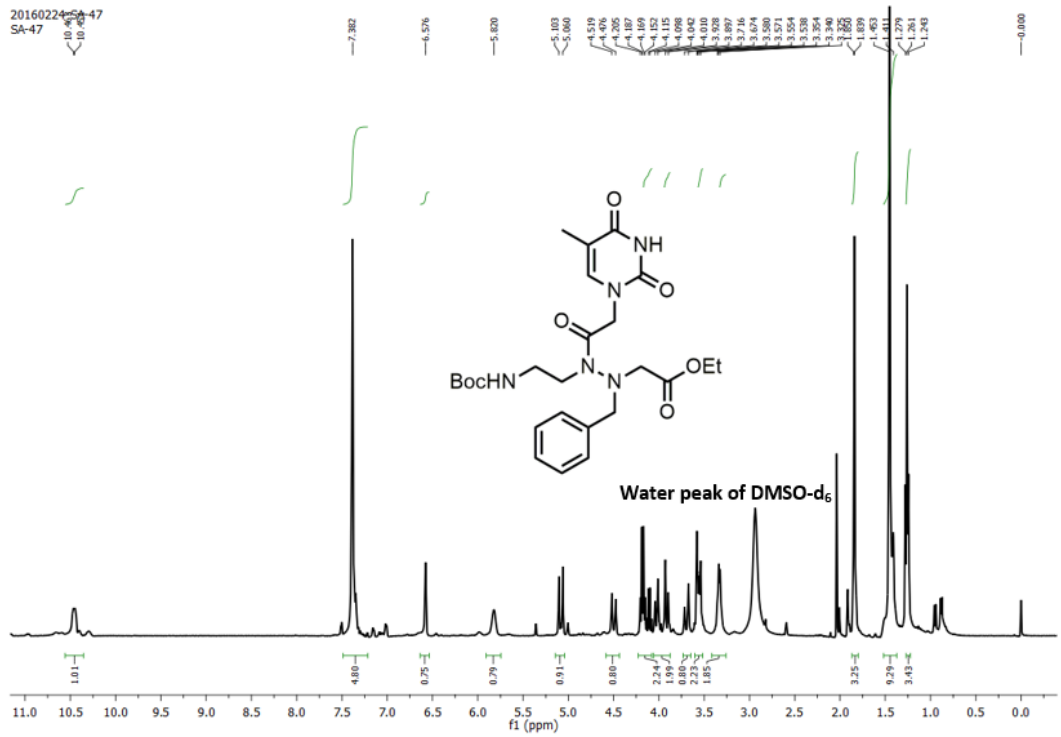
^{1}H NMR of compound 13a in $CDCl_3$ (400 MHz)



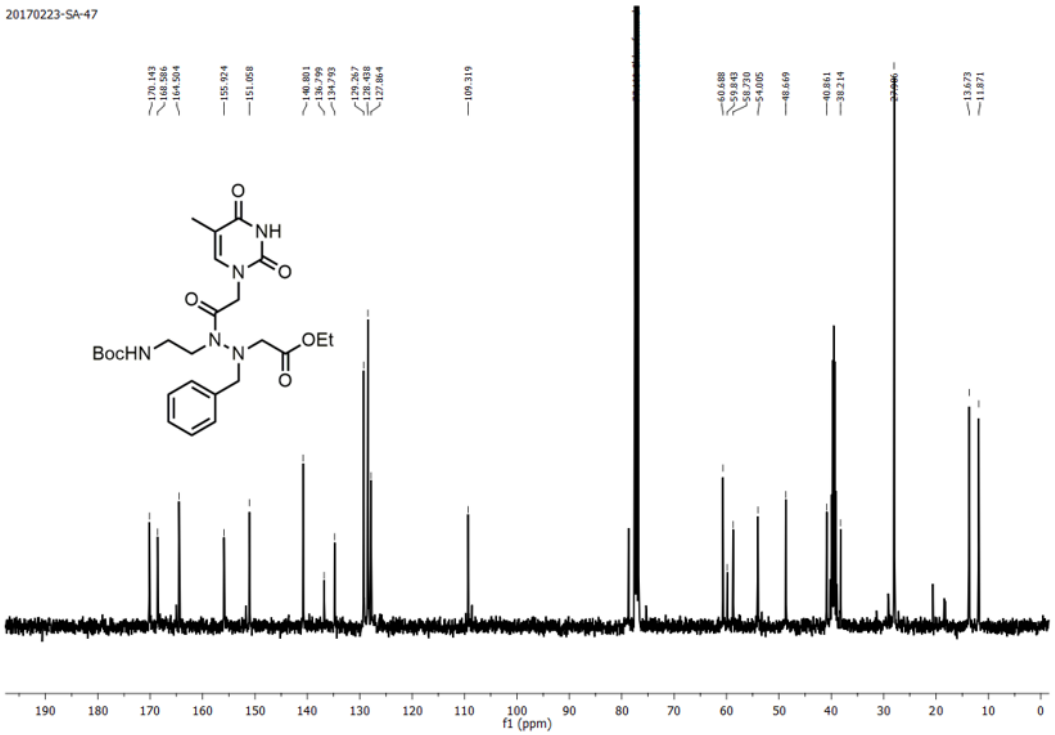
^{13}C {1H} NMR of compound 13a in $CDCl_3$ (101 MHz)



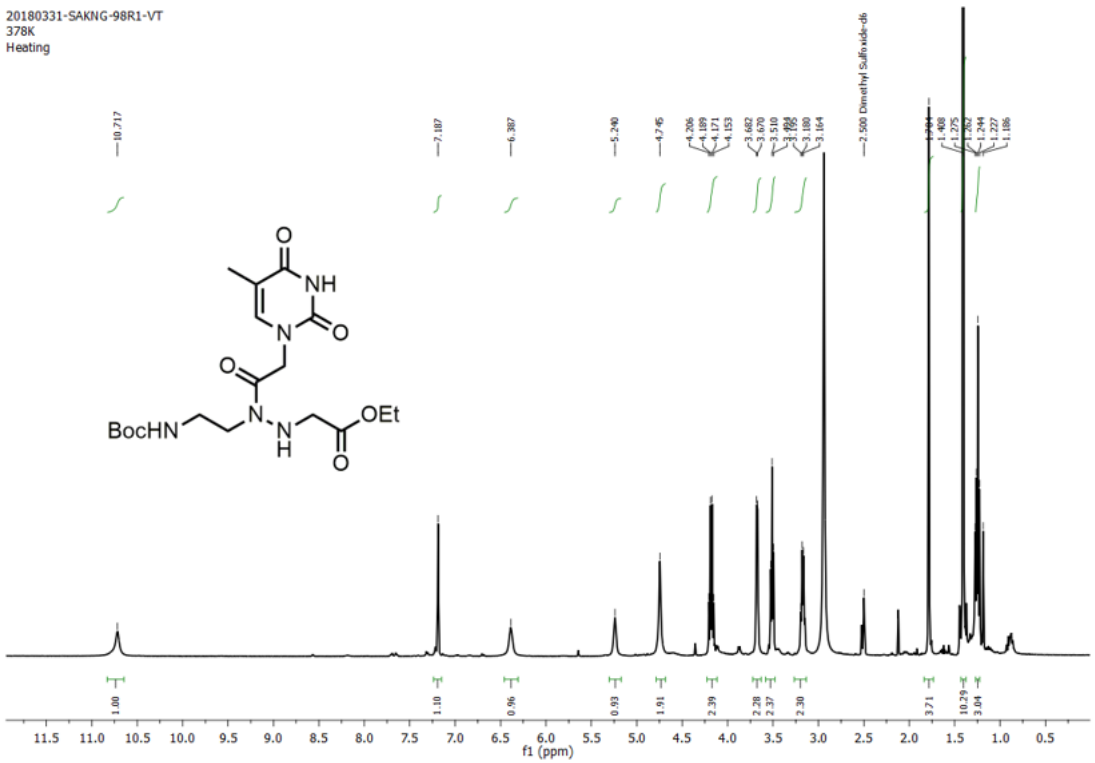
^{1}H NMR of compound 14a in DMSO-d$_6$ (400 MHz)



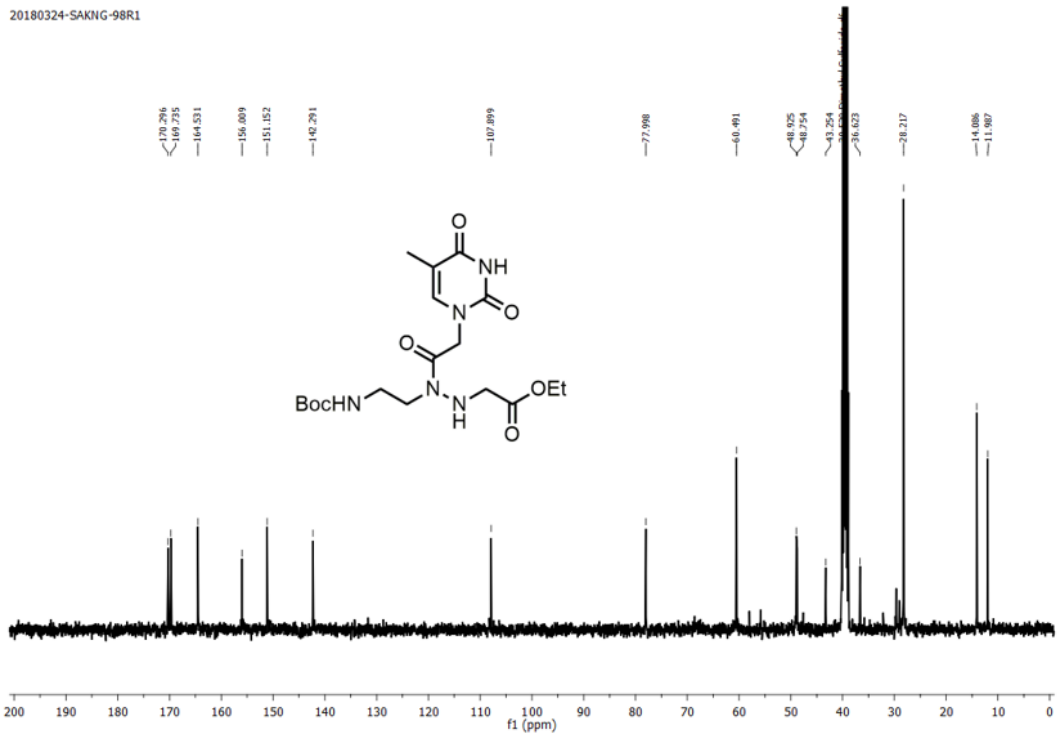
^{13}C {1H} NMR of compound 14a in CDCl$_3$ (101 MHz)



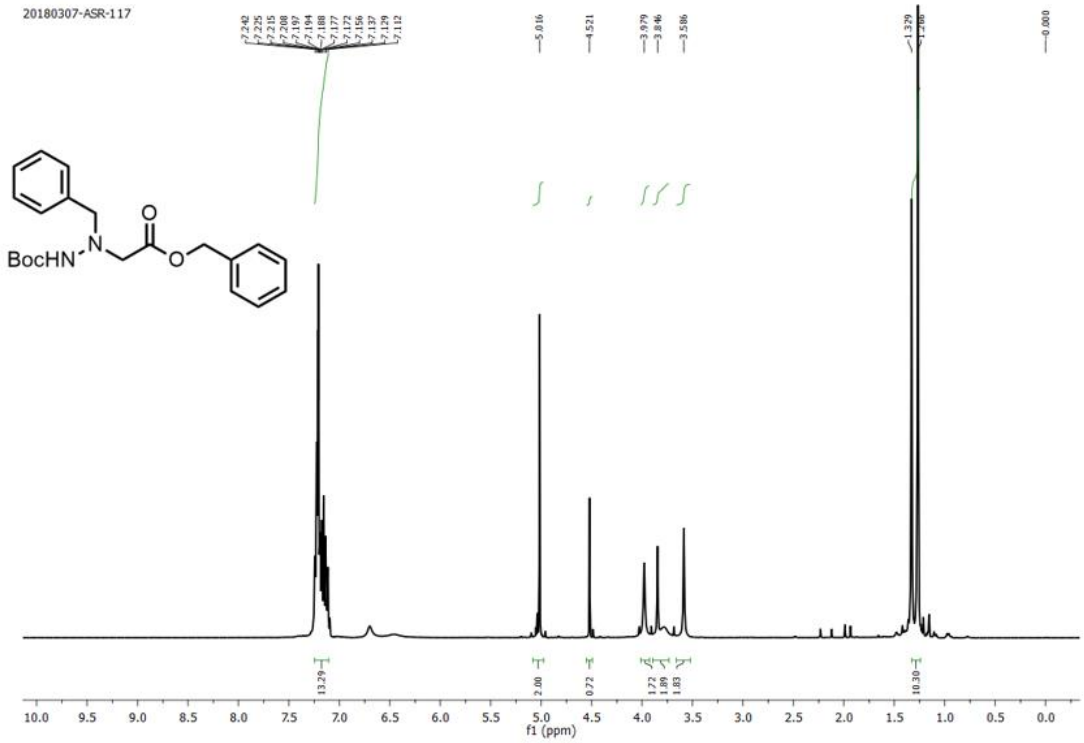
^{1}H NMR of compound 2 in DMSO-d$_6$ (400 MHz)



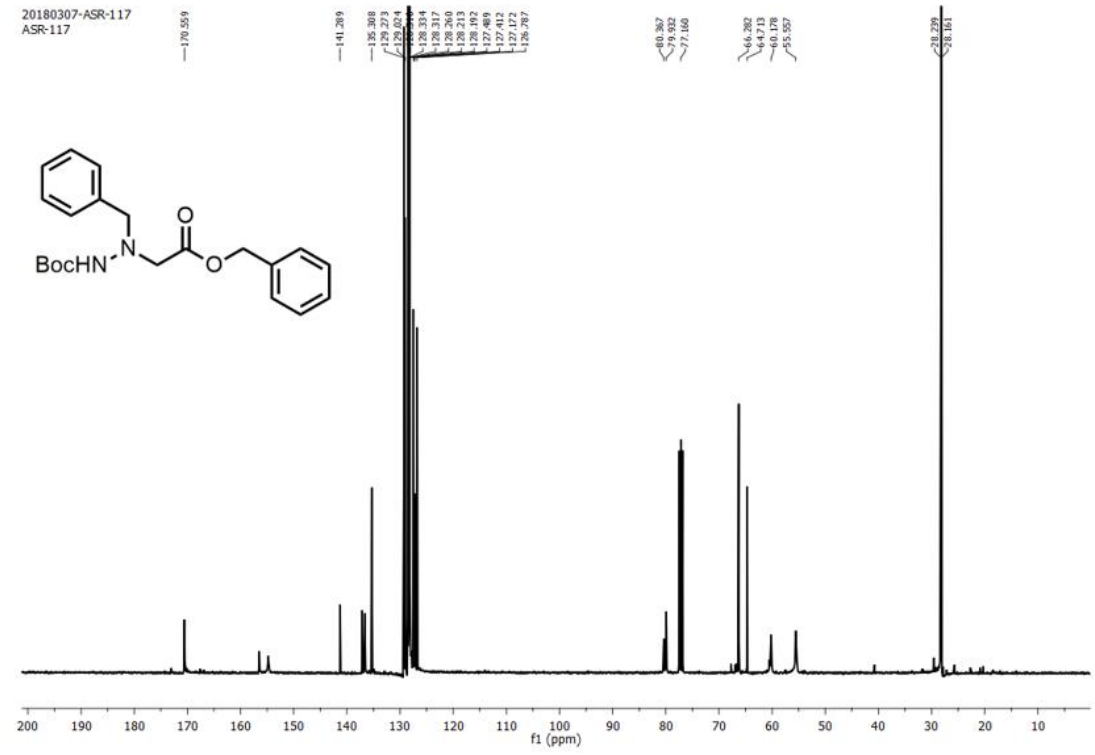
^{13}C {1H} NMR of compound 2 in DMSO-d$_6$ (101 MHz)



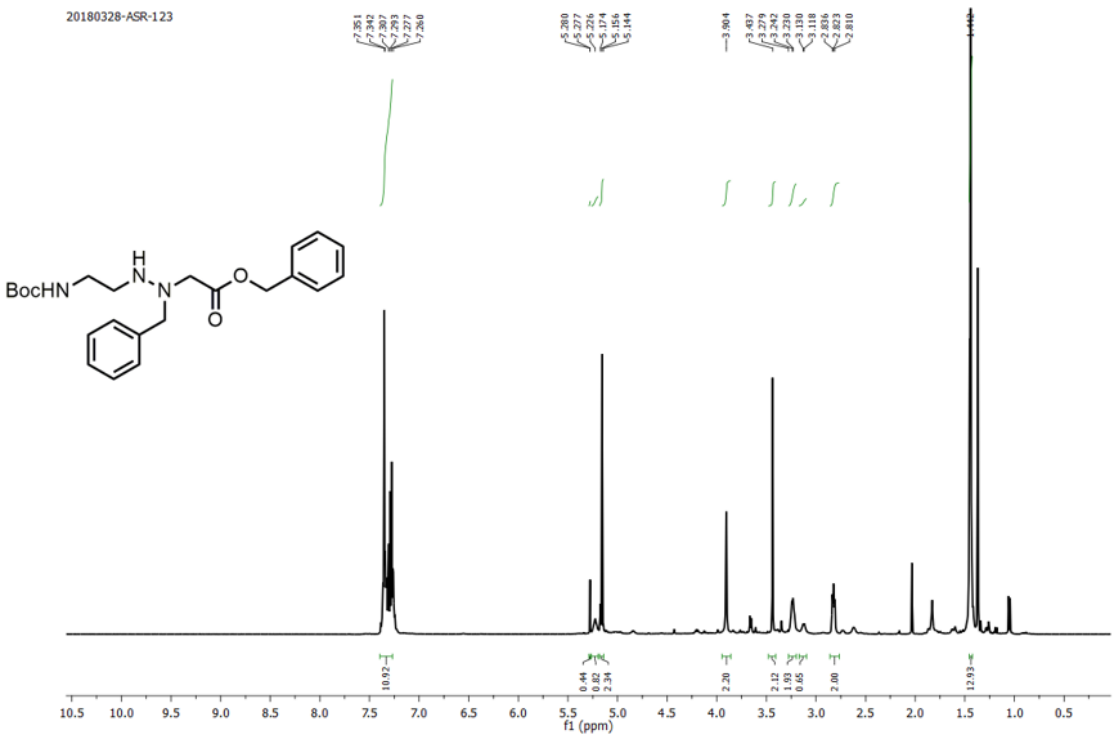
^{1}H NMR of compound 12b in $CDCl_3$ (400 MHz)



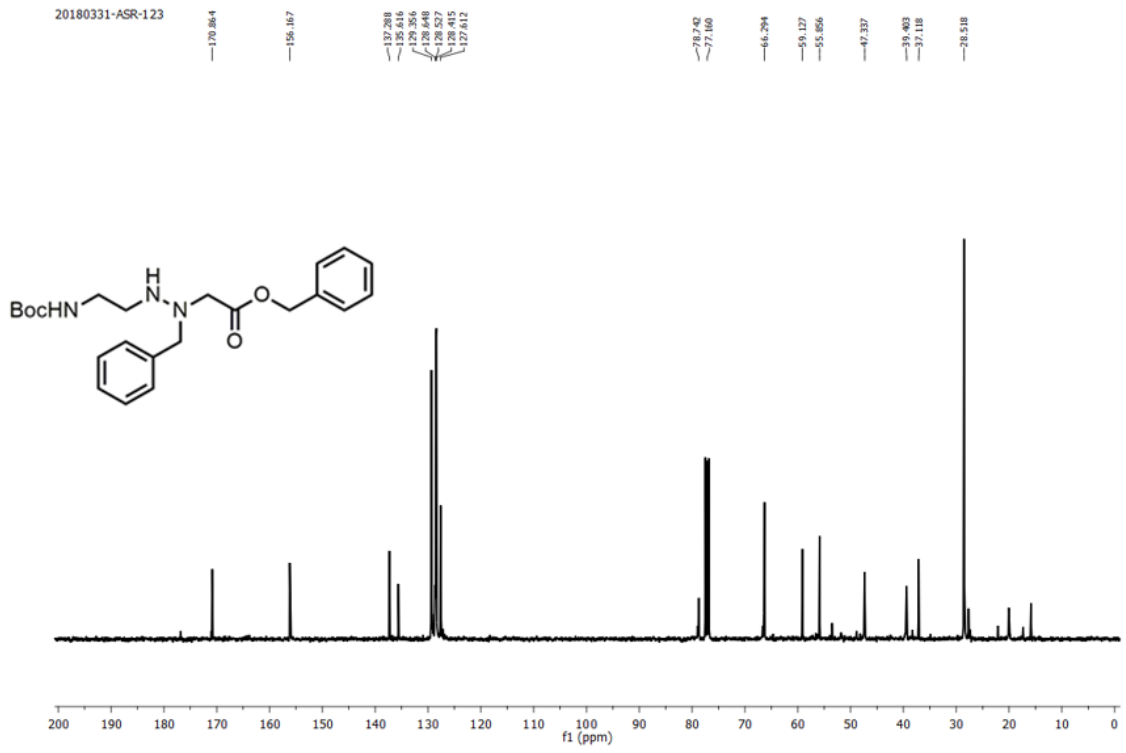
^{13}C {1H} NMR of compound 12b in $CDCl_3$ (101 MHz)



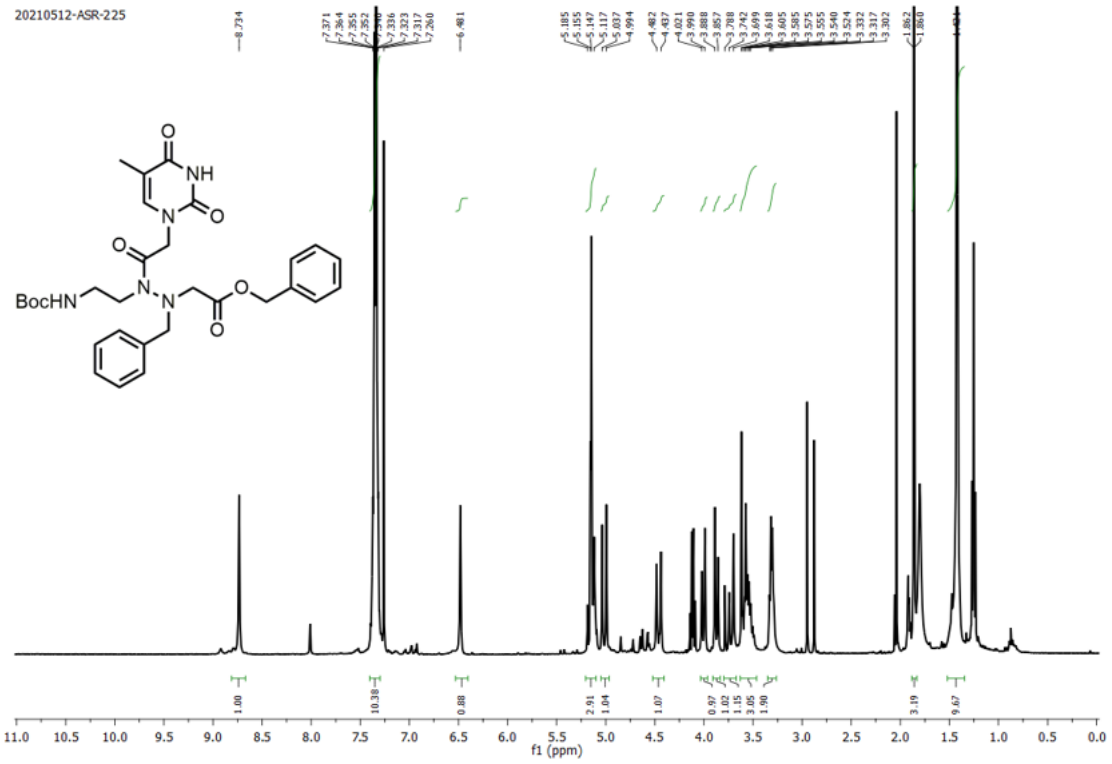
^{1}H NMR of compound 13b in $CDCl_3$ (400 MHz)



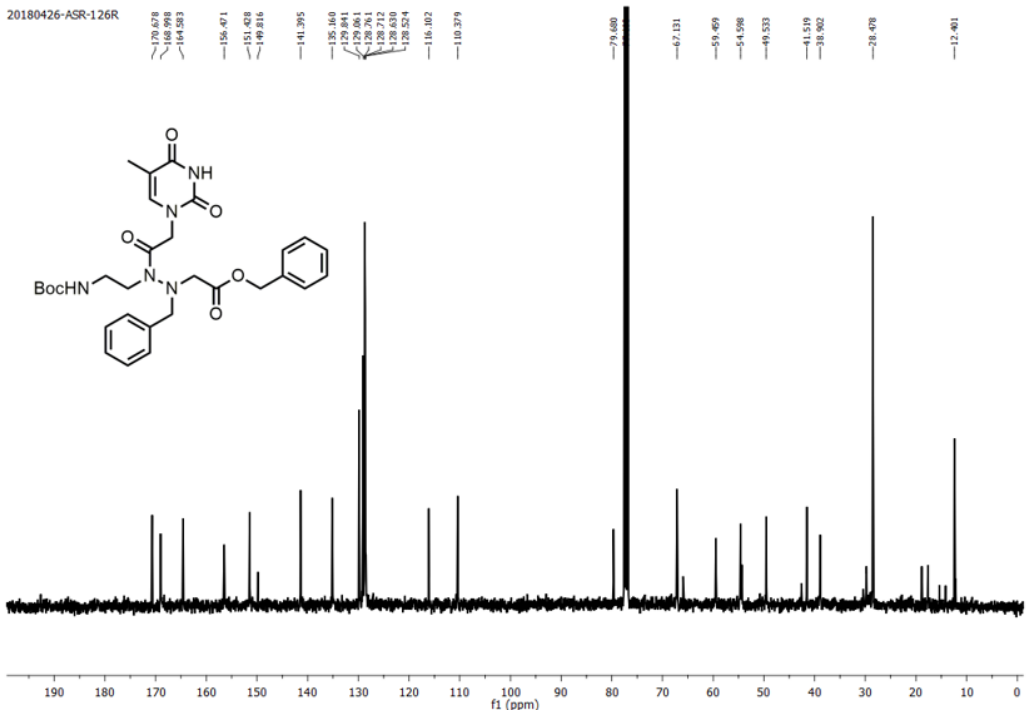
^{13}C {1H} NMR of compound 13b in $CDCl_3$ (101 MHz)



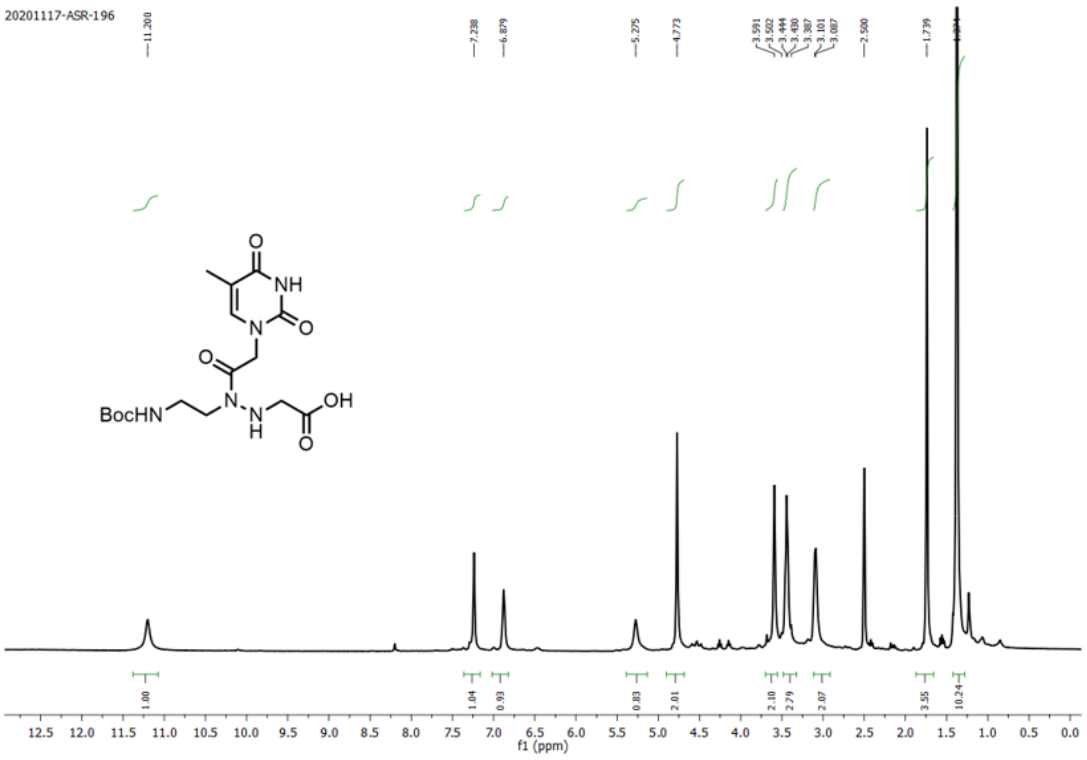
^{1}H NMR of compound 14b in $CDCl_3$ (400 MHz)



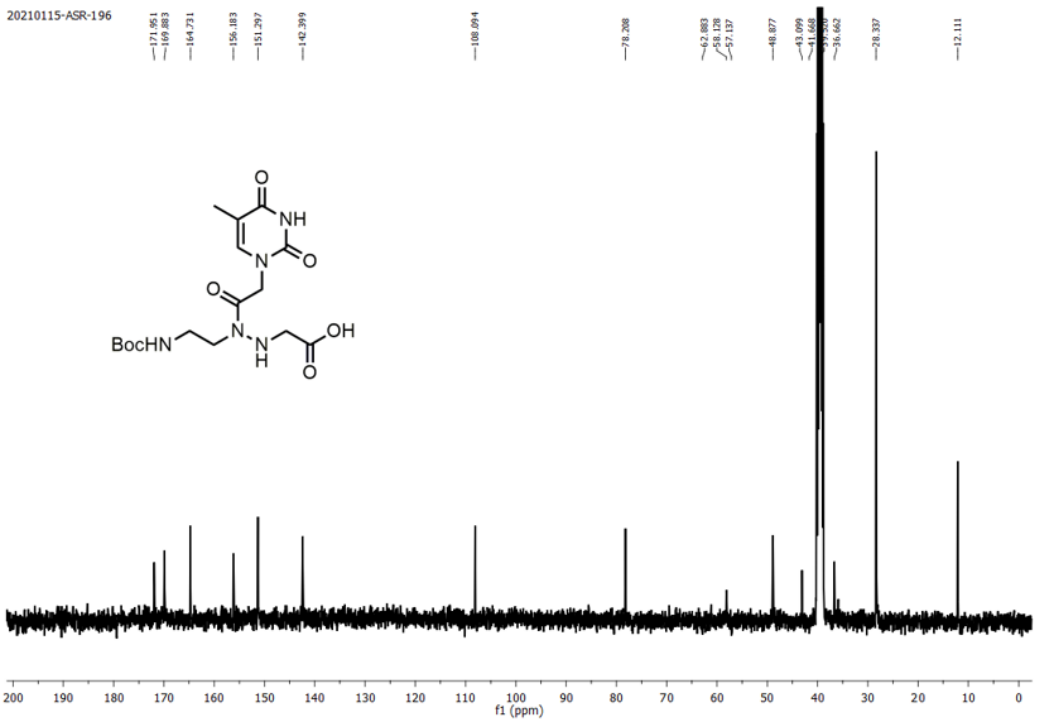
^{13}C {1H} NMR of compound 14b in $CDCl_3$ (101 MHz)



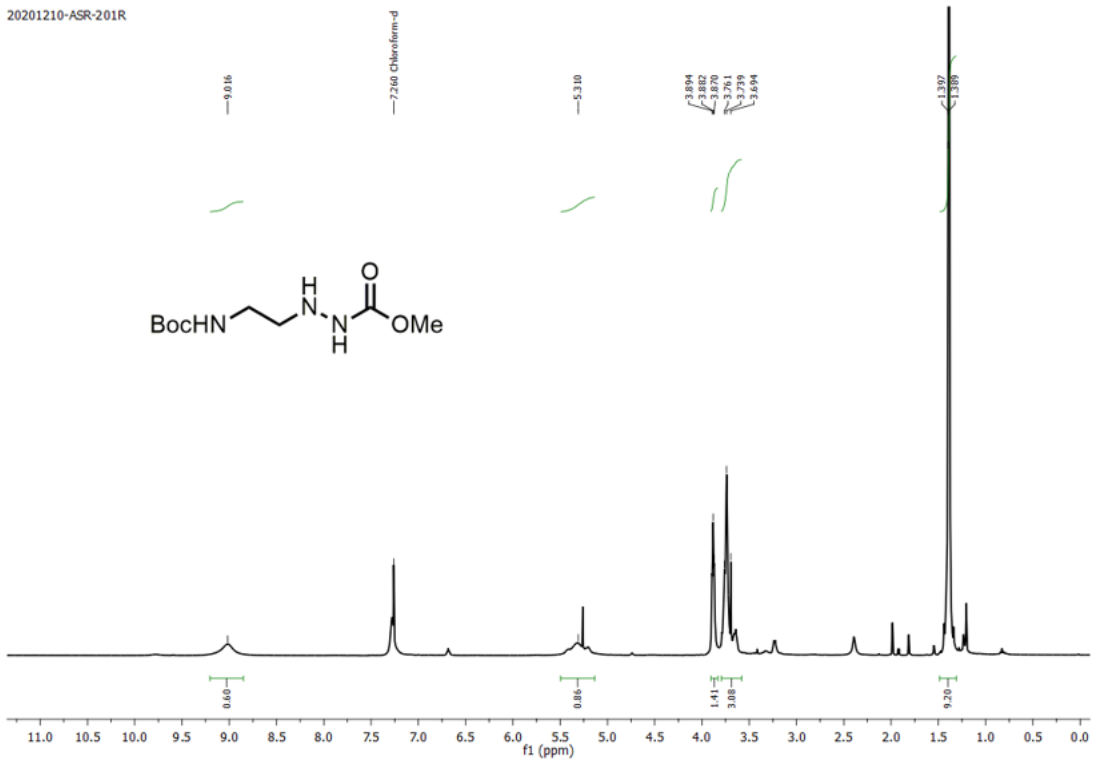
^{1}H NMR of compound 15 in DMSO-d$_6$ (400 MHz)



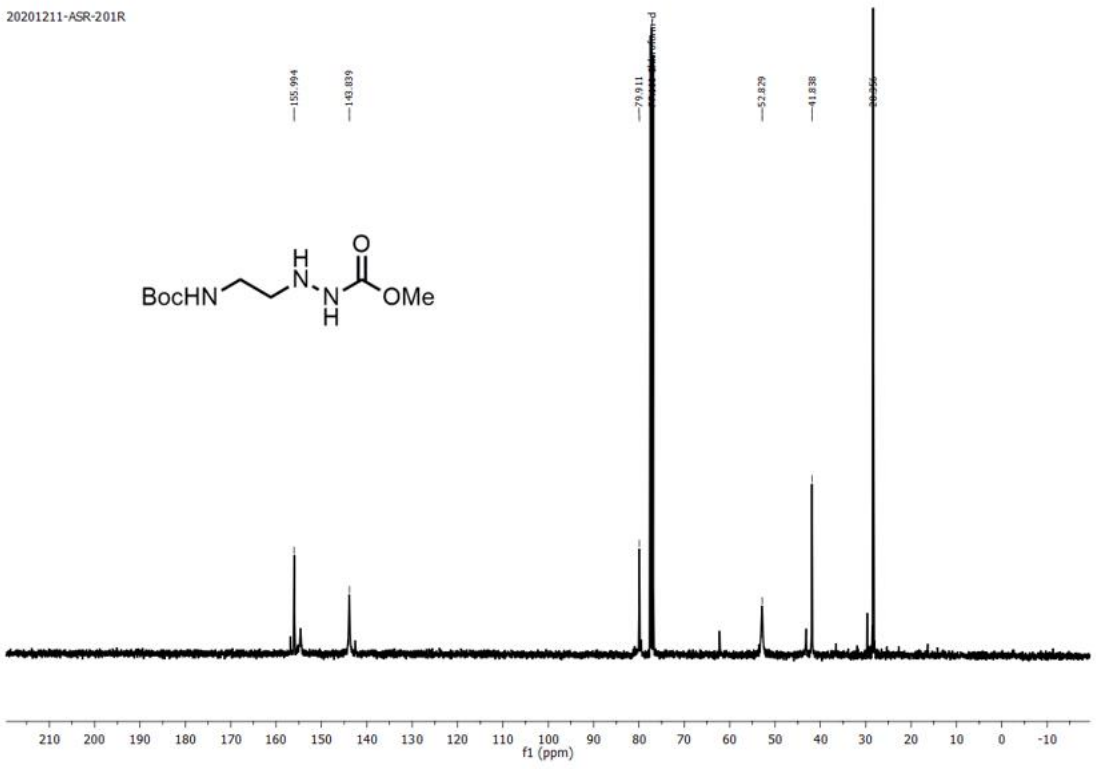
^{13}C {1H} NMR of compound 15 in DMSO-d$_6$ (101 MHz)



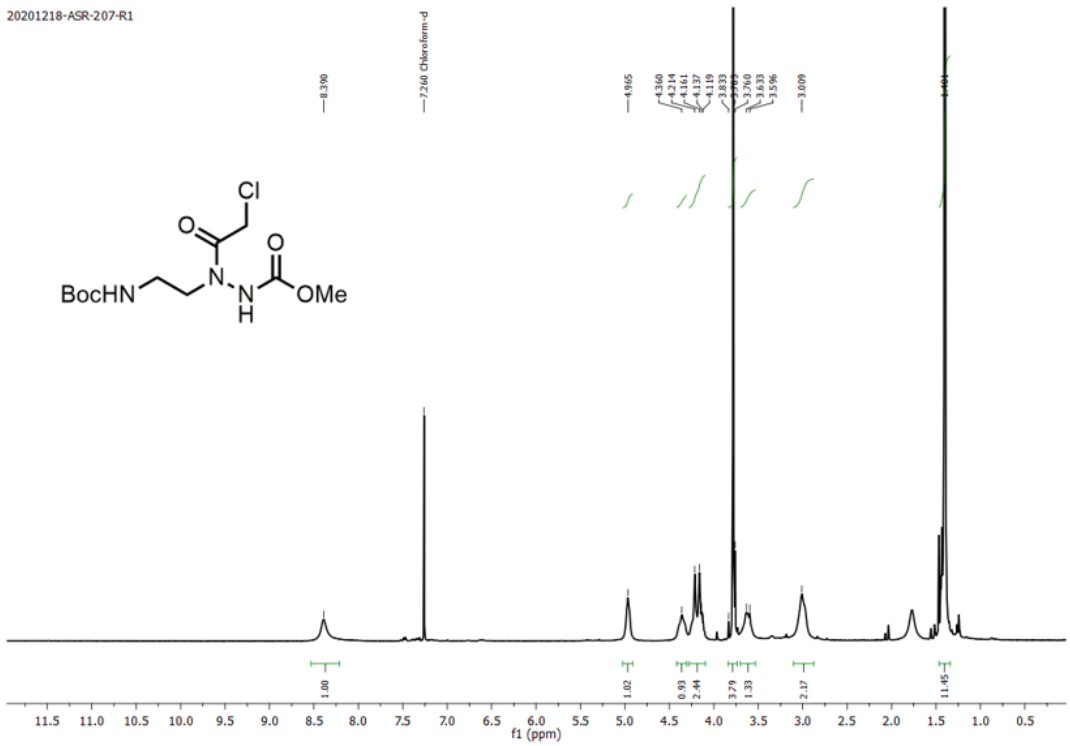
^{1}H NMR of compound 6 in $CDCl_3$ (400 MHz)



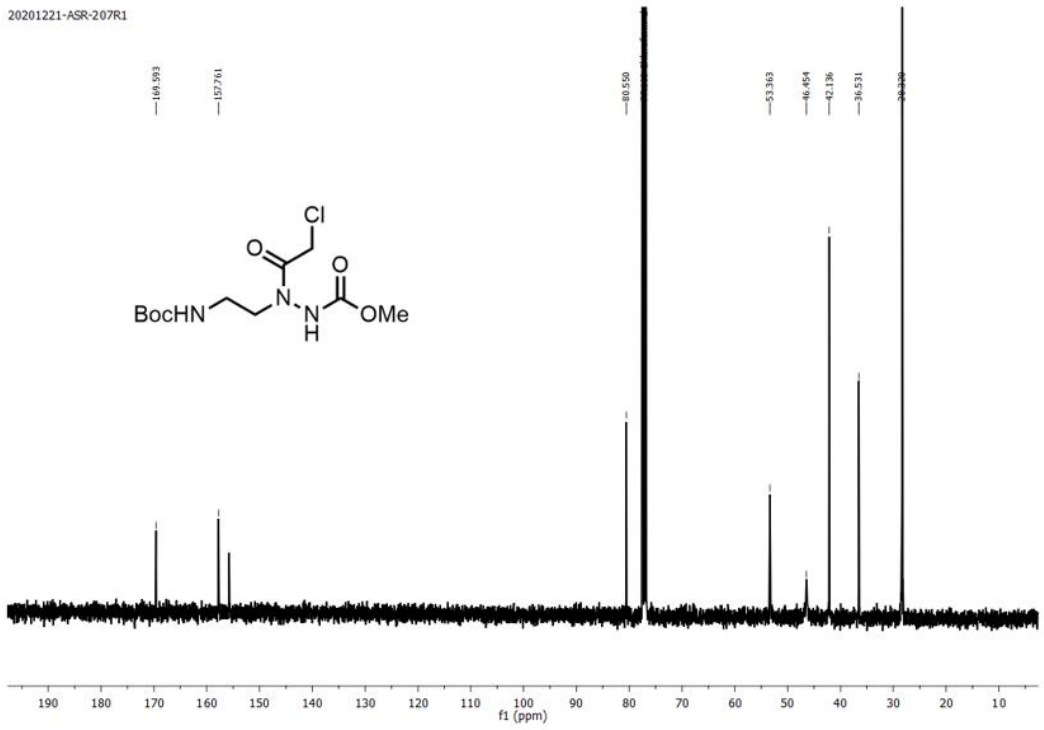
^{13}C {1H} NMR of compound 6 in $CDCl_3$ (101 MHz)



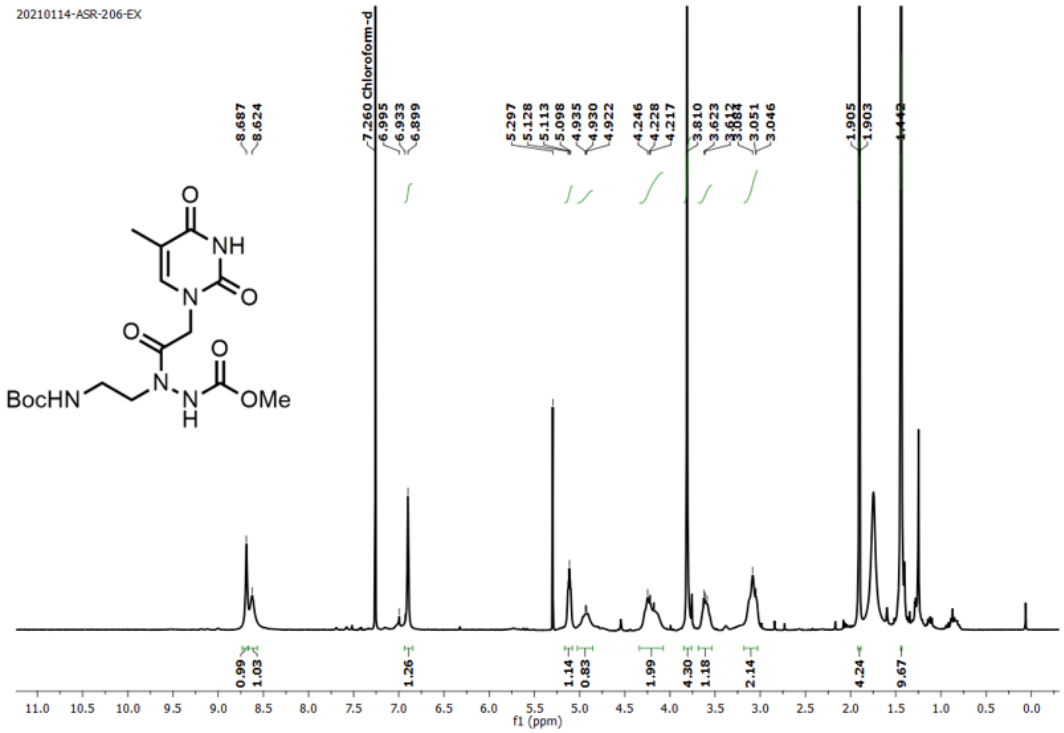
^{1}H NMR of compound 7 in $CDCl_3$ (400 MHz)



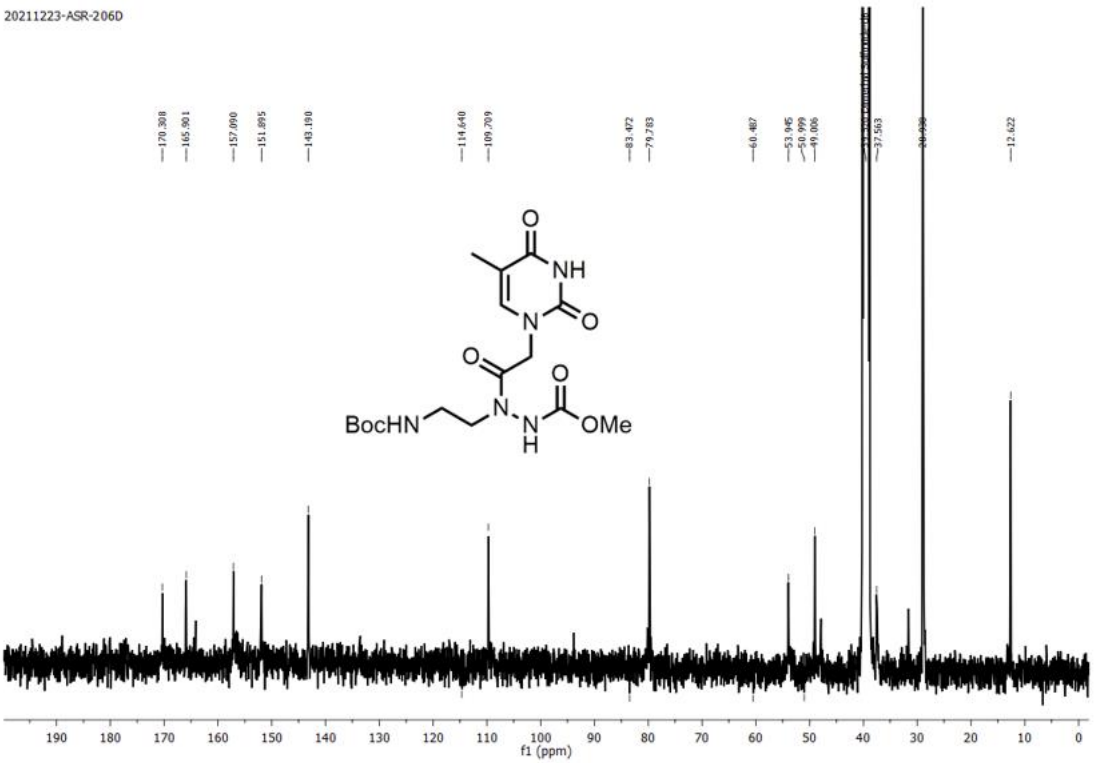
^{13}C {1H} NMR of compound 7 in $CDCl_3$ (101 MHz)



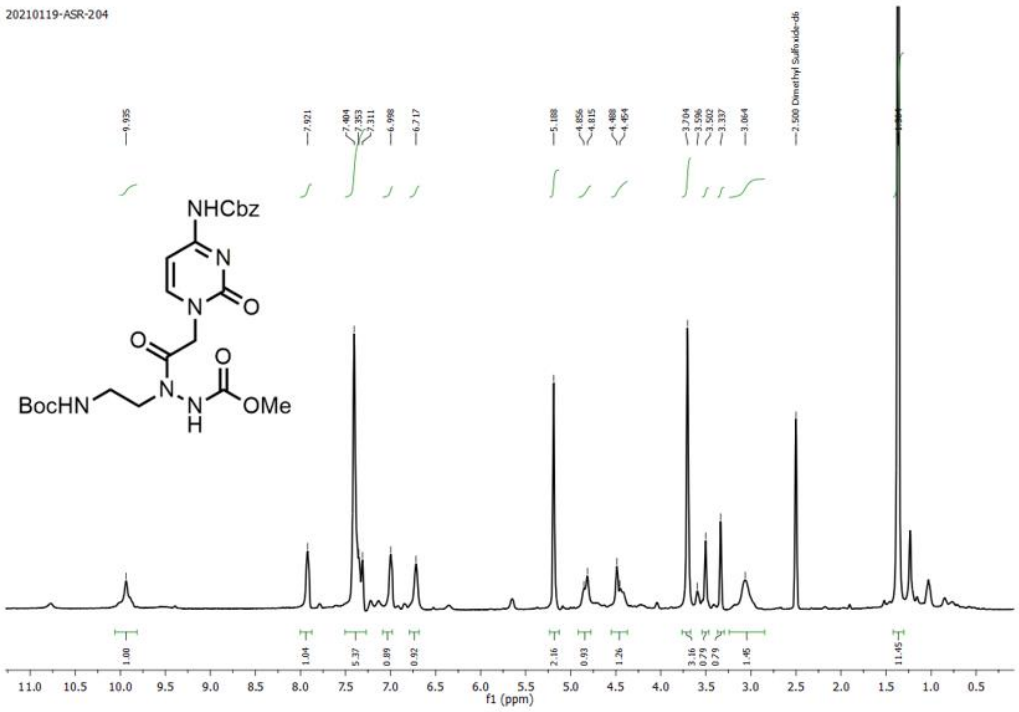
^{1}H NMR of compound 1 in $CDCl_3$ (400 MHz)



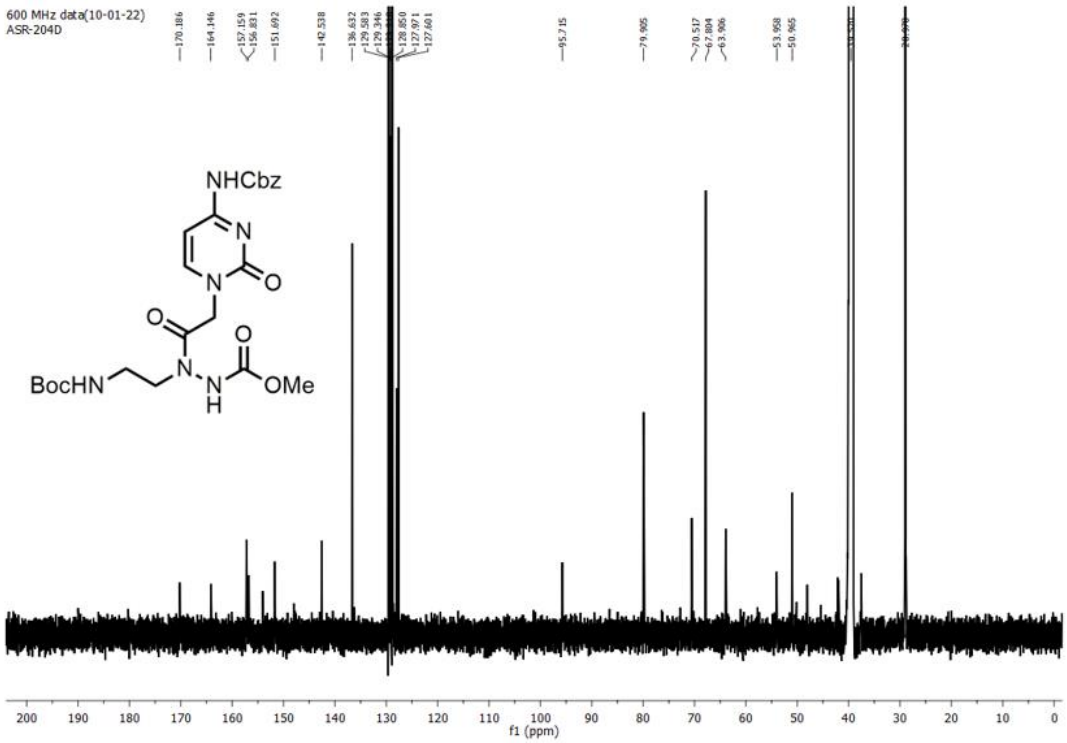
^{13}C {1H} NMR of compound 1 in $CDCl_3$ (101 MHz)



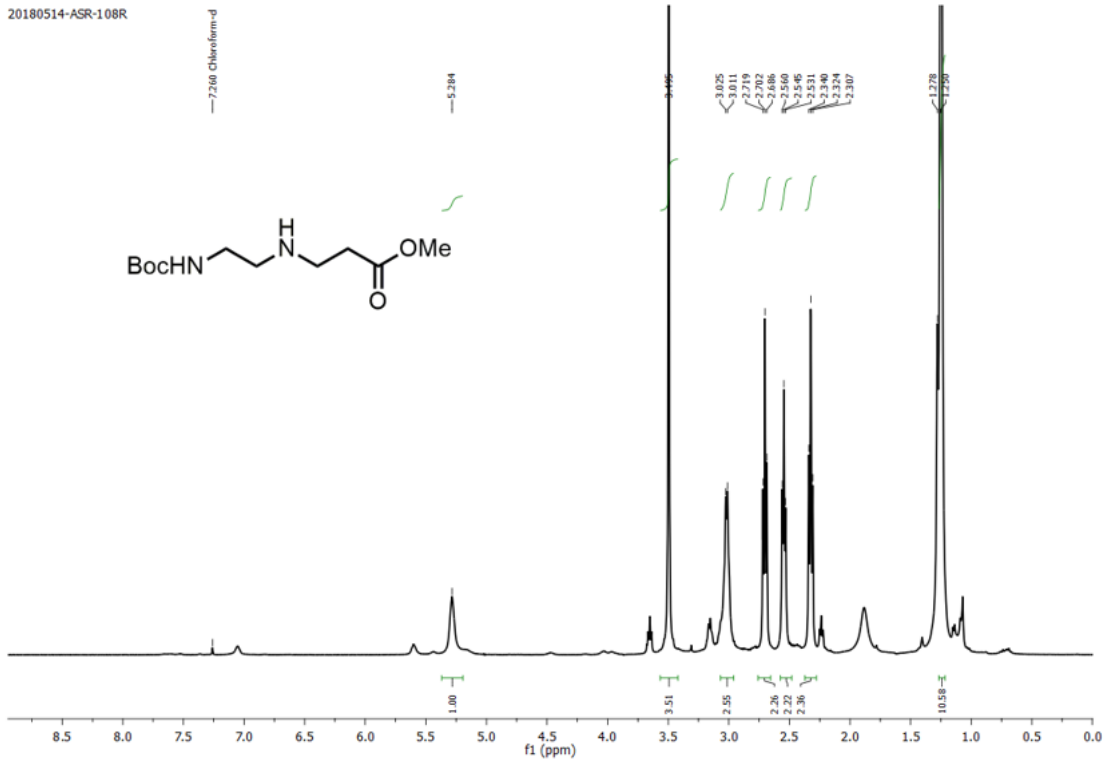
^{1}H NMR of compound 1C in DMSO-d$_6$ (400 MHz)



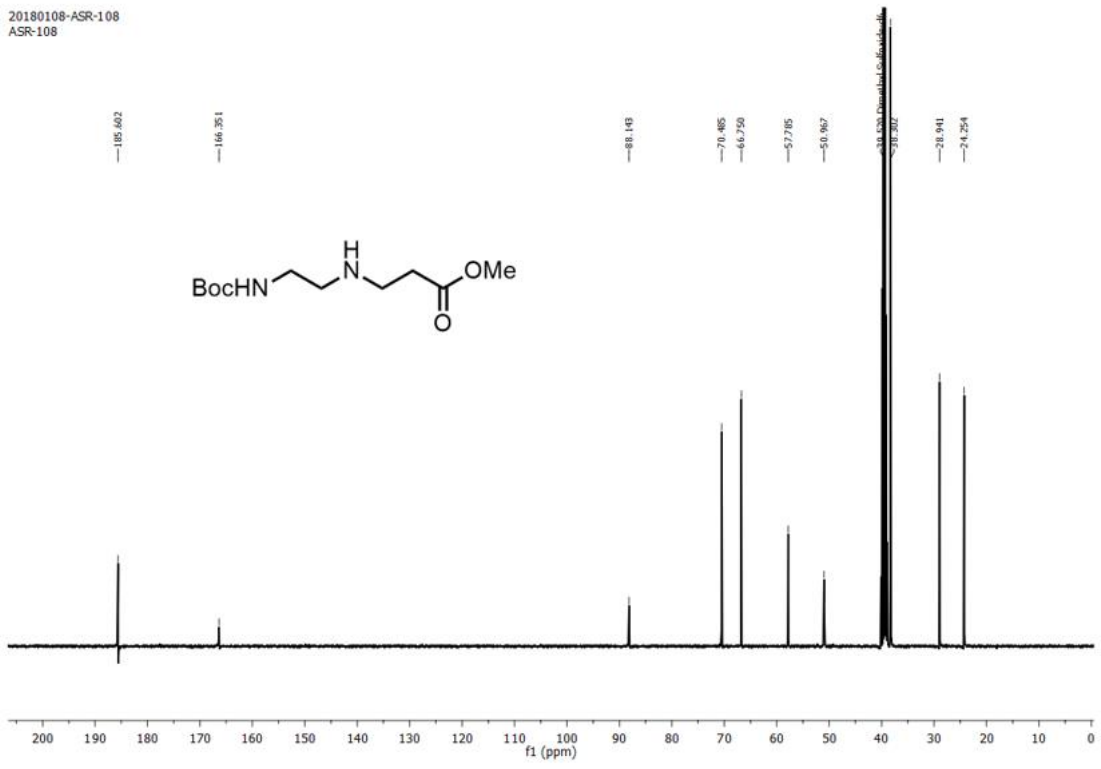
^{13}C {1H} NMR of compound 1C in DMSO-d$_6$ (101 MHz)



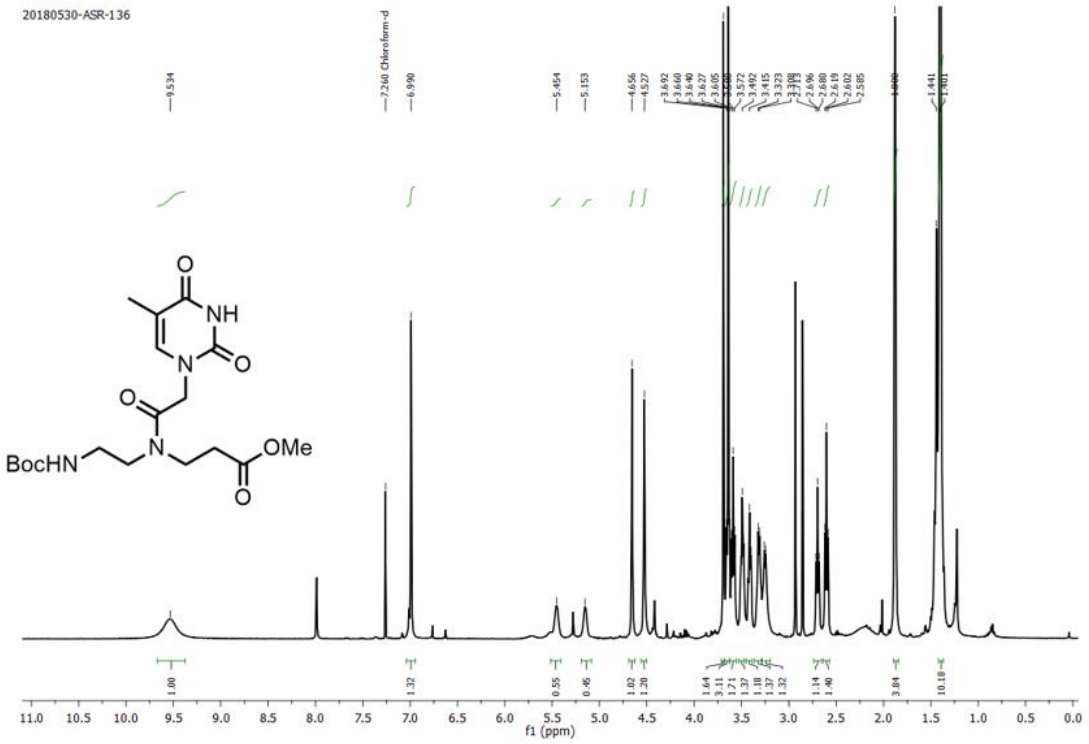
^{1}H NMR of compound 18 in $CDCl_3$ (400 MHz)



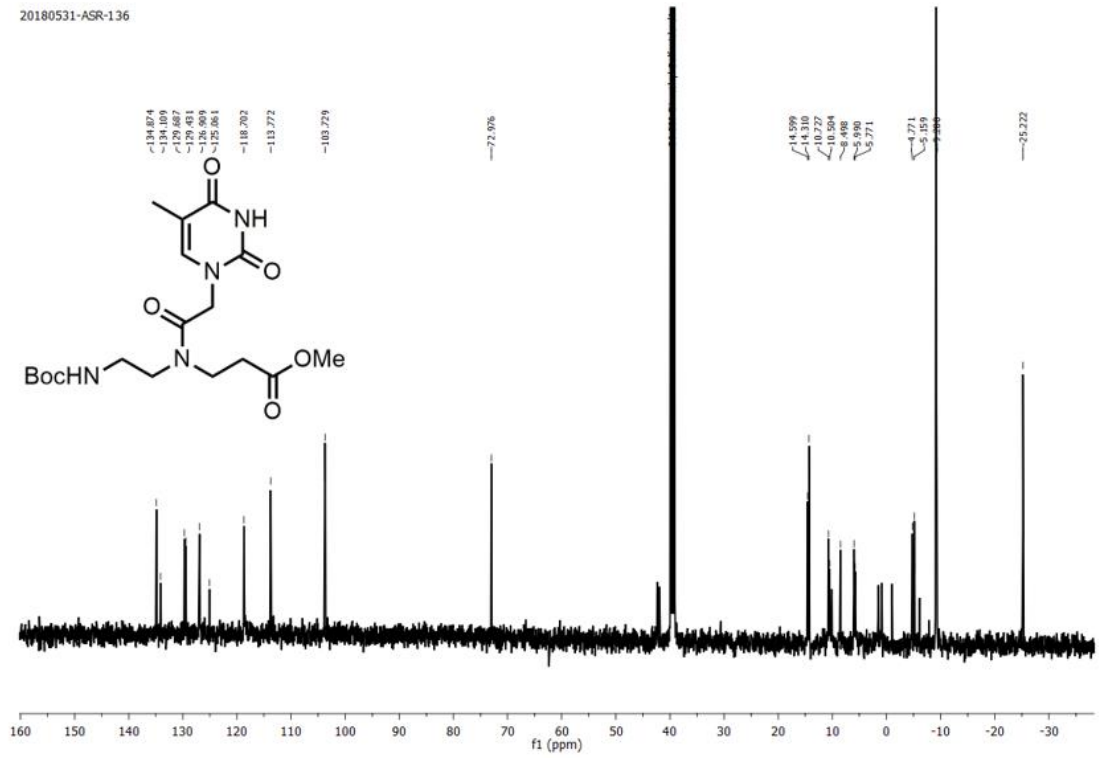
^{13}C {1H} NMR of compound 18 in $CDCl_3$ (101 MHz)



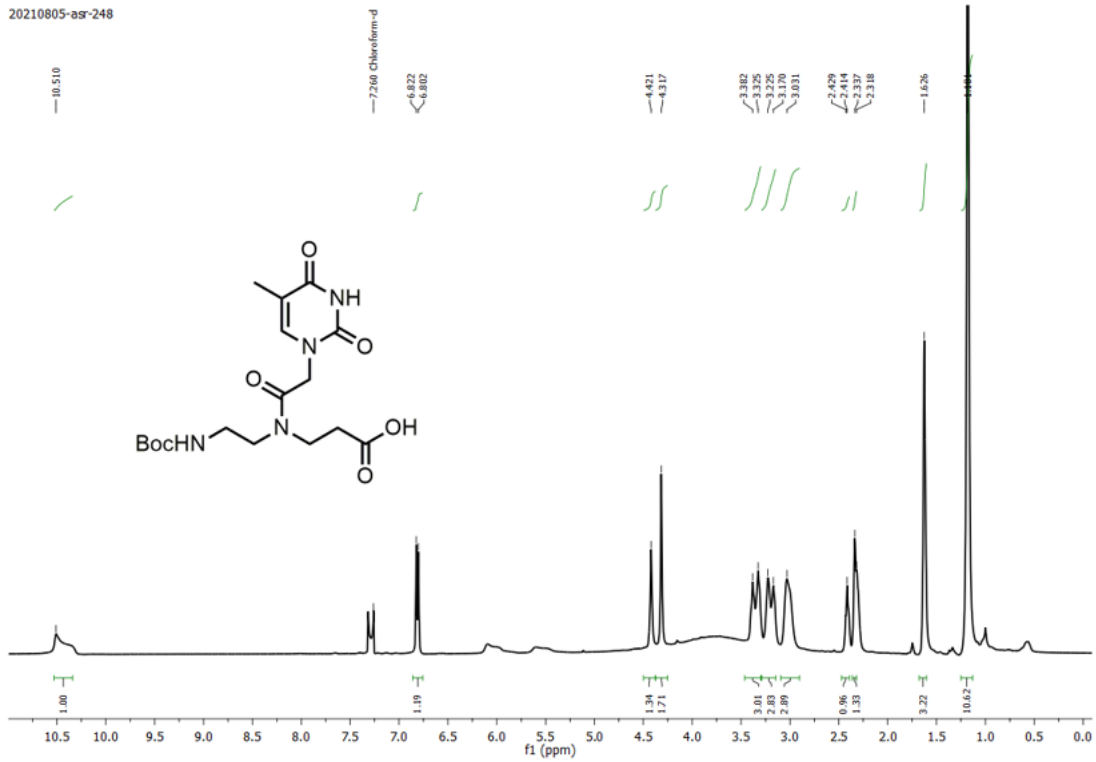
^{1}H NMR of compound 4 in $CDCl_3$ (400 MHz)



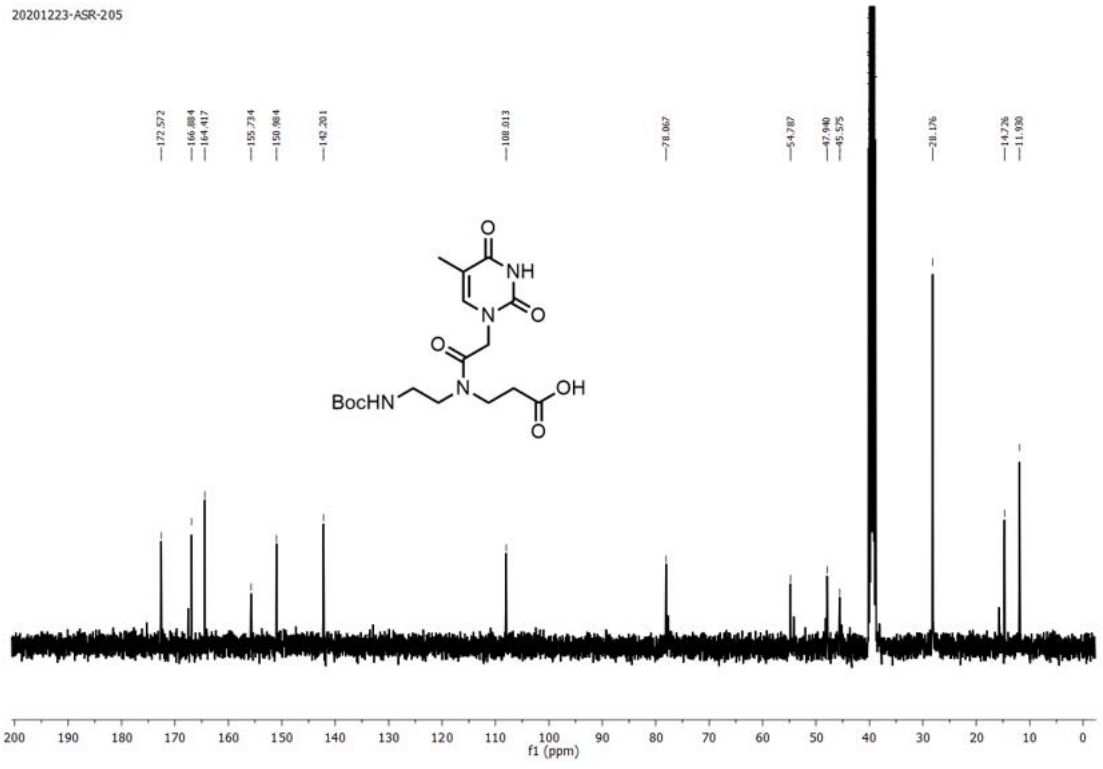
^{13}C {1H} NMR of compound 4 in $CDCl_3$ (101 MHz)



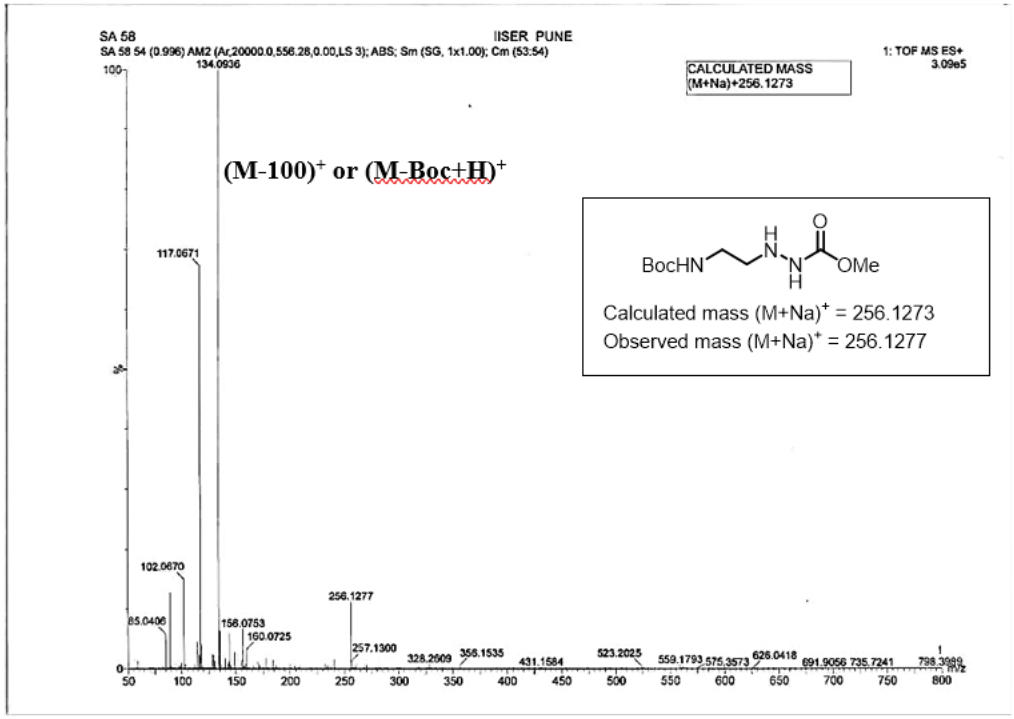
^{1}H NMR of compound 20 in $CDCl_3$ (400 MHz)



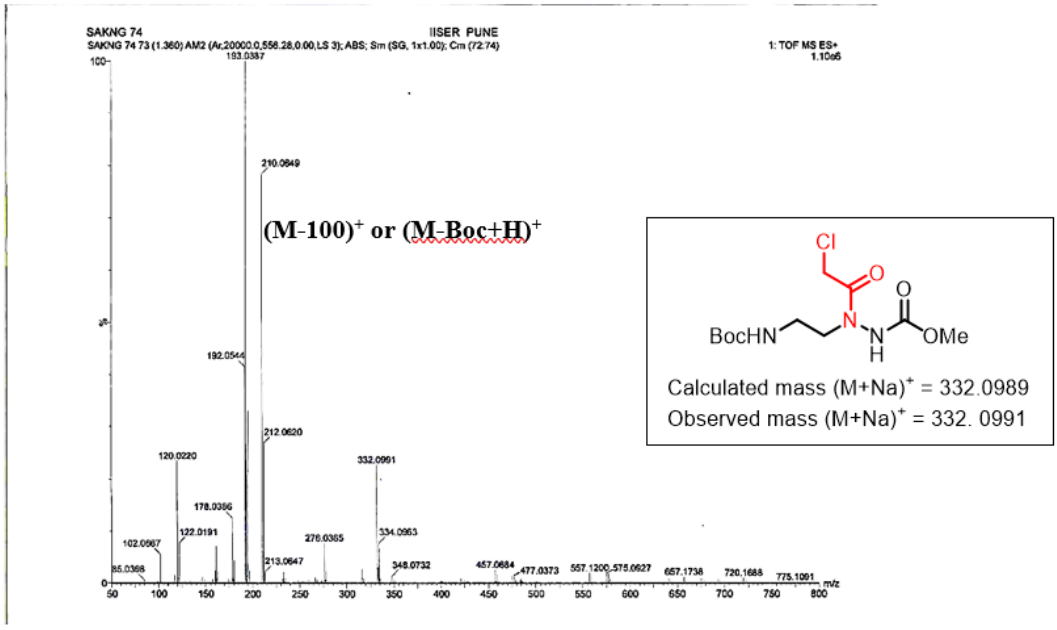
^{13}C {1H} NMR of compound 20 in $CDCl_3$ (101 MHz)



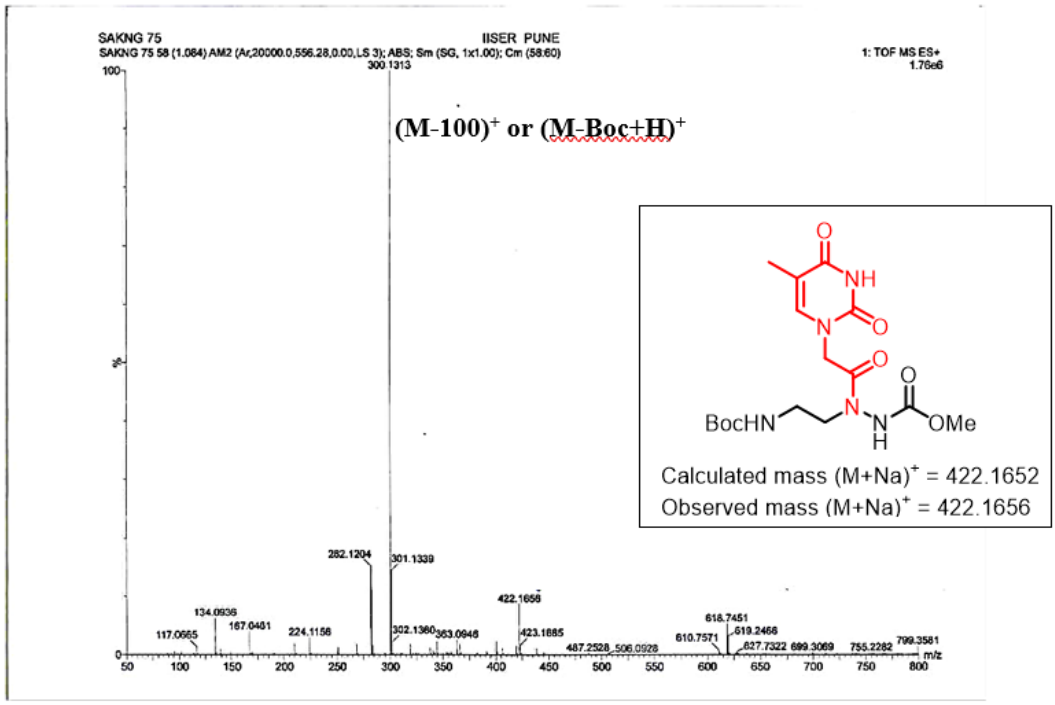
Mass spectra of compound 6



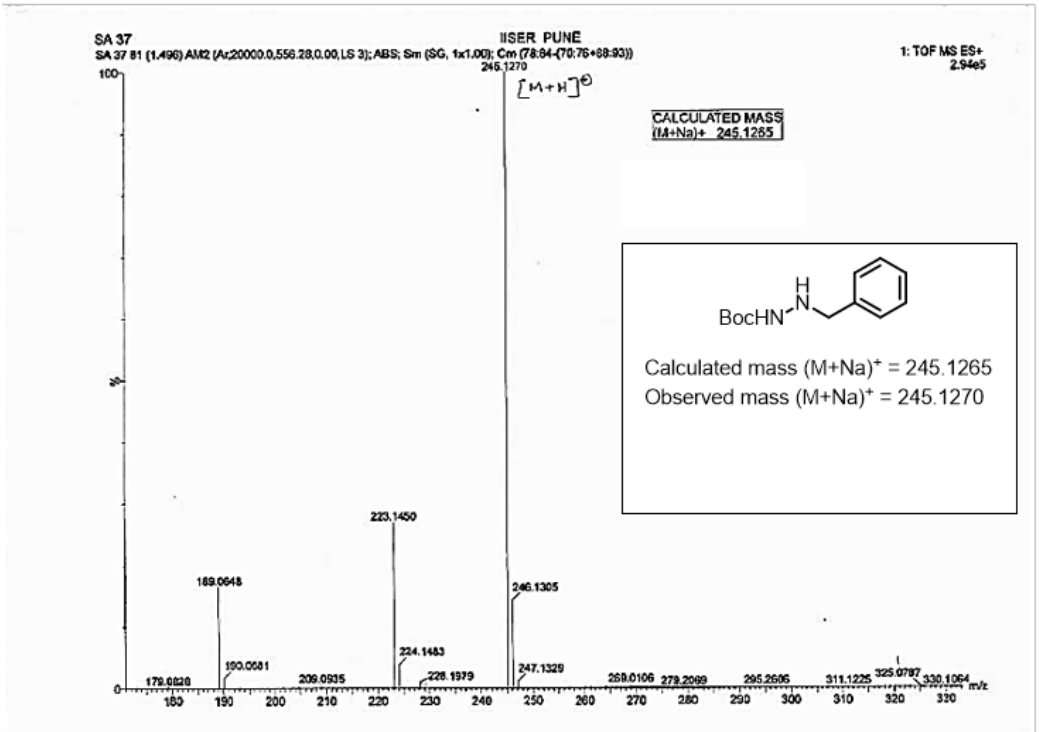
Mass spectra of compound 7



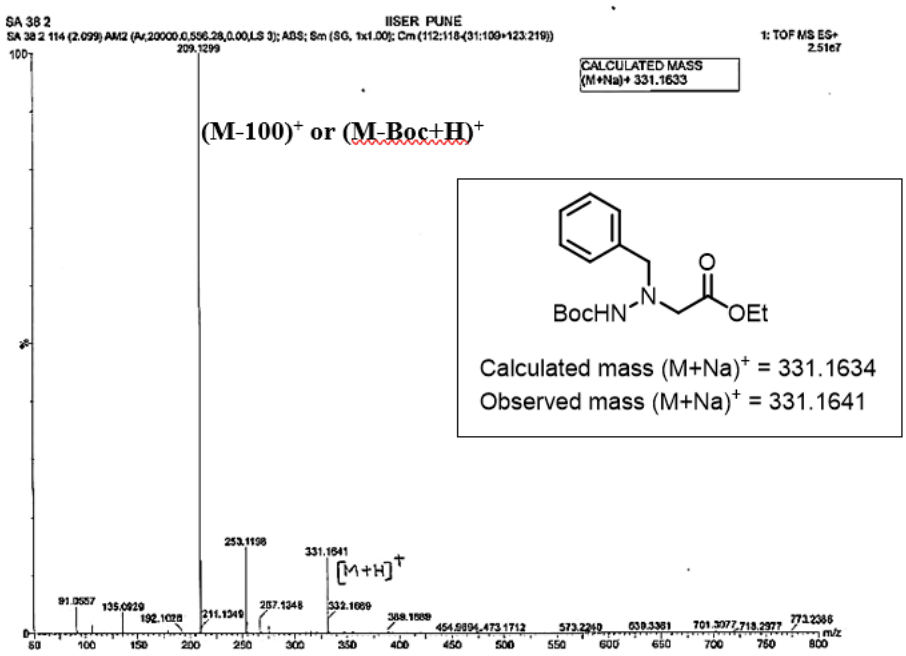
Mass spectra of compound 1



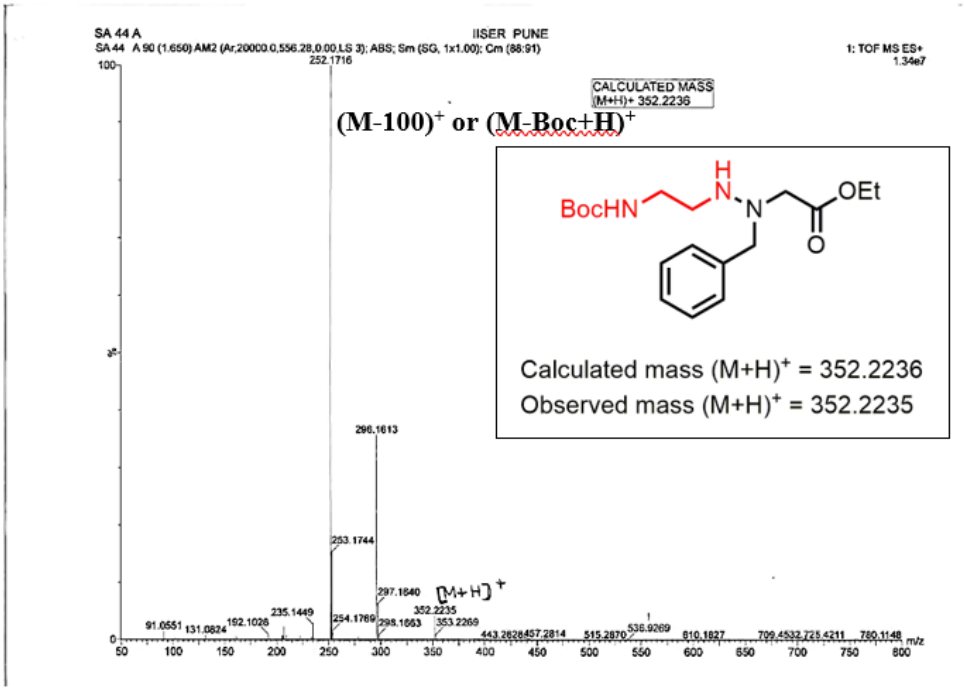
Mass spectra of compound 11



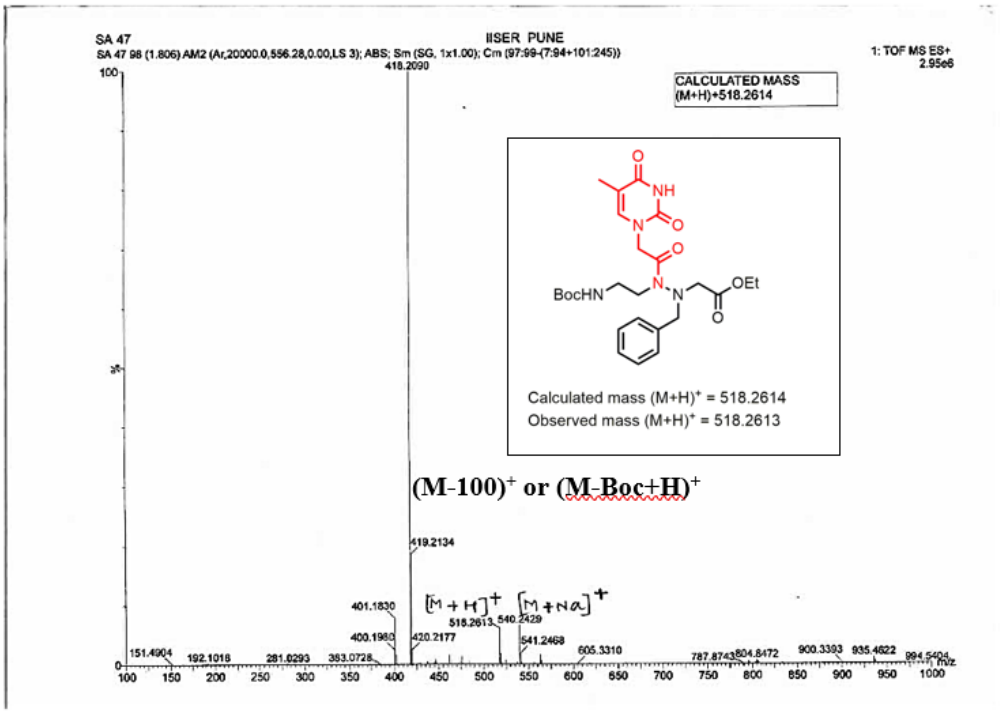
Mass spectra of compound 12a



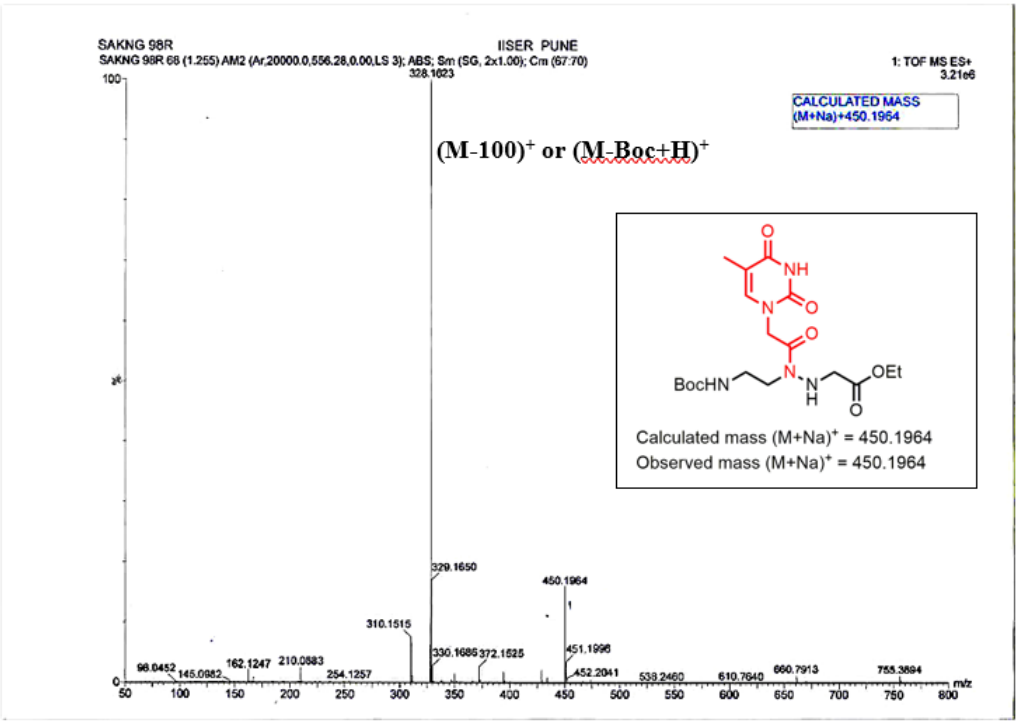
Mass spectra of compound 13a



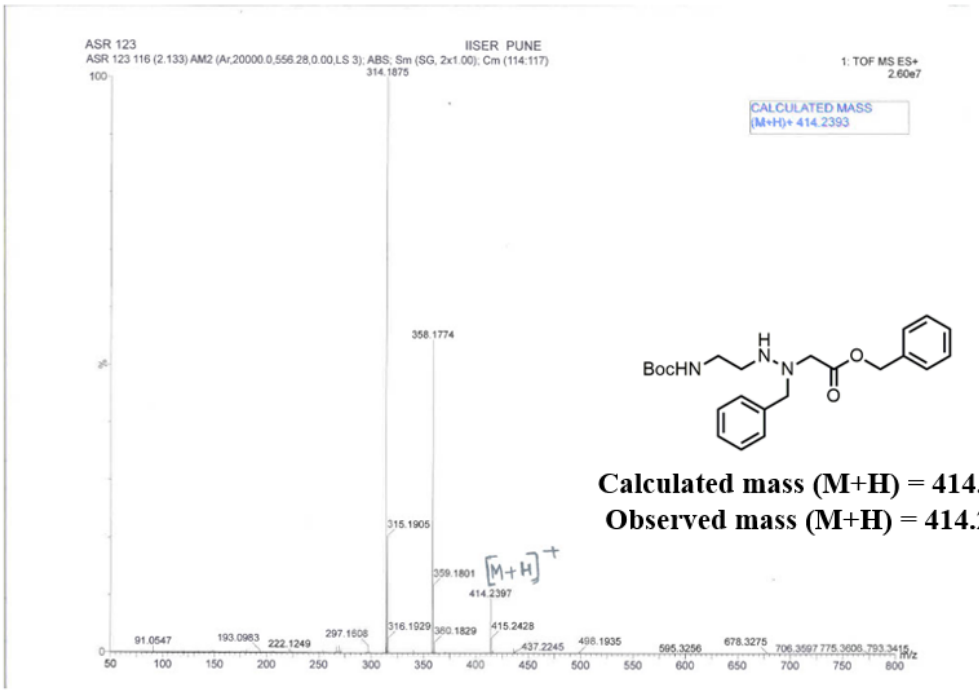
Mass spectra of compound 14a



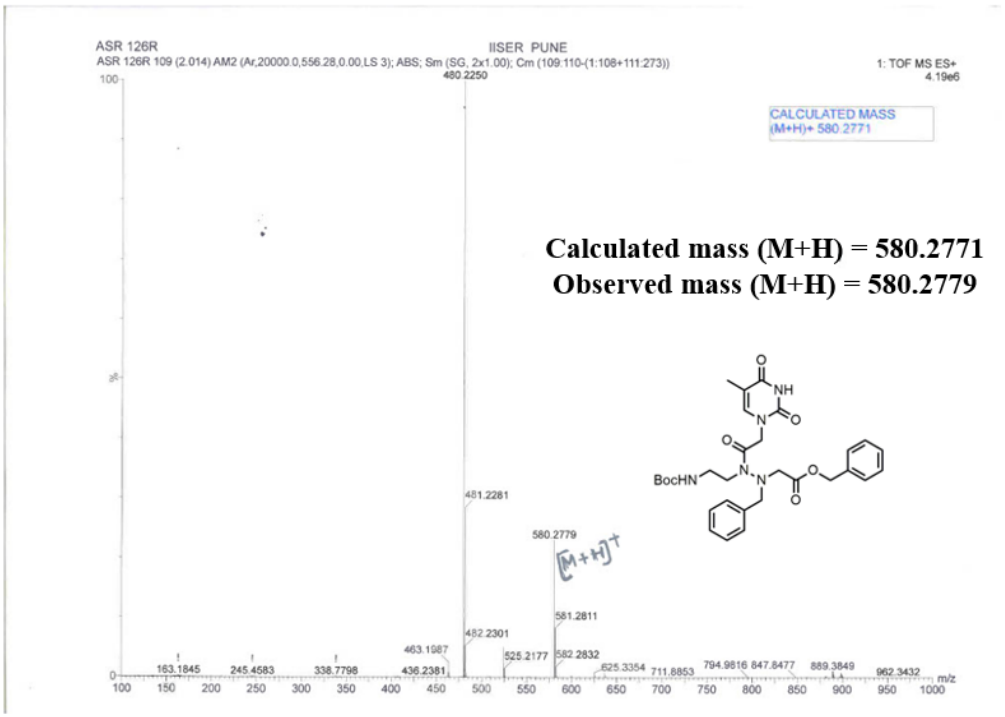
Mass spectra of compound 2



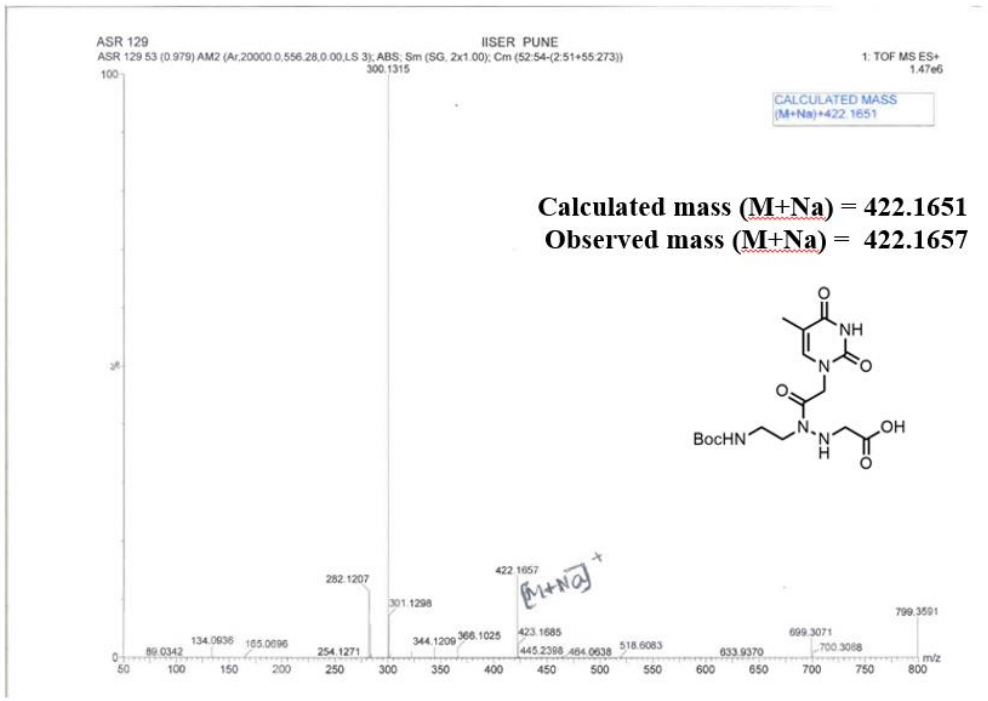
High resolution Mass spectra of 13b



High resolution Mass spectra of 14b

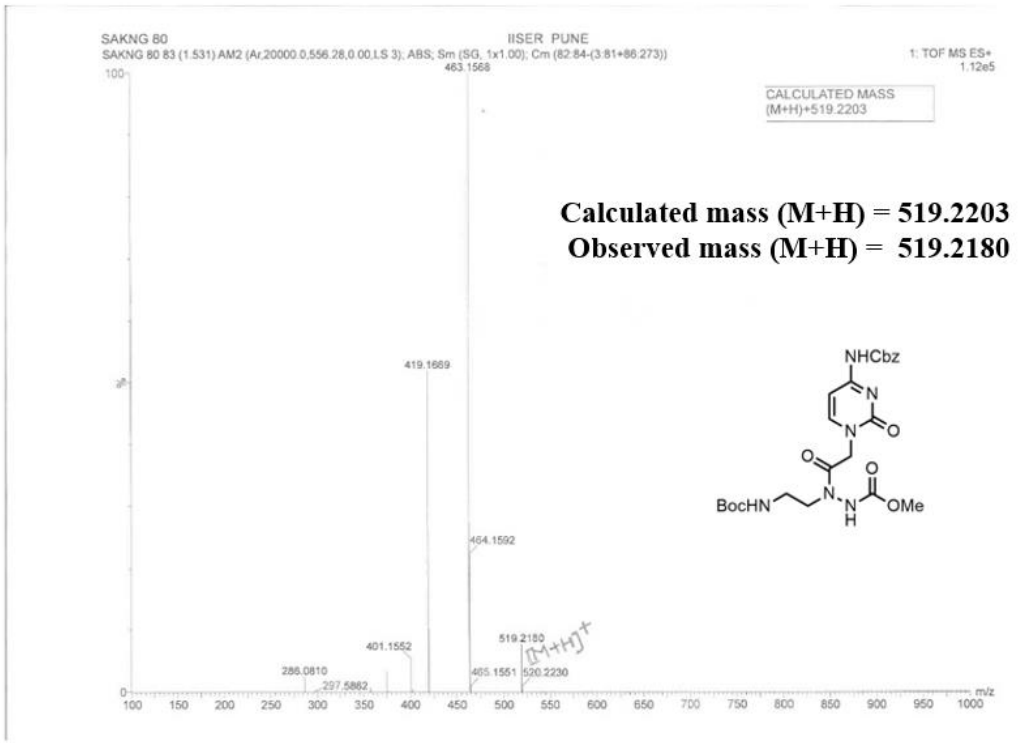


High resolution Mass spectra of 15



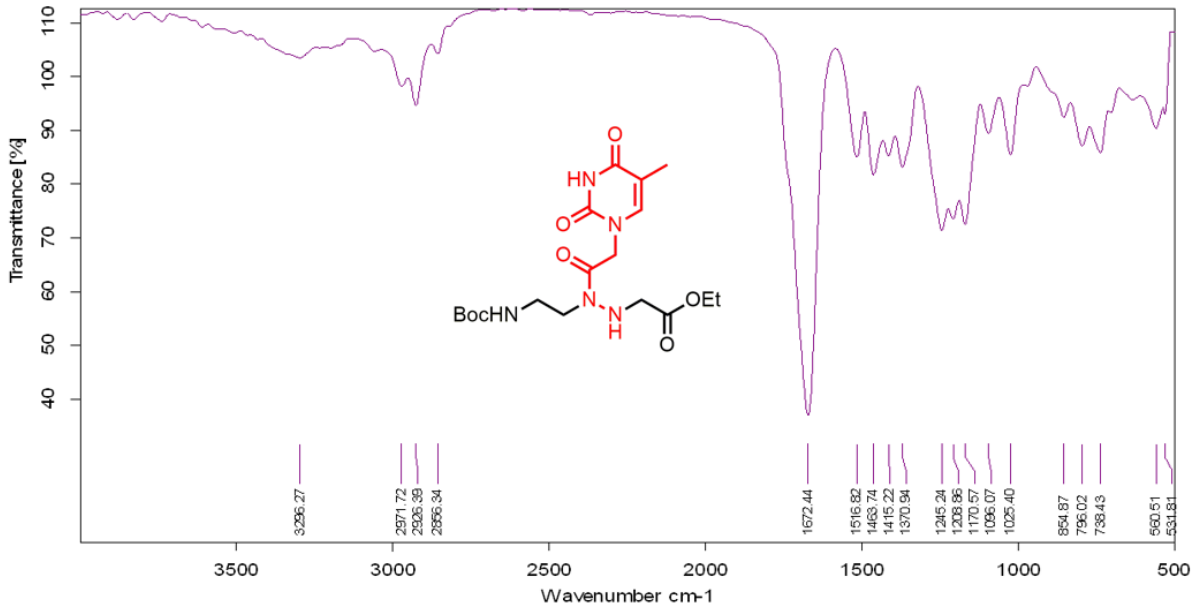
Calculated mass (M+Na) = 422.1651
Observed mass (M+Na) = 422.1657

High resolution Mass spectra of 1C

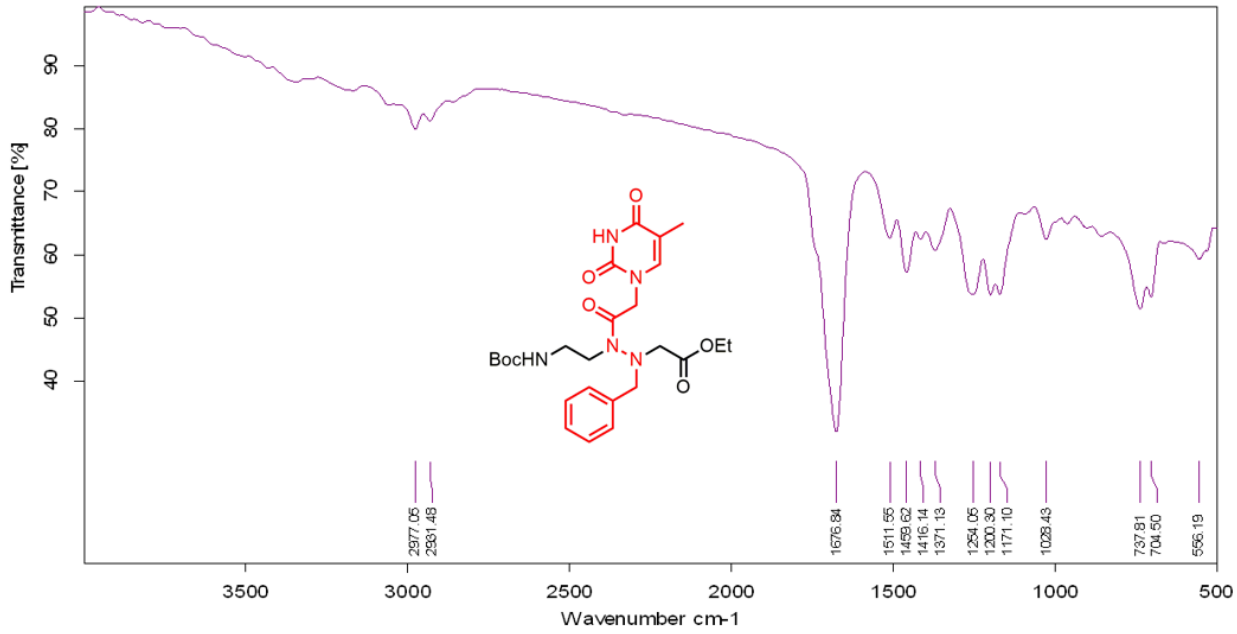


Calculated mass (M+H) = 519.2203
Observed mass (M+H) = 519.2180

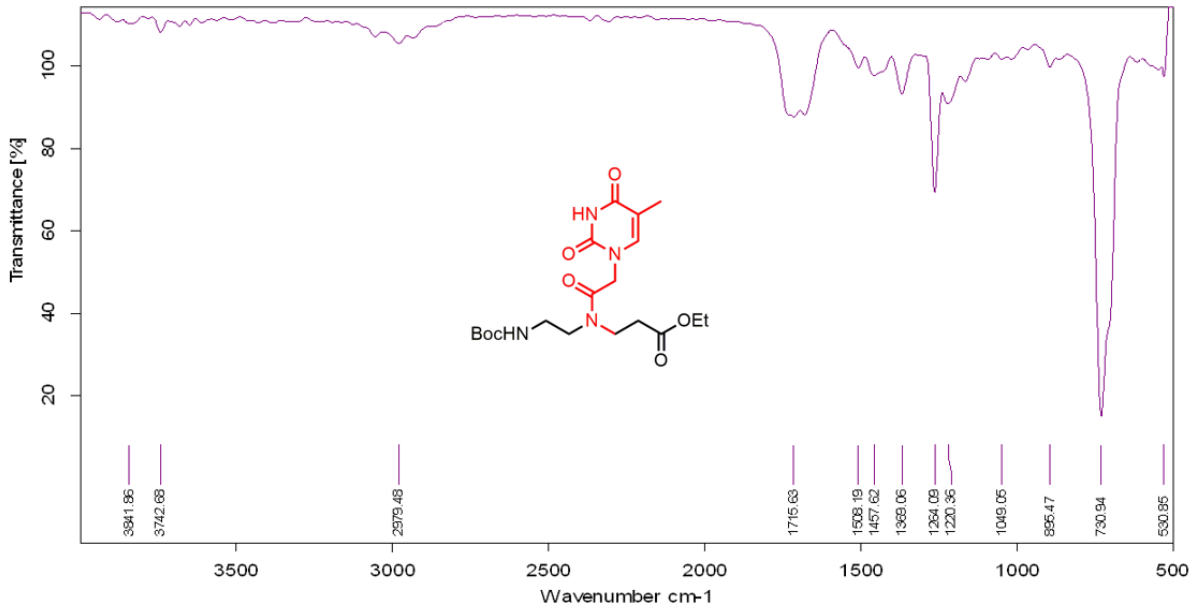
IR spectra of compound 2 in chloroform



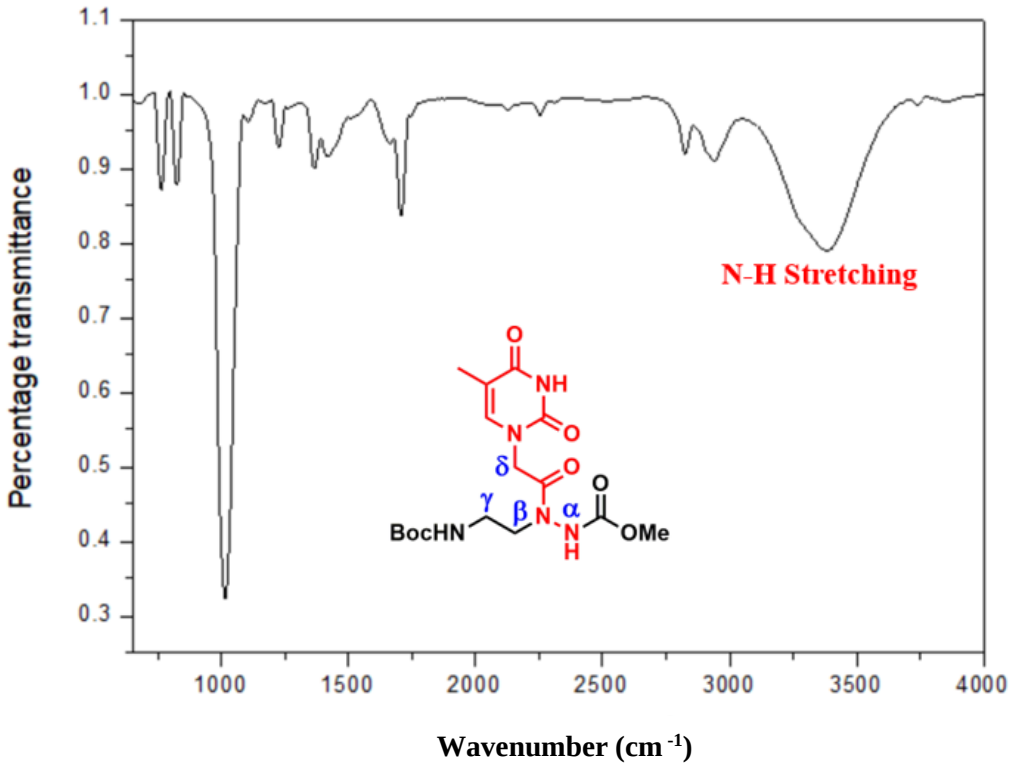
IR spectra of compound 14a in chloroform

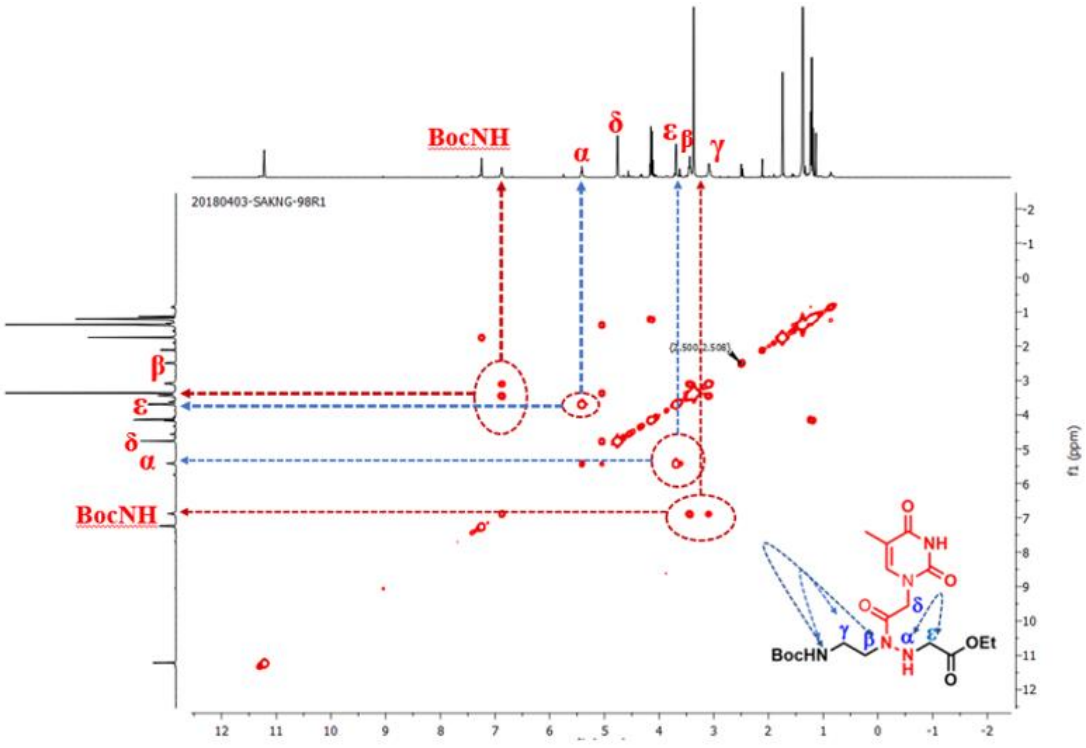


IR spectra of compound 4 in chloroform

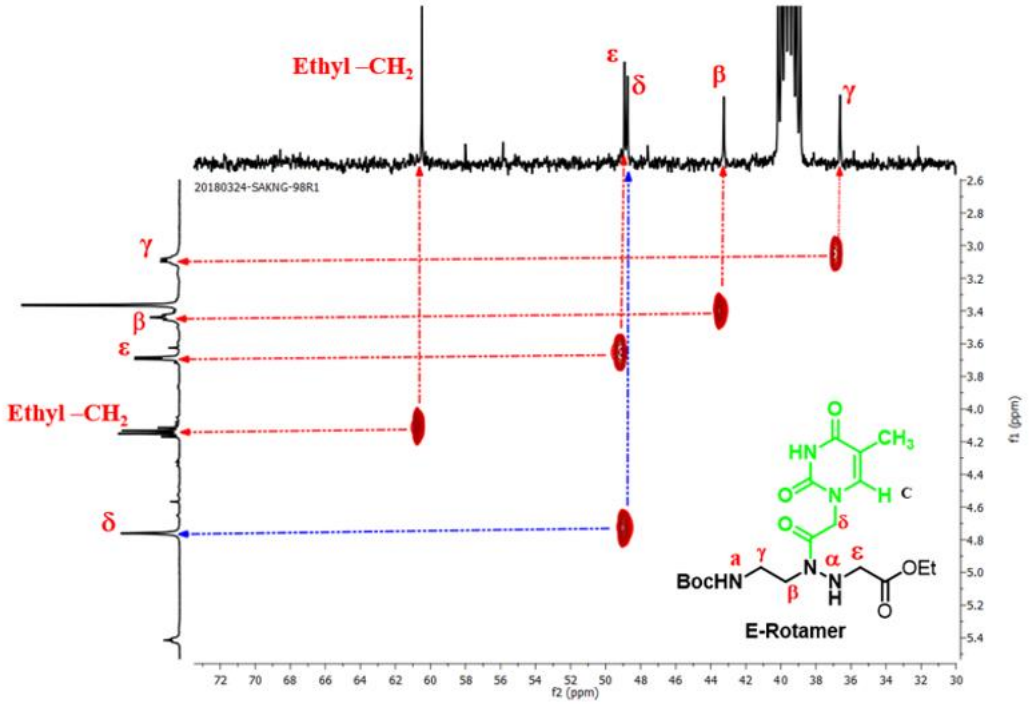


IR spectra of compound 1 in chloroform

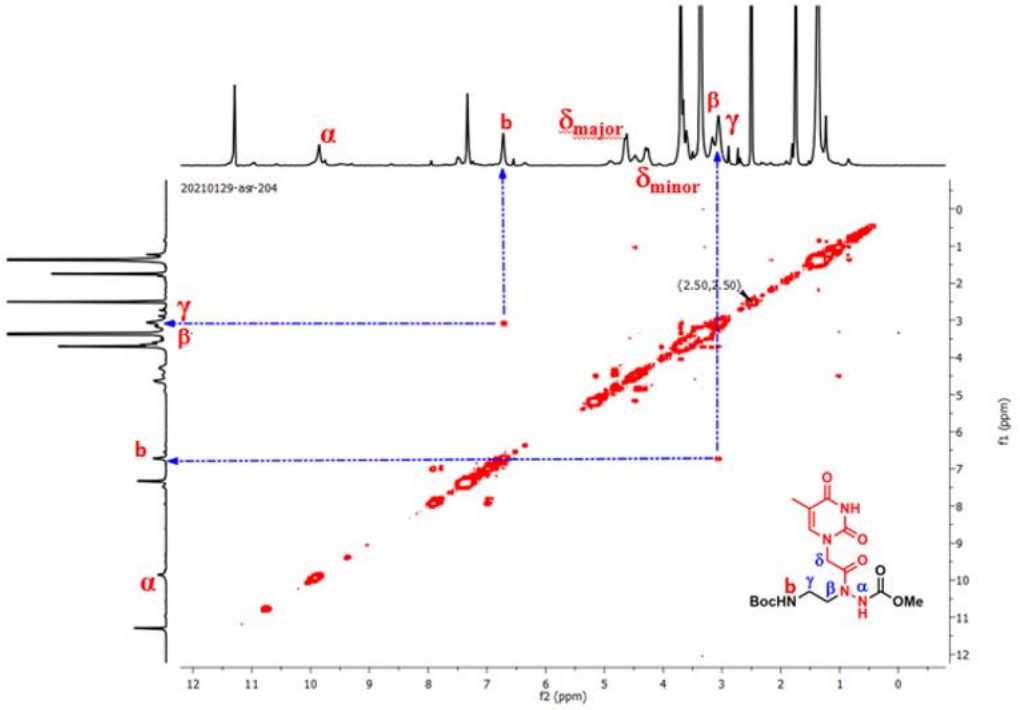




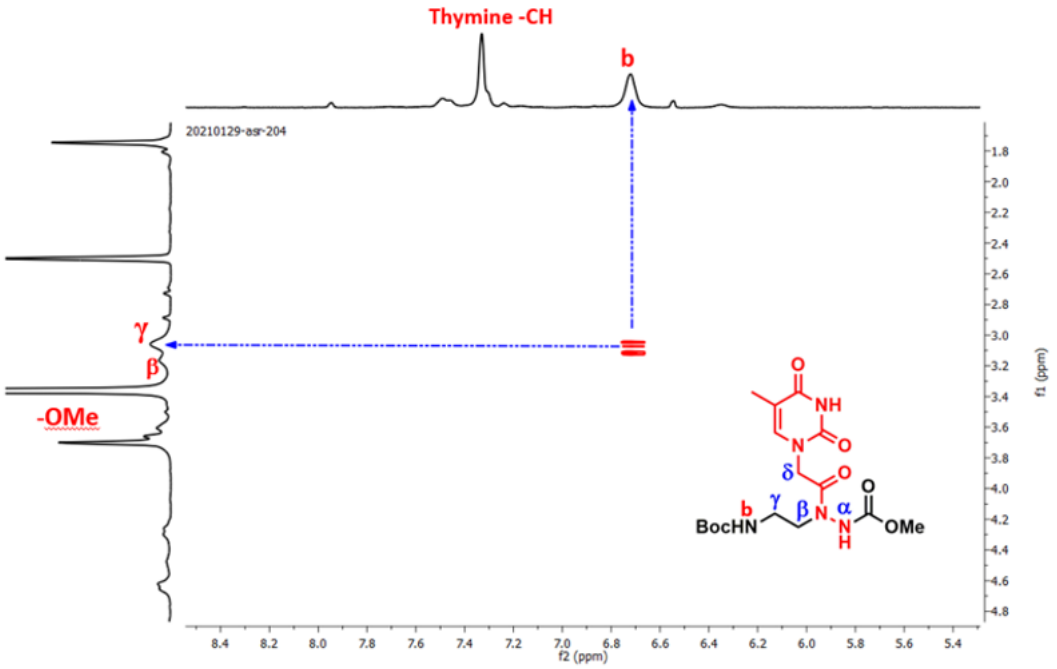
^{1}H-^{1}H TOCSY spectra of β-ala-*aza*-PNA-T **(2)** in DMSO-d_6 (500 MHz, *Full spectra***)**



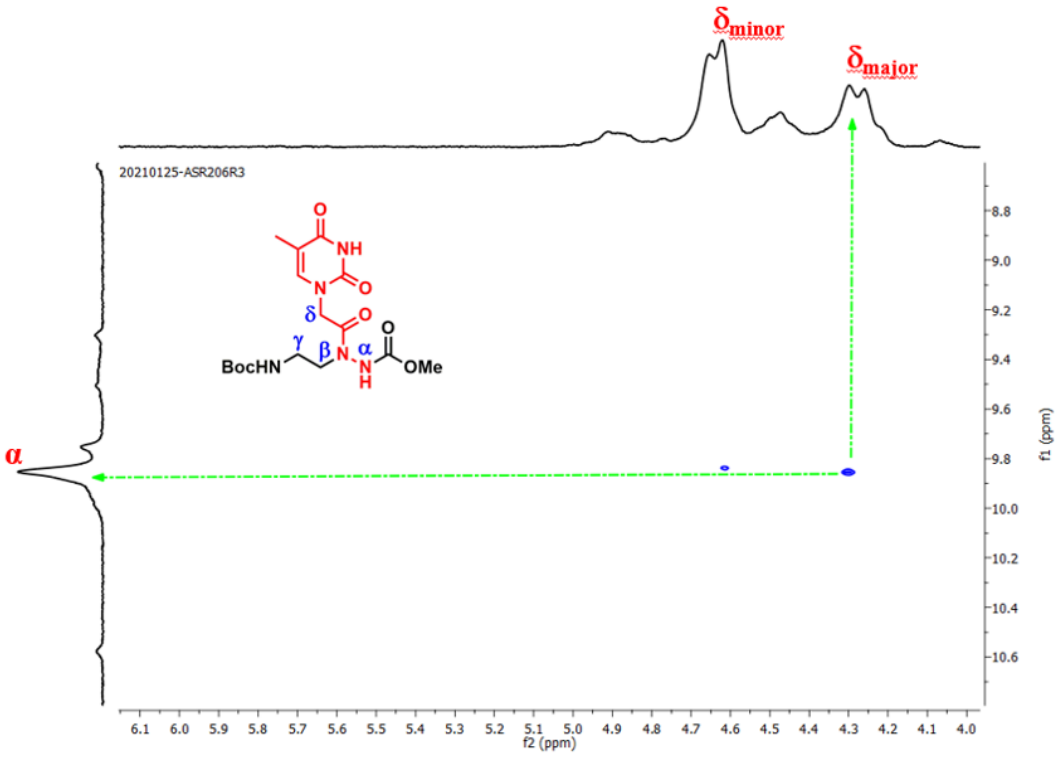
^{1}H-^{13}C HSQC spectra of β-ala-*aza*-PNA-T **(2)** in DMSO-d_6 (500 MHz, *Enlarged spectra***)**



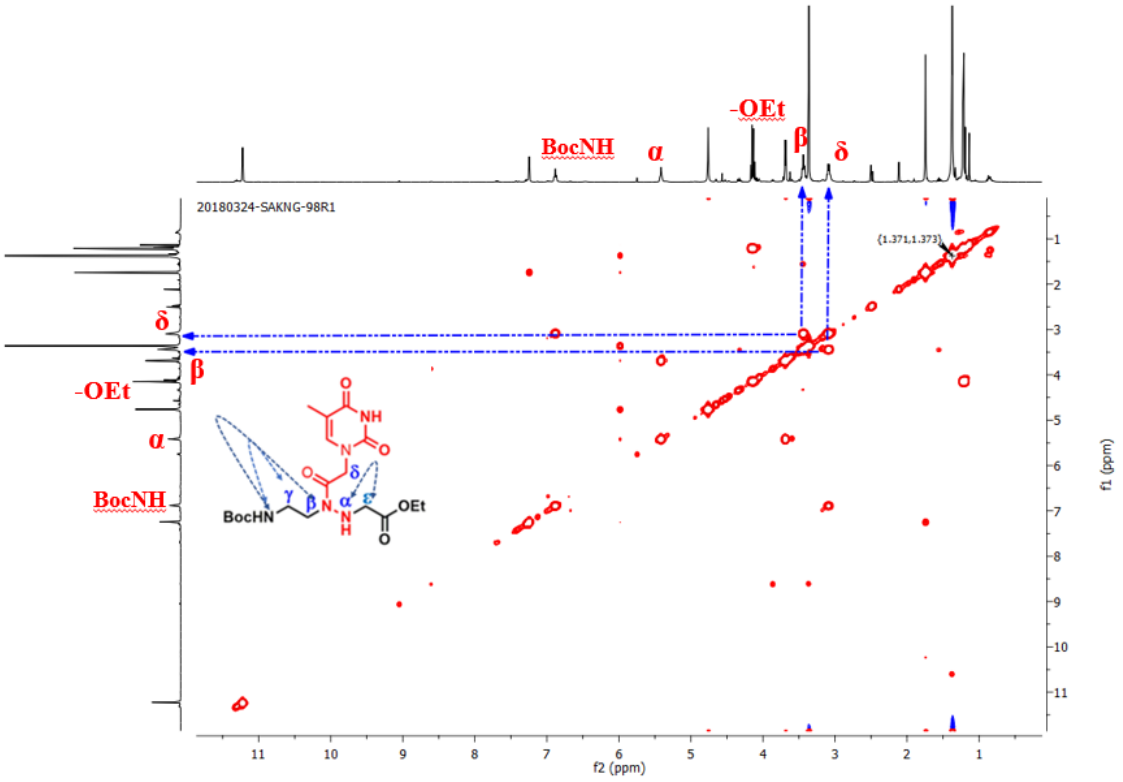
^{1}H-^{1}H TOCSY spectra of α-*aza*-PNA-T **(1)** in DMSO-d$_6$ (500 MHz, *Full spectra)*



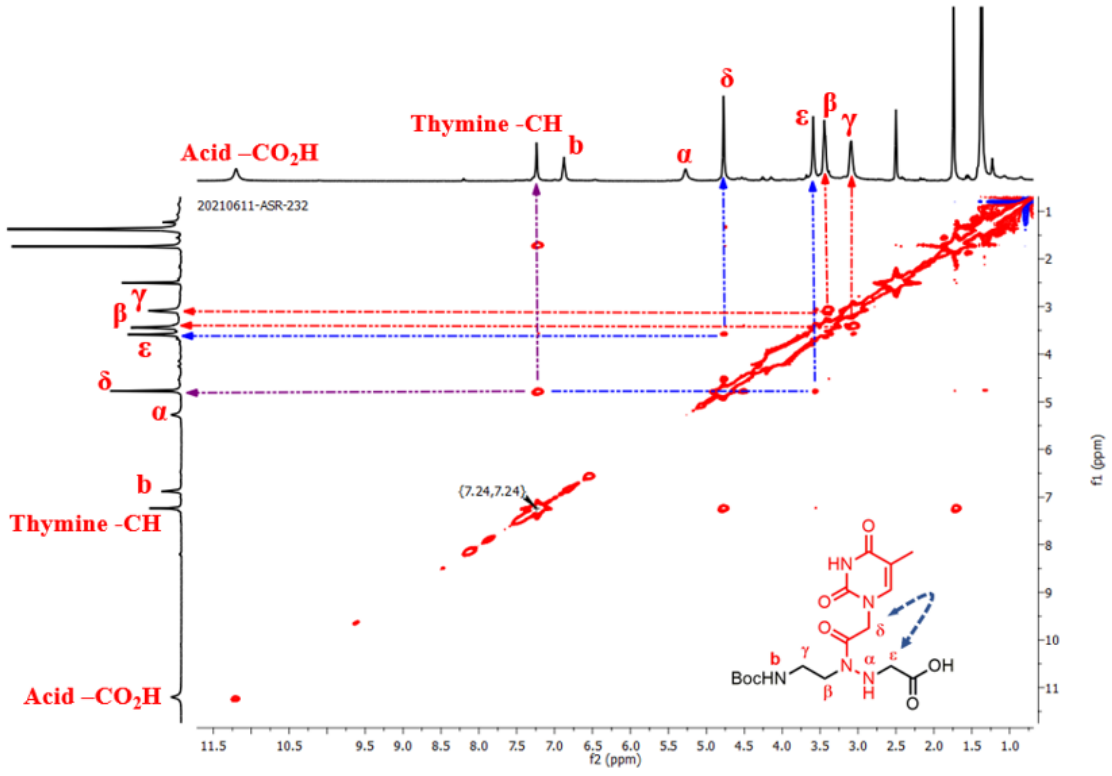
^{1}H-^{1}H TOCSY spectra of α-*aza*-PNA-T (**1**) in DMSO-d$_6$ (500 MHz, *Enlarged spectra*)



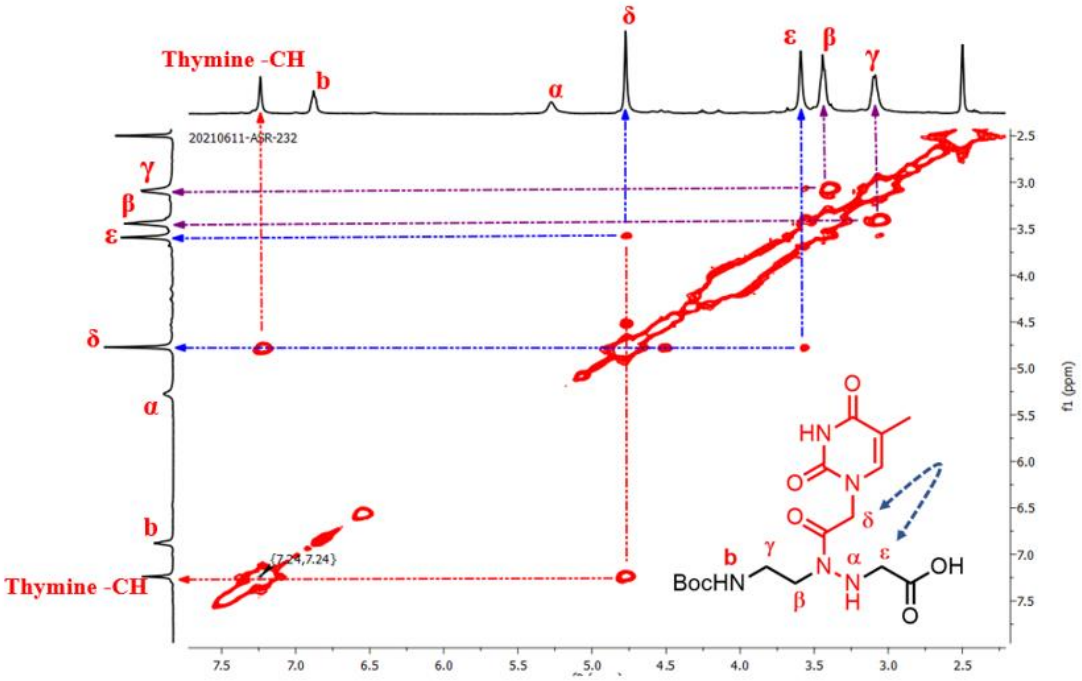
^{1}H-^{1}H NOSEY spectra of α-*aza*-PNA-T **(1)** in DMSO-d_6 (500 MHz, *Enlarged spectra*)



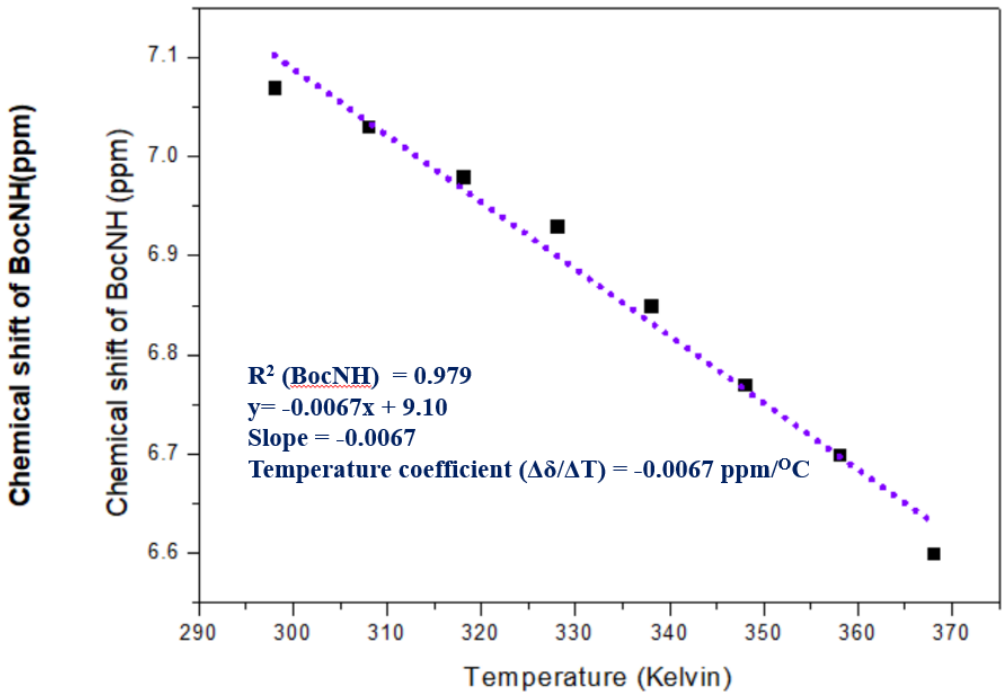
^{1}H-^{1}H COSY spectra of β-ala-*aza*-PNA-T **(2)** in DMSO-d$_6$ (500 MHz, *Full spectra*)



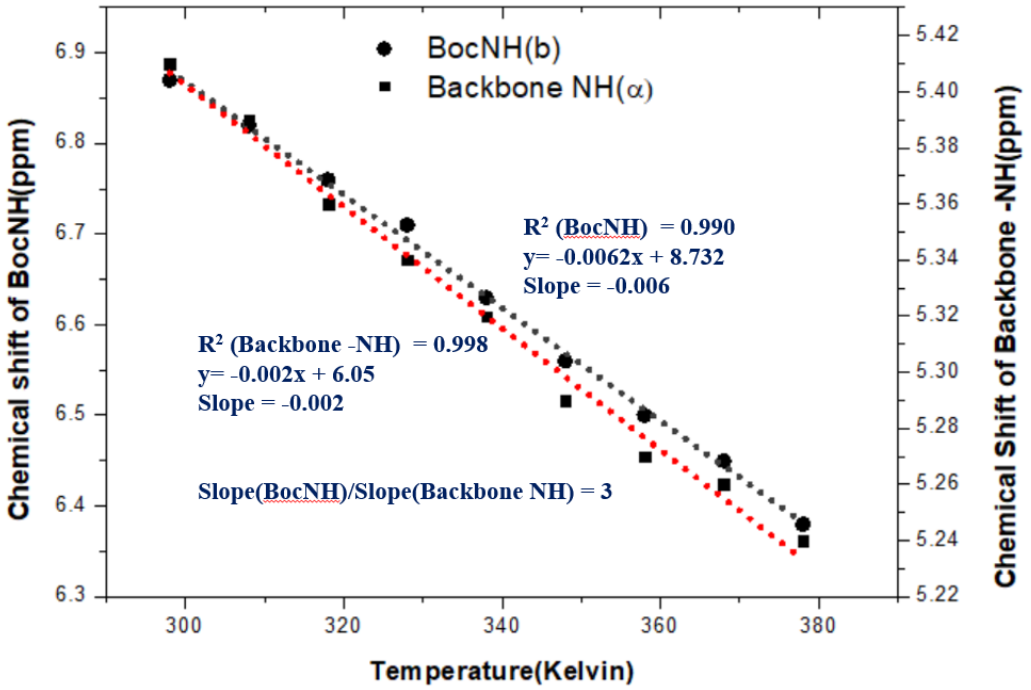
^{1}H-^{1}H NOSEY spectra of β-ala-*aza*-PNA-T **(2)** in DMSO-d$_6$ (500 MHz, *Full spectra*)



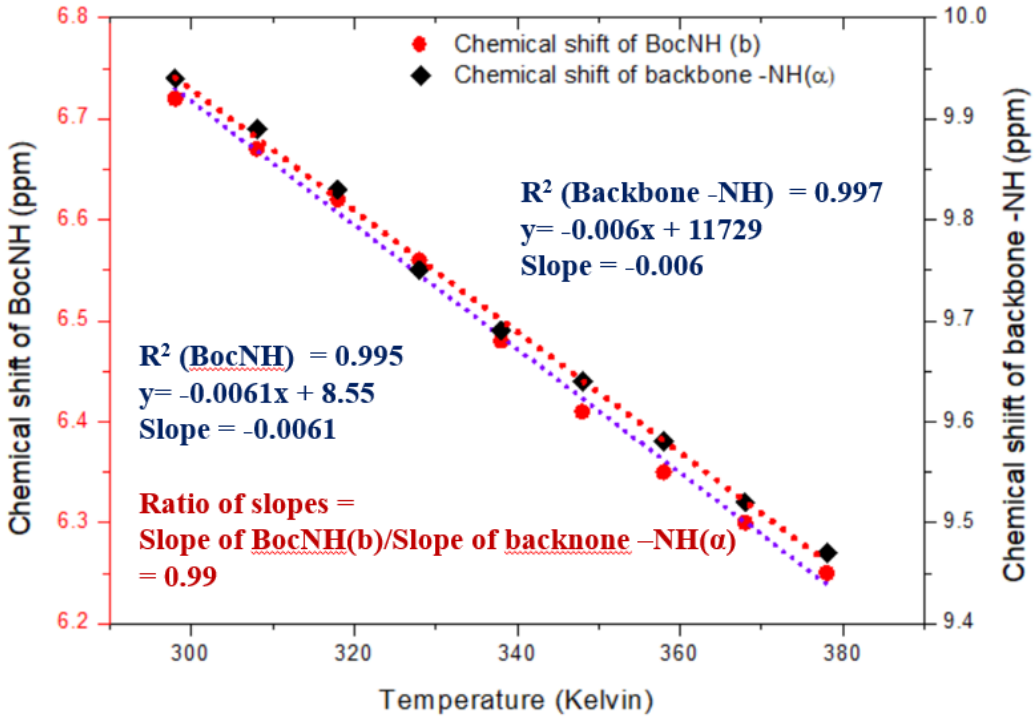
^{1}H-^{1}H NOSEY spectra of β-ala-*aza*-PNA-T (**2**) in DMSO-d_6 (500 MHz, *enlarged spectra*)



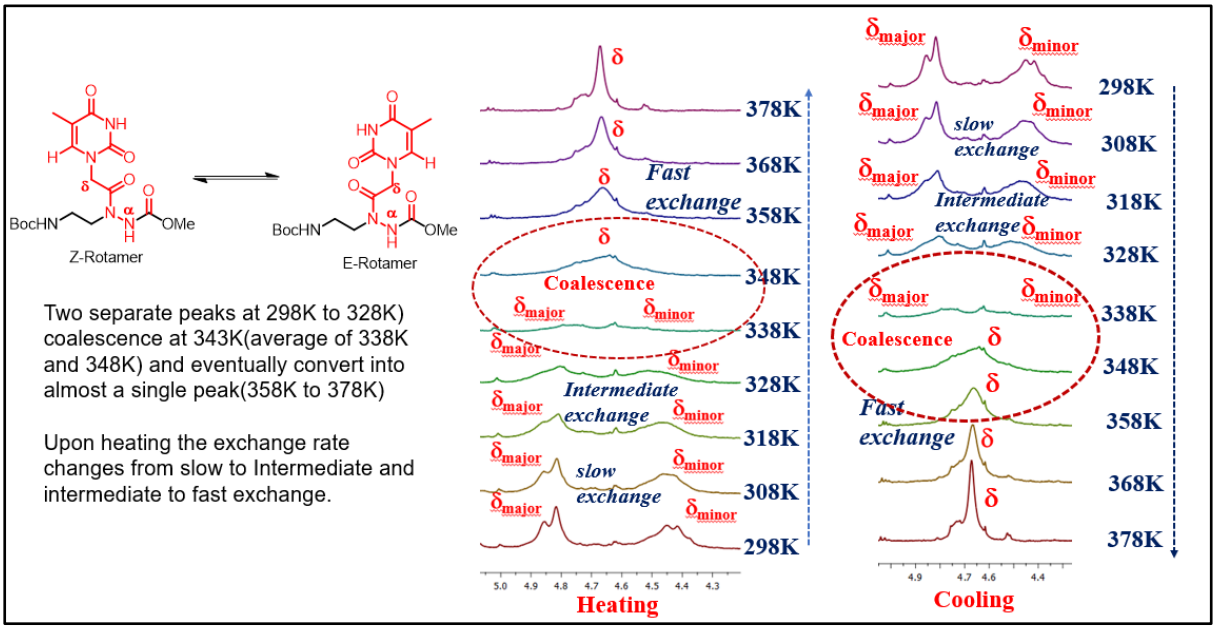
Graphical representation of Variation of Chemical shift of Backbone -NH (α) and BocNH of β-ala-PNA-T **4** (b) Slope of the line (Δδ/ΔT) indicates *temperature coefficient*.



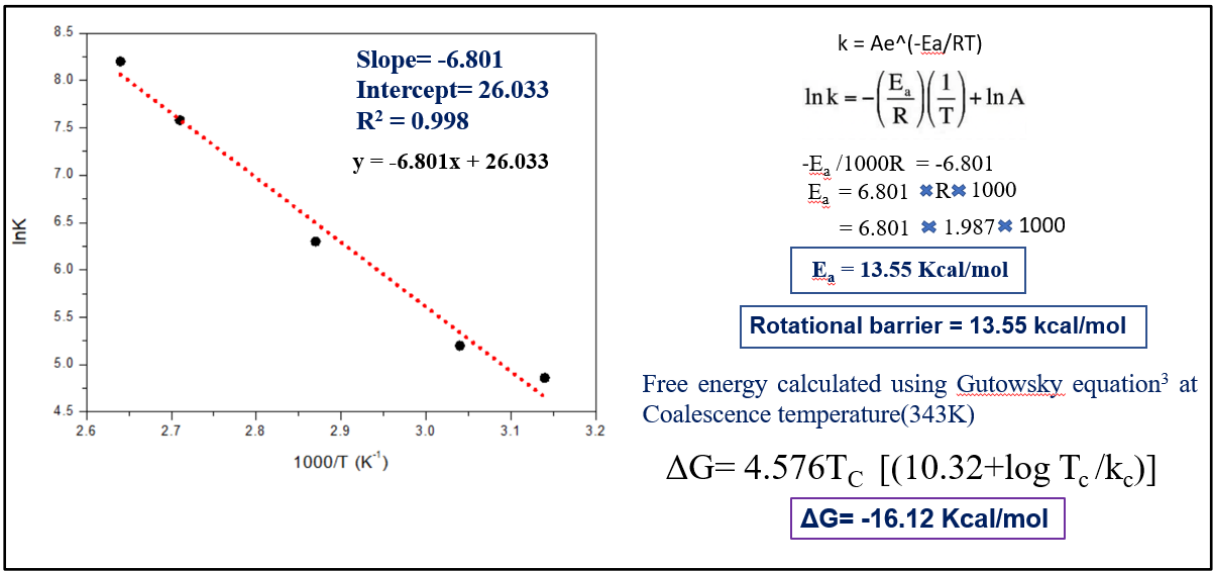
Graphical representation of Variable temperature NMR (DMSO-Heating) of β-ala-*aza*-PNA-T **2** and variation of chemical shift of backbone -NH (α) and BocNH (b)**.** Slope of the line (Δδ/ΔT) indicates *temperature coefficient*.



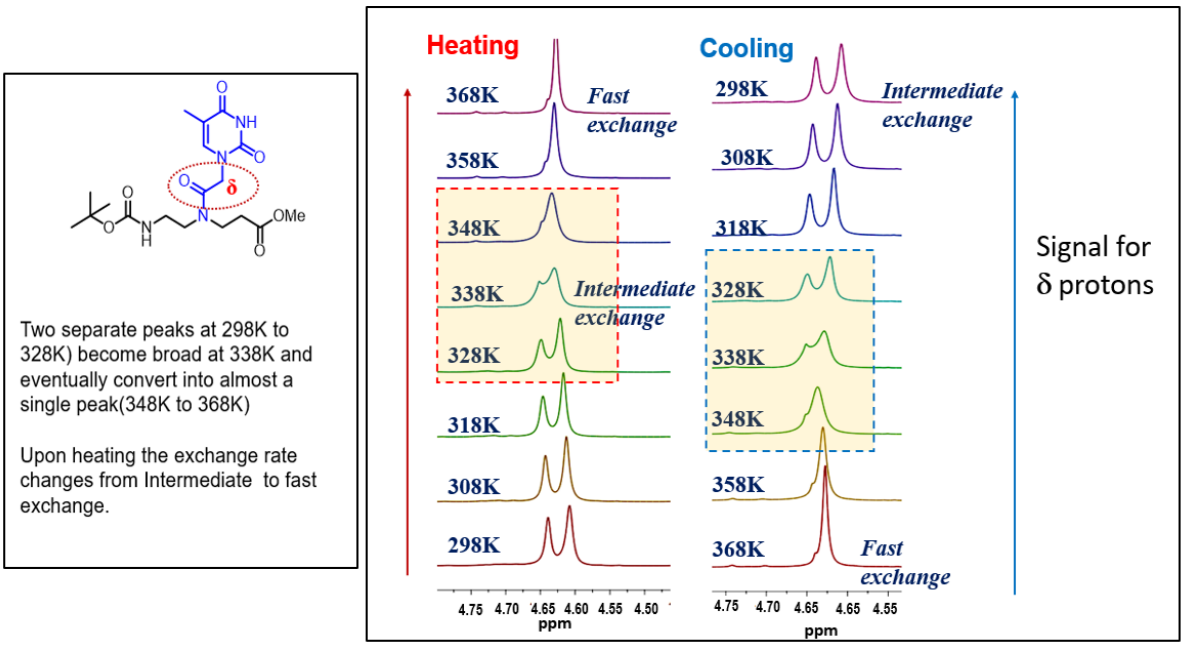
Graphical representation of Change of chemical shift of BocNH (b) and backbone -NH (α) during the Temperature dependent (Heating in DMSO-d_6) NMR of α-*aza*-PNA-T monomer **1.** Slope of the line (Δδ/ΔT) indicates *temperature coefficient.*



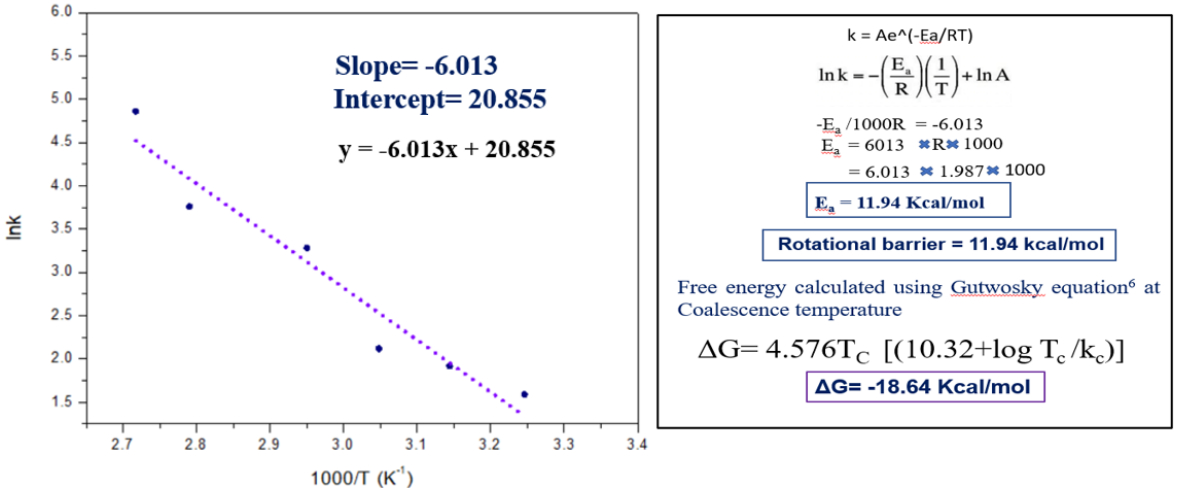
Change of the peak pattern of δ-CH_2 during the temperature dependent NMR (Heating in DMSO-d_6) of α-*aza*-PNA-T **1**. Initial two separate rotamers (298K to 338K) coalescence at 343K (average of 338K and 348K)



Calculation of Rotational energy barrier (E_a) using Arrhenius equation and Free energy of activation using Gutwosky equation from the temperature dependent ^{1}H NMR data[34] (of α-*aza*-PNA-T). The Energy of activation was calculated is 13.55 Kcal/mol and free energy of activation was -16.12 kcal/mol.



Change of the peak pattern of δ-CH_2 during the temperature dependent NMR (Heating in DMSO-d_6) of β-ala-PNA-T **4**. Initial two separate rotamers (298K to 338K) coalescence at 343K (average of 338K and 348K), and a single peak predominates at higher temperature (358K to 368K).



Calculation of Rotational energy barrier (E_a) using Arrhenius equation and Free energy of activation using Gutwosky equation from the temperature dependent ^{1}H NMR data.[34] The Energy of activation was calculated is 11.94 Kcal/mol and free energy of activation was -18.64 kcal/mol.

Chapter 3

α/β-*Gemdimethyl(gdm)* Guanine PNA monomers and dimers: Rotamer dependent fluorescence

Chapter-3

α/β-*Gemdimethyl(gdm)* Guanine PNA monomers and dimers: Rotamer dependent fluorescence

3.1. Introduction

3.1.1. Introduction to luminescence[1]

It is essential to provide light or another form of energy to generate emission from a material.[1] The process of generation of emission from a material is broadly classified into two categories i) incandescence and ii luminescence). The process of incandescence[1] involves heating a metal/material to a high temperature, which results emission of light from the heated metal. Examples of incandescence includes heating an iron rod to a high enough temperature which results in emission of red light. Glowing of sun and stars in our universe follow the incandescence mechanism.[1] Luminescence involves all forms of cool light i.e., emission of light from a cold substance.[1] Plausible reason of luminescence is the movement of electrons of an optical light emitting material from higher energy states to lower energy states. The process of luminescence can be broadly classified into five categories as (i) chemiluminescence (ii) electroluminescence (iii) triboluminescence (iv) bioluminescence and (v) luminescence due to absorption of external energy sources. Chemiluminescence is the emission of light due to a chemical reaction like oxidation reaction under very low temperature. Electroluminescence appears when an electric charge releases light during the separation of adhesive surfaces. Triboluminescence is the emission of light due to rubbing or crushing of crystals. Bioluminescence is the generation and emission of light energy in a living organism ex., blue light emission by firefly and glowworm. Luminescence resulting from absorption of external energy sources can be divided into two categories: (i) fluorescence, in which the material absorbs light in the ultraviolet region but emits in the visible region of the electromagnetic spectrum.[1] It is an instantaneous process with a lifetime of $<10^{-8}$ seconds, and the fluorescence emission persists as long as the source of UV-radiation kept on and (ii) phosphorescence, which is the emission of light from materials due to the absorption of UV-light, visible light, or any form of energy. Lifetime of phosphorescence process is higher than 10^{-8} seconds and is a non-instantaneous process, and the phosphorescence appears after absorbing the energy and in absence of the energy source.[1]

3.1.2 Fluorescence and phosphorescence

The mechanism of fluorescence and phosphorescence is explained by popular energy diagram termed Jablonski diagram (Figure 3.1).[2] When materials absorb UV radiation, a photon excites the electrons from lower energy ground state (S_0) to the higher energy excited state (S_1, Figure 3.1). The electrons in the excited state S_1 participate in various competitive pathways for relaxation. The excited electrons may relax by dissipating the energy through heat known as non-radiative emission, through which some energy of the excited electrons is lost.[2] On the other hand, the excited electrons of the fluorophore may relax by returning to its original ground state S_0 (Figure 3.1). The energy released during this transition (S_1 to S_0) appears in the visible region of the electromagnetic spectrum as emission of light known as fluorescence, with wavelength in visible region and hence the color. The efficiency of the fluorescence emission is determined by the quantum yield which is the ratio of the number of photons emitted to the number of photons absorbed.[1,2] The maximum theoretical quantum yield is 1 (100%), but however, some energy is always lost due to non-radiative transition and intersystem crossing

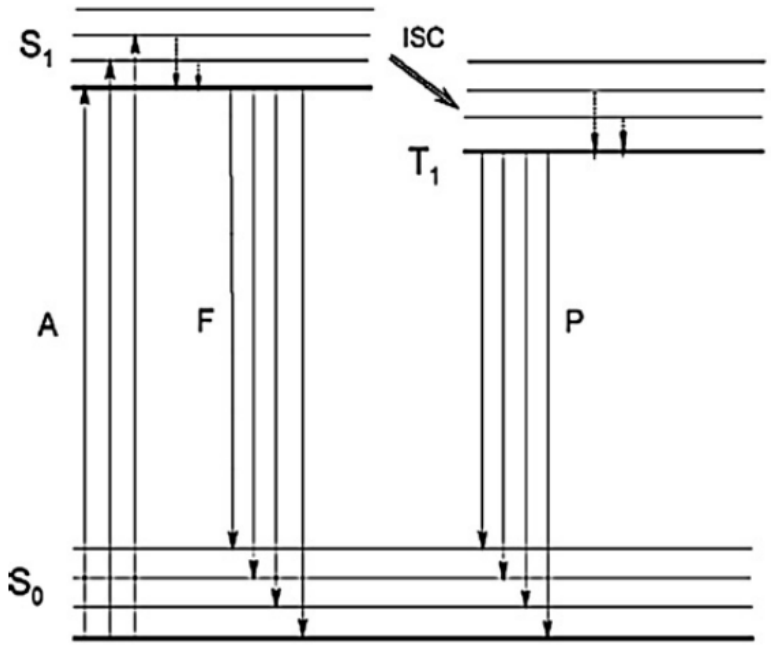


Figure 3.1. Jablonski diagram explaining fluorescence and phosphorescence (A= absorption, F = fluorescence and P= phosphorescence)

(phosphorescence) with the quantum yield always being less than 1. The average time which the excited electrons stay in the excited state S_1 before emitting the photons is known as fluorescence lifetime. The kinetics of excitation and de-excitation in fluorescence follows first

order kinetics $[S_1] = [S_1]_0 e^{-\Gamma t}$ where $[S_1]$ = concentration of excited state electrons at time t, $[S_1]_0$ = concentration of ground state electrons and Γ = decay rate (1/fluorescence lifetime).[1] The total decay rate of fluorescence is the summation of radiative and non-radiative process and expressed as $\Gamma_{total} = \Gamma_{rad} + \Gamma_{nrad}$ where Γ_{rad} = decay rate due to radiative process and Γ_{nrad} = decay rate due to non-radiative process.[1]

Apart from the primary fluorescence process, the excited state electrons at S_1 may relax through phosphorescence process.[2] The excited electrons convert into stable triplet state by intersystem crossing (ISC, Figure 3.1) to the triplet state T_1 and return to original S_0 state (Figure 3 .1). The energy released during the transition (T_1 to S_0) appears in the visible region of the electromagnetic spectrum with the observed emission of light in visible region known as phosphorescence.[1,2] According to the selection rules, a singlet-to-singlet transition is allowed (S_1 to S_0) but triplet to singlet transition (T_1 to S_0) is forbidden by quantum mechanics.[1,2] Hence the forbidden T_1 to S_0 de-excitation transition process is slower ($>10^{-8}$ s) and so phosphorescence appears slowly after the excitation process and may continue even after the source of the excitation is removed.[1,2]

3.1.3 Fluorescent nucleobases

Natural nucleobases (A/T/G/C) and the sugar phosphate backbone of the DNA are non-fluorescent. The synthetically modified fluorescent nucleobases incorporated in the oligomers along with the standard natural nucleobases are useful to study the dynamic structure and function of DNA and RNA. The synthetic fluorescent nucleobases can be either canonical nucleobases or non-canonical.[3] The canonical nucleobases retain the ability of the Watson-Crick complementary base pairing and show fluorescence properties, which is useful for the study of hybridization of derived oligomers. In contrary, the non-canonical nucleobases may or may not retain the base pairing ability of natural nucleobases but have potential to show high emissive properties in the derived oligomers and monomers.[3] 2-Aminopurine (2-AP) is the first canonical synthetic fluorescent nucleobase by Stryer et al[4] and the first synthetic non-canonical highly emissive fluorescent nucleobase ethenoadenine was reported by Leonard *et al*.[5] Later in 1990 the first synthetic fluorescent canonical thymine nucleobase was reported.[6] In 1996, Kool *et al.*[7] reported the synthesis and application of pyrene and other fluorophore attached nucleotides. After 2000, the work in the field of synthetic emissive nucleobases progressed very rapidly including development of discriminating emissive nucleobase by Saito and Okamoto.[8] The design of these nucleobases involves attachment of pyrene with thymine nucleobase, extending the conjugation in thymine by fusing the pyrimidine ring with π-

conjugated molecule and extending the conjugation in purines by fusing the ring with π-conjugated molecules.[8]

3.1.4 Canonical fluorescent purine nucleobases

For adenine, the C2 / C8 carbon and for guanine the N7 /C8 are the possible sites for modification to afford the emissive canonical nucleobases,[3] examples including pyrene coupled modified adenine and guanine (Figure 3.2a).[3] Another modification involves fusing the six membered ring of the purines with pyrazine ring (Figure 3.2b).[3] These modified purines exhibit significant bathochromic shift (100 nm) compared to the natural nucleobases and with higher quantum yields.[9] The modification or rearrangement of the substituents present on purine ring involves the most useful reporter nucleobase 2-aminopurine (2-AP),[4] where the 6-amine group is shifted to C2. Other examples include 2,6-diamino purine (DAP)[10] and 8-aza-guanine (Figure 3.2d).[11] 2-AP and 2,6-diamino purine exhibit 50 nm and 30 nm redshift in fluorescence compared to the original adenine.

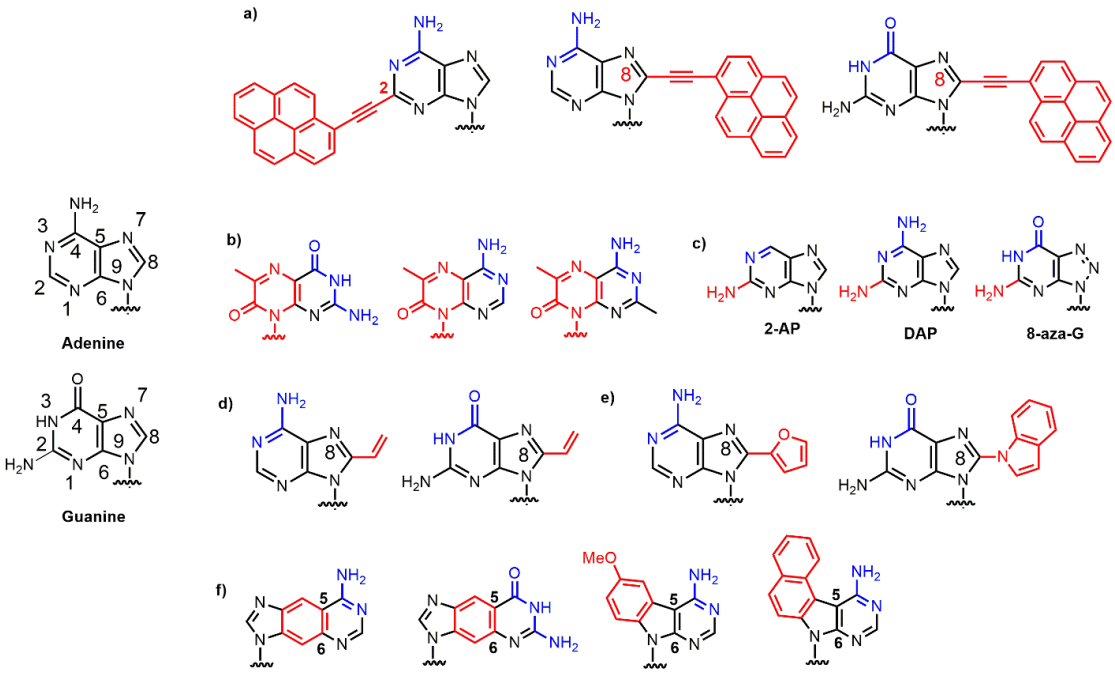


Figure 3.2. Modified canonical emissive purine nucleobases

Vinylation of adenine and guanine at C8 show 20 nm red shift in the emission spectra of the monomer with increment in the quantum yield (Figure 3.2d).[12] Substitution of single heteroatom heterocycles such as thiophene, furan, indole etc at C8 or C6 position of the guanine (Figure 3.2e), introduction of six membered ring in between the 5 and 6 membered ring of the purines, or addition of another ring at C7 and C8 of purines (Figure 3.2f) bring dramatic red

shifts of emission maximum and improve the quantum yield of the derived monomer or oligomer.[13, 14]

3.1.5 Canonical fluorescent pyrimidine nucleobases

Incorporation of thiophene or furan at the C5 position of thymine and cytosine improves the optical properties of the natural nucleobases (Figure 3.3a).[3] C5 furan substituted cytosine shows UV absorption maxima at 310 nm and emission maxima at 443 nm.[15] Deamination of cytosine followed by incorporation of methyl group at C5 affords the fluorescent pyrimidine derivative m^5K (Figure 3.3a) that exhibits emission maxima to 400 nm.[16] Introduction of six membered ring structure or naphthyl ring at the C4 and C5 position of the cytosine (Figure 3.3b) increases the conjugation with significant red shift in emission and increment in quantum yield.[17] Conjugation of fluorophores such as pyrene, fluorenone or boron-dipyrromethene (BODIPY) to pyrimidines nucleobases (Figure 3.3c) at C5,[18,19] fusion of another ring structure to C5-C6 or C4-C5 bonds of the pyrimidine ring (Figure 3.3d) resulted in significant enhancement of optical properties of natural nucleobases.[3]

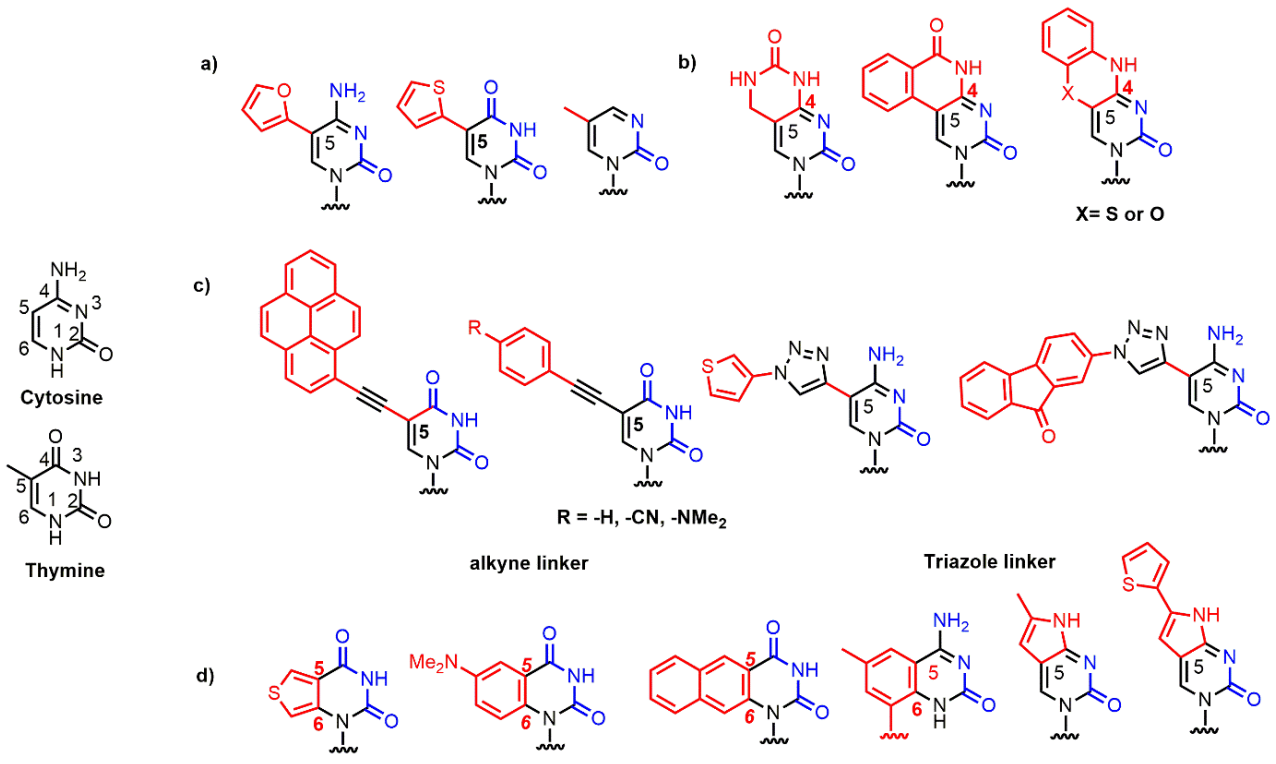


Figure 3.3. Modified canonical emissive pyrimidine nucleobases

Thiophene fused to uracil at C5-C6 bond exhibits absorption maxima at 304 nm and emission maxima at 412 nm with increased quantum yield of 0.48.[20] Naphthalene fused to

thymine at C5-C6 bond shows emission maxima at 434 nm and quantum yield of 0.82.[21] However, naphthalene fused purines (Figure 3.2f) show poor optical properties compared to the naphthalene fused pyrimidines. This difference is attributed due to the non-planar structure of naphthalene fused purines compared to the planar and rigid tricyclic structure of naphthalene fused pyrimidines.[21] Incorporation of modified fluorescent nucleobases results in interesting emissive properties of the derived monomer and oligomer.

3.1.6 Fluorescence of compounds that lack conventional chromophore

Traditional fluorophores (pyrene, aromatic organic dyes etc) contain a network of conjugated π-electrons which lead to fluorescence properties.[22] Many π-conjugated molecules also exhibit bright fluorescence in solid state but the fluorescence gets quenched in solution. Recent studies suggest that molecules containing non-conjugated C-N, C-O and C-C bonds show bright fluorescence either in the solid/solution state or in the aggregated state, although these molecules lack the presence of conventional fluorophores.[23] Such molecules exhibiting *clusterisation triggered emission (CTE)* are termed *clusteroluminogens.* There are three categories of clusteroluminogens: (i) polymers (natural and synthetic) (ii) small molecules (with or without aromatic rings) and (iii) metal clusters.[23] Basic structural and spectral features of *clusteroluminogens* include (i) non-conjugated structure with functional groups and π-electrons separated by sigma bonded spacer (ii) polymers that are non-fluorescent in low concentration but show bright fluorescence in aggregated state, (iii) materials that exhibit excitation-dependent emission spectra with increase in excitation wavelengths showing red shifted emission band and (iv) polymers and nano particles that exhibit size dependent emission spectra. *Clusterisation triggered emission (CTE)* could lead to appearance of room temperature phosphorescence.[23] In green fluorescence protein (GFP),[24] 4-(*p*-hydroxybenzylidene) imidazolidin-5-one chromophore is the origin of the optical properties of protein.[25] Polypeptide (ValGlyGlyLeuGly) showed green light emission, [26] whose intensity strongly depends on the degree of hydration.[26] Even biopolymers like starch, cellulose show faint blue fluorescence in solid or aggregated states.[27] Sodium alginate (SA) irradiated at 312 nm exhibits blue fluorescence but removal of UV light reveals yellow-green room temperature phosphorescence emission.[28] The PANAM dendrimer, whose basic structure includes presence of amido, alkyl and amine group and devoid the presence of any chromophore[29] shows fluorescence arising from n - π^* interaction from the nitrogen of amido group to carbonyl of the amido group.[29] Fluorescence seems to be the inherent property of dendrimers and hyper branched synthetic polymers[30] and some hetero atom (O,S) based polymers including poly(ethylene glycol)

(PEG), polydimethylsiloxane (PDMS).[31] PMV irradiated at 365 nm shows concentration dependent blue fluorescence emission.[32] Wang *et al.* reported the synthesis and optical properties of the copolymers.[33] Low molecular weight copolymer poly[(maleic anhydride)-alt-(vinyl pyrrolidone)] (PMVP) shows bright yellow fluorescence in powder compared to the high molecular weight copolymer PMVP-2 that shows sky blue colour emission in powder.[33] High purity and defined structure of the polymers is essential for the study of optical properties.[34] In this context, small molecule *clusteroluminogens* with aromatic rings or lacking aromatic rings (Figure 3.4) that show luminescence activity have assumed importance. 1,1,2,2-tetraphenyl ethane (s-TPE) with four isolated phenyl groups connected via sigma bonded ethyl spacer[35] shows aggregation induced emission and red shifted emission spectra similar to conventional conjugated fluorophores.[35] The mechanism of the origin of fluorescence in s-TPE and other examples such as compound **1** and tetrathienylethene (TTE) Figure 3.4) remain unclear.[36] Intramolecular S…S interaction predominates at aggregated states for TTE along with the through bond conjugation between peripheral thiophene and central double bond.[37] The fluorescence of compound **1** is explained by intramolecular through space conjugation (TSC). Figure 3.4 shows the chemical structure of some *clusteroluminogens* consisting of isolated phenyl and other chromophores.

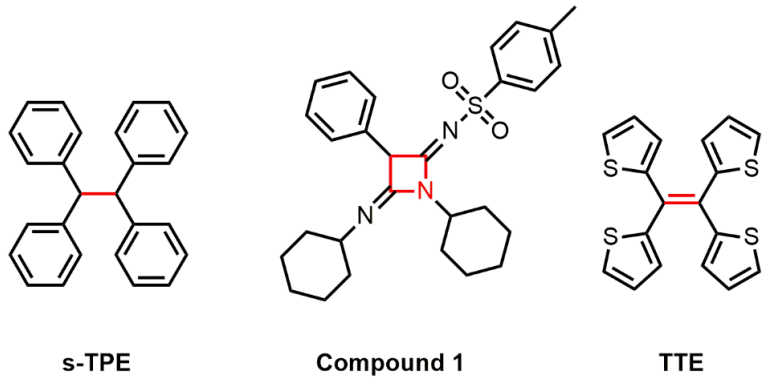


Figure 3.4. Molecular structure of s-TPE, Compound 1 and TTE

Poly (amino acids) are the first class of *clusteroluminogens* lacking aromatic rings[38] that show fluorescence in the aggregated state. Recrystalized L-isoleucine showed the yellow-green fluorescence emission.[39] Compound **2** with simple aliphatic oxime structure exhibits concentration dependent bright blue fluorescence emission[40] at 450 nm and the intensity of is enhanced with increasing concentration (Figure 3.5).[40] The succinimide derivatives **3** and **4** (Figure 3.5) show clusteroluminescence properties.[41] Compound **5** exhibited completely unexpected fluorescence (Figure 3.5).[42] The triethyl amine (Compound **6**, Figure 3.5) exhibited

blue fluorescence emission.[43] A long chain primary amine (compound **7)** [44] and the polyol xylitol[45] (compound **8**, Figure 3.5) also exhibited *clusteroluminescence* property.

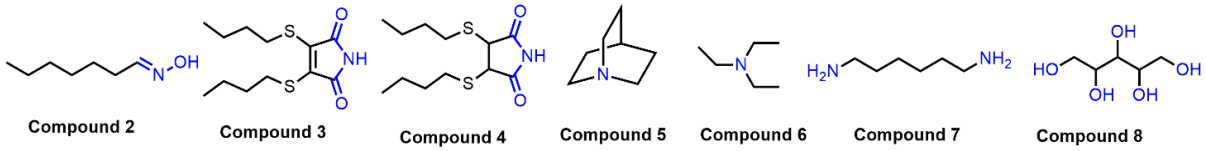


Figure 3.5. Molecular structure of *clusteroluminogens* lacking aromatic rings

3.1.7 Origin of fluorescence in clusteroluminogens

Fluorescence emission from a *clusteroluminogens* is explained based on the two mechanisms (i) formation of *clusters* and (ii) through space conjugation (TSC) induced interactions.[46] The *Clusterisation Triggered Emission* (CTE) phenomena arising from band gap (S_n/Tn to S_o) of the isolated, crosslinked, and clustered molecules is shown in Figure 3.6.[46] In the isolated molecule, the donors (lone pair of N, O or S, double bond, or aromatic group etc) are connected by a non-conjugated linker. The distance between the donor and acceptor group for n-σ*, n-π*, π-π* interaction is sufficiently large to facilitate nonbonding interactions and hence have large band gap (S_n/Tn to S_o, E_1).[46]

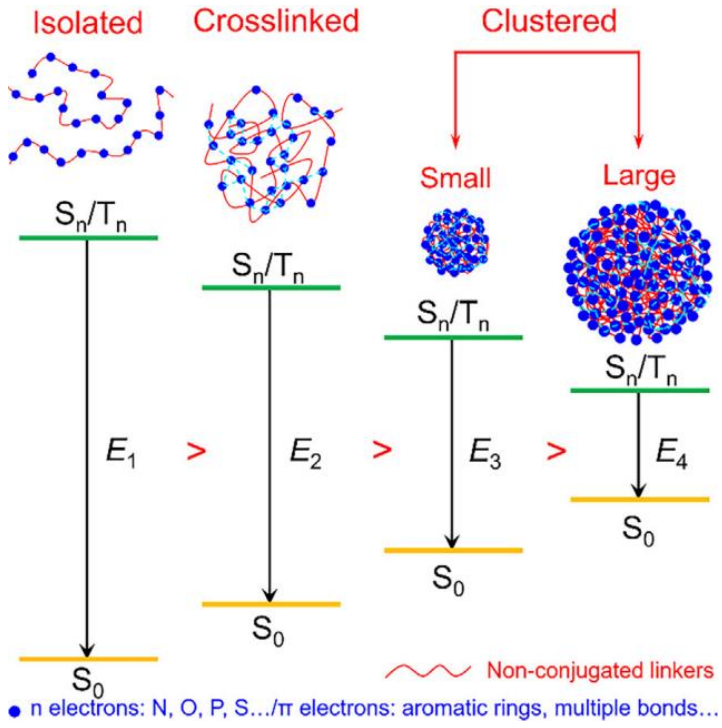


Figure 3.6. Schematic representation of the process of clusters formation accompanied with the *clusterization*-triggered emission of *clusteroluminogens.*

Crosslinking or aggregation brings the donor and acceptor groups within the bonding distance and hence participate in the n-σ*, n-π*, π-π* interaction to form a network of interacting molecules.[46] This reduces the S_n/Tn to S_o band gap in crosslinked polymers (E_2). Small and large *clusterization* further facilitate the interactions as the distances between the donor and acceptor molecules (groups) are very close in clusters compared to the isolated molecules or crosslinked polymers.[46] In clusters the interacting molecules form a large network like structure which subsequently stabilize the cluster by reducing the band gap to minimum (E_3 and E_4) resulting in a red shifted emission spectra of the polymeric *clusteroluminegens.*

The other phenomena of through space conjugation (TSC) induced interactions in *clusteroluminogens* is shown on Figure 3.7. The dipoles of C=O and C-N bonds are oriented and polarized to generate a partial positive charge on carbon and partial negative charge on nitrogen or oxygen.[46] This partial polarization facilitates the inter and intra molecular electrostatic interactions in space. Apart from the electrostatic interaction, hydrogen bonding interaction between the donor and acceptor results a network like structure in solution and solid state.[46] Other major secondary interactions like dipole-dipole interaction also contribute to the stabilization of the molecular dipoles in aggregated state or solid state. n-σ*, n-π*, π-π* interactions result an extensive overlap of orbitals in the cluster and aggregated state.[46] These interactions along with the n-σ*, n-π*, π-π* interactions reduce the band gap between S_n/Tn to S_o and hence red shifted emission spectra results.

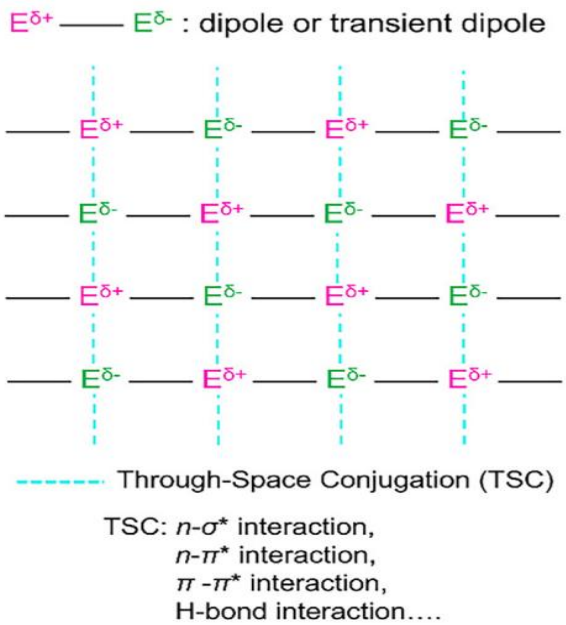


Figure 3.7. Through space conjugation mechanism to explain the fluorescence emission of *clusterolumiogens.*

3.2. Rationale and objective for the present work

In 2015 Gazit *et.al* reported the single crystal structure of GC di-PNA and unusual optical properties of GC dimer.[47] The crystal structure of GC di-PNA exhibits Watson-Crick hydrogen bonding between the two complementary strands with effective π-π stacking interaction between the nucleobases.[47] A notable feature of the crystal structure is the orientation of the tertiary amide bond in the solid state. The carbonyl group of the tertiary amide is oriented towards the glycine side of the backbone in both the G and C monomers suggesting stabilization of Z-rotamer in the solid state. Surprisingly, the powder of GC-dimer dissolved in buffer solution showed excitation dependent emission spectra, although it lacked the presence of any conventional fluorophore in the molecule.[48] Apart from the fluorescence, it exhibited electroluminescence properties. The possible reason for the fluorescence of *aeg*-PNA-GC dimer could be the stabilizing Watson-crick hydrogen bonding interactions, packing of the molecule and the efficient stacking interaction between the nucleobases,[47] possibly driven by the Z-rotamer.[47] In a follow up work published by Romanelli *et a.l* [49] it was shown that the GC di PNA aggregates exhibited bright fluorescence emission from the nanostructures formed by Fmoc GC-di PNA.

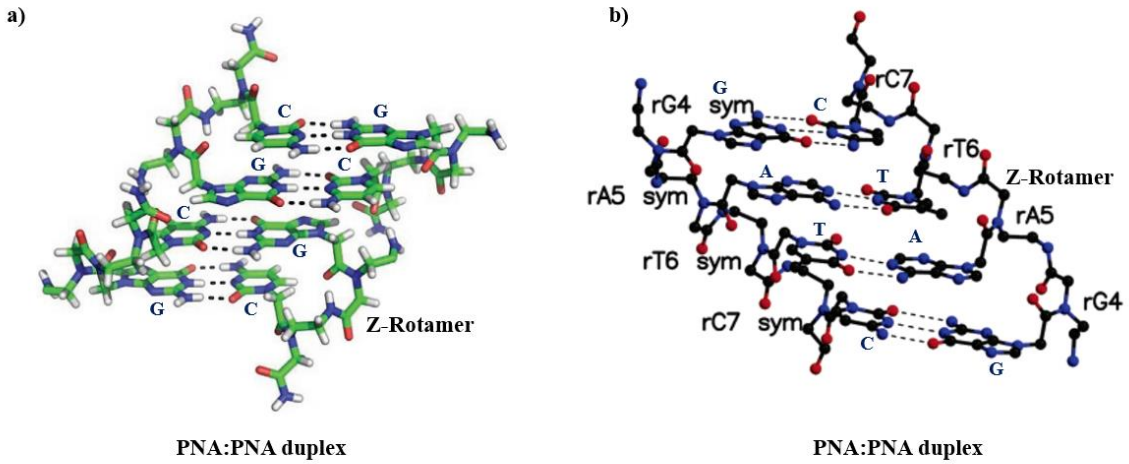


Figure 3.8. Crystal structure of PNA: PNA duplexes formed by two different sequences

In 2005, Nielsen *et. al*[50] have reported the crystal structure of PNA:PNA duplex and PNA_2:PNA triplex network. The crystal structure revealed the duplex to be right-handed and in each PNA monomer steps the rotamer is Z. In the PNA:PNA dimer duplexes reported by Gazit (Figure 3.8), the tertiary carbonyl group is oriented towards the glycine side of the PNA

backbone suggesting a Z-Rotamer orientation as in PNA monomer steps.[50] Thus though *aeg*-PNA monomers exhibited a mixture of Z and E-rotamers, in the derived oligomeric hybrids they prefer all Z-rotamer. With this literature background, it was thought interesting to construct *aeg* PNA monomer / dimer analogues that prefer E-rotamer and study the non-bonding properties including crystal packing, stacking interactions, optical properties etc induced exclusively by E-rotamers.

Recently Ganesh *et al.*[5]. reported exclusive stabilization of rotamers in either E or Z form in solution by incorporating sterically bulky *gemdimethyl (gdm)* group individually at α, β and γ positions of the *aeg*-PNA backbone (Figure 3.9). It was shown that the β-*gemdimethyl (gdm)*-PNA exclusively stabilized the E-rotamer, while the α-*gemdimethyl (gdm)*-PNA stabilized the Z-rotamer in solution. The γ-*gdm*-PNA gave a mixture of Z and E-rotamers, with

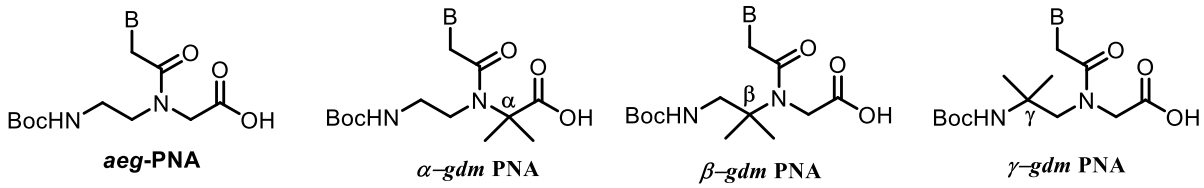


Figure 3.9. Structures of γ/α/β-*gemdimethyl* monomer and *aeg*-PNA monomer

E as major rotamer. The identity of the rotamers were proved by expected cross peaks in ^{1}H NMR NOESY spectra and supported by computational studies and crystal structure for the E-rotamer from γ-*gdm*-PNA.[51] It was shown that the reason for the selective stabilization of Z and E-rotamers is the avoidance of steric clash between the *gemdimethyl(gdm)* group and the δ methylene group of the PNA backbone and favourable CH…O interaction.[51] The γ-*gemdimethyl*-*aeg*-PNA-T and *aeg*-PNA-T show a mixture of rotamers in solution as the *aeg*-PNA lacks any *gemdimethyl* substitution, while in γ-gemdimethyl-PNA the dimethyl groups are far from the t-amide center to have any steric selectivity.[51]

It was therefore thought that it would be interesting to tune and modify the optical properties of Gazit GC-di-*aeg*-PNA by incorporation of *gemdiimethyl* moiety at C_{α} and C_{β} positions on *aeg* PNA backbone. The orientation of the nucleobases present in such *gemdimethyl* PNA monomer would be different among E and Z rotamers of PNA and hence would have different stacking attributes in the derived dimers as well. This may reflect in their fluorescence properties that may be rotamer dependent. Apart from locking in E or Z form, *gemdimethyl (gdm)* group is also known to induce helicity as seen in peptides containing

unnatural amino acid 2-aminoisobutyric acid (α-*aib*).[52] The specific effect of E and Z rotamers on the biophysical and optical properties of the derived PNAs have not hitherto received

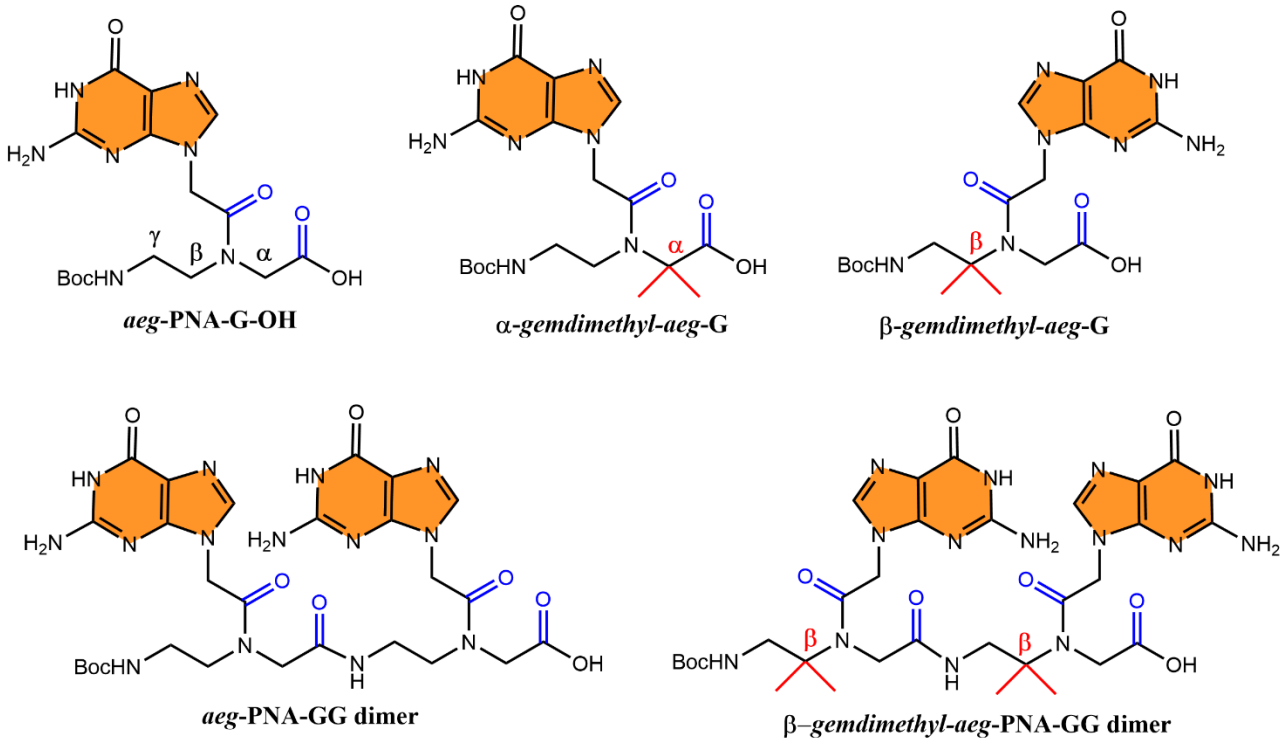


Figure 3.10. Target β/α-*gemdimethyl*-G-PNA monomers and dimers

attention. With this rationale, the target monomers and dimers of α/β-*gemdimethyl*-PNA-G monomers and dimers (Figure 3.10) were synthesized and the optical properties of the synthesized PNA monomers and dimers were studied by UV-visible and fluorescence spectroscopy. To evaluate the effect of aggregation on the fluorescence properties of PNA monomers, fatty acid (dodecanoic acid) chains are introduced in the N-terminal of the PNA monomers.

The specific objectives of this chapter are

(i) Synthesis and characterization of β/α-*gemdimethyl*-G-PNA monomers and dimers (Figure 3.10)

(ii) Conformational study of the derived monomers using 2D-NOSEY and gradient COSY.

(iii) Study of optical properties of monomer and dimers using UV-visible and fluorescence spectroscopy.

(iv) Evaluation of the effect of aggregation on the optical properties of β/α-*gemdimethyl*-G-PNA monomers.

(v) Excitation dependent emission spectral study of the synthesized monomers and dimers.

(vi) Fluorescence lifetime measurement of the PNA monomers.

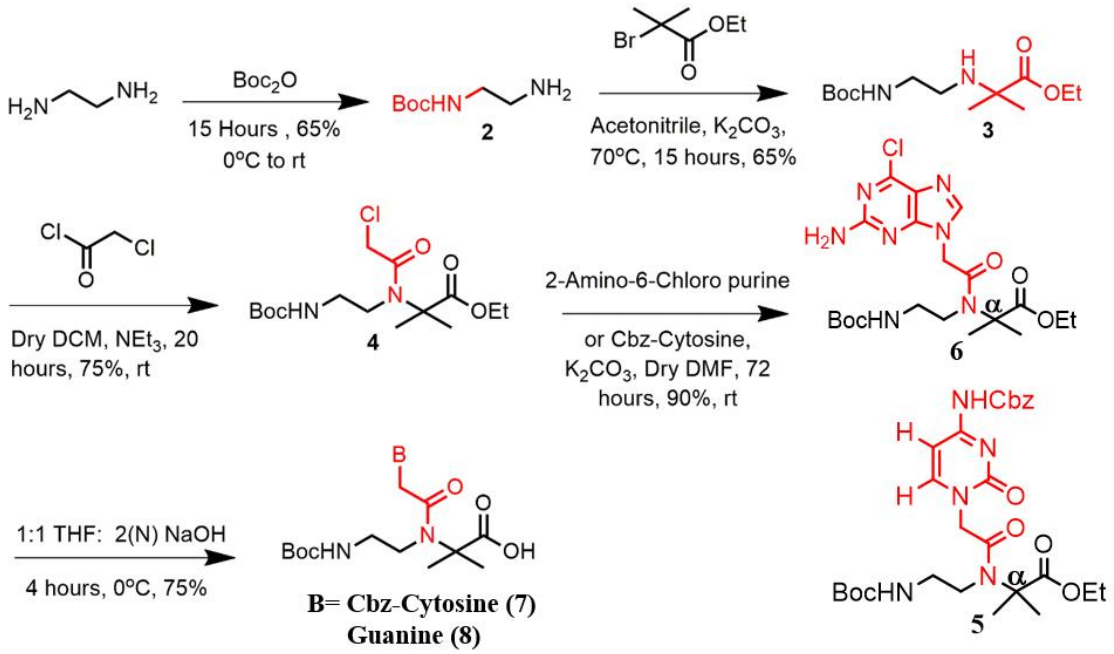
(vii) Aggregation induced emission study of the synthesized PNA monomers.

viii) Langmuir Blodgett isotherm study of the synthesized PNA monomers

3.3. Results and Discussion

3.3.1 Synthesis of α,α-*gemdimethyl*-*aeg*-G/C-PNA monomers

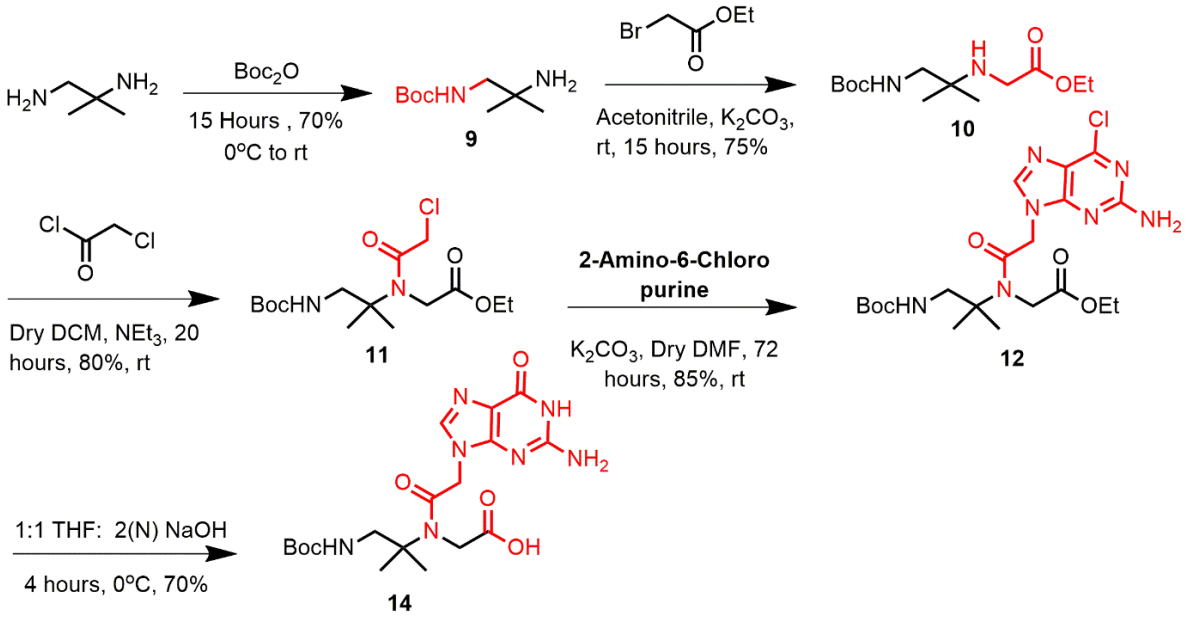
The commercially available ethylene diamine was treated with Boc_2O to afford the mono N-Boc-protected ethylene diamine (**2**), which was N-alkylated with ethyl 2-bromo isobutyrate to afford the N(Boc)-N'(α,α-*gemdimethyl*)ethyl acetate (**3**) (Scheme 1). N'-acylation of free amine backbone (**3**) with chloroacetyl chloride gave acylated N'-acetylchloro t-amide **4**.[51] This was treated with 2-amino-6-chloropurine or N^4(Cbz)-cytosine to afford the corresponding monomer esters **5** and **6** respectively. The hydrolysis of esters using 1:1 THF and aq. NaOH produced the desired α,α-*gemdimethyl*-*aeg*-PNA-G/C monomers **7** and **8**. These were characterized by ^{1}H NMR, ^{13}C NMR and mass spectrometry.



Scheme 1: Synthetic scheme of α,α-*gemdimethyl*-*aeg*-G/C-PNA monomers

3.3.2 Synthesis of β,β-*dimethyl-aeg*-G-PNA monomer

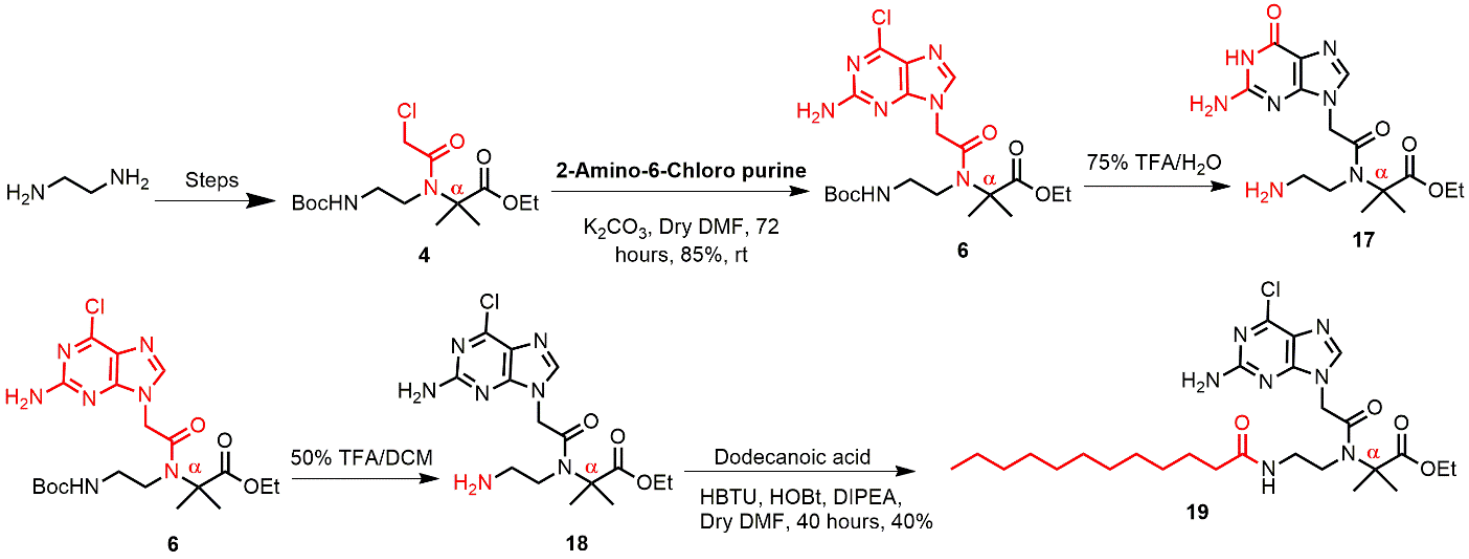
The commercially available 1,2-diamino-2-methylpropane was treated with Boc_2O to afford the mono NH(Boc)-1,2-diamino-2-methylpropane (**9**). N'-Alkylation of **9** with ethyl bromoacetate afforded the required NH(Boc)-aminoethyl-β,β-dimethyl N'-ethyl glycinate (**10**) (Scheme 2). N'-Acylation of **10** with the chloroacetyl chloride gave the NH(Boc)-N'-acetylchloro t-amide **11**.[51] This was treated with 2-amino-6-chloro purine to afford the monomer ester **12** (Scheme 2**)** which was hydrolyzed using aq. NaOH:THF to produce the desired β,β-*gemdimethyl-aeg*-G-PNA monomer (**14**). The monomers **7**, **8** and **14** were characterized by ^{1}H NMR, ^{13}C NMR and mass spectrometry. The appearance of characteristic peaks for nucleobases as well as for the backbone confirms the structures of β, β-*gemdimethyl-aeg*-G-PNA monomers.



Scheme 2: Synthetic scheme of β-*gemdimethyl-aeg*-G-PNA monomers

3.3.3 Synthesis of dodecanoyl conjugated β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (13) and β-*gemdimethyl-aeg*-G PNA monomer (16)

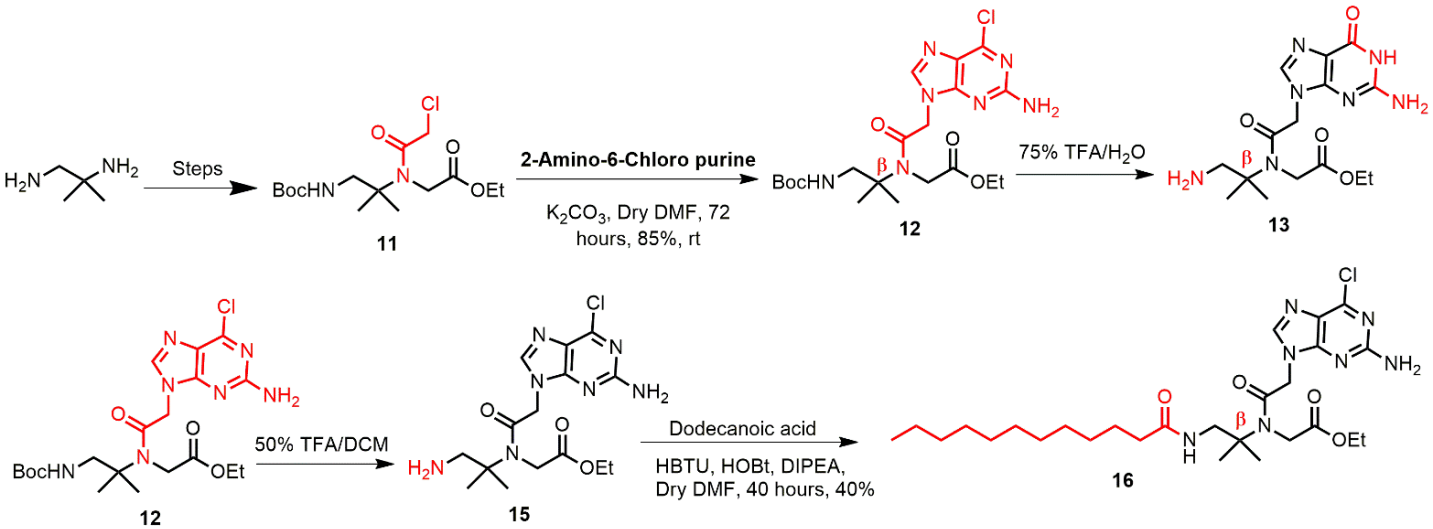
The β,β-gemdimethyl-6-chloro compound **12** was reacted with 75% TFA/water mixture to yield the free amine guanine monomer ester (**13**) (Scheme 3). Compound **12** upon treatment with 50% TFA/DCM to afforded the 6-Cl-2-aminopurine monomer **15** that was coupled with dodecanoic acid in presence of coupling agent HBTU to afford the dodecanoyl conjugated monomer **16**.



Scheme 3: Synthetic scheme for dodecanoyl conjugated β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**13)** and β-*gemdimethyl-aeg*-G PNA monomer (**16**)

3.3.4. Synthesis of dodecanoyl conjugated α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA (17) and α-*gemdimethyl-aeg*-G PNA monomer (19)

The NHBoc-α-gemdimethyl-6-chloro compound **6** was treated with 75% TFA/water mixture to yield the free amine **17** (Scheme 4). Compound **6** was treated with 50% TFA/DCM to afford the NH_2-α,α-gemdimethyl-6-chloro-2-amino purine-*aeg*-PNA monomer (**18**). The free amine in 18 was coupled with dodecanoic acid in presence of coupling agent HBTU to



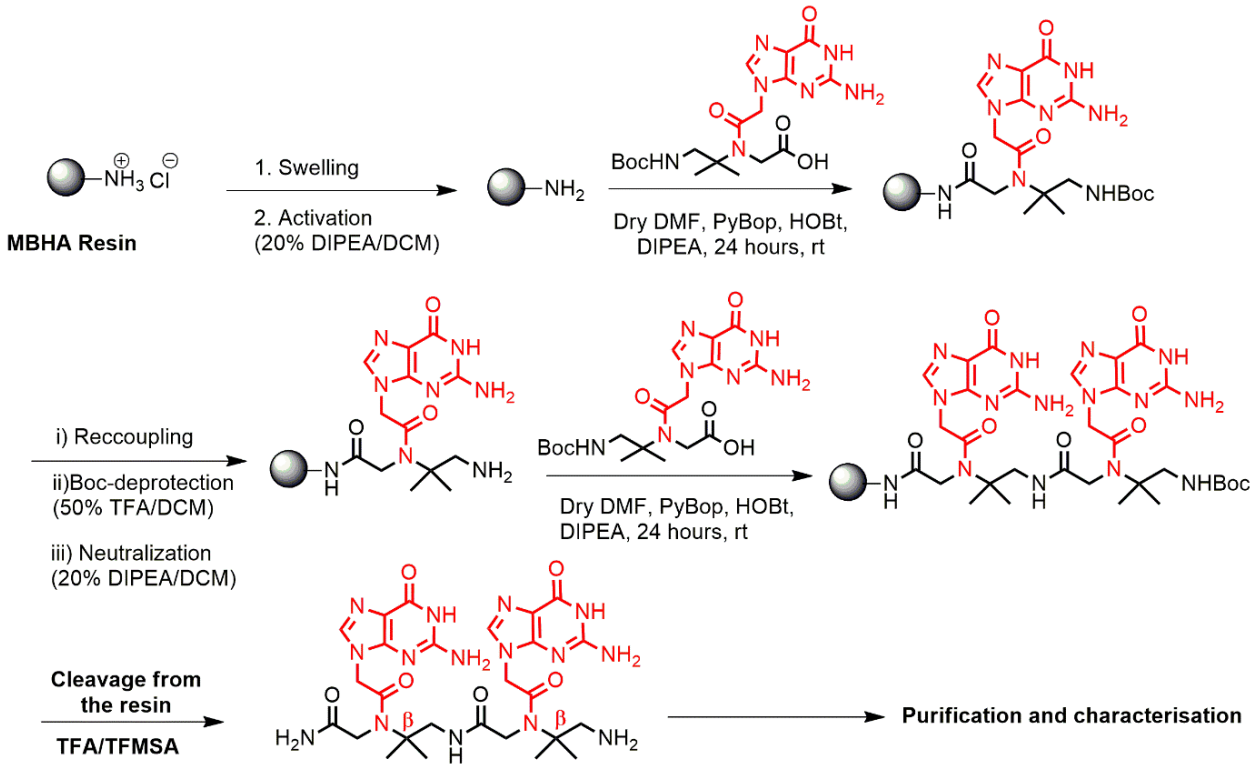
Scheme 4: Synthetic scheme for dodecanoyl conjugated α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**19)** and α-*gemdimethyl-aeg*-G PNA monomer (**17**)

afford the N^1(dodecanoyl)-α,α-gemdimethyl-6-chloro-2-aminopurine *aeg* PNA monomer (**19**). The obtained dodecanoyl conjugated α/β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA

monomer (**19, 16)** and α-*gemdimethyl-aeg*-G PNA monomer (**17, 13**) were characterized by ^{1}H NMR, ^{13}C NMR and mass spectrometry.

3.3.5 Synthesis of dimers

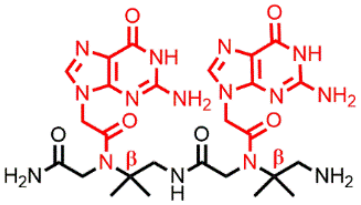
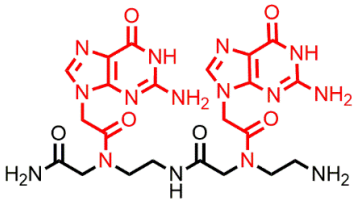
The dimers of β-*gemdimethyl-aeg*-G-PNA and *aeg*-G-PNA were synthesized using standard solid phase peptide synthesis using Boc strategy (Scheme 5).[53] The MBHA resin (loading value 1.19 mmol/gm) was swelled over DCM and activated using DIPEA in DCM. The *aeg*-G-PNA-OH monomer (**20**) was coupled with the free amino group present on the MBHA resin in presence of HBTU as a coupling agent. After coupling, the N-Boc group was deprotected using 50% TFA/DCM to generate free amine on resin and the second coupling with *aeg*-G-PNA-OH monomer gave the required (*aeg*-G-PNA)$_2$ dimer (**23**) on the resin.



Scheme 5: Strategy for the synthesis of PNA dimers

Similar procedure was followed starting with MBHA resin and coupling with the β-*gemdimethyl-aeg*-G-PNA-OH monomer (**14**) in presence of PyBop as a coupling agent, since HBTU gave lower yields. The coupling step was repeated again with β-*gemdimethyl-aeg*-G-PNA-OH (**14**) to improve the yield. This was followed by deprotection of Boc and second coupling with β-*gemdimethyl-aeg*-G-PNA-OH (**14**) monomer to generate the dimer. The resin bound PNA dimers β,β-*gemdimethyl-aeg*-PNA-GG **22)** and (*aeg*-PNA-GG **23** (Table 3.1) and

were cleaved by using TFA/TFMSA mixture, purified by reverse phase HPLC and characterized by MALDI-TOF mass spectrometry.[54] The structure of synthesized PNAs and their observed mass with the HPLC retention time are shown in Table 3.1.

Sr. No	Structure of the Dimer	Calculated mass	Observed mass	HPLC Retention time (minutes)
1	 $C_{26}H_{37}N_{15}O_6$ **22**	**694.2688** **(M+K)$^+$**	**693.2264** **(M+K)$^+$**	**12.36**
2	 $C_{22}H_{29}N_{15}O_6$ **23**	**622.2323** **(M+Na)$^+$**	**622.5656** **(M+Na)$^+$**	**10.50**

*Gradient used for HPLC purification (50% ACN/water, 95% water/ACN, C_{18} column)

Table 3.1. Structures, mass spectra and HPLC data of dimers.

3.3.6. Conformational study of the monomers by ^{1}H NMR

It is essential to establish the conformations of PNA monomers as it is rationalized that orientation of the monomers may influence the optical properties of the derived β/α-*gemdimethyl*-G-PNA monomers. As illustrated in the Chapter 2, the presence of rotamers can be detected by the appearance of double peaks for each methylene protons at β, α and δ in ^{1}H NMR spectra.[51] In Z-rotamer, the δ-CH_2 is closer to β/γ protons of the backbone and hence the Z-Rotamer is characterised by the NOE cross peak between side chain δ-CH_2 and β/γ protons of the backbone (Figure 3.11).[51] In the E-rotamer the δ-methylene is closer to α-methylene protons and characterised by the NOE cross peak between them (Figure 3.11).[51]

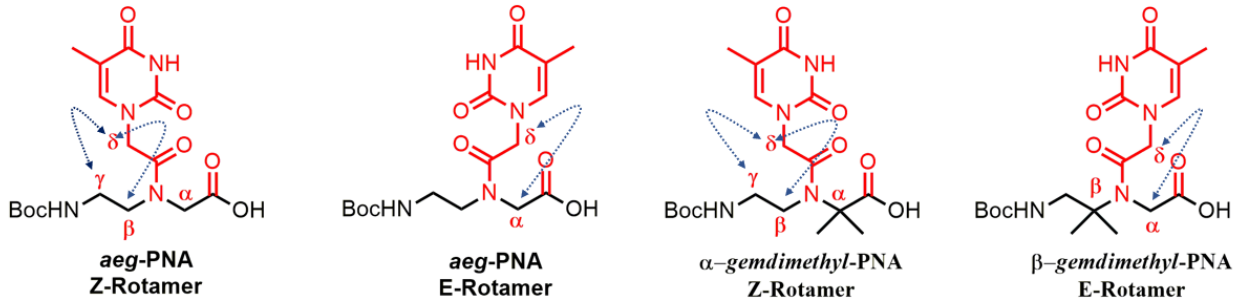


Figure 3.11. Rotamers of *aeg*-PNA monomer, β/α-*gemdimethyl*-*aeg*-PNA monomer and the possible NOSEY cross peaks

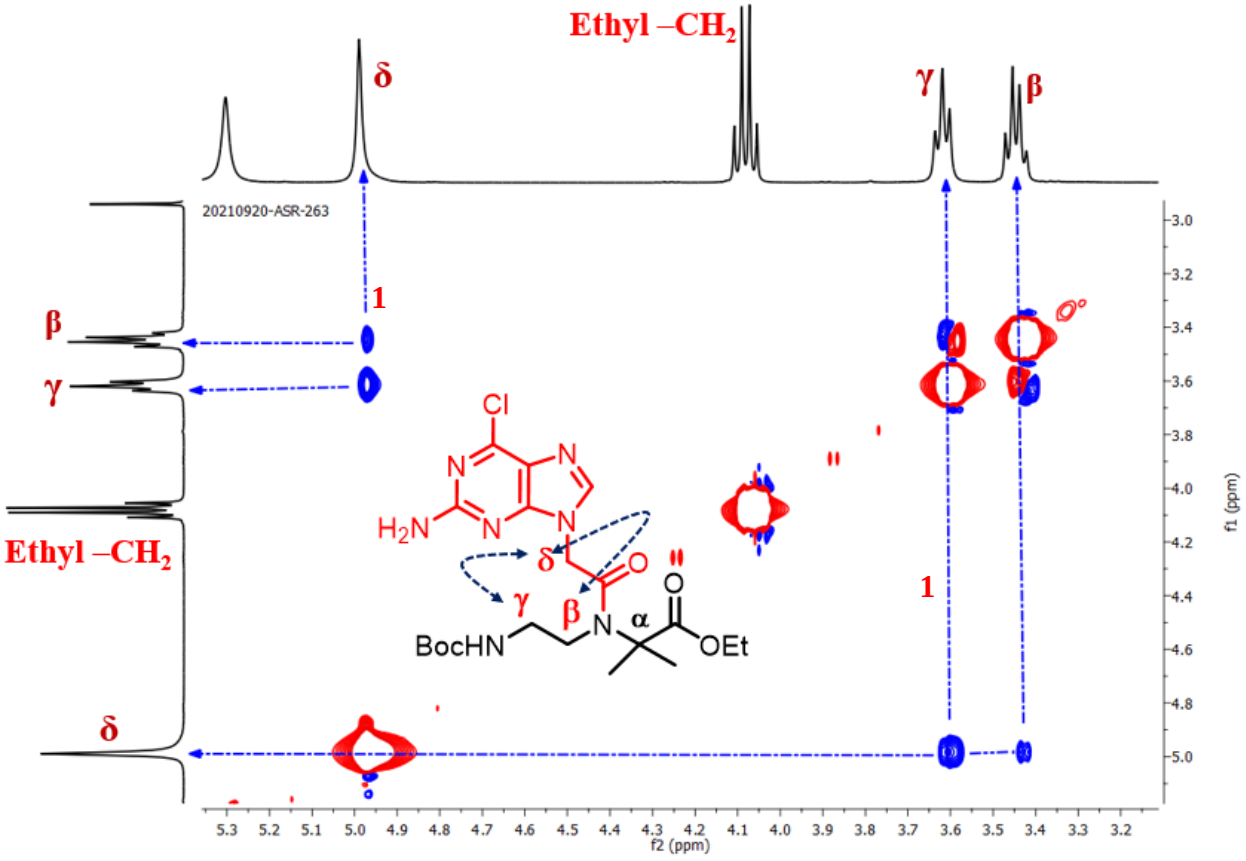


Figure 3.12. ^{1}H-^{1}H NOSEY of α,α-*dimethyl*-(6-Cl-2-amino)-*aeg*-purine PNA (**6**, DMSO-d_6)

The primary ^{1}H NMR analysis of α,α-*dimethyl*-*aeg*-G-PNA (**8**), α,α-*dimethyl*-*aeg*-(6-Cl-2-aminopurine)-PNA (**6**) and α,α-*dimethyl*-(N^4-Cbz)-*aeg*-C-PNA (**5**) revealed a single rotamer in solution. Any ambiguity regarding the chemical shift assignment of α, β and γ-methylene is ruled out as *gemdimethyl* group is present at α-C or β-C, making chemical shift assignments easier. The partial ^{1}H-^{1}H NOSEY spectra of the α,α-*dimethyl*-*aeg*-(6-Cl-2-aminopurine) PNA (**6**) in DMSO-d_6 is shown in Figure 3.12. The side chain δ-CH_2 shows cross peak with the β/γ-methylene protons of the backbone (cross peak 1, blue dotted line) suggesting

the presence of Z-rotamer for α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA (**6**). α,α-*dimethyl-aeg*-G-PNA (**8**) and α,α-*dimethyl-aeg*-(N^4-Cbz)-C PNA (**5**) also show single Z-rotamer with the side chain δ-CH_2 showing a cross peak with the β/γ-methylene protons of the backbone.

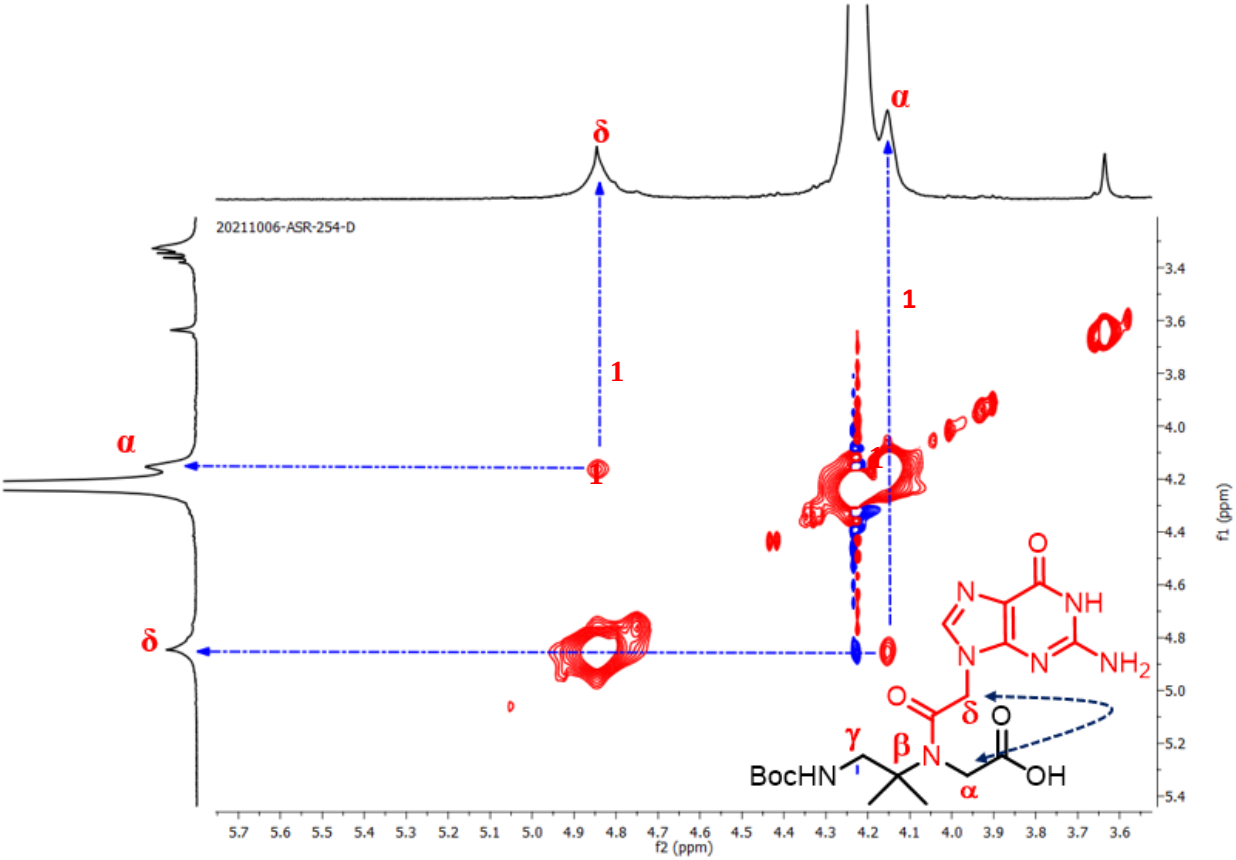


Figure 3.13. ^{1}H-^{1}H NOSEY of β-*gemdimethyl-aeg*-G PNA (**14**) in DMSO-d_6

The primary ^{1}H NMR analysis of β-*gemdimethyl-aeg*-G PNA monomer (**14**) and β,β-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**12**) revealed a single rotamer in solution. (Figure 3.13). In gradient COSY spectra, no cross peak was observed between the methylene protons. The partial ^{1}H-^{1}H NOSEY spectra of the β-*gemdimethyl-aeg*-G PNA (**14**) (Figure 3.13) shows cross peak of side chain δ-CH_2 with the α-CH_2 of the backbone (cross peak 1, blue dotted line) suggesting the presence of only E-Rotamer. β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA (**12**) also shows only E-rotamer with the side chain δ-CH_2 having cross peak with the α-CH_2 of the backbone.

The primary ^{1}H NMR analysis of decanoyl conjugated monomers **16** and **19** also shows single rotamer in solution. The backbone β/γ and α-CH_2 protons were assigned using the established protocol. The partial ^{1}H-^{1}H NOSEY spectra of the N-decanoyl β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**16**) (Figure 3.14), shows the side chain δ-CH_2 having

cross peak with the α-CH_2 of the backbone suggesting preferential stabilization of E-Rotamer cross peak 1, blue dotted line). On the other hand, the corresponding N(decanoyl)-α-*gemdimethyl-aeg*-(6-Cl-2-amino purine) PNA monomer (**19**) shows the opposite trend with the side chain δ-CH_2 showing cross peak with the β/γ-CH_2 of the backbone, suggesting a Z-rotamer in solution. In summary, the fatty acid conjugates of the α/β-*gemdimethyl-aeg*-PNA monomers (**16** and **19**) retain single rotamer nature suggesting a steric effect of *gemdimethyl* substitution towards the preferential stabilization of E/Z rotamers in solution.

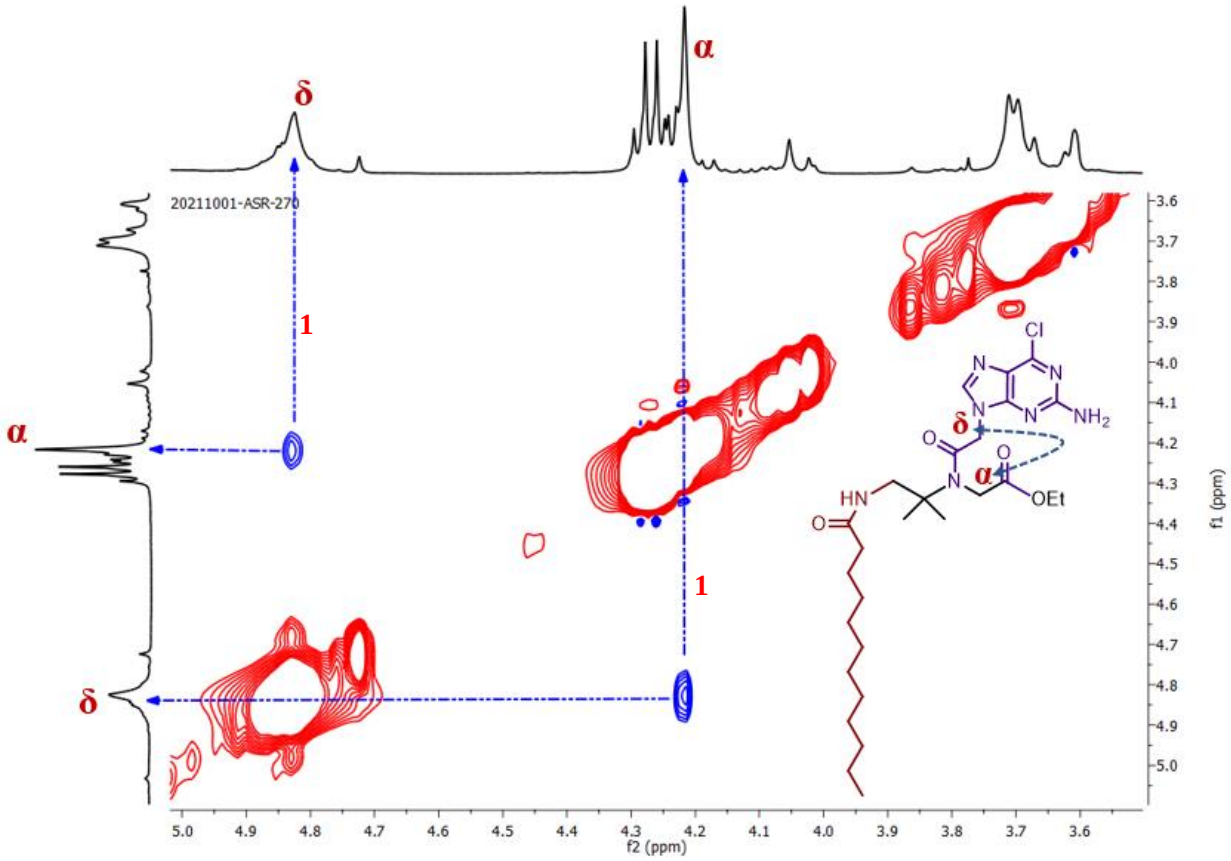


Figure 3.14. ^{1}H-^{1}H NOSEY of N-(decanoyl)-β-*gemdimethyl-aeg*-(2-Cl-6-aminopurine) PNA monomer (**16**) in $CDCl_3$

The free amine of β-*gemdimethyl-aeg*-G-PNA monomer (**13**) exhibited rotamer mixture in solution suggesting the role of N-terminal amide group for the stabilization of E-rotamer of β-*gemdimethyl-aeg*-PNA G monomer (**14**) in solution. The β-*gemdimethyl* group folds the monomer in an extended γ-turn fashion and perhaps facilitate H-bonding between N-terminal NH and ester carbonyl group.[61] Such a possibility is supported by the crystal structure of N-acetyl chloro compound **11** (Figure 3.15) in which the terminal carbamate NH and ester carbonyl group are in vicinity of the H-bonding distance 2.187 Angstrom, with the rotamer in E-form. In comparison, the corresponding free amine of α-*gemdimethyl-aeg*-G-PNA monomer

(**17**) exhibited a single Z-rotamer in solution. which the *aeg*-PNA-G-OH monomer (**20**) lacks. Table 3.2 shows the list of monomers and their preferred rotamers in solution.

Sr. No	Monomer (compound number)	Rotamer	Preferred rotamer
1	α-*gemdimethyl-aeg*-PNA-G **(8)**	Single rotamer	**Z-Rotamer**
2	α- *gemdimethyl*- (6-Cl-2-amino) purine PNA **(6)**	Single rotamer	**Z-rotamer**
3	α- *dimethyl*-(N^4-Cbz)-C *aeg* PNA **(5)**	Single rotamer	**Z-Rotamer**
4	β-*gemdimethyl-aeg*-PNA-G **(14)**	Single rotamer	**E-Rotamer**
5	β- *gemdimethyl*- (6-Cl-2-amino) purine PNA **(12)**	Single rotamer	**E-Rotamer**
6	N-decanoyl β, β-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer **(16)**	Single rotamer	**E-Rotamer**
7	N-decanoyl α, α-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer **(19)**	Single rotamer	**Z-Rotamer**
8	Free amine α-*gemdimethyl*-G-PNA **(17)**	Single rotamer	**Z-Rotamer**
9	Free amine β-*gemdimethyl*-G-PNA **(13)**	**Mixture**	**Mixture of rotamer**

Table 3.2. *gemdimethyl-aeg*-PNA monomers and preferred rotamers

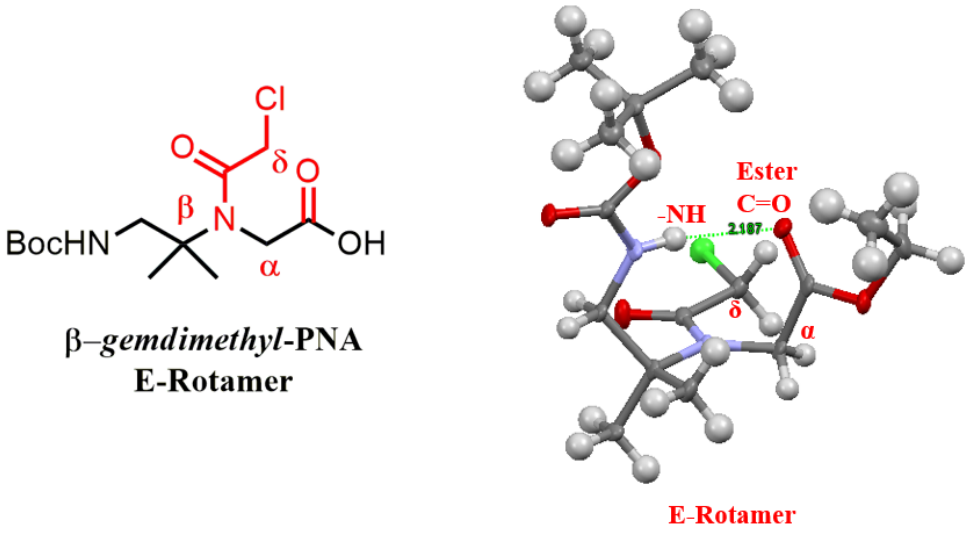


Figure 3.15. Crystal and molecular structure of **11**

3.3.7 Fluorescence spectroscopic study of the monomers

To examine the effect of rotamers on the optical properties of the α/β-*gemdimethyl-aeg*-G PNA monomers (**8** and **14**), fluorescence studies of various synthesized monomers were

undertaken. UV-visible spectra of unsubstituted *aeg*-PNA-G monomer (**20**) shows characteristic peak of nucleobase at 254 nm (Figure 3.16a). Upon excitation at 220-240 nm, emission bands appeared at 302 nm (Figure 3.16b) but at higher excitation wavelengths, the 302 nm band disappeared and very low intensity emission bands were seen around 370 nm. This suggested negligible fluorescence property of the *aeg*-PNA-G (**20**) monomer in solution.

Interestingly, the UV-visible spectra of β-*gemdimethyl-aeg*-PNA G monomer (**14**) in ethanol shows a strong absorption band at 306 nm along with the characteristic UV band for guanine at 246 nm (Figure 3.17a). The 306 nm UV band is characteristic for β-*gemdimethyl* substituted *aeg*-G PNA monomer and it is absent in unsubstituted *aeg*-PNA-G monomer (**20**) that lacked any UV band beyond 290 nm. The excitation dependent emission spectra of β-*gemdimethyl-aeg*-G-PNA monomer (**14**) in ethanol revealed interesting emission band patterns (Figure 3.17b). Excitation at wavelengths 307-337 nm led to emission band at 370 nm and as λ_{ex} was increased to higher excitation wavelengths (357-407 nm), the intensity of λ_{em} at 307 nm gradually decreased with emergence of new emission band with maxima at 475 nm (Figure 3.17b). At λ_{ex} of 327 nm the emission bands at 370 nm and 475 nm were simultaneously observed (Figure 3.17b). At excitations beyond 370 nm, emission band at 475 nm increased and got saturated at excitation beyond 400 nm. These results suggested strong fluorescence emission property of the β-*gemdimethyl-aeg*-G PNA monomer (**14**) with visible green fluorescence in solution. *Thus, the β-gemdimethyl-aeg-G PNA monomer* (**14**) *that lacks conventional fluoro-chromophore shows intense fluorescence property in solution unlike unsubstituted aeg-PNA monomer* (**20**).

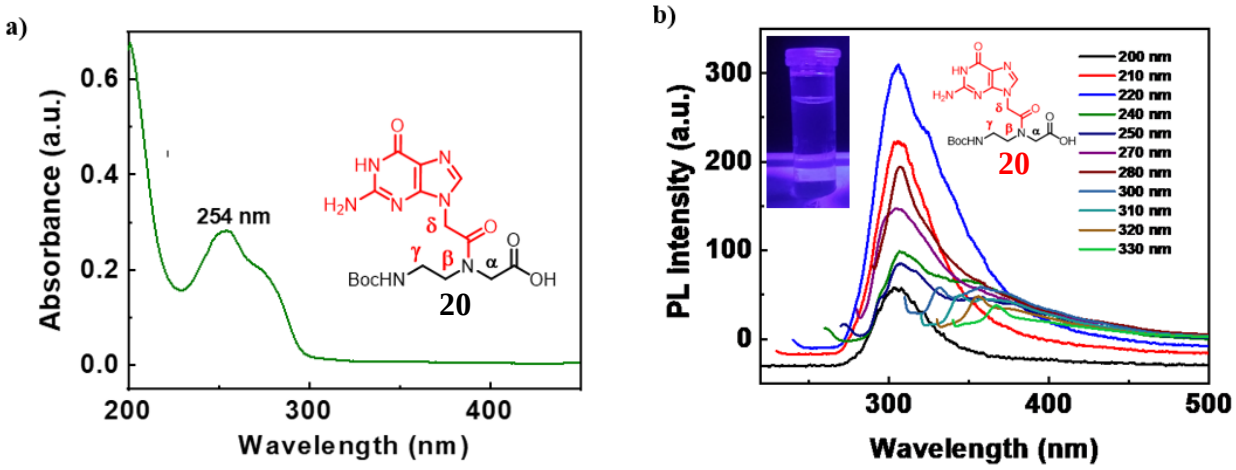


Figure 3.16. (a) UV-Visible spectra of *aeg*-PNA-G (**20**, concn. 200 μM) (b) Fluorescence spectra *aeg*-PNA-G monomer λ_{ex} 220-240 nm (concn, 200 μM); Solvent ethanol

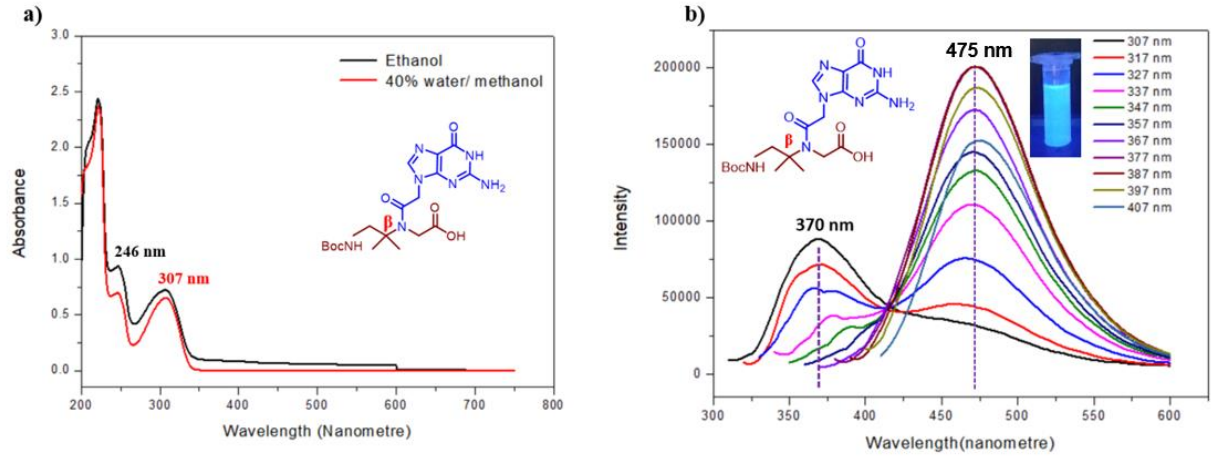


Figure 3.17. a) UV-Visible spectra of β-*gemdimethyl-aeg*-PNA G (**14**, concn, 210 µM), (b) Fluorescence spectra β-*gemdimethyl-aeg*-PNA-*G* monomer (concn, 210 µM). (Solvent-Ethanol)

The UV-visible spectra of α-*gemdimethyl-aeg*-G-PNA monomer (**8**) exhibits a strong absorption band at 308 nm and the characteristic band for nucleobase at 245 nm (Figure 3.18a). The excitation dependent emission spectra of α-*gemdimethyl-aeg*-G-PNA monomer (**8**) revealed an entirely different emission band pattern. Excitation at lower wavelengths (250-290 nm) showed a weak emission maximum at 340 nm, which gradually disappears as excitation wavelength is increased. This was accompanied by appearance of vibrationally resolved multiple emission bands at maxima at 402 nm and 428 nm respectively (Figure 3.18b), and λ_{ex} at 370 nm showed maximum emission intensity and visible observation of blue

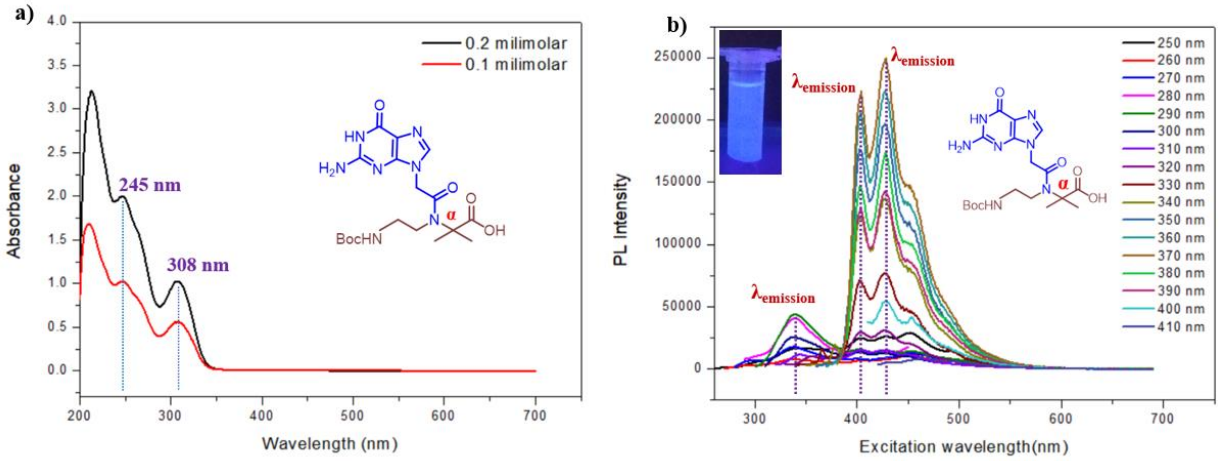


Figure 3.18. a) UV-Visible spectra of α-*gemdimethyl-aeg*-PNA-G (**8**, concn, 200 µM) b) Fluorescence spectra α-*gemdimethyl-aeg*-G-PNA monomer (concn, 220 µM, Solvent-Ethanol)

fluorescence in solution. The different patterns of emission bands for α-*gemdimethyl-aeg*-G-PNA (**8**) and β-*gemdimethyl-aeg*-G-PNA (**14**) monomers that have exclusive Z and E rotamers respectively is significant. From NMR, it was observed that E and Z-rotamers have different patterns of intramolecular hydrogen bonding, which may elicit different stacking and dipole-dipole interactions and hence differing through space conjugation (TSC). This may result in different band gap between the ground and excited states for β and α-*gemdimethl*-aeg-G-PNA monomers (**14** and **8**) leading to different patterns of emission bands. The stoke shift decreases with increasing the excitation wavelength for both β and α-*gemdimethyl-aeg*-PNA-G monomers (See appendix-II). Thus, incorporation of *gemdimethyl* group in the β/α-position of *aeg*-PNA-G (**20**) induces fluorescence properties in the resulted β/α-*gemdimethyl-aeg*-G monomers (**14** and **8**).

3.3.7a *Effect of solvents*: Fluorescence emission measurements of β-*gemdimethyl-aeg*-G-PNA monomer (**14**) in different solvents such as DMSO, ethanol and 40% water-methanol were done to examine the effect of solvent polarity (Figure 3.19). The emission intensity decreased in the order of polarity DMSO > ethanol > 40% water/methanol (Figure 3.19) and the λ_{em} maxima decreased in the reverse order: ethanol (337 nm), 40% water/methanol (336 nm) and DMSO (334 nm). Interestingly the corresponding β-*gemdimethyl-aeg*-(6-C-2-aminopurine)-PNA (**12**) monomer is non fluorescent.

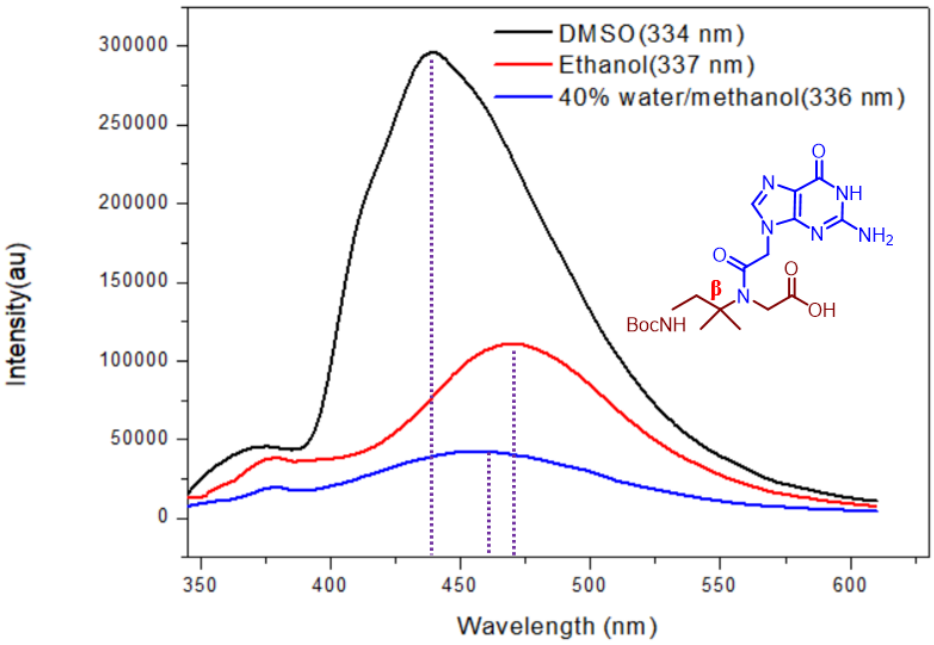


Figure 3.19. Solvent dependent fluorescence of β-*gemdimethyl-aeg*-G-PNA monomer (**14**)

However, α-*gemdimethyl-aeg*-(6-chloro-2-aminopurine)-PNA monomer (**6**) and α-*gemdimethyl-aeg*-(N^4-Cbz-C) PNA monomers (**5**) exhibited fluorescence spectra. The UV-visible spectra of the α-*gemdimethyl-aeg*-(6-chloro-2-aminopurine) PNA monomer (**6**) displayed absorption bands 221 nm, 248 nm and 308 nm (Figure 3.20a). The excitation dependent emission spectra of α-*gemdimethyl-aeg*-(6-chloro-2-aminopurine)-PNA monomer (**6**) showed emission bands with λ_{em} at 325 nm with λ_{ex} in the region 270-300 nm. At higher excitation wavelengths of 320-400 nm, this band disappeared and new vibrationally resolved multiple emission bands appeared at 402 nm and 428 nm respectively (Figure 3.20b). The Stokes shift decreases with increasing excitation wavelength (see appendix). Thus α-*gemdimethyl-aeg*-(6-chloro-2-aminopurine)-PNA monomer (**6**) emitted fluorescence, while the corresponding β-analog (**12**) was non-fluorescent.

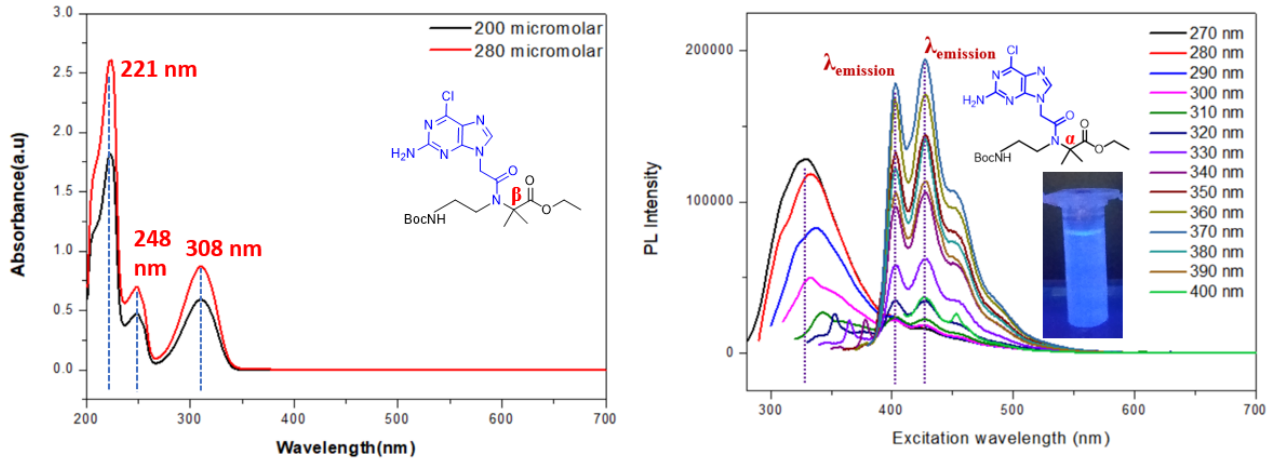


Figure 3.20. a) UV-Visible spectra of α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA (**6**,concn, 200 μM) b) Fluorescence spectra *α-gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (concn, 200 μM); Solvent: EtOH.

Fluorescence emission of α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**6**) was measured in different solvents of variable polarity such as DMSO, ethanol, 50% water/ethanol and 50% ethanol/DCM at constant excitation wavelength. Higher fluorescence emission was seen in more polar solvents (50% ethanol/DCM and DMSO) compared to less polar solvents (Figure 3.21). The emission maxima also vary with the solvent with 50% ethanol/water showing higher emission followed by DMSO, ethanol and 50% ethanol/DCM. Ethanol exhibits higher emission maxima compared to 50% water/ethanol for the peak I_1 but shows similar maxima for I_2, with the emission maxima decreasing in the order DMSO>50% ethanol/DCM> ethanol> 50% ethanol/water. These results suggest that fluorescence emission

of α-*gemdimethyl*-*aeg*-(6-Cl-2-aminopurine) PNA monomer (**6**) is sensitive to the change of solvent environment and around the fluorophore.

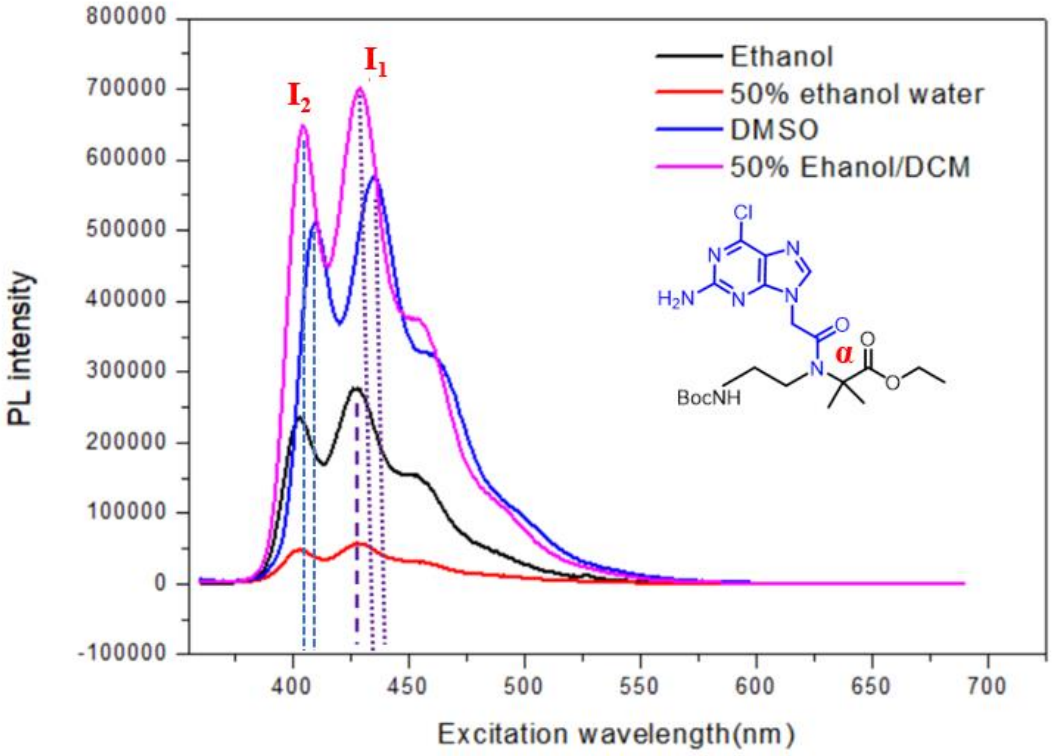


Figure 3.21. Effect of solvent on the fluorescence emission of α,α-*dimethyl*-aeg-(6-Cl-2-aminopurine) PNA monomer (**6**) monomer (concn, 220 μM)

3.3.8 Fluorescence emission of *gemdimethyl* pyrimidine PNA monomers

The commercially available *aeg*-C(N^4-Cbz)-PNA monomer (**21**) is non-fluorescent. The UV-Visible spectra of the synthesized *aeg*-C(N^4-Cbz)-PNA monomer (**21**) shows absorption bands 296 nm and 241 nm band which are characteristic bands for nucleobase (Figure 3.22b). Unlike purines where the *gemdimethyl* substitution in the PNA backbone shows unique 308 nm band (Figure 3.17a, 3.18a) compared to the *aeg*-analogue, the pyrimidines show UV spectral pattern similar to *aeg* PNA monomers (Figure 3.22), suggesting differential electronic properties of purines and pyrimidines in the *gemdimethyl* substituted PNA backbone. The influence of α-*gemdimethyl* substitution into *aeg*-PNA backbone is reflected in the fact that α-*gemdimethyl*-*aeg*-C(N^4-Cbz)-PNA monomer (**5**) is fluorescent compared to *aeg*-C(N^4-Cbz)-PNA monomer (**21**) which is non-fluorescent. The initial λ_{em} at 360 nm at lower λ_{ex} (260-310 nm) disappears at a higher λ_{ex} (330-400 nm) and results in appearance of vibrationally resolved multiple emission bands at λ_{em} 402 nm and 428 nm respectively (Figure 3.23a). Excitation at 370 nm exhibits maximum emission intensity at 428 nm (Figure 3.23a). In

contrary, the *aeg*-C(N^4-Cbz)-PNA (**21**) shows emission maxima of 350 nm when excited at 250-310 nm and the emission band disappears at higher λ_{ex} (330-400 nm, Figure 3.23b).

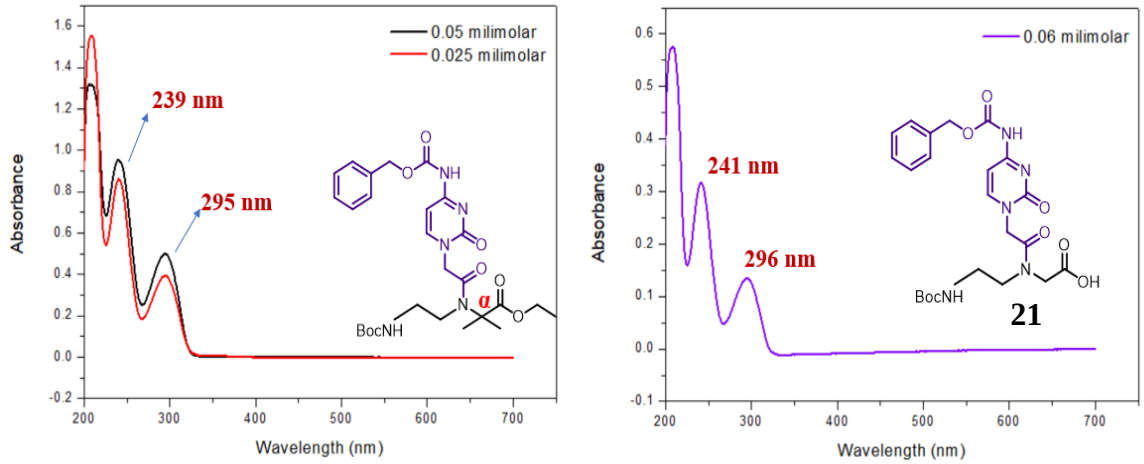


Figure 3.22. a) UV-Visible spectra of α-*gemdimethyl*-C(N^4-Cbz)-*aeg*-PNA monomer (**5**, concn, 50 μM). (b) UV-visible spectra of *aeg*-C(N^4-Cbz)-PNA monomer (**21**, concn, 60 μM); Solvent-ethanol

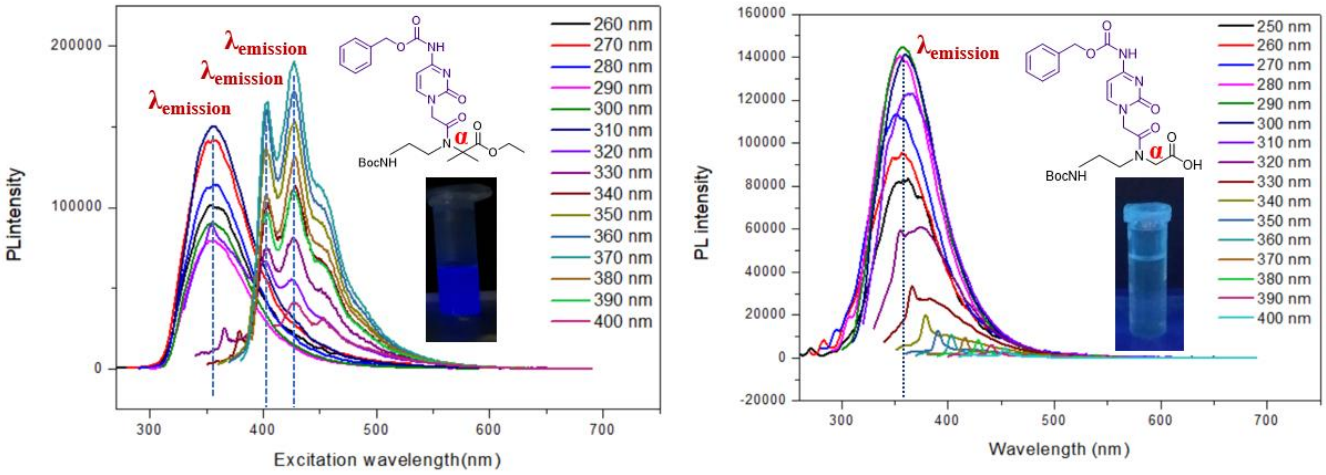


Figure 3.23. a) Fluorescence emission spectra of α-*gemdimethyl-aeg*-C(Cbz)-PNA monomer (**5**, conc, 200 μM). b) Fluorescence emission spectra of *aeg*-C(N^4-Cbz)-PNA monomer (**21**, conc. 200 μM)

3.3.8a. Fluorescence spectra of *gemdimethyl* purine PNA monomers (free amine): To examine the role and effect of N-protecting groups, fluorescence of N(Boc) deprotected *gemdimethyl* purine PNA monomers (**13** and **17**) were studied. The UV-visible spectra of free amine β-*gemdimethyl-aeg*-G-PNA monomer (**13**) and free amine α-*gemdimethyl-aeg*-G-PNA (**17**) shows a single absorption band at 254 nm and 205 nm. The strong 308 nm band observed for protected α and β-*gemdimethyl(gdm)-aeg*-G PNA monomers (**8** and **14**) completely disappeared (Figure 3.24). The excitation dependent emission spectra of free amine β-*gemdimethyl-aeg*-G PNA monomer (**13**) shows emission band at 360 nm in the lower excitation

wavelength (310 nm), and at higher excitation wavelength (330-370 nm) shifts to 425 nm (Figure 3.25a).

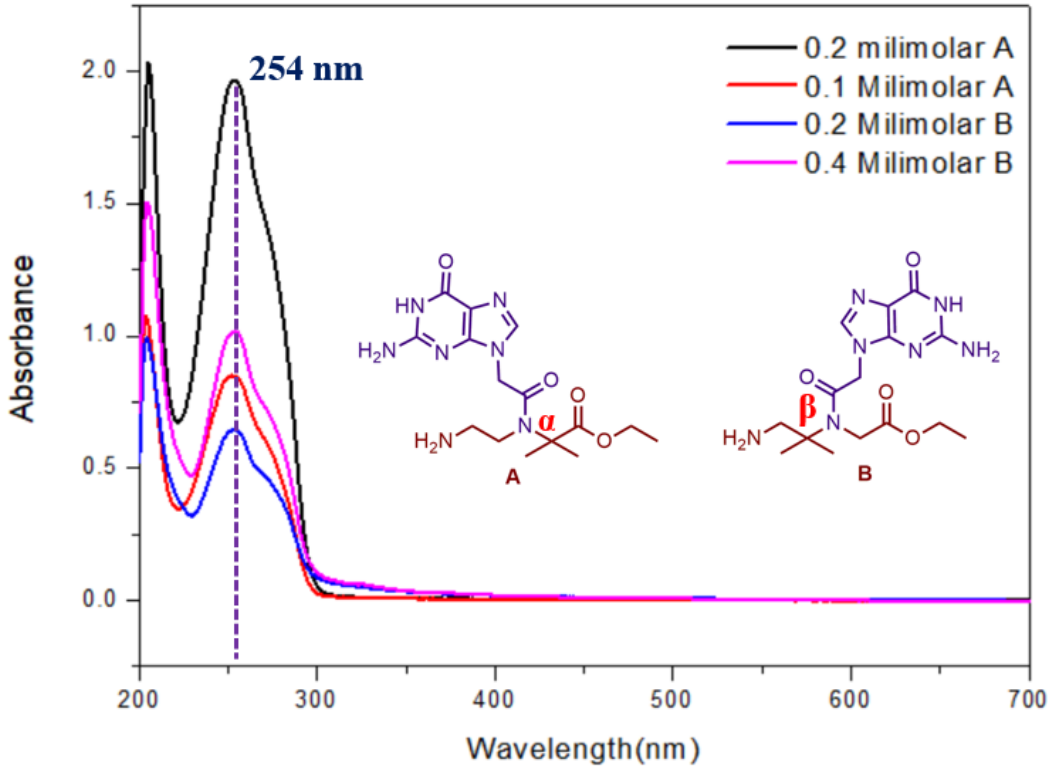


Figure 3.24. UV-Visible spectra of free amine α and β-*gemdimethyl-aeg*-G PNA monomers in ethanol (**13** and **17**, conc, 200 µM and 220 µM)

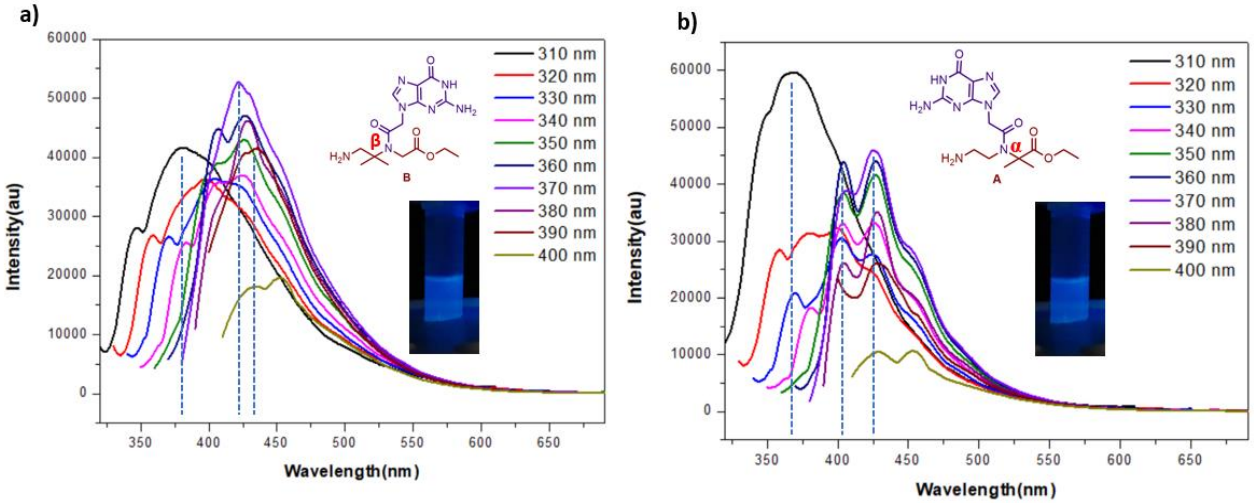


Figure 3.25. (a) Fluorescence emission spectra of free amine β-*gemdimethyl-aeg*-G-PNA monomer in ethanol (13, Concn. 210 µM). (b) Fluorescence emission spectra of free amine α-*gemdimethyl-aeg*-G-PNA monomer in ethanol (17, Concn. 210 µM).

Similarly, the excitation dependent emission spectra of free amine α-*gemdimethyl-aeg*-G PNA monomer (**17**) shows emission band at 370 nm in the lower λ_{ex} (310 nm) and at higher λ_{ex} (320-400 nm) leads to vibrationally resolved multiple emission spectral bands with λ_{em} at 402 and 428 nm respectively (Figure 3.25b). The reduced intensity of the fluorescence emission of the α/β-*gemdimethyl-aeg*-G PNA free amine monomers (**17** and **13**) may indicate lesser and ineffective TSC interaction compared to the protected monomer. The pattern of emission spectra remains unaffected in the deprotected monomers. The results suggest the important role of N-terminal NH and C=O group in the fluorescence emission property of α and β-*gemdimethyl-aeg*-G monomers (**8** and **14**). Irregular spectral pattern of free amine β-*gemdimethyl(gdm)-aeg*-G PNA monomer (**13**) may arise due to the presence of both E and Z rotameric forms in solution.

3.3.9. Effect of fatty acid conjugation on the fluorescence properties of α/β-*gemdimethyl-aeg*-G(Cl) monomers

Fatty acids and surfactants have a unique property of micelle formation and aggregation induced by micellization could influence the fluorescence emission of α/β-*gemdimethyl(gdm)-aeg*-(6-Cl-2-amino-purine)/G PNA monomers (**19** and **16**). The fatty acid conjugated monomers of N-decanoyl-β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) and N-decanoyl α-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) display UV-visible spectra as shown in the Figure 3.26. The N-decanoyl β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) exhibits UV bands at 308 nm, 245 nm and 221 nm (Figure 3.26). In contrary N-decanoyl α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) *exhibits* band at only 296 nm and 245 nm and devoid of bands at > 300 nm (Figure 3.26). The different rotameric forms of N-decanoyl α-*gdm-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) (100% Z-Rotamer) and N- decanoyl β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) *(100% E-Rotamer)* monomers may be the reason to this difference in UV bands and also the different influences of aggregation by fatty acid component.

In fluorescence spectra, N-decanoyl-β-*gdm*-(6-Cl-2-aminopurine)-PNA monomer (**16**) and N-decanoyl-α-*gdm*-(6-Cl-2-aminopurine)-PNA monomer (**19**) exhibit excitation dependent emission spectra, with α-*gdm* monomer (**19**) showing vibrationally resolved spectrum with λ_{em} at 402 nm and 428 nm with λ_{ex} at 372 nm excitation showing maximum intensity of the fluorescence emission (Figure 3.27a). Similarly, the excitation dependent emission spectra of N-decanoyl β-*gdm*-(6-Cl-2-aminopurine)-PNA monomer (**16**) exhibits

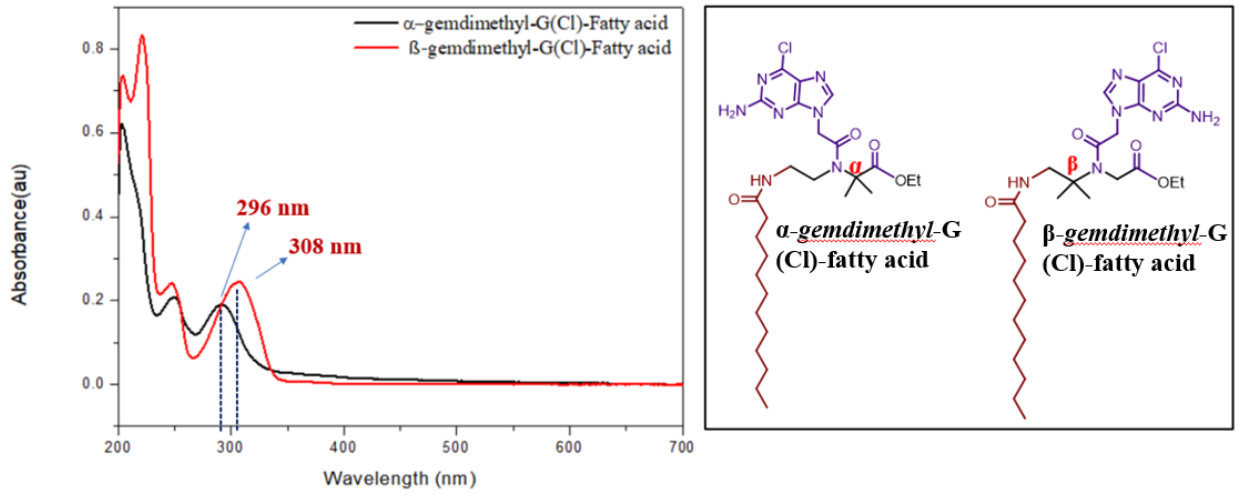


Figure 3.26. UV-Visible spectra of N-decanoyl β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**, concn. 210 µM) and N-decanoyl α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**, concn. 200 µM); Solvent, Ethanol

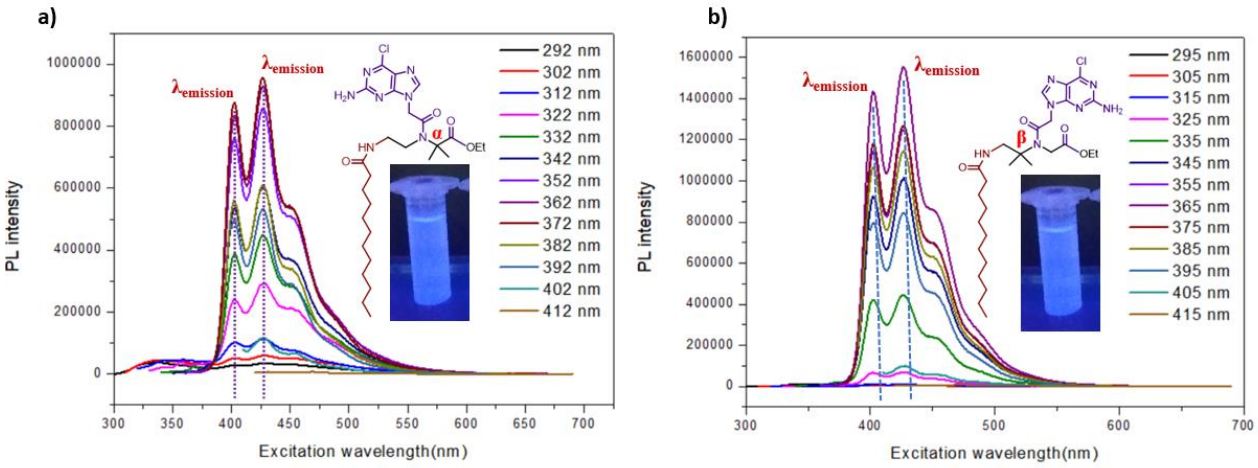


Figure 3.27. a) Fluorescence emission spectra of N-decanoyl α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**, concn. 230 µM). b) Fluorescence emission spectra of N-decanoyl β,β-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**, concn. 230 µM, Solvent -Ethanol)

vibrationally resolved emission spectrum with emission maxima at 402 nm and 428 nm (Figure 3.27b), with λ_{ex} at 365 nm showing maximum emission intensity. β-*gdm*-(6-Cl-2-aminopurine)-monomer (**12**) which is non-fluorescent, showed florescence emission band at 402 nm and 428 nm upon fatty acid attachment with bright blue fluorescence in the visible region of electromagnetic spectrum. The fluorescence intensity of α-*gdm*-(6-Cl-2-amino)-

monomer (**6**) and N-decanoyl β-*gdm*-(6-Cl-2-aminopurine)-PNA monomer (**16**) is enhanced upon fatty acid conjugation. Unlike α-*gemdimethyl-aeg*-G-PNA (**8**) and β-*gemdimethyl-aeg*-G PNA (**14**) monomers in which the rotamers E and Z exhibit different patterns of fluorescence emission spectra, the corresponding fatty acid conjugates N-decanoyl β/α-*gdm*-(6-Cl-2-aminopurine)-PNA monomer (**16** and **19**) show similar pattern of emission spectra suggesting that fluorescence emission in fatty acid conjugates is mostly driven by the aggregation of fatty acid component.

Aggregation could increase the intensity of fluorescence emission (aggregation induced emission, AIE) or may quench the fluorescence emission (aggregation caused quenching, ACQ). Various established protocols are reported in literature to determine the AIE and ACQ.[55] One of the approaches involves monitoring fluorescence emission in DMSO (at constant concentration of fluorophore) with slow incremental (10%) addition of water from 100% DMSO to 100% water, keeping the total volume same.[55] AIE materials show a sudden increment of fluorescence intensity after a definite percentage of water/DMSO mixture. ACQ materials exhibit a quenching of fluorescence in all mixtures. The measurement of fluorescence

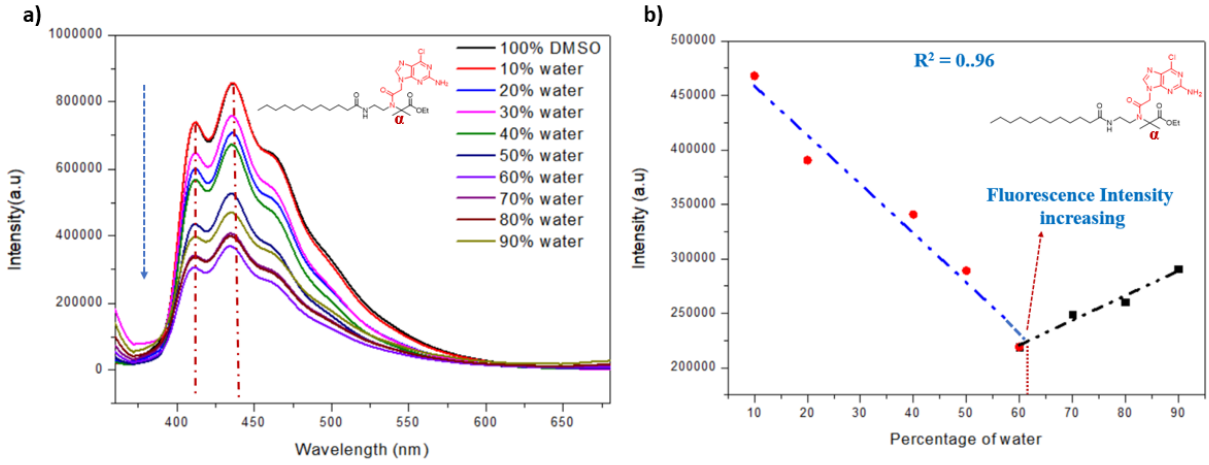


Figure 3.28. (a) Change of fluorescence intensity of N-decanoyl α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) in various water/DMSO mixture (concn. 230 μM) (b) Graphical representation of aggregation caused quenching (ACQ) N-decanoyl α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) (concn. 230 μM), Solvent Ethanol

emission of N-decanoyl α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) in the various water/DMSO mixture shows initial quenching of fluorescence, upto 60% water:DMSO, followed by a small increase of fluorescence intensity thereafter upto 100% water (Figure 3.28) These results suggest that the aggregation causes the quenching of

fluorescence intensity for N-decanoyl α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**).

Figure 3.29 shows the variation of fluorescence intensity of N-decanoyl-β-*gemdimethyl -aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) in various water/DMSO mixture. The intensity of fluorescence decreased with increased amount of water upto 60% water/DMSO and thereafter a small increase of intensity was observed. Thus, even the N-decanoyl β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) monomer also shows aggregation induced quenching (ACQ) in solution (Figure 3.29).

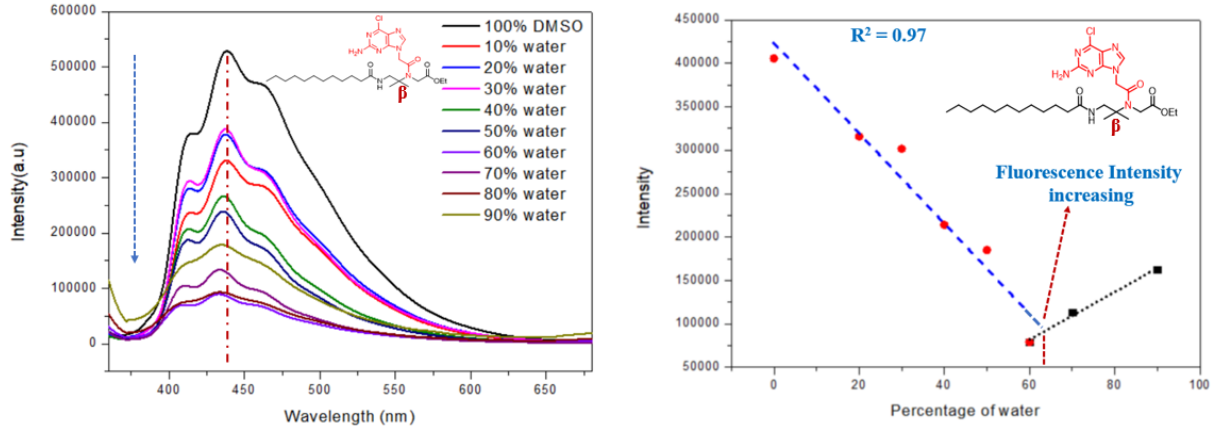


Figure 3.29. a) Change of fluorescence intensity of N-decanoyl β, β-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) in various Water/DMSO mixture (Concn, 220 μM) (b) Graphical representation of Aggregation caused Quenching (ACQ) of N-decanoyl β, β-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**).

3.3.10. Langmuir-Blodgett monolayer formation by N-decanoyl conjugated α/β-*gdm*-PNA monomers (collaboration with *Prof. B.L.V. Prasad, CeNS Bangalore*)

Amphiphiles having sufficiently polar group and a non-polar tail are insoluble in water form Langmuir monolayer at the air-water interface.[56] Solutions of such amphiphiles prepared in a volatile solvent (e.g., chloroform) is placed on liquid water and evaporation of the volatile solvent results in an ordered arrangement of amphiphile at the solvent/water interface to form the Langmuir monolayer.[56] Compression of monolayer increases the free energy of the system and at a critical pressure, the monolayer collapses. The fatty acid conjugated N-dodecanoyl-β/α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomers (**16** and **19**) are amphiphilic in nature with N-terminus having a long chain fatty acid and the C-terminus having polar C=O and C-N groups. The fatty acid conjugated N-decanoyl-β/α-*gdm* PNA monomers **16** and **19**

were dissolved in chloroform and the surface pressure (Π)-area per molecule (A_m) isotherm of **16** and **19** are shown in Figure 3.30.

The monolayers of N-decanoyl α/β-*gdm* PNA monomers **16** and **19** were compressed at a rate of 10 mm/min. The N-decanoyl α-*gdm*-PNA monomer (**19**) exhibited higher area per molecule (Limiting A_m ~75 $Å^2$ /molecule) compared to the N-decanoyl β-*gdm*-PNA monomer (**16**) (Limiting A_m ~60 $Å^2$/molecule). The collapse pressure of ~18 mN/m for N-decanoyl β-*gdm*-PNA monomer (**16**)) is marginally higher ~15 mN/m seen for N-decanoyl α-*gdm*-PNA monomer (**19**) suggesting better stability of monolayer formed by monomer **16** (Figure 3. 31). Typical values of surface compressional modulus (C_s^-) in the range 0−25 mN/m suggest low density liquid phase, 25−50 mN/m indicate liquid-expanded phase and above this in the range 100−250 mN/m belong to the liquid-condensed phase. Calculation of surface compressional modulus (C_s^-) for N-dodecanoyl β-*gdm*-PNA monomer (**16**) and N-dodecanoyl α-*gdm*-PNA monomer (**19**) point to the presence of liquid expanded phase (Figure 3.31a). To access the mechanical stability of the film, compression-expansion cycles were carried out. The films were compressed to its collapse pressure (~18 mN/m for **16** and ~15 mN/m for **19**) and decomposed to 0 mN/m to study the mechanical stability of the films formed by N-dodecanoyl β-*gdm*-PNA monomer (**16**) and N-dodecanoyl α-*gdm*-PNA monomer (**19**).

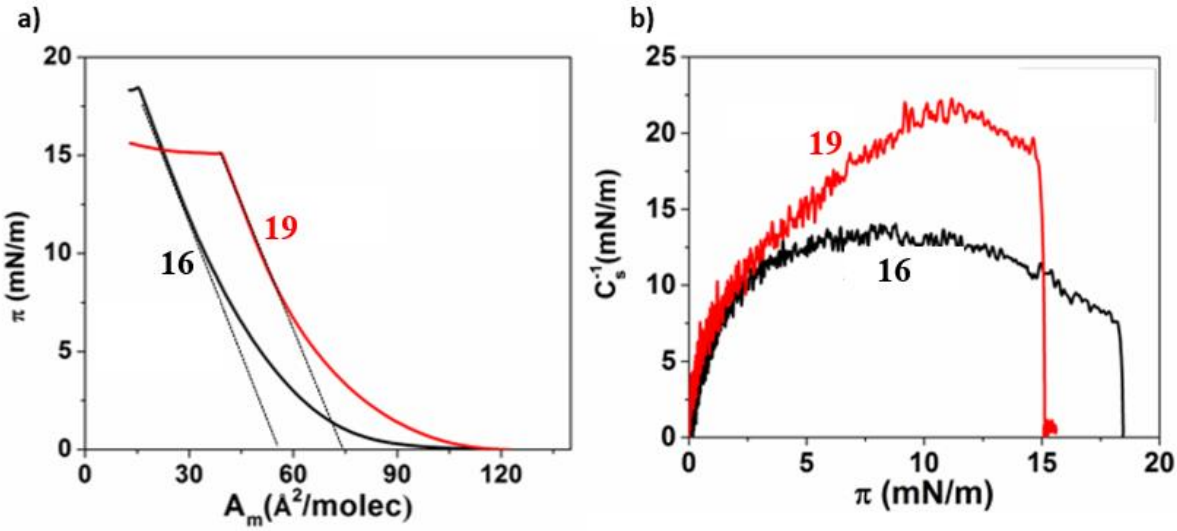


Figure 3.30. (a) Surface pressure (π) – area per molecule (A_m) (298 K) isotherm for N-decanoyl-β /α-*gdm*-(6-Cl-2-aminopurine)-*aeg* PNA monomers (**16** and **19**) (conc, 0.5 mg/ml) (b) Magnitude of compressional modulus (C_s^-) of N-decanoyl-β/α-*gdm*-(6-Cl-2-aminopurine)-*aeg* PNA monomers (**16** and **19**)

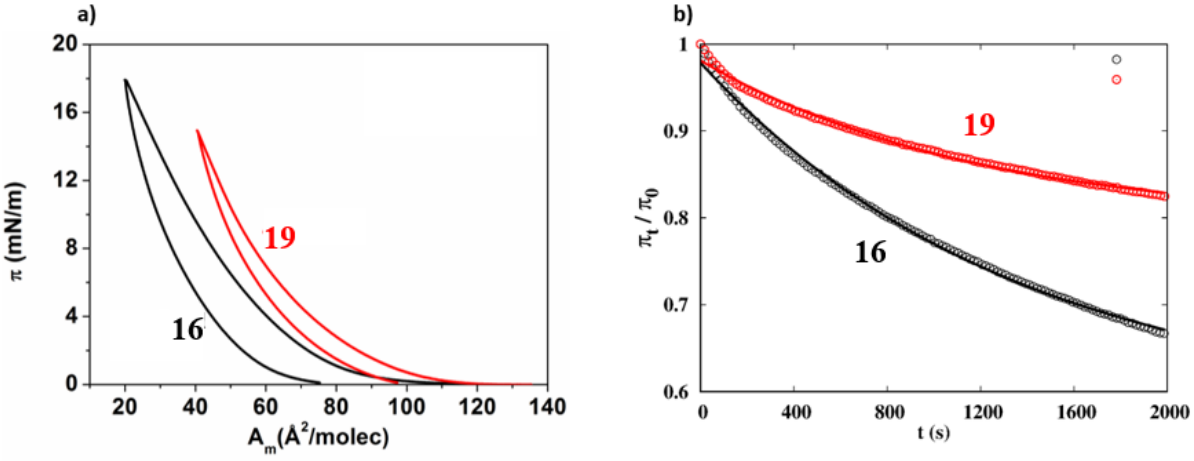


Figure 3.31. a) Compression-expansion cycles of the film of N-decanoyl β-*gdm*-PNA monomer (**16**) and N-decanoyl α-*gdm*-PNA monomer (**19**) monomer b) Relaxation profile of of PNA monomers (**16**) and (**19**) showing faster relax for **19**.

N-decanoyl-α-*dmg*-PNA monomer (**19**) exhibited less hysteresis compared to N-decanoyl-β-*dmg*-PNA monomer (**16**) suggesting better stability of the film formed **16** (Figure 3.31a). Relaxation studies were performed by keeping the area constant to understand the properties of the film. The films of **19** and 16 were compressed to a target surface pressure (8 mN/m) and the compression was stopped after that. The results obtained were fitted with the following equation to yield the equation for relax mechanism of the film. The results suggest a faster relaxation for **19** (1185. 8 second) compared to the **16** (1358.2 second, Figure 3.31b).

Fitted to an expression

$$\mathbf{f(x) = a + b*\exp(-t/\tau)}$$

$$\mathbf{\tau = 1358.2 +/- 7.6\ s\ for\ 16}$$

$$\mathbf{\tau = 1185.8 +/- 10.6\ s\ for\ 19}$$

In Summary, N-decanoyl β, β-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) forms a more stable monolayer at the air-water interface compared to the N-decanoyl α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**). Calculation of surface compressional modulus (C_s^-) of **16** and **19** suggested presences of liquid-expanded phase for the fatty acid conjugated N-decanoyl- α/β-*gdm-aeg*-(6-Cl-2-aminopurine) PNA monomers (**16** and **19**). Less

hysteresis, faster relaxation, more stable monolayer formation for N-decanoyl α-*gdm*-PNA monomer (**19**) compared to N-decanoyl β-*gdm*-PNA monomer (**16**) indicate higher preferences for monolayer by E rotamer (**16**, 100% E) compared to Z rotamer (**19**, 100% Z). This may be a consequence of differential effects on the orientation of polar head groups and non-polar tails dictated by the E/Z-rotamer preferences of the two monomers in solution.

3.3.11. Fluorescence lifetime measurement of monomers and Red Edge Excitation Shift (REES) effect[48]

In general, fluorescence emission from fluorophores is independent of the applied excitation wavelength.[48] However, abundant literature reports exist on the excitation dependent emission fluorescence and phosphorescence of polar fluorophores in polar viscous solvents.[48] Increasing excitation wavelength shifts the emission maxima towards the red edge of the UV-visible spectrum and this shift is termed red edge excitation shift (REES).[48] REES effect could be explained by a two-state model. Exciting the fluorophore at the center of UV-absorption band (λc) results in initial excited state F (Figure 3.32). The solvent molecules initially around F state coming from ground state, then reorient for effective interaction with the excited F state with a rate constant of *k* to a relaxed (R) state, that has lower energy difference between the ground and excited states (Figure 3.32). The corresponding excitation wavelength is λ_R and the rate of relaxation (*k)* depends on the nature and extent of interaction of relaxed fluorophore with the reoriented nearby solvent molecules in response to the newly formed excited state dipoles. The solvent relaxation is a highly complex phenomena and the emission wavelength (I_F) observed for F state being shorter compared to the emission wavelength (I_R) of R state (Figure 3.32).

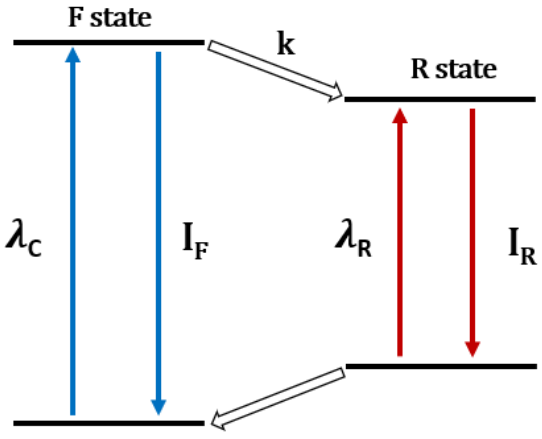


Figure 3.32. Two state diagrams proposed for the explanation of REES effect[48]

At lower temperature, the rate of solvent relaxation (*k*) is significantly slower than the rate of fluorescence decay (Γ, Γ>>*k*) and leads to a blue shifted emission (Figure 3.33a). Alignment of solvent dipoles in the reoriented solvent relaxed (R) state results in stabilization of excited state hence lower in energy than F state (Figure 3.33). But the destabilization of the ground state results in higher energy of the ground state in R state. Excitation in the red-edge at low temperature preferably selects the molecules which has solvent relaxed environment around the ground and excited state, consequently results in red-shifted emission spectra (Figure 3.33a).[48]

At higher temperature the rate of solvent relaxation (*k*) is significantly faster than the rate of fluorescence decay (Γ, Γ<<*k*) and excitation at the wavelength (λc) of central UV-absorption results in emission from the relaxed (R) state as higher solvent relaxation rate (*k)* facilitates the solvent relaxation prior to emission. Hence red shifted emission results for excitation at the wavelength of central UV-absorption (λc, Figure 3.33b). For excitation at red-edge at higher temperature an equilibrium is established between the fluorescence emission and solvent relaxation of R state prior to emission. Hence a red shifted emission results for excitation at red-edge at higher temperature.[48]

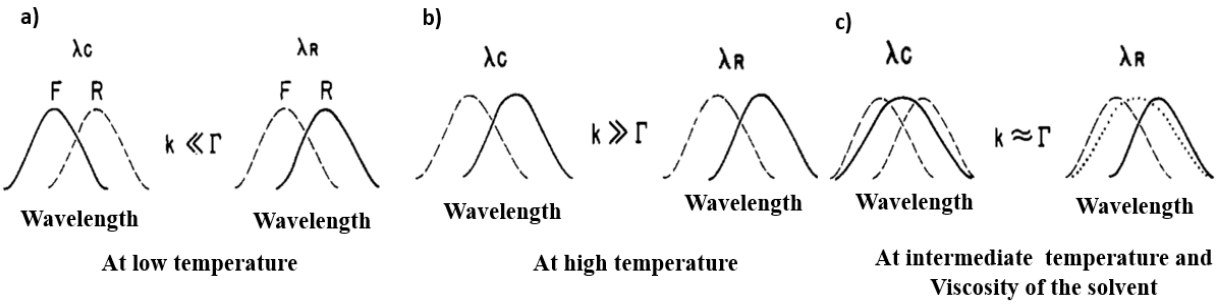


Figure 3.33. (a) REES effect[48] at *low temperature* (b) REES effect at *higher temperature* (c) REES effect *at intermediate temperature* and viscosity of the solvent. The *solid lines* represent the *observed spectra*, and the *dashed lines* represent the *emission spectra of either* the F or the R state.

In the intermediate range of temperature and viscosity of the solvent the rate of fluorescence decay (Γ) is comparable to the rate of solvent relaxation (*k, k ≃ Γ).* For the excitation at the wavelength (λc) of central UV-absorption band the contribution from both solvent relaxation and fluorescence decay results in an intermediate spectrum of I_F and I_R. For the excitation at red-edge wavelength (λ_R), the solvent relaxed states are selectively excited

resulting in a red shifted spectrum. Compared to the excitation at the wavelength (λc) of central UV-absorption where the emission results majorly from F state and relaxed state, the emission band of the excitation at red-edge results majorly from relaxed state. Hence a narrow emission spectrum is expected for excitation in the red edge wavelength (Figure 3.33c).[48] The fluorescence lifetime of fluorophores in different solvents thus depend on the viscosity since stabilization of "R" state depends on the viscosity of the solvent.

To examine the role of viscosity in the excitation dependent emission spectra of fluorophores, fluorescence emission of β-*gdml-aeg*-G-PNA monomer (**14**) was recorded in ethanol and 40% methanol/water. The excitation dependent emission spectra recorded in ethanol exhibited 370 nm emission band at lower λ_{ex} region (307 nm-327 nm) but at higher λ_{ex} (337 nm-407 nm), shifted to λ_{em} band at 475 nm (Figure 3.34a). Similar spectra recorded in 40% water/methanol shows λ_{em} at 365 nm at lower excitation wavelength (306 nm-326 nm), but at higher λ_{ex}, the emission maxima varied from 462 nm - 475 nm (Figure 3.34b). This distinct pattern of emission in different solvents indicated the important role of solvent viscosity in REES effect.

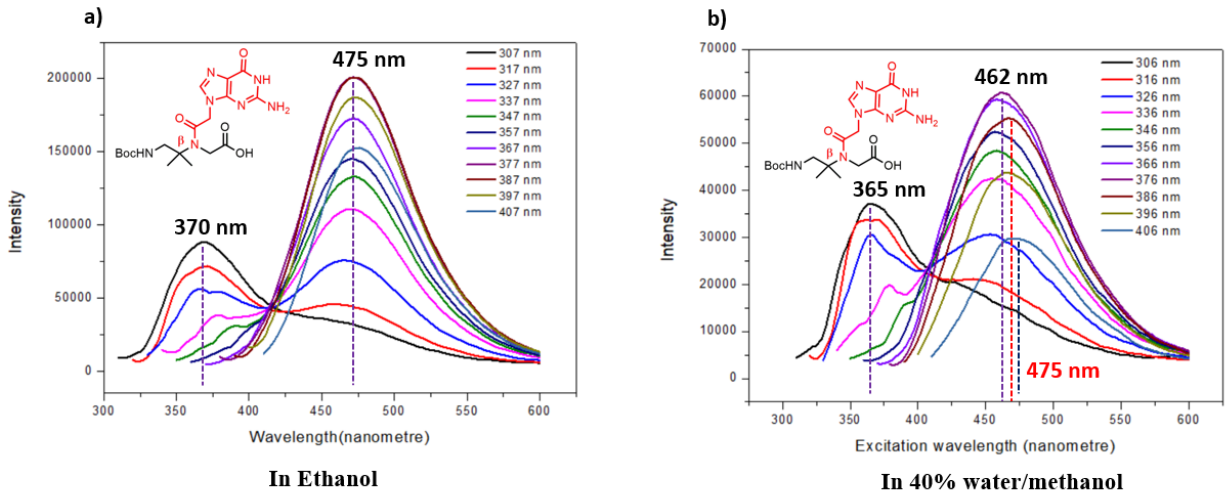


Figure 3.34. Excitation dependent emission spectra of β-*gemdimethyl-aeg*-G-PNA monomer (**14**) in (a) *ethanol* and (b) *40% water/methanol* (Concn, 210 μM).

3.3.11a *Fluorescence lifetime*: The fluorescence lifetime is the time spent by a fluorophore in the excited state before decaying through the emission of photons prior to return in the ground state.[48] The fluorescence lifetime is short in the range of nanoseconds to picoseconds but may go upto 10^{-8} seconds. The fluorescence lifetime depends on the local environment around the fluorophore. The fluorescence lifetimes of β-*gdm-aeg*-G PNA (**14**) monomer in ethanol and

40% water/methanol were performed using time correlated single photon counting at λ_{ex} of 370 nm. The average fluorescence lifetime is calculated using the standard equation reported in the literature.[57] β-*gdm-aeg*-G PNA monomer (**14**) exhibited a fluorescence lifetime of *3.379* nanoseconds (Figure 3.35a) in ethanol and a slightly shorter 3.12 nanoseconds in 40% water/methanol (Figure 3.35b). This result along with the trends of excitation dependent emission spectra in the ethanol and 40% water/methanol confirms the red Edge excitation shift (REES) effect in the fluorescence emission of β-*gdm-aeg*-G-PNA monomer (**14**). Figure 3.35 shows the solvent dependent fluorescence lifetime measurement of monomer **14**.

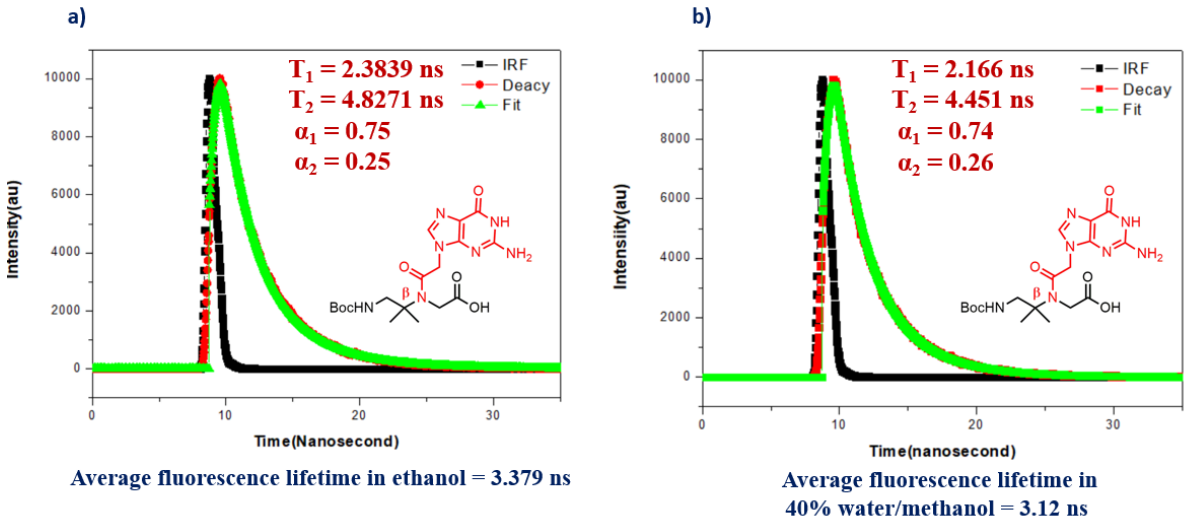


Figure 3.35. a) Fluorescence lifetime of β-*gdm*-G-PNA monomer (**14**) in (a) ethanol (*3.379 ns*) and (b) 40% water/methanol (*3.12 ns*).

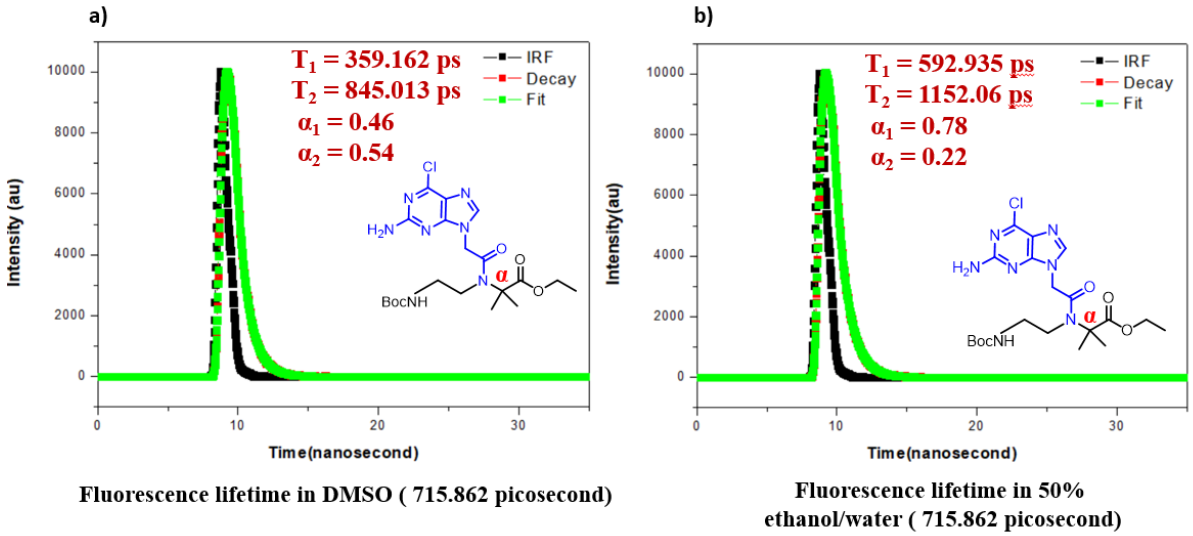


Figure 3.36. Fluorescence lifetime of α- *gdm-(6-Cl-2-aminopurine)* PNA monomer (**6**) in (a) DMSO (*715.96 picosecond*) and (b) 50% ethanol/water *(769.21 picosecond).*

The fluorescence lifetime measurement of monomer α-*gdm*-(6-Cl-2-aminopurine) PNA (**6**) in solvents with varying viscosity exhibited lifetime in picosecond range (Figure 3.36). It exhibited a fluorescence lifetime of *752.61* picosecond in ethanol, 715.92 picosecond in DMSO, 769.21 picosecond in 50% ethanol/DCM and 790.92 picosecond in 50% ethanol/water. The lifetime decreased in solvent mixtures in the order 50% water/ethanol > 50% ethanol/DCM > Ethanol > DMSO. Thus, the viscosity of the solvent around the fluorophore influences the excitation dependent emission spectrum pattern of the fluorophores and the fluorescence lifetime, thus explaining the REES effect observed in α- *gdm* PNA monomer (**6**).

3.3.12 Fluorescence emission of PNA dimers

The UV-visible spectra of the (β-*gdm*-*aeg*-G)$_2$-PNA dimer (**22**) and (*aeg*-G)$_2$-PNA dimer (**23**) were recorded in ethanol (Figure 3.37). The (β-*gdm*-*aeg*-G)$_2$ PNA dimer (**22**) exhibited UV-visible band at 270 nm and 247 nm (Figure 3.37 a) and (*aeg*-G)$_2$ PNA dimer (**23**) shows UV-Visible band at 270 nm and 252 nm (Figure 3.37b). The strong 308 nm band observed β-*gdm*-*aeg*-G PNA (**14**) monomer (Figure 3.37) was absent in both the dimers. The appearance of 308 nm band may arise by preference for a single rotamer in solution as in β-*gdm*-G-*aeg* PNA (**14**, E-rotamer), while *aeg*-PNA-G (**20**) with mixture of major Z and minor E rotamers do not show band at 308 nm. It is possible that both (β-*gdm*-*aeg*-G)$_2$ PNA dimer (**22**) and (*aeg*-G)$_2$ PNA (**23**) dimers have rotamer mixtures in solution and hence lack the 308 nm UV-absorption band.

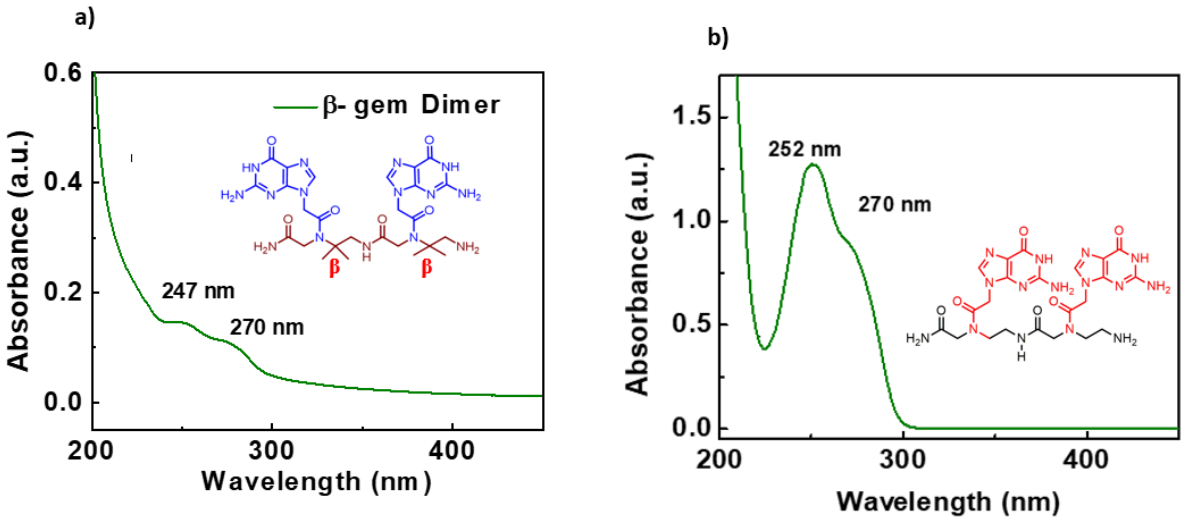


Figure 3.37. UV-Visible spectra of (a) (β-*gdm*-*aeg*-G)$_2$ PNA dimer (**22**) (Concn, 50 μM) and (b) (*aeg*-G)$_2$ PNA dimer (**23**) (Concn, 200 μM); Solvent, ethanol

The fluorescence emission spectra of (β-*gdm*-*aeg*-G)$_2$ PNA dimer (**22**) were recorded in ethanol and in phosphate buffers at pH 7.65, 9.98 and 12.75 (Figure 3.38). In phosphate buffer at pH 7.65, the 325 nm emission band observed upon excitation at λ_{ex} of 250 nm shifted to lower wavelength when λ_{ex} was increased from 260 nm to 290 nm and no significant emission band was observed beyond 300 nm (Figure 3.38a). The emission spectra of (β-*gdm*-*aeg*-G)$_2$ PNA dimer (**22**) in ethanol exhibited weak emission at 450 nm (λ_{ex} 270 nm) and

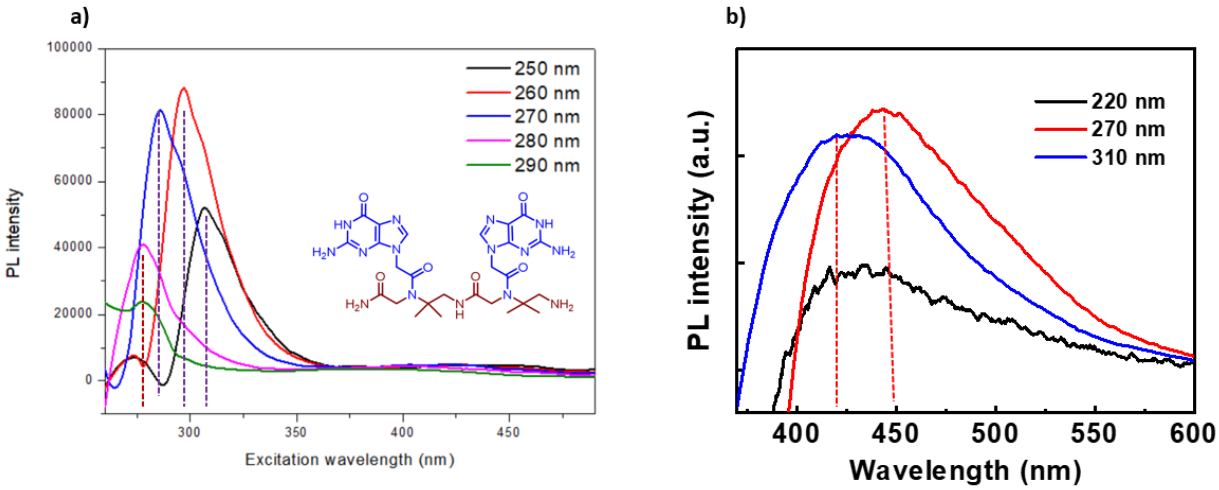


Figure 3.38. Emission spectra of (β-*gdm*-G)$_2$ PNA dimer (**22**) (a) phosphate buffer, 10 mM, pH 7.65 (b) in ethanol (Concn, 100 µM)

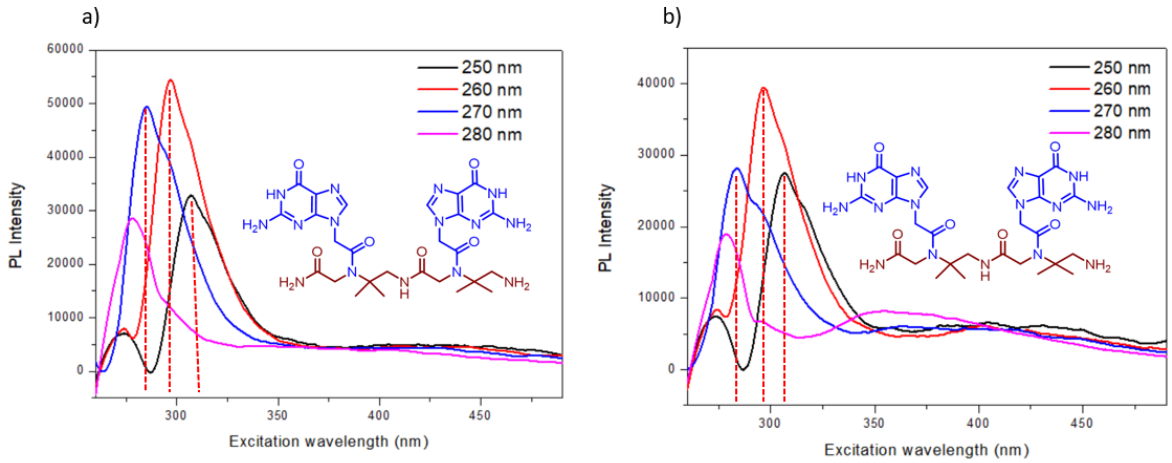


Figure 3.39. Emission spectra of (β-*gdm*-*aeg*-G)$_2$ PNA dimer (**22**) (a) phosphate buffer (10 mM), pH 9.98, (b) phosphate buffer (10 mM), pH 12.75; Concn,100 µM

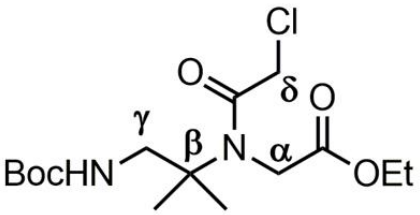
increasing λ_{ex} shifts the λ_{em} to lower wavelengths 420 nm (Figure 3.38b). The weak emission of dimers compared to the β-*gdm*-*aeg*-G PNA monomer may be due to difference in rotamer preferences and hence conformations in the monomer and dimer. (β-*gdm*-*aeg*-G)$_2$ PNA dimer

(**22**) and (*aeg*-G)$_2$ PNA dimer (**23**) may have a mixture of rotamers in solution for two t-amides in same molecule. Stabilization by complementary base pairing (Watson Crick pairing) and other structural motif (G-quadruplex) formation is ruled out for GG dimers and hence weak emission properties may result in ethanol and in phosphate buffer. In comparison, the emission spectra at alkaline pH were more pronounced at higher pHs of 9.98 and 12.75, (Figure 3.39) but again exhibited shift of emission maxima to lower wavelengths upon increasing excitation to longer wavelengths (blue shift), the reasons being not so clear at the moment from this work.

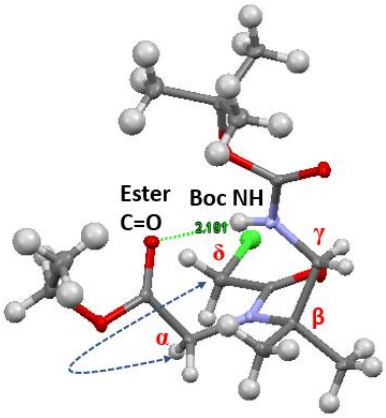
3.3.13. Crystal structure of the N-acetylchloro-β-*gdm*-aeg methyl ester (11)

N-acetylchloro-β-*gdm*-aeg methyl ester (**11**), a common intermediate in the synthesis of target PNA monomers crystallized from MeOH/DCM (50%). The crude powder of **11** from the was dissolved in MeOH and kept at 4^{o}C in glass vials which were sealed by para film with a few small holes and placed into a ventilated fridge to evaporate the solvents slowly. Colorless needle type crystals were obtained after 2 ~ 3 days, which were used for collecting single crystal structure data.

In the crystal structure, the orientation of the side chain δ-methylene group of the PNA backbone. was towards the ethylene diamine side of the PNA backbone suggesting a E-rotamer in the solid state. The structure showed eight membered intramolecular hydrogen bonding between N-terminal BocNH and C-Terminal ester carbonyl group of PNA backbone with the distance between the ester carbonyl and N-terminal BocNH of 2.191 Å, suggesting the γ-turn like folding in the β-*gdm*-*aeg* PNA backbone.[61] This result supports the NOSEY observation and confirmed the E-rotamer stabilization of β-*gdm*-*aeg*-PNA-G (**14**) monomer in solution. The structure of **11**, bears striking resemblance to the N-acetylchloro-a-*aza*-*aeg* (**7**, chapter 2). The N-acetylchloro intermediate of β-*gdm*-aeg-PNA-G (**11**) exhibits intramolecular hydrogen bonding between N-Terminal BocNH and ester carbonyl group suggesting major role of *gemdimethyl* group for the stabilization of γ-turn like orientation.[61] On the other hand the N-acetyl chloro intermediate of α-*aza*-PNA-T (**7**, chapter 2) exhibit intramolecular hydrogen bonding between the backbone α-NH and Boc C=O still γ-turn like orientation.[61] It would be interesting to compare the crystal structures of the N-acetylchloro intermediates with crystal strcutures of β-*gdm*-*aeg*-PNA-G (**11**) and α-*aza*-PNA-T (**7**, chapter 2). Unfortunately, attempts to obtain good crystals of these compounds did not succeed. The table 3.3 shows crystallographic information of the N-acetyl chloro intermediate of β-*gemdimethyl*-*aeg*-PNA-G (**11**).

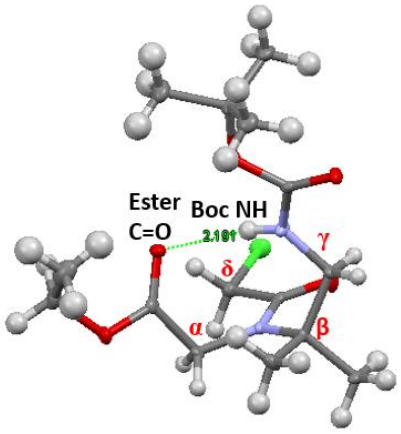


Molecular structure of compound 11
(E-Rotamer)

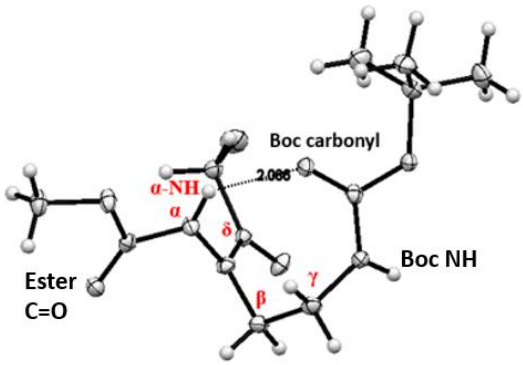


Crystal structure of compound 11
(E-Rotamer)

Figure 3.39. Molecular structure and crystal structure of **compound 11 (Chapter 3)**



Crystal structure of compound 11
(E-Rotamer)



Crystal structure of Compound 7

(*α- aza-* Chloro compound, chapter 2, compound 7)

Figure 3.40. Comparison of crystal structure of the N-acetyl chloro intermediate of β-*gemdimethyl-aeg*-PNA-G (**11**) and α-*aza*-PNA-T (**7**, Chapter 2)

Crystal data and structure refinement of compound 11

Identification code	try a
Empirical formula	C_{15} H_{27} N_2 O_5 Cl
Formula weight	350.1608
Crystal system, space group	Monoclinic, P21/c
Refines used	8575
Measurement temperature	296 K

Unit cell dimension	a = 12.14 angstrom, α = 90 degree b = 9.14 angstrom, β = 98 degree c =15.13 angstrom, γ = 90 degree
Cell measurement theta max	65.46 degree
Cell measurement theta min	3.57 degree
Crystal description	Block crystal
Crystal color	Colorless
Crystal density	1.387
Crystal size	0.13
Absorption coefficient	2.092
Absorpt correction type	Multi scans
Absorpt process details	SADBAS
Radiation wavelength and type	0.7107 angstrom and MoK\a
Diffraction source	Fine focused seal tube
Diffraction measurement device	Bruker APEX-II CCD
Diffraction method	Omega and Phi Scan
Diffraction reflection number	17241
Diffraction reflection h min	-14
Diffraction reflection h max	14
Diffraction reflection l min	-11
Diffraction reflection l max	11
Diffraction reflection k min	-17
Diffraction reflection k min	17
Diffraction reflection theta min	3.57
Diffraction reflection theta min	65.27
Total reflection number	3088
Computing data collection	Bruker Apex-II
Computing cell refinement	Bruker SAINT
Computing data reduction	Bruker SAINT

Computing structure solution	SHELXT 2014/5 (sheldrick, 2014)
Computing data refinement	SHELX-L 2018/3 (sheldrick, 2018)
Computing molecular graphics	Bruker SHELXTL
Computing publication material	Bruker SHELXTL
Refine ls extinction method	SHELX-L 2018/3 (sheldrick, 2018)
Refine ls extinction coefficient	0.000 (5)
Refine ls R factor all	0.0803
Refine ls R factor gt	0.0645

Table 3.3. Single crystal data of compound **11**

3.4 Discussion

The role of t-amide group connecting the backbone of *aeg*-PNA with nucleobases and the generated rotamers, in defining the PNA secondary structure has not been investigated well. It was recently found that *gem* dimethyl substitution at Cα leads to exclusive Z-rotamer and at C_β gave only the E-rotamer.[51] Such selectivity arises from a mix of evasion of steric clash of *gemdimethyl* group with the δ-CH_2 and favourable C-H•••O interactions.[51] In the present work, another important property related to the rotamers around the t-amide group has been unravelled. It was found that the α/β-*gdm-aeg* PNA-G monomers (**8** and **14**) inspite of lacking fluorophores are found to be fluorescent. Interestingly, it is related to the type of rotamer: the α-*gdml-aeg* PNA-G monomer (**8**) with exclusive Z-rotamer gives a bluish fluorescence, while the β-*gdm-aeg* PNA-G monomer (**14**) that shows exclusive E-rotamer exhibits green fluorescence. To our knowledge, these are perhaps the first examples in nucleic acid chemistry reporting fluorescence emission in DNA analogues without any base modifications. The *gem*-dimethyl substiututed PNA monomers show a band at 308 nm in the UV spectra, that is absent in the unsubstituted *aeg*-PNA-G (**20**). Further, the α/β-*gdm-aeg*-PNA monomers (**8** and **14**) exhibit excitation dependent emission spectra and red edge excitation shift (REES).[48] The β-*gdm-aeg*-PNA-G monomer (**14**) that exclusively shows the E-rotameric form (100% E rotamer) exhibits green fluorescence emission with broad vibrationally non resolved emission band centered at 475 nm. On the other hand, α-*gdm-aeg*-PNA-G monomer (**8**) forming only the Z-rotamer exhibits blue fluorescence with sharp vibrationally resolved emission bands

centered at 402 and 438 nm, characteristic of excimers. The origin of fluorescence of these non-conjugated molecules could be explained on the basis of recently proposed phenomena of *clusteroluminogens*[23] through space conjugation (TSC) and lower band gap between the ground and excited states.[46] Dipole-dipole and ion-dipole interactions, intra residue hydrogen bonding, interaction between σ, n and π electrons of C-N, C-O, bonds in a network like structure of the aggregates leads to short band gap between the ground and excited states. The N-terminal Boc-deprotected monomers show weak fluorescence emission suggesting the importance of amide terminal NH involved in H-bonding with ester carbonyl to stabilize a folded geometry necessary for efficient TSC. Solvent dependent experiments ruled out the possibility of aggregation induced emission in β/α-*gdm*-aeg-PNA-G monomers (**14** and **8**). Solvent dependent fluorescence emission of the β/α-*gdm*-*aeg*-PNA G (**14** and **8**) and β/α-*gdm*-*aeg*-(6-Cl-2-amino purine) PNA monomers (**16** and **19**) show characteristic shifts of wavelengths of emission maxima and intensity changes depending on the polarity of the solvent. The observance of REES effect suggested role of solvation around the excited fluorophore. α-*gdm*-aeg-PNA-C(N^4-Cbz) monomer (**5**) also exhibits similar UV-absorption band as *aeg*-C(N^4-Cbz)-PNA monomer (**21**), but shows bright blue fluorescence with sharp vibrationally resolved emission band centered at 402 and 438 nm unlike the latter derivative. Thus, it is imperative that the *gemdimethyl* substitution on the backbone pre-organizes the single sigma bond orientations for favorable σ-n-π electrons interactions leading to short band gap and fluorescence emission. Fatty acid conjugation *of gdm* PNA monomers irrespective of the position of *gemdimethyl* substitution (α and β) exhibit bright blue emission with similar pattern of vibrationally resolved emission band centered at 402 nm and 428 nm. N-decanoyl α/β-*gdm*-(6-Cl-2-aminopurine)-PNA monomers (**16** and **19**) having somewhat amphiphilic structure form Langmuir monolayer at the air/water interface. The α-PNA monomer (**19**) exhibits higher area per molecule compared to the N-decanoyl β-PNA monomer (**16**)**.** The collapse pressure of N-decanoyl β-*gdm*-(6-Cl-2-aminopurine)-PNA monomers (**16**) is marginally higher than the N-decanoyl α-*gdm*-(6-Cl-2-aminopurine)-PNA monomers (**19**) suggesting better stability of monolayer formed by N-decanoyl β-*gdm*-(6-Cl-2-aminopurine)-PNA monomers (**16**). The hysteresis experiments indicated the N-decanoyl α-*gdm*-(6-Cl-2-aminopurine)-PNA monomers (**19**) to have less hysteresis compared to the N-decanoyl β-*gdm*-(6-Cl-2-aminopurine)-PNA monomers (**16**). Solvent dependent fluorescence lifetime measurement indicated the lifetime dependence on viscosity of solvent strongly supporting solvation mediated REES effect. In contrast to the monomers, the **GG** PNA dimers exhibit very weak emission in both organic solvent and phosphate buffer and suggesting non

correlation of rotameric effects in the two PNA units due to the rotameric mixture eventually affecting the base stacking and backbone geometry unfavorable for TSC interactions necessary for fluorescence.

Surprisingly, the α/β-*gdm*-*aeg*-PNA-G monomers (**8** and **14**) exhibited fluorescence in solid powder form as well (Figure 3.41) retaining the emission shown in solution. The solid-state fluorescence may have practical applications in device fabrications. The (β-*gdm*-aeg-G)$_2$ PNA dimer (**22**) did not show any significant fluorescence in solution or in solid state. Our results on fluorescence of monomers assume importance in light of work by Gazit, who showed that only dimeric PNA exhibited fluorescence properties. Further work using theoretical calculations is necessary to fully understand the serendipity observations of fluorescence in α/β-*gdm*-*aeg* PNA monomers (**8** and **14**) reported in this chapter.

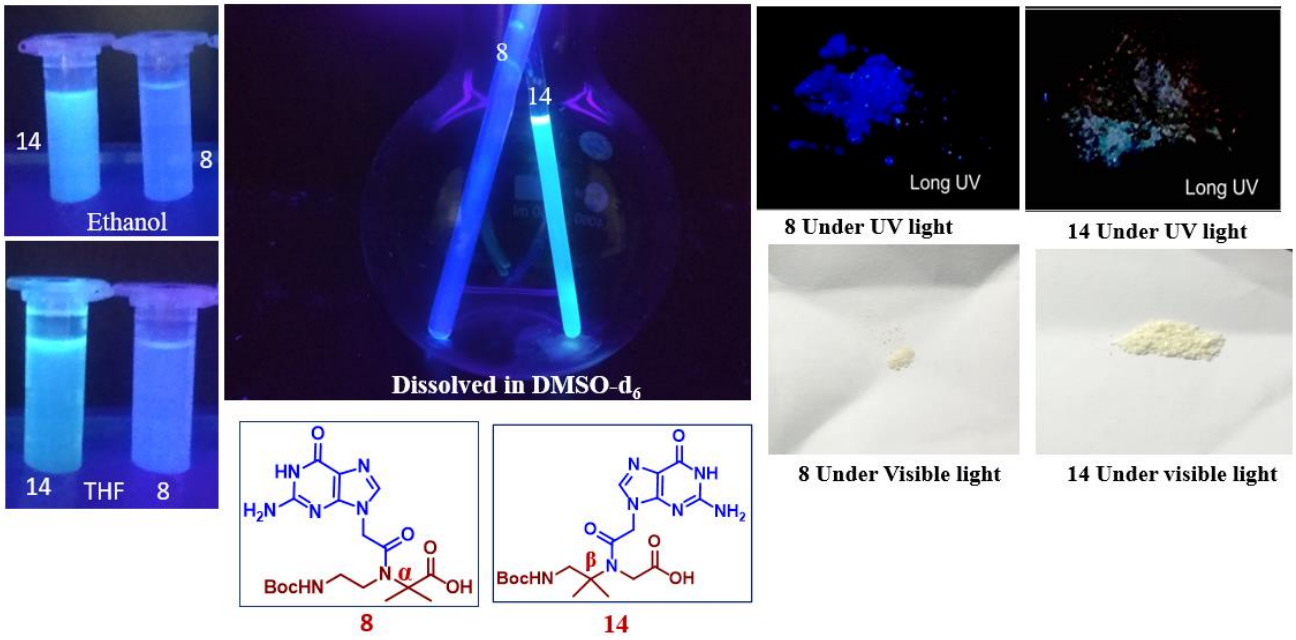


Figure 3.41. Solution and solid-state fluorescence of β-*gdm*-*aeg*-PNA-G monomer (**14**) and α-*gdm*-*aeg*-PNA-G monomer (**8**)

3.5. Summary

This chapter reports on unexpected fluorescence properties observed in α/β-*gdm*-*aeg*-PNA G monomers (**8** and **14**) which arise from the rotamer induced backbone structure that has favorable single bond TSC that leads to the observed phenomena.

(i) Synthesis and characterization the of β/α-*gdm*-*aeg*-PNA monomers (**8** and **14**) having the bases G, 6-Cl-2-aminopurine and C(N^4-Cbz) (**5**, **6**)

(ii) NMR studies by 1D and 2D COSY/NOESY techniques suggest exclusive E-rotamer stabilization for β-*gdm*-*aeg*-G PNA monomer (**14**) and only Z-rotamer stabilization for α-*gdm*-*aeg*- G- PNA monomer (**8**).

(iii) β/α-*gdm*-*aeg*-PNA G (purine) monomers (**12** and **6**) exhibit characteristic UV-absorption band at 308 nm and rotamer dependent fluorescence emission in solution and solid state compared to the non-fluorescent *aeg*-PNA-G monomer.

(iv) α-*gdm*-*aeg*-PNA-C(N^4-Cbz) (pyrimidine) monomer (**5**) lacks 308 nm band in UV spectra, but shows characteristic blue fluorescence emission with vibrationally resolved emission band at 402 nm and 428 nm.

(v) Fatty acid conjugation at N-terminus N1-(dodecanoyl)- β/α-*gdm*-*aeg* PNA-G and (6-Cl-2-aminopurine) monomers (**16** and **19**) exhibit characteristic UV-absorption band at 308 nm and shows blue fluorescence with vibrationally resolved emission band at 402 nm and 428 nm.

(vi) All the *gdm*-PNA monomers exhibit excitation dependent emission spectra (REES) in solution.

(vii) Results of aggregation induced emission study shows Aggregation Caused Quenching (ACQ) for all the monomers with increasing aggregation of the monomer.

(viii) Solvent dependent emission spectra and fluorescence lifetime measurements of monomers in various solvents strongly support the REES effect in the fluorescence emission of α/β-*gdm*-*aeg*-PNA Monomers.

(ix) Langmuir Blodgett study of fatty acids conjugated β/α-*gdm*-*aeg*-G /(6-Cl-2-aminopurine) PNA monomers (**16** and **19**) shows monolayer formation.

(x) (β-*gdm*-aeg-G)$_2$ PNA dimer (**22**) and (*aeg*-G)$_2$ PNA dimer (**23**) exhibit very weak fluorescence emission compared to the monomers.

(xi) The origin of fluorescence emission of N-Terminal Boc protected β/α-*gdm*-*aeg*-PNA G monomers (**14** and **8**) is due to TSC mechanism for the origin of fluorescence in these non-conjugated molecules.

3.6. Experimental procedures

This section describes the detailed synthesis of monomers and the spectral characterisation of the rationally designed α/β-*gdm-aeg* PNA monomers (**8** and **14**), dimers (**22** and **23**) and fatty acid conjugated monomers (**16** and **19**).

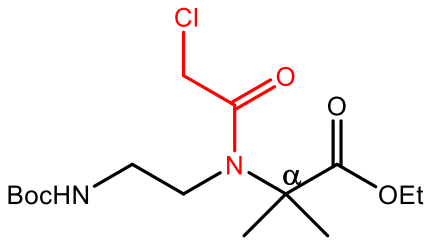
3.6.1 General information

Laboratory grade chemicals were used for setting up standard chemical reactions. Chemicals were purchased from Sigma Aldrich and Spectrochem. For moisture sensitive reactions the reagents were redistilled to ensure the maximum purity of the reagents. Solvents used for performing chemical reactions were dried using standard procedures available in the literature. Reactions were monitored using Thin Layer Chromatography (TLC). Standard UV-Lamp, Iodine spray and spraying with ninhydrin solution followed by heating the TLC plate were used to analyse the TLC. Standard procedure of work up involves extracting the desired organic compound by washing the aqueous layer with organic solvent (DCM, EtOAc) followed by washing of the combined organic layer with brine. The organic layer was dried over anhydrous sodium sulphate (Na_2SO_4) or calcium chloride ($CaCl_2$ for DCM) and evaporated under reduced pressure to afford the desired organic compound. Precoated silica gel GF_{254} (Merck 5554) TLC plates were used for TLC analysis. Standard chemical reactions were performed in room temperature, wherever required temperature was mentioned. Silica gel (60-120 or 100-120 mesh) were used to perform the column chromatography. Double distilled EtOAc, hexane, MeOH and DCM were used as a mobile phase to perform the column chromatography. ^{1}H NMR, ^{13}C NMR and 2D-NMR experiments were performed in Bruker-400 MHz, JEOL-400 MHz, Bruker-500 MHz, and Bruker-600 MHz NMR spectrometers. NMR solvents were purchased from Sigma Aldrich. Deuterated chloroform ($CDCl_3$) and DMSO-d_6 were used as an NMR solvent. Chemical shifts were reported in δ ppm with respect to the internal standard TMS or deuterated NMR solvent. Mass spectra of the reaction intermediates were recorded in an Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using TiO_2 or 2,5- dihydroxybencoic acid (DHB) as matrix. To further confirm the mass spectra of small molecules (Mass 100-1200) obtained in MALDI, High Resolution Mass Spectrometry (HRMS) was performed for each reaction intermediates. IR spectra of the samples were recorded in a Bruker IR spectrometer. The obtained crystal was mounted on top of a cryoloop and transferred into the cold nitrogen stream (100 K) of a Bruker-AXS D8 Venture diffractometer. Data collection and reduction was performed using the Bruker

software suite APEX.[58] The final unit cell was obtained from the xyz centroids of 9824 reflections after integration. A multiscan absorption correction was applied, based on the intensities of symmetry-related reflections measured at different angular settings (SADABS). The structures were solved by direct methods using SHELXT[59] and refinement of the structure was performed using SHLELXL.[60] The hydrogen atoms were generated by geometrical considerations. SHIMADJU UV-2600i instrument was used for the measurement of UV Data of the monomers. A pair of square quartz cuvette with maximum volume 3.5 ml, path length of 10 mm (1 cm), external dimensions of 12.5×12.5 mm and height of 45 mm (4.5 cm) was used for the measurement with a wavelength range 200 nm-800 nm. Before every measurement a blank experiment was performed with the solvent of interest and a standard baseline correction was performed with respect to the solvent of interest to ensure the accuracy of the obtained data. Emission spectra were taken on a Horiba Jobin Yvon Fluorolog-3 spectrofluorometer at various excitation wavelengths. Excitation wavelength range of 280 nm to 420 nm was used for the measurement of emission spectra of monomers at 25^{o}C, emission and excitation slits were set at 2.5 nm for every measurement. Measurements were performed in a 1 cm rectangular quartz fluorescence cuvette. All spectra were normalized so that the emission maxima and minima were identical. A blank emission spectrum of the solvent was taken before the experiment to ensure the accuracy of the obtained data. Time-resolved spectra were captured using time-correlated single photon counting with a pulsed LED light source at 370 nm. The emission slit was set at 10 nm. A stirred solution of 5 mg/ml of the monomers in various organic solvents was used for the measurements of lifetime of the fluorescence.

3.6.2 Experimental procedures and spectral data

Ethyl 2-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-chloroacetamido)-2-methylpropanoate (4)[51]

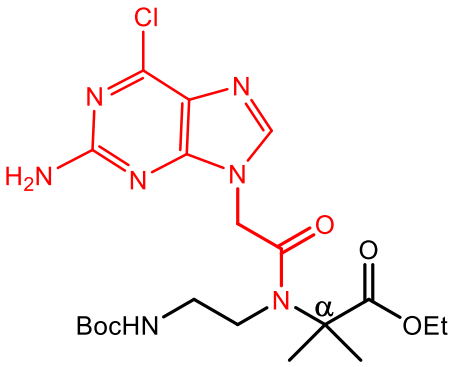


4

N-acetylchloro-α-*gdm-aeg* methyl ester (**4**) was synthesized from the known amine backbone **3** using the reported procedure.[51] The amine (**3**, 3.5 gm, 12.7 mmol) was dissolved

in dry DCM (45 ml) and triethylamine (NEt_3 ,7.1 ml, 50 mmol) was added in the mixture. The solution was cooled to 0^oC and chloroacetyl chloride (4.0 ml, 50 mmol) was added to the reaction mixture drop by drop (*Caution-Addition of chloroacetyl chloride produces heat, add the reagent slowly and drop by drop*). The reaction mixture was stirred at room temperature for 30 h. The reaction was quenched by the addition of aq. $NaHCO_3$ and the desired compound was extracted by DCM (3 x 150 ml) . The combined organic layer was dried over $MgSO_4$ and DCM was evaporated to give the desired compound as yellow oil. The crude oil was purified by flash chromatography using 35% EtOAc in hexane as eluent to afford **4** as yellow oil (3.5 gm, 78% yield), R_f = 0.65 (50% EtOAc in hexane). ^1H NMR (400 MHz, $CDCl_3$), δ 5.18 (t, *J* = 6.1 Hz, 1H), 4.15 (t, *J* = 4.8 Hz, 2H), 4.13 – 4.06 (m, 2H), 3.53 (t, *J* = 6.8 Hz, 2H), 3.31 (dd, *J* = 13.0, 6.5 Hz, 2H), 1.49 (s, 6H), 1.42 (s, 9H), 1.24 (t, *J* = 7.2 Hz, 3H), ^{13}C NMR (101 MHz, $CDCl_3$), δ 174.2, 167.6, 156.2, 80.2, 61.1, 61.5, 43.1, 41.1, 41.4, 28.5, 24.1, 14.1. The synthesized compound was used for next step of without further analysis.

Ethyl2-(2-(2-amino-6-chloro-9H-purin-9-yl)-N-(2-((tert-butoxycarbonyl) amino)ethyl)acetamido)-2-methylpropanoate (6)

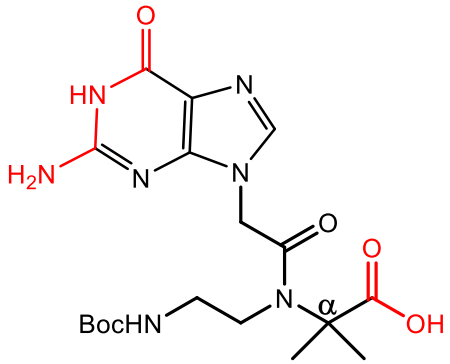


6

2-Amino-6-Chloropurine (264 mg, 1.5 mmol) was dissolved in dry DMF (6 ml) and anhydrous K_2CO_3 (215 mg, 1.5 mmol) and anhydrous Cs_2CO_3 (46 mg, 0.1 mmol) was added to the reaction mixture. The reaction mixture was stirred at room temperature for 5 min, and N-acetylchloro-α-*gdm-aeg* methyl ester (**4**, 500 mg, 1.4 mmol) was added to the mixture. Resulting mixture was stirred at room temperature for additional 72 h. The reaction mixture was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml) . The combined organic layer was washed with brine (3 x 150 ml) and the same was dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer **6** as a yellow hygroscopic solid. The crude was purified by flash chromatography using 90% EtOAc in hexane as eluent to afford **6** as

white solid (550 mg, 80% yield), R_f = 0.45 (80% EtOAc in Hexane). ^{1}H NMR (400 MHz, CHLOROFORM-D) δ 7.81 (s, 1H), 5.49 (s, 1H), 5.26 (s, 2H), 4.99 (s, 2H), 4.09 (q, *J* = 7.1 Hz, 2H), 3.62 (t, *J* = 6.7 Hz, 2H), 3.45 (q, *J* = 6.5 Hz, 2H), 1.49 (s, 6H), 1.43 (s, 9H), 1.16 (t, *J* = 7.1 Hz, 3H),^{13}C NMR (101 MHz, $CDCl_3$) δ 174, 166.5, 159.2, 156.3, 154.1, 151.4, 143.4, 124.9, 80.4, 44.8, 43.5, 41.4, 28.5, 24.3, 14.1. Calculated Mass $(C_{20}H_{30}N_7O_5Cl + H)^+$ 484.2075,Observed Mass 484.2083.

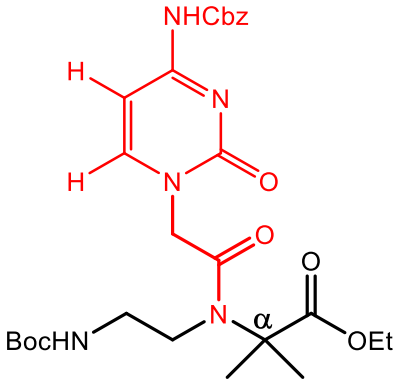
2-(2-(2-amino-6-oxo-1H-purin-9(6H)-yl)-N-(2-((tert-butoxycarbonyl) amino)ethyl)acetamido)-2-methylpropanoic acid (8)



8

α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**6**, 500 mg, 1.1 mmol) was dissolved in 1:1 mixture of 20 ml THF and 20 ml 2(N) NaOH solution and the reaction mixture was stirred at 0^{o}C for 5 h. THF was evaporated and water was added in the crude. Unreacted methyl ester was removed by washing the water layer with EtOAc (3 x 100 ml). The water layer was acidified to pH 3 using 2(N) HCl and the α-*gemdimethyl-aeg*-G-PNA monomer (**8**) was obtained by extracting the same by EtOAc (3 x 150 ml). The combined organic layer was dried over Na_2SO_4 and EtOAc was removed under reduced pressure to obtain the α-*gemdimethyl-aeg*-G-PNA monomer (**8**). The crude was purified by flash chromatography using 10% MeOH in DCM as eluent to afford **8** as white solid (250 mg, 55% yield), R_f = 0.6 (10% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.64 (s, 1H), 3.91 (s, 2H), 3.16 (t, *J* = 6.8 Hz, 2H), 3.03 (t, *J* = 5.9 Hz, 2H), 1.33 (s, 9H), 1.14 (s, 6H), ^{13}C NMR (151 MHz, DMSO-d_6) δ 186.8, 172.1, 160.2, 156.6, 156.3, 149.6, 147.6, 123.8, 89.8, 78.9, 62.5, 28.9, 23.8. Calculated Mass $(C_{18}H_{27}N_7O_6 + Na)^+$ 460.1921,Observed Mass 460.1486.

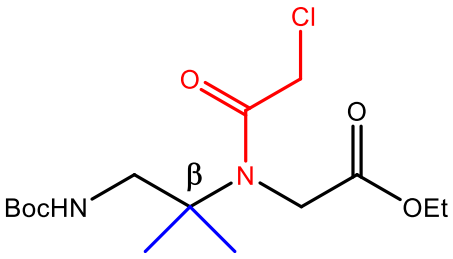
ethyl 2-(2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)-N-(2-((tert-butoxycarbonyl)amino)ethyl)acetamido)-2-methylpropanoate (5)



5

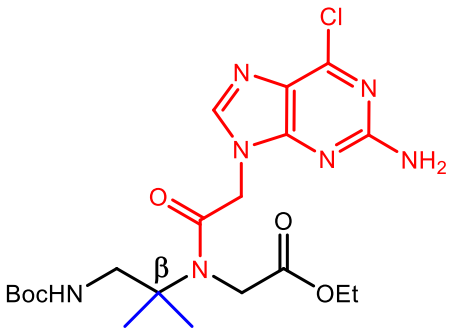
To a suspension of (N^4-Cbz)-cytosine (245 mg, 2.8 mmol) in dry DMF (6 ml) was added anhydrous K_2CO_3 (215 mg, 1.5 mmol) and anhydrous Cs_2CO_3 (46 mg, 0.1 mmol). The reaction mixture was stirred for 5 min, and N-acetylchloro-α-*gdm-aeg* methyl ester (**4**, 500 mg, 1.4 mmol) was added to the mixture. Resulting mixture was stirred at room temperature for additional 48 h. The reaction mixture was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml) . The combined organic layer was washed with brine (3 x 150 ml) and the same was dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the α-*gemdimethyl*-C(N^4-Cbz)-*aeg*-PNA monomer (**5**) as a yellow hygroscopic solid. The crude was purified using column chromatography to afford **5** as white solid (650 mg, 82% yield), R_f = 0.45 (80% EtOAc in Hexane). ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.77 (s, 1H), 7.91 (d, *J* = 7.3 Hz, 1H), 7.48 – 7.33 (m, 5H), 7.01 (s, 1H), 7.00 (s, 1H), 5.18 (d, *J* = 7.2 Hz, 2H), 4.79 (s, 2H), 3.93 (q, *J* = 7.1 Hz, 2H), 3.46 – 3.41 (m, 2H), 3.26 – 3.18 (m, 2H), 1.39 (s, 9H), 1.37 (s, 6H), 1.08 (t, *J* = 7.1 Hz, 3H), ^{13}C NMR (101 MHz, DMSO-d_6) δ 173.2, 166.9, 163.2, 155.7, 155.2,154.6, 151.3 ,150.4, 140, 128.5, 128.1, 127.9, 93.7, 78.1, 66.5, 60.8, 60.1, 50.4, 42.2, 28.2, 23.5, 13.8. Calculated Mass ($C_{27}H_{37}N_5O_8$ + H)$^+$ 560.2720,Observed Mass ($C_{27}H_{37}N_5O_8$ + H)$^+$ 560.2723.

Ethyl-2-(N-(1-((tert-butoxycarbonyl)amino)-2-methylpropan-2-yl)-2-chloroacetamido)acetate (11)



N-acetylchloro-β-*gdm-aeg* methyl ester (**11**) was synthesized from the known amine backbone **10** using the reported procedure.[51] The amine (**10**, 3.5 gm, 12.7 mmol) was dissolved in dry DCM (45 ml) and triethylamine (NEt_3 ,7.1 ml, 50 mmol) was added in the mixture. The solution was cooled to 0°C and chloroacetyl chloride (4.0 ml, 50 mmol) was added to the reaction mixture drop by drop (*Caution-Addition of chloroacetyl chloride produces heat, add the reagent slowly and drop by drop*). The reaction mixture was stirred at room temperature for 30 h. The reaction was quenched by the addition of aq. $NaHCO_3$ and the desired compound was extracted by DCM (3 x 150 ml) . The combined organic layer was dried over $MgSO_4$ and DCM was evaporated to give the desired N-acetylchloro-β-*gdm-aeg* methyl ester (**11**) as white hygroscopic solid. The crude was purified by flash chromatography using 35% EtOAc in hexane as eluent to afford **11** as white crystalline solid (3.5 gm, 78% yield), R_f = 0.64 (45% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.12 (s, 1H), 4.25 (q, *J* = 7.1 Hz, 2H), 4.14 (s, 2H), 3.97 (s, 2H), 3.60 (d, *J* = 6.5 Hz, 2H), 1.41 (s, 9H), 1.34 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 3H). The synthesized compound was used for next step of synthesis without any further analysis.

Ethyl-2-(2-(2-amino-6-chloro-9H-purin-9-yl)-N-(1-((tert-butoxycarbonyl) amino)-2-methylpropan-2-yl)acetamido)acetate (12)

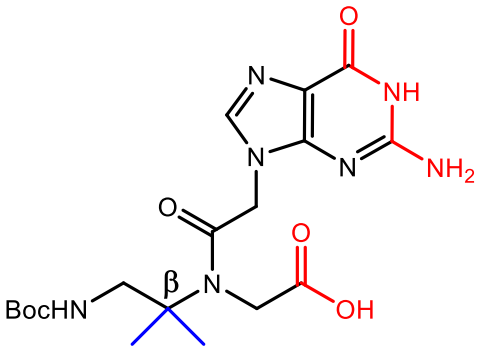


12

2-Amino-6-Chloropurine (264 mg, 1.5 mmol) was dissolved in dry DMF (6 ml) and anhydrous K_2CO_3 (215 mg, 1.5 mmol) and anhydrous Cs_2CO_3 (46 mg, 0.1 mmol) was added to the reaction mixture. The reaction mixture was stirred for 5 min, and N-acetylchloro-β-*gdm-aeg* methyl ester (**11**, 500 mg, 1.4 mmol) was added to the mixture. Resulting mixture was stirred at room temperature for additional 48 h. The reaction mixture was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml) . The combined organic layer was washed with brine (3 x 150 ml) and the same was dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the β,β-*dimethyl-aeg-*

(6-Cl-2-aminopurine) PNA monomer **12** as a white hygroscopic solid. The crude was purified by flash chromatography using 80% EtOAc in hexane as eluent to afford the **12** as white solid (584 mg, 85% yield), R_f = 0.40 (80% EtOAc in Hexane). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.94 (d, *J* = 5.6 Hz, 1H), 6.84 (s, 2H), 6.72 (s, 1H), 4.90 (s, 2H), 4.30 (s, 2H), 4.22 (q, *J* = 7.1 Hz, 2H), 3.38 (s, 2H), 1.39 (s, 9H), 1.26 (t, *J* = 7.1 Hz, 3H), 1.23 (s, 6H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 170.0, 166.7, 166.5, 160.2, 159.3, 157.1, 156.42, 153.9, 153.8, 151.8, 151.8, 143.1, 142.3, 125.1, 80.2, 79.8, 46.8, 46.2, 28.6, 28.5, 25.2, 14.3. Calculated Mass ($C_{20}H_{30}N_7O_5Cl$ + H)$^+$ 484.2075,Observed Mass 484.2082.

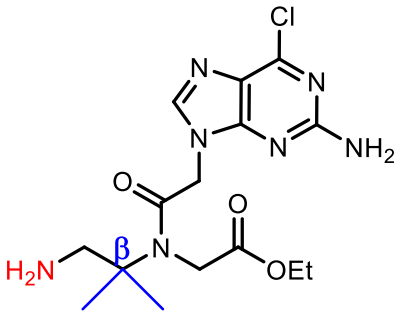
2-(2-(2-amino-6-oxo-1H-purin-9(6H)-yl)-N-(1-((tert-butoxycarbonyl) amino)-2-methylpropan-2-yl)acetamido)acetic acid (14)



14

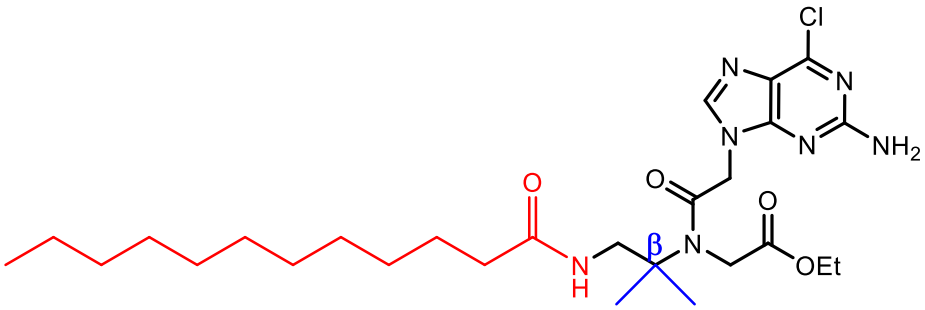
β,β-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**12**, 500 mg, 1.1 mmol) was dissolved in 1:1 mixture of 20 ml THF and 20 ml 2(N) NaOH solution and the reaction mixture was stirred at 0^{o}C for 5 h. THF was evaporated and water was added in the crude. Unreacted methyl ester was removed by washing the water layer with EtOAc (3 x 100 ml). The water layer was acidified to pH 3 using 2(N) HCl and the β-*gemdimethyl-aeg*-G-PNA monomer (**14**) was obtained by extracting the same by EtOAc (3 x 150 ml). The combined organic layer was dried over Na_2SO_4 and EtOAc was removed under reduced pressure to obtain the β-*gemdimethyl-aeg*-G-PNA monomer (**14**). The crude was purified by flash chromatography using 10% MeOH in DCM as eluent to afford **14** as white solid (318 mg, 70% yield), R_f = 0.5 (10% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.93 (s, 1H), 4.85 (s, 2H), 4.15 (s, 2H), 3.35 (dd, *J* = 14.1, 7.1 Hz, 2H), 1.32 (s, 9H), 1.20 (s, 6H), ^{13}C NMR (151 MHz, DMSO-d_6) δ 174.8, 173.4, 168.3, 160.5, 157.7, 155.2, 150.7, 145.7, 123.8, 80.8, 80, 62.5, 60.6, 31.6, 29.1, 25.1, 21.9. Calculated Mass ($C_{18}H_{27}N_7O_6$ +H) $^+$ 438.2102, Observed Mass 438.2112.

ethyl 2-(2-(2-amino-6-chloro-9H-purin-9-yl)-N-(1-dodecanamido-2-methylpropan-2-yl)acetamido)acetate (16)



15

β,β-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**12**, 300 mg, 0.6 mmol) was treated with 50% TFA in DCM and stirred at room temperature for 12 h. After 12 h TFA/DCM was evaporated under reduced pressure give the NH_2-β,β-*gemdimethyl*-6-chloro-2-amino purine-*aeg*-PNA monomer (**15**) as yellow sticky oil. Cooled diethyl ether was added to that yellow sticky oil and resulted white suspension was sonicated to give the NH_2-β,β-*gemdimethyl*-6-chloro-2-amino purine-*aeg*-PNA monomer (**15**) as hygroscopic white solid and used for the next step without further purification.

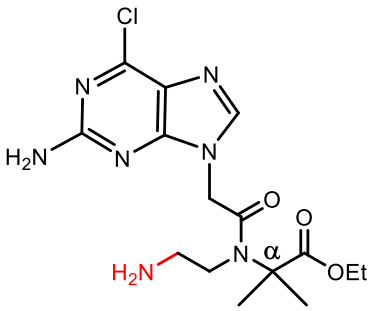


16

Dodecanoic acid (115 mg, 0.5 mmol) was dissolved in dry DMF (4 ml) and followed by HBTU (335 mg, 1.1 mmol), HOBt (210 mg, 1.5 mmol) was added to the reaction mixture. The mixture was stirred for 5 min at room temperature. NH_2-β,β-*gemdimethyl*-6-chloro-2-amino purine-*aeg*-PNA monomer (**15**, 200 mg, 0.5 mmol) was dissolved in 2 ml dry DMF and DIPEA (271 µL, 1.5 mmol) was added to it, the resulted mixture was added to the reaction mixture containing the activated dodecanoic acid. The combined reaction mixture was stirred at room temperature for 48 h. After completion of the reaction, the reaction was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and the same was dried over

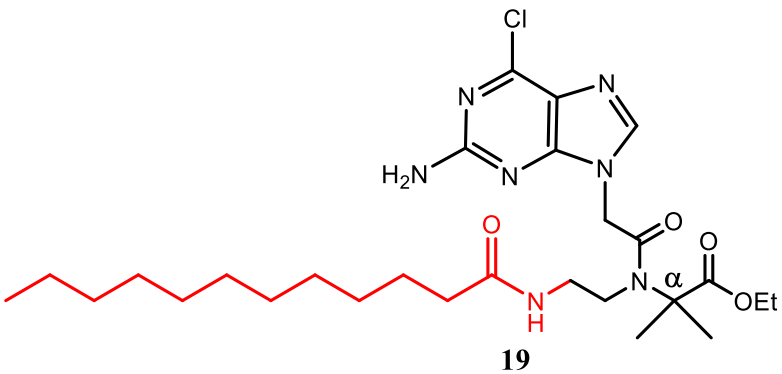
Na_2SO_4. Evaporation of the organic layer under reduced pressure give the N-decanoyl β-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**16**) as white hygroscopic solid. The crude was purified by flash chromatography using 10% MeOH in DCM as eluent to afford **16** as white hygroscopic solid (125 mg, 53% yield), R_f = 0.45 (100% EtOAc). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.86 (d, *J* = 5.5 Hz, 1H), 6.66 (d, *J* = 7.9 Hz, 1H), 5.28 (s, 2H), 4.83 (s, 2H), 4.30 – 4.24 (m, 2H), 4.22 (d, *J* = 5.8 Hz, 2H), 3.75 – 3.65 (m, 2H), 2.15 (t, *J* = 7.6 Hz, 2H), 1.56 (d, *J* = 6.4 Hz, 2H), 1.33 (d, *J* = 4.6 Hz, 6H), 1.22 (s, 19H), 0.85 (t, *J* = 6.7 Hz, 3H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 174, 170.9, 167.2, 159.2, 154.0, 151.4, 143.6, 124.7, 62.6, 46.6, 45.8, 36.9, 32.2, 32.0, 29.8, 29.7, 29.5, 26, 25.5, 22.8, 14.3. Calculated Mass ($C_{27}H_{44}N_7O_4Cl$ +H)$^+$ 566.3221,Observed Mass 566.3228.

Ethyl 2-(2-(2-amino-6-chloro-9H-purin-9-yl)-N-(2-dodecanamidoethyl) acetamido)-2-methylpropanoate (19)



18

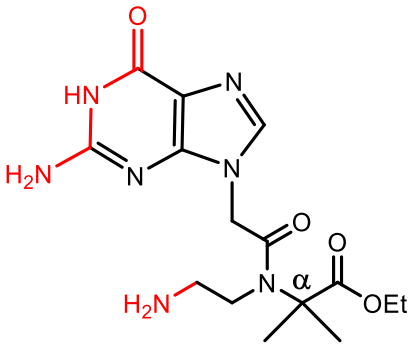
α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**6**, 300 mg, 0.6 mmol) was treated with 50% TFA in DCM and stirred at room temperature for 12 h. After 12 h TFA/DCM was evaporated under reduced pressure give the NH_2-α,α-*gemdimethyl*-6-chloro-2-amino purine-*aeg*-PNA monomer (**18**) as yellow sticky oil. Cooled diethyl ether was added to that yellow sticky oil and resulted white suspension was sonicated to give the NH_2-α,α-*gemdimethyl*-6-chloro-2-amino purine-*aeg*-PNA monomer (**18**) as hygroscopic white solid and used for the next step without further purification.



19

Dodecanoic acid (115 mg, 0.5 mmol) was dissolved in dry DMF (4 ml) and followed by HBTU (335 mg, 1 mmol), HOBt (210 mg, 1.5 mmol) was added to the reaction mixture. The mixture was stirred for 5 min at room temperature. NH_2-α,α-*gemdimethyl*-6-chloro-2-amino purine-*aeg*-PNA monomer (**18**, 200 mg, 0.5 mmol) was dissolved in 2 ml dry DMF and DIPEA (271 µL, 1.5 mmol) was added to it, the resulted mixture was added to the reaction mixture containing the activated dodecanoic acid. The combined reaction mixture was stirred at room temperature for 48 h. After completion of the reaction, the reaction was quenched by addition of water and the desired organic compound was extracted by EtOAc (3 x 150 ml). The combined organic layer was washed with brine (3 x 100 ml) and the same was dried over Na_2SO_4. Evaporation of the organic layer under reduced pressure give the N-decanoyl α-*gemdimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer (**19**) as white hygroscopic solid. The crude was purified by flash chromatography using 10% MeOH in DCM as eluent to afford **19** as white hygroscopic solid (129 mg, 55% yield), R_f = 0.4 (100% EtOAc). δ 7.74 (d, *J* = 9.2 Hz, 1H), 7.52 (q, *J* = 8.7 Hz, 2H), 5.08 (s, 2H), 4.10 – 4.01 (m, 2H), 3.59 (d, *J* = 12.9 Hz, 4H), 2.22 (d, *J* = 15.9 Hz, 2H), 1.60 (d, *J* = 7.7 Hz, 2H), 1.49 (d, *J* = 9.5 Hz, 6H), 1.33 – 1.24 (m, 10H), 1.22 (s, 9H), 0.86 (t, *J* = 6.8 Hz, 3H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 174.5, 166.6, 159.2, 158.9, 156.4, 143.3, 129.1, 125.2, 119.9, 116.2, 109.4, 62.3, 61.4, 45.1, 42.8, 40.1, 36.6, 32.1, 29.8, 29.7, 29.6, 29.5, 29.4, 25.7, 24.2, 22.8, 14.2, 14.1. Calculated Mass $(C_{27}H_{44}N_7O_4Cl$ $+H)^+$ 566.3221,Observed Mass 566.3222.

ethyl 2-(2-(2-amino-6-oxo-1,6-dihydro-9H-purin-9-yl)-N-(2-aminoethyl)acetamido)-2-methylpropanoate (17)

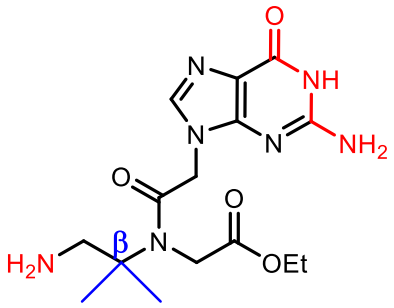


17

α,α-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**6**, 900 mg, 1.8 mmol) was dissolved in 75% TFA in water and stirred at room temperature for 48 h to complete the reaction. After 48 h TFA was evaporated under reduced pressure. The obtained sticky oil is the free amine α-*gemdimethyl-aeg*-G-PNA monomer (**17**). To remove the undesirable impurities,

the obtained sticky oil was treated with cooled diethyl ether and afford **17** as a hygroscopic white solid. The crude solid was purified by flash chromatography using 20% MeOH in DCM as eluent to obtain the free amine α-*gemdimethyl-aeg*-G-PNA monomer (**17**) as hygroscopic white solid (515 mg, 75% yield), R_f 0.3 (50% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.64 (s, 1H), 7.92 (s, 2H), 7.61 (s, 1H), 6.38 (s, 2H), 4.97 (s, 2H), 3.93 (q, *J* = 7.1 Hz, 2H), 3.73 – 3.66 (m, 2H), 3.19 (s, 2H), 1.39 (s, 6H), 1.03 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174, 166.5, 159.2, 156.3, 154, 151.4, 143.4, 124.8, 62.3, 61.6, 44.8, 43.5, 41.4, 24.3, 14.10, Calculated Mass $(C_{15}H_{23}N_7O_4 + H)^+$ 366.1892 observed mass 366.1890.

ethyl N-(1-amino-2-methylpropan-2-yl)-N-(2-(2-amino-6-oxo-1,6-dihydro-9H-purin-9-yl)acetyl)glycinate (13)



13

β,β-*dimethyl-aeg*-(6-Cl-2-aminopurine) PNA monomer (**12**, 900 mg, 1.8 mmol) was dissolved in 75% TFA in water and stirred at room temperature for 48 h to complete the reaction. After 48 h TFA was evaporated under reduced pressure. The obtained sticky oil is the free amine β-*gemdimethyl-aeg*-G-PNA monomer (**13**). To remove the undesirable impurities, the obtained sticky oil was treated with cooled diethyl ether and afford **13** as a hygroscopic white solid. The crude solid was purified by flash chromatography using 20% MeOH in DCM as eluent to obtain the free amine β-*gemdimethyl-aeg*-G-PNA monomer (**13**) as hygroscopic white solid (515 mg, 75% yield), R_f 0.3 (50% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.76 (d, *J* = 64.9 Hz, 1H), 9.01 (s, 1H), 7.67 (d, *J* = 39.9 Hz, 1H), 6.65 (s, 1H), 6.48 (s, 1H), 5.22 (s, 1H), 4.68 (s, 1H), 4.62 (s, 1H), 4.19 (dq, *J* = 21.4, 7.1 Hz, 2H), 4.02 (s, 1H), 3.67 (s, 2H), 3.35 (d, *J* = 6.1 Hz, 1H), 1.33 (s, 6H), 1.29 – 1.18 (m, 3H), ^{13}C NMR (101 MHz, DMSO-d_6) δ 168.2, 167.8, 167, 164.3, 156.7, 154, 153.6, 151.4, 136.9, 116.2, 66.8, 61.9, 61.8, 59.9, 55.7, 24.4, 21.3, 13.8, Calculated Mass $(C_{15}H_{23}N_7O_4 + H)^+$ 366.1892 observed mass 366.1891.

3.7. References

1. Kalyani, T.; Dhoble. S.J. *Renewable and Sustainable Energy Reviews*., **2012**, *16*, 2696–2723.

2. a) Frackowiak. D. *Journal of Photochemistry and Photobiology, B: Biology*., **1998**, *2*, 399 – 408. b) Jablonski, A. Efficiency of anti-Stokes fluorescence in dyes, *Nature*, **1993**, *131* ,839 – 840. c) Jablonski, A. *Z. Phys*., **1935**, *94* ,38 - 44. d) Lewis, G.N.; and Kasha, M. *J. Am. Chem. Soc*., **1994**, *66*, 2100. e) Kasha, M. *Acta Phys. Pol. A*, **1987**, *71*, 661 - 670. f) Jablonski, A. *Bull. Acad. Pol. Sci*., **1958**, *6*, 589 - 593.

3. Xu, W.; Chan. K.M.; Kool, E.T. *Nature Chemistry*., 2017, *9*, 1043-1055.

4. Ward, D. C.; Reich, E.; Stryer, L. *J. Biol. Chem*., **1969**, *244*, 1228–1237.

5. a) Secrist, J. A.; Barrio, J. R.; and Leonard, N. J. *Science*., **1972**, *175*, 646–647. b) Rist, M. J.; Marino, J. P. *Curr. Org. Chem*., 2002, *6*, 775–793.

6. a) Wu, P.; Nordlund, T. M.; Gildea, B.; McLaughlin, L. W. *Biochemistry*., **1990**, *29*, 6508–6514. b) Hawkins, M. E.; Pfleiderer, W.; Mazumder, A.; Pommier, Y. G.; Balis, F. M. *Nucleic Acids Res*., **1995**, *23*, 2872–2880.

7. Ren, R. X.; Chaudhuri, N. C.; Paris, P. L.; Rumney, S.; Kool, E. T. *J. Am. Chem. Soc*., **1996**, *118*, 7671–7678.

8. a) Okamoto, A.; Kanatani, K; and Saito, I. *J. Am. Chem. Soc*., **2004**, *126*, 4820–4827. b) Okamoto, A.; Tainaka, K.; and Saito, I. *J. Am. Chem. Soc*., **2003**, *125*, 4972–4973. c) Okamoto, A.; Tanaka, K.; Fukuta, T.; and Saito, I. *J. Am. Chem. Soc*., **2003**, *125*, 9296–9297.

9. a) Hawkins, M. E.; Pfleiderer, W.; Balis, F. M.; Porter, D.; and Knutson, J. R. *Anal. Biochem*., **1997**, *244*, 86–95. b) Hawkins, M. E.*Cell Biochem. Biophys*., **2001**, *34*, 257–281. c) Hawkins, M. E.; Pfleiderer, W.; Jungmann, O.; Balis, F. M. *Anal. Biochem*., **2001**, *298*, 231–240.

10. Ward, D. C.; Reich, E.; and Stryer, L. *J. Biol. Chem*., **1969**, *244*, 1228–1237.

11. Wierzchowski, J.; Wielgus-Kutrowska, B.; Shugar, D. *Biochim. Biophys. Acta*., **1996**, *1290*, 9–17.

12. a) Gaied, N.B.; Glasser, N.; Ramalanjaona, N.; Beltz1, H.; Wolff, P.; Marquet, R.; Burger, A.; Me´ly1, A. *Nucleic Acids Res.*, **2005**, *33*, 1031–1039. b) Nadler, A.; Strohmeier, J.; Diederichsen, U. *Angew. Chem. Int. Ed*., **2011**, *50*, 5392–5396.

13. a) Mitsui, T.; Kimoto, M.; Kawai, R.; Yokoyama, S.; Hirao, I. *Tetrahedron*., **2007**, *63*, 3528–3537. b) Greco, N. J.; and Tor, Y. *Tetrahedron*., **2007**, *63*, 3515–3527. c) Sun, K. M.; McLaughlin, C. K.; Lantero, D. R.; & Manderville, R. A. *J. Am. Chem. Soc*., **2007**, *129*, 1894–1895. d) Schlitt, K. M.; Millen, A. L.; Wetmore, S. D.; & Manderville, R. A. *Org. Biomol. Chem*., **2011**, *9*, 1565–1571.

14. a) Leonard, N. J.; Sprecker, M. A.; & Morrice, A. G. *J. Am. Chem. Soc*., **1976**, *98*, 3987–3994. b) Okamoto, A.; Tanaka, K.; Fukuta, T.; & Saito, I. *ChemBioChem*., **2004**, *5*, 958–963. c) Okamoto, A.; Tanaka, K.; Fukuta, T.; & Saito, I. *J. Am. Chem. Soc*., **2003**, *125*, 9296–9297.

15. Greco, N. J.; & Tor, Y. *Tetrahedron*., **2007**, *63*, 3515–3527.

16. Wu, P.; Nordlund, T. M.; Gildea, B.; & McLaughlin, L. *Biochemistry*., **1990**, *29*, 6508–6514.

17. Godde, F.; Toulme, J. J.; & Moreau, S. *Biochemistry*., **1998**, *37*, 13765–13775.

18. a) Korshun, V.A.; Prokhorenko, I.A.; Gontarev, S.V.; Skorobogatyi, M.V.; Balakin, K.V.; Manasova, E.V.; Malakhov, A.D.; and Berlin, Y.A. *Nucleosides* Nucleotide.,**1997**, *16*, 1461–1464. b) Ehrenschwender, T.; Wanninger-Weiss, C.; & Wagenknecht, H. A. *Nucleic Acids Symp. Ser*., **2008**, *52*, 349–350.

19. Dodd, D.W.; Swanick, K.N.; Price, J.T.; Brazeau, A.L.; Ferguson, M.J.; Jones, N.D.; and Hudson, R.H.E. *Org. Biomol. Chem*., **2010**, *8*, 663–666.

20. Shin, D.; Sinkeldam, R. W.; & Tor, Y. *J. Am. Chem. Soc*., **2011**, *133*, 14912–14915.

21. Godde, F.; Toulme, J. J.; & Moreau, S. *Biochemistry*., **1998**, *37*, 13765–13775.

22. a) Xu, W.; Lee, M.M.S.; Zhang, Z.; Sung, H.H.Y.; Williams, I.D.; Kwok, R.T.K.; Lam, J.W.Y.; Wang, D.; and Tang, B.Z. *Chem. Sci*., **2019**, *10*, 3494–3501. b) Shen, X.Y.; Wang, Y.J.; Zhang, H.; Qin, A.; Sun, J.Z.; and Tang, B.Z. *Chem. Comm*., **2014**, *50* ,8747–8750. c) Shen, X.Y.; Wang, Y.J..; Zhang, H.; Qin, A.; Sun, J.Z.; and Tang, B.Z. *Chem. Comm*., **2015**,*51*, 4830–4833. d) Liu, S.; Zhou, X.; Zhang, H.; Ou, H.; Lam, J.W.Y.; Liu, Y.; Shi,

L.; Ding, D.; and Tang. B.Z. *J. Am. Chem. Soc*., **2019,** *141*, 5359–5368 e) Zhu, C.; Kwok, R.T.K.; Lam, J.W.Y.; and Tang, B.Z. ACS *Appl. Bio Mater*., **2018,** *1*, 1768–1786.

23. a) Zhang, H.; Zhao, Z.; McGonigal, P.R.; Ye, R.; Liu, S.; Lam, J.W.Y.; Kwok, R.T.K.; Yuan, W.Z.; Xie, J.; Rogach, A.L.; Tang, B.Z. *Materials Today*., **2020**, *32*, 275-292. b) Larson, C.L.; Tucker, S.A. *Appl. Spectrosc*., **2001**, *55*, 679–683. c) Varnavski, O.; Ispasoiu, R.G.; Balogh, L.; Tomalia, D.; Goodson. T. *J. Chem. Phys*., **2001**, *114*, 1962–1965. d) Lee, W.I.; Bae, Y.; and Bard, A.J. *J. Am. Chem. Soc*., **2004**, *126*, 8358–8359. e) Liu, X.; Zeng, Y.; Liu, J.; Li, P.; Zhang, D.; Zhang, X.; Yu, T.; Chen, J.; Yang, G.; and Li, Y. *Langmuir*., **2015,** *31* ,4386–4393. f) Al-Jamal, K.T.; Ruenraroengsak, P.; hartel, N.; & Florence, A.T. *J. Drug Target*., **2006**, *14*, 405–412. g) Marek, P.; Gupta, R.; and Raleigh, D.P. *Chembiochem*., **2008**, *9*, 1372–1374. h) Turoverov, K.K.; Kuznetsova, I.M. *J. Fluoresc*., **2003**, *13*, 41–57. i) Wang, H.; Zhao, E.; Lam, J.M.Y.; Tang, B.Z. *Mater. Today*., **2015**, *18*, 365–377. j) Zhang, Y.; Shen, P.; He, B.; Luo, W.; Zhao, Z; and Tang, B.Z. *Polym. Chem*., **2018**, *9*, 558–564. k) Schilter, D. *Nat. Rev. Chem*., **2017**, *1*, 0097. l) Tomalia, D.A.; Maculewicz, B.K.; Johnson, K.A.M.; Brinkman, H.F.; Janaszewska, A.; Hedstrand , D.M. *Prog. Polym. Sci*., **2019**, *90*, 35–117. m) Zhou, Q.; Wang, Z.; Dou, X.; Wang, Y.; Liu, S.; Zhang, Y.; and Yuan, W. Z. *Mater. Chem. Front*., **2019**, *3*, 257–264. n) Zhang, T.; Zhao, Z.; Ma, H.; Zhang, Y.; and Yuan, W.Z. *Chem. Asian J*., **2019**, *14*, 884–889. o) Yuan, W.Z.; Zhang, Y. *J. Polym. Sci. A*., **2017**, *55*, 560–574. p) Konopka, M; Janaszewska, A.; and Maculewicz, B.K. *Polymers*., **2018**, *10*, 540.

24. a) Shimomura, O.; Johnson, F.H.; Saiga Y. *J. Cell. Physiol*., **1962**, *59*, 223–239. b) Weiss, P.S. *ACS Nano*., **2008**, *2*, 1977.

25. a) Jiang, M.; He, Z.; Zhang, Y.; Sung, H.H.Y.; Lam, J.W.Y.; Peng, Q..; Yan, Y.; Wong, K.S.; Williams, I.D.; Zhao, Y.; and Tang, B.Z. *J. Mater. Chem. C*., **2017**, *5*, 7191–7199. b) Barondeau, D.P.; Putnam, C.D.; Kassmann, C.J.; Tainer, J.A.; and Getzoff, E.D. *Proc. Natl. Acad. Sci. U.S.A*., **2003**, *100*, 12111–12116.

26. Mercato, L.L.; Pompa, P.P.; Maruccio, G.; Torre, A.D.; Sabella, S.; Tamburro, A.M.; Cingolani, R.; and Rinaldi, R. *Proc. Natl. Acad. Sci. U.S.A.*, **2007**, *104*, 18019–18024.

27. Yang, G.Y.; Qiang, T.Y.; Ju, M.; Yiren, Z.; Zhang, Y.W.; Ming, Z.Y.; Zhi, S.J.; and Zhong, T.B. *Sci. China Chem*., **2013**, *56*, 1178–1182.

28. Dou, X.; Zhou, Q; Chen, X; Tan, Y.; He, X.; Lu, P.; Sui, K.; Tang, B.Z.; Zhang, Y.; and Yuan, W.Z. *Biomacromolecules*., **2018**, *19*, 2014–2022.

29. Larson, C.L.; Tucker, S.A. *Appl. Spectrosc*., **2001**, *55*, 679–683.

30. a) Wu, D.; Liu, Y.; He, C.; and Goh, S.H. *Macromolecules*., **2005**, *38*, 9906–9909. b) Lin, S.Y.; Wu, T.H.; Jao, Y.C.; Liu, C.P.; Lin, H.Y.; Wei Lo, L.; and Yang, C.S. *Chem. Eur. J.*, **2011**, *17*, 7158–7161.

31. Wang, Y.; Bin, X.; Chen, X.; Zheng, S.; Zhang, Y.; and Yuan, W.Z. *Macromol. Rapid. Commun*., **2018,** *39*, 1800528.

32. a) Shang, C.; Wei, N.; Zhuo, H.; Shao, Y.; Zhang, Q.; Zhang, Z.; and Wang, H. *J. Mater. Chem. C*., **2017**, *5*, 8082–8090.

33. Shang, C.; Wei, N.; Zhuo, H.; Shao, Y.; Zhang, Q.; Zhang, Z.; and Wang, H. *J. Mater. Chem. C*., **2017,** *5*, 8082–8090.

34. a) Zhao, W.; Cheung, T.S.; Jiang, N.; Huang, W.; Lam, J.W.Y.; Zhang, X.; He, Z.; & Tang. B.Z. *Nat. Commun*., **2019**, *10*, 1595. b) Wang, J.; Mei, J.; Hu, R.; Sun, J.Z.; Qin, A.; and Tang, B.Z. *J. Am. Chem. Soc.*, **2012**, *134*, 9956–9966. c) Peng, H.Q; Liu, B.; Wei, P.; Zhang, P.; Zhang, H.; Zhang, J.; Li, K.; Li, Y.; Cheng, Y.; Lam, J.W.Y.; Zhang, W.; Lee, C. S.; and Tang, B.Z. *ACS Nano*., **2019**, *13*, 839–846. d) Yang, Z.; Qin, W.; Leung, N.L.C.; Arseneault, M.; Lam, J.W.Y.; Liang, G.; Sung, H.Y.; Williams, I.D.; and Tang, B.Z. *J. Mater. Chem. C*., **2016,** *4*, 99–107. e) Shang, X.; Song, I.; Ohtsu, H.; Tong, J.; Zhang, H.; & oh, H.J. *Sci. Rep*., **2017**, *7*, 5508. f) Hu, Y.; Song, F.; Xu, Z.; Tu, Y.; Zhang, H.; Cheng, Q.; Lam, J.W.Y.; Ma, D.; Tang, B.Z. *ACS Appl. Polym. Mater*., **2019**, *1* ,221–229.

35. a) Zhang, Y.; He, B.; Luo, W.; Peng, H.; Chen, S.; Hu, R.; Qin, A.; Zhao, Z.; and Tang, B.Z. *J. Mater. Chem. C*., **2016**, *4*, 9316–9324. b) Martı´nez, A.G.; Barcina, J.O.; Cerezo, F.; and Rivas, R.G. *J. Am. Chem. Soc*., **1998**, *120*, 673–679. c) Chen, O.; Wang, Y.H.; He, B.; Nie, H.; Hu, R.; Huang, F.; Qin, A.; Zhou, X.S.; Zhao, Z.; and Tang, B.Z. *Angew. Chem. Int. Ed*., **2015**, *54*, 4231–4235. d) Zhang,H.; Zheng, X.; Xie, N.; He, Z.; Liu, J.; Leung, N.L.C.; Yingli Niu, Y.; Huang, X.; Wong, K.S.; Kwok, R.T.K.; Sung, H.Y.; Williams, I.D.; Qin, A.; Lam, J.W.Y..; and Tang, B.Z. *J. Am. Chem. Soc*., **2017**, *139*, 16264–16272.

36. Han, T.; Haiqin Deng, H.; Qiu, Z.; Zhao, Z.; Zhang, H.; Zou, H.; Leung, N.L.C.; Shan, G.; Elsegood, M.R.J.; Lam, J.W.Y.; and Tang, B.Z. *J. Am. Chem. Soc*., **2018**, *140*, 5588–5598.

37. a) Esrafili, M.D.; Vakili, M. *J. Mol, Model*., **2014**, *20*, 2291. b) Antonijevic, I.S.; Janjić, G.V.; Milcič, M.; and Zaric, S.D. *Cryst. Growth Des.*, **2016**, *16*, 632–639.

38. Ye, R.; Liu, Y.; Zhang, H.; Su, H.; Zhang, Y.; Xu. L.; Hu, R.; Kwok, R.T.K.; Wong, K.S.; Lam, J.WY.; Goddard, W.A.; and Tang**,** B.Z. *Polym. Chem*., **2017,** *8*, 1722–1727.

39. Chen, X.; Luo, W.; Ma, H.; Peng, Q.; Yuan, W.Z.; and Zhang, Y. *Sci. China Chem*., **2017,** *61*, 351–359.

40. Qin Zhang, Q.; Mao, Q.; Shang, C.; Chen, Y.N.; Peng, X.; Tan, H.; and Wang, H. *J. Mater. Chem. C*., **2017**, *5*, 3699–3705.

41. Yan, J.; Zheng, B.; Pan, D.; Yang, R..; Xu, Y.; Wang, L.; and Yang**,** M. *Polym.* Chem., **2015**, *6*, 6133–6139.

42. Halpern, A.M. *Chem. Phys. Lett*., **1970**, *6*, 296–298.

43. Chu, C.C and Imae, T. *Macromol. Rapid. Commun*., **2009**, *30*, 89–93.

44. H. Zhang, *Aggregation-Induced Emission: Mechanistic Study, Clusteroluminescence and Kinetic Resolution PhD Thesis*, *Hong Kong University of Science and Technology***, 2018.**

45. Wang, Y.; Bin, X.; Chen, X.; Zheng, S.; Zhang, Y.; and Yuan, W.Z. *Macromol. Rapid. Commun*., **2018,** *39*, 1800528.

46. a) Qian, H.; Cousins, M.E.; Horak, E.H.; Wakefield, A.; Liptak, M.D.; and Aprahamian, I. *Nat. Chem*., **2017**, *9*, 83–87. b) Guo, J.; Fan, J.; Lin, L.; Zeng, J.; Liu, H.; Wang, C.K.; Zhao, Z.; and Tang, B.Z. *Adv. Sci.*, **2019**, *6*, 1801629. c) He, Z.; Zhao, W.; Lam, J.W.Y.; Peng, Q.; Ma, H.; Liang, G.; Shuai, Z.; and Tang, B.Z. *Nat Commun*., **2017,** *8*, 416. d) halpern, A.M.; Roebber, J.; and Weiss, K. *J. Chem. Phys*, **1968**, *49*, 1348–1357. e) Ohrn, A. and Karlstro, G. *J. Phys. Chem. A*., **2006**, *110*, 1934–1942. f) N.J. Turro *et al., Modern Molecular Photochemistry of Organic Molecules, University Science Books, Sausalito, Calif., 2010.* h) I.V. Alabugin, *Stereoelectronic Effects: A Bridge Between Structure and Reactivity, John Wiley & Sons, Ltd., Chichester, 2016*.

47. Berger, O.; Abramovich, L.A.; Sakin, M.L.; Grunwald, A.; Peer, Y.L.; Bachar , M.; Buzhansky, L.; Mossou, E.; Forsyth, V.T.; Schwartz, T.; Ebenstein, Y.; Frolow,F.; Shimon, L.J.W.; Patolsky, F.; and Gazit, E. *Nature Nanotechnology*, **2015**, *10*, 353–360.

48. a) Lakowicz, J.R and Nakamoto, S.K. *Biochemistry.*, **1984**, *23*, 3013–3021. (DOI:10.1021/bi00308a026) b) Kabir, M.L.; Wang, F.; and Clayton, A.HA. *Int. J. Mol. Sci*., **2021**, *22*, 2582. c) Chattopadhyay. A. and Haldar, S. *Accounts of Chemical Research*., **2014**, *47*, 12-19. d) Chattopadhyay, A and Mukherjee, S. *J. Phys. Chem. B*., **1999**, *103*, 8180-8185. d). Berger, O.; Abramovich, L.A.; Sakin, M.L.; Grunwald, A.; Peer, Y.L.; Bachar , M.; Buzhansky, L.; Mossou, E.; Forsyth, V.T.; Schwartz, T.; Ebenstein, Y.; Frolow,F.; Shimon, L.J.W.; Patolsky, F.; and Gazit, E. *Nature Nanotechnology*, **2015**, *10*, 353–360.

49. Avitabile, C.; Diaferia, C.; entura, B.D.; Mercurio, F.A.; Leone, M.; Roviello, V.; Saviano, M.; Velotta, R.; Morelli, G.; Accardo, A and Romanelli, A. *Chem. Eur. J*., **2018**, *24*, 4729 – 4735.

50. Petersson, B.; Nielsen, B.B.; Rasmussen, H.; Larsen, I.K.; Gajhede, M.; Nielsen, P.E. and Kastrup, J.S. *J. Am. Chem. Soc*., **2005**, *127*, 1424-1430.

51. Kulkarni, P.; Datta, D.; Ramabhadran, R.O. and Ganesh, K.N. *Org. Biomol. Chem*., **2021**, *19*, 6534.

52. Vass, E.; Hollo´si, M.; Besson, F. and Buchet, R. *Chem. Rev*. **2003**, *103*, 1917−1954.

53. (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.*

54. Muttenthaler, M.; Albericio, F.; Dawson, P. E. *Nat. Protoc*.,**2015**, *10*, 1067.

55. Deshpande, N.U.; Virmani, M and Jayakannan, M. *Polym. Chem*., **2021**, *12*, 1549–1561.

56. Oliveira, O.N.; Caseli, J.L and Ariga, K. *Chem. Rev*.,**2022,** *122*, 6459–6513.

57. a) Lakowicz, J.R. *Principles of Fluorescence Spectroscopy*, 3rd ed., Springer: New York, **1999**. b) Xavier, J.A.M.; Devatha, G.; Roy, S.; Rao, A. and Pillai, P.P. *J. Mater. Chem. A*, **2018**, *6*, 22248–22255.

58. Bruker, (2016). *APEX3* (v2016.1-0), *SAINT* (Version 8.18C) and *SADABS* (Version 2012/1). Bruker AXS Inc., Madison, Wisconsin, USA.

59. Sheldrick, G. M. *Acta Cryst*. **2015**, *A71*, 3-8.

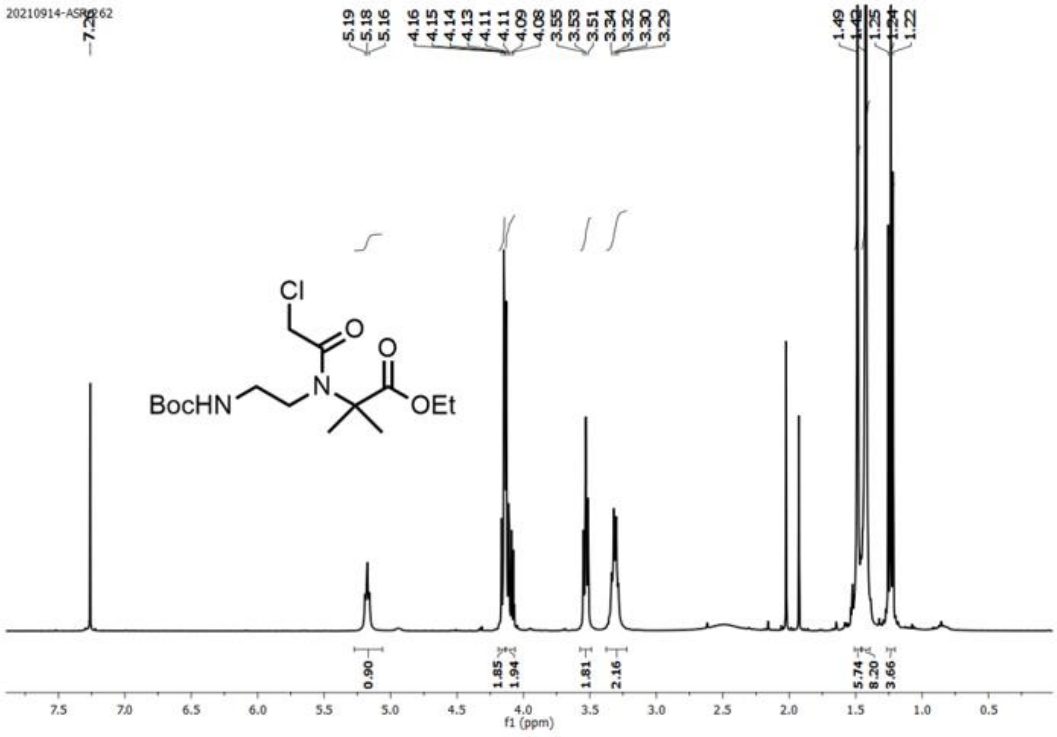
60. Sheldrick, G. M. *Acta Cryst*. **2008**, *A64*, 112-122.

61. Vass, E.; Hollo´si, M.; Besson, F.and Buchet, R. *Chem. Rev*. **2003**, *103*, 1917−1954.

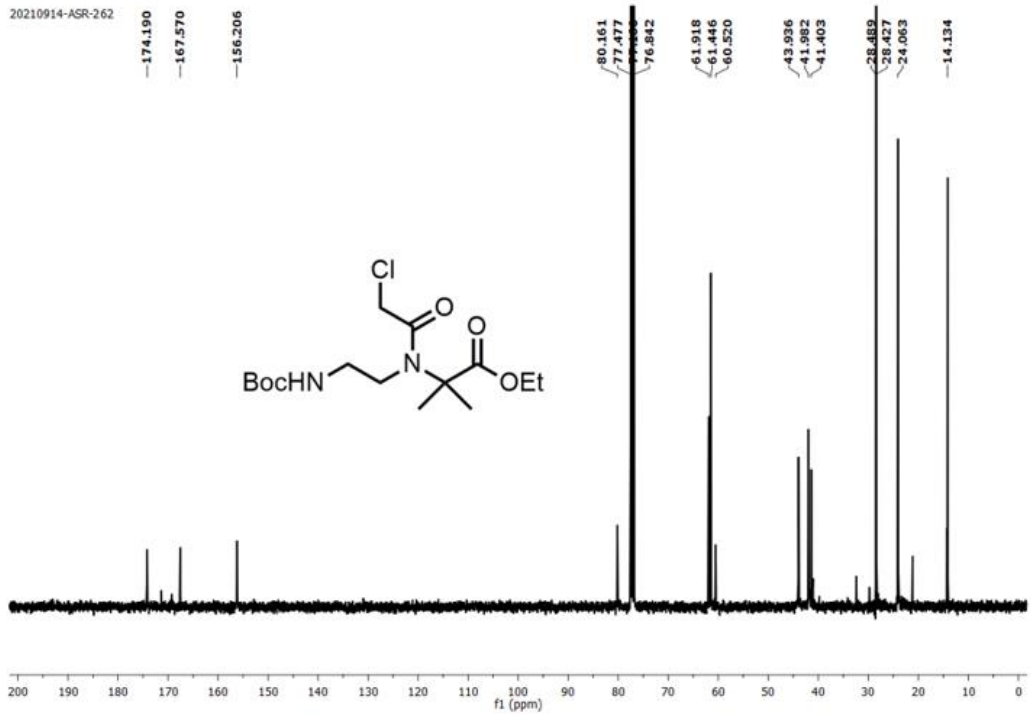
3.7 Appendix-Characterisation data of the synthesized compounds and dimers (Appendix-II)

Entry	Table of contents	Page no
1	**^{1}H NMR, ^{13}C NMR, IR and HRMS spectra of the monomers**	**219-237**
2	**MALDI, HPLC of dimers and emission data of monomers and 2D-NMR of monomers**	**238-245**

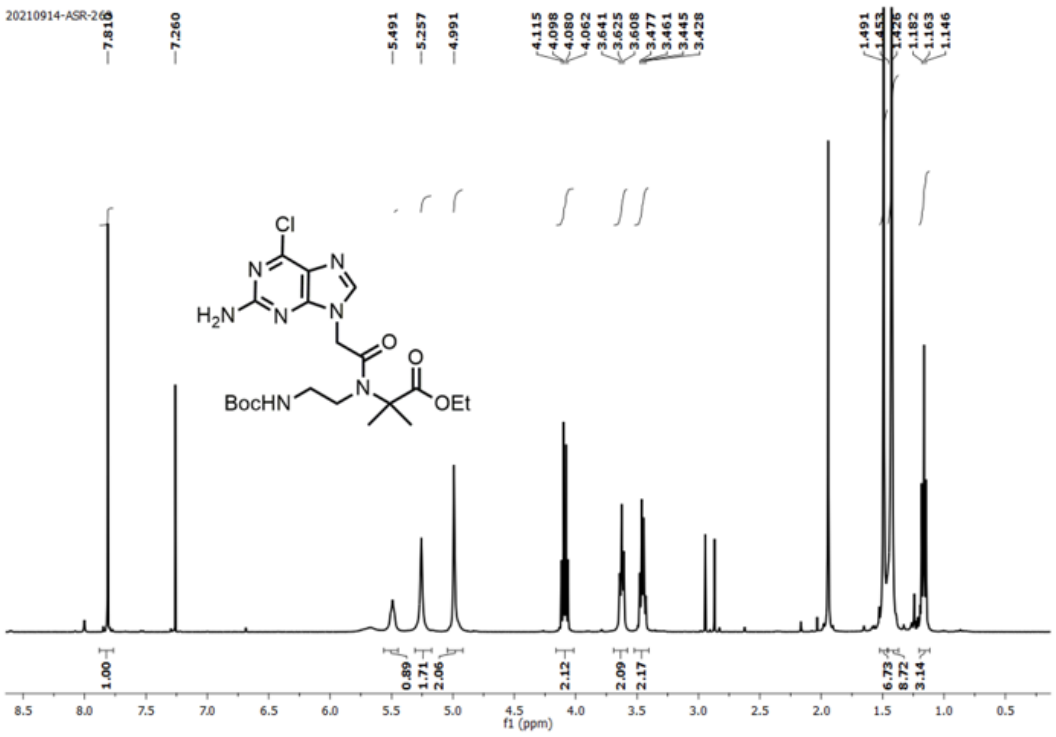
^{1}H NMR of compound 4 in $CDCl_3$ (400 MHz)



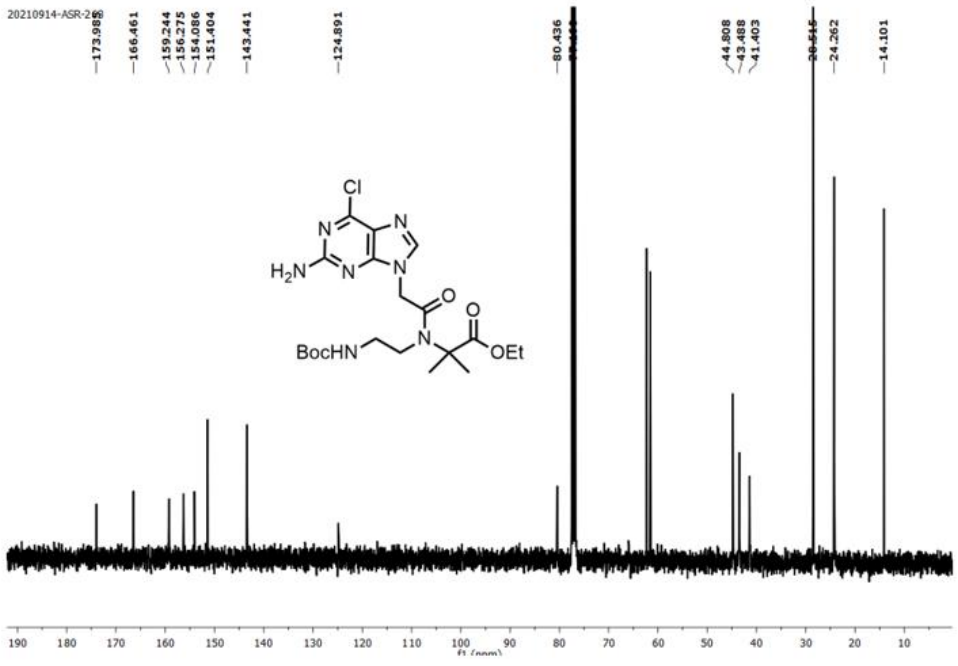
^{13}C {1H} NMR of compound 4 in $CDCl_3$ (101 MHz)



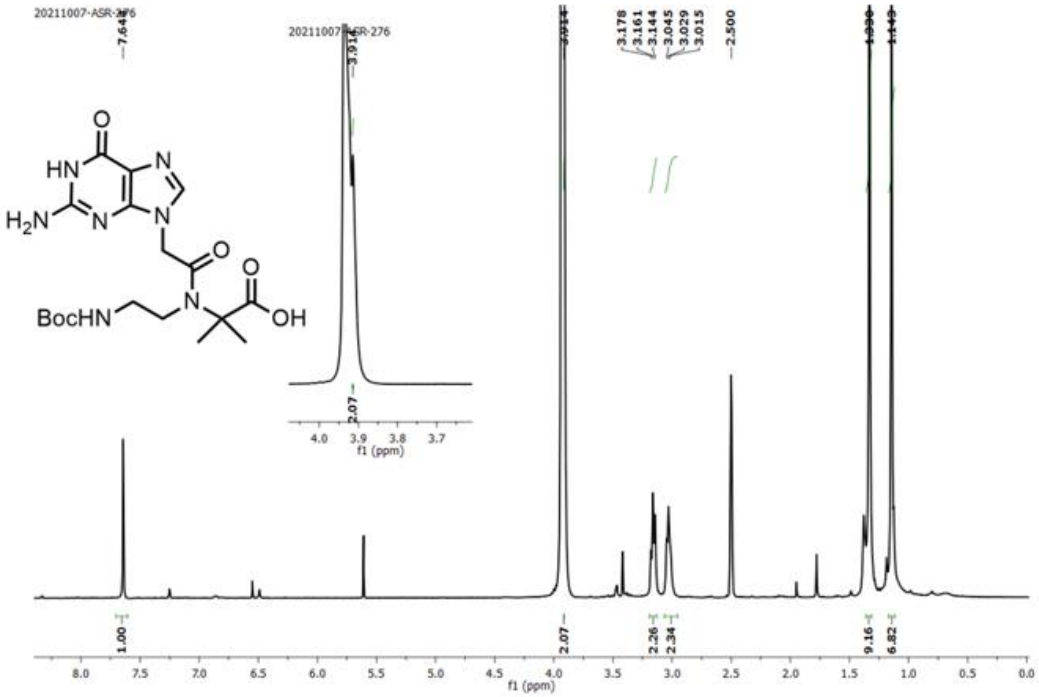
^{1}H NMR of compound 6 in $CDCl_3$ (400 MHz)



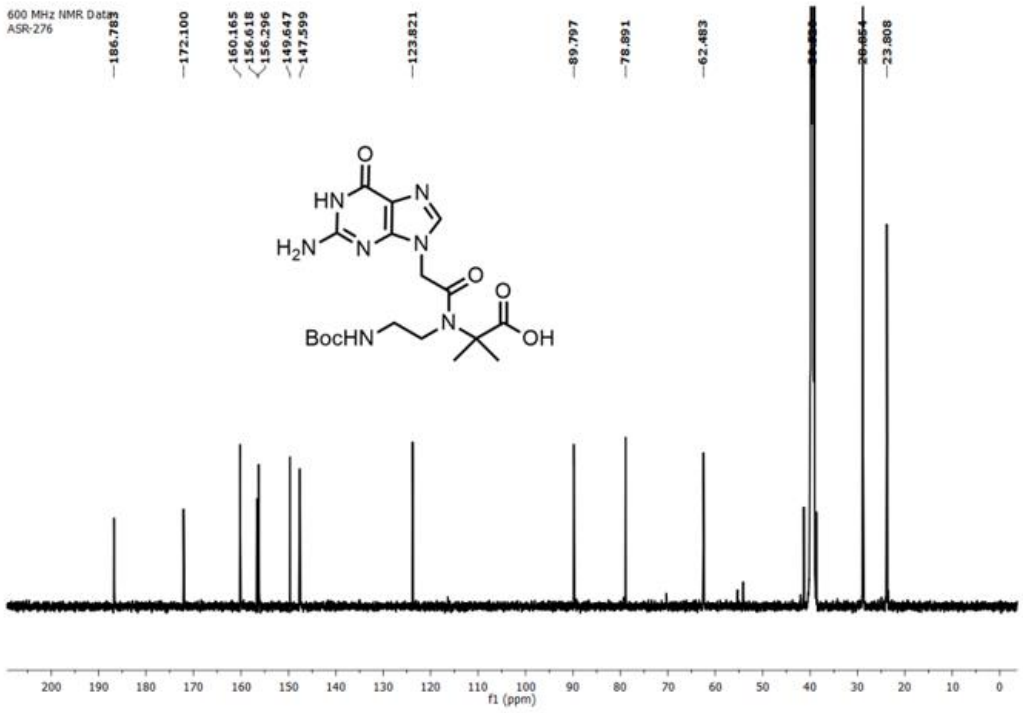
^{13}C {1H} NMR of compound 6 in $CDCl_3$ (101 MHz)



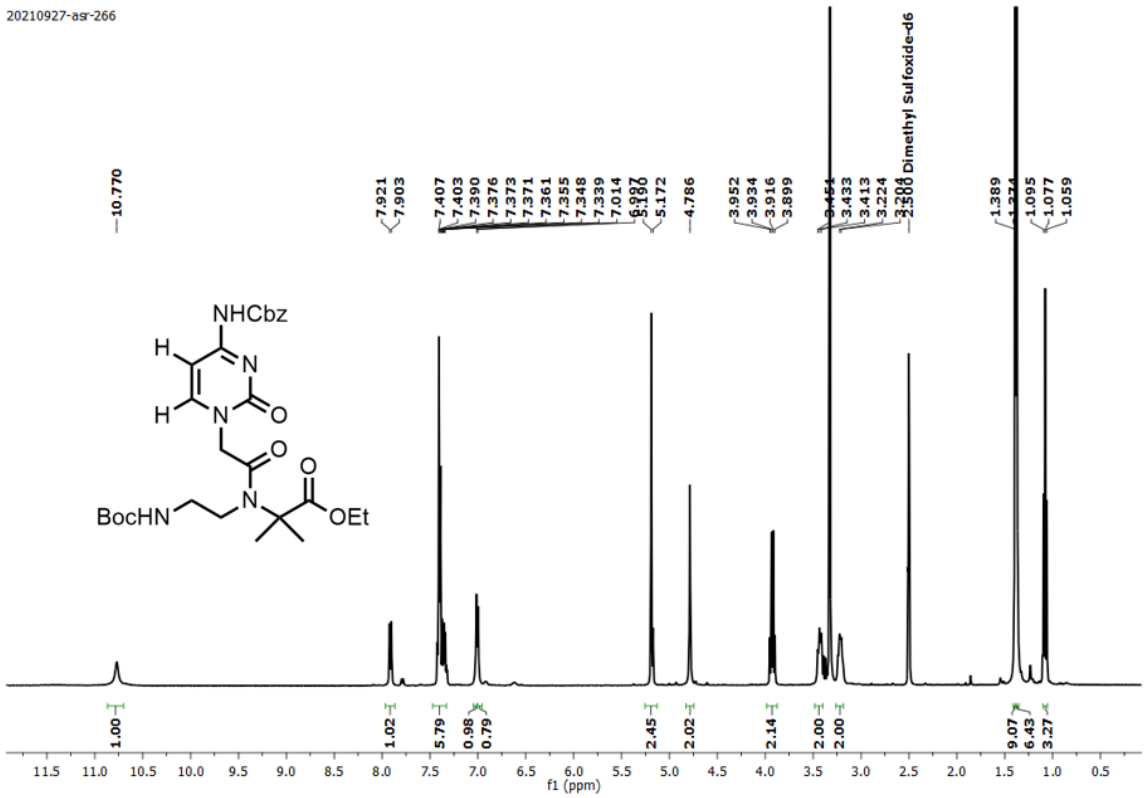
^{1}H NMR of compound 8 in DMSO-d_6 (400 MHz)



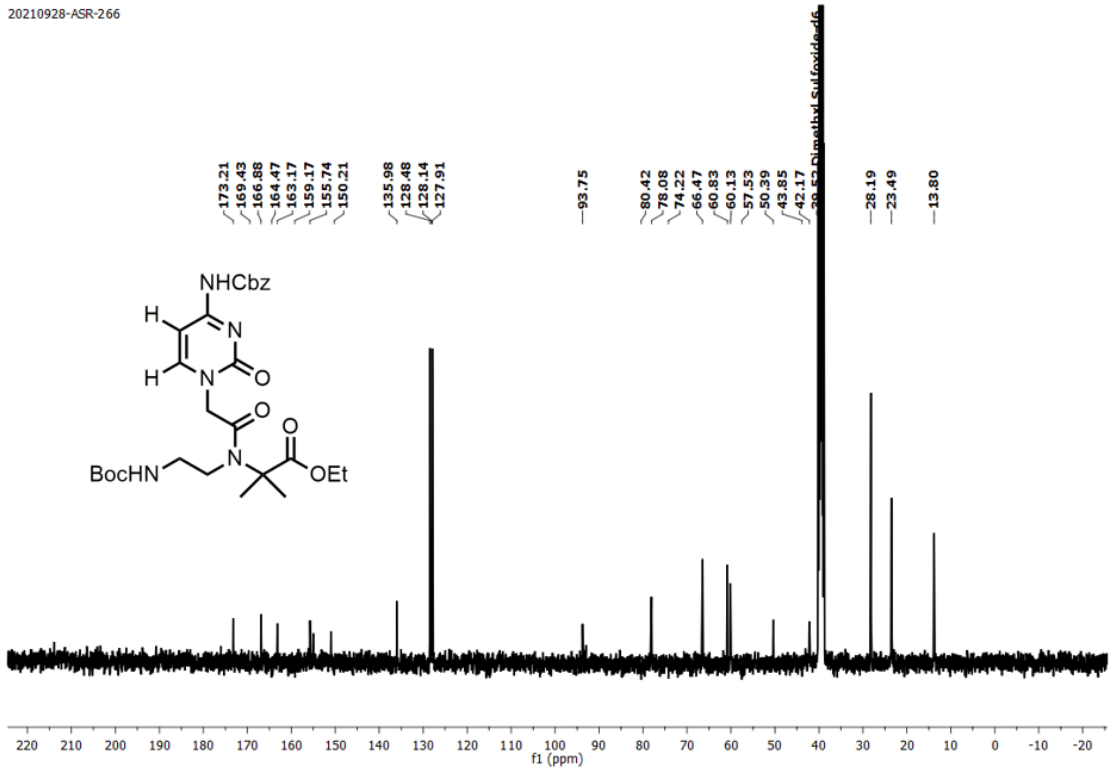
^{13}C {1H} NMR of compound 8 in DMSO-d_6 (151 MHz)



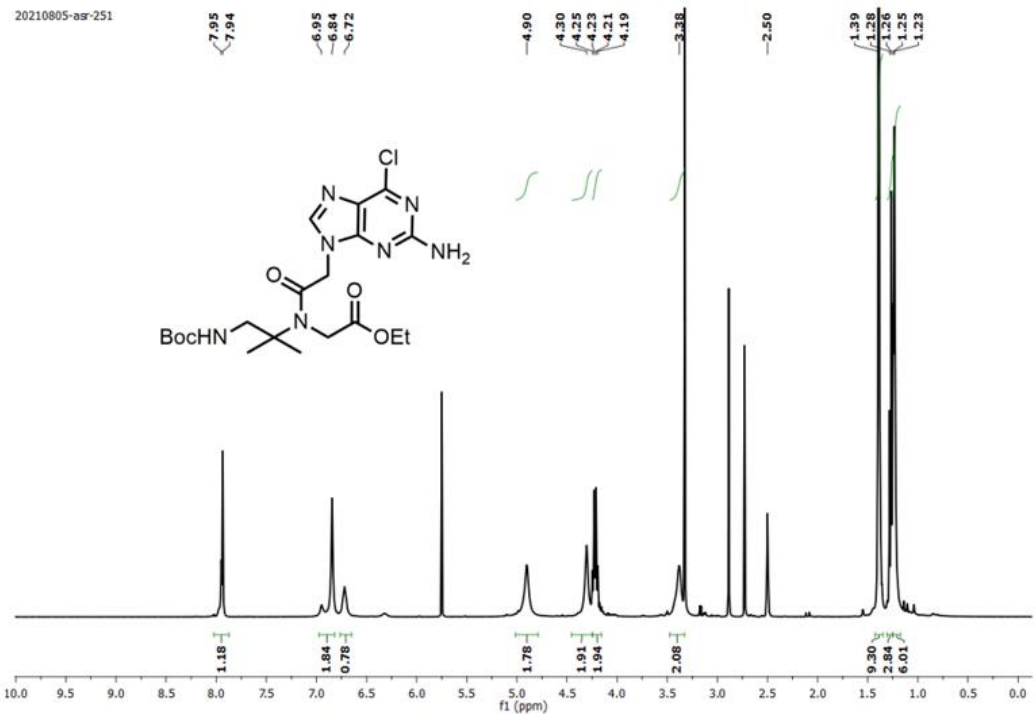
^{1}H NMR of compound 5 in DMSO-d$_6$ (400 MHz)



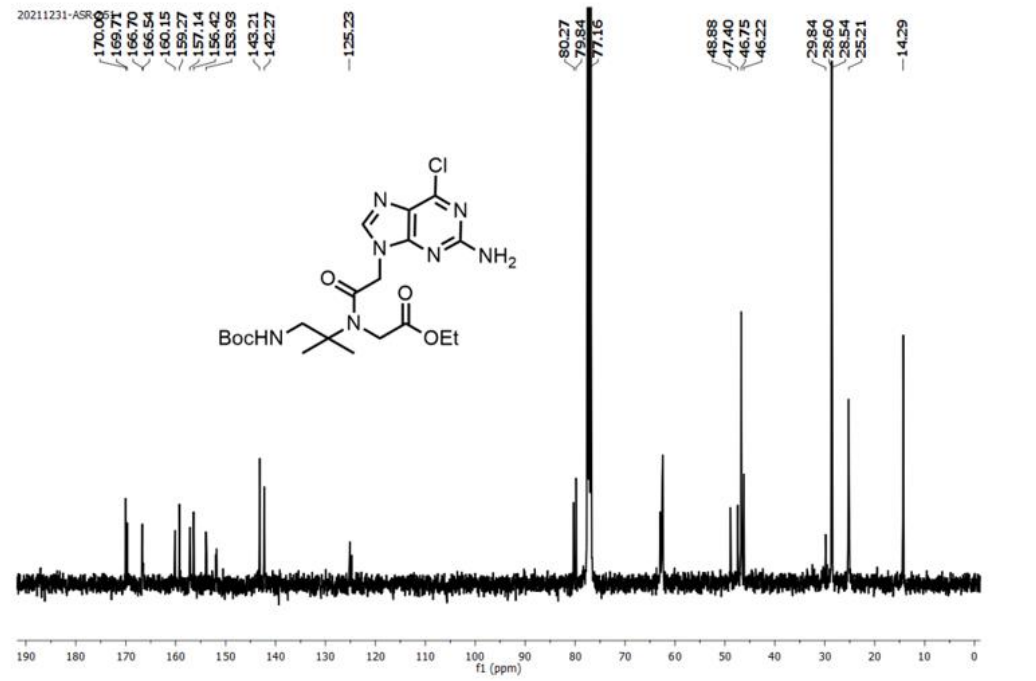
^{13}C {1H} NMR of compound 5 in DMSO-d$_6$ (101 MHz)



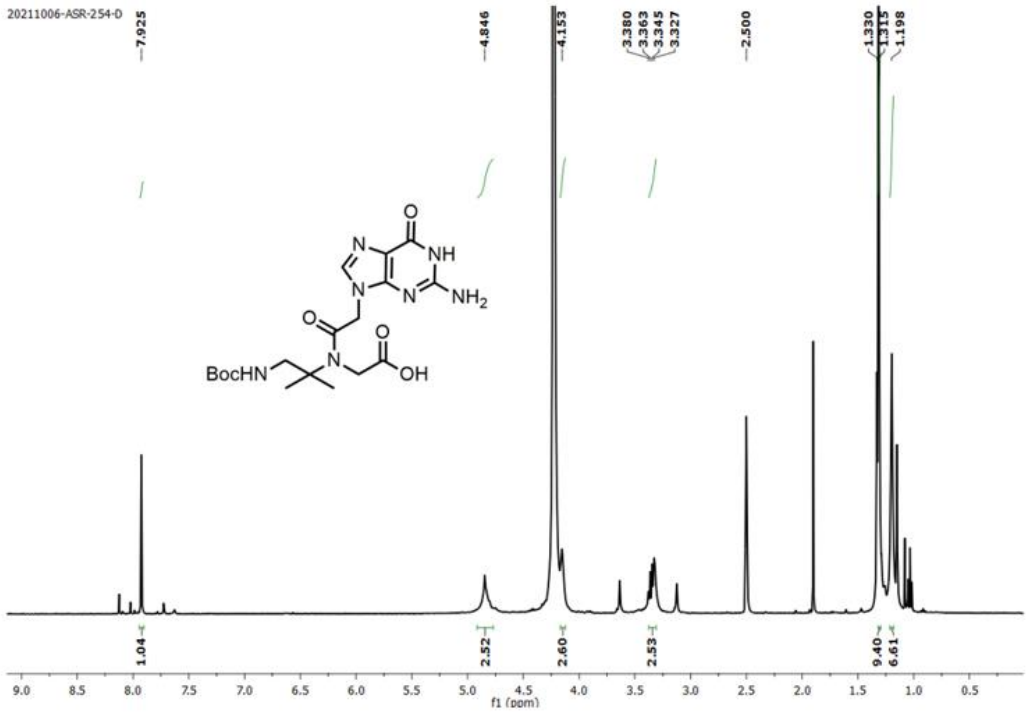
^{1}H NMR of compound 12 in DMSO-d$_6$ (400 MHz)



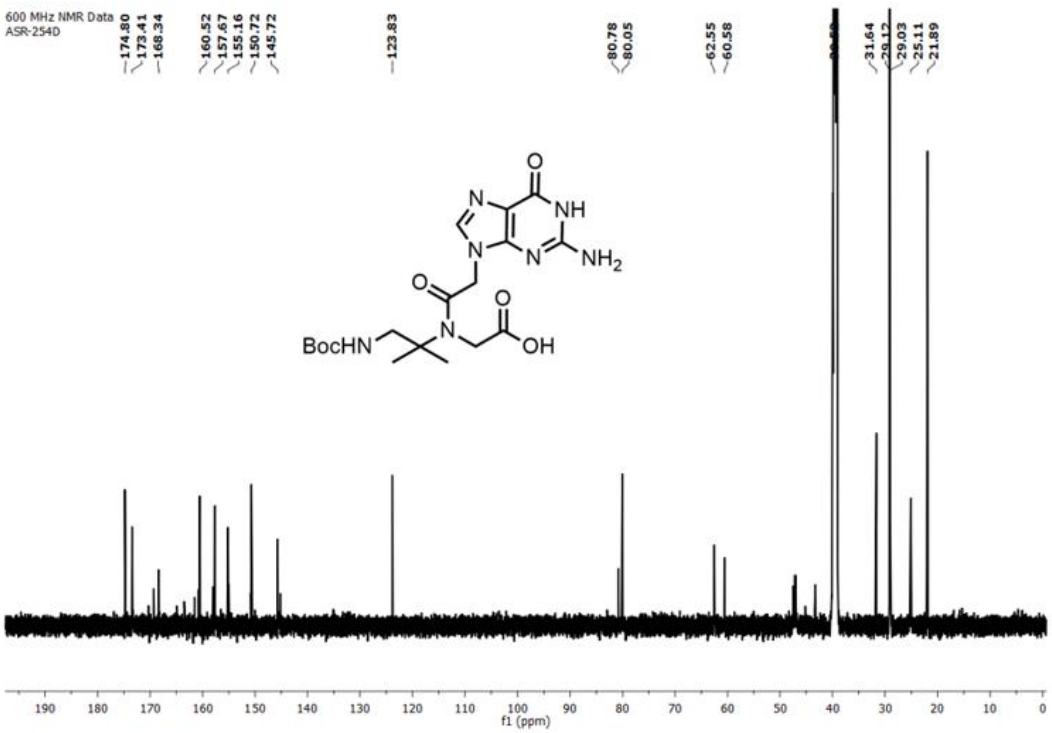
^{13}C {1H} NMR of compound 12 in CDCl$_3$ (101 MHz)



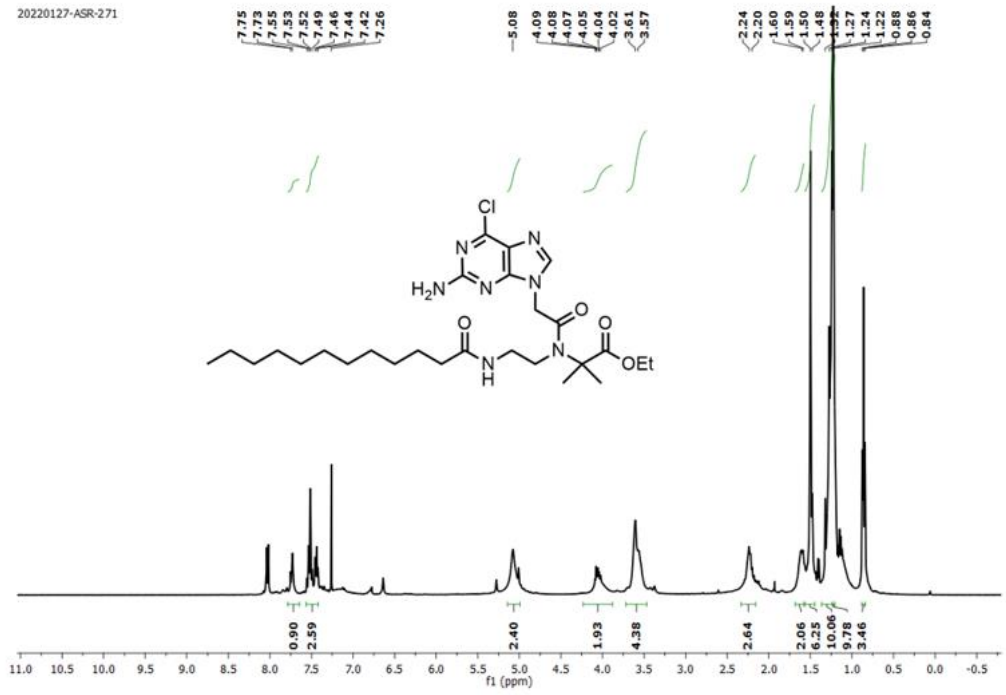
^{1}H NMR of compound 14 in DMSO-d$_6$ (400 MHz)



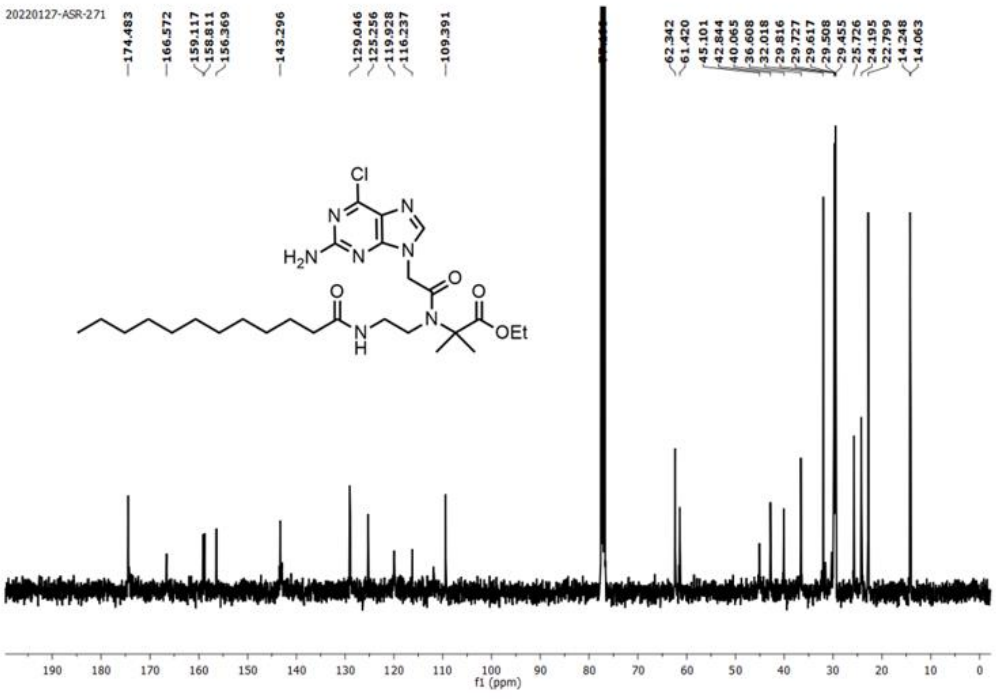
^{13}C {1H} NMR of compound 14 in DMSO-d$_6$ (101 MHz)



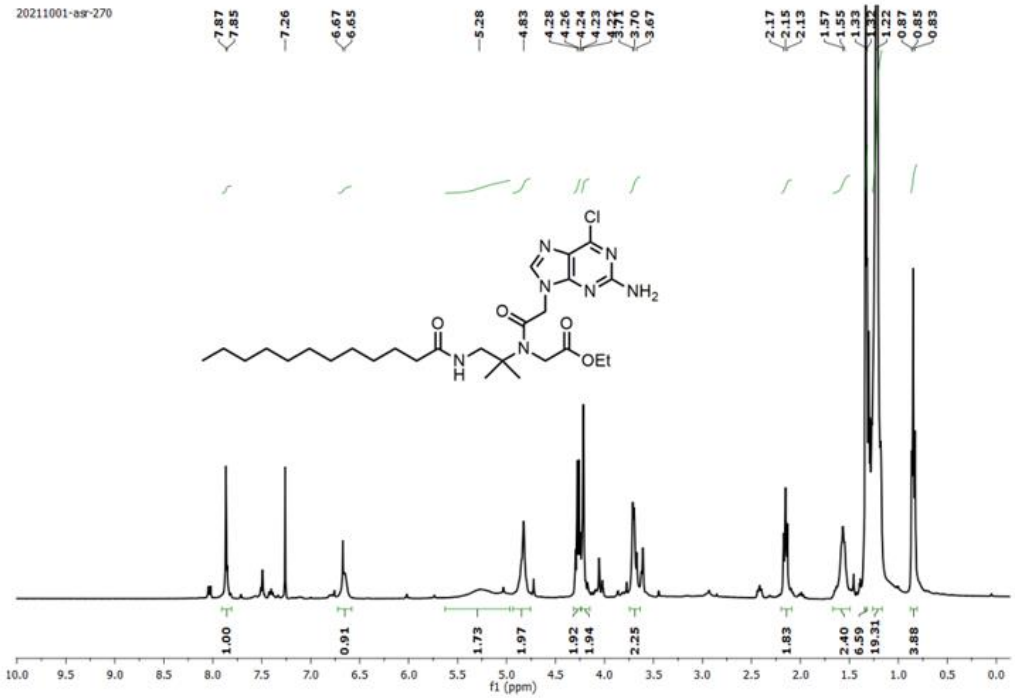
^{1}H NMR of compound 19 in $CDCl_3$ (400 MHz)



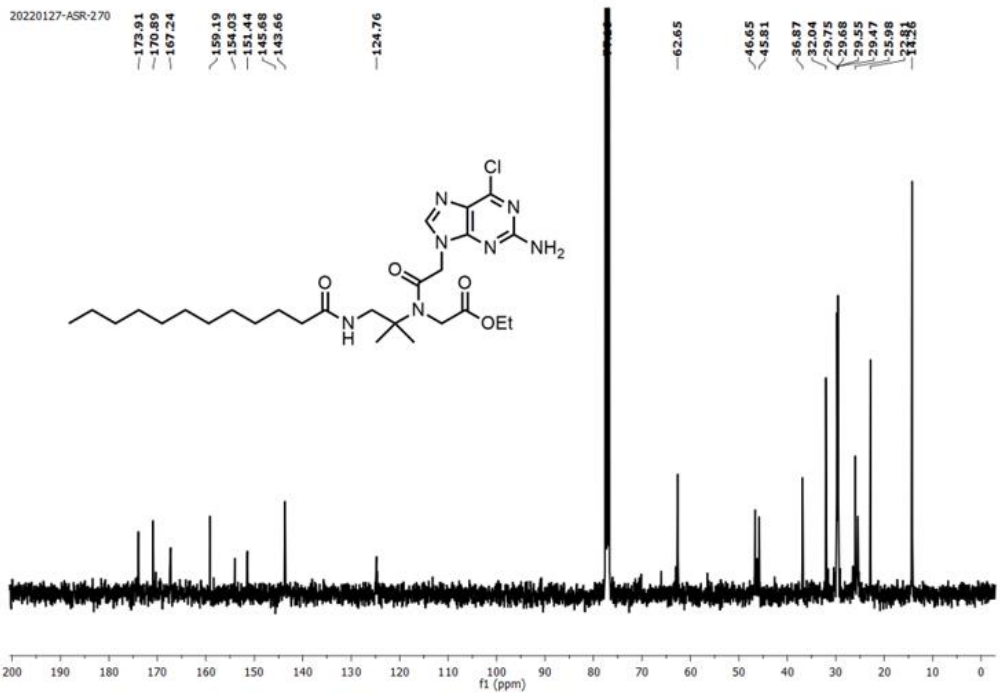
^{13}C {1H} NMR of compound 19 in $CDCl_3$ (101 MHz)



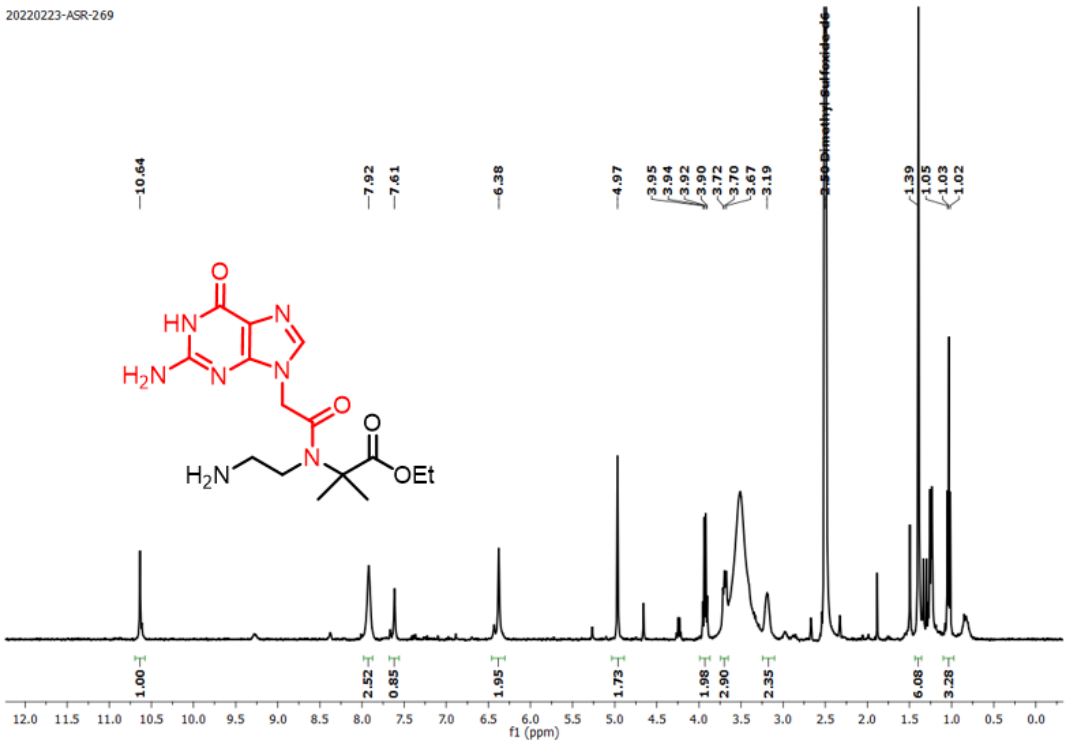
^{1}H NMR of compound 16 in $CDCl_3$ (400 MHz)



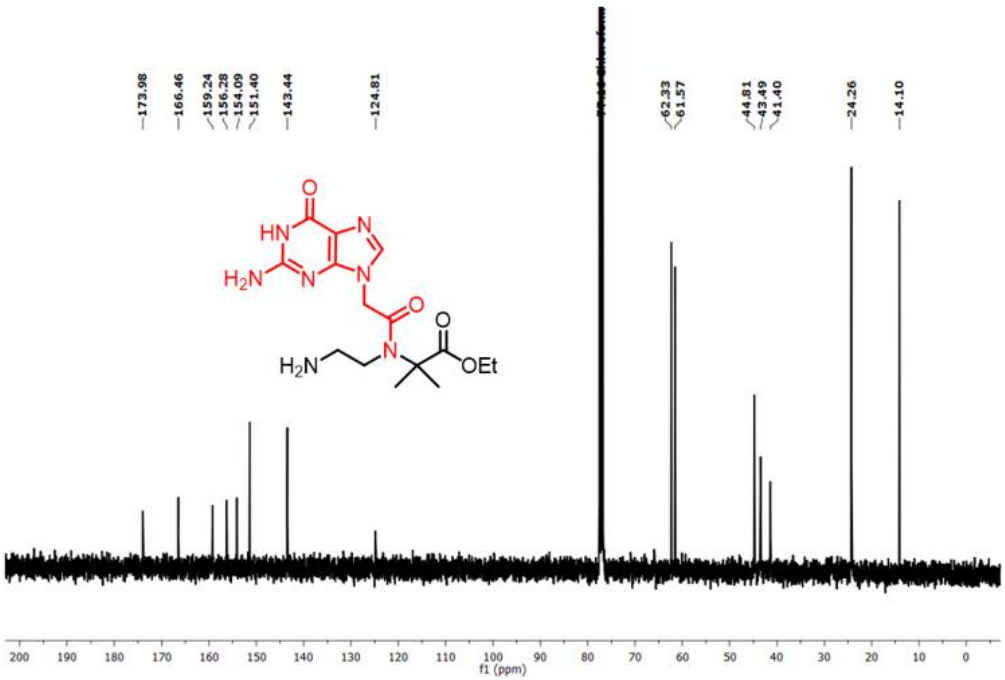
^{13}C {1H} NMR of compound 16 in $CDCl_3$ (101 MHz)



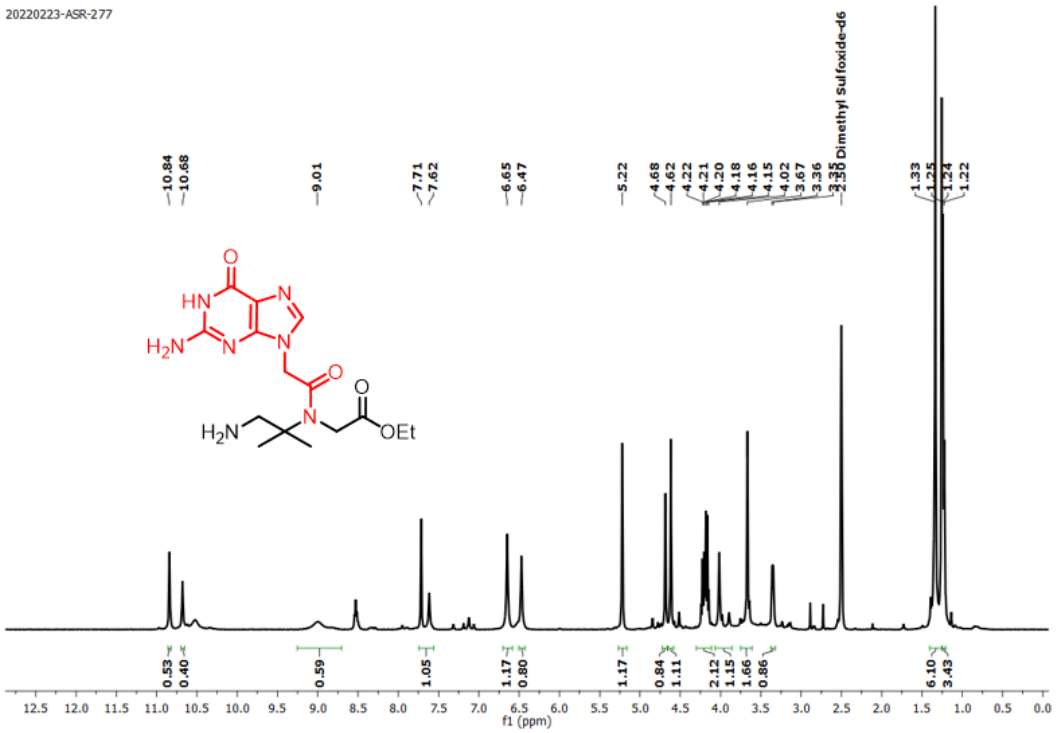
^{1}H NMR of compound 17 in DMSO-d$_6$ (400 MHz)



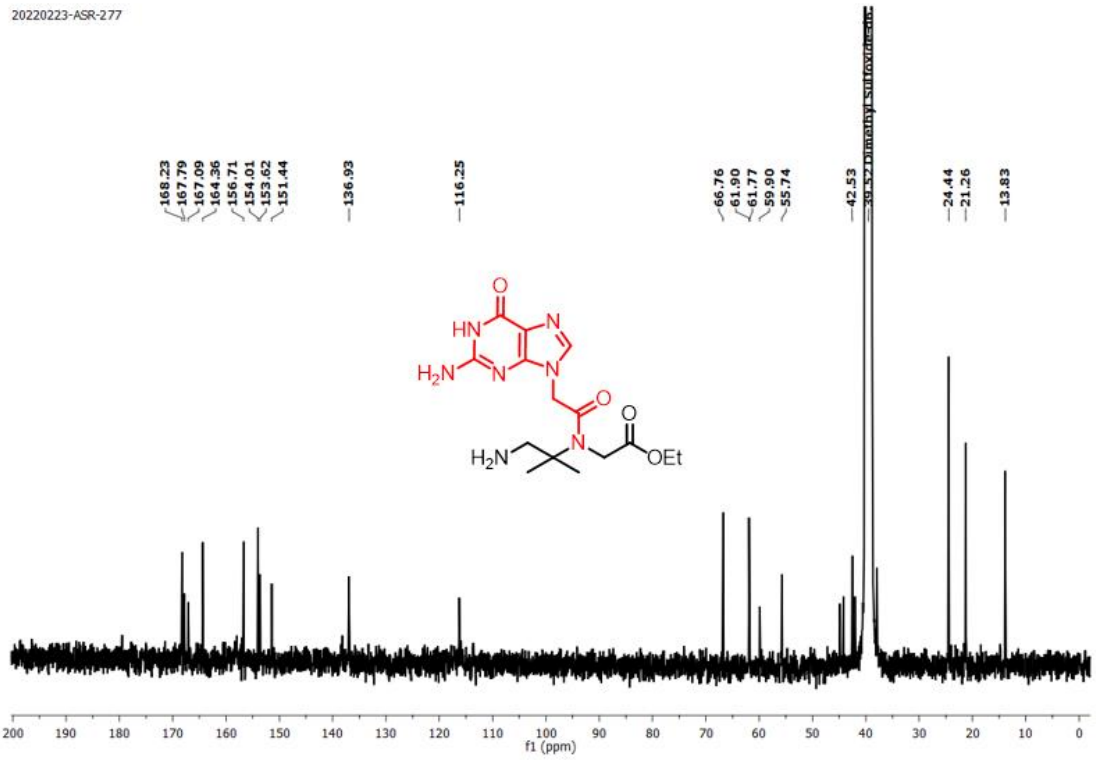
^{13}C {1H} NMR of compound 17 in CDCl$_3$ (101 MHz)



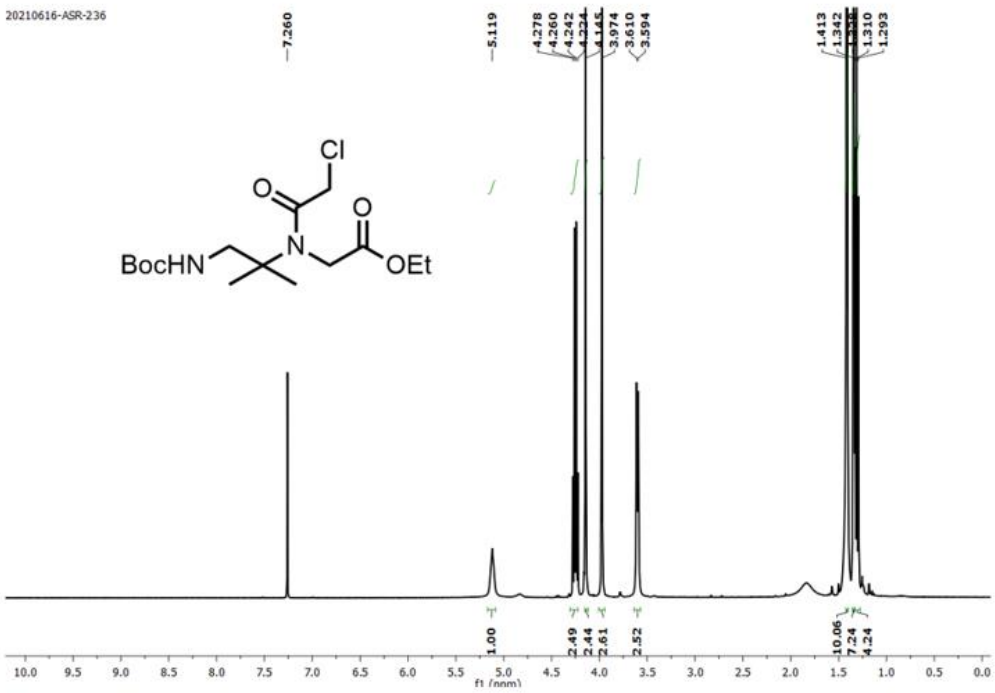
^{1}H NMR of compound 13 in DMSO-d$_6$ (400 MHz)



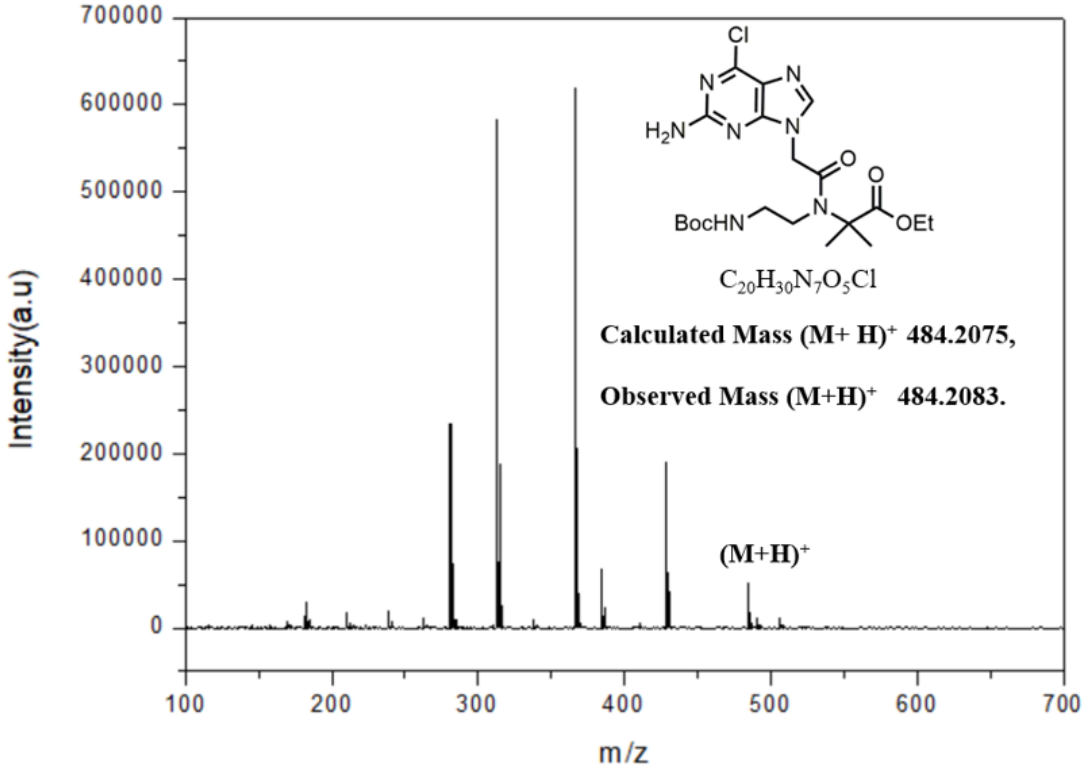
^{13}C {1H} NMR of compound 13 in DMSO-d$_6$ (101 MHz)



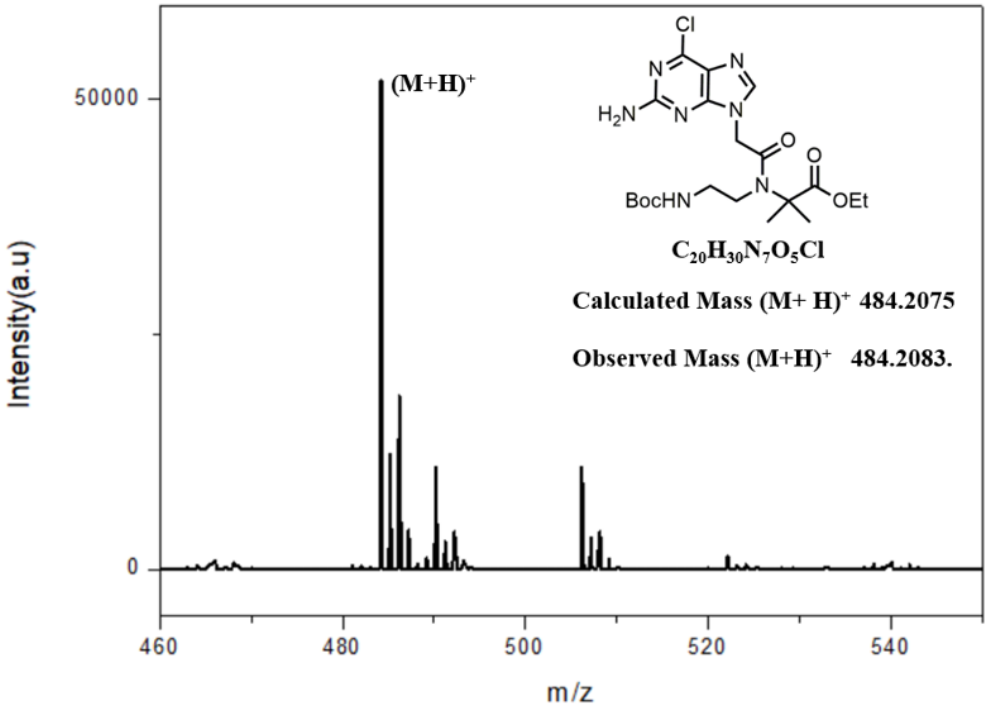
^{1}H NMR of compound 11 in $CDCl_3$ (400 MHz)



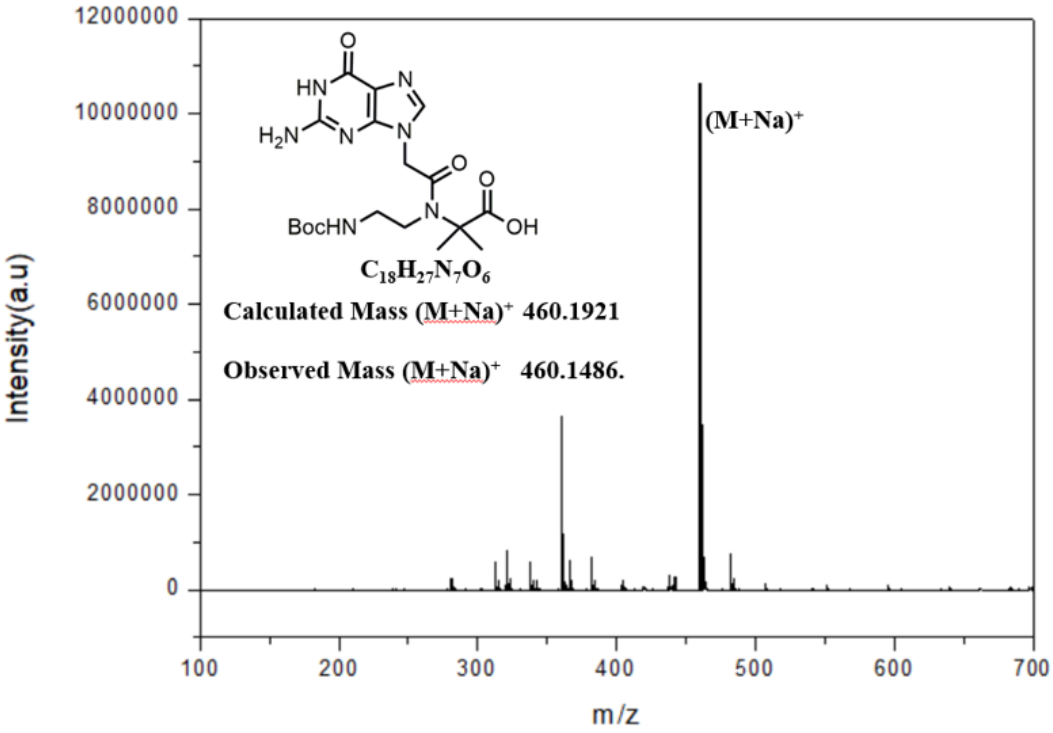
HRMS of compound 6 (Full spectra)



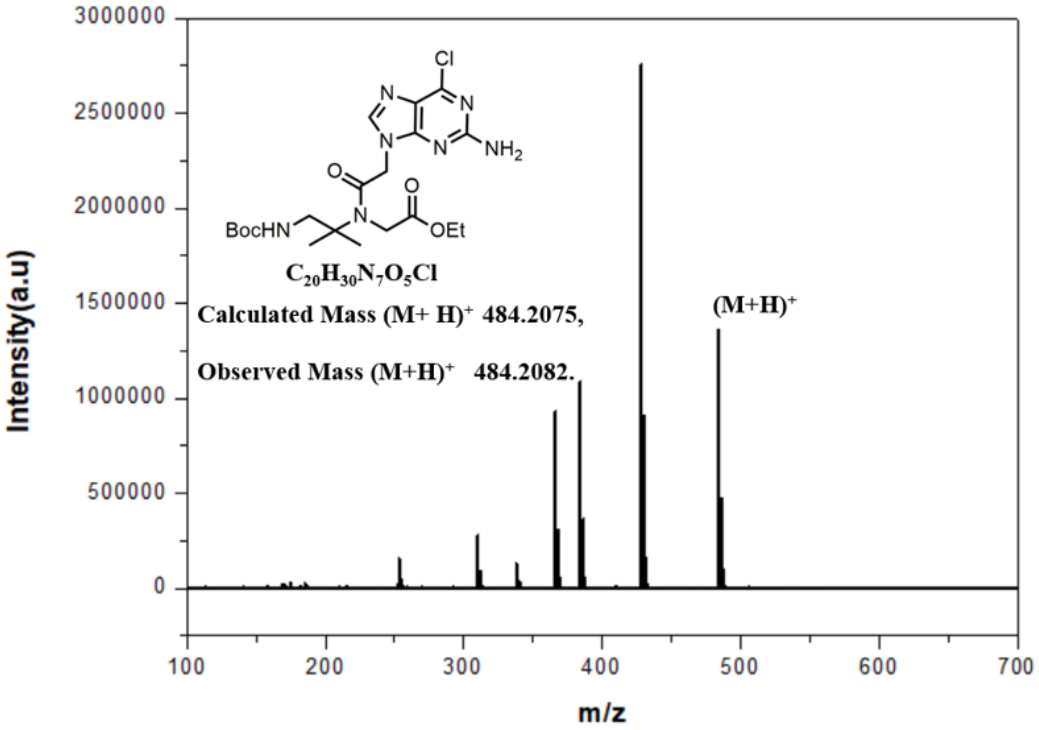
HRMS of compound 6 (enlarged)



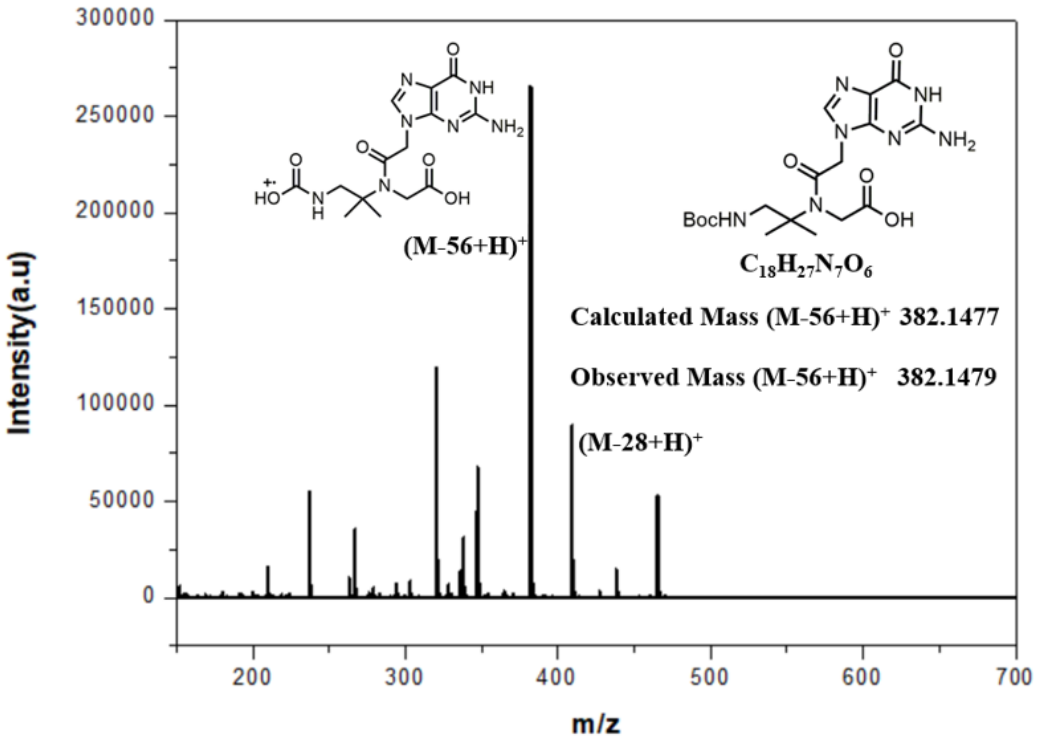
HRMS of compound 8



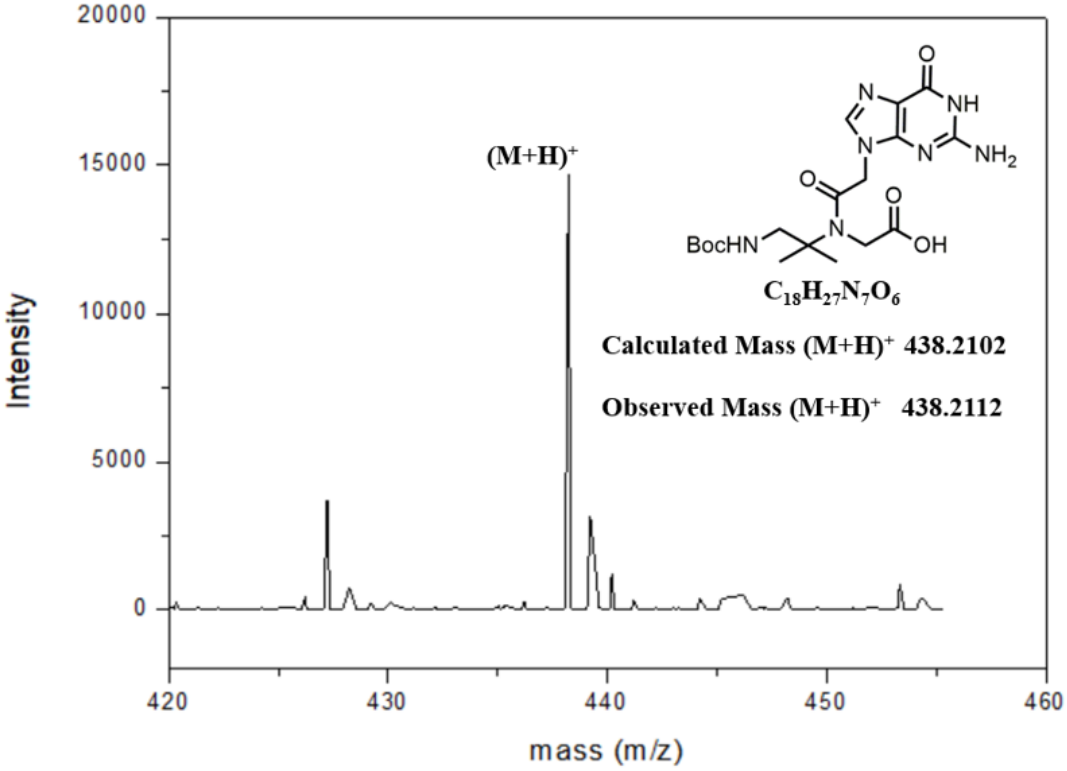
HRMS of compound 12



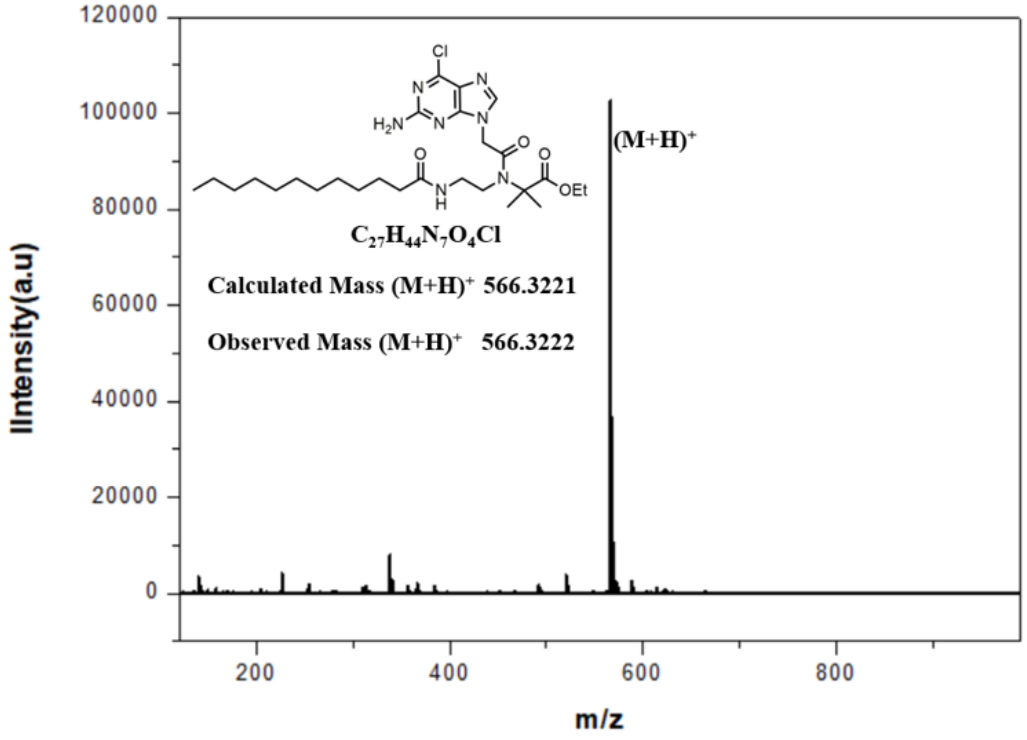
HRMS of compound 14 (Full spectra)



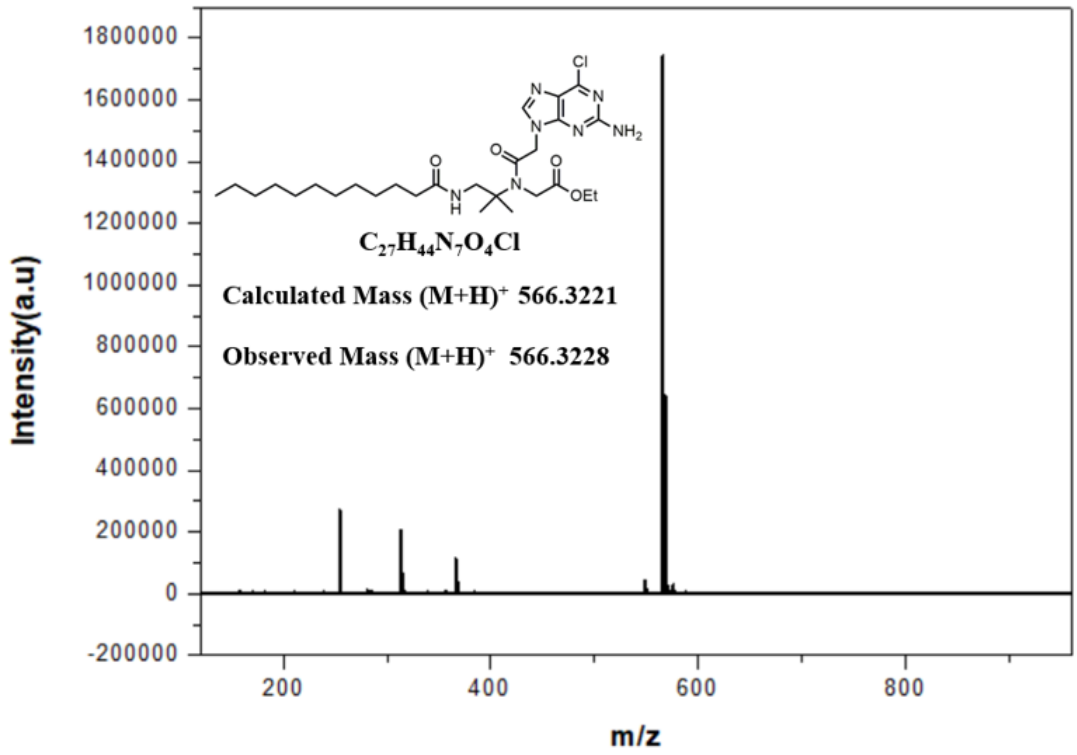
HRMS of compound 14 (Enlarged)



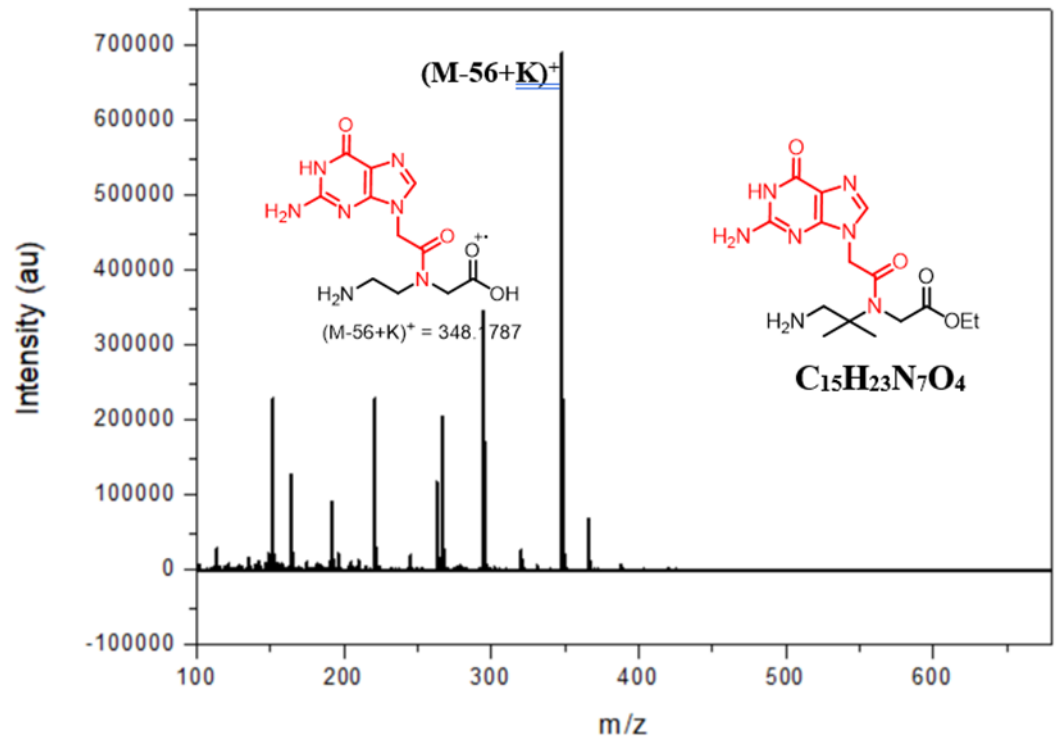
HRMS of compound 19



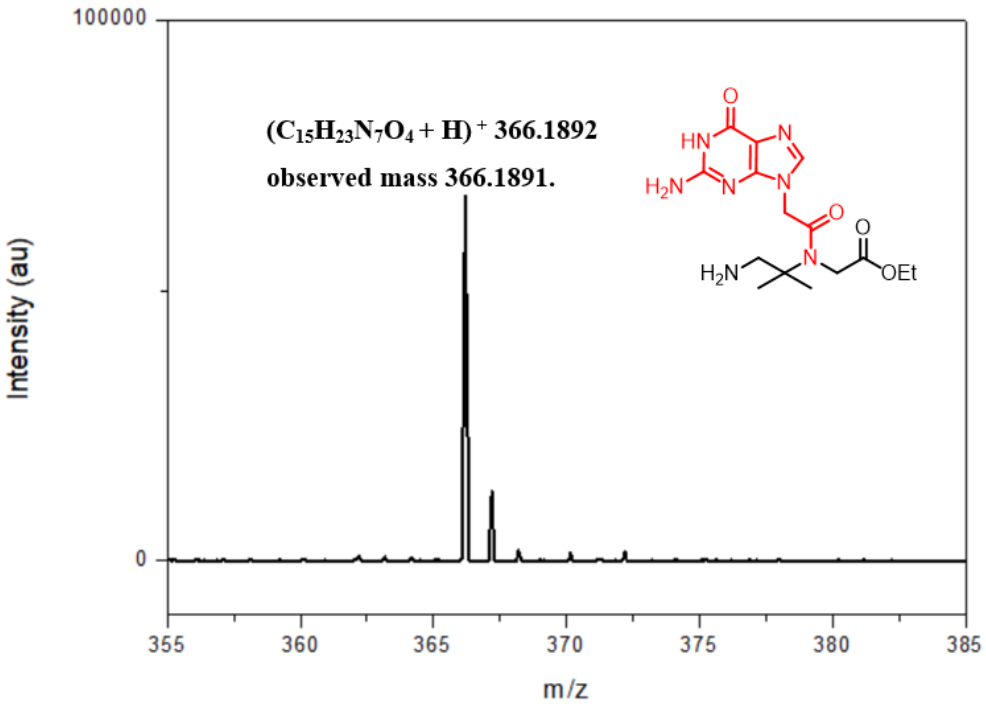
HRMS of compound 16



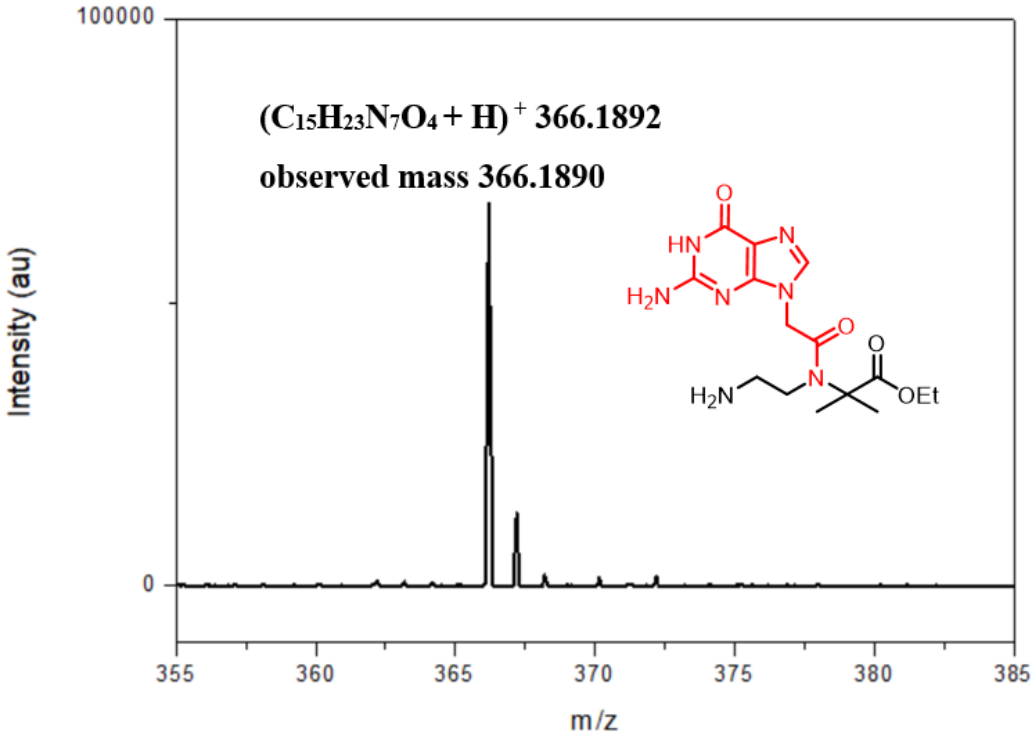
HRMS of compound 13 (Full spectra)



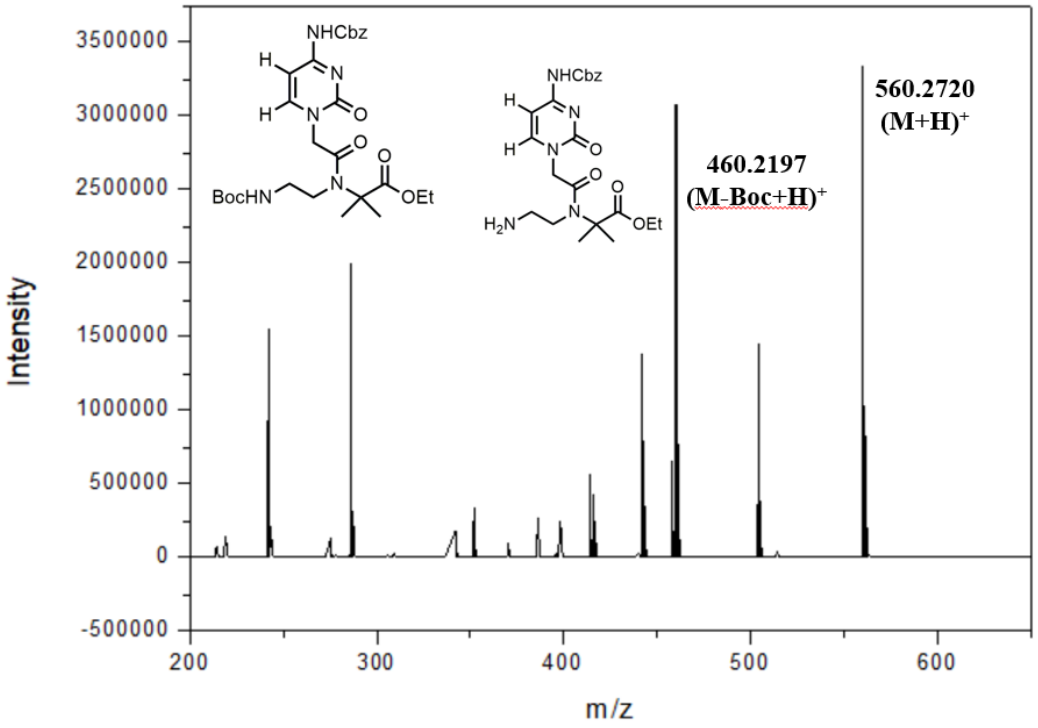
HRMS of compound 13 (enlarged)



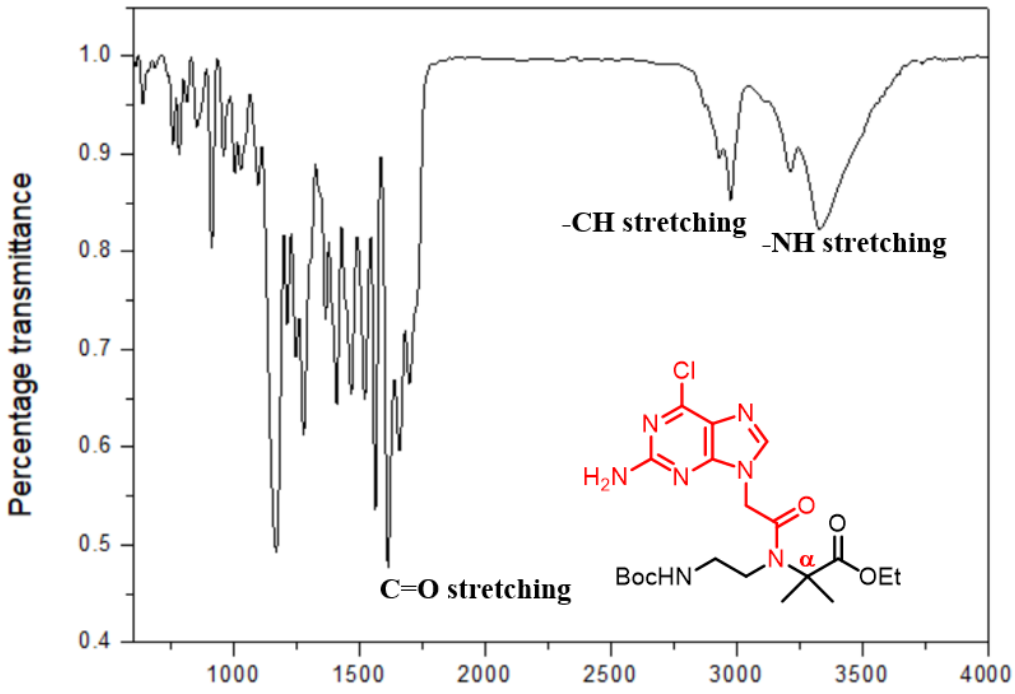
HRMS of compound 17



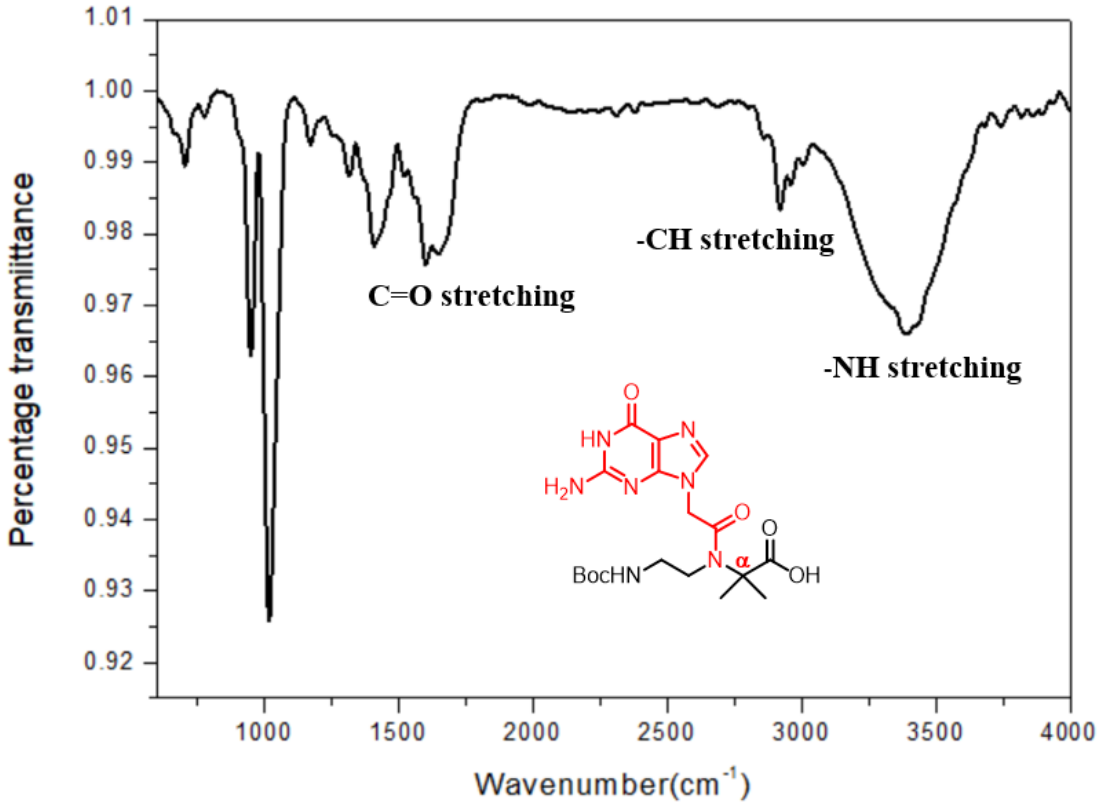
HRMS of compound 5



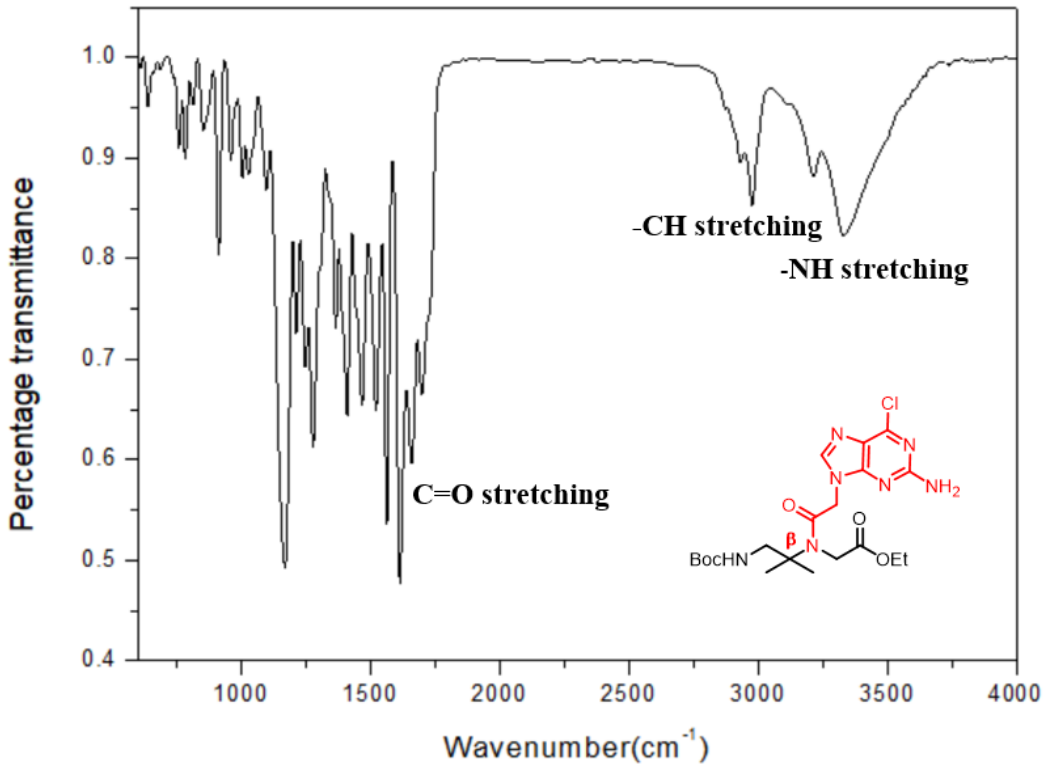
IR Spectra of compound 6 in Chloroform



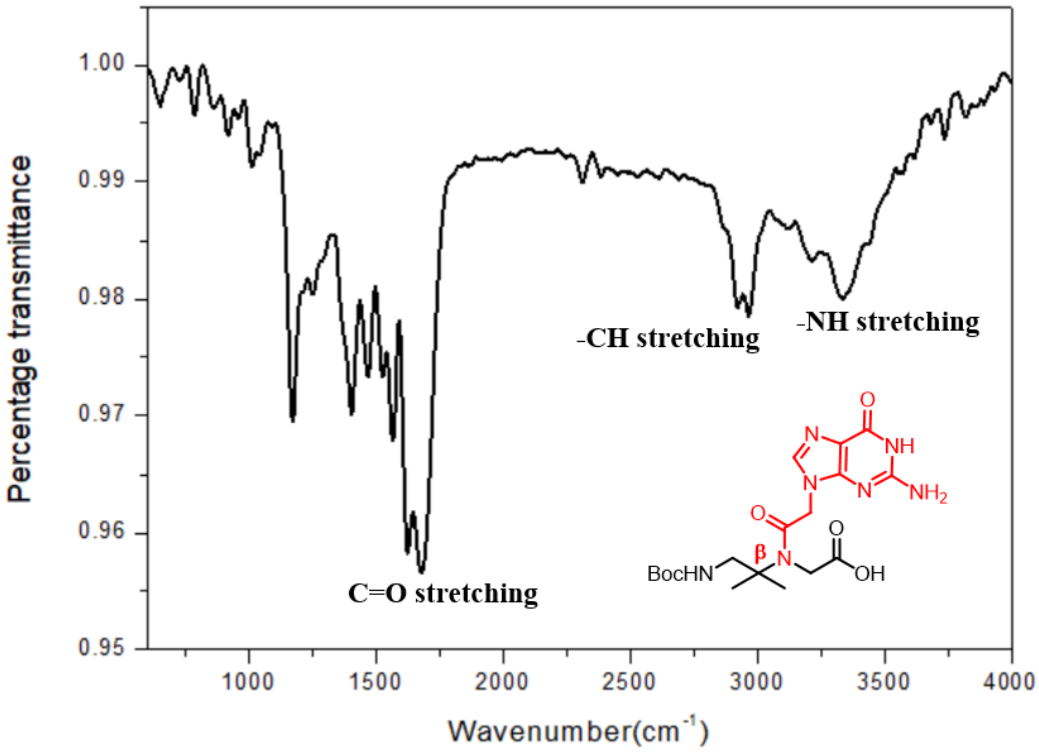
IR Spectra of compound 8 in Chloroform



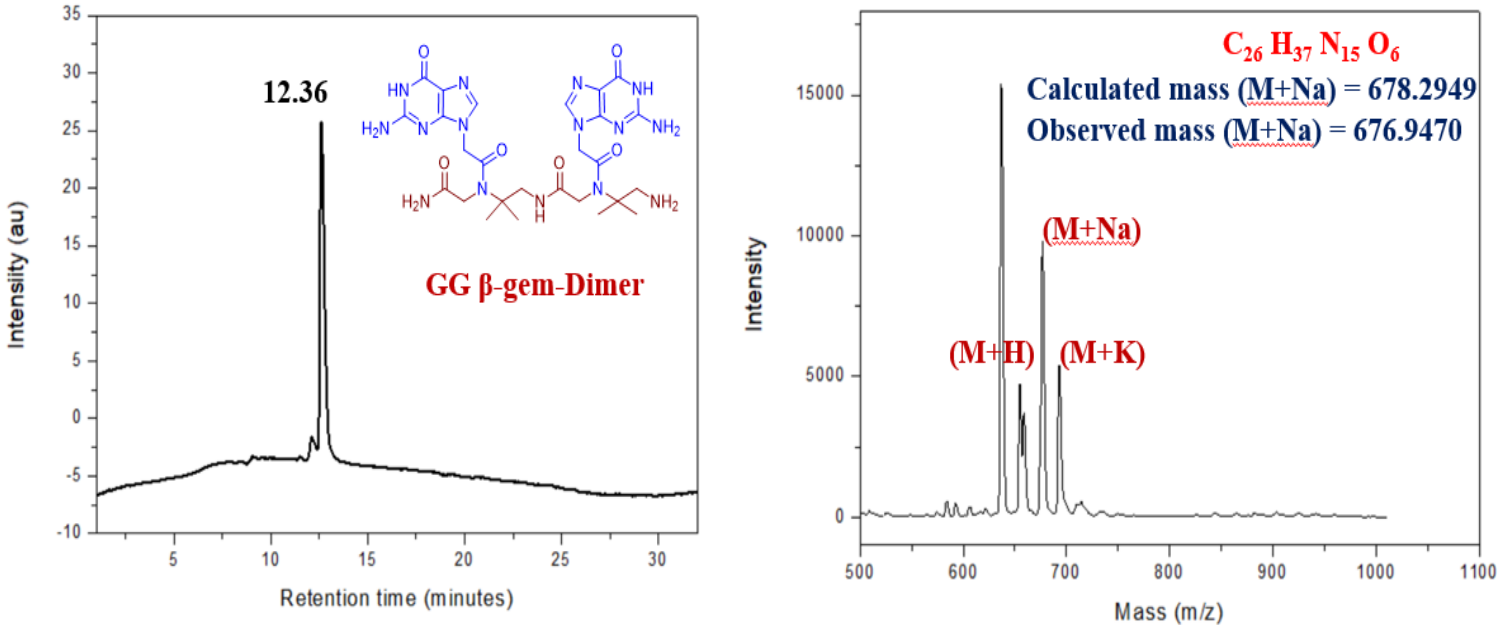
IR Spectra of compound 12 in Chloroform



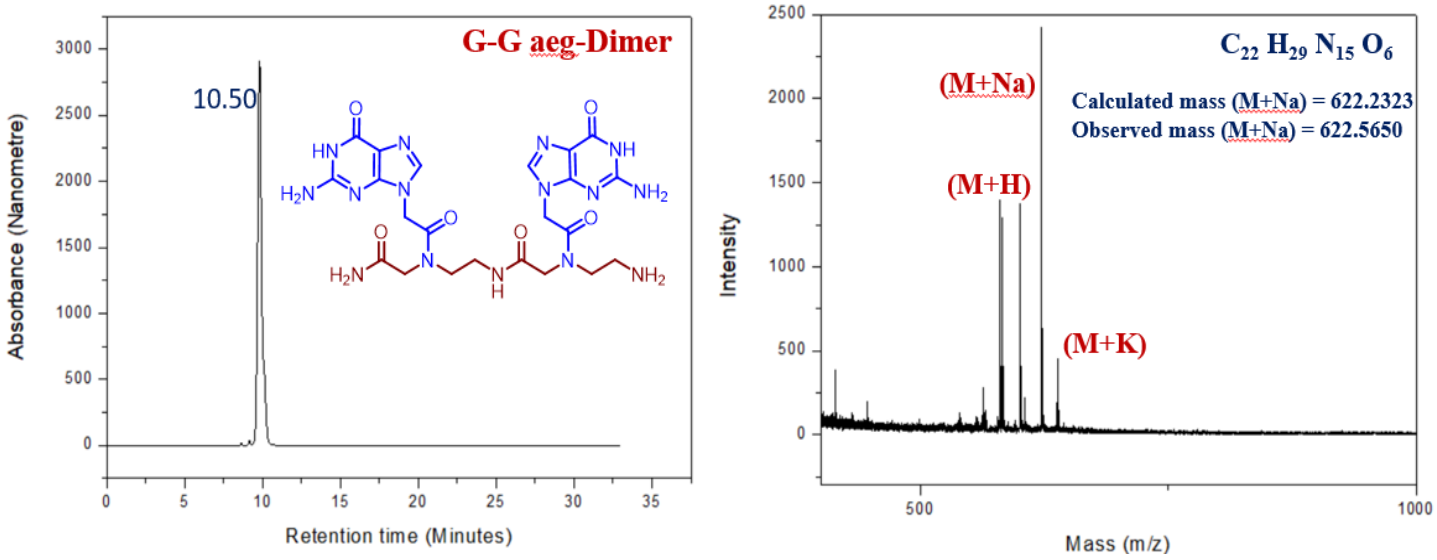
IR Spectra of compound 14 in Chloroform



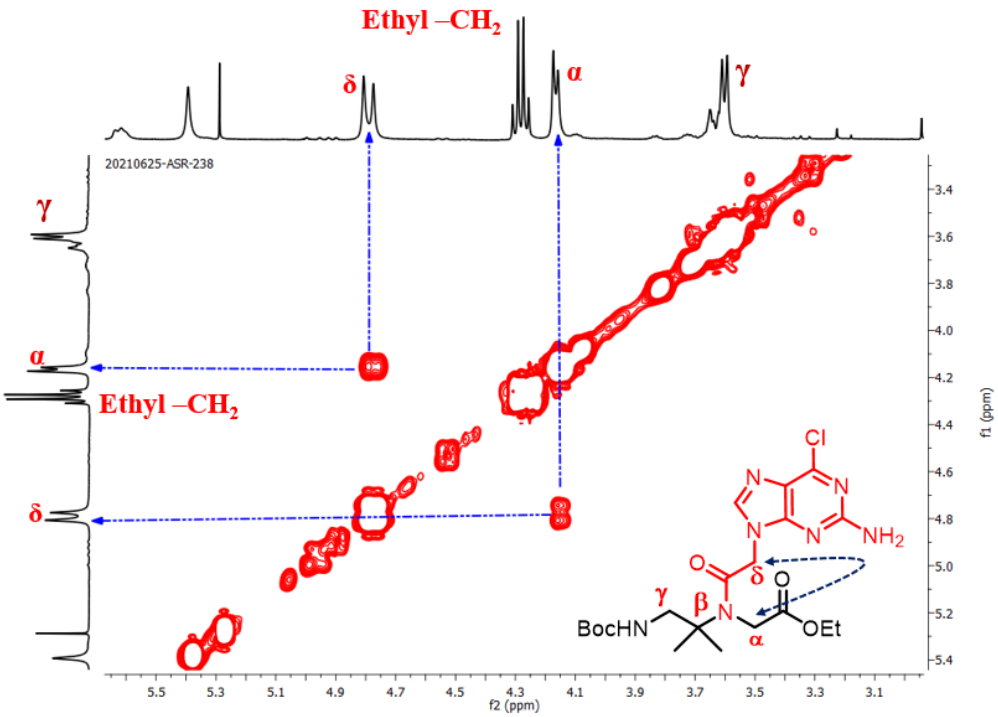
HPLC and MALDI of (β-*gemdimethyl-aeg*-G)$_2$ PNA dimer (22) and (*aeg*-G)$_2$ PNA dimer (23)



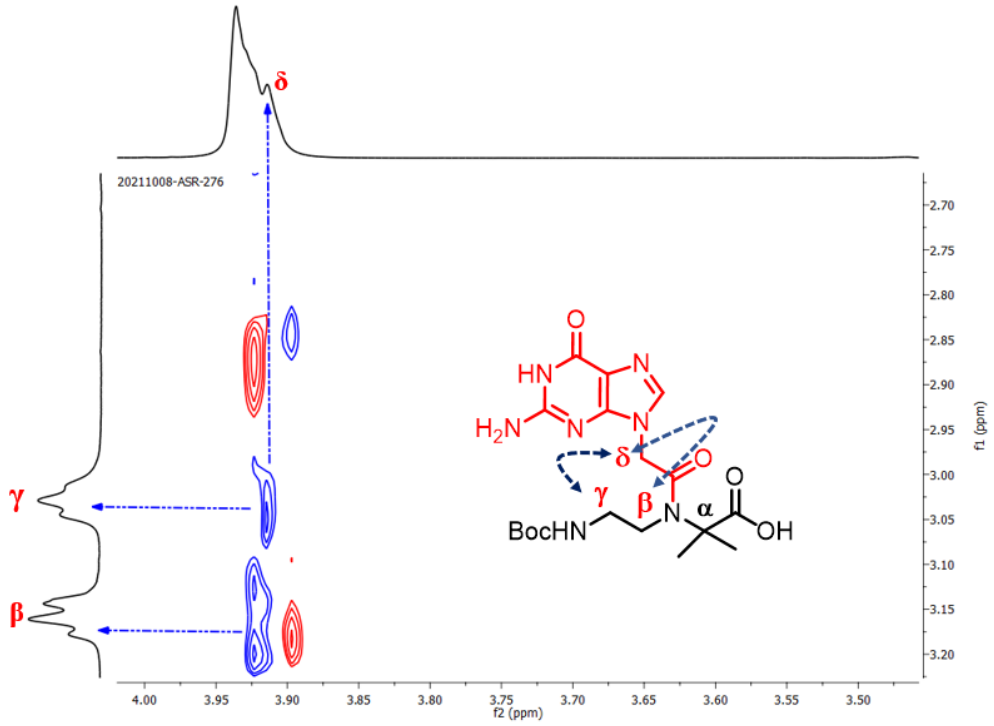
HPLC and MALDI of (β-*gemdimethyl-aeg*-G)$_2$ PNA dimer (**22**)



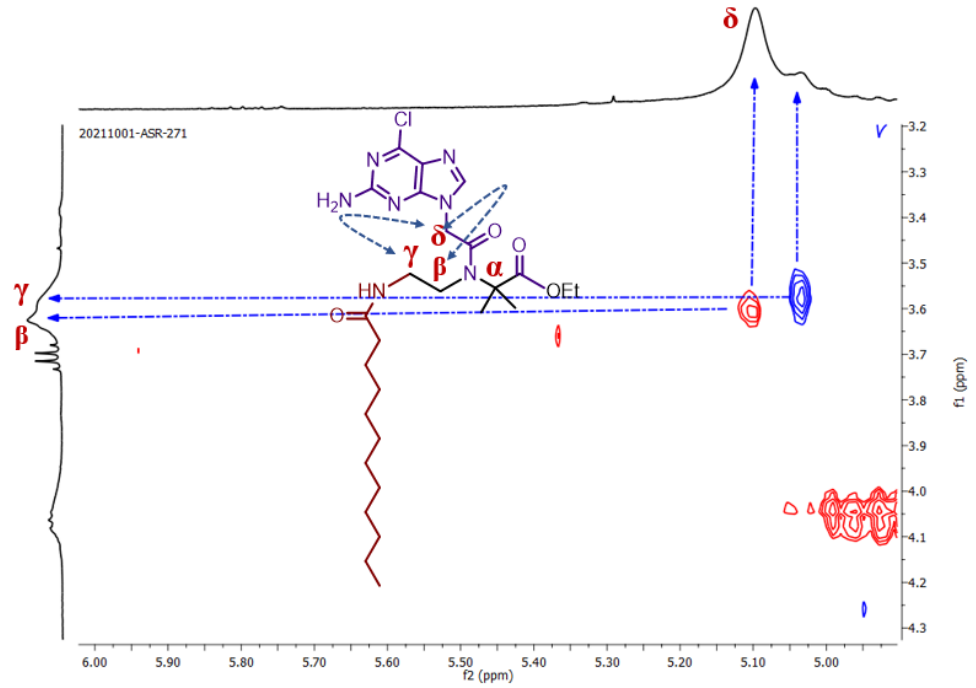
HPLC and MALDI of (*aeg*-G)$_2$ PNA dimer (**23**)



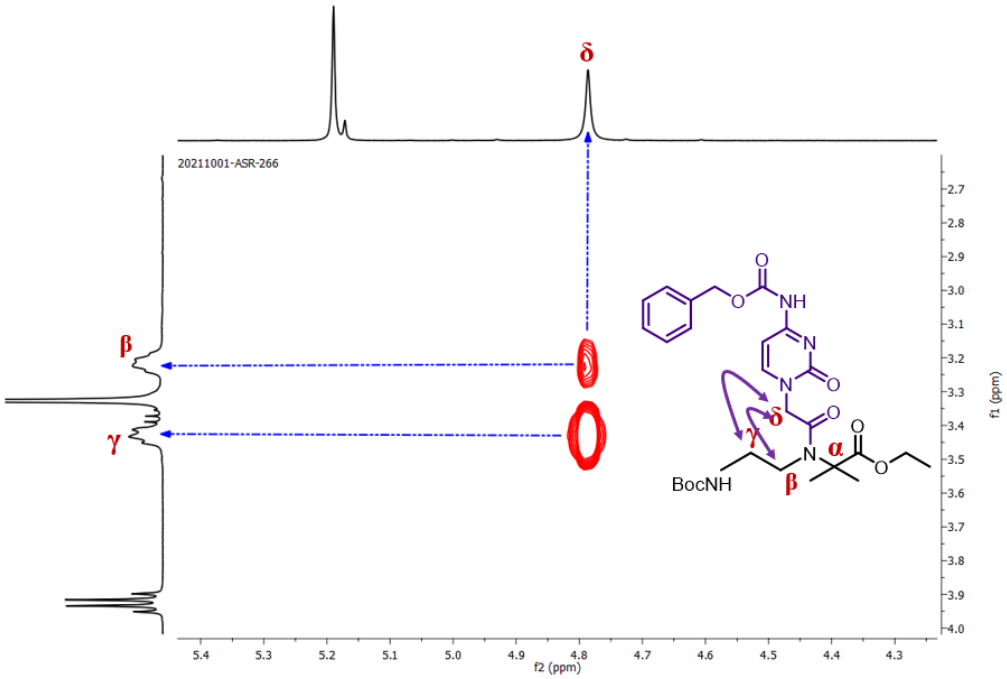
Partial ^{1}H-^{1}H 2D NOSEY spectra of β- *gemdimethyl*- (6-Cl-2-amino) purine PNA **(12)** in $CDCl_3$ (100% E rotamer). δ to α connectivity indicate 100% E-rotamer in solution.



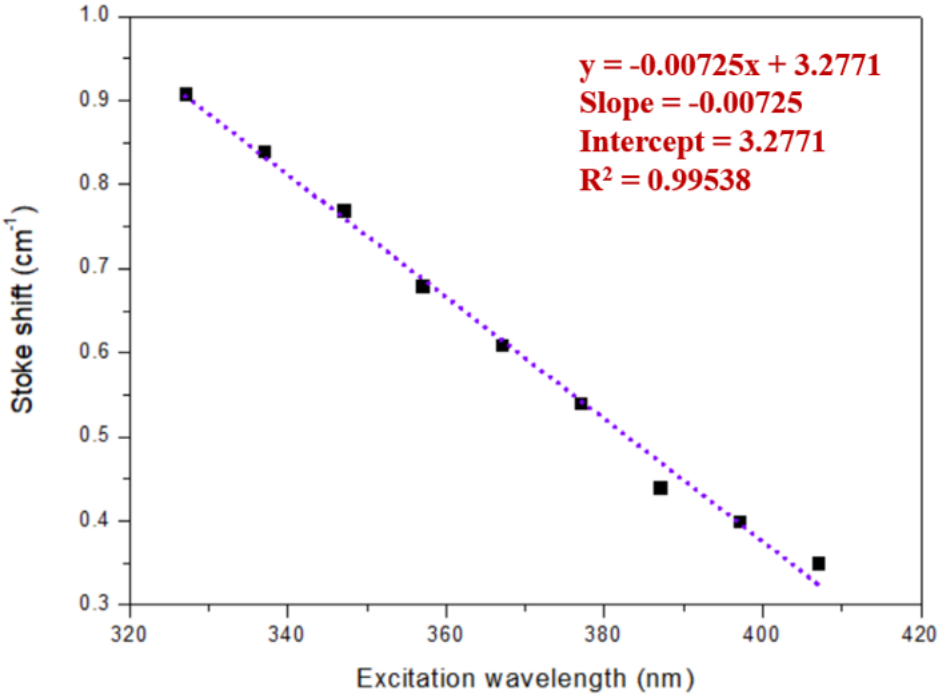
Partial ^{1}H-^{1}H 2D NOSEY spectra of α-*gemdimethyl-aeg*-PNA-G **(8)** in DMSO-d_6 (100% Z rotamer). δ to β/γ connectivity indicate 100% Z-rotamer in solution.



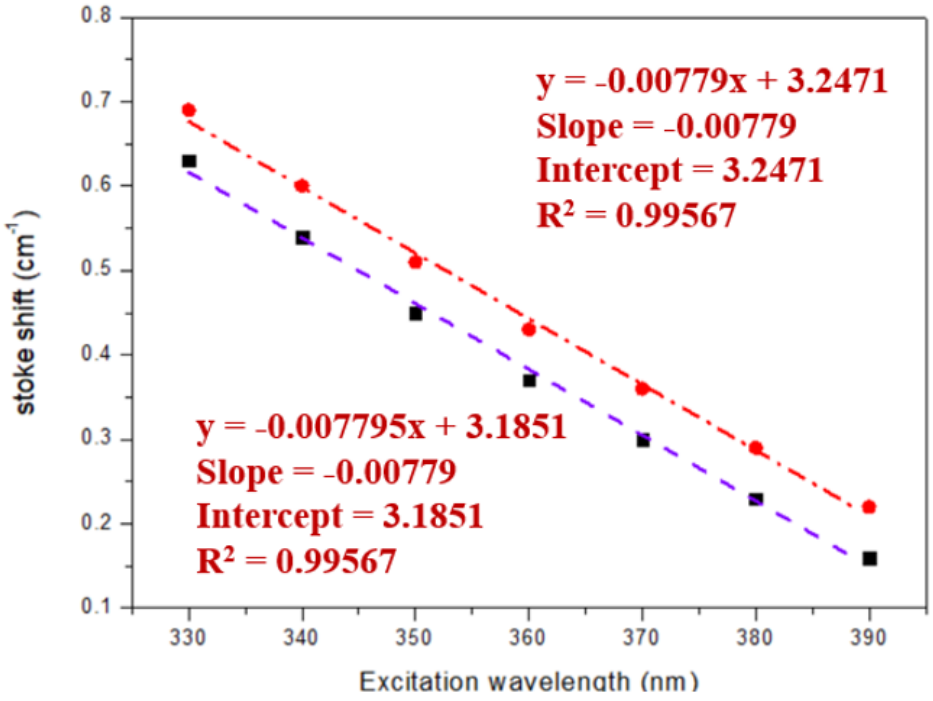
Partial ^{1}H-^{1}H 2D NOSEY spectra of N-decanoyl α, α-*dimethyl-aeg*-(6-Cl-2-aminopurine)-PNA monomer **(19)** in $CDCl_3$ (100% Z rotamer). δ to β/γ connectivity indicate 100% Z-rotamer in solution.



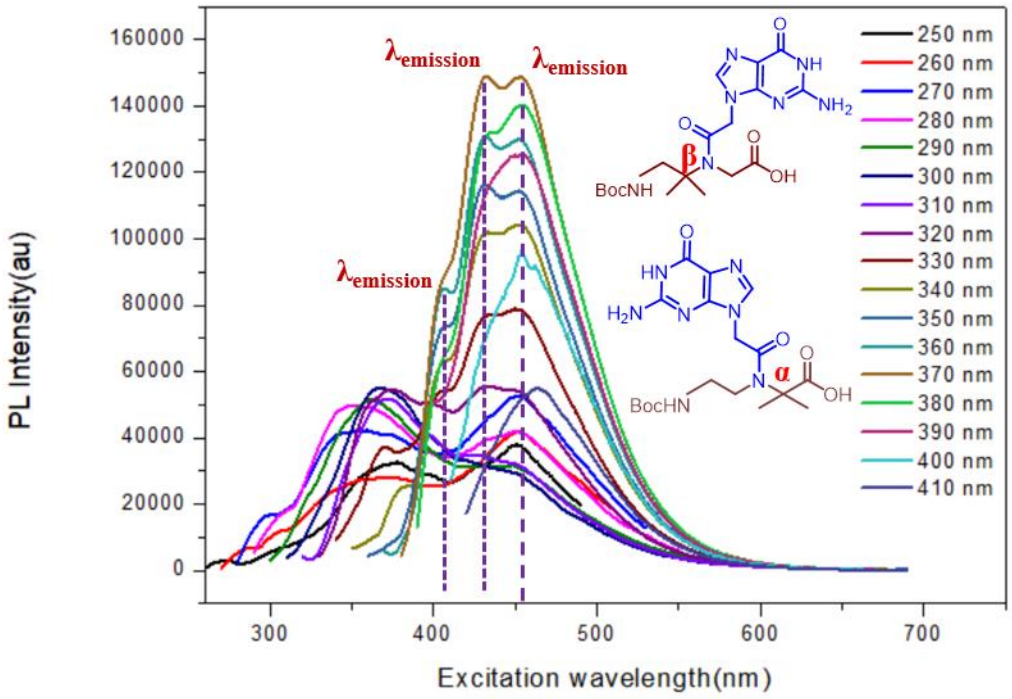
Partial ^{1}H-^{1}H 2D NOSEY spectra of α, α-*dimethyl*-(N^4-Cbz)-C *aeg* PNA **(5)** in $CDCl_3$ (100% Z rotamer). δ to β/γ connectivity indicate 100% Z-rotamer in solution.



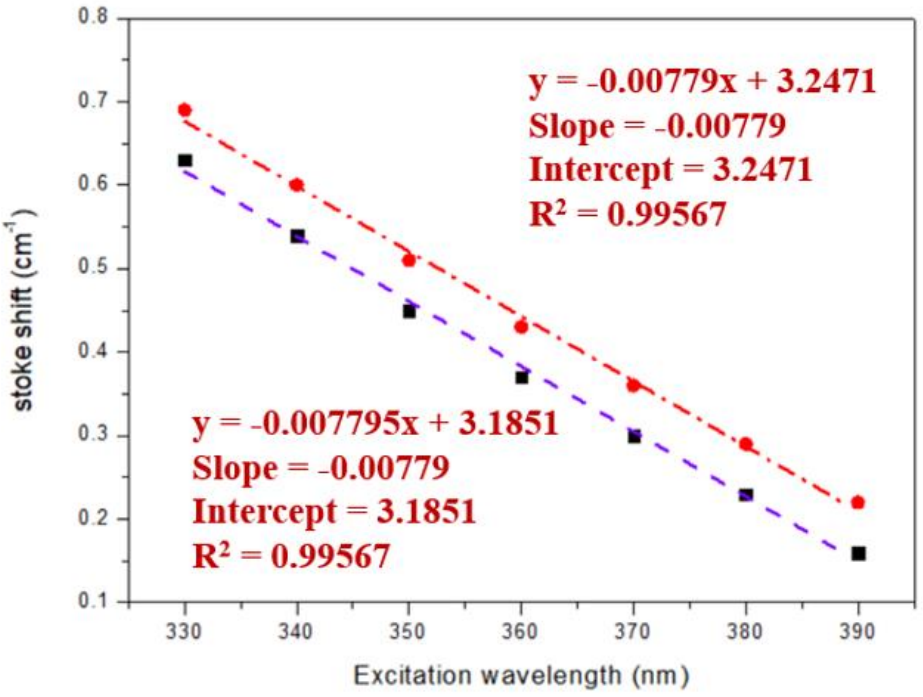
Plot of stoke shift vs Excitation wavelength of β-*gemdimethyl-aeg*-PNA-G **(14)**. Stoke shift decreases with increasing the excitation wavelength of the emission.



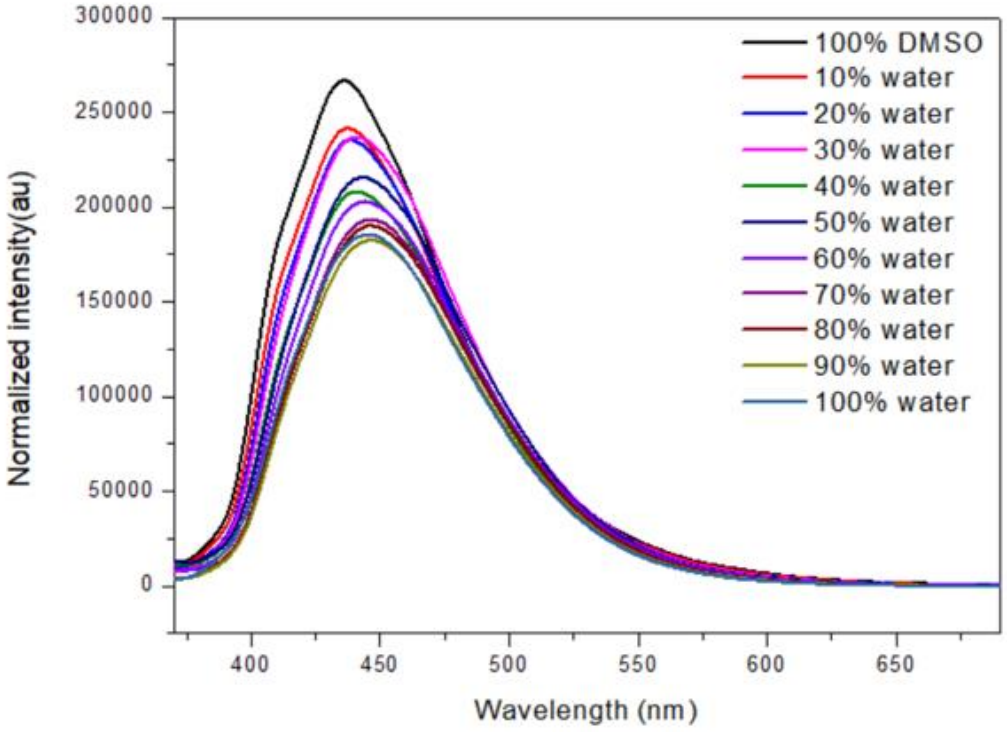
Stoke shift vs excitation wavelength plot of α-*gemdimethyl-aeg*-PNA-G **(8)**. Stoke shift decreases with increasing the excitation wavelength.



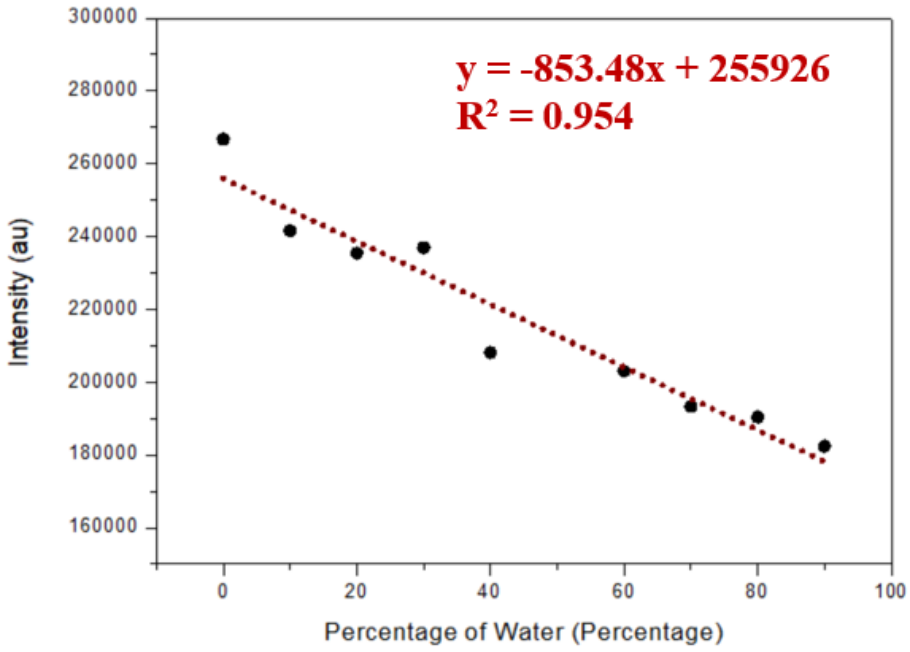
Excitation dependent emission spectra of 1:1 molar ratio of α-*gemdimethyl-aeg*-PNA-G **(8)** and β-*gemdimethyl-aeg*-PNA-G **(14)** (concn, 240 μM)



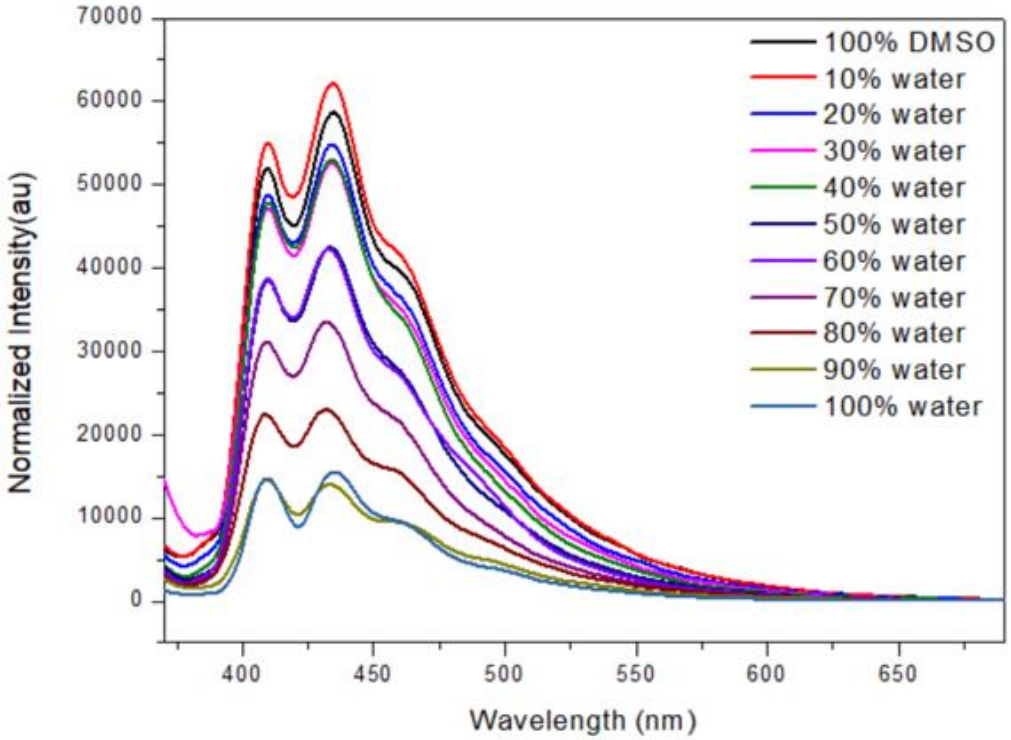
Stoke shift vs excitation wavelength plot of α- *gemdimethyl*- (6-Cl-2-amino) purine PNA (**6**). Stoke shift decreases with increasing the excitation wavelength.



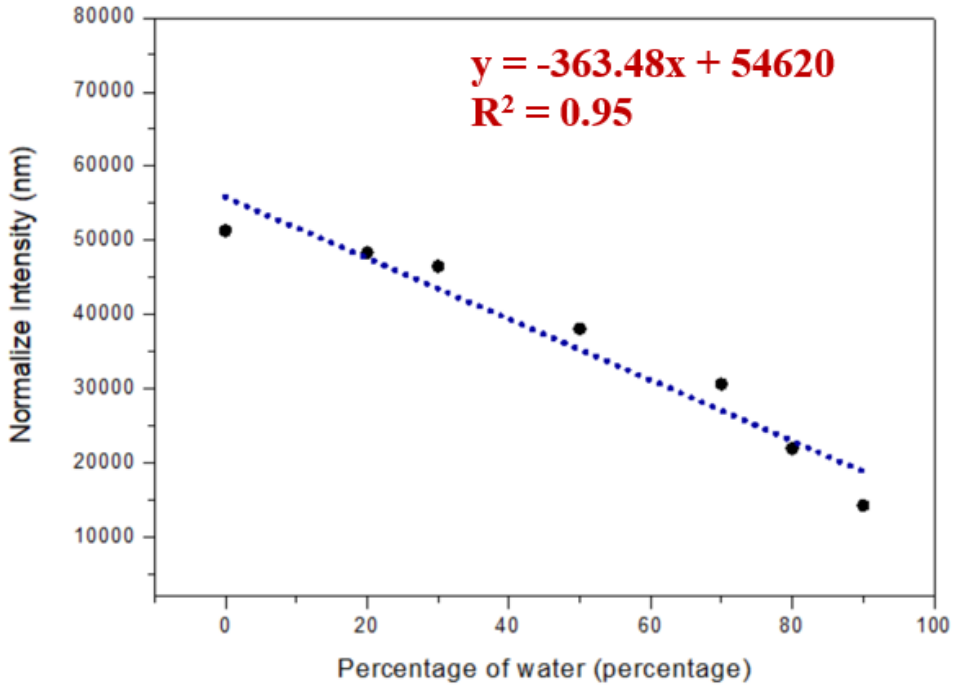
Variation of fluorescence intensity of β-*gemdimethyl-aeg*-PNA-G (**14**) (concn,150 μM) in various proportion of DMSO/Water mixture at 360 nm excitation.



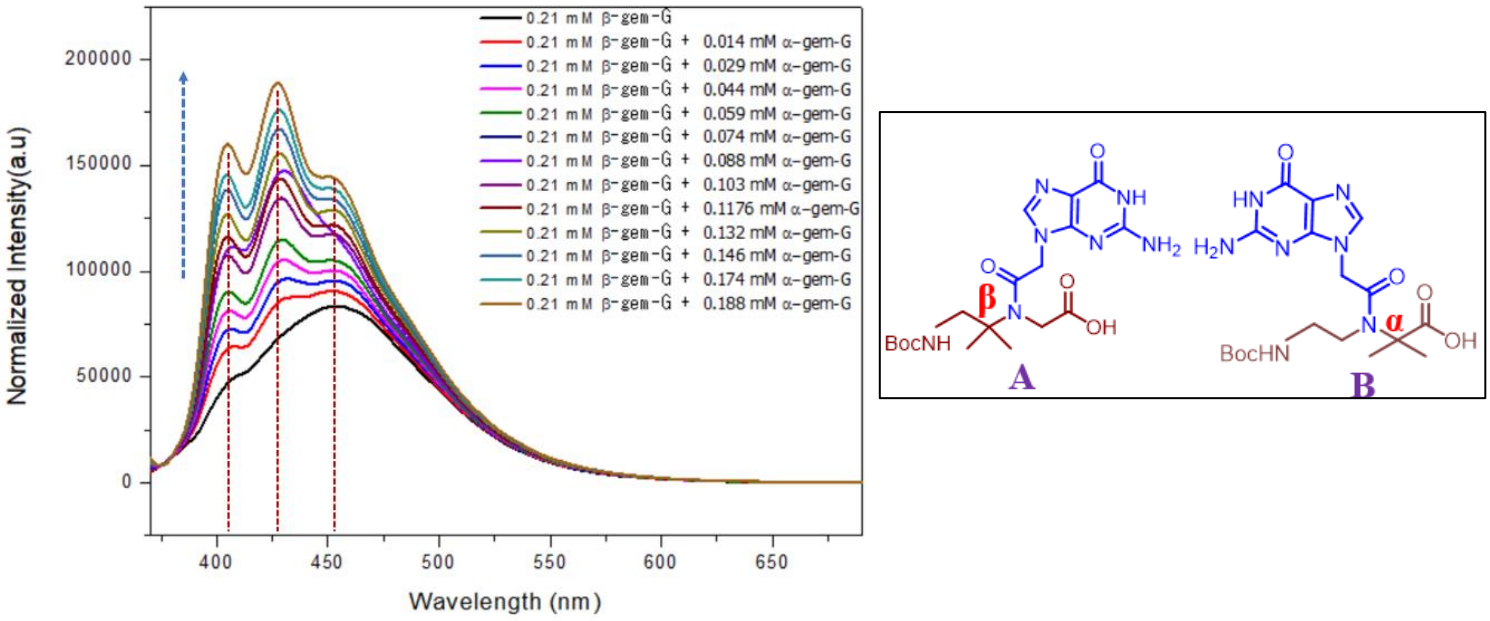
Graphical representation of variation of fluorescence intensity of β-*gemdimethyl-aeg*-PNA-G (**14**) with increasing water percentage suggesting quenching.



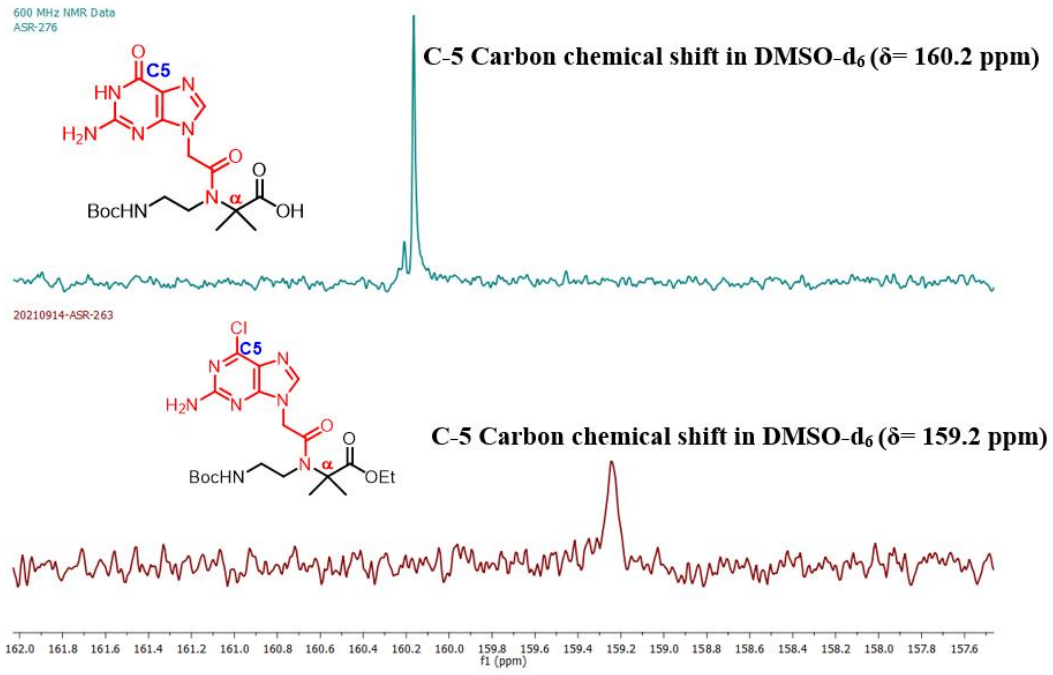
Variation of fluorescence intensity of α-*gemdimethyl-aeg*-PNA-G (**8**) (concn,150 μM) in various proportion of DMSO/Water mixture at 360 nm excitation.



Graphical representation of variation of fluorescence intensity of α-*gemdimethyl-aeg*-PNA-G (**8**) with increasing water percentage suggesting quenching.



Titration of β-*gemdimethyl-aeg*-PNA-G **(14, A)** with increasing amount of α-*gemdimethyl-aeg*-PNA-G **(8, B)**. Initially the emission spectra of the β-*gemdimethyl-aeg*-PNA-G (**14**) predominates but with increasing concentration of α-*gemdimethyl-aeg*-PNA-G (**8**) a combination emission spectrum result.



Comparison of chemical Shift of C-5 carbon of α- *gemdimethyl-* (6-Cl-2-amino) purine PNA (**6**) and α-*gemdimethyl-aeg*-PNA-G (**8**)**.** Comparison shows 1 ppm shift of C5 chemical shift suggesting change of functional group after hydrolysis.

Chapter 4

Optical properties of Bi-modal PNA monomers and oligomers

Chapter 4

Optical properties of Bimodal PNA monomers and oligomers

4.1. Introduction

Peptide Nucleic Acid (PNA) are synthetic DNA mimic where the negatively charged sugar phosphate backbone is replaced by neutral amino ethyl glycine (*aeg*) unit.[1] Nucleobases are connected methylene carbonyl linker in PNA. The nucleobases (A/T/G/C) are separated by six atom spacers in PNA, and the backbone relates to the nucleobase by a two-atom spacer similar as DNA. PNA binds with complementary DNA/RNA and PNA with higher binding affinity compared to the original DNA molecule.[1] Polypyrimidine PNAs forms PNA: DNA: PNA (PNA_2: DNA) triplexes where one PNA strand binds with complementary DNA via Watson-Crick hydrogen bonding while the same DNA strand binds with another PNA via Hoogsteen base pairing to form a triplex.[2] The ability to constitute complex with reverse composition of DNA1:PNA: DNA2 allows targeting of two DNA strands simultaneously by a single PNA strand and has potential future for PNA based biomedical research.

A novel approach in the literature towards this aim is the use of bifacial heterocyclic bases that can bind two bases one on either side in a complementary manner, which was designed by Lehn who coined the term *Janus* bases.[3] These exhibit hydrogen bonding from both faces of the heterocycle, compared to natural nucleobases (A/T/G/C) which exhibit hydrogen bonding from one face of heterocycle.[3] Depending on the base complementarity Janus heterocycles are divided into two broad categories (a) self-complementary Janus heterocycles and (b) non-self-complementary Janus heterocycles. If one face of the heterocycle is complementary to Adenine (complementary of T) and another face is complementary to Cytidine (complementary of G) then the Janus heterocycle (Janus-TG) belongs to the category of non-self-complementary Janus heterocycles as T and G are non-complementary nucleobases (Figure 4.1). The modified DNA oligomers with Janus-TG can target two DNA strands containing C and A

If one face of the heterocycle is complementary to Thymine (complementary of A) and another face is complementary to Adenine (complementary of T) then the Janus heterocycle is a self-complementary Janus heterocycle (Janus-TA) as T and A are complementary

heterocyclic nucleobases (Figure 4.2). Janus-TA modified oligomers can form duplexes by self-complementary base pairing.

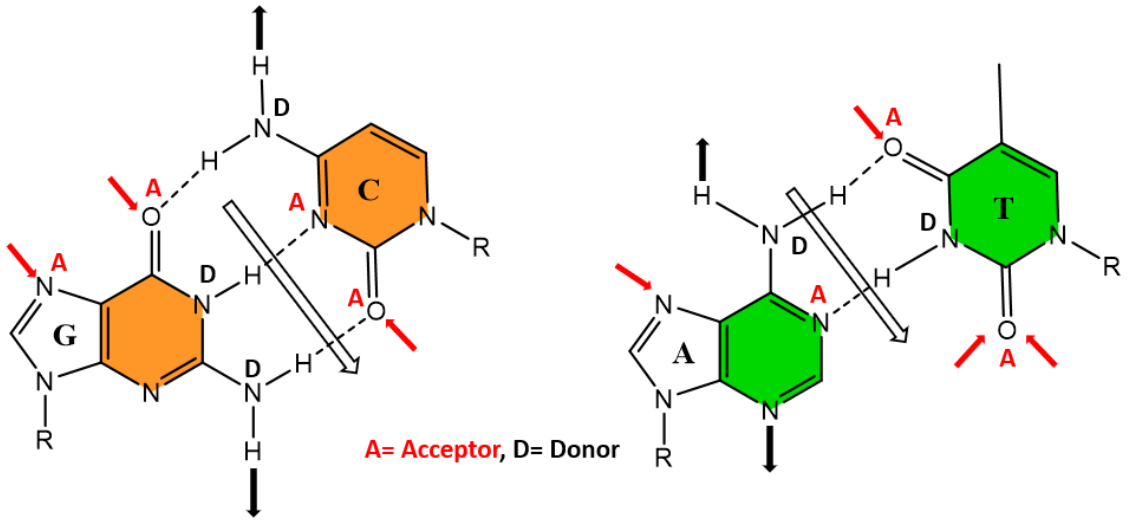


Figure 4.1. Janus-wedge recognising regular DNA base pairs (A: Hydrogen bond acceptor, D= Hydrogen bond donor, R= deoxyribose)

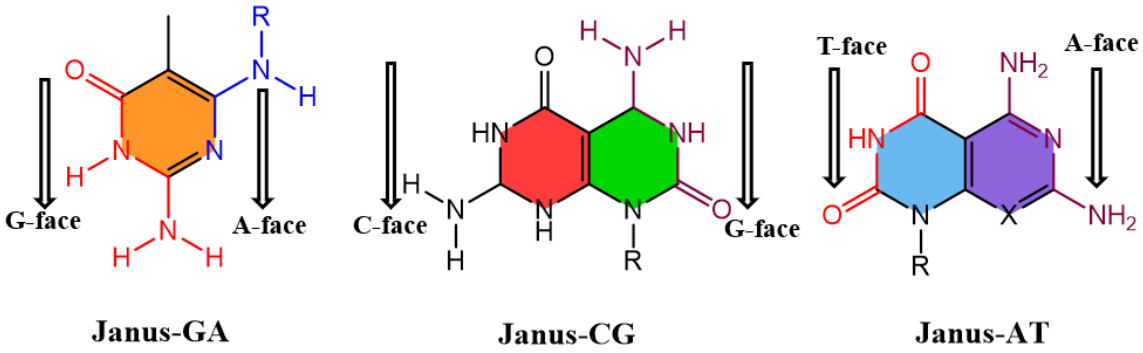


Figure 4.2. Janus heterocycles (R= alkyl group and X= N or CH for Janus-AT)

Several groups have reported the synthesis and properties of Janus-Heterocycles. Lehn *et al.*[4] reported the synthesis and properties of non-self-complementary Janus-GA, Mascal *et al.*[5] reported the synthesis and properties of self-complementary Janus-GC and Perrin *et al.*[6] reported the synthesis and properties of self-complementary Janus-AT. Self-complementarity of the Janus-heterocycles results in interesting H-bonding pattens and non-covalent stacking interactions to form interesting supramolecular structures. Hydrogen bonding and stacking interactions between the nucleobases are the driving force for the formation of supramolecular structures. Fenniri *et al.*[7] showed formation of helical rosette nano-tubes by Janus-GC

heterocycles in water. The property of the helical rosette nanotube could be altered by changing the substituent of the Janus-GC heterocycles. Chirality could be induced in the rosette nanotube by adding chiral amino acids to a Janus-GC heterocycles consisting of crown ether in the side chain. Investigation of the nanostructure of Janus-AT under similar condition results linear sheet-like structure.[8,9]

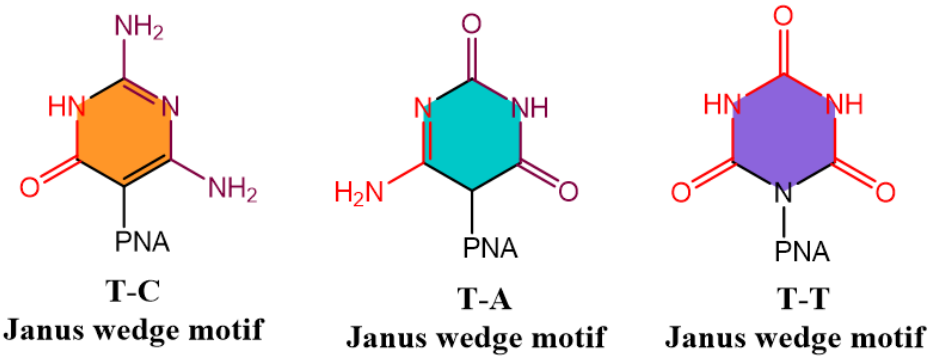


Figure 4.3. Janus wedge motif used for targeting DNA to form PNA_2: DNA triplexes

PNAs containing modified nucleobases were used to target the single stranded DNA or double stranded DNA to form triplexes. Mclaughlin *et al.*[10] targeted T-C mismatch in DNA via a modified nucleobase containing PNA (Janus-wedge motifs, W) having dual thymine and cytosine face in a single nucleobase. Interestingly, the modified PNA shows stable triplex formation with DNA having consecutive eight T-C mismatches. In a follow up work heterocycles containing thymine and adenine face were used to target DNA to form triplexes.[10]

Ganesh *et al.*[11] used cyanuric acid as Janus wedge motif to stabilize the PNA_2: DNA triplexes. Single cyanuric acid modified PNA oligomer exhibited a T_m increment of 12º C in the PNA_2:DNA triplexes suggesting a stronger triplex formation by cyanuric acid modified PNA. All these modifications involve the structurally changed nucleobases retaining the parent backbone of PNA or DNA. In a very original approach in 2020, Ganesh *et al.*[12] presented modification of PNA backbone to target two DNA strands simultaneously, by keeping the natural nucleobases intact (Figure 4.4). This novel backbone modified PNAs termed *Bimodal peptide nucleic acids*[12] (*bm*-PNA) can simultaneously target two DNA strands. In bimodal PNAs, one of the nucleobases is connected to the backbone by methylene carbonyl linker (t-amide, standard PNA) while the second nucleobase is linked at the α-side chain via a triazole linker,[12] or at Cγ of backbone via an extended amide spacer.[13] Such backbone modified bimodal peptide nucleic acids bind two complementary strands of DNA, one from the t-amide

face, and the other from nucleobase on C_α or C_γ sidechain forming two duplexes. The derived double duplexes shows higher thermal stability compared to the isolated tertiary amide PNA:cDNA duplex or the Cα- triazole face PNA^{tz}:cDNA duplex or C_γ-amide PNA^{am}:DNA duplex. Ganesh *et al.*[12] in 2021 extended this idea to assemble the triplex of duplexes (*bm*-PNA-T)$_2$: dA_8 with *bm*-PNA carrying all-thymine on C_α-triazole face.

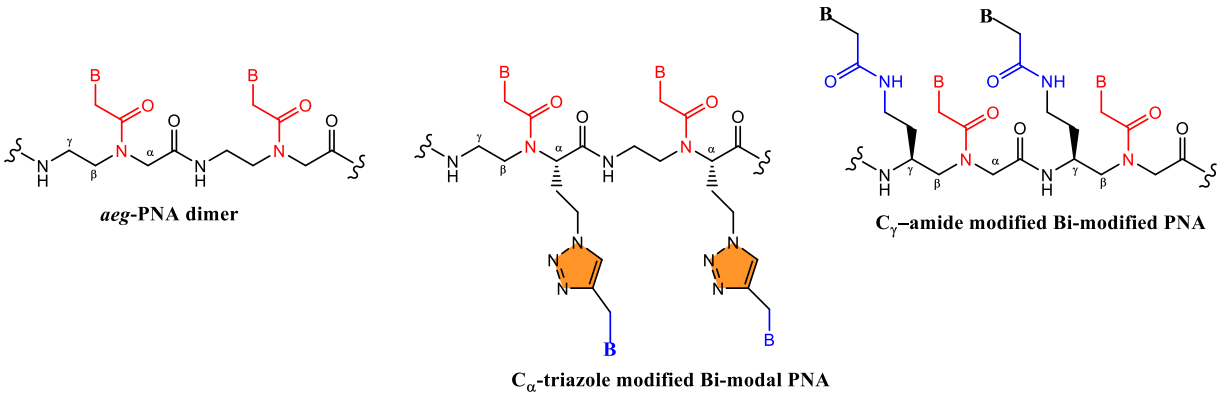


Figure 4.4. Novel Backbone modified Bimodal PNA analogues

The C_γ-modified bimodal PNA analogues are similar to C_α-triazole *bm*-PNA, in which the second string of nucleobases are connected via amide linker at C_γ-site on PNA backbone.[13] These also form double duplexes and triplex of duplexes similar to similar to that C_α-triazole modified bimodal PNA (Figure 4.5). Bimodal PNAs become chiral at C_α and C_γ and it was shown that C_γ(*S*)-*bm*-PNA complexes are more stable than C_α-*bm*-PNA complexes. In such multistranded complex assemblies it does not matter how the supra-molecular complexes assemble-– first duplex and then triplex or vice-versa and the final complex are always the

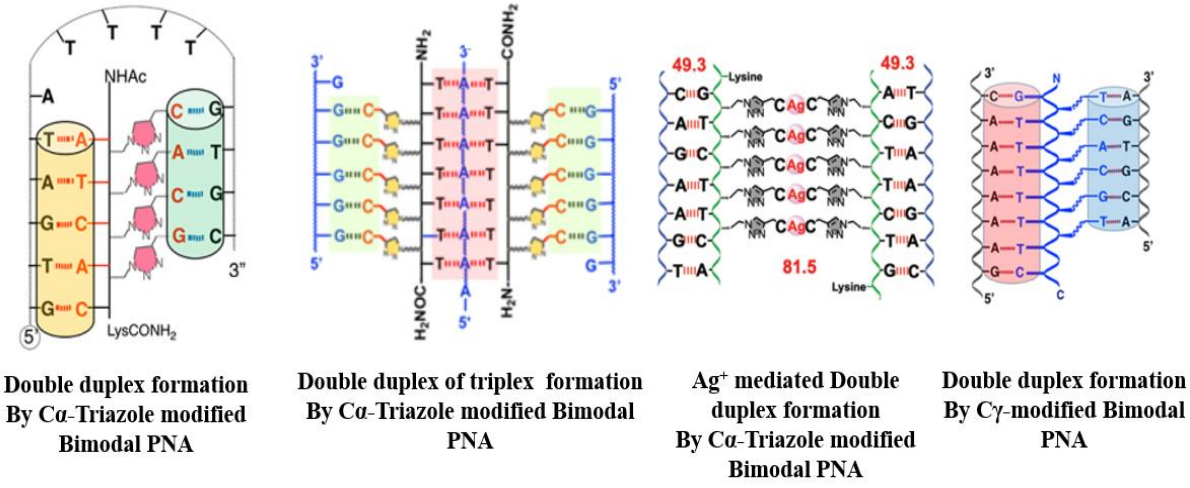


Figure 4.5. Formation of multistranded complexes from bimodal PNAs

same. The C_γ-modified polypyrimidine bimodal PNAs also exhibit cell penetrating properties in MCF 7 and NH 3T3 cells.[14] It was also shown that bimodal PNAs with all C on Cα-triazole side chain or t-amide side can form silver assisted C: Ag^+:C duplex which in turn can form duplex with cDNA from the other face, leading to triple duplexes in which one duplex is held by Ag^+ and the other by WC H-bonds. This approach has the potential to assemble programmed supramolecular complexes such as tetra duplex of tetraplexes.[15]

4.2. Rationale and objective

The solid phase synthesis of bimodal PNAs with mixed base sequence on both sides involve complex synthetic schemes.[12,13] These employ orthogonal protecting groups to enable selective deprotection of Boc at N-terminus for coupling to extend the main chain and selective deprotection of Fmoc-group on sidechain to enable coupling of second nucleobase on side chain.[12] An alternative method is to synthesize *bm*-PNA monomers in which both nucleobases are already present (Figure 4.6) and then use it for sequential coupling. This approach needs synthesis of several possible dinucleobase combinations in the block monomer. This chapter involves synthesis of such monomer blocks for use in bimodal PNA synthesis. Based on the results of previous chapter where some Cα / Cβ disubstituted PNA monomers showed fluorescence, it would be curious to study if Cα-monosubstituted monomers would be fluorescent and if so the nature of bimodal oligomers assembled from such monomers.

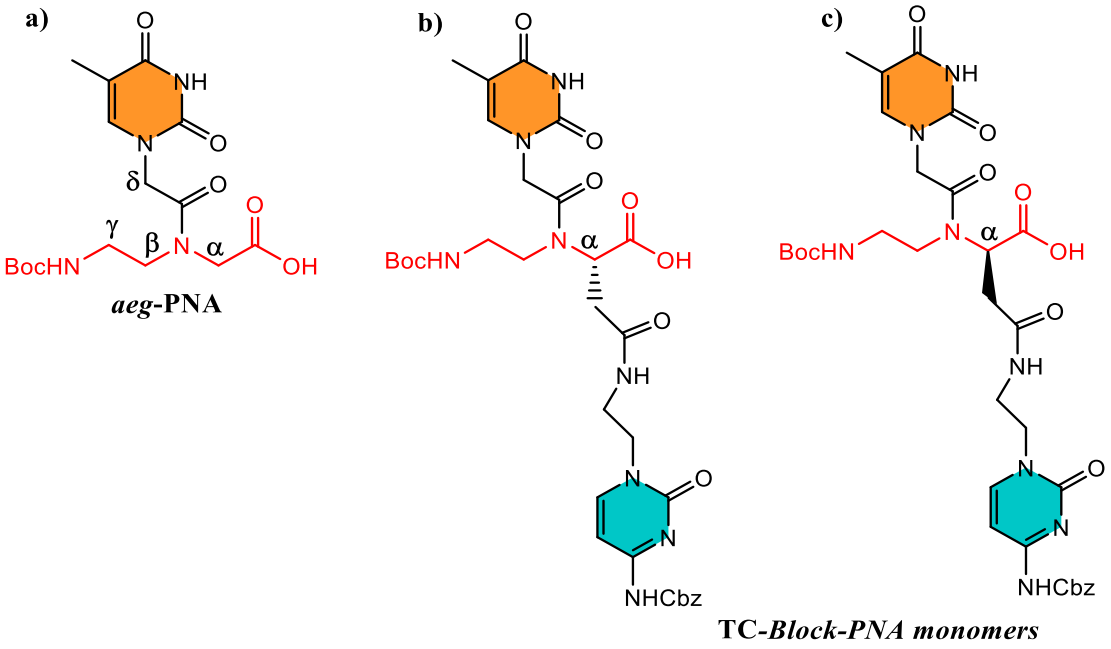


Figure 4.6. a) *aeg*-PNA monomer and (b) and (c) target block bimodal PNA monomers

Fluorescence of purine and pyrimidine based *gemdimethyl* PNA monomers were studied in Chapter 3 of this thesis. Hence fluorescence of pyrimidine-based block bimodal PNA monomer were explored in this chapter to observe the effect of incorporating two different pyrimidine bases in the same *aeg*-PNA based monomer. Apart from that, nature of rotamer also plays a vital role in the fluorescence emission of purine and pyrimidine based *gemdimethyl* PNA monomers. Thus, it would be interesting to study the nature of rotamers in and their effect on the fluorescence emission of designed block bimodal PNA monomers. With this rationale the Cα(*R*/*S*)-(acetamido)-*bm*-PNA monomers (TC-block monomer, Figure 4.6b and 4.6c) were synthesized, and fluorescence properties of the derived monomers and oligomers were studied.

The specific objectives of this chapter are

(a) Synthesis of Cα(*S*)-(acetamido)-*bm*-PNA monomer (Figure 4.6b)

(b) Characterization of the synthesized Cα(*R*)-(acetamido)-*bm*-PNA monomer

(c) Study the rotamer equilibrium and solution state structure of the Cα(*R*)-(acetamido)-PNA monomer.

(d) Study of fluorescence emission properties of the derived PNA monomer

(e) Synthesis and characterization of the fully modified bimodal PNA oligomer

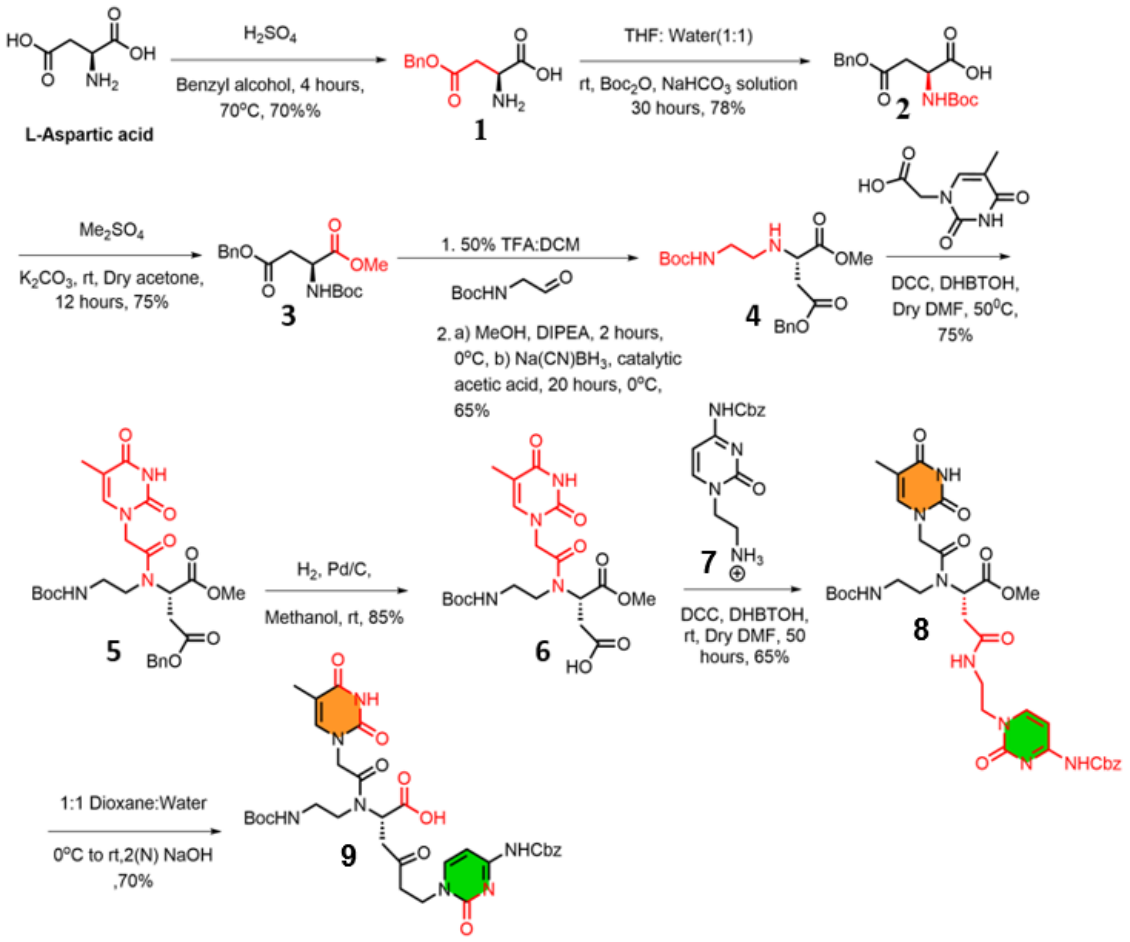
(f) Study of fluorescence emission properties of the synthesized fully modified PNA oligomer.

4.3. Results and discussion

4.3.1 Synthesis of Cα(S) -(acetamido)-*bm*-PNA monomer

The commercially available L-Aspartic acid is treated with benzyl alcohol in presence of sulfuric acid to obtain the selective γ-benzylated amino acid (**1**) (Scheme 1).[16] This was treated with Boc_2O in presence of *aq*. $NaHCO_3$ to afford the NHBoc-protected amino acid (**2**) which was methylated using dimethyl sulphate in presence of anhydrous K_2CO_3 gives the orthogonally protected amino acid derivative (**3**).[17,18] Boc-deprotection of 3 in presence of 50% TFA:DCM, followed by reductive amination in presence of tert-butyl(2-oxoethyl)carbamate gives *in situ* generated Schiff base. The *in situ* reduction of Schiff base with $Na(CN)BH_3$ afforded the required PNA backbone precursor (**4**). DCC mediated coupling of 4 with thymine acetic acid resulted in nucleobase coupled monomer (**5**). Hydrogenolysis of **5** in methanol

afforded the debenzylated Cα-aspartic acid PNA monomer (**6**). DCC mediated coupling of compound **6** with N^4-Cbz-Protected cytosine amine (**7**) gave the C_α-modified block bimodal peptide nucleic acid monomer **8** (Scheme 1). Base mediated hydrolysis of **8** in presence of aq. NaOH provided the free acid monomer (**9**). N^4-Cbz-protected cytosine amine (**7**) was synthesized using reported procedures.[19] The synthesized compounds were characterized by ^{1}H NMR and mass spectrometry.

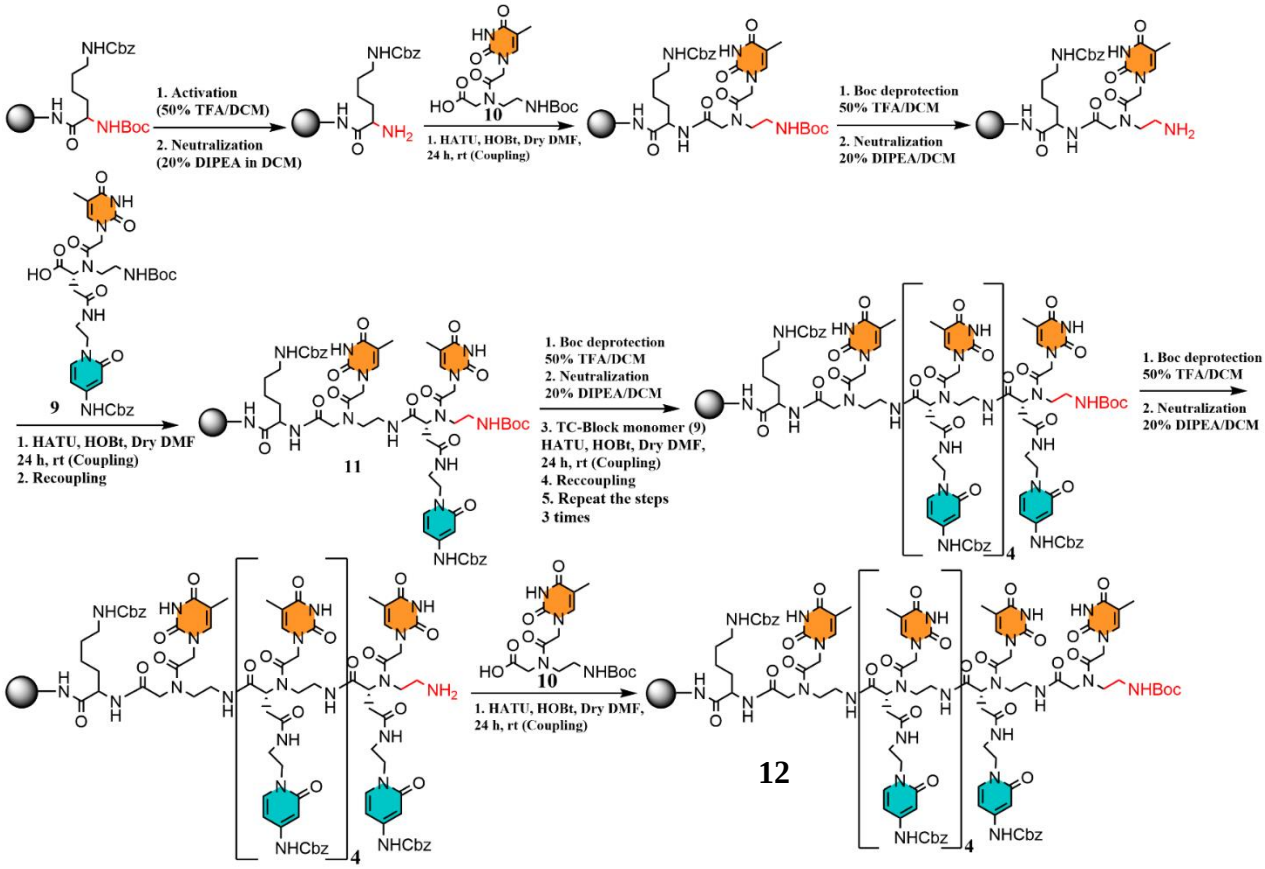


Scheme 1: Synthesis of Cα(S) -(acetamido)-*bm*-PNA monomer (TC-Block PNA monomer)

4.3.2 Solid phase synthesis of *bm*-PNA oligomer

The homo-oligomer of Cα(S) -(acetamido)-*bm*-PNA monomer **9** was synthesized using standard solid phase peptide synthesis for chain extension (scheme 2). MBHA resin derivatized with lysine (loading value-0.22 mmol/gm) was swelled over DCM and activated using DIPEA in After the neutralization process the *aeg*-PNA-T monomer **10** was coupled with the amino group present in the MBHA resin. It is known that α-CH of Cα-modified PNA participates in racemization reaction in presence of HBTU. To avoid the racemization during the coupling in

solid phase, a mixture of HATU and HOBt was employed as coupling agent which prevents the racemization process during coupling reaction.[20] After successful coupling, the Boc protecting group was removed by 50% TFA/DCM followed by neutralization using 20% DIPEA in DCM. Coupling of the free amine with the Cα(*S*)-(acetamido)-*bm*-PNA **9** affords the dimer (**11**) of the *aeg*-PNA-T-bimodal PNA on the resin. The cycle of coupling, deprotection of sidechain and coupling at sidechain process was repeated 4 times to obtain the desired length of the *bm*-PNA oligomer **12**. The final coupling was performed with *aeg*-PNA-T for completion of the target sequence and the synthesis (Scheme 2). The resin bound PNA oligomer (Scheme 3) were cleaved by using TFA/TFMSA, purified by reverse phase HPLC and characterized by MALDI-TOF mass spectrometry.[21]

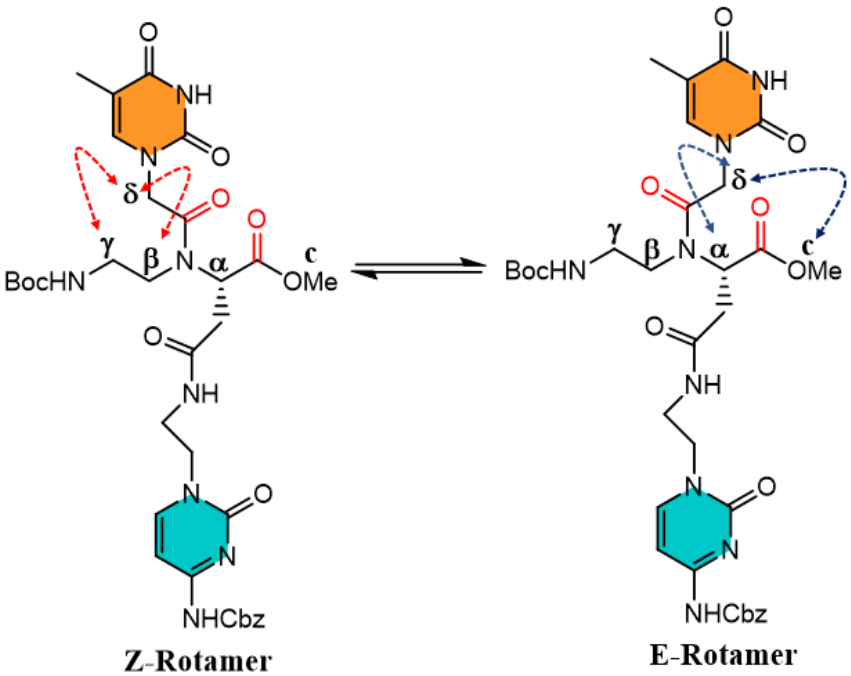


Scheme 2: Solid phase synthesis of *bm*-PNA oligomer **12**

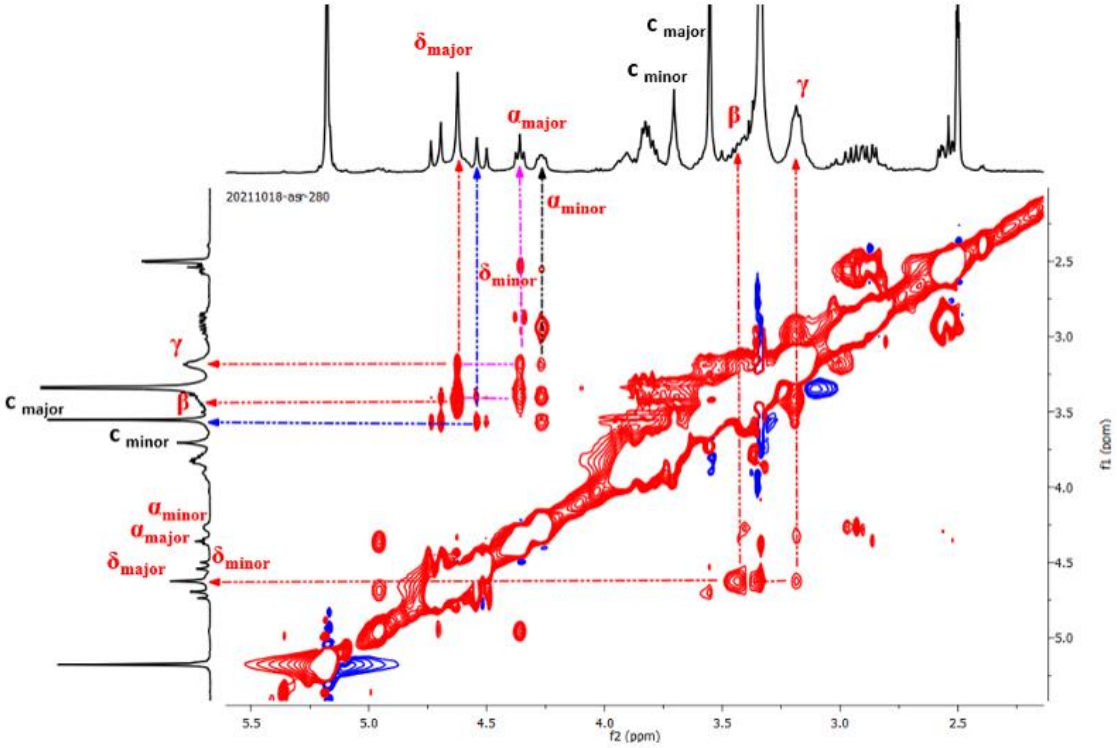
4.3.3. Rotamer preferences of Cα(*S*)-(acetamido)-*bm*-PNA bimodal PNA monomer

As described in Chapter 2, *gemdimethyl* substitution at Cα leads to exclusive Z-rotamer, while substitution at C_β results in exclusive E-rotamer.[22] In the present chapter mono-substitution of C_α-methylene by amide linked nucleobase results in Cα(*S*)-(acetamido)-*bm*-

PNA monomer (**9**). The primary ^{1}H NMR analysis of **8** shows the presence of a mixture of rotamers with the rotamer ratio being 55% (major) and 45% of the (minor) compared to the 75% major and 25% minor rotamer for unsubstituted *aeg*-PNA-T (**10**). The assignment of Z and E-rotamers were done as described earlier by NOE cross-peak between the δ-methylene group with β/γ protons for the Z-Rotamer and δ-methylene group with a proton for the E-rotamer (Figure 4.7a). The partial ^{1}H-^{1}H NOSEY spectra of Cα(*S*)-(acetamido)-*bm*-PNA monomer (**8**) in DMSO-d_6 is shown in Figure 4.7b. The major rotamer of Cα(*S*)-(acetamido)-*bm*-PNA monomer (**9**) exhibits cross peak with β/γ -CH_2 protons suggesting it to be the Z-Rotamer (red line, Figure 4.7b), whereas the minor rotamer shows cross peak with ester methyl groups suggesting the minor rotamer to be E (blue line, Figure 4.7b). This is similar to that observed in unsubstituted *aeg*-PNA (**10**). Thus, the rotamer preference of C_α (*S*)-derivatised *aeg*-PNA as in block monomer **8** remains unaffected. Interestingly the major α-CH proton also shows ^{1}H-^{1}H NOSEY cross peak with backbone β/γ- methylene suggesting a folding of backbone in the major Z-conformation. Similarly, the minor α-CH also shows ^{1}H-^{1}H NOSEY cross peak with β/γ-methylene and ester methyl group suggesting a folded conformation for the minor E-rotamer as well in solution (magenta and black lines, figure 4.7b). *aeg*-PNA (**10**) lacks similar folded conformation in solution and solid state.



(**a**)



(b)

Figure 4.7. (a) E and Z Rotamer assignments (b) Partial ^{1}H-^{1}H NOSEY spectra of Cα(S)-(acetamido-C)-*bm*-PNA T monomer methyl ester (**8**)

4.3.4. UV-Visible spectroscopy and fluorescence study of the monomers

Classical *aeg*-PNA-T (**10**) shows strong UV-absorption band at 254 nm and a shoulder band around 220 nm respectively. The fluorescence emission spectra of *aeg*-PNA-T (**10**) at λ_{ex} of 320 nm and 330 nm exhibit very weak emission bands or absent suggesting non-fluorescent nature of the *aeg*-PNA-T monomer (**10**) in solution. In comparison, UV-visible spectra of Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer (**9**) exhibits a band at 280 nm with shoulder at 300 nm in addition to the strong UV absorption band around 246 nm arising from the absorption of nucleobases (Figure 4.8a). No significant UV-absorption band was observed beyond 300 nm. The Cα(S)-(acetamido-C)-*bm*-PNA-T monomer (**9**) exhibits excitation dependent emission spectra in ethanol. Upon excitation in the region 280-320 nm, the emission maxima at 376 nm slowly red shifted to 400 nm followed by decreasing intensity and shift to 425 nm at higher excitation wavelengths in 330 nm-400 nm region (Figure 4.8b). Figure 4.8 shows the UV and fluorescence spectra of the bimodal PNA monomer.

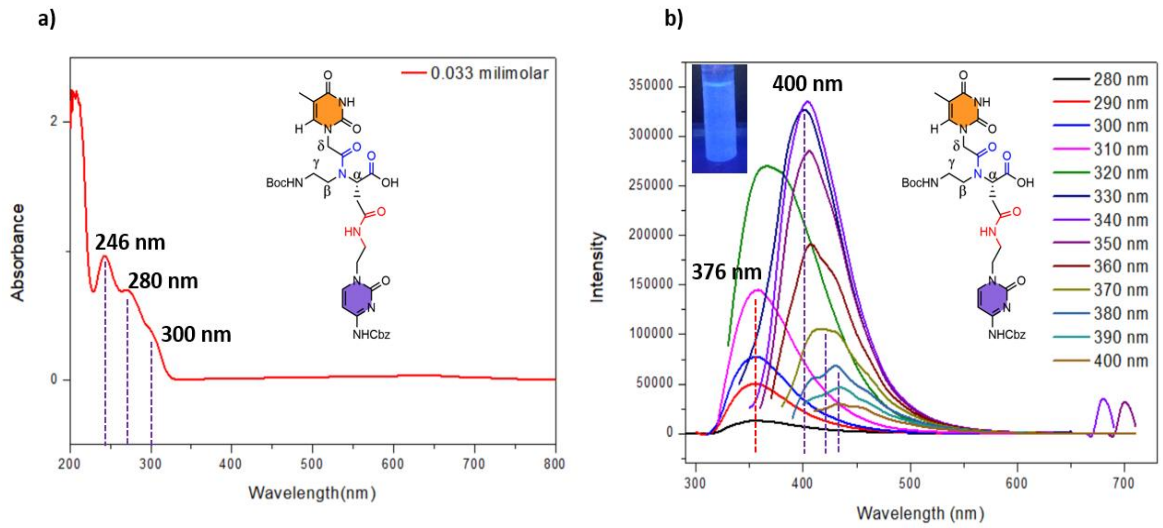


Figure 4.8. a) UV-Visible spectra of Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer **9** (Conc, 33 μM) (b) Fluorescence spectra Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer **9** (Conc, 250 μM); Solvent: ethanol

This result suggested Red Edge Excitation Shift (REES) effect[23] in the emission spectra of Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer (**9**) that shows bright blue fluorescence emission in solution. The pattern of emission band remains same at all excitation wavelengths, similar to that observed with α/β-*gemdimethyl* PNA-G monomers which exhibited selective rotamer (100% E or 100% Z) dependent fluorescence emission. The only difference with Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer (**9**) is that it shows a mixture of rotamers and hence combined fluorescence with broad band and a shoulder.

4.3.5. Solvent dependence of UV spectra and fluorescence emission of homo oligomer

In view of the unexpected fluorescence emission seen in the Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer (**9**) in solution, the fluorescence emission of the Cα(*S*)-(acetamido-C)-*bm*-PNA-T homo oligomer (**12**) was studied in water and phosphate buffer. The UV-Visible spectra of homo oligomer show a strong band at 270 nm at all concentrations and with no significant longer wavelength UV-absorption band at ~300 nm as observed for monomer in water (Figure 4.9a).

The fluorescence emission spectra of the homo-oligomer (**12**) recorded in water exhibited emission maxima at 425 and 430 nm upon excitation at in the region 330 nm - 350 nm (Figure 4.9b). The position of the emission maxima remained almost same at all the excitation wavelengths (330 nm-350 nm) in water. The intensity of the emission maxima

decreased upon increasing the excitation wavelength. However, in phosphate buffer at pH 12.75, the emission spectra of Cα(*R*)-(acetamido-C)-*bm*-PNA-T homo-oligomer (**12**) exhibited excitation dependent fluorescence emission (Figure 4.10), with strong emission at 374 nm at λ_{ex} in 280-310 nm. At excitation beyond 320 nm, the emission red shifted to λ_{em} at 480 nm, reaching maximum for excitation at 380 nm. Interestingly, the emission property disappeared in phosphate buffer at pHs of 7.23 and 10.25 respectively.

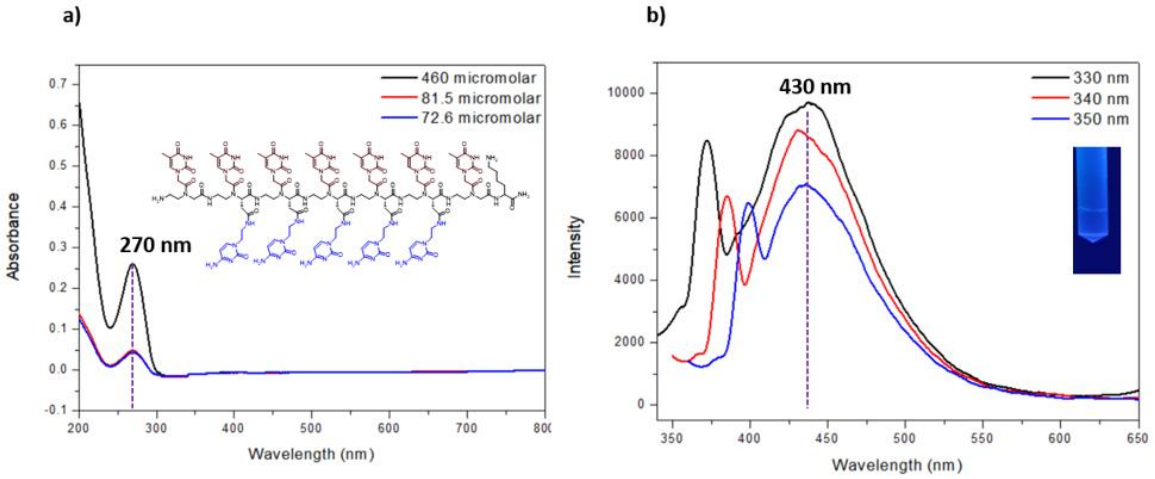


Figure 4.9. a) UV-Visible spectra of the homo oligomer of Cα(S)-(acetamido)-*bm*-PNA homo oligomer **12** (Conc, 250 μM) b) Fluorescence spectra of the homo oligomer of Cα(S)-(acetamido)-*bm*-PNA monomer **12** in water (Conc, 250 μM).

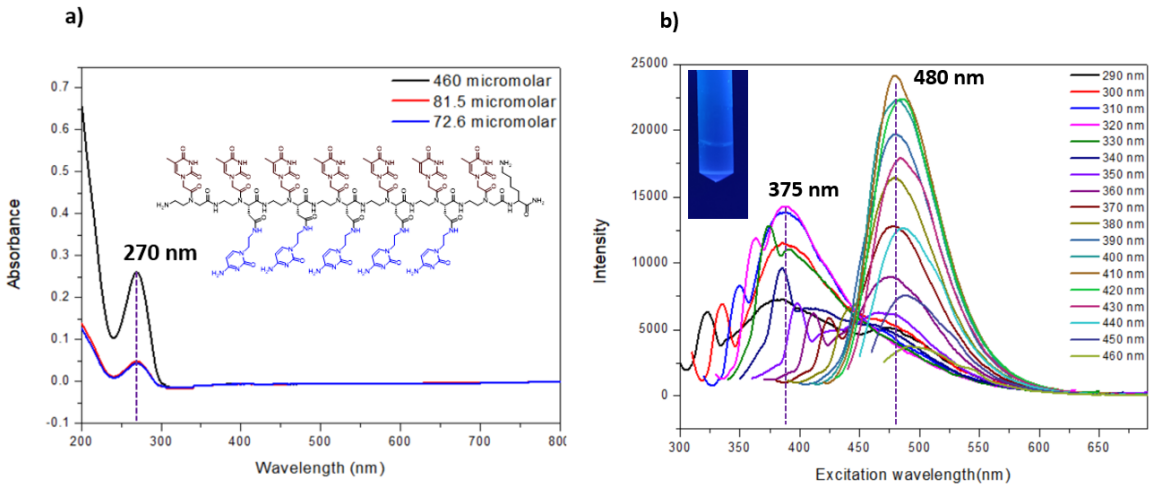


Figure 4.10. a) UV-visible spectra of Cα(*S*)-(acetamido-C)-*bm*-PNA-T homo oligomer **12** (b) Fluorescence spectra of Cα(S)-(acetamido-C)-*bm*-PNA-T homo oligomer **12.** Concentration-250 μM; Solvent, phosphate buffer (10 mM, pH 12.75)

4.4. Discussion

This chapter reports on the synthesis and rotameric studies on Cα(*S*)-(acetamido-C)-*bm*-PNA-T methyl ester (**8**) and Cα-monosubstituted *Janus PNA* monomer acid (**9**). Unlike the corresponding *gemdimethyl* substituted monomer reported in previous chapter, the mono-substituted aeg PNA monomer shows a mixture of rotamers with Z as the major rotamer. This suggests that steric effect imposed by *gemdimethyl* sunstituents have an important role in selective stabilization single rotamer, which is lacking in the present Cα-monosubstituted monomers that show a mixture of rotamers. The synthesis of the corresponding homo oligomer was achieved on solid phase peptide synthesis by repeated single coupling step without the orthogonal protection strategy used in earlier work.[12,13] The purified oligomer (**12**) was characterized by mass spectra (MALDI) data. The monomer (**9**) despite lacking UV spectral band in beyond 290 nm, exhibited fluorescence with emission at 430 nm. The oligomer which also lacked UV absorption beyond 280 nm, also exhibited fluorescence with emission at 480 nm in phosphate buffer at alkaline pH 12.75, but no fluorescence at lower pHs of 10.75 and 7.3. Both the monomer and oligomer exhibited excitation dependent emission spectra and red edge excitation shift (REES) in solution.[23] The multiple emission bands at lower wavelength 375 nm, but a single emission band at 425 nm is reverse of that seen earlier in Cα-*gemdimethyl* substituted monomer having a single rotamer. This may be a consequence of presence of mixture two rotamers in Cα-monosubstituted derivatives, that may slightly differ in solvation in excited state. This is perhaps the first example of fluorescence in PNA oligomer that does not possess a fluorescence active chromophore. The origin of fluorescence in these non-conjugated molecules may arise from correlated electronic excitations of the σ, n and π electrons of C-C, C-N and C-O bonds and TSC interactions that reduce the band gap between the S_1/T_1 to S_0 state.[24,25] n-σ*, n-π*, π-π* interactions result an extensive overlap of orbitals in the cluster and assembled state. The results also point out that the rotamer distribution affects the emission of fluorophores. The α/β-*gemdimethyl-aeg* -G PNA monomers are locked in either Z (100%) or E (100%) rotamers and exhibit emission in lower excitation wavelength as well as higher excitation wavelength. While in the monomers, a mixture of rotamers is present, it may not be so in the derived monosubstituted oligomer which show single emission band at higher excitation wavelength suggesting that the emission originates majorly from a single rotamer in the oligomer of Cα(S)-(*acetamido*-C)-*bm*-PNA-T monomer (**12**). Interestingly, the homo oligomer *bm*-PNA-T_7C_5 (**12**) exhibited fluorescence emission in the range 290 nm-460 nm in phosphate buffer at higher alkaline pH (12.75), but not at lower pH (7.0-10.0). This

suggested that deprotonation of T and C at pH 12.75 and consequent presnec of base anions may be important for the fluorescence phenomenon of the TC block monomer (**9**). This needs further investigations.

4.5. Summary

In summary, this chapter reports

- Synthesis of the Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer (TC block monomer, **9**) from L-aspartic acid.
- NMR studies on rotamer assignment in Cα(*S*)-(acetamido)-*bm*-PNA monomer (**9**) in DMSO-d_6 with 55% Z rotamer and 45% E-rotamer,
- Solid phase synthesis of the homo oligomer of Cα(*S*)-(acetamido-C)-*bm*-PNA-T (*bm*-PNA-T_7C_5, **12**), purification and its characterization
- Fluorescence studies on monomer (**9**) and derived oligomer (**12**) exhibit excitation dependent emission in water and phosphate buffer pH 12.75.
- Intact pattern of emission band at higher excitation wavelength suggest emission from single rotamer for the oligomer of Cα(R)-(acetamido)-*bm*-PNA (jp-T_7C_5, **12**)

4.6. Experimental Section

This section describes the detailed synthesis of monomers and the spectral characterisation of the rationally designed monomers described in this chapter.

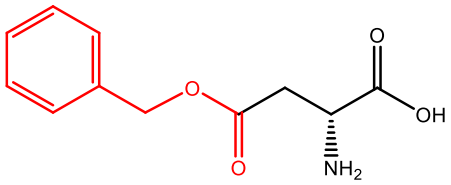
4.6.1. General information

Laboratory grade chemicals were used for setting up standard chemical reactions. Chemicals were purchased from Sigma Aldrich and Spectrochem. For moisture sensitive reactions the reagents were redistilled to ensure the maximum purity of the reagents. Solvents used for performing chemical reactions were dried using standard procedures available in the literature. Reactions were monitored using Thin Layer Chromatography (TLC). Standard UV-Lamp, Iodine spray and spraying with ninhydrin solution followed by heating the TLC plate were used to analyse the TLC. Standard procedure of work up involves extracting the desired organic compound by washing the aqueous layer with organic solvent (DCM, EtOAc) followed by washing of the combined organic layer with Brine. The combined organic layer was dried over Sodium sulphate (Na_2SO_4) or Calcium chloride ($CaCl_2$, For DCM) and evaporated under

reduced pressure to afford the desired organic compound. Precoated silica gel GF_{254} (Merck 5554) TLC plates were used for TLC analysis. Standard chemical reactions were performed in room temperature (rt), wherever required temperature was mentioned. Silica gel (60-120 or 100-120 mesh) were used to perform the column chromatography. Double distilled EtOAc, hexane, MeOH and DCM were used as a mobile phase to perform the column chromatography. ^{1}H NMR, ^{13}C {1H} NMR and 2D-NMR experiments were performed in Bruker 400 MHz, JEOL 400 MHz, Bruker 500 MHz, and Bruker 600 MHz NMR spectrometers. NMR solvents were purchased from Sigma Aldrich. Deuterated chloroform ($CDCl_3$) and DMSO-d_6 were used as an NMR solvent. Chemical shifts were reported in δ ppm with respect to the internal standard TMS or deuterated NMR solvent. Mass spectra of the reaction intermediates were recorded in an Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using TiO_2 or 2,5-dihydroxybencoic acid (DHB) as matrix. To further confirm the mass spectra of small molecules (Mass 100-1200) obtained in MALDI, High Resolution Mass Spectrometry (HRMS) was performed for each reaction intermediates. IR spectra of the samples were recorded in a Bruker IR spectrometer. Oligomers were characterised by Agilent HPLC instrument, C_{18} column and MALDI-TOF spectrometry. SHIMADJU UV-2600i instrument was used for the measurement of UV Data of the monomers and oligomers. A pair of square quartz cuvette with maximum volume 3.5 ml, path length of 10 mm (1 cm), external dimensions of 12.5×12.5 mm and height of 45 mm (4.5 cm) was used for the measurement with a wavelength range 200 nm-800 nm. Before every measurement a blank experiment was performed with the solvent of interest and a standard baseline correction was performed with respect to the solvent of interest to ensure the accuracy of the obtained data. Emission spectra were taken on a Horiba Jobin Yvon Fluorolog-3 spectrofluorometer at various excitation wavelengths. Excitation wavelength range of 280 nm to 420 nm was used for the measurement of emission spectra of monomers at 25^{o}C, emission and excitation slits were set at 2.5 nm for every measurement. Measurements were performed in a 1 cm rectangular quartz fluorescence cuvette. All spectra were normalized so that the emission maxima and minima were identical. A blank emission spectrum of the solvent was taken before the experiment to ensure the accuracy of the obtained data.

4.6.2 Experimental procedures and spectral data

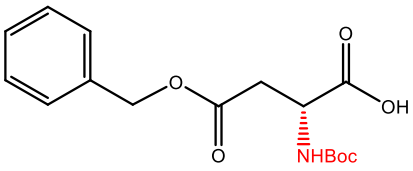
(*R*)-2-amino-4-(benzyloxy)-4-oxobutanoic acid (1)[16]



1

The titled compound was synthesized according to the experimental procedure reported in literature.[17] Original procedure was modified to obtain better yield of the (*R*)-2-amino-4-(benzyloxy)-4-oxobutanoic acid (**1**). L-aspartic acid (3 gm, 22.5 mmol), benzyl alcohol (7 ml, 67.6 mmol) and a mixture of conc. H_2SO_4 (1.3 ml) and water (0.6 ml) were mixed in round bottom flask and heated in an oil bath at 70^o C for 8 h to give a clear yellowish solution. The resulting yellow solution was cooled in room temperature. Ice cold solution of $NaHCO_3$ (4.2 gm) in 25 ml of water was added slowly to the yellowish reaction mixture. Addition of Cold $NaHCO_3$ solution to the reaction mixture results formation of white precipitate of the titled (*R*)-2-amino-4-(benzyloxy)-4-oxobutanoic acid (**1**). Filtration and washing of the white precipitate with water followed by acetone under vacuum affords the titled selective γ-benzylated aspartic acid (**1**) derivative as a white solid. The resulting compound was used for the next step without further purification (3.9 gm, Crude yield 70%).

(*R*)-4-(benzyloxy)-2-(tert-butoxycarbonyl)amino)-4-oxobutanoic acid (2)[17]

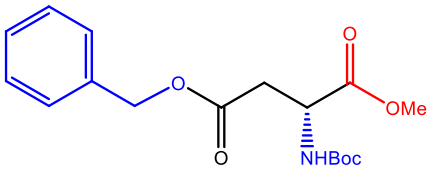


2

The titled compound was synthesized according to the experimental procedure reported in literature.[17] (R)-2-amino-4-(benzyloxy)-4-oxobutanoic acid (**1**, 3 gm ,13.4 mmol) was dissolved in a 1:1 solution of dioxane (18 ml) and water (18 ml) to which was added solid $NaHCO_3$ (15 mmol, 1.3 gm) and Boc_2O (3.7 ml, 16.1 mmol) at 0^o C. After 5 min the reaction mixture was allowed to be stirred at room temperature for overnight. After the completion of the reaction, dioxane was evaporated and the aqueous layer was acidified with aq. HCl (2N) solution to pH 2.0 and the compound was extracted with EtOAc (3 x 100 ml). The combined

organic layer was dried, filtered and concentrated to give a crude and colourless oil. The resulting Boc-protected aspartic acid derivative (**2**) was used for the next step without further purification (3.5 g, yield 78%).

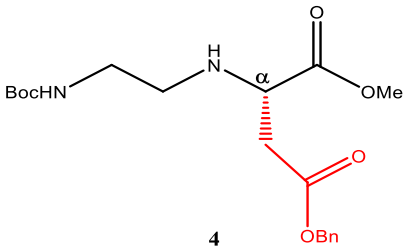
(R)-4-benzyl 1-methyl 2-((tert-butoxycarbonyl)amino)succinate (3)[18]



3

The titled compound was synthesized according to the experimental procedure reported in literature.[18] (*R*)-4-(benzyloxy)-2-((tert-butoxycarbonyl) amino)-4-oxobutanoic acid (**2**, 1.0 g 3.0 mmol) was dissolved in dry Acetone (25 ml) to which solid K_2CO_3 (640 mg, 4.6 mmol) was added. The reaction mixture was stirred for 2 min at room temperature, to which dimethyl sulphate (0.4 ml, 4.6 mmol) was added and allowed to stir at room temperature overnight. After the completion of the reaction acetone was evaporated and water was added. The required compound was extracted using EtOAc (3 x 150 ml). The combined organic layer was dried over Na_2SO_4, filtered, and concentrated in vacuo to give the orthogonally protected aspartic acid derivative **3** as a colourless oil. The crude oil was purified by flash column chromatography using 20% EtOAc in hexane to obtain the titled orthogonally protected aspartic acid derivative **3** as a colourless oil (900 mg, yield 75%), R_f = 0.8 (in 40% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.29 (m, 5H), 5.50 (d, *J* = 8.0 Hz, 1H), 5.16 – 5.07 (d, *J*= 3.3 Hz, 2H), 4.61 – 4.56 (m, 1H), 3.69 (s, 3H), 3.04 (dd, *J* = 16.9, 4.4 Hz, 1H), 2.86 (dd, *J* = 16.9, 4.8 Hz, 1H), 1.44 (s, 9H), ^{13}C NMR (101 MHz, $CDCl_3$), δ 171.5, 170.9, 155.4, 141.1, 135.4, 128.6, 128.52, 128.4, 128.3, 127.5, 127, 80.2, 66.8, 65.2, 58.6, 52.7, 50.1, 37.0, 28.3. Exact mass ($C_{17}H_{23}O_6N$ +Na) $^+$ 360.1423, observed mass 360.1421.

(*S*)-4-benzyl-1-methyl-2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)succinate (4)

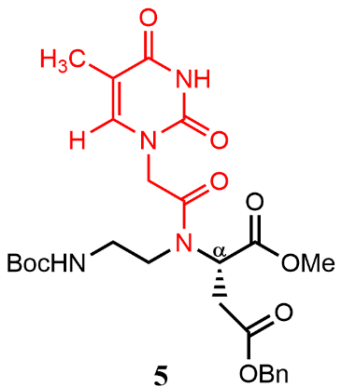


4

(*R*)-4-benzyl 1-methyl 2-((tert-butoxycarbonyl) amino) succinate **3** (3.5 gm) was dissolved in 50% TFA in DCM (10 ml) and the reaction mixture was stirred at room temperature for 3 h. The TFA was evaporated under reduced pressure and the residue was used for next step without further purification.

The free amine of orthogonally protected aspartic acid derivative **3** (3 gm, 8.5 mmol) was dissolved in dry MeOH and to that solution tert-butyl (2-oxoethyl) carbamate (4.0 g, 25.6 mmol) was added and the reaction mixture was stirred at room temperature for 5 min. The reaction mixture was cooled at 0^{o} C. DIPEA (1.8 ml, 10.2 mmol) was added to the reaction mixture and stirred at 0^{o} C for 2 h. After that $NaCNBH_3$ (500 mg 7.9 mmol) and catalytic amount of acetic acid was added to it and the reaction mixture was allowed to stir at 0^{o} C for additional 12 h. After the completion of the reaction MeOH was evaporated under reduced pressure and the analyte was extracted using EtOAc (3 x 150 ml). The combined organic layer was dried over Na_2SO_4, filtered, and evaporated under reduced pressure to give the desired product as a reddish oil. Column chromatography using 40% EtOAc in hexane give the amine backbone **4** as yellow oil (2.8 g, yield 65%), R_f = 0.35 (30% EtOAc in hexane). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 –7.29 (m, 5H), 5.12 (s, 2H), 5.03 (s, 1H), 3.67 (d, *J* = 2.6 Hz, 3H), 3.64 (d, *J* = 1.8 Hz, 1H), 3.26 (dd, *J* = 10.2, 5.2 Hz, 1H), 3.14 (ddd, *J* = 25.2, 12.3, 7.0 Hz, 2H), 2.85 – 2.77 (m, 1H), 2.75 (d, *J* = 5.6 Hz, 1H), 2.68 (d, *J* = 7.4 Hz, 1H), 2.64 (d, *J* = 7.4 Hz, 1H), 2.58 (ddd, *J* = 11.8, 7.0, 3.7 Hz, 1H), 2.40 (s, 2H), 1.42 (d, *J* = 2.7 Hz, 9H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.03, 170.7, 156.8, 156.1, 135.6, 128.6, 128.4, 128.4, 128.3, 79.6, 79.2, 66.7, 62.5, 57.3, 52.2, 47.5, 43.2, 40.2, 38.2, 28.4, 28.5. Exact mass $(C_{19}H_{28}N_2O_6 + K)^+$ 419.1584 observed mass 419.1384.

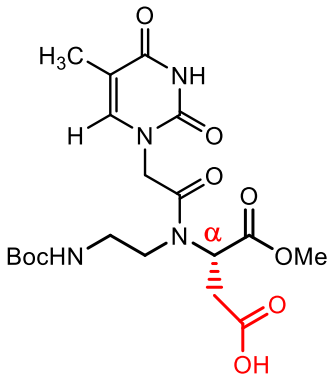
(*S*)-4-benzyl-1-methyl-2-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)succinate (5)



5

To a stirred solution of thymine acetic acid (100 mg, 0.5 mmol) in dry DMF was added N,N′-Dicyclohexylcarbodiimide (DCC, 115 mg, 0.5 mmol) followed by 3-Hydroxy-1,2,3-benzotriazin-4(3H)-one (DHbtOH, 90 mg, 0.5 mmol) was added and the reaction mixture was stirred at room temperature for 1 h. A solution of free amine backbone **4** (165 mg, 0.4 mmol) in dry DMF was added and the reaction mixture was heated at 50^{o} C for 30 h. After the completion of the reaction the desired thymine attached monomer **5** was portioned between brine solution and EtOAc and extracted using EtOAc (3 x 150 ml). The combined organic layer was dried over Na_2SO_4, filtered, and evaporated under vacuo to give mixture of a hygroscopic white solid as a crude. Column chromatography using 5% MeOH in DCM gave benzyl protected Cα-aspartic acid PNA monomer **5** as a hygroscopic white solid (180 mg, yield, 75%) R_f = 0.4 in 100% EtOAc). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.94 (s, 1H), 7.37 – 7.31 (m, 5H), 6.82 (s, 1H), 5.57 (d, *J* = 5.6 Hz, 1H), 5.20 – 5.09 (m, 2H), 4.66 (d, *J* = 16.5 Hz, 1H), 4.28 – 4.18 (m, 2H), 3.70 (s, 3H), 3.67 – 3.57 (m, 1H), 3.50 (dd, *J* = 13.4, 5.0 Hz, 2H), 3.40 (dt, *J* = 15.0, 3.8 Hz, 1H), 3.30 (dd, *J* = 17.4, 5.3 Hz, 1H), 3.26 – 3.14 (m, 1H), 3.08 – 2.98 (m, 1H), 1.89 (s, 3H), 1.43 (s, 9H), ^{13}C NMR (101 MHz**,** $CDCl_3$**),** δ 171.44, 170.49 ,170.0, 167.3, 164.1, 156.1, 150.9, 140.8, 135.6, 128.7, 128.6, 128.5, 80.1, 69.0, 58.3, 53.2, 49.4 ,48.8, 48.2, 38.8, 34.6, 28.5, 12.5. Exact mass ($C_{26}H_{34}N_4O_9$ +H)$^+$ 547.2404 observed mass 547.2403.

(*S*)-3-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)-4-methoxy-4-oxobutanoic acid (6)

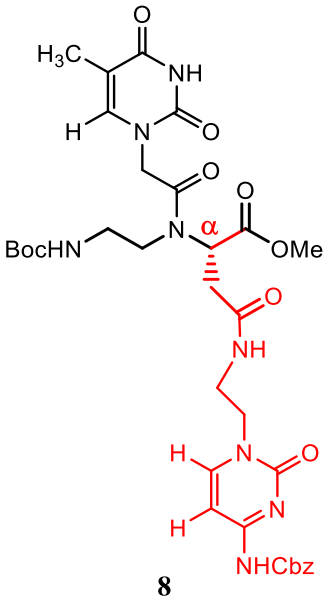


6

Thymine attached backbone **5** was dissolved in MeOH and the mixture was hydrogenated for 4.5 h in presence of Pd/C at H_2 atmosphere (1 atm). The conversion of benzyl ester to free acid was monitored by TLC. After the completion of the reaction the reaction mixture was filtered through celite using methanol and the Pd/C was quenched by adding aq. HCl (2N) solution. The crude product was obtained as a white powder by evaporating the

methanol solution. Purification of the crude compound by column chromatography gave the free acid Cα-aspartic acid PNA monomer **6** as a hygroscopic white solid (710 mg, yield, 85%) R_f = 0.3 in 100% EtOAc.^{1}H NMR (400 MHz, DMSO-d_6), δ 12.45 (s, 1H), 11.28 (s, 1H), 7.32 (s, 1H), 6.94 (s, 1H), 4.59 (d, *J* = 5.1 Hz, 1H), 4.45 – 4.36 (m, 1H), 3.57 (s, 3H), 3.44 (dd, *J* = 9.4, 4.7 Hz, 2H), 3.20 (d, *J* = 3.4 Hz, 2H), 3.03 (dd, *J* = 17.4, 7.8 Hz, 2H), 2.75 – 2.68 (m, 1H), 2.62 (dd, *J* = 17.1, 5.7 Hz, 1H), 1.75 (d, *J* = 0.9 Hz, 3H), 1.38 (s, 9H), ^{13}C NMR (101 MHz, DMSO-d_6), δ 172.04, 171.01, 170.07, 167.3, 164.4, 162.3, 155.7, 150.9, 142.01, 108.1, 79.2, 78.1, 64.9, 62.02, 57.4, 52.1, 51.6, 47.8, 34.4, 34.01, 29.6, 28.1, 25.5, 15.18, 11.9. Exact mass $(C_{19}H_{28}N_4O_9 + H)^+$ 457.1934 observed mass 475.1931.

(*S*)-methyl-4-((2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)ethyl)amino)-2-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)-4-oxobutanoate (8)

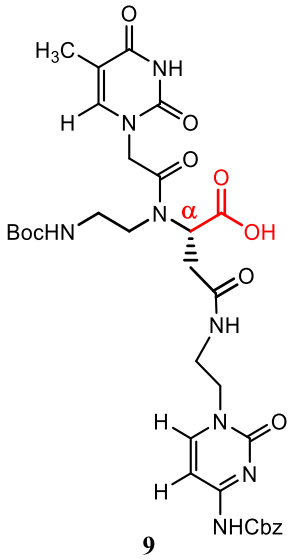


8

Free acid of Cα-aspartic acid PNA monomer **6** (100 mg, 0.2 mmol) was dissolved in dry DMF and to that solution was added N,N′-Dicyclohexylcarbodiimide (DCC, 45 mg, 0.2 mmol,) followed by 3-Hydroxy-1,2,3-benzotriazin-4(3H)-one (DHbtOH, 36 mg, 0.2 mmol) and the solution was stirred at room temperature for 1 h. 2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)ethan-1-aminium salt of trifluoro acetate **7** (130 mg, 0.30 mmol) was added and the reaction mixture was stirred at rt for additional 48 h. After the completion of the reaction the desired compound was extracted with EtOAc (3×150 ml) and washed with brine (3 x 100 ml). The combined organic layer was dried over Na_2SO_4, filtered, and evaporated under vacuo to give mixture of a white solid and yellow liquid as a crude. Column

chromatography using 5% MeOH in DCM gave the Cα(*S*)-(acetamido-C)-*bm*-PNA-T methyl ester **8** as a White hygroscopic solid (105 mg, yield, 65%) R_f = 0.3 in 5% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO-d_6), δ 11.28 (d, *J* = 7.6 Hz, 1H), 10.74 (s, 1H), 7.90 (dd, *J* = 7.3, 2.9 Hz, 1H), 7.39 (d, *J* = 5.2 Hz, 4H), 7.37 – 7.31 (m, 2H), 7.03 – 6.94 (m, 1H), 6.84 (d, *J* = 7.0 Hz, 1H), 5.18 (s, 2H), 4.74 – 4.49 (m, 2H), 4.36 (t, *J* = 6.5 Hz, 1H), 3.82 (dd, *J* = 11.2, 5.7 Hz, 2H), 3.70 (s, 1H), 3.55 (s, 2H), 3.35 (s, 3H), 3.19 (s, 2H), 2.10 (d, *J* = 13.3 Hz, 2H), 1.75 (d, *J* = 5.7 Hz, 3H), 1.38 (d, *J* = 1.7 Hz, 9H), 1.14 (s, 2H). ^{13}C NMR (101 MHz, DMSO-d_6), δ 208.4, 174.8, 174.5, 170.4 ,170.2, 169.8, 167.3, 167.19, 166.9, 164.4, 164.3, 162.9, 162.8, 155.8, 155.0, 153.2, 150.9, 150.2, 142.1, 141.8, 136.0, 128.5, 128.1, 127.9, 108.3, 108.1, 78.2, 78.1, 68.5, 66.5, 55.8, 52.1, 32.1, 30.7, 29.6, 28.2, 12.0, 11.9. Exact mass ($C_{33}H_{42}N_8O_{11}$ +H) $^+$ 727.3051 observed mass 727.3047.

(*S*)-4-((2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)ethyl)amino)-2-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)-4-oxobutanoic acid (9)

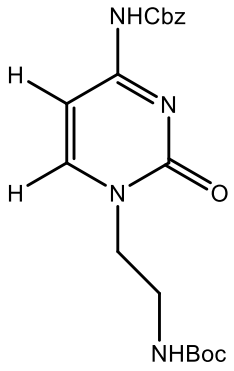


9

Cα(*S*)-(acetamido-C)-*bm*-PNA-T methyl ester **8** (500 mg, 0.7 mmol) was dissolved in THF (20 ml) and aq. NaOH (2N, 20 ml) was added to it. The reaction mixture was allowed to stir at room temperature for 4 h at 0^o C. After the completion of the reaction THF was evaporated under reduced pressure. The resulting crude was washed with EtOAc (3 x 100 ml) to remove the unreacted Cα(*S*)-(acetamido-C)-*bm*-PNA-T methyl ester **8**. Combined aqueous solution was acidified with aq. HCl (2N) to pH 2.0 and the product was extracted with EtOAc (3 x 150 ml). The combined organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to give an oily product. This was purified by column chromatography

using 10% MeOH in DCM to give the desired Cα(*S*)-(acetamido-C)-*bm*-PNA-T monomer **9** as white hygroscopic solid (370 mg, yield, 70%) R_f = 0.2 in 20% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO), δ 11.29 (d, *J* = 14.0 Hz, 1H), 7.91 – 7.83 (m, 1H), 7.39 (d, *J* = 5.9 Hz, 4H), 7.35 (dd, *J* = 5.6, 2.9 Hz, 2H), 7.01 (d, *J* = 4.9 Hz, 0.5H), 6.98 – 6.91 (m, 1H), 6.82 (s, 0.5H), 5.18 (s, 2H), 4.66 (dd, *J* = 22.6, 6.0 Hz, 2H), 3.82 (dd, *J* = 41.7, 9.9 Hz, 4H), 3.33 (s, 2H), 3.21 – 3.10 (m, 2H), 3.08 – 2.95 (m, 2H), 1.73 (d, *J* = 8.0 Hz, 3H), 1.37 (d, *J* = 6.0 Hz, 9H). ^{13}C NMR (101 MHz, DMSO), δ 172.3, 169.3, 168.8, 167.6, 164.4, 162.8, 155.2, 153.1, 151.1, 150.5, 149.7, 142.2, 135.9, 135.8, 128.5, 128.2, 128.0, 127.9, 115.7, 108.1, 108.0, 66.5, 49.2, 28.2, 28.1, 11.9. Exact mass $(C_{32}H_{40}N_8O_{11} + Na)^+$ 735.2714 observed mass 735.2715.

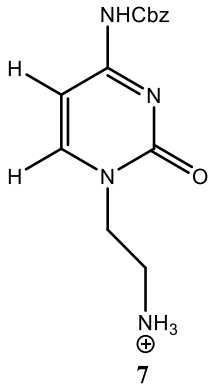
2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)ethan-1-aminium (7)[19]



13

The title compound **7** was synthesized from the Boc-Protected cytosine amine **13** by modifying the literature available procedure.[19] N^4-Cbz- C (2 g, 8.1 mmol) was dissolved in dry DMF (40 ml) and K_2CO_3 (1 gm, 8.1 mmol) was added to that mixture and stirred at room temperature for 5 min. tert-butyl (2-bromoethyl) carbamate (1.8 g, 8.2 mmol) and tert-butylammonium iodide (TBAI, 300 mg, 0.8 mmol) was added to the reaction mixture and heated at 70^o C for 5 h. It was quenched by addition of brine solution and the desired product was extracted using EtOAc (3 x 150 ml). The combined organic layer was washed with brine and dried over anhydrous Na_2SO_4. Evaporation of the EtOAc under reduced pressure resulted in a yellow oil residue. This was purified by column chromatography using 20% MeOH in DCM to give the Boc-Protected cytosine amine **13** as white solid (yield, 63%), R_f = 0.35 in 70% EtOAc/Hexane. ^{1}H NMR (400 MHz, DMSO), δ 10.76 (s, 1H), 7.83 (d, *J* = 7.3 Hz, 1H), 7.44 – 7.32 (m, 5H), 6.95 (s, 1H), 6.90 (t, *J* = 5.1 Hz, 1H), 5.18 (s, 2H), 3.78 (d, *J* = 5.6 Hz, 2H), 3.22 (d, *J* = 5.3 Hz, 2H), 1.32 (s, 9H), ^{13}C NMR (101 MHz, DMSO), δ 162.8, 155.7, 155,

153.2, 150.1, 136.1, 128.5, 128.1, 127.9, 93.7, 77.8, 66.4, 49.8, 28.1. Exact mass ($C_{19}H_{24}N_4O_5$ +H)$^+$ 389.1825 observed mass 389.1832.



7

The Boc-protected cytosine amine **13** (1 g) was dissolved in 50% TFA in DCM and stirred at rt for 6 h. TFA was evaporated under reduced pressure to give the 2-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)ethan-1-aminium **7** as yellow oil. Addition of cold diethyl ether gave the desired amine as hygroscopic white solid. The obtained hygroscopic solid was used for the next step without any further characterisation and purification.

4.7. References

1. a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497- 1501. b) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordon, B.; Nielsen, P. E. *Nature*, **1993**, *365*, 566-568. c) Jensen, K. K.; Qrum, H.; Nielsen, P. E.; Norden, B. *Biochemistry* **1997**, *36*, 5072-5077.

2. a) Hoogsteen, K. *Acta. Cryst*.,**1963**, *65*, 907. b) Felsenfeld, G.; Davies, R. D.; Rich, A. *J. Am. Chem. Soc*., **1957**, *79*, 2023-2024. (c) Le Deon, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucleic Acid Res*., **1987**, *15*, 7749-7760. (d) Helene, C.; Toulene, J. J: *Oligonucleotides: Antisense Inhibitors of Gene Expression*, McMillan Press, London, **1989**. (e) Moser, H. E.; Dervan, P. B. *Science.*, **1987**, *238*, 645-650. (f) Felsenfeld, G.; Rich, A. *Biochem. Biophys. Acta*., **1957**, *26*, 457- 468. (g) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*., **1991**, *254*, 1497-1501. (h) Nielsen, P. E.; Egholm, M.; Buchardt, O. J. *Mol. Recogn*., **1994**, *7*, 165-170.

3. Branda, N.; Kurz, G.; Lehn, J.-M. *Chem. Commun*., **1996**, *2*, 2443-2444. b) Marsh, A.; Nolen, E. G.; Gardinier, K. M.; Lehn, J.-M. *Tetrahedron Lett*., **1994**, *35*, 397- 400.

4. 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.

5. 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) Mascal, Mark.; Hext, Nicholas M.; Warmuth, Ralf.; Moore, Madeleine H.; Turkenburg, Johan P. *Angew. Chem. Int. Ed*., **1996,** *35*, 2204-2206.

6. Asadi, A.; Patrick, B. O.; Perrin, D. M. *J. Org. Chem*., **2007**, *72*, 466- 475.

7. (a) Moralez, J. G.; Raez, 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.; Thiyagarajan, P. *Proc. Natl. Acad. Sci.U.S.A.*, **2002**, *99*, 6487-6492.

8. (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.

9. Chen, D.; Meena; Sharma, S. K.; McLaughlin, L. W. *J. Am. Chem. Soc*., **2004**, *126*, 70-71.

10. Chen, H.; Meena; McLaughlin, L. W. *J. Am. Chem. Soc*., **2008**, *130*, 13190-13191.

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

12. a) Gupta, M. K..; Madhanagopal, B. R. Datta, D. and Ganesh, K. N. *Org. Lett*., **2020**, *22*, 5255−5260. b) Gupta, M. K.; Madhanagopal, B. R. and Ganesh, K. N. *J. Org. Chem*., **2021**, *86*, 414−428.

13. Bhingardeve, P.; Madhanagopal, B. R.; and Ganesh, K. N. *J. Org. Chem.,* **2020**, *85*, 13680−13693.

14. Bhingardeve, P.; Jain, P.; and. Ganesh, K. N. *ACS Omega*., **2021**, *6*, 19757−19770.

15. Todkari, I.; Gupta, M. K.; and Ganesh, K. N. *Chem. Commun*., **2022**, *58*, 4083–4086.

16. Asano, Y.; Umezaki, M.; Li, Y.F.; Tsubota, S.; Lübbehüsen., T.L. *Journal of Molecular Catalysis B.*, **2001,** *12*, 53–59.

17. Eldo, J.; Cardia, J.P.; O'Day, E.M.; Xia, J.; Tsuruta, H.; and Kantrowitz, E.R. *J. Med. Chem*., **2006**, *49*, 5932-5938.

18. Chakraborty, S.; and Tai, D.F. *Tetrahedron Letters*., **2014**, *55*, 2274–2276.

19. Narayanaswamy, N.; Avinash, M.B.; and Govindaraju, T. *New J. Chem*., **2013**, *37*, 1302.

20. a) Corradini, R.; Sforza, S.; Dossena, A.; Palla, G.; Rocchi, R.; Filira, F.; Nastri, F.; and Marchelli, R. *J. Chem. Soc., Perkin Trans*., **2001,** *1*, 2690–2696. b) Mali, S.M.; Kumar, G.M.; Katariya, M.M.; and Gopi, H.N. *Org. Biomol. Chem*., **2014**, *12*, 8462.

21. Orain, D.; Ellard, J.; Bradley, M. *J. Comb. Chem*., **2002**, *4*, 1-16.

22. Kulkarni, P.; Datta, D.; Ramabhadran, R.O.; and Ganesh, K.N. *Org. Biomol. Chem*., **2021**, *19*, 6534.

23. a)Lakowicz, J.R and Nakamoto, S.K. *Biochemistry.*, **1984**, *23*, 3013–3021. (DOI:10.1021/bi00308a026) b) Kabir, M.L.; Wang, F.; and Clayton, A.HA. *Int. J. Mol. Sci*., **2021**, *22*, 2582. c) Chattopadhyay. A. and Haldar, S. *Accounts of Chemical*

Research., **2014**, *47*, 12-19. d) Chattopadhyay, A and Mukherjee, S. *J. Phys. Chem. B*., **1999**, *103*, 8180-8185. d). Berger, O.; Abramovich, L.A.; Sakin, M.L.; Grunwald, A.; Peer, Y.L.; Bachar , M.; Buzhansky, L.; Mossou, E.; Forsyth, V.T.; Schwartz, T.; Ebenstein, Y.; Frolow,F.; Shimon, L.J.W.; Patolsky, F.; and Gazit, E. *Nature Nanotechnology*, **2015**, *10*, 353–360.

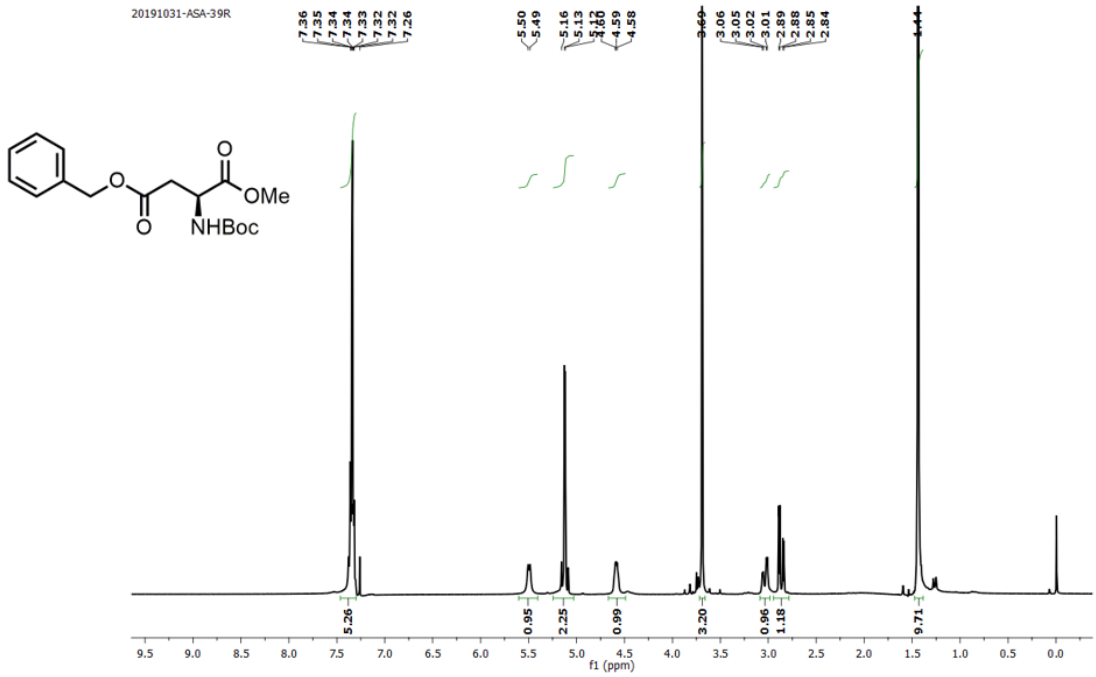
24. a) Zhang, H.; Zhao, Z.; McGonigal, P.R.; Ye, R.; Liu, S.; Lam, J.W.Y.; Kwok, R.T.K.; Yuan, W.Z.; Xie, J.; Rogach, A.L.; Tang, B.Z. *Materials Today*., **2020**, *32*, 275-292. b) Larson, C.L.; Tucker, S.A. *Appl. Spectrosc*., **2001**, *55*, 679–683. c) Varnavski, O.; Ispasoiu, R.G.; Balogh, L.; Tomalia, D.; Goodson. T. *J. Chem. Phys*., **2001**, *114*, 1962–1965. d) Lee, W.I.; Bae, Y.; and Bard, A.J. *J. Am. Chem. Soc*., **2004**, *126*, 8358–8359. e) Liu, X.; Zeng, Y.; Liu, J.; Li, P.; Zhang, D.; Zhang, X.; Yu, T.; Chen, J.; Yang, G.; and Li, Y. *Langmuir*., **2015,** *31* ,4386–4393. f) Al-Jamal, K.T.; Ruenraroengsak, P.; hartel, N.; & Florence, A.T. *J. Drug Target*., **2006**, *14*, 405–412. g) Marek, P.; Gupta, R.; and Raleigh, D.P. *Chembiochem*., **2008**, *9*, 1372–1374. h) Turoverov, K.K.; Kuznetsova, I.M. *J. Fluoresc*., **2003**, *13*, 41–57. i) Wang, H.; Zhao, E.; Lam, J.M.Y.; Tang, B.Z. *Mater. Today*., **2015**, *18*, 365–377. j) Zhang, Y.; Shen, P.; He, B.; Luo, W.; Zhao, Z; and Tang, B.Z. *Polym. Chem*., **2018**, *9*, 558–564. k) Schilter, D. *Nat. Rev. Chem*., **2017**, *1*, 0097. l) Tomalia, D.A.; Maculewicz, B.K.; Johnson, K.A.M.; Brinkman, H.F.; Janaszewska, A.; Hedstrand , D.M. *Prog. Polym. Sci*., **2019**, *90*, 35–117. m) Zhou, Q.; Wang, Z.; Dou, X.; Wang, Y.; Liu, S.; Zhang, Y.; and Yuan, W. Z. *Mater. Chem. Front*., **2019**, *3*, 257–264. n) Zhang, T.; Zhao, Z.; Ma, H.; Zhang, Y.; and Yuan, W.Z. *Chem. Asian J*., **2019**, *14*, 884–889. o) Yuan, W.Z.; Zhang, Y. *J. Polym. Sci. A*., **2017**, *55*, 560–574. p) Konopka, M; Janaszewska, A.; and Maculewicz, B.K. *Polymers*., **2018**, *10*, 540.

25. a) Qian, H.; Cousins, M.E.; Horak, E.H.; Wakefield, A.; Liptak, M.D.; and Aprahamian, I. *Nat. Chem*., **2017**, *9*, 83–87. b) Guo, J.; Fan, J.; Lin, L.; Zeng, J.; Liu, H.; Wang, C.K.; Zhao, Z.; and Tang, B.Z. *Adv. Sci.*, **2019**, *6*, 1801629. c) He, Z.; Zhao, W.; Lam, J.W.Y.; Peng, Q.; Ma, H.; Liang, G.; Shuai, Z.; and Tang, B.Z. *Nat Commun*., **2017,** *8*, 416. d) halpern, A.M.; Roebber, J.; and Weiss, K. *J. Chem. Phys*, **1968**, *49*, 1348–1357. e) Ohrn, A. and Karlstro, G. *J. Phys. Chem. A*., **2006**, *110*, 1934–1942. f) N.J. Turro *et al., Modern Molecular Photochemistry of Organic Molecules, University Science Books, Sausalito, Calif., 2010.* h) I.V. Alabugin, *Stereoelectronic Effects: A Bridge Between Structure and Reactivity, John Wiley & Sons, Ltd., Chichester, 2016*.

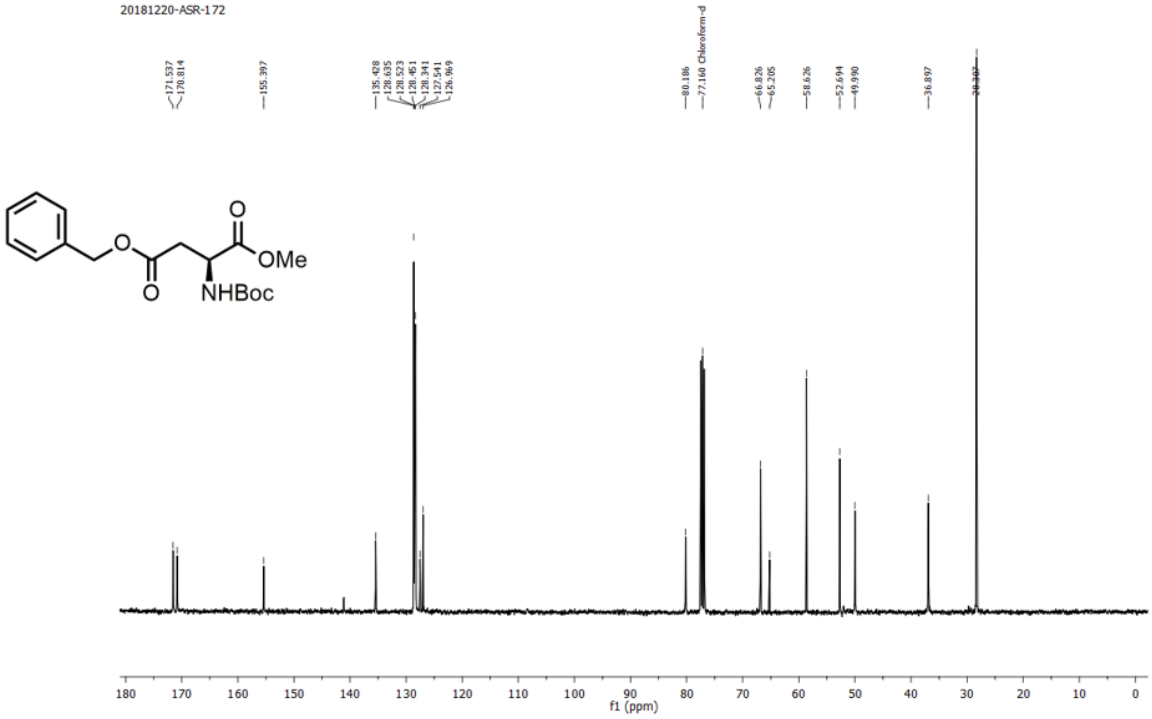
4.8 Characterisation data of the synthesized compounds and monomers (Appendix -III)

Entry	Table of contents	Page no
1	^{1}H NMR, ^{13}C NMR, and HRMS spectra of the monomers	273-283
2	HPLC and MALDI of the oligomer	284

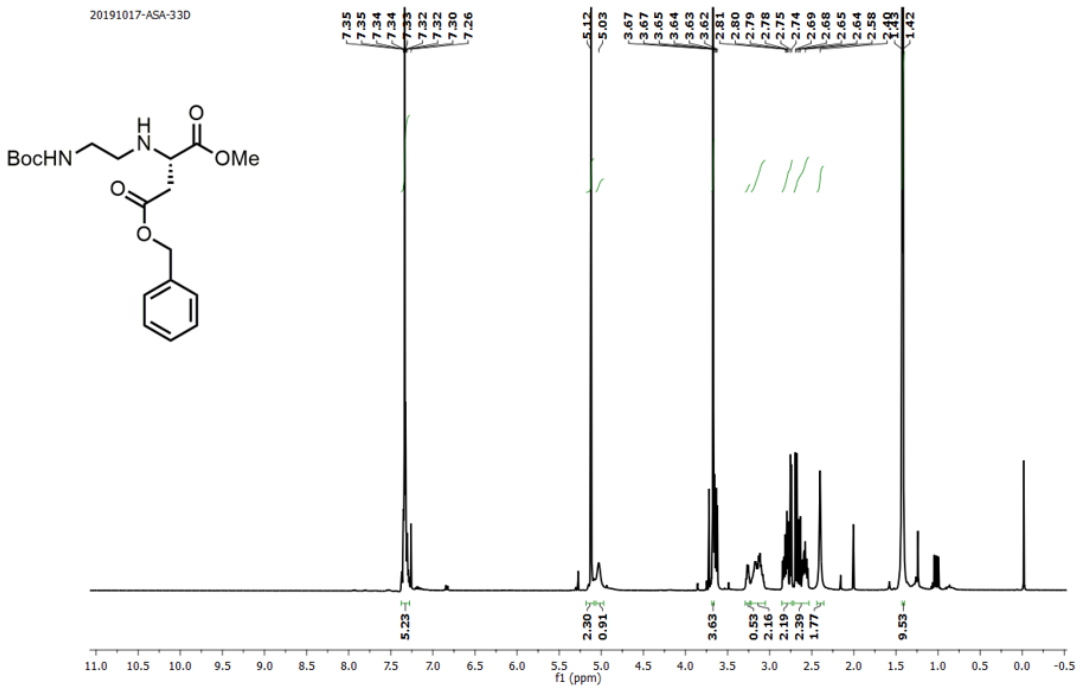
^{1}H NMR of compound 3 in $CDCl_3$ (400 MHz)



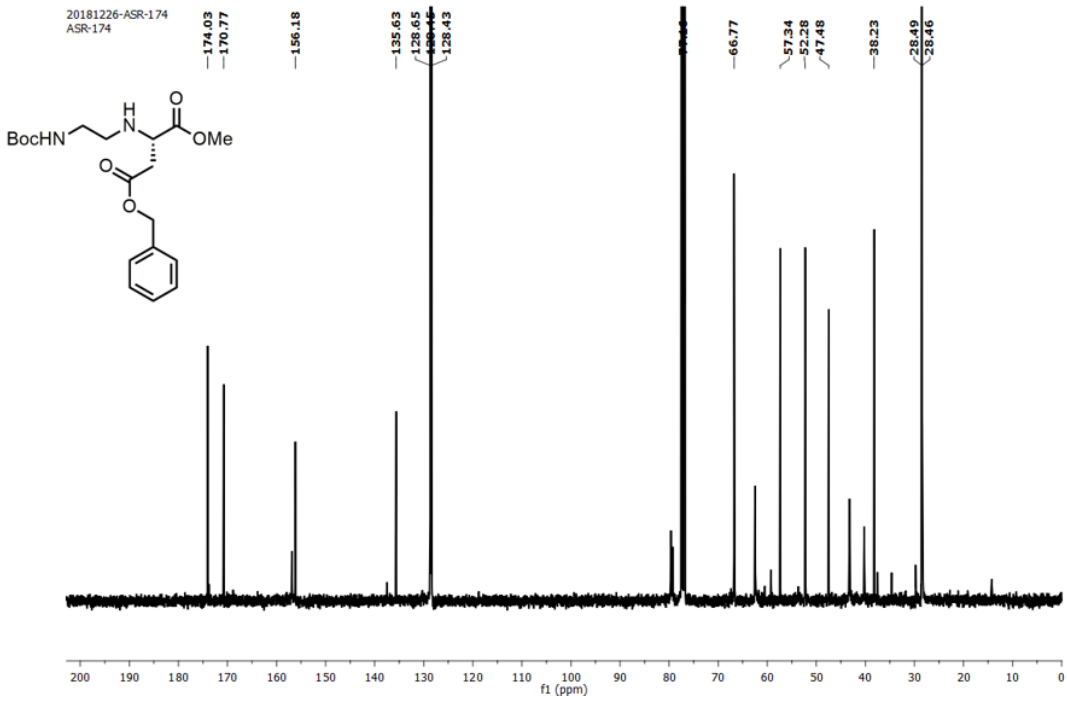
^{13}C {1H} NMR of compound 3 in $CDCl_3$ (101 MHz)



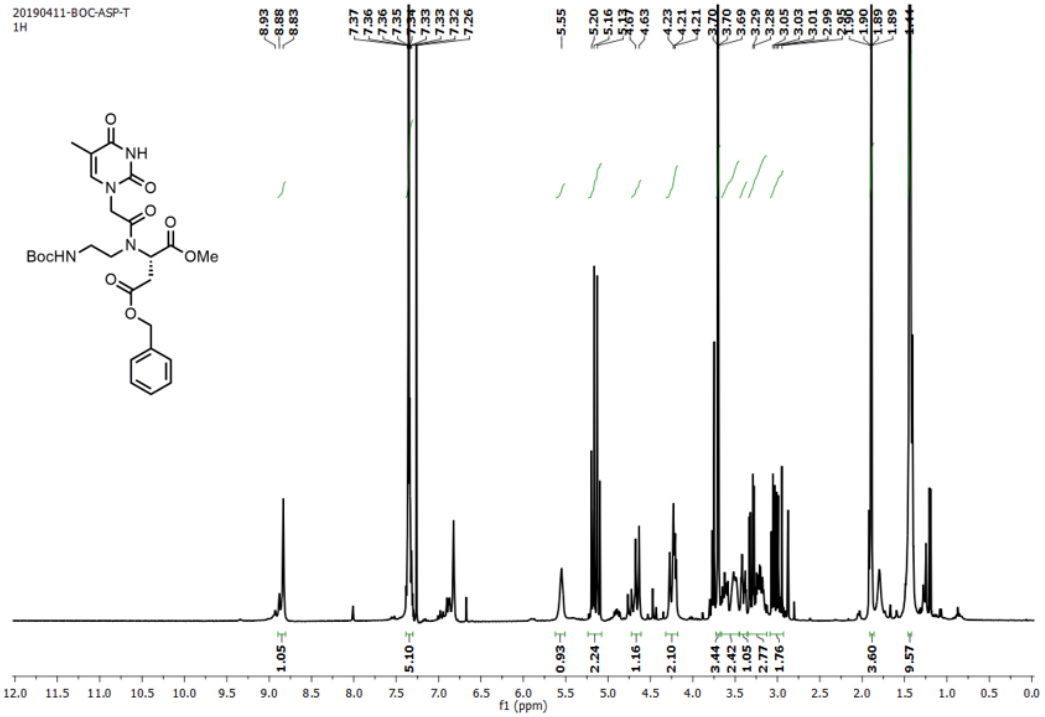
^{1}H NMR of compound 4 in $CDCl_3$ (400 MHz)



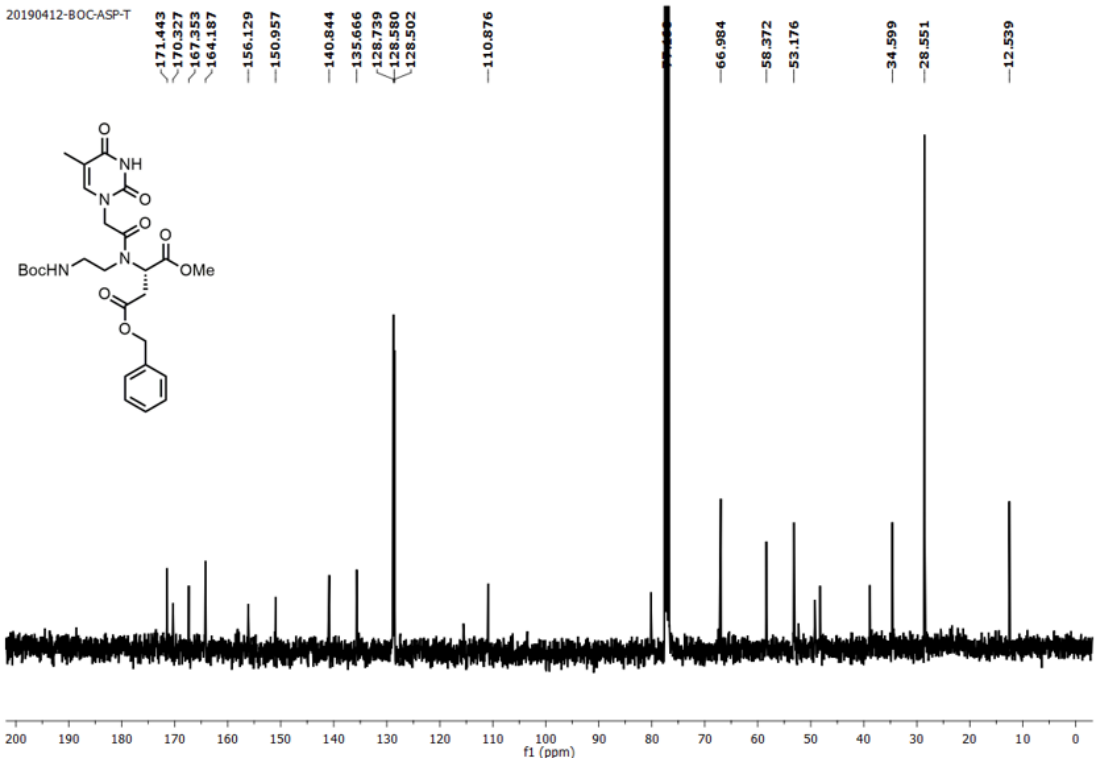
^{13}C {1H} NMR of compound 4 in $CDCl_3$ (101 MHz)



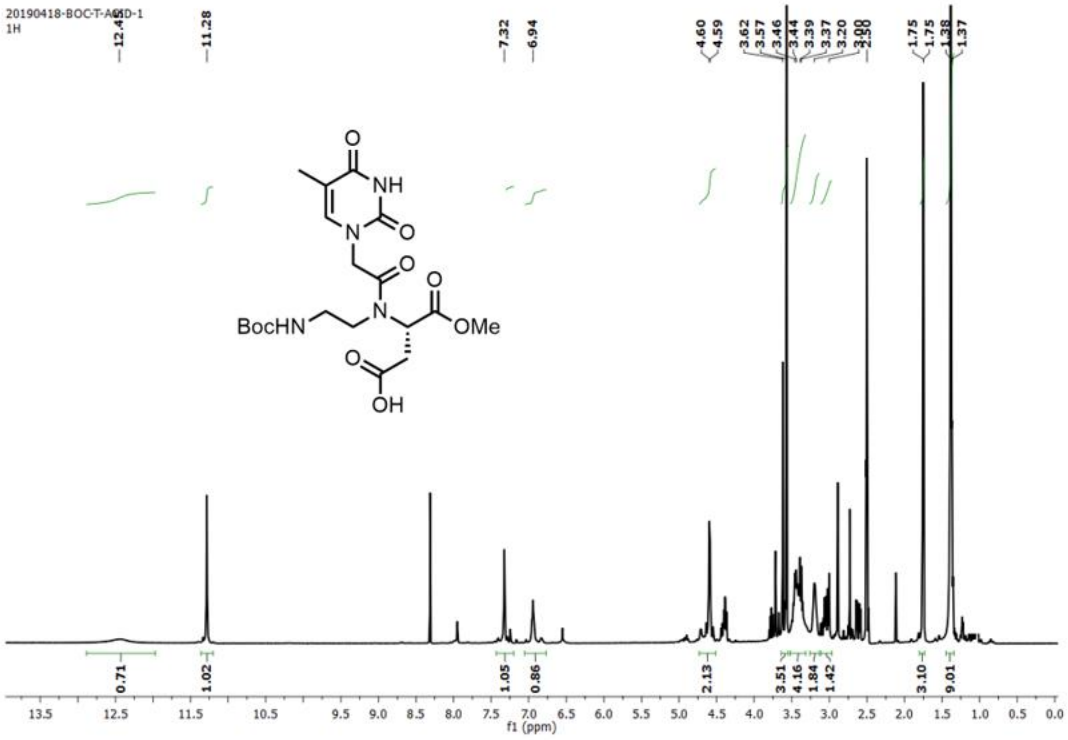
^{1}H NMR of compound 5 in $CDCl_3$ (400 MHz)



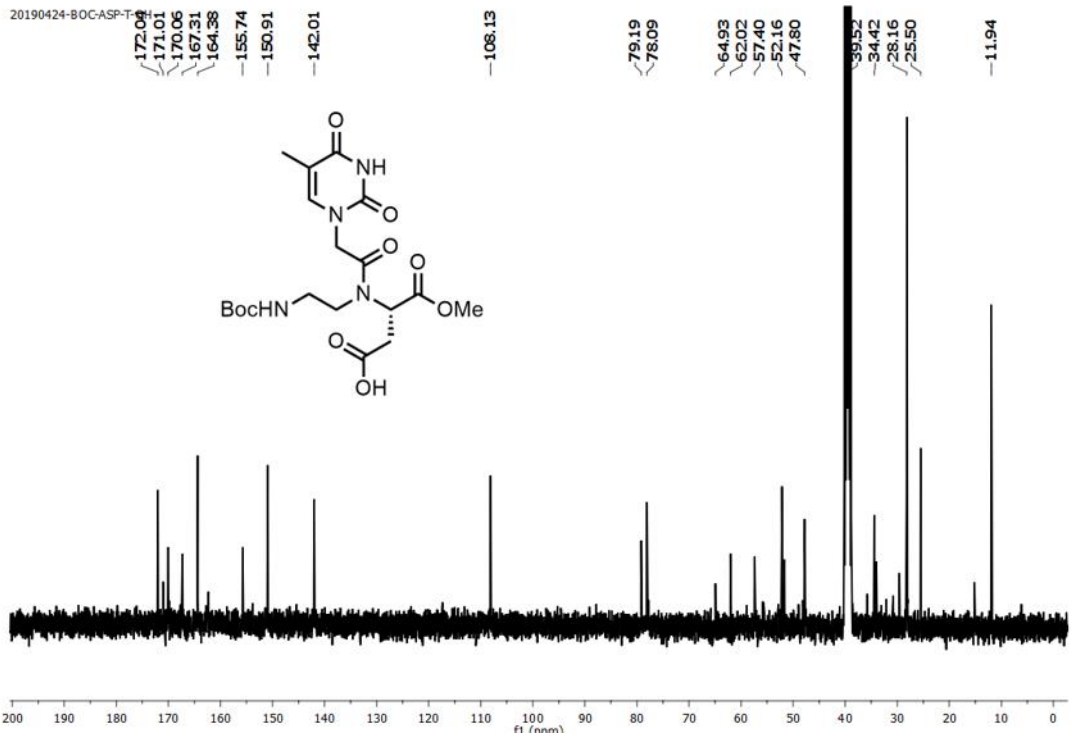
^{13}C {1H} NMR of compound 5 in $CDCl_3$ (101 MHz)



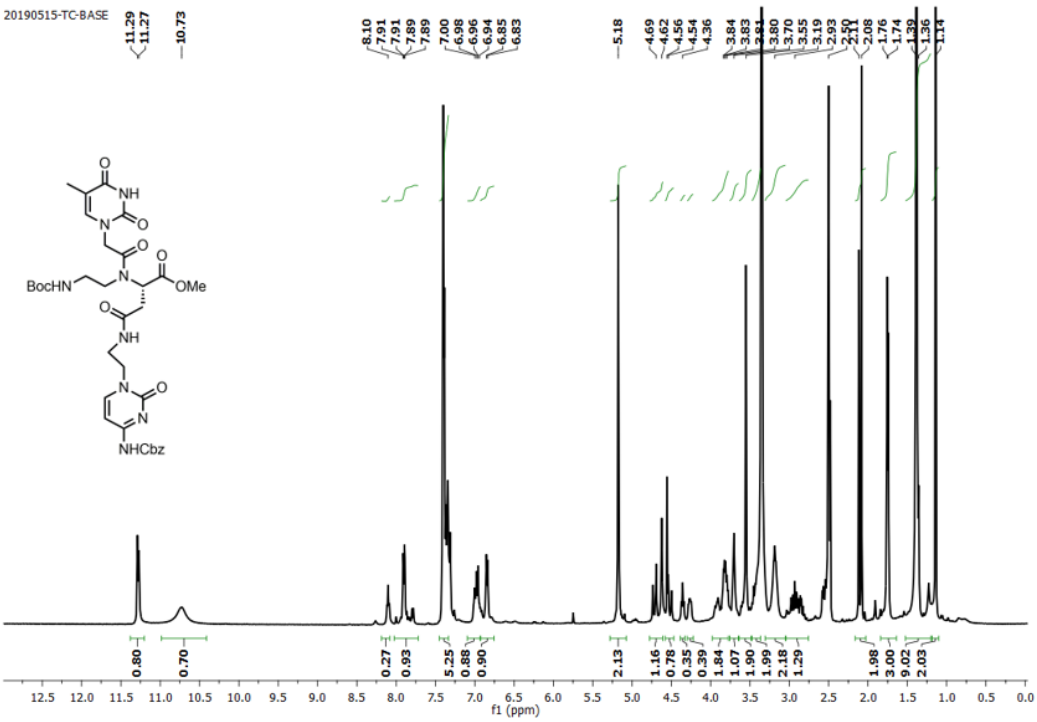
^{1}H NMR of compound 6 in DMSO-d$_6$ (400 MHz)



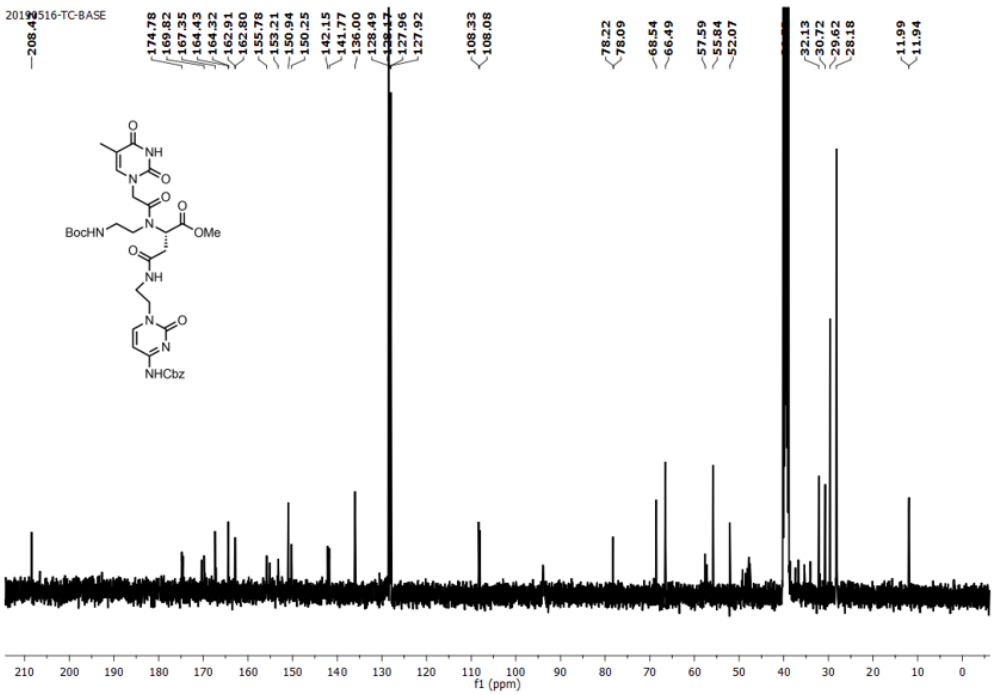
^{13}C {1H} NMR of compound 6 in DMSO-d$_6$ (101 MHz)



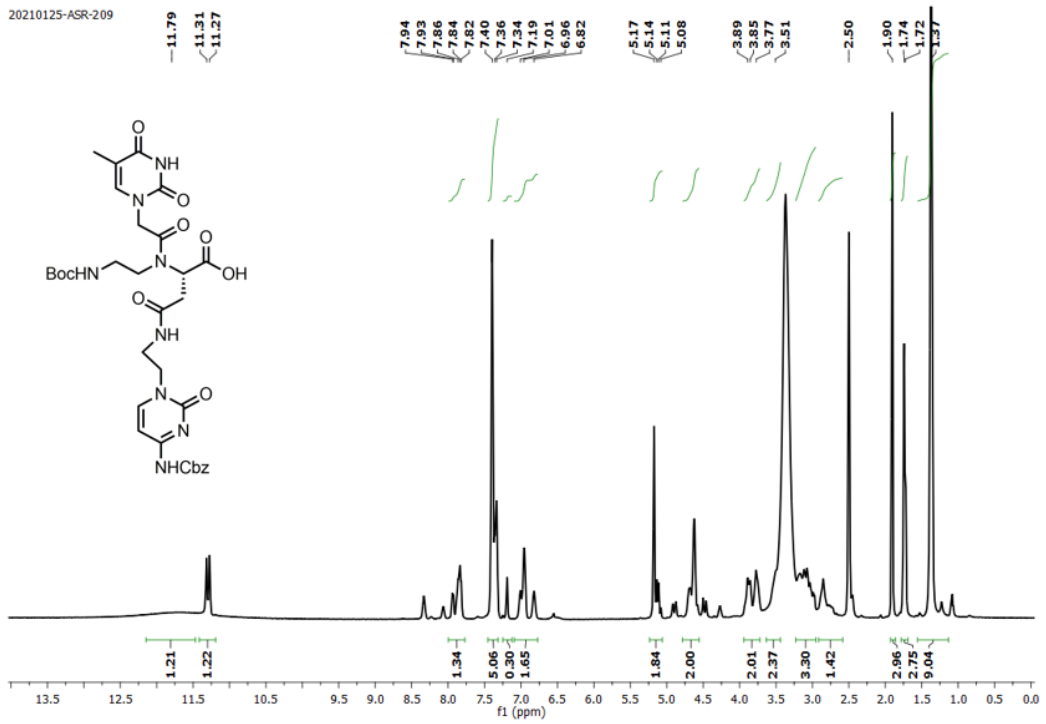
^{1}H NMR of compound 8 in DMSO-d$_6$ (400 MHz)



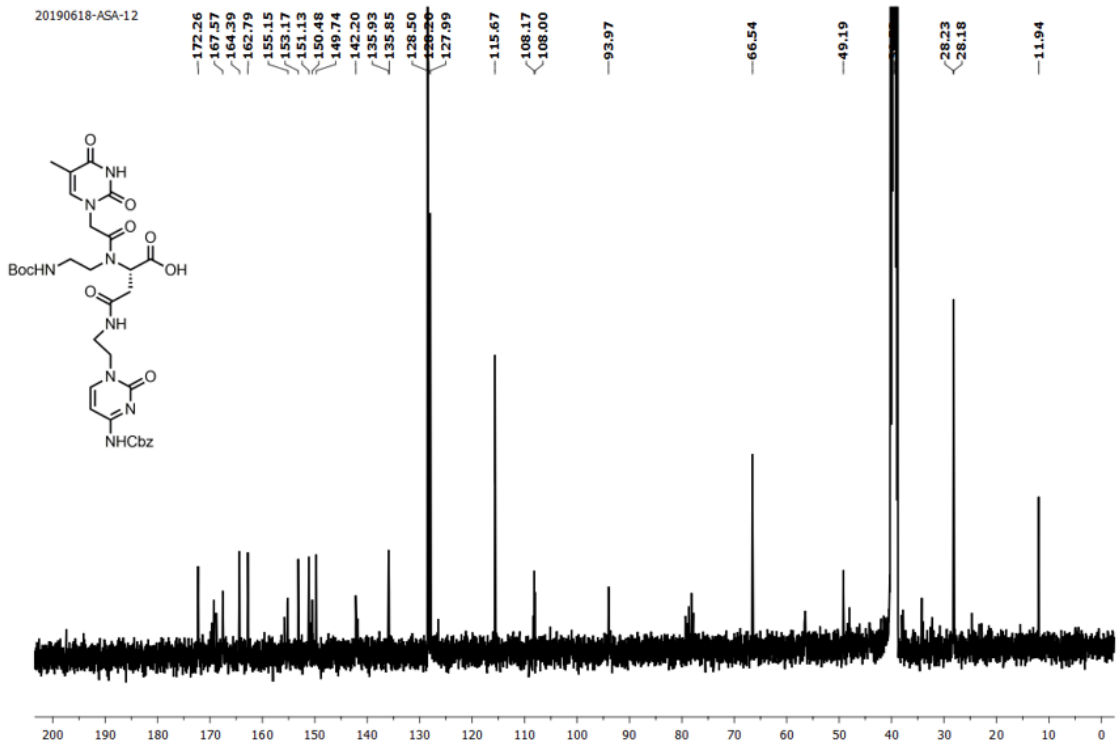
^{13}C {1H} NMR of compound 8 in DMSO-d$_6$ (101 MHz)



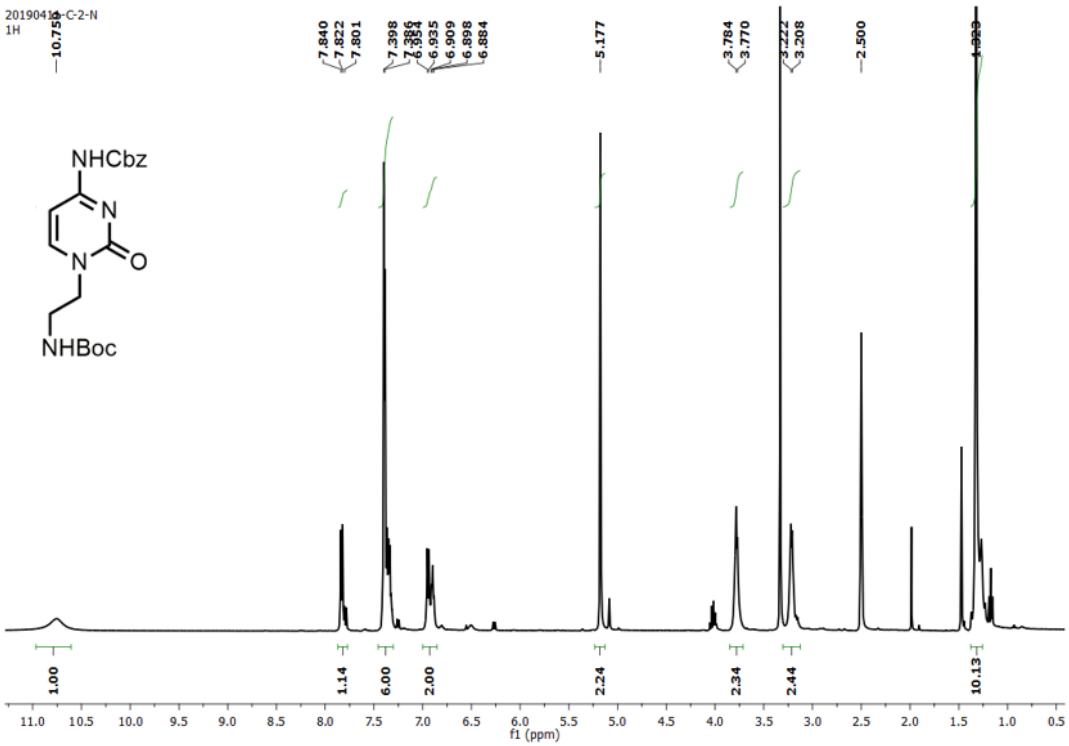
^{1}H NMR of compound 9 in DMSO-d$_6$ (400 MHz)



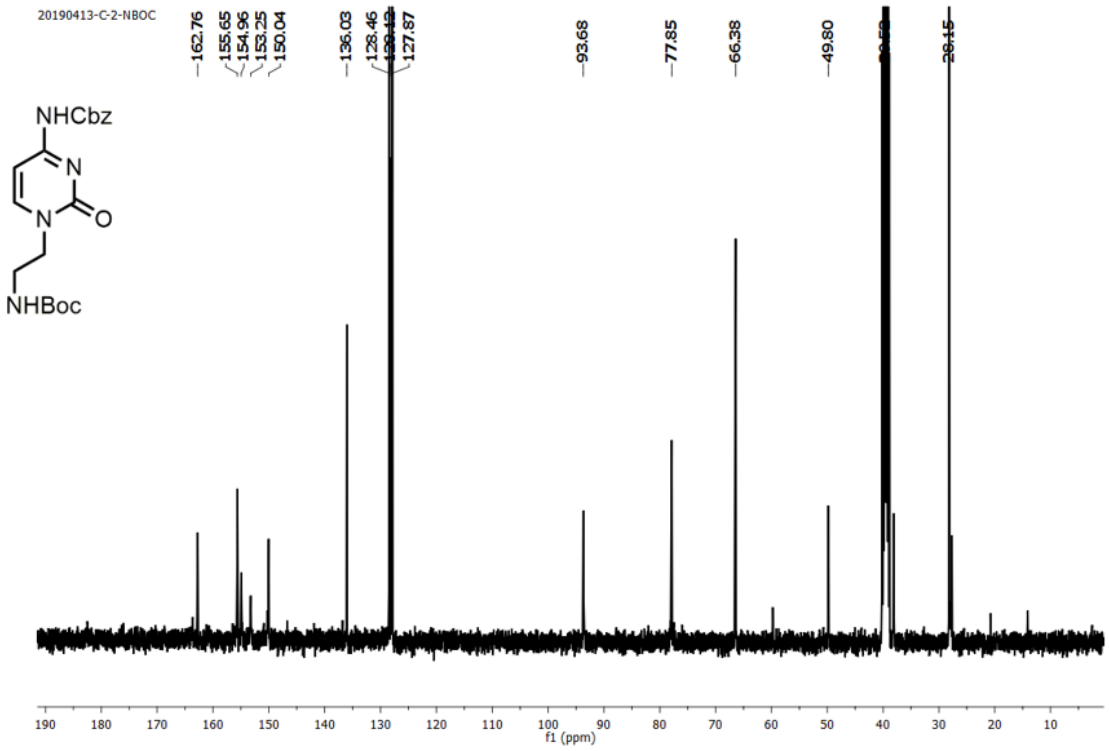
^{13}C {1H} NMR of compound 9 in DMSO-d$_6$ (101 MHz)



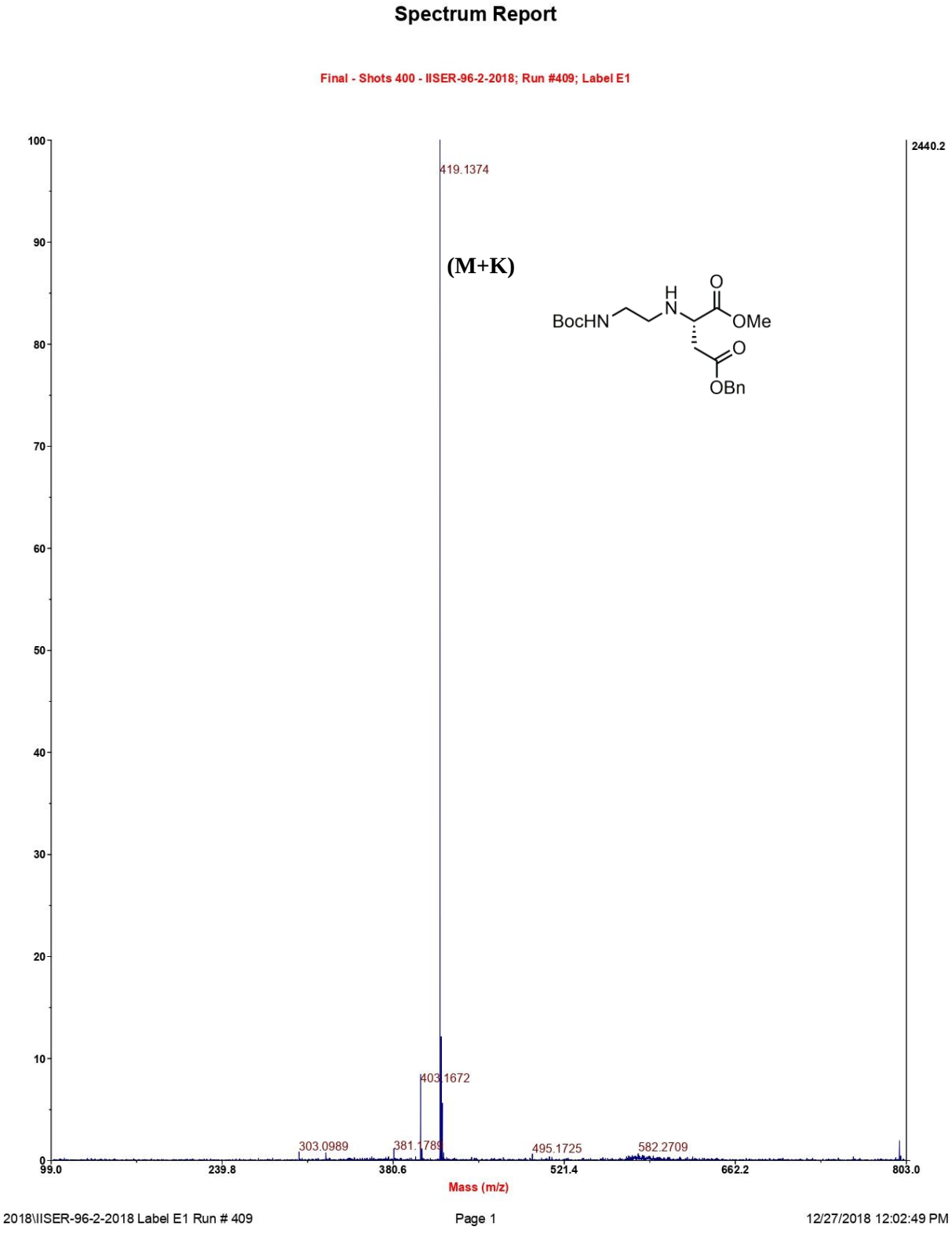
^{1}H NMR of compound 13 in DMSO-d$_6$ (400 MHz)



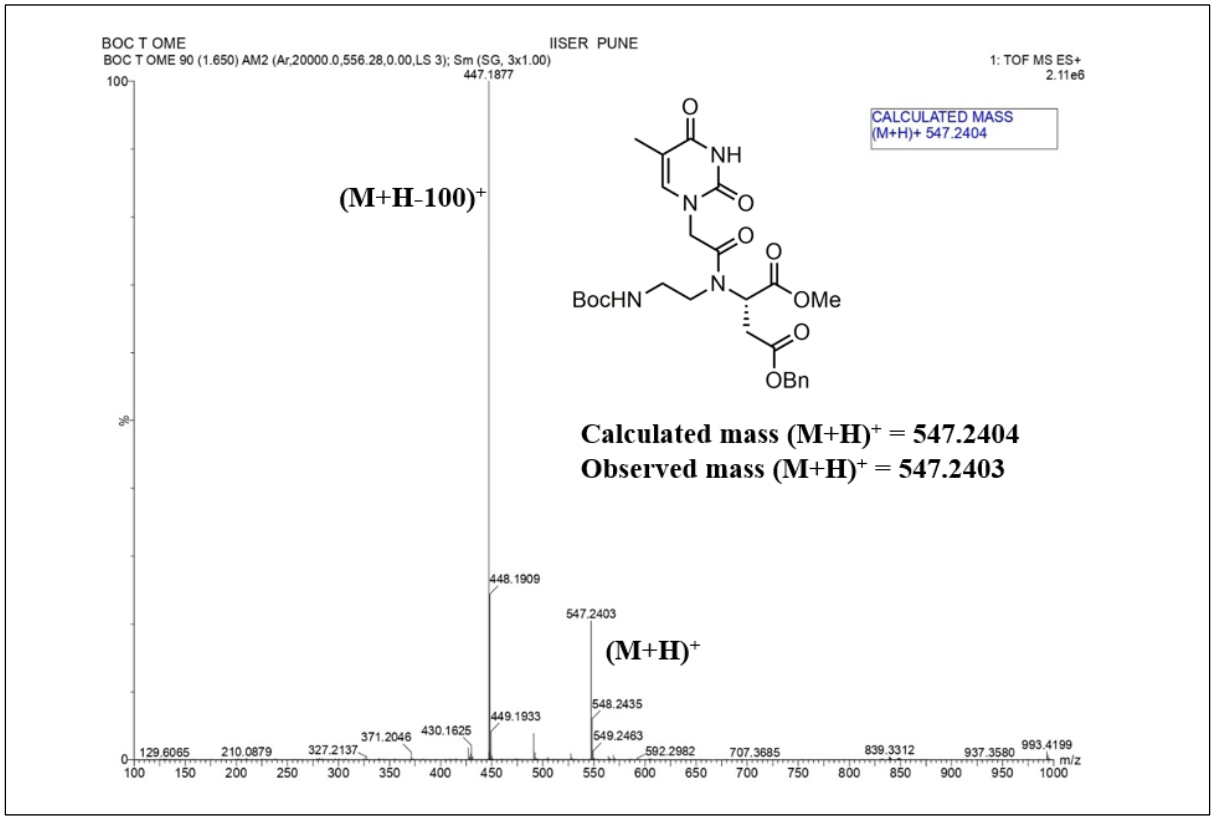
^{13}C {1H} NMR of compound 13 in DMSO-d$_6$ (101 MHz)



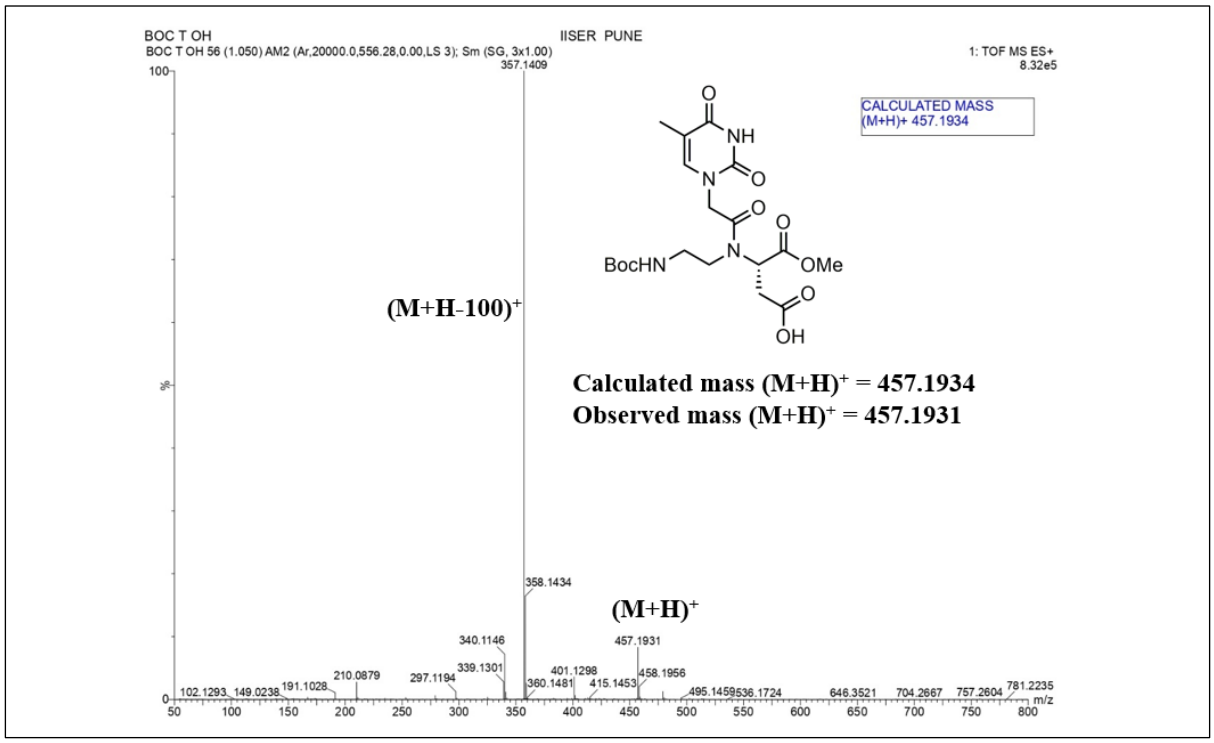
MALDI of compound 4



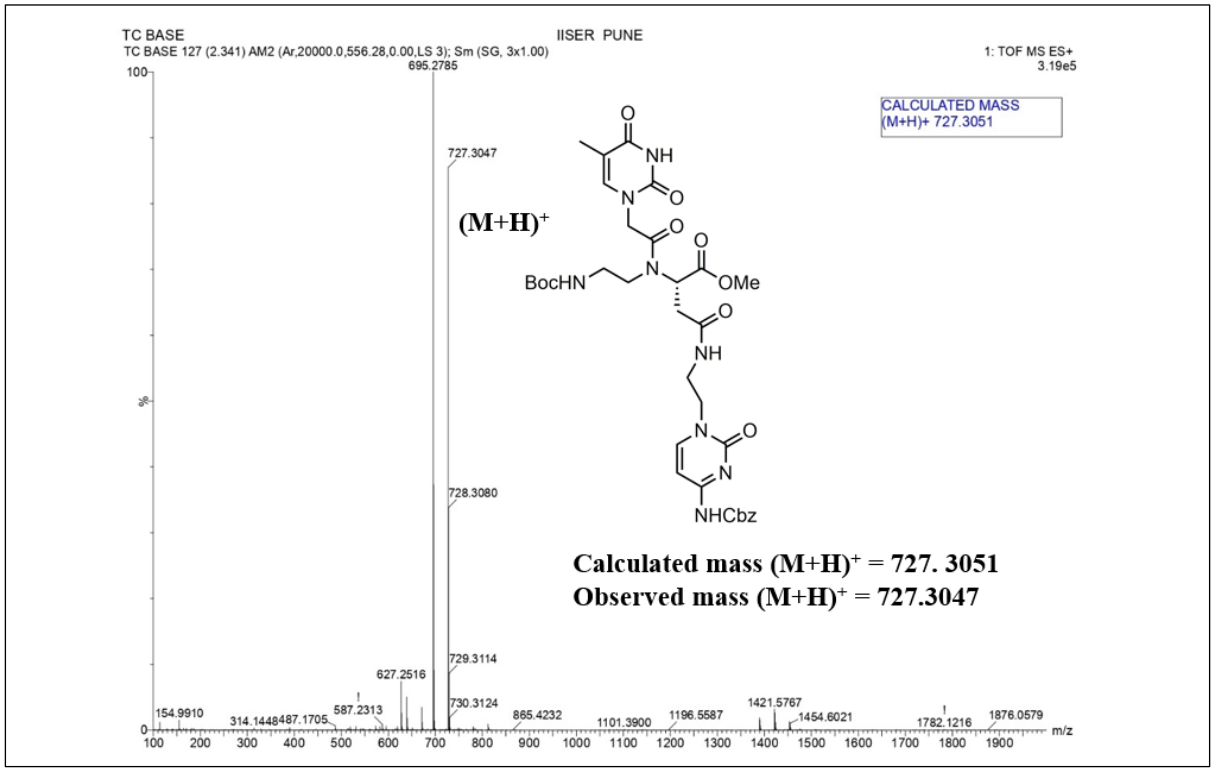
HRMS of compound 5



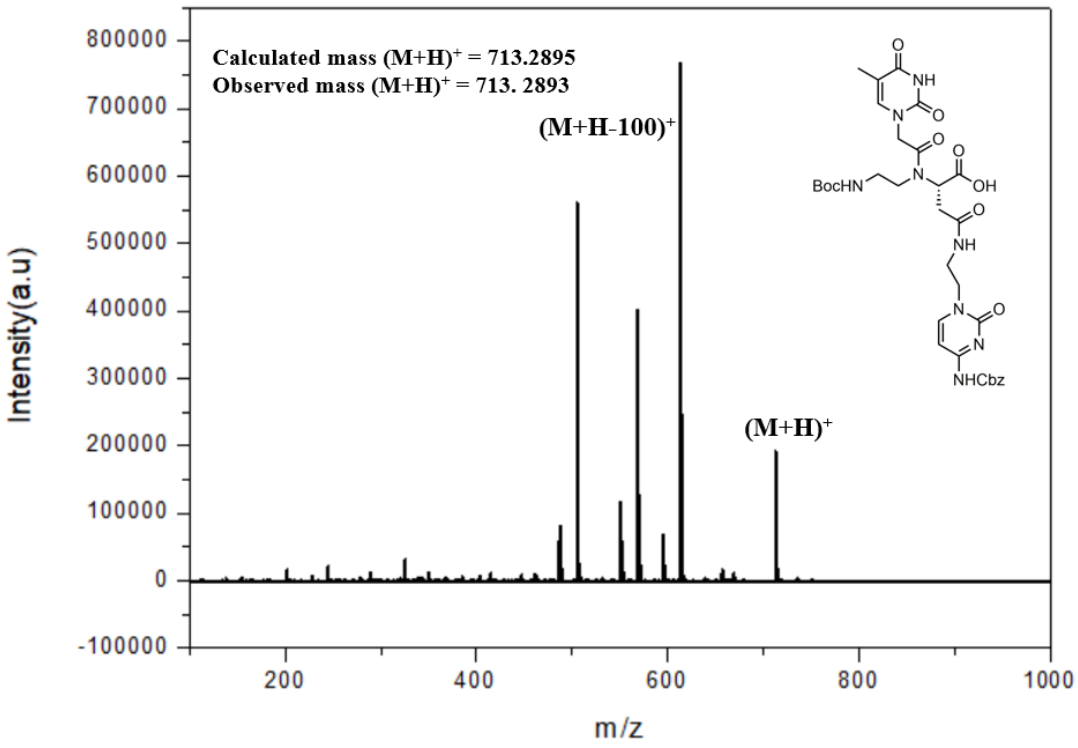
HRMS of compound 6



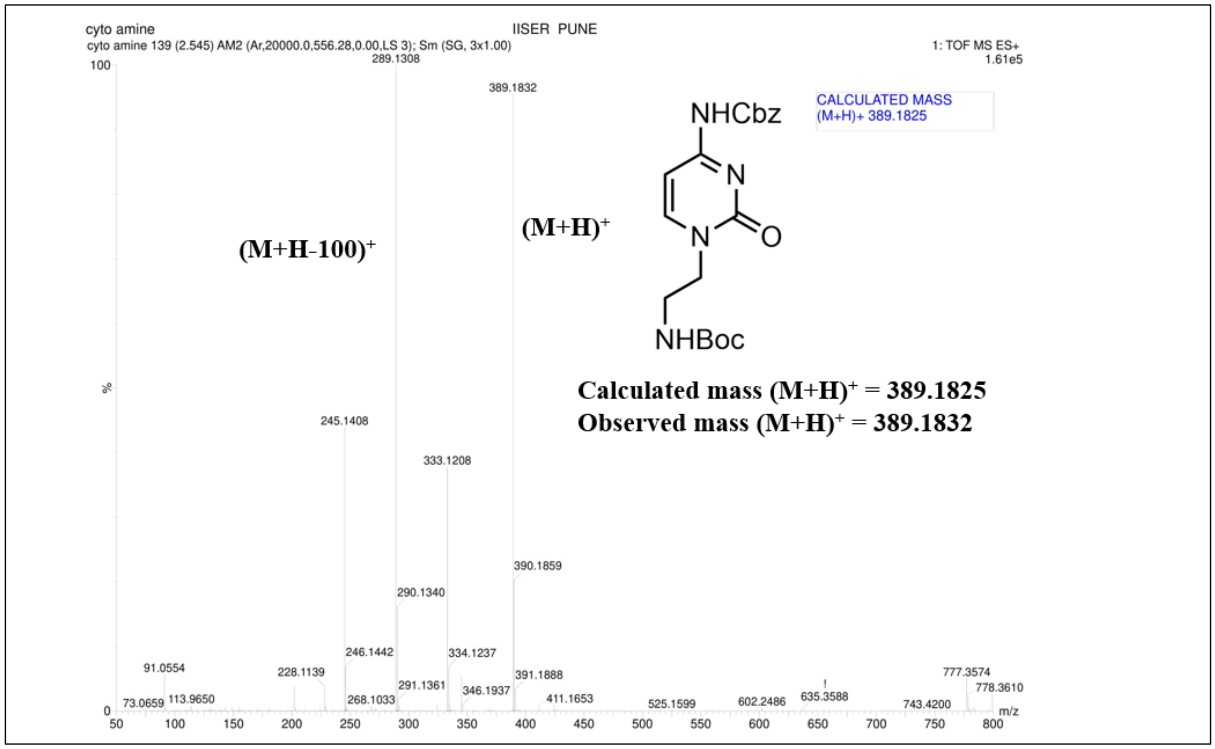
HRMS of compound 8



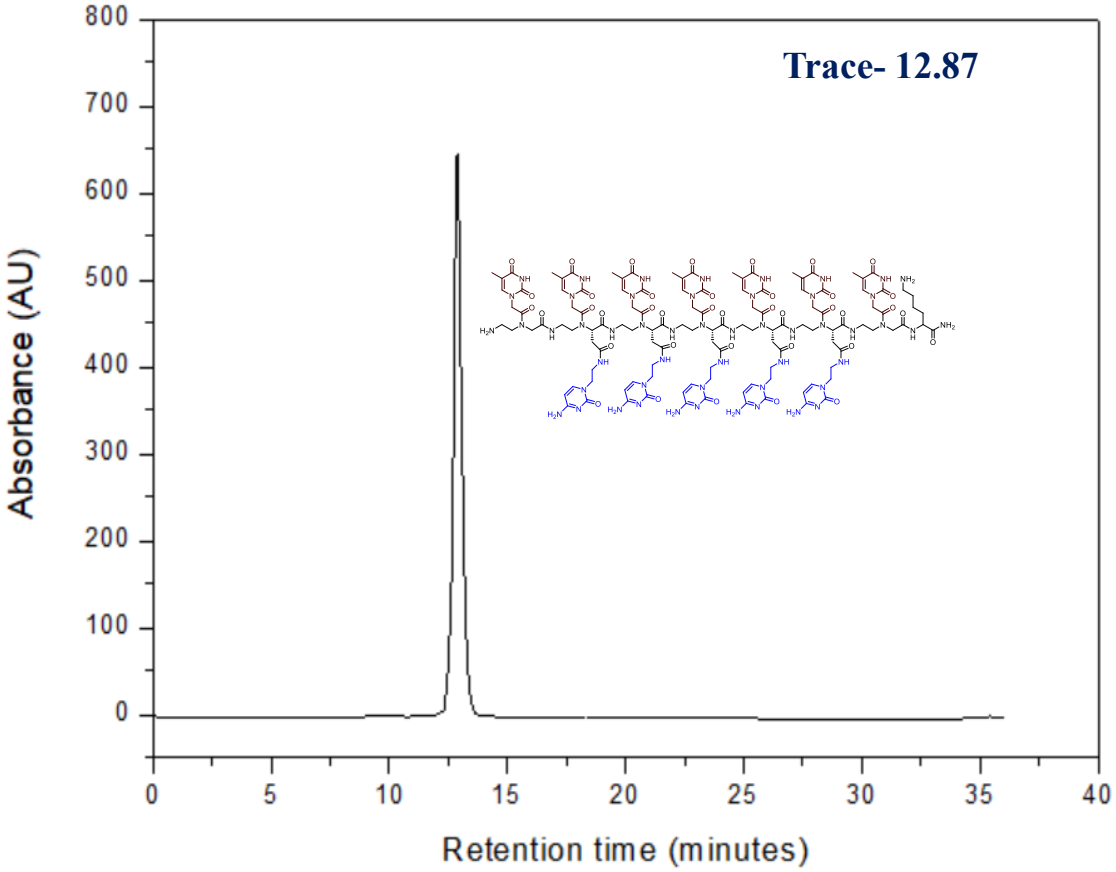
HRMS of compound 9

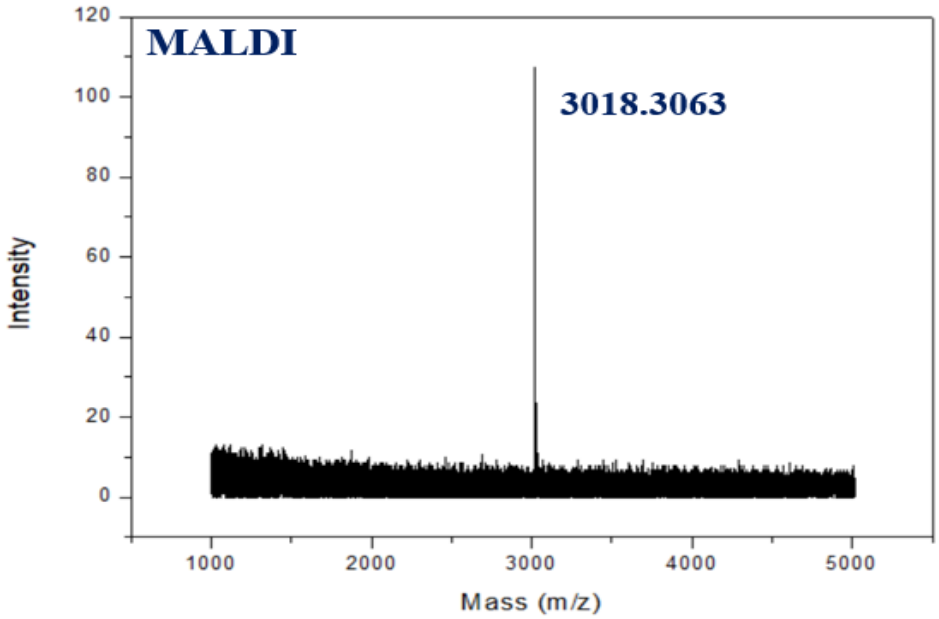


HRMS of compound 13



HPLC and MALDI of *Homo oligomer 12*





Observed mass (M+K) = 3018.3063
Calculated mass (M+K) = 3018.2130
Molecular formula = C_{123} H_{163} N_{51} O_{39}

Chapter 5

Synthesis and characterization of succinimide analogues of Peptide Nucleic Acid

Chapter 5

Synthesis and characterization of succinimide analogues of Peptide Nucleic Acid

5.1. Introduction

Previous chapters have described rational design, synthesis and biophysical studies of backbone modified PNA analogues in which the *gemdimethyl* (*gdm*) groups were introduced at α, β and γ methylene sites and α-CH_2 on backbone replaced either by NH or by a sidechain containing nucleobase (bimodal PNA). Duplex formation of acyclic PNA oligomer with complementary DNA and RNA results loss of entropy upon binding.[1] Loss of entropy during duplex formation is compensated by the enthalpic gain due to hydrogen bond formation and stacking interaction between the nucleobases in PNA and DNA.[1] Apart from that higher rotational barrier around the tertiary carbonyl group for the conversion of Z to E rotamer results distinct kinetics for the binding of PNA with complementary parallel and antiparallel DNA.[2] Figure 5.1 shows the chemical structure of DNA/RNA and PNA.

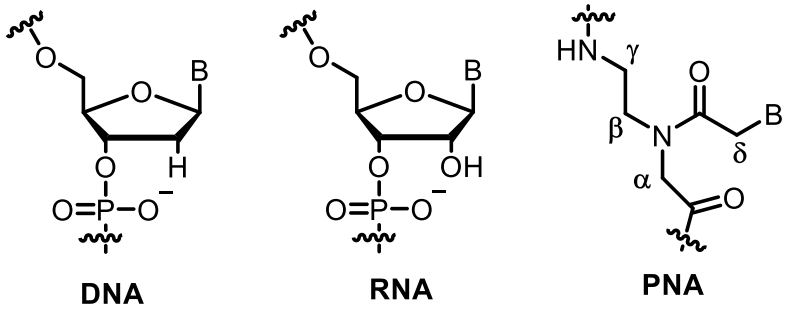


Figure 5.1. Basic chemical structure of DNA, RNA and PNA

Cyclisation of the PNA backbone or incorporation of cyclic molecules in the α, β and γ position of the PNA backbone preorganized the derived PNA oligomers for binding with complementary DNA and RNA.[3] Preorganization results a shift of the equilibrium towards the duplex formation with complementary DNA as entropy loss of the PNA oligomers due to complexation decreases significantly due to cyclisation.[3] Common strategies of preorganization include cyclisation of the *aeg*-PNA backbone in the glycine side, cyclisation in the ethylene diamine side, locking the conformation of PNA monomers by introducing cyclic moieties in the glycine/ethylene diamine side of the backbone etc.[3]

5.2. Biologically active succinimide derivatives

Succinimides are one of the important pharmacophores in medicinal chemistry. *Phensuximide*, *methsuximide* and *ethosuximide* are the succinimide based anti-epileptic drugs.[4] Several other drugs based on succinimide pharmacophore have been approved by US food and drug administration (FDA) and several others are in the clinical trial. First FDA approved drug of smallpox *tecovirimat*, Drug for schizophrenia *lurasidone*, second generation anti-psychotic drug *perospirone* etc. all are based on the pharmacophore *succinimide*. Apart from that succinimide based molecules may act as anticonvulsant, anti-inflammatory, antitumor and anti-microbial agents.[4] Succinimide based drugs are broadly classified into two categories as (a) bio-active natural succinimide drugs and b) synthetic succinimide drugs.[4] Examples of natural bio-active succinimide drugs include *chlorolissoclimide*, *moiramide B* and *nitrosporeusines B*. *Chlorolissoclimide* and *dichlorolissoclimide* acts as cytotoxin against NSCLC, KB, P388 cell lines.[4] *Moiramide B* is a natural succinimide based antibacterial agent with MIC value 0.5 μg/mL. Similar compound with extended methylene *andrimid* has MIC value 2 μg/mL. Structure and activity relation study of *Moiramide B* suggest important role of 4(*S*)-methyl (Me) succinimide fragment in efficient binding to target. Extension of the carbon chain at the N-terminus reduces the potency of *Moiramide B*.[5] Replacing the five membered ring with the six membered succinimide moiety shows reduced activity of the drug.[5] *Nitrosporeusines A* and *B* are thioester containing alkaloids isolated from the arctic actinomycete *Streptomyces nitrosporeus*.[4,6] Yang *et al.*[6] reported the single crystal structure of *nitrosporeusines A and B*, which confirmed the chemical structure of the drug. These exhibit inhibitory action against H1N1 virus in MCDK cells and shows significant anti-inflammatory properties (ED_{50} of 126.8 μM). Replacement of phenol ring by fluorobenzene ring and acetylation of the OH group results improved inhibition property (Compound 1, ED_{50}-34.3 μM). Acetylation of sulphur and OH group (compound **2**) results in significant improved inhibition property (ED_{50}-19.5 μM) compared to original *nitrosporeusines A* and *B*.[6]

Apart from these natural sussinimide compounds, several synthetically modified succinimide drugs have been reported in the literature.[4] Synthetic succinimide drugs can be classified inti the following categories (a) anticonvulsant agents (b) 5-HT receptor ligands (c) anti-inflammatory agents (d) antitumor agents (e) antimicrobial agents (f) antiviral agents (g) enzyme inhibitors and h) ther bioactivities.[4]. Compound **3** acts as an anticonvulsant agent,[4] Compound **4** shows novel anticonvulsant and antinociceptive activity,[4] Compound **5** acts as 5-HT receptor ligands[4] and Compound **6** acts as anti-inflammatory agent. Figure 5.2 and 5.3

shows the chemical structure of the FDA approved and synthesized succinimide based drugs reported in the literature.

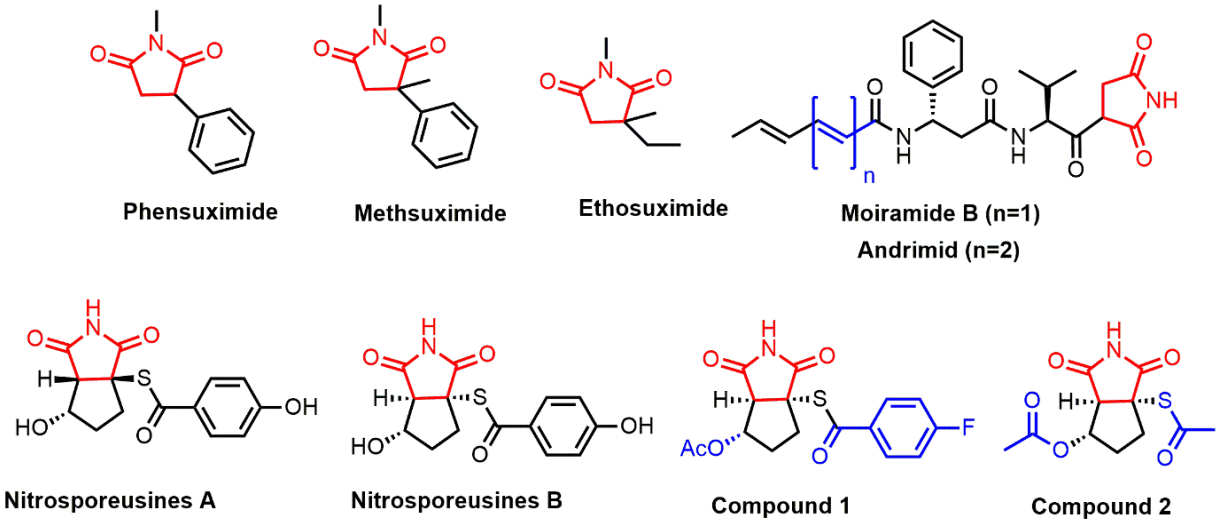


Figure 5.2. Structure of biologically active natural succinimide based drug molecules

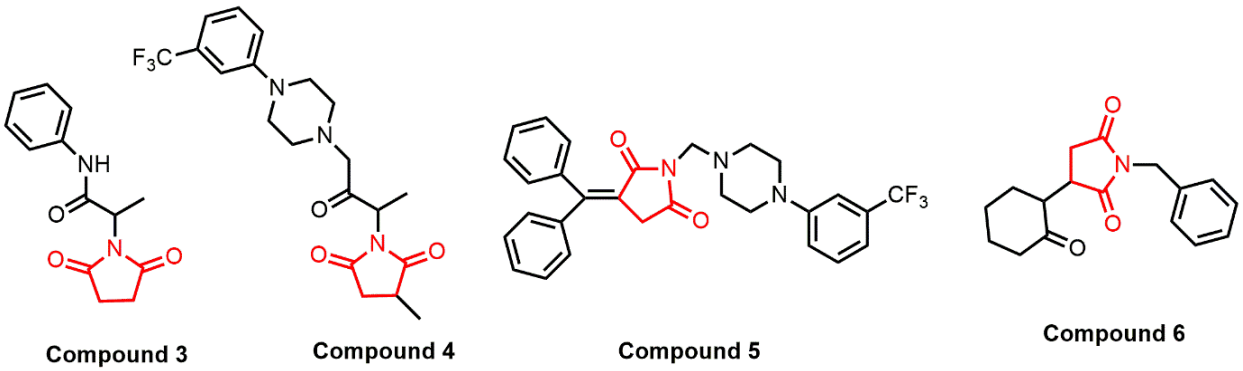


Figure 5.3. Some synthetically modified drugs containing pharmacophore succinimide

5.3. Rationale and objective

Many of the cyclic modifications of Peptide Nucleic Acid involve connecting glycyl a-methylene and β/γ-methylene of the PNA backbone, with side chain δ-methylene group and incorporation of cyclohexane or cyclopentane ring fused to Cβ-Cγ bond in the PNA backbone.[7] In comparison, cyclisation involving the glycyl a-methylene with C-terminal amide group is not known. Cyclisation of suitable C_{α} side chain on PNA backbone with amide group at C-terminus can lead to formation of lactam or succinimide ring. Formation of ring structure at C-terminus could influence the rotamer distribution of the derived *aeg*-PNA. Since many succinimido derivatives also possess antitumor, antimicrobial, antiviral, anti-inflammatory and enzyme inhibition activities (Figure 5.2 and 5.3), diverse nature of substitution around

succinimide pharmacophore fused to PNA backbone gives new opportunities for new succinimide drugs. Compared to the conventional succinimide drug molecules, PNA based drugs will have extra features such as nucleobases that could potentially and selectively interact with the DNA targets via hydrogen bonding. Succinimide based PNA monomers thus would add new repertoire for existing small molecule drugs. With this literature background and rationale, the succinimide derivatives derived from connecting C_{α}-methylene of the PNA backbone to the C-terminus carbonyl of PNA to result in N-substituted succinimide ring (Figure 5.4) were designed.

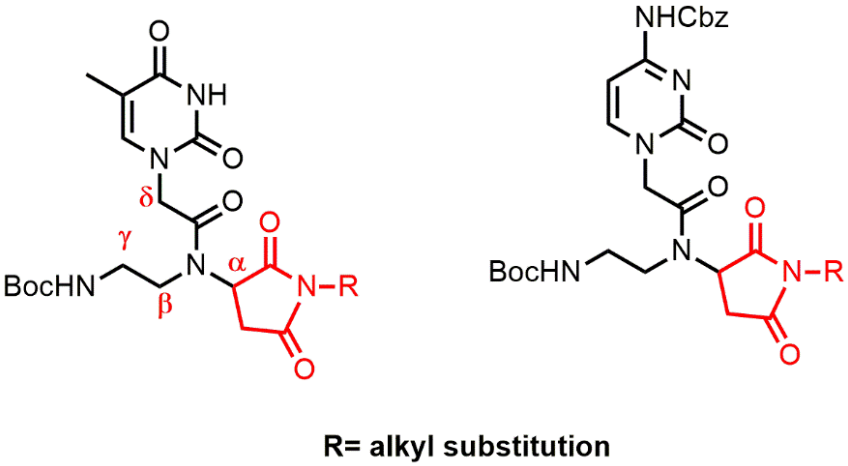


Figure 5.4. Structure of rationally designed succinimide analogues of PNA monomers

The specific objectives of this chapters are

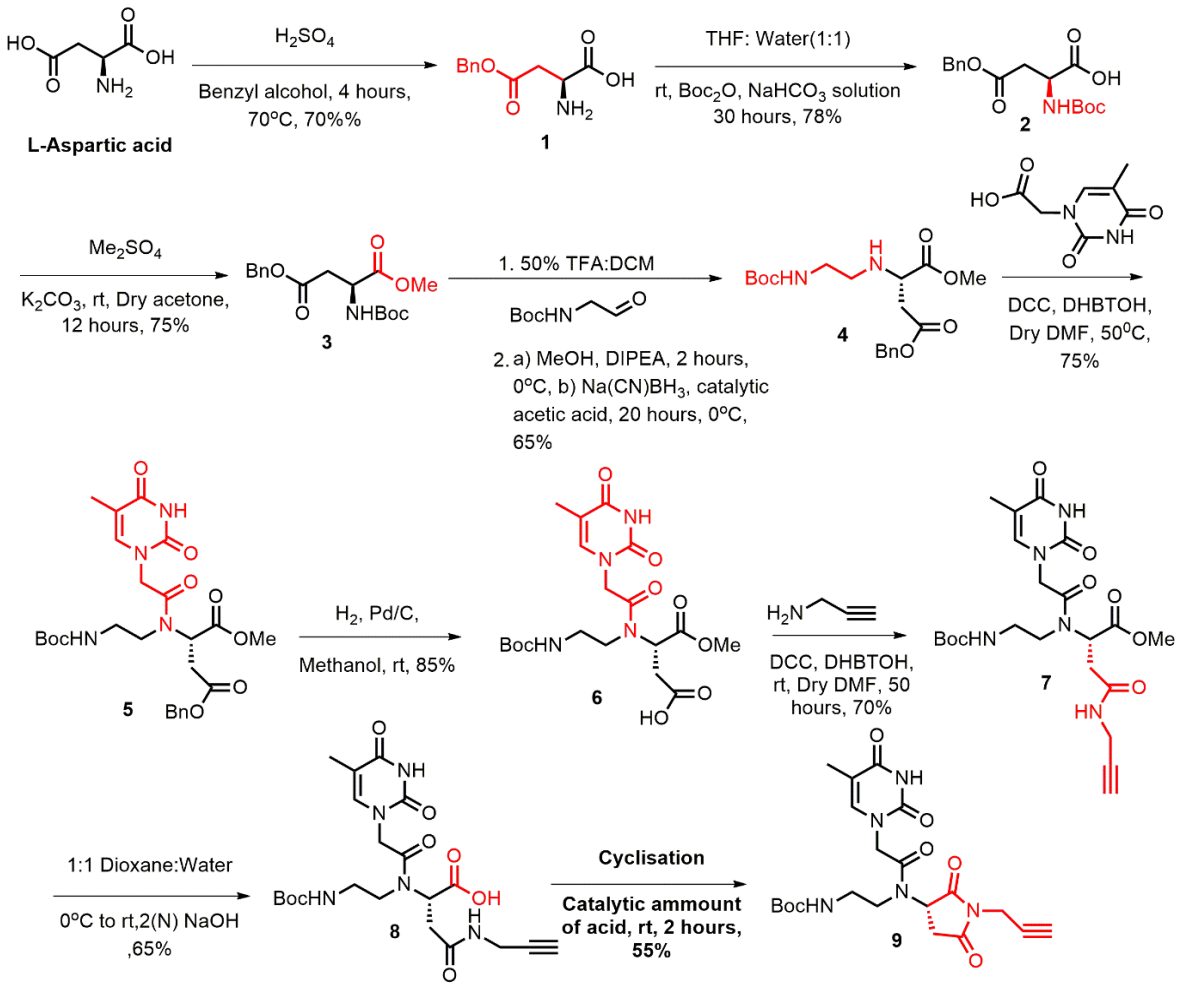
- Synthesis and characterization of rationally designed succinimide analogues of peptide nucleic acid monomers.
- Study the influence of cyclisation on the adjacent t-amide rotamer distribution of the synthesized PNA monomers.
- Study of solution state structure and rotamer distribution of succinimido monomer.
- Single crystal structure of the succinimido PNA monomers

5.4. Result and discussion

5.4.1. Synthesis of Cα-succinimido PNA monomer (9)

The commercially available L-Aspartic acid was treated with benzyl alcohol in presence of sulfuric acid to obtain the selective O-benzylated amino acid (**1**), which was treated with Boc_2O in presence of mild base aq. $NaHCO_3$ to afford the NHBoc-O-(benzyl)-L- aspartic

acid (**2**) (Scheme 1).[8,9] Esterification of **2** with dimethyl sulphate in presence of anhydrous K_2CO_3 gave the orthogonally protected amino acid derivative (**3**).[10] Boc-deprotection of **3** in presence of 50% TFA:DCM, followed by reductive amination in presence of tert-butyl (2-oxoethyl) carbamate gave in situ generated Schiff base that was reduced with $Na(CN)BH_3$ to afford NHBoc-aminoethyl (O-benzyl) methyl aspartate (**4**). DCC mediated coupling of **4** with thymine acetic acid resulted in nucleobase coupled *aeg*-monomer (**5**) which was O-debenzylated by hydrogenolysis in methanol to yield HNBoc-aminoethyl aspartyl methyl ester



Scheme 1: Synthesis of Cα-succinimide of PNA monomer **9**

PNA monomer (**6**). This was treated with propargyl amine to afford the corresponding N-propargyl *aspartamine* PNA ester derivative **7**. The hydrolysis of ester **7** with aq. NaOH provided the free Cα-(N-propargyl-acetamido) PNA monomer acid **8**. Acid mediated cyclisation of **8** in presence of catalytic amount of acid results in sidechain cyclisation with the backbone to afford the Cα- succinimides PNA monomer (**9**) that was characterized by ^1H NMR, mass spectrometry and X-ray crystallography.

5.4.2. NMR studies of the Cα-succinimido PNA monomer (9)

aeg-PNA (**10**) is a flexible molecule with rotational energy barrier (around tertiary N-CO bond) obtained by variable temperature NMR to be 19±2 kcal/mol.[11] The presence of two rotamers of *aeg*-PNA **10** (E and Z) can be detected by ^{1}H NMR spectroscopy and identified by 2D-NOESY experiments as described in earlier chapters. The comparative ^{1}H NMR spectra of *aeg*-PNA (**10**) and C_α-Succinimide modified PNA monomer (**9**) in DMSO-d_6 are shown in (Figure 5.5). C_α-Succinimide modified PNA (**9**) exhibits a mixture of rotamers in $CDCl_3$ and DMSO-d_6 (Figure 5.5), with Z: E rotamer ratio of 1:1 in solution (*aeg*-PNA shows 3:1 ratio). A larger chemical shift difference between the major and minor rotamers of *aeg*-PNA (δ-methylene, 0.17 ppm) suggests slow rate of rotation around t-amide bond in solution. On the other hand, comparatively lesser chemical shift difference was seen between the two rotamers of C_α-succinimide modified PNA **9** (δ-methylene, 0.09 ppm), suggesting a faster exchange rate between the E and Z rotamer in **9**. The backbone methylene protons α, β and γ exhibit a downfield shift compared to *aeg*-PNA **10** in DMSO-d_6. In contrary, δ-methylene protons show

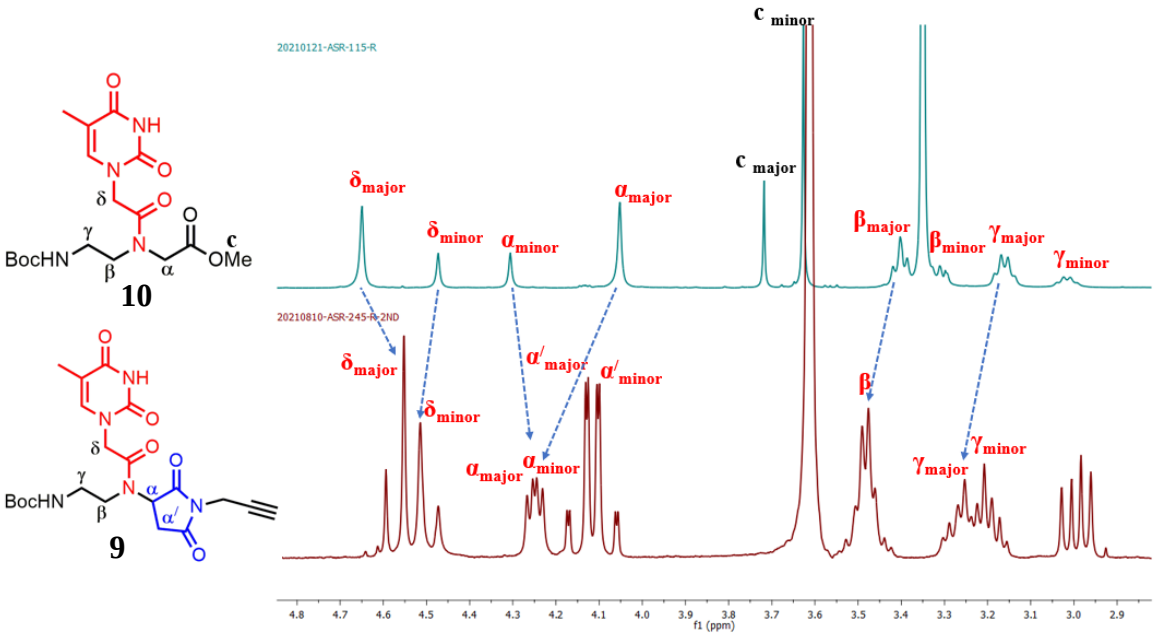


Figure 5.5. comparison of ^{1}HNMR spectra of *aeg*-PNA (**10**) and C_α-succinimide modified PNA (**9**)

upfield shift for one rotamer, but the other rotamer exhibits a small downfield shift compared to *aeg*-PNA (**10**). The 5-CH_3 and CH-6 peaks of thymine appear at 1.73 and 7.09 ppm respectively. The α'-methylene protons in the succinimide ring show multiple (doublet of a doublet) at 4.1 ppm due to both *geminal coupling* and *vicinal coupling* with α-CH peak of

backbone that appears 4.25 ppm reflecting similar multiplicity. The side chain δ-CH_2 appears at 4.55 ppm with PNA backbone β-CH_2 and γ-CH_2 appearing at 3.50 ppm and 3.25 ppm respectively. Succinimide N-substituted propargyl CH_2 and alkyne CH appears at 2.53 ppm and 2.73 ppm respectively. Thus C_α-succinimide cyclization modification results in lowering of rotamer ratio of modified monomer to 1:1. Figure 5.6 shows the full ^{1}H NMR spectra of the C_α-succinimide modified PNA monomer (**9**).

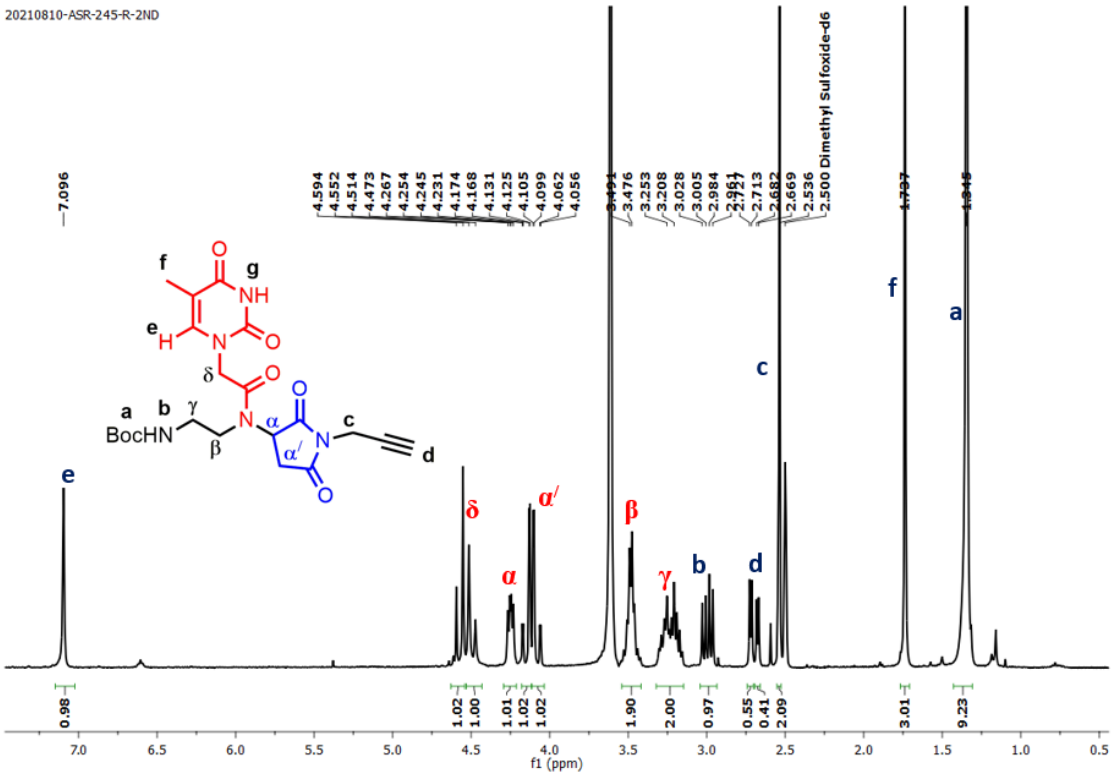


Figure 5.6. ^{1}H NMR of C_α- succinimide modified cyclic PNA monomer (**9**)

Characterisation of the rotamers of C_α-succinimide modified PNA monomer (**9**) were performed by combining the results of 2D-NOSEY (Figure 5.7), Gradient COSY and TOCSY NMR data. δ-methylene at 4.5 ppm of one of the rotamer exhibits cross peak with β-CH_2 protons at 3.4 ppm suggesting the rotamer to be Z-rotamer in DMSO-d_6 (blue line, Figure 5.7). The other rotamer with δ-CH_2 at 4.4 ppm shows a cross peak with Cα at 4.2 ppm suggesting it to be the E-rotamer. Thus, the rotamer preferences of monomers remains unaffected due to C_α-cyclisation of *aeg*-PNA (**10**). Interestingly, α-CH shows NOSEY cross peak with the backbone β-methylene and the propargyl CH suggesting a folded conformation for C_α-succinimide modified PNA (**9**) monomer in solution (red line, figure 5.7), which *aeg*-PNA lacks. X-Ray crystallography further confirms the folded structure of C_α-succinimide modified PNA

monomer (**9**) in solid state. Figure 5.7 shows the NOSEY spectra of the C_{α}-succinimide modified PNA monomer (**9**) in DMSO-d_6.

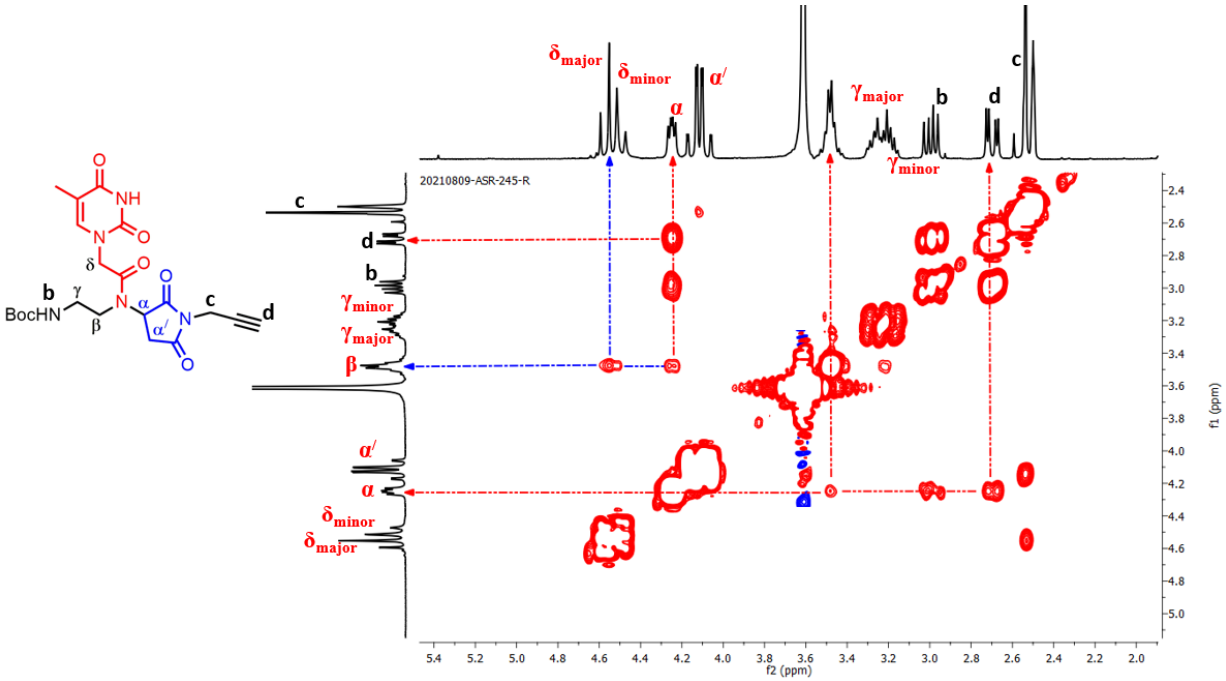


Figure 5.7. Partial ^{1}H-^{1}H NOSEY spectra of C_{α}-succinimide modified PNA monomer in DMSO-d_6

5.4.3 Crystal structure of C_{α}-succinimide modified PNA monomer (9)

The single crystal of C_{α}-succinimide modified PNA monomer (**9**) was obtained by crystallization from 50% methanol/DCM. The sample powders of **9** were dissolved in 50% methanol/DCM to obtain a clear solution, which was sealed in a glass vial, with para film punctured with small holes. Then the vials stored at 4^{o}C and slow evaporation of the solvents lead to crystals. The crystals were mounted on top of a cryoloop and transferred into the cold nitrogen stream (100 K) of a Bruker-AXS D8 Venture diffractometer. Data collection and reduction was performed using the Bruker software suite APEX.[12] The final unit cell was obtained from the xyz centroids of 9824 reflections after integration. A multiscan absorption correction was applied, based on the intensities of symmetry-related reflections measured at different angular settings (SADABS). The structures were solved by direct methods using SHELXT and refinement of the structure was performed using SHLELXL.[13] The hydrogen atoms were generated by the geometrical considerations.

The single crystal X-ray structure of C_{α}-succinimide PNA monomer (**9**) exhibits a folded structure, in which the ethylene diamine side of the PNA backbone folds to give a turn

like structure (Figure 5.8). This is similar to that obtained for β-*gemdimethyl-aeg*-PNA-T (**14**, chapter 3). Another notable feature of the crystal structure is the orientation of t-amide carbonyl which is towards the C_α-succinimide side (glycine side) of the backbone indicating the rotamer to be Z in solid state. This observation also confirms the ^{1}H-H^1 NOSEY results in solution, which suggested one of the rotamers to be Z in the solution. The C_α-succinimide α-CH proton comes close to the backbone β-methylene in crystal structure due to the folded structure of the monomer and hence explains the observance of NOSEY cross peak, suggesting a similar folded structure in solution.

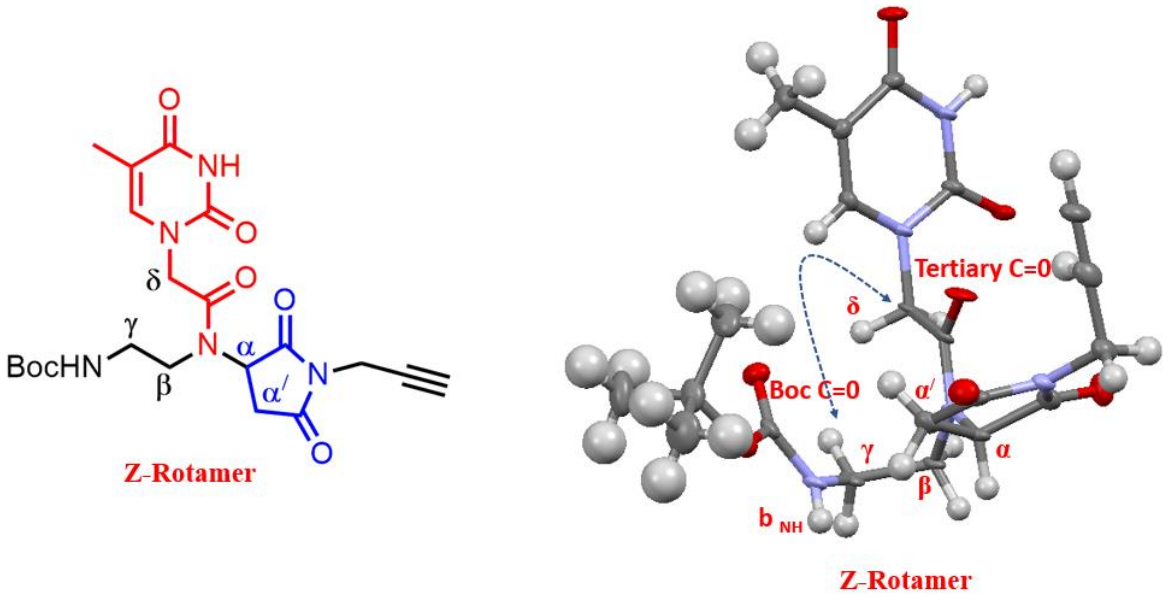
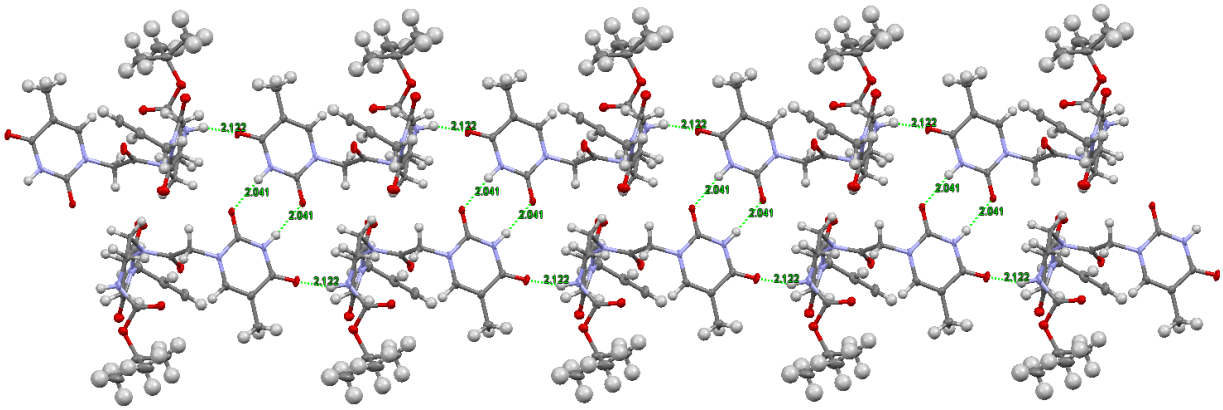


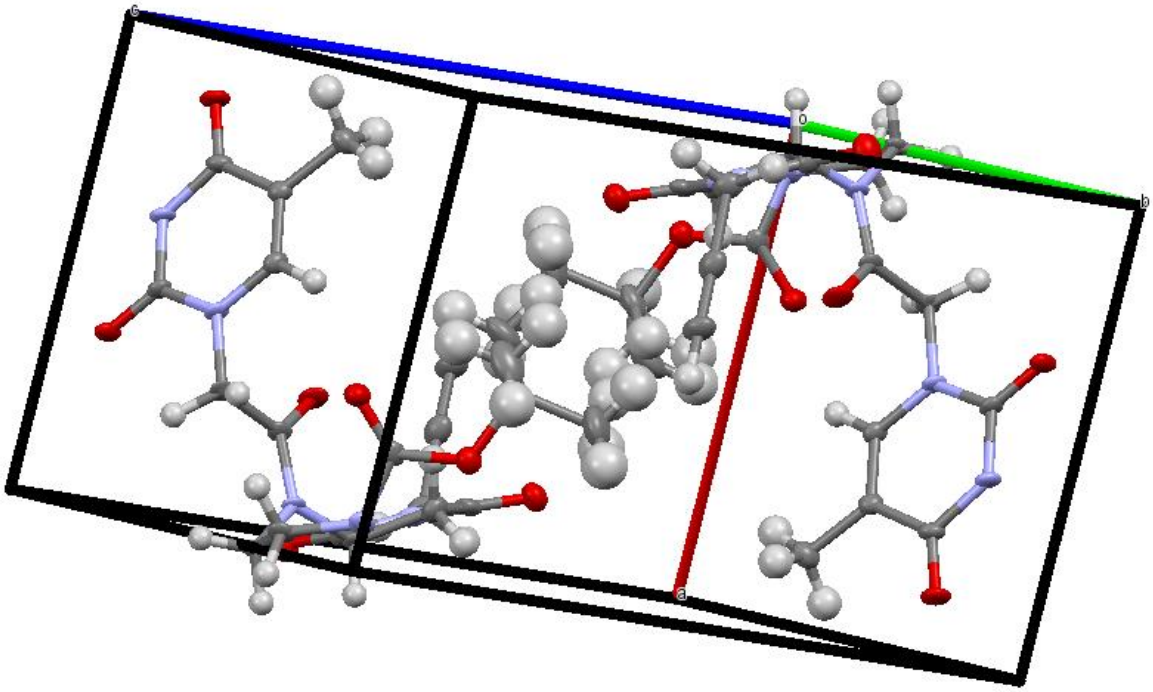
Figure 5.8. Molecular structure and crystal structure of C_α-succinimide modified PNA

The crystal packing of C_α-succinimide modified PNA monomer **9** (Figure 5.9) shows unique packing as the N-Terminal BocNH participates in intermolecular hydrogen bonding with the thymine carbonyl group (NH…O=C distance, 2.12 angstrom) which is an unusual feature, since in other cases, the BocNH participates in hydrogen bonding with the C-terminal acid or ester carbonyl. The ring NH and C=O of thymine participate in the T:T dimeric hydrogen bonding which holds the adjacent T:T chains together through a duplex in solid state. The T:T duplex formation is assisted by the intermolecular hydrogen bonding interaction between the BocNH of the PNA backbone and one C=O group of another thymine in the same chain. Thus C_α-succinimide modified PNA monomer (**9**) exhibit folded structure in solution as well as solid state in the Z-rotamer form suggesting that the cyclized conformation also has the propensity homo base pairing. The packing shows two linear chains of the molecule are held together by T:T base pairing and alternating with two other chains that are held together

by hydrophobic interaction of NH Boc groups. interaction between the N-terminal Boc groups. The following table shows the parameters obtained by the analysis of the crystal.



(**a**)



(**b**)

Figure 5.9. Packing of the C_{α}-succinimide modified PNA monomer in solid state, a) Dimeric T:T duplex structure of the monomer in solid state b) Packing shows hydrophobic interaction between the N-terminal BocNH group.

Crystal data and structure refinement of compound 9

Identification code	TRY a
Molecular formula	C_{21} H_{27} N_5 O_7
Formula weight	461.1910
Crystal system, space group	Triclinic, P-1
Unit cell dimension	a = 10.59 angstrom, α = 106.28º b = 10.80 angstrom, β = 96.88º c =10.33 angstrom, γ = 78.54º
Cell measurement theta max	65.46º
Cell measurement theta min	4.83º
Crystal description	Block crystal
Crystal color	Colorless
Crystal density	1.32
Absorption coefficient	0.845
Absorpt correction type	Multi scans
Absorpt process details	SADBAS
Radiation wavelength and type	1.54 angstrom and Cu/K-α
Diffraction source	Fine focused seal tube
Diffraction measurement device	Bruker APEX-II CCD
Diffraction method	Omega and Phi Scan
Diffraction reflection number	11011
Diffraction reflection h min	-12
Diffraction reflection h max	12
Diffraction reflection l min	-11
Diffraction reflection l max	12
Diffraction reflection k min	-14
Diffraction reflection k min	14
Diffraction reflection theta min	4.82
Diffraction reflection theta max	65.45

Total reflection number	3865
Computing data collection	Bruker Apex-II
Computing cell refinement	Bruker SAINT
Computing data reduction	Bruker SAINT
Computing structure solution	SHELXT 2018/2 (sheldrick, 2018)
Computing data refinement	SHELX-L 2018/3 (sheldrick, 2018)
Computing molecular graphics	Bruker SHELXTL
Computing publication material	Bruker SHELXTL
Refine ls extinction method	SHELX-L 2018/3 (sheldrick, 2018)
Refine ls extinction coefficient	0.00 (6)
Refine ls R factor all	0.0871
Refine ls R factor gt	0.0667

5.5. Discussion

Gazit *et al.* [4] reported the crystal structure of GC di-PNA in 2015. Notable feature of the crystal structure is the complementary Watson Crick base pairing between the Guanine and Cytosine nucleobases to form the PNA:PNA duplex structure in solid state. The π-π stacking interactions between the guanine and cytosine nucleobases stabilize the dimeric structure in solid state. The tertiary carbonyl group is oriented towards the glycine side of the PNA backbone suggesting a Z-rotamer stabilization in the GC dimer as observed for individual monomers in solution state. The backbone is unfolded backbone in the crystal structure.[14] Acyclic PNA monomers lacking *gemdimethyl* group at γ/β-position stabilize the Z-rotamer over E-Rotamer in solution and solid state.[11] *Gemdimethyl* substitution in the γ/β-position stabilizes the E-rotamer over Z-rotamer in solution and solid state.[15] The chains in *gemdimethy* PNA are also unfolded.[15]

The NMR study on Ca-succinimide PNA show that cyclisation of PNA backbone in the Cα side leads to a Z:E rotamer ratio 1:1, compared to 3:1 for *aeg*-PNA. ^{1}H-H^{1} NOSEY correlation data for C_{α}-succinimide PNA monomer **9** exhibits one of the rotamer to be Z-rotamer and another rotamer to be E-rotamer stabilization in DMSO-d_6. However, the ^{1}H-H^{1} NOSEY shows NOSEY cross peak of α-CH with the backbone γ-methylene and the propargyl CH protons suggesting a folded conformation of C_{α}-succinimide PNA monomer in solution

(Figure 5.7). The single crystal X-Ray structure of C_α-succinimide PNA monomer shows Z rotamer in solid state similar as *aeg*-PNA, but, the backbone the modified PNA shows folded structure (Figure 5.8). The folding of the backbone and side chain compacts the structure resulting in preorganization of the PNA for T:T base pairing which holds two chains. This is perhaps the first crystal structure of PNA wherein the Cα side chain is cyclised to the C-terminus of the backbone. It would be interesting to study the PNA oligomers derived from such cyclised PNA units and the effect of stereochemistry on the rotamer preferences and PNA:DNA binding property of the resulted oligomers.

5.6. Summary

The major conclusions are summarised as

- Successful synthesis and characterization of the C_α-succinimide - PNA monomer.
- Rotamer studies by ^{1}H NMR indicating 1:1 a ratio of Z and E-rotamer compared to 3:1 rotameric ration in *aeg*-PNA.
- NOESY spectra hints a folded structure of the modified PNA monomers in solution.
- Single crystal structure study shows Z-Rotamer in a backbone folded state.
- C_α-succinimide PNA monomer shows a folded structure in the solid state confirming the observation in of NMR study in solution.
- The T:T duplex formation of C_α-succinimide PNA monomer across two chains is assisted by the intermolecular hydrogen bonding between BocNH and thymine carbonyl group.

5.7. Experimental Section

This section describes the detailed synthesis of monomers and the spectral characterisation of the rationally designed monomers described in this chapter.

5.7.1. General information

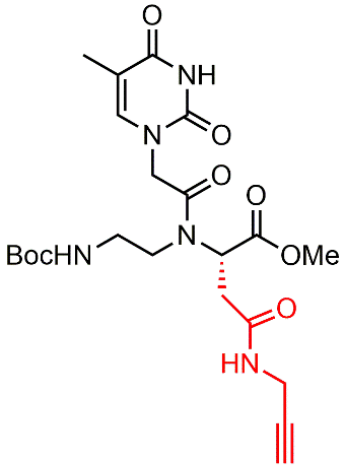
Laboratory grade chemicals were used for setting up standard chemical reactions. Chemicals were purchased from Sigma Aldrich and Spectrochem. For moisture sensitive reactions the reagents were redistilled to ensure the maximum purity of the reagents. Solvents used for performing chemical reactions were dried using standard procedures available in the literature. Reactions were monitored using Thin Layer Chromatography (TLC). Standard UV-

Lamp, Iodine spray and spraying with ninhydrin solution followed by heating the TLC plate were used to analyse the TLC. Standard procedure of work up involves extracting the desired organic compound by washing the aqueous layer with organic solvent (DCM, Ethyl acetate) followed by washing of the combined organic layer with Brine. The combined organic layer was dried over Sodium sulphate (Na_2SO_4) or Calcium chloride ($CaCl_2$, For DCM) and evaporated under reduced pressure to afford the desired organic compound. Precoated silica gel GF_{254} (Merck 5554) TLC plates were used for TLC analysis. Standard chemical reactions were performed in room temperature, wherever required temperature was mentioned. Silica gel (60-120 or 100-120 mesh) were used to perform the column chromatography. Double distilled Ethyl acetate, hexane, Methanol and DCM were used as a mobile phase to perform the column chromatography. ^{1}H NMR, ^{13}C NMR and 2D-NMR experiments were performed in Bruker-400 MHz, JEOL-400 MHz, Bruker-500 MHz, and Bruker-600 MHz NMR spectrometers. NMR solvents were purchased from Sigma Aldrich. Deuterated chloroform ($CDCl_3$) and DMSO-d_6 were used as an NMR solvent. Chemical shifts were reported in δ ppm with respect to the internal standard TMS or deuterated NMR solvent. Mass spectra of the reaction intermediates were recorded in an Applied Biosystems 4800 Plus MALDI-TOF/TOF mass spectrometry using TiO_2 or 2,5- dihydroxybencoic acid (DHB) as matrix. To further confirm the mass spectra of small molecules (Mass 100-1200) obtained in MALDI, High Resolution Mass Spectrometry (HRMS) was performed for each reaction intermediates. IR spectra of the samples were recorded in a Bruker IR spectrometer. Oligomers were characterised by Agilent HPLC instrument, C_{18} column and MALDI-TOF spectrometry.

5.7.2. Experimental procedures and spectral data

Experimental procedures for the synthesis of compound **1** to **6** are described in the chapter 4 (**Section 4.8.2**-experimental procedure), hence synthesis and spectral data of compound **7** to **9** are described here.

(S)-methyl-2-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)-4-(ethynylamino)-4-oxobutanoate (7)

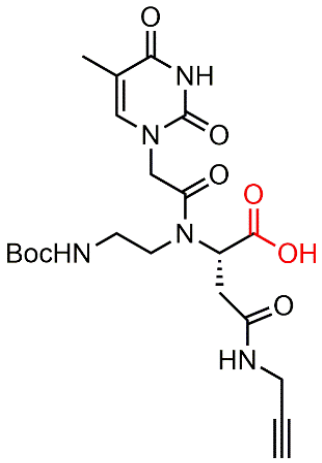


7

Free acid of C_{α}-Substituted Thymine attached monomer **6** (0.2 mmol, 100 mg) was dissolved in Dry DMF and to that solution was added N,N′-Dicyclohexylcarbodiimide (DCC, 0.2 mmol, 45 mg) followed by 3-Hydroxy-1,2,3-benzotriazin-4(3H)-one (DHbtOH, 0.2 mmol, 35 mg) and the solution was stirred at room temperature for 1 h . After 1 h Propargyl amine (0.3 mmol, 21 µL) was added and the reaction mixture was stirred at room temperature for additional 48 hours. After the completion of the reaction the desired compound was partitioned between brine and EtOAc. The titled compound was extracted by EtOAc (3×150 ml). The combined organic layer was dried over Na_2SO_4, Filtered, and evaporated under vacuo to give mixture of a white solid and yellow liquid as a crude. Column chromatography using 5% MeOH in DCM gave the titled compound **7** as a white hygroscopic solid (74 mg, Yield- 65%, R_f = 0.3 in 5% MeOH in DCM). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.20 (s, 1H), 10.01 (d, *J* = 44.7 Hz, 1H), 6.92 (t, *J* = 20.7 Hz, 1H), 5.74 (d, *J* = 83.2 Hz, 1H), 4.86 (dd, *J* = 23.8, 16.8 Hz, 1H), 4.28 (dd, *J* = 37.9, 17.5 Hz, 2H), 4.06 (dd, *J* = 34.5, 16.1 Hz, 3H), 3.76 (d, *J* = 19.6 Hz, 2H), 3.47 (dd, *J* = 13.3, 6.4 Hz, 1H), 3.25 (d, *J* = 36.4 Hz, 2H), 3.00 (d, *J* = 18.0 Hz, 1H), 2.23 (s, 2H), 1.87 (s, 3H), 1.43 (s, 9H), ^{13}C NMR (101 MHz, CHLOROFORM-D) δ 174.4, 172.7, 170.8, 170.1, 167.6, 166.9, 164.6, 164.4, 156.30, 156.1, 151.9, 151.4, 140.9, 140.8 ,140.7, 114.18 ,114.16, 111.1, 80.2, 79.9, 71.8, 71.01, 65.9, 58.8, 58.0, 53.0, 49.4, 39.4, 38.7, 35.8,

33.9, 31.9, 31.7, 29.7, 29.1, 28.5, 28.5, 22.7, 15.32, 12.4, Exact mass 494.2251 $(C_{22}H_{31}N_5O_8+H)^+$, Observed mass 494.2254.

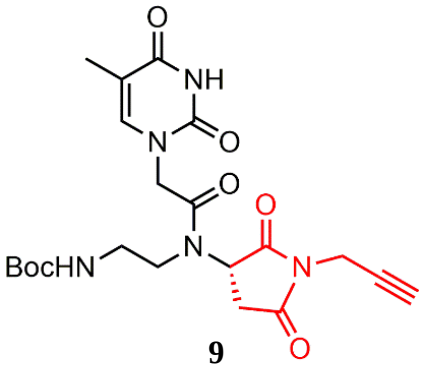
(S)-2-(N-(2-((tert-butoxycarbonyl)amino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)-4-(ethynylamino)-4-oxobutanoic acid (8)



8

Propargyl amine attached thymine monomer **7** (2.1 mmol, 1 gm) was dissolved in 1:1 solution of Dioxane in water (20 ml of Dioxane: 20 ml of H_2O) and 2(N) LiOH (30 ml) solution was added to it. The reaction mixture was allowed to stirred at room temperature for 3 h. After the completion of the reaction Dioxane was evaporated under vacuo and the aqueous solution was acidified with 2(N) HCl to pH 2. The compound was extracted with EtOAc (3×150 ml). The combined organic layer was dried over Na_2SO_4 and concentrated under reduced pressure to give an oily product. DCM was added and evaporated under vacuo to remove the excess EtOAc. Purification of the crude via column chromatography using 10% MeOH in DCM as a eluent affords the compound **8** as hygroscopic solid. The obtained white hygroscopic powder is the titled acyclic monomer **8** (640 mg, Yield-65%, R_f = 0.4 in 20% MeOH in DCM). ^{1}HNMR (400 MHz, CHLOROFORM-D) δ 12.33 (s, 1H), 11.27 (dd, *J* = 17.9, 8.6 Hz, 1H), 8.39 – 8.13 (m, 1H), 7.27 (dd, *J* = 13.8, 6.9 Hz, 1H), 7.05 – 6.75 (m, 1H), 4.71 – 4.50 (m, 2H), 3.95 – 3.74 (m, 2H), 3.19 – 3.01 (m, 3H), 3.00 – 2.80 (m, 2H), 2.48 (s, 1H), 2.44 (d, *J* = 1.5 Hz, 1H), 1.73 (d, *J* = 4.7 Hz, 3H), 1.36 (s, 9H), ^{13}C NMR (101 MHz, DMSO-d_6) δ 171.8, 168.8, 167.6, 164.4, 155.8, 151.1, 142.1, 108.2, 81.0, 78.2, 73.0, 34.53, 28.2, 12.04. , Exact mass 480.2094 $(C_{21}H_{29}N_5O_8+H)^+$ Observed mass 480.2098.

(S)-tert-butyl (2-(N-(2,5-dioxo-1-(prop-2-yn-1-yl)pyrrolidin-3-yl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)ethyl)carbamate (9)



9

Propargyl amine attached thymine monomer **8** (200 mg, 0.4 mmol) was dissolved in MeOH: DCM (1:1) and catalytic amount of glacial acetic acid (0.5 µL) was added in the mixture. The mixture was stirred for 2 h at room temperature. After stirring for 2 h the solution was kept for crystallization at 4^{o} C. White block type crystals of Cα-succinimide modified PNA monomer (**9)** was obtained in the in the solution. Analysis of the mother liquor by ^{1}H NMR shows (Solution) cyclisation of the open chain Propargyl amine attached thymine monomer **8** in presence of glacial acetic acid. In conclusion the crystal and the mother liquor exhibit cyclisation of the backbone. Purification of the cyclised monomer affords the C_{α}-succinimide modified PNA monomer **9** as white solid (107 mg, Yield-55%, R_f = 0.5 in 20% MeOH in DCM). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.10 (s, 1H), 4.53 (q, *J* = 16.7 Hz, 2H), 4.25 (dd, *J* = 9.0, 5.3 Hz, 1H), 4.12 (qd, *J* = 17.2, 2.5 Hz, 2H), 3.48 (dd, *J* = 11.9, 6.0 Hz, 2H), 3.23 (ddd, *J* = 20.8, 14.2, 7.6 Hz, 1H), 2.99 (dd, *J* = 17.9, 9.2 Hz, 1H), 2.70 (dd, *J* = 17.9, 5.3 Hz, 1H), 2.54 (s, 2H), 1.74 (s, 3H), 1.34 (s, 9H), ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.8, 168.8, 167.6, 164.4, 155.8, 151.1, 142.1, 108.2, 81.0, 78.2, 73.0, 34.53, 28.2, 12.04, Exact mass 484.1807 $(C_{21}H_{25}N_5O_7+Na)^+$ Observed mass 484.1804.

5.8. References

1. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc*., **1996**, *118*, 5544-5549.

2. a) Hollenstein, M.; Leumann, C. J. *Org. Lett*., **2003**, *5*, 1987-1990. b) Rose, D. *J. Anal. Chem*., **1993**, *65*, 3545-3549.

3. a) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **1999**, *1*, 1513-1516. b) Vilaivan, T.; Khongdeesameor, C.; Harnyuttanokam, P.; Westwell, M. S.; Lowe, G. *Bioorg. Med. Chem. Lett*., **2000**, *10*, 2541-2545. c) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **2001**, *3*, 1281-1284. d) Sharma, N.; Ganesh, K. N. *Tetrahedron Lett*., **2004**, *45*, 1403-1406. b) Sharma, N.; Ganesh, K. N. *Chem. Commun*., **2003**, *0*, 2484-2485. e) Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc*., **2006**, *128*, 16456 – 16457.

4. Zhao, Z.; Yue, J.; Ji, X.; Nian, M.; Kang, K.; Qiao , H.; Zheng, X. *Bioorganic Chemistry*., **2021**, *108*, 104557.

5. a) Zhao, Z.; Yue, J.; Ji, X.; Nian , M.; Kang, K.; Qiao , H.; Zheng, X. *Bioorganic Chemistry*., **2021**, *108*, 104557. b) Needham, J.; Kelly, M.T.; Ishige, M.; Andersen, R.J.; Andrimid.; and Moiramides. *J. Org. Chem*., **1994**, *59*, 2058–2063. c) Pohlmann, J.; Lampe, T.; Shimada, M.; Nell, P.G.; Pernerstorfer, J.; Svenstrup, N.; Brunner, N.A.; Schiffer, G.; Freiberg, G. *Bioorg. Med. Chem. Lett*., **2005**, *15*, 1189–1192. d) Silvers, M.A.; Robertson, G.T.; Taylor, C.M.; Waldrop, G.L. *J. Med. Chem*., **2014**, 57, 8947–8959.

6. a) A. Yang, L.; Si, Z.; Shi, L.; Tian, D.; Liu, D.; Zhou, P.; Proksch, W. Lin. *Org. Lett*., **2013**, 15, 5366–5369. b) Philkhana, S.C.; Jachak, G.R.; Gunjal, V.B.; Dhage, N.M.; Bansode, A.H.; Reddy, D.S. *Tetrahedron Lett*., **2015**, *56* ,1252–1254.c) Zhao, Y.; Nie, M.; Xu, P.; Li, H.; Xu, L.; Cheng, X.Q; Lei, Y.Q. *Biochem. Biophys. Res. Commun*., **2019**, *515* ,474–480. d) Philkhana, S.C.; Verma, A.K.; Jachak, G.R.; Hazra, B.; Basu, A.; Reddy, D.S. *Eur. J. Med. Chem*., **2017**, 135 ,89–109.

7. a) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **1999**, *1*, 1513-1516. b) Vilaivan, T.; Khongdeesameor, C.; Harnyuttanokam, P.; Westwell, M. S.; Lowe, G. *Bioorg. Med. Chem. Lett*., **2000**, *10*, 2541-2545. c) D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Org. Lett*., **2001**, *3*, 1281-1284. d) Sharma, N.; Ganesh, K. N. *Tetrahedron Lett*., **2004**, *45*, 1403-1406. e) Sharma, N.; Ganesh, K. N. *Chem. Commun*., **2003**, *0*, 2484-2485. f) Englund, E. A.; Xu, Q.; Witschi, M. A.; Appella, D. H. *J. Am. Chem. Soc*., **2006**, *128*, 16456 – 16457.

8. Asano, Y.; Umezaki, M.; Li, Y.F.; Tsubota, S.; Lübbehüsen, T.L. *Journal of Molecular Catalysis B.*, **2001,** *12*, 53–59.

9. Eldo, J.; Cardia, J.P.; O'Day, E.M.; Xia, J.; Tsuruta, H.; and Kantrowitz, E.V. *J. Med. Chem*., **2006**, *49*, 5932-5938.

10. Chakraborty, S.; Tai, D.F. *Tetrahedron Letters*., **2014**, *55*, 2274–2276.

11. a) Brown, S.C.; Thomson, S.A.; Veal, J.M.; Davis, D.G. *Science* **1994**, *265*, 777 - 780. b) Oleszczuk, M.; Motowidlo, S.R.; and Falkiewicz, B. *nucleosides, nucleotides & nucleic acids*, **2001**, *20,* 1399–1402. c) Meltzer, P.C.; Liang, A.Y.; and Matsudaira, P. *J. Org. Chem*. **1995**, *60*, 4305-4308. (d) Ovadia, R.; Lebrunb, A-L.; Barvikd, I.; Vasseurc, J.-J; Baragueyc, C.; Alvarez, K. *Org. Biomol. Chem.*, **2015**, *13*, 11052-11071.

12. Bruker, (2016). *APEX3* (v2016.1-0), *SAINT* (Version 8.18C) and *SADABS* (Version 2012/1). Bruker AXS Inc., Madison, Wisconsin, USA.

13. a) Sheldrick, G. M. *Acta Cryst*. **2015**, *A71*, 3-8. b) Sheldrick, G. M. *Acta Cryst*. **2008**, *A64*, 112-122.

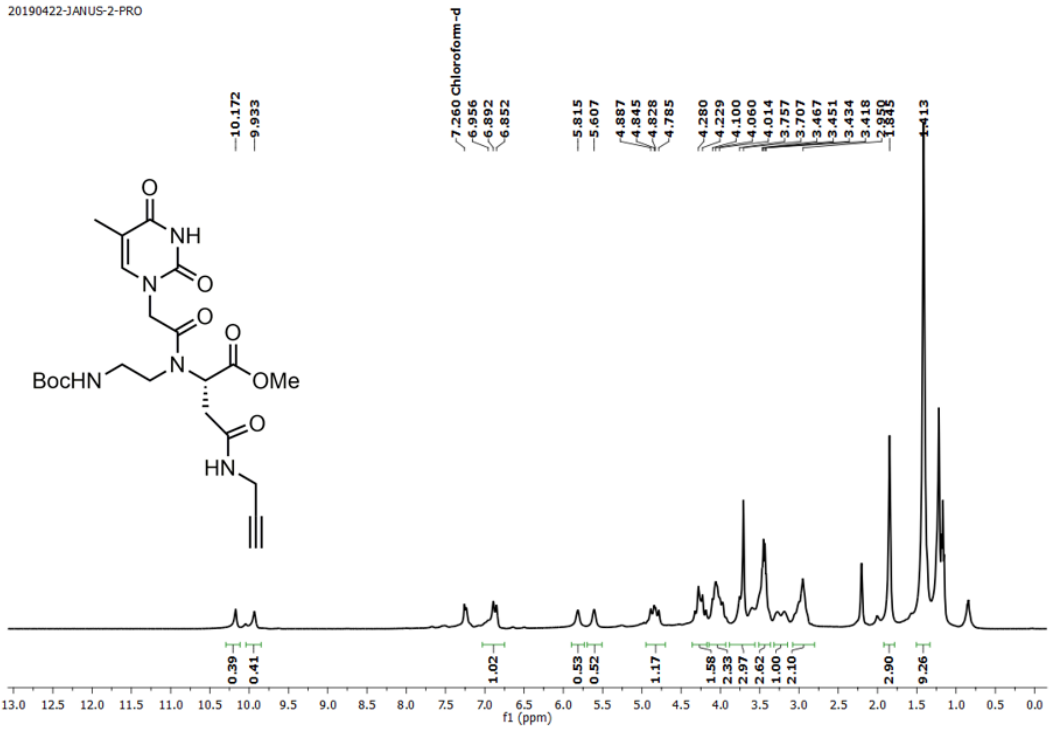
14. Berger, O.; Abramovich, L.A.; Sakin, M.L.; Grunwald, A.; Peer, Y.L.; Bachar , M.; Buzhansky, L.; Mossou, E.; Forsyth, V.T.; Schwartz, T.; Ebenstein, Y.; Frolow,F.; Shimon, L.J.W.; Patolsky, F.; and Gazit, E. *Nature Nanotechnology*, **2015**, *10*, 353–360.

15. Kulkarni, P.; Datta, D.; Ramabhadran, R.O.; and Ganesh, K.N. *Org. Biomol. Chem*., **2021**, *19*, 6534.

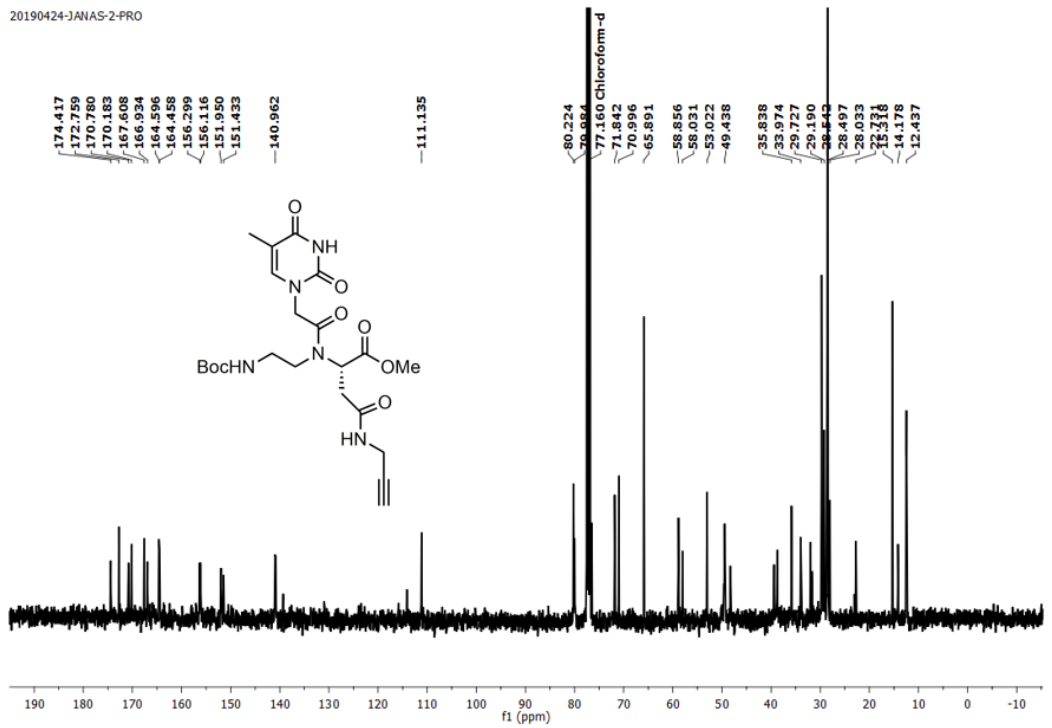
5.9. Characterisation data of the synthesized compounds and monomers (Appendix-IV)

Entry	Table of contents	Page Number
1	**^{1}HNMR, ^{13}CNMR, and HRMS spectra of the monomers**	**304-308**

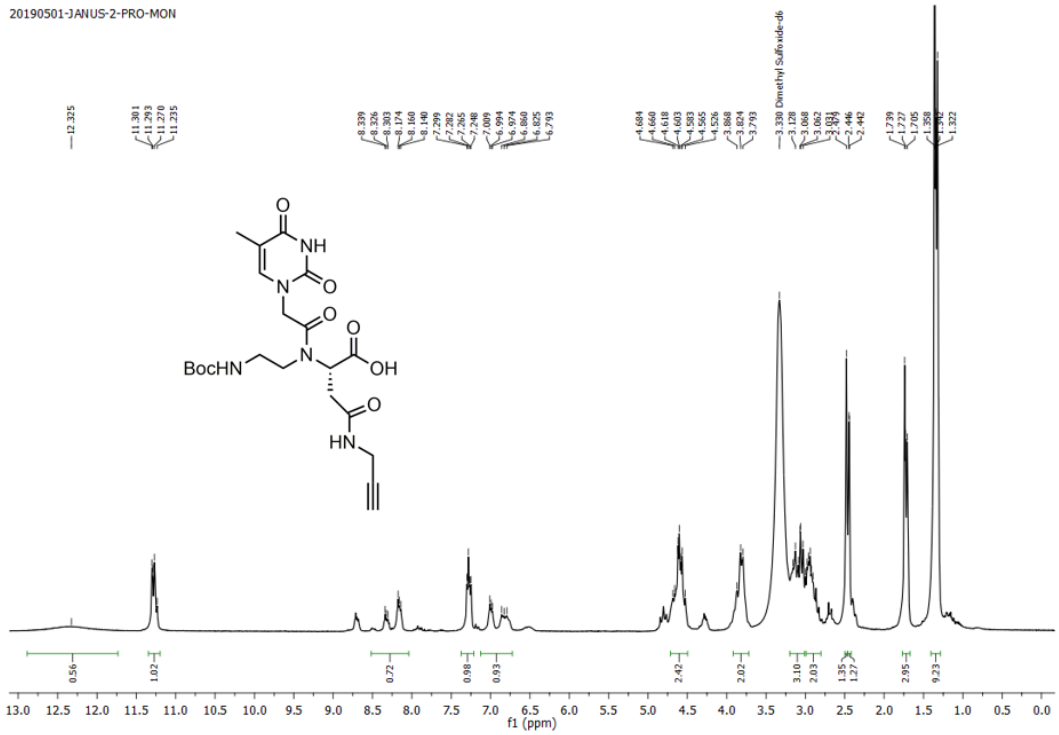
^{1}H NMR of compound 7 in $CDCl_3$ (400 MHz)



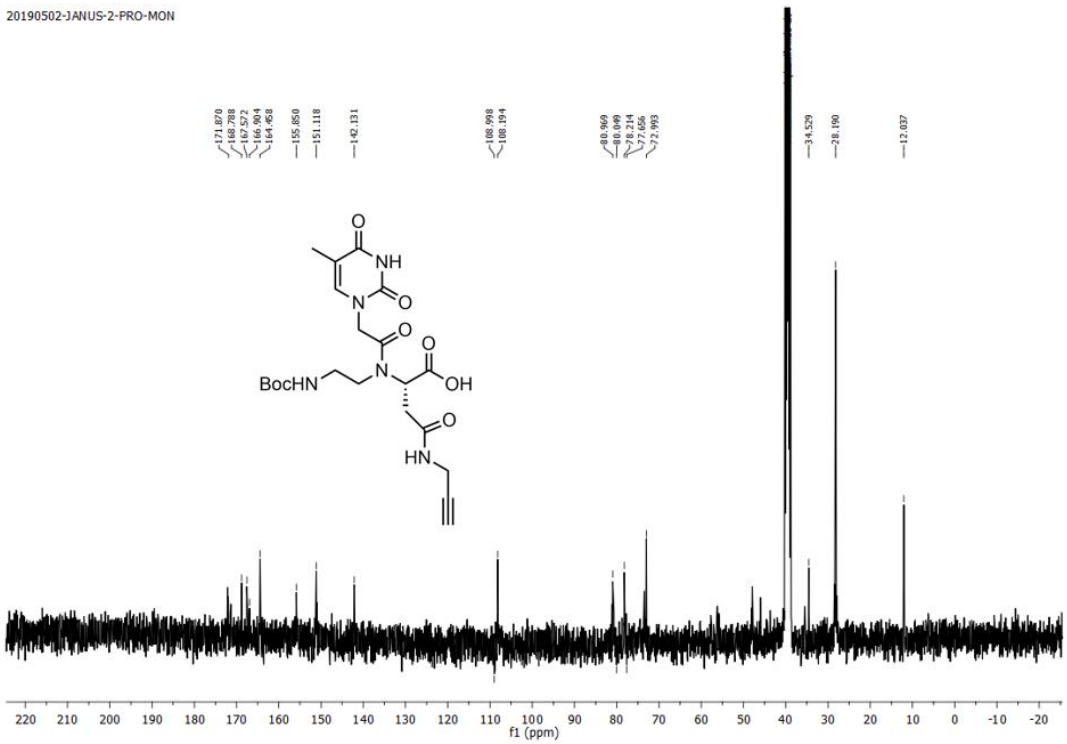
^{13}C {1H} NMR of compound 7 in $CDCl_3$ (101 MHz)



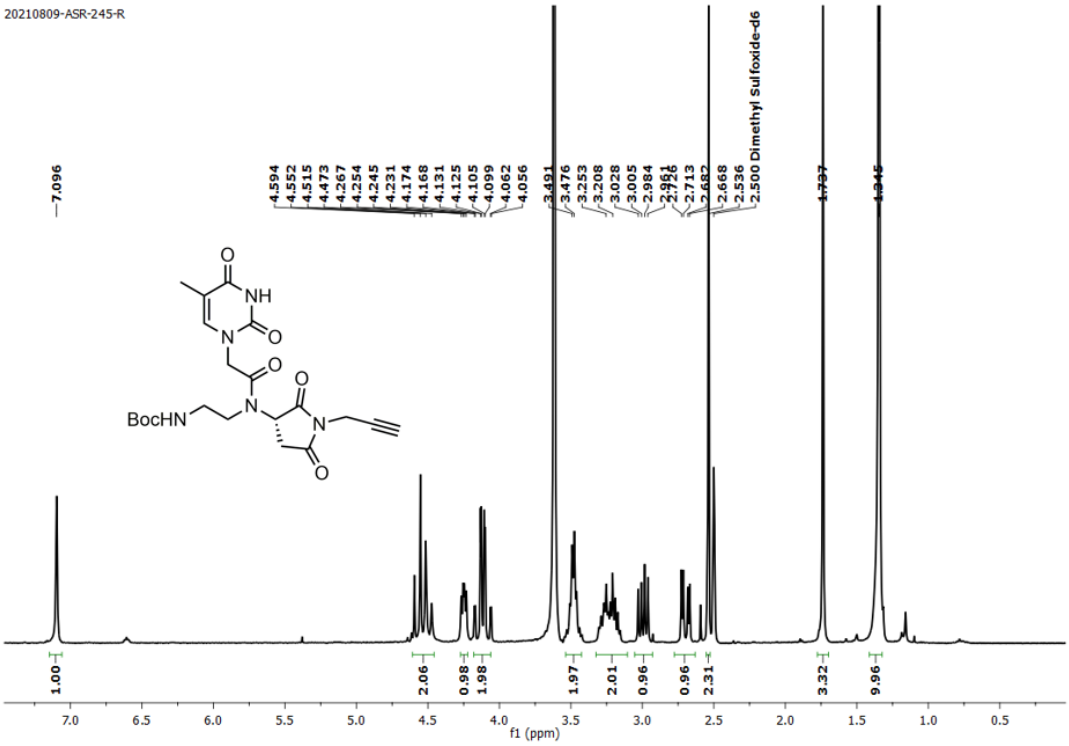
^{1}H NMR of compound 8 in DMSO-d_6 (400 MHz)



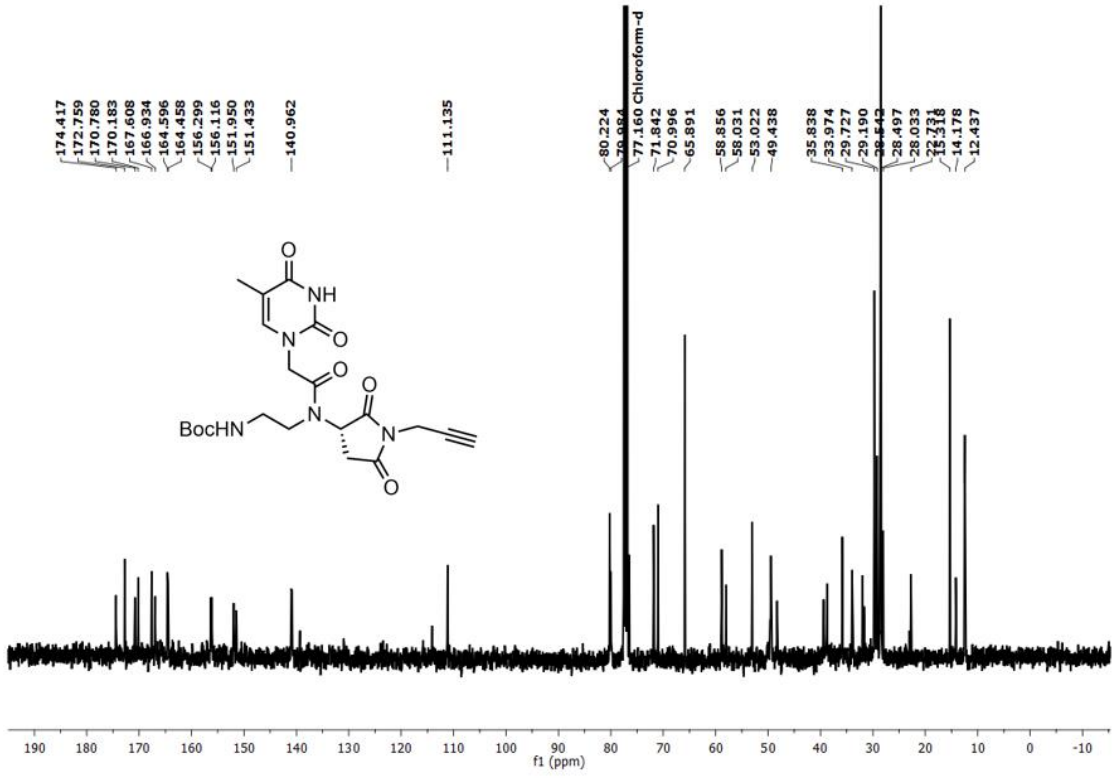
^{13}C {1H} NMR of compound 8 in DMSO-d_6 (101 MHz)



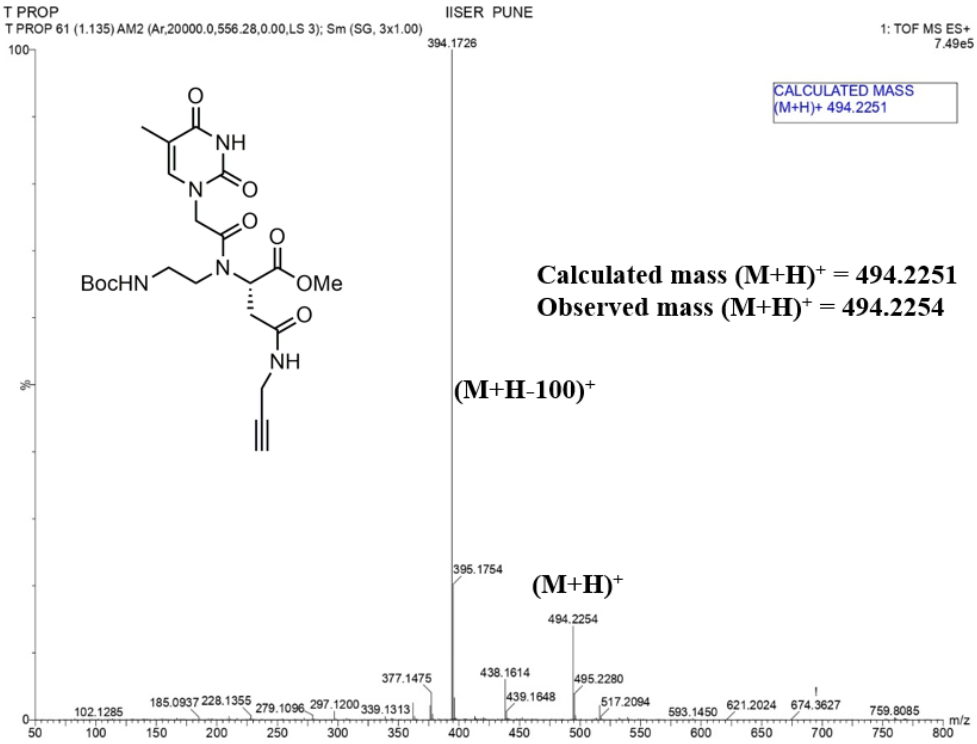
^{1}H NMR of compound 9 in DMSO-d_6 (400 MHz)



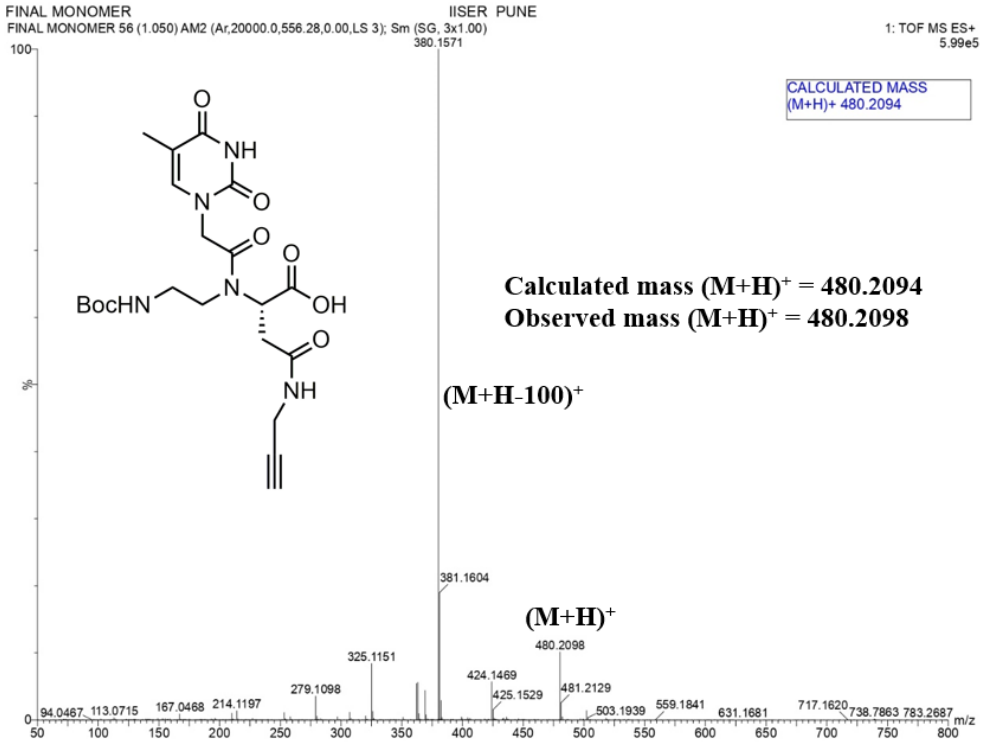
^{13}C {1H} NMR of compound 9 in $CDCl_3$ (101 MHz)



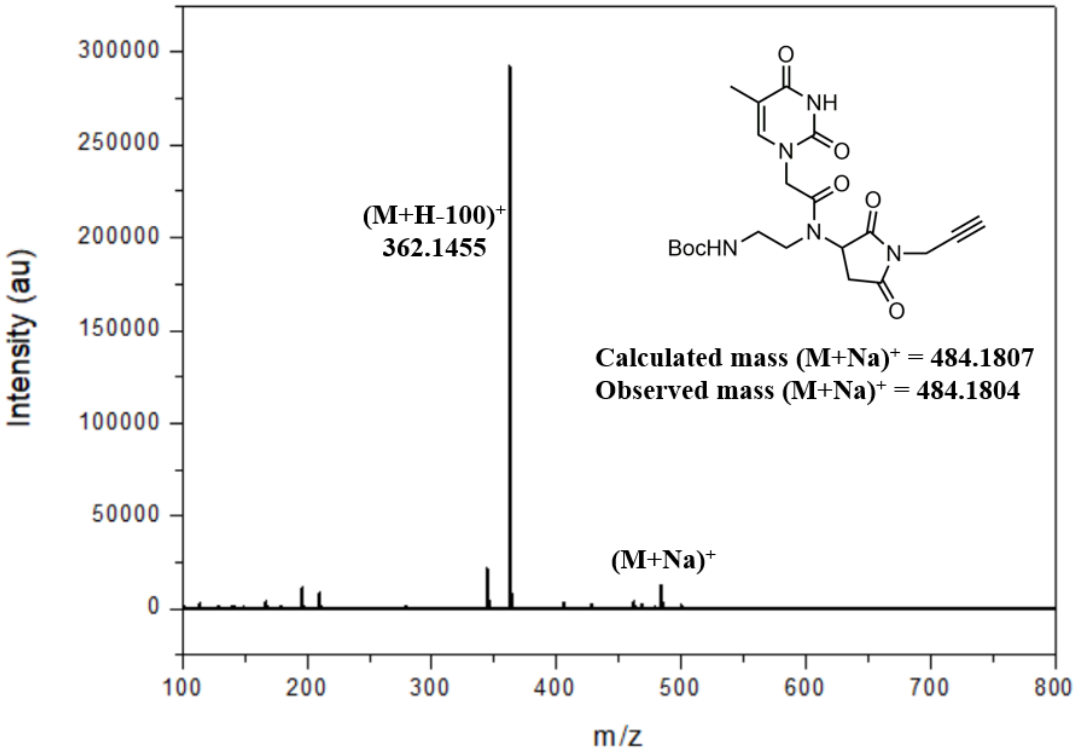
HRMS of compound 7



HRMS of compound 8



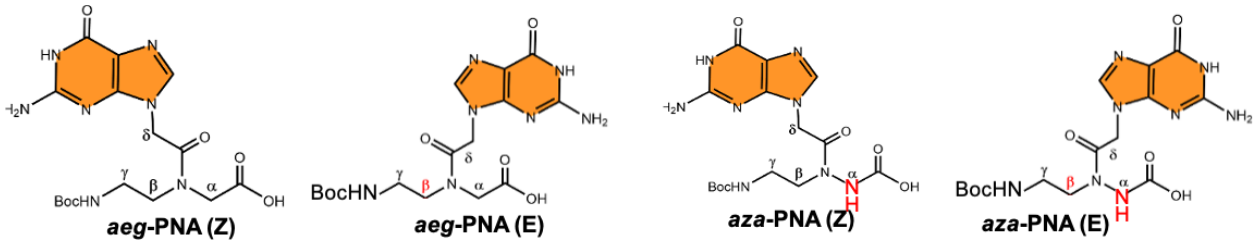
HRMS of compound 9



Synopsis of the thesis and future outlook

Structural Control of Peptide Nucleic Acid Rotamers: Aza-PNA and Optical Properties of Gemdimethyl PNA, Bimodal PNA Monomers and Oligomers

The aim of this thesis is to establish a relation between the structure, property and the orientation of the central tertiary amide carbonyl group of the PNA backbone (rotamers) to physical and optical properties of PNA monomers and derived oligomers. A new and novel type of PNA analogue termed "*aza*-PNA" in which the α-methylene of *aeg*-PNA is replaced by NH group with potential to form intramolecular hydrogen bonding with the N-terminal carbonyl group to fold the PNA backbone in an extended γ-turn like orientation. As this would lead to an urea linkage in the derived PNA oligomer, which is not conducive to form helical structures the corresponding homologue β-ala-*aza*-PNA-T was also designed for study. Both *aza*-T-PNA monomers was studied by ^{1}H NMR spectroscopy. The analysis indicated that α-

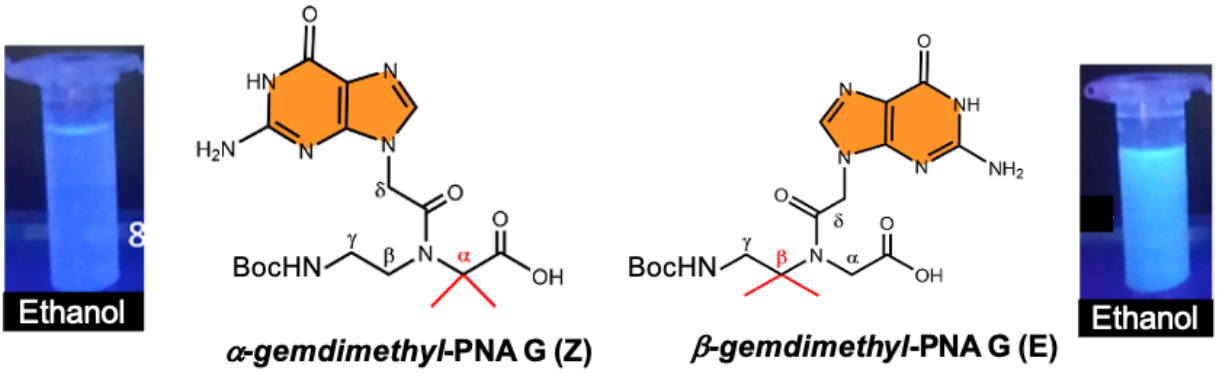


aza-PNA-T had a mixture of Z (major) and E (minor) rotamers while β-ala-*aza*-PNA-T monomer showed a single rotamer identified as E. Solvent and temperature dependent NMR studies suggested that α-NH is involved in intramolecular H-bonding with the backbone carbamate carbonyl (8-membered ring) as H-bonding with either δ-CO or ester CO lead to unfavourable 4 and 5-membered rings. The single crystal structure of the precursor intermediate N-acetylchloro aminoethyl-α-*aza* glycine showed 8-membered hydrogen bonding between the α-NH and N-terminal carbonyl group with rotamer in E-form. The crystal structure strongly supported such a possibility of intramolecular hydrogen bonding and extended γ-turn geometry of the *aza*-PNA backbone. Theoretical DFT calculations supported the experimental observations with 8-membered N_{α}-$H_A\cdots O_B$ bonding in *aza*-PNA monomers. This compacts the PNA backbone in *aza*-PNA compared to *aeg*-PNA with preferential stabilization of E-rotamer. The ^{13}C NMR chemical shifts of carbamate and δCO carbonyls in *aza*-PNA and *aeg*-PNA monomers exhibited upfield shift compared to *deaza* (*aeg*-PNA), suggesting their involvement in H-bonding (δ-CO with the $C_{\beta}H$; carbamate carbonyl with α-NH). The results

in this chapter established that in *aza*-PNAs intramolecular H-bonding leads to folding of backbone and preference for E-rotamer.

In a second approach, α/β-*gemdimethyl*-*aeg*-G PNA monomers were synthesized and ^{1}H NMR studies revealed that α-*gdm*-G-PNA monomer preferred Z-rotamer, while β-*gdm*-G-PNA preferred E-rotamer. Inspired from previous reports on fluorescence emission of *aeg*-PNA GC dimer that prefers Z-rotamer in crystals, the α/β-*gemdimethyl*-*aeg*-G locked exclusive in Z and E rotamers were studied for their fluorescence properties. The UV-spectra α/β-*gemdimethyl*-*aeg*-G monomer revealed a strong band at 309 nm, which is absent in *aeg*-PNA-G spectra. Fluorescence emission spectra of α/β-*gemdimethyl*-*aeg*-G monomer exhibited rotamer dependent blue fluorescence for Z-rotamer in α-*gemdimethyl*-PNA G and light green fluorescence for E-rotamer in β-*gemdimethyl*- PNA-G in both solution and solid state. The β-*gemdimethyl*-*aeg*-G monomer shows a broad emission band with maxima at 475 nm while α-*gemdimethyl*-*aeg*-G monomer shows vibrationally resolved emission bands with maxima at 402 nm and 428 nm respectively. Excitation dependent fluorescence study suggested these exhibit REES effect with fluorescence dependent on excitation wavelength. Solvent dependent emission study of the monomers shows that fluorescence emission of *gemdimethyl* PNA monomers are sensitive to microenvironment of the surroundings around the fluorophore. N-terminal carbonyl group seems to be important since fluorescence emission was weak in the *gemdimethyl*-PNA-amine monomers. The N-decanoyl-β-*gdm*-(6-Cl-2-aminopurine)-PNA monomer retained the 309 nm band after fatty acid conjugation but in N-decanoyl-α-*gdm*-(6-Cl-2-aminopurine)-PNA monomer the band shifted to 296 nm. With fatty acid conjugates, both α and β-*gemdimethyl*-PNA-G monomers exhibited bright blue fluorescence with vibrationally resolved emission bands with maxima at 402 and 428 nm. Further experiments ruled out the possibility of Aggregation Induced Emission (AIE) and suggested aggregation actually caused quenching (ACQ) in both α/β-*gemdimethyl*-*aeg*-G monomers and their fatty acid conjugates. N-decanoyl-β-*gdm*-(6-Cl-2-aminopurine)-PNA monomer having 100% E-rotamer forms more stable Langmuir monolayer at the air/water interface compared to the N-decanoyl-α-*gdm*-(6-Cl-2-aminopurine)-PNA monomer having 100% Z-rotamer. The derived dimers (β-*gdm*-*aeg*-G)$_2$-PNA and (*aeg*-G)$_2$-PNA dimer lack 309 nm UV-absorption band and show very weak emission in buffer and organic solvents. Thus, based on current theory of fluorescence of molecules without chromophore, the origin of fluorescence in α/β-*gemdimethyl*-PNA-G monomers lies in the superposition of n-π^* bonds in C-C, C-N and CO groups with through space conjugation. Such superpositions are dependent on molecular geometry which is

influenced by Z or E-rotamers in α/β-*gemdimethyl*-Guanine PNA monomers and fatty acid conjugated *gemdimethyl* monomers.



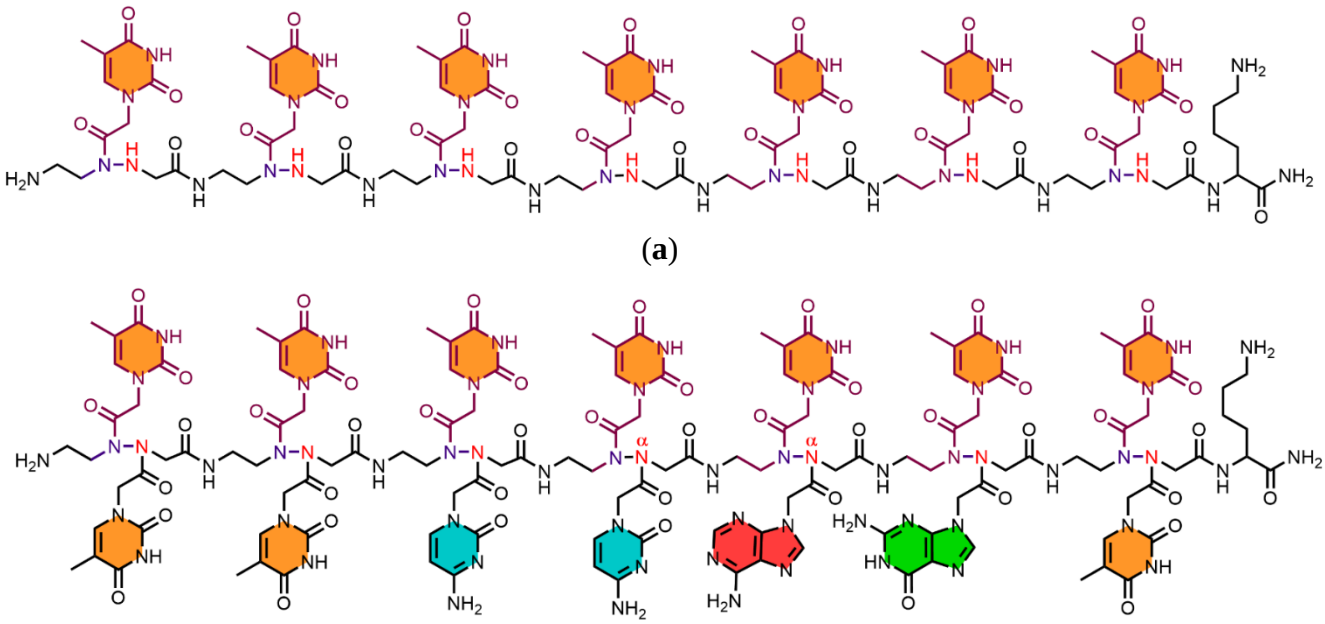
The studies were then directed towards Cα-monosubstituted *bimodal* PNA monomers and oligomers. *Bi-modal* PNAs are novel synthetic analogues which are structurally endowed to target two DNA strands simultaneously. The reported methods of synthesis is slightly complicated and involves orthogonal protection. A new method to synthesize oligomers was explored through pre-synthesis of *bm*-PNA di-nucleobase block monomers having two nucleobases present in the monomer, which are then used for sequential coupling. The synthesis of one such monomer block *bm*-PNA-TC is reported here and it was used for synthesis of corresponding homo-oligomeric bimodal PNA. The Cα(*S*)-(acetamido)-*bm*-PNA TC monomer exhibited a mixture of rotamers in solution with major Z-rotamer (57%) and minor E-rotamer (43%). In NOSEY, cross peak of α-CH with the β/γ-methylene suggests a possibly folded backbone for the Cα(*S*)-(acetamido)-*bm*-PNA TC monomer. This monomer exhibited bright blue fluorescence emission with a broad emission band having maxima at 400 nm and 425 nm. It also showed excitation dependent emission spectra (REES effect) for both monomer and the oligomer. The homo-oligomer (*bm*-PNA TC)$_n$ showed fluorescence in phosphate buffer (pH 12.5) with emission maxima at 480 nm, but no fluorescence at neutral pH. This chapter demonstrates that bimodal PNA monomers and oligomers are fluorescent without possessing any fluorophore and so belong to the class of *clusterluminogens.*

In a final attempt to establish the connectivity between the PNA structure and the rotamer preferences of PNA, C$_\alpha$-Succinimide modified PNA monomers in which the α-methylene undergoes cyclisation with the C-terminal C=O group to form N-substituted succinimide ring on the C-terminal of the PNA backbone was synthesized. Succinimide is a well-known pharmacophore in the field of medicinal chemistry and several FDA approved drugs are based on substitution around the succinimide pharmacophore. ^{1}H NMR analysis of

N-substituted C_α-succinimide modified PNA monomers reveal a rotamer mixture (1:1) with a higher E to Z exchange rate compared to acyclic *aeg*-PNA (3:1). ^{1}H-^{1}H NOSEY the NOSEY cross peak of α-CH with the β/γ-methylene suggest folded conformations E and Z rotamers, Crystal structure of the cyclic C_α-Succinimide PNA exhibited Z-rotamer as in *aeg*-PNA. The crystal structure showed T: T base pairing and intermolecular hydrogen bonding between the N-terminal BocNH and thymine C=O group. Overall, this chapter demonstrates how cyclisation of backbone around the C-terminus affects the rotamer preferences.

Future outlook: The work presented in this thesis sheds new light on the role of rotamers and how stercoral variations in backbone bring selectivity in stabilization of Z or E rotamers. The new PNA analogue *aza*-PNA shows intramolecular H-bond in backbone and has potential to give PNA oligomers with new helical properties, which could be useful in developing new PNA applications. *aza*-PNA, *Bi-modal* PNA, *gemdimethyl* PNA and succinimide PNA have unique properties structural and optical (UV, fluorescence) properties.

1. Novel *aza*-PNAs possesses interesting γ-turn motif which in unique in the PNA literature. Fully modified homo and hetero oligomers derived from *aza*-PNAs could exhibit interesting cell permeation ability and bio-physical properties with complementary DNA and RNA as it contains a free -NH group and a γ-turn motif in the monomer and oligomer.



(**a**)

(**b**)

Figure 1. (a) Fully modified β-ala-*aza*-PNA-T oligomer (b) Possible *aza*-bimodal PNA analogues

On the other hand, the β-ala-*aza*-PNA-T could act as scaffold for *aza-bimodal* PNA. *bimodal* PNAs are novel synthetic analogue of PNA which may simultaneously target two DNA strands. Classical *bi-modal* PNA involves derivatization at Cα or Cγ side chain with an active chiral center which may undergo racemization in presence of coupling agent. *aza-bimodal* would reveal the effect of preorganization resulted due to chirality on the PNA: DNA binding and may have potential biological application.

2. Most of work recent work reported till now for *bi-modal* lacks conformational constraint in the system and have a mixture of rotamers. Fixing the rotamers of *bi-modal* PNA into a single rotamer could potentially open up versatile application of *bi-modal* PNA as literature suggests constrained PNA selectively binds complementary DNA over RNA or vice versa. It could be possible to design a *bi-modal* PNA in which the tertiary amide face is selective towards complementary DNA while the Cα or Cγ face is selective towards the complementary RNA or vice versa. Apart from that imposing conformational constraint could induce interesting fluorescence properties of the *bi-modal* PNA monomers and oligomer which may have interesting biological application.

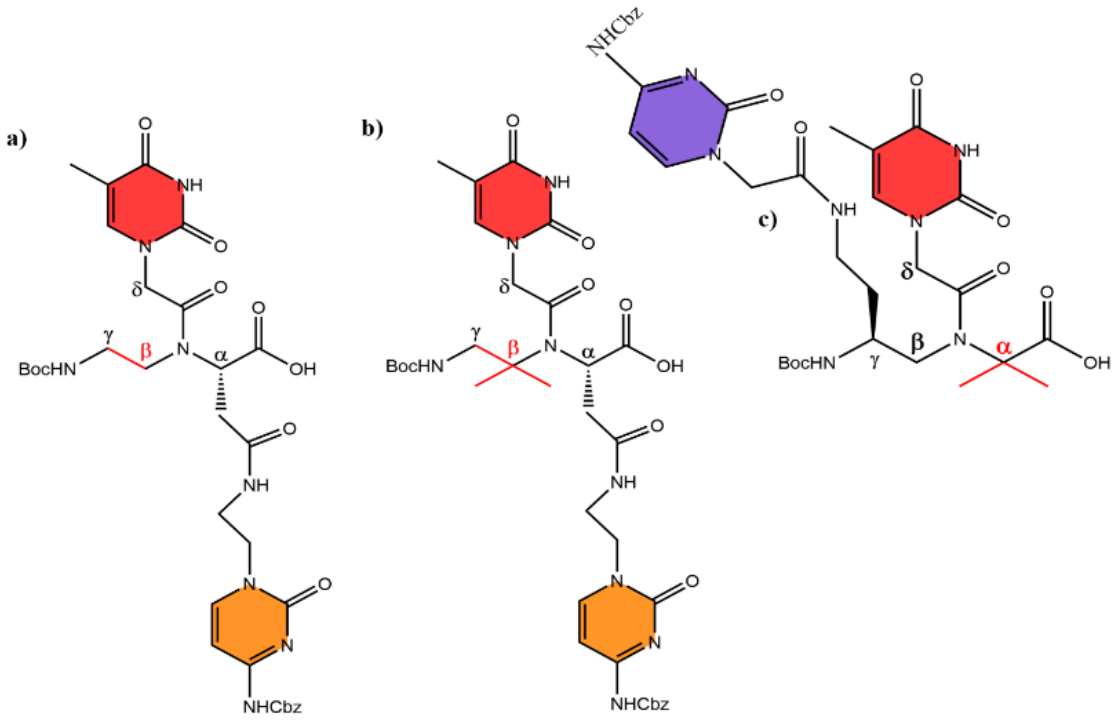


Figure 2. (a) Classical *bimodal* PNA TC-block PNA monomer, (b) Rationally designed Cα *bi-modal* PNA monomers for selective binding and fluorescence study, (c) Rationally designed Cγ *bi-modal* PNA monomers for selective binding and fluorescence study.

3. Fluorescence of *gemdimethyl* monomer reported in the thesis does not involve metal coordination. Conjugation of the PNA backbone with a ligand like pyridine or quinoline could induce metal coordination ability of the resulting monomers. Metal coordination of the derived monomer could result interesting fluorescence properties of the *gemdimethyl* monomer and could act as fluorescence-based metal sensor.

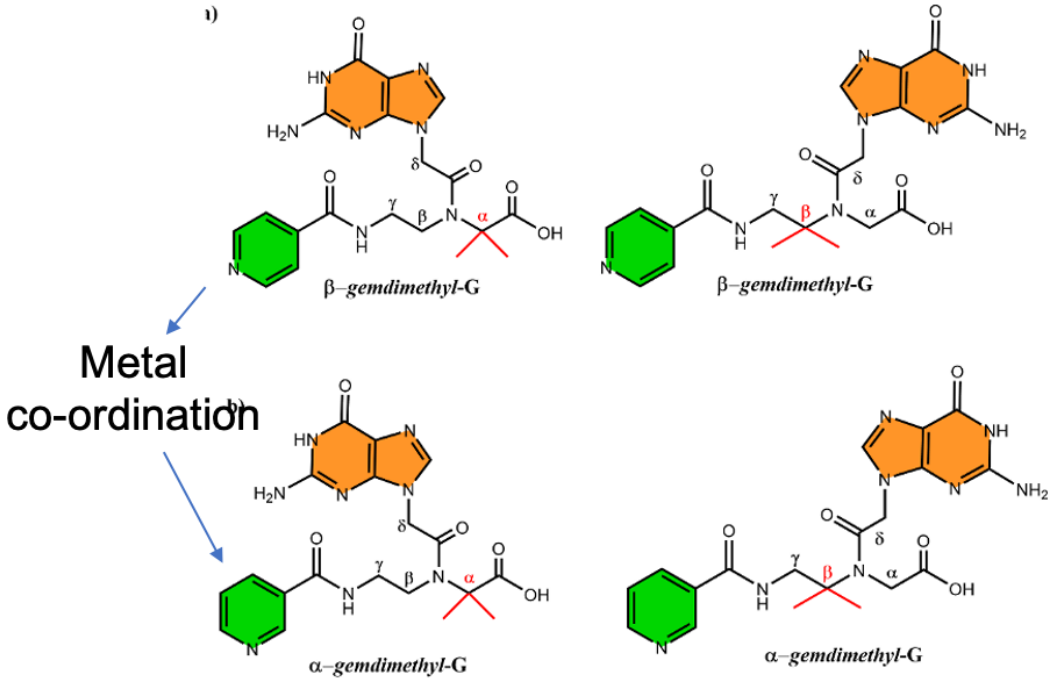


Figure 3. a) β-*gemdimethyl-aeg*-G PNA based metal sensor b) α-*gemdimethyl-aeg*-G PNA based metal sensor

Apart from that it could be possible to design β/α-*gemdimethyl* Guanie/Cytosine based oligomers which could form G-quadruplex or i-motif in standard condition and may exhibit interesting fluorescence emission. Depending on the type and orientation (parallel and antiparallel) of G-quadruplex or i-motif the oligomer could exhibit distinct fluorescence emission properties.

4. Primary results of succinimide modified PNA suggest interesting structural and preorganization properties. It could be interesting to study the PNA: DNA binding and other

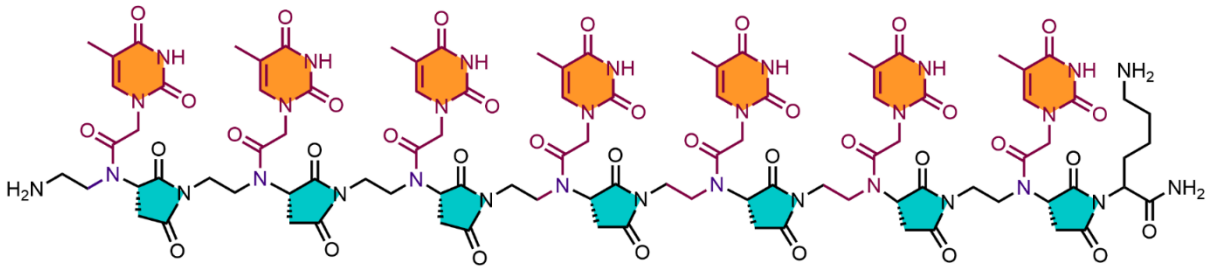


Figure 5. Succinimide modified PNA oligomer for the study of PNA: DNA binding

properties of the derived oligomers of succinimide modified PNA oligomer (Figure 4).