Designed nucleic acid mimics with internucleobase distance-complementarity

Thesis Submitted to
The University of Pune for degree of

Doctor of Philosophy
in
Chemistry

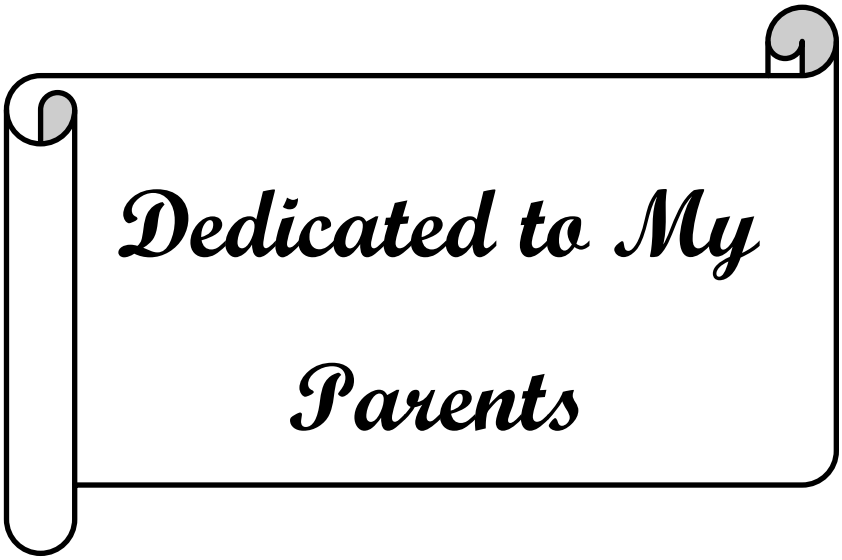
By

SACHIN GOKHALE

Research Supervisor
Dr. (Mrs.) Vaijayanti A. Kumar

Division of Organic Chemistry
National Chemical Laboratory
Pune-411008

July 2010


Dedicated to My
Parents

CERTIFICATE

This is to certify that the work presented in the thesis entitled "**Designed nucleic acid mimics with internucleobase distance-complementarity**" submitted by Sachin S. Gokhale, was carried out by the candidate at the National Chemical Laboratory Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

Dr. (Mrs.) Vaijayanti A. Kumar **July 2010**

(Research Supervisor)

Division of Organic Chemistry

National Chemical Laboratory

Pune-411008

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled "**Designed nucleic acid mimics with internucleobase distance-complementarity**" submitted for the award of degree of ***Doctor of Philosophy*** in Chemistry to the University of Pune has not been submitted by me to any other university or institution. This work was carried out by me at the National Chemical Laboratory, Pune, India. Such materials as obtained from other sources have been duly acknowledged in the thesis.

Sachin S. Gokhale **July, 2010**

National Chemical Laboratory

Pune- 411 008

Acknowledgement

It gives me an immense pleasure to express my deep and sincere gratitude to my research supervisor Dr. (Mrs.) Vaijayanti A. Kumar for all the advice, guidance and encouragement during every stage of this work. She made me realise the importance of doing quality research, from working table to formulation of ideas to presentation of results. The confidence she had in me, willingness to share new ideas, excitements helped me in a real sense to shape my research career.

Thank you doesn't seem sufficient but it is said with all respect

Dr. (Mrs.) Moneesha deserves special thanks. Her encouragement, understanding and suggestions were invaluable during my stay in the lab and went a long way towards the completion of this thesis.

My sincere thanks also go out to Mrs. Anita Gunjal for her help during the course of this work. I am also thankful to Dr. (Mrs.) V. S. Pore, Prof. K. N. Ganesh, Prof. D. D. Dhavale, Dr. Thulasiram and Dr. Anil Kumar, Dr. Sourav pal for their advice and various kinds of help.

I am also thankful to Mrs. M. V. Mane for HPLC analysis, Dr. Mahesh Kulkarni, Mr. S. S. Deo, and Mrs. Shanta Kumari for MALDI-TOF and LC-MS analysis. The kind support from NMR group is greatly acknowledged and I am specially thankful to Mrs. Usha D. Phalgune. I am also Thankful to Dr. Souvik Maiti (IGIB, New Delhi) and his students for their help.

This page would be incomplete without the mention of my seniors and colleagues who have gone out of their way to help in various capacities and have been my extended family throughout the tenure of my life in NCL. Thank you Dr. Khirud, Madhuri, Seema, Namrata, Harshit, Harshal, Parmeshwar, Venu, Kiran, Manoj, Anjan, Tanaya.

I thank to Dr. Ngendra, Dr. Dinesh, Dr. Sunil, Dr. Umashankara, Dr. Raman, Dr. Gourishankar, Dr. Amit, Dr. Nilkanth, Ashwini, Shreedhar, Gitali, Nasrin, Roopa, Pradnya, Manaswini, Mahesh, Deepak, Nitin, Vijay, Satish, Rajendra, Madhuri Patil, Shobana, Bharat, Suresh,

My special thanks to my MSc classmates and friends for their encouragement.

I thank Organic chemistry Division office staff for their help. I thank Mr. Bhumkar for the laboratory assistance.

I am grateful to UGC, New Delhi for research fellowship and Dr. S. Sivaram, Director, NCL for providing infrastructure.

Finally, I would like to acknowledge in a special way my parents, sisters and other family members, who were extremely patient and tolerant towards my erratic hours of work and constantly encouraged me to excel in whatever I did.

Contents

Publication /Symposia i

Abbreviations ii

Abstract vi

1 Chapter 1: Introduction

1.1 Introduction to nucleic acids
- 1.1.1 Primary structures of DNA and RNA 1
- 1.1.2 Sugar puckering in nucleosides 2
- 1.1.3 Base Pairing *via* Hydrogen bonding 3

1.2 Antisense Technologies
- 1.2.1 Introduction 4
- 1.2.2 Non-RNA-degrading mechanisms 5
- 1.2.3 RNA Cleavage mechanism 7

1.3 Modified oligonucleotides 8
- 1.3.1 Modifications of phosphate linkage 9
- 1.3.2 N3'→P5' Phosphoramidate Nucleic Acids 11
- 1.3.3 Locked Nucleic acid (LNA) 13
- 1.3.4 Oligonucleotides with dephosphono backbone 15
- 1.3.5 *xylo* nucleic acids (XNA) 17

1.4 Peptide Nucleic Acids 18
- 1.4.1 PNA-DNA complex formation and their structures 19
- 1.4.2 Inhibition of gene expression 22
- 1.4.3 Chemical Modifications of PNA 23
- 1.4.4 Cellular Uptake of PNA 30
- 1.4.5 Applications of PNA 31

1.5 Spectroscopic methods for studying DNA/RNA Interactions 34

1.6 The present Work 37

1.7 References 40

2 Chapter 2

***N*-(pyrrolidinylethyl)glycyl based pet-PNA and its amino derivative :**

2.1 Synthesis of monomer units and their incorporation in aegPNA sequences 51

- 2.1.1 Introduction 51
- 2.1.2 Rationale and objectives of the present work 52
- 2.1.3 Synthesis of Thymine pet-PNA and Amino pet-PNA monomers, Results and Discussion 55
- 2.1.3a Synthesis of pyrrolidinylethyl thymine (H- pet) PNA monomer 55
- 2.1.3b Synthesis of Aminopyrrolidinylethyl thymine (Am-pet-PNA) monomer 57
- 2.1.3c Synthesis of aegPNA monomers 59
- 2.1.4 Solid phase PNA synthesis 59
- 2.1.4a General protocols for solid phase PNA synthesis 59
- 2.1.4b Synthesis of H-pet, Amino-pet-PNA, Guanidino-pet-PNAs 61

2.1.4c Post synthetic conversion of amino to guanidino groups on solid phase 62
2.1.4d Cleavage of the PNA Oligomers from the Solid Support 63
2.1.5 Purification and MALDI-TOF characterization of oligomers 65
2.15a Purification of PNA oligomers by HPLC 65
2.1.6 Synthesis of complementary oligonucleotides 66
2.1.7 Summary 67
2.1.8 Experimental 68
2.1.9 Appendix 82

2.2 Biophysical studies of H-pet-PNA, Amino-pet-PNA and Guanidino-pet-PNA towards complementary DNA/RNA

2.2.1: Introduction 116
2.2.2 Rationale and objectives of the present work 116
2.2.3 Biophysical spectroscopic techniques for studying PNA- DNA/ PNA-RNA interactions 117
2.2.4 UV melting studies 119
2.2.4a Homopyrimidine PNA-T_8 sequences: UV studies 119
2.2.4b Mixed pyrimidine PNA sequences (b_2a_2): UV studies 123
2.2.4c Mixed purine-pyrimidine PNA sequences: UV-melting studies 125
2.2.5 CD analysis of single stranded PNAs 129
2.2.6 Summary 130

2.3 References 131

3 Chapter 3

3.1 Design and Synthesis of *iso*-TANA dimers, their incorporation into DNA oligomers and biophysical studies with 136

3.1.1 Introduction 136
3.1.2 Rationale, Design of *iso*-TANA and objectives of the present work 137
3.1.3 Methodology, Results and Discussion 141
3.1.3a Synthesis *iso*-thioacetamido linked Thymidine-Thymidine (*iso*-tst), 2'-deoxy-5-Methyl cytidine-Thymidine(*iso*-cst), 2'-deoxy-5-Methyl cytidine–2'-deoxy-cytidine dimers (*iso*-csc) and thioacetamido linked 2'-deoxy-5-Methyl cytidine–2'-deoxy-cytidine dimer (csc) and their phosphoramidite derivatives 141
3.1.3b NMR and CD studies of *iso*-TANA dimers 145
3.1.4 Synthesis, purification and mass spectral analysis of the chimeric *iso*-TANA-DNA oligomers 154
3.1.4a Solid phase synthesis of *iso*-TANA modified DNA oligonucleotides by phosphoramidite method 154
3.1.4b Purification and MALDI-TOF characterization of *iso*-TANA modified oligomers 156
3.1.4c Synthesis of complementary oligonucleotides 157
3.1.5 UV-melting studies *iso*-TANA modified DNA 157
3.1.5a Polypyrimidine sequences 158
3.1.5b Mixed purine-pyrimidine sequences 159
3.1.6 Summary 163

3.1.7 Experimental 164
3.1.8 Appendix 178
3.1.9 References 211

3.2 Synthesis of Modified PNA sequences with *iso*-thioacetamidonucleic acids monomers (*iso*-TANA) and their Biophysical studies 216

3.2.1 Introduction 216
3.2.2 Synthesis of *iso*-TANA monomers results and discussion 217
3.2.2a Synthesis of *iso*-TANA thymine monomer 217
3.2.2b Synthesis of *iso*-TANA 5-methyl cytosine monomer 218
3.2.2c Synthesis of Fmoc protected aegPNA thymine monomer 219
3.2.3 Solid phase synthesis of *iso*-TANA, aegPNA and aegPNA-TANA chimeric oligomers 219
3.2.3a Synthesis of *iso*-TANA homooligomer and *iso*-TANA-PNA chimeric oligomers 221
3.2.3b Cleavage of the iso-TANA and aegPNA-*iso*-TANA oligomers from the solid support 221
3.2.4 Purification and MALDI-TOF characterization of oligomers 221
3.2.5 UV-melting studies *iso*-TANA modified oligomers 222
3.2.6 Summary 224
3.2.7 Experimental 225
3.2.8 Appendix 232
3.2.9 References 246

List of Research Publications:

1. '*iso*-Thioactaamido nucleic acids (*iso* TANA) synthesis and DNA binding studies of Lyxo/ Ribo oligonucleotides **Sachin Gokhale**.; Vaijayanti Kumar.* *Nucleic acid Symp.ser,* **Sept 2008**, *52***,** 145-146
2. Amino/guanidino-functionalized N-(pyrrolidinyl-2-ethyl)glycine- based pet-PNA: Design, synthesis and binding with DNA/RNA S**achin S. Gokhale**; Vaijayanti A. Kumar***** accepted in *Org. Biomol. Chem*. 2010 (/Org. Biomol. Chem/., 2010, DOI: 10.1039/c004005c)
3. Probing binding preferences of DNA and RNA: Backbone chirality of TANA and *iso*TANA to differentiate DNA versus RNA selectivity **Sachin S. Gokhale**; Khirud Gogoi; Vaijayanti A.Kumar***** (manuscript under preparation)

Patent:

US Patent Application

1. Mercaptoacetamido Nucleic Acids: Five atom mercaptoacetamido linked oligo-2', 3'-dideoxy-3'ribo/lyxo-amino-nucleosides as DNA mimics and their RNA/DNA binding selectivity. K. Gogoi, **S. S. Gokhale**, A. D. Gunjal, Vaijayanti A. Kumar

Symposia Attended/Poster Presentations:

1. Joint Symposium of 18th International Roundtable on Nucleosides, Nucleotides and Nucleic Acids (IRTXVIII) and 35th International Symposium on Nucleic Acids Chemistry (SNAC) held in Kyoto, Japan, September 8-12, 2008: Poster presentation **''*iso* Thioactaamido nucleic acids (*iso* TANA) synthesis and DNA binding studies of Lyxo/ Ribo oligonucleotides**'' **Sachin Gokhale**.; Vaijayanti Kumar.*
2. Attended 4th INSA-KOSEF Symposium in Organic Chemistry. Contemporary organic chemistry and its future directions. National Chemical Laboratory, Pune India, January 2009.
3. Attended 11th CRSI National Symposium in Chemistry National Chemical Laboratory, Pune, India, February 2009.

Abbreviations

β- ala	β-alanine
A	Adenine
Ac	Acetyl
Ac_2O	Acetic anhydride
aeg	Aminoethylglycine
ap	Antiparallel
arg	Arginine
aq.	Aqueous
bep	Backbone extended PNA
Bz	Benzoyl
C	Cytosine
Cat	Catalytic/catalyst
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
2'-dA	Deoxyadenosine
DCA	Dichloroacetic acid
DCM	Dichloromethane
dG	2'-Deoxyguanosine
DIPEA/DIEA	Diisopropylethylamine
DMAP	4',4'-Dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	*N*,*N*-Dimethyl sulfoxide

DNA	2’-deoxyribonucleic acid
ds	Double stranded
EDTA	Ethylenediaminetetraacetic acid
Et	Ethyl
EtOAc	Ethyl acetate
Fmoc	9-Fluorenylmethoxycarbonyl
FT	Fourier Transform
g	gram
G	Guanine
hrs	Hours
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronum hexafluorophosphate
HIV	Human Immuno Difficiency Virus
HOBt	1-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
Hz	Hertz
IR	Infra red
L-	Levo-
LC-MS	LiquidChromatography-Mass Spectrometry
Lys	Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-Methyl benzhydryl amine

MF	Molecular formula
mg	milligram
MHz	Megahertz
mins	minutes
µL	Microliter
µM	Micromolar
mL	milliliter
mM	millimolar
mmol	millimoles
m.p	melting point
Ms	Methanesulfonyl
MS	Mass spectrometry
MW	Molecular weight/Microwave
N	Normal
nm	Nanometer
NMR	Nuclear Magnetic Resonance
p	Parallel
pfp	Pentafluorophenyl
POM	Pyrrolidine –amide linked oligonucleotides mimics
ppm	Parts per million
Pro	Proline
PS-oligo	Phosphorothioate-oligo
Py	Pyridine

PNA	Peptide Nucleic Acid
pet	pyrolidinylethyl
R	Rectus
R_f	Retention factor
RP	Reversed Phase
rt	Room temperature
S	Sinister
SPPS	Solid Phase Peptide Synthesis
t-Boc	*tert*-Butoxycarbonyl
T	Thymine
TANA	Thio acetamido Nucleic Acids
Et_3N	Triethylamine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T_m	Melting temperature
Tos	p-toluene sulfonyl
U	Uridine
UV-Vis	Ultraviolet-Visible

Abstract

The Thesis entitled ‘‘**Designed nucleic acid mimics with internucleobase distance-complementarity**” has been divide into three chapters.

Chapter 1: Introduction to nucleic acids

Chapter 2

Section 2.1: *N*-(pyrrolidinylethyl)glycyl based pet-PNA and its amino derivative: Synthesis of monomer units and their incorporation in aegPNA sequences

Section 2.2: Biophysical studies of H-pet-PNA, Amino-pet-PNA and Guanidino-pet-PNA towards complementary DNA/RNA

Chapter 3

Section 3.1: Design and Synthesis of *iso*-TANA dimers, their incorporation into DNA oligomers and biophysical studies

Section 3.1: Synthesis of *iso*-TANA monomers, their incorporation into PNA oligomers and biophysical studies

Chapter 1: Introduction to natural and modified nucleic acids.

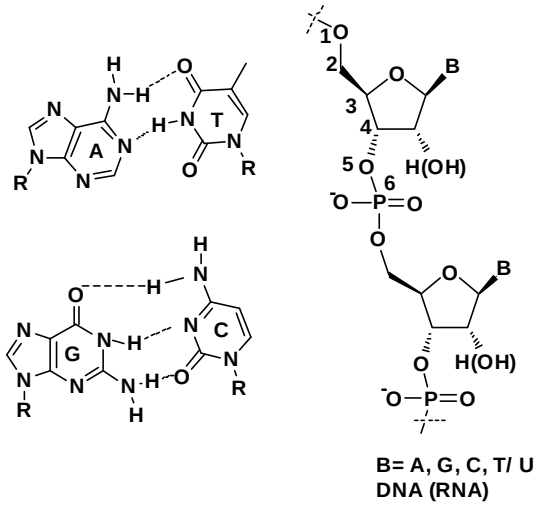


Figure 1: Structure of DNA/ RNA

The potential of modified oligonucleotides to act as antisense agents that can inhibit the expression of a target gene in a sequence-specific manner may be used for

therapeutic and other biological applications. Major challenges for modified oligonucleotides as therapeutic agents are sequence-specific recognition of RNA, low cellular uptake and directional selectivity. In order to overcome these shortcomings many modifications in oligonucleotides have been made. Some of the promising modified oligonucleotides are phosphorothioate, morpholino, PNA etc. In some of the modifications sugar- phosphate backbone was modified or the sugar ring was replaced with other heterocyclic or carbocylic rings, whereas in *aeg* PNA whole DNA backbone was replaced by acyclic, achiral amino ethyl glycyl units. PNA's strong binding affinity towards DNA and RNA and enzymatic stability are major advantages over other modifications. This chapter is focused on PNA modifications and DNA modifications which helped to overcome some shortcomings of earlier modifications. It is found that the six atom phosphate backbone could be replaced by seven atom amide linkage leading to folded or extended geometries to maintain in inter nucleobase distance complementarity. Also ribose sugars in DNA and RNA have a preference for either N or S conformation. This conformational state always has an impact on the binding affinities towards DNA/RNA. This conformational change occurs due to the substituents on the sugar ring especially 3' and 2' substituents. This substituent effect also has been discussed in this Chapter.

Chapter 2

Section 2.1: *N*-(pyrrolidinylethyl) glycyl based pet-PNA and its amino derivative : Synthesis of monomer units and their incorporation in aegPNA sequences

Introduction

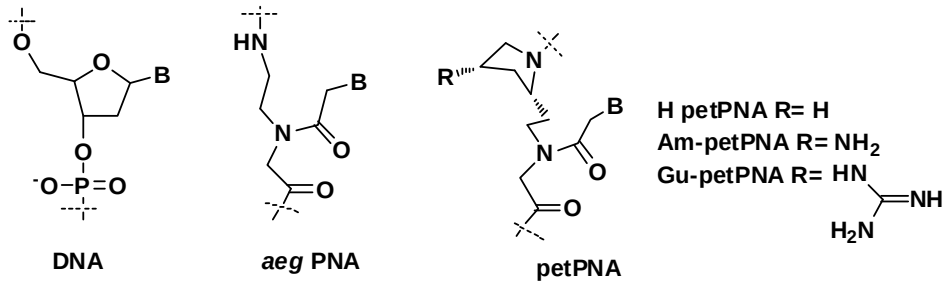


Figure 2: Structure of DNA, aegPNA and petPNA

Peptide nucleic acids (PNA, Figure 2) are a promising class of modified oligonucleotides, where sugar phosphate backbone of DNA is replaced by achiral *N*-(2-

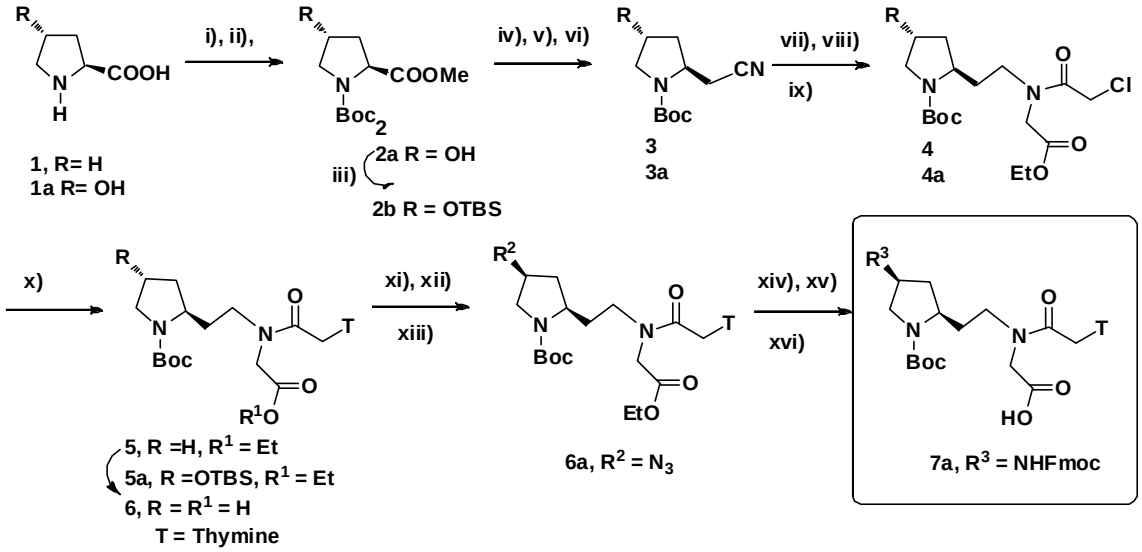
amino ethyl) glycine units. PNA hybridizes with complementary DNA/ RNA strand through Watson-Crick hydrogen bonding. Major advantage of PNA is its resistance towards enzymatic degradation; this advantage makes PNA an attractive reagent for biotechnological applications. PNA suffers from some drawbacks such as low water-solubility, low cell permeability and ambiguity in directional selectivity which restricts its use as an antisense drug. To overcome these issues many modifications in PNA were suggested. Positively charged natural and unnatural peptides like TAT and poly argenine were found to have good cell penetrating capacities. Conjugation of these peptides with PNA has successfully transported PNA into cells. But this conjugation cannot release PNA so efficiently because of problem of endosomal trapping of these peptide conjugates. Thus PNA backbone itself carried groups that help in cellular uptake, would be desirable.

GPNA is a recent such kind of example of modified PNA which has shown comparable binding with complementary RNA as of aegPNA. As GPNA has an argenine side chain in backbone, it has shown good cellular permeability. Introduction of chirality and cyclic structures are known to reduce the conformational flexibility of acyclic, achiral *aeg* PNA. In this chapter we present the synthesis of *N*-(pyrrolidinyl ethyl) glycyl PNA (pet-PNA) and corresponding amino, guanidino pet-PNA monomer blocks, their incorporation into PNA sequences and biophysical studies with complementary DNA/ RNA.

Previously reported pmg PNA showed destabilized complexes towards complementary DNA/ RNA, possibly due to the constrain of the pyrrolidyl ring in PNA. This effect in pmgPNA was compensated by backbone extension in petPNA. The various extended backbone pyrrolidyl PNA reported from our laboratory and other groups showed preferential binding to RNA over DNA.

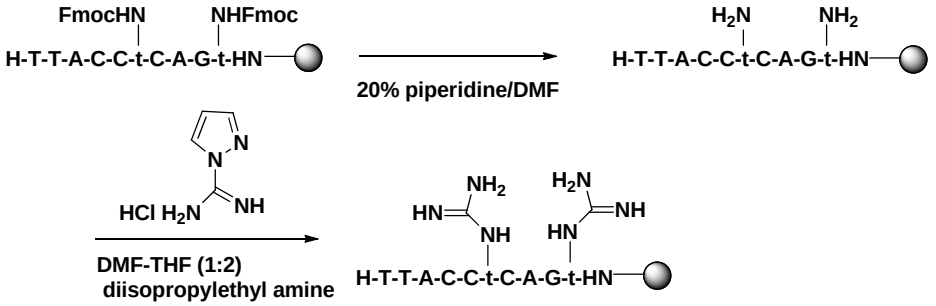
Synthesis of PNA monomers and solid phase PNA synthesis

A simple synthetic scheme was followed towards synthesis of H- pet and amino-pet-PNA monomers. Incase of H-pet-PNA, proline was used as starting material, while for Amino-pet-PNA monomer 4-*trans*-hydroxy-L- proline was used. In both the synthesis cyano functionality was used to increase the carbon chain length. These Boc protected monomers were used for solid phase PNA synthesis, Boc- aegPNA monomers were synthesized with by reported procedure.



Scheme-1: Reagents and conditions i) $SOCl_2$/MeOH, ii) Boc_2O/ 1,4-dioxane/water, iii) TBSCl/ imidazole, iv) $NaBH_4$-LiCl/dry THF-EtOH, v) MsCl/Py 0^oC, vi) NaCN/ DMSO, 60^oC, vii) H_2/Raney-Ni/ MeOH, viii) $BrCH_2COOEt$/ TEA-MeCN, ix) $ClCH_2COCl$ /$NaHCO_3$, 1,4-dioxane /water, x) thymine / K_2CO_3/ DMF, x) 1M TBAF in THF, xi) MsCl/ Py, 0^oC, xii) NaN_3/ DMF, 60^oC, xiii) H_2/Pd-C, MeOH, xvi) 1N LiOH/ MeOH, xv) Fmoc-succinimide / $NaHCO_3$, 1,4-dioxane/ water

H-pet-PNA, Amino-pet-PNA(Am-pet-PNA) were synthesized on Solid support using standard solid phase peptide synthesis. Guanidino-pet-PNA (Gu-pet-PNA) were synthesised from Am-pet-PNAs using the procedure in scheme 2.



Scheme-2: Global guanidinylation of PNA on solid support

PNA sequences which were synthesized were purified by RP- HPLC. Purity of sequences was checked by RP-HPLC and sequences were characterized by MALDI-TOF mass spectrometry.

Table.1 PNA Oligomers synthesized, their HPLC t_R and MALDI-TOF mass characterization.

Entry	Sequence	HPLC t_R (min)	Mass calcd/obsd
PNA 1	H-TTTTTTTT-Lys-NH_2	8.90	2273.93/2274.91
	H-TTTTTTTt^R-Lys-NH_2		
PNA2	R = H	9.48	2328.3/2329.20
PNA3	R= NH_2	8.48	2343.99/2366.13(M-1+Na^+)
PNA4	R= NH-C(= NH) NH_2	8.66	2386.0/2386.57
	H-TTTTt^RTTT-Lys-NH_2		
PNA5	R = H	9.21	2328.3/2329.2
PNA6	R= NH_2	14.16	2343.99/2343.34
PNA7	R= NH-C(= NH)NH_2	14.58	2385.02/2382.33
PNA	H-$t^H t^H t^H t^H t^H t^H t^H t^H$ -Lys-NH_2	13.27	2707.92/2706.64
PNA9	H-CTTCTTCCTT-Lys-NH_2	11.19	2746.4/2741.15
PNA10	H-Lys-Lys-Lys-Lys-CTTCTTCCTT-Lys-NH_2	12.00	3271.1/3366.39
	H-CTt^RCTt^RCC$t^R t^R$-Lys-NH_2		
PNA11	R= NH_2	13.47	3026.40/3034.50
PNA12	R= NH-C(= NH) NH_2	14.64	3194.35/3201.40
PNA13	H-TTACCTCAGT-Lys-NH_2	9.04	2804.16/2827.07(M+Na^+)
PNA14	H-TTACCTCAGt^R-Lys-NH_2 R= H	8.77	2859.32/2859.38
	H-TTACCt^RCAGt^R-Lys-NH_2		
PNA15	R=H	9.44	2912.27/2910.43
PNA16	R= NH_2	8.51	2942.29/2940.6
PNA17	R= NH-C(= NH) NH_2	8.24	3028.34/3024.5

Section 2.2: Biophysical studies of H-pet-PNA, Am-pet-PNA, Gu-pet-PNA towards complementary DNA and RNA

UV-melting studies of PNA complexes with complementary DNA/RNA.

These sequences were tested for binding studies with complementary DNA/ RNA and

The results are tabulated as below.

Table 2A: T_8 PNA sequences and their binding studies with complementary DNA/ RNA, mismatch DNA/ RNA

Entry	Sequence	DNA-1	RNA-1	DNA-5	RNA-5
PNA-1	H-TTTTTTTT-Lys-NH_2	41.6	41.2	30.0	29.8
	H-TTTTTTTt^R-Lys-NH_2				
PNA -**2**	R= H	26.7	48.7	nd	20.9
PNA-**3**	R= NH_2	45.7	49.9	26.8	28.5
PNA -**4**	R= NH-C(= NH) NH_2	51.6	58.7	32.6	31.8

DNA-1: 5’ GCAAA AAAACG 3’, **RNA-1**: 5’GCAAA AAAACG 3’, **DNA-5**: GCAAATAAAACG3’
RNA-5: 5’GCAAAUAAACG 3’ nd= not detectable

Table 2B: T_8 PNA sequences and their binding studies with complementary DNA/ RNA, mismatch DNA/ RNA

Entry	Sequence	DNA-1	RNA-1	DNA-5	RNA-5
PNA-1	H-TTTTTTTT-Lys-NH_2	41.6	41.2	30.0	29.8
	H-TTTTt^RTTT-Lys-NH_2				
PNA- 5	R= H	27.2	37.4	nd	16.5
PNA- 6	R= NH_2	40.7	36.8	19.4	18.2
PNA- 7	R= NH-C(= NH) NH_2	36.0	36.4	21.4	18.9
PNA- 8	H-$t^H t^H t^H t^H t^H t^H t^H t^H$ -Lys-NH_2	52.9	40.6	-	-

DNA-1: 5' GCAAAAAAAACG 3', **RNA-1**: 5'GCAAAAAAAACG 3', **DNA-5**: GCAAATAAAACG3' **RNA-5**: 5'GCAAAUAAAACG 3' nd= not detectable

Table 3: Mixed pyrimidine PNA sequences and their binding studies with complementary DNA/ RNA, mismatch DNA/ RNA

Entry	Sequence	DNA-2	RNA-2	DNA-6	RNA-6
PNA-9	H-CTTCTTCCTT-Lys-NH_2	52.0	55.6	42.6	45.2
PNA-10	H-Lys-Lys-Lys-Lys-CTTCTTCCTT-Lys-NH_2	63.4	59.4	57.8	53.1
	H-CTt^RCTt^RCC$t^R t^R$-Lys-NH_2				
PNA-11	R= NH_2	63.7	65.5	49.9	57.6
PNA-12	R= NH-C(= NH) NH_2	64.8	68.3	54.5	59..2

DNA-2: 5' AAGGAAGAAG 3', **RNA-2**: 5' AAGGAAGAAG **DNA-6**: 5' AAGG**T**AGAAG 3' **RNA-6**: 5' AAGGA**C**GAAG 3'

Table 4: Mixed purine–pyrimidine PNA sequences and their binding studies with complementary DNA/ RNA, mismatch DNA/ RNA.

Entry	Sequence	DNA-3	RNA-3	*p*DNA- 4	DNA-7	RNA-7
PNA-13	H-TTACCTCAGT-Lys-NH_2	54.2	54.5	53.9	47.5	46.5
PNA-14	H-TTACCt^RCAGT-Lys-NH_2 R= H	57.6	52.5	55.4	51.2	38.8
	H-TTACCt^RCAGt^R-Lys-NH_2					
PNA-15	R= H	63.2	48.1	59.3	50.8	34.5
PNA-16	R= NH_2	66.9	52.1	59.5	58.1	43.8
PNA-17	R= NH-C(= NH) NH_2	68.9	54.2	59.9	60.0	46.1

DNA-3:5'ACTGAGGTAA3', ***p*DNA-4**:5'AATGGAGTCA3', **RNA-3**: 5' UGUAACUGAGGUAAGAGG3', **DNA-6**: 5' ACTGTGG**T**AA3' **RNA-6**: 5' UGUAACUGCGGUAAGAGG3'

Chapter 3

Section 3.1: Design and Synthesis of *iso*-TANA dimers, their incorporation into DNA oligomers and biophysical studies

Modified backbone nucleic acids are gaining importance for their applications in chemistry, biology and medicine. Natural nucleic acids containing phosphate backbone cannot be used as drug due to their susceptibility to enzymatic hydrolysis. Therefore oligonucleotides (ONs) with different backbones were designed. For designed modified ONs to be effective leads as drugs they need to have enhanced strength of hybridization with target RNA, good cellular permeability and stability towards degradative enzymes. Modified ONs containing amide backbone instead of phosphate backbone showed better binding towards complementary RNA than DNA.

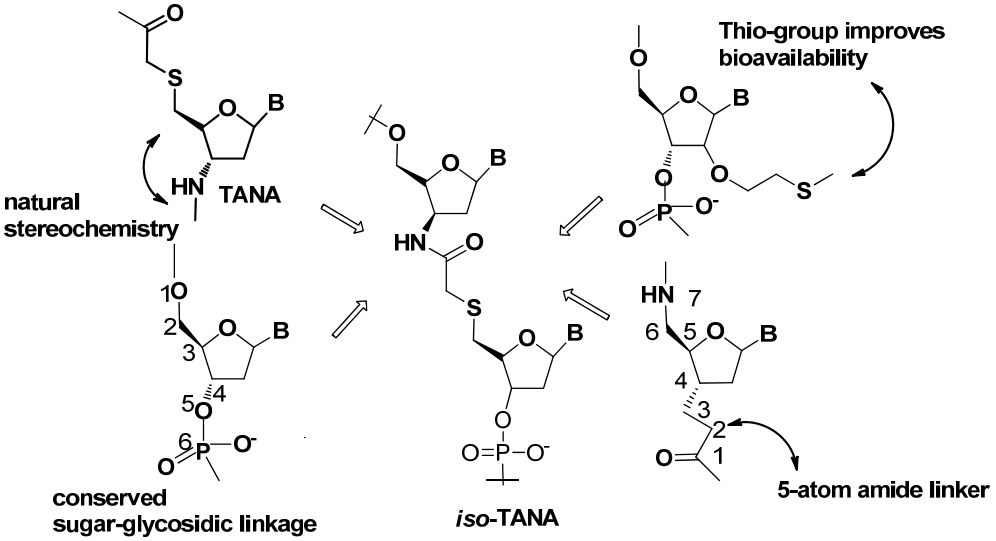


Figure 3: Design of *iso*-TANA

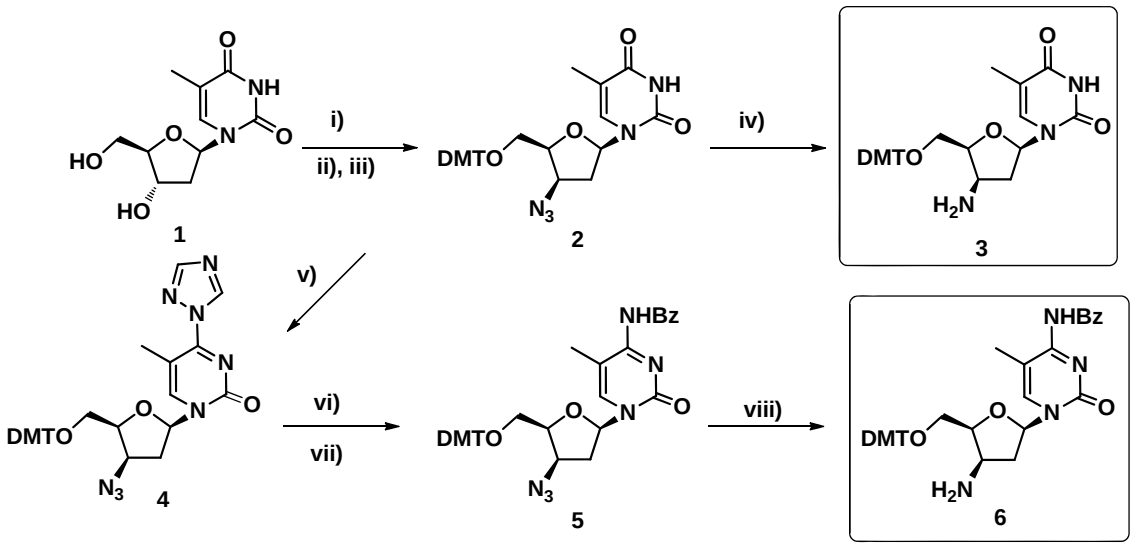
Thioacteamido linked oligonucleotides (TANA) which were reported form our laboratory showed very good binding selectivity towards complementary RNA over DNA. Sugar puckering is known to have an effect on base stacking interactions in oligonucleotides. In DNA and RNA majority of the sugars are in S and N form respectively. The 3'substituent in ribose/2'-deoxy ribose sugar has a major impact on the conformational preference N/S of sugars. In this chapter *iso*-TANA which is an isomer of TANA containing the 3' amino part in unnatural stereochemistry is studied. *iso*-TANA also contains one extra carbon atom in the backbone in order to compensate the shorter amide bond in comparison to the phosphate bond. Literature also reports that backbone -extended PNA and DNA analogues favour the RNA -binding over DNA.

In the design of *iso*-TANA the sugar glycosidic linkage has been conserved. Also presence of thio group could improve the bioavailability. This chapter has been divided into two Sections. In Section 3A, is described synthesis of three dimers viz. thymidine-thymidine, 2'-deoxy-5-methyl cytidine-thymidine-and 2'-deoxy-5-methyl cytidine-2'-deoxycytidine, their incorporation into DNA oligomers and biophysical studies of their complexes with DNA/ RNA. Section 3.2 consists of synthesis of thymine and 5-methyl cytosine monomers and solid phase synthesis of homo *iso*-TANA and chimeric PNA-*iso*-TANA oligomers. Biophysical studies of complexes of these oligomers with complementary DNA/RNA are also described.

*i*so-thioacetamido linked thymidine-thymidine, 2'-deoxy-5-methyl cytidine thymidine and 2'-deoxy-5-methyl cytidine-2'-deoxycytidine dimers were synthesised by solution phase coupling of the 5'-thio acid part and 3'-*xylo* configured amino modified nucleosides.

1) Synthesis of xylo configured amino nucleosides

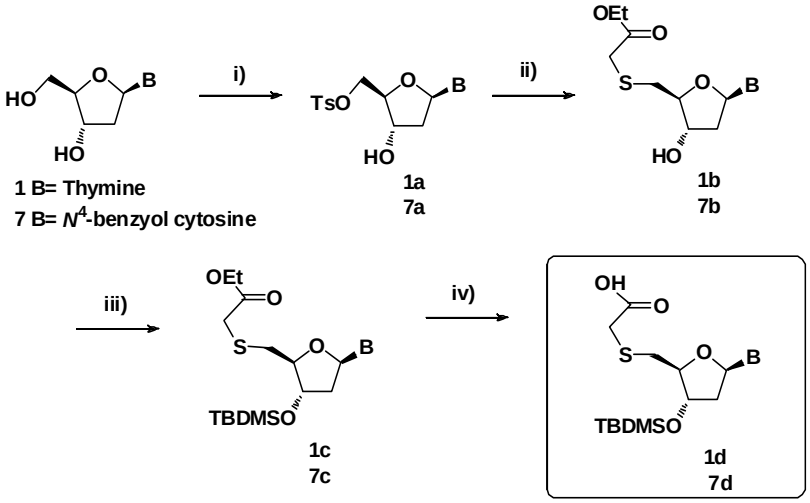
3'-amino *xylo* configured modified nucleoside were synthesized by reduction of corresponding azido derivative of nucleosides. The inversion at 3' was achieved by nucleophilic substitution of 3'-O-mesyl with sodium azide. The 3'-deoxy*xylo*-3'-azido thymidine was converted to the 2', 3'-dideoxyxylo-3'-azido 5-methyl cytidine derivative. The exocylic amino group of 5-methyl cytosine was protected with benzyol and which was subsequently reduce to give amino derivative.



Scheme 3.1: Reagents and conditions , i) DMTrCl/pyridine ,87%, ii) MsCl/ Et_3N, DCM, 85%, iii) NaN_3 /DMF, 60^oC, 45 %, iv) H_2/ Pd-C, MeOH, 86%,v) 1,2,4-triazole, $POCl_3$, Et_3N, MeCN, 88%, vi) conc. NH_3, 1-4 dioxane, 90%, vii) BzCl, pyridine, 75%, viii) H_2S/ pyridine, Et_3N, 78%

2) Synthesis of 5'-mercaptoacetic acid derivatives of thymidine and 2'-deoxy cytidine

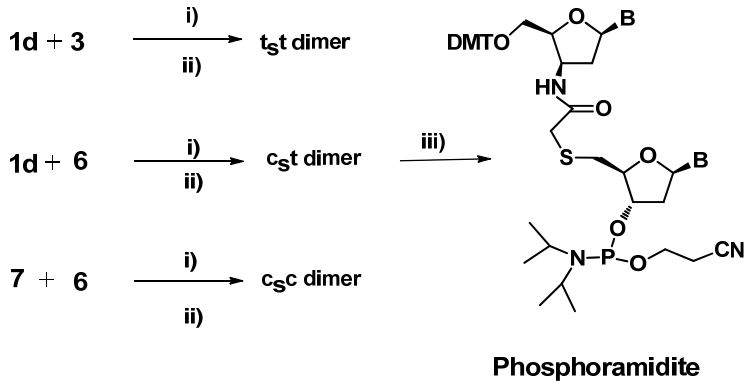
The mercapto acetic derivatives of thymidine and 2'-deoxy cytidine were synthesized from thymidine and 2'-deoxy cytidine. The synthetic scheme is given in scheme-3.2



Scheme-3.2 reagents and conditions: i) TsCl/ pyridine, 60%, ii) ethyl mercapto acetic acid/NaH/DMF, 80%, iii) TBSCl/imidazole/DMF, 90%, iv) B = Thymine 1N LiOH/MeOH, 88%, B= N^4-benzyol cytosine 1N NaOH/ THF, 0°C, 75%

3) Formation of amide linked dithymidine (*iso*-t_St), 2'-deoxy-5-methylcytidine-thymidine (*iso*-c_St) and 2'-deoxy-5-methyl cytidine-2'-deoxy cytidine (*iso*-c_Sc) dimers and their phospharamidite derivatives

The amino nucleosides from scheme-3.1 and acid part from scheme3.2 were coupled by standard solution phase coupling. After removal of protecting groups, these dimers were converted into their phospharamidite derivatives.



Scheme-3.3 reagents and conditions: i) HBTU/ HOBt, DIPEA, ACN:DMF (5:1), ii) 1N TBAF in THF, iii) *N*,*N'*-diisipropylcyanoethyl chlorophosphine, DIPEA, DCM

NMR and CD studies of *iso*-TANA dimers

In the present Section the studies towards finding out the conformational preference of the deoxyribose ring in 3'-acetamido *5'-O*-TBS-2'-3'dideoxy-xylofuranosyl nucleosides and *iso*-TANA dimers are determined by NMR, and the results are compared with the corresponding 3'-acetamido *5-O*-TBS-2'-3'dideoxy ribofuranosyl nucleosides and and TANA dimers . 3'-acetamido *5'-O*-TBS-2'-3'dideoxy-xylofuranosyl nucleosides were synthesized from intermediates described previously in the earlier section. The %S character for sugar ring was calculated from empirical formula

% S = (Σ H1'-9.8)/5.9x100 where ΣH1' = $J_{1'2'}$+ $J_{1'2''}$

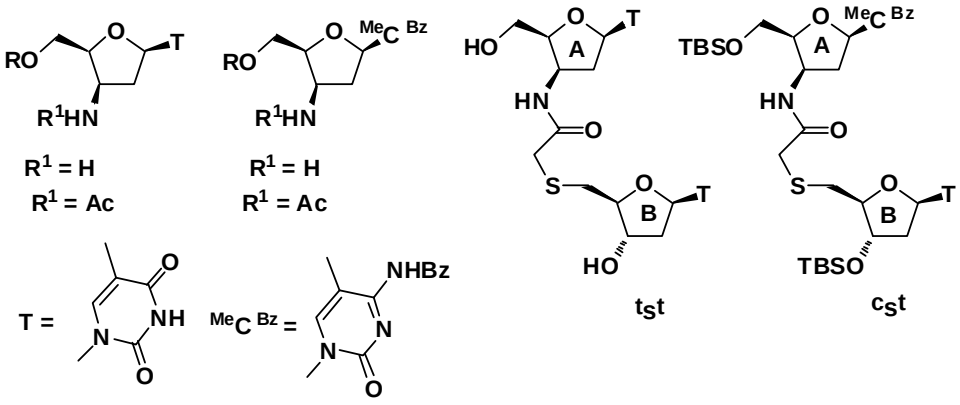


Table 1: Conformational analysis of iso-TANA monomers and dimers

Xylo	3'-OH T	3'-NH_2 T	3'-NHAcT	t_st	
				A	B
%S	16	26.7	67.1	57.8	65.9
xylo	3'-OH C	3'-NH_2C	3'-NHAcC	c_st	
				A	B
%S	≤ 5	34.4	62.4	46.4	63.7

CD analysis

To study the change in the backbone sugar configurations in TANA (t_St/c_St) and *iso*-TANA dimers on the base stacking interactions CD analysis was carried out. CD analysis showed that different base-stacking interactions were present due to change in the stereochemistry of at 3'-amino part of thioacetamido linkage. CDof *iso*-TANA t_St dimer resembles the CD for phosphate linked xylo dithymidine dimer.

Synthesis, purification and mass spectral analysis of the chimeric *iso*TANA-DNA oligomers and their binding studies with complementary DNA and RNA sequences

The phosphoramidite derivatives of *iso*-TANA dimers obtained in scheme-3.3 were incorporated into DNA oligomers by automated synthesis on DNA synthesizer

Following are the ON sequence synthesized containing various *iso*-TANA dimers.

Table 2: ON sequence synthesized containing various *iso*-TANA dimers

Entry	ON Sequence	HPLC t_R (min)[a]	[a]Mass calc/observed
1	CGT Tt_St TTT TGC **ON-1**	7.4	3606.54/ 3604.85
2	CGT Tt_St TTt_StGC **ON-2**	7.7	3596.6/ 3595.28
3	CGt_St t_St t_St t_St GC **ON-3**	9.0	3576.8/ 3575.54
4	TCT C t_St TCT T **ON-4**	8.1	2928.03/ 2929.29
5	5' TCT C t_St Tc_StT 3' **ON-5**	8.9	2935.08/ 2937.54
6	5' Tc_St CTT TCT T 3' **ON-6**	8.2	2942.05/ 2943.96
7	5' Tc_St CTT Tc_St T 3' **ON-7**	8.4	2949.10/ 2950.44
8	5' Tc_St c_StT Tc_St T 3' **ON-8**	9.0	2956.15/ 2955.49
9	5' TCAc_StAGATG 3' **ON-9**	8.8	3036.12/3036.55
10	5'CCT Ct_St ACC TCA GTT ACA 3' **ON-10**	8.2	5366.7/ 5370.44
11	5'CCT Ct_St ACC TCA Gt_St ACA 3' **ON-11**	8.6	5356.90/ 5353.93
12	CCT CTTAc_ScTCAGTTACA 3' **ON-12**	9.1 9.3[b]	5376.65/5376.85 5376.65/5382.82[b]
13	GAA GGG c_StT t_StG AACT c_StT **ON-13**	9.3	5841.03/5837.80
14	GAA GGG c_StT t_StG AAc_St c_StT **ON-14**	9.6	5849.08/ 5849.36

a HPLC and MALDI-TOF mass analysis of sequence containing *iso*-TANA unit

b HPLC and MALDI-TOF mass analysis of sequence containing TANA unit

UV-melting studies of *iso*-TANA-modified oligomer complexes with complementary DNA/RNA.

The synthesized sequences were tested for binding studies with complementary DNA/ RNA. Binding strengths of TANA and *iso*-TANA dimers modified oligomers with complementary DNA/ RNA have been compared.

Table 3: T_m (°C) values of *iso*-TANA ONs: DNA/RNA complexes

Entry	ON Sequences(5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C (TANA)	
		DNA-1	RNA-1	DNA-1	RNA-1
	CGT TTT TTT TGC	32.5	26.7	40.0	32.0
ON-1	CGT Tt_st TTT TGC	21.8	nd	23.7	32.3
ON-2	CGT Tt_st Tt_st TGC	nd	25.0	nd	50.0
ON-3	CGtst t_st t_st t_stGC	nd	29.8	nd	47.8

DNA-1 = 5' GCAAAAAAACG 3'; **RNA-1** = 5' GCAAAAAAACG 3'

Table 4: T_m (°C) values of *iso*-TANA ONs: DNA/RNA complexes

Entry	ON Sequences(5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C (TANA)	
		DNA-2	RNA-2	DNA-2	RNA-2
	CCTCTTACCTCA GTTACA	54.6	53.8	54.6	54.7
ON-10	CCTCt_st ACCTCAGTTACA	49.3	46.7	39.6	47.5
ON-11	CCTCt_stACCTCAGt_stACA	42.6	40.9	43.5	52.8
ON-12	CCTCTTAc_scTCAGTTACA	51.6	51.5	52.2	52.5

DNA-2 = 5' TGTAACTGAGGTAAGAGG 3'; **RNA-2** = 5' UGUAACUGAGGUAAGAGG 3'

Table 5: T_m (°C) values of *iso*-TANA ONs: DNA/RNA complexes

Entry	ON Sequences(5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C (TANA)	
		DNA-3	RNA-3	DNA-3	RNA-3
	GAA GGG CTT TTG AAC TCTT	53.0	54.2	53.6	54.8
ON-13	GAA GGGc_stT t_stGAAC Tc_stT	44.7	39.7	42.7	54.5
ON-14	GAA GGGc_stT t_stG AAc_st c_st T	45.8	34.3	41.6	43.6

DNA-3 = 5' AAGAGTTCAAAAGCCCTTC 3'; **RNA-3** = 5' AAGAGUUCAAAAGCCCUUC 3'

Section 3.2: Synthesis of Modified PNA sequences with *iso*-thioacetamidonucleic acids monomers (*iso*-TANA) and their Biophysical studies

In this Section we have synthesized *iso*-TANA thymine monomer and it was incorporated into aeg PNA backbone, also its homoligomer also has been synthesized. The biophysical studies of these sequences wee carried with DNA RNA.

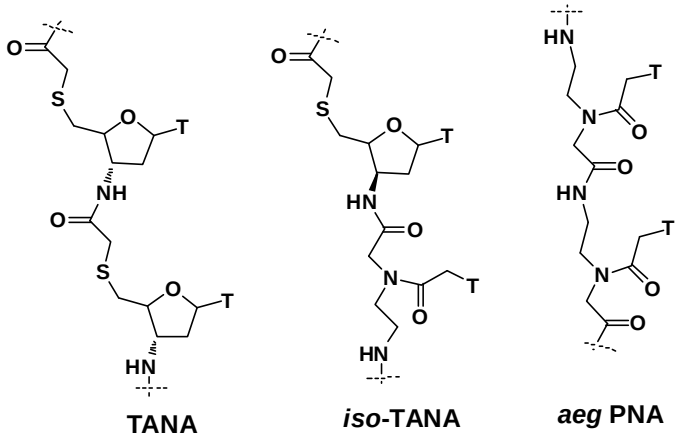
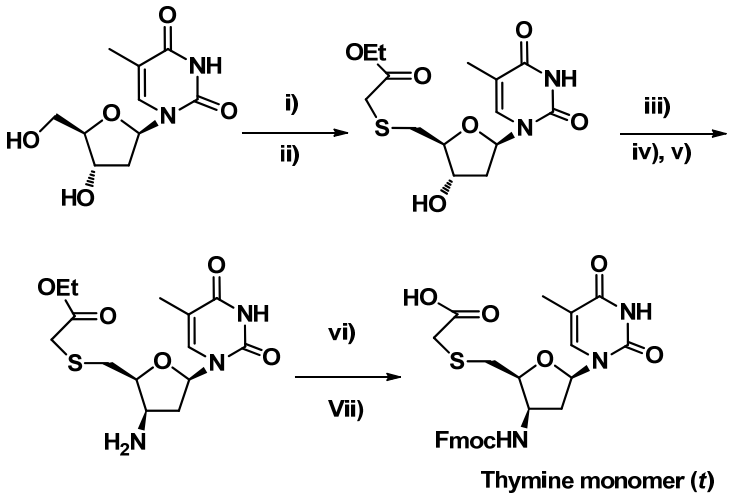


Figure 1: Structure of TANA, PNA and *iso*-TANA

The synthesis of thymine monomer thymine monomer was carried out as per the scheme-3.2.1



Scheme 3.2.1 : Reagents and conditions) TsCl/Py ,60%, ii) ethyl mercapto acetic acid/ NaH-/ DMF, 87%, iii) MsCl/ Et_3N-DCM, 90%,iv) NaN_3/ DMF, 60 °C, 47%, v) H_2S/ Pyridine, TEA 82%, vi) 1N LiOH / MeOH, 78%, vii) Fmoc -succinimide, $NaHCO_3$ in 1,4-Dioxane/ water, 64 %.

Following are the sequences synthesized by standard Fmoc solid phase peptide synthesis strategy.

Entry	Sequence	HPLC t_R (min)	Mass calc/obsd.
T-1	H-*ttttttt*-βAlanine	13.1	2467.6/2488.91 ($M+Na^+$)
T-2	H-TTTT*t*TTT-βAlanine	9.7	2249.4/2278.54($M+6+Na^+$)
T-3	H-TTTTTTT*t*-βAlanine	10.0	2249.4/2274.91($M+1+Na^+$)
T-4	H-TTT*t*TTT*t*-βAlanine	10.6	2280.6/2279.35 ($M+Na^+$)
T-5	H-T*t*T*t*T*t*Tt-βAlanine	11.1	2343.00/2343.99

UV-melting studies of synthesized oligomers complexes with complementary DNA/RNA

Entry	Sequences	UV *T*m °C	
		DNA	RNA
1	H-TTTTTTTT-βAlanine **aeg T_8**	41.6	41.2
2	H-*tttttttt*-βAlanine **T_1**	nd	37.6
3	H-TTTTtTTT-βAlanine **T_2**	nd	37.2
4	H-TTTTTTTt-βAlanine **T_3**	nd	34.5
5	H-TTTtTTTt-βAlanine **T_4**	nd	37.3

Chapter 1
Introduction

Chapter 1

1.1 Introduction to nucleic acids

1.1.1 Primary structures of DNA and RNA

DNA and RNA are the macromolecules composed of chains of monomeric nucleotides (Figure 1). These molecules are the universal carriers of the genetic information in biological systems. Each nucleotide is comprised of three components: a nitrogenous hetrocyclic base, which is either a purine or pyrimidine; a pentose sugar and a phosphate group. DNA and RNA differ in the structure of the sugar in their nucleotide-DNA contains deoxyribose, while RNA is made up of ribose sugars. The nitrogenous bases found in the two nucleic acids are different: adenine, cytosine, and guanine are found in both RNA and DNA, thymine occurs only in DNA while uracil only occurs in RNA (Figure 2). Nucleic acids are usually either single-stranded or double-stranded, though structures with three or more strands are also known. Watson and Crick discovered the double helical structure of DNA in 1953.[1] (Figure 1).

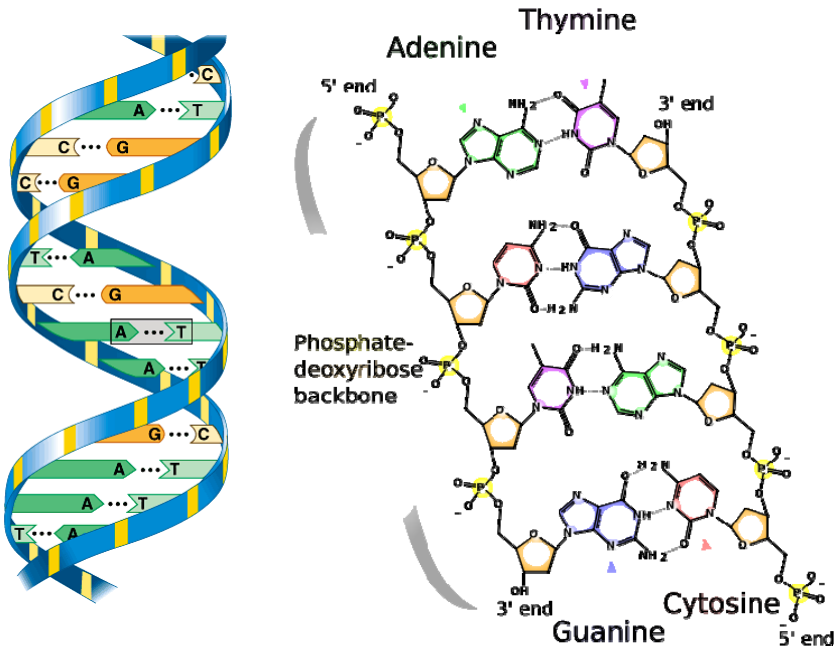


Figure 1: Double helical structure of DNA

In contrast, RNA is usually single stranded, but any given strand may fold back upon itself to form secondary structures as in tRNA and rRNA. Within cells, DNA is usually double stranded, though some viruses have single stranded DNA as their genetic

material. Retroviruses have single-stranded RNA as their genetic material. The sugars and phosphates in nucleic acids are connected to each other in a chain, linked by shared oxygens, forming a phosphodiester bond. In conventional nomenclature, the carbons to which the phosphate groups attach, are 3'-end and the 5'-end carbons of the sugar. Bases are attached through *N*-1 of pyrimidines and *N*-9 of purines to the 1'- carbon of ribose through β-glycosyl bond.

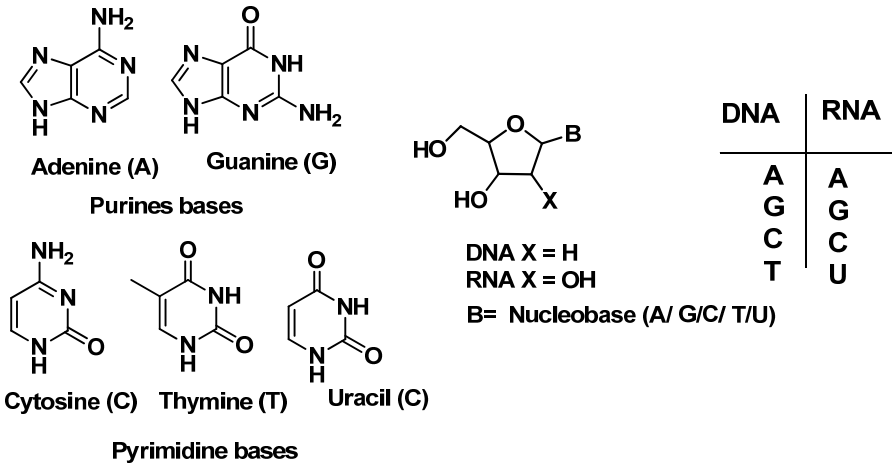


Figure 2: Structures of nucleoside and nucleobases of DNA and RNA

1.1.2 Sugar puckering in nucleosides

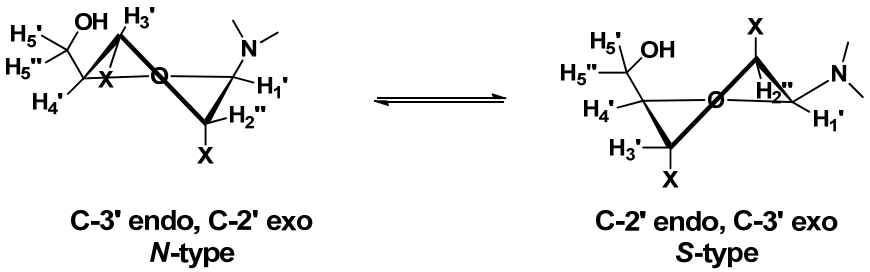


Figure 3: Puckering of sugar ring

In natural nucleic acids, pentose sugars are puckered or twisted to give preferred helical conformations. These pentose sugar moieties are puckered in order to minimize the non-bonded interactions between their substituents. This '*puckering*' is described by identifying the major displacement of the carbons *C*-2' and *C*-3' from the median plane of *C*1'-*O*4'-*C*4'. Thus, if the *endo*-displacement of *C*-2' is greater than the *exo*-displacement of C-3', the conformation is called *C*2'-*endo* and so on for other atoms of the ring (Figure 3). The *endo*-face of the furanose is on the same side as *C*5' and the base; the *exo*-face is on the opposite face to the base. Sugar rings in natural nucleic

acids are found to be in equilibrium between these two forms (Figure 3) *viz*. N- type (C-3'-endo) and S-type (C2'-endo). The 2'-endo (S-type) sugar conformation is preferred in DNA whereas the 3'-endo (N-type) form is preferred in RNA due to the presence of 2'-hydroxy group in ribose sugar. In DNA, the 3'- substituent of the sugar ring and anomeric effect of nucleobase affect the sugar puckering in the nucleoside.[2]

1.1.3 Base Pairing *via* Hydrogen bonding

The N-H groups of the nucleobases are potent hydrogen bond donors, while the sp^2-hybridized electron pairs on the oxygens of the base C=O groups and that on the ring nitrogens are hydrogen bond acceptors. In Watson-Crick pairing, there are two hydrogen bonds in an A:T base pair and three in a C:G base pair (Figure 4).[1]

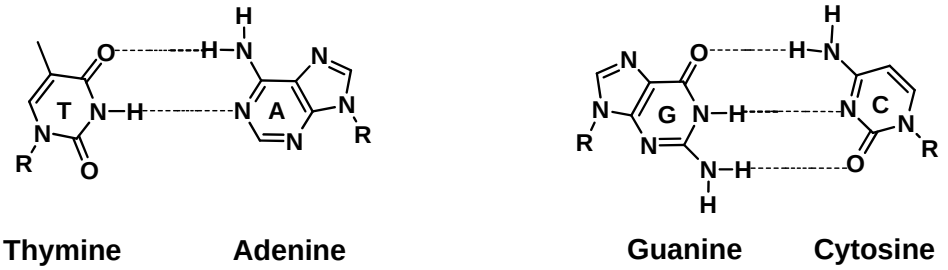


Figure 4: Watson and Crick hydrogen- bonding scheme for A: T and G: C base pair

While Watson-Crick base pairing is the dominant pattern between the nucleobases, other significant pairings are Hoogsteen (HG)[3] and Wobble base pairs.[4]

A Hoogsteen A:T base pair (Figure 5) applies the *N*-7 position of the purine base (as a hydrogen bond acceptor) and C6 amino group (as a donor), which bind the Watson-Crick (*N*3-*O*4) face of the pyrimidine base. Hoogsteen pairs have quite different properties from Watson-Crick base pairs. The angle between the two glycosydic bonds (ca. 80° in the A:T pair) is larger and the C1′–C1′ distance (ca. 8.6 Å) is smaller than in the regular geometry. In some cases, called reversed Hoogsteen base pairs, one base is rotated 180° with respect to the other.

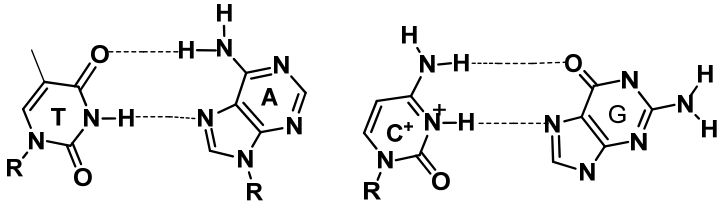


Figure 5: Hoogsteen hydrogen- bonding scheme for A: T and G: C base pair

Hoogsteen base-pairing allows sequence specific binding of the pyrimidine third strands in the major groove of Watson-Crick purine:pyrimidine duplexes to form triple-helical structures (poly (dA):2poly (dT)) and (poly (rG):2poly (rC)). The triplexes are present in three-dimensional structures of transfer RNA.

In the wobble base pairing (Figure 6), a single purine base is able to recognize pyrimidines (e.g. G:U, where U = uracil) and have importance in the interaction of messenger RNA (*m*-RNA) with transfer RNA (*t*-RNA) on the ribosome during protein synthesis (codon-anticodon interactions). Several mismatched base pairs and anomalous hydrogen bonding patterns have been seen in X-ray studies of synthetic oligodeoxynucleotides.[5]

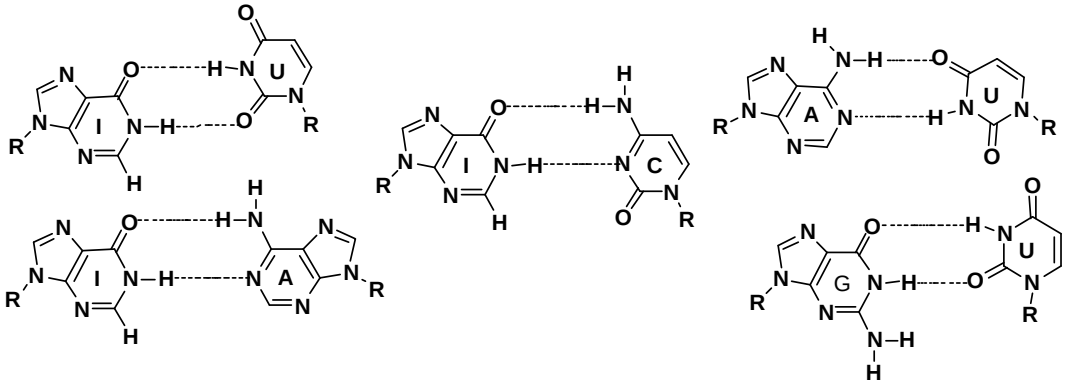


Figure 6: Wobble base pair for Ionosine and Uracil

1.2 Antisense Technologies

1.2.1 Introduction

Drugs consisting of oligonucleotide analogues capable of recognizing RNA through a Watson-Crick base-pairing mechanism, thereby arresting cellular processes at the transcription level are known as antisense agents.[6]

The potential of oligodeoxynucleotides to act as antisense agents that inhibit viral replication in cell culture was discovered by Zamecnik and Stephenson in 1978.[7] Efficient methods for gene silencing have been receiving increased attention in the era of functional genomics, since sequence analysis of the human genome and the genomes of several model organisms revealed numerous genes, whose function is not yet known. As Bennett and Cowsert pointed out in their review article,[8] AS-ONs combine many desired properties such as broad applicability, direct utilization of sequence information,

rapid development at low costs, high probability of success and high specificity compared to alternative technologies for gene functionalization and target validation.

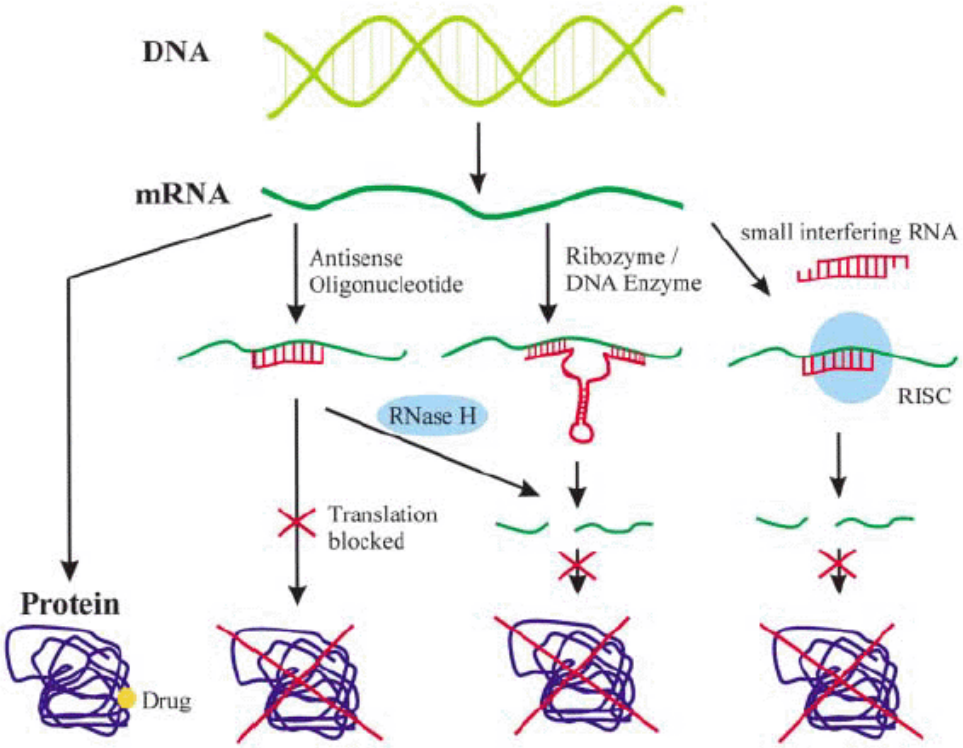


Figure 7: Different antisense strategies[8b]

The wide variety of ONs has been designed to modulate RNAs through a diverse set of post-binding mechanism (Figure 7). These mechanisms can be broadly categorized as a) those involve binding to the RNA and interference with its function without promoting RNA degradation (e.g. inhibition of translation initiation, disruption of RNA structure Figure 8) and b) those that promote degradation of the RNA either through endogenous enzymes, such as RNase H, or Argonaute 2 (RNA interference), or cleavage mechanism designed into the oligonucleotides (Figure 8).[9]

1.2.2 Non-RNA-degrading mechanisms

Blocking translation of mRNA by oligonucleotides is the prototypical antisense mechanism and is often referred to as translation or hybridization arrest. Such ONs are usually designed to bind to, or are adjacent to, the translation initiation region of an mRNA. Although evidence suggest that ONs utilize such mechanisms in the cell free assays, there is limited evidence that oligonucleotides designed to work through translation arrest mechanism actually do so in cell culture or *in vivo*.[9]

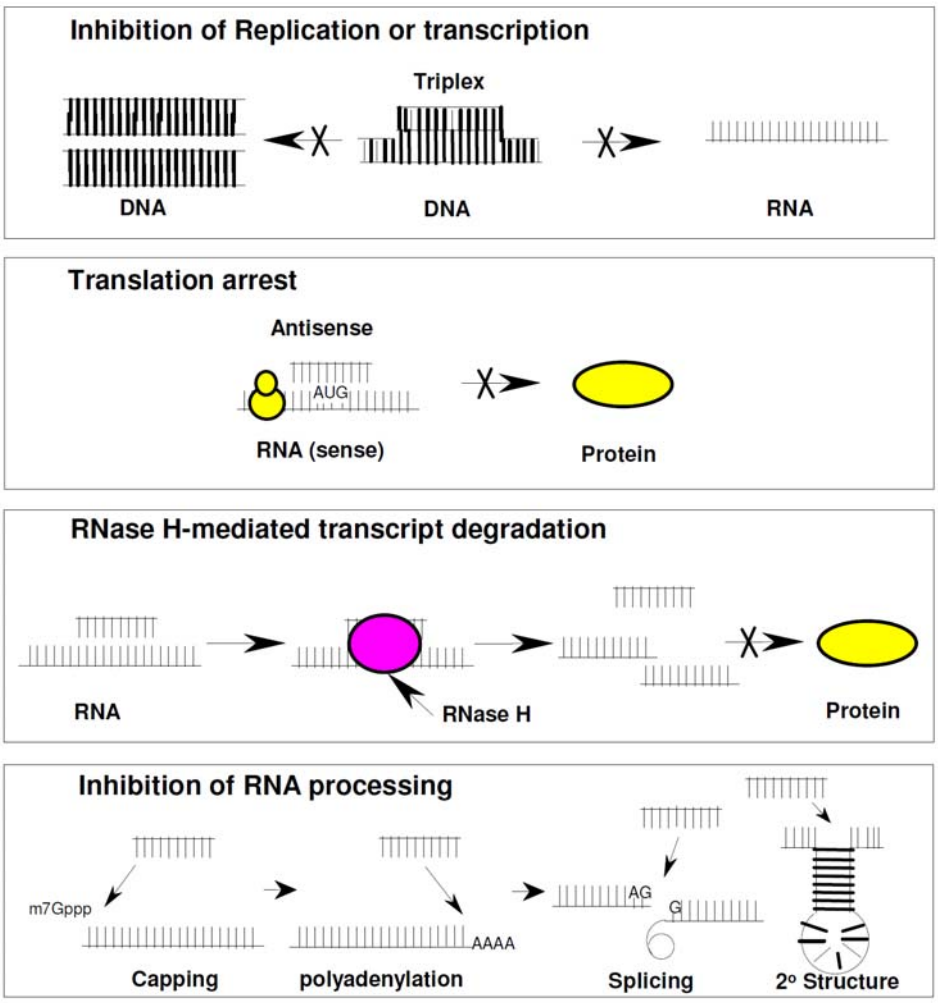


Fig 8: Potential mechanisms of action of antisense oligonucleotides

Splice Correction: Corrective antisense approach

Recent research opens up an exciting possibility for the application of AS-ONs as corrective antisense[10] in the cases of certain diseases caused by genetic mutations. In this approach the AS-ONs can help restoring the viable protein production by acting at pre-mRNA levels for splice corrections[10] (Figure 9) or to yield mRNA that is translated into viable proteins.

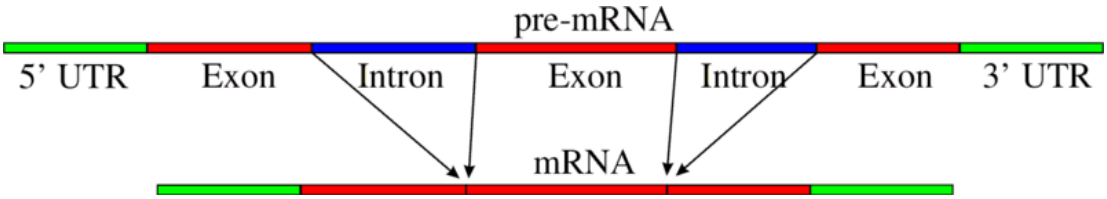


Figure 9: Simple illustration of exons and introns in pre-mRNA.The mature mRNA is formed by splicing. (UTR: Untranslated region)

Precursor mRNA, more commonly termed pre-mRNA, is an incompletely processed single strand of messenger ribonucleic acid (mRNA), synthesized from a DNA template in the nucleus of a cell by transcription. Pre-mRNA includes two different types of segments, exons and introns. Most of exons encode protein, while introns do not and must be excised before translation. This process is called splicing.[11] Spliceosomes, small proteins found in the nucleus and composed of protein and RNA, perform the excision. An exon is any region of DNA within a gene that is transcribed to the final mRNA molecule, rather than being spliced out from the transcribed RNA molecule. Introns are sections of DNA co-linear to the mRNA sequence that will be spliced out after transcription, but before the mRNA is translated.

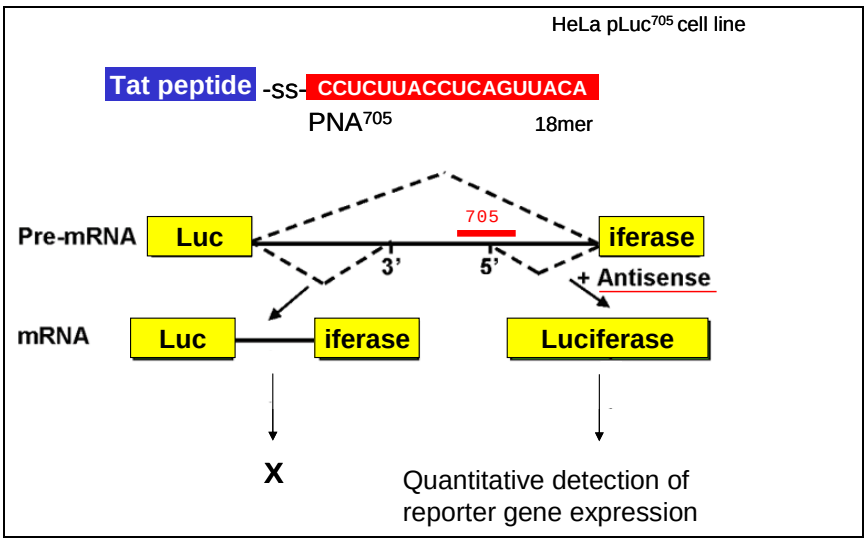


Figure 10: Kole's splice correction assay

Splicing can be experimentally modified so that targeted exons are excluded from mature mRNA transcripts by blocking the access of splice-directing small nuclear ribonucleoprotein particles (snRNPs) to pre-mRNA using oligonucleotides like PNA, LNA or Morpholino antisense oligos.[12] This has become a standard technique in developmental biology. Antisense ONs like PNA, LNA, Morpholino oligos etc can also be targeted to prevent molecules that regulate splicing (e.g. splice enhancers, splice suppressors) from binding to pre-mRNA, altering patterns of splicing (Figure10).[13]

1.2.3 RNA Cleavage mechanism

ONs that work through the RNase H-dependant cleavage mechanism are the best understood class of antisense ONs, accounting for the majority of drugs in the development. RNase H is a family of enzymes present in all mammalian cells that

mediates the cleavage of RNA in an RNA-DNA heteroduplex.[14] RNase H recognises an RNA-DNA heteroduplex, cleaving the RNA strand, resulting in a 5'-phosphate on the product and release of the intact DNA strand (oligonucleotides). The RNase H mechanism has proven to be robust antisense mechanism and is broadly exploited as both a research tool and a potential human therapeutic. ONs that work through an RNA interference mechanism are another example of an antisense mechanism that induces cleavage of the target RNA through endogeneous enzyme. ONs that work through the RNA interference (siRNA) mechanism appear to mimic endogenous small RNAs present in cells, which naturally regulate expression of the targeted gene (Figure 8).

1.3 Modified oligonucleotides

The unmodified antisense oligonucleotides are intended to enter the cell where they can pair with, and so inactivate, the complementary *m*-RNA sequences. But their inability to permeate cell membrane as they carry anionic charge results from repulsive interaction with the cell lipid layer. Further they are degraded by intracellular enzymes such as exo-nucleases. It became very clear that the nuclease susceptibility of natural oligonucleotides based on an unmodified (deoxy)-ribose phosphodiester (PO) backbone, would preclude profound antisense activity in biological systems at concentrations relevant in the context of potential therapeutic applications.[15] This necessitated the chemical modification of oligonucleotides (Figure 11) and represents the key element in development of antisense therapeutics.

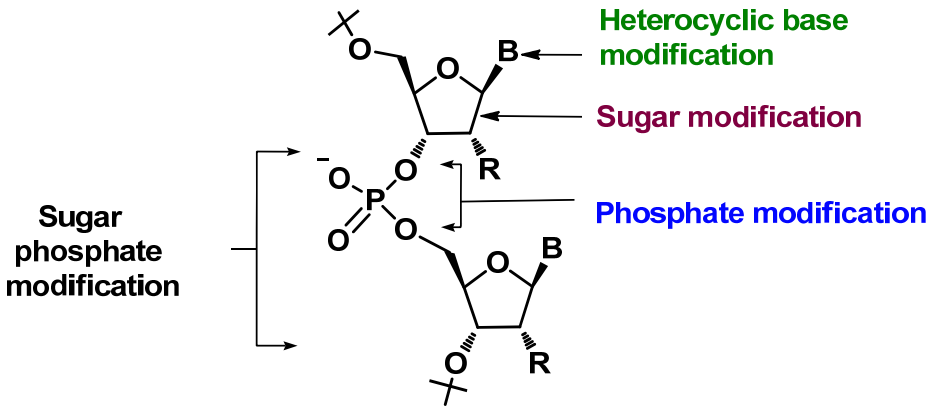


Figure 11: Different sites of modification for DNA

To be broadly useful in antisense approaches, modifications must retain, and preferably enhance, the ability to recognize their target RNAs by Watson-Crick base pairing. This

property is commonly measured by melting temperature of oligonucleotides that contains the modification of interest and is duplxed with complementary strand of RNA. The most widely used antisense drugs utilize the cleavage of the target RNA to achieve their desired effects, using endogenous nucleases such as RNase H or Ago 2.Many studies have shown that modifications that are not substrate for these nuclease are poor antisense drugs , despite being tight binders to their complementary RNA, which is presumably their receptors. It is therefore crucial to examine the effects of modifications on these terminating events and to understand these effects when optimising oligonucleotides designs.

1.3.1 Modifications of phosphate linkage

The first generation modifications focused attention on oligonucleotides having phosphorothioates (PS, Figure 12a),[16] phosphorodithioates, methylphosphonates (Figure 12b)[17] and phosphoramidates (Figure 12c)[18] replacing the anionic phosphate diester linkages and to 2'-OH protection of the ribonucleotides as second-generation antisense oligonucleotides.

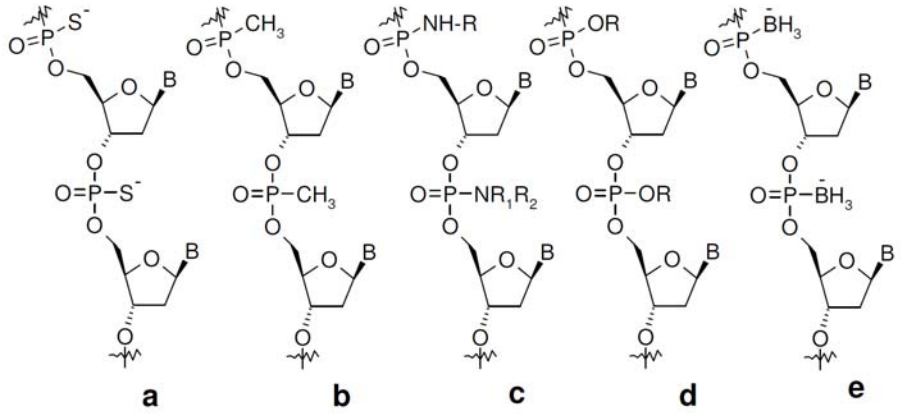


Figure 12: First generation modified antisense oligonucleotide (replacement of 'O' in P-O linkage

Phosphorothioates (PS-oligos): The PS-oligos are oligonucleotides in which one of the non-bridged oxygen was replaced by sulfur and the backbone thus retains the negative charge, but with reduced charge density compared to phosphodiester analogue (Figure 12a). PS-oligos can be easily synthesized on a commercial DNA synthesizer, but the avidity of this analog for the complementary DNA was lower than natural oligonucleotide. The non-stereospecific synthesis of phosphorothioates results in the formation of a mixture of diastereomers and the conformational heterogeneity leads to

lowering of the melting temperature. As a consequence, the search for a stereo controlled synthesis has been undertaken,[19] and synthesized P-chiral analogues of PS-oligos were used for the biological studies of antisense targets inside living cells. PS-oligos are substrates to RNase H and these first generation antisense agents have been extensively tested in human clinical trials against numerous targets. The only antisense agent approved by FDA (Vitravene) so far is based on PS-oligos. However, these oligos have tendency to induce non-specific effects, through binding to cellular and extracellular proteins[20] as well as cleavage of non-target *m*-RNAs that are only partially complementary by RNaseH. Replacement of both non-bridging phosphodiester oxygens leads to non-chiral phosphorodithioates. Compared with the corresponding phosphorothioates, the duplex stability of phosphorodithioate modified oligonucleotides with DNA decreased, and they exhibit less non-specific protein binding and more resistance to nuclease.

Methyl Phosphonates[9] (Figure 12 b): Methyl phosphonates have one of the non bridging oxygen atoms replaced by a methyl group and are neutral in charge. Although these ONs were found to be nuclease resistant, this modification does not support RNase H activity. Additionally, high methylphosphonate content in an oligomers leads to loss of affinity towards its complementary RNA and to poor solubility. Interestingly, removal of the negative charge did not improve the cellular uptake.

Oligonucleotide Phosphoramidates (Figure 12 c): Phosphoramidates are another well-studied class of backbone modifications.[21] They are resistant to hydrolysis by nucleases. With DNA targets, oligonucleotide phosphoramidates (Figure 11c, R_1 = H, R_2 = -CH_3, -CH_2CH_2OMe, $(CH_2CH_2)_2O$) exhibit rather poor hybridization characteristics (ΔT_m = -0.1 to –2.3ºC, at pH 7.2).[22] Oligonucleotide phosphoramidates (R_1= H, R_2 = CH_2CH_2N-$(CH_2CH_2)_2O$) form weak duplexes at neutral pH and more stable duplexes under acidic conditions (pH 5.6), due to protonation of the terminal amine. Oligophosphoramidate with (R_1 = H, R_2 = $CH_2CH_2NMe_2$), is however reported to hybridize to DNA targets at both neutral and acidic pH. Hybridization of oligonucleotides consisting of alternating phosphoramidite/phosphodiester linkage to RNA targets is similar to the corresponding interaction with wild type oligonucleotides.[23] The T_m of these complexes turned out to be independent of the ionic strength of the buffer, in contrast to the behaviour of unmodified oligonucleotides with

DNA or RNA targets, which shows a strong T_m increase, with increasing salt concentration.

Oligo-phosphotriesters (Figure 12d)**:** Even though phoshotriesters are common intermediates in oligonucleotide synthesis, little data about their hybridization behaviour are available. It has been shown that ethyl phosphotriester (Figure 11d, R = -CH_2CH_3) modified oligonucleotides form substantially less stable self-complementary duplexes compared to wild type.[24] A single modification in a self complementary sequence (which therefore had a backbone modification in each strand of the duplex) resulted in a T_m decrease of –4ºC and –11ºC, depending on whether the Sp or the Rp diastereomer was used.

Oligo-boranophosphonates (Figure 12e): These are derived by replacing one of the non-bridging oxygen atoms in the phosphodiester group of DNA with borane (BH_3).[25] The boron phosphate diester is isoelectronic with phosphodiesters, isosteric with methylphosphonate group and is chiral. These negatively charged oligos are highly water soluble, but more liphophilic than DNA. NMR and CD studies show that replacing the phosphodiester linkage in dinucleotides with the boranophosphodiester results in only a slight change in configurational characteristics, such as sugar pucker, acyclic torsional angles and base stacking.[26] Boranophosphate DNA is considerably more stable to various nuclease enzymes than native DNA and overall more stable than phosphorothioate DNA. The discovery that oligonucleotides boranophosphodiesters can activate *E.coli* RNase H and induce cleavage of RNA is encouraging.[27]

1.3.2 N3'→P5' Phosphoramidate Nucleic Acids

The substitution of the internucleoside phosphodiester for the N3'→P5' phosphoramidates (Figure 13 I) was initially thought as an alternative to biologically stable phosphorothioate nucleic acid analogues.[28] Unlike phosphorothioates, the N3'→P5' phosphoramidates did not give rise to the stereoisomers at phosphorus that complicated the binding patterns with complementary nucleic acids. The N3'→P5' phosphoramidate internucleoside linkage dramatically changed the oligonucleotides hybridization properties. The uniformly N3'→P5' phosphoramidate-modified ONs, either with polypyrimidine or mixed base sequence, were found to bind very strongly to complementary RNA sequences. The stability of their complexes with DNA also was better but was much less compared to RNA.[29] The duplex formed with the RNA target

by the N3'→P5' phosphoramidates was far more stable than the one formed by even the native RNA oligomer.

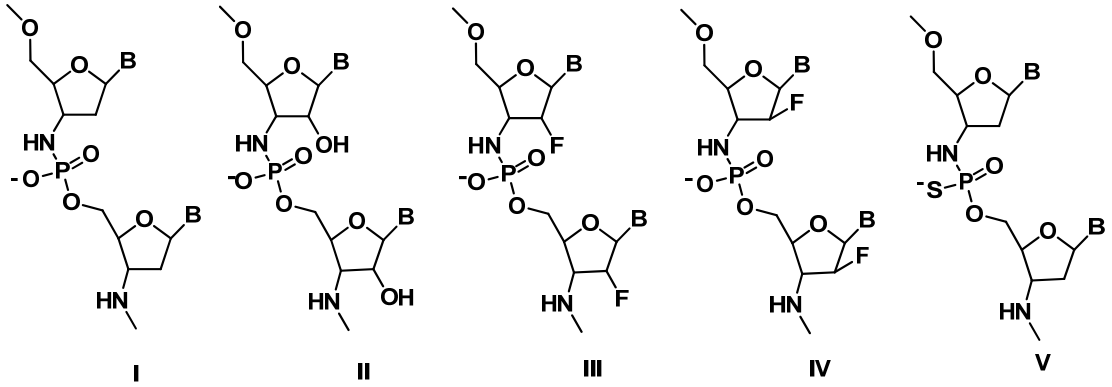


Figure 13: N3'→P5' phosphoramidate internucleoside

The difference in thermal stability between matched and single mismatched sequences was the same for N3'→P5' phosphoramidates as in the case of native phosphodiester linkages.

The NMR studies of the monomeric 3'- amino nucleosides indicated increased H3'-C3'-C4'-H4' dihedral angle, characteristic of C3'-endo sugar ring conformation, and consequently an A-form duplex structure. The possibility of an additional interstrand hydrogen bond mediated by a water molecule, between 3'-amino and 2'-hydroxy or phosphate groups through the narrow minor groove of the A-form N3'→P5' phosphoramidate ON:RNA duplex, was also envisaged as a possible reason for RNA selectivity. The 2'-ribo-fluoro substitution in phosphodiester and phosphorothioate ONs (Figure 13 III) stabilized the duplex with both DNA and RNA sequences due to C3'-*endo* or N-type sugar puckering. In contrast, stability of 2'-arabino-fluoro ON (Figure 13 IV): DNA/RNA duplexes was found to be significantly lower, the introduction of 2'-ribo-fluoro substituent in N3'→P5' phosphoramidates rendered very high stability for the complexes with RNA and the effect of 2'-ribofluoro and 3'-amino was found to be synergistic in stabilizing the 3'-*endo* sugar pucker.[30] Whereas the most stable complexes with RNA were formed by fully modified 2'-arabinofluoro phosphoramidate ONs.

Oligoribophosphoramidates (Figure 13 II)[31] are very resistant to enzymatic hydrolysis by snake venom phosphodiesterase. These compounds form stable duplexes with complementary natural phosphodiester DNA and RNA strands, as well as with 2'-deoxy

N3'→P5' phosphoramidates. Furthermore, the oligopyrimidine riboN3'→P5' phosphoramidate formed an extremely stable triplex with an oligopurine/oligopyrimidine DNA duplex with Δ*Tm* +14.3ºC relative to the 2'-deoxy N3'→P5' phosphoramidate counterpart. The properties of the oligoribonucleotide N3'→P5' phosphoramidates indicate that these compounds can be used as hydrolytically stable structural and functional RNA mimetics. The increase in the thermal stability of duplexes and triplexes of ribophosphoramidates relative to the 2'-deoxyphosphoramidate counterparts, as well as the parent phosphodiester, is due to the further increase in the population of N-type sugar conformations for the 3'-amino-2'-hydroxyl nucleosides, which is determined by a cooperative and additive effect of 3'-amino and 2'-hydroxyl groups on furanose puckering. Additionally, improved hydration of the phosphoramidate duplexes, due to the presence of the 3'-amino group as an additional donor and acceptor of hydrogen bonds, as well as 2'-hydroxyl, may contribute to the increase in thermal stability.

N3'→P5' thio phosphoramidates (Figure 13 V) ONs contain the combination of phosphorothioates and phosphoramidtes.[32] These ONs showed comparable binding with complementary RNA as phosphoramidite.[33] Conjugates of these ONs with lipids showed good cellular uptake and target specificity *in vitro*. *In vivo* theses ONs show good bioavailability and efficient bio-distribution to all organs. A palmitoyl-conjugated thiophosramidate-modified oligonucleotide, GRN163L, has shown promise as an antitelomerase agent.[34]

1.3.3 Locked Nucleic acid (LNA)/ Bridged Nucleic acids (BNA) and their analogues

LNA is generally considered to be an RNA analogue in which the ribose sugar moiety is locked by an oxymethylene bridge connecting the C(2') and C(4') atoms which conformationally restricts LNA monomers into *N*-type sugar puckering.[35] Structural studies by NMR spectroscopy have shown LNA-containing ONs to fit into an A-type duplex geometry.[36] Several unique properties make LNA a promising analogue in the field of nucleic acids research. LNA ONs posses extremely high binding affinity to complementary DNA and RNA ONs as evidenced by thermal denaturation studies, *i.e*. increase in melting temperature (T_m) of +2 to +8 per LNA monomer compared to unmodified duplexes.[37, 35a, b] In addition, LNAs, as the sequence length can be reduced,

also display an improved mismatch discrimination (or improved Watson-Crick base-pairing selectivity) relative to unmodified nucleic acids.

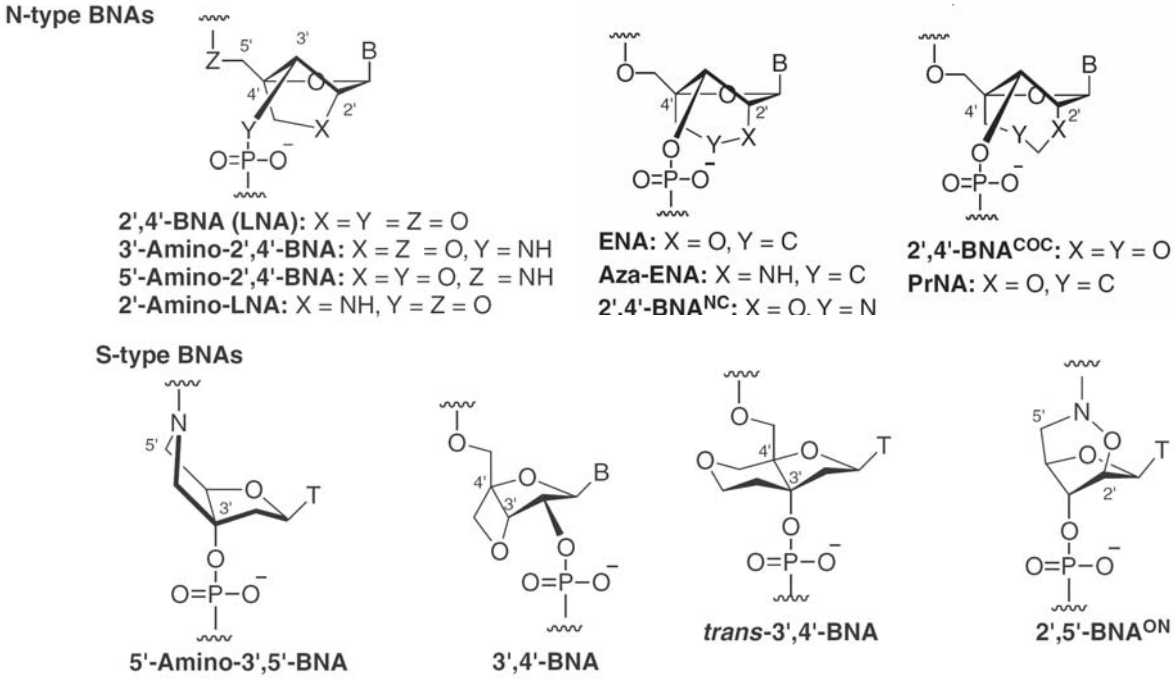
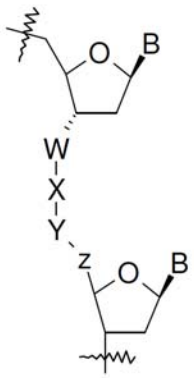


Figure 14: Structures of bridged nucleic acids

Furthermore, LNA-modified ONs show high stability in the biological systems. (*i.e*. resistance to enzymatic degradation).[38] Due to the above properties, various LNA analogues are reported in the literature.[39] 3'-Amino-2'-4'-BNA,[40] 5'-amino-2'-4'-BNA,[41] and 2'-amino-LNA[42] were designed by replacing an oxygen with a nitrogen either in the backbone or in the bridge moiety. These nucleic acids showed hybridizing affinity similar to that of 2'-4'-BNA and the nuclease resistance of 3'-amino-2',4'-BNA was augmented abundantly. Ethylene-bridge nucleic acid (ENA)[43] with six-membered bridged structure was synthesized by koizumi *et al.* Which showed comparable duplex forming ability, and triplex formation and nuclease resistance were somewhat improved compared to LNA.[44] Aza-ENA shows similar to slightly decrease affinity.[45] 2'-aminomethylene-bridged nucleic acids (2'-4'-BNA^{NC}),[46] designed by introducing a fuctionlizable N atom on the bridge, and possesses similar slightly better affinity. Nucleic acids with seven-membered bridged structure such as 2', 4'-BNA^{COC} was also synthesized which showed RNA selective hybridizing profile.[47] Among the S-type nucleic acids: 5'-amino-3', 5'-BNA,[48] 3', 4'-BNA,[49] *trans*-3'-4'-BNA,[50] and 2'5'-BNA^{ON}[51] were developed.

1.3.4 Oligonucleotides with dephosphono backbone

Backbone replacements for the use of antisense oligonucleotides involve elimination of phosphorous atom from the phosphotriester backbone (Figure 14). A common problem for all anionic analogs is the ineffective permeation of cellular membranes. Anionic ONs are taken up by endoscopes, but are unable to cross the endosomal membrane in the absence of cationic lipids. Based on this observation, neutral isosteres of the phosphodiester linkage (Figure 15) have been devised.[52]



	W	X	Y	Z
1	CH_2	SO_2	CH_2	CH_2
2	NH	SO_2	CH_2	CH_2
3	O	SO_2	NH	CH_2
4	O	CH_2	O	CH_2
5	S	CH_2	O	CH_2
6	CH_2	NCH_3	O	CH_2
7	NR	CO	CH_2	CO
8	CH_2	CH_2	NH	CH_2
9	CH_2	CO	NR	CH_2
10	CH_2	NH	CO	CH_2
11	CO	NH	CH_2	CH_2
12	O	CO	O	CH_2
13	O	CO	NH	CH_2
14	O	CH_2	C	C

Figure 15: Non-phosphorus neutral backbone modifications

An increasing number of neutral ODN analogs have been developed that do not contain stereogenic phosphorous centers. Substitution of the PO linkage with neutral SO_2 group resulted in a series of sulfonyl containing linkages. Sulfones 1, sulphonamides 2 and sulfamates 3 have been prepared as ON analogs, but very little *in vitro* and cellular data are available. The formacetal linkage 4 exhibits somewhat inferior sequence-specific binding affinities. Slightly increased RNA binding properties were observed with the 3'-thioformacetal 5 and N-methylhydroxylamine 6 linkages. 5'-thioformacetal containing ODNs do not hybridize as well as the unmodified ODNs to either RNA or DNA. Replacement of the PO backbone by amide groups 7-11 resulted in neutral and achiral linkages that were tested for RNA binding and nuclease stability. Analougs 9 and 11 were identified as having good binding affinity to the RNA targets and high stability towards cellular nucleases. Modifications of the 2'-position into OMe provided ODNs with even greater binding affinity and nuclease stability. Carbonate 12 and carbamate (Figure-15 13) linkages have also been reported as replacements for the PO backbone. The carbonate linkage has been prepared as dimer, but no biochemical

data available. The 5'-N-carbamate linkage 13 is chemically stable, and cytidine hexamers were found to bind complementary DNA and RNA with high affinity, while thymidine hexamer cabamate ODN bound nucleic acid targets with relatively low affinity. Replacement of the PO linkage with an acetylenic bond 14 resulted in depressed RNA affinity. Modifications that do not contain PO linkages as part of the actual backbone structure automatically eliminate the susceptibility to nuclease degradation. They have to be incorporated into oligonucleotides *via* 3'-activated backbone-modified di-nucleotide analogs. The most promising modification that has emerged from the replacement of PO linkage by a neutral carboxylic amide is amide (A, Figure 15) from CIBA group.[53] Modified oligonucleotides with alternating arrangement of PO and amide A linkages exhibit modest increase in RNA binding affinity (0.5ºC/modified dimer unit) over the unmodified wild-type duplexes. Various types of hydroxylamine-based internucleoside linkages have been studied by the ISIS group.[54]

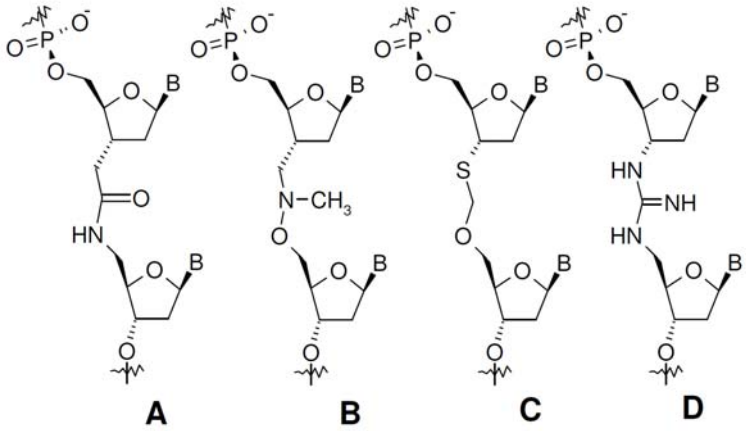


Figure 16: Modification of the internucleoside PO linkage in ONs

The most attractive modification appears to be the methyl (methylimino) (MMI) linkages (B, Figure 16), whose incorporation into oligodeoxynucleotides leads to slight enhanced RNA binding affinity (0.1-0.5ºC/modified dimer unit) and increases nuclease resistance. Another successful approach to backbone modified oligonucleotide analogs with enhanced RNA affinity consists in the replacement of phosphodiester group by 3'-thioformacetal (3'-SCH_2-5') linkages C (Figure 16). An increase in T_m of 0.8ºC/modified dimer unit has been reported with complementary RNA, in comparison to the unmodified wild type duplex. A novel type of DNA analogue, which incorporates internucleoside guanidine groups in place of natural PO linkages, termed deoxyribonucleic guanidine [D, (DNG), Figure 15], has recently been reported.[55] For

fully modified analogue DNG-T_5, whose 1:1 complex with poly rA was reported to exhibit incredible thermal stability (T_m = 100ºC at ionic strength μ= 0.22).

1.3.5 *xylo* nucleic acids (XNA)

2'-deoxy-*xylo* nucleic acids (dXNA) (Figure 17 **X)** form complxes with DNA/RNA complements that substantially differ from the unmodified in structure due to inversion of the configuration at C3'of the sugar ring.[56] The various DNA/RNA modifications containing xylo sugars were reported in the literature.[57] The dXNA monomers predominantly adopt an *N*-type furanose conformation as opposed to the S-type of their 2'-deoxy-ribo counterparts.[56] Notably, oligomers containing dXNA monomers possess higher stability towards exonucleases.[55, 58] The ONs containing single dXNA monomer destabilized duplex with DNA whereas stabilized duplex with RNA as good as unmodified ONs. The increased binding towards RNA has been attributed to the tendency of these nucleotides to adopt a C3'-endo conformation, giving thermally stable A-type duplexes. The introduction of a few dXNA monomers into a DNA strand had a very negative influence on hybridization properties.[59] The corresponding conformationally locked xylo LNA (Figure 17, **L**) showed preferential binding with RNA than DNA.[60]

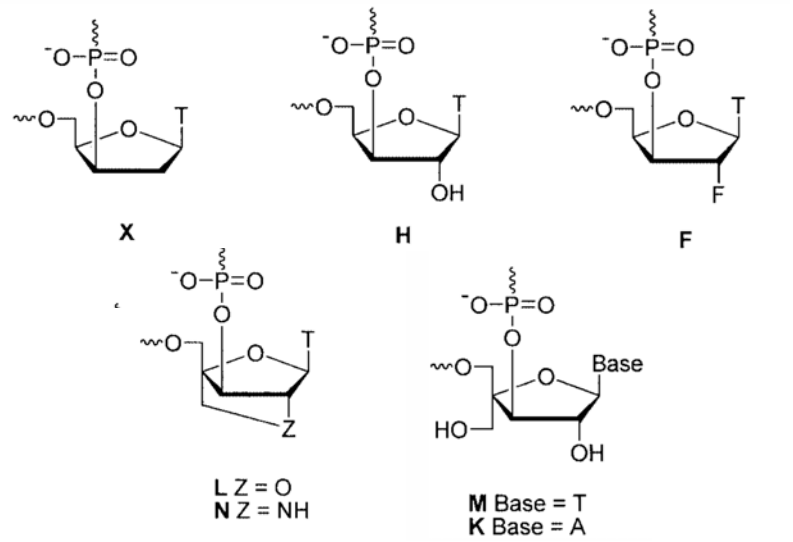


Figure 17: Structures of *xylo* nucleic acids

The change from 2'-deoxy xylosesugar ring to 2'hydroxy (Figure 17, **H**) in xylose nucleic acid increases the stability of the duplexes with RNA. Whereas the

introduction of one 2'-fluro xylonucleosides (Figure 17, **F**) in DNA leads to decreased affinity towards unmodified DNA, while the effect is less detrimental towards RNA.

The 2'-amino –xylo-LNA monomer (Figure 17, **N**) observed to be stabilize the duplex when compared to other sugar –modified XNAs.[57] The incorporation of xylo monomer containing 4'-*C*-hydroxymethyl group (Figure 17, **M**, **K)** in ONs pointing towards the minor groove, has a negative on the duplex stability(with one or three incorporations). The unfavourable steric interactions due to the 4'-C-alkyl branch may be responsible for destabilization.

1.4 Peptide Nucleic Acids

Peptide nucleic acid (PNA), a DNA mimic resulting from the sugar-phosphate backbone replacement by a pseudopeptide backbone, was introduced by Nielsen *et al*. In 1991.[61] The structure of PNA is remarkably simple consisting of repeating *N*-(1-aminoethyl)-glycine units linked by amide bonds (Figure 18). The purine (A, G) and pyrimidines (C,T) bases are attached to the backbone through methylene carbonyl linkages and hence PNA is hybrid of peptides and nucleic acids in a rare structural combination. PNAs do not contain any sugar moieties or phosphate groups. It was therefore a surprise that PNA in many respects mimicked the behavior of DNA, and in some applications demonstrated superior properties. Homopyrimidine PNAs bind to complementary DNA and RNA to form triplexes with slightly increased affinity to RNA over DNA, while PNAs with both pyrimidine and purine bases form very stable duplexes with DNA/RNA, with equal affinity but more than that of either DNA:DNA

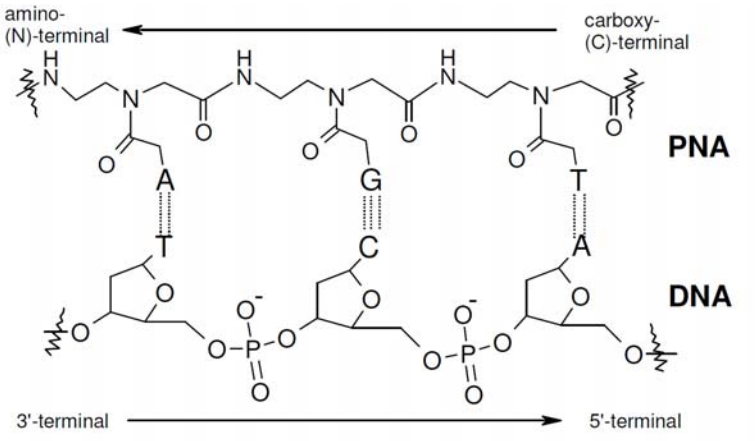


Figure 18: An antiparallel PNA:DNA duplex showing chemical structure of PNA (upper strand) as compared to that of DNA (lower strand)

or DNA:RNA complexes. Interestingly PNA can bind to DNA duplex to form triplex or duplex by displacing one of the DNA strand having same sequence of bases.

By convention, PNAs are depicted like peptides, with the *N*-terminus (corresponding to 5'-end of DNA) at the first (left) position and the *C*-terminus (3'-end of DNA) at the right. PNA is set apart from DNA in that the backbone of PNA is acyclic, achiral and neutral. PNAs can bind to complementary nucleic acids in both antiparallel and parallel orientation, unlike natural DNA duplexes that always prefer antiparallel orientation. However, the antiparallel orientation is strongly preferred, and the parallel duplex has been shown to have different structure.[62] Because of very favourable hybridization properties[63] and high chemical and bio-stability,[64] it was regarded as a very promising lead for developing efficient antisense agents and medical drugs.[65]

1.4.1 PNA-DNA complex formation and their structures

Duplex formation with complementary DNA and RNA:PNAs obey Watson-Crick rules of hybridization with complementary DNA and RNA. Antiparallel PNA-DNA hybrids are considerably more stable than the corresponding DNA-DNA complexes.[61] The increased stability results in an increase in T_m of approximately 1ºC/base. Antiparallel PNA-RNA duplexes are even more stable compared to DNA-RNA hybrids, and PNA-DNA duplexes.[61] The stability of parallel PNA-DNA and PNA-RNA duplexes is almost exactly the same as that of (antiparallel) DNA-DNA and DNA-RNA duplexes respectively. An interesting aspect of PNA-DNA duplex formation is, the T_m decreases with increase in salt concentration (ionic strength) which is contrast to that of DNA-DNA duplex, for which increase in T_m with salt concentration observed.[66] Base pair mismatches result in a reduction of the T_m value of 8-20ºC.[61] This discrimination is, in some cases, approximately double that observed for DNA-DNA duplexes.

Triplex formation: Homopyrimidine PNAs and PNAs with high pyrimidine/purine ratio bind to target DNA normally by formation of unusually stable PNA_2-DNA triplexes. However, in case of C-rich PNAs and GC-rich DNA duplexes, PNA-DNA_2 triplexes are observed. Base pairing mismatches result in a drop in melting temperature of 14-25ºC.[67] The sequence specificity of triplex formation is based on the selectivity of formation of the intermediate PNA-DNA duplex, whereas binding of the third strand

contributes only slightly to selectivity. In contrast, the analogous PNA_2-DNA triplexes are not formed by homopurine PNA.[68] PNAs also form stable PNA_2-RNA triplexes with RNA.[69]

Quadraplex formation: The novel supramolecular architecture of G-quartets has led to the development of interesting and functional noncovalent assemblies such as G-wires,[70] ion-channels[71] and self-assembled ionophores.[72] PNAs have been developed to mimic Watson-Crick and Hoogsteen base-pairing, and they are expected to be in proper register for participating in G-tetrad formation as well. In an attempt to use this mode of molecular recognition homologous G-rich PNA and DNA oligomers hybridize to form a PNA_2-DNA_2 quadruplex. The hybrid quadruplex exhibits high thermodynamic stability and expands the range of molecular recognition motifs for PNA beyond duplex and triplex formation.

PNA targeting double stranded DNA: Homopyrimidine PNAs complex with target DNA duplexes to form triplexes. An interesting and peculiar property of PNA is that, it can bind to duplex DNA to form triplex by displacing one of the strands of DNA duplex. Other binding modes for PNA have been demonstrated (Figure 19), where it can bind to DNA duplex through unusual mechanism called strand invasion. Strand invasion based on PNA-DNA formation (Figure 19c) appears to be limited to PNAs that form extremely stable PNA-DNA, such as very purine rich PNAs.[73] Classical triplex formation with a single PNA Hoogsteen strand (Figure 19a) has been observed as a kinetic intermediate.[74] However, PNA:DNA_2 triplexes are much less stable than

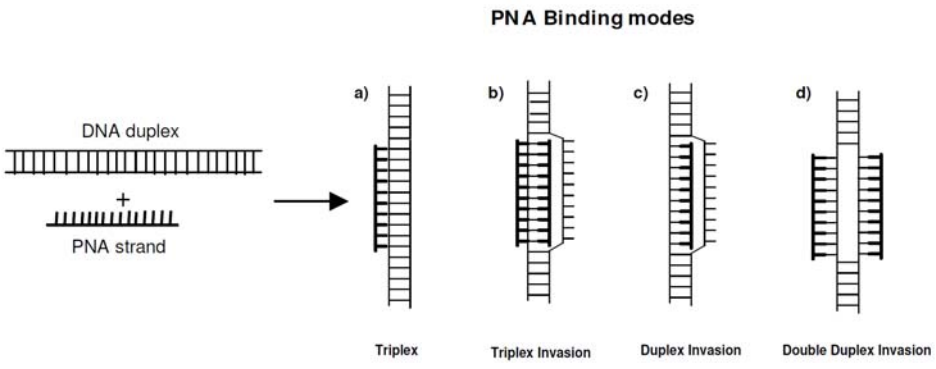


Figure 19: Schematic representation of PNA binding modes for targeting double stranded DNA. Thick structures signify PNA

The corresponding triplex invasion complexes (Figure 19b). Double duplex invasion complexes (Figure 19d) can also be formed using pseudo-complementary PNA

containing diaminopurine-thiouracil 'base pairs' that sterically destabilize the competing PNA-PNA duplex.[75]

Structure of PNA-DNA duplexes: Detailed structural information has been obtained from the NMR spectroscopic study of two antiparallel PNA-DNA duplexes (Figure 19). The DNA strand is in a conformation similar to the B-form, with a glycosidic *anti*-conformation, and the deoxyribose in C2'-*endo* form. A more recent NMR study showed that an octameric antiparallel PNA-DNA duplex contained elements of both A-form and B-form. The primary amide bonds of the backbone are in *trans* conformation and the carbonyl oxygen atoms of the backbone-nucleobase linker point towards the carboxy-terminus of the PNA strand. The CD patterns of antiparallel PNA-DNA complexes are similar to DNA-DNA CD and indicate the formation of right handed helix.[76]

Structure of PNA-RNA duplexes: The structural information of PNA-RNA complex has also been obtained from NMR solution structure study (Figure 20).[77] All bases form Watson-Crick base pairs, the glycosidic torsion angle in the RNA strand indicates an *anti*-conformation, and the ribose sugars are in the 3'-*endo* form.

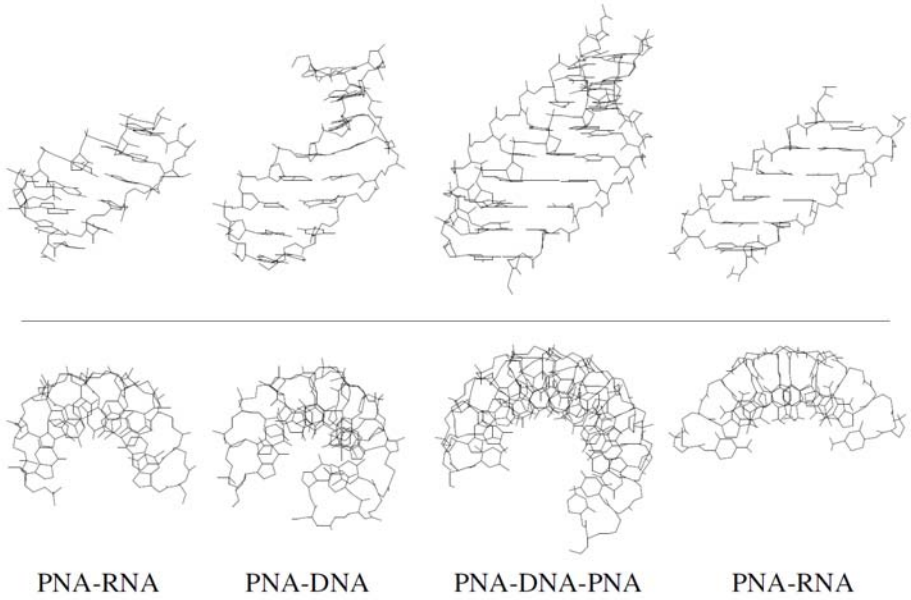


Figure 20: Structures of various PNA complexes in side view (upper panel) or top view (lower panel)

The RNA strand thus resembles an A-form structure. The tertiary amide bonds all are in the *cis* conformation. The carbonyl group of the tertiary amide in PNA backbone is isosteric to the C2'-hydroxyl group, which increases the solvent contact of the carbonyl

oxygen atom. The CD spectra of antiparallel PNA-RNA duplexes also indicate the formation of a right-handed helix with geometry similar to the A- or B-form.

Structure of PNA_2-DNA triplexes: The structural information of PNA_2-DNA triple-helix was obtained from the X-ray crystal structure analysis of the complex formed by bis-(PNA) and its complementary antiparallel DNA (Figure 20).[78] The nucleobases of the PNA strand bind to the DNA by Watson-Crick pairing and Hoogsteen hydrogen bonding. The structure is different, however, from both A-form and B-form DNA, and forms a "P-helix" with 16 bases per turn. The DNA phosphate groups are hydrogen bonded to the PNA backbone amide protons of the Hoogsteen strand. These hydrogen bonds, together with additional Van-der Waals contact and the lack of electrostatic repulsion are the main factors responsible for the enormous stability of the triplex. The deoxyribose of DNA strand is in C3'-*endo* conformation like A-form and bases lie almost perpendicular to the helix axis, which is characteristic of B-form DNA. The crystal structure is in agreement with the CD spectra of PNA_2-DNA triple-helices measured in solution.[79] The X-ray structure of self-complementary PNA-PNA duplex bears a strong similarity to the P-form of PNA_2-DNA triplex (Figure 20).[80]

1.4.2 Inhibition of gene expression

Antisense: Unlike antisense oligonucleotides, PNA-RNA hybrids are not substrates for RNase H65,[81] and therefore PNAs must exert their antisense effect by other mechanisms, such as direct steric blocking of ribosomes or essential translation factors. The targets around the AUG initiation codon in general seems sensitive.[65,82] In contrast, it has been found by *in vitro* translation experiments that mixed purine/pyridine sequence PNAs, which form PNA-RNA duplexes with the target, did not arrest translation-elongation, indicating that the PNA was displaced by the moving ribosome.[65] However, when homopurine targets are present in the translated region of gene, these can be targeted by homopyrimidine bis-(PNA), which forms extremely stable PNA_2-DNA triplexes that are indeed capable of arresting the ribosome during elongation.[65,78] Therefore, as with general antisense gene targeting a sensible first choice would be targets close to or overlapping the AUG initiation site.

Antigene: Homopurine regions of 8 bp or more in length in double-stranded DNA (dsDNA) can be targeted by homopyrimidine PNAs via the formation of extremely stable PNA triplex strand invasion complexes.[60,83] Because two PNAs-one W-C bound

and the other Hoogsteen bound are required to form these complexes, most often bis-PNAs in which the two parts are chemically linked are employed for dsDNA targeting.[84] PNA triplex invasion complexes occlude protein binding (transcription factors, restriction enzymes) to proximal or overlapping DNA sites that have sufficient stability to arrest the elongating RNA polymerase[78, 85] or DNA polymerase, thereby making PNAs good candidates for antigene reagents.[86] Binding to dsDNA is, however, very sensitive to even moderate ionic strength,118a and by only employing bis-PNAs, stable binding is possible at 140 mM KCl.[87] Thus the possibilities of using antigene PNAs *in vivo* seems slim.

Inhibition of Replication: The elongation of DNA primers by DNA polymerases can be inhibited by PNAs in cell free systems. Consequently the inhibition of DNA replication by PNAs should be possible via DNA duplex invasion under physiological pH or if the DNA is single stranded during the replication process as in the case of extra chromosomal mitochondrial DNA.

1.4.3 Chemical Modifications of PNA

1.4.3a Preorganization through conformational constraints

The structure of the classical PNA monomer (1) has been subjected to a variety of rational modifications with the aim of understanding the structure activity relations in this class of DNA mimics. The modifications also address the drawbacks of PNA oligomers and to improve the properties for various applications within medicine, diagnostics, molecular biology etc. These drawbacks include low aqueous solubility, ambiguity in DNA binding orientation and poor membrane permeability. Structurally, the analogues can be derived from modifications in the ethylenediamine or glycine sector of the monomer, linker to the nucleobase, the nucleobase itself or a combination of the above.

The strategic rationale behind the modifications[88] are (i) introduction of chirality into the achiral PNA backbone to influence the orientational selectivity in complementary DNA binding, (ii) rigidification of PNA backbone *via* conformational constrain to pre-organize the PNA structure and entropically drive the duplex formation, (iii) introduction of cationic functional groups directly in the PNA backbone, in a side chain substitution or at the *N* or *C* terminus of the PNA, (iv) modulate nucleobase pairing either by modification of the linker or the nucleobase itself and (v) conjugation

with ‘transfer’ molecules for effective penetration into cells. In addition to improving the PNA structure as above for therapeutics, several modifications are directed towards their applications in diagnostics.

The earliest and the simplest of the modifications involved extension of the PNA structure with a methylene group individually in each of the structural sub-units (aminoethyl,[89] glycine[90] and base linker[88]) of the PNA monomer. These resulted in PNAs with *N*-(2-aminoethyl)-β-alanine 2 and *N*-(3-aminopropyl) glycine 3 backbone and ethylene carbonyl linked nucleobase 4. However, these modifications resulted in a significant lowering of T_m of the derived PNA:DNA hybrids. The deleterious consequences of such subtle changes to the PNA structure suggested the high structural organization to which the original PNA structure is inherently tuned for interaction with

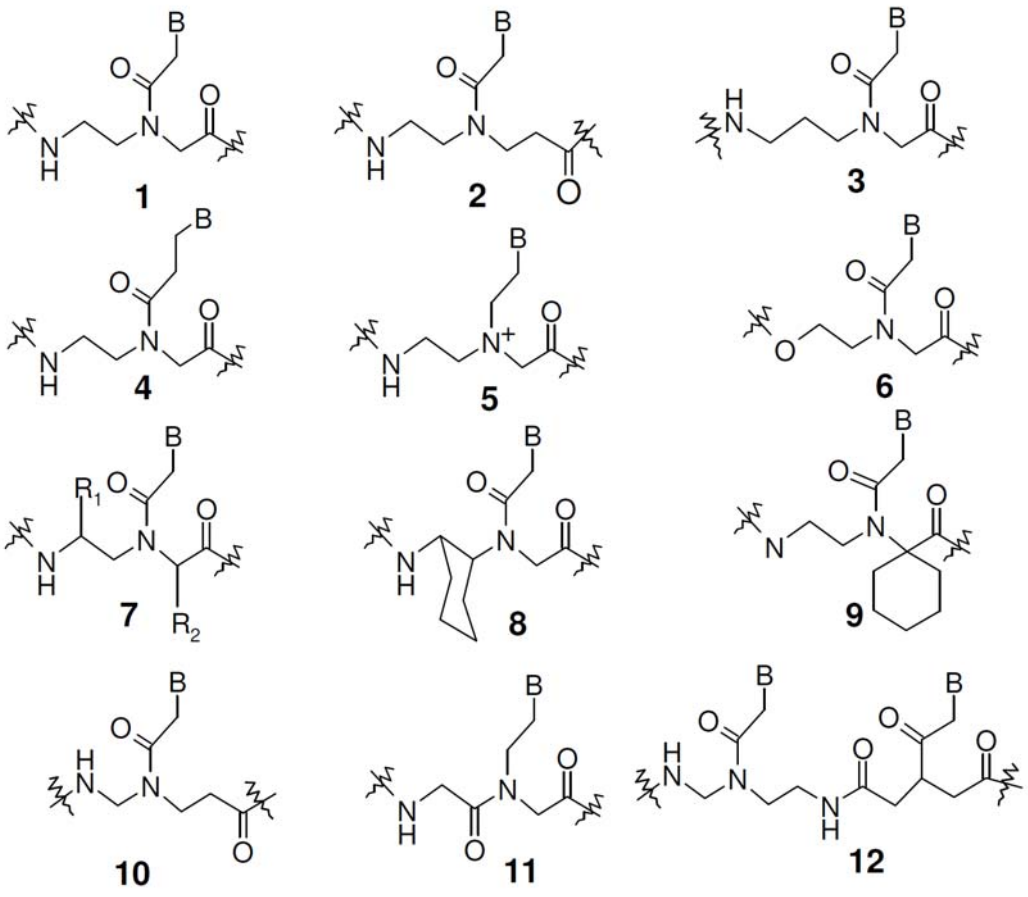


Figure 21: Chemical modifications of PNA

DNA. The replacement of the tertiary amide carbonyl by a methylene group leading to a flexible, cationic tertiary amine monomer 5 resulted in a large

destabilization of the PNA:DNA hybrids.[91] The necessity of such a pseudo rigid amide group pointed to the importance of constrained flexibility in the backbone.

Further rigidification of the PNA backbone has been attempted by introduction of alkyl substituents individually or simultaneously in the aminoethyl 7a or glycine[83,92,93] 7b segments or in both. Suitable substitutions may also lead to generation of cyclic structures with 1,2-cyclohexylamino- (8) and spirocyclohexyl (9) rings in monomers.[94,95] While PNAs that bear (*S,S*) cyclohexyl ring in the aminoethyl part hybridize with complement DNA similar to the unmodified PNA, those derived from (*R,R*) cyclohexyl moiety (8b) lacked such a property.[89] Thermodynamic data showed that DNA binding of the *SS*-isomer PNA was accompanied by a reduced loss in entropy, but this was counter-balanced by a decreased gain in enthalpy[89]A number of modifications generated by substitution of glycine component by other α-amino acids, leading to chiral PNA 7 (R_1=H) having hydrophobic, hydrophilic or charged α-substituents have been reported.[88] PNA oligomers incorporating chiral monomers retained the hybridization properties though less efficiently, with tolerance for small and medium substituents at the glycine-α position. Substitution with L/D-alanine showed a slight preference for antiparallel binding with DNA, with D-alanine being slightly better than L-alanine.[87] Among other replacements, only those derived from D-lysine exhibited DNA hybridization properties as good as that of original PNA. The incorporation of chiral monomers enhanced the sequence selectivity of PNA oligomers in hybridization, with maximum for D-glutamic acid and D-lysine substitutions. The lysine-modified oligomers were also more readily soluble in aqueous systems. In general, the different substituents caused equal or lower destabilization of PNA:RNA hybrids as compared to PNA:DNA hybrids.

Another type of modification involved interchange of various CO and NH groups on the peptide linkages leading to retro inverse[96] 10, peptoid[97] 11 and heterodimeric[98] 12 analogues. In all these systems, the inter-base residue separations are similar to the unmodified PNA, but accompanied by inversion of intra and inter residue amide bonds. Except for the heterodimer analogue, these exhibited a lower potency for duplex formation with complementary DNA/RNA suggesting that in addition to geometric factors, other subtle requirements such as hydration and dipole-dipole interactions, etc

influencing the microenvironment of the backbone, may be involved in effecting efficient PNA:DNA hybridization.

1.4.3b Pre organization through rigid five membered heterocycles

Some of the relatively successful conformationally preorganized modifications[99] so far are based on introduction of methylene/ethylene groups to bridge the aminoethyl-glycyl backbone and methylene carbonyl side chain to generate diverse five or the six-membered nitrogen hetetocyclic analogues. The cyclic analogues where the nucleobases are directly attached to the ring have defined nucleobase orientation, overcoming the rotamer problem. It also concomitantly introduces chiral centers, which may impart directional selective binding of PNA with chiral DNA/RNA.

The naturally occuring amino acid *trans*-4-hydroxy-L-proline, a five-membered nitrogen heterocycle, is a versatile, commercially available starting material for creating structural diversity to mimic DNA/PNA structures. From this amino acid a wide variety of chiral, constrained and structurally preorganized PNAs have been synthesized. Depending on the synthetic approach and on the presence of the tertiary amine group in the monomers, the modifications afford either positively charged or uncharged cyclic PNA analogs. The different cyclic PNAs proposed showed that the right stereochemistry and conformation is really important for binding abilities towards nucleic acids and in some cases even for discriminate between RNA and DNA.

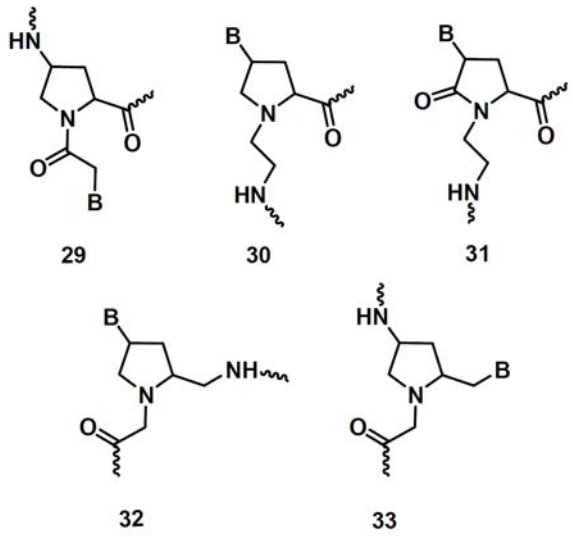


Figure 22: Structures of conformationally preorganized five and six membered heterocyclic PNA monomers

For example, in the case of the cyclic PNA analogue *N*-(thymin-1-yl-acetyl)-4-aminoproline 29 (Figure 20), all stereoisomer's were synthesized, and the L-*trans*-4-aminoprolyl isomer, was shown to bind to DNA with higher affinity, while the L-*cis* isomer and the D-*trans*, which could not adopt the same spatial arrangement, showed reduced performances. The same model can be applied to the (2*R*,4*S*)-stereoisomer of the aminoethylprolyl PNA57 30 (*aep*PNA), Effect of a chiral monomer on DNA affinity might be dependent on whether it is used for the synthesis of all the PNA oligomer or it is inserted in an oligomer made of achiral monomers.

The aminoethylprolyl-5-one (*aepone*) thymine monomers[100] 31 were synthesized and incorporated into *aeg*-PNA-T8 backbone at different positions. The *aepone*-PNAs showed remarkable stabilization of derived PNA_2:DNA triplexes compared to *aeg*-PNA. In particular the synthesis of pyrrolidine-based chiral positively charged PNA 32 (Figure 22) the derived (2*R*,4*S*) stereomeric homoadenylate oligomer formed a stable complex with both DNA and RNA. PNAs containing the other pyrrolidine stereoisomers are reported but they do not show any considerable improvement in binding affinity.

Introduction of a methylene bridge between the α' and β carbon yields another pyrrolidine-PNA[101] 33. Diastereomeric monomers bearing T, A, C, G, nucleobases were introduced in PNA oligomers and the complexation with DNA and RNA sequences was studied. It was found that: (2*R*,4*S*) homopyrimidine PNA stabilize PNA_2:DNA triplexes, (2*S*,4*R*) stereoisomers in mixed sequences affects enhanced DNA duplex stability, (2*S*,4*S*) and (2*R*,4*R*) remarkably enhance PNA:RNA duplex stability. All the pyrrolidine modifications show preference for the anitparallel DNA binding. Thus in this case the effect of stereochemistry is not very clear, and relative conformational freedom of the five-membered ring can account for this.

1.4.3c Modified Nucleobases

There is increasing interest in modulating and expanding the recognition motifs of standard base pairs. Employing non-natural nucleobase ligands in place of natural nucleobases would help understand the recognition process in terms of various factors contributing to the event such as hydrogen bonding and internucleobase stacking. Further, new recognition motifs may also have potential applications in diagnostics and nanomaterial chemistry. This when coupled with high affinity and strand invasion

properties offered by PNA would add a new dimension to PNA applications. The nonstandard nucleobases employed so far with PNA are limited, compared to the repertoire of backbone modifications described earlier. 2, 6-Diaminopurine[102] (36, Figure 23) offers increased affinity and selectivity for thymine and pseudoisocytosine[103] 37 is a very efficient mimic of protonated cytosine for triplex formation. 2-Aminopurine[104] 38 hydrogen bonds with U and T in reverse Watson-Crick mode and has the advantage of being inherently fluorescent to enable study of kinetic events associated in hybridization. The E-base[105] 39 was rationally designed for recognition of A:T base pair in the major groove and form a stable triad with T in the central position.

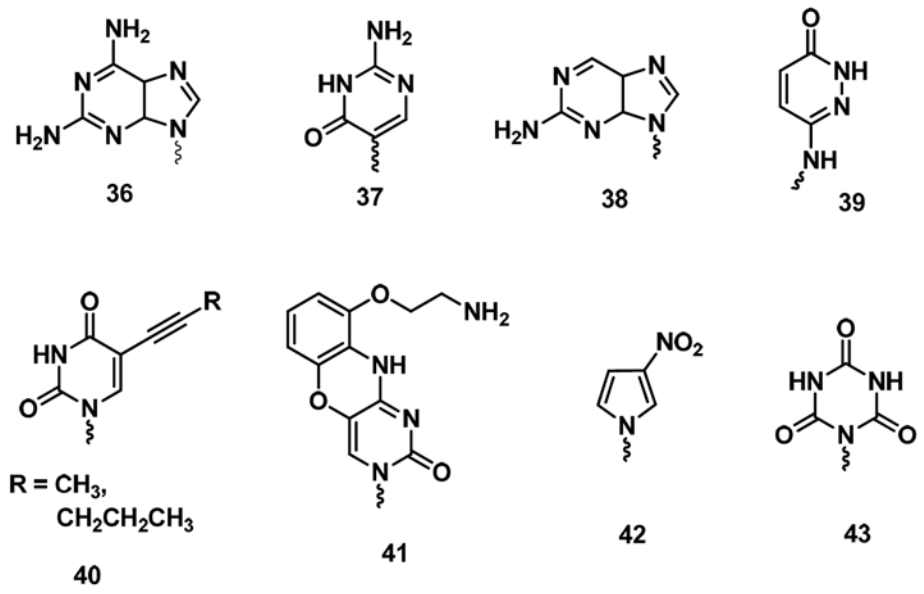


Figure 23: Structures of modified nucleobases

In order to increase the stability of complexes formed by PNA with target nucleic acids, bases that possess a larger surface area (for greater hydrophobic/stacking interactions), make additional H-bonds or that are positively charged have been prepared. A wide variety of 5-substituted uracils were synthesized and their ability for triplex formation has been studied[106] (40, Figure 23). The G-clamp base 41 was developed to build in specific, additional bonding interactions with guanine.[107] Unnatural heterocycles 3-nitropyrrole 42 have been used as potential universal bases in PNA.[108] The synthesis of cyanuryl PNA monomer containing cyanouric acid 43 as the base was achieved by direct N-monoalkylation of cyanuric acid with *N*-(2-Bocaminoethyl)-*N*′-(bromoacetyl)glycyl ethyl ester.[109] The monomer was incorporated

as a T-mimic into PNA oligomers and biophysical studies on their triplexes/duplex complexes with complementary DNA oligomers indicated unusual stabilization of PNA:DNA hybrids when the cyanuryl unit was located in the middle of the PNA oligomer.

PNA-DNA Chimerae: The successful applications of the remarkable DNA binding properties of PNAs are sometimes (sequence/length dependent) hampered by their tendency to self-aggregate and poor aqueous solubility. Overcoming these limitations and imparting other abilities for therapeutic applications such as cellular uptake, RNase H activation properties, have been addressed by designing covalent hybrids or chimaeras of PNA with DNA, functional peptides and other effector molecules. Three types of PNA-DNA chimeras (Figure 23) are in place (i) 5'-DNA-linker X- PNA -*pseudo*-3' [110, 111] 24 (ii) *pseudo*-5'-DNA-linker X- DNA-3' [112] 25 and (iii) *pseudo*-5'-PNA-linker X- DNA-3'[105] 26. Synthetic protocols have been developed with protecting groups compatible for carrying out on-line synthesis of both PNA and DNA to generate the chimeras. Several interesting properties were noticed in such covalent hybrids such as co-operative stabilizing effects against proteases and nucleases, enhanced water solubility and duplex/triplex stabilities dependent on the structure of chimera and the linker. The linker can also be a fragment of DNA to generate PNA-(5')-DNA-(3')-PNA chimera which formed stable duplexes with both DNA and RNA with a lower stability than corresponding DNA:DNA and DNA:RNA duplexes.[113]

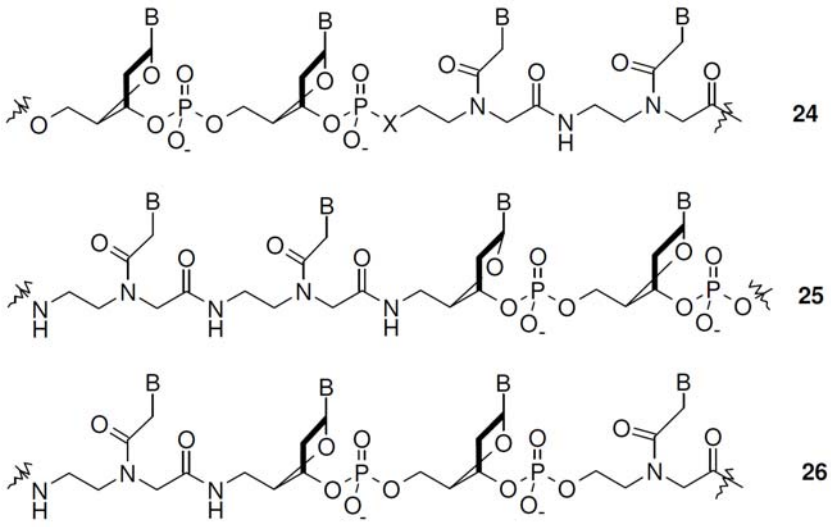


Figure 24: PNA-DNA chimeras

1.4.4 Cellular Uptake of PNA

Peptide nucleic acids suffer from poor aqueous solubility and cellular uptake, because of its neutral backbone. The solubility problem was solved by selecting lysine amino acid as a spacer chain during the solid phase synthesis of PNAs and also by synthesizing PNA analogues that possesses positive charge in the backbone by default. In an effort to equip PNAs with a lipophilic tail that would confer liposome affinity, PNA lipid (adamantyl) conjugates (Figure 25b) have been prepared and studied their liposome mediated cellular uptake.[114]

The properties of these conjugates are influenced by PNA sequence and results were encouraging and in several cases diffuse cytoplasmic uptake was observed when using cationic liposomes as carriers. However certain peptides are internalized very efficiently by cells.[115] The remarkable uptake properties of HIV-1 Tat transduction domain is the result of short basic sequences of (GRKKRRQRRR). They are amphiphatic –helices with high content of basic amino acid residues (Lys, Arg).

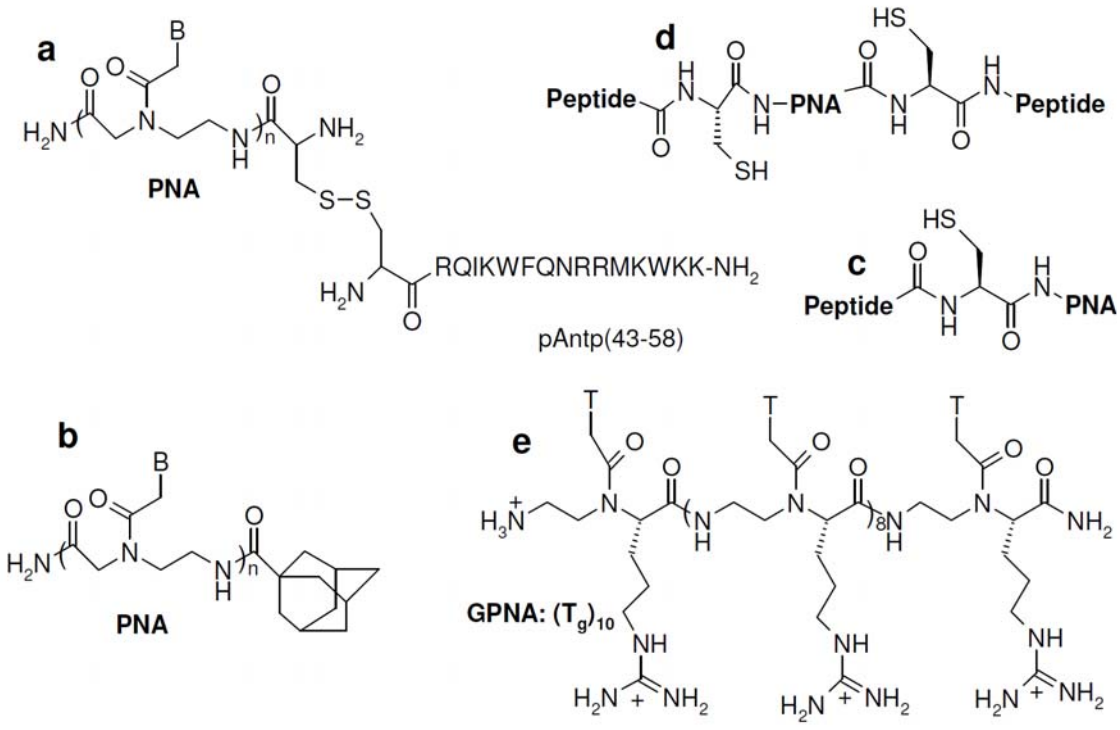


Figure 25: PNA-peptide conjugates (a, c, d). PNA-lipid (adamantyl) conjugate (b). Cationic PNAs (GPNA, e)

These peptides are conjugated to one or both the end of PNA by an amide or disulfide linker (Figure 25a, c, d).[116] The results on cellular uptake of such PNA-peptide

conjugates are impressive. In another strategy PNA containing arginine side chains (GPNA, Figure 25e) exhibited remarkable uptake properties.[117] Interestingly, the cellular uptake of these GPNAs is not by endocytosis-driven nor receptor mediated perhaps through membrane flipping.

1.4.5 Applications of PNA

Antigene and antisense applications: Peptide nucleic acids have promise as candidates for gene therapeutic drug design. They require well-identified target and well characterized mechanism for their cellular delivery PNAs are chemically and biologically stable molecules and have significant effect on replication, transcription (antigene), and translation processes, as revealed from in vitro experiments. Moreover no sign of any general toxicity of PNA has so far been observed.

Interaction of PNA with enzymes. *RNase H*: The antisense oligonucleotides with an RNase H activity (e.g., PS-oligos) is considered a better antisense molecule (inhibitor) than one without the activity (methylphosphonates and HNAs).[107]

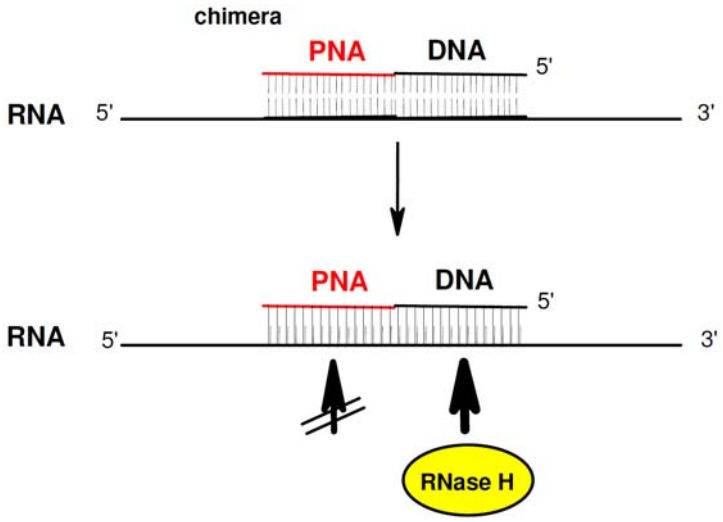


Figure 26: Schematic representation of RNase H-mediated cleavage of activity after the binding of a PNA-DNA chimera to an RNA target

Despite their remarkable DNA binding properties, PNAs generally are not capable of stimulating RNase H activity on duplex formation with RNA. However, recent studies have shown that DNA/PNA chimeras are capable of stimulating RNase H activity. On formation of a chimeric RNA double strand, PNA/DNA can activate the RNA cleavage activity of RNase H (Figure 26). Cleavage occurs at the ribonucleotide

parts base-paired to the DNA part of the chimera. Moreover, this cleavage is sequence specific in such a way that certain sequences of DNA/PNA chimerae are preferred over others.[107] They are also reported to be taken up by cells to a similar extent as corresponding oligonucleotides.[107] Thus, PNA/DNA chimerae appear by far the best potential candidates for antisense PNA constructs.

Diagnostics: The properties of PNAs in forming complexes with DNA like, high thermal stability, better sequence discrimination, strand invasion of dsDNA, and more importantly they can alter the gel electrophoretic mobility of complementary nucleic acids considerably on binding on account of their neutral character made them as the choice for the use in diagnostics.

The "PCR clamping" method for the detection of point mutations is based on the ability to form strong PNA-DNA complex and inability to function as primers in the polymerase chain reaction (PCR). Targeting the PNA oligomer, even partially, against the primer binding site can block the formation of the PCR product. Skilful choice of the length of the primer can even allow discrimination of alleles, which differ only in one base pair. Discrimination of point mutations in the Ki-*ras* gene has been carried out by PCR clamping, where the amplification of the wild-type DNA is inhibited, relative to mutated gene, which binds with lower affinity (Figure 27). Demers *et al*. have used PNAs to suppress the preferential amplification of small allelic PCR products during copying of VNTR *loci*, which leads to false genotypic patterns. The quantitation of telomeric repeats is also possible with PNAs. The use of PNAs immobilized on surfaces as sequence specific DNA biosensors.[118]

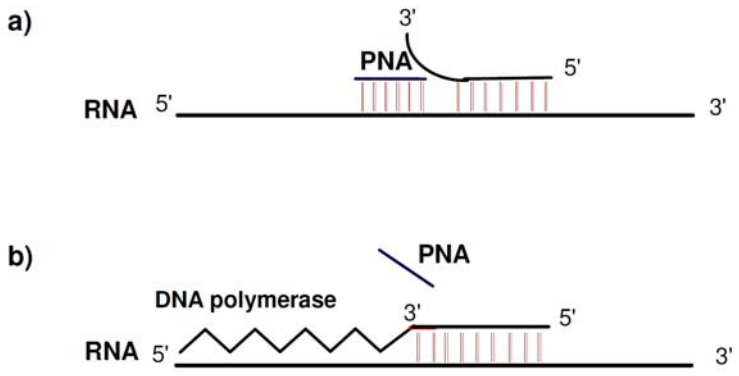


Figure 27: 'PCR clamping' technique. (a) inhibition of wild-type DNA amplification and (b) amplification of mutated DNA, by PNA

PNAs as tools in biotechnology: PNAs can be used for the modulation of enzymatic cleavage. After strand displacement by PNA in a DNA duplex, the displaced DNA strand can be selectively cleaved by nuclease S1.[119] The single strand specific nuclease S1, in combination with two PNAs, be transformed into an "artificial restriction enzyme" that cleaves both DNA strands, and whose recognition sequence is determined by the PNA sequences employed (Figure 28). Conversely PNAs can also be used to block DNA cleavage by restriction enzymes.[120]

Furthermore, PNAs can be used to prevent the methylation of DNA sequence specifically. On dissociation of the PNA-DNA complex only the modified recognition sequence can be cleaved by the appropriate methylase sensitive restriction enzyme. This principle of "rare genome-cutters" has been exploited in a similar way with triple helix forming oligonucleotides.[121]

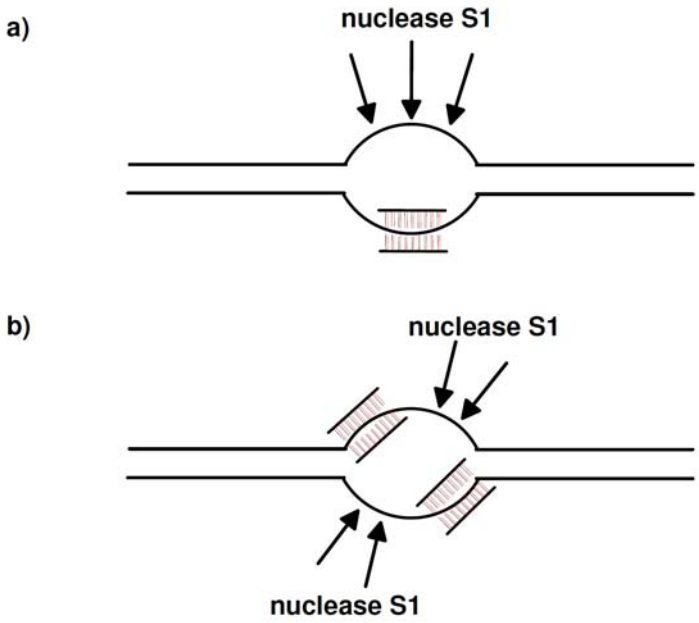


Figure 28: Artificial restriction enzymes: (a) single stranded cleavage by PNA (b) double stranded cleavage by PNA clamping

Nucleic acid purification: Based on its unique hybridization properties, PNAs can also be used to purify target nucleic acids. PNAs carrying six histidine residues have been used to purify target nucleic acids using nickel affinity chromatography. Also, biotinylated PNAs in combination with streptavidin-coated magnetic beads may be used to purify *Chlamydia trachomatis* genomic DNA directly from urine samples. However, it appears that this simple, fast, and straightforward 'purification by hybridization'

approach has drawbacks. It requires the knowledge of a target sequence and depends on a capture oligomer to be synthesized for each different target nucleic acid. Such target sequences for the short pyrimidines PNA, i.e., the most efficient probe for strand invasion, are prevalent in large nucleic acids. Thus short PNAs can also be used as generic capture probes for purification of large nucleic acids. It has been shown that a biotin-tagged PNA-thymine heptamer could be used to efficiently purify human genomic DNA from whole blood by a simple and rapid procedure.

1.5 Spectroscopic methods for studying DNA/RNA Interactions

The ability of antisense oligonucleotides to bind in vitro the target DNA/RNA, can be detected by using various biophysical techniques as follows

UV-Spectroscopy: Nucleic acid complexes have lower UV absorption than that predicted from the sum of their constituent base extinction coefficients that is usually measured at 260 nm. This phenomenon known as "hypsochromicity" results from coupling of the transition dipoles between neighboring stacked bases and is larger in amplitude for A-U and A-T pairs.[122] As a result, the UV absorption of a DNA duplex increases typically by 20-30 % when it is denatured. This transition from a helix to an unstacked, strand-separated coil has a strong entropic component and so is temperature dependent. The mid-point of this thermal transition is known as the melting temperature (T_m). Such dissociation of nucleic acid helices in solution to give single stranded DNA/RNA is a function of base composition, sequence, and chain length as well as of temperature, salt concentration, and pH of the solvent (buffer). In particular, early observations of the relationship between T_m and base composition for different DNAs showed that A-T pairs are less stable than G-C pairs, a fact which is now expressed in a linear correlation between T_m and the gross composition of a DNA oligomer by the equation:

$$T_m = X + 0.41\ (\%C+G)\ ^\circ C$$

The constant X is dependent on salt concentration and pH and has a value of 69.3ºC for 0.3 M sodium ions at pH 7.

A second consequence is that the steepness of the transition curve also depends on base sequence. Thus, melting curves for homooligomers have much sharper transitions than those for random-sequence oligomers. This is because A-T rich regions melt first

to give unpaired regions which then extend gradually with rising temperature until; finally, even the pure G-C regions have melted. Short homooligomers melt at lower temperatures and with broader transitions than longer homooligomers. For example for poly (rA)n-poly (rU)n, the octamer melts at 9ºC, and the undecamer at 20ºC, and long oligomers at 49ºC in the same sodium cacodylate buffer at pH 6.9. Consequently, in the design of synthetic self complementary duplexes for crystallization and X-ray structure determination, G-C pairs are often placed at the ends of hexamers and octamers to stop them 'fraying'. Lastly, the marked dependence of T_m on salt concentration is seen for DNA from *Diplococcus* 21 *pneumoniae* whose *T*m rises from 70ºC at 0.01 M KCl to 87ºC for 0.1 M KCl and to 98ºC at 1.0 M KCl. The converse of melting is the renaturation of two, separated, complementary strands to form a correctly paired duplex.

Duplex melting: Duplexes exhibit a single transition into single strands accompanied by an increase in UV absorption termed as *hyperchromicity*. According to the 'all or none model'[123] the UV absorbance value at any given temperature is an average of the absorbance of duplex and single strands. A plot of absorbance against temperature gives a sigmoidal curve in case of duplexes and the midpoint of the sigmoidal curve (Figure 29) called as the 'melting temperature' (T_m) is the temperature (equilibrium point) at which the duplex and the single strands exist in equal proportions. *Triplex melting*: In the case of triplexes, the first dissociation leads to melting of triplex generating the duplex (WC duplex) and third strand (Hoogsteen strand), followed by the duplex dissociation to form two single strands. The DNA triplex melting shows characteristic double sigmoidal transition (Figure 29) and UV melting temperature for each transition is obtained from the first derivative plots. The lower melting temperature (T_m1) corresponds to triplex to duplex transition while the second higher melting temperature (T_m2) is the transition of duplex to single strands.

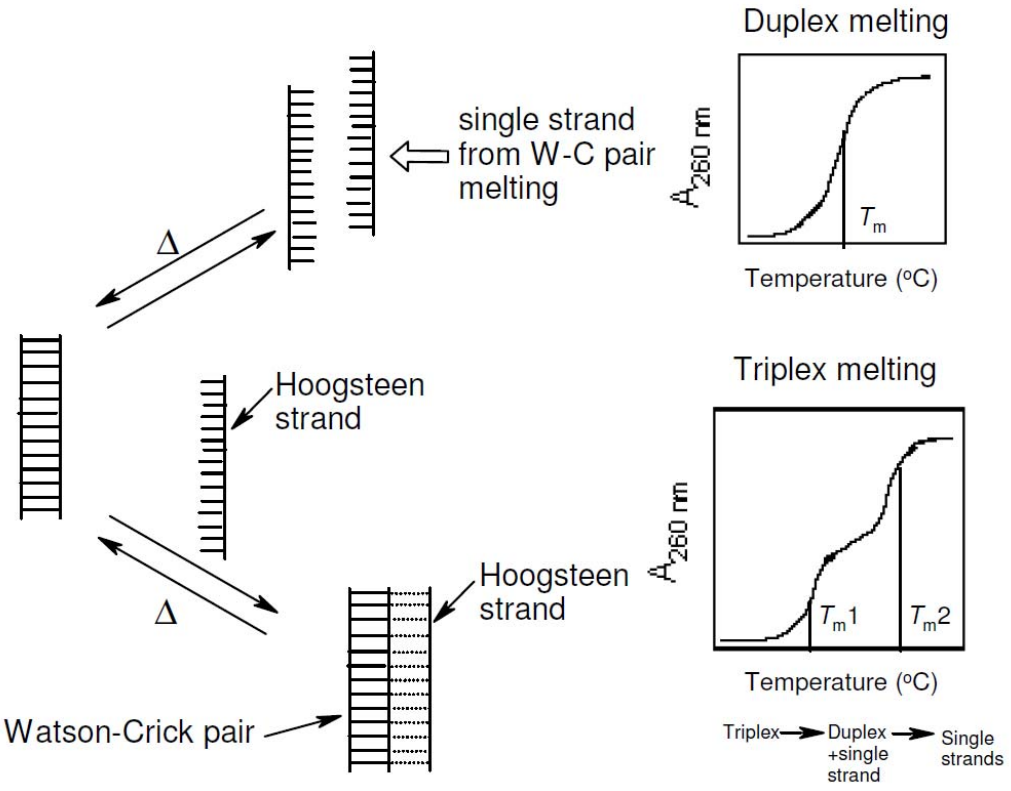


Figure 29: Schematic representation of DNA duplex and Triplex

Stoichiometry: The binding stoichiometry of oligonucleotide to target DNA/RNA by UV-titration (mixing): This is derived from Job's plot. The two components of the complex are mixed in different molar ratios, so that the total concentration of each mixture should be constant, i.e., as the concentration of one strand decreases, concentration of the second strand increases and the UV-absorbance of each mixture is recorded. The absorbance decreases in the beginning to a minimum and then again increases. The molar ratio of the two strands at which absorbance reached minimum indicates the stoichiometry of complexation.

Circular Dichroism (CD): Circular dichroism is a spectroscopic technique that depends on the property of optical activity, i.e., chiral molecules that interact with the right and left circularly polarized lights differently. Circularly polarized light occurs in two non-super imposable right and left circularly polarized forms, which are related to each other as mirror images.

The most common applications of CD include probing the structure of biological macromolecules, in particular determining the α-helical content of proteins and conformational analysis of nucleic acids and interaction with other ligands.

Single stranded DNAs are structurally less defined than duplex DNAs and their CD signal is smaller. CD spectra of RNAs show that naturally occurring RNAs to be mainly well structured with extensive duplex regions (stem-bulb).

In an isolated nucleotide, the chirality originates from the ribose sugar and influences the base absorption. In the normal spectral region (down to 180 nm), the chiral ribose-phosphate backbone has no transitions and in the CD spectrum of nucleic acids one detects the CD induced into base transitions as a result of their coupling with the backbone transitions.[124] The magnitude of $\Delta\epsilon_{max}$ is of the order of 2 mol^{-1} dm^3 cm^{-1} at 270 nm and the purine bases have a negative signal whereas the pyrimidines have a positive signal. In the case of CD spectrum of DNA and RNA polynucleotide with stacked bases, the magnitude is larger at 270 nm and significantly larger at 200 nm than that of individual bases. The spectrum is dominated by the induced CD transitions of the bases from their coupling among each other, due to their stacking on top of each other in a chiral (helical) fashion.

The simplest application of CD to DNA structure determination is for identification of polymorph present in sample.[125] The CD signature of the B-form DNA as read from longer to shorter wavelength is a positive band centered at 275 nm, a negative band at 240 nm, with cross over around 258 nm. These two bands arise not from a simple degenerate exciton coupling, but as a result of superimposition of all coupling transitions from all the bases.

If B-DNA is compacted and the bases tilted and radially displaced from the centre of the helix then A-form DNA results. A-DNA is characterized by a positive CD band centered at 260 nm that is larger than the corresponding B-DNA band, a fairly intense negative band at 210 nm and a very intense positive band at 190 nm. The 250-230 nm region is also usually fairly flat though not necessarily zero. Naturally occurring RNAs adopt the A-form if they are duplex.

CD denaturation (melting) and CD-Job's plots are equally important as they give *T*m and binding stoichiometry of DNA/RNA complexes respectively.

1.6 The present work

The various DNA and PNA modifications have been reviewed in the preceding sections. All these modifications exhibit excellent properties in some way but lack some

other important properties. PNA which is most promising modification of DNA has the major drawbacks like poor water solubility, inefficient cell uptake; self aggregation and ambiguity in directionality of binding restrict its applications. To use PNA as antisense agent, these problems need to be addressed.

Chapter 2: In this Chapter we have described the petPNA (pyrrolininyl ethyl PNA), a modified analogue of PNA and its biophysical studies with DNA and RNA. This Chapter has been divided into two sections

Section 1: The design of petPNA contains a cyclic structure which may address the problem of directionality of binding with target DNA/RNA. Backbone extended cyclic PNA analogues are known to increase the binding preferences of PNA with RNA than DNA. The petPNA presented in this work is also a cyclic extended version of aminoethyl segment of aeg PNA which may improve the binding selectivity of PNA. The presence of cationic groups like amino or guanidino in the backbone of petPNA may improve the cellular uptake of PNA. This Section deals with the synthesis of H-petPNA and Amino-petPNA monomer, their incorporation in the aegPNA backbone. The simple procedure for conversion of amino-petPNA to guanidino-petPNA on solid support also has been described in this Section.

Section 2: The study of the effect of introduction of cyclic, chiral pet PNA monomer into aeg PNA backbone on binding with DNA/RNA has been presented in this section. This section also describes the effect of amino or guanidino groups on the binding efficacy of petPNA modified PNA with DNA and RNA.

Chapter 3: This chapter deals with *iso*-Thoiacetamido nucleic acids(*iso*-TANA) a DNA analogue where phosphate backbone is replaced by thioacetoamido linkage and with inverted stereochemistry at 3'-amino group of the linkage. This chapter also is divided into two sections.

Section 1: The synthesis of various pyrimidine *iso*-TANA dimers has been described in this section. NMR studies were carried out to find out the effect of the change in the stereochemistry at 3'-amino group on the sugar puckering and the comparison of these results with corresponding TANA dimers is presented in this section. The CD studies were carried out for investigating the change in the base stacking interactions of dimers due to change in the stereochemistry at 3'-amino substitution in the thioactamido

linkage and are described in this section. Incorporation of these dimer blocks in chimeric amide-phosphate backbone, characterization of the resulting oligomers and their preference of binding with DNA/RNA in comparison with TANA modified oligomers is presented in this chapter.

Section 2: This section deals with the synthesis of *iso*-TANA thymine monomer and synthesis its homoligomers. The synthesis of thymine octamer sequences of aeg PNA containing *iso*-TANA thymine monomer also has been given. The biophysical studies of these sequences with DNA and RNA has been discussed in this Section.

1.7 References

1 Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acid. A structure for deoxyribose nucleic acid. *Nature*, **1953**, *171*, 737-738.

2 Thibaudeau, C.; Plavec, J.; Garg, N.; Papchikhin, A.; Chattopadhyaya, J. How does the electronegativity of the substituent dictate the strength of the guache effect ? *J.Am.Chem. Soc.* **1994**, *116*, 4038-4043.

3 a) Hoogsteen, K. The crystal and molecular structure of a Hydrogen-bonded complex between 1-methyl thymine and 9-methyl adenine *Acta. Cryst*. **1963**, *65*, 907.

4 Crick, F. H. C. Codon-anticodon pairing. The Wobble hypothesis *J. Mol. Biol*. **1966**, *19*, 548-555.

5 Seeman, N. C.; Rosenberg, J. M.; Rich, A. Sequence specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA*. **1976**, *73*, 804-807

6 **(**a) Weiss; B. Antisense Oligodeoxynucleotides and Antisense RNA: Novel Pharmacological and Therapeutic Agents, CRC Press, Boca Raton, FL, 1997 (ed.). (b) Stein; C. A Does antisense exist? *Nat. Med*. **1995**, *1*, 1119-1121.

7 Zamecnik; P. C., Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligonucleotide. *Proc. Natl. Acad. Sci. U.S.A*. **1978**, *75*, 280-284.

8 a)Bennett; C. F.; Cowsert, L .M. Antisense oligonucleotides as a tool for gene functionalization and target validation. *Biochim. Biophys. Acta* **1999,** *1489*, 19-30. b) Kurreck, J. Antisense technologies improvement trough novel chemical modifications *Eur. J. Biochem.* **2003**, *270*, 1628-1644.

9 Bennett, C. F.; Swayze, E. E. RNA targeting therapeutics: Molecular mechanisms of antisense oligonucleotides as a therapeutic platform *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259-293.

10 a) Dominski, Z. and Kole, R. *Proc. Nat. Acad. Sci U.S.A*. **1993**, *90*, 8673-8677. (b) Sazani, P. and Kole, R. *J. Clin. Invest.* **2003**, *112,* 481–486.

11 Gilbert, W Why genes in pieces?. *Nature* **1978**, *271,* 5645

12 (a) Draper, B.W.; Morcos, P.A.; Kimmel, C.B. Inhibition of zebrafish fgf 8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis* **2001**, *30*, 154-156. (b) Sazani, P. *et al*.Nuclear antisense effects of neutral, anionic and cationic oligonucleotide analogs. *Nucleic Acids Res.* **2001**, *29*, 3965-3974.

13 Zamecnik, P.C.; Stephenson, M.L. Inhibition of Rous-sarcoma virus-replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 280–284.

14 Cerritelli, S. M.; Crouch, R. J. Ribonucelase H: the enzymes in eukaryotes. *FEBS J.* **2009**, *276*, 1494–505.

15 (a) Crooke, S. T.; Lebleu, B. Antisense Research and Applications. Boca Raton, FL: CRC Press, Inc. **1993**. (b) Crooke, S. T. Progress in antisense therapeutics. *Med. Res. Rev*.**1996**, *16*, 319-344. (c) Crooke, S. T.; Bennet, C. F. Progress in antisense oligonucleotide therapeutics. *Annu. Rev. Pharmacol. Toxicol*. **1996**, *36*, 107-129. (d) Milligan, J. F.; Matteucci, M. D.; Martin, J. C. Current concepts in antisense drug design. *J. Med. Chem*. **1993**, *36*, 1923-1937. (e) Pierga, J. –Y.; Magdelenat, H. Applications of antisense oligonucleotides in oncology. *Cell.Mol. Biol.* **1994**, *40*, 237-261.

16 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.

17 Millar, P. S. Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): Oligodeoxynucleotides-Antisense Ihibitors of Gene Expression. London: *Macmillan Press*, **1989**, p. 79.

18 Froehler, B.; Ng, P.; Matteucci, M. Phosphoramidate analogs of DNA: Synthesis and thermal stabilities of heteroduplexes. *Nucleic Acid Res*. **1988**, *16*, 4831-4839.

19 (a) Stec, W. J.; Zon, G. Synthesis, separation, and stereochemistry of diastereomeric oligodeoxyribonucleotides having a 5'-terminal internucleotide phosphorothioates linkage.*Tetrahedron Lett*. **1984**, *25*, 5275. (b) Stec, W. J.; Zon, G: Stereochemical studies of the formation of chiral internucleotide linkages by phosphoramidite coupling in the synthesis of oligodeoxyribonucleotides. *Tetrahedron Lett*. **1984**, *25*, 5279-5282.

20 Stein, C. A. Does antisense exist? *Nat. Med*. **1995**, *1*, 1119-1121.

21 Millar, P. S: Non-ionic antisense oligonucleotides. In Cohen, J. S. (ed.): Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression. *London: Macmillan Press*, **1989**, p. 79.

22 Letsinger, R. L.; Singman, C. N.; Histand, G.; Salunke, M: Cationic ligonucleotides. *J. Am.Chem. Soc*. **198**8, *110*, 4470-4471.

23 Jung, P. M.; Histand, G.; Letsinger, R. L: Hybridization of alternating cationic/anionic oligodeoxyribonucleotides to RNA segments. *Nucleosides Nucleotides* **1994**, *13*, 1597-1605.

24. Summers, M. F.; Powell, C.; Egan, W.; Byrd, R. A.; Wilson, W. D.; Zon, G. Alkyl phospotriester modified oligodeoxyribonucleotides. VI. NMR and UV spectroscopic studies of ethyl phosphotriester (Et) modified Rp-Rp and Sp-Sp duplexes, (d(GGAA(Et)TTCC))2. *Nucleic Acid Res*. **1986**, *14*, 7421-7437.

25 (a) Sood, S.; Shaw, B. R.; Spielvogel, B. F. Boron-containing nucleic Acids. Synthesis of oligonucleoside boranophosphates. *J. Am. Chem. Soc*. **1990**, *112*, 9000. (b) Shaw, B. R.; Madison, J.; sood, S.; Spielvogel, B. F; Oligonucleotide boranophosphate (borane phosphate). In Agrawal, S. (ed): *Methods in Molecular biology*, vol *20*: Protocols for Oligonucleotides and Analogs. Synthesis and properties. Totowa, NJ. Humana Press, Inc., **1993**, pp. 225-243. (c) Sergueev, D. S.; Shaw, B. R: H-phosphonate approach for solid pase synthesis of oligodeoxyribinucleotide boranophosphates and their characterization. *J. Am. Chem. Soc*. **1998**, *120*, 9417-9427.

26 Li, H.; Huang, F.; Shaw, B. R. Conformational studies of dithymidine boranomonophosphate diastereoisomers. *Bioorg. Med. Chem*. **1997**, *5*, 787-795.

27 (a) Higson, A. P.; Sierzchala, A.; Brummel, H.; Zhao, Z.; Caruthers, M. H. Synthesis of an oligothymidylates containing boranophosphate linkages. *Tetrahedron Lett*. **1998**, *39*, 3899-3902. (b) Rait, V. K.; Shaw, B. R. Boranophosphates support the RNase H Cleavage of polyribonucleotides. *Antisense Nucleic Acid Drug Dev*. **1999**, *9*, 53-60.

28 (a) Gryaznov, S.; Chen, J.-K. Oligoribonucleotide N3'-N5' phosphoramidates: Synthesis and hybridization properties. *J. Am. Chem. Soc*. **1994**, *116*, 3143-3144. (b) Zielinska, D.; Pongracz, K.; Gryaznov, S. M. A new approach to oligonucleotide N3'-P5' phosphoramidates building blocks. *Tet. Lett*. **2006**, *47*, 4495-4499.

29 Tereshko, V.; Gryaznov, S.; Egli, M. Consequence of replacing the DNA 3'- Oxygen by an amino group. High resolution crystal structure of a fully modified N3'-N5' phosphoramidates DNA dodecamer duplex. *J. Am. Chem. Soc*. **1998**, *120*, 269-283.

30 Schultz, R. G.; Gryaznov, S. M. Oligo-2'-fluoro-2'-deoxynucleotide N3'-P5' phosphoramidates: synthesis and properties. *Nucleic Acids Res.* **1996**, *24*, 2966-2973.

31 Schultz, R. G.; Gryaznov, S. M. arabino-fluorooligonucleotide N3'-P5' phosphoramidates: synthesis and properties. *Tet. Lett*. **2000**, *41*, 1895-1899.

32 Gryazonv, S. M. Oligonucleotides N3'-P5' phospharamidates as potemtial therapeutic agents *Chemistry & Biodiversity* **2010**, *7*, 477-493.

33 Pongracz, K.; Gryaznov, S. Oligonucleotide N3'-P5' thiophoramidates : synthesis and properties *Tetrahedron Lett*. **1999**, *40*, 7661.

34 Hochreiter A.E.; Xiao, H.; Goldblatt, E. M.; Gryaznov, S.M;, Miller, K. D. Telomerase template antagonist GRN163L disrupts telomere maintenance, tumor growth, and metastasis of breast cancer. *Clin. Cancer Res.* **2006**, *12*, 3184–92.

35 a) Singh, S. K.; Koshkin, A. A.; Wengel, J.; Nielsen, P. *Chem. Commun.* **1998**, 455. b) Koshkin, A. A.; Singh, S. K.; Nielsen, P.;. Rajwanshi, V. K; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. *Tetrahedron* **1998**, *54*, 3607. c) Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T. *Tetrahedron Lett.* **1997**, *38*, 8735.

36 a) Petersen, M.; Bondensgaard, K.; Wengel, J.; Jacobsen, J. P. *J. Am. Chem. Soc.* **2002**, *124*, 5974. b) Nielsen, K. E.; Rasmussen, J.; Kumar, R.; Wengel, J.; Jacobsen, J. P.; Petersen, M. *Bioconjugate Chem.* **2004**, *15*, 449.

37 a) Koshkin, A. A.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V. K.; Singh, S. K. ; Wengel, J. *J. Am. Chem. Soc.***1998**, *120*, 13252. b) Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.-i.; Morio, K.-i.; Doi, T.; Imanishi, T. *Tetrahedron Lett.* **1998**, *39*, 5401. c) Wengel, J. *Acc. Chem. Res.* **1999**, *32*, 301.

a) Vester, B.; Wengel, J. *Biochemistry* **2004**, *43*, 13233. b) Jepsen, J. S.; Sørensen, M. D.; Wengel, J. *Oligonucleotides* **2004**, *14*, 130.

39 Abdur Rahman, S. M.; Imanishi, T.; Obika, S. Synthesis of several types of briged nucleic acids. *Chem. Lett.* **2009**, *38*, 512-517.

39 a) Obika, S.; Onoda, M,; Andoh, J.; Imanishi, T.; Morita, K.; Koizumi, M. *Chem. Commun.* **2001**, 1992. b) Obika, S.; Rahman, A. M. A.; Song, B.; Onada. M.; Koizumi, M.; Morita, K.; Iminishi, T. *Bioorg. Med. Chem.* **2008**, *16*, 9230.

41 Obika, S.; Nakagawa, O.; Hiroto, A.; Hari, Y.; Imanishi, T. *Chem. Commun.* **2003**, 2202.

42 Singh, S. K.; Kumar, R.; Wengel, J. *J. Org. Chem.* **1998**, *63*, 10035.

43 Morita, K.; Takagi, M.; Hasegawa, C.; Kaneko, M.; Tsutsumi, S.; Sone, J.; Ishikawa, T.; Imanishi, T.; Koizumi, M. *Bioorg. Med. Chem.* **2003**, *11*, 2211.

44 Koizumi, M.; Morita, K.; Daigo, M.; Tsutsumi, S.; Abe, K.; Obika, S.; Imanishi, T. *Nucleic . Acids Res.* **2003**, *31*, 3267.

45 Varghese, O. P.; Barman, J. ; Pathmasiri, W.; Plashkevych ,O.; Honcharenko, D.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2006**, *128*, 15173.

46 a) Rahman, S. M. A.; Seki, S.; Obika, S.; Haitani S.; Miyashita, K.; Imanishi, T.; *Angew. Chem., Int. Ed.* **2007**, *46*, 4306. b) Miyashita, K.; Rahman, S. M. A.; Seki, S.; Obika, S.; Imanishi, T. *Chem. Commun.* **2007**, 3765. c) Rahman, S. M. A.; Seki, S.; Utsuki, K.; Obika, S.; Miyashita, K.; Imanishi, T. *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 1625. d) Rahman, S. M. A.; Seki, S.; Obika, S.; Yoshikawa, H.; Miyashita, K.; Imanishi, T. *J. Am. Chem. Soc.* **2008**, *130*, 4886.

47 a) Hari, Y.; Obika, S.; Ohnishi, R.; Eguchi, K.; Osaki, T.; Ohishi, H.; Imanishi, T. *Bioorg. Med. Chem.* **2006**, *14*, 1029. b) Mitsuoka, Y.; Kodama, T.; Ohnishi, R.; Hari, Y.; Imanishi, T.; Obika, S. *Nucl. Acids Res.* **2009**, *37*, 1225.

48 Obika, S.; Sekiguchi, M.; Somjing, R.; Imanishi, T. *Angew. Chem., Int. Ed.* **2005**, *44*, 1944.

49 a) Obika, S.; Morio, K.; Nanbu, D.; Imanishi, T. *Chem. Commun.* **1997**, 1643. b) Obika, S.; Morio, K.; Nanbu, D.; Hari, Y.; Itoh, H.; Imanishi, T. *Tetrahedron* **2002**, *58*, 3039.

50 a) Sekiguchi, M.; Obika, S.; Harada, Y.; Osaki, T.; Somjing, R.; Mitsuoka, Y.; Shibata, N.; Masaki, M.; Imanishi, *J. Org. Chem.* **2006**, *71*, 1306. b) Osaki, T.; Obika, S.; Harada, Y.; Mitsuoka, Y.; Sugaya, K.; Sekiguchi, M.; Roongjang, S.; Imanishi, T. *Tetrahedron* **2007**, *63*, 8977.

51 Kodama, T.; Matsuo, C.; Ori, H.; Miyoshi, T.; Obika, S.; Miyashita, K.; Imanishi, T. *Tetrahedron* **2009**, *65*, 2116.

52 Halford, M. H.; Jones, M. S: Synthetic analogues of polynucleotides. *Nature* 1968, *217*, 638.

53 De Mesmaeker, A.; Lebreton, J.; Waldner, A.; Fritsch, V.; Wolf, R. M. Replacement of the phosphodiester linkage in oligonucleotides: comparison of two structural amide isomers. *Bioorg. Med. Chem. Lett*. **1994**, *4*, 873-878. (b) De Mesmaeker, a.; Waldner, A.; Lebreton, J.; Hoffmann, P.; Fritsch, V.; Wolf, R. M.; Freier, S. M. Amides as a new type of backbone modification in oligonucleotides. *Angew.Chem. Int. Ed.* **1994**, *33*, 226-229. (c) Idziak, I.; Just, G.; Damha, M. J.; Giannaris, P. A. Synthesis and hybridization

properties of amide-linked thymidine dimers incorporated into oligodeoxynucleotides. *Tetrahedron Lett*. **1993**, *34*, 5417-5420.

54 Debart, F.; Vasseur, J.-J.; Sanghvi, Y.S.; Cook, P. D: Synthesis and incarporation of a methyleneoxy(methylimino)-linked thymidine dimer into antisense oligonucleosides. *Bioorg.Med. Che. Lett.* 1992, *2*, 1479-1482.

55 (a) Browne, K. A.; Dempcy, R. O.; Bruice, T. C. Binding studies of cationic thymidyl-deoxyribinucleic guanidine to RNA homopolynucleotides. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7051-7055. (b) Dempcy, R. O.; Browne, K. A.; Bruice, T. C. Synthesis of the polycation thymidyl DNG, its fidelity in binding polyanionic DNA/RNA, and the stability and nature of the hybrid complexes. *J. Am. Chem. Soc*.**1995**, *117*, 6140-6141.

56 Seethe reference no. 15 from Chapter 3-Section-1

57 Babu, B. R.; Raunak; Poopeiko, N. E.; Juhl, M.; Bond, A.; Parmar, V. S.; Wengel, J. XNA(xylo Nucleic acid): Asummary and new derivatives *Eur. J. Org. Chem.* **2005**, 2297-2321.

58 Seela, F.; Heckel, M.; Rosemeyer, H. *Helv. Chim. Acta* **1996**, *79*,1451.

59 Poopeiko, N. E.; Dahl, B. M.; Wengel, J. *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1147.

60 a)Rajwanshi, V. K.; Hakansson, A. E.; Dahl, B. M.; Wengel, J. *Chem. Commun.* **1999**, 1395. b) Rajwanshi, V. K.; Håkansson, A. E.; Kumar, R.; Wengel, J. *Chem. Commun.* **1999**, 2073.

61 Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*, 1497-1501.

61 Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordon, B.; Nielsen, P. E. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature*, **1993**, *365*, 566-568.

63 Jensen, K. K.; Qrum, H.; Nielsen, P. E.; Norden, B. Hybridization kinetics of peptide nucleic acids (PNA) with DNA and RNA studied with BIAcore Technique. *Biochemistry* **1997**, *36*, 5072-5077.

64 Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sonnichsen, S. H.; Nielsen, P. E. Stability of peptide nucleic acids in human serum and cellular extracts. *Biochem. Pharmacol*. **1994**, *48*, 1310.

65 (a) Good, L.; Nielsen, P. E. Progress in developing PNA as gene targeted drugs. *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: Synthetic polyamide Nucleic Acids with Unusual Binding Properties. *Angew. Chem. Int. Ed*. **1998**, *37*, 2796-2823.

66 Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graeslund, A: Ionic effects on the stability and conformation of peptide nucleic acid (PNA) complexes. *J. Am. Chem. Soc*. **1996**, *118*, 5544-5552.

67 Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. Peptide nucleic acids (PNA). Oligonucleotide analogues with an achiral peptide backbone. *J. Am. Chem. Soc*. **1992**, *114*, 9677-9678.

68 Nielsen, P. E.; Christensen, L. Strand displacement binding of a duplex forming Homopurine PNA to a homopyrimidine duplex DNA target. *J. Am. Chem. Soc*. **1996**, *118*, 2287-2288.

69 (a) Bonham, M. A. *et al. Nucleic Acid Res*. **1995**, *23*, 1197-1203. (b) Knudsen, H.; Nielsen, P. E. Antisense properties of duplex and triplex forming PNA. *Nucleic Acid Res*. **1996**, *24*, 494-500.

70 Marsh, T. C.; Vesenka, J.; Henderson, E. G-Wires. Self-Assembly of a Telomeric Oligonucleotide, d (GGGGTTGGGG), into Large Superstructures. *Biochemistry* **1994**, *33*, 10718-10724.

71 Forman, S. L.; Fettinger, J. C.; Pieraccini, S.; Gottarelli, G.; Davis, J. T. Toward Artificial Ion Channels: A Lipophilic G-Quadruplex. *J. Am. Chem. Soc.* **2000**, *122*, 4060-4067.

72 West, R. T.; Garza, L. A.; Winchester, W. R.; Walmsley, J. A. Conformation, hydrogen bonding and aggregate formation of guanosine 5'-monophosphate and guanosine in dimethylsulfoxide *Nucleic Acids Res.* **1994**, *22*, 5128-5134.

73 Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Peptide nucleic acids (PNA). DNA analogues with a polyamide backbone. In Antisense Research and Application. Crook, S. And Lebleu, B. (eds). CRC Press, Boca Raton, **1993**, pp. 363-373.

74 (a) Praseuth, D. *et al.* Peptide Nucleic Acids directed to the promoter of the α-chain of the interleukin-2-receptor. *Biochim. Biophys. Acta*. **1997**, *1309*, 226-238. (b) Wittung, P.; Nielsen, P. E.; Norden, B. Extended DNA-recognition repertoire of PNA. *Biochemistry* **1997**, *36*, 7973-7979.

75 Lohse, J.; Dahl, O.; Nielsen, P. E. Double duplex invation by peptide nucleic acid: A general principle for recognition of double stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11804.

76 (a) Leijo, M. *et al*. Structural characterization of PNA-DNA duplexes by NMR. Evidence for DNA in a B-like conformation. *Biochemistry* **1994**, *22*, 9820-9825. (b) Eriksson, M.; Nielses, P. E. Solution structure of a peptide nucleic acid-DNA duplex. *Nat. Struct. Biol*. **1996**, *3*, 410-413.

77 Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. NMR Solution structure of a peptide nucleic acid complexed with RNA. *Science* **1994**, *265*, 777-780.

78 Bets, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. A nucleic acid triple helix formed by a peptide nucleic acid-DNA complex. *Science* **1995**, *270*, 1838-1841.

79 Kim, S. K.; Nielsen, P.E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. Right handed triple helix formed between peptide nucleic acid PNA-T8 and poly dA shown by linear and circular dichroism spectroscopy. *J. Am. Chem. Soc*. **1993**, *115*, 6477-6481.

80 Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. Crystal structure of peptide nucleic acid (PNA) duplex at 1.7 A resolution. *Nat. Struct. Biol*. **1997**, *4*, 98-101.

81 Hanvey, J. C. *et al*. Antisense and antigene properties of peptide nucleic acids. *Science*, **1992**, *258*, 1481.

82 (a) Gambacorti-Passerini, C. *et al*. *Blood* **1996**, *88*, 1411. (b) Mologni, L. *et al*. Additive antisense effects of different PNAs on the in vitro translation of the PML/RAR gene. *Nucleic Acid Res*. **1998**, *26*, 1934.

83 Nielsen, P. E.; Egholm, M.; Buchardt, O. Evidence for PNA2-DNA triplex structure upon binding to dsDNA by strand displacement. *J Mol. Recogn*. **1994**, *7*, 165-170.

84 Egholm, M. *et al*. Efficient pH independent sequence specific DNA binding by pseudoisocytosine-containing bis-PNA. *Nucleic Acid Res*. **1995**, *23*, 217-222.

85 Nielsen, P. E.; Egholm, M.; Buchardt, O. Sequence specific translation arrest by PNA bound to the template strand. *Gene* **1994**, *149*, 139-145.

86 Vickers, T. A. *et al*. Inhibition of NF-kB specific transcriptional activation by PNA strand invasion. *Nucleic Acid Res*. **1995**, *23*, 3003-3008.

87 Griffith, M. C. *et al*. Single and bispeptide nucleic acids as triplexing agents: binding and stoichiometry. *J. Am. Chem. Soc*. **1995**, *117*, 831.

88 Dueholm, K. L.; Nielsen, P. E. Chemistry, properties and applications of PNA (peptide nucleic acid). *New. J. Chem*. **1997**, *21*, 19-31.

88 Hyrup, B.; Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. Modification of the binding affinity of peptide nucleic acid (PNA). PNA with extended backbones consisting of 2-aminoethyl-β-6-γ-alanine or 3- aminopropylglycine units. *J. Chem. Soc. Chem.Commun*. **1993**, 518-519.

89 Hyrup, B.; Egholm, M.; Nielsen, P. E.; Wittung, P.; Nordon, B.; Buchardt, O. Structure-activity studies of the binding of modified peptide nucleic acids (PNA) to DNA. *J. Am. Chem.Soc*. **1994**, *116*, 7964.

90 Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. A flexible and positively charged PNA analogue with an ethylene-linker to the nucleobase: Synthesis and hybridization properties. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 1083-1088.

91. Dueholm, K.; Peterson, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. Peptide nucleic acid (PNA) with chiral backbone based on alanine. *Bioorg. Med. Chem. Lett*. **1994**, *4*, 1077-1080.

92 (a) Haaima, G.; Lohse, A.; Buchardt,O.; Nielsen, P. E. Peptide nucleic acids (PNA) containing thymine monomers derived from chiral amino acids: Hybridization and solubility properties of D-lysine PNA. *Angew. Chem. Int. Ed*. **1996**, *35*, 1939. (b) Puschl,

A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. Peptide nucleic acids (PNAs) with functional backbone. *Tetrahedron Lett*. **1998**, *39*, 4707-4710.

93 Lagriffoule, P.; Buchardt, O.; Wittung, P.; Nordan, B.; Jensen, K. K.; Nielsen, P. E. Peptide nucleic acids (PNAs) with conformationally constrained, chiral cyclohexyl derived backbone. *Chem. Eur. J*. **1997**, *3*, 912-919.

94 Maison, W.; Schlemminger, I.; Westerhoff, O.; Martens, J. *Bioorg. Med. Chem. Lett*. **1999**, *9*, 581.

95 Kortz, A. H.; Larsen, S.; Buchardt, O.; Nielsen, P. E. Retro-Inverse PNA: Structural implications for DNA binding. *Bioorg. Med. Chem. Lett*. **1998**, *6*, 1983.

96 Almorsson, O.; Bruice, T. C. Peptide nucleic acid (PNA) conformation and polymorphism in PNA-DNA and PNA-RNA hybrids. *Proc. Natl. Acad. Sci. U.S.A*. **1993**, *90*, 9542-9546.

97 Lagriffoule, P.; Egholm, M.; Nielsen, P. E.; Buchardt, O. The synthesis, co-oligomerization and hybridization of a thymine-thymine heterodimer containg PNA. *Bioorg. Med. Chem.Lett*. **1994**, *4*, 1081-1084.

98 (a) Kumar, V. A.; Ganesh, K. N. Conformationally constrained PNA analogues: Structural evolution toward DNA/RNA binding selectivity. *Acc. Chem. Res*. **2005**, *38*, 404-412. (b) Kumar, V. A. Structural preorganization of peptide nucleic acid: chiral cationic analogues with five-or six-membered ring structures. *Eur. J. Org. Chem.* **2002**, 2021-2032.

99 Sharma, N.; Ganesh, K. N. Regioselective oxidation of N-alkyl pyrrolidines to pyrrolidin-5-enes by $RuCl_3/NaIO_4$. *Tetrahedron Lett.* **2004**, *45*, 1403-1406.

100 D'Costa, M.; Kumar, V. A.; Ganesh, K. N. Synthesis of 4(S)-(N Boc-amino)- 2(S/R)-(thymin-1-yl)-pyrrolidine-N-1-acetic acid: a novel cyclic PNA with constrained flexibility. *Tetrahedron. Lett.* **2002**, *43*, 883-886.

101 Haaima, G.; Hansen, H. F.; Christensen, L.; Dahl, O.; Nielsen, P. E. Increased DNA binding and sequence discrimination of PNA oligomers containing 2,6- diaminopurine. *Nucleic Acids Res.* **1997**, *25*, 4639-4643.

102 Egholm, M.; Christensen, L.; Deuholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P.E. Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA. *Nucleic Acids Res.* **1995**, *23*, 217-222.

104 (a) Gangamani, B. P.; Kumar V. A. 2-Aminopurine peptide nucleic acids (2-*ap*PNA): intrinsic fluorescent PNA analogues for probing PNA-DNA interaction dynamics. *JCS Chem. Commun.* **1997**, 1913-1914. (b) Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. Spermine conjugated peptide nucleic acids (spPNA): V and fluorescence studies of PNA-DNA hybrids with improved stability. *Biochem. Biophys. Res. Commun.* **1997**, *240*, 778-782.

104 Eldrup, A. B.; Dahl, O.; Nielsen, P. E. A novel peptide nucleic acid monomer for recognition of thymine in triple-helix structures *J. Am. Chem. Soc.* 1997, *119*, 11116-11117.

105 (a) Wojciechowski. F.; Hudson. R. H. E. Nucleobase modifications in peptide nucleic acids. *Cur. Top. Med. Chem.* **2007**, *7,* 667-679. (b) Bajor, Z.; Sagi, G.; Tegyey, Z.; Kraicsovits, F. PNA-DNA Chimeras containing 5-alkynyl- yrimidine PNA units. Synthesis, binding properties, and enzymatic stability. *Nucleosides & Nucleotides*, **2003**, *22*, 1963-1983. (d) Hudson, R. H. E.; Li, G.; Tse, J. The use of Sonogashira coupling for the synthesis of modified uracil peptide nucleic acid. *Tetrahedron Lett*. **2002**, *43*, 1381-1386.

107 (a) Rajeev, K. G.; Maier, M. A.; Lesnik, E. A.; Manoharan, M. High-Affinity Peptide Nucleic Acid Oligomers Containing Tricyclic Cytosine Analogues. *Org. Lett.* **2002**, *4*, 4395-4398. (b) Ausin, C.; Ortega, J. A.; Robles, J.; Grandas, A.; Pedroso, E. Synthesis of Amino- and Guanidino-G-Clamp PNA Monomers. *Org. Lett.* **2002**, *4*, 4073-4075. (b) Christensen, L.; Hansen, H. F.; Koch, T.; Nielsen, P. E. Inhibition of PNA triplex formation by N4-benzoylated cytosine. *Nucleic Acids Res.* **1998**, *26*, 2735-2739.

108 Challa, H.; Styers, M. L.; Woski, S. A. Nitroazole universal bases in peptide nucleic acids. *Org. Lett.* **1999**, *1*, 1639-1641.

109 Vysabhattar, R.; Ganesh, K. N. Cyanuryl peptide nucleic acid: Synthesis and DNA complexation properties. *Tetrahedron Lett*. **2008**, *49*, 1314-1318.

110 Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W: PNA. Synthetic polyamide Nucleic Acids with Unusual Binding Properties. *Angew. Chem. Int. Ed*. **1998**, *37*, 2796-2823.

111 Petersen, K. H.; Buchardt, O.; Nielsen, P. E. Synthesis and oligomerization of Nδ-Boc-Nα-) thymin-1-yl-acetyl)-ornithine. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 793-796.

112 Bergamann, F.; Bannwarth, W.; Tam, S. Solid phase synthesis of directly linked PNADNA-hybrids. *Tetrahedron Lett*. **1995**, *36*, 6823.

113 Van der Laan, A. C. *et al*. A convenient automated synthesis of PNA-(5')-DNA (3')-PNA chimera. *Tetrahedron Lett*. **1997**, *38*, 2249-2252.

113 Knudsen, H. PhD. Thesis university of Cophenhagen, **1997**: T. Ljungstrom, H.; Knudsen, H.; Nielsen, P. E.

115 (a) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. The third helix of the Antennapedia Homeodomain translocates through biological membranes. *J. Biol. Chem*. **1994**, *269*,10444-10450. (b) Richard, J. P. *et al.* Cell-penetrating Peptides: A reevaluation of the mechanism of cellular uptake. *J. Biol. Chem*. **2003**, *278*, 585-590.

116 (a) Pooga, M. *et al*. *Nature Biotechnol*. **1998**, *16*, 857. (b) de Koning, M. C. *et al.* Synthetic developments towards PNA-peptide conjugates. *Curr. Open. Chem. Biol*. **2003**, *7*, 734-740. (c) Mier, W. *et al*: Peptide-PNA conjugate: Targeted transport of antisense therapeutics in to tumors. *Angew. Chem. Int. Ed*. **2003**, *42*, 1968-1971.

117 Zhou, P. *et al*. Novel binding and efficient cellular uptake of guanidine-based peptide nucleic acids (GPNA). *J. Am. Chem. Soc*. **2003**, *125*, 6878-6879.

118 Weiler, J. *et al*. Hybridization based DNA screening on Peptide nucleic acid (PNA) oligomer arrays. *Nucleic Acid Res*. **1997**, *25*, 2792-2799.

119 Demidov, V. Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Nielsen, P. E. Sequence selective double strand DNA cleavage by PNA targeting using nuclease S1. *Nucleic Acid Res*. **1993**, *21*, 2103-2107.

120 Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence specific inhibition of restriction enzyme cleavage by PNA. *Nucleic Acid Res*. **1993**, *21*, 197-200.

121. Strobel, S. A.; Dervan, P. B. Single-site enzymatic cleavage of yeast genomic DNA mediated by triplex formation. *Nature*, **1991**, *350*, 172-174.

121 Ornstein, R. L.; Macelroy, R. D. Nucleic acid constituent interactions. *Proc. Indian. Acad. Sci. Sect. B*, **1978**, *87b*, 135-145.

122 Cantor, C. R.; Schimmel, P. R. (Eds) Biophysical Chemistry part III, 1971, W. H. Freeman and Company, New Yok.

123 (a) Callis, P. R: Electronic states and luminescence of nucleic acid systems. *Ann. Rev. Phys. Chem*. **1983**, *34*, 329-357. (b) Ho, P. s.; Zhou, G.; Clark, L. B: Polarized electronic spectra of Z-DNA single crystals. *Biopolymers* **1990**, *30*, 151-163.

124 (a) Stryer, L. *Biochemistry*, 3rd ed.; New York: W. H. freeman and Company, **1988**. (b) Egli, M.; williums, L. D.; Gao, Q.; Rich, A. *Biochemistry* **1991**, *30*, 11388. (c) Calladine, C. R.; Drew, H. R. *Understanding DNA, The molecule and how it works*; Cambridge: Academic Press LTd., **1992**. (d) Beveridge, D. L,; Jorgensen, W. L. Ann. NY Acad. Sci. 1986, 482. (e) Gassner, R. V.; Frederick, C. A.; Quigley, G. J.; Rich, A.; Wang, A. H.-J. J. The molecular structure of the left handed Z-DNA double helix at 1.0 A atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *Biol. Chem*. **1989**, *264*, 7921.

Chapter 2

(Section-1)

***N*-(pyrrolidinylethyl)glycyl based pet-PNA and its amino derivative: Synthesis of monomer units and their incorporation in aegPNA sequences**

(Section-2)

Biophysical studies of H-pet-PNA, Amino-pet-PNA and Guanidino-pet-PNA towards complementary DNA/RNA

Chapter 2

2.1 Section I: *N*-(pyrrolidinylethyl)glycyl based pet-PNA and its amino derivative : Synthesis of monomer units and their incorporation in aegPNA sequences

2.1.1 Introduction

Antisense technology has emerged as a powerful tool over last two decades for controlling gene expression where synthetic oligonucleotides are used to target disease-causing DNA/RNA (Figure 1).[1] Several synthetic oligonucleotide analogues have been developed for their applications in antisense therapeutics.[2] Peptide nucleic acid (PNA, Figure 1) is one of the promising nucleic acid mimics in which the natural sugar – phosphate backbone is completely replaced by achiral, acyclic repeating *N*-(2-aminoethyl)glycine units.[3] PNAs are uncharged DNA mimics that bind to complementary DNA/RNA sequences with high affinity and sequence specificity. In PNA, the natural nucleobases are attached to the uncharged pseudopeptide backbone *via* methylene carbonyl linkers.

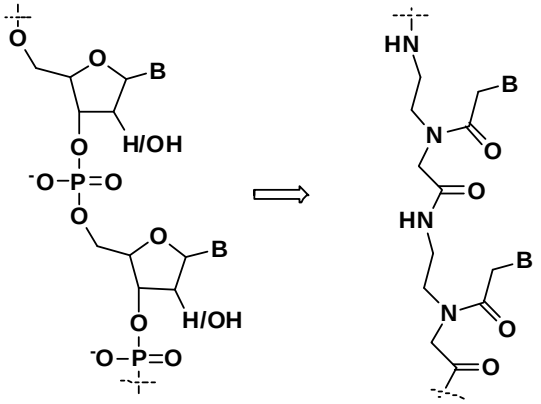


Figure 1: Basic structure of DNA/RNA and PNA

PNA hybridizes to complementary DNA/RNA sequences *via* specific Watson-Crick base complementation to form duplexes for mixed purine-pyrimidine sequences and triplexes *via* Watson-Crick and Hoogsteen base pairing for homopyrimidine/homopurine sequences.[4] The complexes of PNA with DNA/RNA sequences show thermal stabilities much higher than the corresponding DNA:DNA/ DNA:RNA complexes,[5] independent of sequence compositions. PNAs and their

analogues are also resistant to proteases and nucleases. Due to these exceptional properties, PNAs have major applications as tools in molecular biology,[6] as lead compounds for gene targeted drugs *via* antisense/antigene technology,[7] for diagnostics and biosensors,[8] and as building blocks for designing PNA supramolecular constructs.[9]

The current limitations of PNA are poor water solubility, lack of cell permeability and less discrimination in directionality between parallel and antiparallel orientation. These limitations can be systematically overcome with rationally modified PNA analogues.[10] For PNA to find applications *in vivo* as therapeutic agents some of these shortcomings need to be addressed.

2.1.2 Rationale and objectives of the present work

PNA is a very promising DNA mimic**,** but its optimum utility was stymied by poor water solubility, ambiguity in directional selectivity as well as negligible capacity to cross the cellular membrane.[11] The cellular uptake of PNA can be improved with various means like microinjection, electroporation, co-transfection with DNA, conjugation with lipophilic moieties and conjugation with positively charged peptides.[12] The conjugation of PNA with natural and unnatural positively charged peptides such as TAT and other guanidine rich peptides has improved cellular uptake of PNA.[13] Also, the antisense activity of PNA increased when PNA was conjugated to cell penetrating peptides or fatty acids. These conjugates suffer a problem of entrapment in the endosomal cavities, and to address this problem some endosomal realising agents such as chloroquine have been used in *in vitro* studies.[13c] Moreover positively charged peptides often are cytotoxic,[14] show off-target effects and may not be useful for *in-vivo* studies.[15] As an alternative to peptide conjugation, PNAs could be designed to be intrinsically positively charged and various cationic PNAs derived from positively charged amino acids like lysine[16](Figure 2) and arginine[17] (Figure 3).

PNA synthesised from D-Lysine showed stabilizing effect towards DNA and RNA as compared to L-Lysine derived PNA. The position of amino side chain in the PNA backbone also showed different structural pre-organisation, the γ–substituted PNA was shown to exert better stability towards DNA/RNA than the α–substituted PNA.[16c]

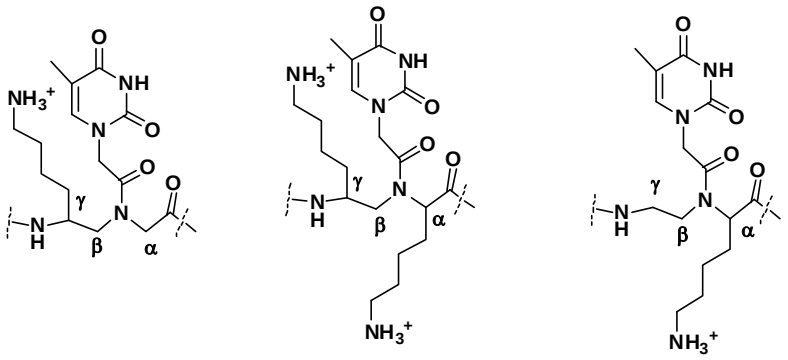


Figure 2: Lysine derived PNA analogues

GPNA reported by Ly *et al*[17].is one of the most promising modification of aegPNA which was derived from arginine. GPNAs derived from D- arginine showed increased cellular uptake with binding strength with DNA/RNA as good as control PNA. They later reported γGPNA (Figure 3), an analogue of GPNA, which has an extended side chain to carry the guanidino group at γ–position in the aegPNA backbone. The γ–substitution acted as helical director inducing structural pre-organization as studied by CD spectral studies. γ–GPNA also showed stronger binding to RNA as compared to control PNa:RNA complex.[17d]

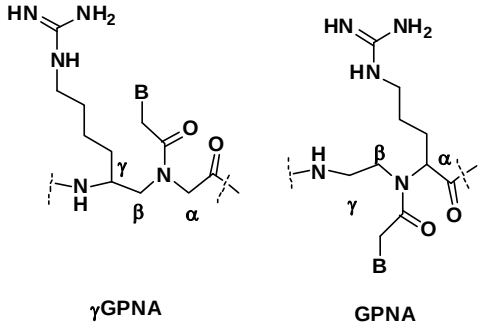


Figure 3: Structures of GPNA and γGPNA

Introduction of chirality and cyclic structures in the PNA backbone are known to reduce possible conformational states of acyclic, achiral single stranded PNAs.[9c] Various cyclic, cationic PNA analogues (Figure 4) have been reported from our lab and from other research groups.[18,10b,10c] Cationic Pyrrolidine PNA (A) a cyclic analogue of aegPNA showed very good binding strength towards both DNA/ RNA.[10c] Whereas pyrrolidine PNA (B) formed the stable complexes with DNA/ RNA, this stability and directional selectivity was dependent upon the stereochemistry in the pyrrolidinyl ring.[10c] In aminoethylprolyl PNA (*aep*-PNA), the nucleobase was directly attached to pyrrolidine ring, a single *aep*-PNA unit carrying individual nucleobases in mixed

purine/pyrimidine sequence exhibited nucleobase dependent binding efficiencies and orientation selectivies toward target DNA oligomers.[19]

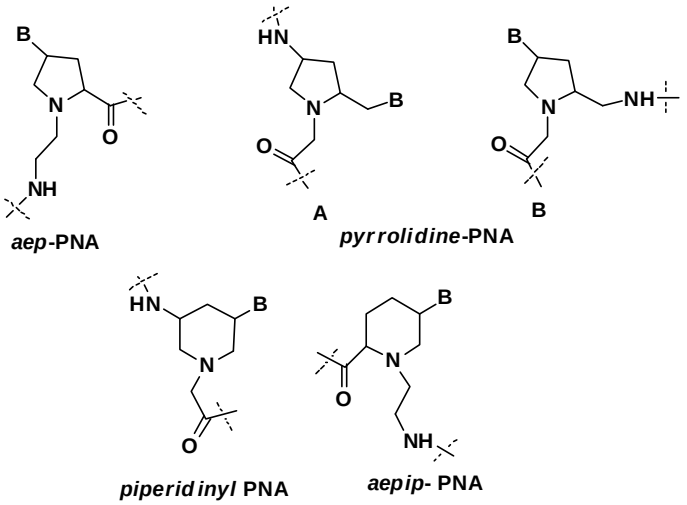


Figure 4: Conformationally constrained cationic PNA analogues

Piperidinyl and *aepip* PNA analogues contained more rigid six membered piperidine rings. In both these cases, the ring nitrogen is capable of protonation at physiological pH and may be favourable for cellular uptake.[10c]

The cyclic pyrrolidinyl/pentose based PNAs with extended backbone have been previously reported to bind sequence specifically only to RNA.[20] The five atom amide backbone replaced by seven atom amide backbone may reduce the conformational flexibility of amide relative to six atom phosphodiester backbone.[20b,21]

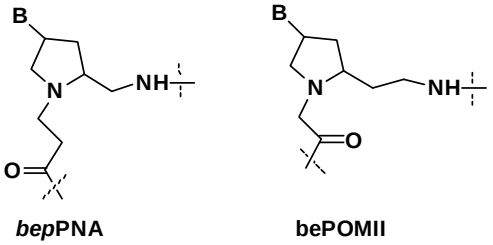


Figure 5: Backbone extended PNA analogues

POM III a pyrrolidine PNA showed kinetic binding preference to RNA over DNA. *bep*PNA (Figure 5) which was also pyrrolidine PNA with backbone extension in glycyl part of aegPNA showed preferential binding to RNA over DNA in both triplexes as well duplexes.[20c, d] bePOM PNAs (Figure 5) also showed hybridization only with complementary RNA but not with DNA.[20f]

Pyrrolidinylethyl PNA (pet-PNA) (Figure 6) a new modified PNA analogue described in this chapter has emerged by converging features of the guanidinylated γGPNA and *N*-(pyrrolidinyl-2-methyl)glycine based pmgPNA.[22] The destabilization due to constraint of pyrrolidine ring in pmgPNA may be compensated by additional flexibility of an extended backbone in pet-PNA. Also, the 4-position of pyrrolidine ring could be derivatized to amino/guanidino functionalised analogues (Figure 7).

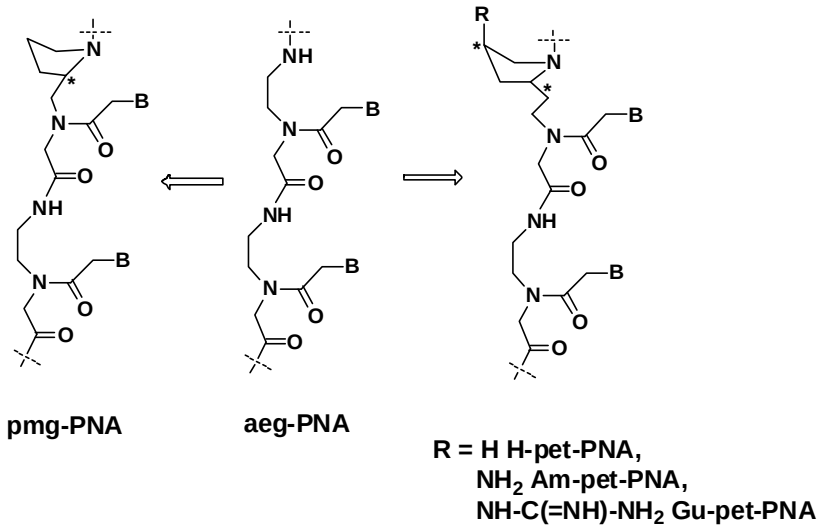


Figure 6: pyrolidinylethyl PNA (pet-PNA)

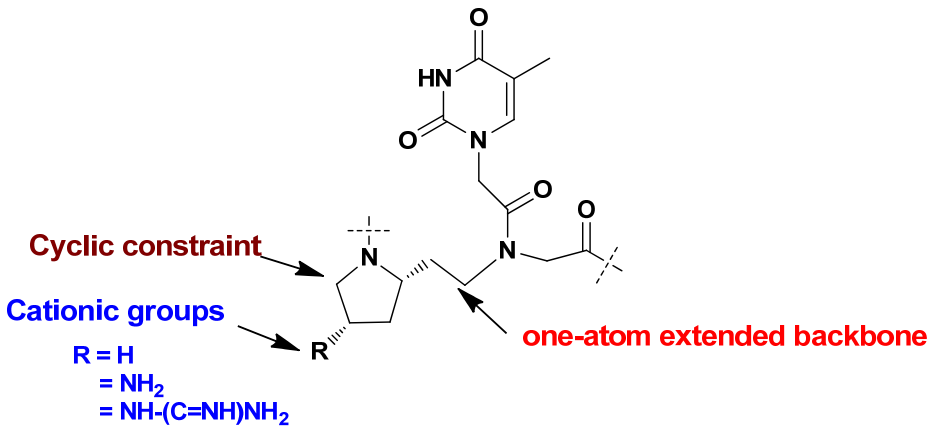


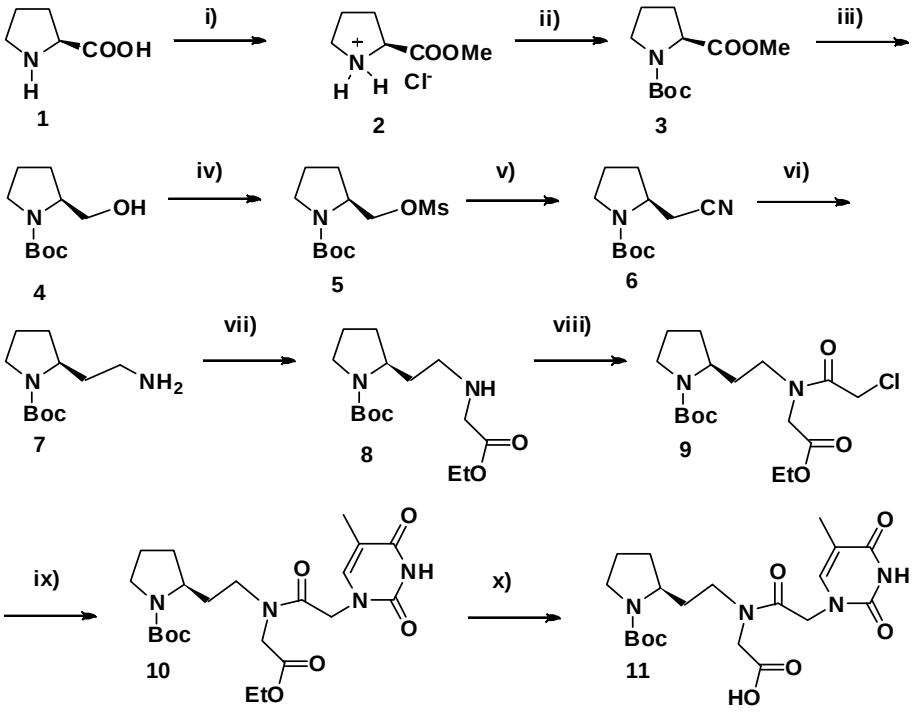
Figure 7: Design of pet-PNA

2.1.3 Synthesis of Thymine pet-PNA and Amino pet-PNA monomers, Results and Discussion

2.1.3a Synthesis of pyrrolidinylethyl thymine (H- pet-PNA) monomer

The naturally occurring amino acid L-proline was used for synthesis of H-pet-PNA monomer (**Scheme-1**). L-proline was converted to proline methyl ester by methanol and thionyl chloride in 90% yield, which on subsequent protection with Boc

gave *N*- Boc proline methyl ester derivative **3**. The *N*-Boc protected Proline ester **3** was reduced to prolinol **4** with $LiBH_4$ generated *in-situ* with $NaBH_4$/LiCl. Prolinol **4** was converted to mesylate derivative **5**, which on nucleophilic substitution with NaCN gave cyano compound **6**. Compound **6** was characterized by IR spectroscopy, which showed $-C\equiv N$ the absorption band at 2210cm^{-1} corresponding to stretching frequency**.** The ^{13}C NMR spectrum showed peak for $-C\equiv N$ carbon at δ117.6.



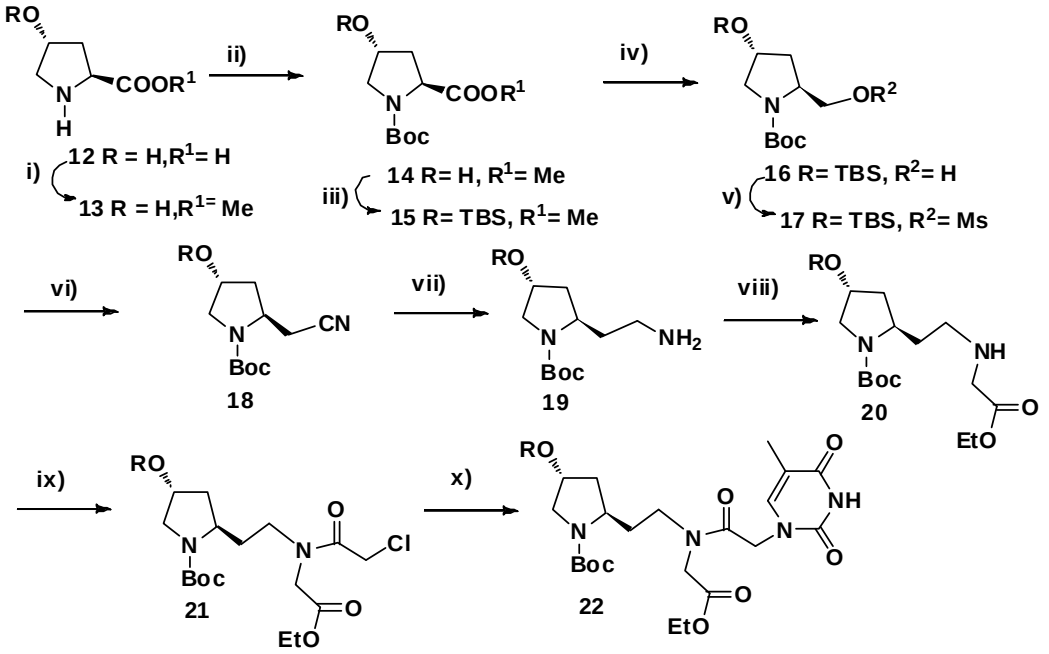
Scheme 1 Reagents and conditions **:** i) MeOH, $SOCl_2$, 90%, ii) Boc_2O, 1-4 dioxane, water, Et_3N, 88%, iii) $NaBH_4$-LiCl, THF-abs. EtOH, 89%, 4) MsCl/ py, 88%, 5) NaCN/DMSO, 60^oC, 75%, 6) H_2/Raney - Ni, MeOH, 90%,7) $BrCH_2COOEt$ /Et_3N, MeCN,70%, 8) $ClCH_2COCl$/ $NaHCO_3$,1,4-dioxane-water, 86%, 9) Thymine/K_2CO_3, dry DMF,74%, 10) 1N LiOH/MeOH, 90%

Cyano derivative **6**, on reduction with Raney-Ni/H_2 gave primary amine **7** with one extra carbon in backbone as compared to aegPNA. The primary amino group of compound **7** was monoalkylated with ethyl bromoacetate to give *sec*. amine **8** with 70% yield. Monoalkylation of primary amine was controlled by slow addition of ethyl bromoacetate and dilution of the reaction mixture with acetonitrile. This compound **8** was acylated with chloroacetyl chloride to give compound **9**. The replacement of chloro group in **9** by thymine nucleophile in presence of K_2CO_3 gave thymine *N*-1 alkylated product **10**. Chloro derivative **9** can also be used for synthesizing other natural or unnatural nucleobases containing derivatives.[23] [*N*-(*N*-Bocpyrrolinyl-2-ethyl)-*N*-(*N*-1-thyminylacetyl)]glycine was synthesized by hydrolysing ester group in **10** with 1N LiOH and methanol to give H-pet thymine monomer **11** which was further used in

solid phase synthesis of PNA. All the compounds were adequately analysed by ^{1}H, ^{13}C NMR and mass spectrometry.

2.1.3b Synthesis of Aminopyrrolidinylethyl thymine (Am-pet-PNA) monomer

4*R*-Hydroxy-L-proline **12** was used as starting material for synthesis of amino pet thymine PNA monomer (**Scheme-2-1**). 4*R*-Hydroxy-L-proline **12** was converted to methyl ester **13** by treatment with $SOCl_2$ in dry methanol and the ring nitrogen was Boc protected to give compound **14**. The 4-Hydroxy group was protected as TBS ether by reaction of TBSCl and imidazole in DMF, in the latter part of scheme this hydroxyl group was converted into amino group.

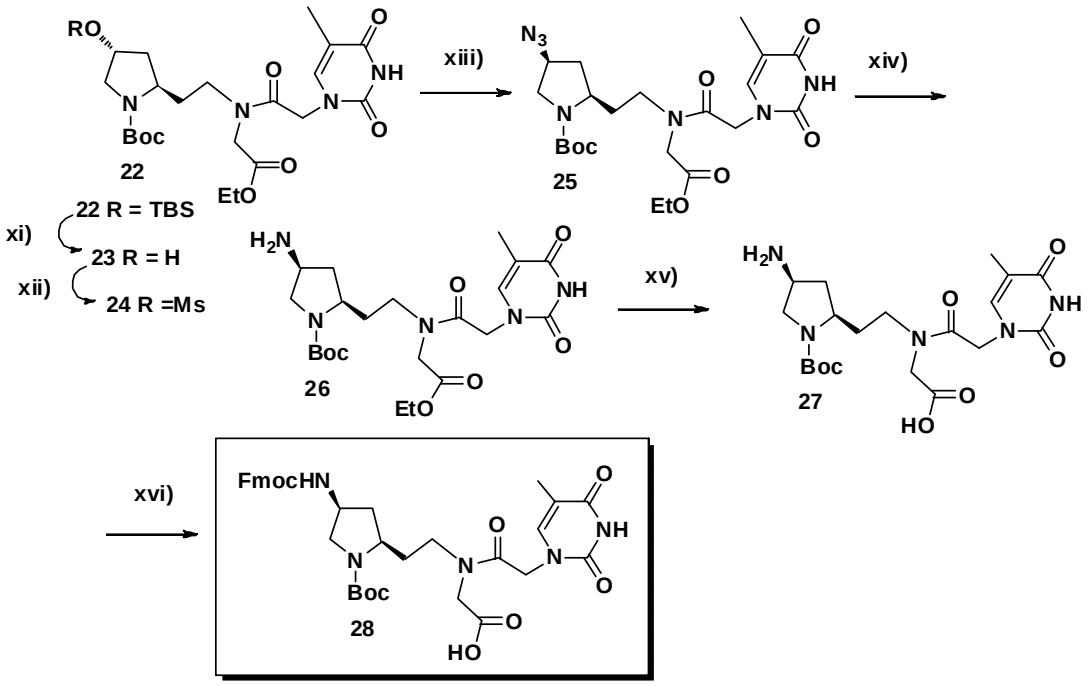


Schem-2.1: Reagents and conditions i) MeOH, $SOCl_2$, 90%,ii) Boc_2O, 1-4 dioxane, water, Et_3N, 88%, iii) TBSCl/Imidazole, DMF, 89%, iv) $NaBH_4$-LiCl, THF-abs. EtOH, 89%, v) MsCl/ py, 88%, vi) NaCN/DMSO, 60°C, 77%,vii) H_2-Raney-Ni Et_3N, MeOH, 90%, viii $BrCH_2COOEt$ /Et_3N, MeCN,68%, ix) $ClCH_2COCl/NaHCO_3$, 1,4-dioxane-water, 75%, x) Thymine/K_2CO_3, dry DMF,72%

The fully protected compound **15** was reduced to alcohol 1**6** with $NaBH_4$-LiCl in dry THF and absolute ethanol. Alcohol was converted to mesylate derivative **17**, which on nucleophilic substitution with NaCN in DMSO at 60° C gave cyano compound **18.** Cyano compound **18** was reduced to amine **19** with H_2/Raney-Ni in presence of triethylamine. Amine **19** has a one extra carbon which is corresponding to ethyl part of aegPNA, this was mono alkylated with ethylbromoacetate in acetonitrile, to give product **20** in 68% yield. Then compound **20** was acylated with chloroacetyl chloride in dioxane–water and pH was maintained at 8.0 with sat.$NaHCO_3$ solution

during addition of chloroacetyl chloride. Thymine was attached *via* nucleophilic substitution of chloro by *N*1 nitrogen of thymine gave compound **22**.

Further deprotection of TBS ether in compound **22** gave free hydroxyl compound **23** (Scheme-2-2). This hydroxyl was converted to mesylate derivative **24** and displacement of mesylate with azide nucleophile gave **25** with inversion of configuration at C4 canter of the pyrrolidine ring. Azide functionality in **25** was ascertained with IR spectroscopy which showed absorption band at 2100 cm^{-1}. Compound **25** was reduced to amine **26** with H_2/Pd-C at 40 psi hydrogen pressure. Compound **26** is amino ester, so it was used without purification for hydrolysis of ester to give amino acid **27**.

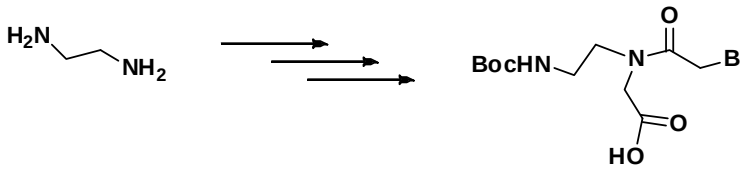


Scheme 2-2: Reagents and conditions xi) 1N TBAF in THF, 90%, xii) MsCl, py, 78%, xiii) NaN_3, DMF, 60°C, 70%, xiv) H_2/ Pd-C, MeOH, 92%, xv) 1N LiOH/ MeOH, 90%, xvi) Fmoc-succinimide/ $NaHCO_3$, 1- 4, dioxane- water, 62%

The exocylic amine in **27** was protected with orthogonal Fmoc protecting group to give amino-pet-PNA monomer **28**. The exocylic amine in monomer **28** served to provide the amino functionality in PNA and could also be conveniently converted to the highly basic guanidino functionality post synthetically on solid support to get guanidino-pet-PNA sequences.

2.1.3c Synthesis of aegPNA monomers

aegPNA monomers were synthesized by reported procedure[24] starting from ethylenediamine with required protection for nucleobases.



aeg T monomer B= Thymine
aeg C monomer B= N^4-cbzcytosine
aeg G monomer B=Guanine, N^2-cbz Guanine
aeg A monomer B=Adenine, N^6-cbzAdenine

Scheme 3: Synthesis of aegPNA monomers

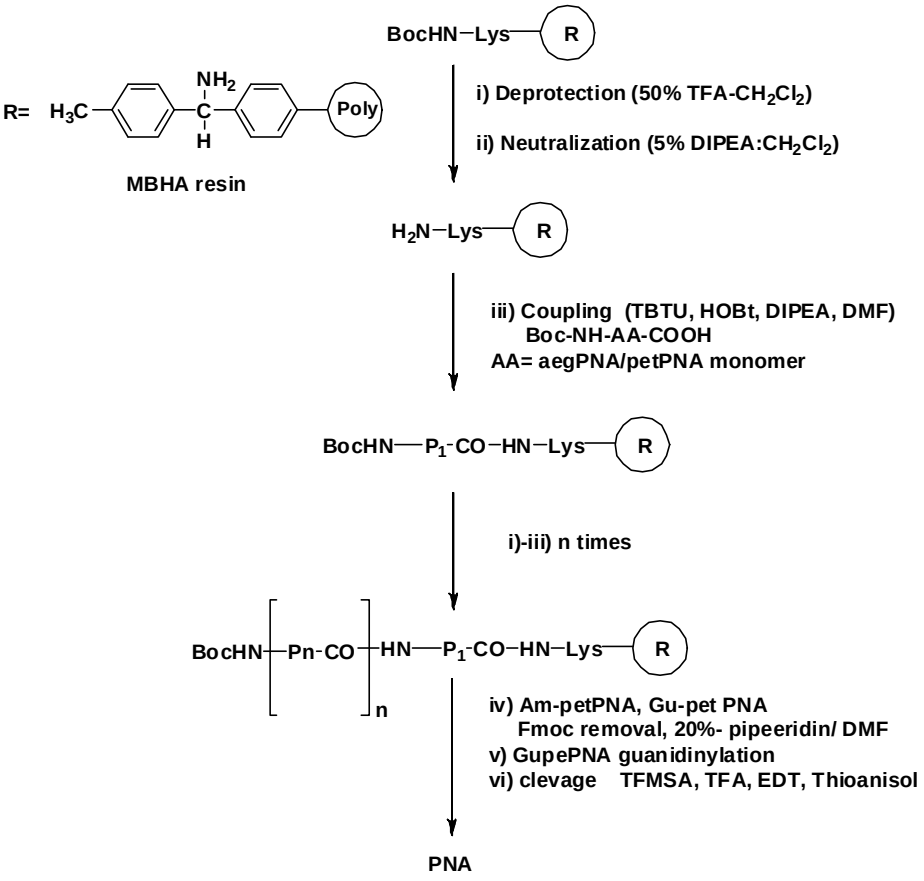
2.1.4 Solid phase PNA synthesis

2.1.4a General protocols for solid phase PNA synthesis

Solid phase peptide synthesis protocols can be applied to the synthesis of PNAs. The ease of handling and synthesis procedures have made possible the synthesis of PNAs including those incorporating a large number of analogues in an endeavour to improve its favourable binding and biological properties.

As in the case with solid phase peptide synthesis, PNA synthesis is also done conveniently from 'C' terminus to the '*N*' terminus (scheme-4). The monomeric units must have their amino functions suitably protected and their carboxylic acid functions free. The most commonly used *N*-protecting groups for solid phase peptide synthesis are the *t*-butyloxycarbonyl (Boc) and the 9-flurorenylmethoxycarbonyl (Fmoc) groups. As Boc protection was suitable for the synthetic scheme used for synthesis of modified monomers, the Boc strategy was used for solid phase PNA synthesis. The pyrrolidinyl ring nitrogen in all the pet-PNA were protected with Boc whereas in Am-pet-PNA monomers, the exocylic amino function was orthogonally protected with Fmoc and the carboxylic acid function was free to enable coupling with the resin linked monomer. The TBTU/ HOBt activation strategy[24c] was employed for the coupling reaction. MBHA resin (4-Methylbenzhydrylamine resin) was selected as solid polymeric matrix on which oligomers were synthesized. The first amino acid is linked to this matrix *via* benzyl amide linkage.[24a] This can be cleaved using standard protocol employing TFMSA-TFA to yield the C-terminal amide.

All the oligomers of the present work were synthesized manually on MBHA resin. L-lysine was selected as linker amino acid. Lysine at 'C' terminus is known to help in improving aqueous solubility of PNA. α*N*-Boc ω*N* ClCbz lysine was linked to the resin through amide bond. The free amine content on the resin was determined by the picrate assay and was found to be 1.75 mmol/g and loading was suitably lowered to approximately 0.30 mmol/g by partial acetylation of amine content using calculated amount of acetic anhydride.[25] Free -NH_2 groups on the resin available for coupling are again estimated before starting synthesis.



Scheme 4: Solid phase Synthesis of PNA

The PNA oligomers were synthesized using repetitive cycles, each comprising the following steps:

i) Deprotection of *N*-protecting Boc group using 50%TFA in CH_2Cl_2

ii) Neutralization of the TFA salt formed with diisopropylethyl amine (DIPEA) (5% DIPEA in CH_2Cl_2) to liberate the free amine.

iii) Coupling of the free amine with the free carboxylic acid group of the incoming monomer (3-4 equivalents). The coupling reaction was carried out in presence of TBTU

and HOBt in DMF or NMP as solvent and DIPEA as base. The deprotection of *N*-Boc amine and subsequent coupling reactions were monitored by the Kaiser test.[26]

2.1.4b Synthesis of H-pet, Amino-pet-PNA, Guanidino-pet-PNAs

i) Homopyrimidine Sequences

The effect of the pet unit on triplex forming ability could also be tested by synthesizing mixed thyminyl or thyminyl-cytosine oligomeric sequences. These are known to form complexes with complementary DNA oligomers in 2:1 PNA:DNA stoichiometry. The control aminoethylglycyl (aeg) PNA T_8 oligomer was first synthesized following the Boc peptide strategy outlined scheme 4. PNA sequences were synthesized incorporating the different pet (pyrrolidinylethyl) PNA thymine units at predetermined positions within the octamer. The series of octamer sequences comprising different aminoethylglycyl PNA-T and/or pyrrolidinylethyl/ amino pyrrolidinyl PNA T units is listed in Table 1. With a view to exploring the end effects of pet-PNA monomeric unit, a single pet unit was introduced at 'C' terminus (Table-1 entries 2-4). Also to explore the effect of pet-PNA unit at the middle of sequence, a single pet unit was incorporated in the middle of sequence (Table 1 entries 5-7). A fully modified oligomer with thymine H-pet-PNA units was synthesized (Table-1 entry 8).

The b_2a_2 sequence chosen for the synthesis of mixed pyrimidine oligomers, is part of *bcr-abl*[27] target gene which is responsible for chronic myelogenous leukemia (CML). The decamer sequence with four Am-pet thymine units at predefined positions (Table 2 entries 3-4) was also synthesized. The control sequence with aeg cytosine and thymine units was also synthesised. Four extra lysine (cationic amino acid) units were attached to the N-terminus of control sequence for comparison with modified sequences (Table 2, entry 2).

ii) Mixed Purine-Pyrimidine sequences

In order to study the duplex formation potential of the different pet-PNA backbone, it was imperative to synthesize mixed purine-pyrimidine sequences. Splice correction of an aberrant β-globin intron (705 site) by PNA conjugates is reported in the literature.[28] With the view of exploring pet-PNA in the splice correction assay, the central 10 mer part for target site 705 was synthesized. The pet-PNA–T monomers were incorporated into the decamer sequence in place of aegPNA-T units. H-pet-PNA unit was introduced in the centre of the sequence to see the effect of this cyclic structure in

middle of sequence (Table 3 entry2). Also H-pet-PNA and Am-pet-PNA units were also introduced at C terminus of the sequence and at the centre of the sequences to get doubly modified PNA oligomers (Table 3 entries 3-5).

2.1.4c Post synthetic conversion of amino to guanidino groups on solid phase

Guanidino functionality is known to be highly basic, due to its high pKa value (*pKa*= 12.5) as compared to amino group. It introduces a positive charge that is maintained over a wide pH range. Polyarginine and other guanidine rich peptides exhibit good cell penetrating properties.[29] Cellular uptake of PNA can be increased by either conjugating PNA to cationic peptide[14] or introduction of a cell transduction group in the PNA backbone.[17] There are two ways to obtain the guanidino function in the PNA backbone; either PNA monomer itself with protected guanidino function can be used or guanidinylation can be carried out after completion of solid phase synthesis. The post synthetic conversion of amino to guanidino group circumvents then necessicity for prior protection of guanidino groups and simplifies the synthetic procedure; hence this method was chosen for guanidinylation. There are several methods reported for conversion of amino to guanidino functionality in solution, but very few reagents (Figure.8) were used for guanidinylation on solid support.

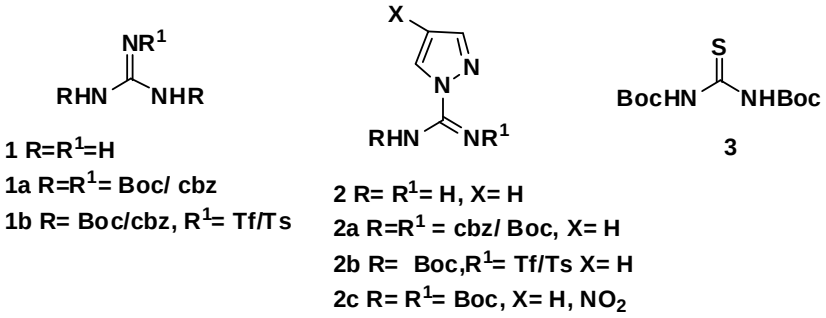
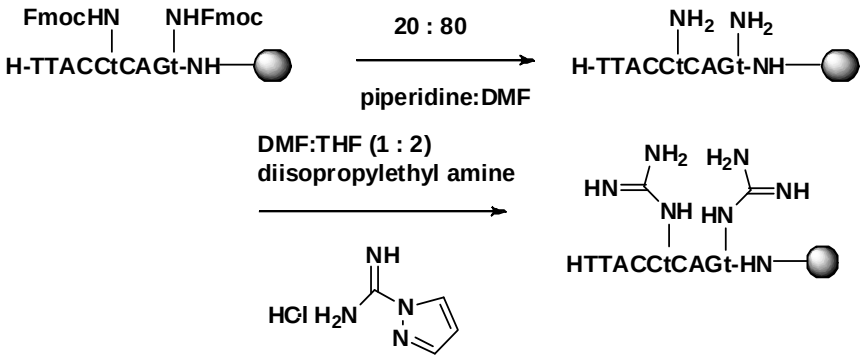


Figure 8: Guanidinylating reagents

The most commonly used guanidinylating reagents include guanidine or its derivatives (Figure 8: 1-1b) with suitable protections,[29] derivatives of pyrazole-1-carboxamidine[30] derivatives of thiourea in combination with/without $HgCl_2$/Mukaiyama's reagent.[31] As guanidinylation was carried out at the end of solid phase synthesis, unprotected Pyrazole-1-carboxamidine hydrochloride was used in present work. Scheme-5 represents the global guanidinilation of the Am-pet-PNA sequences on the solid support. Fmoc protected amino groups in Am-pet- units in the Am-pet-PNA deprotectd by 20% piperidine-DMF solution and resin was then subjected

to guanidinylation by suspending in the solution of pyrazole-1-carboxamidine hydrochloride (10 equivalents per amino group)/DIPEA in DMF-THF (1:2) mixture. Then Gu-pet-PNA was cleaved from resin under standard conditions employing TFMSA-TFA and purified by reverse phase HPLC. The conversion of amino to guanidino functionality on PNA was confirmed by MALDI-TOF mass spectrometry. The representative example of conversion of Am-pet-PNA oligomers to to Gu-pet-PNA is given in Figure **9**.



Scheme 5: Post synthetic conversion of amino to guanidino functionality

2.1.4d Cleavage of the PNA Oligomers from the Solid Support

The oligomers were cleaved from the solid support, using TFMSA in the presence of TFA ('Low, High TFMSA-TFA method"),[32] which yields oligomers with an amide group at their *C*-terminus. In case of Am-pet-PNA exocyclic amino-Fmoc protection was removed with 20% piperidine-DMF before cleavage from support and for Gu-pet-PNA the free amino groups were subsequently converted to guanidino and then oligomers were cleaved with standard cleavage conditions. A cleavage time of 2-2.5 hrs. at room temperature was found to be optimum. The side chain protecting groups for nucleobase were also cleaved during this cleavage process. After cleavage reaction, the oligomer was precipitated with dry diethylether. The oligomers listed in Table 1, 2, 3 were further subjected to RP-HPLC purification.

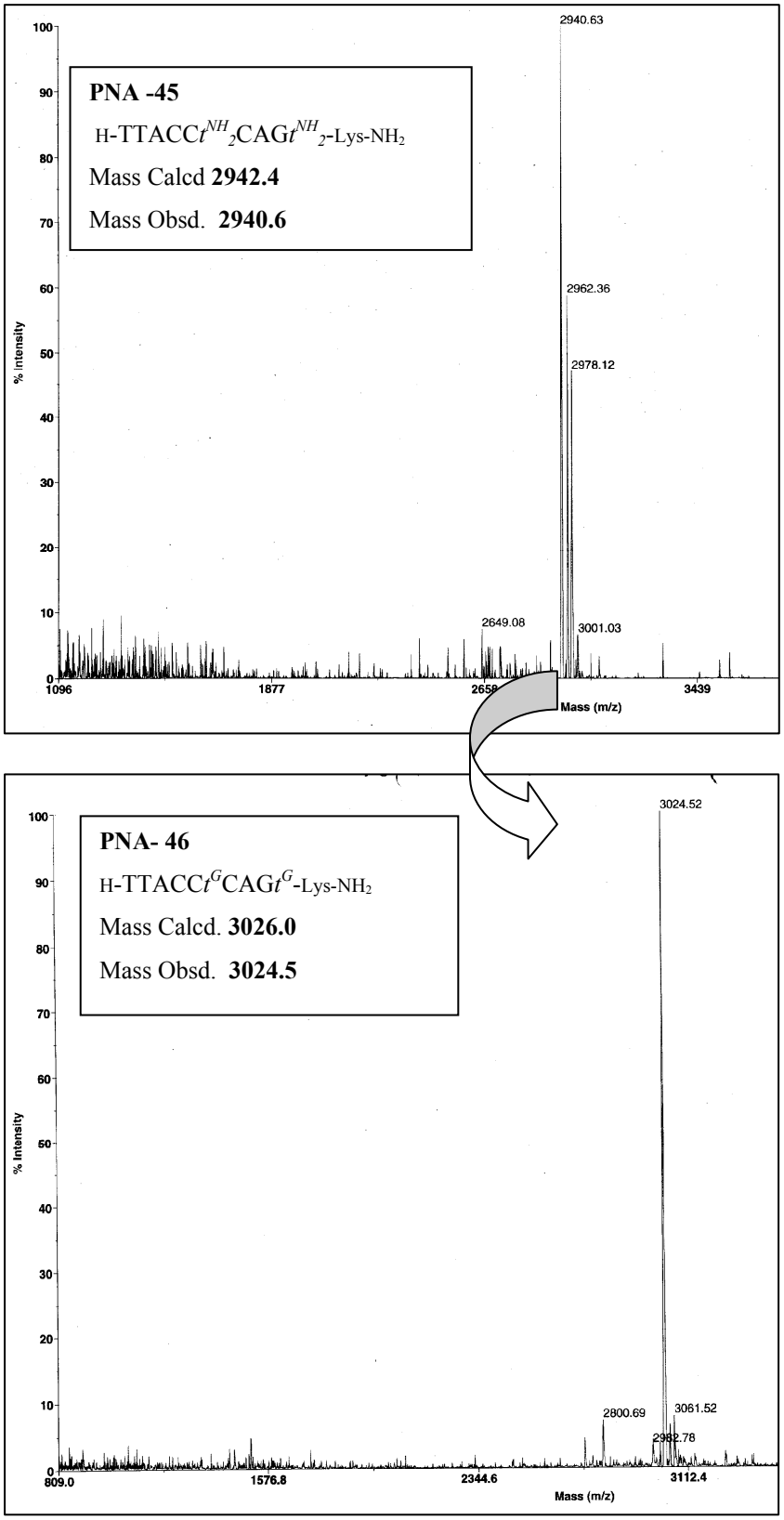


Figure 9: Representative MALDI-TOF spectra for conversion of Am-pet-PNA to Gu-pet-PNA

2.1.5 Purification and MALDI-TOF characterization of oligomers

2.15a Purification of PNA oligomers by HPLC

The purity of the oligomers was checked by analytical RP-HPLC (C18 column, CH_3CN-H_2O- 0.1% TFA system), which showed more than 65-70% purity. The gradient used for HPLC analysis is (A: 95% Water: 5% acetonitrile+0.1% TFA, B: 50% Water: 5% acetonitrile-0.1% TFA) 0-20 min 100% B. The complete conversion of amino groups to guanidino functional group was confirmed by HPLC analysis by co-injection of Amino-pet-PNA and corresponding Guanidino–pet-PNA oligomer. The gradient used for this HPLC analysis is (A: 95% Water: 5% acetonitrile-0.1% TFA, B: 50% Water: 50% acetonitrile-0.1% TFA) 0-20 min- 0%-50%B, 20-30min-50%-100% B. These were subsequently purified by reverse phase HPLC on a C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis.

Table 1: Oligothymine octamers: HPLC analysis and MALDI-TOF mass analysis

Entry	Code	Sequence	HPLC t_R (min)	Mass (calcd./obsd.)
1)	aegPNA **30**	H-TTTTTTTT-Lys-NH_2	8.90	2273.93/2274.91
		H-TTTTTTTt^R-Lys-NH_2		
2)	PNA **31**	R = H	9.48	2328.98/2329.20
3)	PNA **32**	R= NH_2	8.48	2343.99/2366.13($M+Na^+$)
4)	PNA **33**	R= NH-C(= NH) NH_2	8.66	2386.02/2386.57
		H-TTTTt^RTTT-Lys-NH_2		
5)	PNA **34**	R = H	9.21	2328.98/2329.2
6)	PNA **35**	R= NH_2	14.16*	2343.99/2343.34
7)	PNA **36**	R= NH-C(= NH)NH_2	14.58*	2385.02/2382.33
8)	PNA **37**	H-$t^H t^H t^H t^H t^H t^H t^H t^H$ -Lys-NH_2	13.27	2707.92/2706.64

R = H- H-pet-PNA , R = NH_2- Am-pet-PNA , R = NH-C(=NH)NH_2- Gu-pet-PNA

Table 2: Mixed pyrimidine oligomers: HPLC analysis and MALDI-TOF mass analysis

Entry	Code	Sequence	HPLC t_R (min)	Mass (calcd. /obsd.)
1)	aegPNA **38**	H-CTTCTTCCTT-Lys-NH_2	11.19	2746.40/2741.15
2)	aegPNA **39**	H-Lys-Lys-Lys-Lys-aha-CTTCTTCCTT-Lys-NH_2	12.00	3371.60/3366.39
		H-CTt^RCTt^RCC$t^R t^R$-Lys-NH_2		
3)	PNA **40**	R= NH_2	13.47*	3026.40/3034.50
4)	PNA **41**	R= NH-C(= NH) NH_2	14.64*	3194.35/3201.40

aha= amino hexanoiec acid R = H- H-pet-PNA , R = NH_2- Am-pet-PNA , R = NH-C(=NH)NH_2- Gu-pet-PNA

Table 3: Mixed pyrimidine-purine oligomers: HPLC analysis and MALDI-TOF mass analysis

Entry	Code	Sequence	HPLC t_R (min)	Mass (calcd./obsd.)
1)	aegPNA **42**	H-TTACCTCAGT-Lys-NH_2	9.04	2804.16/2827.07($M+Na^+$)
2)	PNA **43**	H-TTACCTCAGt^R-Lys-NH_2 R= H	9.30	2859.22/2859.38
3)	PNA **44**	H-TTACCt^RCAGt^R-Lys-NH_2 R=H	9.60	2912.27/2910.43
4)	PNA **45**	R= NH_2	12.98*	2944.29/2940.6
5)	PNA **46**	R= NH-C(= NH) NH_2	14.12*	3028.34/3024.5

R = H- H-pet-PNA, R = NH_2- Am-pet-PNA, R = NH-C(=NH)NH_2- Gu-pet-PNA

* The gradient use for HPLC analysis is (A: 95% Water: 5% acetonitrile-0.1% TFA, B: 50% Water: 50% acetonitrile-0.1% TFA) 0-20 min- 0%-50%B, 20-30min-50%-100% B

2.1.6 Synthesis of complementary oligonucleotides

The DNA oligonuclotides (Table 4) were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard β-cyanoethyl phosphoramidite chemistry.[33] The oligomers were synthesized in the 3'-5' direction on polystyrene solid support (control pore glass (CPG) resin) followed by ammonia treatment. The oligonucleotides were desalted by gel filtration; their purity ascertained by RP-HPLC on a C18 column (0.1N TEAA: ACN) to be more than 95% and were used without further purification in the biophysical studies. The RNA oligonucleotides (Table 4) were obtained commercially.

Table 4: DNA/RNA oligonucleotides used in present work

Entry	Code	Sequence (5'-3')	Type (complementary PNA)
1)	DNA **50**	GCAAAAAAAACG	complementary to PNA-**30** to PNA-**37**
2)	DNA **51**	GCAAATAAAACG	mismatch to PNA-**30** to PNA-**37**
3)	DNA **52**	AAGGAAGAAG	complementary to PNA-**38** to PNA-**41**
4)	DNA **53**	AAGGTAGAAG	mismatch to PNA-**38** to PNA-**41**
5)	DNA **54**	ACTGAGGTAA	complementary to PNA-**42** to PNA-**46**
6)	DNA **55**	ACTGTGGTAA	mismatch to PNA-**42** to PNA-**46**
7)	DNA **56**	AATGGAGTCA	parallel to PNA-**42** to PNA-**46**
8)	RNA **60**	GCAAAA AAAACG	complementary to PNA-**30** to PNA-**37**
9)	RNA **61**	GCAAAUAAAACG	mismatch to PNA-**30** to PNA-**37**
10)	RNA **62**	AAGGAAGAAG	complementary to PNA-**38** to PNA-**41**
11)	RNA **63**	AAGGACGAAG	mismatch to PNA-**38** to PNA-**41**
12)	RNA **64**	UGUAACUGAGGUAAGAGG	complementary to PNA-**42** to PNA-**46**
13)	RNA **65**	UGUAACUGCGGUAAGAGG	mismatch to PNA-**42** to PNA-**46**

2.1.7 Summary

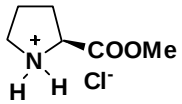
- H-pet-PNA and Amino-pet-PNA thymine monomers have been synthesized and incorporated into aeg PNA sequences.
- The conversion of Amino-pet-PNA sequences to Guanidino-pet-PNA sequences was achieved by global guanidination on solid support.

2.1.8 Experimental

General: The chemicals used were of laboratory or analytical grade. All solvents used were purified according to the literature procedures. Reactions were monitored by TLC. Usual reaction work up involved sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulphate and evaporation of the solvent under vacuum. Melting points of samples were determined in open capillary tubes using Buchi Melting point B-540 apparatus and are uncorrected. IR spectra were recorded on an infrared Fourier Transform spectrophotometer using nujol, chloroform or neat. Column chromatographic separations were performed using silica gel 60-120 mesh (Merck) or 200- 400 mesh (Merck) and using the solvent systems EtOAc/Pet ether or MeOH/DCM. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). TLCs were run in either petroleum ether with appropriate quantity of ethyl acetate or dichloromethane with an appropriate quantity of methanol for most of the compounds. TLCs were visualized with UV light and iodine spray and/or by ninhydrin treatment then heating. ^{1}H and ^{13}C NMR spectra were obtained using Bruker AC-200, AC-400 or AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^{1}H. The optical rotation values were measured on Bellingham-Stanley Ltd, ADP220 polarimeter. Mass spectra were obtained either by LCMS techniques. Oligomers were characterized by RP HPLC (Varian /Waters) C18 column and MALDI-TOF mass spectrometry. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems) MALDI-TOF instrument and the matrices used for analysis were CHCA (α-Cyano-4-hydroxycinnamic Acid) or THAP (2', 4', and 6'-trihydroxyacetophenone). DNA oligomers were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard β-cyanoethyl phosphoramidite chemistry.

Synthesis of proline methyl ester hydrochloride (2)

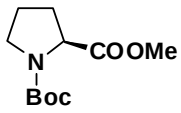
Proline **1** (4 g, 34.78 mmol) was suspended in 40 mL dry methanol and cooled to 0 °C, then $SOCl_2$ (3.0 mL, 41.73 mmol) was added slowly. Then reaction mix was then



stirred at RT for 7-8 hrs. After completion of reaction was confirmed by TLC analysis (*tert*. BuOH: H_2O: ACOH-(9:1:1), methanol was removed under reduced pressure, the residue was dried and desiccated. This compound (5.1 g, 90%) was used as such for Boc protection.

Synthesis of *N*-Boc proline methyl ester (3)

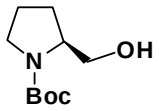
Compound **2** (5 g, 30.21 mmol) was dissolved in 50 mL (1:1 1, 4-dioxane: water) and cooled to 0 °C**.** Triethylamine (12.6 mL, 90.63 mmol) was added, then Boc_2O (7.9 g, 36



.25 mmol) dissolved in 10 mL 1, 4-dioxane was added slowly. Then reaction mixture was stirred for 7-8 hrs, after completion of reaction 1, 4-dioxane was removed under reduced pressure and compound was extracted with ethyl acetate. The compound was purified by column chromatography using EtoAc gradient in pet. ether. (Yield = 6.08 g, 88%).

Synthesis of *N*-Boc prolinol (4)

$NaBH_4$ (2.4g, 65.5 mmol) was suspended in 50 mL dry THF: abs. ethanol (2:1) mix and

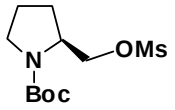


cooled to 0 °C. LiCl (2.77 g, 65.5mmol) was added in 2-3 portions. The above solution was stirred at RT for 1 hr., when the appearance of a milky solution indicates the formation of $LiBH_4$ *in situ*. To the above solution, methyl ester **3** (6 g, 26.2 mmol) dissolved in 20 mL abs. ethanol was added slowly through addition funnel. Then reaction was stirred at RT for 7-8 hrs. Then the pH of reaction mixture was adjusted to 7 by addition of sat. NH_4Cl solution, solvent was removed under vacuum and the residue was extracted with ethyl acetate. The product was purified with column chromatography (pet. ether: ethyl acetate) to give alcohol **3** (4.7 g, 89%).

^{1}H NMR ($CDCl_3$, 200Hz) δ 1.47 (s, 9H), 1.76-2.07 (m, 4H), 3.28-3.48 (m, 2H), 3.58-3.62 (m, 2H), 3.95 (m, 1H).

Synthesis of *N*-Boc-*O*-mesyl – prolinol (5)

Compound **4** (3.6 g, 18.32 mmol) was dissolved in 20 mL dry pyridine and mesyl

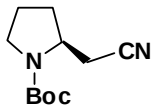


chloride (1.7 mL, 21.98 mmol) was added at 0 °C. Reaction was stirred for 2 h at RT. After completion of the reaction, pyridine was removed under reduceed pressure. Residue was dissolved in 100 mL

ethyl acetate and organic layer was washed with 50 mL 10% $NaHCO_3$ solution. Organic layer was dried over Na_2SO_4 and concentrated to give the crude product. It was used further for next reaction without purification. Crude yield = (4.5 g, 88%).

Synthesis of *N*-Boc- 2- cyanomethyl pyrrolidine (6)

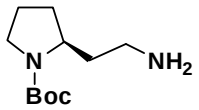
NaCN (6.3 g, 128.8 mmol) was added to the solution of the crude mesylate obtained as above (4.5 g, 16.1 mmol) in 20 mL dry DMSO, reaction was heated to 60˚C for 6 hrs. After completion of reaction DMSO was removed under reduced presssure and residue was dissolved in 100 mL ethyl acetate. Organic layer was washed with 50 mL water followed by a wash with 20 mL brine. Organic layer was concentrated under reduced presssure; crude compound was purified by column chromatography using gradient ethyl acetate–petroleum ether. Yield (2.8 g, 79%).



UV λ_{max}(nm)(CH_3CN) = 197(3.2), IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019,2210,1724,1670. ^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.42 (s, 9H, BocCH_3), 1.77-2.14 (m, 4H, *H*3,3', *H*4, 4'), 2.54-2.77 (m, 2H, C$\underline{H_2}$CN), 3.35-3.38 (m, 2H, *H*5, 5'), 3.94-4.05 (m, 1H, *H*2). ^{13}CNMR (50 MHz, $CDCl_3$) δ = 21.8, 23.3 (*C*3, *C*4), 28.0 (Boc ($\underline{C}H_3)_3$), 30.9 ($\underline{C}H_2$CN), 46.3 (*C*5), 53.4 (*C*2), 79.5 (Boc$\underline{C}(CH_3)_3$), 117.6 ($\underline{C}$N), 154.1 (Boc $\underline{C}$O). m/z(ESI) 233.44 (M+ Na^+) ($C_{11}H_{18}N_2O_2$ reqiures 210.13).

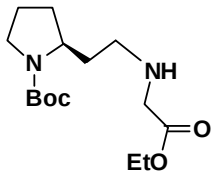
Synthesis of *N*-Boc-2*R*-aminoethyl-pyrrolidine (7)

The cyano derivative **6** obtained (2.5 g, 11.9 mmol) was dissolved in 20 mL methanol and was hydrogenated over Raney-Ni catalyst in presence of dry Et_3N (2.0 mL, 23.88 mmol) with 60 psi hydrogen pressure. After the completion of the reaction, methanol was removed under reduced presssure to give product 2.3 g (yield 90%) which was used further without purification.



Synthesis of Ethyl-[–*N*-(*N*-Bocpyrrolidin-2*R*-ethyl)-glycinate (8)

Amino compound **7** (2.2 g, 10.47 mmol) was dissolved in 25 mL dry MeCN and triethyl amine (2.6 mL, 20.94 mmol) was added with stirring. The reaction mixture was cooled to 0˚C and ethyl bromoactate (0.81 mL, 12.56 mmol) diluted with 10 mL MeCN was added slowly. The reaction was stirred for 2 h at RT. After completion of the reaction, MeCN was removed under reduced presssure; and the residue was dissolved in 150 mL ethyl acetate. Organic layer was washed with 50 mL 10 %



$NaHCO_3$ solution, dried over anhydrous anhydrous sodium sulfate and concentrated under reduced presssure to give ethyl Ethyl-[–*N*-(*N*-Boc–2*R*-pyrrolidinethyl)-glycinate **8** as crude product which was purified by column chromatography using gradient ethyl acetate–petroleum ether (Yield = 2.2 g, 70 %).

UV λ_{max}(nm)(CH_3CN)= 211(3.4), IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019, 1724, 1670.

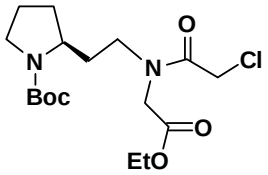
^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.24-1.31(t, 3H *J* = 7.2 Hz, $COOCH_2C\underline{H}_3$), 1.46 (s, 9H, Boc($C\underline{H}_3)_3$), 1.52-1.94 (m, 6H, *H*3, 3', *H*4, 4', $NHCH_2C\underline{H}_2$), 2.59-2.67 (m, 2H, $NHC\underline{H}_2CH_2$), 3.32-3.4 (m, 4H, *H*5, *H*5', $NHC\underline{H}_2COOEt$), 3.85 (m, 1H, *H*2), 4.2-4.24 (m, 2H, *J*=7.2 Hz, $COOC\underline{H}_2CH_3$).

^{13}C NMR (50 MHz, $CDCl_3$) δ = 14.1 ($COOCH_2\underline{C}H_3$), 28.4 (Boc($\underline{C}H_3)_3$), (30.4, 30.8), (*C*3, *C*4), 34.9 ($NCH_2\underline{C}H_2$), 46.2 ($N\underline{C}H_2CH_2$), 46.7 ($NH\underline{C}H_2COOC_2H_5$), 50.8 (*C*5), 55.1 (*C*2), 60.6 ($COO\underline{C}H_2CH_3$), 78.9 (BocC($\underline{C}H_3)_3$), 154.6 (Boc$\underline{C}$O), 172.2 ($\underline{C}OOC_2H_5$).

Mass(ESI) 301.34 ($M+H^+$), 323.34 ($M+Na^+$) ($C_{15}H_{28}N_2O_4$ requires 300.20)

Synthesis of Ethyl-[–*N*-(*N*-Boc-pyrrolidin–2*R*-ethyl)-*N*-chlorocetyl]-glycinate (9)

chloroacetyl chloride (4.3 mL, 34 mmol) was added in 2-3 portions to the solution of the compound **8** (2.04 g, 6.8 mmol) and $NaHCO_3$ (9.9 g, 68 mmol) in 20 mL (1:1 water:1,4-dioxane) at 0˚C. pH of reaction mixture was maintained at 8-9. after complete consumption of the starting material, reaction mixture was concentrated under reduced presssure. Product was extracted in 2x 50 mL dichloromethane. The organic layer was dried over anhydrous sodium sulfate, concentrated and the product was purified by column chromatography using gradient ethyl acetate–petroleum ether to get pure product. Yield product (2.2 g, 84%).



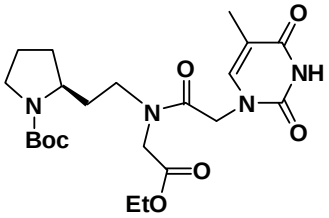
UV λ_{max} (nm) (CH_3CN) = 205(3.7), IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019, 1710, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.23-1.31 (m, 3H, $COOCH_2C\underline{H}_3$), 1.46 (s, 9H, Boc ($C\underline{H}_3)_3$), 1.76-1.99 (m, 6H, *H*3, *H*3', *H*4, *H*4', $NCH_2C\underline{H}_2$), 3.33-3.41 (m, 4H, *H*2, *H*5, $NC\underline{H}_2CH_2$), 3.79 (m, 1H, *H*5'), 4.01-4.25 (m, 6H, $COOC\underline{H}_2CH_3$, $NC\underline{H}_2COOEt$, $COC\underline{H}_2Cl$)

^{13}C NMR (50 MHz, $CDCl_3$) δ = 14.0 ($COOCH_2\underline{C}H_3$), 28.4 (Boc($\underline{C}H_3)_3$), 30.8 (*C*4), 33.7 (*C*3), 40.6 ($NCH_2\underline{C}H_2$), 46.9(*C*5), 47.7 ($N\underline{C}H_2CH_2$), 49.4 ($NH\underline{C}H_2COOC_2H_5$), 54.9 (*C*2), 61.2 ($CO\underline{C}H_2Cl$), 61.8 ($CO\underline{C}H_2CH_3$), 79.3 (Boc$\underline{C}(CH_3)_3$), 154.5 (BocCO), 166.6 ($N\underline{C}OCH_2Cl$), 168.8 ($\underline{C}OOC_2H_5$). Mass (ESI) 377.33 ($M+H^+$),

399.28(M+Na^+), 401.31 ($C_{17}H_{29}ClN_2O_5$ requires 376.17).

Synthesis of Ethyl-[–*N*-(*N*-Boc-pyrrolidin-2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate (10)



The chloroacetyl derivative (2.05 g, 5.50 mmol), K_2CO_3 (0.97 g, 6.61 mmol) and thymine (0.68 g, 6.32 mmol) was suspended in 5 mL dry DMF and stirred at RT for 6 hrs. DMF was removed under reduced presssure and compound extracted in 50 mL ethyl acetate. The crude product was purified by column chromatography using gradient dichloromethane–methanol to give the pure product **10**. (1.9 g yield = 74%).

UV λ_{max} (nm) (CH_3CN) = 206(3.6), 267(3.5),

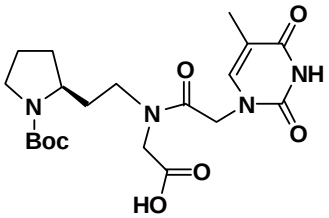
IR, υ (cm^{-1}) ($CHCl_3$) 3128, 3019, 1720, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.23-1.32 (m, 3H, $COOCH_2C\underline{H_3}$), 1.46 (s, 9H, Boc ($C\underline{H_3})_3$), 1.85-2.07 (m, 6H, *H*3, *H*3', *H*4, *H*4', $NCH_2C\underline{H_2}$), 1.91 (s, thymine $C\underline{H_3}$), 3.29-3.43 (m, 4H, *H*2, *H*5, $NC\underline{H_2}CH_2$), 3.73-3.85 (m, 1H, *H*5'), 4.08-4.63 (m, 6H, $COOC\underline{H_2}CH_3$, $NC\underline{H_2}COOEt$, $COC\underline{H_2}$ thymine), 7.03 (s, 1H *H*6 thymine).

^{13}C NMR (50 MHz, $CDCl_3$) δ = 12.1 (thymineCH_3), 13.9($COOCH_2\underline{C}H_3$), 28.3 (Boc($\underline{C}H_3)_3$), 29.9 (*C*4), 30.6 (*C*3), 33.4 ($NCH_2\underline{C}H_2$), 45.9 (*C*5), 47.5 ($N\underline{C}H_2CH_2$), 49.0 ($NH\underline{C}H_2COOC_2H_5$), 54.7 (*C*2), 61.2 ($CO\underline{C}H_2$thymine), 61.9 ($CO\underline{C}H_2CH_3$), 79.3 (Boc$\underline{C}(CH_3)_3$), 110.4 (*C*5thymine), 140.9 (*C*6thymine), 151.0 (*C*2thymine), 154.4 ((BocCO), 164.4 (*C*4thymine), 166.9 ($N\underline{C}OCH_2$thymine), 168.7 ($\underline{C}OOC_2H_5$).

Mass (ESI) 489.60 (M+Na^+) ($C_{22}H_{34}N_4O_7$ requires 466.24)

[–*N*-(*N*-Boc-pyrrolidin-2*R* -ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine (11)



The compound **10** (1.6 g, 3.41 mmol) was dissolved in 2 mL methanol, to which 2 mL 1N LiOH solution in water was added. Reaction was complete in 30 min. Methanol was removed, aqueous layer was neutralized with Dowex H^+ resin and resin was separated by filtration. The aqueous solution was washed with ethyl acetate, concentrated to give thymine monomer **11** (1.4 g yield = 88%).

UV λ_{max} (nm) (CH_3OH) = 201 (4.1), 268(3.8), IR, υ (cm^{-1}) (CH_3OH) 3126, 3022, 1720, 1670.

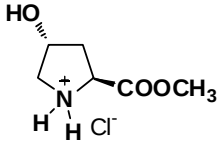
^{1}H NMR (200 MHz, DMSO d_6) δ = 1.38 (s, 9H, Boc ($C\underline{H_3})_3$), 1.62-2.01 (m, 6H, *H*3,

*H*3', *H*4, *H*4', $NCH_2C\underline{H_2}$), 1.77(s, thymine $C\underline{H_3}$), 3.23-3.35 (m, 4H, *H*2, *H*5, $NC\underline{H_2}CH_2$), 3.62-3.86 (rotamers, m, 2H, *H*5', *H*4), 3.96-4.06 (rotamers, s, 2H, $NC\underline{H_2}COOH$), 4.48-4.63 (m, rotomers, $COC\underline{H_2}$thymine) 7.31-7.43 (rotamers, d, 1H, *H*6 thymine).

^{13}C NMR (50 MHz DMSO d_6) δ = 12.1 (thymineCH_3), 28.3 (Boc($\underline{C}H_3)_3$), 30.5 (*C*4), 32.7 (*C*3), 35.0 ($NCH_2\underline{C}H_2$), 45.1(*C*5), 46.1 ($N\underline{C}H_2CH_2$), 47.7 ($NH\underline{C}H_2COOC_2H_5$), 54.8(*C*2), 59.7 ($CO\underline{C}H_2$thymine), 78.6 (Boc$\underline{C}(CH_3)_3$), 108.2 (*C*5thymine), 142.5 (*C*6thymine), 151.2 (*C*2thymine), 153.8 (BocCO), 164.7 (*C*4thymine), 167.3 ($N\underline{C}OCH_2$thymine), 170.6 ($\underline{C}OOH$). $[\alpha]^{20}_D$ = -34.07 (c =1.2.MeOH).

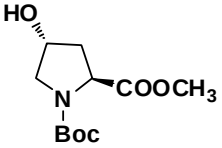
Mass obsd. 445.08 ($M+Li^+$), 461.06 ($M+Na^+$) ($C_{20}H_{30}N_4O_7$ requires 438.22).

Synthesis of *trans*-4 -hydroxy L proline methyl ester hydrochloride (13)



trans-4-hydroxy L Proline **12** (5.6 g, 42.81 mmol) was suspended in 50 mL dry methanol and cooled to 0 °C , then $SOCl_2$ (3.7 mL, 51.37 mmol) was added slowly. Reaction mix was then stirred at RT for 7-8 hrs. After completion of reaction and confirmed by TLC analysis (*tert*.BuOH:H_2O:ACOH-(9:1:1), methanol. Methanol was removed under reduced pressure, residue was dried and desiccated. This compound (7.0 g, 90%) was used as for Boc protection.

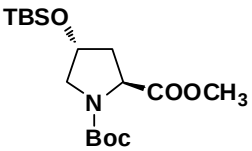
Synthesis of trans -4-hydroxy- *N*-Boc L proline methyl ester (14)



Compound **13** (6.8 g, 37.731 mmol) was dissolved in 60 mL (1:1 1-4 dioxane: water) and cooled to 0 °C**.** Triethyl amine (15.75 mL, 113.21 mmol) was added, then Boc_2O (9.87 g, 45.27 mmol) dissolved in 20 mL 1-4 dioxane added slowly. Reaction mixture was stirred for 7-8 hrs, after completion of reaction 1, 4-dioxane was removed under reduced pressure and compound was extracted with ethyl acetate. Compound was purified by column chromatography with pet. ether and ethyl acetate gradiant. (Yield = 8.2 g, 88%).

Synthesis of *N*-Boc- 4*R*-*O*TBS proline methyl ester (15)

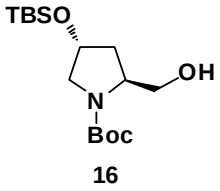
Compound **14** (8.15 g, 22.74 mmol) and imidazole (3.4 g, 50.03mmol) was taken in 15 mL DMF and cooled to 0 °C. TBSCl (3.77g, 25.01mmol) was added to this in 2-3 portions and reaction was stirred at RT for 8 hrs. After completion of reaction, reaction was work it up with



ethyl acetate and organic layer was washed with water and brine. Organic layer was concentrated under vacuum and product was purified with column chromatography (pet. ether: ethyl acetate) yield = 10.7 g, 89 %

^{1}H NMR ($CDCl_3$, 200Hz) δ 0.06(s, 6H), 0.87 (s, 9H), 1.41-1.46 (s, 9H), 1.60 (m, 2H), 1.94-2.25 (m, 2H), 3.28-3.44 (m, 1H), 3.59-3.63 (m, 1H), 3.73 (d, 3H), 4.30-4.41(m, 2H).

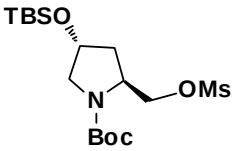
Synthesis of *N*-Boc- 4*R*-*O*TBS prolinol (16)



16

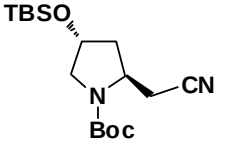
$NaBH_4$ (2.7g, 73.23 mmol) was suspended in 70 mL dry THF: abs. ethanol (2:1) mix and cooled to 0 °C. LiCl (3.1 g, 73.23mmol) was added in 2-3 portions. The above solution was stirred at RT for 1 hr.and appearance of milky solution indicates the formation of $LiBH_4$ *in situ*. To above solution methyl ester (10.5 g, 29.29 mmol) dissolved in 30 mL abs. ethanol was added slowly through addition funnel. Reaction was stirred at RT for 7-8 hrs. Then pH of reaction mixture was adjusted to 7 by addition of sat. NH_4Cl solution, solvent was removed under vacuum and residue was extracted with ethyl acetate. Product was purified with column chromatography (pet. ether: ethyl acetate) to give alcohol **16** (8.7 g, 89%).

^{1}H NMR ($CDCl_3$, 200Hz) δ 0.06 (s, 6H), 0.87(s, 9H), 1.47 (s, 9H), 1.92-2.01 (m, 2H), 3.55-3.69 (m, 5H), 4.12- 4.28(m, 2H), 4.92-4.95(m, 1H).



Synthesis of *N*-Boc- 4*R*-*O*TBS-2*S*-*O*-mesyl – prolinol (17)

Compound **16** (8.6 g, 26.13 mmol) was dissolved in 30 mL dry pyridine and mesyl chloride (2.5 mL,31.36 mmol) was added at 0 °C. Reaction was stirred for 2 h at RT. After completion of the reaction, pyridine was removed under reduced presssure. Residue was dissolved in 100 mL ethyl acetate and organic layer was washed with 50 mL 10% $NaHCO_3$ solution. Organic layer was dried over Na_2SO_4 and concentrated to give the crude product. It was used further for next reaction without purification. Crude yield: (8.7 g, 85 %)



Synthesis of *N*-Boc-4*R*-*O*-TBS -2*S*- cyanomethyl-pyrrolidine (18)

NaCN (8.7 g, 177.8 mmol) was added to the solution of the crude mesylate obtained as above (8.5 g, 22.2 mmol) in 25 mL dry DMSO, reaction was heated to 60˚C for 6 hrs. After completion of reaction DMSO was removed

under reduced presssure and residue was dissolved in 200 mL ethyl acetate. Organic layer was washed with 50 mL water followed by a wash with 20 mL brine. Organic layer was concentrated under reduced presssure; crude compound was purified by column chromatography using gradient ethyl acetate–petroleum ether. Yield (5.6 g, 77%).

UV λ_{max}(nm)(CH_3CN)= 194(3.6), IR, $\upsilon(cm^{-1})$ ($CHCl_3$) 3128, 3019,2208,1724,1670.

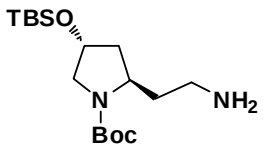
^{1}H NMR (200 MHz, $CDCl_3$) δ = 0.07 (s, 6H, Si-$(CH_3)_2$), 0.86 (s, 9H, Si-C$(CH_3)_3$), 1.46 (s, 9H, Boc $(CH_3)_3$), 1.94-2.18 (m, 2H, *H*3,3'), 2.62-3.11 (m, 2H, CH_2CN), 3.42 (m, 2H, *H*5,5'), 4.12-4.39 (m, 2H, *H*4, *H*2).

^{13}CNMR (50 MHz, $CDCl_3$) δ = -5.0 (Si(CH_3)$_2$), 17.7(SiC$(CH_3)_3$), 25.5 (SiC(CH_3)$_3$), 28.2 (BocCH_3), 39.8 (*C*3),40.8 (CH_2CN), 52.6 (*C*2), 55.6 (*C*5), 69.6 (*C*4), 79.9 (BocC$(CH_3)_3$), 117.3 (CN), 154.5 (BocCO).

Mass (ESI) 363.35 (M+Na^+) ($C_{17}H_{32}N_2O_3Si$ requires 340.21).

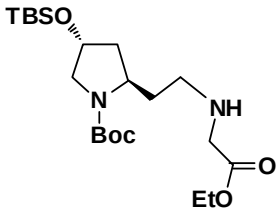
Synthesis of *N*-Boc-4*R*-*O*-TBS–2*R*-aminoethyl- pyrrolidine (19)

The cyano derivative **18** (5 g, 14.74 mmol) was dissolved in 20 mL methanol and was hydrogenated over Raney-Ni catalyst in presence of dry Et_3N (4.09 mL, 47.18 mmol) with 60 psi hydrogen pressure. After the completion of the reaction, raction mixture was filtered and methanol was removed under reduced presssure to give **19** 4.5 g (yield 90%) which was used further without purification.



Synthesis of ethyl –*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-glycinate (20)

Compound **19** (4.2 g, 12.28 mmol) was dissolved in 40 mL dry MeCN and Et_3N (3.14 mL, 24.56 mmol) was added with stirring. The reaction mixture was cooled to 0˚C and ethyl bromoactate (0.95 mL, 14.73 mmol) diluted with 10 mL MeCN added slowly. The reaction was stirred for 2 hrs. at RT. After completion of the reaction, CH_3CN was removed under reduced presssure and the residue was dissolved in 200 mL ethyl acetate. Organic layer was washed with 50 mL 10 % $NaHCO_3$ solution, dried over anhydrous sodium sulfate and concentrated under reduced presssure to give **20** as crude product which was purified by column chromatography using gradient ethyl acetate–petroleum ether. 3.6 g (yield 68 %).



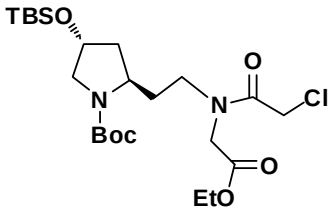
UV λ_{max}(nm)(CH_3CN)= 203(3.8), IR, $\upsilon(cm^{-1})$ ($CHCl_3$) 3128, 3019,2110,1724,1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 0.05 (s, 6H, Si-$(CH_3)_2$), 0.86 (s, 9H, Si-C$(CH_3)_3$), 1.24-1.31 (m, 3H, $COOCH_2C\underline{H_3}$), 1.46 (s, 9H, Boc $(CH_3)_3$), 1.56-2.01 (m, 4H, *H*3, 3', $NCH_2C\underline{H_2}$), 2.57-2.64 (m, 2H, $NC\underline{H_2}CH_2$), 3.34-3.39 (m, 4H, *H*2, *H*5, $NC\underline{H_2}COOEt$), 3.84-3.94(m, 1H, H5'), 4.13-4.24 (q, J, $COOC\underline{H_2}CH_3$), 4.28-4.35 (m, 1H, *H*4).

^{13}C NMR (50 MHz, $CDCl_3$) δ = -4.9(Si($\underline{C}H_3)_2$), 14.1($COOCH_2\underline{C}H_3$), 17.8(Si$\underline{C}(CH_3)_3$), 25.7(SiC($\underline{C}H_3)_3$), 28.3(BocCH_3), 29.5 ($\underline{C}H_2CH_2NH$), 35.5 (*C*3), 40.3 ($HN\underline{C}H_2CH_2$), 46.4 (*C*5), 50.8 ($NH\underline{C}H_2COOC_2H_5$), 54.1 (*C*2), 60.6 ($COO\underline{C}H_2CH_3$), 69.7(*H*4), 79.1(Boc$\underline{C}(CH_3)_3$), 155.0(Boc$\underline{C}$O), 172.2($\underline{C}OOC_2H_5$). Mass (ESI) 431.60 (M+H) ($C_{21}H_{42}N_2O_5Si$ requires 430.28)

Synthesis of Ethyl-[–*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-*N*-chlorocetyl]-glycinate (21)

chloroacetyl chloride (5.2 mL, 40.65 mmol) was added in 2-3 portions to the solution of the compound **20** (3.5 g, 8.13 mmol) and $NaHCO_3$ (6.8 g, 46.51 mmol) in 20 mL (1:1 water:1-4 dioxane) at 0˚C, pH of reaction mixture was maintained at 8-9. After complete conversion of the starting material, reaction mixture was concentrated under reduced presssure. Product was extracted in 2x 50 mL dichloromethane. The organic layer was dried over anhydrous sodium sulfate, concentrated and the product was purified by column chromatography using gradient ethyl acetate–petroleum ether to get pure product **21**. 3.0 g (yield 75%).



UV λ_{max}(nm)(CH_3CN)= 205(3.6), IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019, 2110, 1724, 1670.

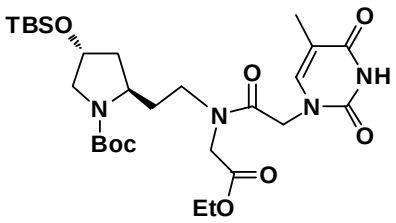
^{1}H NMR (200 MHz, $CDCl_3$) δ = 0.06 (s, 6H, Si-$(CH_3)_2$), 0.86 (s, 9H, Si-C$(CH_3)_3$), 1.24-1.27 (m, 3.0 H, $COOCH_2C\underline{H_3}$), 1.48 (s, 9H, Boc $(CH_3)_3$), 1. 56-2.14 (m, 4H, *H*3, 3', $NCH_2C\underline{H_2}$), 3.25-3.46 (m, 4H, *H*2, *H*5, $NC\underline{H_2}CH_2$), 4.00-4.33 (m, 7H, *H*4, $OC\underline{H_2}CH_3$, $NCOC\underline{H_2}T$, $NC\underline{H_2}COOEt$).

^{13}C NMR (50 MHz, $CDCl_3$) δ = -4.9 (Si($\underline{C}H_3)_2$), 14.0 ($COOCH_2\underline{C}H_3$), 17.8(Si$\underline{C}(CH_3)_3$), 25.5 (SiC($\underline{C}H_3)_3$), 28.3 (BocCH_3), 29.5 ($\underline{C}H_2CH_2N$), 33.7 (*C*3), 41.1 ($N\underline{C}H_2CH_2$), 46.5 (*C*5), 47.7 ($N\underline{C}H_2COOC_2H_5$), 53.4 (*C*2), 54.9 ($NCO\underline{C}H_2Cl$), 61.3 ($COO\underline{C}H_2CH_3$), 70.0 (*C*4), 79.5 (Boc$\underline{C}(CH_3)_3$), 155.1 (Boc$\underline{C}$O), 155.3 ($\underline{C}OCH_2Cl$), 166.6 ($\underline{C}OOC_2H_5$).

Mass (ESI) 507.59 ($M+H^+$), 509.58 ($C_{23}H_{43}ClN_2O_6Si$ requires 506.25)

Synthesis of Ethyl-[–*N*-(*N*-Boc-4*R*-*O*-TBS-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate (22)

The chloroacetyl derivative **21** (3.0 g, 6.32mmol), K_2CO_3 (0.87 g, 5.92 mmol) and



thymine (0.79 g, 6.32 mmol) were suspended in 5 mL dry DMF and stirred at RT for 6 hrs. After completion of reaction DMF was removed under reduced presssure and compound extracted in 50 mL ethyl acetate. The organic layer was dried over anhydrous sodium sulfate, concentrated to give crude product.The crude product was purified by column chromatography using gradient dichloromethane–methanol to give the pure product. 2.6 g (yield 72%).

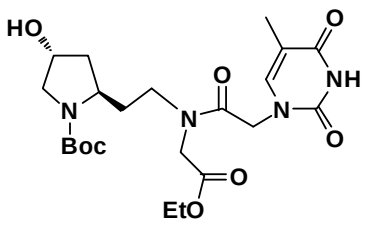
UV λ_{max}(nm)(CH_3CN)= 267 (3.9) IR, $\upsilon(cm^{-1})$ ($CHCl_3$) 3128, 3019, 1724, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) δ = 0.06 (s, 6H, Si-$(CH_3)_2$), 0.86 (s, 9H, Si-C$(CH_3)_3$), 1.26-1.45 (m, 3H, $COOCH_2C\underline{H_3}$), 1.45 (s, 9H, Boc $(CH_3)_3$), 1.66-2.05 (m, 4H, $NCH_2C\underline{H_2}$, *H*3, *H*3'), 1.92 (s, 3H, thymine-$C\underline{H_3}$), 3.25-3.41 (m, 4H, *H*2, *H*5, $NC\underline{H_2}CH_2$), 3.86-3.98(m, 1H, *H*5'), 4.07-4.62 (m, 7H, *H*4, $OC\underline{H_2}CH_3$, $NCOC\underline{H_2}T$, $NC\underline{H_2}COOEt$), 7.03 (s, 1H, *H*6 thymine).

^{13}C NMR (50 MHz, $CDCl_3$) δ = -4.9 ($Si(\underline{C}H_3)_2$), 12.2 (thymine$\underline{C}H_3$), 14.0($COOCH_2\underline{C}H_3$), 17.8($Si\underline{C}(CH_3)_3$), 25.5($SiC(\underline{C}H_3)_3$), 28.3($BocCH_3$), 29.5($\underline{C}H_2CH_2N$), 33.9(*C*3), 40.7($N\underline{C}H_2CH_2$), 45.7(*C*5), 47.5($N\underline{C}H_2COOC_2H_5$), 53.6(*C*2), 54.9($NCO\underline{C}H_2$thymine), 61.3($COO\underline{C}H_2CH_3$), 70.0(*C*4), 79.6(Boc$\underline{C}(CH_3)_3$), 110.4(*C*5thymine), 140.9(*C*6thymine), 151.1(*C*2thymine), 155.1(Boc$\underline{C}$O), 164.3($\underline{C}OCH_2$thyymine), 167.0(*C*4thymine), 168.7($\underline{C}OOC_2H_5$).

Mass (ESI) 597.68 (M+H), 619.71($M+Na^+$) ($C_{28}H_{48}N_4O_8Si$ requires 596.3241).

Synthesis of ethyl-[–*N*-(*N*-Boc-4*R*-hydroxypyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate (23)

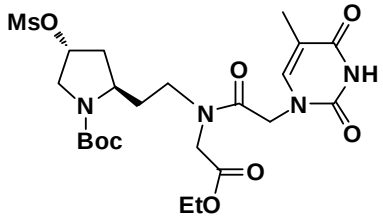


Compound **22** (2.5 g, 3.35 mmol) was dissolved in 10 mL dry THF, 1M TBAF in THF (5.0 mmol) was added. After 30 min, THF was removed under reduced presssure, product extracted in 50 mL dichloromethane, organic layer washed with 20 mL water. After

evaporation of the solvent, the crude product was purified by column chromatography using gradient methanol-dichloromethane to give the pure compound **23** yield (1.75 g, 90%)

Synthesis of ethyl-[–*N*-(*N*-Boc-4*R*-*O*-mesylpyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate (24)

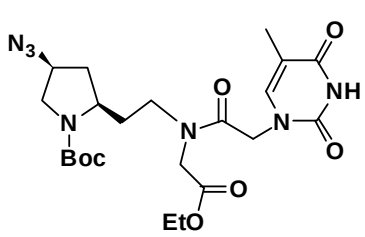
The pure compound **23** (1.7 g, 3.52 mmol) was dissolved in 20 mL dry pyridine and



mesyl chloride (0.4 mL, 5.2 mmol) was added drop wise to this solution at 0˚C while stirring. After 2 hrs., pyridine was removed under reduced presssure; residue was dissolved in 50 mL ethyl acetate. Organic layer washed with 10% $NaHCO_3$ and dried over anhydrous sodium sulfate. The orgaic layer was concentrated to give, the crude product. This was purified by column chromatography using gradient methanol-dichloromethane to give compound **24**. 1.5 g (yield 78%)

Synthesis of Ethyl-[–*N*-(*N*-Boc-4*S*-azido-pyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate (25)

The mesylate obtained in the previous step (1.4 g, 2.5 mmol) and NaN_3 (1.65 g, 25.0



mmol) were suspended in 10 mL dry DMF and stirred at 60 ºC for 6 hrs. After completion of the reaction, DMF was removed under reduced pressure and residue was dissolved in 100 mL ethyl acetate. The organic layer was washed with water, dried over anhydrous sodium sulfate, concentrated and the crude product was purified by column chromatography using gradient methanol-dichloromethane to give **25**. 0.88 g (yield 70%).

UV λ_{max}(nm)(CH_3CN)= 194(3.2), 266(3.8), IR, $\upsilon(cm^{-1})$ ($CHCl_3$) 3128, 3019,2110,1724,1670.

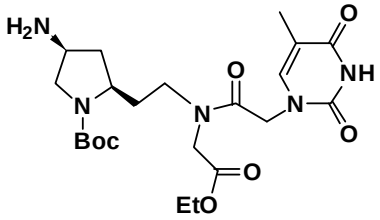
^{1}H NMR (200 MHz, $CDCl_3$) δ = 1.24-1.29 (m, 3H, $COOCH_2C\underline{H_3}$), 1.46 (s, 9H, Boc $(CH_3)_3$), 1.92(s, 3H, thymine CH_3), 1.84-2.35 (m, 4H, *H*3, 3', $NCH_2C\underline{H_2}$), 3.32-3.74 (m, 5H, *H*2, *H*5, 5', $NC\underline{H_2}CH_2$), 4.10-4.64 (m, 7H, *H*4, $COOC\underline{H_2}CH_3$, $NC\underline{H_2}COOEt$, $COC\underline{H_2}$thymine), 7.03 (s, 1H, thymine *H*6).

^{13}C NMR (50 MHz, $CDCl_3$) δ = 12.3 (thymineCH_3), 14.0 ($COOCH_2\underline{C}H_3$), 28.4 (Boc$C(\underline{C}H_3)_3$), 29.6 ($NCH_2\underline{C}H_2$), 33.7 (*C*3), 36.0 ($N\underline{C}H_2CH_2$), 46.0 (*C*5), 47.7

(N<u>C</u>$H_2$$COOC_2H_5$), 54.3 (*C*2), 59.4 (*C*4), 61.4 ($COOCH_2CH_3$), 80.4 (Boc<u>C</u>$(CH_3)_3$), 110.7 (*C*5thymine), 141.1 (*C*6thymine), 151.1 (*C*2thymine), 154.2 (BocCO), 164.3 (*C*4thymine), 167.2 (N<u>C</u>OCH_2thymine), 168.8 (<u>C</u>OOC_2H_5).

Mass (ESI) 508.58 ($M+H^+$), 530.64 ($M+ Na^+$) ($C_{22}H_{33}N_7O_7$ requires 507.24)

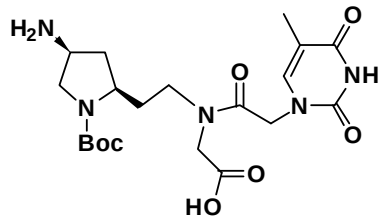
Synthesis of ethyl [–*N*-(*N*-Boc-4*S*-aminopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycinate (26)



Compound **25** (0.8 g 1.6 mmol) and 10% Pd-C (80 mg) were suspended in 10 mL methanol and hydrogenation reaction was carried out at 40 psi hydrogen pressure. After completion of the reaction, the reaction mixture was filtered and solvent was evaporated to give amino derivative (0.69 g, 92%) which was used without further purification in the next step.

Synthesis of [–*N*-(*N*-Boc-4*S*-aminopyrrolidin–2*R*-ethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine (27)

The amino derivative obtained in the previous step (0.6 g, 1.24 mmol) was dissolved in 2 mL methanol and to which was added 2 mL 1N LiOH solution in water. Reaction was complete in 30 min, methanol was removed and aqueous layer was neutralized with Dowex H^+ resin and resin was separated by filtration. The aqueous solution was washed with ethyl acetate, concentrated to give free amino acid derivative **27** (0.5 g, 90%).

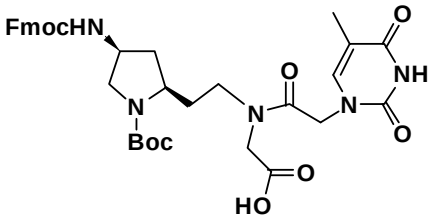


^{1}H NMR (200 MHz, D_2O) δ = 1.33 (s, 9H), 1.64-1.79 (m, 5H), 1.98-2.09 (m, 1H), 2.45-2.59 (m, 1H), 3.80-3.90 (m, 9H), 4.45-4.50 (d, J= 10.06 Hz, 1H), 7.26-7.30 (d, 1H).

^{13}C NMR (50 MHz, D_2O) δ = 11.3, 27.6, 34.8, 44.5, 47.4, 49.0, 49.9, 51.0, 54.1, 54.4, 82.0, 110.6, 143.2, 152.0, 155.6, 166.8, 168.6, 174.8.Mass (ESI) obsd. 454.53 ($M+H^+$), 476.52 ($M+Na^+$) ($C_{20}H_{31}N_5O_7$ requires 453.22) $[\alpha]^{20}_D$ = 15 (c=1.3, MeOH).

Synthesis of [–*N*-(*N*-Boc-4*S*-Fmoc-amino–2*R*-pyrrolidinylethyl)-*N*-(*N*-1-thyminylacetyl)]-glycine (28)

The amino acid obtained in the previous step (0.5 g, 1.1 mmol) was dissolved in 3 mL water-dioxan and $NaHCO_3$ (0.23 g, 2.75 mmol) and Fmoc –succinimide (0.44 g,



1.32 mmol) was added to the reaction mixture. The reaction was stirred for 6 hrs at RT. Then 1,4-dioxane was removed under reduced presssure and aqueous layer was

washed with ethyl acetate until impurity of Fmoc succinimide is removed. Then aqueous layer was neutralized with H^+ dowex resin. The compound was then extracted in 40 mL 5% methanol in ethyl acetate. Organic layer was dried over Na_2SO_4, concentrated and the product was purified by column chromatography using gradient methanol-dichloromethane to give thymine monomer **28** (0.46 g, 61%).

UV λ_{max}(nm)(CH_3OH)= 205(4.4), 265(4.03), 299(3.3),), IR, $\upsilon(cm^{-1})$ (CH_3OH) 1724,1670

^{1}H NMR (200 MHz, DMSO d_6) δ = 1.41 (s, 9H, Boc ($C\underline{H}_3)_3$), 1.55-2.36 (m, 4H, *H*3,3', $NCH_2C\underline{H}_2$) ,1.76 (s, 3H, thymine $C\underline{H}_3$), 2.69 -3.92 (8H, m, *H*2, *H*5,5', $NC\underline{H}_2CH_2$, Fmoc $C\underline{H}_2$, C<u>H</u>), 3.16-3.75 (moisture from DMSO), 4.23-4.63 (m, 5H, *H*4, $COC\underline{H}_2$thymine, $NHC\underline{H}_2COOH$) 7.31-7.92 (m, 9H, Fmoc, Thymine *H*5).

^{13}C NMR (50 MHz, DMSO d_6) δ = 12.3(thymineCH_3), 28.5(BocC($\underline{C}H_3)_3$), 29.4 ($NCH_2\underline{C}H_2$), 31.6 (*C*3), 41.2 ($N\underline{C}H_2CH_2$) 44.4(*C*5), 47.1 (Fmoc<u>C</u>HCH_2OCON), 48.1 ($N\underline{C}H_2COOH$), 50.8 (*C*2), 54.1 (*C*4), 65.8 ($\underline{C}H_2$OFMOC), 79.0 (Boc<u>C</u>$(CH_3)_3$), 108.4 (*C*5 thymine), (FmocAr110.1, 120.5, 121.8, 125.6, 127.5, 128.0, 129.9, 137.9, 139.8, 141.2, 144.3) 142.7 (*C*6 thymine), 151.5 (*C*2thymime), 153.8(BocCO), 156.2(FmocCO), 164.2(*C*4thymine), 167.5($N\underline{C}OCH_2$), 174.4 (<u>C</u>OOH).

Mass (ESI) 698.68 (M+ Na^+), 720.69(M+2Na^+) ($C_{35}H_{41}N_5O_9$ requires 675.29) $[\alpha]^{20}_D$ = 16.7(c=1.2, MeOH)

Oligomer Synthesis: The PNA oligomers were synthesized by standard Boc-solid phase peptide strategy.[3] MBHA resin (4-Methylbenzhydrylamine resin) was selected as solid polymeric matrix on which oligomers were synthesized**.** The free amine content on the resin was determined by the picrate assay and was found to be 1.75 mmol/g and loading was suitably lowered to approximately 0.30 mmol/g by partial acetylation of amine content using calculated amount of acetic anhydride. Free -NH_2 groups on the resin available for coupling are again estimated before starting synthesis. α*N*-Boc ω*N* ClCbz lysine was linked to the resin through amide bond**.** All the monomers and coupling reagent TBTU and HOBt were useed 3 equivalents of

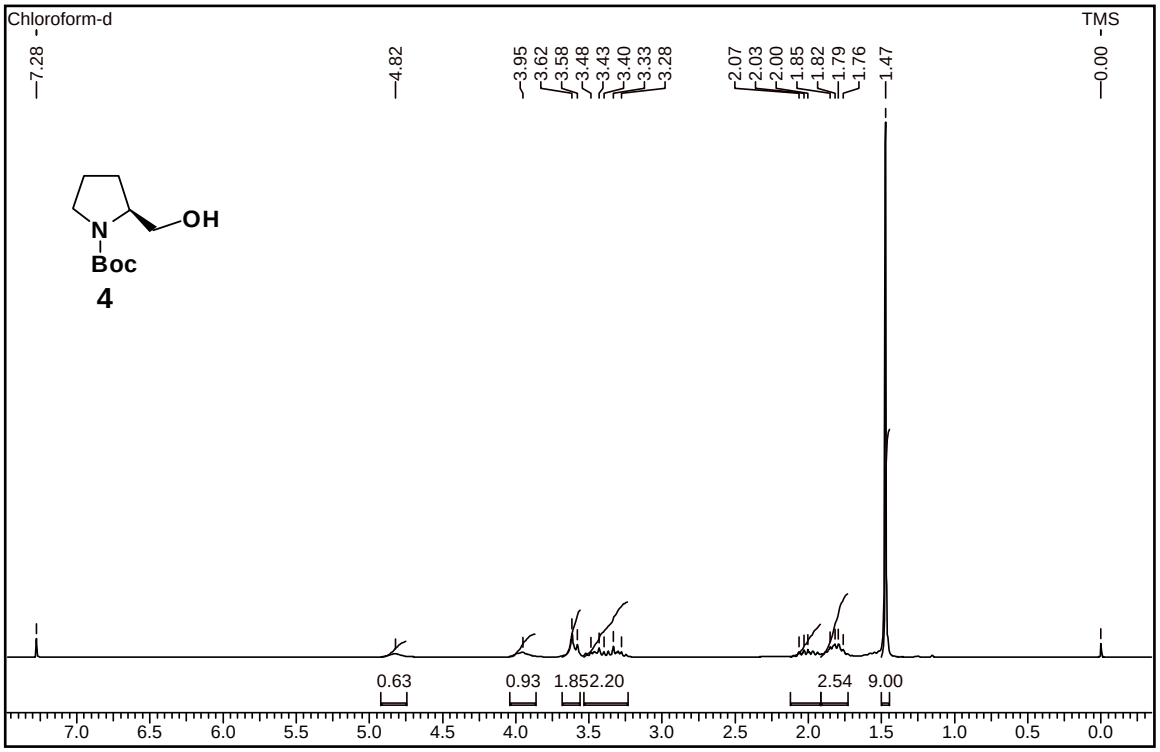
loading value. The completion of coupling was confirmed by kaiser test. PNA oligomers were cleaved with standard TFA-TFASA protocol. Purification of all the PNA oligomers was carried out by Varian dual pump PROSTAR model No.210 HPLC on RP-C18 column with Water: CH_3CN-0.1% TFA system. PNA oligomers were characterized by MALDI-TOF mass spectrometry by using Voyager-DE-STR (Applied Biosystems) MALDI-TOF. A nitrogen laser (337 nm) was used for desorption and ionization. Spectra were acquired in linear mode. The matrix used for analysis was CHCA (α-cyano hydroxyl cinnamic acid).

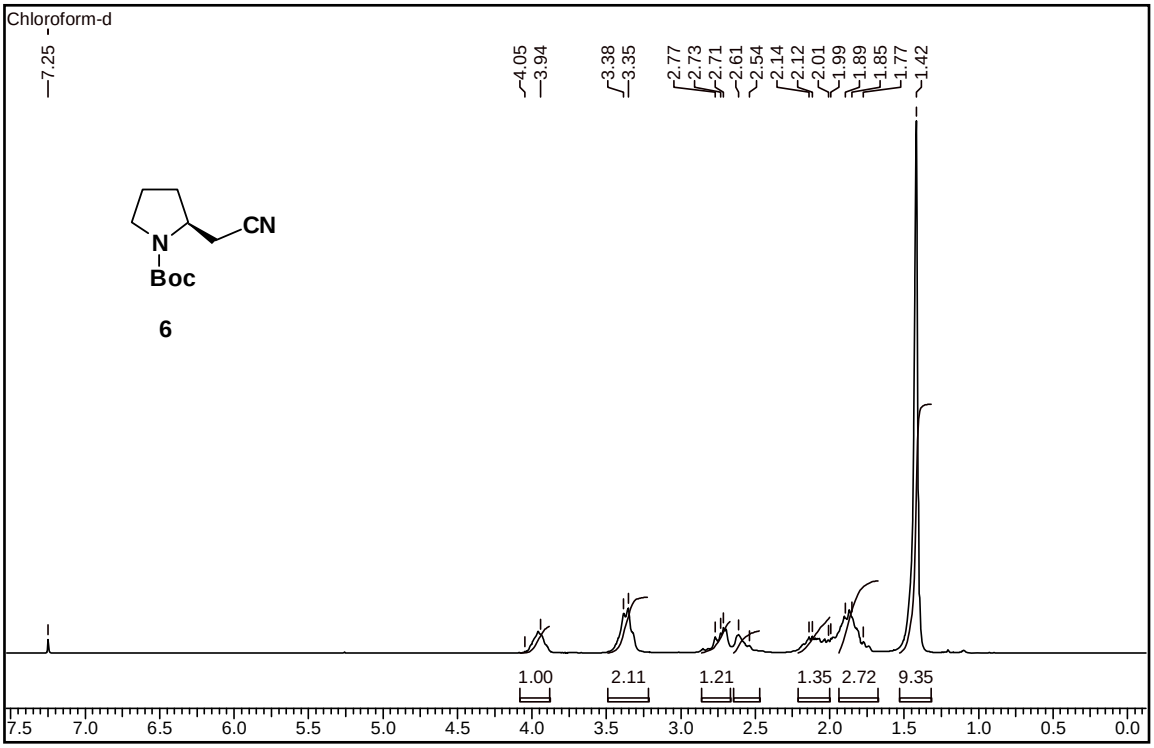
Post synthetic conversion of Am-pet-PNA to Gu-pet*PNA: Am-pet-PNA synthesized MBHA resin (5mg) was suspended in dry DMF. Fmoc protected amino groups in Am-pet units in the Am-pet-PNA deprotectd by 20% piperidine-DMF solution and resin was then subjected for guanidinilation by suspending in the solution of pyrazole-1-caboxamidine hydrochloride (10 eqivalents for each amino group)/DIPEA in DMF-THF (1:2) mixture. After 6-7 hours, the resin was washed with dry DMF and treated with the same guanidination mixture once again to ensure completion conversion from amino to guanidino functionality. The resin was washed with dry DMF and dry DCM and was then desscicated. Gu-pet-PNA was cleaved from this resin with standard conditions employing TFMSA-TFA and purified by reverse phase HPLC. (yield of conversion of amino to guanidino is > 95%). The complete conversion of amino to guanidino groups was cheacked with reverse phase HPLC analysis. This was confirmed by co-injection of amino-pet-PNA and Guanidio-pet-PNA sequences.

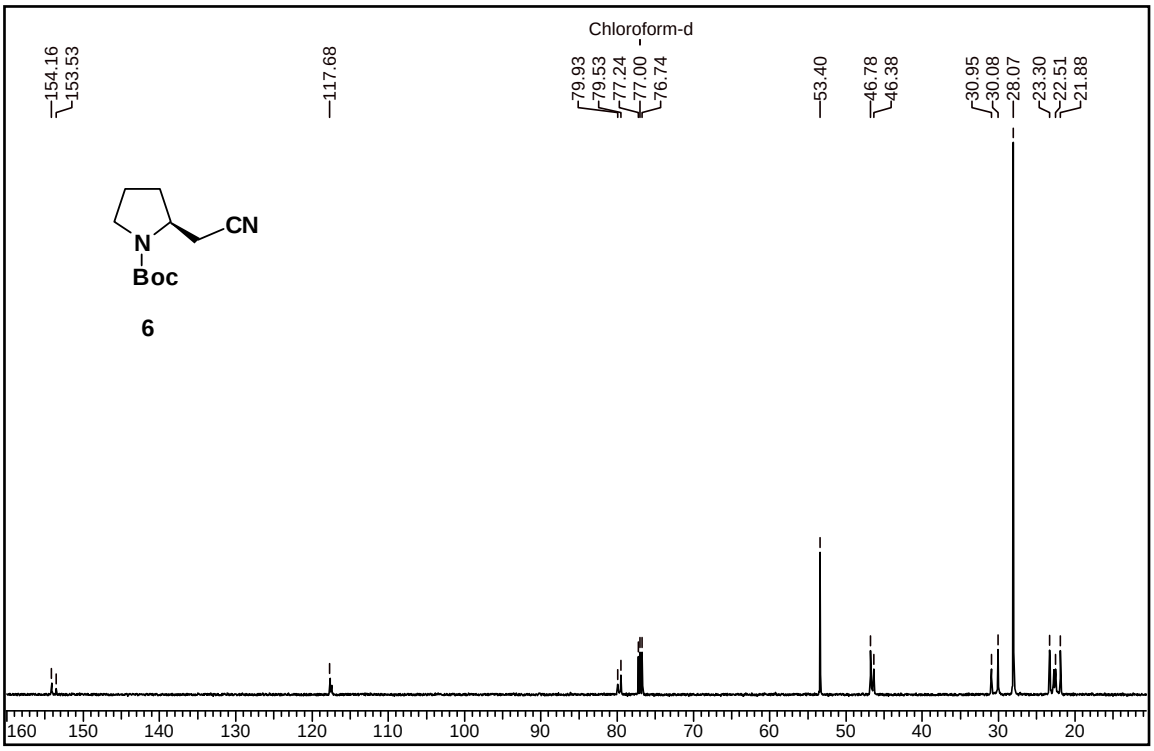
2.1.9 Appendix

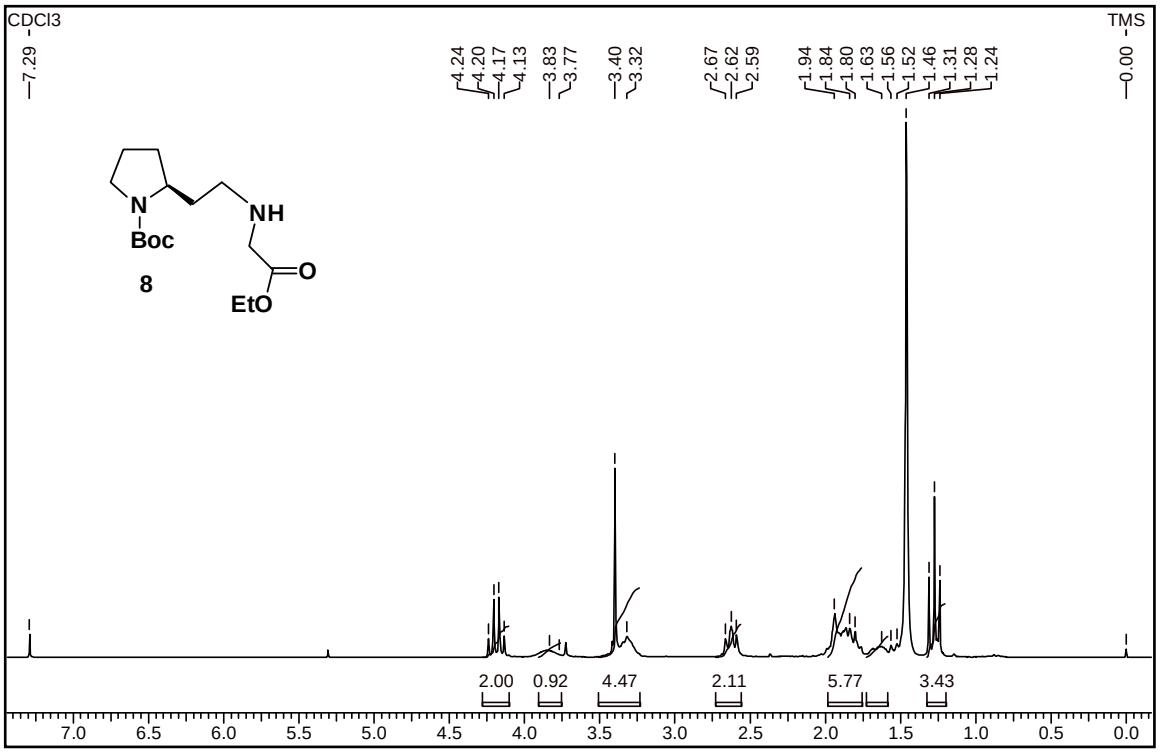
1H and 13C spectra for compound **4**-**28** -Page no 82-100

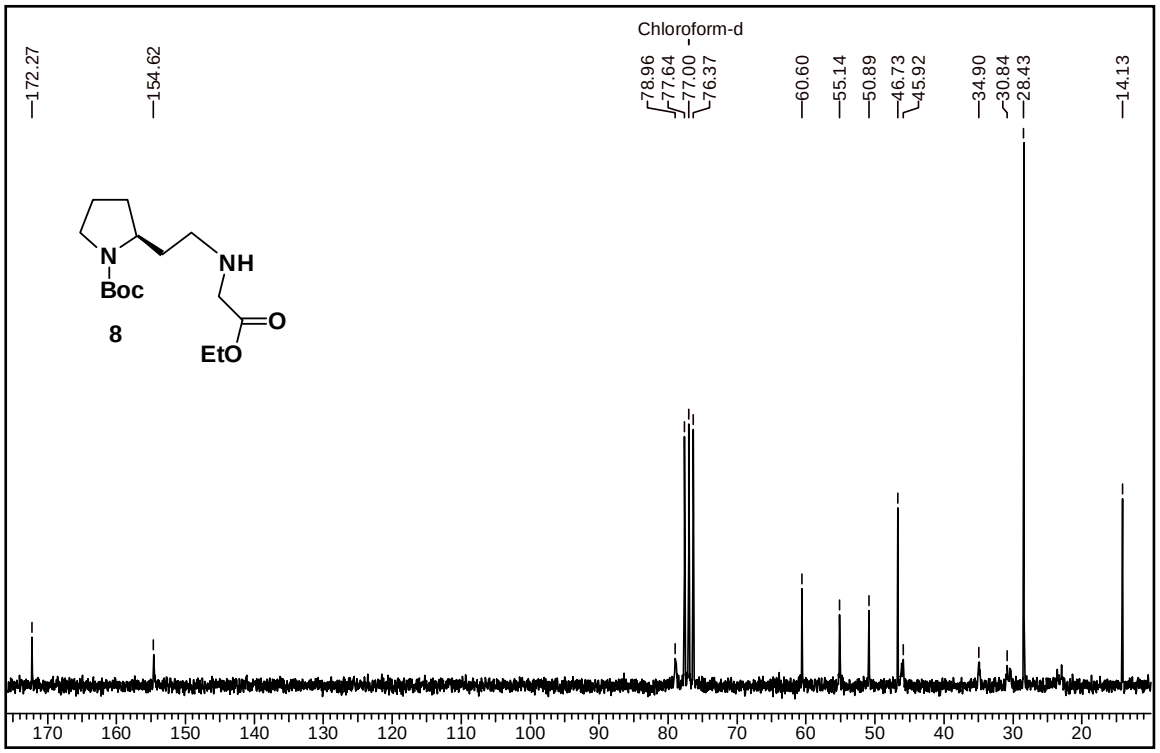
HPLC and MALDI-TOF spectra for PNA sequences-**30**-**46**–Page no- 101-114

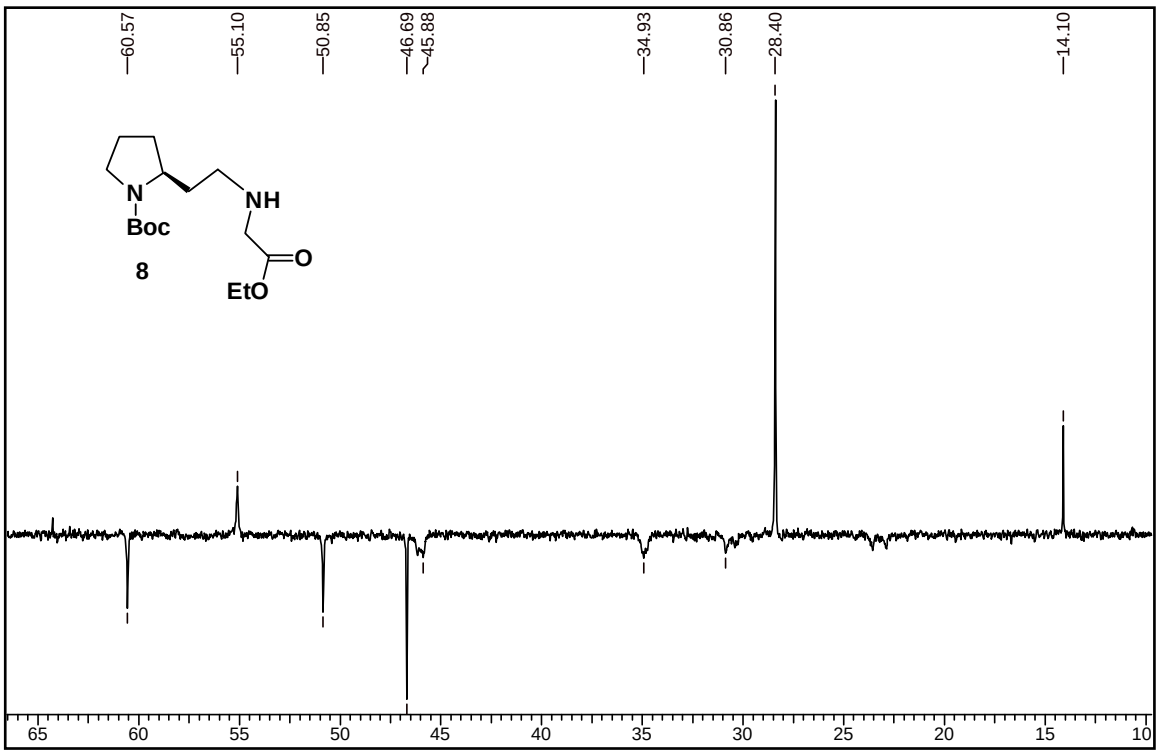

Chloroform-d
TMS
7.28
4.82
3.95
3.62
3.58
3.48
3.43
3.40
3.33
3.28
2.07
2.03
2.00
1.85
1.82
1.79
1.76
1.47
0.00
OH
N
Boc
4
0.63
0.93
1.85
2.20
2.54
9.00
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

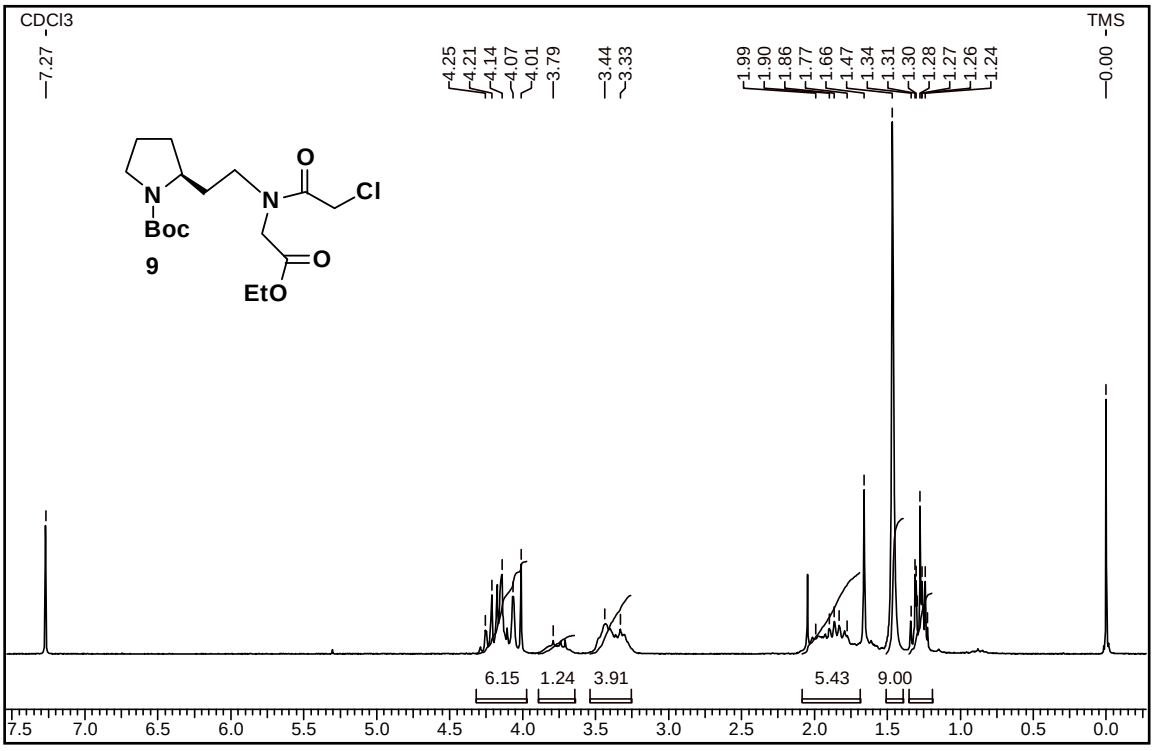

Chloroform-d
7.25
4.05
3.94
3.38
3.35
2.77
2.73
2.71
2.61
2.54
2.14
2.12
2.01
1.99
1.89
1.85
1.77
1.42
CN
N
Boc
6
1.00
2.11
1.21
1.35
2.72
9.35
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

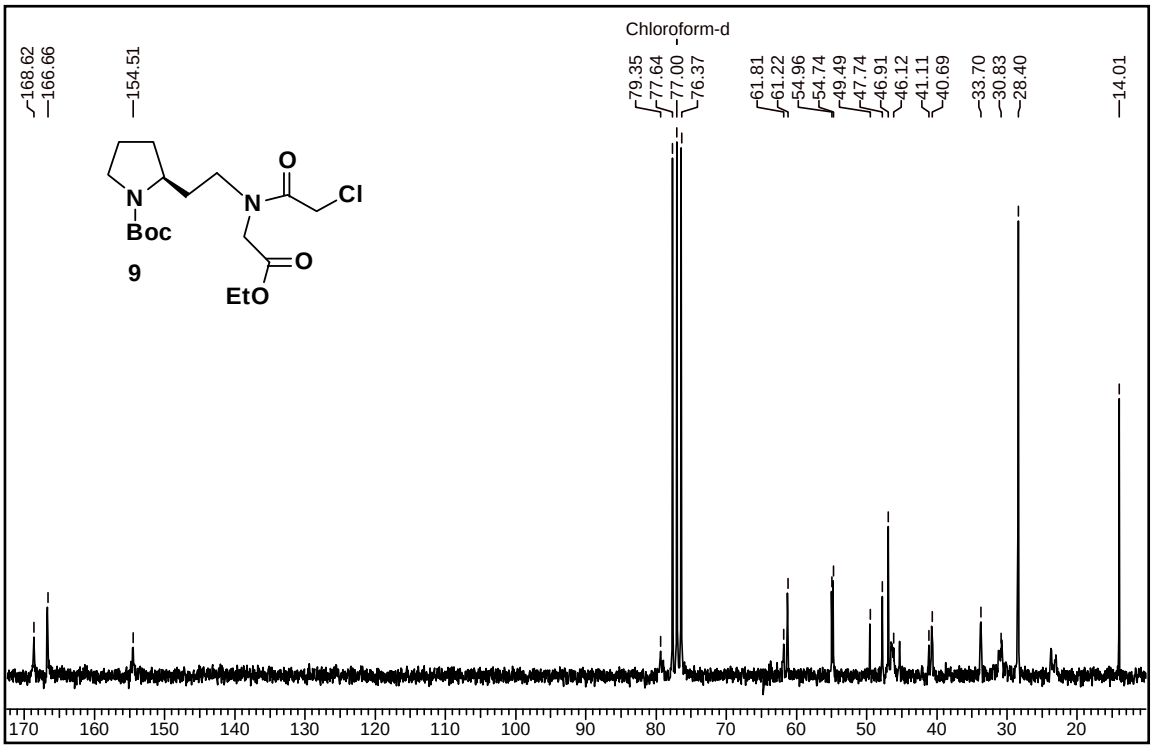

Chloroform-d
154.16
153.53
117.68
79.93
79.53
77.24
77.00
76.74
53.40
46.78
46.38
30.95
30.08
28.07
23.30
22.51
21.88
N
Boc
CN
6
160 150 140 130 120 110 100 90 80 70 60 50 40 30 20

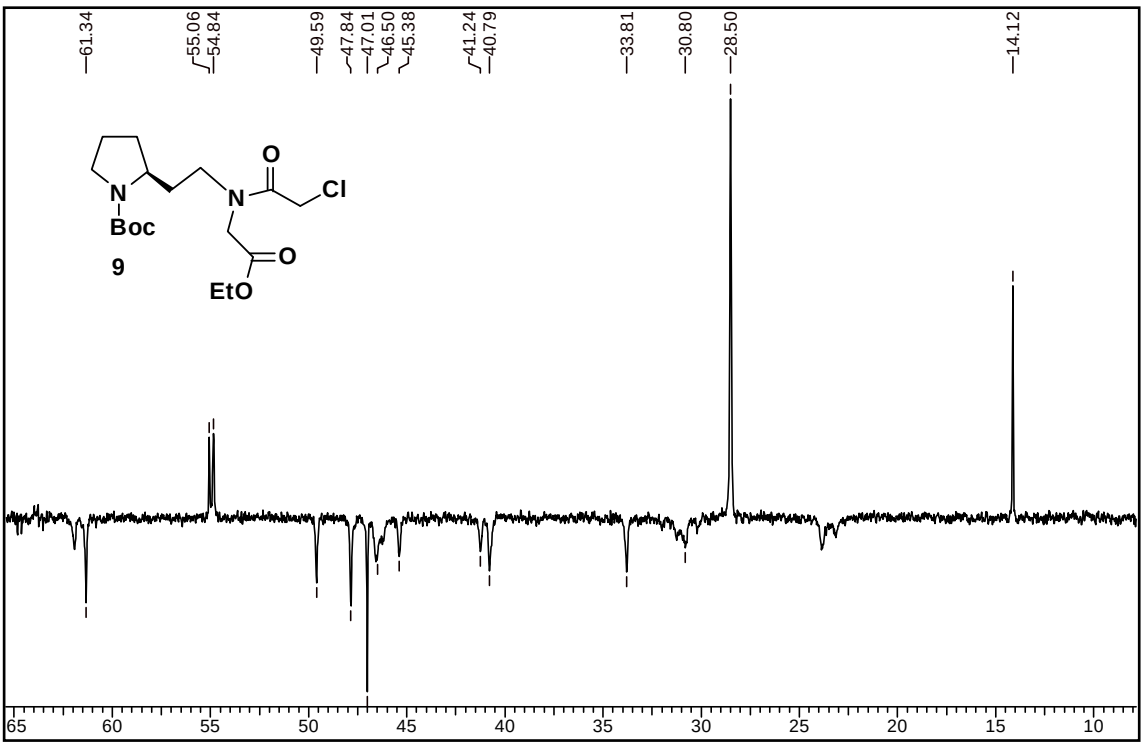

CDCl3
TMS
7.29
4.24
4.20
4.17
4.13
3.83
3.77
3.40
3.32
2.67
2.62
2.59
1.94
1.84
1.80
1.63
1.56
1.52
1.46
1.31
1.28
1.24
0.00
N
Boc
NH
O
EtO
8
2.00 0.92 4.47 2.11 5.77 3.43
7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

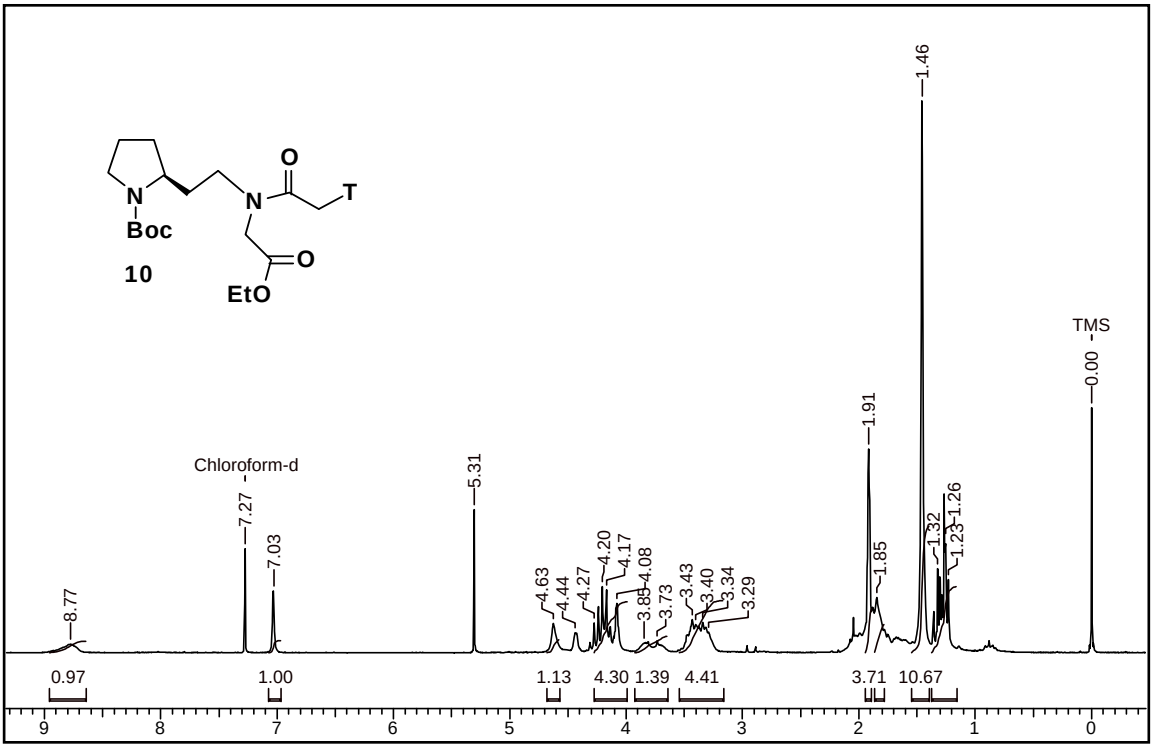

Chloroform-d
172.27
154.62
78.96
77.64
77.00
76.37
60.60
55.14
50.89
46.73
45.92
34.90
30.84
28.43
14.13
N
Boc
NH
O
EtO
8
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20

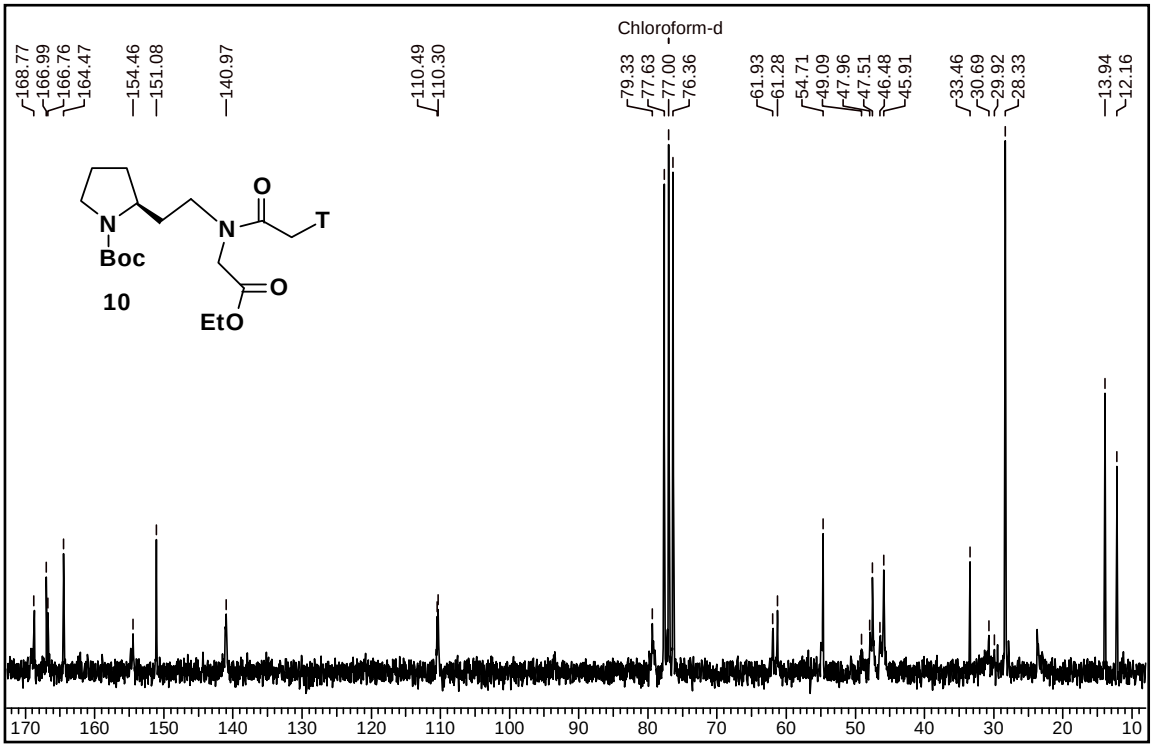

60.57
55.10
50.85
46.69
45.88
34.93
30.86
28.40
14.10
N
Boc
NH
O
EtO
8
65 60 55 50 45 40 35 30 25 20 15 10

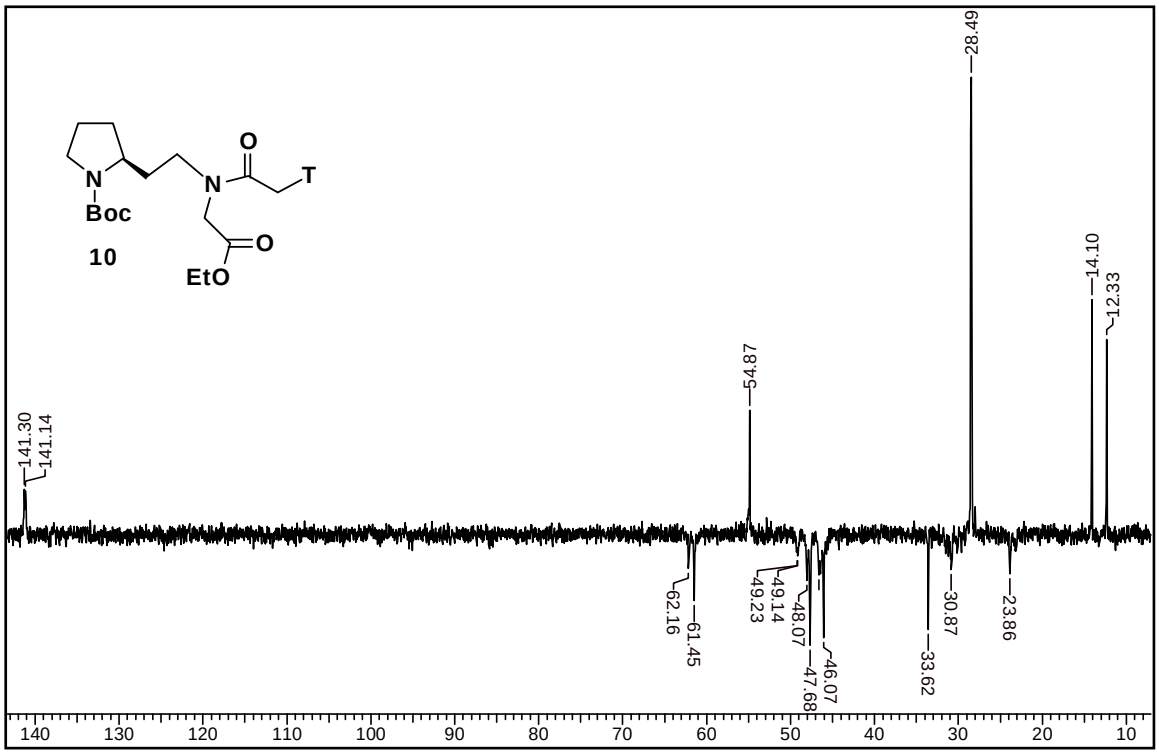

CDCl3
TMS
7.27
4.25
4.21
4.14
4.07
4.01
3.79
3.44
3.33
1.99
1.90
1.86
1.77
1.66
1.47
1.34
1.31
1.30
1.28
1.27
1.26
1.24
0.00
N
Boc
9
N
O
Cl
O
EtO
6.15
1.24
3.91
5.43
9.00
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

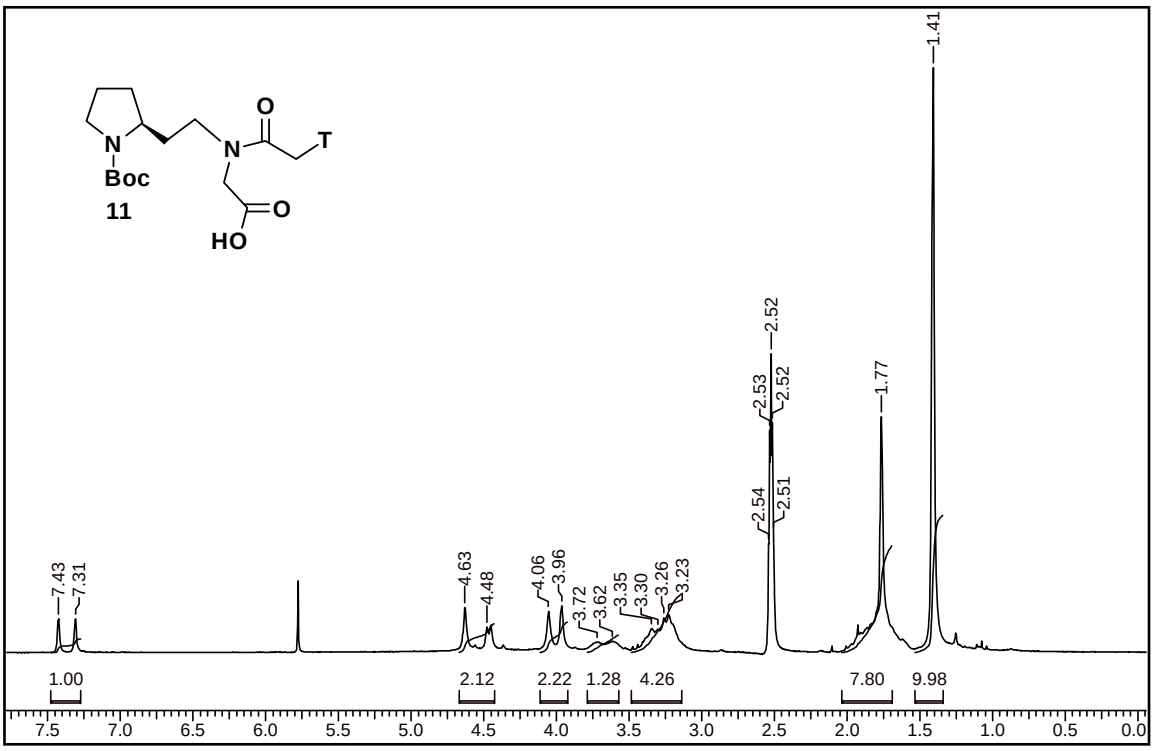

Chloroform-d
168.62
166.66
154.51
79.35
77.64
77.00
76.37
61.81
61.22
54.96
54.74
49.49
47.74
46.91
46.12
41.11
40.69
33.70
30.83
28.40
14.01
N
Boc
9
N
O
Cl
O
EtO
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20

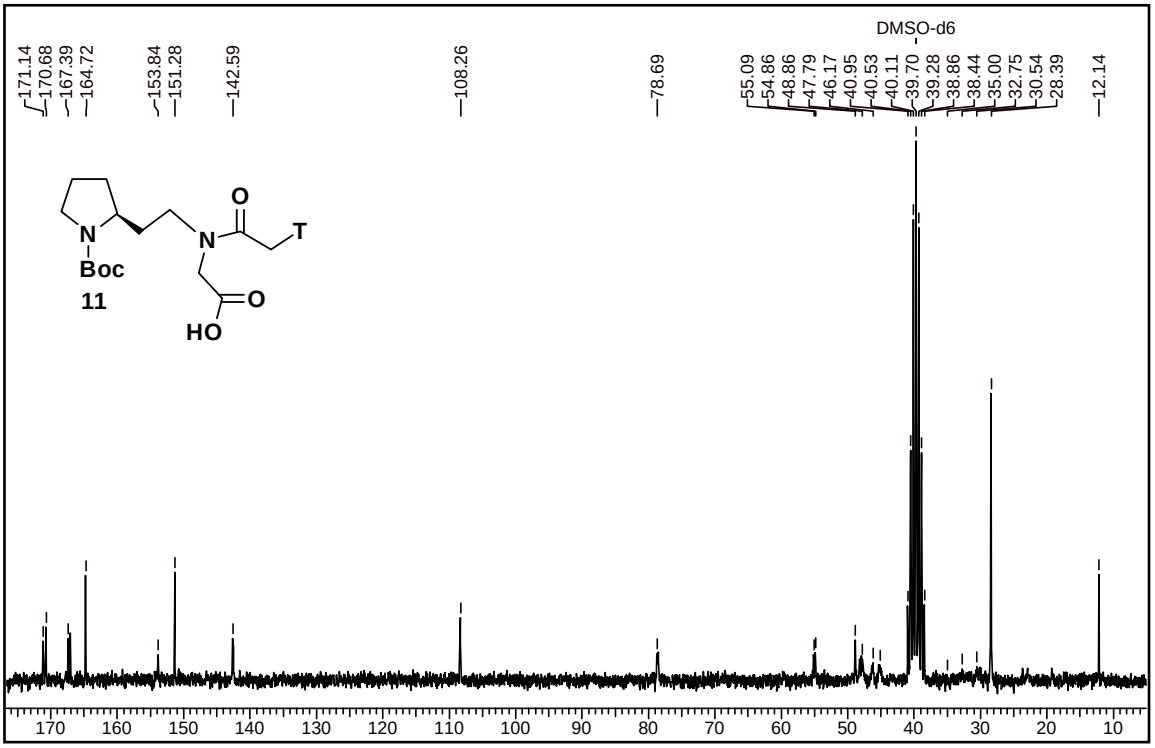

61.34
55.06
54.84
49.59
47.84
47.01
46.50
45.38
41.24
40.79
33.81
30.80
28.50
14.12
O
Cl
N
N
Boc
O
9
EtO
65
60
55
50
45
40
35
30
25
20
15
10

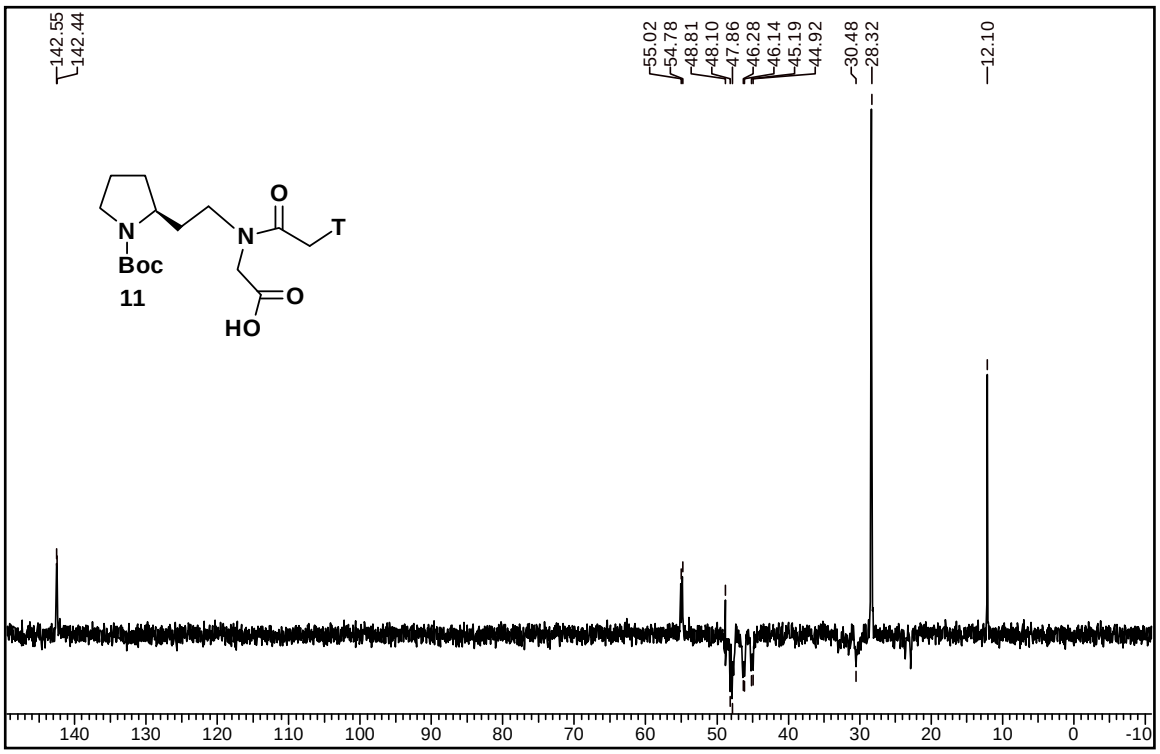

O
T
N
N
Boc
O
10
EtO
1.46
TMS
0.00
1.91
Chloroform-d
5.31
7.27
7.03
8.77
4.63
4.44
4.27
4.20
4.17
4.08
3.85
3.73
3.43
3.40
3.34
3.29
1.85
1.32
1.23
1.26
0.97
1.00
1.13
4.30
1.39
4.41
3.71
10.67
9
8
7
6
5
4
3
2
1
0

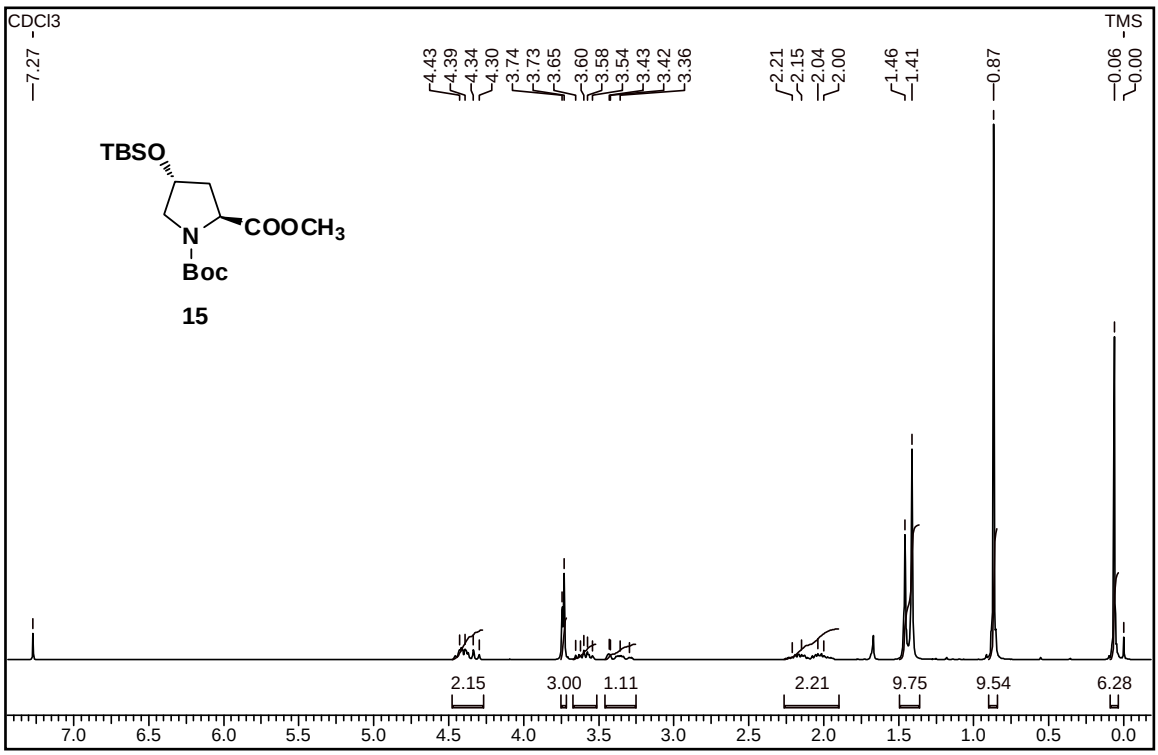

Chloroform-d
168.77
166.99
166.76
164.47
154.46
151.08
140.97
110.49
110.30
79.33
77.63
77.00
76.36
61.93
61.28
54.71
49.09
47.96
47.51
46.48
45.91
33.46
30.69
29.92
28.33
13.94
12.16
N
Boc
10
O
N
T
O
EtO
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

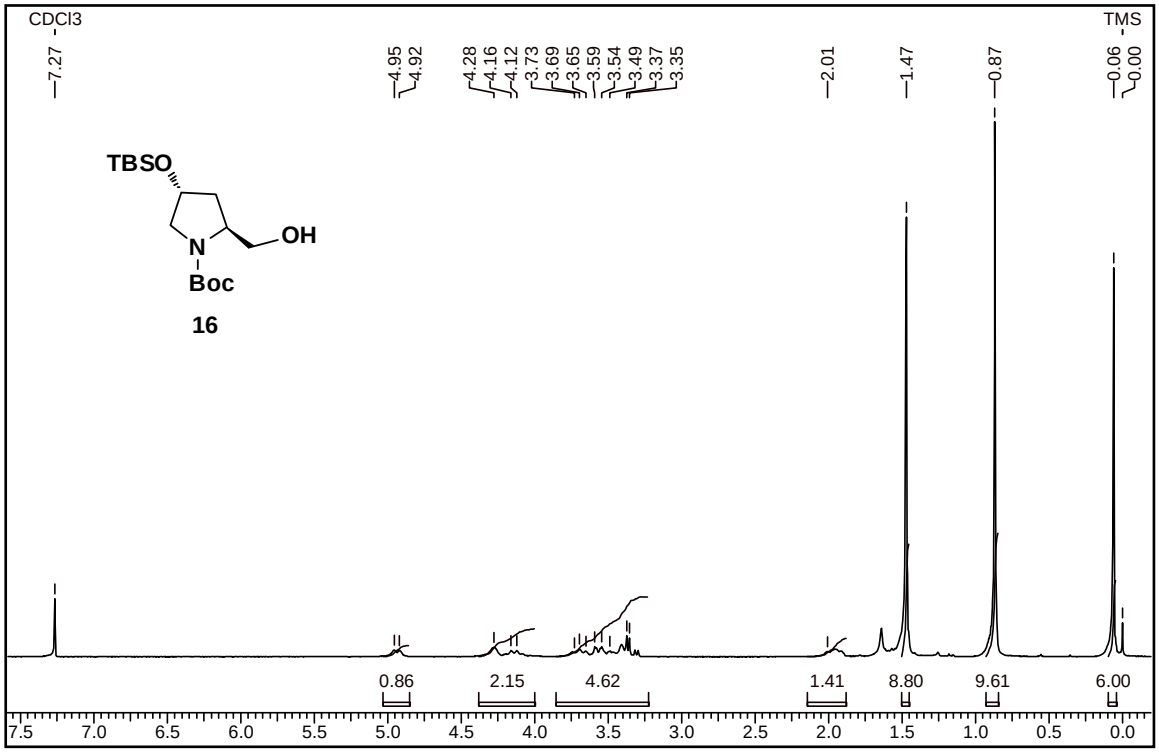

28.49
14.10
12.33
54.87
141.30
141.14
62.16
61.45
49.23
49.14
48.07
47.68
46.07
33.62
30.87
23.86
N
Boc
10
O
N
T
O
EtO
140
130
120
110
100
90
80
70
60
50
40
30
20
10

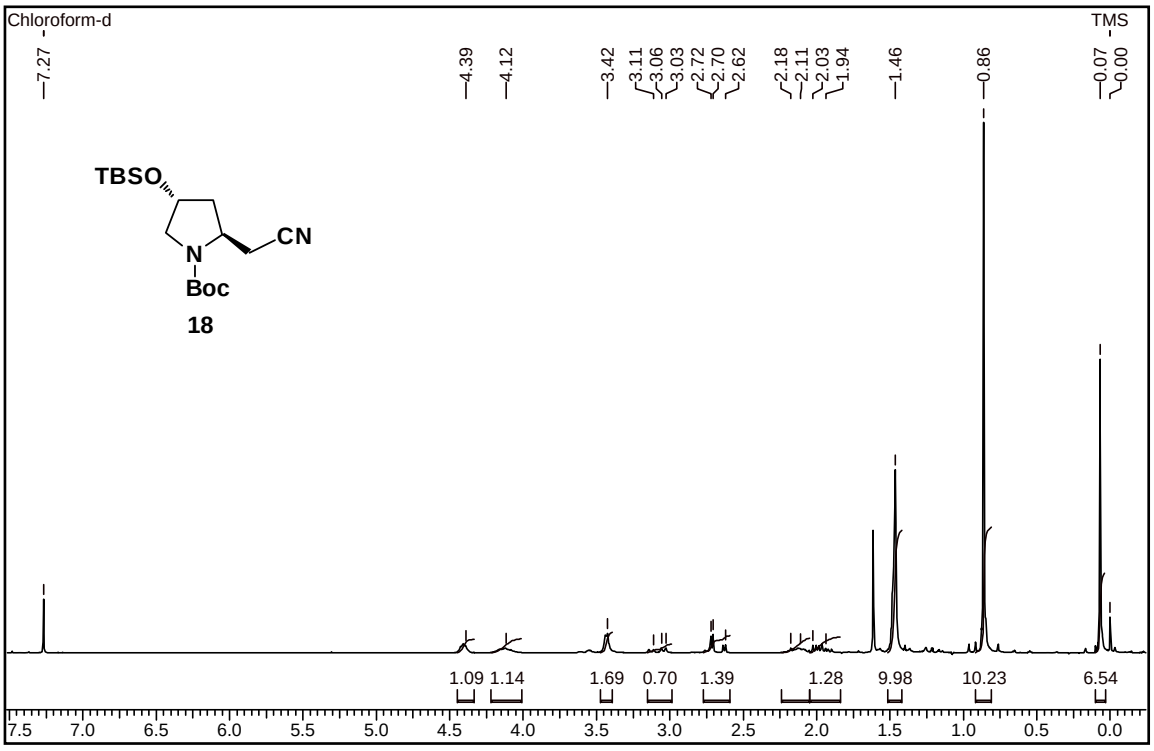

N
Boc
11
O
N
T
O
HO
1.41
2.52
2.53
2.52
1.77
2.54
2.51
7.43
7.31
4.63
4.48
4.06
3.96
3.72
3.62
3.35
3.30
3.26
3.23
1.00
2.12
2.22
1.28
4.26
7.80
9.98
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

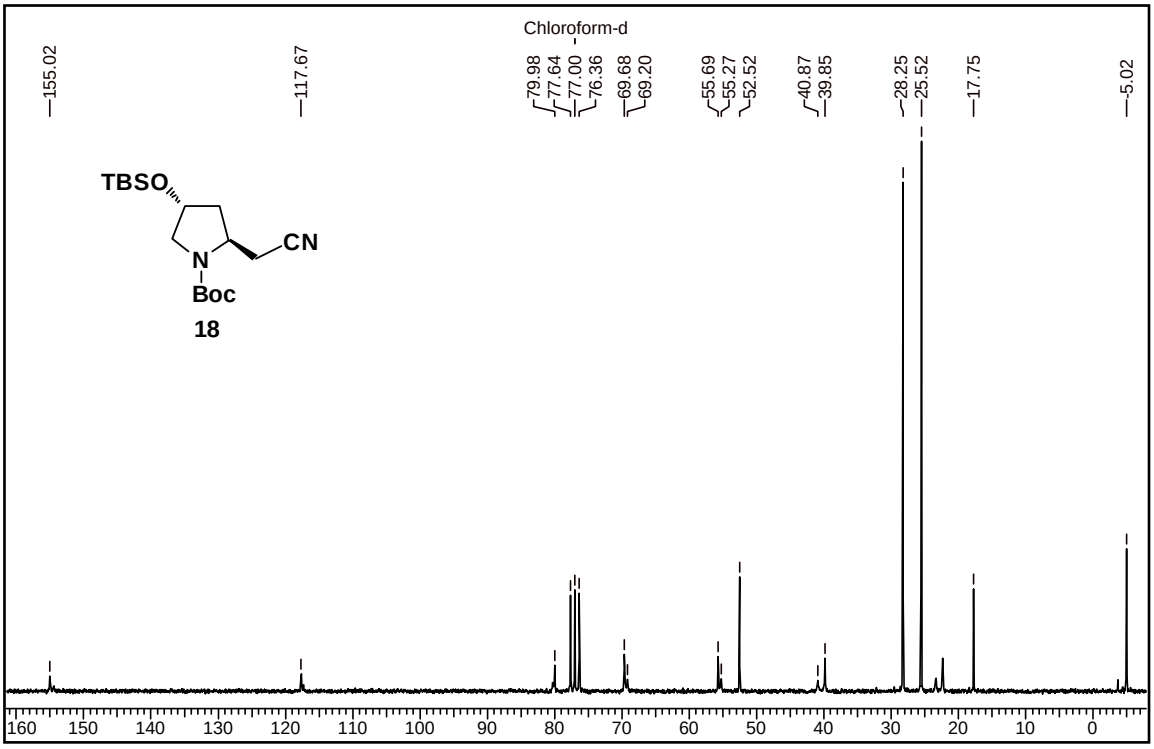

DMSO-d6
171.14
170.68
167.39
164.72
153.84
151.28
142.59
108.26
78.69
55.09
54.86
48.86
47.79
46.17
40.95
40.53
40.11
39.70
39.28
38.86
38.44
35.00
32.75
30.54
28.39
12.14
N
Boc
11
O
N
T
O
HO
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

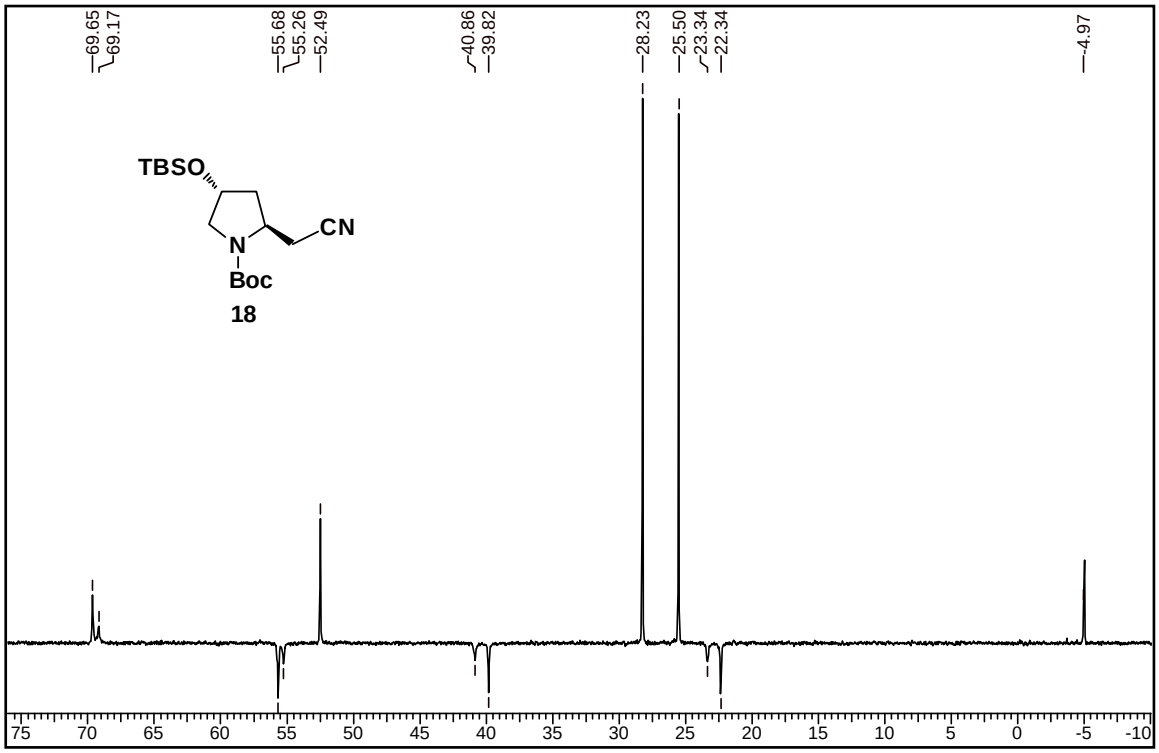

142.55
142.44
55.02
54.78
48.81
48.10
47.86
46.28
46.14
45.19
44.92
30.48
28.32
12.10
N
Boc
11
N
O
T
O
HO
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

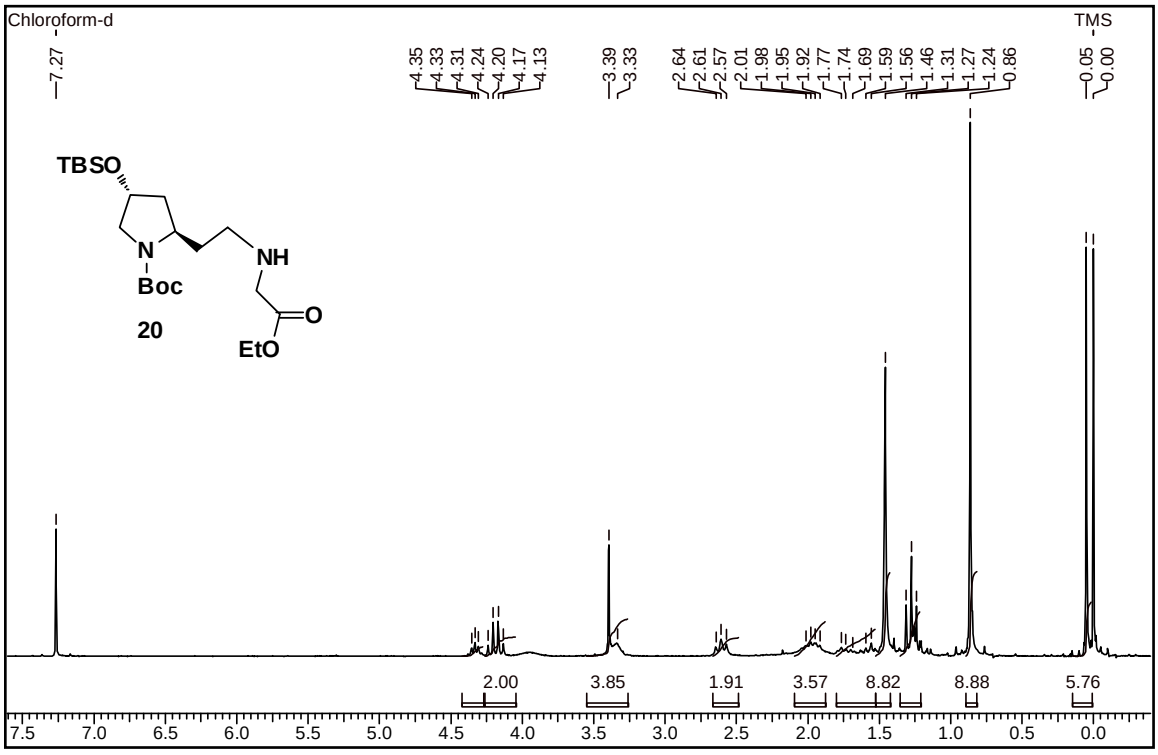

CDCl3
TMS
7.27
4.43
4.39
4.34
4.30
3.74
3.73
3.65
3.60
3.58
3.54
3.43
3.42
3.36
2.21
2.15
2.04
2.00
1.46
1.41
0.87
0.06
0.00
TBSO
N
COOCH3
Boc
15
2.15
3.00
1.11
2.21
9.75
9.54
6.28
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

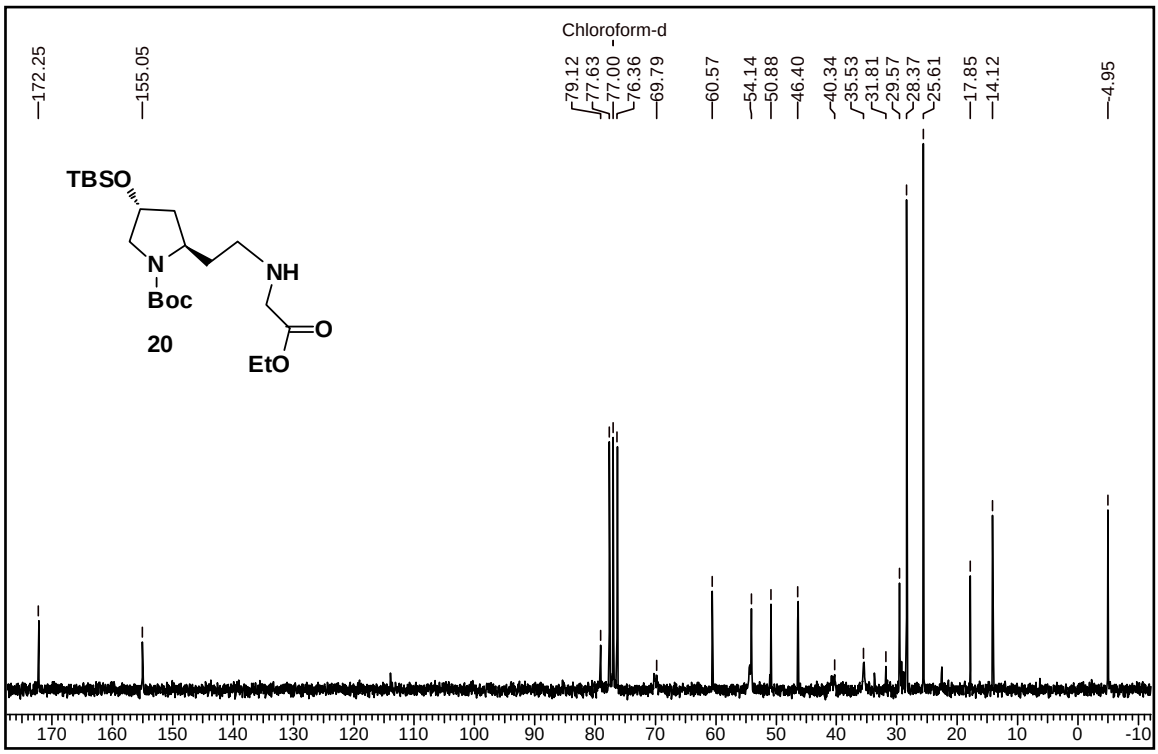

CDCl3
TMS
7.27
4.95
4.92
4.28
4.16
4.12
3.73
3.69
3.65
3.59
3.54
3.49
3.37
3.35
2.01
1.47
0.87
0.06
0.00
TBSO
OH
N
Boc
16
0.86
2.15
4.62
1.41
8.80
9.61
6.00

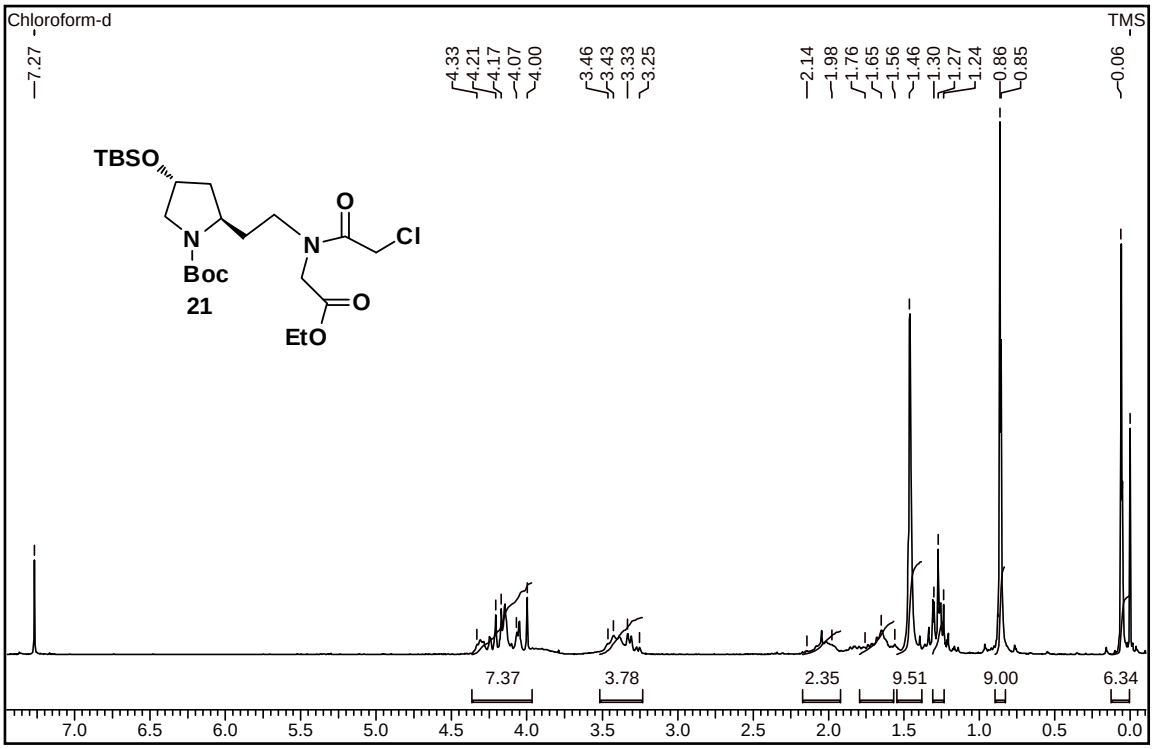

Chloroform-d
TMS
7.27
4.39
4.12
3.42
3.11
3.06
3.03
2.72
2.70
2.62
2.18
2.11
2.03
1.94
1.46
0.86
0.07
0.00
TBSO
CN
N
Boc
18
1.09
1.14
1.69
0.70
1.39
1.28
9.98
10.23
6.54

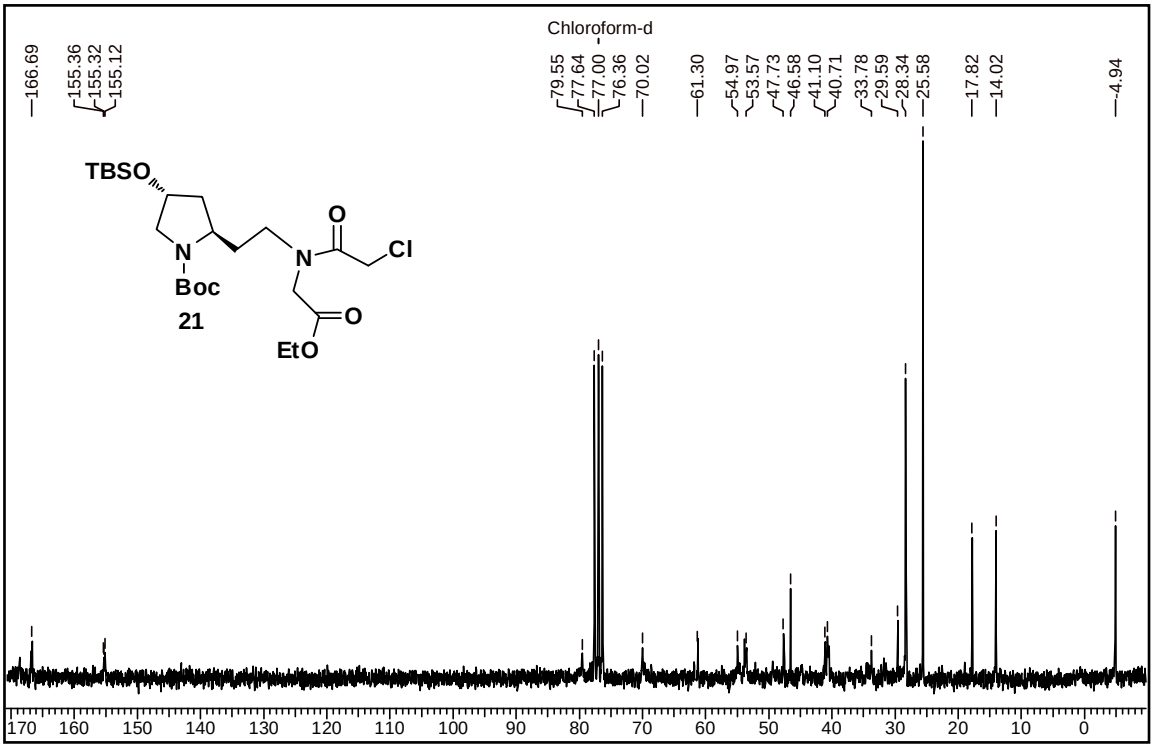

Chloroform-d
155.02
117.67
79.98
77.64
77.00
76.36
69.68
69.20
55.69
55.27
52.52
40.87
39.85
28.25
25.52
17.75
-5.02
TBSO
CN
N
Boc
18
160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

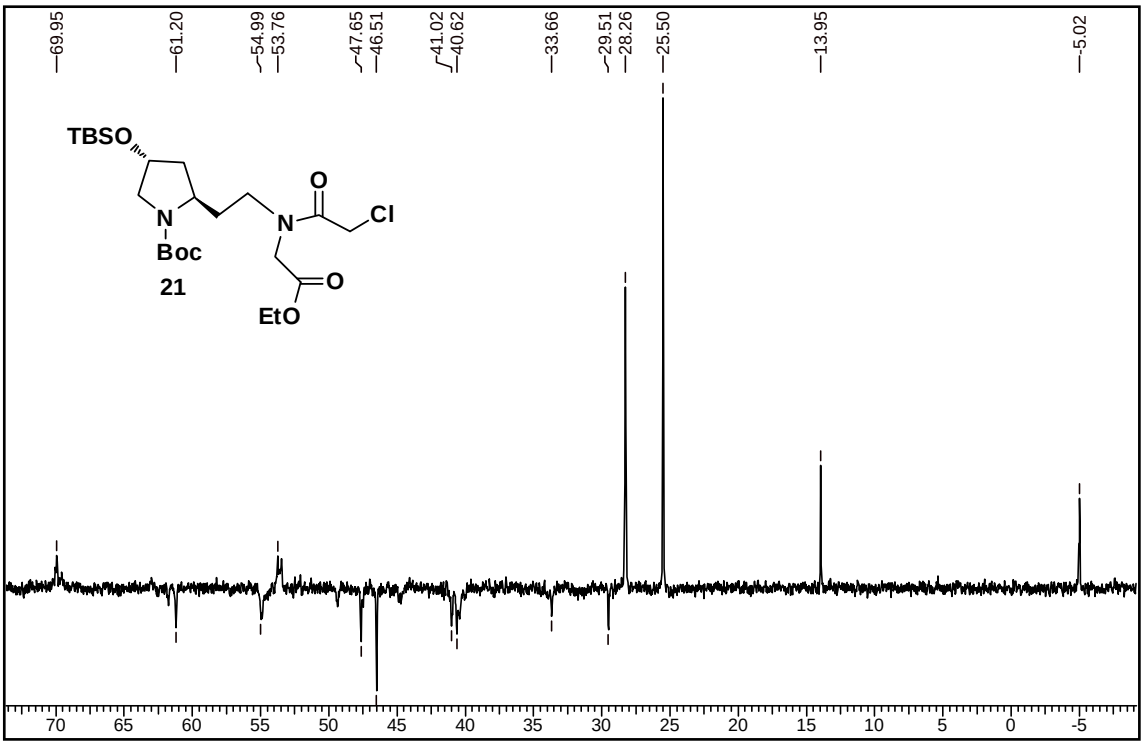

69.65
69.17
55.68
55.26
52.49
40.86
39.82
28.23
25.50
23.34
22.34
-4.97
TBSO
CN
N
Boc
18
75 70 65 60 55 50 45 40 35 30 25 20 15 10 5 0 -5 -10

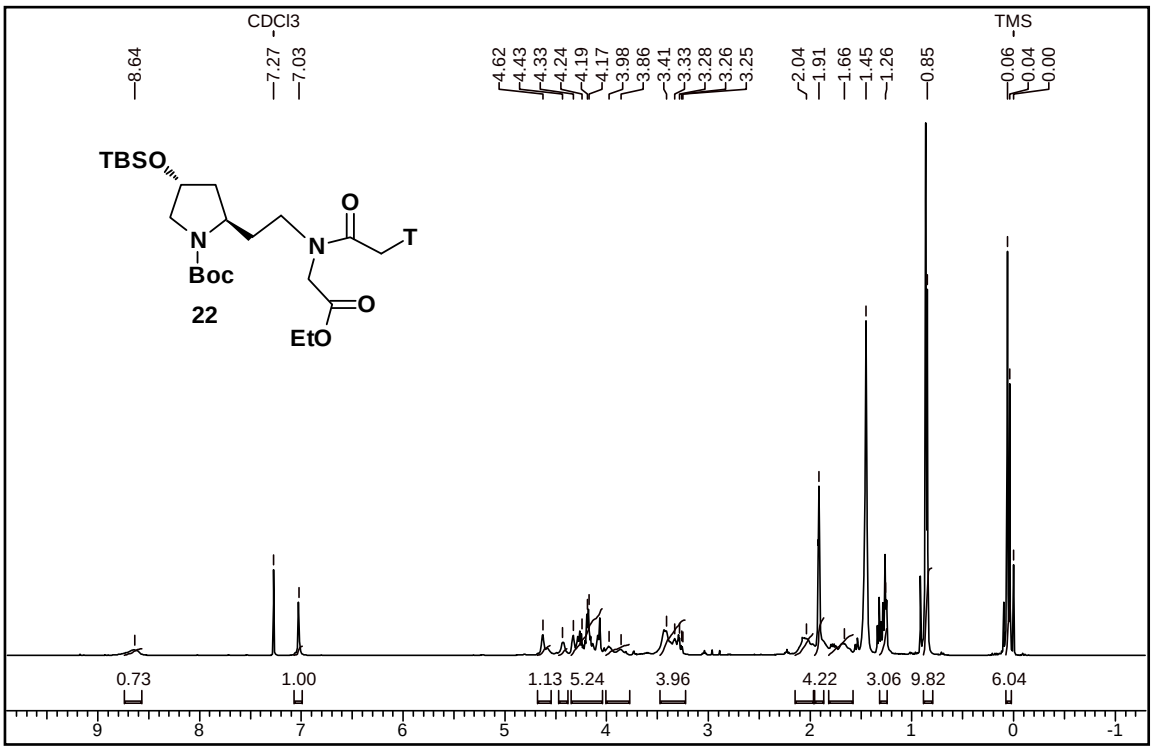

Chloroform-d
TMS
7.27
4.35
4.33
4.31
4.24
4.20
4.17
4.13
3.39
3.33
2.64
2.61
2.57
2.01
1.98
1.95
1.92
1.77
1.74
1.69
1.59
1.56
1.46
1.31
1.27
1.24
0.86
0.05
0.00
TBSO
N
Boc
NH
O
EtO
20
2.00
3.85
1.91
3.57
8.82
8.88
5.76
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

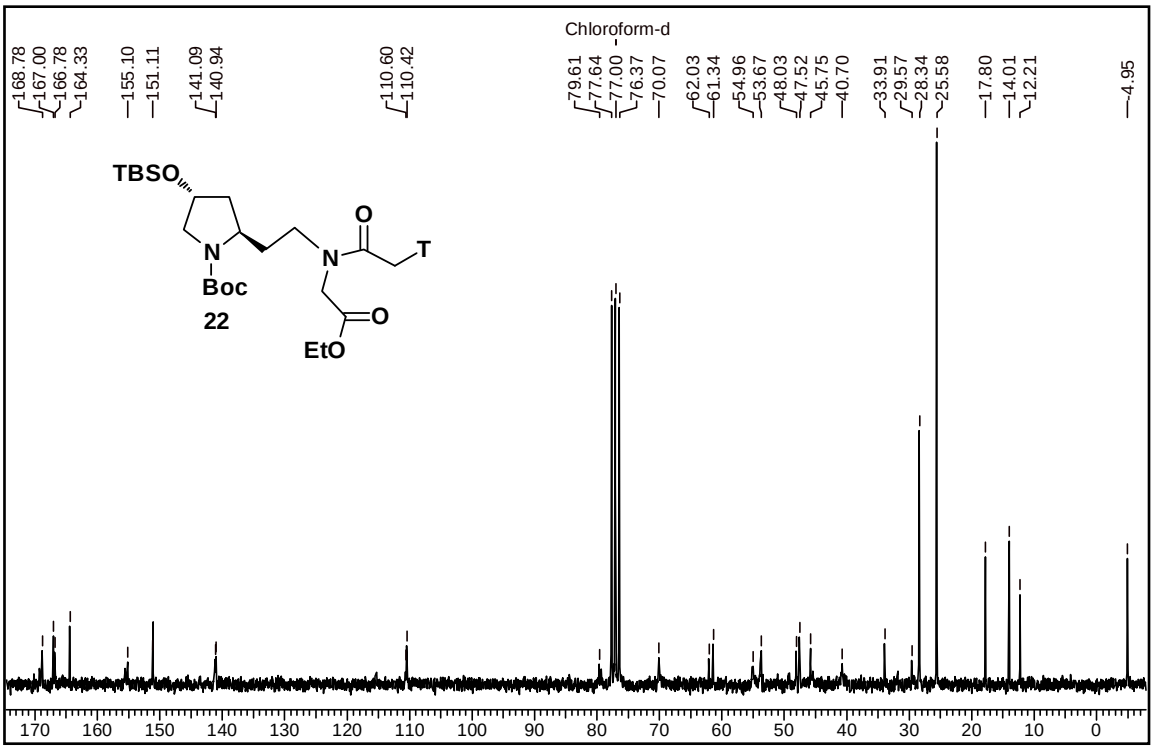

Chloroform-d
172.25
155.05
79.12
77.63
77.00
76.36
69.79
60.57
54.14
50.88
46.40
40.34
35.53
31.81
29.57
28.37
25.61
17.85
14.12
4.95
TBSO
N
Boc
NH
O
EtO
20
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

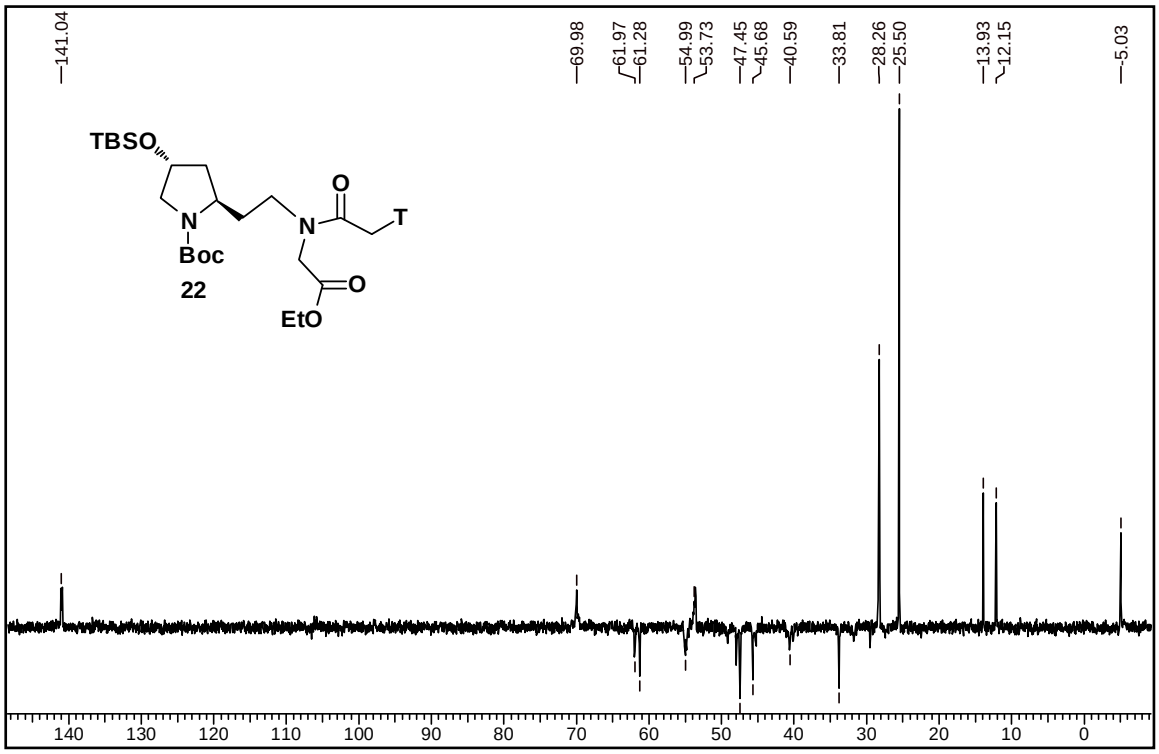

Chloroform-d
TMS
7.27
4.33
4.21
4.17
4.07
4.00
3.46
3.43
3.33
3.25
2.14
1.98
1.76
1.65
1.56
1.46
1.30
1.27
1.24
0.86
0.85
0.06
TBSO
N
Boc
21
O
N
Cl
O
EtO
7.37
3.78
2.35
9.51
9.00
6.34
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

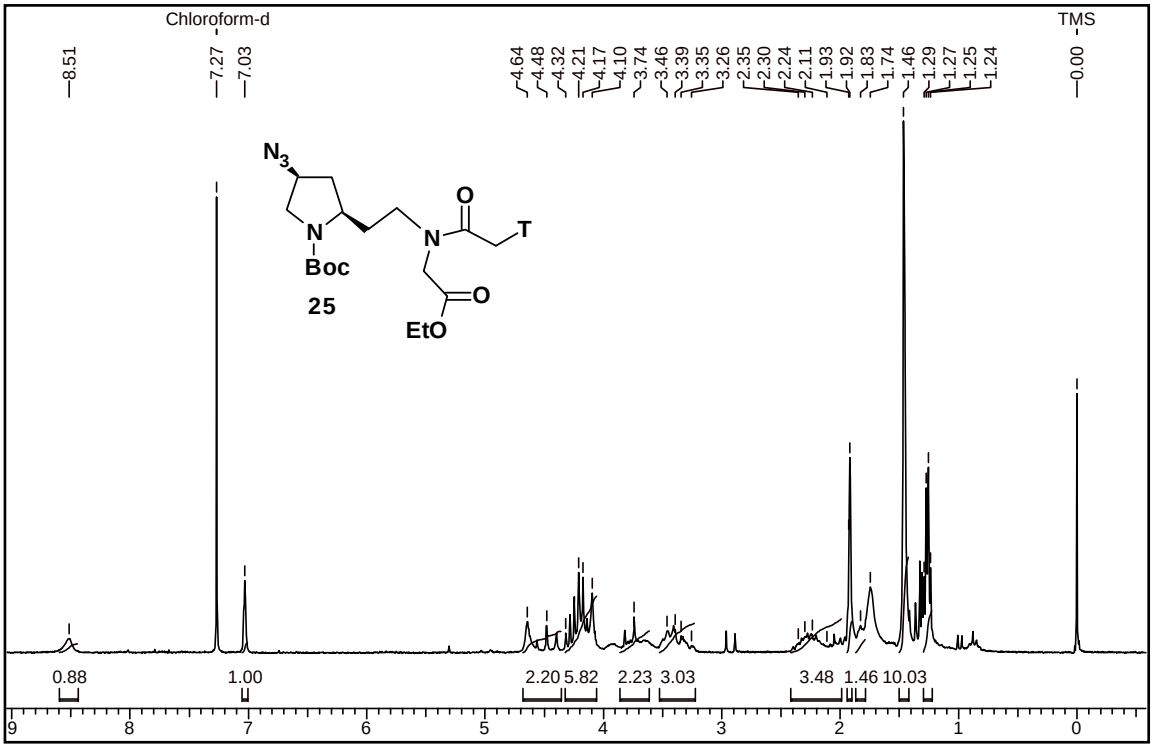

Chloroform-d
166.69
155.36
155.32
155.12
79.55
77.64
77.00
76.36
70.02
61.30
54.97
53.57
47.73
46.58
41.10
40.71
33.78
29.59
28.34
25.58
17.82
14.02
4.94
TBSO
N
Boc
21
O
N
Cl
O
EtO
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

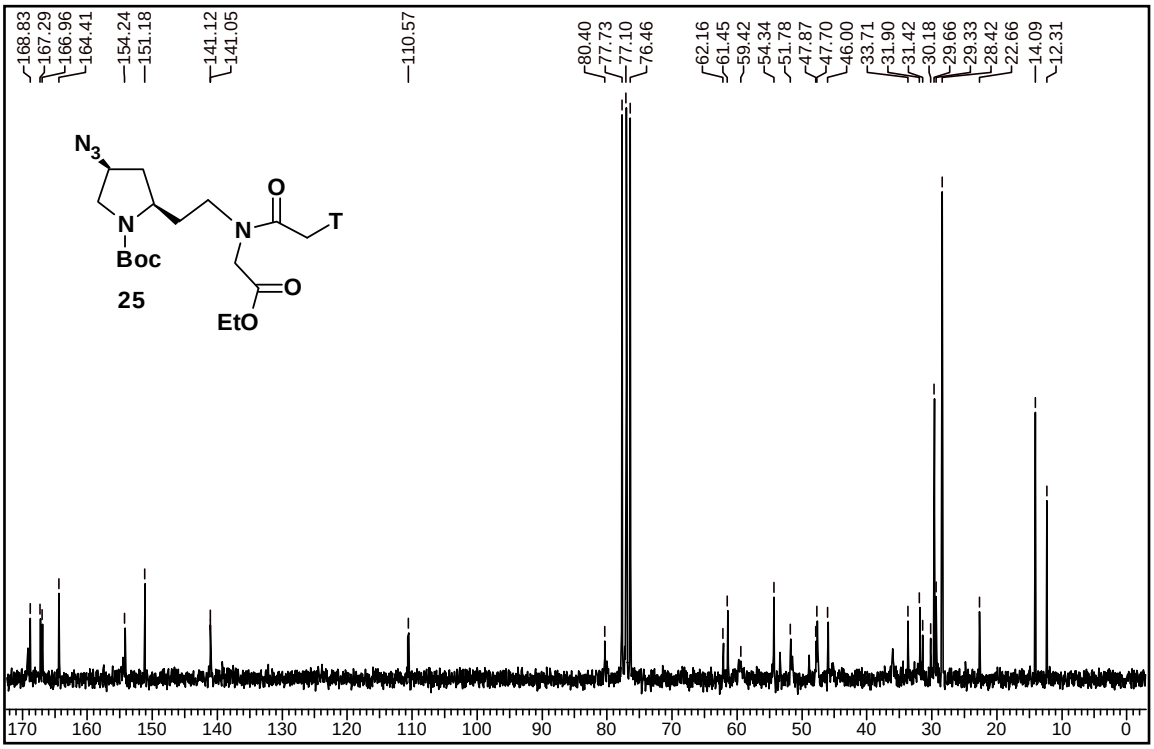

69.95
61.20
54.99
53.76
47.65
46.51
41.02
40.62
33.66
29.51
28.26
25.50
13.95
-5.02
TBSO
N
Boc
21
O
N
Cl
O
EtO
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
-5

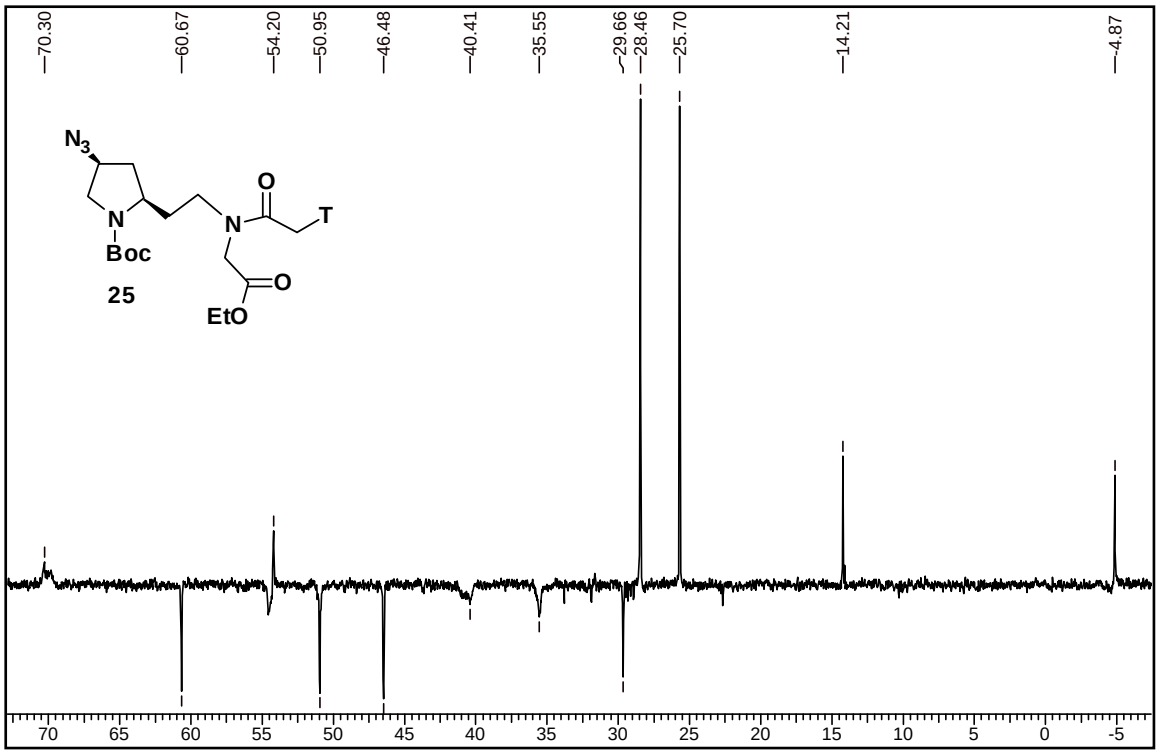

CDCl3
TMS
8.64
7.27
7.03
4.62
4.43
4.33
4.24
4.19
4.17
3.98
3.86
3.41
3.33
3.28
3.26
3.25
2.04
1.91
1.66
1.45
1.26
0.85
0.06
0.04
0.00
TBSO
N
Boc
22
O
N
T
O
EtO
0.73
1.00
1.13
5.24
3.96
4.22
3.06
9.82
6.04
9
8
7
6
5
4
3
2
1
0
-1

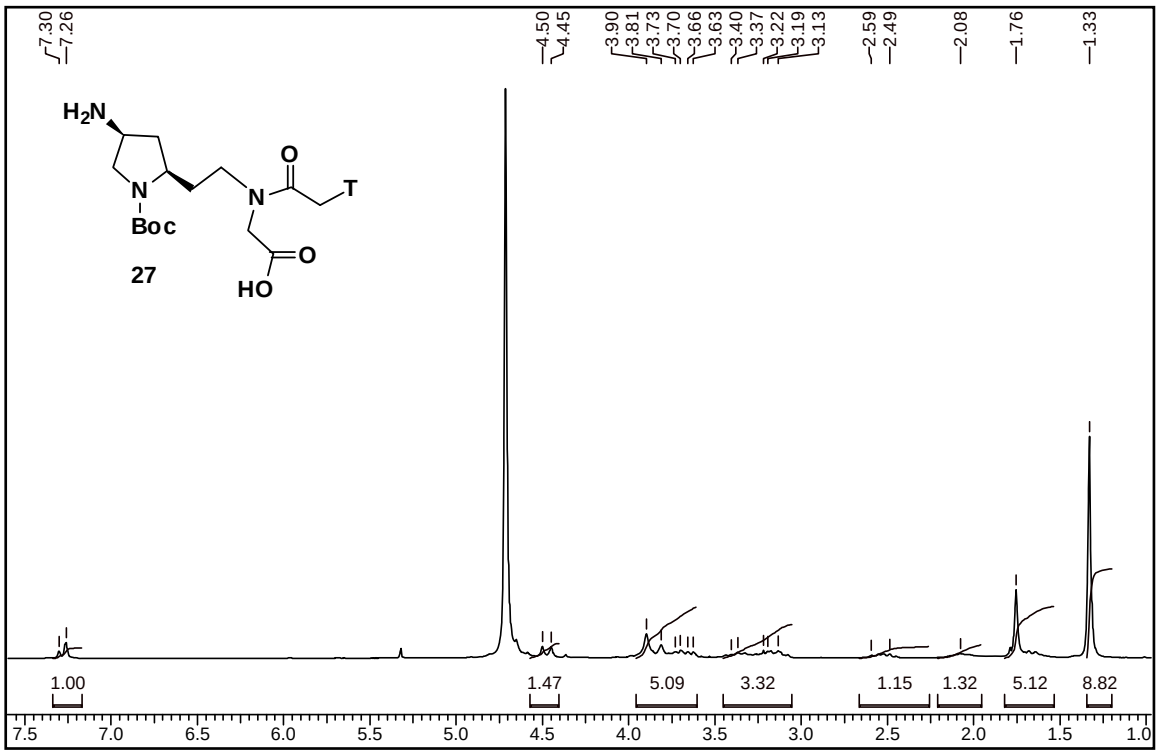

Chloroform-d
168.78
167.00
166.78
164.33
155.10
151.11
141.09
140.94
110.60
110.42
79.61
77.64
77.00
76.37
70.07
62.03
61.34
54.96
53.67
48.03
47.52
45.75
40.70
33.91
29.57
28.34
25.58
17.80
14.01
12.21
4.95
TBSO
N
Boc
22
O
N
T
O
EtO
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

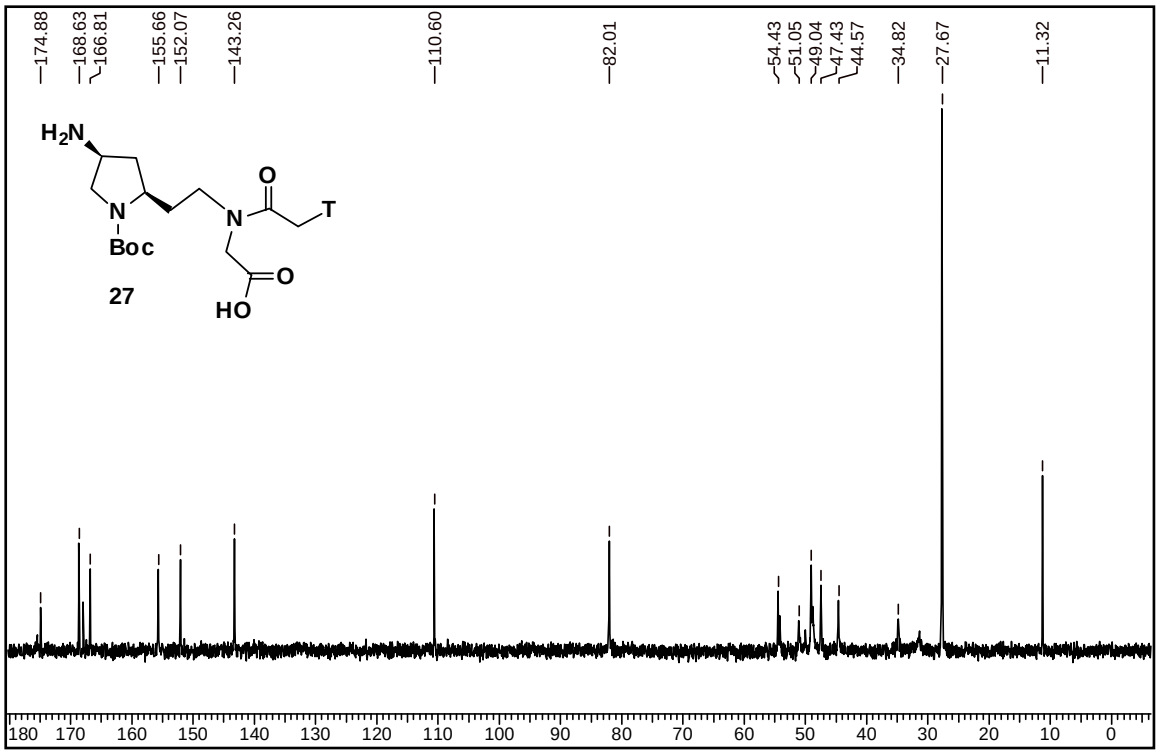

141.04
69.98
61.97
61.28
54.99
53.73
47.45
45.68
40.59
33.81
28.26
25.50
13.93
12.15
5.03
TBSO
N
Boc
22
O
N
T
O
EtO
140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

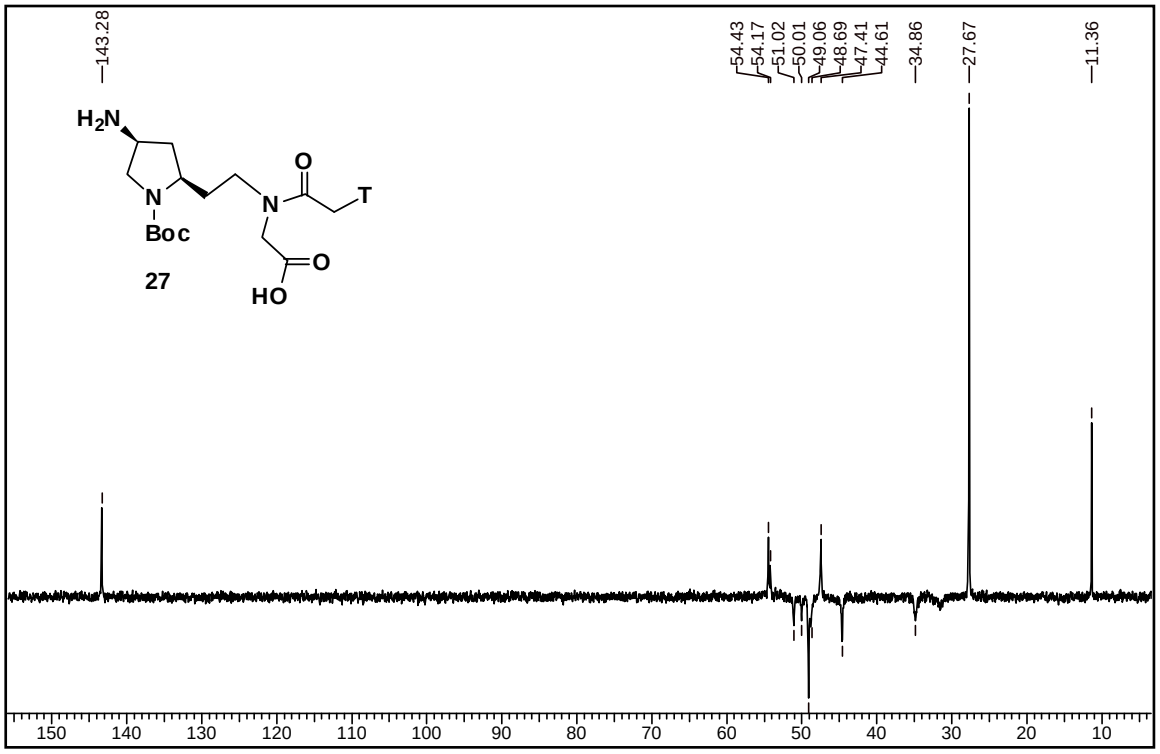

Chloroform-d
TMS
8.51
7.27
7.03
4.64
4.48
4.32
4.21
4.17
4.10
3.74
3.46
3.39
3.35
3.26
2.35
2.30
2.24
2.11
1.93
1.92
1.83
1.74
1.46
1.29
1.27
1.25
1.24
0.00
N3
N
Boc
25
O
N
T
O
EtO
0.88
1.00
2.20
5.82
2.23
3.03
3.48
1.46
10.03
9
8
7
6
5
4
3
2
1
0

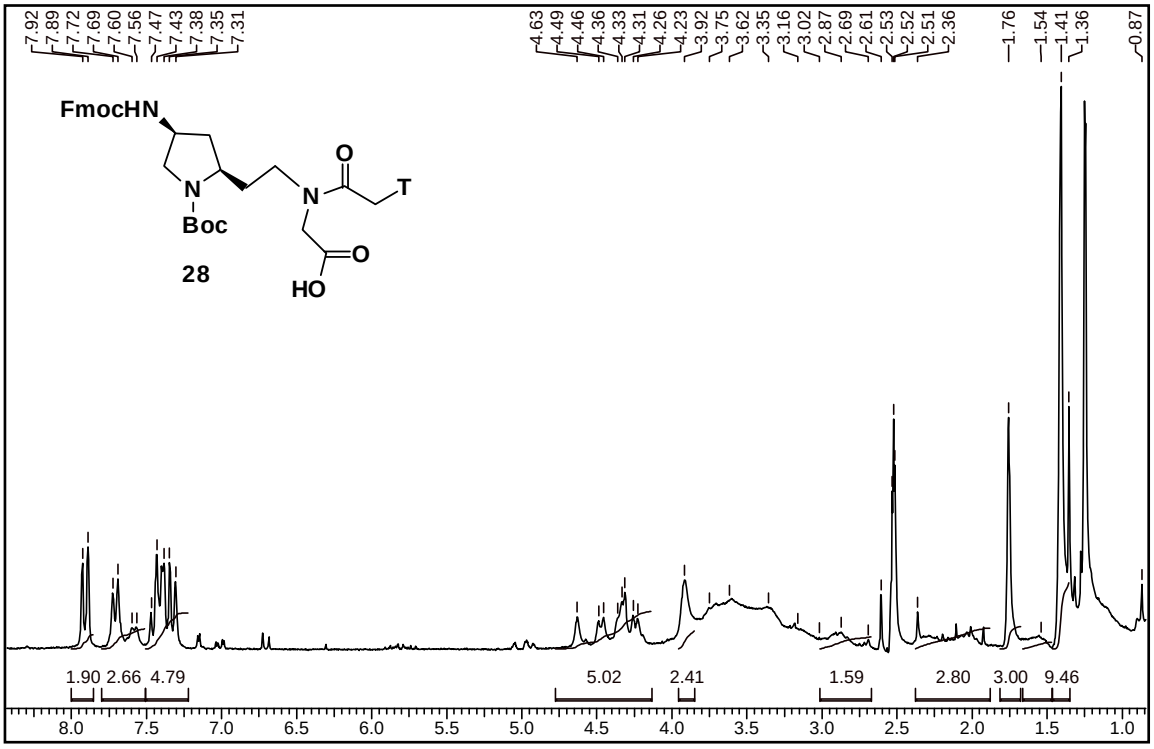

168.83
167.29
166.96
164.41
154.24
151.18
141.12
141.05
110.57
80.40
77.73
77.10
76.46
62.16
61.45
59.42
54.34
51.78
47.87
47.70
46.00
33.71
31.90
31.42
30.18
29.66
29.33
28.42
22.66
14.09
12.31
N3
N
Boc
25
O
N
T
O
EtO
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

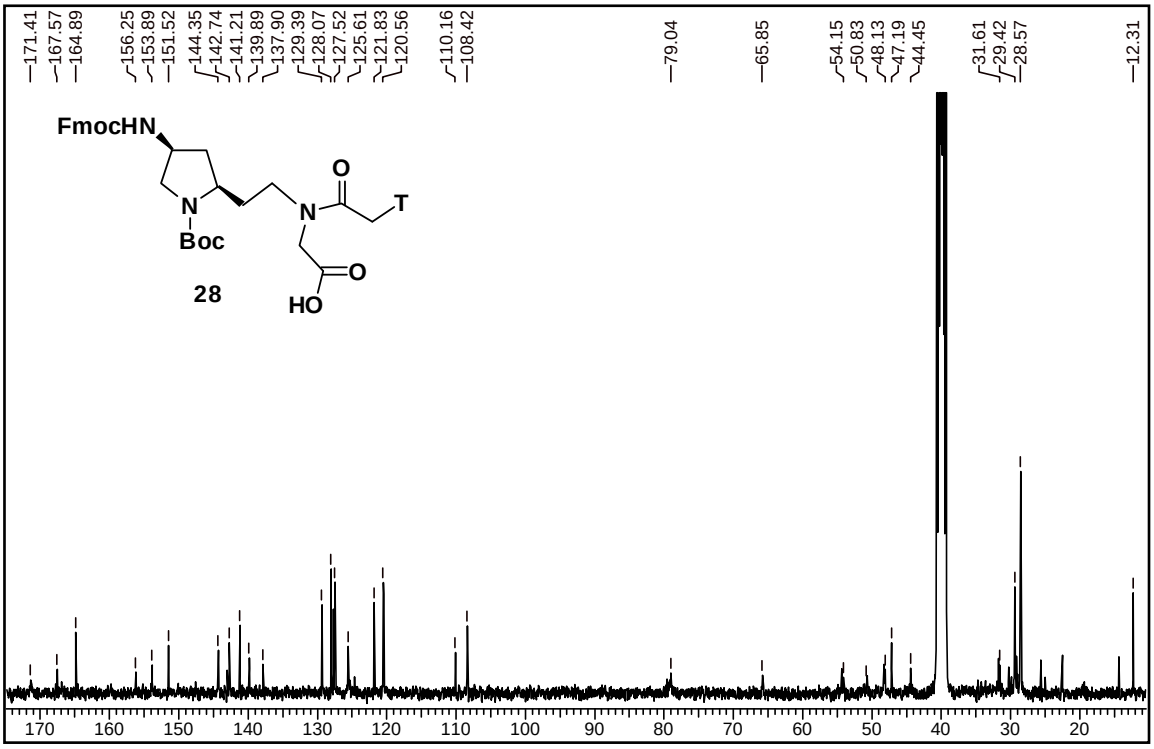

70.30
60.67
54.20
50.95
46.48
40.41
35.55
29.66
28.46
25.70
14.21
4.87
N3
N
Boc
25
O
N
T
O
EtO
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
-5

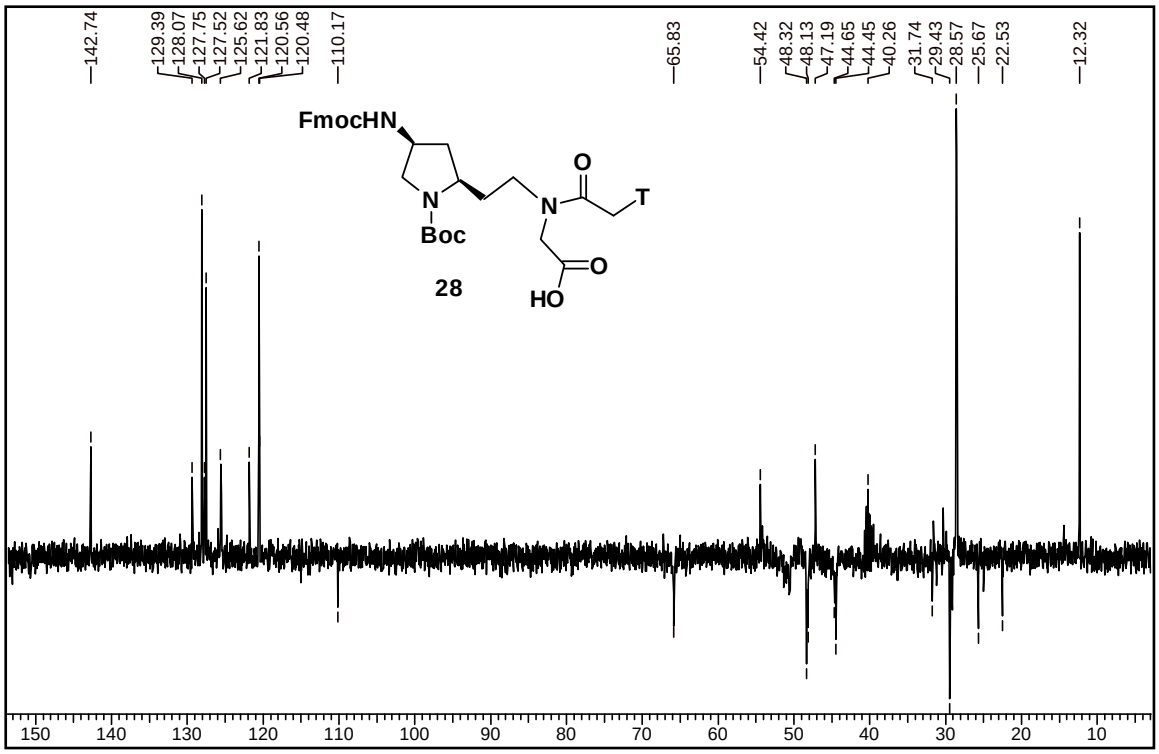

7.30
7.26
4.50
4.45
3.90
3.81
3.73
3.70
3.66
3.63
3.40
3.37
3.22
3.19
3.13
2.59
2.49
2.08
1.76
1.33
H2N
N
Boc
27
O
N
T
O
HO
1.00
1.47
5.09
3.32
1.15
1.32
5.12
8.82
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

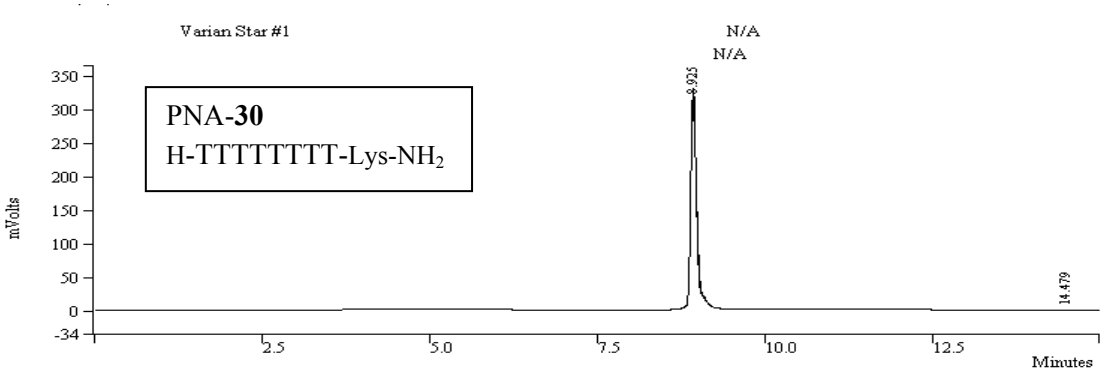

174.88
168.63
166.81
155.66
152.07
143.26
110.60
82.01
54.43
51.05
49.04
47.43
44.57
34.82
27.67
11.32
H2N
N
Boc
O
N
T
O
HO
27
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

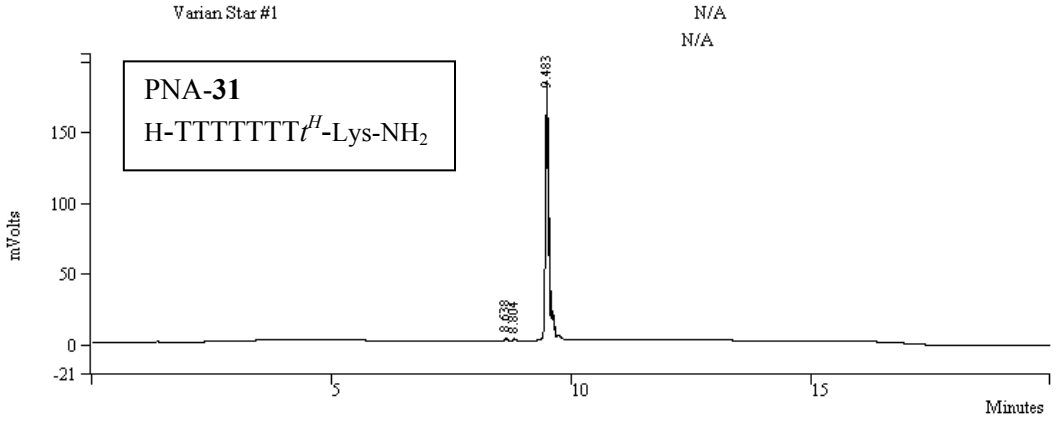

143.28
54.43
54.17
51.02
50.01
49.06
48.69
47.41
44.61
34.86
27.67
11.36
H2N
N
Boc
O
N
T
O
HO
27
150 140 130 120 110 100 90 80 70 60 50 40 30 20 10

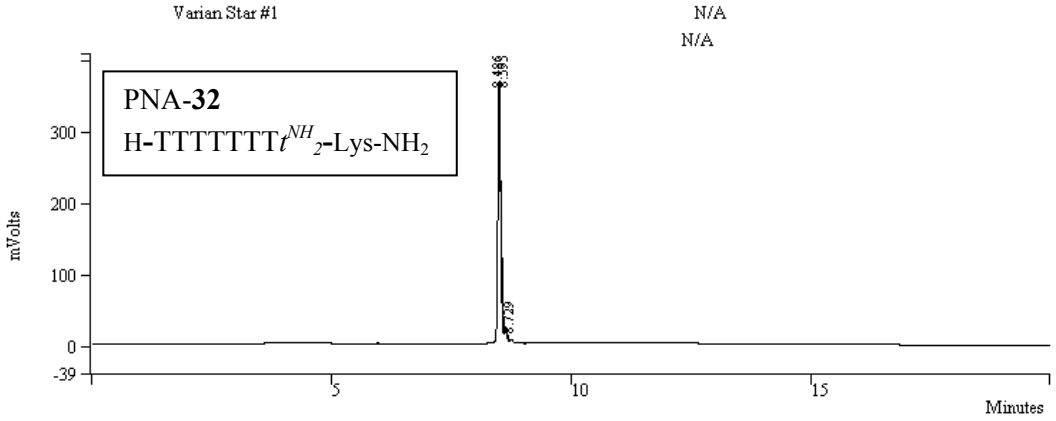

7.92
7.89
7.72
7.69
7.60
7.56
7.47
7.43
7.38
7.35
7.31
4.63
4.49
4.46
4.36
4.33
4.31
4.26
4.23
3.92
3.75
3.62
3.35
3.16
3.02
2.87
2.69
2.61
2.53
2.52
2.51
2.36
1.76
1.54
1.41
1.36
0.87
FmocHN
N
Boc
N
O
T
O
HO
28
1.90
2.66
4.79
5.02
2.41
1.59
2.80
3.00
9.46
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

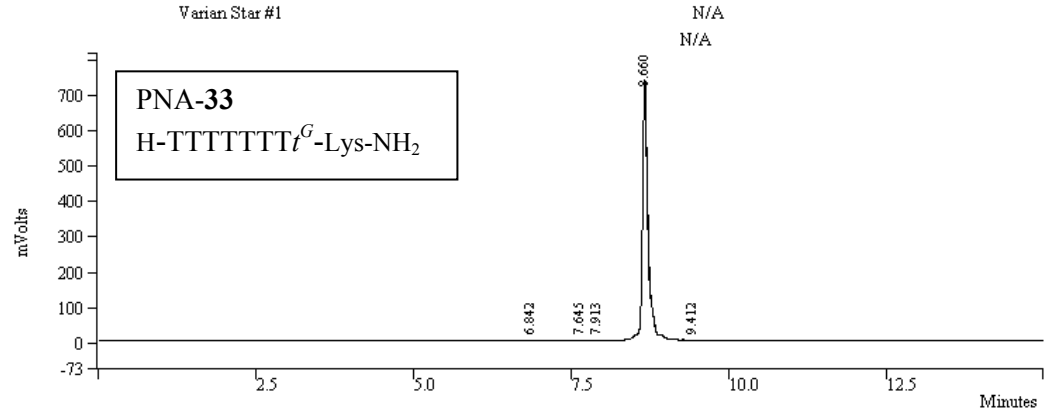

171.41
167.57
164.89
156.25
153.89
151.52
144.35
142.74
141.21
139.89
137.90
129.39
128.07
127.52
125.61
121.83
120.56
110.16
108.42
79.04
65.85
54.15
50.83
48.13
47.19
44.45
31.61
29.42
28.57
12.31
FmocHN
N
Boc
N
O
T
O
HO
28
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20

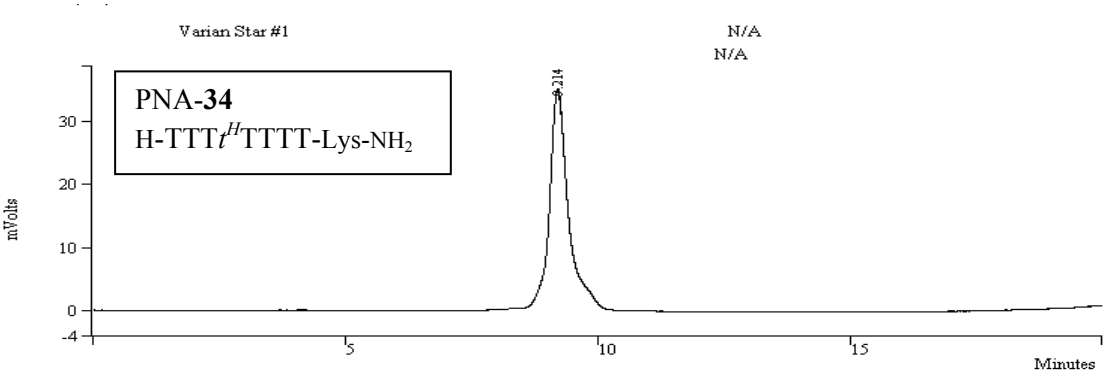

142.74
129.39
128.07
127.75
127.52
125.62
121.83
120.56
120.48
110.17
65.83
54.42
48.32
48.13
47.19
44.65
44.45
40.26
31.74
29.43
28.57
25.67
22.53
12.32
FmocHN
N
Boc
N
O
T
O
HO
28
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

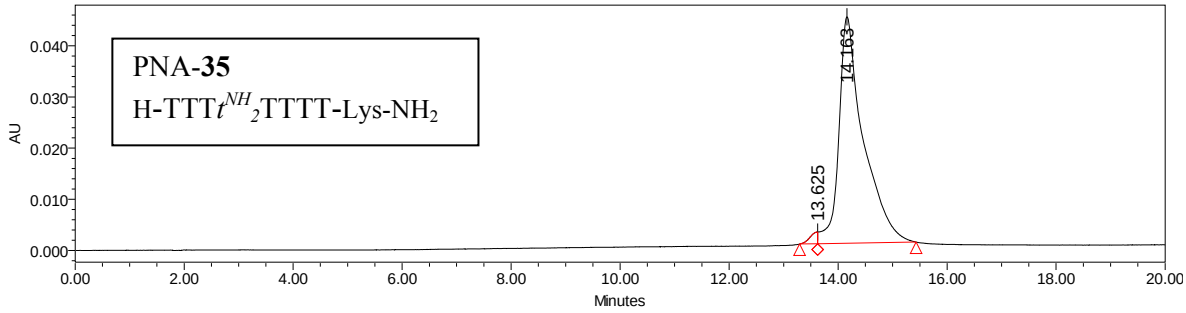

Varian Star #1
N/A
N/A
PNA-30
H-TTTTTTTT-Lys-NH2
mVolts
Minutes
8.925
14.479

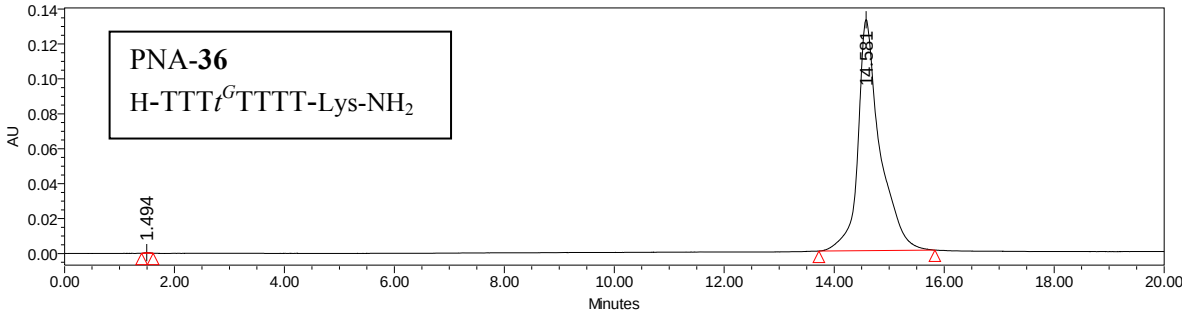

Varian Star #1
N/A
N/A
PNA-31
H-TTTTTTTtH-Lys-NH2
mVolts
Minutes
9.483

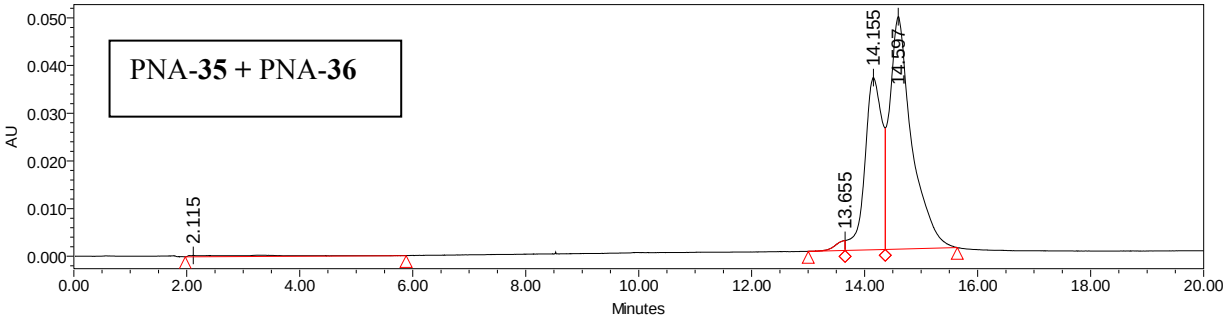

Varian Star #1
N/A
N/A
PNA-32
H-TTTTTTTtNH2-Lys-NH2
mVolts
Minutes

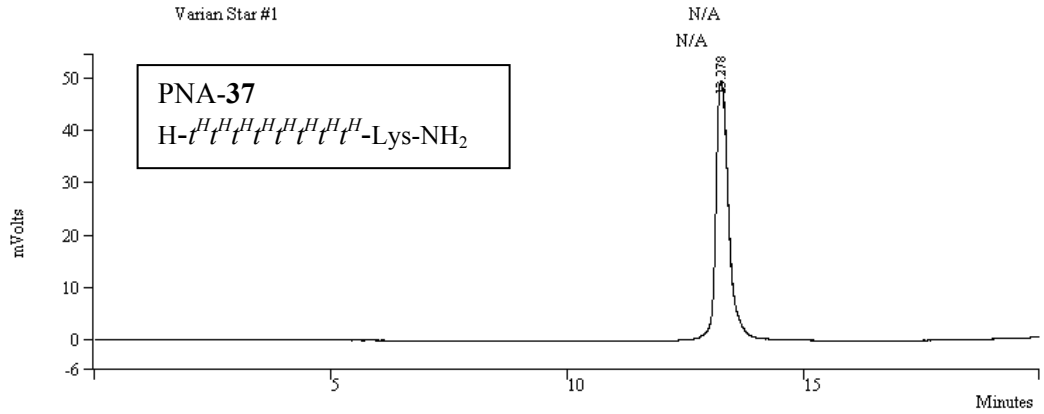

Varian Star #1
N/A
N/A
PNA-33
H-TTTTTTTtG-Lys-NH2
mVolts
Minutes
8.660
6.842
7.645
7.913
9.412

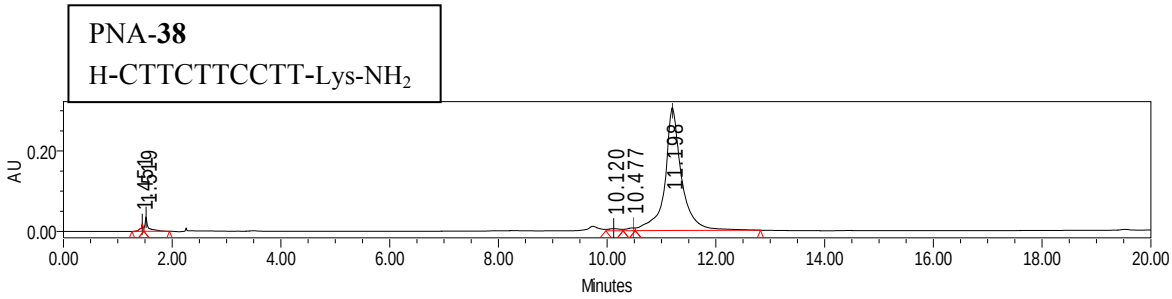

Varian Star #1
N/A
N/A
PNA-34
H-TTTtHTTTT-Lys-NH2
mVolts
Minutes

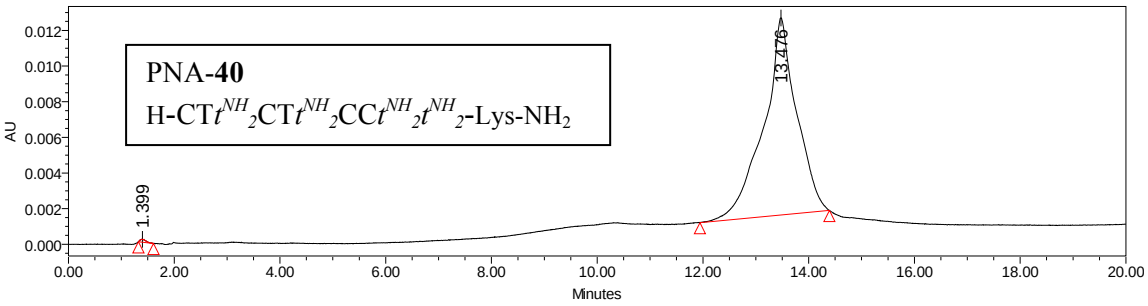

PNA-35
H-TTTtNH2TTTT-Lys-NH2
AU
14.163
13.625
Minutes

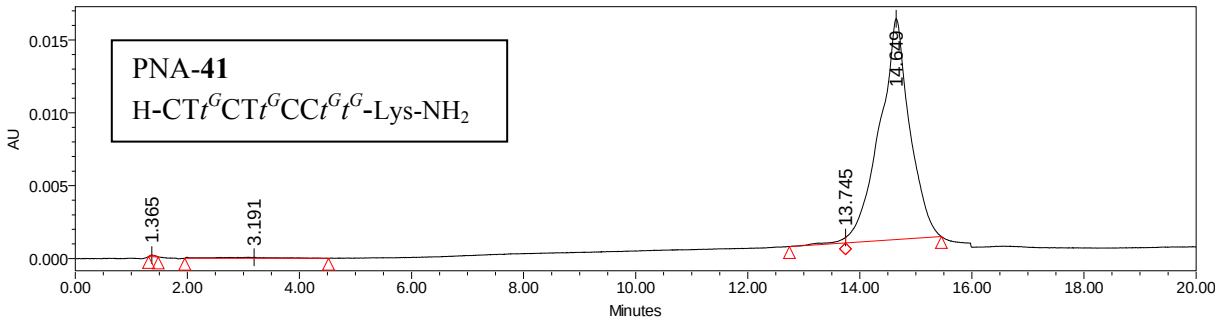

PNA-36
H-TTTtGTTTT-Lys-NH2
AU
14.581
1.494
Minutes

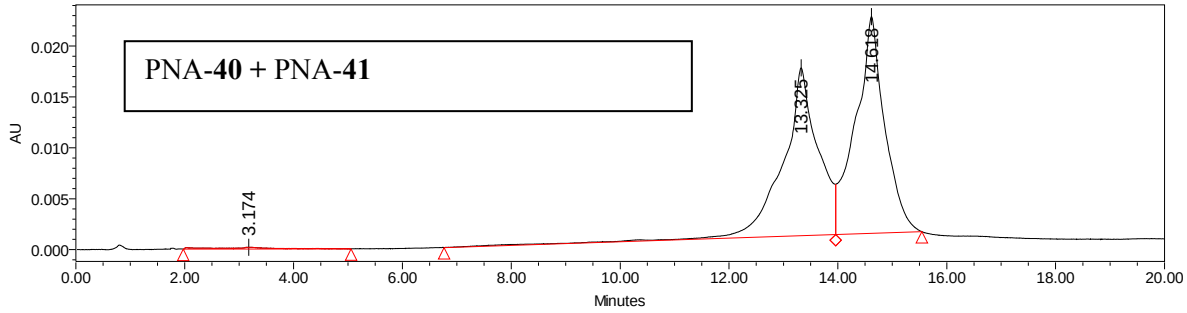

PNA-35 + PNA-36
AU
14.155
14.597
13.655
2.115
Minutes

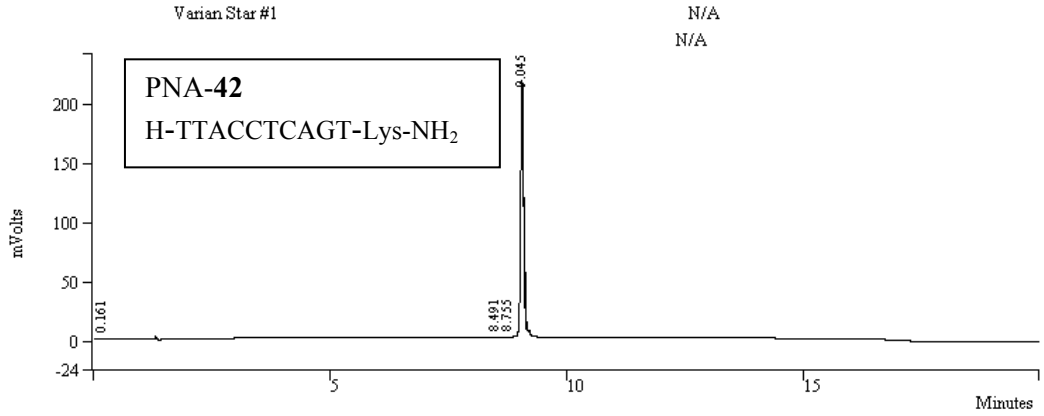

Varian Star #1
N/A
N/A
PNA-37
H-$t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}$-Lys-$NH_2$
mVolts
Minutes

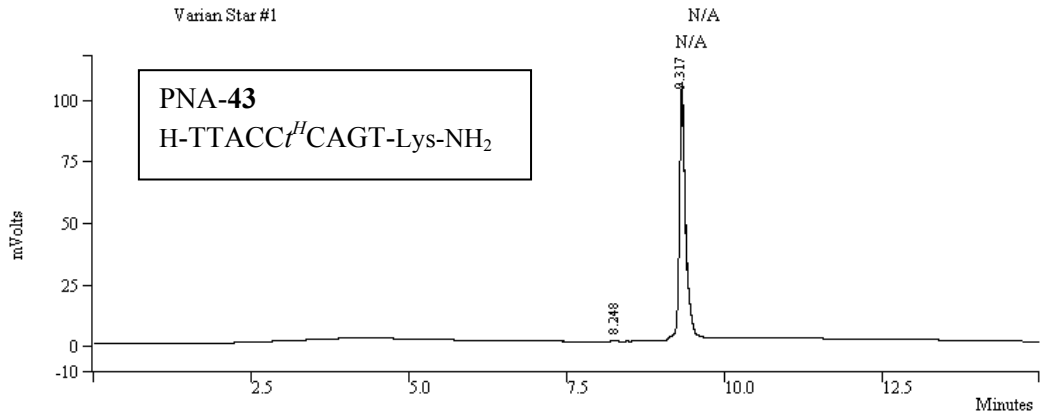

PNA-38
H-CTTCTTCCTT-Lys-NH_2
AU
Minutes

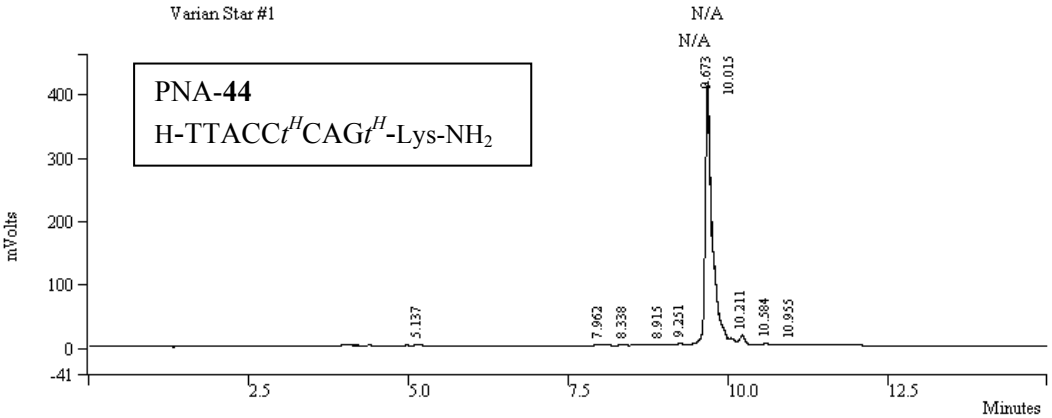

PNA-40
H-CTt^{NH_2}CTt^{NH_2}CC$t^{NH_2}t^{NH_2}$-Lys-NH_2
AU
Minutes

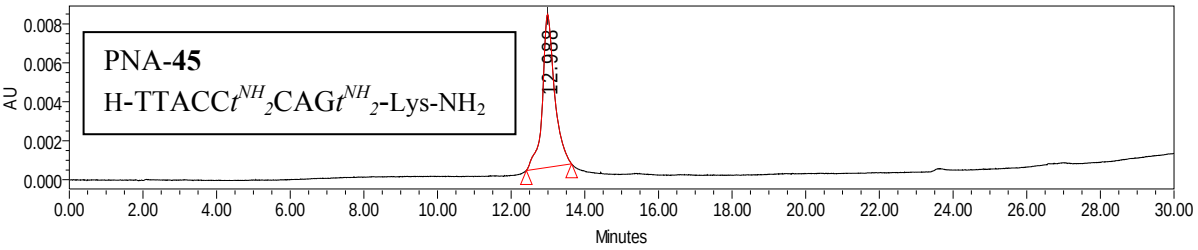

PNA-41
H-CTt^{G}CTt^{G}CC$t^{G}t^{G}$-Lys-NH_2
AU
Minutes

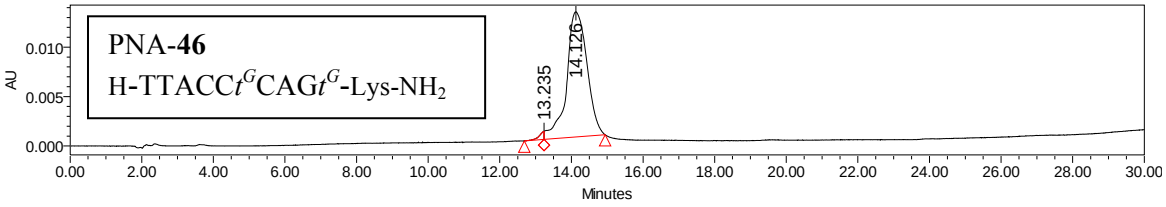

PNA-40 + PNA-41
AU
Minutes
3.174
13.325
14.618

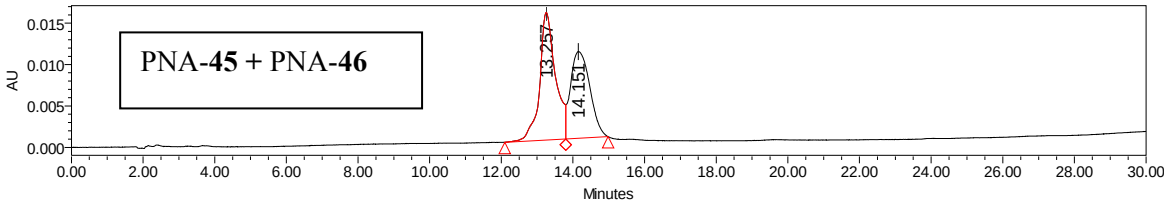

Varian Star #1
N/A
N/A
PNA-42
H-TTACCTCAGT-Lys-NH2
mVolts
Minutes
0.161
8.491
8.755
9.045

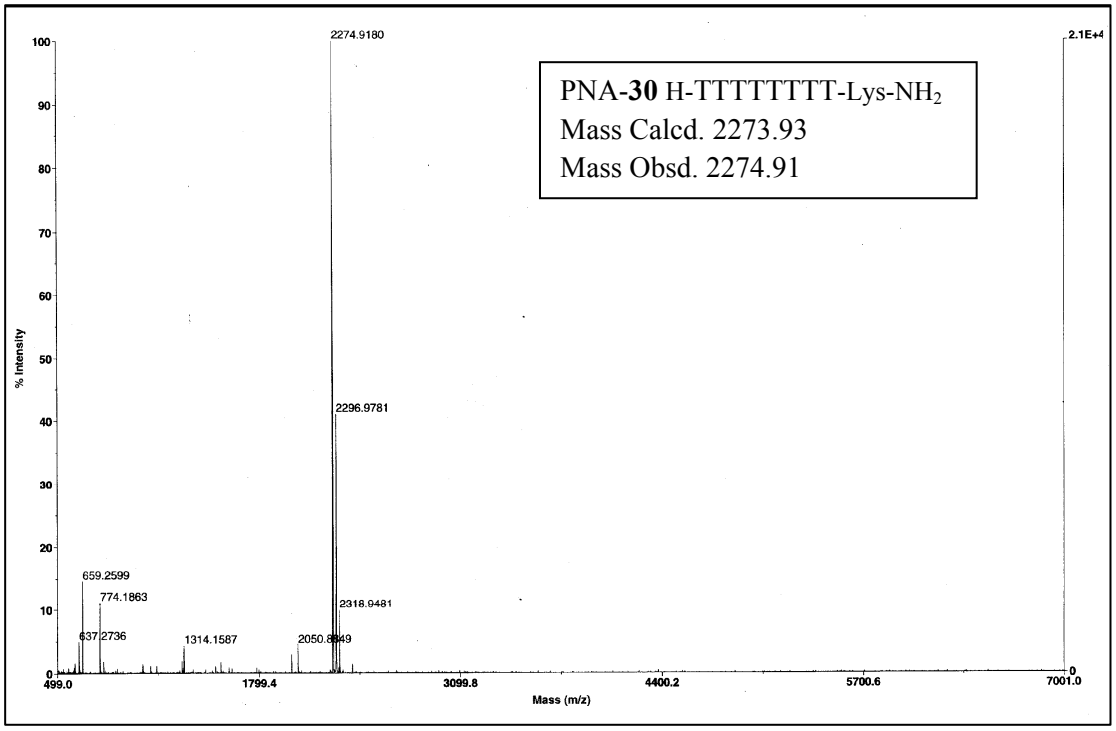

Varian Star #1
N/A
N/A
PNA-43
H-TTACCtHCAGT-Lys-NH2
mVolts
Minutes
8.248
9.317

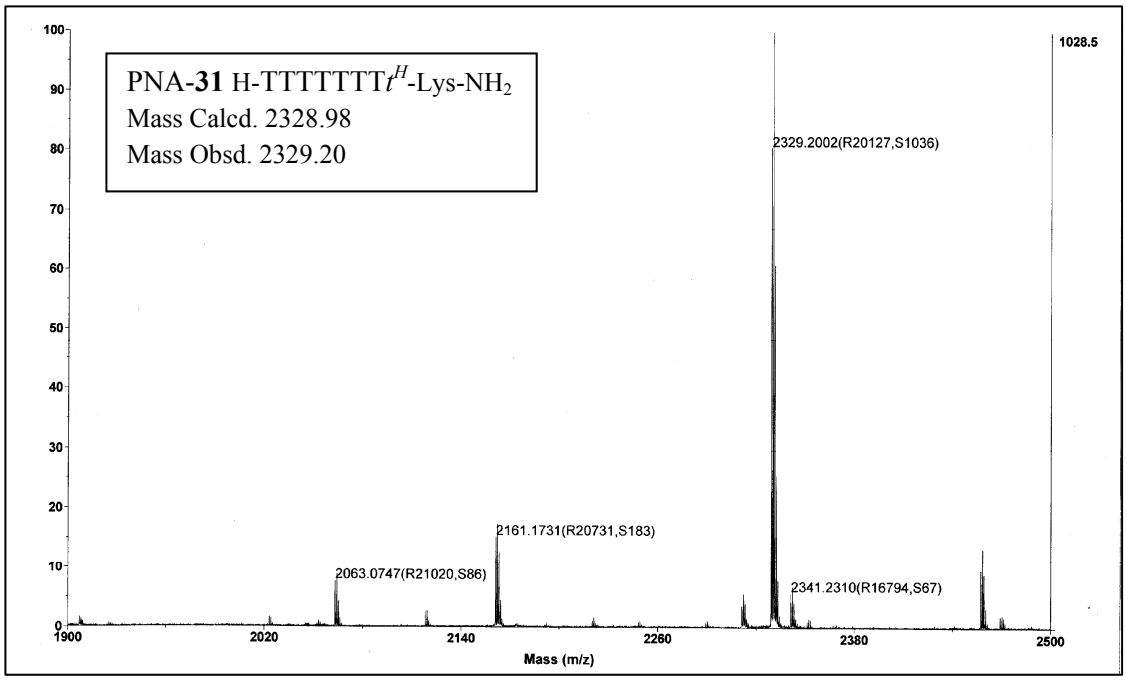

Varian Star #1
N/A
N/A
PNA-44
H-TTACCtHCAGtH-Lys-NH2
mVolts
Minutes
5.137
7.962
8.338
8.915
9.251
9.673
10.015
10.211
10.584
10.955

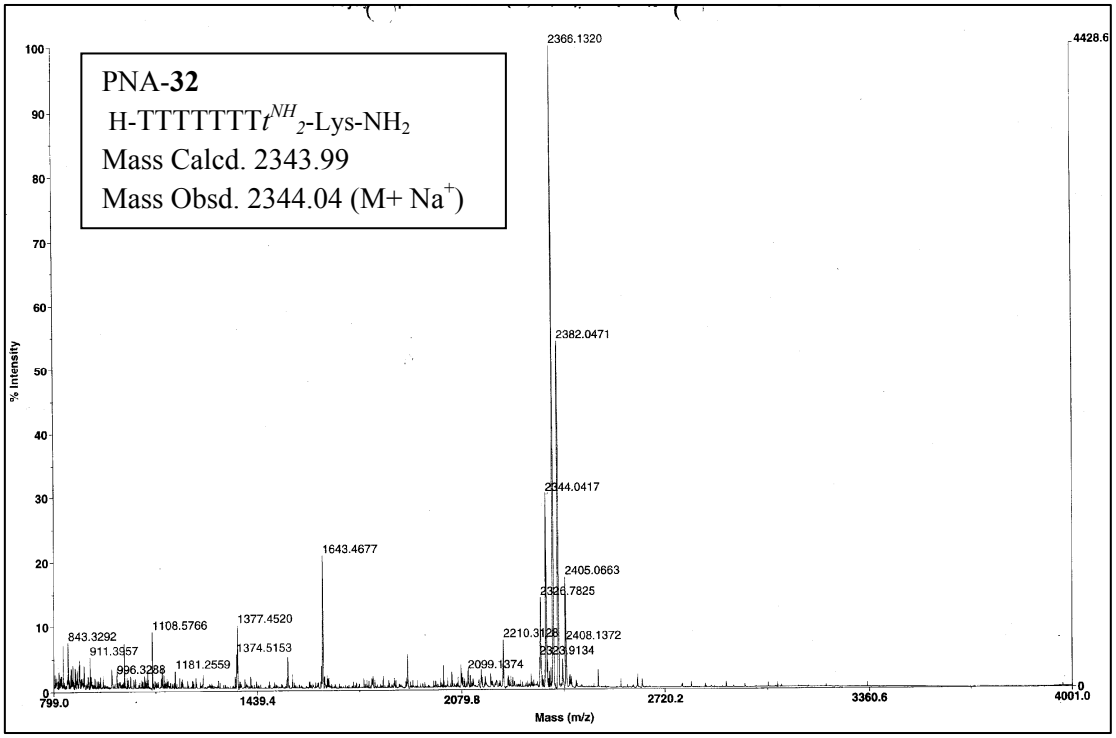

PNA-45
H-TTACCt^NH2CAGt^NH2-Lys-NH2
12.988
AU
Minutes

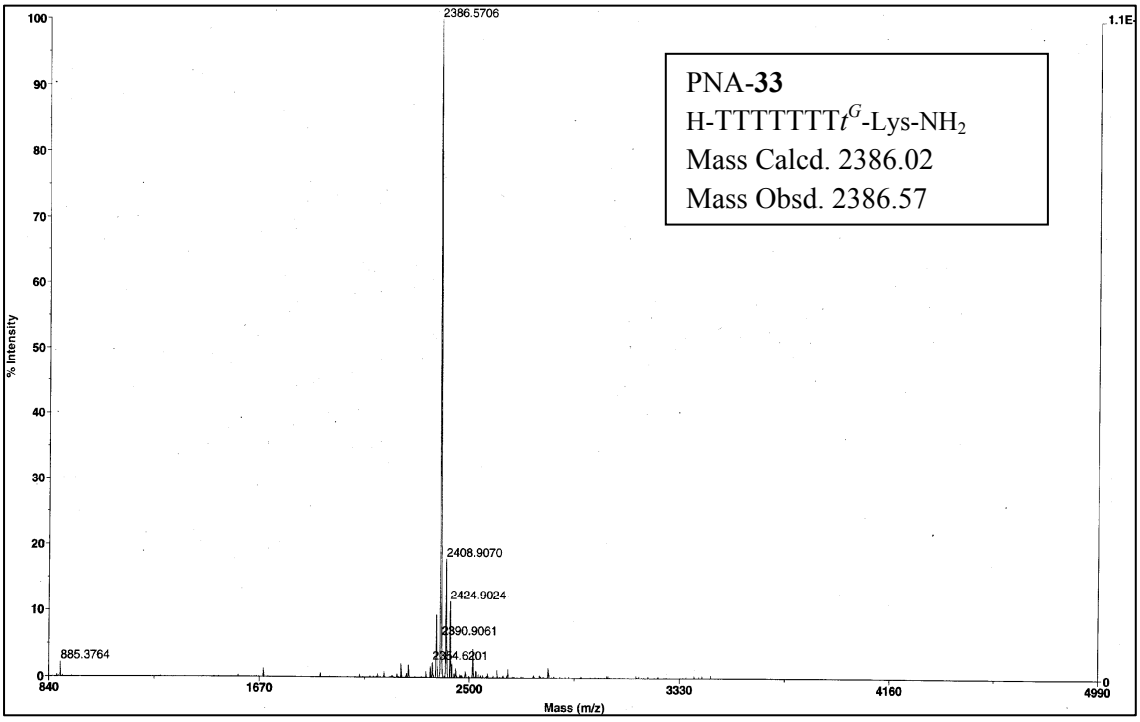

PNA-46
H-TTACCt^GCAGt^G-Lys-NH2
13.235
14.126
AU
Minutes

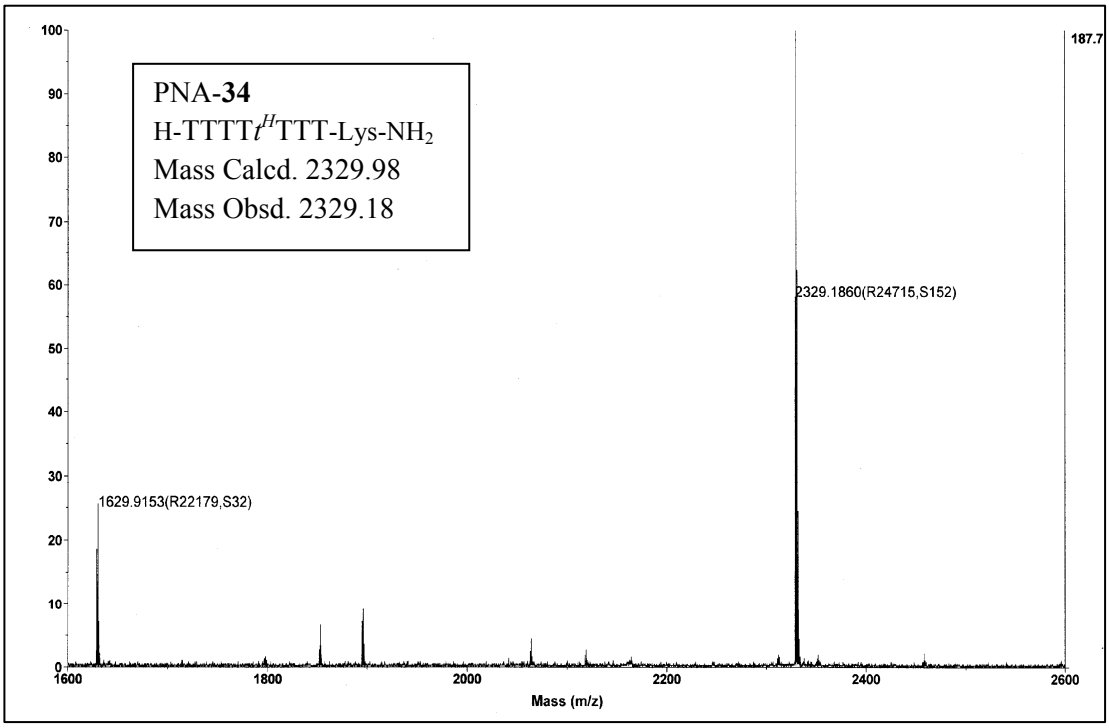

PNA-45 + PNA-46
13.257
14.151
AU
Minutes

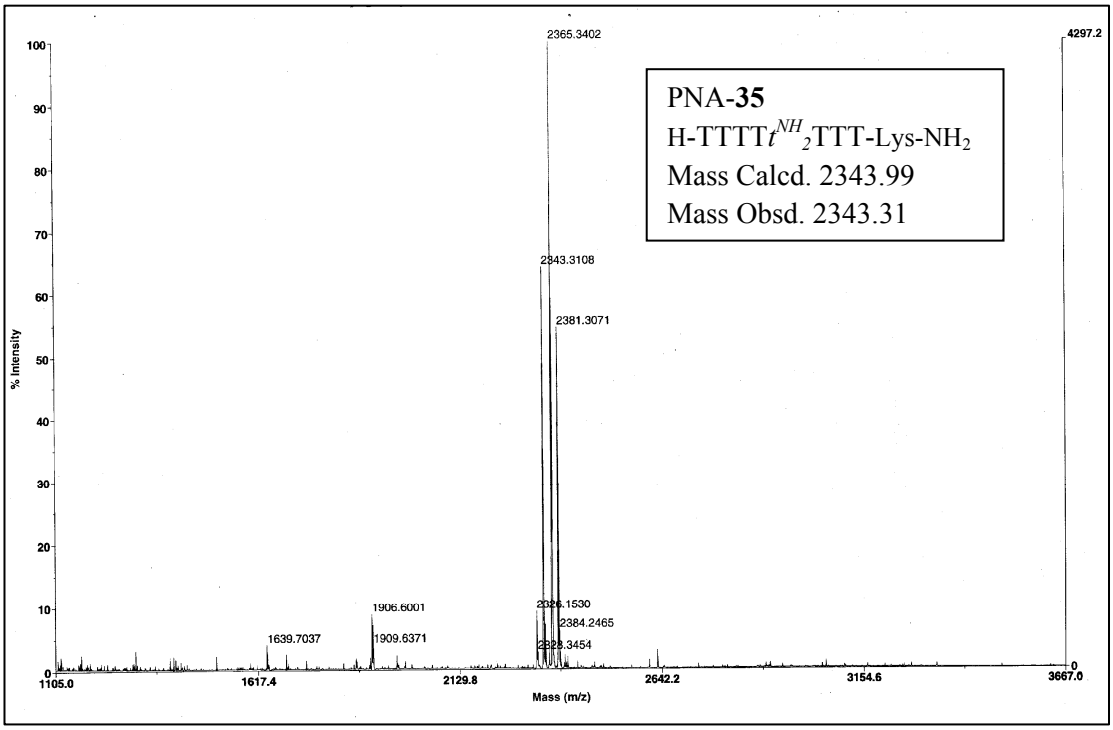

PNA-**30** H-TTTTTTTT-Lys-NH_2
Mass Calcd. 2273.93
Mass Obsd. 2274.91
2274.9180
2296.9781
659.2599
774.1863
637.2736
1314.1587
2318.9481
2050.8849
% Intensity
Mass (m/z)
2.1E+4

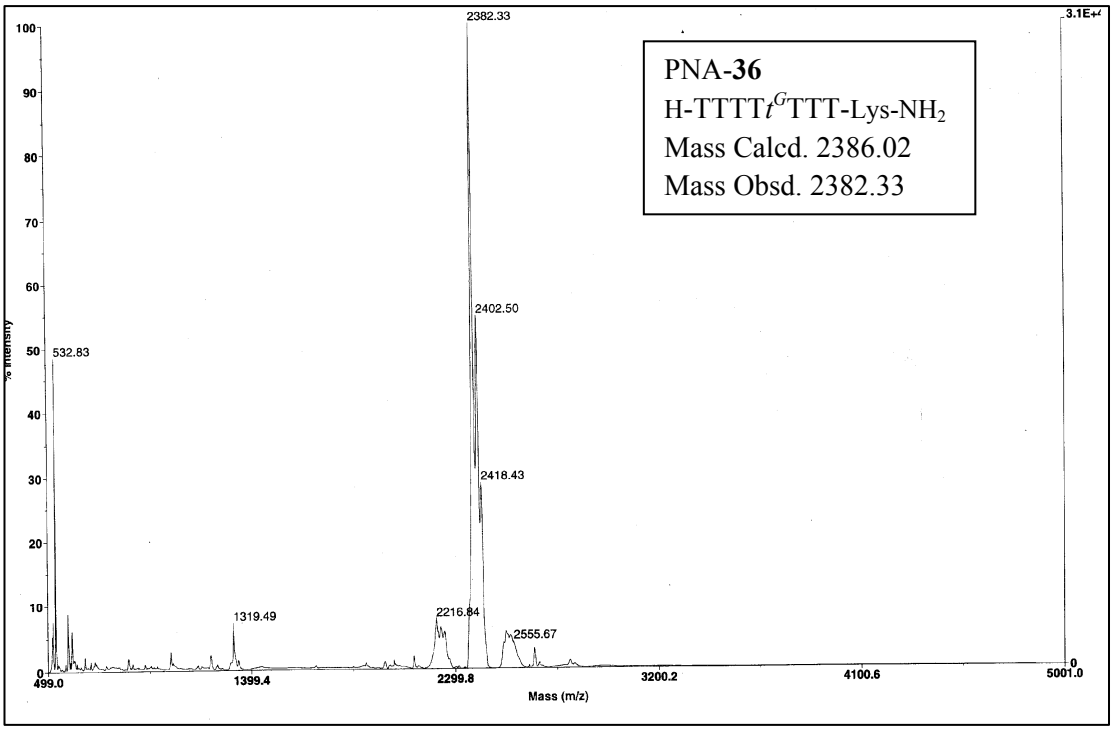

PNA-**31** H-TTTTTTTt^{H}-Lys-NH_2
Mass Calcd. 2328.98
Mass Obsd. 2329.20
1028.5
2329.2002(R20127,S1036)
2161.1731(R20731,S183)
2063.0747(R21020,S86)
2341.2310(R16794,S67)
Mass (m/z)

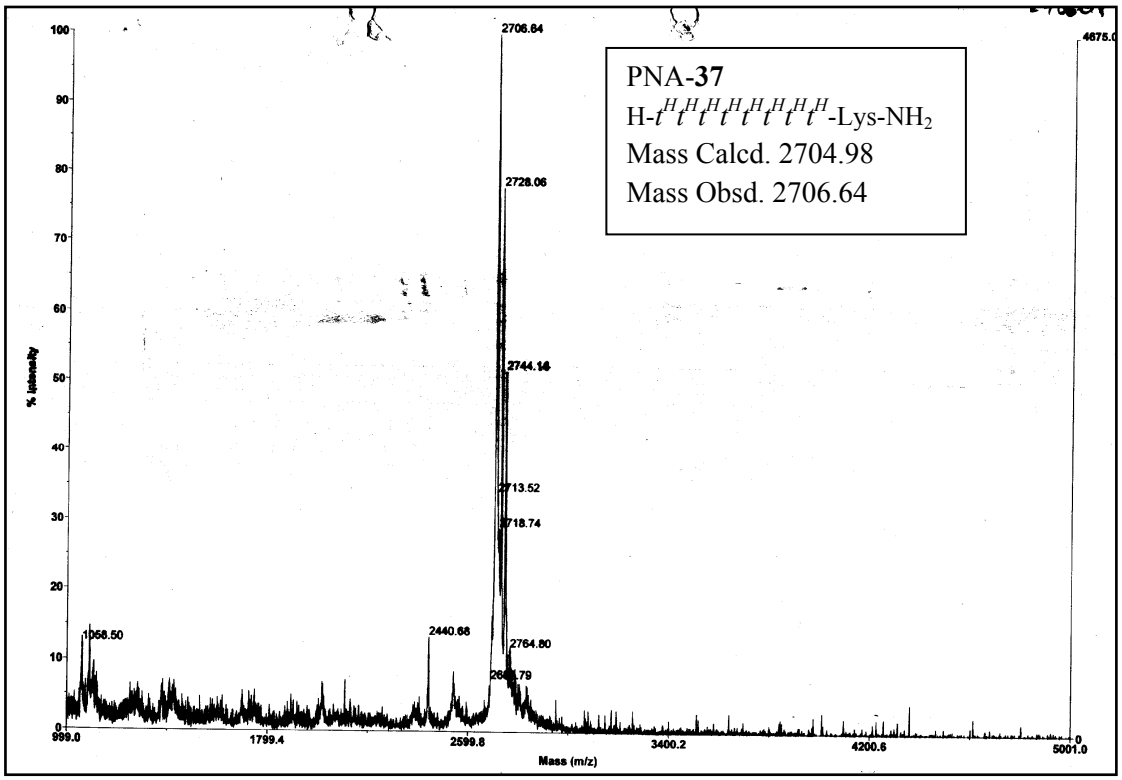

PNA-32
H-TTTTTTTt^NH2-Lys-NH2
Mass Calcd. 2343.99
Mass Obsd. 2344.04 (M+ Na+)
% Intensity
Mass (m/z)
2366.1320
4428.6
2382.0471
2344.0417
1643.4677
2405.0663
2326.7825
1108.5766
1377.4520
843.3292
911.3957
1374.5153
996.3288
1181.2559
2210.3128
2408.1372
2323.9134
2099.1374
799.0
1439.4
2079.8
2720.2
3360.6
4001.0

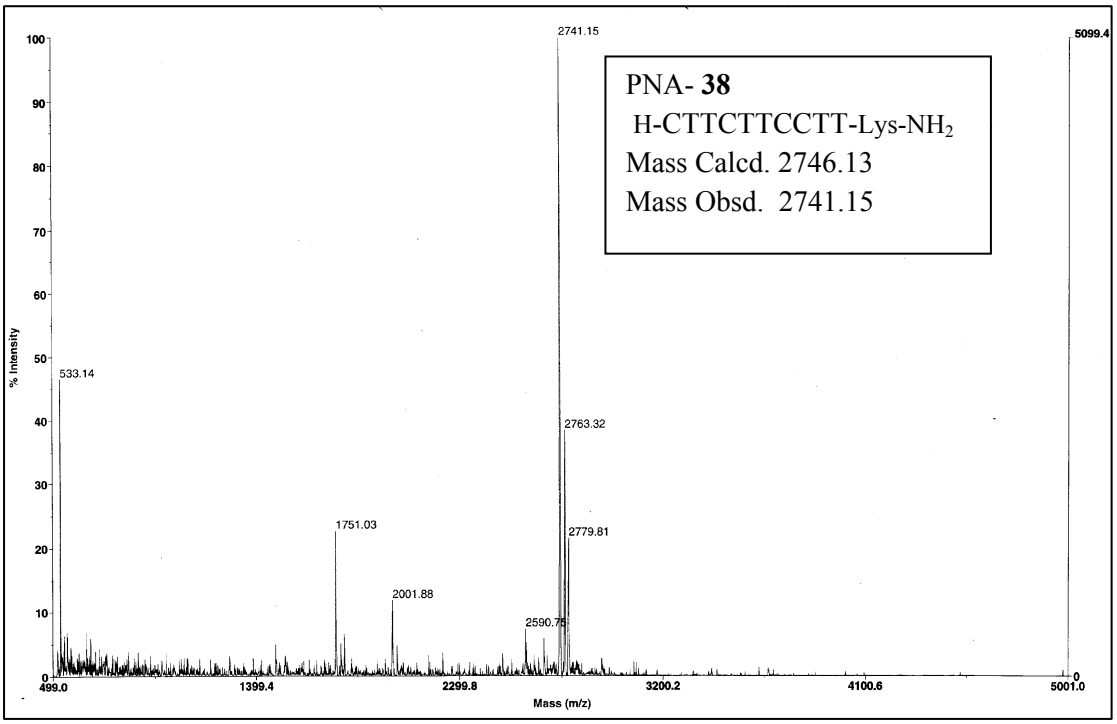

PNA-33
H-TTTTTTTt^G-Lys-NH2
Mass Calcd. 2386.02
Mass Obsd. 2386.57
% Intensity
Mass (m/z)
2386.5706
2408.9070
2424.9024
2390.9061
885.3764
840
1670
2500
3330
4160
4990

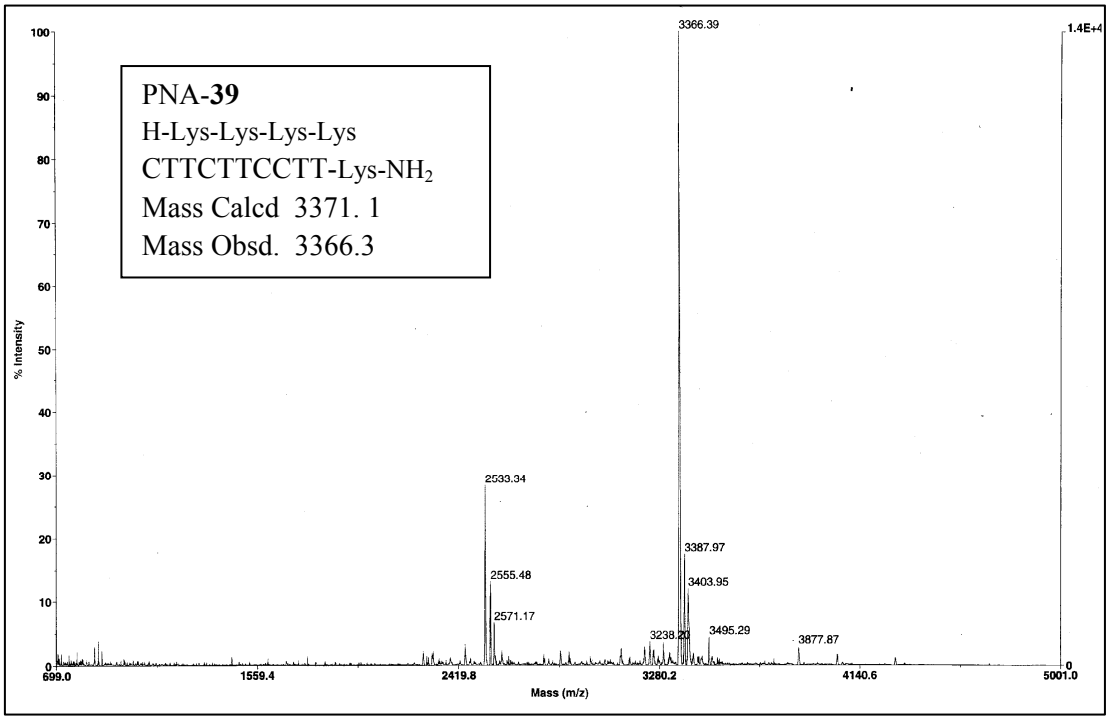

PNA-34
H-TTTTt^HTTT-Lys-NH2
Mass Calcd. 2329.98
Mass Obsd. 2329.18
2329.1860(R24715,S152)
1629.9153(R22179,S32)
187.7
Mass (m/z)

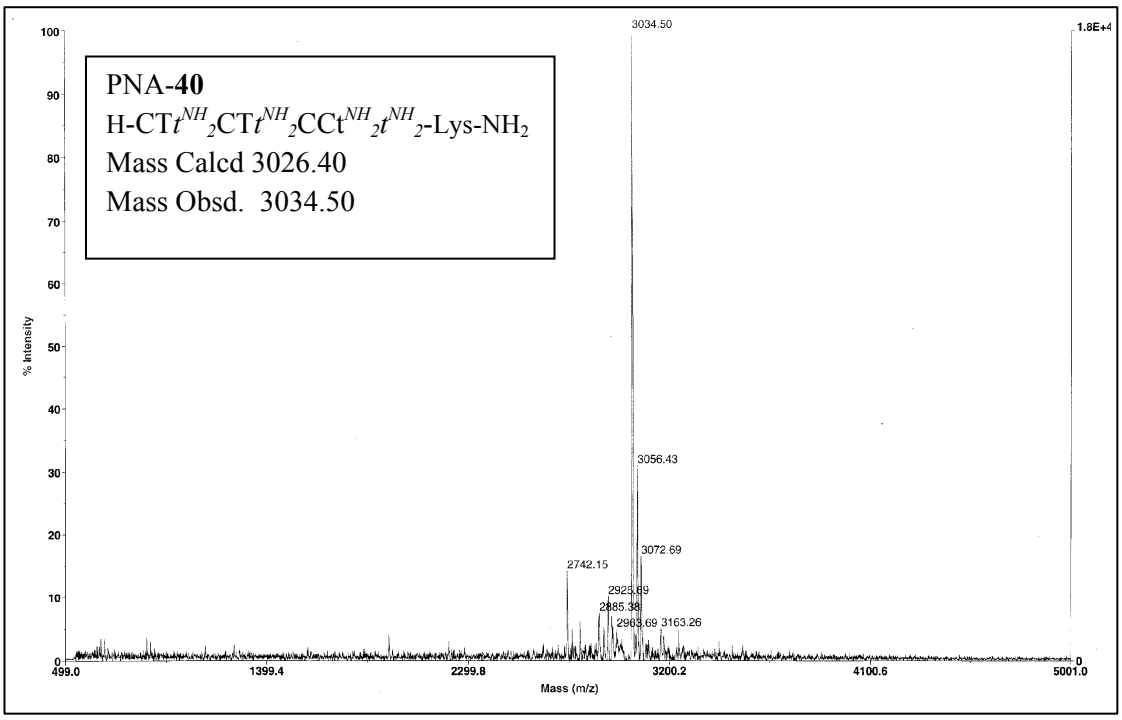

PNA-35
H-TTTTt^NH2TTT-Lys-NH2
Mass Calcd. 2343.99
Mass Obsd. 2343.31
2365.3402
2343.3108
2381.3071
4297.2
1906.6001
1639.7037
1909.6371
2384.2465
% Intensity
Mass (m/z)

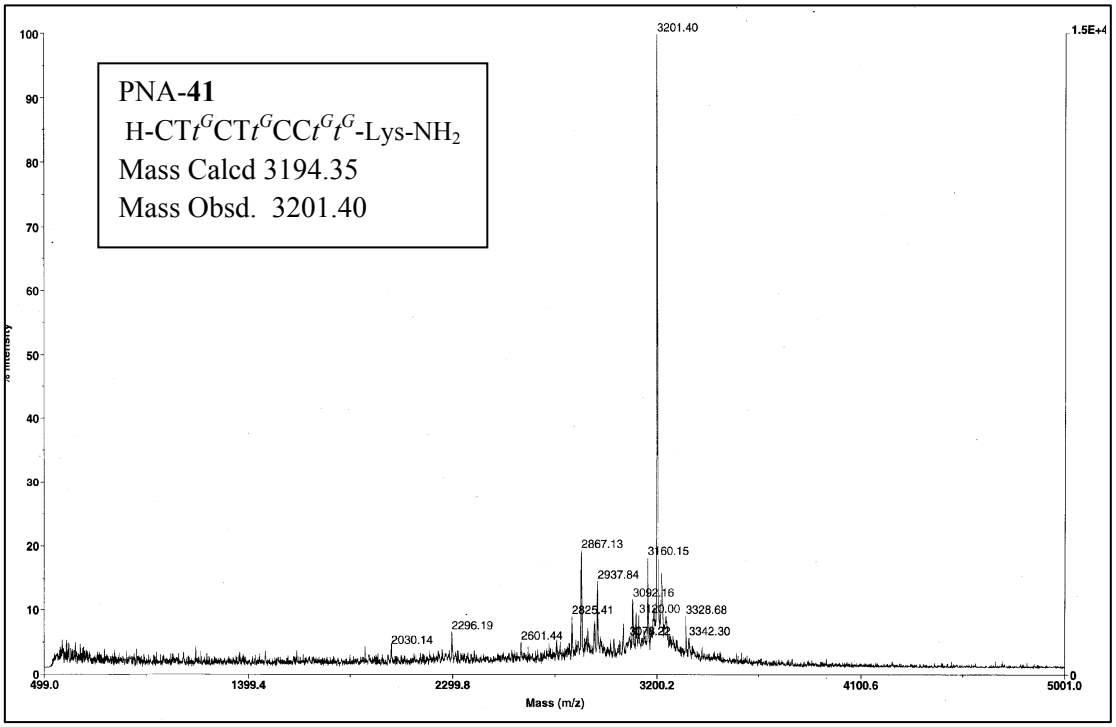

PNA-**36**
H-TTTTt^{G}TTT-Lys-NH_2
Mass Calcd. 2386.02
Mass Obsd. 2382.33
2382.33
2402.50
2418.43
532.83
1319.49
2216.84
2555.67
% Intensity
Mass (m/z)

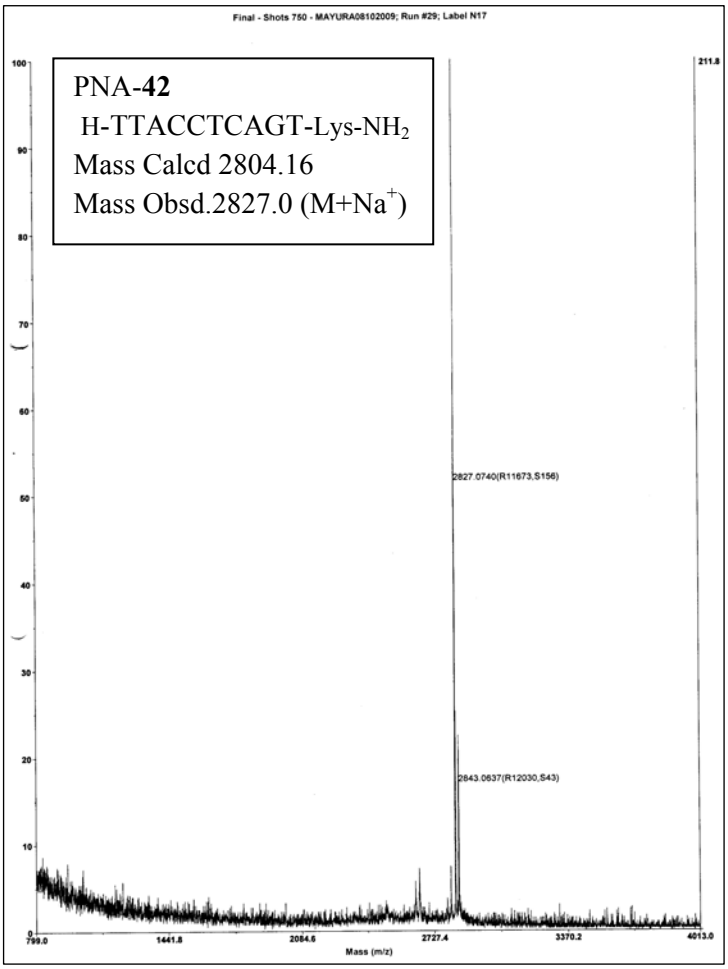

PNA-**37**
H-$t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}t^{H}$-Lys-$NH_2$
Mass Calcd. 2704.98
Mass Obsd. 2706.64
2706.64
2728.06
2744.14
2713.52
2718.74
1058.50
2440.66
2764.80
% Intensity
Mass (m/z)

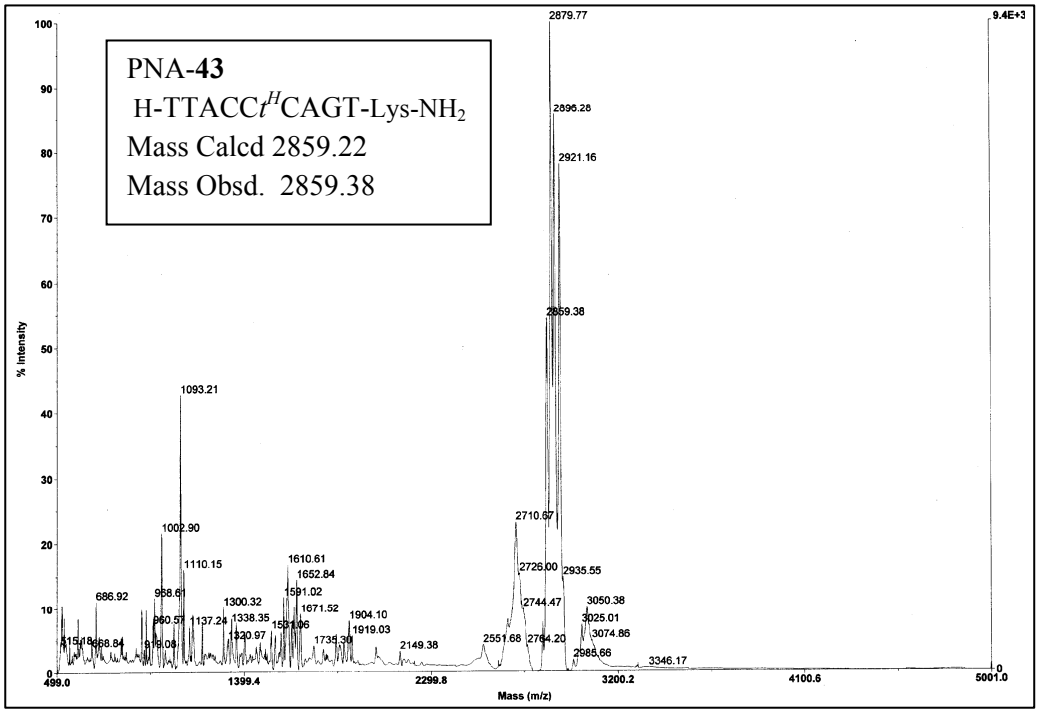

PNA- 38
H-CTTCTTCCTT-Lys-NH2
Mass Calcd. 2746.13
Mass Obsd. 2741.15
% Intensity
Mass (m/z)
2741.15
5099.4
533.14
2763.32
1751.03
2779.81
2001.88
2590.75
499.0
1399.4
2299.8
3200.2
4100.6
5001.0

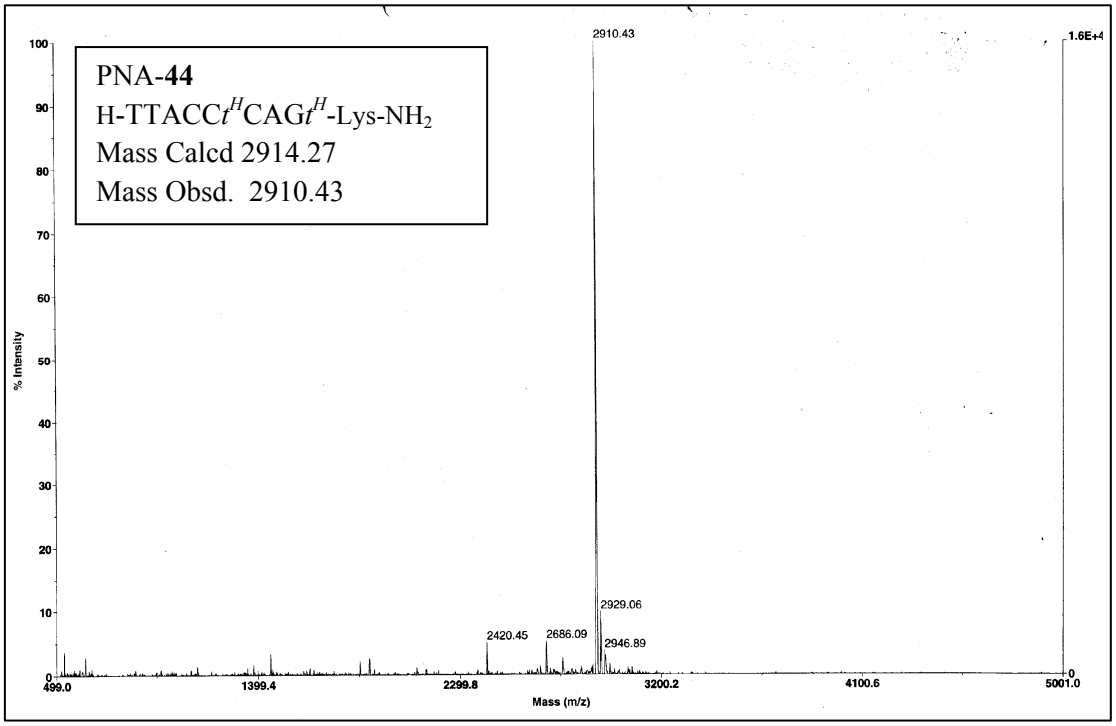

PNA-39
H-Lys-Lys-Lys-Lys
CTTCTTCCTT-Lys-NH2
Mass Calcd 3371. 1
Mass Obsd. 3366.3
% Intensity
Mass (m/z)
3366.39
1.4E+4
2533.34
3387.97
2555.48
3403.95
2571.17
3238.20
3495.29
3877.87
699.0
1559.4
2419.8
3280.2
4140.6
5001.0

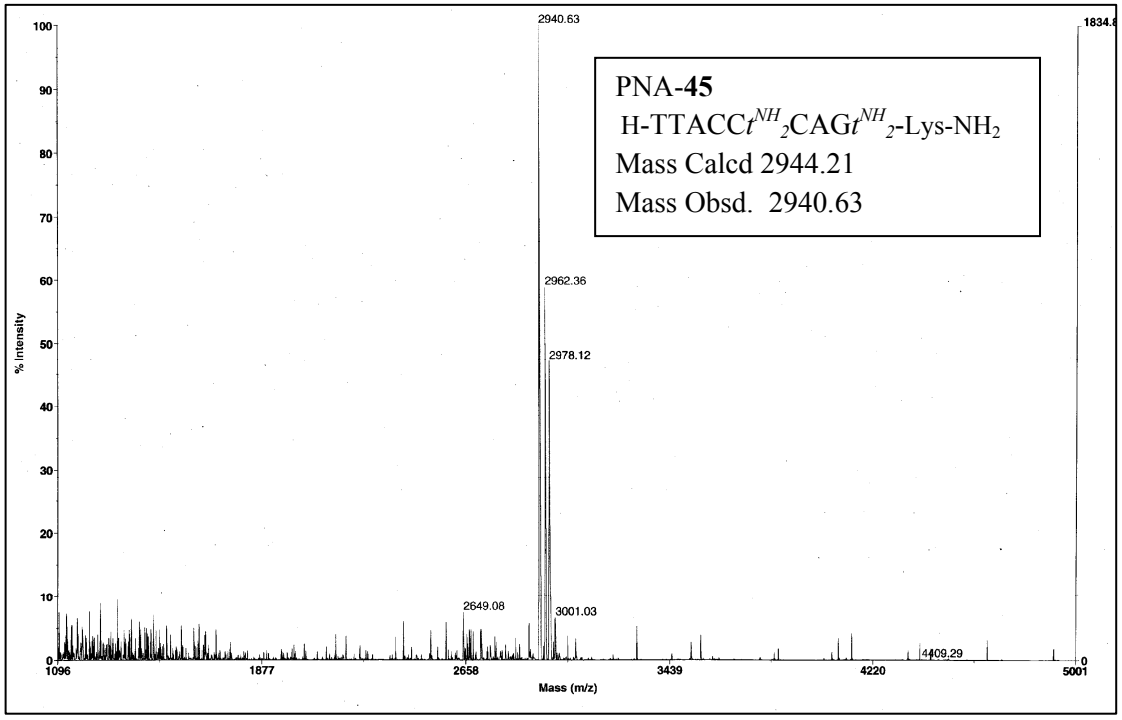

PNA-40
H-CTt^NH2CTt^NH2CCt^NH2t^NH2-Lys-NH2
Mass Calcd 3026.40
Mass Obsd. 3034.50
3034.50
1.8E+4
3056.43
3072.69
2742.15
2925.69
2885.38
2963.69
3163.26
% Intensity
Mass (m/z)
499.0
1399.4
2299.8
3200.2
4100.6
5001.0

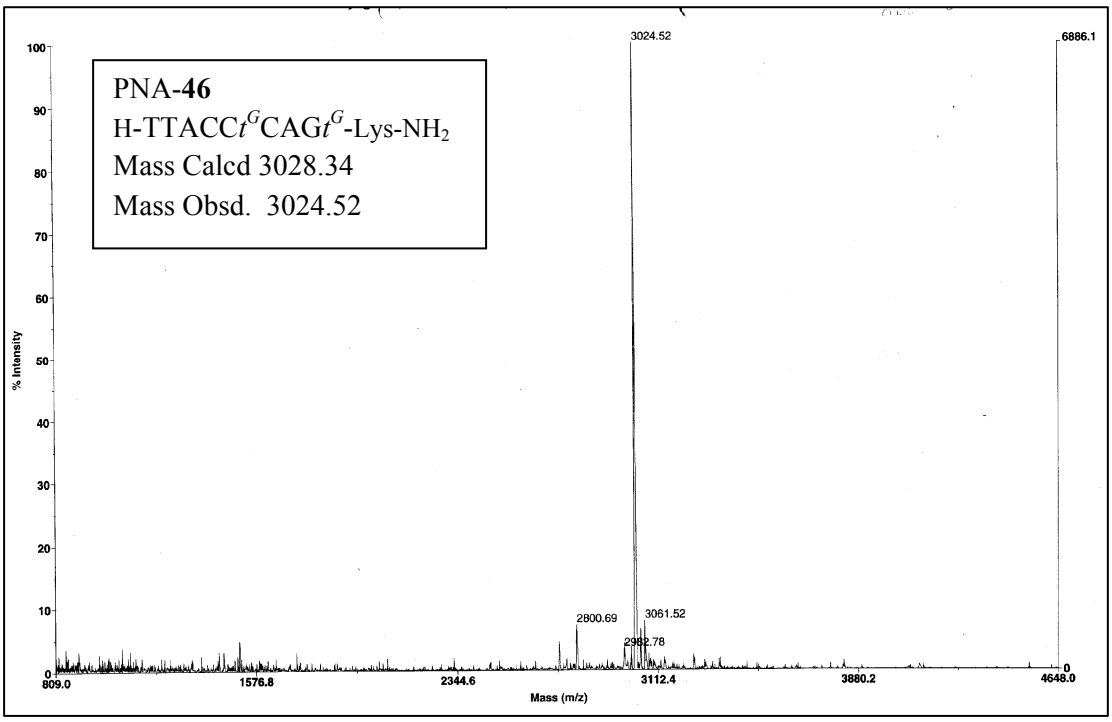

PNA-41
H-CTt^GCTt^GCCt^Gt^G-Lys-NH2
Mass Calcd 3194.35
Mass Obsd. 3201.40
3201.40
1.5E+4
2867.13
3160.15
2937.84
3092.16
2825.41
3120.00
3328.68
2030.14
2296.19
2601.44
3342.30
Mass (m/z)
499.0
1399.4
2299.8
3200.2
4100.6
5001.0


Final - Shots 750 - MAYURA08102009; Run #29; Label N17
PNA-42
H-TTACCTCAGT-Lys-NH2
Mass Calcd 2804.16
Mass Obsd.2827.0 (M+Na+)
2827.0740(R11673,S156)
2843.0637(R12030,S43)
211.8
Mass (m/z)


PNA-43
H-TTACCtHCAGT-Lys-NH2
Mass Calcd 2859.22
Mass Obsd. 2859.38
% Intensity
Mass (m/z)
2879.77
2896.26
2921.16
2859.38
1093.21
1002.90
1110.15
1610.61
1652.84
1591.02
1671.52
686.92
1300.32
1338.35
1904.10
1919.03
2149.38
2710.67
2726.00
2935.55
2744.47
3050.38
3025.01
3074.86
2551.68
2764.20
3346.17


PNA-**44**
H-TTACCt^{H}CAGt^{H}-Lys-NH_2
Mass Calcd 2914.27
Mass Obsd. 2910.43
% Intensity
Mass (m/z)
2910.43
2929.06
2946.89
2420.45
2686.09


PNA-**45**
H-TTACC$t^{NH}{}_2$CAG$t^{NH}{}_2$-Lys-NH_2
Mass Calcd 2944.21
Mass Obsd. 2940.63
% Intensity
Mass (m/z)
2940.63
2962.36
2978.12
2649.08
3001.03


PNA-46
H-TTACCt^GCAGt^G-Lys-NH2
Mass Calcd 3028.34
Mass Obsd. 3024.52
3024.52
6886.1
2800.69
3061.52
2982.78
% Intensity
Mass (m/z)
809.0
1576.8
2344.6
3112.4
3880.2
4648.0

Section II: Biophysical studies of H-pet-PNA, Amino-pet-PNA and Guanidino-pet-PNA towards complementary DNA/RNA

2.2.1: Introduction

The purpose of the designed oligonucleotides is to achieve i) high binding affinity and specificity to target nucleic acid. ii) stability to cellular enzymes iii) a long enough half-life within the system to allow manifestation of its effect and iv) non-toxicity to the system in which the activity is desired. All newly designed and synthesized molecules were hence subjected to various biophysical and biochemical studies to evaluate their potential as antisense agents.

2.2.2 Rationale and objectives of the present work

The preceding section reports the synthesis of chiral, conformationally constrained backbone-extended PNA analogue pet-PNA. These pet-PNA units were introduced at various positions in PNA oligomers in order to study the effect of structural constraint, chirality and cationic nature of pyrrolidinyl ring functionalized with amino and guanidino groups. In this section the studies that measure the influence of modifications on the binding properties to target nucleic acids in the terms of strength, specificity and directionality of binding is presented.

The objective of this Section is to present the results of the systematic study of the different chiral, pet-PNA, Am-pet and Gu-pet-PNA backbone modifications in terms of their ability and strength to form specific base paired complexes (duplexes and triplexes) with DNA/RNA. In order to explore the triplex forming potential, polypyrimidine sequences were constructed and to enable the study of duplexes, mixed purine-pyrimidine sequences were employed with mixed aeg/pet-PNA backbones. Also the effect of cationic substituent in the pyrrolidinyl ring on the strength of binding with the complementary DNA/RNA was studied. The specificity of binding was addressed by performing experiments with mismatched pairing at a single site in the target DNA/RNA sequences. This Section addresses the biophysical studies of pet-PNAs and their hybrids with complementary nucleic acids using techniques such as UV spectroscopy & CD spectrophotometry.

2.2.3 Biophysical spectroscopic techniques for studying PNA- DNA/ PNA-RNA interactions

Monitoring the UV absorption at 260 nm as a function of temperature has been extensively used to study the thermal stability of nucleic acid duplexes and triplexes and consequently, PNA:DNA hybrids as well. Increasing the temperature perturbs these complexes, inducing a structural transition by causing disruption of hydrogen bonds between the base pairs, diminished stacking interactions between adjacent nucleobases and larger torsional motions in the backbone leading to a loss of secondary and tertiary structure. This is reflected in the UV absorption at 260 nm, termed as hyperchromicity. The magnitude of hyperchromicity is a measure of the extent of the secondary structure present in nucleic acids. The process is cooperative and the plot of the absorbance at 260 nm *vs* the temperature is therefore sigmoidal. this also represents a two-state/All or none model for nucleic acid melting i.e., the nucleic acids exist in only two states, either in duplexes or single strands and at varying temperatures, the relative proportion change. A non-sigmoid (e.g. linear) transition with low hyperchromicity is a consequence of non duplexation (non complementation). In many cases, the transitions are broad and the exact T_ms are obtained from the peak in the first derivative plots. This technique has provided valuable information regarding complementary interactions in nucleic acids hybrid involving DNA, RNA and PNA.[34]

The fidelity of base pairing in the PNA:DNA/PNA:RNA complexes can be examined by changing the PNA oligomer with a DNA/RNA strand bearing mismatch at a desired site, preferably opposite the site of modification. The base mismatch leads to the absence of or incorrect hydrogen bonding between the bases and causes a drop in the measured melting temperature. A modification of the PNA structure is considered good if it gives a significantly lower T_m with RNA/ DNA sequences containg mismatches as compared to unmodified PNA. It is to be pointed out that in all biophysical experiments described herein, the modified PNAs were always evaluated by comparision with unmodified control PNA.

Homopyrimidine thymine PNA sequences bind to the complementary homopurine DNA sequences forming PNA_2:DNA triplexes in which it is difficult to distinguish the PNA strands that bind to the central DNA strand by WC hydrogen bonding and by HG hydrogen bonding. Mixed-base sequences form duplexes of antiparallel or parallel orientations that can be selected by proper design of the

complementary DNA sequences. By convention, antiparallel PNA:DNA complexes are defined as those in which the 'N' terminal of PNA faces the 3'-end of the DNA while the 'C' terminal faces the 5'-end. Parallel PNA:DNA complexes are those in which the 'C' terminal of PNA faces the 3'-end of DNA with the 'N' terminal towards the 5'-end of the DNA (Figure 9).

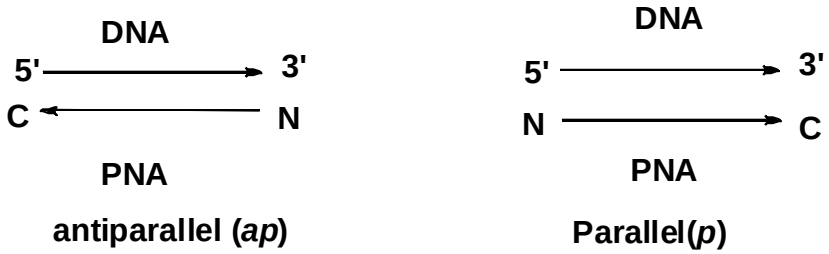


Figure 9: Schematic representation of the antiparallel and parallel modes of complexation of PNA with complementary DNA.

Mismatch studies: The complexes of PNAs were constituted with DNA/RNA containing a mismatch base (Figure 10). Mismatch studies show the sequence-specificity of the PNA sequence; melting temperatures for PNA:mismatch DNA/RNA complexes are typically lower than PNA:DNA/RNA complex. The binding strength is less for mismatch DNA/RNA as now one base is not available for hydrogen bonding.

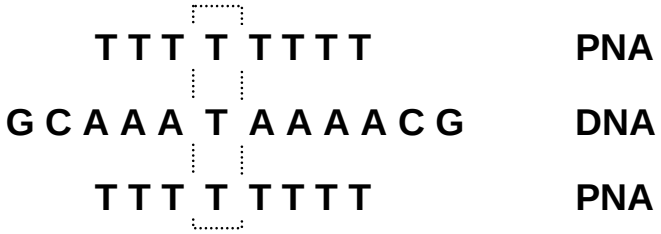


Figure 10: Mismatched PNA_2: DNA complex

Salt effect: Increase in the salt concentration in the medium is known to destabilize PNA:DNA complexes.[35] Positively charged cations in salt compete with cationic functional groups in the PNA backbone to have electrostatic interactions with the phosphate backbone of DNA/RNA. Change in the melting temperature of PNA:DNA/RNA complexes due to increase in the salt concentration of medium gives the information about electrostatic contribution to melting temperature and thus stability of complexes.

Results

2.2.4 UV melting studies

In the present Section studies on PNA:DNA/ PNA:RNA interactions investigated by UV and CD spectroscopic techniques are presented with discussion on effect of PNA modification on duplex/ triplex formation.

2.2.4a Homopyrimidine PNA-T_8 sequences: UV studies

The unmodified PNA T_8 sequence and PNA sequence with H-pet-PNA, Am-pet-PNA and gu-pet-PNA were annealed with DNA (**50**) and RNA (**60**) in 2:1 stoichiometry. Results for sequences with different pet units at the 'C' terminus are tabulated in Table 5. The H-pet-PNA modification was found to be destabilised PNA:DNA triplex. As a lower T_m 27 °C was observed, 15 °C lower than the unmodified PNA:DNA triplex. In case Am-pet and Gu-pet modifications, highly stable complexes were formed with 4 °C and 10 °C increase in T_m respectively as compared to control PNA. In the case of RNA, all the sequences containing pet units stabilized the resulting complexes. These complexes showed increase in T_m depending on the pet unit in the sequence. The stability increased from H-pet to Gu-pet modified unit containing triplexes melting temperatures of 48.7 °C to 58.7 °C ($\Delta T_m \approx 7$ °C-17 °C) respectively. The UV- T_m experiments are summarized in the Figures 10 and 11and data obtained is tabulated in Table 5.

Table 5: UV melting studies for PNA-**30-33** with DNA-**50** and RNA-**60**

Code	Sequence	UV-T_m °C	
		DNA 50	**RNA 60**
aegPNA **30**	H-TTTTTTTT-Lys-NH_2	41.6	42.5
	H-TTTTTTTt^R-Lys-NH_2		
PNA **31**	R = H	26.7	48.7
PNA **32**	R= NH_2	45.7	49.9
PNA **33**	R= NH-C(= NH) NH_2	51.6	58.7

DNA 50 = 5'GCA_8CG 3'; RNA **60** = 5'GCA_8CG 3', *T*m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), of PNA_2-DNA/ PNA_2-RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

Mismatch studies were also carried out to evaluate the sequence specificity of binding and these are tabulated in the Table 6. All the pet-PNA sequences exhibited better mismatch discrimination than control PNA sequence in both the cases of TT mismatch in DNA and TU in mismatch RNA complexes. In the case of both Am-pet and Gu-pet-PNAs decrease in the T_m was in the range of 17 °C-22 °C, much higher than

that observed for unmodified PNA:DNA(ΔT_m= 11.6) and PNA:RNA(ΔT_m= 12.7) triplexes. If the electrostatic interactions of amino/guanidine functions on modified oligomers with phosphates in DNA/RNA were the most important contributors than H-bonding for the higher melting temperatures seen in the Table 5 the mismatch dicrimination would have been much less pronounced. The observed mismatcyh discrimination means that the electrostatic interactions of amino/guanidino groups with phosphates do not overrule the WC or HG hydrogen bonding interactions in triplexes.

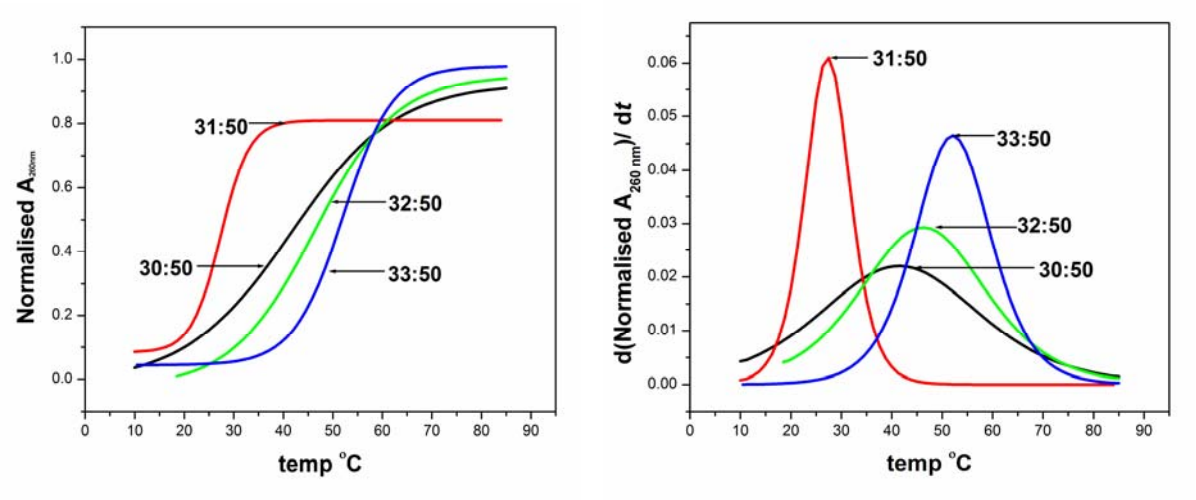


Figure 10: A UV melting curves of PNA-**30-33** with DNA-**50** and **B** corresponding derivative graphs**.**

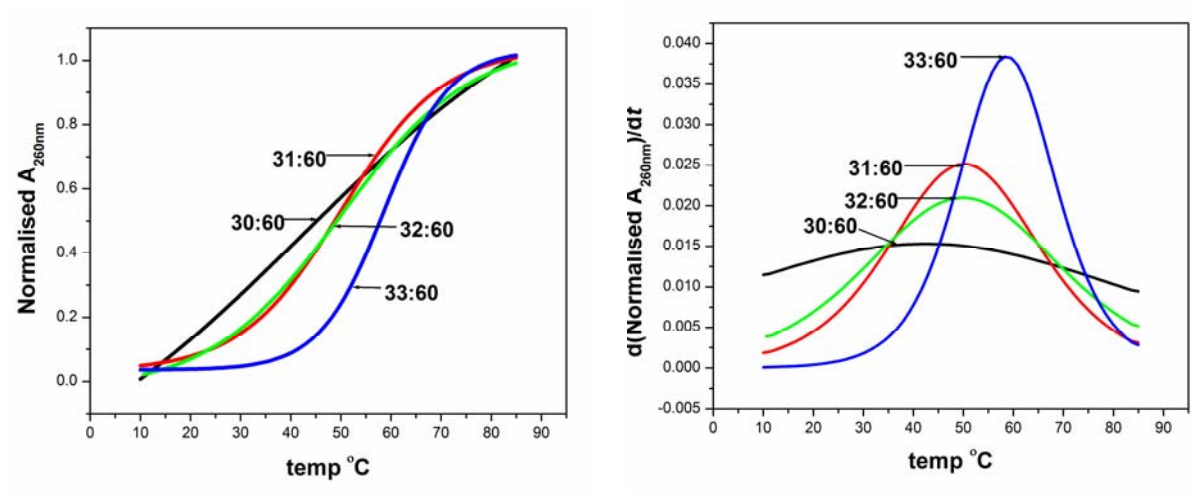


Figure 11: A UV melting curves of PNA-**30-33** with RNA-**60** and **B** corresponding derivative graphs**.**

Table 6: Mismatch UV melting studies for PNA-**30-33** with DNA-**51** and RNA-**61**

Code	Sequence	UV-T_m °C	
		DNA 51	**RNA 61**
aegPNA **30**	H-TTTTTTTT-Lys-NH_2	30.0(-11.6)	29.8 (-12.7)
	H-TTTTTTTt^R-Lys-NH_2		
PNA **31**	R = H	nd	20.9 (-17.8)
PNA **32**	R= NH_2	26.8(-18.9)	28.5(-21.4)
PNA **33**	R= NH-C(= NH) NH_2	32.6(-19.0)	31.8(-26.9)

DNA 51 = 5**'**GCAAATAAAACG 3'; RNA **61** = 5**'**GCAAAUAAAACG 3', *T*m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), PNA_2-DNA/ PNA_2-RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.Values in parentheses denote the difference in the *T*m due to mismatch in the DNA/RNA sequence. nd = not detectable.

UV melting studies of aegPNA T_8 sequences with pet-PNA units in the middle of the aegPNA, (Table 7) were carried out with complementary DNA-**50** and RNA-**60**. PNA sequence **34** with H-pet unit in the middle of the sequence destabilizes PNA: DNA complex in a similar way as the 'C' terminus modified sequence PNA-**31**. In the case of sequence Am-pet (PNA-**35**) and Gu-pet (PNA-**36**) units, the PNA: DNA complexes were marginally destabilized as compared to control the aegPNA**30**. The binding strength of PNAs **35** and **36** with RNA-**60** was slightly les than with DNA-**50.** Both amino and guanidino units destabilized the PNA:RNA complexes by 5-6 °C as compared to control PNA **30**. PNA-**37** which is a homooligomer of H-pet thymine units showed stabilization with DNA-**50** but binds to RNA-**60** almost as well as unmodified aegPNA.

Table 7: UV melting studies for PNA-**34-37** with DNA-**50** and RNA-**60**

Code	Sequence	UV-T_m °C	
		DNA 50	**RNA 60**
aegPNA **30**	H-TTTTTTTT-Lys-NH_2	41.6	42.5
	H-TTTTt^RTTT-Lys-NH_2		
PNA **34**	R = H	27.2	37.4
PNA **35**	R= NH_2	40.7	36.8
PNA **36**	R= NH-C(= NH) NH_2	36.0	36.4
PNA **37**	H-*ttttttt*-Lys-NH_2 *t*=H-pet	52.9	40.6

DNA 50 = 5**'**GCA_8CG 3'; RNA **60** = 5**'**GCA_8CG 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0),of PNA_2-DNA/ PNA_2-RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C. t is pet-PNA unit.

Mismatch studies (Table 8) of these sequences with DNA-**51**and RNA-**61** with single base-pair mismatch (T-T/T-U), displayed large destabilization compared with aegPNA. The sequence specificity of theses sequence was found to be better than control PNA sequence. PNA-**35** and PNA-**36** with cationic pet units also exhibited the decrease in the T_m with mismatched DNA/RNA in the range of 14-22°C.

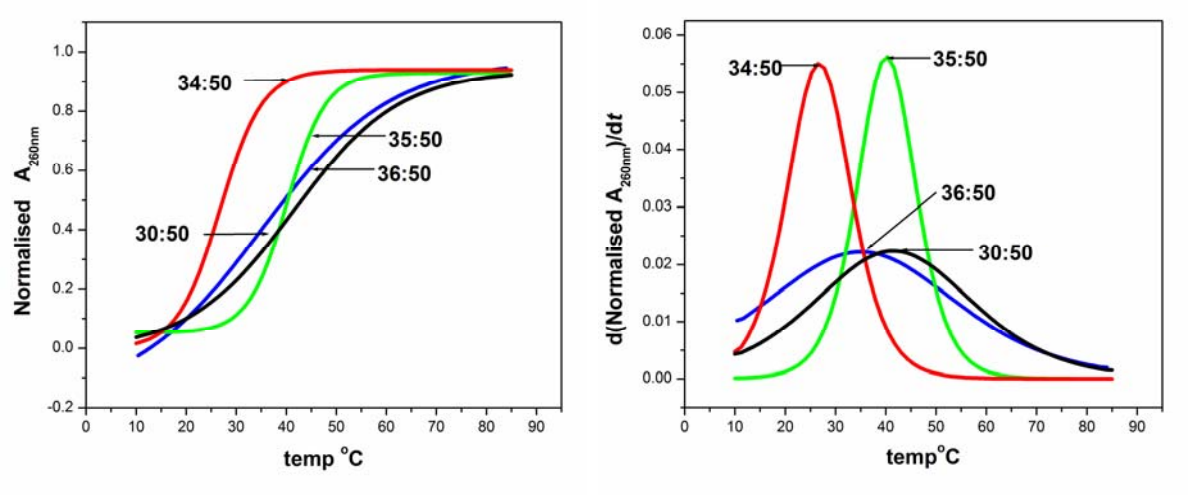


Figure 12: A UV melting curves of PNA-**30, 34-36** with DNA-**50** and **B** corresponding derivative graphs**.**

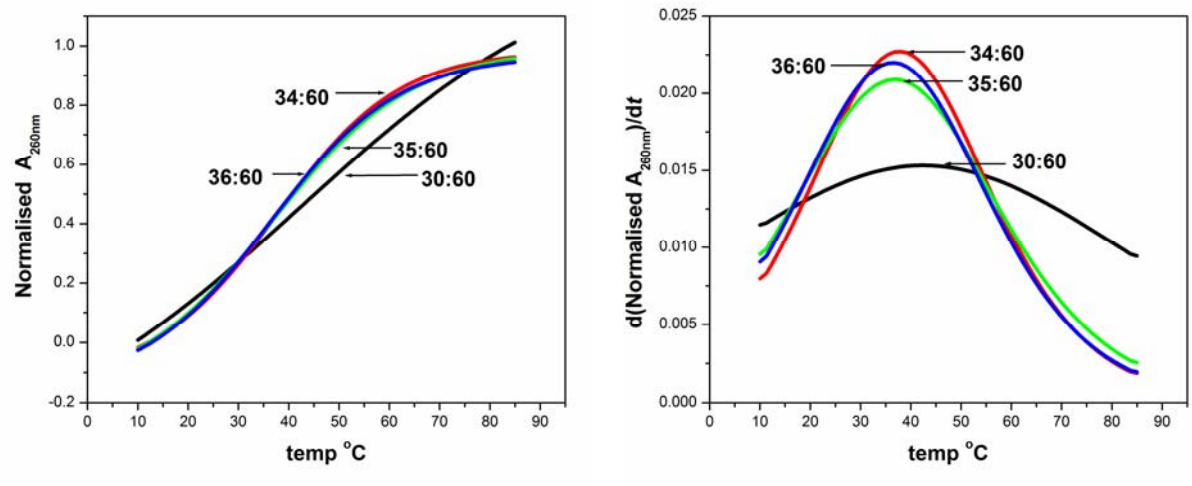


Figure 13: A UV melting curves of PNA-**30, 34-36** with RNA-**60** and **B** corresponding derivative graphs

Table 8: Mismatch UV melting studies for PNA-**34-36** with DNA-**51** and RNA-**61**

Code	Sequence	UV-T_m °C	
		DNA 51	**RNA 61**
aegPNA**30**	H-TTTTTTTT-Lys-NH_2	30.0(-11.6)	29.8 (-12.7)
	H-TTTTt^RTTT-Lys-NH_2		
PNA **34**	R = H	nd	16.5 (-20.9)
PNA **35**	R= NH_2	19.4(-21.3)	18.2(-18.6)
PNA **36**	R= NH-C(= NH) NH_2	21.4(-14.6)	18.9(-17.5)

DNA 51 = 5**'**GCAAATAAAACG 3'; RNA **61 =** 5**'**GCAAAUAAAACG 3', *T*m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), PNA_2-DNA/ PNA_2-RNA complexes. The 4 values reported here are the average of 3 independent experiments and are accurate to ±1.0°C. Values in the parentheses denote the difference in the T_m due to mismatch in the DNA/RNA sequence. nd = not detectable.

2.2.4b Mixed pyrimidine PNA sequences (b_2a_2): UV studies

A sequence complementary to b_2a_2 site in CML cells is a mixed pyrmidine sequence. The binding of this sequence to the target complementary RNA and also to DNA was studied by UV melting studies and the results are tabulated in the Table 9. PNA **40** and **41** contain 4 modified units of amino and guanidino-pet units in the aegPNA backbone respectively. To compare the results of these PNAs, we have carried out UV melting studies of aegPNA **38** with four lysines attached at 'N' terminus (PNA **39**). PNA **39** exhibited increase in the binding with both DNA and RNA compared to the corresponding control PNA **38**. Increase in the melting temperature was about 5 °C - 9 °C, Am-pet-PNA (PNA**40)** and Gu-pet-PNA (PNA**41**) formed complexes with DNA with melting temperatures as good as PNA**39**.

Table 9: UV melting studies for PNA-**38-41** with DNA-**52** and RNA-**62**

Code	Sequence	UV-T_m °C	
		DNA 52	**RNA 62**
aegPNA **38**	H-CTTCTTCCTT-Lys-NH_2	52.0	55.6
aegPNA **39**	H-Lys-Lys-Lys-Lys-CTTCTTCCTT-Lys-NH_2	63.4	59.4
	H-CTt^RCTt^RCC$t^R t^R$-Lys-NH_2		
PNA **40**	R= NH_2	63.7	65.5
PNA **41**	R= NH-C(= NH) NH_2	64.8	68.3

DNA 52 = 5**'** AAGGAAGAAG 3'; RNA **62 =** 5**'** AAGGAAGAAG 3', *T*m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), of PNA_2-DNA/ PNA_2-RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

These sequences also formed strong complexes with complementary RNA, that were even stronger than aegPNA**38** and aegPNA**39** giving melting temperatures 65.5°C and 68.3 °C respectively. This increase in the melting temperature is 10 °C-18 °C more than control aegPNA **38**.As PNA **40** and PNA **41** and also aegPNA **39** showed increase in binding with DNA and RNA as compared to control PNA, mismatch studies (Table 10) were carried out with DNA**53** and RNA**63** with T-T andT-C mismatch respectively. PNA **39** has showed less mismatch discrimination than aegPNA, due to the presence of additional lysines that might have given rise to sequence independent electrostatic interactions with the phosphates of DNA/ RNA. PNA **40** and PNA **41** showed mismatch base discrimination as good as control aegPNA with mismatch DNA. For RNA **63** decreases in the melting temperature was just 1-2 °C less as compared to aegPNA **38**.

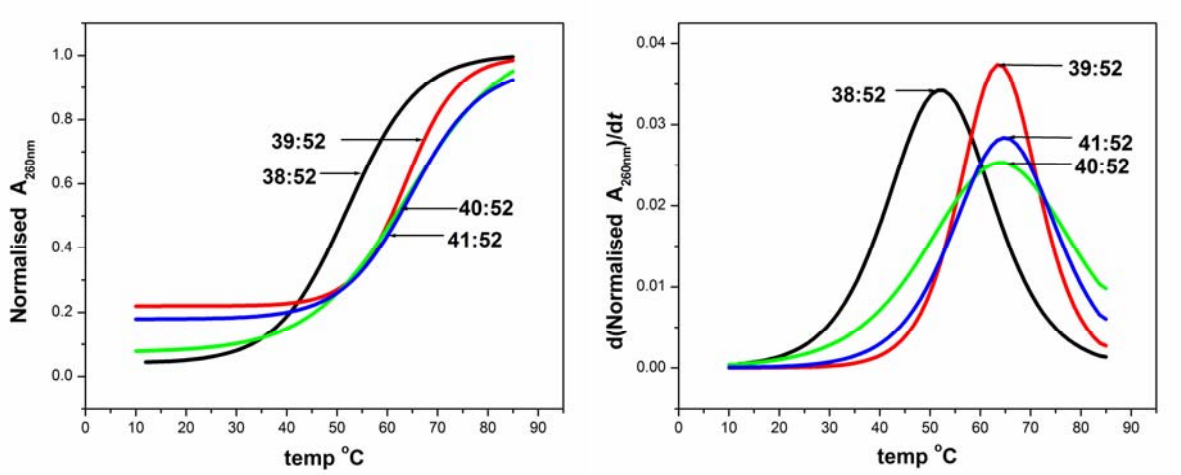


Figure 14: A UV melting curves of PNA **38-41** with DNA-**52** and **B** corresponding derivative graphs**.**

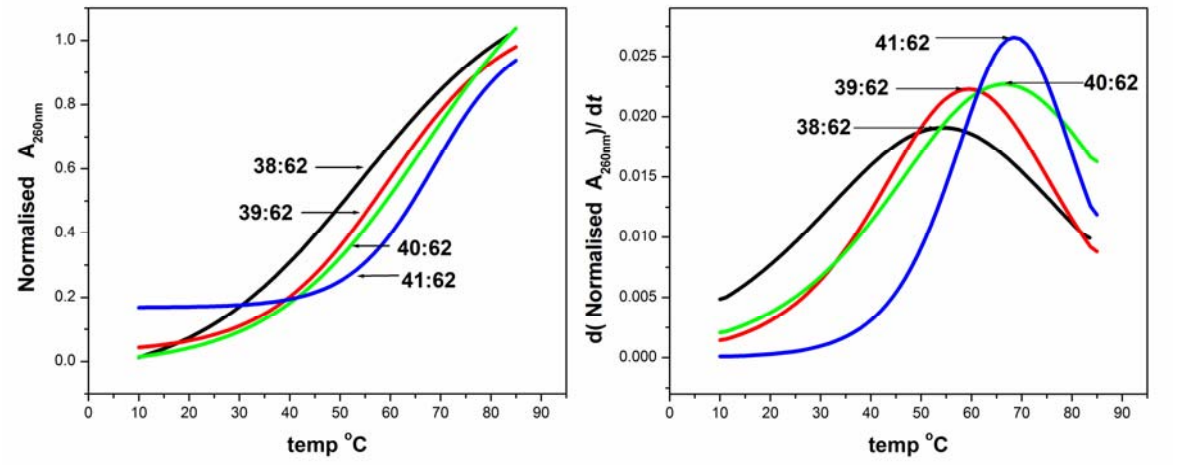


Figure 15: A UV melting curves of PNA **38-41** with RNA-**62** and **B** corresponding derivative graphs**.**

Table 10: Mismatch UV melting studies for PNA-**38-41** with DNA-**53**and RNA-**63**

Code	Sequence	UV-T_m °C	
		DNA 53	**RNA 63**
aegPNA **38**	H-CTTCTTCCTT-Lys-NH_2	42.6 (-9.4)	45.2(-10.4)
aegPNA **39**	H-Lys-Lys-Lys-Lys-CTTCTTCCTT-Lys-NH_2	57.8(-5.6)	53.1(-6.3)
	H-CTt^RCTt^RCC$t^R$$t^R$-Lys-$NH_2$		
PNA **40**	R= NH_2	49.9(-13.8)	57.6(-7.9)
PNA **41**	R= NH-C(= NH) NH_2	54.5(-10.3)	59.2(-9.1)

DNA 53 = 5**'** AAGGTAGAAG 3'; RNA **63 =** 5**'** AAGGACGAAG 3', *T*m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), of PNA_2-DNA/ PNA_2-RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C. Values in the parentheses denote the difference in the T_m due to mismatch in the DNA/RNA sequence.

2.2.4c Mixed purine-pyrimidine PNA sequences: UV-melting studies

The modified unit could show different effects in the mixed base sequences due to the presence of purine as well as pyrimidine bases in the sequence. The mixed base sequences can also be used to evaluate the ability to distinguish between the parallel and antiparallel modes of binding. The chirality of the backbone may be an important factor that influence the directional selectivity of binding. UV melting studies of mixed purine-pyrimidine sequences with DNA **54,** RNA **64** and parallel DNA **56** are summarized in the Table 11.

Table 11: UV melting studies for PNA-**42-46** with a*p*DNA-**54**, RNA-**64** and *p*DNA-**56**

Code	Sequence	UV-T_m °C		
		***ap*DNA 54**	**RNA 64**	***p*DNA 56 (Δ T_m= *ap-p*)**
aegPNA **42**	H-TTACCTCAGT-Lys-NH_2	54.2	54.5	53.9 (0.3)
PNA **43**	H-TTACCTCAGt^R-Lys-NH_2 R= H	57.6	52.5	55.4(1.8)
	H-TTACCt^RCAGt^R-Lys-NH_2			
PNA **44**	R=H	63.2	48.1	59.3(3.9)
PNA**45**	R= NH_2	66.9	52.1	59.5(7.4)
PNA **46**	R= NH-C(= NH) NH_2	68.9	54.2	59.9(9.0)

***ap*DNA 54** = 5' ACTGAGGTAA 3'; RNA **64 =** 5**'** UGUAACUGAGGUAAGAGG 3', ***p*DNA 56** = 5**'** AATGGAGTCA 3' T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), of PNA:DNA/ PNA:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C. *ap*= antiparrallel p = parallel

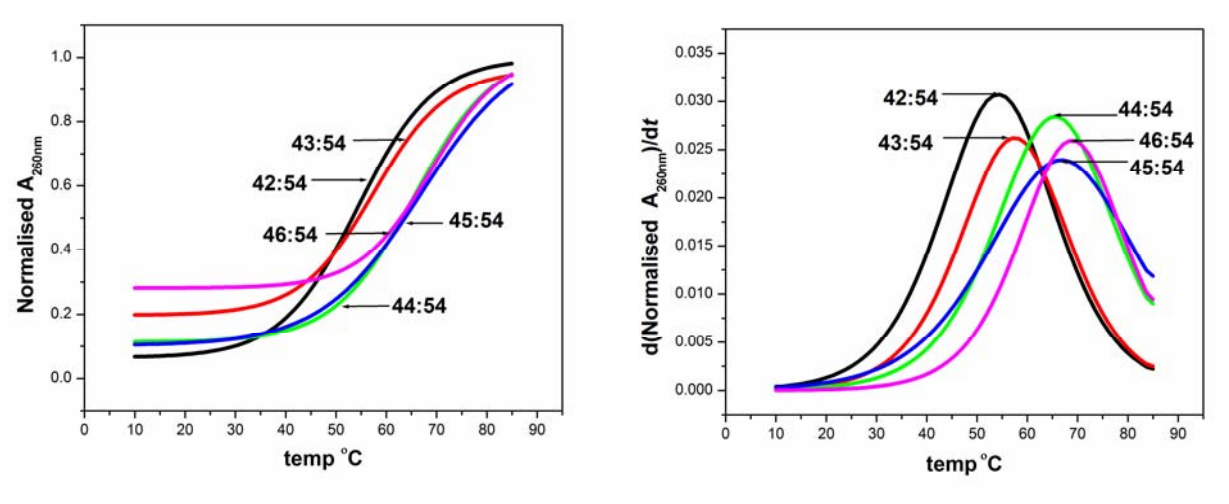


Figure 16: A UV melting Curves of PNA **42-46** with DNA-**54** and **B** corresponding derivative graphs**.**

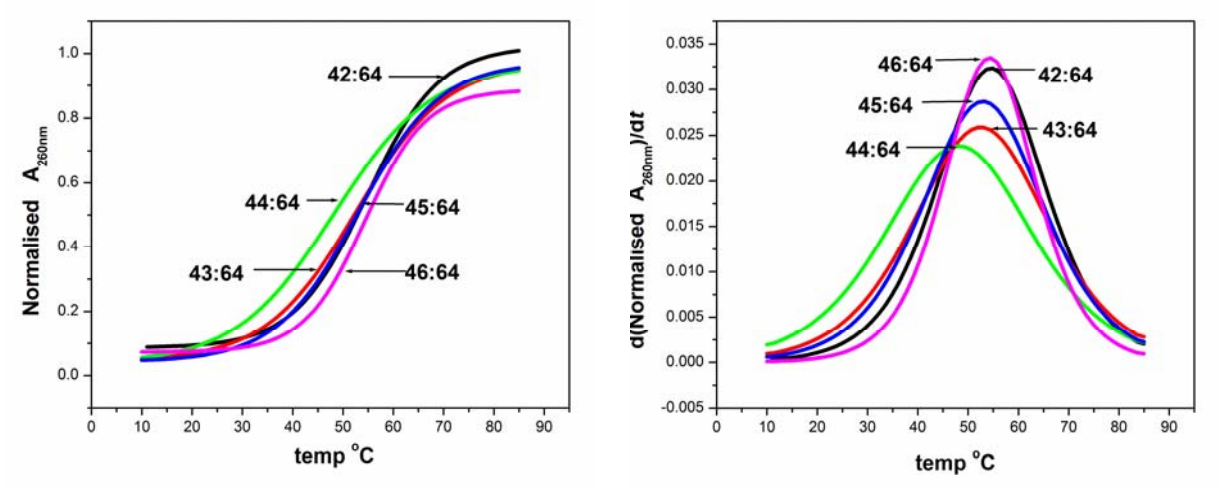


Figure17: A UV melting Curves of PNA **42-46**with RNA-**64** and **B** corresponding derivative graphs**.**

Mixed purine-pyrimidine PNA sequences form strong duplexes with complementary DNA and RNA targets. There is no restriction of nucleobase sequence in the case , as in the triplxes where polypyrimidine are necessary. Therefore theses sequences have wider applicability in the antisence research.

The control aegPNA**42** showed similar binding strength with DNA-**54** and RNA-**64** giving melting temperatures around 54 °C. Replacement of one of the thymine units in PNA-**42** by a H-pet unit in the sequence stabilized the complex with *ap*DNA **54** (ΔT_m = +3.4), whereas a 2 °C destabilization was observed for the duplex with RNA with compared to control PNA (ΔT_m = -2.0). The doubly modified sequence PNA **44** further stabilized the duplex with DNA-**54**. The effect of modified units thus being additive. Binding with RNA was found to be further reduced by 4 °C, thus the

destabilization of the duplexes with RNA was also additive. The doubly modified sequences with Am-pet and Gu-pet units increase the stabilization in both DNA and RNA duplexes. Both sequences bind to RNA as good as the control PNA whereas the incase of DNA duplex, this stabilization has increase by 12 °C to 15 °C as compared to PNA-**42**. The control PNA is known to have less discrimination for directionality in binding. To see the effect of chirality of the modified units on antiparallel (*ap*) *versus* parallel (*p*) orientation preference, binding studies were carried out with *p*DNA**56** and results are shown in Table 11 column 5. Values in the parentheses are the difference in T_m values (*ap*-*p*). It is clear that introduction of modified units induced better binding in *ap* orientation and this directional preference increased from H-pet-PNA(**43**) to gu-pet-PNA(**46**).

Mismatch studies: To delineate the electrostatic interactions and nuclobase fidelity of WC base pairing, mismatch studies of mixed purine-pyrimidine sequences were carried out with DNA **55** with (T-T) and RNA **65** with (U-T) mismatches and the results are shown in Table 12, values in the parentheses indicate the difference in T_m values, from these values it can be seen that mismatch discrimination is better than control PNA for all the modified sequences. Mismatch base discrimination is more in doubly modified aegPNA sequence with H-pet units (PNA **55**) than Am-pet and Gu-pet pet-PNA sequences. This mismatch discrimination could mean that electrostatic interactions may not overrule the hydrogen bonding in the bases

Table 12: Mismatch UV melting studies for PNA-**42-46** with DNA-**55**and RNA-**65**

Code	**Sequence**	**UV-T_m °C**	
		DNA 55	**RNA 65**
aegPNA**42**	H-TTACCTCAGT-Lys-NH_2	47.5(-6.7)	46.5 (-8.0)
PNA **43**	H-TTACCTCAGt^R-Lys-NH_2 R= H	51.2(-6.4)	38.8(-13.8)
PNA **44**	H-TTACCt^RCAGt^R-Lys-NH_2 R=H	50.8(-12.4)	34.5(-13.7)
PNA **45**	R= NH_2	58.1(-8.8)	43.8(-8.3)
PNA **46**	R= NH-C(= NH) NH_2	60.0(-8.9)	46.1(-8.1)

DNA 55 = 5**'** ACTGTGGTAA 3'; RNA **65 =** 5**'** UGUAACUGCGGUAAGAGG 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), of PNA:DNA/ PNA:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C. Values in the parentheses denote the difference in the *T*m due to mismatch in the DNA/RNA sequence.

Salt effect: Increase in the salt concentration reduce the electrostatic interactions of the cationized PNA with anionic phosphates groups of DNA-RNA. Hence theses studies were undertaken to study the effect of salt concentration. UV melting experiments were carried out with RNA-**54** at 200 mM NaCl salt concentration and the results are summarized in Table 13, Values in parentheses indicate the difference in melting temperature due increase in the salt concentration.

Table 13: UV melting studies for PNA-**45, 46**with RNA-**64** at 200 mmol NaCl concentration

Code	**Sequence**	**UV-T_m °C**
		RNA 64
aegPNA**42**	H-TTACCTCAGT-Lys-NH_2	52.8 (-1.7)
	H-TTACCt^RCAGt^R-Lys-NH_2	
PNA **45**	R= NH_2	48.6 (-3.5)
PNA **46**	R= NH-C(= NH) NH_2	49.7 (-4.5)

RNA **64** = 5' UGUAACUGAGGUAAGAGG 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 200 mM NaCl, 0.1 mM EDTA, pH = 7.0), of PNA:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C. a. values in the parentheses denote the difference in the *T*m due to increase in the salt concentration in the RNA sequence.

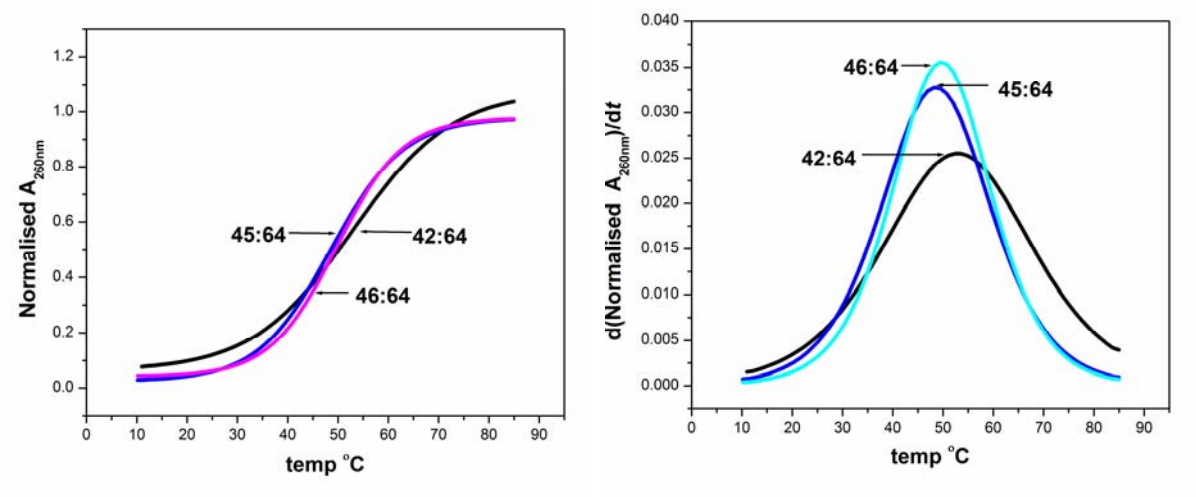


Figure 18: A UV melting Curves of PNA **42, 45, 46** with RNA-**64** at 200 mM NaCl concentration and **B** corresponding derivative graphs**.**

The Salt effect can be seen more prominently when cationic groups are present in backbone of PNA, thus salt effect studies were carried out with Am-pet- and Gu-pet-PNA sequences. Decrease in the melting temperatures were observed for both modified PNA sequences and this decrease was of 3-4 °C. Decrease in the melting temperatures at 200 mM salt concentrations points out to the contribution of electrostatic interactions

to the stability of Am-pet-, Gu-pet-PNA with DNA and RNA. Am-pet- and Gu-pet-PNAs showed good mismatch discrimination. This could mean that electrostatic interactions would not overrule the WC hydrogen bonding interactions in duplexes.

2.2.5 CD analysis of single stranded PNAs

Effect of incorporation of cyclic, chiral units in aegPNA sequences on structural preorganisation was studied by CD analysis of single stranded PNA sequences at 2 μM PNA concentration. This study is presented in Figure 19.

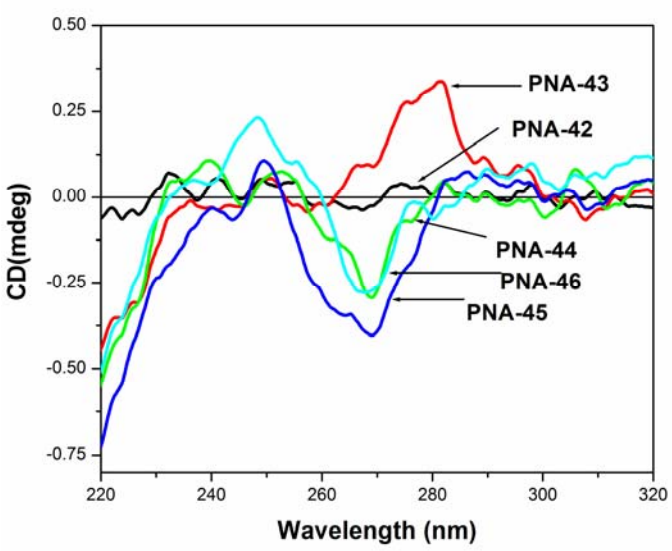


Figure 19: CD analysis of single stranded mixed purine-pyrimidine sequences at 2 μM concentration

The achiral aegPNA**42** did not show any prominent CD signal, while incorporation of one H-pet unit (PNA-**43**) showed a positive band centred at 280nm. Introduction of one more H-pet unit in PNA (PNA-**44**) has led to a completely different changed CD pattern giving a negative band centered at 270 nm and other doubly modified sequences (PNA-**45** & PNA-**46**) also showed same CD pattern.

2.2.6 Summary

- The introduction of pet-PNA unit in the middle of the thyminyl aegPNA octamer destabilized the PNA_2:DNA/PNA_2:RNA complex, whereas the sequences containing pet-PNA units at 'C' terminus of octamer stabilized the complexes with RNA.

- The complexes of thymine-cytosine contaning PNA decamer sequences modified with four Am-pet-PNA/ Gu-pet-PNA units formed the stable complxes with complementary DNA/RNA as compared to control aegPNA sequences.

- In the mixed purine-pyrimidine PNA sequences, the stability of PNA:DNA duplexes increased when pet-PNA units were introduced. The stability of PNA:RNA duplexes decreased with increase in H-pet-PNA units. Whereas the stability of duplexes of Am-pet-PNA and Gu-pet-PNA :RNA was found to be as good as the control PNA:RNA duplexes.

- The introduction of modified units induced better binding in *ap* orientation and this directional preference increased from H-pet-PNA(**43**) to Gu-pet-PNA(**46**).

- The effect of introduction of chiral, cyclic pet-PNA units in aeg sequences PNA on structural preorganization of PNA was observed in CD analysis of single strand PNA.

2.3 References

1 Bennett, C. F.; Swayze, E. E. RNA targeting therapeutics: Molecular mechanisms of antisense oligonucleotides as a therapeutic platform *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259-293.

2 Micklefield, J. Backbone modification of nucleic acids: synthesis, structure and therapeutic applications *Curr. Med. Chem.*, **2001**, *8*, 1157. b) Turner, J. J.; fabini, M.; Arzumanov, A.; Ivanova, G.; Gait, M. J. Targeting the HIV-1 RNA leader sequence with synthetic oligonucleotides and siRNA: chemistry and cell delivery *Biochim.biophys. Acta* **2006**, *1758*, 290. c) Kumar, V. A.; Ganesh, K. N. Structure-Editing of nucleic acids for selective targeting of RNA *Curr.Top. Med. Chem.* **2007**, *7*, 715.

3. Nielsen, P. E.; Egholm, M.; Buchardt, O. (1991) Sequence selective recognition of DNA by strand displacement**.** *Science*, *254*, 1497-1500.

4 Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Frier, S.; Driver, D. A.; Berg R.H.; Kim, S.K.; Norden, B.; Nielsen, P.E. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993**, *365*, 566-568.

5 Nielsen, P.E.; Egholm, M.; Buchardt, O. Peptide nucleic acid (PN(A); ADNA mimic with a peptide backbone . *Bioconjugate Chem.* **1994**, *5*, 3-7.

6 (a) Nielsen, P. E.; Egholm, M. *Peptide Nucleic Acids*. Protocols and applications; Horzon Press; Norfolk, **1999**. (b) Ray, A.; Norden, B. Peptide nucleic acid : its medical and biotechnological applications and promise for the future . FASEB J. **2000**, *14*, 1041-1060.

7 Good, L.; Nielsen, P.E. Antisense Inhibition of gene expression in bacteria by PNA targeted to mRNA. *Nature (Biotechnology)*, **1998**, *16*, 355-358.

8 Nielsen, P.E. Applications of peptide nucleic acids. *Curr. Opin. Biotechnol*. **1999**, *10*, 71-76.

9 (a) Nielsen. P.E. *Pure Appl. Chem*. **1998**, *70*, 105-110. (b) Corey, D.R. Peptide nucleic acids; expanding the scope of nucleic acid recognition. *Trends Biotechnol*. **1997**, *15*, 224-229.10.

10 (a) K. N. Ganesh; Nielsen, P.E. Peptide nucleic acids: Analogs and derivatives . *Curr. Org. Chem.* **2000**, *4*, 931-943. (b) Kumar, V. A. Structural pre-organization of peptide nucleic acids: chiral cationic analogs with five- or six- membered ring structures. *Eur. J.Org. Chem.* **2002**, 2021-2032. (c) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404-413.

11 (a) Bendifallah, N.; Rasmussen, F. W.; Zachar, V.; Ebbesen, P.; Nielsen, P. E.; Koppelhus, U. *Bioconjugate Chem.* **2006**, *17,* 750. (b) Abes, S.; Turner, J. J.; Ivanova, G. D.; Williams, D.; Arzumanov, A.; Clair, P.; Gait, M. J. ; Lebleu, B. *Nucleic Acids Research*, **2007***, 35,* 4495-4501. (c) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed*. **1998**, *37*, 2796-2823.

12 (a) Koppelhus, U.; Awasthi, S. K.; Zachar, V.; Holst, H. U.; Ebbeson, P.; Nielsen, P. E.Cell-dependant diferential cellular uptake of PNA, peptides , and PNA-peptide conjugates *Antisense Nucleic Acid DrugDev*. **2002**, *12*, 51–63.

13 (a) Wolf, Y; Pritz, S; Abes, S; Bienert, M; Lebleu, B; Oehlke, J. Structural requirements for cellular uptake and antisense activity of peptide nucleic acids conjugated with various peptides *Biochemistry* **2006**, *45,* 14944. (b) Abes, S.; Willams, D.; Prevot, P.; Thierry, A.; Gait, M. J. ; Lebleu, B. Endosome trapping limits the effiency of splicing correction by PNA-oligolysine conjugates *J. Controlled Release* **2006**, *110,* 595. (c) Turener, J. J; Fabani, M.; Arzumanov, A.A.; Ivanova, G.; Gait, M.J**.** Targeting the HIV-1RNA leader sequence with synthetic oligonucleotides and siRNA: Chemistry and cell delivery Biochim. Bioohys. Acta , **2006**, 1758, 290. (d) Koppelhus, U.; Shiraishi, T.; Zachar, V.; Pankratova, S.; Nielsen, P. E. Improved Cellular Activity of Antisense peptide Nucleic Acids by Conjugation to a Cationic peptide-

Lipid (CatLip) Domain *Bioconjugate Chem.* **2008**, *19*, 1526-1534 (e) Hassane, F. S.; Ivanova, G. D.; Bolewska-Pedyczak, E.; Abes, R.; Arzumanov, A. A.; Gait, M. J.; Lebleu, B.; Gariepy, J. A Peptide Based Dendrimer that enhances the Splice –Redirecting Activity of PNA Conjugates in Cells *Bioconjugate Chem*. **2009**, *20*, 1523-1530.

14 Koppelhus, U.; Nielsen, P.E.; Cellular delivery of peptide nucleic acid (PNA) *Adv. Drug Deliv. Rev.* **2003**, *55*, 267-280.

15 Jones, S. W; Christison, R; Bundell, K; Catherine, J. V.; Brockbank, S. M. V. ; Newham, P.; Lindsay, M. A. The characterisation of cell-penetrating peptides-mediated peptide delivery *Br. J. Pharmacol* **2005**, *145,* 1093

16 (a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. Peptide Nucleic Acids (PNAs) Containing Thymine Monomers Derived from chiral amino acids: hybridization and Solubility properties of D-lysine PNA *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939-1942. (b) Tedeschi, T.; Sfroza, S.; Coradini, R.; Marchelli, R. Synthesis of new chiral PNAs bearing dipeptide-mimic monomer with two lysine-derived stereogenic centres. *Tetrahedron Lett.* **2005**, *46*, 8395-8399. (c) Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.*, **1999**, 197-204. (d) Tedeschi, T.; Sforza, S.; Dossena, A.; Corradini, R.; Marchelli, R. *Chirality*, **2005**, *17*, S196-S204. (e) Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli, R. *Eur. J. Org. Chem.*, **2007**, 5879-5885.

17 (a) Zhou, P.; Wang, M.; Du, L; Fisher, G.W.; Waggoner, A.; Ly, D.H. Novel Binding and Efficient Cellular Uptake of Guanidine-Based peptide Nucleic Acids (GPNA) *J. Am. Chem. Soc.* **2003**, *125,* 6878. (b) Dragulescu- Andrasi, A.; Zhou, P.; He, G.; Ly, D.H. Cell-permeable GPNA with appropriate backbone streochemistry and spacing binds sequence-specifically to RNA *Chem. Comm.* **2005**, 244-246. (c) Zhou, P.; Dragulescu-Andrasi, A.; Bhattacharya, B.; O'Keefe, H.; VAtat, P.; Hyldig-Nielsen, J. J.; Ly, D. H. Synthesis of cell-permeable peptide nucleic acids and characterization of their hybridization and uptake properties *Bioorg. Med. Chem. Lett.* **2006**, *16,* 4931-4935.(d) Sahu, B.; Chenna, V.; Lathrop, K. L.; Thomas, S. M. ; Zon, G.; Livak, K. J. ; Ly, D. H. Synthesis of Conformationally Preorganized and Cell-permeable Guanidine-Based γ-peptide Nucleic Acids (γGPNAs) *J. Org. Chem.* **2009**, *74,* 1509-1516.

18 (a) Micklefield, J. Backbone modification of nucleic acids: synthesis, structure and therapeutic appilications *Curr. Med. Chem.* **2001**, *8*, 1157. (b) Kurreck, J.Antisense technololigies: improvement through novel chemical modifications *Eur. J. Biochem.* **2003**, *270*, 1628. (c) Kumar, V. A. Ganesh, K. N. Structure-Editing of Nucleic Acids for Selective Targeting RNA *Curr.Top. Med. Chem.* **2007**, *7*, 715.

19 D'Costa, M.; Kumar,V.A.; Ganesh, K. N. Aminoethylprolyl (*aep*) PNA: Mixed purine/ pyrimidine oligomers and binding orientation preferences for PNA:DNA duplex formation . *Org. Lett.* **2001**, *3*, 1281-1284.

20 (a) Pallan, P. S.; von Matt, P.; Wilds, C. J.; Altmann, K.-H.; Egli, M. RNA-binding Affinities and Crystal Structure of Oligonucleotides Containg Five –Atom Amide –Based Backbone Structures *Biochemistry***, 2006,** *45*, 8048-8057. (b) Wilds, C. J.; Minasov, G.; von Matt, P.; Altmann K.-H.; Egli, M. Studies of a chemically modified oligodeoxynucleotide containing a 5-atom amide backbone which exhibits improved binding to RNA *Nucleosides, Nucleosides and Nucleic Acids,* **2001***, 20*, 991-994. (c) Govindaraju, T.; Kumar, V. A. Backbone-extended pyrrolidine peptide nucleic acids (*bep*PNA): design, synthesis and DNA/RNA binding studies *Chem. Commun*., **2005**, 495–497 (d) Govindaraju, T.; Kumar, V. A. Backbone extended pyrrolidine PNA (*bep*PN(A): a chiral PNA for selective RNA recognition *Tetrahedron*, **2006**, *60*, 2321–2330. (e) Gogoi, K.; Kumar, V. A. Sugar-thioacetamide backbone in

oigodeoxyribonucleosides for specific recognartion of nucleic acids *Chem. Commun*. **2008**, 706-708. (f) Worthington, R. J.; Bell, N. M.; Wong, R.; Micklefield, J. RNA-selective cross-pairing of backbone–extended pyrrolidine-amide oligonucleotide mimics (bePOMs) *Org. Biomol. Chem.* **2008,** *6,* 92-103. (g) Gogoi, G.; Kumar, V. A. Chimeric (amino acids + nucleoside- amino acid) peptide oligomers show sequence specific DNA/ RNA recognition *Chem. Commun.* **2008**, 706-708.

21 **(**a) Petersen, G. V.; Wengel, J. Synthesis of thymidine dimers containing piperizine inthe internucleoside linkage and their incorporation into oligodeoxynucleotides *Tetrahedron* **1995**, *51*, 2145-2154. (b) Lauritsen, A.; Wengel, J. Oligodeoxynucleotides containing amide linked LNA-type dinucleotides: synthesis and high-affinity nucleic acid hybridization *Chem. Commn.* **2002**, 530-531. (c) Vilaivan, T.; Lowe, G.A novel pyrrolidinyl PNA showing high sequence specificity and preferential binding to DNA over RNA *J. Am. Chem. Soc*. **2002**, *124*, 9326-9327

22 Slaiatas, A.; Yeheskiely, E. A novel N-(pyrrolidinyl-2-methyl)glycine-based PNA with a strong preference for RNA over DNA *Eur. J. Org. Chem*. **2002**, 2391

23 (a) Gangamani, B.P.; Kumar, V. A. ; Ganesh, K. N. Synthesis of N-alpha-(purinyl/pyrimidinyl acetyl)-4-aminoproline diastereomers with potential use in PNA synthesis *Tetradheron* **1996**, *52,* 15017. (b) Gangamani, B.P.; Kumar, V. A. ; Ganesh, K. N. 2-aminopurine peptide nucleic acids (2-apPN(A): intrinsic fluorescent PNA analogues for probing PNA-DNA interaction dynamics *J. Chem. Soc., Chem. Comm*. **1997**, 1913-1914.

24 (a) Egholm, M.; Buchardt, O.; Nielsen, P. E. Peptide nucleic acids (PNAs): oligonucleotide analogues with achiral peptide backbone . *J. Am. Chem. Soc*. **1992**, *114*, 1895-1897. (b) Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. Recognition of guanine and adenine in DNA by cytosine and thymine containing peptide nucleic acids(PNAs) *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678. (c) Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. Synthesis of peptide nucleic acid monomers containing the four natural bases: Thymine, cytosine, adenine and guanine and their oligomerization *J. Org. Chem.* **1994**, *59*, 5767-5773.

25 Erickson, B. W.; Merrifield, R. B. Solid Phase Peptide Synthesis. In the Proteins Vol. II, 3rded.; Neurath, H. and Hill, R. L. eds.; Academic Press, New York, **1976**, pp 255. b) Merrifield, R. B.; Stewart, J. M.; Jernberg, N: Instrument for automated synthesis of peptides. *Anal. Chem*. **1966**, *38*, 1905-1914.

26 Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color test for detection of free terminal amino groups in solid-phase synthesis of peptides *Anal. Biochem.* **1970**, *34*, 595

27 Gryaznov, S.; Skorski, T.; Cuco, C.; Nieborowska-Skorska, M.; Chiu, C. Y.; Llyod, D.; Chen, J.; Koziolkiewicz, M.; Calabretta, B. Oligonucleotide N3'→P5' as phosphoramidates antisense agents *Nucleic Acids Research* **1996**, *24*, 1508.

28 Kang, S-H.; Cho, M-J.; Kole, R. Up-regulation of luciferase gene expression with antisense oligonucleotides : Implications and Applications in functional assay development *Biochemistry*, **1998**, 37, 6235-6239

29 (a) Feichtinger, K.; Sings, H. L.; Baker, T. J.; Matthews, K.; Goodman, M. Triurethan-protected Guanidines and Triflyldiurethane-Protected Guanidines: New Reagents for Guanidinylation Reactions *J. Org. Chem*. **1998**, *63*, 8432-8439. (b) Baker, T. J.; Rew, Y.; Goodman, M. Novel reagents and reactions for drug design* *Pure Appl. Chem.* **2000**, *72*, 347-354.

30 (a) Bernatowicz, M.S; Wu, Y.; Matsueda, G. R. *J. Org. Chem.* **1992**, *57,* 2497. (b) Yong, Y. F.; Kowalski, J. A.; Thoen, J. C.; Lipton, M. A. A new reagent for solid and solution phase

synthesis of protected guanidines from amines *Tetrahedron Lett.* **1999**, *40*, 53-56. (c) Zhang, Y.; Kennan, A. J. Efficient Introduction of Protected Guanidines in BOC Solid Phase Peptide Synthesis *Org. Lett.* **2001**, *3*, 2341-2344. D0 Mamai, A.; Madalengotia, J. S. Solid-Phase Guanidinylation as a Diversification Strategy of Poly-L-proline Type II Peptide mimic Scaffolds *Org. Lett.* **2001**, *3*, 561-564.

31 Young, Y.F; Kowalski, J. A.; Lipton, M. A. Facile and Efficient Guanylation of Amines Using Thiourea and Mukaiyama's Reagent *J. Org. Chem.* **1997**, *62*, 1540-1542.

32 Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. Solid-Phase synthesis of peptide nucleic acids *J. Peptide Sci*. **1995**, *3*, 175.

33 (a) Gait, M. J. Oligonucleotide synthesis: A practical approach. IRL Press Oxford, UK 217 (b) Agrawal, S. PROTOCOLS FOR OLIGONUCLEOTIDES AND ANALOGS Synthesis and Properties Methods in Molecular Biology. (ed): vol.20 Totowa, NJ,. Humana Press, Inc.

34 Puglisi, J. D.; Tinco, I. absorbancy melting curves of RNA *Jr. Methods Enzymol*. **1989**, 180, 304.

35 Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Gräslund, A. Ionic Effects on the Stability and Conformation of Peptide Nucleic Acid Complexes *J. Am. Chem. Soc.* **1996**, *118*, 5544-5552.

Chapter 3

(Section-1)

Design and Synthesis of *iso*-TANA dimers, their incorporation into DNA oligomers and biophysical studies

(Section-2)

Synthesis of Modified PNA sequences with *iso*-thioacetamidonucleic acids monomers (*iso*-TANA) and their Biophysical studies

Chapter 3

Section 3.1 Design and Synthesis of *iso*-TANA dimers, their incorporation into DNA oligomers and biophysical studies

3.1.1 Introduction

The preorganized structures of natural nucleic acids continue to attract broad interest. Modifications of the nucleic acid backbone influence the geometry of the double helical structures. Such modifications have been introduced to increase the enzymatic stability and the hybridization properties of the artificial nucleic acids, so they can be better antisense or antigene agents.[1] The development of artificial analogues of DNA involves major modifications at sugar phosphate backbone or sugar modifications.[2] The current backbone modifications include second and third generation antisense ONs such as 2'-O-alkyl, 2'-O-methylthio ethyl, LNA, HeNA, Morpholino NA, Peptide Nucleic Acids (*aeg* PNA, Figure 1) or PNA modifications. All these modifications exhibit excellent properties in some way but lack some other important properties. PNA is one of the modification where the complete sugar-phosphate backbone is replaced by amide linkage. PNA forms strong PNA:DNA/PNA:RNA complexes[3]. Replacement of phosphate linkages in DNA by the robust amide bond as in PNA (Figure 1) was used as an alternative linker group between nucleosides, to get the advantages of chirality and 3'-5' directionality of backbone (Figure 1, amide linked DNA). The five-atom amide linker in DNA improved the preference for binding to RNA over DNA.[4]

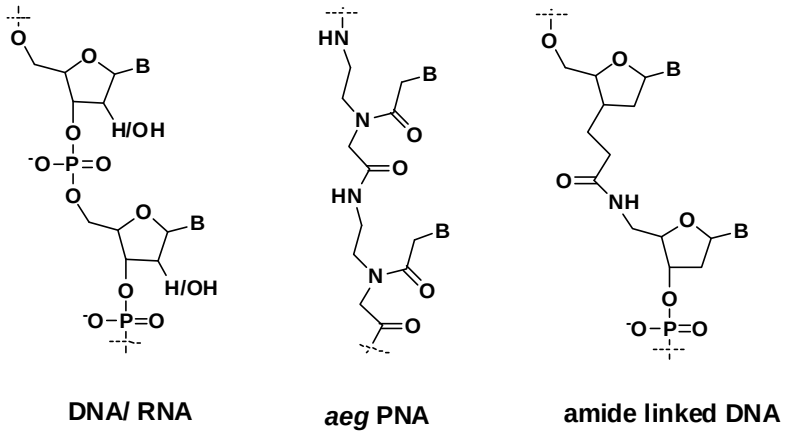


Figure 1A: Structures of DNA, PNA and amide linked DNA

In natural nucleic acids, pentose sugars are puckered or twisted to give preferred helical conformations. These pentose sugar moieties are puckered in order to minimize

non bonded interactions between their substituents. Sugar rings in natural nucleic acids are found to be in equilibrium between two forms (Figure 1B) *viz*. N type (C-3'-endo) and S type (C2'-endo). The 2'-endo (S type) sugar conformation is preferred in DNA whereas the 3'-endo (N-type) form is preferred in RNA due to the presence of 2'-hydroxy group in ribose sugar. In DNA, the 3'- substituent of the sugar ring and anomeric effect of nucleobase affect the sugar puckering in the nucleoside.[5]

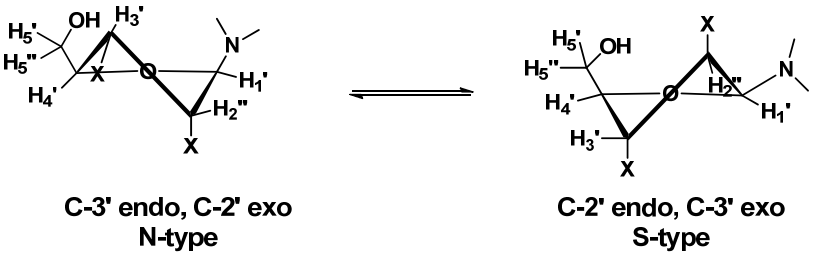


Figure 1B: Puckering of sugar ring

3.1.2 Rationale, Design of *iso*-TANA and Objectives of the present work

Natural oligonucleotides (ON) with phosphodiester linkage are susceptible for hydrolysis with various enzymes like nucleases and this limits their use for various biological applications. PNA has emerged as promising DNA analogue due to its major advantage of stability towards these enzymes. PNA suffers from other drawbacks like less directiontional discrimination[3] and to address this problem, chiral PNAs were designed by introducing chirality in the backbone.[6] Replacement of phosphodiester linkage in DNA/RNA by amide linkage may solve the problem of directional selectivity of PNA and give the stability towards different hydrolytic enzymes to natural nucleic acids. There are reports about various amide-linked oligonucleotides (Figure 2) in the literature.[7]

De Mesmaeker *et al.* reported the four atom amide linked oligonucleotides in which the phosphodiester linkage was replaced with an amide linker (am-1). These oligomers showed higher affinity for target RNA than DNA; also this analogue showed better results than the other five atom amide linked analogous (am-2) for RNA binding.[7e] As am-1 analogue was not investigated in detail, the general effect of this modification on RNA binding is not known. The five atom amide linkers were introduced to compensate for the shorter amide backbone as compared to the phosphate linker in the dinucleoside to maintain the internucleotide distance complementarity. This strategy was successful in increasing affinity to target RNA than DNA.[4,7] Pallen *et*

al. carried out systematic study for RNA binding affinity of modified oligonucleotides with five atom amide linked DNA analogues (Figure 2).[4a] The five atom amide linked oligonucleotides (am-3, am-4) showed increase in the binding affinity with target RNA, than the wild DNA-RNA complex whereas am-5 modification decreased the RNA binding affinity.

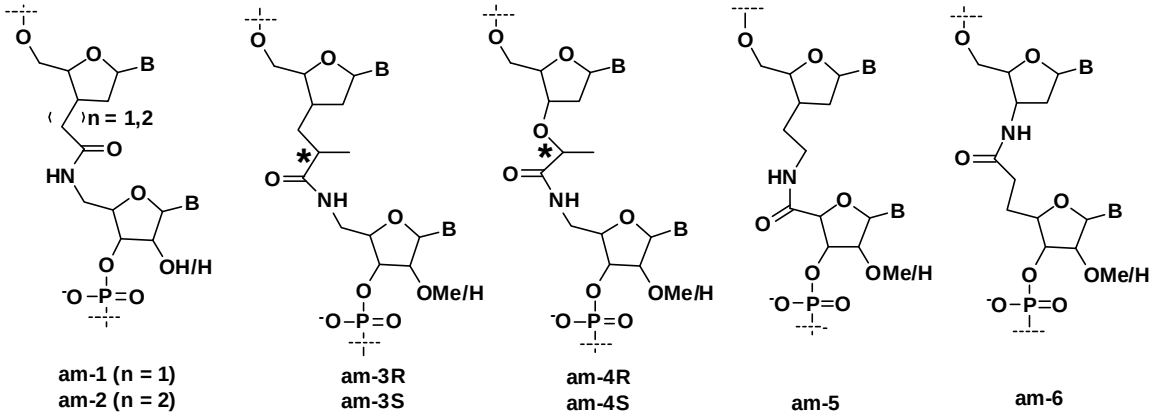


Figure 2: amide linked oligonucleotides

The 2'-methoxy substituent in am-3 and am-4 in the 3'-end sugar of dimer showed favourable effect for binding with RNA irrespective of amide linkage. The effect of the configuration of the stereocenter in the amide linkage also has been studied. The am-3R modification was found to be more favourable for RNA-DNA duplex stability than corresponding S configured backbone modification. The am-4S modification stabilized the DNA-RNA duplex than am-4R configured amide.

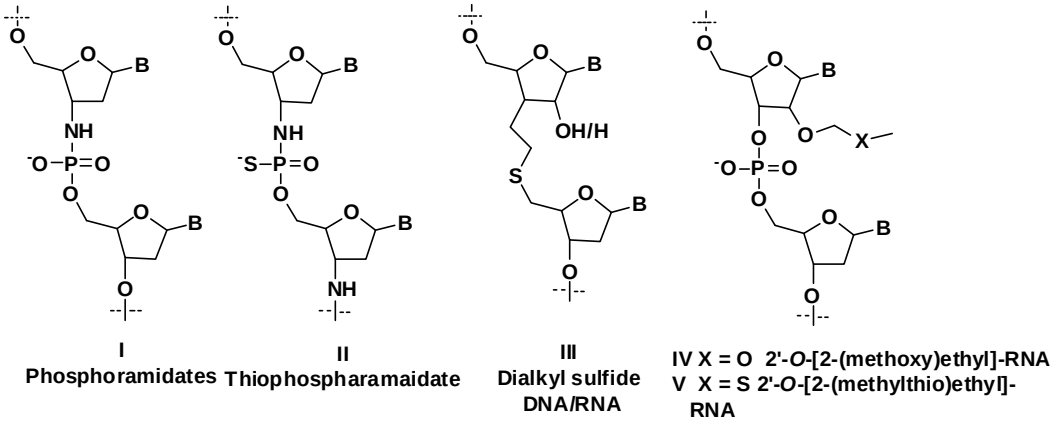


Figure 3: N3'→P5' phospharamidate, N3'→P5' thiophospharamidate, dialkyl sulfide linked DNA/RNA , 2'-O-[2-(methylthio)ethyl] RNA

NMR study of 3'-aminonucleosides showed that 2'-deoxyfuranose ring predominantly is in the N-type (C3'-endo) conformation.[5] The N3'→P5' phospharamidate (Figure 3, I) ONs exhibit preferred binding to RNA which was

attributed to the predominant N-type sugar conformation of the sugar ring due to the 3'-amino substitution.[8] Moreover, circular dichroism (CD) spectroscopy of the phospharmidate duplexes also indicated a general RNA-like A form of the duplexes being formed.[9]

The presence of the thio group in the backbone of oligonucleotides is reported to show improved bioavailability.[10] The 2'-modified, 2'-*O*-[2-(methylthio) ethyl] or 2'-*O*-MTE, oligonucleotides (Figure 3, IV) exhibited high binding affinity to target RNA and also improved binding to human serum albumin compared to the 2'-*O*-MOE modified oligonucleotides.[11] Nuclease digestion of 2'-*O*-MTE oligonucleotides showed that they have limited resistance to exonuclease degradation. N3'→P5' thio phosphoramidates (Figure 3, II) ONs contain the combination of phosphorothioates and phosphoramidates.[11] These ONs showed comparable binding with complementary RNA as phosphoramidate.[12] Conjugates of these ONs with lipids showed good cellular uptake and target specificity *in vitro*. *In vivo,* theses ONs show good bioavailability and efficient bio-distribution to all organs. The alkyl sulfide linked oligonucleotides[13] (Figure 3, III) were reported to form stable complexes with both DNA and RNA but with reduced affinity. The less stability of complexes with RNA was improved by introduction of some ribo nucleoside units in the oligonucleotides.[14]

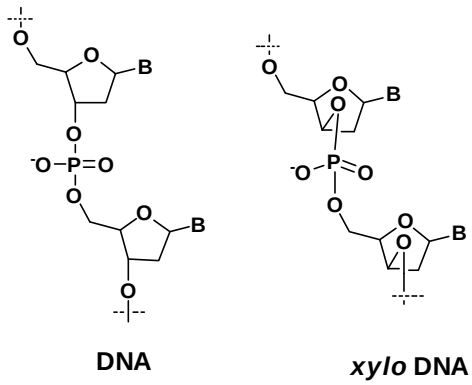


Figure 4: *xylo* configured DNA

As discussed earlier, the anomeric effect of the nucleobase and gauche interactions between 3'-α-substituent and *O*4' are the major factors for S-type conformational preference of sugars in 2'-deoxy nucleosides.[5] In DNA, due to the presence of 3'-OH or 3'-O-phosphate group, sugars prefer to be in S-type conformations. On the other hand, in the *xylo* (3'β-OH) DNA, (Figure 4) the preference

for sugar puckering changes from S-type to structurally more favourable N type.[15] Thus these sugars showed similar preference like 3'-amino substituted nucleosides.

Design of *iso*-TANA: The thioacetamido (TANA, Figure 5) linked oligomers in which the 3'-amino substituent is in ribo configuration, earlier reported from our laboratory have shown preferential binding with RNA over DNA. These oligomers combine the structural features of phosphoramidates, 2'-*O*-MTE ONs, 5-atom amide linker containing ONs. NMR studies of TANA dimers showed that the predominant form of sugar rings in the dimer was S-type. To take advantage of both the factors that favour N-type sugar conformations, i.e 3'-amino substitution in *xylo* configuration, in the present work, we designed the *iso*thioacetamido linked analogue (*iso*-TANA) with inverted stereochemistry at 3'-amino substitution compared with the ribo configured TANA thioacetamido linkage. The aim of design of *iso*-TANA analogue is to study the effect of change in the stereochemistry at 3'amino substitution on the sugar puckering, base stacking interaction and binding affinity with complementary DNA/RNA.

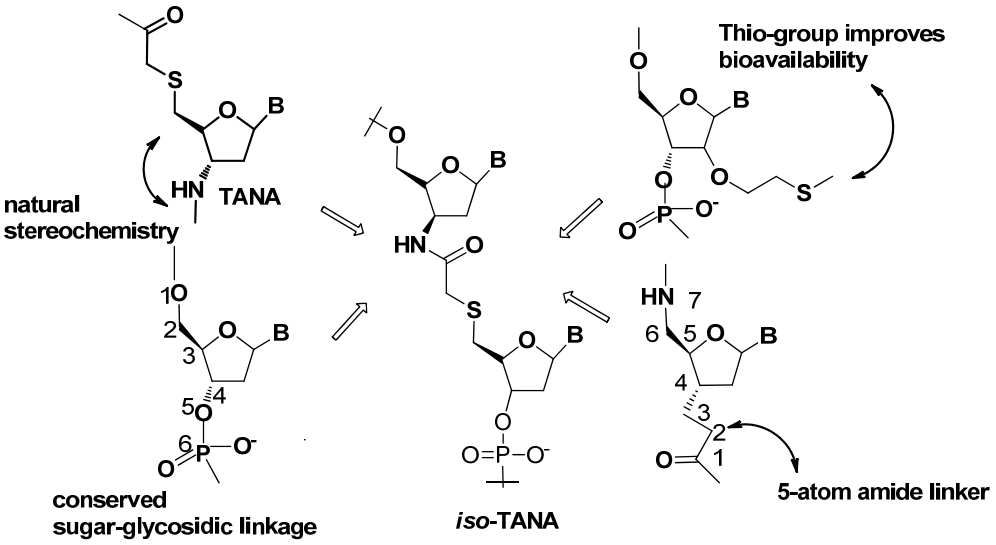


Figure 5: Design of *iso*-TANA

The designed *iso*-TANA linker will have all other attributes of TANA such as 5- atom amide linkages, presence of thio group in backbone of ONs and the base attachment to the sugar ring as in the natural nucleic acids.

Objectives:

1. a) Synthesis of *iso*-TANA dimers *iso*-t_st, *iso*-c_st and *iso*-c_sc and TANA dimer c_sc and their phosphoramidite derivatives

b) NMR studies of the dimer blocks for sugar ring conformational analysis and the CD spectroscopic analysis to study comparative base stacking interactions

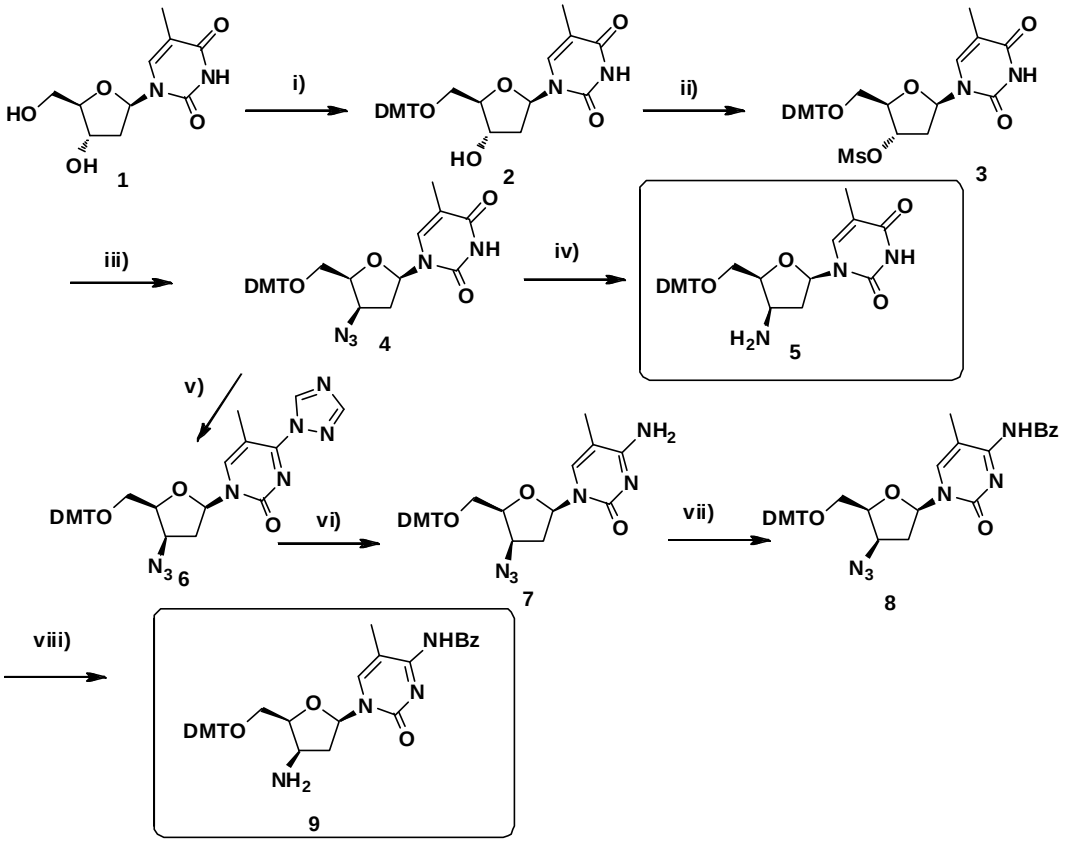
2. Synthesis, purification and mass spectral analysis of the chimeric *iso*-TANA-DNA oligomers and their binding studies with complementary DNA and RNA sequences

3.1.3 Methodology, Results and Discussion

3.1.3a Synthesis *iso*-thioacetamido linked Thymidine-Thymidine (*iso*-t_st), 2'-deoxy-5-Methylcytidine-Thymidine(*iso*-c_st), 2'-deoxy-5-Mehtyl cytidine–2'-deoxy-cytidine dimers (*iso*-c_sc) and thioacetamido linked 2'-deoxy-5-Mehtyl cytidine–2'-deoxy-cytidine dimer (c_sc) and their phosphoramidite derivatives

Synthesis of the *iso*-thioacetamido linked dimers is divided into three parts, 1) synthesis of *xylo* configured 3'- amino nucleosides (Scheme-1), 2) synthesis of 5'mercapto acetic acid derivatives of thymidine and 2'-deoxy cytidine (Scheme-2), and 3) synthesis of dimers by coupling reaction of amine and acid constituents and conversion of these dimers into their phospharamidite derivatives (Scheme-3).

1) Synthesis of *xylo* configured amino nucleosides



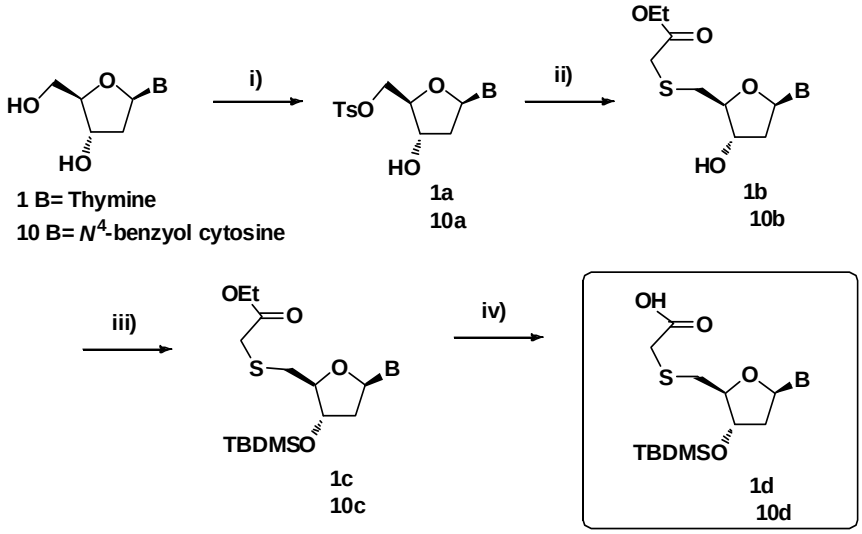
Scheme 1: Reagents and conditions , i) DMTrCl/pyridine , 87%, ii) MsCl/ Et_3N, DCM, 85%, iii) NaN_3/DMF, 60 ^oC, 45 %, iv) H_2/ Pd-C, MeOH, 86%,v) 1,2,4- triazole, $POCl_3$, Et_3N, MeCN, 88%, vi) conc. NH_3, 1-4 dioxane, 90%, vii) BzCl, pyridine, 75%, viii) H_2S/ pyridine, Et_3N, 78%

3’-deoxy-*xylo*-3’-amino thymidine **5** and 3’-deoxy-*xylo*-3’-amino-5-mehtyl-2’-deoxycytidine **9** were synthesized from common natural precursor thymindine **1** (Scheme-1). 5’-OH of thymidine was protected as DMT by reaction with DMT-Cl in pyridine. The protected derivative **2** was converted to its 3’-*O*-mesyl derivative **3**. The 3’-*O*-mesyl derivative was used without further purification in the next step. 3’-deoxy-*xylo*-3’azido derivative **4** was obtained *via* SN^2 displacement of 3-*O*-mesyl group by treatment with sodium azide with inversion of configuration at C3’. The isolated yield of **4** was 45%. Pyrimidine nucleobases are known to interfere in the substitution reactions at 3’ position of sugar, giving rise to side products and reduced yield of the required compound products.[16] Azide functionality in **4** was ascertained by IR spectroscopy which showed absorption band at 2100 cm^{-1}. The azido derivative **4** was reduced to amino derivative **5** by catalytic hydrogenation. The azido derivative **4** was also used for its conversion to 5-methyl-2’-deoxycytidine derivative (Scheme-1). The compound **4** was converted to triazole derivative **6** by reaction with triazole–$POCl_3$ complex. The crude triazole derivative was treated with conc. NH_3 solution to give 5-methyl-2’-deoxy-cytidine derivative **7**. Exocyclic amino group in 5-methyl-cytosine derivative was protected with benzyol group to give N^4-benzoyl-2’, 3’-dideoxy *xylo*-3’-azido-5’-*O*-dimethoxytrityl-5-methyl cytidine **8**. Catalytic hydrogenation of **8** was attempted but reaction was very sluggish and the starting material was recovered after long reaction time. The reduction of azido functionality in **8** was then achieved by hydrogen transfer reaction using H_2S-pyridine complex to give N^4-benzoyl-2’, 3’-dideoxy *xylo*-3’-amino-5’-*O*-dimethoxytrityl-5-methyl cytidine **9**.[17]

2) Synthesis of 5’- mercaptoacetic acid derivatives of thymidine and 2’-deoxy cytidine

The 5’-mercaptoacetic acid derivatives of thymidine and 2’-deoxy cytidine were synthesized from corresponding nucleosides thymidine **1** and 2’-deoxy cytidine **10** (Scheme-2). The 5-’tosyl derivatives **1a**/**10a** were obtained by selective tosylation of thymidine and N^4-benzoyl -2’-deoxy-cytidine. The selective tosylation was achieved under high dilution conditions and by controlled addition of tosyl chloride at 0 °C. Mercapto nucleophile generated *in-situ* by reaction of NaH and ethyl mercaptoacetate was used to substitute the tosyl group from **1a**/**10a** to give mercapto derivative **1b**/**10b**. 3’-OH was protected as TBS ether by reaction with TBSCl in DMF to get **1c**/**10c**. Silyl protection was used to improved the solubility of these compounds. Hydrolysis of the

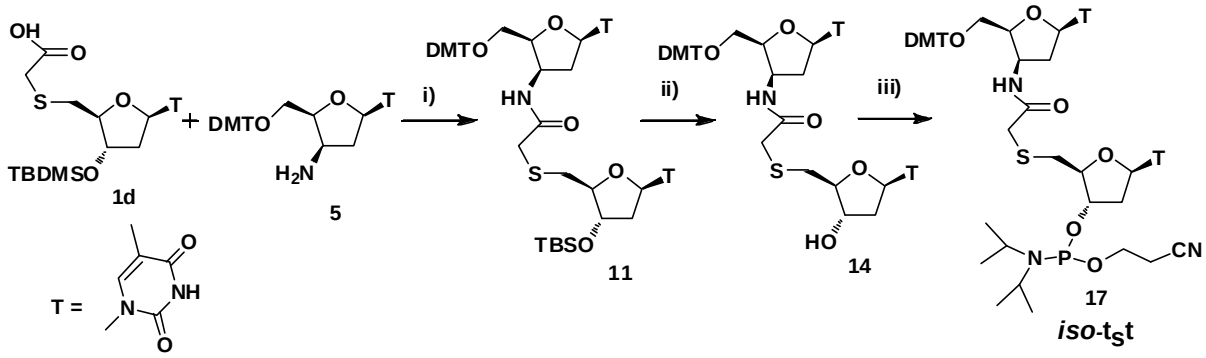
ester group in **1c** was carried out by 1N LiOH in methanol to give 5'-mercapto acetic acid derivative of thymidine **1d**.



Scheme 2 reagents and conditions: i) TsCl/pyridine, 60%, ii) ethyl mercapto acetic acid/NaH/DMF, 80%, iii) TBSCl/imidazole/DMF, 90%, iv) B = Thymine 1N LiOH/MeOH, 88%, B= N^4-benzyol cytosine 1N NaOH/ THF, 0°C, 75%

While attempting ester hydrolysis for compound **10c** under these conditions, benzoyl protection of cytosine exocyclic amino group also got hydrolysed. Hydrolysis of ester group in **10c** was attempted with different conditions like triethyl amine/water, DBU/water, but these conditions were not successful for hydrolysis of ester group. The selective hydrolysis of the ester group in **10c** was achieved with 1N NaOH in THF-water mixture at 0 °C. Under these conditions hydrolysis of ester group was achieved keeping benzoyl protection of the cytosine base intact to give product **10d**.

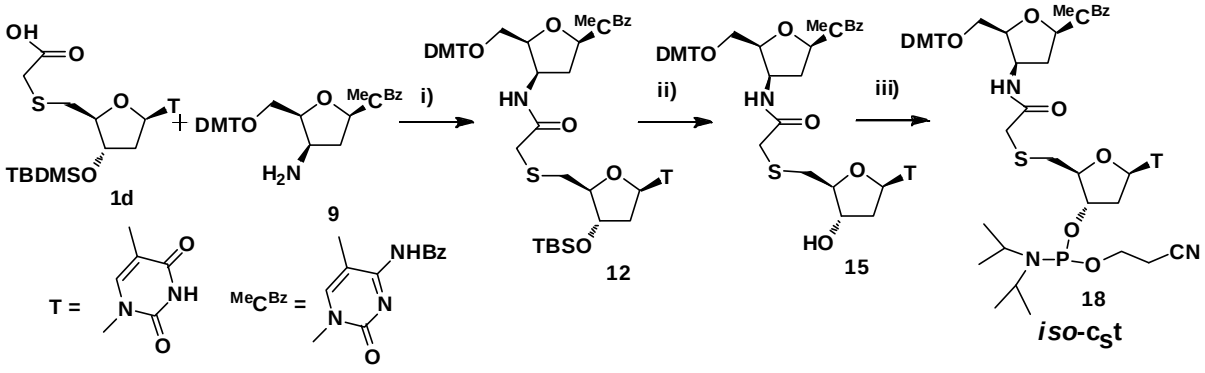
3) Formation of amide linked dithymidine (*iso*-t_st), 2'-deoxy-5-methylcytidine-thymidine (*iso*-c_st) and 2'-deoxy-5-methylcytidine-2'-deoxy cytidine (*iso*-c_sc) dimers and their phospharamidite derivatives



Scheme 3a: Reagents and conditions, i) TBTU/ HOBt, DIPEA, ACN:DMF(5:1), 61% ii) 1N TBAF in THF, 87%, iii) *N*,*N*'-diisopropylcyanoethyl chlorophosphine, DIPEA, DCM, 78%

The three dimers were synthesized by amide formation between the individual *xylo* configured amino component **5**/**9** (scheme-1) with the corresponding mercaptoacetic acid component **1d**/**10d** (scheme-2). Synthetic scheme of these dimers is outlined in scheme-3a-3c.

The protected *iso*-TANA t_St dimer **11** (scheme-3a) was synthesized by coupling of 3'-amino thymidine derivative **5** with 5'-mercaptoacetic acid derivatives of thymidine **1d** with coupling reagent TBTU and HOBt. 5'-mercaptoacetic acid derivatives **1d** was used in excess than the amino derivative **5** to ensure the complete consumption of amino derivative. Dimer **11** was isolated and purified by column chromatography. 3'-*O*–silyl protection in **11** was subsequently removed by treatment with 1N TBAF in THF to give the dimer **14.** 3'-hydroxy in **14** was phosphitilated to get the phospharamidite **17**.

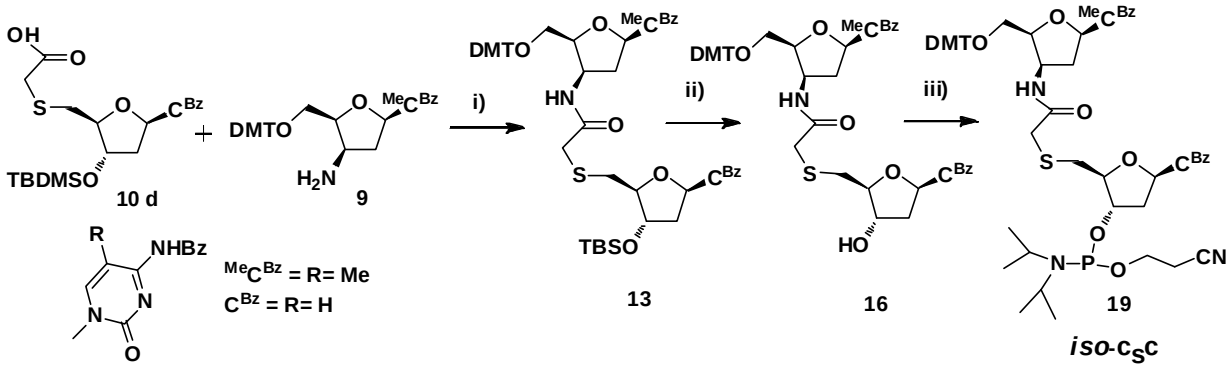


Scheme 3b: Reagents and conditions i) TBTU/ HOBt, DIPEA, ACN, DMF, 58%, ii) 1N TBAF in THF, 80%, iii) *N*,*N*'- diisopropylcyanoethyl chlorophosphine, DIPEA, DCM, 70%

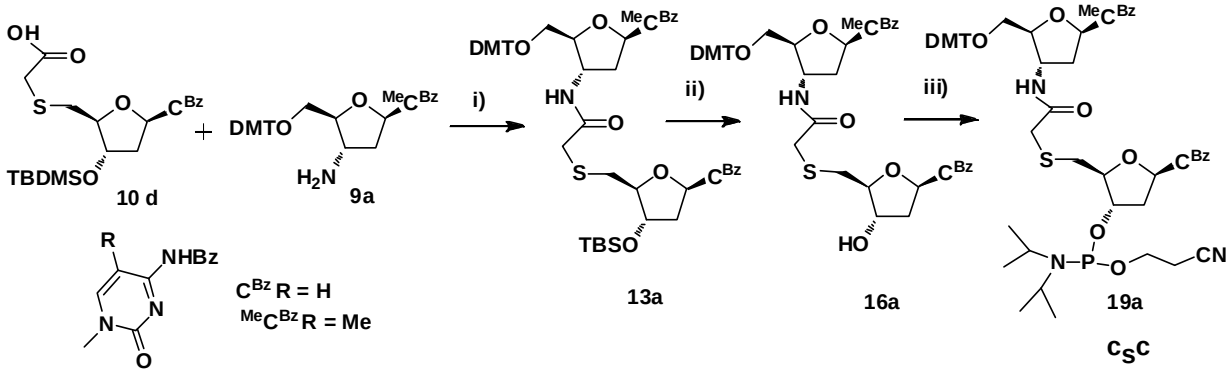
iso-TANA 5-mehtyl cytidine–thymidine dimer **12** (scheme-3b) was synthesized by coupling of 5'-mercapto acetic acid derivative **1d** with N^4-benzoyl-2',3'-dideoxy*xylo*-3'-amino-5'-*O*-dimethoxytrityl-5-methylcytidine derivative **9**.using TBTU and HOBt. Dimer **12** was purified and converted to 3'- hydroxyl free dimer **15** by treatment of 1N TBAF solution. This dimer **15** was converted to phospharamidite derivative **18** by reaction of phosphitylation.

TANA and *iso*-TANA 5-mehtyl cytidine–cytidine dimers (**13** *iso*-c_Sc /**13a** c_Sc Scheme 3c/ Scheme 3d) were synthesized by coupling of 5'-mercaptoacetic acid 2'-deoxycytidine derivative with 2'-3'-dideoxy-3'-amino-5-methyl cytidine (**9a**)[4g] and it's epimeric amino compound (**9**) respectively. (**13**/**13a**) dimers on desilylation gave compounds **16**/**16a** and on subsequent phosphitilation of these dimers gave

corresponding phosphoramidites **19**/**19a**. As the coupling efficiency of acid and amino parts was low, moderate yields of coupling reactions were obtained. These phospharamidite derivatives (**17/18/19/19a**) were used to incorporate dimers into DNA sequences.



Scheme 3c: Reagents and conditions i) TBTU/ HOBt, DIPEA, ACN:DMF(5:1), 54% ii) 1N TBAF in THF, 78%, iii) *N*,*N*′-diisopropylcyanoethyl chlorophosphine, DIPEA, DCM, 70%



Scheme 3d: reagents and conditions, i) TBTU/ HOBt, DIPEA, ACN, DMF, 55%, ii) 1N TBAF in THF, 80%, iii) *N*,*N*′- diisopropylcyanoethyl chlorophosphine, DIPEA, DCM, 73%

3.1.3b NMR and CD studies of *iso*-TANA dimers

In natural nucleic acids, the pentose sugars are puckered or twisted to give preferred helical conformations. These pentose sugar moieties are puckered in order to minimize non bonded interactions between their substituents. This puckering is described by identifying the major displacement of carbon C-2' and C-3' from the median plane of *C*-1'-*O*4'-*C*-4'. The detailed study of X-ray crystal structures of nucleosides and nucleotides revealed that ribo and deoxyribofuranosyl moieties occur preferentially within two distinct major conformational(Figure 6) to as north (C-3' endo) and south (C-2' endo).[18] In solution, N-type and S-type conformations are in rapid equilibrium and are separated by low energy barrier.

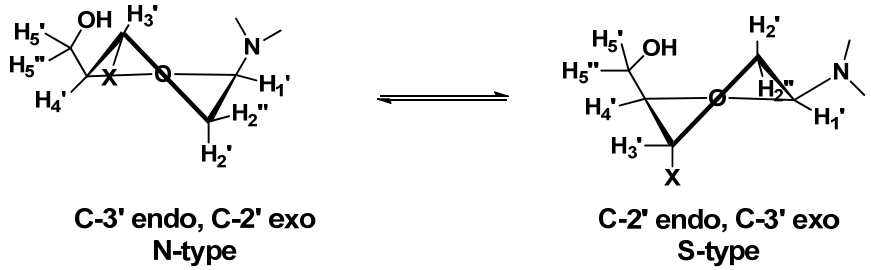


Figure 6: Puckering of sugar ring

The conformation of pentofuranose can be fully described in terms of the phase of pseudorotation (*P*), and puckering amplitude (ψ).[19] For North-type (N) sugars (C3'-endo, C2'-exo), *P* ranges from -1° to 34° and for S-type (S) sugar (C2'-endo, C3'-exo), P ranges from 137°- 194°. The mole fraction of N and S conformers as well as their geometry, expressed by their phase angle of pseudorotation P_N and P_S and puckering amplitude ψ_N and ψ_S, can be calculated from vicinal proton-proton ($^3J_{HH}$) coupling constants $J_{1'-2'}$, $J_{1'-2''}$, $J_{2'-3'}$, $J_{2''-3'}$, $J_{3'-4'}$. These coupling constants can be used as input for pseudorotation analysis of the sugar using PSEUROT.[20] PSEUROT program can not be used in some cases where it is not possible to get the coupling constants accurately over a range of different temperature due to spectral overlap.

Alternatively, a semi empirical method termed as "Sum Rule"[21]can also be employed to have some idea about N↔S equilibrium in the case of sugar conformations. This method was employed in the present work to calculate conformational preference of pentofuranose moieties of thioacetamido linked dimers and other modified nucleosides. The empirical equation to calculate the percentage of S conformer is as below.

% S = (Σ H1'-9.8)/5.9x100 where ΣH1' = $J_{1'2'}$+ $J_{1'2''}$

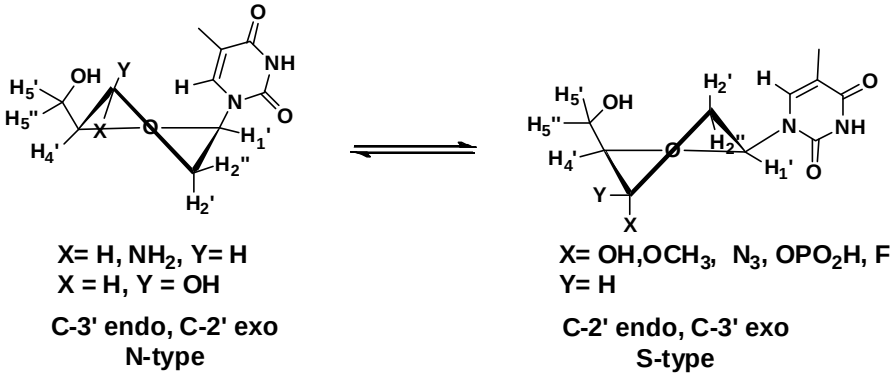
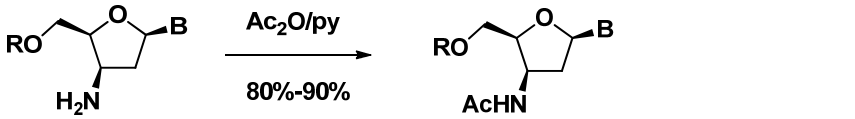


Figure 7: Effect of change in stereochemistry of 3'substituent on sugar conformation

The modified oligonucleotides with propensity for northern sugar conformation could be pre-organized into the required conformation for A-type duplexes and

therefore more likely to form stable hetroduplexes with complementary RNA. On the other hand oligonucleotides with more S type sugar conformation could be preorganize to B type duplex with DNA. Various steric and stereoelectronic effects of sugar skeleton (gauche effect) and the nucleobase (anomeric effect) energetically dictate the pseudorotational equilibrium between two preferred conformations (N/S).[22] The stereoelectronic effect of 3'substituent in case of 2', 3'-dideoxy-3'-substituted nucleosides was studied by Chattopadhyaya *et al*[23]. It was observed that in 2'-deoxy sugars the equilibrium of N↔S was shifted towards S conformation of sugar as electronegativity of substituent increased. This may be due to guache effect of 3'α substituent and O4' which decides the predominance of one of conformer on other.

NMR studies of thymidine has shown that it prefers S-type sugar puckering (58% S) whereas 2'-deoxy-*xylo*thymidine prefers the N-type sugar puckering (84% N type)[24], thus change in stereochemistry at 3'-carbon in nucleoside also shifts N↔ S equilibrium. In the present section the studies towards finding out the conformational preference of the deoxyribose ring in 3'-acetamido *5'-O*-TBS-2'-3'dideoxy-*xylo*furanosyl nucleosides and *iso*-TANA dimers are determined by NMR, and the results are compared with the corresponding 3'-acetamido *5-O*-TBS-2'-3'dideoxy ribofuranosyl nucleosides **20** and **21** (Figure 8) and TANA dimers **22a** and **23a**(Figure 8). 3'-acetamido *5'-O*-TBS-2'-3'dideoxy-*xylo*furanosyl nucleosides were synthesized from intermediates described previously in the earlier section.

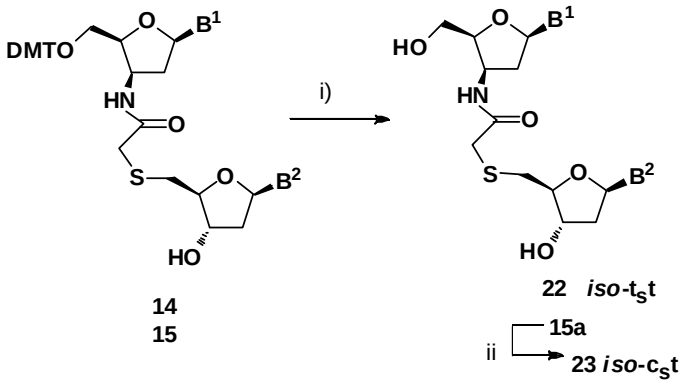


5 B= Thymine
9 B= N^4-benzyol 5-methyl cytosine

20 B= Thymine
21 B= N^4-benzyol 5-methyl cytosine

Scheme 4a: **Synthesis of** 5'-*O*-TBS-2'-3'dideoxy-*xylo*furanosyl nucleosides

The 3'-acetamido derivatives of 2'-3'dideoxy nucleoside were studied to see the effect of conversion of amino group to the amide functional group on sugar puckering of corresponding nucleosides**.** The completely deprotected dimers **22** (*iso*-t_St) and **23** (*iso*-c_St) were used for the NMR studies in water. These dimers were synthesized from previously described 5'-ODMT protected dimers **14** and **15** (Scheme 4b).



Scheme 4b: Reagents and condition, i) 2%-dichloracetic acid in DCM, ii) conc. NH_3/ 40 °C

The assignment of the non-exchangeable proton resonances was achieved by using 2D cosy experiments. The scalar coupling constants were determined from 1D ^{1}H NMR spectra. %S character was calculated the sum rule % S = (Σ H1'- 9.8)/5.9 x 100

The results of NMR studies for 3'-amino amd 3'-acetamido derivatives of 2'-3'-*xylo*furanosyl nucleosides and *iso*-TANA dimers are summarized in Table 1

Table 1: The conformational analysis for 3'-amino amd 3'-acetamido derivatives of 2'-3'-*xylo*furanosyl nucleosides and *iso*-TANA dimers

Entry	Compound	Σ H1' (Hz)	%S	Predominant sugar ring pucker
Nucleosides	***xylo* thymidine**[24]	-	16.0	3'-endo
	5	11.1	22.0	3'-endo
	20	13.7	67.1	2'-endo
	9	11.5	28.8	3'-endo
	21	13.4	62.4	2'-endo
Dimers	**22** **A**	13.2	57.8	3'-endo/2'-endo
	22 **B**	13.6	65.9	2'-endo
	23 **A**	12.5	46.7	3'-endo/2'-endo
	23 **B**	13.5	63.7	2'-endo

The conformational analysis results for corresponding TANA monomers (Figure 8) and dimers **22a, 23a** (Figure 8), with natural counterparts are summarized in Table 2.

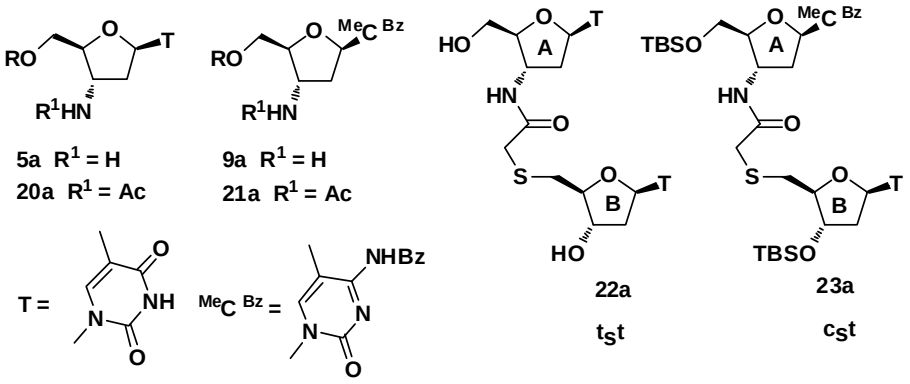


Figure 8: Structures of TANA monomers and dimers

Table 2: The conformational analysis for TANA monomers and TANA dimers and natural nucleoside and dimers[25]

Entry	Compound	Σ H1' (Hz)	%S	Predominant sugar ring pucker
Nucleosides	thymidine	-	66.0	2'-endo
	1c	13.0	54.2	3'-endo/2'-endo
	20a	14.0	71.2	2'-endo
	21a	13.7	66.0	2'-endo
Dimers	TpT A	-	68	2'-endo
	TpT B	-	61	2'-endo
	22a A	13.6	64.4	2'-endo
	22a B	12.5	46.7	3'-endo/2'-endo
	23a A	14.0	71.2	2'-endo
	23a B	13.3	59.3	2'-endo

The sugar rings of 3'-amino nucleoside are known to prefer 3'-endo (N-type) conformation. The sugar rings of 3'-amino nucleosides with **xylo** (**5**, **9**) showed predomint N-type conformation. The shift in equilibrium from N-type sugar conformation to S- type sugar conformation was observed after acetylation of 3'-amino -2'3'-dideoxy nucleosides .

The effect of the change in the stereochemistry at 3'- amino part of thioacetamido linkage was studied by comparing the sugar conformations of *iso*t_st (Table 1)with t_st dimers (Table 2). Sugar ring **A** of t_st dimer **22a** was found to be predominantly in S-type conformation, while in the case of B ring, equal preference for either S-type or N-

type conformation was observed. In the *iso*t_St dimer **22** with unnatural configuration at 3'amino part, both sugars remain in S-type conformation like in the case of phoshophodiester linked dithymidine. The N-acylation of C-3' amino functionality, in either ribo or *xylo* configuration in the dimers, also shifts the nature of sugar pucker to C2'endo (S-type and DNA like).

In *iso*-c_St dimer **23**, sugar **A** ring was found to having equal proportion of S-type and N-type sugar conformation, while sugar ring B was major in S-type conformation. In the case of c_St dimer **23a** both the sugars were observed to be in predominantly in S-type conformation. The NMR analysis thus concludes that the sugar ring conformations are not largely affected either TANA t_St, c_St or *iso*TANA t_St, c_St dimers due to change in the stereochemistry in the amino part of thioacetamido linkage.

CD analysis of TANA monomers and dimers and comparison study of conformational and base-stacking interactions in *iso*-TANA and TANA dimers

Most biological molecules contain a number of electronic units that absorb light nearly independently, called chromophores, which are asymmetrically disposed in the space. Such asymmetric molecules will absorb left circularly polarized light differently from right circularly polarized light. Circular dichroism (CD) is the difference in absorption between left and right circularly polarized light, so that CD spectral bands occur due to the change in the normal electronic absorption bands in an asymmetric molecule.[26]

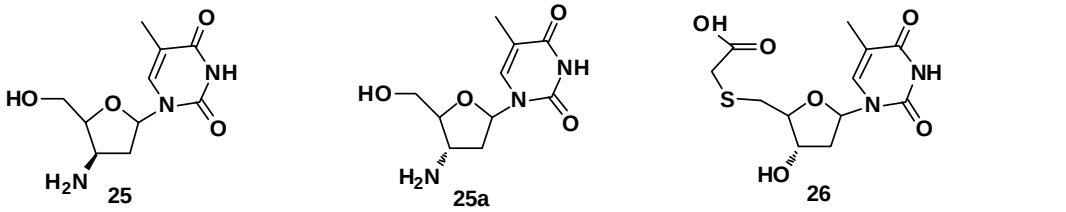


Figure 9: Structures of monomers used for CD analysis

The change in stereochemistry in the thioacetamido linkage of dimers has shown minimal effects on the sugar puckering of dimers in NMR studies. It would be very interesting to study the effect of the change in the chiral center in the backbone sugar on the internucleobase-stacking interactions. It is reported that in *xylo*-configured phosphodiester linked DNA analogues, a clear inversion in the CD spectrum is seen.[24]

This change in CD spectrum i.e. minima at 280 nm is also observed in the case of Z-DNA. To study the change in the backbone sugar configurations in TANA and *iso*-TANA dimers on the base stacking interactions CD analysis was carried out. The CD analysis of monomers was also carried out to see the individual contributions of monomers to the CD pattern of dimers. The protecting groups in all the compounds used for CD analysis were removed to increase the water solubility of compounds. CD analysis of all the compounds was carried out at 100 µmolar concentration. Temperature dependent change in the base-stacking was studied by scanning CD of the dimer solution with increase in temperature from 10 °C-80 °C.

CD spectra of the monomers: CD spectra of all the three monomers (**25**, **25a**, **26**) showed positive band at 270-280 nm. The addition spectra of the two monomer i.e (**25** + **26**) and (**25a**+ **26**) were recorded. Both the addition spectra also give rise to a positive CD signal at 280 nm (Figure 10).

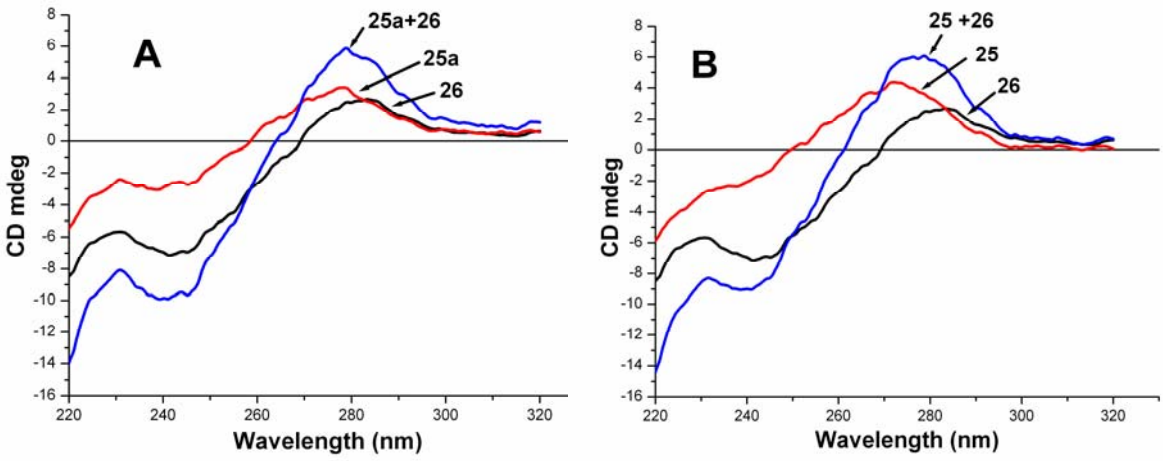


Figure 10: A) CD spectra of monomers **25a** and **26**and their additive CD spectra, B) CD spectra of monomers **25** and **26** and their additive CD spectra

CD analysis for dimers: The CD spectra of the *iso*-t_st /*iso*-c_st dimer block and the t_st /c_st dimer block (Figure 11) was investigated to see the effect of their epimeric nature on the base stacking interactions.

For the t_st dimer block, a strong positive CD band was observed at 280 nm (Figure 12, A). The amplitude in the CD spectrum at 280 nm of the t_st dimer block was much larger as compared to the additive CD of the two individual monomers (Figure 11 A). In contrast, the CD minima at 275-280 nm observed in the case of *iso*-t_st was found to be

opposite in nature to the additive CD of the individual monomers which exhibited a maxima at 280 nm (Figure 12 B).

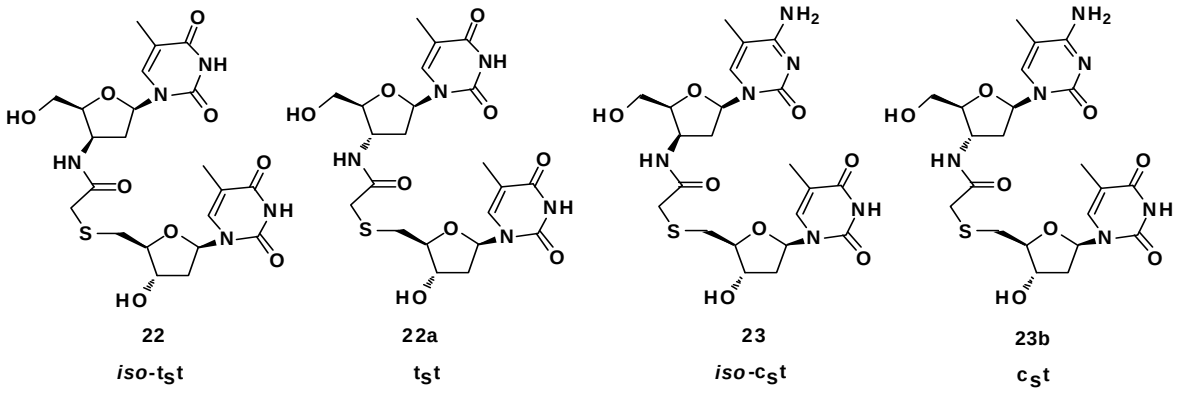


Figure 11: Structures of dimers used for CD analysis

The resulting CD pattern in both the dimers could be thus due to the different stacking interactions due to the adopted secondary conformations of the backbone in t_st and *iso*-t_st. Temperature dependent CD changes were observed in the case of both the dimers. For t_st, the amplitude of positive CD band gradually reduced on heating while for *iso*-t_st dimer the minima found at 280 nm slowly changed to a maximum at 280 nm with increasing temperature. At 80º C, when the dimers would be presumably in the de-stacked state, the CD profiles in both the cases was qualitatively the same as the addition CD spectra of the corresponding monomer units. This observation leads to the conclusion that the CD maximum for the t_st dimer and CD minimum for *iso*-t_st dimer at lower temperature are arising due to the stacking interactions which are different in two epimeric dimers. Very similar patterns of CD differences were observed for TpT and *xylo*- TpT dimer blocks reported earlier.[24] Even in the region of the CD that arises due to the amide absorption at 220 nm is different in two dimers. The minima for t_st dimer (Figure 12 A) whereas maxima for *iso*-t_st dimer at 220 nm in CD spectra was observed (Figure 12 B). The temperature dependant CD change was observed for both dimers. For t_st dimer the amplitude of negative CD band gradually increased on heating while for *iso*-t_st dimer the maxima found at 220 nm slowly changed to minima at 220 nm (Figure 12B) with increasing temperature. The effect of change in configuration of amide part in thioacetamido linkage thus can be observed.

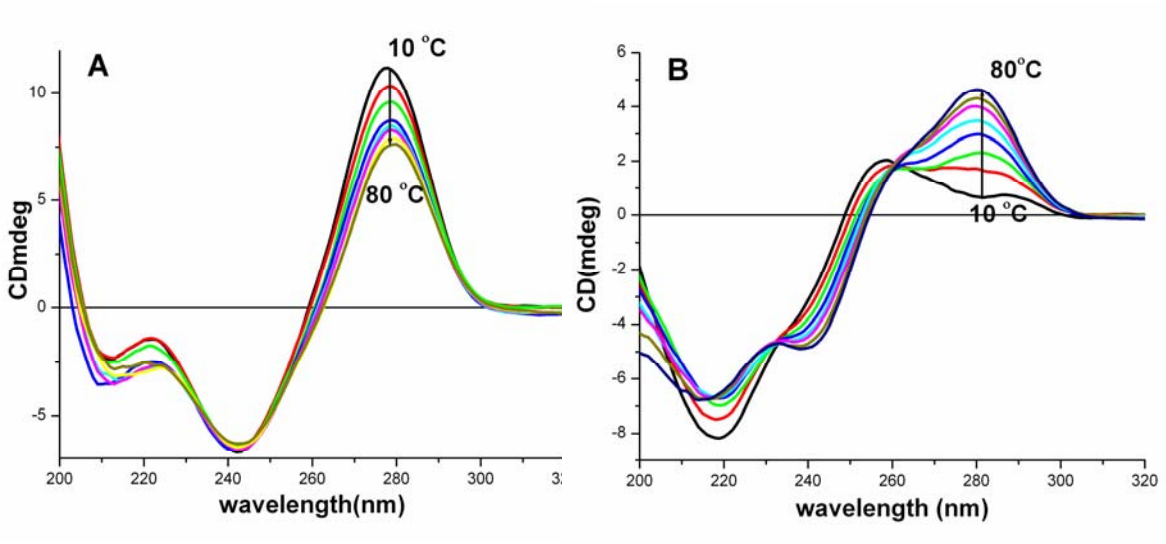


Figure 12: A) Temperature dependant CD spectra of t_st dimer **22a** B) Temperature dependant CD spectra of *iso*-t_st dimer **22**

For the c_st dimer **23**, strong band at 280 nm CD spectrum was observed(Figure 13A). In contrast, the *iso*-c_st dimer **29** showed CD at 280 nm with much less amplitude(Figure 13B).

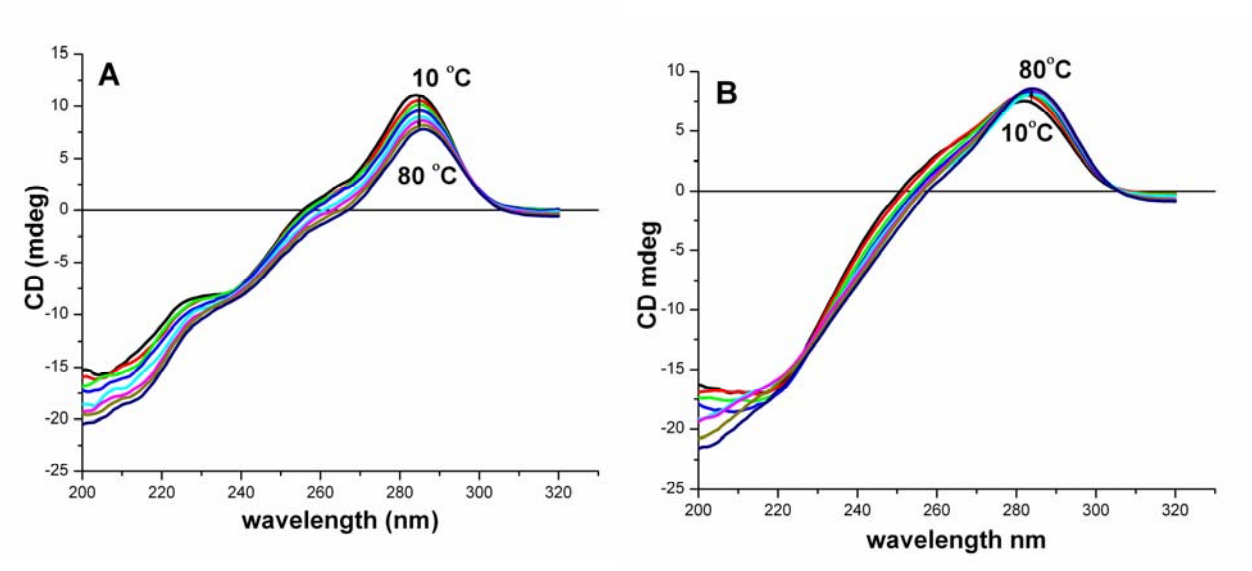


Figure 13: A) Temperature dependant CD spectra of c_st dimer **23a** B) Temperature dependant CD spectra of *iso*-c_st dimer **23**

The decrease in the amplitude of the positve band in CD signal at 280 nm with increasing temperature was observed for c_st dimer. This decrease in the amplitude of CD signal was less as compared to t_st dimer. For *iso*-c_st dimer, increase in the

amplitude CD signal at 280 nm was observed with increase in the temperature (Figure 13).

It is known in the literature that the stacking interaction are maximum in thyminyl dimers and the strength of stacking interactions reduces from thymine ≃ guanine >> adenine ≃ cytosine.[5] CD spectra of the present c_st dimer **23b** and *iso*c_st dimer **23** show less changes on heating when compared with the t_st (**22a)** and *iso*t_st (**22**) but the nature of the CD change i.e decrease in the CD signal amplitude at 280 nm for **23b** and increase in the CD signal amplitude at 280 nm for **23** was observed with increasing temperature. No CD signatutre was observed at 220-240 nm range in the case of c_st **23b** and *iso*-c_st **23** dimers.

The CD spectra of the *xylo* configured TANA dimers thus qualitatively resemble those reported for *xylo*-configured phosphodiester linked nucleic acids. It has been recently observed that the *xylo*-modified oligothymidilate DNA form stable complexes with RNA as compared with DNA,[15e] but the mixed purine-pyrimidine oligomers modified with *xylo*-nucleosides show reduced strength of the complexes with RNA and DNA. [15f] In the next section we report the results of our studies towards the synthesis of *iso*-TANA modified DNA and its effects on DNA/RNA binding selectivity in comparison with the TANA modified DNA reported earlier.

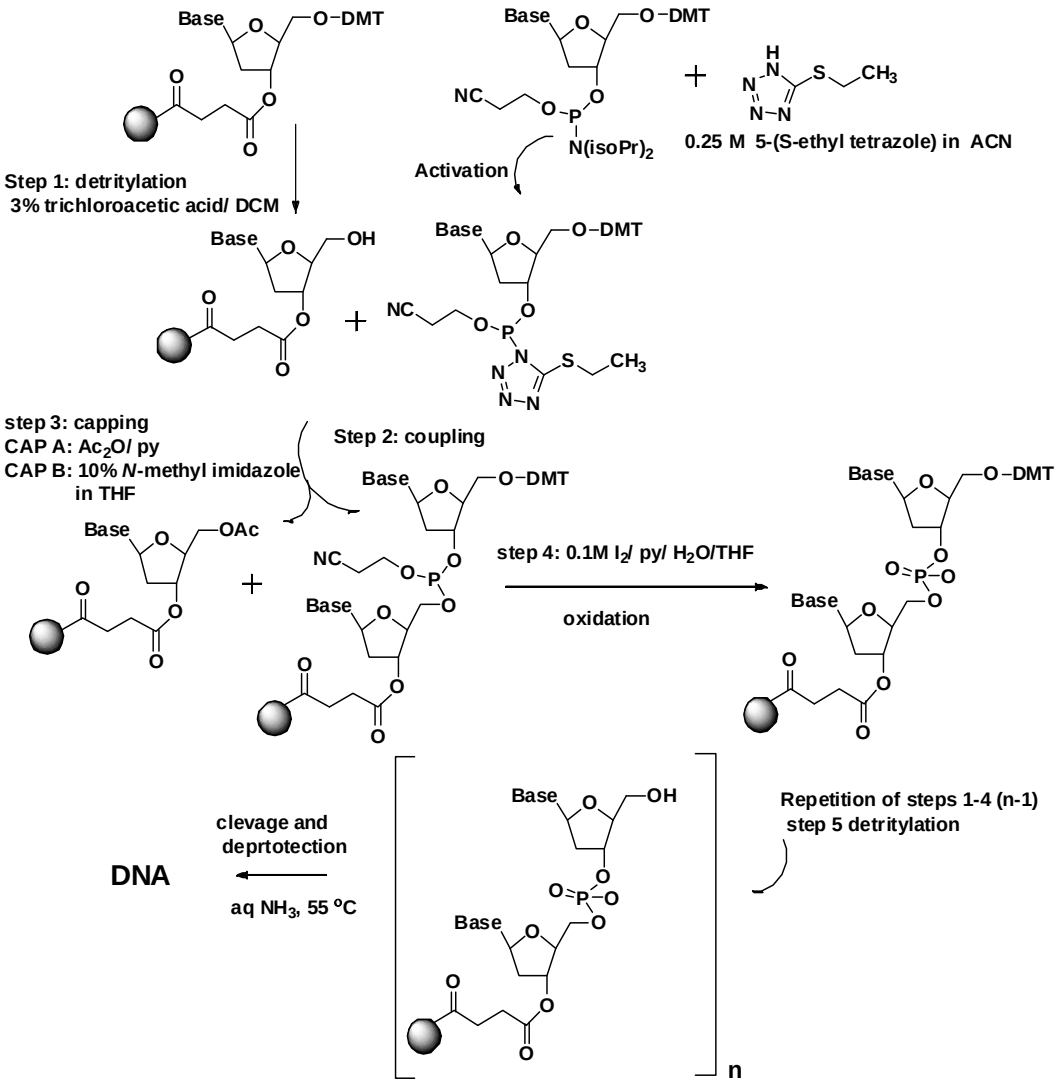
3.1.4 Synthesis, purification and mass spectral analysis of the chimeric *iso*-TANA-DNA oligomers

3.1.4a Solid phase synthesis of *iso*-TANA modified DNA oligonucleotides by phosphoramidite method

General solid phase synthesis of oligonucleotides by β–cyanoethyl phospharamidite method is given in Scheme 5.[27] modified DNA sequences were synthesized using above strategy, incorporating *iso*-TANA dimers at predefined positions. The sequences are listed in Table 3.

Polypyrimidine sequences: Entry 1-3 (Table 3, **ON-1-ON-3**) are the T_8 sequences synthesized incorporating *iso*-t_st dimer. GC clamps were introduced to increase the binding strength and also to avoid the sliding of duplexes. **ON-3** was synthesized where all the thymine units were replaced with *iso*-t_st unit. Mixed polypyrimidine sequences (**ON-4-ON-8**) were synthesized to see the effect of *iso*-TANA dimers on triplex

forming ability of DNA sequences. Thsese sequences were synthesized by incorporating both *iso*-c_st as well as *iso*-t_st dimers in the sequence at predefined positions.



Scheme 5: Solid phase synthesis of DNA

Mixed purine-pyrimidine sequences :In order to study the duplex formation potential of the *iso*-TANA modified DNA sequences with DNA /RNA**, ON-9, ON-10-ON-12** were synthesized by incorporating *iso*-t_st, *iso*-c_st and/or *iso*-c_sc dimers. These sequences are important with respect to their application for splice correction assay.[28] Sequence **ON-10** and **ON-11** are the sequence with single and double *iso*-t_st units respectively. Whereas, sequence **ON-12** contains one *iso*-c_sc unit. We also synthesized sequence **ON-12**, where the modified ribo configured c_sc unit was used for comparison that was

not reported in the earlier reported data. These sequences would give information about the individual effect of isomeric TANA dimers on duplex formation in comparison with previously reported TANA modified sequences.[4g] **ON-13** and **ON-14** are the sequences which are important antisense sequences for tageting b_3a_2 gene present in cancerous human cell lines causing chronic myloid leukema.[29]

3.1.4b Purification and MALDI-TOF characterization of *iso*-TANA modified oligomers

The purity of the oligomers listed in Table 3 was checked by analytical RP-HPLC (C18 column, 0.1 N TEAA Buffer-acetonitrile) which showed more than 75-80% purity. These oligomers were subsequently purified by reverse phase HPLC on C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis. THAP (2,4,6-trihydroxy acetophenone) matrix with diammonium citrate as additive was used for MALDI-TOF mass analysis of oligomers. HPLC retention time and observed values of Mass in MALDI-TOF spectrometry are listed in Table 3.

Table 3: The ON sequences incorporating various *iso*-TANA dimers[a]

Entry	*iso*-TANA Sequence	HPLC[a] t_R (min)	Mass (calc./obsd.)[a]
1	5' CGT Tt_st TTT TGC 3' **ON-1**	7.4	3606.54/3604.85
2	5' CGT Tt_st TTt_stGC 3' **ON-2**	7.7	3596.6/3595.28
3	5' CGt_st t_st t_st t_st GC 3' **ON-3**	9.0	3576.8/3575.54
4	5' $TCTCt_stTCTT$ 3' **ON-4**	8.1	2928.03/2929.29
5	5' TCTC t_st Tc_sT 3' **ON-5**	8.9	2935.08/2937.54
6	5' Tc_stCTT TCTT 3' **ON-6**	8.2	2942.05/2943.96
7	5' Tc_stCTT Tc_stT 3' **ON-7**	8.4	2949.10/2950.44
8	5' Tc_st c_stT Tc_stT 3' **ON-8**	9.0	2956.15/2955.49
9	5' $TCAc_stAGATG$ 3' **ON-9**	8.8	3036.12/3036.55
10	5'CCT $Ct_stACCTCA$ GTTACA 3' **ON-10**	8.2	5366.7/5370.44
11	5'CCT Ct_st ACCTCAGt_st ACA 3' **ON-11**	8.6	5356.9/5353.93
12	5'CCT CTTAc_scTCAGTTACA 3' **ON-12**	9.1	5376.65/5376.85
		9.3[b]	5376.65/5382.82[b]
13	5' GAA GGG$c_stTt_stGAACTc_stT$ 3' **ON-13**	9.3	5841.03/5837.80
14	5' GAA GGG $c_stTt_stGAAc_stc_stT$ 3' **ON-14**	9.6	5849.08/5849.36

a HPLC and MALDI-TOF mass analysis of sequence containing *iso*-TANA unit
b HPLC and MALDI-TOF mass analysis of sequence containing TANA unit

3.1.4c Synthesis of complementary oligonucleotides

The DNA oligomers (Table 2A) were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard *β*-cyanoethyl phosphoramidite chemistry.[27] The oligomers were synthesized in the 3'-5' direction using control pore glass (CPG) solid support. After synthesis, the sequences were cleaved from the support by ammonia treatment to cleave the ester functionality that joins support to the 3'-terminus of the oligomrs and deprotects the exocyclic amino protecting groups used during the synthesis. The oligonucleotides were desalted by gel filtration; their purity ascertained by RP-HPLC on a C18 column (0.1N TEAA:CH_3CN) to be more than 95% and was used without further purification in the biophysical studies. The RNA oligonucleotides (Table 2B) were obtained commercially.

Table 4: DNA/RNA oligonucleotides used in present work

Entry	Sequence(5'→3')	Type
1	CGTTTTTTTTGC **20**	control for **ON-1**-**ON-3**
2	TCTCTTTCTT **21**	control for **ON-4**-**ON-8**
3	TCACTAGATG **22**	control for **ON-9**
4	CCTCTTACCTCAGTTACA **23**	control for **ON-10**-**ON-12**
5	GAAGGGCTT TTGAACTCTT **24**	control for **ON-13**,**ON-14**
6	GCAAAAAAAACG **30**	complementary for **ON-1**-**ON-3**
7	AAGAAAGAGA **31**	complementary for **ON-4**-**ON-8**
8	CATCTAGTGA **32**	Complementary or **ON-9**
9	TGTAACTGAGGTAAGAGG **33**	complementary for **ON-10**-**ON-12**
10	AAGAGTTCAAAAGCCCTTC **34**	complementary for **ON-13**,**ON-14**
11	r(GCAAAAAAAACG) **40**	complementary for **ON-1**-**ON-3**
12	r(AAGAAAGAGA) **41**	complementary for **ON-4**-**ON-8**
13	CAUCUAGUGA **42**	Complementary for **ON-9**
14	UGUAACUGAGGUAAGAGG **43**	complementary for **ON-10**-**ON-12**
15	AAGAGUUCAAAAGCCCUUC **44**	complementary for **ON-13**, **ON-14**

3.1.5 UV-melting studies *iso*-TANA modified DNA

The binding affinity of *iso*-TANA-DNA chimeric ONs (Table 3) with complementary DNA and RNA was investigated by measuring the melting temperatures (T_m) of the duplexes and triplexes. The UV-melting experiments were carried out in phosphate buffer (pH = 7.2) containing 100 mM NaCl. The *iso*-TANA-DNA chimeric ONs were individually hybridized with complementary DNA/RNA strand to obtain complexes. The results of UV-melting experiments of *iso*-TANA-

DNA/RNA complexes were compared with the results of TANA-DNA/RNA complexes. The comparison of these results would give the information about the effect of change in stereochemistry on binding preferences of thioacetamido ONs with DNA and RNA.

3.1.5a Polypyrimidine sequences

Table 5: T_m (°C) values of *iso*-TANA ONs: DNA/RNA complexes

Entry	ON Sequences (5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C TANA[4g]	
		DNA- 30	**RNA- 40**	**DNA- 30**	**RNA- 40**
1	CGTTTTTTTTGC **20**	32.5	26.7	40.0	32.0
2	CGTTt$_s$tTTTTGC **ON-1**	21.8	nd	23.7	32.3
3	CGTTt$_s$tTt$_s$tTGC **ON-2**	nd	25.0	nd	50.0
4	CGt$_s$tt$_s$tt$_s$tt$_s$tGC **ON-3**	nd	29.8	nd	47.8

DNA 30 = 5' GCA$_8$CG 3'; **RNA 40** = 5' GCA$_8$CG 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 100 mM NaCl, pH = 7.2), of *iso*-TANA ONs:DNA/ *iso*-TANA ONs:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

The complex of unmodified DNA-**20** with DNA-**30** has higher melting temperature than with RNA-**40** (entry-1, Table 5). Introduction of one *iso*-t$_s$t unit in DNA-**20** showed destabilization with both DNA and RNA. Increasing the number of modified *iso*-t$_s$t blocks did not show consistent increase or decrease in the Tm for complexes either with DNA or RNA. In contrast, the TANA ONs containing multiple t$_s$t units had displayed consistent increase in the Tm of their complexes with RNA.

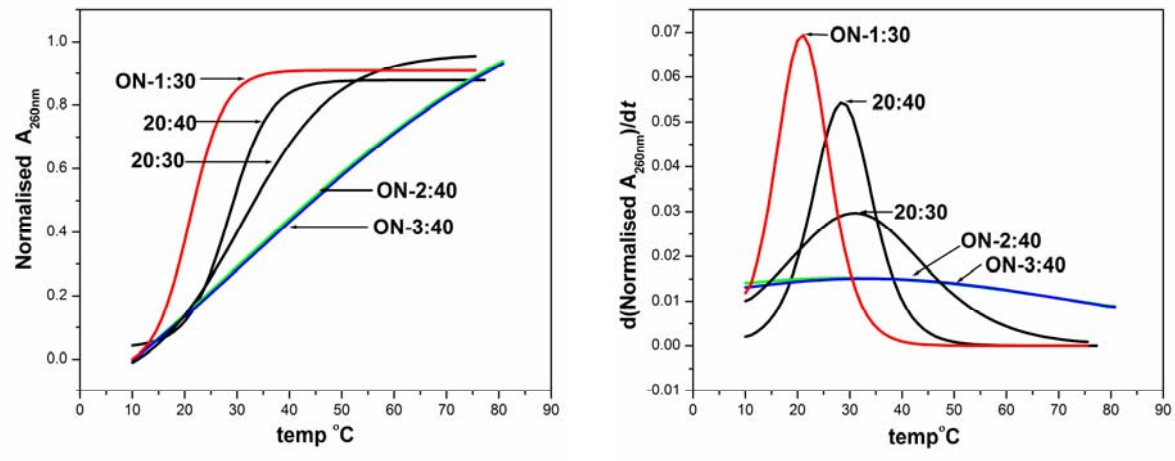


Figure 14: A UV melting curves of ON-**1-3**, DNA-**20** with DNA-**30** and RNA-**40** and **B** corresponding derivative graphs

The UV-Tm studies were then carried out for the mixed cytosine-thymine pyrimidine sequences listed in Table 6. The *iso*-TANA **ON-4-8** did not show any binding with either DNA or RNA whereas the TANA **ON-4-8** showed preferential binding with RNA.

Table 6: T_m (°C) values of *iso*-TANA ONs: DNA/RNA complexes

Entry	ON Sequences (5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C TANA[4g]	
		DNA- 31	RNA- 41	DNA- 31	RNA- 41
1	TCTCTTTCTT **20**	24.0	26.7	23.6	25.4
2	TCTCt_stTCTT **ON-4**	nd	nd	nd	29.6
3	TCTC t_st Tc$_s$tT **ON-5**	nd	nd	nd	33.8
4	Tc$_s$tCTT TCTT **ON-6**	nd	nd	19.5	26.2
5	Tc$_s$t CTTTTc$_s$tT **ON-7**	nd	nd	nd	33.8
6	Tc$_s$tc$_s$tTTc$_s$tT **ON-8**	nd	nd	nd	39.7

DNA 31 = 5' AAGAAAGAGA 3'; **RNA 41** = 5' AAGAAAGAGA 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 100 mM NaCl, pH = 7.2), of *iso*-TANA ONs:DNA/ *iso*-TANA ONs:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

3.1.5b Mixed purine-pyrimidine sequences

The duplexes of control DNA–**23** showed similar binding strength with complementary DNA-**33** and with RNA-**43**. Introduction of one *iso*-t_st unit (**ON-10**) in the DNA-**23** showed preferential binding with DNA-**33** as compared to that with RNA-**43**; although in either case the complexes were less stabilized than the unmodified DNA. The doubly modified **ON-11** further destabilized the complexes with DNA-**33** and RNA-**43**. The t_st modified sequences (**ON-10**, **ON-11**) showed increase in the stability of DNA:RNA complexes with increase in t_st units in the sequence, but this modification destabilized the DNA:DNA duplexes (Table 7).**ON-12** containing a single c_sc unit, in either ribo or *xylo* configration, formed stable complexes with both complementary DNA and RNA. The melting temperatures of these duplexes were as good as DNA-**23**:DNA-**33** and DNA-**23**:RNA-**43** duplexes. **ON-12** with *iso*-c_sc unit showed similar binding efficiency with DNA and RNA like c_sc modified **ON-12**. The ONs with single modified *iso*-t_st/t_st units (**ON-10** with both stereochemistry) destabilized the DNA:DNA and DNA:RNA

duplexes whereas introduction of c_Sc/*iso*-c_Sc unit in DNA-**23** did not destabilize DNA:DNA complexes.

Table 7: T_m (°C) values of *iso*-TANA ONs: DNA/RNA duplexes

Entry	ON Sequences (5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C TANA[4g]	
		DNA-33	RNA-43	DNA-33	RNA-43
1	CCTCTTACCTCAGTTACA **23**	54.6	53.8	54.6	54.7
2	CCTCt_StACCTCATTACA **10**	49.3	46.7	39.6	47.5
3	CCTCt_StACCTCAGt_StACA **11**	42.6	40.9	43.5	52.8
4	CCTCTTAc_ScTCAGTTACA **12**	51.6	51.5	52.2	52.5

DNA 33 = 5' TGTAACTGAGGTAAGAGG 3'; **RNA 43** = 5' UGUAACUGAGGUAAGAGG 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 100 mM NaCl, pH = 7.2), of *iso*-TANA ONs:DNA/*iso*-TANA ONs:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

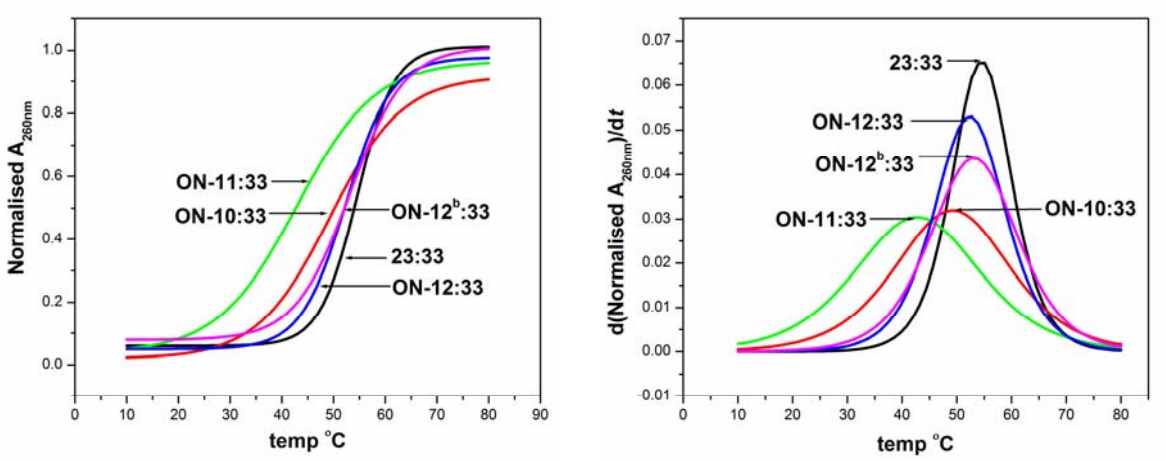


Figure 15: A UV melting curves of ON-**10-12**, DNA-**23** with DNA-**33** and **B** corresponding derivative graphs

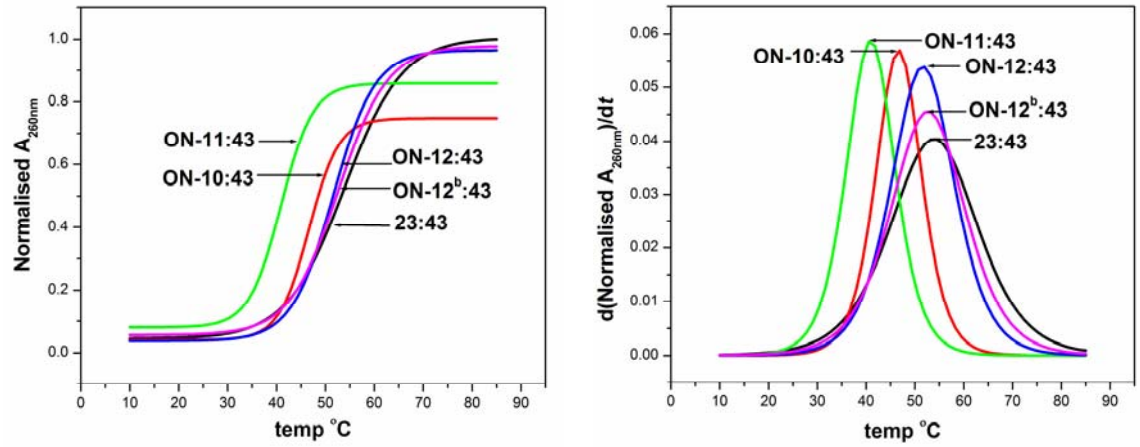


Figure 16: A UV melting curves of ON-**10-12**, DNA-**23** with RNA-**43** and **B** corresponding derivative graphs

We then studied the effect of these modified dimer units in another sequence in which three or more modified units were introduced. The complex of the control DNA-**24** with RNA-**44** showed slightly more melting temperature than with DNA-**34**. The duplexes of **ON-13** with three modified dimer blocks and **ON-14** with four modified dimer blocks (*iso*-t_st and *iso*-c_st) with DNA-**34** and RNA-**44** showed preferential binding with DNA. The complexes with RNA were much more destabilized. Although the *T*m was less when compared to control DNA-**24**: DNA-**34** and DNA-**24**: RNA-**44** complex, the destabilization observed for duplex of **ON-13** containing *iso*-t_st and *iso*-c_st units with DNA -**34** was less than that with **ON-13**:RNA-**44** duplex.

Table 8: T_m (°C) values of *iso*-TANA ONs: DNA/RNA duplexes

Entry	ON Sequences (5'→3')	UV T_m °C (*iso*-TANA)		UV T_m °C TANA	
		DNA-34	RNA-44	DNA-34	RNA-44
1	GAAGGGCTTTTGAACTCTT **24**	53.0	54.2	53.6	54.8
2	GAAGGGc$_s$tTt$_s$tGAACTc$_s$tT **13**	44.7	39.7	42.7	54.5
3	GAAGGGc$_s$tT t$_s$t GAAc$_s$tc$_s$tT **14**	45.8	34.3	41.6	43.6

DNA 34 = 5' AAGAGTTCAAAAGCCCTTC 3'; **RNA 44** = 5' AAGAGUUCAAAAGCCCUUC 3', T_m = melting temperature (measured in the buffer 10 mM sodium phosphate, 100 mM NaCl, pH = 7.2), of *iso*-TANAONs:DNA/*iso*-TANAONs:RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

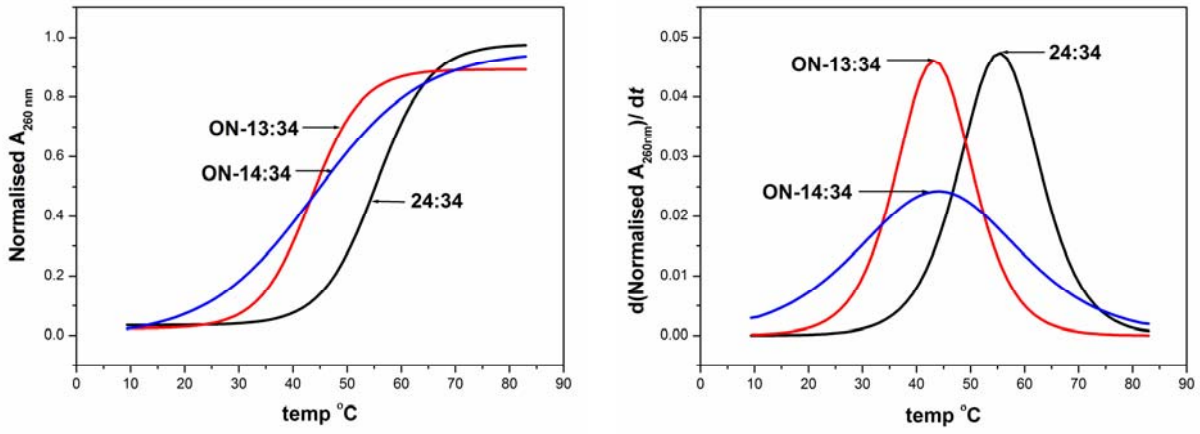


Figure 17: A UV melting curves of ON-**13-14**, DNA-**24** with DNA-**34** and **B** corresponding derivative graphs

Introduction of one more *iso*-c_st unit in **ON-13** stabilized the DNA:DNA duplex with 1°C than **ON-13** whereas destabilized further DNA:RNA duplex with 5-6 °C. **ON-13** with t_st and c_st units showed destabilization with DNA-**34** whereas binding strength to RNA-**44** was as good as control DNA:RNA complex. Unexpectedly, with

additional c_St block in **ON-14,** the duplex of **ON-14**:RNA-**44** was destabilized (ΔT_m= 11 °C) than **ON-13** whereas melting temperature of **ON-14**:DNA-**14** did not change much than **ON-13**:DNA-**34** duplex.

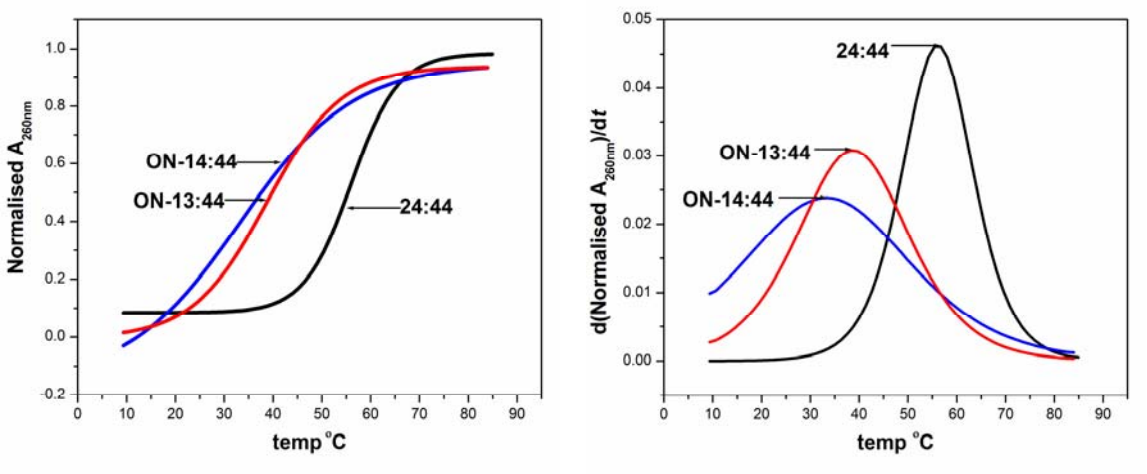


Figure 18: A UV melting curves of ON-**13-14**, DNA-**24** with RNA-**44** and **B** corresponding derivative graphs

A single modified unit of LNA with a locked N-type sugar conformation in an ON is known to effectively stabilize duplex structures with both DNA and RNA.[30] The electronegativity of the 3'-substituents in 3'-*N*-phoramidates was shown to set the sugars in preferred N-type conformation, but unlike LNA, show preferential binding to RNA over DNA.[31] Several other examples in the literature such as 2'-5' DNA[32]/RNA[33] prefer to bind to RNA over DNA and it is not therefore entirely certain which factors differentiate the DNA verses RNA selectivity.[31] Native DNA and RNA prefer to be in S- or N-type sugar conformations giving rise to either B- or A-form structures in equilibrium. In this particular case, however, the ^{1}H NMR studies point out conformational equilibration in either ribo or *xylo* configured 3'-amino or 5'-thioacetamido sugars to be similar to the native 3'-5'phosphate linked DNA. The structural dissimilarity between ribo and *xylo* configured modified TANA and *iso*-TANA units could be the reason for the differential binding selectivities of these isomeric TANA modified DNA. The RNA selectivity of binding seems to be arising from the extended backbone linker in TANA derivative, whereas change in stereochemistry of the sugar affects this distance complementarity and the differently folded *iso*-TANA allows better binding with DNA over RNA. The t_St and c_St dimer blocks were found to be compatible in the DNA backbone to selectively stabilze the ON:RNA complexes.

3.1.6 Summary

- Various epimeric amide linked pyrimidine TANA dimers with inverted stereochemistry at the 3'-amino substituent have been synthesized and their phosphoramidite derivatives were incorporated into DNA oligomers.

- The NMR studies showed that the acetylation of 3'-amino groups in nucleosides shifted the conformational equilibrium for sugar rings from N type to S type in both the stereisomers. The sugar ring conformations are not largely affected in the *iso*-$t_S t$, *iso*-$c_S t$ dimers as compared to corresponding epimeric TANA $t_S t$, $c_S t$ dimers.

- The base-stacking interactions for *iso*-TANA dimers were found to be different than TANA dimers in the CD analysis. CD spectra of *iso*-$t_S t$ dimer resemble the observed CD spectra of phosphate linked *xylo* dithymidine.

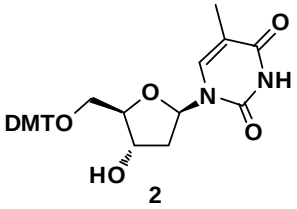
- The binding preferences of *iso*-TANA modified ONs were found to be better with DNA than RNA in contrast to TANA modified ONs. The effect of different base stacking interactions reflected in the CD pattern could be the cause of this difference in binding preferences.

3.1.7 Experimental

General: The chemicals used were of laboratory or analytical grade. All solvents used were purified according to the literature procedures. Reactions were monitored by TLC. Usual reaction work up involved sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulphate and evaporation of the solvent under vacuum. IR spectra were recorded on an infrared Fourier Transform spectrophotometer using nujol, chloroform or neat. Column chromatographic separations were performed using silica gel 60-120 mesh (Merck) and 200- 400 mesh (Merck) and using the solvent systems EtOAc/petroleum ether and MeOH/DCM. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). TLCs were run in either petroleum ether with appropriate quantity of ethyl acetate or dichloromethane with an appropriate quantity of methanol for most of the compounds. TLCs were visualized with UV light and iodine spray and/or by perchloric acid treatment then heating. ^{1}H and ^{13}C spectra were obtained using Bruker AC-200, AC-400 and AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^{1}H. The optical rotation values were measured on Bellingham-Stanley Ltd, ADP220 polarimeter. Mass spectra were obtained either by LCMS techniques. Oligomers were characterized by RP HPLC (Varian /Waters) C18 column and MALDI-TOF mass spectrometry. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems) MALDI-TOF instrument and the matrix used for analysis was THAP (2, 4, 6-trihydroxyacetophenone). DNA oligomers were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard β-cyanoethyl phosphoramidite chemistry.

Synthesis of 5'-*O*-dimethoxytrirylthymidine (2)

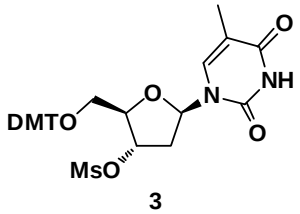
A mixture of thymidine (6 g, 24.7 mmol), dimethoxytrityl chloride (10 g, 29.7 mmol), DMAP (cat) in anhydrous 40 mL pyridine was stirred at RT for 4 hrs. Pyridine was removed under reduced pressure and residue was redissolved in 400 mL ethyl acetate. The organic layer was washed with 10% $NaHCO_3$ solution with (2x 30 mL) followed by water (50 mL) and brine. The organic layer was dried over anhydrous $NaSO_4$ and concentrated to dryness. Column chromatography purification of crude product using 3% methanol-dichloromethane gave pure product **2**, 11.72 g (87%).



2

Synthesis of 3'-*O*-mesyl -5'-*O*-dimethoxytritylthymidine (3)

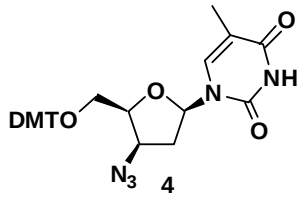
A mixture of **2** (11 g, 20.2 mmol), Et_3N (7.0 mL, 50.5 mmol) in 40 mL anhydrous dichloromethane was cooled to 0˚C and mesyl chloride (2.35 mL, 30.3 mmol) in 10 mL dry dichloromethane) was added dropwise. After 10 min reaction mixture was taken in 100 ml dichloromethane, organic layer was washed with 10% $NaHCO_3$ solution with (2x 30 mL) followed by water (2x50 mL). The organic layer was dried over anhydrous sodium sulphate and concentrated to dryness. Column purification using 2% methanol-dichloromethane gave pure product **3**, 11.3 g, 85%.



3

Synthesis of 3'-deoxy*xylo*-3'-azido -5'-*O*- dimethoxytritylthymidine (4)

Mesyl derivative **3** (11 g, 17.67 mmol) and sodium azide (11.48 g, 176 mmol) in 20 mL anhydrous DMF was stirred at 60˚C for 24 hrs. DMF was removed under reduced pressure and residue was redissolved in ethyl acetate. The organic layer was washed with water (2x 50 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to dryness. Column purification using gradient of ethyl acetate in petroleum ether gave pure **4**, 4.52 g, 45%.



4

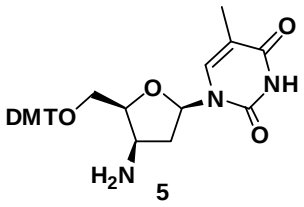
IR, $\upsilon(cm^{-1})$ ($CHCl_3$) 3128, 3019, 2110, 1724, 1670.

^1H NMR (200 MHz, $CDCl_3$) δ 1.84 (s, 3H), 2.12-2.19 (m, 1H), 2.64-2.78 (m, 1H), 3.32-3.40 (m, 1H), 3.56-3.64 (m, 1H), 4.08-4.14 (m, 1H), 4.24-4.29 (m, 1H) 6.12-6.17 (dd, 1H, J =2.9 Hz, J =7.9 Hz), 6.82-6.87 (d, 4H), 7.3-7.48 (m, 10H).

^{13}CNMR (50 MHz, $CDCl_3$) δ 12.5, 38.6, 55.1, 61.1, 61.8, 81.7, 84.1,
86.7, 110.6, 113.0, 126.8, 127.8, 129.9, 135.5, 144.4, 150.5, 158.5, 164.0.
Mass calcd. for $C_{31}H_{31}N_5O_6$ 569.22 obsd. 592.68 (M+ Na^+).

Synthesis of 3'-deoxy*xylo*-3'-amino- 5'-*O*- dimethoxytritylthymidine (5)

Compound **4** (2.0g, 2.6 mmol) was dissolved in 10 mL ethyl acetate and 0.20 g (10%) Pd-C catalyst was added. Then reaction mixture was subjected to catalytic hydrogenation at 40 psi of hydrogen pressure for 4 hrs. The catalyst was filtered and washed with (10 mL x 2) ethyl acetate. Organic layer was concentrated to give and **The** crude product was purified with column chromatography with DCM –MeOH gave compound **5** (1.7 g, 86%).



5

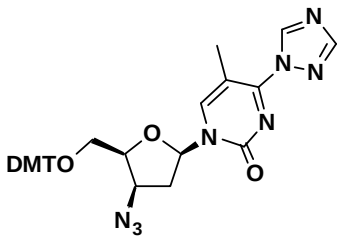
IR, υ(cm^{-1}) ($CHCl_3$) 3398, 3019, 2924,2853,1701,1686.

^{1}H NMR (200 MHz, $CDCl_3$) δ 1.81 (s, 3H), 1.91-2.07 (m, 2H), 2.63-2.78 (m, 1H), 2.92-3.03 (m, 1H) 3.39-3.47 (m, 1H), 3.79 (s, 6H), 4.19-4.27 (m, 1H), 6.31-6.36 (dd,1H, *J*= 2.9 Hz,8. 2 Hz), 6.79-6.84 (m, 4H), 7.24-7.44 (m, 10H).

Mass calcd. for $C_{31}H_{33}N_3O_6$ 543.23 obsd. 566.14 (M+Na^+)

Synthesis of 3'-deoxy*xylo*-3'-azido-5'-*O*-dimethoxytrityl-*C*^{4}-(1,2,4-triazol-1-yl)-pyrimidine 2'deoxynucleoside (6)

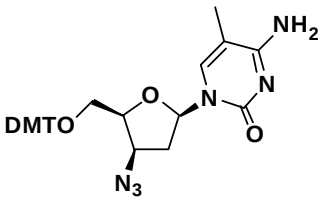
Et_3N(6.1 mL, 43.9 mmol) was added drop wise to a stirred mixture of 1, 2, 4-triazole (3.33 g, 48.3 mmol) and phosphoryl chloride (1.02 mL, 10.97 mmol) in 50 mL CH_3CN at 0˚C. The solution of **4** (2.5g, 4.39 mmol) in 15 mL dry CH_3CN was then added drop wise. The reaction mixture was stirred at room temp for 2.5 hrs. The solvent was evaporated and residue was redissolved in 50 mL ethyl acetate. The organic layers was washed with water (2x50 mL) and brine (2x20 mL), dried over Na_2SO_4 and evaporated to dryness to give triazolide derivative **6** (2.62 g, 88%) as yellow foam, which was used immediately for next reaction.



6

Synthesis of 2',3'-dideoxy*xylo*-3'-azido-5'-*O*-dimethoxytrityl-5-methylcytidine (7)

Concentrated aqueous ammonia (10 mL) was added to the solution of triazolide derivative (2.6 g, 4.08 mmol) in 40 mL 1,4-dioxane. After complete conversion of triazolide derivative to 5-methyl- 2'-deoxy cytidine derivative in 1hr, solvent was removed under reduced pressure and the residue redissolved in dichloromethane and



7

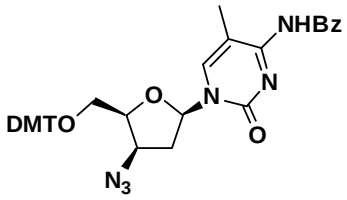
purified by silica gel column chromatography (0-7% methanol in dichloromethane) to afford product (2.12 g, 90%) as white foam.

IR, υ(cm^{-1}) ($CHCl_3$) 3192, 2984, 2827, 2114, 1687.6

^{1}H NMR (200 MHz, $CDCl_3$) δ 1.84 (s, 3H), 2.19-2.27 (m, 1H), 2.66-2.82 (m, 1H), 3.31-3.38 (m, 1H), 3.56-3.63 (m, 1H), 3.80 (S, 6H), 4.19-4.21 (m, 1H), 6.12-6.17 (dd, 1H, *J*= 2.4 Hz, *J*= 7.5 Hz), 6.82-6.87 (d, 4H), 7.23-7.51 (m, 10H).

Mass calcd. for $C_{31}H_{32}N_6O_5$ 568.24 obsd. 569.73 (M+ H^+), 591.73 (M+Na^+)

Synthesis of *N*^{4}-benzoyl-2', 3'-dideoxy *xylo*-3'-azido-5'-O-dimethoxytrityl-5-methylcytidine (8)



8

Benzoyl chloride (0.51 mL, 4.43 mmol) was added drop wise slowly to a solution **7** (2.1 g, 3.69 mmol) in 30 mL pyridine at 0˚C. The mixture was then stirred at room temperature for 4 hrs.. The reaction was quenched with addition of 5 mL 2N NH_3 solution. Reaction mixture was further stirred for 30mins. The solvent was evaporated and residue redissolved in 50 mL ethyl acetate, washed the organic layer with 10% $NaHCO_3$ solution (3x30 mL), water (2x50 mL) and brine (20 mL). The organic layer was then dried over anhydrous sodium sulphate and evaporated to dryness. Compound was purified on column chromatography (5-20% ethyl acetate in petroleum ether) to afford **8**, 1.82 g (75%) as white foam.

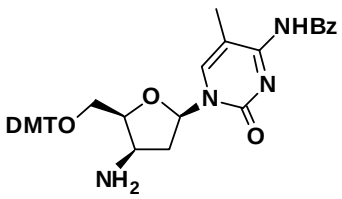
IR, υ(cm^{-1}) ($CHCl_3$) 3192, 3024, 2110, 1715, 1705, 1696, 1685

^{1}H NMR (200 MHz, $CDCl_3$) δ 2.02 (s.3H), 2.17-2.28 (m,1H), 2.66-2.77 (m,1H) 3.34-3.42 (m,1H),3.59-3.67 (m,1H), 3.81 (s,6H), 4.07-4.29 (m,2H), 6.13-6.18 (dd,1H, *J*= 2.5, 7.7 Hz), 6.84-6.88 (d,4H), 7.26-7.52 (m,13H), 8.30-8.33 (m,2H)

^{13}C NMR (50 MHz, $CDCl_3$) δ 13.8, 33.9, 55.1, 61.1, 61.9, 82.2, 84.9, 86.8, 111.3, 113.1, 126.9, 127.9, 128.08, 130.0, 132.3, 135. 5, 137.1, 144.4, 147.9, 158.6, 159.7, 179.4.

Mass Calc. for $C_{38}H_{36}N_6O_6$ 672.26 observed 673.72 (M+ H^+), 695.71 (M+Na^+)

Synthesis of N^4-benzoyl-2', 3'-dideoxy-*xylo*-3'-amino-5'-*O*-dimethoxytrityl-5-ethylcytidine (9)



9

Compound **8** (1.8 g, 2.67 mmol) was dissolved in pyridine (4 mL) and Et_3N 0.5 mL was added to this solution. H_2S gas was bubbled into the solution at 0˚C for 15minutes. Then solution was stirred at room temperature for additional 1 hr. and solvent was removed under reduced pressure . The residue was purified by column chromatography (2-5% methanol-dichloromethane) to afford **9**, 1.4 g (78%).

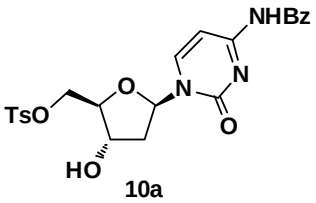
IR, υ(cm^{-1}) ($CHCl_3$) 3189, 3024, 2810, 1718, 1707, 1696, 1685

^{1}H NMR (200 MHz, $CDCl_3$) δ 1.83 (s,3H), 1.88-2.07 (m,2H), 2.6-2.72 (m,1H), 3.07-3.18 (q,1H), 3.39-3.59 (m,2H), 3.8 (s,3H), 4.13-4.18 (m,1H), 6.16-6.22 (dd,1H, *J*= 7.07 Hz), 6.84-6.88 (d,4H), 7.28-7.52 (m,13H), 8.04 (m, 1H), 8.28—8.32 (d, 2H).

^{13}C NMR (50 MHz, $CDCl_3$) δ 13.1, 41.5, 51.1, 61.7, 81.8, 84.7, 86.7, 110.9, 113.2, 127.0, 128.0, 129.9, 132.2, 135.2, 137.0, 138.8, 144.2, 148.0, 158.6, 159.9, 179.3.

Mass calcd. for $C_{38}H_{38}N_4O_6$ 646.27 obsd. 647.87 ($M+H^+$), 669.82 ($M+Na^+$)

Synthesis of 5'-*O*-Tosyl-N^4-Benzoyl 2',3'-dideoxycytidine (10a)



10a

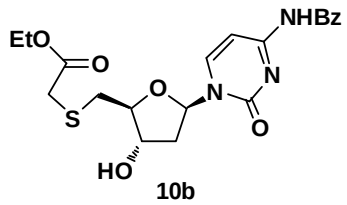
N^4-Benzoyl 2', 3'-dideoxy Cytidine (5 g, 15.09 mmol) was suspended in 50 mL dry pyridine and cooled to 0˚C. Tosyl chloride (4.3 g, 22.6 mmol) diluted with 10 mL pyridine was added to above mix over period of 4hrs at 0 °C. Then reaction mixture was stirred for 1 hr at RT and pyridine was removed under reduced pressure; residue was redissolved in 250 mL of ethyl acetate. Organic layer was washed with 10% $NaHCO_3$ solution and concentrated to give crude product. Compound was purified by column chromatography by 1-3 % methanol – dichloromethane. Yield = (4.3 g 60%)

^{1}H NMR (200 MHz, DMSO-d_6) δ 2.10-2.35 (m.2H), 2.45 (s, 3H), 3.99-4.07(m, 1H), 4.16-4.22(m, 2H), 4.25-4.39 (m, 2H), 6.11-6.18(t, 1H, *J*= 6.32 Hz), 7.31(d, 1H, *J*=7.32), 7.49-7.70(m, 5H) 7.83-8.03(m, 5H)

^{13}C NMR (50 MHz, DMSO-d_6) δ 13.9, 21.0, 33.2, 34.1, 60.7, 69.7, 69.9, 72.4, 83.9, 86.4, 96.2, 127.6, 128.4, 130.2,132.0, 132.7, 133.1, 144.6, 145.2, 154.0, 162.9, 167.5, 170.0.

Mass calcd. for $C_{23}H_{23}N_3O_7S$ 485.1257 obsd. 508.3680 ($M+Na^+$)

Synthesis of ethyl [(2'-3'-dideoxy 5'-*S*-cytidinyl) mercapto] acetic acid (10b)



10b

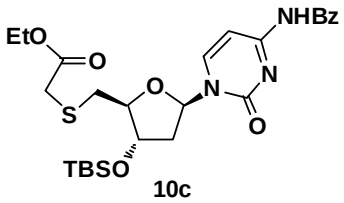
ethyl mercapto acetic acid (1.2 mL, 9.6 mmol) was added dropwise to suspension of NaH (60% suspension in hexane) (0.38 g, 0.96 mmol) and in dry DMF to generate mercapto nucleophile. Tosyl derivative **10a** (3.9 g, 8.03 mmol) in 5 mL dry DMF was added to it at 0˚C. Then reaction was stirred for 1 hr at RT, DMF was removed at in vacuo, residue was redissolved in 200 mL ethyl acetate, organic layer was washed with 50 mL water. Organic layer was concentrated under reduced pressure; compound was purified with column chromatography with 1-3% methanol –dichloromethane. To white solid (Yield 3.0 g = 80 %)

^{1}H NMR (200 MHz, DMSO-d_6) δ 1.18-1.25 (t, 3H), 2.11- 2.40 (m, 2H), 2.91-2.96 (m, 2H), 4.01-4.18 (m, 4H) 6.15-6.22(t, 1H, *J*= 6.44 Hz), 7.39-7.89(m, 4H), 8.01-8.05(m, 2H),8.18-8.22(d, 1H, *J*= 7.45 Hz)

^{13}C NMR (50 MHz, DMSO-d_6) δ 14.2, 33.5, 34.4, 61.0, , 72.6, 86.2, 86.5, 96.6, 128.8, 132.9, 133.3, 145.1, 154.4, 163.2, 167.7, 170.2.

Mass calcd. for $C_{20}H_{23}N_3O_6S$ 433.1308 obsd. 456.3636 ($M+Na^+$)

Synthesis of ethyl *S*-5'-(3'-*O*-TBDMS-2', 3'-dideoxy cytidinyl)-mercaptoacetic acid (10c)



10c

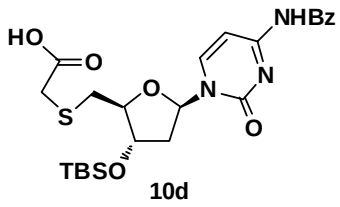
Compound **10b** (2.7 g, 6.46 mmol) and imidazole (0.96 g, 14.22 mmol) was dissolved in dry DMF. TBDMS-Cl (1.07 g, 7.11 mmol) was added at 0˚C. Reaction was stirred at RT for 6 hrs, DMF was removed in vacuo, residue was dissolved in ethyl acetate, and organic layer was washed with water. Organic layer was concentrated invacuo, crude product purified with 20 % ethyl acetate –Petroleum ether to give product (2.9 g, Yield 90%)

^{1}H NMR (200 MHz, $CDCl_3$) δ 0.09 (s, 6H), 0.89 (s, 9H), 1.27-1.34 (t, 3H, *J* = 7.20Hz), 2.05-2.20 (m, 1H), 2.47-2.66 (m, 1H), 2.98-3.36 (m, 2H), 3.24-3.43 (q, 2H, *J* = 23.62 Hz), 4.10-4.27 (m, 4H), 6.19-6.25(dd,1H, *J* = 5.81, 6.06 Hz), 7.48-7.63(m, 4H), 7.89-7.93 (m, 2H), 8.2-8.24(d, 1H, *J*=7.45 Hz)

^{13}C NMR (50 MHz, $CDCl_3$) δ 4.7, 14.0, 17.7, 25.5, 33.8, 41.3, 61.3, 73.2, 86.1, 86.6, 96.5, 127.5, 128.7, 132.9, 144.0, 154.5, 162.2, 166.7, 169.9.

Mass calcd. for $C_{26}H_{37}N_3O_6SSi$ 547.21 obsd. 548.20($M+ H^+$), 570.15($M+ Na^+$)

Synthesis of (5'-*S*-3'-*O*-TBDMS-2', 3'-dideoxy Cytidinyl) mercaptoacetic acid (10d)



10d

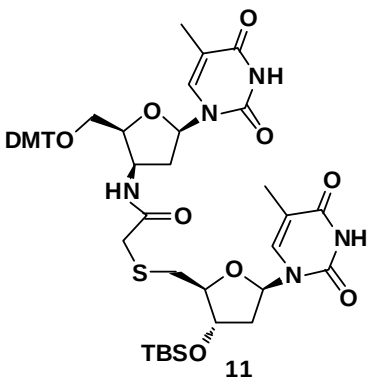
Compound 10c (2.0 g, 3.6 mmol) was suspended in 9:1 THF: water 10 mL, 1 mL 1N NaOH solution was added at 0 ˚C. Reaction was stirred at 0 ˚C up to starting material get finished. Then pH of reaction mixture was adjusted to 7 with cation exchange resin, Then resin was filtered and washed with methanol. Reaction mix was concentrated to dryness at vacuo. Product was desiccated to give foam (1.3 g, 75%)

^{1}H NMR (200MHz, CDCl3) δ 0.00 (s, 6H), 0.78 (s, 9H), , 2.08-2.28 (m, 2H), 2.4 (m, 2H), 2.63-2.91(m, 2H), 3.88-3.394 (m, 1H), 4.26-4.32 (m, 1H), 6.01-6.07(dd,1H, J= 6.31, 6.57Hz), 7.25-7.52(m, 5H), 7.89-7.99(m,2H), 8.11-8.25 (d, 1H, *J*= 7.57 Hz)

^{13}C NMR (50 MHz, DMSO-d_6) δ- 4.6, 17.9, 25.9, 34.1, 34.4, 63.11, 74.3, 86.2, 86.6, 96.7, 128.7, 130.49, 133.0, 133.4, 145.1, 154.5 163.2, 167.7, 171.7

Mass calcd. for $C_{24}H_{33}N_3O_6SSi$ 519.18 obsd. 542.12 ($M+Na^+$)

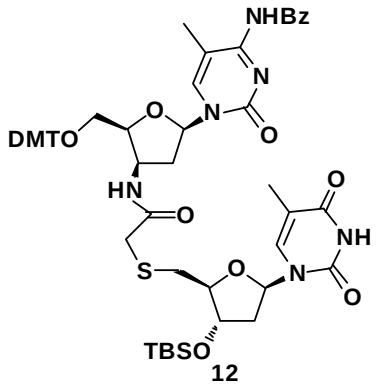
Synthesis of 5'-*O*-DMT-*iso*-t$_s$t-3'-*O*-TBDMS dimer (11)



11

S-5'-(3'-*O*-TBDMS-thymidinyl)-mercapto- acetic acid **1d** (1g, 2.32mmol) TBTU (0.89 g, 2.784mmol) and HOBt (0.15g 1.16 mmol) was taken inmixture of dry CH_3CN:dryDMF(5:1) (5 mL), diisopropylethyl amine (1.00 mL, 5.8 mmol) was added to this solution and stirred for 15 min. Then amino compound **5** (1.25 g, 2.32mmol) dissolved in 4 mL CH_3CN was added into reaction mixture and further stirred for 2 hours. Then reaction mixture was concentrated and product was extracted with ethyl acetate (50 mL) and washed with 5% $NaHCO_3$ solution (2x20 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to get crude product. This was purified by column chromatography with gradient dichloromethane-methanol yield product, 1.34 g (61%).

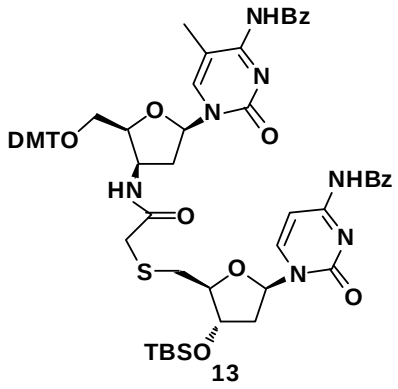
Synthesis of 5'-*O*-DMT-*iso*-c$_s$t-3'-*O*-TBDMS dimer (12)



12

S-5'-(3'-*O*-TBDMS-thymidinyl)-mercapto- acetic acid **1d** (1.3 g, 3.01mmol) TBTU (1.12g, 3.62 mmol) and HOBt (0.2 g 1.50 mmol) was taken in mixture of dry CH_3CN:dryDMF(5:1) (5 mL), diisopropylethyl amine (1.2 mL, 7.52 mmol) was added to this solution and stirred for 15min. Then amino compound **9** (1.74 g, 2.70 mmol) dissolved in 4 mL CH_3CN was added into reaction mixture and further stirred for 2 hours. The reaction mixture was concentrated and product was extracted with (70 mL) ethyl acetate and washed with 5% $NaHCO_3$ solution (2x20 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to get crude product. This was purified by column chromatography with gradient dichloromethane-methanol yield product, 1.68 g (58%).

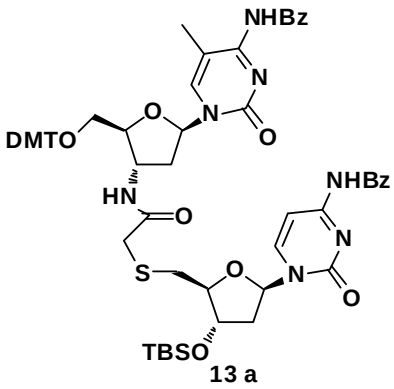
Synthesis of 5'-*O*-DMT- *iso*-c$_s$c-3'-*O*-TBDMS dimer (13)



13

S-5'-(3'-*O*-TBDMS-cytidinyl)-mercapto- acetic acid **10d** (1.5 g, 2.88 mmol) TBTU (1.1 g, 3.45 mmol) and HOBt (0.19 g 1.44 mmol) was taken in mixture dry CH_3CN:dryDMF (5:1) (5 mL), diisopropylethyl amine (1.2 mL, 7.2 mmol) was added to this solution and stirred for 15min. Then amino compound **9** (1.67 g, 2.59 mmol) dissolved in 4 mL CH_3CN was added into reaction mixture and further stirred for 2 hours. The reaction mixture was concentrated and product was extracted with (70 mL) ethyl acetate and washed with 5% $NaHCO_3$ solution (2x20 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to get crude product. This was purified by column chromatography with gradient dichloromethane-methanol yield product, 1.62 g (54 %).

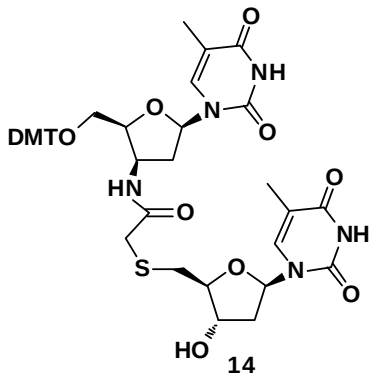
Synthesis of 5'-*O*-DMT-c_sc-3'-*O*-TBDMS dimer (13a)



13 a

S-5'-(3'-*O*-TBDMS-cytidinyl)-mercapto- acetic acid **10d** (1.5 g, 2.88 mmol) TBTU (1.1 g, 3.45 mmol) and HOBt (0.19 g 1.44 mmol) was taken in dry CH_3CN:dryDMF (5:1) (5 mL), diisopropylethyl amine (1.2 mL, 7.2 mmol) was added to this solution and stirred for 15min. Then amino compound **9a** (1.67 g, 2.59 mmol) dissolved in 4 mL CH_3CN was added into reaction mixture and further stirred for 2 hours. The reaction mixture was concentrated and product was extracted with (70 mL) ethyl acetate and washed with 5% $NaHCO_3$ solution (2x20 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to get crude product. Compound **13a** was purified by column chromatography with gradient dichloromethane-methanol yield product, 1.68 g (55%).

Synthesis of 5'-*O*-DMT-*iso*-t_st-3'-OH dimer (14)



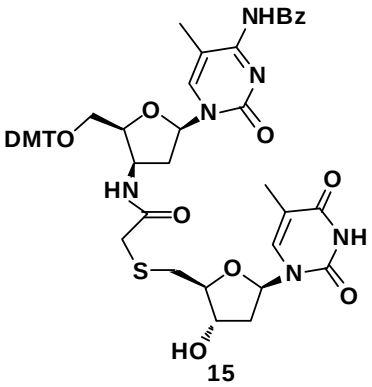
14

Dimer **11** (1.3g, 1.36 mmol) was taken in 10 mL anhydrous THF and 1N TBAF in THF (2.04 mL, 2.04 mmol) was added. Reaction was stirred at RT for 1 h, then THF was removed under reduced pressure and residue redissolved in dichloromethane (50mL). The organic layer was washed with water (2X20mL) followed by brine and dried over anhydrous sodium sulphate and evaporated to dryness. After purification by column chromatography with gradient dichloromethane-methanol gave the pure product **14**, (1.0 g, 87 %).

^{1}H NMR (200MHz, $CDCl_3$) δ 1.50(s,3H), 1.90(s,3H), 2.18-2.42(m,4H),2.67-2.84(m,4H), 3.78(s,6H), 3.92-4.00(m,1H), 4.21-4.34(m,2H), 4.73-5.05(m,1H), 5.98-6.04(dd,1H, *J* = 6.82, 6.07Hz), 6.13-6.20 (t,1H, *J*=6.56Hz), 6.81-6.85 (m,4H), 7.25-7.64(m,12H).

Mass calcd. for ($C_{43}H_{47}N_5O_{11}S$) 841.29 observed 865.66 ($M+1+Na^+$).

Synthesis of 5'-*O*-DMT-*iso*-c_st-3'-OH dimer(15)

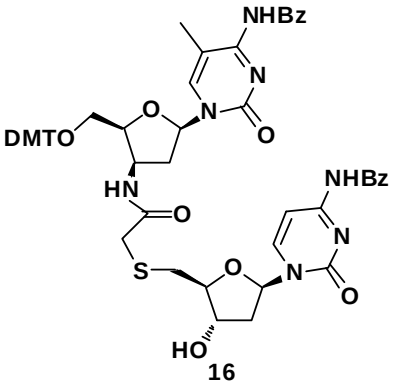


Dimer **12** (1.6 g, 1.51 mmol) was taken in 10 mL anhydrous THF and 1N TBAF in THF (2.2 mL, 2.22 mmol) was added. Reaction was stirred at RT for 1 h, then THF was removed under reduced pressure and residue redissolved in dichloromethane (60mL). The organic layer was washed with water (2X20mL) followed by brine and was dried over anhydrous sodium sulphate and evaporated to dryness. After purification by column chromatography with gradient dichloromethane-methanol gave the pure product **15** (1.14 g, 80%).

^{1}H NMR (200MHz, $CDCl_3$) δ 1.71(s,3H), 1.90 (s,3H), 2.16-2.40(m,3H), 2.73-2.82(m,3H), 3.07-3.24(m,4H), 3.8(s,6H),3.83-3.93(m,1H) 4.24-436 (m,2H), 4.76-4.83 (m,1H), 6.04-6.11(dd, 1H, *J*= 6.84, 7.07), 6.12-6.18(dd,1H, *J*= 6.44, 6.70Hz), 6.84-6.89(m,4H), 7.23-7.53(m,13H), 7.87(s,1H), 8.28-8.30(m,2H).

Mass calcd. for $C_{50}H_{52}N_6O_{11}S$ 945.04 observed 945.71 ($M+ H^+$), 967.68 ($M+Na^+$).

Synthesis of 5'-*O*-DMT-*iso*-c_sc-3'-OH dimer (16):

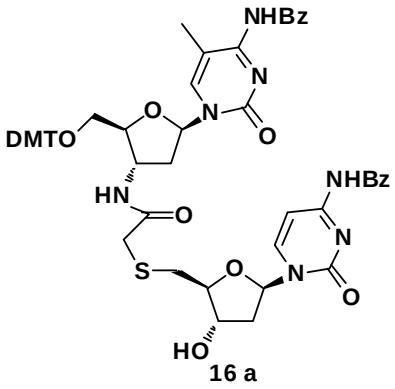


Dimer **13** (1.6 g, 1.39 mmol) was taken in 10 mL anhydrous THF and 1N TBAF in THF (2.1 mL, 2.09 mmol) was added. Reaction was stirred at RT for 1 h, then THF was evaporated under reduced pressure and residue redissolved in dichloromethane (60mL). The organic layer was washed with water (2X20mL) followed by brine and was dried over anhydrous sodium sulphate and evaporated to dryness. After purification by column chromatography with gradient dichloromethane-methanol gave the pure product **16** (1.12 g, 78%).

^{1}H NMR (200 MHz, $CDCl_3$) δ 1.65 (s, 3H), 2.02-2.37(m, 4H), 2.59-2.69(m, 2H), 2.76-2.81(m, 2H), 4.23-4.26(m, 2H0, 4.72-5.05(m, 2H), 6.11-6.20(m, 2H), 6.84-6.88(m, 4H), 7.29-7.5 (m, 19H), 7.88-7.92(m, 4H, 8.07-8.23(m, 1H)

Mass Calcd. for $C_{56}H_{55}N_7O_{11}S$ 1033.36 observed1057.15 ($M+Na^+$)

Synthesis of 5'-*O*-DMT-c_sc-3'-OH dimer(16a):



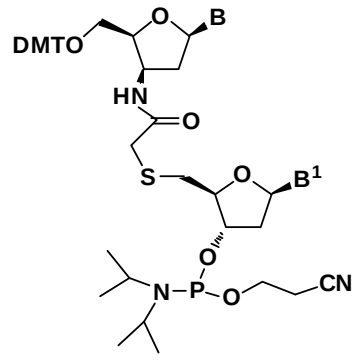
16 a

Dimer **13a** (1.6 g, 1.39 mmol) was taken in 10 mL anhydrous THF and 1N TBAF in THF (2.1 mL, 2.09 mmol) was added. Reaction was stirred at RT for 1 h, then THF was evaporated under reduced pressure and residue redissolved in dichloromethane (60mL). The organic layer was washed with water (2X20mL) followed by brine and was dried over anhydrous sodium sulphate and evaporated to dryness. After purification by column chromatography with gradient dichloromethane-methanol gave the pure product **16a** (1.15 g, 80%).

^{1}H NMR (200MHz, $CDCl_3$) δ 1.5 (s, 3H), 2.02-2.09(m, 3H), 2.36-2.42(m, 2H), 2.91-2.98 (m, 2H), 3.29 (m, 1h), 3.40 (m, 1H), 3.47 9m, 2H), 3.76 (s, 6H), 4.95-5.04 (m, 1h), 5.72-5.84 (m, 1H), 6.2-6.18 (dd, 1H, *J*= 6.31, 6.19), 6.52-6.59 (dd, 1H, *J*= 6.82, 6.32) 6.81-6.86 (m, 4H), 7.30-7.34(m, 6H), 7.42-7.57 (m, 10H), 7.84-7.93 (m, 3H), 8.03-8.07 (m, 1H), 8.24-8.26 (m, 3H)

Mass Calcd. for $C_{56}H_{55}N_7O_{11}S$ 1033.3680 observed1057.2532 ($M+Na^+$)

Synthesis of phospharamidite derivatives of thioacetamido linked dimers.



17 B= B^1= Thymine
18 B =N^4-benzyol 5-mehtyl cytosine
B^1= Thymine
19/19a B =N^4-benzyol 5-mehtyl cytosine
B^1= N^4-benzyol cytosine

A common procedure was followed for synthesis of phospharamidite derivatives of dimers. Compound (**14/15/16/16a**) was dissolved in dry dichloromethane (10 mL) followed by the addition of diisoprpylethylamine (2.5 molar equivalents of dimer) Then 2-cynoethyl-*N*,*N*,-diisopropyl-chloro phosphine (1.2 molar equivalents of dimer) was added to the solution at 0 °C and reaction mixture was stirred at room temperature for 3 hours. The contents were then diluted with DCM and washed with 5% $NaHCO_3$ solution. The organic phase was dried over anhydrous sodium sulphate and concentrated to foam. The residue was redissolved in DCM and precipitated with n-hexane to obtain corresponding phospharamidite derivatives.

Synthesis of 5'-*O*-(4, 4'-dimethoxy) trityl-*iso*-t_St-3'-O-(2-cyanoethyl-*N*,*N*-diisopropylphoshoramidate)-dimer (17)

Compound **17** was synthesized following the procedure as described for (compound **17-19a**). Yield (0.7 g, 78%)

^{13}P NMR (200MHz, $CDCl_3$) δ148.67, 148.73.

Synthesis of 5'-*O*-(4, 4'-dimethoxy) trityl-*iso*-c_St-3'-O-(2-cyanoethyl-*N*,*N*-diisopropylphoshoramidate)-dimer (18)

Compound **18** was synthesized following the procedure as described for (compound **17-19a**). Yield (0.84g, 70%).

^{13}P NMR (200MHz, $CDCl_3$) δ148.78, 148.90

Synthesis of 5'-*O*-(4, 4'-dimethoxy) trityl-*iso*-c_Sc-3'-O-(2-cyanoethyl-*N*,*N*-diisopropylphoshoramidate)-dimer (19)

Compound **19** was synthesized following the procedure as described for (compound **17-19a**). Yield (0.84g, 70%).

^{13}P NMR (200 MHz, $CDCl_3$) δ148.70, 149.19

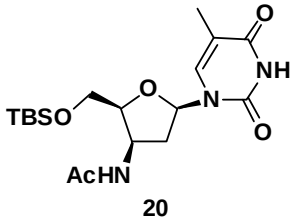
Synthesis of 5'-*O*-(4, 4'-dimethoxy) trityl-c_Sc-3'-O-(2-cyanoethyl-*N*,*N*-diisopropylphoshoramidate)-dimer (19a)

Compound **19a** was synthesized following the procedure as described for (compound **17-19a**). Yield (0.82g, 73 %)

^{13}P NMR (200 MHz, $CDCl_3$) δ148.72, 148.93

Synthesis of 3'-deoxy*xylo*-3'-*N*-acetyl-5'-*O*TBS thymidine (20)

3'-deoxy*xylo* -3'- amino-5'-*O*TBS Thymidine (0.5 g, 1.2mmol) was taken in 5 mL pyridine. Acetic anhydride (0.16 mL, 1.63 mmol) was added at 0 °C and reaction was stirred at RT for 2 hrs. Pyridine was removed under vacuum and product was extracted in ethyl acetate and purified by column chromatography (gradient of MeOH in DCM) to give **20** (Yield = 0.5 g, 90 %)



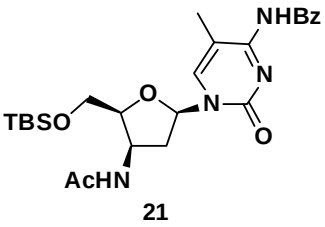
20

^{1}H NMR (200MHz, $CDCl_3$) δ 0.16 (s, 6H), 0.96 (s, 9H), 1.90 (m, 1H), 1.93 (s, 3H), 1.99 (s, 3H), 2.70-2.85 9m, 1H), 3.79-4.05 (m, 2H), 4.10-4.84 (m, 2H), 6.00-5.03 (dd, 1H, J = 7.93 Hz, 6.33 Hz), 6.88-6.9 (m, 1H), 7.41 (s, 1H).

^{13}C NMR (50 MHz, $CDCl_3$) δ -5.48, 12.3, 18.0, 23.1, 25.7, 38.1, 49.9, 62.8, 78.3, 83.6, 110.7, 135.6, 150.4, 134.1, 170.0

Synthesis of 2', 3'-dideoxy*xylo*-3'-*N*-acetyl-5'-*O*TBS 5-methyl cytidine (21)

2', 3'-dideoxy*xylo*-3'-amino-5'-*O*TBS 5-methylcytidine (0.5 g, 1.0 mmol) was taken in 5 mL pyridine. Acetic anhydride (0.14 mL, 1.41 mmol) was added at 0 °C and reaction



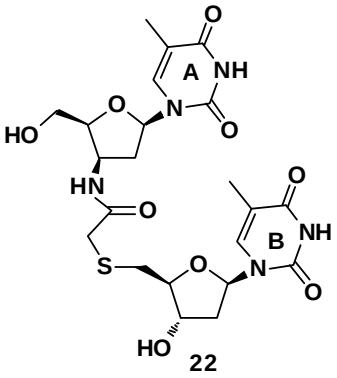
21

was stirred at RT for 2 hrs. Pyridine was removed under vacuum and product was extracted in ethyl acetate and purified by column chromatography (gradient of MeOH in DCM) to give **21** (Yield = 0.48 g, 88 %)

^{1}H NMR (200MHz, $CDCl_3$) δ 0.18 (s, 6H), 0.98 (s, 9H), 1.97-2.15 (m, 1H), 2.00 (s, 3H), 2.12(s, 3H), 2.72-2.77(m, 1H), 3.81-4.08 (m, 2H), 4.14-4.16 (m, 1H), 4.79-4.85 (m, 1H), 6.01-6.03 (dd, 1H, *J* = 6.88 Hz, 6.6 Hz), 6.86-6.07 (s, 1H), 7.43-7.46 (Ar, 4H), 8.31-8.32 (Ar, 2H)

^{13}C NMR (50 MHz, $CDCl_3$) δ-5.4, 13.5, 18.0, 23.2, 25.8, 29.5, 38.4, 49.9, 62.8, 78.7, 84.3, 111.5, 128.0, 129.7, 132.4, 136.7, 147.8, 159.5, 169.8, 179.4

Synthesis of 5'-OH-*iso*-t$_s$t-3'-OH dimer (22)



22

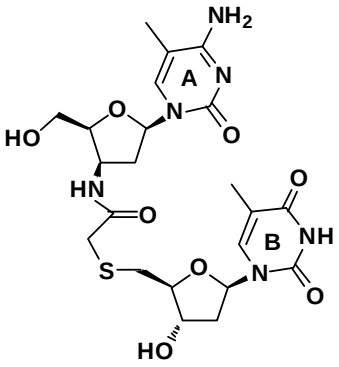
5'-*O*-DMT-*iso*-t$_s$t-3'-OH dimer (200mg, 0.2 mmol) was dissolved in 5 mL 2% dichloroacetic acid and 50 µL triisopropyl silane was added as scavenger. Reaction was stirred for ½ hr. Reaction mixture was concentrated and dietyl ether wash has been given to remove trityl impurities. Yield. (89 mg, 70%)

^{1}H NMR (200MHz, D_2O) δ 1.75(s, 6H, T^A, T^B, C*H*$_3$), 2.04-2.08(m, 1H, T^A *H*2'), 2.21-2.31(m, 2H, T^B *H*2', *H*2'), 2.66-2.73 (m, 2H, T^A *H*2'', T^B*H*5'), 2.86-2.89 (m, 1H, T^B, *H5''*), 3.20-3.28 (dd, 2H, SC*H*$_2$CO, *J*= 15.65 Hz), 3.63-3.75 (m, 2H, T^A, *H5', H5''*), 3.86-3.90(m, 1H, T^B, *H4'*), 4.08-4.11(m, 1H, T^A, *H4'*), 4.23-4.27 (m, 1H, T^B, *H3'*), 4.49-4.53(m, 1H, T^A, *H3'*), 5.86-5.89 (m, 1H, T^A, *H1'J* = 5.87 Hz, 7.34 Hz), 6.06-6.09 (m, 1H, T^B, *H1' J* = 6.85 Hz, 6.84 Hz), 7.35, 7.59 (s, 2H T^A, T^B, *H*6)

LCMS calcd. 562.15 obsd. 562.21 HRMS calcd. for ($C_{22}H_{29}N_5O_9NaS$) 562.1583 obsd. 562.1572

Synthesis of 5'-OH--*iso*-c$_s$t-3'-OH dimer (23)

5'-*O*-DMT-*iso*-c$_s$t-3'-OH dimer (200mg, 0.2 mmol) was dissolved in 5 mL 2% dichloroacetic acid and 50 µL triisopropyl silane was added as scavenger. Reaction was stirred for ½ hr. Reaction mixture was concentrated and dietyl ether wash has been given to remove trityl impurities. The product was subjected for ammonia treatment for removal of N^4benzyol protection of 5-methyl cytosine base. Reacton was concentrated



23

and crude product has been given washes of diethyl ether for reomoval of impurities. Yield. (70 mg, 60%)

^{1}H NMR (200MHz, D_2O) δ δ 1.84, 1.92 (s, 6H, C^A, T^B, C*H*$_3$), 2.10-2.14(m, 1H, C^A *H*2'), 2.31-2.40(m, 2H, T^B *H*2', *H*2'), 2.77-2.83 (m, 2H, C^A *H*2'', T^B*H*5'), 2.91-2.95 (m, 1H, T^B, *H5''*), 3.31-3.38 (m, 2H, SC*H*$_2$CO,), 3.75-3.88 (m, 2H, C^A, *H5', H5''*), 3.95-3.99(m, 1H, T^B, *H4'*), 4.19-4.22(m, 1H, C^A, *H4'*), 4.23-4.35 (m, 1H, T^B, *H3'*), 4.56-4.60(m, 1H, C^A, *H3'*), 5.91-5.94 (m, 1H, C^A, *H1'J*= 5.8 Hz, 6.8 Hz), 6.14-6.18 (m, 1H, T^B, *H1' J*= 6.7 Hz, 6.7 Hz), 7.44, 7.68 (s, 2H C^A, T^B, *H*6)

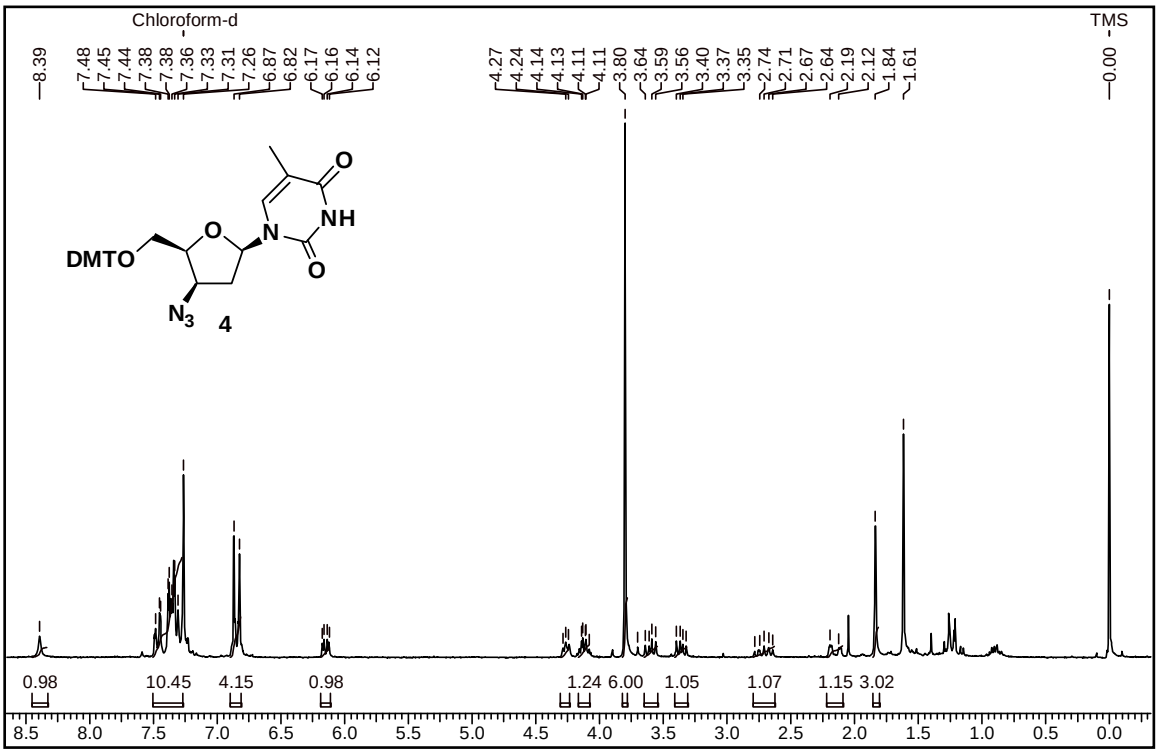
LCMS calcd. 561.17 obsd. 561.20, 585.57 HRMS calcd. for ($C_{22}H_{30}N_6O_8NaS$) calcd. 561.1743 obsd. 562.1747, ($C_{22}H_{30}N_6O_8KS$) calcd. 577.1692 obsed. 577.1679

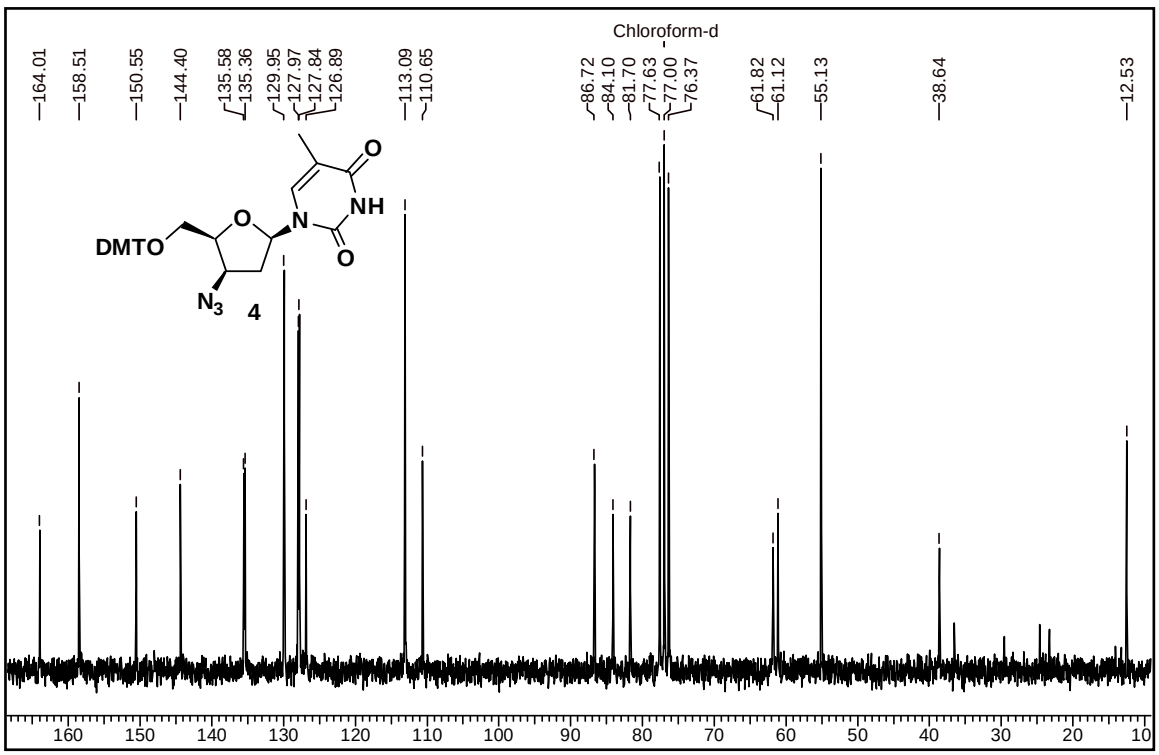
3.1.8 Appendix:

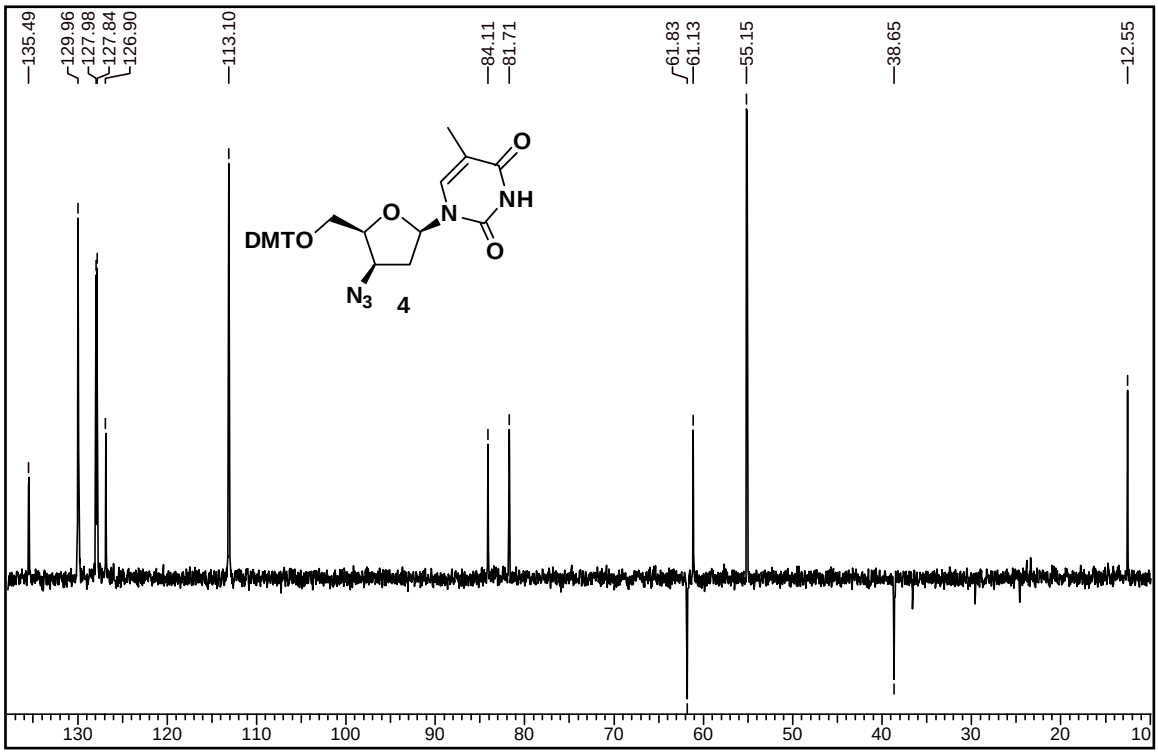
^{1}H and ^{13}C spectra for compounds **1-16a-**, **10a-10d** [31] page no.176-195

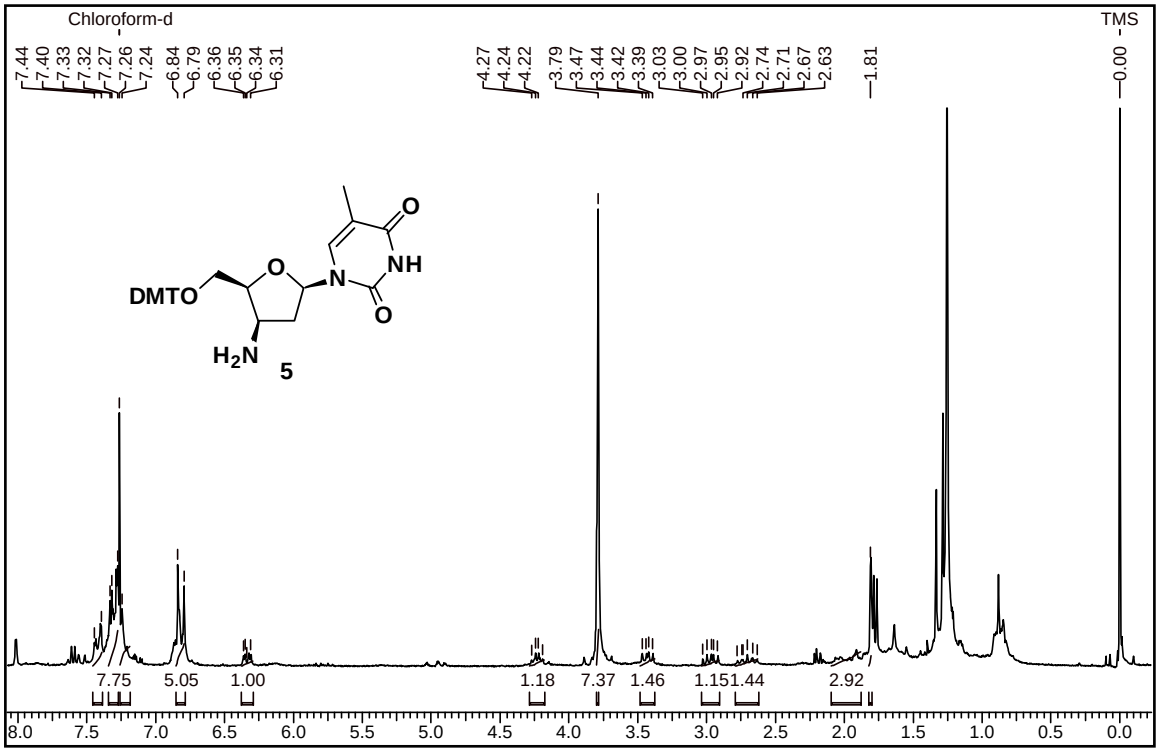
P spectra for compound **17**, **18**, **19**

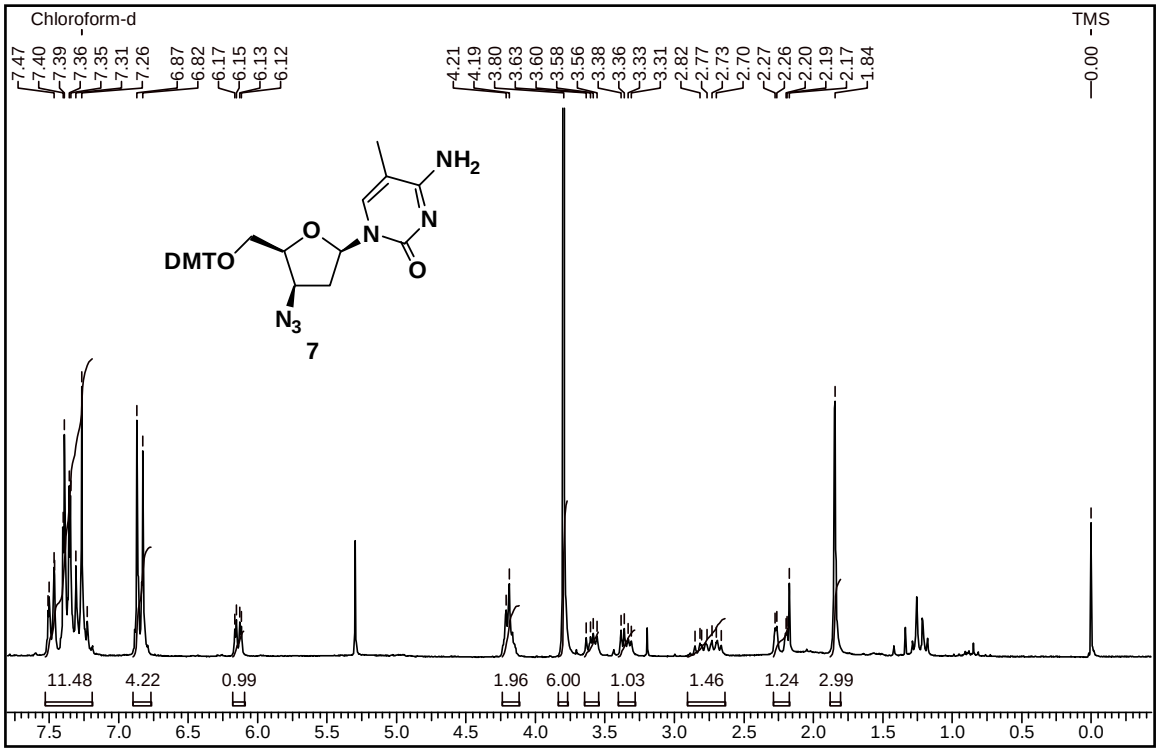
HPLC and MALDI-TOF spectra for **ON-1-ON-14** page no. 196-211

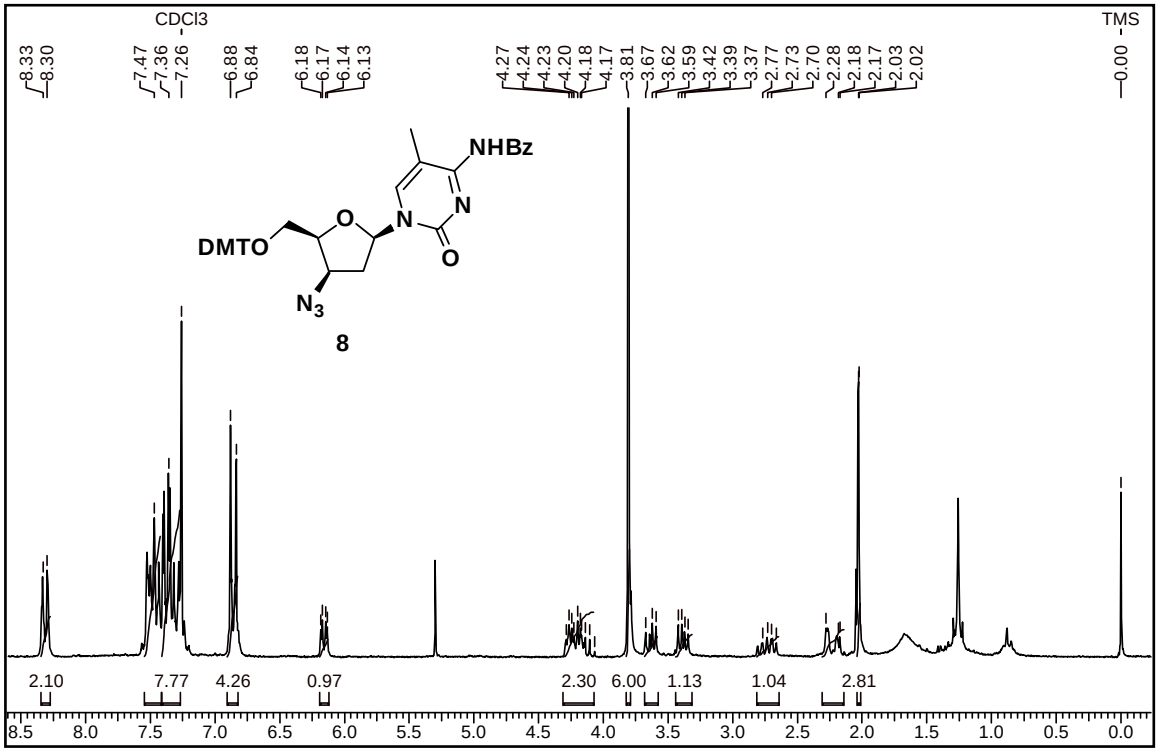

Chloroform-d
TMS
8.39
7.48
7.45
7.44
7.38
7.38
7.36
7.33
7.31
7.26
6.87
6.82
6.17
6.16
6.14
6.12
4.27
4.24
4.14
4.13
4.11
4.11
3.80
3.64
3.59
3.56
3.40
3.37
3.35
2.74
2.71
2.67
2.64
2.19
2.12
1.84
1.61
0.00
O
NH
N
O
O
DMTO
N3
4
0.98
10.45
4.15
0.98
1.24
6.00
1.05
1.07
1.15
3.02
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

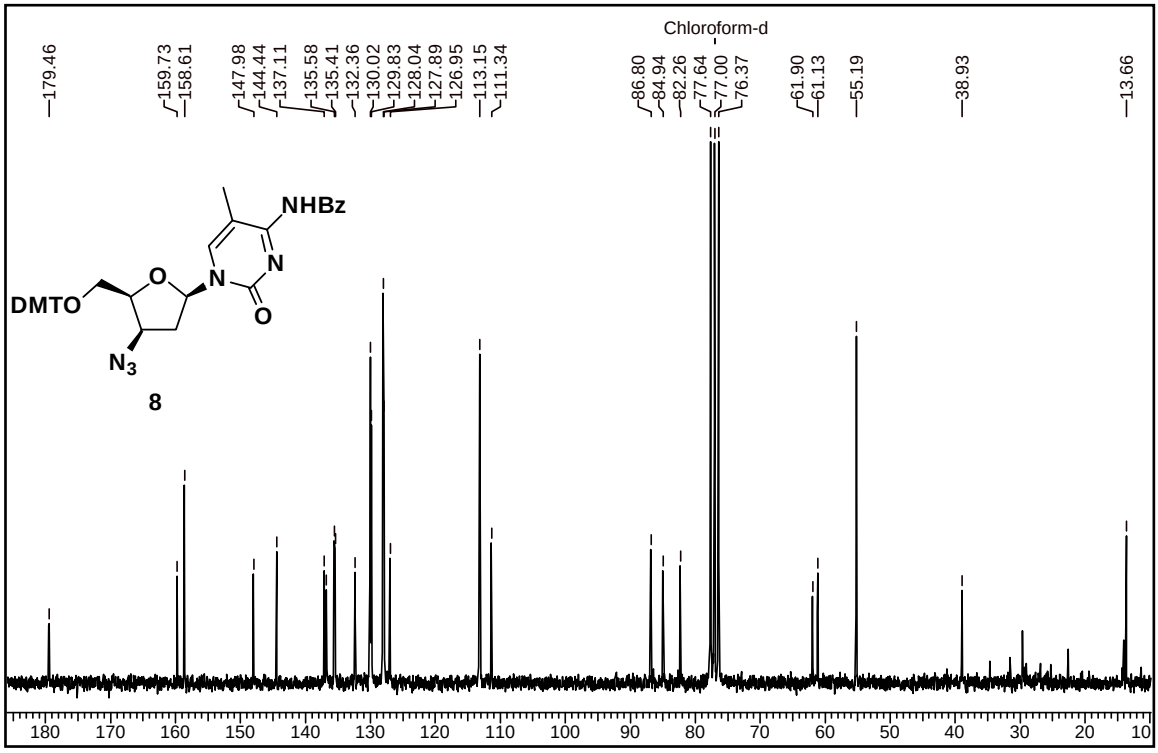

Chloroform-d
164.01
158.51
150.55
144.40
135.58
135.36
129.95
127.97
127.84
126.89
113.09
110.65
86.72
84.10
81.70
77.63
77.00
76.37
61.82
61.12
55.13
38.64
12.53
O
NH
N
O
O
DMTO
N3
4
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

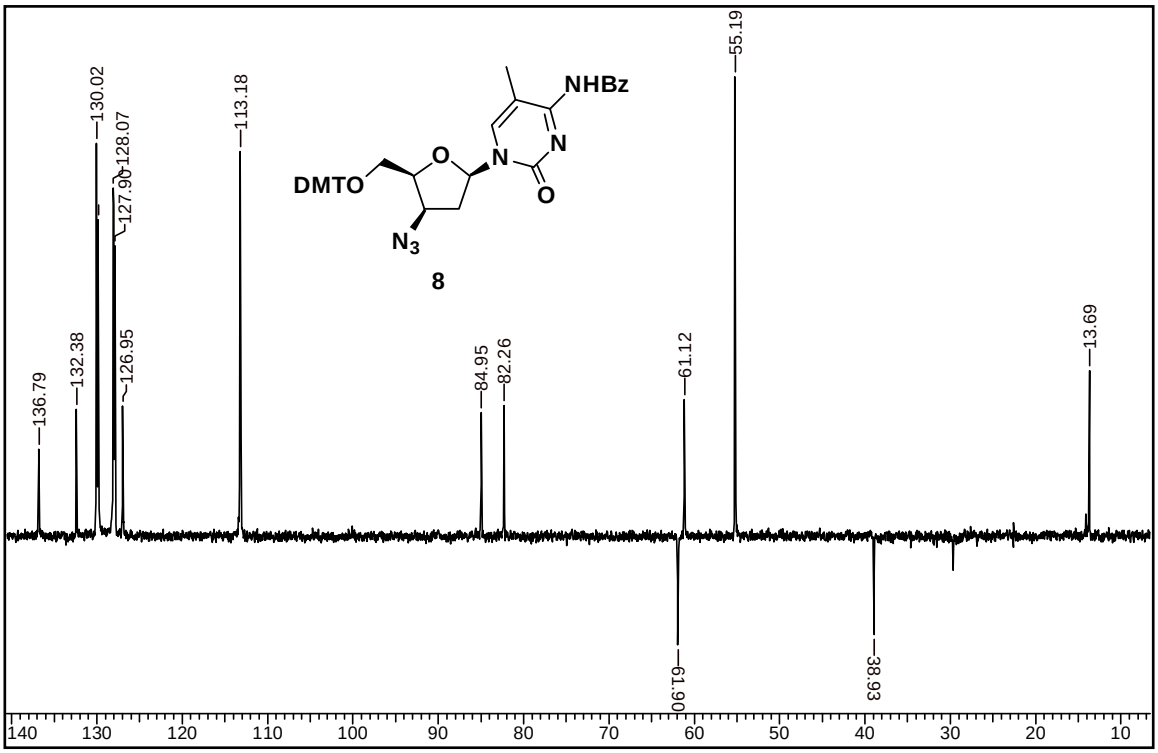

135.49
129.96
127.98
127.84
126.90
113.10
84.11
81.71
61.83
61.13
55.15
38.65
12.55
DMTO
O
NH
N
O
O
N3
4
130
120
110
100
90
80
70
60
50
40
30
20
10

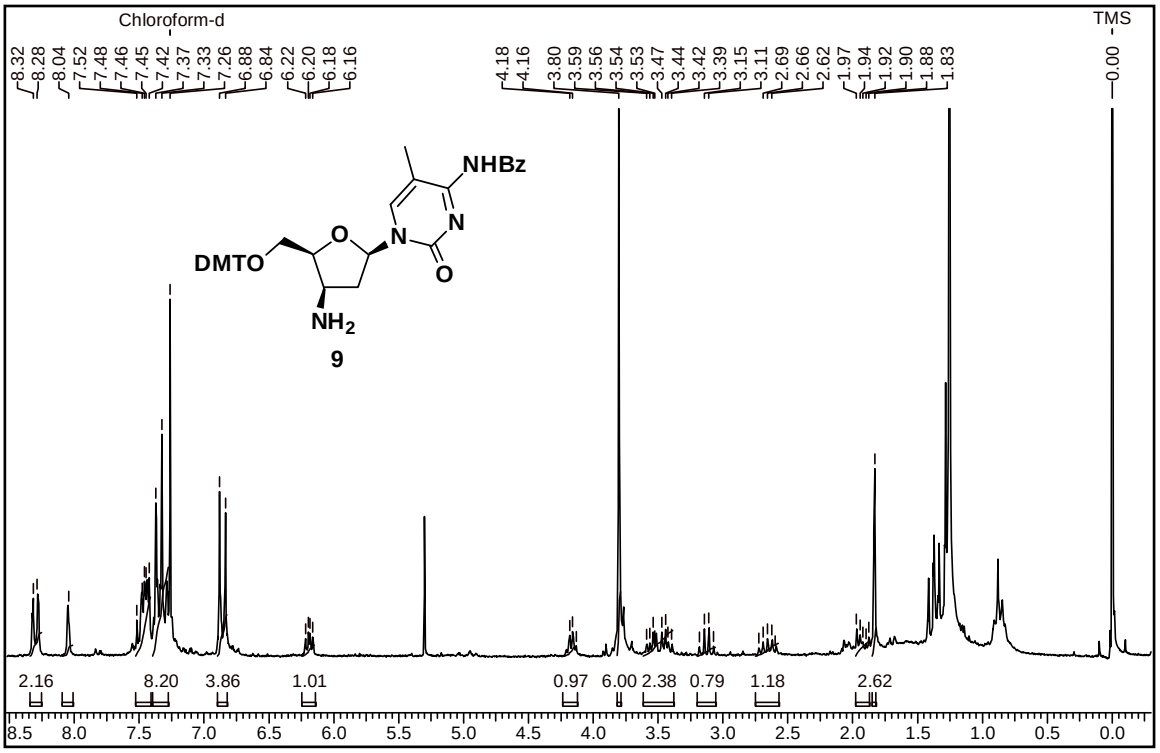

Chloroform-d
TMS
7.44
7.40
7.33
7.32
7.27
7.26
7.24
6.84
6.79
6.36
6.35
6.34
6.31
4.27
4.24
4.22
3.79
3.47
3.44
3.42
3.39
3.03
3.00
2.97
2.95
2.92
2.74
2.71
2.67
2.63
1.81
0.00
DMTO
O
NH
N
O
O
H2N
5
7.75
5.05
1.00
1.18
7.37
1.46
1.15
1.44
2.92
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

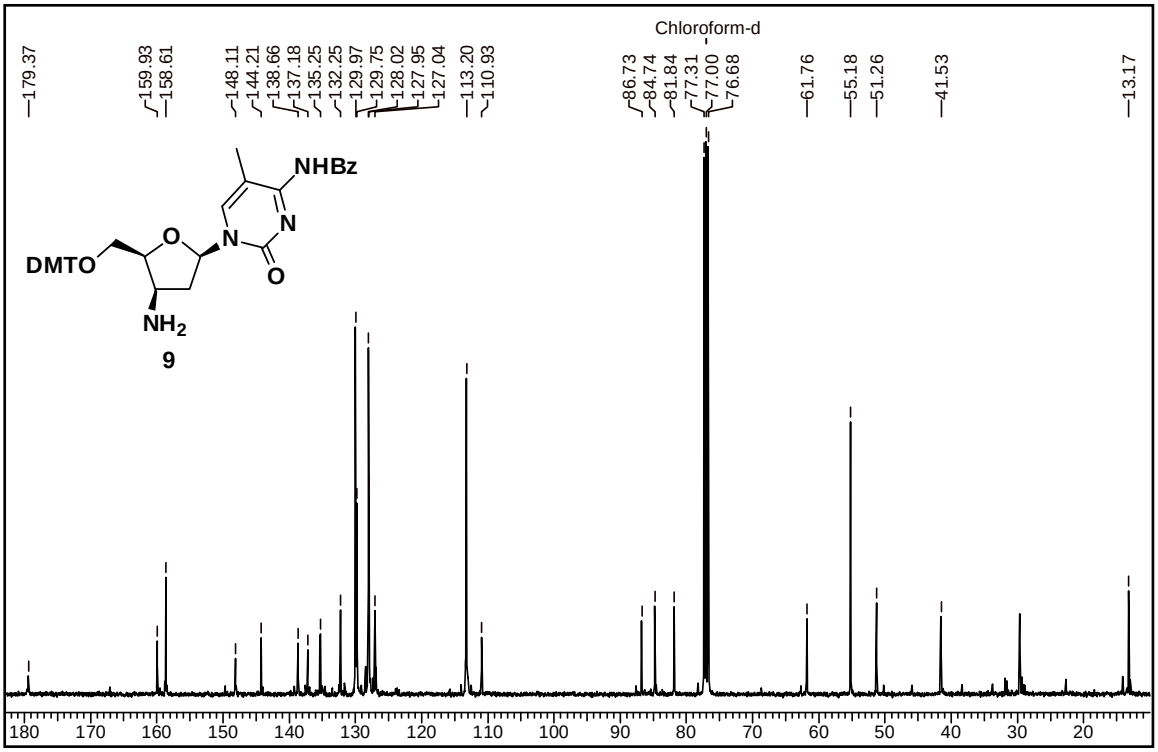

Chloroform-d
TMS
7.47
7.40
7.39
7.36
7.35
7.31
7.26
6.87
6.82
6.17
6.15
6.13
6.12
4.21
4.19
3.80
3.63
3.60
3.58
3.56
3.38
3.36
3.33
3.31
2.82
2.77
2.73
2.70
2.27
2.26
2.20
2.19
2.17
1.84
0.00
NH2
N
O
N
O
DMTO
N3
7
11.48
4.22
0.99
1.96
6.00
1.03
1.46
1.24
2.99
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

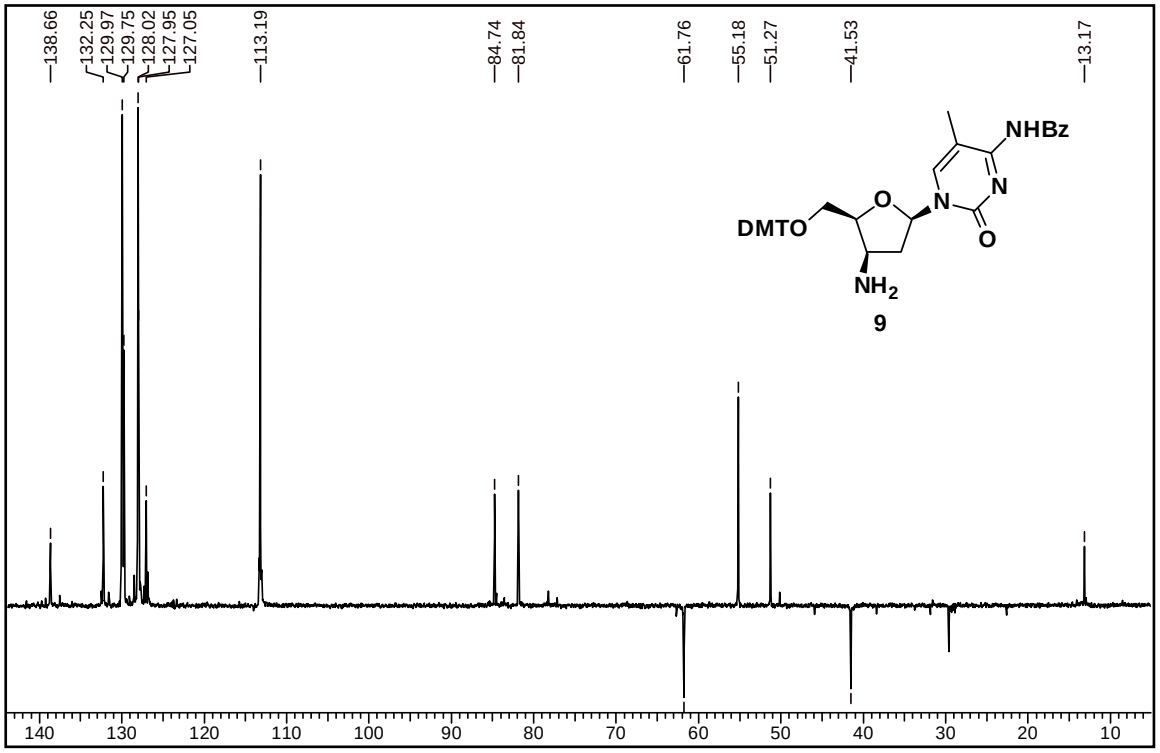

CDCl3
TMS
8.33
8.30
7.47
7.36
7.26
6.88
6.84
6.18
6.17
6.14
6.13
4.27
4.24
4.23
4.20
4.18
4.17
3.81
3.67
3.62
3.59
3.42
3.39
3.37
2.77
2.73
2.70
2.28
2.18
2.17
2.03
2.02
0.00
NHBz
N
O
N
O
DMTO
N3
8
2.10
7.77
4.26
0.97
2.30
6.00
1.13
1.04
2.81
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

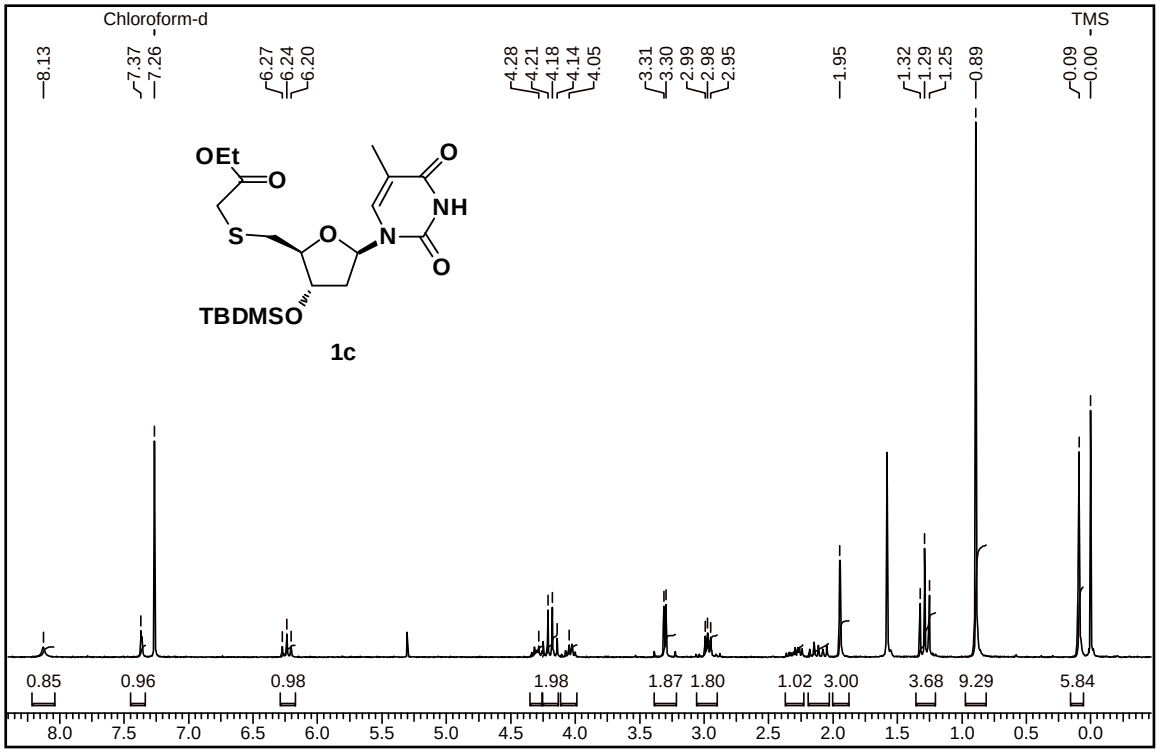

Chloroform-d
179.46
159.73
158.61
147.98
144.44
137.11
135.58
135.41
132.36
130.02
129.83
128.04
127.89
126.95
113.15
111.34
86.80
84.94
82.26
77.64
77.00
76.37
61.90
61.13
55.19
38.93
13.66
NHBz
DMTO
O
N
N
O
N3
8
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

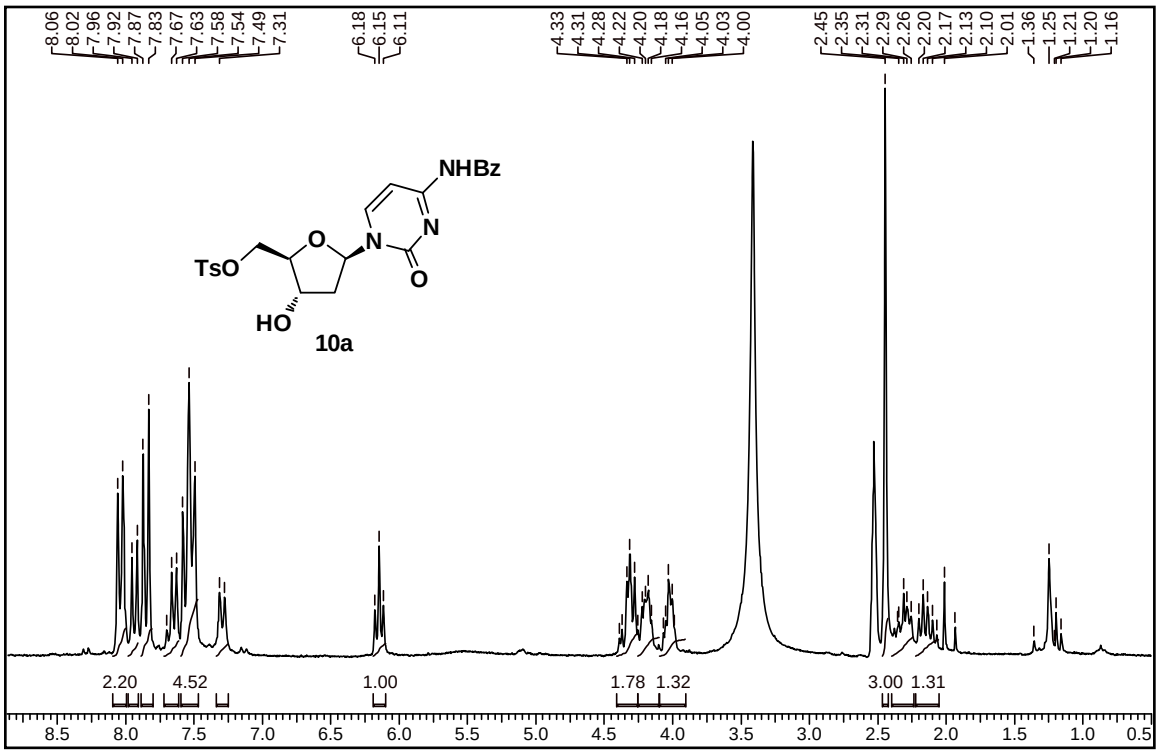

55.19
130.02
113.18
128.07
127.90
NHBz
DMTO
O
N
N
O
N3
8
136.79
132.38
126.95
84.95
82.26
61.12
13.69
61.90
38.93
140
130
120
110
100
90
80
70
60
50
40
30
20
10

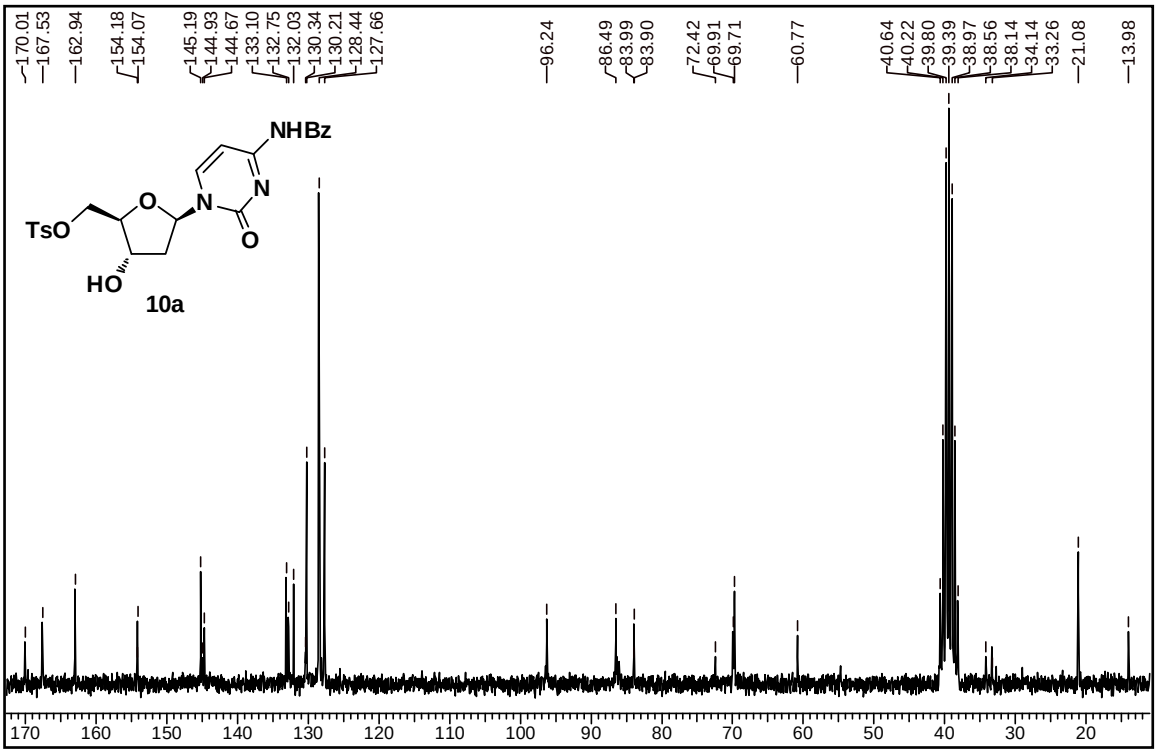

Chloroform-d
TMS
8.32
8.28
8.04
7.52
7.48
7.46
7.45
7.42
7.37
7.33
7.26
6.88
6.84
6.22
6.20
6.18
6.16
4.18
4.16
3.80
3.59
3.56
3.54
3.53
3.47
3.44
3.42
3.39
3.15
3.11
2.69
2.66
2.62
1.97
1.94
1.92
1.90
1.88
1.83
0.00
NHBz
N
N
O
DMTO
O
NH2
9
2.16
8.20
3.86
1.01
0.97
6.00
2.38
0.79
1.18
2.62
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

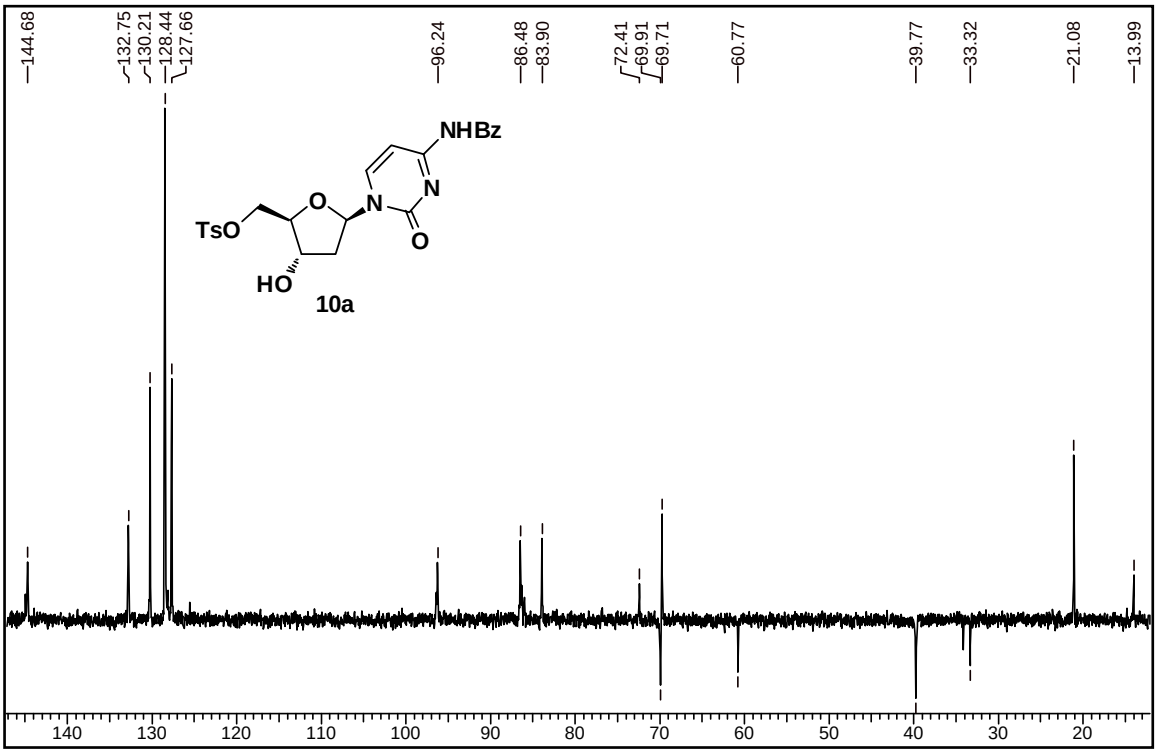

Chloroform-d
179.37
159.93
158.61
148.11
144.21
138.66
137.18
135.25
132.25
129.97
129.75
128.02
127.95
127.04
113.20
110.93
86.73
84.74
81.84
77.31
77.00
76.68
61.76
55.18
51.26
41.53
13.17
NHBz
N
N
O
DMTO
O
NH2
9
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20

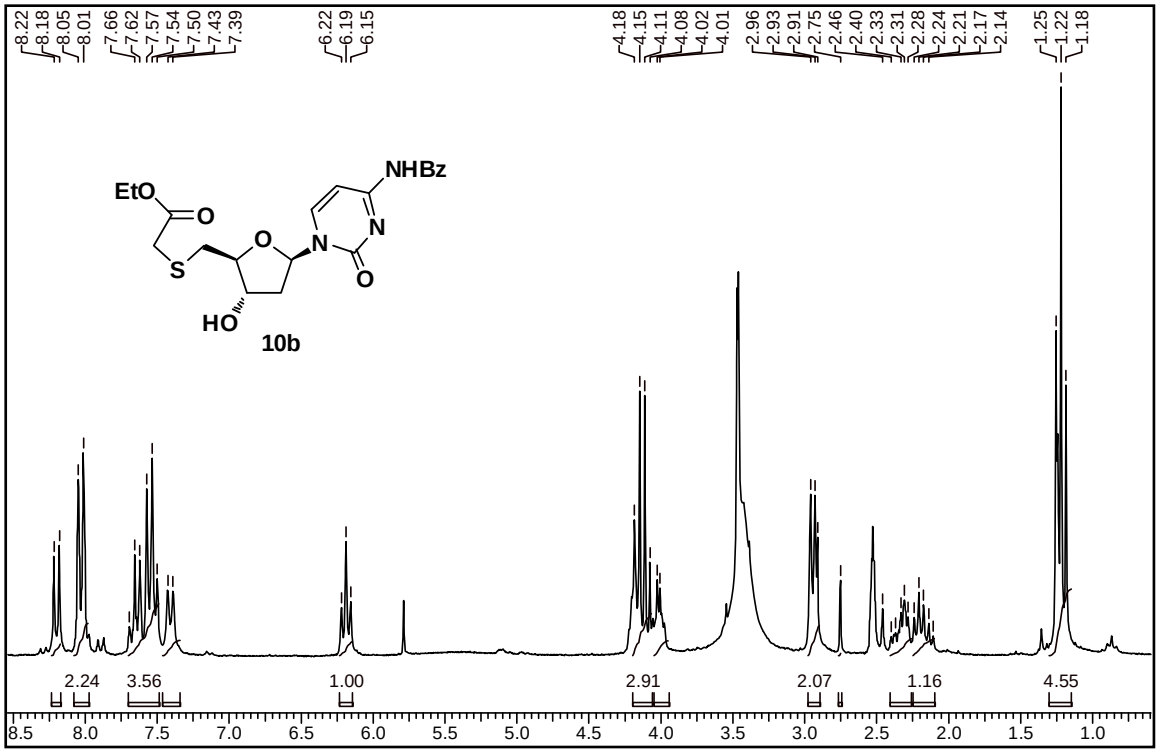

138.66
132.25
129.97
129.75
128.02
127.95
127.05
113.19
84.74
81.84
61.76
55.18
51.27
41.53
13.17
NHBz
N
N
O
DMTO
O
NH2
9
140
130
120
110
100
90
80
70
60
50
40
30
20
10

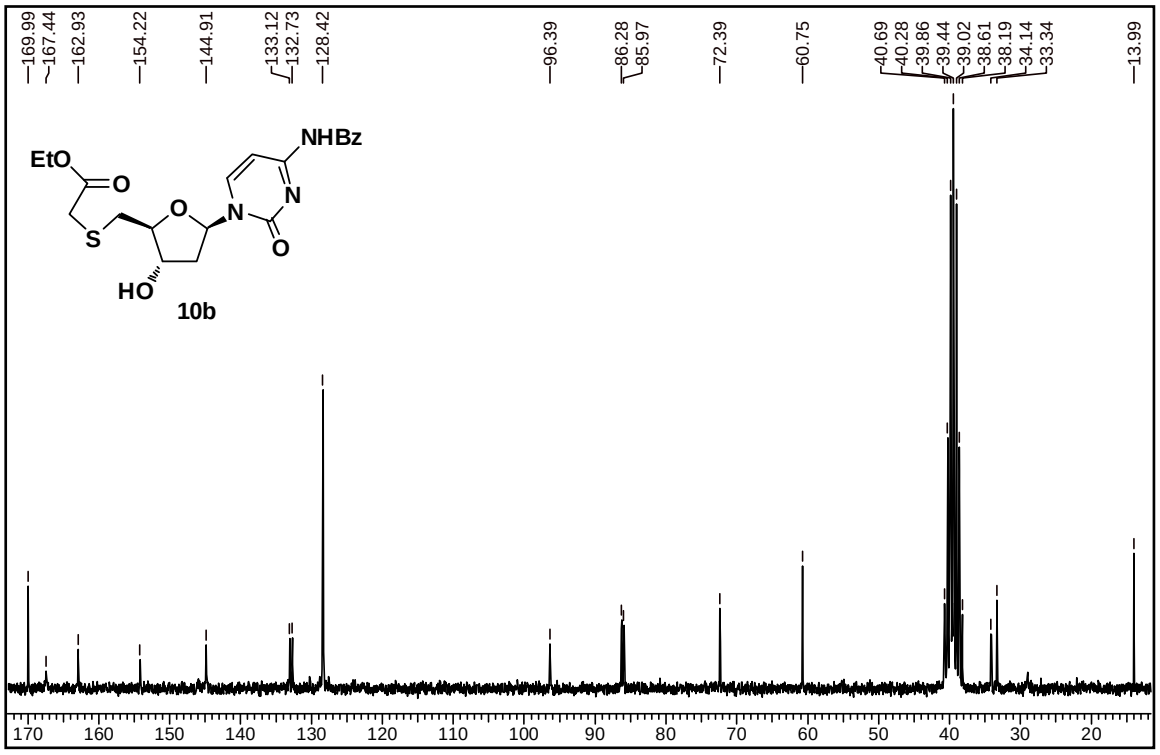

Chloroform-d
TMS
8.13
7.37
7.26
6.27
6.24
6.20
4.28
4.21
4.18
4.14
4.05
3.31
3.30
2.99
2.98
2.95
1.95
1.32
1.29
1.25
0.89
0.09
0.00
OEt
O
O
NH
S
O
N
O
TBDMSO
1c
0.85
0.96
0.98
1.98
1.87
1.80
1.02
3.00
3.68
9.29
5.84
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

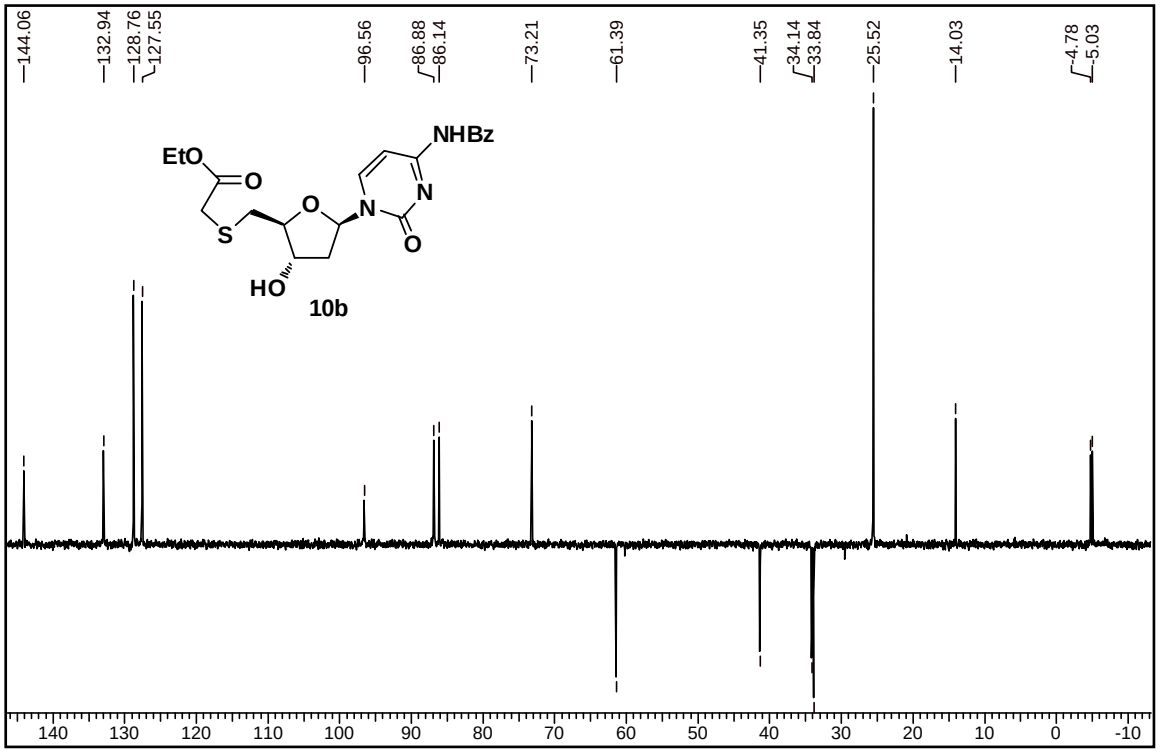

8.06
8.02
7.96
7.92
7.87
7.83
7.67
7.63
7.58
7.54
7.49
7.31
6.18
6.15
6.11
4.33
4.31
4.28
4.22
4.20
4.18
4.16
4.05
4.03
4.00
2.45
2.35
2.31
2.29
2.26
2.20
2.17
2.13
2.10
2.01
1.36
1.25
1.21
1.20
1.16
NHBz
N
N
O
N
O
TsO
HO
10a
2.20
4.52
1.00
1.78
1.32
3.00
1.31
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

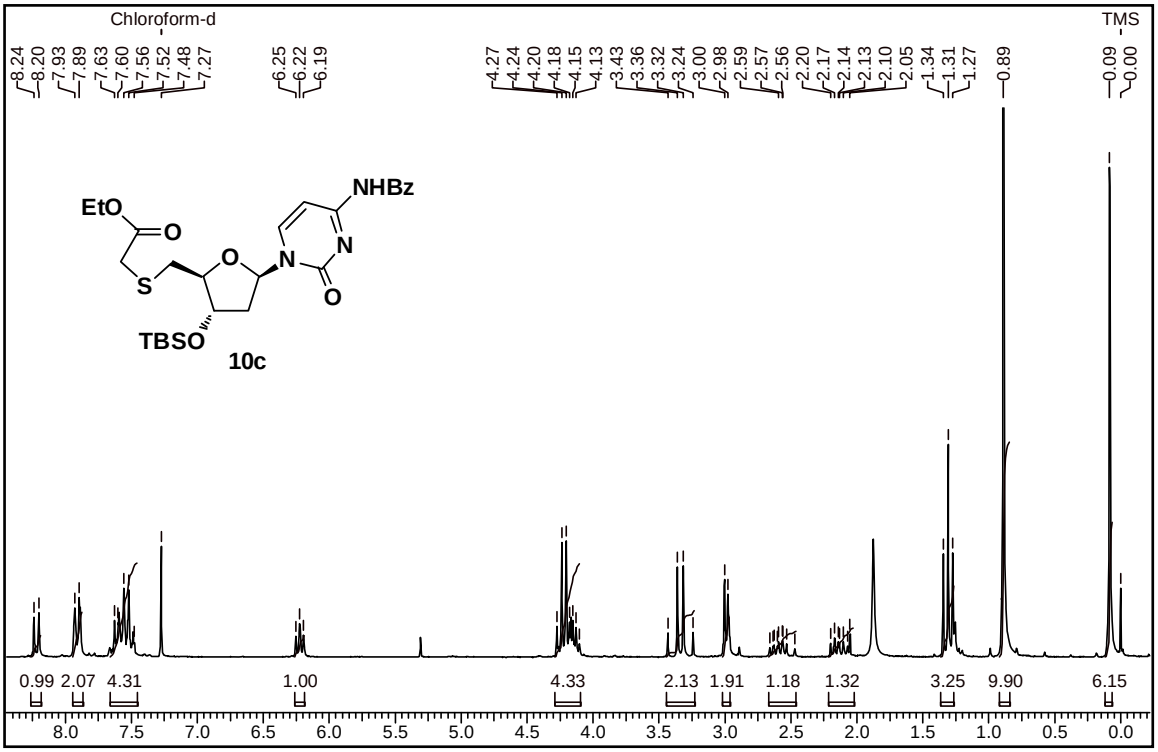

170.01
167.53
162.94
154.18
154.07
145.19
144.93
144.67
133.10
132.75
132.03
130.34
130.21
128.44
127.66
96.24
86.49
83.99
83.90
72.42
69.91
69.71
60.77
40.64
40.22
39.80
39.39
38.97
38.56
38.14
34.14
33.26
21.08
13.98
NHBz
N
N
O
N
O
TsO
HO
10a
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20

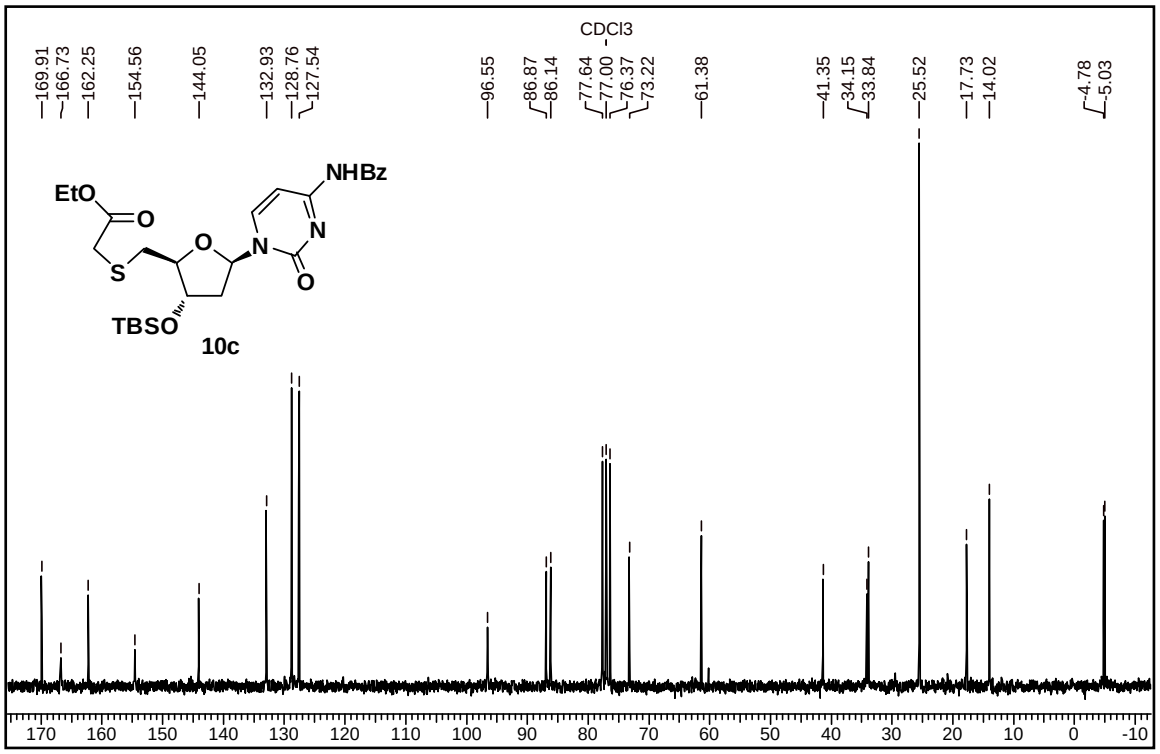

144.68
132.75
130.21
128.44
127.66
96.24
86.48
83.90
72.41
69.91
69.71
60.77
39.77
33.32
21.08
13.99
NHBz
N
N
O
O
TsO
HO
10a
140
130
120
110
100
90
80
70
60
50
40
30
20

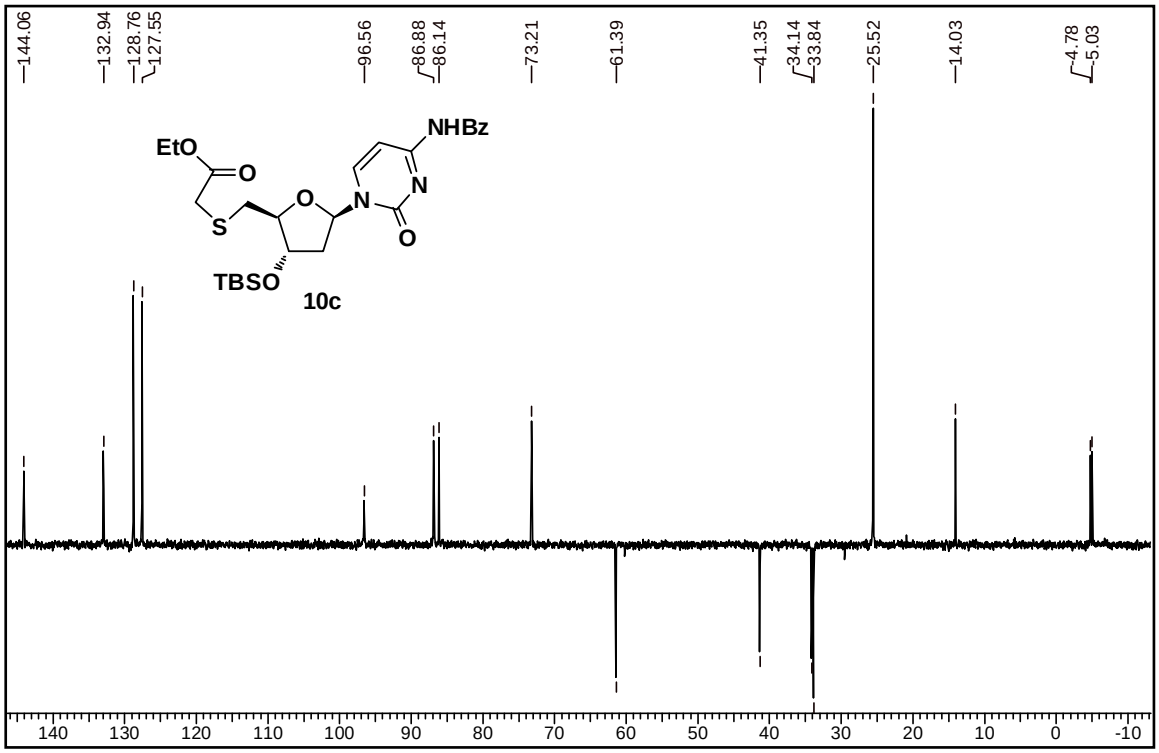

8.22
8.18
8.05
8.01
7.66
7.62
7.57
7.54
7.50
7.43
7.39
6.22
6.19
6.15
4.18
4.15
4.11
4.08
4.02
4.01
2.96
2.93
2.91
2.75
2.46
2.40
2.33
2.31
2.28
2.24
2.21
2.17
2.14
1.25
1.22
1.18
NHBz
EtO
O
N
O
S
O
HO
10b
2.24
3.56
1.00
2.91
2.07
1.16
4.55
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

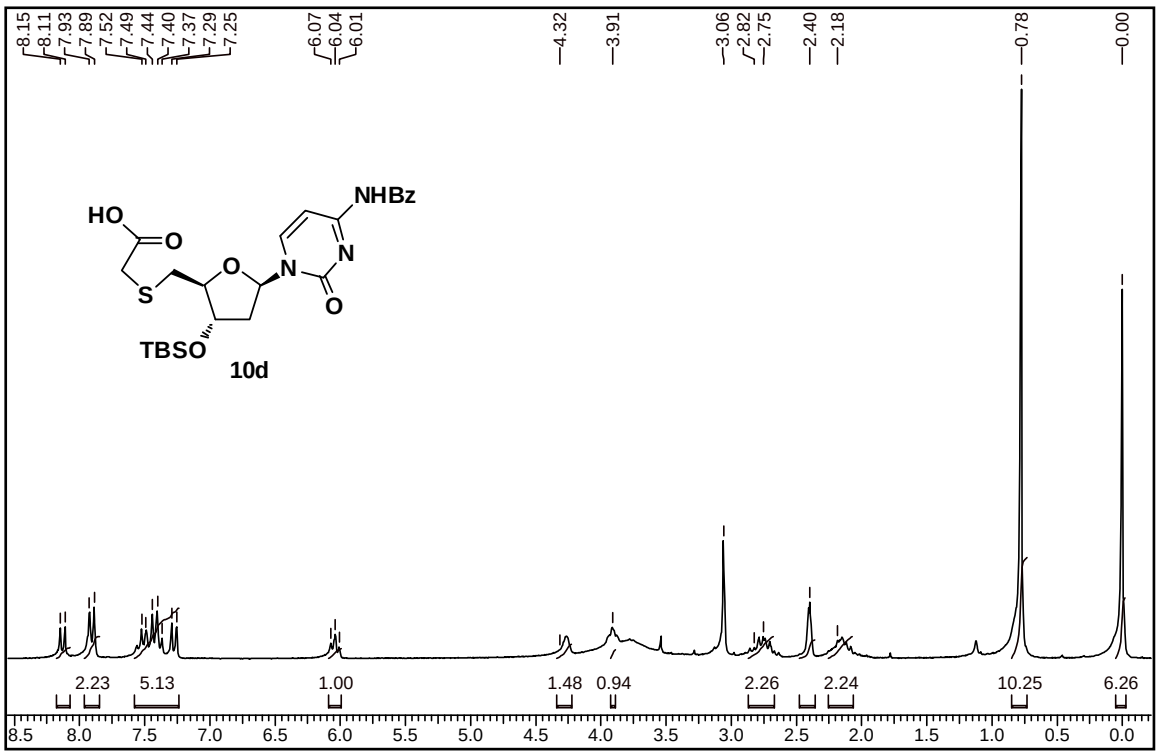

169.99
167.44
162.93
154.22
144.91
133.12
132.73
128.42
96.39
86.28
85.97
72.39
60.75
40.69
40.28
39.86
39.44
39.02
38.61
38.19
34.14
33.34
13.99
NHBz
EtO
O
O
N
N
S
O
HO
10b
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20

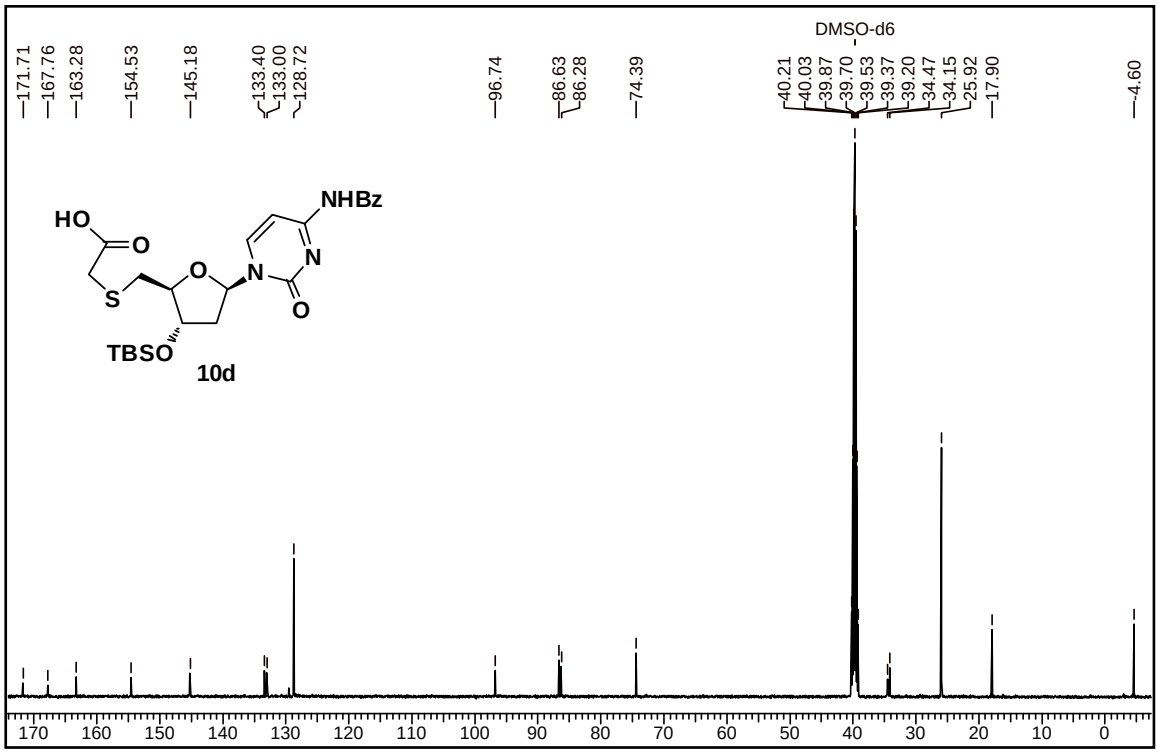

144.06
132.94
128.76
127.55
96.56
86.88
86.14
73.21
61.39
41.35
34.14
33.84
25.52
14.03
4.78
5.03
NHBz
EtO
O
O
N
N
S
O
HO
10b
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

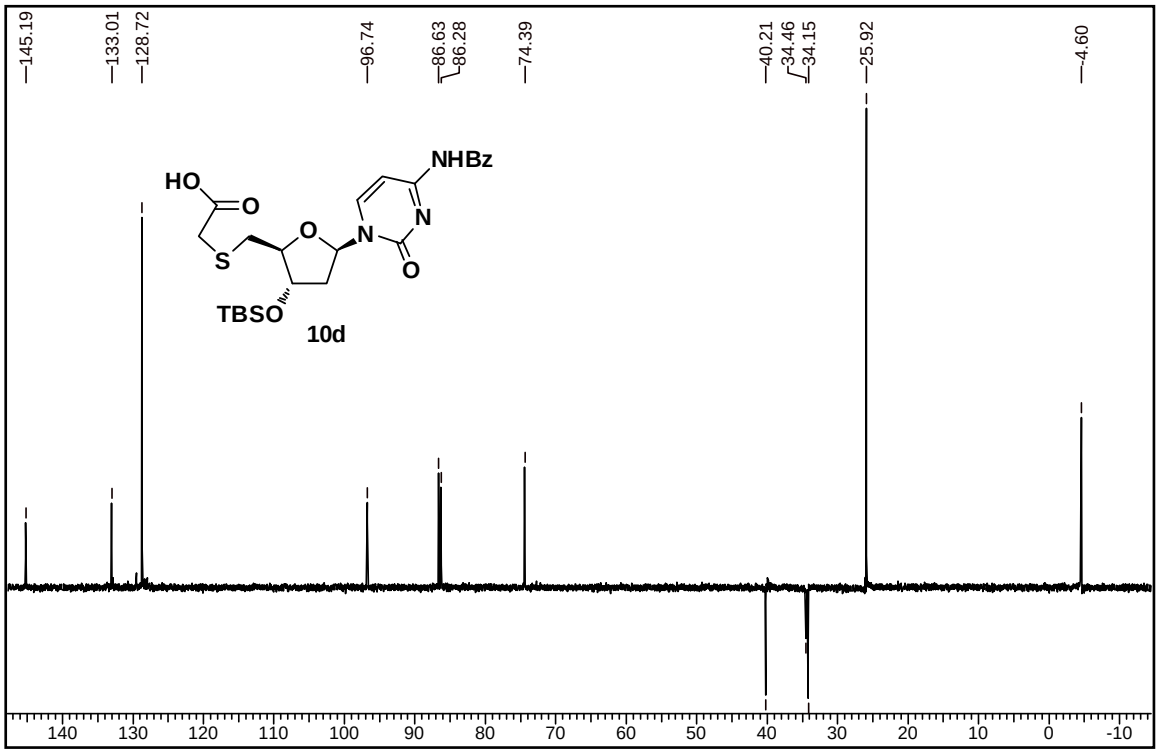

Chloroform-d
TMS
8.24
8.20
7.93
7.89
7.63
7.60
7.56
7.52
7.48
7.27
6.25
6.22
6.19
4.27
4.24
4.20
4.18
4.15
4.13
3.43
3.36
3.32
3.24
3.00
2.98
2.59
2.57
2.56
2.20
2.17
2.14
2.13
2.10
2.05
1.34
1.31
1.27
0.89
0.09
0.00
NHBz
EtO
O
N
N
O
S
O
TBSO
10c
0.99
2.07
4.31
1.00
4.33
2.13
1.91
1.18
1.32
3.25
9.90
6.15
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

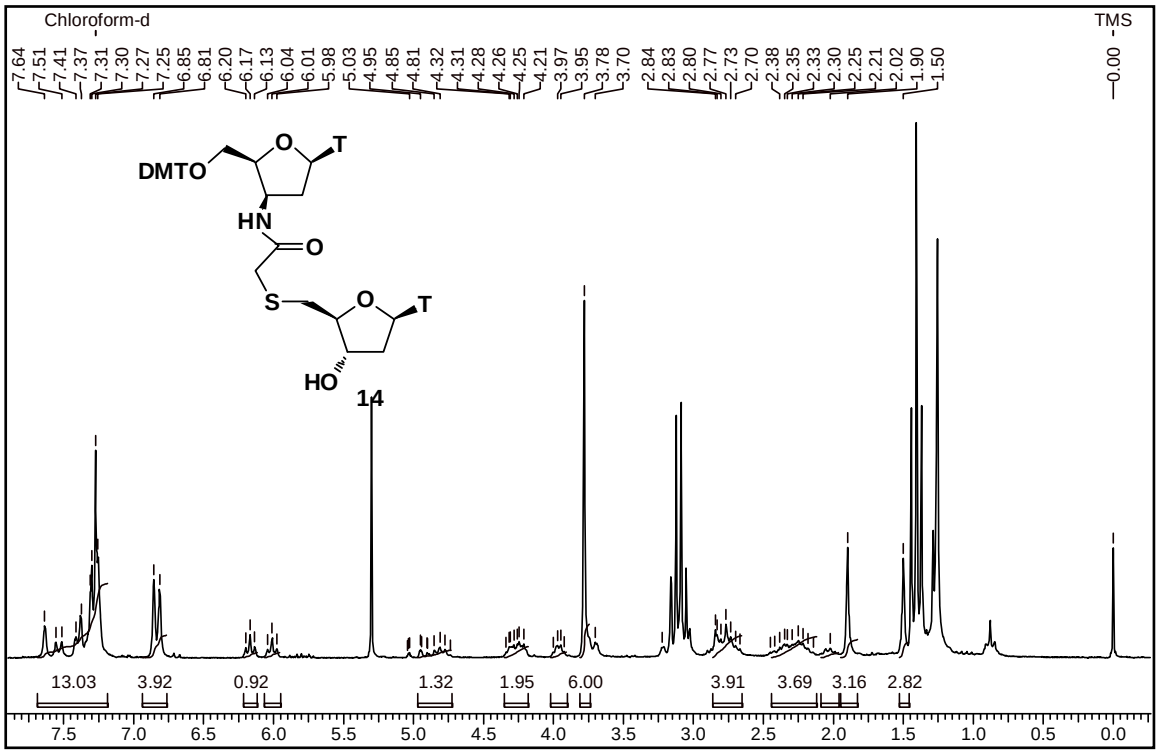

CDCl3
169.91
166.73
162.25
154.56
144.05
132.93
128.76
127.54
96.55
86.87
86.14
77.64
77.00
76.37
73.22
61.38
41.35
34.15
33.84
25.52
17.73
14.02
4.78
5.03
NHBz
EtO
O
N
N
O
S
O
TBSO
10c
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

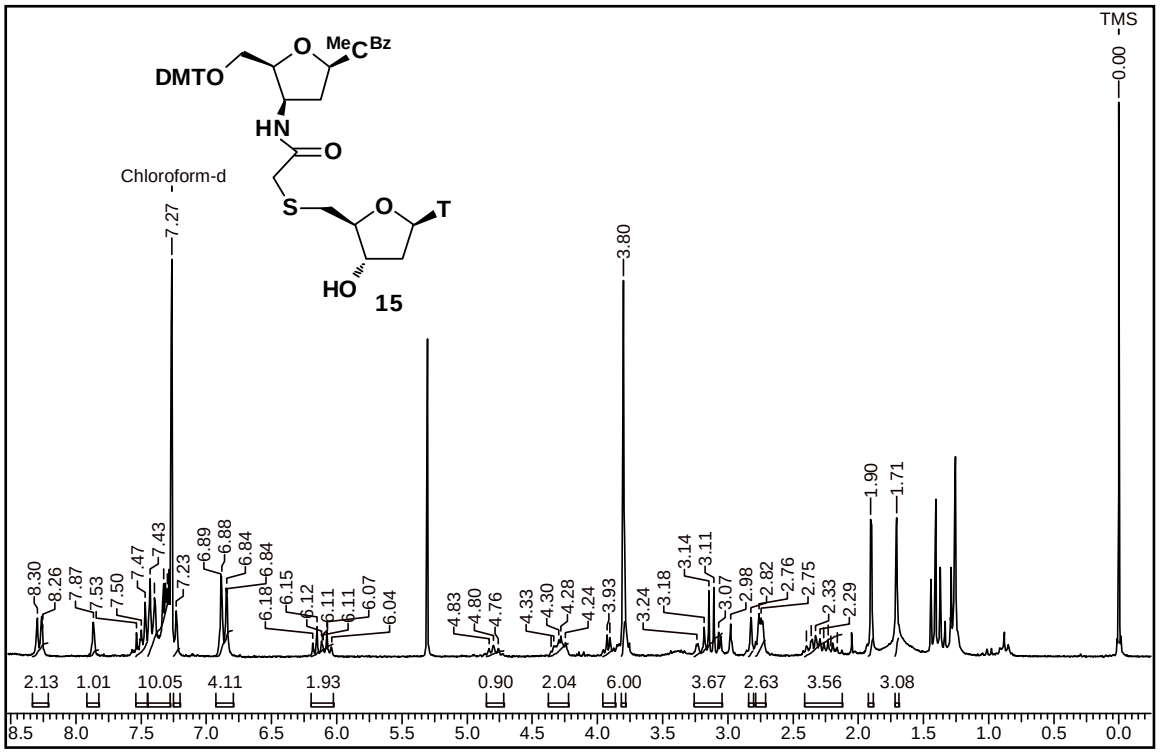

144.06
132.94
128.76
127.55
96.56
86.88
86.14
73.21
61.39
41.35
34.14
33.84
25.52
14.03
4.78
5.03
EtO
O
S
O
N
N
NHBz
O
TBSO
10c
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

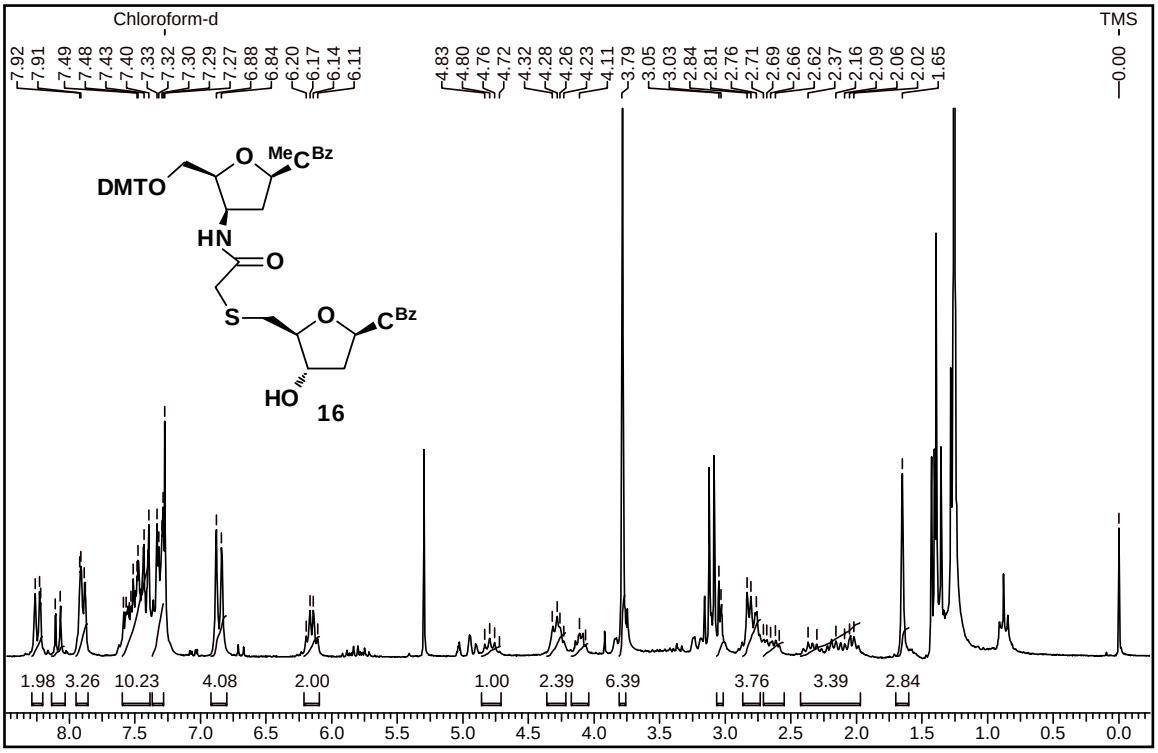

8.15
8.11
7.93
7.89
7.52
7.49
7.44
7.40
7.37
7.29
7.25
6.07
6.04
6.01
4.32
3.91
3.06
2.82
2.75
2.40
2.18
0.78
0.00
HO
O
S
O
N
N
NHBz
O
TBSO
10d
2.23
5.13
1.00
1.48
0.94
2.26
2.24
10.25
6.26
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

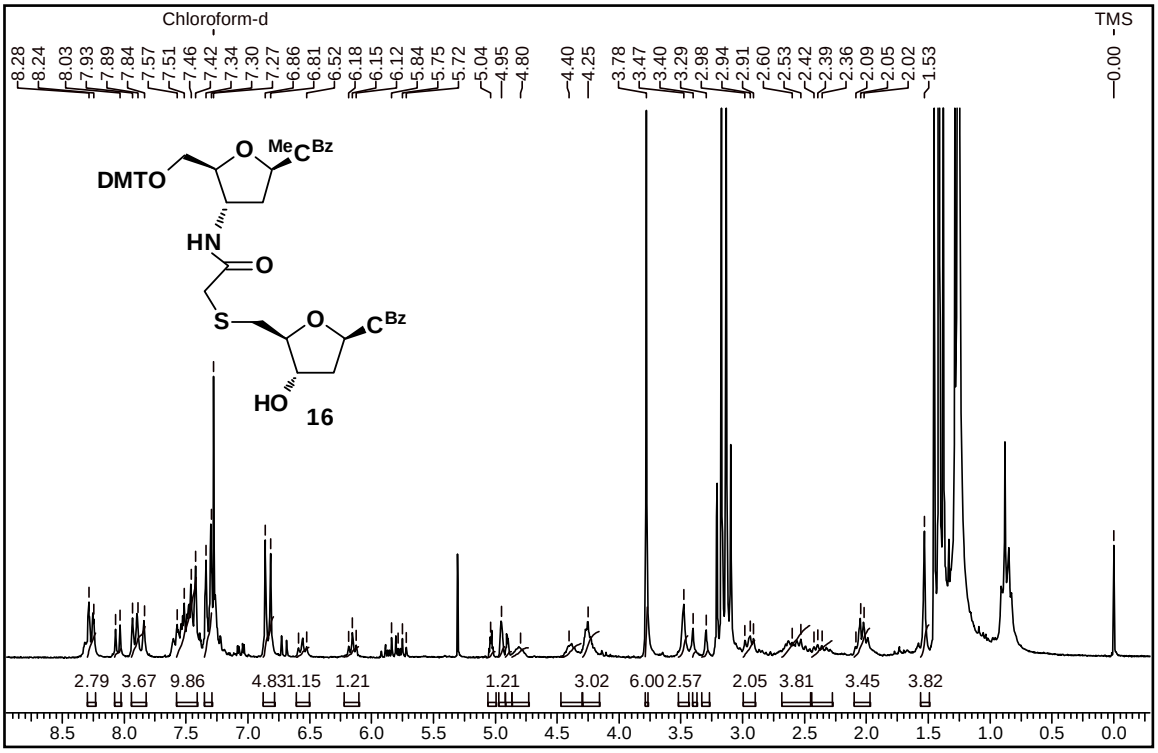

DMSO-d6
171.71
167.76
163.28
154.53
145.18
133.40
133.00
128.72
96.74
86.63
86.28
74.39
40.21
40.03
39.87
39.70
39.53
39.37
39.20
34.47
34.15
25.92
17.90
4.60
HO
O
S
O
N
N
NHBz
O
TBSO
10d
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

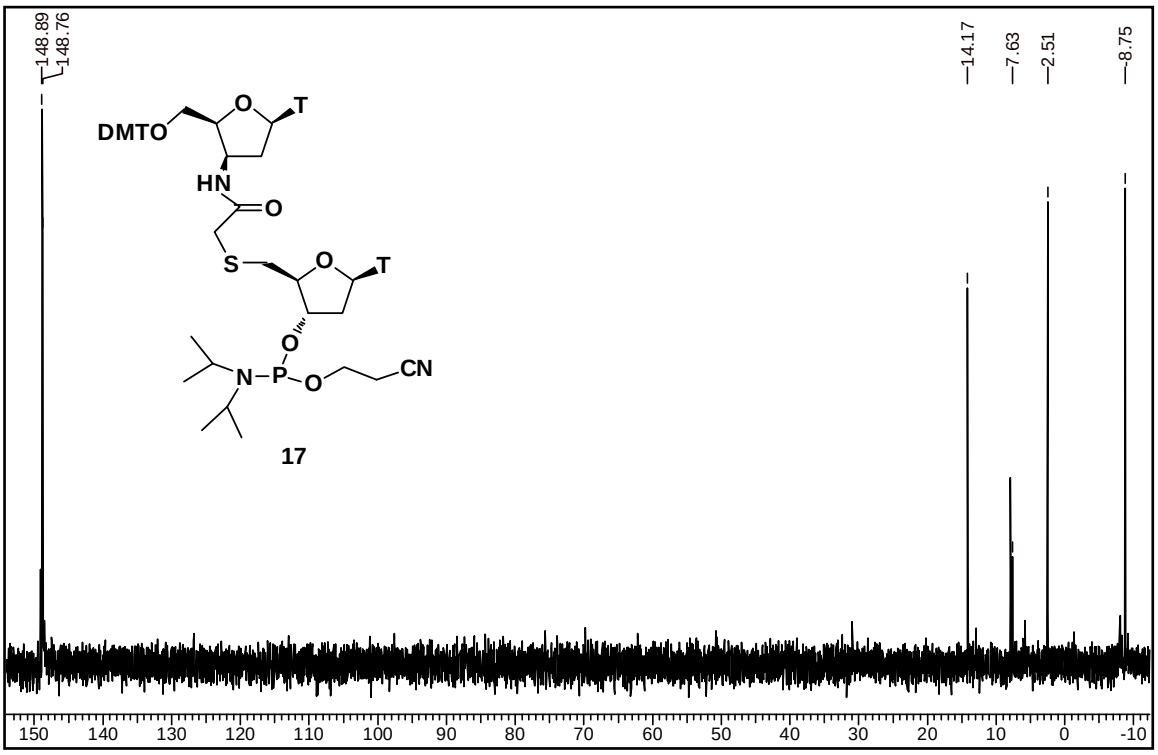

145.19
133.01
128.72
96.74
86.63
86.28
74.39
40.21
34.46
34.15
25.92
4.60
HO
O
S
O
N
N
NHBz
O
TBSO
10d
140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10


Chloroform-d
TMS
DMTO
HN
S
HO
14
13.03
3.92
0.92
1.32
1.95
6.00
3.91
3.69
3.16
2.82

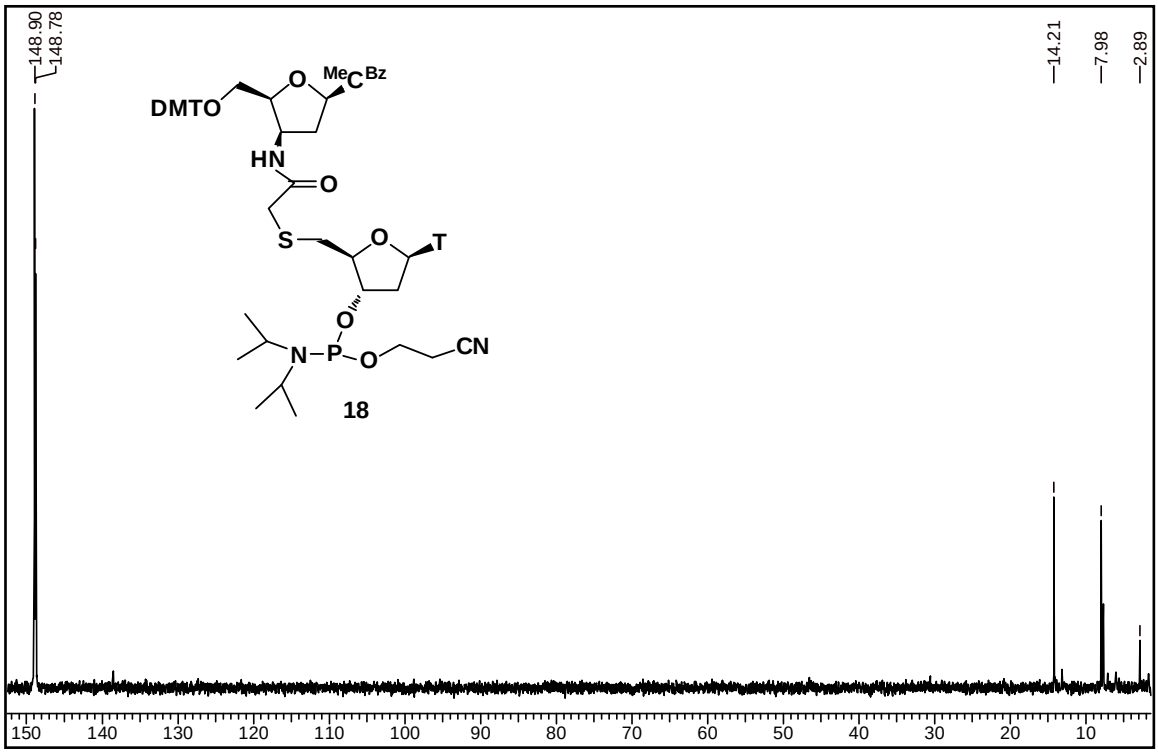

TMS
DMTO
HN
Chloroform-d
S
HO
15
2.13
1.01
10.05
4.11
1.93
0.90
2.04
6.00
3.67
2.63
3.56
3.08


Chloroform-d
TMS
DMTO
HN
O
S
HO
16
1.98 3.26 10.23 4.08 2.00 1.00 2.39 6.39 3.76 3.39 2.84

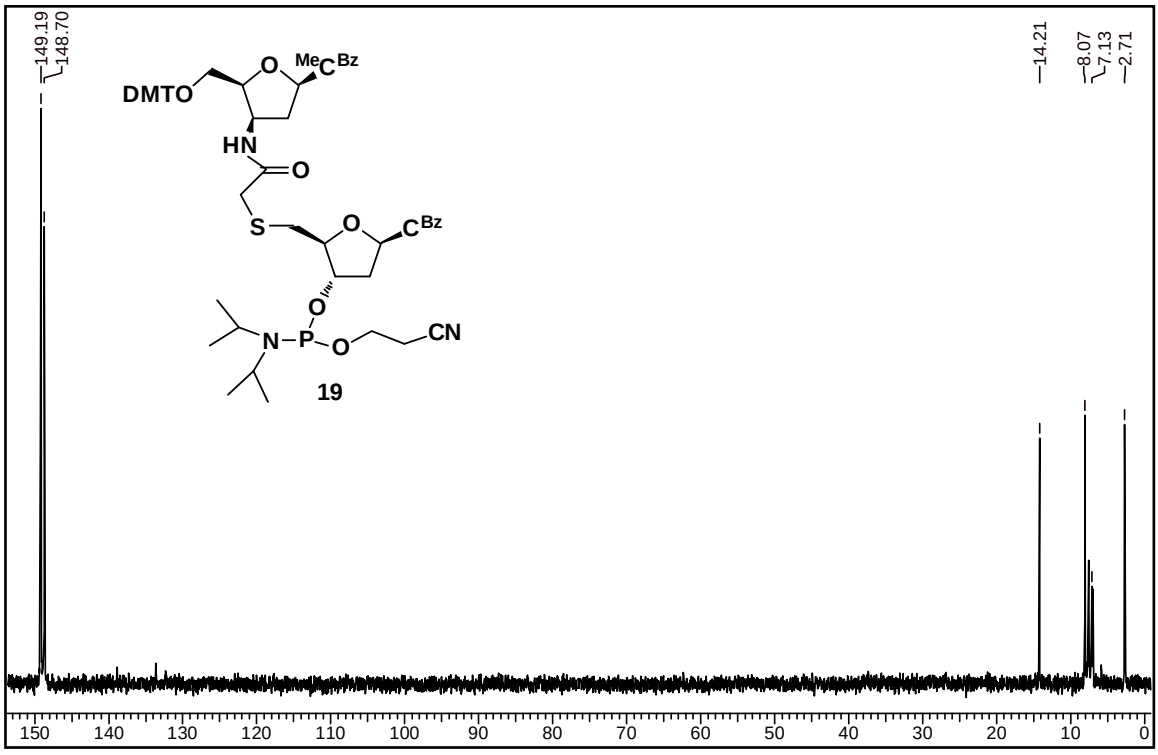

Chloroform-d
TMS
DMTO
HN
O
S
HO
16
2.79 3.67 9.86 4.83 1.15 1.21 1.21 3.02 6.00 2.57 2.05 3.81 3.45 3.82

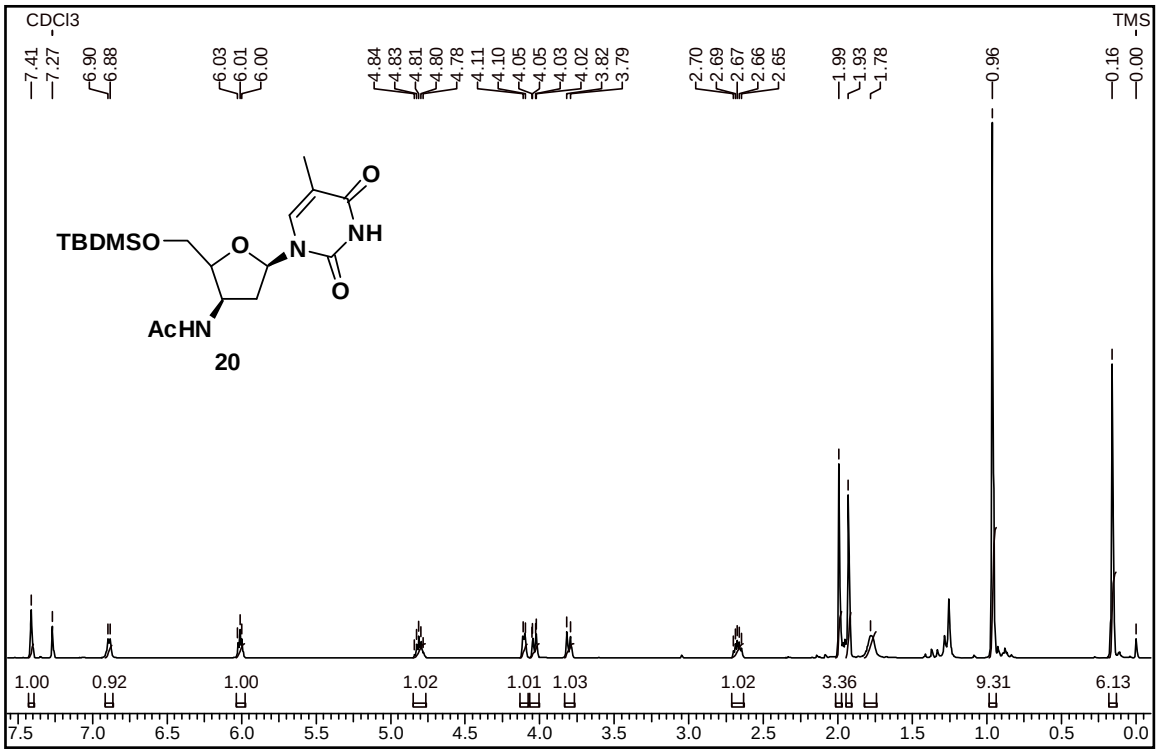

148.89
148.76
14.17
7.63
2.51
8.75
DMTO
O
T
HN
O
S
O
T
O
N
P
O
CN
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

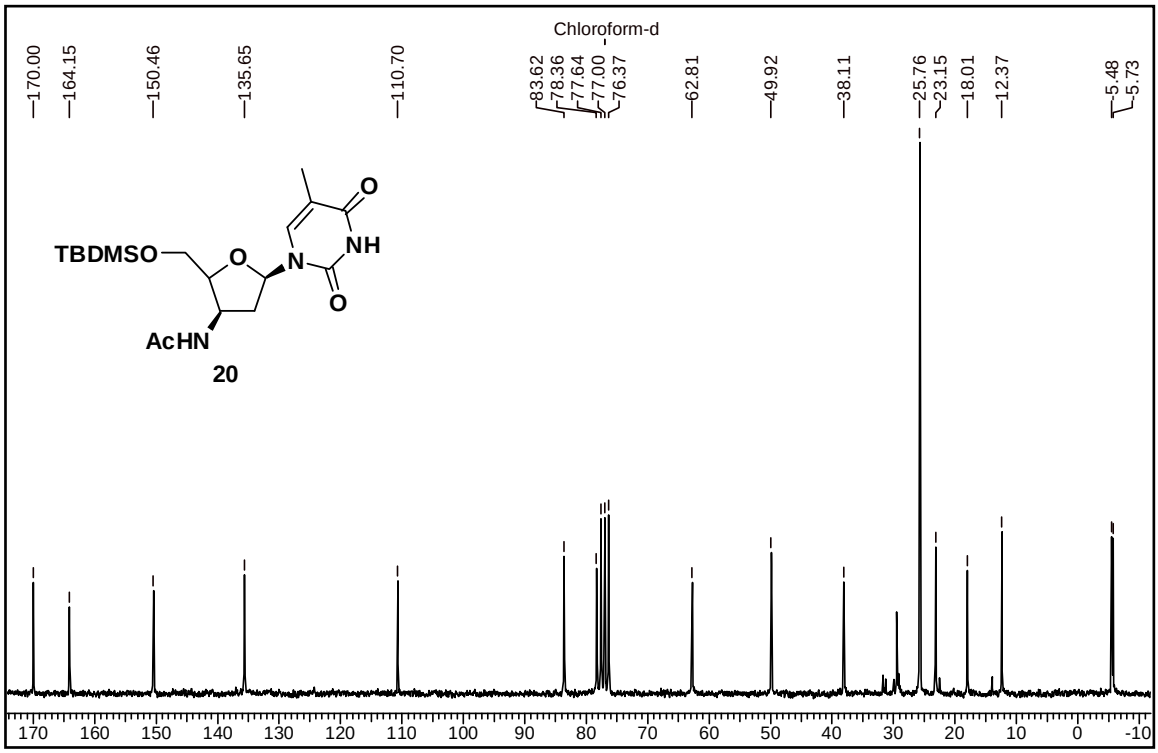
17

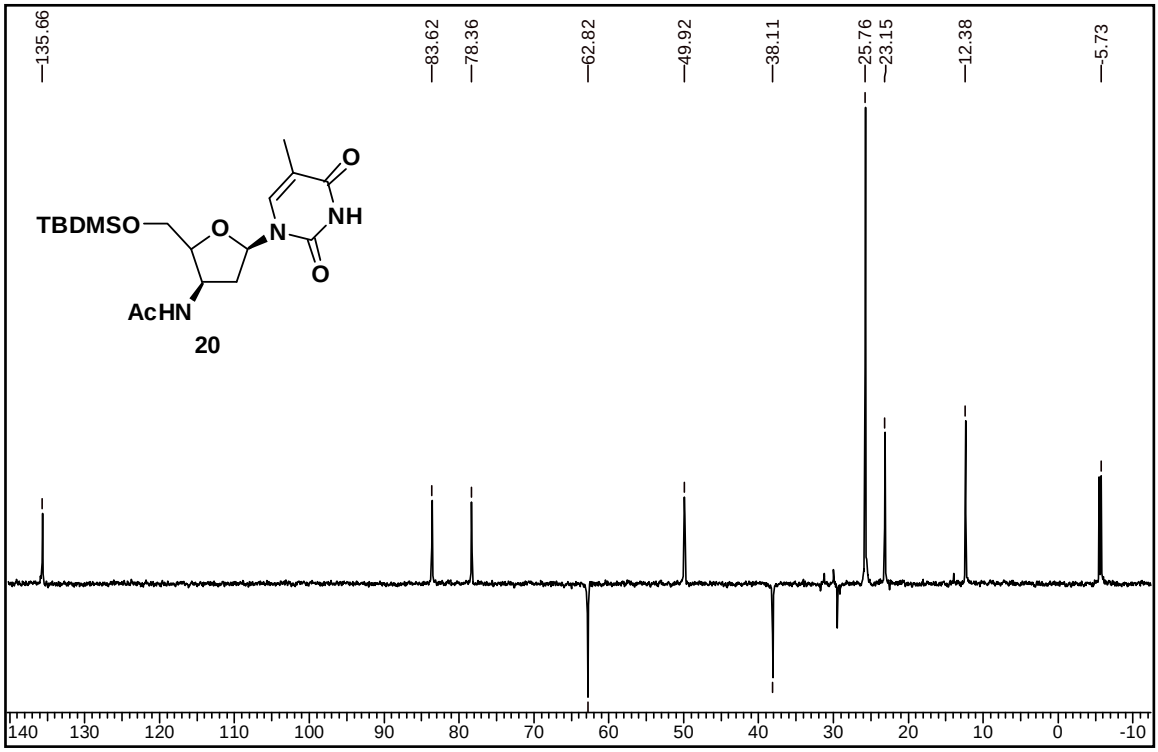

148.90
148.78
14.21
7.98
2.89
DMTO
O
Me
C
Bz
HN
O
S
O
T
O
N
P
O
CN
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

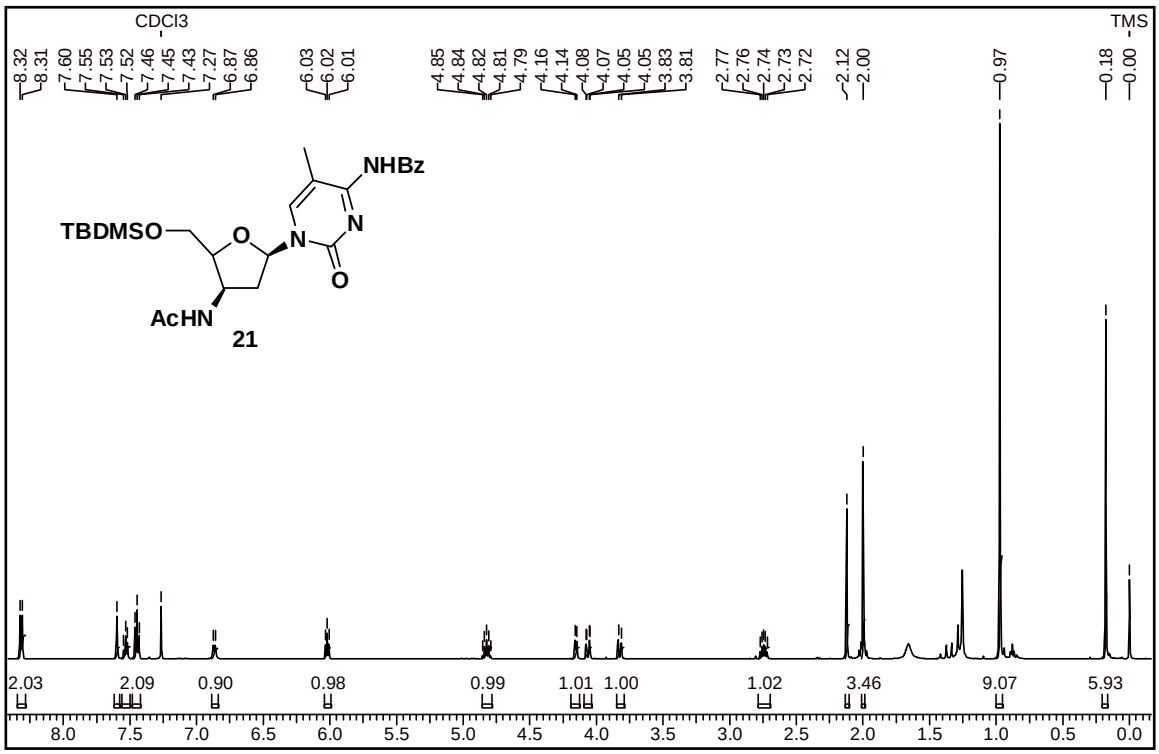
18

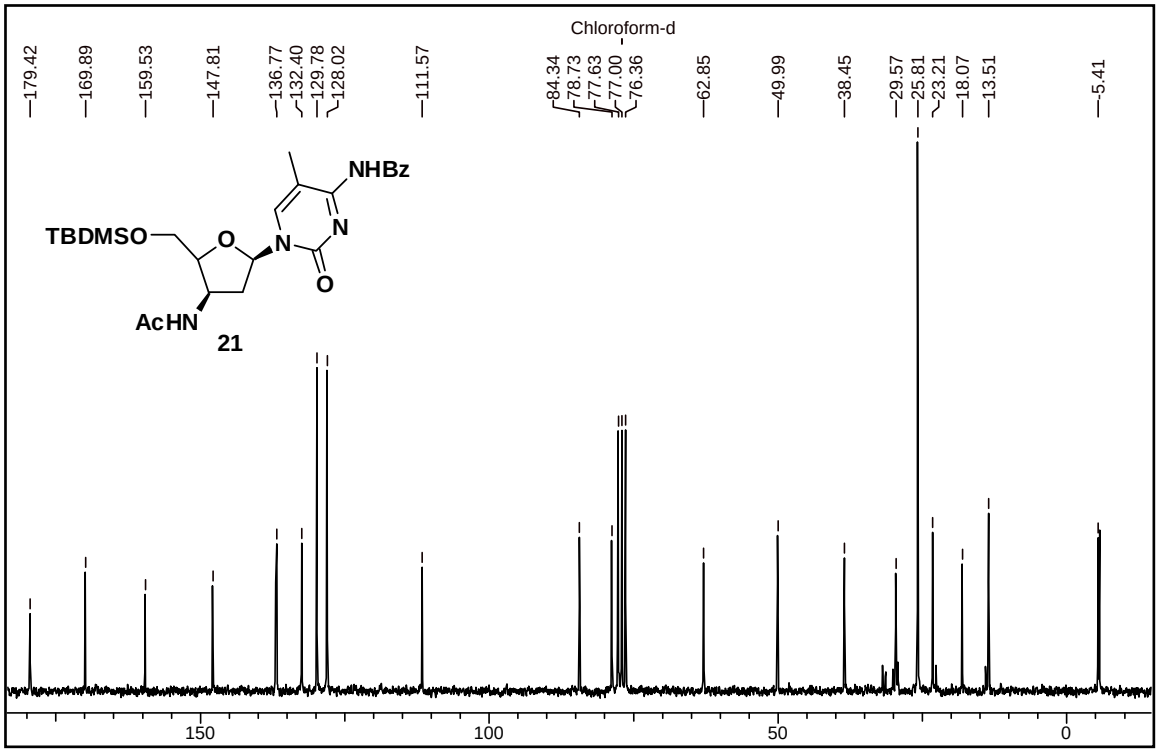

149.19
148.70
14.21
8.07
7.13
2.71
DMTO
Me
Bz
HN
O
S
CN
N
P
19
150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

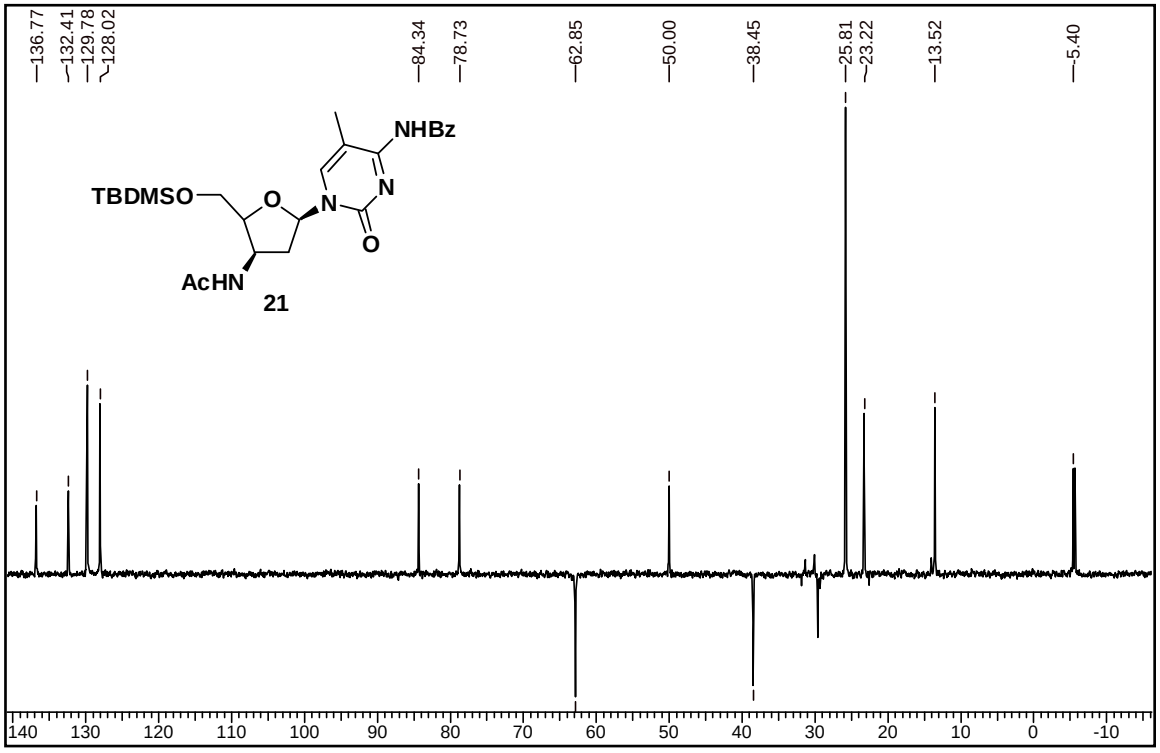

CDCl3
TMS
7.41
7.27
6.90
6.88
6.03
6.01
6.00
4.84
4.83
4.81
4.80
4.78
4.11
4.10
4.05
4.05
4.03
4.02
3.82
3.79
2.70
2.69
2.67
2.66
2.65
1.99
1.93
1.78
0.96
0.16
0.00
TBDMSO
O
NH
N
AcHN
20
1.00 0.92 1.00 1.02 1.01 1.03 1.02 3.36 9.31 6.13
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

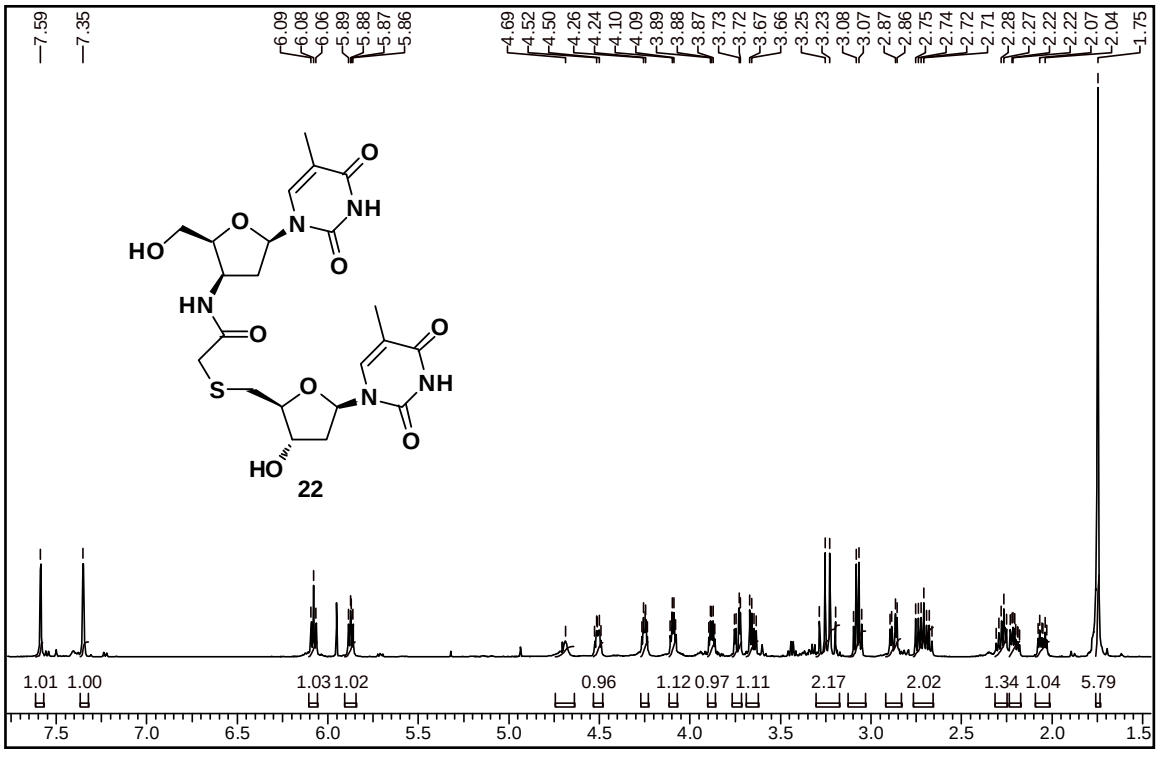

Chloroform-d
170.00
164.15
150.46
135.65
110.70
83.62
78.36
77.64
77.00
76.37
62.81
49.92
38.11
25.76
23.15
18.01
12.37
5.48
5.73
TBDMSO
O
N
NH
O
AcHN
20
170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10

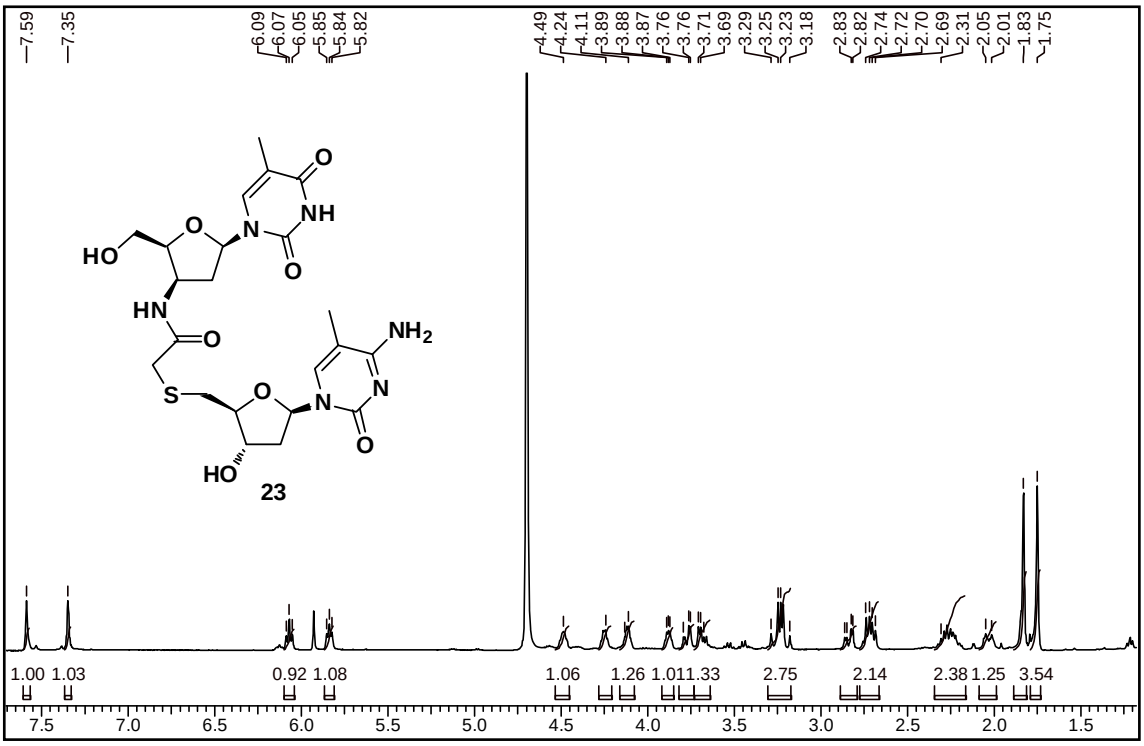

135.66
83.62
78.36
62.82
49.92
38.11
25.76
23.15
12.38
5.73
TBDMSO
O
N
NH
O
AcHN
20
140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10

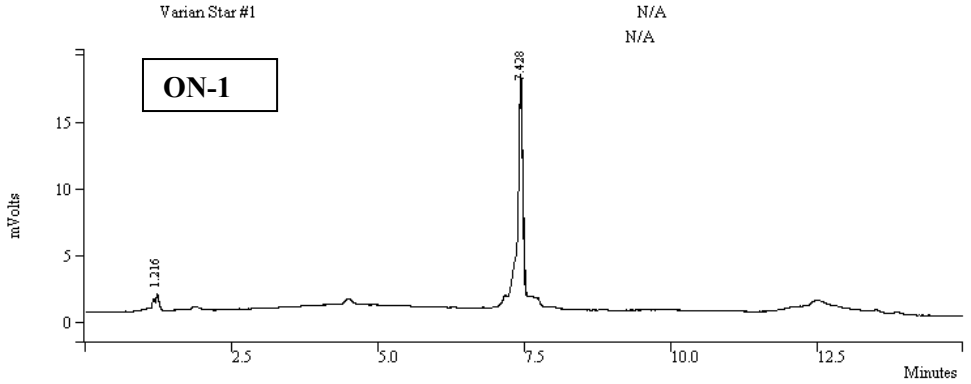

CDCl3
TMS
8.32
8.31
7.60
7.55
7.53
7.52
7.46
7.45
7.43
7.27
6.87
6.86
6.03
6.02
6.01
4.85
4.84
4.82
4.81
4.79
4.16
4.14
4.08
4.07
4.05
4.05
3.83
3.81
2.77
2.76
2.74
2.73
2.72
2.12
2.00
0.97
0.18
0.00
NHBz
TBDMSO
O
N
N
O
AcHN
21
2.03
2.09
0.90
0.98
0.99
1.01
1.00
1.02
3.46
9.07
5.93
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

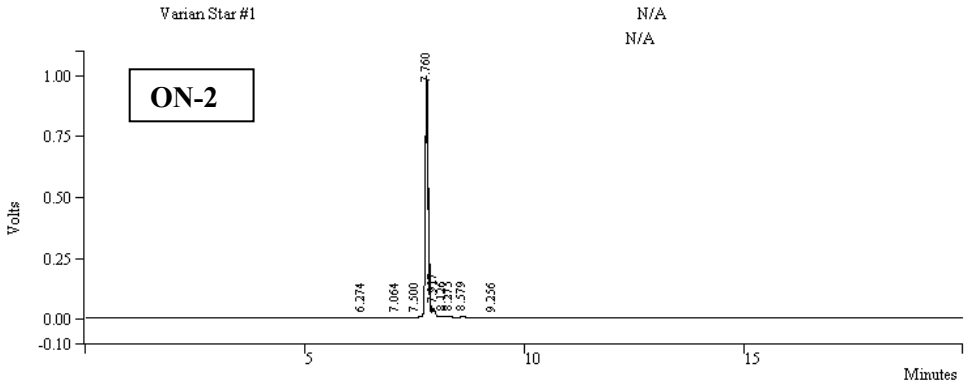

Chloroform-d
179.42
169.89
159.53
147.81
136.77
132.40
129.78
128.02
111.57
84.34
78.73
77.63
77.00
76.36
62.85
49.99
38.45
29.57
25.81
23.21
18.07
13.51
5.41
NHBz
TBDMSO
O
N
N
O
AcHN
21
150
100
50
0

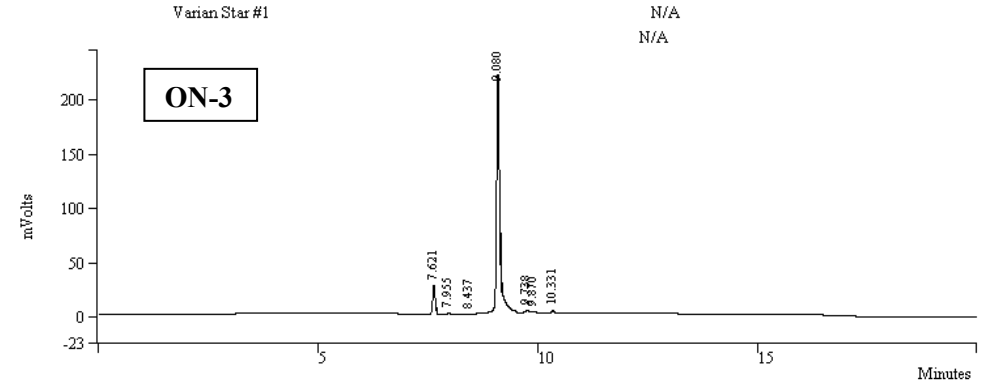

136.77
132.41
129.78
128.02
84.34
78.73
62.85
50.00
38.45
25.81
23.22
13.52
5.40
NHBz
TBDMSO
O
N
N
O
AcHN
21
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

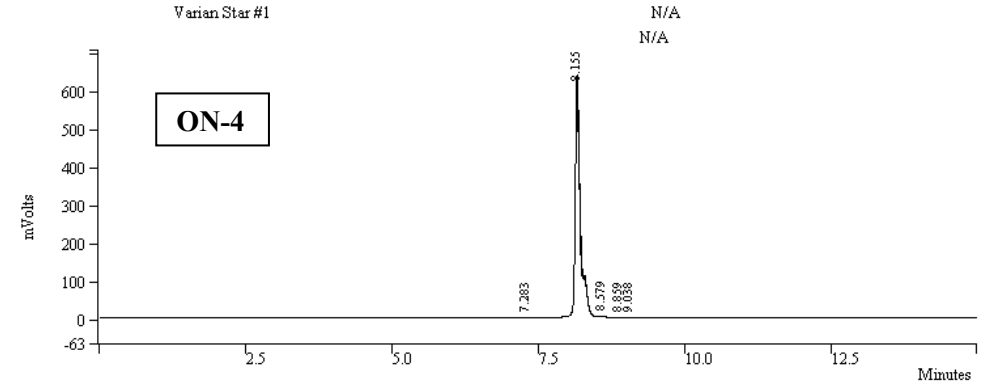

7.59
7.35
6.09
6.08
6.06
5.89
5.88
5.87
5.86
4.69
4.52
4.50
4.26
4.24
4.10
4.09
3.89
3.88
3.87
3.73
3.72
3.67
3.66
3.25
3.23
3.08
3.07
2.87
2.86
2.75
2.74
2.72
2.71
2.28
2.27
2.22
2.22
2.07
2.04
1.75
O
NH
HO
O
N
O
HN
O
S
O
N
NH
O
HO
22
1.01
1.00
1.03
1.02
0.96
1.12
0.97
1.11
2.17
2.02
1.34
1.04
5.79
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5

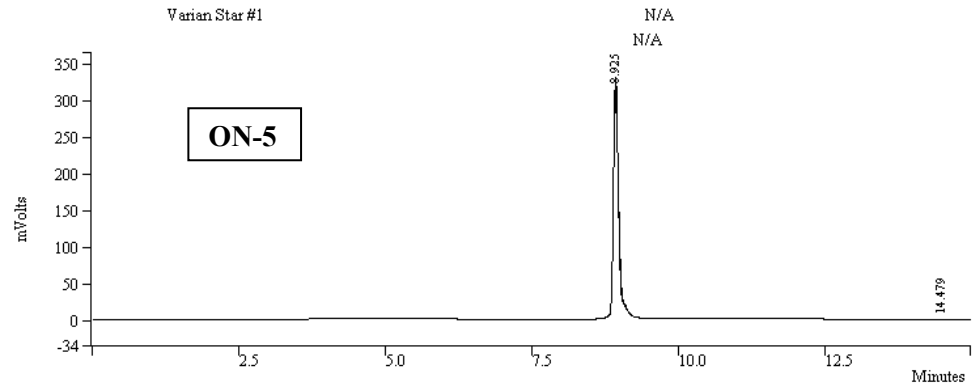

7.59
7.35
6.09
6.07
6.05
5.85
5.84
5.82
4.49
4.24
4.11
3.89
3.88
3.87
3.76
3.76
3.71
3.69
3.29
3.25
3.23
3.18
2.83
2.82
2.74
2.72
2.70
2.69
2.31
2.05
2.01
1.83
1.75
O
NH
O
N
O
HO
HN
O
S
O
N
N
NH2
N
O
HO
23
1.00 1.03 0.92 1.08 1.06 1.26 1.01 1.33 2.75 2.14 2.38 1.25 3.54
7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5

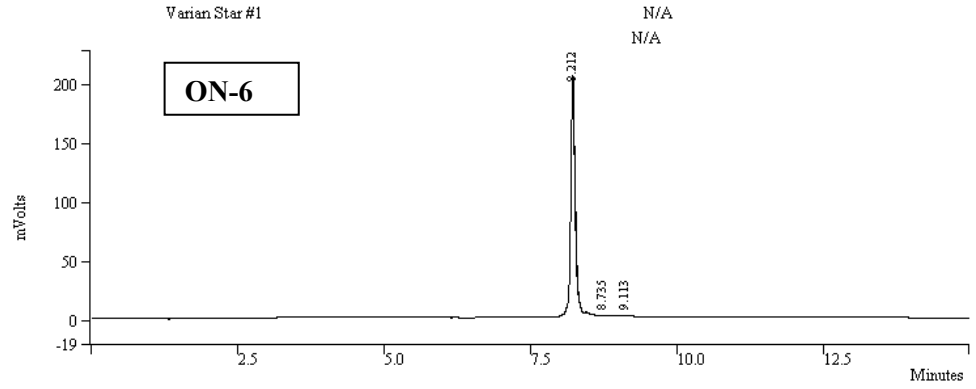

Varian Star #1
N/A
N/A
ON-1
7.428
1.216
mVolts
15
10
5
0
2.5
5.0
7.5
10.0
12.5
Minutes

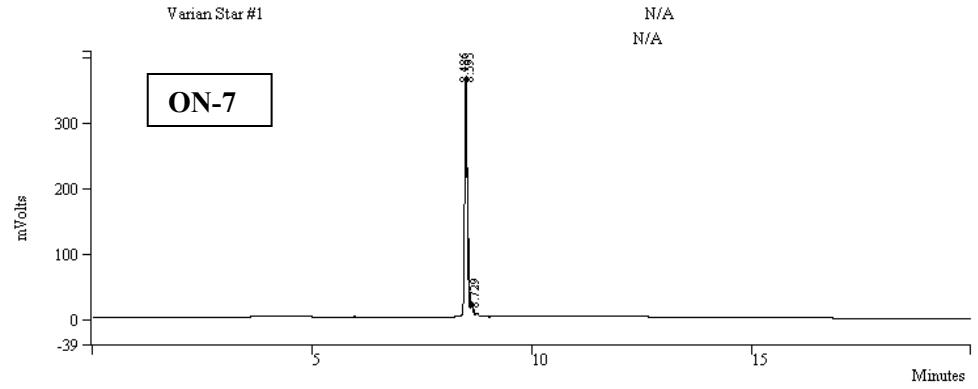

Varian Star #1
N/A
N/A
ON-2
7.760
6.274
7.064
7.500
8.379
9.256
Volts
1.00
0.75
0.50
0.25
0.00
-0.10
5
10
15
Minutes

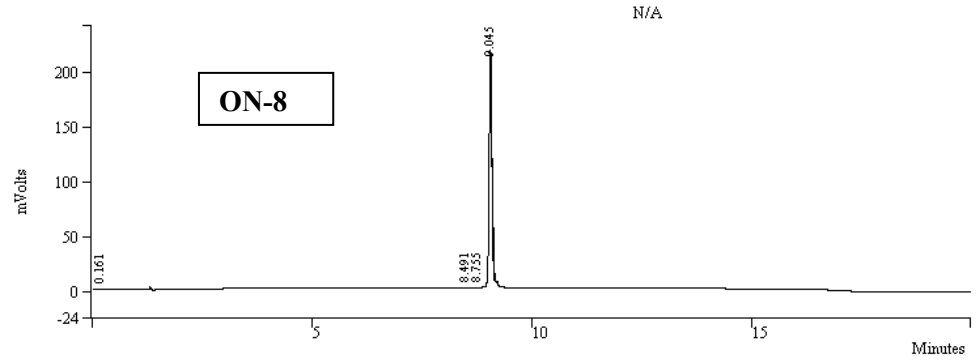

Varian Star #1
N/A
N/A
ON-3
8.080
7.621
7.955
8.437
10.331
mVolts
200
150
100
50
0
-23
5
10
15
Minutes

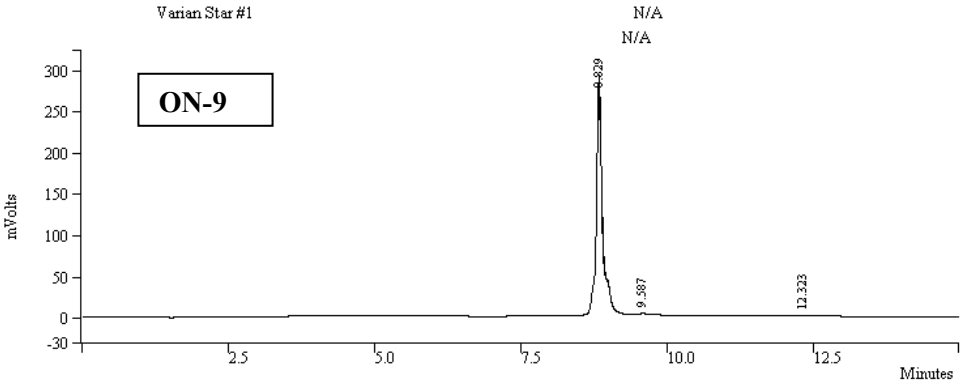

Varian Star #1
N/A
N/A
ON-4
8.155
7.283
8.579
8.838
9.038
mVolts
600
500
400
300
200
100
0
-63
2.5
5.0
7.5
10.0
12.5
Minutes

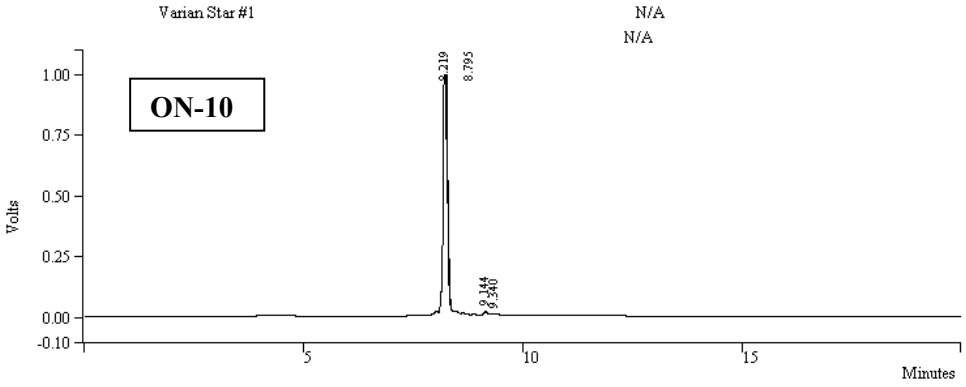

Varian Star #1
N/A
N/A
ON-5
mVolts
Minutes
8.925
14.479

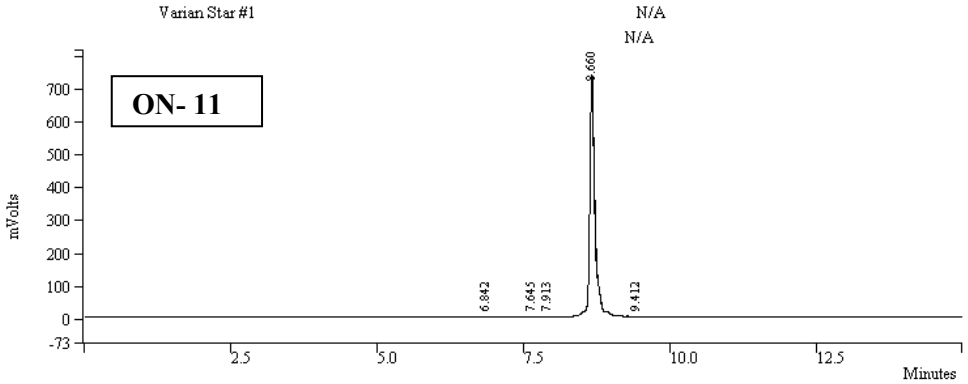

Varian Star #1
N/A
N/A
ON-6
mVolts
Minutes
8.212
8.735
9.113

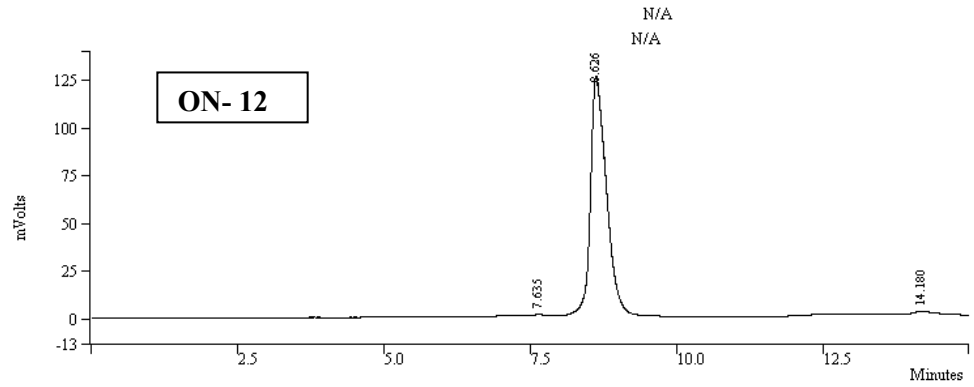

Varian Star #1
N/A
N/A
ON-7
mVolts
Minutes
8.386
8.729

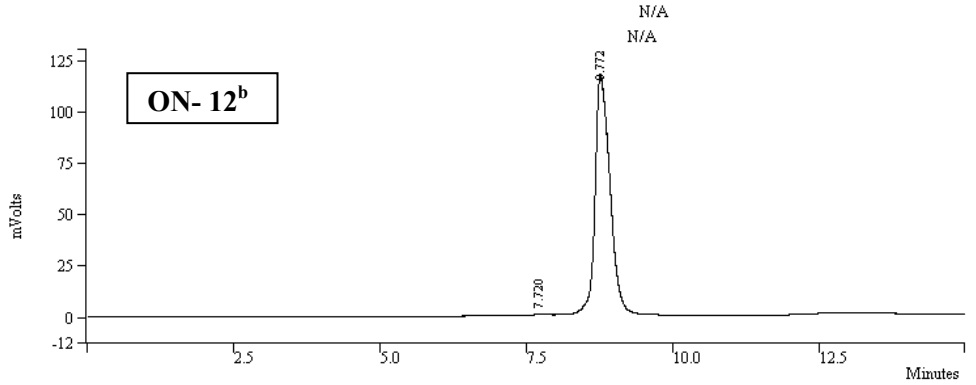

N/A
ON-8
mVolts
Minutes
9.045
0.161
8.491
8.735

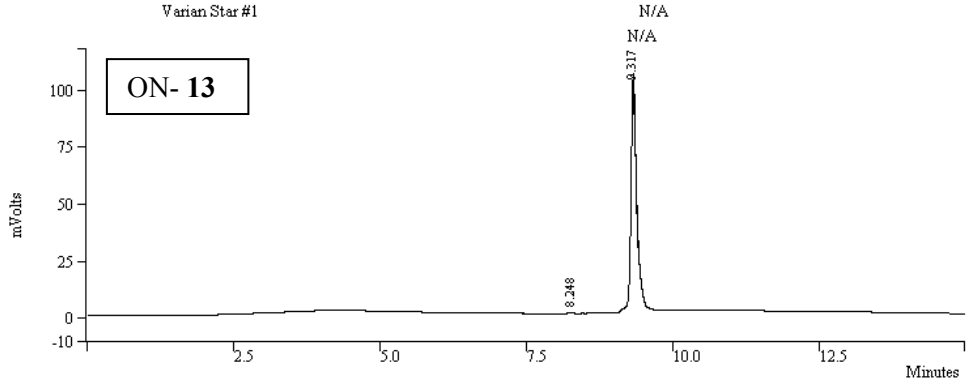

Varian Star #1
N/A
N/A
ON-9
mVolts
Minutes
8.829
9.587
12.323

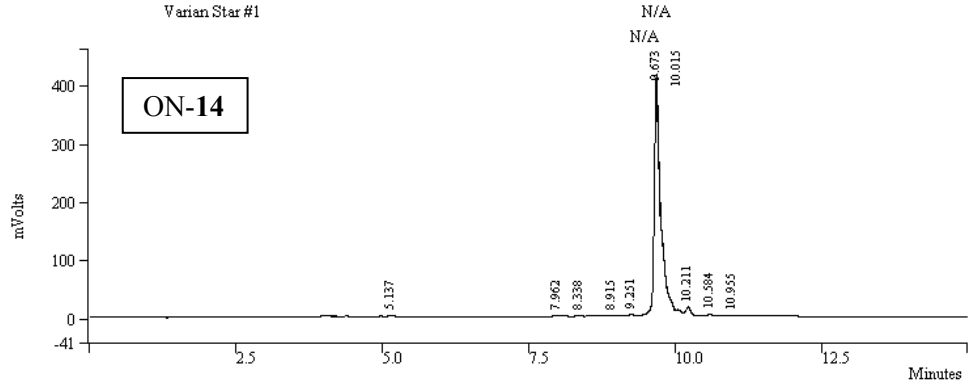

Varian Star #1
N/A
N/A
ON-10
Volts
Minutes
8.219
8.795

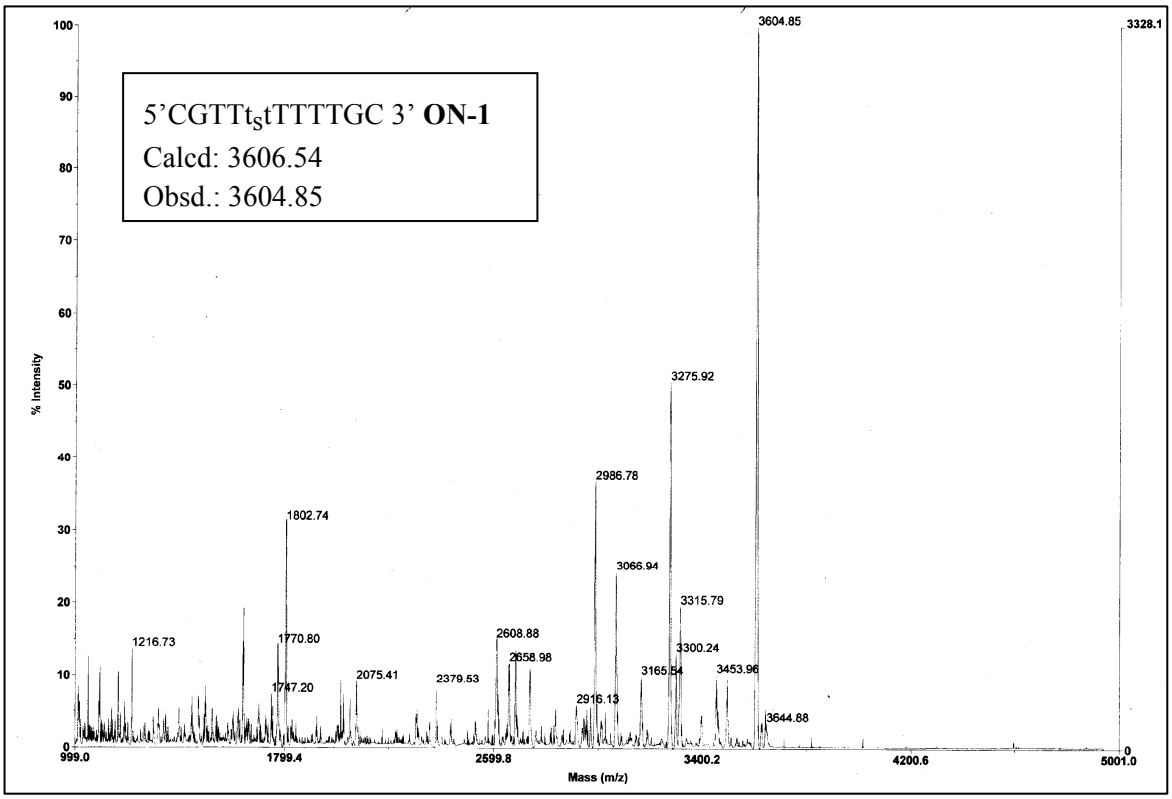

Varian Star #1
N/A
N/A
ON- 11
mVolts
Minutes
8.660
6.842
7.645
7.913
9.412

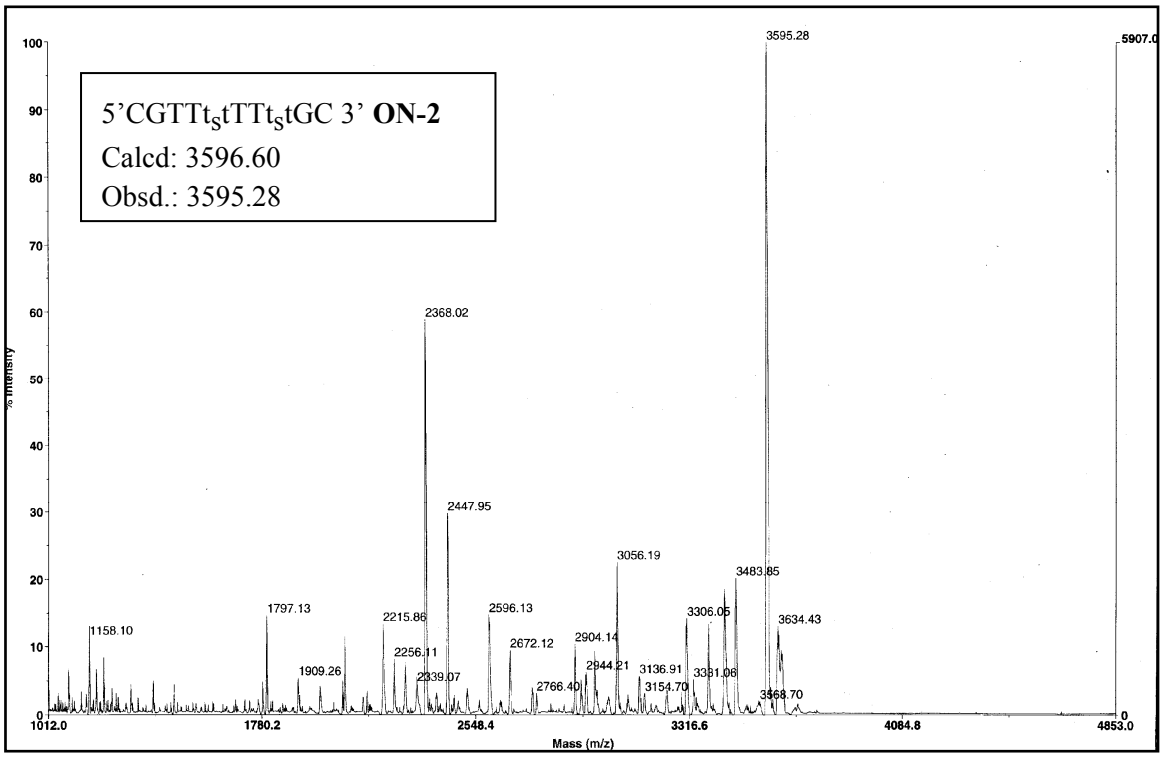

N/A
N/A
ON- 12
mVolts
Minutes
8.626
7.635
14.180

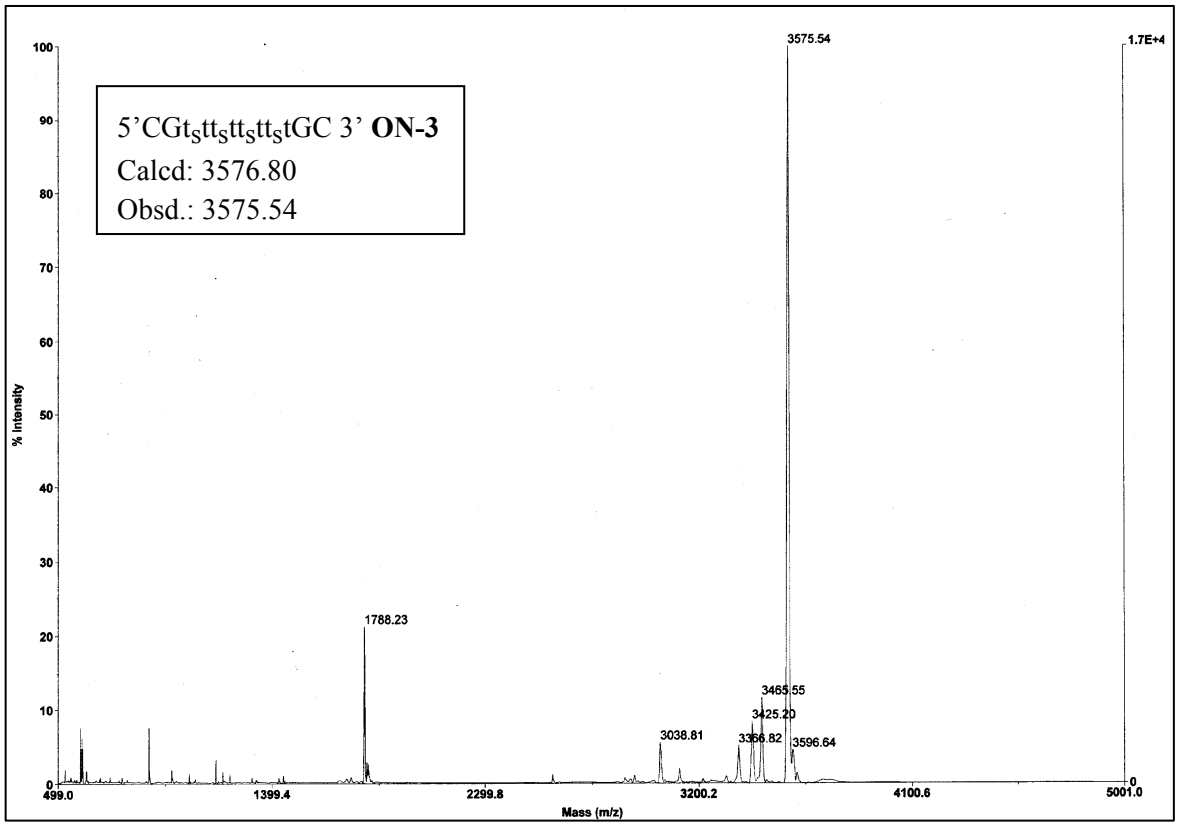

N/A
N/A
ON- 12b
125
100
75
50
25
0
-12
mVolts
8.772
7.720
2.5
5.0
7.5
10.0
12.5
Minutes

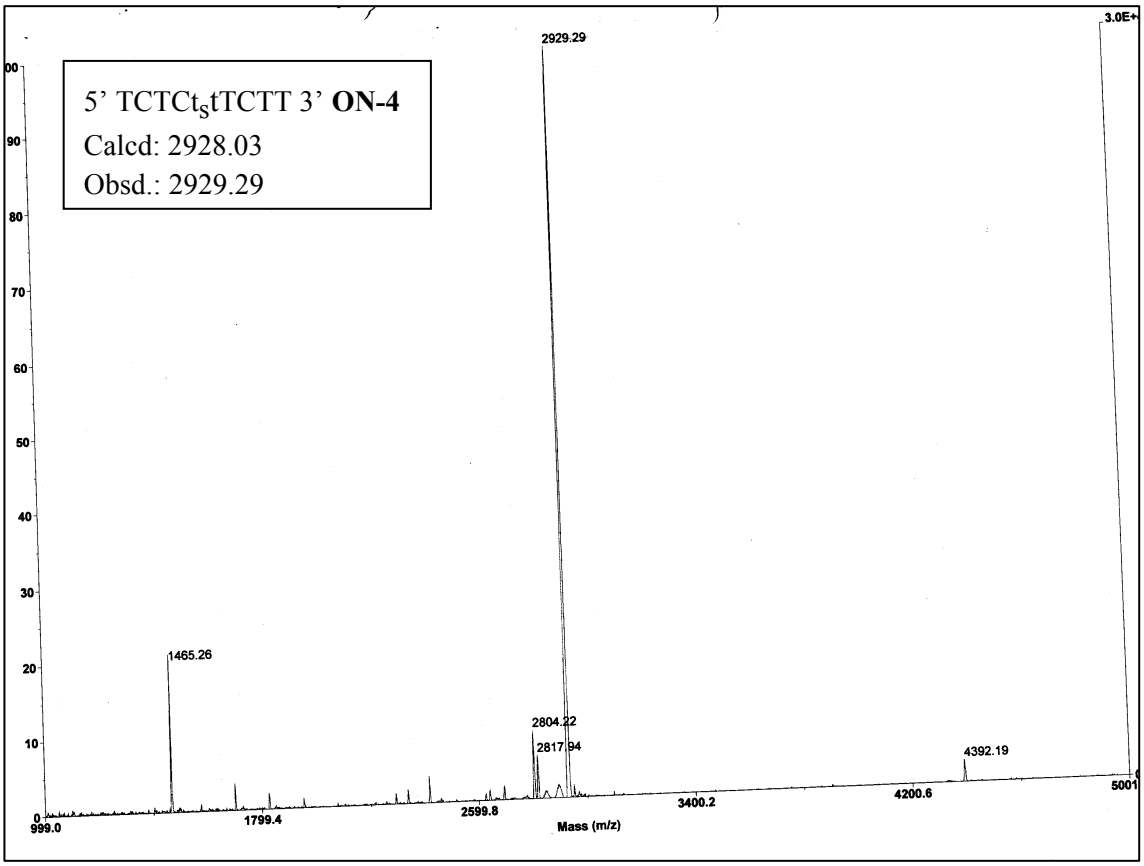

Varian Star #1
N/A
N/A
ON- 13
100
75
50
25
0
-10
mVolts
9.317
8.248
2.5
5.0
7.5
10.0
12.5
Minutes

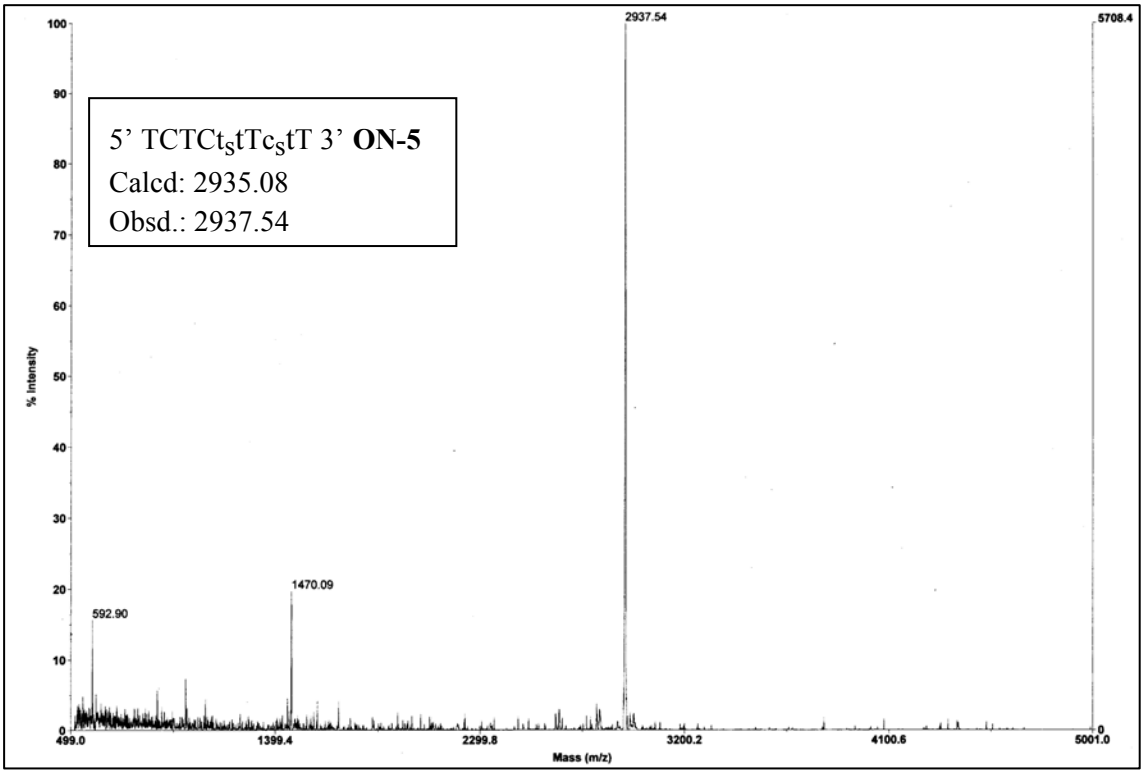

Varian Star #1
N/A
N/A
ON-14
400
300
200
100
0
-41
mVolts
9.673
10.015
5.137
7.962
8.338
8.915
9.251
10.211
10.584
10.955
2.5
5.0
7.5
10.0
12.5
Minutes

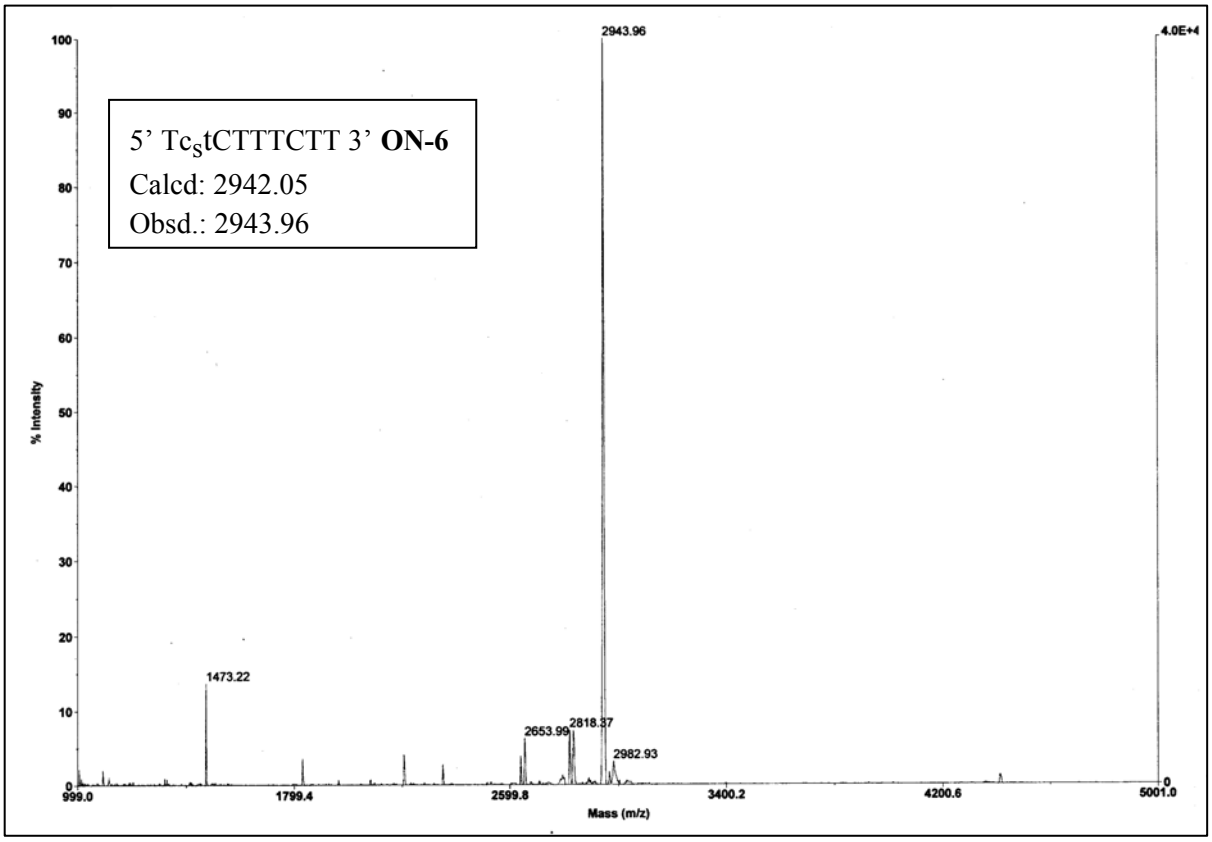

5'CGTTtstTTTTGC 3' **ON-1**
Calcd: 3606.54
Obsd.: 3604.85
% Intensity
Mass (m/z)
3604.85
3328.1
3275.92
2986.78
1802.74
3066.94
3315.79
1216.73
1770.80
2608.88
2658.98
3300.24
3165.54
3453.96
2075.41
2379.53
1747.20
2916.13
3644.88
999.0
1799.4
2599.8
3400.2
4200.6
5001.0

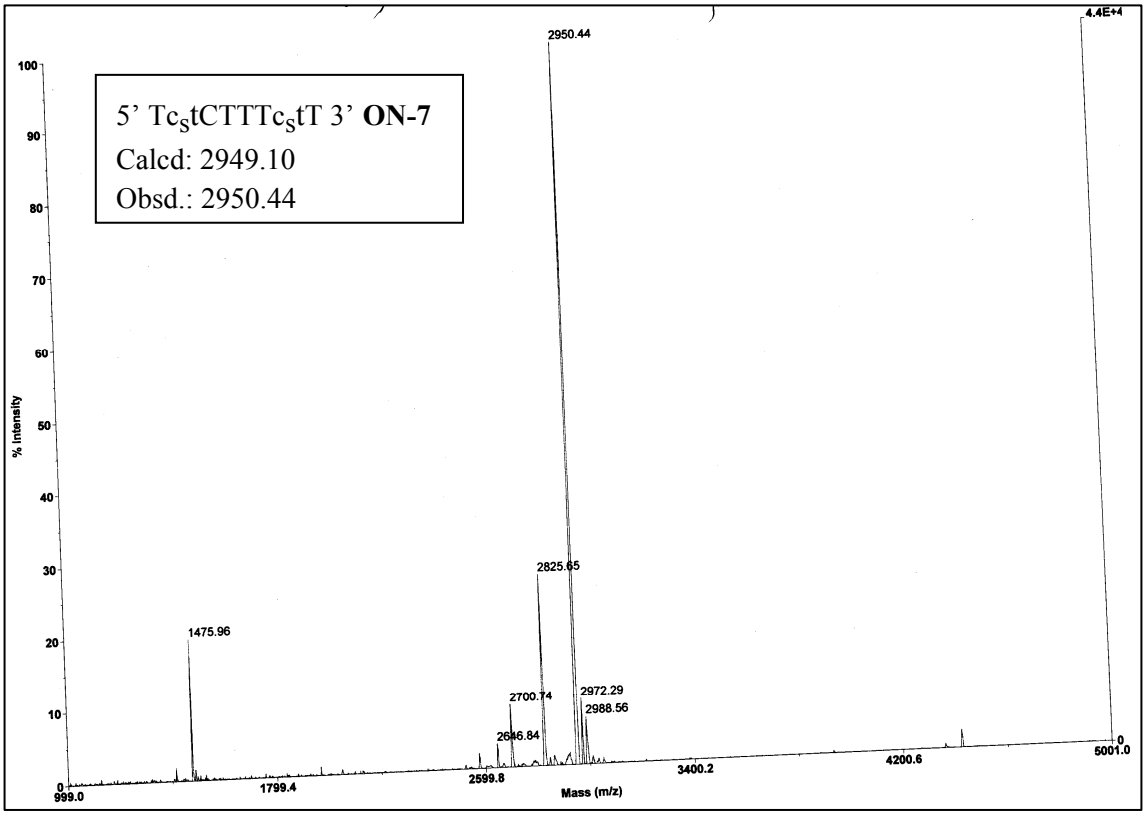

5'CGTTtstTTtstGC 3' **ON-2**
Calcd: 3596.60
Obsd.: 3595.28
% Intensity
Mass (m/z)
3595.28
5907.0
2368.02
2447.95
3056.19
3483.85
1797.13
2215.86
2596.13
3306.05
3634.43
1158.10
2672.12
2904.14
2256.11
1909.26
2339.07
2944.21
3136.91
3331.08
2766.40
3154.70
3368.70
1012.0
1780.2
2548.4
3316.6
4084.8
4853.0

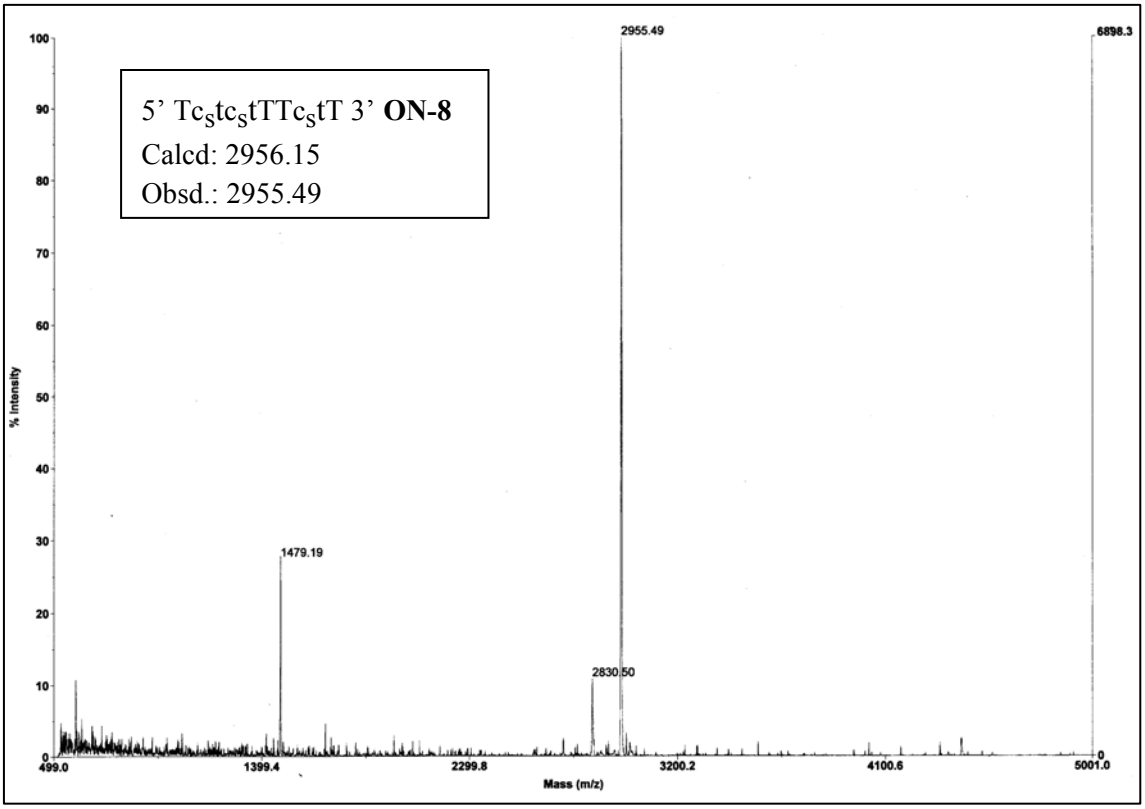

5’CGtstttttttstGC 3’ ON-3
Calcd: 3576.80
Obsd.: 3575.54
% intensity
Mass (m/z)
3575.54
1.7E+4
1788.23
3465.55
3425.20
3038.81
3366.82
3596.64
499.0
1399.4
2299.8
3200.2
4100.6
5001.0

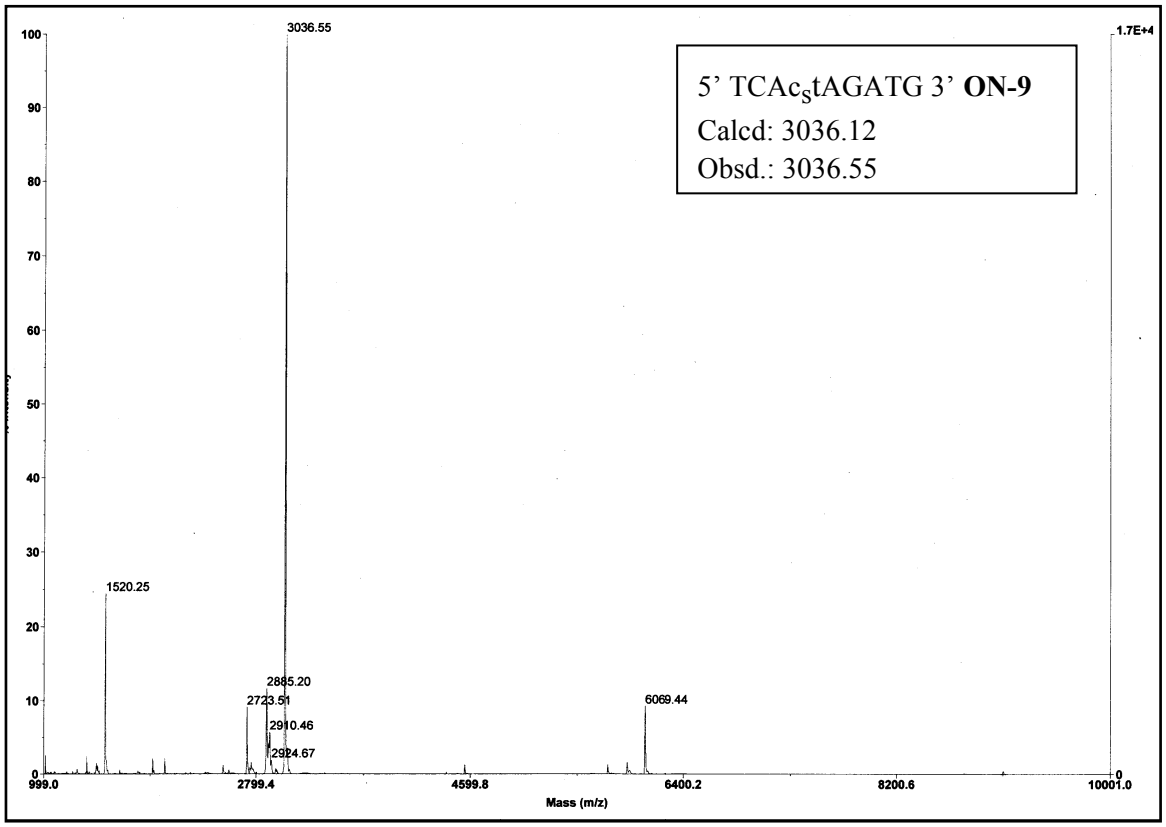

5’ TCTCtstTCTT 3’ ON-4
Calcd: 2928.03
Obsd.: 2929.29
Mass (m/z)
2929.29
3.0E+
1465.26
2804.22
2817.94
4392.19
999.0
1799.4
2599.8
3400.2
4200.6
5001

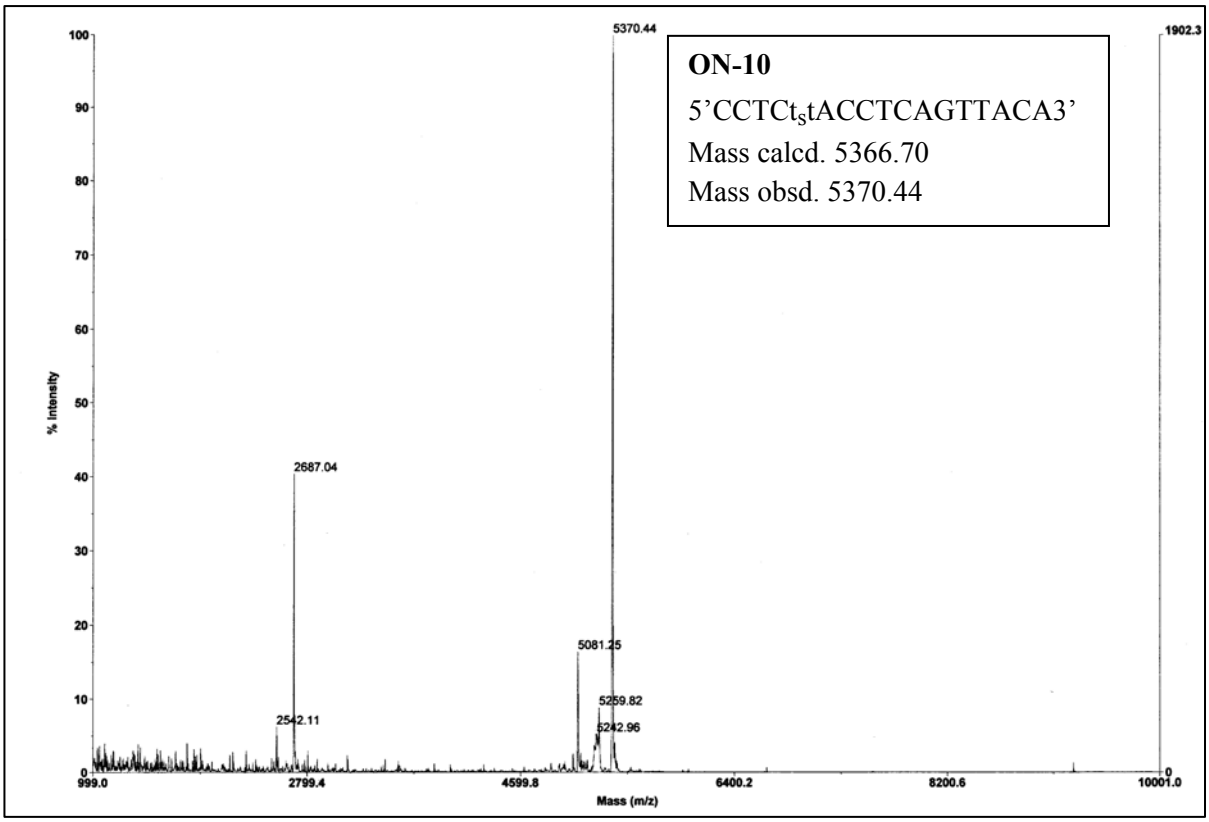

5’ TCTCtstTcstT 3’ ON-5
Calcd: 2935.08
Obsd.: 2937.54
2937.54
5708.4
1470.09
592.90
% Intensity
Mass (m/z)
499.0
1399.4
2299.8
3200.2
4100.6
5001.0

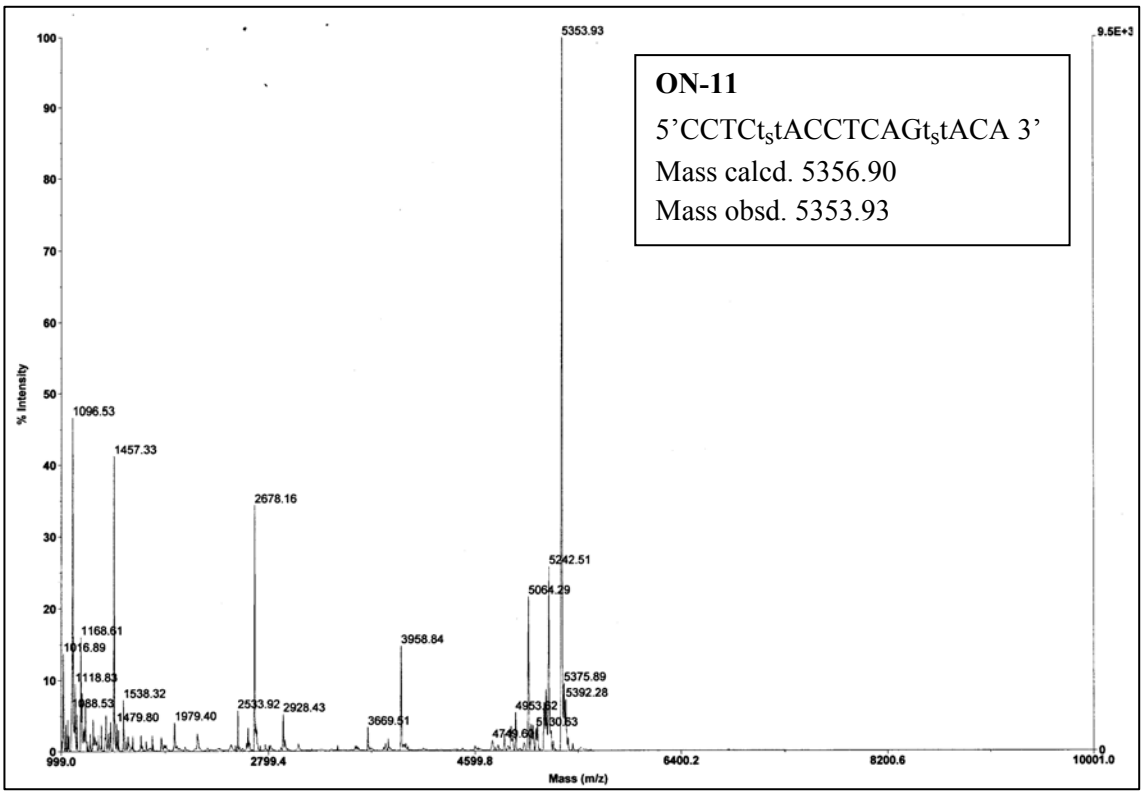

5’ TcstCTTTCTT 3’ ON-6
Calcd: 2942.05
Obsd.: 2943.96
2943.96
4.0E+4
1473.22
2653.99
2818.37
2982.93
% Intensity
Mass (m/z)
999.0
1799.4
2599.8
3400.2
4200.6
5001.0

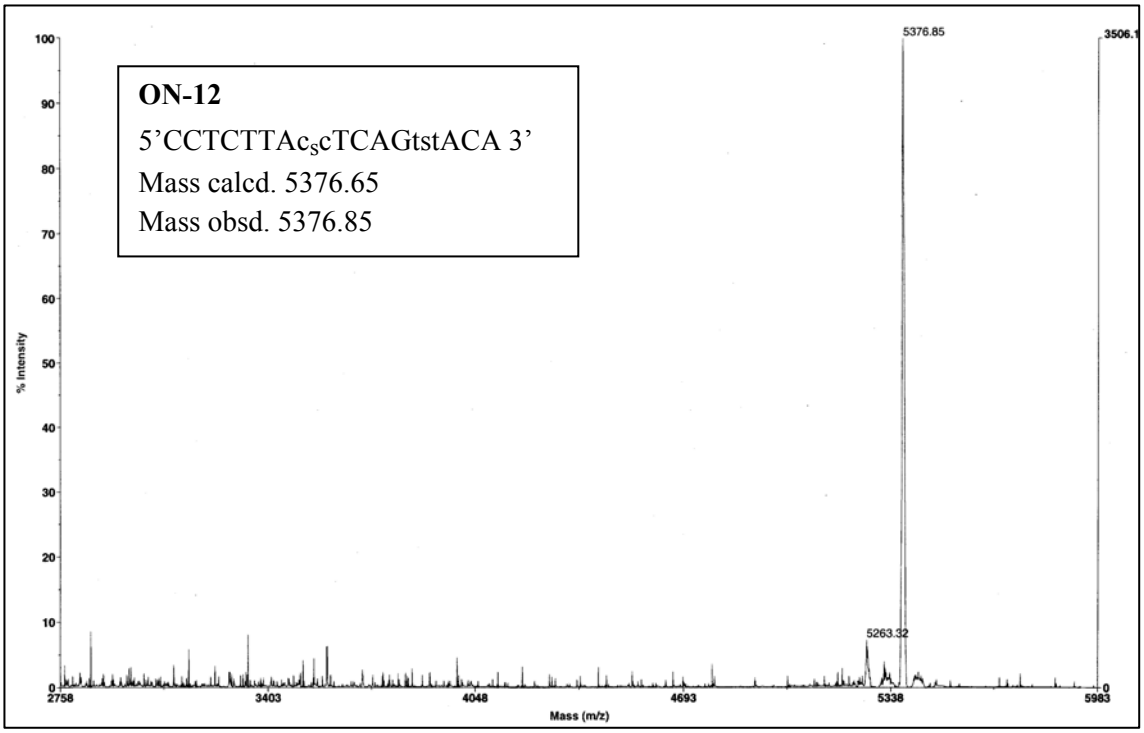

5’ Tc$_S$tCTTTc$_S$tT 3’ **ON-7**
Calcd: 2949.10
Obsd.: 2950.44
2950.44
4.4E+4
2825.65
1475.96
2700.74
2972.29
2988.56
2646.84
% Intensity
Mass (m/z)
999.0
1799.4
2599.8
3400.2
4200.6
5001.0

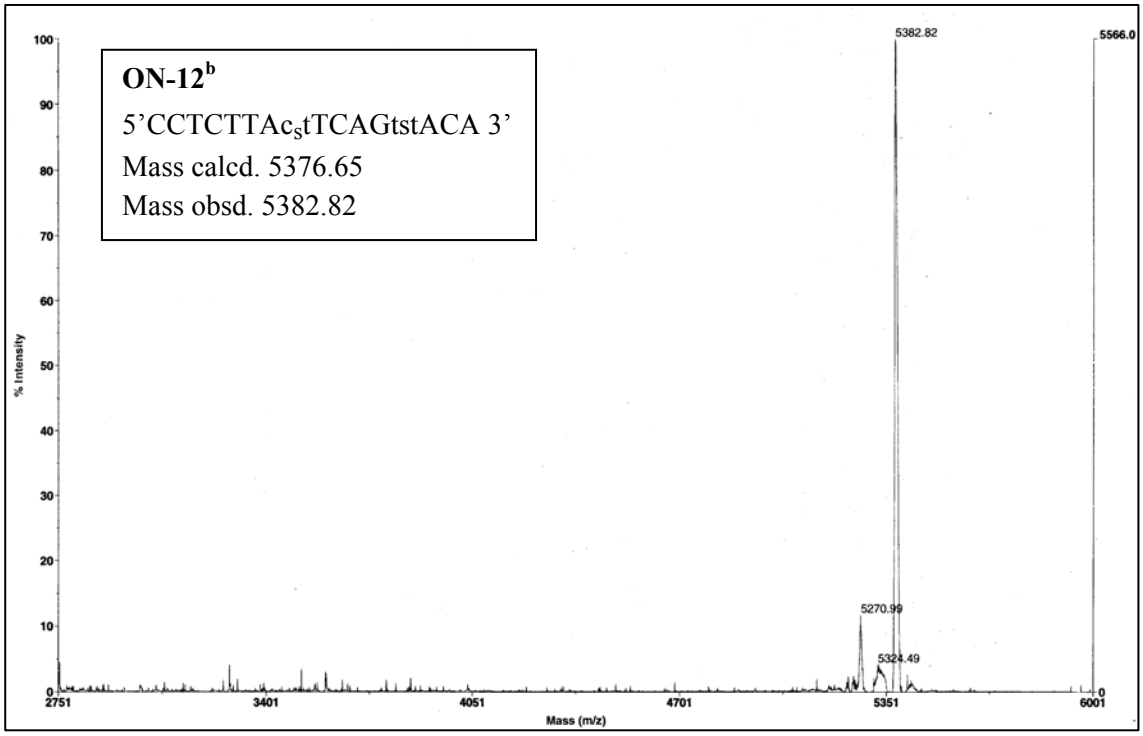

5’ Tc$_S$tc$_S$tTTc$_S$tT 3’ **ON-8**
Calcd: 2956.15
Obsd.: 2955.49
2955.49
6898.3
1479.19
2830.50
% Intensity
Mass (m/z)
499.0
1399.4
2299.8
3200.2
4100.6
5001.0

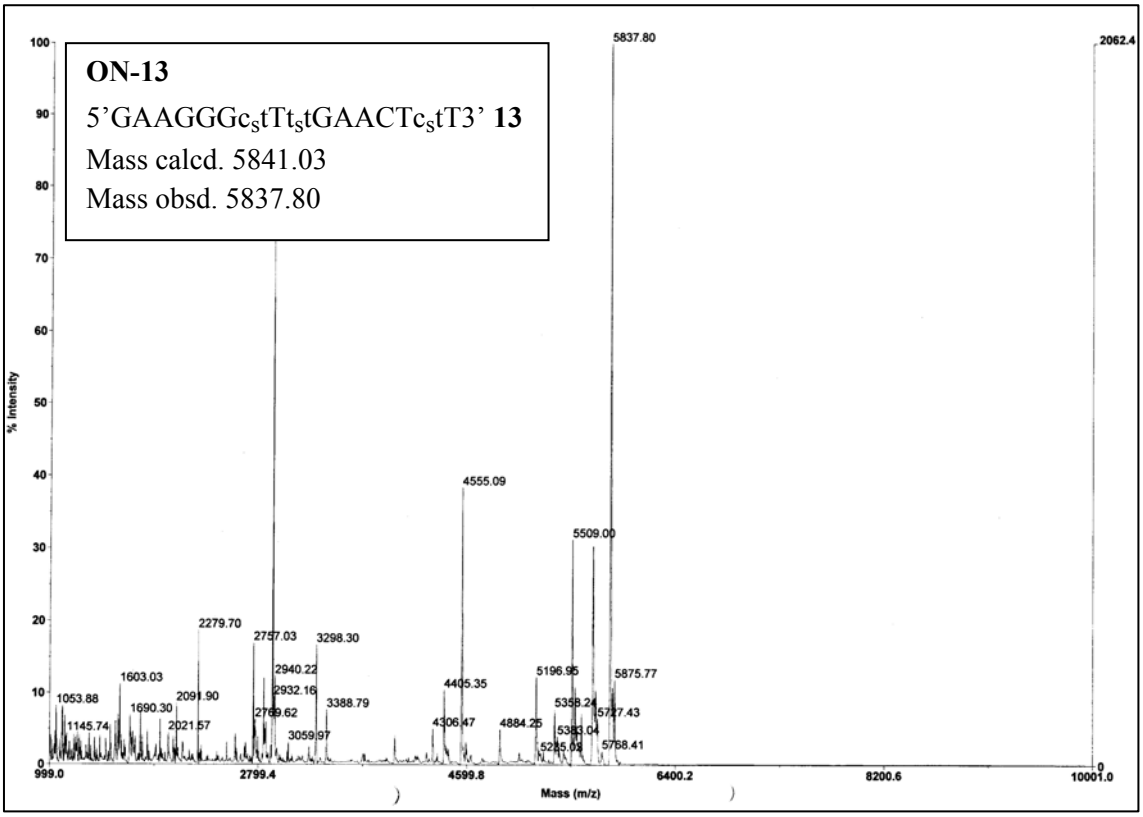

3036.55
1.7E+4
5' TCAc$_s$tAGATG 3' **ON-9**
Calcd: 3036.12
Obsd.: 3036.55
1520.25
2885.20
2723.51
2910.46
2924.67
6069.44
100
90
80
70
60
50
40
30
20
10
0
999.0
2799.4
4599.8
6400.2
8200.6
10001.0
Mass (m/z)

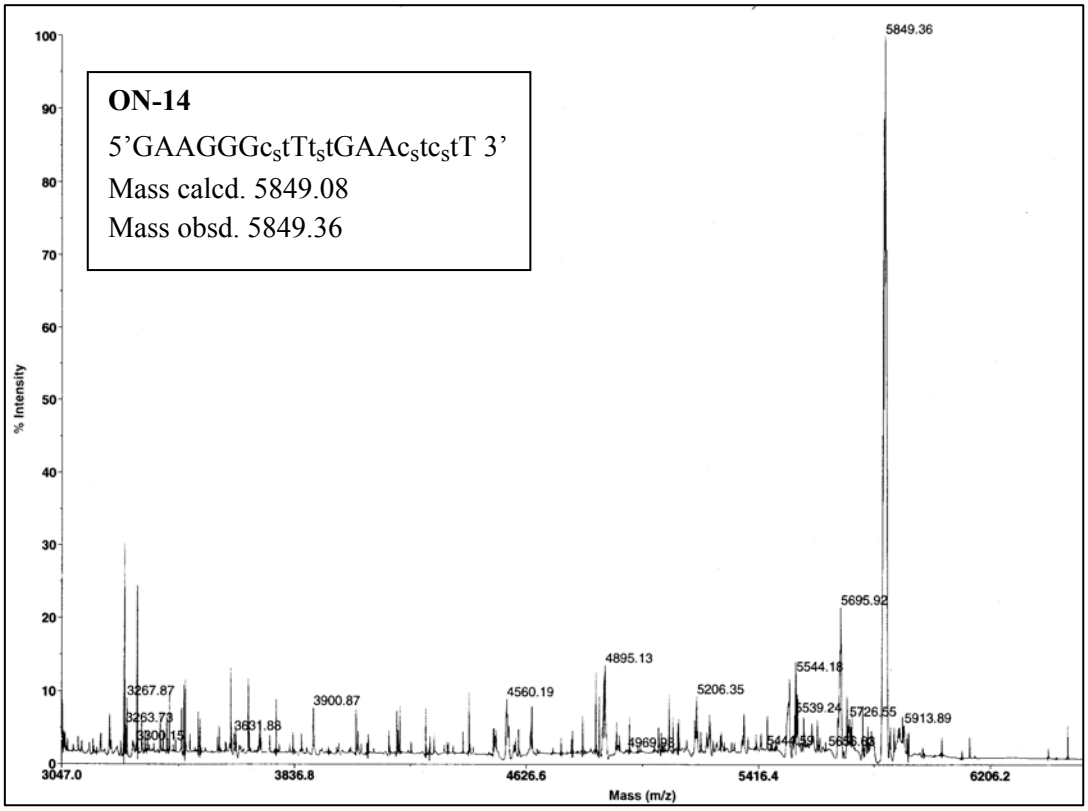

5370.44
1902.3
ON-10
5'CCTCt$_s$tACCTCAGTTACA3'
Mass calcd. 5366.70
Mass obsd. 5370.44
% Intensity
2687.04
5081.25
5259.82
5242.96
2542.11
100
90
80
70
60
50
40
30
20
10
0
999.0
2799.4
4599.8
6400.2
8200.6
10001.0
Mass (m/z)


ON-11
5'CCTCtstACCTCAGtstACA 3'
Mass calcd. 5356.90
Mass obsd. 5353.93
% Intensity
Mass (m/z)
5353.93
1096.53
1457.33
2678.16
5242.51
5064.29
3958.84
5375.89
5392.28
1168.61
1016.89
1118.83
1538.32
1088.53
1479.80
1979.40
2533.92
2928.43
3669.51
4953.62
5130.53
4749.68
999.0
2799.4
4599.8
6400.2
8200.6
10001.0


ON-12
5'CCTCTTAcscTCAGtstACA 3'
Mass calcd. 5376.65
Mass obsd. 5376.85
% Intensity
Mass (m/z)
5376.85
5263.32
2758
3403
4048
4693
5338
5983


ON-12[b]
5'CCTCTTAcstTCAGtstACA 3'
Mass calcd. 5376.65
Mass obsd. 5382.82
% Intensity
Mass (m/z)
5382.82
5566.0
5270.99
5324.49
2751
3401
4051
4701
5351
6001


ON-13
5'GAAGGGcstTtstGAACTcstT3' 13
Mass calcd. 5841.03
Mass obsd. 5837.80
% Intensity
Mass (m/z)
5837.80
2062.4
4555.09
5509.00
2279.70
2757.03
3298.30
2940.22
2932.16
1603.03
1053.88
2091.90
1690.30
3388.79
2769.62
1145.74
2021.57
3059.97
4405.35
5196.95
5875.77
5358.24
5727.43
4306.47
4884.25
5383.04
5768.41
999.0
2799.4
4599.8
6400.2
8200.6
10001.0


ON-14
5'GAAGGGcsstTtststGAAcstcsstT 3'
Mass calcd. 5849.08
Mass obsd. 5849.36
% Intensity
Mass (m/z)
5849.36
5695.92
4895.13
5544.18
5206.35
4560.19
3900.87
3267.87
3263.73
3631.88
5539.24
5726.55
5913.89
3047.0
3836.8
4626.6
5416.4
6206.2

3.1.9 References

1 Kurreck, J. Antisense technologies improvement trough novel chemical modifications *Eur. J. Biochem.* **2003**, *270*, 1628-1644

2 a) Micklefield, J. Backbone modification of nucleic acids: synthesis, structure and therapeutic applications *Curr. Med. Chem.*, **2001**, *8*, 1157. b) Turner, J. J.; fabini, M.; Arzumanov, A.; Ivanova, G.; Gait, M. J. Targeting the HIV-1 RNA leader sequence with synthetic oligonucleotides and siRNA: chemistry and cell delivery *Biochim.biophys. Acta* **2006**, *1758*, 290. c) Kumar, V. A.; Ganesh, K. N. Structure-Editing of nucleic acids for selective targeting of RNA *Curr.Top. Med. Chem.* **2007**, *7*, 715.

3 Kumar, V. A.; Ganesh, K. N. Conformationally constrained PNA analogues: structural evaluation toward DNA/RNA binding selectivity *Acc. Chem. Res.* **2005**, *38*, 404-413.

4 (a) Pallan, P. S.; von Matt, P.; Wilds, C. J.; Altmann, K.-H.; Egli, M. RNA-binding Affinities and Crystal Structure of Oligonucleotides Containg Five –Atom Amide –Based Backbone Structures *Biochemistry***, 2006,** *45*, 8048-8057. (b) Wilds, C. J.; Minasov, G.; von Matt, P.; Altmann K.-H.; Egli, M. Studies of a chemically modified oligodeoxynucleotide containing a 5-atom amide backbone which exhibits improved binding to RNA *Nucleosides, Nucleosides and Nucleic Acids,* **2001***, 20*, 991-994. (c) Govindaraju, T.; Kumar, V. A. Backbone-extended pyrrolidine peptide nucleic acids (*bep*PNA): design, synthesis and DNA/RNA binding studies *Chem. Commun*., **2005**, 495–497 (d) Govindaraju, T.; Kumar, V. A. Backbone extended pyrrolidine PNA (*bep*PN(A): a chiral PNA for selective RNA recognition *Tetrahedron*, **2006**, *60*, 2321–2330. (e) Gogoi, K.; Kumar, V. A. Sugar-thioacetamide backbone in oigodeoxyribonucleosides for specific recognartion of nucleic acids *Chem. Commun*. **2008**, 706-708. (f) Worthington, R. J.; Bell, N. M.; Wong, R.; Micklefield, J. RNA-selective cross-pairing of backbone–extended pyrrolidine-amide oligonucleotide mimics (bePOMs) *Org. Biomol. Chem.* **2008,** *6,* 92-103. g) . Gogoi, K.; Gunjal, A.D.; Phalgune, U. D.; Kumar, V. A. Synthesis and RNA binding selectivity of oligonucleotides modified with five atom thioacetamido Nucleic Acid backone structures *Org. Lett*. **2007**, *9*, 2697-2700.

5 Thibaudeau, C.; Plavec, J.; Garg, N.; Papchikhin, A.; Chattopadhyaya, J. How does the electronegativity of the substituent dictate the strength of the guache effect ? *J.Am.Chem. Soc.* **1994**, *116*, 4038-4043

6 Sfroza, S.; Corradini, R.; Dossena, A.; Marchelli, R. DNA binding of a D-lysine-based chiral PNA: Direction control and mismatch recognition. *Eur. J. Org.chem*.**2000**, 2905-2913

7 a) Matt, P. V.; De Mesmaeker, A.; Pieles, U.; Zurcher, W.; Altmann, K-H. 2'-Deoxyribo-PNAs: A Structurally Novel Class of polyamide Nucleic Acids with Good RNA and DNA Binding Affinity *Tetrahedron Lett* **1999**, *40*, 2899-2902. b) Amide linked oligomers Stork, G.; Zhang, C.; Gryaznov, S.; Schultz, R. Modified oligonucleotides. Effect of 4- *vs* 5-atom chimeric internucleoside linkages on duplex stability *Tetrahedron Lett.* **1995** *36*, 6387-6390. c) De Mesmaeker, A.; Jouanno,C.; Wolf, R. M.; Wendeborn, S. Replacement of the phosphodiester linkage in oligonucleotides by an amide: effect of backbone length on duplex stability with RNA complement *Bioorg. Med. Chem. Lett.* **1997**, *7*, 447-452. d) De Napoli, L., Iadonisi, A., Montesarchio, D., Varra, M., Piccialli, G. Synthesis of thymidine dimers containing a new internucleosidic amide linkage and their incorporation into oligodeoxyribonucleotides, *Bioorg. Med. Chem. Lett.* **1995,** *5*, 1647-1652. e) De

Mesmaekar, A.; Waldner, A.; Lebreton, J.; Hoffmann, P.; Fritsch, V.; Wolf, R. M.; Freier, S.M. Amides as a New Type of backbone Modification in Oligonucleotides *Angew. Chem. Int. Ed. Engl*. **1994**, *33*, 226-229.

8 (a) Ding, D., Gryaznov, S.M., Lloyd, D.H., Chandrasekaran, S., Yao, S.,Ratmeyer, L., Pan, Y.; Wilson, W.D. An oligodeoxyribonucleotides N3'→P5' phosphoramidate duplex forms an A-type helix in solution, *Nucleic Acids Res.,* **1996**, *24*, 354-360. (b) Ding, D., Gryaznov, S. M.; Wilson, W. D. NMR solution structure of the N3'→P5' phosphoramidate duplex d(CGCGAATTCGCG) by iterative relaxation matrix approach, *Biochemistry*, **1998**, *37*, 12083-12093.

9 Gryaznov, S. M., Lloyd, D.H., Chen, J.-K., Schultz, R.G., DeDionisio, L.A., Ratmeyer, L. and Wilson, W.D. Oligonucleotide N3'→P5' phosphoramidates, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5798-5802. b) Gryazonv, S. M. Oligonucleotides N3'→P5' phospharamidates as potemtial therapeutic agents *Chemistry & Biodiversity* **2010**, *7*, 477-493

10 Gryazonv, S. M. Oligonucleotides N3'→P5' phospharamidates as potemtial therapeutic agents *Chemistry & Biodiversity* **2010**, *7*, 477-493

11 Prakash, T. P.; Manoharan, M.; Kawasaki, A. M.; Fraser, A. S.; Lesnik, E. A.; Sioufi, N.; Leeds, J. M.; Teplova M.; Egli, M. 2'-O-[2-(methylthio)ethyl]-modified oligonucleotides:An analogue of 2'-O-[2-(methoxy)-ehtyl]-modified oligonucleotides with improved protein binding properties *Biochemistry* **2002**, *41,* 11642-11648.

12 Pongracz, K.; Gryaznov, S. Oligonucleotide N3'→P5' thiophoramidates : synthesis and properties *Tetrahedron Lett*. **1999**, *40*, 7661.

13 Kawai, S. H.,; Wang, D.; Giannaris, P. A.; Damha, M.; Just, G. Solid phase synthesis and hybridization properties of DNA containing sulfide-linked dinucleoside. *Nucleic Acids Res.* **1993**, *21*, 1473-1479.

14 Meng, B.; Kawai, S. H.; Wang, D.; Just, G.; Giannaris, P. A.; Damha, M. A sulfide linked oligonucleotide analogue with selective hybridization properties. *Angew. Chem Int. Ed. Engl*. **1993**, *32*, 729-731.

15 a) Rosemeyer, H.; Seela, F. 72.1-(2'-deoxy-β-D-xylofuranosyl) thymine building Blocks for solid phase synthesis and properties of oligo(2'-deoxynucleotides) *Helv. Chim. Acta* **1991**, *74*, 748–760. b) Seela, F.; Wörner, K.; Rosemeyer, H. 83. 1-(2'-deoxy-β-xylofuranosyl)cytosine: Base pairing of oligonucleotides with a configurationally altered sugar-phospahte backbone *Helv. Chim. Acta* **1994**, *77*, 883–896. c) Rosemeyer, H.; Krecmerova, M.; Seela, F. 192.9-(2'-deoxy-β-D-xylofuranosyl) adenine building Blocks for solid phase synthesis and properties of oligo (2'-deoxynucleotides) *Helv. Chim. Acta* **1991**, *74*, 2054–2067. d) Seela, F.; Heckel, M.; Rosemeyer, H.122. Xylose-DNA containg the four natural bases *Helv. Chim. Acta* **1996**, *79*, 1451–1461. e) Poopeiko, N. E.; Juhl, M.; Vester, B.; Sorensen, M. D.; Wengel, J. *Xylo*-configured oligonucleotides (XNA, xylonucleic acids): synthesis of conformationally restricted derivatives and hybridization towards DNA and RNA complements Bioorg. Med. Chem. Lett. 2003, 13, 2285-2290. f) Babu, B. R.; Raunak; Poopeiko, N. E.; Juhl, M.; Bond, A.; Parmar, V. S.; Wengel, J. XNA(xylo Nucleic acid): Asummary and new derivatives *Eur. J. Org. Chem.* **2005**, 2297-2321. g) Ramaswamy, A.; Froeyen, M.; Herdewjin, P.; Ceulemans, A. Helical structure of xylose-DNA *J. Am. Chem. Soc.* **2010**, *132*, 587-595.

16 Mastuda, A.; Watanabe, K. A.; Fox, J. J. Nucleosides. 115. Reaction of 3'-*O*-mesylthymidine. Formation of 1-(3-azido-2,3-dideoxy-*β*-D-thero-pentofuranosyl)thymine and its conversion into 6,3'-imino-1-(2,3-dideoxy--*β*-D-thero-pentofuranosyl)thymine *J. Org. Chem.* **1980**, *45*, 3274-3278

17 Saneyoshi, M.; fujii, T.; Kawagchi, T.; Aai, K.; Kimura, S. *Nucleic Acid Chemisry* Wiley New York, *1991*, 4, 67-72.

18 Sundaralingam, M. *Biopolymers* **1969**, *7*, 821

19 a) Altona, C.; Sundaralingam, M. Coformational-analysis of sugar ring in nucleosides and nucleotides-improved method for interpretation of proton magnetic–resonance coupling-constants *J. Am. Chem. Soc.* **1973**, *95*, 2333. b) Altona, C.; Sundaralingam, M. Coformational-analysis of sugar ring in nucleosides and nucleotides-new description using concept of pseudorotation *J. Am. Chem. Soc.* **1972**, *94*, 8205.

20 a) van Wijk, J.; Haasnoot, C.A. G.; de Leeuw, F. A. A. M.; Huckriede, B. D.; Westra Hoeckzema, A.; Altona, C. PSEUROT 6.2 1993, PSEUROT 6.3 1999; Leiden Institute odf Chemistry, Leiden University. b) de Leeuw, F. A. A. M. ; Altona, C. J. computer-assisted pseudorotation analysis of 5-Membered rings by means of proton spin spin coupling –constants- program PSEUROT Comput. Chem. **1983** *4*, 428

21 Rinkel, L. J.; Altona, C. Conformational analysis of the deoxyribofuranose ring in DNA by means of sum of proton-proton coupling constants: a graphical method *J. Biomol. Sruct. Dyn*. **1987**, *4*, 621

22 Plavec, J.; Tong, W.; Chattopadhyaya, J. How do the gauche and anomeric effects drive the pseudorotational equilibrium of the pentafuranose moiety of nucleosides *J. Am. Chem. Soc.* **1993**, *115*, 9734

23 a) Plavec, J.; Garg, N.; Chattopadhyaya, J. How does the steric effect drive the sugar conformation in the 3'-C-branched nucleosides. *J. Chem. Soc. , Chem. Commun.* , **1993**, *12*, 1011. b) Plavec, J; Koole, L. H.; Chattopadhyaya, J.Structural –analysis of 2'-3'-dideoxyinosine, 2'-3'-dideoxyadenosine, 2'-3'-dideoxyguanosine and 2'-3'-dideoxycytidine by 500-MHz H-1-NMR spectroscopy and abinitio molecular-orbital calculations *J. Biochem. Biophys. Methods* **1992**, *25*, 253. c) Plavec, J.; Thibaudeau, C.; Viswanadham, G.; Sund, C.; Chattopadhyaya, J. How does the 3'-phosphate drive the sugar conformation in DNA *J. Chem. Soc. , Chem. Commun*., **1994**, 718-783. d) Glemarec, C.; Reynolds, R. C.; Crooks, P. A.; Maddry, J. A.; Akhtar, M. S.; Montgomery, J. A.; Secrist III, J. A.; Chattopadhyaya, J. Conformational studies of thymidine dimers containing sulfonate and sulfonamide linkages by NMR-spectroscopy *Tetrahedron*, **1993**, *49*, 2287. e) Plavec, J.; Koole, L.H.; Sandstrom, A.; Chattopadhyaya, J. Structural studies of anti-HIV 3'-alpha-flurothymidine and 3'-alpha-aziothymidine by 500MHZ H-1-NMR spectoscopy and molecular mechanics (MM2) calculations *Tetraherdon* **1991**, *47*, 7363. f) Glemarec, C.; Nyllas, A.; Sund, C.; Chattopadhyaya, J. A comparative conformational study of thymidylyl (3'-]5')-5'-thio-5'-thio-5'-deoxythymidine and thymidinylacetamido-[3'(O)-]-5'-deoxythymidine *J. Biochem. Biophys. Methods* **1990**, *12*, 311

24 Rosemeyer, H.; Strodtholz, I.; Seela, F. 2'-deoxy-β-D-xylothymidinyl-(3'-5')-2'-deoxy-β-D-xylothymidylate: stereochemical course of dinucleoside phosphate formation and conformational properties *Bioorg. Med. Chem. Lett.***1992**, *2*, 1201-1206.

25 Ph.D Thesis of Mr. Khirud Gogoi

26 Analysis of circular dichroism spectra Methods of enzymology vol .210,1992, 426-447

27 a) Gait, M. J. Oligonucleotide synthesis: A practical approach. IRL Press Oxford, UK 217 (b) Agrawal, S. PROTOCOLS for oligonucleotides and analogs Synthesis and Properties Methods in Molecular Biology. (ed): vol.20 Totowa, NJ,. Humana Press, Inc.

28 Kang, S-H.; Cho, M-J.; Kole, R. Up-regulation of luciferase gene expression with antisense oligonucleotides : Implications and Applications in functional assay development *Biochemistry*, **1998**, 37, 6235-6239

29 Gryaznov, S.; Skorski, T.; Cuco, C.; Nieborowska-Skorska, M.; Chiu, C. Y.; Llyod, D.; Chen, J.; Koziolkiewicz, M.; Calabretta, B. Oligonucleotide N3'→P5' as phosphoramidates antisense agents *Nucleic Acids Research* **1996**, *24*, 1508.

30 Petersen, M.; Wengel, J. *Trends Biotechnol.* **2003**, *21*, 74.

31 a) Nawrot, B.; Boczkowska, M.; Wojcik, M.; Sochaki, M.; Kazmierski, S.; Stec, W. J. *Nucleic Acids Res.* **1998**, *26*, 2650. b) Gryaznow, S.; Chen, J.-K. *J. Am. Chem. Soc.* **1994**, *116*, 3143

32 Prakash, T. P.; Kraynack, B.; Baker, B. F.; Swayze, E. E.; Bhat, B. *Bioorg. Med. Chem. Lett*. **2006**, *16*, 3238.

33 (a) Giannaris, P. A.; Damha, M. J. *Nucleic Acids Res.* **1993**, *21*, 4742. (b) Wasner, M.; Arion, D.; Borkow, G.; Noronha, A.; Uddin, A. H.; Parnaik, M. A.; Damha, M. J. *Biochemistry* **1998**, *37*, 7478.

Chapter 3

Section II: Synthesis of Modified PNA sequences with *iso*-thioacetamidonucleic acids monomers (*iso*-TANA) and their Biophysical studies

3.2.1 Introduction

Previous work from our laboratory described the synthesis of thioacetomido linked dimers and their incorporation into DNA ONs. DNA oligomers with TANA dimers exhibited preferential binding with RNA over DNA. The homo-oligomeric pyrimidine TANA oligomers showed preferential binding with complementary RNA and the binding efficiency was found to be as good as PNA (Figure 1).[1] PNA oligomers incorporating modified TANA thymine monomers also showed binding selectivity with complementary RNA over DNA. As described in the previous section, the stereochemistry of the sugar ring substituents, when incorporated in PNA may also affect the binding preferences with DNA and RNA.

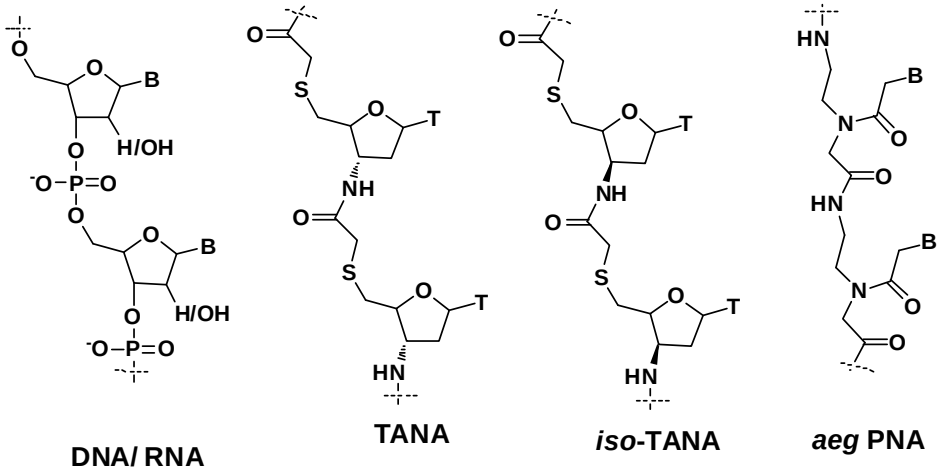


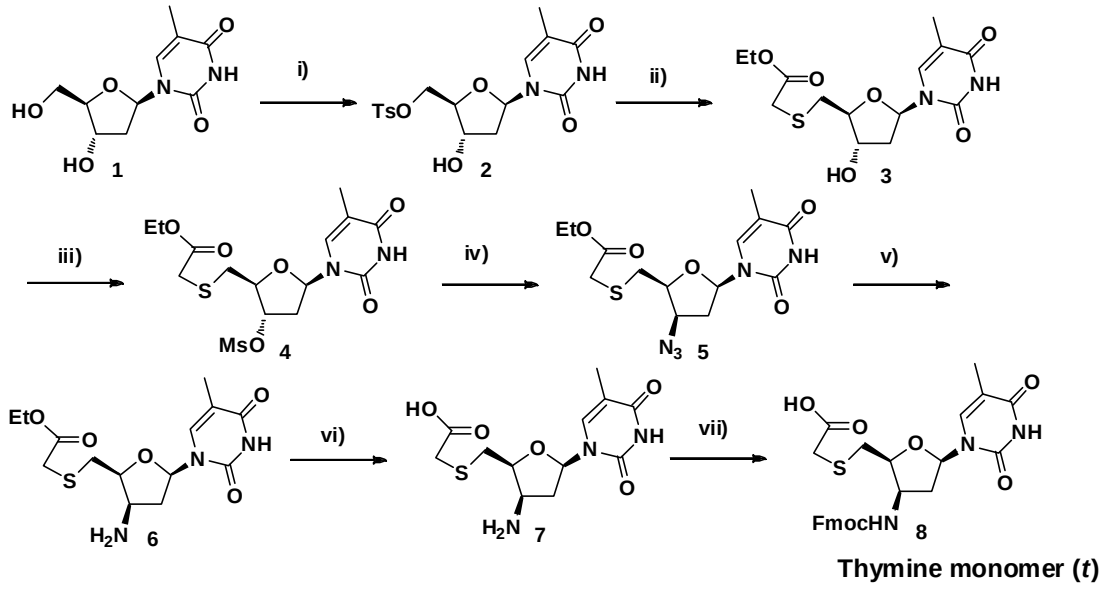
Figure 1: Structure of DNA, PNA and *iso*-TANA

In this section we describe the synthesis of protected *iso*-TANA thymine and 5-methyl cytosine monomers amenable for their incorporation in PNA sequences. The 3'-amino terminus was protected with Fmoc group with a free acid terminus at 5'-end of the sugar. The synthesis of homogeneous backbone oligomers of *iso*-TANA thymine sequence (Figure 1) using these monomeric units, the chimeric PNA oligomers with *iso*-TANA thymine units incorporated in aegPNA at selected positions and their binding preferences with DNA and RNA is presented in this section. The results of UV-melting

experiments will give information about effect of change in stereochemistry in the sugar ring on binding preferences with complementary DNA and RNA.

3.2.2 Synthesis of *iso*-TANA monomers results and discussion

3.2.2a Synthesis of *iso*-TANA thymine monomer

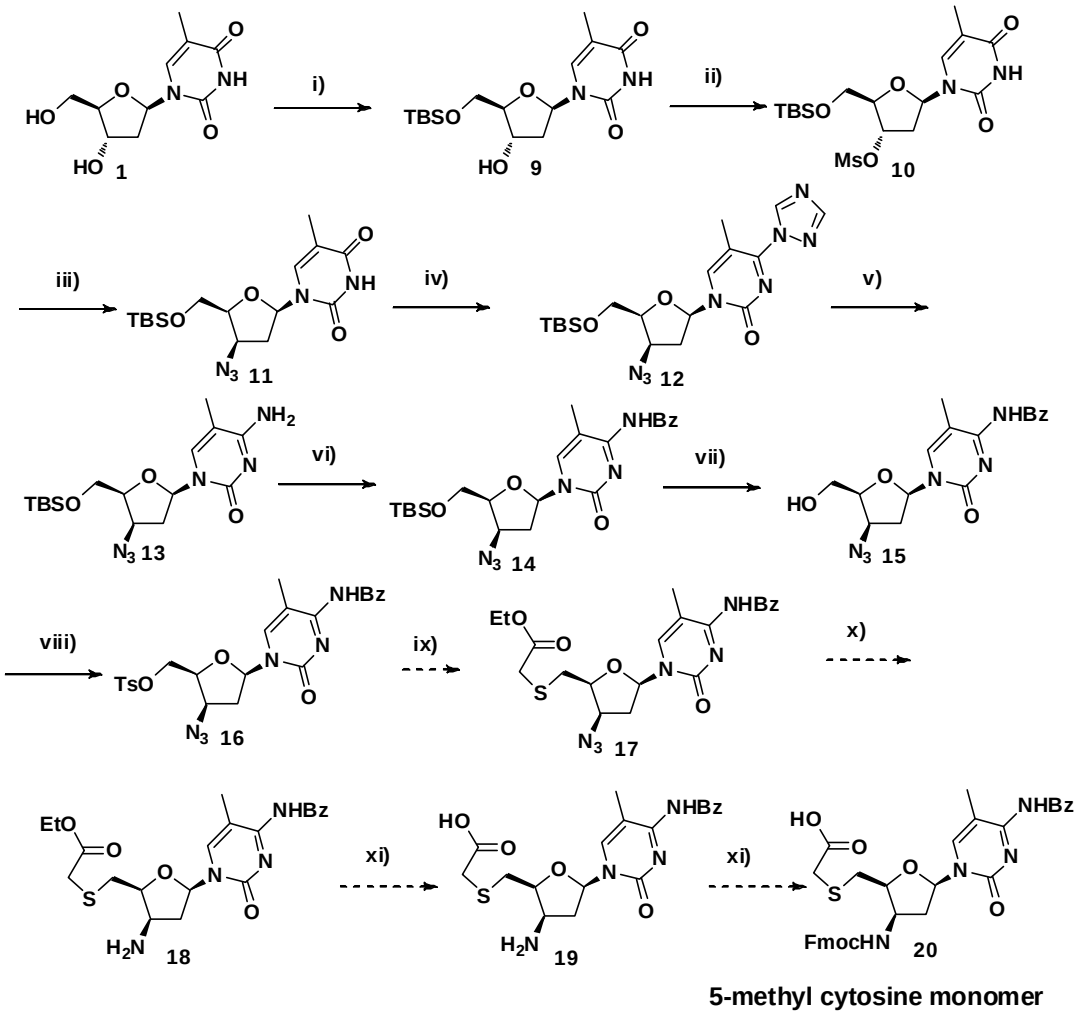


Scheme 1: Reagents and conditions i) TsCl/ Py ,60%, ii) ethyl mercapto acetic acid/ NaH/ DMF, 87%, iii) MsCl/ Et_3N -DCM, 90%, iv) NaN_3/ DMF, 60 ^oC 47%, v) H_2S/ pyridine, Et_3N, 82%, vi) 1N LiOH /MeOH, 78%, vii) Fmoc-succinimide, $NaHCO_3$ in 1,4-dioxane/water, 64 %

The natural nucleoside thymidine was used as starting material for the synthesis of *iso*-TANA thymine monomer. 5'-OH of thymidine was selectively monotosylated to give compound **2**. This selectively was achieved by dilution of reaction mixture with pyridine and slow addition of tosyl chloride at 0 °C. Tosyl derivative **2** was converted to mercapto derivative **3** by nucleophilic substitution of tosyl group with *in-situ* generated anion of ethyl mercapto acetic acid by NaH. Mesylation of 3'hydroxy of compound **3** gave mesyl derivative **4**. To synthesize the azido derivative **5** with inversion of configuration at 3' of thymidine derivative, 3'-*O*-mesyl derivative was subjected for nucleophilic substitution with sodium azide. The moderate yield of this nucleophilic substitution reaction may be due to the formation of other side products due to interference of thymine nucleobase.[2] The azido functional group was characterized by IR band at 2100 cm^{-1}. The conversion of azide group in **5** to amino group could not be achieved by usual hydrogenation conditions, as sulphur containing compounds are known to poison the catalyst. Thus reduction of azido group was carried out with hydrogen transfer reaction using H_2S-pyridine complex to give amino

compound **6**. The ester group in compound **6** was hydrolysed to give amino acid **7**. The synthesis of *iso*-TANA thymine monomer was completed with Fmoc protection of amino group of amino acid **7**. This Fmoc protected monomer **8** was used in solid phase synthesis of homooligomers of *iso*-TANA and aegPNA-*iso*-TANA chimeric oligomers.

3.2.2b Synthesis of *iso*-TANA 5-methyl cytosine monomer

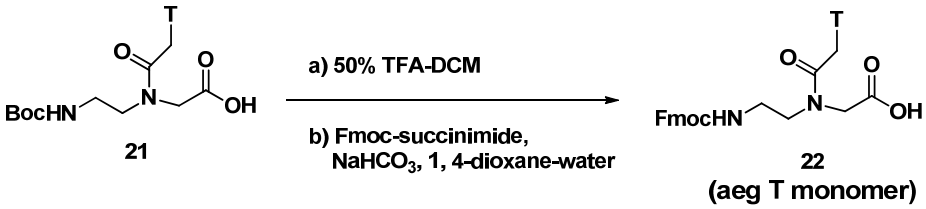


Scheme 2: Reagents and conditions , i) TBDMSCl/ imidazole/ DMF, 80%, ii) MsCl/ Et_3N- DCM 88%, iii) NaN_3/DMF, 60°C, 48 h, 47%, iv) 1,2,4-triazole- $POCl_3/Et_3N$, MeCN, 88%, v) conc. NH_3/1-4 dioxane, 90%, vi) BzCl/pyridine , 78%, vii) 1N TBAF in THF, 90%, viii) TsCl/ Pyridine, 80%, ix) ethyl mercapto acetic acid/NaH, DMF, x) H_2S/pyridine, xi) 1N NaOH/THF, 0^{o}C, xii) Fmoc-succinimide, $NaHCO_3$, 1,4-dioxane/ water

Thymidine was used as starting material for the synthesis of *iso*-TANA 5-methyl cytosine monomer. Selective silyl protection 5'-OH of thymidine was carried out to give 5'-OTBS thymidine **9**. The secondary hydroxyl of 3' was converted to mesyl derivative. 3'O-mesyl group was displaced with NaN_3 to give azido compound **11**. The conversion from mesyl to azide derivative was found to be very sluggish. Compound **11** was transformed into triazolide derivative **12**. The conversion of thymine nucleobase

to 5-methyl cytosine was achieved by treatment of triaole derivative with conc. NH_3 solution. Exocyclic amino group of 5-methyl cytosine was protected with benzoyl to give protected product **14**. Silyl group was deprotected with 1N TBAF to give compound 5'-free hydroxyl compound **15**. 5'-OH group was tosylated and tosyl was substituted with mercapto group. This substitution was carried out by reaction of tosyl derivative with *in-situ* generated anion of ethylmercapto acetic acid. acid with NaH. The remaining synthesis was yet to complete.

3.2.2c Synthesis of Fmoc protected aegPNA thymine monomer



Scheme 3: Synthesis of Fmoc protected aegPNA monomer

Fmoc protected aegPNA thymine monomer **22** was synthesized from Boc protected aegPNA thymine monomer **21**. Boc group of thymine monomer was deprotected with 50% TFA-DCM and this free amine was subjected for Fmoc protection to give Fmoc protected aeg thymine monomer. Fmoc protected aeg thymine monomer was used to synthesize aeg thymine octamer sequence as well as for chimeric aegPNA-*iso*-TANA sequences.

3.2.3 Solid phase synthesis of *iso*-TANA, aegPNA and aegPNA-TANA chimeric oligomers

Solid phase peptide synthesis protocols can be easily applied to the synthesis of *iso*-TANA as well as chimeric aegPNA-*iso*-TANA oligomers. The TANA monomers contain sugar glycosidic bond which are susceptible for acidic hydrolysis. Thus the Boc peptide strategy cannot be employed to synthesize *iso*-TANA oligomers as it contains high acidic conditions for deprotection of Boc group. The *iso*-TANA oligomers were synthesized by Fmoc strategy using Rink-Amide resin as solid support which is depicted in scheme-4. The TBTU/HOBt activation strategy[3] was employed for the coupling reaction. The first amino acid (β-alanine) is linked to this matrix *via* benzyl amide linkage.[4] This can be cleaved with standard protocol 10% TFA-DCM to yield the C-terminal amide. The amine content on the resin was determined by the picrate assay

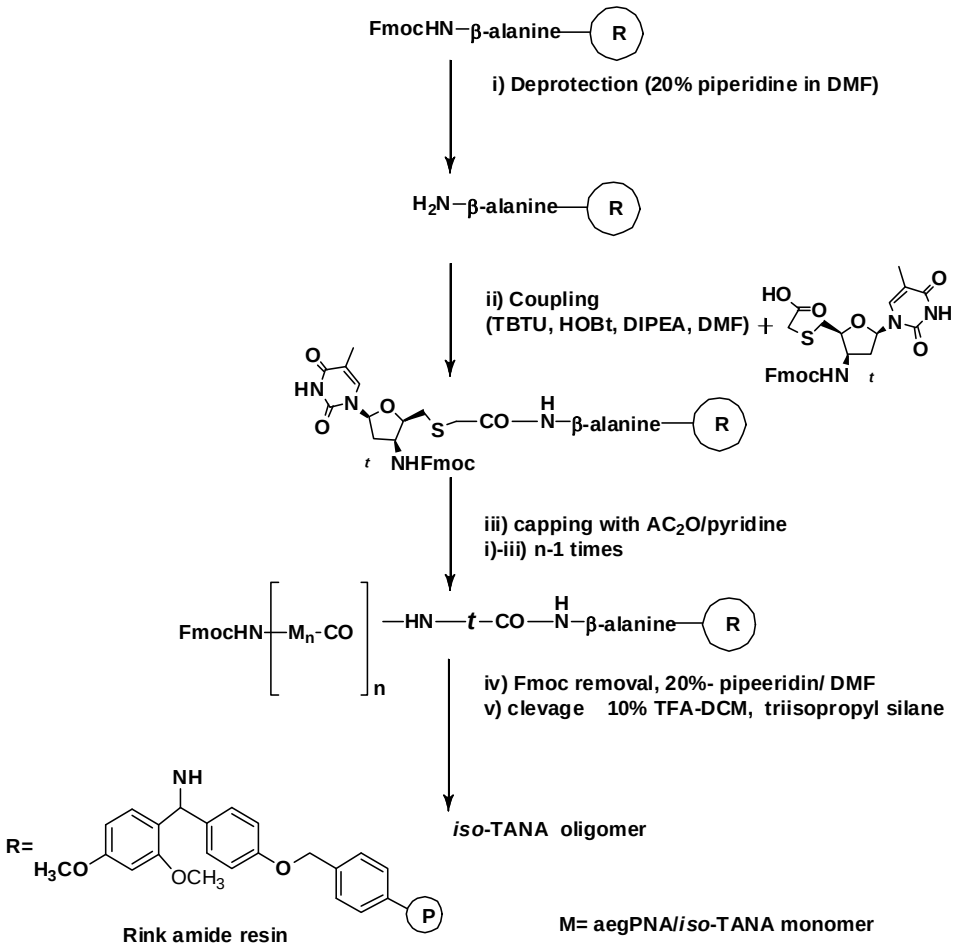
and was found to be 0.65 mmol/g and loading was suitably lowered to approximately 0.25 mmol/g by partial acetylation of amine content using calculated amount of acetic anhydride.[5] Free -NH_2 groups on the resin available for coupling are again estimated before starting synthesis.

The *iso*-TANA and aegPNA-*iso*-TANA oligomers were synthesized using repetitive cycles, each comprising the following steps:

i) Deprotection of *N*-protecting Fmoc group using 20% piperidine in DMF

ii) Coupling of the free amine with the free carboxylic acid group of the incoming monomer (3-4 equivalents). The coupling reaction was carried out in presence of (TBTU) and (HOBt) in DMF solvent and DIPEA as base. The deprotection of *N*-Fmoc protecting group and coupling reaction were monitored by Kaiser's test.[6]

iii) Capping of the unreacted amino groups using AC_2O in pyridine.



Scheme 4: Solid phase synthesis of *iso*-TANA oligomers

3.2.3a Synthesis of *iso*-TANA homooligomer and *iso*-TANA-PNA chimeric oligomers

The effect of *iso*-TANA unit on triplex forming ability could be tested by synthesizing thyminyl oligomeric sequences. These sequences are known to form complexes with complementary adenine DNA/RNA oligomers in 2:1 PNA:DNA/RNA stoichiometry. The control aminoethylglycyl (aeg) $PNAT_8$ oligomer was first synthesized following the Fmoc peptide strategy outlined Scheme 4. PNA sequences were synthesized incorporating the *iso*-TANA thymine units at predetermined position within the octamer. The series of octamer sequences comprising different aminoethylglycyl PNA-T and/*iso*-TANA T units is listed in Table 1. With a view to exploring the end effects of *iso*-TANA monomeric unit, a single *iso*-TANA unit was introduced at 'C' terminus (Table- 1, entry 3). Also to explore the effect of *iso*-TANA unit at middle of sequence, a single *iso*-TANA unit was incorporated in the middle of sequence (Table 1, entry 2). With view to see the combined effect of *iso*-TANA unit in C terminal as well middle positions a sequence with two *iso*-TANA T units was synthesized. A sequence with alternate amide and thioacetamido backbone was synthesized. The contribution of both linkages on the binding of oligomers to nucleic acids can be explored with this sequence. Fully modified oligomer with thymine *iso*-TANA units also has been synthesized (Table-1, entry 1).

3.2.3bCleavage of the iso-TANA and aegPNA-*iso*-TANA oligomers from the solid support

The oligomers synthesized were cleaved from solid support using 10% trifluroacetic acid in DCM . The cleavage time of ½ hr was found to be optimum. After the cleavage of oligomers, solution of oligomers in TFA–DCM was evaporated and oligomers were precipitated out with cold diethyl ether.

3.2.4 Purification and MALDI-TOF characterization of oligomers

The purity of the oligomers was checked by analytical RP-HPLC (C18 column, CH_3CN-H_2O- 0.1% TFA system), which showed more than 65-70% purity. These were subsequently purified by reverse phase HPLC on C18 column. The purity of the oligomers was again ascertained by analytical RP-HPLC and their integrity was confirmed by MALDI-TOF mass spectrometric analysis.

Table 1: Mixed *iso*-TANA and chimeric *iso*-TANA-PNA thymine octamers , HPLC analysis and MALDI-TOF mass analysis

Entry	Sequence	HPLC t_R (min)	Mass calculated/observed
1	H-TTTTTTTT-βAlanine **aegT**	9.0	2218.2/2218.8
2	H-*tttttttt*-βAlanine $\mathbf{T_1}$	13.1	2467.6/2488.91(M+Na^+)
3	H-TTTT*t*TTT-βAlanine $\mathbf{T_2}$	9.7	2249.4/2278.54(M+6+Na^+)
4	H-TTTTTTT*t*-βAlanine $\mathbf{T_3}$	10.0	2249.4/2274.91(M+1+Na^+)
5	H-TTT*t*TTT*t*-βAlanine $\mathbf{T_4}$	10.6	2280.6/2279.35
6	H-T*t*T*t*T*t*T*t*-βAlanine $\mathbf{T_5}$	11.1	2343.00/2343.99

T = aeg thymine unit, *t* = *iso*-TANA thymine unit

3.2.5 UV-melting studies *iso*-TANA modified oligomers

The binding strength of *iso*-TANA T_8 sequences ($\mathbf{T_1}$ to $\mathbf{T_5}$) with complementary DNA and RNA was investigated by UV spectroscopic techniques. The effect of incorporation of *iso*-TANA thymine units in PNA thymine octamer **aegT** sequence on stability of PNA:DNA and PNA:RNA complexes also has been investigated. The results of these studies are presented in this section.

Homopyrimidine *iso*-TANA-T_8 sequences: UV studies

UV-melting studies of *iso*-TANA and chimeric *iso*-TANA–PNA octamers are tabulated in Table 2. Control **aegT** sequence binds with complementary **DNA-1** and **RNA-1** with similar strength (Table-2, entry-1). *iso*-TANA homooligomers $\mathbf{T_1}$ did not showed any binding with **DNA-1** while the complex with **RNA-1** showed slight destabilization as compared to **aegT**. Sequences $\mathbf{T_2}$-$\mathbf{T_5}$ with *iso* TANA unit/s in the **aeg T** also did not show any transition in the melting experiments with complementary **DNA-1**.

Table 2: UV melting studies *iso*-TANA and chimeric *iso*-TANA-PNA thymine octamers with DNA /RNA

Entry	Code	Sequence	UV-T_m °C	
			DNA-1	RNA-1
1	**aegT**	H-TTTTTTTT-βAlanine	41.6	41.2
2	$\mathbf{T_1}$	H-*tttttttt*-βAlanine	nd	37.6
3	$\mathbf{T_2}$	H-TTTT*t*TTT-βAlanine	nd	37.2
4	$\mathbf{T_3}$	H-TTTTTTT*t*-βAlanine	nd	34.5
5	$\mathbf{T_4}$	H-TTT*t*TTT*t*-βAlanine	nd	37.3
6	$\mathbf{T_5}$	H-T*t*T*t*T*t*T*t*-βAlanine	nd	55.0

DNA-1 = 5**'**GCA_8CG 3'; **RNA-1 =** 5**'**GCA_8CG 3', a *T*m = melting temperature (measured in the buffer 10 mM sodium phosphate, 10 mM NaCl, 0.1 mM EDTA, pH = 7.0), *iso*-TANA_2-DNA/ *iso*-TaNA_2-RNA complexes. The values reported here are the average of 3 independent experiments and are accurate to ±1.0°C.

The complex of **T_2** with **RNA-1** showed little destabilization than the control **aegT**:RNA complex. Incorporation of *t* unit at 'C' terminal of **aegT** (**T_3**) destabilized the complex with RNA more than the sequence **T_2** in which the *t* unit was in the centre of the sequence. The doubly modified sequence (**T_4**) showed similar binding like homooligomer **T_1** with **RNA-1**. The alternating aegPNA-*iso*-TANA backbone sequence **T_5** showed increase in binding affinity with **RNA-1** than homo-oligomer **T_1**. The increase in the melting temperature was about 15 °C than the control aegT:**RNA-1** complex.

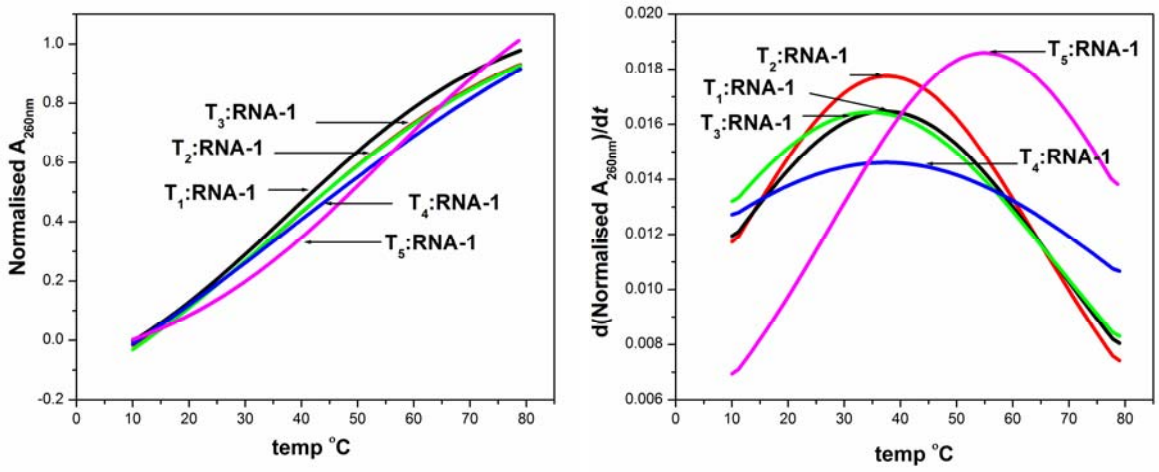


Figure 1:**A** UV melting Curves *iso*-TANA ONs T_1-T_5 with RNA-**1** and **B** corresponding derivative graphs

3.2.6 Summary

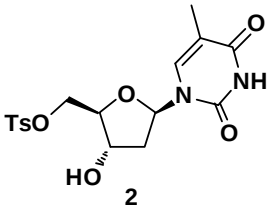
- The iso-TANA thymine monomer has been synthesized and its octmeric homoligomer was synthesized. This monomer also has been incorporated into aegPNA thyminyl octamer sequences.
- The biophysical studies of *iso*-TANA homoligomer and aegPNA-*iso*-TANA chimeric sequences with DNA and RNA were carried out. These sequences preferentially bind to RNA over DNA. The aegPNA sequences containing *iso*-TANA thymine units as well as homoligomer of *iso*-TANA destabilized the complex with RNA as compared to control aegPNA sequences.

3.2.7 Experimental

General: The chemicals used were of laboratory or analytical grade. All solvents used were purified according to the literature procedures. Reactions were monitored by TLC. Usual reaction work up involved sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulphate and evaporation of the solvent under vacuum. IR spectra were recorded on an infrared Fourier Transform spectrophotometer using nujol, chloroform or neat. Column chromatographic separations were performed using silica gel 60-120 mesh (Merck) and 200- 400 mesh (Merck) and using the solvent systems EtOAc/petroleum ether and MeOH/DCM. TLCs were carried out on pre-coated silica gel GF254 sheets (Merck 5554). TLCs were run in either petroleum ether with appropriate quantity of ethyl acetate or dichloromethane with an appropriate quantity of methanol for most of the compounds. TLCs were visualized with UV light and iodine spray and/or by perchloric acid treatment then heating. ^{1}H and ^{13}C spectra were obtained using Bruker AC-200, AC-400 and AC-500 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^{1}H. The optical rotation values were measured on Bellingham-Stanley Ltd, ADP220 polarimeter. Mass spectra were obtained either by LCMS techniques. Oligomers were characterized by RP HPLC (Varian /Waters) C18 column and MALDI-TOF mass spectrometry. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems) MALDI-TOF instrument and the matrix used for analysis was CHCA (α-Cyano-4-hydroxycinnamic Acid). DNA oligomers were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard β-cyanoethyl phosphoramidite chemistry.

Synthesis of 5'-*O*-tosyl thymidine 2

Thymidine **1** (10 g, 41.28 mmol) was suspended in 60 mL dry pyridine and cooled to 0˚C. Tosyl chloride (11.8 g, 61.92 mmol) diluted with 20 mL pyridine was added to above mix over period of 4hrs at 0 °C. Then reaction mixture was stirred for 1 hr at RT and pyridine was removed under reduced pressure. 50 mL Ethyl acetate and 30 mL of sat. $NaHCO_3$ solution was added. The only 5'-monotosylated derivative was precipited out, which was filtered and washed with 50 mL ethyl acetate and 50 mL water to give pure product (yield = 9.8 g 60%)

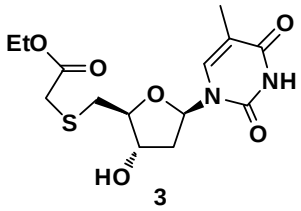


2

^{1}H NMR (200 MHz, DMSO-d_6) δ 1.89 (s, 3H), 2.10-2.25 (m.2H), 2.4 (s, 3H), 3.98-4.1(m, 3H), 4.24 (m, 1H), 6.2 (t, 1H, *J*= 6.32 Hz), 7.33(d, 3H,), 7.78(m, 2H) Mass calcd. for $C_{17}H_{20}N_2O_7S$ 396.09 obsd. 396.46

Synthesis of ethyl [(5'-*S*-thymidinyl) mercapto] acetic acid 3

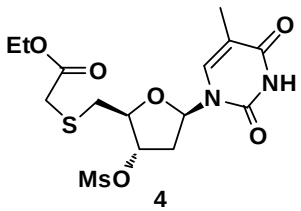
ethyl mercapto acetic acid (2.9 mL, 27.2 mmol) was added drop wise to suspension of NaH (60% suspension in hexane) (1.0 g, 27.2 mmol) and in dry DMF to generate mercapto nucleophile. Tosyl derivative **2** (9.0 g, 22.7 mmol) in 20 mL dry DMF was added to it at 0˚C. Then reaction was stirred for 1 hr at RT, DMF was removed at in vacuo, residue was redissolved in 200 mL ethyl acetate, organic layer was washed with 50 mL water. Organic layer was concentrated under reduced pressure; compound was purified with column chromatography with 1-5% methanol in dichloromethane to white solid (Yield 6.3 g = 81 %).



3

Synthesis of ethyl [(3'-*O*-mesyl-5'-*S*-thymidinyl) mercapto] acetic acid 4

A mixture of **3** (6 g, 17.4 mmol) and Et_3N (7.2 mL, 52.2 mmol) in 100mL anhydrous dichloromethane was cooled to 0˚C and mesyl chloride (2.1 mL, 26.1 mmol) in 10 mL dry dichloromethane) was added dropwise. After 10 min reaction mixture was taken in 100 ml dichloromethane, organic layer was washed with 10% $NaHCO_3$ solution (2x 30 mL) and water (2x50 mL). The organic layer was dried over anhydrous sodium sulphate and concentrated to dryness. Column purification using 3% methanol-dichloromethane gave pure product **4**, 7.3 g, 85%.

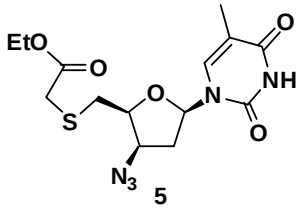


4

Mass calcd. for $C_{15}H_{22}N_2O_8S$ 422.7 obsd. 445.10 (M+ Na^+).

Synthesis of ethyl [(3'-deoxyxylo-3'-azido-5'-*S*-thymidinyl) mercapto] acetic acid 5:

Mesyl derivative **4** (7.0 g, 16.5 mmol) and sodium azide (8.6 g, 132 mmol) in 20 mL anhydrous DMF was stirred at 60˚C for 24 hrs. DMF was removed under reduced pressure and residue was redissolved in 200 mLethyl acetate. The organic layer was washed with water (2x 50 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to dryness. Column purification using gradient of ethyl acetate in petroleum ether gave pure **5**, 2.7 g, 45%.



5

IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019, 2110, 1724, 1670.

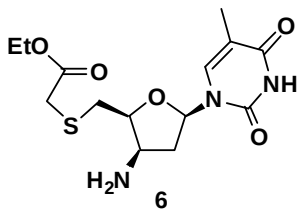
^{1}H NMR (200 MHz, $CDCl_3$) δ 1.19-1.26 (t, 3H, *J* = 7.08 Hz) 1.89 (s, 3H), 2.08-2.17 (m, 1H), 2.64-2.78 (m, 1H), 2.65-2.79 (m, 1H), 2.85-3.20 (m, 2H), 3.16-3.30 (dd, 2H, *J* = 15.03 Hz), 4.07-4.19 (m, 3H), 4.24-4.28 (m, 1H) 6.09-6.14 (dd, 1H, *J* =2.77 Hz, *J* =7.95 Hz), 7.38 (s, 1H).

^{13}CNMR (50 MHz, $CDCl_3$) δ 12.4, 13.8, 31.2, 33.9, 38.4, 61.3, 61.4, 82.0, 83.7, 110.6, 135.2, 150.5, 164.0, 169.9.

Mass calcd. for $C_{14}H_{19}N_5O_5S$ 369.11 obsd. 370.89, 392.93 (M+ Na^+).

Synthesis of ethyl [(3'-deoxyxylo-3'-amino-5'-*S*-thymidinyl) mercapto] acetic acid 6:

Azido compound **5** (2.5 g, 6.77 mmol) was dissolved in pyridine (10 mL) and Et_3N 3 mL was added to this solution. H_2S gas was bubbled into the solution at 0˚C for 5 minutes. Then solution was stirred at room temperature for additional 1 hr. and solvent was removed under reduced pressure. The residue was purified by column chromatography (2-5% methanol- dichloromethane) to afford **6**, 1.8 g (78%).



6

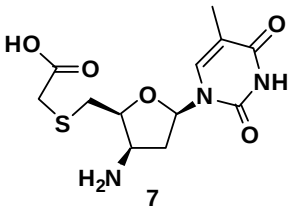
IR, υ(cm^{-1}) ($CHCl_3$) 3398, 3019, 2924,2853,1701,1686.

^{1}H NMR (200 MHz, $CDCl_3$) δ 1.30 (t, 3H, *J* = 7.20 Hz), 1.80-1.82 (m, 1H), 1.93 (s 3H), 2.62-2.76 (m, 1H), 2.91-3.17 (m, 2H), 3.23-3.43 (dd, 2H, *J* = 15.03 Hz) , 3.72-3.78 (m, 1H), 4.03-4.12 (m, 1H), 4.16-4.27 (q, 2H, *J* = 7.20 Hz), 6.12-6.17 (dd, 1H, *J* = 3.16 Hz, 8.08 Hz), 8.01 (s, 1H).

Synthesis of [(3'-deoxyxylo-3'-amino-5'-*S*-thymidinyl) mercapto] acetic acid 7:

Compound **6 (**1.8 g, 5.28 mmol) was dissolved in 5 mL MeOH, 3 mL 1N LiOH was added. The reaction was stirred for 1 hr. pH of reaction mixture was adjusted to 7 with Dowex H^+ resin. Then MeOH was evaporated and aqueous layer was washed with

20 mL ethyl acetate. Aqueous layer was concentrated to give amino acid compound **7**.



7

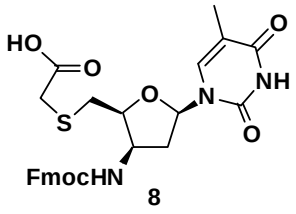
The compound **7** was desiccated and used for next reaction. Yield (1.4 g, 87%)

^{1}H NMR (200 MHz, D_2O) δ 1.78 (s, 3H), 2.34-2.44 (m, 1H), 2.79-2.99 (m, 3H), 3.21-3.25 (d, 2H), 4.01-4.08 (m, 1H), 4.16-4.25 (m, 1H), 5.61-5.68 (dd, 1H, J = 5.44 Hz, 8.47 Hz), 7.45 (s, 1H)

Mass calcd. for $C_{12}H_{17}N_3O_5S$ 315.08 obsd. 316.48, 338.49 ($M+Na^+$)

Synthesis of [(3'-deoxyxylo-3'-Flurenylmethoxycabonylamino-5'-*S*-thymidinyl) mercapto] acetic acid 8:

The amino acid obtained in the previous step (1.3 g, 1.1 mmol) was dissolved in 3 mL water-dioxan and $NaHCO_3$ (0.23 g, 2.75 mmol) and Fmoc –succinimide (0.44 g, 1.32 mmol) was added to the reaction mixture. The reaction was stirred for 6 hrs at RT. Then 1,4-dioxane was removed under reduced presssure and aqueous layer was washed with ethyl acetate until impurity of Fmoc succinimide is removed. Then aqueous layer was neutralized with H^+ dowex resin. The compound was then extracted in 40 mL 5% methanol in ethyl acetate. Organic layer was dried over Na_2SO_4, concentrated and the product was purified by column chromatography using gradient methanol-dichloromethane to give thymine monomer **8** (1.3 g, 62%).



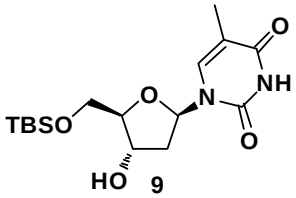
8

^{1}H NMR (200 MHz, $CDCl_3$ + drop of DMSO-d_6) δ 1.91 (s, 3H), 2.05-2.25 (m, 1H), 2.28 (m, 2H), 3.20-3.49 (m, 2H), 4.19-4.49 (m, 4H), 5.67-5.70 (m, 1H), 6.70-6.79 (m, 1H), 7.27-7.77 (m, 9H)

^{13}CNMR (50 MHz, $CDCl_3$ + drop of DMSO-d_6) δ 12.1, 28.7, 34.1, 46.9, 52.0, 66.4, 80.9, 110.6, 119.7, 124.9, 128.9, 127.5, 141.0, 143.5, 150.2, 156.0, 164.4, 172.3

Mass calcd. for $C_{27}H_{27}N_3O_7S$ 537.15 obsd. 560.14 ($M+Na^+$)

Synthesis of 5-*O*-TBS thymidine (9)



9

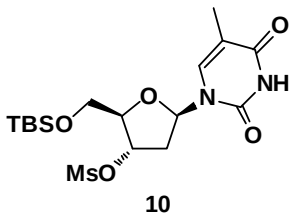
Thymidine **1** (10 g, 41.28 mmol) and imidazole (6.0 g, 90.75 mmol) was suspended in 12 mL of dry DMF. TBS-Cl (6.5 g, 41.25 mmol) was added at 0 °C and reaction was stirred at RT for 7-8 hrs. DMF was removed under vacuume and product was extracted in ethyl acetate. Organic layer was concentrated

and product was purified with column chromatography. Yield (12.1 g, 80 %),

^{1}H NMR (200 MHz, $CDCl_3$) δ 0.11 (s, 6H), 0.92 (s, 9H), 1.91 (s, 3H), 2.05-2.51 (m, 2H), 3.80-4.07 (m, 3H), 4.4-4.46 (m, 1H), 6.38-6.40 (m, 1H), 7.55 (s, 1H), 9.42 (bs, 1H)

Synthesis of 3'-*O*-mesyl -5'-*O*-TBS thymidine (10)

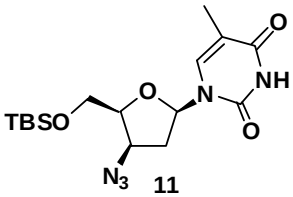
A mixture of **9** (11 g, 30.0 mmol), Et_3N (10.0 mL, 72 mmol) in 40 mL anhydrous dichloromethane was cooled to 0˚C and mesyl chloride (2.7 mL, 36.3 mmol) in 10 mL dry dichloromethane) was added dropwise. After 10 min reaction mixture was taken in 100 ml dichloromethane, organic layer was washed with 10% $NaHCO_3$ solution with (2x 30 mL) followed by water (2x50 mL). The organic layer was dried over anhydrous sodium sulphate and concentrated to dryness. Column purification using 2% methanol-dichloromethane gave pure product **10**, 11.3 g, 85%.



10

Synthesis of 3'-deoxyxylo-3'-azido -5'-*O*-TBS thymidine (11)

Mesyl derivative **10** (11 g, 24.78 mmol) and sodium azide (12.88 g, 198 mmol) in 20 mL anhydrous DMF was stirred at 60˚C for 48 hrs. DMF was removed under reduced pressure and residue was redissolved in ethyl acetate. The organic layer was washed with water (2x 50 mL). The organic layer was dried on anhydrous sodium sulphate and concentrated to dryness. Column purification using gradient of ethyl acetate in petroleum ether gave pure **11**, 3.9 g, 42%.



11

IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019, 2110, 1724, 1670.

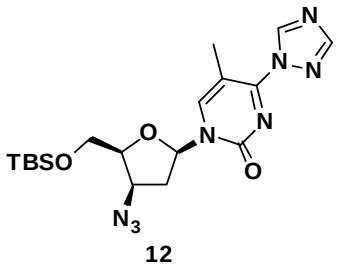
^{1}H NMR (200 MHz, $CDCl_3$) δ 0.12 (s, 6H), 0.92 (s, 9H), 1.96 (s, 3H), 2.13-2.21 (m, 1H), 2.67-2.82 (m, 1H), 3.93-4.09 (m, 3H), 4.26-4.30 (m, 1H), 6.15-6.21 (dd, 1H, J = 3.03 Hz, 7.58 Hz), 7.47 (s, 1H), 8.96 (bs, 1H) .

^{13}CNMR (50 MHz, $CDCl_3$) δ -5.6, 12.6, 18.1, 25.7, 38.3, 53.3, 60.8, 61.0, 82.2, 83.7, 110.8, 135.4, 150.6, 164.0.

Mass calcd. for $C_{16}H_{27}N_5O_4Si$ 381.18 obsd. 404.30 ($M+Na^+$)

Synthesis of 3'-deoxyxylo-3'-azido-5'-*O*-TBS-*C*^{4}-(1,2,4-triazol-1-yl)-pyrimidine 2'deoxynucleoside (12)

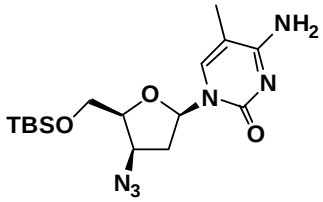
Et_3N (7.0 mL, 50.5 mmol) was added drop wise to a stirred mixture of 1, 2, 4-triazole (1.5 g, 22.7 mmol) and phosphoryl chloride (2.1 mL, 22.7 mmol) in 50 mL CH_3CN at 0˚C. The solution of **11** (3.5 g, 9.1 mmol) in 15 mL dry CH_3CN was then added drop



12

wise. The reaction mixture was stirred at room temp for 2.5 hrs. The solvent was evaporated and residue was redissolved in 50 mL ethyl acetate. The organic layers was washed with water (2x50 mL) and brine (2x20 mL), dried over Na_2SO_4 and evaporated to dryness to give triazolide derivative **12** (3.5 g, 88%) as white foam, which was used immediately for next reaction.

Synthesis of 2',3'-dideoxyxylo-3'-azido-5'-*O*-TBS-5-methylcytidine (13)



13

Concentrated aqueous ammonia (10 mL) was added to the solution of triazolide derivative **12** (3.5 g, 8.1 mmol) in 40 mL 1,4-dioxane. After complete conversion of triazolide derivative to 5-methyl- 2'-deoxy cytidine derivative in 1hr, solvent was removed under reduced pressure and the residue redissolved in dichloromethane and purified by silica gel column chromatography (0-7% methanol in dichloromethane) to afford product (2.77 g, 90%) as white foam.

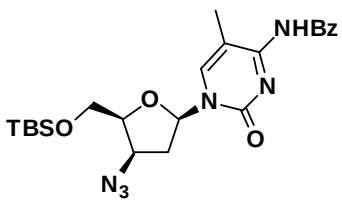
IR, $\upsilon(cm^{-1})$ ($CHCl_3$) 3192, 2984, 2827, 2114, 1687.6

^1H NMR (200 MHz, $CDCl_3$) δ 0.12(s, 6H), 0.92 (s, 9H), 1.96 (s, 3H), 2.08-2.15 (m, 1H), 2.65-2.80 (m, 1H), 3.94-3.98 (m, 2H), 4.04-4.10 (m, 1H), 4..22-4.27 (m, 1H), 6.13-6.18 (dd, 1H, J = 3.03 Hz, 7.58 Hz) , 7.47 (s, 1H).

Mass calcd. for $C_{16}H_{28}N_6O_3Si$ 380.19 obsd. 403.39 ($M+Na^+$)

Synthesis of N^4-benzoyl-2', 3'-dideoxy xylo-3'-azido-5'-O-TBS-5-methylcytidine (14)

Benzoyl chloride (1.01 mL, 8.50 mmol) was added drop wise slowly to a solution **13**



14

(2.7 g, 7.12 mmol) in 30 mL pyridine at 0˚C. The mixture was then stirred at room temperature for 4 hrs. The reaction was quenched with addition of 5 mL 2N NH_3 solution. Reaction mixture was further stirred for 30mins. The solvent was evaporated and residue redissolved in 80 mL ethyl acetate, washed the organic layer with 10% $NaHCO_3$ solution (3x30 mL), water (2x50 mL) and brine (20 mL). The organic layer was then dried over anhydrous sodium sulphate and evaporated to dryness. Compound was purified on column chromatography (5-20% ethyl acetate in petroleum ether) to afford **14**, 2.57 g (75%) as white foam.

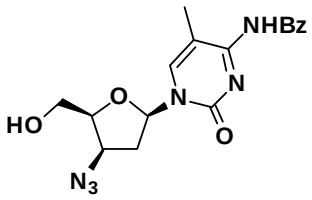
IR, υ(cm^{-1}) ($CHCl_3$) 3192, 3024, 2110, 1715, 1705, 1696, 1685

^{1}H NMR (200 MHz, $CDCl_3$) 0.13 (s, 6H), 0.93 (s, 6H), 2.15 (s, 3H), 2.10-2.24 (m, 1H), 2.68-2.82 (m, 1H), 3.96-4.00 (m, 2H), 4.06-4.12 (m, 1H), 4.28-4.32 (m, 1H), 6.15-6.20 (dd, 1H, *J* = 3.03 Hz, 7.83 Hz), 7.40-7.62 (m, 4H), 8.30-8.34 (m, 2H).

^{13}C NMR (50 MHz, $CDCl_3$) δ -5.5, 13.7, 18.1, 25.7, 38.6, 60.9, 61.0, 82.8, 84.5, 111.5, 128.0, 129.8, 132.3, 136.8, 137.0, 147.9, 159.6, 179.4.

Mass Calc. for $C_{23}H_{32}N_6O_4Si$ 484.22 observed 485.61 ($M+ H^+$), 507.61 ($M+Na^+$).

Synthesis of N^4-benzoyl-2', 3'-dideoxy xylo-3'-azido-5-methylcytidine (15)

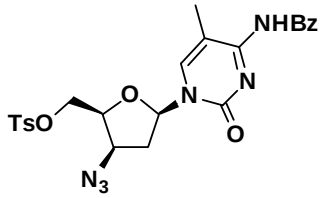


15

Compound **14** (2.5 g, 5.1 mmol) was taken in 10 mL THF and 7.7 mL 1N TBAF in THF was added to this solution. The reaction mixture was stirred for 30 mins and THF was removed under vacuume. The compound was extracted in DCM. The organic layer was concentrated and crude compound was used in the next step without purification. Crude yield = (1.7 g, 90%)

Synthesis of N^4-benzoyl-2', 3'-dideoxy xylo-3'-azido-5'-*O*-tosyl-5-methylcytidine (16)

Tosyl chloride (1.13 g, 5.97 mmol) was added to the solution of compound **15** in 10 mL pyridine at 0 °C. The reaction was monitored by TLC and after completion of reaction, pyridine was under vacuume. The compound was extracted in 100 mL ethyl acetate and organic layer was washed with 50 mL sat. $NaHCO_3$ solution. The organic layer was concentrated and compound was purified with column chromatography (10-30% ethyl acetate in petroleum ether). Yield = 1.2 g, 80%.



16

IR, υ(cm^{-1}) ($CHCl_3$) 3128, 3019, 2110, 1724, 1670.

^{1}H NMR (200 MHz, $CDCl_3$) 2.13 (s, 3H), 2.2-2.31 (m, 1H), 2.47 (s, 3H), 2.72-2.86 (m, 1H), 4.07-4.35 (m, 4H), 6.15-6.20 (d, 1H, *J* =3.54 Hz, 7.58 Hz), 7.37-8.34 (Ar , 10 H)

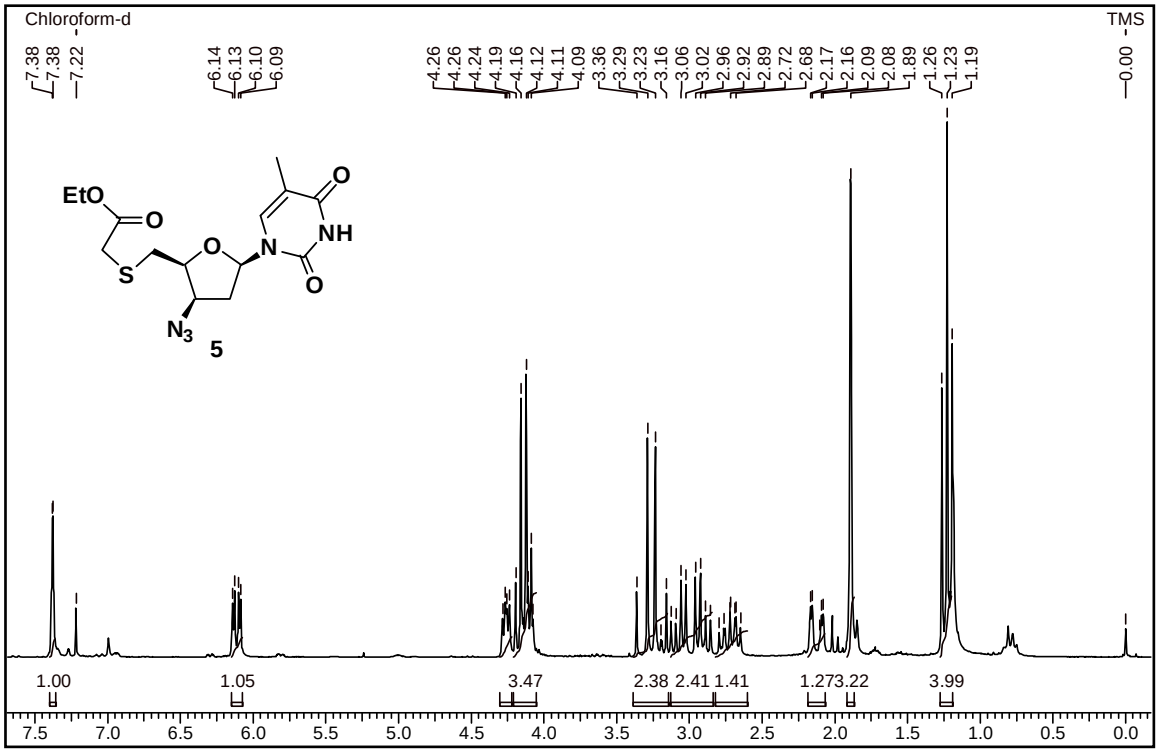
^{13}CNMR (50 MHz, $CDCl_3$) δ 13.5, 21.5, 38.2, 60.3, 66.9, 79.1, 84.4, 111.8, 127.8, 129.9, 132.3, 132.2, 136.2, 136.6, 145.3, 147.8, 159.4, 179.38.

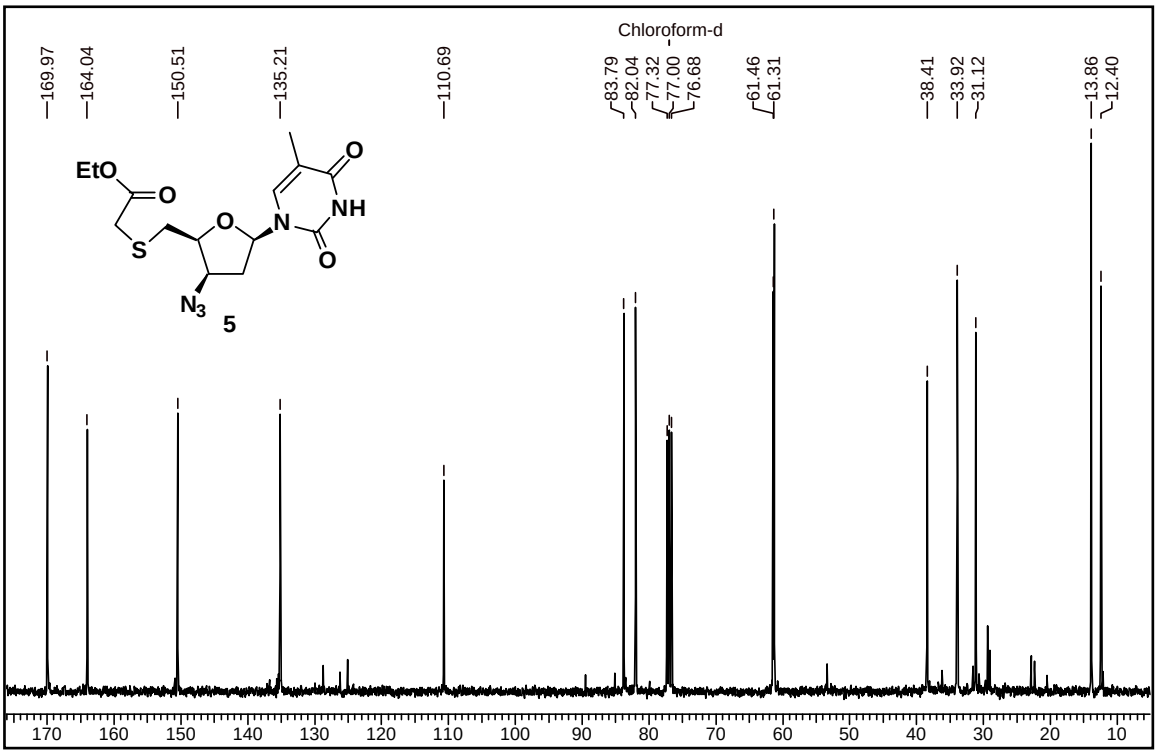
Mass calcd. for $C_{24}H_{24}N_6O_6S$ 524.14 obsd. 547.33 ($M+Na^+$)

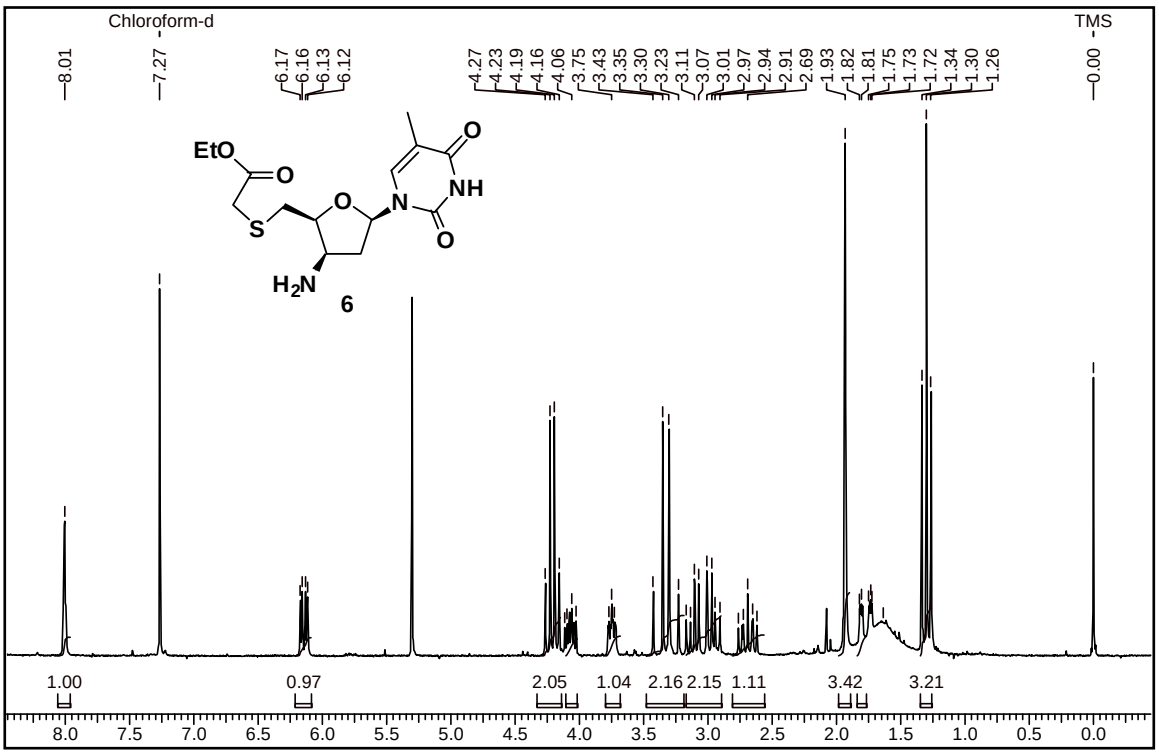
3.2.8 Appendix

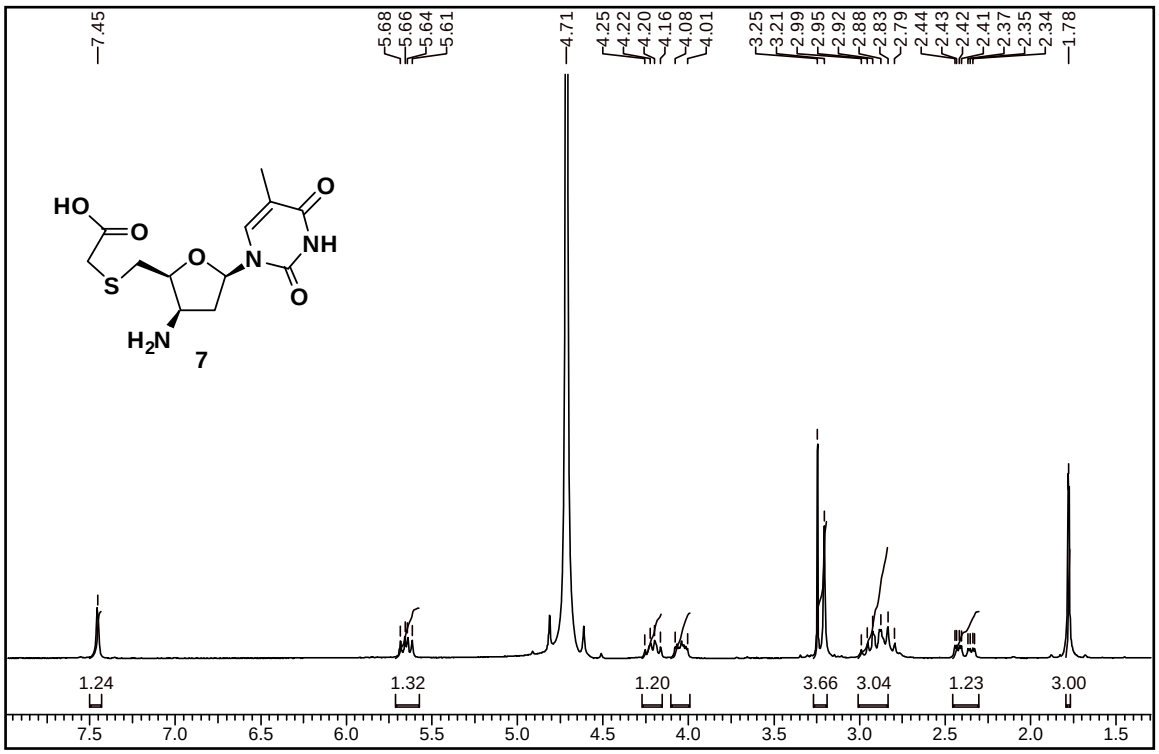
^{1}H and ^{13}C spectra for compound **5-16** page no. 221-229

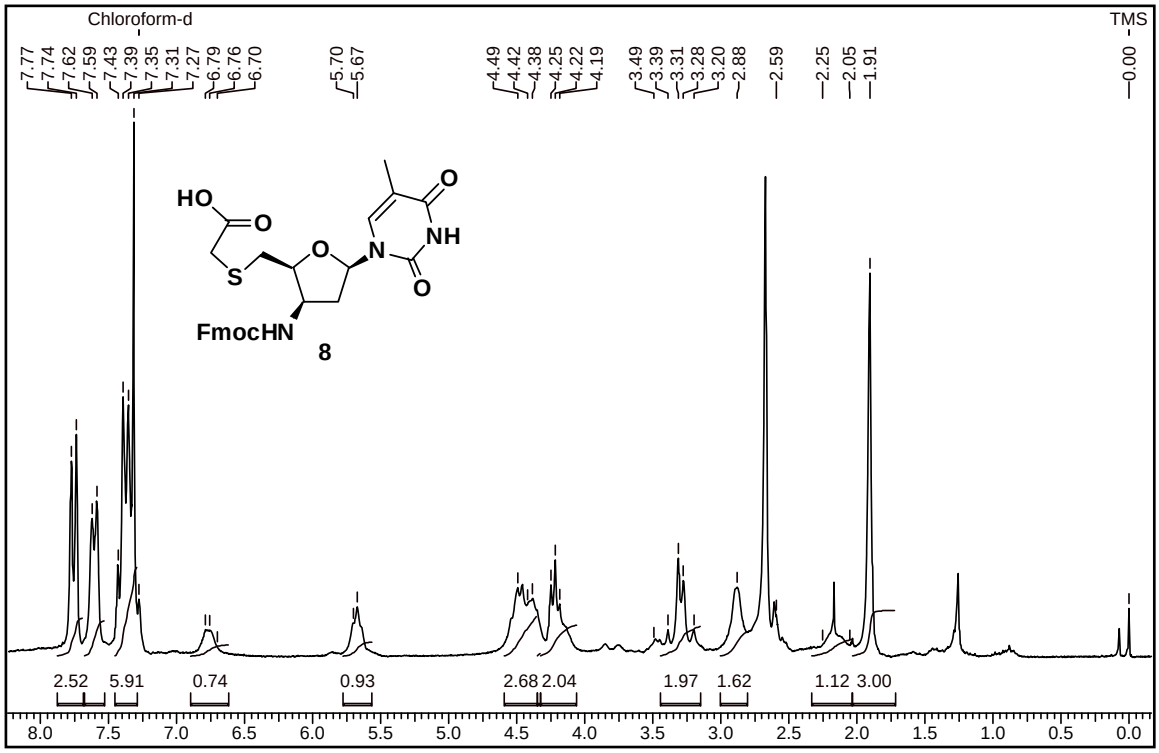
HPLC and MALDI-TOF spectra for T_1-T_5 230-233

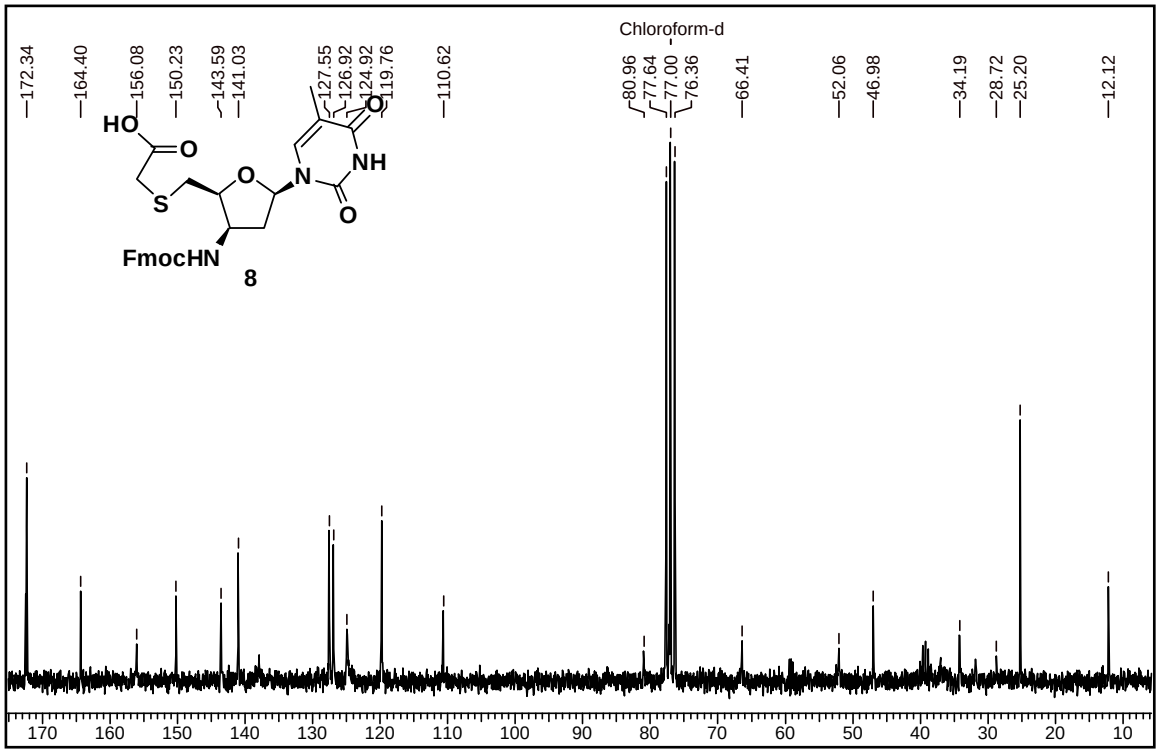

Chloroform-d
TMS
7.38
7.38
7.22
6.14
6.13
6.10
6.09
4.26
4.26
4.24
4.19
4.16
4.12
4.11
4.09
3.36
3.29
3.23
3.16
3.06
3.02
2.96
2.92
2.89
2.72
2.68
2.17
2.16
2.09
2.08
1.89
1.26
1.23
1.19
0.00
EtO
O
S
O
N
NH
O
O
N3
5
1.00
1.05
3.47
2.38
2.41
1.41
1.27
3.22
3.99
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

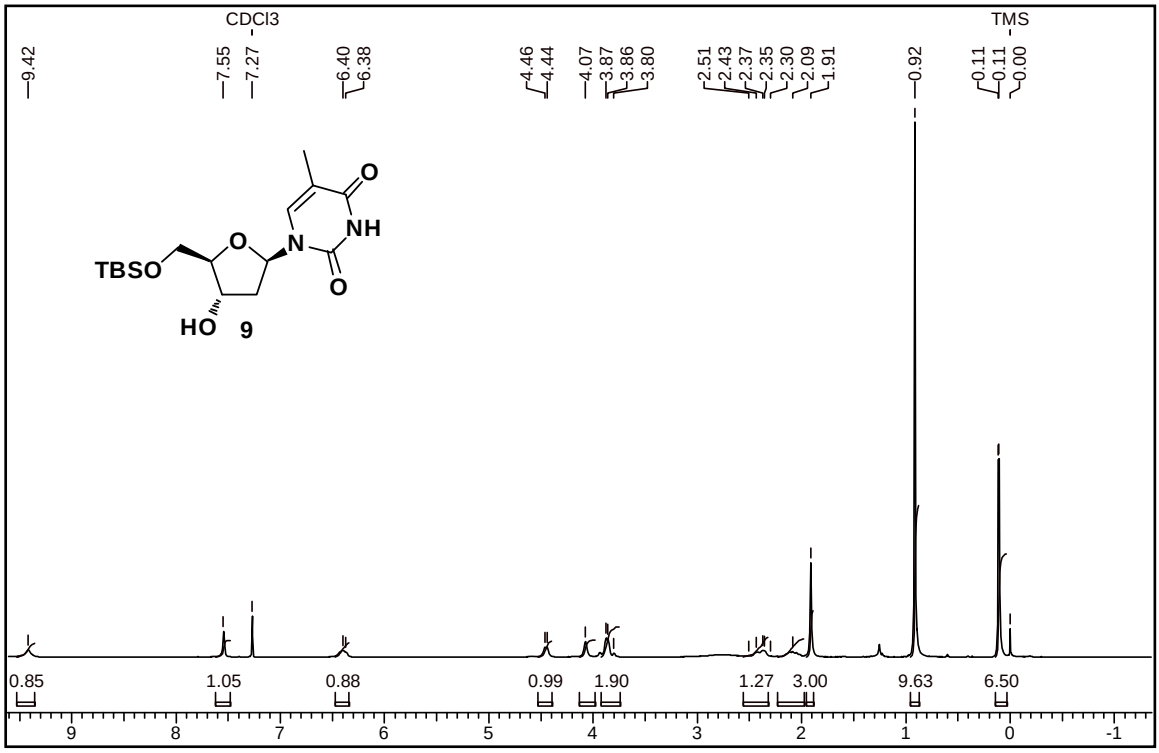

Chloroform-d
169.97
164.04
150.51
135.21
110.69
83.79
82.04
77.32
77.00
76.68
61.46
61.31
38.41
33.92
31.12
13.86
12.40
EtO
O
S
O
N
NH
O
O
N3
5
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

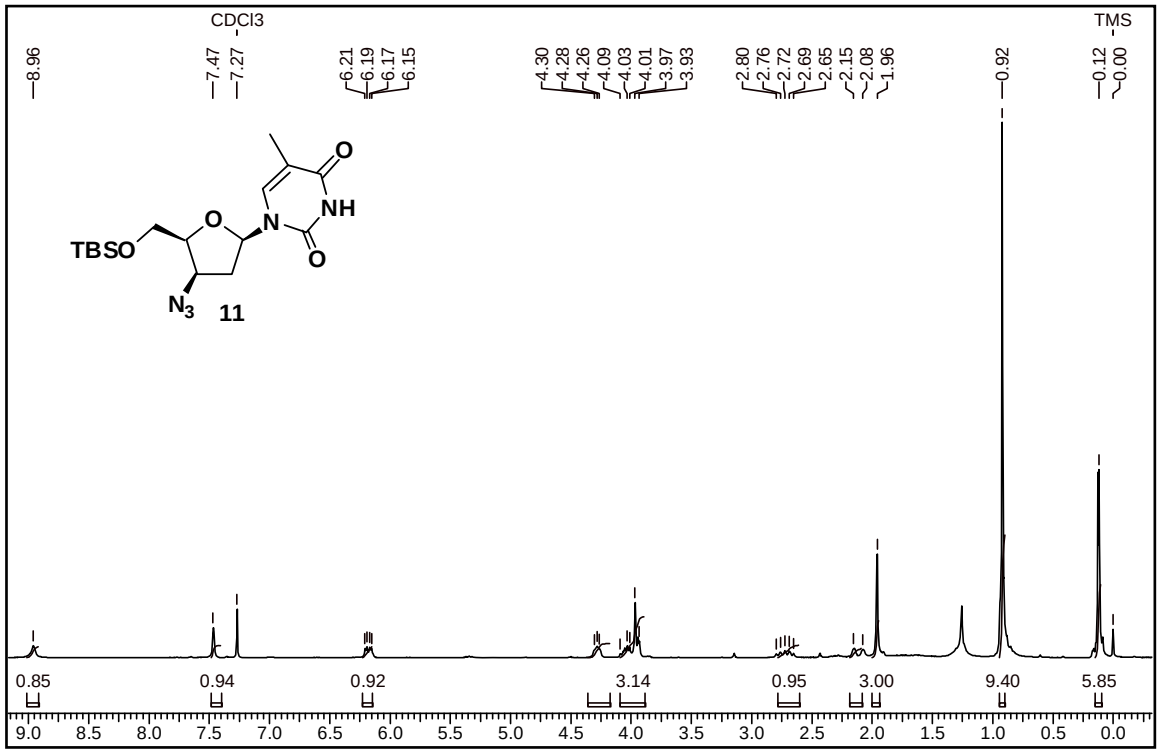

Chloroform-d
TMS
8.01
7.27
6.17
6.16
6.13
6.12
4.27
4.23
4.19
4.16
4.06
3.75
3.43
3.35
3.30
3.23
3.11
3.07
3.01
2.97
2.94
2.91
2.69
1.93
1.82
1.81
1.75
1.73
1.72
1.34
1.30
1.26
0.00
EtO
O
S
O
N
NH
O
O
H2N
6
1.00
0.97
2.05
1.04
2.16
2.15
1.11
3.42
3.21
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

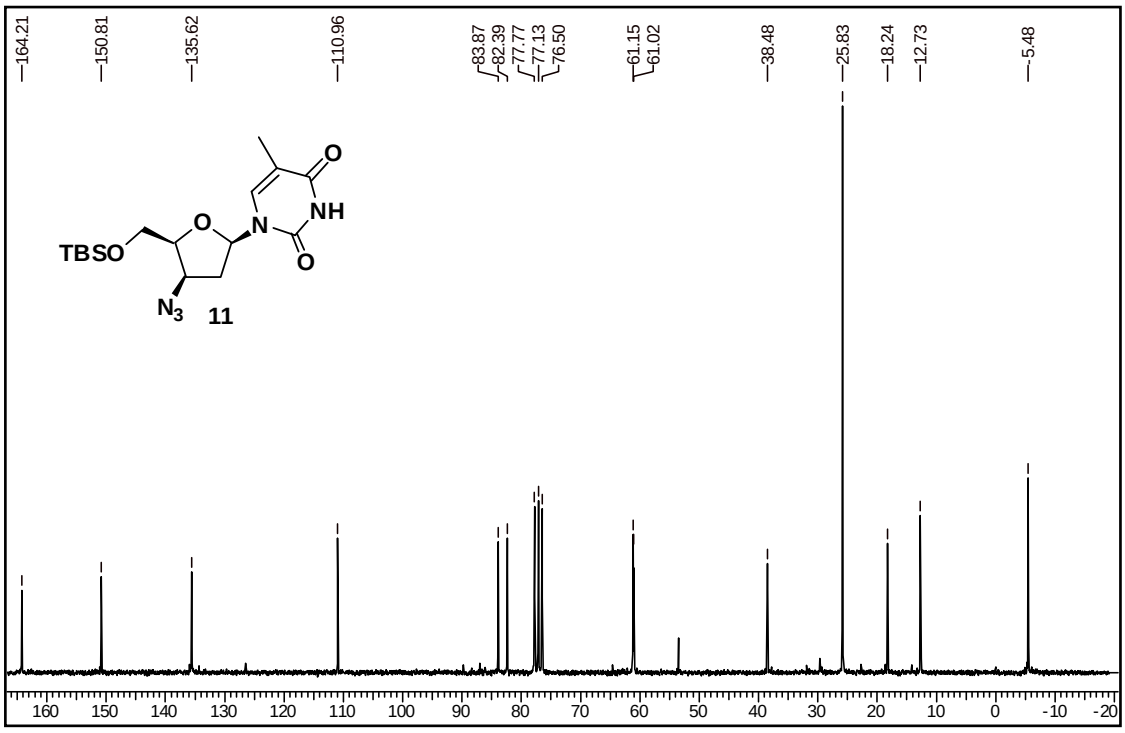

7.45
5.68
5.66
5.64
5.61
4.71
4.25
4.22
4.20
4.16
4.08
4.01
3.25
3.21
2.99
2.95
2.92
2.88
2.83
2.79
2.44
2.43
2.42
2.41
2.37
2.35
2.34
1.78
HO
O
S
O
N
NH
O
O
H2N
7
1.24
1.32
1.20
3.66
3.04
1.23
3.00
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5

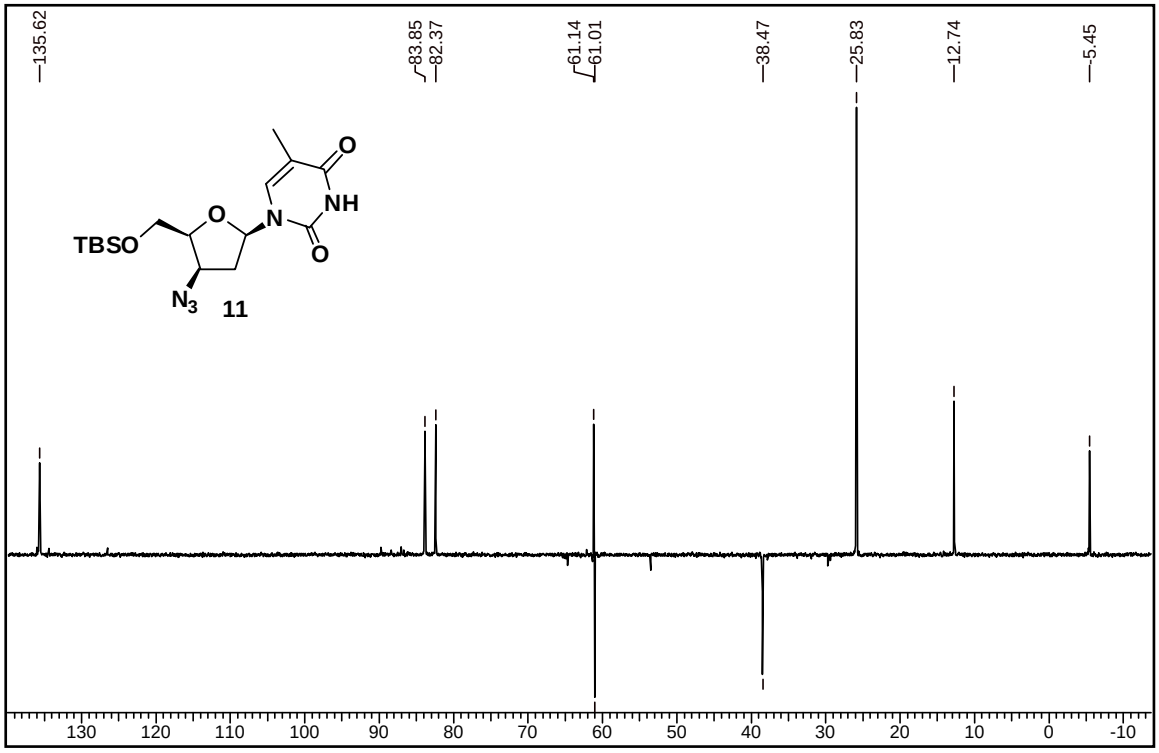

Chloroform-d
TMS
7.77
7.74
7.62
7.59
7.43
7.39
7.35
7.31
7.27
6.79
6.76
6.70
5.70
5.67
4.49
4.42
4.38
4.25
4.22
4.19
3.49
3.39
3.31
3.28
3.20
2.88
2.59
2.25
2.05
1.91
0.00
HO
O
S
O
N
NH
O
O
FmocHN
8
2.52
5.91
0.74
0.93
2.68
2.04
1.97
1.62
1.12
3.00
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

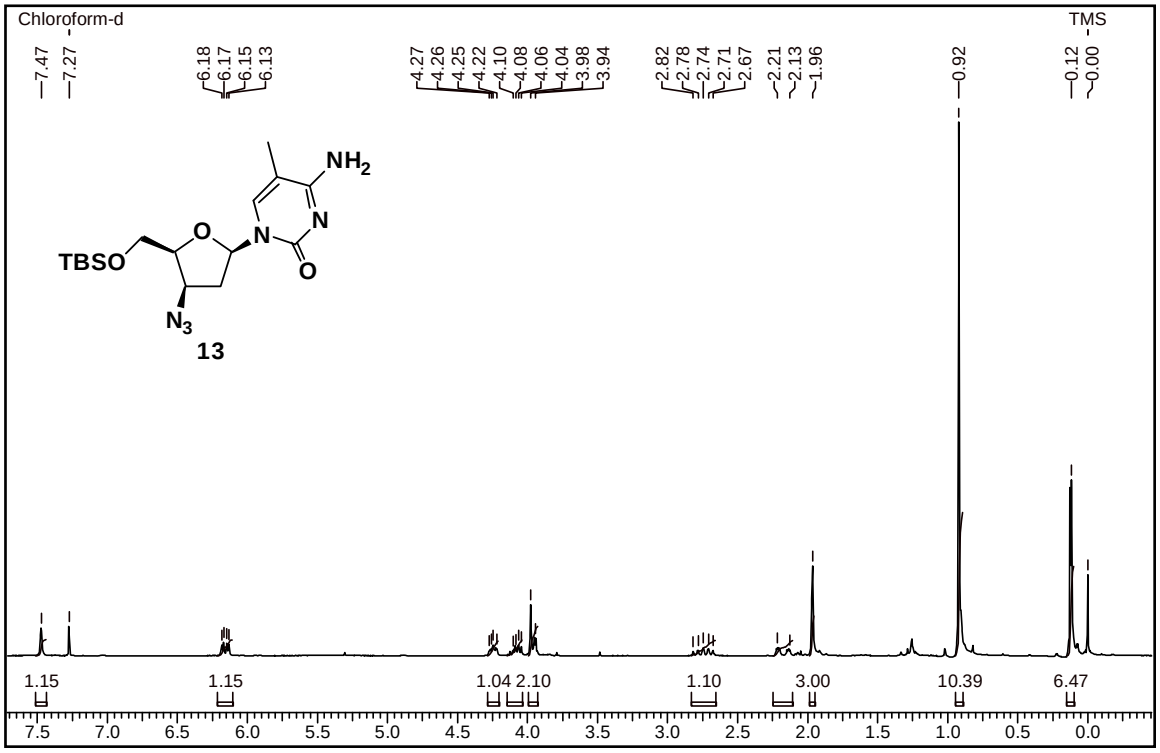

Chloroform-d
172.34
164.40
156.08
150.23
143.59
141.03
127.55
126.92
124.92
119.76
110.62
80.96
77.64
77.00
76.36
66.41
52.06
46.98
34.19
28.72
25.20
12.12
HO
O
S
O
N
NH
O
O
FmocHN
8
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

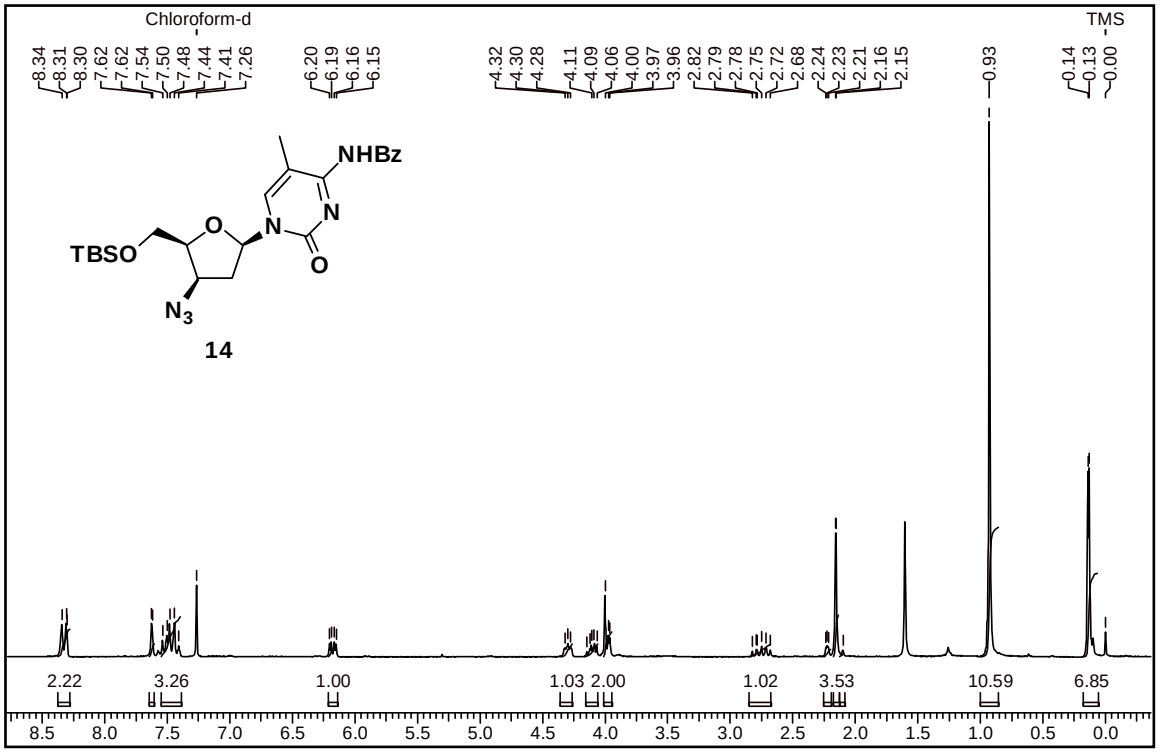

CDCl3
TMS
9.42
7.55
7.27
6.40
6.38
4.46
4.44
4.07
3.87
3.86
3.80
2.51
2.43
2.37
2.35
2.30
2.09
1.91
0.92
0.11
0.11
0.00
O
NH
N
O
O
TBSO
HO
9
0.85
1.05
0.88
0.99
1.90
1.27
3.00
9.63
6.50

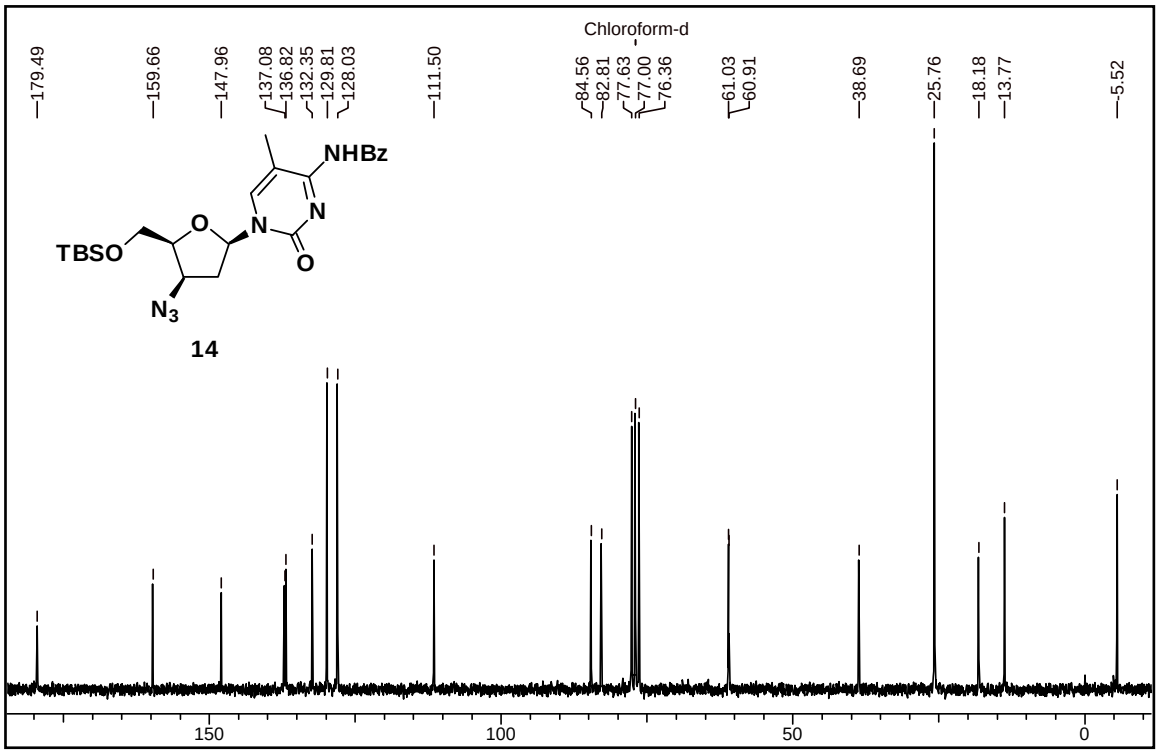

CDCl3
TMS
8.96
7.47
7.27
6.21
6.19
6.17
6.15
4.30
4.28
4.26
4.09
4.03
4.01
3.97
3.93
2.80
2.76
2.72
2.69
2.65
2.15
2.08
1.96
0.92
0.12
0.00
O
NH
N
O
O
TBSO
N3
11
0.85
0.94
0.92
3.14
0.95
3.00
9.40
5.85

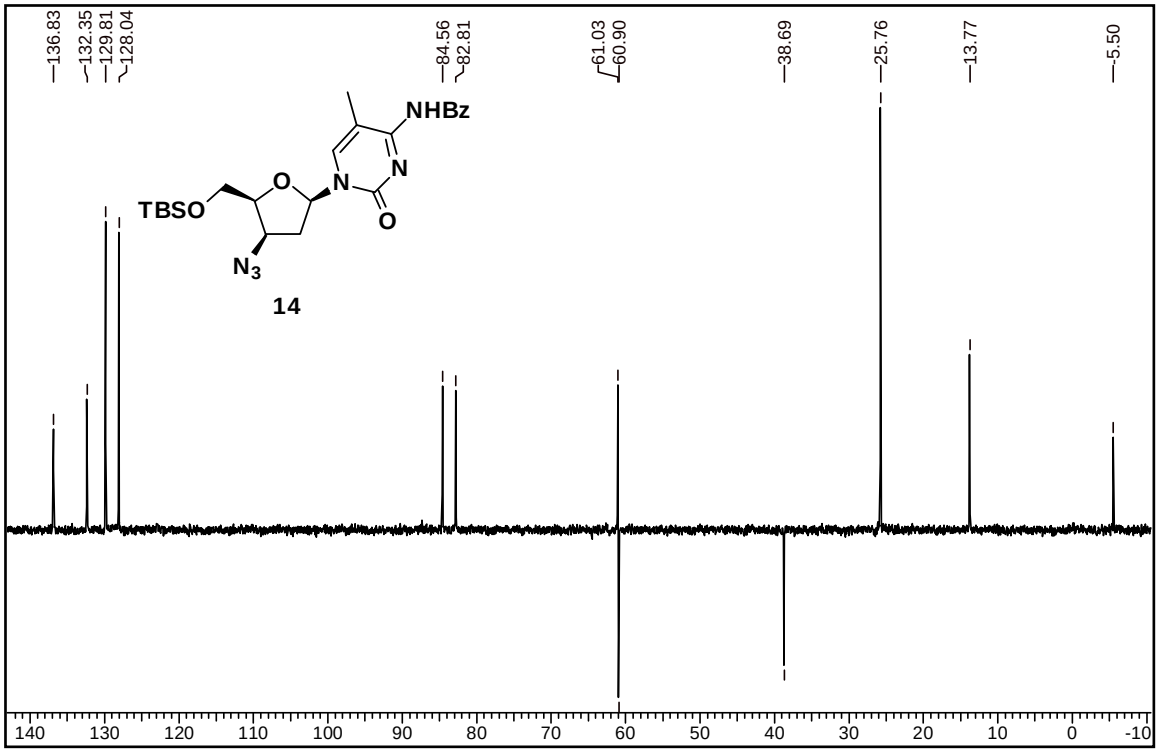

164.21
150.81
135.62
110.96
83.87
82.39
77.77
77.13
76.50
61.15
61.02
38.48
25.83
18.24
12.73
5.48
TBSO
O
NH
N
O
O
N3
11
160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20

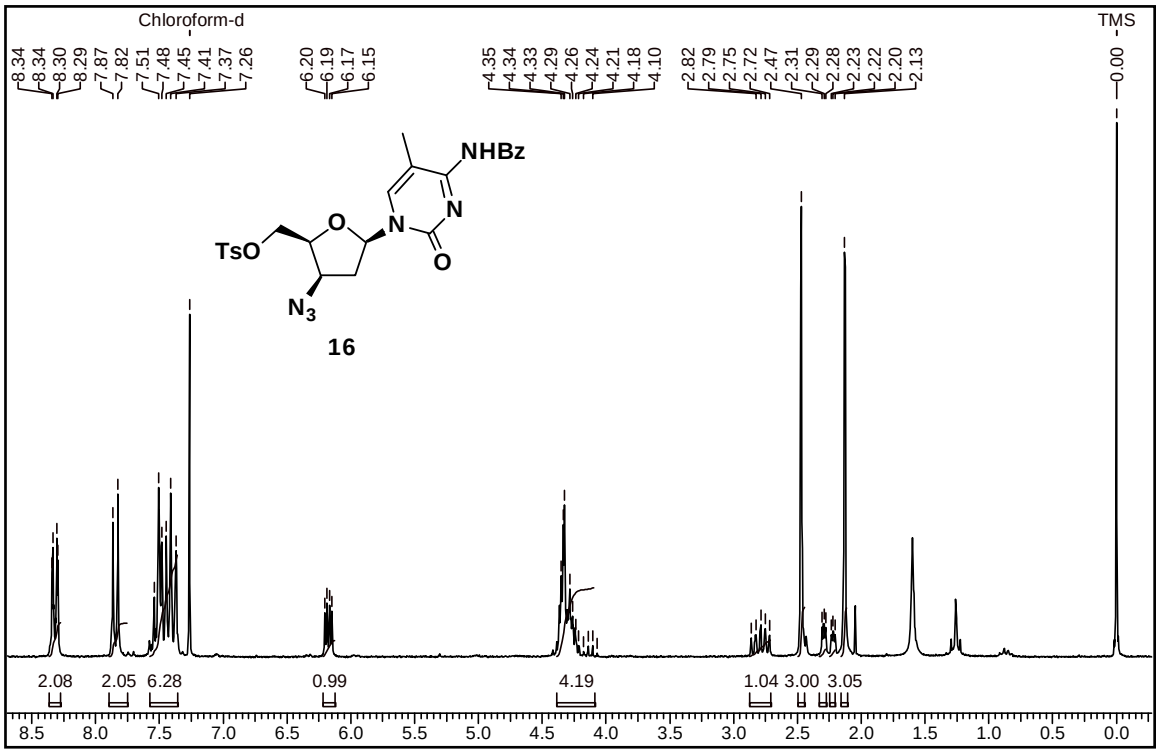

135.62
83.85
82.37
61.14
61.01
38.47
25.83
12.74
5.45
TBSO
O
NH
N
O
O
N3
11
130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10

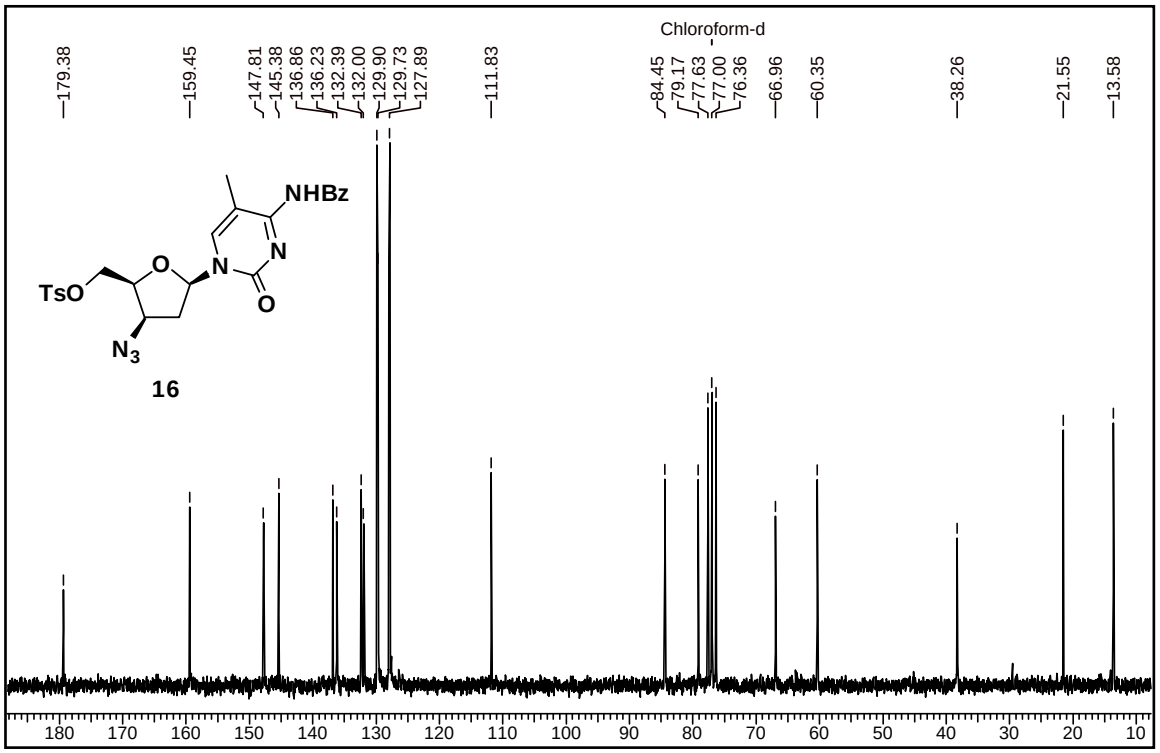

Chloroform-d
TMS
7.47
7.27
6.18
6.17
6.15
6.13
4.27
4.26
4.25
4.22
4.10
4.08
4.06
4.04
3.98
3.94
2.82
2.78
2.74
2.71
2.67
2.21
2.13
1.96
0.92
0.12
0.00
NH2
N
O
N
O
TBSO
N3
13
1.15
1.15
1.04
2.10
1.10
3.00
10.39
6.47
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

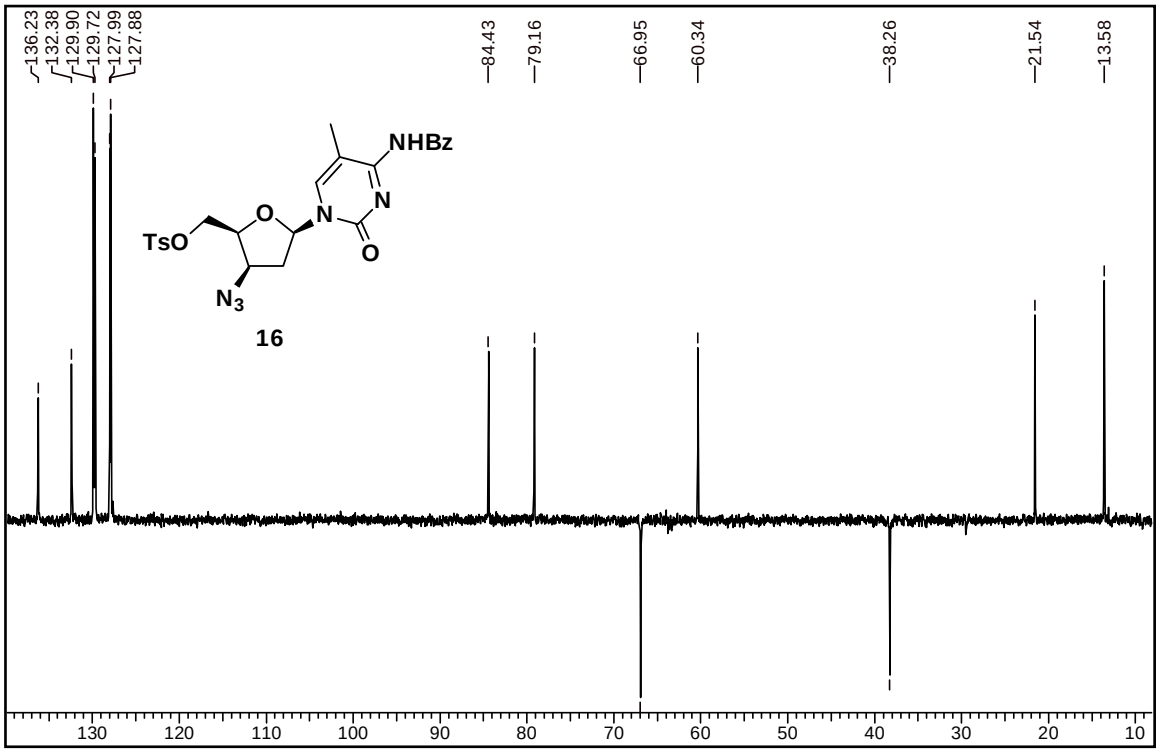

Chloroform-d
TMS
8.34
8.31
8.30
7.62
7.62
7.54
7.50
7.48
7.44
7.41
7.26
6.20
6.19
6.16
6.15
4.32
4.30
4.28
4.11
4.09
4.06
4.00
3.97
3.96
2.82
2.79
2.78
2.75
2.72
2.68
2.24
2.23
2.21
2.16
2.15
0.93
0.14
0.13
0.00
NHBz
N
O
N
O
TBSO
N3
14
2.22
3.26
1.00
1.03
2.00
1.02
3.53
10.59
6.85
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

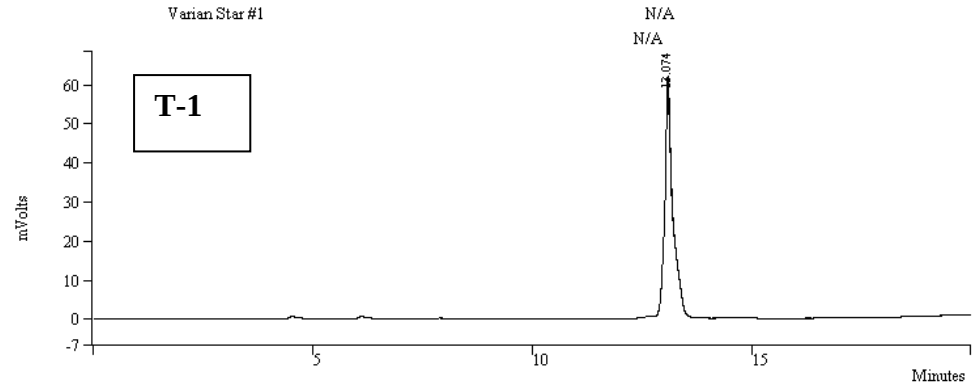

Chloroform-d
179.49
159.66
147.96
137.08
136.82
132.35
129.81
128.03
111.50
84.56
82.81
77.63
77.00
76.36
61.03
60.91
38.69
25.76
18.18
13.77
-5.52
NHBz
TBSO
O
N
N
O
N3
14
150
100
50
0

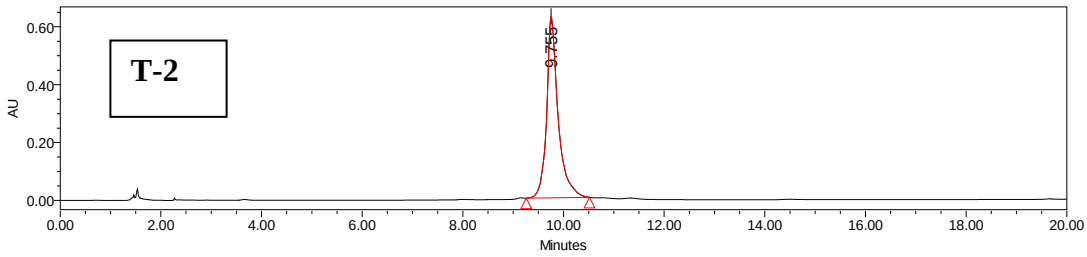

136.83
132.35
129.81
128.04
84.56
82.81
61.03
60.90
38.69
25.76
13.77
-5.50
NHBz
TBSO
O
N
N
O
N3
14
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

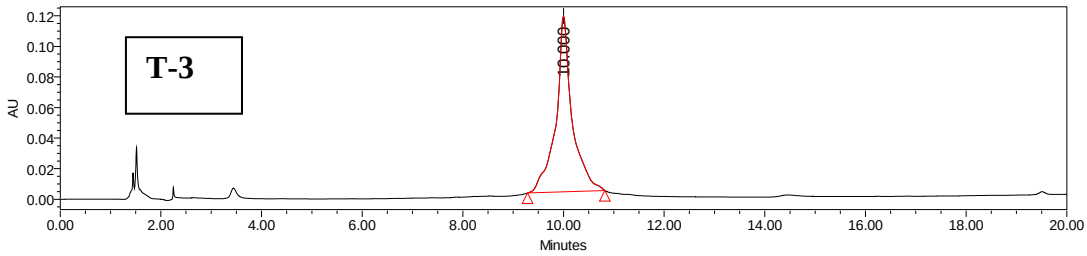

Chloroform-d
TMS
8.34
8.34
8.30
8.29
7.87
7.82
7.51
7.48
7.45
7.41
7.37
7.26
6.20
6.19
6.17
6.15
4.35
4.34
4.33
4.29
4.26
4.24
4.21
4.18
4.10
2.82
2.79
2.75
2.72
2.47
2.31
2.29
2.28
2.23
2.22
2.20
2.13
0.00
NHBz
N
N
O
N
O
TsO
N3
16
2.08
2.05
6.28
0.99
4.19
1.04
3.00
3.05
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

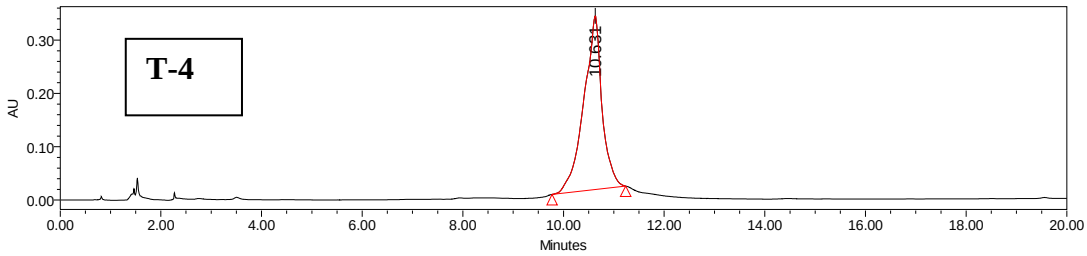

Chloroform-d
179.38
159.45
147.81
145.38
136.86
136.23
132.39
132.00
129.90
129.73
127.89
111.83
84.45
79.17
77.63
77.00
76.36
66.96
60.35
38.26
21.55
13.58
NHBz
N
N
O
N
O
TsO
N3
16
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10

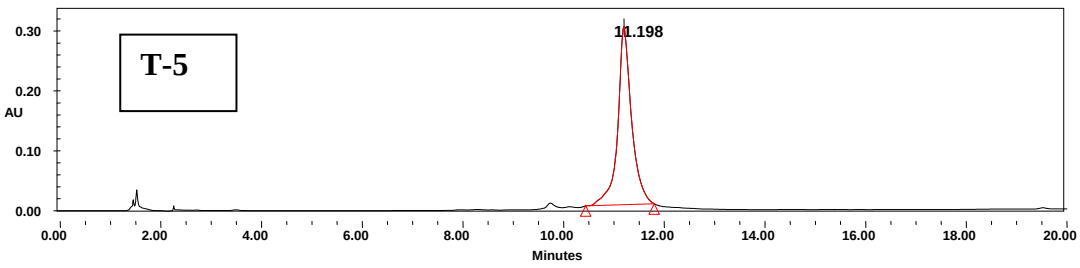

136.23
132.38
129.90
129.72
127.99
127.88
84.43
79.16
66.95
60.34
38.26
21.54
13.58
NHBz
N
N
O
N
O
TsO
N3
16
130
120
110
100
90
80
70
60
50
40
30
20
10

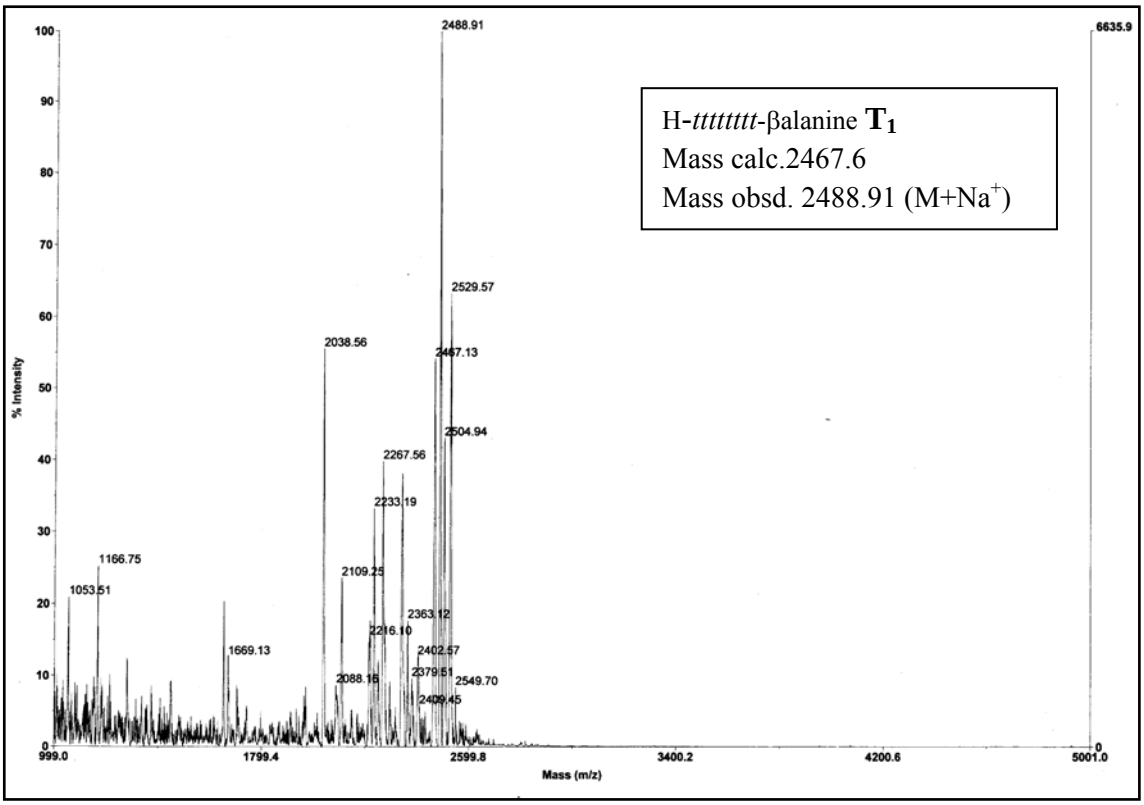

Varian Star #1
N/A
N/A
T-1
mVolts
Minutes

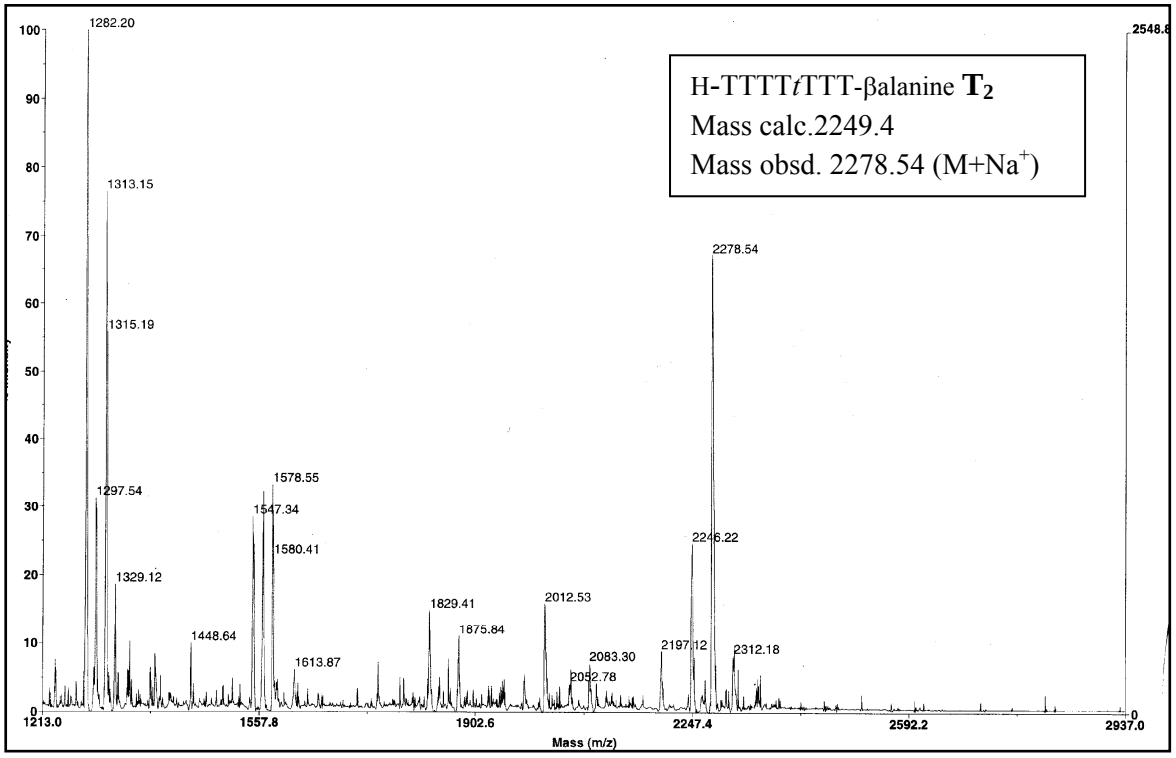

T-2
9.756
AU
Minutes

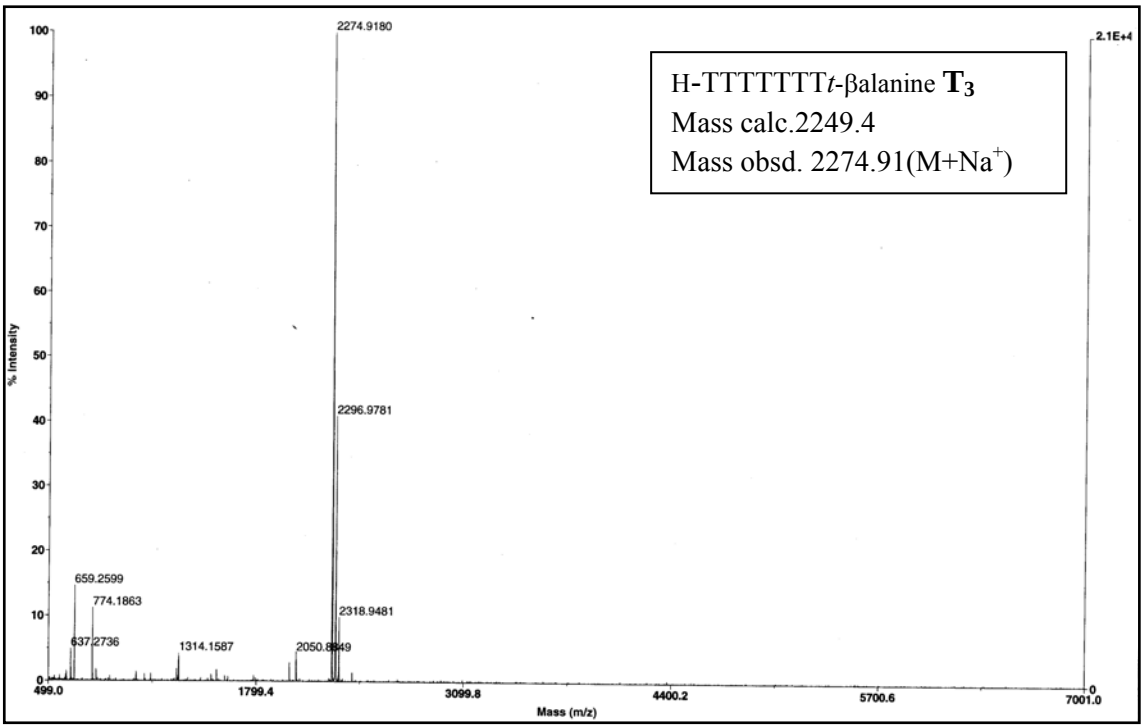

T-3
10.000
AU
Minutes

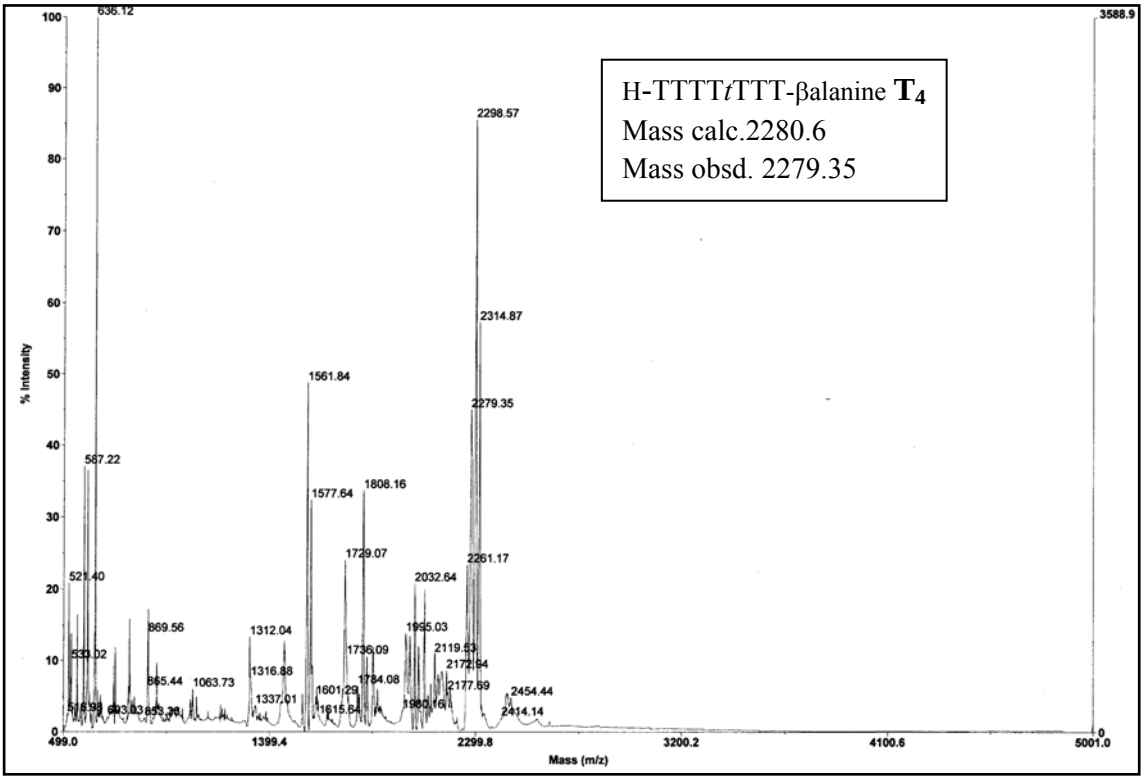

T-4
10.631
AU
Minutes

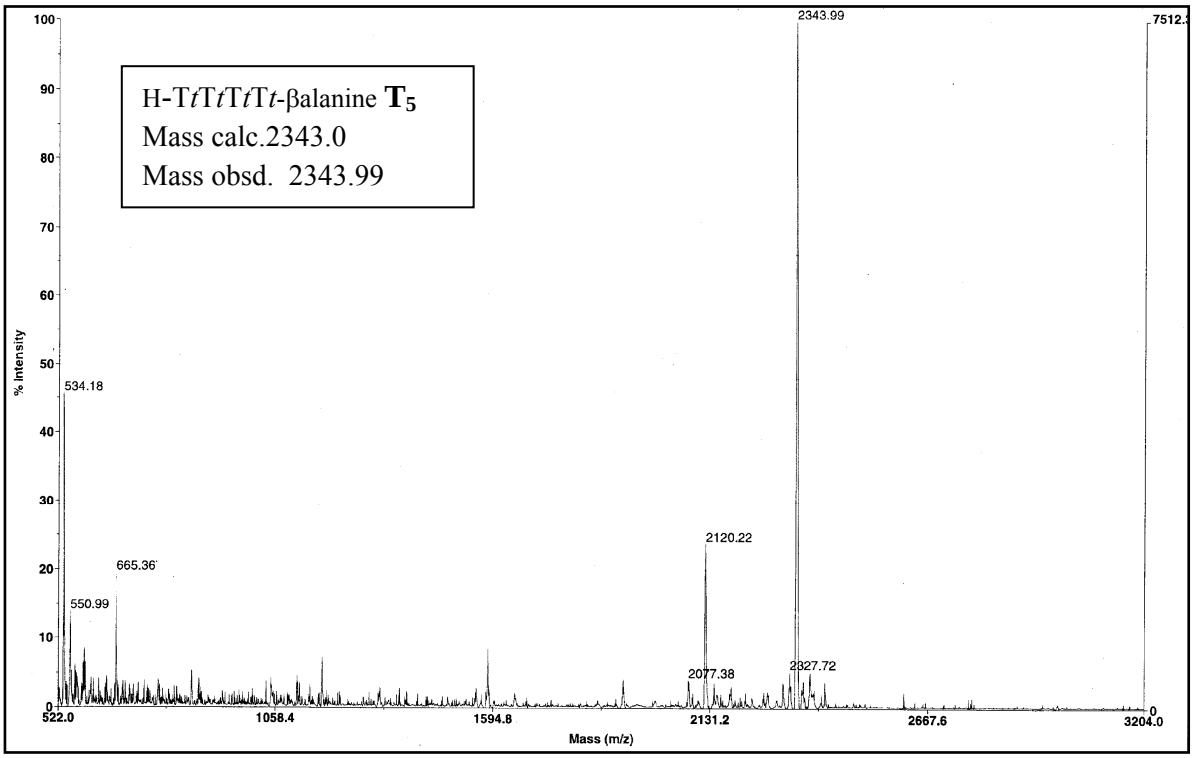

T-5
11.198
AU
Minutes


H-*ttttttt*-βalanine **T1**
Mass calc.2467.6
Mass obsd. 2488.91 (M+Na+)
% Intensity
Mass (m/z)
2488.91
6635.9
2529.57
2038.56
2467.13
2504.94
2267.56
2233.19
1166.75
1053.51
2109.25
2363.12
2216.10
1669.13
2402.57
2379.51
2549.70
2088.18
2409.45
999.0
1799.4
2599.8
3400.2
4200.6
5001.0


H-TTTT*t*TTT-βalanine **T2**
Mass calc.2249.4
Mass obsd. 2278.54 (M+Na+)
Mass (m/z)
1282.20
2548.8
1313.15
2278.54
1315.19
1578.55
1297.54
1547.34
1580.41
2246.22
1329.12
1829.41
2012.53
1448.64
1875.84
2197.12
2312.18
1613.87
2083.30
2052.78
1213.0
1557.8
1902.6
2247.4
2592.2
2937.0


H-TTTTTTT*t*-βalanine $\mathbf{T_3}$
Mass calc.2249.4
Mass obsd. 2274.91(M+Na$^+$)
2274.9180
2296.9781
2318.9481
659.2599
774.1863
637.2736
1314.1587
2050.8849
% Intensity
Mass (m/z)


H-TTTT*t*TTT-βalanine $\mathbf{T_4}$
Mass calc.2280.6
Mass obsd. 2279.35
636.12
2298.57
2314.87
1561.84
2279.35
587.22
1808.16
1577.64
1729.07
2261.17
2032.64
869.56
1312.04
1316.88
1063.73
1337.01
1601.29
1784.08
2454.44
2414.14
% Intensity
Mass (m/z)


H-TtTtTtTt-βalanine T5
Mass calc.2343.0
Mass obsd. 2343.99
% Intensity
Mass (m/z)
534.18
665.36
550.99
2120.22
2077.38
2327.72
2343.99
522.0
1058.4
1594.8
2131.2
2667.6
3204.0

3.2.9 References

1 Gogoi**,** K**.;** Gunjal, A. D.; Kumar, V. A. Sugar-thioacetamide backbone in oligodeoxyribonucleosides for specific recognition of nucleic acids. *Chem. Commun.,* **2006**, 2373-2375.

2 Mastuda, A.; Watanabe, K. A.; Fox, J. J. Nucleosides. 115. Reaction of 3'-*O*-mesylthymidine. Formation of 1-(3-azido-2,3-dideoxy-*β*-D-thero-pentofuranosyl)thymine and its conversion into 6,3'-imino-1-(2,3-dideoxy--*β*-D-thero-pentofuranosyl)thymine *J. Org. Chem.* **1980**, *45*, 3274-3278.

3 Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. Synthesis of peptide nucleic acid monomers containing the four natural bases: Thymine, cytosine, adenine and guanine and their oligomerization *J. Org. Chem.* **1994**, *59*, 5767-5773.

4 Egholm, M.; Buchardt, O.; Nielsen, P. E. Peptide nucleic acids (PNAs): oligonucleotide analogues with achiral peptide backbone . *J. Am. Chem. Soc*. **1992**, *114*, 1895-1897.

5 Erickson, B. W.; Merrifield, R. B. Solid Phase Peptide Synthesis. In the Proteins Vol. II, 3rded.; Neurath, H. and Hill, R. L. eds.; Academic Press, New York, **1976**, pp 255. b) Merrifield, R. B.; Stewart, J. M.; Jernberg, N: Instrument for automated synthesis of peptides. *Anal. Chem*. **1966**, *38*, 1905-1914.

6 Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color test for detection of free terminal amino groups in solid-phase synthesis of peptides *Anal. Biochem.* **1970**, *34*, 595