SYNTHESIS AND BIOPHYSICAL STUDIES OF PNA AND CHIMERIC PNA-DNA ANTISENSE OLIGOMERS WITH FIVE ATOM LINKAGES

THESIS SUBMITTED TO

THE UNIVERSITY OF PUNE

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN CHEMISTRY

BY

KHIRUD GOGOI

Dr. (Mrs.) VAIJAYANTI A. KUMAR

(RESEARCH GUIDE)

ORGANIC CHEMISTRY DIVISION

NATIONAL CHEMICAL LABORATORY

PUNE 411 008

INDIA

DECEMBER 2007

CERTIFICATE

This is to certify that the work presented in the thesis entitled **"Synthesis and biophysical studies of PNA and chimeric PNA-DNA antisense oligomers with five atom linkages"** submitted by Khirud Gogoi, 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 **December 2007**

Research guide

Division of Organic Chemistry

National Chemical Laboratory

Pune 411008

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled " **Synthesis and biophysical studies of PNA and chimeric PNA-DNA antisense oligomers with five atom linkages**", submitted for the Degree of ***Doctor of Philosophy*** to the University of Pune, has been carried out by me at Organic Chemistry Division, National Chemical Laboratory, Pune 411 008, India, under the supervision of Dr. (Mrs.) Vaijayanti A. Kumar. The work is original and has not been submitted in part or full by me for any other degree or diploma to this or any other University.

Khirud Gogoi

National Chemical Laboratory
Pune-411008
December 2007

Dedicated to MY

Parents

ACKNOWLEDGEMENT

It gives me an immense pleasure to express my deep sense of gratitude towards my research guide Dr. (Mrs.) Vaijayanti A. Kumar for all the advice, guidance, support and encouragement during every stage of this work. She made me realize the importance of doing quality research; she taught me each and every aspect of 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. Although this eulogy is insufficient, I preserve an everlasting gratitude for her.

I am very much grateful to Prof. K. N. Ganesh for his encouragement; understanding and suggestions were invaluable during my stay in the laboratory.

My special thanks to Mrs. Anita Gunjal, Dr. A. A. Natu and Dr. (Mrs.) Vandana S. Pore for their help and encouragement during the course of this work.

I am also thankful to Mrs. M. V. Mane and Mrs. S. S. Kunte for HPLC analysis, Dr.Mahesh J. Kulkarni for mass analysis. The kind support from NMR group is greatly acknowledged and I am especially thankful to Mrs. Usha D. Phalgune. My sincere thanks to Dr. N. P. Argade, Dr. Anil Kumar, Dr. V. R. Pedireddy, Dr. B. G. Hazra, Dr. Saurav Pal, Dr. P. L. Joshi, Dr. R. H. Naik and Dr. Dilip D. Dhavale for their advice and various kinds of help. Special thanks go to Mr. S. S. Deo and Mr. Prashant Mane for their valuable helps. I also take this opportunity to thank library staff, glass blowing, stores, purchase, workshop, security, and administrative groups for their support during my PhD work.

At this moment, I remember few of my former teachers, Mr. Amar Jyoti Borthakur, Mr. Prashanta Kumar Bordoloi, Prof. S. K. Bhattacharjee, Mr. P. K. Chetia, Prof. H. K. Das, Prof. P. J. Das, Prof. O. K. Medhi, Prof. D. K. Kakati, Prof. J. N. Ganguly, Prof. B. K. Das, Dr. P. Phukan, Dr. D. K. Das and all other my teachers as they built the foundation for this achievement.

I sincerely thank my seniors and colleagues for their help in various capacities, co-operation and maintaining cheerful atmosphere in the laboratory. Thank you Dr. Moneesha, Dr. Pallavi, Dr. Dinesh, Dr. Nagendra, Dr. Govindaraju, Dr. Praveen, Dr. Umashankara, Dr. Gourishankar, Dr. Sunil, Raman, Madhuri, Amit Patwa, Ashwini, Sridhar, Roopa, Sachin, Pradnya, Manaswini, Mahesh, Parameshwar, Seema, Namrata, Harshit, Deepti and Pankaj. I can never forget the help from my colleagues at Lab No. 201, where I stayed a fruitful year, thank you Bapu, Deepak, Namdev, Neelkantha, Sudhir Sachin and Dr. Sushmita. I also thank Mr. Pawar, Mr. Iyer, Mr. Deepak Jori and especially Mr. Pandurang D. Bhumkar for their assistance.

It gives me immense pleasure to thank Dr. Manash P. Sharma, Dr. Arindam Adikari, Dr. Satyajyoti Senapati, Dr. Ramakanta Barman, Dr. Jadab Sharma, Pranjal Kalita da, Dr. Sasanka Deka, Dr. Pranjal Barua, Sanjib da, Dr. Diganta Sarma, Lakshi da, Gitali Baidew, Sofia Baidew, Rahul, Ankur, Gupta (Upendra) and Rituda for their direct and indirect contributions in my personal and professional life during tenure in NCL.

My special thanks go to my beloved friends Pankajda, Kulakamal Senapati da, Achyutda (Dada Goswami), Anupam Pandit da, Diganta (Kapoor), Pabitra (Puwali), Partha (Baba), Medhi, Archana Bora, Suranjana Chaliha , Ranju for their individual support and encouragement during my carrier.

I wish to thank all my fellow colleagues in NCL, Bhargav, Prassana, Ravindra, Manmath (Manager), Mukesh Gupta, Prabhas, Mehraj Baag, Sailesh, Murali, Manoj, Devraj, Sachin Bhai, Rajendra,Madhuri, Girish Anegudi, Sashi Bairagi, Nilesh, Nishant Gupta, Victor, Sukhenda, Atul, Sanbhag, Debasish G, Swaroop (Swamiji), Kishore (Bhaiya), Deepak Jadav, Sunil Pandey, Puspesh, Mahesh Thakkar, Prashad, Manish, Amit Delori, Satyanarayana, Raju Reddy, Arbind, Santosh, Alpana, Rupali, Sushim, Sreekanth D., Atul Thakur, Akhilesh Tanower, Sufy Siddique, Kiran Toti, Suresh Pujari, Taterao (Maharaj), Raghupathy, Nookaraju, Bhalchandra, Joseph Jolly, Sarvesh for their cheerful company and making my NCL life very lively and enjoyable. I also am indebted to the hostel members and mess workers of G J Hostel and Jagtap family for their most essential help.

I am grateful to Council of Scientific and Industrial Research, Government of India, for awarding the junior and senior research fellowships and Dr. S. Sivaram, Director, and Dr. B. D. Kulkarni, Deputy Director, National Chemical Laboratory to carry out my research works, extending all infrastructural facilities and to submit this work in the form of a thesis for the award of Ph. D degree.

Finally, it has been a difficult task to capture and express my feelings for my family members. I have no words to express my gratitude to my Parents and brothers, without knowing much what I am doing exactly, just wishing me all the time with no expectations. Their patience and sacrifice were always a main source of inspiration and will remain throughout my life. Their blessings and encouragement have always made me an optimist in any unknown areas I had ventured.

Khirud Gogoi

CONTENTS

Publications

1. **Khirud Gogoi**, Anita D. Gunjal and Vaijayanti A. Kumar: Sugar-thioacetamide backbone in oligodeoxyribonucleosides for specific recognition of nucleic acids. ***Chem. Commun.,*** **2006**, 2373-2375. *Highlighted in RSC Chemical Biology,* **2006,** *6.*
2. **Khirud Gogoi**, Anita D. Gunjal, Usha D. Phalgune and Vaijayanti A. Kumar: Synthesis and RNA binding selectivity of oligonucleotides modified with five-atom thioacetamido nucleic acid backbone structures. ***Org. Lett.,*** **2007***, 9,* 2697-2700. *Highlighted by M. Egli at Faculty of 1000 Biology.*
3. **Khirud Gogoi**, Meenakshi V. Mane, Sunita S. Kunte and Vaijayanti A. Kumar: A versatile method for the preparation of conjugates of peptides with DNA/ PNA/ analogue by employing chemo-selective click reaction in water. ***Nucleic Acids Research,*** **2007***,* 35(21): e 139 *DOI:10.1093*/nar/gkm935.
4. **Khirud Gogoi*** and Vaijayanti A. Kumar*: Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers show sequence specific DNA/RNA recognition *Chem. Commun.,* **2007**, DOI: 10.1039/ B716835G.
5. **Khirud Gogoi**, Usha D. Phalgune and Vaijayanti A. Kumar: Conformational analysis of monomer and dimers of thioacetamide nucleic acids by NMR spectroscopy. Manuscript under preparation.

Symposia attended/ Poster presented/ Oral presentation

1. "**Peptide-PNA conjugates by employing chemo-selective click chemistry**" **Khirud Gogoi,** Meena V. Mane, Sunita S Kunte and Vaijayanti A. Kumar: Poster presentation at "Cell Penetrating Peptides" Meeting, Organized by Biochemical Society, Telford, United Kingdom, 2007.
2. **"Conformational Studies of Thioacetamido Nucleic Acids (TANA) by NMR Spectroscopy"** Usha D. Phalgune, Khirud Gogoi, Anita D.Gunjal, Vaijayanti A. Kumar and P. R. Rajamohanan: Poster Abstract accepted for publication in

National Magnetic resonance Society (NMRS) Meeting to be held at New Delhi from 16th to 19th January, 2008.

3. **"Deoxyribo-oligonucleosides: New sugar-thioacetamide backbone for specific recognition of Nucleic acids"** **Khirud Gogoi** and Vaijayanti A. Kumar. Oral Presentation at "Junior National Organic Chemistry Symposium Trust" (JNOST) symposium 2006, Jaipur, India.
4. **"Deoxyribo-Oligonucleotides: New sugar-thioacetamide backbone for specific recognition of nucleic acids"** **Khirud Gogoi** , Anita D. Gunjal and Vaijayanti A. Kumar Poster presentation at ACS-CSIR "Organic Chemistry Chemical Biology" (OCCB) conference 2006, National Chemical Laboratory, Pune, India.
5. "**Synthesis of antisense oligonucleotides and their characterization by MALDI-TOF mass spectrometry**" **Khirud Gogoi**, Mahesh J. Kulkarni and Vaijayanti A. Kumar. Poster presentation at National Mass spectrometry symposium 2007, National Chemical Laboratory, Pune, India.
6. Attended Post NOST mini symposium in Organic Chemistry 2003, National Chemical Laboratory, Pune, India, 2003.

ABSTRACT

The thesis entitled "**SYNTHESIS AND BIOPHYSICAL STUDIES OF PNA AND CHIMERIC PNA-DNA ANTISENSE OLIGOMERS WITH FIVE ATOM LINKAGES**" is divided into 5 chapters as follows:

Chapter 1: Introduction: Modified Nucleic Acids for selective RNA recognition

Novel oligonucleotide analogues that can form stable duplexes or triplexes with nucleic acids are important synthetic targets, because of their use as therapeutic agents. Various types of modified oligonucleotides have been developed over the last two decades as potential diagnostic probes and therapeutics for the antisense and antigene approach. The more recent developments such as splice correcting and exon skipping strategies require

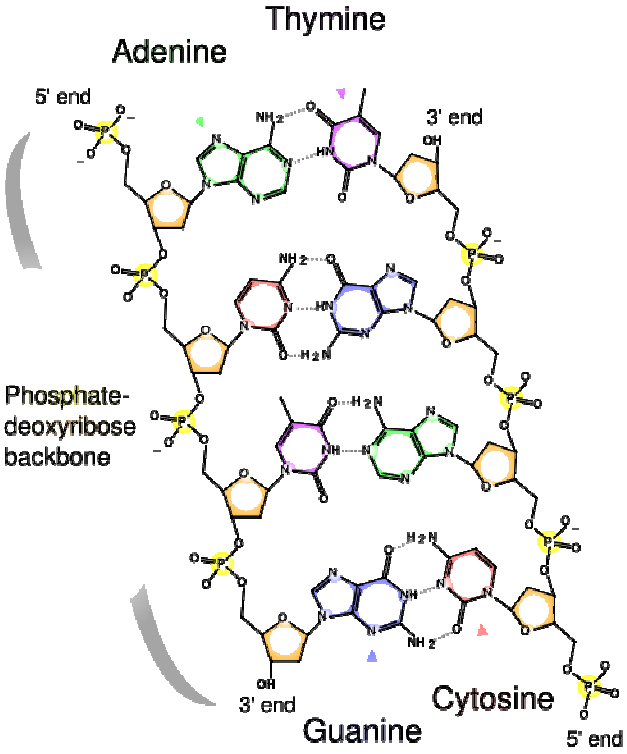


Figure 1: Chemical Structure of DNA

highly robust nucleic acid analogues that are stable under physiological conditions as single strands as well as in the form of duplexes with complementary RNA sequences.

The naturally occurring deoxyribose- (DNA) and ribose (RNA) sugar-phosphate backbones are endowed with considerable differences in their binding affinities towards

themselves. This occurs because of the different sugar conformations prevalent in DNA and RNA and the subtle structural changes accruing from these in hydrogen bonding, base-stacking interactions and hydration of major/minor grooves. The 4-atom phosphodiester linkages and pentose-sugar give immense opportunities for chemical modifications that lead to several backbone modified nucleic acid structures. Changes in the sugar phosphate backbone invariably bring about changes in the complementation properties of the nucleic acids. This chapter is focused on such modifications that impart RNA-selective binding properties to the modified nucleic acid mimics and the rationale behind the said selectivity. It is found that the six-atom sugar-phosphate backbone could be replaced by either one-atom extended or one-atom edited repeating units, leading to the folded or extended geometries to maintain the internucleoside distance-complementarity. Other important contributions come from electronegativity of the substituent groups, hydration in the major/minor groove, base stacking etc.

Chapter 2: **Thioacetamido Nucleic Acids (TANA): Synthesis and biophysical studies**

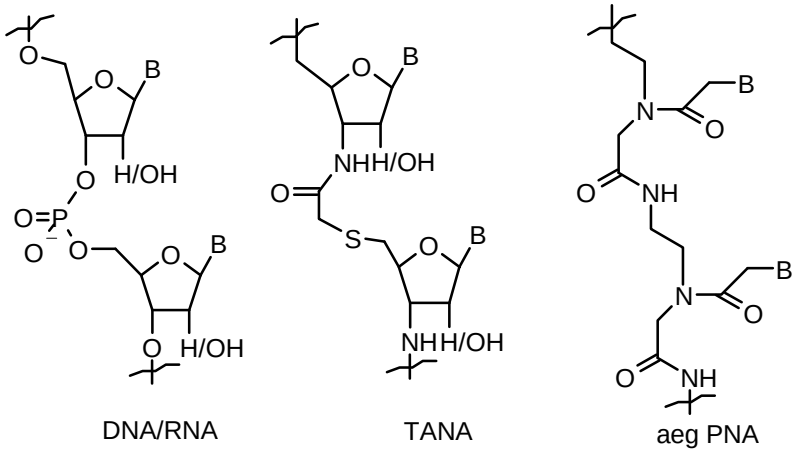


Figure 2: Chemical Structure of DNA/ RNA, Thioacetamido Nucleic Acid (TANA) and Peptide nucleic Acid (PNA).

The use of modified nucleic acids as gene-targeted drugs and as tools in molecular biology is a developing field. The designed ONs, to be effective leads as drugs, need to have enhanced strength of hybridization with target RNA, aqueous solubility, efficiency of cellular uptake and should be accessible by easy synthetic methodologies. Apart from

FDA approved phosphorothioate backbone, the other backbone modifications being developed presently include the second and third generation antisense ONs such as 2'-O-alkyl, 2'-O-methylthioethyl, LNA, HeNA, morpholino NA, aminoethylglycyl peptide nucleic acids (PNA) or several PNA modifications. Among the sugar-amide backbones, there are many examples in the literature suggesting that a five-atom amide linker leading to a seven atom repeating backbone may be more useful because of the reduced conformational flexibility of the amide relative to the six-atom phosphodiester backbone. This postulate has also been supported by X-Ray studies. Some of the analogues with extended seven-atom backbone cross-pair with RNA with high affinity compared to DNA. In the search for an ideal backbone, we comply that 7-atom extended sugar-amide backbone could be more uniform with respect to nucleobase orientation as well as distance complementarity to bind complementary RNA sequences. In this chapter we present the synthesis of thioacetamido nucleic acids (TANA) and the thermal stability studies with complementary DNA and RNA sequences. The strategy of the design, synthesis of the monomer blocks, oligomer synthesis and their complementary RNA recognition using UV-Tm measurements is discussed.

Section A: 2A.1 Rationale of the Design

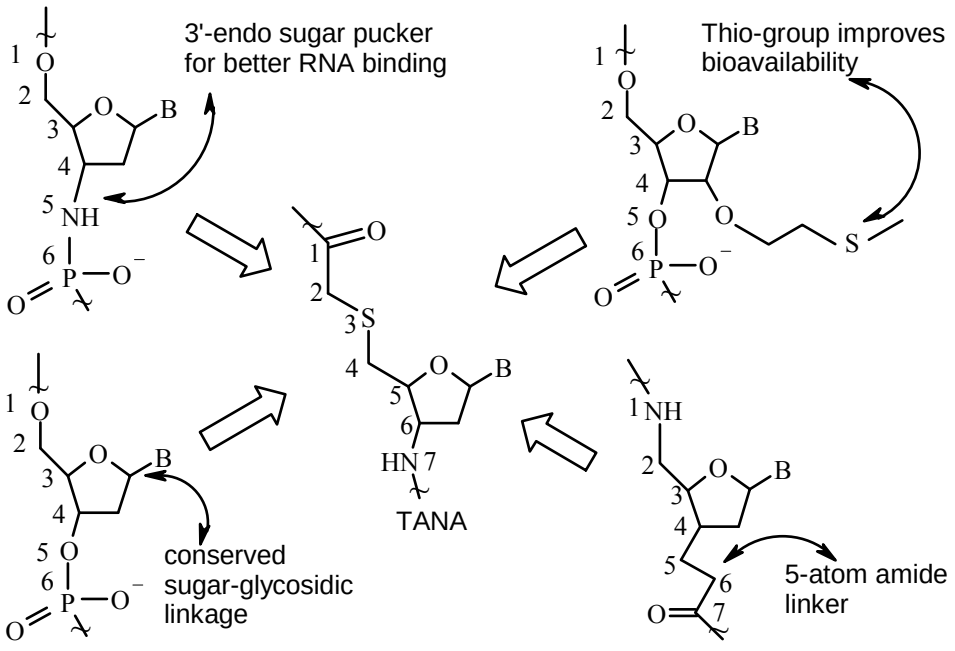


Figure 3: Design of Thioacetamido Nucleic Acid

Thioaceamido nucleic acid (TANA) is a designed DNA/RNA analogue in which the 4-atom phosphodiester linkage is replaced by a 5-atom mercaptoacetamide linkage. The 2'-deoxyribose sugar-phosphate backbone in DNA is replaced by the sugar-thioacetamido backbone considering that 1) In 3'deoxy-3'amino ribose sugar the five-membered heterocyclic ring pucker is preferred to be 3'-endo that is better suited for mRNA recognition 2) The thio functionality in the backbone may facilitate biodistribution of ONs 3) The linker group has no chiral center and the flexibility of the phosphate group in DNA/RNA may be conserved by using a five-atom amide linker 4) The backbone is now uncharged as in PNA 5) TANA homo oligomers as well as TANA-PNA chimeric backbone oligomers may be evaluated.

2A.2: Synthesis of TANA monomers:

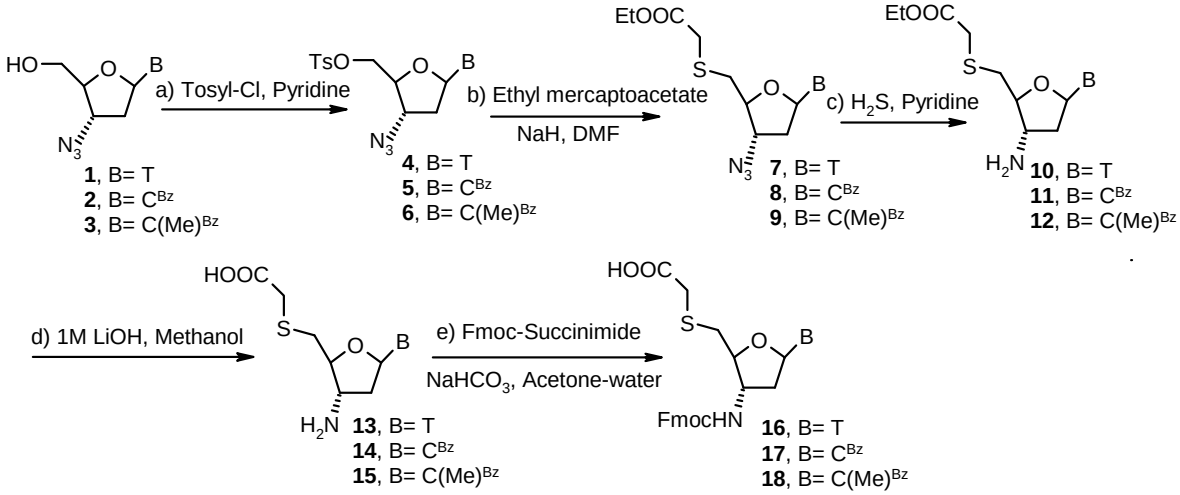


Figure 4: Synthesis of TANA monomers

The synthesis starts from the starting materials 3'-deoxy azido nucleoside, converted to nucleoside 2', 3', 5'-deoxy 3'-amino 5'-thioacetic acid in a simple synthetic pathway shown in Figure **4**.

2A.3. Synthesis of the oligomers:

To test the sequence specific DNA/RNA recognition by the oligomers comprising of the sugar-amino acids, various TANA homo oligomers as well as chimeric TANA-PNA were

synthesized using rink amide resin and Fmoc peptide synthesis strategy. A schematic representation is shown in Figure **5**. All the oligomers were purified by RP-HPLC and characterized by MALDI-TOF mass spectrometry (Table**1**).

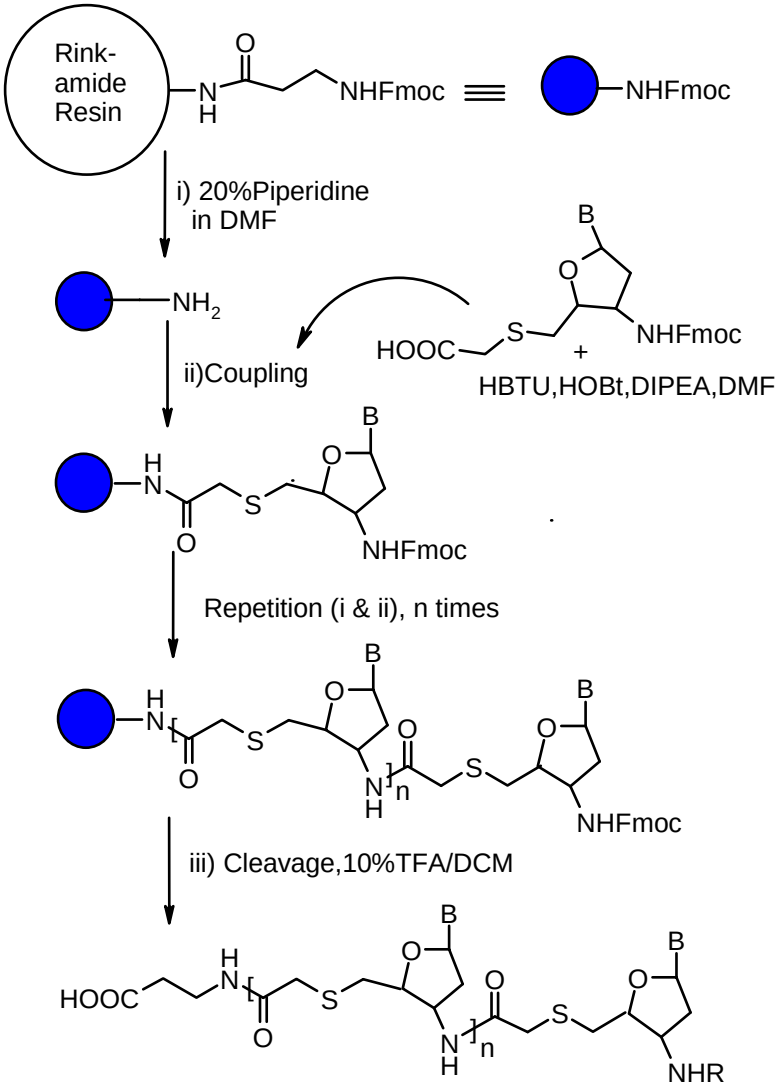


Figure 5. Solid-phase synthesis of TANA Oligomers

2A.4. UV-melting studies of TANA oligomers with DNA and RNA

To investigate the binding selectivity, specificity and discrimination of TANA oligomers towards complementary DNA and RNA, first the stoichiometry of the TANA: DNA and TANA: RNA were determined using Job's method. There was no binding found with complementary DNA. The UV-melting studies were carried out with all the synthesized oligomers and the *Tm* data was compared with the control DNA as well as PNA

oligomers (Table **1**). The stability of the TANA duplexes with DNA/RNA was also studied (Table**1**).

The pyrimidine TANA oligomers were found to impart an unprecedented selectivity for binding RNA over DNA for duplex and triplex formation. The homo TANA oligomers were found to form more stable complexes with RNA than the chimeric TANA-PNA oligomers.

Table 1. TANA Oligomers synthesized, their HPLC and Mass characterization and UV- melting data with complementary DNA and RNA

No	TANA Sequence (Mass calculated/Observed), $t_{R(min)}$	TANA:DNA T_m °C	TANA:RNA T_m °C	TANA:mismatch RNA T_m °C
1	H-**t t t t t t t t**-β ala **19** (2467.6/2491.0), 12.7	**19:29** nd	**19:30** 63.8 (63.5)	**19:31** 49.8
2	H-**t t t t c^m t t t**-β ala **20** (2466.7 /2487.9), 12.5	**20:34** nd	**20:31** 63.1(67.1)	**20:30** 52.9
3	H-t t t t t t **t**-β ala **21** (2249.4/2251.7), 10.0	**21:29** nd	**21:30** 53.5	-
4	H-t t t t **t** t t t-β ala **22** (2249.4/2252.6), 8.3	**22:29** nd	**22:29** 32.4	-
5	H-t **t** t t t t t **t** -β ala **23** (2343.0/2367.9), 10.0	**23:29** nd	**23:30** 25.5	-
6	H-t t t t t t t-β ala **24** (2218.2/2218.9), 7.6	**24:29** 42.6	**24:30** 61.5	**24:31** 52.5
7	3'T T T T T T T T-5' **25**	**25:29** nd	**25:30** 20	-
8	3' T T C T T T C T C T 5' **26**	**26:32** 23.6	**26:33** 25.4	-
9	H-**t t c t t t c t c t**- Lys **27** (3074.5/ 3074.8),	**27:32** nd	**27:33** 56.6	-
10	H-**t t c^m t t t c^m t c^m t**-Lys **28** (3116.57/ 3117.1),	**28:32** nd	**28:33** 57.2	-

t, c, c^m TANA backbone**,** t *aeg*PNA backbone, A,T,G,C DNA backbone. **29** 5'GC AAAAAAAACG 3', **30** r(5'GCAAAAAAAACG3'), **31** r(5'GCAAAACAAACG3'), **32** 3' AGAGAAAGAA 5' **33** r(5' AAGAAAGAGA 3'), **34** 5'GCAAAACAAACG3'. The expriments were performed in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM).Values in the parenthesis indicates *Tm* at pH 5.5.

Section B

2B.1 TANA/ DNA chimeric backbone

As we have discussed in the section A, the homooligomeric pyrimidine TANA ONs were found to bind to complementary RNA sequences significantly better than their DNA counterparts and the binding efficiency was found to be as good as PNA itself. To study the compatibility of the TANA backbone in a regular phosphodiester backbone, we have synthesized and incorporated the thymidine and thymidine-cytidine dimer blocks (**t**s**t** and

cst) connected with a five-atom amide-linker N3'-CO-CH_2-S-CH_2 (TANA) into oligomers. The assessment of the compatibility of the TANA dimer blocks in sugar-phosphate backbone was studied by UV-T_m measurements of the resulting mixed-backbone ON complexes with DNA and RNA.

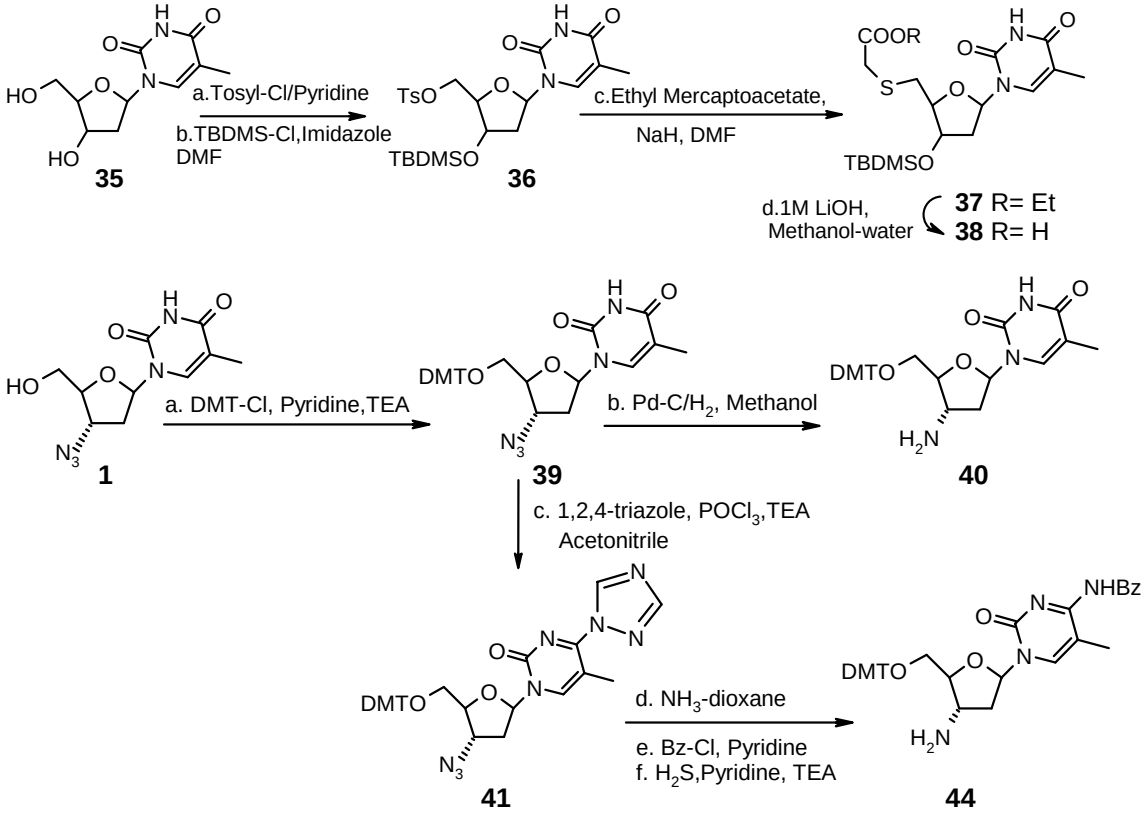


Figure 6. Synthesis of acid and amine synthons.

2B.2 Synthesis of TANA dimer building blocks:

Thymidine **35** was selectively tosylated at 5' position and the 3'-OH was then protected as 3'-*O*-TBDMS group. Treatment of 3'-*O*-protected 5'-*O*-tosyl- thymidine with ethyl mercaptoacetate followed by ester hydrolysis gave the acid synthon **38** (Figure **6**). The amine synthon of thymidine was synthesized from 3'-azidothymidine **1** by protection of the free hydroxy group as 5'-*O*-DMT and reduction of azide to amine in order to get **40**. The 5'-*O*-DMT, 3'-azidothymidine **39** was converted to cytidine derivative by known procedures *via* C4-triazolide **41** followed by amination, then benzoylation and reduction of the azide to give the desired cytidine derivative **44** (Figure **6**). The synthesis of the

dimer building blocks **tst 49** and **cst 50** was then achieved by using peptide coupling chemistry from the monomer units, deprotection and phosphitylation (Figure **7**).

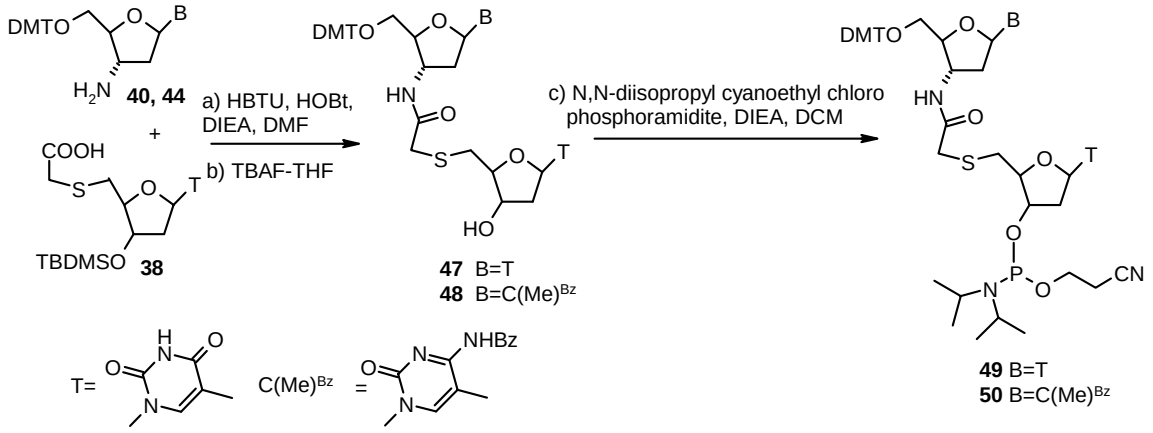


Figure 7: Synthesis of phosphoramidite building blocks.

2B.3 Synthesis of oligomers

A series of chimeric ONs (Table **2** and **3**) containing one to four **tst** and **cst** blocks were synthesized by automated solid phase synthesis using phosphoramidite approach and Applied Biosystems 3900 DNA Synthesizer. After cleavage from the support, the oligomers were purified by gel filtration and reverse phase HPLC. The purity of the oligomers was checked by reverse phase HPLC analysis on a C18 column and characterized by mass spectrometry.

The ONs synthesized were tested for their binding affinity to complementary DNA and RNA sequences in thermal denaturation-UV measurement experiments and the data is summarized in Tables **2** and **3**.

Table 2. Modified ON-pyrimidine sequences and their characteristic molecular mass and melting temperatures of complexes with complementary DNA and RNA

	ON Sequences Mass (calculated/observed)	ON:DNA T_m °C	ON:RNA T_m °C
1	5' CGTTTTTTTGC **51**	**33:63** 40	**33:64** 32.0 (- 8.0)[b]
2	5' CG TT**tst**TTT TGC **52** (3606.5/3606.6)	**34:63** 23.7	**34:64** 32.3 (+ 8.6)
3	5' CGTT **tst** TT **tst** GC **53** (3596.6/5597.0)	**35:63** nd	**35:64** 50.0
4	5' CG **tst tst tst tst** GC **54** **(**3576.8/3576.6)	**36:63** nd	**22:63** 47.8
5	5' TCT CTT TCT T **55**	**37:65** 23.6	**37:66** 25.4 (+ 0.8)

6	5’ TCT C **tst** TCT T **56**	**38:65**	**38:66**
	(2933.1/2933.1)	nd	29.6
7	5’ TCT C **tst** TC **tst** **57**	**39:65**	**39:66**
	(2923.2/2924.0)	nd	33.8
8	5’ T**cst** CTT TCTT **58**	**40:65**	**40:66**
	(2948.2/2948.1)	19.5	26.2
9	5’ T **cst** CTT T **cst** T **59**	**41:65**	**41:66**
	(2938.3/2938.7)	nd	33.8
10	5’ T**cst cst** T T **cst** T **60**	**42:65**	**42:66**
	(2928.4/2928.7)	nd	39.7
11	5’ TCA CTA GAT G **61**	**47:67**	**47:68**
		24.3	24.7 (0.4)
12	5’ TCA **cst** A GAT G 3’ **62**	**48:67**	**48:68**
	(3036.2/ 3037.0)	16.3	29.6 (+13.3)

[a] DNA (**63**, **65, 67**) and RNA (**64**, **66**, **68**) sequences **63** 5’ GCAAAAAAAACG 3’ **65** 5’AAG AAA GAG A3’ **67** 5’CAT CTA GAG A3’ **64** r (5’ GCAAAAAAAACG 3’) **66** r (5’ AAG AAA GAG A 3’), **68** r (5’CAT CTA GAG A3’) values in parenthesis indicate the difference in T_m between complexes with RNA and DNA. Experiments were performed in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM).

Table 3. Modified ON-pyrimidine sequences and their characteristic molecular mass and melting temperatures of complexes with complementary DNA and RNA

Entry	ON Sequence (Mass calculated/observed)	ON:DNA T_m °C	ON:RNA T_m °C
1	5’ GAA GGG CTT TTG AAC TCT T3’ **69**	**69:79** 53.6	**69:80**
2	5’ GAA GGG **cst**T **tst**G AA **cst cst** T3’ **70**	**69:79** nd	**70:80**
3	5’ GAA GGG **cst**T **tst**G AA CT **cst** T 3’ **71**	**71:79** 47.2	**71:80**
4	5’ AGA GTT CAA AAG CCC TTT T3’ **72**	**72:81** 56.1	**72:82**
5	5’ AGA G**tst** CAA AAG CCC **tst**T T3’ **73**	**73:81** 45.5	**73:82**
6	5’AGA G**tst** CAA AAG CC**cst tst** T3’ **74**	**74:81** 46.9	**74:82**
7	5’ CCT CTT ACC TCA GTT ACA3’ **75**	**75:83** 54.6	**75:84** 54.7 (+ 0.1)[b]
8	5’C**cst** C **tst** AC **cst** CAG **tst** ACA3’ **76**	**76:83** 37.3	**76:84** 50.6 (+13.3)
9	5’ CCT C **tst** ACC TCA G TT ACA 3’ **77**	**77:83** 39.6	**77:84** 47.5 (+ 7.9)
10	5’ CCT C **tst** ACC TCA G **tst** ACA 3’ **78**	**78:83** 43.5	**78:84** 52.8 (+ 9.3)

DNA (**79**, **81** and **83**) and RNA (**80**, **82** and **84**) sequences: **79** 5' AAG AGU UCA AAA GCC CUU C 3', **81** AAA AGG GCU UUU GAA CUC U 3', **83** 5’ TGT AAC TGA GGT AAG AGG 3’, **80** r(AAG AGU UCA AAA GCC CUU C 3'), **82** r(AAA AGG GCU UUU GAA CUC U 3'), **84** r (5’ TGT AAC TGA GGT AAG AGG 3’) Values in parenthesis indicate the difference in T_m between complexes with RNA/ DNA. Experiments were performed in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM).

The data shows that the **t**s**t** and **c**s**t** dimer blocks are compatible in the DNA backbone to selectively stabilze the ON: RNA complexes. The RNA selectivity of binding seems to be arising from the extended backbone linker that is probably inherently folded to be competent to bind to RNA over DNA as was found with the reported five-atom linked ON analogues.

Chapter 3: Conformational studies of TANA dimers and monomers by NMR spectroscopy

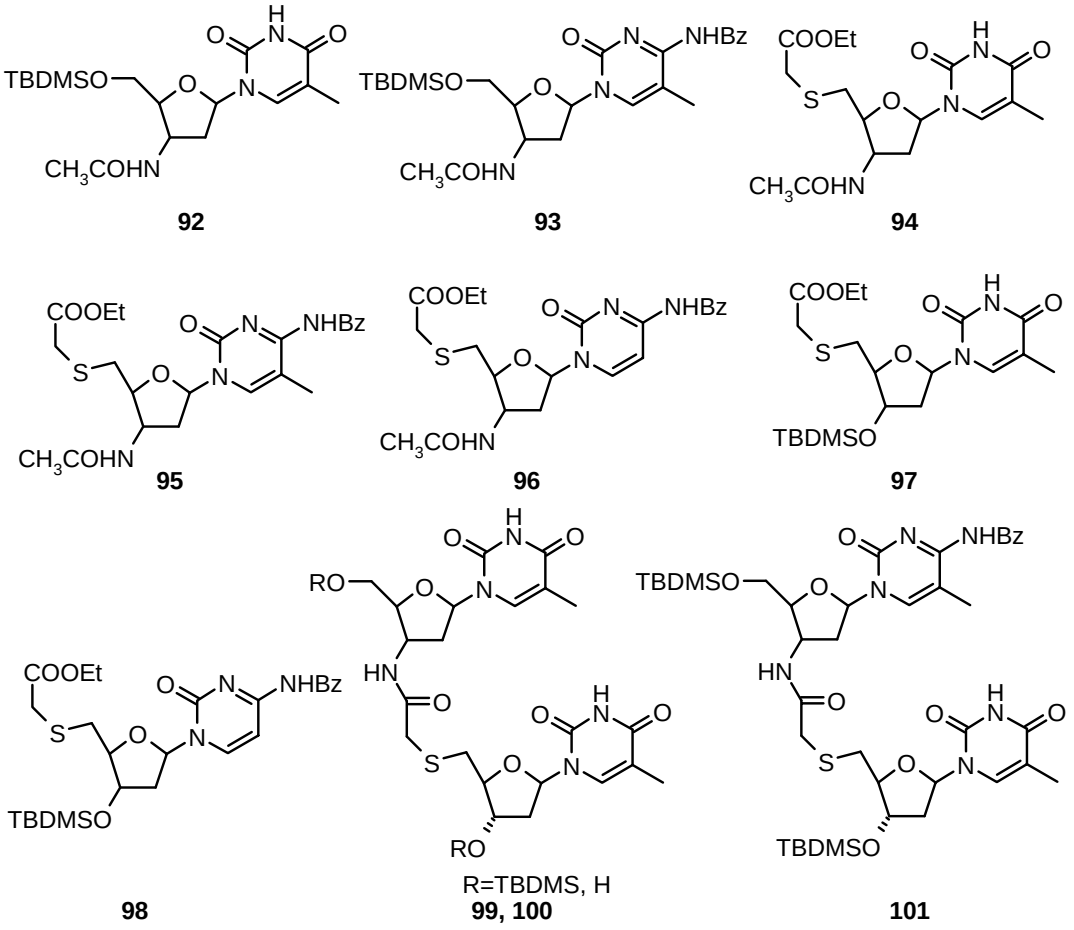


Figure 8. chemical structure of the TANA monomers and dimer used for conformational analysis

We have discussed the RNA binding selectivity of the TANA oligomers in the previous chapter. To understand this extensive RNA binding selectivity of TANA oligomers we have performed NMR conformational analysis of the sugar rings of the TANA dimers and monomers.

The conformation of the pentafuranose ring in a nucleoside moiety can be fully described in terms of the phase angle of pseudorotation (P) and the puckering amplitude (ϕ). From X-ray studies on nucleosides and nucleotides, it was found that ϕ values range from 35º to 45º. For North-type (N) sugars (C3'-endo, C2'-exo), P ranges from -1º to 34º and for S-type(S) sugars (C2'-endo, C3'-exo), P ranges from 137º to 194º. In solution, the sugar ring exists in equilibrium of the two rapidly interconverting conformers N↔S. The mol fraction of N and S conformer as well as their geometry, expressed by their phase angle of pseudorotation P_N and P_S and puckering amplitude ϕ_N and ϕ_S, can be calculated from the vicinal proton-proton ($^3J_{HH}$) coupling constants $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$ and $J_{3'4'}$. These coupling constants were used as an input for the pseudorotation analysis of the sugar using the program PSEUROT.

The assignment of the non- exchangeable proton resonances was achieved by using TOCSY, COSY and NOESY experiments. The coupling constants were determined from 1D spectrum and/or by homonuclear decoupling experiments

Chapter 4: Chimeric (α-amino acid + nucleoside-β- amino acid)$_n$ peptide oligomers for selective RNA recognition

4.1 Rationale of the Design

In this chapter, we describe homothymidine DNA analogues having a conformationally constrained amino acid (proline and N-methyl glycine), positively charged amino acid (L-Lysine) and neutral amino acid (L-methionine) as the backbone. The 2'-deoxyribose sugar phosphate backbone in DNA is replaced by the sugar-amino acid backbone considering that- a) In 3'-deoxy-3'amino ribose sugar the five membered ring pucker is found to be 3'-endo that is better suited for mRNA recognition b) The flexibility of the phosphate group in DNA/RNA may be conserved by using a five-atom amide linker c) Negative charge may be inherently unfavourable and thought should be given to neutral (proline, sarcosine and methionine) or positively charged (Lysine) alternatives.

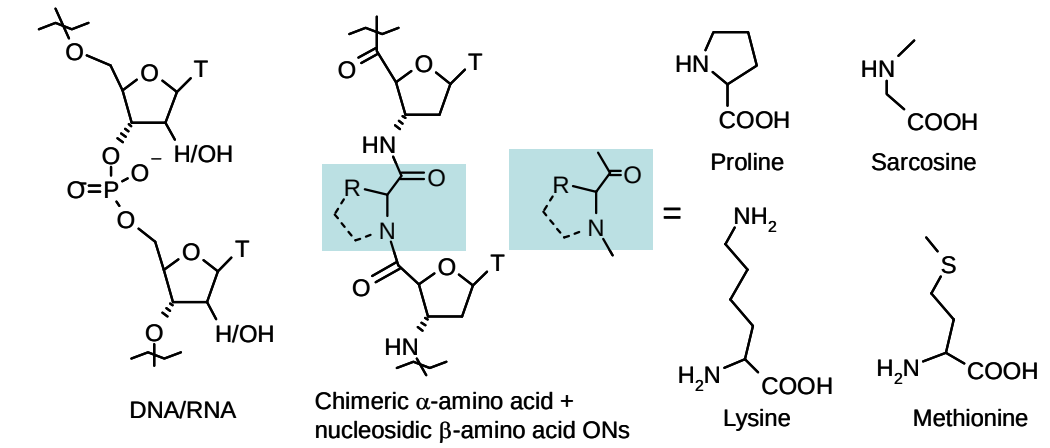


Figure 9. Structure of DNA/RNA, chimeric α- amino acid + nucleoside-β-amino acid ONs and structure of α-amino acids

The synthetic strategy is aimed from assembling these molecules preferably on solid phase, for its subsequent adaptation to the vast and established conventional peptide chemistry. The 5'-hydroxy moiety of 3'-azido thymidine **1** was converted to

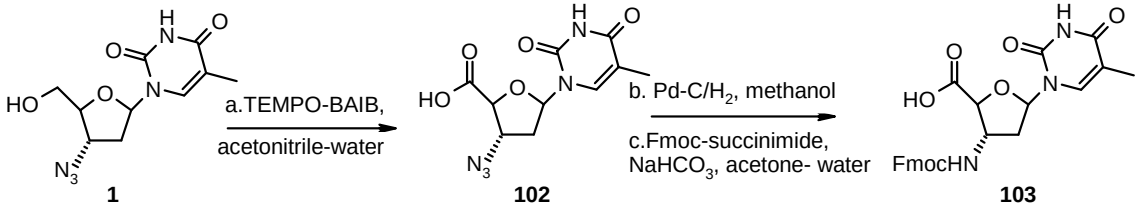


Figure10: Synthesis of nucleoside- β-amino acid

corresponding acid **102** by TEMPO-BAIB oxidation. The 3'-azido functionality was then converted to amine and subsequent Fmoc protection gave the monomer buiding block **103.**

4.2 Synthesis of the oligomers

Four octameric sequences were synthesized using rink amide resin and Fmoc peptide synthesis strategy. The oligomers were cleaved from the solid-support using standard conditions and were purified by gel-filtration followed by RP-HPLC. Finally the purity and integrity of the oligomers were established by MALDI-TOF mass spectrometry (Table **4**).

4.3 UV-melting studies with DNA and RNA

UV melting experiments indicate that the modified oligomers hybridized to DNA and RNA with melting temperatures (*T*m) higher than those of the complexes formed by oligo-T DNA fragment of the same length (Table **4**). Complexes formed by all the modified oligomers with the RNA target had higher stability than similar complexes formed with the DNA target (Table **4**).

Table **4**: The amino acid backbone modified ON sequences and their characteristic molecular mass and melting temperatures of complexes with complementary DNA and RNA

Entry	Sequence (Mass calculated/observed), t_R(min)	DNA **109**	RNA **110**	$\Delta Tm_{RNA-DNA}$
1	5' TTTTTTTT 3' **104**	17.8	15.6	-
2	β-Ala-(Pro-**t**)$_8$-H **105** 2763.7/ 2785.6(+Na^+), 12.7	49.0 (+31.2)	60.8 (+ 45.2)	+ 11.8
3	β-Ala-(Sar-**t**)$_8$-H **106** 2555.4/ 2578.8(+Na^+), 12.3	49.1 (+31.3)	61.6 (+ 46.0)	+ 12.5
4	(Lys-**t**)$_8$-H **107** 2941.1/ 2941.7, 11.7	56.3 (+38.5)	69.0 (+53.4)	+ 12.7
5	β-Ala-(Met-**t**)$_8$-H **108** 3036.3/ 3037.0, 12.6	47.3 (+29.5)	57.3 (+ 41.7)	+ 10.0

T denotes DNA backbone. DNA **109** 5' GCAAAAAAAACG 3', RNA**110**: r (5' GCAAAAAAAACG 3'), Experiments were performed in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM).

Chapter 5: A versatile method for the preparation of conjugates of peptides with DNA/PNA/analogue by employing chemo-selective click reaction in water.

The application of PNA and analogous uncharged DNA mimics as antisense agents is stymied by the fact that PNAs show very low cell-penetration for any observable antisense effect. Several strategies are being developed for the delivery of modified ONs into cells. For uncharged ON mimics such as PNAs, the best option seems to be the covalent conjugation of PNA oligomers with cell penetrating peptides (CPP).The CPPs are mostly positively charged peptides containing lysine or arginine or other peptides having specific cell receptors. The conjugation can be achieved by tedious continuous solid phase synthesis of PNA and peptide. The other methods for conjugation could be post-synthetic via disulfide bridge or more stable thioether linkages at either *C*- or *N*-

terminus. The highly functionalized nature of these biomolecules render them susceptible for side reactions during conjugation and the yield and purity of structurally defined conjugated biomolecules is often low. The current literature clearly indicates the need for a simple and straightforward strategy for generating highly pure ON/PNA-peptide conjugates in high yield.

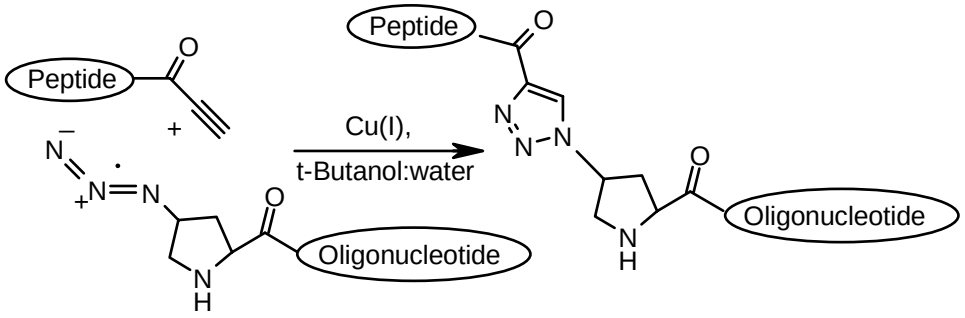


Figure 11. Peptide-alkyne conjugation with 4-azidoprolyl ON using click chemistry

The highly selective orthogonal Cu (I) catalyzed Huisgen 1, 3 dipolar cycloadddition reaction, recognized as 'click chemistry' is highly predictable, very fast and resistant to side reactions, can be carried out in aqueous medium and can be employed post-synthetically on purified units decorated with a variety of functional groups without employing any group protecting strategies. With this background we envisaged a very exciting but simple possibility of synthesizing peptide-DNA/PNA conjugates using click chemistry *i.e.* a specific cycloaddition reaction between terminal azide and alkyne functionalities on PNA/DNA or peptide as per the synthesis design. In this chapter the successful application to generate the cell-penetrating peptide conjugates with DNA/PNA ONs and thioacetamido nucleic acid is described.

The alkyne functionalized peptide and azide functionalized PNA (or TANA) oligomers were synthesized individually on solid phase and cleaved from solid support and the click cycloaddition reaction was performed in solution-phase.The results are shown in Table 5.

Table **5** HPLC and mass spectral data of CPP- PNA/ (TANA), DNA conjugates synthesized.

	Compounds	HPLC	Purity %	Mass calcd	Mass found
1	$(Lys)_6$-alkyne **111**	9.7	91	839.08	837.8
2	$(Lys)_6$triazole-proline **112**	9.59	90	995.23	996.5
3	β-Ala-TTTTTTTT-Pro-N_3 **113**	12.6	99.1	2356.2	2356.6
4	β-Ala-TTTTTTTT-Pro-triazole$(Lys)_6$ **114**	12.29	100	3195.3	3195.35
5	β-Ala-tttttttt-Pro-N_3 **115**	18.5	100	2605.3	2628($+Na^+$)
6	β-Ala-tttttttt-Pro-triazole-$(Lys)_6$ **116**	17.2	100	3444.9	3444.0
7	Lys-TCACTAGATG- Pro-N_3 **117**	11.7	98.5	2991.2	2990.8
8	Lys-TCACTAGATG-Pro-triazole-$(Lys)_6$ **118**	11.08	98	3830.3	3832.5
9	$(Arg\text{-}Aha\text{-}Arg)_4$-alkyne **119**	12.9	100	1772.1	1772.38
10	β-Ala-tttttttt-Pro-triazole-$(Arg\text{-}Aha\text{-}Arg)_4$ **120**	17.2	89	4378.6	4376.29
11	Lys-TCACTAGATG-Pro-triazole-$(Arg\text{-}Aha\text{-}Arg)_4$ **121**	12.7	95.4	4763.3	4763.8
12	5'alkyne**TTGTACTGATAGAGT GTCC** 3' **122**	8.35	95	5878.0	5878.01
13	$(Arg\text{-}Aha\text{-}Arg)_4$- Pro-$N_3$ **123**	-	-	1858.3	1881.2($+Na^+$) 1897.2($+K^+$)
14	$(Arg\text{-}Aha\text{-}Arg)_4$-Pro-triazole-**TTGTACTGATAGAGTGTCC**3' **124**	11.04	81.6	7736.3	7735.09

A, T, G, C *aeg* PNA backbone, t TANA backbone , **A**, **T**, **G**, **C** DNA backbone.

ABREVIATIONS

β- ala	β-alanine
A	Adenine
Ac	Acetyl
Ac_2O	Acetic anhydride
aeg	Aminoethylglycine
ap	Antiparallel
arg	Arginine
aq.	Aqueous
AZT	3'-azido thymidine
bep	Backbone extended pyrrolidine
Bz	Benzoyl
C	Cytosine
Cat	Catalytic/catalyst
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
D-	Dextro
2'-dA	Deoxyadenosine
dC	2'-Deoxycytidine
dU	2'-Deoxyuridine
DCA	Dichloroacetic acid
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DCU	Dicyclohexyl urea
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
h	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	Liquid Chromatography-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
min	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	Nucler Magnetic Resonance
p	Parallel
pfp	Pentafluorophenyl
ppm	Parts per million
Pro	Proline
PS-oligo	Phosphorothioate-oligo
Py	Pyridine
PNA	Peptide Nucleic Acid
Rf	Retention factor
RP	Reversed Phase
SPPS	Solid Phase Peptide Synthesis
t-Boc	Tert-Butoxycarbonyl
T	Thymine
TANA	Thio acetamido Nucleic Acids
TEA	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

Chapter 1

1.1 Nucleic Acids: An Introduction

1.1.1 Primary structures of DNA and RNA

Nucleic acids, DNA and RNA, are complex, high-molecular-weight biochemical macromolecules. These are composed of nucleoside linked through 3'-5'phosphodiester linkages leading to sequences convey genetic information (Figure 1). Nucleic acids are found in all living cells and viruses and these are the most important class of biopolymers dominating the modern bio-molecular science since Watson and Crick discovered double helical structure.[1]

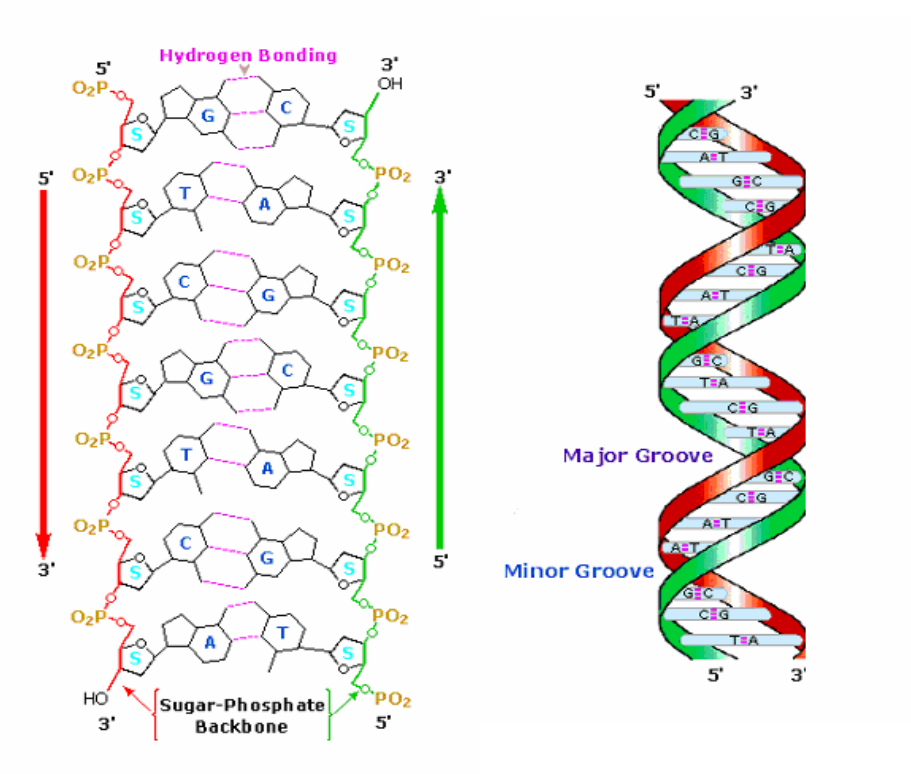


Figure 1. Chemical structure and schematic representation of DNA

Nucleic acids contain all the information required for transmission of genetic information to execute the steps necessary for the synthesis of functional proteins.

Each nucleoside consists of two components: a nitrogenous heterocyclic base, which is either a purine or a pyrimidine and pentose sugar deoxyribose or ribose (Figure 2).The nitrogenous bases adenine, cytosine, and guanine are found in both RNA and DNA, while

thymine only occurs in DNA and uracil only occurs in RNA (Figure 2). A very rare exception is a bacterial virus called PBS1 that contains uracil in its DNA.[2]

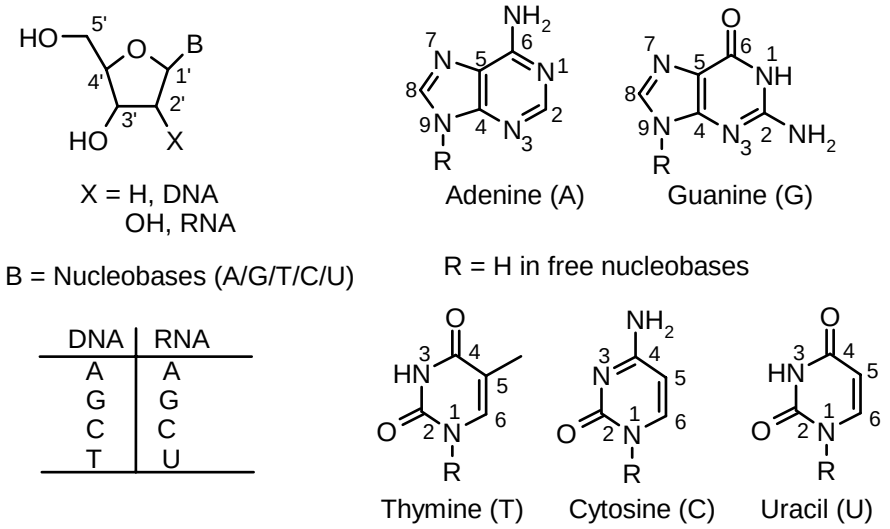


Figure 2. Structures of nucleosides and nucleobases of DNA and RNA

1.1.2 Shapes of Nucleotides

The molecular geometry of an individual nucleotide is very closely related to that of the corresponding nucleotide units in oligomers and nucleic acid helical structure .The details

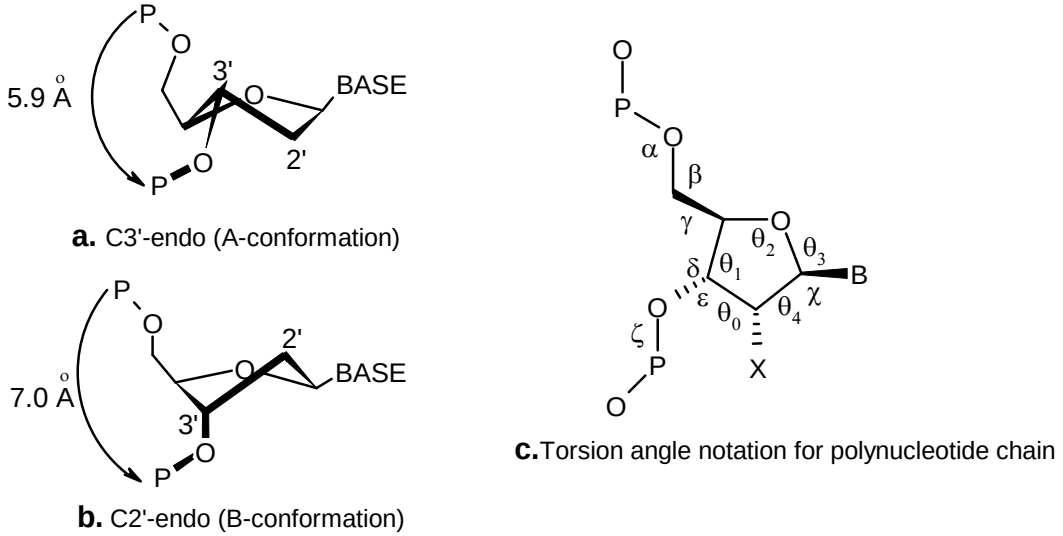


Figure 3. **a)** Structures of C3'- endo and **b)** C2'-endo preferred sugar puckers and **c)** Torsion angle notation for polynucleotide chain

of the conformational structure of nucleotides are accurately defined by the torsional angles α, β, γ, δ, ε and ζ in the phosphate backbone, θ_0 to θ_4, in the furanose ring, and χ for the glycosidic bond (Figure 3). Because many of these torsional angles are interdependent, one can simply describe the shapes of nucleotides in terms of the parameters: the sugar pucker, the *syn-anti* conformation of the glycosidic bond, the orientation of C4'-C5' bond and the phosphodiester backbone.

1.1.3 Sugar Pucker in nucleosides

The pentose sugar rings in nucleosides are twisted or 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 C1'-O4'-C4'. Thus, if the *endo*-displacement of C-2' is greater than the *exo*-displacement of C-3', the conformation is called C2'-*endo* and so on for other atoms of the ring (Figure 3a and 3b). The *endo*-face of the furanose is on the same side as C5' and the base; the *exo*-face is on the opposite face to the base. The sugar puckers are located in the north (N) and south (S) domains of the pseudorotation cycle of the furanose ring.[3] In solution, N and S conformations are in rapid equilibrium and are separated by low energy barrier. The average position of the equilibrium is influenced by (i) the preference of the electronegative substituents at C2' and C3' for axial orientation, (as in the case of RNA the C2'-OH interact with backbone phosphate, C3'-*endo* is preferable conformation), (ii) the orientation of the base (*syn* goes with C2'-*endo*), and (iii) the formation of an intra-strand hydrogen-bond from O2' in one RNA residue to O4' in the next, which favors C3'-*endo*-pucker. The rise in the sugar phosphate backbone for each monomeric unit is 5.9 Å in case of RNA and 7.0 Å for DNA.

1.1.4 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]

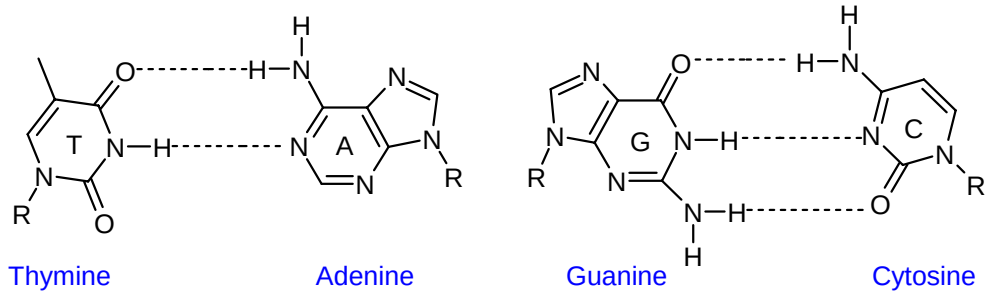


Figure 4. Watson and Crick hydrogen- bonding scheme for A: T and G: C base pair

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

A Hoogsteen A:T base pair (Figure 5) applies the N7 position of the purine base (as a hydrogen bond acceptor) and C6 amino group (as a donor), which bind the Watson-Crick (N3-O4) face of the pyrimidine base. Hoogsteen pairs have quite different properties from Watson-Crick base pairs. The angle between the two glycosylic 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.

Hoogsteen base-pairing allows sequence specific binding of 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.

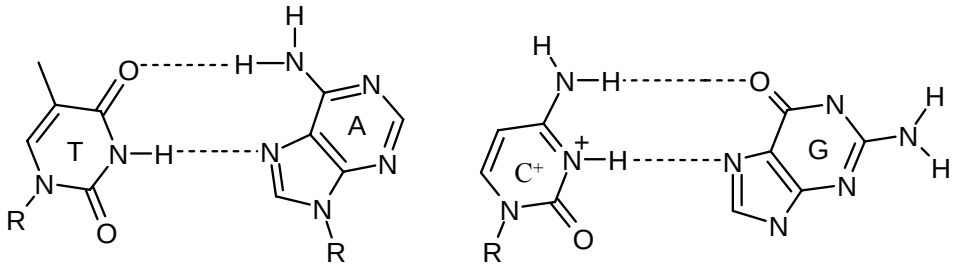


Figure 5. Hoogsteen hydrogen- bonding scheme for A: T and G: C base

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]

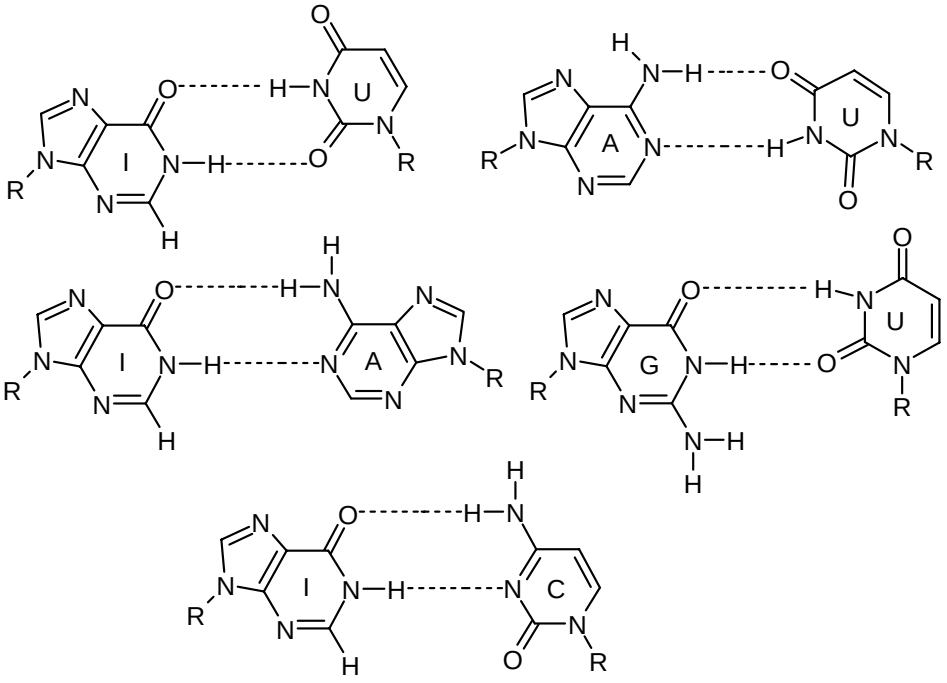


Figure 6. Wobble base pair for Ionosine and Uracil

1.1.5 DNA Secondary structures

Three DNA conformations are believed to be found in nature, A-DNA, B-DNA, and Z-DNA (Figure 7). The "B" form described by James D. Watson and Francis Crick is believed to predominate in cells.[6] B-DNA is a right-handed double helix with a wide major-groove and a narrow minor-groove where the bases are perpendicular to the helical axis. It is 23.7 Å wide and extends 34 Å per 10 bp of sequence. A-DNA and Z-DNA differ significantly in their geometry and dimensions from B-DNA, although still form helical structures. At low humidity and high salt, the favored form is highly crystalline A-DNA while at high humidity and low salt, the dominant structure is B-DNA. In both A and B forms of DNA the Watson-Crick base pairing is maintained by *anti* glycosidic conformation of the nucleobases. The sugar conformation however, is different in both

forms with the B form showing C2'-*endo* puckered sugar and the A form DNA exhibiting C3'-*endo* sugar- pucker. A very unusual form of DNA-duplex is the left-handed Z-DNA.[7]

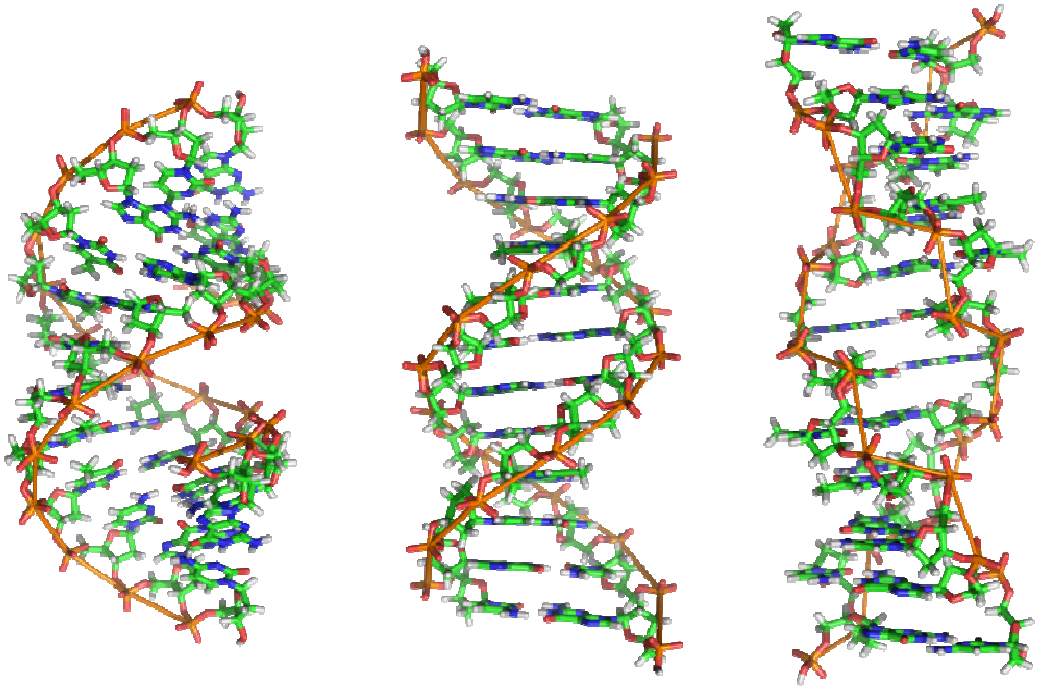


Figure 7. The structures of A, B and Z DNA

This conformation of DNA is stabilized by high concentrations of $MgCl_2$, NaCl and ethanol, and is favored for alternating G:C/G:C sequences. Z-DNA has characteristic zig-zag phosphate backbone and the Watson-Crick base pairing is achieved by purines adopting *syn* glycosidic conformation with C3'-*endo* sugar-pucker.[6] Z-DNA forms excellent crystals. The characterisic features of major DNA conformations are summarized in Table 1**.**

RNA can form double stranded duplexes. The most important structural feature of RNA, that distinguishes it from DNA is the presence of a hydroxyl group at the 2'-position of the ribose sugar. The presence of this functional group enforces the C3'-endo sugar conformation (as opposed to the C2'-endo conformation of the deoxyribose sugar in DNA) that causes the helix to adopt the A-form geometry rather than the B-form most commonly observed in DNA. This results in a very deep and narrow major groove and a shallow and wide minor groove.

Table1. Salient features of the three major forms of DNA and RNA strands

Geometry attribute	A-DNA	B-DNA	Z-DNA
Helix sense	right-handed	right-handed	left-handed
Repeating unit	1 bp	1 bp	2 bp
Rotation/bp	33.6°	35.9°	60°/2bp
Mean bp/turn	11	10.0	12
Inclination of bp to axis	+19°	-1.2°	-9°
Rise/bp along axis	2.3 Å	3.32 Å	3.8 Å
Pitch/turn of helix	24.6 Å	33.2 Å	45.6 Å
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	anti	anti	C: anti, G: syn
Sugar pucker	C3'-endo	C2'-endo	C: C2'-endo, G: C3'-endo
Diameter	25.5 Å	23.7 Å	18.4 Å
Conditions	Low humidity & High salt	Dilute aqueous solutions	High salt & alternating GC sequences
Major groove	Narrow & deep	Wide and deep	Flat
Minor groove	Wide	Narrow and deep	Narrow and deep

Some other DNA conformations are also possible;[8] C-DNA, D-DNA, E-DNA, L-DNA, P-DNA , and S-DNA have been described so far. As mentioned above C-DNA, D-DNA, E-DNA, and P-DNA have not been observed in naturally occurring biological systems.

1.1.6 Structures of RNA

RNA is recognized to have greater structural versatility than DNA.[9] Its chemical reactivity (instability) arises from the extra complexity of the presence of 2'-OH group in ribonucleosides, capable of intramolecular nucleophilic attack on adjacent phosphate. RNA is single stranded but can form complex and unusual structures such as stem and bubble structures, due to the chain folding as a consequence of intramolecular base pairing. An example of folded RNA structure is *t*-RNA, which is the key RNA involved in the translation of genetic information from *m*-RNA to proteins. *t*-RNA contains about 70 bases that are folded such that there are base paired stems, bulges and open loops. The

overall shape of the completely folded *t*-RNA is L-shaped. The presence of the 2'-hydroxy group in RNA hinders the formation of a B-type helix but can be accommodated within an A-type helix. At low ionic strength, A-RNA has 11 base pairs per turn in a right-handed, antiparallel double helix. The sugars adopt a C3'-*endo* pucker and other geometric parameters are all very similar to A-DNA. If the salt concentration is raised above 20%, A'-RNA form is observed which has 12 base pairs per turn of the duplex. Both structures have typical Watson-Crick base pairs, which are displaced by 4.4 Å from the helix axis, hence forming a very deep major-groove and rather shallow minor-groove.

1.1.7 Sense DNA and antisense RNA

A DNA sequence is called "sense" if its sequence is the same as that of a messenger RNA copy that is translated into protein. The sequence on the opposite strand is complementary to the sense sequence and is therefore called the "antisense" sequence.

Antisense RNA is defined as a short RNA transcript that lacks coding capacity, but has a high degree of complementarity to structure of coding RNA which enables the two to hybridize.[10] The consequence is that such antisense, or complementary RNA can act as a repressor of the normal function or expression of the targeted RNA. Such species have been detected in many prokaryotic cells with suggested functions concerning RNA primed replication of plasmid DNA, transcription of bacterial genes, and messenger translation in bacteria and bacteriophages. Quite clearly, such a regulation of gene expression depends on the integrity of RNA duplexes.[10]

1.1.8 Molecular Recognition in the Major and Minor Grooves of Duplex DNA

The major and the minor grooves of DNA differ significantly in their electrostatic potential, hydrogen bonding character,[5] steric effects, hydration[3a] and dielectric strength.[11] A:T and T:A base pairs can accept additional hydrogen bonds from ligands bound in the major groove via the C4 carbonyl of T and N7 of A, while in the minor groove hydrogen bonding occurs through the C2 carbonyl of T and N3 of A (Figure 8). The only hydrogen bond donor in the major groove for the A:T base pair is the N6 amino group of A, while none exists in the minor groove. For C:G and G:C duplexes, the H-bond acceptors in the major groove are N7 and O6 for G and in the minor groove are O2

of C and N3 of G. The hydrogen bond donor in the major groove for C:G is the N4 amino of C and in the minor groove, the N2 amino of G.

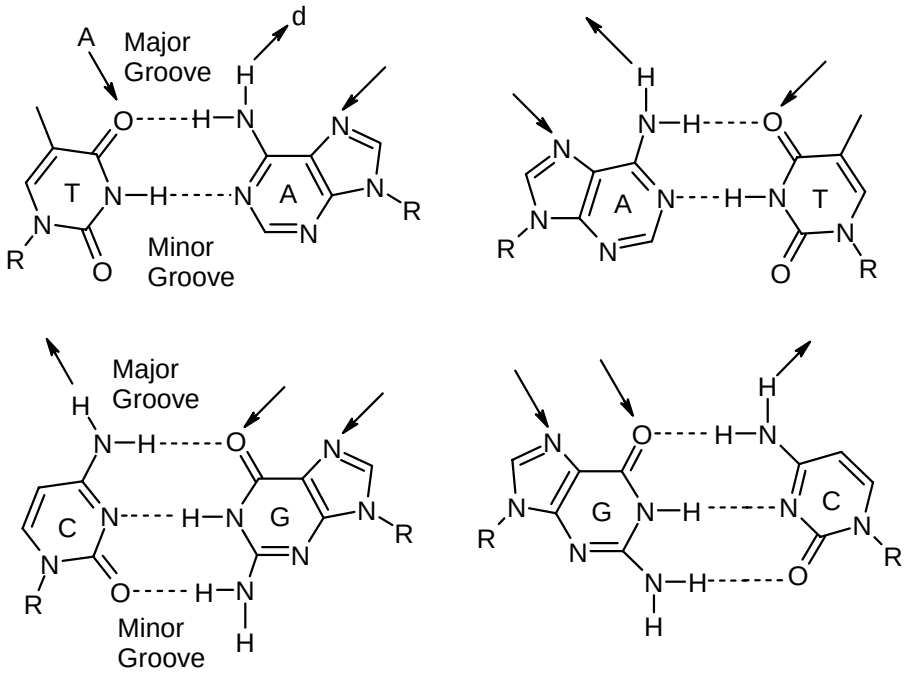


Figure 8. Hydrogen bond donor and acceptor sites in the major groove of duplex DNA at T:A, A:T, C:G and G:C base pairs. Arrows pointing to the atoms indicate acceptor sites; arrows pointing away from the atoms indicate donor sites.

1.1.9 Triplex forming Oligonucleotides (TFO)

Triple-stranded DNA was first described in 1957.[12] The nucleobase thymine binds to adenine in the Watson-Crick base-pair A:T through Hoogsteen hydrogen bonds (Figure 9). An N-3 protonated cytosine, represented as C^+, can also form a base-triplet with a guanine of C-G pair through the Hoogsteen H-bonding (Figure 9). Thus, the triple-helical DNAs using these Hoogsteen pairings consist of two homopyrimidines and one homopurine, and the homopyrimidine third strand is parallel to the homopurine strand. A homopurine third strand can also bind to a homopurine-homopyrimidine duplex using reversed Hoogsteen patterns. In this triplex, a nucleobase A binds to a T-A base pair and a G to a C-G pair. Since the nucleobases on the third strand have to be reversed, the homopurine third strand is antiparallel to the homopurine strand of the original duplex.

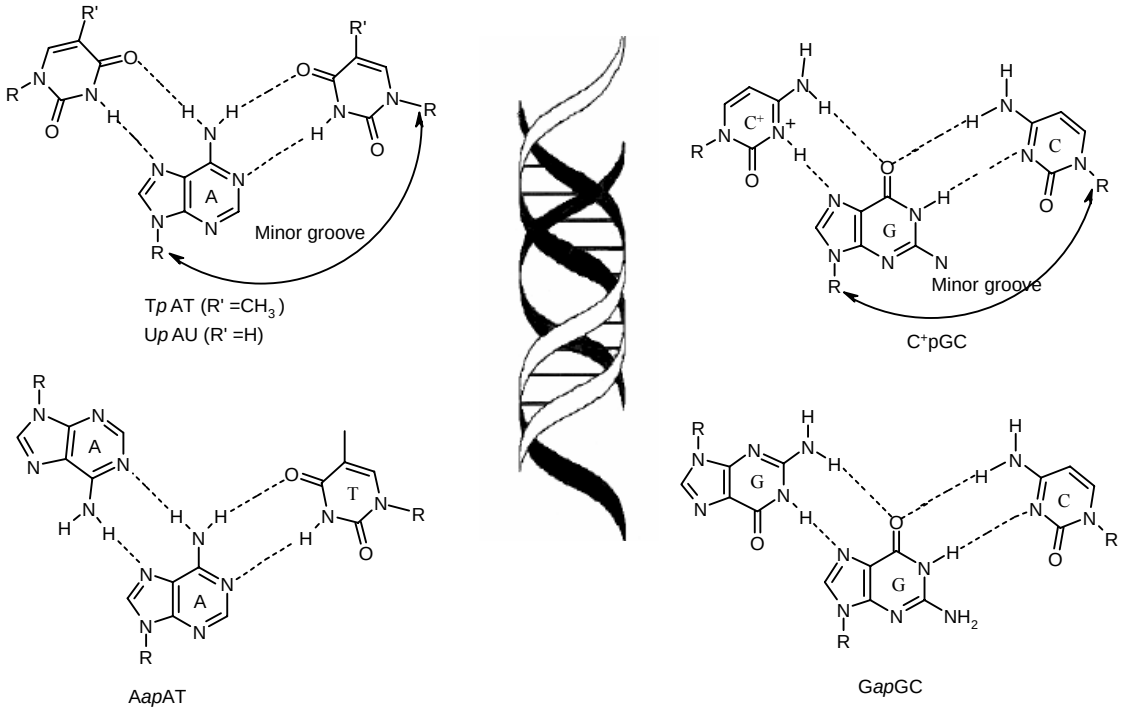


Figure 9. A ribbon model demonstrating the relative position of a TFO in the major groove of DNA, The pyrimidine binding motif- The binding of a TFO in a parallel (*p*) orientation and the purine motif-the binding of TFO in antiparallel (*ap*) orientation.

The widespread occurrence of polypurine/polypyrimidine tracts in eukaryotic DNA suggest that these sequences may have a biological function and consequently, therapeutic significance. The analysis of eukaryotic sequence databases reveals thousands of polypurine/polypyrimidine tracts, with potential for triple helix formation. Triplex formation obeys precise rules imposed by structural constraints. For a given sequence of DNA it is possible to design therapeutic oligonucleotides (ODNs) that will specifically bind to it and thereby inhibit the gene expression, different aspects of triple-stranded structures have been discussed in several reviews.[13] The formation and stabilization of triplexes depends on different types of interactions like electrostatic forces, stacking and hydrophobicity contributions, Hoogsteen hydrogen bonds, and hydration forces.[13]

1.2 Antisense Technologies

1.2.1 Introduction

Drugs consisting of oligonucleotide analogues capable of combining with RNA through a Watson-Crick base-pairing mechanism, thereby arresting cellular processes at the transcription level are known as antisense agents.[10]

The potential of oligodeoxynucleotides to act as antisense agents that inhibit viral replication in cell culture was discovered by Zamecnik and Stephenson in 1978.[14] 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,[15] 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. For example, the widely used approach to generate knock-out animals to gain information about the function of genes in vivo is time-consuming, expensive, labor intensive and, in many cases, non informative due to lethality during embryogenesis.[16] In these cases, antisense technologies offer an attractive alternative to specifically knock down the expression of a target gene. Another advantage of the development of AS-ONs is the opportunity to use molecules for therapeutic purposes, which have been proven to be successful in animal models.

Three types of anti-mRNA strategies can be distinguished.[17] Firstly, the use of single stranded antisense oligonucleotides; secondly, the triggering of RNA cleavage through catalytically active oligonucleotides referred to as ribozymes; and thirdly, RNA interference induced by small interfering RNA molecules (Figure 10).

Figure10 demonstrates the difference between antisense approaches and conventional drugs, most of which bind to proteins and thereby modulate their function. In contrast, antisense agents act at the mRNA level, preventing its translation into protein. Antisense-oligonucleotides (AS-ONs) pair with their complementary mRNA, whereas ribozymes and DNA enzymes are catalytically active ONs that not only bind, but can also cleave, their target RNA. In recent years, considerable progress has been made through the

development of novel chemical modifications to stabilize ONs against nucleolytic degradation and enhance their target affinity. In addition, RNA interference has been established as a third, highly efficient method of suppressing gene expression in mammalian cells by the use of 21–23-mer small interfering RNA (siRNA) molecules.

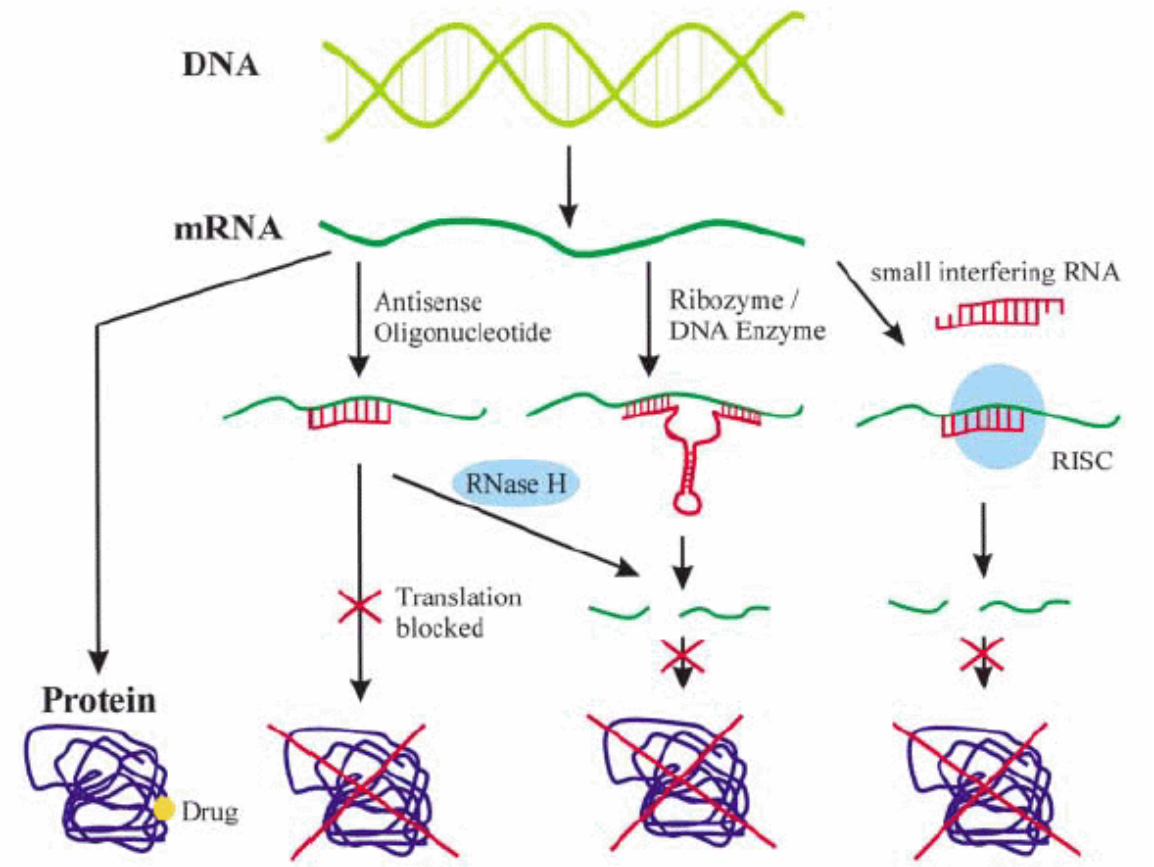


Figure 10. Different antisense strategies

1.2.2 Antisense Oligonucleotides: Disruptive antisense approach

The AS-ONs which can disrupt the translation processes at mRNA level (Figure11), thus causing the disruption or down-regulation of the synthesis of disease-causing functional protein. AS-ONs usually consist of 15-20 nucleotides, which are complementary to their target mRNA. There are two major mechanisms for their antisense activity. The first is that most AS-ONs are designed to activate RNase H, which cleaves the RNA moiety of a DNA/RNA heteroduplex and therefore leads to degradation of the target mRNA. In addition, AS-ONs that do not induce RNase H cleavage can be used to inhibit translation

by steric blockage of the ribosome. When the AS-ONs are targeted to the 5' terminus, binding and assembly of the translation machinery can be prevented.

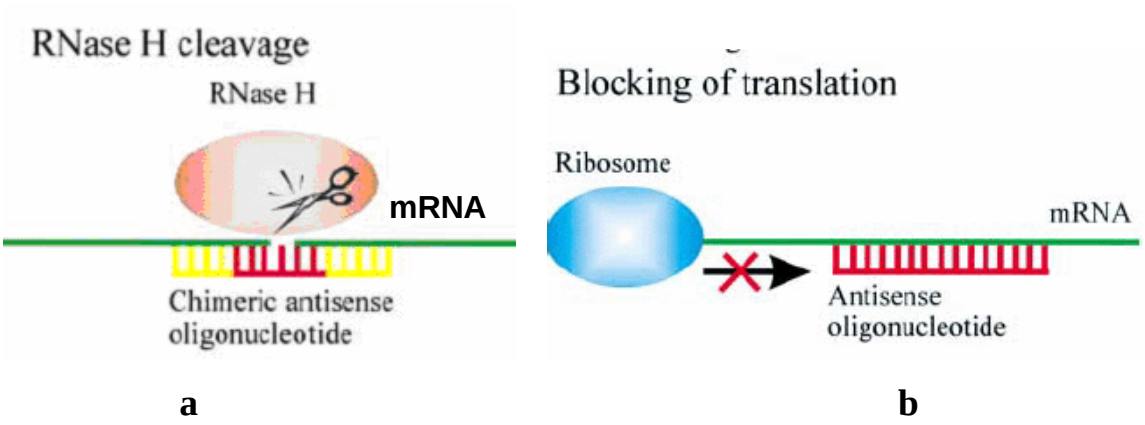


Figure 11. a) RNase H cleavage induced by (chimeric) antisense-oligonucleotides
b) Translational arrest by blocking the ribosome.

1.2.3 Splice Correction: Corrective antisense approach

Recent research opens up an exciting possibility for the application of AS-ONs as corrective antisense[18] 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[18] or to yield mRNA that is translated into viable proteins. The disruptive antisense effects could involve non-antisense, non-specific interactions leading to stimulation of immune response but corrective antisense depends upon specific and stable interactions with mRNA for the desired corrective action.

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. Once pre-mRNA has been completely processed, it is termed "mature messenger RNA", "mature mRNA", or simply "mRNA". 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.[19] 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 colinear to the mRNA sequence that will be spliced out after transcription, but before the mRNA is translated.

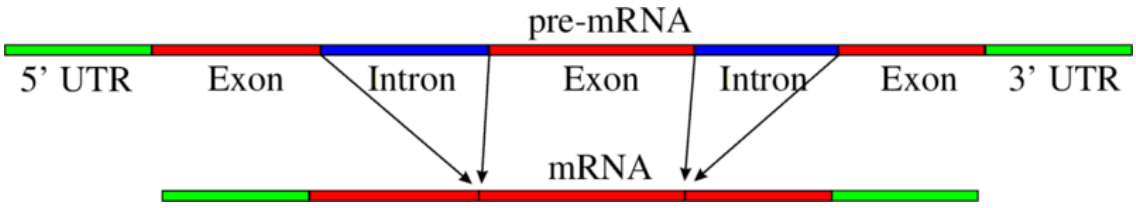


Figure 12. Simple illustration of exons and introns in pre-mRNA.The mature mRNA is formed by splicing.(UTR: Untranslated region)

Splicing involves a two-step biochemical process (Figure 13). Both steps involve transesterification reactions that occur between RNA nucleotides. First, the 2'-OH of a specific *branch-point* nucleotide within the intron that is defined during spliceosome assembly performs a nucleophilic attack on the first nucleotide of the intron at the 5' splice site forming the *lariat intermediate*. Second, the 3'-OH of the released 5' exon then performs a nucleophilic attack at the last nucleotide of the intron at the 3' splice site thus joining the exons and releasing the intron lariat.

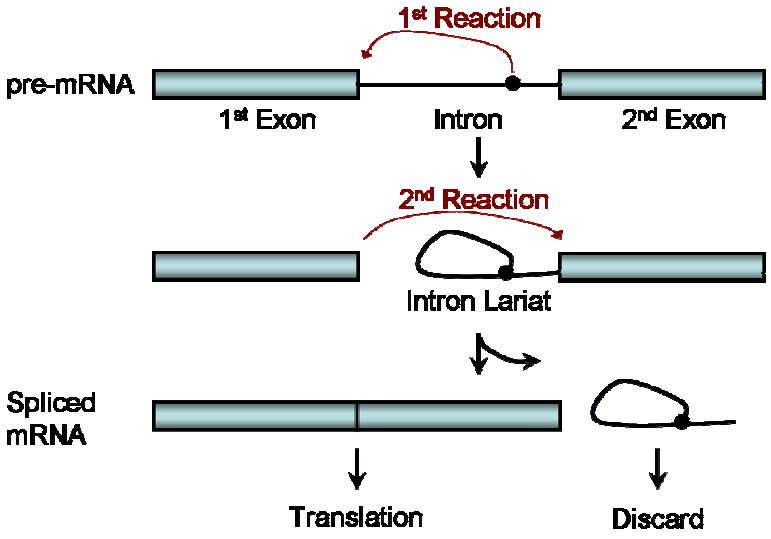


Figure 13. Biochemical mechanism of splicing

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.[20] 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.

1.3 Antisense oligonucleotide modifications: Structure-Editing of Nucleic Acids for Selective Targeting of RNA

One of the major challenges for antisense approaches is the stabilization of ONs, as unmodified oligodeoxynucleotides are rapidly degraded in biological fluids by nucleases. A vast number of chemically modified nucleotides have been used in antisense experiments most of which have been described in the review article by Kumar and Ganesh.[21] In general, three types of modifications of ribonucleotides can be distinguished (Figure 14): analogs with unnatural bases, modified sugars (especially at the 2' position of the ribose) or altered phosphate backbones.

For either corrective or disruptive antisense activity of AS-ONs requires strong and specific interactions with target mRNA sequences. The platform where this interaction occurs may be either cytoplasm or nucleus and AS-ONs may have to compete with isosequential DNA for complex formation. The preferential RNA binding may prove to be more significant in corrective AS drug development as the site of their action is the nucleus rather than the cytoplasm and are basically driven by steric blocking. The recently discovered siRNA[115, 116] therapeutics employs dsRNA as therapeutic entity, and thus the RNA selective ONs may prove to be very effective in their application perspective.

The differences in RNA and DNA backbone structures mainly arise from the presence of 2'-OH function in RNA.[3a] Comprehensive analyses of a two-state North (N, roughly C3'-*endo*) $\leftrightarrow$ South (S, roughly C2'-*endo*) (Figure 3) pseudorotational equilibrium in nucleosides and nucleotides have shown that the sugar conformation is controlled by gauche effects of [O4'-C4'-C3'-O3'/ O4'-C1'-C2'-O2'] fragments, anomeric effect of heterocyclic nucleobase and other steric interactions. The 2'-OH function in RNA gives RNA a distinct propensity to prefer 3'-*endo* sugar conformation. The absence of 2'- OH

function in DNA allows an extended backbone with C2'-*endo* sugar conformation (Figure 3). The changes in sugar pucker mainly affect the backbone torsion angle δ, which on average is ~80º in A form duplex and ~120º in B-form duplex.[22] The DNA:DNA duplex is an extended B-form structure whereas RNA:RNA duplex remains in compact A-form, further stabilized by increased base stacking of axially oriented nucleobases. The relatively low conformational energy barriers of the deoxyribose sugar render flexibility to 2'-deoxyribose sugar to adopt 3'-*endo* (or O4'-*endo*) conformations in the DNA strands and allow DNA: RNA duplexes also to be in A-form. The backbone modified antisense oligonucleotides with engineered desired conformers of sugar moiety bind to complementary DNA or RNA sequences depending upon the flexibility of the

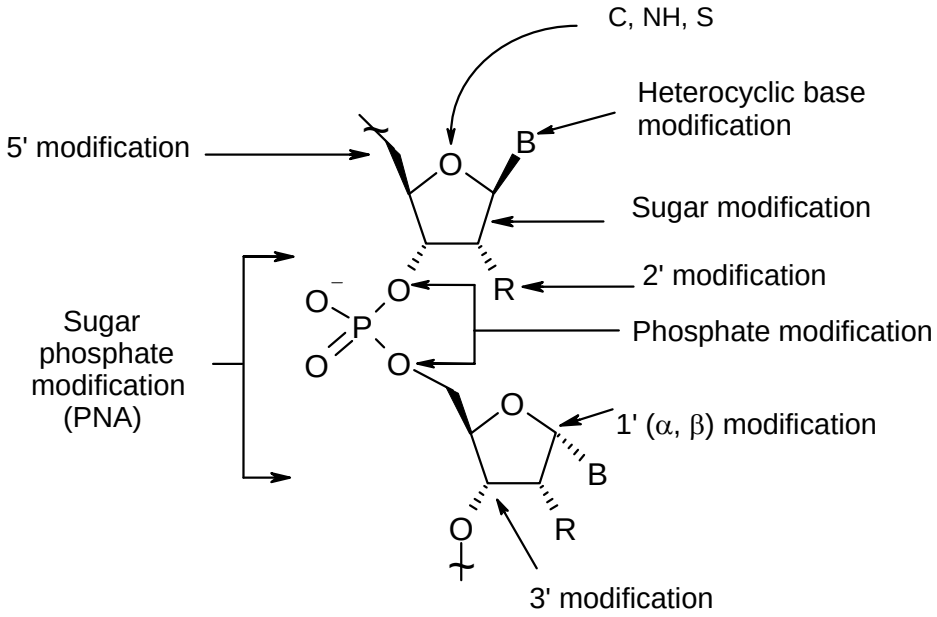


Figure 14. Possible chemical modifications in oligonucleotides.

designed backbone to adopt A, B or intermediate form of the duplex structures. The stability is further enhanced by base-stacking interactions and presence or absence of positive/ negative charges on the modified backbones.

1.3.2 Enantio-DNA and other stereoisomers of DNA/RNA

A very early example towards RNA selectivity was found for short enantio-DNA sequences based on L-sugars (Figure 15). It was found as early as in 1970 by Ts'O that L-ApA interacts with poly(U).[23] Later in 1984, L- $(dUp)_{17}$ dU was shown to have no

interaction with poly(dA).[24] In 1990, Shudo *et al.*[25] synthesized enantiomeric-dA_6 and showed that they bind sequence specifically with poly(U) but not with poly(dT). The thermal stability of selective enantio-DNA:RNA complexes was less than either DNA:RNA or RNA:RNA complexes. Meyer, Jr. *et al.*[26] studied a 15 mer sequence

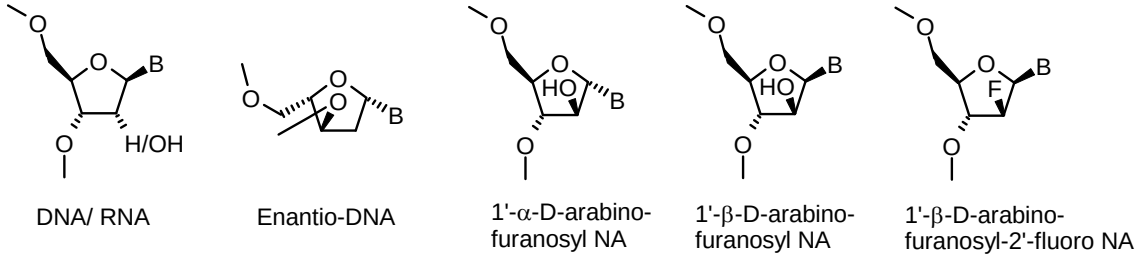


Figure 15. Enantio-DNA and other stetroisomers of DNA and RNA

containing 1'-α-D-arabinofuranosylthymine (Figure 15) that showed binding with complementary RNA as good as the natural DNA:RNA duplex but its complex with poly-dA showed substantially lower stability than the natural counterpart. The presence of arabino-D-sugar, the 2'-stereisomer of ribose in the oligomers (ANA) allowed complex formation with both RNA and DNA, but with reduced stability.[27] This destabilization was presumed to be derived from the steric interference by the β-2'-OH group, which being oriented into the major groove of the helix, could cause slight local deformation or unstacking. Replacement of 2'-OH of ANA by 2'-Fluoro-ANA (Figure 15) led to the stabilization of 2'F-ANA: RNA duplexes. This stability was attributed to the reduced steric interactions of fluorine atom (versus OH groups) and higher pre-organization state of 2'FANA relative to DNA. The CD spectroscopy and NMR analysis of the ANA:RNA and 2'-F-ANA:RNA duplexes indicated resemblance to the corresponding RNA:DNA -like A type duplex, and could be a substrate for RNase H enzyme action. *Ab initio* model calculations suggest that the sugar puckers in 2'F-ANA and ANA strands, in 2'F-ANA/ANA:RNA duplexes adopt a more S-type or C2'-*endo* geometry. Closer examination of pucker values and structural details, however, indicated substitution-dependent differences in the two complexes. The standard arabinose sugar pucker was shown to be almost C2'-*endo* locked due to an internal H-bonding with C5'-oxygen. The lack of such interaction in 2'FANA allowed increased variability and sugar pucker shift towards C4'-*endo,* but the overall structural geometry of 2'F-ANA/

ANA:RNA duplexes was predicted to be similar to the native DNA:RNA substrate of RNase H.[28] Contrary to these studies, the NMR solution structural analysis showed that in either ANA:RNA or 2'F-ANA:RNA duplexes, the ANA sugar residues assumed conformations in C4'-*endo* range.[29] The differential stability of 2'F-ANA:RNA versus ANA:RNA hybrids was attributed to different extents of hydration of the 2'hydroxy versus 2'-fluorine groups.

1.3.3 LNA and LNA analogues

The anomeric-inverted analogue of DNA (α-DNA) (Figure16) has been demonstrated to hybridize efficiently with complementary DNA/RNA, but in parallel orientation.[30] Due to the predominant influence of the 3'-OH gauche effect [O3'-C3'-C4'-O4'] over the anomeric effect, the β-D-2'-deoxy nucleosides sugar moiety preferentially adopted S-type conformations. The sugars in α-D-2'-deoxy nucleosides were predominantly in the S-type conformation due to the co-operative influence of nucleobase as well as the gauche effect.[31]

Conformationally restricted DNA/RNA analogues such as locked nucleic acids (LNA) (Figure 16) stabilized antiparallel complexes with either DNA or RNA by locking the N type sugar conformation in the AS strand. The target DNA/RNA complementary strand could be in S type, N type or intermediate sugar conformation giving rise to A-form or other intermediate structural features to the stable LNA: DNA/RNA antiparallel duplexes.[32] The α-configured LNA was introduced as α-DNA with locked N type sugar conformation (Figure 16). Fully modified α-D-LNA sequences of either oligothymine or mixed oligopyrimidines, displayed unprecedented recognition of parallel complementary RNA sequences whereas the complementary DNA sequences were not recognized.[33]

Chimeric α-D-LNA/α -DNA sequences showed low affinity for RNA. This observation meant that the stabilization of parallel orientation in RNA:α-D-LNA was not fine tuned in chimeric sequences probably because α-DNA units in the sequence remain in S type preferred conformation creating incoherency between N and S type conformations within the same sequence. This suggested that unlike the original LNA, the N-type conformation of α-D-LNA could not influence the conformations of neighbouring α-DNA nucleosides to adopt N-type conformations.

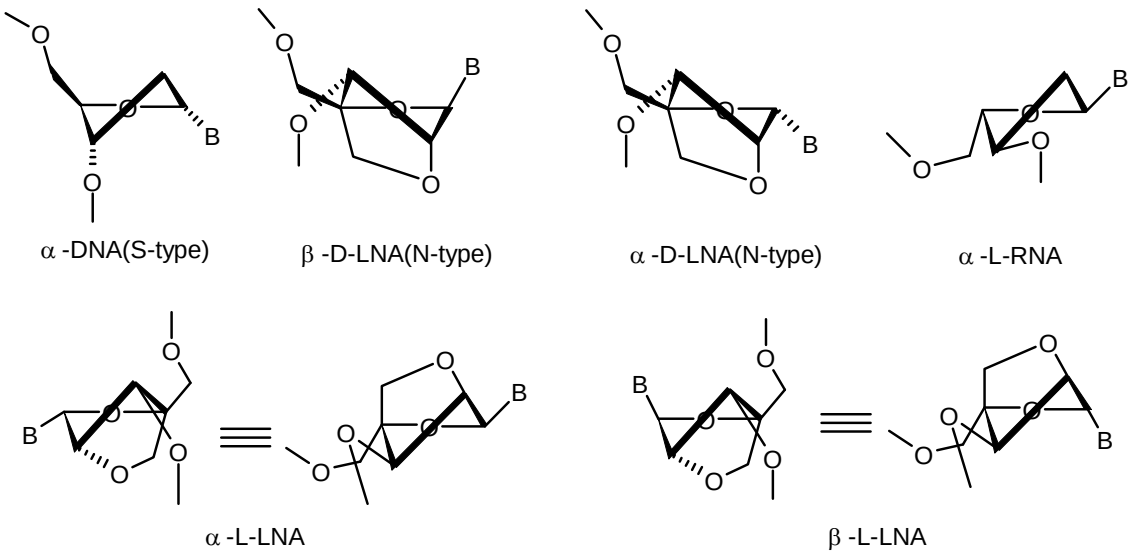


Figure 16. Enantio DNA and other stereoisomers of DNA and RNA

The α-LNA having L-sugar configuration (α-L-LNA) on the other hand was shown to be an excellent structural mimic of DNA (Figure 16).[34] The C3'-*endo* furanose conformation in this locked sugar was equivalent to an N-type conformation because of L-sugar configuration. The α-L-LNA mixmer formed high affinity antiparallel duplexes with RNA and DNA and the thermal stability of the duplexes was only second to that of LNA. The NMR structural studies of a mixmer DNA-α -L-LNA: RNA duplex pointed out that unlike in the case of LNA, the sugar conformations of the deoxyribose sugars remain unperturbed. Molecular dynamics simulations suggested that the α-L-LNA nucleotides slightly organized the 3'-flanking sugar rings in a more S-like conformation and thus favored a more B-type hybrid conformation. Thus if LNA acted as an A-type mimic, α-L-LNA acted as a preferred B-type mimic.[34] On the other hand, the chimeric α-L-RNA /α-L-LNA thyminyl sequences were found to show hybridization only with complementary RNA sequences, and no hybridization with DNA complement was detected.[35] It was quite notable that α-L-LNA displayed increase in RNA binding affinity in chimeric sequences either with α-L-RNA or DNA. With LNA being RNA mimic and α-L-LNA being a DNA mimic, β-L-LNA was deduced to be an α-DNA mimic (Figure 16). Changing the configuration at anomeric center of α-L-LNA led to the synthesis of β-L-LNA, which was expected to mimic the characteristics of α-DNA. Thus, the β-L-LNA

sequences recognized complementary RNA and DNA sequences in parallel orientation with similar stabilities of parallel α-LNA:RNA duplex.[36] Considering the various stereoisomeric possibilities and their conformationally restricted analogues, it could be rationalized that the hybridization properties in pentofuranosyl nucleic acids and preferential recognition of RNA targets is directed by conformational equilibria of the pentofuranoses.

1.3.4 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 17a).[37] 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-stereo specific 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, [38] 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 [39] as well as cleavage of non-target *m*-RNAs that are only partially complementary by RNaseH.[40] 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.

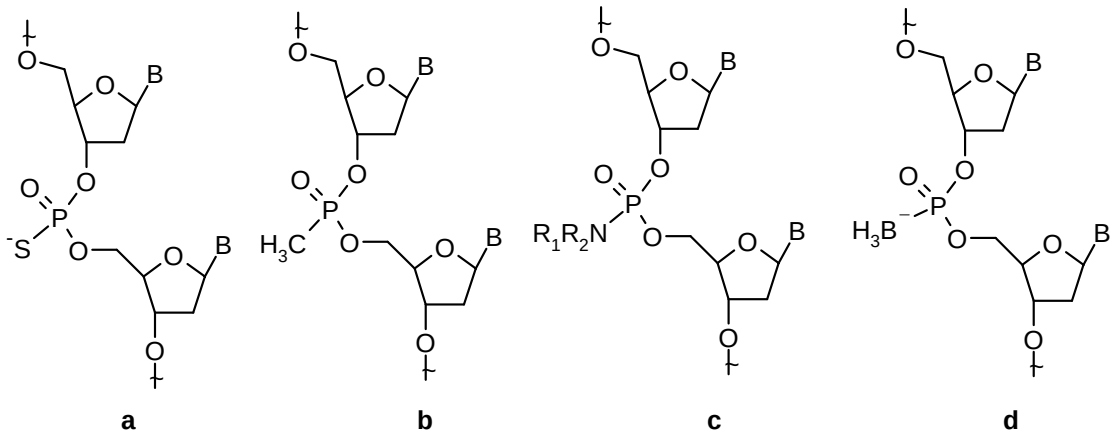


Figure 17. First generation modified antisense oligonucleotide (replacement of 'O' in P-O)

Methylphosphonates: Oligonucleoside methylphosphonates are a class of nonionic oligonucleotide analogs (Figure 17b).[41] These oligomers contain deoxyribonucleoside or 2'-O-methylribonucleoside units that are linked in a 3'-5' manner via methylphosphonate groups. The methylphosphonate internucleotide linkage is similar in structure to the natural phosphodiester (PO) internucleotide bond.[42] One of the non esterified oxygens of PO group is replaced with a methyl group, and this substitution results in the loss of the negative charge. Inclusion of methylphosphonate linkage in the oligomer introduces a new center of asymmetry into the oligonucleotide as Rp or as Sp configuration and consequently the oligomer preparation can contain up to 2^n diastereomers, where *n* is the number of methylphosphonate linkages. Methylphosphonates are resistant to the action of cellular nucleases. The electroneutral methylphosphonate linkage would reduce charge repulsion in methylphosphonate oligos and complementary target nucleic acid more readily penetrate cell membranes, thus enhancing their internalization by cells (also known by the name MATAGENES, masking tape for gene expression).[43] Although oligodeoxyribonucleoside methylphosphonates form stable duplexes with single stranded DNA, hybrids formed with complementary RNA of the same nucleotide sequence generally have lower stability.[44] The reduction in melting value has been ascribed to the presence of mixture of diastereomers. In contrast to normal DNA:DNA/ RNA duplexes, the stability of methylphosphonate-DNA duplexes are not significantly affected by changes in ionic strength. In addition to the poor aqueous solubility, methylphosphonates

cannot induce RNase H activity. For these reasons this modification is less useful as antisense molecule but can serve as the platform for the development of effective antisense oligonucleotides.

Oligonucleotide Phosphoramidates (Figure 17c): Phosphoramidates are another well-studied class of backbone modifications.[45] They are resistant to hydrolysis by nucleases. With DNA targets, oligonucleotide phosphoramidates (Figure 17c, R_1 = H, R_2 = -CH_3, -CH_2CH_2OMe, $(CH_2CH_2)_2O$) exhibit rather poor hybridization characteristics.[46] 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.[47] The *T*m of these complexes turned out to be independent of the ionic strength of the buffer, in contrast to the behavior of unmodified oligonucleotides with DNA or RNA targets, which shows a strong *T*m increase, with increasing salt concentration.

Oligo-boranophosphonates (Figure 17d): These are derived by replacing one of the non-bridging oxygen atoms in the phosphodiester group of DNA with borane (BH_3).[48] 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.[49] Boranophosphate DNA is considerably more stable to various nuclease enzymes than native DNA and overall more stable than phosphorothioate DNA.[48c] The discovery that oligonucleotide boranophosphodiesters can activate *E.coli* RNase H and induce cleavage of RNA is encouraging.[50]

1.3.5 2'-5' RNA/DNA (*iso*RNA/*iso*DNA)

The 2'-5' isomers of RNA are known to occur transiently during RNA splicing reactions [51] and in interferon treated cells.[52] These *iso*RNAs having exclusive 2'-5' phosphodiester bonds showed properties of self-association through Watson-Crick base pairing and interestingly, the 2'-5' linked 3'-deoxy-NA/RNA sequences (Figure 18) could also bind specifically to complementary RNA sequences but show absence or very low binding with complementary DNA.[53] The stability of the 2'-5'- RNA: RNA complexes was less than DNA: RNA or RNA: RNA complexes and the DNA: 2'-5'-linked RNA complex formation was not detected at all. CD studies indicated that the 2'-5'-RNA: RNA duplexes resembled the conformation of A-family duplexes and the structure was found to resemble unmodified DNA: RNA hybrids. Solution NMR studies indicated that the sugar conformations in the 2'-5'- RNA strand in 2'-5'-RNA: RNA duplex remained in a dynamic C2' to C3'*endo* equilibrium with a predominant C2'-*endo* pucker. Using molecular modeling studies, Yatindra *et al.*[54] also predicted a favored C2'-*endo* sugar pucker.

It was found independently by both Breslow [55] and Switzer [56] that 2'-5' linkages in *iso*DNA formed hybrid duplexes with RNA and the stability of these duplexes was as good as unmodified DNA: RNA duplexes. The *iso*DNA completely failed to recognize the corresponding isosequential DNA. CD spectral data and the free energy changes indicated that isoDNA:RNA duplexes are structurally similar to native DNA:RNA duplexes.[56] Comparison of different *iso*DNA:DNA, isoDNA:RNA and *iso*RNA:RNA duplexes was found to be consistent with the probability of adopting A-type structures by these systems. 2'-5' *iso*DNA:DNA complexation was thought to be less stable based on two aspects. First, the reluctance of natural DNA to assume A-type conformation in this particular case and second, the inability of the *iso*DNA strand to override this preference. The 2'-5' RNA/DNA (C2' *endo*) and RNA (C3' *endo*) both adopt compact sugar-phosphate backbone and result in stable duplex structures. The NMR solution structure of 2'-5'DNA: RNA duplex also favored non-B DNA-type compact duplex structure.[57]

Conformational constraint was introduced in the form of a bicyclic 3'-O, 4'-C methylene ribonucleoside, in which predominant S-type conformation (Figure 18) was predicted by NMR and PM3 calculation. Incorporation of this monomer in an oligomer decreased

binding to RNA although RNA versus DNA selectivity was maintained.[58] Conversely, NMR studies revealed that the substitution of 3'OH group in 2'-5' linked RNA by 3'methoxyethoxy (MOE) group resulted in N type conformational preorganization that offered favorable hybridization properties.[59] The inversion of configuration at 3'-OH in

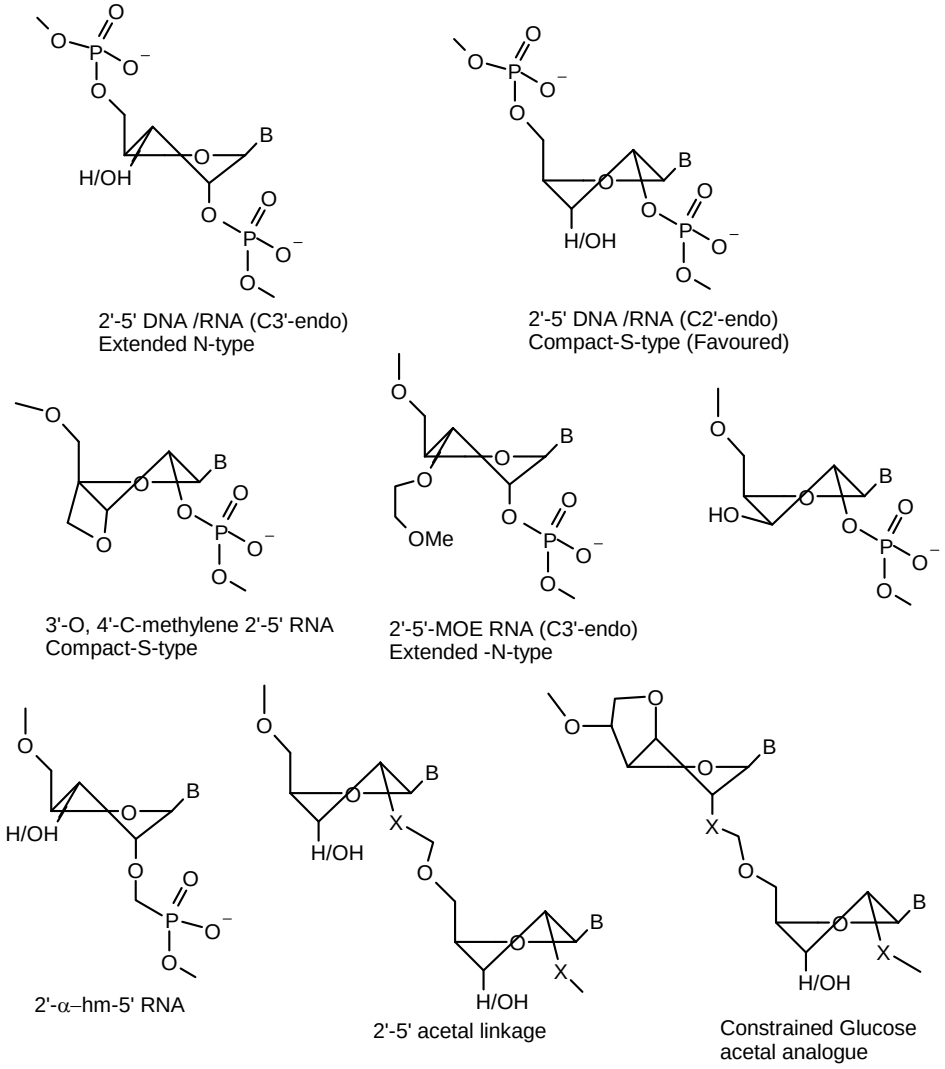


Figure 18. 2'-5' iso DNA/ RNA and their analogues

2'-5' RNA, led to the synthesis of 2'-5' xylose nucleic acids (XNA). The sugars of 2'- 5' XNA adopt C3' *endo* conformation and may therefore adopt an extended (DNA-like) conformation. The pre-organized extended conformation of 2'-5' XNA bound very weakly to the compact RNA target. The conformational incompatibility or spatial mismatching of two strands was thought to be the reason why extended XNA bound

poorly to RNA.[60] In the A-type compact helices of 2', 5'-RNA: RNA duplexes were shown to have smaller interstand phosphate-phosphate distances and could have been the cause of destabilization. This point was considered by Peng and Damha, when they introduced an extra C-atom in 2'-hydroxymethyl- 2', 5'-RNA oligomers, in order to lengthen the backbone and reduce the interstrand P-P repulsion, still keeping compact geometry. This modification unfortunately destabilized the complex with RNA.[61]

The achiral uncharged phosphodiester replacement of 2'-5' connection was explored by using 2'-5' formacetal and thioformacetal linkages. Molecular modeling studies revealed reduced steric space for 2'-5' connection and the shorter non-bulky linkage between 2' and 5' positions was expected to be favored for both A and B conformations. The 2'-5' formacetal ONs resulted in comparable hybridization with both DNA and RNA as predicted by modeling studies.[62] The conformationally restricted 5, 5 bicyclic ribose 2'-5' formacetal was identified as a candidate to pre-order a nucleoside to resemble the bound duplex conformation. Incorporation of this analog in a regular 3'-5' ON led to the destabilization of the duplex. The rigid conformational restriction on the parent ribose ring and loss of flexibility to adopt a conformation competent with duplex formation might have been the cause of this destabilization.[63]

Thus from the current literature it appears that compact A-type conformation in 2'-5' linked RNA could be C2'-endo and would favor the stabilization of the 2'-5' isomeric DNA:RNA structures.

1.3.6 N3'-P5' Phosphoramidate Nucleic Acids

The substitution of the internucleoside phosphodiester for the N3'-P5' phosphoramidates was initially thought as an alternative to biologically stable phosphorothioate nucleic acid analogues.[64] Unlike phosphorothioates, the N3'-P5' phosphoramidates did not give rise to the stereoisomers at phosphorus that complicated the binding patterns with complementary nucleic acids (Figure19). 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.[65] The duplex formed with the RNA target by the N3'-P5' phosphoramidates was

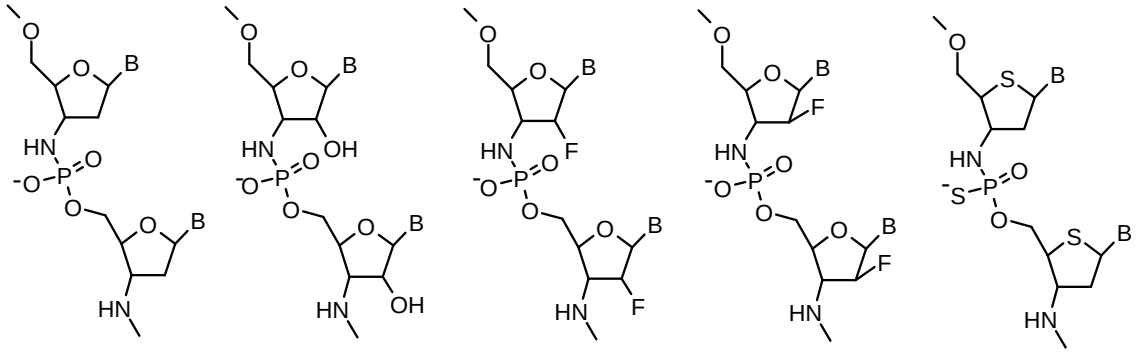


Figure 19. The N3'-P5' phosphoramidate internucleoside

far more stable than the one formed by even the native RNA oligomer. 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 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: 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.[66] Structural analysis of apparently opposing conformational effects of 2'-arabino-fluoro and 3'-amino groups, which shift the conformational equilibrium of the nucleoside sugar puckering to C2'-*endo* (or O4'-*endo*) and C3'- *endo* forms respectively, was studied in 2'-arabino-fluoro N3'-P5' phosphoramidates backbone.[67] NMR analysis revealed that the replacement of 3'-oxygen by a 3'-amino group shifted the 2'- arabino-fluoro furanose

conformational equilibrium towards N-type, similar to that seen for 2'-deoxy-3'-amino nucleosides. The duplexes formed by 2'-arabino-fluoro phosphoramidate and 2'- deoxy phosphoramidate ONs had similar thermal stabilities, which reflected the similarities of their sugar puckering. The most stable complexes with RNA were formed by fully modified 2'-arabinofluoro phosphoramidate ONs.

Oligoribophosphoramidates[68] 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. The increase in melting temperature, ΔTm, was 5-14ºC relative to the 2'-deoxy phosphoramidates for decanucleotides. Also, the thermal stability of the ribophosphoramidate homoduplex was noticeably higher (ΔTm + 9.5ºC) than that for the isosequential 24-deoxy phosphoramidate complex. Furthermore, the oligopyrimidine ribo N3'-P5' phosphoramidate formed an extremely stable triplex with an oligopurine/oligopyrimidine DNA duplex with ΔTm +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.

^{1}H NMR analysis of the α-N3'-P5' phosphoramidate[69] dimer structures indicates significant differences in the sugar puckering of these compounds relative to the β-N3'-P5' phosphoramidates and to the α-phosphodiester counterparts. Additionally, the ability of the α-oligonucleotide N3'-P5' phosphoramidates to form duplexes was studied using thermal denaturation experiments. Thus the N3'-P5' phosphoramidate decamer containing only α-thymidine residues did not bind to poly (A) and exhibited lower duplex

thermal stability with poly(dA) than that for the corresponding β-anomeric phosphoramidate counterpart. A mixed base decamer α-CTTCTTCCTT formed duplexes with the RNA and DNA complementary strands only in a parallel orientation. Melting temperatures of these complexes were significantly lower than for the duplexes formed by the isosequential β-phosphoramidates in antiparallel and parallel orientations respectively. In contrast, the α-decaadenylic N3'-P5' phosphoramidate formed duplexes with both RNA and DNA complementary strands with a stability similar to that of the corresponding β-anomeric phosphoramidate. Moreover, the self complementary oligonucleotide α-ATATATATAT did not form an α:α homoduplex. These results demonstrate the effects of 3'-aminonucleoside anomeric configuration on sugar puckering and consequently on stability of the duplexes.

^{1}H NMR spectroscopic studies demonstrated different sugar conformations of pyrimidine-containing β-phosphoramidates and their α counterparts, being C3'-*endo* and presumably C2'-*endo* respectively. In contrast, 3'-aminodeoxyfuranose conformations of α- and β-adenosine-containing phosphoramidates are more similar to each other and likely correspond to C3'-*endo* puckering. This emphasizes the importance and influence of the anomeric positioning of bases on the 3'-aminonucleoside sugar ring conformation and, consequently, its influence on the oligonucleotide duplex formation properties. Thus pyrimidine-containing α-oligonucleotide phosphoramidates form much less stable parallel duplexes with DNA and RNA strands then their β counterparts in either the parallel or antiparallel orientations, while the thermal stabilities of α- and β-oligoadenylic phosphoramidate duplexes are similar. The pyrimidine-rich β-oligonucleotide N3'-P5' phosphoramidate apparently formed parallel duplexes with complementary DNA and RNA strands, although their thermal stability was much lower than that for the antiparallel duplexes.

Human telomerase is a reverse transcriptase that is expressed in essentially all cancer cells,[70] but not in the vast majority of normal somatic cells. Therefore, the specific inhibition of telomerase activity in tumors might have significant beneficial therapeutic effects. Recent studies showed that oligonucleotide N3'-P5' thio-phosphoramidates as telomerase template antagonists. In biochemical cell-free assays 11-13-mer thio-

phosphoramidate oligonucleotides demonstrated sequence specific and dose dependent inhibition of telomerase with pico-molar IC_{50} values. Optimization of the oligonucleotide sequence and length resulted in the identification of a 13-mer-oligonucleotide thio-phosphoramidate GRN163 as a drug development candidate. In cell cultures GRN163 was able to inhibit telomerase activity in the absence of cationic lipid with ~1 mM IC_{50} values. Telomerase inhibition by GRN163 produced gradual telomere shortening, followed by cellular senescence and/or apoptosis of cancer derived cell lines.

The oligonucleotide thio-phosphoramidates[70] exerted sequence specific telomerase inhibitory activity with IC_{50} values in the pico-molar concentration range in biochemical assays, while in cell culture IC_{50} values were in the high nM- to low-mM concentration range without uptake enhancers. Treatment of cancer cells with the 13-mer thio-phosphoramidate telomerase inhibitor GRN163 resulted in telomere shortening followed by cellular apoptosis. These findings provide rationale for the development of the telomerase template antagonist GRN163 as a novel and specifically designed anticancer chemotherapeutic agent with increased tumor selectivity.

1.3.7 3', 5'-Guanidino linkers

ONs with phosphodiester replacement by guanidine (DNG) and S-methylurea (DNmt) linkages in ONs bind to DNA and RNA with high preference for RNA binding (Figure 20).[71] Predictions based on molecular modeling studies of DNG: DNA and DNG: RNA complexes suggested that DNG in DNG: DNA duplex may primarily retain B-form structure with contracted minor groove due to electrostatic interactions between phosphate and guanidine groups on opposite strands. The DNG: RNA complex, on the other hand adopted A-type structure with furanose sugars in both the strands in C3'-*endo* conformation and also with contracted minor groove. Specific base-pairing was observed when DNG interacted with DNA and RNA, giving highly stable complexes with RNA over DNA, both the hybrids being highly stabilized compared with their natural counterparts. The enhanced stability by electrostatic interactions was supported by experiments at varying ionic environments. The S-methylthiourea linker, similar to DNG was designed to lessen the electrostatic attraction compared to DNG and reduce any nonspecific binding with negatively charged phosphate groups.[72]

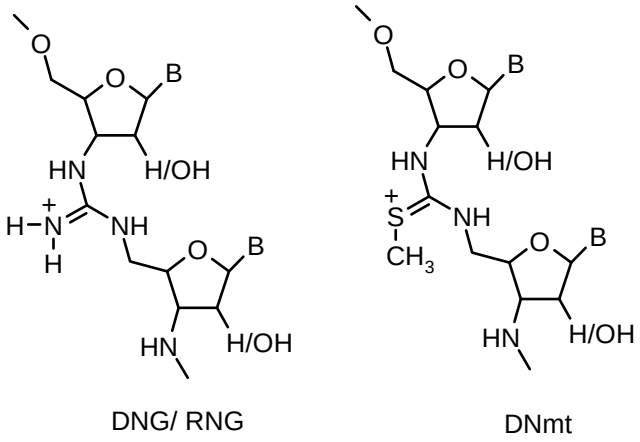


Figure 20. Guanidino and methylthioguanidino DNA/RNA analogues.

The RNG counterpart of DNG containing ribose sugar instead of deoxyribose, sugar was recently shown to bind preferentially to DNA over RNA due to its rigid B-DNA type backbone.[73]

1.3.8 4'-Thio-DNA/RNA

Replacement of deoxyribose by 4'thio-2'deoxyribose (Figure 21) led to 4'thio-DNA,[74] that cross-paired with DNA, RNA and also showed self-pairing. Interestingly, the self-pairing of 4'-thio-DNA showed similar thermal stability as that of the RNA: RNA duplex. The 4'-thio-DNA: DNA stability was similar to RNA: DNA and that of 4'-thio-DNA: RNA was like RNA: RNA, much higher than 4'thio-DNA: DNA duplex. These results indicated that 4'-thio- DNA behaved like RNA despite the absence of 2'-hydroxy group. The CD spectroscopic studies suggested A-form of duplex for 4'-thio-DNA: 4'-thio-DNA, which was similar to RNA: DNA and 4'- thio-DNA: RNA, whereas 4'-thio-DNA: DNA duplex was thought to be an intermediate form between A and B forms. These results, from CD spectroscopy were not in agreement with the crystal structure data of partially modified 4-thio-DNA/DNA duplexes. The crystal structure indicated that such modified ONs fundamentally form B-form of helix. Further studies with groove binders that differentiated RNA and DNA grooves also proved that the 4'-thio-DNA behaved more like RNA than like DNA. It is known that the B form DNA duplex could

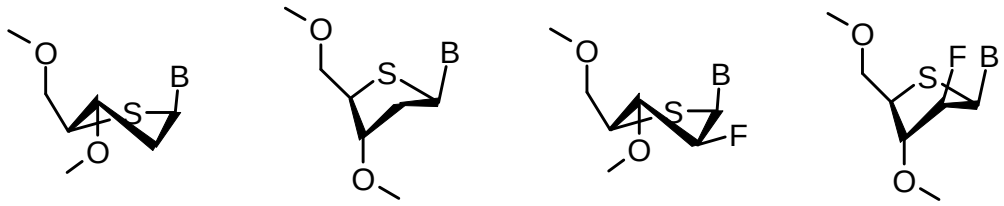


Figure 21. 4'-thio-DNA analogues

be forced to change to A-form helical structures under dehydrating conditions. The substitution of furanose ring oxygen by hydrophobic sulfur in 4'- thio-DNA, could disturb the critical role of 4'-O in stabilizing the minor groove spine of hydration and derive a dehydrated environment. This may be the reason for adopting an A-form conformation. The 4'-thio-deoxynucleoside monomer was found to be in 2'-*endo* sugar conformation in NMR solution studies and also in crystal structure, despite which, it adopted RNA-like conformation in oligomers. Substitution by 2'-arabinofluoro-4'thio sugar in thymidine nucleoside, adopted predominantly N-type conformation in contrast to the S-type conformation in arabinofluorothymidine nucleoside.[75] The distinct conformational switch attained by the replacement of the ring heteroatom was attributed to decrease in the related steric and stereoelectronic effects. Its implications on RNA selectivity are not yet investigated.

1.3.9 Uncharged DNA/RNA analogues

The four-atom amide substitution of four- atom phosphodiester linkage, led to chimeric ONs with alternate amide-phosphodiester groups. Such chimeric ONs bound to RNA with a moderate RNA binding selectivity and with moderately diminished DNA affinity.[76] In contrast, the full replacement of the phosphodiester in the ON by amide linkages, led to the formation of much less stable duplexes with RNA compared to the duplexes with DNA.[77] The higher binding affinity in alternating amide-phosphate backbone has been partially attributed to the absence of a gauche effect arising from oxygen atoms of O4′-C4′-C3′-O3 of the native phosphodiester linkage, causing a C3′-*endo* sugar pucker in 3′- deoxy-3′-carboxymethyl nucleoside derivative. This gives the oligomer a favorable A-form like geometry, in a mixed alternate amide- phosphodiester

backbone. Molecular dynamics simulations of such alternate backbone DNA:RNA complexes favoured exclusive adaptation of C3´-*endo* conformation and A-type DNA:

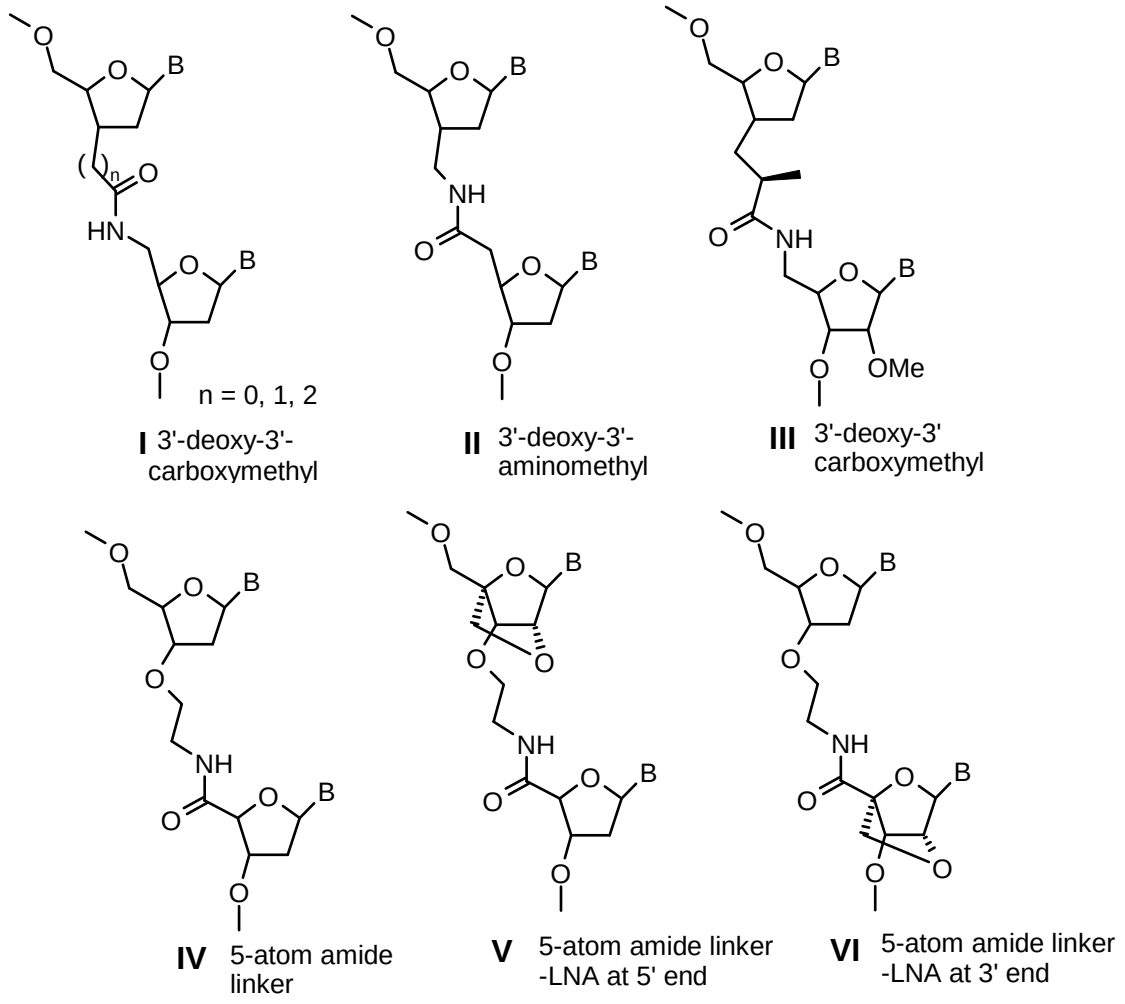


Figure 22. Replacement of phosphodiester by amide linkage

RNA hybrid geometry. The amide (Figure 22, I, n = 1) which formed the most stable backbone could be in constrained planar *trans* conformation and the neighboring phosphate residues (i-1 and i+2) need to compensate one less degree of rotational freedom in the constrained amide.[78] This could also be the reason for lower stability of fully amide-modified DNA: RNA duplexes. The other amide backbones with extended (Figure 22, I, n = 2) and shortened (Figure 22, I, n = 0) amide linkages in a mixed amide-phosphate backbone destabilized the complex with RNA.[77] The amide backbone (Figure

22, I, n = 1 and II) stabilized the complexes with RNA due to the restricted number of low energy conformations compatible with the overall helical structure.[77] In contrast to the above discussion, an additional atom in the five-atom amide linker that provided requisite flexibility was shown to lead to the stabilization of the duplexes with RNA(Figure 22, III).[79] The amide linkers in 2'-*O*-methyl substituted furanose sugar also led to the stabilization of RNA duplexes. The replacement of 3'-*O* with a methylene group and addition of a methyl group in either *R* or *S* configuration (Figure 22, III) also led to stabilization of duplexes.[80] CD studies indicated increase in A- type of duplex structure in these cases. Thermodynamic results on duplex formation of DNA/RNA ONs with multiple amide bonds show less unfavorable entropic component relative to wild type DNA: RNA despite the presence of an additional atom in the backbone. The reduced conformational flexibility of the amide bond leading to favorable pre-organization could be one reason for increased stability of the duplexes. X-ray structure of the duplex with RNA revealed a *trans* amide conformation and that the distance between the neighbouring base pairs was not affected by the longer backbone. An alternate, slightly different five-atom linker was introduced in the dinucleoside. These chimeric amide-phosphate linkages resulted in destabilization of the duplexes (Figure 22, IV). [81] Further, in order to improve the binding efficiency, an LNA nucleoside was introduced as one of the dinucleoside unit in the amide dimer. The LNA monomer improved binding to RNA, when positioned preferably towards the 3'-end of the oligomer (Figure 22, VI).[81] Thus, the LNA monomer could replace the destabilizing effect despite the fact that the linker had an additional atom compared to the natural phosphodiester linkage. The 3'-substituents that could exert conformational restrictions on the ribose ring also showed effective stabilization with RNA targets when positioned towards 3'-end.

The alkyl sulfide internucleoside (Figure 23, I) linkage also were found to adequately replace the phosphodiester linkage in DNA oligomers to form stable complexes with both RNA and DNA but with reduced affinity.[82] A mixed ribo-deoxyribo dimer (Figure 23, II) with alkyl sulfide linkage in an ON showed marked selectivity for binding to RNA and their ability to form duplexes with DNA was abolished.[83] The RNA-binding selectivity exhibited by these oligomers was attributed to the tendency of the 2'-substituted furanose

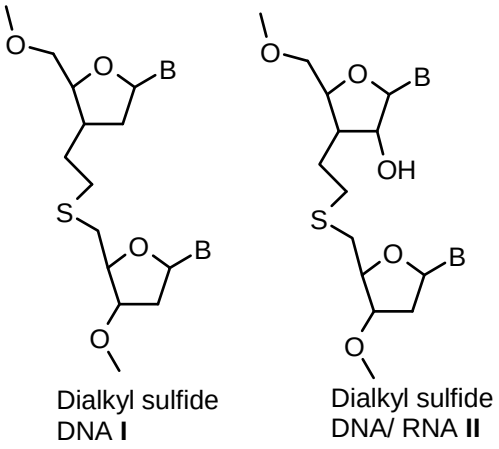


Figure 23. Dialkyl sulfide DNA/ RNA

to adopt a 3'-endo pucker, a conformation that was inconsistent with the B-form structure of helical DNA. The preference of these sugars to remain exclusively in the C3'-endo form was studied by NMR and was attributed to gauche and anomeric effects, making them especially good binders to targeted mRNA.[84]

1.3.10 Hexitol Nucleic Acids (HNA)

The replacement of conformationally flexible deoxyribose by the restricted anhydrohexitol ring resulted in a structurally preorganized HNA (Figure 24, I) to form A-type helices with efficient base stacking.[85] Molecular associations between HNA and RNA were found to be more stable than those between HNA and DNA. ^{1}H NMR analysis of a HNA dimer confirmed the axial orientation of the base moiety with respect to the hexitol ring, and this was used as starting conformation for molecular dynamics study of HNA/RNA and HNA/DNA duplexes. Both complexes showed an A-type geometry and very similar hydrogen bonding patterns between base pairs. The presence of 2'OH group could be responsible for the differential hydration patterns in the minor grove of RNA: HNA and DNA: HNA duplexes. The DNA strand in DNA: HNA complex was found to be more flexible than the RNA strand in RNA: HNA duplexes.[85] In contrast, the crystal structure of HNA: RNA hybrid suggested less-pronounced rigidity of the backbone. The differences in hydration of the HNA and RNA backbones were also observed in the

RNA: HNA duplex crystal structure. The hydration of the HNA strand promoted tighter bridging of the adjacent phosphate groups by water molecules and helped alteration in

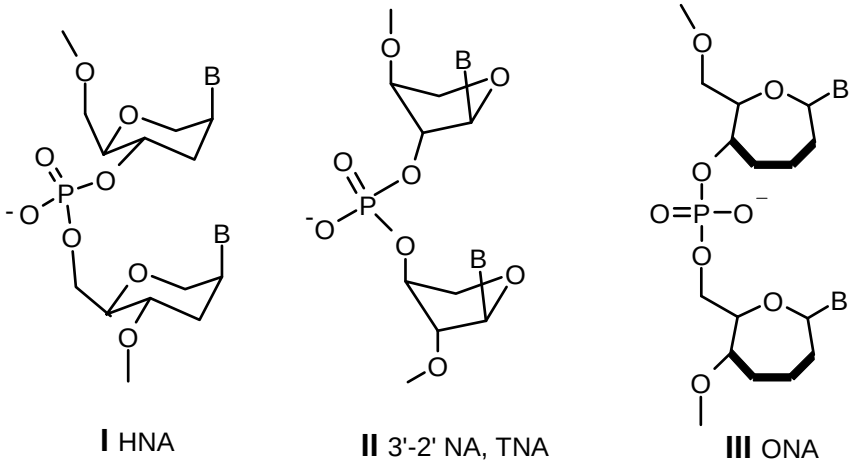


Figure 24. HNA, TNA and ONA

adjacent P-P distances that were close to RNA than DNA. This reinforcing effect upon hydration could be the second reason for overall duplex stability of RNA: HNA duplex in addition to the conformational preorganization imposed by the six membered rings.[86]

1.3.11 L-Threofuranosyl Nucleic Acid (TNA)

3'-2' Nucleic acids (Figure 24, II) cross paired both with RNA and DNA, but the pairing with RNA was found to be stronger in comparison with DNA. The backbone is shorter than both in DNA or RNA as the sugar moiety in TNA contains only four atoms between two adjacent P atoms as compared to five atoms in either DNA or RNA.[87] The detailed crystal structure study of the three dimensional structure provided insight into the origins preferred pairing between TNA and RNA relative to that between DNA and TNA.[87, 88] It was found that the 3' and 2' phosphate groups preferred to be in quasi diaxial orientation which brought the P..P distance of around 5.8 Å. This distance is similar to that between A form of duplexes leading to very stable TNA: RNA structures. Cross pairing between TNA and DNA was permissible because of the allowed conformational adjustments of DNA strand in TNA: DNA duplex and not vice-versa. The close and specific distance

complementarity could be the structural parameter that determined the capability of TNA ONs for preferential cross pairing.

1.3.12 Oxepane Nucleic Acid (ONA)

Recently Sabatino and Damha have developed a new seven membered DNA analog namely oxepane nucleic acid (ONA) (Figure 24, III).[89] These ONs exhibited cross-pairing with complementary RNA to give a duplex whose conformation evaluated by CD spectroscopy very closely matched that of the natural DNA/RNA hybrid. Furthermore, these ONs confer resistance to nucleases while retaining the ability to direct RNase H-mediated cleavage of a target RNA.

1.3.13 Six and seven membered carbocyclic analogs

In case of carbocyclic pyranosyl analogues, cyclohexanyl-nucleic acid (CNA, Figure 25) was prepared in both enantiomeric (D/L) forms and D-CNA hybridizes to complementary RNA as compared to DNA with reduced affinity.[90] A similar analogue is cyclohexenyl-nucleic acid (CeNA, Figure 25), the dehydrogenated form of CNA. Complementary recognition of RNA by the CeNA is more efficient as pure DNA and half efficient as HNA.[91]

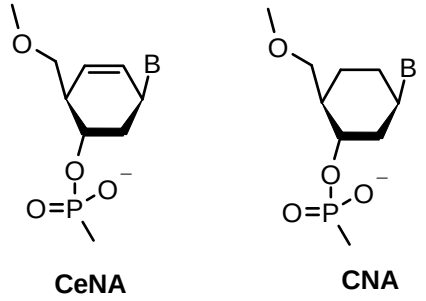


Figure 25. Carbocyclic oligonucleotide analogues.

1.3.14 Sugar-Phosphate modified oligonucleotides

Morpholino nucleotides and peptide nucleic acids are the most prominent modified nucleic acid mimics that involve the replacement of the sugar-phosphate backbone.

1.3.15 Morpholino nucleotides

Summerton *et al*. prepared novel oligonucleotide analogues from ribonucleotide derived morpholine units, linked by carbamate groups (Figure 26a).[92] Cytosine hexamer with stereoregular backbone was prepared in solution phase and was shown to bind poly dG with very high affinity. Solubility characteristics of the resulting oligomer have been improved by terminal conjugation with polyethylene glycol. Fully modified morpholino oligomers where phosphorodiamidate groups are shown to be more effective antisense

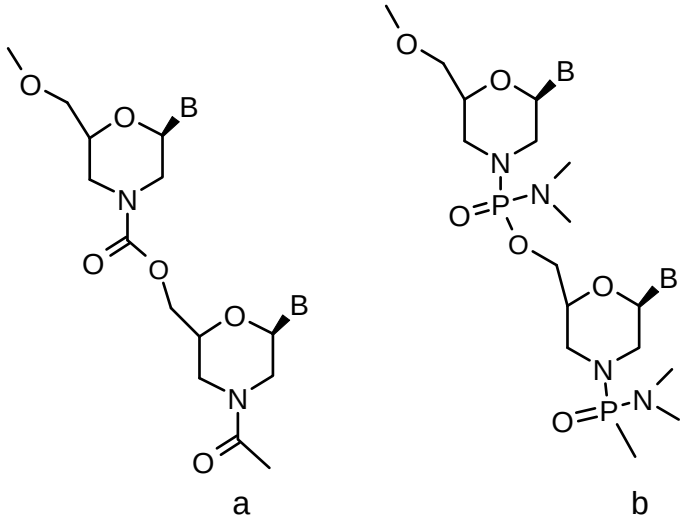


Figure 26. Morpholino ONs linked by **a**) carbamate, **b**) phosphorodiamidate

agents than iso-sequential PS-oligos in cell-free systems and in various cultured cells (Figure 26b). This is attributed to the fact that morpholino oligos, unlike PS-oligos, do not bind to non-specific proteins and are much more sequence specific in their binding to DNA. In one notable study, morpholino oligos were shown to be more effective than PS-oligos as sequence specific antisense inhibitors of tumor necrosis factor-α (TNF-α) in mouse macrophages, despite poor uptake into these cells.[93]

1.3.16 Peptide Nucleic Acids

Peptide nucleic acids (PNAs) (Figure 27) are neutral, achiral DNA mimics that bind to complementary DNA/RNA sequences with high affinity and sequence specificity.[94] In PNA the natural nucleobases are attached via methylene carbonyl linkers to an uncharged

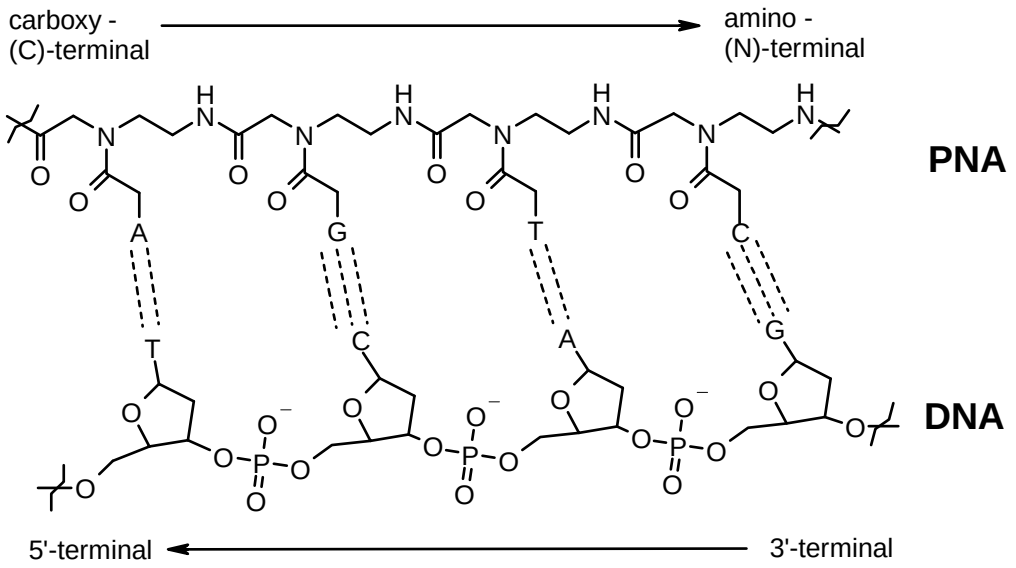


Figure 27. An antiparallel PNA: DNA duplex showing chemical structure of PNA (upper strand) as compare to that DNA (lower strand).

pseudopeptide backbone composed of repeating *N*-(2-aminoethyl)glycyl units. PNA hybridizes to complementary DNA/RNA sequences via specific base complementation to form duplexes for mixed sequences and triplexes for homopyrimidine/homopurine sequences.[95] The complexes of PNA with DNA/RNA sequences generally show thermal stabilities higher than the corresponding DNA-DNA/RNA complexes,[96] depending on sequence. PNAs and their analogues are resistant to proteases and nucleases. Due to these exceptional properties PNAs have major applications as a tool in molecular biology,[97] as lead compounds for development of gene targeted drugs via antisense/antigene technology,[98] for diagnostics and biosensors,[99] and as building blocks for designing PNA supramolecular constructs.[100]

The main limitations of PNAs are its poor water solubility and lack of cell permeability coupled with ambiguity in DNA/RNA recognition arising from its equally facile binding in a parallel/antiparallel fashion (Figure 27) with the complementary nucleic acid sequences. These limitations are being systematically addressed with rationally modified PNA analogues.[101] The solubility of PNAs was improved through conjugation with cationic ligands such as lysine[96] at the N/C terminus of the PNA, or guanidine in backbone[102] without compromising the hydrogen-bonding specificity. The cell

penetrability has been improved by conjugation with various transfer molecules such as cell-penetrating peptides.[103] The equal binding of PNA in parallel and antiparallel orientations to DNA/RNA reduces its target specificity, which can have serious implications for therapeutic applications depending on the other gene that gets affected. This issue was addressed by introduction of chirality into achiral PNA backbones to effect orientational selectivity in complementary DNA/ RNA binding.[104] The ambiguity in DNA/RNA binding would lead to recognition of nonspecific target sequences, and achieving PNA specificity in binding is a desirable goal for successful therapeutic applications.

The relatively high binding affinity of PNAs toward natural oligonucleotides is attributed to the lack of electrostatic repulsion between the uncharged PNA backbone and the negatively charged sugar-phosphate backbone of DNA/RNA.[96] The single-stranded form of PNA, being acyclic, is conformationally flexible in its different structural segments. Consequently, formation of PNA:DNA/RNA complexes is accompanied by conformational changes in PNA to gain enthalpic advantage by hydrogen-bonding and base-stacking interactions accompanied by an undesirable decrease in entropy.[105] Any further increase in the conformational freedom in *aeg*-PNA through extended backbone[106] or side-chain structures[107] lead to substantial reduction in the stability of the resulting PNA: DNA/RNA complexes.

1.3.17 Conformationally constrained PNA- analogues for selective RNA recognition

Introduction of chirality in PNA analogues could attain selectivity for parallel versus antiparallel binding with nucleic acids. An early example of PNA analogue that induced kinetic selectivity for RNA binding was the POM-PNA, a pyrrolidine-amide oligonucleotide mimic (Figure 28, II).[108] The pyrrolidine ring in the 2*R,* 4*R* stereochemistry was selected as it was expected to adopt the sterically less demanding *trans* relative configuration making the system a stereochemical match with native nucleic acids. X-ray crystallographic data for a similar protonated pyrrolidine ring system showed that the alkyl substituents adopted the *trans* relative stereochemistry with close conformational resemblance to native RNA like C3' *endo* sugar conformation.[108a] It was found that short pentameric thyminyl- POM sequences bind much faster to RNA than to

DNA. Introduction of an additional methyl group to restrict the conformational freedom in the backbone led to destabilizing effect for binding to either DNA or RNA.[108b] The binding of oligomers of enantiomeric pyrrolidinyl-PNA with 2*S*, 4*S* stereochemistry (Figure 28**,** III**,** n = 1) were found to destabilization of complexes with either DNA or RNA.[109] A five-atom amide linker between the nucleosides that stabilized the complexes

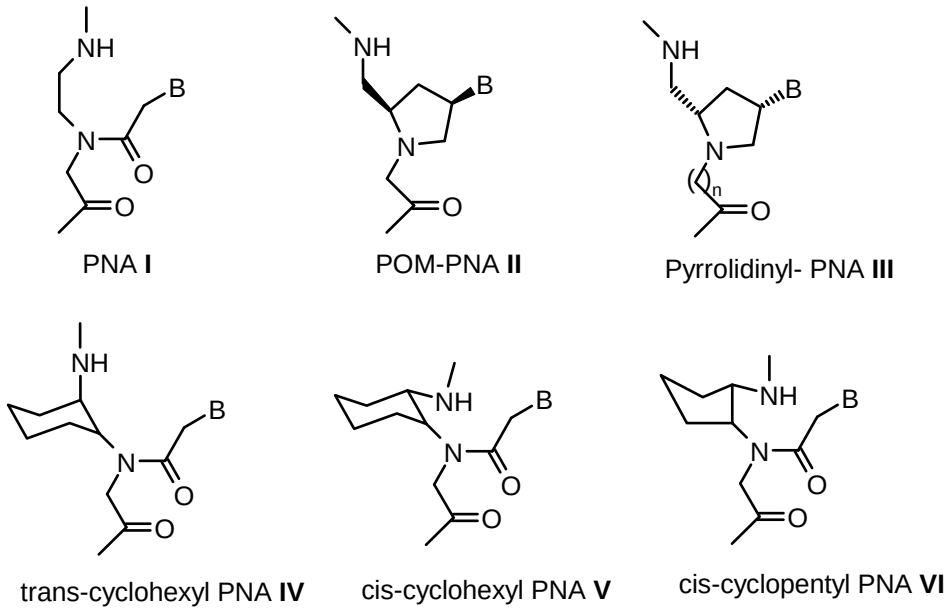


Figure 28. Peptide Nucleic Acid (PNA) and its conformational constrained analogues.

formed with RNA to the 2*S*, 4*S*-pyrrolidinyl-PNA and arrived at the backbone extended pyrrolidinyl PNA (*bep*PNA) (Figure 28, III**,** n = 2). Indeed, alternating PNA:*bep*PNA backbone formed very stable complexes with the target RNA.[110] The most interesting results were obtained when structural constraint in the aminoethyl segment of PNA was introduced in the form of cyclopentyl/cyclohexyl ring systems. As early as in 1996, it was shown by Nielsen *et al.* that a *trans* cyclohexyl system in *aeg*PNA destabilized the PNA:DNA/RNA complexes (Figure 28, IV).[111] Comparison of the X-ray and NMR structures of PNA:DNA, PNA:RNA and PNA_2:DNA complexes revealed that the dihedral angle β could play a very important role in deciding the overall stability of these complexes.[112] The angle β remains near 60-65º in RNA:PNA duplexes and in PNA_2:DNA triplexes. The angle β was then frozen in the range 60-65º by adopting the

cis stereochemistry in the conformationally rigid cyclohexanyl PNA analogue (Figure 28, V). X-ray studies on either of the *cis R,S* and *S,R* stereoisomeric cyclohexanyl PNA monomers confirmed that the angle β remains in the range of 60-65º.[112] Introduction of one to three *cis* cyclohexanyl units in the *aeg*PNA led to very strong stabilization of PNA: RNA duplexes and triplexes at the same time the PNA:DNA duplex was considerably destabilized.[113] In contrast, the *cis*/*trans*-cyclopentanyl PNA, consisting of a conformationally flexible five-membered ring stabilized the duplex/triplex structures with both DNA and RNA (Figure 28, VI)[113,114]

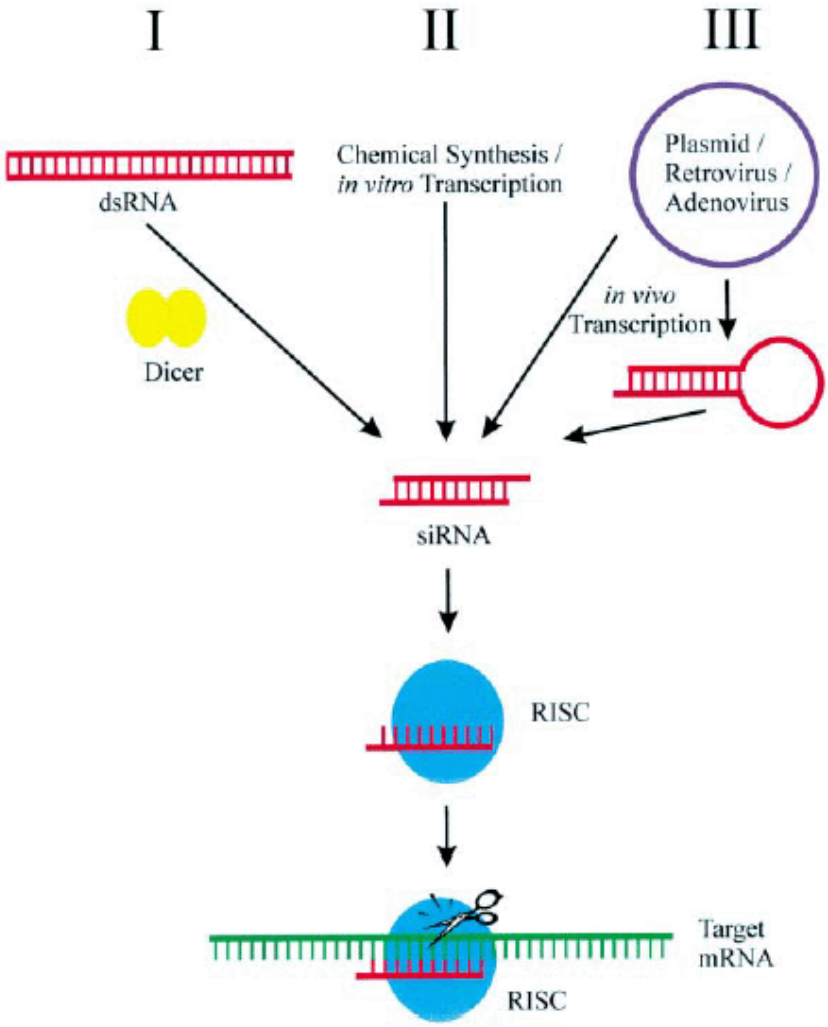
1.4 RNA Interference Oligonucleotides

Only recently, research in the antisense field increased in impact by the discovery of RNA interference (RNAi). This naturally occurring phenomenon as a potent sequence-specific mechanism for post-transcriptional gene silencing was first described for the nematode worm *Caenorhabditis elegans.*[115] RNA interference is initiated by long double-stranded RNA molecules, which are processed into 21-23 nucleotides long RNAs by the Dicer enzyme (Figure 29). This RNase III protein is thought to act as a dimer that cleaves both strands of dsRNAs and leaves two-nucleotide, 3' overhanging ends. These small interfering RNAs (siRNAs) are then incorporated into the RNA-induced silencing complex (RISC), a protein-RNA complex, and guide a nuclease, which degrades the target RNA.

This conserved biochemical mechanism could be used to study gene functions in a variety of model organisms, but its application to mammalian cells was hampered by the fact that long double-stranded RNA molecules induce an interferon response. It was therefore a revolutionary breakthrough, when Tuschl and coworkers could show that 21 nucleotide-long siRNA duplexes with 3' overhangs can specifically suppress gene expression in mammalian cells.[116] This finding triggered an enormous number of studies using RNAi in mammalian cells, as it is thought to provide a significantly higher potency compared to traditional antisense approaches.

Interestingly, not only short double-stranded RNA molecules but also short hairpin RNAs (shRNAs), i.e. fold-back stem-loop structures that give rise to siRNA after intracellular processing, can induce RNA interference.[117] This opened up the possibility

of constructing vectors expressing the interfering RNA for long-term silencing of gene expression in mammalian cells.[118]



.

Figure 29. **Gene silencing by RNA interference (RNAi).** RNAi is triggered by siRNAs, which can by generated in three ways. (I) Long double-stranded RNA molecules are processed into siRNA by the Dicer enzyme; (II) chemically synthesized or *in vitro* transcribed siRNA duplexes can be transfected into cells; (III) the siRNA molecules can be generated *in vivo* from plasmids, retroviral vectors or adenoviruses. The siRNA is incorporated into the RISC and guides a nuclease to the target RNA

An alternative approach to prolong siRNA-mediated inhibition of gene expression is the introduction of modified nucleotides into chemically synthesized RNA, despite the fact that even unmodified short double-stranded RNA revealed an unexpectedly high stability in cell culture and *in vivo*. For certain applications, however, further enhancement of the

siRNA stability might be desirable. Therefore, modified nucleotides were introduced to the ends of both strands.[119] A siRNA with two 2'-*O*-methyl RNA nucleotides at the 5' end and four methylated monomers at the 3' end was as active as its unmodified counterpart and led to a prolonged silencing effect in cell culture.

The first promising *in vivo* experiments with siRNA have already been performed and further therapeutically important genes are expected to be targeted soon. No toxic reactions after siRNA application have been observed in the studies performed to date, but care has to be taken to rule out severe side-effects of long-term induction of RNAi before trials can be started to treat human diseases. Because silencing of gene expression by siRNAs is similar to traditional antisense technology, researchers will be able to benefit from the lessons learned for more than a decade such as the requirement to use proper controls to proof a specific knock-down of gene expression and a careful analysis of possible unspecific effects mediated by the immune system.

1.5 Ribozymes

Ribozymes are ribonucleic acid (RNA) molecules with enzymatic activity that have a great potential as therapeutic entities because of their ability to either cleave deleterious RNAs or repair mutant cellular RNAs.[120] They form base pair specific complexes and catalyze the hydrolysis of specific phosphodiester bonds, causing RNA strand cleavage. Differences exist between ribozymes in size and structure and, although most naturally occurring ribozymes cleave intramolecularly in a *cis* linkage, the RNA component of RNase-P, which is involved in the processing of pre-t RNA molecules, acts in *trans,* i.e., intermolecularly.[121]

Like its protein counterpart, a catalytic RNA or ribozyme greatly accelerates the rate of a biochemical reaction and shows extraordinary specificity with respect to the substrates it acts upon and the products it produces. Ribozymes can cleave the normally unreactive bonds of a phosphodiester linkage in an RNA molecule resulting in a 3’ hydroxyl (3’OH) and 5’ phosphate (5’OH) and 3’ or 2’ 3’-cyclic phosphate.[122] There are several different classes of ribozymes: the self-splicing group I and group II introns; RNase P; and several distinct catalytic motifs found in the small pathogenic RNAs.

For ribozymes to become realistic therapeutic agents several obstacles need first to be overcome. These obstacles are the efficient delivery to a high percentage of the cell population, efficient expression of the ribozyme from a vector or intracellular ribozyme concentration, colocalization of the ribozyme with the target, specificity of ribozyme for the desired mRNA, and an enhancement of ribozyme-mediated substrate turnover. Despite these reservations, results with ribozymes so far look promising, particularly in the HIV-1 studies. Stable ribozymes with high catalytic activity were obtained by systematically modifying naturally occurring ribozymes[123] or by in vitro selection techniques. Several ribozymes are currently being investigated in clinical trials.[17]

1.6 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:

1.6.1. 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. As a result, the UV absorption of 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 melting temperature (T_m) is defined as the temperature at which half of the DNA strands are in the double-helical state and half are in the "random-coil" states (Figure 30). The melting temperature depends on both the length of the sequence as well as the base composition of the sequence.

Methods of estimating T_m

Several formulas are used to calculate T_m values.[124] Some formulas are more accurate in predicting melting temperatures of DNA duplexes. Here we will discuss the most widely used method of determination of T_m.

Nearest neighbour method

This is far more accurate method [125] used to predict melting temperatures of nucleic acid duplexes. Although GC content plays a large factor in the hybridization energy of double-stranded DNA, interactions between neighbouring bases along the helix means that stacking energies are significant. The nearest-neighbour model accounts for this by considering adjacent bases along the backbone two at a time. Each of these has enthalpic and entropic parameters, the sums of which determine melting temperature according to the following equation:

$$\Delta T_m = [\ \Delta H / \{\ \Delta S + R \ln (2C_1 - C_2/2)\}] - 273.15\ ^\circ C$$

where ΔH is the standard enthalpy and ΔS is the standard entropy for formation of the duplex from two single strands,

C_1 is the initial concentration of the single strand that is in excess (usually probe, primer),

C_2 is the initial concentration of the complementary strand that is limiting (usually target), R is the universal gas constant 1.987 cal mol $^{-1}K^{-1}$.

Recently, Warren A. Kibbe from Northwestern University has developed OligoCalc as a web-accessible, client-based computational engine for reporting DNA and RNA single-stranded and double stranded properties, including molecular weight, solution concentration, melting temperature, estimated absorbance coefficients, inter-molecular self-complementarity estimation and intramolecular hairpin loop formation.[126] OligoCalc has a familiar 'calculator' look and feel, making it readily understandable and usable. OligoCalc incorporates three common methods for calculating oligonucleotide-melting temperatures, including a nearest-neighbor thermodynamic model for melting temperature (Figure 30). Since it first came online in 1997, there have been more than 900,000 accesses of OligoCalc from nearly 200,000 distinct hosts, excluding search engines. OligoCalc is available at http://basic.north western.edu/biotools/OligoCalc.html, with links to the full source code, usage patterns and statistics at that link as well.

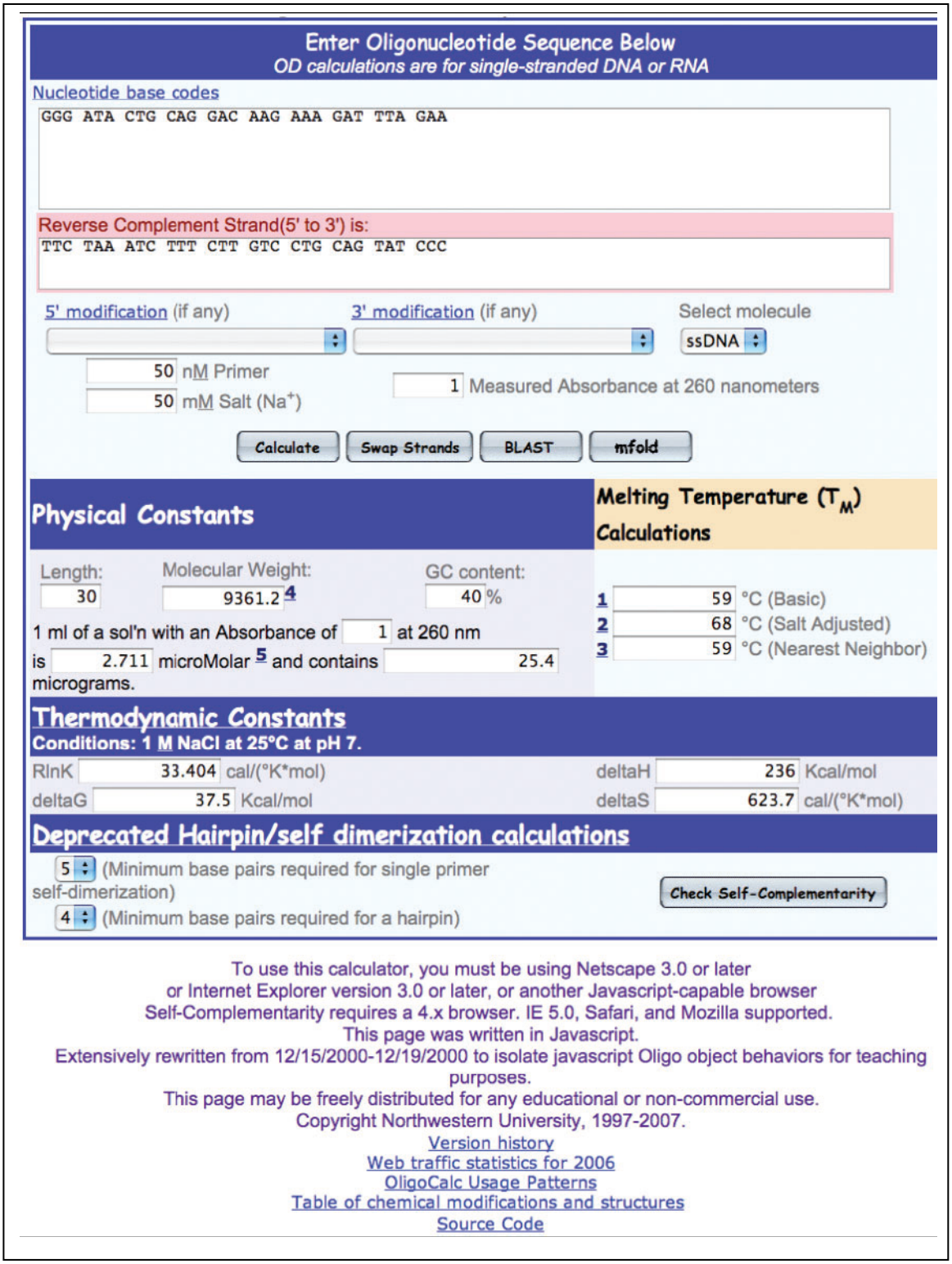


Figure 30. Entry and main calculation screen for OligoCalc

The DNA melting is readily monitored by measuring its absorbance at a wavelength of 260 nm. A plot of absorbance against the temperature of measurement gives a sigmoidal

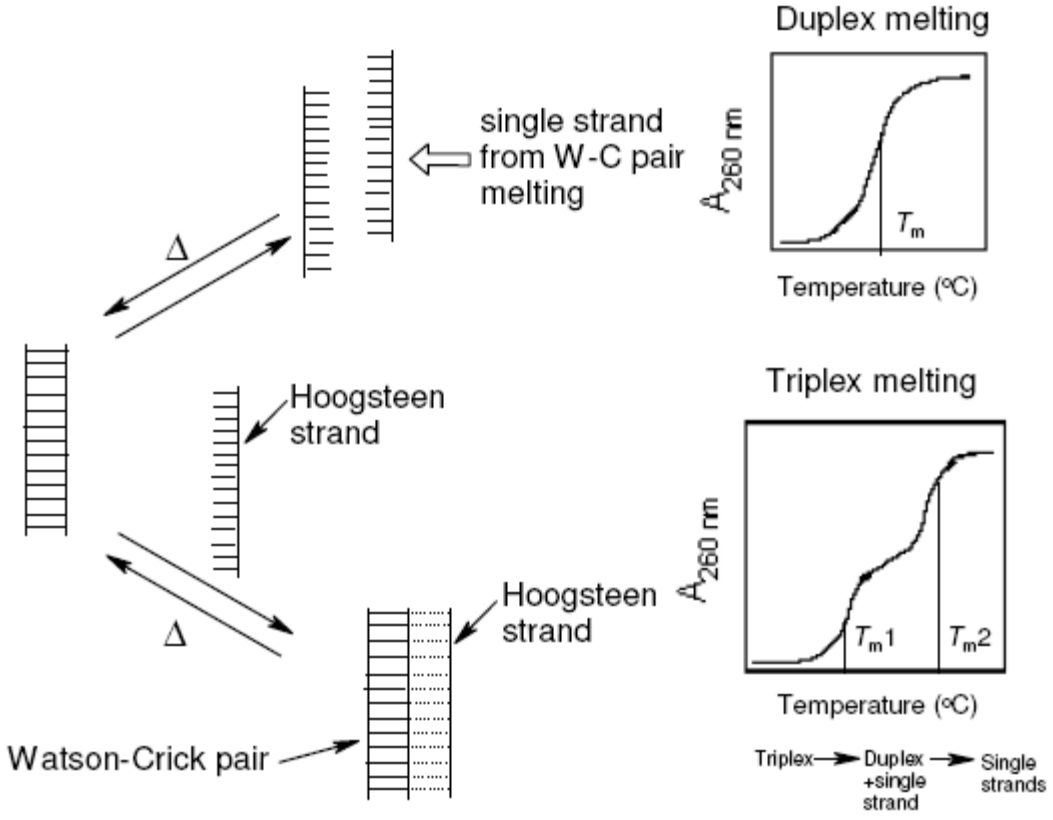


Figure 31. Schematic representation of DNA duplex and triplex melting

curve in the case of duplexes and the midpoint of the transition gives the *Tm*. In case of triplexes, the first dissociation leads to the duplex (Watson-Crick duplex) and the third strand (Hoogsteen strand), followed by the duplex dissociation at higher temperature into two single strands. The DNA triplex melting shows a characteristic double sigmoid transition with separate melting temperature (*Tm*) for each transition. The lower melting temperature corresponds to the triplex ↔ duplex transition while the second transition gives the *Tm* of the duplex ↔ single strand (Figure 31).

Stoichiometry: The binding stoichiometry of oligonucleotide to target DNA/RNA can be determined by UV-titration (mixing) and this can be 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.

1.6.2 Circular Dichroism (CD)

Circular dichroism (CD) is a form of spectroscopy based on the difference in absorption of left- and right-handed circularly polarized light as a function of wavelength. Since chiral molecules interact with the right and left circularly polarized lights differently; in a CD experiment, equal amounts of left and right circularly polarized light are radiated into a (chiral) solution. One of the two types is absorbed more than the other one and this wavelength dependent difference of absorption is measured, yielding the CD spectrum of the sample. CD is particularly useful for studying chiral molecules and has very special significance in the characterization of biomolecules (including the secondary structure of proteins and the handedness of DNA). The commonly used units in current literature are the mean residue ellipticity (degree cm^2 $dmol^{-1}$) and molar ellipticity [θ], the difference in molar extinction coefficients, [θ] = 3298 ($\Delta\varepsilon$).

1.6.3 Application of CD spectroscopy to biological molecules

In general, this phenomenon will be exhibited in absorption bands of any optically active molecule. As a consequence, circular dichroism is exhibited by biological molecules, because of the dextrorotary (e.g. some sugars) and levorotary (e.g. some amino acids) molecules they contain. Even more important is that a secondary structure will also impart a distinct CD to its respective molecules. Therefore, the alpha helix of proteins and the double helix of nucleic acids have CD spectral signatures representative of their structures. The far-UV (ultraviolet) CD spectrum of proteins can reveal important characteristics of their secondary structure. CD spectra can be readily used to estimate the fraction of a molecule that is in the alpha-helix conformation, the beta-sheet conformation, the beta-turn conformation, or some other (e.g. random coil) conformation.

The simplest application of CD to DNA structure determination is for identification of polymorph present in sample.[127] The CD signature of the B-form DNA is a positive band centered at 275 nm, a negative band at 240 nm, with cross over around 258 nm. 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.

The left handed Z-form DNA characterized by a negative CD band at 290 nm and positive band at 260 nm. Care must be taken in using these signatures to identify Z form DNA, since the same DNA in A-Form has a negative band at 295 nm and a positive band at 270 nm. A more definitive signal is the large negative CD signal in the 195-200 nm region for Z-DNA, where as B-form DNA CD is near zero or positive in this region. The sign of the CD signature tells the handedness of the condensed DNA particle, with a negative signal above 250 nm corresponding to a left handed helix. CD spectra are easy to measure that often the simplest technique to probe DNA conformational changes as a function of ionic strength, solvent, ligand concentration etc. When the twist direction is reversed from B-DNA to Z-DNA, [128] an inversion of the CD spectrum might be expected. The salt dependence of DNA CD was correlated with changes in the helix winding angle (the angle between the neighboring bases) deduced from measurements of super-helicity in closed circular DNA.[129] The CD at 275 nm is highly sensitive to the winding angle.[130] On going from dilute salt to 3.0 M CsCl, for example, the 275 nm band is nearly disappeared. CD denaturation (melting) and CD-Job's plots are equally important as they give T_m and binding stoichiometry of DNA/RNA complexes respectively.

1.7 The Present work

The introduction has given an overview on the concept of gene therapeutic techniques, various types of modified antisense oligonucleotides to improve the activity of natural oligonucleotides and in particular the phosphate backbone modified nucleic acids and their applications.

1.7.2 Definition of problem: An important problem from application perspective that has not been adequately dealt is that of non-discrimination of identical DNA and RNA sequences. An analysis of the X-ray studies as well as some previously reported work from our laboratory suggest that the the flexibility of the five-atom amide linker is conserved taking into account the shorter amide bond compared to the phosphodiester linkage which leads to the better RNA binding selectivity. The RNA selectivity of binding seems to be arising from the extended backbone linker that is probably inherently folded to be competent to bind to RNA over DNA.

The present work in the thesis focused on introducing RNA/DNA discrimination into phosphate backbone modified by five atom extended backbone modified ONs.

The thesis comprises studies towards the design, synthesis and DNA/RNA recognition properties of oligonucleotides in which the 4-atom phosphodiester backbone is replaced by five atom amide linkages. The thesis also discusses a convenient method of preparation of cell penetrating peptide-PNA/oligonucleotide conjugates by click-chemistry and conformational studies of deoxy ribose rings on modified monomer and dimer building blocks. The following chapters discuss the synthesis of designed monomers, oligomers synthesis using solid phase synthesis and study of the discrimination properties using various biophysical techniques.

Chapter 2: Section A: This chapter describes the synthesis and biophysical studies of Thioacteamido Nucleic Acids (TANA).

Section B: This section describes the synthesis and RNA binding selectivity of chimeric oligonucleotides modified with five-atom TANA backbone structures

Chapter 3: This chapter reports the conformational studies of TANA monomers and dimers by NMR spectroscopy.

Chapter 4: This chapter describes the synthesis and biophysical studies of (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers.

Chapter 5: This chapter describes the synthesis of cell penetrating peptide-oligonucleotides conjugates by click-chemistry.

1.8 References

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

2 Takahashi, I. and Marmur, J. (1963). *Nature,* **197**, 794-795.

3 (a) Saenger, W. (1984). Principles of nucleic acid structure. Springer-Verlag, New York. (b) Lescrinier, E., Froeyen, M. and Herdewijn, P.(2003) Survey and summary: Difference in conformational diversity between nucleic acids with a six-membered 'sugar' unit and natural 'furanose' nucleic acids. *Nucleic Acid. Res*., **31**, 2975-2989.

4 (a) Hoogsteen, K. (1963) The crystal and molecular structure of a Hydrogen-bonded complex between 1-methyl thymine and 9-methyl adenine. *Acta. Cryst*., **65**, 907. (b) Crick, F. H. C. (1966) Codon-anticodon pairing. The Wobble hypothesis. *J. Mol. Biol*., **19**, 548-555.

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

6 (a) Richmond, et al., (2003) The structure of DNA in the nucleosome core. *Nature*, **423***,* 145-150. b) Leslie, A.G., Arnott., S., Chandrasekaran, R. and Ratliff, R.L. (1980) Polymorphism of DNA double helices. *J. Mol. Biol.*, **143,** 49-72. c) Wahl, M. and Sundaralingam, M.(1997) Crystal structures of A-DNA duplexes. *Biopolymers,* **44 ,** 45-63. d) Ghosh. A. and Bansal, M. (2003). A glossary of DNA structures from A to Z. *Acta Crystallogr D Biol Crystallogr.,* **59**,620–626.

7 Wang, A. H. J., Quigley, G. J., Kalpaks, F. J., Van der M. G., Van Boom, J. H. and Rich, A. (1981) Left-handed double helical DNA: variations in the backbone conformation. *Science,* **211**, 171-176.

8 (a) Hayashi, G., Hagihara, M. and Nakatani, K. (2005). Application of L-DNA as a molecular tag. *Nucleic Acids Symp Ser. (Oxf),* **49**, 261-262. (b) Vargason, J. M., Eichman, B. F. and Ho, P. S. (2000), The extended and eccentric E-DNA structure induced by cytosine methylation or bromination. *Nature Structural Biology,* **7**, 758-761. (c) Wang, G., Vasquez, K. M. (2006). Non-B DNA structure-induced genetic instability. *Mutat Res.,* **598**, 103-119 (d) Allemand, et al (1998). Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Natl. Acad. Sci. USA.,* **24**, 14152-14157.

9 Berg, J. M., Tymoczko J. L. and Stryer, L. (2002). Biochemistry, 5th Edition, WH Freeman and Company, 781-808.

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

11 Barawkar, D. A., Rajeev, K. G., Kumar, V. A. and Ganesh, K. N. (1996) Triplex formation at physiological pH by 5-Me-dC-N4-(spermine) [X] oligonucleotides: non protonation of N3 in X and X*G:C triad and effect of base mismatch/ionic strength on triplex stabilities. *Nucleic Acid Res.*, **24**, 1229-1237.

12 Felsenfeld, G., Davies, R. D. and Rich, A. (1957) Formation of a three stranded polynucleotide molecule. *J. Am. Chem. Soc*., **79**, 2023-2024.

13 (a) Wells, R. D. and Harvey, S. C. (1988) Unusual DNA Structures, Springler-Verlag, New York. (b) Helene, C (1991) *Anti-Cancer Drug Design*, **6**, 569-584. (c) Frank-Kamenetskii, M. D. (1992) Protonated DNA structures. *Methods in Enzymology*, **211**, 180-191. (d) Soyfer, V. N., Potaman, V. N. (1996) Triple Helical Nucleic Acids, Eds., Springler-Verlag, New York.

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

15 Bennett, C. F. and Cowsert, L .M. (1999) Antisense oligonucleotides as a tool for gene functionalization and target validation. *Biochim. Biophys. Acta,* **1489**, 19-30.

16 Driver, S.E., Robinson, G .S., Flanagan, J., Shen, W., Smith, L.E.H., Thomas, D.W. and Roberts, P.C. (1999) Oligonucleotidebased inhibition of embryonic gene expression. *Nat. Biotechnol*., **17**, 1184-1187.

17 Kurreck, J. (2003) *Eur. J. Biochem.,* **270**, 1628–1644.

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

19 Gilbert, W (1978) Why genes in pieces?. *Nature,* **271,** 5645, 501.

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

21 Kumar, V. A. and Ganesh, K. N. (2007) Structure-Editing of Nucleic Acids for Selective Targeting of RNA. *Curr. Topics in Med Chem.*, **7**, 715-726.

22 Eschenmoser, A. and Dobler, M. (1992) Why pentose and not hexose nucleic acids? Part I. Introduction to the problem, conformational analysis of oligonucleotide single strands containing 2', 3'-dideoxyglucopyranosyl building blocks ('homo-DNA') and reflections on the conformation of A- and B-DNA. *Helvetica Chimica Acta,* **75**, 218-259.

23 Tazawa, I., Tazawa, S., Stempel, L. M. and Ts'o, P. O. P. (1970) Conformation and interaction of dinucleoside mono- and diphosphates. 3. L-adenylyl-(3'-5')-Ladenosine and L-adenylyl-(2'-5')-L-adenosine. *Biochemistry,* **9**, 3499 -3514.

24 Anderson, D. J., Reischer, R. J., Taylor, A. J. and Wechter, W. J. (1984) Synthesis of L- $(dUp)_{17}dU$ and the absence of interaction of the compound with poly(dA).*Nucleosides Nucleotides,* **3**, 499-403.

25 Shudo, K., Fujimori, S. (1990) Enantio-DNA recognizes complementary RNA but not complementary DNA. *J. Am. Chem. Soc*, **112**, 7436-7438.

26 Adams, D. A., Petrie, C. R. and Meyer, Jr. R. B. (1991) Preparation and hybridization of oligonucleotides containing 1-α-D-arabinofuranosylthymine. *Nucleic Acids Res*., **19**, 3647-3651.

27 Damha, M. J., Wilds, C. J., Noronha, A., Brukner, I., Borkow, G., Arion, D. and Parniak, M. A.(1998) Hybrids of RNA and Arabinonucleic Acids (ANA and 2'FANA) Are Substrates of Ribonuclease H. *J. Am. Chem. Soc.,* **120** , 12976-12977.

28 (a) Venkateswarlu, D. and Ferguson, D. M. (1999) Effects of C2'-sbstitution on arabinonucleic acid structure and conformation. *J. Am. Chem. Soc.*, **121,** 5609-5610. (b) Minasov, G., Teplova, M., Nielsen, P., Wengel, J. and Egli, M. (2000) Structural Basis of Cleavage by RNase H of Hybrids of Arabinonucleic Acids and RNA. *Biochemistry*, **39,** 3525-3532.

29 Denisov, A. Y., Noronha, A. M., Wilds, C. J., Trempe, J.-F., Pon, R. T., Gehring, K. and Damha, M. J.(2001) Solution structure of an arabinonucleic acid (ANA)/RNA duplex in a chimeric hairpin: comparison with 2'-fluoro- ANA/RNA and DNA/RNA hybrids. *Nucleic Acids Res.*, **29**, 4284 -4293.

30 Gagnor, C., Rayner, B., Leonetti, J.-P., Imbach, J.-L., Lebleu B. (1989) α-DNA IX: parallel annealing of α-anomeric oligodeoxyribonucleotides to natural mRNA is

required for interference in RNase H mediated hydrolysis and reverse transcription. *Nuceic Acids Res.*, **17,** 5107-5115.

31 Thibaudeau, C., Fiildesi, A. and Chattopadhyaya, J. (1998) The Quantitation of the Competing Energetics of the Stereoelectronic and Steric Effects of the 3'-OH and the Aglycone in the α- versus β-D- & -L--2'-deoxyribonucleosides by ^{1}H-NMR. *Tetrahedron*, **54,** 1867-1900.

32 Petersen, M. and Wengel, J. (2003) LNA: A versatile tool for therapeutics and genomics. *Trends Biotechnol*., **21,** 74-81.

33 (a) Nielsen, P. and Dalskov, K. (2000) α-LNA, locked nucleic acid with α-D-configuration. *Chem. Commun*., 1179-1180. (b) Nielsen, P., Christensen, N. K. and Dalskov, J. K. (2002) α-LNA (Locked Nucleic Acid with α -D-Configuration) Synthesis and Selective Parallel Recognition of RNA. *Chem. Eur. J.*, **8**, 712-722.

34 Petersen, M., Håkansson, A.E., Wengel, J. and Jacobsen, J. P. (2001) α-L-LNA (α–I-*ribo* Configured Locked Nucleic Acid) Recognition of RNA. A Study by NMR Spectroscopy and Molecular Dynamics Simulations. *J. Am. Chem. Soc*, **123,** 7431-7432.

35 Keinicke, L., Sørensen, M. and Wengel, J. (2002) α-L-RNA (α-L-ribo configured RNA): Synthesis and RNA selective hybridization of α-L-RNA/α-L-LNA chimera. *Bioorg. Med. Chem. Lett.,* **12**, 593-596.

36 Christensen, N. K., Bryld, T., Sørensen, M. D., Arar, K., Wengel, J. and Nielsen, P. (2004) Parallel nucleic acid recognition by LNA stereoisomers β-L-LNA and α-D-LNA: Studies in the mirror image world. *Chem. Commun*., 282-283.

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

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

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

40 Summerton, J., Stein, D., Huang, S. B., Matthews, P., Weller, S., Partridge, M. (1997) Morpholino and phosphorothioates antisense oligomers compared in cell-free and in-cell systems. *Antisense Nucleic Acid Drug Dev*., **7**, 63-70.

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

42 Checko, K. K., Linder, K., Saenger, W. and Miller, P. S. (1983) Molecular structure of deoxyadenylyl-3'-methylphosphonate-5'-thymidine dihydrate, (d-ApT.2H_2O), a dinucleoside monophoshate with neutral phosphodiester backbone. An X-ray crystal study. *Nucleic Acid Res.*, **11**, 2801-2814.

43 Miller, P. S., Yano, J., Yano, E., Carroll, C., Jayaraman, K. and Ts'o. P. O. P (1979) Nonionic nucleic acid analogues: Synthesis and characterization of dideoxyribonucleoside methylphosphonates. *Biochemistry*, **18**, 5134-5142.

44 Kean, J. M., Kipp, S. A., Miller, P. S., Kulka, M., Aurelian, L (1995) Inhibition of herpes simplex virus replication by antisense oligo-2'-O-methylribonucleoside methylphosphonates. *Biochemistry*, **34**, 14617-14620.

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

46 Letsinger, R. L., Singman, C. N., Histand, G. and Salunke, M. (1988) Cationic oligonucleotides. *J. Am. Chem. Soc*., **110**, 4470-4471.

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

48 (a) Sood, S., Shaw, B. R. and Spielvogel, B. F. (1990) Boron-containing nucleic Acids. Synthesis of oligonucleoside boranophosphates. *J. Am. Chem. Soc*., **112**, 9000. (b) Shaw, B. R., Madison, J., Sood, S. and Spielvogel, B. F (1993) 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., pp. 225-243. (c) Sergueev, D. S. and Shaw, B. R (1998) H-phophonate approach for solid phase synthesis of oligodeoxyribinucleotide boranophosphates and their characterization. *J. Am. Chem. Soc*., **120**, 9417-9427.

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

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

51 Wallace J. C. and Edmonds, M. (1983) Polyadenylated nuclear RNA contains branches. *Proc. Nati. Acad. Sci. U.S.A.*, **80**, 950-954.

52 Kerr, I. M. and Brown, R. E. (1978) pppA2'p5'A2'p5'A: An Inhibitor of Protein Synthesis Synthesized with an Enzyme Fraction from Interferon-Treated Cells. *Proc. Nati. Acad. Sci. U.S.A.,* **75**, 256-260.

53 (a) Giannaris, P. A. and Damha, M. J.(1993) Oligoribonucleotides containing 2',5'-phosphodiester linkages exhibit binding selectivity for 3',5'-RNA over 3',5'-ssDNA. *Nucleic Acids Res.*, **21**, 4742-4749. (b) Wasner, M., Arion, D., Borkow, G., Noronha, A., Uddin, A. H., Parnaik, M. A. and Damha. M. J. (1998) Physicochemical and biochemical properties of 2'-5'- linked RNA and 2'-5'-RNA: 3'-5'-RNA hybrid duplexes. *Biochemistry*, **37**, 7478-7486.

54 Lalitha, V. and Yathindra N. (1995) Even nucleic acids with 2', 5'-linkages facilitate duplexes and structural polymorphism: Prospects of 2', 5'- oligonucleotides as antigene/antisense tool in gene regulation. *Curr. Sci*., **68**, 68-75.

55 Sheppard, T. L. and Breslow, R. C. (1996) Selective binding of RNA, but not DNA by complementary 2'-5'-linked DNA. *J. Am. Chem. Soc*., **118**, 9810-9811.

56 Prakash, T. P., Jung, K.-E. and Switzer, C. (1996) RNA recognition by 2'-structural isomer of DNA. *Chem. Commun*., 1793-1794.

57 Premraj, B. J., Raja, S., Bhavesh, N., Shi, K., Hosur, R. V., Sundarlingam, M. and Yathindra, N. (2004) Solution structure of 2',5' d(G4C4) Relevance to topological restrictions and nature's choice of phosphodiester links. *Eur. J. Biochem*, **271**, 2956-2966.

58 Obika, S., Morio, K.-I., Hari, Y. and Imanishi, T. (1999) Preparation and properties of 2'-5'- linked oligonucleotide analogues containing 3'-O, 4'C-methyleneribonucleosides. *Biorg. Med. Chem. Lett.*, **9**, 515-518.

59 Polak, M., Manoharan, M., Inamati, G. B., Plavec, J. (2003) Tuning of conformational preorganization in model 2'5'- and 3'5'-linked oligonucleotides by 3' and 2' methoxyethyl modification. *Nucleic Acids Res.*, **21**, 2066-2076.

60 Agha, K. A. and Damha, M. J. (2003) Synthesis and binding properties of a homopyrimidine 2', 5'-linked xylose nucleic acid. (2', 5' XNA) *Nucleosides, Nucleotides Nucleic Acids*, **22**, 5-8, 1175-1178.

61 Peng, C. G. and Damha, M. J.(2005) Synthesis and hybridization studies of oligonucleotides containing 1-(2-deoxy-2-α-C-hydroxymethyl-β-D-ribofuranosyl) thymine (2'- α -hm-dt). *Nucleic Acids Res.*, **33**, 7019-7028.

62 Pudlo, J. S., Cao, X., Swaminathan, S. and Matteucci, M. D. (1994) Deoxyoligonucleotides containing 2', 5' acetal linkages: synthesis and hybridization properties. *Tet. Lett.*, **35**, 9315-9318.

63 Zou, R. and Matteucci, M. D. (1996) Synthesis and hybridization properties of oligonucleotide analog containing glucose-derived conformation-restricted ribose moiety and 2'-5' formacetal linkages. *Tet. Lett.*, **37**, 941-944.

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

65 Tereshko, V., Gryaznov, S. and Egli, M. (1998) 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.*, **120**, 269-283.

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

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

68 Gryaznov S. M., Winter, H. (1998) RNA mimetics: oligoribonucleotide N3'-P5' phosphoramidates. *Nucleic Acids Res.* **26**, 4160-4167.

69 Pongracz, K. and Gryaznov S. M. (1998) α-Oligodeoxyribonucleotide N3'-P5'-phosphoramidates: synthesis and duplex formation. *Nucleic Acids Res.,* **26**, 1099-1106.

70 Akiyama, M. *et. al*, (2003) Effects of oligonucleotide N3'-P5' thio-phosphoramidate (GRN163) targeting telomerase RNA in human multiple myeloma cells. *Cancer Res*., **63**, 6187-6194.

71 Dempcy, R. O., Browne, K. A. and Bruice, T. C. (1995) Synthesis of polycation thyminyl DNG, its fidelity in binding polyanionic DNA/RNA, and the stability and nature of the hybrid complexes. *J. Am. Chem. Soc*., **117**, 6140-6141.

72 Arya, D. P. and Bruice, T. C. (1998) Replacement of negative phosphodiester linkages of DNA by positive S-methylthiourea linkers: A novel approach to putative antisense agents. *J. Am. Chem. Soc*., **120**, 6619-6620.

73 Park, M., Toporowski, J. W. and Bruice, T. C. (2006) Ribonucleic guanidine demonstrates an unexpected marked preference for complementary DNA rather than RNA. *Bioorg. Med. Chem*, **14,** 1743-1749.

74 Inoue, N., Minakawa, N. and Matsuda, A. (2006) Synthesis and properties of 4'-thioDNA: Unexpected RNA-like behaviour of 4'-thio DNA. *Nucleic Acids Res*, **34**, 3476-3483.

75 Watts, S., K., Choubdar, N., Mario Pinto, B. and Damha, M. J. (2006) Synthesis and conformational analysis of 2'-Fluoro-5-methyl-4'-thioarabinouridine. *J. Org. Chem*., **71**, 921-925.

76 (a) Lebreton, J., De Mesmaeker, A., Waldner, A., Fritsch, V., Wolf, R. M. and Freier, S. M. (1993) Synthesis of thymidine dimer derivatives containing an amide linkage and their incorporation into oligodeoxyribonucleotides. *Tet. Lett.*, **34**, 6383-4386., (b) Lebreton, J. Waldner, A., Fritsch, F., Wolf, R.M., De Mesmaeker A. (1994) Comparison of two amides as backbone replacement of the phosphodiester linkage in oligodeoxynucleotides. *Tet. Lett.*, **35 ,** 5225-5228.

77 von Matt, P., De Mesmaeker, A., Pieles, U., Zurcher, W. and Altman, K.-H. (1999) 2'- deoxyribo-PNAs: A structurally novel class of polyamide nucleic acids with good RNA and DNA binding affinity. *Tet. Lett.*, **40**, 2899- 2902.

78 Nina, M., Fonne-Pfister, R., Beaudegnies, R., Chekatt, H., Jung, P. M. J., Murphy-Kessabi, F., Mesmaeker, A. and Wendeborn, S. (2005) Recognition of RNA by

amide modified backbone nucleic acids: Molecular dynamics simulations of DNA-RNA hybrids in aqueous solution. *J. Am. Chem. Soc*., **127**, 6027- 6038.

79 Wilds, C. J., Minasov, G., Natt, F., Matt von, P., Altmann, K.-H. and Egli, M. (2001) Studies of a chemically modified oligonucleotide containing a 5-atom backbone which exhibits improved binding to RNA. *Nucleosides Nucleotides Nucleic Acids*, **20**, 991-994.

80 Pallan, P. S., Matt von, P., Wilds, C., Altmann, K.-H., Egli, M. (2006) RNA binding affinities and crystal structure of oligonucleotides containing five-atom amide-based backbone structures. *Biochemistry*, **45**, 8048-8057.

81 Lauristen, A., Wengel, J. (2002) Oligonucleotides containing amide linked LNA type dinucleotides: synthesis and high affinity. *Chem. Commun.*, 530- 531.

82 Kawai, S. H., Wang, D., Giannaris, P. A., Damha, M. and Just, G. (1993) Solid phase synthesis and hybridization properties of DNA containing sulfide-linked dinucleoside. *Nucleic Acids Res.*, **21**, 1473-1479.

83 Meng, B., Kawai, S. H., Wang, D., Just, G., Giannaris, P. A. and Damha, M. (1993) A sulfide linked oligonucleotide analogue with selective hybridization properties. *Angew. Chem Int. Ed. Engl.*, **32**, 729-731.

84 Damha, M. A., Meng, B., Kawai, S. H., Wang, D., Yannopoulos, C. G. and Just, G. (1995) Structural basis for RNA binding selectivity of oligonucleotide analogue containing alkylsulfide internucleoside linkages and 2'-substituted 3'-deoxyribonucleosides. *Nucleic Acids Res.*, **23**, 3967-3973.

85 De Winter, H., Lescrinier, E., Van Aerschot, A. and Herdewijn, P.(1998) Molecular dynamics simulation to investigate differences in minor groove hydration of HNA/RNA hybrids as compared to HNA/DNA complexes. *J. Am. Chem. Soc.*, **120,** 5381-5394.

86 Maier, T., Przylas, I., Strater, N., Herdewijn, P. and Saenger, W. (2005) Reinforced HNA backbone hydration in the crystal structure of a decameric HNA/RNA hybrid. *J. Am. Chem. Soc.*, **127**, 2937-2943.

87 Schöning, K.-U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R. and Eschenmoser, A. (2000) Chemical etiology of nucleic acid structure: The α-threofuranosyl-(3'→2') oligonucleotide System. *Science*, 1347-1351.

88 Pallan, P. S., Wilds, C. J., Wawrzak, Z., Krishnamurthy, R., Eschenmoser, A. and Egli, M. (2003) Why does TNA cross pair more strongly with RNA than with

DNA? An answer from X-ray analysis. *Angew. Chem Int. Ed. Engl*., **42**, 5893-5895.

89 Sabatino, D., Damha, M. J. (2007) Oxepane Nucleic Acids: Synthesis, Characterization, and Properties of Oligonucleotides Bearing a Seven-Membered Carbohydrate Ring. *J. Am. Chem. Soc.*, **129**, 8259-8270.

90 Wang, J., Verbeure, B., Luyten, I., Lescrinier, E., Froeyen, M., Hendrix, C., Rosemeyer, H., Seela, F., Arschot, A. V. and Herdewijn, P. (2000) Cyclohexene nucleic acids (CeNA): Serum stable oligonucleotides that activate RNase H and increase duplex stability with complementary RNA. *J. Am. Chem. Soc*., **122**, 8595-8602.

91 Mourinsh, Y., Rosemeyer, H., Esnouf, R., Mevedovici, A., Wang, J., Ceulemans, G., Lescrinier, E., Hendrix, C., Busson, R., Sandra, P., Seela, F., Aerschot, A. V. and Herdewijn, P. (1999) Synthesis and purity properties of oligonucleotides containing 3-hydroxy-4-hydroxymethyl- 1-cyclohexanyl nucleosides. *Chem. Eur. J*., **5**, 2139-2150.

92 (a) Summerton, J., Stein, D., Huang, S. B., Matthews, P., Weller, S. and Partridge, M (1997) Morpholino and phosphorothioates antisense oligomers compared in cell-free and in-cell systems. *Antisense Nucleic Acid Drug Dev*., **7**, 63-70. (b) Summerton, J. and Weller, D (1997) Morpholino antisense oligomers: Design, preparation, and properties. *Antisense Nucleic Acid Drug Dev.*, **7**, 187-195.

93 Toyler, M. F. *et al* (1998) Effect of TNF-a antisense oligomers on cytokine production by primary murine alveolar macrophages. *Antisense Nucleic Acid Drug Dev*., **8**, 199-205.

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

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

96 Nielsen, P. E., Egholm, M. and Buchardt, O. (1994) Peptide nucleic acid (PNA): A DNA mimic with a peptide backbone. *Bioconjugate Chem.*, **5**, 3-7.

97 (a) Nielsen, P. E. and Egholm, M. (1999) *Peptide Nucleic Acids. Protocols and Applications*; Horizon Press: Norfolk. (b) Ray, A. and Norden, B. (2000) Peptide

nucleic acid: its medical and biotechnological applications and promise for the future. *FASEB J.*, **14**, 1041-1060.

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

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

100 Nielsen, P. E. (1998) *Pure Appl. Chem.*, **70**, 105-110. (b) Corey, D. R. (1997) Peptide nucleic acids: expanding the scope of nucleic acid recognition. *Trends Biotechnol.*, **15**, 224-229.

101 (a) Ganesh, K. N. and Nielsen, P. E. (2000) Peptide nucleic acids: Analogs and derivatives. *Curr. Org. Chem.*, **4**, 931-943. (b) Kumar, V. A. (2002) Structural pre-organization of peptide nucleic acids: chiral cationic analogues with five- or six-membered ring structures. *Eur.J. Org. Chem.*, 2021-2032. (c) Kumar, V. A. and Ganesh K. N. (2005) Conformationally Constrained PNA Analogues: Structural Evolution toward DNA/RNA Binding Selectivity. *Acc. Chem. Res.*, **38**, 404-412. (d) Uhlmann, E., Breipohl, G. and Will, D. (1998) PNA: Synthetic polyamide nucleic acids with unusual binding properties. *Angew. Chem., Int. Ed. Engl.*, **37**, 2796-2823.

102 Zhou, P., Wang, M., Du, L., Fischer, G. W., Waggoner, A. and Ly, D. H. (2003) Novel binding and efficient cellular uptake of guanidine based peptide nucleic acids. *J. Am. Chem. Soc.*, **125**, 6878-6879.

103 (a) de Koning, M. C., van der Marcel, G. and Overhand, M. (2003) Synthetic developments towards PNA-peptide conjugates. *Curr. Opin. Chem. Biol.*, **7**, 734-740. (b) Eriksson, M., Nielsen, P. E. and Good, L. (2002) Cell permeabilization and uptake of antisense peptide nucleic acid into E. coli. *J. Biol. Chem.*, **277**, 7144-7147.

104 Sforza, S., Corradini, R., Dossena, A. and Marchelli, R. (2000) DNA binding of a D-lysine-based chiral PNA: Direction control and mismatch recognition. *Eur. J. Org. Chem.*, 2905-2913.

105 Tomac, S., Sarkar, M., Ratilainen, T., Wittung, P., Nielsen, P. E., Norden, B. and Graslund, A.(1996) Ionic effects on the stability and conformation of peptide nucleic acids. *J. Am. Chem. Soc.,* **118**, 5544-5549.

106 (a) Kumar, V. A., (2002) *Eur. J. Org. Chem*, 2021-2032. (b) Hyrup, B., Egholm, M., Nielsen, P. E., Wittung, P., Norden, B. and Buchardt, O.(1994) Structure

activity studies of binding of modified peptide nucleic acids (PNAs) to DNA. *J. Am. Chem. Soc.*, **116**, 7964-7970.

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

108 (a) Hickman, D. T., Tan, T. H. S., Morral, J., King, P. M., Cooper, M. A. and Mickelfield, J. (2003) J. *Org. Biomol. Chem.*, **1**, 3277-3292. (b) Khan, A. I., Tan T. H. S.A. and Micklefield J. (2006) Stereospecific backbone methylation of pyrrolidine–amide oligonucleotide mimics (POM) *Chem. Commun.*, 1436-1438.

109 Kumar, V. A., Pallan, P. S., Meena and Ganesh, K. N. (2001) Pyrrolidine nucleic acids: DNA/PNA oligomers with 2-hydroxy/ aminomethyl-4-(thymin-1yl) pyrrolidine-N-acetic acid. *Org. Lett.*, **3**, 1269-1272.

110 (a) Govindaraju, T. and Kumar, V. A. (2005) Backbone Extended Pyrrolidine Peptide Nucleic Acids (*bep*PNA): Design, Synthesis and DNA/RNA binding Studies**.** *Chem. Commun.*, 495-497. (b) Govindaraju, T., Kumar, V. A. (2006) *Tetrahedron*, **60**, 2321-2330.

111 Lagriffoule, P., Wittung, P., Eriksson, M., Jensen, K.K., Norden, B., Buchardt, O. and Nielsen, P.E. (1997) Peptide nucleic acids with a conformationally constrained chiral cyclohexyl-derived backbone. *Chem. Eur. J.*, **3**, 912-919.

112 Govindaraju, T., Gonnade, R. G., Bhadbhade, M. M., Kumar, V. A. and Ganesh, K. N. (2003) (*1S, 2R/1R, 2S*)-2-Aminocyclohex-1-yl glycyl PNA Thymine monomers: Synthesis and Crystal Structures. *Org. Lett.*, **5**, 3013-3016.

113 (a) Govindaraju, T., Kumar, V. A. and Ganesh, K. N. (2005) (*SR/RS*)-Cyclohexanyl PNAs: Conformationally preorganized PNA analogues with unprecedented preference for duplex formation with RNA. *J. Am. Chem. Soc*, **127***,* 4144-4145. (b) Govindaraju, T., Madhuri, V., Kumar, V. A. and Ganesh, K. N. (2006) Cyclohexanyl peptide nucleic acids (*ch*PNA) for preferential RNA binding: Effective tuning of dihedral angle β in PNA for DNA/RNA discrimination. *J. Org. Chem.*, **71**, 14-21.

114 Govindaraju, T., Kumar, V. A. and Ganesh K. N. (2004) (1*S*, 2*R*/1*R*, 2*S*)-*cis*-Cyclopentyl PNA (*cp*PNA) as constrained PNA analogues: Synthesis and Evaluation of *aeg-cp*PNA chimera and stereopreferences in hybridization with DNA/RNA. *J. Org. Chem.,* **69**, 5725-5734.

115 Fire, A., Xu, S.Q., Montgomery, M. K. , Kostas, S. A., Driver, S. E. and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature,* **391**, 806-811.

116 Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in culturedmammalian cells. *Nature*, **411**, 494-498.

117 (a) Yu, J . -Y., De Ruiter, S. L. and Turner, D. L. (2002) RNA interference by expression of short-interfering and hairpin RNAs in mammalian cells. *Proc. Natl Acad. Sci. U.S.A.*, **99**, 6047-6052. (b) Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. and Conklin, D.S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.*, **16**, 948-958.

118 (a) McManus, M.T. and Sharp, P. A. (2002) Gene silencing in mammals by small interfering RNAs. *Nat. Rev.*, **3**, 737-747. (b) Tuschl, T. (2002) Expanding small RNA interference. *Nat. Biotechnol.*, **20**, 446-448.

119 Amarzguioui, M., Holen, T., Babaie, E. and Prydz, H. (2003) Tolerance for mutations and chemicalmodifications in a siRNA. *Nucleic Acids Res.,* **31**, 589-595.

120 Sullenger, B. A. and Cech,T. R. (1994) Ribozyme-mediated repair of defective mRNA by targeted *trans*-splicing. *Nature,* **371**, 619.

121 (a) James, H. A. and Gibson I (1998) The therapeutic potential of ribozymes. *Blood*, **91**, 371-382. (b) James, H. A. and Turner, P. C. (1995) Ribozymes. *Essays Biochem.,* **29,** 175.

122 Cedergren, R. (1990) RNA-The catalyst. *Biochem Cell Biol*, **68**, 903.

123 Beigelman, L. et al (1995) Chemical modification of hammerhead ribozymes. *J. Biol. Chem.,* **270**, 25702-25708.

124 (a) Breslauer, K. J. et al. (1986). Predicting DNA Duplex Stability from the Base Sequence *Proc. Natl. Acad. Sci. USA.* **83**, 3746-3750. (b) Rychlik, W. et al. (1990) *Nucleic Acids Res.* **18**, 6409-6412. (c) Owczarzy R., Vallone P.M., Gallo F.J., Paner T.M., Lane M.J. and Benight A.S (1997). Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers,* **44**, 217-239.

125 SantaLucia, J. Jr. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA ,* **95**,1460-1465.

126 Kibbe, W. A. (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.*, **35**, W43–W46.

127 (a) Stryer, L. (1988) *Biochemistry*, 3rd ed.; New York: W. H. freeman and Company, (b) Egli, M., Williams, L. D., Gao, Q. and Rich, A. (1991) *Biochemistry*, **30**, 11388. (c) Calladine, C. R. and Drew, H. R. (1992) Understanding DNA, The molecule and how it works; Cambridge: Academic Press LTd. (d) Beveridge, D. L. and Jorgensen, W. L. (1986) *Ann. NY Acad. Sci*., 482. (e) Gassner, R. V., Frederick, C. A., Quigley, G. J., Rich, A. and Wang, A. H.-J. J. (1989) The molecular structure of the left handed Z-DNA double helix at 1.0 Å atomic resolution. Geometry, conformation, and ionic interactions of d (CGCGCG). *Biol. Chem*., **264**, 7921. (f) Berova, N., Nakanishi, K. and Woody R.W., Eds. *Circular Dichroism: Principles and Applications*; Wiley-VCH: New York, 2000; pp 703-736.

128 Gray, D. M., Hung, S. - H. and Johnson, K. H. (1995) *Methods Enzymol.*, **246**, 19-34.

129 Anderson, P. and Bauer, W. (1978) Super coiling in closed circular DNA: dependence upon ion type and concentration. *Biochemistry*, **17**, 594-601.

130 Sprecher, C. A., Baase, W. A. and Johnson, W. C. Jr (1979) Conformation and circular dichroism of DNA. *Biopolymers*, **8**, 1009-1019.

Chapter 2

Section I: Thioacetamido Nucleic Acids (TANA): Synthesis and biophysical studies

2.1.1 Introduction

Last two decades have witnessed an upsurge in the synthesis of several modified nucleic acid derivatives with intentions to arrive at nucleic acid analogues having the potential to be therapeutically suitable and commercially viable.[1] The more recent developments such as splice correcting [2] and exon skipping [2] strategies require highly robust nucleic acid analogues that are stable under physiological conditions as single strands as well as in the form of duplexes with complementary RNA sequences. Unmodified natural nucleic acids were initially used as antisense nucleic acids. The relatively rapid degradation of native DNA warranted synthesis of designed molecules modified to enhance their stability towards hydrolytic enzymes. The additional demands of the design being enhanced strength of hybridization with target RNA. In addition, the modified ONs may have improved other properties such as aqueous solubility, efficiency of cellular uptake, degradation of RNA by RNase-H enzyme and above all easy synthetic methodologies. Also the design and synthesis should fulfill the requirement of the enantiomeric purity and integrity of the oligonucleotides. The challenge of meeting all these requirements in any one molecule is a demanding task. Satisfying one condition could be mostly realized at the price of another. The design of phoshorothioate ONs, which is the only FDA approved antisense drug, leads to an inseparable mixture of diastereomers due to chirality at phosphorous. The other currently developing backbone modifications include second and third generation antisense ONs such as 2'-O-alkyl, 2'-O-methylthioethyl, LNA, HeNA, Morpholino NA, PNA or PNA modifications. All these modifications fall in the Venn diagram of modified ONs that exhibit excellent therapeutic properties in some way but lack some other important properties.[3] In chapter 1 we have discussed most of these modifications. The enzymatic stability of peptide nucleic acids (PNAs) gain importance for such applications.[4] The unacharged PNAs (Figure 1) are poorly soluble in water and bind to DNA in either parallel or antiparallel orientation.[5] Also, PNAs require assistance in the form of covalent conjugation with cell penetrating peptides (cpp)[6] or additonal positive charges on PNA[7] to cross the cell membranes for their activity. The replacement

of the internucleoside phosphate linkages by the robust amide bond as in PNA could be an alternate solution so that the advantages of chirality and 3'-5' directionality of the sugar be maintained in the analogue along with the enzymatic stability of the amide bond.

Several four- and five-atom amide-linked oligonucleotide analogues are known in the literature,[3] most of them we have discussed in Chapter 1. The five-atom amide linkers were introduced to compensate the shorter amide bond in the dinucleoside as compared with the four-atom phosphate linkers to maintain the internucleoside distance complementarity. The strategy was found to be successful as some of the five-atom-linked dimers when substituted in ON sequences lead to significantly higher RNA affinities compared with that of the native DNA. It has been shown by CD spectroscopy and crystal structure that the longer amide backbones do not disrupt the duplex geometries.[8] A recent report from our laboratory having a five atom between the nucleosides that stabilized the complexes formed with RNA to the 2*S*, 4*S*-pyrrolidinyl-PNA and arrived at the backbone extended pyrrolidinyl PNA (*bep*PNA) (Figure 1). The alternating PNA:*bep*PNA backbone are forming a very stable complexes with the target RNA.[9]

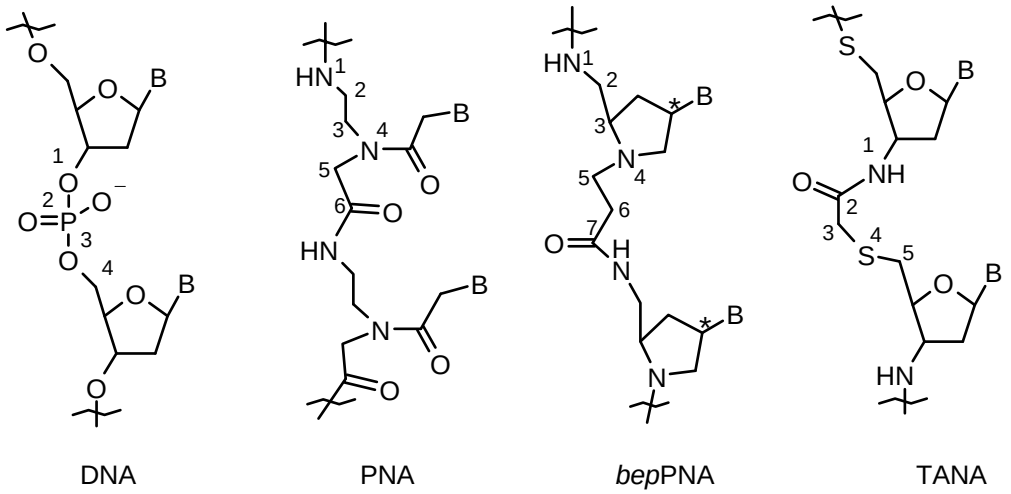


Figure 1. Structure of DNA, PNA, *bep* PNA and TANA

In this chapter we present the essence of our important preliminary results on a new backbone modification, thioacetamido nucleic acids (TANA). The strategy of the design, synthesis of the monomer blocks, oligomer synthesis and and their complementary RNA recognition using UV-Tm measurements is discussed.

2.1.2 Rational and objectives of present work

The 2'-deoxyribose sugar-phosphate backbone in DNA is replaced by the sugar-thioacetamido backbone (Figure 2) considering that 1) In 3'deoxy-3'amino ribose sugar the five-membered heterocyclic ring pucker is prefered to be 3'-endo that is better suited for mRNA recognition 2) The solid phase synthesis and scale-up methodology for an amide bond formation are very well established 3) The thio functionality in the backbone is known to facilitate biodistribution of ODNs 4) The linker group has no chiral center and the flexibility of the five-atom amide linker is conserved taking into account the shorter amide bond compared to the phosphodiester linkage 5) The backbone is now uncharged as in PNA 6) Introduction of an additional atom is much simpler as compared to some earlier literature reports as described in Chapter 1.

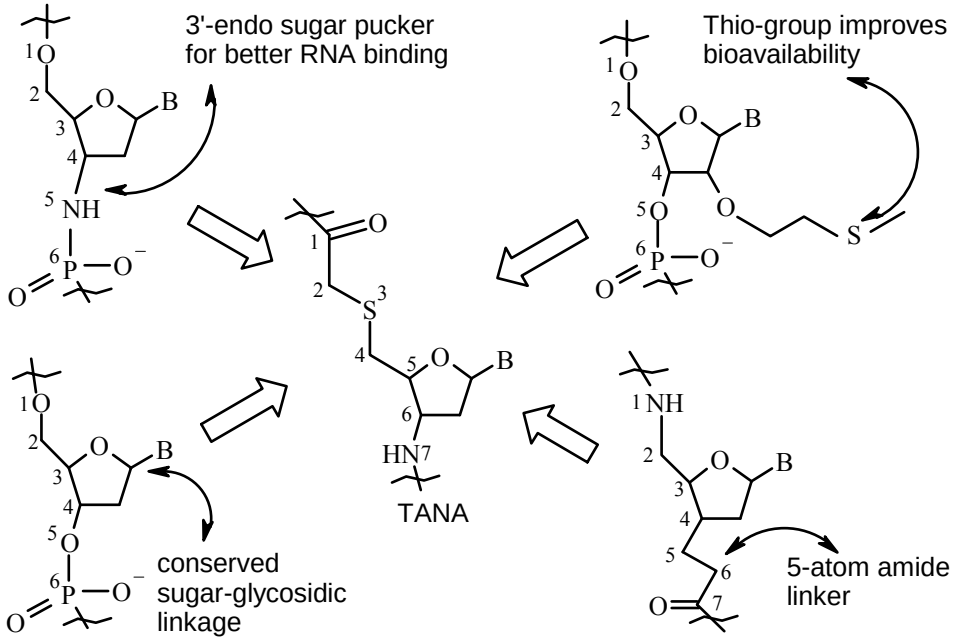


Figure 2. Design of thioacetamido nucleic acids(TANA)

It was demonstrated first by a high-resolution proton NMR study that the 3'-aminonucleoside 2'-deoxyfuranose conformation is predominantly C3'-endo, or N-type, for the free 3'-aminonucleosides[10] and also for 3'-aminonucleosides incorporated into the single and double stranded oligonucleotide N3'→P5' phosphoramidates.[11] In contrast, the 3'-hydroxy-2'-deoxynucleosides prefer C2'-endo, or S-type, conformations under

similar experimental conditions. Moreover, circular dichroism (CD) spectroscopy of the phosphoramidate duplexes has indicated a general RNA-like A-form of the duplexes being formed.[12] This change in the helix geometry is likely due to N-type of 3'-aminonucleoside sugar puckering. These findings were confirmed by the X-ray analysis of crystals of the self-complementary Dickerson- Drew dodecamer CGCGAATTCGCG containing only N3'→P5' internucleoside phosphoramidate linkages.[13] This phosphoramidate duplex adopts a RNA-like A-type double helix, despite being made from 2'-deoxynucleosides. Additionally, X-ray analysis revealed a more extensive hydration pattern of the oligonucleotide phosphoramidate sugar- phosphate backbone compared to that for phosphodiesters. The better hydration of the phosphoramidate complex is due to the presence of 3'-NH groups, which are capable of accepting and donating hydrogen bonds from water. Formation of the hydrogen bonds between 3'-amino groups and ammonium chloride salt ions, which were not observed for the cognate phosphodiester duplex, was also seen in the phosphoramidate crystal.[13] Spatial geometry of substituents at the 3'-nitrogen is more planar than tetrahedral, suggesting N→P electronic conjugation, resulting in increased rigidity of the sugar phosphate backbone.[13]

The presence of thio group in the oligonucleotide is known to improve the bioavailability of the ONs. The 2'-modified, 2'-*O*-[2-(methylthio) ethyl] or 2'-*O*-MTE, oligonucleotides (Figure 3, II & III) exhibited high binding affinity to target RNA and improved binding to human serum albumin compared to the 2'-*O*-MOE modified oligonucleotides.[14] Nuclease digestion of 2'-*O*-MTE oligonucleotides showed that they have limited resistance to exonuclease degradation. The crystal structure of a decamer DNA duplex containing the 2'-*O*-MTE modifcation provides insight into the improved RNA binding affinity, protein binding affinity and limited resistance of 2'-*O*-MTE modified oligonucleotides to exonuclease degradation.

The alkyl sulfide internucleoside (Figure 3, IV) linkage also was found to adequately replace the phosphodiester linkage in DNA oligomers to form stable complexes with both RNA and DNA but with reduced affinity.[15] A mixed ribo-deoxyribo dimer (Figure 3, IV) with alkylsulfide linkage in an ON showed marked selectivity for binding to RNA.[16]

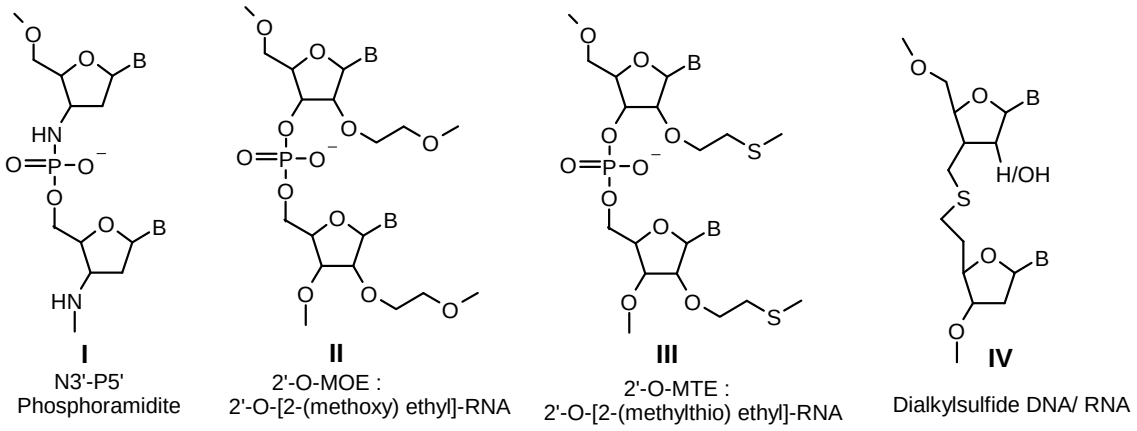


Figure 3. N3'-P5' phosphoramidite, 2'-O-MOE, 2'-O-MTE and dialkyl sulfide DNA/RNA

An additional atom in the five-atom amide linker that provided requisite flexibility was shown to lead to the stabilization of the duplexes with RNA (Figure 4, I).[17] The amide linkers in 2'-O-methyl substituted furanose sugar also led to the stabilization of RNA duplexes. The replacement of 3'-O with a methylene group and addition of a methyl group in either *R* or *S* configuration (Figure 4, I) also led to stabilization of RNA: DNA duplexes.[18]

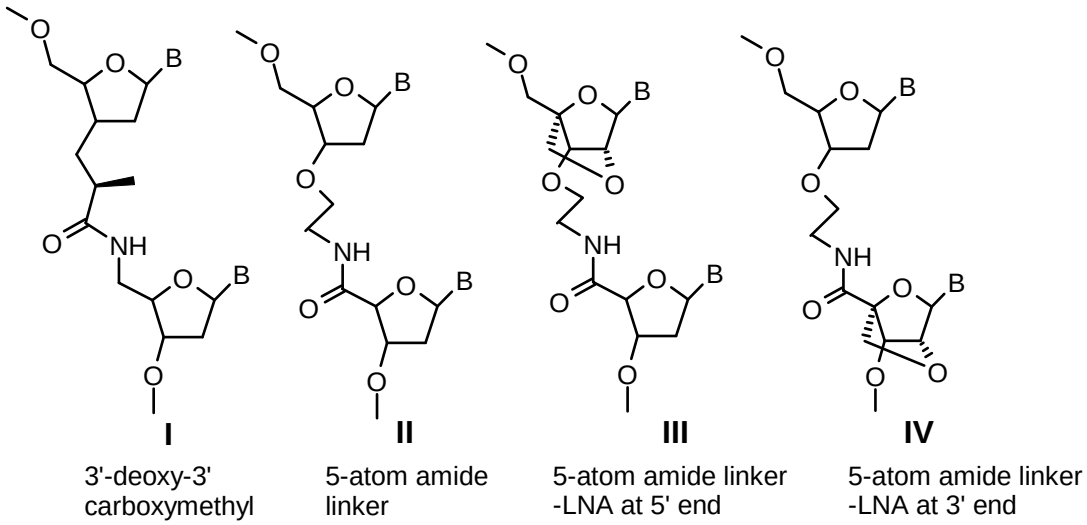


Figure 4. Replacement of phosphodiester by amide linkage

CD studies indicated increase in A-type of duplex structure in these cases. Thermodynamic results on duplex formation of DNA/RNA ONs with multiple amide bonds show less unfavorable entropic component relative to wild type DNA:RNA despite the presence of an additional atom in the backbone. The reduced conformational

flexibility of the amide bond leading to favorable pre-organization could be one reason for increased stability of the duplexes. X-ray structure of the duplex with RNA revealed a *trans* amide conformation and that the distance between the neighboring base pairs was not affected by the longer backbone. An alternate, slightly different five-atom linker was introduced in the dinucleoside. These chimeric amide-phosphate linkages resulted in destabilization of the duplexes (Figure 4, II).[19] Further, in order to improve the binding efficiency, an LNA nucleoside was introduced as one of the dinucleoside unit in the amide dimer. The LNA monomer improved binding to RNA, when positioned preferably towards the 3'-end of the oligomers (Figure 4, IV). Thus, the LNA monomer could replace the destabilizing effect despite the fact that the linker had an additional atom compared to the natural phosphodiester linkage. The 3'-substituents that could exert conformational restrictions on the ribose ring also showed effective stabilization with RNA targets when positioned towards 3'-end.

The chapter deals with chemical synthesis and biophysical studies of the homo TANA oligomers and TANA monomers incorporated at specific positions into *aeg*PNA backbone.

The specific objectives of this chapter are

i) Synthesis of TANA monomers (T, C, C^{Me}, A and G).

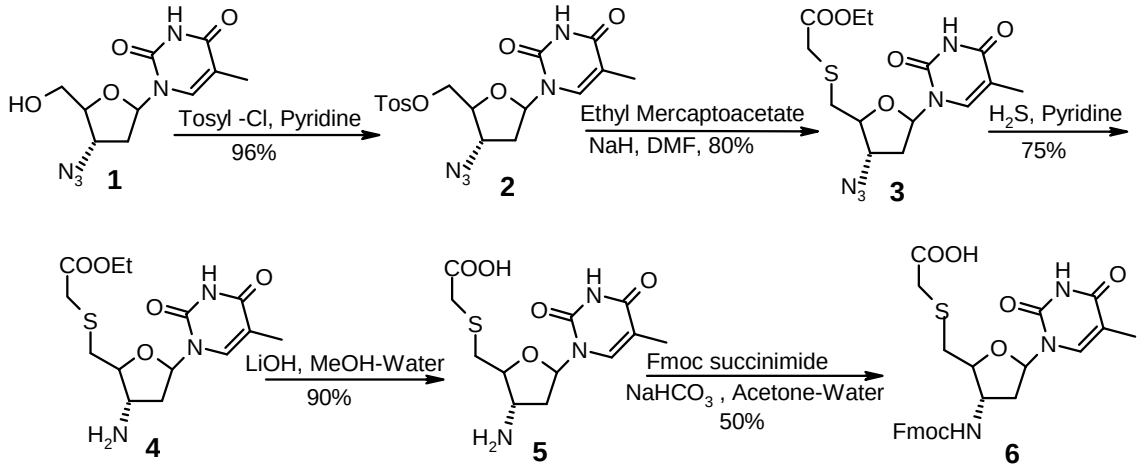
ii) Solid phase synthesis of TANA homo oligomers and oligomers incorporating *aeg*PNA monomers and TANA monomer.

iii) Cleavage from the solid support and purification, characterization of the oligomers.

iv) The binding studies of TANA oligomers and TANA-*aeg*PNA chimeric oligomers with DNA/RNA using UV and CD spectroscopy.

2.1.3 Synthesis of TANA monomers, Results and discussion

2.1.3.1 Synthesis of thyminyl TANA monomer



Scheme 1. Synthesis of thyminyl TANA monomer

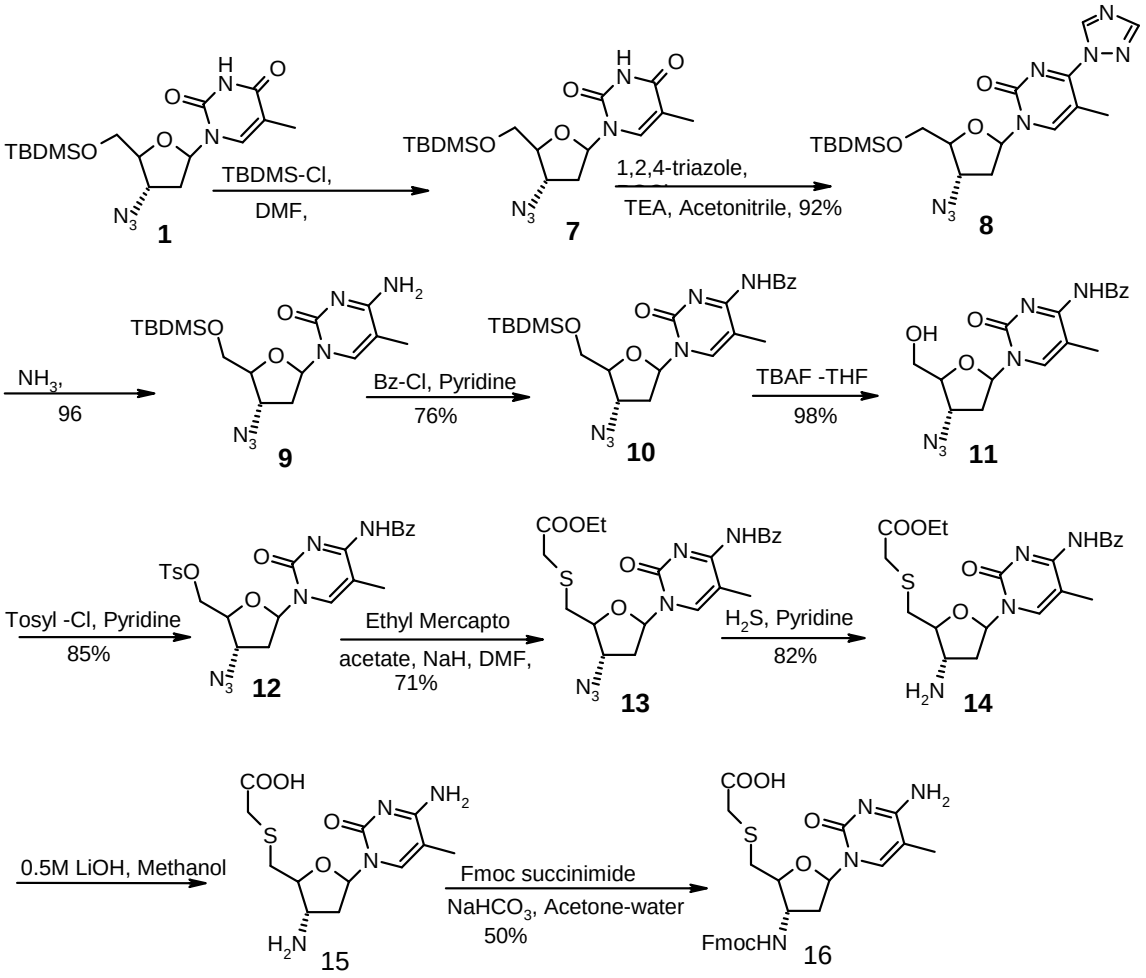
The monomer synthons 3'-Fmoc-amino-5'-thioacetic acid derivative of thymidine **6** was derived from the starting material 3'-azido thymidine(AZT) **1** (Scheme 1). AZT could be synthesized in the laboratory starting from thymidine by following reported procedures.[20] The 5'-OH of AZT **1** was converted to the tosyl derivative **2**, by treatment with tosyl-chloride in pyridine at room temperature. The formation of the product was confirmed by appearance of the peaks due to the tosyl methyl protons at δ at 2.4 ppm and aromatic protons at 7.8 ppm (dd) in the ^{1}H NMR spectrum. Nucleophilic displacement of the 5'O-tosylate in **2** by ethyl mercaptoacetate in presence of NaH generated 5'-ethyl mercapto ester derivative **3**. The formation of the product was confirmed by disappearance of the peaks due to tosyl and appearance of new peaks due to the mercapto ethyl ester group in ^{1}H NMR. Also, there is upfield in chemical shift of the of the 5', 5''-H from 4.2 ppm to 3.0 ppm due to the substitution of the 5'O by less electronegative S- atom. The azide group was converted to 3'-amino group in **4** using pyridine-H_2S in 75% yield.[21] The H_2S gas was generated by addition of dilute HCl to Na_2S and gas produced was passed through the reaction mixture. The reaction takes 45 minutes to 1 hour for completion. The compound **4** was isolated by removal of pyridine and column purification. The conversion of azide to amine was confirmed by disappearance of the azide peak at 2108 cm^{-1} in the IR spectrum. The other methods of azide reduction using catalytic hydrogenation failed due to poisoning of the catalyst due to sulfur. The reduction of the azide functionality with PPh_3 /H_2O method also did not give good yield of the required

product. The product **4** is not stable for longer time, so the ester function of **4** was hydrolyzed immediately to get the free sugar-amino acid **5**. The hydrolysis was done using 1:1 1M aqueous LiOH and methanol. The free amine group of **5** was then protected as Fmoc to get the monomer synthon **6**. The Fmoc protection was done using Fmoc- succinimide and $NaHCO_3$ as base in 1:1 acetone- water as the solvent system. The Fmoc protected monomer **6** was desiccated and used for solid-phase oligomers synthesis.

2.1.3.2 Synthesis of 5-methyl cytosinyl TANA monomer

The 5-methyl cytosine monomer **16** was synthesized starting from AZT **1** (Scheme 2). The 5'-OH of AZT **1** was protected as 5'-*O-tert.*butyldimethylsilyl ether to get **7.** The nucleobase was then transformed to 5-methyl cytosine derivative **9** via C^4-1,2,4- triazolo derivative followed by replacement of the triazolide with ammonia to get **9**.[21] The formation of C^4-1,2,4- triazolide proceeds by the *in situ formation* of tri (l H-1, 2, 4-triazol-l-yl) phosphoryl oxide by the reaction of phosphoryl chloride with 3 mol equiv. each of 1,2,4-triazole and triethylamine in acetonitrile solution.[22] The tris (phosphoryl**)**-triazolide in turn phosphorylates C^4-OH tautomer in presence of triethyl amine. The 1, 2, 4-triazole internally substitutes at C4-O-P (=O) $(1, 2, 4 \text{ triazole})_2$ to get compound **8**. The 1, 2, 4- triazolyl group in **8** readily undergo nucleophilic substitution at C-4 when it is treated with aqueous ammonia, ethanolic dimethylamine, or morpholine in dioxan solution at room temperature.[23]

N-benzoylation of the exocyclic amine of **9**, gave benzoylated product **10**, which was then 5'*O*- desilylated to get the compound **11**. Tosylation of the 5'OH gave **12**, followed by nucleophilic substitution gave the ethyl mercaptoester **13**. Reduction of the 3'-azide function by H_2S-Pyridine gave the primary amine **14.** The product **14** is not stable for longer time, so the ester function of **14** was immediately hydrolyzed to get the free sugar-amino acid **15**. It was found that during the basic hydrolysis of the ester group the exocyclic amino group was also deprotected. To examine the stability of the exocyclic amine protection, different molar concentration of LiOH were tested to standardize the optimum hydrolysis condition. But it was observed that at low molarity of LiOH 0.2M-0.4M, the reaction does not proceed or proceeds very slowly. The hydrolysis reaction was



Scheme 2: synthesis of 5-methyl cytosine TANA monomer

performed with 0.5M LiOH and it took 2 hour for complete consumption of the starting material **14**. ^{1}H NMR spectrum of the product showed that the change of molarity of LiOH from 0.2 M to 0.5 M was enough for deprotection the exocyclic benzoyl protecting group. The hydrolysis reaction was performed with the change of solvent from methanol to THF and dioxane, but the same result of deprotection of the exocyclic benzoyl group was observed. A possible explanation of the instability of the exocyclic benzoyl group against hydroxylic bases is depicted in Figure 5.

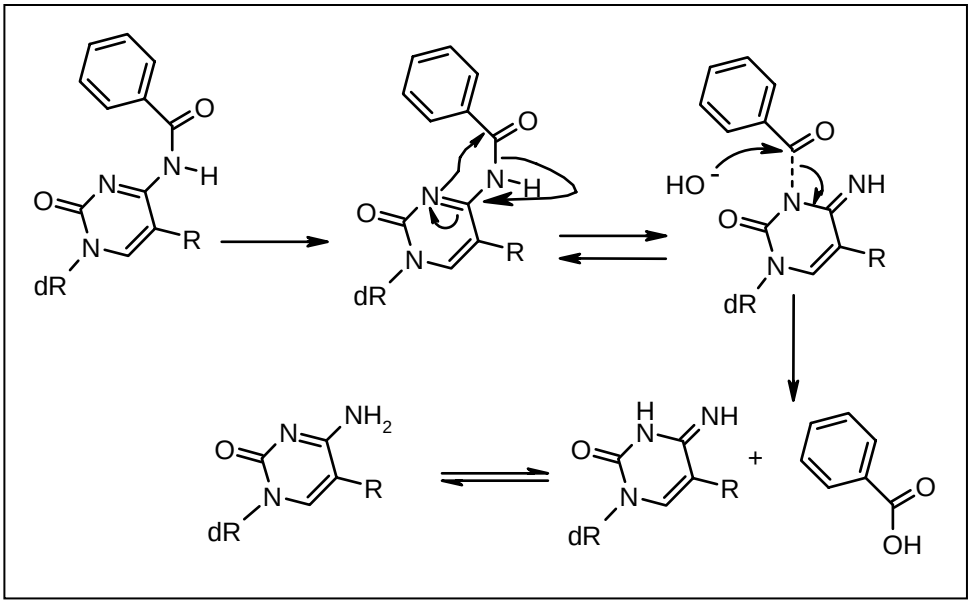
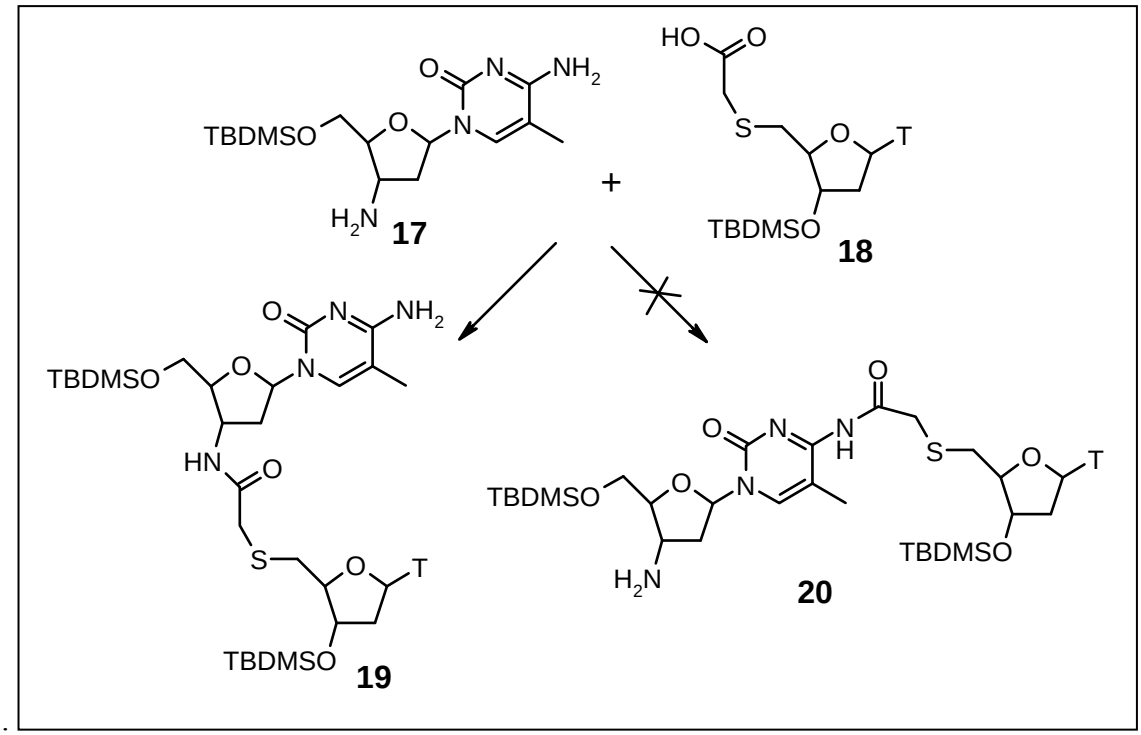


Figure 5. Mechanism of benzoyl deprotection by anchimeric assistance

Protection of 3'-amino group of **15** was then accomplished to get Fmoc-amino acid 5-methyl cytosine derivative **16**. Protection of 3'-amino group was confirmed by negative ninhydrin test and downfield shift of the 3'-H compared to the starting material **14.** The monomer **16** was desiccated used for solid phase oligomer synthesis.

2.1.3.3 Evaluation of reactivity of 3'-amino group vs. C4-amino group

The 5-methyl cytosine TANA monomer is without exocyclic amino protected group and in solid phase peptide coupling of the monomer **16** either the 3'-amino or the C4- amine may form amide bond with the corresponding acids. Therefore the reactivity of the 3'-NH_2 vs. the C4-NH_2 group was evaluated (Scheme 3) .This was done by solution phase coupling reaction between 5'-*O tert.* butyl dimethyl silyl 2', 3'-dideoxy 3'-amino 5-methyl cytidine **17** and 3'-*O tert.* butyl dimethyl silyl thymidine ethyl mercaptoactic acid **18**, using HBTU, HOBt, DIEA in acetonitrile. The NMR spectrum of the only product formed confirms exclusive formation of the product **19** by coupling at the 3'-amino group and not at the exocyclic amine group. ^{1}H NMR spectroscopy showed a downfield chemical shift of the 3'H after the coupling reaction.



Scheme 3a. Reactivity of 3'-amino group vs. C4-amino group of 5-methyl cytosine

Scheme 3b. Reactivity of 3'-amino group vs. C4-amino group of 5-methyl cytosine

The insufficient reactivity of the C4-exocyclic amino group was further confirmed by the coupling reaction between 5'-O-dimethoxy trityl 2', 3'-dideoxy 3'-azido 5-methyl cytidine **21** and 3'-*O tert.* butyl dimethyl silyl thymidine ethyl mercaptoactic acid **18.** The rate of the reaction was found to be very slow, only 5-10 % product formation was observed on TLC after 24h (Scheme 3b).

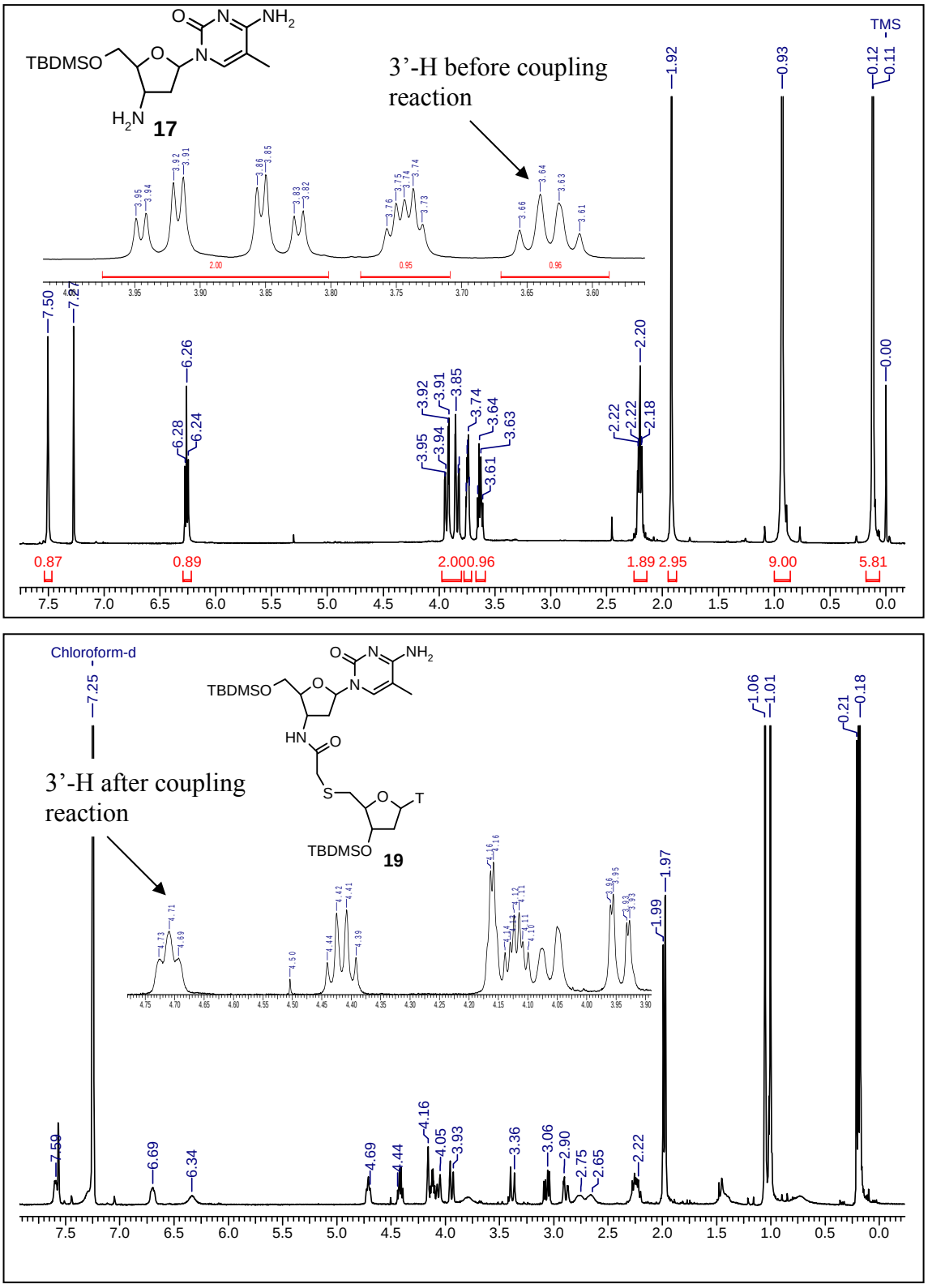
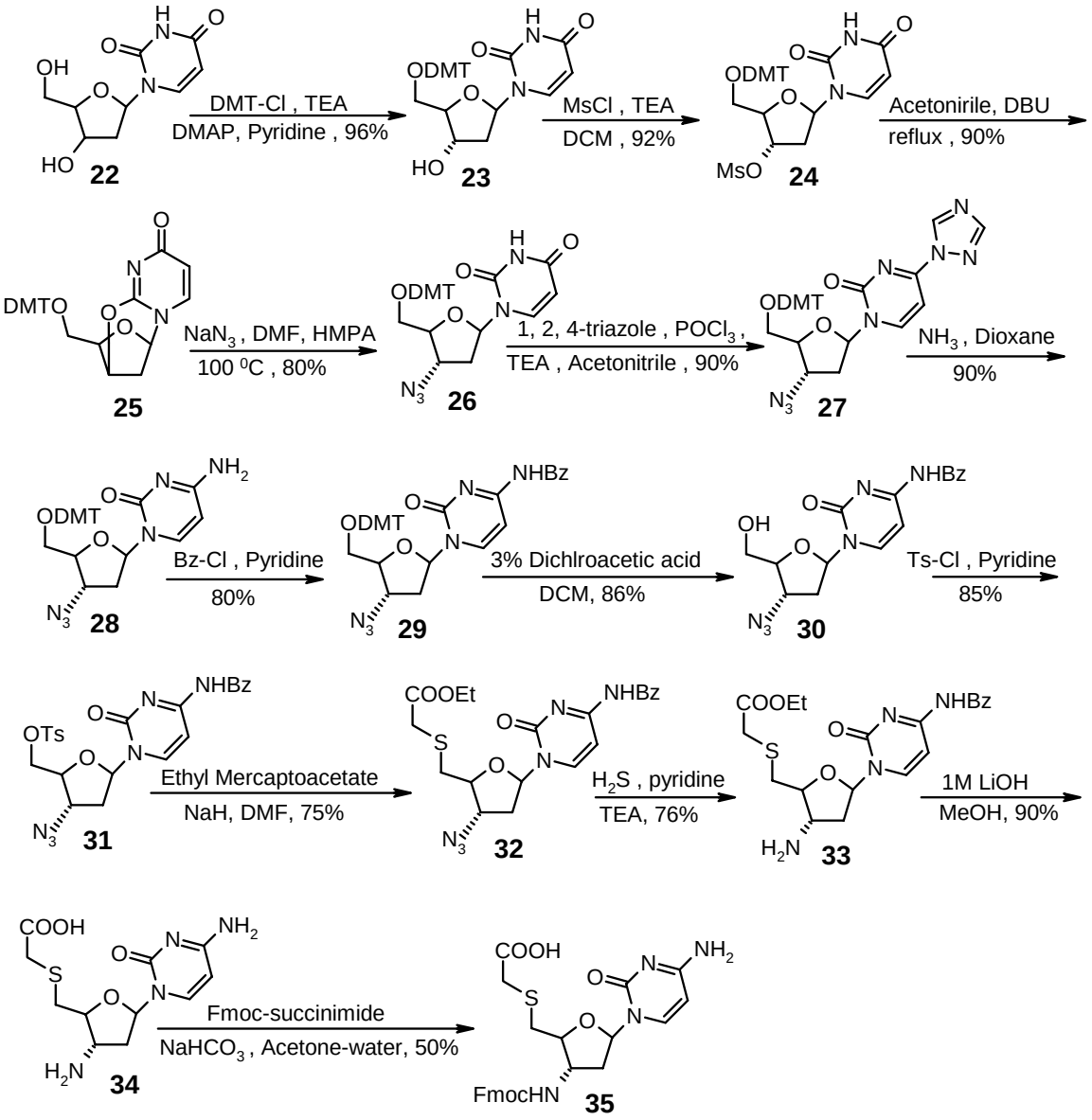


Figure 6. ^{1}H NMR chemical shift of 3'H proton in **17** and **19**

2.1.3.4 Synthesis of cytosinyl TANA monomer

The synthesis of cytosinyl TANA monomer starting from deoxy uridine is depicted in Scheme 4. The starting material used was 2'-deoxyuridine **22**. The 5'-OH in **22** was

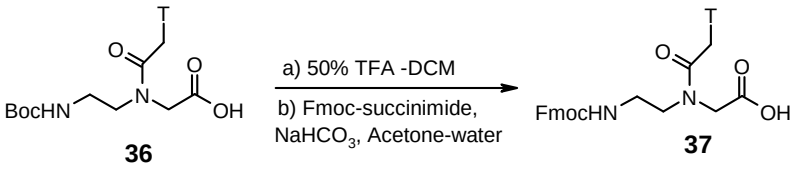


Scheme 4. Synthesis of cytosinyl TANA monomer.

protected as the dimethoxy trityl ether to give **23**. The 3'-OH was converted to mesylate **24**, which on reflux with DBU in acetonitrile gave 2', 3'-anhydronucleoside **25**. The anhydro compound was transformed to the 3'-azido derivative **26**, by treatment with NaN_3 in DMF and HMPA. The nucleobase uracil was then converted to cytosine derivative **28** via C^4-1, 2, 4- triazolo derivative **27** followed by replacement of the

triazolide with ammonia to get **28**.[21] N-benzoylation of the exocyclic amine of **28**, gave the benzoylated product **29**, which on 5'-*O*- detritylation to give the compound **30**. The detritylation was done using 3% dichloroacetic acid in DCM and triethylsilane was used as a scavenger for the trityl cations formed. Tosylation of the 5'OH of **30** gave tosylated derivative **31**, which on nucleophilic substitution using ethyl mercaptoacetate/NaH gave the ethyl mercaptoester **32**. The reduction of the 3'- azide by H_2S-Pyridine gave the amine **33.** The product **33** is not stable for longer time, so the ester function in **34** was immediately hydrolyzed to get the free sugar-amino acid **35**. In this case also the benzoyl group of the exocyclic amine was found to be unstable against basic hydrolysis. Finally protection of the 3'-amine to give the Fmoc protected cytosinyl TANA monomer **35**. The monomer **35** was desiccated and used for solid phase oligomer synthesis (Scheme 4).

2.1.3.5 Synthesis of Fmoc thyminyl *aeg* PNA monomer



Scheme 5. Synthesis of Fmoc *aeg* PNA monomer

The Boc protected *aeg*PNA monomer **36** was synthesized following literature procedure.[24] The Boc group was deprotected using 50% TFA-DCM and amine formed was protected as Fmoc to give Fmoc protected thyminyl *aeg*PNA monomer building block **37** (Scheme 5).

2.1.4 Solid Phase Synthesis of TANA, *aeg*PNA and *aeg*PNA - TANA chimeric oligomers

2.1.4.1 General principle of Solid Phase Peptide Synthesis (SPPS)

In contrast to the solution phase method, the solid phase method devised by Merrifield, [25] offer great advantages. In this method, the *C*-terminal amino acid is linked to an insoluble matrix such as polystyrene beads having reactive functional groups, which also acts as a permanent protection for the carboxylic acid (Figure 7). The next N^{α}-protected amino acid is coupled to the resin bound amino acid either by using an active pentaflurophenyl

(pfp) or 3-hydroxy-2, 3-dihydro-4-oxo-benzotriazole (Dhbt) ester or by an *in situ* activation with carbodiimide reagents or other coupling reagents. The excess amino acid is washed out and the deprotection and coupling reactions are repeated until the desired peptide is achieved. The need of purify the intermediates at every step is obviated. Finally, the resin bound peptide and the side chain protecting groups are cleaved in one step. The advantage of the solid phase synthesis are (i) all the reactions are performed in a single vessel minimizing the loss due to transfer, (ii) large excess of monomer carboxylic acid component can be used to achieve in high coupling efficiency, (iii) excess reagents can be removed by simple filtration and washing steps and (iv) the method is amenable to automation.

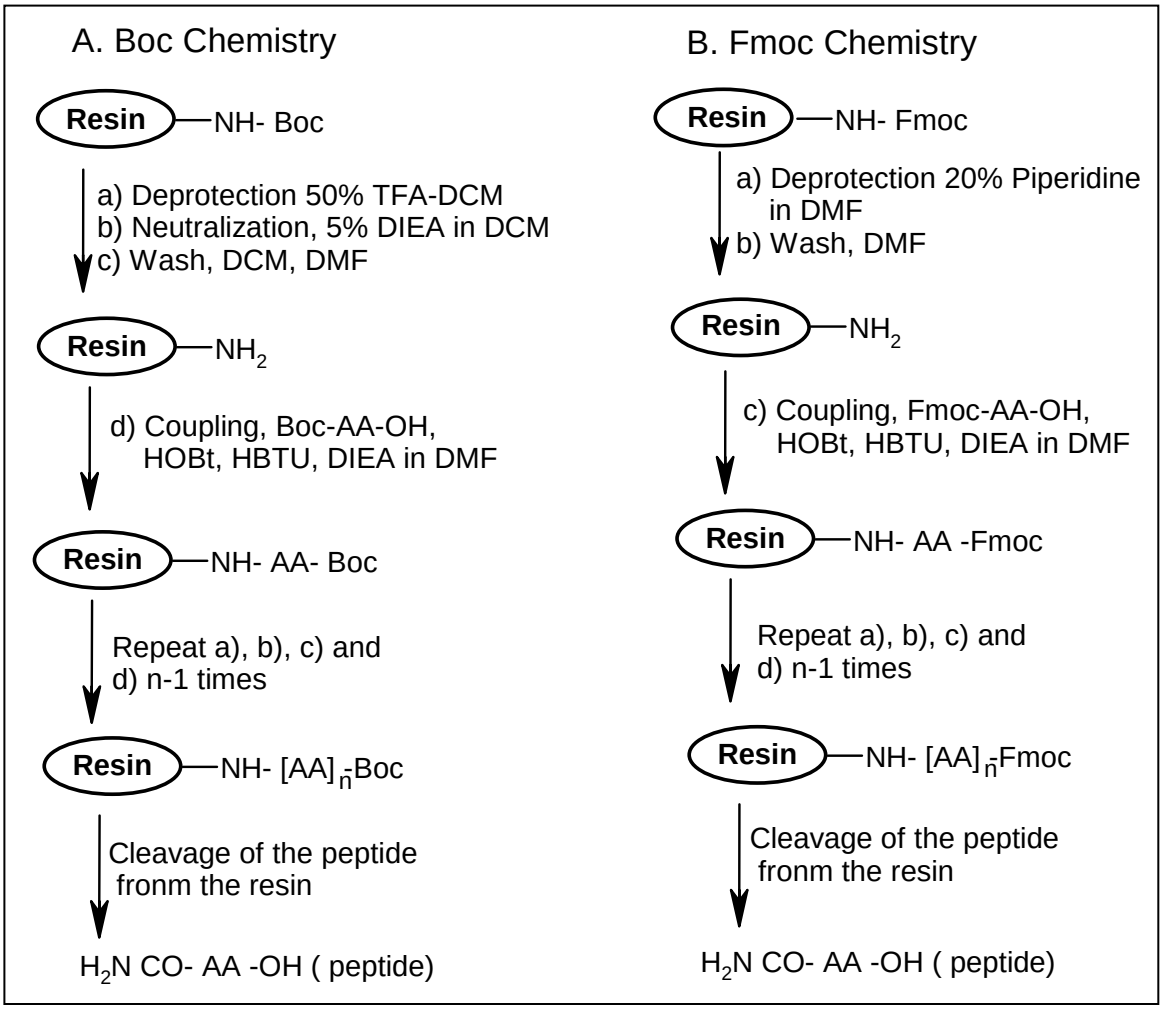


Figure 7. Schematic representation of SSPS **A**) Boc-chemistry and **B**) Fmoc-chemistry

Currently, two chemistries are available for the routine synthesis of peptides by solid phase method (Figure 7). When R. B. Merrifield invented SPPS in 1963, it was according to the t-Boc method.This method involves the use of *t*-butyloxycarbonyl (*t*- Boc) group as N^{α}-protection that could be removed under acidic conditions such as 50% TFA in DCM.[26] The reactive side chains of amino acids were protected with groups stable to *t*-Boc deprotection conditions (such as Cbz, benzyl etc.) and removable under strongly acidic conditions using HF in dimethyl sulfide or TFMSA in TFA. Alternatively a base labile protecting group strategy could be used involving fluorenylmethyloxycarbonyl (Fmoc) group for N^{α}- protection, which is stable to acidic conditions but can be deprotected efficiently with a secondary base such as piperidine. This method was introduced by Carpino in 1972.[27] Coupled with mild acid sensitive side chain protection, this method offers a second strategy for SPPS. In both chemistries, the linker group that joins the peptide to the resin is chosen such that the side chain protecting groups and the linker are cleaved in one step at the end of the peptide synthesis.

2.1.4.2 Resins for solid phase peptide synthesis

Merrifield resin:

Merrifield resin (Figure 8) has been used for more than thirty years for the synthesis of peptides and are the standard supports for the SPPS of small or medium sized peptides using the Boc strategy. The resin consists of of divinylbenzene cross-linked polystyrenes beads that have been functionalized with chloromethyl groups. Attachment of Boc amino acids is generally achieved by heating the resin, in DMF, with the appropriate carboxylic acid cesium salt in the presence of KI or dibenzo-18-crown-6. Release of peptide or carboxylic acid is normally effected by treatment of the resin with HF or TFMSA.

Trityl resins:

The advantage of trityl based resin (Figure 8) is that any nucleophilic functionality can be linked to these resins under extremely mild conditions. Most of the side-chains of the trifunctional amino acids can be attached to the 2-chlorotrityl resin. Furthermore, symmetrical bifunctional compounds, such as diamines, diols, diphenols and diacids, are in effect monoprotected by this process, allowing one end of the molecule to be

selectively modified. Cleavage of products from the supports takes place under very mild conditions, owing to the high stability of the trityl cations. Furthermore, trityl cations are extremely poor electrophiles and as such do not undergo alkylation side-reactions.[28]

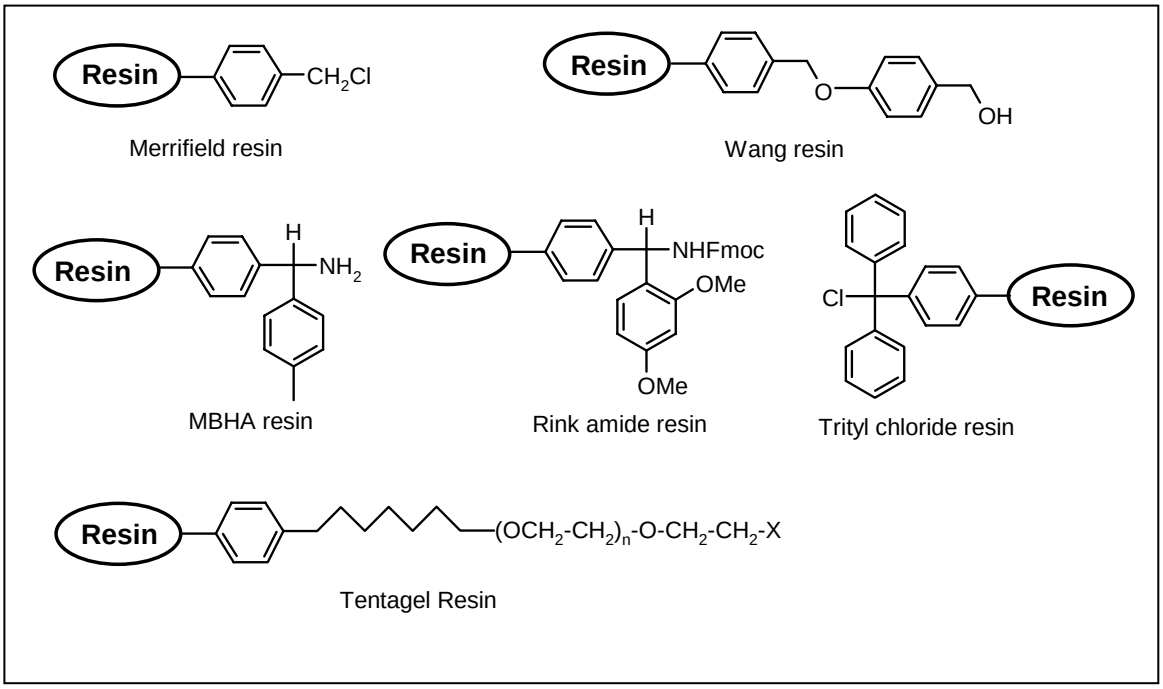


Figure 8. Resins commonly used for solid phase peptide synthesis

MBHA resin:

MBHA resin (Figure 8) is based on 100-200 mesh, 1% DVP cross-linked polystyrene functionalized with 4-methylphenyl-aminomethyl group. These resins are the most widely used resins for the synthesis of peptides by the Boc strategy.[29] Cleavage of the peptides from the resin can be achieved by treatment with HF, TFMSA or HBF_4. In Fomc SPPS, this resin has been used in conjunction with a low-high acid cleavage, where side-chain protecting groups were first removed prior to cleavage from the resin with TFMSA. The excellent swelling properties of this resin make it an ideal matrix onto which to attach TFA-labile linkers for combinatorial synthesis.

Rink amide resins:

Rink amide resin (Figure 8) is highly acid sensitive because of the benzhydrylamine linker is joined to the support through a benzylic ether bond. Lacking any amide groups, Rink amide resin, has a broader range of chemical compatibility than other rink Amide

resins. Breakdown of the linker during cleavage of the with high concentration of TFAs can , however, occur, leading to the formation of highly colored by-products, this problem can be minimized through the use of low concentrations of TFA, or by the addition of trialkylsilanes to the cleavage mixture.[30]

Wang resin:

Wang resin (Figure 8) is the standard support for Fmoc SPPS by the batch-wise method. The resins consist of chloromethyl polystyrene modified with 4-hydroxybenzyl alcohol. Attachment of amino acids to hydroxy functionalized supports is normally achieved by DMAP catalyzed esterification with the appropriate symmetrical anhydride, and by using MSNT/ MeIm or 2, 6-dichlorobenzoyl chloride activation. Wang resin has very good physical properties and is well suited to SPOS. This support is extremely useful in SPOS for the immobilization of carboxylic acids, alcohols and phenols. Release of these functionalities is usually effected by mild acidolysis with TFA.

2.1.4.3 Coupling reagents

Carbodiimides have been some of the most popular *in situ* activating reagent in peptide synthesis. Dicyclohexylcarbodiimide (DCC), first described in the 1950s, remains the most popular and is a particularly appropriate choice for the apolar environment of polystyrene resins.[31] The principal limitation in using carbodiimides is the dehydration of Asn and Gln residues. The addition of HOBt to the reaction mixture will prevent dehydration and has the added benefit of acting as catalyst. The principal drawback to the use of DCC is the formation of the DCM-insoluble dicyclohexylurea (DCU) during activation and acylation. Other carbodiimides used in SPPS such as diisopropylcarbodiimide (DIPCDI), t-butyl-methylcarbodiimide and t-butylethylcarbodiimide form ureas which are more soluble in DCM than DCU.

In situ activating agents are widely accepted because they are easy to use; they give fast reactions, even between sterically hindered amino acids, and their use in generally free of side reactions. Most are based on phosphonium or uranium salts which in presence of a tertiary base, can smoothly convert protected amino acids to a variety of activated

species. The most commonly employed BOP, PyBOP, HBTUand TBTU, generate OBt ester, and these have wide application in routine SPPS and solution synthesis for difficult couplings (Figure 9). BOP should be handled with great care as the by- product formed during activation is highly carcinogenic. BOP can be substituted by PyBOP without loss of performance.

Coupling reagents are also available which generate esters that are more reactive than OBt. Two such reagents are HATU [32] and HCTU [33], which in presence of base converts carboxylic acids to corresponding OAt and 6-ClOBt esters respectively. Such esters are more reactive than their OBt counterparts owing to the lower pKa of HOAt and HOCt compared to HOBt. Furthermore, HOAt has the added advantage of the pyridine nitrogen which provides anchimeric assistance to the coupling reaction.[32] HATU is generally

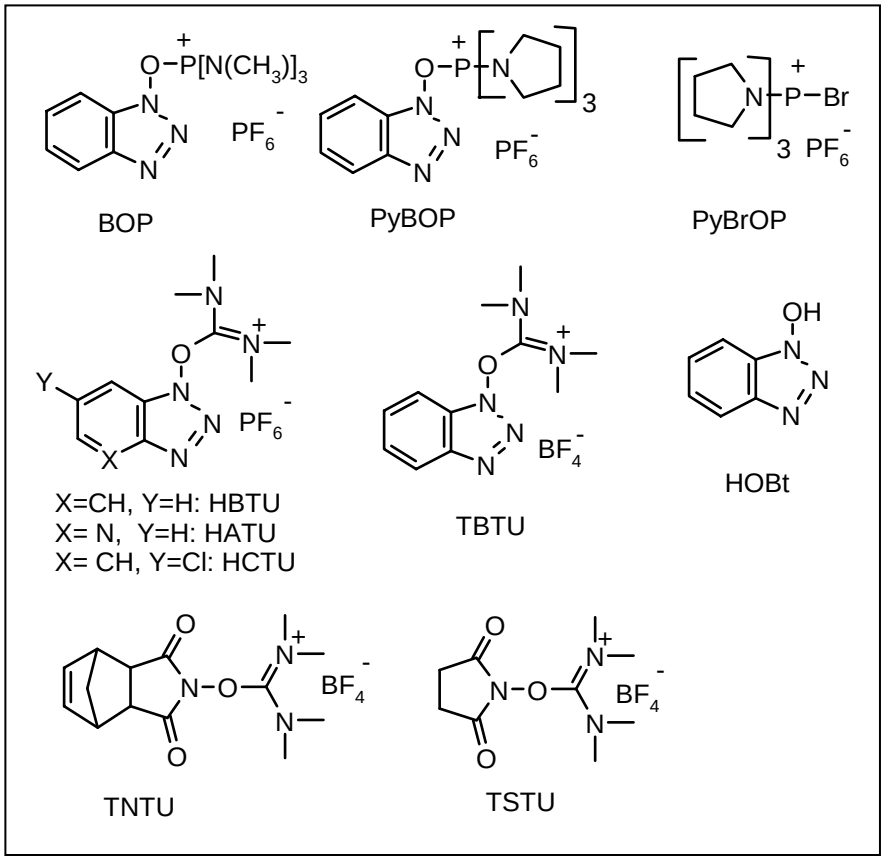


Figure 9. Coupling reagents used for Peptide synthesis

regarded as the most efficient coupling reagent. However, owing to its cost, it use is generally recommended for difficult and hindered couplings and the synthesis of long peptides.

The coupling of N-methyl amino acids can be difficult and slow process. PyBOP[34] allows efficient, fast and enantimerization- free couplings, providing a significant improvement over existing reagents.Coupling efficiencies and rates have been investigated for several esters and reagents.[35] The following order was found:

BOP/HOBt> DIPCDI/HOBt> DIPCDI/ODhbt>DIPCDI/OPfp.

2.1.4.4 Resin Tests

The most widely used qualitative test for the presence or absence of free amino groups (deprotection/coupling) was devised by Kaiser.[36] The test is simple and quick; however, it should be noted that some deprotected amino acids do not show the expected dark blue color typical of free primary amino groups(e.g., serine, aspargine, aspartic acid) and that proline, being a secondary amino acid, does not yield a positive reaction. Furthermore, occasionally false negative tests are observed, particularly with strongly aggregated sequences, which show that one has to be very careful when assessing the extent of deprotection or coupling on solid phase. Other methods such as picric acid monitoring [37] or mass spectrometry[38] are also available. For iminoacids like proline, the chloranil test [39] is recommended. Recently a new method for detecting resin-bound primary and secondary amines has been described.[40] Resins are treated with 2, 3-dichloro-5-nitro-1, 4-naphthoquinone in DMF or DCM in the presence of 2, 6- di –*tert.* butylpyridine.

2.1.4.5 Synthesis of TANA and chimeric TANA-PNA Oligomers

Solid phase peptide synthesis protocols can be easily applied to the synthesis of oligomeric TANAs from the *C*-terminus to *N*-terminus. Since TANA monomer building block consists of sugar glycosidic bond which is not compatible against the Boc deprotection condition. The TANA oligomers were synthesized by Fmoc strategy using rink amide resin as the solid support. Commercially available rink amide resin having the loading value 0.5-0.7 mmol/g was reduced to 0.25 mmol/g by capping with acetic anhydride in 5% DIPEA-DCM solution and then the resin was functionalized with Fmoc-

β-alanine or Fmoc-Lys(Boc)-OH, by *in situ* activation with HBTU/HOBt /DIEA in DMF. The loading value was estimated by picric acid assay [41] and it was found to be ~

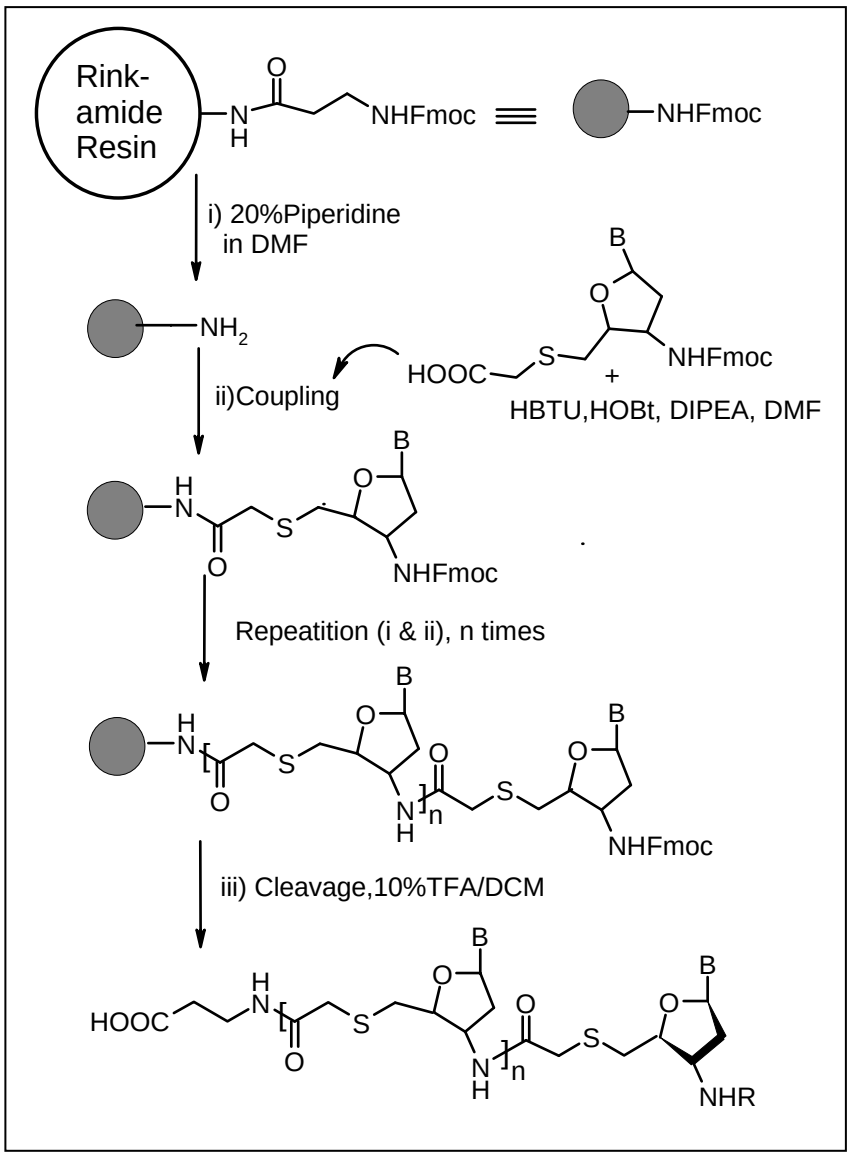


Figure 10. Solid-phase synthesis of TANA oligomers

0.25 mmol/g. The TANA oligomers were synthesized using repetitive cycles (Figure 10), each comprising the following steps:

(i) Deprotection of the *N-Fmoc* group using 20% piperidine in DMF.

(ii) Coupling of the free amine with the free carboxylic acid group of the incoming monomer (3 to 4 equivalents). The coupling reaction was carried out in DMF with HBTU

as coupling reagent in the presence of DIEA and HOBt as catalysts. The deprotection of the *N-Fmoc* protecting group and the coupling reactions were monitored by Kaiser's test.[36] The *Fmoc*-deprotection step leads to a positive Kaiser's test, where in the resin beads as well as the solution are blue in color. On the other hand, upon completion of the coupling reaction, the Kaiser's test is negative, the resin beads remaining colorless.

(iii) Capping of the unreacted amino groups using Ac_2O/5% DIEA in CH_2Cl_2 in case coupling does not go to completion. Figure 10 represents a typical Solid phase peptide synthesis of the TANA oligomers.

Various TANA oligomers were synthesized for the present study of DNA/RNA recognition are shown in Table1. The unmodified *aeg*PNA homooligomer T8, **43** (Table1, entry 6, *aeg*PNA) was synthesized using *Fmoc*- protected *aeg*PNA monomer **37**. This was used as the control sequence for comparing the hybridization and discrimination properties of TANA oligomers. The thyminyl TANA monomer **6** was incorporated into PNA oligomers using *Fmoc* chemistry on β-alanine derivatized rink amide resin to give the chimeric oligomers **40**, **41** and **43**. Various homo thymine TANA oligomeric sequences were synthesized **38**, **39**, **44** and **45** for study of DNA/RNA recognition.

Table1. TANA and TANA-aeg PNA chimeric sequences, their HPLC and mass characterization

Entry	TANA Sequence	HPLC t_R(Min)	Mass	
			Calculated	Observed
1	H-**t t t t t t t t**-β ala **38**	12.7	2467.6	2491.0 (M+ Na^+)
2	H-**t t t t c^m t t t**-β ala **39**	12.5	2466.7	2487.9 (M+ Na^+)
3	H-t t t t t t **t**-β ala **40**	10.0	2249.4	2251.7 (M+ 2)$^+$
4	H-t t t t **t** t t t-β ala **41**	8.3	2249.4	2252.6 (M+ 3)$^+$
5	H-t **t** t **t** t t **t** -β ala **42**	10.0	2343.0	2367.9 (M+ Na)$^+$
6	H-t t t t t t t-β ala **43**	7.6	2218.2	2218.9 (M^+)
7	H-**t t c t t t c t c t**- Lys **44**	13.7	3074.5	3074.8 (M^+)
8	H-**t t c^m t t t c^m t c^m t**-Lys **45**	12.3	3116.6	3117.1 (M^+)

t, c, c^m TANA backbone**,** t *aeg*PNA backbone

Polypyrimidine (thymine, cytosine and 5-methyl cytosine) octamers and decamers as mentioned above are synthesized to study the effect of sugar-backbone modification on their duplex and triplex-forming ability. The classical homopyrimidine *aeg*PNAs are known to complex with the complementary oligonucleotide DNA/RNA in a 2:1 PNA:

DNA/RNA stoichiometry.[42] In order to study the duplex formation potential and in particular DNA/RNA discrimination of the sugar-phosphate backbone modified TANA monomeric units, it was imperative to synthesize mixed sequences. For this purpose the homo mixed sequences **39, 44** and **45** were synthesized (Table 1).

2.1.4.6 Cleavage of the TANA and TANA-PNA chimeric ONs from the Solid Support

The oligomers were cleaved from the solid support, using 10% Trifluoroacetic acid in DCM and triisopropyl silane as the scavenger. A cleavage time of 0.5 h at room temperature was found to be optimum. The side chain protecting groups (the Boc protection in case of the sequences **44** and **45**) were also cleaved during this cleavage process. After cleavage reaction, the TFA-DCM mixture was evaporated by passing nitrogen gas and the oligomer was precipitated by addition of dry diethyl ether.

2.1.4.7 Purification of the TANA Oligomers

All the cleaved oligomers were subjected to initial gel filtration. The purity of the oligomers was checked by analytical RP-HPLC (C18 column, CH_3CN-H_2O system), which shows more than 65-70% purity. These were subsequently purified by reverse phase HPLC on a semi preparative 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.

2.1.4.8 Synthesis of Complementary Oligonucleotides

The DNA oligonucleotides **38**-**48** (Table 10) were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard β-cyanoethyl

Table 2. DNA/ RNA Oligonucleotides used in the present work

Entry	Sequence (5' to 3')	Type(corresponding TANA-PNA)
1	GCAAAAAAAACG **46**	Complementary for **38**, **40**, **41**, **42** and **43** Mismatch for **39**
2	GCAAAAGAAAACG **47**	Complementary for **39** Mismatch for **38**, **40**, **41**, **42** and **43**

3	r(GCAAAAAAAACG) **48**	Complementary for **38**, **40**, **41**, **42** and **43** Mismatch for **39**
4	r(GCAAAAGAAACG) **49**	Complementary for **39** Mismatch for **38**, **40**, **41**, **42** and **43**
5	TTTTTTTT **50**	Control DNA sequence
6	TCTCTTTCTT **51**	Control DNA sequence
7	AAG AAA GAGA **52**	Complementary for **44** and **45**
8	r(AAG AAA GAGA) **53**	Complementary for **44** and **45**

phosphoramidite chemistry.[43] The oligomers were synthesized in the 3' to 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 to be more than 98% and was used without further purification in the biophysical studies of TANAs. The RNA oligonucleotides (Table 2) were obtained commercially.

2.1.5 Biophysical Studies of TANA: DNA/RNA Complexes

To investigate the binding selectivity, specificity and discrimination of sugar phosphate backbone modified TANA oligomers towards complementary DNA and RNA, the stoichiometry of the TANA:DNA/RNA was first determined using Job's method.[44] The UV-melting studies were then carried out with all the synthesized oligomers and the *T*m data was compared with the control *aeg*PNA T8 and control DNA T8. The CD spectra of single strands and corresponding complexes with complementary DNA/RNA were recorded.

2.1.5.1 Binding Stoichiometry: UV-mixing Curves

Ultraviolet (UV) absorption and circular dichroism (CD) measurements are extremely useful to determine the stoichiometry and stability of duplexes and triplexes. The stoichiometry of the paired strands may be obtained from the mixing curves, in which the optical property at a given wavelength is plotted as a function of the mole fraction of each strand.

Various stoichiometric mixtures of TANA **38** with DNA **46** and RNA **48** were made with relative molar ratios of (TANA **38**:DNA **46** /RNA **48**) strands of 0:100, 10:90,

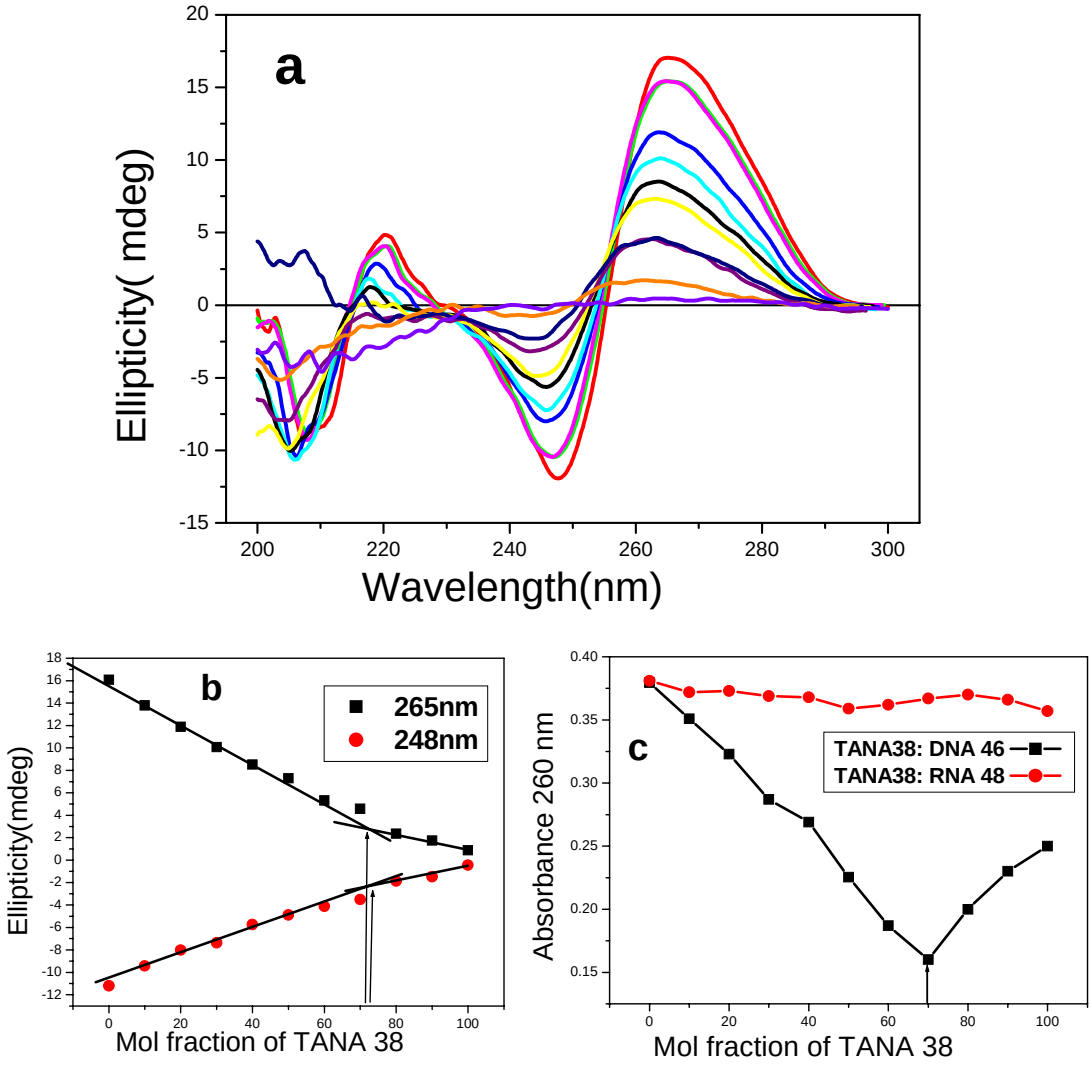


Figure 11. a) CD-curves for TANA **38** and the complementary RNA **48** r(CGCA$_8$CGC) mixtures in the molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 100:0 (Buffer, 10 mM Sodium phosphate pH 7.0, 100 mM NaCl, 0.1 mM EDTA), b) CD-Job plot corresponding to 265 and 248 nm, c) UV-job plot corresponding to above mixtures of different molar ratios, absorbance recorded at 260 nm.

20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 100:0, all at the same strand concentration 2 µM in sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA (10 mM, pH 7.0). The samples with the individual strands were annealed and the CD spectra were

recorded (Figure 11a). The CD spectra at different molar ratios show a isodichroic point (wavelength of equal CD magnitude) around 265 nm. CD values around such wavelengths are therefore especially useful for plotting mixing curves. This systematic change in the CD spectral features upon variable stoichiometric mixing of TANA and RNA components can be used to generate Job's plot which was indicated a binding ratio of 2:1 for the $TANA_2$: RNA complexes, expected for triplex formation. The mixing curve in Figure 11b was plotted using the CD data at 265 and 248 nm from the spectra in Figure 11a, which indicated the formation of triplex of TANA**38**: RNA **48**.

The mixing curves were plotted, absorbance of the mixtures of TANA **38** and DNA **46** and RNA **48** at fixed wavelength (λmax 260 nm) against mole fractions of TANA **38**. There is a drastic shift in the λmax value of the TANA **38**- RNA **48** mixture when the concentration of TANA **38** in the mixtures increased to 60 to 70%, which further increases behind those proportions. This supports the formation of (TANA **38**)$_2$: RNA **48** triplexes. In case of TANA **38** - DNA **48** mixtures there was no significant change of UV absorbance was observed which rule out the complex formation between TANA and DNA.

2.1.5.2 UV-Melting studies of TANA and TANA-PNA chimeric ONs -DNA/RNA complexes. ***Thermal Stability of triplexes***

To establish the binding stoichiometry of the TANA oligomers (**38** and **39)** Job's plot study was undertaken with complementary DNA/RNA sequences (Figure 11). The binding stoichiometry could be clearly established to be 2:1 between sequence **38** and RNA **48** using both CD and UV spectroscopy (Figure 11). The TANA:DNA binding could not be observed in the UV Job's plot (Figure 11c). All the binding studies of the sequences listed in Table 2 with complementary DNA/RNA were carried out in 2:1 stoichiometry using UV-T_m experiments. The complexes of TANA **38** with DNA (3'-GCAAAAAAAACG-5') **46** and RNA r (3'-GCAAAAAAAACG-**5') 48** were used to study their comparative binding efficiency. The TANA **39** has a 5-methyl cytosine unit and its binding with complementary DNA (3'-GCAAAAGAAACG-5') **47** and RNA sequence r(3'-GCAAAAGAAACG-5') **49** was studied. For single nucleotide mismatch studies the UV-*T*m measurements were carried out using sequences **38:49** and **39:48**.

The results are summarized in Table 3. For control, the T_m of complexes of DNA 3' TTTTTTTT 5' **50** and *aeg*PNA **43** with DNA **46** and RNA **48** were measured under

Table 3. UV-T_m values in ^oC of $(TANA)_2$:DNA/RNA triplexes

Entry	Sequence	TANA: DNA	TANA:RNA	TANA: Mismatch RNA
1	H-**t t t t t t t t**-β ala **38**	nd[a]	63.8 (63.5)[b]	49.8 (-14)[c]
		38: 46	**38: 48**	**38:49**
2	H-**t t t t c t t t**-β ala **39**	nd[a]	63.1 (67.1)[b]	52.9(-10.2)[c]
		39: 47	**39:49**	**39:48**
3	3'T T T T T T T T-5' **50**	nd[a]	20	nd
		50: 46	**50:48**	**50:49**

[a] nd not detected, [b] values in the parenthesis denote the melting temperature at pH= 5.5, [c] values in the parenthesis denote the decrease of melting temperature due to single base mismatch hybridization

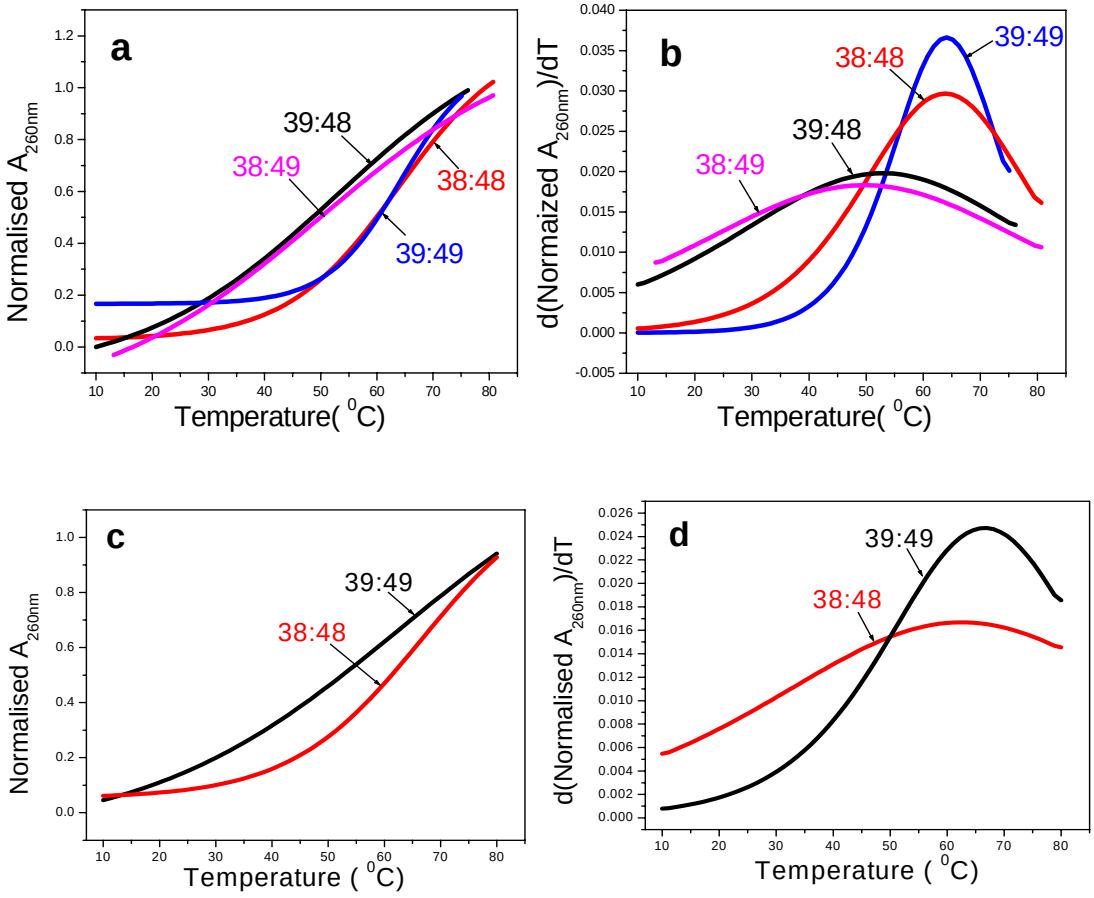


Figure 12. a.UV-melting Curves of **38** and **39** with RNA **48** and **49 (38:48, 39:49, 38:49** and **39:48) b.** Corresponding first derivative Curves**. c.** UV-melting curves of **38** and **39** with RNA **48** and **49** at pH= 5.5 **d.** corresponding first derivative curves.

identical buffer conditions. It is evident from the above experiments that the homopyrimidine TANA backbone oligomers (**38** and **39**) bind to the homopurine complementary RNA sequences (**48** and **49**, respectively) and show a single base-mismatch discrimination of about 10-14^{o} C. The thermal stability of **39:49** was increased by 4^{o}C at pH 5.5 in which a C^{+}GC base triple is formed.[45] The complex of TANA oligomer **38** with cRNA **48** was highly stable compared to the DNA:RNA complex (**50**:**49**) and the stability was comparable with that of *aeg*PNA_2:RNA (**43**:**48**). The formation of TANA: complementary DNA was not observed.

Table 4. UV-T_m values in ^{o}C of $(TANA\text{-}PNA)_2$:DNA/RNA triplexes

Entry	Sequence	DNA **46**	RNA **48**	RNA **49**
1	H-t t t t t t t-β ala **43**	42.6	61.5	52.5
2	H-t t t t t t **t**-β ala **40**	nd	53.5	-
3	H-t t t t **t** t t t-β ala **41**	nd	32.4	-
4	H-t **t** t **t** t t **t** -β ala **42**	nd	25.5	-
5	3'T T T T T T T T-5' **50**	nd	20	nd

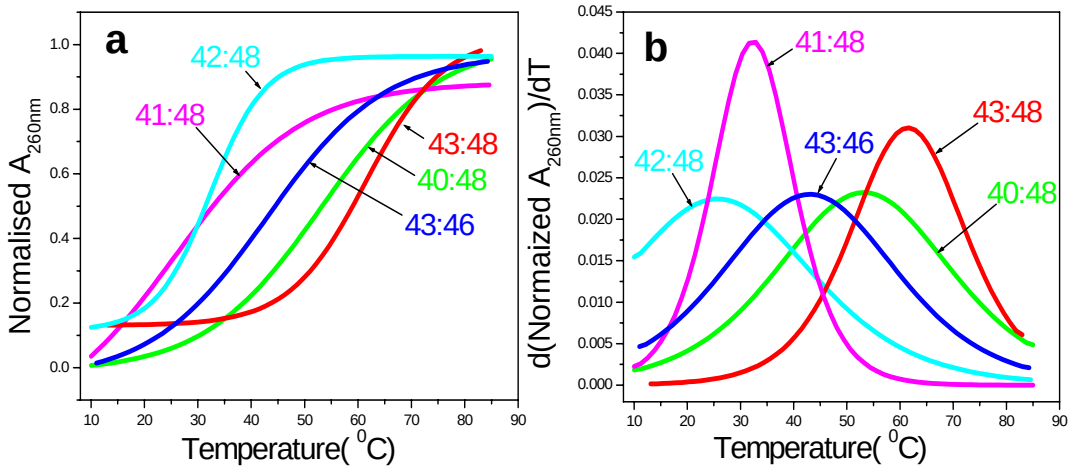


Figure 12. a.UV-melting Curves of **40**, **41**, **42** and **43** with complementary RNA **48** and **43: 48 b.** Corresponding first derivative Curves**.**

To test the compatibility of the TANA backbone in *aeg*PNA:TANA mix-backbone oligomers, the sequences **40-42** were synthesized. Sequence **40** has a TANA monomer unit at C-terminus, sequence **41** has the modified unit at the central position and sequence **42** is an alternate *aeg*-TANA sequence. The formation of complex of TANA-*aeg*PNA chimeric ONs **40-42** with complementary DNA **46** was not observed. The

homopyrimidine TANA-aegPNA chimeric ONs backbone oligomers **40**-**42** bind to the homopurine complementary RNA sequence **48** but all the complexes of oligomers **40**-**42** with RNA **48** were destabilized compared to the control *aeg*PNA:RNA complex **43**:**48**. The results are summarized in Table 4.

2.1.5.3 *Thermal Stability of duplexes*

To study the effect of thioacetamido nucleic acids designs on corresponding duplexes, two mixed homooligomeric TANA decamers **44** and **45** were synthesized. To establish the binding stoichiometry of the mixed TANA oligomers **44** Job's plot study was undertaken with complementary RNA sequences. (Figure 13) The binding stoichiometry was found as 1:1 between sequence **44** and complementary RNA **53** using UV spectroscopy (Figure 13).

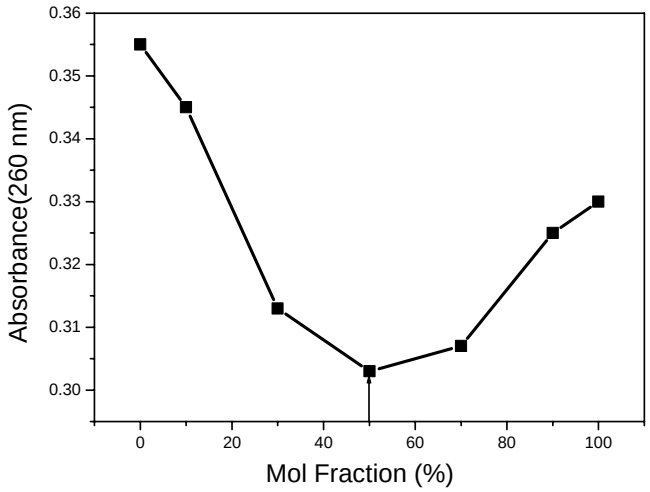


Figure 13. UV-Job's plot for complexes of TANA ON **44** and the complementary RNA **53** in the molar ratios of 0:100, 10:90, 30:70, 50:50, 70:30, 90:10, 100:0 (Buffer, 10 mM Sodium phosphate pH 7.0, 100 mM NaCl, 0.1 mM EDTA).

The *T*m values of mixed TANA ONs **44** and **45** hybridized with DNA **52**/RNA**53** bindings were determined from temperature-dependent UV absorbance and summarized in Table 5 (Figure 14). The experiments were performed with 1:1 stoichiometric ratio of TANA:complementary DNA/RNA. For control, the T_m of complexes of DNA 3' TTCTTTCTCT 5' **51** with complementary DNA **52** and RNA **53** were measured under

identical buffer conditions. It was found that there is no complex formation of TANA **44** and **45** with complementary DNA **52**.

Table 5. UV-T_m values in ^oC of TANA: DNA/RNA duplexes

Entry	Sequence	DNA **52**	RNA **53**
1	3' TTCTTTCTCT 5' **51**	23.6	25.4
2	H-**ttctttctct**-Lys **44**	nd	56.6 (+31.2)
3	H-**ttc^{m}tttc^{m}tc^{m}t**-Lys **45**	nd	57.2 (+31.8)

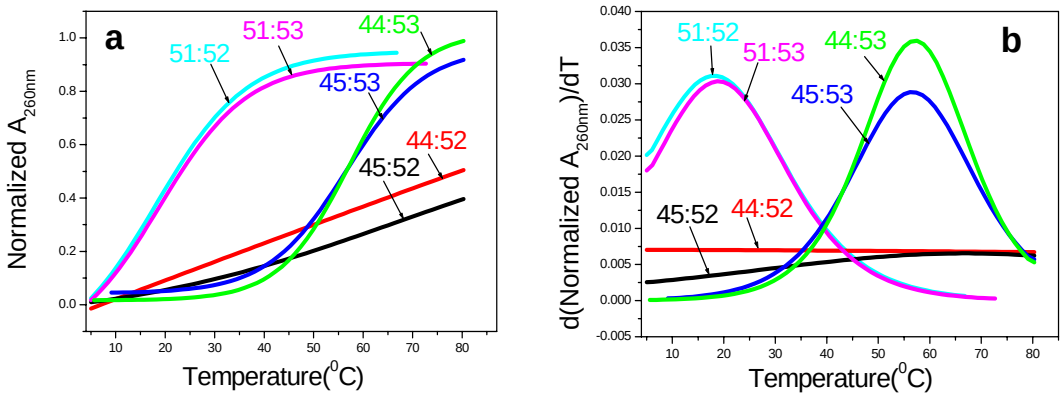


Figure 14. a.UV-melting Curves of **44, 45** and **51** with complementary DNA **52** and RNA **53** **b.** Corresponding first derivative Curves**.**

The TANA ONs **44** and **45** are showing excellent stability of the complex formed with their complementary RNA **53.** Both the TANA ONs **44** and **45** show almost similar melting temperatures (56.6 and 57.2 ºC) and the melting temperature was found to be 31-32ºC more than the control DNA **51**: RNA **53** complex. The UV melting studies were carried out at 2:1 stoichiometric ratio of TANA ONs **44** and **45**: complementary RNA **53** and pH = 5.5, but there were no significant changes of melting temperature were observed.

2.1.5.4 CD spectroscopic studies of TANA: DNA/ RNA complexes

CD spectra of the single stranded TANA ONs **38**, **40**, **41** and **42** show a positive CD band at 260-270 nm regions and a strong negative CD band at 200-210 nm regions (Figure 15 A). CD spectra of ONs **38**, **40, 41** and **42** with their complementary RNA **48** show that there were no so much difference observed in the CD spectra of the single stranded RNA

48 and its complexes **38:48**, **40:48**, **41:48** and **42:48** (Figure 15 C). In all these cases there are two positive CD bands at 265nm and 220 nm and a negative CD band at 250 nm. Figure 14B shows the CD spectra of aeg PNA **43** with its complementary RNA **48.** The results suggest that the nature of the complexes of TANA ONs with complementary RNA is not identical with the complexes formed by *aeg* PNA (Figure 15 B and C).

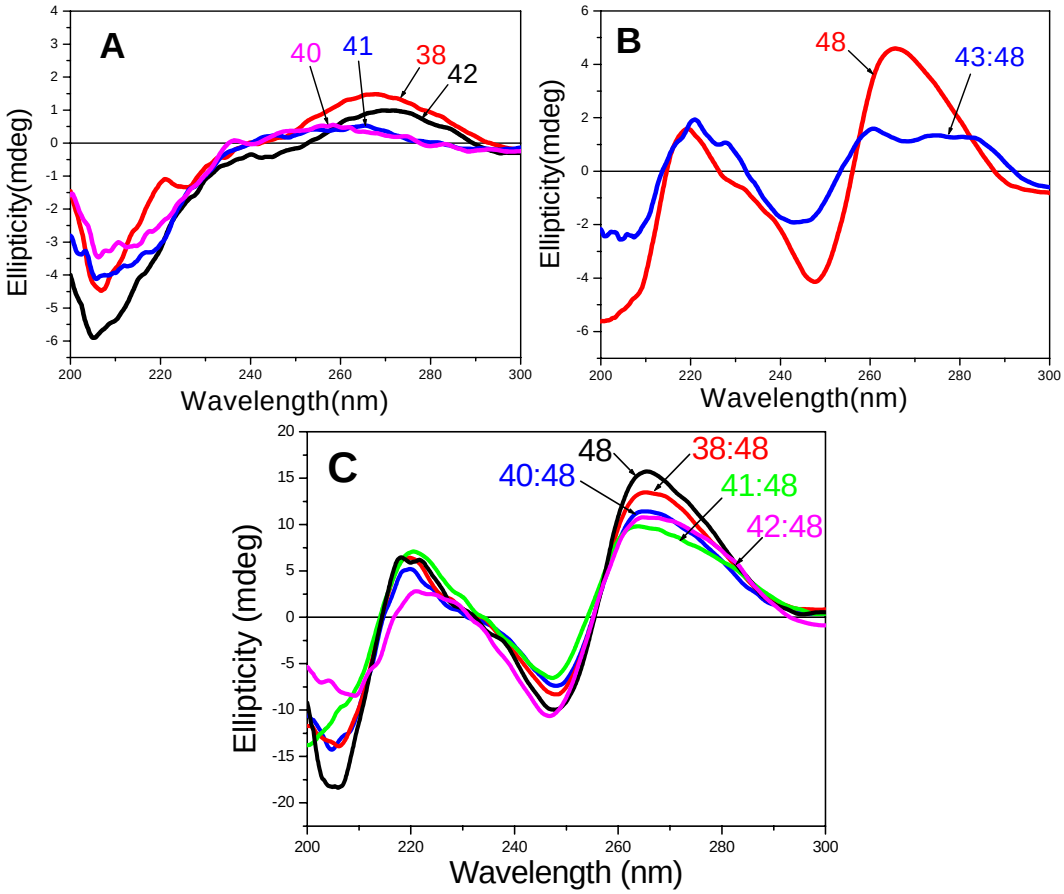


Figure 15. CD spectra of **A.** single stranded TANA ONs **38**, **40**, **41** and **42**. **B**. complex **43:48** and **48** **C**. complexes **38:48**, **40:48**, **41:48,** **42:48** and **48.** (10 mM sodium phosphate buffer, pH=7.0, 100 mM NaCl, 0.1 mM EDTA, 20°C).

The single stranded CD spectra of TANA ONs **44** and **45** showing a wide negative CD band in the region of 240-270 nm (Figure 16 A). The CD spectra of ONs **44** and **45** with the complementary RNA **53** show two positive CD bands at 270 nm and 220 nm and a negative band at 250 nm (Figure 16 B), which is characteristics for duplex DNA: RNA CD spectra.

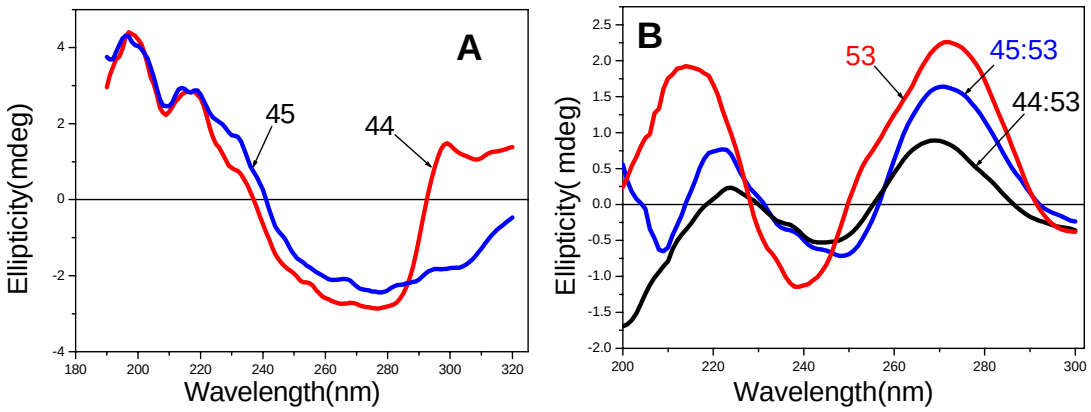


Figure 16. CD spectra of **A.** single stranded TANA ONs **44** and **45**. **B**. complexes **44:53** and **45:53**(10 mM sodium phosphate buffer, pH=7.0, 100 mM NaCl, 0.1 mM EDTA, 20°C).

2.1.6 Synthesis of trityl protected TANA monmers: synthesis of oligomers by trityl based solid phase synthesis

In earlier sections, studies of the effect of the TANA backbone with the nucleobases thymine, cytosine and 5-methyl cytosine are discussed confirm that pyrimidine sequences all give triplexes. To study the effect of TANA backbone on duplex formation, it is essential to synthesize the mixed purine-pyrimidine oligomers comprising all the nucleobases. We described solid phase synthesis of pyrimidine oligomers using Fmoc strategy and Rink amide resin was used as the solid support. Cleavage of oligomers from support requires the use of 10% trifluoro acetic acid. The purine nucleosides are readily depurinated in presence of acids like TFA, and 10% TFA will be good enough for significant depurination. So we have to choose a strategy where the loss of nucleobases by depurination is minimum. Instead of Fmoc protecting group for amino functions, we now propose trityl as the 3'- NH_2 protecting group for which, the deprotection conditions require only 3% DCA - DCM and CPG (control pore glass) as the solid support. The cleavage of oligomers from CPG solid support can be done by treatment with aqueous concentrated ammonia. A schematic diagram for this proposed strategy is depicted in Figure 15. In this section we describe the synthesis of trityl protected TANA monomers which would be useful synthons for oligomer synthesis.

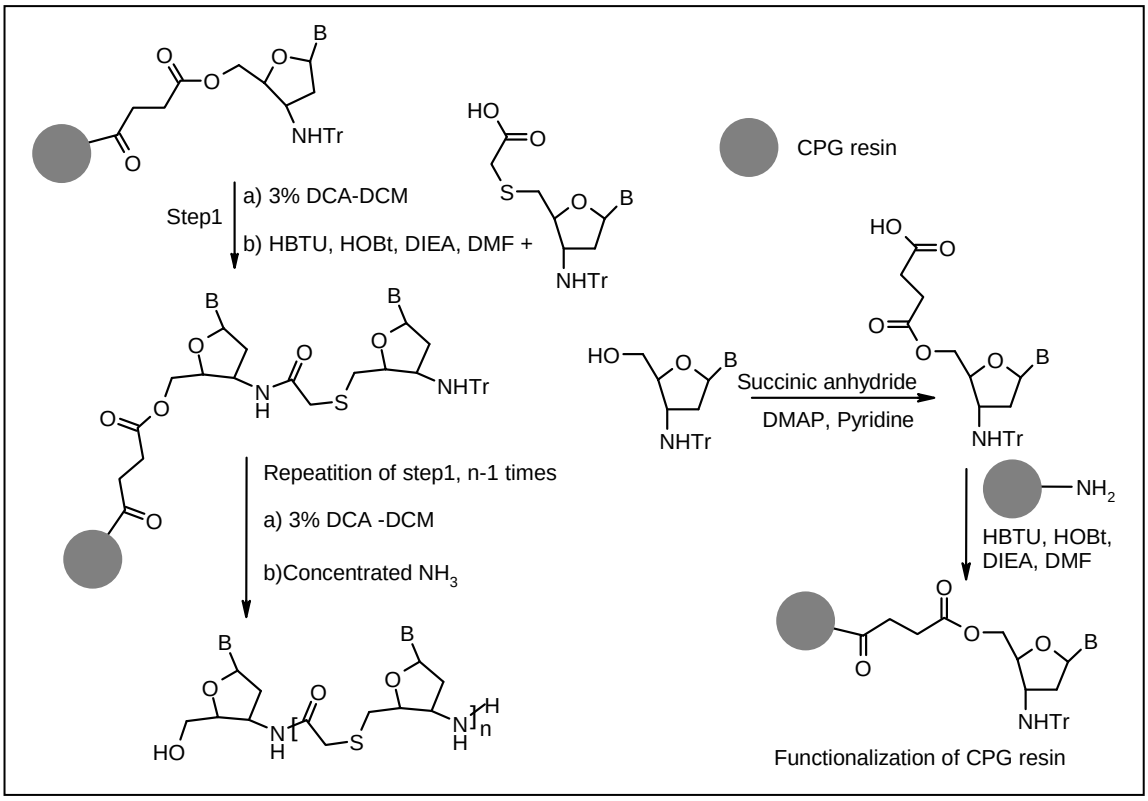
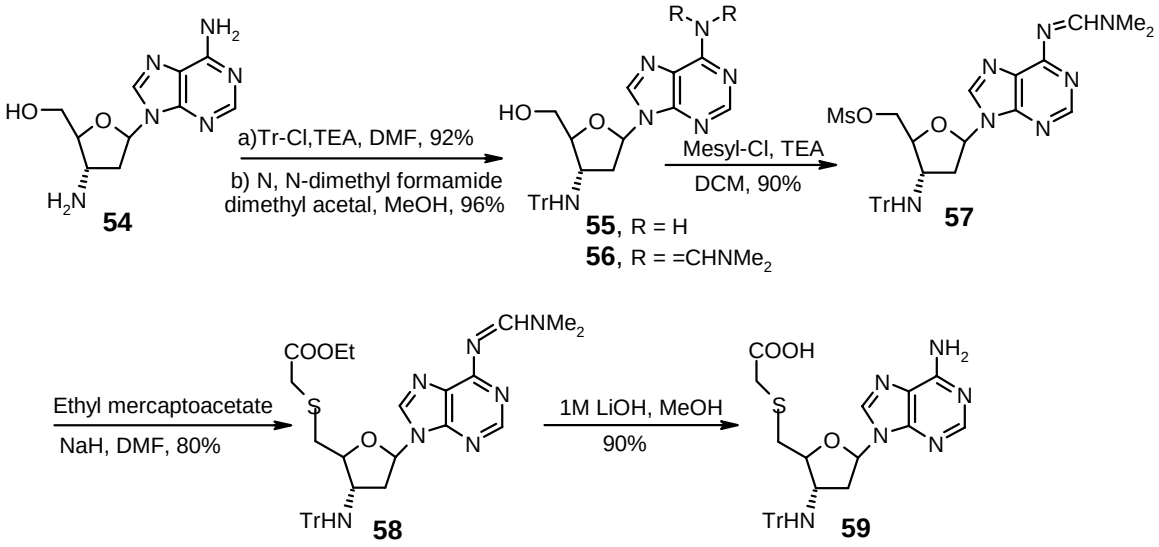


Figure 15. Trityl based solid phase peptide synthesis of TANA ONs

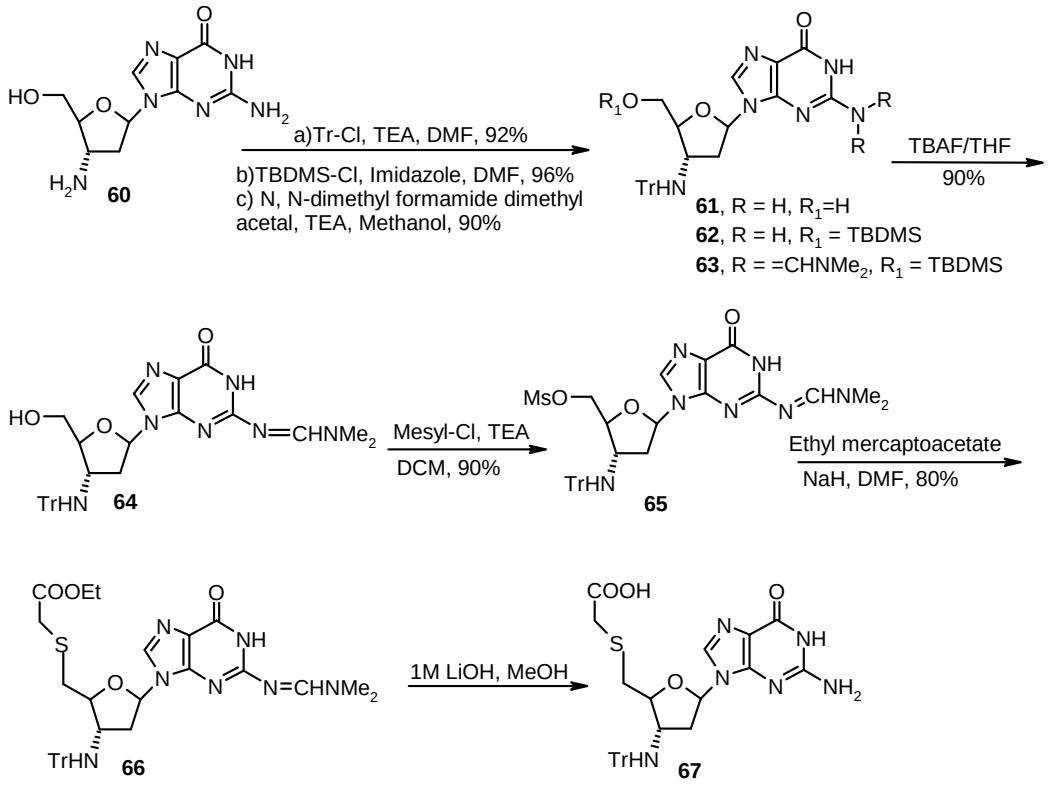
2.1.6.1 Synthesis of trityl protected deoxyadenosinyl TANA monomer



Scheme 6. Synthesis of trityl protected deoxy adenosine TANA monomer

Synthesis of deoxy adenosinyl monomer (Scheme 6) was done starting from commercially available 3'-amino 2', 3'-dideoxy adenosine **54**.[47] The 3'-amino group of **54** was protected as its trityl derivative **55** by reaction with trityl chloride and triethyl amine in DMF.[48] The exocyclic C6-NH_2 group was protected as N, N dimethyl formamidine **56** by reaction with N, N-dimethyl formamide dimethyl acetal in methanol. Then conversion of 5'-OH of **56** to corresponding mesylate **57**, which on nucleophilic substitution with ethyl mercaptoacetate to gave the corresponding ethyl mercapto ester **58**. The ester functionality of **58** was hydrolyzed using 1M LiOH in methanol to give the monomer building block **59**, and it was found that the exocyclic amidine protection was also cleaved during the hydrolysis.

2.1.6.2 Synthesis of trityl protected deoxyguanosinyl TANA monomer

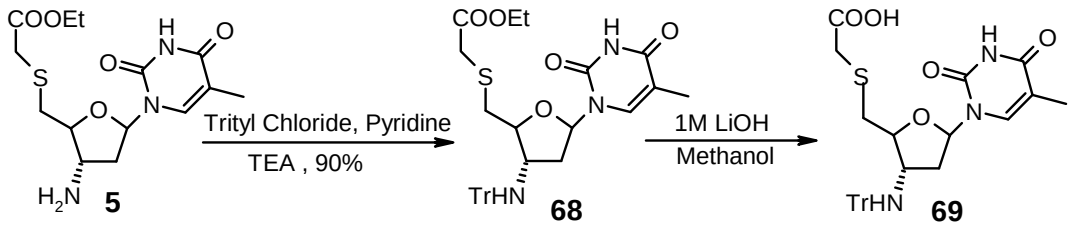


Scheme 7. Synthesis of trityl protected deoxy guanosine TANA monomer

Synthesis of deoxy guanosinyl monomer (Scheme 7) was done starting from commercially available 3'-amino 2', 3'-dideoxyguanosine **60**.[47] The 3'-amino group of

60 was protected as its trityl derivative **61** by reaction with trityl chloride and triethyl amine in DMF.[48] The compound **60** is highly polar and partially soluble in DMF at room temperature, so the substrate was fully solubilize in DMF by keeping the reaction mixture at 60ºC for 30 minutes before addition of trityl chloride. Then it was attempted to protect the exocyclic C2-NH_2 group as its dimethyl formamidine derivative in methanol, but it could not succeeded due to insolubility of the substrate **61** in methanol. So the 5'- OH of **61** was converted to the silyl ether derivative **62** to make the compound less polar. Then dimethyl formidine protection to give **63**, followed by desilylation to give compound **64**. Conversion of 5'-OH of **64** to corresponding mesylate **65**, which on nucleophilic substitution with ethyl mercaptoacetate to gave the corresponding ethyl mercapto ester **66**. The ester functionality of **66** was hydrolyzed using 1M LiOH in methanol to give the monomer building block **67**, and it was found that the exocyclic amidine protection was also cleaved during the hydrolysis.

2.1.6.3 Synthesis of trityl deoxy thyminyl TANA monomer



Scheme 8. Synthesis of trityl protected thyminyl TANA monomer

Synthesis of trityl protected thyminyl TANA monomer **69** (Scheme 8) was reached from the intermediate **5** discussed in Scheme 1. The 3'- NH_2 of **5** was converted to its trityl derivative followed by hydrolysis to give the corresponding monomer synthon **69**.

The synthesis of cytosinyl monomer is on progress and the oligomers synthesis remains to be done and the method is not fully established, so the results are not reported here.

2.1.7 Conclusion

In conclusion, we have developed a novel class of neutral ONs having extended backbone NH3'-CO-CH_2-S-CH_2 to improve the binding affinity and selectivity towards DNA/RNA recognition. The homooligomeric pyrimidine ONs were found to bind to

complementary RNA sequences significantly better than their DNA counterparts and the binding efficiency was found to be as good as PNA itself. However, the chimeric TANA-*aeg*PNA sequences were found to be detrimental for RNA/DNA binding. The stability of the complexes of homogeneous backbone TANA with RNA over DNA is a very valuable result from application perspective. Further studies on the mixed purine-pyrimidine sequences are necessary to draw clear conclusions. The NMR solution structure and molecular modeling studies are also necessary to elucidate the structure-activity relationship exhibited by TANA oligomers that is responsible for the unusual RNA recognition.

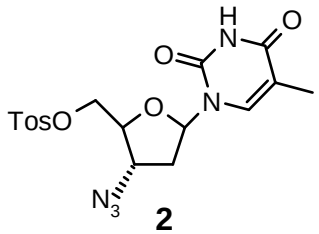
Summary

- Neutral, one atom extended thioacetamido nucleic acids (TANA) were designed for specific recognition of nucleic acids.
- The thioacetamide nucleic acids with preorganized geometry results in optimization of internucleobase distance complementarity.
- The TANA ONs are showing unprecedented RNA binding selectivity to form triplexes and duplexes.

2.1.8 Experimental

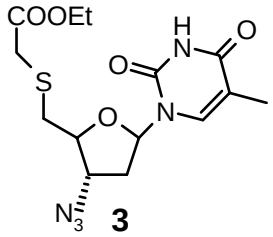
***General*:** The chemicals used were of laboratory or analytical grade. All solvents used were purified according to the literature procedure. Reactions were monitored by TLC. Usual reaction work up involved sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulfate 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) and 200- 400 mesh (Merck) and using the solvent systems EtOAc/Pet 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 spraying perchloric acid solution and 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 MALDI-TOF or LCMS techniques. Oligomers were characterized by RP HPLC (Hewlett Packard) C18 column and MALDI-TOF mass spectrometry. The MALDI-TOF spectra were recorded on Voyager-De-STR (Applied Biosystems) MALDI-TOF instrument and the matrixes used for analysis were CHCA (α-Cyano-4-hydroxycinnamic Acid), THAP (2', 4', and 6'-trihydroxyacetophenone) and HPA (3-Hydroxy picolinic acid) and diammonium citrate was used as additive when THAP and HPA were used as matrices. CD spectra were recorded on a JASCO J715 spectropolarimeter. UV-melting experiments were carried out on Perkin-Elmer *Lambda-35* UV-spectrophotometer.

Synthesis of 3'-azido-5'-O-tosyl thymidine (**2)** A solution of 2.68 g (10 mmol) of **1**, AZT (3'-azido thymidine) in 20 ml of dry pyridine was cooled to 0º C. p-toluenesulfonyl (tosyl) chloride (2.3 g, 1.2 mmol) dissolved in 5 ml of pyridine was then added drop wise. The reaction mixture was stirred at room temperature for four hours. The mixture was concentrated *in vacuo* and redissolved in 100 ml of ethyl acetate. The organic layer was washed with 10% $NaHCO_3$ solution (3 x 30 ml), water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml) and then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column chromatography (30% ethyl acetate in petroleum ether) to afford 2.7 g **2** (85%) as a white foam. IR, ν (cm^{-1}) ($CHCl_3$); 3128, 3018, 2109, 1724, 1670 cm^{-1}. ^{1}H NMR: ($CDCl_3$) δ1.89 (s, 3H, 5-CH_3), 2.3-2.4 (m, 2H; 2', 2''-H), 2.41 (s, 3H; tosyl-CH_3), 3.98 (m,1H; 4'-H), 4.18-4.2 (m,1H; 3'-H), 4.26 (d, 2H; 5',5''-H), 6.20 (t, 1H; 1'-H), 7.33 (dd, 3H; 5-H, tosyl m-H), 7.78 (dd, 2H; tosyl -H),10.14 (s, b, 1H; 3-NH).



2

Synthesis of Ethyl S-5'-deoxy-5'- (3'-azido-thyminyl)-mercapto acetate (**3)** NaH (60% solution in hexane) 0.24 g (6 mmol) was taken of in 5 ml of anhydrous DMF and to it was added 0.66 ml (6 mmol) ethyl mercapto acetate. The mixture was stirred for 15 minutes and 2.1 g (5 mmol) of **2** dissolved in 2 ml of DMF was then added. After 45 minutes the mixture was concentrated *in vacuo* and redissolved in 50ml ethyl acetate. The organic layer was washed with water (3 x 20 ml), saturated NaCl (2 x 15 ml) and then dried over Na_2SO_4, evaporated to dryness and purified by column chromatography (30% ethyl acetate in petroleum ether) to afford 1.45 g (76%) of **3**. IR, ν(cm^{-1}) ($CHCl_3$) 3186.18, 3018.39, 2108.05, 1712.67, 1704.96, 1693.38, 1681.81. ^{1}HNMR: ($CDCl_3$, 200 MHz) δ 1.26-1.34 (t, 3H; -CH_3), 1.94 (s, 3H; 5-CH_3), 2.39-2.46 (m, 2H; 2', 2''-H), 3.03-3.08 (m, 2H; -CH_2), 3.3-3.44 (m, 2H; -SCH_2), 4.02-4.27 (m, 4H; 4'-H, 3'-H; 5', 5''-H), 6.12-6.18 (t,1H;1'-H), 7.33(s, 1H; 5-H), 9.96 (s, b,1H, 3-NH).

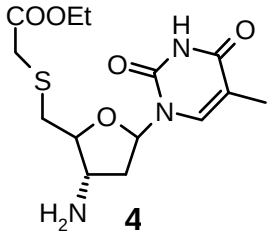


3

^{13}C NMR (50 MHz, $CDCl_3$) δ 11.7, 13.4, 33.4, 33.7, 36.2, 60.8, 61.7, 82.4, 84.3, 110.5, 135.3, 149.9, 163.6, 169.50. $[\propto]_{20}^{D}$= + 42.5˚ (c 0.5, $CHCl_3$).

Synthesis of ethyl-S-5'-deoxy-5'- (3'-amino-thyminyl)-mercapto acetate 4

The Ethyl S-(3'-azido-thyminyl)-mercapto acetate **3** (1.2 g, 3.2 mmol)**,** was dissolved in 15% triethylamine in pyridine (5 ml) and H_2S gas was bubbled into the solution at 0º C for 30 minutes. The solution was then stirred at room temperature for an additional 30 minutes and solvent removed *in vacuo*. The residue was purified by flash silica gel (200-400 mesh) column chromatography (5-10% methanol in DCM) to afford 0.9g **4** (81%).The compound was found to be unstable for long time and immediately used for the next reaction. IR, ν(cm^{-1}) ($CHCl_3$) 3398, 3018, 2923, 2852, 1701 and 1685.

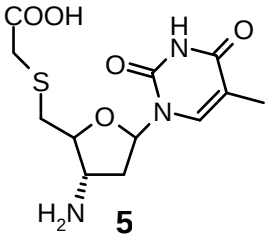


4

^{1}HNMR: ($CDCl_3$, 200 MHz) δ 1.23-1.30 (t, 3H; -CH_3), 1.89 (s, 3H; 5-CH_3), 2.18-2.39 (m, 2H; 2', 2''-H), 3.06-3.17 (m, 2H, 5',5''-CH_2), 3.24-3.41 (m, 2H, -SCH_2), 3.7--3.77 (m, 1H, 4'-H), 4.07-4.24 (m, 3H; 3'-H, CH_2), 6.24-6.4 (t, 1H; 1'-H), 7.39 (s, 1H; 5-H), 7.72 (b,1H; NH).

Synthesis of *S*-5'-deoxy-5'- (3'-amino-3'deoxy-thyminyl)-mercapto acetic acid (5)

The ethyl-S-(3'-amino-thyminyl)-mercapto acetate **4** (0.6 g,1.67 mmol) was dissolved in 4 ml methanol and to it was added 2 ml of 2 M NaOH solution. The mixture was stirred at room temperature for 1 h. The excess NaOH was neutralized by Dowex-$50H^+$ resin which were then filtered off. Concentration of the filtrate *in vacuo* afforded 0.53 g of **5** as a white solid, which was immediately used for the next reaction.^{1}H NMR (D_2O, 200MHz) δ: 1.66 (s, 3H; 5-CH_3), 2.16(m, 2H; 2',2''-H), 2.71-2.76 (m, 2H, -SCH_2), 3.2-3.25 (m, 2H; 5''5''-H), 3.88 (m,1H; 4'-H), 4.20 (m, 1H; 3'-H), 6.0 (t, 1H; 1'-H), 7.3 (s, 1H; 6-H).

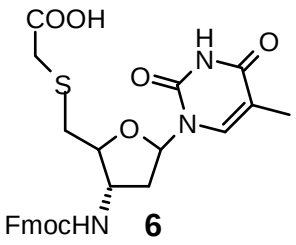


5

^{13}C NMR (D_2O, 50 MHz) δ 11.5, 34.1, 34.3, 37.7, 72.7, 85.1, 111.5, 137.4, 151.5, 166.3, 174.6.

_S_-5'-deoxy-5'- (3'-Fluorenylmethoxycarbonylamino-3'-deoxy-thyminyl)-mercapto acetic acid 6

_S_-(3'-amino-3'deoxy-thyminyl)-mercapto acetic acid **5** (0.45 g, 1.26 mmol) was dissolved in 5 ml acetone-water (50:50) and to it was added 0.53 g (6.3 mmol) of $NaHCO_3$, followed by 0.51 g (1.51 mmol) of (_N_-9-fluorenylmethoxycarbonyloxy) succinimide. After stirring for three hours acetone was removed and the reaction mixture was diluted with 2 ml water. The water layer was extracted with ethyl acetate to remove all the nonpolar impurities. The water layer was then neutralized by adding dilute HCl up to pH=7. The aqueous layer was then extracted with ethyl acetate. The organic layer was washed with water (1x10ml) and saturated NaCl solution (1x10ml). The organic layer was dried over Na_2SO_4, evaporated to dryness to afford 0.4g of **6** (50%) as pure white amorphous solid. M.p.147-150°C. $[\alpha]_D^{20}$ = +16 (c 0.5; CH_3OH) IR, ν (cm^{-1}) 3326, 2927, 2358, 1704, 1677, 1529, 1463.

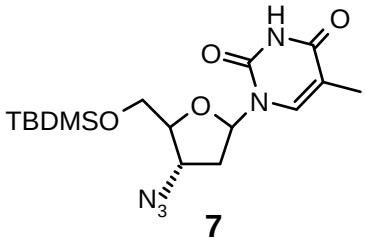


6

^{1}HNMR: (DMSO-d_6, 200 MHz) δ: 1.78 (s, 3H; 5-CH_3), 2.06-2.28 (m, 2H; 2',2''-H), 3.15-3.38 (m, 2H; S-CH_2), 3.4-4.32 (m, 6H; 4'-H, 3'-H; 5',5''-H,-CH_2 Fmoc), 6.06-6.12 (t, 1H; 1'-H), 7.27-7.96 (m, 10H; 6-H, Fmoc-H), 11.33 (bs, 1H, -NH).

^{13}C (DMSO-d_6, 50 MHz) δ: 12.6, 34.6, 36.4, 47.2, 53.7, 62.6, 64.3, 83.2, 83.8, 110.5(C), 120.6, 125.6, 127.6, 128.2, 136.7, 141.3, 144.2, 150.9, 156.4, 164.4, 172.0. Anal Calcd (%) for $C_{27}H_{27}N_3O_7S$: C 60.19; H 5.2; N 7.8, S 5.9 Found C 60.7; H 5.02; N 7.4; S 6.03. MS, M+Na^+, Calculated 560.33; Observed, MALDI-TOF 560.38, LCMS 560.06.

Synthesis of 3'-azido-5'-_O-tert_. Butyldimethylsilyl Thymidine 7



7

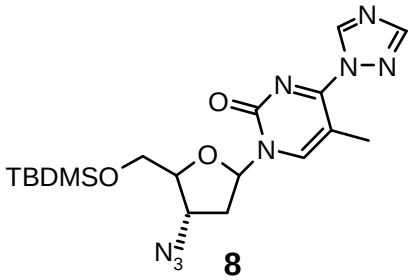
A mixture of AZT **1** (2.68 g, 10 mmol)), Imidazole(1.71 g, 25 mmol) and TBDMS-Cl (1.8 g, 12 mmol) in 10 ml DMF was stirred for six hours at room temperature. The mixture was concentrated _in vaccuo_ and redissolved in 100 ml ethyl acetate. The organic layer was washed with 10% $NaHCO_3$ solution (3 x 30 ml), water (2 x 50 ml) and

saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column filtration (30% ethyl acetate in petroleum ether) to afford 3.7g (96%) as white foam.

^{1}H NMR (200 MHz, $CDCl_3$) δ: 0.11(s, 6H; -CH_3), 0.91 (s, 9H, -CH_3), 1.90 (s, 3H; 5-CH_3), 2.13-2.27 (q, 1H, 2'-H), 2.36-2.48 (m, 1H; 2''-H), 3.74-3.82 (q, 1H, 4'-H), 3.9-3.96(m, 2H; 5',5''-H), 4.18-4.26 (m, 1H; 3'-H), 6.18-6.25 (t, 1H; 1'-H), 7.429 (s,1H; 6-H), 9.22 (s, b,1H; 3-NH).

Synthesis of 3'-azido-5'-*O-tert.* butyldimethylsilyl-C^4-(-1, 2, 4-triazol-1-yl)-pyrimidine 2'-deoxyribonucleoside 8

Triethylamine (12.7 ml, 91 mmol) was added dropwise to a stirred mixture of 1, 2, 4-triazole (7 g, 100 mmol) and phosphoryl chloride (2 ml, 22 mmol) in 50ml CH_3CN at 0ºC. A solution of **7** (3.5 g, 9.1 mmol) in 15 ml dry CH_3CN was then added to the reaction mixture, and stirred at room temperature for 2.5 hours. Triethylamine (4.4 ml) and water (1.1 ml) were added at 0ºC and the mixture was stirred for further 10 minutes. The solvent was evaporated and the residue was redissolved in 50 ml ethyl acetate. The organic layer was washed with water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness to give 3.6g (92%) homogeneous (by TLC analysis) of **8** as yellow foam.



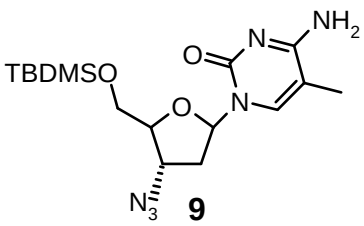
8

^{1}HNMR ($CDCl_3$, 200 MHz) δ: 0.10 (s, 6H; -CH_3), 0.87 (s, 9H; -CH_3), 2.27-2.37 (m, 1H; 2'-H), 2.42 (s, 3H, -CH_3), 2.7-2.82 (m, 1H, 2''-H), 3.77-3.84 (q, 1H; 4'-H), 3.99-4.09 (m, 2H; 5',5''-H), 4.13-4.21(m, 1H; 3'H), 6.12-6.18 (t, 1H; 1'-H), 8.08 (s, 1H, 6-H), 8.19-8.22 (d, 2H, triazol-H).

Synthesis of 3'-azido-5'-*O-tert.*butyldimethylsilyl-(-5-methylcytidine) 9

Concentrated aqueous ammonia (10 ml) was added to the solution of **8** (3.5 g, 8 mmol) in 50 ml dioxane. The reaction mixture was stirred at room temperature for 2.5 hours. The solvents were evaporated *in vacuo* and the residue redissolved in CH_2Cl_2 and purified by

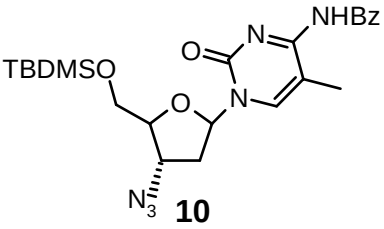
silica-gel column chromatography (eluted with 0-5%methanol in CH_2Cl_2) to afford **9**(3 g, 96%) as white foam.



9

^{1}H NMR ($CDCl_3$, 200MHz) δ: 0.13 (s, 6H; -CH_3), 0.93 (s, 9H; -CH_3), 1.96 (s, 3H; -CH_3), 2.16-2.29 (m, 1H; 2'-H), 2.49-2.61 (m, 1H; 2''-H), 3.77-3.85 (m, 1H; 4'-H), 3.94-4.00 (m, 2H; 5',5''-H), 4.14-4.24 (m, 1H; 3'-H), 6.22-6.28 (t, 1H; 1'-H), 7.56 (s, 1H; 6-H).

5'-*O*-*tert*.butyldimethylsilyl-N^4-benzoyl-3'-azido-2',3'-dideoxy-5-methylcytidine 10



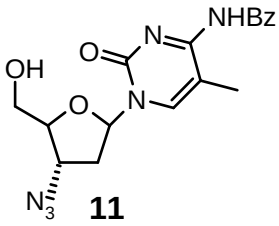
10

Benzoyl chloride (2.6 ml, 22.7 mmol) was added dropwise slowly to a solution of **9** (2.9 g, 7.56 mmol) in 40 ml pyridine at 0ºC. The mixture was then stirred at room temperature overnight. Then 8 ml water was added and the reaction mixture was stirred for 5 min at 0ºC, followed by addition of concentrated ammonia (8 ml) at 0ºC and stirring for another 30 minute. The solvent was evaporated *in vacuo* and the residue was taken in 50ml ethyl acetate. The organic layer was washed with 10% $NaHCO_3$ solution (3 x 30 ml), water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column filtration (10% ethyl acetate in petroleum ether) to afford (2.9 g, 76%) as a white foam. ^{1}H NMR ($CDCl_3$, 200 MHz) δ: 0.11(s, 6H; -CH_3), 0.91(s, 9H; -CH_3), 2.04-2.05 (d, 3H; -CH_3), 2.15-2.30 (m, 1H; 2'-H), 2.4-2.52 (m, 1H; 2''-H), 3.73-3.79(m, 1H; 4'-H), 3.89-3.98 (m, 2H; 5',5''-H), 4.16-4.24 (m,1H, 3'-H), 6.13-6.2 (t, 1H,1'-H), 7.38-7.58 (m, 5H; 6-H, benzoyl), 8.24-8.29 (m, 2H; Benzoyl), 13.25 (b s, 1H; -NH).

Synthesis of 3'-azido-*N*^{4}-benzoyl-5-methylcytosine 11

To a solution of 2.8 g (5.47 mmol) of **10** in 10ml dry THF, 6.6ml 1M (6.6 mmol) solution of TBAF in THF was added and the reaction mixture was stirred for 2 hours at room temperature. The solvent was removed and residue was redissolved in 50 ml ethyl acetate. The organic layer was washed with 2N HCl (2 x 20 ml), 10% $NaHCO_3$ solution

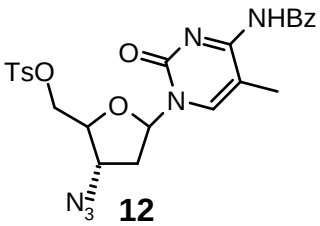
(3 x 30 ml), water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column filtration (40 % ethyl acetate in petroleum ether) to afford 2.1g (98%) as a white solid. ^{1}HNMR ($CDCl_3$, 200MHz) δ: 2.07 (s, 3H; -CH_3), 2.37-2.61 (m, 2H; 2',2'' -H), 3.81-4.07 (m, 3H; 4'-H, 5',5''-H), 4.39-4.47(m, 1H; 3'-H), 6.1-6.16 (t, 1H; 1'-H), 7.39-7.54 (m, 3H; Ph), 7.68 (s, 1H; 6-H), 8.27-8.32 (m, 2H; Ph), 13.23 (bs, 1H; NH).



11

Synthesis of 3'-azido-5'-*O*-tosyl-N^4-benzoyl-5-methylcytosine 12

Solution of 1.86 g (5 mmol) of **11** in 10 ml of dry pyridine was cooled to 0ºC. p-toluene sulfonyl (tosyl) chloride (1.2 g, 1.2 mmol) dissolved in 3 ml of pyridine was then added dropwise. The reaction mixture was stirred at room temperature for five hours. Pyridine was removed under reduced pressure and the residue was redissolved in 50 ml of ethyl acetate. The organic layer was washed with 10% $NaHCO_3$ solution (3 x 15 ml), water (2 x 20 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column chromatography (20% ethyl acetate in petroleum ether) to afford 2.2 g **12** (85%)as a white foam.

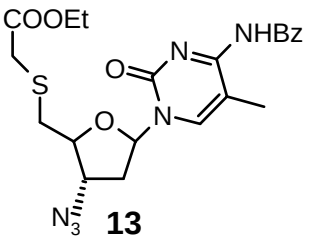


12

^{1}H NMR ($CDCl_3$, 200 MHz) δ: 2.13 (s, 3H; CH_3), 2.3-2.56 (m, 2H; 2',2''-H), 2.45 (s, 3H; CH_3 tosyl), 3.99-4.39 (m, 4H; 4'-H, 5',5''-H, 3'-H), 6.18-6.24 (t, 1H; 6-H), 7.36-7.56 (m, 6H; 6-H, Ph, tosyl), 7.78-7.82 (m, 2H; tosyl), 8.28-8.33 (m, 2H; Ph),13.2 (bs, 1H; NH).

Synthesis of ethyl *S*-5'-deoxy-5'- (3'-azido-N^4-benzoyl-5-methylcytosinyl) - mercapto acetate 13 NaH (60% solution in hexane) (0.18 g, 4.4 mmol) was suspended of in 5 ml of anhydrous DMF and to it was added 0.45 ml (4.4 mmol) ethyl mercapto acetate. The mixture was stirred for 15 minutes and then 2.1 g (4 mmol) of **12** dissolved in 2 ml of

DMF was added. After 45 minutes the mixture was concentrated *in vacuo* and redissolved in 50ml ethyl acetate. The organic layer was washed with water (3 x 20 ml) and saturated NaCl (2 x 15 ml). The organic layer was dried over Na_2SO_4, evaporated to dryness and residue purified by column chromatography (12 % ethyl acetate in petroleum ether) to afford 1.4 g (71%) of **12**.



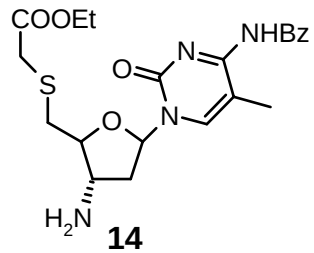
13

^{1}H NMR ($CDCl_3$, 200 MHz) δ: 1.26-1.33 (t, 3H; -CH_3), 2.12-2.13 (d, 3H; 5-CH_3), 2.31-2.60 (m, 2H; 2', 2''-H), 2.97-3.16 (q, 2H; -CH_2), 3.25-3.41 (m, 2H; -SCH_2), 4.04-4.28 (m, 4H; 4'-H, 5',5''-H, 3'-H), 6.11-6.17 (t, 1H; 1'-H), 7.39-7.57 (m, 5H; 6-H, Ph), 8.28-8.33 (m, 2H; Ph), 13.25 (bs, 1H; -NH).

^{13}C NMR ($CDCl_3$, 50 MHz) 13.3, 1.9, 34.0, 34.4, 37.2, 61.4, 62.0, 83.2, 85.4, 112.07, 127.9, 129.7, 132.3, 136.4, 136.9, 147.5, 159.3, 169.7, 179.3.

$[\alpha]_D^{20}$ = +44°(c 0.5; CH_3OH).

Synthesis of ethyl *S*-5'-deoxy-5'- (3'-amino-*N*^{4}-benzoyl-5-methylcytosinyl)-mercapto thioacetate 14 Compound **13 (**1.3 g, 2.65 mmol) was dissolved in 15% triethylamine in pyridine (5 ml) and H_2S gas was bubbled into the solution at 0º C for 15 minutes. The solution was then stirred at room temperature for an additional 30 minutes and solvent removed *in vacuo*. The residue was purified by flash silica gel (200-400 mesh) column chromatography (2-5 % methanol in DCM) to afford 1g **14** (82%).



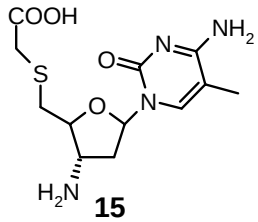
14

^{1}H NMR ($CDCl_3$, 200 MHz) δ: 1.26-1.33 (t, 3H; -CH_3), 2.13-2.14 (d, 3H; 5-CH_3), 2.24-2.32 (m, 2H; 2',2''-H), 3.06-3.09 (m, 2H, -CH_2), 3.26-3.43 (m, 2H, -SCH_2), 3.46-3.57 (m, 1H, 4'-H), 3.84-3.94 (m, 1H; 3'-H), 4.15-4.26 (m, 2H; 5'5''-H), 6.17-6.23 (t, 1H; 1'-H), 7.39-7.57 (m, 3H; Ph), 7.64 (s, 1H; 6-H), 8.29-8.34 (m, 2H; Ph).

^{13}C NMR ($CDCl_3$, 50 MHz)δ: 13.4, 14.0, 34.1, 34.2, 41.5, 53.7, 61.4, 84.8, 86.3, 111.72, 128.0, 129.7, 132.3, 136.7, 147.7, 159.6, 170.1, 179.3.

Synthesis of S-5'-deoxy-5'- (3'-amino-*N*^{4}-benzoyl-5-methylcytosinyl)-mercapto acetic acid 15

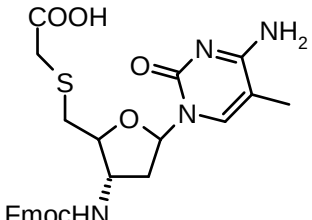
Compound **13** (0.9 g, 1.94 mmol) was dissolved in 2 ml methanol and to it was added 2 ml of 0.75 M LiOH solution. Stirred the mixture at room temp for 2 hour. The excess LiOH was neutralized by Dowex-50H$^+$ resin which were then filtered off. Concentration of the filtrate *in vacuo* afforded 0.6 g of **15** as a white solid, which was immediately used for the next reaction.



15

Synthesis of S-5'-deoxy-5'- (3'-Fluorenylmethoxycarbonylamino-*N*^{4}-benzoyl-5-methylcytosinyl) - mercapto thioacetic acid 16

Compound **15** 0.6 g (1.8 mmol) was dissolved in 5ml acetone-water (50:50) and to it was added 0.76 g (9 mmol) of $NaHCO_3$, followed by 0.73g (2.2 mmol) of (N-9-fluorenylmethoxycarbonyloxy) succinimide. After stirring for three hours acetone was evaporated. The reaction mixture was diluted with 3 ml water. The water layer was extracted with ethyl acetate to remove all the nonpolar impurities. Then neutralized the water layer by adding dilute HCl up to pH=7. Then extracted with ethyl acetate. Washed the organic layer with water (1 x 10 ml) and saturated NaCl solution (1 x 10 ml). The organic layer was dried over Na_2SO_4, evaporated to dryness afforded 0.48g of **16** (50%) as pure white amorphous solid. M.p.160-163°C $[\alpha]_D^{20}$= +18 (c 0.5; CH_3OH)



16

IR, ν(cm^{-1}), (nujol) 3238, 2923, 2854, 1714, 1666.

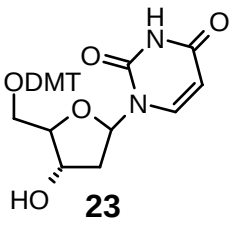
^{1}H NMR: ($CDCl_3$ + CD_3OD, 200 MHz) δ: 2.08 (s, 3H; 5-Me), 2.26-2.55 (m, 2H; 2', 2''-H), 3.2(m, 2H; -CH_2, Fmoc), 3.22-3.44 (m, 2H; -SCH_2), 3.7-4.39 (m, 4H; 4'-H, 5',5''-H, 3'-H), 6.09 (s, 1H; 1'-H), 7.29-8.07 (m, 10H; 6-H, Fmoc-H). ^{13}C (DMSO-d_6) δ: 13.4, 34.0, 35.9, 46.6, 53.9, 62.7, 66.06, 83.9, 83.9, 110.7, 120.7, 125.7, 127.9, 128.4, 137.0, 141.9, 144.5, 151.2, 156.5, 166.2, 172.3.

MS, M, M+Na$^+$ Calculated 536.60, 559.39; Observed MALDI-TOF, 559.43; LCMS 537.06.

5'-*O*-DMTr 2', 3' deoxyuridine 23

2.3 g (10 mmol) of 2', 3' deoxyuridine **22**, DMTr-Cl (4.1 g, 12 mmol), DMAP (122 mg, 0.1 mmol) were dissolved in 60 ml pyridine and to it added TEA (2.8 ml, 20 mmol). The

reaction mixture was stirred at room temperature for six hours. Removed the pyridine *in vacuo* and redissolved the residue in 250 ml ethyl acetate. Washed the organic layer with sodium bicarbonate (2 x 30 ml) followed by water (2 x 50 ml) and then brine (2 x 30 ml).The organic layer was kept over anhydrous Na_2SO_4, removed *in vacuo* to give **23** as a white foam which was desiccated and used for the next reaction without any purification.

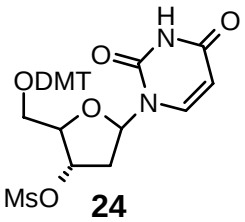


23

Yield: 4.8 g, 90 %

5'-*O*-DMTr 3'-*O*- mesyl 2', 3' deoxyuridine 24

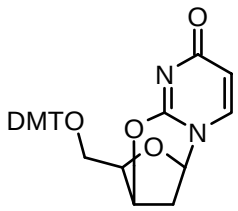
A solution of **17** (4.8 g, 9 mmol) in 50 ml DCM and 3.8 ml TEA (27 mmol) was cooled to 0˚C. Methane sulfonyl (mesyl) chloride (1.4 ml, 18 mmol, dissolved in 10 ml DCM) was then added drop wise. The mixture was stirred for 10 minutes. After the disappearance of the starting material (TLC analysis), the reaction was quenched by adding 30 ml 10% $NaHCO_3$ and was diluted with 50 ml DCM. The organic layer was dried over Na_2SO_4, evaporated to dryness to give **24** which was used immediately for the next reaction. Yield 5.5 g (95 %, crude).



24

5'-*O*-DMTr 3'-*O*- anhydro deoxyuridine 25

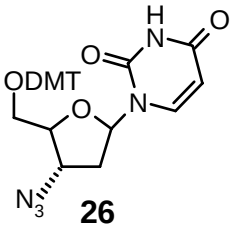
A solution of **24** (5.5 g, 9 mmol) in 100 ml anhydrous acetonitrile and 4 ml DBU (27 mmol) was refluxed for 3.5 hours. The solvent was evaporated to dryness and the residue was redissolved in 100 ml EtOAc. The organic layer was washed with water (2 x 50 ml) followed by brine. The organic layer was kept over anhydrous Na_2SO_4, evaporated to dryness to give **25**, which was used immediately for the next reaction. Yield 4 g (86 %).



25

5'-*O*- DMTr 3'azido deoxyuridine 26

4.0 g (8 mmol) **25** was dissolved in 50 ml anhydrous DMF and HMPA (1:1) and to it added 3.5g (56 mmol) NaN_3. The reaction mixture was stirred at 100ºC for six hours. The

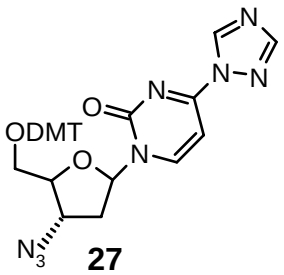


26

solvent was evaporated *in vacuo* and redissolved the residue in 100 ml EtOAc. The organic layer was washed with water (2 x 50 ml) followed by brine (2 x 20 ml) and kept over anhydrous Na_2SO_4. The solvent was evaporated to dryness and the product was purified by column chromatography using 1% MeOH- DCM as the solvent system to give **26**, yield 3.5 g (80%).

Synthesis of 3'-azido-5'-*O*-dimethoxy trityl -C^4-(-1, 2, 4-triazol-1-yl) - 2'-deoxy uridine 27

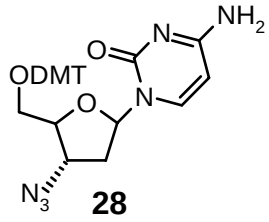
Triethylamine (8.8 ml, 63 mmol) was added dropwise to a stirred mixture of 1, 2, 4-triazole (4.8 g, 69 mmol) and phosphoryl chloride (1.5 ml, 15.8 mmol) in 50 ml CH_3CN at 0ºC. A solution of **26** (3.5 g, 6.3 mmol) in 15ml dry CH_3CN was then added to the reaction mixture, and stirred at room temperature for 2.5 hours. Triethylamine (4.0 ml) and water (1.0 ml) were added at 0ºC and the mixture was stirred for further 10 minutes. The solvent was evaporated and the residue was redissolved in 50 ml ethyl acetate. The organic layer was washed with water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness to give 3.6 g (94 %) homogeneous (by TLC analysis) of **27** as yellow foam.



27

Synthesis of 3'-azido-5'-*O*- dimethoxy trityl - (2'-deoxycytidine) 28

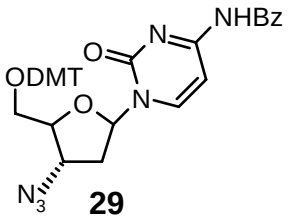
Concentrated aqueous ammonia (10 ml) was added to the solution of **27** (3.6 g, 8 mmol) in 50 ml dioxane. The reaction mixture was stirred at room temperature for 2.5 hours. The solvents were evaporated *in vacuo* and the residue redissolved in CH_2Cl_2 and purified by silica-gel column chromatography (eluted with 0-5% methanol in CH_2Cl_2) to afford **28** (3 g, 90 %) as white foam.



28

5'-*O*-dimethoxy trityl -N^4-benzoyl-3'-azido-2', 3'-dideoxycytidine 29

Benzoyl chloride (2.5 ml, 21.6 mmol) was added dropwise slowly to a solution of **28** (3 g, 5.4 mmol) in 40 ml pyridine at 0ºC. The mixture was then stirred at room temperature overnight. Then 8 ml water was added and the reaction mixture was stirred for 5 min at 0ºC, followed by addition of concentrated ammonia (8ml) at 0ºC and stirring for another 30 minute. The solvent was evaporated *in vacuo* and the residue was taken in 50ml ethyl acetate. The organic layer was washed with 10% $NaHCO_3$ solution (3 x 30 ml), water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column filtration (10% ethyl acetate in petroleum ether) to afford **29** (2.9 g, 76%) as a white foam.



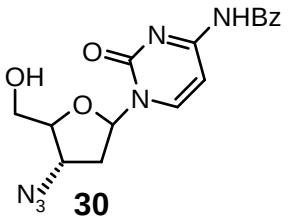
29

IR, ν(cm^{-1}) (CHCl3) 3189, 3024, 2810, 1718, 1709, 1692, 1683.

^{1}HNMR: ($CDCl_3$, 200 MHz) δ 2.5 (m, 1H), 2.6 (m, 1H), 3.8 (dd, 1H), 4.1 (m, 2H), 4.3 (m, 1H), 3.8 (s, 6H), 6.2 (t, 1H), 6.87 (m, 4H), 7.32-8.3 (m, 15H).

N^4-benzoyl-3'-azido-2', 3'-dideoxycytidine 30

5.7 g (7.1 mmol) of N-benzoyl, 5'-*O*-DMT 3'-*O*-TBDMS thymidine **29** was dissolved in 20 ml 3 % DCA-DCM solution and to it added 1 ml triethyl silane as the scavenger. The reaction mixture was stirred for 1 hour. The reaction mixture was diluted by addition of 20 ml DCM and neutralized the excess acid with 5% $NaHCO_3$ solution (2 ml). The organic layer was washed with 20 ml water and kept it over anhydrous Na_2SO_4 and evaporated to dryness. Column purification using 1% Methanol-DCM to afford 2.7 g pure **30** yield 76%.

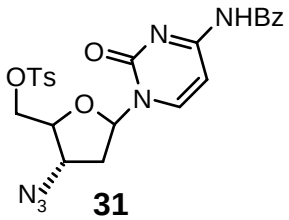


30

^{1}HNMR ($CDCl_3$, 200 MHz)δ: 2.5 (m, 1H), 2.6 (m, 1H), 3.8 (dd, 1H), 4.1 (m, 2H), 4.3 (m, 1H), 3.8 (s, 6H), 6.2 (t, 1H), 7.4-7.6 (m, 4H), 7.9-8.0 (d, 2H), 8.6 (d, 1H).

Synthesis of 3'-azido-5'-*O*-tosyl-*N*^{4}-benzoyl-2',3'-dideoxy cytosine 31

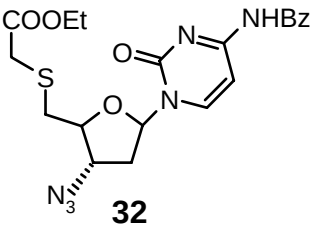
This compound was synthesized by following the same procedure starting from **30** as for the preparation of **2** and **12.**



31

^{1}H NMR: ($CDCl_3$, 200 MHz) δ: 2.33-2.47(1H, m; 2'-H), 2.49 (3H, s ;-CH_3 tosyl), 2.63-2.77 (1H, m ; 2''-H), 4.06-4.46 (4H, m; 5',5'',4',3'-H), 6.12-6.18 (1H, t; 1'-H),7.3-7.79 (11H, m; H5, H6, tosyl, Ph), 8.62(1H, b; NH).

Synthesis of ethyl *S*- *S*-5'-deoxy-5' - (3'-azido-*N*^{4}-benzoyl-2', 3'-dideoxy cytosinyl) - mercapto acetate 32



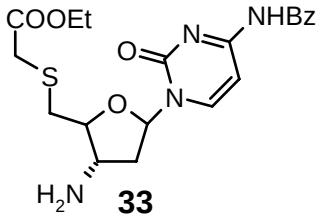
32

This compound was synthesized by following the same procedure starting from **31** as preparation of **3** and **13.**

^{1}H NMR: ($CDCl_3$, 200 MHz) δ 1.27-1.34 (3H, t ; CH_3), 2.33-2.48 (1H, m; 2'-H), 2.65-2.78 (1H, m; 2''-H), 3.06-3.08 (2H, d ; 5',5''-H), 3.25-3.43 (2H, q ; -SCH_2), 4.09-4.27 (4H, m ; 4'H,-CH_2, 3'- H), 6.10-6.16(1H, q ;1'-H), 7.38-8.15 (8H, m ; H-5, H-6, Ph, NH).

Synthesis of ethyl *S*- *S*-5'-deoxy-5' - (3'-amino-*N*^{4}-benzoyl-2', 3'-dideoxy cytosinyl)-mercapto acetate 33

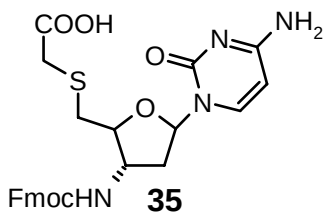
This compound was synthesized by following the same procedure starting from **32** as for the preparation of **4** and **14.**



33

^{1}H NMR: ($CDCl_3$, 200MHz) δ 1.25-1.34 (3H, t ; CH_3), 2.35-2.45 (2H, m; 2', 2''-H), 3.08-3.11(2H, d; 5',5''-H), 3.25-3.46 (3H, m; -SCH_2, 3'-H), 3.94 -4.03 (1H, m; 4'-H), 4.16-4.27 (2H, q; -CH_2), 6.15-6.2 (1H, q; 1'-H), 7.39-8.24 (7H, m; H-5, H-6, Ph).

Synthesis of *S*- *S*-5'-deoxy-5' - (3'-Fluorenylmethoxy carbonylamino-2', 3'-dideoxy-cytosinyl) - mercapto acetic acid 35



35

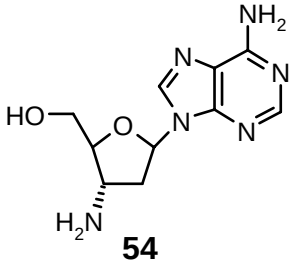
This compound **34** was synthesized by following the same procedure starting from **33** as for the preparation of **15** and was immediately used for the next reaction.

The procedure of preparation of **35** was same as the preparation of **6** and **16** starting from **34**.

^{1}H NMR: ($CDCl_3$ + 1 drop DMSO-d_6, 200MHz) δ: 2.18-2.47 (2H, m), 2.99-3.1 (2H, m), 3.22-3.42 (2H, q), 4.17-4.24 (2H, m), 4.44-4.46 (2H, m), 6.18-6.24 (1H, t), 6.66-6.69 (1H,d),7.28-7.78 (9H, m).

3'-amino 2', 3'-dideoxy adenosine 54

This nucleoside was procured from Metkinen Oy (Finland). This compound was prepared by an enzymatic trans-glycosylation process starting from 3'-amino-3'-deoxy thymidine, which was obtained via reduction of readily available 3'-azidothymidine.[47]

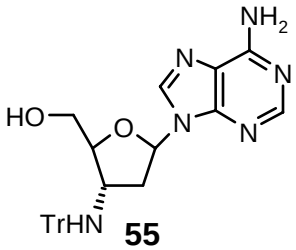


54

^{1}H NMR: D_2O δ 2.3-2.36 (m, 1H), 2.57-2.62(m, 1H), 3.59-3.66 (m, 2H), 3.75-3.78 (dd, 1H), 3.83-3.86 (m, 1H), 6.17-6.19 (q, 1H), 7.9 (s, 1H), 8.1(s, 1H).

3'-NH trityl 2', 3'-dideoxy adenosine 55

2 g of 3'-amino 2', 3'-dideoxyadenosine **54** (8.0 mmol) was suspended in 120 ml anhydrous DMF and to it added 1.2 ml (8.0 mmol) of triethyl amine. The mixture was stirred at room temperature for 10 minutes. Then 2.24 g (8.0 mmol) of trityl chloride was added in two portions and the reaction mixture was stirred for two hours. After the disappearance of the starting material (TLC analysis), the reaction mixture was poured into 120 ml of ice cold water and placed into the freezer (-10ºC) for 1 hour. The desired product **55** was filtered-off as a white solid, desiccated and used for the next reaction without any purification. Yield: 3.7 g, 94%.

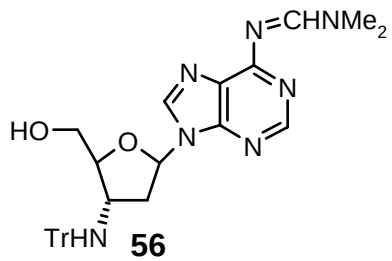


55

^{1}H NMR δ 1.7-1.8 (m, 1H), 2.45-2.6 (m, 1H), 3.3 (d, 1H), 3.68-3.76 (m, 2H), 6.1-6.2 (q, 1H), 7.1-7.3 (m, 10H), 7.5-7.55 (m, 5H), 7.76 (s, 1H), 8.0 (s, 1H), 8.2 (s, 1H).

N, N-dimethyl formamidine 3'-NH trityl 2', 3'-dideoxy adenosine 56

3.7 g (7.5 mmol) of 3'-NH trityl 2', 3'-dideoxy adenosine **55** was dissolved in 50 ml dry methanol and to it added 2 ml (15 mmol) of N, N-dimethylformamide dimethyl acetal. Stirred the reaction mixture for 6 hours. Then evaporated the methanol in vacuo and



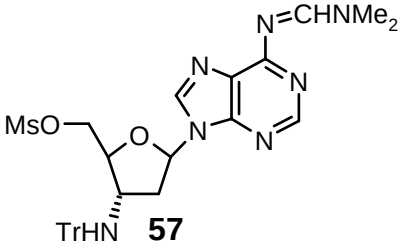
56

dissolved the residue in 150 ml DCM, washed the organic layer with water (2 x 20 ml). Kept the organic layer over anhydrous Na_2SO_4 and evaporated. Column purification using 2% MeOH-DCM to afford pure **56**, yield 3.9 g, 95%.

^{1}H NMR δ 1.7-1.8 (m, 1H), 2.54-2.7 (m, 1H), 3.2(s, 3H), 3.25 (s, 3H), 3.7 (m, 3H), 5.8-5.9 (d, 1H), 6.17-6.25 (q, 1H), 7.16-7.33 (m, 9H), 7.5 (m, 6H), 7.9 (s, 1H), 8.4 (s, 1H), 8.9 (s, 1H).

N, N-dimethyl formamidine 3'-NH trityl 5'*O*-mesyl 2', 3'-dideoxy adenosine 57

1.5 g (2.7 mmol) N, N-dimethyl formamidine 3'-NH trityl 2', 3'-dideoxy adenosine **56** was dissolved in 20 ml dry DCM and to it added 0.8 ml(5.5 mmol) of TEA, cooled the mixture to 0ºC. Then added mesyl chloride 0.26 ml (3.3 mmol) dissolved in 5 ml of DCM drop wise by a syringe. After 10 minutes TLC showed complete conversion of starting material and the reaction was quenched by addition of 2 ml water, the mixture was diluted with the addition of 20 ml more DCM. The organic layer was separated and washed with 5% NaHCO3 (2 x 20 ml). The organic layer was kept over anhydrous Na_2SO_4 and evaporated to dryness to give **57,** 1.6 g, yield 94%. The product **57** was kept for desiccation and used for the next reaction without any purification.

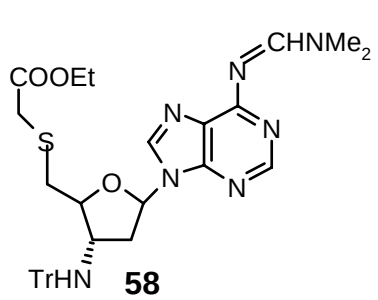


57

^{1}H NMR δ 1.76-1.82 (m, 2H), 2.77 (s, 3H), 3.2 (s, 3H), 3.24 (s, 3H), 3.62 (m, 1H), 3.94-4.0 (m, 1H), 4.24-4.4 (m, 2H), 6.21-6.27 (t, 1H), 7.16-7.34 (m, 9H), 7.5-7.55 (m, 6H), 7.84 (s, 1H), 8.5 (s, 1H), 9.0 (s, 1H).

Ethyl-S-5'-deoxy-5'- [N, N-dimethyl formamidine 3'-NH trityl 2', 3' dideoxyadenosinyl]-mercapto acetate 58

NaH (60 % solution in hexane) 0.105 g (2.63 mmol) was taken of in 3 ml of anhydrous DMF and to it was added 0.66 ml (6 mmol) ethyl mercapto acetate. The mixture was stirred for 15 minutes and 1.5 g (2.4 mmol) of **57** dissolved in 2 ml of DMF was then



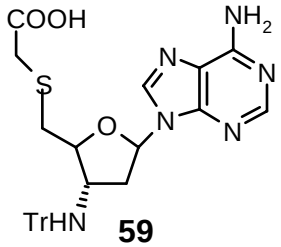
58

added. After 1 hour the mixture was concentrated *in vacuo* and redissolved in 50 ml ethyl acetate. The organic layer was washed with water (3 x 20 ml) and saturated NaCl (2 x 15 ml). The organic layer was dried over Na_2SO_4, evaporated to dryness and purified by column chromatography (1% MeOH in DCM) to afford 1.1 g (71%) of **58**. 649.6

^{1}H NMR δ 1.25-1.29 (t, 3H), 1.8 (m, 2H), 3.2 (s, 3H), 3.25(s, 3H), 3.6 (m, 1H), 4.0-4.34 (m, 5H), 6.21-6.27 (t, 1H), 7.17-7.34 (m, 9H), 7.51-7.55 (m, 6H), 7.84 (s, 1H), 8.5 (s, 1H), 9.0 (s, 1H).

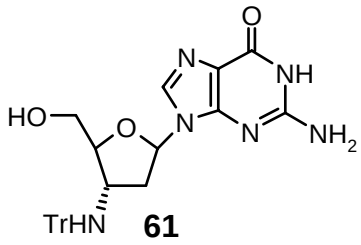
S-5'-deoxy-5'-[3'-NH trityl 2', 3'-dideoxyadenosinyl]-mercapto acetic acid 59

The ethyl-S-(3'-NH trityl 2',3'-dideoxyadenosyl)-mercapto acetate **58** (1 g, 1.54 mmol) was dissolved in 4 ml methanol and to it was added 2 ml of 1M LiOH solution. The mixture was stirred at room temp for 1 h. The excess LiOH was neutralized by Dowex-50H$^+$ resin which were then filtered off. Concentration of the filtrate *in vacuo* afforded **59** as a white solid. The solid obtained was triturated repeatedly with diethyl ether to remove the impurities and kept for desiccation, yield 0.8g, 92%.



59

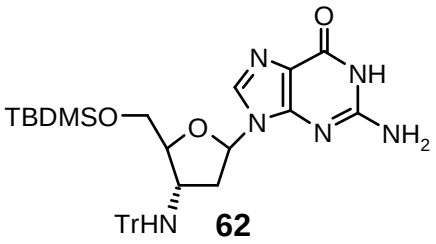
3'-NH trityl 2', 3'-dideoxy guanosine 61 2 g of 3'-amino 2', 3'-dideoxyguanosine **60** (8.0 mmol) was suspended in 120 ml anhydrous DMF and to it added 1.2 ml (8.0 mmol) of triethyl amine. The mixture was stirred at 60ºC for 30 minutes. Then 2.24g (8.0 mmol) of trityl chloride was added in two portions and the reaction mixture was stirred for another two hours at room temperature. After the disappearance of the starting material (TLC analysis), the reaction mixture was poured into 120 ml of ice cold water and placed into the freezer (-10ºC) for 1 hour. The desired product **61** was filtered-off as a white solid, desiccated and used for the next reaction without any purification. Yield: 3.7g, 94%.



61

3'-NH trityl 5'-*O*-TBDMS 2', 3'-dideoxy guanosine 62

3.7 g (7.3 mmol) 3'-NH trityl 2', 3'-dideoxy guanosine **61**, Imidazole (1.5 g, 22 mmol) and TBDMS-Cl (1.3 g, 8.7 mmol) were dissolved in 20 ml anhydrous DMF and stirred for six hours. The DMF was removed *in vacuo* and redissolved the mixture in 200 ml ethyl acetate. The organic layer was washed with water (3 x 50 ml) and then brine (2 x 30 ml).The organic layer was kept over anhydrous sodium sulphate and evaporated to dryness *in vacuo*. Then column purification was done using 3% MeOH-DCM as the solvent system to give 4.2 g pure **62** yield 93 %.

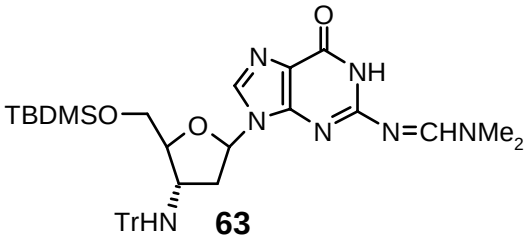


62

^{1}H NMR: δ 0.06(s, 6H), 0.94 (s, 9H), 1.68-1.77 (m, 2H), 3.6 (m, 1H), 3.74-3.92 (m, 3H), 6.06-6.12 (t, 1H), 7.26-7.53 (m, 9H), 7.63-7.66 (m, 6H).

N, N-dimethyl formamidine 3'-NH trityl 5'-*O*-TBDMS 2', 3'-dideoxy guanosine 63

4.15 g (6.7 mmol) of 3'-NH trityl 5'-*O*-TBDMS 2', 3'-dideoxy guanosine **62** was dissolved in 50 ml dry methanol and to it added TEA 2.8 ml (20 mmol) and 1.8 ml N, N-dimethylformamide dimethyl acetal (13.3 mmol). The reaction mixture was stirred for 6 hours. Then evaporated the methanol in *vacuo* and dissolved the residue in 150 ml DCM, washed the organic layer with water (2 x 20 ml). Kept the organic layer over anhydrous Na_2SO_4 and evaporated. Column purification using 2% MeOH-DCM to afford pure **63**, yield 4.3 g, 95%.



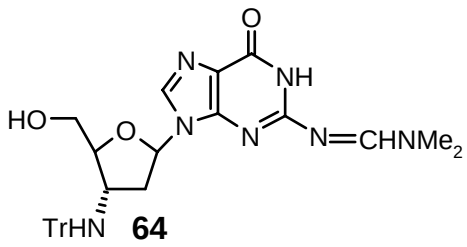
63

^{1}H NMR $CDCl_3$ δ 0.08(s, 6H), 0.94 (s, 9H), 1.73-1.86 (m, 2H), 3.21 (s, 3H), 3.27 (s, 3H), 3.54-3.61 (m, 1H), 3.76-3.92 (m, 2H), 3.96-4.02 (m, 1H), 6.21-6.27 (t, 1H), 7.28-7.45 (m, 9H), 7.65-7.69 (m, 6H), 8.67 (s, 1H), 9.24 (s, 1H).

N, N-dimethyl formamidine 3'-NH trityl 2', 3'-dideoxy guanosine 64

4.25 g (6.3 mmol) N, N-dimethyl formamidine 3'-NH trityl 5'-*O*-TBDMS 2', 3'-dideoxy guanosine **63** was dissolved in 15 ml anhydrous THF and to it added 3 g TBAF (9.4 mmol). The reaction mixture was stirred for 1h. The solvent was evaporated *in vacuo* and

redissolved the residue in 100 ml ethyl acetate. The organic layer was washed with water (2 x 30 ml), followed by brine (1 x 20 ml). The organic layer was kept over anhydrous Na_2SO_4 and evaporated *in vacuo*. Column purification using 3% Methanol- DCM to afford 3.3 g **64** Yield 93 %.

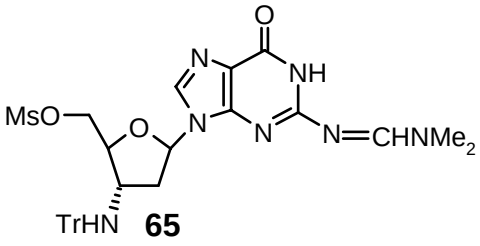


64

^{1}H NMR $CDCl_3$ δ 2.04-2.17 (m, 2H), 3.02 (s, 3H), 3.07 (s, 3H), 3.6 (m, 1H), 3.7-3.8 (m, 3H), 6.03-6.1 (t, 1H), 7.2-7.32 (m, 9H), 7.52-7.59 (m, 6H), 8.4 (s, 1H), 8.97 (s, 1H).

N, N-dimethyl formamidine 3'-NH trityl 5'-*O*-mesyl 2', 3'-dideoxy guanosine 65

1.5 g (2.7 mmol) N, N-dimethyl formamidine 3'-NH trityl 2', 3'-dideoxy guanosine **64** was dissolved in 20 ml dry DCM and to it added 0.8 ml(5.5 mmol) of TEA, cooled the mixture to 0ºC. Then added mesyl chloride 0.26 ml (3.3 mmol) dissolved in 5 ml of DCM drop wise by a syringe. After 10 minutes TLC showed complete conversion of starting material and the reaction was quenched by addition of 2 ml water, the mixture was diluted with the addition of 20 ml more DCM. The organic layer was separated and washed with 5% $NaHCO_3$ (2 x 20 ml). Kept the organic layer over anhydrous Na_2SO_4 and evaporated to dryness to give **65,** 1.6 g, yield 94%.The product **65** was kept for desiccation and used for the next reaction without any purification.



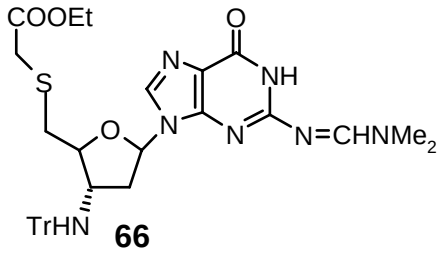
65

^{1}H NMR $CDCl_3$ δ 1.96-2.04 (s, 2H), 2.78 (s, 3H), 3.04 (s, 3H), 3.07 (s, 3H), 3.55 (m, 1H), 3.95 (m, 1H), 4.3-4.43 (m, 2H), 6.05 (t, 1H), 7.16-7.33 (m, 9H), 7.51-7.55 (m, 6H), 8.47 (s, 1H), 9.4 (s, 1H).

Ethyl-S-5'-deoxy-5'- [N, N-dimethyl formamidine 3'-NH trityl 2', 3'-dideoxy guanosinyl]-mercapto acetate 66

NaH (60% solution in hexane) 0.105 g (2.63 mmol) was taken of in 3 ml of anhydrous DMF and to it was added 0.66 ml (6 mmol) ethyl mercapto acetate. The mixture was stirred for 15 minutes and 1.5 g (2.4 mmol) of **65** dissolved in 2 ml of DMF was then

added. After 1hour the mixture was concentrated *in vacuo* and redissolved in 50 ml ethyl acetate. Washed the organic layer with water (3 x 20 ml) and saturated NaCl (2 x 15 ml). The organic layer was dried over Na_2SO_4, evaporated to dryness and purified by column chromatography (1% MeOH in DCM) to afford 1.1 g (71%) of **66.**

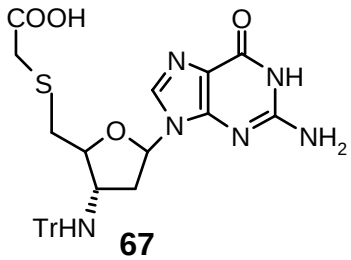


66

^{1}H NMR $CDCl_3$ δ 1.18-1.25 (t, 3H), 1.4-1.6 (m, 2H), 2.9 (s, 3H), 2.95 (s, 3H), 3.06-3.13 (m, 6H), 3.45 (m, 1H), 4.1 (q, 2H), 6.0 (t, 1H), 7.15-7.32 (m, 9H), 7.52-7.55 (m, 6H), 8.0 (s, 1H), 8.5 (s,1H).

_S_-5'-deoxy-5'-[3'-NH trityl 2', 3'-dideoxyguanosinyl]-mercapto acetic acid 67

The ethyl-S-(3'-NH trityl 2',3'-dideoxyguanosyl)-mercapto acetate **66,** (1 g, 1.54 mmol) was dissolved in 4 ml methanol and to it was added 2ml of 1M LiOH solution. The mixture was stirred at room temp for 1 h. The excess LiOH was neutralized by Dowex-50H$^+$ resin which were then filtered off. Concentration of the filtrate *in vacuo* afforded **6** as a white solid. The solid obtained was triturated repeatedly with diethyl ether to remove the no polar impurities and kept for desiccation, yield 0.8 g, 92%.



67

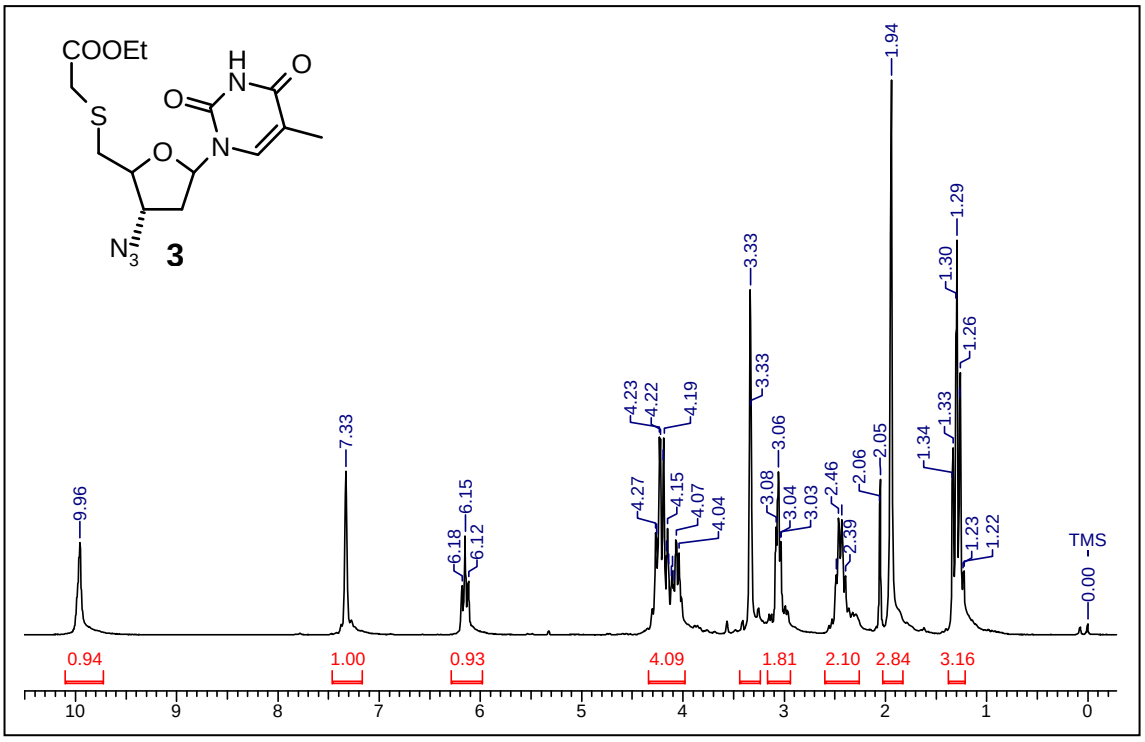
2.1.9 Appendix

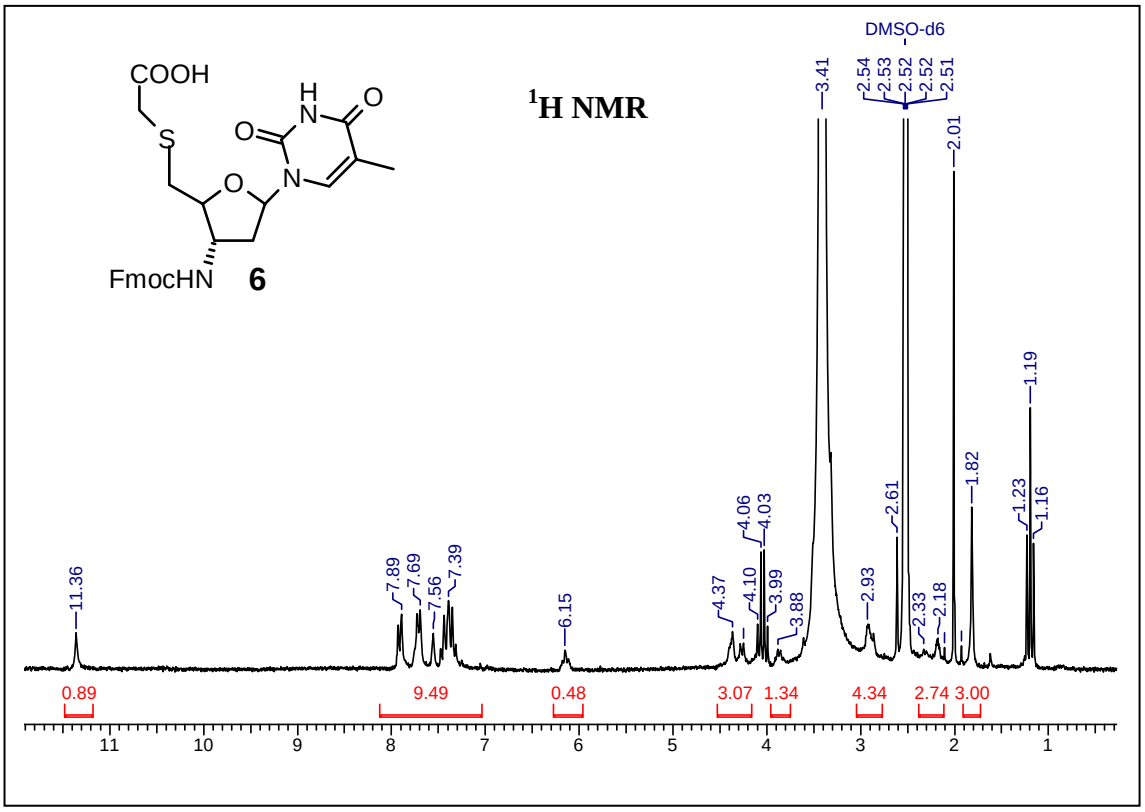
^{1}H NMR of compound **3, 6, 11, 13, 14, 16, 29, 30, 31, 32, 33, 35, 56, 57, 58 & 66**

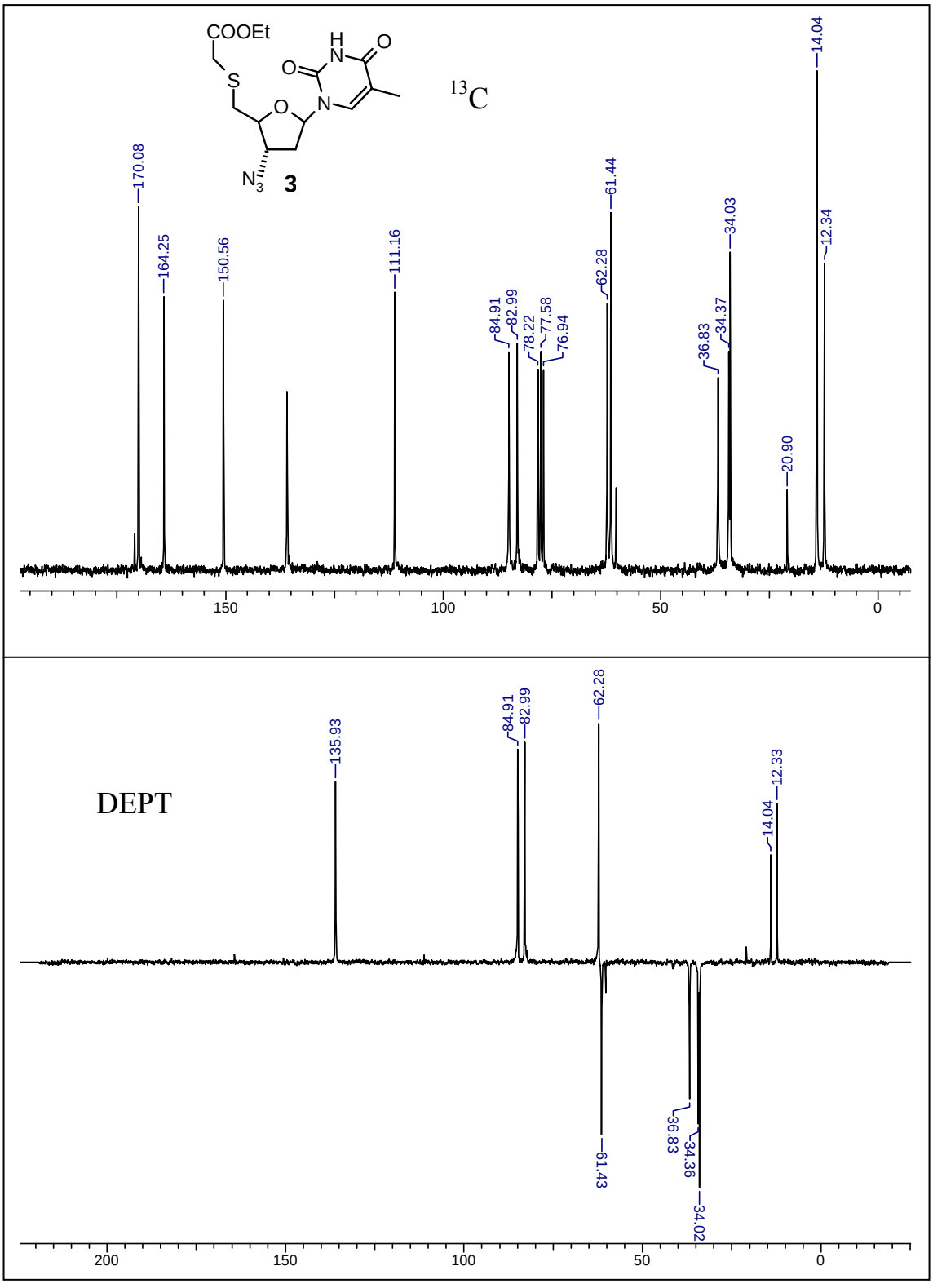
^{13}C & DEPT NMR of compound **3, 6, 13, 14 & 16**

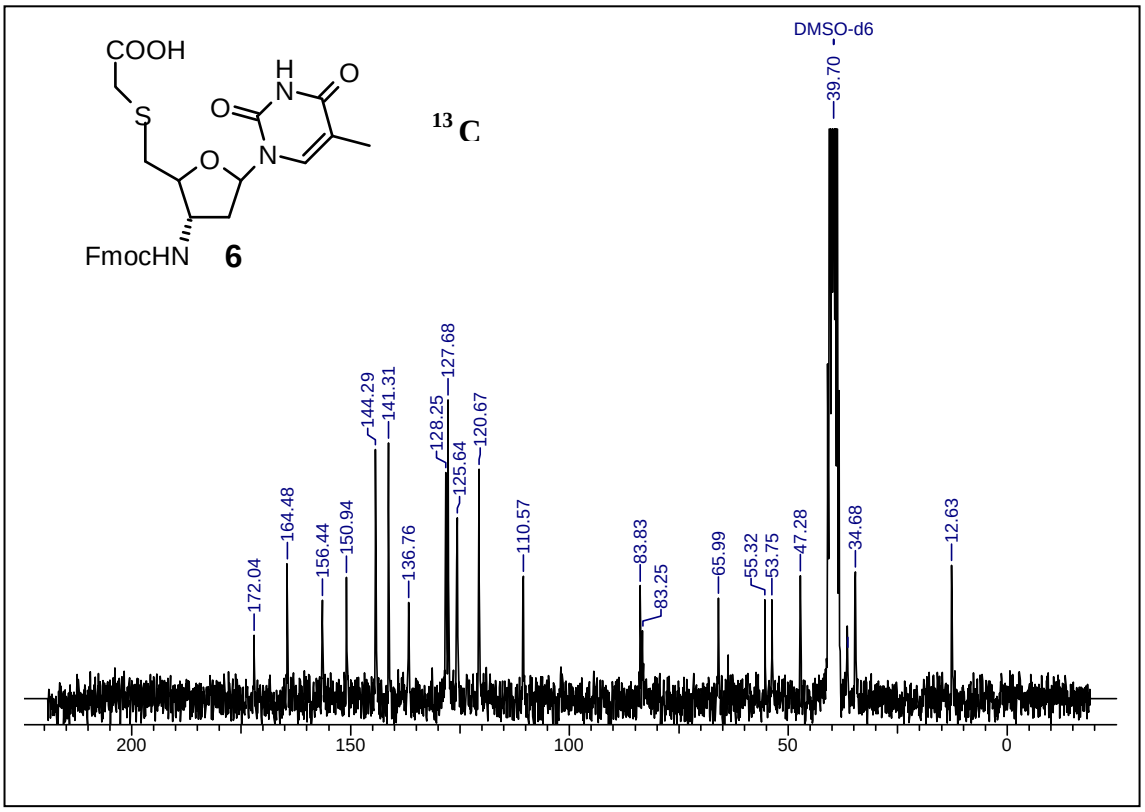
Mass spectra of compound **3, 4, 13, 6, 16, 38, 39, 40, 41, 42, 42 & 43**

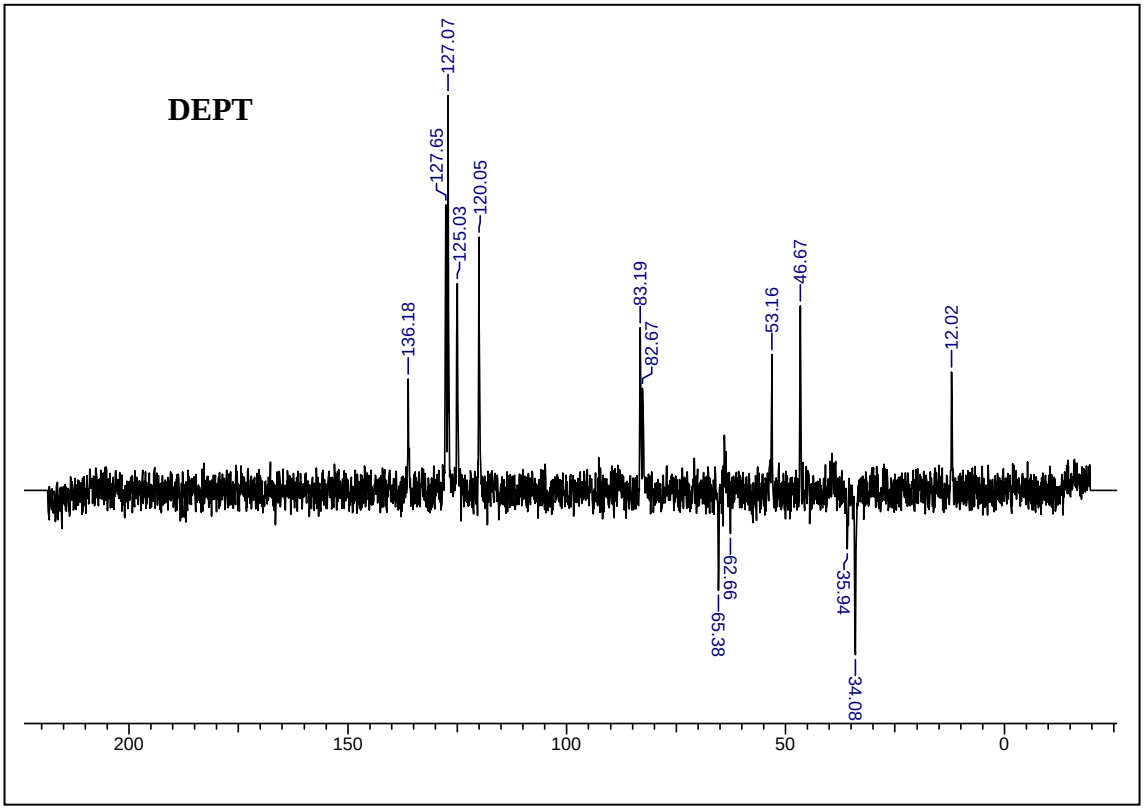
RP-HPLC of compound **38-45**

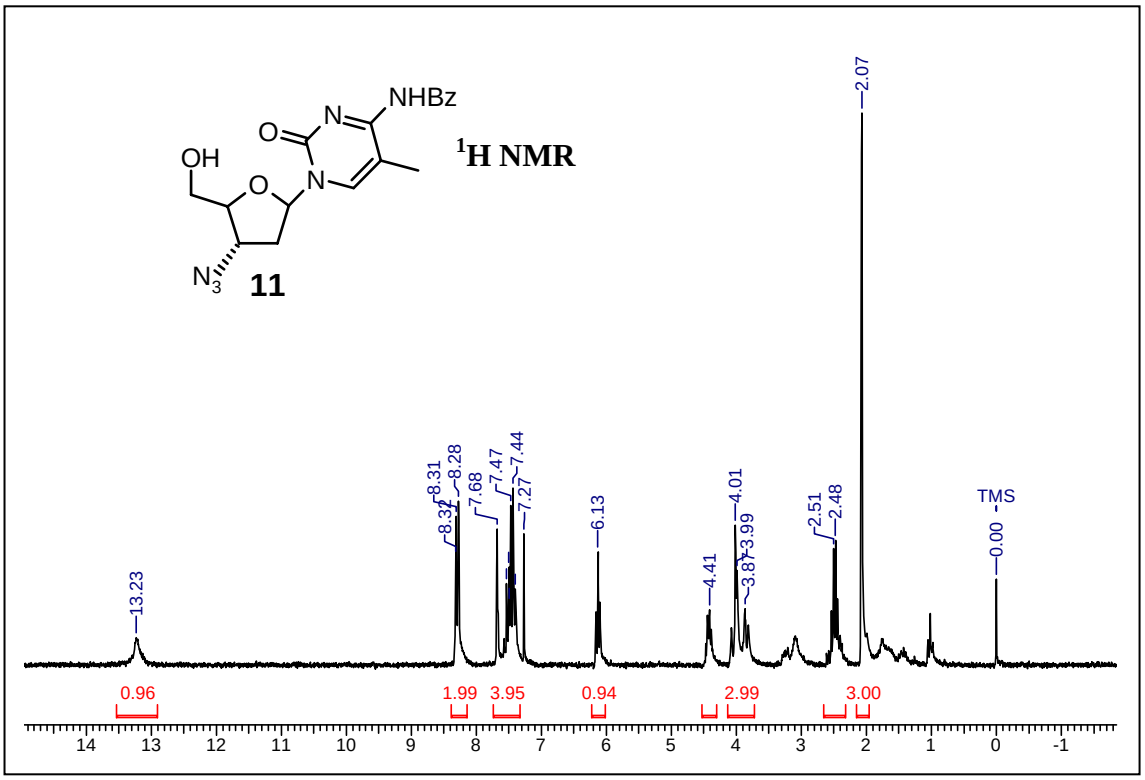

COOEt
S
O
N
H
O
O
N
N3
3
9.96
7.33
6.18
6.15
6.12
4.27
4.23
4.22
4.19
4.15
4.07
4.04
3.33
3.33
3.08
3.06
3.04
3.03
2.46
2.39
2.06
2.05
1.94
1.34
1.33
1.30
1.29
1.26
1.23
1.22
TMS
0.00
0.94
1.00
0.93
4.09
1.81
2.10
2.84
3.16
10
9
8
7
6
5
4
3
2
1
0

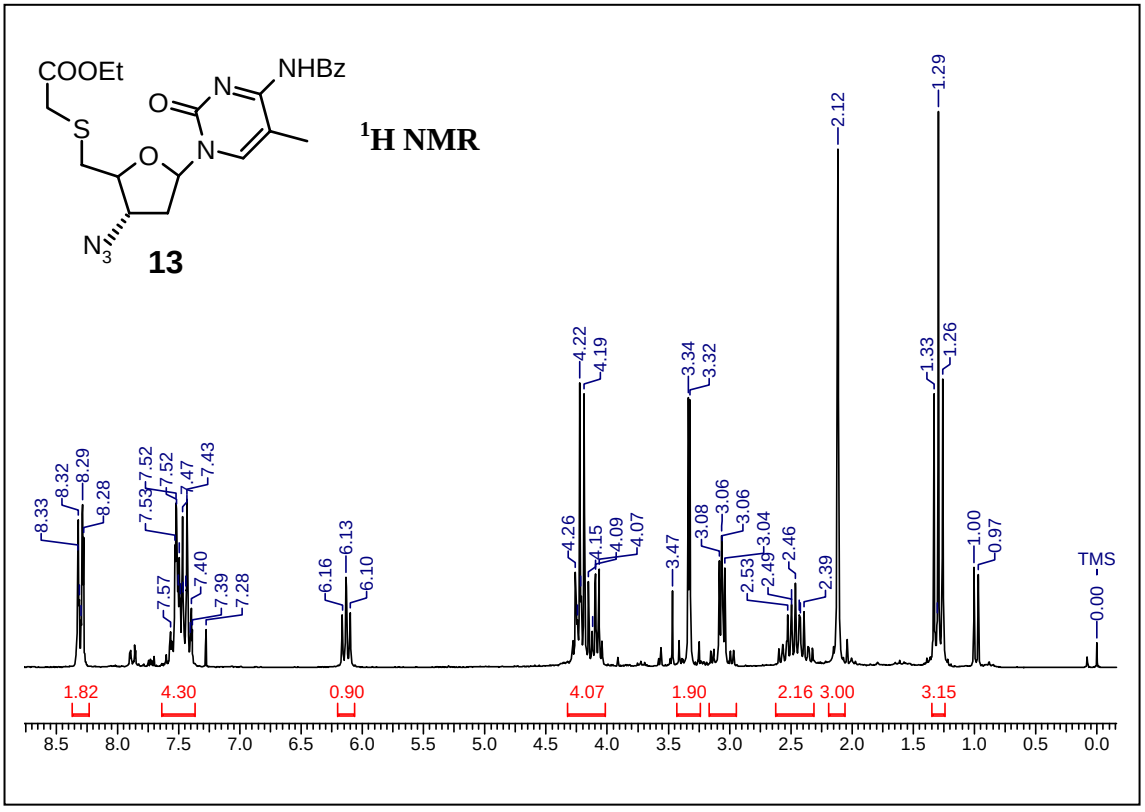

COOH
S
O
N
H
O
O
N
FmocHN
6
1H NMR
DMSO-d6
11.36
7.89
7.69
7.56
7.39
6.15
4.37
4.10
4.06
4.03
3.99
3.88
3.41
2.93
2.61
2.54
2.53
2.52
2.52
2.51
2.33
2.18
2.01
1.82
1.23
1.19
1.16
0.89
9.49
0.48
3.07
1.34
4.34
2.74
3.00
11
10
9
8
7
6
5
4
3
2
1

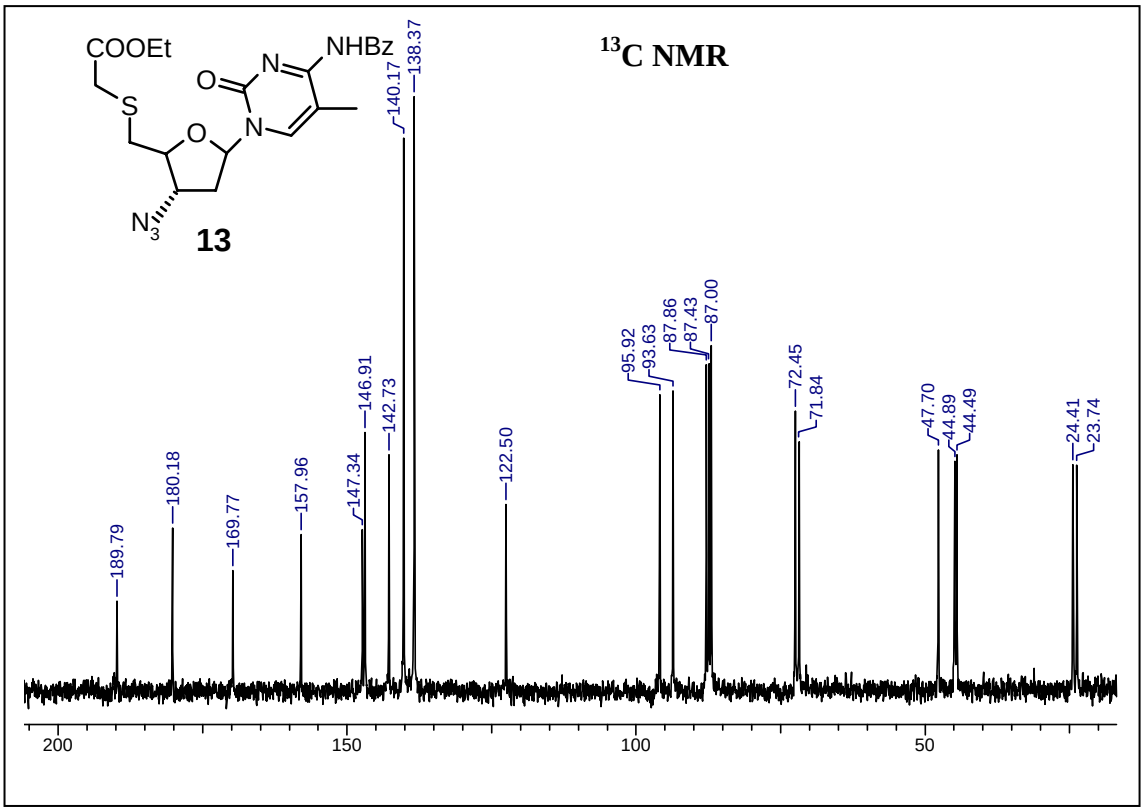

COOEt
S
O
H
N
O
O
N
O
N3
3
13C
170.08
164.25
150.56
111.16
84.91
82.99
78.22
77.58
76.94
62.28
61.44
36.83
34.37
34.03
20.90
14.04
12.34
150
100
50
0
DEPT
135.93
84.91
82.99
62.28
14.04
12.33
61.43
36.83
34.39
34.02
200
150
100
50
0

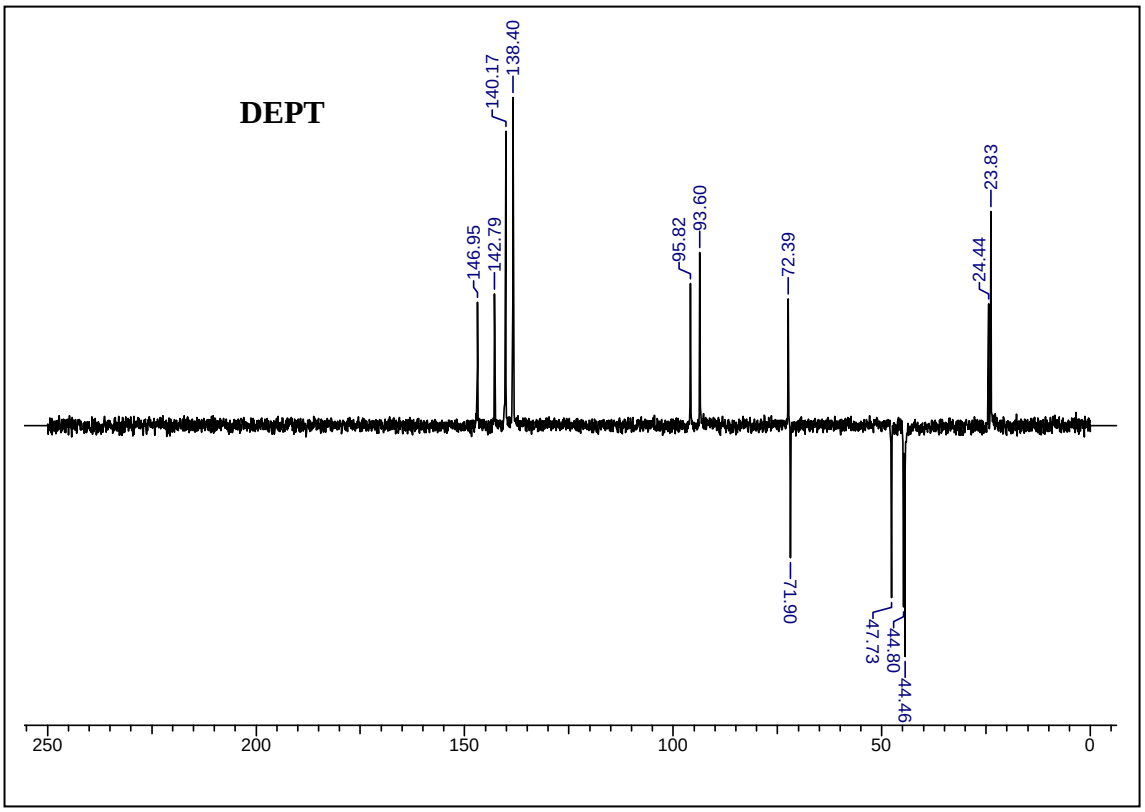

COOH
S
O
H
N
O
O
N
FmocHN
6
^{13}C
DMSO-d6
39.70
172.04
164.48
156.44
150.94
144.29
141.31
136.76
128.25
127.68
125.64
120.67
110.57
83.83
83.25
65.99
55.32
53.75
47.28
34.68
12.63
200
150
100
50
0

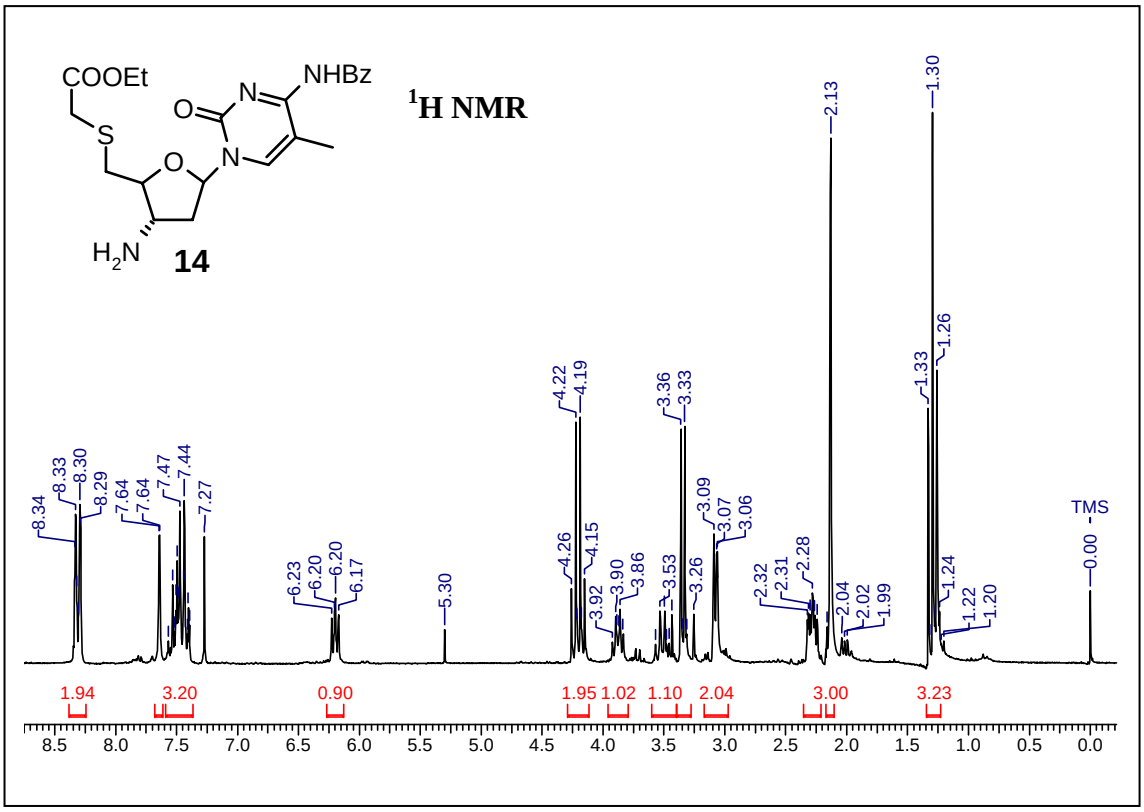

DEPT
127.07
127.65
125.03
120.05
136.18
83.19
82.67
53.16
46.67
12.02
62.66
63.38
35.94
34.08
200
150
100
50
0

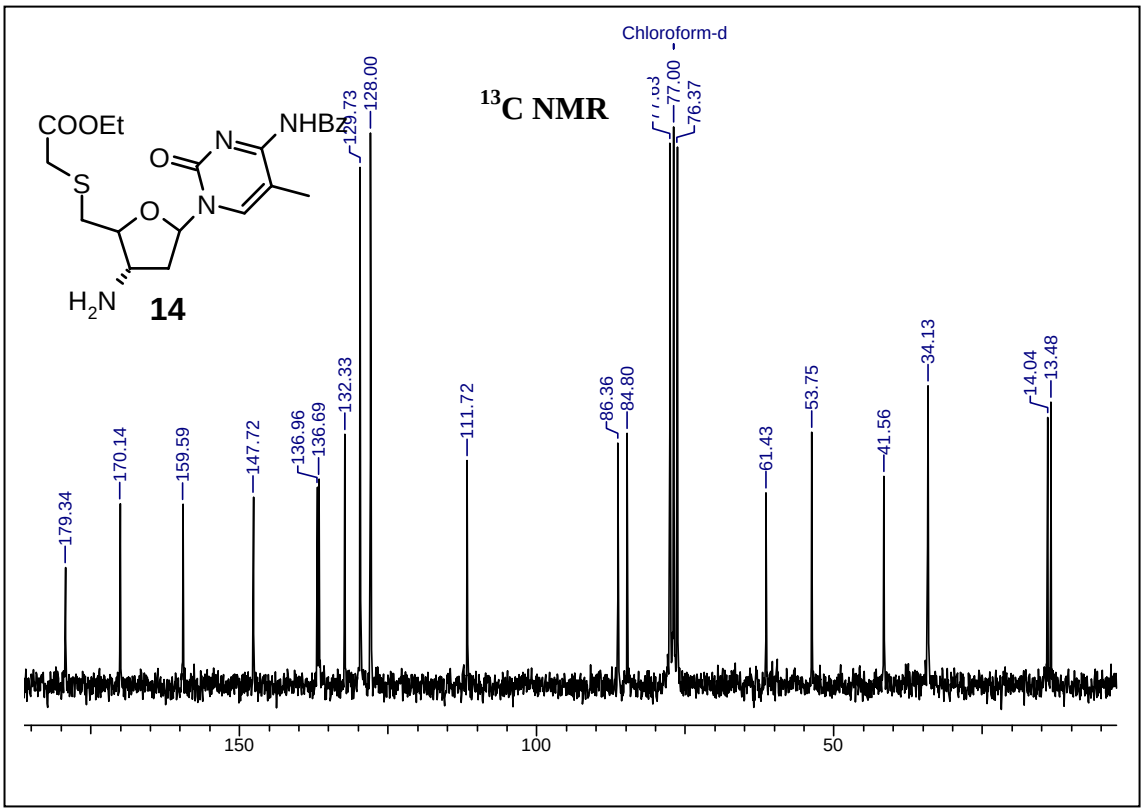

NHBz
OH
^{1}H NMR
11
2.07
13.23
8.31
8.28
7.68
7.47
7.44
7.27
6.13
4.41
4.01
3.99
3.87
2.51
2.48
TMS
0.00
0.96
1.99
3.95
0.94
2.99
3.00

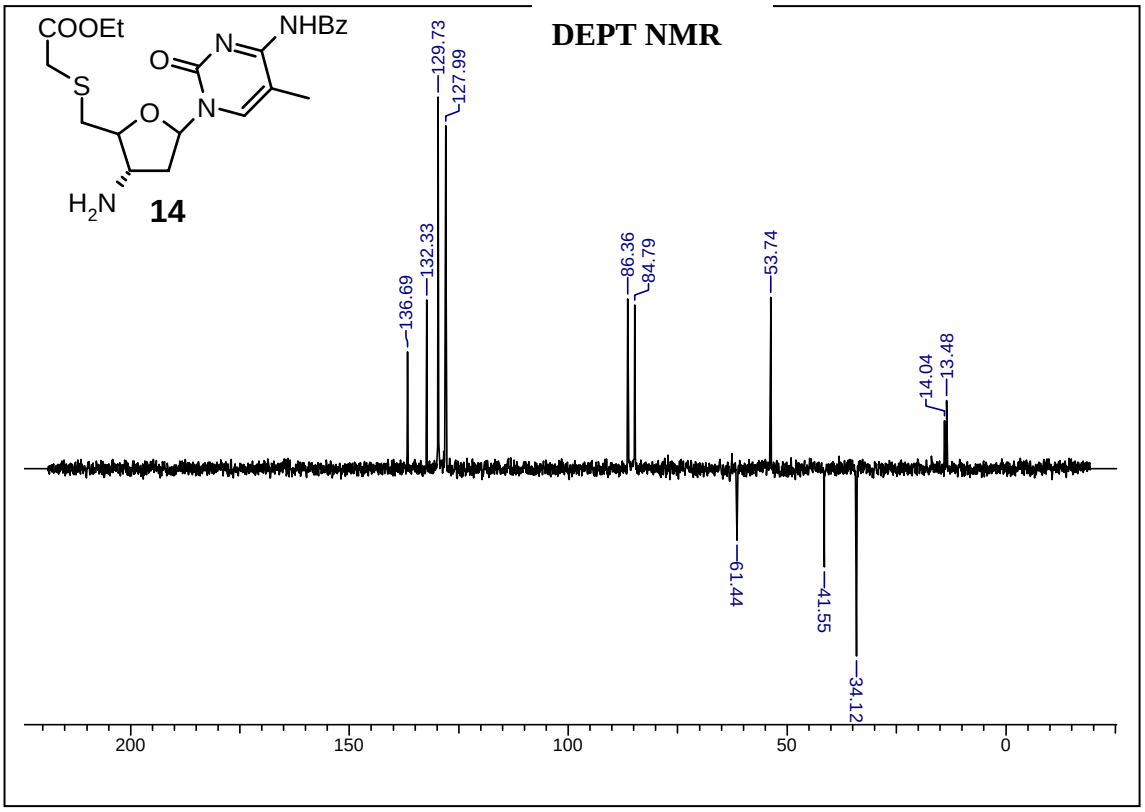

COOEt
NHBz
^{1}H NMR
13
2.12
1.29
4.22
4.19
3.34
3.32
1.33
1.26
8.33
8.32
8.29
8.28
7.53
7.52
7.52
7.47
7.43
7.57
7.40
7.39
7.28
6.16
6.13
6.10
4.26
4.15
4.09
4.07
3.47
3.08
3.06
3.06
3.04
2.53
2.49
2.46
2.39
1.00
0.97
TMS
0.00
1.82
4.30
0.90
4.07
1.90
2.16
3.00
3.15

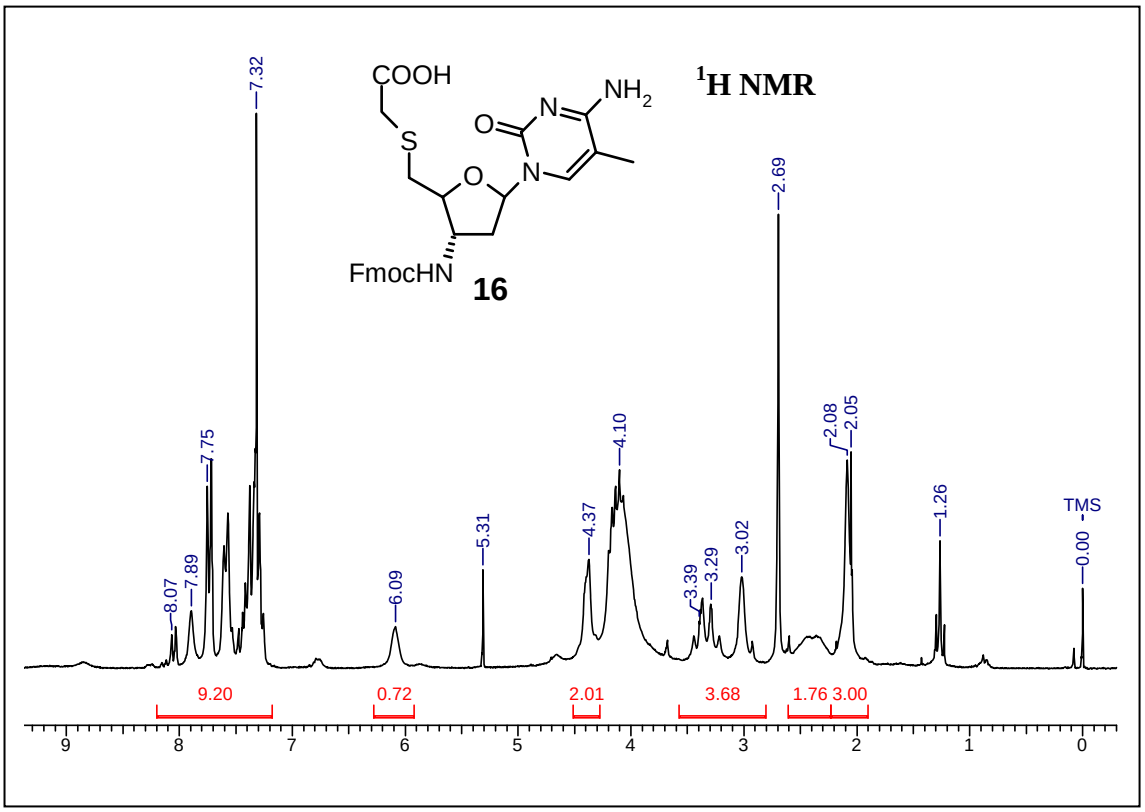

COOEt
NHBz
O
N
S
O
N
N3
13
13C NMR
189.79
180.18
169.77
157.96
147.34
146.91
142.73
140.17
138.37
122.50
95.92
93.63
87.86
87.43
87.00
72.45
71.84
47.70
44.89
44.49
24.41
23.74
200
150
100
50

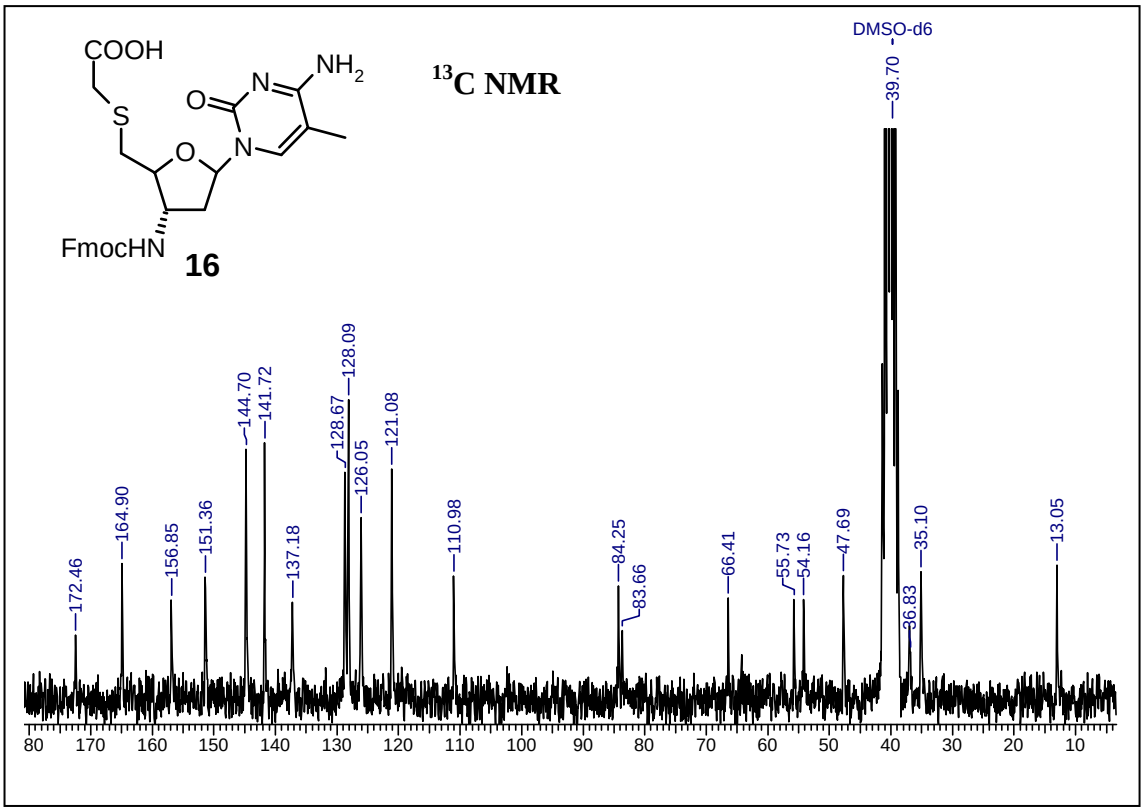

DEPT
146.95
142.79
140.17
138.40
95.82
93.60
72.39
24.44
23.83
71.90
47.73
44.80
44.49
250
200
150
100
50
0

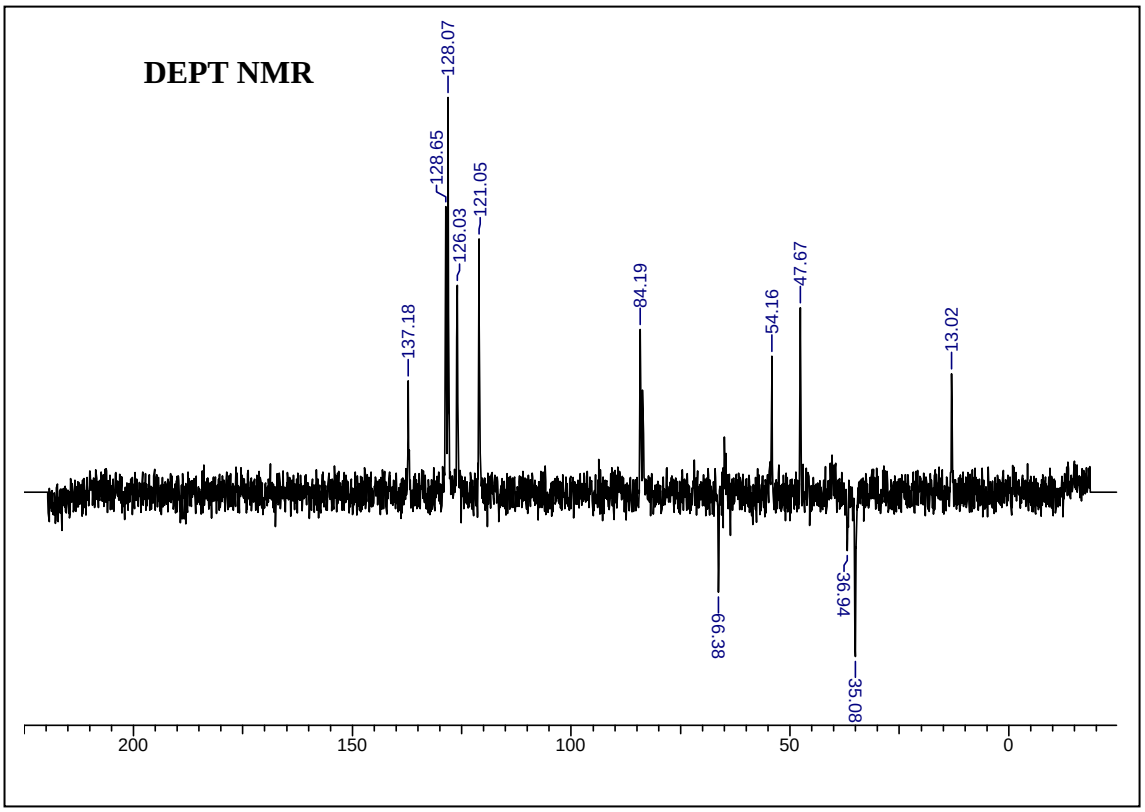

COOEt
NHBz
O
N
S
O
N
H_2N
14
^{1}H NMR
8.34
8.33
8.30
8.29
7.64
7.64
7.47
7.44
7.27
6.23
6.20
6.20
6.17
5.30
4.26
4.22
4.19
4.15
3.92
3.90
3.86
3.53
3.36
3.33
3.26
3.09
3.07
3.06
2.32
2.31
2.28
2.13
2.04
2.02
1.99
1.33
1.30
1.26
1.24
1.22
1.20
TMS
0.00
1.94
3.20
0.90
1.95
1.02
1.10
2.04
3.00
3.23
8.5
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

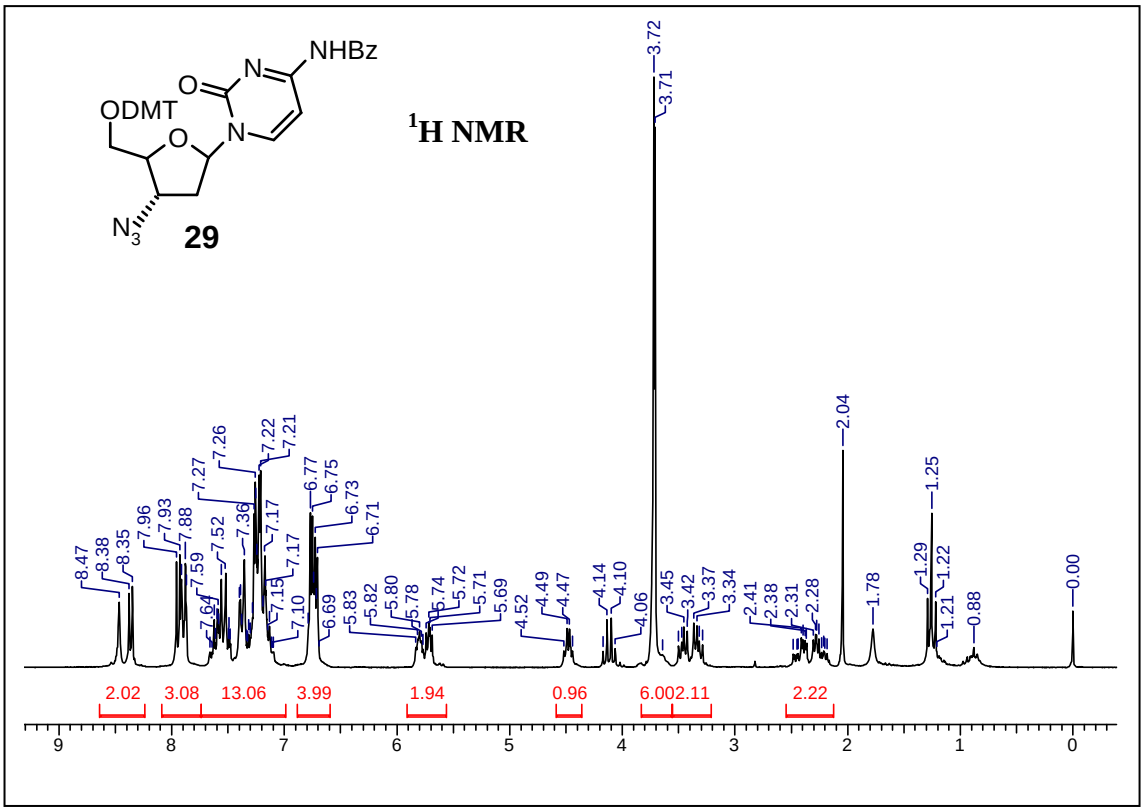

COOEt
NHBz
O
N
S
O
N
H_2N
14
^{13}C NMR
Chloroform-d
179.34
170.14
159.59
147.72
136.96
136.69
132.33
129.73
128.00
111.72
86.36
84.80
77.00
76.37
61.43
53.75
41.56
34.13
14.04
13.48
150
100
50

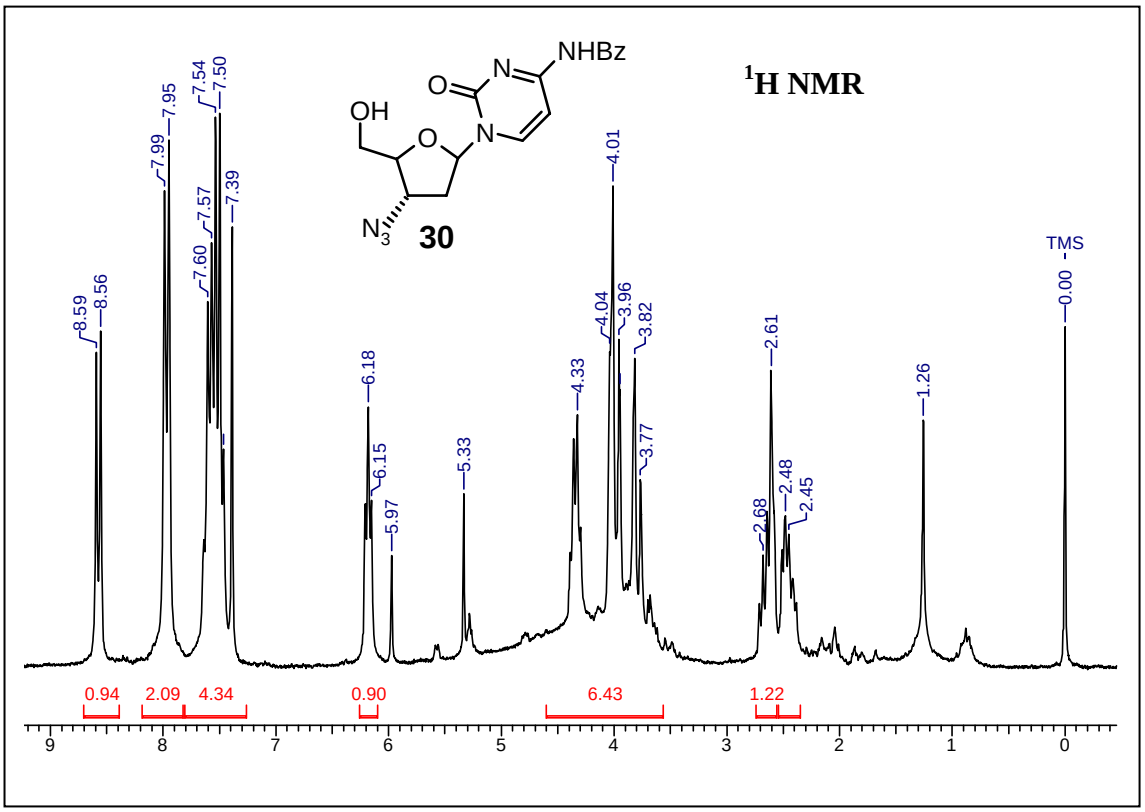

DEPT NMR
COOEt
NHBz
O
N
S
O
N
H_2N
14
129.73
127.99
136.69
132.33
86.36
84.79
53.74
14.04
13.48
-9.14
-41.55
-84.12
200
150
100
50
0

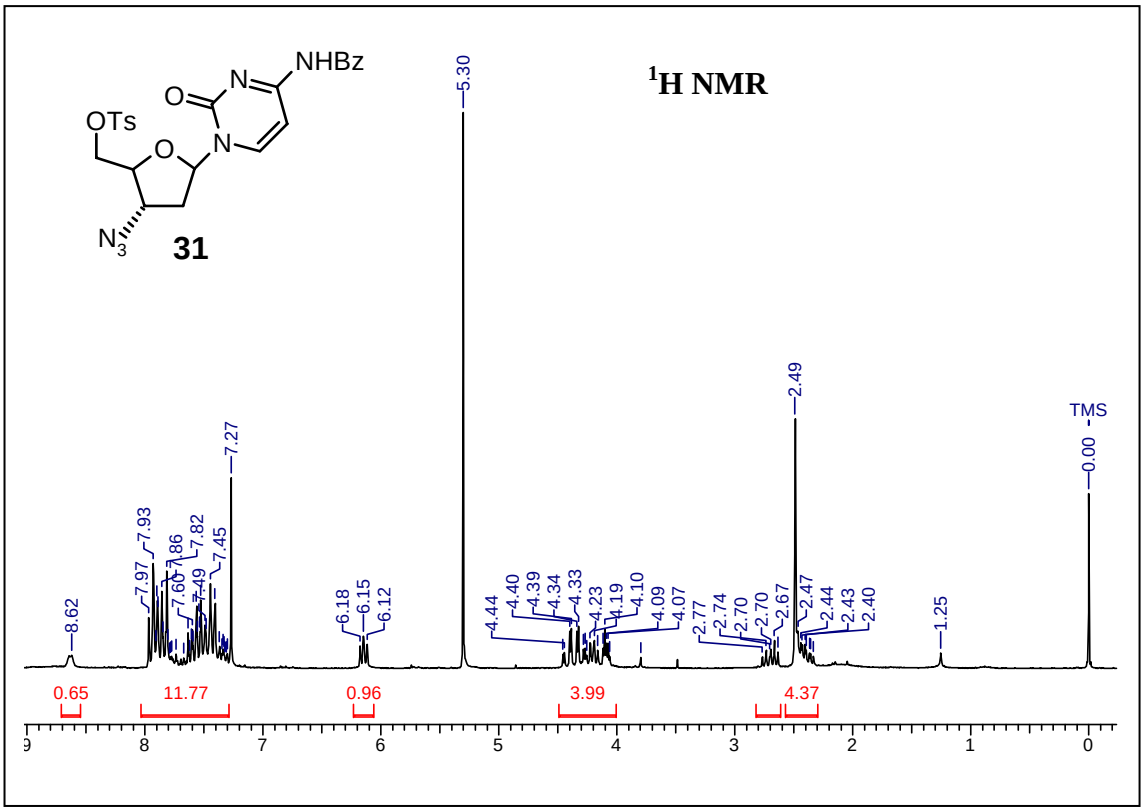

^{1}H NMR
COOH
NH_2
O
N
S
O
N
FmocHN
16
7.32
2.69
7.75
4.10
2.08
2.05
8.07
7.89
6.09
5.31
4.37
3.39
3.29
3.02
1.26
TMS
0.00
9.20
0.72
2.01
3.68
1.76
3.00
9
8
7
6
5
4
3
2
1
0

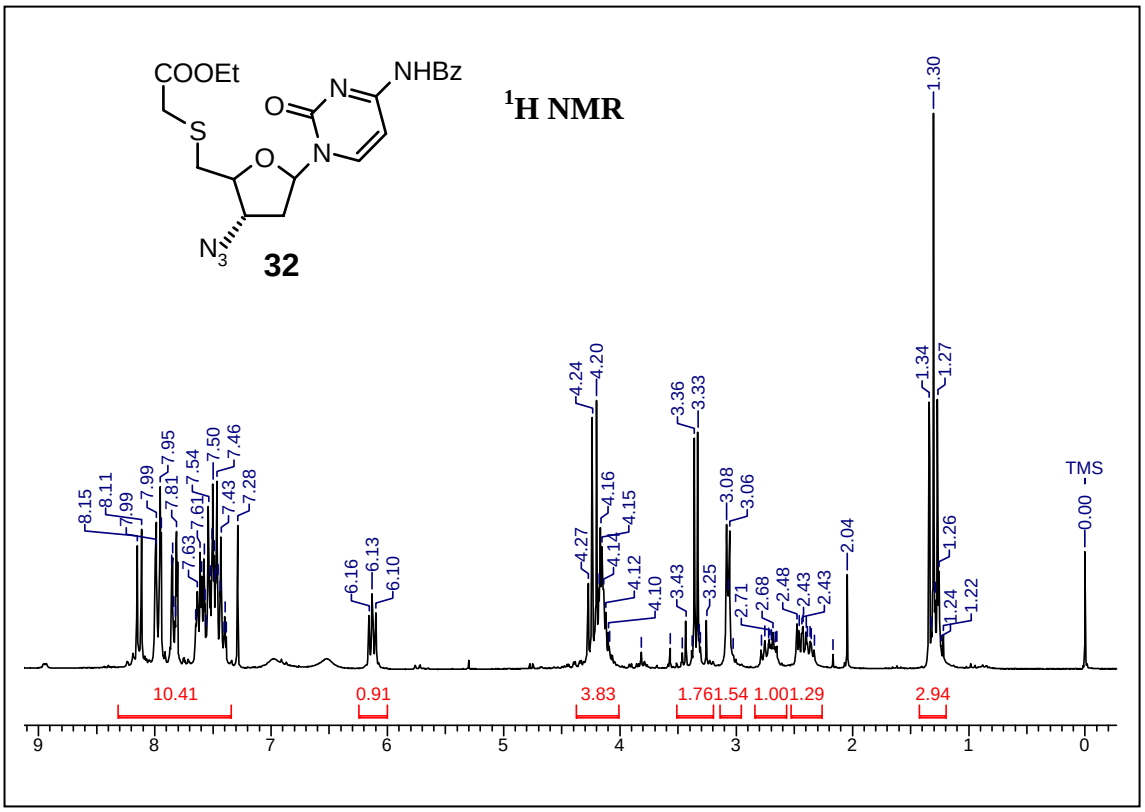

COOH
S
O
N
NH2
O
N
FmocHN
16
13C NMR
DMSO-d6
39.70
172.46
164.90
156.85
151.36
144.70
141.72
137.18
128.67
128.09
126.05
121.08
110.98
84.25
83.66
66.41
55.73
54.16
47.69
36.83
35.10
13.05
80 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10

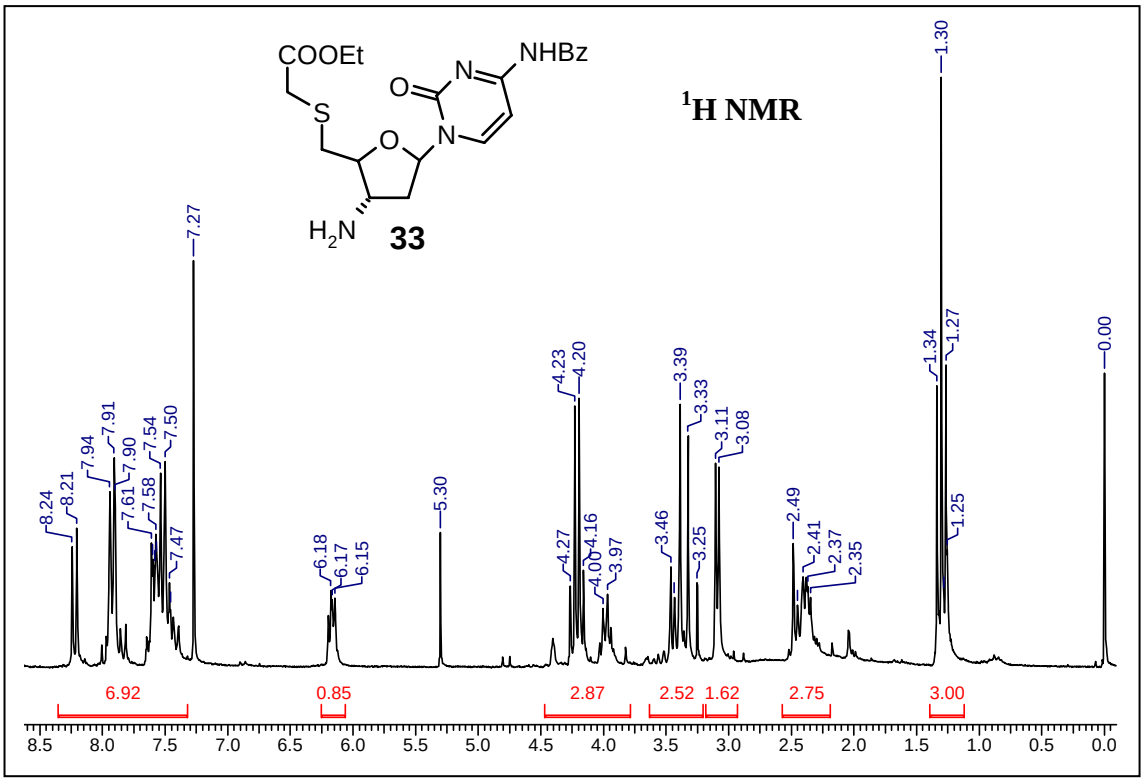

DEPT NMR
128.07
128.65
126.03
121.05
137.18
84.19
54.16
47.67
13.02
66.38
36.94
35.08
200 150 100 50 0

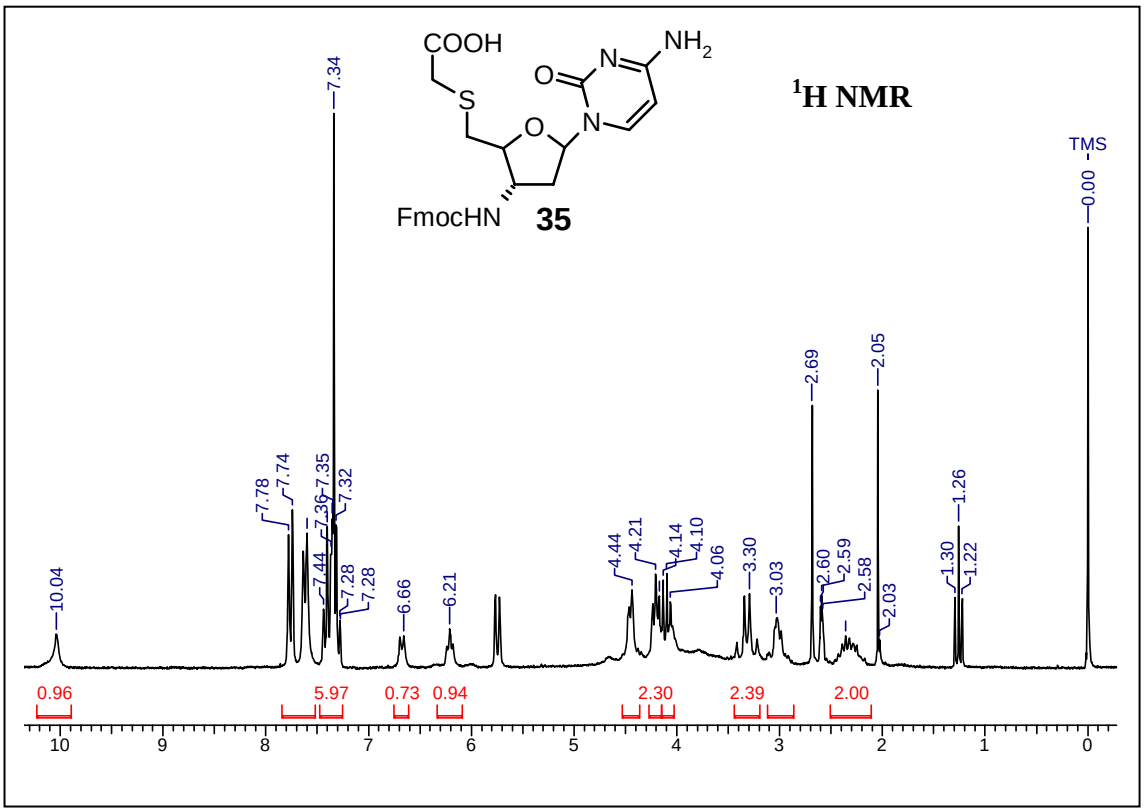

NHBz
N
O
ODMT
O
N
N3
29
1H NMR
3.72
3.71
2.04
1.25
1.29
1.22
1.21
0.88
1.78
0.00
2.02
3.08
13.06
3.99
1.94
0.96
6.002.11
2.22

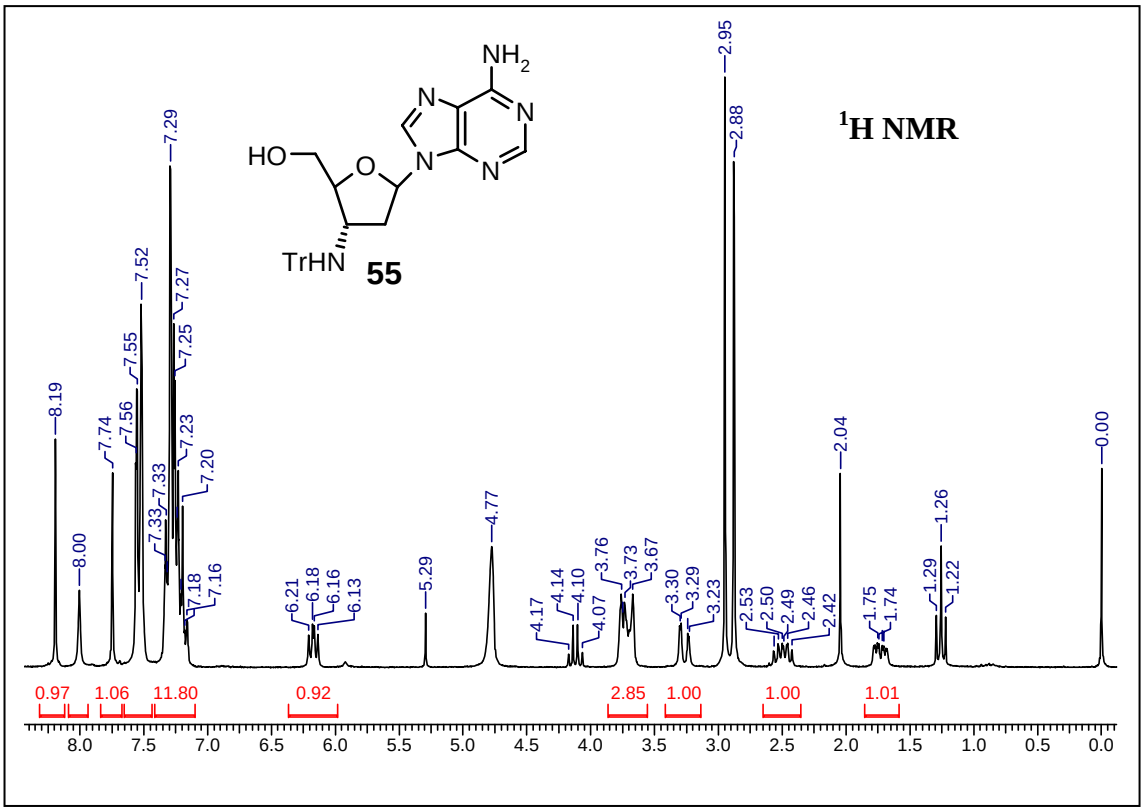

NHBz
N
O
OH
O
N
N3
30
1H NMR
TMS
0.00
8.59
8.56
7.99
7.95
7.54
7.50
7.57
7.60
7.39
6.18
6.15
5.97
5.33
4.33
4.04
4.01
3.96
3.82
3.77
2.61
2.68
2.48
2.45
1.26
0.94
2.09
4.34
0.90
6.43
1.22

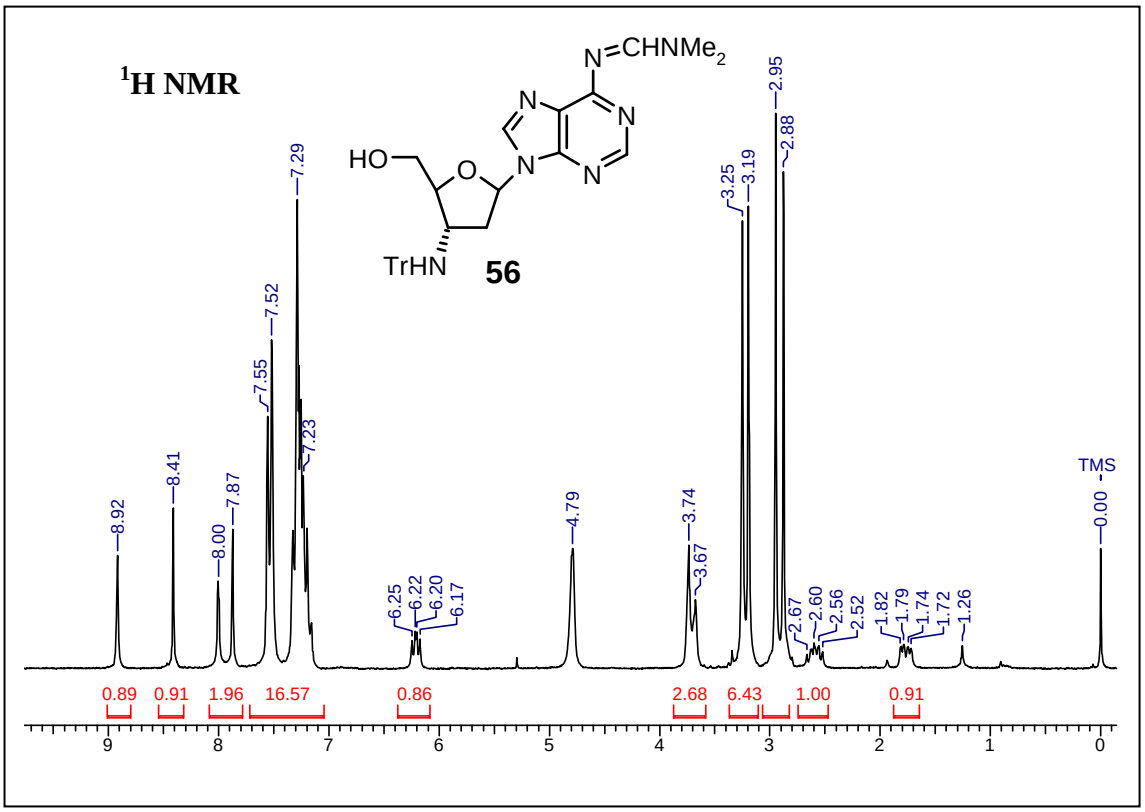

NHBz
OTs
N3
31
1H NMR
5.30
2.49
TMS
0.00
7.27
7.93
7.97
7.86
7.82
7.60
7.49
7.45
8.62
6.18
6.15
6.12
4.44
4.40
4.39
4.34
4.33
4.23
4.19
4.10
4.09
4.07
2.77
2.74
2.70
2.70
2.67
2.47
2.44
2.43
2.40
1.25
0.65
11.77
0.96
3.99
4.37

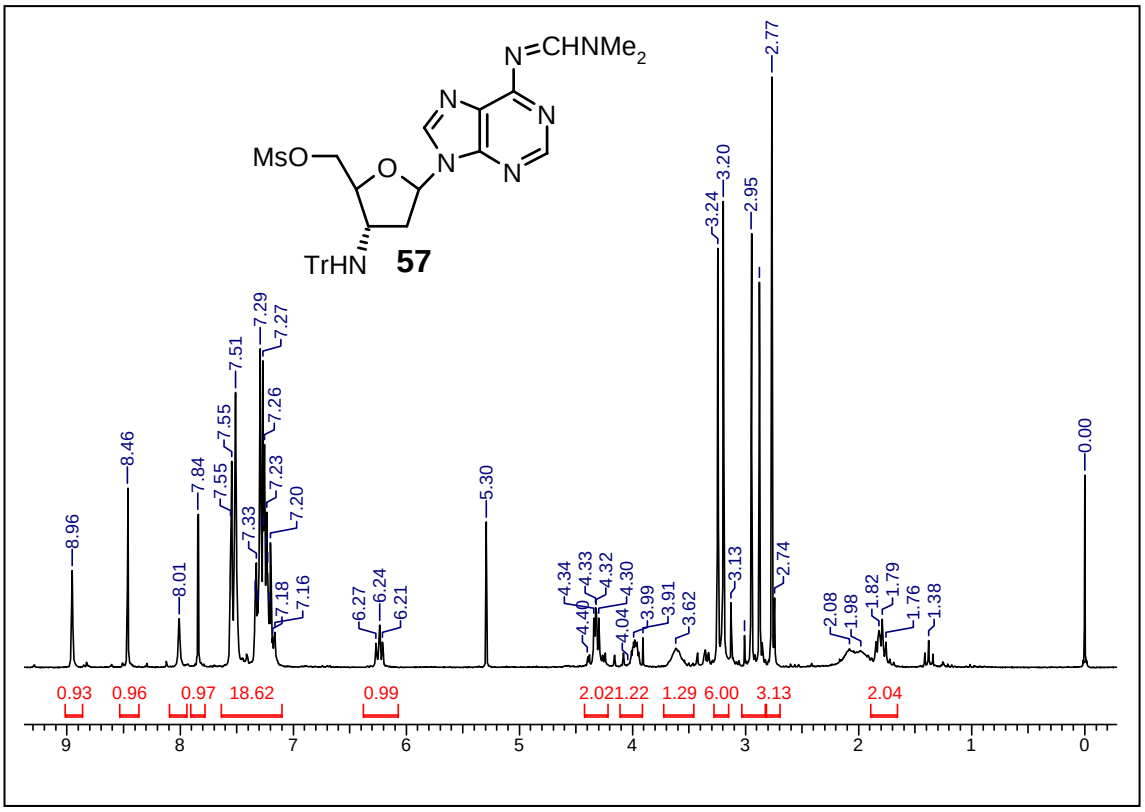

COOEt
NHBz
S
N3
32
1H NMR
1.30
TMS
0.00
8.15
8.11
7.99
7.99
7.95
7.81
7.54
7.61
7.63
7.50
7.46
7.43
7.28
6.16
6.13
6.10
4.24
4.20
4.27
4.16
4.15
4.14
4.12
4.10
3.43
3.36
3.33
3.25
3.08
3.06
2.71
2.68
2.48
2.43
2.43
2.04
1.34
1.27
1.26
1.24
1.22
10.41
0.91
3.83
1.76
1.54
1.00
1.29
2.94

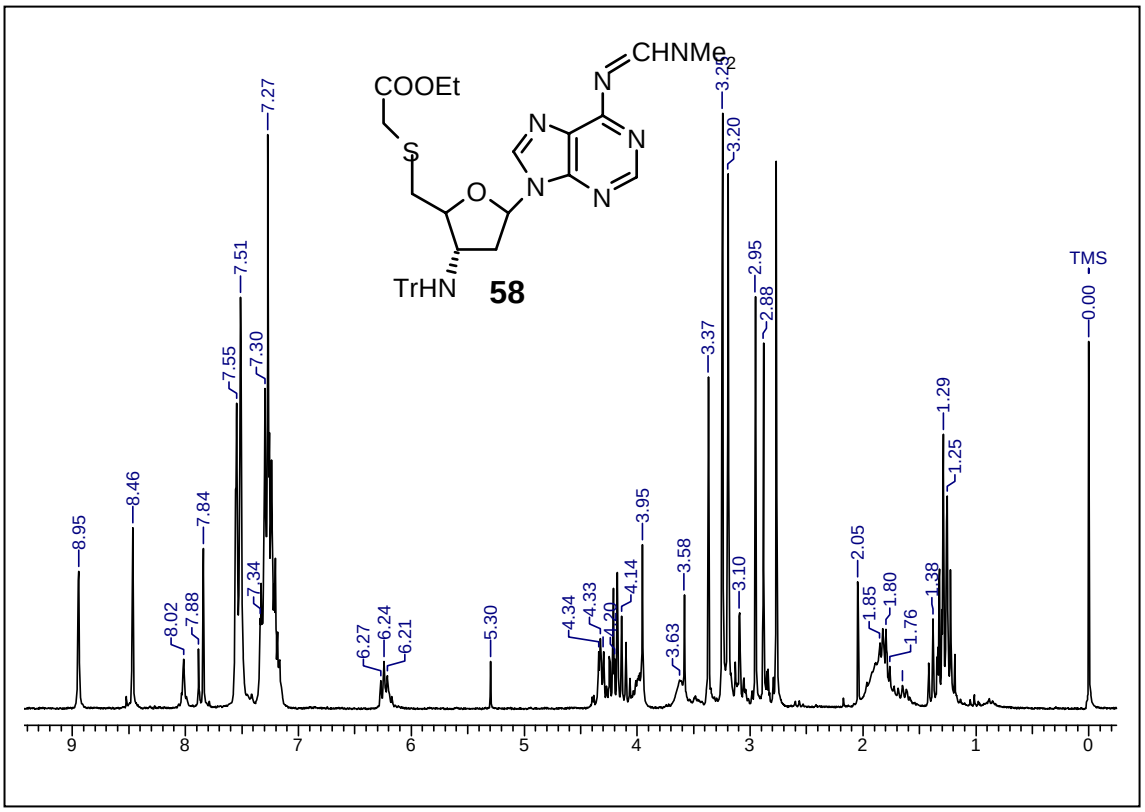

COOEt
NHBz
O
N
S
O
N
H_2N
33
^{1}H NMR
8.24
8.21
7.94
7.91
7.90
7.61
7.58
7.54
7.50
7.47
7.27
6.18
6.17
6.15
5.30
4.27
4.23
4.20
4.16
4.00
3.97
3.46
3.39
3.33
3.25
3.11
3.08
2.49
2.41
2.37
2.35
1.34
1.30
1.27
1.25
0.00
6.92
0.85
2.87
2.52
1.62
2.75
3.00

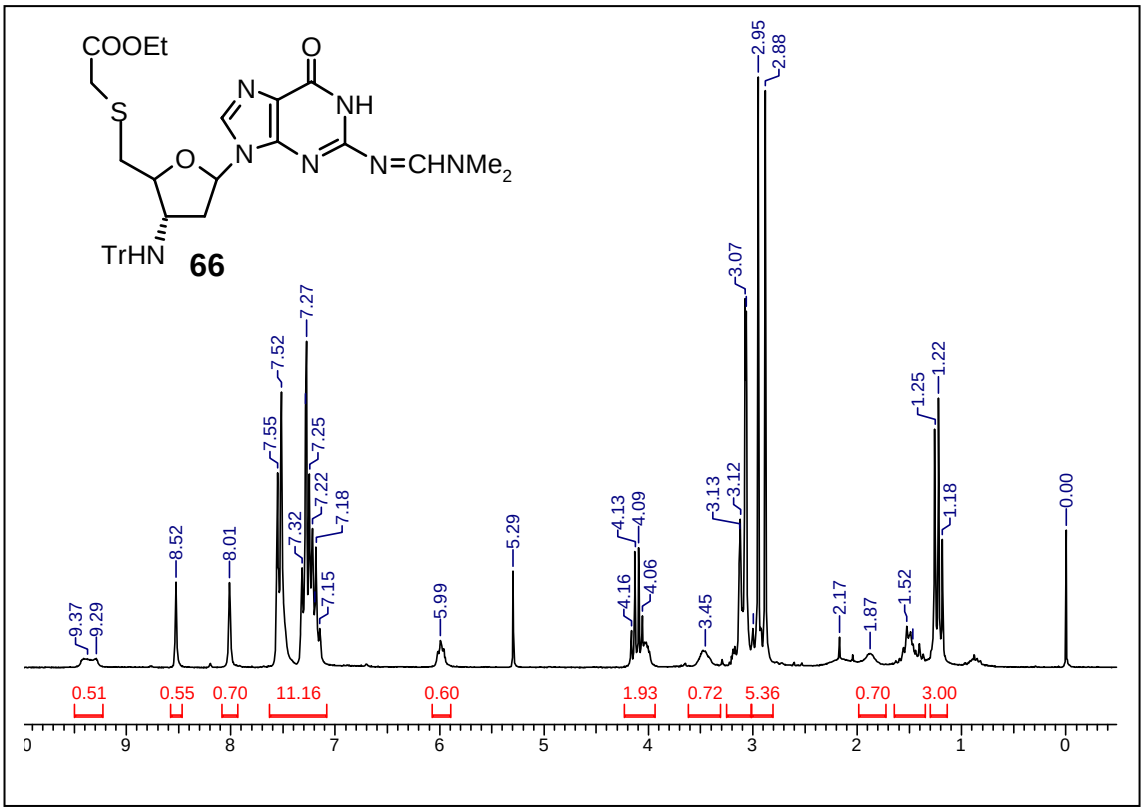

COOH
NH_2
O
N
S
O
N
FmocHN
35
^{1}H NMR
TMS
10.04
7.78
7.74
7.36
7.35
7.34
7.32
7.44
7.28
7.28
6.66
6.21
4.44
4.21
4.14
4.10
4.06
3.30
3.03
2.69
2.60
2.59
2.58
2.05
2.03
1.30
1.26
1.22
0.00
0.96
5.97
0.73
0.94
2.30
2.39
2.00

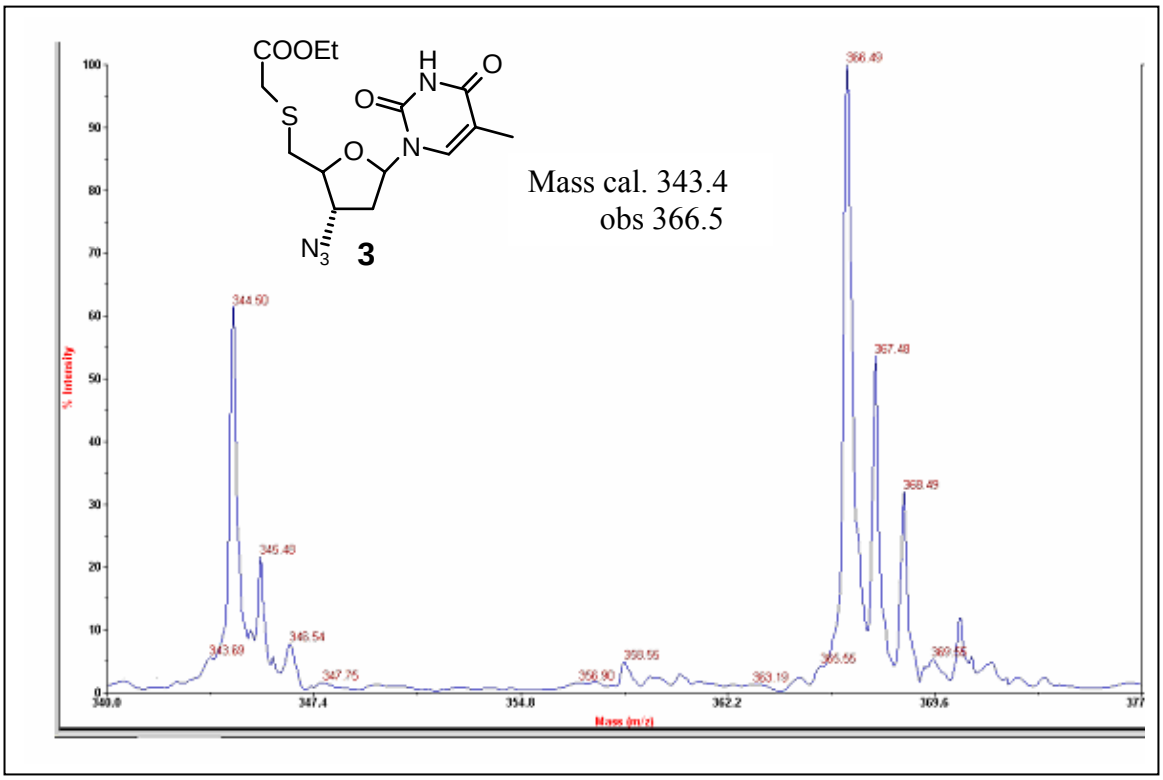

^{1}H NMR
NH2
N
N
N
N
HO
O
TrHN
55
8.19
8.00
7.74
7.56
7.55
7.52
7.33
7.33
7.29
7.27
7.25
7.23
7.20
7.18
7.16
6.21
6.18
6.16
6.13
5.29
4.77
4.17
4.14
4.10
4.07
3.76
3.73
3.67
3.30
3.29
3.23
2.95
2.88
2.53
2.50
2.49
2.46
2.42
2.04
1.75
1.74
1.29
1.26
1.22
0.00
0.97
1.06
11.80
0.92
2.85
1.00
1.00
1.01

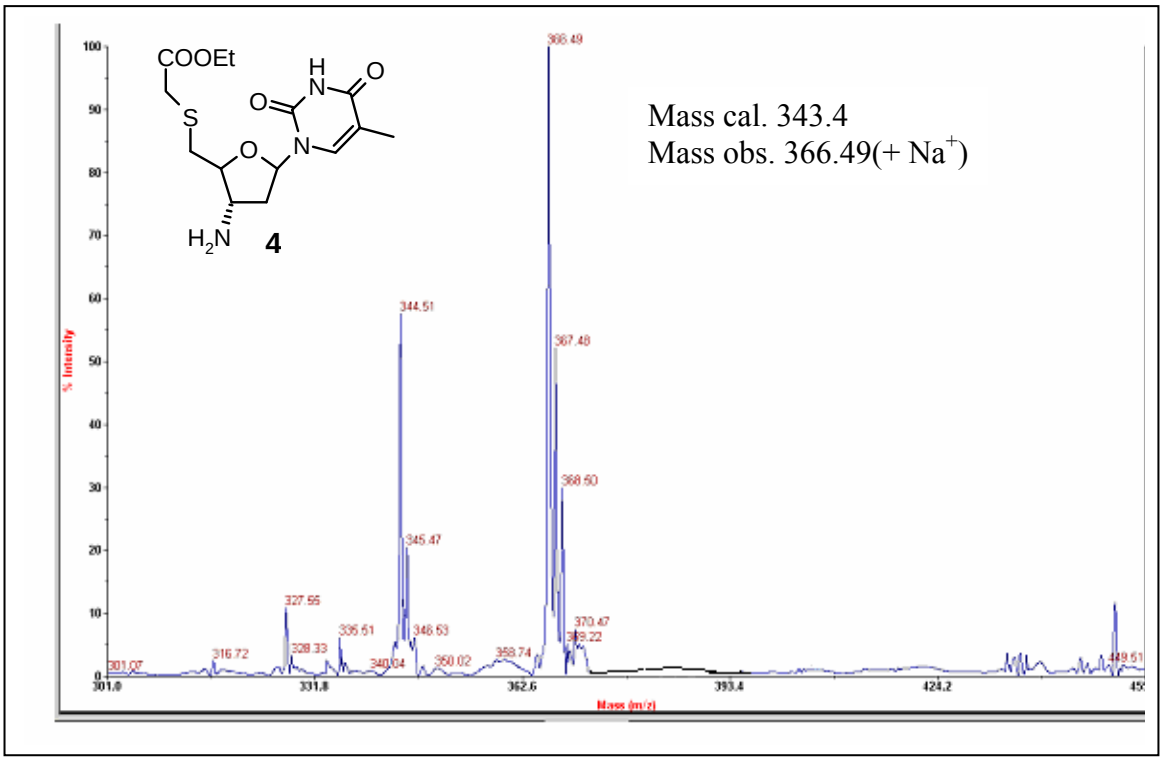

^{1}H NMR
N=CHNMe2
N
N
N
N
HO
O
TrHN
56
8.92
8.41
8.00
7.87
7.55
7.52
7.29
7.23
6.25
6.22
6.20
6.17
4.79
3.74
3.67
3.25
3.19
2.95
2.88
2.67
2.60
2.56
2.52
1.82
1.79
1.74
1.72
1.26
TMS
0.00
0.89
0.91
1.96
16.57
0.86
2.68
6.43
1.00
0.91

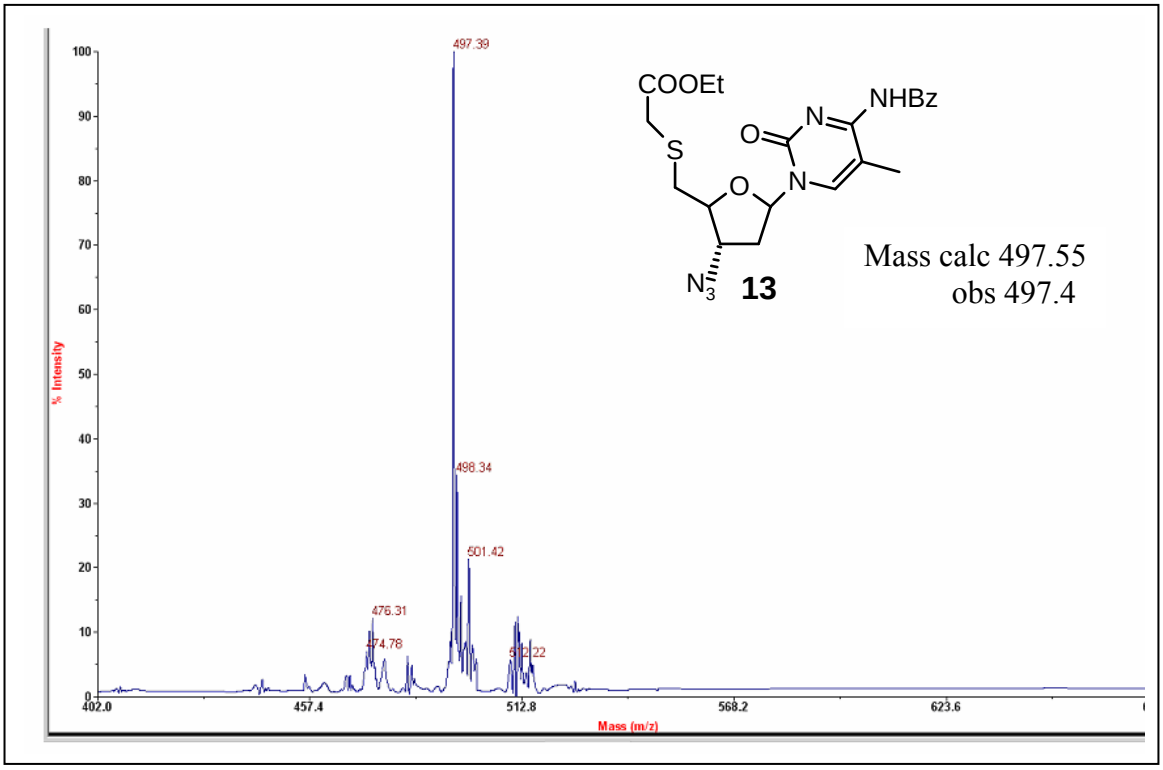

N=CHNMe2
MsO
TrHN
57
8.96
8.46
8.01
7.84
7.55
7.55
7.51
7.33
7.29
7.27
7.26
7.23
7.20
7.18
7.16
6.27
6.24
6.21
5.30
4.34
4.33
4.32
4.30
4.40
4.04
3.99
3.91
3.62
3.24
3.20
3.13
2.95
2.77
2.74
2.08
1.98
1.82
1.79
1.76
1.38
0.00
0.93
0.96
0.97
18.62
0.99
2.02
1.22
1.29
6.00
3.13
2.04

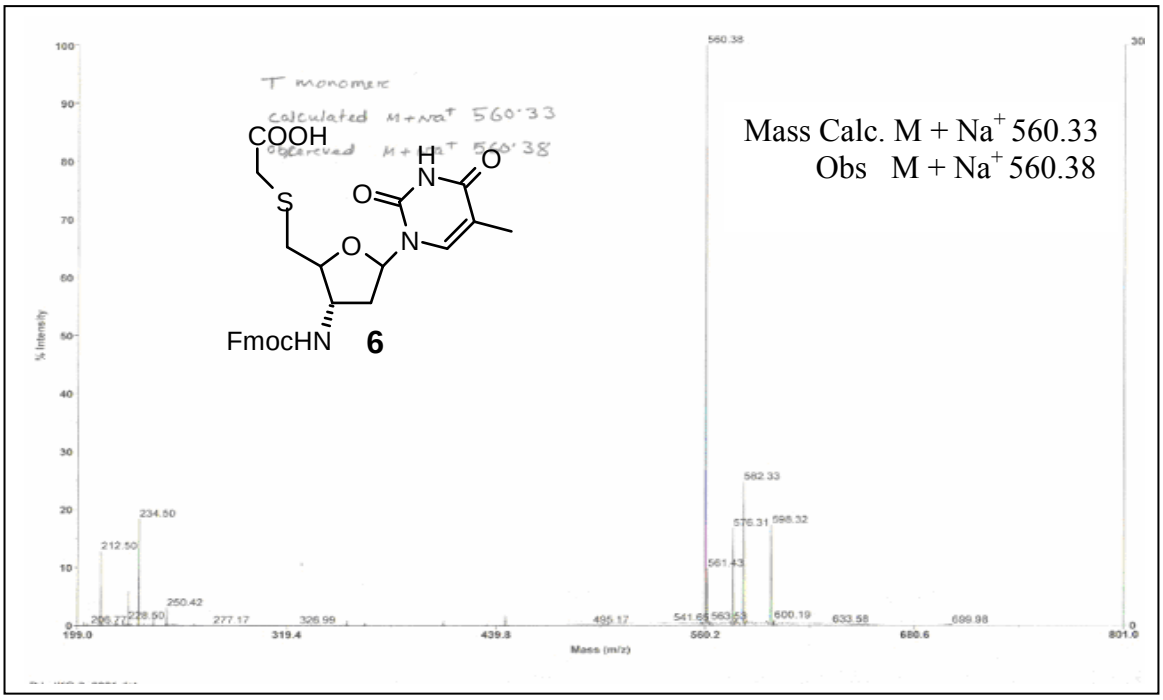

CHNMe2
COOEt
S
TrHN
58
8.95
8.46
8.02
7.88
7.84
7.55
7.51
7.34
7.30
7.27
6.27
6.24
6.21
5.30
4.34
4.33
4.20
4.14
3.95
3.63
3.58
3.37
3.25
3.20
3.10
2.95
2.88
2.05
1.85
1.80
1.76
1.38
1.29
1.25
TMS
0.00

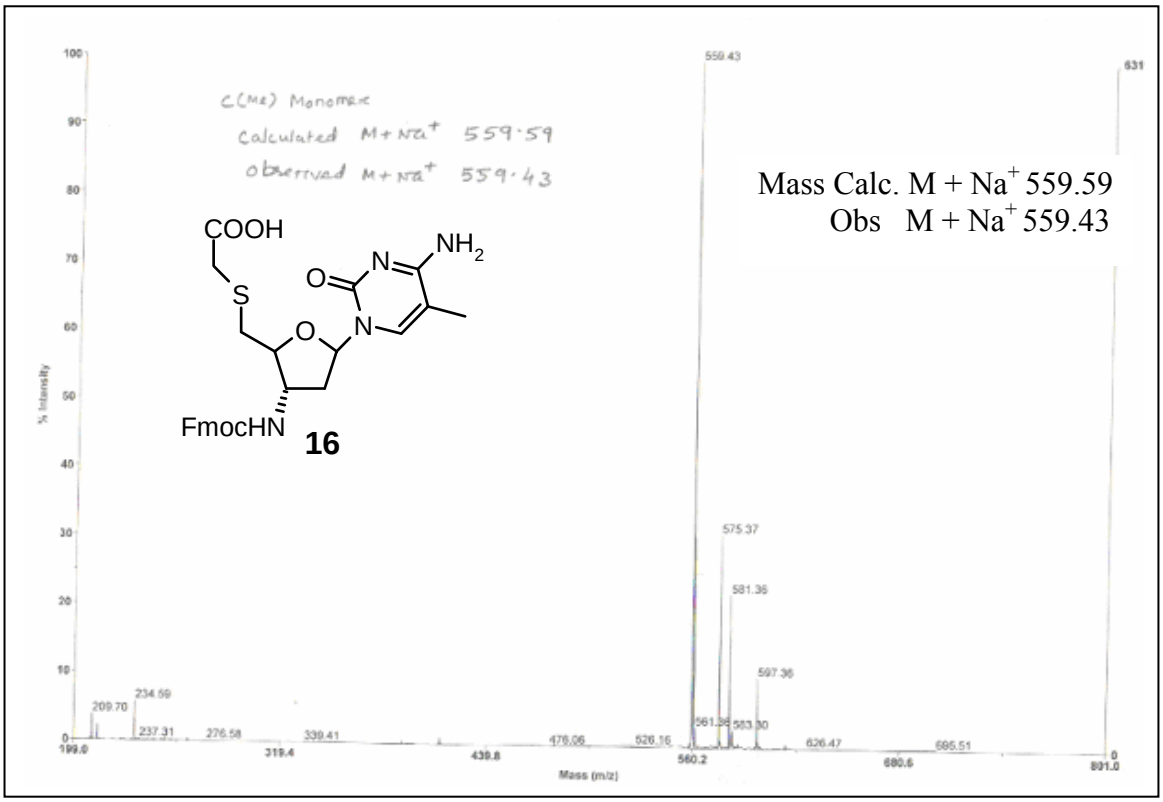

COOEt
S
O
N
NH
N
N
N=CHNMe2
TrHN
66
9.37
9.29
8.52
8.01
7.55
7.52
7.32
7.27
7.25
7.22
7.18
7.15
5.99
5.29
4.16
4.13
4.09
4.06
3.45
3.13
3.12
3.07
2.95
2.88
2.17
1.87
1.52
1.25
1.22
1.18
0.00
0.51
0.55
0.70
11.16
0.60
1.93
0.72
5.36
0.70
3.00

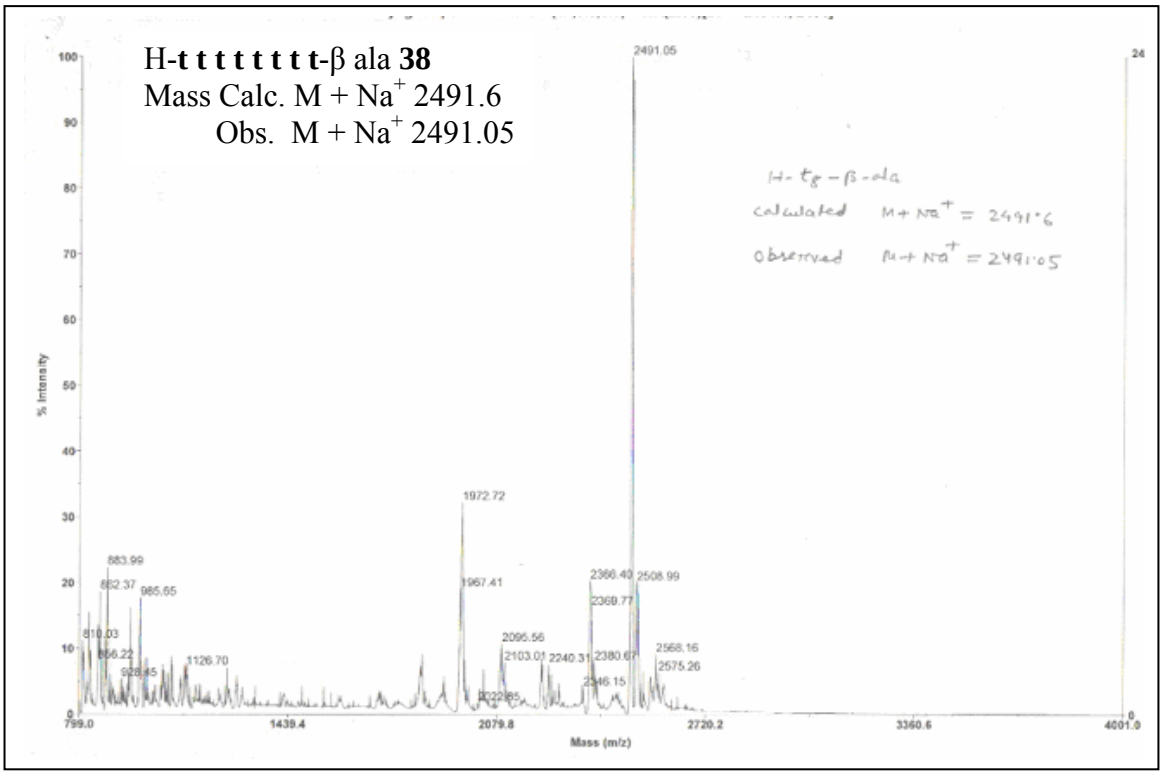

COOEt
S
O
H
N
O
O
N
N3
3
Mass cal. 343.4
obs 366.5
% Intensity
344.50
345.49
346.54
343.69
347.75
358.55
358.90
363.19
365.55
366.49
367.48
368.49
369.55
Mass (m/z)

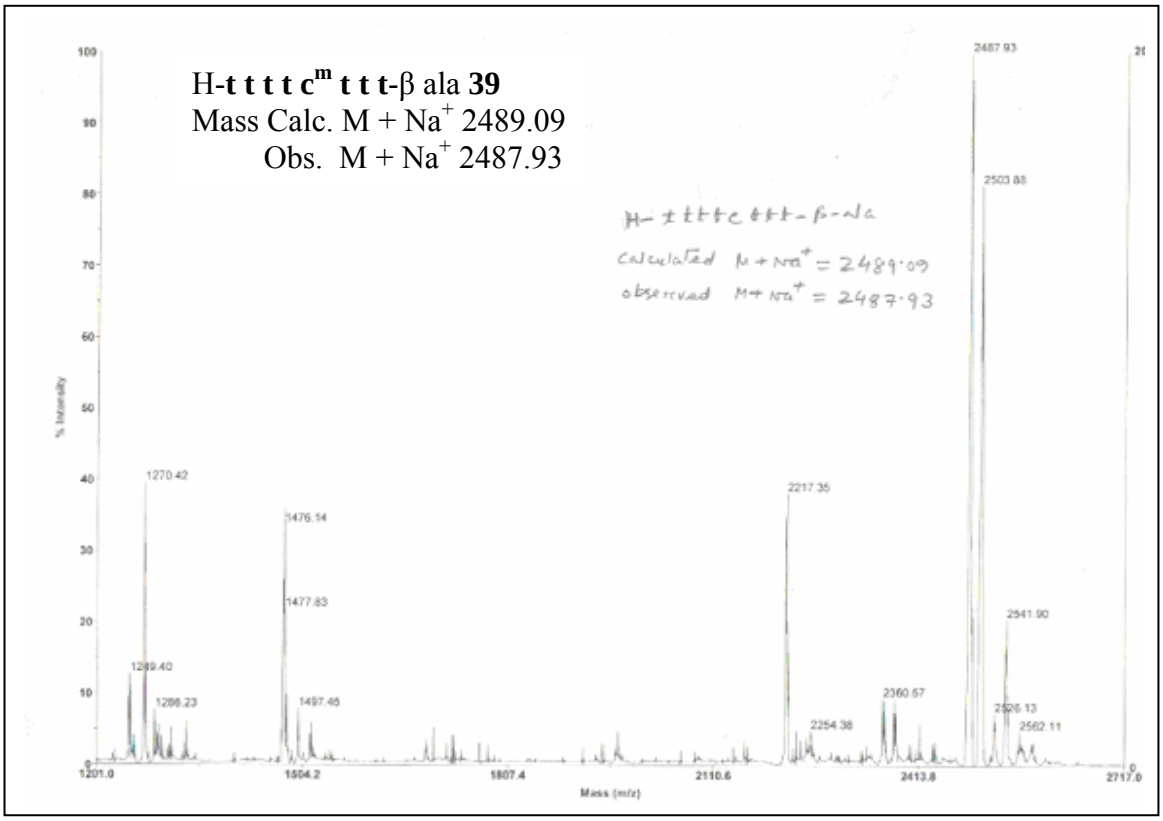

COOEt
S
O
H
N
O
O
N
H2N
4
Mass cal. 343.4
Mass obs. 366.49(+ Na+)
% Intensity
Mass (m/z)

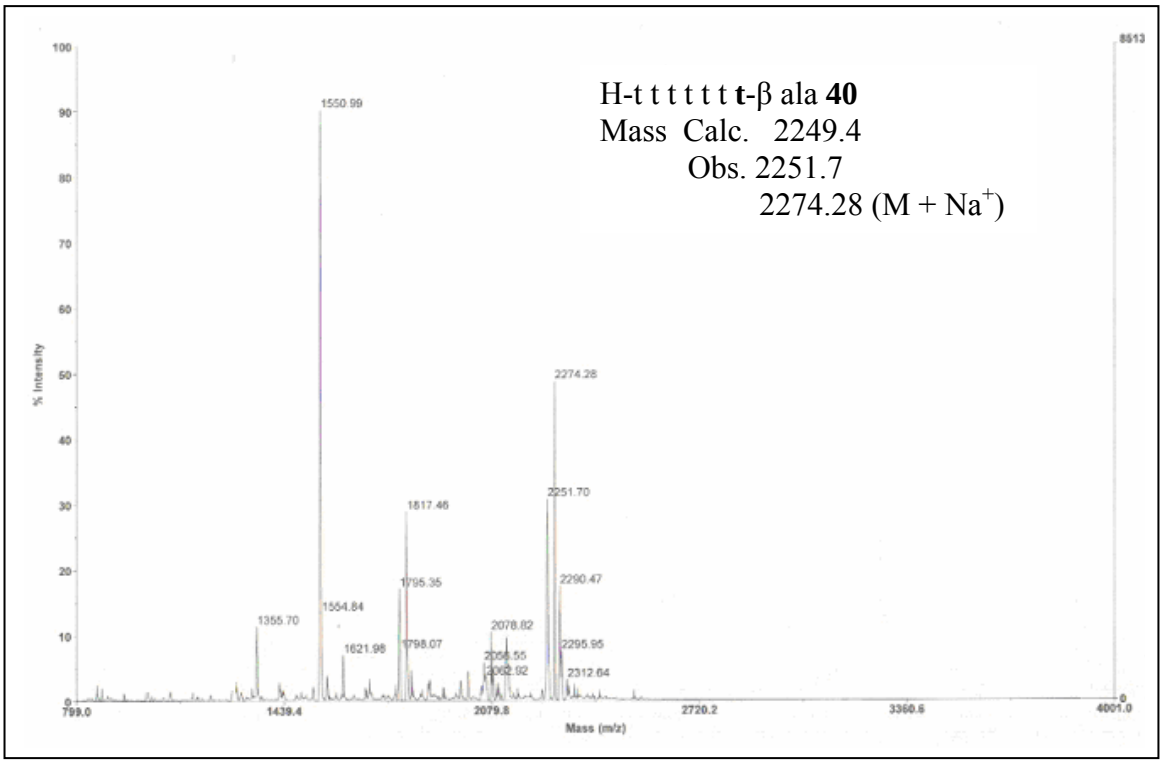

COOEt
S
O
N
NHBz
O
N
N3
13
Mass calc 497.55
obs 497.4
% Intensity
Mass (m/z)

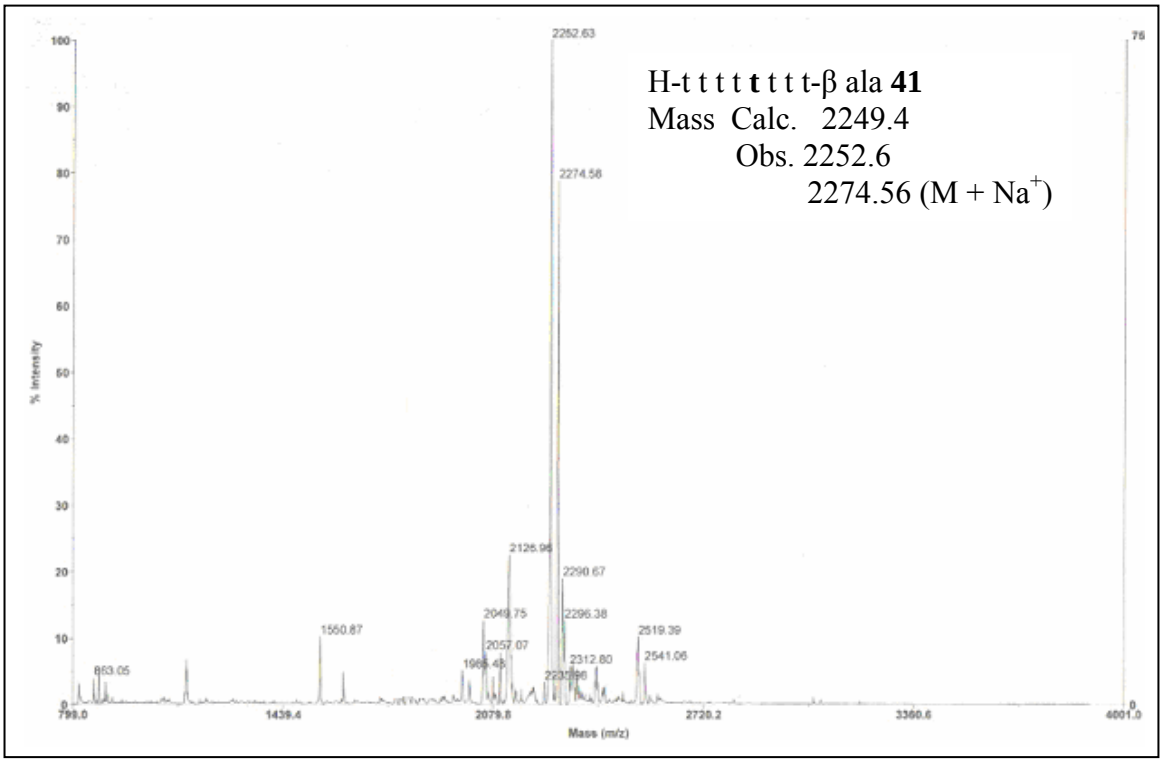

T monomer
calculated M+Na+ 560.33
observed M+Na+ 560.38
COOH
S
O
N
O
N
O
FmocHN
6
Mass Calc. M + Na+ 560.33
Obs M + Na+ 560.38
% Intensity
Mass (m/z)

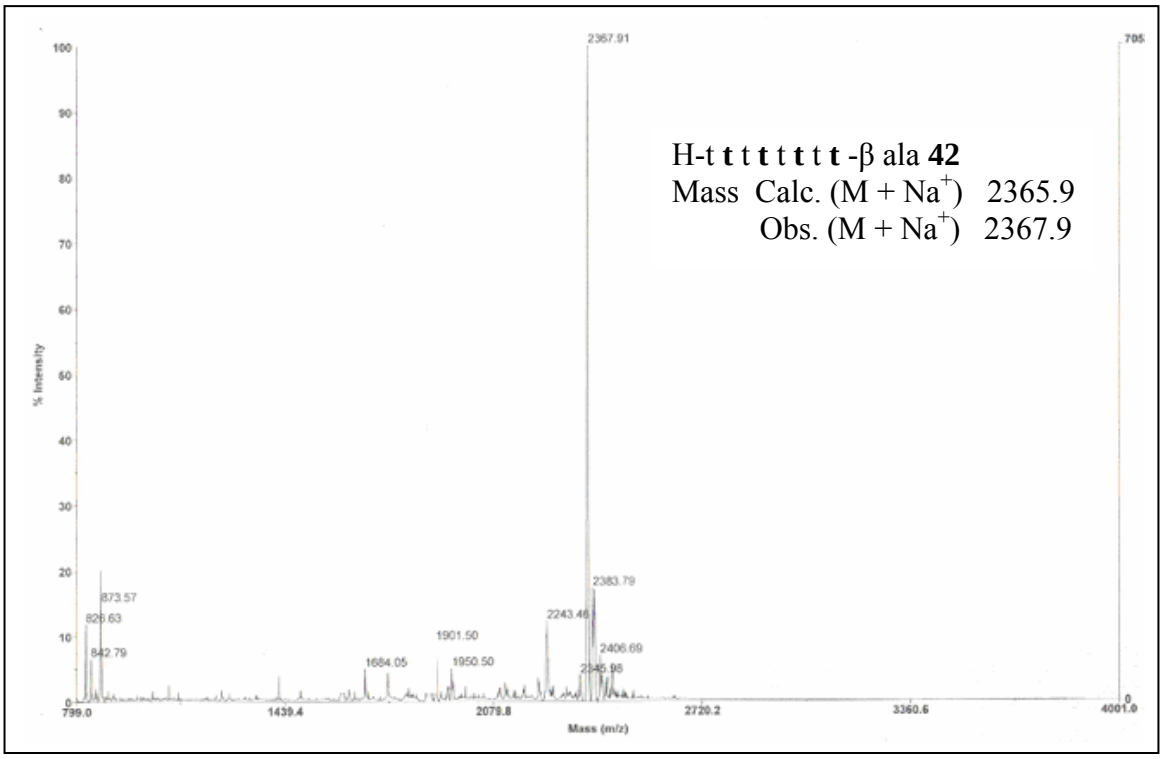

C(Me) Monomer
Calculated M+Na+ 559.59
observed M+Na+ 559.43
COOH
S
O
N
O
N
NH2
FmocHN
16
Mass Calc. M + Na+ 559.59
Obs M + Na+ 559.43
% Intensity
Mass (m/z)

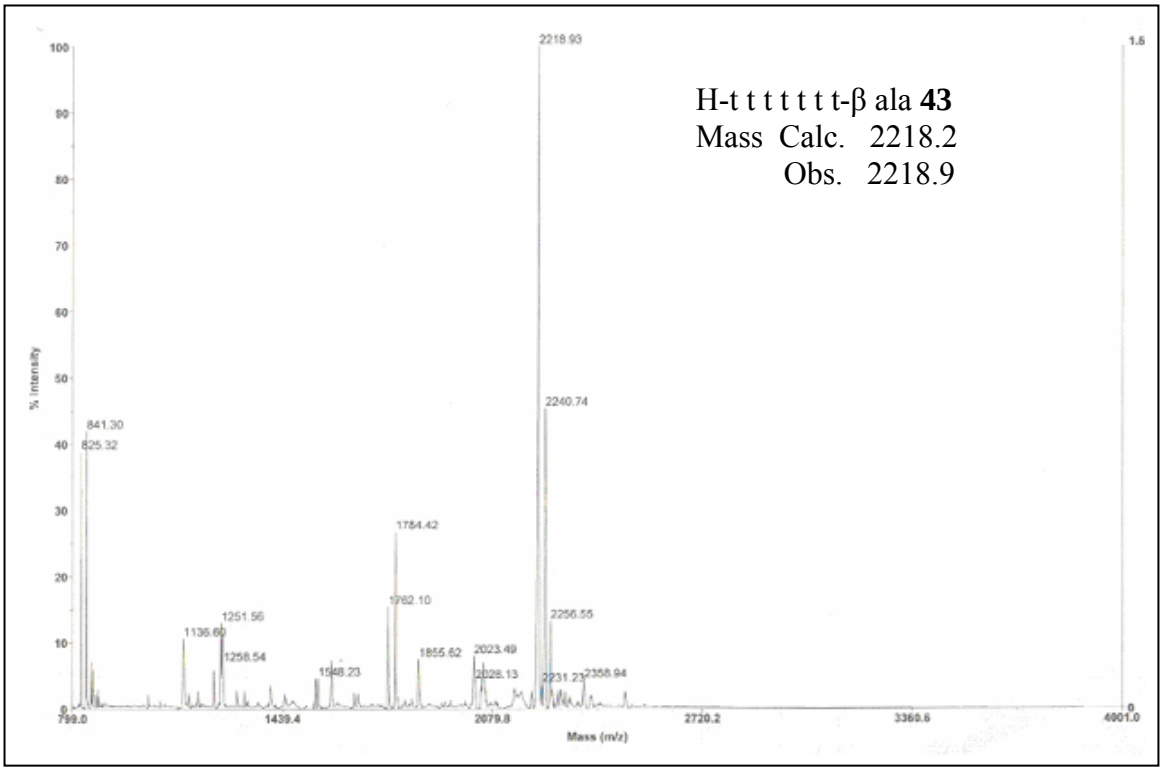

H-t t t t t t t t-β ala 38
Mass Calc. M + Na+ 2491.6
Obs. M + Na+ 2491.05
H- tg - β - ala
calculated M + Na+ = 2491·6
observed M + Na+ = 2491·05
% Intensity
Mass (m/z)
2491.05
1972.72
1967.41
883.99
852.37
985.65
810.03
856.22
928.45
1126.70
2095.56
2103.01
2240.31
2366.40
2369.77
2380.67
2346.15
2508.99
2568.16
2575.26
799.0
1439.4
2079.8
2720.2
3360.6
4001.0


H-t t t t c^m t t t-β ala 39
Mass Calc. M + Na+ 2489.09
Obs. M + Na+ 2487.93
H- t t t t c t t t - β - ala
calculated M + Na+ = 2489·09
observed M + Na+ = 2487·93
% Intensity
Mass (m/z)
2487.93
2503.88
1270.42
1476.14
1477.83
1269.40
1286.23
1497.46
2217.35
2254.38
2360.57
2541.90
2526.13
2562.11
1201.0
1504.2
1807.4
2110.6
2413.8
2717.0


H-t t t t t t **t**-β ala **40**
Mass Calc. 2249.4
Obs. 2251.7
2274.28 (M + Na+)
% Intensity
Mass (m/z)
1550.99
2274.28
2251.70
1817.46
1795.35
2290.47
1554.84
1355.70
1621.98
1798.07
2078.82
2295.95
2312.64
799.0
1439.4
2079.8
2720.2
3360.6
4001.0


H-t t t t **t** t t t-β ala **41**
Mass Calc. 2249.4
Obs. 2252.6
2274.56 (M + Na+)
% Intensity
Mass (m/z)
2252.63
2274.56
2126.96
2290.67
2296.38
2519.39
2049.75
2057.07
2312.80
2541.06
1550.87
863.05
799.0
1439.4
2079.8
2720.2
3360.6
4001.0


H-t **t** t t **t** t t **t** -β ala **42**
Mass Calc. (M + Na+) 2365.9
Obs. (M + Na+) 2367.9
% Intensity
Mass (m/z)
2367.91
873.57
826.63
842.79
1684.05
1901.50
1950.50
2243.46
2383.79
2406.69
799.0
1439.4
2079.8
2720.2
3360.6
4001.0


H-t t t t t t t-β ala **43**
Mass Calc. 2218.2
Obs. 2218.9
% Intensity
Mass (m/z)
2218.93
841.30
825.32
1251.56
1136.60
1258.54
1548.23
1784.42
1762.10
1855.62
2023.49
2028.13
2240.74
2256.55
2358.94
799.0
1439.4
2079.8
2720.2
3360.6
4001.0

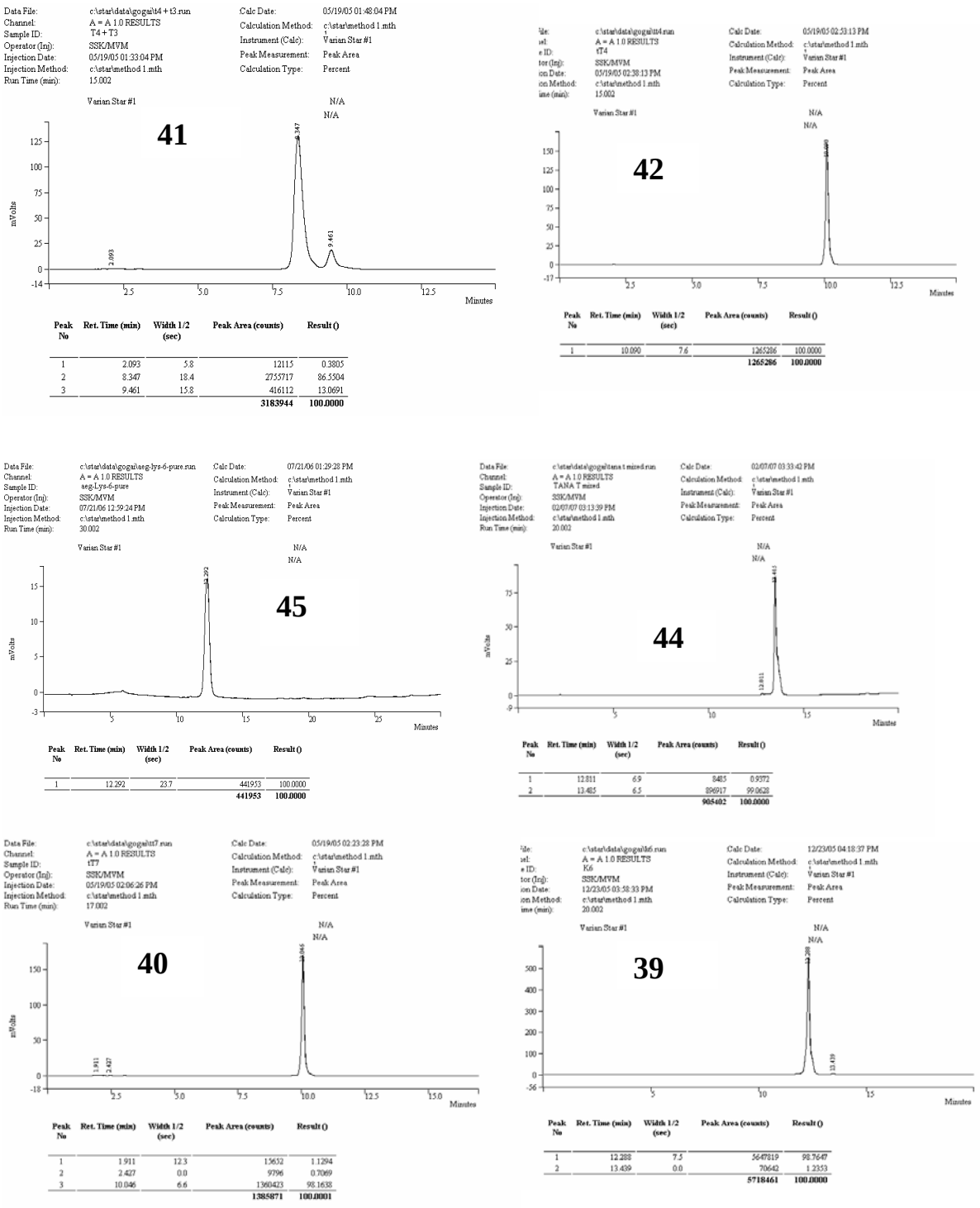


Figure. RP-HPLC of TANA, chimeric *aeg*PNA- TANA ONs

Chapter 2

Section II Synthesis and RNA binding selectivity of oligonucleotides modified with five-atom TANA backbone structures

2.2.1 Introduction

In the previous section, a convenient synthetic methodology is described to convert pyrimidine nucleosides to a sugar-amino acid monomer unit by using thiol functionality in ethyl mercaptoacetate as a requisite nucleophile. The homooligomeric pyrimidine ONs were found to bind to complementary RNA sequences significantly better than their DNA counterparts and the binding efficiency was found to be as good as PNA itself. However, introduction of these amino acid nucleoside derivatives in PNA sequences were found to be detrimental for RNA/DNA binding. In this section, we present the synthesis and incorporation of thymidine and thymidine-cytidine dimer blocks (**tst** and **cst)** connected with a five-atom amide-linker NH3'-CO-CH_2-S-CH_2 (TANA) (Figure 1) into oligonucleotide oligomers. The retention of some negatively charged phosphate groups would be advantageous for water solubility. The assessment of the compatibility of the TANA dimer blocks in sugar-phosphate backbone was studied by UV-T_m measurements of the resulting mixed-backbone ON complexes with DNA and RNA.

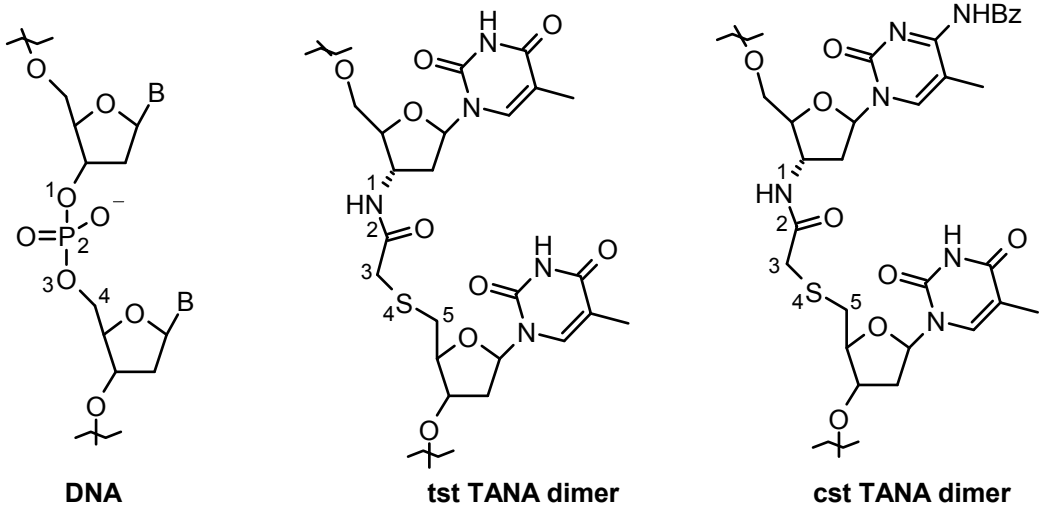


Figure 1. Structure of DNA and TANA

2.2.2 Solid phase synthesis of DNA Oligonucleotides by phosphoramidite method

General solid-phase chemical synthesis of ONs by β-cyanoethyl phosphoramidite method is shown in Figure 2.[43]

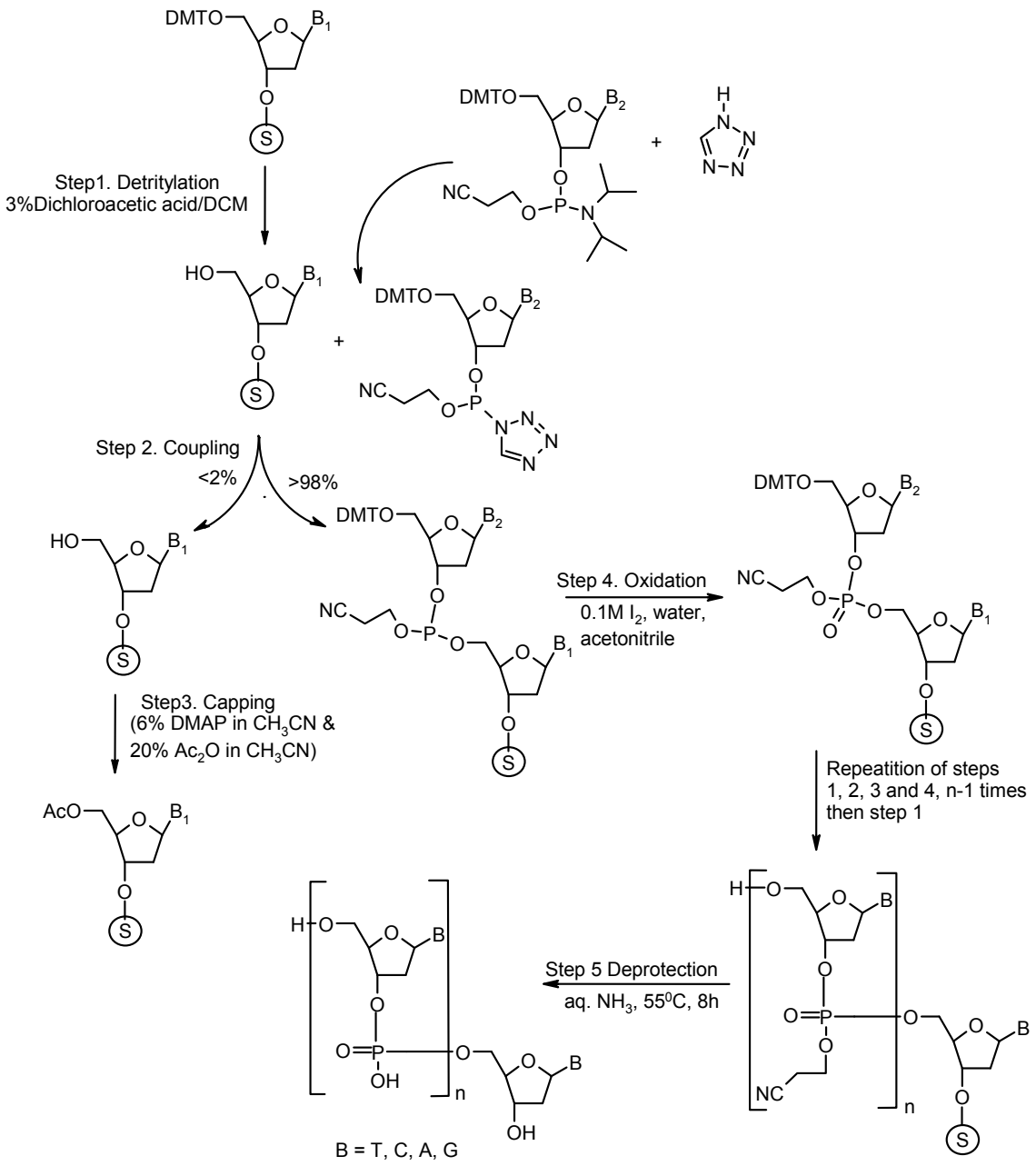


Figure 2 Solid-phase syntheses of oligodeoxynucleotides by phosphoramidite method.

2.2.3 Results and discussions

2.2.3.1 Synthesis of tst and cst dimer building blocks

The retro synthetic analysis of the dimer building block is depicted in Figure 3.

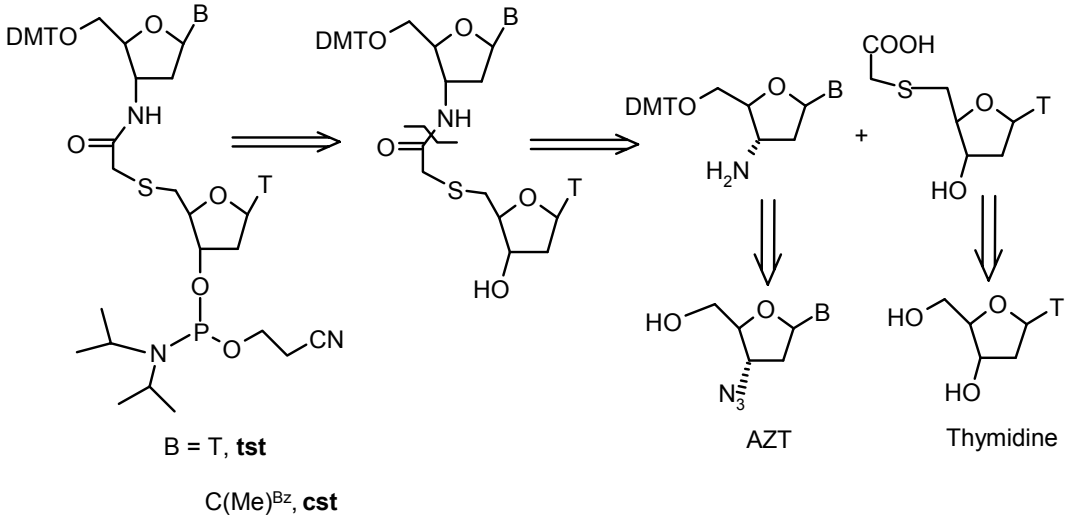
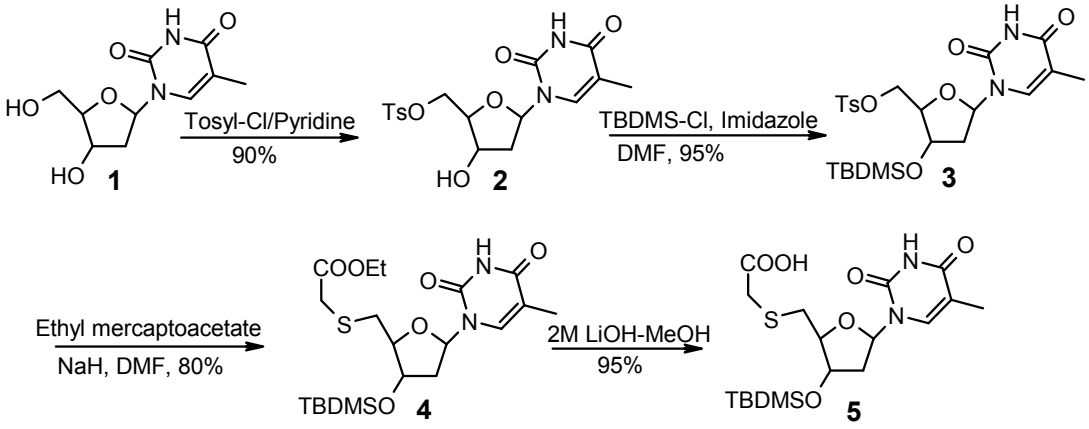


Figure 3. Retro synthetic pathway of TANA dimer building block

The retro synthetic analysis shows that the **tst** and **cst** dimer building blocks can be synthesized starting from 3'-azido thymidine and thymidine. The dimer blocks can easily be created by the coupling between the the 3'-amino nucleoside and thymidine 5'-mercapto acetic acids.

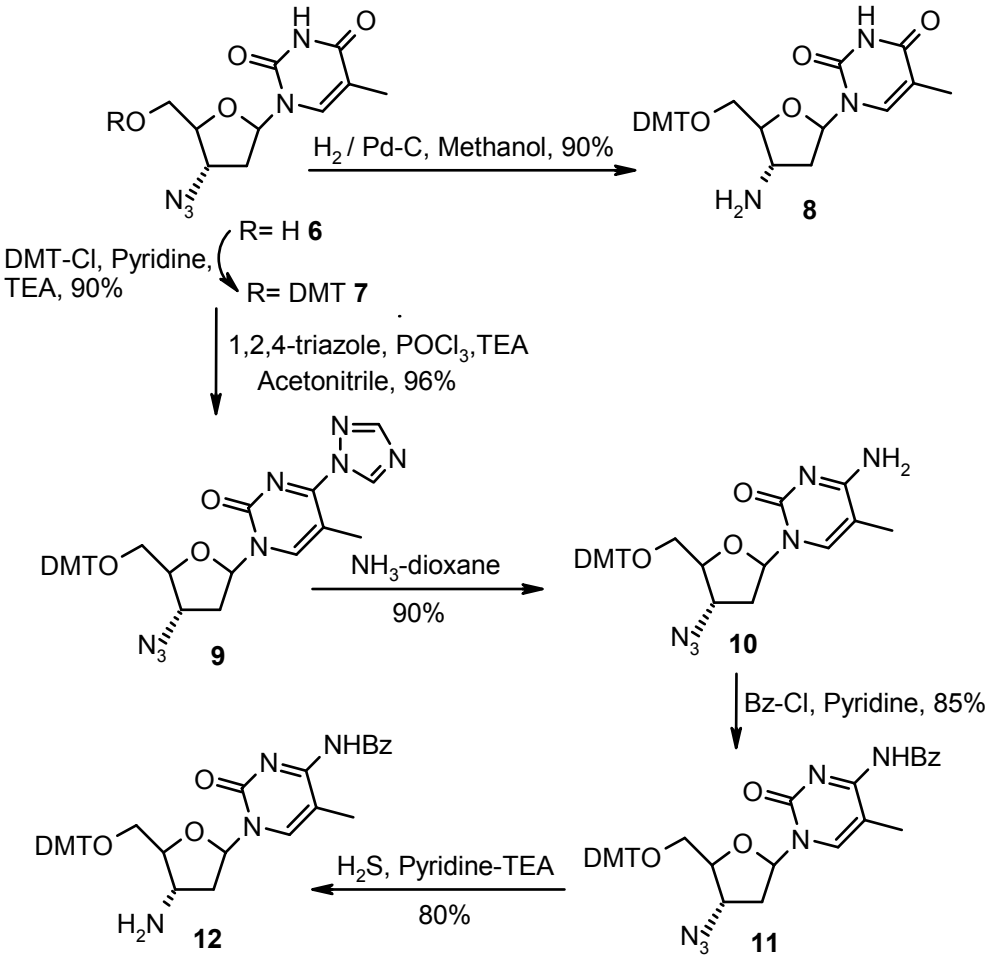
2.2.3.2 Synthesis of acid monomer synthon

Synthesis of thymidine 5'- mercapto acetic acid monomer synthon (Scheme 1) was done starting from thymidine **1**. The 5'-OH of thymidine **1** was selectively tosylated by reaction with tosyl chloride in pyridine to give **2**. High dilution and low temperature (~ 0ºC) was maintained to prevent the formation of ditosylated product. A small amount of ditosylated product (5%) was always formed which could be easily separated from the product by crystallization. The 3'-OH was then protected as 3'-*O*-TBDMS group to give **3**. Nucleophilic displacement of 5'-*O*- tosylate in **3** by ethyl mercaptoacetate gave the 5'-mercapto ethyl ester derivative **4**. Hydrolysis of the ethyl ester using 2M LiOH to give the acid synthon **5** (Scheme 1).



Scheme 1. Synthesis of thyminyl 5'-mercapto acetic acid monomer synthon

2.2.3.3 Synthesis of 3'-amine monomer synthons

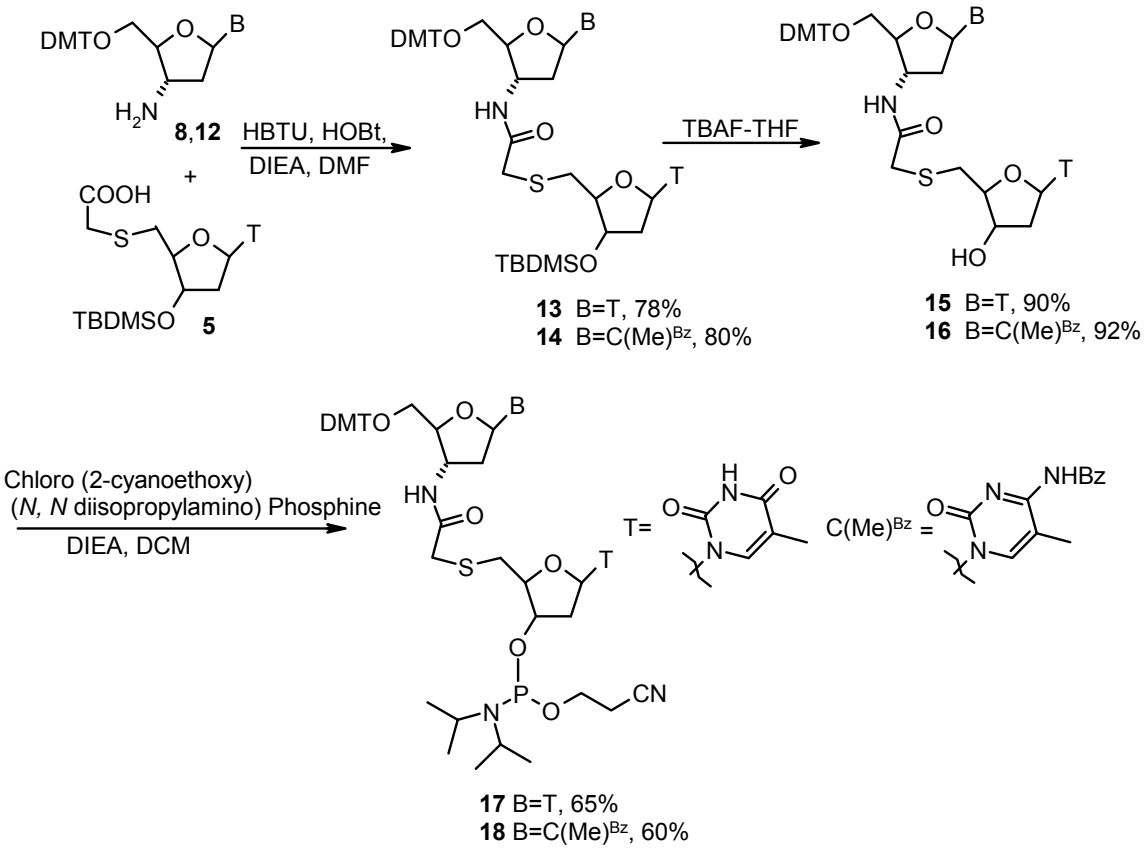


Scheme 2. Synthesis of amine synthons.

The amine synthon of thymidine **8** was obtained from 3'-azidothymidine **6** by protection of the free hydroxy group as 5'-*O*-DMT in **7** and reduction of azide to amine by catalytic hydrogenation using Pd-C in methanol to give **8**. The amine synthon of 5-methyl-2'-deoxycytidine was achieved by conversion of 5'-*O*-DMT, 3'-azidothymidine **7** to 5-methyl-2'-deoxycytidine derivative by known procedures *via* C4-triazolide **9**[22] followed by amination **10**, benzoylation **11** and reduction of the azide gave the desired cytidine derivative **12** (Scheme 2).

2.2.3.4 Synthesis of tst and cst dimer building blocks

The synthesis of the dimer building blocks **tst 13** and **cst 14** was then achieved by using peptide coupling chemistry from the monomer units, using the coupling reagents HBTU/ HOBt/ DIEA in acetonitrile. The formation of the product was confirmed by ^{1}H NMR as



Scheme 3. Synthesis of **tst** and **cst** phosphoramidite dimer building blocks

well as mass spectroscopy. Then desilylation of the 3'-OTBDMS by using TBAF in THF to give the desilylated products **15** and **16**. Then phosphitilation of the 3'- OH of **15** and **16** were done using the reagent chloro (2-cyanoethoxy) (*N, N*-diisopropylamino)-phosphine using the base diisopropyl ethyl amine and DCM as the solvent to give the phosphoramidites **17** and **18**. The formation of the products was confirmed by ^{31}P NMR spectroscopy.

2.2.3.5 Synthesis of DNA Oligomers

A series of chimeric ONs (Table 1) containing one to four **tst** and **cst** blocks were synthesized by automated solid phase synthesis using phosphoramidite approach and Applied Biosystems 3900 DNA Synthesizer.[43]

Table1. TANA ONs synthesized and their HPLC and mass characterization

Entry	Sequence (5' to 3')	HPLC[a] (min)	Mass	
			calcd	found
1	CG TT**tst**TTT TGC **19**	7.9	3606.5	3606.6
2	CGTT **tst** TT **tst** GC **20**	8.3	3596.6	5597.0
3	CG **tst tst tst tst** GC **21**	9.2	3576.8	3576.6
4	TCT C **tst** TCT T **22**	8.6	2933.1	2933.1
5	TCT C **tst** TC **tst** **23**	9.2	2923.2	2924.0
6	T**cst** CTT TCTT **24**	8.0	2948.2	2948.1
7	T**cst** CTT T**cst** T **25**	8.8	2938.3	2938.7
8	T**cst cst**T T**cst** T **26**	9.0	2928.4	2928.7
9	TCA **cst** A GAT G **27**	8.3	3036.2	3037.0
10	CTC TTC CTC C **tst** C **28**		3784.7	3784.4
11	CCT C **tst** ACC TCA G TT ACA **29**	8.9	5366.7	5366.8
12	CCT C **tst** ACC TCA G **tst** ACA **30**	9.3	5356.9	5356.9
13	C **cst** C **tst** AC **cst** CAG **tst** ACA **31**	9.9	5367.2	5367.7
14	GAA GGG **cst**T **tst**G AA **cst cst** T **32**	10.2	5846.6	5846.8
15	GAA GGG **cst**T **tst**G AA CT **cst** T **33**	9.6	5842.4	5841.3
16	AGA G **tst** CAA AAG CCC **tst**T T **34**	8.9	5774.2	5774.4
17	AGA G **tst** CAA AAG CC**cst tst** T **35**	9.4	5779.3	5779.7
18	GC**tst** CC TTC CTT CG **36**	9.4	4153.9	4147.9
19	GC**tst** CC **tst**C CTT CG **37**	9.9	4144.05	4142.5
20	GC**tst** CC **tst**C C**tst** CG **38**	10.3	4134.2	4139.6
21	G**cst** TCC TTC CTT CG **39**	9.7	4169	4163.3

The **tst** dimer was subjected to aq. NH_3 treatment for 8h at 55ºC to check the stability of amide linkage towards the standard cleavage conditions. TLC showed no degradation of

the product and hence the stability of the mercapto acetamide backbone. The dimer phosphoramidites **tst** and **cst** was incorporated into ON sequences at the desired positions by automated DNA synthesis to yield the chimeric ONs shown in Table 1. The coupling yields of the chimeric ONs were >90% as quantified by RP HPLC. The synthesized ONs were cleaved form the solid support with conc. NH_4OH, lyophilized and desalted to get the crude ONs. The purity of these crude ONs were checked by RP-HPLC and found to be more than >90%.

All the ONs synthesized were purified by RP-HPLC and characterized by MALDI-TOF mass spectroscopy. Some of the representative RP-HPLC of the modified TANA ONs is presented Figure 4.

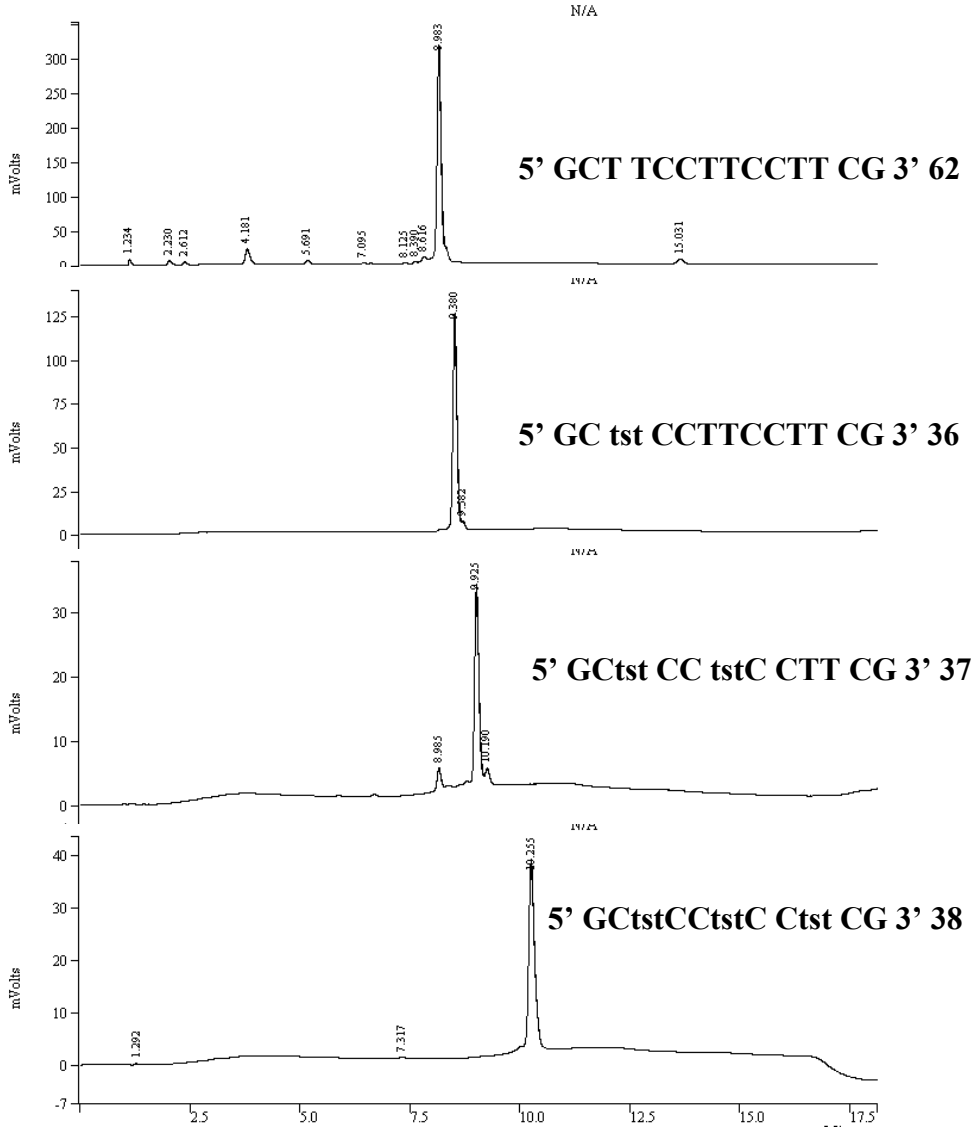


Figure 4. Variation of HPLC t_R with respect to the number of modifications

In case of ON **62,** when there is no TANA modification the HPLC retention time is 9 minute, for ONs **36**, **37** and **38** having single, double and triple modifications respectively, the HPLC t_R are 9.4, 9.9 and 10.3 minutes respectively, this clearly indicates hydrophobicity of the ONs increases with the number modifications. This is due to the replacement of the negative phosphate backbone by neutral TANA backbone.

2.2.3.6 Biophysical studies of TANA ONs

The DNA ONs (Table 2) were synthesized on Applied Biosystems ABI 3900 High Throughput DNA Synthesizer using standard β-cyanoethyl phosphoramidite chemistry.[43] The RNA ONs were obtained commercially.

Table 2. DNA and RNA sequences used for UV melting and CD studies

No	Sequence(5' to 3')	Type
1	CGTTTTTTTTGC **40**	Control for **19**, **20** and **21**
2	GCAAAAAAAACG **41**	Complementary for **19**, **20** and **21**
3	GCAAAACAAAACG **42**	Mismatch for **19**, **20** and **21**
4	r(GCAAAAAAAACG) **43**	Complementary for **19**, **20** and **21**
5	r(GCAAAAUAAACG) **44**	Mismatch for **19**, **20** and **21**
6	TCT CTT TCT T **45**	Control for **22**, **23**, **24**, **25** and **26**
7	AAG AAA GAG A **46**	Complementary for **22**, **23**, **24**, **25** and **26**
8	r (AAG AAA GAG A) **47**	Complementary for **22**, **23**, **24**, **25** and **26**
9	TCA CTA GAT G **48**	Control for **27**
10	CAT CTA GAG A **49**	Complementary for **27**
11	r(CAT CTA GAG A) **50**	Complementary for **27**
12	CTC TTC CTC CTT C **51**	Control for **28**
13	GAA GGA GGA AGA G **64**	Complementary for **28**
14	CCT CTT ACC TCA GTT ACA **52**	Control for **29**, **30** and **31**
15	TGTAACTGA GGT AAG AGG **53**	Complementary for **29**, **30** and **31**
16	r(UGU AACUGAGGUAAGAGG) **54**	Complementary for **29**, **30** and **31**
17	r(UGUAACUGACGUAAG AGG) **55**	Mismatch for **29**, **30** and **31**

18	GAA GGG CTT TTG AACTCTT **56**	Control for **32** and **33**
19	AAG AGT TCA AAA GCC CTTC **57**	Complementary for **32** and **33**
20	r(AAGAGU UCAAAAGCCCUUC) **58**	Complementary for **32** and **33**
21	AGAGTTCAA AAG CCC TTTT **59**	Control for **34** and **35**
22	AAAAGGGCTTTTGAACTCT **60**	Complementary for **34** and **35**
23	r(AAAAGGGCUUUUGAACUCU) **61**	Complementary for **34** and **35**
24	GCT TCC TTC CTT CG **62**	Control for **36**, **37**, **38** and **39**
25	CGA AGG AAG GAA GC **63**	Complementary for **36**, **37**, **38** and **39**

2.2.3.7 UV-Melting Studies of TANA-DNA Chimeric Oligonucleotides

The binding affinity of the TANA-DNA chimeric oligonucleotides (Table1) to their complementary nucleic acid sequences was investigated by measuring the melting temperatures (*Tm*) of the duplexes and triplexes. The *Tm* experiments of duplexes were carried out in phosphate buffer (pH 7.0) containing 100 mM NaCl. The TANA-DNA chimeric ONs were individually hybridized with the complementary DNA and RNA (Table 2) strands, to obtain duplexes.

The unmodified sequence GCT_8CG **40** was found to form complexes with both cDNA and RNA with higher melting temperature for ON: DNA complex over ON: RNA (ΔT_m= -8). Introduction a single thioacetamido linked **tst** dimer unit in **19** reversed this selectivity and the ON:RNA duplex **19:43** was more stable (ΔT_m= +8.6) than its complex

Table 3. *T*m(ºC) values of TANA ONs: DNA / RNA duplex

Entry	DNA Sequence	DNA **41**	RNA **43**	Δ*T*m (RNA-DNA)
1	5’ CGTTTTTTTTGC 3’ **40**	40	32.0	- 8.0
2	5’ CG TT**tst**TTT TGC 3’ **19**	23.7	32.3	8.6
3	5’ CGTT **tst** TT **tst** GC 3’ **20**	nd	50.0	-
4	5’ CG **tst tst tst tst** GC 3’ **21**	nd	47.81	-

with cDNA (**19:41**). A cumulative effect was observed in stabilizing the ON:RNA complex **20:43** and at the same time the complex with DNA **20:41** was destabilized, when two units of modified **tst** dimer were incorporated in the ON. Similarly, the ON sequence **21** with alternating phosphate-TANA linkers exhibited binding only with RNA 4**3** and no observable melting transition with complementary DNA **41** (Table 3, Figure 5).

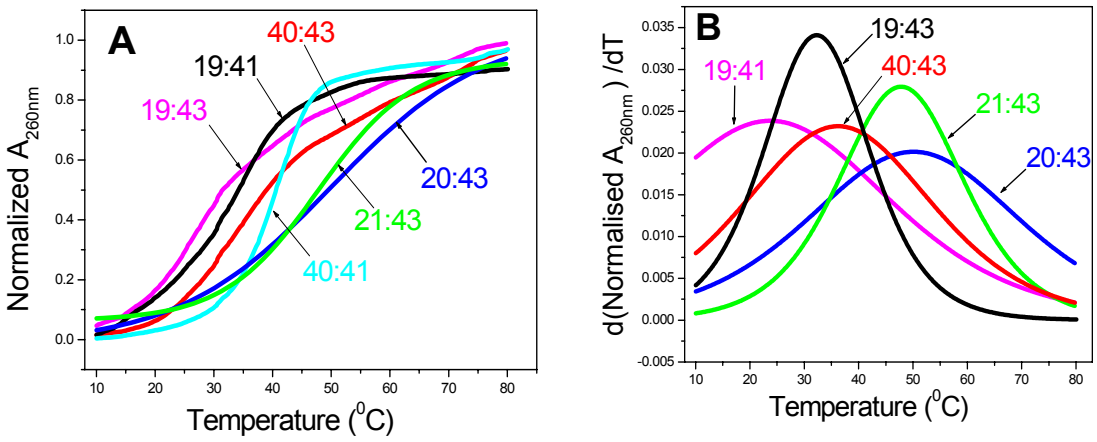


Figure 5. A. UV-melting Curves of **19**, **20**, **21** and **40** with complementary DNA **(41)** and cRNA **(43) B.** Corresponding first derivative Curves**.**

The unmodified mixed pyrimidine ON **45** formed complexes with either complementary DNA or RNA and exhibited almost equal binding strength. The preferential binding to RNA was consistently observed when one or two phosphate linkers were replaced by TANA either by incorporation of either one or two modified (**tst** or **cst)** dimer units, almost independent of their position in the ON. Inclusion of two or more modified units in ONs leads to the significant stabilization of their complexes with complementary RNA while complexes with the DNA counterpart did not show detectable transition. (Table 4). For sequences **22**, **23**, **25** and **26** there was no transition was observed with their complementary DNA. Sequence **24** having a single modification at the 5'- end the ON: RNA duplex **24:47** was found to be more stable (ΔT_m= +9) than its complex with complementary DNA **24: 46**.

Table 4. *T*m(ºC) values of TANA ONs: DNA / RNA duplex

Entry	Sequence	DNA **46**	RNA **47**	Δ*T*m (RNADNA)
1	5’ TCT CTT TCT T 3’ **45**	23.6	27.3	3.7
2	5’ TCT C **tst** TCT T 3’ **22**	nd	31	-
3	5’ TCT C **tst** TC **tst** 3’ **23**	nd	33.8	-
4	5’ T**cst** CTT TCTT 3’ **24**	18.5	27.5	9
5	5’ T**cst** CTT T**cst** T 3’ **25**	nd	32.6	-
6	5’ T**cst cst**T T**cst** T 3’ **26**	nd	39.7	-

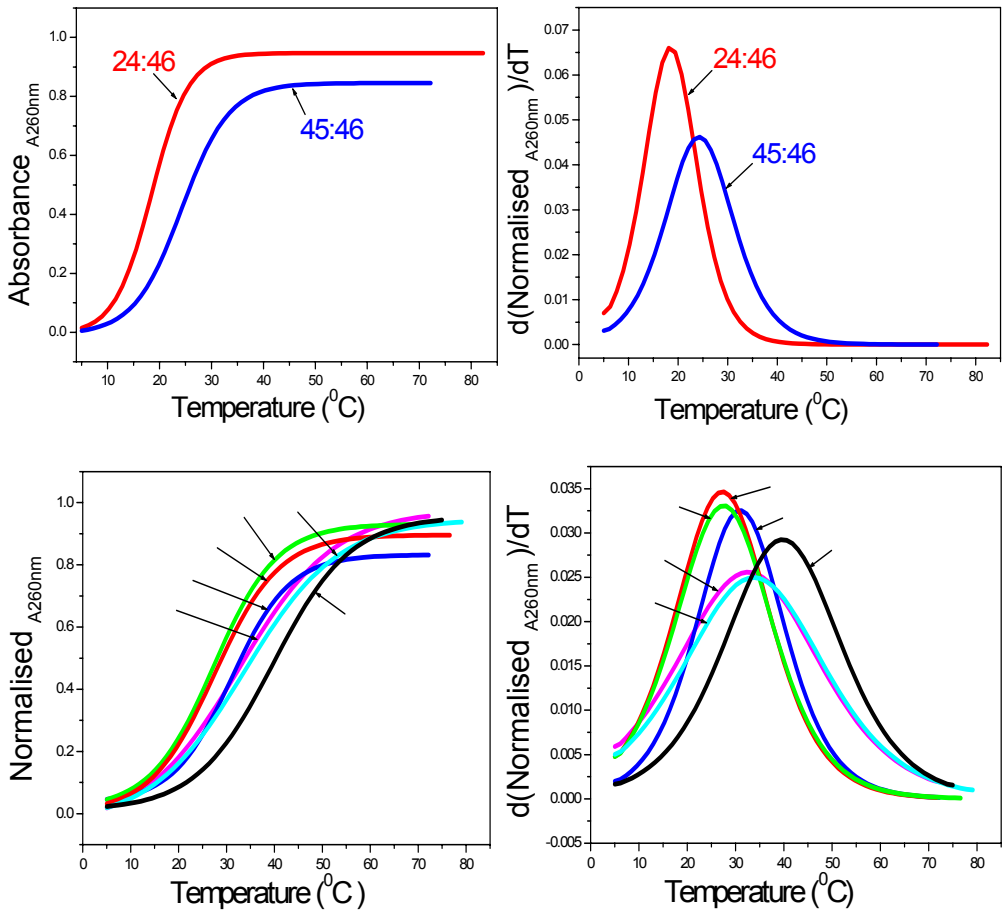


Figure 6. UV-melting Curves of **C**. **24** and **45** with complementary DNA **(46) D.** Corresponding derivative curves**. E. 22**, **23**, **24**, **25**, **26** and **45** with complementary RNA **(47) F.** Corresponding first derivative Curves**.**

To verify the usefulness of these modified units in mixed purine-pyrimidine sequence context, we synthesized the unmodified 18 mer sequence **52** (Table 5). The 18mer

sequence **52** was modified by introduction of one **tst** unit in **29**, two **tst** units in **30** and four **tst** and **cst** units in **31.**

The unmodified 18 mer **52** recognized both complementary DNA **53** and RNA **54** with equal affinity (ΔT_m= + 0.1). One **tst** unit caused destabilization of complexes with both

Table 5. *T*m(ºC) values of TANA ONs: DNA / RNA duplex

Entry	Sequence (5' to 3')	DNA 53	RNA **54**	$\Delta T_{m(RNA-DNA)}$
1	CCT CTT ACC TCA GTT ACA **52**	54.6	54.7	0.1
2	CCT C **tst** ACC TCA G TT ACA **29**	39.6	47.5	7.9
3	CCT C **tst** ACC TCA G **tst** ACA **30**	43.5	52.8	9.3
4	C **cst** C **tst** AC **cst** CAG **tst** ACA **31**	37.3	50.6	13.3

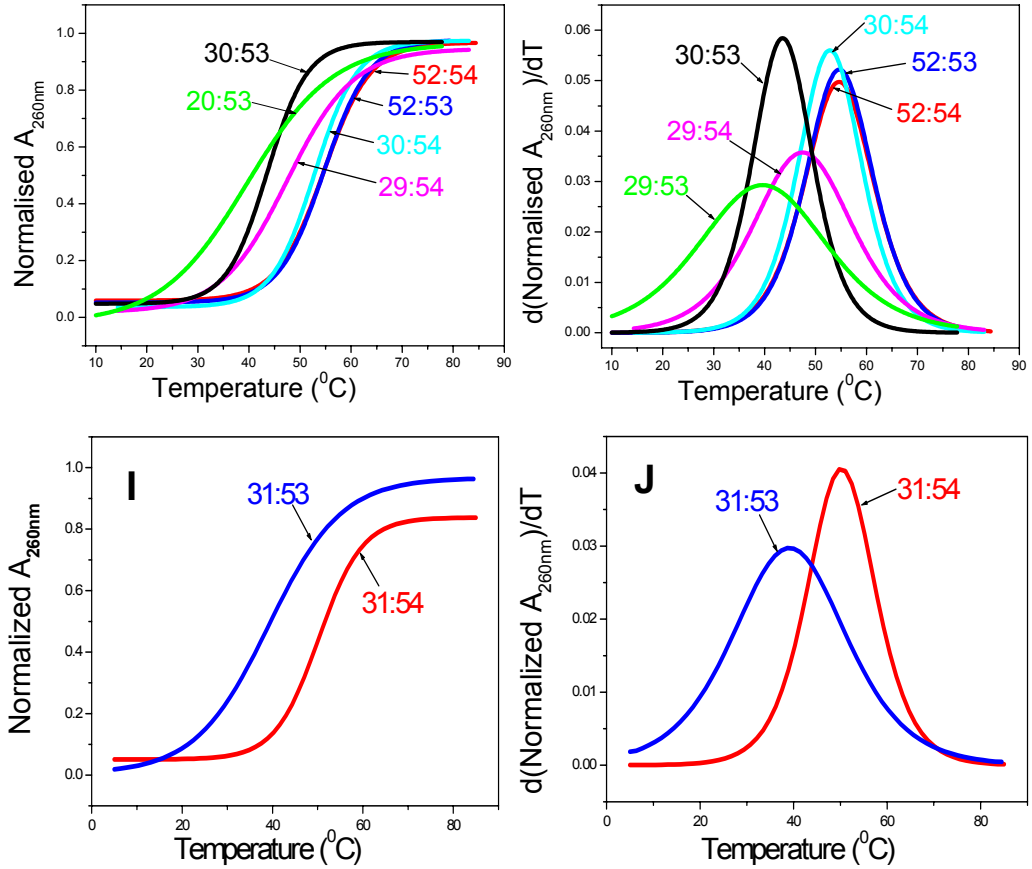


Figure 7. G & **I**. UV-melting Curves of **29, 30, 31** and **52** with complementary DNA (**53**) and RNA (**54**) **H & J.** Corresponding derivative curves**.**

DNA (**29:53**) and RNA (**29:54**), but to a much less extent with RNA than with DNA. The discrimination between RNA versus DNA recognition was observed with ΔT_m about 8 ^oC. Introduction of two **tst** units in **30** increased this stabilization of ON: DNA/RNA (**30:53**/**30:54**) complexes and RNA versus DNA discrimination to 9.3 ^oC. The increase of RNA versus DNA discrimination was further verified by introducing two **tst** and two **cst** units in **31** where the stabilization of ON: DNA/ RNA (**31: 53**/ **31:54**) complexes and RNA versus DNA discrimination was found to be 13.3 ^oC (Table 5).

Table 6. *T*m(ºC) values of TANA ONs: DNA / RNA duplex

Entry	Sequence(5' to 3')	ON :DNA	ON :RNA	$\Delta T_{m(RNA-DNA)}$
1	TCA CTA GAT G **48**	24.3 **48:49**	24.7 **48:50**	0.4
2	TCA **cst** A GAT G **27**	16.3 **27:49**	29.6 **27:50**	13.3
3	CTC TTC CTC CTT C **51**	46.1 **51:64**	-	-
4	CTC TTC CTC C **tst** C **28**	43.5(-2.6) **28:64**	-	-
5	GCT TCC TTC CTT CG **62**	- **62:63**	-	-
6	GC**tst** CC TTC CTT CG **36**	- **36:63**	-	-
7	GC**tst** CC **tst**C CTT CG **37**	- **37:63**	-	-
8	GC**tst** CC **tst**C C**tst** CG **38**	- **38:63**	-	-
9	G**cst** TCC TTC CTT CG **39**	- **39:63**	-	-

In a 10 mer ON **27**, a single **cst** unit caused stabilization of the ON:RNA complex **27:50** compared to control unmodified complex **48:50**, and destabilized the ON:DNA complex **27:49** (Table 6, entry 1 and 2, Figure 8)**.** Thus the modified units consistently destabilized the complex formation of the modified ONs with cDNA and stabilized the ON: RNA complexes.

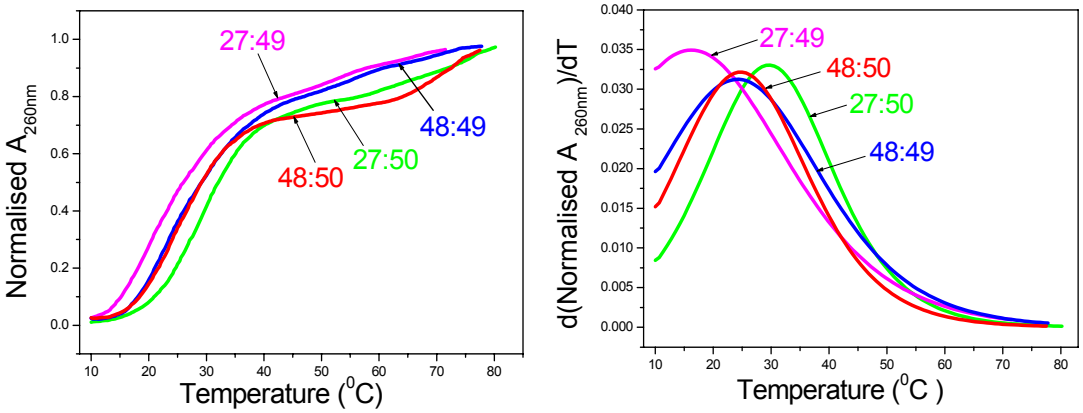


Figure 8. K. UV-melting Curves of **27** and **48** with complementary DNA (**49**) and RNA (**50**) **L.** Corresponding derivative curves

In the 13 mer ON **28**, a single **tst** unit caused destabilization of the ON: DNA complex **28:64** by 2.6ºC in comparison to the control **51:64** (Table 6, entry 3 and 4, Figure 9).

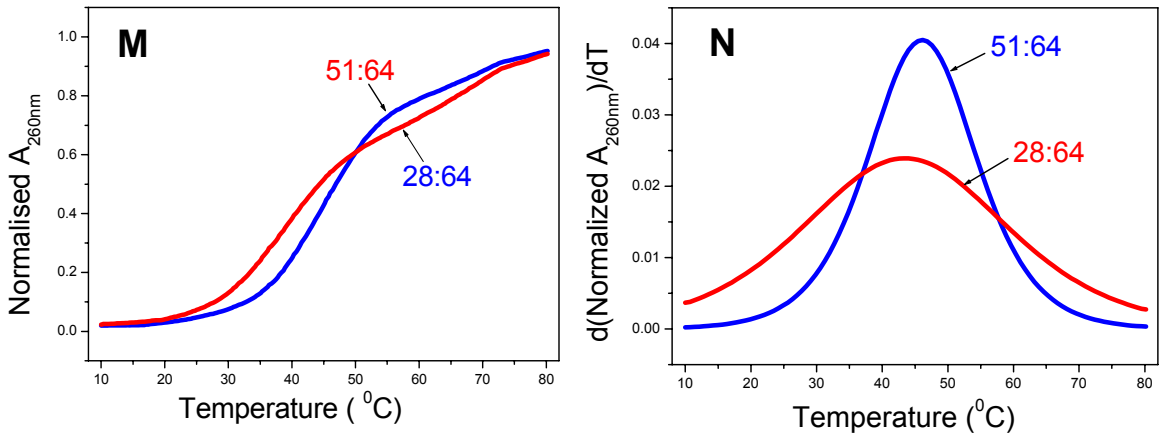


Figure 9. **M.**UV-melting Curves of **28** and **51** with complementary DNA (**64**) **N.** Corresponding derivative curves**.**

As we have seen that the chimeric TANA-ONs exhibit excellent hybridization affinity toward complementary RNA than DNA. So it is necessary to study the energetics, counterion effects and the hydration contribution of the incorporation of TANA modifications by a combination of spectroscopic and calorimetric techniques. To study the hybridization thermodynamics of TANA-based ONs, a series of 14 mer ONs **62**, **36**,

37 and **38** containing zero, single, double and triple **tst** TANA modifications respectively and ON **39** containing single **cst** TANA modification were synthesized (Table 6, entries 5-9). These studies are currently in progress and results will be reported elsewhere.

Chronic myeloid leukemia (CML) is probably the most extensively studied human malignancy. The discovery of the Philadelphia (Ph) chromosome in 1960 as the first consistent chromosomal abnormality associated with a specific type of leukemia was a breakthrough in cancer biology.[48] The ONs **56** is the antisense sequence and **59** is the sense sequence (Table 7, entry 1 and 4) for bcr/abl gene which is responsible for chronic myeloid leukemia. To study the antisense activity of the TANA-DNA ONs the ONs **32**, **33**, **34** and **35** (Table 7, entry 2, 3, 5 and 6) have been synthesized. The antisense activity and UV melting studies are currently in progress and results will be reported elsewhere.

Table 7. *T*m (ºC) values of TANA ONs: DNA / RNA duplex

Entry	Sequence (5' to 3')	DNA	RNA	$\Delta Tm_{(RNA-DNA)}$
1	GAA GGG CTT TTG AAC TCT T **56**	53.6	-	
		56: 57	**56: 58**	
2	GAA GGG **cst**T **tst**G AA **cst cst** T **32**	-	-	
		32: 57	**32: 58**	
3	GAA GGG **cst**T **tst**G AA CT **cst** T **33**	-	-	
		33: 57	**33: 58**	
4	AGA GTT CAA AAG CCC TTT T **59**	56.1	-	
		59: 60	**59:61**	
5	AGA G **tst** CAA AAG CCC **tst**T T **34**	45.5	-	
		34: 60	**34:61**	
6	AGA G **tst** CAA AAG CC**cst tst** T **35**	46.9	-	
		35: 60	**35: 61**	

2.2.3.8 Mismatch UV-melting studies

The thermal stabilities of TANA ONs with RNA containing single mismatch at the middle position of the sequences were studied (Table 8, Figure 10)

For ONs **19**, **20** and **21** having single, double and triple **tst** modifications, there is decrease of UV-melting temperature 9ºC, 16ºC and 22ºC respectively when they are hybridized with one single base mismatch RNA **44** (Table 8, entries 1-3, Figure 10).

Table 8. Mismatch UV-melting studies of modified ONs.

Entry	ON Sequences (5'to 3')	ON: cRNA T_m °C	ON:Mismatch RNA T_m °C
1	CG TT**tst**TTT TGC **19**	**19:35** 32.3	**19:44** 23.3 (-9)
2	CGTT **tst** TT **tst** GC **20**	**20:35** 50.0	**20:44** 33.9 (-16.1)
3	CG **tst tst tst tst** GC **21**	**21:35** 47.8	**21:44** 25.4 (-22.4)
4	CCT C **tst** ACC TCA G TT ACA **29**	**29:54** 47.5	**29:55** 35.7 (-11.8)
5	CCT C **tst** ACC TCA G **tst** ACA **30**	**30:54** 52.8	**30:55** 39.7 (-13.1)
6	C **cst** C **tst** AC **cst** CAG **tst** ACA **31**	**31:54** 50.6	**31:55** 32.3 (-18.3)

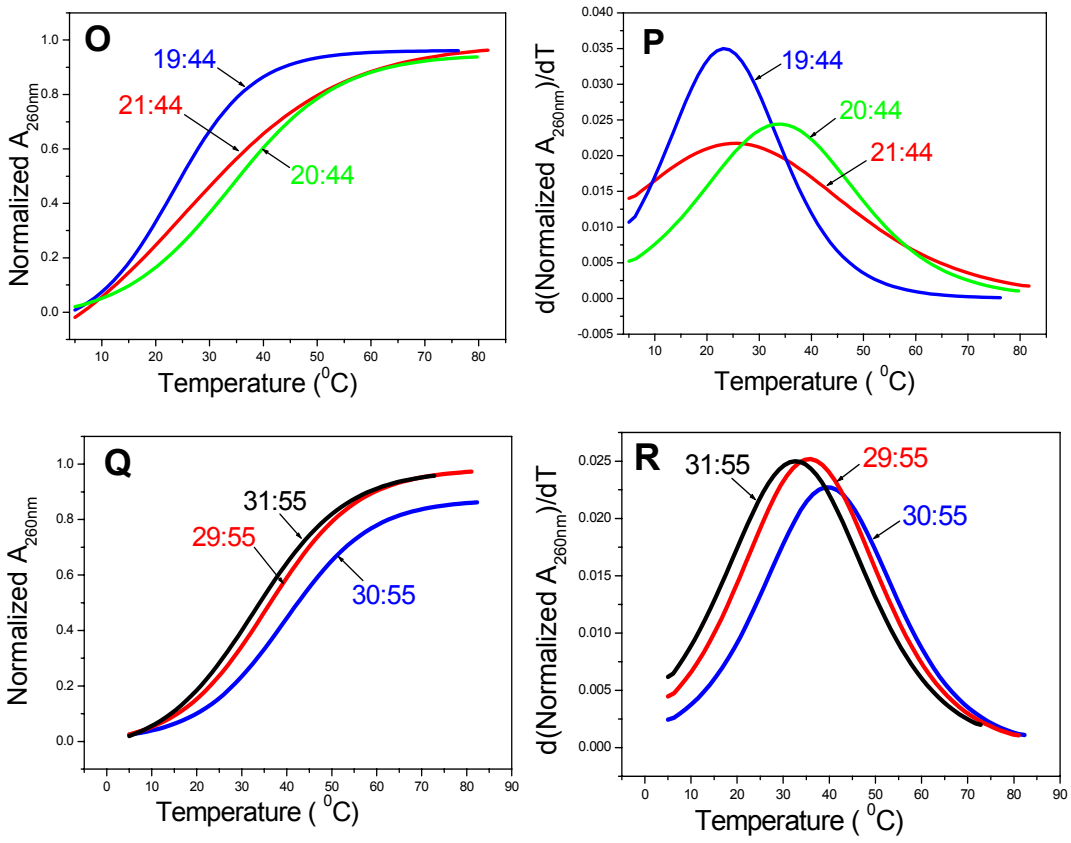


Figure 10. O & **Q**. Mismatch melting curves of ONs **19**, **20**, **21, 29, 30** & **31** with mismatch RNA **44** & **56** **P** & **R.** Corresponding derivative curves.

Similar effect of decrease of melting temperature was observed for the 18 mer ONs **29**, **30** and **31** with the mismatch RNA **56** (Table 8, entries 4-6, Figure 10). These results clearly indicate that the TANA ONs hybridize with their complementary RNA through Watson-Crick H-bonding.

2.2.4. Conclusion

Synthesis of thioaceamide nucleic acids **tst** and **cst** dimer phosphoramidites was done successfully and a fully automated solid-phase synthetic procedure for the incorporation of **tst** and **cst** dimer blocks into DNA sequences has been developed.

UV-melting studies of all the TANA-DNA chimeric oligomers show selective binding and more stability with RNA over DNA. The RNA binding selectivity increases with the number of modifications. The RNA selectivity of binding seems to be arising from the extended backbone linker that is probably inherently folded to be competent to bind to RNA over DNA as was found with the reported five-atom linked ON analogues.

As the **t**s**t** and **c**s**t** dimer blocks were found to be compatible in the DNA backbone to selectively stabilize the ON:RNA complexes. Further work to exploit their utility is currently in progress in our laboratory. The preferential sequence independent RNA binding ability of these evolved modified ONs will find applications in current antisense research and the studies are currently in progress.

Summary

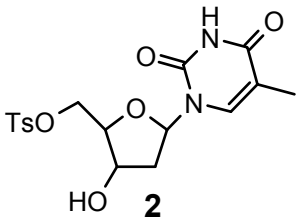
- Neutral, one atom extended thioacetamido nucleic acids (TANA) were designed for specific recognition of nucleic acids.
- The thioacetamido nucleic acids with preorganized geometry results in optimization of internucleobase distance complementarity.
- The TANA-DNA chimeric ONs are showing unprecedented RNA binding selectivity.

2.2.5 Experimental

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 spectrometer. Column chromatographic separations were performed using silica gel 60-120 and 230-400 mesh, Ethyl acetate, Petroleum ether, Dichloromethane and Methanol as the solvent system. ^{1}H and ^{13}C spectra were obtained using Bruker AC 200, 400 and 500 NMR spectrometers. The optical rotation values were measured on Bellingham-Stanley Ltd, ADP220 polarimeter.

Synthesis of 5'-*O*-tosyl thymidine 2

A solution of thymidine **1**, 2.42 g (10 mmol) in 30 ml anhydrous pyridine was cooled to 0 ˚C. Tosyl chloride (2.3 g, 12 mmol) dissolved in 10 ml of pyridine was then added dropwise during 4 hours. The reaction temperature was maintained at 0˚C during the addition. The reaction mixture was stirred for additional 4 h at room temperature. Pyridine was removed *in vacuo* and the residue was dissolved in 100 ml ethyl acetate and the organic layer was washed with 10% $NaHCO_3$ solution followed by water .The tosyl thymidine **2** slowly crystallized out. The product was filtered and dried under vacuum. Yield 3.2 g (80%).



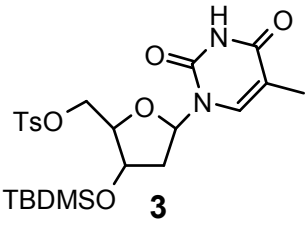
2

M.p. 160-164ºC. ^{1}HNMR: ($CDCl_3$ +1-drop DMSO-d6), (200 MHz) δ1.89 (s, 3H), 2.1-2.25 (m, 2H), 2.4 (s, 3H), 3.98-4.1 (m, 3H), 4.24 (m, 1H), 6.2(t, 1H), 7.33 (d, 3H), 7.78 (d, 2H), 10.14 (s, 1H).

Mass calculated 396.39 observed 396.46.

Synthesis of 5'*O*-tosyl, 3'*O-tert* butyl dimethyl silyl thymidine 3

A solution of 5'*O*- tosyl thymidine **2** (3.1g, 7.8 mmol), *tert*-butyl dimethyl silyl chloride (1.4 g, 9.4 mmol), imidazole (1.33 g, 19.6 mmol) in 10 ml anhydrous DMF was stirred for 6 h. DMF was removed the *in vacuo* and the residue was dissolved in 200 ml ethyl acetate. The organic layer was washed with water (3 x 50 ml) followed by brine (2 x 30 ml). The organic layer was dried



3

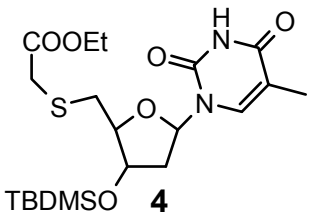
over anhydrous Na_2SO_4 and solvent was then removed *in vacuo*. The residue was redissolved in 20 ml DCM and subjected to filtration column purification (1:1 Pet Ether: Ethyl acetate). Pure compound **3** was obtained in 95% yield, 3.8 g.

^{1}HNMR: ($CDCl_3$), (200 MHz) δ 0.06 (s, 6H), 0.86 (s, 9H) 1.89 (s, 3H), 2.1 (m, 1H), 2.29 (m,1H), 3.98-4.1 (m, 3H), 4.24 (m,1H), 6.2 (t, 1H), 7.33 (d, 3H), 7.8 (d, 2H), 9.7 (s, 1H).

Mass calculated 511.2 observed 511.87

Synthesis of ethyl-[*S*- 5'-(3'-*O*-TBDMS- Thymidinyl)-mercapto]-acetate 4

Solution of NaH (0.35 g, 60% solution in hexane, 8.7 mmol.) and ethyl mercapto acetate (1 ml, 8.7 mmol) in 5 ml anhydrous DMF was stirred for 30 minutes. 5'*O*-tosyl, 3'*O-tert* butyl dimethyl silyl thymidine **3** (3.7 g, 7.23 mmol) dissolved in 5 ml DMF was then added slowly. The reaction mixture was stirred for 1 h at RT. DMF was removed under reduced pressure, and the residue was dissolved in 200 ml ethyl acetate. The organic layer was washed with water (3 x 50 ml) followed by washing with brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and solvent removed *in vacuo*. The product was subjected to column chromatography purification (1:1 Pet Ether: Ethyl acetate) to give pure **4**, 2.75 g, 80 %.



4

M.p. 122-124 ºC $[\propto]_{25}^{D}$= +19.4˚ (c 0.5, $CHCl_3$).

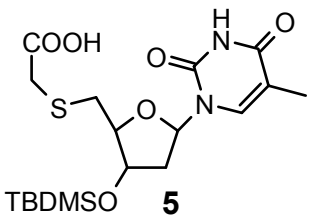
^{1}HNMR: ($CDCl_3$), (400MHz) δ 0.05 (s, 6H), 0.86 (s, 9H) 1.9 (s, 3H), 2.09 (m,1H), 2.26 (m,1H), 2.9-3 (m, 2H), 3.22-3.31 (q, 2H), 4.0 (q, 1H), 4.16 (q, 2H), 4.26 (m,1H) 6.2 (t, 1H), 7.33 (d,1H), 9.4 (b, 1H)

^{13}C NMR (100MHz) δ -4.8, -4.7, 12.5, 14.1, 17.9, 25.7, 34.2, 34.3, 40.5, 61.5, 73.5, 84.7, 85.7, 111.2, 135.4, 150.1, 163.7, 170.0.

Mass calculated 458.64 observed 497.4 ($+K^+$).

Synthesis of *S*- 5'-(3'-*O*-TBDMS-thymidinyl)-mercapto-acetic acid 5

5 ml 2M NaOH solution was added to the solution of **4** (2.6 g, 5.47 mmol) in 10 ml methanol. The reaction mixture was stirred for 30 minutes. The sodium salt of the acid and excess NaOH present in the solution was neutralized by DOWEX 50 H^+ resin. The



5

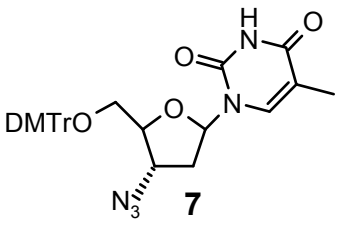
resin filtered and washed with 2:1 mixture of methanol:water. The filtrate was concentrated under vacuum and the dried by vacuum desiccation to give **5,** 2.3 g, (95%)

M.p. 134-136ºC. ^{1}H NMR (D_2O) δ 0.06 (s, 6H), 0.85 (s, 9H), 1.66 (s, 3H), 2.16 (m, 2H), 2.71-2.76 (m, 2H), 3.22 (q, 2H), 3.88 (dd, 1H), 4.2 (m, 1H), 5.99 (t, 1H), 7.3 (s, 1H).

Mass calculated 430.6 observed 453.11(+ Na^+).

Synthesis of 3'-deoxy-3'-azido- 5'-*O*-dimethoxytrityl- thymidine 7

A mixture of AZT **6** (2.68 g, 10 mmol), dimethoxytrityl chloride (4 g, 12 mmol), DMAP (1 mmol), triethylamine (7 ml, 50 mmol) in 50 ml anhydrous pyridine was stirred at RT for 6 h. Pyridine was removed *in vacuo* and residue redissolved in 200 ml DCM. The organic layer was washed with 10 % $NaHCO_3$ solution (2 x 30 ml) followed by water (2 x 50 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated to dryness. Column purification using 1% methanol-dichloromethane gave pure **7,** 5.1 g, 90%.

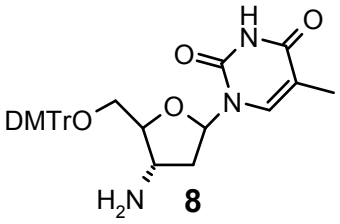


7

IR, ν(cm^{-1}) ($CHCl_3$) ; 3128, 3019, 2110, 1724, 1670 cm^{-1} ^{1}HNMR: ($CDCl_3$, 200 MHz) δ 1.51 (s, 3H), 2.4-2.5 (m, 2H), 3.3-3.5 (m, 2H), 3.8 (s, 6H), 4.0 (d, 1H), 4.4 (q, 1H), 6.25 (t, 1H), 6.8-8.6 (m,14H) Mass calculated 569.6 observed 568.3.

Synthesis of 3'-deoxy-3'-amino 5'-*O*-dimethoxytrityl- thymidine 8

Compound **6** (1.5 g, 2.6 mmol) was dissolved in 10 ml methanol and to it was added 0.15 g (10%) Pd-C catalyst. Then mixture was subjected to catalytic hydrogenation at 40 Psi of hydrogen pressure for 3.5 h. The catalyst was removed by filtration over celite and concentration of the filtrate *in vacuo* give **8,** 1.3 g, 90%.

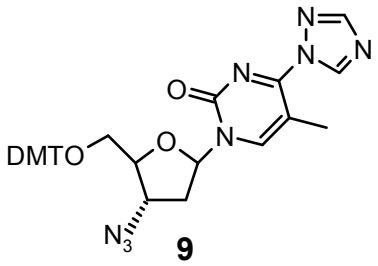


8

IR, ν(cm^{-1}) ($CHCl_3$) 3398, 3018, 2924, 2853, 1701, and 1686. ^{1}HNMR: ($CDCl_3$, 200 MHz) δ 1.52 (s, 3H), 2.17-2.3 (m, 2H), 3.38 (dd, 1H), 3.48 (dd, 1H), 3.7-3.8 (m, 2H) 3.8 (s, 6H), 6.29 (t, 1H), 6.86 (d, 4H), 7.2-7.5 (m, 9H), 8.64 (s, 1H). Mass calculated 543.6 observed 543.7.

Synthesis of 3'-deoxy-3'-azido-5'-*O*-dimethoxytrityl-C^4-(-1, 2, 4-triazol-1-yl)-pyrimidine 2'-deoxyribonucleoside 9

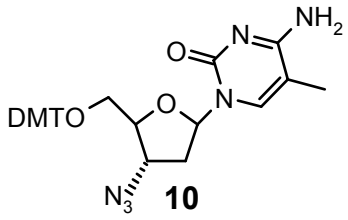
Triethylamine (12.7 ml, 91 mmol) was added dropwise to a stirred mixture of 1, 2, 4-triazole (7 g, 100 mmol) and phosphoryl chloride (2 ml, 22 mmol) in 50ml CH_3CN at 0ºC. The solution of **7** (3.5 g, 9.1 mmol) in 15 ml dry CH_3CN was then added dropwise. The reaction mixture was stirred at room temperature for 2.5 h. Triethylamine (4.4 ml) and water (1.1 ml) were added at 0ºC and the mixture was stirred for another 10 minutes. The solvent was evaporated and the residue was re-dissolved in 50 ml ethyl acetate. The organic layer washed with water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml), dried over Na_2SO_4 and evaporated to dryness to give 3.6 g (92%) **9** as yellow foam, which was used immediately for the next reaction.



9

Synthesis of 2',3'-dideoxy-3'-azido-5'-dimethoxytrityl-5-methylcytidine 10

Concentrated aqueous ammonia (10 ml) was added to the solution of **9** (3.5 g, 8 mmol) in 50 ml dioxane. The reaction mixture was stirred at room temperature for 2.5 hours. The solvents were evaporated *in vacuo* and the residue redissolved in CH_2Cl_2 and purified by silica-gel column chromatography (eluted with 0-5% methanol in CH_2Cl_2) to afford **10** (3 g, 96%) as white foam.



10

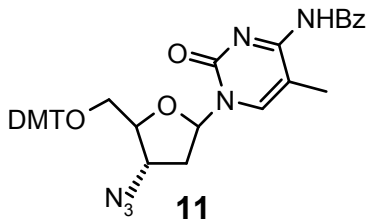
IR, ν (cm^{-1}) 3192, 2984, 2827, 2114, 1687.6 ^{1}HNMR: ($CDCl_3$, 200MHz) δ 1.52 (d, 3H), 2.44 (m, 1H), 2.6 (m, 1H) 3.33 (dd, 1H), 3.56 (dd, 1H), 3.8 (s, 6H) 4.0 (m, 1H), 4.3 (m, 1H), 6.25 (t, 1H), 6.84 (d, 4H), 7.3-7.5 (m, 9H), 7.75 (s, 1H).

Mass Calculated 568.6 observed 568.4.

5'-*O*-dimethoxy-trityl-N^4-benzoyl-2', 3'-dideoxy-3'-azido -5- methylcytidine 11

Benzoyl chloride (2.6 ml, 22.7 mmol) was added drop wise slowly to a solution of **10** (2.9 g, 7.56 mmol) in 40 ml pyridine at 0ºC. The mixture was then stirred at room temperature overnight. Then 8 ml water were added and the reaction mixture was stirred

for 5 min at 0ºC, followed by addition of concentrated ammonia (8 ml) at 0ºC and stirring for another 30 minute. The solvent was evaporated *in vacuo* and the residue redissolved in 50 ml ethyl acetate. Washed the organic layer with 10 % $NaHCO_3$ solution (3 x 30 ml), water (2 x 50 ml) and saturated aqueous NaCl (2 x 20 ml). The organic layer was then dried over Na_2SO_4 and evaporated to dryness. The residue was redissolved in 10 ml DCM and purified by silica gel (60-120 mesh) column filtration (10 % ethyl acetate in petroleum ether) to afford (2.9 g, 76 %) **11** as white foam.

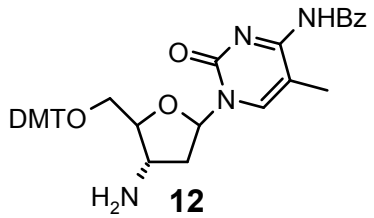


11

IR, ν(cm^{-1}) ($CHCl_3$) 3189, 3024, 2110, 1715, 1705, 1696, 1685 ^{1}HNMR: ($CDCl_3$, 200 MHz) δ1.7 (d, 3H), 2.5 (m, 2H), 3.36 (dd, 1H), 3.58 (dd, 1H), 3.8 (s, 6H), 4.0 (m,1H), 4.3 (m,1H), 6.28 (t, 1H), 6.87 (m, 4H), 7.3-8.3 (m, 15H). Mass calculated 673.33 observed 673.31, 695.27 (+ Na^+).

5'-*O*-dimethoxy-trityl-N^4-benzoyl-2',3'-dideoxy-3'-amino-5-methylcytidine 12

Compound **11 (**2.8 g, 4.1 mmol) was dissolved in 15% triethylamine in pyridine (5 ml) and H_2S gas was bubbled into the solution at 0º C for 15 minutes. The solution was then stirred at room temperature for an additional 30 minutes and solvent removed *in vacuo*. The residue was purified by flash silica gel (200-400 mesh) column chromatography (2 - 5 % methanol in DCM) to afford **12**, 2.2 g (80%).



12

IR, ν (cm^{-1}) ($CHCl_3$) 3189, 3024, 2810, 1718, 1709, 1692, 1683.

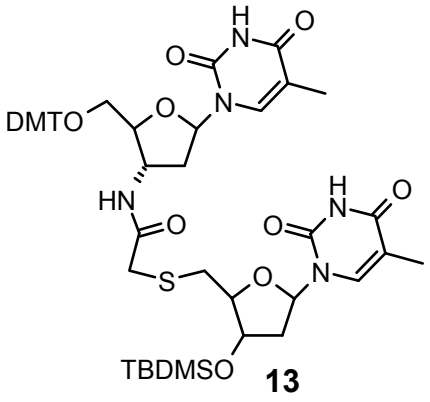
^{1}HNMR: ($CDCl_3$, 200MHz) δ1.71 (d, 3H), 2.26 (m, 1H), 2.4(m, 1H), 3.38(dd, 1H), 3.56(dd, 1H), 3.76 (m, 2H), 3.8 (s, 6H), 6.24 (t, 1H), 6.87 (m, 4H), 7.32-8.3 (m, 15H).

Mass calculated 647.33, observed 647.32, 669.29 (+ Na^+).

Synthesis of 5'-*O*-DMT -tst-3'-*O*-TBDMS Dimer 13

To compound **5** (1 g, 2.32 mmol) in dry acetonitrile (5 ml), HBTU (1 g, 2.8 mmol), DIPEA (1.2 ml, 7 mmol) and HOBt (0.16 g, 1.2 mmol) were added and stirred for 15 min. compound **8** (1.25 g, 2.3 mmol) was dissolved in 3 ml acetonitrile and was then

added into the reaction mixture and further stirred at room temperature for 1h. The reaction mixture was concentrated to dryness, dissolved in ethyl acetate (30 ml) and washed with 5% $NaHCO_3$ solution (2 x 10 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated to get the crude product. This was purified by column chromatography using CH_2Cl_2/ MeOH to get pure product **13,** 1.7 g, 78%.



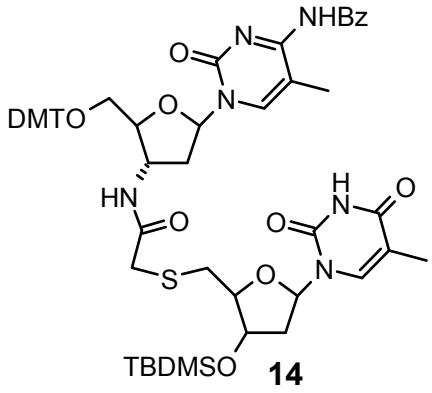
13

^{1}HNMR: ($CDCl_3$, 200 MHz) δ 0.07 (s, 6H), 0.87 (s, 9H), 1.82 (s, 3H),1.84 (s, 3H), 2.3-2.4 (m, 4H), 2.8-2.96 (m, 2H), 3.3 (q, 2H), 3.66-3.82 (m, 2H), 3.94 (m, 1H), 3.99 (m, 1H), 4.36 (m, 1H), 4.41 (m,1H), 6.13-6.2 (m, 2H), 6.8-6.84 (d, 4H), 7.17-7.69 (m, 11H), 8.63 (d, 2H).

Mass calculated 955.96 observed 979.0 ($+Na^+$), 994.23 ($+K^+$).

Synthesis of 5'-*O*-DMT- cst-3'-*O*-TBDMS Dimer 14

To compound **5** (1 g, 2.32 mmol) in dry acetonitrile (5 ml), HBTU (1 g, 2.8 mmol), DIPEA (1.2 ml, 7 mmol) and HOBt (0.16 g, 1.2 mmol) were added and stirred for 15 min. Compound **12** (1.5 g, 2.3 mmol.) dissolved in 4 ml acetonitrile was then added into the reaction mixture and further stirred at room temperature for 1h. The reaction mixture was concentrated to dryness, dissolved in ethyl acetate (30 ml) and washed with 5% $NaHCO_3$ solution (2 x 10 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated to get the crude product. This was purified by column chromatography using CH_2Cl_2 / MeOH to get pure product **15**, 2 g, 80%.



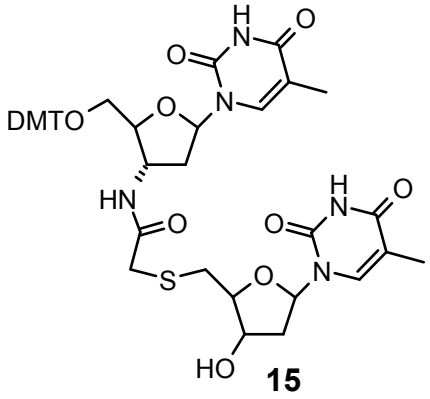
14

^{1}HNMR: ($CDCl_3$, 200 MHz) δ 0.08 (s, 6H), 0.88 (s, 9H), 1.58 (s, 3H), 1.9 (s, 3H), 2.15-2.34 (m, 2H), 2.46-2.53 (m, 2H), 2.85 (m, 2H), 3.28 (q, 2H), 3.47 (m, 2H), 3.96 (m, 1H), 4.1(d, 1H), 4.29 (m, 1H), 4.76 (m, 1H), 6.06 (t, 1H), 6.4 (t, 1H), 6.82-6.87 (d, 4H), 7.24-

7.72 (m, 15H), 7.82 (s,1H), 8.25 (s,1H), 8.29 (s,1H). Mass calculated 1059.18 observed 1081.44 (+ Na^+).

Synthesis of 5'-*O*-DMT tst-3'-OH Dimer 15

A solution of **13** (1.6 g, 1.7 mmol) and TBAF (0.65 g, 2.5 mmol) in 15 ml anhydrous THF was stirred at RT for 1h. THF was evaporated *in vacuo* and residue re-dissolved in DCM (50 ml). The solution was washed with water (2 x 20 ml) followed by brine. The organic layer was dried over anhydrous Na_2SO_4 and evaporated to dryness. Silica gel column chromatography gave the pure product **15**, 1.25 g, 90 %.

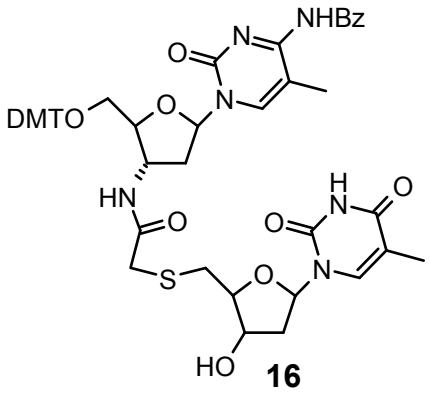


15

^{1}H NMR ($CDCl_3$, 400MHz) δ 1.4 (s, 3H), 1.9 (s, 3H), 2.27 (m, 1H), 2.4 (m, 1H), 2.46 (m, 2H), 2.92 (m, 2H), 3.32 (q, 2H), 3.46 (m, 2H), 3.79 (s, 6H), 4.08-4.14 (m, 2H), 4.38 (m, 1H), 4.75 (m, 1H), 6.2 (t, 1H), 6.4 (t, 1H), 6.84 (d, 4H), 7.2-7.8 (m, 9H), 7.64 (s, 1H), 8.0 (b, NH).

Mass calculated 841.93 observed 841.4.

Synthesis of 5'-*O*-DMT-cst -3'-OH Dimer 16



16

A solution of **14** (1.8 g, 1.67 mmol) and TBAF (0.67 g, 2.54 mmol) in 15 ml anhydrous THF was stirred at RT for 1h. The solvent was evaporated *in vacuo* and residue was re-dissolved in DCM (50 ml). The solution was washed with water (2 x 20 ml) followed by brine. The organic layer was dried over anhydrous Na_2SO_4 and evaporated to dryness. Silica gel column chromatography gave the pure product **16**, 1.48g, 92%.

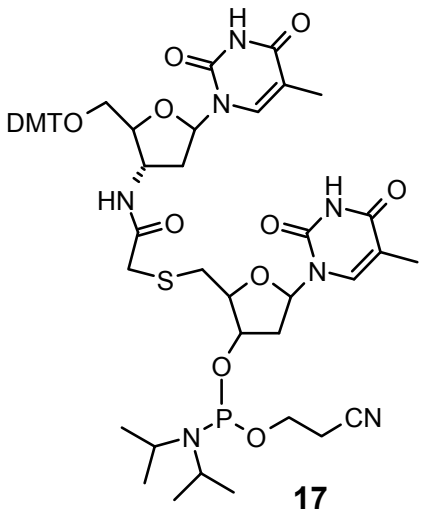
^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.6 (s, 3H), 1.9 (s, 3H), 2.35 (t, 2H), 2.5 (m, 2H), 2.8-2.9 (m, 4H), 3.5 (m, 2H), 3.8 (s, 6H), 4.02 (m, 1H), 4.1 (m, 1H), 4.4 (m, 1H), 4.73 (m, 1H),

6.1 (t, 1H), 6.4 (t, 1H), 6.85 (d, 4H), 6.8 (d, 4H), 7.2-7.53 (m, 13H), 7.8 (s, 1H), 8.25 (d, 2H).

Mass calculated 945.05 observed 968.2(+Na^+).

5'-*O*-(4, 4'-dimethoxy) trityl -tst- 3'-*O*-(2-cyanoethyl-N, N-diisopropylphosphoramidite)-dimer 17

Compound **15** (0.9 g, 1.07 mmol) was dissolved in dry DCM (10 ml) followed by the addition of diisopropyl ethyl amine (0.3 ml, 1.6 mmol) and chloro (2-cyanoethoxy) (*N, N* diisopropylamino)-phosphine (0.3 ml, 1.3 mmol) and the reaction mixture was stirred at room temperature for 2 h. The contents were then diluted with dry DCM and washed with 5% $NaHCO_3$ solution. The organic phase was dried over anhydrous Na_2SO_4 and concentrated to foam. The residue was dissolved in DCM and precipitated with hexane to obtain **17** (0.73 g, 65%). The phosphoramidite **17** was dried overnight over P_2O_5 and KOH in a desiccator before applying on DNA synthesizer. TLC shows two close moving spots for two diasteromers (R_f =0.5, 2% methanol-dichlromethane)



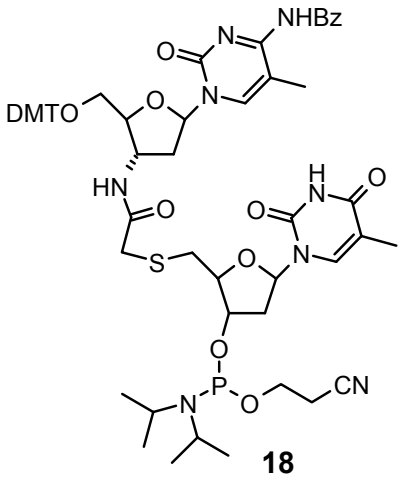
17

^{31}P NMR ($CDCl_3$) δ 149.0, 149.08. ^1H NMR ($CDCl_3$, 400 MHz) δ 1.2 (d, 12H), 1.41 (s, 3H), 1.92 (s, 3H), 2.47 (m, 4H), 2.47 (s, 2H), 2.78 (m, 2H), 2.96 (m, 2H), 3.29(m, 2H), 3.5-3.61(m, 6H), 3.73(m, 1H), 3.8 (s, 6H), 3.85 (m, 1H), 4.13-4.24 (m, 4H), 4.5 (m, 1H), 4.7 (m, 1H), 6.1(t, 1H), 6.4 (t, 1H), 8.85 (d, 4H), 7.2-7.4 (m, 10H), 7.6 (s, 1H).

Mass calculated 1042.1 observed 1041.5.

5'-*O*-(4, 4'-dimethoxy) trityl-cst-3'-*O*- (2-cyanoethyl-N, N-diisopropylphosphoramidite) dimer 18

Compound **16** (1g, 1.05 mmol) was dissolved in dry DCM (10 ml) followed by the addition of diisopropyl ethyl amine (0.3 ml, 1.6 mmol) and chloro (2-cyanoethoxy) (*N, N*-diisopropylamino)-phosphine (0.3 ml, 1.4 mmol) and the reaction mixture was stirred at room temperature for 2 h. The contents were then diluted with dry dichloromethane and washed with 5% $NaHCO_3$ solution. The organic phase was dried over anhydrous



18

Na_2SO_4 and concentrated to foam. The residue was dissolved in DCM and precipitated with hexane to obtain **18** (0.72 g, 60%). The phosphoramidite **18** was dried overnight over P_2O_5 and KOH in a desiccator before using on DNA synthesizer. TLC shows two close moving spots for two diastreomers (R_f =0.6, 2% methanol-dichlromethane)

^{31}P NMR ($CDCl_3$) δ 148.91, 147.03. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 1.2 (d, 12H), 1.6 (s, 3H), 1.9 (d, 3H), 2.4-2.5 (m, 4H), 2.7 (m, 2H), 2.8 (m, 1H), 2.9-3.0 (m, 2H), 3.2-3.35 (m, 2H), 3.45-3.65 (m, 5H), 3.73 (m, 1H), 3.8 (s, 6H), 3.9 (m, 1H), 4.1-4.3 (m, 3H), 4.5 (m, 1H), 4.8 (m, 1H), 6.1(q, 1H), 6.4 (t, 1H), 6.85 (d, 4H), 7.2-7.5 (m, 13H), 7.8 (s, 1H), 8.3 (d, 2H).

Mass calculated 1145.3 Observed 1145.06.

Oligomer Synthesis:

Base protected standard nucleoside phosphoramidites (A, T, C, G) and nucleoside derivatized controlled pore glass supports were used. The DNA synthesis was carried out on an Applied Biosystems 3900 DNA Synthesizer at 40 nmol scale. Dry solvents were used for synthesis. The commercially available amidites (0.1M) were dissolved in dry acetonitrile while 0.15 M solution was prepared for dimer phosphoramidite **17** and **18** and 4A molecular sieves were added to it to remove traces of moisture. The solid phase synthesis protocol is summarized in Figure 2. The oligomers were cleaved from the support by treatment with concentrated aqueous NH_3 for 8h. The cleaved oligomers were subjected to initial gel filtration and then purified by RP-HPLC.

HPLC Analysis

To check the purity of unmodified and chimeric oligonucleotides, Merck- Lichrosphere Lichrocart 100RP-18 (250 x 4 mm) endcapped (5 μm) column was used. A gradient elution method with A to B in 20 min. was used, where buffer A was 5%CH_3CN in 0.1M TEAA (triethyl ammonium acetate) and buffer B was 30% CH_3CN in 0.1M TEAA with

flow rate 1.5ml/min. The HP 1050 multiwavelength UV detector was used which was set at 254nm.

MALDI-TOF mass spectrometry

Mass spectral analysis was performed on a 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 matrixes used for analysis were THAP (2', 4', and 6'-trihydroxyacetophenone) for ONs having mass < 5000 and HPA (3-Hydroxy picolinic acid) for ONs having mass >5000. Diammonium citrate was used as additive. The samples were prepared by mixing 10 µl of matrix (0.5 M solution), 5 µl of diammonium citrate (0.1 M) and 1 µl ONs (OD 5-10/ml or 150-300 µg / ml) and spotted 1 µl on a stainless steel plate for analysis.

UV-*T*m measurements

The concentration was calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases (i.e., T, 8.8 cm^2 / µmol; C, 7.3 cm^2 / µmol; G, 11.7 cm^2 /µmol and A, 15.4 cm^2 /µmol). The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM) and were annealed by keeping the samples at 90°C for 5 minutes followed by slow cooling to room temperature (annealing). Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Perkin-Elmer *Lambda 35 UV-VIS* spectrophotometer scanning from 5 to 85°C at ramp rate of 0.5°C per minute. The data were processed using Micro cal Origin 6.0 and *T*m values derived from the derivative curves.

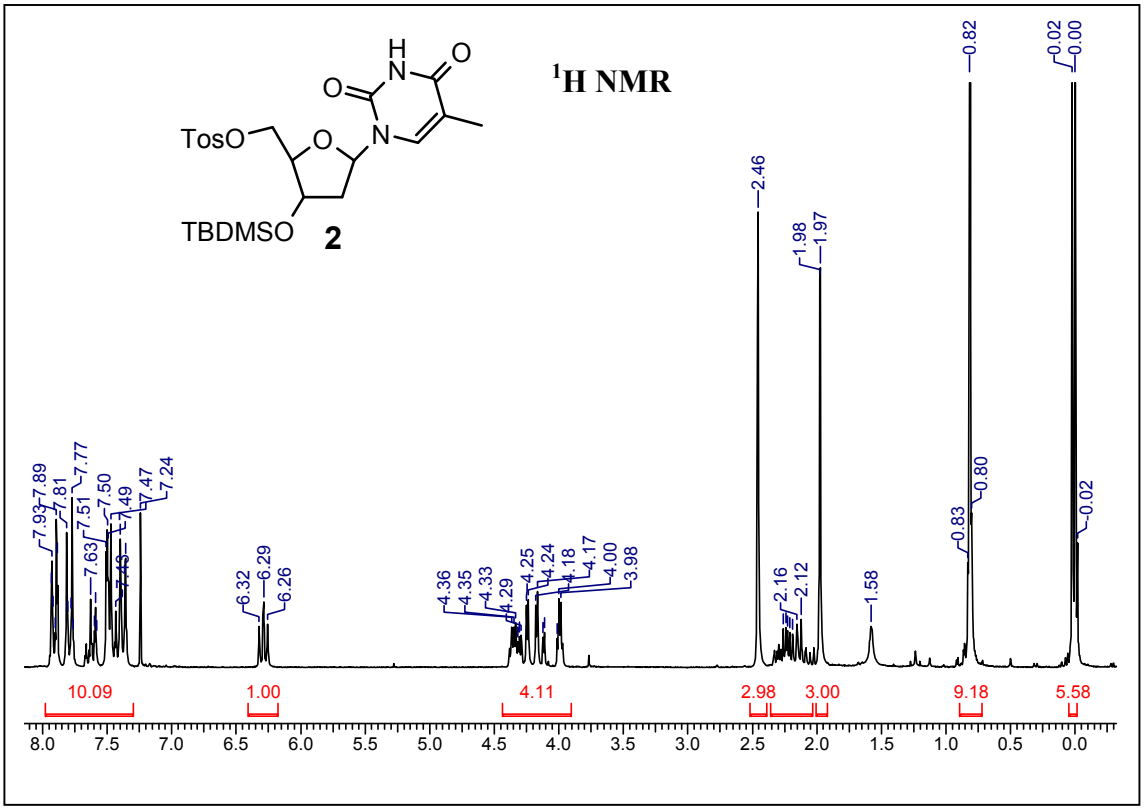
2.2.6 Appendix

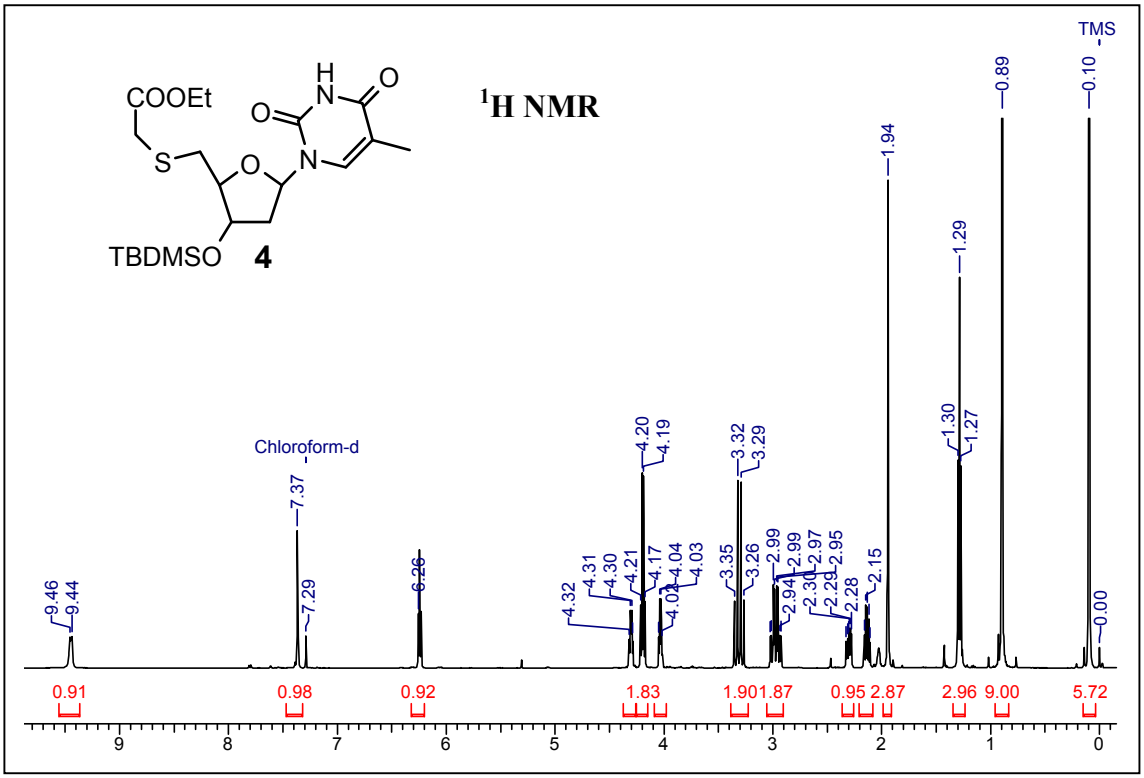
^{1}H NMR of compound **2, 4**, **13**, **14**, **15**, **16**, **17**& **18**

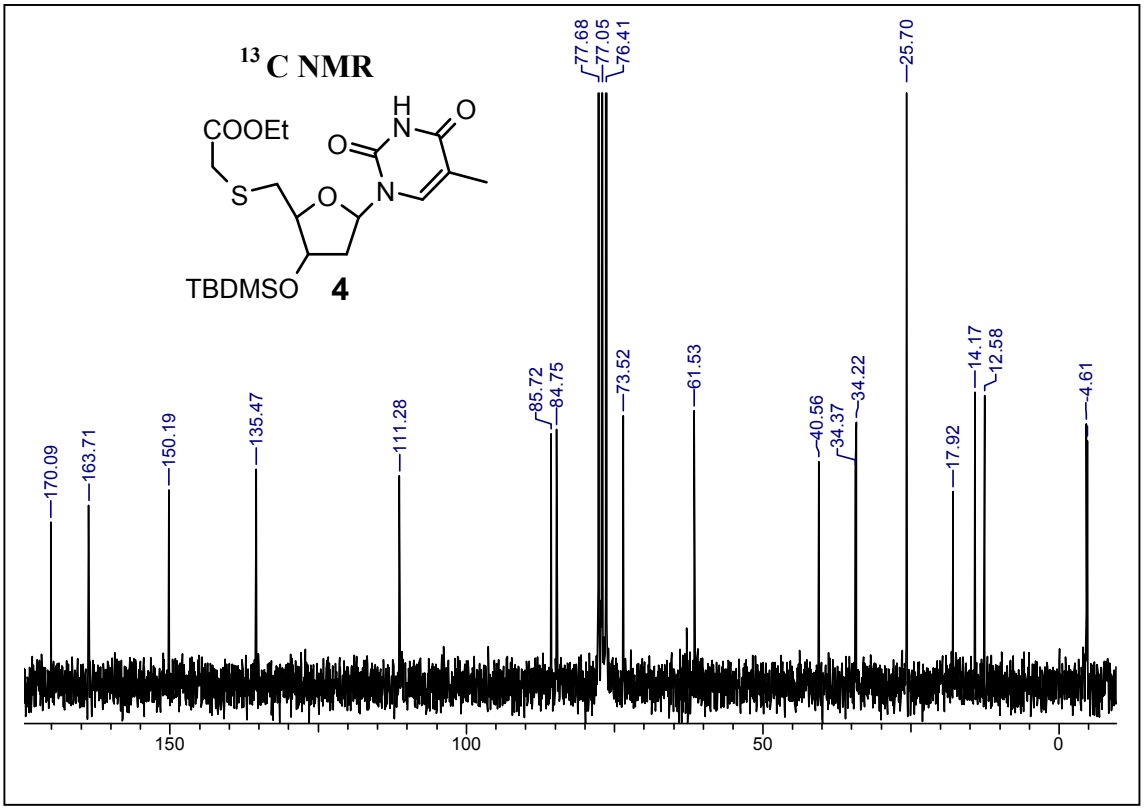
^{13}C NMR of compound **4**

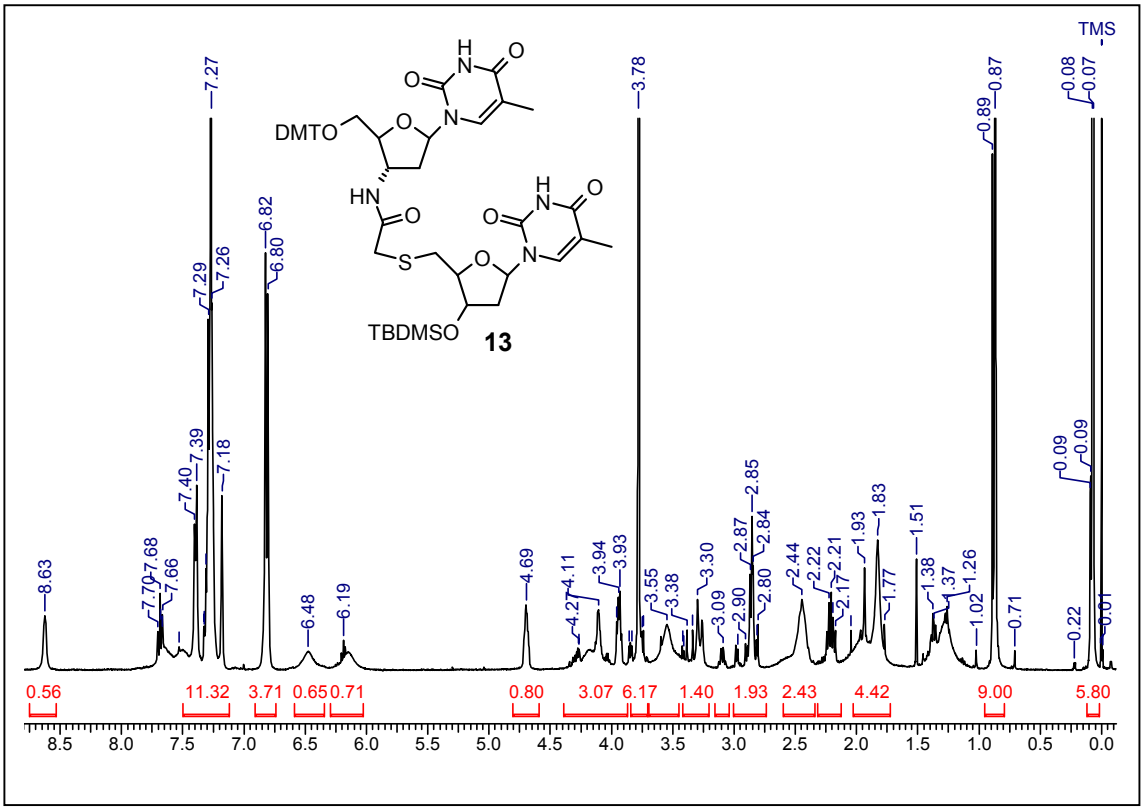
^{31}P NMR of compound **17** & **18**

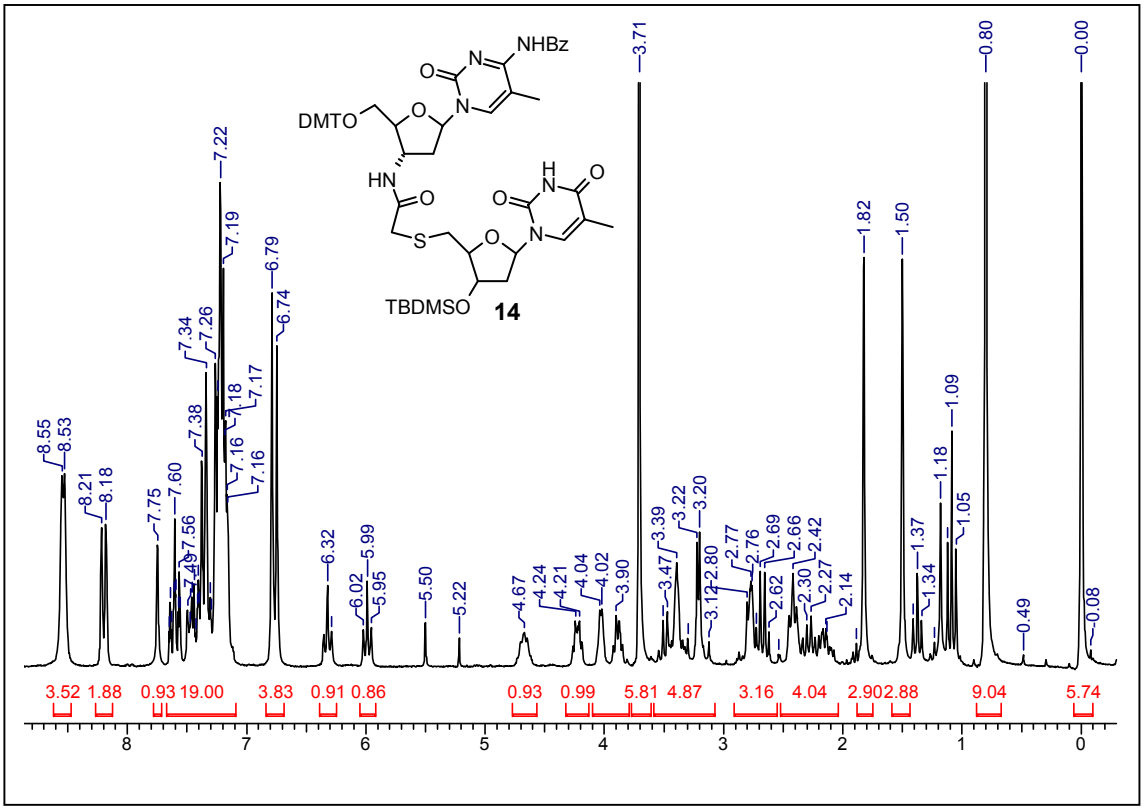
Mass of compound **4**, **5**, **13**, **15**, **17**, **14**, **16**, **18**, **19**, **20**, **21**, **22**, **23**, **24**, **25**, **26**, **27**, **29**, **30**, **31**, **32**, **33**, **34**, **35**, **36**, **37** & **38**

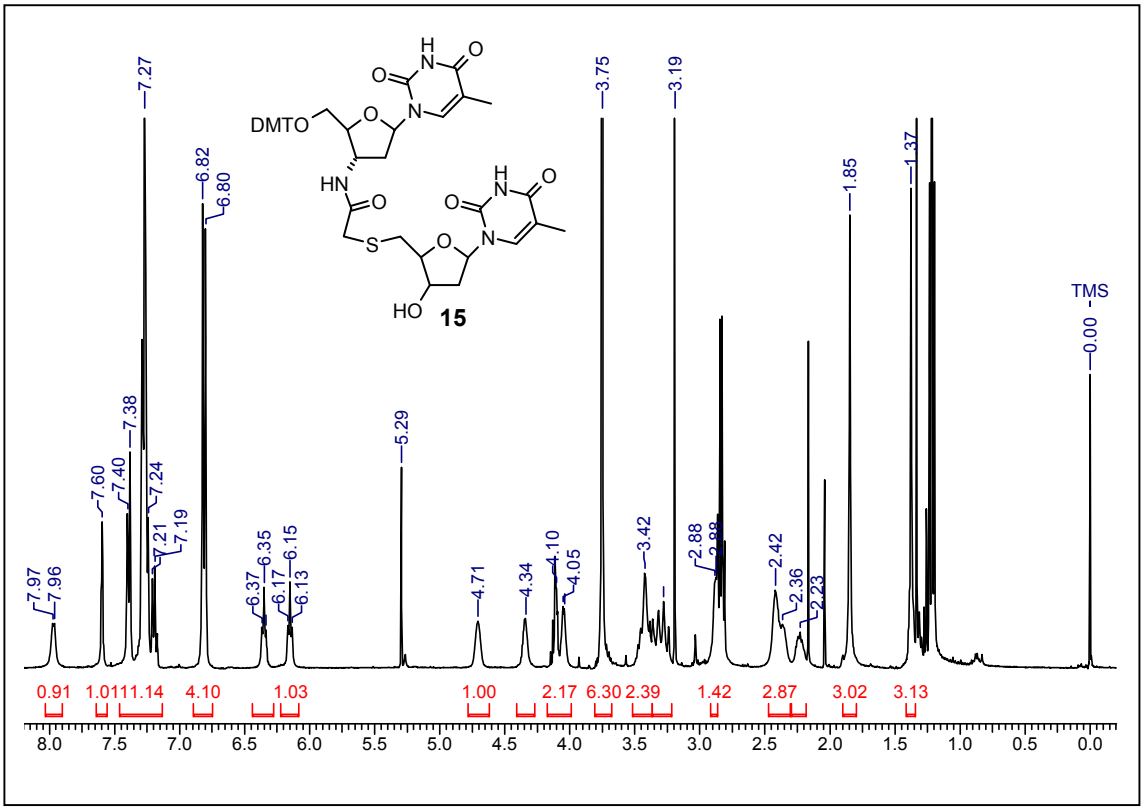

TosO
TBDMSO
2
^{1}H NMR
7.93 7.89 7.81 7.77 7.63 7.51 7.50 7.49 7.47 7.24
6.32 6.29 6.26
4.36 4.35 4.33 4.29 4.25 4.24 4.18 4.17 4.00 3.98
2.46 2.16 2.12 1.98 1.97 1.58
0.82 0.83 0.80 0.02 0.00
10.09 1.00 4.11 2.98 3.00 9.18 5.58

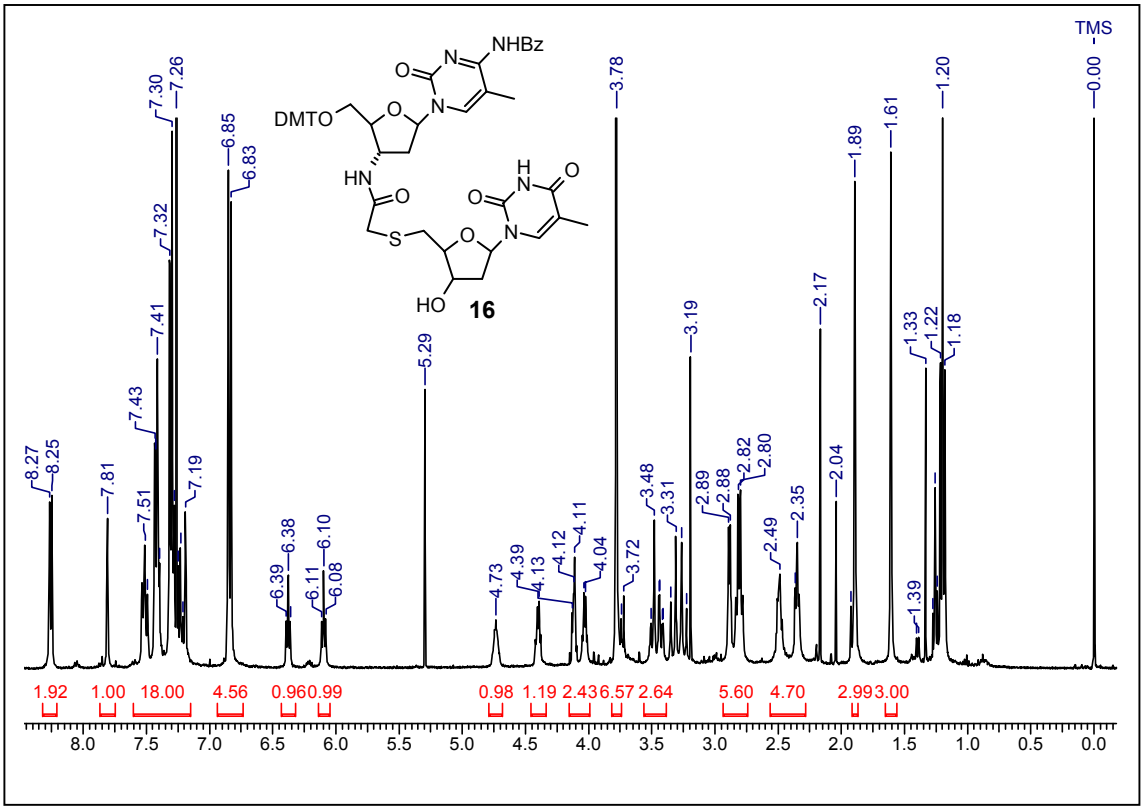

COOEt
S
TBDMSO
4
^{1}H NMR
TMS
Chloroform-d
9.46 9.44
7.37 7.29
6.26
4.32 4.31 4.30 4.21 4.20 4.19 4.17 4.04 4.03
3.35 3.32 3.29 3.26
2.99 2.99 2.97 2.95 2.94 2.30 2.29 2.28 2.15
1.94 1.30 1.29 1.27
0.89 0.10 0.00
0.91 0.98 0.92 1.83 1.90 1.87 0.95 2.87 2.96 9.00 5.72

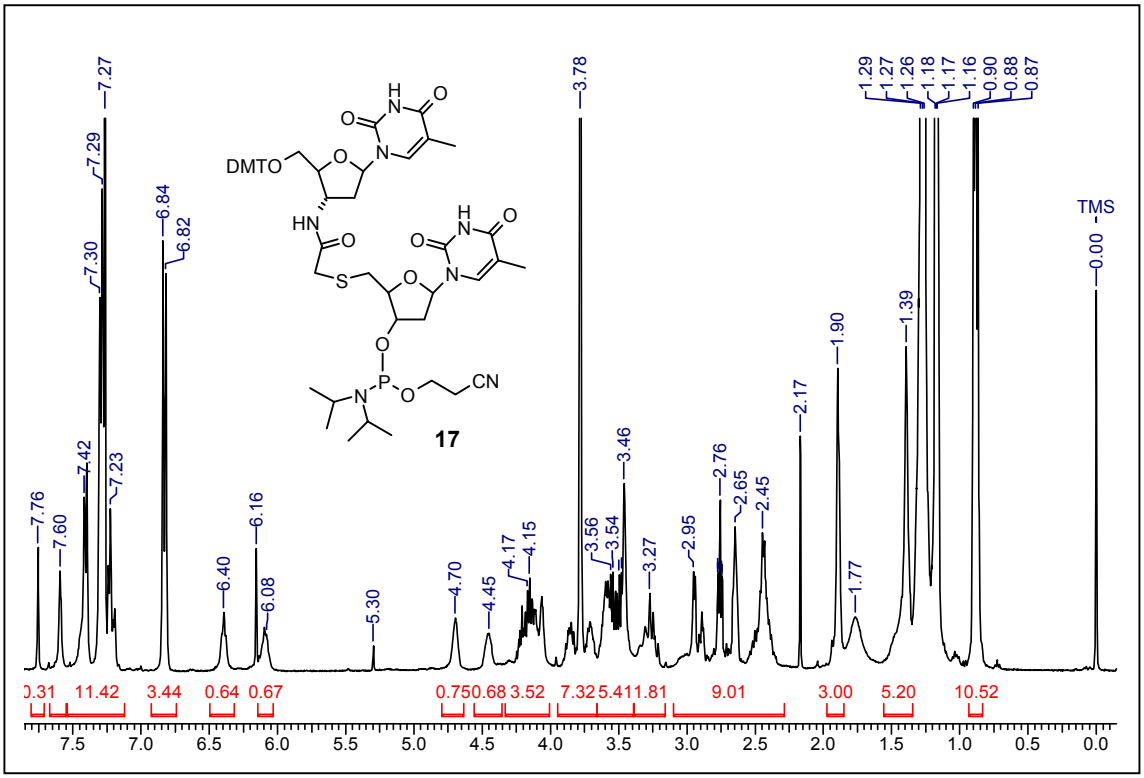

13C NMR
COOEt
S
O
N
H
O
TBDMSO
4
170.09
163.71
150.19
135.47
111.28
85.72
84.75
77.68
77.05
76.41
73.52
61.53
40.56
34.37
34.22
25.70
17.92
14.17
12.58
4.61
150
100
50
0

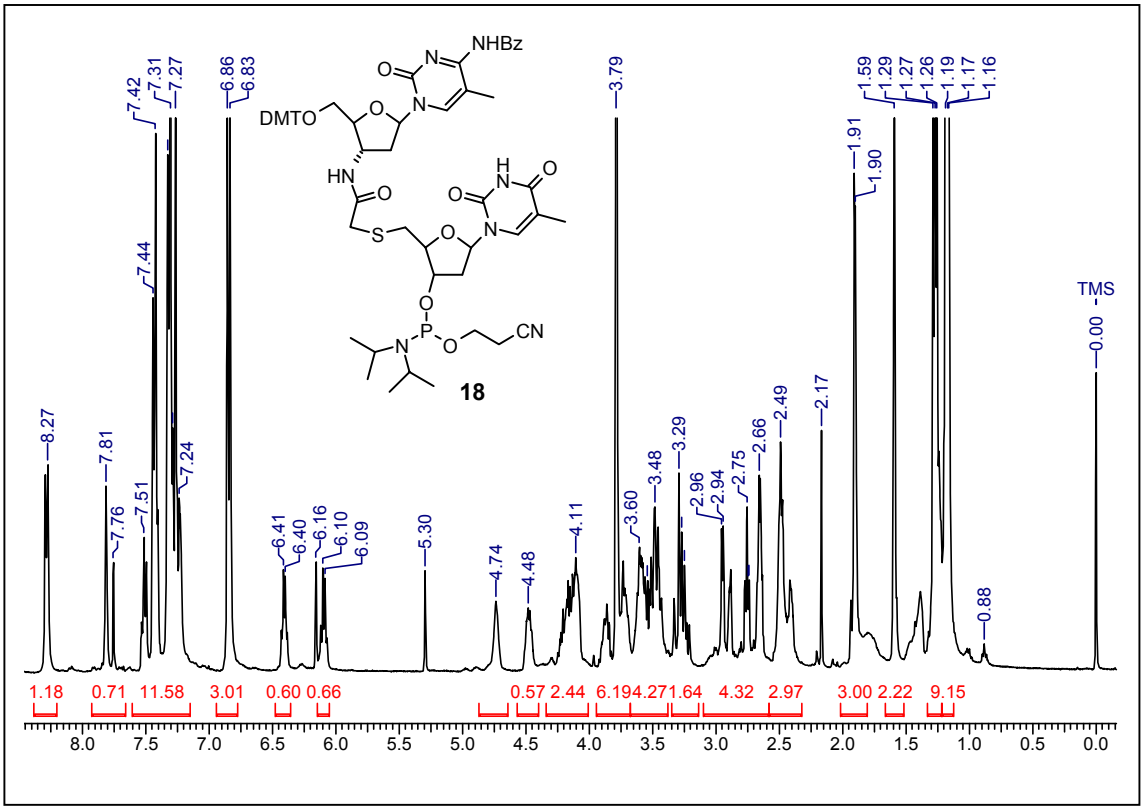

TMS
DMTO
HN
S
TBDMSO
13
8.63
7.70
7.68
7.66
7.40
7.39
7.29
7.27
7.26
7.18
6.82
6.80
6.48
6.19
4.69
4.27
4.11
3.94
3.93
3.78
3.55
3.38
3.30
3.09
2.90
2.87
2.85
2.84
2.80
2.44
2.22
2.21
2.17
1.93
1.83
1.77
1.51
1.38
1.37
1.26
1.02
0.89
0.87
0.71
0.22
0.09
0.08
0.07
0.01
0.56
11.32
3.71
0.65
0.71
0.80
3.07
6.17
1.40
1.93
2.43
4.42
9.00
5.80
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

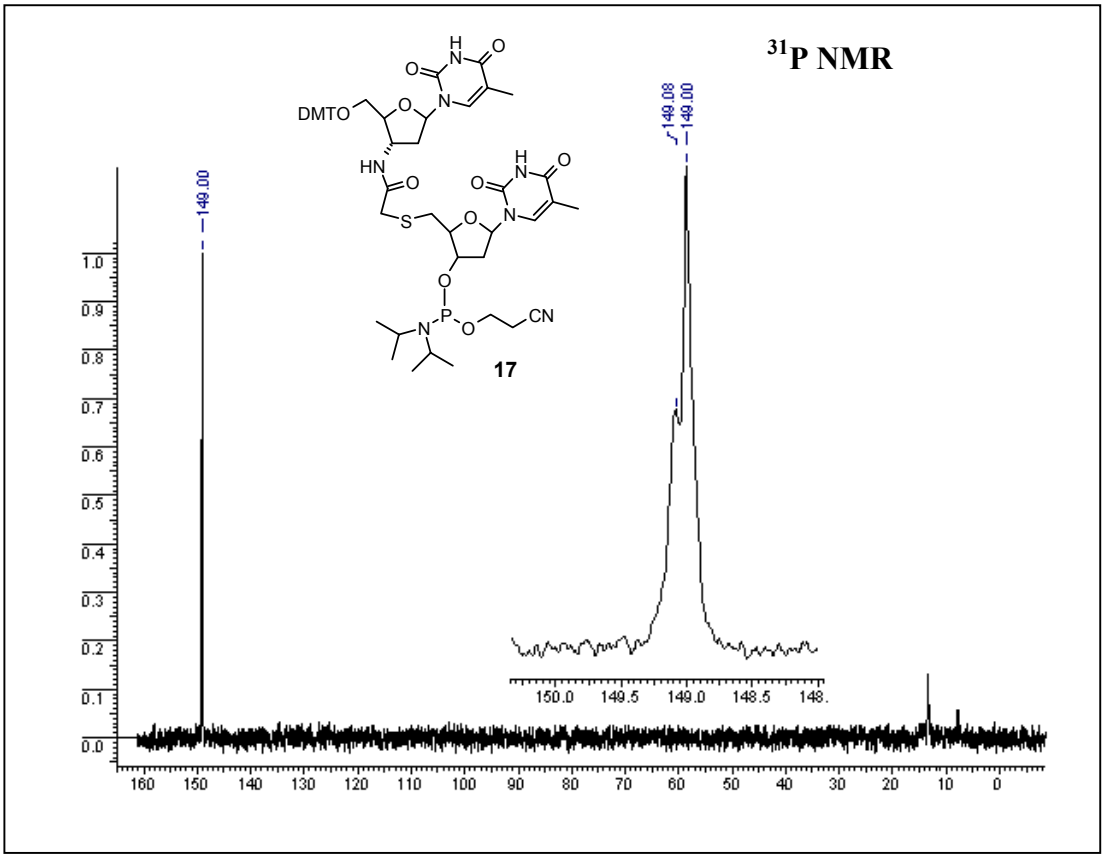

NHBz
DMTO
HN
TBDMSO
14
8.55
8.53
8.21
8.18
7.75
7.60
7.56
7.49
7.38
7.34
7.26
7.22
7.19
7.18
7.17
7.16
7.16
6.79
6.74
6.32
6.02
5.99
5.95
5.50
5.22
4.67
4.24
4.21
4.04
4.02
3.90
3.71
3.47
3.39
3.22
3.20
3.14
2.80
2.77
2.76
2.69
2.66
2.62
2.42
2.30
2.27
2.14
1.82
1.50
1.37
1.34
1.18
1.09
1.05
0.80
0.49
0.08
0.00
3.52 1.88 0.93 19.00 3.83 0.91 0.86 0.93 0.99 5.81 4.87 3.16 4.04 2.90 2.88 9.04 5.74

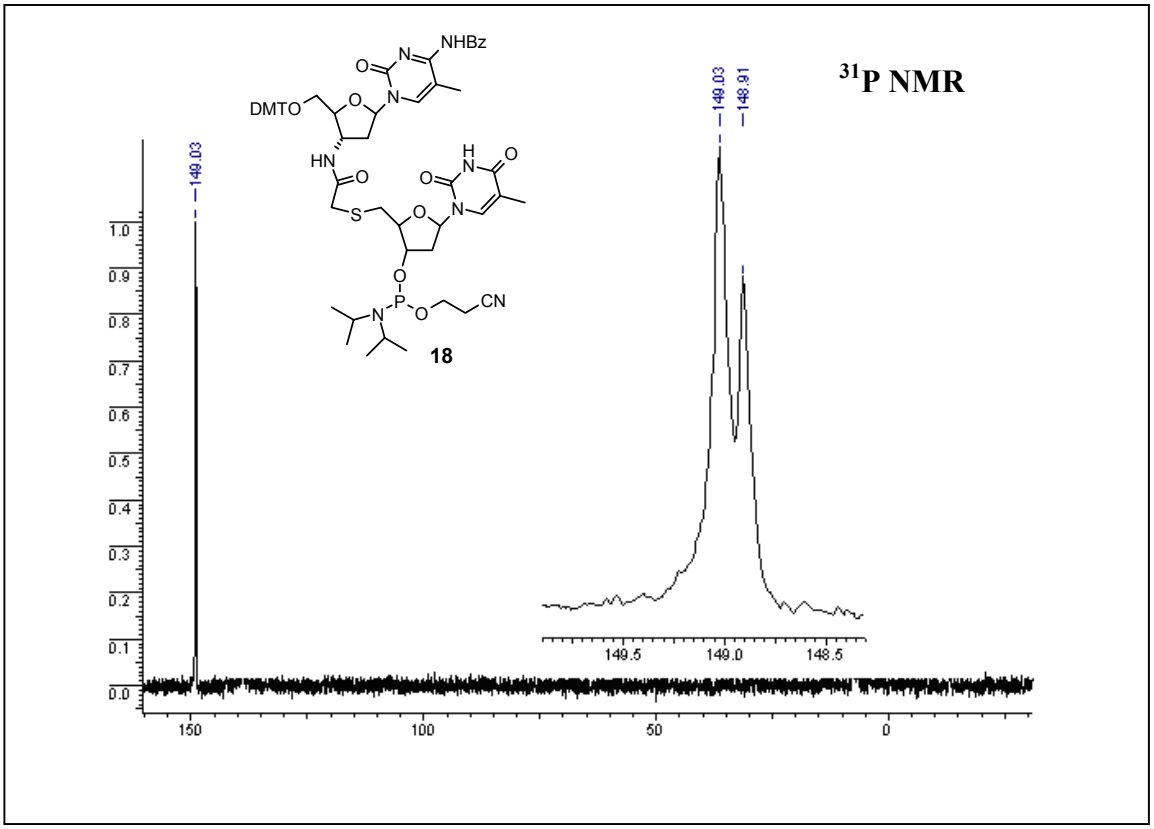

DMTO
HN
HO
15
7.97
7.96
7.60
7.40
7.38
7.27
7.24
7.21
7.19
6.82
6.80
6.37
6.35
6.17
6.15
6.13
5.29
4.71
4.34
4.10
4.05
3.75
3.42
3.19
2.88
2.88
2.42
2.36
2.23
1.85
1.37
TMS
0.00
0.91 1.01 11.14 4.10 1.03 1.00 2.17 6.30 2.39 1.42 2.87 3.02 3.13

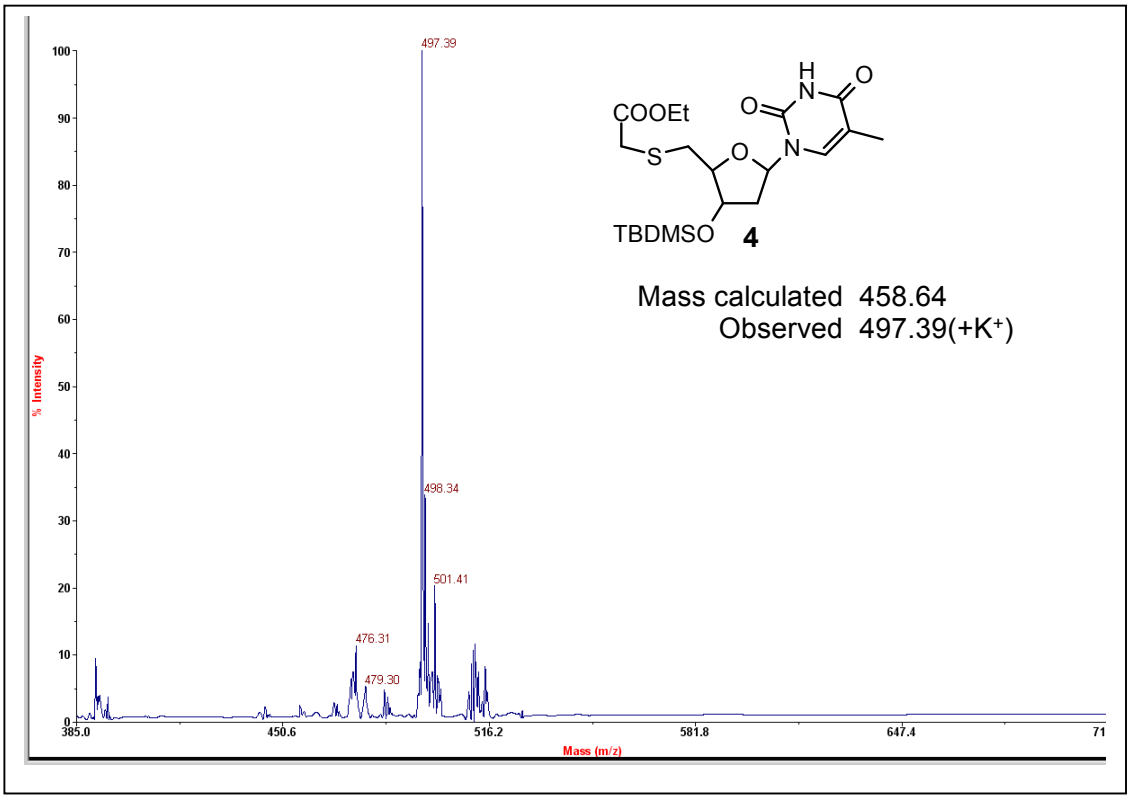

NHBz
DMTO
HN
HO
16
TMS
8.27
8.25
7.81
7.51
7.43
7.41
7.32
7.30
7.26
7.19
6.85
6.83
6.39
6.38
6.11
6.10
6.08
5.29
4.73
4.39
4.13
4.12
4.11
4.04
3.78
3.72
3.48
3.31
3.19
2.89
2.88
2.82
2.80
2.49
2.35
2.17
2.04
1.89
1.61
1.39
1.33
1.22
1.20
1.18
0.00
1.92 1.00 18.00 4.56 0.96 0.99 0.98 1.19 2.43 6.57 2.64 5.60 4.70 2.99 3.00

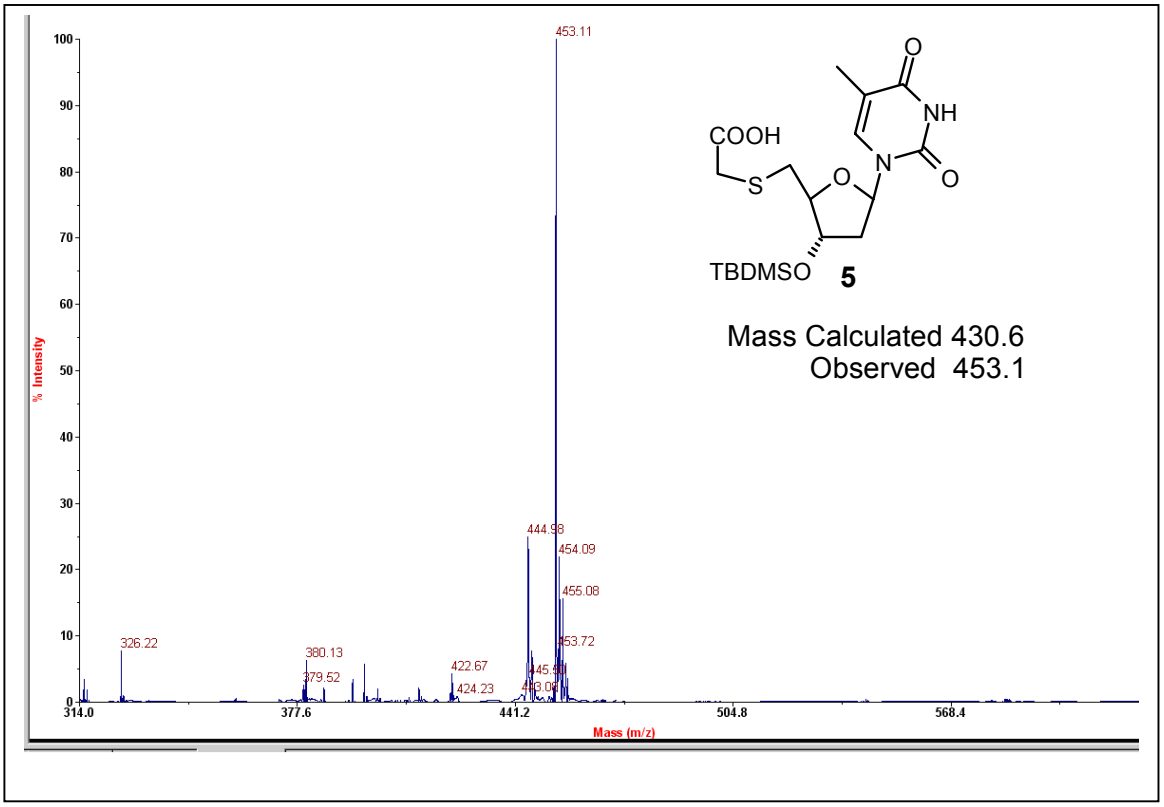

DMTO
HN
CN
17
TMS
7.76
7.60
7.42
7.30
7.29
7.27
7.23
6.84
6.82
6.40
6.16
6.08
5.30
4.70
4.45
4.17
4.15
3.78
3.56
3.54
3.46
3.27
2.95
2.76
2.65
2.45
2.17
1.90
1.77
1.39
1.29
1.27
1.26
1.18
1.17
1.16
0.90
0.88
0.87
0.00
0.31 11.42 3.44 0.64 0.67 0.75 0.68 3.52 7.32 5.41 1.81 9.01 3.00 5.20 10.52

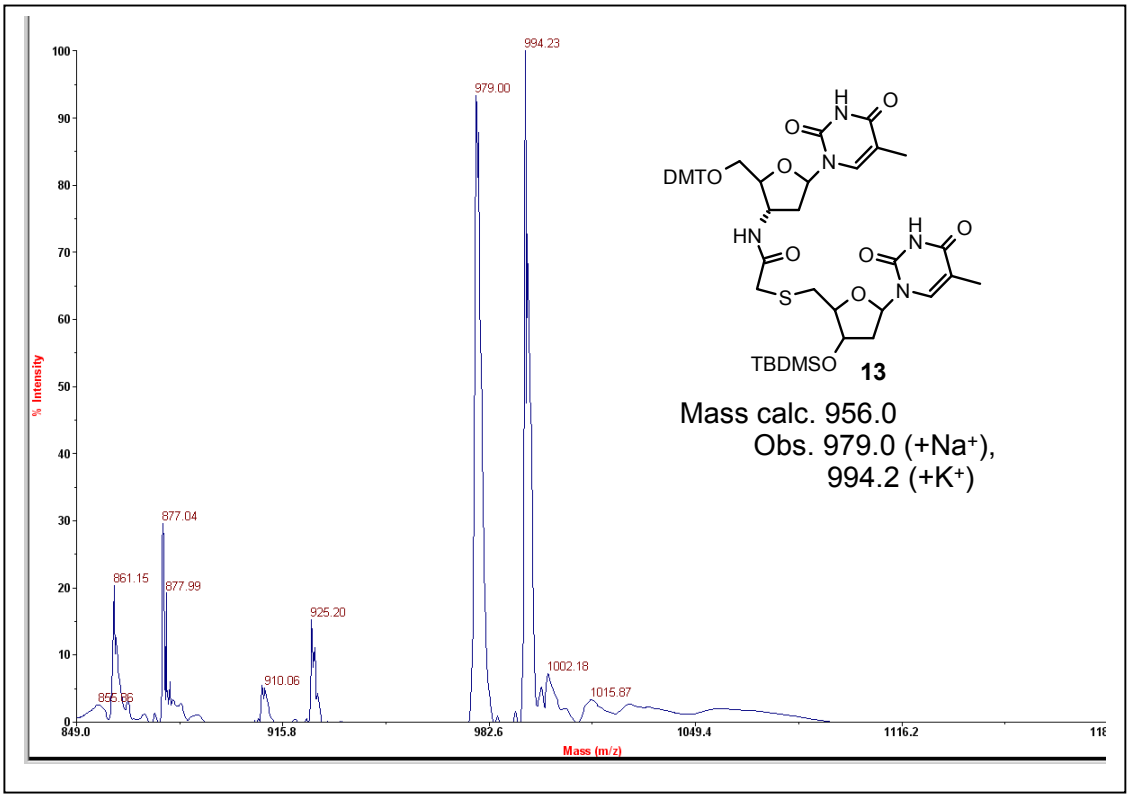

NHBz
DMTO
HN
CN
18
8.27
7.81
7.76
7.51
7.44
7.42
7.31
7.27
7.24
6.86
6.83
6.41
6.40
6.16
6.10
6.09
5.30
4.74
4.48
4.11
3.79
3.60
3.48
3.29
2.96
2.94
2.75
2.66
2.49
2.17
1.91
1.90
1.59
1.29
1.27
1.26
1.19
1.17
1.16
0.88
TMS
0.00
1.18
0.71
11.58
3.01
0.60
0.66
0.57
2.44
6.19
4.27
1.64
4.32
2.97
3.00
2.22
9.15

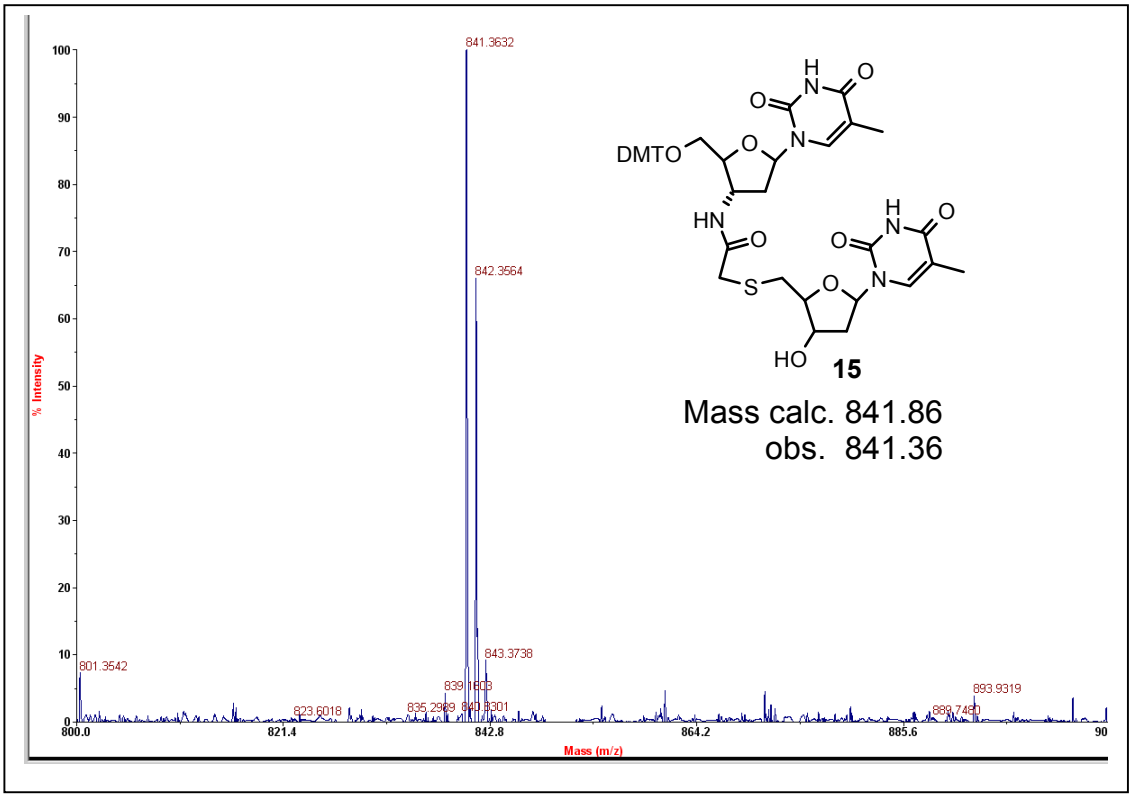

31P NMR
DMTO
HN
CN
17
149.00
149.08
149.00

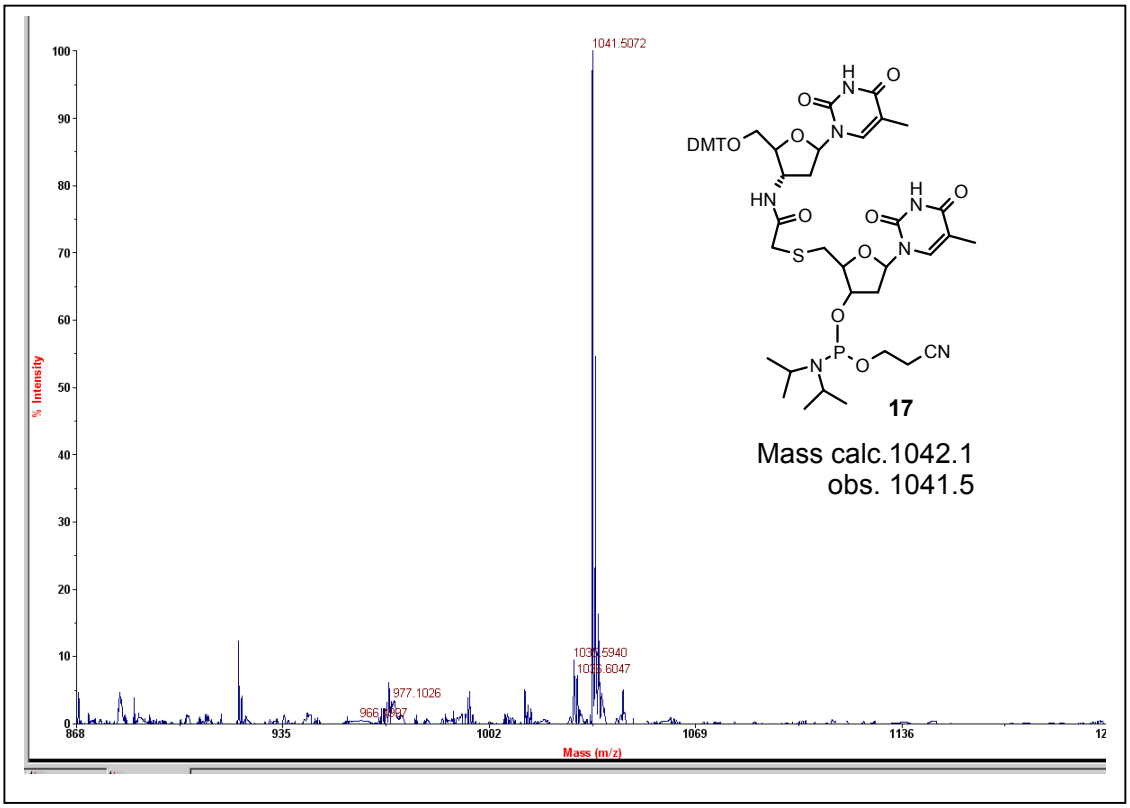

NHBz
DMTO
HN
S
O
P
CN
18
^{31}P NMR
149.03
148.91
149.5
149.0
148.5
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
150
100
50
0

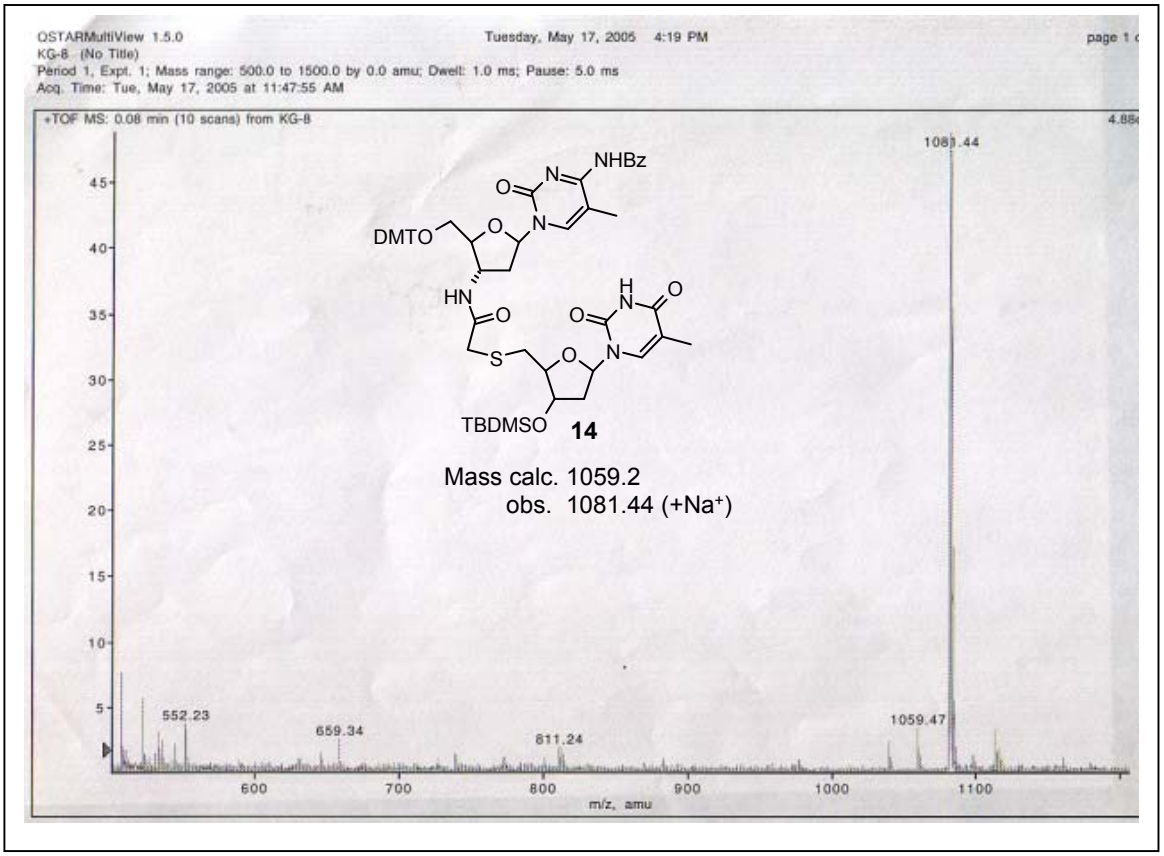

497.39
498.34
501.41
476.31
479.30
% Intensity
Mass (m/z)
385.0
450.6
516.2
581.8
647.4
COOEt
S
TBDMSO
4
Mass calculated 458.64
Observed 497.39(+K⁺)

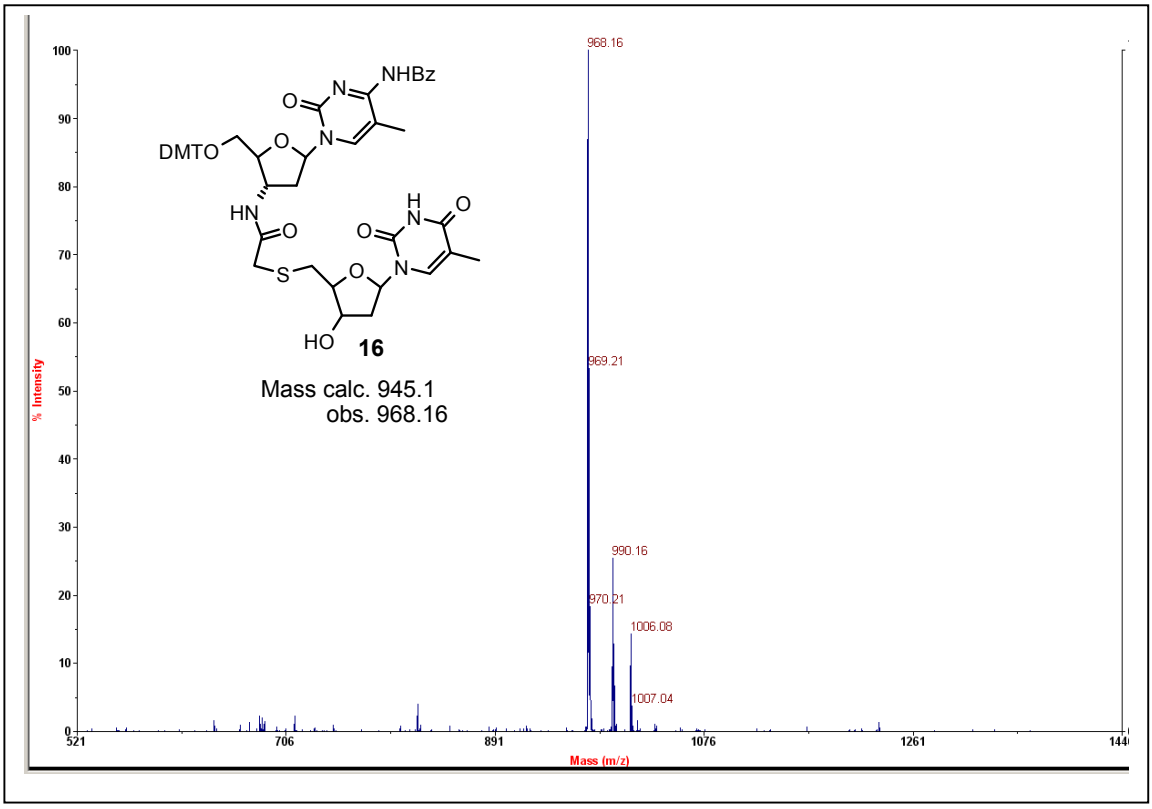

453.11
444.98
454.09
455.08
453.72
326.22
380.13
379.52
422.67
424.23
% Intensity
Mass (m/z)
314.0
377.6
441.2
504.8
568.4
COOH
S
O
N
NH
O
O
TBDMSO
5
Mass Calculated 430.6
Observed 453.1

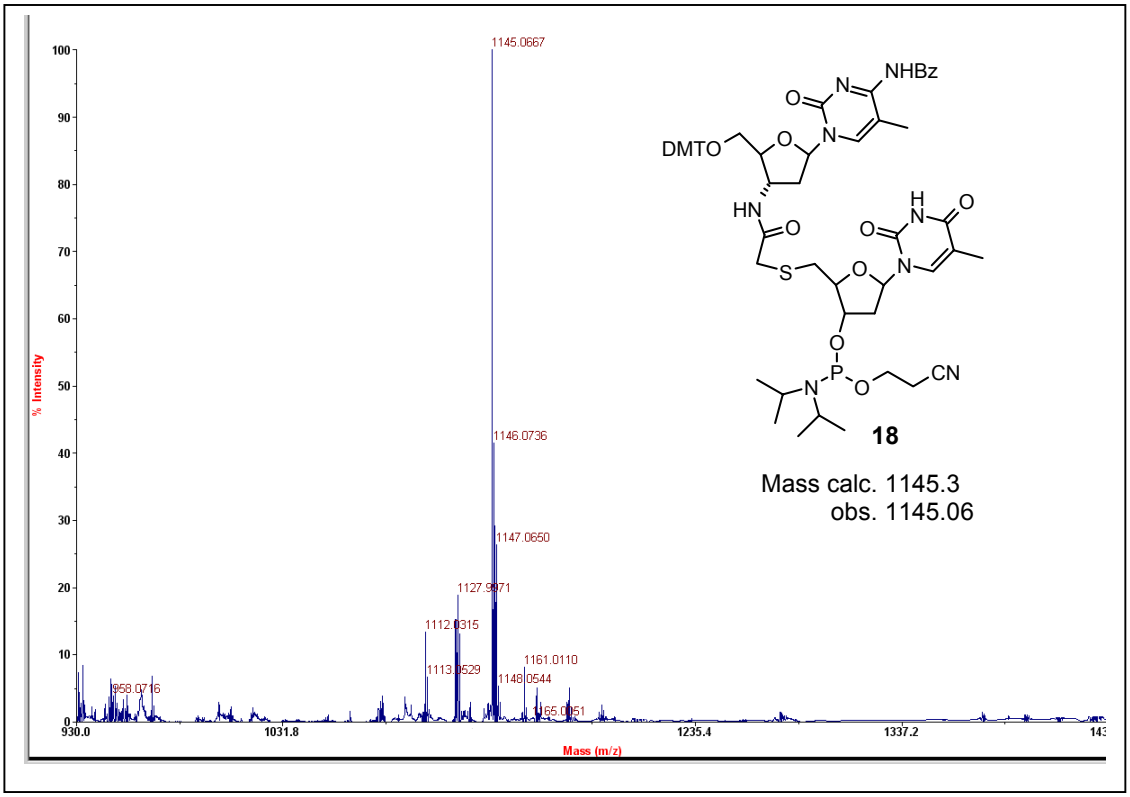

994.23
979.00
877.04
861.15
877.99
925.20
910.06
1002.18
1015.87
855.86
% Intensity
Mass (m/z)
849.0
915.8
982.6
1049.4
1116.2
DMTO
HN
S
O
N
H
TBDMSO
13
Mass calc. 956.0
Obs. 979.0 (+Na+),
994.2 (+K+)

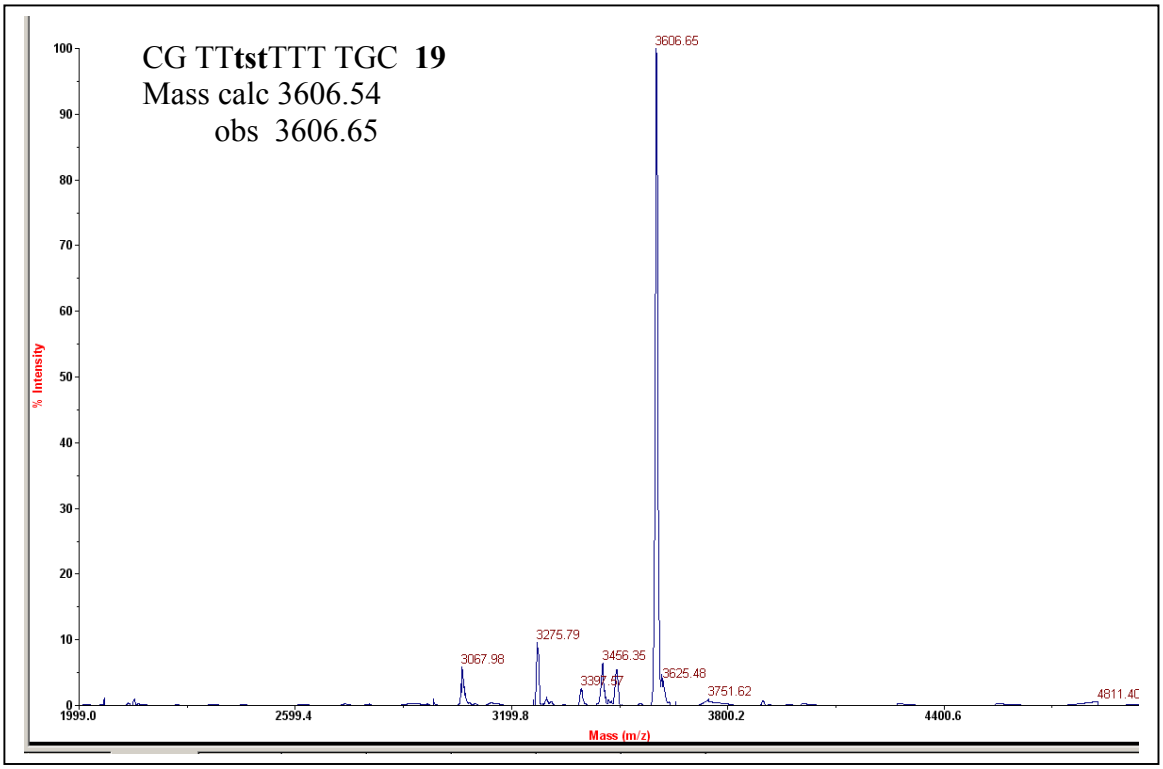

841.3632
842.3564
843.3738
801.3542
823.6018
893.9319
% Intensity
Mass (m/z)
DMTO
HN
S
HO
15
Mass calc. 841.86
obs. 841.36

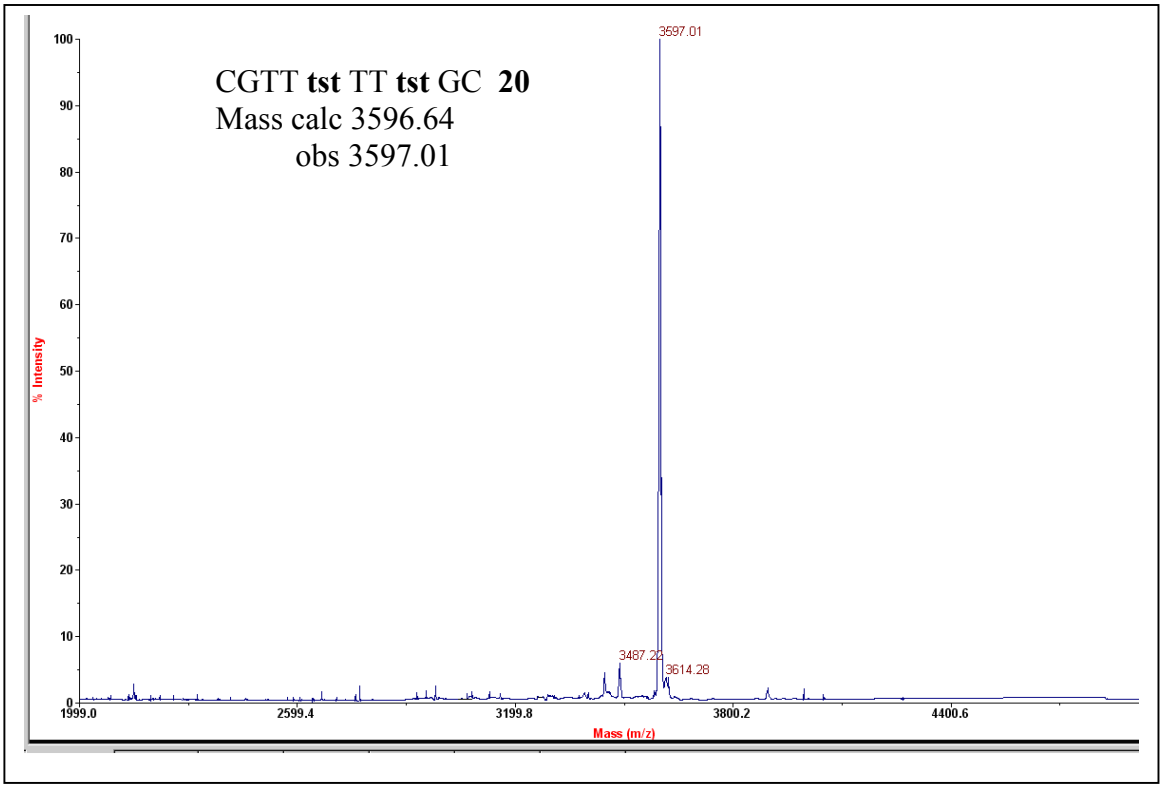

1041.5072
977.1026
% Intensity
Mass (m/z)
DMTO
HN
S
CN
17
Mass calc.1042.1
obs. 1041.5

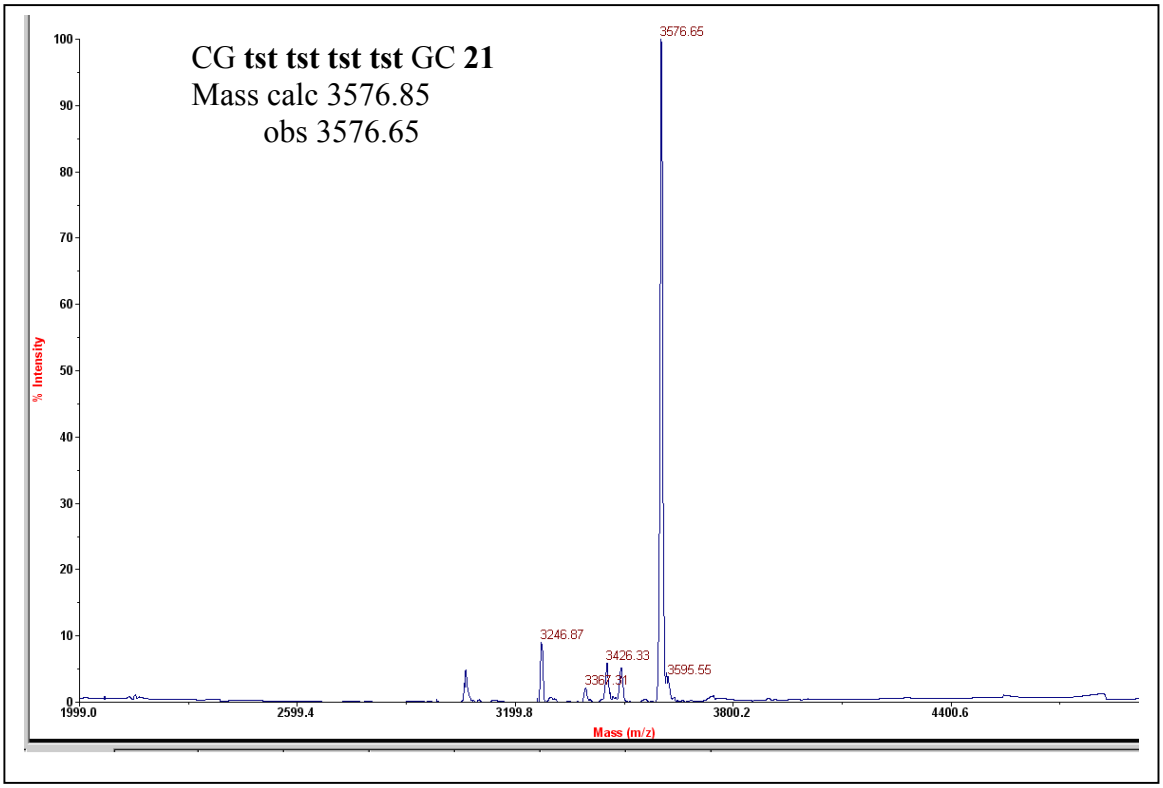

QSTARMultiView 1.5.0
Tuesday, May 17, 2005 4:19 PM
KG-8 (No Title)
Period 1, Expt. 1; Mass range: 500.0 to 1500.0 by 0.0 amu; Dwell: 1.0 ms; Pause: 5.0 ms
Acq. Time: Tue, May 17, 2005 at 11:47:55 AM
+TOF MS: 0.08 min (10 scans) from KG-8
1081.44
552.23
659.34
811.24
1059.47
m/z, amu
NHBz
DMTO
HN
TBDMSO
14
Mass calc. 1059.2
obs. 1081.44 (+Na+)

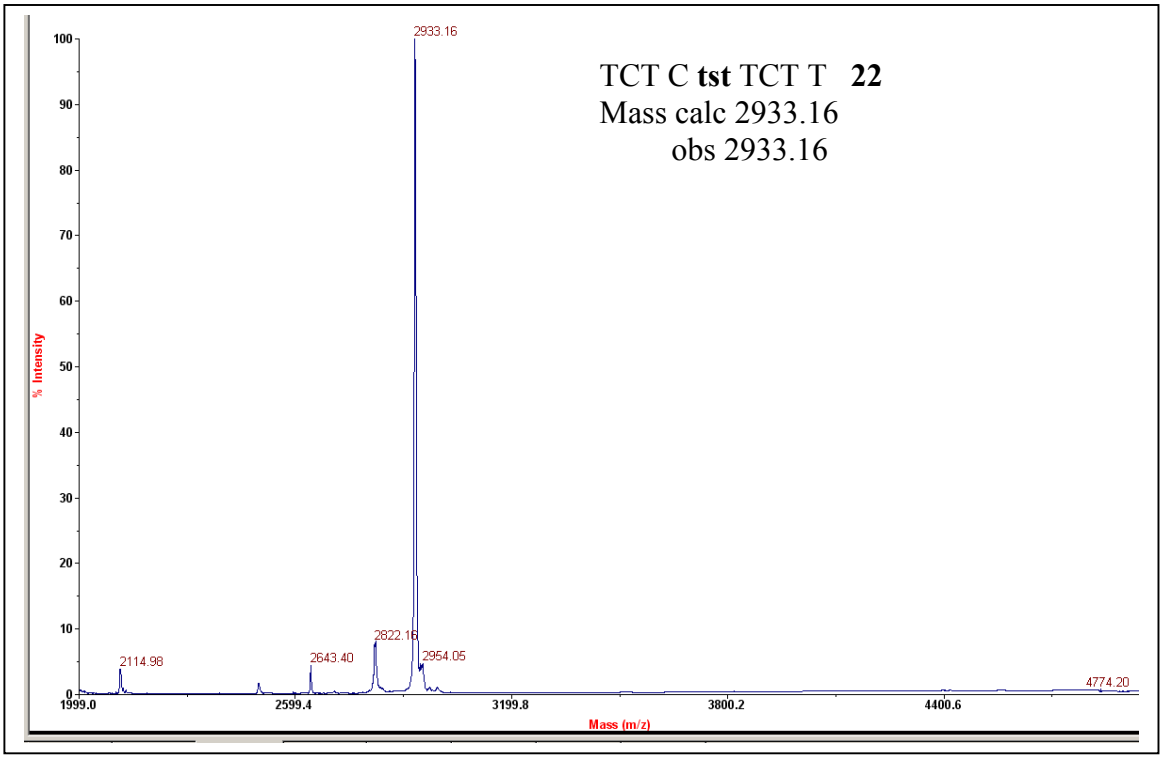

% Intensity
968.16
969.21
990.16
970.21
1006.08
1007.04
Mass (m/z)
NHBz
DMTO
HN
HO
16
Mass calc. 945.1
obs. 968.16


1145.0667
1146.0736
1147.0650
1127.9971
1112.0315
1161.0110
1148.0544
958.0716
% Intensity
Mass (m/z)
930.0
1031.8
1235.4
1337.2
NHBz
DMTO
HN
CN
18
Mass calc. 1145.3
obs. 1145.06

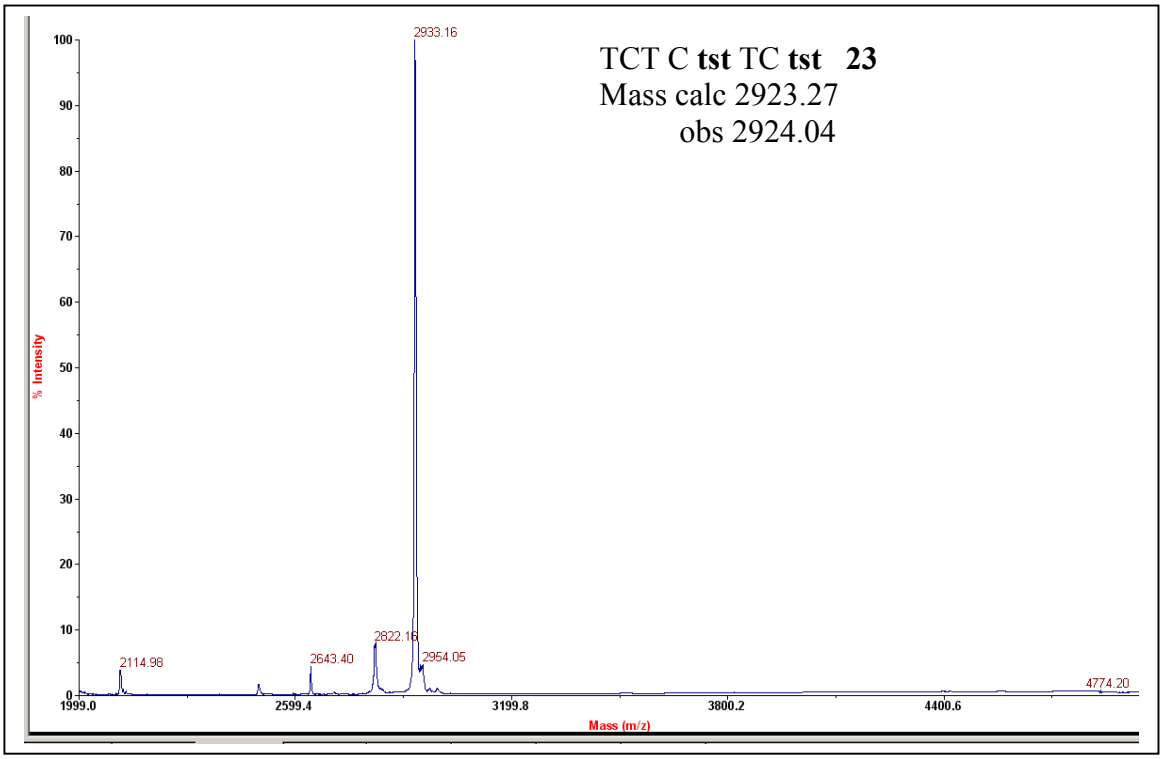

CG TT**tst**TTT TGC **19**
Mass calc 3606.54
obs 3606.65
3606.65
3275.79
3067.98
3456.35
3625.48
3751.62
% Intensity
Mass (m/z)
1999.0
2599.4
3199.8
3800.2
4400.6


CGTT **tst** TT **tst** GC **20**
Mass calc 3596.64
obs 3597.01
3597.01
3487.22
3614.28
% Intensity
Mass (m/z)
1999.0
2599.4
3199.8
3800.2
4400.6

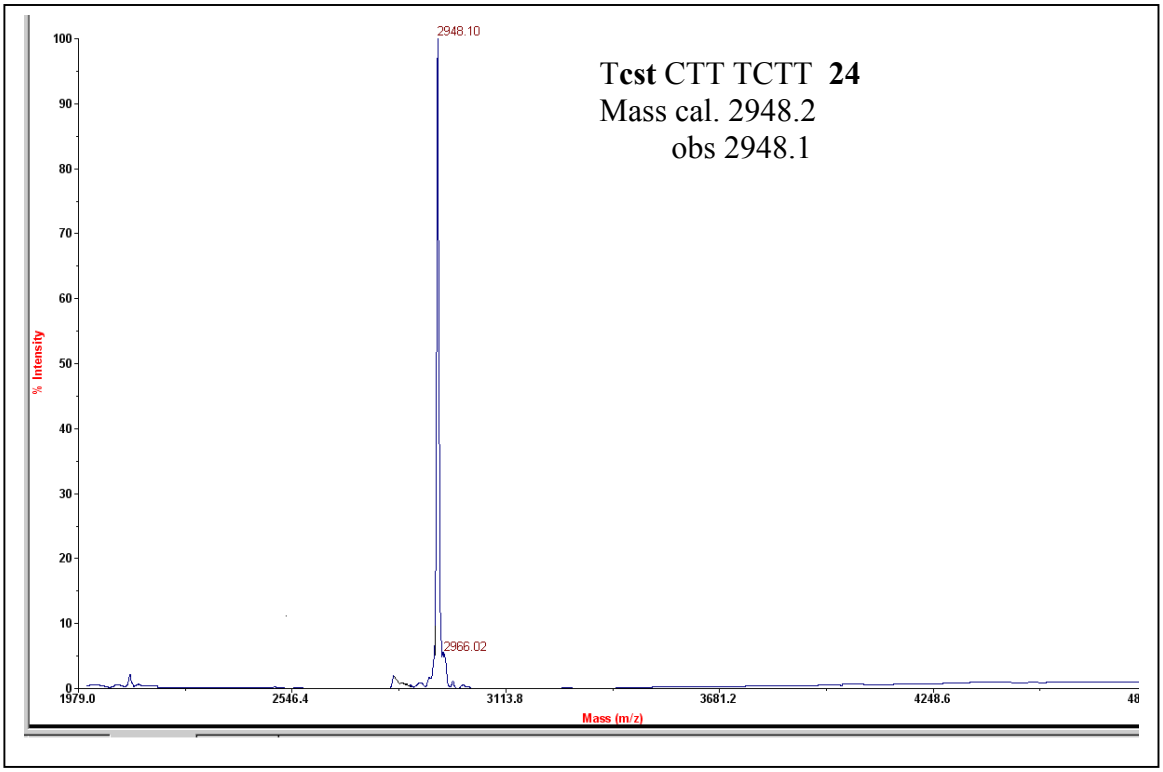

CG **tst tst tst tst** GC **21**
Mass calc 3576.85
obs 3576.65
3576.65
3246.87
3426.33
3595.55
% Intensity
Mass (m/z)
1999.0
2599.4
3199.8
3800.2
4400.6

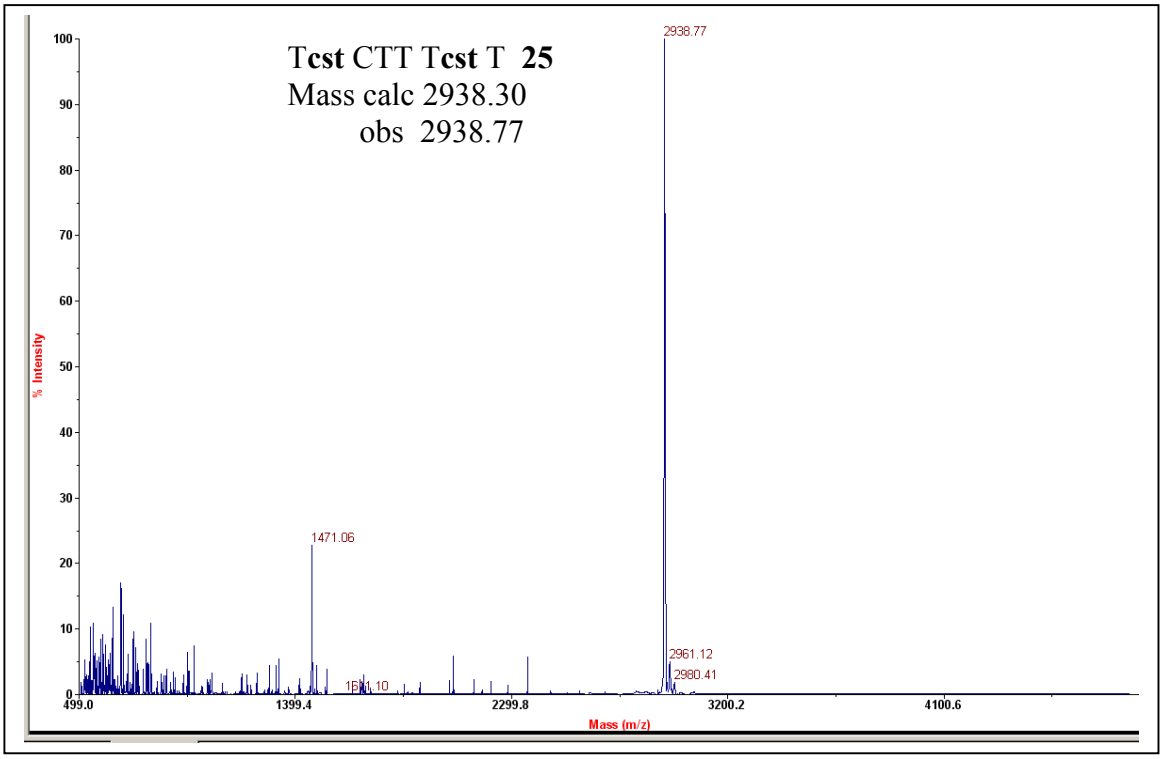


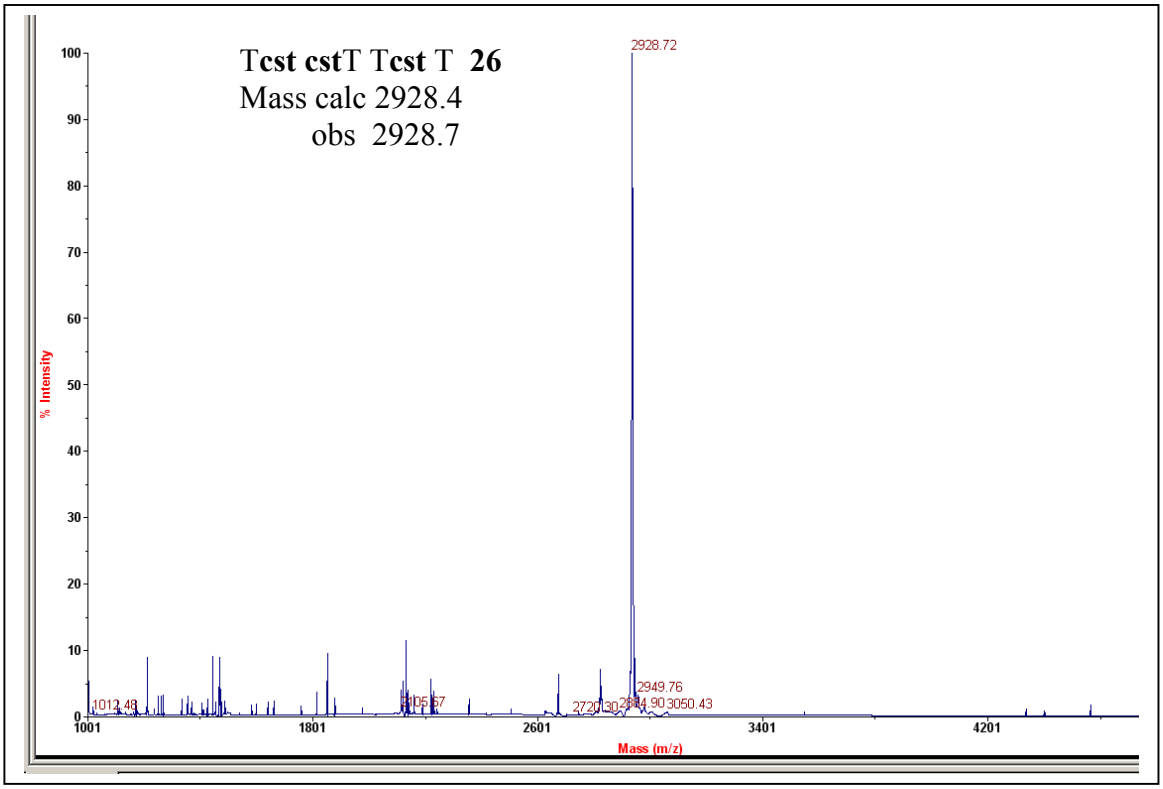
TCT C **tst** TCT T **22**
Mass calc 2933.16
obs 2933.16

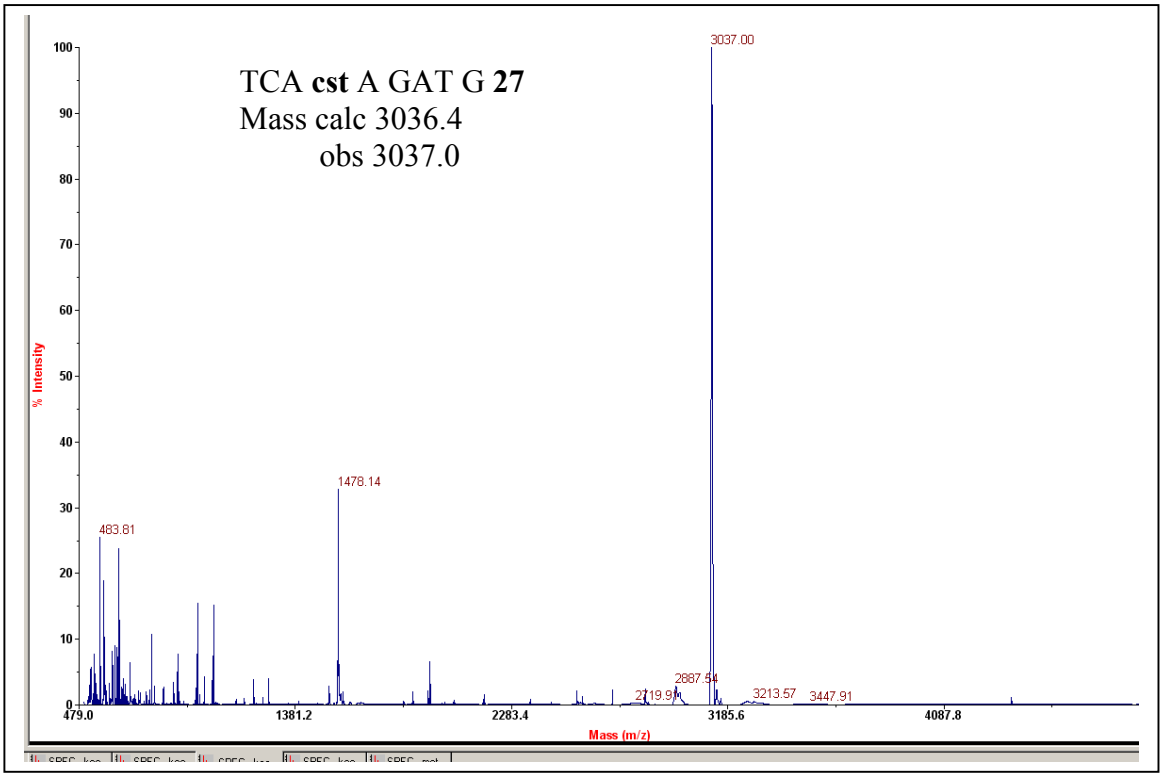


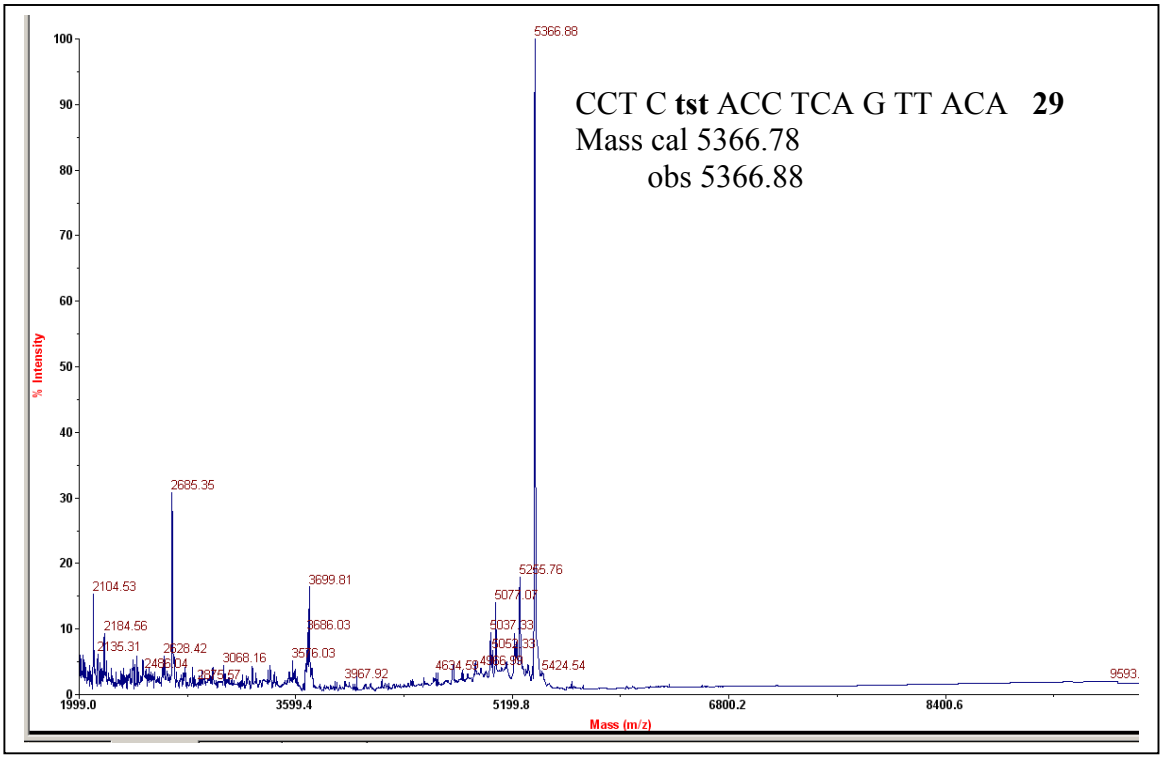
TCT C **tst** TC **tst** **23**
Mass calc 2923.27
obs 2924.04

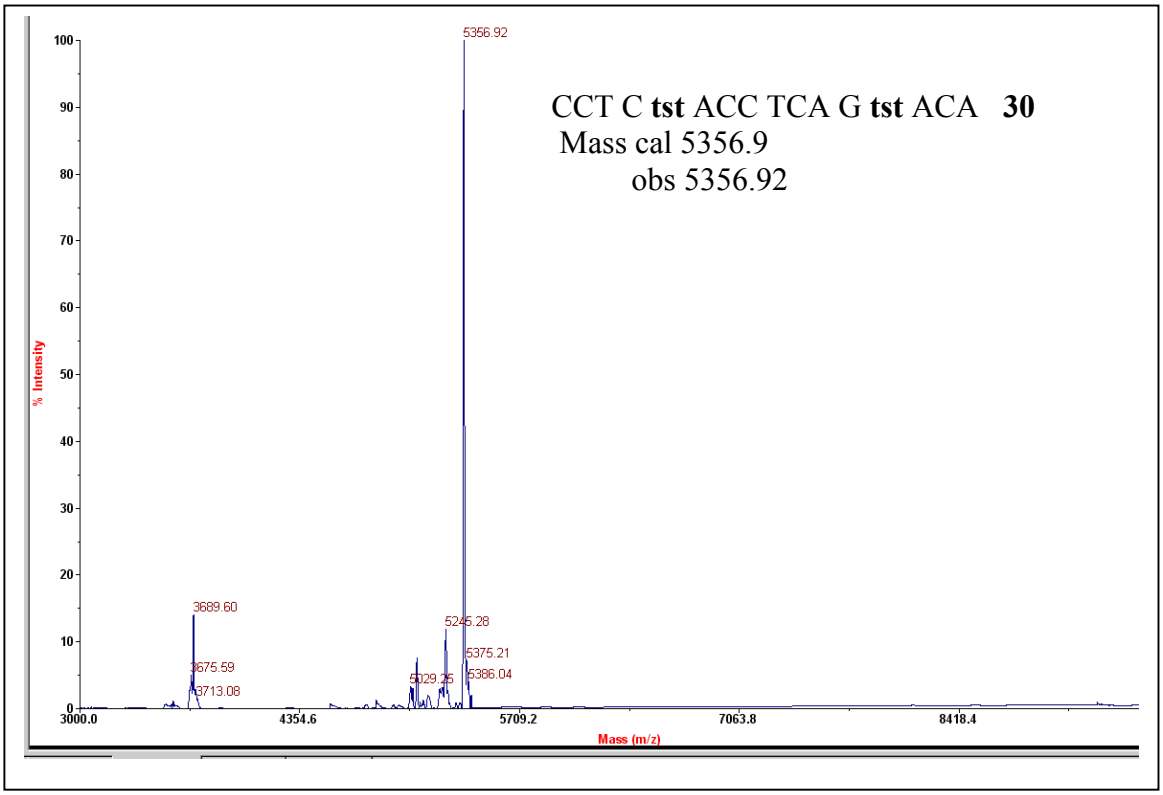

Tcst CTT TCTT 24
Mass cal. 2948.2
obs 2948.1
2948.10
2966.02
% Intensity
Mass (m/z)
1979.0
2546.4
3113.8
3681.2
4248.6

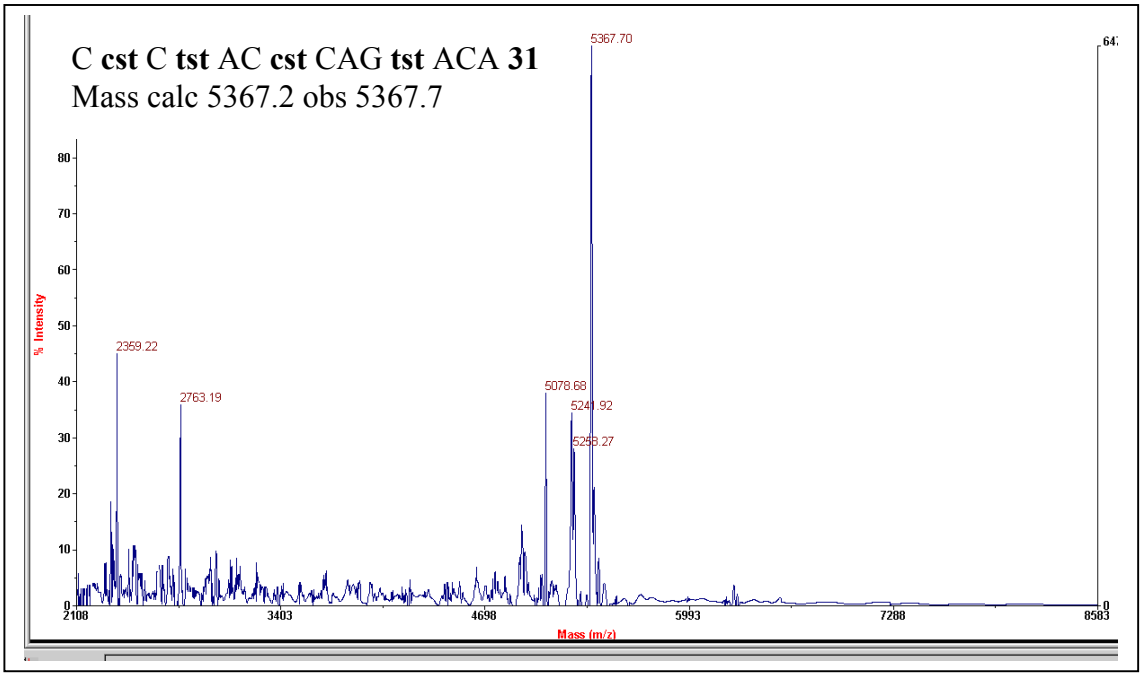

Tcst CTT Tcst T 25
Mass calc 2938.30
obs 2938.77
2938.77
1471.06
2961.12
2980.41
% Intensity
Mass (m/z)
499.0
1399.4
2299.8
3200.2
4100.6

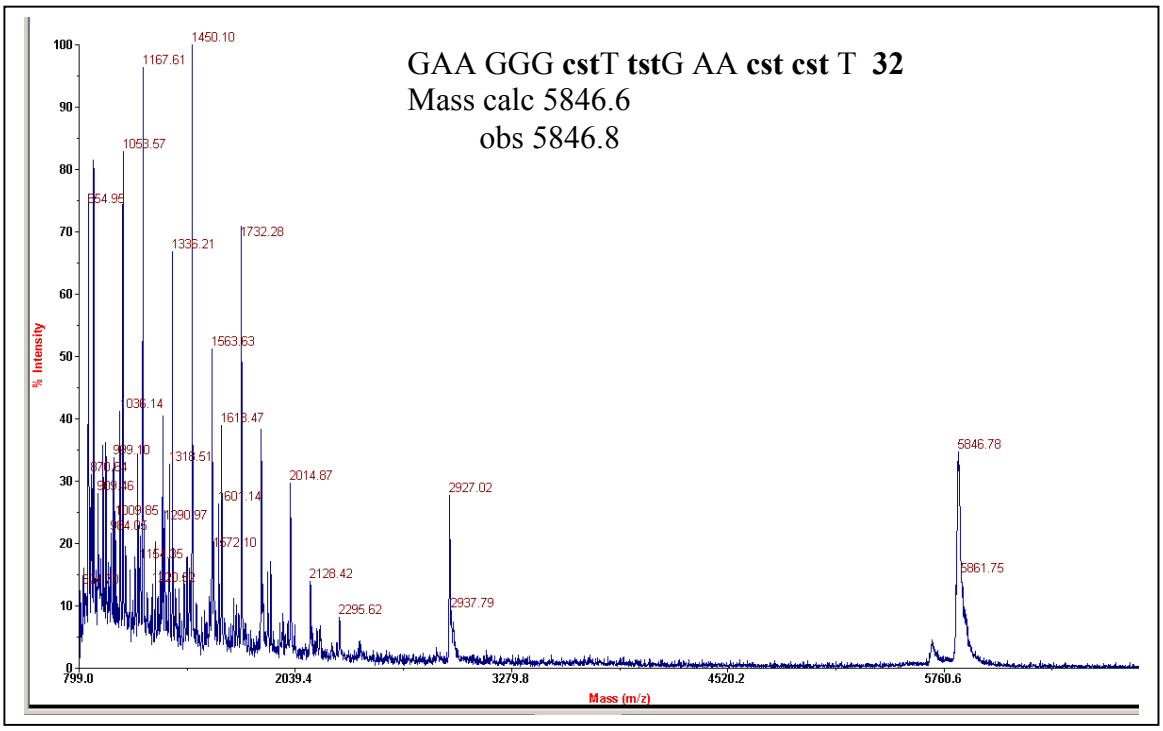

Tcst cstT Tcst T 26
Mass calc 2928.4
obs 2928.7
2928.72
2949.76
% Intensity
Mass (m/z)

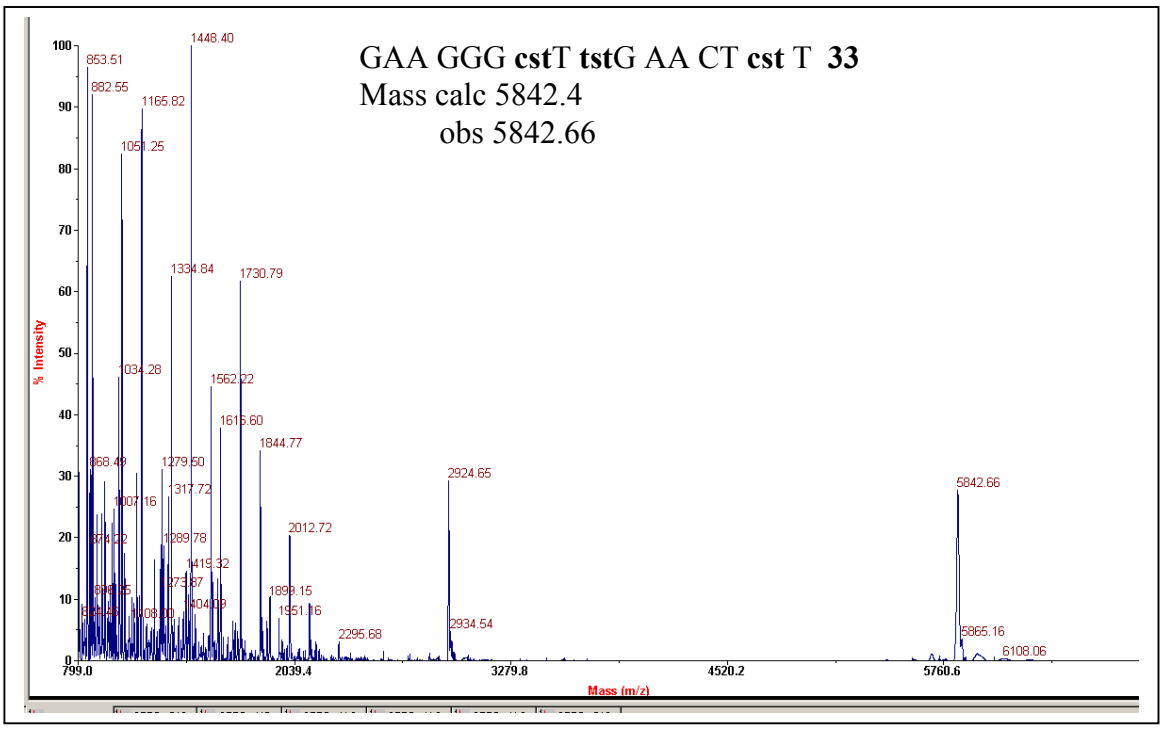

TCA cst A GAT G 27
Mass calc 3036.4
obs 3037.0
3037.00
1478.14
483.81
% Intensity
Mass (m/z)

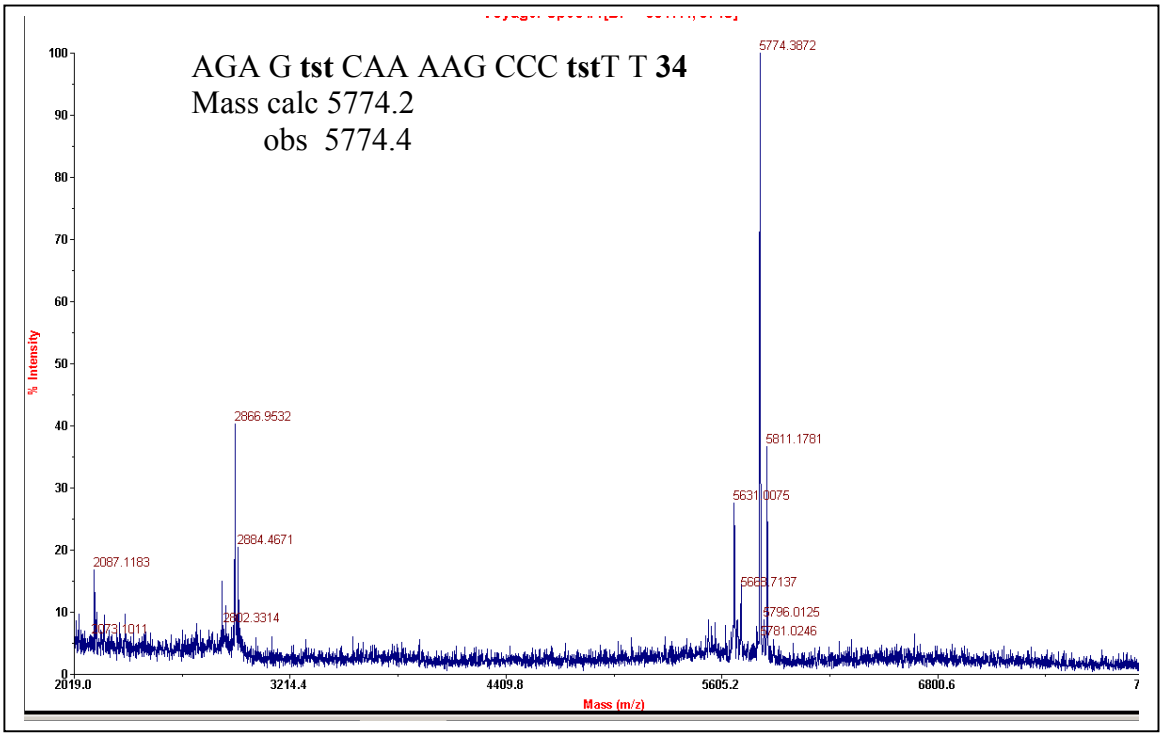

CCT C **tst** ACC TCA G TT ACA **29**
Mass cal 5366.78
obs 5366.88
% Intensity
Mass (m/z)
5366.88
2685.35
2104.53
3699.81
5255.76
5077.07
2184.56
3686.03
5037.33
2135.31
2628.42
3068.16
3516.03
3967.92
5424.54
1999.0
3599.4
5199.8
6800.2
8400.6

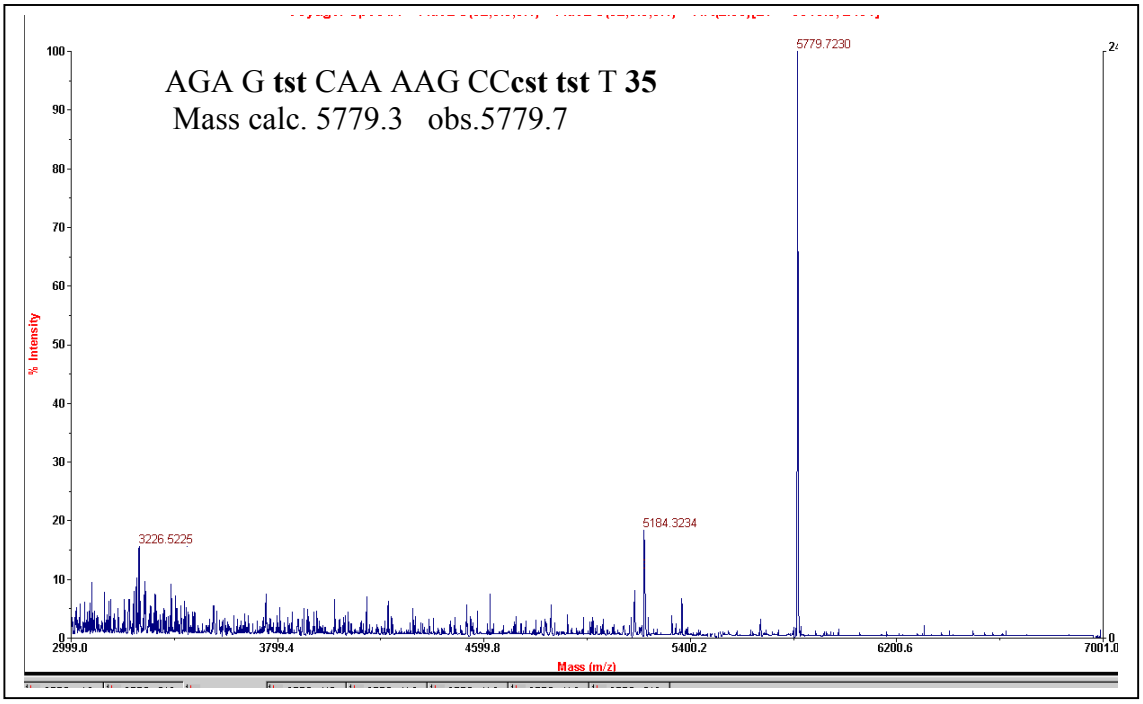

CCT C **tst** ACC TCA G **tst** ACA **30**
Mass cal 5356.9
obs 5356.92
% Intensity
Mass (m/z)
5356.92
3689.60
5245.28
5375.21
3675.59
5386.04
3713.08
5029.25
3000.0
4354.6
5709.2
7063.8
8418.4

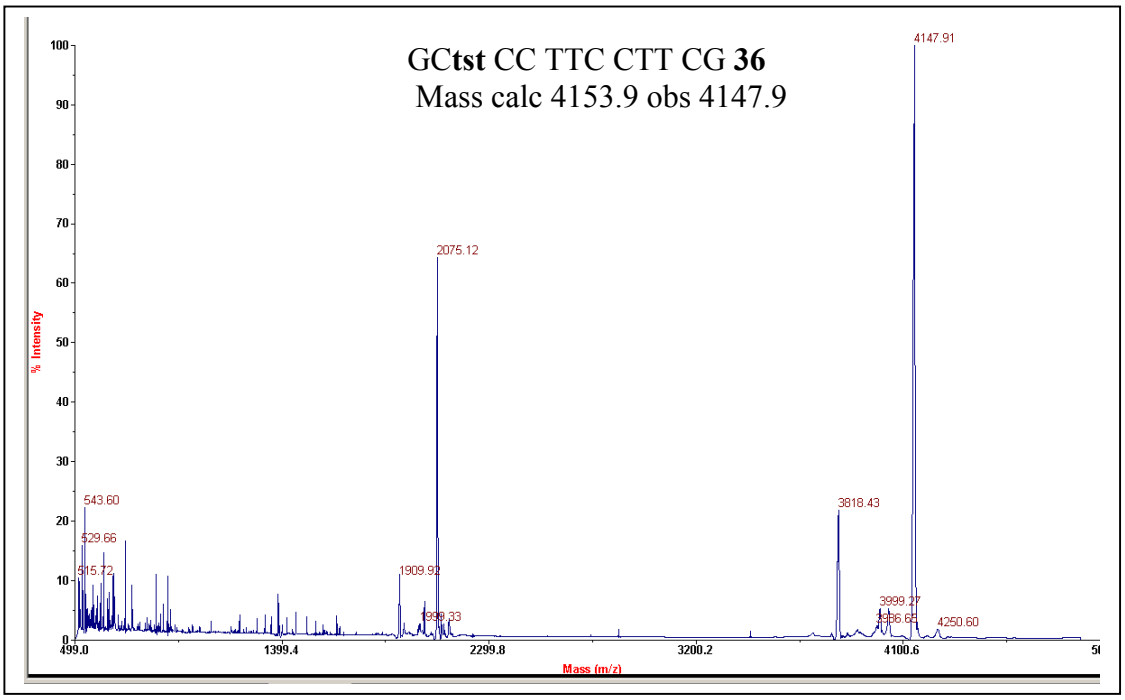

C cst C tst AC cst CAG tst ACA 31
Mass calc 5367.2 obs 5367.7
% Intensity
Mass (m/z)
5367.70
2359.22
2763.19
5078.68
5241.92
5258.27
2108
3403
4698
5993
7288
8583

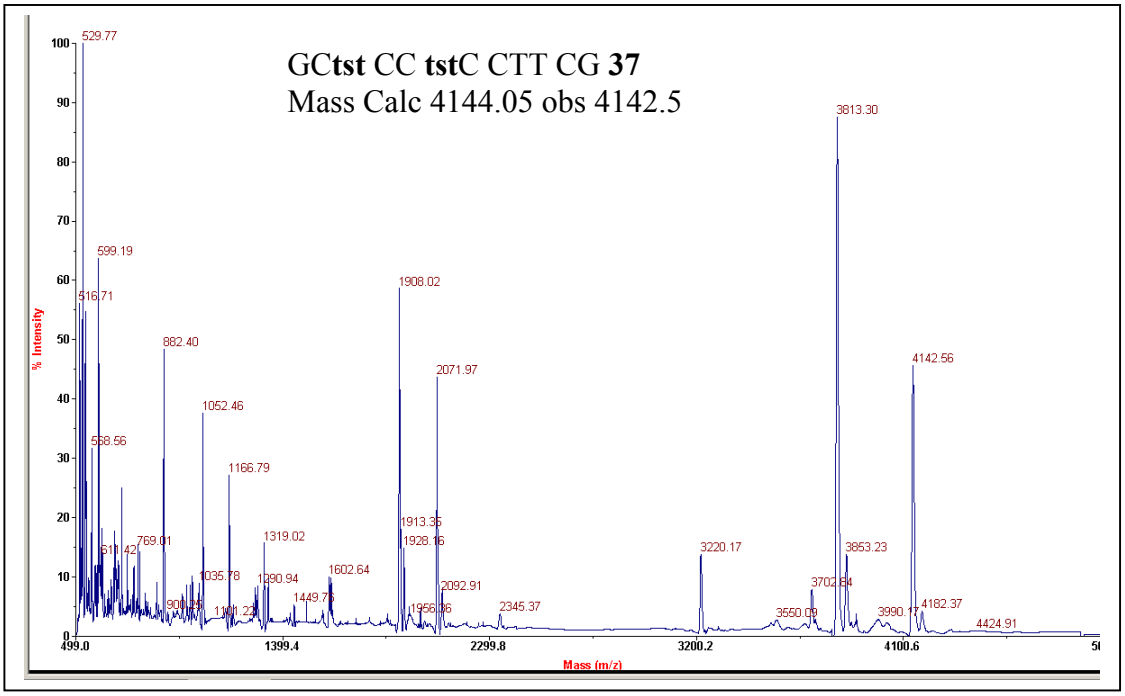

GAA GGG cstT tstG AA cst cst T 32
Mass calc 5846.6
obs 5846.8
% Intensity
Mass (m/z)
1450.10
1167.61
1053.57
854.95
1336.21
1732.28
1563.63
1036.14
1618.47
1318.51
2014.87
2927.02
2128.42
2295.62
2937.79
5846.78
5861.75
799.0
2039.4
3279.8
4520.2
5760.6

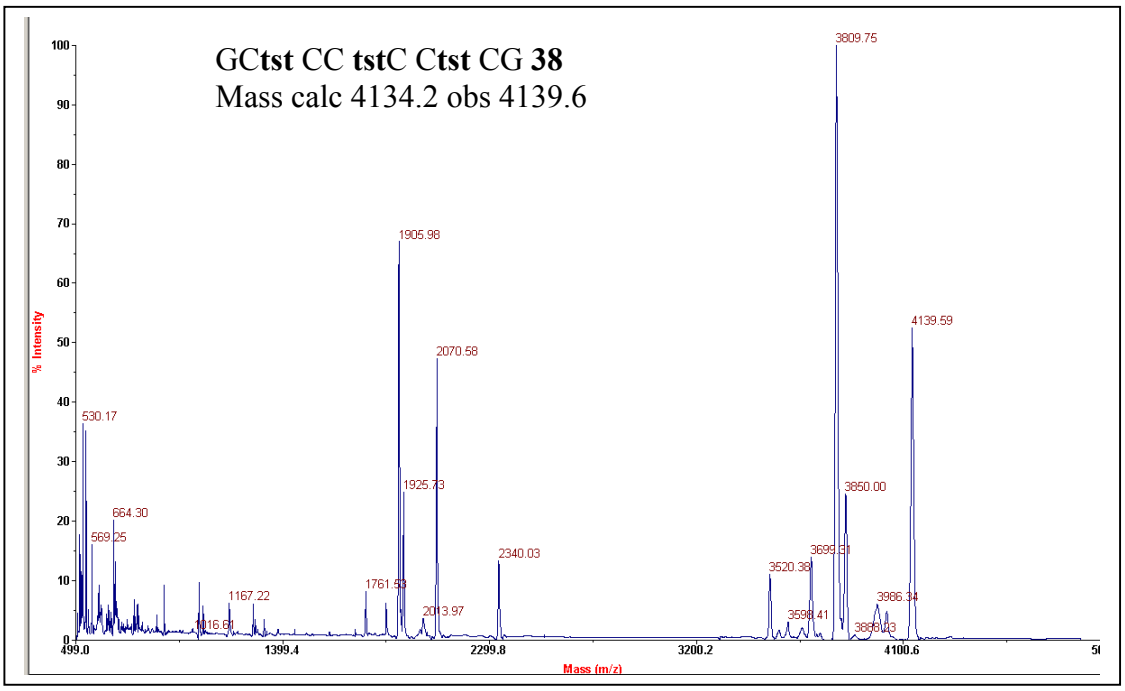

GAA GGG **cst**T **tst**G AA CT **cst** T **33**
Mass calc 5842.4
obs 5842.66
% Intensity
Mass (m/z)
1448.40
853.51
882.55
1165.82
1730.79
1334.84
1562.22
1616.60
1844.77
2924.65
5842.66
2012.72
2295.68
2934.54
5865.16
6108.06


AGA G **tst** CAA AAG CCC **tst**T T **34**
Mass calc 5774.2
obs 5774.4
% Intensity
Mass (m/z)
5774.3872
2866.9532
5811.1781
5631.0075
2884.4671
2087.1183
5663.7137
5796.0125
5781.0246
2802.3314


AGA G tst CAA AAG CCcst tst T 35
Mass calc. 5779.3 obs.5779.7
5779.7230
5184.3234
3226.5225
% Intensity
Mass (m/z)


GCtst CC TTC CTT CG 36
Mass calc 4153.9 obs 4147.9
4147.91
2075.12
543.60
529.66
515.72
1909.92
1998.33
3818.43
3999.27
4250.60
% Intensity
Mass (m/z)


GC**tst** CC **tst**C CTT CG **37**
Mass Calc 4144.05 obs 4142.5
% Intensity
Mass (m/z)
529.77
599.19
516.71
882.40
1908.02
2071.97
1052.46
568.56
1166.79
1319.02
1913.35
1928.16
1602.64
2092.91
2345.37
3220.17
3813.30
3853.23
4142.56
4182.37
4424.91


GC**tst** CC **tst**C C**tst** CG **38**
Mass calc 4134.2 obs 4139.6
% Intensity
Mass (m/z)
3809.75
1905.98
4139.59
2070.58
530.17
1925.73
3850.00
664.30
569.25
2340.03
1167.22
1761.53
2013.97
3986.34

2.2.7 References Section I & II

1. (a) Uhlmann, E. and Peyman, A. (1990) *Chem.Rev.*, **90**, 543-584. (b) Freier, S. M. and Altmann, K.-H. (1997) *Nucleic Acids Res.*, **25**, 4429-4443. (c) Micklefield, J. (2001) *Curr. Med. Chem.*, **8**, 1157-1179 (d) Kurreck, J. (2003) *Eur. J. Biochem.*, 270, 1628-1644. (e) Turner, J. J., Fabani, M., Arzumanov, A. A., Ivanova, G. and Gait, M. J. (2006) *Biochim. Biophys. Acta.*, **1758**, 290-300. (f) Kumar, V. A. and Ganesh, K. N. (2007) *Curr. Topics in Med Chem.*, **7**, 715-726.
2. (a) Dominski, Z. and Kole, R. (1993) *Proc. Nat. Acad. Sci U.S.A.*, **90**, 8673-8677. (b) Sazani, P. and Kole, R. (2003) *J. Clin. Invest.,* **112***,* 481–486.
3. References quoted in Chapter1.
4. (a) Egholm, M., Buchardt, O. and Nielsen, P. E. (1992) *J. Am. Chem. Soc.*, **114**, 1895-1897. (b) Nielsen, P. E., Egholm, M. and Buchardt, O. (1994) *Bioconjugate Chem.*, **5**, 3-7.
5. Uhlmann, E., Peyman, A., Breipohl, G. and Will, D. W. (1998) *Angew. Chem. Int. Ed*., **37**, 2796-2823.
6. Bendifallah, N., Rasmussen, F. W., Zachar, V., Ebbesen, P., Nielsen, P. E.and Koppelhus, U. (2006) *Bioconjugate Chem.*, **17***,* 750-758.
7. (a) Zhou, P., Wang, M., Du, L., Fisher, G. W., Waggoner, A. and Ly, D. H. (2003) *J. Am. Chem. Soc*., **125**, 6878 - 6879. (b) Englund, E. A. and Appella, D. H. (2007) *Angew. Chem. Int. Ed.*, **46**, 1414-1418.
8. (a) Pallan, P. S., von Matt, P., Wilds, C. J., Altmann, K.-H. and Egli, M. (2006) *Biochemistry***, 45**, 8048-8057. (b) Wilds, C. J., Minasov, G., von Matt, P., Altmann K.-H. and Egli, M. (2001) *Nucleosides, Nucleosides and Nucleic Acids,* **20**, 991-994.
9. (a) Govindaraju, T. and Kumar, V. A. (2005) *Chem. Commun*., 495- 497 (b) Govindaraju, T. and Kumar, V. A. (2006) *Tetrahedron*, **60**, 2321–2330.
10. Thibaudeau, C., Plavec, J., Garg, N., Papchikhin, A. and Chattopadhyaya, J. (1994*) J. Am. Chem. Soc.* **116**, 4038-4043.
11. (a) Ding, D., Gryaznov, S.M., Lloyd, D.H., Chandrasekaran, S., Yao, S.,Ratmeyer, L., Pan, Y. and Wilson, W.D. (1996) An oligodeoxyribonucleotides N3'→P5' phosphoramidate duplex forms an A-type helix in solution, *Nucleic Acids Res.,* **24**, 354-360. (b) Ding, D., Gryaznov, S. M. and Wilson, W. D. (1998) NMR solution structure of the N3'→P5' phosphoramidate duplex d(CGCGAATTCGCG) by iterative relaxation matrix approach, *Biochemistry*, **37**, 12083-12093.

12. Gryaznov, S. M., Lloyd, D.H., Chen, J.-K., Schultz, R.G., DeDionisio, L.A., Ratmeyer, L. and Wilson, W.D. (1995) Oligonucleotide N3'→P5' phosphoramidates, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5798-5802.
13. Tereshko, V., Gryaznov, S. and Egli, M. (1998) Consequences of replacing the DNA 3P-oxygen by an amino group: high-resolution crystal structure of a fully modified N3'→P5' phosphoramidate DNA dodecamer duplex. *J. Am. Chem. Soc.,* **120**, 269-283.
14. Prakash, T. P., Manoharan, M., Kawasaki, A. M., Fraser, A. S., Lesnik, E. A., Sioufi, N., Leeds, J. M., Teplova M. and Egli, M. (2002) *Biochemistry,* **41***,* 11642-11648.
15. Kawai, S. H., Wang, D., Giannaris, P. A., Damha, M. and Just, G. (1993) Solid phase synthesis and hybridization properties of DNA containing sulfide-linked dinucleoside. *Nucleic Acids Res.*, **21**, 1473-1479.
16. Meng, B., Kawai, S. H., Wang, D., Just, G., Giannaris, P. A. and Damha, M. (1993) A sulfide linked oligonucleotide analogue with selective hybridization properties. *Angew. Chem Int. Ed. Engl*., **32**, 729-731.
17. Wilds, C. J., Minasov, G., Natt, F., Matt von, P., Altmann, K.-H. and Egli, M. (2001) Studies of a chemically modified oligonucleotide containing a 5-atom backbone which exhibits improved binding to RNA. *Nucleosides Nucleotides Nucleic Acids*, **20**, 991-994.
18. Pallan, P. S., Matt von, P., Wilds, C., Altmann, K.-H. and Egli, M. (2006) RNA binding affinities and crystal structure of oligonucleotides containing five-atom amide-based backbone structures. *Biochemistry*, **45**, 8048-8057
19. Lauristen, A. and Wengel, J. (2002) Oligonucleotides containing amide linked LNA type dinucleotides: synthesis and high affinity. *Chem. Commun.*, 530- 531.
20. Chen, J.-K., Schultz, R.G., Lloyd, D.H. and Gryaznov, S.M.(1995) Synthesis of oligodeoxyribonucleotide N3'→P5' phosphoramidates. *Nucleic Acids Res.,* **23**, 2661-2668.
21. Saneyoshi, M., Fujii,T., kawagchi,T., Awai, K. and Kimura, S. (1991) In Townsend, L.B. and Tipson, R. S.(Eds), *Nucleic Acid Chemistry.* Wiley, NewYork, NY, **4**, 67-72.
22. Divakar, K. J. and Reese, C. B. (1982) *J. Chem. Soc. Perkin Trans*., **1**, 1171-1176.
23. Kraszewski, A. and Stawinski, J. (1980) *Tetrahedron Lett.*, **21**, 2935.
24. (a) Egholm, M., Buchardt, O. and Nielsen, P. E. (1992) Peptide nucleic acids (PNAs): Oligonucleotide analogues with an achiral peptide backbone. *J. Am. Chem. Soc*., **114**, 1895-1897. (b) Egholm, M., Nielsen, P. E., Buchardt, O. and Berg, R. H. (1992) Recognition of guanine and adenine in DNA by cytosine and thymine containing peptide nucleic acids (PNAs). *J. Am. Chem. Soc*., **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. and Buchardt, O. (1994) Synthesis of peptide nucleic acid monomers containing the four natural bases: Thymine, cytosine, adenine, and guanine and their oligomerization. *J. Org. Chem*., **59**, 5767-5773.

25. Merrifield, R. B. (1963) Solid phase synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc*. **85**, 2149-2154.
26. (a) Mckay, F. C. and Albertson, N. F. (1957) New amine-masking groups for peptide synthesis. *J. Am. Chem. Soc*., **79**, 4686-4690. (b) Anderson, G. W. and McGreoger, A. C. (1957) *t*-Butyloxycarbonyl aminoacids and their use in peptide synthesis. *J. Am. Chem. Soc*., **79**, 6180-6183.
27. Carpino, L.A. and Han, G.Y. (1972) The 9-fluorenylmethoxycarbonyl amino-protecting group. *J.Org. Chem.* **37**, 3404-3409.
28. Bernhardt, A., Drewello, M. and Schutkowski, M. (1997) The solid-phase synthesis of side-chain-phosphorylated peptide-4- nitroanilides. *J. Peptide Res.,* **50**,143- 152.
29. Matsueda, G. R. and Stewart, J. M. (1981) A p-methylbenzhydrylamine resin for improved solid-phase synthesis of peptide amides. *Peptides*, **2**, 45- 50.
30. Rink, H. (1987) Solid-phase synthesis of protected peptide fragments using a trialkoxy-diphenyl-methylester resin. *Tet. Lett*., **28**, 3787-3790.
31. Sheehan, J. C. and Hess, G. P. (1955) A New Method of Forming Peptide Bonds. *J. Am. Chem. Soc*., **77**, 1067.
32. (a) Carpino, L. A. (1993) 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive J. *Am. Chem. Soc.,* **115**, 4397-4398. (b) Albericio, F., Bofill, J. M., Faham, A. E.- and Kates, S. A. (1998) Use of Onium Salt-Based Coupling Reagents in Peptide Synthesis. *J. Org. Chem*., **63**, 9678.
33. Marder, O. *et. al.* (1990) *Chimica Oggi*, 37.
34. Coste, J., Dufour, M. N., Pantaloni, A. and Castro B. (1990) Brop: A new reagent for coupling N-methylated amino acids. *Tet. Lett*, **31**, 669-672.
35. Hudson, D. (1988) Methodological implications of simultaneous solid-phase peptide synthesis. 1. Comparison of different coupling procedures, *J. Org. Chem*, **53**, 617-624.
36. Kaiser, E., Colescott, R. L., Bossinger, C. D. and Cook, P. I. (1970) Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.,* **34**, 595-598.
37. Oded, A. and Houghten, R. A. (1990) *Pept. Res*, **3**, 42.
38. Fronenot, J. D. *et. al.* (1991) *Pept. Res.*, **4**, 19.

39. Vojkovsky, T. (1995) *Pep. Res*., **3**, 42.

40. Blackburn, C. (2005) Solid-phase synthesis of 2-amino-3-chloro-5- and 8-nitro-1, 4-naphthoquinones: a new and general colorimetric test for resin-bound amines. *Tet Lett*., **46**, 1405-1409.

41. Erickson, B. W. and Merrifield, R. B (1976) Solid Phase Peptide Synthesis. In the Proteins Vol. II, 3rd ed.; Neurath, H. and Hill, R. L. eds.; Academic Press, New York, pp 255.

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

43. (a) Gait, J. M. (1984) Oligonucleotide synthesis: A practical approach. IRL Press Oxford, UK 217. (b) Agrawal, S (1993) in Protocols for oligonucleotides and analogs: Synthesis and properties. *Methods in Molecular Biology*. Agrawal, S. (ed): vol 20: Totowa, NJ. Humana Press, Inc.

44. (a) Job, P. (1928) *Ann. Chim*., **9**, 113-203. (b) Cantor, C. R. and Schimmel, P. R. (1980) *Biophys. Chem.*, *Part III,* 624.

45. Kim, S. K., Nielsen, P. E., Egholm, M., Buchardt, O., Berg R. H. and Norden, B.(1993) *J. Am. Chem. Soc*., **115**, 6477-6481.

46. Zaitseva *et. al.* (1994) Chemical- enzymatic synthesis of 3'-amino-2',3'-dideoxy-β-D-ribofuranosides of natural heterocyclic bases and their 5'-monophosphates. *Nucleosides& Nucleotides*, **13***,* 819-834.

47. Zielinska, D., Pongracz, K. and Gryaznov, S.M. (1996) A new approach to oligonucleotide N3' → P5' phosphoramidate building blocks. *Tett. Lett.,* **47**, 4495-4499.

48. (a) Nowell, P. and Hungerford, D. (1960) A minute chromosome in human chronic granulocytic leukemia. *Science*, **132**, 1497. (b) Deininger, M. W. N., Goldman J. M. and Melo J. V. (2000) The molecular biology of chronic myeloid leukemia*. Blood*, **96**, 3343-3356.

Chapter 3

Conformational studies of TANA monomers and dimers by NMR spectroscopy

3.1 Introduction

The furanose ring is centrally located in the sugar phosphate backbone of nucleic acids.[1] These pentose sugar rings are twisted or 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 C1'-O4'-C4'. Thus, if the *endo*-displacement of C-2' is greater than the *exo*-displacement of C-3', the conformation is called C2'-*endo* and so on for other atoms of the ring (Figure 1). The *endo*-face of the furanose is on the same side as C5' and the base; the *exo*-face is on the opposite face to the base. The sugar puckers are located in the north (N) and south (S) domains of the pseudorotation cycle of the furanose ring.[1] In solution, N-type and S-type conformations are in rapid equilibrium and are separated by low energy barrier.

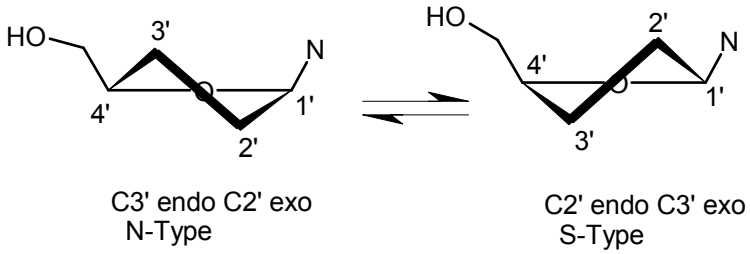


Figure 1. Puckering of sugar ring

The conformation of the pentofuranose ring in a nucleoside moiety can be fully described in terms of the phase angle of pseudorotation (P) and the puckering amplitude (ϕ).[2] From X-ray[3] studies on nucleosides and nucleotides, it was found that ϕ values range from 35º to 45º. For North-type (N) sugars (C3'-endo, C2'-exo), P ranges from -1º to 34º and for S-type (S) sugars (C2'-endo, C3'-exo), P ranges from 137º to 194º. In solution, the sugar rings exist in equilibrium of the two rapidly interconverting conformers N↔S. 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 ϕ_N and ϕ_S, can be calculated from the vicinal proton-proton ($^3J_{HH}$) coupling constants $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$ and $J_{3'4'}$.[2,4] These

coupling constants can be used as an input for the pseudorotation analysis of the sugar using the program PSEUROT.[5] Alternatively, a semi empirical method termed as "Sum rule" [6] can also be employed to have some idea about N↔S equilibrium in case of sugar conformations.

3.1.1 Calculation of ribose ring conformation

3.1.1.1 Pseurot 5.4.1 program

The vicinal coupling constant ($^3J_{x\text{-}y}$) was used to derive the pseudorotation phase angle (P) and puckering amplitude (ϕ_m) of pyrrolidine rings in different compounds by using PSEUROT programming version (5.4.1),[5a] based on the relation between $\phi_{(H\text{-}H)}$ and P as equation (1)

$$\phi_{(H\text{-}H)} = A * \phi_{Max} * \cos (P + \text{phase}) + B \text{---------- (1)}$$

Where $\phi_{(H\text{-}H)}$ = torsional angle between two adjacent hydrogens, P = pseudorotation angle, ϕ_{Max} = puckering amplitude and A and B are constants. For the 5-membered ribose ring the parameters phase angles P and B values of corresponding vicinal proton pairs were used as in the software (PSEUROT 5.1.4). The role of electronegativities of substituents in 5- membered rings is important in determining the puckering of rings. The electronegativity values of different substituents in ribose ring shown in Table 1 and 2 were used in PSEUROT 5.4.1 program to find the conformation of the ribose rings.

The following group electronegativities (L values) are recommended assuming a reasonably normal nucleoside or nucleotide. L values of lone pair bearing atoms are somewhat solvent dependent.

Table 1. Group electronegativities of substituents or ring carbons of ribose ring

Substituents or ring carbon	L value
H (in CH_2 or CH groups)	0.00
C1' (C-C/O)	0.62
C2'(Ribo etc.)	0.62
C2' (2' deoxy)	0.67
C3'	0.62
C4'	0.62
C5'	0.68
CH3	0.80
N-bases (A, G, C, U, T)	0.56

Table 2. Group electronegativities (L value) substituents (solvent dependent)

Substituent on ribose ring	L(D_2O)	L($CDCl_3$)	L(DMSO)
OH	1.25	1.34	1.42
OR	1.26	1.40	1.42
But OR/OR	1.41	1.41	1.42
OAr	1.34	1.42	1.47
OPO_2 OR	1.25	---	---
OCOR	1.17	1.17	1.22
NR_2	1.01	1.12	1.20
NHR	1.02	1.16	1.22
NHAr	1.12	1.16	1.20
NH_2	1.10	1.19	1.27
NHCOR	0.81	0.86	---
NHCOH	0.53	0.54	---
A, G, C, T, U	0.56	0.56	0.56

The NMR measured (experimental) values of $^3J_{H\text{-}H}$ of all protons in proline ring in all compounds were used as input values of PSEUROT 5.1.4 program. The input parameters PSEUROT 5.4.1 are $^3J_{H\text{-}H}$ (experimental), phase angle, A and B along with electronegativities values of substituents as L_1, L_2, L_3 and L_4 in ribose ring and are shown in Table 3.

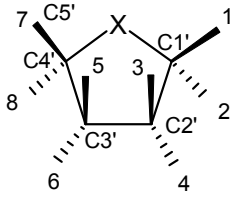


Table 3. General pattern of Input values of PSEUROT 5.4.1

Coupling	Phase(º)	A	B (º)	L_1	L_2	L_3	L_4
1 - 3	-144	1	0	x	2	4	C3’
1 - 4	-144	1	-120	x	2	C3’	3
2 – 3 (1’2’)	-144	1	+120	1	x	4	C3’
2 – 4 (1’2”)	-144	1	0	1	x	C3’	3
3 – 5 (2’3’)	0	1	0	C1’	4	6	C4’
3 – 6	0	1	-120	C1’	4	C4’	5
4 – 5 (2”3’)	0	1	+120	3	C1’	6	C4’
4 – 6	0	1	0	3	C1’	C4’	5
5 – 7	+144	1	0	C2’	6	8	x
5 – 8 (3’4’)	+144	1	-120	C2’	6	x	7
6-8	+144	1	0	5	C2’	x	7

For ribo, deoxyribo and proline more accurate A and B values are available from X-ray data.

The phase angle, A and B along with electronegativities from Tables 1 & 2 and J_{exp}, were used as the input in PSEUROT program to obtain the output values as $^{3}J_{x\text{-}y}$ (calculated), P_N, ϕ_N, P_S, ϕ_S, MF (mole fraction) and dihedral angle ($\phi_{1\text{-}3}$) of all vicinal pairs of ribose ring for two most probable conformers. A typical format of **PSEUROT 5.4.1** is given follows, as five steps:

Step1: Parameters in pseudorotation relation as vicinal coupling pairs, phase angle of equation karplus equation, A, B, electronegativities of substitution and no. of substitution.

Step 2: Input data

J --> 1'-2' 1'-2" 2'-3' 2"-3' 3'-4'

Step 3: First estimates

Conformer 1: P_N and ϕ_N

Conformer 2: P_S and ϕ_S

Mole fraction of conformer 2 (MF2) = 0.500

Step 4: Final outputs

1'-2' Jexp; Jcal; ΔJdiff
1'-2" Jexp; Jcal; ΔJdiff
2'-3' Jexp; Jcal; ΔJdiff
2"-3" Jexp; Jcal; ΔJdiff
3'-4' Jexp; Jcal; ΔJdiff
MF1 MF2 RMS

Conformer 1(*N*):		Conformer 2 (*S*):	
P =		P =	
ϕ		ϕ	
ϕ 1'-2' ;	J1'-2'	ϕ 1'-2' ;	J1'-2'
ϕ 1'-2";	J1'-2"	ϕ 1'-2" ;	J1'-2"
ϕ 2'-3' ;	J2'-3'	ϕ 2'-3' ;	J2'-3'
ϕ 2"-3';	J2"-3'	ϕ 2"-3';	J2"-3'
ϕ 3'-4' ;	J3'-4'	ϕ 3'-4' ;	J3'-4'

Step 5: Error analysis

Overall RMS

Standard deviation on parameters

Correlation matrix of parameters

AR. 1 2 3 4 5

End of program PSEUROT

3.1.1.2 Sum Rule

In some cases, spectral overlap made it impossible to determine all these coupling constants accurately over a range of different temperatures. In such cases, PSEUROT program can not be performed and a more uniform graphical method devised by Altona, [6] which relies on analysis of the relative magnitude of the $J_{1',2'}$ coupling constants and the sums of the H1' coupling constants (Σ1') which is related to the percentage of S conformer (% S) by the "sum rule" (equation 2).

$$\% S = (\Sigma Hl'-9.8)/ 5.9 \quad \text{where } \Sigma l' = J_{1',2'} + J_{1',2''} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

3.2 Rationale for the present work and objectives

The aim of this work is to study the deoxyribose ring conformation in the thioacetamido nucleic acids (TANA) monomers and dimers described in chapter 2. The specific binding of the TANA ONs with complementary RNA could be because of the nature of the 3'- and 5'-substituents of the deoxyribose sugar that would affect the sugar ring pucker expressed as the extent of the N↔S equilibrium. The modified oligonucleotides with prospensity for northern sugar conformations could be preorganized into the required conformation for A-type duplexes, and are therefore more likely to form stable heteroduplexes with complimentary RNA.[7]

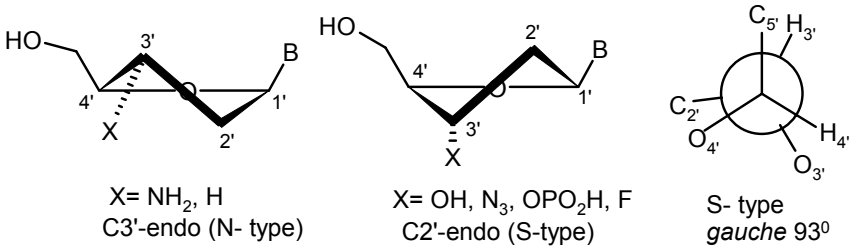


Figure 2. Effect of 3'-substituents on sugar ring conformation and Newmann projection of S-type conformation

The nature of the 3'-α- substituents has an influence on the conformation of the sugar ring. Upon substitution of the 3'-oxygen by a 3'-azide or amine group, the N↔S equilibrium is shifted toward more population N-type pseudorotamer (Figure 2).[9] This effect is also noticeable when the phosphate functionality is replaced by sulfonamide

functionality, affected the pseudorotamer equilibrium towards less population of N-type pseudorotamer.[9] This may be partly explained by the *gauche* effect which means that the favored adopted structure could be the one in which the O4' and 3'-α-substituents are in a *gauche* orientation. In S-type 2'-deoxysugars, the O4' and 3'- α substituents are in a *gauche* orientation, and the preference in such cases is for S conformation. The higher electronegativity of oxygen versus nitrogen leads to a greater preference for a X3'-C3'-C4'-O4' *gauche* orientation leading to greater prospensity for % S-type pucker. The *gauche* effect increases with the higher polarization of the C3'-X bond and therefore the % S increases with the increased electronegativity of the 3'-substituent and vice versa. The *gauche* effect was clearly observed in 3'-α-substituted -2'-deoxynucleosides such as 3'- α-flurothymidine (90% S-type) and 3'- α-azidothymidine (50% S-type).[8,9] This fact suggests that in 2'-deoxy sugars the polarity of the C3'-X3' bond dictates the S-type/N-type conformational preference based on the strength of the gauche effect (Figure 2).[9]

J. Chattopadhyaya and coworkers demonstrated that the 3'-hydroxy-2'-deoxynucleosides prefer C2'-endo, or S-type conformations. The shift of N↔S sugar conformation equilibrium more towards S-type for the corresponding 3'-O-monophosphates suggests the preference of the S sugar conformer in the native B-DNA.[10] The gauche effect between O 4' and 3'-OPO_3H group is responsible for this equilibrium shift toward S-type. In contrast, high-resolution proton NMR study demonstrated that the 3'-amino 2',3'-

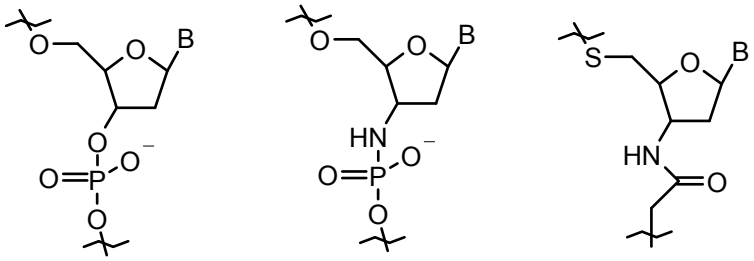


Figure 3. Structure of DNA, N3'-P5' Phosphoramidates and TANA

dideoxy nucleoside the furanose conformation is predominantly C3'- endo, or N-type, for the free 3'-aminonucleosides [11] and also for 3'-aminonucleosides incorporated into the single and double stranded oligonucleotide N3'→P5' phosphoramidates (Figure 3).[12]

These 2'-deoxy 3'-amino phosphoramidates are known to form stable complexes with complementary RNA (Figure 3).[12c]

In the present study of TANA oligomers, the 3'-amino substituent is less elctroneagitive than oxygen in the native 2'-deoxy sugars and would give rise to a comparatively more northern (N-type) C3'-endo ribose conformation. It is therefore anticipated that the electrically neutral acetamido group at 3'-position might adopt similar N-type conformation as in the case of N3'- P5' phosphoramidate oligomers, at the same time alleviating the electrostatic repulsion which exists between the negatively charged phosphate groups in the complementary strands of the target DNA oligomers.

The substitution of oxygen by sulfur at 5'-position will also have some effects on the ring conformation and thus it is important to study the conformational preference of the deoxyribose in TANA monomers and dimers which would indicate the effect of modified TANA backbone on conformational preferences in oligomers. It is thus essential to know the exact coupling constants $J_{1'2'}$.

The specific objectives of this work are:

i) Synthesis of TANA monomers and dimers with appropriate protecting groups.

ii) NMR assignment of deoxyribose ring protons of TANA derivatives by a combination of 1D proton decoupling and 2D-NMR experiments such as COSY, TOCSY and NOESY.

iii) Determination of ^{3}J coupling constants from 1D proton spectra and homonuclear proton decoupling experiment.

iv) Determination of the conformation of deoxyribose sugar ring of TANA derivatives by comparative *J* values obtained from pseurot (PSEUROT 5.4.1) program as well as by employing the sum rule.

3.3 Present Work

To determine the conformational preferences of the deoxyribose ring in TANA monomers and dimers, it is essential to obtain the corresponding coupling constants required for the calculations using Pseurot program as well as "sum rule". The TANA monomers and dimers chosen for the conformational analysis are shown in Figures 4 and 9, respectively. All the compounds were synthesized from the intermediates described in Chapter 2. Compounds **1** and **2** were synthesized to study the effect of 3'-NH substitution on the sugar ring. Analysis of compounds **6** & **7** would indicate the effect of 5'-S-CH_2COOEt substituents and that of compounds **3**, **4** and **5** would give the effect of 3'-& 5'-modifications. As the TANA oligomers are linked by an amide linkage, the 3'- NH_2 group in compounds **1-5** was protected as acetamide derivative. The silyl protected dimers **8** and **10** and 5', 3' OH free dimer **9** were synthesized by following the same procedures as in chapter 2.

3.3.1 Assignments of ^{1}H NMR spectra of TANA monomers

Conformational analysis was done for the thyminyl, cytosinyl and 5-methyl cytosinyl monomers depicted in Figure 4.

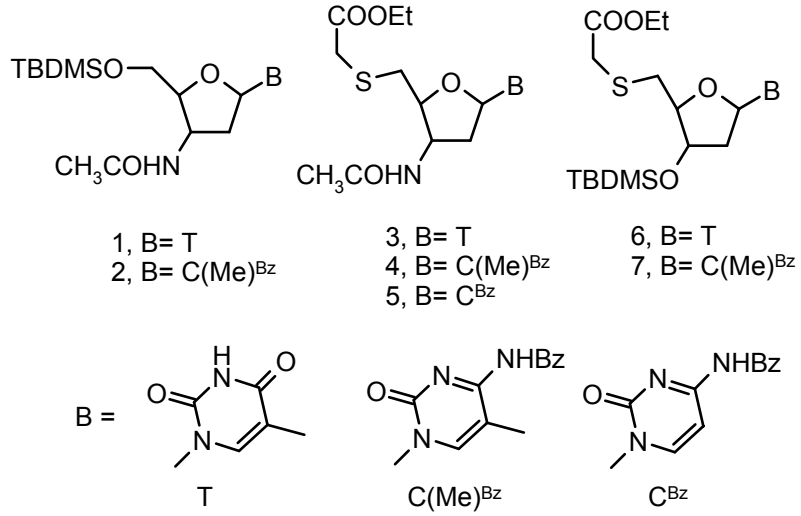


Figure 4. Structure of TANA monomers for conformational studies.

The assignment of the non- exchangeable proton resonances was achieved by using 2D COSY experiments (Table 4). The H2'/H2" and H5'/H5" protons were assigned

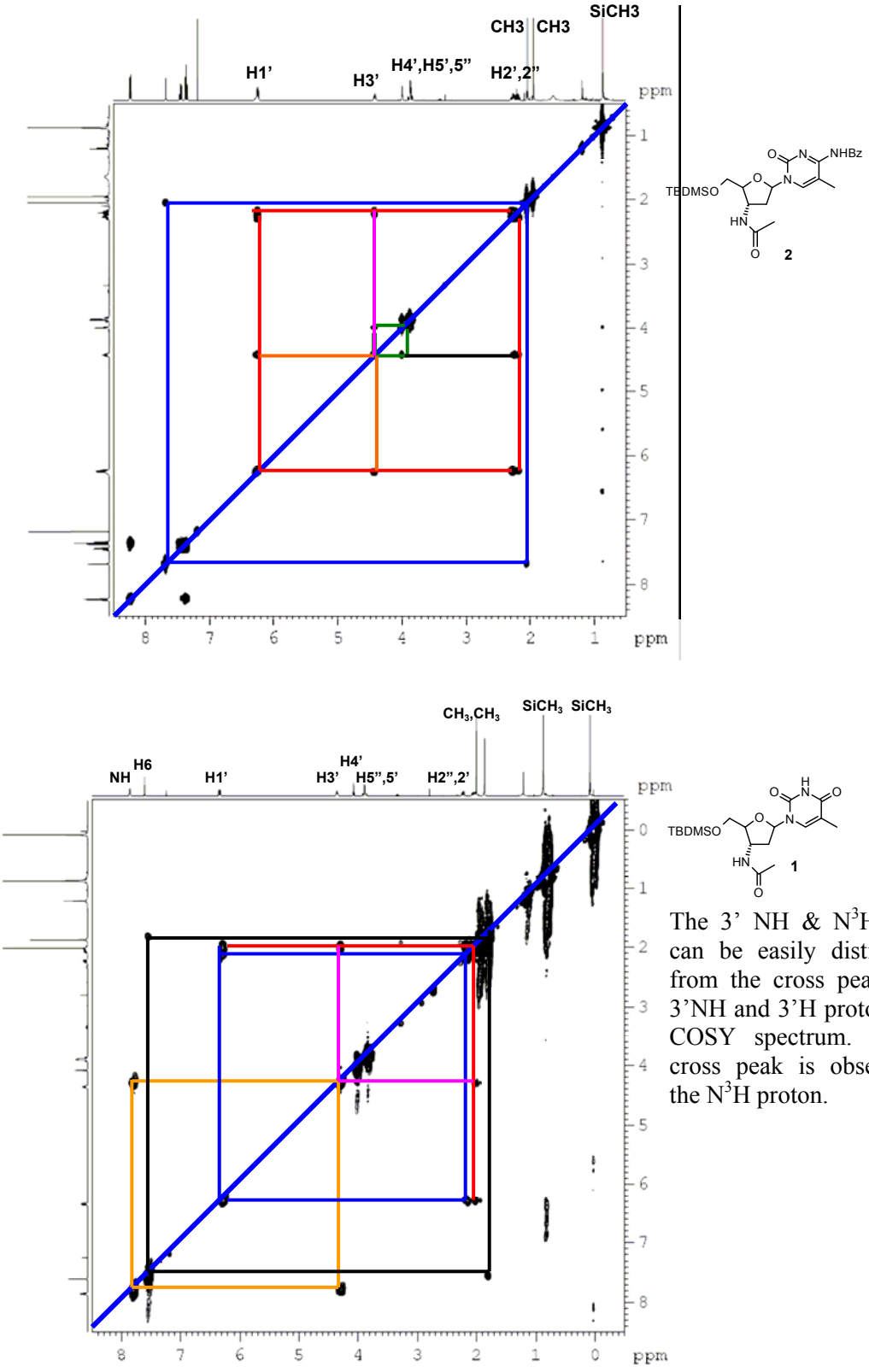
according to the Remin and Shugar rule[13] where downfield proton at C2' or C5'were assigned as H2" or H5" while the upfield protons are assigned as the H2' or H5'. The coupling constants were determined from 1D ^{1}H NMR spectra and from homonuclear proton decoupled spectra where spectral overlap was severe (Table 5).

Table 4. Chemical shift value of TANA monomers

Entry	Compound	1'	2'	2''	3'	4'	5'	5''	S CH_2	H5	H6	CH_3	C_6H_5	CO CH_3	Si-CH_3	3' NH
1	**2**	6.3	2.1	2.2	4.4	4.0	3.8	3.9	-	-	7.7	1.9	7.4, 8.2	2.0	0.06, 0.89	
2	**1**	6.3	2.0	2.2	4.3	4.0	3.8	3.9	-	-	7.6	1.8	-	2.0	0.06, 0.85	
3	**3**	6.3	2.2	2.3	4.4	4.1	3.1	3.1	3.3	-	7.5	1.9	-	2.0	-	10.1
4	**6**	6.2	2.1	2.3	4.3	4.0	2.9	3.0	3.3	-	7.3	1.9	-	-	0.06, 0.86	9.4
5	**7**	6.2	2.2	2.6	4.2	4.1	3.0	3.0	3.3	-	-	-	7.5, 8.2	-	0.08, 0.89	-
6	**4**	6.3	2.3	2.4	4.4	4.1	3.0	3.1	3.3	-	7.9	2.2	7.5, 8.3	2.0	-	7.3
7	**5**	6.3	2.3	2.6	4.5	4.3	3.0	3.1	3.3	-	-	-	7.6, 8.3	2.0	-	7.3

Table 5. $^{3}J_{HH}$ ^{1}H NMR coupling constant of TANA monomers

Entry	Compound	1'2'	1'2"	2'2"	2'3'	2"3'	3'4'	4'5'	4'5"	5'5"
1	**1**	6.2	7.8	12.8	6.7	4.6	3.9	1.8	-	13.5
2	**2**	7.8	5.9	13.8	7.0	3.1	5.6	2.6	2.3	11.5
3	**3**	6.8	7.0	13.7	8.5	4.2	5.4	6.0	2.0	11.4
4	**4**	6.4	6.3	13.8	7.0	7.2	6.6	6.3	4.0	14.2
5	**5**	6.1	6.2	13.8	7.1	5.8	6.1	6.6	3.9	14.1
6	**6**	6.5	6.5	13.5	6.8	4.3	4.4	5.5	5.0	14.0
7	**7**	5.9	6.0	13.8	6.2	5.5	5.4	6.0	5.5	-



The 3' NH & N^3H protons can be easily distinguished from the cross peak due to 3'NH and 3'H protons in the COSY spectrum. No such cross peak is observed for the N^3H proton.

Figure 5. COSY spectrum of **2** and **1**

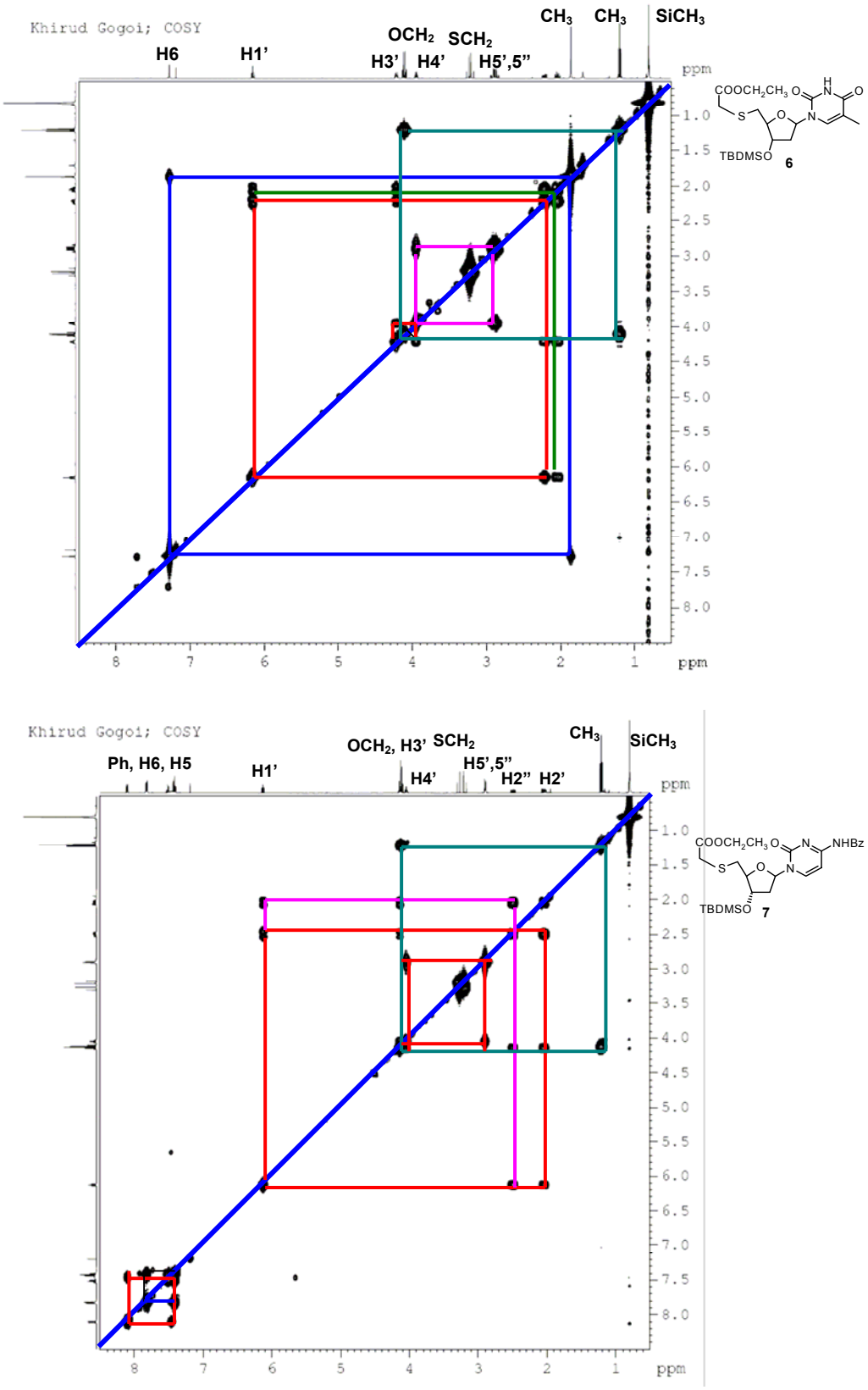


Figure 6. COSY spectrum of **6** and **7**

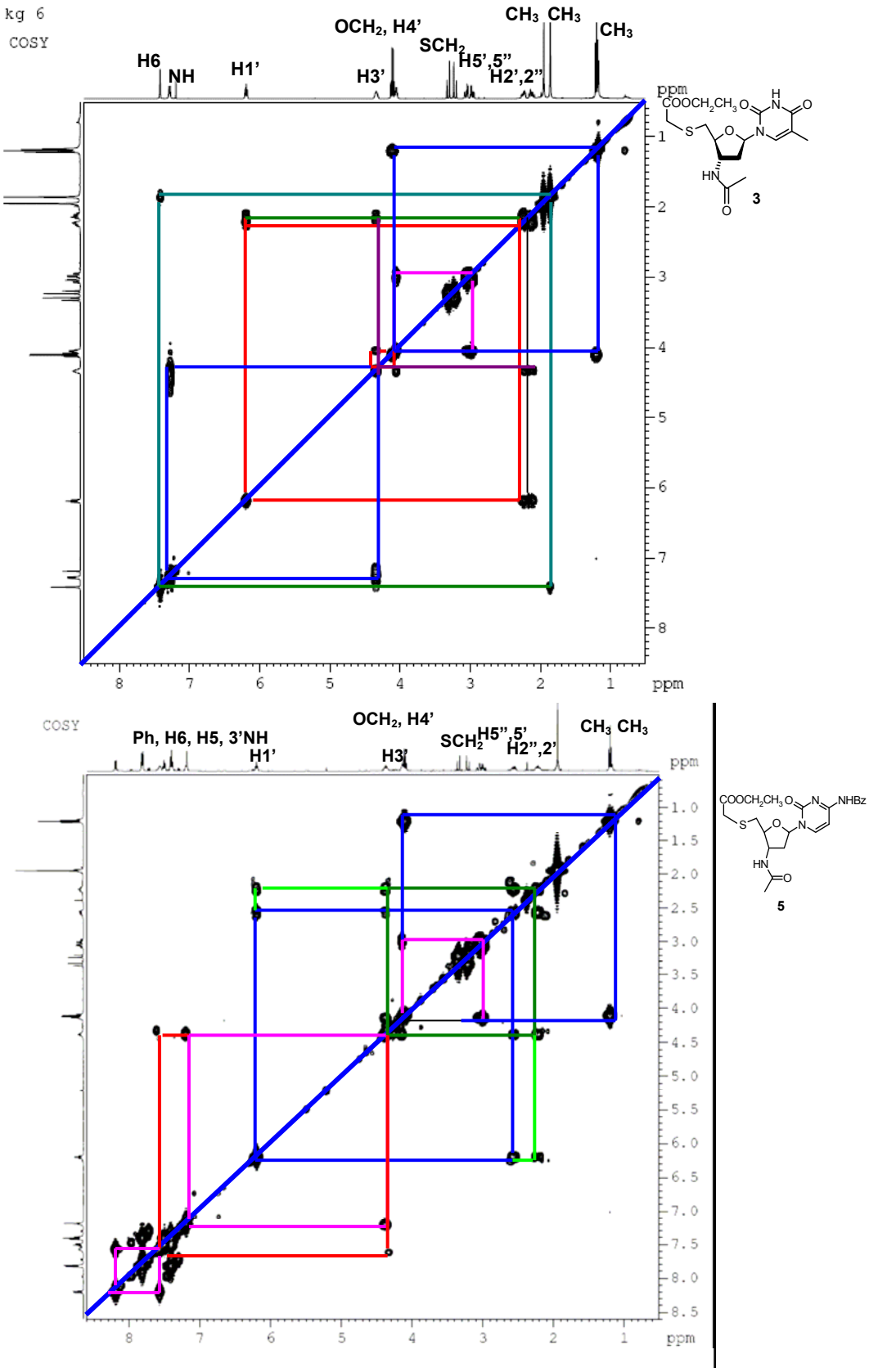


Figure 7. COSY spectrum of **3** and **5**

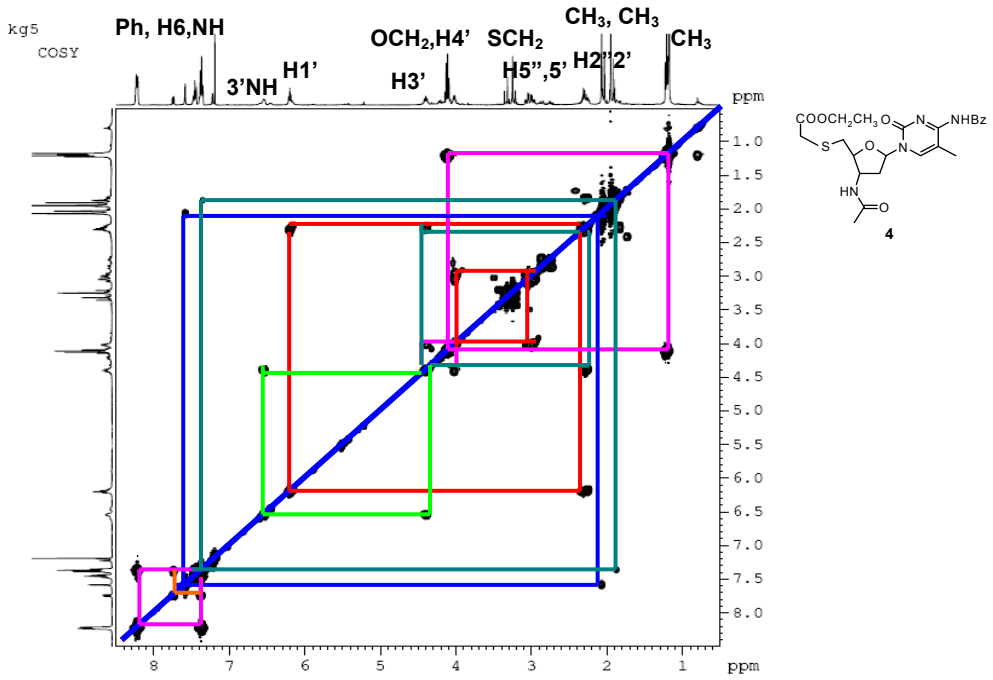


Figure 8. COSY spectrum of **4**

3.3.2 Assignments of ^{1}H NMR spectra of tst 8, 9 and cst 10 dimers

The 1D ^{1}H NMR spectra of **tst 9** and **cst 10** (Figure 9) were fully assigned with the aid of 2D COSY, TOCSY and NOESY experiments (Table 6). The two spin systems of the 5'- and 3'-terminal ribose rings were assigned by TOCSY and COSY experiments and distinguished by noticeable shielding of the H5'/5" protons adjacent to the less electronegative sulphur atom of the mercaptoacetamide linkage. The H2'/H2" and H5'/H5" protons were assigned according to the Remin and Shugar rule.[13] The conformation about the glycosidic bond was determined from 2D NOESY experiments. A pyrimidine nucleotide has a *syn* conformation when a strong nOe between the H6 and H1' together with a weak nOe between H6 and H2' and H3' is observed. A nucleotide with an S-type sugar is in *anti* conformation when a strong nOe between H6 and H2'is observed while a nucleotide with an N-type sugar has an *anti* conformation when a strong nOe between H6 and H3' is observed.[9] In the case of the **cst** dimer **10**, there are strong nOe cross peaks between H6-H3' observed which indicate the anti glycosidic conformations of the nucleobases (Figure 13). In **tst** dimer **9,** nOe cross peaks between H6-H3' observed only in the 3'-end (B ring) (Figure 17), which indicate the *anti* glycosidic conformation. Some of the major nOe interactions observed for **tst** and **cst**

dimers **8** and **10** were depicted in Figures 14 and 18. The coupling constants of tst **9** and cst **10** were determined from 1D spectra and from homonuclear proton decoupled spectra where spectral overlap was severe (Table 7).

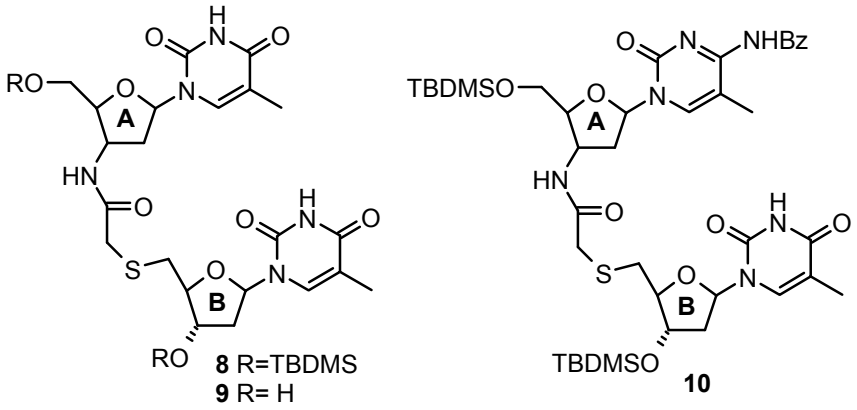


Figure 9. Structure of TANA dimers for conformational studies

Table 6. ^{1}H NMR chemical shifts (ppm) of **tst 8** (entry 3), **9**(entry 2) and **cst 10** (entry 1) dimer

Entry	Sugar	1'	2'	2''	3'	4'	5'	5''	S CH_2	H6	CH_3	C_6H_5	Si-CH_3	SiC $(CH_3)_3$	3' NH	3 NH
1	**cst 10 A**	6.3	2.2	2.4	4.4	4.0	3.8	3.9	-	7.9	2.1	7.4, 8.2	0.01	0.8	7.5	
	cst 10 B	6.0	2.2	2.3	4.2	3.9	2.8	2.9	3.2	7.2	1.8	-	0.01	0.8		13.3
2	**tst 9 A**	6.2	2.3	2.4	4.4	4.0	2.8	2.9	-	7.6	1.8					
	tst 9 B	6.1	2.3	2.4	4.3	3.9	3.7	3.8	3.3	7.4	1.8					
3	**tst 8 A**	6.6	2.7	2.2	4.6	4.1	4.0	3.9	3.3	7.5	1.9	-	0.13	0.98	7.5	1.3
	tst 8 B	6.2	2.6	2.2	4.3	4.0	3.0	2.8		7.5	1.9	-	0.10	0.93		11.5

Table 7. ^{1}H NMR $^{3}J_{HH}$ coupling constants (Hz) and % S (Sum rule) of **cst 10** (entry 1)**, tst 9** (entry 2) dimer

Sugar	1'2'	1'2"	2'2"	2'3'	2"3'	3'4'	4'5'	4'5"	5'5"
cst 10 A	8.5	5.5	13.3	7.3	2.9	4.4	2.0	2.3	11.5
cst 10 B	6.4	6.9	13.8	5.6	4.4	2.3	4.0	7.3	14.0
tst 9 A	6.8	6.8	13.1	7.0	5.0	6.6	4.8	2.5	12.8
tst 9 B	6.5	6.0	13.1	6.5	4.5	5.0	7.5	4.5	14.1

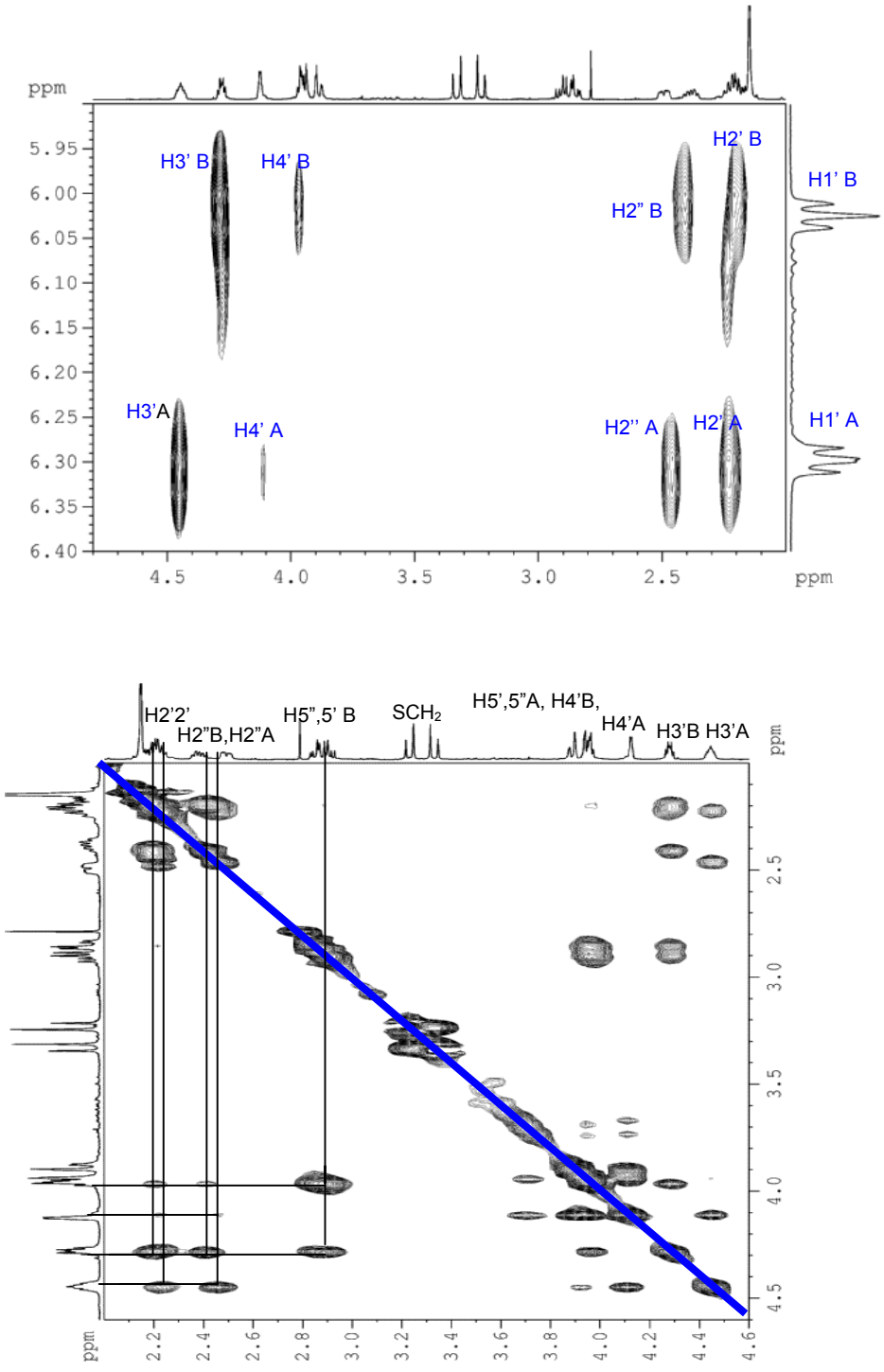


Figure10. Expanded TOCSY spectrum of **cst** dimer **10**

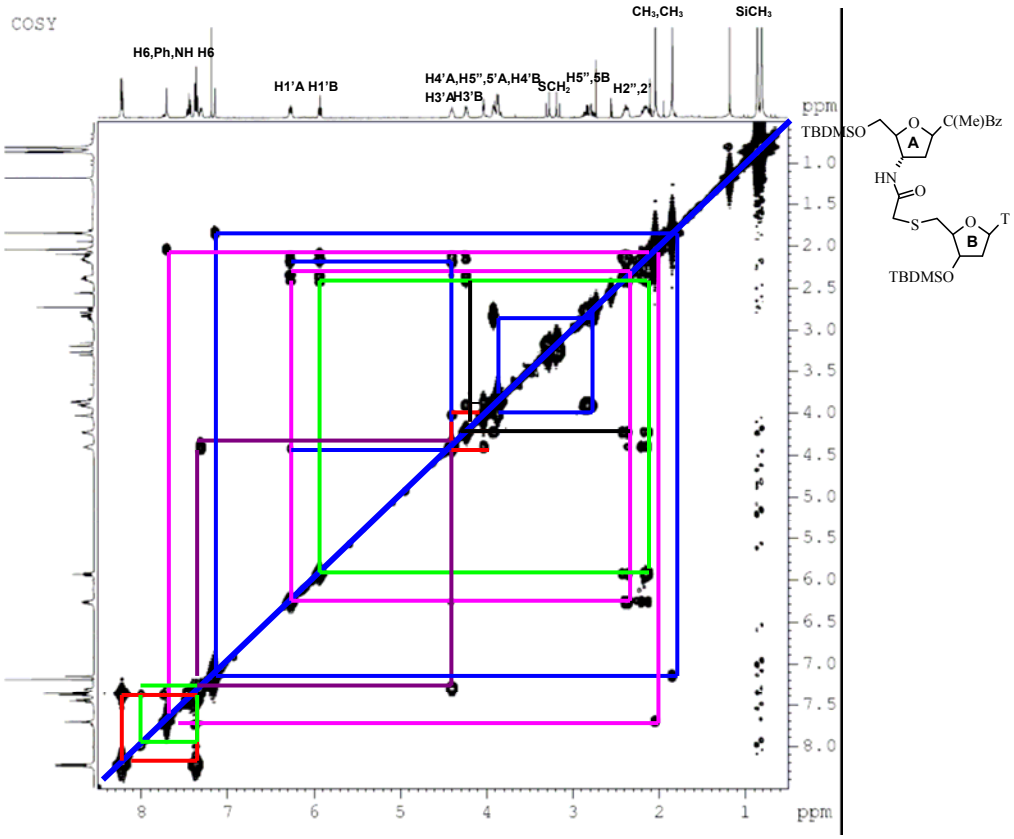


Figure 11. Fully assigned COSY spectrum of **cst** dimer **10**

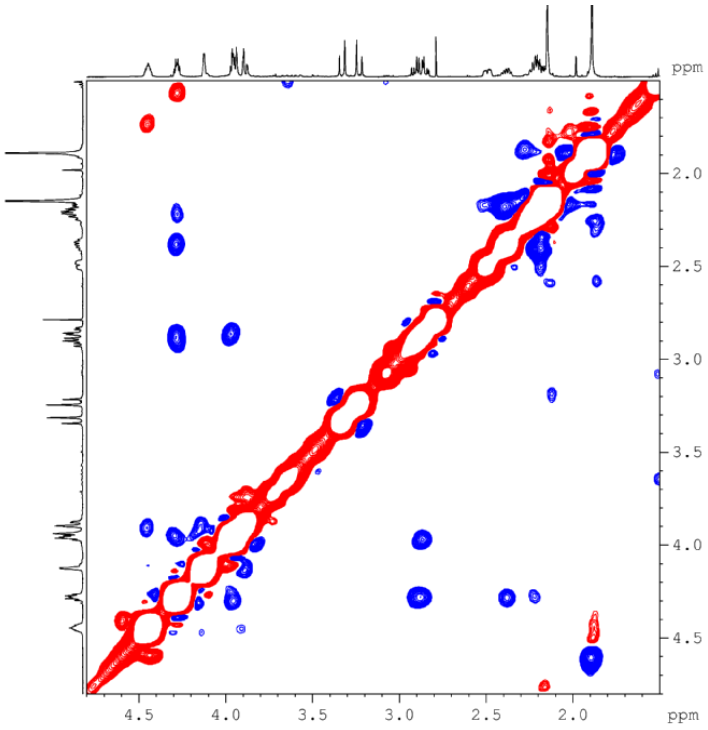
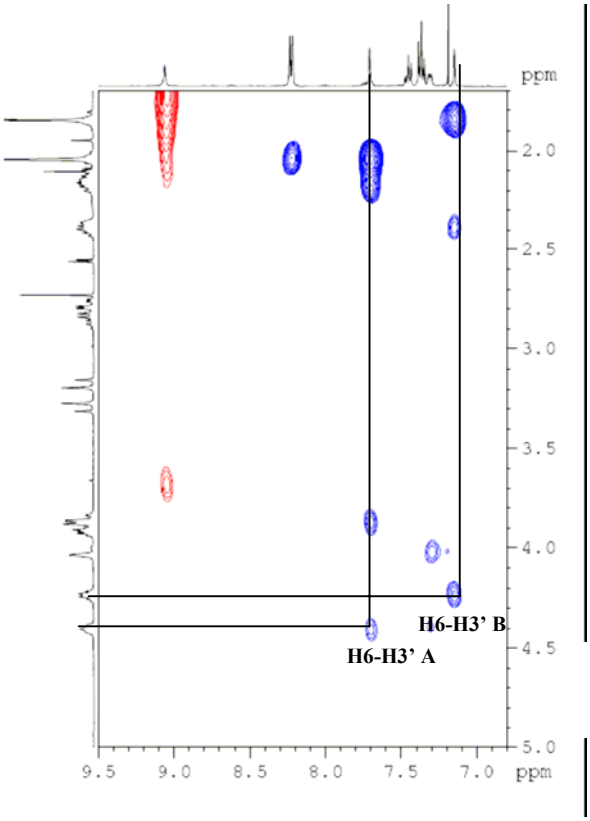


Figure 12. Expanded NOESY spectrum of **cst** dimer **10**



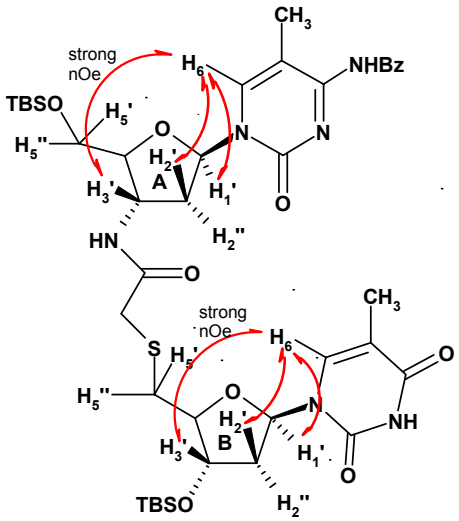


Figure 14. Schematic representation of the some of long range nOes observed for **cst** dimer **10**

Figure 13. Two-dimensional NOESY spectra of **cst** dimer **10**. The nOe cross peaks between H6-H3' indicate anti conformation. The F1 axis represents the H2'-H5" protons while the F2 axis represents H6-protons.

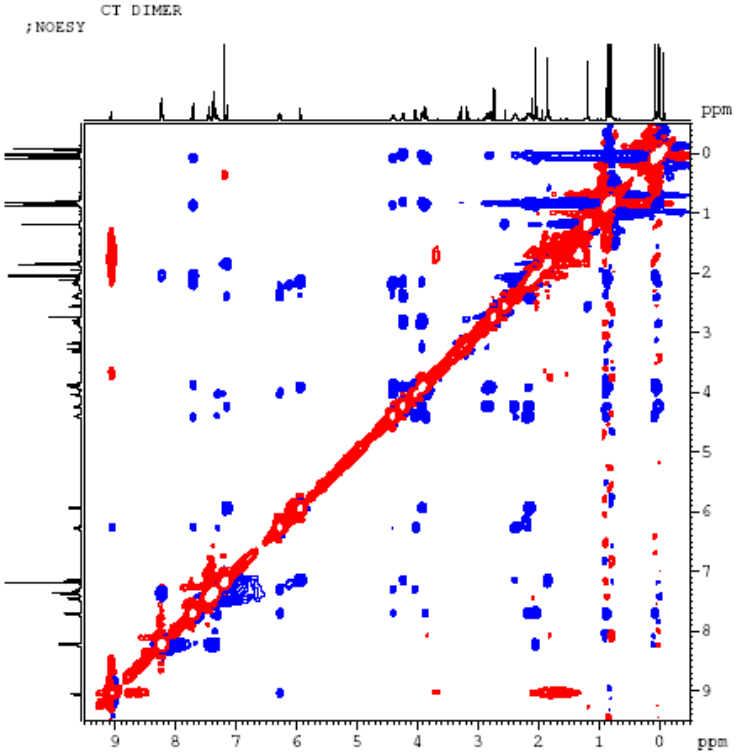


Figure 15. NOESY spectrum of **cst** dimer **10**

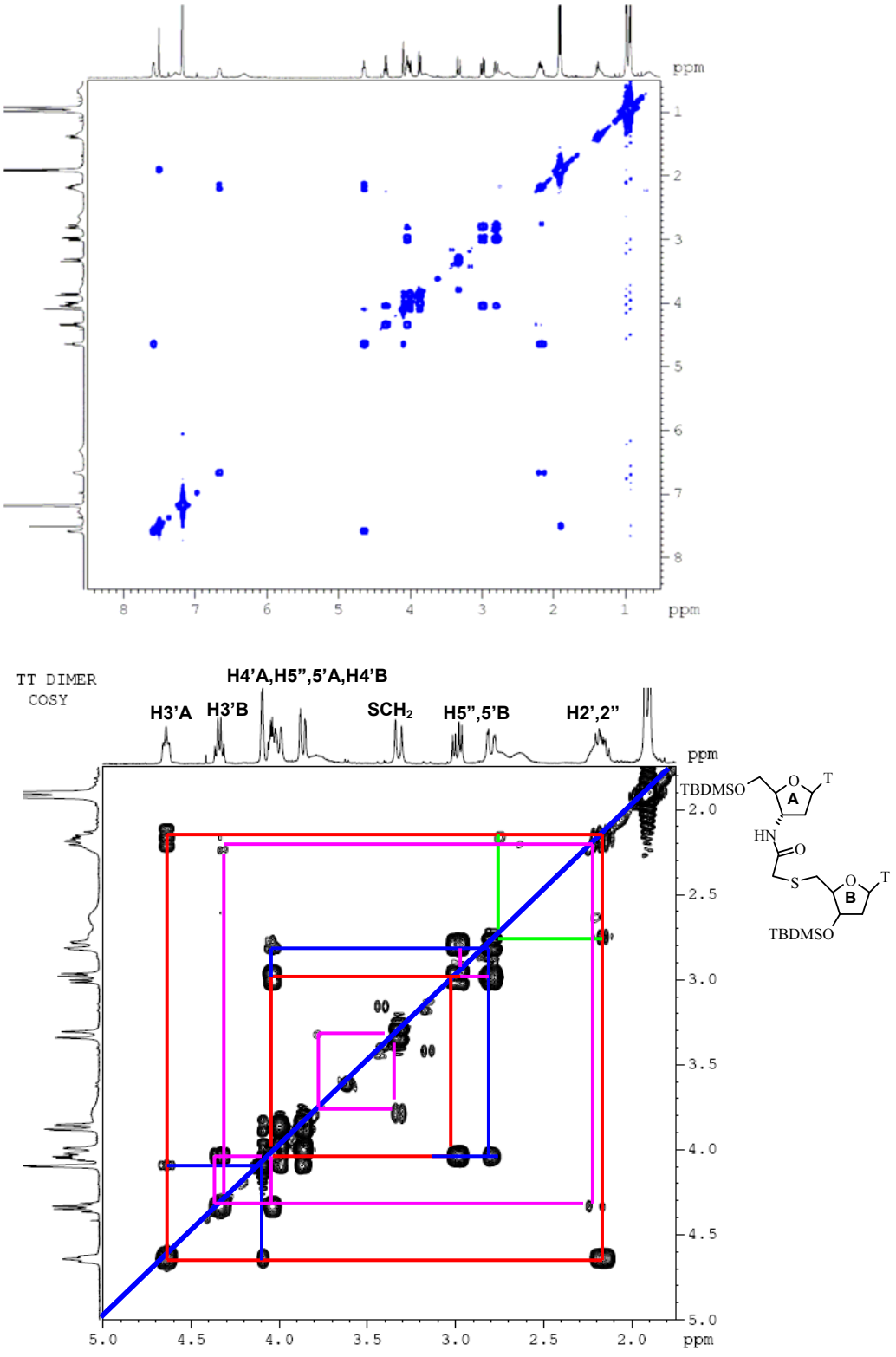


Figure 16. COSY spectrum and expanded assigned COSY spectrum of **tst** dimer **8**

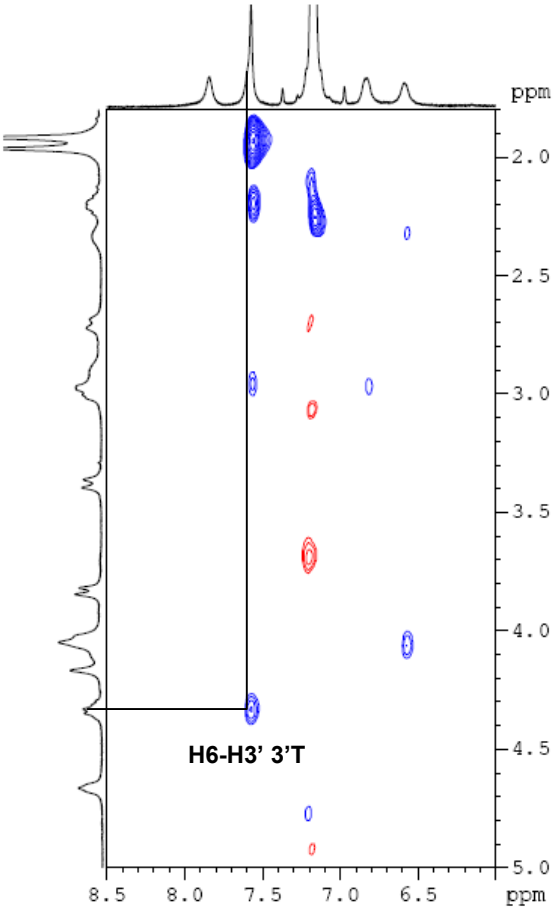


Figure 17. Two-dimensional NOESY spectra of **tst** dimer **8**. The nOe cross peaks between H6-H3' indicate anti conformation

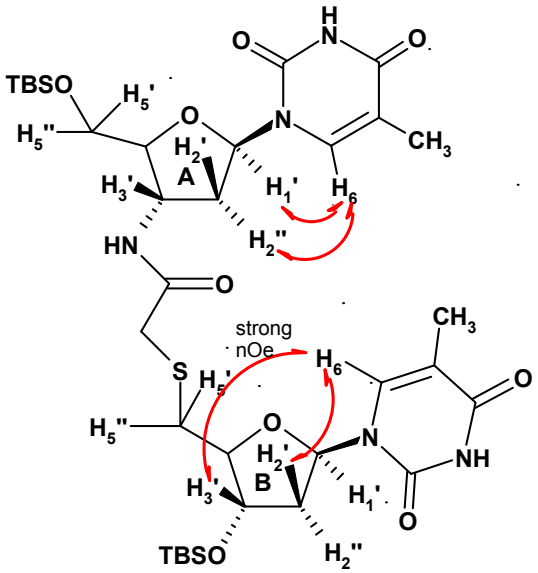


Figure 18. Schematic representation of some of the long -range nOes seen in **tst** dimer **8**

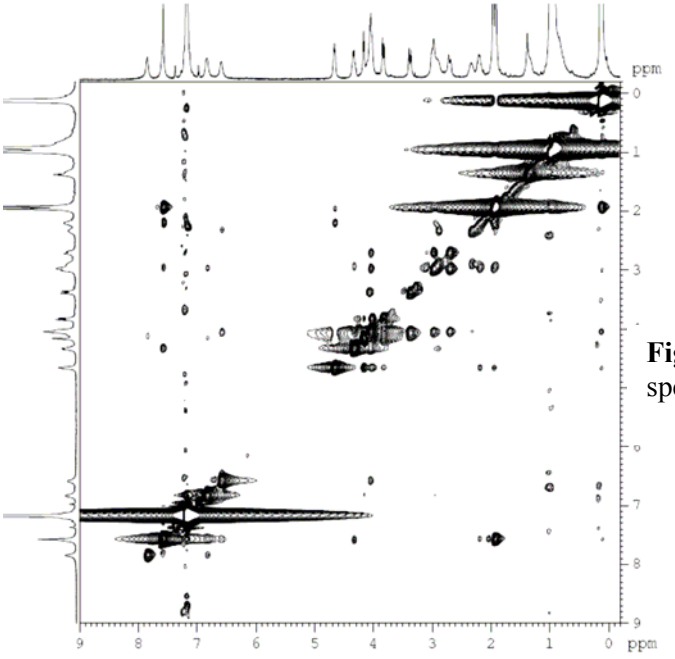
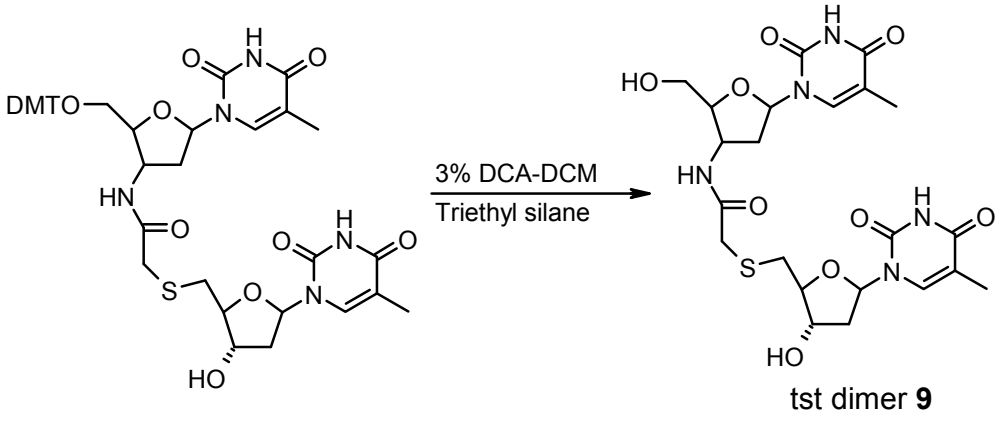


Figure 19. NOESY spectrum of **tst** dimer **8**.

In the ^{1}H NMR spectrum of **tst** dimer **8**, the peaks were not found to be well resolved. The solvent was changed from chloroform-d to benzene-d_6, but no significant improvement in the ^{1}H NMR spectrum was observed. The NMR experiment for the **tst** dimer **8** was performed at variable temperature and significant splitting of the peaks was observed when the spectrum was recorded at 343K. But the peaks due to the H1'-protons and H2', H2" were still unresolved.

All our attempts to improve the splitting of the peaks in the ^{1}H NMR spectrum of disilylated **tst** dimer **8** were failed and we were not able to calculate the coupling constants. Therefore the NMR experiments were performed with unprotected dinucleoside **9**. The 5', 3'-OH free dinucleotide was obtained by deprotection of 5'-ODMT group of 5'-*O*-DMT **tst**-3'-OH dimer (Synthesis is described in Chapter 2). The deprotection was done using 3% dichlroacetic acid in DCM using triethyl silane as the scavenger (Scheme 1). The free **tst** dinucleoside **9** formed was purified by successive trituration of the crude product with diethyl ether.



Scheme 1. Synthesis of 5', 3'-OH free **tst** dinucleoside dimer block

The ^{1}H NMR spectrum was recorded using D_2O as the solvent. All the peaks were found to be well resolved for the 5', 3'- OH free **tst** dinucleoside **9** (Figure 20).

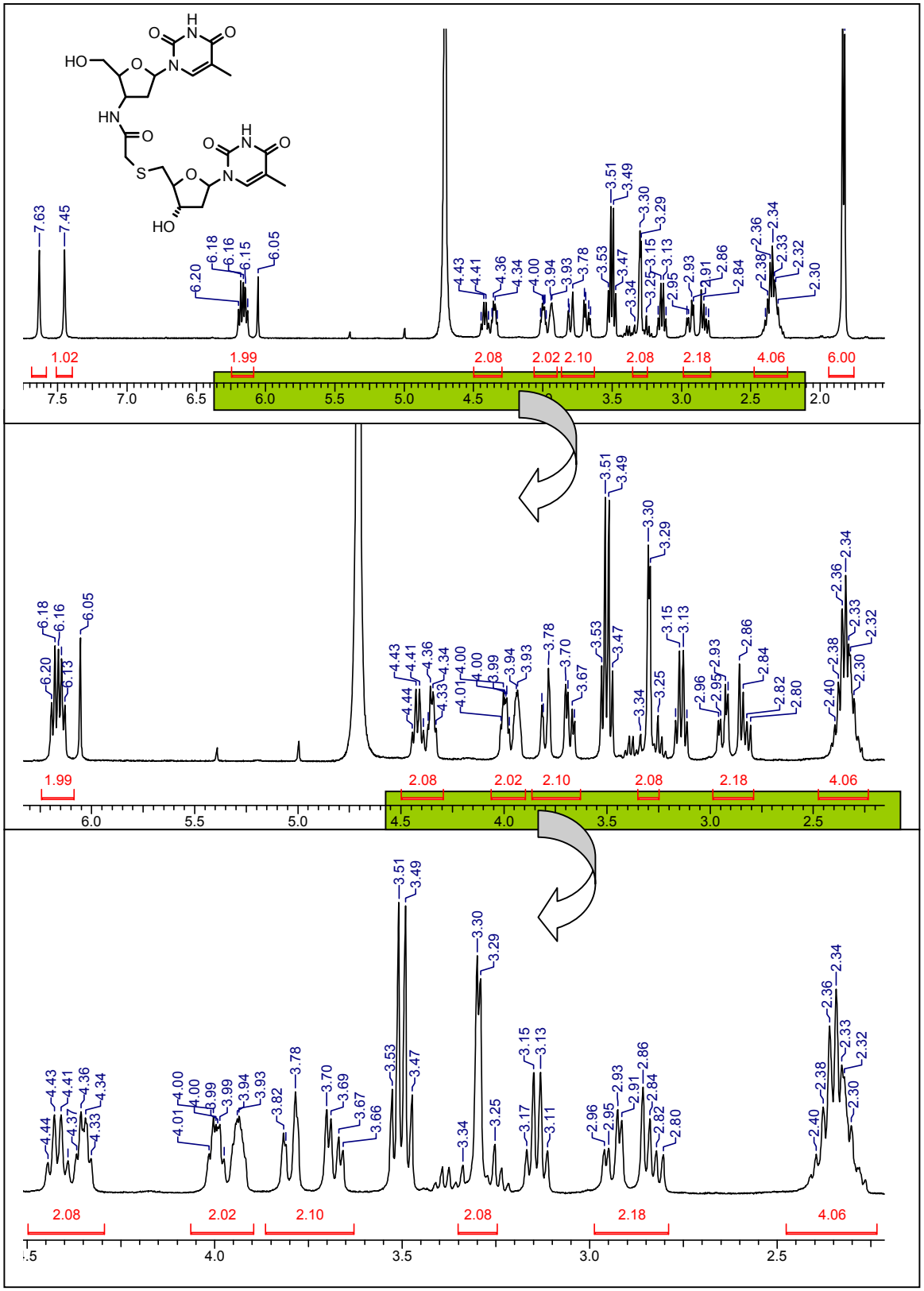
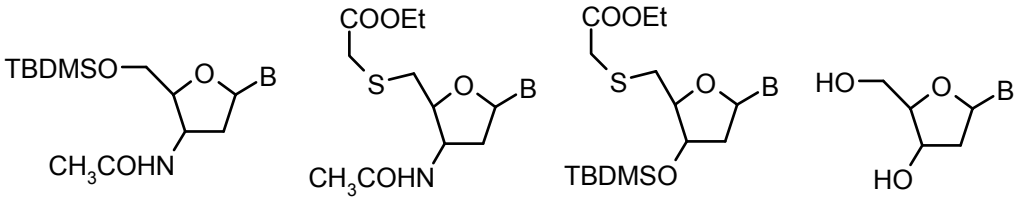


Figure 20. ^{1}H NMR of 5', 3'-OH free **tst** dinucleoside **9**

3.4 PSEUROT and Sum rule results

The validity of PSEUROT is reflected in the difference between J_{exp} and J_{calcd} (ΔJ, Hz), which should be in 0.0 ± 0.9 Hz and the least root mean square (rms) in range of 0.0-0.9. In this program, the mole fraction (MF) of the two probable conformers of ribose ring are denoted as MF1 (*N* conformer; P = 0º) and MF2 (*S* conformer, P = 180º), which are in equilibrium with the fractional ratio of conformers being 1:1.

The result of PSEUROT analysis as well as 'sum rule' for the TANA monomers are summarized in Table 8 and 9, and those for the dimers are summarized in Table 10.



1, B= T
2, B=C(Me)^{Bz}

3, B= T
4, B=C(Me)^{Bz}
5, B= C^{Bz}

6, B= T
7,B=C^{Bz}

11, B= T
12, B=C

Table 8. Summary of PSEUROT * output data of TANA monomers

Compound	P_N[a]	ϕ_N[b]	P_S[c]	ϕ_S[d]	MF1[e] (%N-type)	MF2[f] (%S-type)	ΔJ[g]	rms[h]	%S (Sum rule)
1	27.8	36.0	155.4	36.0	36	64	0.83	0.320	66
3	-42.1	35.0	102.0	36	31	69	0.39	0.301	68
6	-22.5	35.0	129.5	34	41	59	0.78	0.237	54
11[i]						66			66

[a] P_N pseudorotation angle for N-conformer, [b] ϕ_N Puckering amplitude for N conformer (P=0), [c]P_S pseudorotation angle for S-conformer, [d] ϕ_S Puckering amplitude for S conformer(P=180), [e]MF1 mole fraction of N conformer (P=0), [f]MF2 mole fraction of S conformer (P=180), [g] ΔJ sum of the difference between the five experimental $^3J_{HH}$ coupling constants and their calculated values, [h] rms Root mean square.

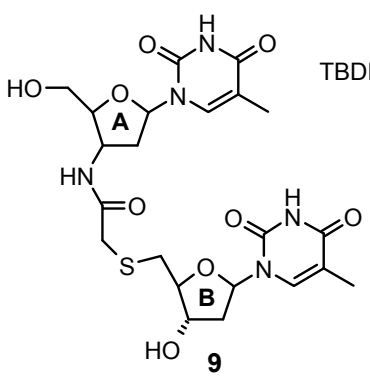
*The PSEUROT program was performed by keeping ϕ_N and ϕ_S constant while P_N, P_S and MF2 were redefined.[i] taken from the reference 10.

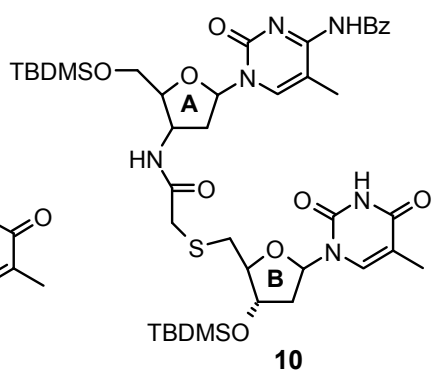
Table 9. Summary of PSEUROT output data of TANA monomers

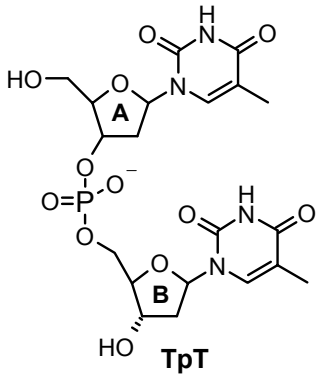
Compound	P_N	ϕ_N	P_S	ϕ_S	MF1 (%N-type)	MF2 (%S-type)	ΔJ	rms	%S (Sum rule)
2	15.2	36.0	121.9	36.0	27	73	-0.75	0.357	71
4	18.8	36.0	132.5	35.0	52	48	0.78	0.485	49
5	9.6	36.0	129.8	36.0	50	50	0.26	0.25	42
7	-0.6	36.0	140.1	36	51.5	48.5	0.47	0.28	36
12[i]						63			63

[i]taken from the reference 10.

Table 9. Summary of PSEUROT output data of TANA dimers







Compound	P_N	ϕ_N	P_S	ϕ_S	MF1 (%N-type)	MF2 (%S-type)	ΔJ	rms	%S (sum rule)
tst 9 A	28	36	121	36	45	55	0.14	0.36	64
tst 9 B	3.9	36	127	34	42	58	-0.17	0.288	46
cst 10 A	7.5	36	131	36	22	78	-0.22	0.196	71
cst 10 B	-17	33	182	35	42	58	0.47	0.288	60
TpT [i] **A**	-3	35	153	35	26	74	-	-	68
TpT [i] **B**	3	35	155	30	46	66	-	-	61

[i] taken from the reference 9.

3.5 Discussion

The proton NMR spectra of the compounds **1-7** (Figure 4) and **8-10** (Figure 9) were completely assigned by a combination of 1D ^{1}H-^{1}H decoupling and 2D-COSY ^{1}H NMR spectroscopic techniques (Table 4 for **1-7** and Table 6 for **8-10**) and all the vicinal coupling constants of the ribose ring were determined from combination of 1D ^{1}H NMR spectra and homonuclear ^{1}H decoupling experiments (Table 5 for **1-7** and Table 7 for **8-10**). These experimentally measured coupling constants were used as inputs into the PSEUROT program developed by Altona *et. al.*[5] to analyze the conformations adopted by 5-membered deoxy ribose rings. Also the % S and % N conformations were calculated by following the semiempirical sum rule by Altona *et. al.*[6] From this analysis, it is seen that significant decrease of % S-type character was observed for the TANA dimers and monomers as compared to that found in natural sugar-phosphate backbone. The results are tabulated in Table 8, 9 and 10.

3.5.1 Conformational studies of TANA monomers

Conformational analysis was done for the thyminyl, cytosinyl and 5-methyl cytosinyl TANA derivatives **3-7** and thyminyl and 5-methylcytosinyl 3'-acetamide derivatives **1** & **2** depicted in Figure 4.

The N-type ↔ S-type equilibrium of the sugar rings of the thyminyl TANA derivatives **3** and **6**, 5'-*O*-TBDMS 3'-NH acetyl thymidine **1** was calculated using both "sum rule" and PSEUROT program (Table 8) and the results were compared with the reported data for the natural deoxy nucleoside thymidine **11** (Table 8).[10] This gives the effect of 3'-NH substitution on the sugar ring, the effect of 5'-S-CH_2COOEt substituents and the effect of 3'-& 5'-modifications respectively.

The results showed that S-type conformation of the sugar rings in the compounds **1** (64% PSEUROT, 66% sum rule) and **3** (69% PSEUROT, 68% sum rule) is almost similar to their natural counterpart thymidine **11** (66%) (Table 8). A significant decrease of S-type conformation was observed for the compound **3** (59% PSEUROT, 54 % sum rule) (Table 8), when 3'-substituent is O-TBDMS. The shift towards N-type conformation could be result of 5'-S-CH_2COOEt group, which is not known in the literature.

The N-type ↔ S-type equilibrium of the sugar rings of the TANA derivatives **4**, **5** and **7**, and N4-benzoyl 5'-*O*-TBDMS 3'-NH acetyl 5-methyl 2', 3'-dideoxycytidine **2** was calculated using both "sum rule" and PSEUROT program (Table 9). The results were compared with the reported data for the natural deoxy nucleoside cytidine **12** (Table 9).[10] The results show a shift toward N-type conformation for the protected cytosinyl TANA derivatives **5** and **7** and 5-methylcytosinyl TANA derivative **4**. In contrast, comparison of **12** and **2**, suggests the equilibrium shift toward S-type upon N4-benzoyl protection of base.This shift in conformation towards N-type could be the result of 5'-S-CH_2COOEt group, which is not known in the literature.

3.5.2 Conformational studies of TANA dimers

The 1D ^{1}H NMR spectra of **tst 9** and **cst 10** were fully assigned with the aid of 2D COSY, TOCSY and NOESY experiments (Table 6). The coupling constants of tst **9** and cst **10** were determined from 1D spectra and from homonuclear proton decoupled spectra where spectral overlap was severe (Table 7).

Before going for PSEUROT program we have determined the % S-type conformation of the sugar rings of **tst** and **cst** dimer by the 'Sum rule'.[6] A significant decrease of % S-type conformation of the 3' end (B-ring) of **tst 9** dimer (61% to 46%, Table 10) was observed from the natural **TpT** dimer (68 % for A ring and 61% for B ring) by "sum rule". To verify these results, we have performed the PSEUROT program. The results obtained are comparable to the results as obtained by sum rule and show similar trend of decrease of %S-type conformation. There is significant decrease of % S-type conformation of the B-sugar ring of both **tst** and **cst** dimer (58% and 58%, Table 10) in comparison to the reported **TpT** dimer (66% for B-ring).[9] Similar trend of decrease of %S-type conformation of 5' end (A-ring) of **tst** dimer (55%) was observed (for A-ring of TpT dimer %S-type conformation is 74%).[9] The results obtained for the 5' end sugar ring (A sugar ring) of **cst** dimer (78% S-type conformation) is not in the similar trend as the other sugar rings. It might be the presence of exocyclic benzoyl protecting group, so for comparative studies, it will be required to do the PSEUROT analysis for the fully deprotected **cst** dimer.

Using the sum rule as well as PSEUROT program, we established that replacement of the phosphodiester linkage with mercaptoacetamide linkage results in a significant shift in the N-S conformational equilibrium from a predominantly S to N in 3'-terminal ribose sugar. NOESY experiments also support these observations. Both 3'-terminal nucleosides of **tst** and **cst** show strong H6 to H3' cross-peaks, characteristic of a nucleoside which has a predominantly N ribose conformer, with the base *anti* relative to its ribose ring (Figures 13 and 17)[9] whereas the 5'-terminal ribose ring of **tst** dimer such nOe interactions were not observed.

3.6 Conclusion

The TANA monomers and dimers were synthesized and protected appropriately for NMR conformational studies. DQF-COSY, COSY, TOCSY and NOESY experiments were perfomed to assign the proton resonances.The coupling constants were determined from the 1D spectrum and by homonuclear decoupling experiments.The conformational analysis of the pentose sugar rings of the monomers and dimers were determined by PSEUROT program version 5.4.1. The studies show that ONs having the mercaptoactamido group which is neutral might adopt a conformation in between a conformation of DNA and RNA that contribute to the thermodynamic stability of the resulting heteroduplexes formed with complementary RNA, by preorganizing the ONs into the required conformation for a preferred A-type duplex.

Summary

- The TANA monomers and dimers were synthesized and protected appropriately for NMR conformational studies.
- DQF-COSY, COSY, TOCSY and NOESY experiments were performed to assign the proton resonances.
- The coupling constants were determined from the 1D spectrum.
- The conformational analysis of the pentose sugar rings of the monomers and dimers were determined by PSEUROT program version 5.4.1 and “sum rule”.
- Introduction of mercapto functionality at 5’ position increases the % N-type conformation of the ribose sugar ring.

3.7 References

1. (a) Saenger, W. (1984). Principles of nucleic acid structure. Springer-Verlag, New York. (b) Lescrinier, E., Froeyen, M. and Herdewijn, P.(2003) *Nucleic Acid. Res*., **31**, 2975-2989.
2. (a) Sundaralingam, M., (1969) *Biopolymers,* **7**, 821 and references therein. (b) Altona, C. and Sundaralmgam, M. (1972) *J. Am. Chem.Soc*., **94**, 8205.
3. de Leeuw, H. P. M. , Haasnoot, C.A. G. and Altona, C. (1980) *Isr. J. Chem.* ,**20**,108.
4. Altona, C. and Sundaralmgam, M. (1973) *J. Am. Chem. Soc.*, **95**, 2333.
5. (a) van Wijk, J., Haasnoot, C. A. G., de Leeuw, F. A. A. M., Huckriede, B. D., Westra Hoekzema, A. and Altona, C. *PSEUROT 6.2 1993*, *PSEUROT 6.3 1999*; Leiden Institute of Chemistry, Leiden University. (b) de Leeuw, F. A. A. M. and Altona, C. J. (1983) *Comput. Chem.*, **4**, 428. (c) Altona, C. (1982) *Recl. Trav. Chem. Pays-Bas*, **101**, 413.
6. Rinkel, L. J. and Altona, C. (1987) *J. Biomol. Struct. Dyn.,***4***,* 621
7. (a) Herdewijn, P. (1996) *Liebigs Ann.,* 1337. (b) Marquez, V. E., Siddiqui, M. A., Ezzitouni, A., Russ, P., Wang, J., Wagner, R. W. and Matteucci, M. D.(1996) *J. Med. Chem*., **39**, 3739.
8. (a) Plavec, J., Koole, L.H., Sandstrom, A. and Chattopadhyaya, J. (1991) *Terahedron*, **47**, 7363. (b) Cheng, D. M. and Sarma, R.H. (1977) *J.Am.Chem.Soc.*, **99**, 7333.
9. (a) Glemarec, C., Reynolds, R. C., Crooks, P. A., Maddry, J. A., Akhtar, M. S., Montgomery, J. A., Secrist III, J. A. and Chattopadhyaya, J.(1993) *Tetrahedron,* **49**, 2287. (b) Plavec, J., Koole, L.H., Sandstrom, A. and Chattopadhyaya, J. (1991) *Terahedron*, **47**, 7363. (b) Glemarec, C., Nyllas, A., Sund, C. and Chattopadhyaya, J. (1990) *J Biochem Btophys Methods*, **12**, 311.
10. Plavec, J., Thibaudeau, C., Viswanadham, G., Sund, C. and Chattopadhyaya, J. (1994) *J. Chem..Soc., Chem. Commun.*, 718-783.
11. Thibaudeau, C., Plavec, J., Garg, N., Papchikhin, A.; Chattopadhyaya, J. (1994*) J. Am. Chem. Soc.* **116**, 4038-4043.
12. (a) Ding, D., Gryaznov, S.M., Lloyd, D.H., Chandrasekaran, S., Yao, S., Ratmeyer, L., Pan, Y.and Wilson, W.D. (1996) *Nucleic Acids Res.,* **24**, 354-360. (b) Ding, D., Gryaznov, S. M. and Wilson, W. D. (1998) *Biochemistry*, **37**, 12083-12093. (c) Gryaznov, S. and Chen, J.-K. (1994) *J. Am. Chem. Soc*., **116**, 3143-3144.
13. Remin, M. and Shuger, D. (1972) *Biochem. Biophys. Res. Commun.,* **48**, 636.

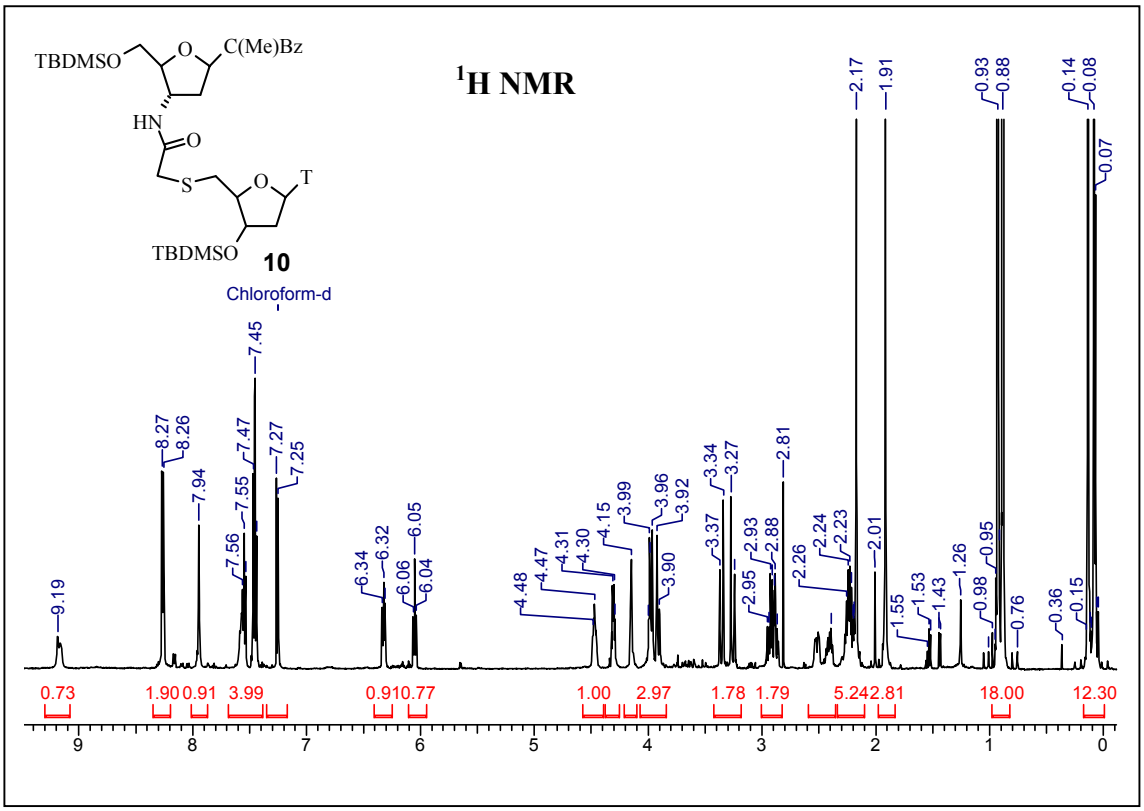
3.8 Appendix

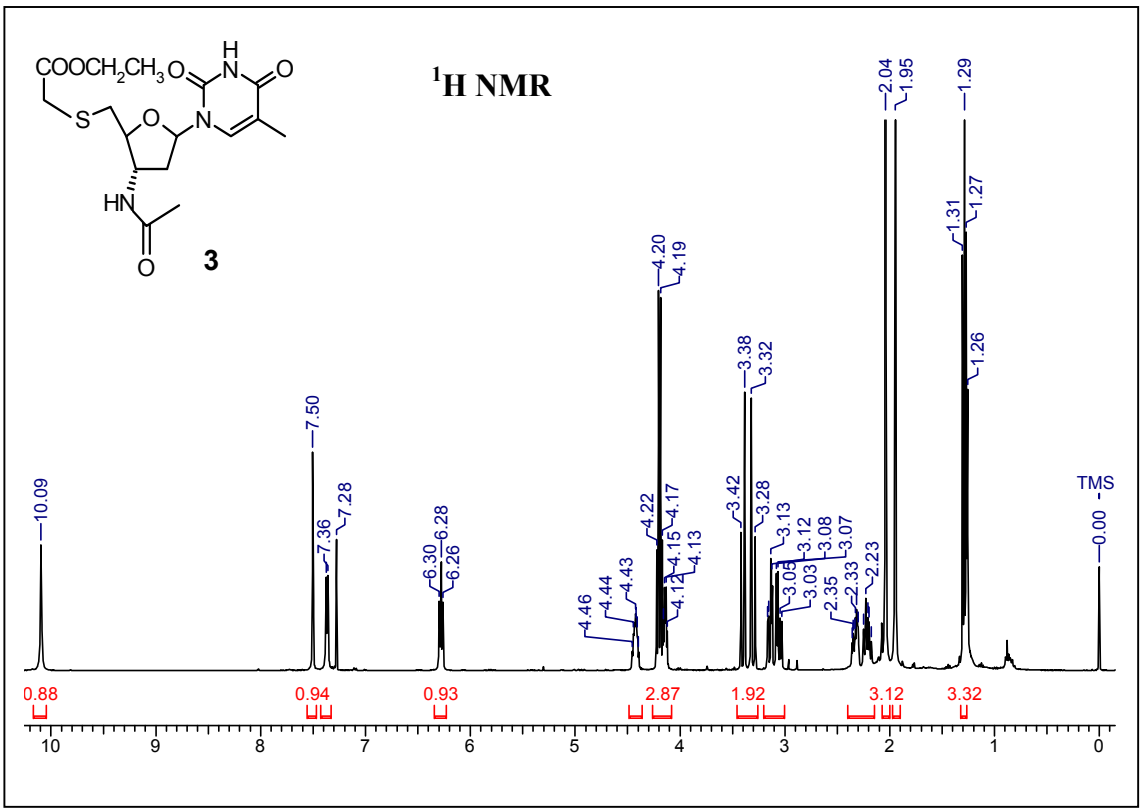
3.8.1 Typical format of PSEUROT

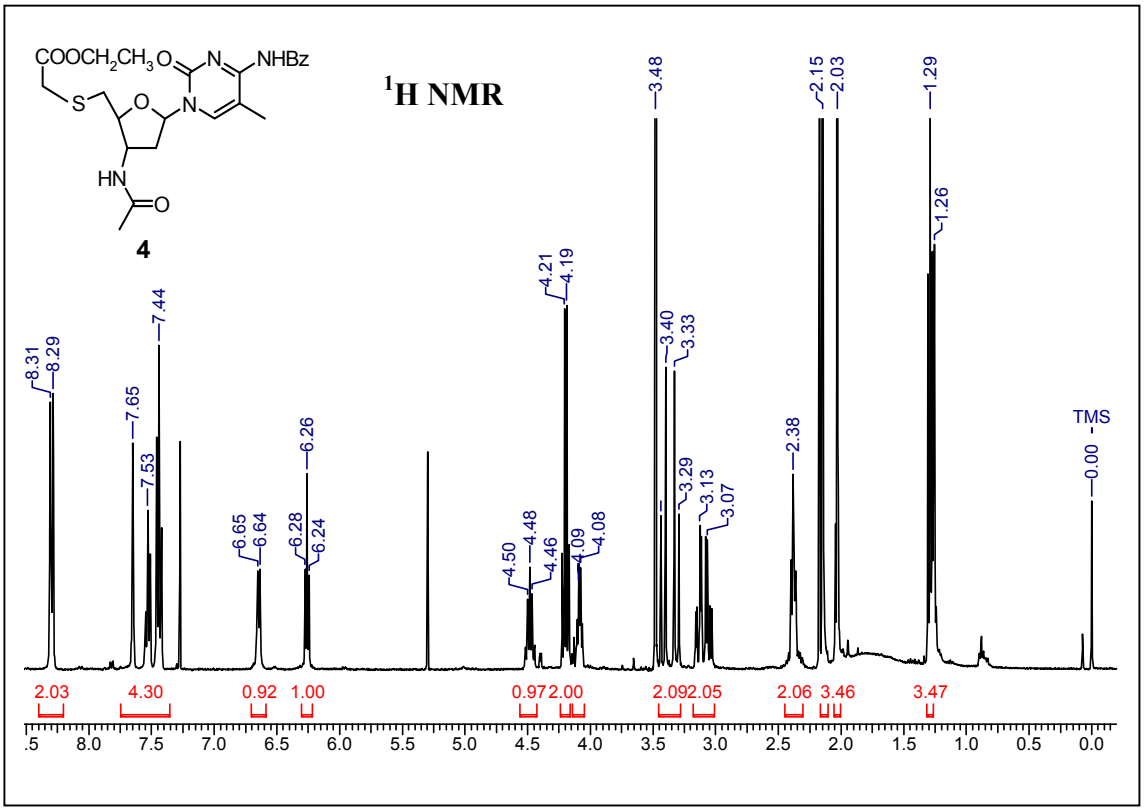
```
 C:\PSEUROT>TYPE PS54.OUT   | MORE
 1CASE NR:    1
 TITLE  :KG9
                  PSEUROT-5.4
P A R A M E T E R S   I N   P S E U D O R O T A T I O N   R E L A T I O N S

0COUPLING CONSTANTS DEFINED BY:

 NAME        FASE       A         B        ELECTRONEGATIVITIES      # OF SUBSTITUENTS

 1'-2'      -144.0      1.030     120.0     .560     1.420     .000      .620      3
 1'-2"      -144.0      1.020      .0       .560     1.420     .620      .000      3
 2'-3'       .0         1.060      .0       .620     .000      .620      .620      3
 2"-3'       0          1.060     120.0     .000     .620      .860      .620      3
 3'-4'      144.0       1.090    -120.0     .620     .860      1.420     .680      4


0PARAMETERS USED IN GENERALIZED KARPLUS EQUATION:

          CONST       SUM       SUMSQU     12  34     13  24     14  23

          7.0100     -.5800      .0000     -.2400      .0000      .0000    CONSTANT
         -1.0800      .0000      .0000      .0000      .0000      .0000    COSPHI
          6.5400     -.8200      .0000      .0000      .0000      .2000     COS2
          -.4900      .0000      .0000      .0000      .0000      .0000     COS3
           .0000      . 0000     .0000      .0000      .0000      .0000     COS4
           .0000      .0000      .0000      .0000      .0000      .0000     COS5

                    .0000      .0000              .0000      SINPHI
                    .0000      .6800              .0000       SIN2
                    .0000      .0000              .0000       SIN3
                    .0000      .0000              .0000       SIN4
                    .0000      .0000              .0000       SIN5



I N P U T   D A T A
     J -->   1'-2'     1'-2"     2'-3'     2"-3'     3'-4'
 25C?          6.10      6.20      7.10      5.80      6.10
0CASE     1 HAS   1 SETS ->   5 OBSERVATIONS 0MAX STEPSIZE              =    .100
 CONVERGENCE CRITERIUM        =   5.000E-04
 MAXIMUM NUMBER OF ITERATIONS =  25
1CASE NR:    1
 TITLE  :KG9
                  PSEUROT-5.4
F I R S T   E S T I M A T E S

 CONFORMER 1:       P =     .0 DEGREES =    .000 RAD  ---> TO BE REFINED
              PHI =   36.0 DEGREES =    .628 RAD  ---> FIXED PARAMETER

 CONFORMER 2:       P = 180.0 DEGREES =   3.142 RAD  ---> TO BE REFINED
              PHI =   36.0 DEGREES =    .628 RAD  ---> FIXED PARAMETER

                 25C?     : MF2 =    .000      ---> TO BE REFINED
0***** ITERATION CONVERGED *****
1CASE NR:    1
 TITLE: KG9
                  PSEUROT-5.4
```

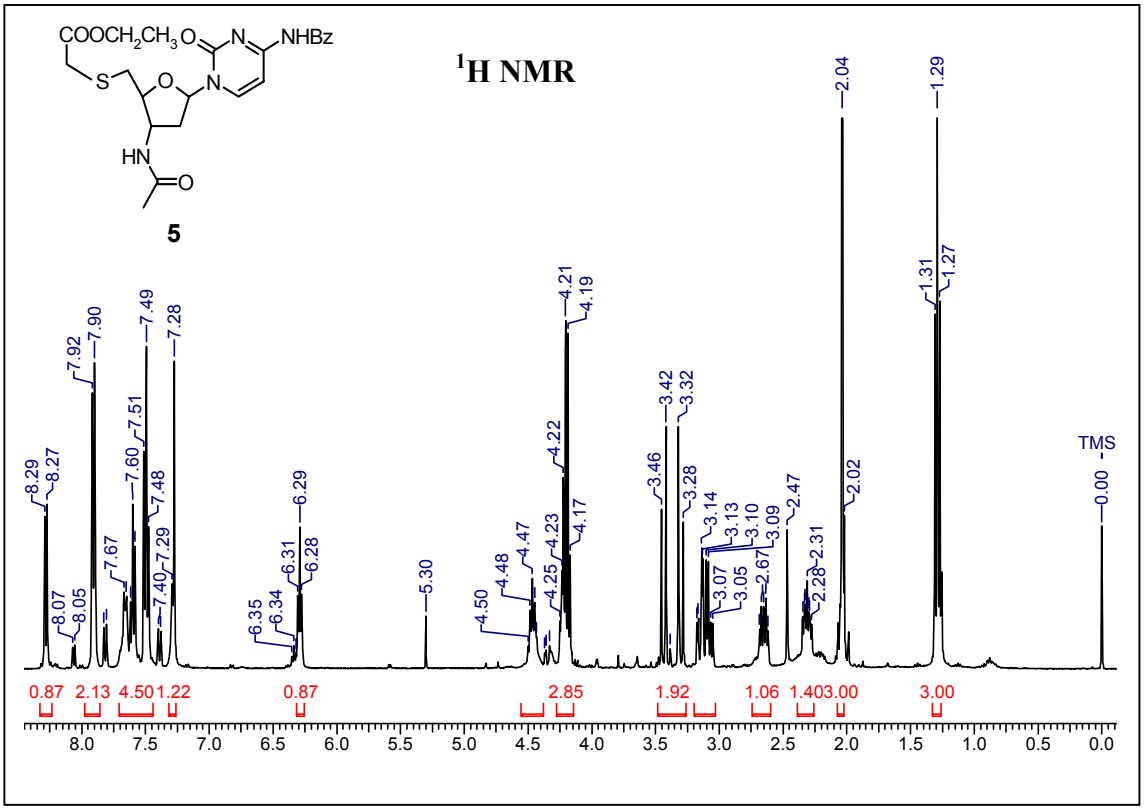
```
F I N A L   O U T P U T

0          1'-2'              1'-2"              2'-3'      2"-3'              3'-4'
         JEXP   JCAL   DIFF   JEXP   JCAL   DIFF   JEXP   JCAL   DIFF   JEXP   JCAL   DIFF   JEXP   JCAL   DIFF
 25C?    ! 6.10   5.79    .31 ! 6.20   6.49   -.29 ! 7.10   7.30   -.20 ! 5.80   5.56    .24 ! 6.10   5.90    .20  !
0          MF1      MF2      RMS
 25C?       .498     .502     .250
0CONFORMER 1:                          CONFORMER 2:
       P =   9.6   DEGREES   =   .167 RAD          P =  129.8   DEGREES =  2.266 RAD
     PHI =  36.0   DEGREES   =   .628 RAD        PHI =  36.0   DEGREES =   .628 RAD
 PHI1'-2' =  94.0 ---> J1'-2' =  1.05        PHI1'-2' =  156.0 ---> J1'-2' = 10.49
 PHI1'-2" =  -25.7 ---> J1'-2" =  7.79        PHI1'-2" =  35.6 ---> J1'-2" =  5.20
 PHI2'-3' =  37.6 ---> J2'-3' =  6.72        PHI2'-3' =  -24.4 ---> J2'-3' =  7.87
 PHI2"-3' =  157.6 ---> J2"-3' = 10.37        PHI2"-3' =  95.6 ---> J2"-3' =  .80
 PHI3'-4' = -155.1 ---> J3'-4' =  8.85        PHI3'-4' = -117.4 ---> J3'-4' =  2.98
1CASE NR:   1
 TITLE  :KG9
                PSEUROT-5.4
E R R O R   A N A L Y S I S

0OVERALL RMS  =   .250E+00
0STANDARD DEVIATIONS IN PARAMETERS:

    0     .201     .134     .030
0CORRELATION MATRIX OF PARAMETERS
0PAR.      1        2        3
 1     1.000
 2      .464     1.000
 3     -.364     -.555     1.000
0*** END OF PGM PSEUROT ***
```

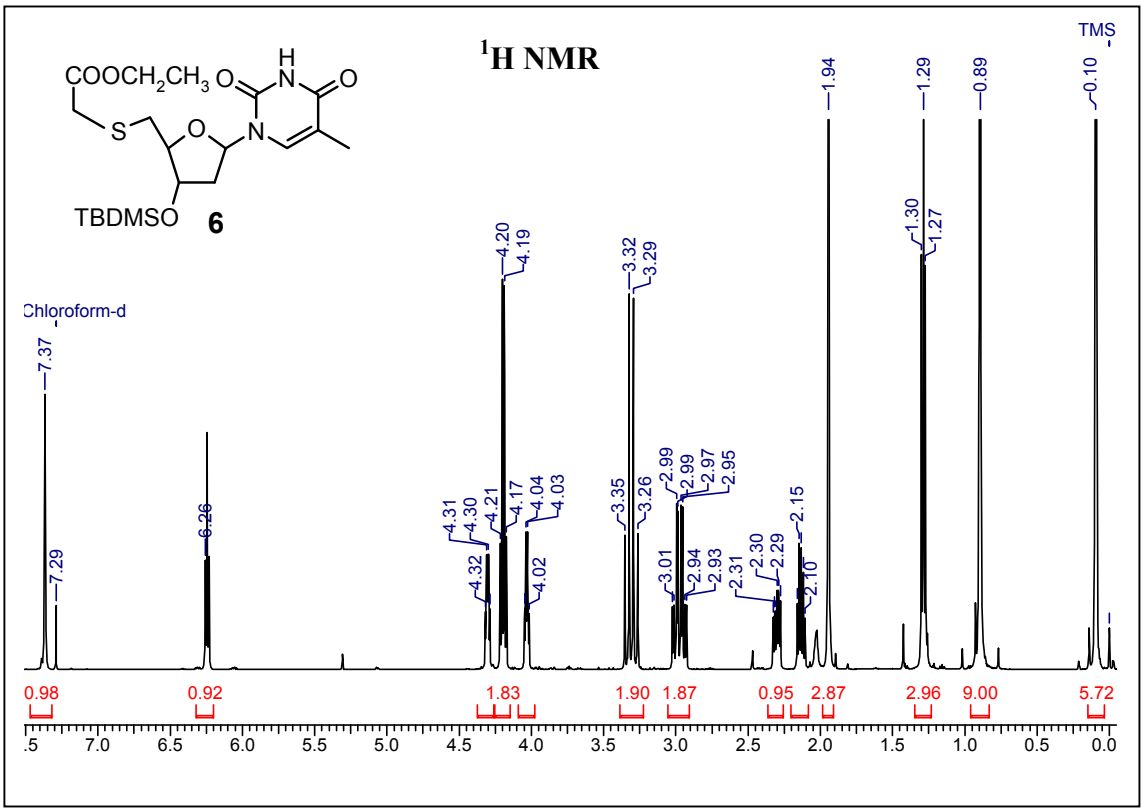
3.8.2 ^{1}H NMR of compounds 1- 10

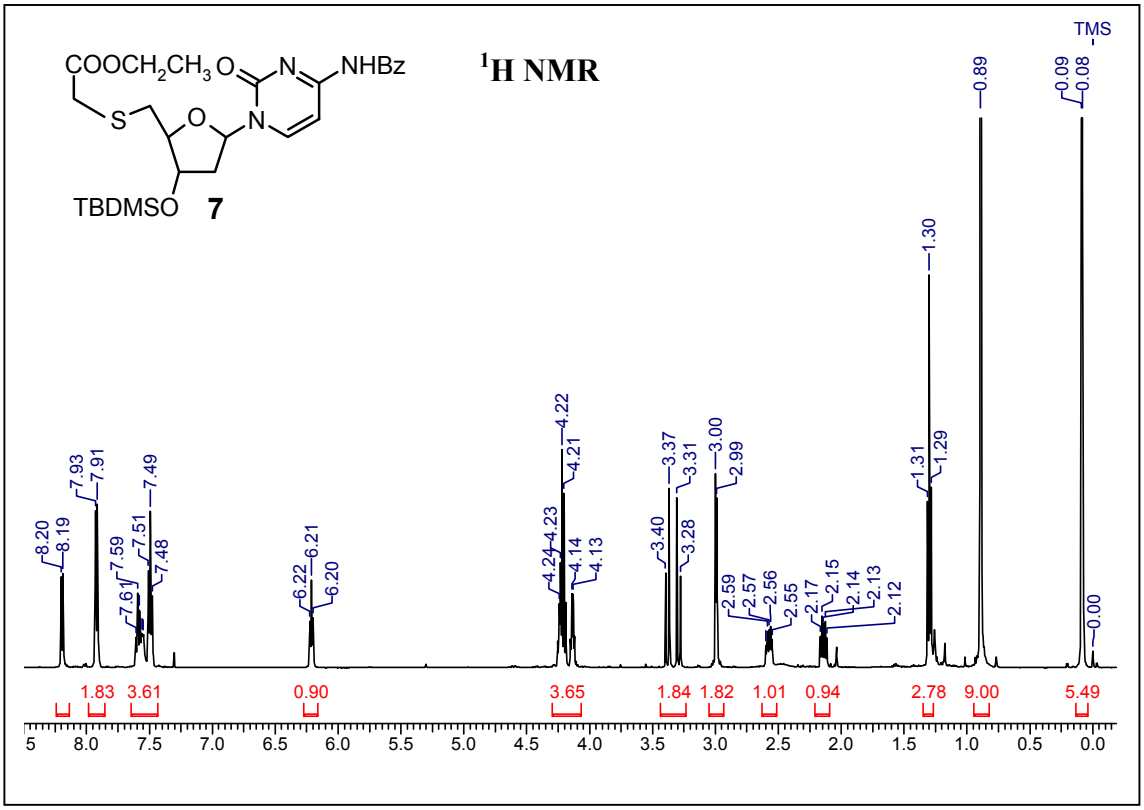


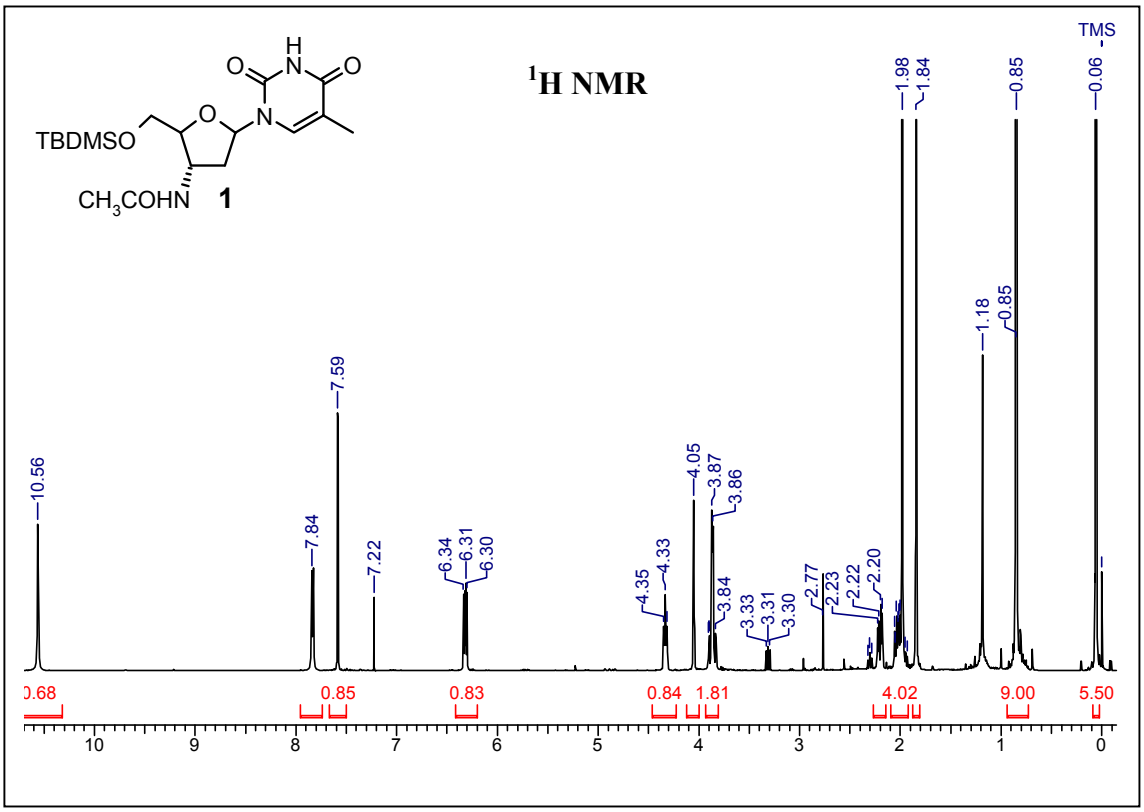

COOCH2CH3
O
H
N
O
S
O
N
HN
O
3
1H NMR
10.09
7.50
7.36
7.28
6.30
6.28
6.26
4.46
4.44
4.43
4.22
4.20
4.19
4.17
4.15
4.13
4.12
3.42
3.38
3.32
3.28
3.13
3.12
3.08
3.07
3.03
3.03
2.35
2.33
2.23
2.04
1.95
1.31
1.29
1.27
1.26
TMS
0.00
0.88
0.94
0.93
2.87
1.92
3.12
3.32
10
9
8
7
6
5
4
3
2
1
0

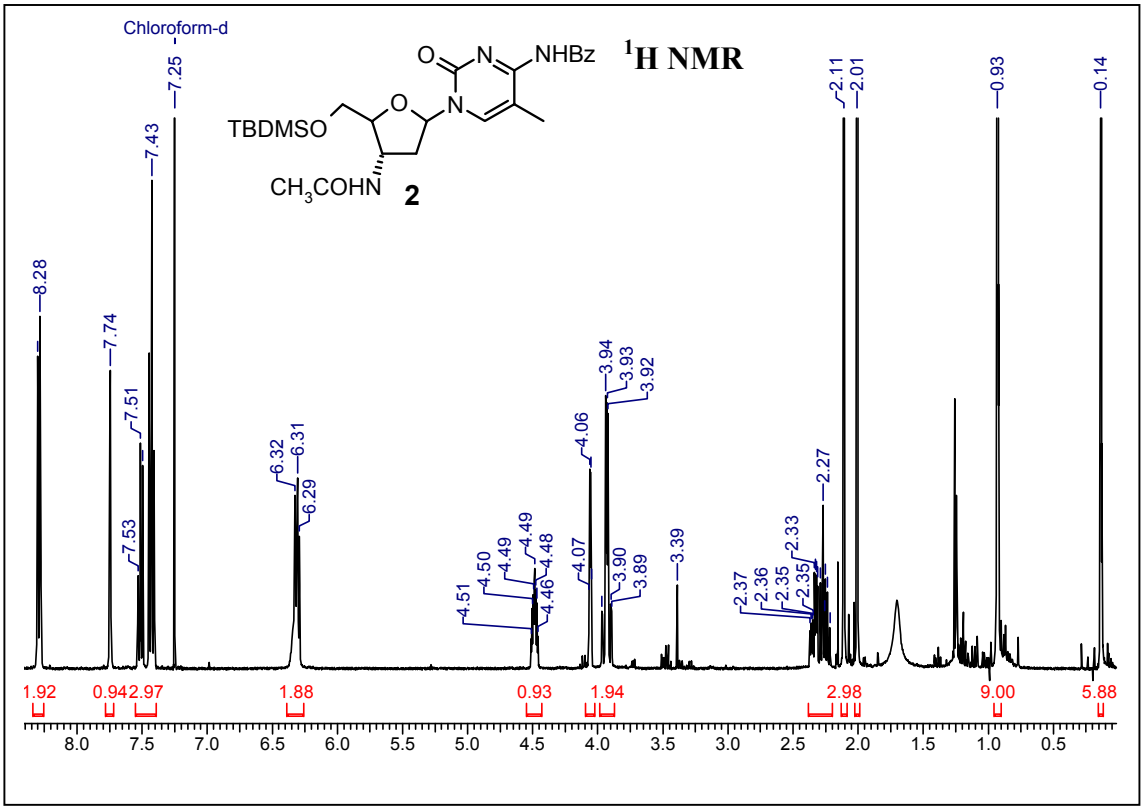

COOCH2CH3
O
N
NHBz
S
O
N
HN
O
4
1H NMR
8.31
8.29
7.65
7.53
7.44
6.65
6.64
6.28
6.26
6.24
4.50
4.48
4.46
4.21
4.19
4.09
4.08
3.48
3.40
3.33
3.29
3.13
3.07
2.38
2.15
2.03
1.29
1.26
TMS
0.00
2.03
4.30
0.92
1.00
0.97
2.00
2.09
2.05
2.06
3.46
3.47
.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

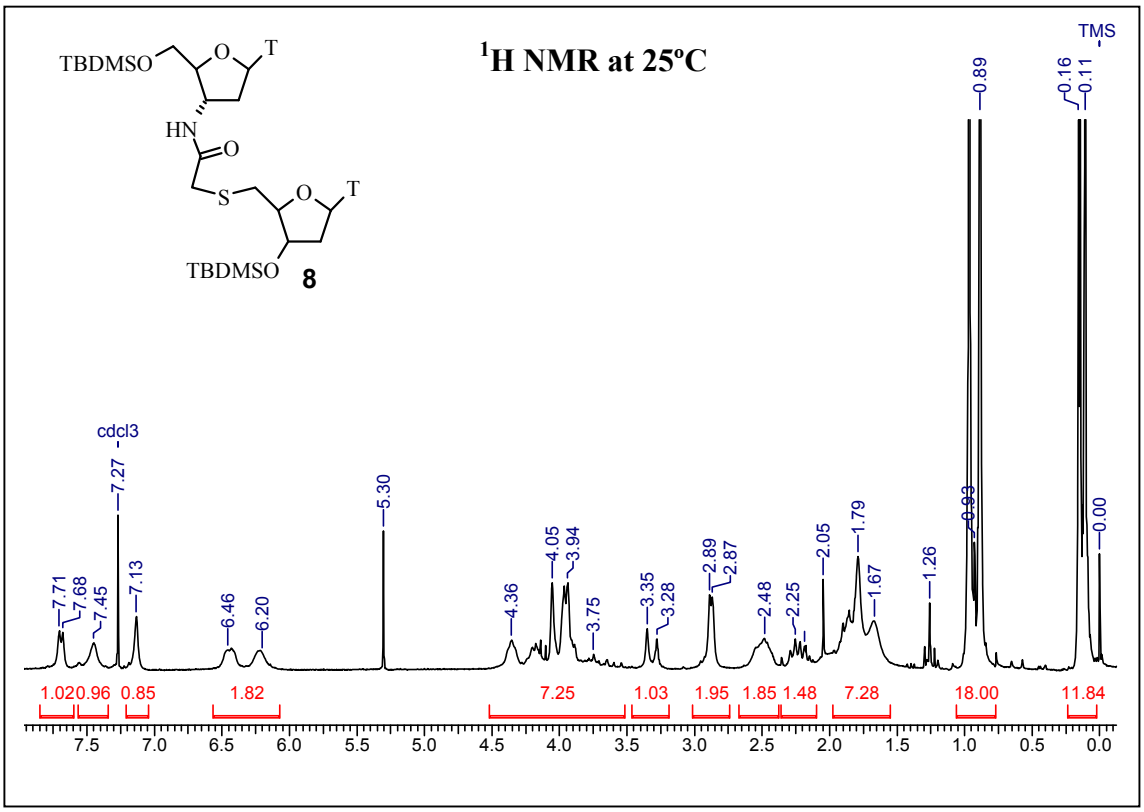

COOCH2CH3
O
N
NHBz
S
O
N
HN
O
5
1H NMR
TMS
0.87 2.13 4.50 1.22 0.87 2.85 1.92 1.06 1.40 3.00 3.00


COOCH2CH3
O
H
N
O
S
O
N
TBDMSO
6
1H NMR
TMS
Chloroform-d
0.98 0.92 1.83 1.90 1.87 0.95 2.87 2.96 9.00 5.72


COOCH2CH3
O
N
NHBz
S
O
N
TBDMSO
7
1H NMR
TMS
0.89
0.09
0.08
1.30
1.31
1.29
4.22
4.21
4.23
4.24
4.14
4.13
3.37
3.31
3.40
3.28
3.00
2.99
2.59
2.57
2.56
2.55
2.17
2.15
2.14
2.13
2.12
8.20
8.19
7.93
7.91
7.59
7.61
7.51
7.49
7.48
6.22
6.21
6.20
0.00
1.83
3.61
0.90
3.65
1.84
1.82
1.01
0.94
2.78
9.00
5.49


H
N
O
O
N
O
TBDMSO
CH3COHN
1
1H NMR
TMS
1.98
1.84
0.85
0.06
1.18
0.85
10.56
7.59
7.84
7.22
6.34
6.31
6.30
4.05
3.87
3.86
4.35
4.33
3.84
3.33
3.31
3.30
2.77
2.23
2.22
2.20
0.68
0.85
0.83
0.84
1.81
4.02
9.00
5.50


Chloroform-d
^{1}H NMR
TBDMSO
NHBz
CH$_3$COHN
2
8.28
7.74
7.53
7.51
7.43
7.25
6.32
6.31
6.29
4.51
4.50
4.49
4.49
4.48
4.46
4.07
4.06
3.94
3.93
3.92
3.90
3.89
3.39
2.37
2.36
2.35
2.33
2.27
2.11
2.01
0.93
0.14
1.92
0.94
2.97
1.88
0.93
1.94
2.98
9.00
5.88


^{1}H NMR at 25°C
TBDMSO
HN
S
TBDMSO
8
TMS
cdcl3
7.71
7.68
7.45
7.27
7.13
6.46
6.20
5.30
4.36
4.05
3.94
3.75
3.35
3.28
2.89
2.87
2.48
2.25
2.05
1.79
1.67
1.26
0.89
0.16
0.11
0.00
1.02
0.96
0.85
1.82
7.25
1.03
1.95
1.85
1.48
7.28
18.00
11.84

Chapter 4

Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers show sequence specific DNA/RNA recognition

4.1 Introduction

In chapter 2 we have described that a five atom thioacetamido linker leading to a seven atom repeating backbone that is more useful for RNA recognition because of the reduced conformational flexibility of the amide relative to phosphodiester backbone. The TANA ONs showed sequence specific binding with RNA and provided a proof of design concept. In chapter 3 we have described extensive NMR studies of the TANA dimers and monomers and which indicated that there is marginal increase in the N-type sugar conformations over that of the native DNA which may be responsible for better RNA recognition.

In this chapter we present our important preliminary results on a backbone modification in oligonucleotides comprising [α + β-peptide] backbone using natural α-amino acids alternating with β-amino acid components, derived from thymidine (Figure 1). The strategy of the design, synthesis of the monomer blocks, oligomers synthesis and their complementary DNA/RNA recognition using UV-Tm measurements is discussed.

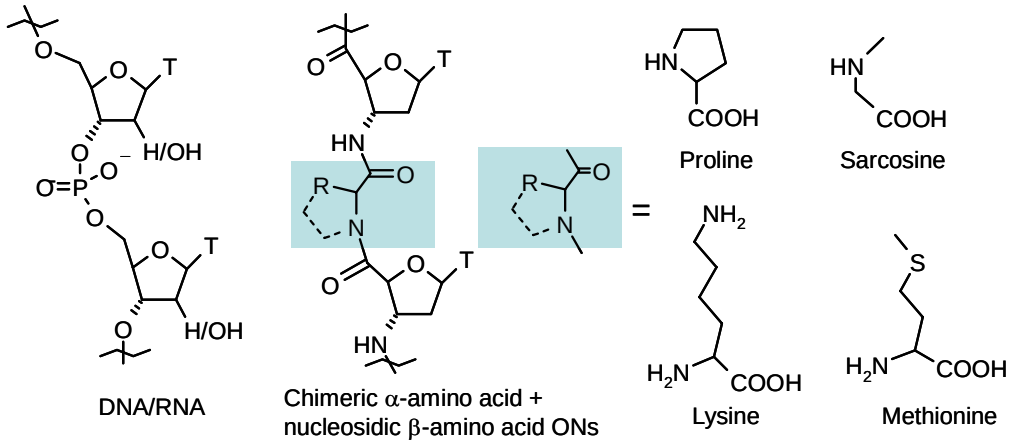


Figure 1. Structure of DNA/RNA, chimeric α- amino acid + nucleoside-β-amino acid ONs and structure of α-amino acids.

4.2 Design of chimeric α- amino acid + nucleoside-β-amino acid Oligonucleotides

The modification is aimed at including various structural features from different sources as shown in Figure 2 that impart important properties for its applications using antisense principle and gene targeting.

Martin Egli and co workers suggested that ONs having a five atom amide (linker) leading to a seven atom repeating backbone was found to be more useful for better RNA

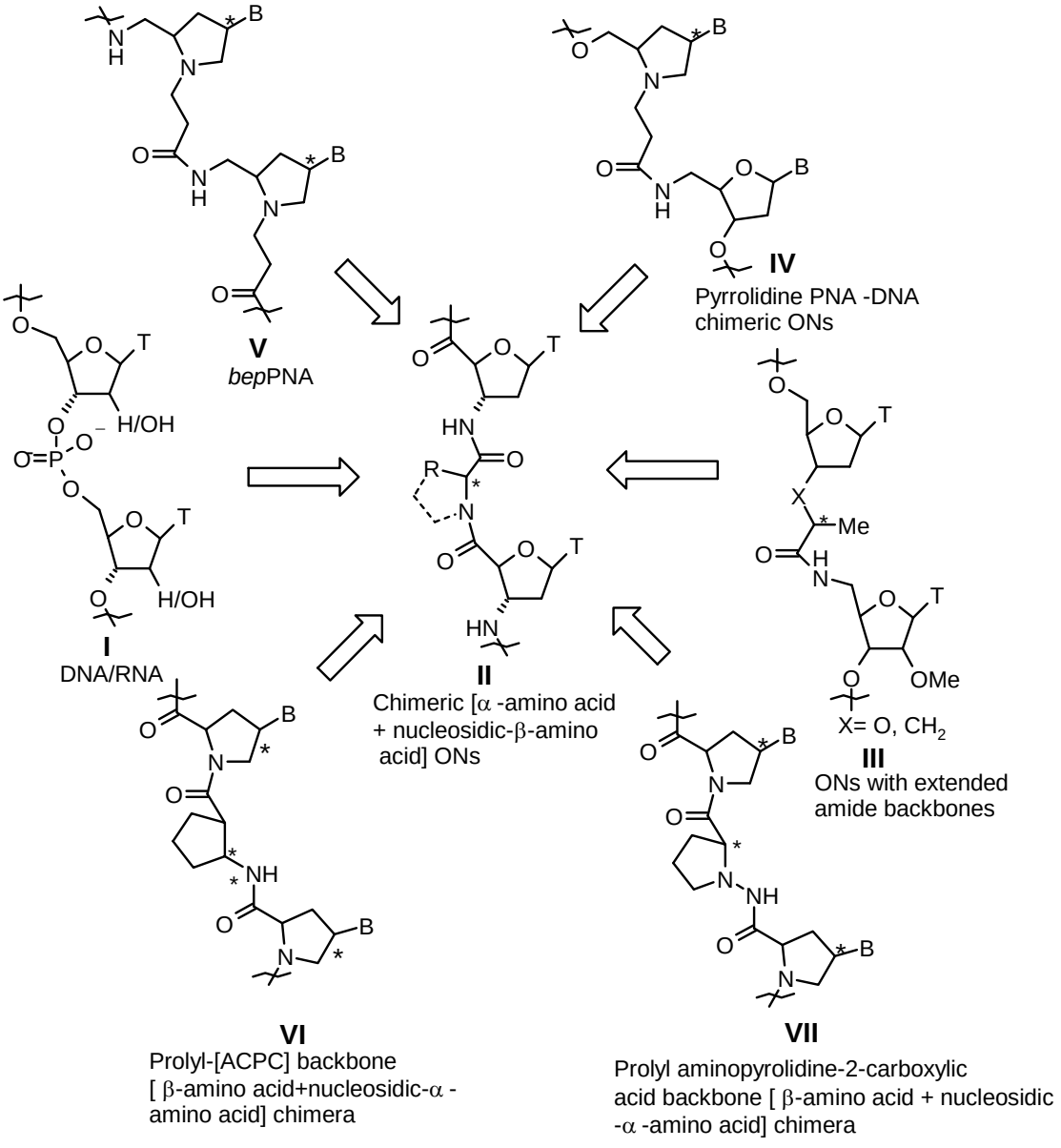


Figure 2. Design of chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide ONs

recognition because of the reduced conformational flexibility of the amide relative to phosphodiester backbone (Figure 2, Amide III).[1] An additional atom in the five-atom amide linker that provided requisite flexibility showed better stabilization of the duplexes with RNA and CD studies indicated increase in A- type of duplex structure.[2] Thermodynamic results on duplex formation of DNA/RNA ONs with multiple amide bonds show less unfavorable entropic component relative to wild type DNA:RNA despite the presence of an additional atom in the backbone.[2] The reduced conformational flexibility of the amide bond leading to favorable pre-organization could be one reason for increased stability of the duplexes. X ray structure of the duplex with RNA revealed a *trans* amide conformation and that the distance between the neighbouring base pairs was not affected by the longer backbone.[1,2]

The backbone extended peptide nucleic acids (*bep*PNA), cationic, chiral PNA analogues with an extra atom in the backbone (*bep*PNA) was developed in our laboratory (Figure 2, V).[3] The (2*S*, 4S) geometry of the pyrrolidine ring, and an additional carbon atom in the backbone of homopyrimidine-*bep*PNAs resulted in the optimization of the inter-nucleobase distance, such that selective binding to complementary RNA over DNA was observed in the triplex mode. It was evident from circular dichroism studies that oligomers with mixed aminoethylglycyl-*bep* (*aeg*-*bep*) repeating units, and also *bep*PNA with homogeneous backbone attained structures quite different from those of *aeg*PNA_2: RNA/DNA complexes. The *bep*PNA, when incorporated in a duplex forming mixed purine-pyrimidine sequence, also showed a preference for binding complementary RNA over DNA. More over, the preliminary results from our laboratory on the chimeric phosphate-amide extended backbone revealed that 2*R*, 4R pyrrolidine-amide chimera were accommodated better in triplex forming sequences (Figure 2, IV).[4]

Vilaivan and co-workers have described a series of novel conformationally rigid pyrrolidinyl peptide nucleic acids (PNA) based on D-prolyl-2-aminocyclopentanecarboxylic acid (ACPC) backbones (Figure 2, VI). Investigation of the binding properties of possessing different stereochemistry at the ACPC part with DNA revealed that the PNA containing (1S, 2S)-ACPC can form a very stable 1:1 complex with the complementary DNA in a sequence-specific manner.[5] A mixed-base, α-amino acid containing, pyrrolidinyl peptide nucleic acid (PNA) binds strongly and

selectively to complementary DNA in an exclusively antiparallel fashion. The PNA-DNA binding specificity strictly follows the Watson-Crick base-pairing rules.[6] Also they have developed a novel pyrrolidinyl peptide nucleic acids comprising alternate sequences of nucleobase-modified D-proline and β-amino acid spacers selected from L-aminopyrrolidine-2- carboxylic acid, D-aminopyrrolidine-2-carboxylic acid, (1*R*, 2*S*)-2-aminocyclopentane carboxylic acid and β-alanine. Gel-binding shift assay revealed that the homothymine PNA decamer bearing D-aminopyrrolidine-2-carboxylic acid spacer binds with $(dA)_{10}$ (Figure 2, VII).[7]

In this chapter, we describe synthesis and DNA/RNA binding studies of amide linked oligothyminyl DNA analogues (Figure 3) comprising conformationally constrained α-amino acids (L-proline **2** and sarcosine **3,** Figure 1), positively charged α-amino acid (L-lysine **4,** Figure 3) and a neutral α-amino acid (L-methionine **5,** Figure 3) in the backbone, alternating with a β-amino acid nucleoside component derived from thymidine (**1**, Figure 3). The advantages could be a) In 3'-deoxy-3'-amino- 2'deoxyribose sugar the five membered ring pucker is in preferred 3'-endo conformation that is better suited for m-RNA recognition.[8] b) The new synthesis of nucleoside-β-amino acid is quite simple using recently reported TEMPO/BAIB method [9] c) The flexibility of the five atom internucleoside amide linker could be better fitting for replacement of phosphate group for RNA binding taking into account the shorter amide bond compared to the phosphodiester linkage.[1] d) Negatively charged phosphate group can be replaced by neutral (proline, sarcosine and methionine) or positively charged[10] (Lysine) alternatives.

The specific objectives of this chapter are

i) Synthesis of Fmoc protected nucleoside -β-amino acid.

ii) Solid phase synthesis of chimeric [α-amino acid + nucleosidic -β-amino acid] oligomers.

iii) Cleavage from the solid support and purification, characterization of the oligomers.

iv) The binding studies of the modified ONs with DNA/RNA using UV and CD spectroscopy.

4.3 Results and discussions

Our synthetic strategy was aimed for assembling the molecules (Figure 3) on solid support, for its subsequent adaptation to well established conventional peptide chemistry.[11]

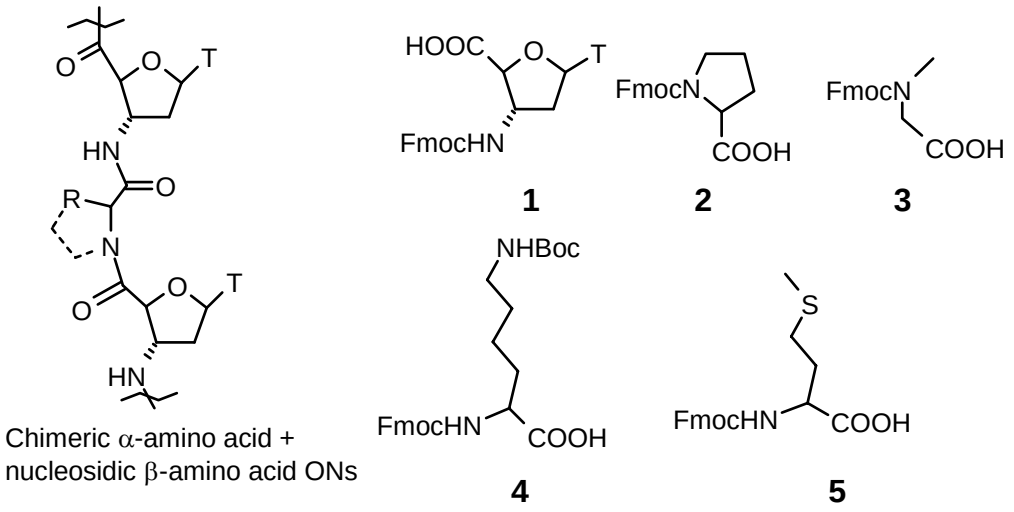


Figure 3. Generalized structure of chimeric α-amino acid + nucleoside -β-amino acid ONs and the monomer building blocks.

4.3.1 Synthesis of thyminyl sugar-amino acid monomer unit: *Synthesis of nucleoside 5'-carboxylic acid*

There are relatively few methods describing the oxidation of the 5'-hydroxymethylene of nucleosides to 5'-carboxylates. One of the most widely applied methods for affecting the oxidation of the 5'-hydroxyl of unprotected nucleosides employs molecular oxygen and a platinum catalyst.[12] However, this method affords relatively low yields when applied to 2', 3'-isopropylidene-protected nucleosides.[13]

The 5'-hydroxymethylene group of nucleosides can be easily oxidized using potassium permanganate under strongly alkaline reaction conditions.[14] This limits the potassium permanganate methodology to purine-containing nucleosides.

Verma and co-workers developed a method utilizing ruthenium trichloride and sodium periodate under Sharpless conditions (Figure 4) to obtain the 5'-carboxylic acids of 2',3'-isopropylidenepurine- containing nucleosides in high yield.[15] Unfortunately, this method has only limited application among purines as the complete loss of nucleobases occurred

in case of pyrimidines. Extension of ruthenium trichloride-mediated oxidation to 2', 3'-isopropylidenepyrimidine-containing nucleosides requires the use of both alkaline conditions and potassium persulfate (Figure 4).[16, 18] The main drawback of this reaction is the use of strongly alkaline conditions which makes it unsuitable for adapting it to other amino-protected nucleosides. The products were also often contaminated with inorganic salts.

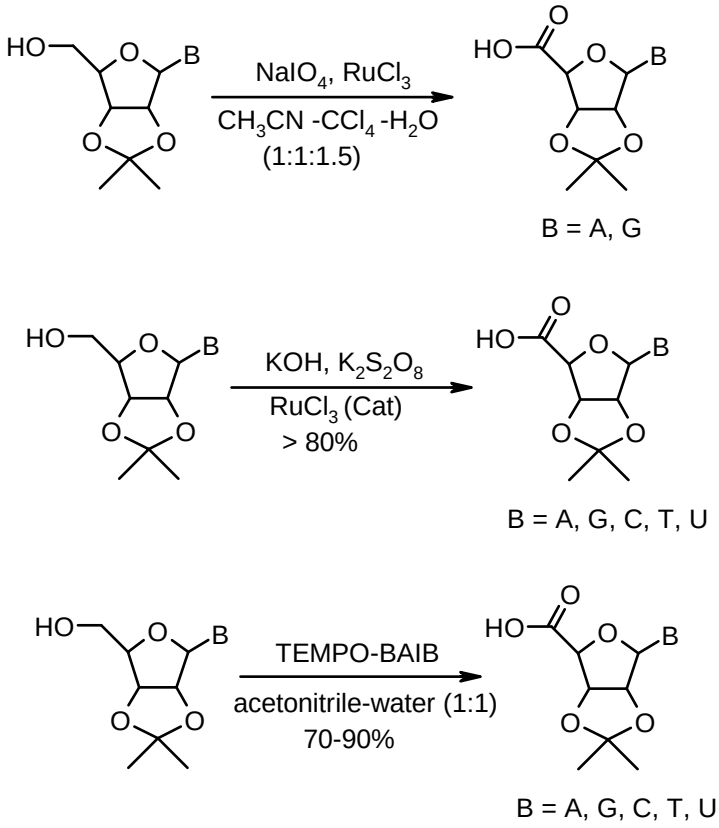
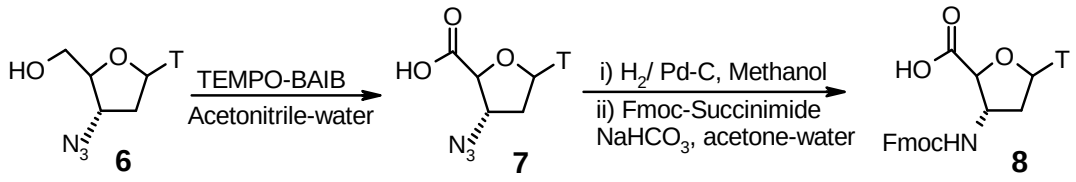


Figure 4. Some methods of synthesis of nucleoside 5'-acid

A recent method of oxidation of alcohols to ketones and aldehydes using catalytic amounts of 2,2,6,6-tetramethyl-1-piperidinyloxyl (TEMPO) and stoichiometric amounts of an organic oxidant, [bis(acetoxy)- iodo]benzene (BAIB)[9a] has drawn much attention because of its mildness and efficiency. The active oxidant is an *N*-oxoammonium salt generated by dismutation of TEMPO; BAIB is necessary to regenerate TEMPO by oxidizing the corresponding hydroxylamine of TEMPO. The reaction generates acetic acid and iodobenzene as byproducts and is different from most other TEMPO mediated oxidations in that it avoids inorganic salt contaminants.[9a] In addition, *N*-oxoammonium

salt-mediated oxidations are compatible with double and triple bonds, esters, ethers, acetals, epoxides, amides, halides, and azides. Finally, protecting groups such as TBDMS, THP, MOM, Boc, Cbz, benzyl, and acetyl are also stable to the reaction conditions.[17] In the presence of high concentrations of water, *N*-oxoammonium salts convert aliphatic alcohols to their respective carboxylic acids (Figure 4).[9b] This method of general procedure for the production of 5'-carboxylic acids of nucleosides is very mild and the desired products were obtained in good yields and separated from TEMPO and reaction byproducts by trituration with diethyl ether. The mildness of this reaction and its tolerance of acid sensitive, base sensitive and oxidatively labile functional groups should make it an attractive method for the oxidation of primary alcohols to carboxylic acids for both nucleosides. We have applied this methodology to convert the 5'-hydroxy methylene group of 3'-azido thymidine **6** to the corresponding acid **7**. Compound **6** were easily converted to the corresponding acid **7** under TEMPO-BAIB oxidation conditions and product was recovered by trituration with ether and acetone (Scheme 1). The formation of the product was confirmed by disappearance of the peaks due to the 5'-H protons in the ^{1}H NMR spectra.



Scheme 1. Synthesis of thyminyl sugar-amino acid monomer unit. (TEMPO 2,2,6,6-tetramethyl-1-piperidinyloxyl; BAIB [bis(acetoxy)-iodo]benzene; Fmoc fluorenylmethoxycarbonyl)

The 3'-azido group in **7** was then hydrogenated using Pd-C to amine which was subsequently Fmoc protected to give the monomer building block **8**.

4.3.2 Oligomer synthesis

The solid support used was Rink amide resin that was derivatized with L-lysine using peptide coupling reagents. Commercially available Rink amide resins having the loading value 0.5- 0.7 mmol/ g. The loading value was reduced to 0.25 mmol/ g by capping with acetic anhydride, which is found to be suitable to build a peptide oligomer. The loading

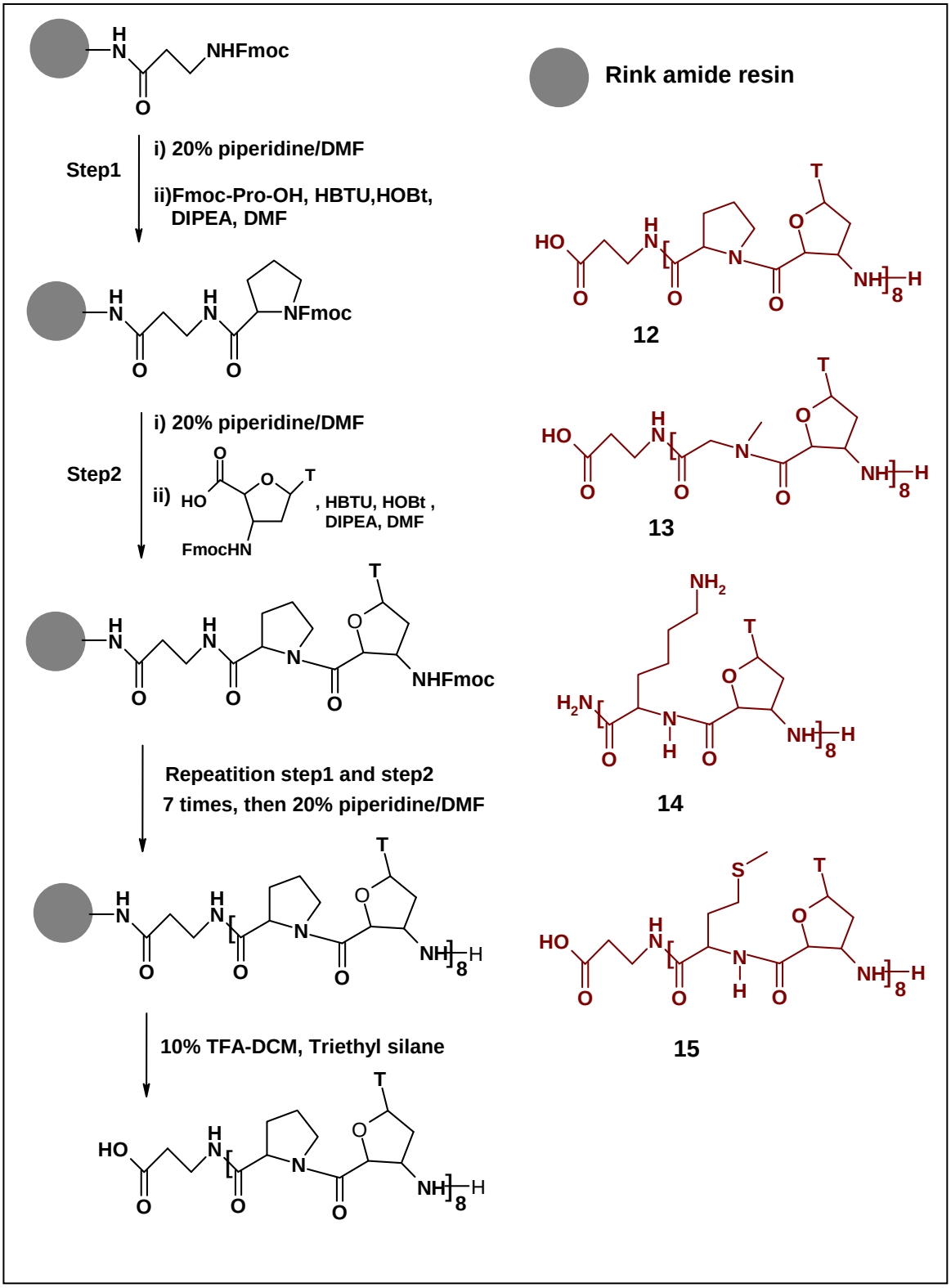
value of the resin was determined by the non-destructive picrate assay.[19] The synthesis was carried out by repetitive coupling cycles as shown in Scheme 2, each cycle comprising i) deprotection of the Fmoc amino group using 20% Piperidine in dichloromethane to generate the amine group ii) coupling of the incoming protected N^{α}-Fmoc-amino acids (**2**, **3**, **4**, **5**, Figure 3) with the resin bound L-lysine α- amino groups using coupling reagents HBTU, HOBt and DIEA as a base in DMF as a solvent iii) deprotection of the Fmoc amino group using 20% Piperidine in dichloromethane to generate the amine group iv) coupling of 3' NH Fmoc protected thyminyl β-amino acid **1.** The deprotection and coupling steps were monitored by the ninhydrin method using Kaiser's test.[20] The coupling cycle is repeated according to the length of the required oligomer.

4.3.3 Cleavage of the chimeric ONs from the Solid Support

The oligomers were cleaved from the solid support, using 10% Trifluoroacetic acid in DCM and triisopropyl silane as the scavenger. A cleavage time of 0.5 h at room temperature was found to be optimum. The side chain protecting groups (the Boc protection in case of the sequences **14**) were also cleaved during this cleavage process. After cleavage reaction, the TFA-DCM mixture was evaporated by passing nitrogen gas and the oligomer was precipitated by addition of dry diethyl ether.

Table 1. Amino-acid backbone modified ONs and their HPLC and MALDI-TOF mass characterization.

Entry	Sequence	Mol Formula	HPLC t_R (min)	Mass	
				calculated	observed
1	β-ala-(Pro-**t**)$_8$-H **12**	$C_{123}H_{145}N_{33}O_{42}$	12.7	2763.7	2785.6(+Na^+) 2801.6(+K^+)
2	β-ala-(Sar-**t**)$_8$-H **13**	$C_{107}H_{134}N_{33}O_{42}$	12.3	2555.4	2578.8(+Na^+) 2594.9(+K^+)
3	(Lys-**t**)$_8$-H **14**	$C_{128}H_{186}N_{40}O_{41}$	11.7	2941.1	2941.7
4	β-ala-(Met-**t**)$_8$-H **15**	$C_{123}H_{147}N_{33}O_{42}S_8$	12.6	3036.3	3037.0



Scheme 2. Solid-Phase synthesis of amino acid backbone modified ONs and structure of ONs synthesized.

4.3.4 Purification of the ONs

All the cleaved oligomers were subjected to initial gel filtration. The purity of the oligomers was checked by analytical RP-HPLC (C18 column, CH_3CN-H_2O system), which shows more than 65-70% purity. These were subsequently purified by reverse phase HPLC on a semi preparative 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 (Figure 5, 6 and 7).

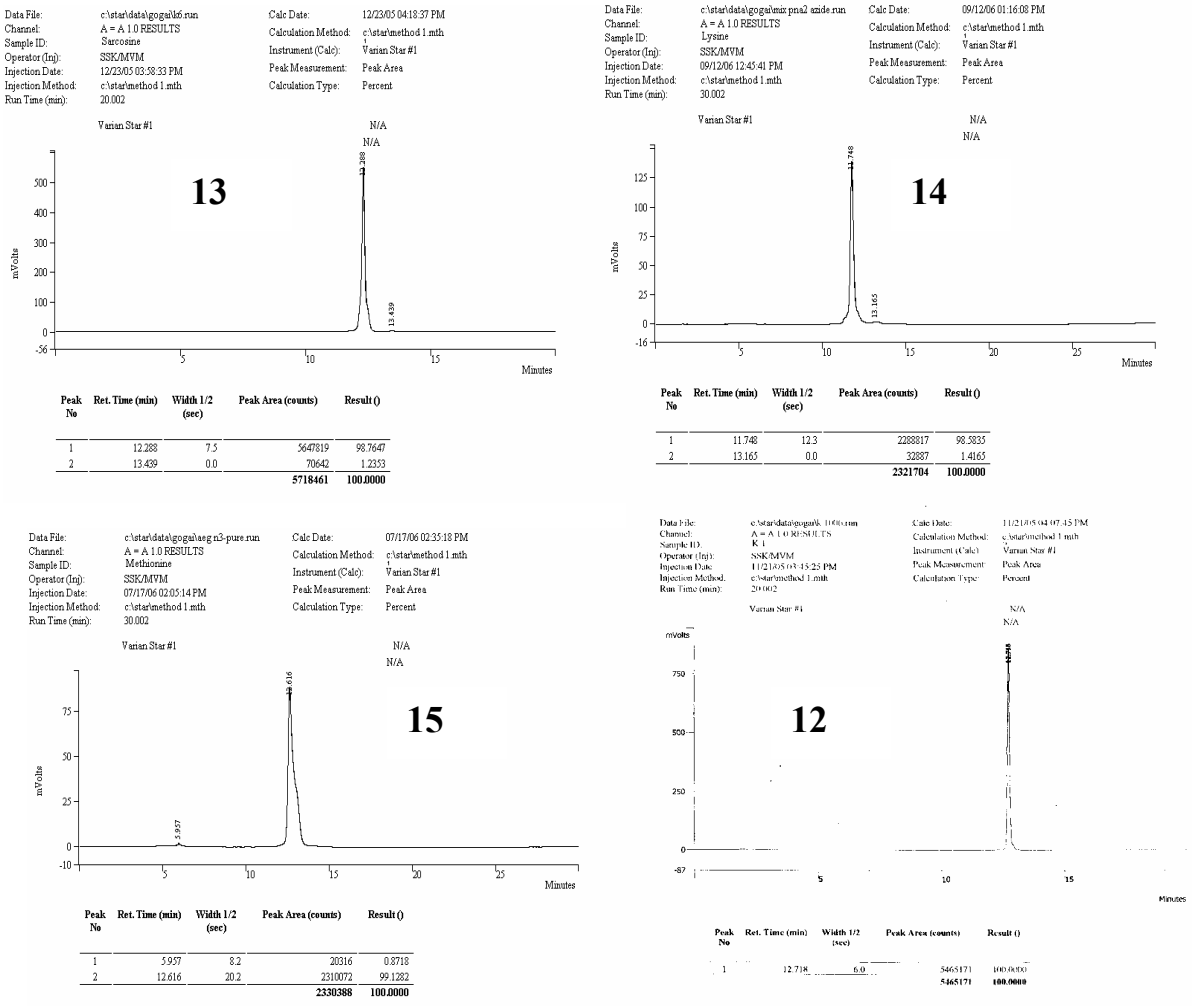


Figure 5. RP HPLC of ONs **12**, **13**, **14** and **15**.

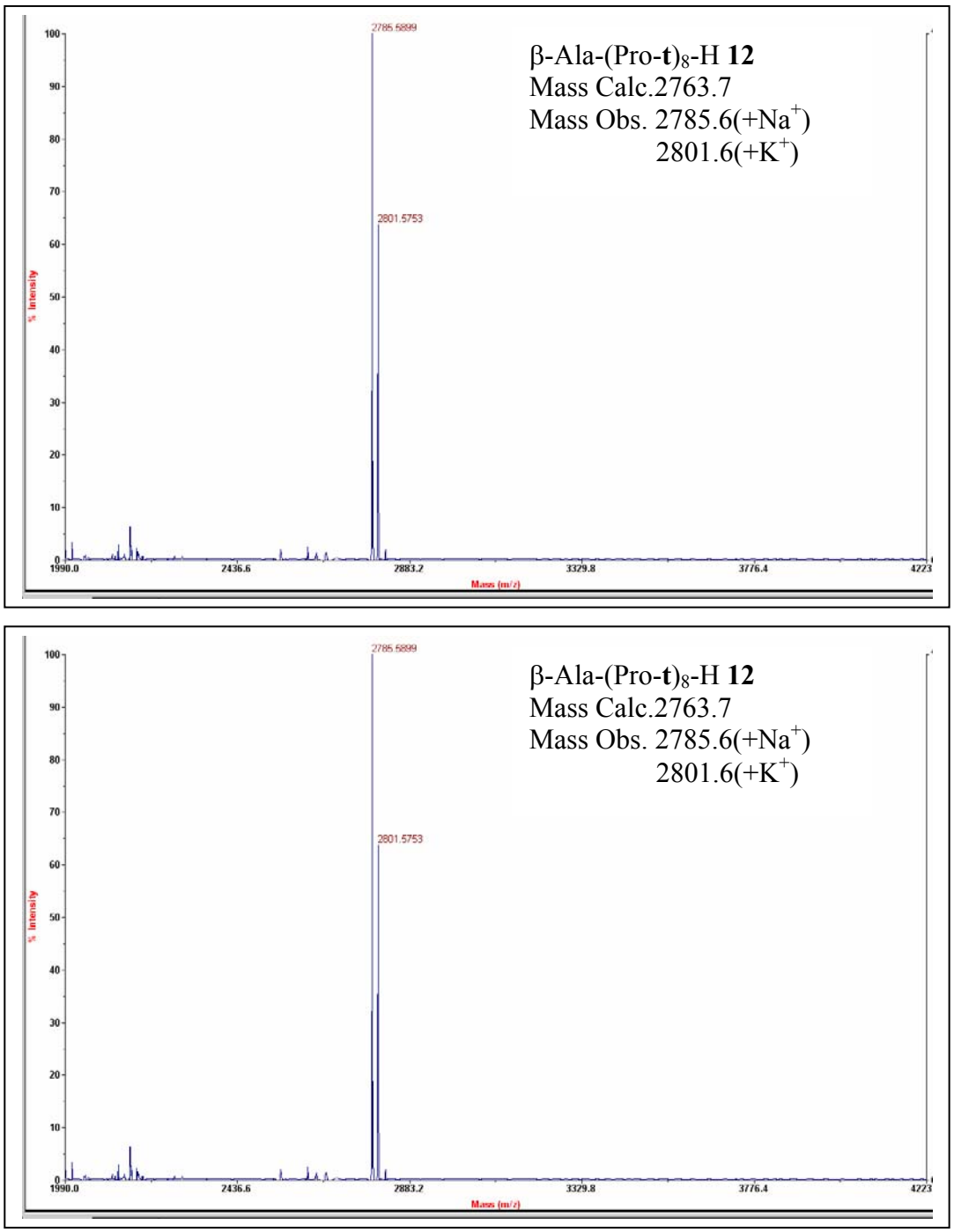
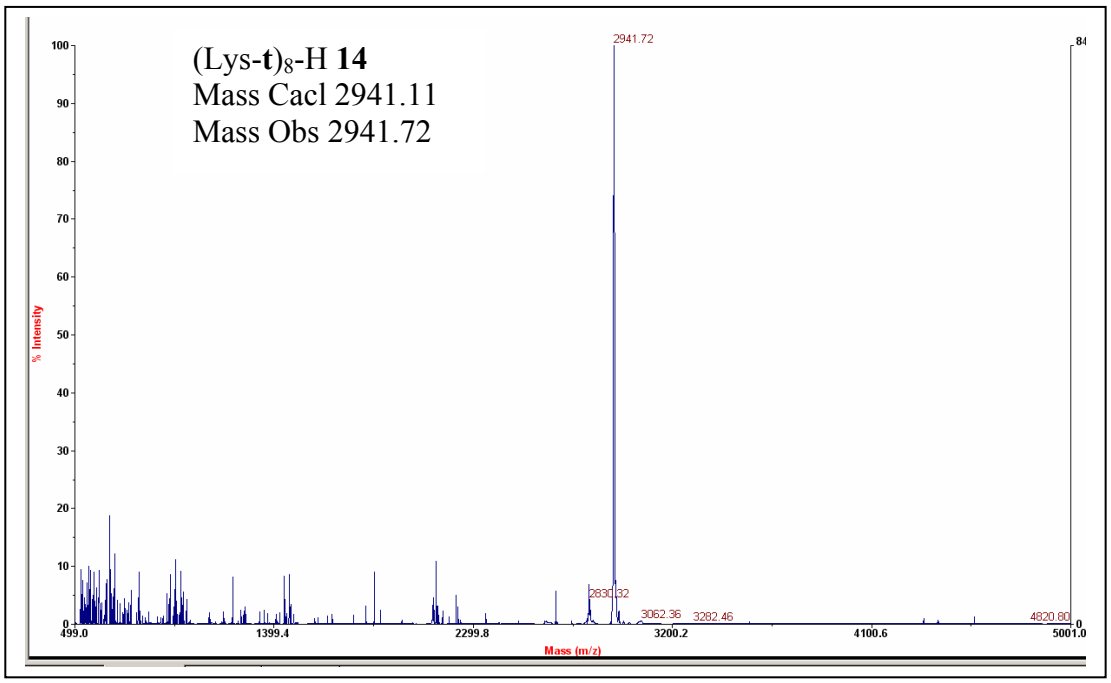


Figure 6. MALDI-TOF mass of **12** and **13**



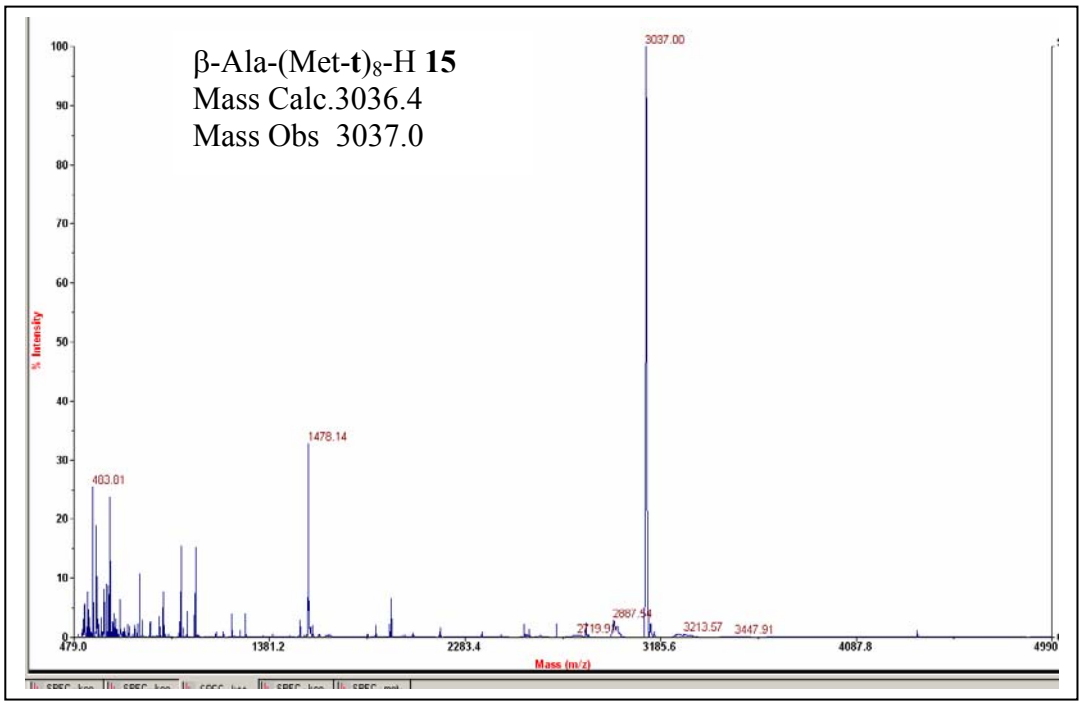


Figure 7. MALDI-TOF mass of **14** and **15**

4.3.5 Binding Stoichiometry: UV-mixing Curves

Ultraviolet (UV) absorption and circular dichroism (CD) measurements are extremely useful to determine the stoichiometry and stability of duplexes and triplexes.[21] The stoichiometry of the paired strands may be obtained from the mixing curves, in which the optical property at a given wavelength is plotted as a function of the mole fraction of each strand.

Various stoichiometric mixtures of ONs **12**, **13**, **14** and **15** with DNA **9** and RNA **10** were made with relative molar ratios of (ON :DNA **9** /RNA **10**) strands of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 100:0, all at the same strand concentration 2 μM in sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA (10 mM, pH 7.0). The samples with the individual strands were annealed and the UV-absorbance at 260 nm was recorded. The UV-absorbance at different molar ratios especially useful for plotting mixing curves. This systematic change in the UV-absorbance upon variable

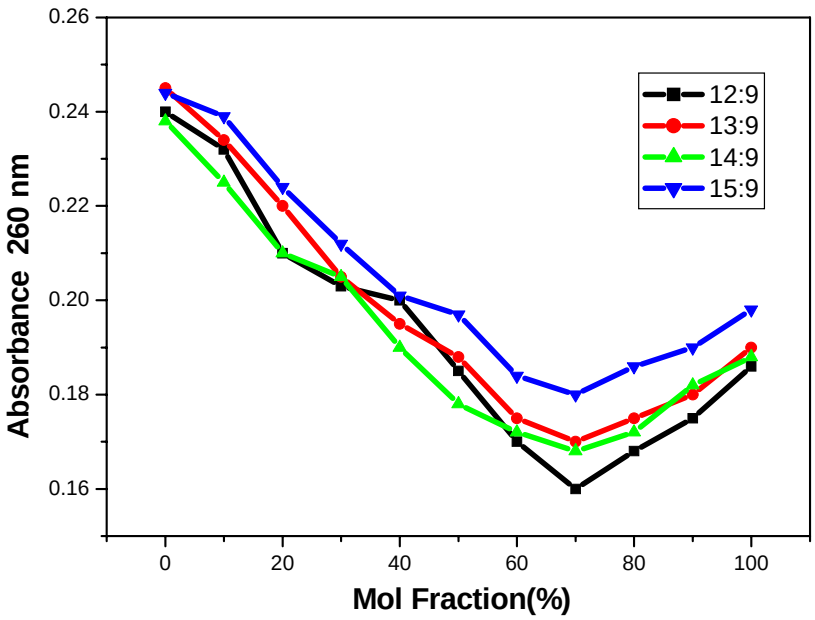


Figure 8. Job's plot of modified ONs **12**, **13**, **14** and **15** with complementary DNA **9** at 25ºC in the relative molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, and 100:0 (10mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 0.1 mM EDTA.

stoichiometric mixing of ONs and DNA/RNA components can be used to generate Job's plot. There is a drastic shift in the λ_{max} value of the ON- DNA **9** /RNA **10** mixture when

the concentration of ONs **12**, **13**, **14** and **15** in the mixtures increased to 60 to 70%, which further increases behind those proportions. This supports the formation of (ONs **12**, **13**, **14** and **15**)$_2$: DNA**9**/ RNA **10** triplexes (Figure 8 and 9).

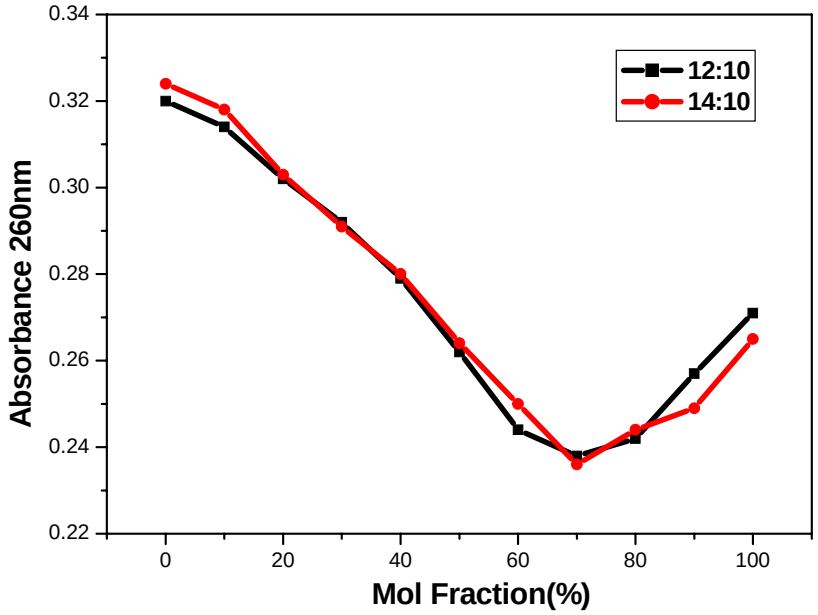


Figure 9. Job's plot of modified ONs **12** and **14** with complementary RNA **10** at 25ºC in the relative molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, and 100:0 (10mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 0.1 mM EDTA

4.3.6 UV-Melting studies of Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers -DNA/RNA complexes. *Thermal Stability of triplexes*

To establish the binding stoichiometry of the Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers (**12**, **13**, **14** and **15)** Job's plot study was undertaken with complementary DNA/RNA sequences (Figure 8 and 9).[22] The binding stoichiometry was clearly established as 2:1 between sequence **12**, **13**, **14** and **15** with complementary **DNA 9** and RNA **10** using UV spectroscopy (Figure 8 and 9). All the binding studies of the sequences listed in Table 1 with complementary DNA (Table 2, Figure 10 A & B) /RNA (Table 3, Figure 11 A & B) were carried out in 2:1 stoichiometry using UV-T_m experiments. For single nucleotide mismatch studies the UV-*T*m measurements were

carried out using the single nucleobases mismatch sequences DNA **16** (Table 2, Figure 10 C & D) and RNA **17** (Table 3, Figure 11C & D).

Table 2. *T*m(ºC) values of ONs: DNA / Mismatch DNA

Entry	ON Sequence	DNA **9**	DNA **16**	Δ*T*m (mismatch)
1	5' **TTTTTTTT** 3' **11**	17.8	-	-
2	β-Ala-(Pro-**t**)$_8$-H **12**	49.0	37.5	11.5
3	β-Ala-(Sar-**t**)$_8$-H **13**	49.1	35.3	13.8
4	(Lys-**t**)$_8$-H **14**	56.3	38.5	17.8
5	β-Ala-(Met-**t**)$_8$-H **15**	47.3	32.9	14.4

DNA **9** 5'GCAAAAAAAACG 3' DNA **16** 5' CAAAATAAACG 3'(Single nucleobase mismatch sequence)

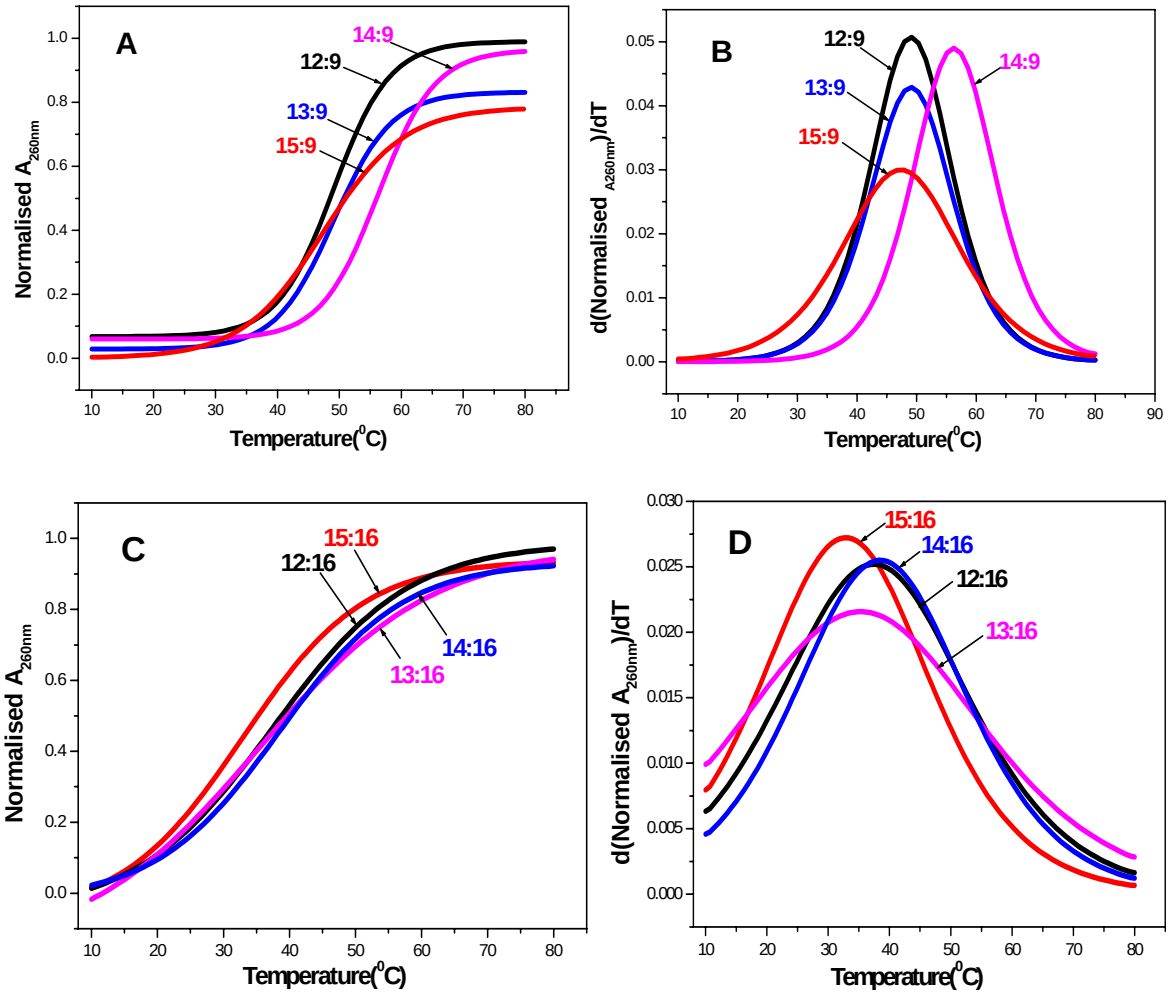


Figure 10. A.UV-melting Curves of **12**, 1**3, 14** and **15** with complementary DNA **9** and **B.** Corresponding first derivative Curves. **C.** UV-melting Curves of **12**, **13, 14** and **15** with mismatch DNA **16** and **D**. Corresponding first derivative Curves.

Table 3. *T*m(ºC) values of ONs: RNA / Mismatch RNA

Entry	ON Sequence	RNA**10**	RNA **17**	Δ*T*m (mismatch)
1	5' **TTTTTTTT** 3' **11**	15.6	-	-
2	β-Ala-(Pro-**t**)$_8$-H **12**	60.8	44.6	16.2
3	β-Ala-(Sar-**t**)$_8$-H **13**	61.6	45.6	16.0
4	(Lys-**t**)$_8$-H **14**	69.0	55.5	13.5
5	β-Ala-(Met-**t**)$_8$-H **15**	57.3	41.2	16.1

RNA **10** **r**(5' GCAAAAAAAACG 3') RNA **17** **r**(5' GCAAAAUAAACG 3') (single nucleobase mismatch sequence).

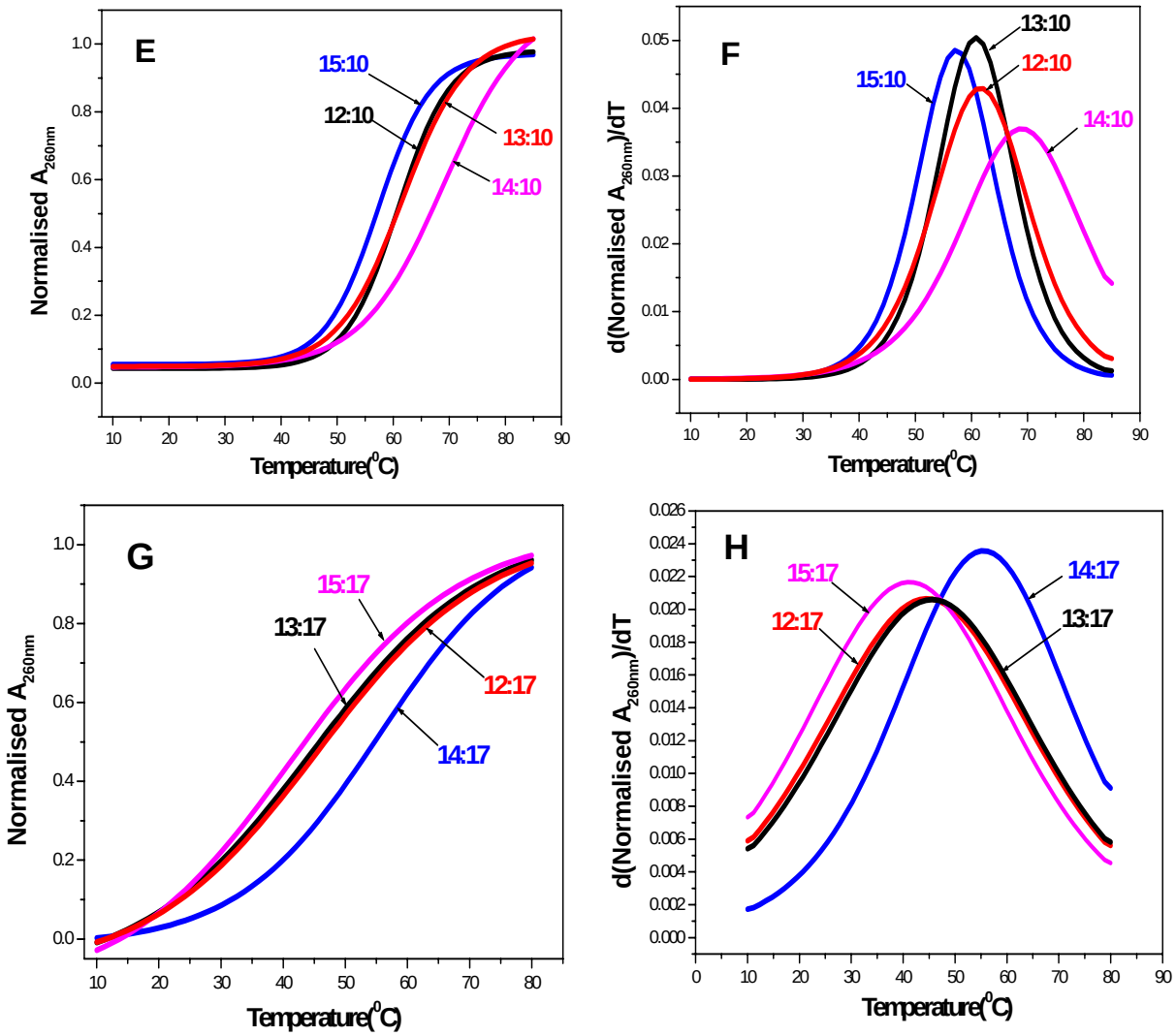


Figure 11. E.UV-melting Curves of **12**, **13, 14** and **15** with complementary RNA **10** and **F.** Corresponding first derivative Curves. **G**.UV-melting Curves of **12**, **13, 14** and **15** with mismatch RNA **17** and **H.** Corresponding first derivative Curves.

Table 4. Comparison of UV-*T*m values in (ºC) of chimeric modified ONs: complementary DNA/RNA

Entry	Sequence	DNA **9**	RNA **10**	$\Delta Tm_{(RNA-DNA)}$
1	TTTTTTTT **11**	17.8	15.6	-
2	β-Ala-(Pro-**t**)$_8$-H **12**	49.0	60.8	+ 11.8
3	β-Ala-(Sar-**t**)$_8$-H **13**	49.1	61.6	+ 12.5
4	(Lys-**t**)$_8$-H **14**	56.3	69.0	+ 12.7
5	β-Ala-(Met-**t**)$_8$-H **15**	47.3	57.3	+ 10.0

The hybridization properties of the amino acid modified oligomers **12**, **13**, **14** and **15** with complementary ONs DNA **9** GCAAAAAAAACG or RNA **10** r(GCAAAAAAAACG) were determined by thermal denaturation studies. UV melting experiments indicate that oligomers **12**, **13**, **14** and **15** hybridized to complementary DNA and RNA with melting temperatures (*T*m) higher than those of the complexes formed by oligothymidinyl-DNA fragment **11** of the same length (Table 2 & 3). Complexes formed by all the modified oligomers **12**, **13**, **14** and **15** with the RNA target showed higher stability (10-12ºC) than similar complexes formed with the DNA target (Table 4). The positively charged lysine oligomer **14** formed most stable complexes with both DNA (**14:9**) (Table 2, entry 4) and RNA (**14:10**) (Table 3, entry 3). The UV melting experiments of ON 1**4** with complementary DNA **9** and RNA **10** were performed at pH 5.5 but no significant change in melting temperature could be detected indicating that the side chain lysine amine group could be protonated even at physiological pH. The UV- melting experiment of all the modified ONs **12**, **13**, **14** and **15** were carried out using the single base mismatch DNA **16** and RNA **17**. A single base mismatch caused fall in *T*m (11-18 ºC) in the case of either DNA or RNA (Table 2 and 3) similar as in the case of PNA.[23] A single nucleotide mismatch studies revealed that the lysine peptide tolerates mismatch pair better than the other oligomers. The hybrid formed between ON **14** and mismatch RNA **17** was found posses highest mismatch tolerance (Table 3, entry 4).

Thus, the complexes between chimeric (α-amino acid + nucleoside-β amino acid) backbone ONs (**12, 13, 14** and **15**) and their complementary DNA/RNA (**9**/**10**) oligonucleotides were found to be highly stable as compared to DNA: DNA (**11:9**) or DNA: RNA (**11:10**) duplexes. The presence of positively charged α- amino acid (L-Lysine) contributes further to the strong binding properties. The undesired mismatch

tolerance in lysine containing ONs towards the mismatch RNA **17** could be the effect of electrostatic interactions between protonated lysine side chain amino groups with phosphate negative charge on complementary RNA. The stability of complexes of sarcosine and proline analogues, in which one amide hydrogen is absent, suggest the absence of contribution of known secondary structures of the (α-amino acid + β amino acid) peptide scaffolds [24] to the overall stability of the complexes. The 5-atom amide linked oligomers in this simple (α-amino acid + nucleoside-β amino acid) backbone exhibit much higher stability of complexes with both complementary DNA and RNA), complexation with RNA being favoured over DNA. (Table 4) The basis of this selectivity could be similar to the other 5-atom internucleoside amide likages as described in the introduction.[1]

4.3.7 CD spectroscopic studies of chimeric modified ONs: complementary DNA/RNA

The CD spectra of individual chimeric modified ONs **12**-**15** show a negative band in the 260- 290 nm regions (Figure 10 A). The CD spectra of complexes of **12**, **13**, **14** and **15** with complementary DNA **9** are nearly identical (Figure 10 B). The single stranded DNA **9** is showing positive band at wavelengths 280 nm, 240 nm and 210 nm and a strong negative band at 250 nm. In the case of the ONs **12, 13, 15** and complementary DNA **9** complexes, there are positive bands at 280-290 nm and 220-230 nm and two negative bands at 270 nm and 250 nm. The 250 nm band is stronger than the 270 nm band. In the case of ON **14**: DNA **9**, there are positive bands at 280 nm and 220 nm and negative bands at 250 nm and 210 nm. The CD spectra of the ONs with complementary DNA obtained were not so informative to predict the geometry of the complexes formed.

The single stranded RNA **10** is showing positive band at wavelengths 270 nm, 230 nm and 220 nm and negative bands at 250 nm and 210 nm (Figure 10 C). In case of complexes of ONs **12** and **13** with complementary RNA **10**, there are positive bands at 280 nm and 220-230 nm and two negative bands at 250 nm (Figure 10 C). In case of complex of ONs **14** with complementary RNA, there are two positive bands at 265 nm and 275 nm of equal intensity (Figure 10 C). The complex of ON **15** with complementary

RNA is also showing almost similar type of CD graph as ONs **12**, **13** and **14** (Figure 10 C). It is known in literature that the CD spectrum of the 2:1 hybrid formed between an

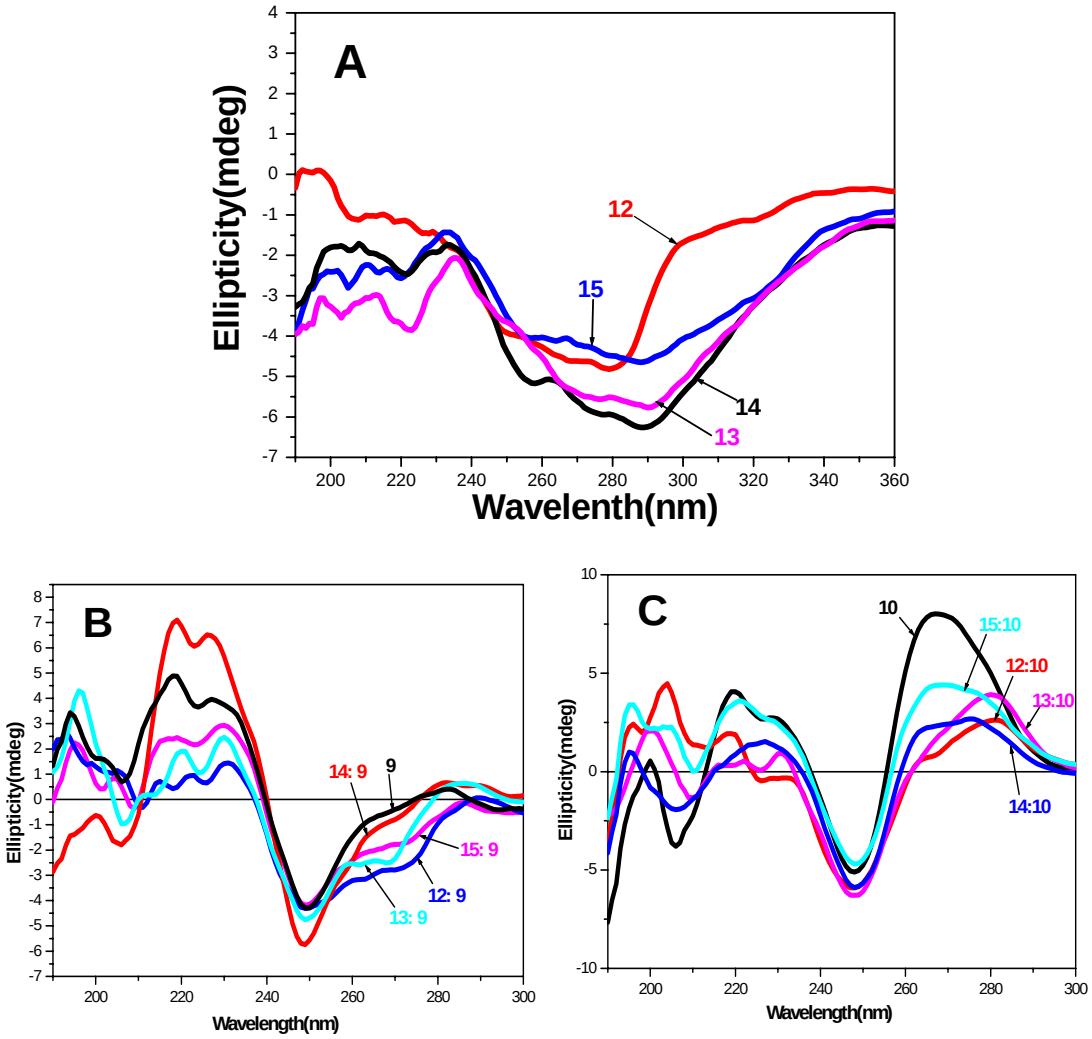


Figure 12. CD curves of **12**, **13**, **14** and **15 A.** single strands **B**. Complexes with complementary DNA and **C**. Complexes with complementary RNA

ON and complementary RNA showed ellipticity minima at 240-250 nm and ellipticity maxima at 280-290 nm is due to the formation of triple helical structure.[16] Thus CD spectrum of the complexes of **13**, **14** and **15** with the complementary RNA showed similar CD signature substantiating the formation of triple helical structures.

4.4 Conclusion

In conclusion, we have designed and synthesized a novel class of amino acid backbone extended DNA modification to improve the binding affinity and selectivity towards DNA/RNA recognition. The complementation studies with DNA/RNA reveals that the α -amino acid + nucleoside- β -amino acid chimeric ONs are showing unprecedented RNA binding selectivity over DNA. The use of naturally occurring α-amino acids in conjunction with easily accessible nucleoside derived β-amino acid described in this chapter, thus provides a very simple peptide scaffold to create nucleobase sequence for DNA/RNA recognition. From an application perspective, this amide backbone modified DNA analogue appears to have properties important for developing gene therapeutic drugs. Further studies on the mixed sequences having other than thymine monomer are necessary to draw clear conclusions. The studies on the mixed purine-pyrimidine sequences as well as incorporation of negatively charged and D-amino acids into the backbone to study the stereoelectronic requirements of the complexes are currently in progress in our laboratory.

Summary

- Synthesis of 3'-deoxy 3'-amino thymine 5'-carboxylic acid starting from 3'-azido thymidine using the TEMPO-BAIB oxidation as the key reaction.
- Synthesis of chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers using Fmoc solid phase peptide synthesis protocols. The four amino acids incorporated are L-Proline, Sarcosine, L-Lysine and L-Methionine.
- The Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers form triplex with their complementary DNA and RNA.
- The Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers form more stable complex with their complementary RNA than DNA.

4.5 References

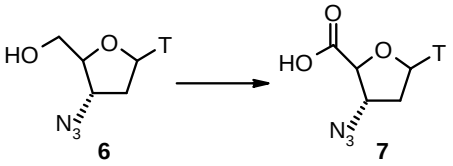
1. Wilds, C. J., Minasov, G., Natt, F., Matt von, P., Altmann, K.-H. and Egli, M. *Nucleosides Nucleotides Nucleic Acids, 2001*, **20**, 991-994.
2. Pallan, P. S., Matt von, P., Wilds, C., Altmann, K.-H.and Egli, M. *Biochemistry, 2006*, **45**, 8048-8057.
3. (a) Govindaraju, T. and Kumar, V. A. *Chem. Commun*., *2005*, 495-497. (b) Govindaraju, T. and Kumar, V. A. *Tetrahedron*, *2006*, **62**, 2321-2330.
4. Kumar, V. A. and Meena *Nucleosides Nucleotides Nucleic Acids*, *2003*, **22**, 1101-1104.
5. Suparpprom, C., Srisuwannaket, C., Sangvanich P. and Vilaivan, T. *Tett. Lett*., *2005*, **46**, 2833-2837.
6. Vilaivan, T. and Srisuwannaket, C. *Org. Lett.*, *2006*, **8**, 1897-1900.
7. (a) Vilaivan, T., Suparpprom, C., Harnyuttanakorn, P.and Lowe, G. *Tet. Lett, 2001,* **42**, 5533-5536. (b) Vilaivan, T. and Lowe, G. *J. Am. Chem. Soc*. *2002*, **124**, 9326-9327.
8. Thibaudeau, C., Plavec, J., Garg, N., Papchikhin, A. and Chattopadyaya, J. *J. Am. Chem. Soc*., *1994*, **116**, 4038-4043.
9. (a) De Mico, A., Margarita, R., Parlanti, L., Vescovi, A. and Piancatelli, G. *J. Org. Chem., 1997*, **62**, 6974. (b) Epp, J. B. and Widlanski, T. S. *J. Org. Chem*., *1999*, **64**, 293-295.
10. (a) Zhou, P., Wang, M., Du, L., Fisher, G. W., Waggoner, A. and Ly, D. H. *J. Am. Chem. Soc*., *2003*, **125**, 6878-6879. (b) Dempcy, R.O., Brown, K. A. and Bruice, T. C. *J. Am. Chem. Soc*., *1995*, **117**, 6140- 6141. (c) Deglane, G., Abes, S., Michel, T., Prevot, P., Vives, E., Debart, F., Barvik, I., Lebleu, B. and Vasseur, J.-J. *ChemBioChem,* 2006, **7**, 684-692.
11. Nielsen, P. E. and Egholm, M. In Peptide Nucleic Acids (PNA). Protocols and Applications; Nielsen, P, E., Egholm, M. Eds. Horizon scientific Norfolk. CT, 1999.
12. Moss, G. P., Reese, C. B., Schofield, K. and Todd, A. R. *J. Chem. Soc.*, *1963*, 1149.
13. Harper P. J. and Hampton A. *J. Org. Chem., 1970*, **35**, 1688.
14. (a) Harmon, R. E., Zenarosa, C. V. and Gupta, C. V. *Chem. Ind. (London) 1969*, 1141. (b) Hutchison, A. J., Williams, M., de Jesus, R., Yokoyama, R., Oei, H. H., Ghai, G. R., Webb, R. L., Zoganas, H. C., Stone, G. A. and Jarvis, M. F. *J. Med. Chem., 1990*, **33**, 1919.
15. Singh, A. K. and Varma, R. S. *Tet. Lett. 1992*, **33**, 2307.
16. Varma R. S. and Hogan, M. E. *Tet. Lett. 1992*, **33,** 7719.
17. De Nooy, A. E. J., Besemer, A. C.and van Bekkum, H. *Synthesis 1996*, 1153.

18. Varma, R. S., Hogan, M. E., Revankar, G. R. and Rao, T. S. *U. S. Patent* 5175266 *1992*.
19. (a) Erickson, B.W. and Merrifield, R. B. Solid Phase Peptide Synthesis. In the Proteins Vol. II, 3rd ed.; Neurath, H. and Hill, R. L. eds. Academic Press, New York, *1976*, pp 255. (b) Merrifield, R. B., Stewart, J. M. and Jernberg, N. *Anal. Chem*., *1966*, **38**, 1905-1914.
20. (a) Kaiser, E., Colescott, R. L., Bossinger, C. D. and Cook, P. I. *Anal. Biochem*., *1970*, **34**, 595-598. (b) Kaiser, E., Bossinger, C. D., Cplescott, R. L. and Olsen, D. B. *Anal. Chim. Acta*., *1980*, **118**, 149-151. (c) Sarin, V. K., Kent, S. B. H., Tam, J. P. and Merrifield, R. B. *Anal. Biochem*., *1981*, **117**, 147.
21. (a) Berova, N., Nakanishi, K. and Woody, R.W. Eds. *Circular Dichroism: Principles and Applications*; Wiley-VCH: New York, *2000*; pp 703-736. (b) Gray, D. M., Hung, S. -H. and Johnson, K. H. *Methods Enzymol*, *1995*, **246**, 19-34.
22. Job, P. *Ann. Chim.*, *1928*, **9**, 113-203.
23. Egholm, M., Buchardt, O., Nielsen, P. E. and Berg, R. H. *J. Am. Chem. Soc., 1992*, **114**, 1895-1897.
24. Hayen, A., Schmitt, M. A., Ngassa, F. N., Thomasson, K. A. and Gellman, S. H. *Angew. Chem. Int. Ed*., *2004*, **43**, 505-510.

4.6 Experimental

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 spectrometer. Column chromatographic separations were performed using silica gel 60-120 and 230-400 mesh, Ethyl acetate, Petroleum ether, Dichloromethane and Methanol as the solvent system. ^{1}H and ^{13}C spectra were obtained using Bruker AC 200 (200MHz) NMR spectrometers.

Synthesis of 3'-azido thymidine 5'-carboxylic acid 7

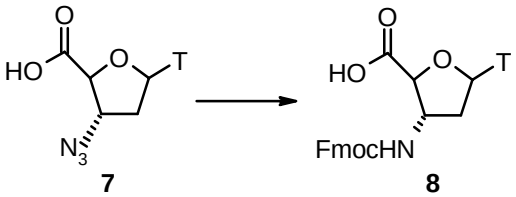


BAIB (3.9 g, 12.4 mmol) and TEMPO (180 mg, 1.12 mmol) and 3'-azido thymidine **6** (1.5 g, 5.6 mmol) were combined in a reaction vessel and to the mixture was added 12 ml of 1:1 acetonitrile-water solution. The reaction mixture was stirred for 3 h. Solvents were removed *in vacuo* and the resulting residue was triturated sequentially with diethyl ether and acetone, filtered and dried *in vacuo*. Yield: 1.22g, 78%.

M.P. 174-177 ºC FT-IR (Nujol) 3440, 2110, 1748, 1721, 1694, 1610 cm^{-1}

^{1}H NMR (DMSO-d_6) δ 1.78 (s, 3H), 2.2 (m, 2H), 4.35 (d, 1H), 4.64 (m, 1H), 6.2 (t, 1H), 8.36 (s, 1H).

Synthesis of 3'-NH Fmoc thymidine 5'-carboxylic acid 8



3'-azido thymidine 5' carboxylic acid **7 (**1.2 g, 4.26 mmol) was dissolved in 10 ml methanol and to it was added 120 mg (10%) Pd-C catalyst. The mixture was subjected to hydrogenation at 40 Psi of hydrogen pressure for 3.5 h. The catalyst was filtered over

celite and concentration of the filtrate *in vacuo* gave 1.1 g, 97% of 3'-deoxy amino thymidine 5'-carboxylic acid.

The amino acid 3'-deoxy amino thymidine 5'-carboxylic acid (1.1 g, 4.3 mmol) was treated with Fmoc-succinimide (1.74 g, 5.17 mmol) and $NaHCO_3$ (1.81 g, 21.5 mmol) in 10 ml 1:1 acetone-water. After stirring for five hours, acetone was evaporated and the reaction mixture was diluted with 5 ml water. The water layer was extracted with ethyl acetate to remove all the nonpolar unwanted impurities. The water layer was neutralized by adding dilute HCl to bring the pH = 7.0. The water layer was then extracted with ethyl acetate. The organic layer was washed with water (1 x 10 ml) and saturated NaCl solution (1 x 10 ml), dried over Na_2SO_4 and evaporated to dryness to afford 1.25 g of **8** (61%) as pure white amorphous solid. M.P.147-150°C.

^{1}H NMR ($CDCl_3$ + 1 drop DMSO-d_6, 200 MHz) 1.87 (s, 3H), 2.0-2.2 (m, 1H), 2.3-2.4 (m, 1H), 4.1-4.5 (m, 4H), 6.46 (t, 3H), 7.2-7.4 (m, 5H), 7.6-7.73 (m, 5H), 8.1(s, 1H), 10.4 (s, 1H).

^{13}C NMR (DMSO-d_6, 50 MHz) 12.4, 36.2, 46.6, 54.5, 65.6, 81.5, 85.1, 109.3, 120.1, 125.1, 127.1, 127.6, 136.4, 140.7, 143.8, 150.5, 155.5, 163.7, 172.2.

Mass calc. 477.42 obs. 499.62 (+ Na^+)

HPLC Analysis

Peptide purifications were performed on waters DELTAPAK-RP semi preparative C18 column attached to a Hewlett Packard 1050 HPLC system and JASCO-UV 970 variable-wavelength detector. An isocratic elution method with 10 % CH_3CN in 0.1 % TFA/H_2O was used with flow rate 1.5 ml/min. and the eluent was monitored at 254 nm. The purity of the oligomers was further assessed by RP C18 analytical HPLC column (25 x 0.2 cm. 5 μm) with gradient elution: A to 50 % B in 30 min, A = 0.1% TFA in H_2O, B = 0.1 % TFA in CH_3CN:H_2O 1:1 with flow rate 1 ml/min. The purities of the so purified oligomers were found to be > 98 %.

MALDI-TOF mass spectrometry

Mass spectral analysis was performed on a 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 hydroxy cinnamic acid).The samples were prepared by mixing 5 µl of matrix (0.5 M solution) and 1 µl ONs (OD 5-10/ml) and spotted 1 µl on a stainless steel plate for analysis.

UV-*T*m measurements

The concentration was calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases (i.e., T, 8.8 cm^2 / µmol; C, 7.3 cm^2 / µmol; G, 11.7 cm^2 / µmol and A, 15.4 cm^2 / µmol). The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM) and were annealed by keeping the samples at 90°C for 5 minutes followed by slow cooling to room temperature (annealing). Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Perkin-Elmer *Lambda 35 UV-VIS* spectrophotometer scanning from 5 to 85°C at ramp rate of 0.5°C per minute. The data were processed using Microcal Origin 6.0 and *T*m values derived from the derivative curves.

Circular Dichroism

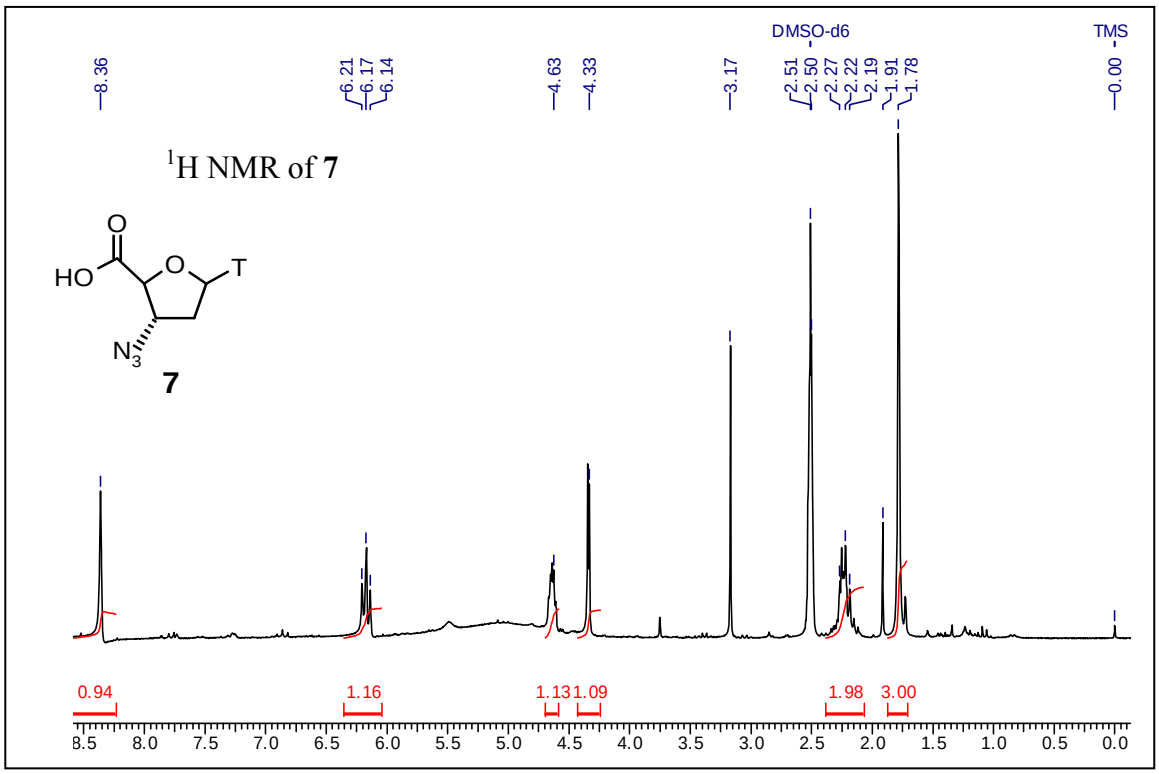
CD spectra were recorded on a JASCO J-715 spectro polarimeter. The CD spectra of the ON complexes and the relevant single strands were recorded in 10 mM sodium phosphate buffer, 100 mM NaCl, 0.1 mM EDTA, pH 7.0. The CD spectra were recorded as an accumulation of 5 scans from 300 to 190 nm using a 1 cm cell, a resolution of 0.1 nm, band-width of 1.0 nm, sensitivity of 2 mdeg, response 2 sec and a scan speed of 50 nm/min and most of the experiments were performed at 20ºC.

4.7 Appendix

^{1}H NMR of **7** and **8**

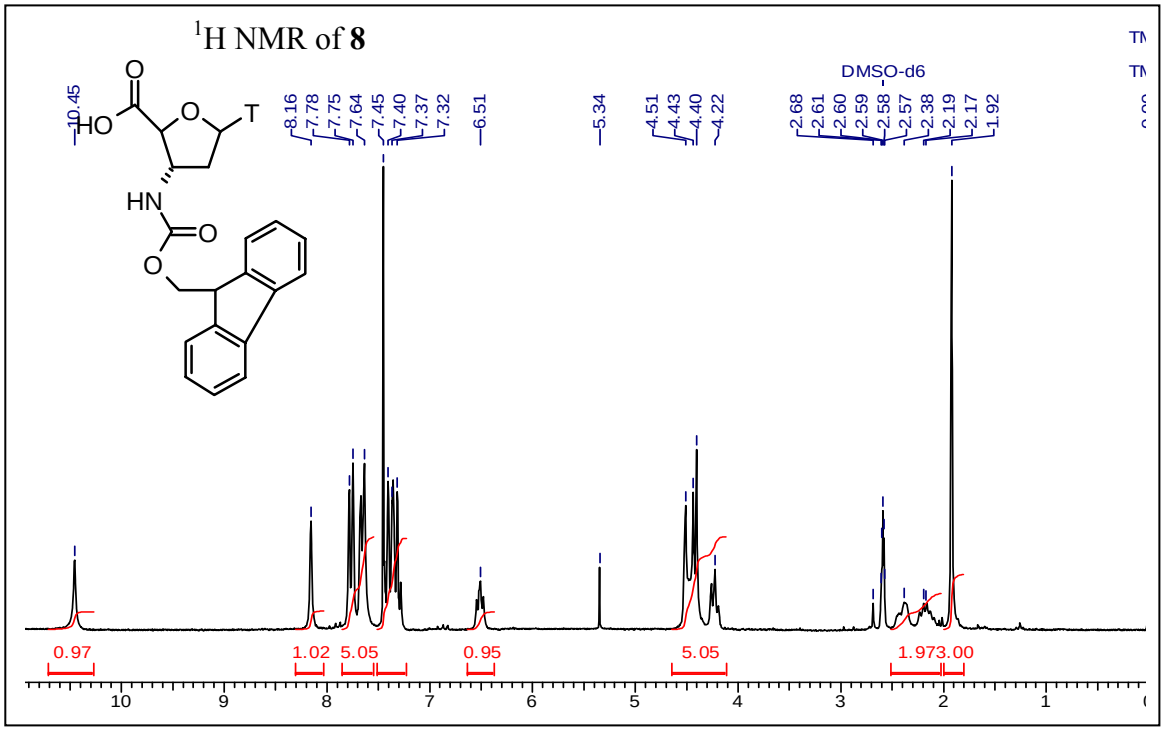
^{13}C NMR and DEPT NMR of **8**

Mass Spectra of **8**

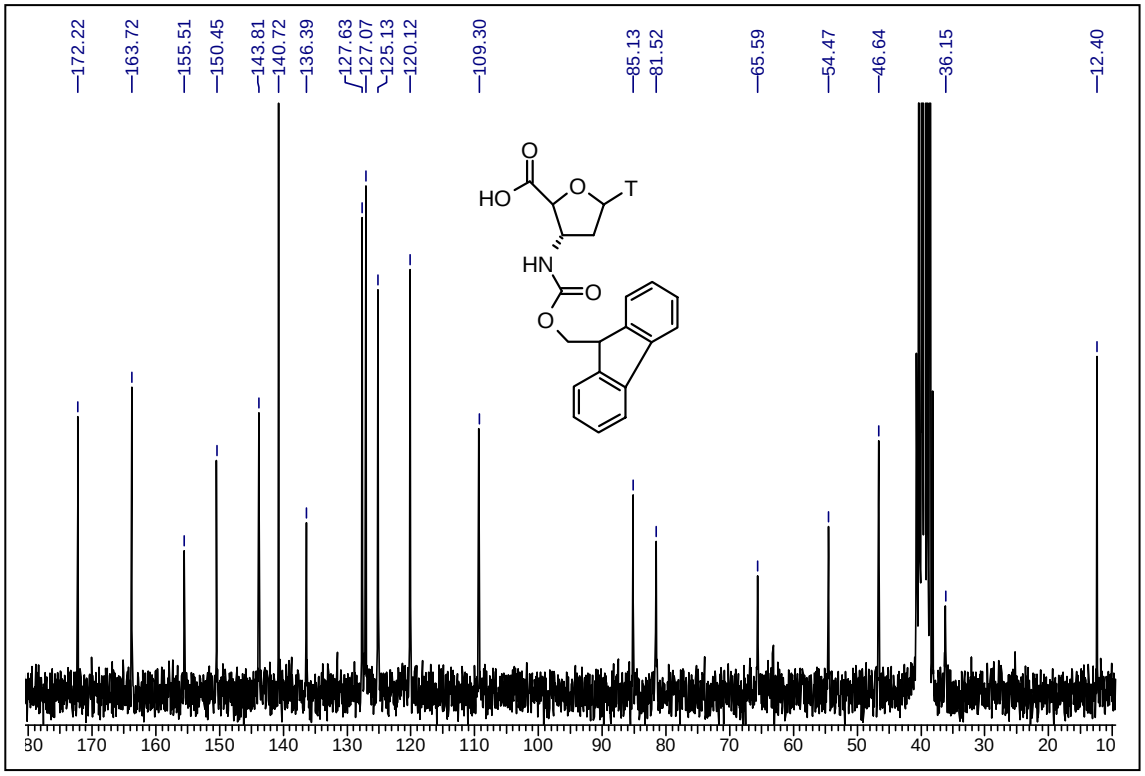


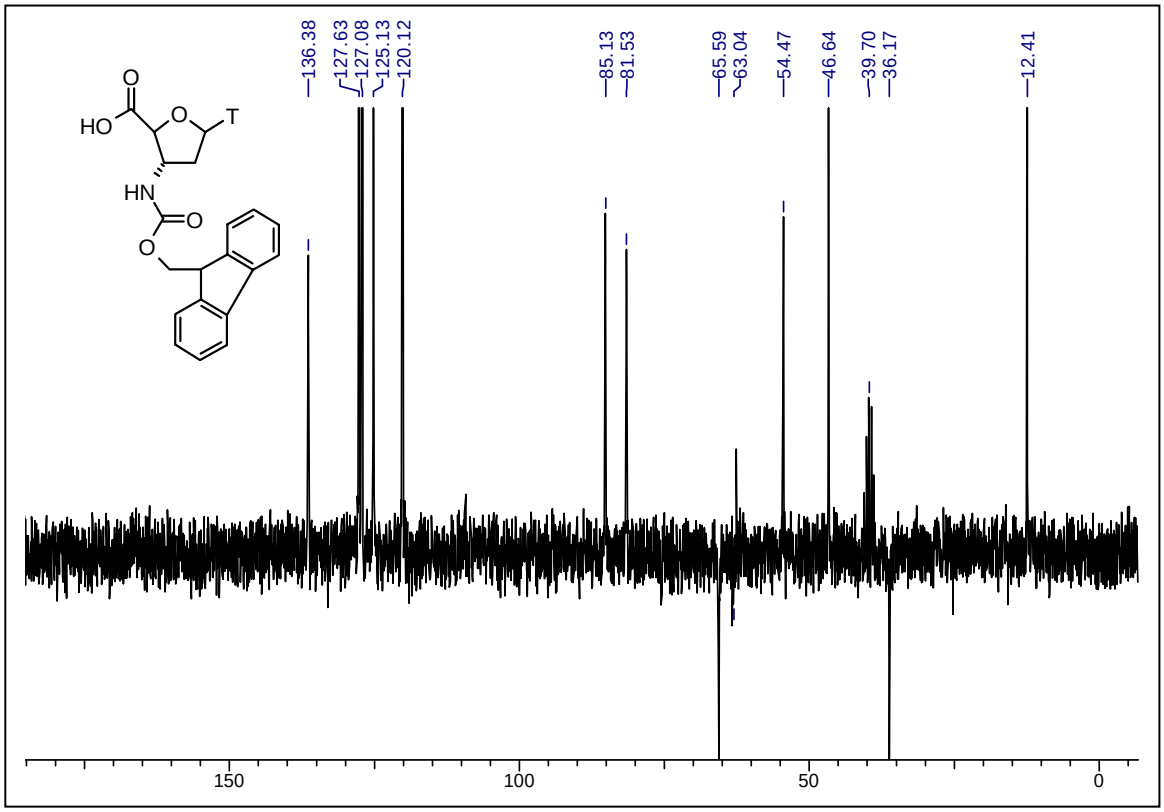
^{1}H NMR of **7**

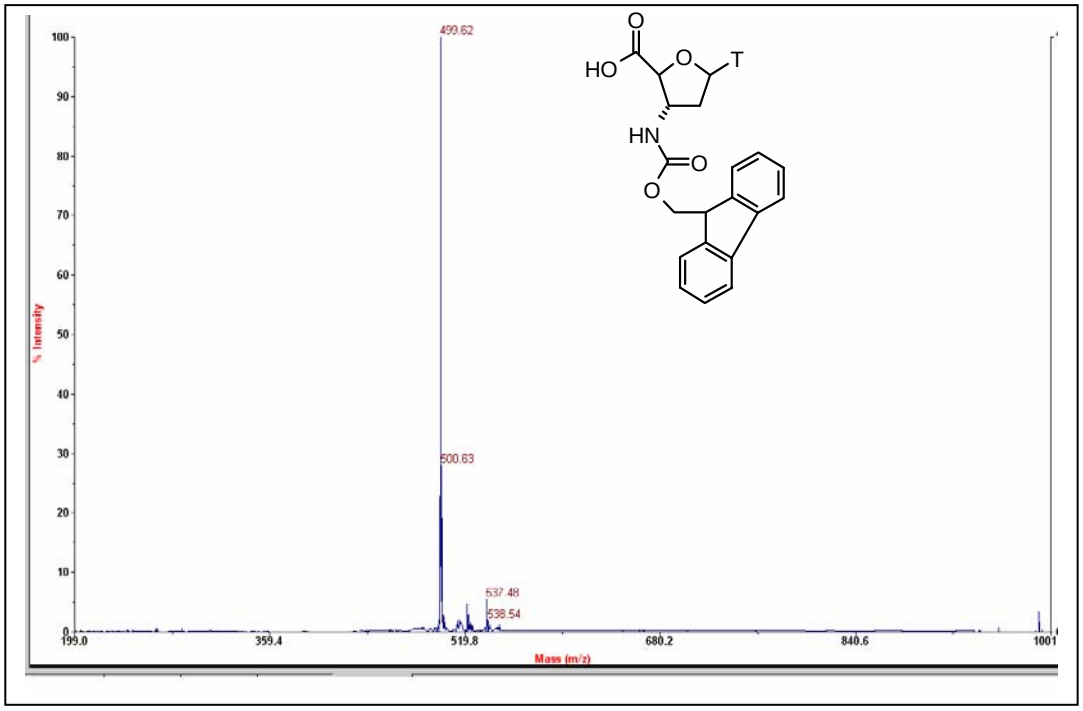
7



^{1}H NMR of **8**


172.22
163.72
155.51
150.45
143.81
140.72
136.39
127.63
127.07
125.13
120.12
109.30
85.13
81.52
65.59
54.47
46.64
36.15
12.40
HO
O
T
HN
O
O
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10


136.38
127.63
127.08
125.13
120.12
85.13
81.53
65.59
63.04
54.47
46.64
39.70
36.17
12.41
HO
O
T
HN
O
O
150
100
50
0


499.62
500.63
537.48
538.54
% Intensity
Mass (m/z)
199.0
359.4
519.8
680.2
840.6
1001
HO
O
T
HN
O
O

Chapter 5

A versatile method for the preparation of conjugates of peptides with DNA/PNA/analogue by employing chemo-selective click reaction in water

5.1 Introduction

Synthetic oligonucleotides (ONs) have interesting therapeutic potentials in various biological applications.[1] Uncharged, achiral peptide nucleic acids (PNAs) are DNA mimics show unprecedented affinity towards complementary RNA and DNA sequences.[2] PNA and other modified oligonucleotides (ONs) [3] are currently being developed as DNA mimics to target disease causing mRNA, [4] using the principle of antisense action. This is gaining further importance because of corrective antisense therapies that do not require activation of RNaseH enzyme or cleavage of the target mRNA.[5] The success or failure of any such candidate in antisense therapeutics depends on a number of factors such as sequence specific recognition of target mRNA, intracellular stability, water solubility [2a] and cell penetration.[6] Application of PNA and analogous uncharged DNA mimics is stymied by the fact that PNAs show very low cell- penetration for any observable antisense effect.[7] Since ONs are poorly taken up by cells, several transfection methods have been developed. However, most of these techniques are not exempt of undesired effects such as toxicity and immunogenicity. About 10 years ago, several small peptides which are able to pass through the plasma membrane were discovered.[8] Interestingly, these peptides, named cell penetrating peptides (CPP) or protein transduction domains (PTD),[9] improve the cellular delivery of cargoes such as peptides and proteins,[8,9] drugs,[10] microorganisms [11] or nucleic acids.[12] For most of the CPP, a covalent bond with the cargo is required to promote transfection and the coupling of CPP to nucleic acids is very challenging because of incompatibilities in their respective chemistry.[13]

Cell-penetrating peptides (CPPs) are defined as peptides with a maximum of 30 amino acids, which are able to enter cells in a seemingly energy-independent manner, thus being able to translocate across membranes in a non-endocytotic fashion [Endocytosis: is a process whereby cells absorb material (molecules such as proteins) from the outside by engulfing it with their cell membrane. It is used by all cells of the body because most substances important to them are large polar molecules and thus cannot pass through the

hydrophobic plasma membrane. The function of endocytosis is the opposite of exocytosis].[14] The first CPP, which was reported in 1994, derives from the third helix of the Antennapedia protein homeodomain[15] and was originally named pAntennapedia (pAntp). Today, this peptide is more commonly referred to as penetratin and together with peptides such as the HIV protein derived transactivating regulatory protein (TAT)[16] and transportan,[17] is one of the most extensively investigated CPP. Examples of sequences of known CPPs are listed in Table 1.

Table 1. Selection of known CPP sequences

Name	Sequence	Reference
Penetratin (pAntp)	RQIKIWFQNRRMKWKK	Derossi *et al.* (1994)[15]
HIV TAT peptide 48–60	GRKKRRQRRRPPQ	Vivés, E., *et al.* (1997)[16]
MAP	KLALKLALKALKAALKLA-amide	Oehlke *et al.* (1998)[18]
Transportan	GWTLNSAGYLLGKINLKALAALAKKIL-amide	Pooga *et al.* (1998)[17]
Transportan 10	AGYLLGKINLKALAALAKKIL-amide	Soomets *et al.* (2000)[19]
R7 peptide	RRRRRRR	Rothbard *et al.* (2000)[10]
pVEC	LLIILRRRIRKQAHAHSK-amide	Elmquist *et al.* (2001)[20]
MPG peptide	GALFLGWLGAAGSTMGAPKKKRKV-amide	Morris *et al.* (1997)[21]
R_6-Penetratin-C-NH_2		Gait *et al.* (2005) [22]
RXR		Moulton *et al.* (2004)[23]

5.1.2 Various methods of synthesis of CPP-ON conjugate

A variety of chemical linkages have been used to link the peptide and oligonucleotide fragments.[24] The choice of a chemical linkage is mainly determined based on its easy synthetic accessibility. In a few biological experiments, the nature of the linkage between

the two fragments is believed to play a role. A study on antisense activities of PMO-peptide conjugates containing chemically variable linkers concludes that the linker length and not the nature of the linkage (type of bond) plays a role in enhancing their antisense effects.[23a] It was also established that the target RNA cleavage efficiencies of PNA-peptide conjugates with two different linkers are strongly dependent on the type of spacer connecting the PNA and the peptide.[25] Thus, the exact role or importance of the nature of chemical linkages is still to be explored.

Two different strategies have been adopted for the synthesis of Peptide-ON conjugates (POCs),[26] they are on-line solid-phase synthesis (divergent method) and fragment conjugation (convergent) methods. In the solid-phase methods, the peptide and oligonucleotide fragments are assembled, sequentially on the same solid support, until the final step. In the fragment conjugation, the peptide and oligonucleotide fragments synthesized individually are linked post synthetically.

The solid-phase divergent method becomes tedious for synthesis of long conjugates. In general, the POCs obtained by the solid-phase method contain not more than 10-15 nucleotide or peptide sequences. The type and length of peptide oligonucleotides required for biological studies are mostly prepared by fragment conjugation. On the other hand, peptides and oligonucleotides of any length could be conjugated using the (Liquid-phase fragment conjugation) LPFC method, provided other conditions, such as solubility, can be varied.

5.1.3 Fragment Conjugation (Convergent Methods)

In the fragment conjugation (FC) method for synthesis of POCs the peptide and oligonucleotide are assembled separately on their own solid supports and they are finally conjugated to each other. A reactive functional group is attached at the desired site of conjugation. The point of conjugation in peptides may be the C- or N-terminus or side chains of the amino acids. In oligonucleotides, the point of conjugation may be the nucleoside, base, or even internucleosidic linkages. If the post synthetic conjugation is done with one of the oligomers still joined to the solid phase, it is called the solid-phase fragment conjugation (SPFC) method. The conjugates are then cleaved, purified, and characterized. Alternatively, if the conjugation is effected post synthetically, after

complete isolation and purification of the peptides and oligonucleotides, it is called fragment conjugation in the liquid phase (LPFC). The general reactions used to prepare some of the functional groups are summarized in Figure1.

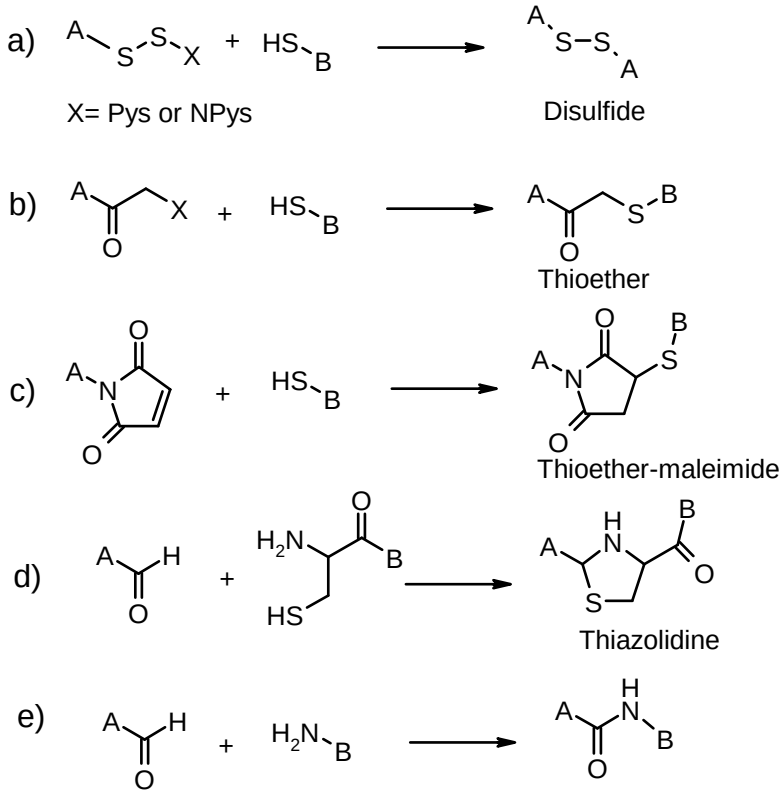


Figure 1. Schematic representation of the most commonly used synthetic protocols to synthesize POCs; A and B can be oligonucleotides or peptides; Py: pyridyl and 4-NO_2-Py: 4-nitropyridyl derivatives.

In fact, most of the POCs required for biological studies have been prepared by this method. As a first step, in the LPFC method, the peptides and oligonucleotides containing protected, but reactive functional groups, are prepared and purified separately. The pure fragments are subsequently conjugated to each other through the reactive functionality.

5.1.4 Maleimide-thiol Protocol

The maleimide-thiol protocol (Figure 1, C) is one of the most commonly used strategies to link the oligonucleotide and peptide analogues. In this strategy, oligonucleotides or peptides containing a maleimide functional group are treated with the peptides or

oligonucleotides containing a thiol group, to yield the corresponding conjugates, linked through a maleimidothioether group.[27]

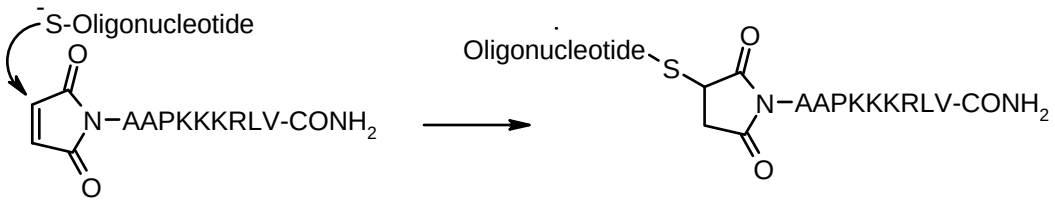


Figure 2. Synthesis of POM by Maleimide-thiol Protocol.

5.1.5 Conjugation through Oxime, Thiazolidine Formation, and Related Reactions

Dumy and co-workers have described a convergent strategy for the synthesis of POCs utilizing chemoselective ligation of peptide to oligonucleotide through oxime and thiazolidine formation (Figure 3).[28] The conjugation through oxime formation was performed by treating an oxyamine containing oligonucleotide or peptide with an aldehyde containing peptide or oligonucleotide (Figure 3a and 3b).

Ligation by thiazolidine formation was achieved by treating a peptide containing a cysteine residue with an oligonucleotide containing an aldehyde function (Figure 3c).

Gait and co-workers have synthesized a series of 2'-deoxyoligonucleotides and 2'-OMe RNAs carrying one or more 2'-aldehyde groups (Figure 4a).[29] These oligonucleotides were successfully coupled to one or more peptides containing an N-terminal cysteine, aminooxy, or hydrazide group to give corresponding peptide-oligonucleotide conjugates in good yields (Figure 4b). This facile conjugation method allows specific coupling of unprotected oligonucleotides containing aldehyde groups to unprotected N-terminally modified peptides and other small molecules, in aqueous conditions. Using this strategy, a 12-mer, 2'-OMe RNA, complementary to the HIV-1 TAR RNA stem-loop was conjugated to copies of an 8-mer model laminin peptide.[29a]

The TAR RNA binding was not affected much upon conjugation. But, the *in vitro* HIV-1 inhibition activity of the conjugate is similar to that of the unconjugated 2'-OMe RNA.

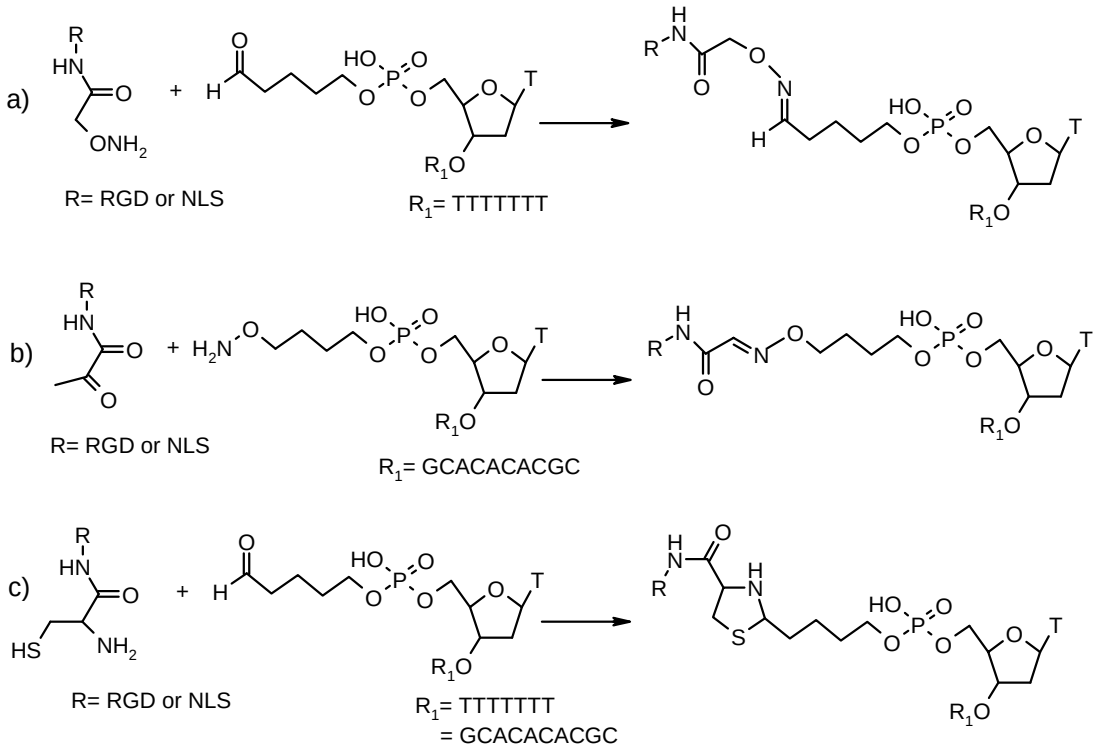


Figure 3. Synthesis of POC by Oxime, Thiazolidine Formation, and Related Reactions

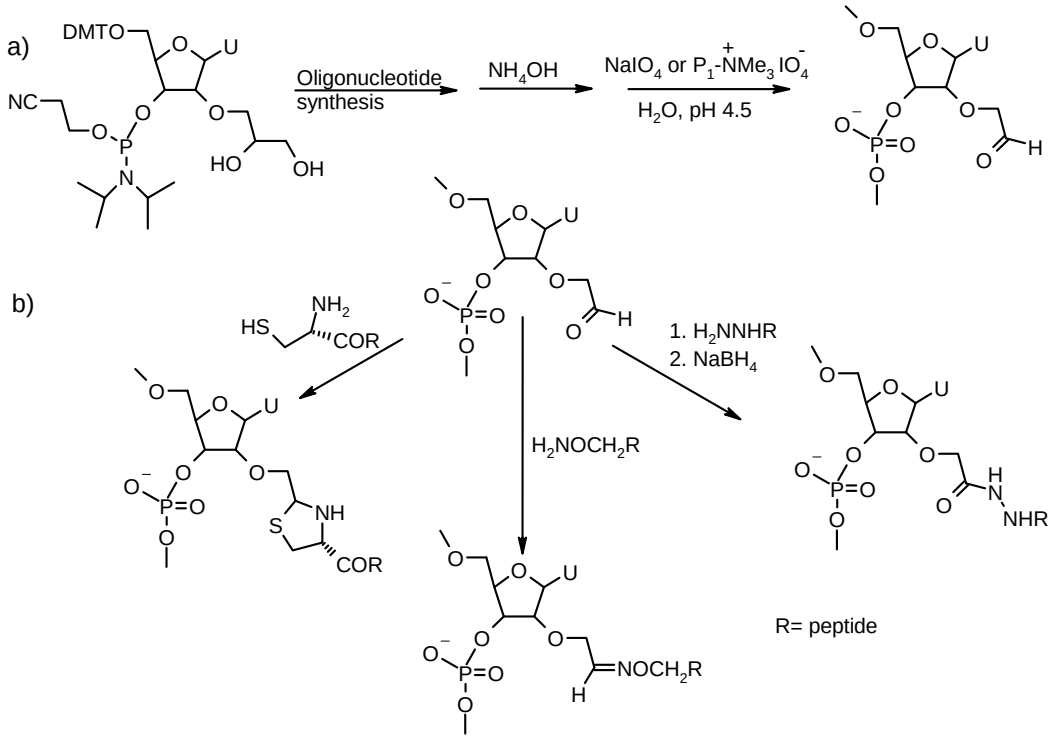


Figure 4. Synthesis of POC via 2'*O* - Oxime, Thiazolidine Formation.

5.1.6 Conjugation through a Thioether Linkage

The oligonucleotide thiols have been successfully conjugated to N-terminal-R-bromoacetyl peptides in a single step coupling reaction. The oligonucleotide and peptide fragments of the resulting conjugates are linked through a thioether bond. For example, the 3'-amino and 5'-thiol modified dodecanucleotide is treated with the (bromoacetyl) pentapeptide to yield the corresponding conjugates in good yields (Figure 5).[30]

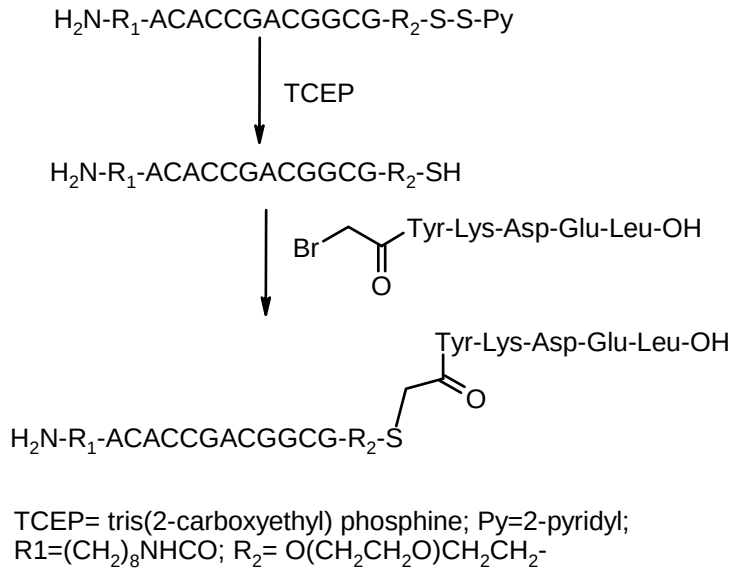


Figure 5. POC via the formation of thioether linkage

5.1.7 Conjugation through a Disulfide Linkage

In addition to thioether and thioester linkages, the disulfide linkage is also more frequently used in oligonucleotide peptide conjugation chemistry.[31]

In most cases, the peptide thiol is prepared by introduction of the cysteine residue at the

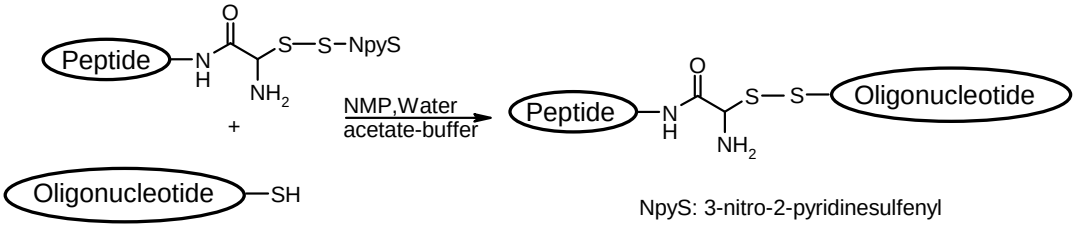


Figure 6. POC via the formation of disulphide linkage

C-terminus. The oligonucleotide thiols are obtained by attaching an O-protected mercaptoproponal phosphoramidite building block followed by reducing with DTT. The peptide thiols are activated by treatment with 2, 2'-dipyridyldisulfide (to form PyS derivatives) or 3-nitro-2, 2'-dipyridyldisulfide (to form NpyS) in TEAA buffer, in ethanol solvent. The coupling is effected using the disulfide reducing agent, DTT, in TEAA buffer.[32]

Peptide-PNA conjugates can also be synthesized by using the same route.[33] In a typical method, PNA (1 equiv) and NpyS-peptide (1.3 equiv) were mixed in an NMP-H_2O solvent system, in the presence of sodium acetate buffer (pH 5.0). The solution was vortexed and incubated at 40 °C for 3 h in the dark. The reaction was quenched by addition of 5% degassed aqueous TFA, and the resulting conjugate was purified by high-performance liquid chromatography. As can be seen in the applications section, the disulfide linkage is one of those most commonly used for preparation of POCs of biological significance.

5.1.8 Peptide oligonucleotide conjugation through Diels-Alder Reactions

The Diels-Alder reaction between diene-modified oligonucleotides and maleimide-

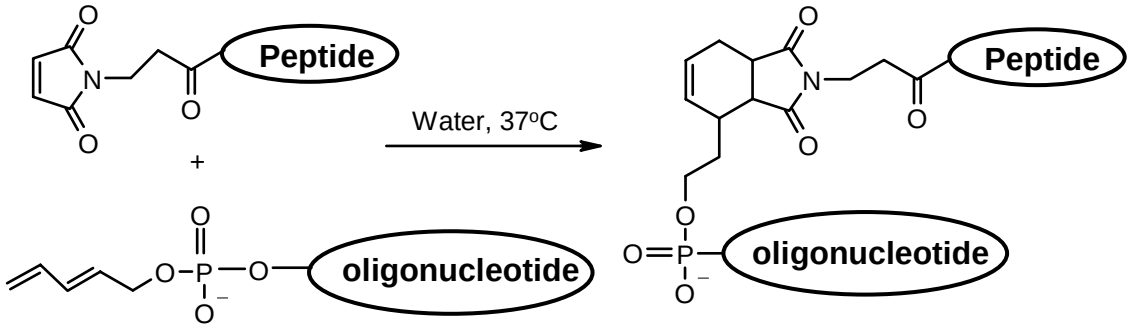


Figure 7. Preparation of peptide–oligonucleotide conjugates by Diels-Alder reaction in water.

derivatized peptides afforded peptide-oligonucleotide conjugates with high purity and yield.[34] Synthesis of the reagents was easily accomplished by on-column derivatization of the corresponding peptides and oligonucleotides. The cycloaddition reaction was carried out in mild conditions, in aqueous solution at 37ºC. The speed of the reaction was found

to vary depending on the size of the reagents, but it can be completed in 8-10 h by reacting the diene-oligonucleotide with a small excess of maleimide-peptide.

5.1.9 Preparation of Peptide Conjugates of PNA and PMO by Convergent Methods

Most of the fragment conjugation (FC) approaches described in the above section extended for preparation of conjugates of modified oligonucleotide analogues as well. Most of the PNA-peptide conjugates reported so far have been prepared by liquid-phase fragment conjugation methods. The frequently used strategy involves conjugation of a Cys-peptide to a Cys-PNA by means of disulfide-bridge. Alternatively, treatment of a Cys-peptide with a maleimide derivatized PNA oligomer yields a PNA-peptide conjugate with a thioether bond.[35] The key aspects of the synthesis and biological activities of conjugates of PMOs with synthetic peptides have recently been reviewed.[31b]

5.1.10 SiRNA-Peptide Conjugates for Gene Silencing

RNA interference (RNAi) is found to be a powerful biological process for specific silencing of gene expression in diversified eukaryotic cells.[36] In this process, a small synthetic RNA fragment (of about 21-23 nucleotides) blocks the expression of target mRNA by binding with its complementary sequences of target mRNA. These RNA fragments are called siRNAs. RNAi has tremendous potential for functional genomics, drug discovery through *in vivo* target validation, and development of novel gene-specific medicines. However, similar to the case of DNAs, the therapeutic value of siRNAs is hindered by poor cellular uptake, limited stability in blood, and nonspecific immune stimulation. To address these problems, different strategies, including linking the siRNAs to ligand-specific, sterically stabilized nanoparticles, have been actively studied.

Similar to its DNA counterparts, RNAs are also shown to exhibit enhanced cellular uptake upon conjugation to peptides. Recently, Scaria and co-workers have shown that functional siRNA can be delivered into cells by using RGD peptide derived nanoparticles.[37] These self assembled nanoparticles were constructed with PEG-PEI conjugated to an RGD peptide. The RGD peptide ligand Arg-Gly-Asp is attached at the distal end of the PEG, to target the tumor cells. The nanoparticles of different size form complexes, called nanoplexes, with the siRNAs to be delivered.

The cellular delivery and RNAi activity of siRNAs upon conjugation to Tat peptide (Tat 47-57) and Tat derived oligocarbamate has been studied.[38] A series of covalent conjugates of Tat and Tat derived oligocarbamates with the siRNAs were prepared using the maleimido-cysteine protocol in the liquid phase.

The membrane permeant peptides (MPPs), namely, penetratin and transportan, have also been conjugated to improve cellular uptake of siRNAs.[39] The peptides were linked to siRNA by means of easily reducible disulfide bonds. The single stranded RNAs were annealed together to form a duplex siRNA, and then, the duplexes were mixed with MPP in equimolar amounts to get the corresponding conjugate. The conjugation of MPPs is shown to be a better and relatively inexpensive method for effective delivery of siRNAs to different cell types. In addition, MPPs may complement the use of transfection with cationic liposomes. The gene inhibition using MPP-siRNAs was highly specific.

5.2 Rational and Objective of the present work

The literature cited above gives an overall account for the synthesis of conjugates of oligonucleotides and ON analogues eg., peptide nucleic acids (PNA).A recent review[24] also summarizes the present methods of synthesizing oligonucleotide- peptide conjugates that are common to ONs and their analogues and points out the need to develop straightforward methods to synthesize such conjugates. The highly functionalized nature of these biomolecules and their mimics such as PNA, render them susceptible for side reactions during conjugation employing most of these methods and yield and purity of structurally defined conjugated biomolecules is often low. The (4+2) Diels-Alder cycloaddition approach was employed recently for the conjugation of DNA and CPP.[34] This involved the reaction between diene and dienophile present on the respective biomolecules to get the conjugates. The maleimide dienophile used in this reaction is susceptible for Michael addition reactions with other nucleophilic centers on peptides or ONs and may give rise to side reactions. The current literature clearly indicates the need for a simple and straightforward strategy for generating highly pure ON/PNA-peptide conjugates in high yield.[24]

Meanwhile, we find that the applications of highly selective orthogonal Cu (I) catalyzed Huisgen 1, 3 dipolar cycloadddition reaction (more commonly called, Cu catalyzed azide

alkyne cycloaddition reactions), recognized as 'click chemistry'.[40] Sharpless and his coworkers have described this reaction as ''click chemistry'' and ones like it that proceed in very high yields and that almost invariably work. There are, in fact, too few transformations that fit these criteria in organic chemistry! Click reactions in general have been reviewed,[40d,e] and there is one review specifically on copper-catalyzed azide–alkyne cycloadditions.[41] This Cu catalyzed azide-alkyne cycloaddition reaction is expanding the scope of synthesis of variety of other bioconjugates such as DNA-glycoconjugates,[42] peptide-protein conjugates,[43] carbohydrate-vaccine conjugates,[44] or protein modification and protein microarray fabrication.[45] This reaction between high-energy organic azides and terminal alkynes can give rise to unlimited array of inert triazole containing architectures (Figure 8). The reaction is highly predictable, fast and resistant to side reactions. Addition of Cu (I) accelerates the reaction.[40d,e] Some of these reactions could be carried out in aqueous medium and can be employed post-synthetically on purified units decorated with a variety of functional groups, without additional functional group protecting strategies. Recently, click reaction has also found applications for the synthesis of circular DNA, [46, 47] and DNA-template directed oligonucleotide strand ligation.[47]

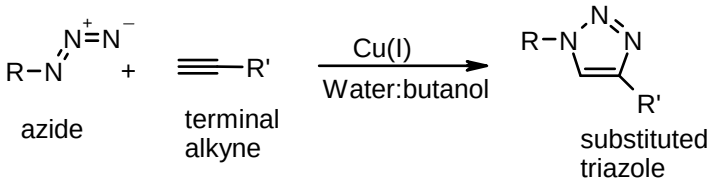


Figure 8. Cu catalyzed Huisgen [3+2] cycloaddition reaction between azide and terminal alkyne groups to give inert triazoles

The mechanistic interpretation (Figure 9) is supported by kinetic studies,[48] product distributions for specialized substrates, and DFT calculations.[49]

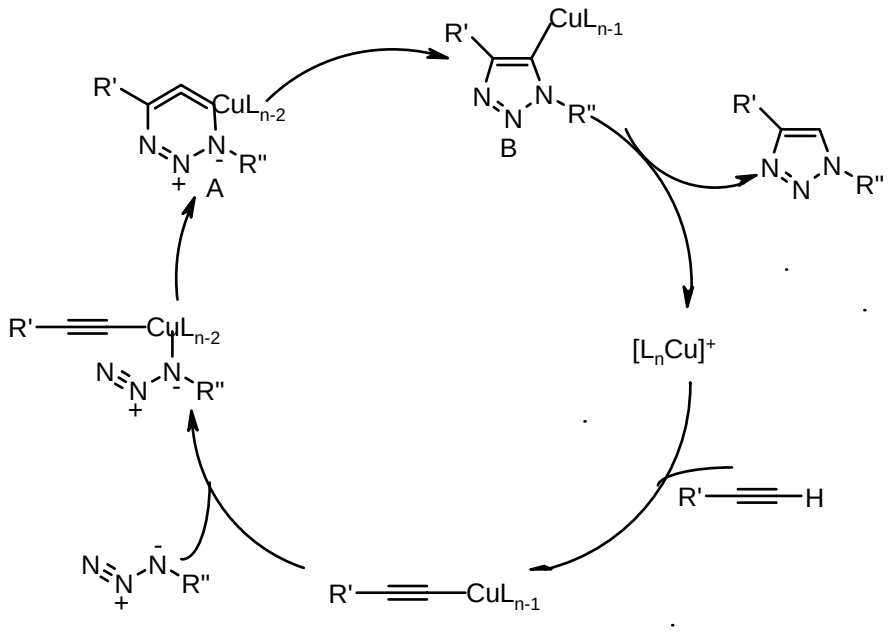


Figure 9. Putative mechanism of the copper-mediated click reaction of alkynes and azides from DFT calculations.

The objectives in this work are

i) Synthesis of azide and alkyne functionalized PNA/ DNA and peptide oligomers.

ii) Specific Cu catalyzed azide alkyne cycloaddition reaction between azide and alkyne functionalized PNA/DNA and peptide oligomers.

iii) Purification and characterization of the conjugates.

5.3 Results and Discussion

The aim of this study was to develop a simple and chemo-selective method for the conjugation of peptides and oligonucleotide mimics such as PNA and TANA (Figure 10). The azide and acetylene carrier units were chosen so that these could be installed online during the solid supported peptide or PNA/TANA synthesis at the *N*/*C*-terminus through an amide linkage that is stable towards conditions employed for the deprotection and cleavage of the peptide/ON from the solid support.

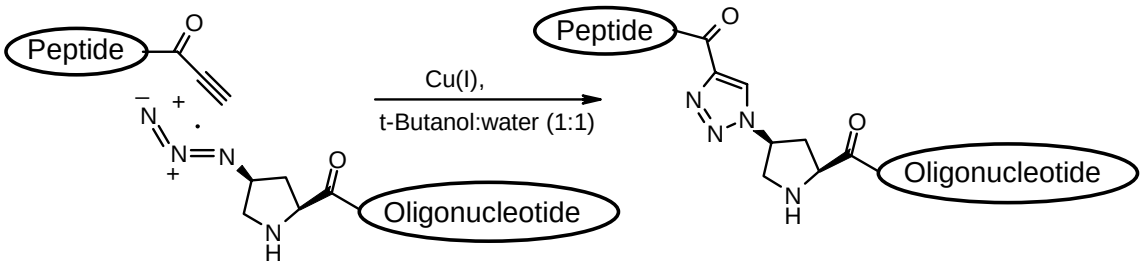


Figure 10. Synthesis of peptide-PNA conjugates by click chemistry

The candidates chosen for this purpose were propynoic acid and (2*S*, 4*S*) - 4-azidoproline (Figure 11).[50] DNA ON was functionalized at 5'-end using 5'-*O*-propynyl-thymidine-phosphoramidite (Figure11).

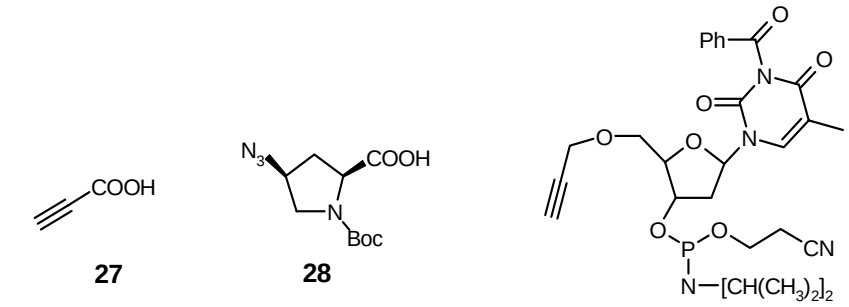


Figure 11. Alkyne and azide carrier units

5.3.1 Synthesis of azide functionalized PNA (/TANA) oligomers and alkyne functionalized peptide oligomers

The 4-azidoproline was chosen in this strategy, as it could open up the possibility to introduce the azide group at either *C/N*-terminus or in the center of the PNA or peptide sequences as desired. This could enable the conjugation of peptide at either *C*- or *N*-terminus of PNA or peptide. To assess the methodology and its application potential, we choose three PNA sequences, a homopyrimidine PNA-T8, a mixed purine-pyrimidine 10 mer PNA sequence and the recently developed homothyminyl-oligomer with uncharged thioacetamido nucleic acid (TANA) backbone. All these sequences were synthesized on solid-phase using both Boc and Fmoc strategy.

The azide functionalized PNA (TANA) oligomers **3** and **5** were synthesized on Rink-amide resin following the standard procedures [51] of Fmoc based solid-phase peptide synthesis. The last coupling reaction was done using *N*-Boc-(2*S*, 4*S*)-4-azidoproline monomer unit. The azide functionalized mixed purine-pyrimidine PNA oligomer **4** was synthesized following the standard procedures of Boc-based solid-phase peptide synthesis.[51] The final coupling reaction was done with *N*-Boc-(2*S*,4*S*)-4-azidoproline. Similarly the alkyne functionalized lysine peptide **1** and arginine-rich peptide **2,** were synthesized on Rink-amide resin following the standard procedures of solid-phase peptide synthesis. The final coupling was done with propynoic acid.

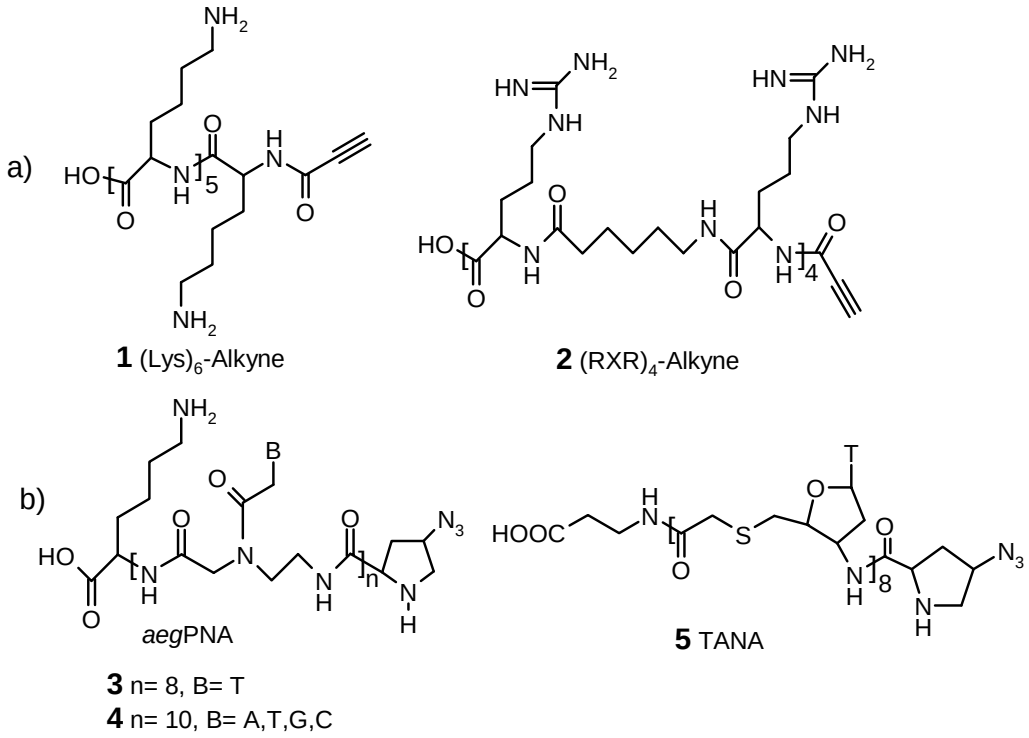
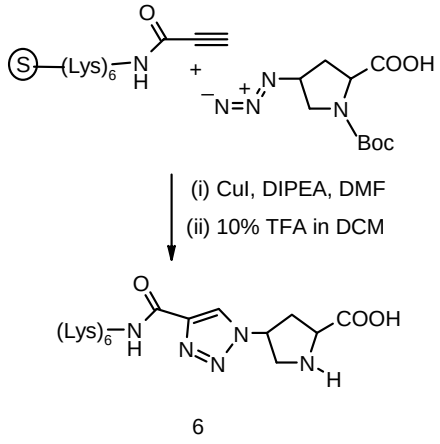


Figure 12. a) peptide alkyne sequences and b) *aeg*PNA and TANA azide sequences used for click reaction.

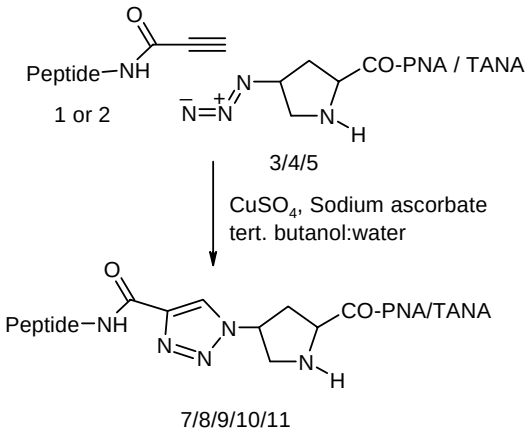
The installation of reactive alkyne and azide functionalities on either the peptide or ONs, was followed by the experiments to study the feasibility of the click cycloaddition reaction. The peptide **1**, attached to the solid support with all the side chain lysine amino- and terminal carboxy- groups protected, was treated with 3 equivalents of *N*-Boc(2*S*,4*S*)-

4-azidoproline in DMF in the presence of CuI. The product **6** was isolated after cleavage from the support. There was negligible difference in RP-HPLC t_R but mass showed conjugation of proline unit through the triazole linkage (Scheme 1, Table 1, entry 2).



Scheme 1. Peptide-alkyne conjugation with 4-azidoproline using click chemistry on solid support

After the formation of the desired product was established, the solution phase reactions of PNAs with *N*-terminal azidoproline (**3** and **4**) or TANA (**5**) and *N*-terminal-alkyne



Scheme 2. Peptide-alkyne conjugation with 4-azidoprolyl ON using click chemistry in solution

substituted lysine peptide **1** and arginine rich peptide **2,** were carried out in *tert.* butanol:water medium in presence of $CuSO_4$ and sodium ascorbate as a catalyst (Scheme 2). The CuI catalyst was found to be less efficient as compared to $CuSO_4$-sodium ascorbate under these conditions. The reactions were followed by HPLC (Figure 13). The

reaction took longer time for PNA, TANA substrates (16-32h) to completion with approximately 1 equivalent of each substrate. With approximately 3 equivalent excess of **1** or **2** complete conversion to the conjugated products (**6**-**8** and **10, 11**) was observed within 2h (Table 2).

Table 2 The RP-HPLC-t_R[a] and MALDI-TOF mass characterization and purity found by HPLC of the peptide, PNA and TANA sequences and peptide-PNA and peptide-TANA conjugates

	Compounds	HPLC[a]	Purity %	Mass	
				calcd	found
1	HO-$(Lys)_6$-alkyne **1**	9.7	91	839.08	837.8
2	HO-$(Arg\text{-}Aha\text{-}Arg)_4$-alkyne **2**	12.9	100	1772.1	1772.38
3	HO-β-ala-TTTTTTTT-Pro-N_3 **3**	12.6	99.1	2356.2	2356.6
4	HO-Lys-TCACTAGATG- Pro-N_3 **4**	11.7	98.5	2991.2	2990.8
5	HO-β-ala-tttttttt-Pro-N_3 **5**	18.5	100	2605.3	2628($+Na^+$)
6	HO-$(Lys)_6$-triazole-Pro **6**	9.59	90	995.23	996.5
7	HO-β-ala-TTTTTTTT-Pro-triazole$(Lys)_6$-OH **7**	12.29	100	3195.3	3195.35
8	HO-Lys-TCACTAGATG- Pro- triazole-$(Lys)_6$-OH **8**	11.08	98	3830.3	3832.5
9	HO-β-ala-tttttttt-Pro-triazole-$(Lys)_6$ **9**	17.2	100	3444.9	3444.0
10	HO-Lys-TCACTAGATG- Pro- triazole-$(Arg\text{-}Aha\text{-}Arg)_4$-OH **10**	12.7	95.4	4763.3	4763.8
11	HO-β-ala-tttttttt-pro-triazole-$(Arg\text{-}Aha\text{-}Arg)_4$ -OH **11**	17.2	89	4378.6	4376.29

[a] t_R-RP-HPLC retention time in minutes. T, C, G, A denote *aeg*PNA units and t denotes TANA unit, Pro-N_3 denotes (2*S*, 4S)-4-azidoproline unit, alkyne denotes acetylene carboxylic acid, Aha denotes ε-aminohexanoic acid.

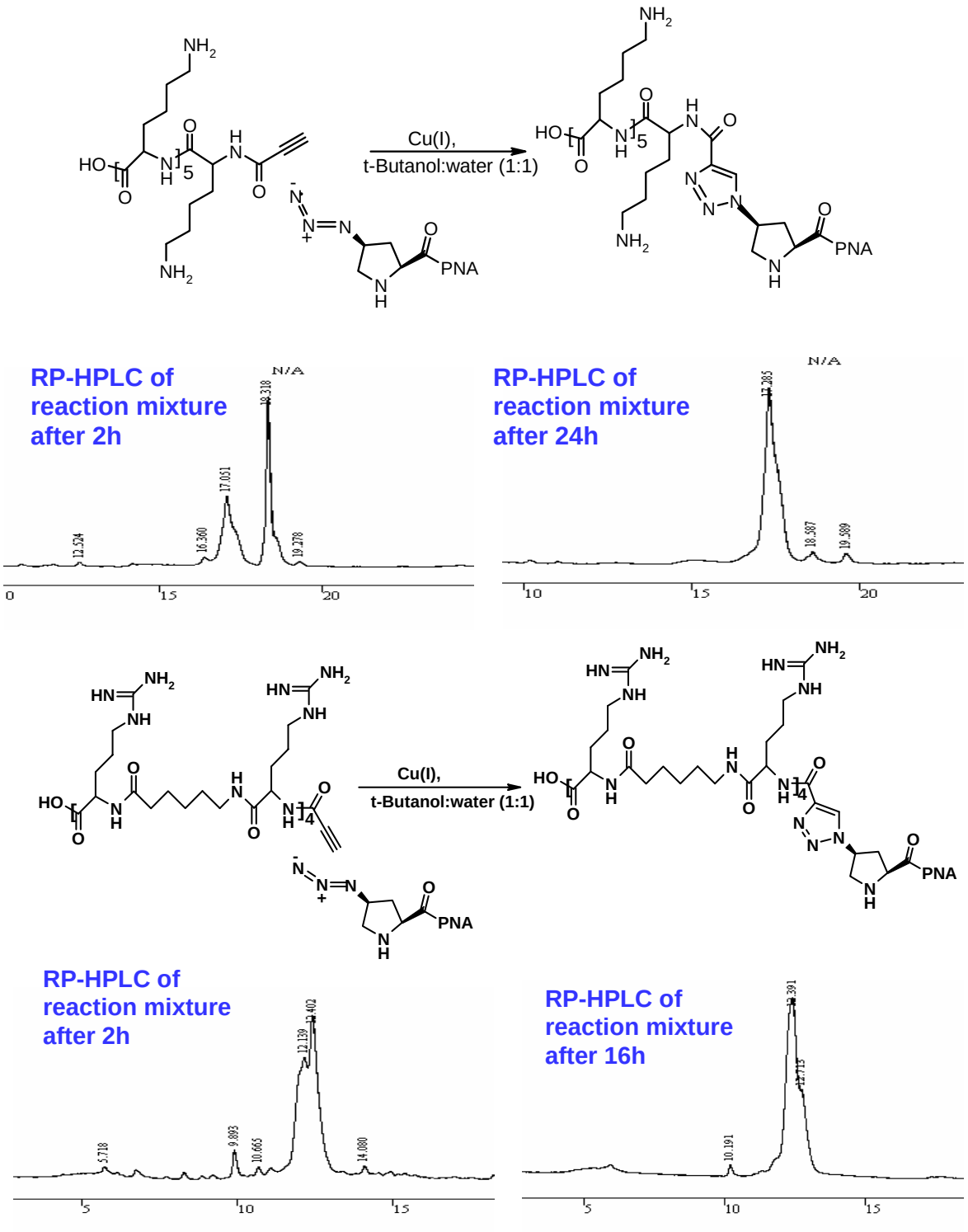
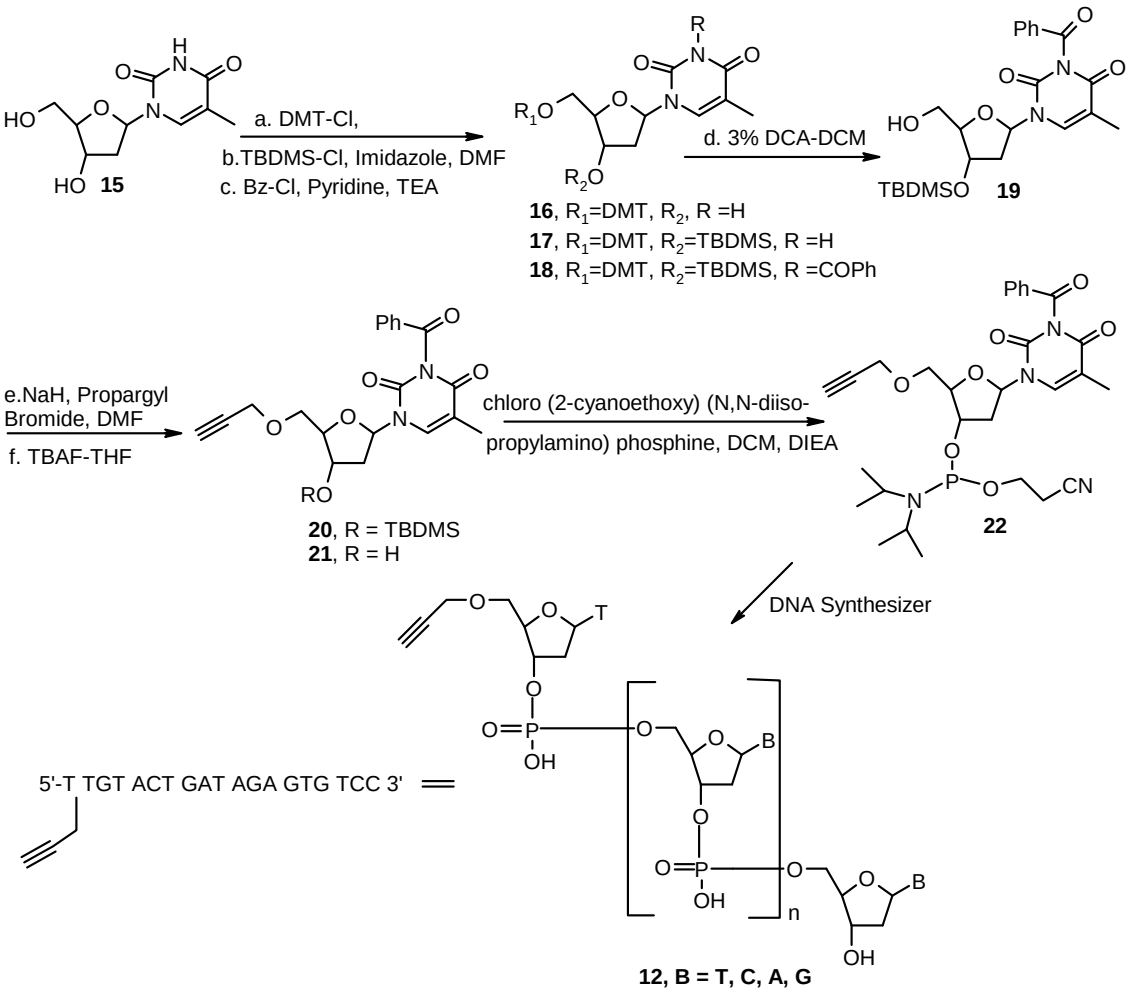


Figure 13. Progress of peptide- ON conjugation monitored by RP-HPLC.

5.3.2 Synthesis of alkyne functionalized DNA oligomers

The DNA sequence **12** was synthesized using commercially available monomeric units using phosphoramidite chemistry on automated DNA synthesizer. The last coupling was done using 5'-*O*-(propynyl-*N*-*3*-benzoyl-thymidin)-3'-*O*-(*N*,*N*-diisopropylamino-*O*-cyanoethyl -phosphoramidite) **22**. The synthesis of the 5'-alkyne modified phosphoramidite building block **22** was synthesized as shown in Scheme 3.

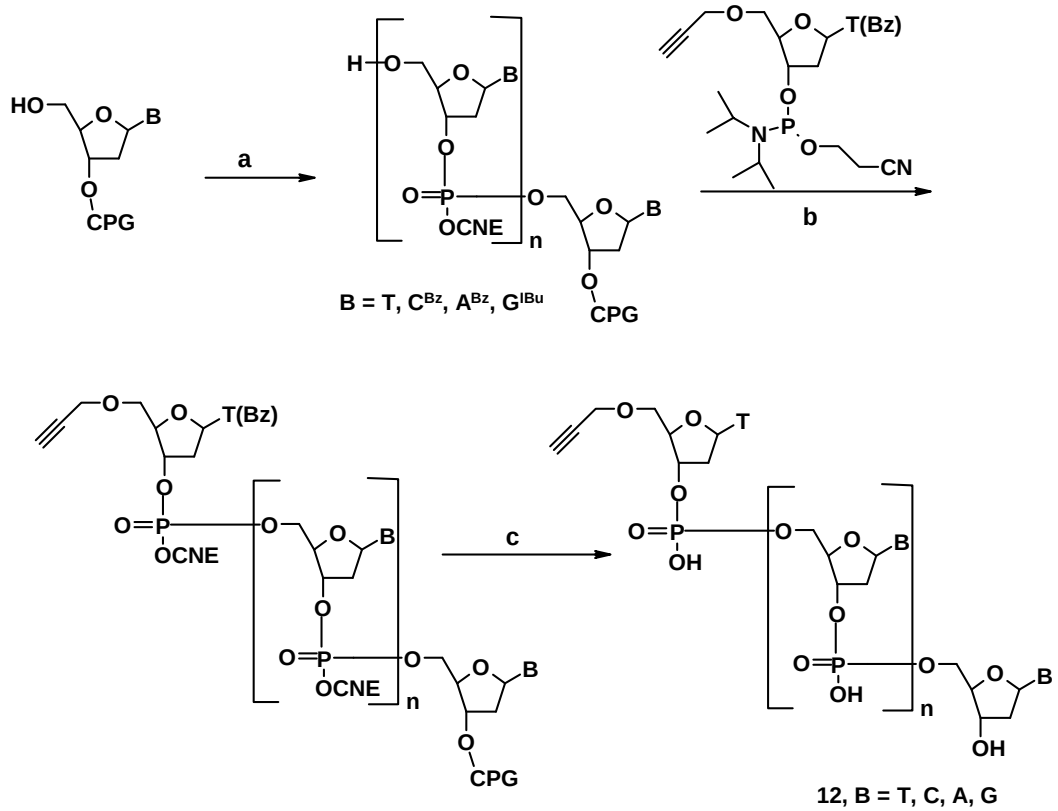


Scheme 3. Syntheis of 5'-alkyne modified phosphoramidite building block **22**

The 5'-OH of thymidine **15** was protected as the dimethoxy trityl ether **16** and 3'-OH was protected as it TBDMS ether **17**. Benzoylation of the N3- position using benzoyl chloride in pyridine to give the fully protected derivative **18**. Then deprotection of the DMTr ether using 3% dichloroacetic acid in DCM and triethyl silane as the scavenger to give

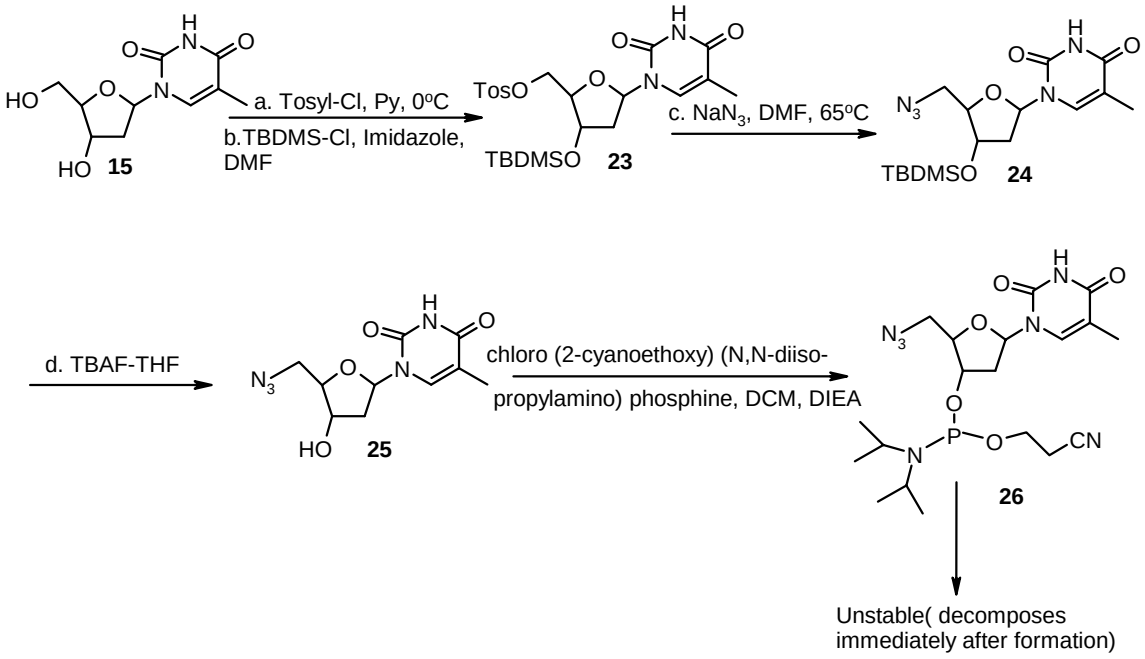
the 5'OH compound **19**. The alkyne functionality was introduced into the 5'- position by 5'O-alkylation using NaH and propargaryl bromide to give **20**. The formation of the product was confirmed by appearance of peak at δ 2.5 ppm due to the alkyne proton in the ^{1}H NMR spectra. Deprotection of the 3' OTBDMS, **21** and subsequent phosphitylation to give the phosphoramidite building block **22**, which was incorporated into the terminal end of the 19 mer DNA sequence as shown in Scheme 4.

To install the azide functional group on DNA oligomers, we attempted to synthesize 5'deoxy-5'-azido-thymidinyl-3'-*O*-phosphoramidite **26**, but this was found to be quite unstable (Scheme 5). Synthesis of the 5'-azide functionalized thymidine derivative was done starting from thymidine. The 5'-OH was selectively converted to it corresponding tosylate by reaction with tosyl chloride in pyridine by maintaining ice cold temperature



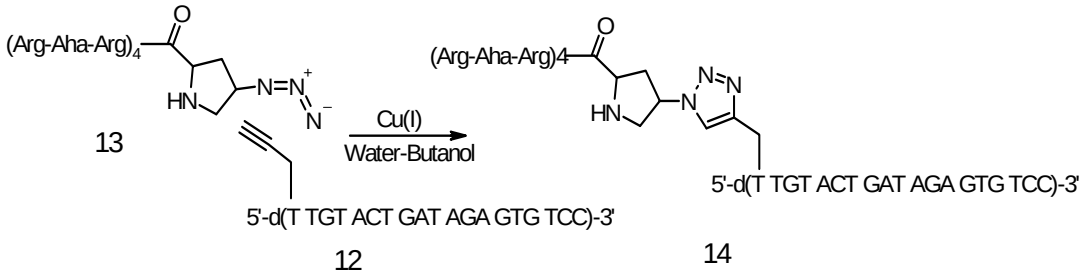
Scheme 4. Solid-Phase synthesis of alkyne functionalized oligonucleotide **12** a) oligonucleotide assembly using standard 3'- phosphoramidite nucleoside derivatives, b) coupling of N-benzoyl 5' [(prop-2-ynyloxy) methyl] 3'-*O*-(2-cyanoethyl-N, N-diisopropylphosphoramidite) Thymidine **23,** followed by capping and oxidation (I_2) c) final deprotection with conc. aq. NH_3.

and high dilution. The 3'-OH was protected as 3'-OTBDMS ether. Then the azide functionality was introduced into the 5' -position by nucleophilic substitution of the tosylate by sodium azide in DMF. The deprotection of the 3'-OTBDMS ether followed by phosphitylation to give the corresponding phosphoramidite. But the phosphoramidite formed was found to be quite unstable. There was immediate generation of five to six closely moving spots on TLC after work up of the reaction. The newly generated spots could not be isolated for characterization. The generation of these new spots may be the due to the intramolecular interaction between the 5'-azide and the 3'-trivalent phosphorous which leads to the Aza- Wittig type of reaction.



Scheme 5. Attempt to synthesize of 5'-azide modified phosphoramidite building block **26**

The DNA substrate **12** was converted to the peptide-DNA conjugate **14** within 1h with 3 equivalents of the peptide (Scheme 6). Cu (I) catalysis is known to degrade DNA due to the presence of hydroxy radicals and addition of Cu (I) stabilizing ligands such as polytriazoles is often required to circumvent the DNA degradation.[47] The arginine peptide used in this reaction could possibly have contributed to the stabilization of Cu



Scheme 6. Peptide-azide conjugation with 5'-*O* alkyne ON using click chemistry in solution

Table 3. The RP-HPLC-t_R[a] and MALDI-TOF mass characterization and purity found by HPLC of the peptide and ON sequences and peptide-ON conjugates

	Compounds	HPLCt t_R[a]	Purity %	Mass	
				calcd	found
1	5'-alkyne d(TTGTACTGATAGAGTGTCC)3' **12**	8.35	95	5878.04	5878.01
2	HO- (Arg-Aha-Arg)$_4$- Pro-N_3 **13**	13.8	95	1858.3	1881.2(+Na^+) 1897.2(+K^+)
3	HO- (Arg-Ahx-Arg)$_4$- pro- triazole-5'-d(TTGTACTGATAGAGTGTCC)-3' **14**	11.04	81.6	7736.3	7735.09

[a] t_R-RP-HPLC retention time in minutes. T, C, G, A denote ON unit, Pro-N_3 denotes (2*S*, 4S)-4-azidoproline unit, alkyne denotes 5'- O alkyne unit, Aha denotes ε-aminohexanoic acid.

(I)[52] similar to the Cu (I) binding ligands. Additionally, presence of *tert.* butanol as solvent in this reaction might stabilize DNA as *tert*. butanol is known as a scavenger of hydroxy radicals.[47] The yields of peptide-DNA conjugates may be further improved by the use of Cu (I) stabilizing ligands such as polytriazoles[47] along with *tert*. butanol as hydroxy radical scavenger.[53] The results are tabulated in Table 3.

The peptide-ON conjugates with either the DNA (**14**), *aeg*PNA backbone (**7, 8** and **10**) or TANA backbone (**9, 11**) were obtained in high purity using this simple cycloaddition reaction. The reaction seems to be very specific without interfering with other functional groups present, either on peptide or DNA/PNA/TANA segments of the conjugates.

5.4 Conclusions

The peptide-DNA/PNA/TANA conjugation chemistry presented will prove to be a unique solution for CPP conjugation to antisense/siRNA therapeutics. In this fast developing field, application is mainly restrained by the problems faced in cellular delivery of modified oligonucleotides. The ease of conjugation and compatibility of the strategy with various substrates and functional groups will prove to be an attractive alternative to present methods for the synthesis of peptide-PNA conjugates. It is sequence-independent with respect to either peptide or PNA/DNA or other backbone modified DNA mimics (such as TANA), which are being developed for antisense applications. The reactive azide and alkyne groups can be installed on either ON or peptide and the work presented, has a potential for general application in nucleic acid-peptide conjugation. This could be a method of choice when a peptide of choice needs to be conjugated with several ON sequences or different modifications of the same sequence for studying their bioactivity. Alternatively, the chosen ON may be conjugated with several peptides to establish the biological efficacy of different peptides. Some recent work was carried out to evaluate the oligonucleotide-peptide conjugates as a potential alternative to the use of transfection agents to improve the efficacy of siRNA.[54] It will be very interesting to study the effects of the chemistry of this new linkage on the intracellular performance of the conjugates.[4b]

Summary

- Syntheis of azide functionalized PNA/TANA oligomers and alkyne functionalized peptide oligomers were done by solid-phase peptide synthesis.
- The click cycloaddition was performed in water: *tert*.Butanol as the solvent system and CuSO4 and sodium ascorbate as the catalyst.
- The progress of the reactions was monitored by RP-HPLC and the conjugates formed were characterized by MALDI-TOF mass spectrometry.
- Synthesis of 5'-*O*-alkyne functionalized monomer phosphoramidite was done and it was incorporated into the DNA sequence. The synthesis of the azide functionalized Peptide oligomer was done by solid-phase peptide synthesis.
- The Peptide-DNA conjugate formed was characterized by MALDI-TOF mass spectrometry.

5.5 References

1. Uhlmann, E. and Peymann A. (1990) Antisense oligonucleotides: A new therapeutic principle. *Chemical Reviews*, **90**, 544-584.
2. (a) Nielsen, P. E., Egholm, M., Berg, R. H. and Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **254**, 1497-1500., (b) Braasch, D. A. and Corey, D. R. (2002) Novel antisense and peptide nucleic acid strategies for controlling gene expression. *Biochemistry*, **41**, 4503-45102.
3. (a) Uhlmann, E., Peymann, A., Breipohl, G. and Will, D. W. (1998) PNA: Synthetic polyamide nucleic acids with unusual binding properties. *Angew. Chem. Int. Ed. Engl.*, **37**, 2796-2823., (b) Kurreck, J. (2003) Antisense technologies: Improvement through novel chemical modifications. *Eur. J. Biochem.*, **270**, 1628-1644. (c) Micklefield, J. (2001) Backbone modification of nucleic acids: Synthesis, structure and therapeutic applications. *Curr. Med. Chem.*, **8**, 1157-1179. (d) Kumar, V. A. (2002) Structural preorganization of peptide nucleic acids: Chiral cationic analogues with five- or six-membered ring structures. *Eur. J. Org. Chem.*, 2021-2032 (e) Kumar V. A. and Ganesh, K. N. (2005) Conformationally constrained PNA analogs: Structural evolution towards DNA/RNA binding selectivity. *Acc. Chem. Res.*, **38**, 404-412.
4. (a) Crooke, S. T. (1998) S. T. Crooke, (Ed.) Antisense research and application, Springer, Berlin, pp 1-50., (b) Opalinska, J. B. and Gewirtz, A.M. (2002) Nucleic-acid therapeutics: basic principles and recent applications. *Nat. Rev. Drug Discovery*, **1**, 503-515.
5. Dominski, Z. and Kole, R. (1993) Restoration of correct splicing in Thalassemic pre-mRNA by antisense oligonucleotides. *Proc. Nat. Acad. Sci. U.S.A.*, **90**, 8673-8677.
6. Turner, J. J., Fabani, M., Arzumanov, A. A., Ivanova, G. and Gait, M. J. (2006) Targeting the HIV-1 RNA leader sequence with synthetic oligonucleotides and si-RNA: Chemistry and cell delivery. *Biochim. Biophys. Acta*, **1758**, 290-300.
7. Manoharan, M. (2002) Oligonucleotide conjugates as potential antisense drugs with improved uptake, biodistribution, targeted delivery, and mechanism of action. *Antisense Nucleic Acid Drug Dev.*, **12**, 103-128.
8. (a) Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G. and Prochiantz, A. (1996) Cell Internalization of the Third Helix of the Antennapedia Homeodomain is receptor-independent. *J. Biol. Chem.*, **271**, 18188-18193. (b) Vives, E., Brodin, P. and Lebleu, B.(1997) A truncated HIV-1 Tat protein basic domain rapidly translocates

through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem*., **272**, 16010-16017.

9. (a) Dietz, G. P. and Bahr, M. (2004) Delivery of bioactive molecules into the cell: the Trojan horse approach. *Mol. Cell Neurosci*., **27**, 85-131., (b) Trehin, R. and Merkle, H. P. (2004) Chances and pitfalls of cell penetrating peptides for cellular drug delivery. *Eur. J. Pharm. Biopharm.*, **2**, 209-223.
10. Rothbard, J. B., Garlington, S., Lin, Q., Kirschberg, T., Kreider, E., McGrane, P. L., Wender, P. A. and Khavari, P. A. (2000) Conjugation of arginine oligomers to cyclosporin A facilitates topical delivery and inhibition of inflammation. *Nat. Med*., **6**, 1253-1257.
11. Eguchi, A., Akuta, T., Okuyama, H., Senda, T., Yokoi, H., Inokuchi, H., Fujita, S., Hayakawa, T., Takeda, K., Hasegawa, M. and Nakanishi, M. (2001) Protein transduction domain of HIV-1 Tat protein promotes efficient delivery of DNA into mammalian cells. *J. Biol. Chem*., **276**, 26204-26210. (b) Gratton, J. P., Yu, J., Griffith, J. W., Babbitt, R. W., Scotland, R. S., Hickey, R., Giordano, F. J. and Sessa, W. C. (2003) Cell-permeable peptides improve cellular uptake and therapeutic gene delivery of replication-deficient viruses in cells and in vivo. *Nat. Med*., **9**, 357-362.
12. (a) Astriab-Fisher, A., Sergueev, D., Fisher, M., Shaw, B. R. and Juliano, R. L. (2002) Conjugates of antisense oligonucleotides with the Tat and Antennapedia cell-penetrating peptides: Effects on cellular uptake, binding to target sequences, and biologic actions *Pharm. Res*., **19**, 744. (b) Moulton, H. M., Nelson, M. H., Hatlevig, S. A., Reddy, M. T. and Iversen, P. L. (2004) Cellular uptake of antisense Morpholino oligomers conjugated to arginine-rich peptides. *Bioconjugate Chem*., **15**, 290-299.
13. (a) Tung, C. H. and Stein, S. (2000) Preparation and applications of peptide-oligonucleotide conjugates. *Bioconjugate Chem*., **11**, 605-618. (b) Turner, J. J., Arzumanov, A. A. and Gait, M. J. (2005) Synthesis, cellular uptake and HIV-1 Tat-dependent trans-activation inhibition activity of oligonucleotide analogues disulphide-conjugated to cell-penetrating peptides. *Nucleic Acids Res*., **33**, 27-42. (c) Rogers, F. A., Manoharan, M., Rabinovitch, P., Ward, D.C. and Glazer, P. M. (2004) Peptide conjugates for chromosomal gene targeting by triplex-forming oligonucleotides. *Nucleic Acids Res*., **32**, 6595-6604.
14. (a) Lundberg, P. and Langel, Ü. (2003) A brief introduction to cell-penetrating peptides. *J. Mol. Recogn*., **16**, 227-233. (b) Järver, P. and Langel, Ü. (2004) The use of cell-penetrating peptides as a tool for gene regulation. *Drug Discovery Today*, **9**, 395-402.

15. Derossi, D. et al. (1994) The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem*., **269**, 10444-10450.

16. Vivés, E. et al. (1997) A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem*., **272**, 16010-16017.

17. Pooga, M., Hällbrink, M., Zorko, M. and Langel, Ü. (1998) Cell penetration by transportan., *FASEB J*., **12**, 67-77.

18. (a) Oehlke, J., Krause,E., Wiesner, B., Beyermann, M. and Bienert, M. (1997) Extensive cellular uptake into endothelial cells of an amphipathic beta-sheet forming peptide. *FEBS Lett*., **415**,196-199. (b) Oehlke, J., Scheller, A., Wiesner, B., Krause, E., Beyermann, M., Klauschenz, E., Melzig, M. and Bienert, M. (1998) Cellular uptake of an alpha-helical amphipathic model peptide with the potential to deliver polar compounds into the cell interior nonendocytically.*Biochim. Biophys. Acta*, **1414**, 127-139.

19. Soomets, U., Lindgren, M., Gallet, X., Hallbrink, M., Elmquist, A., Balaspiri, L., Zorko, M., Pooga, M., Brasseur, R. and Langel, U. (2000) Deletion analogues of transportan. *Biochim. Biophys. Acta.*, **1467**,165-176.

20. Elmquist, A., Lindgren, M., Bartfai, T. and Langel, U. (2001) VE-cadherinderived cell-penetrating peptide, pVEC, with carrier functions. *Exp. Cell. Res*., **269**, 237–244.

21. Morris, M.C., Depollier, J., Mery. J., Heitz, F. and Divita, G. (2001) A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat. Biotechnol*., **19**, 1173-1176.

22. Turner, J. J., Ivanova, G. D., Verbeure, B., Williams, D., Arzumanov, A. A., Abes, S., Lebleu, B. and Gait, M. J. (2005) Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitors of HIV-1 Tat-dependent trans-activation in cells. *Nucleic Acids Res.*, **33**, 6837–6849.

23. (a) Moulton, H. M., Nelson, M. H., Hatlevig, S. A., Reddy, M. T. and Iversen, P. T. (2004) Cellular uptake of antisense morpholino oligomers conjugated to arginine-rich peptides. *Bioconjugate Chem*., **15**, 290-299. (b) Abes, S., Moulton, H. M., Clair, P., Prevot, P., Youngblood, D. S., Wu, R. P., Iversen P. L. and Lebleu, B. (2006) Vectorization of morpholino oligomers by the (R-Ahx-R)$_4$ peptide allows efficient splicing correction in the absence of endosomolytic agents. *J. Control Release*, **116**, 304-313.

24. Venkatesan, N. and Kim, B. H. (2006) Peptide Conjugates of Oligonucleotides: Synthesis and Applications. *Chem. Rev*., **106**, 3712-3761.

25. Petersen, L., de Koning, M. C., Van Kuik-Romeijn, P., Weterings, J., Pol, C. J., Platenburg, G., Overhand, M., Van der Marel, G. A. and Van Boom, J. H. (2004) Synthesis and in vitro evaluation of PNA-Peptide-DETA conjugates as potential cell penetrating artificial ribonucleases. *Bioconjugate Chem*., **15**, 576 -582.

26. (a) Eritja, R. In Solid-Phase Synthesis; Kates, S. A., Albericio, F., Eds.; Marcel Dekker: New York, NY, 2000; pp 529-548. (b) Virta, P., Katajisto, J., Niittymaki, T. and Lonnberg, H. (2003) Solid-supported synthesis of oligomeric bioconjugates. *Tetrahedron*, **59**, 5137. (c) Stetsenko, D. A. and Gait, M. J. (2005) *Methods Mol. Biol*., **288**, 205.

27. Tung, C. H., Rudolph, M. J. and Stein, S. (1991) Preparation of oligonucleotide-peptide conjugates. *Bioconjugate Chem*., **2**, 464-465.

28. (a) Forget, D., Boturyn, D., Defrancq, E., Lhomme, J. and Dumy, P. (2004) Highly efficient synthesis of peptide-oligonucleotide conjugates: chemoselective oxime and thiazolidine formation. *Chem. - Eur. J*., **7**, 3976-3984. (b) Edupuganti, O. P., Renaudet, O., Defrancq, E. and Dumy, P. (2004) The oxime bond formation as an efficient chemical tool for the preparation of 3′, 5′-bifunctionalised oligodeoxyribonucleotides. *Bioorg. Med. Chem. Lett.,* **14**, 2839-2842.

29. (a) Zatsepin, T. S., Stetsenko, D. A., Arzumanov, A. A., Romanova, E. A., Gait, M. J. and Oretskaya, T. S. (2002) Synthesis of peptide-oligonucleotide conjugates with single and multiple peptides attached to 2'-aldehydes through thiazolidine, oxime, and hydrazine linkages. *Bioconjugate Chem*., **13**, 822-830. (b) Zatsepin, T. S., Romanova, E. A., Stetsenko, D. A., Gait, M. J. and Oretskaya, T. S. (2003) Synthesis of 2'-modified oligonucleotides containing aldehyde or ethylenediamine groups. *Nucleosides, Nucleotides Nucleic Acids*, **22**, 1383-1385.

30. Arar, K., Aubertin, A.-M., Roche, A.-C., Monsigny, M. and Mayer, R. (1995) Synthesis and antiviral activity of peptide-oligonucleotide conjugates prepared by using *N*-alpha.-(bromoacetyl) peptides. *Bioconjugate Chem*., **6**, 573-577.

31. (a) Corey, D. R. (2004) *Methods Mol. Biol*., **283**, 197, (b) Zatsepin, T. S., Turner, J. J., Oretskaya, T. S. and Gait, M. J.(2005) *Curr. Pharm. Des*., **11**, 3639.

32. Eritja, R., Pons, A., Escarceller, M., Giralt, E. and Albericio, F. (1991) Synthesis of defined peptide-oligonucleotide hybrids containing a nuclear transport signal sequence. *Tetrahedron*, **47**, 4113-4120.

33. Tripathi, S., Chaubey, B., Ganguly, S., Harris, D., Casale, R. A. and Pandey, V. N. (2005) Anti-HIV-1 activity of anti-TAR polyamide nucleic acid conjugated with various membrane transducing peptides. *Nucleic Acids Res*., **33**, 4345-4356.

34. Marchan, V., Ortega, S., Pulido, D., Pedroso, E. and Grandas, A. (2006) Diels-Alder cycloadditions in water for the straightforward preparation of peptide oligonucleotide conjugates. *Nucleic Acids Res*., **34**, e24.

35. (a) Awasthi, S. K. and Nielsen, P. E. (2002) *Methods Mol. Biol.,* **208**, 43. (b) Bendifallah, N., Rasmussen, F. W., Zachar, V., Ebbesen, P., Nielsen, P. E. and Koppelhus, U. (2006) Evaluation of cell-penetrating peptides (CPPs) as vehicles for intracellular delivery of antisense peptide nucleic acid (PNA). *Bioconjugate Chem*., **17**, 750-758.

36. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C.(1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806-811.

37. Schiffelers, R. M., Ansari, A., Xu, J., Zhou, Q., Tang, Q., Storm, G., Molema, G., Lu, P. Y., Scaria, P. V. and Woodle, M. C. (2004) Cancer siRNA therapy by tumor selective delivery with ligand-targeted sterically stabilized nanoparticle**.** *Nucleic Acids Res*., **32**, e149.

38. Chiu, Y.-L., Ali, A., Chu, C.-Y., Cao, H. and Rana, T. M. (2004) *Chem. Biol.*, **11**, 1165.

39. Muratovska, A. and Eccles, M. R. (2004) *FEBS Lett*., **558**, 63.

40. (a) Huisgen, R. (1984) 1, 3-dipolar cycloaddition chemistry (Ed.: A. Padwa), Wiley, New York (b) Tornøe, C. W., Christensen, C. and Meldal, M. (2002) Peptidotriazoles on solid phase: [1, 2, 3]-triazoles by regiospecific copper (I)-catalyzed 1, 3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem*., **67**, 3057-3064. (c) Narayan, S., Muldoon, J., Finn, M. G., Fokin, V. V., Kolb, H. C. and Sharpless, K. B. (2005) "On Water": Unique reactivity of organic compounds in aqueous suspension. *Angew. Chem. Int. Ed.*, 44, 3275-3279. (d) Kolb, H. C., Finn, M. G. and Sharpless, K. B. (2001) Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed*., **40**, 2004-2021. (e) Kolb, H. C. and Sharpless, K. B. (2003) *Drug Discovery Today*, **8**, 1128-1137.

41. Bock, V. D., Hiemstra, H. and van Maarseveen, J. H. (2006) *Eur. J. Org. Chem*., 51-68.

42. Bouillon, C., Meyer, A., Vidal, S., Jochum, A., Chevolot, Y., Cloarec, J.-P., Praly, J.-P., Vasseur, J.-J. and Morvan, F. (2006) Microwave assisted "Click" chemistry for the synthesis of multiple labeled-carbohydrate oligonucleotides on solid support. *J. Org. Chem.*, **71**, 4700-4702.

43. Meier, J. L., Mercer, A. C., Rivera, Jr. H. and Burkart, M. D. (2006) Synthesis and evaluation of bioorthogonal pantetheine analogues for in vivo protein modification. *J. Am. Chem. Soc.*, **128**, 12174-12184.

44. Wan, Q., Chen, J., Chen, G. and Danishefsky, S. J. (2006) A potentially valuable advance in the synthesis of carbohydrate-based anticancer vaccines through extended cycloaddition chemistry. *J. Org. Chem.*, **71**, 8244 -8249.

45. Lin, P.-C., Ueng, S.-H., Tseng, M.-C., Ko, J.-L., Huang, K.-T., Yu, S.-C., Adak, A. K., Chen, Y.-J. and Lin, C.-C. (2006) Site-Specific protein modification through Cu^{I}-catalyzed 1,2,3-triazole formation and its implementation in protein microarray fabrication. *Angew. Chem. Int. Ed.*, **45**, 4286-4290.

46. Nakane, M., Ichikawa, S. and Matsuda A. (2005) Synthesis of cross-linked circular DNAs using Huisgen reaction. *Nucleic Acids, Symp. Ser.*, **49**, 189-190.

47. Kumar, R., El-Sagheer, A., Tumpane, J., Lincoln, P., Wilhelmsson, L. M. and Brown, T. (2007) Template-directed oligonucleotide strand ligation, covalent intramolecular DNA circularization and catenation using click chemistry. *J. Am. Chem. Soc.*, **129**, 6859-6864.

48. Rodionov, V. O., Fokin, V. V. and Finn M. G. (2005) *Angew. Chem. Int. Ed.*, **44**, 2210-2215.

49. Himo, F., Lovell, T., Hilgraf, R., Rostovtsev, V. V., Noodleman, L., Sharpless, K. B. and Fokin, V. V. (2004) *J. Am. Chem. Soc.*, **127**, 210–216.

50. Gangamani, B. P., Kumar, V. A. and Ganesh, K. N. (1996) Synthesis of N^{α}-(puinyl/pyrimidinyl acetyl)-4-aminoproline diastereomers with potential use in PNA synthesis. *Tetrahedron*, **52**, 15017-15030.

51. Nielsen, P. E. and Egholm, M. in Peptide nucleic acids (PNA). Protocols and Applications; Nielsen, P, E., Egholm, M. Eds. Horizon scientific Norfolk: CT, 1999.

52. Bluhm, B. K., Shields, S. J., Bayse, C. A., Hall, M. B. and Russell, D. H. (2001) Determination of copper binding sites in peptides containing basic residues: a combined experimental and theoretical study. *Int. J. Mass Spectrom.,* **204,** 31–46.

53. Weller, R. L. and Rajski, S. R. (2005) DNA methyltransferase-moderated click chemistry. *Org. Lett.*, **7**, 2141-2144.

54. Turner, J. J., Jones, S., Fabani, M., Ivanova, G., Arzumanov, A. A. and Gait, M. (2007) RNA targeting with peptide conjugates of oligonucleotides, siRNA and PNA. *Blood Cells, Molecules and Diseases*, **38**, 1-7.

5.6 Experimental

The chemicals used were of laboratory or analytical grade. All solvents used were purified according to the literature procedure. 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 sulfate 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. Column chromatographic separations were performed using silica gel 60-120, 200-400 mesh (Merck), solvent systems gradient 10-25% EtOAc/Pet ether and pure DCM to 3% 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 spraying with percholric acid.

^{1}H and ^{13}C spectra were obtained using Bruker AC-200 and AC-400 NMR spectrometers. The chemical shifts are reported in delta (δ) values and referred to internal standard TMS for ^{1}H and deuterated NMR solvent chloroform-*d* itself for ^{13}C.

N^{α}-Fmoc L-amino acids and resins for peptide and PNA synthesis were obtained from Novabiochem (Fmoc= 9-fluorenyl methoxy-carbonyl). The Fmoc amino acid used were Fmoc -Lys(Boc)-OH, Fmoc-Arg(Mtr)-OH and N^{ε}-Fmoc-Aha-OH.(Mtr = 4-methoxy-2,3,6-trimethylbenzenesulfonyl, Boc = *tert*-butoxycarbonyl, Aha = 6-aminohexanoic acid (or ε-amino caproic acid). 6-aminohexanoic acid was obtained from Aldrich chemicals and *N*-Fmoc protection was done using standard procedure. The concentrations of DNA/PNAs and their conjugates were determined spectrophotometrically and the concentration of the peptides was assumed approximately. Propynoic acid was procured from Trade (TCI) Mark, Tokyo Kasei. *N*-Boc-(2*S*, 4*S*)-4-azidoproline was synthesized according to the reported procedure.[50] The Boc and Fmoc protected PNA monomers were obtained from Applied Biosystems, USA. TANA monomers were synthesized in the laboratory following the procedures described in chapter 2. Propyne substituted 19 mer DNA sequence was synthesized using the phosphoramidite approach and an Applied Biosystems 3900 DNA synthesizer.

Reverse phase high-performance liquid chromatography analyses were carried out on VARIAN Analytical Semi-prep HPLC system consisting of Varian Pro-star 210 Binary solvent delivery system. Linear gradients of A: 0.1% TFA in water and B: 55/45: Acetonitrile/Water, 0.1% TFA (Linear gradient from A to B in 30 mins Flow- 1.5 ml/min.). Rainin Dynamax UV D-II Absorbance Detector Star Ver.5 at detection wavelength 254 nm or 220 nm was employed during the experimentation. Chromatography Workstation Rheodyne 7725I with manual injector and Lichrocart Lichrispher 100RP-18,250 x 4 mm id. Particle size-5 μm column were used.

Mass spectral analysis was performed on a Voyager-De-STR (Applied Biosystems) MALDI-TOF. A nitrogen laser (337 nm) was used for desorption. The matrixes used for analysis were CHCA (α-Cyano-4-hydroxycinnamic Acid), THAP (2', 4', and 6'-trihydroxyacetophenone) and HPA (3-Hydroxypicolinic acid). Diammonium citrate was used as additive when THAP and HPA were used as matrix.

Solid-phase synthesis of azide functionalized PNA (TANA) oligomers 3, 4 and 5

The azide functionalized PNA (TANA) oligomers **3** and **5** were synthesized on Rink-amide resin (100 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis (20 min treatment with 20% piperidine in *N*,*N*-dimethyl formamide and reaction with 3 equivalents of Fmoc-PNA monomer, HBTU, HOBt and DIPEA for 6 h, were used for the deprotection and coupling steps, respectively). The last coupling reaction was done with *N*-Boc-(2*S*, 4S)-4-azidoproline. The progress of the coupling reaction was tested by Kaiser test at each step. Cleavage and deprotection were effected by reaction with TFA/DCM/TIS (10:85:5) for 30 min. The resulting oligomer was precipitated by addition of cold ether and purification was done by gel filtration followed by RP HPLC and characterized by MALDI-TOF mass spectrometry. (Table 2, entry 3 and 4).

The azide functionalized mixed PNA oligomer **4** was synthesized using side chain *N*-CBz protected Boc-L-Lysine functionalized MBHA resin (100mg, loading 0.2 mmol/g) following the standard procedures of solid-phase peptide synthesis (30 min treatment with 50% TFA in DCM and reaction with 3 equivalent of Boc-PNA monomer, HBTU, HOBt and DIPEA for 6 h, were used for the deprotection and coupling steps,

respectively). The last coupling reaction was done with *N*-Boc-(*2S, 4S*)-4-azidoproline. Cleavage and deprotection were effected by the reaction with TFA-TFMSA, thioanisole, ethanedithiol for 2 h. The resulting oligomer was precipitated by addition of cold ether and purification was done by gel filtration followed by RP HPLC and characterized by MALDI-TOF mass spectrometry. (Table 2, entry 5).

Solid-phase synthesis of alkyne functionalized Lysine peptide HO-$(Lys)_6$-alkyne 1

The alkyne functionalized lysine peptide was synthesized on Rink-amide resin (100 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis (20 min treatment with 20% piperidine in *N,N*-dimethyl formamide and the reaction with 3 eqivalent of Fmoc-Lys(Boc)-OH, HBTU, HOBt and DIPEA for 4 h, were used for the deprotection and coupling steps, respectively). The last coupling was done with propynoic acid using HBTU, HOBt and DIPEA as the coupling agent. The peptide was cleaved from the resin by treatment with TFA/DCM/TIS (10:85:5) for 30 min. and the deprotection of side chain was effected by 50% TFA-DCM (TIS used as scavenger) for 1 h. The resulting peptide was precipitated by addition of cold ether and was purified by RP HPLC and characterized by MALDI-TOF mass spectrometry. (Table 2, entry 1)

Click reaction on solid phase: Synthesis of HO-$(Lys)_6$-triazole-proline 6

20 mg of resin bound $(Lys)_6$-alkyne **1**, *N*-Boc- (*2S,4S*)-4-azidoproline **28** (5 mg), CuI (1 mg), DIPEA(5µl) and DMF(0.3 ml) were reacted together in a reaction vessel for 8 h (Scheme 1). The excess reagents were washed out with DMF followed by DCM. 5 mg of resin was cleaved by using the standard conditions and the product formed was purified by HPLC and characterized by MALDI-TOF mass spectrometry. (Table 2, entry 6).

Solution phase reaction between PNA-proline azides 3, 4 and 5 and HO-$(Lys)_6$-alkyne 1: **Synthesis of conjugates 7, 8 and 9**

General procedure:

HO-β-ala-TTTTTTTT-Pro-N_3 **3** (or HO-Lys-TCACTAGATG-Pro-N_3 **4** or HO-β-ala-t-t-t-t-t-t-t-t-Pro-N_3 **5**) (1 µmol) and HO-$(Lys)_6$-alkyne **1** (approximately 3 µmol) were dissolved in 50 µl water:*tert.* butanol (1:1). $CuSO_4.5H_2O$ (1.0 equivalent, 1µmol, 10 µl of

a 100 mM solution in water) and freshly prepared solution of sodium ascorbate (4 equivalents, 4 μmol, 8 μl of 500 mM solution in water) was then added. The mixture was stirred in a spinix vortex at room temperature (Scheme 3). The reaction mixture was analyzed by HPLC after 2h. HPLC showed complete consumption of the starting material in the case of **5** giving product **9** (Table 2, entry 9). In the case of either **3** or **4** the differences in HPLC-t_R was not very clear from the starting materials. In these cases the product (**7** or **8**) after HPLC purification were characterized by MALDI-TOF mass spectrometry (Table 2, entry 7 and 8). Complete conversion of the starting material **3** and **4** was established by the disappearance of the corresponding mass peak in MALDI-TOF analysis.

Solid-phase synthesis of the peptide HO-(Arg-Aha-Arg)$_4$-alkyne 2

The arginine peptide **2** was synthesized on Rink-amide resin (200 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis: 20 min treatment with 20% piperidine in *N,N*-dimethyl formamide and reaction with 3 equivalent of Fmoc-Arg(Mtr)-OH (or Fmoc-ε amino hexanoic acid), HBTU, HOBt and DIPEA for 4 h, were used for the deprotection and coupling steps, respectively. The final coupling was done with propynoic acid using HBTU, HOBt and DIPEA as the coupling agent. Cleavage of the peptide from resin was effected by reaction with TFA/DCM/TIS (10:85:5) for 30 min. The deprotection of side chain protecting group (Mtr) was effected by 100% TFA and thioanisole as the scavenger for 5 hours. The resulting peptide was precipitated by addition of cold ether and oligomer was purified by RP-HPLC. The product was further characterized by MALDI-TOF mass spectrometry. (Table 2, entry 2).

Solution phase reaction between PNA (TANA)-proline azides 4, 5 and HO-(Arg-Aha-Arg)$_4$-alkyne 2: **Synthesis of conjugates 10 and 11**

General procedure:

HO-Lys-TCACTAGATG-Pro-N_3 **4** or HO-β-ala-t-t-t-t-t-t-t-t-Pro-N_3 **5** (1 μmol) and (Arg-Aha-Arg)$_4$-alkyne **1** (3 μmol, approximately) were dissolved in 50 μl water:*tert.* Butanol (1:1). $CuSO_4.5H_2O$ (1.0 equivalent, 1μmol, 10 μl of a 100 mM solution in water) and freshly prepared solution of sodium ascorbate (4 equivalents, 4 μmol, 8 μl of 500

mM solution in water) was then added (Scheme 3). The mixture was stirred in a spinix vortex at room temperature. The reaction mixture was analyzed by RP- HPLC after 2h. HPLC showed complete conversion of stating materials to the products. The products (**10** and **11**) formed were purified by RP-HPLC and characterized by MALDI-TOF mass spectrometry (Table 2, entry 10, 11).

Synthesis of alkyne substituted DNA oligomer 12

The DNA sequence **12** was synthesized using commercially available monomeric units using phoaphoramidite chemistry on automated DNA synthesizer. The last coupling was done using 5'-*O*-(propynyl-*N*-*3*-benzoyl-thymidin)-3'-*O*-(*N*,*N*-diisopropylamino-*O*-cyanoethyl-phosphoramidite) **22** (Figure 11). The ON **12** was deprotected and cleaved from the solid support. It was further purified by RP-HPLC and characterized by MALDI-TOF mass spectrometry. (Table 3, entry 1)

Solid-phase synthesis of the peptide HO-(Arg-Aha-Arg)$_4$-azide 13

The arginine peptide **13** was synthesized on Rink-amide resin (200 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis: 20 min treatment with 20% piperidine in *N, N*-dimethyl formamide and reaction with 3 equivalent of Fmoc-Arg(Mtr)-OH (or Fmoc-ε amino hexanoic acid), HBTU, HOBt and DIPEA for 4 h, were used for the deprotection and coupling steps, respectively. The final coupling was done with *N*-Boc-4-azido-proline **28** using HBTU, HOBt and DIPEA as the coupling agent. Cleavage of the peptide from resin was effected by reaction with TFA/DCM/TIS (10:85:5) for 30 min. The deprotection of side chain protecting group (Mtr) was effected by 100% TFA and thioanisole as the scavenger for 5 hours. The resulting peptide was precipitated by addition of cold ether and was purified by RP HPLC. The product was further characterized by MALDI-TOF mass spectrometry. (Table 3, entry 2)

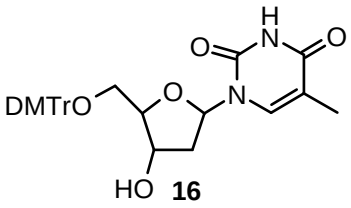
Solution phase reaction between DNA-alkyne 12 and HO-(Arg-Aha-Arg)$_4$-azide 13: Synthesis of conjugate 14

Alkyne functionalized oligonucleotide **12** (1 µmol) and azide functionalized peptide **13 (**3 µmol, approximately) were dissolved in 50 µl water:*tert*. Butanol (1:1). $CuSO_4.5H_2O$ (0.6 equivalents, 0.6 µmol, 6µl of a 100 mM solution in water) and freshly prepared solution of sodium ascorbate (2.4 equivalents, 4 µmol, 6 µl of 500 mM solution in water) was then added (Scheme 6). The mixture was stirred in a spinix vortex at room temperature. After 2 h, a small portion of the reaction mixture was analyzed by RP- HPLC. HPLC showed complete conversion of stating materials to the product. The product **14** formed was purified by RP- HPLC and characterized by MALDI-TOF mass spectrometry. (Table 3, entry 3).

Synthesis of 5'-O-propynyl-thymidin-3'-O-phosphoramidite 22

5'ODMTr Thymidine 16

2.42 g (10 mmol) of thymidine **15**, DMTr-Cl (4.1g, 12 mmol), DMAP (122 mg, 0.1 mmol) were dissolved in 60 ml pyridine and to it added TEA (2.8 ml, 20 mmol). The reaction mixture was stirred at room temperature for six hours. The solvent was removed *in vacuo* and redissolved the residue in 250 ml ethyl acetate. The organic layer was washed with sodium bicarbonate (2 x 30 ml), water (2 x 50 ml) and brine (2 x 30 ml) and then kept over anhydrous Na_2SO_4. The solvent was removed *in vacuo* to give **16** as white foam which was desiccated and used for the next reaction without any purification.

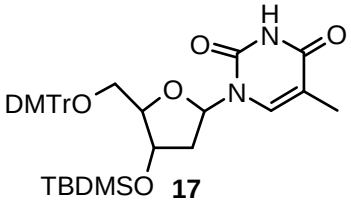


16

Yield: 4.8 g, 88%

5'ODMT 3'OTBDMS Thymidine 17

(4.8 g, 8.81 mmol) 5'ODMT thymidine **16**, Imidazole (1.5 g, 22 mmol) and TBDMS-Cl (2 g, 10.6 mmol) were dissolved in 30 ml anhydrous DMF and stirred for six hours. The solvent was removed *in vacuo* and redissolved the mixture in 250 ml ethyl acetate. The organic layer was washed with water (3 x 50 ml) and brine (2 x 30 ml) and then kept over anhydrous sodium sulphate. The solvent was

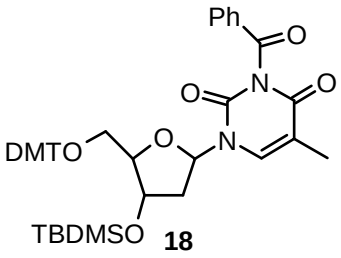


17

evaporated to dryness *in vacuo*. Then column purification was done using 40% ethyl acetate-pet ether as the solvent system to give 5.96 g pure **17** yield 97%.

N-benzoyl, 5'-ODMT 3'OTBDMS thymidine 18

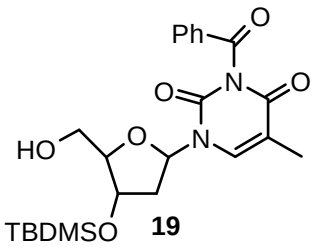
(5.8 g, 8.3 mmol) of 5'ODMT 3'OTBDMS thymidine **17** was dissovled in 50 ml anhydrous pyridine and to added it DIPEA (3 ml,16.6 mmol). The reaction mixture was cooled to 0ºC. Then added drop wise benzoyl chloride (1.9 ml, 16.6 mmol) by a syringe. Stirred the reaction mixture for five hours. The solvent was evaporated *in vacuo* and redissolved the residue in 200 ml ethyl acetate. The organic layer was washed with 5% aquous $NaHCO_3$ solution (2 x 20 ml), followed by water (2 x 50 ml) and then brine (2 x 30 ml).The organic layer was kept over anhydrous sodium sulphate and evaporated to dryness *in vacuo*. Column purification was done using 20-30% ethyl acetate- pet ether to afford 5.8 g pure **18** Yield 87%.



18

N-benzoyl 3'OTBDMS thymidine 19

5.7 g (7.1mmol) of N-benzoyl, 5'-ODMT 3'OTBDMS thymidine **18** was dissolved in 20 ml 3% DCA-DCM solution and to it added 1ml triethyl silane as the scavenger. The reaction mixture was stirred for 1 hour. The reaction mixture was diluted by addition of 20 ml more DCM and neutralized the excess acid with 5% $NaHCO_3$ solution (2 ml). The organic layer was washed with 20 ml more water and kept it over anhydrous Na_2SO_4 and evaporated to dryness. Column purification was done using 1% Methanol-DCM to afford 2.7 g pure **19** yield 76%.



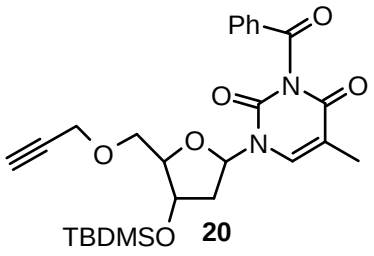
19

^{1}H NMR δ 0.08 (s, 6H), 0.88 (s, 9H), 1.95 (s, 3H), 2.2-2.3 (m, 2H), 3.7-3.9 (m, 3H), 4.5 (m, 1H), 6.2(t, 1H), 7.4-7.7 (m, 4H), 7.9 (d, 2H).

$C_{23}H_{32}N_2O_6Si$ Mass calculated 460.6 observed 461.1

N-benzoyl 5' [(prop-2-ynyloxy) methyl] 3'OTBDMS thymidine 20

1.5 g (3.25 mmol) of compound **19** was dissolved in 4 ml anhydrous DMF and cooled to 0ºC and (0.16 g, 60% solution, 3.9 mmol) NaH was slowly added in two portions. The reaction mixture was stirred for 30 minutes, and propynyl bromide 0.6 ml (7.8 mmol) then added dropwise. The reaction mixture was stirred for another 20 minutes at 0ºC. The reaction was quenched by addition of 2 ml water. The solvent was evaporated *in vacuo* and redissolved the residue in 150 ml ethyl acetate, washed the organic layer with water (2 x 50 ml) followed by brine (2 x 20 ml).The organic layer was kept over anhydrous Na_2SO_4 and evaporated *in vacuo*. Column purification using 30% ethyl acetate-pet ether gave 1.2g 5'-*O*-propynyl-*N*3-benzoyl 3'-*O-tert.* butyldimethysilyl thymidine **(**yield 74%).

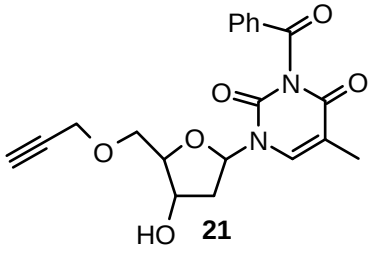


20

^{1}H NMR δ 0.08 (s, 6H), 0.88 (s, 9H), 1.7 (d, 2H), 1.99(d, 1H), 2.1-2.53 (m, 3H), 3.6-4.7 (m, 6H), 6.3 (m, 1H), 7.5-7.9 (m, 6H).

N-benzoyl 5' [(prop-2-ynyloxy) methyl] thymidine 21

1.1 g (2.2 mmol) 5'-*O*-propynyl-*N*3-benzoyl 3'-*O-tert.* butyldimethysilyl thymidine **20** was dissolved in 15 ml THF and to it added 1 g TBAF (3.3 mmol). The reaction mixture was stirred for 1h. THF was removed and the residue was redissolved in 100 ml ethyl acetate. The organic layer was washed with water (2 x 30 ml), followed by brine (1 x 20 ml). The organic layer was kept over anhydrous Na_2SO_4 and evaporated *in vacuo*. Column purification using 2% Methanol- DCM to afford 0.8 g, **21** Yield 95%.

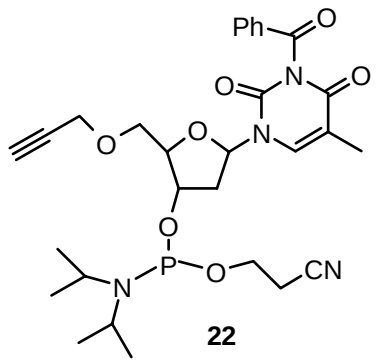


21

^{1}H NMR δ 1.67 (s, 2H), 1.96 (d, 1H), 2.1-2.6 (m, 3H), 3.6-3.7 (m, 6H), 6.3-6.4 (m, 1H), 7.4-8.05 (m, 6H). $C_{20}H_{20}O_6N_2$ Mass calculated 384.4 observed 384.6

N-benzoyl 5' [(prop-2-ynyloxy) methyl] 3'-*O*-(2-cyanoethyl-N, N diisopropylphosphoramidite) Thymidine 22

Compound **21** (0.7g, 1.82 mmol) was dissolved in dry DCM (10 ml) followed by the addition of di-isopropyl ethyl amine (0.39 ml, 2.2 mmol) and chloro (2-cyanoethoxy) (*N,*



22

N-diisopropylamino)-phosphine (0.52 ml, 2.2 mmol) and the reaction mixture was stirred at room temperature for 1h. The contents were then diluted with 10 ml dichloromethane and washed with 5% $NaHCO_3$ solution (1 x 10 ml). The organic phase was dried over anhydrous Na_2SO_4 and concentrated to foam. The residue was dissolved in DCM and precipitated with hexane to obtain **22** (0.73 g, 69%). The phosphoramidite **22** was dried overnight over P_2O_5 and KOH in a desiccator before using on DNA synthesizer. TLC shows two close moving spots for two diastreomers (R_f=0.5, 1% methanol-dichloromethane)**.**

^{31}P NMR δ 148.98, 149.1. C_{29} H_{35} N_4O_7P Mass calculated 582.6 observed 583.3

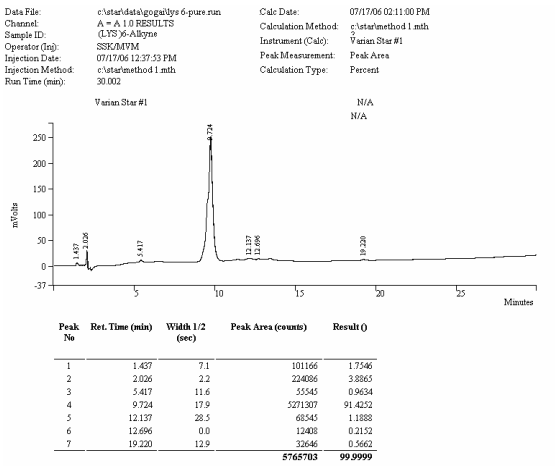
5.7 Appendix

RP-HPLC of the starting materials and products **1- 14**

MALDI-TOF Mass of the starting materials and products **1- 14**.

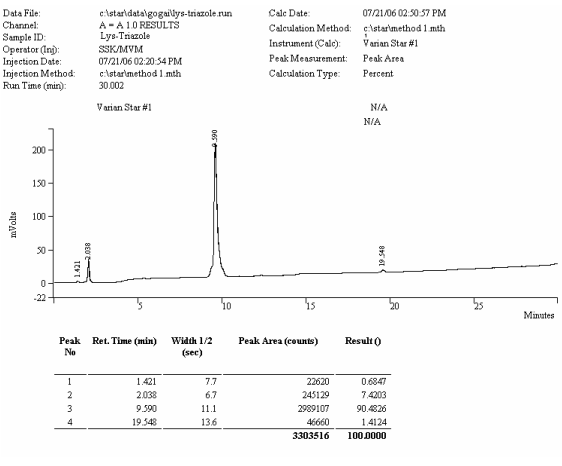
^{1}H NMR of **19**, **20** and **21**

^{31}P NMR of **22**



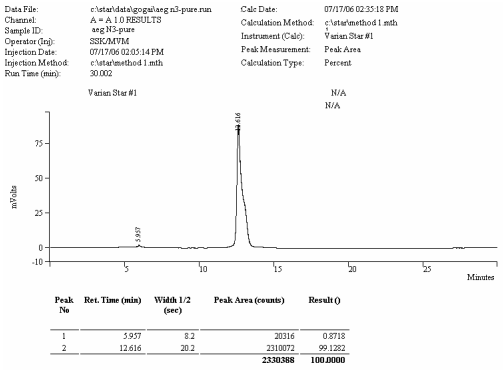
Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	1.437	7.1	101166	1.7546
2	2.026	2.2	224086	3.8865
3	5.417	11.6	55545	0.9634
4	9.724	17.9	5271307	91.4252
5	12.137	28.5	68545	1.1888
6	12.696	0.0	12408	0.2152
7	19.220	12.9	32646	0.5662
			5765703	**99.9999**

RP-HPLC of $(Lys)_6$-alkyne **1**
Detector (UV set at 220nm)



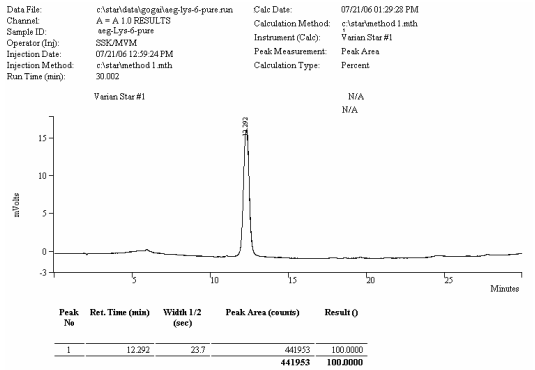
Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	1.421	7.7	22620	0.6847
2	2.038	6.7	245129	7.4203
3	9.590	11.1	2989107	90.4826
4	19.548	13.6	46660	1.4124
			3303516	**100.0000**

RP-HPLC of $(Lys)_6$-triazole-proline **6**
(Detector –UV set at 220nm)



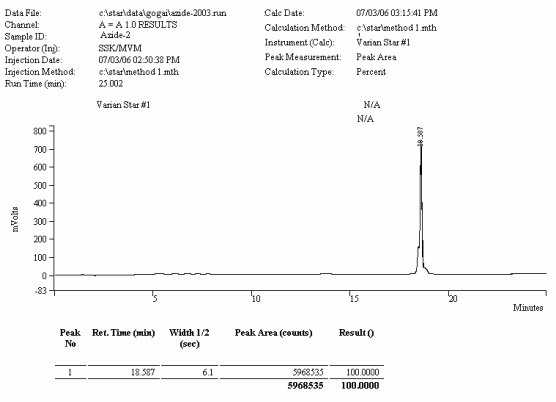
Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	5.937	8.2	20316	0.8718
2	12.616	20.2	2310072	99.1282
			2330388	**100.0000**

RP-HPLC β-ala-TTTTTTTT-Pro-N_3 **3**
Detector –UV set at 254nm



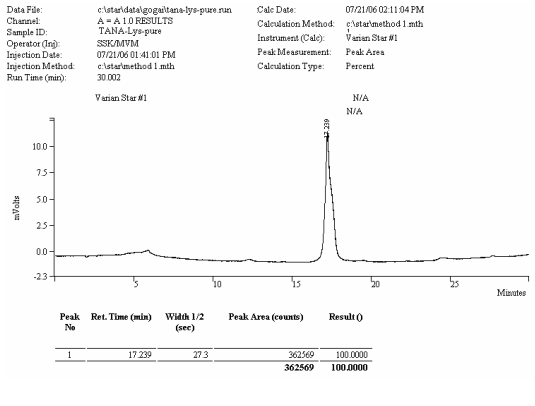
Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	12.292	23.7	441953	100.0000
			441953	**100.0000**

RP-HPLC β-ala-TTTTTTTT-Pro-triazole-$(Lys)_6$
7 Detector –UV set at 254 nm



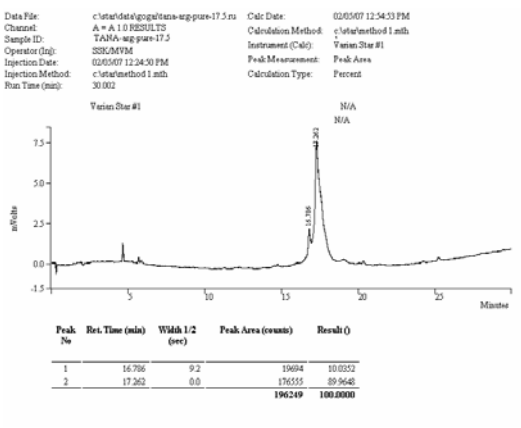
Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	18.587	6.1	5968535	100.0000
			5968535	**100.0000**

RP-HPLC β-ala-t-t-t-t-t-t-t-t-t-Pro-N_3 **5**
Detector –UV set at 254nm

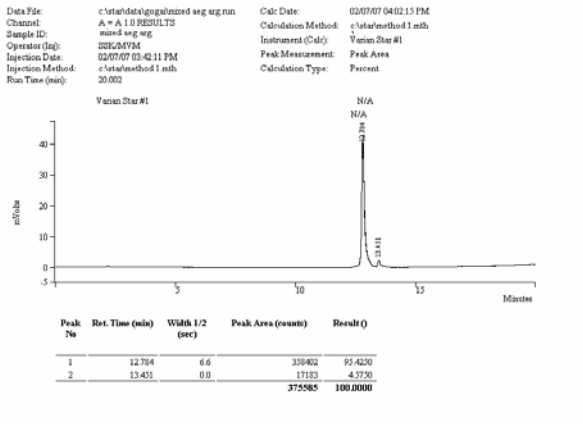


Peak No	Ret. Time (min)	Width 1/2 (sec)	Peak Area (counts)	Result ()
1	17.239	27.3	362569	100.0000
			362569	**100.0000**

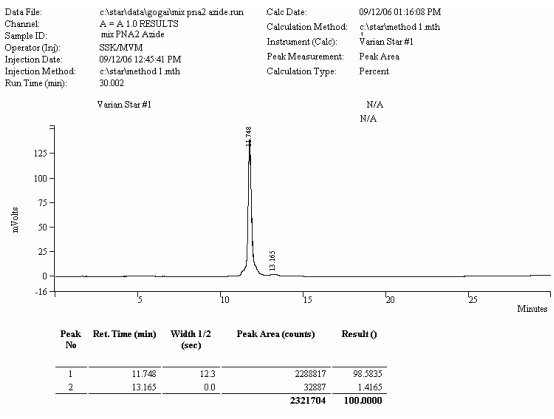
β-ala-t-t-t-t-t-t-t-t-t-Pro- triazole-$(Lys)_6$ **9**
Detector –UV set at 254nm



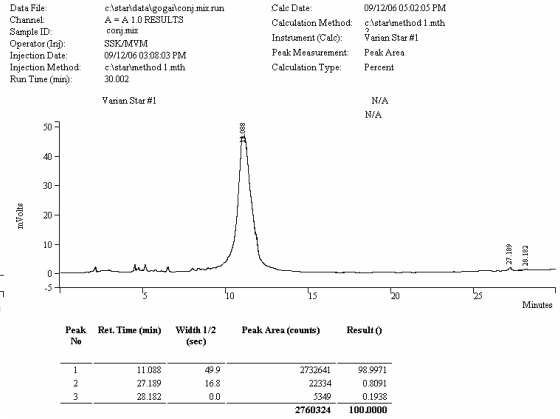
RP-HPLC of β-ala-t-t-t-t-t-t-t-t-Pro-triazole-(Arg-Aha-Arg)$_4$ **10** Detector – UV set at 254 nm



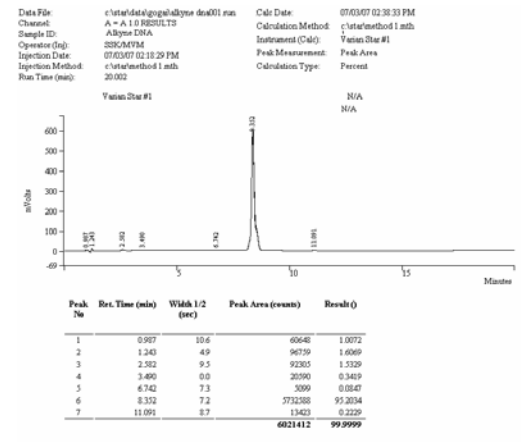
RP-HPLC of Lys-TCACTAGATG-Pro-triazole - (Arg-Aha-Arg)$_4$ **11** Detector – UV set at 254nm



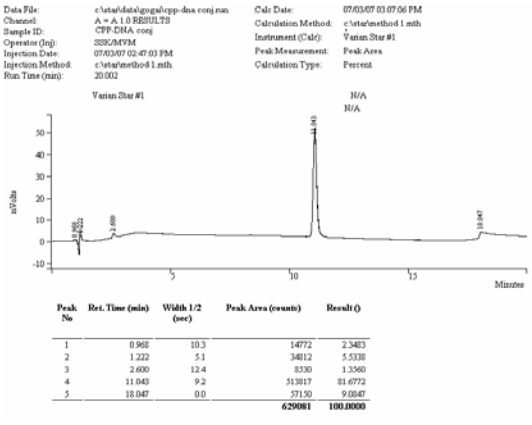
RP-HPLC of Lys-TCACTAGATG-Pro-N_3 **4** Detector –UV set at 254nm



RP-HPLC of Lys-TCACTAGATG-Pro-triazole - $(Lys)_6$ **8** Detector –UV set at 254nm

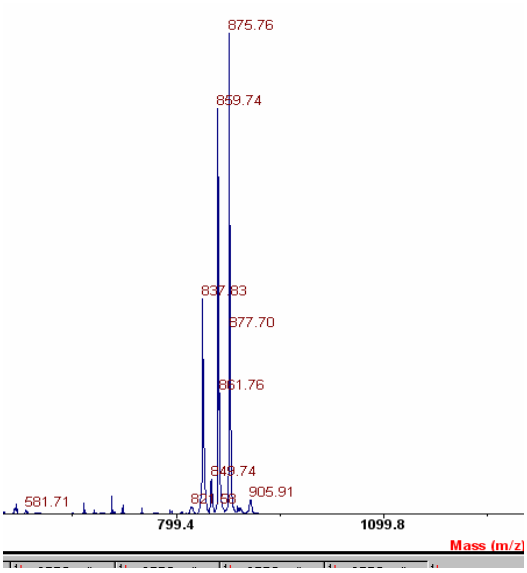


RP-HPLC of **12** detector set at 254 nm

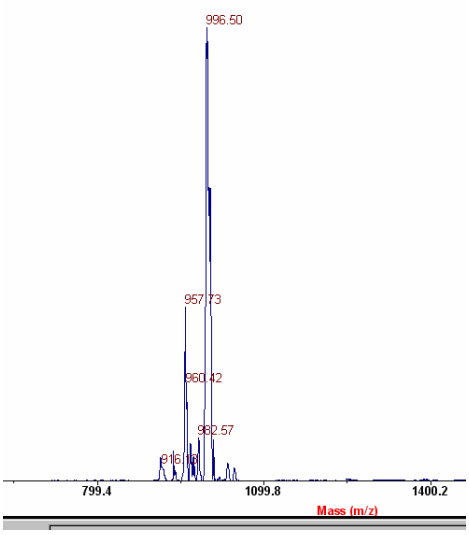


RP-HPLC of **14** detector set at 254 nm

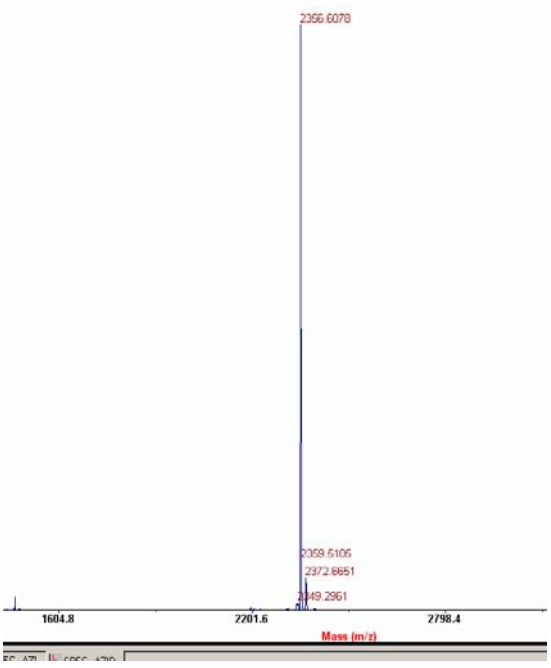
MALDI-TOF MASS OF PEPTIDES AND OLIGOMERS



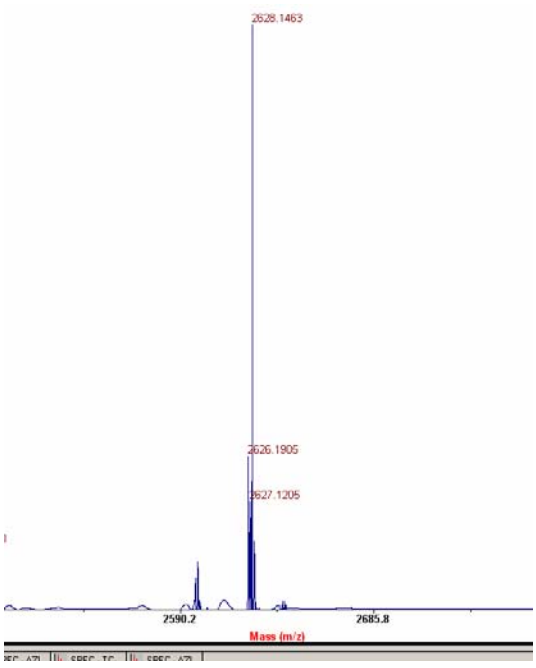
((Lys)$_6$-alkyne **1**
calculated 839.08 observed 837.83



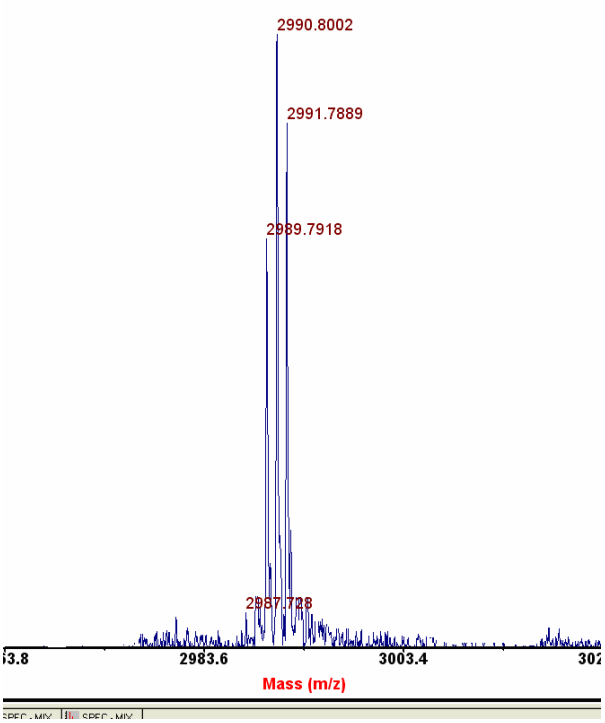
(Lys)$_6$-triazole-proline **6**
Calculated 995.23 Observed 996.5



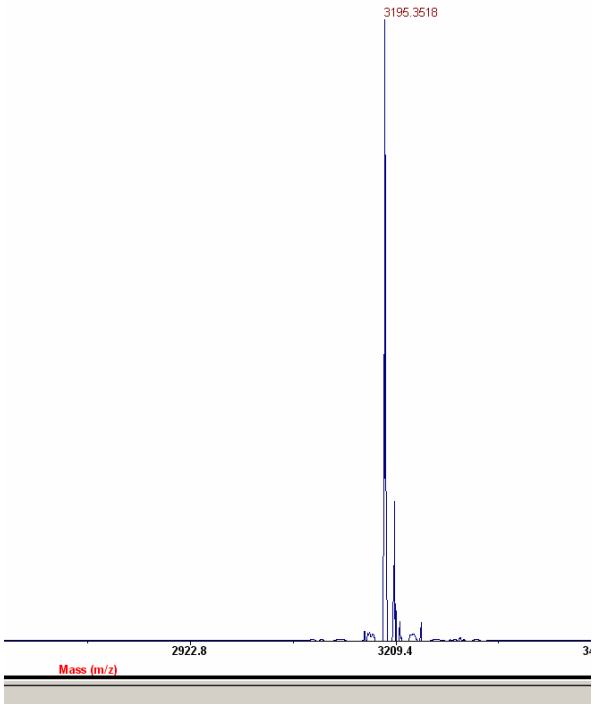
β-ala-TTTTTTTT-Pro-N_3 **3**
Calculated 2356.21
Observed 2356.6



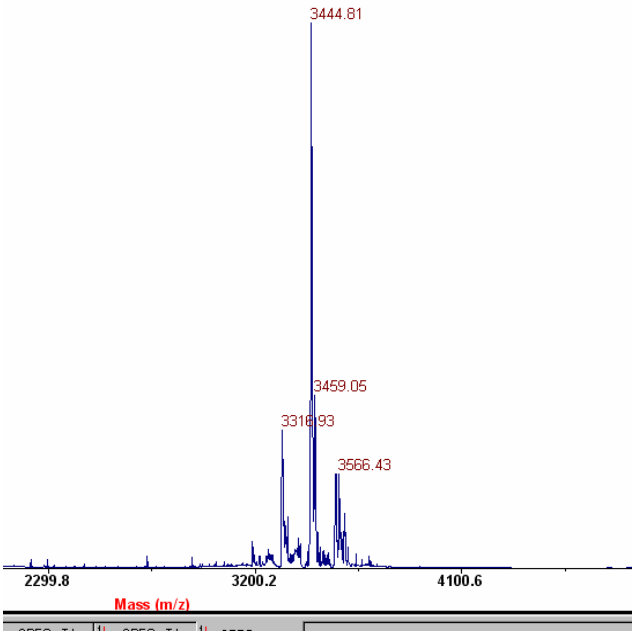
β-ala-ttttttt-Pro-N_3 **5**
Calculated 2605.88
Observed 2628.1(+Na^+)



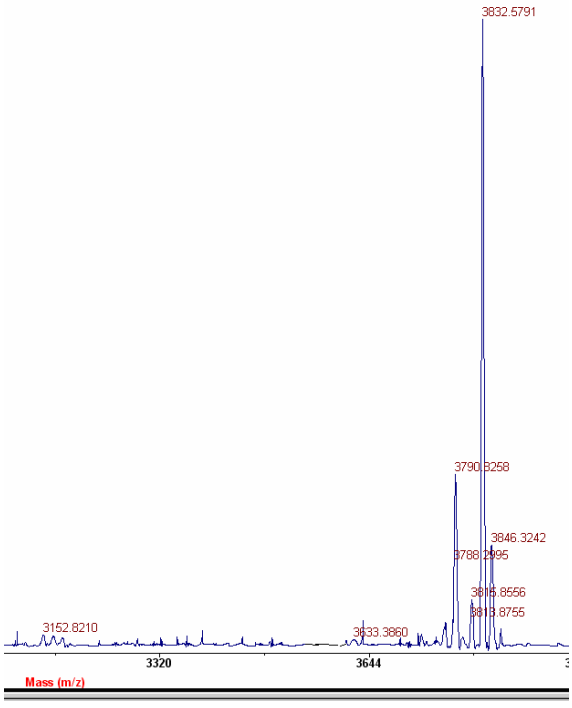
Lys-TCACTAGATG-Pro-N_3 **4**
Calculated 2991.2 Observed 2990.8



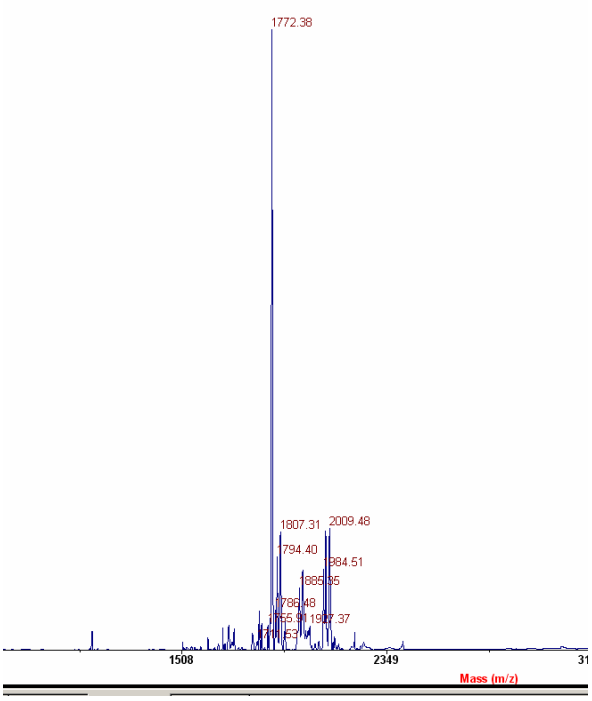
β-ala-TTTTTTTT-Pro-triazole-$(Lys)_6$ **7**
Calculated 3195.3 Observed 3195.35



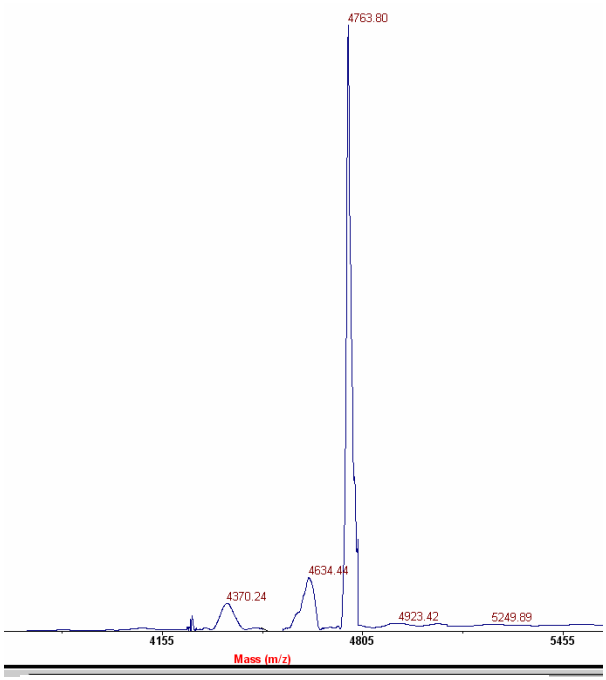
β-ala-t-t-t-t-t-t-t-t-Pro- triazole-$(Lys)_6$ **9**
Calculated 3444.98 Observed 3444.81



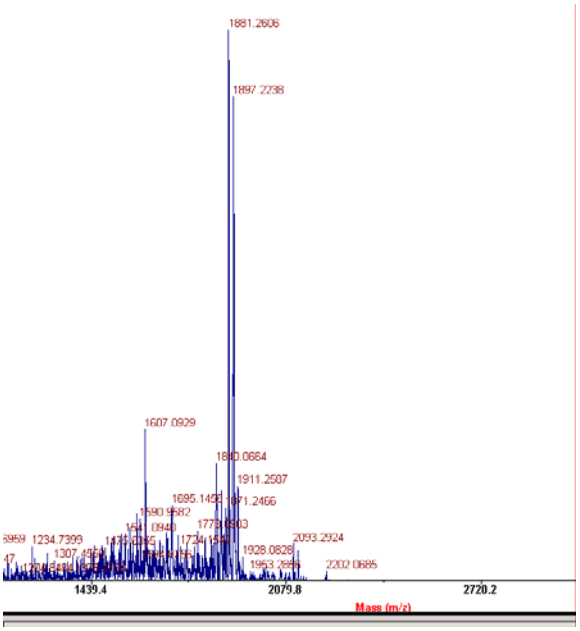
Lys-TCACTAGATG-Pro- triazole-$(Lys)_6$ **8**
Calculated 3830.3 Observed 3832.58



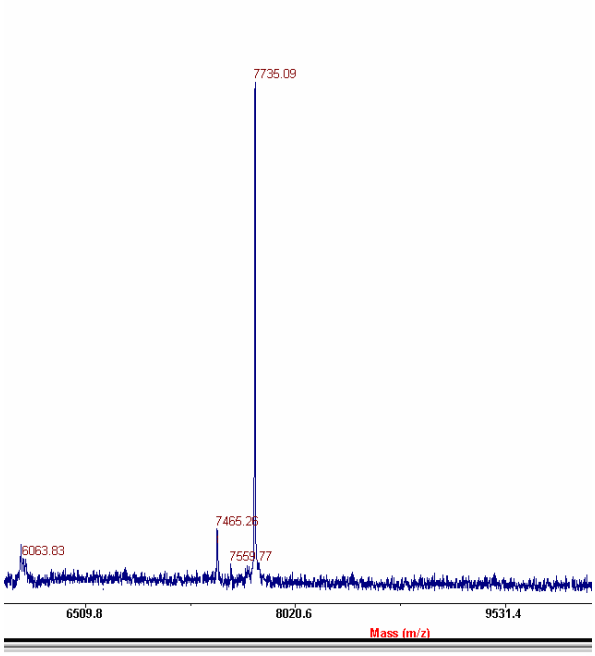
(Arg-Hex-Arg)$_4$-alkyne **2**
Calculated 1772.18 Observed 1772.38



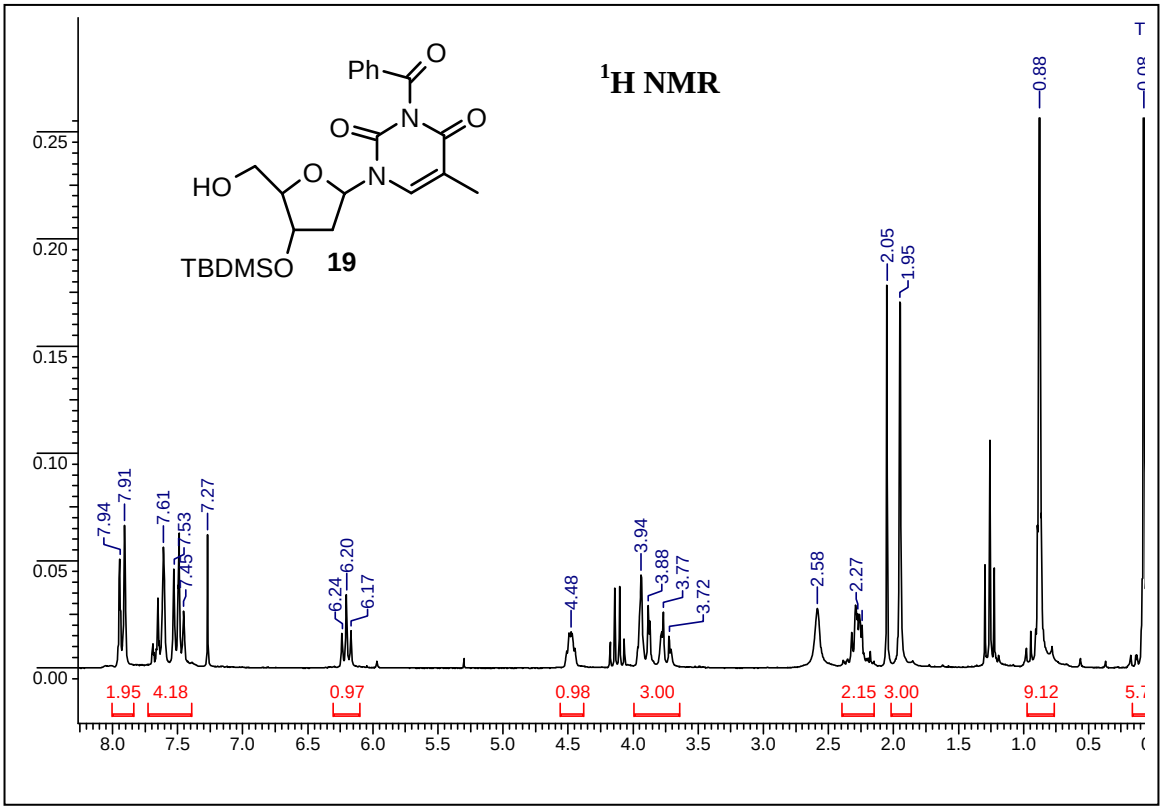
Lys-TCACTAGATG-Pro- triazole-
(Arg-Hex-Arg)$_4$ **10**
Calculated 4763.38 Observed 4763.8

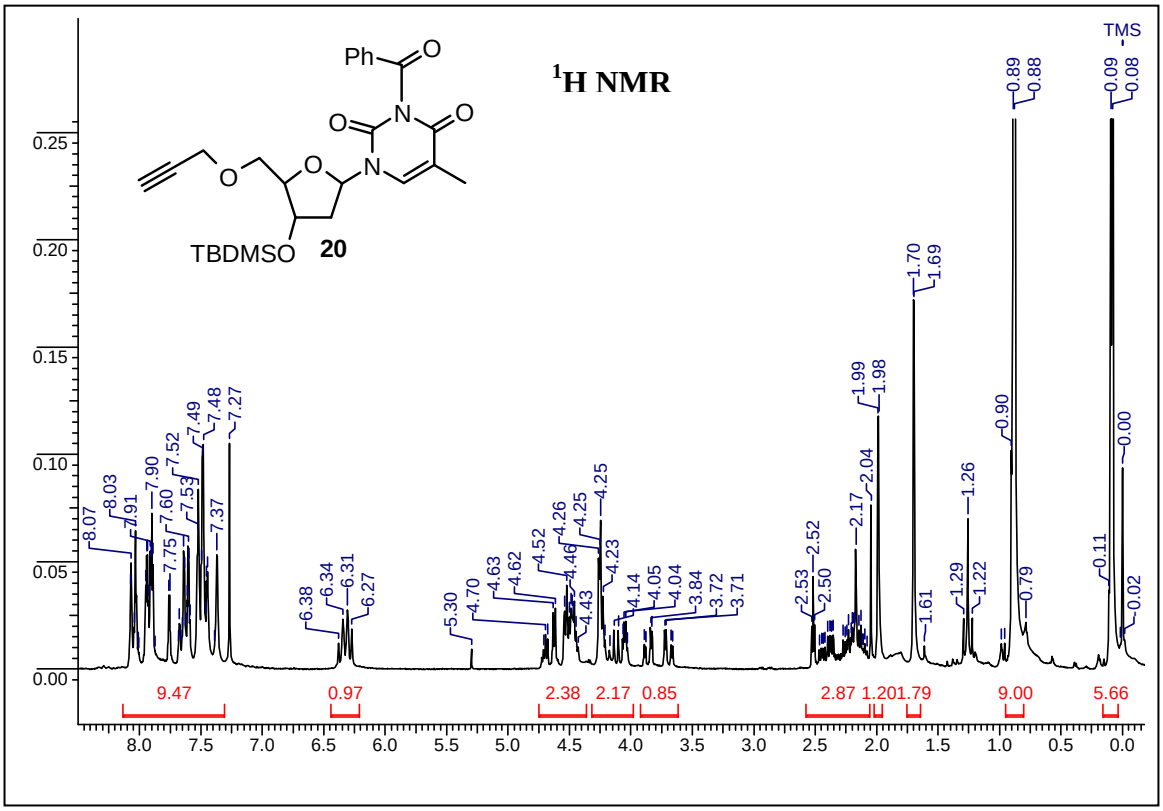


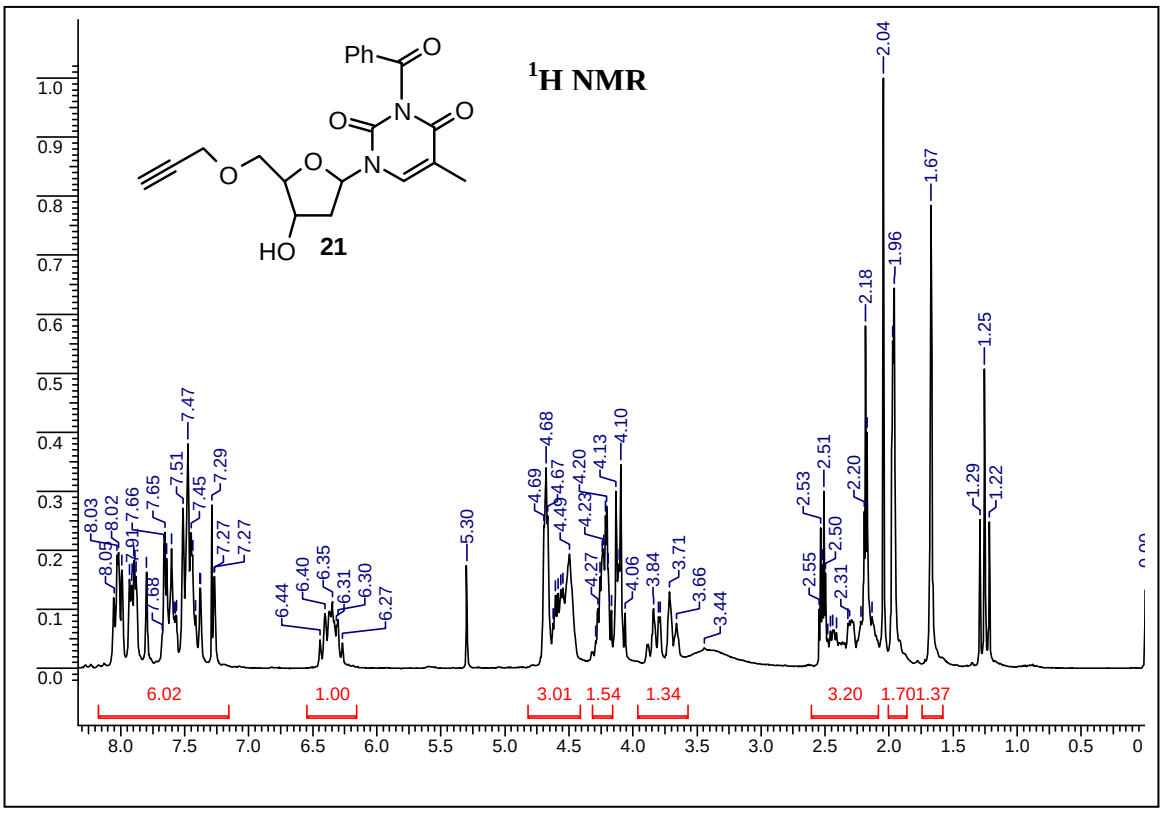
HO- (Arg-Aha-Arg)$_4$- Pro-N$_3$ **13**

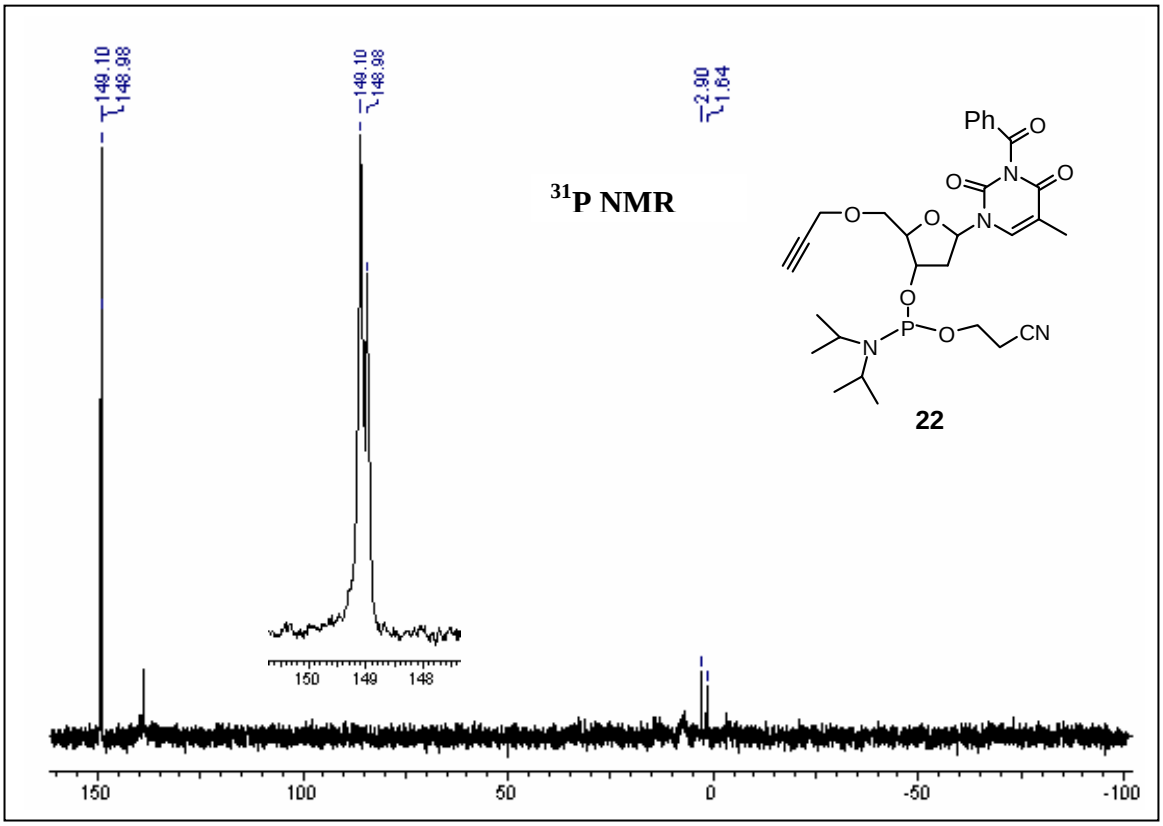


HO- (Arg-Ahx-Arg)$_4$- pro- triazole-
5’-d(TTGTACTGATAGAGTGTCC) 3’ **14**


Ph
O
N
O
O
HO
O
N
TBDMSO
19
^{1}H NMR
7.94
7.91
7.61
7.53
7.49
7.27
6.24
6.20
6.17
4.48
3.94
3.88
3.77
3.72
2.58
2.27
2.05
1.95
0.88
1.95
4.18
0.97
0.98
3.00
2.15
3.00
9.12


Ph
O
N
O
O
O
N
TBDMSO
20
^{1}H NMR
TMS
8.07
8.03
7.91
7.90
7.60
7.53
7.52
7.49
7.48
7.37
7.27
6.38
6.34
6.31
6.27
5.30
4.70
4.63
4.62
4.52
4.46
4.43
4.26
4.25
4.23
4.14
4.05
4.04
3.84
3.72
3.71
2.53
2.52
2.50
2.17
2.04
1.99
1.98
1.70
1.69
1.61
1.29
1.26
1.22
0.90
0.89
0.88
0.79
0.11
0.09
0.08
0.00
0.02
9.47
0.97
2.38
2.17
0.85
2.87
1.20
1.79
9.00
5.66


^{1}H NMR
Ph
O
N
O
O
N
O
O
HO
21
8.03
8.02
7.66
7.65
7.51
7.47
7.45
7.29
7.27
7.27
6.44
6.40
6.35
6.31
6.30
6.27
5.30
4.69
4.68
4.67
4.49
4.23
4.20
4.13
4.10
4.06
3.84
3.71
3.66
3.44
2.55
2.53
2.51
2.50
2.31
2.20
2.18
2.04
1.96
1.67
1.29
1.25
1.22
6.02
1.00
3.01
1.54
1.34
3.20
1.70
1.37


^{31}P NMR
149.10
148.98
149.10
148.98
2.90
1.64
Ph
O
N
O
O
N
O
O
O
P
O
N
CN
22

Sugar–thioacetamide backbone in oligodeoxyribonucleosides for specific recognition of nucleic acids†

Khirud Gogoi, Anita D. Gunjal and Vaijayanti A. Kumar*

Received (in Cambridge, UK) 16th March 2006, Accepted 4th April 2006
First published as an Advance Article on the web 25th April 2006
DOI: 10.1039/b603958h

The amide linkage being shorter than the natural phosphate linkage, an additional atom is introduced into oligodeoxyribonucleosides (ODNs) with sugar–thioacetamide backbone that show very good RNA recognition properties.

The use of modified nucleic acids as gene-targeted drugs and as tools in molecular biology is a developing field.[1] The designed oligodeoxyribonucleosides (ODNs) to be effective leads as drugs need to have enhanced strength of hybridization with target RNA, aqueous solubility, efficiency of cellular uptake, degradation of RNA by RNAs-H enzyme and also easy synthetic methodologies. The current phosphorothioate ODN drugs are a mixture of diastereomeric ODNs due to the chirality at the phosphorus. The design and synthesis should also fulfil the requirement for enantiomeric purity to obtain homogeneous ODNs. The backbone modifications being developed presently include the second and third generation antisense ODNs such as 2′-*O*-alkyl,[2] 2′-*O*-methylthioethyl,[3] LNA,[4] HeNA,[5] morpholino NA,[6] aminoethylglycyl PNA (Fig. 1, *aeg*PNA)[7] or several PNA modifications.[8] The polyamide backbone ODNs such as *aeg*PNA are promising because of the ease of their synthesis and strong binding to the target complementary nucleic acids. Among the sugar–amide backbones, there are many examples in the literature suggesting that a five-atom amide linker leading to a seven-atom repeating backbone may be more useful because of the reduced conformational flexibility of the amide relative to the six-atom phosphodiester backbone.[9] This postulate has also been supported by X-ray studies.[9] Some of the analogues with extended seven-atom backbones cross-pair with RNA with high affinity compared to DNA.[9] Our recent work on PNA modifications (Fig. 1, *bep*PNA) suggested that ODNs containing an extra atom in the pyrrolidine–amide backbone stabilized the complexes with RNA over DNA.[10] In the search for an ideal backbone, we comply that seven-atom extended sugar–amide backbones could be more uniform with respect to nucleobase orientation as well as distance complementarity to bind complementary RNA sequences. We employ a very easy synthetic methodology that fulfills the above mentioned criteria and is much simpler to execute compared to several other reported oligonucleotide analogues so far.[1,11] In this communication we present the synthesis of thioacetamido nucleic acids (TANA) and the thermal stability studies with complementary DNA and RNA sequences. The strategy of the design, synthesis of the monomer blocks, oligomer synthesis and their complementary RNA recognition using UV-T_m measurements is discussed.

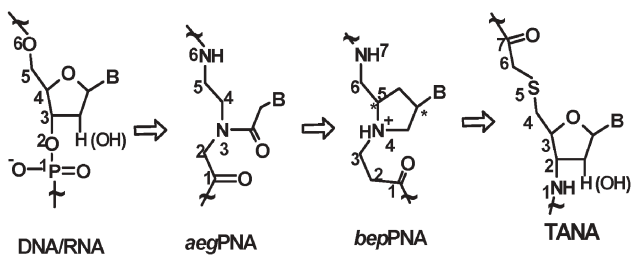
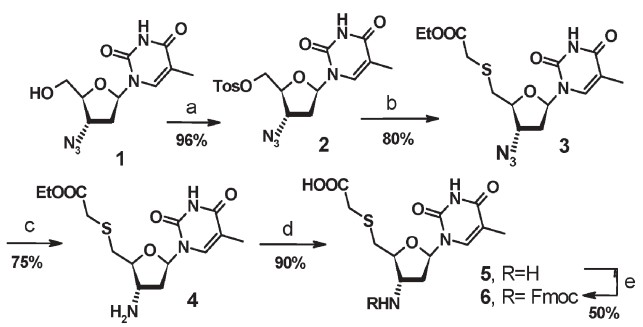


Fig. 1 Designed TANA as a DNA/RNA mimic.

The design is based on the following considerations: (1) in 3′-deoxy-3′amino ribose sugar the five-membered heterocyclic ring pucker is preferred to be 3′-endo[12] which is better suited for mRNA recognition; (2) the solid phase synthesis and scale-up methodology for an amide bond formation are very well established; (3) the linker group has no chiral center; and (4) the flexibility of the phosphate group in DNA/RNA may be conserved by using a five-atom amide linker.

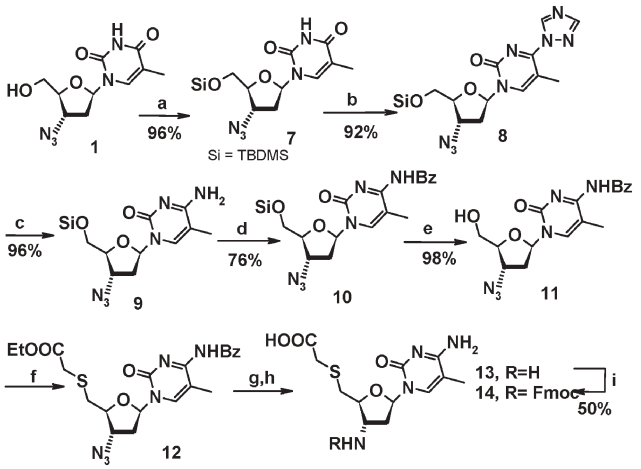
The monomer synthons 3′-Fmoc-amino-5′-thioacetic acid derivative of thymidine **6** and 3′-Fmoc-amino-5′-thioacetic acid derivative of 2′-deoxy cytidine **14** were synthesized from AZT as shown in Scheme 1 and Scheme 2, respectively. The AZT, **1** was converted to its 5′-tosyl derivative **2**. The 3′-azido-5′-tosyl thymidine **2** was converted to 5′-ethyl thioacetate **3** by nucleophilic displacement with ethyl thioacetate in the presence of sodium hydride in very good yields. The azide group was converted to a



Scheme 1 Synthesis of the thyminyl–sugar–amino acid monomer unit. Reagents: (a) tosyl chloride–pyridine; (b) ethyl thioacetate, NaH, DMF; (c) H_2S–pyridine–15%TEA; (d) aqueous LiOH (2 M), MeOH; and (e) Fmoc–succinimide, $NaHCO_3$, acetone–water.

Division of Organic Chemistry (Synthesis), National Chemical Laboratory, Pune, 411008, India. E-mail: va.kumar@ncl.res.in; Fax: +91 20 25602623; Tel: +91 20 25602623

† Electronic supplementary information (ESI) available: ^1H, ^{13}C NMR, mass spectral data for compounds **6** and **14**, HPLC profiles and mass spectra for the modified oligomers **15–20**, and Job's plot and melting curves for the complexes with RNA. See DOI: 10.1039/b603958h



Scheme 2 Synthesis of the 5-methylcytosinyl–sugar–amino acid monomer unit. Reagents: (a) TBDMSCl, imidazole, DMF; (b) 1,2,4 triazole, $POCl_3$, TEA, acetonitrile; (c) conc. NH_3, dioxane; (d) BzCl, pyridine; (e) TBAF, THF; (f) i. tosyl chloride–pyridine and ii. ethyl thioacetate, NaH, DMF; (g) H_2S–pyridine–15%TEA; (h) aqueous LiOH (2 M), MeOH; (i) Fmoc–succinimide, $NaHCO_3$, acetone–water.

3′-amino group in **4** using pyridine–H_2S in 75% yield.[13] The ester function was then hydrolyzed to get the free sugar–amino acid **5**. The free amine group was then protected as Fmoc to get **6**.

The 5-methyl cytosine monomer **14** was synthesized starting from AZT (Scheme 2). AZT **1** was protected as 5′-*O*-*tert*-butyldimethylsilyl ether to get **7**. The nucleobase was then transformed to 5-methyl cytosine derivative **9** *via* C^4-1,2,4 triazolo derivative **8** followed by replacement of the triazolide with ammonia to get **9**.[14] The exocyclic amino group was benzoyl protected to get **10**. 5′-*O*-desilylation, tosylation and conversion to 5′-thioethylacetate derivative of cytosine **11** was accomplished as in the case of synthesis of thyminyl monomer. The azide group was converted to 3′-amino group using pyridine–H_2S in 75% yield to get **12**.[13] The ester function was then hydrolyzed to get free sugar–amino acid **13**. It was found that during the basic hydrolysis of the ester group the exocyclic amino group was deprotected. Protection of the 3′-amino group was then accomplished to get Fmoc–amino acid 5-methyl cytosine derivative **14**. All the new compounds were characterized using spectroscopic techniques such as IR, ^1H, ^{13}C NMR and MALDI-TOF mass spectral analysis.† The protected thymine monomer **6** and cytosine monomer **14** were utilized in the solid phase oligomer synthesis.

To test the sequence specific DNA/RNA recognition by the oligomers comprising the sugar–amino acids **6** and **14**, two 8-mer sequences **15** and **16** were synthesized using rink amide resin and Fmoc peptide synthesis strategy. During the synthesis of sequence **16**, excess of 5-methylcytosine monomer was used in the coupling stage and capping with acetic anhydride was done in order to ensure that the free C^4-amino group remains protected during further synthesis. In order to test the compatibility of these modified monomers in *aeg*PNA, three sequences with mixed TANA–PNA backbone (**17–19**) and control *aeg*PNA **20** were synthesized using Fmoc based solid phase peptide synthesis procedures. The oligomer sequences (Table 1) were cleaved from the solid support using standard conditions and the sequence **16** was further treated with concentrated NH_3 to deprotect the C^4-amino protecting group of the cytosine residue. The uncharged backbone polypyrimidine PNA sequences are known to bind to complementary DNA/RNA in a triplex mode.[15] To establish the binding stoichiometry of the TANA oligomers (**15** and **16**) Job's plot study[16] was undertaken with complementary DNA/RNA sequences.† The binding stoichiometry could be clearly established to be 2 : 1 between sequence **15** and RNA **22** using UV spectroscopy.† The TANA : DNA binding could not be observed in the Job's plot. All the binding studies of the sequences listed in Table 1 with complementary DNA/RNA were carried out in 2 : 1 stoichiometry using UV-T_m experiments. The complexes of TANA **15** with DNA (3′-GCAAAAAAAACG-5′) **21** and RNA r(3′-GCAAAAAAAACG-5′) **22** were used to study their comparative binding efficiency. The TANA **16** has a 5-methyl cytosine unit and its binding with complementary DNA (3′-GCAAAAGAAACG-5′) and RNA sequence r(3′-GCAAAAGAAACG-5′) **23** was studied. For single nucleotide mismatch studies the UV-T_m measurements were carried out using sequences **15** : **23** and **16** : **22** (see Supplementary information,† normalized UV absorbance *vs.* temperature plots). The results are summarized in Table 2. For control, the T_m of complexes of DNA 3′-TTTTTTTT-5′ **24** and *aeg*PNA **20** with DNA **21** and RNA **22** were measured under identical buffer conditions. It is evident from the above experiments that the homopyrimidine TANA backbone oligomers (**15** and **16**) bind to the homopurine complementary RNA sequences (**22** and **23**, respectively) and show a single base-mismatch discrimination of about 10–14 °C. The thermal stability of **16** : **23** was increased by

Table 1 The TANA and TANA–PNA oligomer sequences

Sequence[a]	RP-HPLC[b]	$(Mass_{calc})/(Mass_{obs})$
H-**t t t t t t t t**-βala **15**	12.7	2467.6/2491.0 (M + Na^+)
H-**t t t t c t t t**-βala **16**	12.3	2466.1/2487.9 (M + Na^+)
H-t t t t t t t **t**-βala **17**	10.0	2249.4/2251.7 (M + H^+)
H-t t t **t** t t t t-βala **18**	8.3	2249.4/2252.6 (M + H^+)
H-t **t** t t t t t **t**-βala **19**	10.1	2342.9/2367.9 (M + Na^+)
H-t t t t t t t t-βala **20**	7.6	2218.2/2218.9 (M + H^+)

[a] **t** and **c** denote TANA backbone, t denotes *aeg*PNA backbone. [b] Retention time in minutes.

Table 2 UV-T_m values[a] in °C of $(TANA)_2/(TANA\text{-}PNA)_2$: DNA/RNA triplexes[b]

No	Sequence	DNA **21**	RNA **22**	RNA **23**
1	H-**t t t t t t t t**-βala **15**	nd[d]	63.8 (63.5)[c]	49.8
2	H-**t t t t c t t t**-βala **16**	nd[e]	52.9	63.1 (67.1)[c]
3	H-t t t t t t **t**-βala **17**	nd	53.5	—
4	H-t t t t **t** t t t-βala **18**	nd	32.4	—
5	H-t **t** t t t t **t**-βala **19**	nd	25.5	—
6	H-t t t t t t t t-βala **20**	42.6	61.5	52.5
7	3′-TTTTTTTT-5′ **24**	nd	20	nd

[a] T_m = melting temperature (measured in buffer: 10 mM sodium phosphate, pH 7.0 with 100 mM NaCl and 0.1 mM EDTA). Measured from 10 to 90 °C at ramp 0.2 °C min^{-1}. UV-absorbance measured at 260 nm. All values are an average of 3 independent experiments and accurate to within ±0.5 °C. [b] T denotes DNA backbone. [c] Values in parentheses denote melting temperature at pH 5.5. [d] nd = not detected. [e] With complementary DNA sequence (3′-GCAAAAGAAACG-5′).

4 °C at pH 5.5 in which a C^+GC base triple is formed.[15] The complex of TANA oligomer **15** with cRNA **22** was highly stable compared to the DNA : RNA complex (**24** : **22**) and the stability was comparable with that of *aeg*PNA_2 : RNA (**20** : **22**). The formation of cDNA : TANA complexes was not observed. To test the compatibility of the TANA backbone in *aeg*PNA : TANA mix-backbone oligomers, the sequences **17–19** were synthesized. Sequence **17** has a TANA monomer unit at C-terminus, sequence **18** has the modified unit at the central position and sequence **19** is an alternate *aeg*-TANA sequence. All the complexes of oligomers **17–19** with RNA **22** were destabilized compared to the control *aeg*PNA : RNA complex **20** : **22**. Contrary to the earlier reported *bep*PNA backbone,[11] the mixed backbone containing sugar amide and *aeg*PNA seems to be incompatible to form regular helical structures and mixed backbone ODNs **17–19** show significant decrease in binding with cRNA. The stability of the complexes of homogeneous backbone TANA with RNA over DNA is a very valuable result from an application perspective. The synthesis of mixed purine–pyrimidine backbone sequences and the study of compatibility of the TANA dimers in a regular phosphodiester backbone will be interesting and the work in this direction is currently in progress.

This communication presents the results of selective RNA recognition by a homogeneous DNA analogue and the ease of its synthesis that has potential to be extended to purine nucleosides. The thio function in the backbone may have an additional advantage for better bioavailability of these modified oligodeoxyribonucleosides.[3] The simplicity of the approach described here is equally appealing as is the selectivity.

K. G. thanks CSIR, New Delhi for a research fellowship. V. A. K. thanks Director, NCL for financial assistance.

Notes and references

1 D. A. Braasch and D. R. Corey, *Biochemistry*, 2002, **41**, 4503–4510; P. E. Nielsen and M. Egholm, in *Peptide Nucleic Acids (PNA). Protocols and Applications*, ed. P. E. Nielsen and M. Egholm, Horizon Scientific, Norfolk, CT, 1999; S. T. Crooke, in *Antisense Research and Application*, ed. S. T. Crooke, Springer, Berlin, 1998, vol. 131, pp. 1–50; J. Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628–1644.
2 M. Manoharan, *Biochim. Biophys. Acta*, 1999, **1489**, 117–130.
3 T. P. Prakash, M. Manoharan, A. M. Kawasaki, A. S. Fraser, E. A. Lesnik, N. Sioufi, J. M. Leeds, M. Teplova and M. Egli, *Biochemistry*, 2002, **41**, 11642–11648.
4 M. Petersen and J. Wengel, *Trends Biotechnol.*, 2003, **21**, 74–81.
5 E. Lescrinier, R. Esnouf, J. Schraml, R. Busson, H. A. Heus, C. W. Hilbers and P. Herdewijn, *Chem. Biol.*, 2000, **7**, 719–731.
6 J. Summerton, *Biochim. Biophys. Acta*, 1999, **1489**, 141–158.
7 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500; P. E. Nielsen, M. Egholm and O. Buchardt, *Bioconjugate Chem.*, 1994, **5**, 3–7.
8 V. A. Kumar, *Eur. J. Org. Chem.*, 2002, 2021–2032; V. A. Kumar and K. N. Ganesh, *Acc. Chem. Res.*, 2005, **38**, 404–412.
9 J. Wilds, G. Minasov, F. Natt, P. Von Matt, K.-H. Altamann and M. Egli, *Nucleosides Nucleotides Nucleic Acids*, 2001, **20**, 991–994; G. V. Petersen and J. Wengel, *Tetrahedron*, 1995, **51**, 2145–2154; A. Lauritsen and J. Wengel, *Chem. Commun.*, 2002, 530–531; T. Vilaivan and G. Lowe, *J. Am. Chem. Soc.*, 2002, **124**, 9326–9327.
10 T. Govindaraju and V. A. Kumar, *Chem. Commun.*, 2005, 495–497; T. Govindaraju and V. A. Kumar, *Tetrahedron*, 2006, **60**, 2321–2330.
11 S. M. Frier and K.-H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429–4443.
12 C. Thibaudeau, J. Plavec, N. Garg, A. Papchikhin and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 1994, **116**, 4038–4043.
13 M. Saneyoshi, T. Fujii, T. Kawagchi, K. Awai and S. Kimura, in *Nucleic Acids Chemistry*, ed. L. B. Townsend and R. S. Tipson, Wiley, New York, 1991, vol. 4, pp. 67–72.
14 K. J. Divakar and C. B. Reese, *J. Chem. Soc., Perkin Trans. 1*, 1982, 1171–1176.
15 S. K. Kim, P. E. Nielsen, M. Egholm, O. Buchardt, R. H. Berg and B. Norden, *J. Am. Chem. Soc.*, 1993, **115**, 6477–6481.
16 P. Job, *Ann. Chim.*, 1928, **9**, 113–203.

ORGANIC LETTERS

2007
Vol. 9, No. 14
2697−2700

Synthesis and RNA Binding Selectivity of Oligonucleotides Modified with Five-Atom Thioacetamido Nucleic Acid Backbone Structures

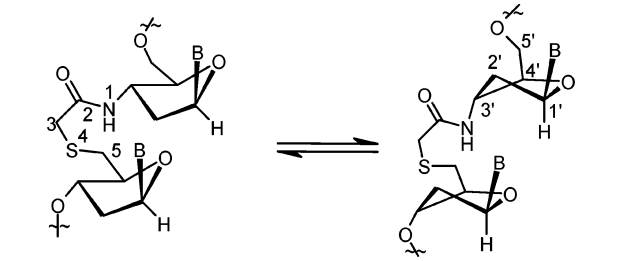
Khirud Gogoi, Anita D. Gunjal, Usha D. Phalgune, and Vaijayanti A. Kumar*

Division of Organic Chemistry, National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, India

va.kumar@ncl.res.in

Received April 27, 2007

ABSTRACT



Convenient chemical synthesis and incorporation of dithymidine and thymidine–cytidine dimer blocks connected with a five-atom amide linker N3′–CO–CH_2–S–CH_2 into oligonucleotides (ONs) are reported. The UV–T_m experiments for binding affinities of these mixed backbone ONs with complementary DNA and RNA sequences revealed important results such as significantly higher RNA-binding selectivity as compared with complementary DNA. NMR studies of the dimer blocks suggested a marginal increase in the N-type sugar conformations over that of the native DNA.

The last two decades have witnessed an upsurge in the synthesis of several modified nucleic acid derivatives. The intentions have been to synthesize therapeutically suitable and commercially viable nucleic acid analogues.[1] The more recent developments such as splice correcting and exon skipping strategies require highly robust nucleic acid analogues[2] that are stable under physiological conditions as single strands as well as in the form of duplexes with complementary RNA sequences. The enzymatic stability[3] of peptide nucleic acids (PNAs) gains importance for such applications.[2] The uncharged PNAs (Figure 1) are poorly soluble in water and bind to DNA in either parallel or antiparallel orientation.[4] Also, PNAs require assistance in the form of covalent conjugation with cell-penetrating peptides (cpp)[5] or additional positive charges on PNA[6,7] to cross the cell membranes for their activity. The replacement of a few or several internucleoside phosphate linkages by the robust amide bond as in PNA could be an alternate solution so that the advantages of chirality and 3′−5′

(1) (a) Micklefield, J. *Curr. Med. Chem.* **2001**, *8*, 1157. (b) Kurreck, J. *Eur. J. Biochem.* **2003**, *270*, 1628. (c) Turner, J. J.; Fabani, M.; Arzumanov, A. A.; Ivanova, G.; Gait, M. J. *Biochim. Biophys. Acta* **2006**, *1758*, 290. (d) Kumar, V. A.; Ganesh, K. N. *Curr. Top. Med. Chem.* **2007**, *7*, 715.

(2) (a) Dominski, Z.; Kole, R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8673. (b) Sazani, P.; Kole, R. *J. Clin. Invest.* **2003**, *112*, 481.

(3) (a) Egholm, M.; Buchardt, O.; Nielsen, P. E. *J. Am. Chem. Soc.* **1992**, *114*, 1895−1897. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioconjugate Chem.* **1994**, *5*, 3.

(4) (b) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem., Int. Ed.* **1998**, *37*, 2796.

(5) Bendifallah, N.; Rasmussen, F. W.; Zachar, V.; Ebbesen, P.; Nielsen, P. E.; Koppelhus, U. *Bioconjugate Chem.* **2006**, *17*, 750.

(6) (a) Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6878. (b) Englund, E. A.; Appella, D. H. *Angew. Chem., Int. Ed.* **2007**, *46*, 1414.

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

10.1021/ol070990u CCC: $37.00
Published on Web 06/07/2007

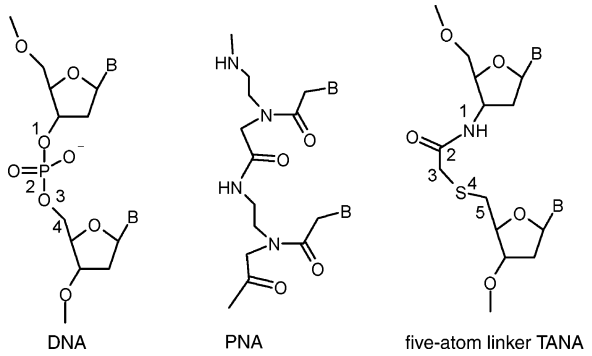


Figure 1. Structures of DNA, PNA, and TANA.

directionality of the sugar are maintained in the analogue along with the enzymatic stability of the amide bond. The retention of some negatively charged phosphate groups would be advantageous for water solubility. Cell-penetration experiments then could be performed using cationic lipids,[8] complexation, or conjugation with cpp.[8] Several four- and five-atom amide-linked oligonucleotide analogues and those containing thioether linkages are known in the literature.[1,9] The five-atom amide linkers were introduced to compensate the shorter amide bond in the dinucleoside as compared with the four-atom phosphate linkers to maintain the internucleoside distance complementarity.[9] The strategy was found to be successful as some of the five-atom-linked dimers when substituted in oligonucleotide (ON) sequences lead to significantly higher RNA affinities compared with that of the native DNA. It has been shown by CD spectroscopy and crystal structure that the longer amide backbones do not disrupt the duplex geometries.[10] The application of these ONs with superior binding properties could be stymied by the multistep synthetic procedures to arrive at the desired monomers. Our continuing efforts toward finding probable optimum oligonucleotide mimics resulted in a new five-atom thioacetamido nucleic acid (TANA, Figure 1) backbone.[11] The convenient synthetic methodology to convert pyrimidine nucleosides to a sugar−amino acid monomer unit by using thiol functionality in ethyl mercaptoacetate as a requisite nucleophile was reported in our previous communication. The homooligomeric pyrimidine ONs were found to bind to complementary RNA sequences significantly better than their DNA counterparts, and the binding efficiency was found to be as good as PNA itself. However, introduction of these amino acid nucleoside derivatives in PNA sequences was found to be detrimental for RNA/DNA binding. In this communication, we present the synthesis and incorporation of thymidine and thymidine−cytidine dimer blocks (**tst** and **cst**) connected with a five-atom amide linker N3′−CO−CH_2−S−CH_2 (TANA) (Figure 2) into oligonucleotide oligomers. The assessment of the compatibility of the TANA dimer blocks in a sugar−phosphate backbone was studied by UV−T_m measurements of the resulting mixed-backbone ON complexes with DNA and RNA.

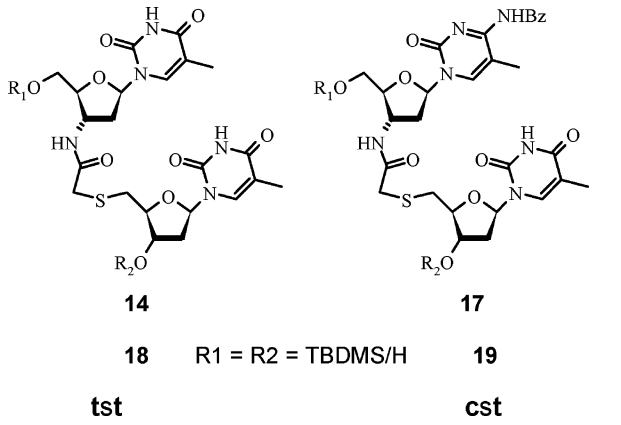
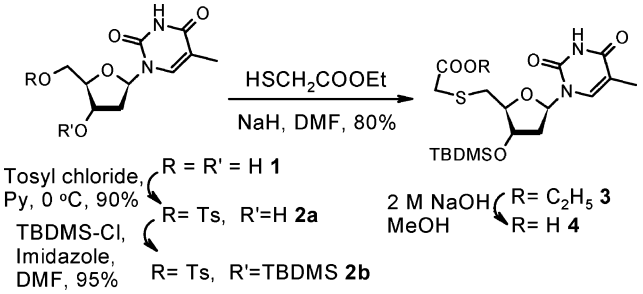


Figure 2. **tst** and **cst** building blocks.

Thymidine was selectively tosylated at the 5′ position, and the 3′-OH was then protected as a 3′-*O*-TBDMS group. Treatment of 3′-O-protected 5′-*O*-tosyl-thymidine with ethyl mercaptoacetate followed by ester hydrolysis gave the acid synthon **4** (Scheme 1). The amine synthon of thymidine was synthesized from 3′-azidothymidine **5** by protection of the free hydroxy group as 5′-*O*-DMTr in **6** and reduction of azide to amine to get **7**. The 5′-*O*-DMTr, 3′-azidothymidine **6** was converted[12] to a cytidine derivative by known procedures via C4-triazolide (**8**) followed by amination (**9**) and benzoylation (**10**), and reduction of the azide gave the desired

Scheme 1. Synthesis of a Thymidine Monomer Unit



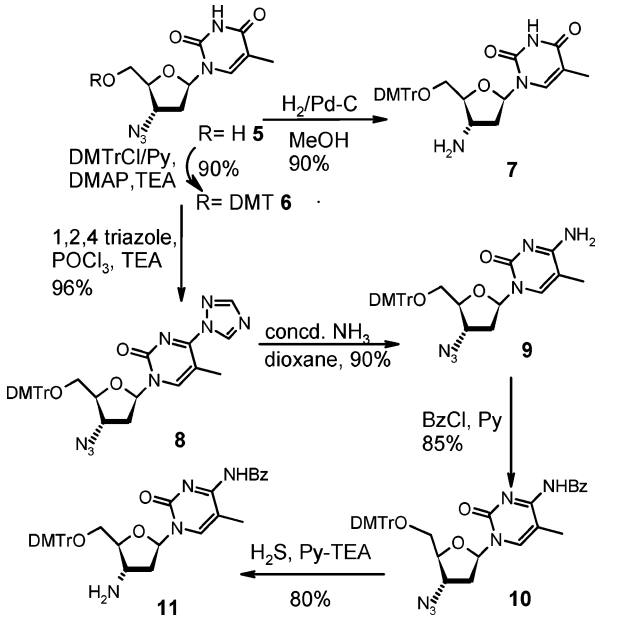
(8) (a) Venkatesan, N.; Hyean Kim, B. *Chem. Rev.* **2006**, *106*, 3712. (b) Turner, J. J.; Jones, S.; Fabani, M. M.; Ivanova, G.; Arzumanov, A. A.; Gait, M. J. *Blood Cells, Mol. Dis.* **2007**, *38*, 1.

(9) (a) De Mesmaeker, A.; Lesueur, C.; Btvikrre, M.-O.; Waldner, A.; Fritsch, V.; Wolf, R. M. *Angew. Chem., Int. Ed.* **1996**, *35*, 2790. (b) Nina, M.; Fonne-Pfister, R.; Beaudegnies, R.; Chekatt, H.; Jung, P. M. J.; Murphy-Kessabi, F.; Mesmaeker, A.; Wendeborn, S. *J. Am. Chem. Soc.* **2005**, *127*, 6027. (c) Govindaraju, T.; Kumar, V. A. *Chem. Commun.* **2005**, 495. (d) Govindaraju, T.; Kumar, V. A. *Tetrahedron* **2006**, *60*, 2321. (e) Meng, B.; Kawai, S. H.; Wang, D.; Just, G.; Giannaris, P. A.; Damha, M. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 729. (f) Damha, M. A.; Meng, B.; Kawai, S. H.; Wang, D.; Yannopoulos, C. G.; Just, G. *Nucleic Acids Res.* **1995**, *23*, 3967.

(10) (a) Pallan, P. S.; von Matt, P.; Wilds, C. J.; Altmann, K.-H.; Egli, M. *Biochemistry* **2006**, *45*, 8048. (b) Wilds, C. J.; Minasov, G.; von Matt, P.; Altmann, K.-H.; Egli, M. *Nucleosides, Nucleosides Nucleic Acids* **2001**, *20*, 991.

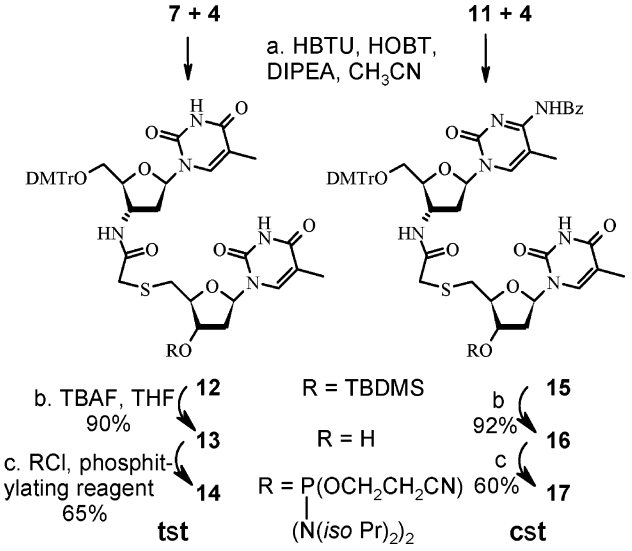
(11) Gogoi, K.; Gunjal, A. D.; Kumar, V. A. *Chem. Commun.* **2006**, 2373.

Scheme 2. Synthesis of a Cytidine Monomer Unit



cytidine derivative **11** (Scheme 2). The synthesis of the dimer building blocks **tst 14** and **cst 17** was then achieved by using peptide-coupling chemistry from the monomer units, deprotection, and phosphitylation (Scheme 3). All the new

Scheme 3. Synthesis of **tst** and **cst** Dimer Blocks



compounds were characterized by ^{1}H NMR, ^{13}C NMR, and mass spectral analysis. The structures of the dimer blocks **18** and **19** (Figure 2) were extensively studied by using 2D NOESY, COSY, and TOCSY ^{1}H spectral analysis. The sugar ring conformations in the dinucleosides were elucidated using an empirical method of analysis of the vicinal coupling constants $^3J_{1'2'}$ and $^3J_{1'2''}$.[13] The population of the 2′-*endo* sugar conformations (% S) was calculated from the equation used in this analysis.[20]

A series of chimeric ONs (Tables 1 and 2) containing one to four **tst** and **cst** blocks were synthesized by automated solid-phase synthesis using the phosphoramidite approach and an Applied Biosystems 3900 DNA Synthesizer. After cleavage from the support, the oligomers were purified by

Table 1. Modified ON−Pyrimidine Sequences and UV−T_m of Complexes with Complementary DNA and RNA[a]

	ON sequences mass (calcd/obsd)	ON:DNA T_m (°C)	ON:RNA T_m (°C)
1	5′ CGTTTTTTTTGC **33**	**33:34**	**33:35**
		40	32.0 (−8.0)[b]
2	5′ CG TT**tst**TTT TGC **20**	**20:34**	**20:35**
	(3606.5/3606.6)	23.7	32.3 (+8.6)
3	5′ CGTT **tst** TT **tst** GC **21**	**21:34**	**21:35**
	(3596.6/5597.0)	nd	50.0
4	5′ CG **tst tst tst tst** GC **22**	**22:34**	**22:35**
	(3576.8/3576.6)	nd	47.81
5	5′ TCT CTT TCT T **39**	**39:40**	**39:41**
		23.6	25.4 (+0.8)
6	5′ TCT C **tst** TCT T **23**	**23:40**	**23:41**
	(2933.1/2933.1)	nd	29.6
7	5′ TCT C **tst** TC **tst 24**	**24:40**	**24:41**
	(2923.2/2924.0)	nd	33.8
8	5′ T**cst** CTT TCTT **29**	**29:40**	**29:41**
	(2948.2/2948.1)	19.5	26.2
9	5′ T **cst** CTT T **cst** T **30**	**30:40**	**30:41**
	(2938.3/2938.7)	nd	33.8
10	5′ T**cst cst** T T **cst** T **31**	**31:40**	**31:41**
	(2928.4/2928.7)	nd	39.7

[a] DNA (**34**, **40**) and RNA (**35**, **41**) sequences **34** 5′ GCAAAAAAAACG 3′, **40** 5′AAG AAA GAG A3′, **35** r (5′ GCAAAAAAAACG 3′), and **41** r (5′ AAG AAA GAG A 3′). [b] Figures in parentheses indicate the difference in T_m between complexes with RNA and DNA.

Table 2. Modified ON−Purine−Pyrimidine Sequences and UV−T_m of Complexes with Complementary DNA and RNA[a]

	ON sequence (mass cald/obsd)	ON:DNA T_m (°C)	ON:RNA T_m (°C)
1	5′ CCT CTT ACC TCA GTT ACA	**36:37**	**36:38**
	36	54.6	54.7 (+0.1)
2	5′ CCT C **tst** ACC TCA G TT ACA	**26:37**	**26:38**
	26 (5366.7/5366.8)	39.6	47.5 (+7.9)
3	5′ CCT C **tst** ACC TCA G **tst** ACA	**27:37**	**27:38**
	27 (5356.9/5356.9)	43.5	52.8 (+9.3)
4	5′ TCA CTA GAT G	**42:43**	**42:44**
	42	24.3	24.7 (0.4)
5	5′ TCA **cst** A GAT G 3′	**32:43**	**32:44**
	32 (3036.2/3037.0)	16.3	29.6 (+13)

[a] DNA (**37**, **43**) and RNA (**38**, **44**) sequences **37** 5′ TGT AAC TGA GGT AAG AGG 3′, **43** 5′CAT CTA GAG A3′, **38** r (5′ UGU AAC UGA GGU AAG AGG 3′), and **44** r (5′CAU CUA GAG A3′). [b]Figures in parentheses indicate the difference in T_m between complexes with RNA and DNA.

(12) Divakar, K. J.; Reese, C. B. *J. Chem. Soc., Perkin Trans. 1* **1982**, 1171.

gel filtration and reverse-phase HPLC. The purity of the oligomers was checked by reverse-phase HPLC analysis on a C18 column and was characterized by mass spectrometry. These ONs were tested for their binding affinity to complementary DNA and RNA sequences in thermal denaturation−UV measurement experiments and the data are summarized in Tables 1 and 2. The unmodified sequence GCT_8CG **33** was found to form complexes with both cDNA and RNA with a higher melting temperature for the ON:DNA complex over ON:RNA ($\Delta T_m = -8$).[14] Introduction of a single thioacetamido-linked **tst** dimer unit in **20** reversed this selectivity, and the ON:RNA duplex, **20:35**, was more stable ($\Delta T_m = +8.6$) than its complex with cDNA (**20:34**). A cumulative effect was observed in stabilizing the ON:RNA complex, **21:35**, and at the same time, the complex with DNA (**21:34**) was destabilized, when two units of modified **tst** dimer were incorporated in the ON. Similarly, the ON sequence **22** with alternating phosphate−TANA linkers exhibited binding only with RNA **35** and no observable melting transition with cDNA **34**. The unmodified mixed pyrimidine ON **39** formed complexes with either cDNA or RNA and exhibited almost equal binding strength. The preferential binding to RNA was consistently observed when one or two phoshate linkers were replaced by TANA by incorporation of either one or two modified (**tst** or **cst**) dimer units, almost independent of their position in the ON. Inclusion of two or more modified units in ONs leads to the significant stabilization of their complexes with complementary RNA, whereas complexes with the DNA counterpart did not show a detectable transition (Table 1, entries 3, 4, 7, 9, and 10). To verify the usefulness of these modified units in the mixed purine−pyrimidine sequence context, we synthesized two different unmodified sequences **36** and **42** (Table 2). The 18mer sequence **36** was modified by introduction of one **tst** unit in **26** and two **tst** units in **27**. The unmodified 18mer **36** recognized both cDNA **37** and RNA **38** with equal affinity ($\Delta T_m = +0.1$). One **tst** unit caused destabilization of complexes with both DNA (**36:37**) and RNA (**36:37**), but to a much lesser extent with RNA than with DNA. The discrimination between RNA and DNA recognition was observed with a ΔT_m of about 8 °C. Introduction of two **tst** units in **27** increased this stabilization of ON:DNA/RNA (**27:37**/**27:38**) complexes and RNA vs DNA discrimination to 9.3 °C. In a 10mer ON **32**, a single **cst** unit caused stabilization of the ON:RNA complex (**32:44**) compared to the control of the unmodified complex (**42:44**) and destabilized the ON:cDNA complex (**32:43**, Table 2, entries 4 and 5). The binding was found to be sequence specific as a single mismatch in the target RNA highly destabilized the modified DNA:RNA complexes.[20] Thus, the modified units consistently destabilized the complex formation of the modified ONs with cDNA and stabilized the ON:RNA complexes.

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.[15] The electronegativity of the 3′-substituents in 3′-*N*-phoramidates was shown to set the sugars in a preferred N-type conformation but, unlike LNA, to show preferential binding to RNA over DNA.[14,16] Several other examples in the literature such as 2′−5′ DNA[17]/RNA[18] prefer to bind to RNA over DNA, and it is not therefore entirely certain which factors differentiate the DNA vs RNA selectivity.[1d] 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 the 3′-amino or 5′-thioacetamido sugars to be similar to the native 3′−5′ phosphate linked DNA.[20] The structural similarity between unmodified and modified RNA:DNA complexes was also evident by CD studies.[20] The RNA selectivity of binding seems to be arising from the extended backbone linker that is probably inherently folded to be competent to bind to RNA over DNA as was found with the reported five-atom-linked ON analogues.[10,19] The **tst** and **cst** dimer blocks were found to be compatible in the DNA backbone to selectively stabilize the ON:RNA complexes. Further work to exploit their utility is currently in progress in our laboratory. The preferential sequence-independent RNA binding ability of these evolved modified ONs will find applications in current antisense research.

Acknowledgment. K.G. thanks CSIR, New Delhi, for a senior research fellowship, and V.A.K. thanks DST, New Delhi, for a research grant. We acknowledge a generous gift of phosphitylating reagent from Innovassynth Technologies (I) Ltd., Khopoli, India.

Supporting Information Available: Experimental and spectral data of the compounds in Schemes 1−3. ^{1}H NMR of dinucleosides. Mass and ^{31}P spectra of **14** and **17**. Mass spectra and CD and UV−T_m plots of the synthetic ONs with complementary DNA/RNA/mismatched RNA. This material is available free of charge via the Internet at http://pubs.acs.org.

OL070990U

(13) Rinkel, L. J.; Altona, C. J. *Biomol. Struct. Dyn.* **1987**, *4*, 621.
(14) Nawrot, B.; Boczkowska, M.; Wojcik, M.; Sochaki, M.; Kazmierski, S.; Stec, W. J. *Nucleic Acids Res.* **1998**, *26*, 2650.
(15) Petersen, M.; Wengel, J. *Trends Biotechnol.* **2003**, *21*, 74.
(16) Gryaznow, S.; Chen, J.-K. *J. Am. Chem. Soc.* **1994**, *116*, 3143.
(17) Prakash, T. P.; Kraynack, B.; Baker, B. F.; Swayze, E. E.; Bhat, B. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3238.
(18) (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.
(19) Alternatively, in duplexes, the extra distance between sugars might be compensated by compact N-type sugar conformations that favor RNA binding (we thank an annonymous referee for this useful suggestion).
(20) See Supporting Information.

Nucleic Acids Research Advance Access published November 2, 2007

Nucleic Acids Research, 2007, 1–7
doi:10.1093/nar/gkm935

A versatile method for the preparation of conjugates of peptides with DNA/PNA/analog by employing chemo-selective click reaction in water

Khirud Gogoi, Meenakshi V. Mane, Sunita S. Kunte and Vaijayanti A. Kumar*

Division of Organic Chemistry, National Chemical Laboratory, Pune, India, 411008

Received August 14, 2007; Revised September 21, 2007; Accepted October 11, 2007

ABSTRACT

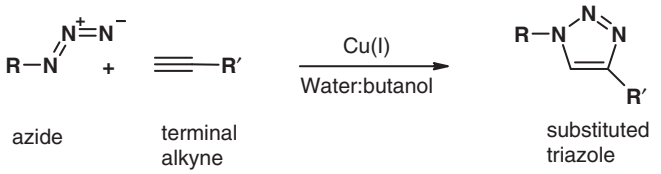
The specific 1,3 dipolar Hüisgen cycloaddition reaction known as 'click-reaction' between azide and alkyne groups is employed for the synthesis of peptide–oligonucleotide conjugates. The peptide nucleic acids (PNA)/DNA and peptides may be appended either by azide or alkyne groups. The cycloaddition reaction between the azide and alkyne appended substrates allows the synthesis of the desired conjugates in high purity and yields irrespective of the sequence and functional groups on either of the two substrates. The versatile approach could also be employed to generate the conjugates of peptides with thioacetamido nucleic acid (TANA) analog. The click reaction is catalyzed by Cu (I) in either water or in organic medium. In water, ~3-fold excess of the peptide-alkyne/azide drives the reaction to completion in 2 h with no side products.

INTRODUCTION

Uncharged, achiral peptide nucleic acids (PNAs) are DNA mimics that show unprecedented affinity towards complementary RNA and DNA sequences(1,2). PNA and other modified oligonucleotides (ONs) (3–7) are currently being developed as DNA mimics to target disease-causing mRNA (8,9), using the principle of antisense action. This is gaining further importance because of corrective antisense therapies that do not require activation of RNase H enzyme or cleavage of the target mRNA (10). The success or failure of any such candidate in antisense therapeutics depends on a number of factors such as sequence-specific recognition of target mRNA, intracellular stability, water solubility (3) and cell penetration (11). Application of PNA and analogous uncharged DNA mimics is stymied by the fact that PNAs show very low cell penetration for any observable antisense effect (12). Several strategies are being developed for the delivery of modified ONs into cells (13,14). For uncharged ON mimics such as PNAs, the best option seems to be the covalent conjugation of PNA oligomers with cell-penetrating peptides (CPP) (15–17). Conjugation of peptides to DNA and DNA analogs with sugar-phosphate backbone ONs is also gaining importance for their biological applications (13,14). The CPPs are mostly positively charged peptides containing lysine (18) or arginine (19,20) or other peptides having specific cell receptors (21). Several other uses of peptide conjugation with ONs are known in the literature(13). The conjugation is achieved by tedious continuous solid-phase synthesis of PNA and peptide (22–25). The other methods for conjugation could be post-synthetic via disulfide bridge (26,27) or more stable thioether linkages (21) at either C- or N- terminus. A recent review (13) not only summarizes the present methods of synthesizing ON–peptide conjugates that are common to ONs and their analogs but also points out the need to develop straightforward methods to synthesize such conjugates. The highly functionalized nature of these biomolecules and their mimics such as PNA, render them susceptible for side reactions during conjugation and yield and purity of structurally defined conjugated biomolecules is often low. The (4 + 2) Diels-Alder cycloaddition approach was employed recently for the conjugation of DNA and CPP (28). This involved the reaction between diene and dienophile present on the respective biomolecules to get the conjugates. The maleimide dienophile used in this reaction is susceptible for Michael addition reactions with other nucleophilic centers on peptides or ONs and may give rise to side reactions. The current literature clearly indicates the need for a simple and straightforward strategy for generating highly pure ON/PNA–peptide conjugates in high yield (13).

Meanwhile, we find that the applications of highly selective orthogonal Cu (I) catalyzed Hüisgen 1, 3 dipolar cycloadddition reaction, recognized as 'click chemistry' (29–32), are expanding the scope of synthesis of variety of other bioconjugates such as DNA-glycoconjugates (33), peptide–protein conjugates (34), carbohydrate–vaccine conjugates (35) or protein modification and protein

*To whom correspondence should be addressed. Tel: +91 20 25902340; Fax: +91 20 25902624; Email: va.kumar@ncl.res.in



Scheme 1. Huisgen [3 + 2] cycloaddition reaction between azide and terminal alkyne groups to give inert triazoles.

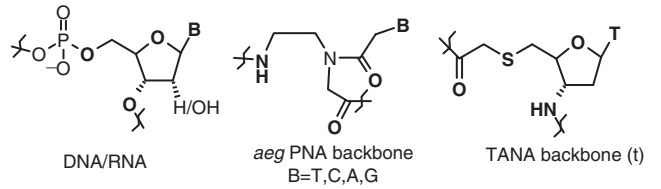


Figure 1. DNA/RNA and PNA and TANA.

microarray fabrication (36). This reaction between high-energy organic azides and terminal alkynes can give rise to unlimited array of inert triazole containing architectures (Scheme 1). The reaction is highly predictable, fast and resistant to side reactions. Addition of Cu(I) accelerates the reaction (31,32). Some of these reactions could be carried out in aqueous medium and can be employed post-synthetically on purified units decorated with a variety of functional groups, without additional functional group protecting strategies. Recently, click reaction has also found applications for the synthesis of circular DNA (37,38), and DNA-template-directed ON strand ligation (38). With this background we envisaged a very simple possibility of synthesizing peptide–DNA/PNA conjugates using click chemistry i.e. a specific cycloaddition reaction between terminal azide and alkyne functionalities on PNA/DNA or peptide as per the synthesis design. These two functional groups are absent in the biomolecules of current interest and no predictable side reaction can be envisaged. This approach has not been previously applied to the synthesis of such conjugates and has potential to lead to a variety of CPP–DNA/PNA conjugates from various combinations of azide and alkyne derivatized peptides-ON substrates in aqueous solution without going into the rigors of continuous synthesis. In this paper, we report successful application of this proposal to generate the CPP conjugates with DNA/PNA ONs and thioacetamido nucleic acid (TANA, Figure 1) (39,40).

MATERIALS AND METHODS

N^{α}-Fmoc L-amino acids and resins for peptide and PNA synthesis were obtained from Novabiochem(Fmoc, 9-fluorenyl methoxy-carbonyl). The Fmoc amino acid used were Fmoc -Lys(Boc)-OH, Fmoc-Arg(Mtr)-OH and N^{ε}-Fmoc-Aha-OH.(Mtr, 4-methoxy-2,3, 6-trimethylbenzenesulfonyl; Boc, *tert*-butoxycarbonyl; Aha, 6-aminohexanoic acid (or ε-aminocaproic acid). 6-aminohexanoic acid was obtained from Aldrich chemicals and *N*-Fmoc protection was done using standard procedure. The concentrations of DNA/PNAs and their conjugates were determined spectrophotometrically and the concentration of the peptides was assumed approximately. Propynoic acid was procured from Trade (TCI) Mark, Tokyo Kasei. *N*-Boc-(2*S*, 4*S*)-4-azidoproline was synthesized according to the reported procedure (41). The Boc and Fmoc protected PNA monomers were obtained from Applied Biosystems, USA. TANA monomers were synthesized in the laboratory following the reported procedures (39). Propyne substituted 19-mer DNA sequence was synthesized using the phosphoramidite approach and an Applied Biosystems 3900 DNA synthesizer.

Reverse phase high-performance liquid chromatography analyses were carried out on VARIAN Analytical Semi-prep HPLC system consisting of Varian Pro-star 210 Binary solvent delivery system. Linear gradients of A: 0.1% TFA in water and B: 55/45: Acetonitrile/Water, 0.1% TFA (Linear gradient from A to B in 30 min Flow-1.5 ml/min.). Rainin Dynamax UV D-II Absorbance Detector Star Ver.5 at detection wavelength 254 nm or 220 nm was employed during the experimentation. Chromatography Workstation Rheodyne 7725I with manual injector and Lichrocart Lichrispher 100RP-18 250 × 4 mm id. Particle size-5 μm column were used.

Mass spectral analysis was performed on a Voyager-De-STR (Applied Biosystems) MALDI-TOF. A nitrogen laser (337 nm) was used for desorption. The matrixes used for analysis were CHCA (α-Cyano-4-hydroxycinnamic Acid), THAP (2′, 4′,and 6′-trihydroxyacetophenone) and HPA (3-hydroxypicolinic acid). Diammonium citrate was used as additive when THAP and HPA were used as matrix.

Solid-phase synthesis of azide functionalized PNA oligomers 2, 3 and 4

The azide functionalized PNA oligomers **2** and **3** were synthesized on Rink-amide resin (100 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis (42) (20 min treatment with 20% piperidine in *N,N*-dimethyl formamide and reaction with 3 equivalents of Fmoc-PNA monomer, HBTU, HOBt and DIPEA for 6 h, were used for the deprotection and coupling steps, respectively). The last coupling reaction was done with *N*-Boc-(2*S*,4*S*)-4-azidoproline. The progress of the coupling reaction was tested by Kaiser test at each step. Cleavage and deprotection were effected by reaction with TFA/DCM/TIS (10:85:5) for 30 min. The resulting oligomer was precipitated by addition of cold ether and purification was done by gel filtration followed by RP–HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 3, 6).

The azide functionalized mixed PNA oligomer **4** was synthesized using side chain *N*-Cbz protected Boc-L-Lysine functionalized MBHA resin (100 mg, loading 0.2 mmol/g) following the standard procedures of solid-phase peptide synthesis (30 min treatment with 50% TFA in DCM and reaction with three equivalents of Boc-PNA

Table 1. The RP-HPLC-t_R and MALDI-TOF mass characterization and purity found by HPLC of the peptide, PNA and TANA sequences and peptide–PNA and peptide–TANA conjugates

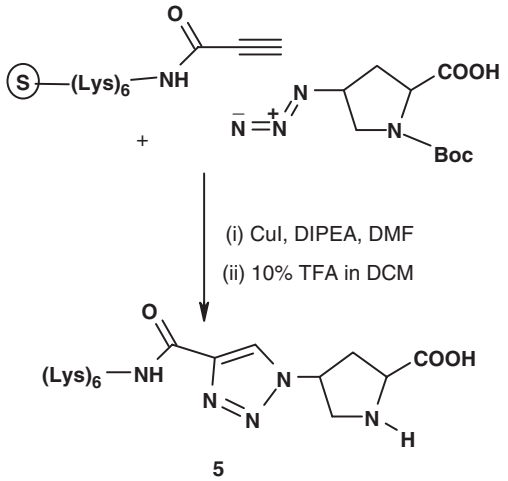
	Compounds	HPLC[a]	Purity (%)	Mass	
				Calcd.	Found
1	HO-$(Lys)_6$-alkyne **1**	9.7	91	839.08	837.8
2	HO-$(Lys)_6$-triazole-Pro **5**	9.59	90	995.23	996.5
3	HO-β-ala-TTTTTTTT-Pro-N_3 **2**	12.6	99.1	2356.2	2356.6
4	HO-β-ala-TTTTTTTT-Pro-triazole$(Lys)_6$-OH **6**	12.29[b]	100	3195.3	3195.35[b]
5	Mixture of **2**+**6** when reaction	12.4	41		
	was still incomplete	12.1	50		
6	HO-β-ala-tttttttt-Pro-N_3 **3**	18.5	100	2605.3	2628(+Na^+)
7	HO-β-ala-tttttttt-Pro-triazole-$(Lys)_6$ **7**	17.2	100	3444.9	3444.0
8	Mixture of **3**+**7** when reaction	18.3	48.7		
	was still incomplete	17.05	43.5		
9	HO-Lys-TCACTAGATG- Pro-N_3 **4**	11.7	98.5	2991.2	2990.8
10	HO-Lys-TCACTAGATG- Pro- triazole-$(Lys)_6$-OH **8**	11.08[b]	98	3830.3	3832.5[b]
11	Mixture of **4**+**8** when reaction was still incomplete.	11.6, 11.1			
12	HO-$(Arg\text{-}Aha\text{-}Arg)_4$-alkyne **9**	12.9	100	1772.1	1772.38
13	HO-β-ala-tttttttt-pro-triazole-$(Arg\text{-}Aha\text{-}Arg)_4$ -OH **10**	17.2	89	4378.6	4376.29
14	HO-Lys-TCACTAGATG- Pro- triazole-$(Arg\text{-}Aha\text{-}Arg)_4$-OH **11**	12.7	95.4	4763.3	4763.8
15	5′-alkyne-d(TTGTACTGATAGAGTGTCC)3′ **12**	8.35	95	5878.04	5878.01
16	HO- $(Arg\text{-}Aha\text{-}Arg)_4$- Pro-$N_3$ **13**	13.8	95	1858.3	1881.2(+Na^+)
					1897.2(+K^+)
17	HO- $(Arg\text{-}Ahx\text{-}Arg)_4$- pro- triazole- 5′-d(TTGTACTGATAGAGTGTCC) 3′ **14**	11.04	81.6[c]	7736.3	7735.09

[a]t_R-RP-HPLC retention time in minutes. T, C, G, A denote *aeg*PNA units or DNA units and t denotes TANA unit, Pro-N_3 denotes (2*S*,4*S*)-4-azidoproline unit, alkyne denotes acetylene carboxylic acid, Aha denotes ε-aminohexanoic acid.
[b]Although the HPLC retention time difference between the starting material and the product was small, the reaction was found to be essentially complete as no corresponding peak for starting oligomer was observed in the mass spectrum of the product (Supplementary Data).
[c]HPLC after purification shows a single peak without any starting materials.

monomer, HBTU, HOBt and DIPEA for 6 h, were used for the deprotection and coupling steps, respectively). The last coupling reaction was done with *N*-Boc-(2*S*,4*S*)-4-azidoproline. Cleavage and deprotection were carried out by the reaction with TFA-TFMSA, thioanisole, ethanedithiol for 2 h. The resulting oligomer was precipitated by addition of cold ether and purification was done by gel filtration followed by RP–HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 9).

Solid-phase synthesis of alkyne functionalized lysine peptide HO-$(Lys)_6$-alkyne 1

The alkyne functionalized lysine peptide was synthesized on Rink-amide resin (100 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis (42) (20 min treatment with 20% piperidine in *N,N*-dimethyl formamide and the reaction with 3 eqivalent of Fmoc-Lys(Boc)-OH, HBTU, HOBt and DIPEA for 4 h, were used for the deprotection and coupling steps, respectively). The last coupling was done with propynoic acid using HBTU, HOBt and DIPEA as the coupling agent. The peptide was cleaved from the resin by treatment with TFA/DCM/TIS (10:85:5) for 30 min and the deprotection of side chain was effected by 50% TFA-DCM (TIS used as scavenger) for 1 h. The resulting peptide was precipitated by addition of cold ether.



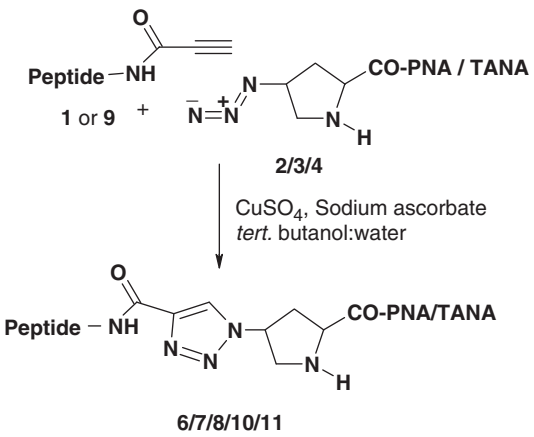
5

Scheme 2. Peptide–alkyne conjugation with 4-azidoproline using click chemistry on solid support.

The product was purified by HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 1).

Click reaction on solid phase: synthesis of HO-$(Lys)_6$-triazole-proline 5

Twenty milligrams of resin bound to protected $(Lys)_6$-alkyne **1**, *N*-Boc-(2S,4S)-4-azidoproline (5 mg) (37), CuI(1 mg), DIPEA(5 μl) and DMF(0.3 ml) were reacted together in a reaction vessel for 8 h (Scheme 2). The excess



Scheme 3. Peptide–alkyne conjugation with 4-azidoprolyl ON using click chemistry in solution.

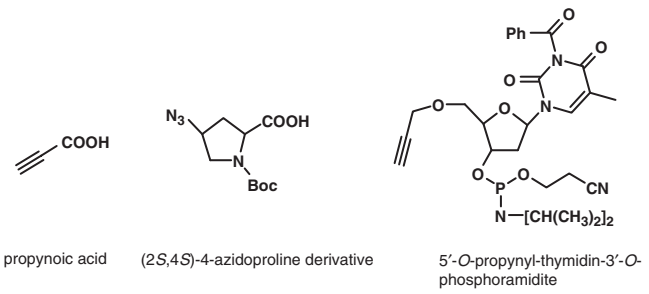


Figure 2. Alkyne and azide carrier units.

reagents were washed out with DMF followed by DCM. Five milligrams of resin was cleaved by using the standard conditions and the product formed was purified by HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 2).

Solution phase reaction between PNA-proline azides 2, 3 and 4 and HO-$(Lys)_6$-alkyne 1: synthesis of conjugates 6, 7 and 8

General procedure. HO-β-ala-TTTTTTTT-Pro-N_3 **2** (or HO-Lys-TCACTAGATG-Pro-N_3 **4** or HO-β-ala-t-t-t-t-t-t-t-t-Pro-N_3 **3**) (1 μmol) and HO-$(Lys)_6$-alkyne **1** (~3 μmol) were dissolved in 50 μl water:*tert*. butanol(1:1). $CuSO_4.5H_2O$ (1.0 equivalent, 1 μmol, 10 μl of a 100 mM solution in water) and freshly prepared solution of sodium ascorbate (4 equivalents, 4 μmol, 8 μl of 500 mM solution in water) was then added. The mixture was stirred in a spinix vortex at room temperature (Scheme 3). The reaction mixture was analyzed by HPLC after 2 h. HPLC showed complete consumption of the starting material in the case of **3** giving product **7** (Table 1, entry 7). In the case of either **2** or **4** the differences in HPLC-t_R was not very clear from the starting materials. In these cases the product (**6** or **8**) after HPLC purification were characterized by MALDI-TOF mass spectrometry (Table 1, entry 4 and 10). Complete conversion of the starting material **2** and **4** was established by the disappearance of the corresponding mass peak in MALDI-TOF analysis (Supplementary Data).

Solid-phase synthesis of the peptide HO-$(Arg\text{-}Aha\text{-}Arg)_4$-alkyne 9

The arginine peptide **9** was synthesized on Rink-amide resin (200 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis: 20 min treatment with 20% piperidine in *N,N*-dimethyl formamide and reaction with 3 equivalent of Fmoc-Arg(Mtr)-OH (or N^{ε}-Fmoc-Aha-OH), HBTU, HOBt and DIPEA for 4 h, were used for the deprotection and coupling steps, respectively. The final coupling was done with propynoic acid using HBTU, HOBt and DIPEA as the coupling agent. Cleavage of the peptide from resin was effected by reaction with TFA/DCM/TIS (10:85:5) for 30 min. The deprotection of side-chain protecting group (Mtr) was effected by 100% TFA and thioanisole as the scavenger for 5 h. The resulting peptide was precipitated by addition of cold ether and oligomer was purified by RP–HPLC. The product was further characterized by MALDI-TOF mass spectrometry (Table 1, entry 12).

Solution phase reaction between PNA- proline azides 3, 4 and HO-$(Arg\text{-}Aha\text{-}Arg)_4$-alkyne 9: synthesis of conjugates 10 and 11

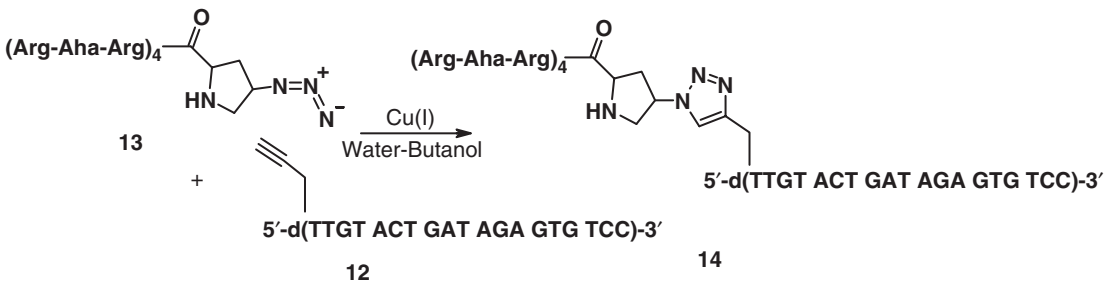
General procedure. HO-Lys-TCACTAGATG-Pro-N_3 **4** or HO-β-ala-t-t-t-t-t-t-t-t-Pro-N_3 **3** (1 μmol) and $(Arg\text{-}Aha\text{-}Arg)_4$-alkyne **9** (~3 μmol) were dissolved in 50 μl water:*tert*. Butanol (1:1). $CuSO_4.5H_2O$ (1.0 equivalent, 1 μmol, 10 μl of a 100 mM solution in water) and freshly prepared solution of sodium ascorbate (4 equivalents, 4 μmol, 8 μl of 500 mM solution in water) was then added (Scheme 3). The mixture was stirred in a spinix vortex at room temperature. The reaction mixture was analyzed by RP–HPLC after 2 h. HPLC showed complete conversion of stating materials to the products. The products (**10** and **11**) formed were purified by RP–HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 13 and 14).

Synthesis of alkyne substituted DNA oligomer 12

The DNA sequence **12** was synthesized using commercially available monomeric units using phoaphoramidite chemistry on automated DNA synthesizer. The last coupling was done using 5′-*O*-(propynyl-*N*-*3*-benzoyl-thymidin)-3′-*O*-(*N,N*-diisopropylamino-*O*-cyanoethyl -phosphoramidite) (Figure 2). (Supplementary Data). The ON **12** was deprotected and cleaved from the solid support. It was further purified by RP–HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 15).

Solid-phase synthesis of the peptide HO-$(Arg\text{-}Aha\text{-}Arg)_4$-azide 13

The arginine peptide **13** was synthesized on Rink-amide resin (200 mg, loading 0.3 mmol/g) following the standard procedures of solid-phase peptide synthesis: 20 min treatment with 20% piperidine in *N,N*-dimethyl formamide and reaction with 3 eqivalent of Fmoc-Arg(Mtr)-OH (or N^{ε}-Fmoc-Aha-OH), HBTU, HOBt and DIPEA



Scheme 4. DNA–alkyne conjugation with 4-azidoprolyl-peptide using click chemistry in solution.

for 4 h, were used for the deprotection and coupling steps, respectively. The final coupling was done with *N*-Boc-4-azido-proline using HBTU, HOBt and DIPEA as the coupling agent. Cleavage of the peptide from resin was effected by reaction with TFA/DCM/TIS (10:85:5) for 30 min. The deprotection of side chain protecting group (Mtr) was effected by 100% TFA and thioanisole as the scavenger for 5 h. The resulting peptide was precipitated by addition of cold ether and was purified by RP–HPLC. The product was further characterized by MALDI-TOF mass spectrometry (Table 1, entry 16).

Solution phase reaction between DNA-alkyne 12 and HO-(Arg-Aha-Arg)$_4$-azide 13: synthesis of conjugate 14

Alkyne functionalized ON **12** (1 μmol) and azide functionalized peptide **13** (~3 μmol) were dissolved in 50 μl water:*tert*. Butanol (1:1). $CuSO_4.5H_2O$ (0.6 equivalent, 0.6 μmol, 6 μl of a 100 mM solution in water) and freshly prepared solution of sodium ascorbate (2.4 equivalents, 4 μmol, 6 μl of 500 mM solution in water) was then added (Scheme 4). The mixture was stirred in a spinix vortex at room temperature. After 1 h, a small portion of the reaction mixture was analyzed by RP–HPLC. HPLC showed complete conversion of stating materials to the product. The product **14** formed was purified by RP–HPLC and characterized by MALDI-TOF mass spectrometry (Table 1, entry 17).

RESULTS AND DISCUSSION

The aim of this study was to develop a simple and chemo-specific method for the conjugation of peptides and ON mimics such as PNA and TANA. The azide and acetylene carrier units were chosen so that these could be installed online during the solid supported peptide or PNA/TANA synthesis at the *N*/*C*-terminus through an amide linkage that is stable towards conditions employed for the deprotection and cleavage of the peptide/ON from the solid support. The candidates chosen for this purpose were propynoic acid and (2*S*,4*S*)- 4-azidoproline (Figure 2) (41). DNA ON was functionalized at 5′-end using 5′-*O*-propynyl-thymidine-phosphoramidite (Figure 2), (Supplementary Data). The 4-azidoproline was chosen in this strategy, as it could also open up the possibility to introduce the azide group at either *C*/*N*-terminus or in the center of the PNA or peptide sequences as desired. This could enable the conjugation of peptide at either C- or N-terminus of PNA or peptide. To assess the methodology and its application potential, we chose three PNA sequences, a homopyrimidine PNA-T8, a mixed purine-pyrimidine 10-mer PNA sequence and a recently developed homothyminyl-oligomer with uncharged thioacetamido nucleic acid (TANA) backbone (39). The reaction medium was *tert*.-butanol:water that solubilizes all the reactants and is also capable of scavenging hydroxy radicals (44). The peptide sequence chosen was the lysine (18) or the arginine rich peptide (19) that were known to act as vehicle for PNAs to penetrate through the cell membrane. The azide functionalized PNA oligomers **2** and **3** were synthesized on Rink-amide resin following the standard procedures (42) of Fmoc based solid-phase peptide synthesis. The last coupling reaction was done using *N*-Boc-(2*S*,4*S*)-4-azidoproline monomer unit. The azide functionalized mixed purine-pyrimidine PNA oligomer **4** was synthesized following the standard procedures of Boc-based solid-phase peptide synthesis (42). The final coupling reaction was done with *N*-Boc-(2*S*,4*S*)-4-azidoproline. The alkyne functionalized lysine peptide **1** and arginine-rich peptide **9**, were synthesized on Rink-amide resin following the standard procedures of solid-phase peptide synthesis. The final coupling was done with propynoic acid. To install the azide functional group on DNA oligomer, we attempted to synthesize 5′deoxy-5′-azido-thymidinyl-3′-*O*-phosphoramidite, but this was found to be quite unstable. We then synthesized 5′-*O*-propynyl-*N*3-benzoylthymidinyl-3′-*O*-phosphoramidite (Supplementary Data). The 5′-end alkyne substituted DNA sequence **12** was synthesized using automated DNA synthesizer when the last coupling was with 5′-*O*-propynyl-*N*3-benzoylthymidinyl-3′-*O*-phosphoramidite. N-terminal azide containing arginine-rich peptide **13** was synthesized as **9**, and *N*-Boc-(2*S*,4*S*)-4-azidoproline was used for the last coupling reaction. The installation of reactive alkyne and azide functionalities on either the peptide or ON, was followed by the experiments to study the feasibility of the click cycloaddition reaction. The peptide **1**, attached to the solid support with all the side chain lysine amino- and terminal carboxy- groups protected, was treated with 3 equivalents of *N*-Boc(2*S*,4*S*)-4-azidoproline in DMF in the presence of CuI. The product **5** was isolated after cleavage from the support. There was negligible difference in RP–HPLC t_R but mass showed conjugation of proline unit through the triazole linkage (Scheme 2, Table 1, entry 5).

After the formation of the desired product was established, the solution phase reactions of PNAs with N-terminal azidoproline (**2** and **4**) or TANA (**3**) and N-terminal-alkyne substituted lysine peptide **1** and arginine-rich peptide **9**, were carried out in *tert*. butanol: water medium in presence of $CuSO_4$ and sodium ascorbate as a catalyst (Scheme 3). The CuI catalyst was found to be less efficient as compared to $CuSO_4$-sodium ascorbate under these conditions. The reactions were followed by HPLC. The reaction took longer time for PNA, TANA substrates (16–32 h) to completion with-approximately one equivalent of each substrate. With approximately three equivalents excess of **1** or **9** complete conversion to the conjugated products (**6–8** and **10, 11**) was observed within 2 h. The DNA substrate **12** was converted to the peptide–DNA conjugate **14** within 1 h with 3 equivalents of the peptide (Scheme 4). Cu(I) catalysis is known to degrade DNA due to the presence of hydroxy radicals and addition of Cu(I) stabilizing ligands such as polytriazoles is often recommended to circumvent the DNA degradation (38). Arginine side chain imine nitrogen is also known to efficiently stabilize Cu(I) (43). It is therefore possible that the arginine peptide used in this reaction could have contributed to the stabilization of Cu(I) similar to the Cu(I)-binding ligands as small percentage of DNA degradation was observed. Additionally, presence of *tert*.-butanol as solvent in this reaction would circumvent DNA degradation as *tert*.-butanol is known scavenger of hydroxy radicals (44). The yields of peptide–DNA conjugates may be further improved by the use of Cu(I)-stabilizing ligands such as polytriazoles (38) along with *tert*.-butanol as hydroxy radical scavenger (44). The results are tabulated in Table 1 (entry 4, 7, 10, 13 and 14). The peptide–ON conjugates with either the DNA (**14**), *aeg*PNA backbone (**6, 8** and **11**) or TANA backbone (**7, 10**) were obtained in high purity using this simple cycloaddition reaction. The reaction seems to be very specific without interfering with other functional groups present, either on peptide or DNA/PNA/TANA segments of the conjugates.

CONCLUSIONS

The peptide–DNA/PNA/TANA conjugation chemistry presented in this paper will prove to be a unique solution for CPP conjugation to antisense/siRNA therapeutics. In this fast-developing field, application is mainly restrained by the problems faced in cellular delivery of modified ONs. The ease of conjugation and compatibility of the strategy with various substrates and functional groups will prove to be an attractive alternative to present methods for the synthesis of peptide–PNA conjugates. It is sequence-independent with respect to either peptide or PNA/DNA or other backbone modified DNA mimics (such as TANA), which are being developed for antisense applications. The reactive azide and alkyne groups can be installed on either ON or peptide and the work presented, has a potential for general application in nucleic acid–peptide conjugation. This could be a method of choice when a peptide of choice needs to be conjugated with several ON sequences or different modifications of the same sequence for studying their bioactivity. Alternatively, the chosen ON may be conjugated with several peptides to establish the biological efficacy of different peptides. Some recent work was carried out to evaluate the ON–peptide conjugates as a potential alternative to the use of transfection agents to improve the efficacy of siRNA (45). It will be very interesting to study the effects of the chemistry of this new linkage on the intracellular performance of the conjugates (9,19).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

K.G. thanks CSIR, New Delhi and V.A.K. thanks DST, New Delhi and CEFIPRA, New Delhi for supporting this work. The initial work was presented as a poster in the Cell Penetrating Peptides meeting held at Telford, UK during 9–11 May, 2007. Funding to pay the Open Access publication charges for this article was waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Nielsen,P.E., Egholm,M., Berg,R.H. and Buchardt,O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **254**, 1497–1500.
2. Braasch,D.A. and Corey,D.R. (2002) Novel antisense and peptide nucleic acid strategies for controlling gene expression. *Biochemistry*, **41**, 4503–4510.
3. Uhlmann,E., Peymann,A., Breipohl,G. and Will,D.W. (1998) PNA: synthetic polyamide nucleic acids with unusual binding properties. *Angew. Chem. Int. Ed. Engl.*, **37**, 2796–2823.
4. Kurreck,J. (2003) Antisense technologies: improvement through novel chemical modifications. *Eur. J. Biochem.*, **270**, 1628–1644.
5. Micklefield,J. (2001) Backbone modification of nucleic acids: synthesis, structure and therapeutic applications. *Curr. Med. Chem.*, **8**, 1157–1179.
6. Kumar,V.A. (2002) Structural preorganization of peptide nucleic acids: chiral cationic analogues with five- or six-membered ring structures. *Eur. J. Org. Chem.*2021–2032.
7. Kumar,V.A. and Ganesh,K.N. (2005) Conformationally constrained PNA analogs: structural evolution towards DNA/RNA binding selectivity. *Acc. Chem. Res.*, **38**, 404–412.
8. In Crooke,S.T. (ed), *Antisense Research and Application*, Springer, Berlin, pp. 1–50.
9. Opalinska,J.B. and Gewirtz,A.M. (2002) Nucleic-acid therapeutics: basic principles and recent applications. *Nat. Rev. Drug Discovery*, **1**, 503–515.
10. Dominski,Z. and Kole,R. (1993) Restoration of correct splicing in Thalassemic pre-mRNA by antisense oligonucleotides. *Proc. Natl Acad. Sci. USA*, **90**, 8673–8677.
11. Turner,J.J., Fabani,M., Arzumanov,A.A., Ivanova,G. and Gait,M.J. (2006) Targeting the HIV-1 RNA leader sequence with synthetic oligonucleotides and si-RNA: chemistry and cell delivery. *Biochim. Biophys. Acta*, **1758**, 290–300.
12. Manoharan,M. (2002) Oligonucleotide conjugates as potential antisense drugs with improved uptake, biodistribution, targeted delivery, and mechanism of action. *Antisense Nucleic Acid Drug Dev.*, **12**, 103–128.
13. Venkatesan,N. and Hyean Kim,B. (2006) Peptide conjugates of oligonucleotides: Synthesis and applications. *Chem. Rev.*, **106**, 3712–3761.

14. Fraley,A.W., Pons,B., Dalkara,D., Nullans,G., Behr,J.-P. and Zuber,G. (2006) Cationic oligonucleotide-peptide conjugates with aggregating properties enter efficiently into cells while maintaining hybridization properties and enzymatic recognition. *J. Am. Chem. Soc.*, **128**, 10763–10771.
15. Albertshofer,K., Siwkowski,A.M., Wancewicz,E. V., Esau,C.C., Watanabe,T., Nishihara,K.C., Kinberger,G.A., Malik,L., Eldrup,A. B. *et al*. (2005) Structure-activity relationship study on a simple cationic peptide motif for cellular delivery of antisense peptide nucleic acid. *J. Med. Chem.*, **48**, 6741–6749.
16. Bendifallah,N., Rasmussen,F.W., Zachar,V., Ebbesen,P., Nielsen,P.E. and Koppelhus,U. (2006) Evaluation of cell-penetrating peptides (CPPs) as vehicles for intracellular delivery of antisense peptide nucleic acid (PNA). *Bioconjugate Chem.*, **17**, 750–758.
17. Maier,M.A., Esau,C.C., Siwkowski,A.M., Wancewicz,E.V., Albertshofer,K., Kinberger,G.A., Kadaba,N.S., Watanabe,T., Manoharan,M. *et al*. (2006) Evaluation of basic amphipathic peptides for cellular delivery of antisense peptide nucleic acids. *J. Med. Chem.*, **49**, 2534–2542.
18. Sazani,P., Kang,S.H., Maier,M.A., Wei,C., Dillman,J., Summerton,J., Manoharan,M. and Kole,R. (2001) Nuclear antisense effects of neutral, anionic and cationic oligonucleotide analogs. *Nucleic Acids Res.*, **29**, 3965–3974.
19. Rothbard,J.B., Kreider,E., VanDeusen,C.L., Wright,L., Wylie,B.L. and Wender,P.A. (2002) Arginine-rich molecular transporters for drug delivery: role of backbone spacing in cellular uptake. *J. Med. Chem.*, **45**, 3612–3618.
20. Rothbard,J.B., Garlington,S., Lin,Q., Kirschberg,T., Kreider,E., McGrane,P.L., Wender,P.A. and Khavari,P.A. (2000) Conjugation of arginine oligomers to cyclosporin- A facilitates topical delivery and inhibition of inflammation. *Nat Med.*, **6**, 1253–1257.
21. Turner,J.J., Ivanova,G.D., Verbeure,B., Williams,D.A., Arzumanov,A., Abes,S., Lebleu,B. and Gait,M.J. (2005) Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitors of HIV-1 Tat-dependent trans-activation in cells. *Nucleic Acids Res.*, **33**, 6837–6849.
22. Diaz-Mochon,J.J., Bialy,L. and Bradley,M. (2004) Full orthogonality between Dde and Fmoc: the direct synthesis of PNA-peptide conjugates. *Org. Lett.*, **6**, 1127–1129.
23. Diaz-Mochon,J.J., Bialy,L., Waston,J., Anchez-Martin,R.M. and Bradley,M. (2005) Synthesis and cellular uptake of cell delivering PNA-peptide conjugates. *Chem. Commun.*, 3315–3318.
24. Kaihatsu,K., Huffmann,K.E. and Corey,D.R. (2004) Intracellular uptake and inhibition of gene expression by PNAs and PNA-peptide conjugates. *Biochemistry*, **43**, 14340–14347.
25. Mayfield,L.D. and Corey,D.R. (1999) Automated synthesis of peptide nucleic acids and peptide nucleic acids-PNA conjugates. *Anal.l Biochem.*, **268**, 401–404.
26. Mier,W., Eritja,R., Mohammed,A., Haberkorn,U. and Eisenhut,M. (2003) Peptide-PNA conjugates: targeted transport of antisense therapeutics into tumors. *Angew. Chem. Int. Ed. Engl.*, **42**, 1968–1971.
27. Turner,J.J., Arzumanov,A.A. and Gait,M.J. (2005) Synthesis, cellular uptake and HIV-1 Tat-dependent *trans*-activation inhibition activity of oligonucleotide analogues disulphide-conjugated to cell-penetrating peptides. *Nucleic Acids Res.*, **33**, 27–42.
28. Marchan,V., Ortega,S., Pulido,D., Pedroso,E. and Grandas,A. (2006) Diels-Alder cycloadditions in water for the straightforward preparation of peptide oligonucleotide conjugates. *Nucleic Acids Res.*, **34**, e24.
29. Huisgen,R. (1984) In Padwa,A (ed), *1, 3-dipolar Cycloaddition Chemistry*, Wiley, New York.
30. Tornøe,C.W., Christensen,C. and Meldal,M. (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.*, **67**, 3057–3064.
31. Narayan,S., Muldoon,J., Finn,M.G., Fokin,V.V., Kolb,H.C. and Sharpless,K.B. (2005) "On Water": Unique reactivity of organic compounds in aqueous suspension. *Angew. Chem. Int. Ed.*, **44**, 3275–3279.
32. Kolb,H.C., Finn,M.G. and Sharpless,K.B. (2001) Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed.*, **40**, 2004–2021.
33. Bouillon,C., Meyer,A., Vidal,S., Jochum,A., Chevolot,Y., Cloarec,J.-P., Praly,J.-P., Vasseur,J.-J. and Morvan,F. (2006) Microwave assisted "Click" chemistry for the synthesis of multiple labeled-carbohydrate oligonucleotides on solid support. *J. Org. Chem.*, **71**, 4700–4702.
34. Meier,J.L., Mercer,A.C., Rivera,H.Jr. and Burkart,M.D. (2006) Synthesis and evaluation of bioorthogonal pantetheine analogues for in vivo protein modification. *J. Am. Chem. Soc.*, **128**, 12174–12184.
35. Wan,Q., Chen,J., Chen,G. and Danishefsky,S.J. (2006) A potentially valuable advance in the synthesis of carbohydrate-based anticancer vaccines through extended cycloaddition chemistry. *J. Org. Chem.*, **71**, 8244–8249.
36. Lin,P.-C., Ueng,S.-H., Tseng,M.-C., Ko,J.-L., Huang,K.-T., Yu,S.-C., Adak,A.K., Chen,Y.-J. and Lin,C.-C. (2006) Site-specific protein modification through Cu^{I}-catalyzed 1,2,3-triazole formation and its implementation in protein microarray fabrication. *Angew. Chem. Int. Ed.*, **45**, 4286–4290.
37. Nakane,M., Ichikawa,S. and Matsuda,A. (2005) Synthesis of cross-linked circular DNAs using Huisgen reaction. *Nucleic Acids Symp. Ser.*, **49**, 189–190.
38. Kumar,R., El-Sagheer,A., Tumpane,J., Lincoln,P., Wilhelmsson,L.M. and Brown,T. (2007) Template-directed oligonucleotide strand ligation, covalent intramolecular DNA circularization and catenation using click chemistry. *J. Am. Chem. Soc.*, **129**, 6859–6864.
39. Gogoi,K., Gunjal,A. D. and Kumar,V.A. (2006) Sugar–thioacetamide backbone in oligodeoxyribonucleosides for specific recognition of nucleic acids. *Chem. Commun.*2373–2375.
40. Gogoi,K., Gunjal,A. D., Phalgune,U.D. and Kumar,V.A. (2007) Synthesis and RNA binding selectivity of oligonucleotides modified with five-atom TANA backbone structures. *Org. Lett.*, **9**, 2697–2700.
41. Gangamani,B.P., Kumar,V.A. and Ganesh,K.N. (1996) Synthesis of N^{α}-(puinyl/pyrimidinyl acetyl)-4-aminoproline diastereomers with potential use in PNA synthesis. *Tetrahedron*, **52**, 15017–15030.
42. Nielsen,P.E. and Egholm,M. (1999) Peptide nucleic acids (PNA). In Nielsen,P.E. and Egholm,M. (eds), *Protocols and Applications.*, Horizon scientific Norfolk, CT.
43. Bluhm,B.K., Shields,S.J., Bayse,C.A., Hall,M.B. and Russell,D.H. (2001) Determination of copper binding sites in peptides containing basic residues: a combined experimental and theoretical study. *Int. J. Mass Spectrom.*, **204**, 31–46.
44. Weller,R.L. and Rajski,S.R. (2005) DNA methyltransferase-moderated click chemistry. *Org. Lett.*, **7**, 2141–2144.
45. Turner,J.J., Jones,S., Fabani,M., Ivanova,G., Arzumanov,A.A. and Gait,M. (2007) RNA targeting with peptide conjugates of oligonucleotides, siRNA and PNA. *Blood Cell. Mol. Dis.*, **38**, 1–7.

Chemical Communications

b716835g

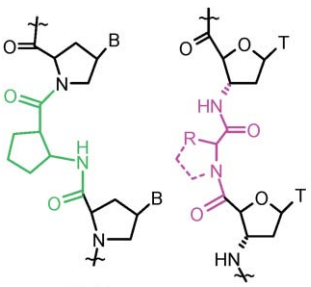
COMMUNICATION

1

Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers show sequence specific DNA/RNA recognition

Khirud Gogoi and Vaijayanti A. Kumar

An α/β-peptide backbone oligonucleotide comprising natural α-amino acids alternating with a β-amino acid component derived from thymidine sequence specifically recognizes and binds to deoxy- and ribo-oligoadenylates in triplex mode.



Please check this proof carefully. Our staff will not read it in detail after you have returned it.
Translation errors between word-processor files and typesetting systems can occur so the whole proof needs to be read. Please pay particular attention to: tabulated material; equations; numerical data; figures and graphics; and references. If you have not already indicated the corresponding author(s) please mark their name(s) with an asterisk. Please e-mail a list of corrections or the PDF with electronic notes attached — do not change the text within the PDF file or send a revised manuscript.

Please bear in mind that minor layout improvements, e.g. in line breaking, table widths and graphic placement, are routinely applied to the final version.

We will publish articles on the web as soon as possible after receiving your corrections; no late corrections will be made.

Please return your **final** corrections, where possible within **48 hours** of receipt, by e-mail to: proofs@rsc.org

Electronic (PDF) reprints will be provided free of charge to the corresponding author. Enquiries about purchasing paper reprints should be addressed via: http://www.rsc.org/Publishing/ReSourCe/PaperReprints/. Costs for reprints are below:

Reprint costs

No of pages	Cost for 50 copies	Cost for each additional 50 copies
2–4	£180	£115
5–8	£300	£230
9–20	£600	£480
21–40	£1100	£870
>40	£1700	£1455

Cost for including cover of journal issue:
£50 per 50 copies

Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers show sequence specific DNA/RNA recognition†

Khirud Gogoi* and Vaijayanti A. Kumar*

Received (in Cambridge, UK) 31st October 2007, Accepted 27th November 2007
First published as an Advance Article on the web
DOI: 10.1039/b716835g

An α/β-peptide backbone oligonucleotide comprising natural α-amino acids alternating with a β-amino acid component derived from thymidine sequence specifically recognizes and binds to deoxy- and ribo-oligoadenylates in triplex mode.

Novel oligonucleotide (ON) analogues that can form stable duplexes or triplexes with target nucleic acids (NA) are important synthetic objectives because of their use as therapeutic agents in antisense and antigene strategies.[1] Various types of modified oligonucleotides have been developed over the last two decades as potential diagnostic probes and antisense and antigene therapeutics.[2] The more recent developments such as splice correcting[3] and exon skipping[3] strategies require highly robust nucleic acid analogues that are stable under physiological conditions as single strands as well as in the form of duplexes with complementary RNA sequences. The study of unnatural de-phosphono nucleic acid oligomers is gaining in importance for such applications. The replacement of the internucleoside sugar–phosphate linkages by the robust sugar–amide bond could be advantageous as in PNA.[4,5] This would maintain the chirality and 3′–5′ directionality of the backbone, along with the enzymatic stability of the amide bond. Several four- and five-atom amide-linked (6/7 atom repeating units) deoxy[6,7] and ribo[8,9] ON analogues are known in the literature. The other relevant examples of 6/7-atom repeating units are those of pyrrolidine-based NA analogues[10] such as pyrrolidinyl[11]/POM-PNA[12] and *bep*PNA (Fig. 1)[13] or those comprising prolyl-aminopyrrolidine-2-carboxylic acid[14] and prolyl-2-aminocyclopentanecarboxylic acid (prolyl-ACPC backbone) (Fig. 1).[15,16]

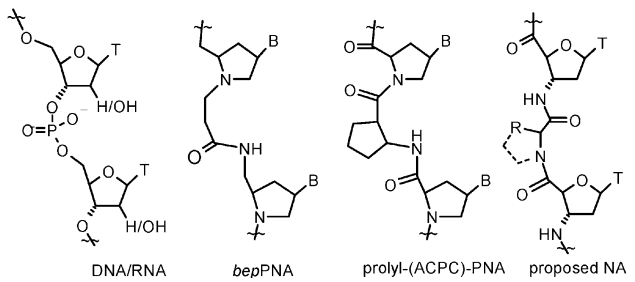


Fig. 1 Structure of DNA/RNA and amino acid backbone modified ONs.

Division of Organic Chemistry, National Chemical Laboratory, Pune, 411008, India. E-mail: k.gogoi@ncl.res.in; va.kumar@ncl.res.in; Fax: +91-20-25902624; Tel: +91-20-25902340

† Electronic supplementary information (ESI) available: Experimental procedures; ^{1}H, ^{13}C, mass spectra of compounds **7** and **8**; MALDI-TOF mass spectra of ONs **12–15**; UV melting experiments and CD curves. See DOI: 10.1039/b716835g

The prolyl-ACPC backbone seems very interesting as this alternating α/β amino acid backbone exhibited preferential binding with complementary DNA but not with complementary RNA sequences. We perceived that alternating nucleoside-β-amino acids with natural α-amino acids (Fig. 2) would give rise to similar alternating α/β amino acid backbone scaffolds. A literature search revealed that similar work on phosphate free oligonucleoside analogues is reported in a patent.[17] The internucleoside amino acids used were only glycine and lysine in that study and only binding studies with complementary DNA oligomers were reported. In the light of recent developments in this area, we thought it worthwhile to reconstitute the sugar–amino acid backbone incorporating different natural amino acids for RNA binding studies.

In this communication, we report the synthesis and DNA/RNA binding studies of amide linked oligothyminyl DNA analogues (Fig. 1) comprising conformationally constrained α-amino acids (L-proline **2** and sarcosine **3**, Fig. 2), a positively charged α-amino acid (L-lysine **4**, Fig. 2) and a neutral α-amino acid (L-methionine **5**, Fig. 2) in the backbone, alternating with a β-amino acid nucleoside component derived from thymidine (**1**, Fig. 2). The potential advantages are as follows. (a) In 3′-deoxy-3′-amino-2′-deoxyribose sugar the five-membered ring pucker is in a preferred 3′-endo conformation that is better suited for m-RNA recognition.[18] (b) The new synthesis of nucleoside-β-amino acid is quite simple using a recently reported TEMPO–BAIB method.[19] (c) The flexibility of the five atom internucleoside amide linker could be better fitting for replacement of a phosphate group for RNA binding taking into account the shorter amide bond compared to the phosphodiester linkage.[7] (d) A negatively charged phosphate group can be replaced by neutral (proline, sarcosine and methionine) or positively charged[20] (lysine) alternatives.

Our synthetic strategy was aimed at assembling these molecules on a solid support, for its subsequent adaptation to well established conventional peptide chemistry.[21] Oxidation of the 5′-hydroxy moiety of 3′-azidothymidine **6** to obtain **7** was reported earlier[17,22a] by Varma *et al.* using $RuCl_3$–$Na_2S_2O_8$–KOH. The use of strongly alkaline conditions in this reaction makes it unsuitable

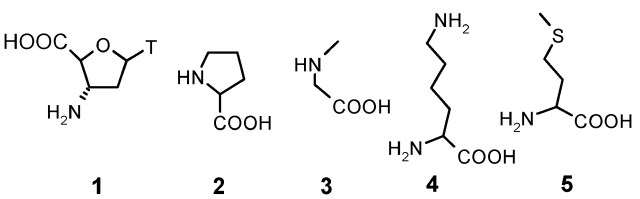
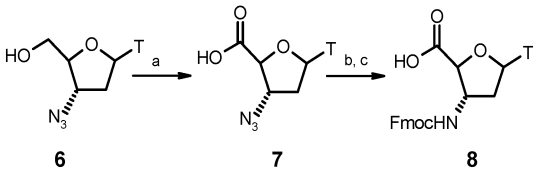


Fig. 2 Structures of monomer units.



Scheme 1 Synthesis of thyminyl sugar amino acid monomer unit. *Reagents and conditions*: (a) TEMPO–BAIB, acetonitrile–water; (b) Pd/C–H_2, methanol; (c) Fmoc-succinimide, $NaHCO_3$, acetone–water. TEMPO 2,2,6,6-tetramethyl-1-piperidinyloxyl; BAIB [bis(acetoxy)-iodo]benzene; Fmoc fluorenylmethoxycarbonyl.

for adaptation to other amino-protected nucleosides. The products were also often contaminated with inorganic salts. Oxidation using Sharpless conditions works well only for purine nucleosides and not with pyrimidine nucleosides.[22b] A recently published TEMPO–BAIB method[19a] used for 2′,3′-isopropylidene ribonucleosides could be a more general procedure for nucleosides[19b] as it is reported to be compatible with several protecting groups and other functionalities such as azide. The products isolated are not contaminated with any inorganic salts. Thus, **6** was easily converted to the corresponding acid **7** under TEMPO–BAIB oxidation conditions and the product was recovered by trituration with ether and acetone (Scheme 1). The 3′-azido group in **7** was then hydrogenated using Pd–C to an amine which was subsequently Fmoc protected to give the monomer building block **8**. The monomer **8** along with protected α-L-amino acids (**2**, **3**, **4** and **5**) was used to synthesize four octameric NA sequences (**12**, **13**, **14** and **15**) on a rink amide resin support using an Fmoc peptide synthesis strategy. For sequences **12**, **13** and **15**, β-alanine was used as a linker to the resin and for sequence **14** no linker was used.[23] All the coupling reactions were monitored by the Kaiser test. The oligomers were cleaved from the solid support using standard conditions and were purified by gel-filtration followed by RP-HPLC. Finally the purity and integrity of the oligomers were established by MALDI-TOF mass spectrometry (ESI†).

The hybridization properties of the amino acid modified oligomers **12**, **13**, **14** and **15** with complementary DNA or RNA ONs GCAAAAAAAACG **9** or r(GCAAAAAAAACG) **10** were determined by thermal denaturation studies.[24] UV melting experiments indicate that oligomers **12**, **13**, **14** and **15** hybridized to complementary DNA and RNA with melting temperatures (T_m) higher than those of the complexes formed by oligothymidinyl-DNA fragment **11** of the same length (Fig. 3, Table 1). Complexes formed by all the modified oligomers **12**, **13**, **14** and **15** with the RNA target showed higher stability than similar complexes formed with the DNA target (Table 1). The positively charged lysine oligomer **14** formed the most stable complexes with both DNA (**14:9**) and RNA (**14:10**). The UV melting experiments of ON **14** with complementary DNA **9** and RNA **10** were performed at pH 5.5 but no significant change in melting temperature could be detected, indicating that the side chain lysine amine group could be protonated even at physiological pH. A single base mismatch caused a fall in T_m (11–18 °C) in the case of either DNA or RNA targets (ESI†) similar to the case of PNA.[4c] A UV–Job's plot of the modified oligomers with complementary DNA and RNA clearly showed a 2 : 1stoichiometry of the complex (ESI†).

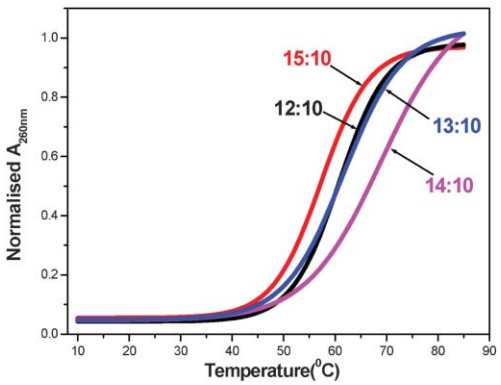


Fig. 3 UV melting graph of ONs **12**, **13**, **14** and **15** with RNA **10** (in 10 mM sodium phosphate buffer, pH = 7.0, 100 mM NaCl, and 0.1 mM EDTA). RNA **10**: r(5′ GCAAAAAAAACG 3′).

The CD spectra of single strands **12–15** showed ellipticity minima at ~280 nm with no positive CD band (ESI†). The CD spectrum of the 2 : 1 hybrid formed between **12** and complementary RNA **10** showed ellipticity minima at 246 nm and ellipticity maxima at 260 and 281 nm (Fig. 4). The CD spectra of the complexes of **13**, **14** and **15** with complementary RNA showed a similar CD signature (Fig. 4), substantiating the formation of triple helical structures.[25]

Thus, the complexes between chimeric (α-amino acid + nucleoside-β-amino acid) backbone ONs (**12**, **13**, **14** and **15**) and their complementary DNA/RNA (**9/10**) oligonucleotides were found to be highly stable as compared to DNA:DNA (**11:9**) or DNA:RNA (**11:10**) duplexes. The presence of a positively charged α-amino acid (L-lysine) contributes further to the strong binding properties (7–8 °C). The stability of complexes of sarcosine and proline analogues, similar to the prolyl-ACPC PNA, in which one amide hydrogen is absent, suggests the absence of a contribution from the known secondary structures of the (α-amino acid + β-amino acid) peptide scaffolds[26] to the overall stability of the complexes. The 5-atom amide linked oligomers in this simple (α-amino acid + nucleoside-β-amino acid) backbone exhibit much higher stability of complexes with both complementary DNA and RNA, complexation with RNA being favoured over DNA. The basis of this selectivity could be similar to that for the other 5-atom internucleoside amide linkages.[7,27]

Table 1 UV-T_m[a] values (in °C) of chimeric modified ONs:complementary DNA/RNA

No.	Sequence	DNA **9**	RNA **10**	ΔT_mRNA–DNA
1	TTTTTTTT **11**	17.8	15.6	—
2	β-Ala-(Pro-**t**)$_8$-H **12**	49.0	60.8	+11.8
3	β-Ala-(Sar-**t**)$_8$-H **13**	49.1	61.6	+12.5
4	(Lys-**t**)$_8$-H **14**	56.3[b]	69.0[b]	+12.7
5	β-Ala-(met-**t**)$_8$-H **15**	47.3	57.3	+10.0

[a] T_m = melting temperature measured in 10 mM sodium phosphate buffer, pH 7.0 with 100 mM NaCl and 0.1 mM EDTA, from 5 to 85 °C at ramp 0.5 °C. All values are an average of three independent experiments and accurate to within ±0.5 °C. T denotes DNA backbone. **t** denotes amide backbone. DNA **9**: 5′ GCAAAAAAAACG 3′, RNA **10**: r(5′ GCAAAAAAAACG 3′). [b] UV melting experiments were performed at pH 5.5 also but no significant change in melting temperatures was detected.

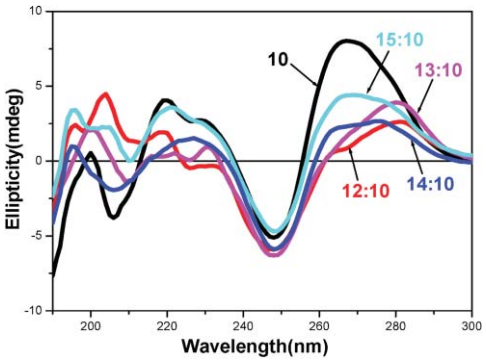


Fig. 4 CD graph of ONs **12**, **13**, **14** and **15** with RNA **10** (in 10 mM sodium phosphate buffer, pH = 7.0, 100 mM NaCl, and 0.1 mM EDTA). RNA **10**: r(5′ GCAAAAAAAACG 3′).

The use of naturally occurring α-amino acids in conjunction with an easily accessible nucleoside derived β-amino acid described in this communication thus provides a very simple peptide scaffold to create a nucleobase sequence for DNA/RNA recognition. Further studies on the mixed purine–pyrimidine sequences as well as incorporation of negatively charged and D-amino acids into the backbone to study the stereoelectronic requirements of the complexes are currently in progress in our laboratory.

KG thanks CSIR, New Delhi for a senior research fellowship. VAK thanks the Department of Science and Technology, New Delhi and the Director, National Chemical Laboratory for research grants.

Notes and references

1 (*a*) E. Uhlmann and A. Peyman, *Chem. Rev.*, 1990, **90**, 543–584; (*b*) S. M. Freier and K.-H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429–4443.
2 (*a*) J. Micklefield, *Curr. Med. Chem.*, 2001, **8**, 1157–1179; (*b*) J. Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628–1644; (*c*) J. J. Turner, M. Fabani, A. A. Arzumanov, G. Ivanova and M. J. Gait, *Biochim. Biophys. Acta*, 2006, **1758**, 290–300; (*d*) V. A. Kumar and K. N. Ganesh, *Curr. Top. Med. Chem.*, 2007, **7**, 715–726.
3 (*a*) Z. Dominski and R. Kole, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 8673–8677; (*b*) P. Sazani and R. Kole, *J. Clin. Invest.*, 2003, **112**, 481–486.
4 (*a*) P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500; (*b*) P. E. Nielsen, M. Egholm and O. Buchardt, *Bioconjugate Chem.*, 1994, **5**, 3–7; (*c*) M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895–1897.
5 E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem., Int. Ed.*, 1998, **37**, 2796–2823.
6 (*a*) A. De Mesmaeker, C. Lesueur, M.-O. Btvikrre, A. Waldner, V. Fritsch and R. M. Wolf, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 2790–2794.
7 (*a*) P. S. Pallan, P. von Matt, C. J. Wilds, K.-H. Altmann and M. Egli, *Biochemistry*, 2006, **45**, 8048–8057; (*b*) C. J. Wilds, G. Minasov, P. von Matt, K.-H. Altmann and M. Egli, *Nucleosides, Nucleotides Nucleic Acids*, 2001, **20**, 991–994.
8 E. Rozners, D. Katkevica, E. Bizdena and R. Stromberg, *J. Am. Chem. Soc.*, 2003, **125**, 12125–12136(*a*) E. Rozners and Y. Liu, *J. Org. Chem.*, 2005, **70**, 9841–9848; (*b*) Q. Xu, D. Katkevica and E. Rozners, *J. Org. Chem.*, 2006, **71**, 5906–5913.
9 T. K. Chakraborty, D. Koley, S. Prabhakar, R. Ravi and A. C. Kunwar, *Tetrahedron*, 2005, **40**, 9506–9512.
10 V. A. Kumar, *Eur. J. Org. Chem.*, 2002, 2021–2032.
11 (*a*) V. A. Kumar, Meena, P. S. Pallan and K. N. Ganesh, *Org. Lett.*, 2001, **3**, 1269–1272; (*b*) V. A. Kumar and Meena, *Nucleosides, Nucleotides Nucleic Acids*, 2003, **22**, 1101–1104.
12 R. J. Worthington, P. O'Rourke, J. D. Morral, T. H. S. Tan and J. Mickelfield, *Org. Biomol. Chem.*, 2007, **5**, 249–259.
13 (*a*) T. Govindaraju and V. A. Kumar, *Chem. Commun.*, 2005, 495–497; (*b*) T. Govindaraju and V. A. Kumar, *Tetrahedron*, 2006, **62**, 2321–2330.
14 (*a*) T. Vilaivan and G. Lowe, *J. Am. Chem. Soc.*, 2002, **124**, 9326–9327; (*b*) T. Vilaivan, C. Suparpprom, P. Harnyuttanakorn and G. Lowe, *Tetrahedron Lett.*, 2001, **42**, 5533–5536.
15 C. Suparpprom, C. Srisuwannaket, P. Sangvanich and T. Vilaivan, *Tetrahedron Lett.*, 2005, **46**, 2833–2837.
16 T. Vilaivan and C. Srisuwannaket, *Org. Lett.*, 2006, **8**, 1897–1900.
17 R. S. Varma, M. E. Hogan, G. R. Revankar and T. S. Rao, *US Patent* 5175266 (1992).
18 C. Thibaudeau, J. Plavec, N. Garg, A. Papchikhin and J. Chattopadyaya, *J. Am. Chem. Soc.*, 1994, **116**, 4038–4043.
19 (*a*) J. B. Epp and T. S. Widlanski, *J. Org. Chem.*, 1999, **64**, 293–295; (*b*) A. De Mico, R. Margarita, L. Parlanti, A. Vescovi and G. Piancatelli, *J. Org. Chem.*, 1997, **62**, 6974.
20 (*a*) P. Zhou, M. Wang, L. Du, G. W. Fisher, A. Waggoner and D. H. Ly, *J. Am. Chem. Soc.*, 2003, **125**, 6878–6879; (*b*) R. O. Dempcy, K. A. Brown and T. C. Bruice, *J. Am. Chem. Soc.*, 1995, **117**, 6140–6141; (*c*) G. Deglane, S. Abes, T. Michel, P. Prevot, E. Vives, F. Debart, I. Barvik, B. Lebleu and J.-J. Vasseur, *ChemBioChem*, 2006, **7**, 684–692.
21 P. E. Nielsen and M. Egholm, in *Peptide Nucleic Acids (PNA). Protocols and Applications*, ed. P. E. Nielsen and M. Egholm, Horizon Scientific, Norfolk, CT, 1999.
22 (*a*) R. S. Varma and M. E. Hogan, *Tetrahedron Lett.*, 1992, **33**, 7719–7722; (*b*) A. K. Singh and R. S. Varma, *Tetrahedron Lett.*, 1992, **33**, 2307–2310.
23 Stereoelectronically benign β-alanine was used as a spacer from the support during the synthesis of sequences **12**, **13**, and **15** for exerting a uniform steric effect at the C-terminus where amino acids were uncharged. In sequence **14**, positively charged L-lysine was preferred at the C-terminus because L-lysine at the C-terminus of peptide nucleic acids is known to have stabilizing effects on complexes with DNA and RNA due to electronic contributions (see reference 4).
24 J. D. Puglisi and I. Tinoco, Jr., *Methods Enzymol.*, 1989, **180**, 304–325.
25 (*a*) *Circular Dichroism: Principles and Applications*, ed. N. Berova, K. Nakanishi and R. W. Woody, Wiley-VCH, New York, pp. 703–736; (*b*) D. M. Gray, S.-H. Hung and K. H. Johnson, *Methods Enzymol.*, 1989, **246**, 19–34.
26 A. Hayen, M. A. Schmitt, F. N. Ngassa, K. A. Thomasson and S. H. Gellman, *Angew. Chem., Int. Ed.*, 2004, **43**, 505–510.
27 (*a*) K. Gogoi, A. D. Gunjal and V. A. Kumar, *Chem. Commun.*, 2006, 2373–2375; (*b*) K. Gogoi, A. D. Gunjal, U. D. Phalgune and V. A. Kumar, *Org. Lett.*, 2007, **9**, 2697–2700.

Authors Queries

Journal: **Chemical Communications**

Paper: **b716835g**

Title: **Chimeric (α-amino acid + nucleoside-β-amino acid)$_n$ peptide oligomers show sequence specific DNA/RNA recognition**

Please note that for ChemComm, communications must not exceed the three page limit unless this was agreed with the Editorial office before acceptance. Proofs that exceed this limit must be reduced before they can be published. If this is the case please indicate on your proof how this reduction can be achieved. Possible areas for shortening are: condensing lengthy introductions and/or discussion, excluding extensive data and experimental details (these may be transferred to ESI), omitting or reducing the size of artwork, removing a conclusions paragraph if this repeats earlier statements and reducing the number of references.

Editor's queries are marked like this... [1], and for your convenience line numbers are indicated like this... 5.

Query Reference	Query	Remarks
1	For your information: You can cite this article before you receive notification of the page numbers by using the following format: (authors), Chem. Commun., (year), DOI: 10.1039/b716835g.	