CHEMICAL SYNTHESIS AND STRUCTURAL STUDIES ON NON SUGAR-PHOSPHATE BACKBONE NUCLEIC ACIDS: OLIGOETHYLENOXY LINKED BIS-THYMINES AND PROLYL NUCLEIC ACIDS.

THESIS

SUBMITTED TO

THE UNIVERSITY OF POONA

FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

IN CHEMISTRY

BY

B. P. GANGAMANI

NATIONAL CHEMICAL LABORATORY

PUNE-411 008

SEPTEMBER 1997

Dedicated to My Parents and Shiva

CERTIFICATE

Certified that the work incorporated in the thesis entitled "**Chemical Synthesis and Structural Studies on Non Sugar-Phosphate Backbone Nucleic Acids: Oligoethylenoxy Linked Bis-Thymines and Prolyl Nucleic Acids**" submitted by **Mrs. B. P. Gangamani** was carried out at National Chemical Laboratory, Pune, by the candidate under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

K N Ganesh

(K. N. GANESH)
Research Guide
Head, Division of Organic Chemistry (Synthesis)
National Chemical laboratory
Pune-411 008.

September 1997

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled "**Chemical Synthesis and Structural Studies on Non Sugar-Phosphate Backbone Nucleic Acids: Oligoethylenoxy Linked Bis-Thymines and Prolyl Nucleic Acids**" submitted for the degree of Doctor of Philosophy in Chemistry to the University of Poona has not been submitted by me for a degree to any other university or institution.

Gangamani B P

National Chemical Laboratory
Pune- 411 008.
September 1997

(B. P. Gangamani)

ACKNOWLEDGMENTS

I take this opportunity to express the deep sense of gratitude for my guide, Dr. K. N. Ganesh . He has been extremely patient and understanding during the course of this work. His inspiration and encouragement has been instrumental in completion of this work. The fruitful scientific discussions with him have taught me the method of designing, executing and analysing problems in a true scientific spirit.

I am thankful to the Director, National Chemical Laboratory, for giving me the opportunity to work in this institute and for providing me with the necessary facilities.

My special thanks for Dr. (Mrs.) V.A. Kumar. She was always there to help me and encourage me both, scientific and personal problem. Her suggestions and encouragement has been an immense help during my stay at NCL. I will cherish the memories of my association with her.

I wish to thank Dr. C.G. Suresh, Div. of Biochemical Sciences, NCL, for the X-Ray crystallography studies. His sence of purpose and methodical way of working has taught me to do things better. I would also like to acknowledge Dr. Uday Maitra, IISc, Bangalore with whom I started my research career and learnt some of the basic principles and laboratory techniques. I also wish to thank Dr. Ganesh Pandey for allowing me to conduct the photodimerisation studies in his laboratory.

My sincere thanks are due to Dr. Vairamani, IICT, Hyderabad, for providing the FAB mass spectra. I wish to thank Prof. Balaram, MBU, IISc, Bangalore, and Dr. C.D. Mohan , CCMB, Hyderabad, for allowing me to record CD spectra. I thank Das & Ganesh for their help in recording the CD spectra.

My thanks to Dr. A. A. Natu, Dr. Shashidhar, and Dr. Pathak, and Mrs Anita Gunjal for their kind help during the course of this work. I wish to acknowledge Mr. Samual, Mr. Sathe, Dr. Sanjayan and Mrs. Phalgune for their assistance in recording the NMR; Mrs. Kunte and Mrs. Mane for the HPLC analysis and Mr. Mulla and Mrs. Santakumari for the mass spectra.

This page of acknowledgment would be incomplete without the mention of my colleagues, who have helped me in various capacities throughout my work and in bringing out this thesis. Thank you Moneesha, Leena, Ramesh, Vasant, Rajeev, Gopal, Vipul, Anand, Paradeep, Vallabh, Dinesh, Prakash, Rajendra, Meena, Dimpan, Nagmani, Praveen, Tanya, Bindu, and Shakti. I thank Pawar, Sunil and Makhar for their assistance. The friendly attitude of Saraswathi, Bando, Kalpana, Devi, Sudha, Sneha, Neelima, Badri, Sheshu, Hegde, Siddharth, Tripathi, Prasad, Joshi made my stay at NCL pleasant. My special thanks to my friend Achamma Thomas who was there when I needed.

I thank the UGC for providing me with the fellowship to carry out this work.

Last but not the least, I would like to thank my parents and my in-laws who have constantly encouraged me to excell. I can not forget to mention the kind co-operation extended to me by my husband Shiva. He has been extremely understanding and encouraging throughout the course of this work. Finally a word of fond acknowledgment for my son Mohith for being such a playful child.

Gangamani B.P.

CONTENTS

LIST OF PUBLICATIONS:

1. Synthesis and X-ray crystal structure of novel *trans-syn* thymine photodimers: Effect of a polyoxyethylene spacer chain on photodimer stereochemistry. **Gangamani B.P.**, Suresh C. G. and Ganesh K. N. *J. Chem. Soc. Chem. Commun.* **1994**, 2275-2276.

2. The rare *trans-syn* thymine photodimers: Structure of polyoxyethylene-linked bisthymines and the derived *trans-syn* thymine photodimers. Comparison of the stereochemistry before and after photodimerization. Suresh C. G., **Gangamani B.P.**, and Ganesh K. N. *Acta. Cryst.* **1996**, *B-52*, 376-383.

3. Synthesis of N^{α}-(purinyl/pyrimidinyl acetyl)-4-aminoproline Diastereomers with potential use in PNA synthesis. **Gangamani, B. P.**, Kumar, V. A. and Ganesh, K. N. *Tetrahedron* **1996** *52*, 15017-15030.

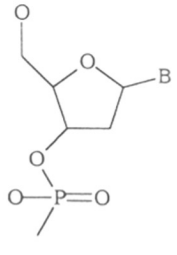
4. 2-Aminopurine-Peptide Nucleic Acids (2-*ap*PNA): Intrinsic fluorescent PNA analogs for probing PNA-DNA interaction dynamics. **Gangamani, B. P.**, Kumar, V. A. and Ganesh, K. N. *J. Chem. Soc. Chem. Commun.* **1997** (in print)

5. Peptide Nucleic Acid (PNA)-spermine conjugates: Synthesis and formation of PNA-DNA hybrids with enhanced stability. **Gangamani, B. P.**, Kumar, V. A. and Ganesh, K. N. *J. Org. Chem.* **1997** (communicated)

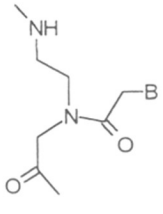
ABSTRACT

Chapter 1: Introduction

The antisense and antigene strategy for treatment of diseases at the level of gene expression has attracted a wide attention in medicinal chemistry. The interference with gene expression can be accomplished by the binding of an oligonucleotide to DNA or RNA. In order to make the oligonucleotides stable to cellular conditions, many modified oligonucleotides have been designed, synthesized and studied. Among these phosphorothioates and more recently peptide nucleic acids (PNA) have emerged as promising candidates. PNA is a DNA mimic in which the sugar phosphate backbone is replaced by 2-N-(aminoethyl)glycine carrying the acetylated nucleobase. PNA has proved to be a highly potent DNA mimic capable of binding to complementary DNA via strand displacement and form very stable complexes. The hybridization properties and the orientational preference are highly dependent on the sequences. Homopyrimidines bind as $(PNA)_2$:DNA complex and mixed sequences form PNA:DNA duplex. They bind DNA in both orientations, antiparallel complexes being more stable than the parallel complexes. In the design of PNA the distance between the two adjacent nucleobases are similar to that in DNA and the backbone amide linkages give a helical twist. This is reflected in the formation of PNA:PNA hybrids which are helical as indicated by CD studies. The structural aspect of PNA complementation with DNA/RNA/PNA and their hybridization properties are discussed.



I



II

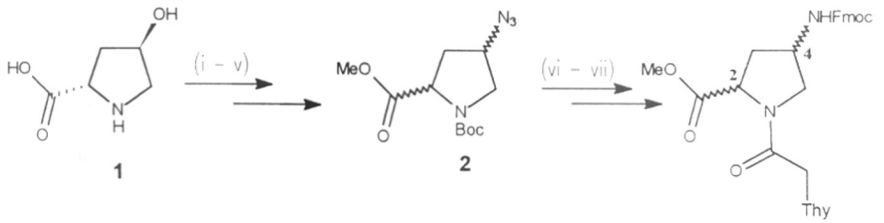
Their poor solubility and unimpressive cellular uptake of PNA has led to modification of the PNA backbone. Among these, modifications to introduce a chiral center on the backbone has some influence of binding and to obtain better orientational preferences. Many attempts made to improve PNA properties such as binding affinity, specificity and cellular uptake by various chemical modifications of PNA backbone are discussed. PNA-peptide and PNA-DNA chimeras and their properties are also described. Finally unraveling the structure-activity relations of PNA have facilitated the design of new DNA analogues.

The high stability of PNA-DNA/RNA complexes has brightened the scope of PNA emerging as a powerful biotechnological tool, a number of biological application oriented studies are discussed here. The future aspects of PNA like molecules as one of the most exclusive drugs is bright as evident by properties of number of modifications and chimeras studied reflect. PNA has opened up a new era culminating in interfacing a number of aspects of molecular genetics, medicinal chemistry and evolutionary biology.

Chapter 2: Synthesis of N^{α}-(Purinyl/Pyrimidinyl acetyl)-4-Amino Proline with Potential use in Chiral PNA synthesis

This chapter describes the modification of achiral peptide nucleic acids to influence its properties in a desirable way. The design of a PNA analogue involved the introduction a methylene group between α-C of the glycine and β-C of ethylenediamine of the same monomeric unit of PNA. This in turn generated a five membered heterocyclic ring along the backbone carrying two asymmetric centers at C4 and C2. The synthetic strategy adopted to build all four diastereomeric monomeric units of 4-(*R*/*S*)-amino-N-(acetyl nucleobase)-2-(*R*/*S*)-proline starting from a single starting material 4-*trans*-hydroxy-L-proline is outlined. The four diastereomers were obtained by designing four different pathways with different inverstion steps to obtain the desired sterochemistry. Scheme 1 represents a model synthesis of these monomers.

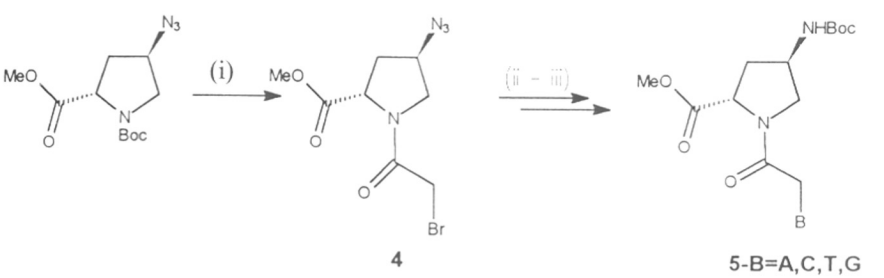
Scheme 1



(i) Ac_2O/AcOH-HCl, (ii) MeOH/$SOCl_2$ (iii) $BocN_3$/TEA, (iv) MsCl/Py or TsOMe/DEAD, (v) NaN_3/DMF (vi) H_2-Pd,FmocCl, (vii) 50%TFA/CH_2Cl_2, (viii) Thy-acetic acid, HOBT,DCC

An alternative methodology used to introduce any of the nucleobases (A,G,C and T) via a common precursor is designed. It is demonstrated that this method of using a common precursor to obtain 4*R*-amino-N(ACTG-1/9 acetyl)-2*S*-proline was much simpler and can be obtained in better overall yields (Scheme 2). Solution phase synthesis was used to prepare two dimers differing in stereochemistry at one center and their structural properties studied. PNA monomer carrying a fluorescent nucleobase was also prepared for incorporation into oligomers and study the hybridization and self-organization of PNA using a intrinsic fluorophore.

Scheme 2



(i) 50%TFA/CH_2Cl_2, $BrCH_2COCl$, (ii) Thymine/K_2CO_3, (iii)H_2-Pd, $BocN_3$

Chapter 3: Solid Phase Synthesis of Prolyl Nucleic Acids and their Biophysical Studies.

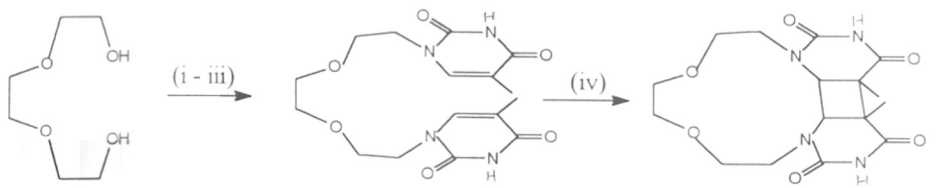
The enantiospecific monomers described in the previous section were used in the synthesis of homopyrimidine prolyl nucleic acid (PrNA) oligomers. The synthesis was accomplished using well established methods adopting Boc strategy. Merrifield resin was used as the solid support and the couplings were carried out using TBTU and DIPCDI as the *in situ* coupling agents. The peptide oligomers were cleaved of the resin either using TFMSA/TFA to obtain the terminal carboxylic acid or aminolysis with polyamines to obtain corresponding C-terminal amides.

The homopyrimidine PrNA consisting of homochiral units of L-*trans*, L-*cis* and D-*trans* did not show any cooperative binding with complementary DNA. The introduction of the chiral unit at the end of PNA and in the middle of the PNA was accomplished by introducing the enantiomeric pair L-*trans* and D-*trans*. The effect of these chiral monomers either at the end or in the middle is studied be UV and CD melting experiments. These show a preferencial binding to complementary DNA, as D-*trans* binds in an antiparallel orientation and L-*trans* binds in parallel orientation in 100 mM NaCl salt conditions. The binding in low salt concentration is observed to be better and consistent with the salt dependency of the PNA. The introduction of spermine at the C-terminus increases the solubility of the oligomers in water and also stabilizes the duplex formed preferentially in antiparallel orientation. The formation of duplex of the mixed sequence carrying a prolyl unit in the end and in the middle was also followed by CD.

Introduction of 2-aminopurine in PNA oligomer in the middle of a 10 mer sequence provides a handle to study various properties of PNA by following fluorescence changes. In the present study the binding constants for the PNA-DNA and spermine-PNA:DNA complexes were determined using fluorescence titration and found that the *sp*PNA binds 20 times better than unmodified PNA. The kinetics of duplex formation was also followed by a change in fluorescence intensity with time upon mixing PNA and DNA. 2-amino purine being an intrinsic fluorophore could be used to observe the changes in single strand upon heating by measuring the anisotropy at various temperature.

Chapter 4: Synthesis of Polyoxyethylene-bis-thymines and thier Photodimerisation.

The damaged caused by UV radiation on the hereditary material DNA is well established. The most susceptible sites to UV radiation in DNA are the adjacently stacked pyrimidine bases. The cyclobutane type of photodimers caused as a product of UV radiation are the basic cause of various mutation leading to skin cancer. All cells have effective repair mechanism to over come such damage but the depletion in ozone has increased the changes of mutations to radiation exposure.



1 2 3

Scheme 1: (i) $SOCl_2$/Py, (ii) NaI/Acetone, (iii) Thymine, HMDS, TMSCl, (iv) hν/Acetophenone

This chapter describes the synthesis of polyoxyethylene bisthymines where the two thymines are linked through various polyethlene glycol, to observe the effect of the spacer chain in presence of alkali metal ions. Though no observable effect of the metal atom was found in this study, it was observed that irrespective of the spacer chain the product obtained from photodimerisation was the unusual and rare *trans-syn* isomer. A comparison of two compounds, one with triethyleneglycol and other with catacholethylene linker, which differ in their spacer chain indicates interaction differences in cyclobutane conformation as a function of the nature of the spacer chain. The crystal structure of the two of the bisthymines and their photoproducts was solved and the results give us an insight to assess the structural effects and stereochemical outcome of photodimerization of model compounds.

ABBREVIATIONS

β-ala	β-Alanine
2-*ap*	2-aminopurine
A	Adenine
aeg	Aminoethylglycine
aq	aqueous
ala	Alanine
ap	Antiparallel
C	Cytosine
CBz	Benzyloxycarbonyl
CD	Circular Dichroism
dA	2'-Deoxyadenosine
dC	2'-Deoxycytidine
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DCU	Dicyclohexyl urea
DEAD	Diethylazodicarboxylate
dG	2'-Deoxyguanosine
DHP	Dihyropyran
DIAD	Diisopropylazodicarboxylate
DIPCDI	Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMAP	N,N-dimethyl-4-aminopyridine
DMF	Dimethylformamide
DNA	2'-deoxynucleic acid
eda	ethylenediamine
EDTA	Ethylenediaminetetraacetic acid
F-moc	9-Fluorenylmethoxycarbonyl

FPLC	Fast Protein Liquid Chromatography
G	Guanine
gly	Glycine
his	Histidine
HOBT	1-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
IR	Infra-red
MF	Merifield resin
MS	Mass Spectrometry / Mass spectrum
NMR	Nuclear magnetic resonance
p	Parallel
Pet-ether	Petroleum ether
PPh_3	Triphenylphosphine
PNA	Peptide nucleic acid
PrNA	Prolyl nucliec acid
Pro	Proline
RNA	Ribonucleic acid
sp	Spermine
T	Thymine / 2'-Deoxythymidine
t-Boc	*t*-Butoxylcarbonyl
TBDMS	*t*-Butyldimethylsilyl
TBTU	2-(1H-benzotriazole)-1,1,3,3-tetrabutyl uronium tetrafluoroborate
TEA	Triethylamine
TEAA	Triethylammoniumacetate
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethane sulfonic acid
THF	Tetrahydrofuran
THP	Tetrahydropyranyl
UV-Vis	Ultraviolet-Visible

CHAPTER 1

Introduction

1.1 Introduction

Molecular recognition is a fundamental principle in biology. The classical examples include enzyme-substrate, antigen-antibody, hormone-receptor and drug-DNA as well as protein-DNA and protein-protein interactions. DNA is the basic genetic material which consists of two complementary strands held together by Watson-Crick hydrogen bonds through A:T and C:G base pairs.[1] Much of the biological functions of DNA are exercised through molecular recognition. Ever since Zamechnik and Stephensen[2] proposed a radical new concept in medicinal chemistry, that of "antisense" oligonucleotides[3] as potential therapeutic agents, an enormous amount of research work has originated to search for the most potent DNA mimic. The prime requisites of such molecules are binding to the target with higher specificity and resistant to enzymatic degradation. The basis of genetic diseases is the gene mutations which in turn either overexpress or code for nonfunctional proteins. Antisense drugs are modified synthetic oligonucleotides that work by interfering with mRNA. Excess protein production can be effectively blocked by small stretches of oligonucleotides which bind to mRNA that serves as a template to synthesize the protein. Since the RNA template is known as the sense strand, the complementary sequence that binds to it is called antisense strand.

Interference with gene expression can be accomplished by binding of oligonucleotides to duplex DNA through Hoogsteen hydrogen bonds which leads to the formation of triple helix;[4,5] a higher order structure of DNA. The double stranded genetic DNA itself can thus act as a target for oligonucleotides or their analogues and successful triplex formation takes place at homopurine stretches of DNA.[6] This therapeutic approach is known as "antigene" therapeutics and is limited to homopurine stretches.[7]

The major limitation in using natural oligonucleotides as therapeutic agents is that they are rapidly degraded by cellular nucleases. Furthermore, the ability of the DNA strand to cross a cellular membrane is poor. As a result, a significant number of modifications of the oligonucleotide structure have been made to make them stable towards cellular enzymes.[8] The various possible sites of modifications on DNA sequence are shown in Figure 1.

The phosphate modifications such as phosphorothioates (**2a**), phosphorodithioates (**2b**), methylphosphonates (**2c**), phosphoramidates (**3**) and phosphotriesters (**4**) form the first generation 'antisense' oligonucleotides (Figure 2) which have already shown promising results.[9] Among these, the phosphorothioates are currently in phase III clinical trials.[10,11] There are various modifications on the sugar as well as on the base moiety which give highly stable complexes as well as remarkable stability towards cellular degradations.

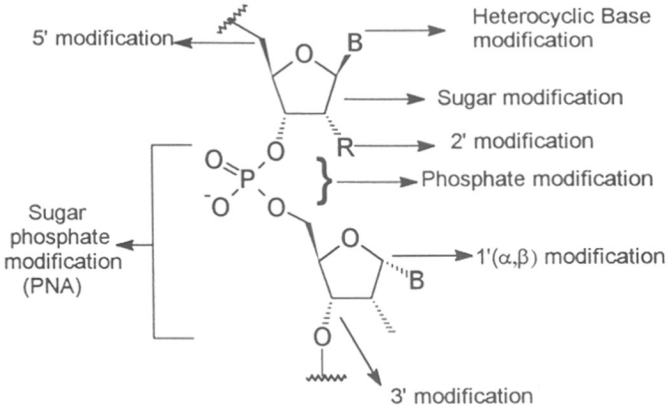


Figure 1: Position and types of oligonucleotide modifications.

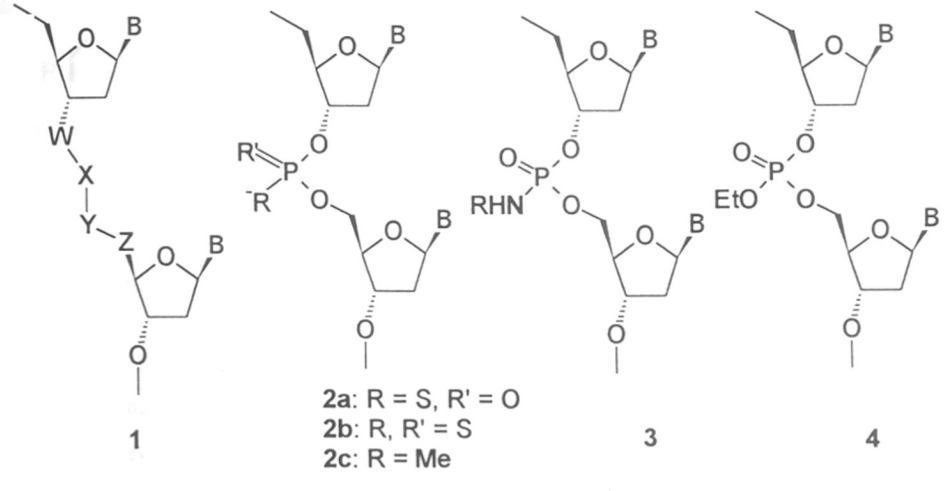


Figure 2: Phosphate modifications.

These conceptual developments along with structural modifications (chemical/ biological) have encouraged the search for better analogues of DNA. Most of the modifications concentrated on replacement of phophodiester linkage by other four atom chains W-X-Y-Z as in structure **1** (Figure 2) are summarized in Figure 3.[8] Few of these modifications have shown fairly good binding with DNA/RNA but none of them have shown the potency to be an outstanding drug.

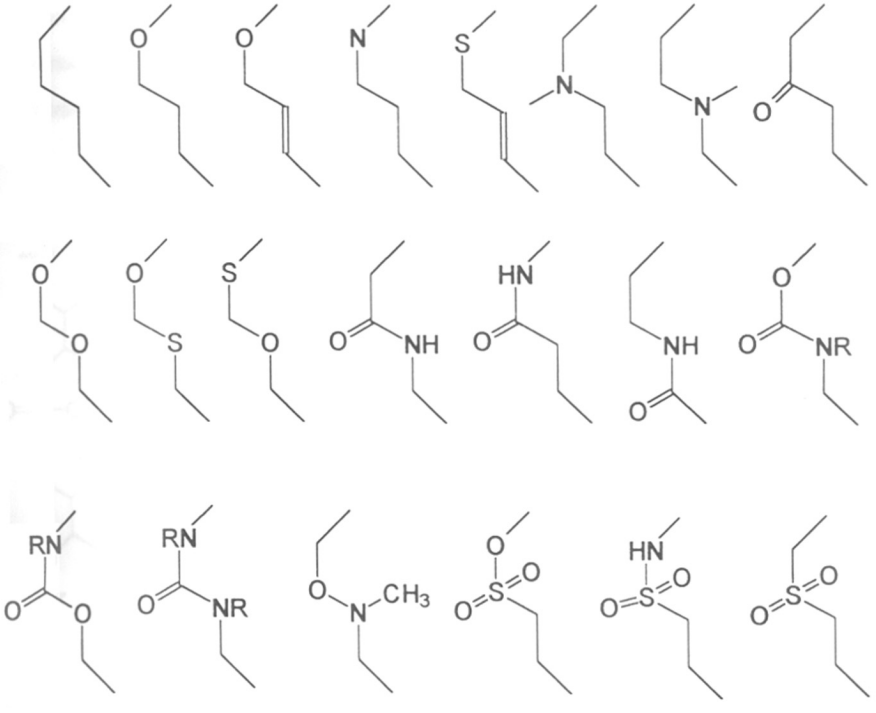


Figure 3: Oligonucleotide backbone modifications of the X-Y-Z-W part structure in **1**

The replacement of ribose sugar by hexose had a major setback since the corresponding oligonucleotides **5** do not show any cooperative binding with natural DNA.[12] They still show linear complexation with a complementary hexose nucleotide oligomer. The carbocylic analogues also form linear duplexes instead of double helical structure which are structurally different. Only a few attempts to replace the ribose-

phosphate backbone have been so far successful. These include morpholino oligomers (Figure 4) wherein the monomers are linked through the carbamate linkages[13] (structure 6) and show good stability towards complementary oligonucleotides. The stability of the these oligomers is attributed to the neutral nature of the backbone. The replacement of the carbamate linkage with phosphoramidate **7** (Figure 4) exhibits higher stability on hybridization with complementary DNA.[8]

Recently, DNA analogues with amide or peptide linkages **8** was shown to form sequence-specific highly stable complexes with DNA.[14] Peptide nucleic acid (PNA) is homomorphous with DNA, it is achiral and has a neutral backbone. The binding efficiency is attributed to the neutral nature of the backbone.

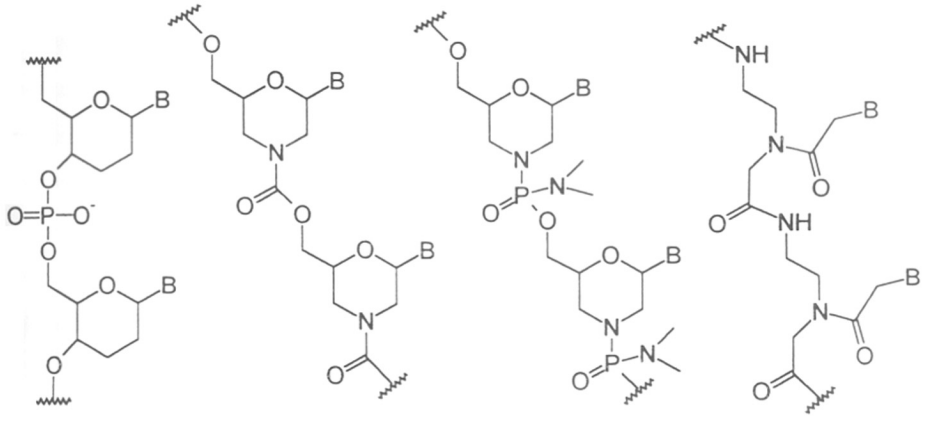


Figure 4

Some exclusive characteristics of PNA such as strand displacement, have important biological applications in therapeutics and diagnostics (*see Section 1.4*).[15] Since the monomeric units are achiral and the synthesis of the oligomers involves standard solid phase peptide synthesis, it is easy to obtain these oligomers in large scale. Homopyrimidine PNA sequences give highly stable complexes with DNA due to the formation of 1:2 (DNA:PNA) triplex complexes. The molecular modeling studies[16,17] have

shown that PNA single strand has an ordered structure that allows the bases to be properly stacked in a helical form. The neutral nature of the backbone makes it less soluble in water and results in the formation of aggregates. This problem was overcome by introduction of a lysine amino acid at the C-terminus.[14]

Triplex formation of $(PNA)_2$:DNA in parallel orientation

Triplex formation of $(PNA)_2$:DNA in antiparallel orientation

PNA:DNA Duplex formation in parallel orientation

PNA:DNA Duplex formation in antiparallel orientation

The PNA binding to DNA with NH_2 terminus pointing towards 5' end of DNA is designated as parallel and NH_2 end pointing towards 3' end is known as antiparallel. PNA is known to bind to DNA in both antiparallel and parallel orientation. In the triplex form parallel orientation is preferred over antiparallel, but in duplex, antiparallel is preferred to parallel mode. The following section discusses the structural features of PNA complexes with DNA and PNA.

1.2 Structure and hybridization properties

1.2.1a ***(PNA)$_2$:DNA triplexes***: Homopyrimidine PNA sequences form 2:1 (PNA:DNA) triplex complexes with the DNA strand by strand displacement (see later).[18] The UV thermal melting experiments do not indicate a biphasic transition as in the case of DNA triplex formation; instead they show a monophasic transition. This is an indication that the Watson-Crick base pairing and Hoogsteen base pairing are of equal strength and both dissociate simultaneously. The C-rich sequences show pH dependence as observed for DNA triplexes even in (PNA)$_2$:DNA triplex formation, in which cytosine needs to be protonated to bind in the Hoogsteen mode,[19] but the stoichiometry of the C-rich sequences even at pH 9 was determined to be (PNA)$_2$:DNA. Another important feature is that the triplexes are more stable in parallel orientation as compared to antiparallel orientation.[20] The stability of the complex is proportional to the length of the sequence with an increase in T_m of 1 °C per base pair. Circular and linear dichroism studies have indicated that (PNA)$_2$:DNA triplexes are very similar to conventional DNA triple-helix in forming right handed helix.[21] The stability of these triplexes is highly dependent on the salt concentrations. At low salt concentrations highly stable triplexes are formed while at higher salt concentrations they are slightly destabilized.[22] The stability of the complexes at high salt concentration is equivalent to DNA, which is an indication that the stability at low salt concentration is due to neutral nature of the backbone. The kinetic studies indicate instant triplex formation at low concentration while a pronounced hysterisis at higher salt conditions pointed to a very slow rate of triplex formation.[21] The presence of a covalently linked intercalator at the end of PNA stabilizes the triple helix.[23] On the other hand, the groove binders and externally added intercalators destabilize the complex.[24] The external addition of spermine to the PNA-DNA complex leads to precipitation of the complex at higher concentration of the polyamine.[22]

1.2.1b ***Bis-PNAs***: Since the PNAs are made of neutral backbone their solubility and hybridization are heavily dependent on the salt concentration. The rate of hybridization is very low at physiological salt concentration although they form stable complexes. In order to improve this, two PNA strands which form components of triple helical structure with

DNA were linked together to reduce entropy and convert the binding into a bimolecular process. Bis-PNAs synthesized using a spacer linker like triethyleneglycol or a lysine analogue[25] do form highly stable complexes with the kinetics of complexation increasing under physiological salt concentrations.

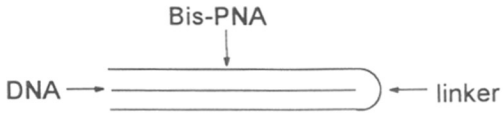


Figure 5: Bis PNA binding to DNA the two PNA strands are held together though a linker

The introduction of isocytosine in the third strand of DNA in place of cytosine is known to give rise to stable DNA triple-helix at physiological pH.[26] The bis-PNA having an antiparallel clamp was prepared and these show superior DNA binding properties.[27] Even in the bis-PNA, the parallel mode of binding to DNA is preferred and the third strand adopts the normal parallel orientation. When isocytosine is present in the Watson-Crick-motif, the triplex shows pH dependency and when it is present in the Hoogsteen motif there is no observable pH dependency. The Figure 6 shows the hydrogen-bonding pattern of cytosine and pseudocytosine when present in the third strand.

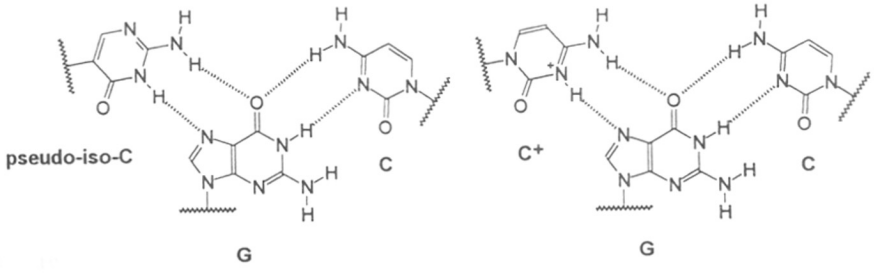


Figure 6

Bis-PNA linked through a short peptide chain was synthesized and the crystal structure of this bis-PNA:DNA triplex is solved.[28] The studies on these systems revealed the occurrence of different structural motif which is different from the known stable helical forms that nucleic acids normally adopt. The structure of the triplex was a hybrid of B and A forms of DNA and designated as P-form. The P-form helix has a base tilt similar to B-form DNA, where the paired bases are displaced from the helix axis even more than in A form DNA. The structure of this triplex arises from 2:1 complexation of PNA with DNA. Interestingly, the phosphates of the DNA backbone formed hydrogen bonds to amide protons present in the PNA backbone of the Hoogsteen strand. In a recent report, Wittung *et al.*[29] have demonstrated that a cytosine rich homopyrimidine PNA adds to double stranded polynucleotide targets through Hoogsteen strand forming PNA(py): DNA(pu):DNA(py) motif of triplex. This development can be exploited to form efficient antigene agents.

1.2.2 PNA:DNA and PNA:RNA duplexes

PNA sequences made of all four base pairs are shown to form stable duplexes with DNA.[30] Though the antiparallel orientation is the preferred mode of binding, they do form stable duplexes in parallel orientations also. The CD spectra of these complexes show base pair complementarity and are similar to the right handed helical B-DNA.[31] The detailed structural information about the helical arrangement of such PNA:DNA hybrids have been revealed by NMR solution structural studies using 2-D NOESY and 2-D COSY experiments.[32] A hexameric PNA, GAACTC formed a 1:1 complex with complementary antiparallel RNA with Watson-Crick-CRICK base pairing similar to the A-form of RNA duplexes.[33] The results show that no intramolecular hydrogen bonding is present on the PNA backbone as proposed by Bruice *et. al.*[16] The A-form conformation places the PNA tertiary amide carbonyl group in a position isoteric to the RNA C2', facilitating maximum exposure of the backbone carbonyl to water. Similarly the solution structure of PNA:DNA duplex with antiparallel orientation was solved by NMR studies. The PNA:DNA complex adopts a B-like structure with 2'-deoxysugar in the C2'-*endo* conformation. The observed base pairing was more like the Watson-Crick base pairing in

DNA. A mixed sequence PNA containing homopurines was also shown to bind dsDNA by initial formation of duplex with strand displacement,[34] that was detected by photofootprinting experiments. Furthermore, the alternating TG-PNA sequences were shown to bind DNA targets and form a strand displacement complex containing PNA:DNA duplex.[29] The strand invasion binding of PNA is a highly sequence dependent process and more studies are required to elucidate a general criteria for strand invasion duplex formation.

1.2.3 PNA:PNA duplexes and PNA:PNA:PNA triplexes

The basic molecular recognition criterion in DNA and its analogues is the base pairing pattern. PNA forms stable duplex and triplex with DNA and interestingly similar isomorphous structures can be formed with complementary PNA. Nielsen *et al.*[35] have shown that such PNA:PNA duplexes are formed with high stability (more stable than PNA:DNA duplex) and these exhibit helical structures similar to the DNA double helix. The binding of these complementary PNA also follow the same rule as in DNA where the antiparallel mode of binding shows much higher stability compared to the parallel mode. The formation of these helical structures, as studied by the change in absorbance as a function of time, is instantaneous. The presence of a C-terminal lysine residue induces chirality to the helical structure and this was observed by CD wherein L and D lysine at the carboxy terminus give exactly opposite handedness to the helical PNA.[36] The evolution of CD signal of the helix formation as a function of time is much slower compared to the kinetics observed by UV-absorption as a function of time. This was explained to be a result of the slow reorientation of the instantly formed duplex due to the induction of chirality by the terminal lysine. The PNA duplex with no chiral moiety failed to show CD band because of the presence of a racemic mixture of both right and left handed helices. The presence of a chiral lysine at the terminus does not affect the binding properties of the PNA with DNA. The $(PNA)_3$ triplexes[37] [$(PNA\text{-}T_{10})_2$:PNA-A_{10}] are also formed and found to be more stable than the PNA_2:DNA triplexes. The introduction of chiral cyclohexane within the backbone reduces the affinity but no preferential binding between the two enantiomers was observed (*see later section 1.3.1*).

1.2.4 Strand Displacement

The most interesting property of PNA is the formation of PNA_2:DNA triplexes containing homopurines through strand displacement.[14] The high stability of the complexes does not favor dissociation. Since PNAs can bind in either orientation, the formation of the triplex takes place readily. These triplexes are so stable that the homopyrimidine/ homopurines bind to DNA duplex by formation of a D-loop (Figure 7) which was shown unambiguously by photofootprinting experiments.[14] The unpaired DNA strand present in the D-loop is highly susceptible to attack by S1 nuclease and oxidation with permanganate.[38] The strand displacement has also been visualized by electron microscopy and furthermore, it was shown that the binding of PNAs to closed-circular DNA results in unwinding of the double helix by approximately one turn per ten bases. The formation of D-loop by strand displacement is highly salt dependent, being efficient at low salt concentration (less than 50 mM).[39] However, once the strand displacement complex is formed, the salt concentration can be raised to more than 0.5 M NaCl without disrupting the complex. The low salt concentration might have a role in destabilizing the dsDNA and initiate PNA binding. A very slow dissociation of PNA from dsDNA was monitored by adding excess of an oligonucleotide complementary to PNA strand in a strand displacement complex.

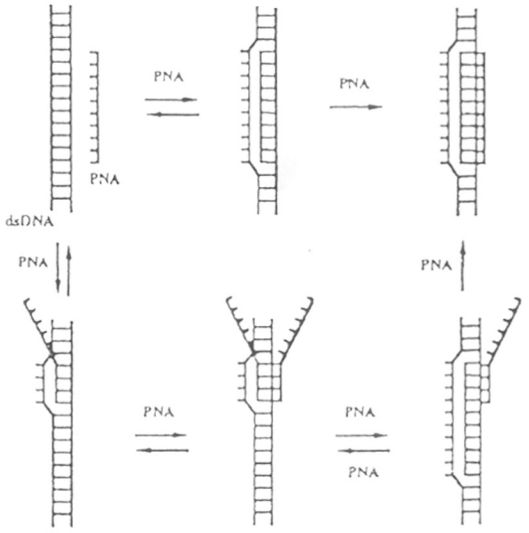


Figure 7: Formation of D-loop by PNA strand invasion.

The design of bis-PNA where the two PNA strands are clamped by a positively charged lysine-aminohexyl linker lead to considerably faster and more efficient binding to the duplex.[25] These strand displacement complexes were formed at relatively higher salt concentration (100 mM NaCl). The cationic charge on the linker increases the solubility and its interaction with the negatively charged DNA backbone leads to higher stability and faster formation of the duplexes.

1.3 PNA Modifications and PNA Conjugates

1.3.1 PNA backbone modifications

The low selectivity barrier between binding of PNA in both parallel and antiparallel direction has led to chemical modifications of the backbone to obtain better orientational preferences. The most simple modifications on the PNA backbone directed to improve its binding properties and hence making it a better DNA mimic, was reported by Nielsen *et. al.*[40,41] The PNA backbone was extended by inserting a methylene group in the PNA monomer **9** (Figure 8) at different positions; (i) in aminoethyl part as in **10** or (ii) in the glycine part as in **11** or (iii) the acetyl linker part **12**.

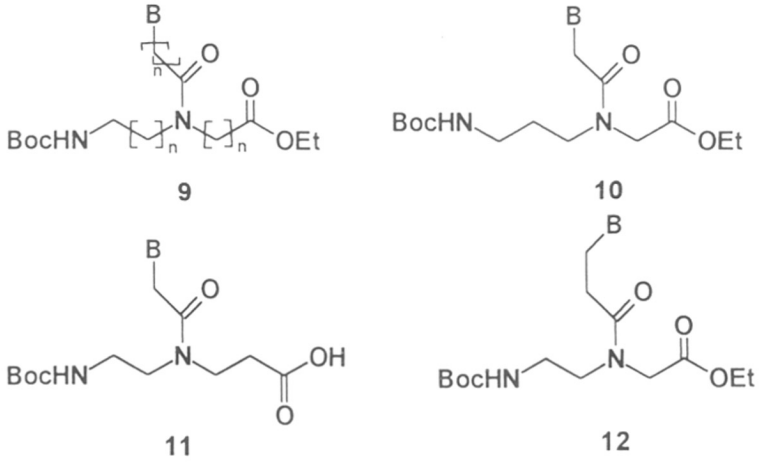


Figure 8

The oligomers prepared using these extended backbone units did not form any stable complexes with complementary duplexes. Incorporation of a single monomeric unit into PNA oligomer also showed a large destabilization though the selectivity was maintained. From this it is evident that the aminoethylglycine backbone is the most optimum structure. The methylene group insertion showed more destabilization when present in the acetyl linker part as in structure **12**.

To observe the effect of the amide linkage on the backbone, heterodimer **13** was synthesized by inverting the amide linkages[42] as shown in Figure 6 (retro inverse PNA[43]). The preparation of heterodimer **13** involved the synthesis of appropriately protected monomeric units. The incorporation of this dimeric unit in PNA has shown some interesting properties, the most important aspect being its slightly better stability as compared to the PNA. This implies that the PNA backbone structure is amenable for modifications without disturbing the binding efficiency and recognition properties.

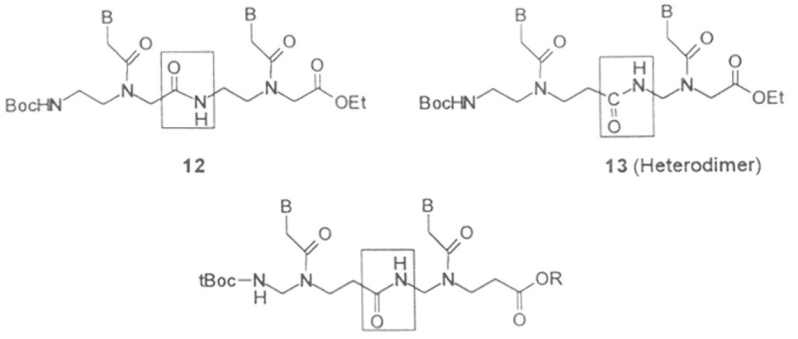
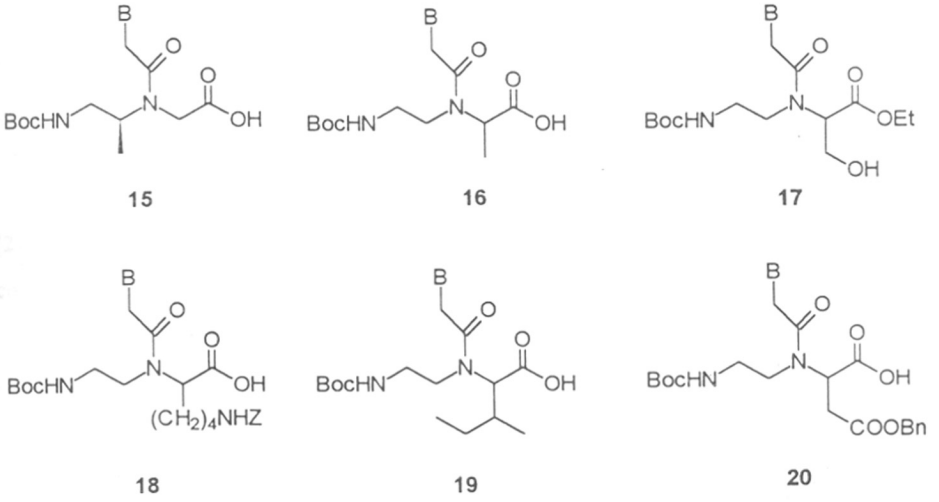


Figure 9

The above observation that certain structural modifications on the PNA backbone do not affect the binding properties, led to the introduction of chirality into the backbone to improve the orientational preferences. The introduction of alanine[44,45] in the backbone was envisaged to control the direction of binding. Such a replacement with L-alanine **16** led to a slight destabilization whereas D-alanine **16** showed marginal stability. In both cases antiparallel orientation was preferred over parallel mode. Though introduction of

methyl group in the backbone does not alter the hybridization properties, bulky replacements like isoleucine **19** were shown to destabilize the duplex formation.[46] A further study where glycine was replaced with L or D lysine as in **18** also showed stabilization of duplexes with D-lysine incorporation, indicating that D-amino acids are perhaps better tolerated than the L-counterparts.[46] Though the introduction of lysine increased solubility, the electrostatic interaction did not show much contribution in duplex stability. The introduction of acidic amino acid as in **20** did show destabilization. The incorporation of a single backbone modified PNA monomer having the nucleobase thymine and a chiral cyclohexyl moiety in its backbone into PNA-T_{10} showed preferential selectivity in forming triplex with the purine strand of PNA-A_{10} as compared to the corresponding DNA-A_{10} strand.[37]



15 **16** **17**

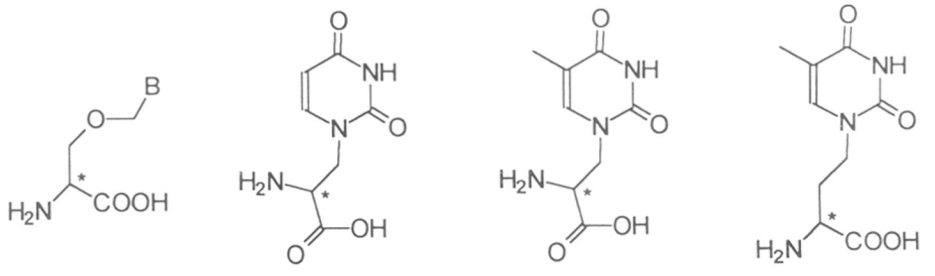
18 **19** **20**

Figure 10

1.3.2 PNA with nucleobase attached to α-C of glycine

The replacement of the sugar-phosphate backbone by a peptide carrying nucleobase was envisaged in the early 1970s.[47] The developments in this direction were motivated by isolation of naturally occurring peptides containing a uracil base known as willuridine (**22**). There are quite a few reports on the synthesis of polythymine peptides in

the 70's but unfortunately, systematic binding studies with DNA were not done at that time. After many efforts finally Nielsen *et. al.*[16] invented an optimum replacement for sugar-phosphate backbone with peptidic analogue. In this section, the efforts to synthesize PNA analogues using amino acid units carrying the nucleobase on the α-position are discussed.



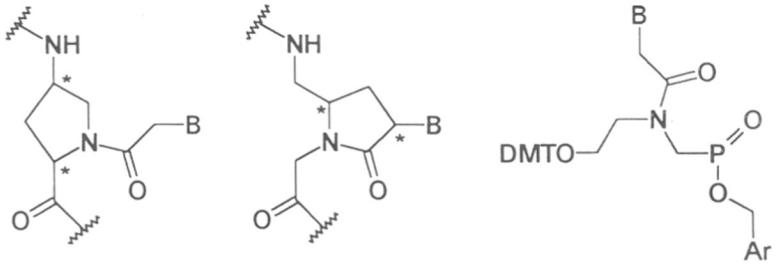
21 B=T **22** **23** **24**

Figure 11

One of the interesting modifications is the preparation of Boc(serCH_2-T)-OMe (**21**), which contains the O-C-N linkage to the nucleobase.[48] This amino acid monomer coupled alternatively with glycine gave tetrapeptide with two thymine units. Taddie *et. al.*[49] have used homoalanyl unit **24** and prepared both A and T monomers which were assembled alternatively with glycine to obtain a homochiral oligomer that forms a self - complementary duplex. They showed that a 10-mer duplex had much lower T_m compared to PNA:PNA duplexes. The work was extended by Shah *et al.*[50] where they have prepared homopyrimidine oligomers similar to those reported by Taddie *et al.*[49] but with alternating *R* and *S* configuration using **24**. These heterochiral PNA oligomers do form stable duplexes with DNA.

Diederichsen *et al.*[51] have used a thymine analogue of willuridine **23** which is also an alanyl thymine amino acid to prepare homooligomers. These have alternating R and S configuration but do not contain a glycine unit as in the earlier cases. They form linear self-complementary complexes with alternating A and T units. An important observation is

that the geometry is favorable for the formation of non-WC type A-A complexes which are more stable than the corresponding WC type A-T complexes. Further, to improve upon the A-T binding ability, an oligomer consisting of two alternating structures were prepared using **23** and **24**, with a heterostereochemistry.[52] These were shown to form stable self-complementary complexes with A-T type of base pairing.



25 **26** **27**

Figure 12

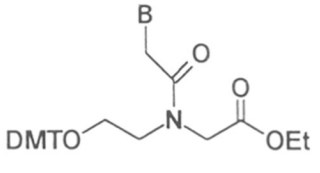
Increasing the flexibility of the backbone was also attempted to study the binding efficiency of these PNA analogues to DNA.[53-55] But these oligomers with higher flexibility do not form any stable duplexes. When the present work was in progress[56] two reports appeared where 4-*trans*-hydroxy-L-proline was used to prepare chiral analogues of PNA. Lowe *et al.*[57] have adopted the modification where the base is directly attached to the 4-position of proline ring as in **26** and built peptides with the proline moiety and glycine in alternating positions, although the biophysical studies of these oligomers are not reported. In another report which followed our publication[56] Jordon *et al.*[58] have prepared proline oligomers with 4-aminoproline carrying nucleobase with an acetyl linker on the N1 position as in structure **25**. They have found some interesting binding properties employing homopyrimidine oligomers which are built with alternate prolyl unit and PNA unit. They have not reported any studies directed with hetero-oligomers as reported in the present work. Another interesting development is the reversal of the peptide linkage to the phosphate linkage. Achiral-acylic PNA monomers **27** are assembled to obtain an oligomer

with phosphonate linkage.[59] These compounds are water soluble due to the negative charges but their binding efficiency was not evaluated.

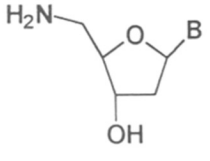
1.3.3 PNA-Analogues

1.3.3a PNA:DNA Chimeras: It is a well established fact that PNAs bind to complementary DNA with high efficiency and specificity and are stable towards cellular enzymes (protease and nuclease). A combination of PNA and DNA may reduce the self - aggregation of the PNA and further increase its bioavailability.[60-63] Such chimeras might also provide RNase H activation upon complexation with mRNA in antisense applications or even function as a primer for DNA polymerase for use in PCR. Nielsen *et. al.*[61] have used linker **28** (Figure 13) to conjugate DNA with PNA. However, this type of head to tail orientation (5'DNA3'-N-PNA-OH) did not show any significant thermal stability. The duplexes formed by such chimeras showed very low stability compared to DNA:DNA duplexes.

Bergman *et al.*[60] have used 5'-aminonucleoside **29** (Figure 13) as the linker unit and assembled the chimera using Fmoc amino acid with appropriate protocols on CPG resin. They synthesized different chimeras with varying length of PNA and DNA stretches.[61] The introduction of PNA unit in the center and chimera containing 50% PNA caused destabilization. On the other hand, the introduction of PNA unit on either/both ends of DNA does not show any destabilizing effect. Such capping processes render the DNA oligomers stable to nucleases while keeping their binding properties intact.



28



29

Figure 13

Uhlmann *et al.*[64] have used a combination of both the linker units **28** and **29** mentioned earlier and synthesized PNA:DNA chimeras. They have demonstrated that these chimeras have preferential binding in antiparallel orientation compared to parallel orientation. The utility of these PNA as antisense agents is still not clear because of their poor cellular uptake, while PNA:DNA chimeras can cross the cell membrane with equal efficiency as an oligonucleotide and are resistant to 3'-exonucleases activity, due to the presence of PNA at the 3'end.

1.3.3b ***PNA-Peptide Chimeras*****:** Covalent linkage of the peptide to PNA can be achieved with great ease since both involve peptide bond formation. Such PNA-peptide chimeras can be designed for various purposes. PNA-peptide chimera containing basic amino acids like lysine and arginine in the peptide part show enhanced stability of the resulting PNA:DNA complexes. The extra stabilization is due to the interaction of the positive charge on the peptide with the negatively charged DNA phosphate backbone. The PNA and peptide segments act independently in such chimeras. The peptide segment chosen by Koch *et al.*[65] for the formation of PNA-peptide chimera is a substrate for protein kinase and independent enough to get phosphorylated by the protein kinase.

Another interesting aspect is covalent linking of PNA to nuclease enzyme. This was achieved through introduction of a cystine unit on the surface of S1-nuclease enzyme.[66] The PNA conjugation is achieved through disulfide exchange. Such conjugates can be site specifically targeted to cleave at a specific site of either the genomic stretch or of mRNA. The recognition is achieved by the PNA fragment and the cleavage by the nuclease enzyme. This may have major applications as genome cutter which is discussed later.

The introduction of many small molecule ligands like intercalator, at the PNA terminus also has shown enhancement in stability. Covalent linkage of artificial nucleases can be easily achieved and hence the PNA can be used for site-specific cleavage of DNA.[66] Fluorescent labeling can be achieved by linking fluorophore to PNA and such oligomers can be used to study the accumulation of PNA in cell media. These application oriented aspects are discussed in the next section.

1.4 Potential Biological Applications of PNA

The primary goal of design and synthesis of PNA is for their potential biological applications, with the advantage being stability and efficient antigene action by binding to dsDNA. But the actual oligomer showed more promising properties by strand displacement binding to the dsDNA.[16] Since the PNA:DNA binding affinity is very high, a lower concentration of PNA is required to get the desired effect. They bind to supercoiled DNA with a higher affinity resulting in uncoiling.[67] The exceptional high stability of PNA_2-DNA triplexes enable strand displacement to take place upon targeting of double stranded DNA. The homopyrimidine PNA oligomers displace the pyrimidine strand of complementary dsDNA targets upon formation of PNA_2-DNA triplex with homopurine strand which is shown to be salt dependent (<50mM) and has slight preference for parallel binding. Similar strand displacement complexes are reported to be formed by mixed sequence of PNA with dsDNA resulting in to the PNA-DNA duplex and displacement of other strand of target DNA.[37]

1.4.1 Stability towards cellular enzymes

Any analogue of DNA to be an efficient antisense or antigene drug should be stable towards cellular enzymes. PNA has non-natural amide linkages which are not recognized by the proteases present in the cytoplasm and hence they are not degraded.[68] Their stability was checked in human blood serum, *Escherichia coli*, *Micrococcus luteus* and cytoplasmic extracts from mouse Ehrlich ascites tumor cells. When compared to a normal control peptide which was degraded completely, no significant degradation of PNA was observed. Thus PNA has sufficient biostability to be used as a drug. PNAs are not attacked by exonucleases and the introduction of a PNA unit at the 3'end of oligomer acts as end capping and provides resistance to exonucleoase activity.[61]

1.4.2 Trascription arrest by PNA

The binding of PNA to dsDNA is highly sequence specific and once bound, they form a strand displacement complex. It was found that PNA-T_{10}-lys-NH_2 bound downstream from the promoter region caused transcription elongation arrest at the site

only where the PNA was bound to the template strand in presence of RNA polymerase T_3. There was no effect when the PNA was bound to the non-template strand. This selectivity indicated that independent targeting of mRNA with antisense sequence or dsDNA with sense sequence is possible with PNA *in vivo*.[69] It was also shown that the primer extension of *Taq* polymerase is arrested by PNA-T_{10} binding on the template DNA. The transcription arrest was significantly low in case of PNA containing one mismatch and virtually no transcription arrest occurred with PNA containing two mismatches. The best effect was obtained by using PNA-T_{10} and the arrest was less efficient upon the introduction of C monomers as in PNA-T_5CT_4 or PNA- $T_2CT_2CT_4$.[70] This was attributed to a weak binding of C monomer in the triplex form at physiological pH due to its non-protonation. PNA-T_8 and T_6 show lower efficiency of transcription arrest since the binding is weak. From the arrested transcripts it was observed that the polymerase transcribed two or three nucleotides into the PNA binding site. These results show that efficient transcription elongation arrest can be achieved by PNA targeting of the template DNA strand. In most of these experiments the PNA:DNA strand displacement complex were formed in low salt concentration and then transferred to the transcription buffer.

Similar transcription elongation arrest with eukaryote RNA polymerase enzyme was also observed. Again stable complexes of PNA with RNA-A_{10} sequence (PNA_2:RNA) was shown to cause inhibition of reverse transcriptase (RNaseH, MMLV).[70] This type of inhibition is not exhibited by DNA forming triplex with RNA. While the binding of PNA to dsDNA is highly salt dependent, its hybridization with ssDNA and RNA is not significantly affected by salt concentration.

1.4.3 PNAs as transcription promoters

The strand displacement binding of PNA has unlimited scope. The D-loop formed by the binding of PNA to dsDNA are very much like transcription initiation complexes. Transcription initiation complexes are formed by helicases by unwinding of the helix with polymerase and transcription factors. The D-loop formed by PNA-T_{10} binding is around 12 base pairs which is similar to the natural complexes (Figure 14). It is shown that RNA polymerase of *E. coli* recognizes the D-loop as the template and transcribes the DNA

strand.[71] The efficiency of the transcription is increased by having more than one PNA binding to the DNA strand adjacent to each other. These results also imply that the PNA can be used as a general sequence specific gene activator (i.e., as synthetic transcription factor). It is not certain under what conditions the translation arrest or transcription is preferred when the PNA binds to dsDNA. This can be successfully used by manipulating the length of sequence, where longer PNA sequences are required for trancription promoters.

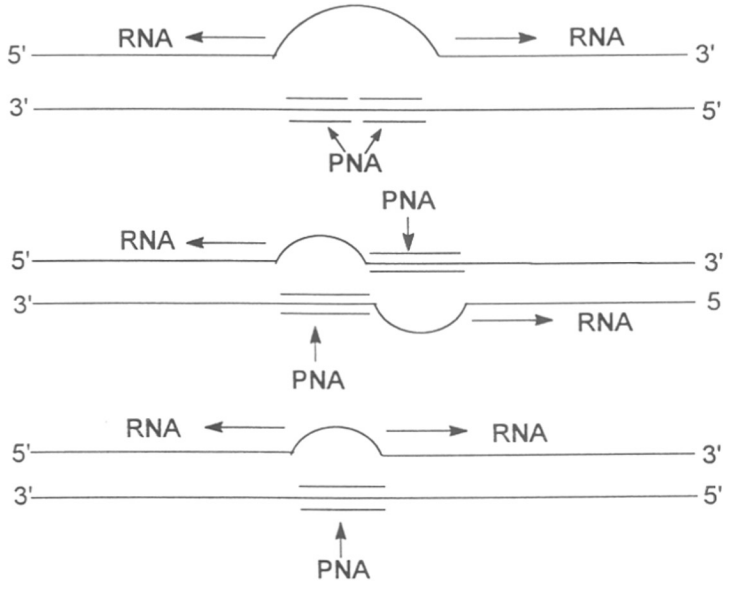


Figure 14: Schematic representation of the D-loop which acts as the transcription promoter.

1.4.4 Translation arrest by PNA

One of the most potential application of PNA as an antisense agent is the ability of PNA to inhibit the *in vitro* translation.[72] Incubation of PNA with the mRNA containing complementary sequence showed the concentration dependent termination of translation elongation products while the random PNA control sequence with the same base composition of target PNA did not cause any significant translation arrest. The antisense effect caused by PNA was attributed to the steric block of PNA:RNA complex, as such

complexes are not the substrates of RNase H. In contrast, the antisense effect caused by oligonucleotides is due to RNase activation.

The major hurdle in use of PNA in antisense/antigene approach is its extremely poor cell permeability. When PNAs (10-20 mer) were encapsulated in liposomes, they were found to have efflux rates of 5 to 11 days .[73] These results suggested that passive diffusion of unmodified PNA over the lipid membrane is not likely to be an effective way of transport in biological cells. The cell culture studies of PNA mediated inhibition of expression of T-Ag gene of SV40 were done by employing microinjection of PNA followed by detection of protein by indirect immunochemistry. Whether the observed inhibition is caused by an antisense or an antigene effect is not clear but the high ionic strength in cells most likely prevents binding of PNA to ds DNA and therefore argues for binding to RNA.

1.4.5 PNA directed PCR clamping

The higher stability and sequence specificity of PNA-DNA complexes compared with DNA-DNA complexes has been exploited in expanding the array of PCR based applications.[74] PNA can be directed at any site of template target DNA for different purpose. Competition between the primer and PNA for binding site can inhibit or even abolish PCR amplification. The careful design of PNA and primer sequences along with optimization of PCR conditions are shown to identify single base mutations.[74] The targeting of PNA sequence adjacent or away to primer site blocks the extension by DNA polymerase leading to inhibition of PCR amplification. The most efficient and discriminative clamping was observed with homopyrimidine sequences due to the extraordinary stable PNA_2-DNA triplex formation causing inhibition of polymerase at any PNA binding site.[75]

The extension of PNA primers by polymerase was studied recently by Lutz *et. al.*[76] The DNA polymerases from phage T4 and T7, *Thermus aquaticus, Thermus flavus*, *Thermus ubiquitos*, *Pyrococcus furiosus* and *DeepVent* polymerase as well as HIV reverse transcriptase do not recognize the PNA primers and no primer extension products were observed. However, DNA polymerase I (*E. coli*) and Vent DNA polymerase

(*Thermococcus litoralis*) did produce primer extended products using PNA (Figure 15) as the primer. At the same time PNA:DNA chimeras containing 50% PNA were not recognized by the enzymes as they were blocked at such hybrid structures. The recognition of PNA by polymerase is 100 times lower compared to its recognition of DNA primers, nevertheless primer extensions do take place.

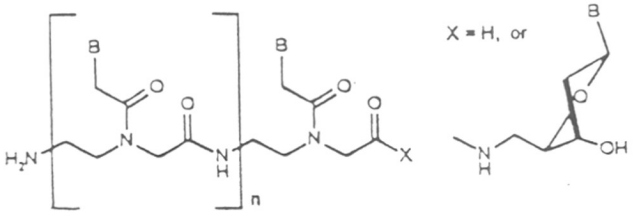


Figure 15: Primers used in PCR reaction recognised by polymerase.

1.4.6 Isolation of specific active genes by PNA strand invasion

The use of biotinylated PNA has been employed in the isolation of specific active genes by strand invasion. The chromatin fragments of interest containing CAG triplets[77] were challenged with complementary PNA at low ionic strength and PNA-DNA hybrids were isolated by density gradient centrifugation followed by capture on streptavidin coated magnetic particles. The precise mechanism of this molecular binding is unclear as under the conditions employed PNA is not expected to engage in stable Watson-Crick-Hoogsteen type triplex formation, and no control experiments were performed.

1.4.7 PNA as an artificial restriction enzyme

The strand displacement complex of PNA provides an opportunity to use these complexes and cut the DNA at a desired position. PNA in combination with nuclease S1 have been used as artificial restriction enzymes.[78] Once the strand displacement complex is formed, the displaced strand of DNA is susceptible to nuclease activity and is easily

digested by S1 nuclease into well defined products.[79] Depending on the strand to be cleaved, the PNA can be selectively directed to the other strand. Tethering of artificial nuclease can also give similar effect as demonstrated by attachment of gly-gly-his to PNA. In the presence of Ni, this PNA with gly-gly-his, binds to complementary DNA and cleaves the DNA duplex part close to its position.

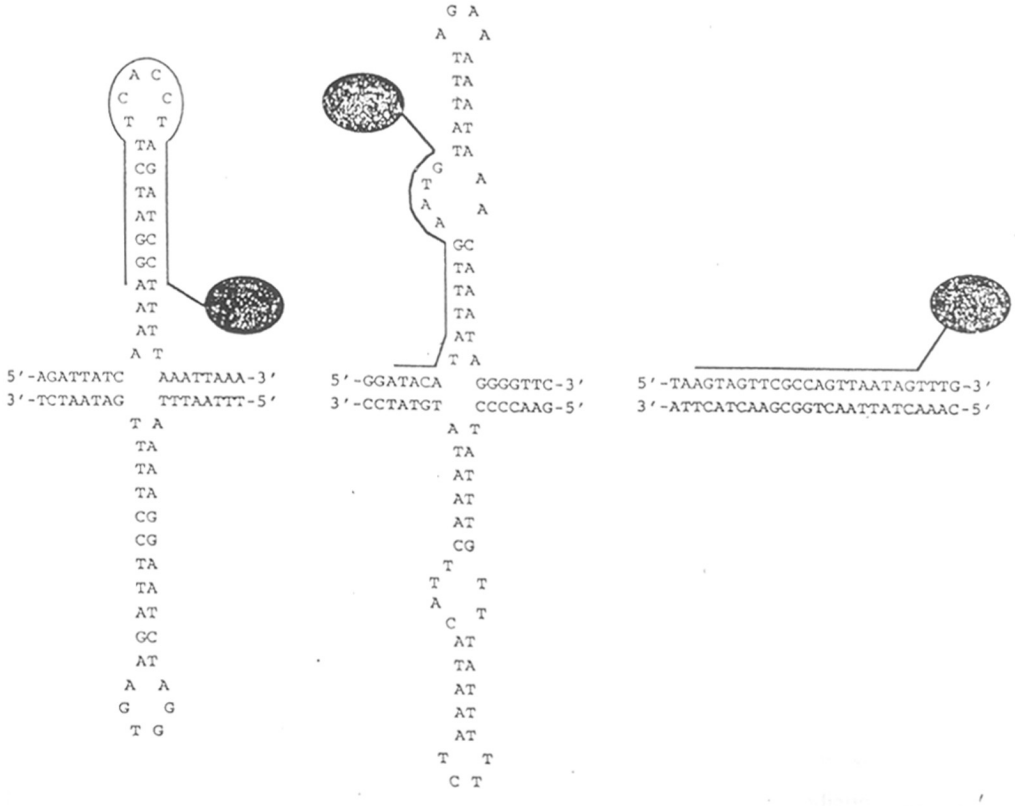


Figure 16: A nuclease attached to PNA as artificial restriction enzyme.

1.4.8 PNA as inhibitor of human telomerase

The telomerase enzyme responsible for the synthesis of G-rich sequences at the end of chromosomes, known as telomeres, has RNA component as a template.[80] The high

expression of telomerase in tumor cells has aroused interest for its use as a target in tumor inhibition. The *in vitro* studies by Corey *et al.*[81] using 'telomeric repeat amplification protocol (TRAP)' has shown that PNA can inhibit the telomerase activity by binding to RNA component of enzyme in picomolar to nanomolar range. They were found to be 10-50 times more efficient in inhibition compared to analogous phosphorothioate (PS) oligomers. In contrast to high selectivity of inhibition by PNAs, PS oligomers inhibit telomerase in non-sequence selective fashion. The results demonstrated that PNAs can control enzymatic activity of ribonucleoproteins and posses important advantages relative to PS oligomers in both the affinity and the specificity of recognition.

1.4.9 PNA as primordial genetic material

There is a constant speculation about the evolutionary origin of life. The prevailing theory is that among the biomolecules, DNA/RNA/Proteins RNA preceded others.[83] This concept is accepted because RNA has catalytic activity and bridges the gap between DNA and proteins. RNA being highly fragile, it is unlikely that prebiotic life could have relied on RNA. This has led to the proposition that other genetic systems which can be easily formed may have preceded the RNA world. It is shown that the PNA can be synthesized by using RNA as a template strand and *vice versa*.[82] This shows that a transition between genetic systems can occur without loss of information. The recent report that the PNA can be recognized by some polymerases and reverse-transcriptases also build the hope of Nielsen's speculation that PNA as a primordial genetic material could be proved right in near future.

1.5 DNA T-T photodimers

The deleterious effects of ultraviolet (UV) radiation on the genetic material of living cells are well known.[83] The most susceptible sites in DNA to UV radiation, are the adjacently stacked pyrimidine bases.[84] Any mutation in DNA due to photodamage, if not repaired correctly, would lead to cell death or cancer. The skin cancer is known to be caused by damage to DNA by sunlight.[85] Patients suffering from the genetic disorder, *Xeroderma pigmentosum*, characterized by a defective repair mechanism of DNA photodamage are reported to be most sensitive to sunlight.[83] Presently, the research in the

field of photochemistry of nucleic acid has acquired special status due to a fast depleting ozone layer and the consequent mutagenic and carcinogenic effects of solar UV radiation increasingly felt by the biosphere. The most significant photoproducts of DNA, both in terms of easy formation and adverse effects, are the cyclobutane dimers of thymine.[86] The possible cyclobutane photodimers from DNA based on their stereochemistry can be classified into *cis-syn* and *trans-syn* isomers.[87] The former is an irradiation product of either duplex or single stranded DNA (ssDNA) whereas the latter is formed only from ssDNA or at the junctions of DNA chains with different conformations.[84] Organisms have developed various repair mechanisms to combat DNA damage.[88,89] In a most ingenious method, the energy of the damaging radiation, in near UV region, is being utilized by DNA photolyase enzyme system to monomerize the pyrimidine photodimers.[90] Photolyases were often found to have comparatively higher or even exclusive specificity towards *cis-syn* rather than *trans-syn* thymine dimers.[89] As a result, studies have been oriented heavily on understanding of the mechanism of comparatively easily available *cis-syn* thymine dimers of natural dinucleotides or model compounds. Such studies are not many for *trans-syn* thymine photodimers. Recently, it has been demonstrated that *E. coli* DNA photolyase does repair *trans-syn* thymine dimers, but with 10^4 fold less efficiently compared to repair of *cis-syn* thymine dimers.[91] Such a subdued mechanism present in the cell for the repair of *trans-syn* dimers deem them extremely lethal to life.

1.6 Present Work

The above resume outlines the properties and utility of PNA, the sugar-backbone replaced DNA mimic. The properties of PNA have demonstrated that a sugar-phosphate backbone is not a prerequisite for obtaining helical duplexes governed by Watson-Crick base pairing. The high stability of both PNA-DNA, PNA-RNA and especially PNA_2-DNA complexes have already made PNA an useful tool in diagnostics, molecular biology and related fields. It can act as a versatile handle for various ligand and site-specific DNA targeting.

The poor cell permeability of PNA and bioavailability have been the major hurdle for the prospects of PNA as a drug. They are known to bind in both antiparallel and

parallel mode, with antiparallel mode being more stable. Above all PNAs have low solubility and strand displacements occur at low concentrations (50 mM). To overcome these drawbacks, a study of properties of PNA analogues with modified backbone is necessary and would greatly contribute to unravel the structure-activity relations of PNA with DNA.

Chapter 2: In view of the above discussion, this chapter reports on synthesis of chiral building blocks for PNA oligomers envisaged to impart control of rigidity into PNA oligomer are discussed. The introduction of chirality was carried out by binding the β-C of ethylenediamine unit of PNA with α-C of glycine unit of the same monomer to obtain a 5-membered heterocyclic ring with two chiral centers. All the four diastereomeric chiral monomers were synthesized from easily available 4-*trans*-hydroxy-L-proline with nucleobase attached at the N1 position through a acetyl spacer. This chapter also describes an alternative methodology where a common precursor is used to alkylate any nucleobase to obtain PrNA monomers. These chiral monomeric units were introduced into PNA to modulate the binding orientation.

Chapter 3: This chapter describes the synthesis of various PrNA oligomers by solid phase synthesis. The effect of PNA containing a modified chiral unit at the N-terminus and in the middle of the sequences are studied. Spermine is a naturally occurring polyamine with four cationic site and hence spermine analogues of PNA were prepared to increase the solubility and studies directed towards their biophysical properties are discussed. The fluorescent base 2-aminopurine was introduced into PNA oligomer and used to study hybridization of the PNA:DNA duplexes. The kinetics and dynamics of duplex formation were studied by the change in fluorescence intensity.

Chapter 4: Oligoethylenoxy glycols are used to link two thymines and the effect of the linkers towards photodimerization are reported in this chapter. The cyclobutane photoproduct of the bisthymines are characterized by single crystal X-ray structure analysis. The comparison of the structures of the two photodimers one with triethylenoxy chain and other having a catechol unit in the spacer chain along with their photodimers is discussed to assess the structural effects and stereochemical outcome on photodimerization of thymines.

1.7 References

1. Watson, J. D.; Crick, F. H. *Nature* **1953**, *171*, 737.
2. Zamechnik, P. C.; Stephensen, M. L. *Proc. Natl. Acad. Sci. USA*. **1978**, *75*, 280.
3. Sanghvi, Y. S. *Antisense Research and Application*: Eds. Crooke, S.T. and Lebleu, B. 1993, CRC, Boca Raton.
4. Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.
5. Moser, H. E.; Dervan, P. B. *Science* **1987**, *288*, 645.
6. Thoung, N. T.; Helene, C. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666.
7. Helene, C. *Anti-cancer. Drugs. Res.* **1991**, *6*, 569.
8. Nielsen, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167.
9. Milligan, J.; Matteucci, M. D.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 1923.
10. Matteucci, M. D.; Wagner, R. W. *Nature*, **1996**, *384* (supp.), 2103.
11. Glazer, V. *Genet. Eng. News* **1996**, *16*, 1,16,17 and 21.
12. Eschenmoser, A.; Loewenthal, E. *Chem. Soc. Rev.* **1992**, *21*, 1.
13. Stirchak, E. P.; Summerton, J. E.; Weller, D. D. *Nucleic Acids Res.* **1989**, *17*, 6129.
14. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
15. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5.
16. Almarsson, O.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA.*, **1993**, *90*, 9542.
17. Chen, S. M.; Mohan, V.; Kiely, S. J.; Griffith, M. C.; Griffith, R. H. *Tetrahedron Lett.* **1994**, *35*, 5105.
18. Egholm, M.; Buchardt, O.; Nielsen, P. E; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895.
19. Egholm, M.; Nielsen, P. E; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677.
20. Lesnik, E. A.; Risen, M. R.; Driver, D. A.; Griffith, M. C.; Sprankle, K.; Freier S. M. *Nucleic Acids Res.* **1997**, 25, 568.
21. Kim, S. H.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

22. Tomac, S. Sarkar, M.; Wittung, P.; Nielsen, P. E.; Norden, B.; Grashlund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544.

23. Engels, W. J.; Meier, C. *Angew. Chem. Int. Ed. Engl.* **1992,** *31*, 1008.

24. Wittung , P.; Kim, S. K.; Buchardt, O.; Nielsen, P. E; Norden, B. *Nucleic Acids Res.* **1994**, 22, 5371.

25. Griffith, M. C.; Risen, L. M.; Greig, M. J.; Lesnik, E.A.; Sprankle, K. G.; Griffey, R. H.; Kiely, J. S.; Freier, S. M. *J. Am. Chem. Soc.* **1995**, *117*, 83.

26. Ts'O, P. O. P.; Ono, A.; Kan L. J. *J. Am. Chem. Soc.* **1991**, *113*, 4032.

27. Egholm, M.; Christensen, L.; Dueholm, L. K.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, *23*, 217.

28. Betts, L.; Josey, J. A.; Veal, J. M. Jordan, S. R. *Science*, **1995**, *270*, 1838.

29. Wittung, P.; Nielsen, P. E; Norden, B. *Biochemistry,* **1997**, *36*, 7973.

30. Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc. Chem. Commun.* **1993**, 800.

31. Egholm, M. *Nature*, **1993**, *365*, 2103.

32. Leijon, M.; Graslund, A.; Nielsen, P. E.; Buchardt, O.; Norden B.; Kristensen, S. M.; Eriksson, M. *Biochemistry,* **1994**, *33*, 9820.

33. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science*, **1994**, *265*, 777.

34. Wittung, P.; Nielsen, P.E.; Norden B. *J. Am. Chem. Soc.* **1996**, *118*, 7049.

35. Wittung, P. Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B. *Nature*, **1994**, *368*, 561.

36. Wittung, P. Eriksson, P.; Lyng R.; Nielsen P.E.; Norden, B. *J. Am. Chem. Soc.* **1995**, *117*, 1167.

37. Wittung, P.; Nielsen, P. E.; Norden, B. *J. Am. Chem. Soc.* **1997**, *119*, 3189.

38. Cherny, D. Y.; Beltserkovskii, B. P.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Berg, R. H.; Nielsen, P. E. *Proc. Natl. Acad. Sci. USA*. **1993**, *90*, 1667.

39. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioconjugate Chem.* **1994**, *5*, 3.

40. Hyrup, B. Egholm, M.; Rolland, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993**, 518.

41. Hyrup, B.; Egholm, M.; Nielsen, P. E.; *J. Am. Chem. Soc.* **1994**, *116*, 7964.

42. Lagriffoul, P.-H.; Egholm, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1088.

43. Krotz, A. H.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *Tetrahedron Lett.* **1994**, *35*, 6937 and 6941.

44. Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett.* **1994**, *35*, 5173.

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

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

47. Doel, M. T.; Jones, A. S.; Walker, R. T.; *Tetrahedron.* **1974**, *30*, 2755.

48. Garner, P.; Yoo, J. U. *Tetrahedron Lett.* **1993**, *34*, 1275.

49. Lenzi, A.; Reginato, G.; Taddei, M. *Tetrahedron Lett.* **1995**, *36*, 1713, 1717.

50. Shah, V. J.; Kuntz, I. D.; Kenyon, G. L. *Bioorg. Chem.* **1996**, *24*, 194.

51. Diederichsen, U. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 445.

52. Diederichsen, U.; Schmitt, H. W. *Tetrahedron Lett.* **1996**, *37*, 475.

53. Huang, S.-B.; Nelson, J. S.; Weller, D. D. *J. Org. Chem.* **1991**, *56*, 6007.

54. Savitri, D.; Leumann, C.; Scheffold, R.; *Helv. Chem. Acta* **1996**, *79*, 288.

55. Lioy, E.; Kesser, H. *Liebigs Ann.* **1996**, 201.

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

57. Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. I* **1997**, 539, 547 and 555.

58. Jordon, S.; Schwemler, C.; Kosch, W.; Kretschmer, A.; Schwenner, E.; Stropp U.; Mielke, B. *Bioorg. Med. Chem. lett.* **1997**, *7*, 681 and 687.

59. van der Laan, A. C.; Stromberg, R; van Boom, J. H. Kuyl-Yeheskiely, E.; Efimov V. A.; Chakhmakhcheva, O. G. *Tetrahedron Lett.* **1996**, *37*, 7857.

60. Bergmann, F.; Bannwarth, W.; Tam, S. *Tetrahedron Lett.* **1995**, *36*, 6823.

61. Stetsenko, D. A.; Lubyako, E. N.; Potapov, V. K.; Azhikina T. L.; Sverdlov, E. D. *Tetrahedron Lett.* **1996**, *37*, 3571.

62. van der Laan, A. C.; Brill, R.; Kuimelis, R. G.; Kuyl-Yeheskiely, E.; van Boom, J. H.; Andrus, A.; Vinayak, R. *Tetrahedron Lett.* **1997**, *38*, 2249.

63. Petersen, K. H.; Jensen, K. D.; Buchardt, O.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1995**, *6*, 1119.

64. Uhlmann, E.; Will, D. W.; Breipohl, G.; Langner, D.; Ryte, A. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939.

65. Koch, T.; Naesby, M.; Wittung, P.; Jorgensen, M.; Larsson, C.; Buchardt, O.; Stanley, C. J.; Norden, B.; Nielsen, P. E.; Orum, H. *Tetrahedron Lett.* **1995**, *36*, 6933.

66. Norton, J. C.; Waggenspack, J. H.; Varnum, E.; Corey, D. R. *Bioorg. Med. Chem.* **1995**, *3*, 437.

67. Smulevitch, S. V.; Simmons, C. G.; Norten, J. C.; Wise, T. W.; Corey, D. R. *Nature Biotech.* **1996**, *14*, 1700.

68. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchard, O.; Nielsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1310.

69. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Anti cancer Drug Design* **1993**, *8*, 53.

70. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Gene*, **1994**, *149*, 139.

71. Mollegaard, N. E.; Buchardt, O.; Egholm, M.; Nielsen, P. E. *Proc. Natl. Acad. Sci. USA.*, **1994**, *91*, 3892.

72. Buchardt, O.; Egholm, M.; Berg, R. H.; Nielsen, P. E. *TIBTECH*, **1993**, *11*, 384.

73. Harvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481.

74. Orum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley C. *Nucleic Acids Res.* **1993**, *21*, 5332.

75. Demers, D. B.; Curry, T.; Egholm, M.; Sozer, A. C. *Nucleic Acids Res.* **1993**, *23*, 3050.

76. Lutz, M. J.; Benner, S. A.; Hein, S.; Breipohl, G.; Uhlmann, E. *J. Am. Chem. Soc.* **1997**, *119*, 3177.

77. Boffa, L. C.; Morris, P. L.; Carganeto, E. M.; Louissaint, M.; Allfrey, V. G. *J. Biol. Chem.* **1996**, *271*, 13228.

78. Demidov, V. V.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchard O. *Nucleic Acids Res.* **1995**, *21*, 2103.

79. Footer, M.; Egholm, M.; Kron, S.; Coull, J. M.; Matsudaira, P. *Biochemistry* **1996**, *35*, 10673.

80. Gambacorti-Passerini, C. *et al. Blood* **1996**, *88*, 1411.

81. Bohler, C.; Nielsen, P. E.; Orgel, L. E. *Nature* **1995**, *376*, 578.

82. Nielsen P.E. *Origins Life Evol. Biosphere* **1993**, *23*, 323.

83. Setlow, R. B. *Nature* **1978**, *271*, 713.

84. Taylor, J-S. *J. Chem. Edu.* **1990**, *67*, 835.

85. Taylor, J-S. *Acc.Chem. Res.* **1994**, *27*, 76.

86. Fisher, G. J.; Johns, H. E. *Photochemistry and Photobiology of Nucleic Acids* ed. S.Y.Wang, **1976**, Vol 1, 225-294.

87. Cadet, J.; Vigny, P. *Bioorganic Photochemistry*, ed Morrison H, **1990**, Vol*1*, 1-272. New York John Wiley and sons.

88. Sancar, A.; Tang, M-S. *Photochem. Photobiol.* **1993**, *57*, 905.

89. Sancar, A; Sancar, G.B. *Ann. Rev. Biochem.* **1988**, *57*, 29.

90. Kim, S-T.; Sancar, A. *Photochem. Photobiol.* **1993**, *57*, 895.

91. Kim, S-T., Malhotra, K.; Smith, C.A.; Taylor, J-S.; Sancar, A. *Biochemistry*, **1994** *32*, 7065.

CHAPTER 2

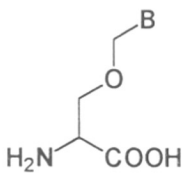
Synthesis of N^{α}-(Purinyl/Pyrimidinyl acetyl)-4-Amino Proline with Potential Use in Chiral PNA Synthesis

2.1 Introduction

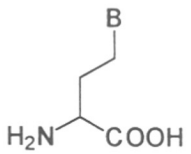
There is a constant quest to search for better drugs that are specific and have less side effects. In this aspect, Peptide Nucleic Acid (PNA) invented by Nielsen *et al.*,[1] have carved a place as one of the forerunners in antisense/antigene therapeutics. PNA has emerged as a novel nucleic acid mimic that binds DNA/RNA with high efficiency and maintains the sequence specificity. These properties, along with their stability towards nucleases and proteases and ease of synthesis in large scale has made PNAs very attractive candidates for antisense technology.[2,3] Recently there is an influx of reports on modification of PNA to overcome some of the drawbacks observed in PNA as already discussed in the previous chapter. Attempts aimed at structural modifications to improve PNA properties for modulating DNA/RNA recognition and facilitating biological applications have centered around modification of PNA backbone[4,5] and introduction of asymmetry into achiral PNA (Figure 1). Nielsen *et al.* have introduced a chiral amino acid,[6] D/L-lysine, at the end of the PNA **1a** and have shown that this induces chirality in the whole molecule to attain a helical structure, whose sense depends on the chirality of the amino acid. Koch *et al.*[7] have prepared PNA-peptide chimera, **1b** in which the peptide chain induces chirality into the molecule. Bergmann *et al.*[8] have synthesized DNA-PNA chimeras[9,10] **1c** on CPG resin. In all these cases, the asymmetry is achieved by terminal modification. Garner,[11] Taddie[12] and Diederichsen[13] have introduced the nucleobase through different spacers on α–C of glycine **2a-c**, and peptides were prepared using these monomers alternately with other amino acids like glycine. Dueholm *et al.*[14] and Kosynkina *et al.*[15] have introduced chirality by incorporation of various amino acids into the PNA backbone itself as shown in structure **3a,b**. Nielsen *et al.*[16] have also prepared a flexible, positively charged analogue by attaching the nucleobase to N(2-aminoethyl)glycine unit through an ethyl side chain as in **4**. While the present work was in progress,[17] Lowe *et al.*[18] reported introduction of the nucleobase at 4-position of 4-hydroxy-L-proline as in **5** and used this in combination with glycine to build PNA. Most of these modifications have shown interesting, but not substantial alterations, in the properties of the resulting peptides as compared to PNA.

H_2N–PNA–CONHR

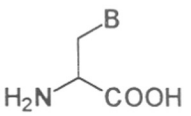
1a : R = amino acid (Ref. 4)
1b : R = Peptide (Ref. 5)
1c : R = Oligonucleotide (Ref. 6)



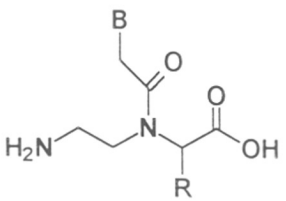
2a (Ref. 9)



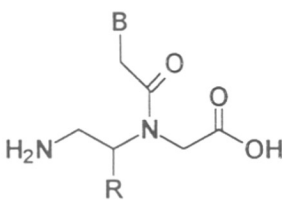
2b (Ref. 10)



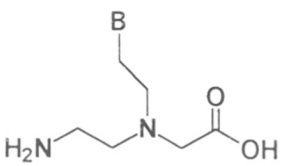
2c (Ref. 11)



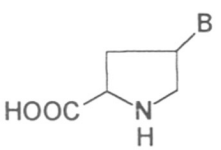
3a (Ref. 12)



3b (Ref. 13)



4 (Ref. 16)



5 (Ref. 17)

Figure 1

2.1.1 Rationale for present work

In the present work it was envisioned to incorporate chirality and controlled rigidity simultaneously into the PNA backbone to modulate the bio/physico-chemical properties. This, in principle can be done by bridging the α/β-C of ethylenediamine unit of PNA with the α"-C of glycine unit (**A** and **B,** Figure 2) or the α'/β'-C on sidechain to which nucleobase is connected (**C** and **D,** Figure 2). These types of structural

modifications introduce conformational constrains into the backbone and create asymmetric centers. While **A/B** approach creates two asymmetric centers in the backbone, the approaches **C/D** generate only one chiral center in the backbone and one in the side chain. A common feature of all these modifications from a practical point is that, they can all be derived from the easily available *trans*-4-hydroxy-L-proline **6**. For the sake of simplicity and straightforwardness, the approach **A** has been explored in this work.

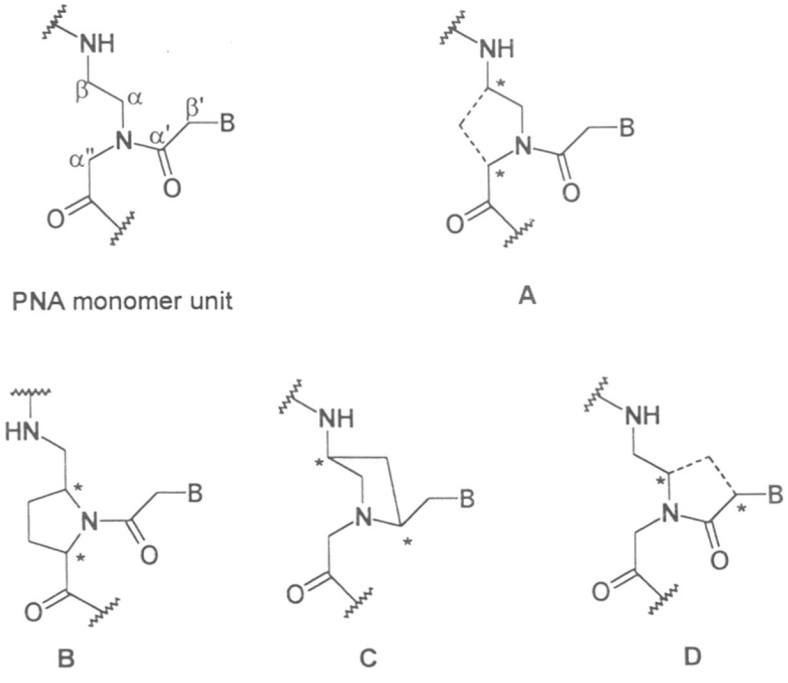


Figure 2

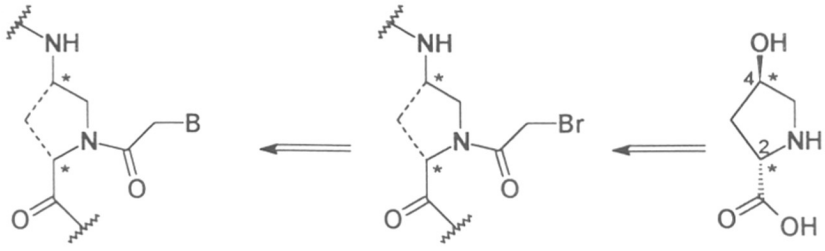
The objectives of this chapter are:

(i) Synthesis of N^{α}-(thymin-1-yl)-4-aminoproline diastereomers for use in PNA synthesis.

(ii) Alternative method for introduction of any heterocyclic (A,T,C,G) base through a common precursor.

(iii) Synthesis of dipeptides using prolyl monomers and

(iv) Synthesis of PNA monomers (A,T,C,G).



B = A,C,T,G,2-*ap* 6

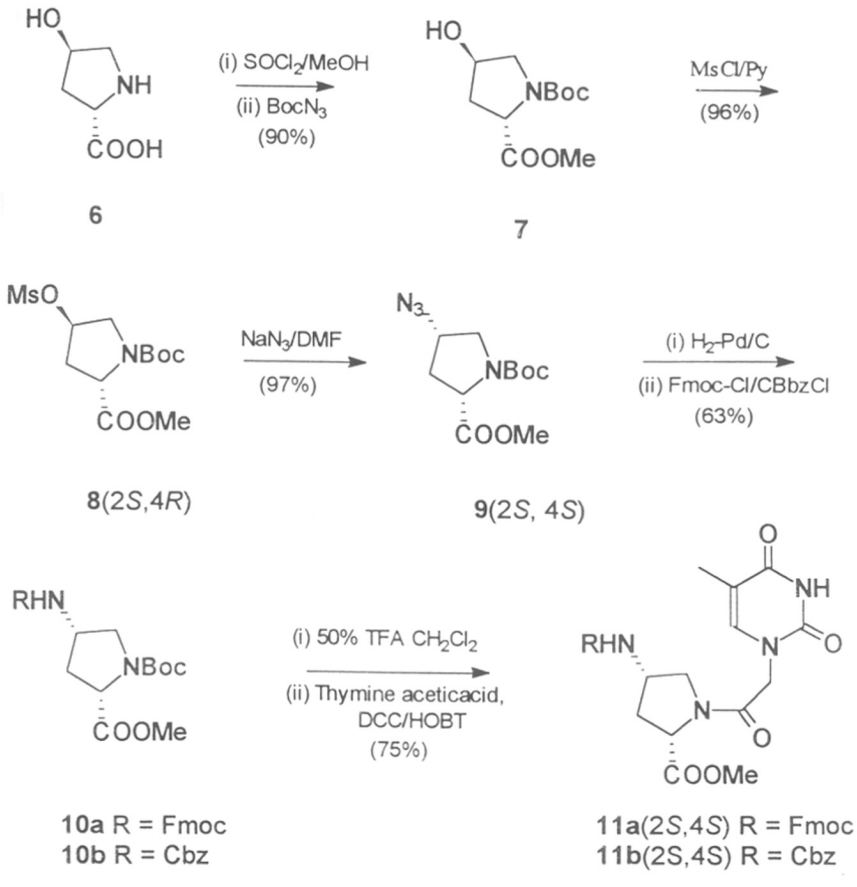
2.2 Present work

2.2.1 Synthesis of (2*R*/*S*,4*R*/*S*)-4-N(Fmoc)proline methyl ester (10)

trans-4-Hydroxy-L-proline is a naturally occurring amino acid of collagen. It has a five membered heterocyclic ring with two chiral centers at C2 and C4, apart from having three functional groups (N1, 2-COOH and 4-OH), which are amenable to easy modification. Further, the oxidation of 4-hydroxy to 4-keto enables funtionalisation of C3 and C5 which are α to the 4-keto moiety. The possibility of such derivatisation has made this compound versatile. This is reflected in several reports on synthesis of various chiral natural products and chiral molecules using this as starting material, well summarised in a recent review by Remuzon.[19] In the present work, the synthesis of all four chiral building blocks for prolyl PNA was achieved by sequential and systematic inversion at C-2 and C-4 centers leading to all 4 diastereomers of 4-hydoxyproline.

2.2.2 2*S*,4*S*-N(Fmoc)-N$^{\alpha}$(*t*-Boc)proline methyl ester

The synthesis of (2*S*,4*S*)-4-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(thymin-1-yl) proline methyl ester (11) was achieved in 6 steps (Scheme 1) starting from *trans*-4-hydroxy-L-proline (6). The compound **6** was converted to its methyl ester by dissolving it in refluxing methanol in the presence of $SOCl_2$. The resulting proline methyl ester was isolated as hydrochloride and then protected at the ring nitrogen using Boc-azide to obtain (2*S*,4*R*)-4-hydroxy N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (**7**). The compound **7** was converted to its 4-O-mesyl derivative **8** in very good yields by treatment with mesylchloride in pyridine.



Scheme 1

The appearance of a weak peak in IR at 2400cm^{-1} and a downfield shift of δ from 4.3 ppm to 5.2 ppm H4 in the ^{1}H NMR indicated the formation of mesyl derivative **8**. The 4*R* mesylate **8** was stirred with sodium azide in DMF to give the 4*S*-azido derivative **9** in almost quantitative yield (98%). This step is accompanied by an inversion of configuration at 4-position due to S_N2 displacement reaction. The azide **9**, shows a characteristic peak in IR at 2105 cm^{-1}, along with the disappearance of H4 peak at 5.2 ppm and appearance of a new peak at 4.2 ppm in ^{1}H NMR. The 2*S*,4*S*-azide **9** after hydrogenation gave the corresponding amine using 10% Pd/C as catalyst. The free amine thus obtained was protected using Fmoc-Cl by following the standard procedure to obtain N4 protected

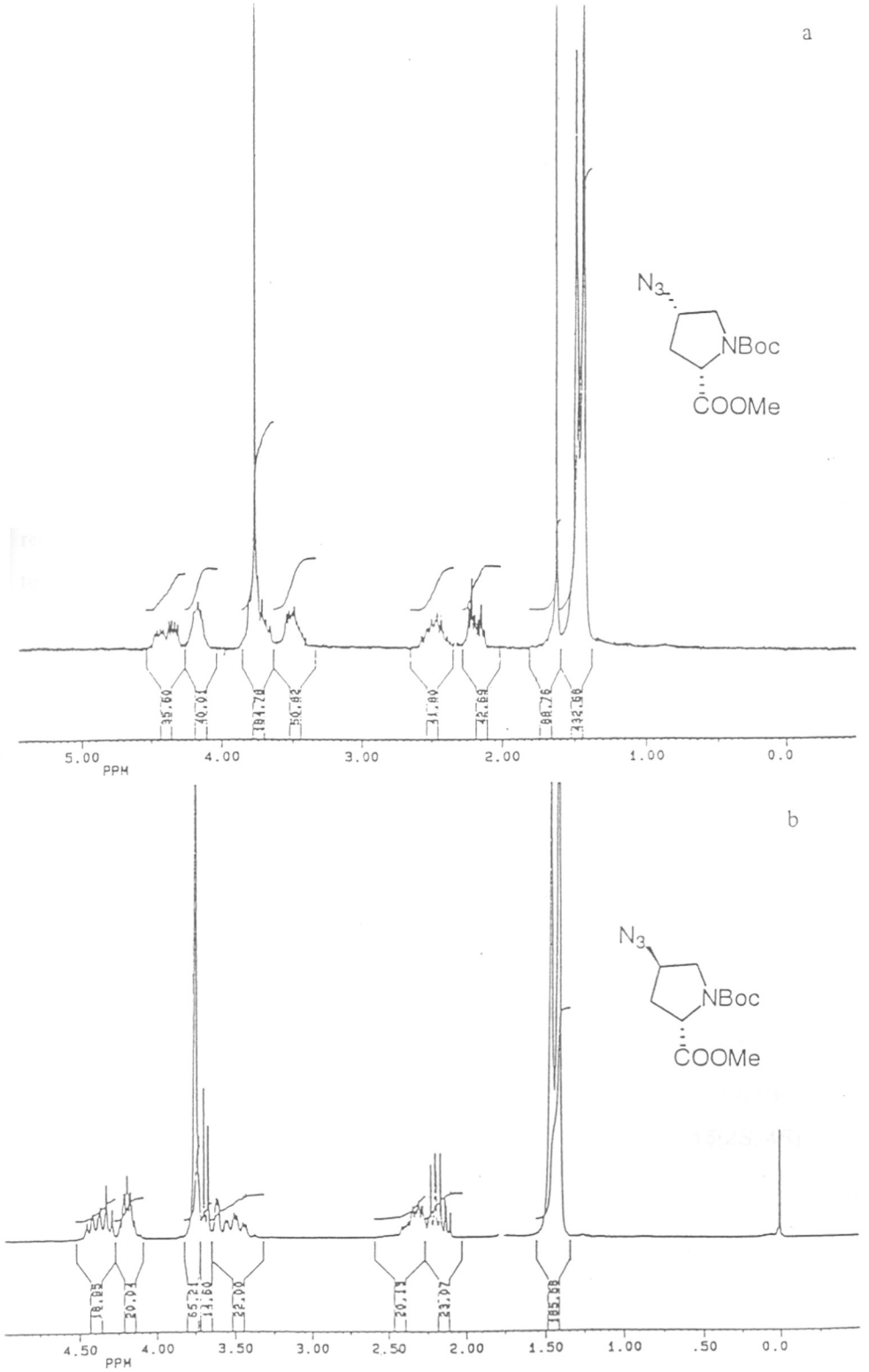
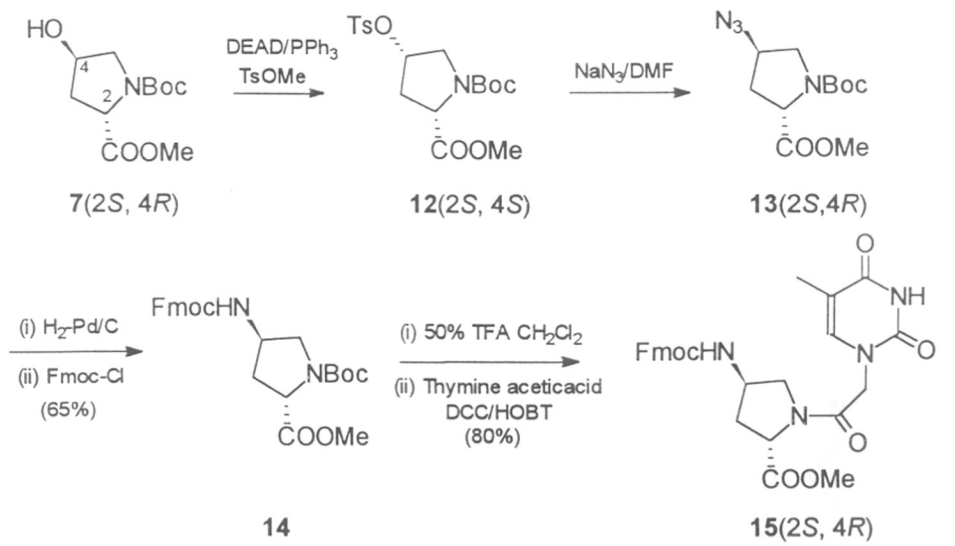


Figure 3: ^{1}H NMR spectra of compounds a: **9**(2*S*,4*S*) and b: **13**(2*S*,4*R*) in $CDCl_3$.

amino(N1-Boc)proline methyl ester **10** in moderate yield. This pathway with a single inversion step results in the overall transformation of 2*S*,4*R* proline (**6**) to 2*S*,4*S* proline derivative (**10**).

2.2.3 2*S*,4*R*-N(Fmoc)-N$^{\alpha}$(*t*-Boc)proline methyl ester (14)

The synthetic methodology to prepare **14** from *trans*-4-hydroxy-L-proline (**6**) consists of two successive inversion steps, giving rise to overall retention of configuration at C4 position of the proline ring (Scheme 2). The (2*S*,4*R*)-4-hydroxy-N$^{\alpha}$(t-butoxycarbonyl)proline methyl ester (**7**) was converted to 2*S*,4*S*-tosylate **12** under Mitsunobu conditions using DEAD, triphenylphosphine and methyl *p*-toluenesulfonate following the reported procedure.[23] This step goes through inversion of configuration at C4. The diethylcarbodihydrazide formed during the reaction could not be completely removed by column chromatography and hence the impure tosylate was directly converted to azide which could be purified by successive silica gel column chromatographies.

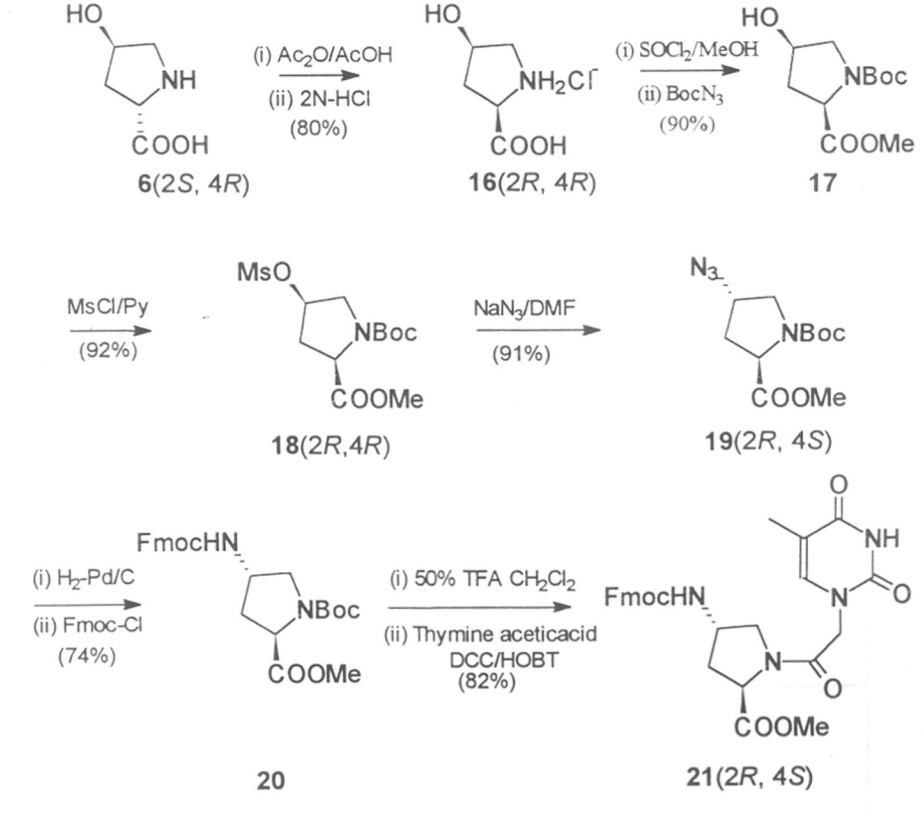


Scheme 2

The 2*S*,4*S* tosylate **12** was treated with sodium azide in DMF to obtain the 2S,4*R*-azide **13** that was purified and characterized. This step is again associated with S_N2 inversion of the 4*S*-tosylate to 4*R*-azide leading to an overall retention of the configuration at C4 from compound **7**. The hydrogenation of 2S,4*R*-azide **13** gave the free amine, which was treated with Fmoc-Cl as described earlier, to obtain the protected monomer **14** in moderate yield.

2.2.4 2*R*/4*S*-N(Fmoc)-N$^{\alpha}$(*t*-Boc)proline methyl ester (20)

The previous procedures gave compounds with both (*R*, *S*) stereochemistries at 4-position while that at C2 was same as the starting material i.e., 2*S*. The synthesis of compounds with 2*R* configuration could be achieved by initial epimerization at C2.



Scheme 3

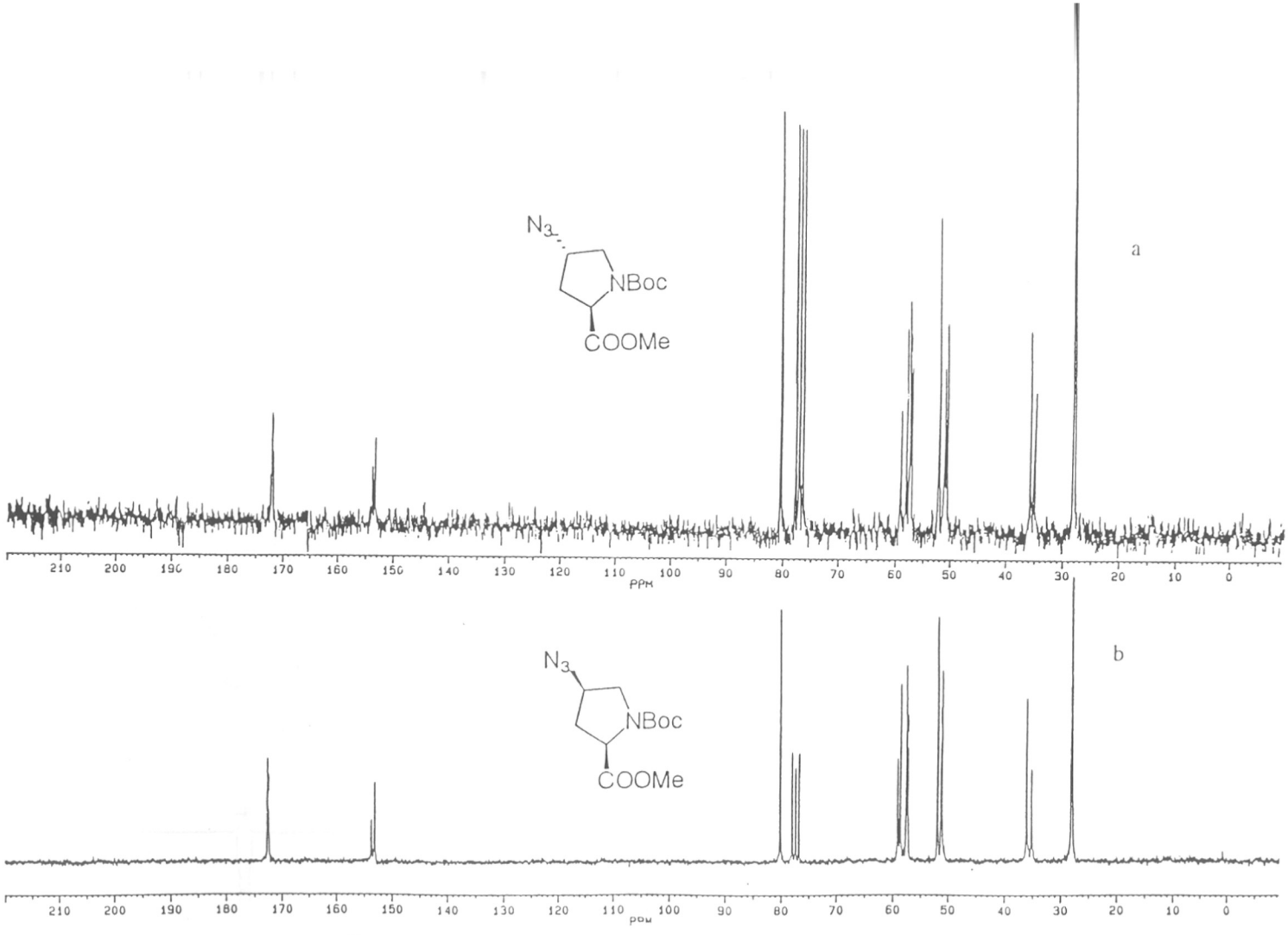


Figure 4: ^{13}C NMR spectra of compounds a: **19**(2*S*,4*R*) and b: **23**(2*R*,4*R*) in $CDCl_3$.

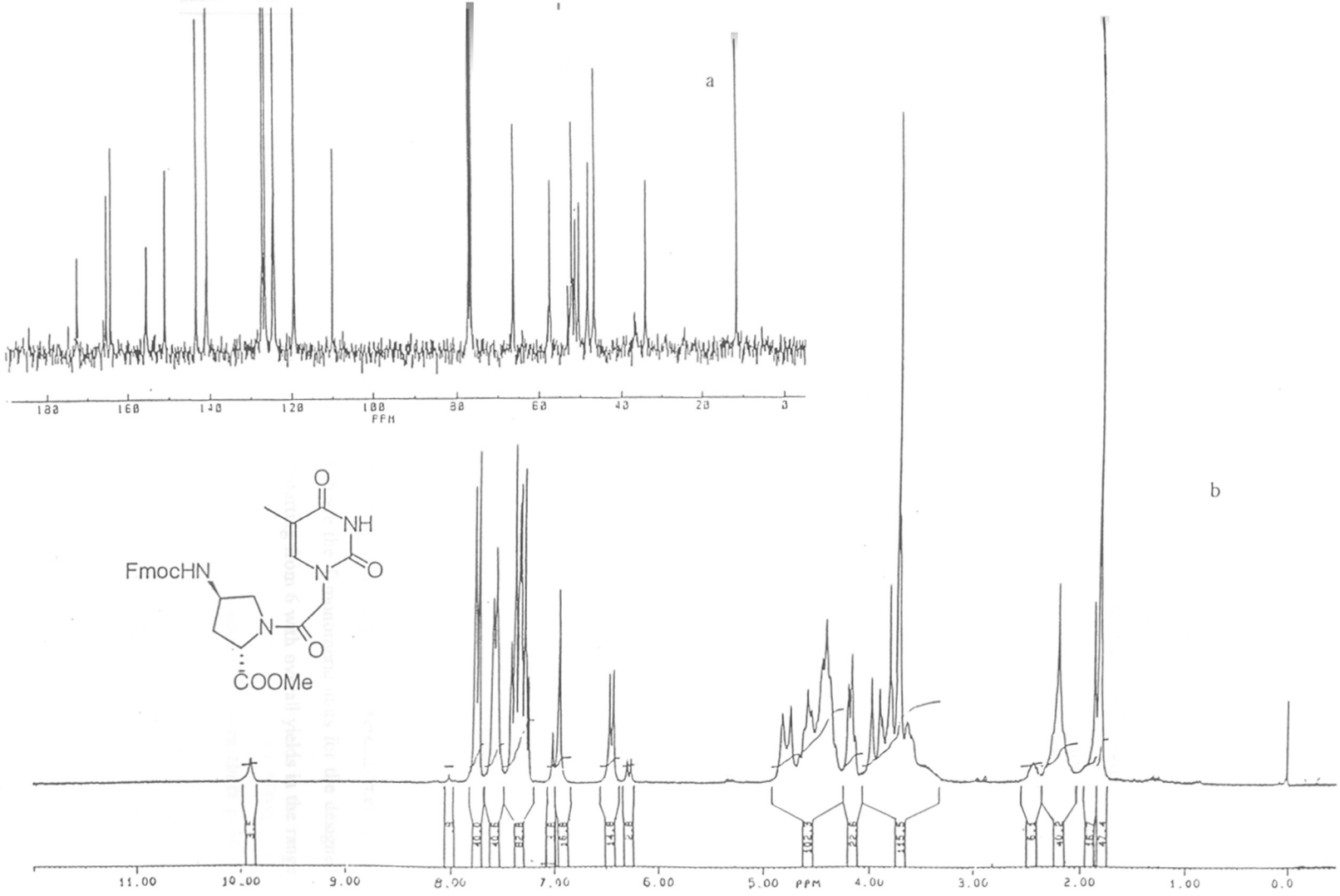


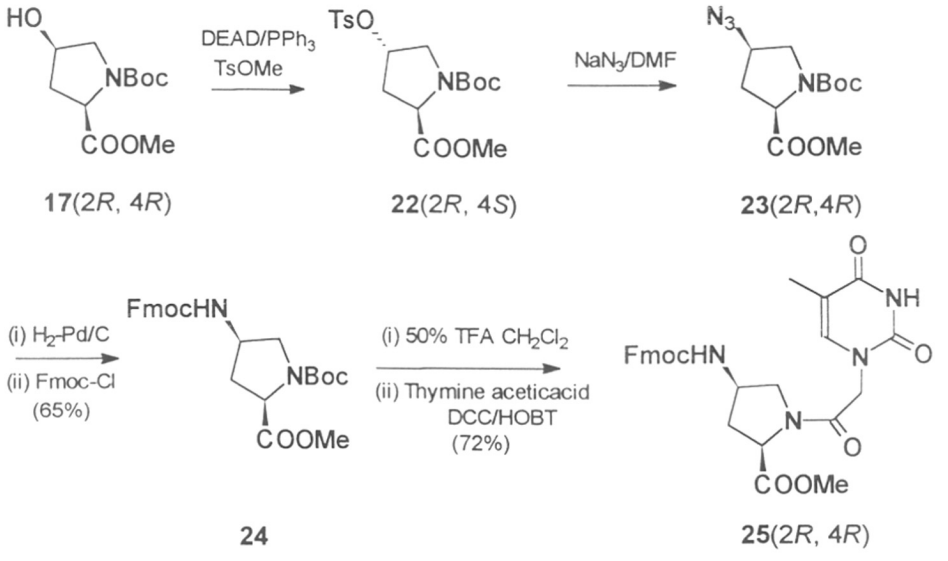
Figure 5: a:^{13}C NMR and b:^{1}H NMR spectra of compound **15**(2*S*,4*R*) in $CDCl_3$.

This was done by treating 2*S*,4*R*-hydroxyproline with acetic acid and acetic anhydride followed by hydrolysis of the resulting lactone with 2N HCl to give **16** in good yield (Scheme 3). A similar set of reactions, as earlier described for the L isomer (2*S*,4*R*, **6**) were used to obtain 4*S*,2*R*-derivatives **20**; (i) esterification of carboxyl group in **16**, (ii) N^{α}-protection with Boc-azide gave the Boc methyl ester **17**, (iii) conversion of **17** into 4*R*-mesyl derivative gave mesylate **18**, (iv) treatment of this mesylate **18** with sodium azide/DMF (displacement with inversion) to give 4*S*-azido derivative **19** and (v) reduction of **19** to 4*S*-amino compound and subsequent protection with FmocCl to obtain fully protected **20**.

2.2.5 2*R*/4*R*-N(Fmoc)-N^{α}(t-Boc)proline methyl ester (24)

The last of the epimer **24** (4*R*,2*R* derivative) was obtained from **17** by following a path as in Scheme 4 identical to Scheme 2, which involved two successive inversion steps; (i) tosylation of 4*R*-hydroxy **17** function under Mitsunobu conditions yielded 4*S*-tosyl derivative **22,** (ii) compound **22** in subsequent reaction with sodium azide underwent an inversion to yield the 4*R*-azido **23**, with an overall retention of configuration at C4, (iii) the azide **23** was transformed to the fully protected 4*R*,2*R* aminoproline **(24)**.

Thymine was treated with bromoethylacetate in presence of K_2CO_3 to obtain thymin-1-yl ethylacetate, which was hydrolyzed with 2N NaOH to obtain thymin-1-yl acetic acid. The N^{α} Boc group of **10** (scheme 1), **14** (scheme 2), **20** (scheme 3), **24** (scheme 4) were deprotected with 50% TFA/DCM to give the free secondary amine which was then coupled with thymin-1-yl acetic acid, following DCC-HOBT method,[22] well established in literature for peptide coupling, to obtain the desired compounds **11, 15, 21** and **25** respectively in good yield (85%). Thus all possible stereoisomers of 4-amino-N^{α}(thymin-1-yl acetyl) prolines which are the T-monomeric units for the designed 4-aminoprolyl nucleic acids were obtained starting from **6** with overall yields in the range of 28-30%.



17(2*R*, 4*R*) 22(2*R*, 4*S*) 23(2*R*,4*R*)

24 25(2*R*, 4*R*)

Scheme 4

2.3 Alternative method for monomer synthesis *via* a common precursor

The above strategy involving condensation of prolyl ring NH with N-carboxy-methylated nucleobase although successful for synthesis of T-monomers, presents problems for synthesis of monomers of other nucleobases, in particular, that of purines. In case of coupling of thymin-1-yl aceticacid, the yields are better with free amine of **10** and **14,** than with the corresponding TFA-salts and Fmoc protection reaction yields were also low. This could be due to relative stereochemical disposition of substituents at 2 and 4 which may sterically hinder the coupling of ring NH with thymine acetic acid. An alternative route was explored towards a general synthesis of prolyl nucleic acid monomers which consisted of direct coupling of individual nucleobases (T, C, 6-Cl-2-aminopurine and A) with suitably protected N^{α}-(bromoacetyl)proline. Such a strategy has been reported to be efficient in large scale synthesis of monomers of open chain peptide nucleic acids.[25]

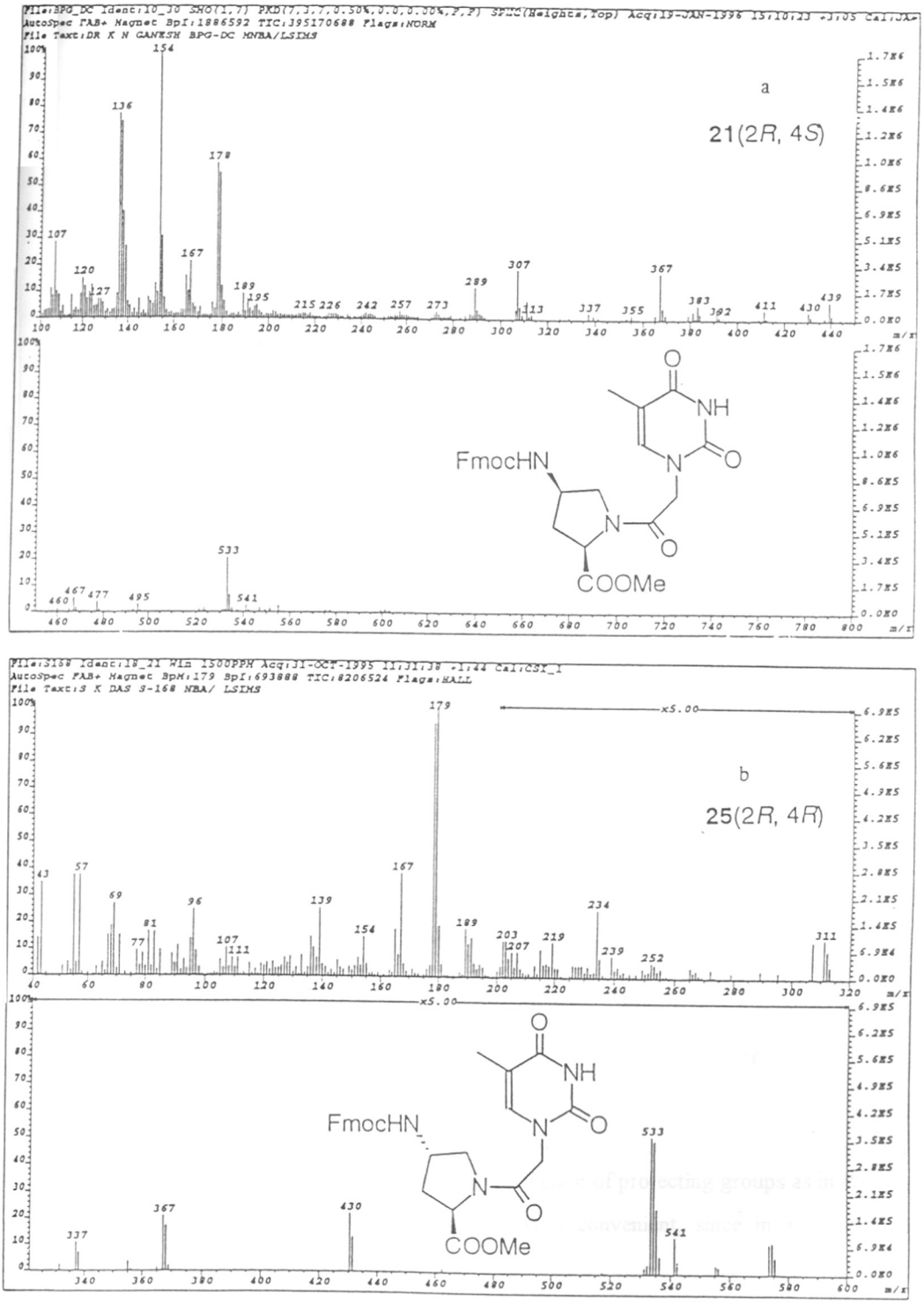
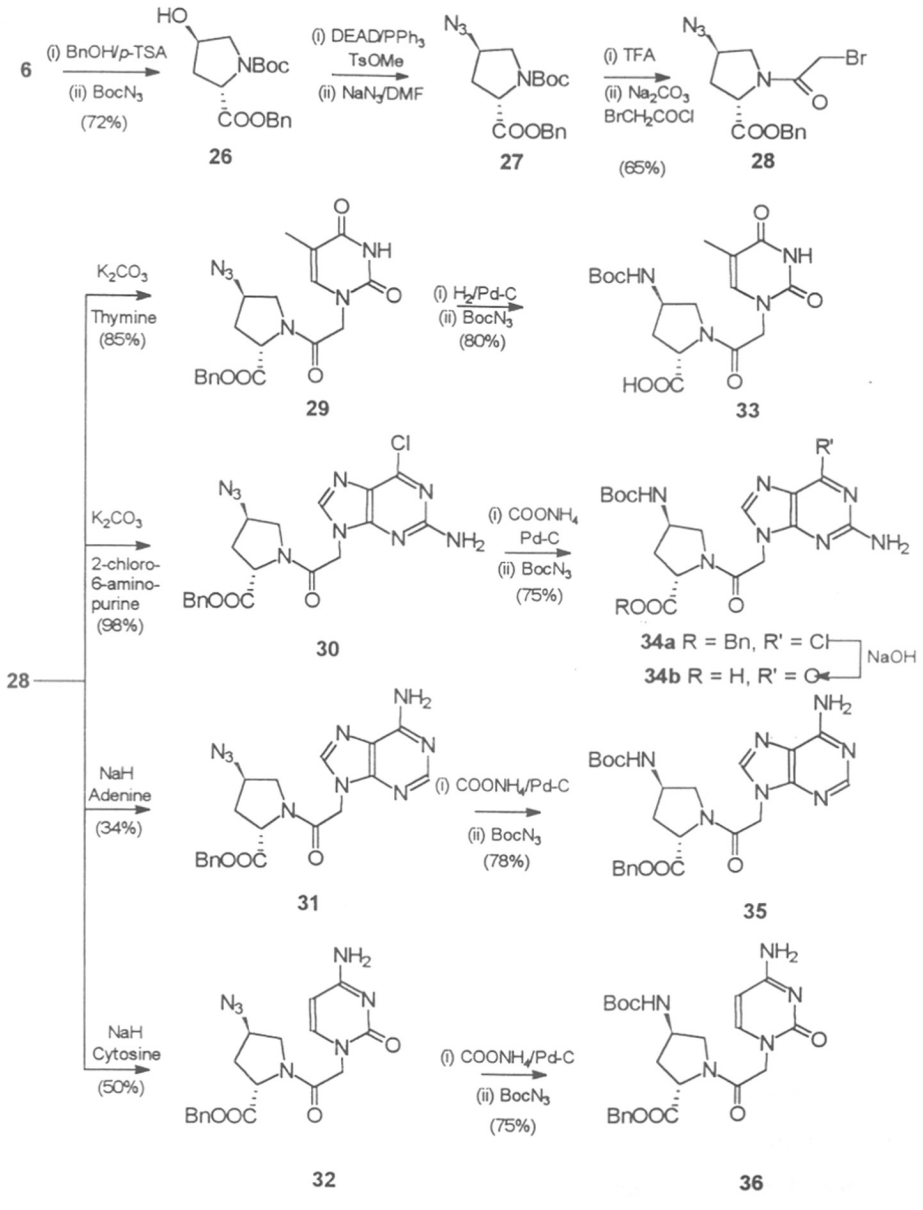


Figure 6: FAB-MS of compounus a: 21(2*R*,4*S*) and b: 25(2*R*,4*R*) in MNBA/LSIMS and NBA/LSIMS respectively.



Scheme 5

From a synthetic point of view, the combination of protecting groups as in the 4*R*-azido-N^{α}-(bromoacetyl)proline benzyl ester (**28**) is convenient, since in a subsequent

catalytic hydrogenation step, simultaneous transformation of 4-azido to 4-amino and deprotection of 2-benzyl ester to carboxylic acid can be achieved to get the free amino acid. It is also possible to selectively reduce 4-azido to 4-amino by catalytic transfer hydrogenation using ammonium formate/Pd-C without affecting the benzyl ester moiety. With this working rationale, 4*R*-hydroxy-L-proline **6** was converted into its benzyl ester by reaction with benzyl alcohol and *p*-toulenesulphonic acid (Scheme 5) followed by Boc protection of N^{α} using Boc-azide to get **26** which was transformed in two further steps to the 4*R*-azido derivative **27**. The removal of N^{α}-Boc with TFA afforded the free amine, which upon treatment with bromoacetylchloride in presence of aq. Na_2CO_3, afforded the desired common intermediate **28**. This reaction gave poor yields when triethylamine was employed as a base as per literature procedure[25] and the colored impurities from reaction mixture could not be easily removed by routine purification. The individual alkylation of the nucleobases thymine and 2-amino-6-chloropurine (precursor for guanine) with **28** was done by reaction of the two components in DMF in presence of K_2CO_3 as a base, followed by chromatographic purification to obtain **29** and **30** respectively, in excellent yields (95%). The alkylation of adenine and cytosine were achieved by reaction of the respective sodium salt (prepared by using NaH) with **28** to give corresponding azides **31** and **32** in moderate yields (Scheme 5). While the alkylation of thymine and cytosine is regiospecific, providing only N1-substitution products, that of purines is known[26] to be non-regiospecific leading to significant N7-substituted product along with the desired major product from N9-substitution. In the present work with adenine, N9-substituted product (**31**) was major. The N7-substitued product (<5%) was easily removed by chromatographic purification. For synthesis of G-monomer (**34b**), direct alkylation of guanine is not a facile reaction and hence 2-amino-6-chloropurine was alkylated with **28** to obtain exclusively the N9 regioisomer **30**. The compound **29** was hydrogenated to the 4-amino-2-carboxylic acid derivative with 10% Pd-C and the amino acid derivative was protected with Boc-azide to obtain **33**. Compounds **30-32** were then hydrogenated using ammonium formate and 10% Pd-C and protected with Boc-azide to get the protected monomers **34a-36** respectively. **34a** was then hydrolyzed to the desired G-monomer **34b** by 1N NaOH.

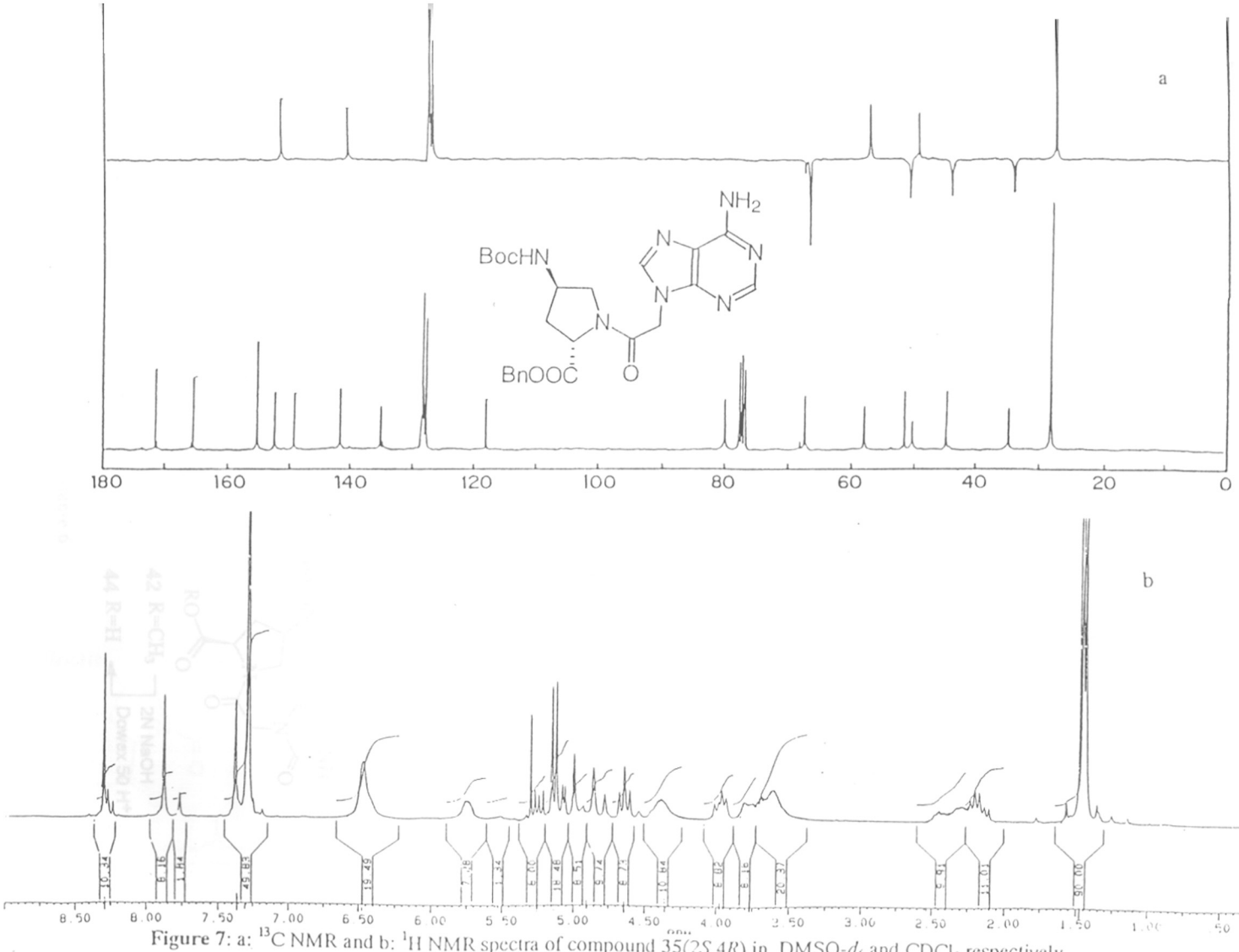
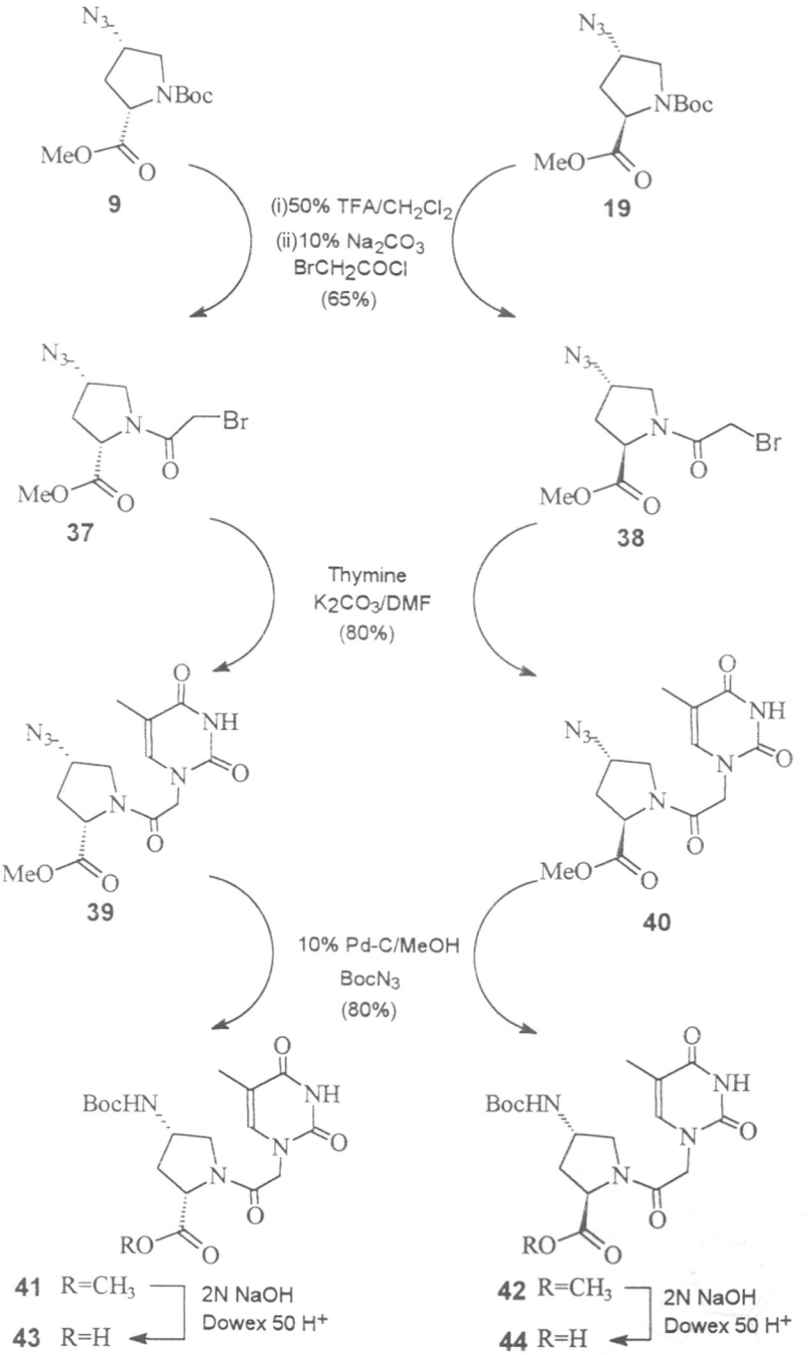


Figure 7: a: ^{13}C NMR and b: ^{1}H NMR spectra of compound 35(2*S*,4*R*) in DMSO-d_6 and $CDCl_3$ respectively.


N3
NBoc
MeO
O
9
19
(i)50% TFA/CH2Cl2
(ii)10% Na2CO3
BrCH2COCl
(65%)
Br
37
38
Thymine
K2CO3/DMF
(80%)
NH
39
40
10% Pd-C/MeOH
BocN3
(80%)
BocHN
RO
41 R=CH3
43 R=H
42 R=CH3
44 R=H
2N NaOH
Dowex 50 H+

Scheme 6

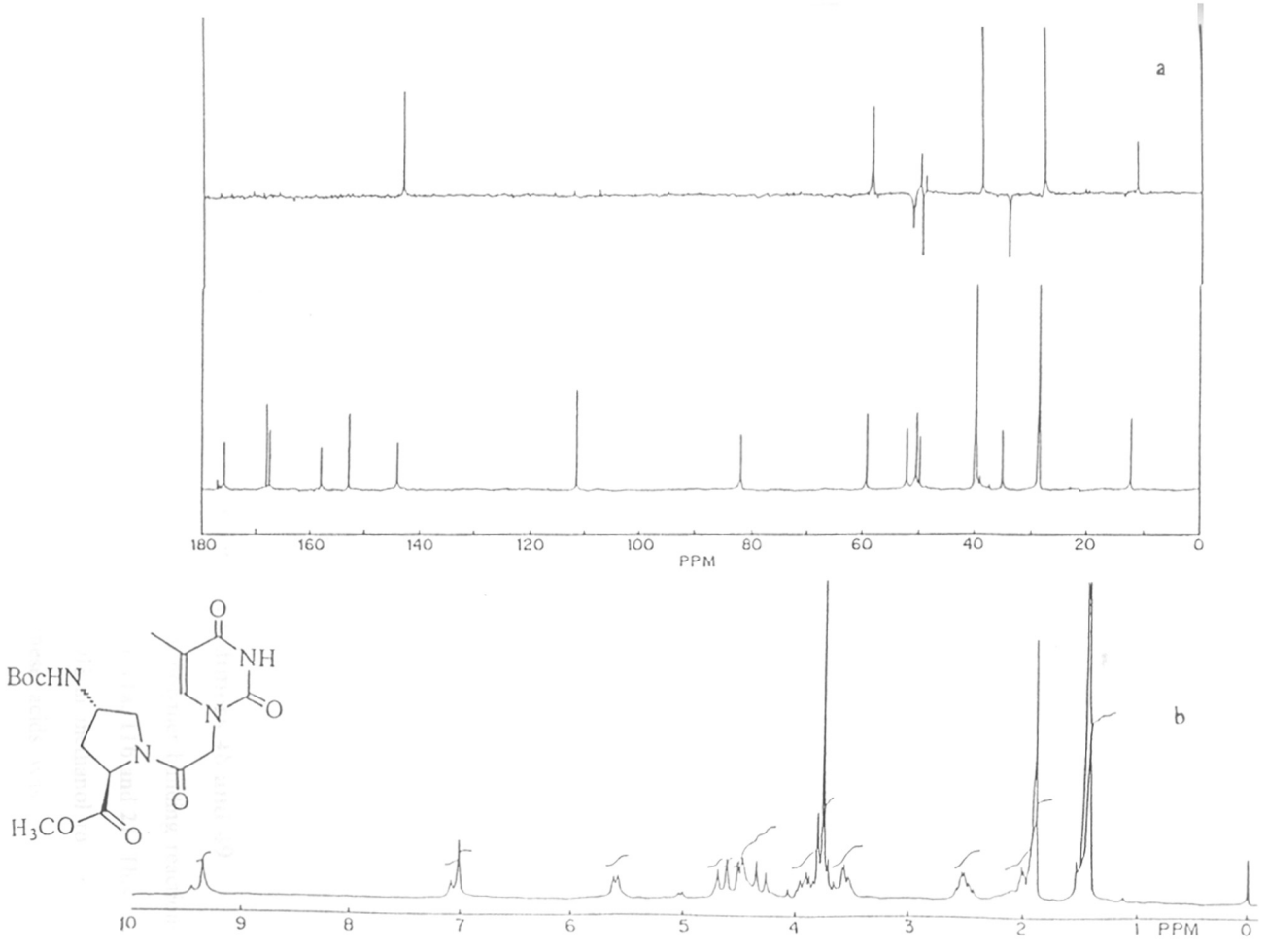


Figure 8: a: ^{13}C NMR and b: ^{1}H NMR spectra of compound **42** in $CDCl_3$

This formally completed the synthesis of all the nucleobase 4*R*,2*S* monomers required for building prolyl nucleic acids. This methodology coupled with that described in section 2.2 enables the synthesis of all possible 4-aminoproline stereoisomers, corresponding to N1 (T and C) and N9 (A and G) substitution with the four nucleobases.

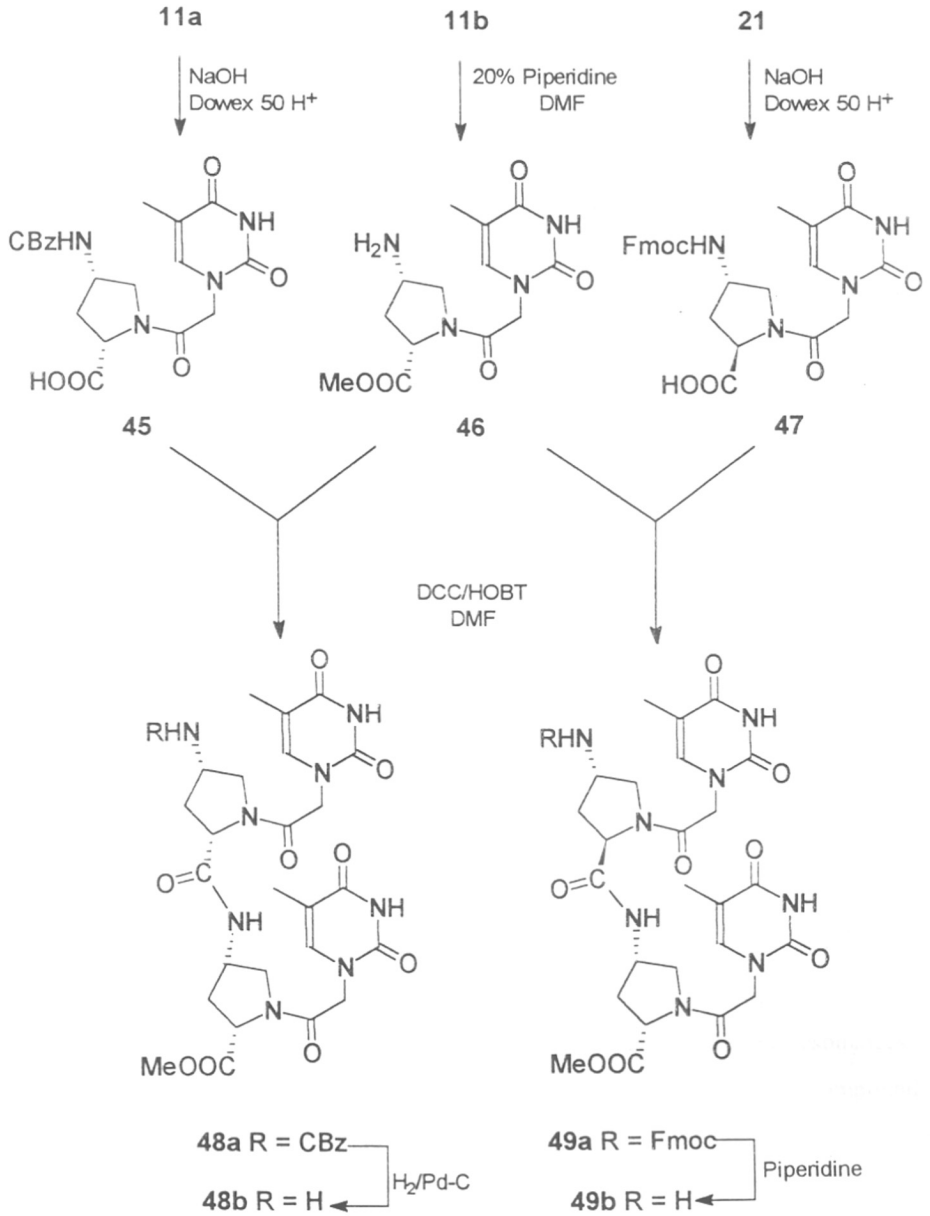
2.3.1 Synthesis of T monomers for solid phase peptide synthesis

It was planned to synthesize homooligomers of these prolyl monomers and test their hybridizing capability with DNA oligomers. Solid phase peptide synthesis was adopted for this purpose (*see Chapter 3*) and the Boc-protected monomers were synthesized as follows. The azides **9** and **19** were deprotected at N^{α} with TFA as discussed earlier and the resulting amine was reacted with excess bromoacetylchloride in presence of aqueous Na_2CO_3 in dioxane to obtain the bromo derivative **37** and **38** respectively in 68% yield (Scheme 6). Compounds **37** and **38** were stirred with thymine and K_2CO_3 to obtain the 4-azido thyminyl prolines **39** and **40** respectively in good yield. These azides **39** and **40** were further hydrogenated to the free amine and protected with Boc-azide to obtain the 4-N(*t*-butoxycarbonyl)-N^{α}(thymin-1-yl) proline methyl esters (**41**, **42**) which were purified by column chromatography. The hydrolysis of the esters **41** and **42** with 2N NaOH to obtain the corresponding acids **43** and **44** gave a very low yield when neutralized with 4N HCl as reported earlier.[22] Alternatively when Dowex-50 H^+ was used for neutralization of the resulting Na salt of the acid, the yields were excellent. The methyl ester **41** and **42** were similarly hydrolyzed to obtain pure acids **43** and **44**. The optical rotation of **44** obtained from NaOH hydrolysis and **33** obtained by hydrogenation gave equal but opposite rotation, indicating no observable racemization. These were used for solid phase peptide synthesis.

2.4 Synthesis and characterization of chiral PNA dimers 48 and 49

To examine the efficacy of synthesized monomers in polymer building reactions, the syntheses of dipeptides **48** and **49** were accomplished from **11a, 11b** and **21**. This was achieved by deprotection of **11a** and **21** with 2N aq. NaOH in methanol to give the corresponding acids **45** and **47**. The neutralization of these acids was conveniently

achieved by adding Dowex-50 H^+, to bring the reaction mixture to pH 7. This method of neutralization resulted in excellent yields (Scheme 7).



Scheme 7

The deprotection of C4 amino function in **11b** with piperidine gave the common amino component **46**. The acid component **45** was condensed with amine **46** in the presence of DCC-HOBT to give the dipeptide **49a** (NH-4*S*,2*S*;4*S*,2*S*-OH) with all *cis* conformation. The acid component **47** was then condensed with the amine **46** in presence of DCC-HOBT to give the prolyl PNA dipeptide **48a** (NH-4*S*,2*R*;4*S*,2*S*-OH) with one *trans* conformation (Scheme 7). The structural integrity of these two dipeptides **48a** and **49a** was established by ^{1}H NMR and the diagnostic mass peaks (M+1) at 723 and 811 respectively in FAB-MS. The dipeptide **48a** was deprotected by hydrogenation to obtain the free amine **48b**. Similarly the dipeptide **49a** upon deprotection with piperidine gave the free amine **49b** which was characterized by ^{1}H NMR.

The coupling yield of these dipeptides was observed to be very low. In this context Farese *et al.*.[26] have also reported an alternative method where they couple the aminoethylglycine unit on the PNA monomer and then couple the thymine acetic acid to the backbone secondary amine of the same. This method gave better coupling yield for solution phase synthesis of short peptides as the yields of direct coupling of monomers are observed to be low perhaps due to the bulky base group.

2.4.1 Spectral properties: ^{1}H and ^{13}C NMR

All compounds including those obtained as the intermediates in earlier pathways were unambiguously characterized for structural purity by ^{1}H and ^{13}C NMR. In general NMR spectra of N1-acyl substituted compounds (both *t*-Boc and carboxymethylated nucleobases) showed characteristic presence of *syn* and *anti* rotamer mixtures originating from the restricted rotation about the tertiary amide bond (>N-CO). The energy barriers for such rotation are high, leading to resolution of *syn* and *anti* rotamers at ambient temperature on NMR time scale, as shown by a doubling of relevant resonances. An equilibrium composition of approximately 80:20 was noticed in most of the compounds in $CDCl_3$, although the assignment of specific resonances to individual rotamers was not possible due to complex and overlapping spectral patterns for prolyl ring protons in 200 MHz spectra. Meltzer *et al.*[25] have shown by NOE correlation experiments that both E and Z form of isomers exist due to the tertiary amide linkages (Figure 11), and the ratio of

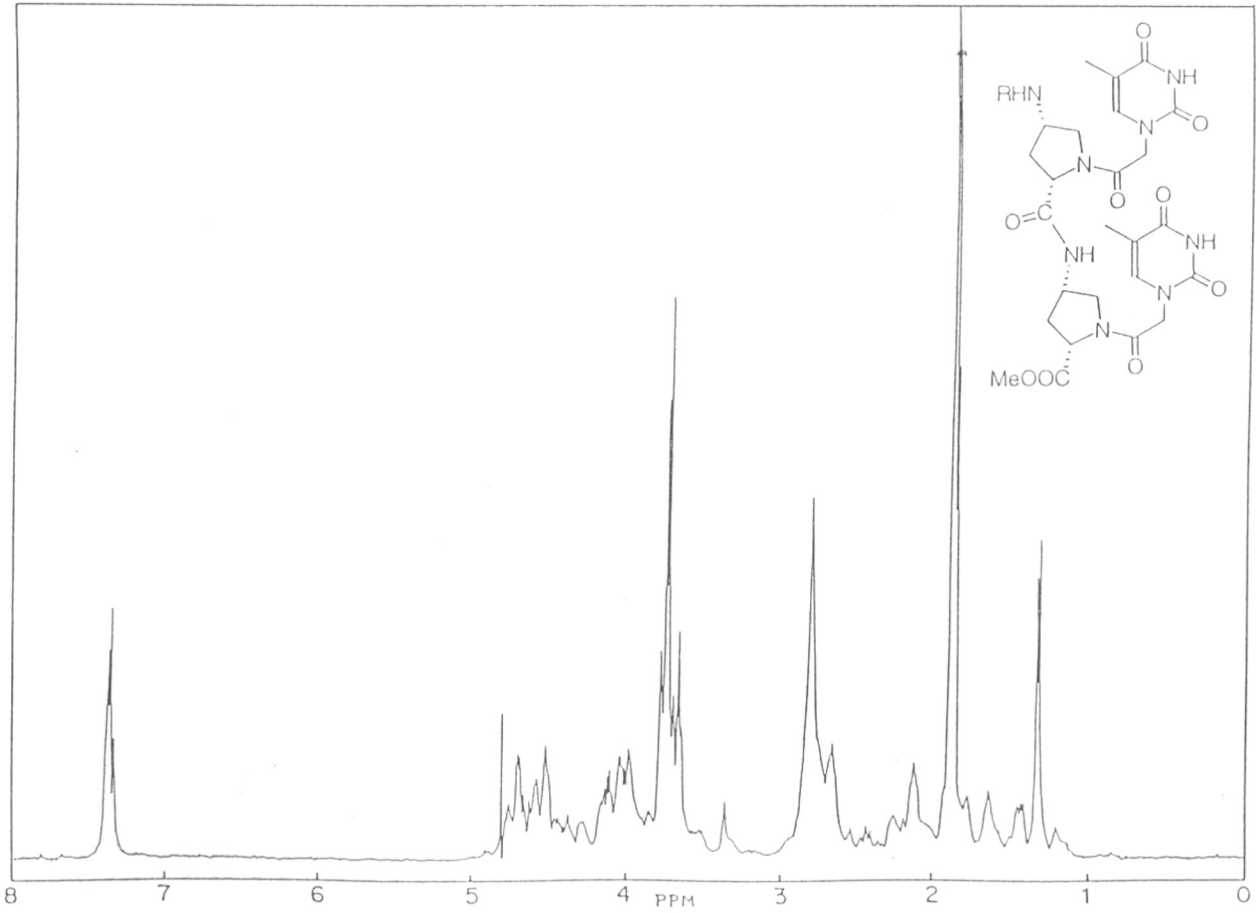


Figure 9: ^{1}H NMR (500 MHz) spectrum of compound **48b** in D_2O.

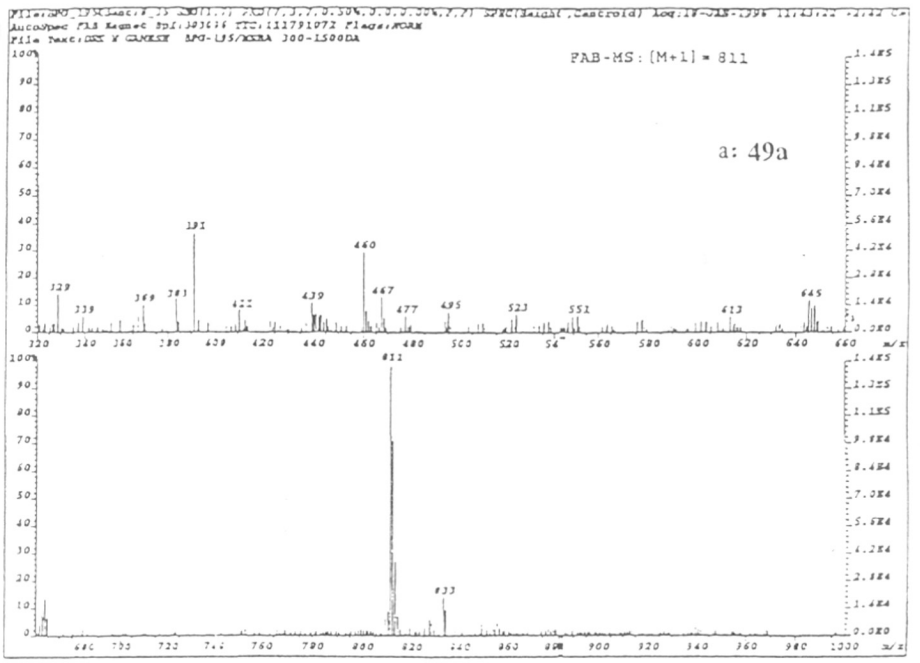
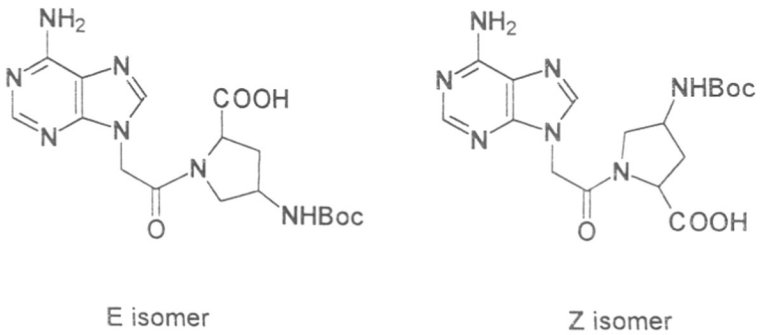


Figure 10: FAB-MS of dipeptides a: 49a and b: 48a in MNBA.

the two depends on the solvent. In D_2O, E and Z enantiomers are in 1:1 ratio and in DMSO-d_6 the E isomer predominates (80:20). The enantiomeric pairs **(9:23)/(13:19)** and **(11a:25)/ (15:21)** exhibited equal but opposite optical rotations, thus establishing the optical purity of compounds maintained in all these pathways.



E isomer Z isomer

Figure 11

^{13}C NMR is valuable in identification of N7/N9 regioisomers in alkylation reactions of purines. In case of 7- and 9- methyladenines it has been established that the N9 isomer exhibits downfield shifts of 6 and 4 ppm for C5 and C6 respectively and the

Table 1: ^{13}C NMR Chemical shifts

Compound	C2	C4	C5	C6	C8
7-methyl adenine	152.4	159.8	111.7	151.9	145.9
9-methyl adenine	152.5	149.9	118.7	155.9	141.4
compound **31**	152.7	149.7	118.5	155.5	141.1
compound **35**	152.4	149.3	118.2	155.4	141.6
compound **55**	152.3	149.6	118.3	155.4	141.2

upfield shifts of 10 and 4.5 ppm for C4 and C8 compared to the corresponding chemical shifts of N7 regioisomer. This pattern was also noticed with A-monomer of peptide-nucleic acids. The ^{13}C NMR spectral data of compounds **31** and **35** prepared presently by alkyation of adenine was consistent with the data reported for N9-isomer (Table 1).

2.4.2 Circular dichroism of dipeptides 48b and 49b

For a preliminary examination of the structural effects induced by different relationships of the chiral centers in dipeptides **48b** and **49b**, the CD spectra of the prolyl PNA are shown in Figure 12. These show characteristically different Cotton effects in the region 240-300 nm, with **48b** having only a negative CD band at 265 nm and **49b** both negative and positive CD bands at 288 and 255 nm respectively. These bands though weak, are completely absent in open chain PNA molecules[27] and must arise in **48b** and **49b** due to asymmetry in nucleobase stacking induced by the chiral peptide backbone. Substantial differences are also noticed in relative CD patterns in the region 210-230 nm, originating from stereochemical differences in the peptide linkage. The two dipeptides differ in prolyl ring stereochemistry at only one center and it is remarkable that this single change affects the sign, amplitude and band position of the CD spectrum. A detailed discussion on CD data of PNA oligomers is presented in the next chapter.

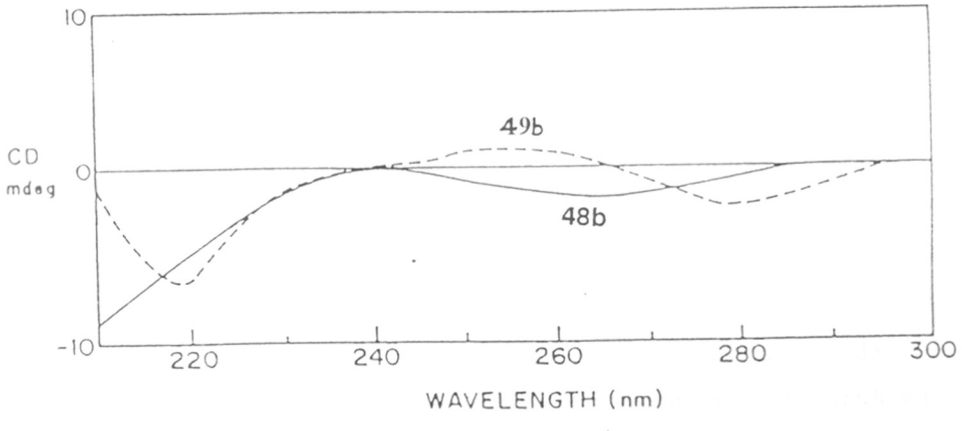


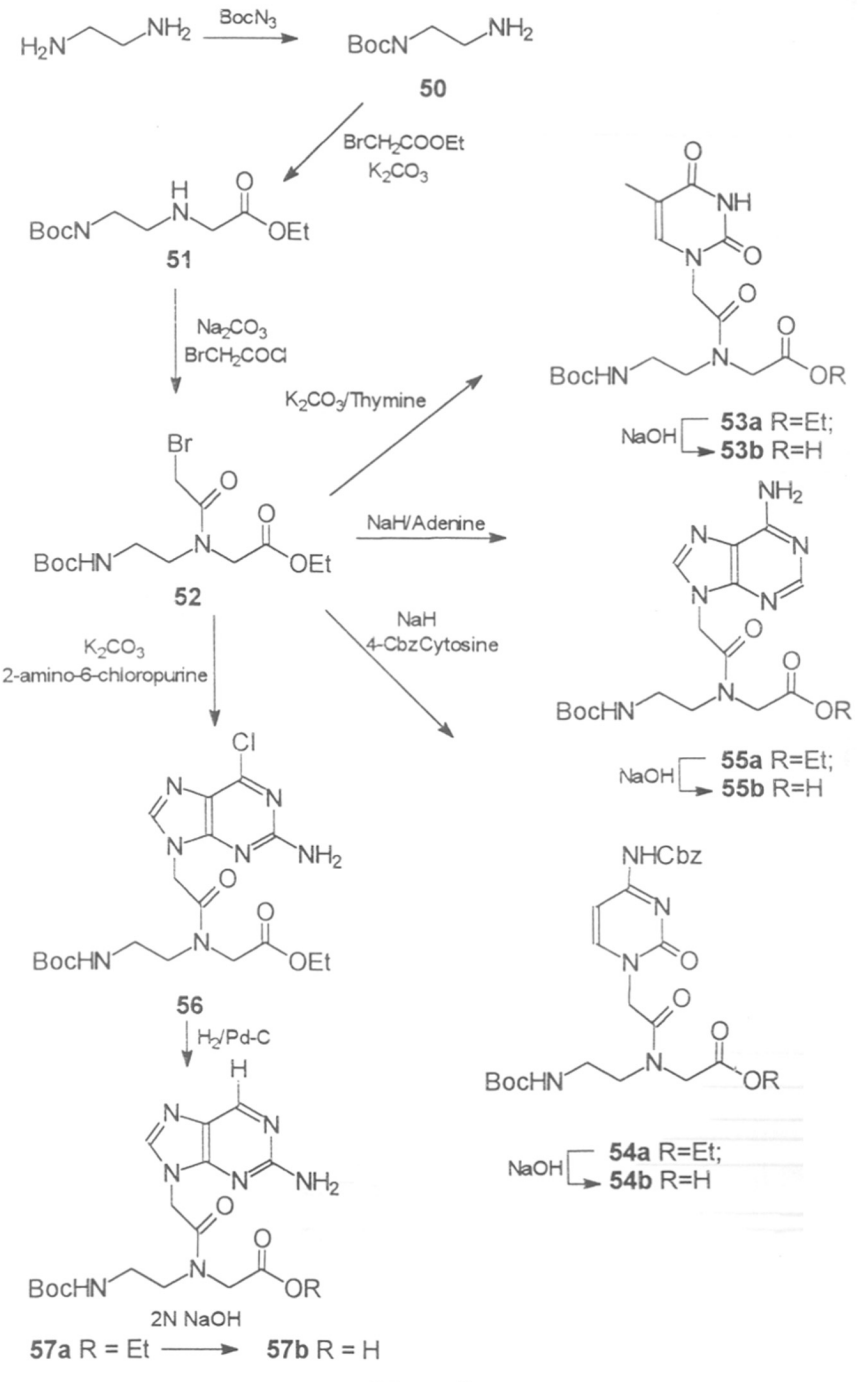
Figure 12: CD spectra of PrNA dipeptides 48b and 49b.

2.5 Synthesis of PNA monomers

PNA is known to show self-ordered helicity due to induced helical structure arising from terminal chiral center of the lysine amino acid. Hence, incorporation of the prolyl nucleic acid moiety on the terminal end of the PNA oligomer would be interesting. With this rationale, PNA monomers (**53** - **56**) were synthesized by a shorter route as shown in the Scheme 8. The advantage of this scheme is that purification was carried out at a later stage when **52** was obtained by column chromatography (60% from Boc azide).

Diaminoethane was converted to its mono-N-*t*-Boc derivative **50** by treating it with 0.1 equivalent of Boc-azide under high dilution conditions. The di-*t*-Boc derivative being insoluble in water can be easily removed. When compound **50** was alkylated efficiently with ethylbromoacetate in presence of one equivalent of K_2CO_3 as base under dilute condition, it was possible to get aminoethylglycine **51** in good yield. This was further treated with bromoacetyl chloride to yield the corresponding bromo derivative **52**. In this case the method followed by Meltzer *et al.* [25] using TEA as base gave poor yield and the resulting material was highly colored. Alternatively using aq. Na_2CO_3 in dioxane as the base and by adding excess bromoacetyl chloride the corresponding bromocompound was obtained.[17] The colorless oil obtained after chromatographic purification was used in alkylation of the base. The intermediate **52** can be stored at 0 °C and used as required.

The alkylation of **52** with thymine and cytosine is established to be regiospecific. Thymine was reacted with **52** using K_2CO_3 as base to obtain **53a** in high yield. In case of cytosine, the N4-amino group was protected as caboxybenzyl derivative, and used for alkylation employing NaH as the base to provide N1-substituted product **55a**. Although adenine is known to form both N7 and N9-substitution, N7-alkylation product was not present when NaH was used as the base to form sodium-adenalide, which was then reacted with **52** to obtain **52a** in moderate yield.


BocN3
50
BrCH2COOEt
K2CO3
51
Na2CO3
BrCH2COCl
52
K2CO3/Thymine
NaH/Adenine
NaH
4-CbzCytosine
2-amino-6-chloropurine
53a R=Et;
53b R=H
55a R=Et;
55b R=H
54a R=Et;
54b R=H
NaOH
56
H2/Pd-C
2N NaOH
57a R = Et
57b R = H

Scheme 8

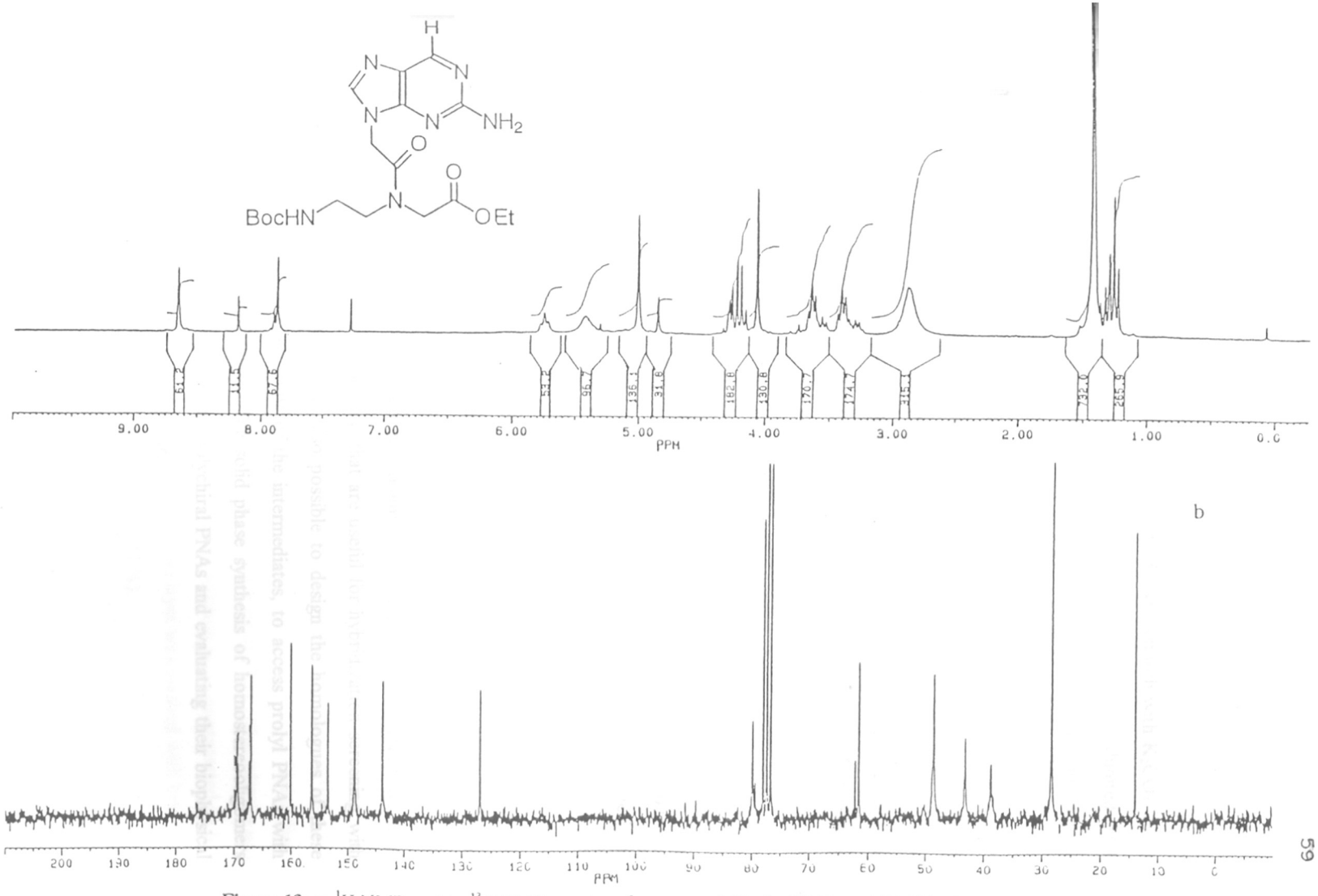


Figure 13: a: ^{1}H NMR and b: ^{13}C NMR spectra of compound 57a in $CDCl_3$ and DMSO-d_6 respectively.

The alkylation of 2-amino-6-chloropurine with **52** was smooth with K_2CO_3 as the base to yield the corresponding **56** in excellent yields. All the compounds exhibited ^1H & ^{13}C NMR data which is consistent with the reported data.[22] 2-Amino-6-chloropurine-aeg **52** was hydrogenated with 10% Pd-C to yield 2-amino purine derivative **57a** which is a fluorescent monomer. The compounds **53a**, **54a**, **55a**, and **57a** were hydrolysed to the corresponding acids **53b**, **54b** **55b**, and **57b**, which were used for solid phase synthesis. The fluorescent **57b** was further incorporated in PNA using solid phase synthesis to obtain a fluorescent PNA. Since the solid phase peptide synthesis was carried out manually and the completion of the reaction was estimated, the capping step was avoided. The 6-amino of **54b** and 2-amino of **57b** did not require protection as they were tested to be unreactive during peptide coupling conditions.

2.6 Conclusions

In this chapter, the design and preparation of stereochemically defined monomer building blocks, based on 4-aminoproline backbone, which are required for synthesis of conformationally rigid, polychiral PNA is demonstrated. The possible effects of stereostructural constrains on the backbone is preliminarily seen as induced chirality in base stacking in the model dipeptides which exhibit significant differences in their CD patterns upon change of stereochemistry even at a single site on backbone. The conformationally defined building blocks reported here with different nucleobases (A,T,C and G) when used in appropriate rational combinations, would lead to a library of chiral PNAs with diverse backbone geometry, that are useful for hybridization screening with appropriate DNA/RNA targets. It is also possible to design the homologues of these monomers (see Figure 2) from some of the intermediates, to access prolyl PNAs with controlled extension of backbones. The solid phase synthesis of homostereooligomers from these monomeric units to generate polychiral PNAs and evaluating their biophysical properties are discussed in the next chapter.

2.7 Experimental

All reagents were obtained from commercial sources and used without further purification. NaH was obtained from Aldrich as a 60% suspension in paraffin oil and the paraffin coating was washed off with pet-ether before use to remove the oil. All the solvents were dried according to literature procedures. Infrared spectra were recorded on a Perkin Elmer 599B instrument. ^{1}H NMR (200 MHz), ^{13}C NMR (50 MHz) spectra were recorded on Bruker ACF200 spectrometer fitted with an Aspect 3000 computer. All chemical shifts are with reference to TMS as an internal standard and are expressed in δ scale (ppm). The values given are directly from the computer printout. TLCs were carried out on (E.Merck 5554) precoated silicagel 60 F254 plates. TLCs were visualised with UV light and/or ninhydrin spray, followed by heating. Optical rotations were measured on JASCO DIP-181 polarimeter and CD spectra were obtained on a JASCO J600 instrument. All TLCs were run using DCM containing appropriate amounts of methanol. In NMR spectra that show splitting of peaks due to presence of rotameric mixtures, arising from the tertiary amide linkage, the major rotamer is designated as ma and the minor rotamer as mi. The ratio of major:minor rotamers is 80:20 unless otherwise mentioned. In cases where minor isomer is <10% (not prominent), only the peaks of major rotamer are reported. Melting points of the compounds reported are uncorrected.

***2S,4R*-Hydroxy-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (7)**: *trans*-4-Hydroxy-L-proline **6** (5.2 g, 39.6 mmol) was suspended in dry methanol and cooled in an ice bath. Thionyl chloride (3.1 mL, 41.6 mmol) was added slowly with vigorous stirring and the resulting solution was refluxed for 6 hr. The solvent was removed under reduced pressure using KOH trap and the solid obtained was further dried under high vacuum to obtain the methyl ester hydrochloride as white crystalline solid (6.8 g, 95%). This methyl ester hydrochloride (6.5 g, 35 mmol) was dissolved in 50 mL of dioxane:water (1:1) and Boc-azide (5.8 g, 41 mmol) was added dropwise followed by 8.5 ml of TEA. The mixture was stirred at ambient temperature for 16 hr. The resulting solution was concentrated to half its volume, extracted with ether (50 mL x 4), the ethereal layer was washed with brine and the solvent evaporated to yield compound **7** (8.01 g, 90 %).

IR. (neat): 3400, 2980-2900, 1735 and 1670 cm^{-1}. ^1H NMR, ($CDCl_3$ +D_2O) δ: 4.48-4.40 (m, 2H, H2 and H4), 3.72 (s, 3H, $COOCH_3$), 3.60-3.44 (m, 2H, H5A & H5B), 2.4-2.32 (m, 1H, H3A), 2.08-2.00 (m, 1H, H3B) and 1.45-1.42 (d, 9H, 3 x CH_3).

2*S*,4*R*-O-Methysulphonate-N$^{\alpha}$(*t*-butoxycarbonyl)-proline methylester (8): Mesyl chloride (1.15 mL, 14.6 mmol) was added dropwise into a solution of compound **7** (3.0 g, 12.2 mmol) in dry pyridine (50 mL) at 0 °C with constant stirring. After the addition was complete, the ice-bath was removed and stirring continued at ambient temperature for 2 h.. Pyridine was removed under reduced pressure and the residue was taken in water and extracted with DCM. The organic layer was washed with dil. aq $KHSO_4$ (2 x 50 mL), brine (50 mL) and dried over Na_2SO_4. The organic solvent was evaporated to yield a colorless oily substance which solidified at low temperature. The mesylate **8** (3.75 g, 96 %) obtained was used without further purification.

IR (nujol): 3019, 2400 (w, OMs), 1740, 1690, 1405, and 1215 cm^{-1}. ^1H NMR, ($CDCl_3$) δ: 5.32-5.22 (m, 1H, H4), 4.52-4.35 (m, 1H, H2), 3.85-3.73 (m, 2H, H5A & H5B), 3.75 (s, 3H, $COOCH_3$), 2.72-2.52 (br, 1H, H3A), 2.35-2.17 (m, 1H, H3B), 1.48 & 1.43 (s, 9H, 3 x CH_3).

2*S*,4*S*-Azido-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (9): The mesylate **8** (3.2 g, 9.9 mmol) was taken in dry DMF (20 mL) to which NaN_3 (1.6 g, 24.7 mmol) was added and the mixture was stirred at 50 °C for 4.5 hr.[20] The removal of DMF under vacuum, gave a residual oil which was suspended in water (100 mL) and the product was extracted with ether (60 mL x 3). The residue (2.45 g) obtained after removal of ether was purified by silica gel column chromatography using pet-ether/ethylacetate (0-50 %) eluant system to yield the azide **9** as a colourless oil (2.54 g, 94 %).

IR. (neat): 3010, 2105 (s, N_3), 1730, 1702 and 1399 cm^{-1}. ^1H NMR, ($CDCl_3$) δ: 4.43 (dq, J=6.0, 8.8, 5.3 Hz, H2), 4.33 (br-m, 1H, H4), 3.76 (s, 3H, OCH_3), 3.74 (m, 1H, H5A), 3.48 (m, 1H, H5B), 2.47 (m, 1H, H3A), 2.17 (dt, J=4.3, 1.3 Hz, 1H, H3B), 1.47 (s, 9H, $C(CH_3)_3$). ^{13}C NMR, ($CDCl_3$) δ: 172.2 (ma) and 171.9 (mi) (COOMe), 154.0 (mi) and 153.4 (ma) (OCON<), 80.5 (CMe_3), 59.3 (mi) and 58.3 (ma) (C4), 57.7 (ma) and 57.4

(mi) (C2), 52.3 (OCH_3), 51.3 (mi) and 50.8 (ma) (C5), 36.0 (ma) and 35.1 (mi) (C3), 28.3 ($C(CH_3)_3$). $[\alpha]_D^{20}$ = -31.3° (c = 0.4, MeOH).

2*S*,4*S*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (10a): The azide 9 (2.0g, 7.4 mmol) was dissolved in methanol (10 mL) and 10 % Pd-C (0.2 g) was added. The mixture was agitated in Parr hydrogenation apparatus for 4 h, under 40 psi H_2 pressure. Pd-C was filtered off over celite. Methanol was removed to give *cis*-4-amino-L-proline derivative in quantitative yield. This was taken in dioxane (7.5 mL) containing 10% aq. Na_2CO_3 (10 mL) and cooled in an ice bath. Fmoc chloride (2.1 g, 8.5 mmol) was slowly added in portions at 0 °C and stirring was continued at 0 °C for 4 h.The reaction mixture was further stirred at ambient temperature for 16 hr.[21] The reaction mixture was then concentrated to half its volume and water (50 mL) was added before extracting with ether (50 mL x 3). The pale yellow residue after removal of ether was purified on a silica gel column using pet-ether/ethylacetate gradient as eluant to obtain **10a** (2.2 g, 63 %) as a white solid.

^{1}H NMR, ($CDCl_3$) δ: 7.78 (d, J=8 Hz, 2H, Ar), 7.60 (d, J=8 Hz, 2H, Ar), 7.45-7.3 (m, 4H, Ar), 5.85 (br, 1H, NH-Boc), 4.45-4.15 (m, 5H, H4, H2, CH & OCH_2 of Fmoc), 3.80 (2s, 3H, $COOCH_3$), 3.80-3.48 (m, 2H, H5A & H5B), 2.62-2.43 (m, 1H, H3A), 2.10-1.95 (m, 1H, H3B), 1.52 &1.47 (s, 9H, 3 x CH_3).

2*S*,4*S*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(thymin-1-yl)proline methyl ester (11a): Compound **10** (2.02 g, 4.3 mmol) was taken in 50 % TFA/DCM (5 mL), and stirred at ambient temperature for 1 hr. The solvent was removed under reduced pressure using KOH trap and the free amine was taken in water and neutralized with 10 % aq. $NaHCO_3$ followed by repeated extraction with ethylacetate (50 mL x 4). The combined organic layer was dried over anhydrous Na_2SO_4 and the solvent removed to obtain the free amine. The residual oil was desiccated over KOH for 12 h. The free amine (4.3 mmol), HOBT (0.54 g, 4.3 mmol) and thymine acetic acid (0.866 g, 4.8 mmol) were dissolved in dry DMF (10 mL). The solution was cooled in an ice bath, DCC (0.9 g, 4.5 mmol) was added and the reaction mixture was stirred for 1 h in an ice bath and then for 3 h at RT.[22] The reaction mixture was filtered to remove DCU and washed thoroughly with cold DCM.

The collective filtrate was washed successively with 10% aq. Na_2CO_3 (50 mL x 2), dil. $KHSO_4$ (50 mL x 2) and brine (40 mL). The organic layer was separated, dried over anhydrous Na_2SO_4 and the solvent was removed under vacuum. The oily residue was purified by silica gel column chromatography using DCM/MeOH (0-10 %) gradient as the eluant to obtain the title compound **11** in pure form (1.69 g, 74.9%).

TLC: 10% MeOH/CH_2Cl_2 , R_f: 0.4. ^{1}H NMR, ($CDCl_3$) δ: 8.62 (s, 1H, *T* NH), 7.76-7.26 (m, 8H, Fmoc), 7.07 (mi) & 7.05 (ma) (s, 1H, *T* H6), 5.98 (ma) (d, J = 11 Hz, C4 NH) & 5.30 (mi, br, C4 NH), 4.73-4.16 (m, overlapping signals, 7H, H2, H4, *T* CH_2, Fmoc CH and OCH_2), 4.0-3.57 (br-m, 2H, H5), 3.78 (ma) & 3.76 (mi) (s, 3H, $COOCH_3$), 2.62-2.45 (br-m, 1H, H3B), 2.21-2.04 (m, 1H, H3A) & 1.93 (s, 3H, *T* CH_3). ^{13}C NMR, ($CDCl_3$) δ: 172.7(COO), 166.1 (mi) & 165.4 (ma)(>NCO), 164.4 (NCO), 155.8 (*T* C2), 151.3 (*T* C4), 141.2 (*T* C6), 143.5, 127-124, 119 (C Ar), 110.4 (*T* C5), 66.5 (CH_2), 57.9 (OCH_3), 52.6 (C2), 51.9 (*T* CH_2), 50.8 (C4), 48.6 (C5), 46.85 (CH), 34.5 (C3), 12.2 (*T* CH_3). $[\alpha]_D^{20}$ (c = 0.4, MeOH) = -38.8°, FAB-MS:[M+1] = 533.

2*S*,4*S*-N(Carboxybenzyl)-N$^{\alpha}$(thymin-1-yl)proline methyl ester (11a): The compound 10b was deprotected with TFA and coupled with thymine acetic acid using DCC/HOBT coupling method, following the procedure for **11a** to obtain 11b in good yield.

^{1}H NMR, ($CDCl_3$) δ: 8.92 (s, 1H, *T* NH), 7.26 (m, 5H, Ar), 7.07 (mi) & 7.05 (ma) (s, 1H, *T* H6), 5.98 (ma) 4.98 (s, 2H, Ar), 4.62-4.16 (m, overlapping signals, 4H, H2, H4, *T* CH_2) 3.98-3.57 (br-m, 2H, H5), 3.73 (ma) & 3.69 (mi) (s, 3H, $COOCH_3$), 2.61-2.42 (br-m, 1H, H3B), 2.27-2.06 (m, 1H, H3A) & 1.93 (s, 3H, *T* CH_3).

2*S*,4*S*-O-*p*-Toluenesulfonyl-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (12): Proline methyl ester **7** (3.5 g, 14.3 mmol) was taken in dry THF (100 mL) along with PPh_3 (4.6 g, 17.5 mmol). To the stirred mixture DEAD (3.1g, 17.8 mmol) in 20 mL of THF was added slowly at 0 °C and stirred for 2 min., followed by dropwise addition of methyl *p*-toluenesulfonate (3.3 g, 17.1 mmol) in 20 mL of THF.[23] The resulting solution was stirred at 0 °C for 20 min. and then at room temperature overnight. Solvent was removed under vacuum and the residue was directly purified by repeated silica gel column

chromatographies. The purification was repeated to remove diethylhydrazine dicarboxylate to obtain the pure tosylate **12.**

IR: (neat), 3092, 1750, 1705, 1592 and 1402 cm^{-1}. ^1H NMR, ($CDCl_3$) δ: 7.78 (d, J=8.0 Hz, 2H, Ar), 7.38 (d, J=8.0 Hz, 2H, Ar), 5.05 (br, 1H, H4), 4.63-4.30 (m, 1H, H2), 3.78-3.65 (m, 2H, H5A & H5B), 3.70 (s, 3H, $COOCH_3$), 2.52-2.21 (m, 2H, H3A & H3B), 2.45 (s, 3H, $ArCH_3$), 1.52 & 1.40 (s, 9H, 3 x CH_3).

2*S*,4*R*-Azido N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (13): The pure tosylate **12** was taken in DMF and reacted with NaN_3 (1.5 g, 22 mmol) as reported for the preparation of compound **8** to obtain the corresponding azide **13** (1.3 g) in pure form after column chromatography.

IR (neat): 3010, 2105 (s, N_3), 1730, 1702 and 1399 cm^{-1}. ^1H NMR, ($CDCl_3$) δ: 4.36 (dt, J = 6.9, 7.6 Hz, H2), 4.19 (m, 1H, H4), 3.74 (s, 3H, OCH_3), 3.64 (dd, J = 5.3, 12 Hz, H5A), 3.47 (dt, J = 3.8, 12 Hz, H5B), 2.32 (m, 1H, H3A), 2.16 (m, 1H, H3B), 1.46 (ma) and 1.41 (mi) [s, 9H, $C(CH_3)_3$]. ^{13}C NMR, ($CDCl_3$) δ: 172.5 (ma) and 172.2 (mi) (CO), 153.6 (mi) and 152.9 (ma) (NCO), 80.0 ($CCH_3)_3$, 59.0 (mi) and 58.5 (ma) (C4), 51.7 (OCH_3), 51.1 (mi) and 50.9 (ma) (C5), 35.9 (ma) and 35.0 (mi) (C3), 28.8 ($C(CH_3)_3$).

$[\alpha]_D^{20}$ = -60.9 ° (c = 0.4, MeOH).

2*S*,4*R*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (14): The azide **13** (1 g, 3.7 mmol) was hydrogenated with 10 % Pd-C (0.1 g) in methanol and the amine obtained was further protected with Fmoc-Cl by following the procedure reported for **10** to obtain the required compound **14** (1.08 g, 63 %) after column chromatographic purification.

^1H NMR, ($CDCl_3$) δ: 7.78 (d, J=8 Hz, 2H, Ar), 7.60 (d, J=8 Hz, 2H, Ar), 7.45-7.30 (m, 4H, Ar), 5.85 (br, 1H, NH), 4.45-4.15 (m, 5H, H4, H2, CH & OCH_2 of Fmoc), 3.80 (2s, 3H, $COOCH_3$), 3.80-3.48 (m, 2H, H5A & H5B), 2.62-2.43 (m, 1H, H3A), 2.10-1.95 (m, 1H, H3B) 1.48 & 1.45 (s, 9H, $C(CH_3)_3$).

2*S*,4*R*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(thymin-1-yl)proline methyl ester (15): The above compound **14** (0.8 g, 1.7 mmol) was deprotected with 50 % TFA/DCM (5 mL) to obtain the corresponding free amine (0.5 g, 79 %, 1.36 mmol) which was coupled with

thymine acetic acid (0.26 g, 1.45 mmol) in the presence of HOBT (0.18 g, 1.38 mmol) and DCC (0.3 g, 1.4 mmol) in DMF (20 mL) to afford the required compound **15** (0.53 g, 79.8 %).

^{1}H NMR, ($CDCl_3$), δ: 9.75 (s, 1H, *T* NH), 7.76-7.25 (m, 8H, Fmoc), 7.02 (mi) & 6.96 (ma) (s, 1H, *T* H6), 6.34 (ma) & 6.12 (mi) (d, 1H, J = 7.4 Hz, C-4 NH), 4.82-3.89 (m, 7H, H2, H4, *T* CH_2, Fmoc CH, OCH_2), 3.72 (s, 3H, $COOCH_3$), 3.80-3.66 (m, 2H, H5A,B), 2.44 (mi) & 2.17 (ma) (br-m, 2H, H3A,B), 2.1 (s, 3H, *T* CH_3). ^{13}C NMR, ($CDCl_3$), δ: 172.2 (COO), 166.3 (mi) & 165.5 (ma) (>NCO), 164.3 (NCOO), 156.1 (*T* C2), 151.6 (*T* C4), 141.3 (*T* C6), 143.2, 127.8-124.5, 119.8 (all C-Ar) 110.8 (*T* C5), 66.3(CH_2), 57.9 (OCH_3), 52.6 (C2), 51.4 (*T* CH_2), 50.7 (C4), 48.9 (C5), 46.2(CH), 35.0 (C3), 12.2 (*T* CH_3). $[\alpha]_D^{20}$ = -22.5 (c = 0.4, MeOH). FAB-MS: [M+1] = 533.

4*R*-*allo*-Hydroxy-D-proline hydrochloride (16): To a solution of acetic anhydride (28 mL) and glacial acetic acid (85 mL) at 50 C was added 4-hydroxy proline **6** (7.2 g, 54.9 mmol) and the mixture was refluxed for 6 h. After 6 h the mixture was cooled and solvents removed under reduced pressure to get a thick oil.[24] This oil was taken in 2N HCl (100 mL) and refluxed for 3.5 hr. The reaction mixture was treated with activated charcoal and filtered. The aqueous solution was concentrated and the 4*R*-*allo*-hydroxy-D-proline hydrochloride **16** was crystallized from water as fine white needles (7.2 g, 79%).
mp. = 115 °C (*lit.*153-153.5 °C for free amine). ^{13}C NMR, (D_2O) δ: 172.6, 69.7, 59.1, 54.3 and 37.6.

2*R*,4*R*-Hydroxy-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (17): The (2*R*,4*R*)-4-hydroxyproline hydrochloride **16** (6.5 g, 35.9 mmol) was protected as the corresponding Boc-methyl ester as described for the L-isomer **7** to the Boc-proline methyl ester **17** (6.98 g, 80 %).
IR (nujol): 3450, 2950, 2860, 1730, and 1660 cm^{-1}. ^{1}H NMR, ($CDCl_3$) δ: 4.39-4.26 (m, 2H, H2, H4), 3.79 (d, 3H, $COOCH_3$), 3.66-3.49 (m, 2H, H5), 2.37-2.25 (m, 1H, H3A), 2.12-2.04 (m,1H, H3B) and 1.46 & 1.45 (s, 9H, 3 x CH_3). ^{13}C NMR, ($CDCl_3$), δ: 70.4, 69.4, 57.5, 55.1, 54.6, 52.3, 52.1, 38.5, 37.8, 28.2 & 28.1.

2*R*,4*R*-O-Methylsulfonyl N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (18): The Boc-methyl ester **17** (2.4 g, 10 mmol) was reacted with mesyl chloride (1.2 mL, 15 mmol) in pyridine as in case of **8** followed by usual work up to obtain the mesylate **18** as a thick oil (2.92 g, 92 %).

IR (nujol): 3019, 2400(w, OMs), 1740, 1690, 1405, and 1215 cm^{-1}. ^{1}H NMR, ($CDCl_3$) δ: 5.32-5.22 (m, 1H, H4), 4.52-4.35 (m, 1H, H2), 3.85-3.73 (m, 2H, H5A & H5B), 3.75 (s, 3H, $COOCH_3$), 3.05 (s, 3H, SO_2CH_3), 2.72-2.52 (br, 1H, H3A), 2.35-2.17 (m, 1H, H3B), 1.48 & 1.43 (s, 9H, $C(CH_3)_3$).

2*R*,4*S*-Azido-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (19): Mesylate **18** (2.42 g, 7.4 mmol) and NaN_3 (1.25 g, 18.7 mmol) were stirred in DMF followed by standard workup to obtain the azide **19** (1.82g, 91 %) following the procedure for the preparation of compound 9.

^{1}H NMR, ($CDCl_3$) δ: 4.36 (dt, J = 6.9, 7.6 Hz), 4.19 (m, 1H, H4), 3.74 (s, 3H, OCH_3), 3.64 (dd, J = 5.3, 12 Hz, H5A), 3.47 (dt, J = 3.8, 12 Hz, H5B), 2.32 (m, 1H, H3A), 2.16 (m, 1H, H3B), 1.46 (ma) and 1.41 (mi) (s, 9H, $C(CH_3)_3$). ^{13}C NMR, ($CDCl_3$) δ: 172.5 (ma) and 172.2 (mi) (CO), 153.6 (mi) and 152.9 (ma) (NCO), 80.0 ($C(CH_3)_3$), 59.04 (mi) and 58.5 (ma) (C4), 51.7 (OCH_3), 51.1 (mi) and 50.9 (ma), (C5), 35.9 (ma) and 35.0 (mi) (C3), 28.80($C(CH_3)_3$). $[\alpha]_D^{20}$ = +60.9 (c = 0.4, MeOH).

2*R*,4*S*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (20): The azide **19** (1.03 g, 3.8 mmol) was reduced using H_2, 10 % Pd-C (0.1 g) and the resulting amine was protected with Fmoc-Cl (1.225 g, 4.7 mmol) by following the procedure used for the L-isomer **10** to get the protected 4-amino-D-proline derivative **20** (1.32 g, 74 %).

^{1}H NMR, ($CDCl_3$) δ: 7.78 (d, J=8 Hz, 2H, Ar), 7.60 (d, J=8 Hz, 2H, Ar), 7.45-7.30 (m, 4H, Ar), 5.85 (br, 1H, NH), 4.45-4.15 (m, 5H, H4, H-, CH & OCH_2 of Fmoc), 3.80 (2s, 3H, $COOCH_3$), 3.81-3.48 (m, 2H, H5A & H5B), 2.62-2.43 (m, 1H, H3A), 2.10-1.95 (m, 1H, H3B), 1.48 & 1.45 (s, 9H, $C(CH_3)_3$).

2*R*,4*S*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(thymin-1-yl)proline methyl ester (21): Compound **20** (1.1 g, 2.3 mmol) after Boc-deprotection with TFA/DCM gave the

corresponding free amine (0.71 g, 1.9 mmol) which was then coupled with thymine acetic acid (0.368, 2.1 mmol), in the presence of HOBT (0.25 g, 1.9 mmol) and DCC (0.45 g, 2.2 mmol) to give the desired compound **21** (0.79g, 82%) as in case of compound **11**.

^{1}H NMR, ($CDCl_3$) δ: 9.82 (s, 1H, *T* NH), 7.76-7.25 (m, 8H, Fmoc), 7.02 (mi) & 6.94 (ma) (s, 1H, *T* H6), 6.40 (ma) & 6.18 (mi) (d, J = 7.4 Hz, C4 NH), 4.83-3.89 (m, 7H, H2, H4, *T* CH_2, Fmoc CH, CH_2), 3.71 (ma) & 3.73 (mi) (s, 3H, OCH_3), 3.78-3.72 (m, 2H, H5A,B), 2.44 (mi) & 2.22 (ma) (brm, 2H, H3A,B), 2.11 (s, 3H, *T* CH_3). ^{13}C NMR, ($CDCl_3$), δ: 171.8 (COO), 165.2 (mi) & 164.6 (ma) (>NCO), 163.2 (NCOO), 156.4 (*T* C2), 151.9 (*T* C4), 141.5 (*T* C6), 143.3, 127.8-124.5, 119.5 (all CAr), 111.3 (*T* C5), 66.3 (CH_2), 58.2 (OCH_3), 52.9 (C2), 51.9 (*T* CH_2), 51.0 (C4), 49.2 (C5), 46.8 (CH), 35.4 (C3), 12.6 (*T* CH_3). $[\alpha]_D^{20}$ = +22.8 ° (c = 0.4, MeOH). FAB-MS: [M+1] = 533.

2*R*,4*S*-O-*p*-Toluenesulfonate-N$^{\alpha}$(t-butoxycarbonyl)proline methyl ester (22): The Boc-methyl ester **17** (2.41 g, 9.7 mmol) was reacted with PPh_3 (3.16 g, 14.3 mmol), DEAD (2.10 g, 12.2 mmol) and methyl *p*-toluenesulfonate (2.25 g, 12.5 mmol) as in case of **12** to obtain the tosylate **22.**

IR(neat): 3092, 1750, 1705, 1592 and 1402 cm^{-1}. ^{1}H NMR, ($CDCl_3$) δ: 7.78 (d, J=8.0 Hz, 2H, Ar), 7.38 (d, J=8.0 Hz, 2H, Ar), 5.05 (br, 1H, H4), 4.63-4.30 (m, 1H, H2), 3.78-3.65 (m, 2H, H5A & H5B), 3.70 (s, 3H, OCH_3), 2.52-2.20 (m, 2H, H3A & H3B), 2.45 (s, 3H, $ArCH_3$), 1.52 & 1.43 (s, 9H, 3 x CH_3).

(2*R*,4*R*)-4-Azido-N$^{\alpha}$(t-butoxycarbonyl)proline methyl ester (23): The tosylate **22** was reacted with NaN_3 (1.05 g, 15 mmol) in DMF (20 mL) to obtain the azide **23** (1.36 g) following the procedure as reported for compound **9**.

IR. (neat): 3010, 2105 (s, N_3), 1730, 1702 and 1399 cm^{-1}. ^{1}H NMR, ($CDCl_3$) δ: 4.43 (dq, J=6.0, 8.8, 5.3 Hz, H2), 4.33 (brm, 1H, H4), 3.76 (s, 3H, $COOCH_3$), 3.74 (m, 1H, H5A), 3.48 (m, 1H, H5B), 2.47 (m, 1H, H3A), 2.17 (dt, J=4.3, 1.3 Hz, 1H, H3B), 1.49 & 1.43 (s, 9H, $C(CH_3)_3$). ^{13}C NMR, ($CDCl_3$) δ: 172.2 (ma) and 171.9 (mi) (COO), 154.0 (mi) and 153.4 (ma) (OCON<), 80.5 ($C(CH_3)_3$), 59.3 (mi) and 58.3 (ma) (C4), 57.7 (ma) and

57.4 (mi) (C2), 52.3 (OCH_3), 51.3 (mi) and 50.8 (ma) (C5), 36.0 (ma) and 35.1 (mi) (C3), 28.3 ($C(CH_3)_3$). $[\alpha]_D^{20}$ = +31.3 (c = 0.4, MeOH).

2*R*,4*R*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(*t*-butoxycarbonyl)proline methyl ester (24): The azide **23** (1.1 g, 4 mmol) was reduced using H_2-10% Pd-C (0.1 g) and the resulting amine was protected with Fmoc-Cl (1.2 g, 4.5 mmol) by following the procedure used for the L-isomer **10** to get the protected aminoproline **24** (1.28 g, 68 %).

^{1}H NMR, ($CDCl_3$) δ: 7.78 (d, J=8 Hz, 2H, Ar), 7.60 (d, J=8 Hz, 2H, Ar), 7.45-7.30 (m, 4H, Ar), 5.85 (br, 1H, NH), 4.45-4.15 (m, 5H, H4, H2, CH & OCH_2 of Fmoc), 3.80 (mi) & 3.78 (ma) (s, 3H, $COOCH_3$), 3.80-3.48 (m, 2H, H5A & H5B), 2.62-2.43 (m, 1H, H3A), 2.10-1.95 (m, 1H, H3B), 1.52 & 1.47 (s, 9H, $C(CH_3)_3$).

2*R*,4*R*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(thymin-1-yl)proline methyl ester (25): The free amine (0.62 g, 80%, 1.6 mmol) obtained from compound **24** (0.98 g, 2.1 mmol) after Boc-deprotection with TFA/DCM, was coupled with thymine acetic acid (0.31 g, 1.75 mmol), HOBT (0.2 g, 1.6 mmol) and DCC (0.35 g, 1.7 mmol) to afford the required compound **23** (0.74 g, 85 %) by following the procedure as in case of **11**.

^{1}H NMR, ($CDCl_3$) δ: 7.76-7.26 (m, 8H, Fmoc), 7.08 (mi) & 7.01 (ma) (s, 1H, *T* H6), 6.05 (ma) & 5.40 (mi) (d, 2H, J = 7.4 Hz, C4 NH), 4.75-4.15 (m, 7H, H2, H4, *T* CH_2, Fmoc CH, CH_2), 3.98-3.56 (br-m, 2H, H5), 3.77 (ma) & 3.75 (mi) (s, 3H, OCH_3), 2.50 (br-m, 2H, H3A,B), 1.93 (s, 3H, *T* CH_3). ^{13}C NMR, ($CDCl_3$) δ: 172.7 (COO), 166.2 (mi), & 165.5 (ma) (>NCO), 164 (NCOO), 156.0 (*T* C2), 151.6 (*T* C4), 141.4 (*T* C6), 143.5, 127.4-124.7, 119 (C-Ar), 110.6 (*T* C5), 66 (CH_2), 58.1 (OCH_3), 52.8 (C2), 52.0(*T* CH_2), 51.0 (C4), 48.9 (C5), 46.8 (CH), 34.8 (C3), 12.3 (*T* CH_3). $[\alpha]_D^{20}$ = +38.63 , (c = 0.4, MeOH), FAB-MS: [M+1]=533.

2*R*,4*S*-Hydroxy-N$^{\alpha}$(*t*-butoxycarbonyl)proline benzyl ester 26: To *trans*-4-hydroxy-L-proline **6** (3.35 g, 25.0 mmol) was added benzyl alcohol (25 mL), *p*-toluenesulfonic acid (6.1g, mmol) and toluene (100 mL) and the mixture was refluxed with continuous removal of water using Dean-Stark appartus. The mixture was cooled and dry diethyl ether was added, upon which, white crystals of *p*-toluenesulfonate salt of proline benzyl ester

separated. The solvent was filtered off and the crystalline salt was taken in DMSO (30 mL) triethylamine (9.5 mL) and Boc-azide (3.5 g, 0.25 mmol) was added. The mixture was stirred overnight at ambient temperature. The resultant mixture was taken in equal amount of water, extracted with ether (50 mL x 3), and the organic layer was washed sequentially with $KHSO_4$ (50 mL x 2), water (50 mL), 10% aq. Na_2CO_3 (50 mL x 2) and water (50 mL). Upon concentration of ether layer compound **26** was obtained.

^{1}H NMR, ($CDCl_3$) δ: 7.25 (s, 5H, ArH), 5.31(m, 2H, $ArCH_2$), 4.45 (m, 1H, H2), 4.23 (m, 1H, H4), 3.65 (m, 2H, H5A,B), 2.30-2.55 (m, 2H, H3), 1.38 and 1.28 (s, 9H, $C(CH_3)_3$).

2S,4*R*-Azido-N$^{\alpha}$-(*t*-butoxycarbonyl)proline benzyl ester 27: Compound **26** (1.7 g, 7 mmol) was taken with *p*-toluenemethylsulfonate (1.65 g, 8 mmol), PPh_3 (2.3 g, 8.7 mmol) and DEAD (1.55 g, 8.9 mmol) in THF (60 mL) and stirred at RT for 6 hr.[22] The resulting mixture was concentrated and purified by silicagel column chromatography and the fractions containing the tosylate were collected. The solvent was removed by evaporation and the product (1.8 g, 4.6 mmol) and stirred with NaN_3 (1.05 g, 15 mmol) in DMF to obtain compound **27** (1.62 g, 92%).

^{1}H NMR, ($CDCl_3$) δ: 7.35 (s, 5H, Ar<u>H</u>), 5.15 (m, 2H, Ar<u>CH</u>$_2$), 4.45 (mi) and 4.34 (ma) (2 x t, 1H, J = 7.5 Hz, <u>H</u>2), 4.13 (brm, 1H, <u>H</u>4), 3.66 (m, 1H, <u>H</u>5A) 3.42 (mi, dd, J=3.5, 11.4 Hz, <u>H</u>5B), 2.30 (m, 1H, <u>H</u>3A), 2.13 (m, 1H, <u>H</u>3B) 1.48 & 1.42 (s, 9H, C(<u>CH</u>$_3$)$_3$).

2*S*,4*R*-Azido-N$^{\alpha}$-(bromoacetyl)proline benzyl ester 28: Compound **27** (3.5 g) was treated with 50% TFA in DCM for 1 h and the residue obtained after evaporation of solvent was dissolved in dioxane (15 mL) containing 10% aq. Na_2CO_3 (25 mL, pH 8.0). The mixture was cooled in ice bath and bromoacetyl chloride was added in two portions (2.5 eq. each).[25] The pH was then adjusted to 8.0 by addition of aq. Na_2CO_3. The mixture was concentrated to half its volume and extracted with DCM (50 mL x 4). The dried organic layer upon concentration and column chromatographic purification gave **28** (3.1 g, 80%) as a colorless oil.

^{1}H NMR, ($CDCl_3$) δ: 7.37 (s, 5H, Ar<u>H</u>), 5.16 (m, 2H, Ar<u>CH</u>$_2$), 4.72 (mi) and 4.63 (ma) (dt, 1H, J = 7.0 Hz, <u>H</u>2), 4.30 (m, 1H, <u>H</u>4), 3.92 (m, 1H, <u>H</u>5A), 3.65 (m, 1H, <u>H</u>5B), 3.81 (s, 2H, C<u>H</u>$_2$Br), 2.36 (m, 1H, <u>H</u>3A), 2.18 (m, 1H, <u>H</u>3B). MS: (M+1) = 368

2*S*,4*R*-Azido-N$^{\alpha}$-(thymin-1-yl acetyl)proline benzyl ester 29: A mixture of **28** (1.25 g, 0.3 mmol), thymine (0.6 g, 0.3 mmol) and solid K_2CO_3 (0.7 g, 0.3 mmol) in dry DMF (20 mL) was stirred at 25 C for 3 h. The residue obtained after aqueous work up was purified by column chromatography using DCM-MeOH (0-10%) as eluant to obtain **29** (yield 90%), as a white solid.

^{1}H NMR, ($CDCl_3$) δ: 9.45 (brs, 1H, NH), 7.38 (mi) and 7.32 (ma) (s, 5H, Ar-H), 7.05 (ma) and 6.85 (mi) (s, 1H, *T*-CH), 5.20 (m, 2H, $ArCH_2$), 4.84 (mi, q, J = 5.8, 7.8 Hz, H-2), 4.65 (ma, q, J = 13.7, 2.0 Hz, H2), 4.60 (d, 1H, J=16.2, *T* CHA), 4.37 (d, 1H, J=16.2 Hz, *T* CHB), 4.35 (ma) and 4.20 (mi) (m, 1H, H4), 3.92 (dd, 1H, J = 11.6, 5.8 Hz, H5A), 3.70 (dt, J = 11, 4.0 Hz, 1H, H5B), 2.35 (m, 2H, H3A,B), 1.95 (s, 3H, *T* CH_3). ^{13}C NMR, ($CDCl_3$) δ: 170.8 (COO), 166.2 (CONH), 156.1 (*T*-C2), 151.6 (*T*-C4), 141.3 (*T*-C6), 136.0, 128.8, 128.4, 128.0 (all Ar-C), 111.1 (*T*-C5), 67.2 (Ar-CH_2), 58.2 (C4), 52.4 (*T*-CH_2), 48.8 (C5), 34.6 (C3), 12.5 (*T*-CH_3).

2*S*,4*R*-Azido-N$^{\alpha}$-(2-amino-6-chloro-purin-9-yl acetyl)proline benzyl ester 30: A mixture of 2-amino-6-chloro-purine (0.73 g, 4.3 mmol), compound **28** (1.1g, 4.3 mmol) and K_2CO_3 (0.46 g, 4.3 mmol) was taken in DMF (10 mL) and stirred at RT for 6 hr. The solution was filtered and the filtrate was evaporated to yield a foam which was purified by column chromatography to obtain **30** (1.28 g) as pure material in 96% yield.

M.P. = 93 C. ^{1}H NMR, ($CDCl_3$) δ: 8.00 (mi) and 7.80 (ma) (s, H8), 5.31 (br-s, 2H, NH_2), 5.23 (mi, m, $ArCH_2$), 5.11 (dd, J=12.2, $ArCH_2$), 4.78-4.62 (m, 1H, H2), 4.60 (mi) and 4.61 (ma, t, J = 7.6 Hz, NCH_2), 4.35 (ma) and 4.20 (mi, m, H4), 3.89 (ma) and 3.81 (mi, dd, 1H, J = 3.1, 12.9 Hz, H5A), 3.64 (dd, J = 3.3, 10.1 Hz, H5B), 2.40 (mi), and 2.37 (ma, dq, 1H, J = 14.0, 7.5 Hz, H3A), 2.17 (dq, 1H, J = 14, 7.5 Hz, H3B). ^{13}C NMR (DMSO-d_6) δ: 170.2 (COO), 164.4 (>NCO), 159.1 (*Pu*-C2), 153.5 (Pu C4), 150.5 (Pu C6), 142.8 (Pu C8), 134.7, 128.2-127.4 (Ar C), 123.4 (Pu C5), 66.5 (Ar-CH_2), 59.2 (C2), 57.5 (CH_2), 50.9 (C4), 44.0 (C5), 33.7 (C3). $[\alpha]_D^{20}$ = -42.5 (c = 0.4, MeOH).

2*S*,4*R*-Azido-N$^{\alpha}$-(adenin-9-yl acetyl)proline benzyl ester 31: To NaH (0.04 g, 2.6 mmol) in dry DMF, adenine (0.34 g, 2.6 mmol) was added and stirred at 75 C for 15 min till the effervescence ceased. The flask was cooled, bromo compound **28** (0.65 g, 2.6

mmol) in DMF (5 mL) was added and the mixture stirred at 75 C for 1hr. DMF was removed under vacuum and the residue was taken in water, extracted with DCM and purified by column chromatography using MeOH/DCM mixture to obtain the pure compound **31** in 36% (0.25 g) yield.

^{1}H NMR, ($CDCl_3$) δ: 8.31 (ma) and 8.30 (mi) (s, 1H, *A* H8), 7.93 (ma) and 7.81 (mi) (s, 1H, *A* H2), 7.38 (mi) and 7.33 (ma) (s, 5H, ArH), 6.36 (ma) and 6.26 (mi) (brs, 2H, *A* NH_2), 5.25 (mi) and 5.15 (ma) (2 x dd, J =12.5 Hz, 2H, ArCH_2), 4.95 (ma, s, *A* CH_2) and 4.92 (mi, t, J = 7.5 Hz, *A* CH_2), 4.65 (t, 1H, J = 9 Hz, H2), 4.35 (ma) and 4.18 (mi) (m, 1H, H4), 3.92 (q, 1H, J = 5.2, 10 Hz, H5A), 3.71 (q, 1H, J=2.5, 10 Hz, H5B), 2.35 (m, 2H, H3A,B). ^{13}C NMR, ($CDCl_3$) δ: 170.6 (COOBn), 165.2 (mi) and 164.7 (ma) (>NCO), 155.5 (*A* C-6 & NCOO), 152.7 (*A* C2), 149.7 (*A* C4), 141.1 (*A* C8), 134.9-128.9 (all Ar-C), 118.5 (*A* C5), 68.02 (mi) and 67.1 (ma) ($ArCH_2$), 59.3 (C2), 57.9 (ma) and 57.2 (mi) (C4), 51.3 (*A* CH_2), 44.3 (ma) and 43.8 (mi) (C5), 36.4 (mi) and 34.08 (ma) (C3).

$[\alpha]_D^{20}$ = - 44.0 (c = 0.3, MeOH). [M+1]= 421.

2*S*,4*R*-azido-N$^{\alpha}$-(cytosin-1-yl acetyl)proline benzyl ester 32: Reaction of cytosine (0.75g, 6 mmol) in DMF with NaH (0.09 g, 5.9 mmol) and compound **28** (1.47 g, 5.8 mmol) following the above procedure gave **32** in 50% (0.38 g) yield after column purification.

M.P. = 85 C, ^{1}H NMR, ($CDCl_3$+D_2O) δ: 7.45-7.15 (m, 6H, ArH and *C* H6), 5.9(d, J = 8.0 Hz, *C* H5), 4.95-5.53 (m, 2H, $ArCH_2$), 4.82-4.45 (overlapping multiplets, 3H, N-CH_2, H2), 4.35 (brm, 1H, H4), 3.95 (brm, 1H, H5B), 3.70 (brm, 1H, H5A), 2.05-2.55 (overlapping multiplet, 2H, H3A,B). ^{13}C NMR, ($CDCl_3$) δ: 171.3 (COO), 166.7 (NCO), 155.0 (C2), 147.7 (*C*-C6), 136.2 (C5), 129.2-128.0 (Ar-C's), 94.6 (C4), 66.6 (Ar-CH_2), 60.0 (C2), 58.1 (C4), 52.0 (CH_2), 50.1 (C5), 34.4 (C3). $[\alpha]_D^{20}$ = -51.3 (c = 0.3, MeOH).

2*S*,4*R*-N-(*t*-butoxycarbonyl)-N$^{\alpha}$-(thymin-1-yl acetyl) proline 33: Compound **29** (2 g, 5.4 mmol) in MeOH was hydrogenated under pressure (30 psi) using 10%Pd-C (200 mg) as catalyst for 15 h. The catalyst was filtered off and the free amino acid product was isolated by evaporation of solvent. This compound (1.4 g, 5 mmol) was dissolved in dioxane:water (1:1, 20 mL), treated with Boc-azide (0.8 mL) and stirred at 25 °C for 24

h. with pH maintained at 9.0 by addition of 4N NaOH. The reaction mixture was concentrated to 10 mL, neutralized with Dowex $50H^+$ and filtered. Removal of solvent from the filtrate afforded the required product which was purified by crystallization from MeOH.

^{1}H NMR, (D_2O), δ: 7.39 (d, J = 1.2 Hz, *T* H6), 4.70 (t, J = 12.1 Hz, N-CH_2), 4.52 (t, J = 9.2 Hz, H2), 4.36 (m, 1H, H4), 4.20 (mi) and 3.95 (ma) (dd, H5B), 2.25-2.50 (bt, 2H, H3A,B), 1.90 (s, 3H, *T* CH_3), 1.46 (s, 9H, $C(CH_3)_3$). ^{13}C NMR, (DMSO-d_6) δ: 175.6 (COOMe), 167.3 (NCO), 157.8 (*T*-C5), 152.7 (*T*-C4), 143.8 (*T*-C6), 111.4 (*T*-C5), 59.1 (OCH_3), 52 (C2), 50.6 (*T*-CH_2), 50.3 (C4), 49.8 (C5), 35.0 (C3), 12.3 (T-CH_3). $[\alpha]_D^{20}$ -14.7 (c = 0.3, MeOH).

2*S*,4*R*-N-(*t*-butoxycarbonyl)-N$^\alpha$-(A/G/C-yl acetyl)proline benzyl ester 34a-36: The 4*R*-azido compounds **29-31** (1.0 g) were individually dissolved in methanol (50 m L) and reacted with ammonium formate (4 eq) and 10% Pd-C (0.1 g) for 24 h, after which the catalyst was removed by filtration over celite. The filtrate was evaporated under vacuum and the residue dissolved in water:dioxane (1:1, 15 mL). This was treated with Boc-azide (1.5 eq) and TEA (1.5 eq) at 50 °C for 15 h, after which the products were isolated by chromatographic purification (yields 60-70%).

2*S*,4*R*-N-(*t*-butoxycarbonyl)-N$^\alpha$-(2-amino-6-chloropurin-9-yl acetyl)proline benzyl ester 34a: ^{1}H NMR, ($CDCl_3$) δ: 7.85 (ma) and 7.70 (mi) (s, 1H, *G* H8), 7.30 (m, 5H, ArH), 4.97-5.45 (overlapping signals, 4H, NH_2 and ArCH_2), 4.82 (dd, 2H, NCH_2), 4.65 (t, J = 4.2 Hz, H2), 4.40 (brm, 1H, H4), 4.05 (brm, 1H, H5A), 3.65 (brm, 1H, H5B), 2.45 (mi, m) and 2.22 (ma, t) (2H, H3A,B), 1.47 (s, 9H, $C(CH_3)_3$). $[\alpha]_D^{20}$ = -20.5 (c = 0.2, MeOH).

2*S*,4*R*-N-(t-butoxycarbonyl)-N$^\alpha$-(guanin-9-yl acetyl)proline benzyl ester 34b: A mixture of **34a** (0.5g, 1 mmol) and 1N aq. NaOH (2 mL) was stirred at RT for 2 h after which it was neutralized with ion-exchange resin, Dowex $50H^+$. MeOH (5 mL) was added, the resin was filtered off and the product that slowly precipitated was collected (yield 60%).

^{1}H NMR, (DMSO-d_6:D_2O) δ: 7.70 (s, 1H, *G* H8), 4.92 (q, 2H, J = 19.5 Hz, N-CH_2), 4.28 (q, 1H, J = 8.1, 4.8 Hz, H2), 4.18 (t, 1H, J = 7.2 Hz, H4), 3.84 (q, 1H, J = 9.6, 5.8 Hz, H5A), 3.40 (q, 1H, J = 9.6, 7.2, H5B), 2.10 (dq, 1H, H3A,B), 1.40 (s, 9H, C(CH_3)$_3$).

2*S*,4*R*-N-(*t*-butoxycarbonyl)-N$^\alpha$-(Adenin-9-yl acetyl)proline benzyl ester 35: ^{1}H NMR, ($CDCl_3$) δ: 8.29 (s, 1H, *A* H8), 7.82 (ma) and 7.72 (mi) (s, 1H, H2), 7.32-7.36 (m, 5H, ArH), 6.48 (brs, 2H, NH_2), 5.75 (brd, ma) and 5.52 (br, mi, 1H, NHBoc), 5.25 (mi) and 5.13 (ma) (dd, 2H, ArCH_2), 4.90 (dd, J = 15.6 Hz, 2H, *A* CH_2), 4.64 (t, 1H, J = 5.8 Hz, H2), 4.36 (br, 1H, H4), 3.96 (dd, 1H, H5A), 3.78 (brm, 1H, H5B), 2.50-2.05 (brm, 2H, H3A,B), 1.49 (s, 9H, C(CH_3)$_3$). ^{13}C NMR, ($CDCl_3$) δ: 171.4 (ma) and 171.1 (mi) (COOR), 165.5 (CONH), 155.4 (C6 and OCON<), 152.4 (*A* C2), 149.3 (*A* C4), 141.6 (*A* C8), 135.1-128 (Ar), 118.2 (*A* C5), 79.8 (CMe_3), 67.9 (mi) and 67.1 (ma) (ArCH_2), 57.8 (C2), 51.4 (C4), 50.0 (*A* NCH_2), 44.7 (C5), 34.8 (C3), 28.3 (C(CH_3)$_3$), $[\alpha]_D^{20}$= -20.7 (c = 0.3, MeOH).

2*S*,4*R*-N-(*t*-butoxycarbonyl)-N$^\alpha$-(cytosin-1-yl acetyl)proline benzyl ester 36: ^{1}H NMR, ($CDCl_3$+D_2O) δ: 7.25-7.40 (overlapping signals, ArH, *C* H6), 5.90 (d, J = 6.36 Hz, 1H, *C* H5), 5.10 (overlapping q, 2H, ArCH_2), 4.55-4.95 (brm, 2H, N-CH_2), 4.15-4.50 (brm, 2H, H3A,B), 1.45 (s, 9H, C(CH_3)$_3$). $[\alpha]_D^{20}$ = - 49.6 (c = 0.3, MeOH).

General Procedure for preparation of 4-azido-N$^\alpha$-(bromoacetyl)proline methyl ester 37, 38: The 4-azido Boc proline methyl esters (9, & 19) were deprotected with 50% TFA/DCM and the TFA was removed by repeated evaporation with ether under vacuum using KOH trap. The free amines were taken in 10% Na_2CO_3 in dioxane/water 1:1, and bromoacetylchloride (2.5 eq, each) was added in two lots. The starting material was consumed in 5 min and the reaction mixtures were brought to pH 8.0 by adding 10% aq. Na_2CO_3, and the required bromo compounds were extracted using DCM. Further column purification yielded the required materials **37** and **38** respectively in good yield (68-75%).

2*S*,4*S*-azido-N$^\alpha$-(bromoacetyl)proline methyl ester 37: ^{1}H NMR, ($CDCl_3$) δ: 4.72-4.45 (br-m, 2H, H2, H4), 3.80-3.55 (m, 2H, H5A,B), 3.82 (s, 3H, $COOCH_3$), 3.72 (s, 2H, $COCH_2Br$), 2.40-2.01 (m, 2H, H3A,B). ^{13}C NMR ($CDCl_3$), δ : 172.3 (COO), 165.9

(CON), 58.3 (O-CH_3), 55.7 (ma), 55.2 (mi) ($COCH_2Br$), 52.44 (mi) 52.38 (ma) (C2), 51.8 (C4), 37.62 (C5), 27.0 (C3). MS: (m/e) = 290.

2*R*,4*S*-azido-N^{α}-(bromoacetyl)proline methyl ester 38: ^1H NMR, ($CDCl_3$,) δ: 4.75-4.6 (br-m, 1H, H2), 4.45-4.35 (m, 1H, H4), 4.05-3.98 (m 1H, 5HA), 3.82(s, 2H, $COCH_2Br$), 3.76 (s, 3H, $COOCH_3$), 3.71-3.60 (m, 1H, H5B), 2.50-2.25 (m, 2H, H3A,B). ^{13}C NMR, ($CDCl_3$) δ : 172.2 (COO), 165.8 (CON), 58.4 (O-CH_3), 55.7 (ma), 55.2 (mi)($COCH_2Br$), 52.4 (C2), 51.8 (C4), 37.7 (C5), 27.2 (C3). MS: (M+1) = 291

4-Azido-N^{α}-(thymin-1-yl acetyl)proline methyl ester 39, 40: *General procedure:* A mixture of compound **37/38** (1.25 g, 0.3 mmol), thymine (0.6 g, 0.3 mmol) and solid K_2CO_3 (0.7 g, 0.3 mmol) in dry DMF (20 mL) was stirred at 25 °C for 3 h. The solvent was removed and the product **39/40** respectively was obtained after usual workup. *trans* Isomer **40** was purified easily by crystallizing from methanol (yield 90%). *cis* Isomer **39** was purified by column chromatography by elution with DCM-MeOH to obtain product as white solid.

2*S*,4*S*-Azido-N^{α}-(thymin-1-yl acetyl)proline methyl ester 39: ^1H NMR, ($CDCl_3$) δ : 9.2 (b, 1H, NH), 7.01 (s, 1H, *T*-6H), 4.72-4.63 (m, 1H, H2), 4.46 (d, 1H, J = 16 Hz, *T*CHA), 4.34-4.25 (m,1H, H4), 4.23 (d, 1H, J = 16 Hz, *T*CHB), 3.81-3.73 (m, 1H, H5A), 3.72 (s, 3H, OCH_3), 3.68-3.54 (m, 1H, H5B), 2.52-2.43 (m, 1H, H3A), 2.13-2.02 (m, 1H, H3B), 1.97 (s, 3H, *T*-CH_3), 1.45 (s, 9H, $C(CH_3)_3$).

2*R*,4*S*-Azido-N^{α}-(thymin-1-yl acetyl)proline methyl ester 40 : IR (neat): 3400, 2952, 2923, 2105, 1735, 1700, 1665, 1642, 1462 and 1441 cm^{-1}. ^1H NMR, ($CDCl_3$) δ : 6.95 (s, 1H, *T*H6), 4.48 (d, 1H, J=18, *T* CHA), 4.20 (d, 1H, J=18, *T* CHB), 4.76-4.65 (mi) & 4.48-4.38 (ma) (m, 1H, H2), 4.30-4.20 (ma) & 4.15-4.06 (mi) (m, 1H, H4), 3.85-3.75 (m, 1H, H5A), 3.71 (mi) & 3.63 (ma) (s, 3H, $COOCH_3$), 3.65-3.53 (m, 1H, H5B), 2.42-2.25 (mi) & 2.23-2.05 (ma) (m, 2H, H3A,B), 1.88 (s, 3H, *T*-CH_3). ^{13}C NMR, ($CDCl_3$) δ: 175.6 (COOMe), 167.3 (NCO), 157.8 (*T*-C5), 152.7 (*T*-C4), 143.8 (*T*-C6), 111.4 (*T*-C5), 59.1 (OCH_3), 52.0 (C2), 50.6 (*T*-CH_2), 50.3 (C4), 49.8 (C5), 35.0 (C3), 12.3 (*T*-CH_3).

N4-(Butoxycarbonyl)-N^{α}-(thymin-1-yl acetyl)proline methyl ester 41, 42: *General procedure*: The azido compounds **39**, **40** were converted to the corresponding amino compounds by hydrogenation using 10% Pd-C in methanol. The free amine was protected with Boc-azide to get the Boc derivative **41** and **42**, which were purified by column chromatography to obtain pure material.

2*S*,4*S*-N-(*t*-Butoxycarbonyl)-N^{α}-(thymin-1-yl acetyl)proline methyl ester 41: ^{1}H NMR, ($CDCl_3$) δ: 9.40 (br, 1H, NH), 7.06 (s, 1H, *T*H6), 5.60 (d, OCONH), 4.68 (d,1H, J=16, *T* CHA), 4.33 (d, 1H, J=16, *T* CHB), 4.52-4.42 (m, 2H, H4, H2), 3.78 (mi) & 3.73 (ma) (s, 3H, $COOCH_3$), 3.98-3.80 (m, 1H, H5A), 3.60-3.51 (m, 1H, H5B), 2.53-2.42 (m, 1H, H3A), 2.17-2.0 (m, 1H, H3B), 1.97 (s, 3H, T-CH_3), 1.49 (s, 9H, 3 x CH_3).

2*R*,4*S*-N-(Butoxycarbonyl)-N^{α}-(thymin-1-yl acetyl)proline methyl ester 42: ^{1}H NMR, ($CDCl_3$) δ: 7.05 (s, 1H, *T*H6), 4.80 (d,1H, J=18, *T* CHA), 4.10 (d, 1H, J=18, *T* CHB), 4.70-4.55 (m, 1H, H2), 4.50-4.42 (m, 1H, H4), 3.95-3.83 (m, 1H, H5A), 3.73 (s, 3H, $COOCH_3$), 3.65-3.55 (m, 1H, H5B), 2.30-2.05 (m, 2H, H3A,B), 1.88 (s, 3H, *T*-CH_3), 1.45 (s, 9H, 3 x CH_3). ^{13}C NMR ($CDCl_3$) δ: 175.6 (COOMe), 167.9 (>NCO), 157.8 (*T*-C5), 152.7 (*T*-C4, NCOO), 143.8 (*T*-C6), 111.4 (*T*-C5), 81.8 ($C(CH_3)_3$), 59.1 (OCH_3), 52.0 (C2), 50.6 (C3), 50.3 (C4), 49.8 (C5), 35.0 (C3), 28.6 ($C(CH_3)_3$), 12.3 (*T*-CH_3).

Hydrolysis of Boc-methyl esters of 43, 44 *General procedure*: The Boc-methyl esters of **41** and **42** were hydrolysed using cold 2N aq NaOH in methanol. Hydrolysis was complete within 5-15 min. The solution was neutralized using Dowex-$50H^{+}$until the pH of the solution was 7.0 and filtered. The filtrate was concentrated and the resultant pale yellow solid was precipitated from methanol/pet-ether mixture.

2*S*,4*S*-N-(*t*-Butoxycarbonyl)-N^{α}-(thymin-1-yl acetyl)proline 43: ^{1}H NMR, (D_2O) δ: 7.37 (s, 1H, *T*H6), 4.50 (d, 1H, J=18, *T* CHA), 4.10 (d, 1H, J=18, *T* CHB), 4.35-4.20 (m, 1H, H2), 4.07-3.95 (m, 1H, H4), 3.80-3.65 (m,1H, H5A), 3.60-3.45 (m, 1H, H5B), 2.88-2.75 (mi) & 2.72-2.50 (ma) (m, 2H, H3A), 2.35-2.15 (mi) & 2.15-2.0 (ma) (m, 2H, H3,B), 1.88 (s, 3H, *T*-CH_3), 1.45 (s, 9H, 3 x CH_3).

2*R*,4*S*-N-(*t*-Butoxycarbonyl)-N$^{\alpha}$-(thymin-1-yl acetyl)proline 44: ^{1}H NMR, (D_2O) δ: 7.37 (s, 1H, *T*-H6), 4.50 (d, 1H, J=18, *T*C-HA), 4.12 (d, 1H, J=18, *T*C-HB), 4.72-4.55 (m, 1H, H2), 4.51-4.42 (m, 1H, H4), 3.95-3.83 (m,1H, H5A), 3.73 (s, 3H, $COOCH_3$), 3.65-3.55 (m, 1H, H5B), 2.31-2.05 (m, 2H, H3A,B), 1.88 (s, 3H, *T*-CH_3), 1.45 (s, 9H, $(CH_3)_3$).

2*S*,4*S*-N(Carboxybenzyl)-N$^{\alpha}$(thymin-1-yl)proline 45: The methyl ester **11b** (100 mg) was hydrolysed with 2N NaOH and neutralised with Dowex 50H$^+$ to obtain the free acid **45** (72 mg).

^{1}H NMR, (DMSO-d_6) δ: 7.25 (s, 5H, Ar), 6.9 (s, 1H, *T*-H6), 5.01 (s, 2H, Ar), 4.50-4.07 (b-m, 4H, *T*C-H_2, H2, H4 overlapping multiplets), 3.80-3.65 (m,1H, H5A), 3.60-3.45 (m, 1H, H5B), 2.72-2.50 (bm, 2H, H3A), 2.16 (m, 2H, H3,B), 1.88 (s, 3H, *T*-CH_3). FAB MS(M+1) = 431.

2*S*,4*S*-Amino-N$^{\alpha}$(thymin-1-yl)proline methyl ester 46: The compound **11a** (200 mg) was deprotected by stirring with 20% piperidine/DMF solution (2 mL). The solvent was removed and desiccated over KOH.

^{1}H NMR, (D_2O) δ: 7.25 (s, 1H, *T*H6), 4.80 (d,1H, J=18, *T* CHA), 4.10 (d, 1H, J=18, *T* CHB), 4.70-4.55 (m, 1H, H2), 4.50-4.42 (m, 1H, H4), 3.95-3.83 (m, 1H, H5A), 3.73(s, 3H, $COOCH_3$), 3.65-3.55 (m, 1H, H5B), 2.50 (br-m, 1H, H3B), 2.04 (m, 1H, H3A) 2.30-1.88 (s, 3H, *T*-CH_3). MS(M+1) = 311.

2*S*,4*S*-N(Fluorenylmethoxycarbonyl)-N$^{\alpha}$(thymin-1-yl)proline 47: The methyl ester **21** (120 mg) was hydrolysed with 2N NaOH and neutralised with Dowex 50H$^+$ to obtain the free acid **45**(89 mg).

^{1}H NMR, (D_2O) δ: 7.76-7.26 (m, 8H, Fmoc), 7.27 (mi) & 7.25 (ma) (s, 1H, *T* H6), 5.98 (ma) (d, J = 11 Hz, C4 NH) & 5.30 (mi, br, C4 NH), 4.73-4.16 (m, overlapping signals, 7H, H2, H4, *T* CH_2, Fmoc CH and OCH_2), 4.0-3.57 (br-m, 2H, H5), 2.05 (m, 2H, H3A,B), 1.93 (s, 3H, *T* CH_3).

Dipeptide Nucleic acids 48 and 49: *General method*: The thymine monomers **45** and **47** (carboxyl component, 0.12 mmol) were condensed separately with the amino components

48 (30 mg, 0.1 mmol) in DMF (500 μL) in presence of DCC (25 mg, 0.12 mmol) and HOBT (15 mg, 0.1 mmol) at 25 °C for 4 h. The usual work up afforded the protected dipeptide products **48a** and **49a** respectively, which were quantitatively deprotected in one step using either H_2/Pd-C or piperidine:DMF to yield the free dipeptides **48b** and **49b** respectively.

48b: ^{1}H NMR, (D_2O), 500 MHz, δ: 7.38 (2H, *T* C<u>H</u>), 4.70-4.28 (8H, *T* C<u>H</u>$_2$ x 2, <u>H</u>4, <u>H</u>2 x 2), 4.16-3.67 (m, 4H, <u>H</u>5A,B x 2), 3.75 (s, 3H, OC<u>H</u>$_3$), 2.80-2.12 (m, 4H, <u>H</u>3A,B x 2), 1.88 (s, 6H, *T* C<u>H</u>$_3$ x 2). **UV** λ_{max} = 269 nm.

49b: ^{1}H NMR, (D_2O), 500 MHz, δ: 7.42,7.38 (2H, *T* C<u>H</u>), 4.80-4.40 (8H, *T* C<u>H</u>$_2$ x 2, <u>H</u>4, <u>H</u>2 x2), 4.20-3.70 (m, 4H, <u>H</u>5A,B x 2), 3.77, 3.73 (s, 3H, OC<u>H</u>$_3$), 2.80-2.12 (m, 4H, <u>H</u>3A,B x 2), 1.89,1.87 (s, 6H, *T* C<u>H</u>$_3$ x 2). **UV** λ_{max}=280 nm.

48a: ^{1}H NMR, (DMSO- d_6) δ: 11.30 (s, 2H, *T* N<u>H</u> x 2), 5.03 (s, 2H, CBz C<u>H</u>$_2$), 4.57-3.81 (m, 8H, <u>H</u>2, <u>H</u>4 x 2, *T* C<u>H</u>$_2$ x 2), 3.56 (s, 3H, OC<u>H</u>$_3$), 3.40-3.20 (br, 4H, <u>H</u>5A,B x 2 H), 2.30-2.55 (br, 4H, <u>H</u>3A,B x 2), 1.75 (s, 6H, *T* C<u>H</u>$_3$ x 2). FAB-MS: [M+1] = 723.

49a: ^{1}H NMR, (DMSO- d_6) δ: 7.95-7.25 (m, 10H, Fmoc-Ar, *T* C<u>H</u>), 4.67-4.16 (m, 10H, <u>H</u>2, <u>H</u>4 x 2, Fmoc C<u>H</u>, C<u>H</u>$_2$, *T* C<u>H</u>$_2$ x 2), 3.91-3.54 (m, 4H, <u>H</u>5A,B x 2), 3.70 (s, 3H, OC<u>H</u>$_3$), 2.65-2.20 (br, 4H, <u>H</u>3A,B x 2), 1.82,1.80 (s, 6H, *T* C<u>H</u>$_3$ x 2). FAB-MS: [M+1] = 811.

N1-(*t*-Butoxycarbonyl)-1,2-diaminoethane 50:

1,2-Ethylenediamine (20 g, 0.33 mol) was taken in dioxane:water (500 mL, 1:1), and cooled in an ice bath. Boc azide (5g, 35 mmol) in dioxane (50 mL) was slowly added with stirring and the pH was maintained at 10.0 by continuous addition of 4N NaOH. The mixture was stirred for 8 hr. and the resulting solution was concentrated to 100 mL. The N1,N2-diboc derivative not being soluble in water, precipitated, and it was removed by filtration. The corresponding N1-monoBoc derivative **45** was obtained by repeated extraction from the filtrate with ethylacetate. Removal of solvents yielded the monoBoc ethylenediamine **50** (3.45 g, 63%).

^{1}H NMR, ($CDCl_3$) 90 MHz δ: 5.21 (b, 1H, NH), 3.32 (t, 2H, J= 8 Hz,), 2.54 (t, 2H, J=8 Hz), 1.42 (s, 9H).

Ethyl N-(2-*t*-Butoxycarbonyl aminoethyl)glycinate 51: The N1-mono-Boc-1,2-diaminoethane **50** (3.2 g, 20 mmol) was treated with bromoethyl acetate (2.25 mL, 20 mmol) in acetonitrile (100 mL) in presence of K_2CO_3 (2.4 g, 20 mmol) and the mixture was stirred at ambient temperature for 5 h. The solid that separated was removed by filtration and filtrate was evaporated to obtain **51** (4.3 g, 83%) as a colorless oil.

^{1}H NMR, ($CDCl_3$) δ: 5.02 (br, 1H,), 4.22 (q, 2H, J=8 Hz,), 3.35 (s, 2H,), 3.20 (t, 2H, J=6 Hz,), 2.76 (t, 2H, J=6 Hz,), 1.46 (s, 9H), 1.28 (t, 3H, J=8,)

Ethyl N-(*t*-Butoxycarbonyl aminoethyl)-N-(bromoacetyl)glycinate 52: Compound **51** (4.0 g, 14 mmol) was taken in 10% aq. Na_2CO_3 (75 mL) and dioxane (60 mL). Bromoacetyl chloride (6.5 mL, 0.75 mol) was added in two shots with rigorous stirring.[25] The reaction was complete within 5 min. The reaction mixture was brought to pH 8.0 by adding more 10% aq. Na_2CO_3 and concentrated to remove dioxane.[22] The product was extracted from the aqueous layer with DCM and was purified by column chromatography to obtain **52** as a colorless oil in good yield (4.2 g, 80%).

^{1}H NMR ($CDCl_3$) δ: 5.45 (br, 1H), 3.28 (q, 2H, J = 8 Hz), 4.14 (s, 2H), 4.00 (s, 2H), 3.53 (t, 2H), 3.28 (q, 2H), 1.46 (s, 9H), 1.23 (t, 3H, J = 8 Hz). [M+1]= 380.

N-(*t*-Butoxycarbonyl aeg)thymine Ethyl Ester 53a: Compound **52** (4.0 g, 11.6 mmol) was stirred with anhydrous K_2CO_3 (1.56 g 11.8 mmol) in DMF with thymine (1.4 g, 11.2 mmol) to obtain the desired compound in good yield. DMF was removed under reduced pressure and the oil obtained was purified by column chromatography.

^{1}H NMR, ($CDCl_3$) δ: 9.00 (br, 1H, NH), 7.05 (mi) 6.98 (ma) (s, 1H, *T*-CH), 5.65 (ma), 5.05 (mi) (br, 1H, NH), 4.58 (ma), 4.44 (mi) (s, 1H, *T*-CH_2), 4.20 (mi), 4.05 (ma) (s, 1H, *T*CH_2), 4.25 (m, 2H, OCH_2), 3.55 (m, 2H), 3.36 (m, 2H), 1.95 (s, 3H, *T*-CH_3), 1.48 (s, 9H), 1.289 (m, 3H). ^{13}C NMR, ($CDCl_3$) δ: 170.8, 169.3, 167.4, 164.3, 156.2, 151.2, 141.1, 110.2, 79.3, 61.8, 61.2, 48.5, 48.1, 47.7, 38.4, 28.1, 13.8, 12.2.

N-(*t*-Butoxycarbonyl aeg)-N4-(benzyloxycarbonyl)cytosine Ethyl Ester 54a: A mixture of NaH (0.25 g, 6.2 mmol) and N4-(bezyloxycarbonyl)cytosine (1.24 g, 6.2 mmol) was taken in DMF and stirred at 75 °C till the effervescence ceased. The mixture was cooled and **52** (2.0 g, 6.1 mmol) was added to obtain cytosine monomer **54a** (1.62 g, 50%) in moderate yield.

^{1}H NMR, ($CDCl_3$) δ: 7.65 (d, 1H, J=8, H6), 7.35 (s, 5H, Ar), 7.25 (d, 1H, J=8, H5), 5.70 (br, 1H, NH), 5.20 (s, 2H, $ArCH_2$), 4.71 (ma), & 4.22 (br-s, 2H), 4.15 (q, 2H), 4.05 (s, 2H), 3.56 (m, 2H), 3.32 (m, 2H), 1.48 (s, 9H), 1.25(t, 3H).

N-(*t*-Butoxycarbonyl aeg)Adenine Ethyl Ester 55a: NaH (0.246 g, 6.1 mmol) was taken in DMF (15 mL) and adenine (0.8 g, 6.1 mmol) was added. The mixture stirred at 75 °C till the effervescence ceased and the mixture was cooled before adding **52**(2.0 g, 6.1 mmol). The reaction mixture was heated at 75 °C for 1 h and DMF was removed. The thick oil was taken in water and extracted with ethylacetate and purified by column chromatography to obtain the product **55a**.

^{1}H NMR, ($CDCl_3$) δ: 8.32 (s, 1H), 7.95 (mi), 7.90 (ma) (s, 1H), 5.93 (ma) & 5.80 (mi), (br, 2H), 5.13 (ma) & 4.95 (mi), 4.22 (mi) & 4.05 (ma) (s, 2H), 4.20 (m, 2H), 3.65 (ma) & 3.55 (mi), (m, 2H), 3.40 (ma) & 3.50 (mi),(m, 2H), 1.42 (s, 9H), 1.25(m, 3H).

N-(*t*-Butoxycarbonyl aeg)-2-amino-6-chloropurine Ethyl Ester 56: A mixture of 2-amino-6-chloropurine (1.14 g, 6.8 mmol), K_2CO_3 (0.932 g, 7 mmol) and **52** (2.4 g, 7 mmol) were taken in DMF (20 mL) and stirred at RT for 4 h, K_2CO_3 was removed by filtration and DMF was removed under reduced pressure. The resulting residue was purified by column chromatography to obtain **56** in excellent yield (2.65 g, 98%).

^{1}H NMR, ($CDCl_3$) δ: 7.89 (mi) & 7.85 (ma) (s, 1H), 7.30 (s, 1H), 5.80 (br, 1H), 5.18 (br, 2H), 5.02 (ma) & 4.85 (mi)(s, 2H), 4.18 (mi) & 4.05 (ma) (s, 2H), 3.65 (ma) & 3.16 (mi), (m, 2H), 3.42 (ma) & 3.28 (mi) (m, 2H), 1.50 (S, 9H), 1.26 (m, 3H).

N-(*t*-Butoxycarbonyl aeg)-2-amino-purine Ethyl Ester 57a: Hydrogenation of **56** with 10% Pd-C, at 40 psi H_2 pressure for 6-8 h gave the required compound **57a** in good yield after column purification (80%).

^{1}H NMR, ($CDCl_3$) δ: 8.65 (s, 1H), 7.93 (mi) & 7.90 (ma) (s, 1H), 5.80 (br, 1H), 5.40 (br 1H), 5.00 (ma) & 4.85 (mi) (s, 2H), 4.17 (mi) & 4.05 (ma) (s, 2H), 4.25 (m, 2H), 3.67 (ma) & 3.56 (mi) (m, 2H), 3.4 (ma), & 3.32 (mi) (m, 2H), 1.45 (s, 9H), 1.25 (m, 3H). ^{13}C NMR ($CDCl_3$) δ: 169.8 (mi) & 169.6 (ma) (COO), 167.3 (ma) & 166.9 (mi) (CON<), 159.9 (*A*C2), 156.1 (OCON), 153.4 (*A*-C4), 148.6 (*A*-C6), 143.7 (*A*-C8), 126.8 (*A*-C5), 79.7 (ma) & 79.2(mi) ($C(CH_3)_3$), 62.13 (mi) & 61.4 (ma) (OCH_2), 48.47, 43.0, 38.65(all CH_2), 28.25[($CH_3)_3$], 13.91 (OCH_2CH_3).

Hydrolysis of ethyl ester of PNA monomers 53a-55a, 57a: *General method:* The ethyl esters **48a-52a** (2g) were hydrolysed using 2N aq NaOH (5 mL) in methanol (5 mL) and the resulting acid was neutralized with activated Dowex-50H$^+$ till the pH of the solution was 7.0. The resin was removed by filtration and the filtrate was concentrated to obtain the resulting Boc-protected acid in excellent yield (>85%).

N-(2-*t*-Butoxycarbonyl-aminoethyl)-N-(thymin-1-yl)glycine 53b: ^{1}H NMR, (D_2O) δ: 7.28 (s, 1H), 4.80 (merged with HDO) (ma) & 4.72 (mi), 4.28 (mi) & 4.15 (ma) (s, 2H), 4.20 (m, 2H), 3.55 (m, 2H), 3.34 (ma) & 3.20 (mi) (m, 2H), 1.90 (s, 3H), 1.45 (s, 9H). ^{13}C NMR, (D_2O) δ: 167.5, 156, 151.3, 142.4, 108.7, 78.5, 48.05, 47.4, 41.0, 38.5, 28.5 & 12.2.

N-(2-*t*-Butoxycarbonyl-aminoethyl)-N-(N^{4}benzyloxycarbonyl cytosin-1-yl)glycine 54b: ^{1}H NMR (D_2O) δ: 7.89 (ma) & 7.87 (mi) (d, 1H, J= 8), 7.41 (m, 5H, Ar), 7.18 (ma) & 7.14 (mi) (d, 1H, j=8), 5.19 (s, 2H, Ar), 4.78 (merged with HDO) (ma), 4.72(mi) & 4.28 (mi), 4.17 (ma) (s, 2H), 4.23 (m, 2H), 3.57 (m, 2H), 3.34 (ma), 3.20 (mi) (m, 2H), 1.45 (s, 9H). ^{13}C NMR, (D_2O) δ: 163.1, 154.9, 153.2, 150.7, 136.0, 128.5, 128.1, 127.9, 93.9, 66.5. 48.7, 47.6, 41.2, 38.5, 28.5,

N-(2-*t*-Butoxycarbonyl-aminoethyl)-N-(adenin-9-yl)glycine 55b: ^{1}H NMR, (D_2O) δ: 8.32 (s, 1H), 8.23 (s, 1H), 5.38 (ma) & 5.2 (mi), 4.2 (ma) & 4.03 (mi)(s, 2H), 3.65 (ma) & 3.5 (mi) (t, J=7Hz, 2H), 3.41 (ma) & 3.20 (mi) (t J=7Hz, 2H), 1.40 (ma) & 1.28 (mi) (s, 9H). ^{13}C NMR, (D_2O) δ: 175, 168.9, 158.4, 152.3, 149.6, 148.1, 145.2, 118.3, 81.6, 51.8, 50.2, 49.0, 48.6, 45.7, 45.5, 38.5 & 28.4.

N-(2-*t*-Butoxycarbonyl-aminoethyl)-N-(2-amino purin-9-yl)glycine 57b: M.P. = 207 °C, ^{1}H NMR, (D_2O) δ: 8.70 (s, 1H), 8.33 (s, 1H), 5.31 & 5.12 (s, 2H), 4.17 & 4.07 (s, 2H), 3.70 (t, 2H), 3.58 (t, 2H), 3.40 (t, 2H), 3.23 (t, 2H), 1.42 (d, 9H). ^{13}C NMR, (DMSO-d_6), δ: 171.6 (mi) & 171.07 (ma) (C=OO), 167.3 (ma) & 166.9 (mi)(C=ON<), 159.9 (*A*-C2), 156.1 (OC=ON), 153.4(-*A*C4), 148.6(*A*-C6), 143.7(*A*-C8), 126.8(*A*-C5), 78.7 (C=$(CH_3)_3$), 50.2, 48.9, 43.0, 38.65 (all NC=H_2), 28.25 (C($C=H_3$)$_3$).

2.8 References

1. Nielsen P.E.; Egholm, M.; Berg, R.H.; Buchardt, O. *Science*, **1991**, *254*, 1497.
2. Harvey, J.C.; Peffer, N.J.; Bisi, J.E.; Thomson, S.A.; Cadilla, R.; Josey, J.A.; Ricca, D.J.; Hassman, C.F.; Bonham, M.A.; Au, K.G.; Carter, S.G.; Bruckenstein, D.A.; Boyd, A.L.; Noble, S.A.; Babiss L.E. *Science*, **1992**, *258*, 1481.
3. Nielsen, P.E.; Egholm, M.; Berg, R.H.; Buchardt, O. *Anticancer. Drug Design.* **1993**, *8*, 53.
4. Hyrup B. Egholm, M.; Rolland, M.; Nielsen P.E.; Berg, R.H.; Buchardt, O. *J. Chem. Soc., Chemical commun.*, **1993**, 518.
5. Hyrup B.; Egholm, M.; Nielsen, P.E.; *J. Am. Chem. Soc.* **1994**, *116*, 7964.
6. Egholm, M.; Buchardt, O.; Chiristensen, L.; Behrens, C.; Freier, S.M.; Driver, D.A.; Berg, R.H.; Kim, S.K.; Norden, B.; Nielsen, P.E. *Nature*, **1993**, *365*, 566.
7. Koch, T.; Naesby, M.; Wittung, P.; Jorgensen, M.; Larsson, C.; Buchardt, O.; Stanley, C.J.; Norden, B.; Nielsen, P.E. *Tetrahedron Lett.* **1995**, *36*, 6933.
8. Bergmann, F.; Bannwarth, W.; Tam, S. *Tetrahedron Lett.* **1995**, *36*, 6823.
9. Stetsenko, D.A.; Lubyako, E.N.; Potapov, V.K.; Azhikina T.L.; Sverdlov, E.D. *Tetrahedron Lett.* **1996**, *37*, 3571.
10. van der Laan, A.C.; Brill, R.; Kuimelis, R.G.; Kuyl-Yeheskiely, E.; van Boom, J.H.; Andrus, A.; Vinayak, R. *Tetrahedron Lett.*, **1997**, *38*, 2249.
11. Garner, P.; Yoo, J.U. *Tetrahedron Lett.*, **1993**, *34*, 1275.
12. Lenzi, A.; Reginato, G.; Taddei, M. *Tetrahedron Lett.*, **1995**, *36*, 1713, 1717.
13. Diederichsen, U.; Schmitt, H.W. *Tetrahedron Lett.*, **1996**, *37*, 475.
14. Dueholm, K. L.; Peterson, K.H.; Jensen, D.K.; Nielsen, P.E.; Egholm, M.; Buchardt, O. *Bioorg. Med. Chem. Lett.*, **1994**, *4*, 1077.
15. Kosynkina, L.; Wang, W.; Liang, T.C. *Tetrahedron Lett.*, **1994**, *35*, 5173.
16. Hyrup, B. Egholm, M., Buchardt, O.; Nielsen, P.E; *Biorg. Med. Chem. Lett.*, **1996.** *6*, 1083.
17. Gangamani B.P.; Kumar V.A.; Ganesh K.N.; *Tetrahedron*, **1996**, *52*, 15017.
18. Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. I.*, **1996,** 539.

19. Remuzon P. *Tetrahedron* **1996**, *52*, 13803.

20. Kaspersen F.M.; Pandit U.K. *J. Chem. Soc. Perkin Trans. I.* **1975,** 1617.

21. Bodanszky, M. and Bodanszky, A. *The Practice of Peptide Synthesis*, Springer - Verlag, 1984.

22. Dueholm, K.L.; Egholm, M.; Beherns, C.; Christensen, L.; Hansen, H.F.; Vulpius, T.; Petersen, K. H.; Berg, R.H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.*, **1994**, *59*, 5767.

23. Peterson, M.L.; Vince, R. *J. Med. Chem.*, 1991, *34*, 2787.

24. Stille, J.R.; Frietschel, S. J.; Baker, G.L. *J. Org. Chem.*, **1981**, *46*, 2954.

25. Meltzer, P.C.; Liang, A.Y.; Matsudaira, P. *J. Org. Chem.*, **1995**, *60*, 4305.

26. Chenon, M-T,; Pugnire, R.J.; Grant, D.M.; Panzica R.P.; Townsend, L.B. *J. Am. Chem. Soc.*, **1975**, *97*, 4627.

27. Farese A.; Patino, N.; Condom, R.; Dalleu, S.; Guedj, R.; *Tetrahedron lett.* **1996**, *37*, 1413.

28. Kim, S.H.; Nielsen, P.E.; Egholm, M.; Buchardt, O.; Berg, R.H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

CHAPTER 3

Solid Phase Synthesis of Prolyl Nucleic Acids and Their Biophysical Studies

3.1 Introduction:

Conventional drugs are designed and targeted against proteins. When the specificity is not high, the treatment requires a higher dosage which may lead to toxicity due to interaction with non-specific targets. Alternatively, it is more attractive to specifically inhibit the mRNA which produces the particular protein or the active gene that *trans*cribes the mRNA.[1] Oligonucleotides by recognizing the complementary DNA/RNA sequences cause antigene or antisense inhibition of gene or mRNA. In order to make the oligonucleotides stable towards nucleases, they have to be chemically modified.[2] The major drawback in modified oligonucleotides[3] is the large scale synthesis, poor cell uptake and non-specific binding to cellular enzymes.[4] In recent times, PNA a structural mimic of DNA in which the sugar phosphate backbone is replaced by 2-N-(aminoethyl)glycine linkages has emerged as a novel DNA analogue.[5,6] It has various advantages over other oligonucleotide modifications: (i) stability to cellular degradations, (ii) hybridization to complementary sequences with higher affinity and specificity, (iii) low non-specific interaction with cellular contents,[4] (iv) easy synthesis in large amounts by adopting solid phase peptide synthesis. A major limitation of PNAs from an application perspective are their poor solubility in aqueous medium due to intermolecular self-aggregation[7] and intramolecular self-organization[8] from base stacking. Further, being achiral, PNA binds to complementary DNA in both parallel and antiparallel modes.[9,10]

3.1.1 Rationale for the present work and objective

In the previous chapter, synthesis of prolyl nucleic acid monomers/dimers which are conformationally restricted and chiral analogues of PNA monomer is reported. The introduction of these chiral monomer analogues into PNA oligomer was envisaged to impart specific directionality in DNA/PNA binding and hence differentiate between parallel and antiparallel modes of binding. The conjugation of polycations such as spermine to oligonucleotides has been found to influence the physicochemical properties of DNA duplexes and triplexes.[11-13] The conjugation of PNA with a polyamine such as spermine may increase the solubility of the PNA and enhance the binding efficiency due to the

terminal cationic charges. In view of this, the following issues mainly concerned with the synthesis and evaluation of PNA analogues are addressed in this chapter.

(i) Syntheses of homopyrimidine oligomeric prolyl nucleic acids (PrNA) and mixed PNA-PrNA sequences.

(ii) Biophysical studies of PrNA, PrNA-DNA hybrids using UV, fluorescence and CD spectroscopic techniques.

3.2 Synthesis of homopyrimidine PrNAs and modified PNA

The synthesis of peptides using solid phase synthesis protocols has become a routine practice.[14] Using these methods and appropriate monomers PNA and modified PNAs have been synthesized. The ease of handling the solid phase synthesis and the high success has lead to synthesis of a large number of analogues to modulate their biophysical properties.

In the present work, the standard PNA monomers used have amino-function of eda unit protected as *t*-Boc and employ DCC/HOBT activation coupling strategy.[15] The use of Fmoc strategy has a drawback in PNA synthesis, as small amount of acyl migration has been observed under basic condition from the tertiary amide to the free amine formed during piperidine deprotection.[16] The *t*-Boc protected monomer units N4-(*t*-Boc)N$^{\alpha}$(thymin-1yl acetyl)prolines (**1-3**) were conveniently prepared on large scale as described in *Chapter II, section 2.4*. The readily available Merrifield resin was chosen as the polymeric matrix to build the model PrNA oligomers. After the first amino acid anchoring, it can be used for both *t*-Boc, as well as Fmoc strategy. The first amino acid is linked to the resin through benzyl ester linkage and an advantage of using this resin is that, it can either be cleaved with a strong acid to obtain the C-terminal carboxylic acid, or it can be cleaved with an amine to get a C-terminal amide.

All the peptide oligomers were synthesized on Merrifield resin, with β-alanine chosen as the first amino acid linker. The rationale for this is that β-alanine is achiral and hence does not interfere in the spectral properties of PrNA which carry chiral monomer units. As it has a short alkyl chain, its contribution to hydrophobic nature of the PNA is negligible. Merrifield resin was therefore functionalized with N-*t*-Boc-β-alanine following

the cesium salt method[17] to obtain the benzyl ester linkage between the resin and the first amino acid. The functionalized resin was assayed to determine the loading value of first amino acid by picric acid method.[18]

Solid Phase Synthesis of PrNA Analogues BOC Strategy

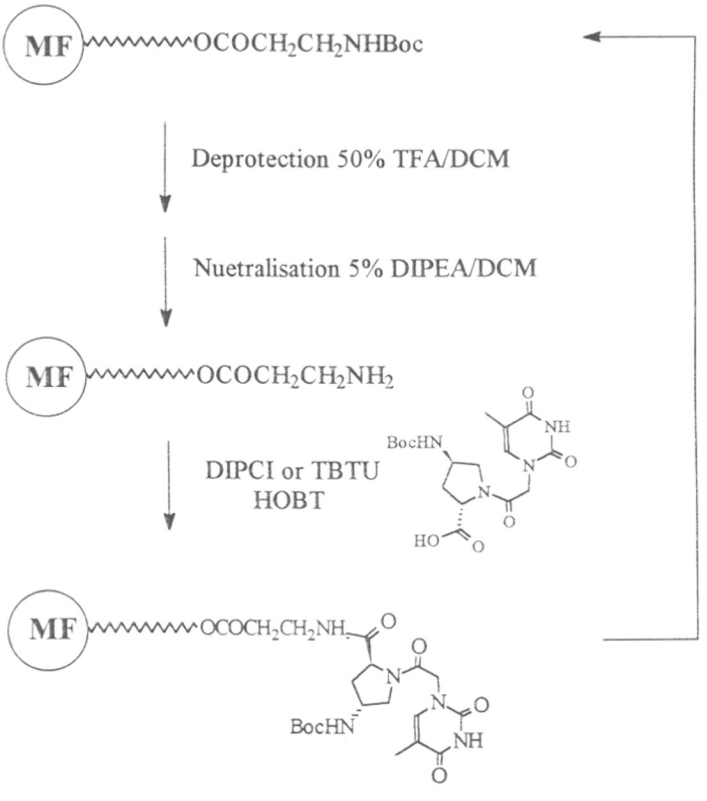


Figure 1

The stepwise synthesis of PNA oligomers was carried out in the C→N direction. The deprotection of N-terminal *t*-Boc group of β-alanine on the resin was accomplished by using 50% TFA/DCM to release the free amino function. The resin was neutralized with

5% DIPEA/DCM solution before the next coupling. The free amino terminus was coupled with 4 mole equivalents of *t*-Boc-monomers (**1-7**) in the presence of 2-(1H-benzotriazole)-1,1,3,3-tetrabutyl uronium tetrafluoroborate (TBTU) or diisopropyl carbodiimide (DIPCDI) in DMF, used along with 1-hydoxybenzotriazole (HOBT) as a racemization suppressant. The deprotection of *t*-Boc gives the free amine which gives a blue color with Kaiser's reagent. The completion of the coupling was verified by the free amino group assay using Kaiser's reagent,[19] which shows no change in color implying the completion of coupling. The cycle was repeated for each monomer unit to assemble the desired sequences as shown in the the flow diagram (Figure 1).

3.2.1 Synthesis of homopyrimidine Prolyl Nucleic Acids:

The synthesis of model homooligomers of unmodified PNA and its analogous PrNA war first carried out manually. The target was a PNA-T_6 which would be the control for the homooligomers of PrNA. The unmodified PNA-T_6 was assembled on Merrifield resin preloaded with β-alanine, following the standard *t*-Boc-strategy. TBTU was used as the coupling reagent and repeated cycles of the coupling gave MF-PNA-T_6-NH-*t*-Boc (**8**).

Each of the prolyl monomers **1**, **2**, and **3** in Figure 2 were oligomerized to their corresponding hexamers using *t*-Boc-strategy. Merrifield resin preloaded with *t*-Boc-β-alanine was taken, and the initial deprotection of the *t*-Boc group by TFA gave the free amine. The couplings were carried out in DMF using TBTU as the coupling agent and HOBT was added to suppress the racemization. Repeating the same set of reactions on the polymer support for each isomers gave the resins, MF(L-*cis*-Pro-T)$_6$-NH-*t*-Boc (**9**), MF(L-*trans*-Pro-T)$_6$-NH-*t*-Boc (**10**) and MF(D-*trans*-Pro-T)$_6$-NH-*t*-Boc (**11**). The capping step was not employed in the present synthetic strategy as the synthesis was carried out manually and coupling efficiency was very high as observed by HPLC (Figure 3). For the same reason the nucleobases containing exocyclic amino functions in 5 and 6 were used without protection.

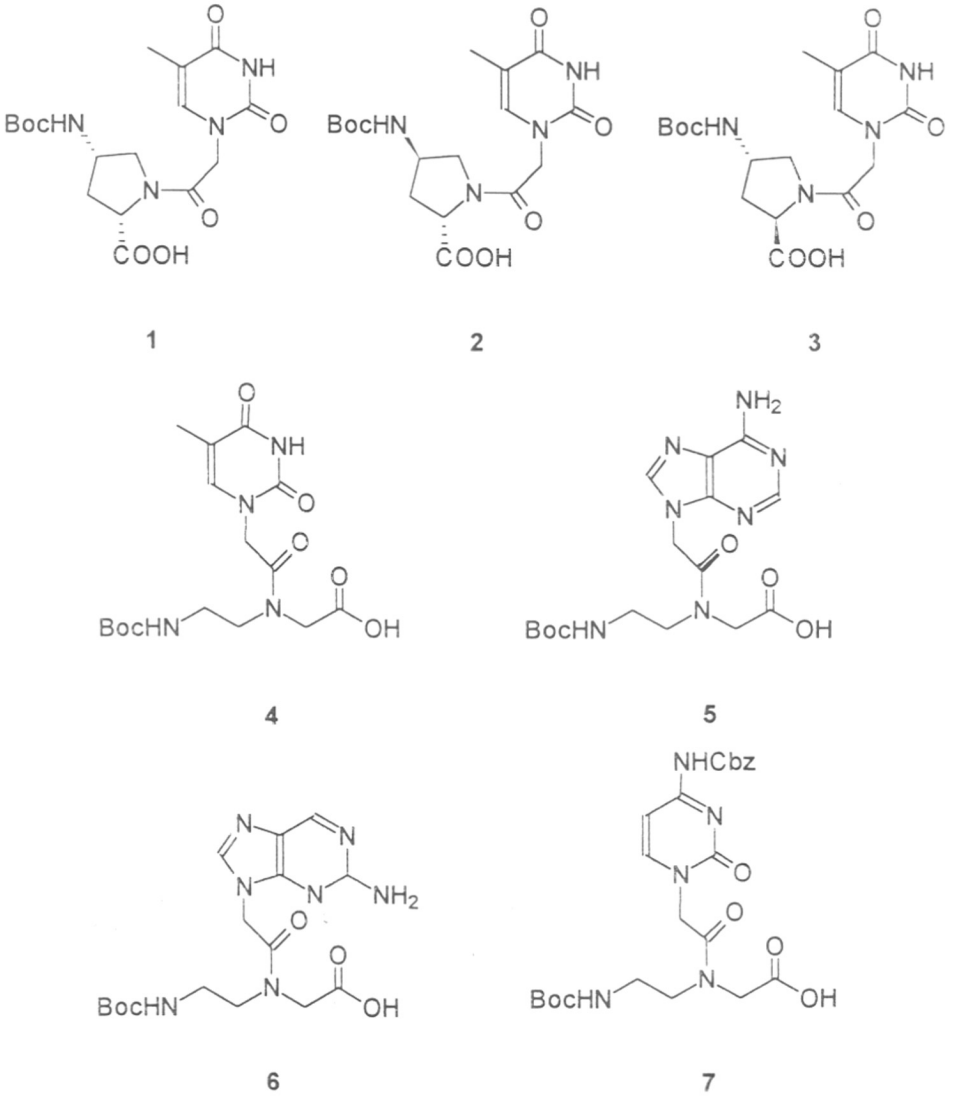


Figure 2

At the end of every cycle, the purity of the product and the efficiency of the coupling was estimated by removing a small portion of the resin at every cycle followed by cleavage of the oligomers from the resin with TFMSA/TFA and subjecting the oligomers to HPLC analysis. The Figure 3 shows representative HPLC chromatograms to

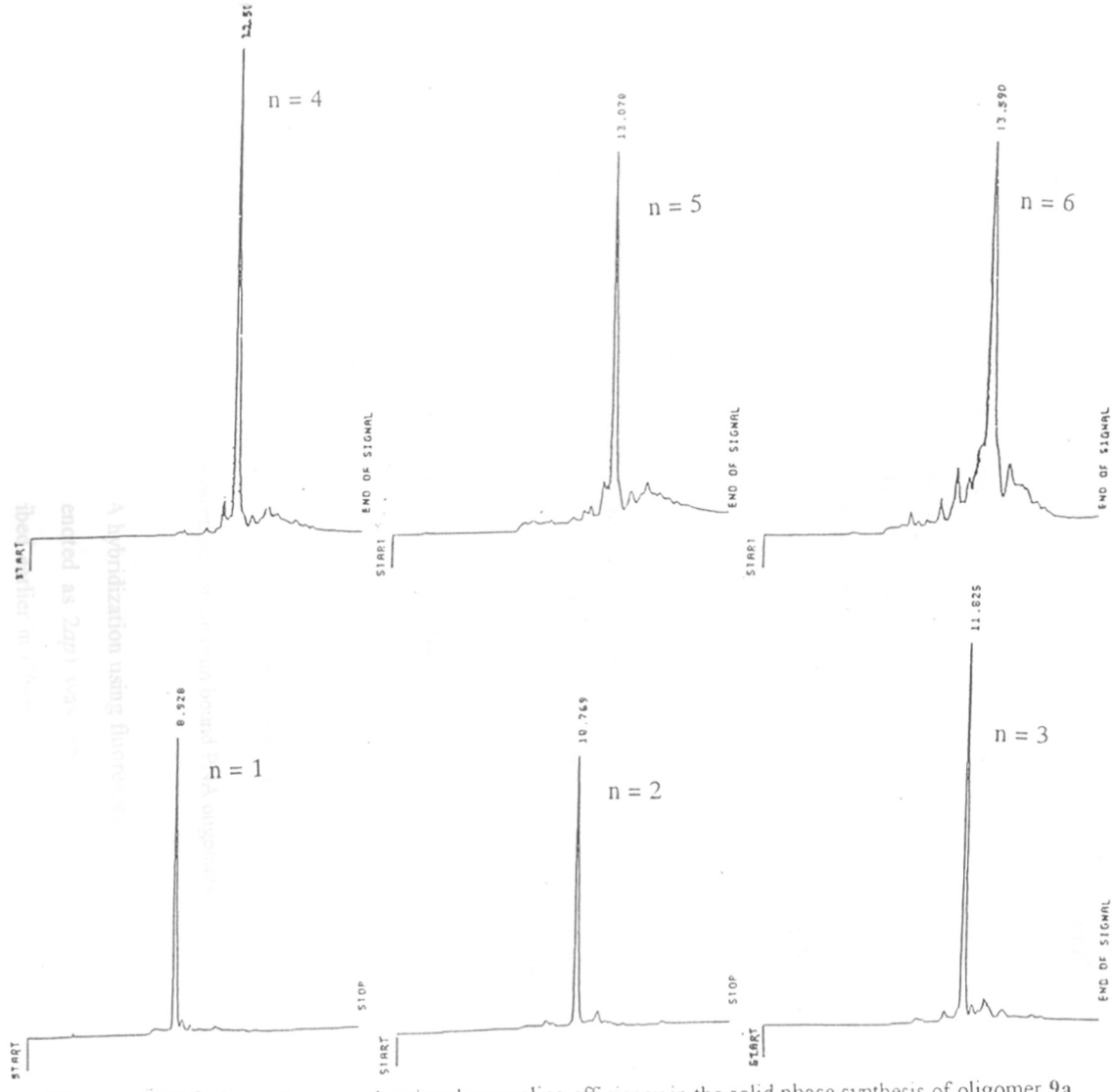


Figure 3: HPLC chromatograms showing the coupling efficiency in the solid phase synthesis of oligomer **9a**.

assay the efficiency after each coupling which was more than 90%. The addition of each monomer unit progressively increased the elution time of the product.

3.2.2 Synthesis of PNA carrying one prolyl monomer unit

The linking of chiral lysine at the C-terminus of PNA induces chirality.[9,20] To examine such effects induced by prolyl monomers, the synthesis of an unmodified 9-mer followed by addition of monomers **2** and **3** at the N-terminus as the 10th unit was undertaken. Merrifield resin with β-alanine spacer was used to prepare the required oligomers in a batch synthesis. The first five aeg PNA units required for (PNA-TTATT) were assembled stepwise using monomers **4** and **5** to obtain MF-β-ala-(aeg-TTATT) resin (**12**). A part of PNA-bound resin **12** was taken and extended by four more units to get MF-β-ala-(aeg-$T_2AT_2AT_2A$) resin **13**. The resin **13** was divided into three portions, which were separately coupled with monomers **4**, **1** and **3** to obtain the resin bound oligomers **14**, **15** and **16** respectively.

For synthesis of the PNA oligomers containing prolyl monomers in the middle of the sequence, the resin bound oligomers **14**, **15** and **16** were extended by 3 units on the solid phase by using aeg-C and aeg-T monomers to obtain the resins bound PNA oligomers **17**, **18** and **19** respectively. In this way PNA oligomers containing the enantiomeric prolyl nucleic acid units **2** (L-*trans*-Pro-T) and **3** (D-*trans*-Pro-T) at N-terminus as well as at the interior were obtained.

Schematic representation of the resin bound PNA oligomers.

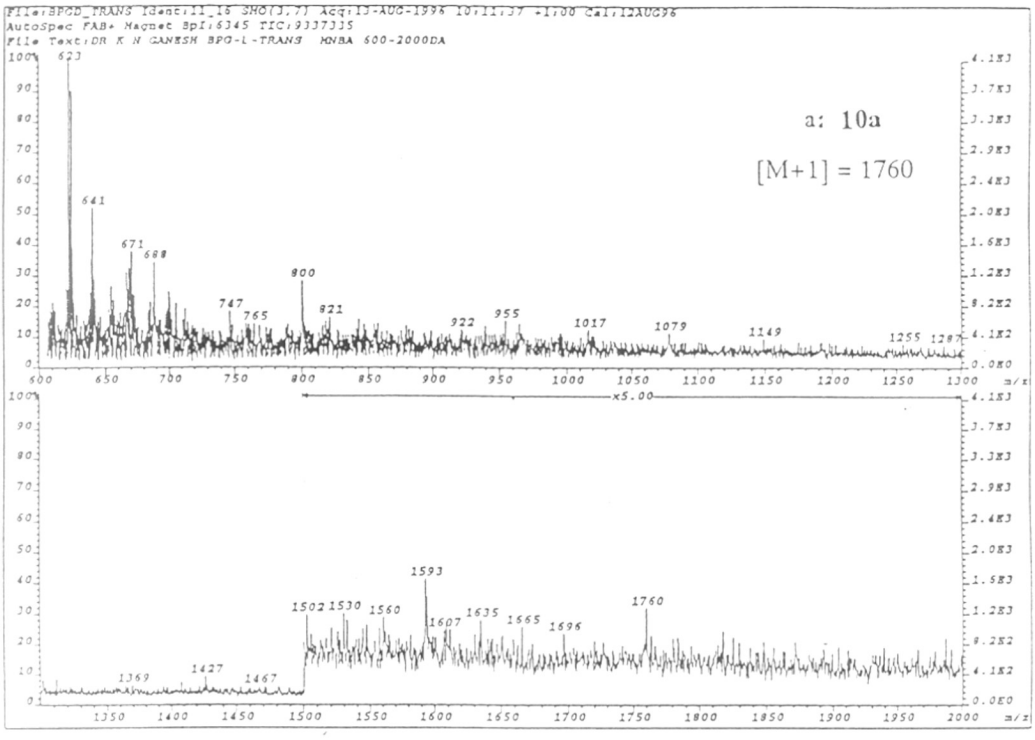
To monitor the PNA-DNA hybridization using fluorescence, the PNA containing a fluorophore 2-aminopurine **6** (denoted as 2*ap*) was synthesized. The synthesis of the corresponding monomer is described earlier in *Chapter II, section 2.5*. In this monomer the 2-amino of the purine needs no protection since it is unreactive to peptide coupling conditions. The synthesis of fluorescent MF(aeg-T_2AT_2A*T_2AT) resin (**20**) was achieved

by incorporation of the fluorescent base **6** (A*) in the PNA oligomer at 6th position of the 10-mer mixed sequence. The earlier prepared resin bound oligomer **12** was deprotected and coupled with *t*-Boc-aeg-2-aminopurine **6** using DIPCDI as the coupling reagent on the solid support followed by *t*-Boc deprotection and the continuation of the cycle using other bases (aeg-TTAT) to obtain the resin **20**.

3.2.3 Cleavage of PNA resins by TFMSA to obtain C-terminal carboxylic acid

The oligopeptides were cleaved from the resin (**8-11** and **14-20**) using strong acids like TFMSA/TFA[21] to obtain PNAs with terminal carboxylic acid **8a-11a** and **14a-20a** following in the standard protocol. The oligomers **8a-11a** and **14a-20a** containing the sequences are shown below. These oligomers were purified on semipreparative reverse phase C-8 FPLC column using Buffer condition 2 (*see section 3.6*) and the purity checked by analytical reverse phase C-18 HPLC and by molecular weight characterized by FAB MS. Some of the representative HPLC chromatograms of the purified oligomers are shown in Figure 5.

PNA	Sequences
8a	NH_2(aeg-T)$_6$-$CONHCH_2CH_2COOH$
9a	NH_2(L-*cis*-Pro-T)$_6$-$CONHCH_2CH_2COOH$
10a	NH_2(L-*trans*-Pro-T)$_6$-$CONHCH_2CH_2COOH$
11a	NH_2(D-*trans*-Pro-T)$_6$-$CONHCH_2CH_2COOH$
14a	NH_2(aeg-$TAT_2AT_2AT_2$)-$CONHCH_2CH_2COOH$
15a	NH_2(D-*trans*-Pro-T{aeg-$AT_2AT_2AT_2$})-$CONHCH_2CH_2COOH$
16a	NH_2(L-*trans*-Pro-T{aeg-$AT_2AT_2AT_2$})-$CONHCH_2CH_2COOH$
17a	NH_2(aeg-$CTCTAT_2AT_2AT_2$)-$CONHCH_2CH_2COOH$
18a	NH_2(aeg-CTC{D-*trans*-Pro-T}$AT_2AT_2AT_2$)-$CONHCH_2CH_2COOH$
19a	NH_2(aeg-CTC{L-*trans*-Pro-T}$AT_2AT_2AT_2$)-$CONHCH_2CH_2COOH$
20a	NH_2(aeg-$TAT_2A^*T_2AT_2$)-$CONHCH_2CH_2COOH$



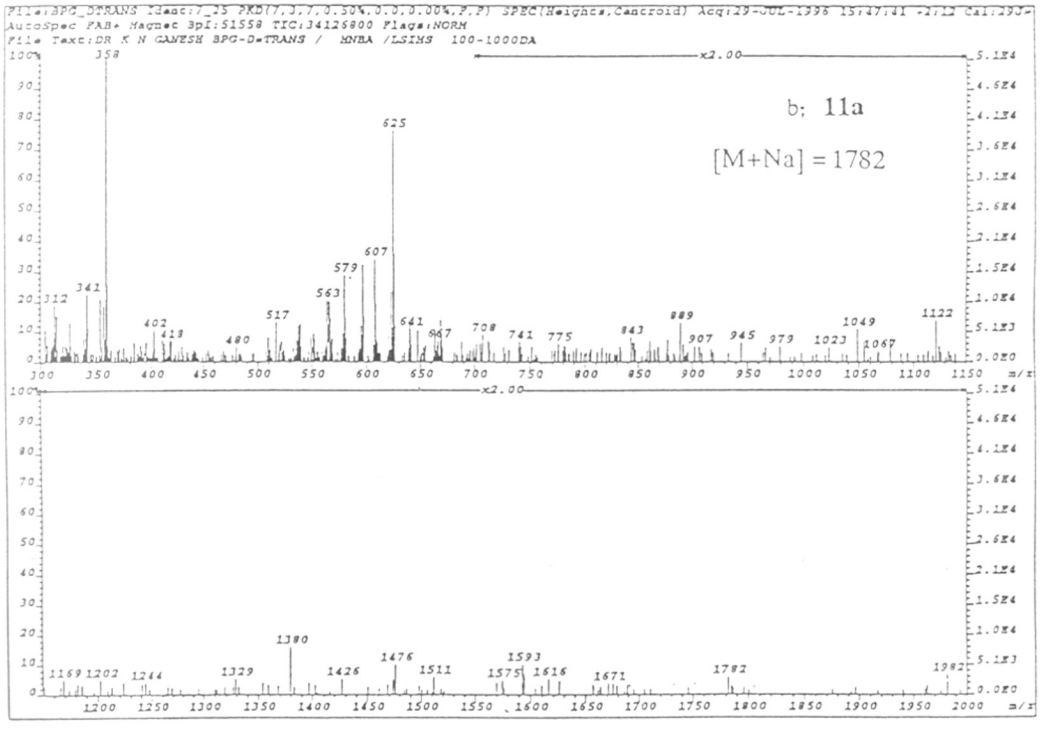


Figure 4: FAB-MS of oligomers a: **10a** and b: **11a** in MNBA and MNBA/LSIMS respectively.

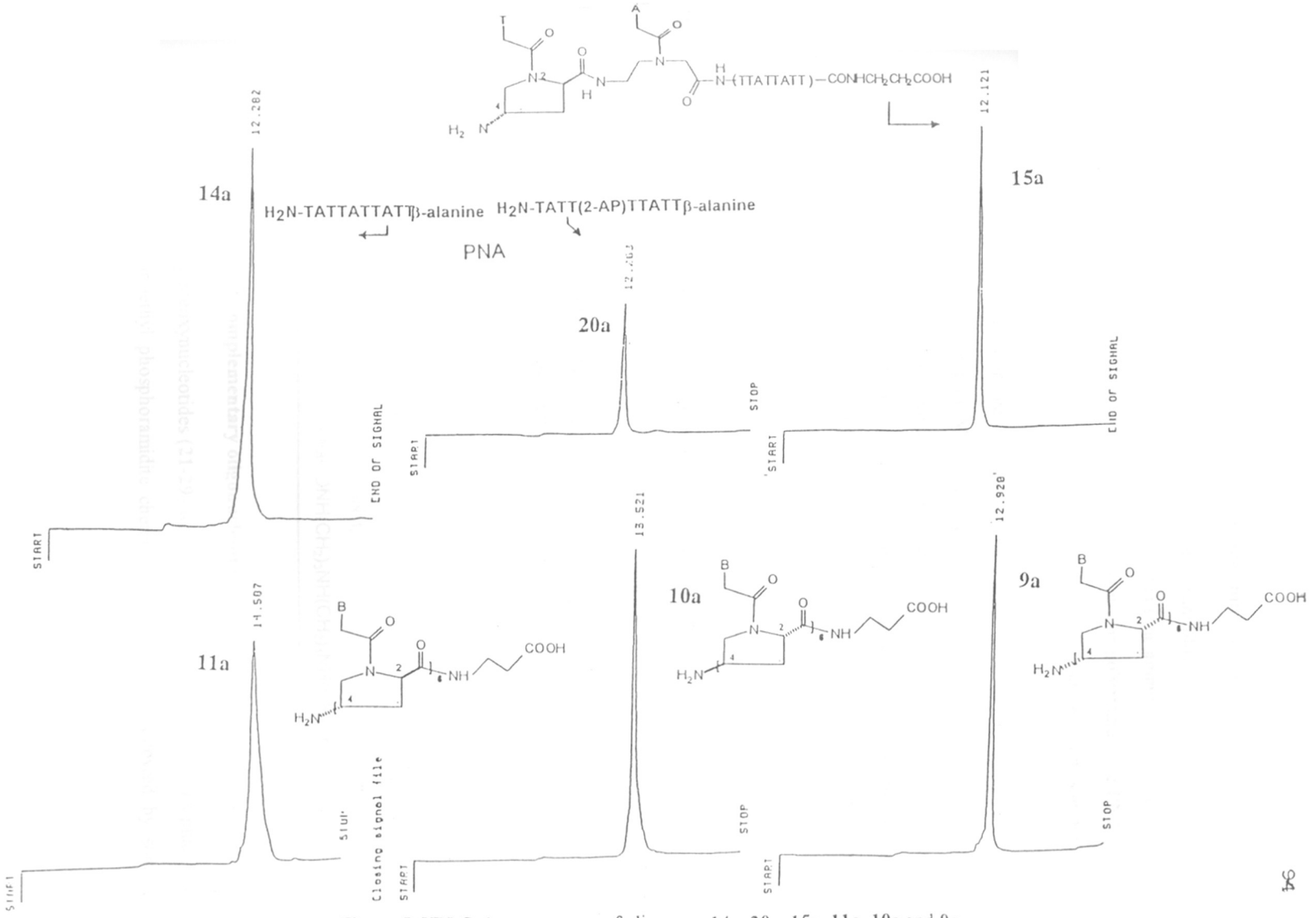


Figure 5: HPLC chromatograms of oligomers **14a**, **20a**, **15a**, **11a**, **10a** and **9a**

3.2.4 Cleavage with spermine to obtain PNA-Spermine conjugates

The benzyl ester linkage between PNA and resin allows cleavage by an amine to obtain PNAs with C-terminal amides.[22] PNA's carrying an amine or a polyamine at C-terminus is possible by reaction of the resin with diamine or polyamine. The PNA **8b** and **8c** were obtained by cleavage of peptide from the MF-resin **8** with spermine and propanediamine respectively by *trans*amidation reaction. Similarly, PNA's with single prolyl units at N-terminus or middle but C-terminus bearing polyamine function **14b-20b** were obtained by treatment of MF-resin (**14-20**) with spermine. The PNA-polyamine analogues (**8b**, **8c**, **14b-20b**) were purified by reverse phase FPLC. In case of spermine the product is expected to arise by the attack of primary amine rather than the secondary amino group due to steric reasons as has been established in earlier studies.[23] Figure 6 shows some of the chormatograms of the purified *sp*-PNA oligomers.

No	Sequence
8b	NH_2(aeg-T)$_6$-β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
8c	NH_2(aeg-T)$_6$-β-alaCONH$(CH_2)_3NH_2$
14b	NH_2(aeg-TAT$_2$AT$_2$AT$_2$)- -β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
15b	NH_2(D-*trans*-Pro-T{aeg-AT$_2$AT$_2$AT$_2$})-β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
16b	NH_2(L-*trans*-Pro-T{aeg-AT$_2$AT$_2$AT$_2$})-β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
17b	NH_2(aeg-CTCTAT$_2$AT$_2$AT$_2$)-β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
18b	NH_2(aeg-CTC{D-*trans*-Pro-T}AT$_2$AT$_2$AT$_2$)-β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
19b	NH_2(aeg-CTC{L-*trans*-Pro-T}AT$_2$AT$_2$AT$_2$)-β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$
20b	NH_2(aeg-TAT$_2$A*T$_2$AT$_2$)- -β-alaCONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$

3.2.5 Synthesis of complementary oligonucleotides

The oligodeoxynucleotides (**21-29**) were synthesised on Pharmacia GA plus, using standard β-cyanoethyl phosphoramidite chemistry on CPG-resin followed by standard

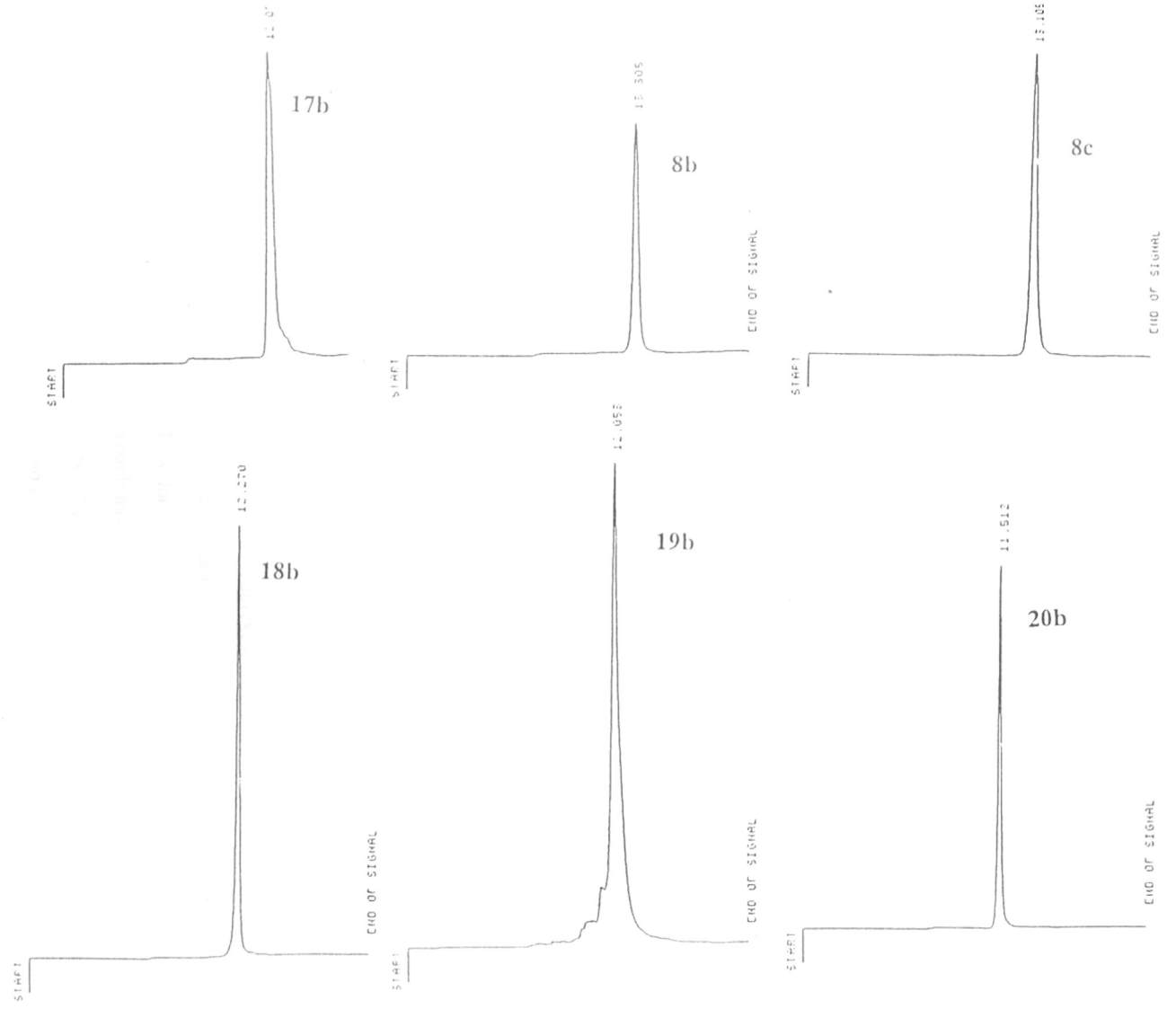


Figure 6: HPLC chromatograms of oligomers **17b**, **8b**, **8c**, **18b**, **19b** and **20b**.

ammonia deprotection.[12] The purity of the oligonucleotides was checked by HPLC and was more than 95-96% and were used directly for biophysical studies. Since PNAs form hybrids with DNA in both parallel and antiparallel orientations, the DNA sequences were designed to obtain specifically, parallel PNA:DNA (5'end adjacent to N-terminus, 5'/N) or antiparallel PNA:DNA (3' end adjacent to N-terminus, 3'/N) hybrids. The sequences **24** and **25** were designed specially to form parallel and antiparallel duplexes with 10-mer PNA sequences while sequences **26** and **27** were prepared to form parallel and antiparallel 13-mer sequences. PNA 10-mers cleaved with spermine were envisaged to show better stability with longer oligonucleotides and hence sequences **28** and **29** were prepared.

DNA	Sequences
21	5'd TTT TTT 3'
22	5'd AAA AAA 3'
23	5'd AAA AAA AAA A 3'
24	5'd AAT AAT AAT A 3'
25	5'd ATA ATA ATA A 3'
26	5'd AAT AAT AAT AGA G 3'
27	5'd GAG ATA ATA ATA A 3'
28	5'd GCG AAT AAT AAT A 3'
29	5'd AAT AAT AAT ACC G 3'

3.3 Biophysical studies of PNA, PrNA and their hybrids with DNA

3.3.1 UV absorbance - Temperature studies

The thermal stabilities of DNA duplexes, triplexes and other structures and associated *trans*itions are conveniently monitored by measuring temperature dependence of their UV absorbance. This has given valuable information about complementary interactions and such an approach has been extended to study PNA interactions as well.[25] The following section reports UV-temperature studies on PNA:DNA duplexes to understand the effect of chiral modifications and conjugates on thermal stability.

3.3.1a UV-thermal stability studies homooligomers PrNA: DNA: The investigation of the effect of backbone modifications on PNA self-structure and the stability of PNA:DNA hybrids, were carried out by temperature dependent UV absorbance measurements. The PNA:DNA duplexes were constituted from appropriate complementary strands for carrying out UV-T_m studies. The homooligomers (**8a-11a**) were annealed with dA_6 (*for details, see section 3.6*) in phosphate buffer (10 mM, pH 7.3) for use in melting studies. The unmodified PNA-T_6 (**8a**):dA_6 complex showed a sigmoidal melting curve (T_m = 19.5 °C). However, similar experiments using DNA hybrids of prolyl hexamers **9a**, **10a** and **11a** did not show any biphasic *trans*ition and gave an eventless linear change in absorbance vs temperature plot. The homooligomers of PrNA alone did not show any change in absorbance with temperature unlike the heterooligomers.

3.3.1b UV thermal studies on heterooligomers PNA:DNA: Table 1 summarizes the T_m data obtained for UV melting experiments of various PNA-DNA hybrids designed to form either parallel (N/5') or antiparallel duplexes (N/3'). The establishment of correct duplexes were indicated by sigmoidal *trans*itions and confirmed by peaks in first derivative plots. The corresponding data from non-duplexes did not fit into this pattern. As expected, the unmodified PNA (**14a**) formed duplex with complementary DNA in both parallel and antiparallel orientations, with parallel mode slightly destabilizing than the antiparallel duplexes. The results in Table 1 indicate that PNA **15a** with terminal D-*trans*-Pro-T shows a better stabilization ($\Delta T_m \approx 3$ °C) compared to L-*trans*-Pro-T **16a**. While both of them bind to complementary DNA **24** in antiparallel orientation (**15a**:**24** and **16a**:**24**) the L-*trans*-oligomer (**16a**) did not show any binding in parallel mode. Interestingly the N-terminal prolyl monomers offered a better overall stability (antiparallel/parallel) to the unmodified control PNA **14a**. The introduction of fluorescent 2-aminopurine (see-later) in the backbone showed stable duplex formation (**20a**:**24**) but with slight destabilization (ΔT_m = 3 °C) compared to the control (**14a:24/25**) consistent with earlier observation.[26] Not much difference was noticeable in the T_m of parallel and antiparallel orientations, similar to that seen with control PNA:DNA duplex. Earlier

Table 1: UV melting data of modified PNA/DNA duplexes

Entry	PNA sequence	DNA sequence	Orientation	T_m(°C)
1	14a	24	ap	28.5
2	14a	25	p	28.0
3	15a	24	ap	38.5
4	15a	25	p	31.5
5	16a	24	ap	35.6
6	16a	25	p	---
7	20a	24	ap	25.5
8	20a	25	p	25.0
9	14a	--	--	31.5[a]
10	20a	--	--	29.0[a]

10 mM sodium phosphate buffer at pH 7.3 and 3 μM strand concentration.
[a] Maximum of the first derivative of sigmoidal melting curve, but not a T_m by definition. (ap:antiparallel orientation p:parallel orientation)

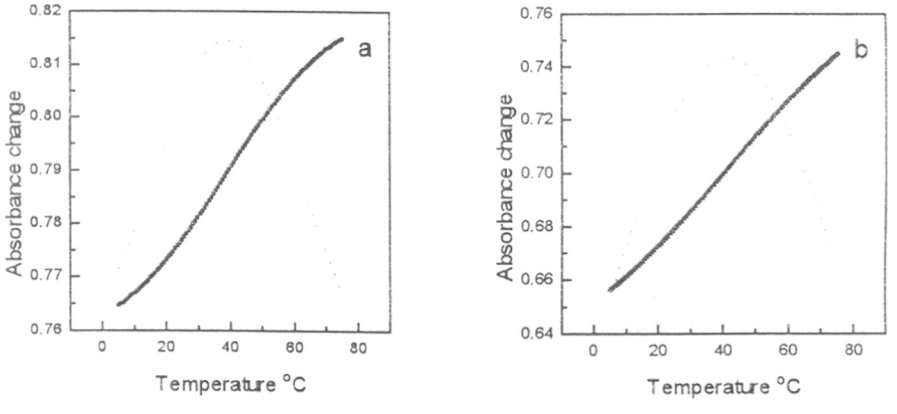


Figure 7: Melting profiles with first derivative plots of PNA:DNA duplexes (a): **15a**:**24** and (b) **16a**:**25** (10 mM Phosphate pH 7.3).

Table 2 summarizes the T_m values obtained when UV melting experiments of PNA containing modified chiral unit in the middle were carried out with the corresponding

complementary DNA oligomers. In case of enantiomeric prolyl monomers **2** and **3** incorporated PNAs, interesting selectivity were noticed with respect to both stereochemistry of the modification as well as the directionality of binding (Table 2).

The PNA-DNA duplexes stabilities are highly sensitive to the presence of salts and hence T_m were measured at two different salt concentrations. The D-*trans*-Pro-T-PNA **18a** formed hybrid with complementary DNA **26** only in antiparallel combination, while no organized structure formed with parallel complementary DNA (**18a**:**27**). In contrast, the L-*trans*-Pro-T-PNA **19a** formed hybrid with complementary DNA only in a parallel orientation (**19a**:**27**), but not in antiparallel combination (**19a**:**26**). However the T_m of both parallel (**19a**:**27**), and antiparallel hybrid (**18a**:**26**), at 100 mM salt concentration were identical. The duplex formation (stoichiometry) of these was established by mixing curve using CD (*see section 3.5.2e*). Fluorescence titration also indicated that the PNA:DNA complexes (**20a**/**20b**:**28**) are formed in 1:1 ratio(*see section 3.3.2a*).

Table 2: UV melting data of modified PNA/DNA duplexes

Entry	PNA sequence	DNA sequence	Orientation	T_m [a]	T_m [b]
1	17a	26	ap	28	--
2	17a	27	p	27.5	--
3	18a	26	ap	33.5	36
4	18a	27	p	---	33
5	19a	26	ap	---	--
6	19a	27	p	33.5	34

(a) 10 mM sodium cacodylate, 100 mM NaCl, 0.1 mM EDTA at pH 7.3 and 3 μM strand concentration. (b) 10 mM phosphate at pH 7.4 (ap: antiparallel, p: parallel)

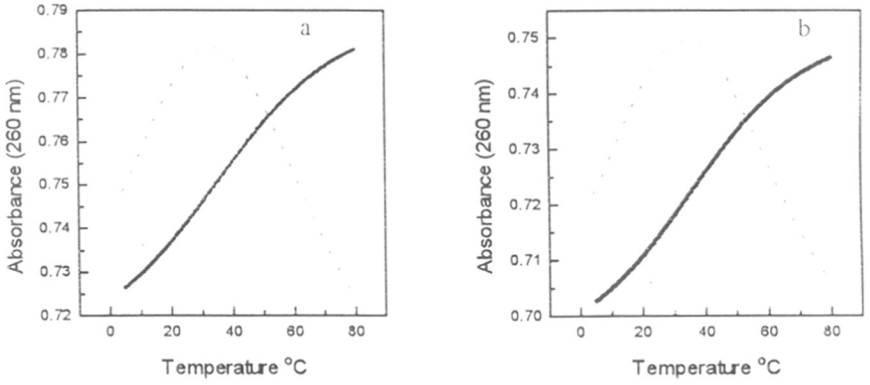
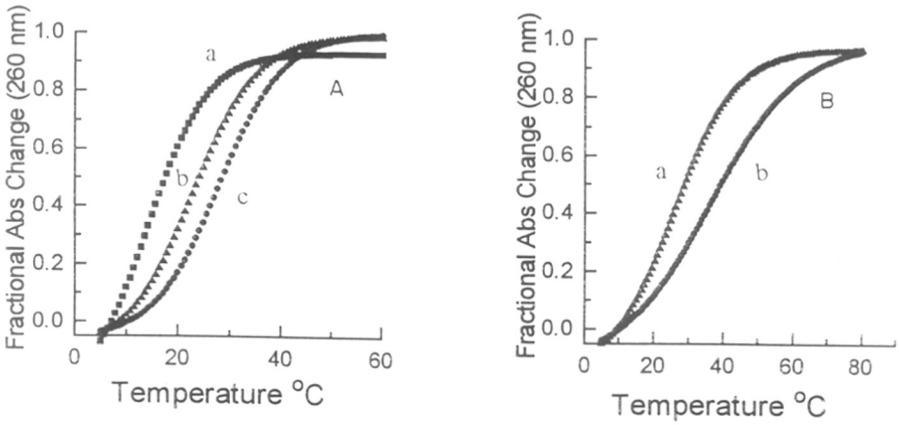


Figure 8: Melting profiles with first derivative plots of duplexes (a) **18a**:**26** and (b) **19a**:**27** .Buffer: 10 mM phosphate at pH = 7.3, with 6 μM strand concentration.

The non-terminal modifications within the oligomer thus enhance T_m by 5-6 °C compared to the parent oligomer (**17a**:**26**/**27**). The selectivity observed at a salt concentration of 100 mM did not change when hybridization melting experiments were performed in absence of salt. It is well known that higher salt concentrations lower T_m of unmodified PNA-DNA hybrids,[28] and for the present PrNA oligomers, the absence of salt enhanced the T_m of antiparallel D-*trans*-Pro-T-PNA hybrids (ΔT_m = + 2.5 °C), while not much change was observed in case of parallel hybrids.

3.3.1c ***Polyamine PNA conjugates and their effect PNA/DNA hybridization:*** Homooligomers **8a-11a** precipitate out in presence of high concentrations of salt and/or bivalent cations like Mg^{2+} due to self-aggregation. This was not the case in peptides cleaved from the resin **8** with 1,3-diamino propane to obtain peptide with a C-terminal amide **8c**. The covalent linking of polycations such as spermine significantly enhances the stability of DNA duplexes and triplexes due to electrostatic acceleration of hybridization of anionic oligomer. To examine analogous effects in PNA:DNA duplex formation, spermine conjugated to the carboxy terminus of PNAs (**8b**, **14b-17b**, obtained by cleavage of corresponding resin with spermine), these PNA conjugates were hybridized with complementary DNA sequences for T_m studies.



A: Triplex melting of (a) **8a:22** (■); (b) **8c:22** (▲); (c) **8b:22** (●) B: duplex melting of (a) **17a:26** (▲) (b) **17b:26**(●)

Figure 9: Melting profiles of spermine PNA analogues.

The UV melting profiles of various PNA-DNA hybrids constituted from appropriate complementary single strands indicated biphasic *trans*itions with T_ms shown in Table 3. As per expectation, the homopyrimidine PNA-T_6 **8a-8c** form triplexes with DNA **22** in a 2:1 stoichiometry. The linking of either a spermine (**8b:22**) or diamine (**8c:22**) to the C-terminus of PNA **8a** lead to a significant enhancement in the T_m of hybrids over that of control PNA:DNA hybrid (**8a:22**). Further, spermine conjugation gave a much higher stability (ΔT_m = 4.5 °C) in duplex **8b:22** compared to the diamine-PNA:DNA duplex (**8c:22**).

It is known that in general mixed sequence PNA oligomers **17a** and **17b** bind to complementary DNA better in antiparallel mode than in parallel mode.[9] In accordance with this, the mixed PNA oligomer **17a** formed hybrid with complementary antiparallel DNA (**26**) in 1:1 stoichiometry with a T_m of 28.5, while no complementation was seen with the corresponding parallel DNA strand (**27**).

Table 3: UV melting data of PNA-spermine conjugates

Entry	PNA sequence	DNA sequence	Orientation	Stoichiometry PNA:DNA	T_m °C
1	**8a**	22	-	2:1	17.5
2	**8c**	22	-	2:1	23
3	**8b**	22	-	2:1	27.5
4	**17a**	26	ap	1:1	28.5
5	**17a**	27	p	1:1	--
6	**17b**	26	ap	1:1	38.6
7	**17b**	27	p	1:1	--

Buffer: 10 mM cacodylate, 100 mM NaCl, 0.1 mM EDTA at pH 7.3.
(ap:antiparallel, p: parallel)

Upon spermine conjugation to **17a**, the resulting PNA oligomer **17b**, retained same selectivity with formation of only antiparallel duplex (**17b:26**). Significantly, the *sp*-PNA conjugate (**17b**) showed a highly remarkable increase in T_m (ΔT_m = 10 °C) upon binding to DNA **26** compared to the unconjugated PNA:DNA hybrid (**17a:26**). The PNA:DNA duplex (**17a/17b:27**) corresponding to parallel orientation failed to show sigmoidal *trans*itions in UV absorption which increased linearly with temperature. The PNA oligomers exhibit pronounced self-ordering in single strand forms, particularly with high purine contents.[29] The unmodified mixed sequence PNA **17a** alone upon heating showed a sigmoidal *trans*ition with a T_m of 27.5 °C. The conjugation of spermine at C-terminus (**17b**) did not show any noticeable effect on this self-melting, which had a profile similar to that of PNA **17a** (T_m = 27.5). Thus, in sharp contrast to the large differences seen with PNA-DNA hybrids, the effect of spermine conjugation on self-order of PNA is negligible.

3.3.2 Fluorescent studies

2-Aminopurine, an isomer of nonfluorescent adenine (6-aminopurine) has intrinsic fluorescence.[30] When present in a DNA double helix it is known to form a stable base pair with thymine (Figure 10) without significantly affecting the local and global structure of

B-form duplex.[31] The incorporation of 2-aminopurine (2-*ap*) into PNA would lead to fluorescent PNA. There are a few reports of attaching other fluorophores

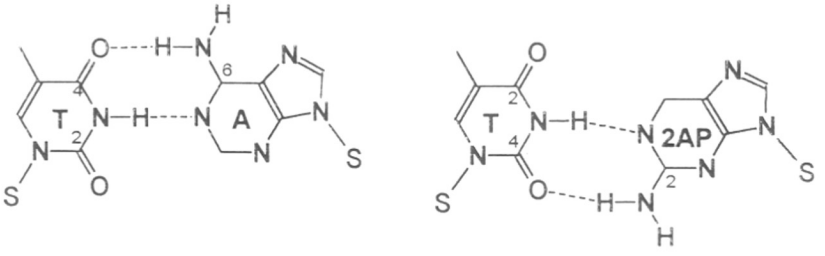


Figure 10

to PNA mainly at N-terminus and using them as labels to detect various properties of PNA.[7,32] The incorporation of intrinsically fluorescent 2-aminopurine into PNA has many advantages.[33] It can be selectively excited in presence of a normal base (A, C, T or G) making it ideal for (i) using as a label, (ii) studying local environment changes resulting from PNA binding with other molecules and (iii) measuring the kinetics of duplex formation and study dynamics of interaction.[34] The incorporation of this base into PNA would therefore add a new dimension to the existing repertoire of PNA properties and potential applications. The following section describes PNA-DNA hybrid formation by monitoring changes in 2-aminopurine fluorescence.

3.3.2a ***PNA:DNA duplex formation: Monitoring by fluorescence*****:** The 2-aminopurine PNA monomer (**6**) and the PNA oligomers (**20a**, **20b**) exhibited florescence emission with maxima at 367 nm upon excitation at 308 nm, similar to that reported earlier[30] for 2-aminopurine base and its nucleoside. Upon incorporation of **6** into PNA, the corresponding PNA **20a** and **20b** did not show any change in the fluorescence observables. Hydrogen bond formation of 2-aminopurine with thymine is known to decrease the quantum yield of the 2-aminopurine.[34] The fluorescent PNA **20a** upon binding to its complementary DNA **24** effected a decrease in fluorescence intensity indicating a successful hybridization. Similar behavior was obtained with *sp*-PNA conjugate **20b** incorporating 2*ap* upon addition of complementary DNA **28**. The duplex formation was substantiated by a fluorescence titration experiment in which stoichiometric

addition of complementary DNA **28** to a fluorescent PNA solution was accompanied by a gradual decrease in fluorescence intensity upto an addition of 1 equivalent of DNA, beyond which the intensity did not vary much (Figure 11). The fluorescent 2*ap*-*sp*-PNA **20b** also showed similar characteristics upon DNA titration but the magnitude of intensity quenching for the **20b** (40%) was almost double that of **20a** (20%).

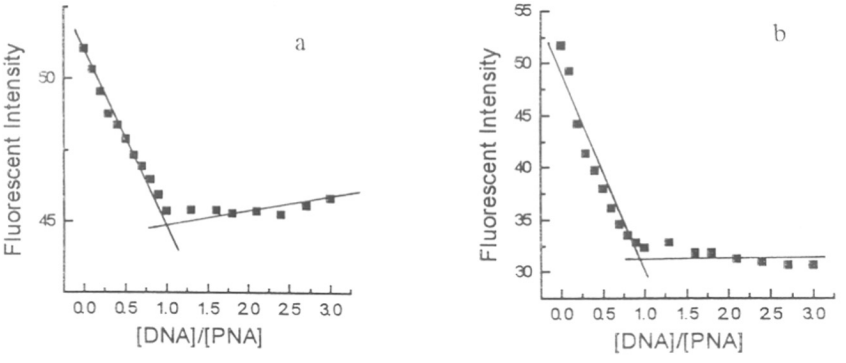


Figure 11: Fluorescence titration of DNA (**28**) with (a) PNA (**20b**) and (b) PNA (**20a**). (3 μM PNA in 10 mM Phosphate at pH 7.3).

The association constants for PNA:DNA binding were calculated using a Scatchard plot. The titration data gave an intrinsic binding constants of 2.0 x 10^5 for **20a** and 4.2 x10^6 for **20b** (details section 3.6). This amounts to a 20 times stronger association of *sp*-PNA (**20b**) than unmodified PNA (**20a**). The association constants calculated in this way were similar to the reported values for DNA:PNA binding from UV data.[9]

3.3.2b Temperature dependent fluorescence: The strength of PNA self-organization and PNA-DNA binding was also studied by temperature dependent fluorescence experiments. Upon heating, both fluorescence intensity of PNA alone and PNA:DNA duplex showed a gradual increase. The fluorescence intensity reached a saturation around 35 °C for **20a** and at 60 °C for **20b** beyond which heating resulted in a small decrease of intensity (Figure 12). The percentage increase in total intensity for PNA-DNA melting (60%) was much larger than that observed for PNA self-melting (25%) under identical temperature range.

The monomer **6** did not show any noticeable change (<2%) in fluorescence intensity upon heating in the temperature range of 5-50 °C.

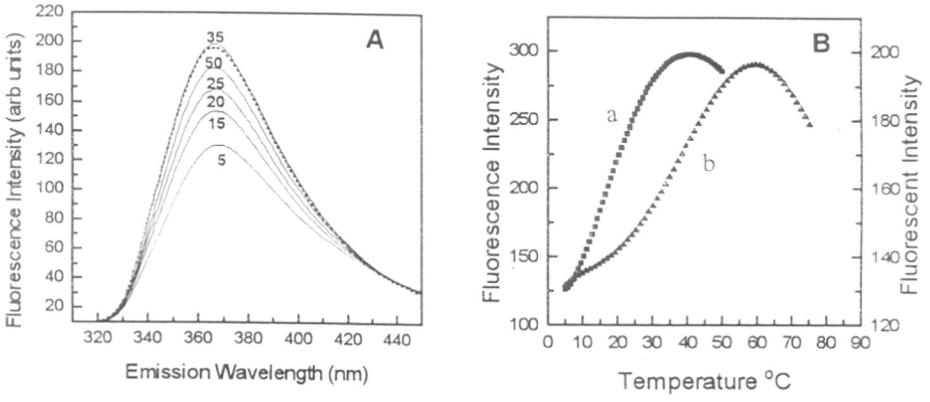


Figure 12: A: Fluorescent spectra at different temperatures B: Melting profile of duplex (a) **20a**:**28** (■) (b)**20b**:**28** (▲).

In consistence with UV-T_m experiments, the self-melting of PNA (**20a**) as seen from fluorescence is higher than that of corresponding PNA-DNA duplex. The T_ms obtained from temperature dependent absorbance (25 °C) was slightly lower than (35 °C) that from the fluorescence temperature data. The fluorescence T_m of spermine PNA:DNA (**20b**:**28**) hybrid was 15 °C more than that of PNA:DNA (**20a**:**24**) duplex hybrid agreeing with the UV-T_m results.

***3.3.2c Kinetics of PNA:DNA hybridization*:** The kinetics of PNA:DNA hybridization process was examined by monitoring the fluorescence emission decay at 367 nm as a function of time after mixing 2-*ap* PNA (**20a**) with the complementary, antiparallel DNA (**28**). The emission intensity decreased exponentially over a period of 6 minutes and there after remained constant. In contrast, either **20a** alone or its mix with DNA dA_{10} (**23**) containing 3 mismatches did not exhibit any effect on the PNA fluorescence which remained constant over this period (Figure 13). The time-dependent decrease in fluorescence of **20a** immediately upon mixing with DNA, may be attributed to specific formation of 2-*ap*-PNA:DNA hybrid. The fluorescence decay profile observed in Figure 13 could be fitted into a double exponential function with time constants of 46 seconds

and 290 seconds, indicating that the fluorescent base 2-*ap* in PNA monitors at least two types of local events, a fast duplex formation by complementary hydrogen bonding, followed by a slow reorganization of the helix after duplex formation.

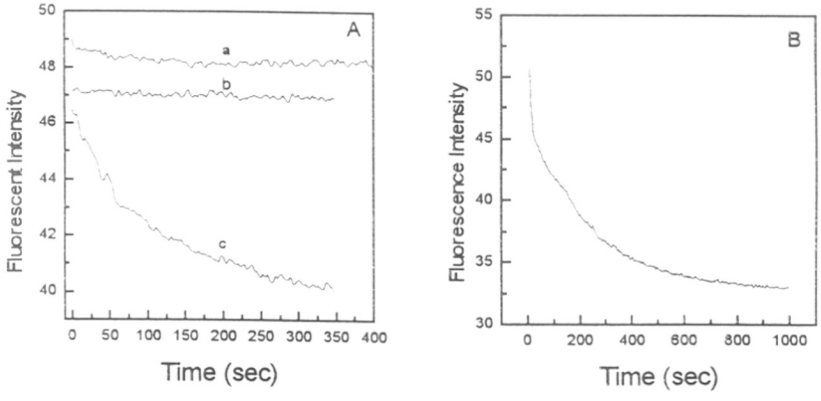


Figure 13: Kinetics measured by fluorescence λ_{Ex}(308nm) and λ_{Em}(367nm). Buffer: 10 mM Phosphate at pH 7. A: (a) binding of PNA **20a** to DNA **23** (b) PNA **20a** (c) binding of PNA **20a** to DNA **24**. B: binding of PNA **20b** to DNA **24.**

A similar behavior was seen in experiments with 2-*ap*-*sp*-PNA:DNA duplex (**20b**:**28**). The fluorescence decay curve was exponential but with a much faster rate of decay in the initial 10 seconds (20% loss) followed by a slower decay (Figure 14). The time constants from a double exponential fitting were 9 seconds and 250 seconds, in contrast to the earlier 40 seconds and 290 seconds for unmodified PNA:DNA duplex. The *sp*-PNA **20b** alone does not show any change in fluorescence with time in agreement with our earlier results. Thus fluorescence from 2-*ap*-PNA is thus a good monitor of hybridization strength and efficiency.

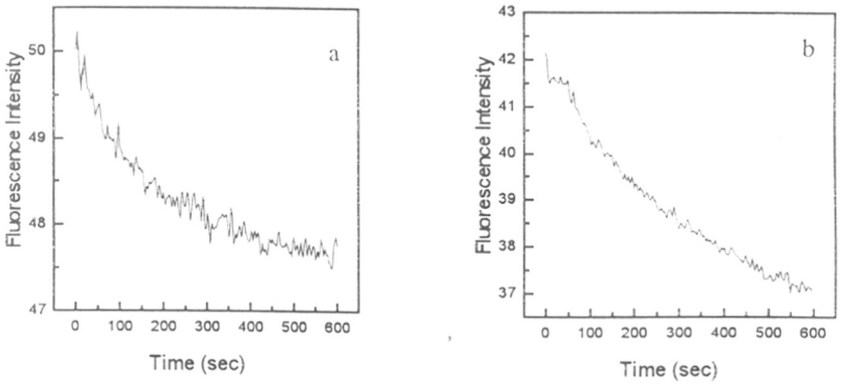


Figure 14: Kinetics measured by fluorescence λ_{Ex} 308 nm and λ_{Em} 367 nm. Buffer: 10 mM phosphate at pH 7.0. (a) PNA **20a**: DNA **24** (b) PNA **20a**: DNA **28.**

Strong ionic strength dependency are observed for binding of the PNA to DNA which was also dependent on the sequence under study.[R] The fluorescence experiments were therefore carried out in presence of 100 mM NaCl. The initial drop in fluorescence is much lower in 100 mM salt concentration as compared to that in absence of salt. From the Figure 13 it is evident that the fast binding component is not observed at 100 mM NaCl concentration and the overall rate of binding is much slow compared with the data in absence salt.

*3.3.2d **Fluorescence Anisotropy***: Another fluorescence observable capable of giving useful information on molecular environment is fluorescence anisotropy. The base 2-*ap* in **20a** senses the local changes in conformation/environment accompanying the self-melting of PNA from an ordered chain to a random coil. This process also reorients the fluorophore, leading to a decrease in anisotropy. A higher fluorescence anisotropy and polarization values imply that the fluorophore is more rigid.[34] The polarization and anisotropy values of 2-*ap* fluorescence were measured as a function of temperature for both PNA **20a** and PNA:DNA duplex (**20a**:**28**) and the results are shown in Table 7. An appreciable decrease of both was observed with increasing temperature till about 40 °C after which it remained constant.

Table 4: Anisotropy and fluorescence polarization values for PNA:DNA duplexes

Temperature	Anisotropy		Polarization	
	20a	**20a:28**	**20a**	**20a:28**
15 °C	0.73	0.65	0.95	0.95
20 °C	0.60	0.56	0.84	0.81
35 °C	0.40	0.55	0.68	0.8
50 °C	0.40	0.55	0.68	0.8

The observed change in anisotropy values overall suggests that 2-*ap* base faithfully reports the environment and structural changes in PNA upon mixing with DNA and/or heating, thus substantiating its utility in monitoring such dynamic events.

3.4 Discussion

The remarkable properties of PNAs had led to many types of chemical alteration of backbones. These comprise, extension of backbone in eda unit, in the glycine unit, introduction of amino acids (chirality) into the backbone, shifting of nucleobase from the central N to the glycyl α-carbon and DNA-PNA chimeras. The single purpose of all these modifications has been to bring directional/orientational selectivity into PNA:DNA interaction and improve their solubility properties for better cell uptake, without compromising on their hybridization strength. In the present work, the approach was to distend the backbone and simultaneously introduce chirality by linking the β-methylene of eda to α-C of the glycine unit. This would result in conformational restriction of PNA backbone, thereby affecting its properties such as self-organization and DNA hybridization. Since the proposed modification also introduces 2 chiral centers and hence four possible stereoisomers in the backbone, it was hoped that chirality based selectivity could be imparted into the backbone. The previous section has described the results

obtained in PNA:DNA hybridization experiments and in the following section the implications of these results are discussed.

3.4.1 UV-temperature studies of PNA oligomers

3.4.1a Prolyl nucleic acid homooligomers (PrNA): The prolyl nucleic acid homo oligomers (**9a-11a**) irrespective of their monomer stereochemistry do not show any sigmoidal transitions in UV absorbance upon heating. The observed gradual melting over a broad temperature for (**9a-11a**) suggests that these do form structures involving base stacking leading to hyperchromicity. However, overall, they are not well defined unique structures and the transitions are non-cooperative. The structures are random coils but with distributed clusters of base stacks. This is unlike the corresponding aeg-PNA hexamers (**8a**) which exhibited sigmoidal transitions. A similar gradual, non cooperative melting seen for prolyl homooligomer PNA duplexes with complementary DNA indicates the absence of double/triple helical structures. The linear absorbance increase again corresponds to individual melting of random coils of PNA and DNA. The incapability of the prolyl nucleic acid homooligomers to form ordered helical structures either alone or in presence of complementary DNA is perhaps due to the severe restriction in backbone conformation arising from the introduction of methylene bridge between eda and gly components of PNA monomer unit and the resulting unfavorable conformation of 5-membered pyrrolidine rings. To understand such effects, a systematic introduction of the designed conformational restriction, one at a time into the backbone was necessary and the studies directed towards this objective is discussed in next section.

3.4.1b Prolyl Nucleic Acids (PrNA):DNA hybrids: In contrast to the homooligomers, single substitution of prolyl nucleic acid monomers within aeg-PNA gave very interesting results. The N-terminal substitution of D-*trans*-Pro-T **3** or L-*trans*-Pro-T **2** in a PNA oligomer led to significant stabilization of the corresponding PNA:DNA duplexes (**15a**:**24/25** and **16a**:**24/25** respectively) by 3 °C to 10 °C. The D-*trans*-Pro-T-PNA (**15a:24**) formed antiparallel PrNA:DNA duplexes with highest stability and similar to that observed with unmodified PNA heterooligomers, the antiparallel duplex T_m was higher than that of parallel duplex by 7 °C. In case of L-*trans*-Pro-T-PNA (**16a**), the

enhancement in T_m over that of unmodified PNA was much smaller (≈3°C) for the antiparallel duplex and no duplexation was seen in parallel combination. In comparison, the T_Ms of internally substituted 13-mer PrNA of D-*trans*-Pro-T (**18a**) or L-*trans*-Pro-T (**19a**) were higher than the corresponding unmodified PNA oligomers **17a**. An interesting feature was observed with respect to orientational selectivity of D-*trans*-Pro-T and L-*trans*-Pro-T containing PNAs, with internal substitution. While D-*trans*-Pro-T-PNA (**18a**) formed duplex (**18a:26**) with complementary DNA only in antiparallel direction, the L-*trans*-Pro-T-PNA (**19a**) formed duplex only in parallel orientation with complementary DNA (**27**). Thus introduction of a single chiral unit into achiral PNA backbone has manifested in imparting orientational preferences. The presence of 100 mM NaCl had no effect on antiparallel L-*trans*-Pro-T-PNA:DNA duplex (**19a:26**) whereas the parallel D-*trans*-Pro-T-PNA:DNA duplex (**18a:26**) was stabilized by 3 °C. While this work was in progress a related report appeared in literature describing incorporation of L-*trans*, L-*cis* and D-*trans*-Pro-T monomers alternatively with unmodified aeg-PNA monomers. The DNA-duplex melting T_m studies revealed that in these alternating 8-mer PNA heterooligomers, the L-*trans*-Pro-T-PNA analogue exhibited 6-7 °C increase in T_m compared to the equivalent unmodified T_m. The corresponding D-*trans*-Pro-T-PNA oligomer was destabilized considerably. The present results give an insight into the orientational preferences (D-*trans*-Pro-T-PNA in antiparallel and L-*trans*-Pro-T-PNA in parallel) and the effect of salts which selectively stabilize antiparallel orientation. Overall, the heterooligomeric PNA consisting of D/L-*trans*-Pro-T units either at N-terminus or within the PNA backbone show stronger binding to complimentary DNA than original PNA and are selective in their orientation with complementary DNA in PrNA-DNA duplexes.

3.4.1c ***Spermine-PNA:DNA hybrids*****:** The UV-absorbance-T_m data of polyamine PNA-DNA hybrids clearly establish that the stability of PNA-DNA duplexes can be considerably enhanced by polyamine conjugation at the C-terminus. The conjugation of naturally occurring tetraamine, spermine was better than that of 1,3-propanediamine, since it has more positive charges. Among the sequences synthesized for the present study, it was seen that the mixed *sp*-PNA conjugate **17b** formed a hybrid with complementary DNA **26** only

in an antiparallel direction (N/3') while there was no duplexation (**17b:27**) in parallel mode (N/5'). Further, an enhancement of 10 °C was observed for the *sp*-PNA:DNA hybrid (**17b:26**) as compared to the control (**17a:26**). The *sp*-PNA oligomers alone exhibited a self-melting which was as good as that of unmodified PNA, with biphasic *trans*ition and similar T_m values. Thus, in sharp contrast to enhancement seen in *sp*-PNA:DNA hybrids, the effect of conjugated spermine on self-order seems to be negligible.

The enhanced stability results on DNA hybrids involving PNA-spermine conjugates are novel, especially in view of the report[35] that external addition of spermine in 10-100 μM concentrations destabilized antiparallel (N/3') PNA-DNA hybrids by 5-8 °C and parallel (N/5') PNA-DNA hybrids by 2-5 °C. At higher concentrations of spermine (>1mM), PNA-DNA hybrids precipitated out of solution. This effect of spermine and other cations (Na^+, Mg^{2+}) at concentrations >1M, where electrostatic interactions saturate, was attributed to the predominance of hydrophobic interactions in stabilizing PNA complexes.[35] In contrast to the externally added spermine (10 μM), it was revealed that the covalently conjugated spermine (concentration equivalent to that of PNA, 4 μM) stabilized both 2:1 (**8b**:**22**, **8c**:**22**) and 1:1 (**17b**:**26**) hybrids, in presence or absence of 100 mM NaCl. For PNA decamers containing cationic lysine at carboxyl end, an increase in T_m of antiparallel PNA-DNA duplex by 3 °C was reported.[36] The conjugation of either 1,3-propanediamine or spermine effected a much higher stability for both triplex (6-mer) and duplex (13-mer), irrespective of the PNA oligomer length. The PNA-spermine conjugates described here contain spermine linked to C-terminus of PNA via a β–alanyl spacer to delink any direct effects of conjugation on PNA conformation (unlike that seen with D/L-lysine). The ineffectiveness of spermine conjugation on PNA self-order suggests that the observed stability of hybrids is a consequence of its interaction with DNA. The extra stabilization of hybrids from PNA-spermine conjugates may arise from a specific association of spermine chain with terminal bases/phosphate of DNA through hydrogen bonding/electrostatic interactions and thereby minimize the destabilizing end-fraying effects. In addition, the linked spermine may also accelerate the initial events in hybridization such as nucleation by electrostatic interaction.[13] The PNA-spermine

conjugates are charged molecules and were more soluble in aqueous medium than the neutral PNA oligomers, similar to the lysine-PNA oligomers.[37]

3.4.2 Fluorescence spectroscopy of 2-*ap*-PNA

DNA by itself is not fluorescent, the introduction of 2-*ap* into DNA does not affect the local and global conformation of B-form duplex. This fluorescent isomer of non-fluorescent adenine is known to form two hydrogen bonds with thymine when present in a DNA duplex. The properties of 2-*ap* are well recorded in the literature.[30] The introduction of 2-*ap* into PNA gave fluorescent oligomers **20a** and **20b** which were used in structural dynamics study.

3.4.2a PNA:DNA duplex formation: The stoichiometry of binding of PNA to DNA was determined by monitoring the change in fluorescence intensity with added DNA. The 2-*ap*-PNA (**20a**) showed a small but significant change in the fluorescence intensity (52-45 arbitrary units) on addition of DNA (**28**) and the quenching was saturated with one equivalence of DNA. There was a tremendous change in the quenching of the spermine analogue 2-*ap*-*sp*-PNA (**20b**) where the fluorescence intensity dropped from 52 to 30 arbitrary units. The comparatively slow rate of duplex formation in **20a** is due to presence of the anionic C-terminal carboxylic acid. As the spermine analogues have terminal positive charges, these interact with the negatively charged DNA to form much stable complexes. This was observed by determining the binding constants. The 2-*ap*-*sp*-PNA (**20b**) binds with 20 times higher efficiency compared to 2-*ap*-PNA (**20a**). Similar results were observed when a DNA oligomer was tagged with cationic peptide. The binding efficiency determined here are consistent with those reported in the literature.[9]

3.4.2b Temperature dependent fluorescence: The blue excitation peak of 2-*ap* free base in water shows that a reduction in ground state energy upon hydrogen-bonding which is greater than that in the excited state energy due to increased dielectric interaction. The lack of any emission spectral shift of the 2-*ap*-PNA (**20a**, **20b**) decamers with temperature is consistent with the earlier observation.[34] The effects of hydrogen-bonding on the ground state of 2-*ap* in the melted helix is about equal to the reduction of the excited state energy by a combination of dielectric and relaxation effects. There was only a 2 nm blue shift in

emission spectra of the PNA:DNA (**20a**:**28**) complex on heating. This indicated that at higher temperatures (50 °C), the 2-*ap*-PNA is in hydrophilic environment. The fluorescent intensity showed a small decrease both in 2-*ap*-PNA (**20a**) 2-*ap*-PNA:DNA duplex beyond 40 °C and this is due to collisional quenching with the solvent molecules. The increase in fluorescent intensity of 2-*ap*-PNA (**20a**, **20b**) with temperature implies that they are highly ordered structures and 2-*ap*-PNA is in a relatively hydrophobic environment in PNA oligomer.

*3.4.2c **Kinetics and Anisotropy of PNA:DNA hybridization***: The kinetics of PNA:DNA hybridization process was examined by monitoring the fluorescence emission decay at 367 nm as a function of time after mixing 2-*ap*-PNA **20a** with the complementary antiparallel DNA **28**. The emission intensity exponentially decreased over a period of 6 minutes. The PNA **20a**/**20b** alone or their mix with dA_{10} does not show any change in fluorescence intensity. This time dependent decrease in fluorescence of **20a**/**20b** upon mixing with complementary DNA may hence be attributed to specific formation of PNA:DNA duplexes. Similar kinetic behavior has been previously observed as for PNA:PNA and PNA:DNA hybrids by monitoring the evolution of a CD signal arising from helical propagation upon duplex formation with the complementary strand. The decay profiles presently seen could be fitted into a double exponential function with time constants of 46 and 290 seconds indicating that fluorescent base 2-*ap* in PNA monitors at least two types of local events, a fast duplex formation by complementary hydrogen bonding followed by a slow reorganization of the helix after duplex formation. The *sp*-2-*ap*-PNA **20b** also shows a quenching when complementary DNA is added and the change in fluorescent intensity upon addition of DNA is much higher as compared to **20b**. The fast decay is over in 9 seconds while the slow decay takes about 250 seconds. This indicates a faster kinetics of duplex formation in case of **20b** which can be attributed to the presence of positive charges in the C-terminus that favorably interact with the negative charges on DNA backbone.

Structural studies of DNA/PNA using fluorescence are best done without perturbing their structures, which is possible only by the use of intrinsic fluorescence from

the nucleobase. In this context, 2-aminopurine is an ideal fluoroprobe since it senses both steady state and dynamic conformational changes. The measurement of fluorescence anisotropy at different temperatures showed an appreciable decrease with increasing temperature till about 40 °C and remained constant thereafter. This behavior indicates that 2-*ap*-PNA, like PNA is considerably self-organized at low temperatures and melts at higher temperature. Thus, the 2-*ap* residue in **20a** senses the local changes in conformation/environment accompanying the self-melting of PNA from an ordered chain to a random coil, a process that also reorients the fluorophore.

3.5 Circular dichroism studies

Circular dichroism (CD) measurements have been in use for over 20 years to study the conformations of nucleic acids in solution.[38,39] Although detailed structural information such as from X-ray crystallography or NMR spectroscopy is not available from CD spectra, the CD spectrum in solution can provide a reliable determination of the overall conformational states when compared with CD spectra of reference samples. The information from CD spectrum is complementary to that from absorption spectrum. Whereas the percentage absorption change at 260 nm on denaturation of a DNA duplex is much larger than the percentage change of the long wavelength CD bands, the different secondary structures of DNA can be more easily distinguished by CD spectra.

3.5.1 CD spectral studies of PNA:DNA duplexes

Although, PNA is non-chiral, its complex with DNA becomes chiral and CD would therefore have a great potential in characterizing PNA-DNA hybrids. Chirality induction into PNA strand can be achieved by linking amino acids[9], peptides[40] and oligonucleotides[41] at the terminus or by incorporation of chiral amino acids in place of glycine in the PNA backbone.[37,42] One of the earlier reports on PNA-DNA complementation, Wittung[43] documented the CD spectra of PNA-DNA and PNA-RNA hybrids in comparison with the corresponding DNA-DNA and DNA-RNA complexes. The CD spectra of DNA-DNA and antiparallel PNA-DNA/PNA-RNA duplexes were quite similar suggesting that PNA hybrids are right handed helices with a base pair

geometry not much different from A or B form DNA helix.[38] Interestingly, CD spectra and structure of parallel versus antiparallel PNA-DNA duplexes were distinctly different.

CD was next used to demonstrate that two complementary PNA strands can hybridize to one another to form a helical duplex.[43] The presence of L or D lysine residue at the carboxy terminus seeded preferential chirality development of which was kinetically followed by evolution of a band at 220 nm. The CD curves from L and D lysine linked PNA exhibited exact mirror image relationships of the resultant CD curves. The kinetics experiments indicated that helix formation between the two complimentary PNA oligomers is relatively slow compared with DNA-DNA or PNA-DNA duplex formation. From a CD study of chemically linked PNA tetramer duplexes of all 16 combinations of two bases proximal to a carboxy teminal lysine residue, it was shown that the base next to the chiral center must be either a G or C for efficient induction of helicity. The propagation length of induced chirality in PNA duplexes was found to be ~10 base pairs. The chiral communication between the amino acid and PNA oligomer was proposed to be the through a combination of dipole interactions between "amide triads" and hydrophobic interactions. When the chiral center was attached at N-terminus of PNA, a significantly different CD spectrum was obtained which corresponded to opposite helicity and with somewhat lower amplitude compared to that with L-lysine at C-terminus.

CD titration of H-T_8-lys-NH_2 (PNA) with DNA showed formation of PNA_2:DNA right handed triplex, supported by additional use of linear dichroic spectroscopy.[20,44] The effect of salts on conformation of PNA-DNA duplexes was followed by CD which indicated that the antiparallel N/3' PNA-DNA is similar to DNA B-helix but quite different from that of parallel N/5' PNA-DNA duplex. The addition of salt (NaCl) did not affect the CD spectrum of antiparallel N/3' PNA-PNA, PNA-DNA and DNA-DNA duplexes while parallel N/5' PNA duplex showed substantial changes upon addition of salt. The CD spectra of parallel N/5' duplex had significant contributions from single stranded PNA and DNA at high salt concentrations indicating dissociation.

Temperature dependent CD spectral studies[45,46] on PNA-hexamer with alternating D-homoalanyl adenine/L-alanyl thymine indicated self-pairing in PNA-PNA duplex to form a linear antiparallel complex. At 0°C, the hexamer showed a strong Cotton effect

with a maximum at 274 nm and minimum at 255 nm and both bands were of low intensity at high temperatures 60°C.

PNA composed of an ethylene linker between the backbone and nucleobase rather than the natural methylene carboxyl linker upon DNA complexation exhibited a T_m lower by 24 °C compared to unmodified PNA.[29] These results were obtained by temperature dependent CD studies of PNA-DNA duplexes, by following signals due to CD of DNA.

3.5.2 Present work: CD studies on prolyl PNA:DNA hybrids

The application of CD spectroscopy to analyze PNA-DNA duplexes though as yet limited, has given useful insights into induced asymmetry in PNA conformation by the presence of a chiral center at the terminus or inside the backbone. However the interpretation of CD of PNA containing chiral centers within the backbone is not straightforward as evident by the following discussion of present results. A successful and meaningful interpretation of CD of biopolymers results also require assignment of CD bands to specific electron *trans*itions and a systematic analysis of CD of various dimers trimers etc. In the absence of such information the analysis is only qualitative. In the present series of compounds, depending on the sequences available and the CD instrument time (all CD data done from external source) only limited data have been accumulated and hence the following discussion is necessarily constrained.

3.5.2a CD of single stranded Prolyl Nucleic Acid homooligomers (9a-11a): Figure 15 shows the CD spectra recorded for PNA-T_6 homooligomers derived from L-*trans*, D-*trans* and L-*cis* monomers. As expected, the D-*trans*-Pro-T_6 (**11a**) and L-*trans*-Pro-T_6 (**10a**) give perfect mirror-image profiles. L-*trans* oligomer exhibited a positive CD band at 272 nm and a small negative band at 242 nm with crossover at 258 nm. A slight displacement of this CD spectra has occurred due to base line drifts, but D-*trans* hexamer gave a CD pattern with exactly opposite sign to that of L-*trans* hexamer. In comparison, the L-*cis* hexamer showed a faint positive band at 280 nm with a large negative band at 255 nm. This indicates that the sign and amplitude of CD bands are significantly influenced by the stereochemistry of the monomeric units in the modified PNA-oligomer. As these did not

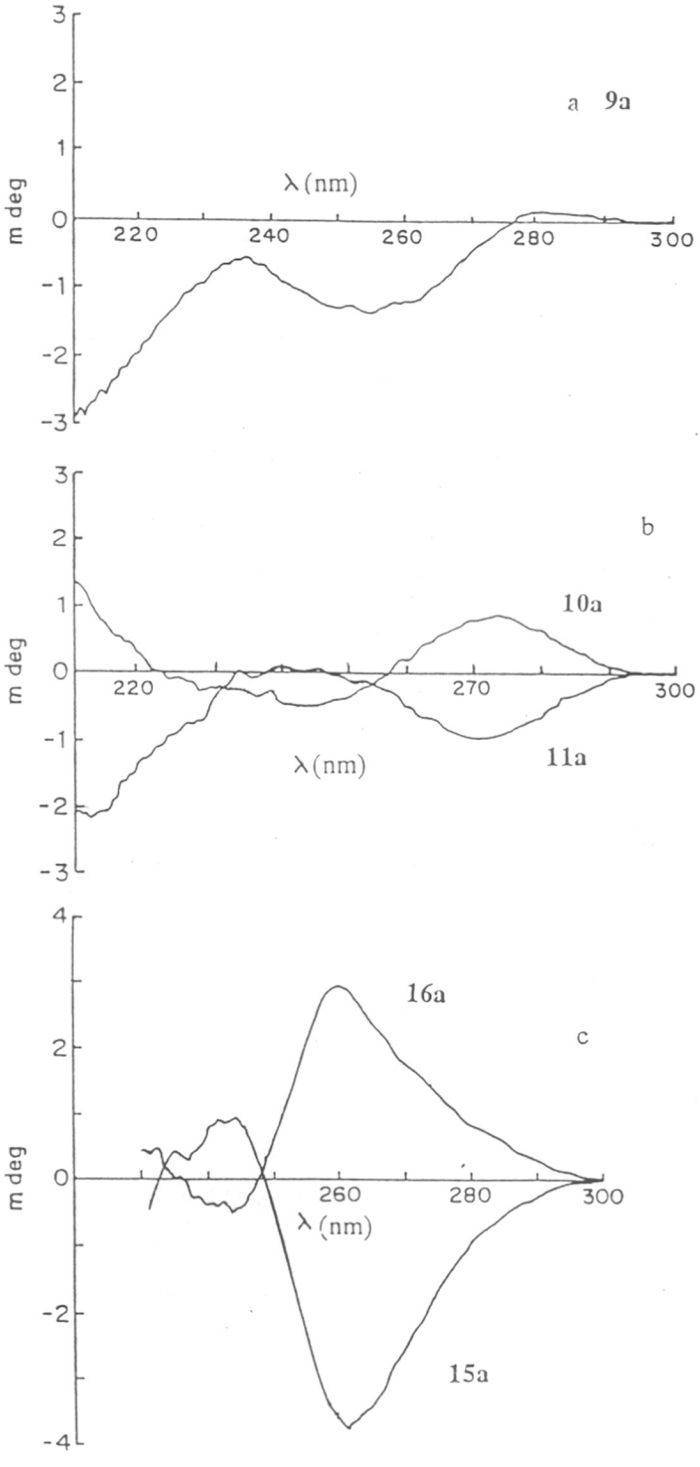


Figure 15: CD spectra of oligomers a: **9a**; b: **10a** & **11a** and c: **15a** & **16a**.

exhibit satisfactory duplexation with DNA (UV-T_m) no further studies on homooligomers were pursued.

3.5.2b CD spectra of heterooligomeric 10-mer prolyl PNAs: Figure 15 shows CD of 10-mer heterooligomers (**15a**) and (**16a**) containing a single unit of chiral D-*trans*-Pro-T **3** or L-*trans*-Pro-T **2** monomers respectively at the N-terminus of aeg-PNA oligomer sequence. The patterns in terms of signs of CD bands are similar to the corresponding homooligomeric hexamer sequences *(see section 3.5.2a)* but considerable differences were noticed in the wavelength positions of CD bands. The 10-mer L-*trans*-Pro-T-PNA **16a** exhibited a large positive band at 262 nm with a much reduced negative band at 242 nm and a crossover at 248 nm while the 10-mer D-*trans*-Pro-T-PNA **15a** showed opposite signs in the corresponding CD bands. Such significant differences in the positions of CD bands between the homooligomeric hexamer sequences (**10a** and **11a**) and N-terminal L/D-*trans*-Pro-PNAs indicate completely different patterns of base stacking in those two PNA oligomers. Since these CD spectra are similar in their sign relationships for the same stereomonomers (**10a** and **16a** both with L-*trans*-Pro-T, **11a** and **15a** both with D-*trans*-Pro-T), it appears that the induced handedness of helical base stacking are same for specific N-terminal chiral monomer, be it in homo or hetero oligomers. More information on these aspects requires a systematic analysis of CD spectra of suitable model oligomers.

3.5.2c CD spectra of D-trans-Pro-T (18a) and L-trans-Pro-T (19a) PNA 13-mer heterooligomers: The two single stranded 13-mer heterooligomers (**18a**) and (**19a**) containing one 4-aminoproline unit and differing in stereochemistry at both centers (C2 and C4) gave CD spectra which are mirror-image profiles (Figure 16). L-*trans*-Pro-T-PNA (**19a**) gave a CD with positive amplitudes at 262 nm and 220 nm while the D-*trans*-Pro-T-PNA (**18a**) gave a mirror image CD with negative bands at 262 nm and 220 nm. Both have a high positive band towards near-UV region below 205 nm. The temperature dependent changes in their CD spectra in the region 5-50 °C is shown in Figure 16 and these as well show mirror image CD curves as a function of temperature. The magnitude of bands at 262 nm in both showed a decrease towards X-axis, becoming almost saturated after 40 °C. These bands also became broad with a 6 nm bathochromic shift from 260 nm

to 268 nm. In comparison to this, the 220 nm showed a relatively less intensity change but no shift in maximum. Both CD spectra showed X-axis contact at 246 nm, that remained constant over the temperature range.

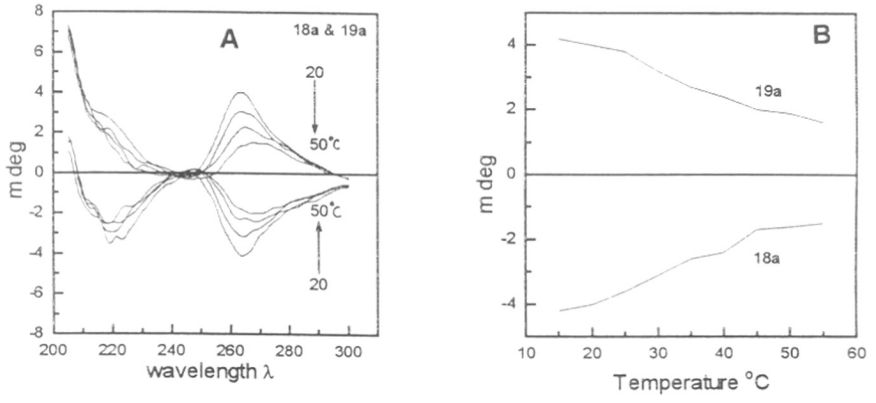


Figure 16: **A:** CD spectra of D-*trans*-PrNA **18a** and L-*trans*-PrNA **19a** at different temperatures. **B**: Melting profiles of **18a** and **19a.**

Since PNA without prolyl units is CD *trans*parent, the observed bands at 262 nm in single stranded Pro-PNA (**18a** & **19a**) must arise due to asymmetry induced in base stacking by the prolyl unit. The destacking of both prolyl PNAs at higher temperature leads to a decrease in the amplitude and shift of 262 nm in the signal. Going by the CD patterns observed in case of lysine conjugates,[20] the L-*trans*-Pro-T-PNA **19a** perhaps has a right handed helical structure while D-*trans*-Pro-T-PNA **18a** a left handed structure. Since the prolyl unit is located in the interior of the backbone, the C4 configuration may influence the helicity of the structure towards N-terminal side of PNA (only 3 more bases) while the C2 configuration may influence that towards carboxy terminus side (up to 9 bases). With the present set of experiments, it is difficult to know whether such a situation exists and if so how to differentiate between these two possibilities and their relative contribution to the overall CD. A systematic study of circular dichroism as a function of the position of prolyl substituent and its stereochemistry may give more insight.

3.5.2d CD spectral study of PNA-DNA duplexes: Table 5 lists some of the CD observables for various prolyl PNA:DNA duplexes in parallel and antiparallel modes. As mentioned in the introduction section, CD studies of DNA hybrids with nonchiral PNA oligomers can be followed through CD bands of DNA component. Figure 17 and figure 18 show CD spectra of 10-mer and 13-mer PNA hybrids respectively in parallel orientation and these are similar but not identical to B-form DNA hybrids. The 13-mer PNA:DNA hybrids (**18a:26/27, 19a:26/27**) showed a longer positive band at 260 nm compared to the negative band at 240 nm, while in case of 10-mers (**15a:24/25, 16a:24/25**), both bands were of similar intensity. Upon heating, in both cases, the positive band at 260 nm decreased in intensity without much changes in the intensity of negative band. The cross over point on wavelength axis exhibited a shift to longer values by 12 nm for 13-mer as compared to only 2-3 nm for 10-mer. The 10-mer hybrid is only a AT sequence

Table 5 : CD parameters for PNA:DNA duplexes

Duplex	Temperature	Positive band(λ_{max})	Negative band(λ_{max})	Crossover
15a:24	15°	4.4 (263)	-2.35 (244)	250.5
	50°	3.8 (266.5)	-4.60 (247)	256.0
15a:25	15°	4.0 (263)	-1.5 (244)	250.0
	50°	2.8 (266)	-2.4 (246)	256.0
16a:24	15°	1.8 (266)	-5.1 (247)	261.0
	50°	1.7 (268)	-3.8 (247)	260.0
16a:25	15°	0.8 (272)	-3.5 (252)	266.0
	50°	0.8 (270)	-1.6 (248)	260.0
B-DNA		(270 and 220)	248	258
A-DNA		250	240	248

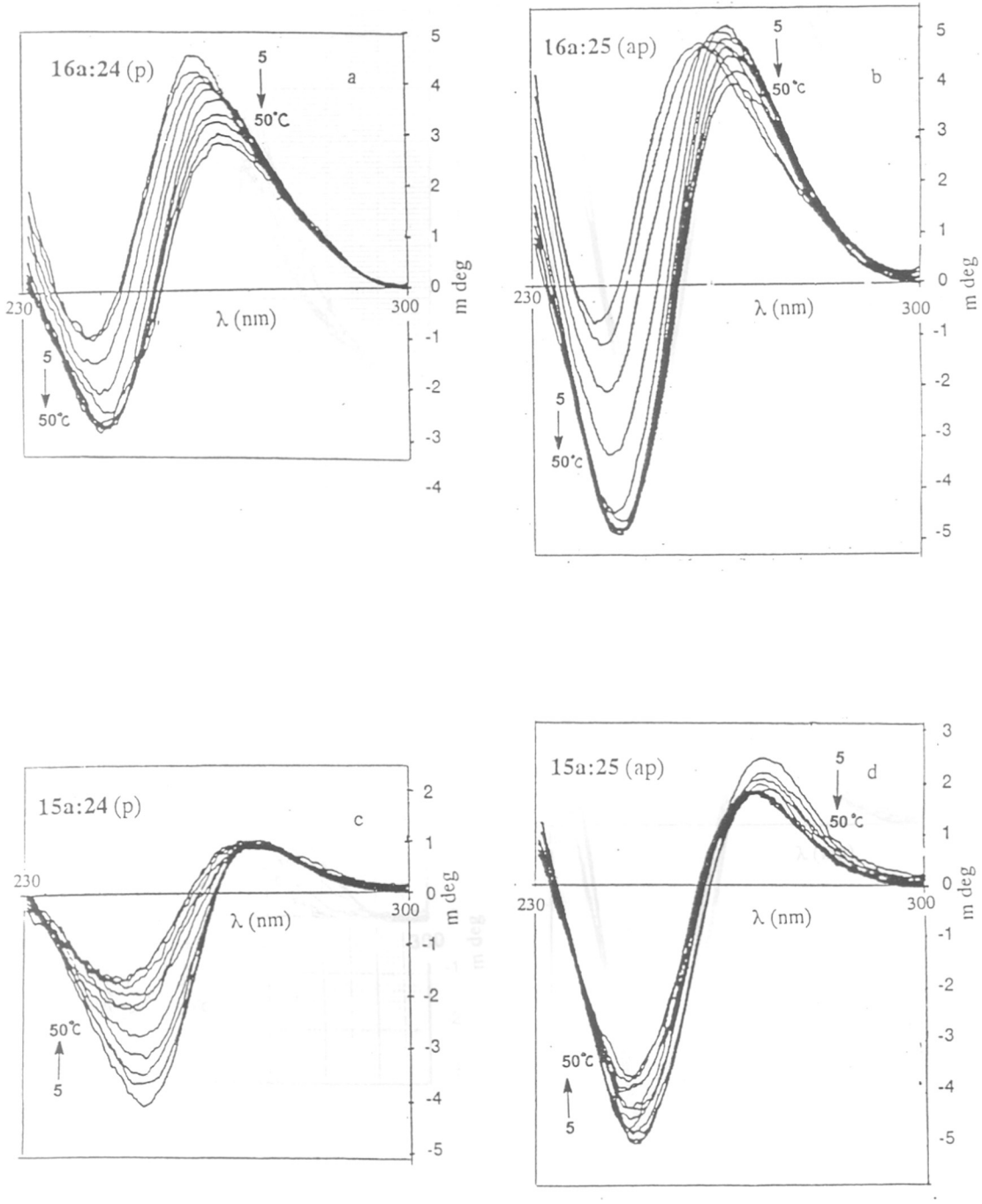


Figure 17: CD spectra of duplexes a: **16a:24** (p); b: **16a:25** (ap); c: **15a:24** (p) and d: **15a:25** (ap).

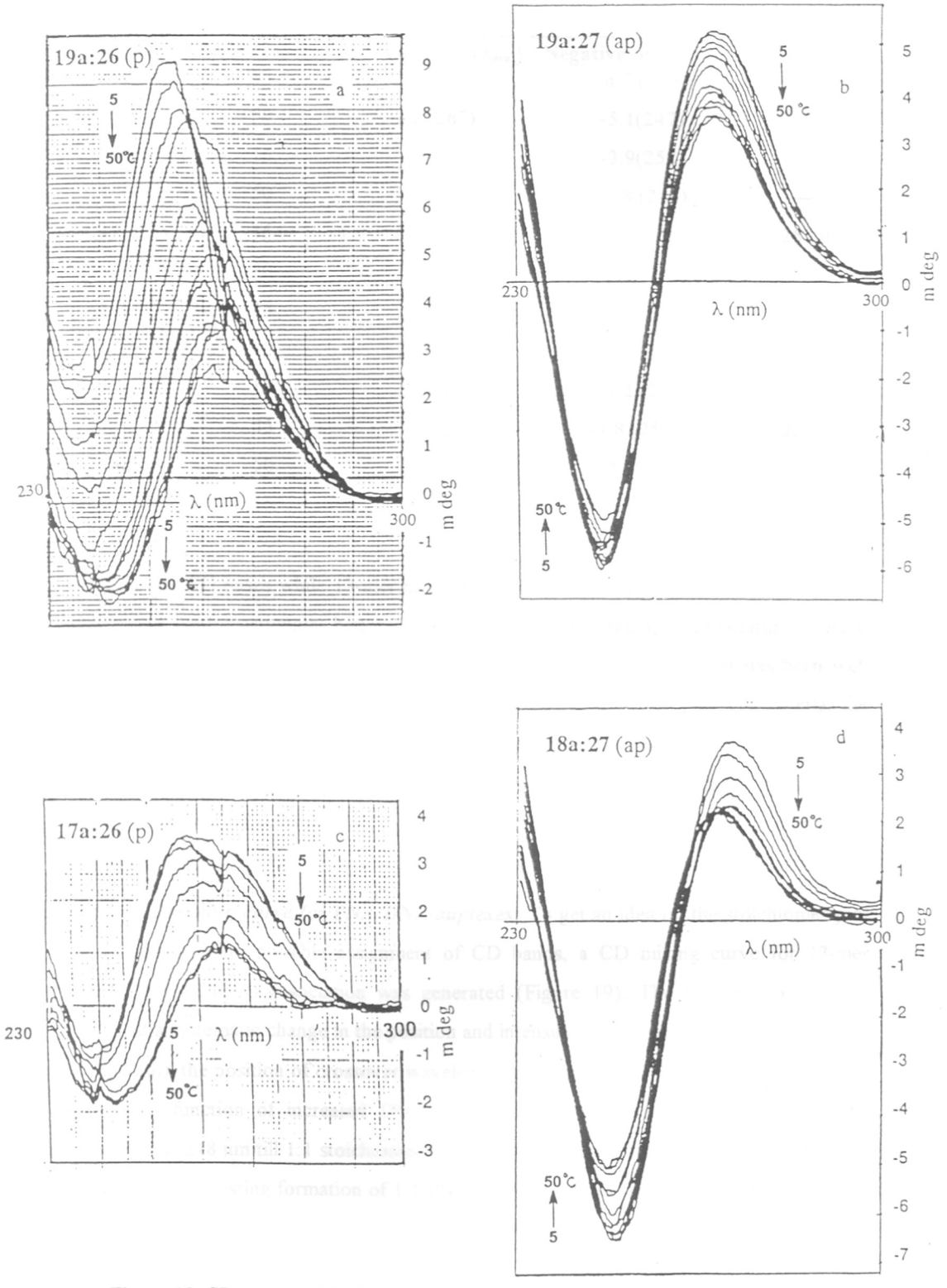


Figure 18: CD spectra of duplexes a: **19a:26** (p); b: **19a:27** (ap); c: **17a:26** (p) and d: **18a:27** (ap).

Duplex	Temperature	Positive band(λ_{max})	Negative band(λ_{max})	Crossover
18a:26	15	2.5 (271)	-6.7 (249)	262.0
	50	2.2 (267)	-5.1(247)	260.0
18a:27	15	--	-3.9(252)	--
	50	--	-1.8 (248)	--
19a:26	15	3.5 (267)	-5.6 (256)	261.0
	50	4.8 (267)	-5.2 (247)	258.0
19a:27	15	7.8 (255)	0.9 (235)	--
	50	3.2 (263)	-2.2 (248)	250.0
17a:27	15	3.2	-1.2 (239)	245.0
	45	1.2	-1.8 (252.5)	255.0
14a:25	50	1.6	-2.4 (247)	257.0
	15	1.3	-1.9 (247)	257.0

without any CG bases while 13-mer has terminal CTC.GAG triplet which makes it more stable and conformationally better defined. Accordingly, the CD-temperature showed significant changes. A shift of 3-5 nm in the long wavelength crossover has been well recognized as a good indicator of denaturation in DNA duplexes.[39] By this criteria, the observed temperature dependent shift of 12 nm in crossover band at 245 nm of 13-mer PNA:DNA (**17a:27**) can be used to determine CD-T_m which was 25 °C. In comparison, the UV-T_m for this duplex was 28 °C. In tune with the shift in crossover, both positive and negative bands of 13-mer PNA-DNA (**17a:27**) duplex also shifted to higher wavelengths.

*3.5.2e **CD mixing curve for PNA:DNA duplexes:*** To get an idea on the stoichiometry of complexation and possible assignment of CD bands, a CD mixing curve for 13-mer PNA:DNA (**17a:27**) association was generated (Figure 19). The CD titration spectra revealed a systematic change in the position and intensity of long wavelength positive band (>260 nm), the position of crossover wavelength and the intensity of negative band (~240 nm) as a function of increased DNA concentration. An isodichroic point that was observed at 248 nm till 1:1 stoichiometry, disappeared with addition of DNA beyond one equivalent, suggesting formation of 1:1 PNA:DNA complex. The spectral profile showed

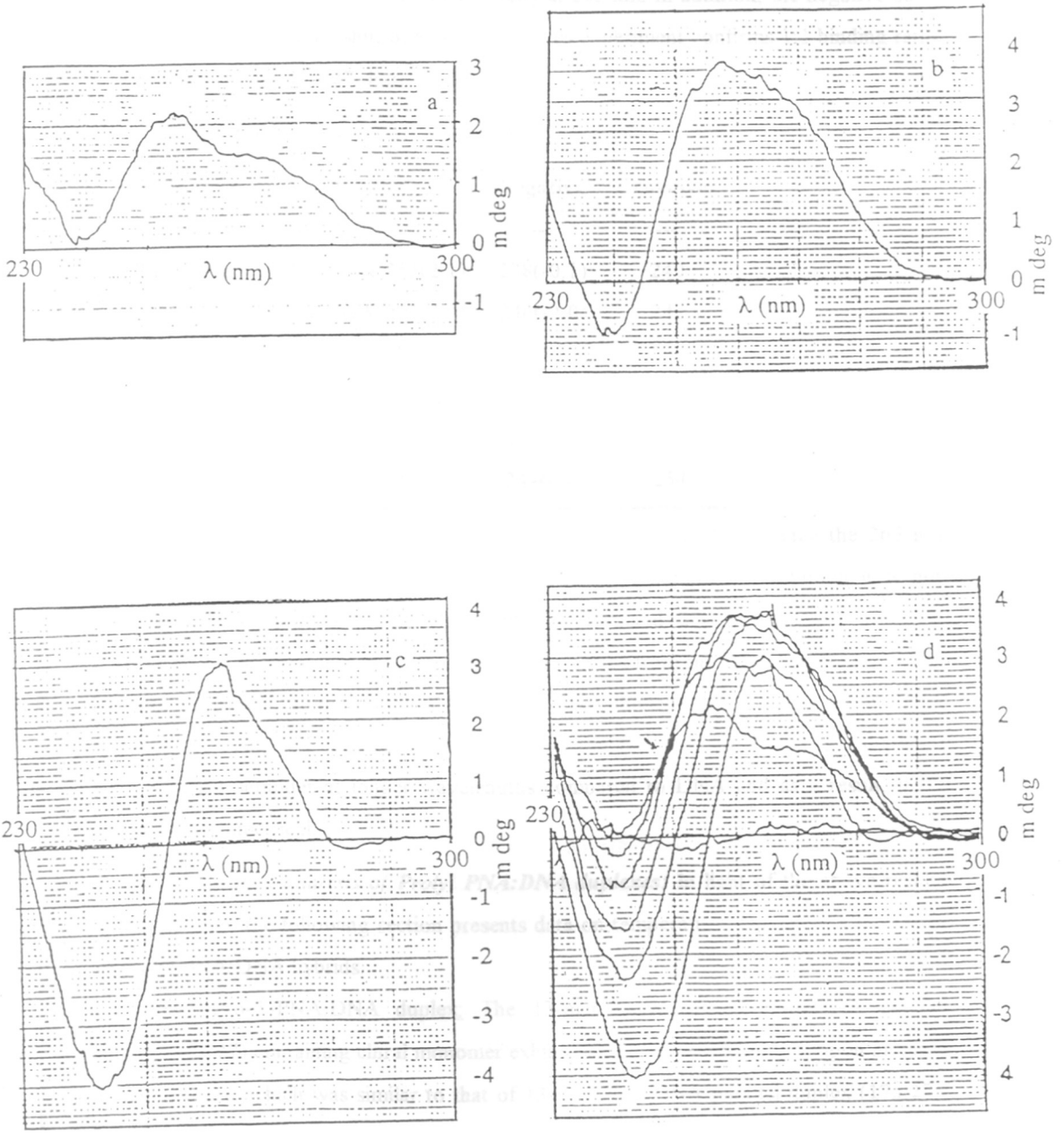


Figure 19: CD mixing experiment spectra of PNA **17a**:DNA **27** in a: 80: 20, b: 50: 50, c: 0: 100 and d: various ratios.

interesting patterns. At PNA:DNA ratios of 1:4, with all DNA perhaps present as duplex, two positive bands at 255 and 270 nm and a negative band at 238 nm were observed. With increase in DNA concentration, the 255 and 270 nm bands started moving towards each other and merged beyond 1:1 stoichiometry at 265 nm. In addition, the negative band at 238 nm enhanced with a shift of about 4 nm. This experiment confirms 1:1 binding ratio in PNA:DNA duplex **17a:27**. Further, formation

Table 6 :CD data on PNA:DNA mixing experiment.

% DNA	Positive	Negative	Crossover
0 %	---	---	---
20 %	270(1.4) & 255(2.1)	238(-0.1)	240
34 %	265(2.8) & 257(3.0)	240(-2.0)	242
50 %	265(sh) & 257(3.6)	240	244
66 %	264(sh) & 257(3.6)	240	247
80 %	265(3.6) & 257(sh)	242(2.2)	249
100 %	265(3.0)	244(4.4)	254

of complex is accompanied by changes in the conformation of DNA since the 265 nm observed with 100% DNA is absent in the PNA:DNA complex. Although it is not possible at this stage to definitively assign the 255 and 270 nm positive bands in PNA:DNA 4:1 spectra, it is likely that the 255 nm arises from the DNA component while 270 nm originates from the helical stacks of PNA nucleobases resulting upon DNA duplexation. This has supporting evidence from the fact that all PNA-DNA duplexes show a broad shoulder at longer wavelengths compared to DNA and as a result, these bands are asymmetrical.

3.5.2f CD spectral analysis of Prolyl PNA:DNA duplexes: In light of the above results and discussion, the following section presents data on temperature dependent CD spectra of prolyl PNA:DNA hybrids.

(i) L-*trans*-Pro-T-PNA:DNA duplex: The 13-mer L-*trans*-Pro-T-PNA:DNA parallel duplex (**19a**:**27**) containing chiral monomer exhibited a CD behavior pattern upon increase in temperature, which was similar to that of 13-mer PNA:DNA parallel duplex (**17a**:**27**).

A systematic shift in the long λ band at 255 nm accompanied by a decrease in intensity and a shift in cross over point on wavelength was noticed. The spectral changes are dominated in the positive amplitude region of 260 nm as both L-*trans*-Pro-T-PNA **19a** and DNA (**27**) have strong positive bands in this region. In contrast to this, L-*trans*-Pro-T-PNA:DNA (**19a**:**27**) antiparallel duplex showed a strong negative band around 245 nm which started decreasing in intensity after T_m while the positive band at 265 nm showed a systematic decrease in intensity with temperature.

(ii) D-*trans*-Pro-T-PNA:DNA antiparallel duplex: The low temperature CD spectrum is dominated a large negative band arising from D-*trans* unit of PNA at 248 nm. Increase in temperature effected changes in intensity of both positive and negative bands but in different ways. Below T_m, the positive bands starts decreasing, becoming constant at 20 °C and beyond T_m, as determined by UV, where the negative band starts showing changes. Shifts of about 2 nm were noticed in each case. The 10-mer D-*trans*-Pro-T-PNA:DNA (**18a:26**) antiparallel duplex showed an identical melting behavioral change.

3.5.2g Conclusion on CD spectral analysis: The results and discussion presented in this section does not lead to definitive conclusions about correlation of the stereochemistry of the backbone monomers with conformation of derived oligomers and their duplexation with DNA. The complication in analysis arise due to the presence of strong bands from the chiral PNA which show changes in ellipticity with temperature. However, the different CD patterns and their response to temperature changes are distinctly a certain function of the monomer stereochemistry and clearly establish that the induced handedness in helical base stacks widely differ in various stereoisomeric PNAs. A systematic design of sequences and assignment of CD bands to electronic *trans*itions and base stacking geometries may allow a better insight into the conformational aspects. *Finally, it is sufficient to point out that the inclusion of a single backbone constrained, distended chiral monomeric unit such as prolyl monomer as a part of a PNA chain does lead to stabilization of PNA:DNA hybrid as seen from temperature dependent UV and CD spectroscopy.*

3.6 Conclusions

In this chapter the synthesis of novel chiral pyrimidine homooligomers of PrNA are reported using chiral building blocks. The biophysical studies (UV and CD) show that they have contrain on the backbone due to which the binding with complementary DNA is not efficient. Further, when these chiral monomers were introduced either at the end or in the middle of aeg-PNA strand, they have shown improved binding properties and orientational preference. The PNA containing a L-*trans*-Pro-T unit in the middle preferentially binds in parallel orientation with complementary DNA and PNA containing a D-*trans*-Pro-T preferentially stabilizes the antiparallel orientation. The PNA-spermine conjugates exhibit remarkably high selectivity in antiparallel binding (with complementary DNA) without affecting the self-order or parallel orientation. The presence of a cationic spermine chain at PNA terminus may have additional utility for improving the poor cell uptake of PNA oligomers in view of the presence of polyamine receptors on cell surface. It is demonstrated that the incorporation of intrinsically fluorescent 2-aminopurine does not effect the binding property of the PNA with the complementary DNA and further this fluorophore can be used to monitor the structural/ conformational variations during hybrid formation and melting. Apart from their utility in diagnostics, these PNA probes may lead to newer capabilities such as study of cellular uptake and intracellular distribution of PNA by employing fluorescence microscopy and as sequence specific DNA biosensors.

3.7 Experimental

Functionalization of the resin: N-*t*-Boc β-alanine was dissolved in ethanol (10 mL) and neutralized with saturated aqueous solution of cesium carbonate.[21] The solvent ethanol was evaporated under vacuum and the residue was dried by co-evaporating with toluene as an azeotrope. The cesium salt of β-alanine and Merrifield resin (3 g, 0.7 m eq./gm) were suspended in DMF (3 mL) and slowly stirred at 60 °C in an oil bath. After 24 h., the resin was filtered, washed successively with DMF (20 mL), water (20 mL), DMF (20 mL), methanol (50 mL), DCM (50 mL) and dried under vacuum.

Estimation of the amino acid loading:[18a] The funtionalized dry resin (5 mg) was taken in a sintered funnel and soaked in DCM for 1 h. The solvent was filtered off and treated with 50% TFA/DCM (2 x 1 mL) for 20 min., followed by filtration, washing with DCM, neutralization with 5% TEA/DCM and further washing with DCM. This resin with free amine was treated with 0.1 M picric acid in DCM (3 x 2 mL) and washed with DCM to remove unbound picric acid. The picrate bound to amino group was eluted with 5% TEA/DCM followed by DCM (3 x 2 mL, 2 min.), An aliquot (0.2 mL) of picrate eluant was diluted to 2 mL with ethanol and optical density measured at 358 (ε_{358}=14,500). From this the loading was calculated by dilution factor to be 0.35 meq./g.

General Method for solid phase synthesis

All the peptides were assembled using solid phase peptide synthesis. Merrifield resin[18] preloaded with β-alanine (0.35 m eq/g) was used in all the oligomer PrNA and PNA preparations. All *t*-Boc-protected amino acids were cleaved with 50% TFA/DCM. Benzyloxycarbonyl group was used in side chain protection of cytosine monomer which was removed in the end by TFMSA treatment.

Synthetic protocol for solid phase synthesis: The resin was preswollen (overnight) and the following steps were performed for each cycle.[15]

Protocol : (i) The Merrifield resin (100 mg, dry weight) was washed with DCM 2 x1 min, 2 mL; (ii) *t*-butoxycarbonyl deprotection with 50% TFA/DCM (1:1 v/v), 1 x 2 min, 1 x 30 min, 2 mL; (iii) DCM wash 4 x 20 sec. 2 mL; (iv) neutralization with DIPEA/DCM (1:19 v/v) 2 x 3 min, 2 mL; (v) DCM wash 3 x 20 sec, 2 mL and DMF wash 3 x 20 sec, 2 mL;

(vi) addition of 4 eq. of DIPCDI or TBTU, 4 eq. of HOBT (with chiral prolyl-amino acids) and 4 eq. of the required *t*-Boc-amino acid; (vii) filtration of reagents followed by washing of resin with DMF 3 x 20 sec, 2 mL and DCM 3 x 20 sec, 2 mL. This cycle was repeated for every amino acid.

Kaiser's test:[19] The reagent comprises of (a) 0.5 g of ninhydrin dissolved in 10 mL of ethanol and (b) 8 g of phenol in 2 mL of ethanol and (c) 0.2 mL of a 0.001 M aqueous solution of potassium cyanide in 9.8 mL of pyridine. The test was carried out twice in each cycle, after step (iv) in which the positive result shows that there is a free amino group for coupling, and after step (vi), the negative results show that the coupling is efficient. If the coupling was incomplete at any stage, the resin was washed thoroughly with DMF and resubjected to coupling conditions.

A few beads of resin are taken in a test-tube and three drops of each solution was added and the solution heated at 100 °C for 2-5 min. The formation of intense purple-blue color implies the presence of free amine and a blank or negative test shows a yellow colored solution.

TFMSA Cleavage:[18b] The peptide resin (20 mg) was taken in a glass vial with thioanisole (18 μL) 1,2 ethanedithiol (13 μL) in an ice bath and stirred for 10 min. TFA (200 μL) is added and after equilibration for 10 min, TFMSA (30 μL) is added slowly. After 1 h the cleavage mixture is filtered through a sintered funnel and washed with TFA (0.5 mL). TFA was removed and the residue was taken in water (5 mL) washed with ether repeatedly to remove the scavengers. The water was lyophilized to obtain the crude PrNA or PNA.

Purification Methodology: The crude PNA/PrNA was dissolved in 200 μL of water and passed through a 5 mL syringe packed with 2.5 mL of G-15/G-25 sephadex matrix and eluted with water. The void volume was collected and further purified by FPLC using semipreparative C-8 reverse phase column and the buffer used were as follows; Buffer A: 0.1% TFA in water; Buffer B: 0.1% TFA in 50% acetonitrile/water.

Aminolysis: Cleavage with Spermine and 1,3-Propanediamine[22]

The peptide resin (20 mg) after *t*-Boc deprotection (unless otherwise mentioned) was taken along with 80 mg of the amine (spermine/1,3-propanediamine) in a 0.5 mL screw-capped vial and heated at 60 °C for 42 hr. The reaction mixture was diluted with water and filtered through sintered funnel and concentrated. The crude mixture was passed through G15/G25 sephadex column and eluted with water to remove the excess amine and further purified by FPLC as described earlier.

Synthesis of H(aeg T_6)-β-ala-MF resin (8)

The PNA oligomer was assembled on MF resin (100 mg, 0.035 mmol) preloaded with *t*-Boc-β-alanine. The stepwise manual synthesis was carried out according to the protocol 1 using aeg-T (42 mg, 0.1 mmol), DIPCDI (20 μL, 0.1 mmol) in 0.5 mL DMF for each coupling. The efficiency of coupling was monitored by Kaiser's test, which was negative indicating the completion of the reaction. The resin was cleaved with TFMSA to obtain the C-terminal carboxylic acid **8a**, further cleavage with spermine and 1,3-propanediamine gave the corresponding amides **8b** and **8c**.

Synthesis of H(L-*cis*-Pro-T)-β-ala-MF resin (9)

About 300 mg of MF resin preloaded with β alanine (0.3 meq/g) was taken in the peptide reaction vessel and soaked overnight, in DCM. The incorporation of the monomer followed the protocol 1 and coupling steps involved addition of 3 eq. TBTU (98 mg, 0.3 mmol), 4 eq. *t*-Boc(L-*cis*-ProT)-OH (125 mg, 0.3 mmol) and HOBT (41 mg, 0.3 mmol) were added in 1.0 mL of DMF. The resin 9 (25 mg) was cleaved by standard procedure using TFMSA and purified by gel filteration and FPLC purification using gradient A. The purity of the product **9a** was confirmed by single peak in HPLC and FAB-MS:

Synthesis of H(L-*trans*-Pro-T)$_6$-β-ala-MF resin (10)

The PrNA was assembled on β-ala-MF resin (200 mg) with stepwise incorporation of the PrNA monomers following the protocol 1. Coupling step involved use of 4 eq. TBTU (97 mg, 3.0 mmol), 4 eq. *t*-Boc(L-*trans*-Pro-T)-OH (125 mg, 0.3 mmol) and HOBT (40 mg, 0.3 mmol) were added in 1.0 mL of DMF. H(L-*trans*-Pro-T)$_6$-OH **10a**

was obtained by cleaving the resin (25 mg) by following TFMSA method described earlier. The peptide was purified by gel filtration and FPLC.

Synthesis of H(D-*trans*-Pro-T)-β-ala-MF resin (11)

The synthesis was started on 200 mg of *t*-Boc-β-ala-MF resin preswollen overnight on DCM and the couplings were carried out as in the protocol 1 using 4 eq. TBTU (98 mg, 0.3 mmol), 4 eq. *t*-Boc(D-*trans*-ProT)-OH (125 mg, 0.3 mmol) and HOBT (41 mg, 0.3 mmol) were added in 1.0 mL of DMF. The resin **11** (25 mg) was cleaved using TFMSA and the resulting peptide **11a** was purified by FPLC.

Synthesis of *t*-Boc(aeg-T_2AT_2)-β-ala-MF resin (12)

The synthesis of the 5-mer was carried out on 510 mg of resin as described in protocol 1, using 3 eq. DIPCDI (80 μL, 0.72 mmol) and 3 eq. PNA monomer (209 g, 0.75 mmol) in DMF (2 mL). After the 5th coupling the resin was washed with DMF and DCM and dried for further use.

Synthesis of *t*-Boc(aeg-$AT_2AT_2AT_2$)-β-ala-MF resin (13)

The synthesis was initiated on 500 mg of *t*-Boc(aeg-T_2AT_2)-β-ala-MF resin which was preswollen overnight in DCM. The synthesis of resin **13** followed protocol 1, using 3 eq. DIPCDI (80 μL, 0.72 mmol) and 3 eq. PNA monomer (200 mg, 0.75 mmol) in DMF (3 mL). The resin was washed with DCM and DMF and dried.

Synthesis of *t*-Boc(aeg-$TAT_2AT_2AT_2$)-β-ala-MF resin (14)

The synthesis was carried out on 100 mg of *t*-Boc(aeg-$AT_2AT_2AT_2$)-β-ala-MF resin (**13**) which was preswollen overnight in DCM. The synthesis of followed protocol 1, using 3 eq. DIPCDI (18 μL) and 3 eq. PNA monomer (40 mg) in DMF (0.5 mL). Cleavage of resin **14** with TFMSA and spermine gave the corresponding peptides **14a** and 14b which were further purified before use by FPLC.

Synthesis of N4-*t*-Boc(D-*trans*-Pro-T-aeg-$AT_2AT_2AT_2$)-β-ala-MF resin (15) and N4-*t*-Boc(L-*trans*-Pro-T-aeg-$AT_2AT_2AT_2$)-β-ala-MF resin (16):

The synthesis was separately performed on 100 mg of resin (**13**) each and coupled with DIPCDI (18 μL), HOBT (14 mg) and *t*-Boc(D-*trans*-Pro-T)/*t*-Boc(L-*trans*-Pro-T)(40 mg) respectively to obtain **15** and **16** respectively. The resin **15** and **16** were cleaved

with TFMSA to obtain the C-terminal acid derivatives **15a** and **16a** which were further purified by usual methods. Cleavage of resin with spermine **15a** and **16a** gave the desired spermine conjugated PNA analogues **15b** and **16b** which were further purified by known methods.

Synthesis of *t*-Boc(aeg-CTCAT$_2$AT$_2$AT$_2$)-β-ala-MF resin (17)

The synthesis was achieved on 50 mg of preswollen resin **14** by addition on aeg-C(CBz)/aeg-T monomer following the protocol 1. After the synthesis was over the resin was washed and dried and divided into two portions. One portion was cleaved with TFMSA method to obtain the terminal acid **17a** and the other was cleaved with spermine to obtain terminal spermine amide **17b**. Both were purified by standard procedures for further use.

Synthesis of *t*-Boc(aeg-CTC{D-*trans*-Pro-T}aeg-AT$_2$AT$_2$AT$_2$)-β-ala-MF resin (18) and *t*-Boc(aeg-CTC{L-*trans*-Pro-T}aeg-AT$_2$AT$_2$AT$_2$)-β-ala-MF resin (19)

The synthesis was separately performed on 50 mg of resin **15** and **16** each and coupled with DIPCDI (10 μL) and aeg-T (25 mg)/aeg-C(CBz)(30 mg) respectively to obtain **18** and **19** respectively. The resin **18** and **19** were cleaved using TFMSA method to obtain the oligomers **18a** and **19a**. Aminolysis of resin **18** and **19** with spermine yielded the spermine analogues **18b** and **19b**. All these oligomers were purified by FPLC.

Synthesis of fluorescent PNA *t*-Boc-(aeg-TATT{2AP}TTATT)-β-ala-MF (20)

The synthesis was initiated on 100 mg of *t*-Boc(aeg-T$_2$AT$_2$)-β-ala-MF resin (**12**) which was preswollen overnight in DCM. The synthesis of **20** followed protocol 1, using 3 eq. DIPCDI (18 μL) and 3 eq. PNA monomer (41 mg) in DMF (0.5 mL). The resin **20** was cleaved with TFMSA to obtain the terminal carboxylic acid **20a** and purified by standard methods. Cleavage with spermine and deprotection of the terminal *t*-Boc group yielded the fluorescent PNA-spermine analogue **20b** which was further purified by gel filtration and HPLC.

Oligonucleotide Synthesis, purification

All oligonucleotides were synthesized on 0.2 μM or on 1.3 μM scale on a Pharmacia GA plus DNA synthesizer using controlled pore glass support and base

protected 5'-O-(4,4'-dimethoxytrityl)deoxyribonucleoside-3'-O-[(diisopropylamino)-β-cyanoethylphopsphoramidite] monomers which was followed by deprotection with aqueous NH_3. All oligonucleotides were purified by gel filtration (NAP 25) and their purity checked on reversed phase HPLC C18 column using buffer system A: 5% CH_3CN in 0.1M triethylammoniumacetate (TEAA) and B:30% CH_3CN in 0.1 M TEAA using a gradient A to B of 1.5 %/min at a flow rate of 1 mL /min. All olives showed purity higher than 90% were used without further purification.

Melting experiments

Duplex and triplex melting experiments were carried out in the following buffers, Buffer A: 10 mM Phosphate, Buffer B: 10 mM cacodylate, 100 mM NaCl, 0.1 mM EDTA. Appropriate oligonucleotides and PNA each at a strand concentration of 3-6 μM based on UV absorbance calculated using molar extinction coefficients at 260 nm, A = 15.4, C =7.3, G =11.7, T = 8.8 $cm^2/\mu M$, were mixed, heated to 80°C for 3 min, cooled to room temperature and then stored at 4°C overnight. The A_{260} at various temperature were recorded using Perkin Elmer Lamda 15 UV/VIS spectrometer, fitted with temperature programmable Julabo water circulator, with a heating rate of 0.5 °C/min over the rage 5-80°C. Dry nitrogen gas was flushed in the spectrophotometer chamber to prevent condensation at low temperature. Microcal Origin software was used for data analysis. All the curves were fitted with a Boltzmann sigmoidal fit and the melting temperature determined from the peak of the differential curves.

Fluorescent Spectroscopy

Fluorescence measurements were done on a Perkin Elmer model LS-50B spectrometer attached to a Julabo water bath circulator for variable temperature. The fluorescent DNA samples dissolved in the buffer were excited at 308 nm and the emission monitored at 367 nm using a spectral band width of 5 nm. The kinetic data were collected for 15-30 min with a time constant of 1 sec. and 5 nm band width at 20 °C. The temperature was kept constant by circulating water through water jacketed cuvette holder. Emission wavelength was fixed at 367 nm with excitation at 308 nm. For fluorescence polarization and anisotropy measurements the sample was prepared with a ratio of

PNA:DNA 1:2 to ensure complete binding. The samples were excited at 308 nm and fluorescence signal at 367 nm monitored through crossed polarizers. The experiment was repeated at different temperature (15, 25, 35 and 50 °C)

The binding affinities of PNA to DNA was calculated from the fluorescence titration curves.[47] By using the fluorescence data, concentration of free probe was determined using eq. 2, where C_T is the concentration of the probe added, C_F is the concentration of the free probe, and I and I_o are the fluorescence intensities in the presence and in the absence of DNA, respectively. *P* is the ratio of observed fluorescence quantum yield of the bound to that of the free probe. The value of *P* was obtained from a plot of I / I_o Vs 1[DNA] such that it is the limiting fluorescence yield given by the y-intercept. The amount of bound probe (C_B) at any concentration was equal to $C_T - C_F$.

$$C_F = C_T(I / I_o - P)/ (1-P) \qquad (1)$$

Scatchard analysis was done according to the modified equation of McGhee and von Hippel [17]. Accordingly a plot r/C_F Vs r, where r is equal to C_B/[DNA], was drawn.

$$r/C_f = K_i(1-nr)[(1-nr)/[1-(n-1)r]]^{n-1} \qquad (2)$$

In equation (3), K_i is the intrinsic binding constant and n is the binding size in base pairs. These values were obtained from the best fit of the data by linear equation fit.

Circular Dichroism

Circular Dichroism spectra were recorded on either JASCO 500a spectrometer or JASCO 715 spectrometer attached to a Julabo water circulator for variable temperature. The samples were scanned with a scan speed of 20 nm/min, slit width of 1 nm, and time constant of 1 sec, in 210 - 300 nm range. Each spectrum was taken as an average of 4 scans using 5 mm cell.

3.8 References:

1. Milligan, J.; Malleucci, M. D.; Marten, J. C. *J. Med. Chem.* **1993,** *36*, 1923.
2. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, 90, 544.
3. Nielsen, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167.
4. Hamilton, S. E.; Iyer, M.; Norton, J. C.; Corey, D. R. *Bioorg. Med Lett.* **1996**, *6*, 2897
5. Hyrup, B.; Nielsen, P. E. *Chem. Soc. Rev.* **1997**, 73.
6. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.*, **1996**, *4*, 5.
7. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science*, **1991**, *254*, 1497.
8. Lagriffoul, P-H.; Egholm, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1088-1082.
9. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freler, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature,* **1993**, 566.
10. Lesnik, E. A.; Risen, M. R.; Driver, D. A.; Griffith, M. C.; Sprankle, K.; Freier, S. M. *Nucleic. Acids. Res.* **1997**, *25*, 568.
11. Schmid, N.; Behr, P. *Tetrahedron Lett.* **1993**, *36*, 1447.
12. Barawkar, D. A.; Rajeev, K. G.; Kumar, V. A.; Ganesh, K. N. *Nucleic Acids Res.* **1996**, *24*, 1229.
13. Sund, C.; Puri, N.; Chattopadhyaya, J. *Tetrahedron* **1996**, *52*, 12275.
14. Grant, G. A. *Synthetic Peptides: A User's Guide*, **1992** University of Wisconsin Biotechnological Centre, Biotechnical Resource Series, Wilt Freeman and Company, New York.
15. Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H., Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767.
16. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K.; Hansen, H. F.; Koch, C.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Peptide Sci.* **1995**, *3*, 175.
17. Gisin, B. F. *Helv. Chem. Acta.* **1970**, *56*, 1476.

18. Merrifield, R. B.; Stewart, J. M.; Jernberg, N. *Anal. Chem.* **1966**, *38*, 1905. (b) Erickson, B. W.; Merrifield, R. B. *Solid Phase Peptide Synthesis. In The Proteins.* Vol. II., 3rd Ed., H. Neurath and R. L. Hill, eds., Academic Press, New York, **1976**, pp. 255.

19. Kaiser, E.; Colescott, R.L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.

20. Kim, S. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

21. Hodges, R. S.; Merrifield, R. B. *Anal. Biochem.* **1975**, *65*, 241.

22. Baird, E. E.; Dervan, P. *J. Am. Chem. Soc.* **1996**, *118*, 6141.

23. Kanavarioti, A.; Baird, E. E.; Smith, P. J. *J. Org. Chem.*, **1995**, *60*, 4873-4883.

24. Prakash, T.P.; Barawkar, D.,A., Kumar, V. A.; Ganesh, K. N. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1733.

25. Puglisi, J. D.; Tinoco, I. Jr. *Methods Enzymol.* **1989**, *180*, 304.

26. Xu, D.; Evans, K. O.; Nordlund, T. M. *Biochemistry*, **1994**, *33*, 9592.

27. Buchardt, O.; Egholm, M.; Nielsen, P. E.; Berg, R. H. *WO 92/20/20702*, **1992**, 107.

28. Wang, J.; Palecek, E.; Nielsen, P. E.; Rivas, G.; Cai, X.; Shiraish, H.; Dontha, N.; Lere, D.; Farias, P. A. M. *J. Am. Chem. Soc.* **1996**, *118*, 7667.

29. Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1083.

30. Evans, K.; Xu, D.; Kim, Y.; Nordlund, T. M. *J. Fluoresc.* **1992**, *2*, 209.

31. Hochstrasser, R. A.; Carver, T. E.; Sowers, L. C.; Millar, D. P. *Biochemistry* **1994**, *33*, 11971. (b) Frey, M. W.; Sowers, L. C.; Millar, D. P.; Benkovic, S. J.*Biochemistry*, **1995**, *34*, 9185.

32. Kremsky, J. N. et al. *Tetrahedron Lett.* **1996**, *37*, 4313.

33. Millar, D. P. *Curr. Opin. Struct. Biol.* **1996**, *6*, 322.

34. Nordlund, T. M.; Andersson, S.; Nisson, L.; Rigler, R.; Graslund, A.; McLaughlin, L. W. *Biochemistry*, **1989**, *28*, 9095.

35. Tomac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Norden, B.; Graslund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544-5552.

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

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

38. Gray, D. M.; Ratliff, R. L.; Vaughan, M. R. *Methods Enzymol.* **1992**, *211*, 389.

39. Gray, D. M.; Hung, S-H.; Johnson, K. H. *Methods Enzymol.* **1995**, *246*, 19.

40. Koch T.; Naesby M.; Wittung P.; Jorgensen M.; Larsson C.; Buchardt O.; Stanley C.J.; Norden B.; Nielsen P.E.; Orum H.; *Tetrahedron Lett.*, 1995, *36*, 6933-6936.

41. (a) Bergmann, F.; Bannwarth, W.; Tam, S. *Tetrahedron Lett.* **1995**, *36*, 6823. (b) Petersen, K. H.; Jensen, K. D.; Buchardt, O.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1995**, *6*, 1119.

42. (a) Kosynkina, L.; Wang, W.; Liang, T. C. *Tetrahedron Lett.* **1994**, *35*, 5173. (b) Dueholm, K. L.; Peterson, K. H.; Jensen, D. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioorg. Med. Chem. Lett.*, **1994**, *4*, 1077.

43. Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Norden, B. *Nature* **1994**, *368*, 561.

44. Wittung P.; Eriksson P.; Lyng R.; Nielsen P.E.; Norden B. *J. Am. Chem. Soc.* **1995**, *117*, 10167.

45. Diederichsen, U. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 445.

46. Diederichsen, U.; Schmitt, H. W. *Tetrahedron Lett.* **1996**, *37*, 475.

47. McGhee, J. D.; von Hittel, P. H. *J. Mol. Biol.*, **1974**, *86*, 469.

CHAPTER 4

Synthesis of Polyoxyethylene-bis-Thymines and Their Photodimerization

4.1 Introduction

UV light induces adjacent stacked pyrimidine bases in DNA to undergo a photochemically allowed [$2\pi_s+2\pi_s$] cycloaddition of C5-C6 double bond to produce cyclobutane linked pyrimidine photodimers.[1] Among the various possible isomers of thymine photodimers, only two are normally formed in DNA: the *cis-syn* in both duplex and single stranded DNA (ssDNA) and the *trans-syn* formed only in ssDNA, the ratio of the products being 7:1 in favor of the *cis-syn* isomer.[2] The structures of these isomeric dimers of thymine are shown in Figure 1.[3] The dimers designated *cis-syn* and *trans-syn* are stereoisomers and since they are not mirror images, they are diastereomers. The *cis-syn* isomer has a mirror plane of symmetry and thus is not chiral, but is a *meso* compound. In contrast, the *trans-syn* dimer can exist as two mirror-image isomers that are of course, chiral. Thus, the *cis-syn* dimer is a diastereomer of each of the enantiomeric *trans-syn* dimers. In contrast, the *cis-anti* and *trans-anti* dimers do not have the same connectivity as the *cis-syn* and *trans-syn* dimers, so the *syn* and *anti* sets of isomers are not stereoisomers but rather regioisomers. The *trans-anti* dimer has a center of inversion and thus is not chiral. The *cis-anti* dimer exists as a pair of enantiomers, each of which is a diastereomer of the *trans-anti* dimer.

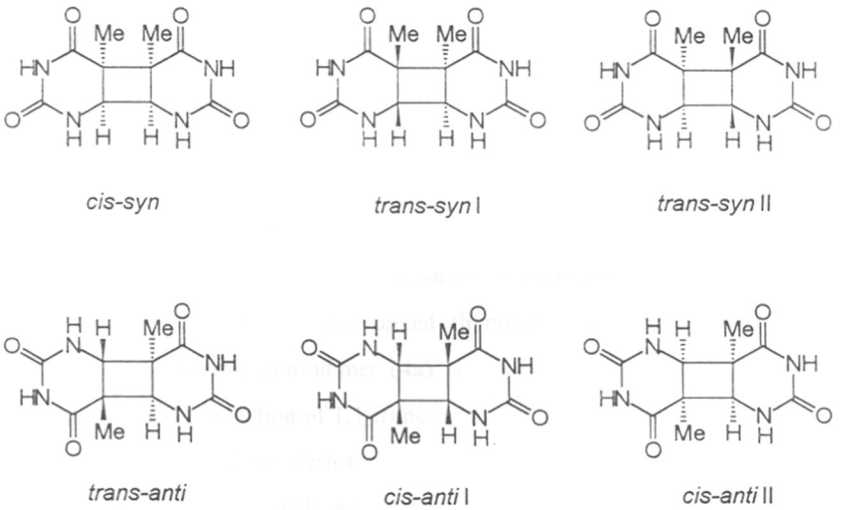


Figure 1

4.1.1 Background on T-T photodimers

Formation of *cis-syn* photodimer proceeds with a high quantum yield and causes lethal damage to cells by UV irradiation. The presence of *cis-syn* thymine photodimers results in structural distortion of the corresponding DNA region which blocks replication and transcription leading to cell death. The high frequency of skin cancer in patients with *Xeroderma pigmentosum,* a condition characterized by deficient capabilities for repairing UV-damaged DNA, shows dramatically that UV-induced lesions are directly involved in cancer development.[4] All cells contain efficient repair systems that are essential for the cells to survive exposure to the UV component of the sunlight.[5] The depletion of ozone in the stratosphere is a major cause of tremendous increase in the rate of skin cancer.[2] Ozone's remarkable effectiveness at reducing DNA damage is due to the fact that it has an absorption spectrum almost identical to that of DNA. More detailed calculations have concluded that for every 1% decrease in ozone level there will be a 4% increase in the skin cancer rate. DNA photolyase,[6] a DNA repair enzyme,[7] reverts pyrimidine photodimers with high specificity for the *cis-syn* photodimer.[8] Since they are not substrates for most DNA photolyases,[9] the *trans-syn* isomers have biological significance as lethal mutagenic lesions. It has been demonstrated recently that *Escherichia coli* DNA photolyase repairs *trans-syn* dimers 10^4 fold less effectively than *cis-syn* isomers.[10] The affinity of photolyase for *trans-syn* is identical to its non-specific affinity for DNA and the 10^4-10^5 fold lower affinity seen is sufficient to explain the failures in earlier detection of *trans-syn* by photorepair system.

Attempts to understand and model this reaction at simple nucleobase level have resulted in extensive research on the chemistry of *cis-syn* photodimers from natural and modified dinucleotides and on models possessing two thymine bases linked via non-sugar phosphate spacers. 1,1'-Trimethylene linked thymine photodimer[11] along with 3,3'-trimethylene linked thymine photodimer (**4a**) is extensively used by Begley *et al*[12] for mechanistic studies. The reaction of 1,1'trimethylene *cis-syn* thymine dimer with O-xylene dibromide gave the O-xylene derivative whose structure was determined by X-ray. Abbreviated dinucleosides[13] such as 5'-deoxy-5'(thymin-1-yl)thymidine (**2**) and the corresponding uridine derivatives were prepared and irradiated to produce internal *cis-syn*

photodimer and *trans-syn* photodimer in 1:1 ratio. N-Alkylated thymines are also used in determining the rate of photodimerization studies.

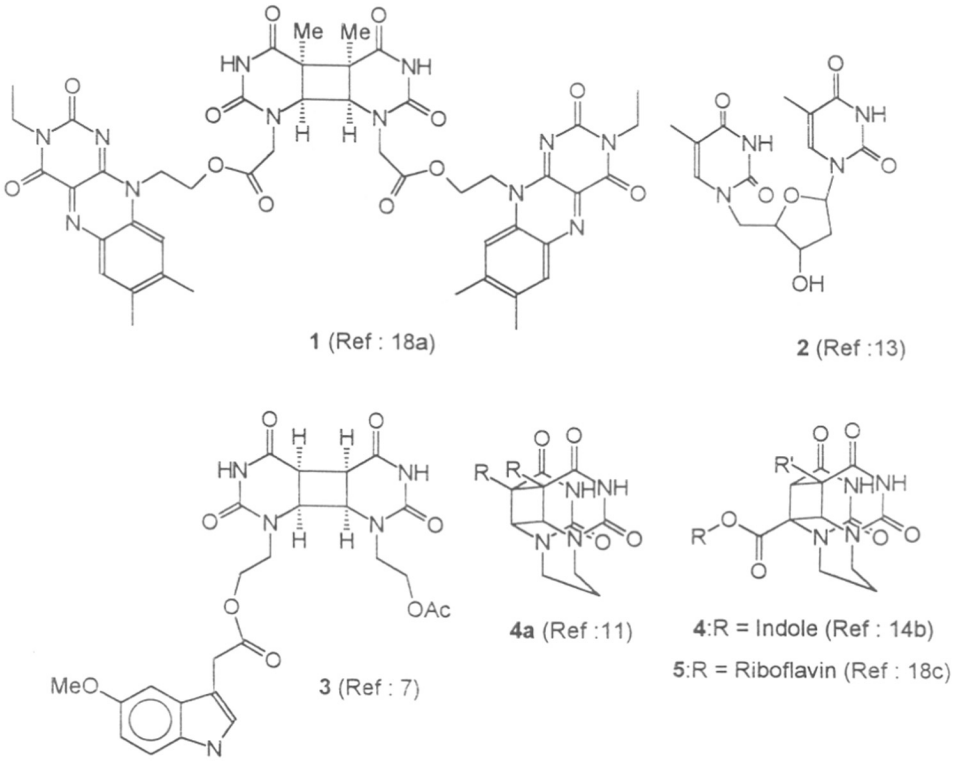


Figure 2

A number of studies on model compounds have been carried out to resolve the mechanism of photosplitting of thymine dimers by photosensitization. The sensitisers used in thymine dimer photosplitting are indole[14] (**3** and **4**), tryptophan,[15,16] quinone,[17] anthraquinone suphonate[11b] and riboflavin (**5**).[18] Rose *et al.*[20] have used such model compounds carrying photosensitizer to cleave the dimer. But all these models are composed of the *cis-syn* thymine dimer. There are very few reports on *trans-syn* thymine photodimers. Molecular recognition of *cis-syn* thymine photodimers using diaminopyridine

moieties as receptor molecules has been reported.[19] The attachment of a senstizer to such receptor molecules which bind to pyrimidine dimers with characteristic hydrogen bonding pattern showed efficient splitting of the photodimer.[20]

Extensive studies have been reported on site-specific incorporation of natural dinucleotide thymine photodimers *cis-syn*[21] and *trans-syn*[22] into oligonucleotides and used as substrates to examine the efficiency of repair enzymes to recognize them. These oligonucleotides are structurally distorted[23] by the presence of photodimers and these distortions were studied by solution structure determination.

Antibodies raised against a heterogeneous mixture of photoproducts obtained by DNA irradiation have been used to quantify induction of DNA photodamage and efficiency of repair mechanism of cells, using radio and enzyme linked immunoassays.[24] These antibodies (polyclonal and monoclonal) cannot discriminate between photoproducts *cis-syn* and *trans-syn* (I and II). In a recent study[24] antibodies raised against specific *cis-syn* and *trans-syn* damages were shown to distinguish between 2+2 and 6+4 products.

Devergne *et al.*[25] have studied the photodimerization properties of bisthymines linked through triethyleneoxy chain. The crystal structure of this compound showed that the pyrimidine base were close enough to form photodimer.[26] Although the photodimers were reported, the products were not characterized in terms of their stereochemistry.

4.1.2 Objectives

In view of the above background work, it was sought to link two thymine units through variable polyoxyethylene linkers for a systematic study of the effect of spacers on photodimerization properties.[27] The work components are

(i) synthesis of oligooxyethylene bisthymines

(ii) photodimerization and effect of linkers

(iii) structure determination by X-ray crystallography.

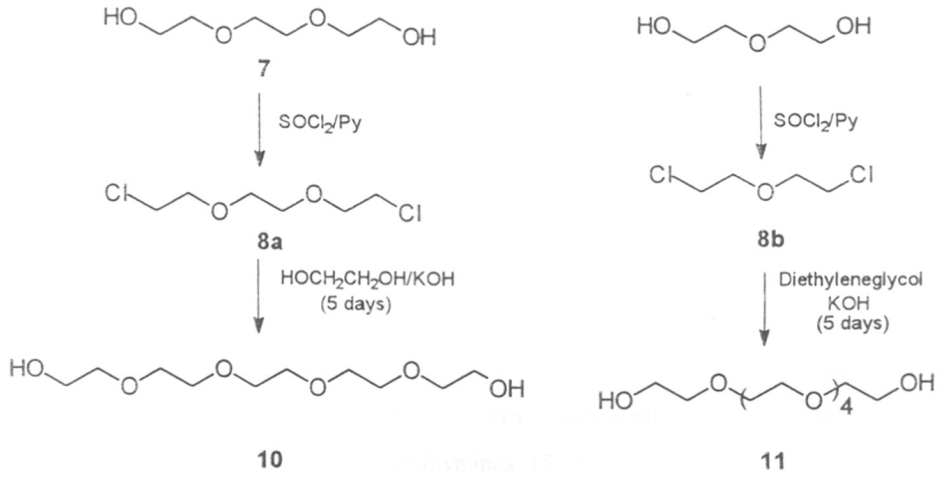
4.2 Present Work

Triethyleneoxybisthymine is known to photoisomerise under 295 nm exciting light with an efficiency 10 times lower compared to 1,1'trimethylenebisthymine.[25] Being polyoxy ether system it can co-ordinate alkali metal ions which may predispose the

thymine units for photodimerization. It was shown that the presence of lithium ion does not significantly affect the photoreaction and the stereochemistry of the product obtained from the photodimerization was not deciphered. In the present work oligoethyleneoxy bisthymines were prepared to investigate the effect of the oligoethylene glycol (n = 3-6) linker on photodimerization in presence of various salts.

4.2.1 Synthesis of oligoethyleneoxybisthymines and catecholethylenebisthymines

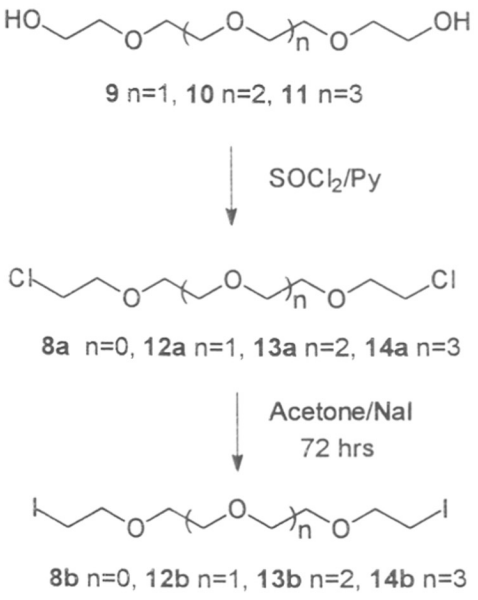
The commercially available polyethyleneglycols are contaminated with appreciable amounts of lower analogues. Hence the higher oligoetheyleneoxy glycols (**10** and **11**) were prepared starting from lowers analogues. The pentaethyleneglycol (**10**) was prepared from triethyleneglycol dichloride (**8a**) by treating it with excess of ethyleneglycol in presence of KOH (Scheme 1). Similarly, treatment of excess diethyleneglycol with 1,5-dichloro-3-oxapentane (**8b**) in presence of KOH gave the hexaethyleneglycol **11**. The oligomers **10** and **11** were purified by distillation under high vacuum and their purity checked by GC and MS.



Scheme 1

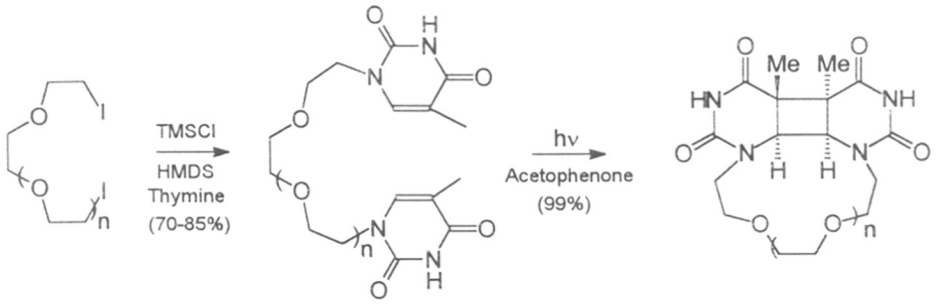
The oligoethyleneoxyglycols (**7**, **9-11**) were converted to their corresponding dichlorides (**8a**, **12a-14a** respectively) using thionyl chloride in pyridine. The alkylation of thymine in presence of a base may lead to N1 and N3 monoalkylated and N1,N3-

dialkylated products. It is well established in the literature# that the alkylation of 2,4-O-disilylthymine yields only the N1-alkylated products. Thymine was refluxed with TMSCl in HMDS to obtain the 2,4-O-disilylthymine which was not isolated and reacted further with the polyethyleneoxy alkylhalides to obtain the thymin-1-yl derivatives.



Scheme 2

The dichloride **8a** when reacted with thymine in HMDS and TMSCl gave poor yields of the corresponding bisthymines, the major product being mono-thymine oligoethyleneoxychloride. Hence the chloro compounds **8a**, **12a-14a** were converted to the corresponding iodo compounds (**8b**, **12b-14b** respectively) by refluxing them with NaI in acetone. These iodo compounds (**8b**, **12b-14b**) were treated with thymine in HMDS and TMSCl to obtain the required bisthymines **15-18** in good yields (75-80%). The monoalkylated derivatives were formed easily and the desired dialkylated derivatives were obtained in good yields only after continuous refluxing for 48-72 h. All the bisthymines were well characterized by ^{1}H and ^{13}C NMR.



8b n=0
12b n=1
13b n=2
14b n=3

15 n=0
16 n=1
17 n=2
18 n=3

19 n=0
20 n=1
21 n=2
22 n=3

Scheme 3

To observe the effect of the constraint in the linker, it was envisaged to introduce a double bond in one of the ethylene group of the bis thymine. Introduction of a benzene ring in the triethyleneglycol-linked bisthymine would give some rigidity to the backbone and hence might influence the stereochemistry of the resulting photodimer. With this rationale, catechol was used to prepare the catechol-ethylenebisthymine derivative **25**. 2-chloroethanol was protected with dihydropyran to obtain 2-chloroethyloxytetrahydropyran which was treated with catechol in presence of KOH in DMSO, to obtain catechol-O,O-*bis*ethane-2-oxytetrahydropyran (**23**). The reagent triphenylphosphine-bromine is known to give the corresponding bromocompounds from the THP-ethers. The compound **23** was derivatized to yield the corresponding bromocompound **24** using this reaction. The dibromide was reacted with thymine in HMDS and TMSCl to obtain catechol-O,O-ethane-2-*bis*thymine (**25**). The compound **25** was crystallized from methanol by slow evaporation at ambient temperature to obtain highly pure X-ray quality crystals.

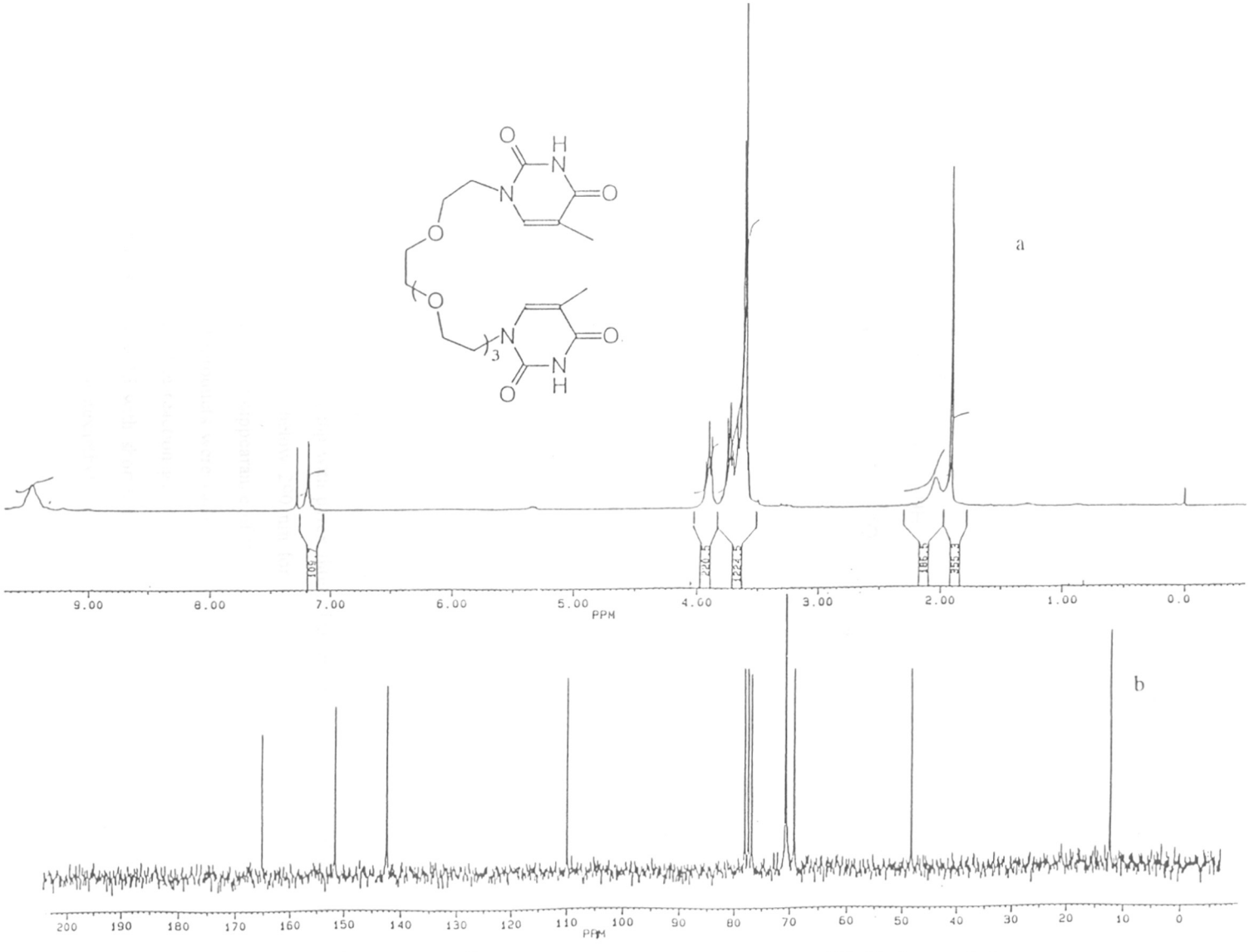
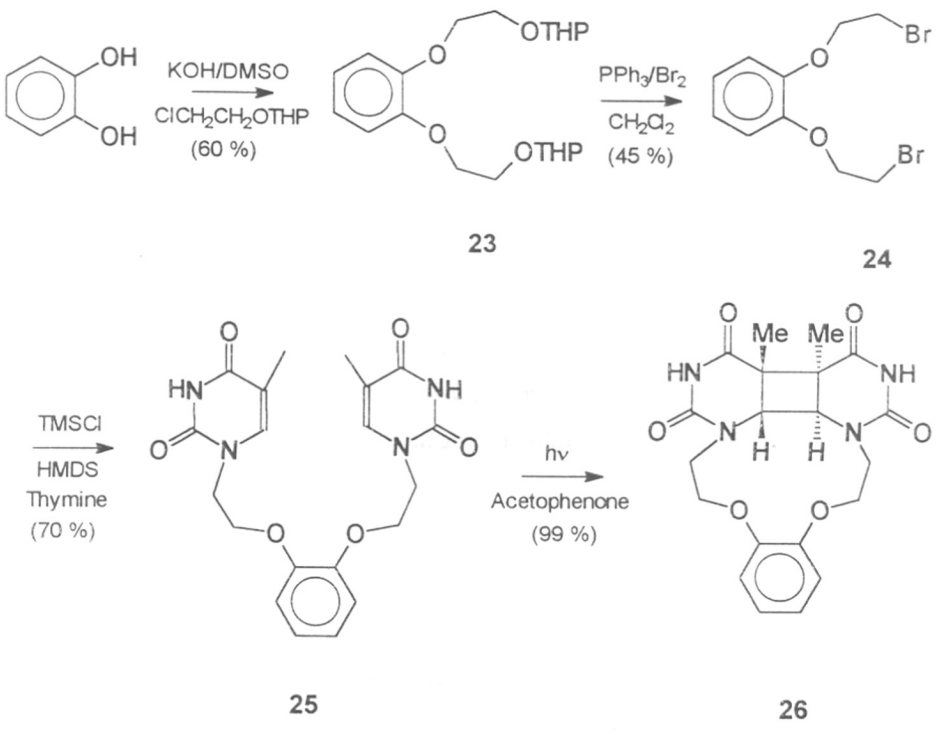


Figure 3. a: ^{1}H NMR and b: ^{13}C NMR spectra of compound **18** in $CDCl_3$ and DMSO-d_6 respectively.



Scheme 4

4.2.2 T-T photodimerization of poyethylenoxy bisthymines

The solubility of the bis thymines (**15-18**, **25**) in water was very low and hence these were taken in mixture of acetonitrile:water (7:3, v/v) for photodimerization. Acetophenone was used as a photosensitizer. The photodimerization was carried out in presence of Hanovia P-451 lamp with pyrex filter. The compounds **15-18** and **25** were irradiated with UV radiation below 290 nm for 4-6 h and the product formation was followed by following the disappearance of the starting material in the reaction mixture using TLC. The photoproducts were isolated by passing through a pad of silica gel to remove acetophenone. The reaction gave almost quantitative yield. It was observed that the bisthymines **15** and **25** with shorter triethyleneoxy spacer dimerised faster than the longer tetra **16**, penta **17** or hexaethyleneoxybis thymines **18**. The rate of dimerisation of

bisthymines with longer spacer chain was not only very slow but incomplete even after 12 h of irradiation. In all cases, the most interesting feature of photodimerzation was the exclusive formation of a single isomer as evident from HPLC. The orientation of the two thymine groups is identical in all the compounds as the spacer chain does not seem to affect the stereochemistry of the photoproduct. The photoproduct **21** having a penta ethyleneoxy linker still shows the formation of *trans-syn* photodimer as determined by X-ray structure. The presence of cations like Na^+, K^+ or Li^+ does not affect the stereochemistry of the products. The open chain oligoethyleneglycols are known to complex alkali metal ion, though weaker, compared to cyclic crown ethers. The orientation of the bulky thymine group might not allow proper organization of the spacer chain for metal ion complexation.

4.2.3 Spectral properties

On photodimerization, the C5-C6 double bond of thymine disappears and the molecule is no longer aromatic. The only choromophore present is the amide function which shows a UV band at 230 nm. The formation of a dimer was evident by the disappearance of 260 nm band in UV absorption spectrum. There is not much change in the 230 nm band implying that the thymines are in keto form (Figure 4). This was further confirmed by X-ray crystallography.

The saturation of C6-C5 double bond due to photodimerization causes drastic changes in both ^1H and ^{13}C NMR spectra. ^1H NMR showed an upfield shift for H6 from 7.10 ppm to 3.40 ppm resulting from saturation of the double bond. The 3.80-3.40 ppm region exhibited a complex multiplet mostly due to through space coupling effects of the ethylene chains which are compactly packed after dimerisation. This is also observed in the crystal packing of the dimers (see later). T-CH_3 also shows an upfield shift from 1.98 ppm to 1.50 ppm.

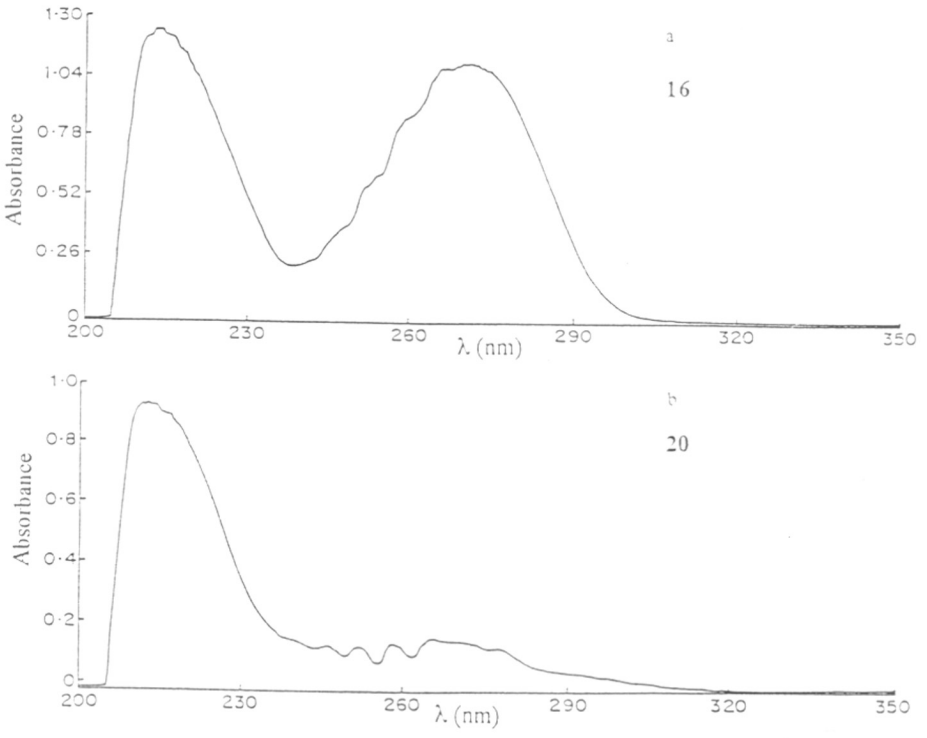


Figure 4: UV spectrum of compound **16** before dimerization and dimerized product **20**

The formation of dimer was also reflected in ^{13}C NMR spectrum, where a drastic shift in C6 and C5 δ values again due to saturation of the C6-C5 double bond. The C-6 has shifted from 108 ppm to 63 ppm and C5 has shifted from 142 ppm to 48 ppm. The T-CH_3 shows a downfield shift from 12 ppm to 21 ppm. Though the formation of the dimer was very well established, the absolute stereochemistry of the photodimers was solved only by X-ray structure.

4.3 X-ray structural studies of T-T photodimers

The refined atomic co-ordinates for non-H atoms are listed in Table 2-5 for compounds **15**, **19**, **25** and **26** respectively. The corresponding *ORTEP*II diagrams[28] for molecular structures along with atom labeling are shown in Figures 7 and 8 and selected interatomic distances and bond angles are tabulated in Table 3. Some dihedral angles relevant to our discussion appear in Table 7. Hydrogen bond parameters and information

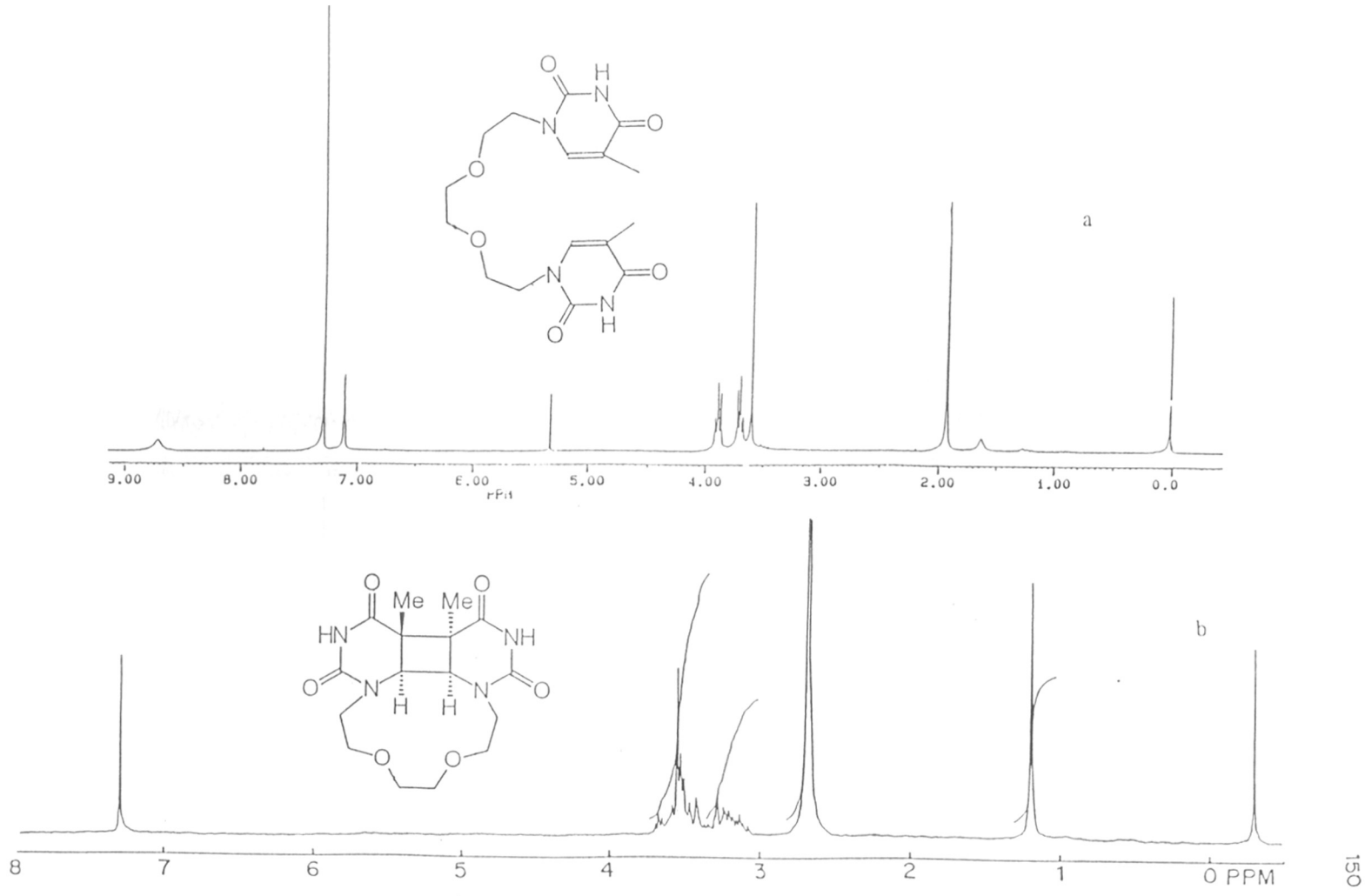


Figure 5: ^{1}H NMR spectra of compounds a: **15** and b: **19** in $CDCl_3$.

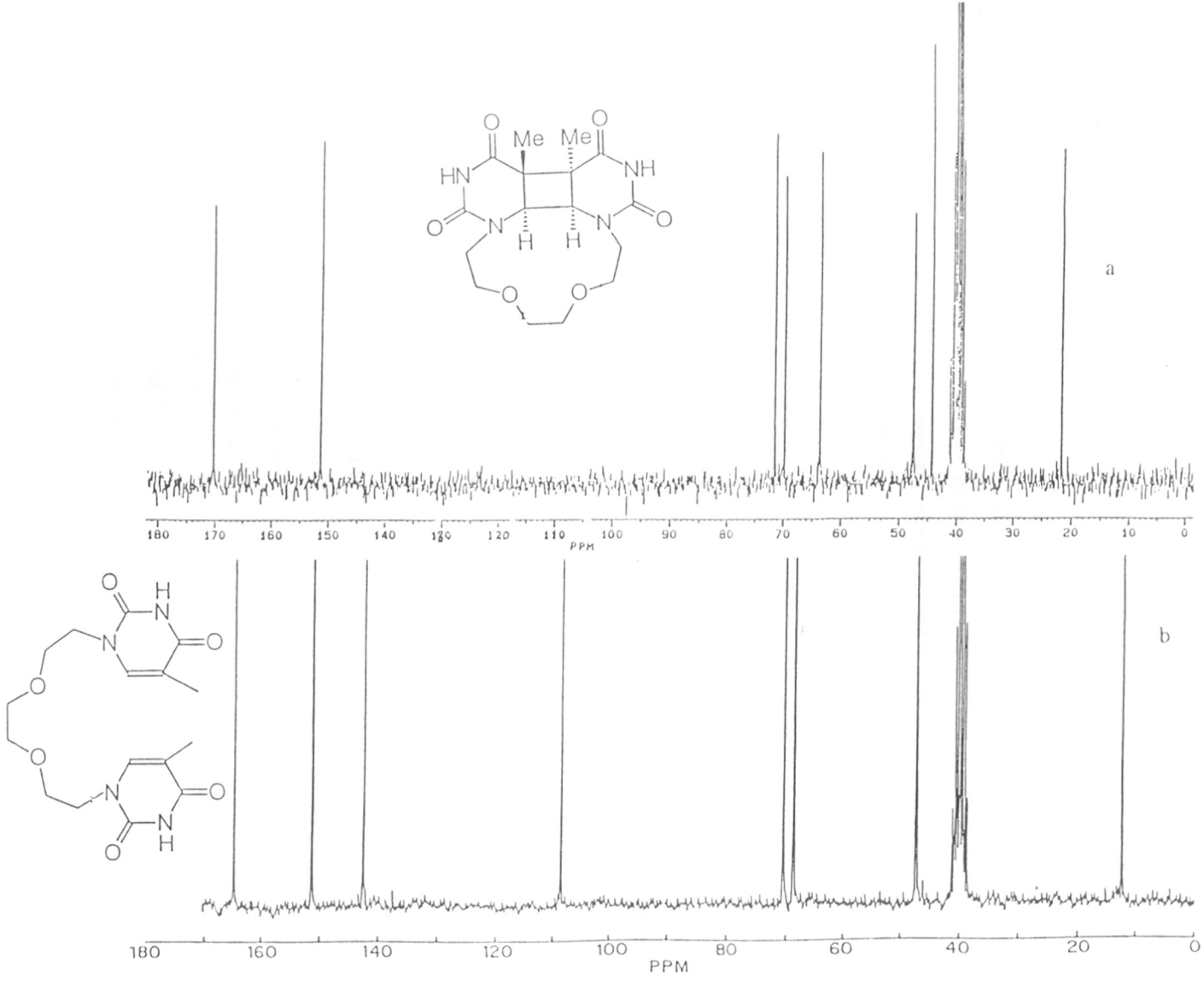


Figure 6: ^{13}C NMR spectra of compounds a: **19** and b: **15** in DMSO-d_6.

on short atomic contacts for the structures reported here are provided in Table 8. Although for each of the compounds one half of the molecule is identical to the other, it was interesting to note that none depicted this symmetry in their crystal structures. In each structure, the asymmetric unit was formed by a complete molecule. The thymines were in keto form in all compounds and the pyrimidine rings were planar in **15** and **25**. The pyrimidine bond lengths and angles were comparable to that of thymine in its crystal structure.[29] The C4-C5 bond, although single, acquired partial double-bond character due to the neighboring double bond and is still the longest intraring bond in the pyrimidine, while the C5-C6 double bond is the shortest. Others have intermediate lengths. Out of the six intraring bond angles, those at positions 1,3 and 6 were wider than the rest. In the crystal structure of **15,** the thymine rings were stacked one above the other in an infinite mode of packing and could perhaps be compared with ssDNA (Figure 9). The mutual orientation of the thymines connected by the spacer chain was such that the projection of one onto the plane of the other showed a rotation of 53° about an axis perpendicular to the plane. This gave a helical twist to the polyoxyethylene chain. The value of this rotation angle was slightly higher compared with the angle of rotation between adjacent thymines in B-DNA, which was only 36°. The two thymine planes were almost parallel, making an angle of 10°. In **25**, where the spacer chain contained a rigid catechol group, the thymines were disposed almost perpendicular to each other 74°, as well as with the planar catechol. Perhaps the presence of this rigid group was a deterrent for the spacer chain in **25** to twist itself in order to bring the thymine in a stacked position.

The photoproducts of **15** and **25** are cyclobutane linked dimers of thymine and were established to be the *trans-syn* isomers by determination of their crystal structures. Belonging to the centrosymmetric lattice, the crystals containing them are both enantiomers D and L of the *trans-syn* dimer. The formation of the photodimer was through inter-pyrimidine bonds between C5 and C6 atoms, respectively, of the two thymine bases. Thus, when the geometries of the compounds were compared before and after photodimerization, the main difference, and an obvious corollary was in the stereochemistry of C5 and C6. Both atoms have changed from a planar to a tetrahedral configuration. Other features observed on dimerization were the stretching of bonds, the

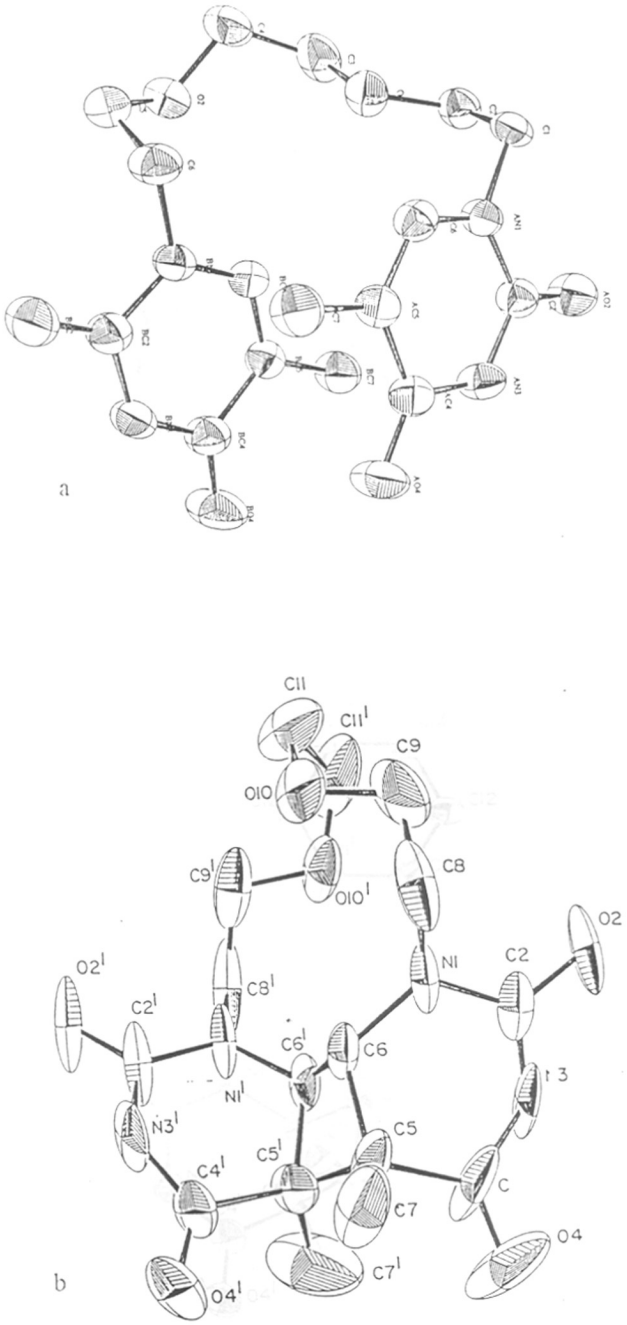


Figure 7: Ellipsoid plots showing the structure and atom labeling for (a) **15** and (b) **19**.

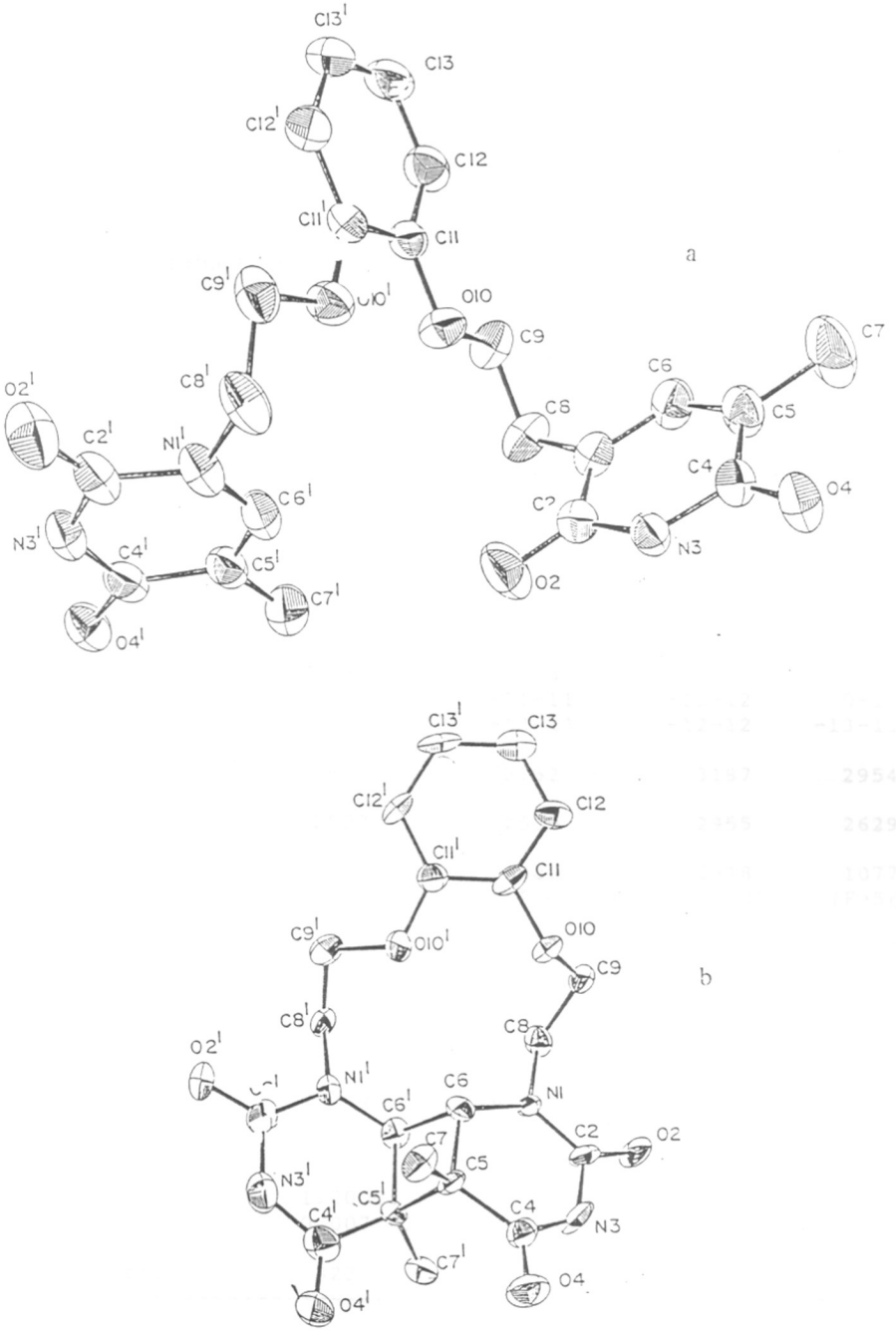


Figure 8: Ellipsoid plots showing the structure and atom labeling for (a) **25** and (b) **26**.

Table 1. Data Collection and structure refinement parameters.

	15	19	25	26
Crystal size (mm)	0.1X0.12X0.5	0.1X0.3X0.6	0.25X0.4X0.65	0.02X0.12X0.25
Reflections for cell (Number)	25	25	25	25
2θ range(^{0})	17.3-32.8	14.2-27.8	15-39.9	10-30
$2\theta_{max}$	47	47	47	46
h	0-9	0-8	0-9	0-11
k	-10-10	-11-11	-12-12	0-16
l	-13-13	-13-13	-12-12	-13-13
Measured reflections	2801	2752	3187	2954
Unique reflections	2597	2535	2955	2629
Observed reflections	1499 (F>5σ(F))	1482 (F>5σ(F))	2818 (F>0.0)	1077 (F>5σ(F))
$(shift/e.s.d.)_{max}$	0.012	0.005	0.090	0.016
$\Delta\rho_{max}/\Delta\rho_{min}$	0.15/-0.18	0.37/-0.50	0.32/-0.22	0.32/-0.29
R	0.0343	0.0953	0.0443	0.0628
wR	0.0385	0.1057	0.0570	0.0520
Weighting Scheme $a/[\sigma^2(Fo)+b(Fo)^2]$				
a value	1.0000	3.9769	1.0	2.7820
b value	0.007562	0.002200	0.003516	0.000001
No. of parameters	323	235	359	359

Table 2. Fractional coordinates (X 10^4) for nonhydrogen atoms and their corresponding isotropic displacement parameters (X 10^4 Å^2) for **15**.

Atom	X	Y	Z	U_{eq}
O2	4345(3)	2822(2)	-788(2)	509(9)
O4	569(3)	-1083(2)	1008(2)	535(9)
O2′	2345(3)	539(3)	5131(2)	641(11)
O4′	-931(3)	1768(3)	2570(2)	780(12)
O10	7638(3)	4654(2)	2279(1)	463(9)
O10′	6920(3)	5281(3)	4554(2)	519(9)
N1	5511(3)	1990(3)	650(2)	349(9)
N3	2505(3)	826(3)	109(2)	404(10)
N1′	3755(3)	2159(3)	3863(2)	380(9)
N3′	740(3)	1141(3)	3803(2)	445(11)
C2	4132(4)	1942(3)	-66(2)	372(12)
C4	2113(4)	-204(3)	913(2)	373(11)
C5	3632(4)	-162(3)	1587(2)	364(11)
C6	5241(4)	931(3)	1428(2)	366(12)
C7	3377(5)	-1292(4)	2449(3)	490(14)
C2′	2287(4)	1236(4)	4321(2)	434(13)
C4′	530(4)	1942(4)	2928(2)	444(12)
C5′	2132(4)	2940(3)	2515(2)	378(11)
C6′	3646(4)	2983(3)	2987(2)	373(11)
C7′	2034(5)	3866(4)	1576(3)	511(15)
C8	7311(4)	3222(4)	572(2)	388(12)
C9	7679(4)	4814(4)	1141(2)	425(13)
C11	8301(5)	6131(4)	2938(3)	554(15)
C11′	8537(5)	5812(5)	4102(3)	624(16)
C9′	6468(5)	3893(4)	5123(3)	526(15)
C8′	5504(4)	2369(4)	4374(3)	482(13)

Table 3. Fractional coordinates (X 10^4) for nonhydrogen atoms and their corresponding isotropic displacement parameters (X 10^3 Å^2) for **19**.

Atom	X	Y	Z	U_{eq}
O2	2540(8)	1050(8)	-583(4)	91(3)
O4	3630(17)	-3627(9)	1169(6)	216(7)
O2′	-2563(7)	-562(8)	5481(4)	88(3)
O4′	2324(13)	-4138(8)	5084(5)	135(5)
O10	-229(10)	3090(6)	2264(4)	90(3)
O10′	-2807(7)	2182(7)	1725(4)	78(3)
N1	2020(7)	478(7)	1430(4)	50(3)
N3	2490(8)	-1224(9)	414(4)	74(3)
N1′	-1522(7)	-645(8)	3587(4)	56(3)
N3′	139(11)	-2112(8)	5137(5)	72(3)
C2	2310(9)	166(11)	374(6)	55(3)
C4	3004(15)	-2461(12)	1333(7)	116(5)
C5	2781(12)	-2207(10)	2533(6)	77(4)
C6	1673(8)	-647(7)	2506(5)	41(3)
C7	4620(14)	-2718(11)	3006(7)	105(5)
C2′	-1407(10)	-1076(11)	4776(6)	68(4)
C4′	1262(16)	-3079(10)	4590(7)	86(5)
C5′	1073(16)	-2750(8)	3274(6)	82(4)
C6′	-34(10)	-1180(8)	2786(5)	51(3)
C7′	554(29)	-3953(13)	3110(13)	176(11)
C8	2489(12)	1737(13)	1460(6)	78(5)
C9	1079(18)	3110(11)	1242(7)	97(5)
C11	-1880(21)	4123(10)	2026(8)	134(7)
C11′	-2881(16)	3646(14)	1352(8)	114(6)
C9′	-3639(10)	1788(14)	2875(7)	79(5)
C8′	-3328(10)	184(15)	3210(6)	88(6)

Table 4. Fractional coordinates (X 10^4) for nonhydrogen atoms and their corresponding isotropic displacement parameters (X 10^4 $Å^2$) for **25**.

Atom	X	Y	Z	U_{eq}
O2	4478(2)	2451(2)	5512(2)	707(7)
O4	253(2)	331(1)	6719(1)	515(6)
O2′	7159(2)	9135(1)	8589(1)	609(6)
O4′	9132(2)	7111(1)	5096(1)	491(5)
O10	7203(1)	3175(1)	8418(1)	423(5)
O10′	6589(2)	5435(1)	9354(1)	436(5)
N1	4990(2)	1086(1)	6549(1)	406(6)
N3	2403(2)	1356(1)	6135(1)	375(5)
N1′	6343(2)	6928(1)	7715(1)	436(6)
N3′	8150(2)	8058(2)	6848(1)	419(6)
C2	4009(2)	1694(2)	6036(2)	415(7)
C4	1716(2)	516(2)	6697(2)	387(6)
C5	2818(2)	-70(2)	7240(2)	534(8)
C6	4372(2)	237(2)	7127(2)	532(8)
C7	2197(4)	-995(4)	7896(5)	1033(18)
C2′	7216(2)	3109(2)	7783(2)	446(7)
C4′	8268(2)	6974(2)	5875(2)	390(6)
C5′	7359(2)	5761(2)	5889(2)	407(6)
C6′	6451(2)	5792(2)	6794(2)	428(7)
C7′	7468(3)	4530(2)	4886(2)	554(8)
C8	6729(2)	1398(2)	6526(2)	490(7)
C9	7659(2)	1976(2)	7823(2)	463(7)
C11	7860(2)	3821(2)	9623(1)	357(6)
C12	8798(2)	3348(2)	10351(2)	489(7)
C13	9377(3)	4099(2)	11578(2)	586(9)
C13′	9023(3)	5300(2)	12091(2)	618(9)
C12′	8085(3)	5778(2)	11374(2)	519(8)
C11′	7508(2)	5054(2)	10144(1)	374(6)
C9′	6268(3)	6713(2)	9798(2)	468(7)
C8′	5348(3)	6895(2)	8712(2)	517(8)

Table 5. Fractional coordinates (X 10^4) for nonhydrogen atoms and their corresponding isotropic displacement parameters (X 10^3 $Å^2$) for **26**.

Atom	X	Y	Z	U_{eq}
O2	-485(8)	4779(5)	6291(5)	52(3)
O4	2743(7)	3438(5)	4769(5)	48(3)
O2′	1426(8)	-508(5)	7579(5)	55(3)
O4′	2106(8)	893(5)	4527(5)	61(3)
O10	1584(6)	3741(4)	9309(4)	27(2)
O10′	1369(7)	1907(5)	9401(5)	49(3)
N1	368(7)	3474(5)	7069(5)	24(3)
N3	1078(9)	4020(7)	5534(7)	39(4)
N1′	678(8)	965(6)	7378(6)	33(3)
N3′	1681(10)	234(7)	6042(7)	45(4)
C2	246(10)	4128(7)	6316(7)	30(4)
C4	2010(10)	3375(7)	5449(7)	30(4)
C5	1955(9)	2551(6)	6141(6)	24(3)
C6	1170(10)	2648(7)	7103(7)	27(4)
C7	3345(11)	2225(9)	6525(9)	40(4)
C2′	1254(11)	198(7)	7063(8)	39(4)
C4′	1628(10)	943(8)	5336(8)	42(4)
C5′	958(9)	1785(6)	5653(6)	24(3)
C6′	386(10)	1783(7)	6725(8)	29(4)
C7′	-63(12)	2083(10)	4733(8)	38(4)
C8	-365(10)	3601(8)	7966(7)	30(4)
C9	365(10)	4149(7)	8866(8)	29(4)
C11	1588(9)	3313(7)	10299(7)	32(4)
C12	1720(9)	3862(8)	11210(8)	35(4)
C13	1753(10)	3434(9)	12207(8)	40(4)
C13′	1759(10)	2491(8)	12264(7)	40(5)
C12′	1650(10)	1943(8)	11357(7)	37(4)
C11′	1539(9)	2364(7)	10367(7)	32(4)
C9′	939(14)	975(8)	9374(8)	45(5)
C8′	25(11)	852(8)	8330(8)	32(4)

most notable being N1-C6, C4-C5 and C5-C6, along with their pairs in the twin pyrimidine rings and the alteration of some bond angles. There was substantial reduction in the bond angles C5-C6-N1 and C4-C5-C6 and their pairs when **15** was converted to **19**. However, such an effect was confined to C5-C6-N1 pair when **25** changed to its photoproduct **26**. All these changes could be attributed to the saturation of the C5-C6 bond as a result of photodimerization. Again, the widest intraring pyrimidine angle, that at N3, has further widened on photodimerization; this effect being more pronounced in the case of **26**. While the intraring angles of cyclobutane assumed nearly equal values in both **19** and **26**, the other bond angles around C5 and C6 atoms have a similar situation only in **26** whereas their magnitudes varied considerably in **19** (Tables 6). In both the cases the angle pairs C5-C4-O4 displayed contraction on photodimerization, hinting at a plausible weak interaction between the fourth atom pairs C7 and O4.

On examining the stereochemistry of the cyclobutane rings in **19** and **26** the interpyrimidine bonds C5-C5' and C6-C6' were found to have an asymmetry in their lengths, with the former being longer than the latter. This could well be due to the bulk of the methyl groups at the C5 position. Similar asymmetry has also been reported in the crystal structures of other photodimers with bulky groups at such positions.[30] All the intraring cyclobutane angles were around 90°. It is a consequence of the asymmetry in the intrapyrimidine bond lengths that the angles opposite C5-C5' were slightly wider than the angles included by this bond. Cyclobutyl rings were twisted with intraring dihedral angles of 20° and 10° for **19** and **26** respectively. This twist, for identical enantiomers, is opposite in a sense for **19** and **26**, with the effect that the methyl group attached at C5 is brought closer in **19** while placed farther away in **26**. This is also reflected in the interatomic distances of methyl carbons in **19** and **26** (Tables 6). On the surface, this was surprising because the mutual disposition of the methyl groups in **19** only increased the strain on the cyclobutane ring. A possible explanation for such an anomalous folding of the cyclobutane ring in this structure could be the influence of the spacer chain. What turned out to be really interesting was that a rigid group such as catechol, present in the spacer chain, only eased the strain on the *trans-syn* thymine photodimer stereochemistry. Consolidation of these arguments needs more such examples and any biological

Table 6. Selected interatomic distances (Å) and bond angles (°).

	15	19	25	26
N1 -C2	1.378(3)	1.363(10)	1.368(2)	1.343(11)
N1'-C2'	1.372(3)	1.372(8)	1.368(2)	1.353(12)
C2 -N3	1.370(4)	1.349(15)	1.385(2)	1.409(12)
C2'-N3'	1.367(4)	1.385(10)	1.375(2)	1.426(12)
N3 -C4	1.381(4)	1.391(10)	1.370(2)	1.363(12)
N3'-C4'	1.383(4)	1.345(11)	1.378(2)	1.362(13)
C4 -C5	1.448(4)	1.514(13)	1.441(3)	1.494(12)
C4'-C5'	1.441(4)	1.542(12)	1.439(3)	1.492(13)
C5 -C6	1.347(4)	1.539(11)	1.342(3)	1.562(12)
C5'-C6'	1.343(4)	1.542(10)	1.336(3)	1.556(12)
C6 -N1	1.377(3)	1.465(7)	1.372(3)	1.461(11)
C6'-N1'	1.376(3)	1.442(8)	1.382(2)	1.461(12)
C2 -O2	1.215(3)	1.230(9)	1.211(2)	1.212(10)
C2'-O2'	1.229(3)	1.222(9)	1.226(2)	1.221(11)
C4 -O4	1.232(3)	1.188(15)	1.237(2)	1.228(11)
C4'-O4'	1.215(4)	1.203(11)	1.230(2)	1.202(11)
C5 -C5'		1.600(15)		1.587(12)
C6 -C6'		1.537(12)		1.540(12)
C7...C7'		3.75(3)		3.93(2)
C6 -N1 -C2	121.5(2)	118.3(7)	120.4(2)	127.0(8)
C6'-N1'-C2'	121.1(2)	120.3(5)	121.6(2)	124.8(8)
N1 -C2 -N3	114.3(2)	116.8(6)	114.7(2)	114.5(9)
N1'-C2'-N3'	114.9(2)	116.8(6)	114.2(2)	116.5(9)
C2 -N3 -C4	127.8(3)	127.7(8)	127.6(2)	129.9(9)
C2'-N3'-C4'	127.1(3)	128.6(8)	127.4(2)	129.2(10)
N3 -C4 -C5	114.8(3)	115.5(9)	115.0(2)	115.0(10)
N3'-C4'-C5'	115.0(2)	115.1(7)	115.2(2)	115.2(9)
C4 -C5 -C6	118.0(3)	110.5(6)	117.6(2)	116.9(8)
C4'-C5'-C6'	118.2(3)	111.4(7)	118.2(2)	118.8(8)
C5 -C6 -N1	123.3(3)	122.4(5)	124.8(2)	114.1(8)
C5'-C6'-N1'	123.5(3)	115.4(5)	123.3(2)	115.0(8)
N1 -C2 -O2	122.7(3)	123.6(10)	123.7(2)	124.8(9)
N1'-C2'-O2'	122.9(3)	122.9(7)	123.5(2)	125.2(10)
N3 -C2 -O2	123.0(3)	119.4(8)	121.6(2)	120.7(9)
N3'-C2'-O2'	122.2(3)	121.3(6)	122.3(2)	118.3(10)
N3 -C4 -O4	124.9(3)	121.7(9)	119.6(2)	121.2(9)
N3'-C4'-O4'	119.7(3)	123.1(8)	118.8(2)	122.0(10)
C5 -C4 -O4	124.9(3)	122.8(8)	125.4(2)	123.4(10)
C5'-C4'-O4'	125.2(3)	121.8(8)	126.0(2)	122.8(10)
C5'-C5 -C6		87.7(6)		89.3(7)
C5 -C5'-C6'		86.5(6)		88.1(6)
C5 -C6 -C6'		88.7(6)		89.5(7)
C5'-C6'-C6		89.9(6)		91.3(7)

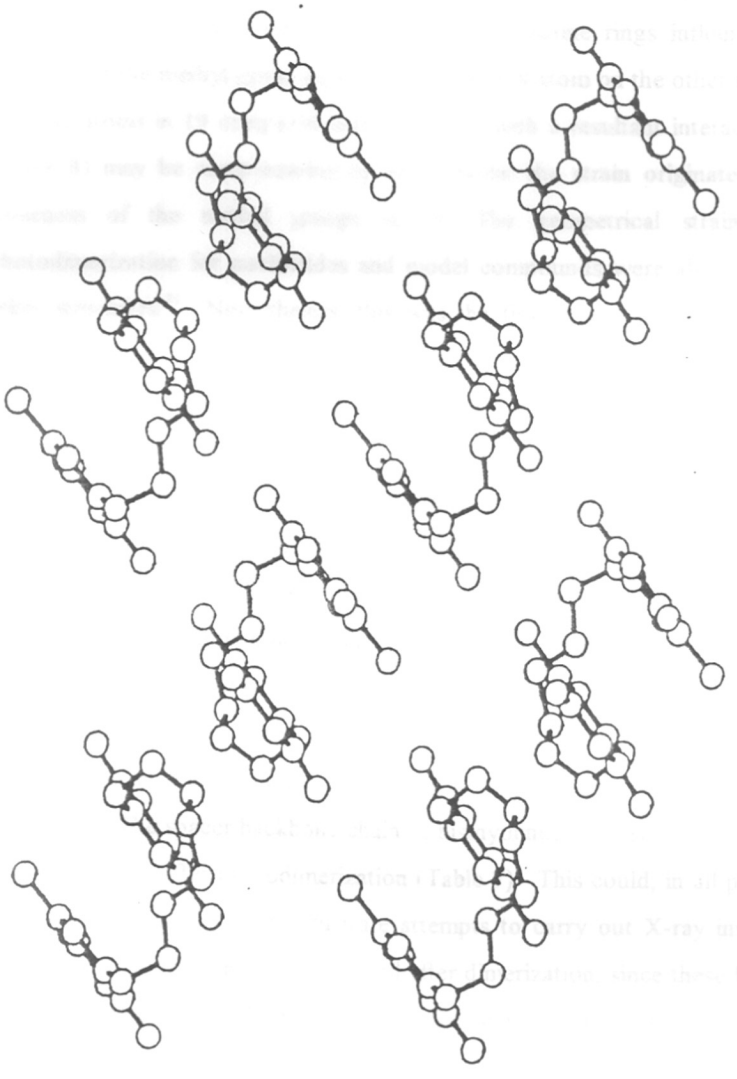


Figure 9: Packing of the molecules in the crystal structure of **25**

implications of this phenomena has to be explored further. The relative twists between the two pyrimidine rings in the photodimer defined as the average deviation of two torsional angles, C4-C5-C5'-C4' and N1-C6-C6'-N1', from the undistorted values of 120°, were 29.3 and 12.8 for **19** and **26**, respectively. The twists defined here, taken together with the twists of the cyclobutane rings referred to earlier, were also opposite in a sense for **19** and **26**. These relative twists between the pyrimidine rings influence the interatomic distances of the methyl group on one ring to the O4 atom on the other ring. The closeness of these atoms in **19** compared with **26**, along with a resultant interaction between them (Table 8) may be compensating to some extent the strain originated from the mutual closeness of the methyl groups in **19**. The geometrical strains at the site of photodimerization for nucleotides and model compounds were also reported for several other structures.[31] Nevertheless, this was the first time a comprehensive study of the stereochemistry of the photoproducts, in comparison, was accomplished through the determination of three dimensional structures of both, the parent compound and its photodimer.

The planar geometry of the pyrimidine rings was altered on photodimerization. This has been most significant in the case of **19**. Here the pyrimidine rings assume a boat conformation, with N3 and C6 located at bowsprit positions (Figure 10a and 10b). The pyrimidine rings assume a boat conformation in **26**. However, the difference is that here at the bowsprit positions the atoms were C2 and C5 in one ring, but N1' and C4' in the other (Figures 10c and 10d). Distortion of the planarity of pyrimidines could also be due to saturation at C5-C6 bonds, as has been reported for other cyclobutane dimers of thymines.[32] The spacer backbone chain of bisthymines has also undergone drastic changes in conformation after photodimerization (Table 7). This could, in all probability, turn out to be one aspect which may frustrate attempts to carry out X-ray investigations on the same crystals of bisthymines before and after dimerization, since these large displacements of the spacer atoms could easily rupture the single crystals. Crystal densities calculated for pre and post-dimerized compounds showed an apparent increase in close packing for photodimers. This could be attributed more to the shrinking of the molecules on dimerization rather than to any tighter packing, as no indications were available for any

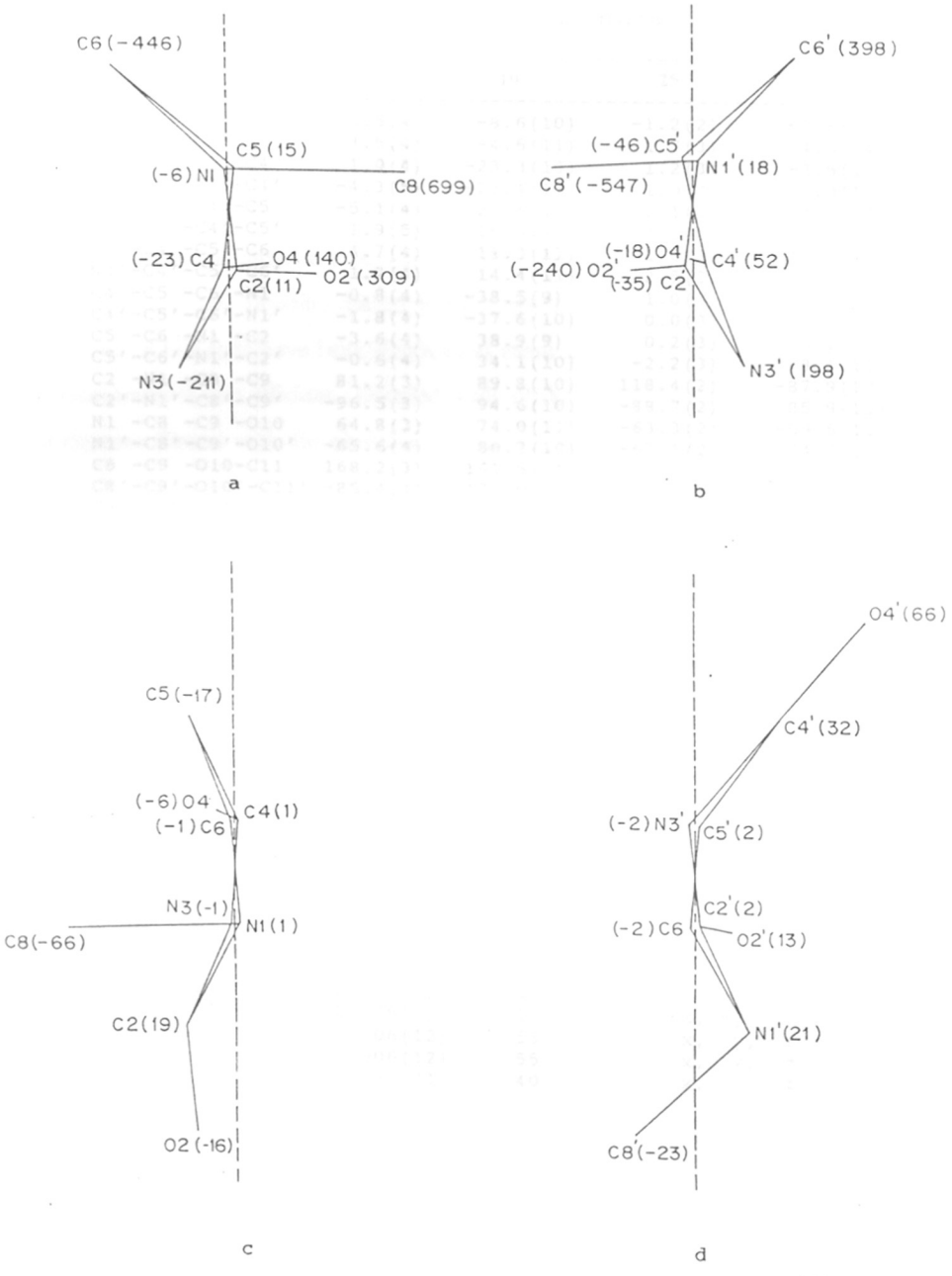


Figure 10: Projection of pyrimidine atoms (a) & (b) in **19** and (c) & (d) in **26**, perpendicular to a mean plane (indicated by dotted lines) calculated with respect to the four atoms having the lowest deviation from a least square plane for all atoms in the respective thymine rings. The numbers in parentheses are the deviations of respective atoms from the mean plane in units of 10^{-3} Å.

Table 7. Pyrimidine and the spacer backbone chain dihedral angles in the structures to show the large shifts in their magnitudes accompanied by confomational changes on dimerisation.

	15	19	25	26
C6 -N1 -C2 -N3	3.5(4)	-8.6(10)	-1.2(2)	-5.6(13)
C6'-N1'-C2'-N3'	3.5(4)	-4.6(11)	1.7(3)	4.6(14)
N1 -C2 -N3 -C4	1.0(4)	-23.3(13)	1.2(3)	-2.5(15)
N1'-C2'-N3'-C4'	-4.3(5)	-23.4(13)	1.0(3)	2.0(16)
C2 -N3 -C4 -C5	-5.1(4)	20.5(13)	-0.1(3)	14.8(15)
C2'-N3'-C4'-C5'	1.9(5)	16.6(14)	-3.1(3)	-3.9(16)
N3 -C4 -C5 -C6	4.7(4)	11.3(11)	-1.0(3)	-18.5(12)
N3'-C4'-C5'-C6'	1.2(4)	14.4(11)	2.5(3)	-0.4(13)
C4 -C5 -C6 -N1	-0.8(4)	-38.5(9)	1.0(3)	12.1(11)
C4'-C5'-C6'-N1'	-1.8(4)	-37.6(10)	0.0(3)	5.7(12)
C5 -C6 -N1 -C2	-3.6(4)	38.9(9)	0.2(3)	0.2(13)
C5'-C6'-N1'-C2'	-0.6(4)	34.1(10)	-2.2(3)	-8.1(13)
C2 -N1 -C8 -C9	81.2(3)	89.8(10)	118.4(2)	-87.9(10)
C2'-N1'-C8'-C9'	-96.5(3)	94.6(10)	-88.7(2)	85.9(11)
N1 -C8 -C9 -O10	64.8(3)	74.0(11)	-63.3(2)	-59.5(11)
N1'-C8'-C9'-O10'	-65.6(4)	80.2(10)	-67.1(2)	64.6(11)
C8 -C9 -O10-C11	168.2(3)	-161.5(9)	177.1(2)	-104.7(9)
C8'-C9'-O10'-C11'	-85.4(4)	-175.0(9)	175.9(2)	143.2(9)
C9 -O10-C11-C11'	-168.1(3)	73.8(11)	174.2(2)	103.6(10)
C9'-O10'-C11'-C11	131.8(3)	64.4(12)	-176.1(2)	-162.4(9)
O10-C11-C11'-O10'	-83.5(4)	40.7(13)	-1.5(2)	-5.5(13)

Table 8. Parameters corresponding to hydrogen bonds and other possible interatomic interactions.

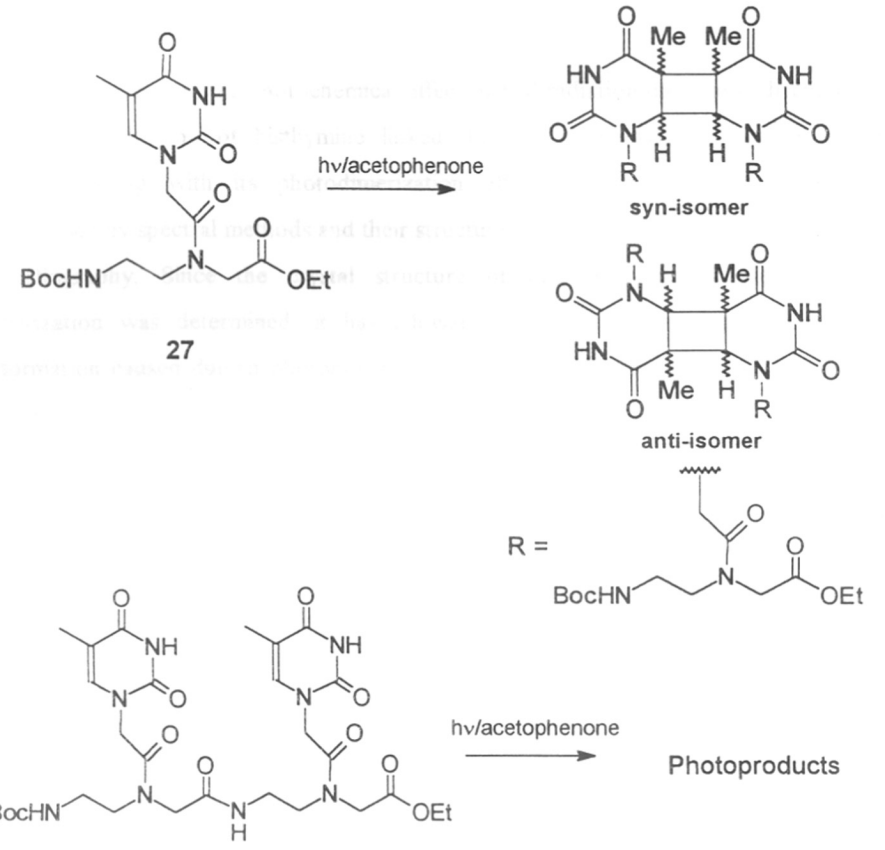
Donor-H...Acceptor	D...A(Å)	angle(H-D...A)(°)	Symmetry of A
15			
N3 -H...O4	2.942(4)	10(3)	-x, -y, -z
N3'-H...O2'	2.905(3)	5(2)	-x, -y, -z+1
C6 -H...O4'	3.153(4)	9(2)	x+1,y, z
C6'-H...O10'	3.061(3)	53(2)	x, y, z
C8-H1...O2	2.752(4)	64(2)	x, y, z
C8-H2...O4'	3.392(5)	21(2)	x+1,y, z
C8'-H1...O2'	2.776(4)	58(2)	x, y, z
19			
N3 -H...O10'	2.966(9)	17	-x, -y, -z
N3'-H...O10	2.976(8)	19	-x, -y, -z+1
C7-H1...O4'	3.006(12)	53	x, y, z
C7-H2...O4'	3.006(12)	55	x, y, z
C7'-H3...O4	3.023(21)	40	x, y, z
25			
N3 -H...O4'	2.930(2)	12(1)	-x+1,-y+1,-z+1
N3'-H...O4	2.888(2)	5(2)	x+1, y+1, z
C6'-H...O10'	3.073(3)	58(1)	x, y, z
C8-H1...O2	2.775(3)	52(1)	x, y, z
C9-H1...O4	3.191(3)	40(2)	x+1, y, z
C8'-H1...O2'	2.744(3)	57(2)	x, y, z
26			
N3 -H...O2	2.897(11)	16(5)	-x, -y+1,-z+1
N3'-H...O10	2.895(12)	8(6)	1/2-x,1/2+y-1,1/2-z+1
C6 -H...O10	3.194(11)	41(5)	x, y, z
C6 -H...O10'	3.085(11)	37(5)	x, y, z

additional intermolecular short contacts being propped up in the structures of photoproducts. From the magnitudes of the bond lengths of C2-O2 and C4-O4 the keto form of pyrimidines could be assumed. Thus, the only potential proton donor N3, sans N1 which is involved in the linkage, in each of the pyrimidine oxygen or a spacer chain oxygen of a symmetry-related molecule (Table 8). Few noticeable short carbon oxygen contacts, including those already discussed, mostly intramolecular, were expected to provide extra stability to individual molecules as well as to the overall packing in their respective crystals (Table 8). *trans-syn* Thymine dimers which were the main photoproducts of ssDNA on irradiation of UV light have been identified as more lethal to the cell, because their repair by photolyases is less efficient. Detailed study of the stereochemical interaction of *trans-syn* thymine dimers with the photolyases has been hampered by the sparse yield of this isomer in any particular experiment. In this context we have shown that the *trans-syn* thymine is the exclusive product upon UV irradiation of hitherto unknown bisthymines linked through novel polyoxyethylene spacer chains. The study reported here would help in the design and exclusive synthesis of rare *trans-syn* photodimers and provide a tool for understanding and modeling the repair mechanisms of DNA or even in designing DNA photolyases targeted at *trans-syn* thymine photodimers in particular.

4.4 Peptide nucleic acid T-T photodimers:

As evident from previous sections in the oligoethyleneglycol bisthymine, two terminal thymines oriented in an array favour the exclusive formation of cis-trans photodimer. The PNA unit having 11 atoms between the two bases is analogous with the tetraethyleneglycol bisthymine (11 atom spacer). The intervening amide bonds in PNA, both in the backbone and the sidechain, confer confomational rigidity to some extent which leads to specific base overlaps even in single stranded form. The latter has supporting evidence from CD studies of oligomers of PNA as well (*see chapter 3, section 3.5*). To examine such relative orientational effects of PNA bases on selectivity, if any, in photodimerization, the protected PNA monomer (**27**) was irradiated in the presence of sensitizer acetophenone. The formation of such dimer was evident from TLC, UV and ^{1}H NMR. But when the PNA dinucleotide (**28**) was irradiated under similar conditions, a complex mixture of products was observed. The behavior of these products on TLC and

their spectral properties indicated that dimers are formed, but the products have not been fully characterized. Further studies are required for systematic analysis of photoproducts.



27

28

Very recently the photodimerization[33] of single stranded PNA dimer containing a thymine and a 4-thiothymine has been reported. The irradiation of these dimers resulted in the formation of various photoadducts due to hydrogen abstraction from the 5-methyl group leading to 5+4 methylene adducts.

This study is very interesting and important as it is well established in the literature that cyclobutane pyrimidines are formed upon UV irradiation of TpT in cellular DNA (B form) whereas, in the bacterial spore DNA, 5-thyminyl-5,6-dithymine derivative is formed

as the spore cellular DNA is in A-form. A successful analysis and interpretation of photoproducts may also throw light on the backbone conformational preferences in DNA.

4.5 CONCLUSION

In this chapter the photochemical effect of UV radiation damage is discussed. The synthetic preparation of bisthymine linked through oligoethylenoxy spacer chain is described along with its photodimerization effect. The photo dimers were well chracterised by spectral methods and their structure was indisputably estabilished by X-ray crystallography. Since the crystal structure of the bisthymines before and after dimerization was determined, it has allowed us to discuss in detail the structural deformation caused due to photodimerization. The dimerisation of PNA dipeptides was attempted to study the effect of the backbone.

4.6 Experimental

For general remarks on TLC, NMR etc. *see chapter 2, section 2.7.*

Synthesis of 3,6,9,12-tetraoxa tetradecane-1,14-diol (10)

Powdered KOH (30 g, 0.5 mol) was taken in ethylene glycol (100 g, 1.6 mol) and stirred slowly till all the KOH dissolved. Triethyleneglycol dichloride (20 g, 0.1 mol) was added and the mixture stirred at RT for 5 days. Excess ethyleneglycol was removed by distillation and the thick oil was distilled under vacuum in a kugelrohr twice to obtain pentaethyleneglycol **10** (15 g, yield 65%). GC purity 99%, MS: m/e 238, base peak 89 ($HOCH_2CH_2OCH_2CH_2$-) ^1H NMR, (D_2O) δ: 3.72 (s)

Synthesis of 3,6,9,12,15-pentaoxa heptadecane-1,17-diol (11)

Powdered KOH (30 g, 0.5 mol) was taken in diethylene glycol (106 g, 1.0 mol), diethylenglycol dichloride (15 g, 0.1 mol) was added and the mixture stirred for 5 days at 60 C. The excess diethyleneglycol was removed by distillation and the thick oil was distilled under vacuum in a kugelrohr twice to obtain 10 g of hexaethyleneglycol **11** (yield 65%). GC purity 95%, MS: m/e 284, base peak 88 (-$OCH_2CH_2OCH_2CH_2$-100%), 89 (91%) ^1H NMR, (D_2O) δ: 3.72 (s)

General synthesis of oligoethyleneoxydichlorides

Oligoethyleneglycol (10 mmol) was taken in 4 eq. of pyridine and cooled in an ice bath. To this well-stirred ice-cold mixture, thionyl chloride (60 mmol) was added slowly through a dropping funnel, taking care that the temperature does not exceed 30 C. The mixture was stirred at ambient temperature for 4-5 h, poured into ice-cold water and extracted with DCM, dried over anhydrous Na_2SO_4 and distilled under vacuum to get pure compound in 70-85 % yield.

1,8-dichloro 3,6-dioxaoctane (8a): To a mixture of triethyleneglycol **7** (20.4 g, 0.13 mol) and pyridine (45 mL) was added thionyl chloride (22 mL) to get the product **8a** in 22 g, (86 %) after distillation.

IR: Thin film (neat): 2880 (s), 1520 (m), 1490 (m), 1150 (s, O-CH_2), 758 and 670 (C-Cl), ^{1}H NMR, ($CDCl_3$) 90 MHz: 3.67 (m)

1,11-dichloro-3,6,9-trioxaundecane: To tetraethylene glycol **9** (10.3 g, 53 mmol) in pyridine (18 mL) was added thionyl chloride (17.5 mL) to get the corresponding dichloride **12a** in 75 % yield.

IR: Thin film (neat): 2880 (s), 1520 (m), 1490 (m), 1150 (s, O-CH_2), 758 and 670 (C-Cl), ^{1}H NMR, ($CDCl_3$) 90 MHz: 3.67 (m)

1,14-dichloro-3,6,9,12-tetraoxa tetradecane (13a): Thionyl chloride (7.0 mL) was added to pentaethylene glycol **10** (10.2 g, 42.8 mmol) and pyridine (15 mL) to get the product **13a** in 76 % yield.

IR: Thin film (neat): 2880 (s), 1520 (m), 1490 (m), 1150 (s, O-CH2), 758 and 670 (C-Cl), ^{1}H NMR, ($CDCl_3$) 90 MHz: 3.67 (m)

1,17-dichloro-3,6,9,12,15-pentaoxa heptadecane (14a): Thionyl chloride (6.4 mL) was added to hexaethylene glycol **11** (11.2 g, 39.7 mmol) and pyridine (13 mL) to get the product (**14a**)in 70% yield.

IR: Thin film (neat): 2880 (s), 1490 (m), 1360 (m), 1310 (m), 1150 (s, O-CH_2), 755 and 670 (C-Cl), ^{1}H NMR, ($CDCl_3$) 80 MHz : 3.67(m)

General synthesis of diiodocompounds from dichloro compounds (8b, 12b-15b)

The dichlorides **8a, 12a-15a** (10 mmol) and NaI (50 mmol) were taken in 10 mL of acetone and the mixture was refluxed for 3 days under N_2 atmosphere. The reaction mixture was cooled, filtered and solvents were removed and the thick brown solution of corresponding diiodides **8b, 12b-15b** were used as such without purification.

Catechol-O,O-di(ethane-2-tetrahydropyranether (23)

KOH (25 g) was dissolved in DMSO (75 mL) and to this mixture catechol (3.0 g, 27.2 mmol) was added followed by 1-chloroethane-2-tetrahedrafuranether (12.35 g, 75 mmol). The mixture was stirred at 60°C for 3 days. The reaction mixture was diluted with

water (150 mL) and extracted with ether (75 mL x 3). The solvent was removed to obtain the crude oil, which was further purified by column chromatography using pet-ether as eluant to obtain the product **23** in yield 56 % (6.2 g)

^{1}H NMR, ($CDCl_3$) (80 MHz) δ: 7.05-6.90 (m, 4H, Ar), 4.60 (br, 2H, 6H-THP), 3.92-3.21 (m, 12H), 1.82-1.40 (m, 12H).

Catechol-O,O-diethane-2-bromide (24)

Compound **23** was taken in DCM (25 mL), PPh_3 (7.2 g, 27 mmol) was added and the mixture was cooled in an ice-bath. Bromine (2.5 mL, 27 mmol) was added dropwise to this mixture and stirred for 15 min. The solvent was evaporated and the residue was loaded on silica gel column and eluted with pet-ether to obtain the dibromide **24** in 45% yield (2.01 g) M.P. = 66°C

^{1}H NMR, ($CDCl_3$) δ: 6.97 (s, 4H), 4.36 (t, J = 6.3, 4H), 3.67 (t, J = 6.3, 4H). MS = m/e (324).

General Procedure synthesis of oligoethyleneoxybis thymine

A mixture of thymine (1.0 g) and trimethylsilyl chloride (TMSCl, 4 mL) in HMDS (15 mL) was refluxed with constant stirring under nitrogen atmosphere for 20 h. The reaction mixture was cooled, oligoethyleneoxydiiodides (1.0 g) was added and the mixture was refluxed for 24-72 h. The progress of the reaction was followed by TLC. The reaction mixture was cooled to ambient temperature and HMDS was removed under vacuum and the residue was taken in methanol, adsorbed on silica gel and purified on silicagel column. Small amounts of monothymine oligoethyleneoxyiodo compounds were also isolated.

1,8-(bisthymin-1-yl)-3,6-dioxaoctane (15): Thymine (0.56 g, 4.46 mmol), HMDS (10 mL) and TMSCl (1.4 mL, 15.9 mmol) were refluxed for 20 h. The diiodide **8b** (0.4 g, 1.24 mmol) was added to this and refluxed for 48 h to obtain **15** in 77 % yield (0.43 g).

M.P: 158 °C ^{1}H NMR, ($CDCl_3$ + DMSO-d_6 5:1 v/v) δ: 9.05 (br, 2H, NH), 7.10 (d, 2H, T-H6), 3.80 (t, J = 5Hz, 4H, O-CH_2), 3.69 (t, J = 5 Hz, 4H, T-NCH_2), 3.47 (s, 4H, O-CH_2), 1.93 (d, 6H, T-CH_3). ^{13}C NMR (DMSO-d_6) δ: 164.6 (T-C2), 151.3 (T-C4), 142.5 (T-C5), 108.6 (T-C6), 70.3 (O-CH_2), 68.2 (-O-CH_2), 47.3 (N-CH_2), 12.2 (T-CH_3).

1,11-(bisthymin-1-yl)-3,6,9-trioxaundecane (16): Thymine (1.072 g, 8.4 mmol) and TMSCl (2.8 mL, 31.8 mmol) in HMDS (18 mL) were refluxed for 20 h, diiodide **12b** (1.2 g, 2.94 mmol) was added and refluxing continued for 32 h. Compound **16** was obtained in 70 % yield (0.84 g).

M.P. = 148 °C. ^{1}H NMR, ($CDCl_3$) δ: 8.95 (br, 2H, NH), 7.16 (d, 2H, T-H6), 3.93 (t, J= 5Hz, 4H, ArO-CH_2), 3.76 (t, J=5 Hz, 4H, T-NCH_2), 3.65 (s, 8H, O-CH_2), 1.94 (d, 6H, T-CH_3). ^{13}C NMR, ($CDCl_3$ + DMSO-d_6 1:1 v/v) δ: 167.8 (T-C2), 153.8 (T-C4), 145.1 (T-C5), 111.3 (T-C6), 72.5 (O-CH_2), 70.6 (-O-CH_2), 49.1 (N-CH_2), 13.2 (T-CH_3).

1,14-(bisthymin-1-yl)-3,6,9,12-tetraoxa tetradecane (17): A mixture of thymine (1.029 g, 8.2 mmol) and TMSCl (2.7 mL 30.6 mmol) in HMDS (15 mL) was refluxed for 20 h followed by addition of diiodide **13b** (1.14 g, 2.53 mmol) and subsequent refluxing of the mixture yielded the desired product **17** in 73 % yield (0.832 g).

M.P. = 98 °C. ^{1}H NMR ($CDCl_3$) δ: 9.50 (br, 2H, NH), 7.12 (d, 2H, T-H6), 3.95 (t, J= 5Hz, 4H, ArO-CH_2), 3.75 (t, J=5 Hz, 4H, T-NCH_2), 3.65 (s, 12H, O-CH_2), 1.94 (d, 6H, T-CH_3). ^{13}C NMR, (CD_3COCD_3) δ: 165.3 (T-C2), 152.1 (T-C4), 143.2 (T-C5), 109.5(T-C6), 71.4 (O-CH_2), 69.6 (-O-CH_2), 48.5 (N-CH_2), 12.4 (T-CH_3).

1,17-(bisthymin-1-yl)-3,6,9,12,15-pentaoxa heptadecane (18): Thymine (1.02 g, 8.2 mmol), HMDS (15 mL), TMSCl (2.7 mL, 30.6 mmol) and diiodide **14b** (1.0 g, 2.01 mmol) were refluxed as earlier described to obtain bis thymine **18** in 65% yield (0.65 g).

^{1}H NMR ($CDCl_3$) δ: 9.40 (br, 2H, NH), 7.15 (d, 2H, T-H6), 3.93 (t, J= 5Hz, 4H, ArO-CH_2), 3.78 (t, J=5 Hz, 4H, T-NCH_2), 3.72 & 3.7 (s, 16H, O-CH_2), 1.94 (d, 6H, T-CH_3). ^{13}C NMR, ($CDCl_3$) δ: 165.15 (T-C2), 152.15 (T-C4), 142.37 (T-C5), 109.8 (T-C6), 70.68 (O-CH_2), 69.15 (O-CH_2), 48.3 (N-CH_2), 12.35 (T-CH_3).

Catechol-O,O-bis(ethane-2-thymine) (25): The desired product **25** was prepared from refluxing a mixture of thymine (0.5 g, 3.89 mmol) and TMSCl (1.37 mL, 12 mmol) in HMDS (15 mL) for 20 h followed by addition of dibromide **24** (0.4 g, 1.23 mmol). The refluxing was continued for 72 h to obtain the bis compound after workup in 70% yield (0.28 g).

^{1}H NMR, ($CDCl_3$ +DMSO-d_6 5:1 v/v) δ: 9.80 (br, 2H, NH), 7.08 (d, 2H, T-H6), 6.75 (m, 4H, Ar), 4.06 (t, J = 5Hz, 4H, ArO-CH_2), 3.92 (t, J = 5 Hz, 4H, T-NCH_2), 1.93 (d, 6H, T-CH_3). ^{13}C NMR, (DMSO-d_6) δ: 164.6 (T-C2), 151.4 (T-C4), 149.4 (O-C=C), 142.3 (T-C5), 122.1 (O-C=C), 108.6 (T-C6), 66.9 (-O-CH_2), 47.2 (N-CH_2), 12.1 (T-CH_3).

General Procedure for Photodimerization

About 25-30 mg of bis thymine was taken in 10 mL of 7:3 v/v water:acetonitrile and irradiated for 4-6 h. with hanovia lamp PC-451 with pyrex filter (λ = 290) in the presence of acetophenone (7 eq.) as the sensitizer. The reaction was followed by TLC and when all the starting material was consumed, the solvents were evaporated and the dimer was crystallized out of methanol.

***1,8-(bisthymin-1-yl-5'5''-6'6''-photodimer)-3,6-dioxaoctane*:** ^{1}H NMR, ($CDCl_3$ + DMSO-d_6) δ: 8.90 (br, 2H, NH), 3.68-3.10 (m, 14H), 1.23 (s, 6H, T-CH_3). ^{13}C NMR, (DMSO-d_6) δ: 170.4 (C2), 151.5 (C4), 72.0 (O-CH_2), 70.0 (O-CH_2), 63.8 (C6), 48.1 (N-CH_2), 44.5 (C5), 21.8 (T-CH_3).

***1,11-(bisthymin-1-yl-5'5''-6'6''-photodimer)-3,6,9-trioxaundecane*:** ^{1}H NMR, ($CDCl_3$+ DMSO-d_6) δ: 8.7 (br, 2H, NH), 3.78-3.1 (m, 18H), 1.33 (s, 6H, T-CH_3).

***1,14-(bisthymin-1-yl-5'5''-6'6''-photodimer)-3,6,9,12-tetraoxa tetradecane*:** ^{1}H NMR, ($CDCl_3$) δ: 8.40 (br, 2H, NH), 4.08-3.41 (m, 22H), 1.39 (s, 6H, T-CH_3). ^{13}C NMR, (DMSO-d_6) δ: 169.7 (C2), 150.5 (C4), 70.0 (O-CH_2), 69.0 (O-CH_2), 63.0 (C6), 46.1 (N-CH_2), 45.5 (C5), 20.8 (T-CH_3).

***1,17-(bisthymin-1-yl-5'5''-6'6''-photodimer)-3,6,9,12,15-pentaoxaheptadecane*:** ^{1}H NMR ($CDCl_3$+ DMSO-d_6) δ: 8.90 (br, 2H, NH), 3.82-3.30 (m, 26H), 1.51(s, 6H,T-CH_3).

***Catechol-O,O-bis(ethane-2-thymin-1-yl-5'5''-6'6''-photodimer)*:** ^{1}H ($CDCl_3$+DMSO-d_6) δ: 8.60 (br, 2H, NH), 6.82 (s, 4H, Ar), 3.40-3.15 (m, 10H), 1.33 (s, 6H, T-CH_3).

X-ray structure: Methodology

Plate crystals of **15** and **25** were readily obtained from methanol solutions of the respective compounds on evaporation. Compound **26** was also crystallized from methanol as thin plates under controlled evaporation. Crystals suitable for X-ray analysis of **19** could

be prepared only by repeated seeding of the initially grown very thin long crystals in methanol-water solution. The resultant crystals were not of the excellent quality as reflected in the slightly high R factor for **19**. However, we decided to use them for analysis in the absence of better crystals. All crystals were colorless. X-ray intensity data for all compounds was collected on an Enraf Nonius CAD4 P.C. controlled single crystal diffractometer using Zirconium filtered Mo Kα radiation (λ=0.71068 °A). Data collection was carried out at 295 K employing ω-2θ scan with scan width (0.80 + 0.35 tan θ)°. Data were assimilated in 96 steps with 16 steps on either side being used for background estimation. The intensities of three reflections monitored at intervals of 3600 seconds for the first three and 5000 seconds for the last showed no significant variation throughout data collection, ruling out any crystal damage especially to photoproducts on incidence of X-rays. Further, the unit cell estimation before and after data collection showed little difference. Lorentz and polarization corrections were applied to the reflection data during data reduction. However, no absorption correction was applied. The crystal structures were solved by direct methods and Fourier calculations in SHELXS-86[34a] and refined by full matrix least-squares method in F using SHELX76.[34b] After the input and refinement of non-hydrogen atoms, treating temperature factors anisotropic, the hydrogens were either fixed at ideal geometries or located from a difference map phased on refined non-hydrogen atom positions. Hydrogens were assigned the isotropic thermal parameters of the atoms to which they were bonded. In all cases except **19**, the hydrogen parameters were refined. Nonetheless, no hydrogens, though included for structure factor calculation, were refined in **19**, during last cycles as some of them tend to assume unacceptable values. Parameters pertaining to data collection and structure refinement were summarized in Table 1. Computations were carried out on a Magnum RISC computer with unix operating system.

4.7 Reference

1. (a) Cadet, J.; Vigny, P. *Bioorganic Photochemistry*, edited by Morrison, H. **1990**, Vol *1*, 1-272, New York:John Wiley & Sons. (b)Carell, T. *Angew. Chem. Int. Ed. Engl.* **1995,** *34*, 2491.
2. Taylor, J-S. *Acc. Chem. Res.* **1994**, *27*, 76.
3. Heelis, P. F.; Hartman, R. F.; Rose, S. D. *Chem. Soc. Rev.* **1995**, 289.
4. Taylor, J-S. *J. Chem. Edu.* **1990**, *67*, 835.
5. Culotta, E.; Koshland, D. E. *Science.* **1994**, *266*, 1926.
6. Sancar, A. *Biochemistry*, **1994**, *33*, 2.
7. Begley, T. P. *Acc. Chem. Res.* **1994**, *27*, 394.
8. Kim, S-T.; Sancar, A. *Photochem. Photobiol.* **1993**, *57*, 895, 905.
9. Park, H-W.; Kim, S-T,; Sancar, A.; Deisenhofer, J. *Science.* **1995**, *268*, 1866, 1858.
10. Kim, S-T.; Malhotra, K.; Smith, C. A.; Taylor, J-S.; Sancar, A. *Biochemistry*, **1993**, *32*, 7065.
11. Browne, D. T.; Eisinger, J.; Leonard, N. J. *J. Am. Chem. Soc.* **1968**, *90*, 7302.
12. (a) Begley, T.P.; Austin, R.; McMordie, S.; Altmann, E.; *J. Am. Chem. Soc.* **1993**, *115*, 10370. (b) Begley, T. P.; Austin, R.; McMordie, S. *J. Am. Chem. Soc.* **1992**, *114*, 1886.
13. Logue, M. W.; Leanard, N. J. *J. Am. Chem. Soc.* **1972**, *94*, 2843.
14. (a) Kim, S-T.; Hartman, R. F.; Rose, S. D. *Photochem. Photobiol.* **1990**, *52*, 789. (b) van Camp, J. R.; Hartman, R. F.; Rose, S. D. *Photochem. Photobiol.* **1987**, *45*, 365.
15. Essenmacher, C.; Kim, S-T.; Atamian, M.; Babcock, G. T.; Sancar, A. *J. Am. Chem. Soc* **1993**, *115*, 1602.
16. Rokita, S.,E.; Walsh, C. T. *J. Am. Chem. Soc.* **1984**, *106*, 4589.
17. Roth, H. D.; Lamola, A. A.; *J. Am. Chem. Soc* **1972**, *94*, 1013.
18. (a) Carell, T.; Epple, R.; Gramlich, V. *Angew. Chem. Int. Ed. Engl.* **1996**. *35*, 620. (b) Hartmann, R. F.; Rose, S. D. *J.Org. Chem.* **1992**, *57*, 2302. (c) Hartman, R. F.; Rose, S. D. *J. Am. Chem. Soc.* **1992**, *114*, 3559. (d) Heelis, P. F. Hartman, R. F.; Rose S. D. *Photochem. Photobiol.* **1993**, *57*, 1053.
19. Hirst, S. C.; Hamilton, A. D. *Tetrahedron Lett.* **1990**, *31*, 2401.
20. Rose, S. D.; Goodman, M. S. *J. Org. Chem.* **1992**, *57*, 3268.

21. (a) Taylor, J-S.; Nadji, S. *Tetrahedron*, **1991**, *47*, 2579. (b) Kemmink, J.; Boelens, R., Koning, T.; van der Marel, G.A.; van Broom, J. H.; Kaptein R. *Nucleic Acids Res.* **1987**, *15*, 5123.

22. Smith, C. A.; Taylor, J-S. *J. Biol. Chem.* **1993**, *268*, 11143. (b) Taylor, J-S.; Brockie, I. R. *Nucleic Acids Res.* **1988**, *16*, 5123.

23. (a) Taylor, J-S.; Garrett, D. S.; Brockie, I. R.; Svoboda, D. L.; Telser, J. *Biochemistry*, **1990**, *29*, 8858. (b) Rycyna, R. E.; Wallace, J. C.; Sharma, M.; Aiderfer, J. L. *Biochemistry*, **1988**, *27*, 3152.

24. Zhao, X.; Taylor, J-S. *J. Am. Chem. Soc.* **1994**, *116*, 8870.

25. Castellan, A.; Desvergne J-P. *Photochem. Photobiol.* **1981**, *34*, 183.

26. Castellan, A.; Desvergne, J-P. Bideau, J-P.; Bravic, G.; Courseille, C.; *Mol. Crysl. Liq. Cryst.*, **1983**, *93*, 103.

27. Gangamani, B. P.; Suresh, C. G.; Ganesh, K. N. *J. Chem. Soc. Chem. Commun.* **1994**. 2275.

28. Johnson, C.K. *ORTEP*II. **1976**, Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.

29. Gerdil, R. *Acta Cryst.* **1961**, *14*, 333.

30. (a) Leonard, N. J.; Golankiewiez, K.; McGredie, R. S.; Johnson, S. M.; Paul, I. C. *J. Am. Chem. Soc.* **1969**, *91*, 5855. (b) Birnbaum, G. I. *Acta Cryst.* **1972**, B-*28*, 1248. (c) Jankun, E. S.; Maluszynska, H.; Kaluski, Z.; Wiez, K. G. *Acta Cryst.* **1977**, B-*33*, 1624. (d) Koziol, A. E.; Rajchel, A. *Acta Cryst.* **1982**, B-*38*, 1001.

31. Karle, I. L. *Photochemistry and Photobiology of Nucleic Acids*, edited by Wang, S.Y., **1976**, Vol *1*, 483. New York : Academic Press.

32. Cadet, J.; Voituriez, L.; Hruska, F. E.; Grand, A. *Biopolymers*, **1985**, *24*, 897

33. Clivio, P.; Guillaume D.; Adeline, M-T.; Fourrey, J-L. *J. Am. Chem. Soc.* **1997**, *119*, 5255.

34. (a) Sheldrick, G.M. *SHELX*86. *Program for Crystal Structure Determination*. **1985**, University of Gottingen, Germany. (b) Sheldrick, G.M. *SHELX*76. *Program for Crystal Structure Determination*. **1976**, University of Cambridge, England.