" SYNTHESIS AND EVALUATION OF CONFORMATIONALLY CONSTRAINED PYRROLIDYL POLYAMIDE NUCLEIC ACIDS "

THESIS SUBMITTED TO

THE UNIVERSITY OF POONA

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

CHEMISTRY

BY

MONEESHA D'COSTA

NATIONAL CHEMICAL LABORATORY

PUNE 411008

FEBRUARY 2001

CERTIFICATE

This is to certify that the work presented in the thesis entitled "**SYNTHESIS AND EVALUATION OF CONFORMATIONALLY CONSTRAINED PYRROLIDYL POLYAMIDE NUCLEIC ACIDS**" submitted by Moneesha D'Costa was carried out by the candidate at the National Chemical Laboratory Pune, under my supervision. Such materials as obtained from other sources have been duly acknowledged in the thesis.

February 2001

(K. N. Ganesh)
Research Guide
Head, Division of Organic Chemistry (Synthesis)
National Chemical Laboratory
Pune 411 008

CANDIDATE'S DECLARATION

I hereby declare that the thesis entitled "SYNTHESIS AND EVALUATION OF CONFORMATIONALLY CONSTRAINED PYRROLIDYL POLYAMIDE NUCLEIC ACIDS**" submitted for the degree of Doctor of Philosophy in Chemistry to the University of Poona has not been submitted by me to any other university or institution. This work was carried out at the National Chemical Laboratory, Pune, India.**

National Chemical Laboratory (Moneesha D'Costa)

Pune 411 008

February 2001

ACKNOWLEDGEMENTS

It gives me great pleasure to express my gratitude to my supervisor, Dr. K. N. Ganesh. His encouragement and scientific temperament furthered my interest in Chemistry and its relevance to biological systems. His tireless enthusiasm was always a source of inspiration.

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

My sincere thanks also go out to Mrs. Anita Gunjal and Dr. A. A. Natu for their help during the course of this work.

I take this opportunity to place on record my thanks to Dr. Vairamani, IICT, Hyderabad and Prof. Balram, IISc, Bangalore for their help in recording the MALDI-TOF spectra.

The kind help extended by the NMR facility fraternity- Mr.Samuel, Mrs. Phalgune, Mr. Sathe and Dr. Sanjayan also must be acknowledged here.

Special gratitude also goes to Mrs. M. V. Mane and Mrs. S. S. Kunte for carrying out HPLC analysis, Mrs. Shantakumari for mass analysis and Mrs. Kale for recording IR spectra.

This page would be incomplete without the mention of my seniors and colleagues who have gone out of their way to help in various capacities and have been my extended family throughout the tenure of my work in NCL. Thank you Joshi, Prakash, Dinesh, Rajeev, Vasant, Gopal, Anand, Gangamani, Pradeep, Vipul, Ramesh, Pallavi, Meena, Vallabh, Leena, Dimpan, Nagamani, Nagendra and Dinesh (Jr.). I also thank Pawar, Bhumkar and Sunil for their assistance.

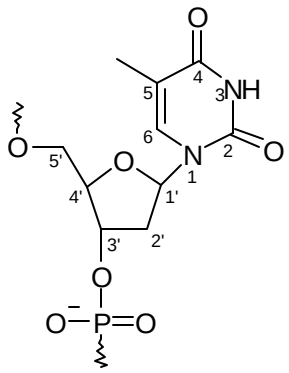
I am grateful to the Director, NCL, for giving me the opportunity to work in this institute and making the many facilities available for carrying out research.

CSIR, New Delhi, is acknowledged for valuable support in the form of a Fellowship.

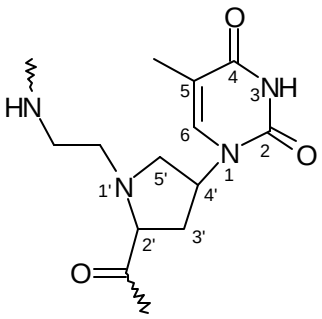
Finally, I'd like to acknowledge in a special way my parents, brother and my aunts and uncles, who were extremely patient and tolerant towards my erratic hours of work and constantly encouraged me to excel in whatever I did.

Moneesha

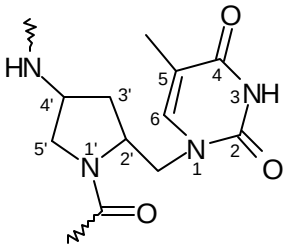
CONTENTS



DNA **I**



*Aep*PNA **II**



Pyrrolidyl PNA **III**

The aminoethylprolyl PNA discussed in this thesis can be considered as a DNA analogue as it very closely resembles the DNA structure. In DNA, the atoms of the nucleobase are numbered 1, 2, 3, … etc., while the sugar unit 'C' atoms are designated as 1', 2', 3', 4' and 5'. An analogous numbering system can be extended to the *aep*PNA unit **II** in which the pyrrolidine ring atoms would be numbered as 1', 2', 3', 4' and 5'. However, for the sake of simplicity and convenience, these atoms have been referred to as 1, 2, 3, 4 and 5 throughout this thesis. The clarity in differentiating the atoms of the pyrrolidyl unit and those of the nucleobase is in no way compromised. The exocyclic amino groups of the nucleobase are referred as the superscripts of the nucleobase carbon atoms to which they are attached.

The pyrrolidine ring numbering convention followed for *aep*PNA **II** is retained in discussing the pyrrolidyl PNA **III**.

LIST OF PUBLICATIONS

1. Conformationally Restrained Chiral PNA Conjugates. Synthesis and DNA Complementation Studies.
   Nucleosides & Nucleotides **1999**, *18*, 1409-1411.
   B. P. Gangamani, **M. D'Costa**, V. A. Kumar and K. N. Ganesh
2. Aminoethylprolyl Peptide Bucleic Acids (*Aep*PNA): Chiral PNA Analogues that form DNA:(*aep*PNA)$_2$ Triplexes.
   Organic Letters **1999**, *1*, 1513-1516.
   Moneesha D'Costa, Vaijayanti A. Kumar and Krishna N. Ganesh
3. Engineering Preferences of Hairpin PNA Binding to Complementary DNA: Effect if N7G in *aeg*/*aep* PNA Backbone.
   Nucleosides, Nucleotides and Nucleic Acids **2001**, 0000. *Accepted for publication.*
   Vaijayanti Kumar, **Moneesha D'Costa** and K.N. Ganesh
4. Aminoethylprolyl (*aep*) PNA: Mixed Purine/Pyrimidine Oligomers and Binding Orientation Preferences for PNA:DNA Duplex Formation

Organic Letters **2001**, submitted.

Moneesha D'Costa, Vaijayanti Kumar and Krishna N. Ganesh

5. Engineering Orientation Selectivity in PNA:DNA Binding: Effect of N7G in *aeg*/*aep* PNA Hairpin Binding.

Manuscript to be communicated

Moneesha D'Costa, Vaijayanti Kumar and Krishna N. Ganesh

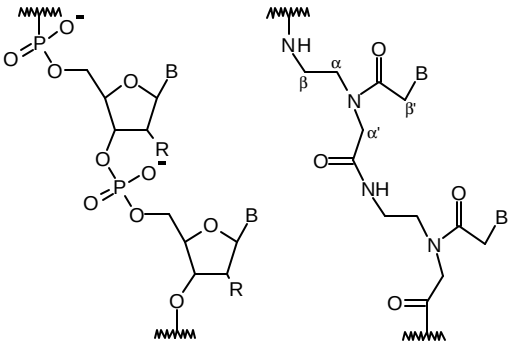
6. Chiral Charged Peptide Nucleic Acid Oligomers from Cyclic Monomers-I.
   US Patent Application filed.
   V. A. Kumar, **M. D'Costa** and K. N. Ganesh
7. Chiral Charged Peptide Nucleic Acid Oligomers from Cyclic Monomers-II.
   European Patent to be filed.
   V. A. Kumar, **M. D'Costa** and K. N. Ganesh

ABSTRACT

The thesis entitled " **SYNTHESIS AND EVALUATION OF CONFORMATIONALLY CONSTRAINED PYRROLIDYL POLYAMIDE NUCLEIC ACIDS** " is divided into four chapters as follows.

Chapter 1: Introduction

The search for effective antigene/antisense agents has drawn considerable attention to DNA/RNA analogues in the past decade. Peptide Nucleic Acids (PNA), the most prominent outcome of this search, is a class of compounds in which the entire charged sugar-phosphate backbone of DNA is replaced by a neutral and achiral polyamide backbone consisting of N- (2-aminoethyl)-glycine units. The nucleobases are attached to the backbone through a conformationally rigid tertiary acetamide linker group. The internucleobase distances in PNA are conserved, allowing its binding to the target DNA/RNA sequences with high sequence specificity and affinity. PNA is also stable to cellular enzymes like nucleases and proteases. However, major limitations to the therapeutic application of PNAs are their poor solubility in aqueous media due to self-aggregation, insufficient cellular uptake and ambiguity in orientational selectivity of binding.



Various attempts have been made to address these problems in an effective manner by many researchers in the recent past. Introduction of chirality into PNA by linking chiral amino acids, peptides, as well as oligonucleotides in the PNA backbone was aimed at improving the orientational selectivity of binding. Chirality to PNA was also imparted by using chiral amino acids in the backbone itself. Most of these efforts had only marginal desired effects on the hybridization properties of PNA. When a five-membered pyrrolidine ring was introduced in PNA to impart structural rigidity and chirality for orientational selectivity, it was found to be ineffective in improving the PNA:DNA binding in a desired manner. Introduction of positive charge in the PNA backbone improved its aqueous solubility. Very recently, considerable interest has been shown in making positively charged PNAs as they are expected to possess superior potential for strand invasion at complementary sequences. Introduction of a positive charge in the PNA backbone was achieved either

by (i) conjugation of cationic peptides to PNA or by (ii) replacement of the tertiary amide linkage to the nucleobase in PNA by a flexible ethylene linker (Figure **1a**) or replacing the amide linkage in the backbone by a guanidinium linkage (Figure **1b**). Although this improved the aqueous solubility of PNA, the detrimental effect on the stability of the PNA:DNA complexes stressed the importance of rigid pre-organization of PNA structure for effective binding to ss/ds DNA.

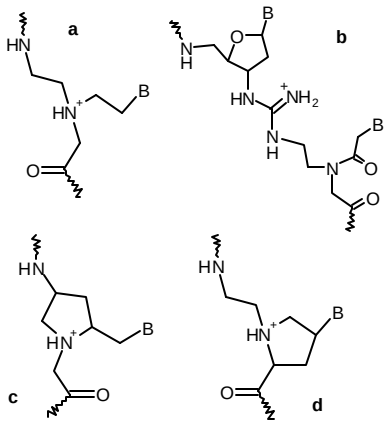


Figure 1

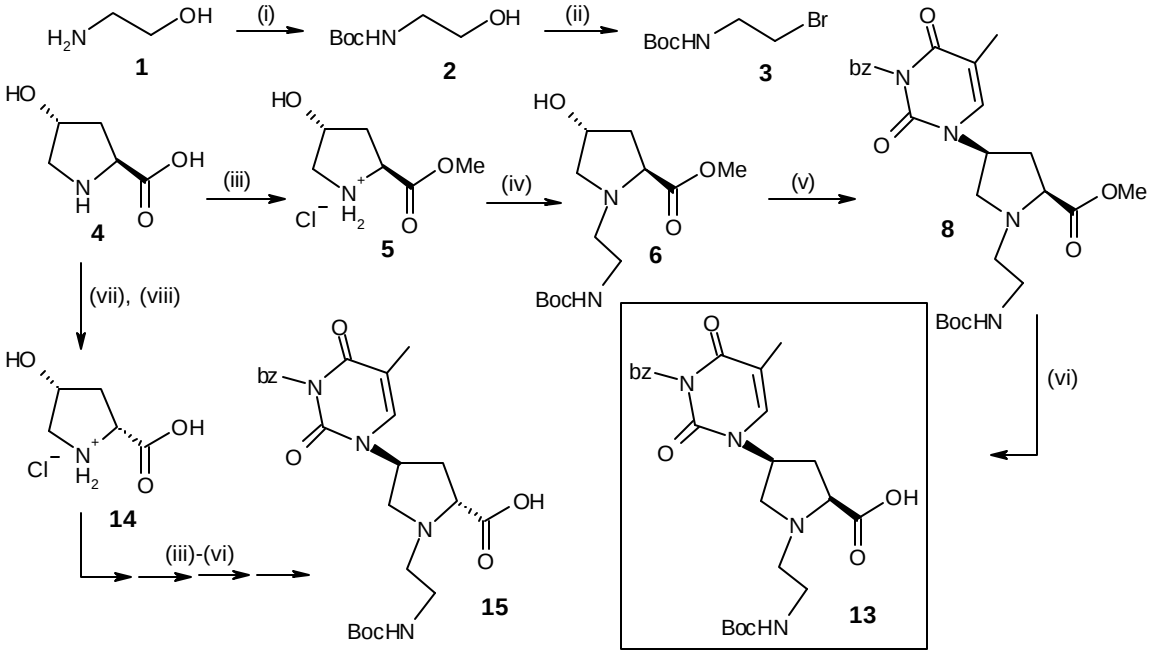
This chapter gives an overview of various modifications of PNAs carried out to improve their application as therapeutic agents. In this work, PNA modifications were designed to impart controlled flexibility, while at the same time, introduce chirality and positive charges in the backbone. The rationale for the design of chiral, pyrrolidine based new PNAs (Figure **1c**, **d**) is presented.

Chapter 2: Chemical synthesis of 1-(N-Boc-aminoethyl)-4-(*S*)-(purinyl/pyrimidinyl)-2-(*S/R*)-proline: cyclic monomers for synthesis of conformationally constrained positively charged PNAs

This chapter describes the modification of neutral, achiral PNA with an aim to improve its hybridization properties to complementary DNA/RNA sequences. A new, chiral aminoethylprolyl (*aep*) backbone is introduced, as against the aminoethylglycyl (*aeg*) backbone of PNA. The five-membered pyrrolidine ring of proline now replaces the tertiary amide linker to the nucleobase in PNA. The nucleobase is attached to the C4 position of the pyrrolidine ring, which has a tertiary ring nitrogen that is protonated at physiological pH. The number of atoms connecting two successive nucleobases is the same as in the case of DNA or PNA. The synthesis of two of the four possible stereoisomers of 1-(N-Boc-aminoethyl)-4-(*S*)-(purinyl/pyrimidinyl)-2-(*S/R*)-proline is described. In addition to the four naturally occurring nucleobases (A, T, G, C), the synthesis of the corresponding monomer bearing the unnatural fluorescent nucleobase, 2-aminopurine is also described.

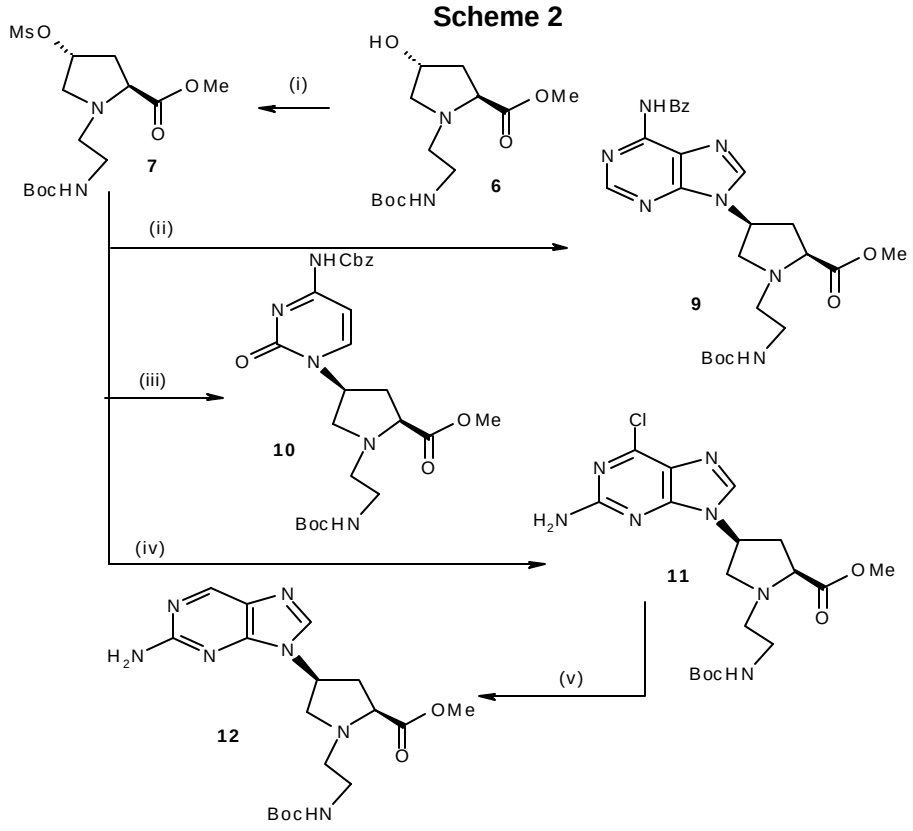
The synthesis of the 1-(N-Boc-aminoethyl)-4-(*S*)-(purinyl/pyrimidinyl)-2-(*S*)-prolines was achieved starting from the naturally occurring 4-(*R*)-hydroxy-2-(*S*)-proline **1**. It involved esterification of the carboxylic acid function followed by alkylation of the

Scheme 1



(i) Boc-N_3, dioxane/H_2O (ii) CBr_4, DIAD, PPh_3, benzene (iii) $SOCl_2$, MeOH (iv) BocHN-$(CH_2)_2$-Br, K_2CO_3, CH_3CN/DMF (v) N3-Bz-T, DIAD, PPh_3, THF (vi) NaOH, MeOH/H_2O (vii) Ac_2O, AcOH (viii) HCl

Scheme 2



(i) Ms-Cl, pyridine (ii) N^6-Bz-A, K_2CO_3, 18-crown-6, DMF (iii) N^4-Cbz-C, K_2CO_3, 18-crown-6, DMF (iv) 2-amino-6-chloropurine, K_2CO_3, 18-crown-6, DMF (v) H_2, Pd-C, MeOH

pyrrolidine ring nitrogen as the key step. The monomer **8** bearing thymine as the

nucleobase was obtained by a Mitsunobu inversion at C4 through substitution of the 4-hydroxyl group by N3-benzoylthymine (**Scheme 1**).

The monomeric units bearing the other nucleobases, *viz*. adenine, cytosine and 2-amino-6-chloropurine (**9-11**) were prepared from 1-(N-Boc-aminoethyl)-4-(*R*)-hydroxy-2-(*S*)-proline methyl ester **6** *via* a mesylation step, followed by alkylation (**Scheme 2**). The 2-aminopurine monomer **12** was prepared from the 2-amino-6-chloropurine monomer **11** by hydrogenation. The corresponding N-Boc-protected monomeric units with free carboxylic acids were obtained after hydrolysis of the ester function. The monomers of 1-(N-Boc-aminoethyl)-4-(*S*)-(purinyl/pyrimidinyl)-2-(*R*)-proline were synthesized from 4-(*R*)-hydroxy-2-(*R*)-proline **14**, which in turn was prepared from 4-(*R*)-hydroxy-2-(*S*)-proline **1** by inversion of the C2-stereocenter, following a similar series of reactions. The monomers were all unambiguously characterized by NMR and mass spectroscopy.

The monomers with different stereochemistries were incorporated into oligomers at different, pre-determined positions using solid phase peptide synthesis employing the Boc-protection strategy. The peptide oligomers were cleaved from the Merrifield support using TFMSA/TFA, which yielded oligomers bearing a 'C'-terminal β-alanine residue with a free carboxylic acid. These were suitably purified and characterized by MALDI-TOF mass spectrometry prior to studying their binding properties.

Chapter 3: Biophysical studies of 1-(N-Boc-aminoethyl)-4-(*S*)-(purinyl/pyrimidinyl)-2-(*S/R*)-proline-containing oligomers

PNA sequences

1	H- T T T T T T T t -(β-Ala)-OH
2	H- T T T t T T T t -(β-Ala)-OH
3	H- T t T t T t T t -(β-Ala)-OH
4	H- t t t t t t t t -(β-Ala)-OH
5	H- T T T T T T T T -(β-Ala)-OH
6	H- G T A G A T C A C T-(β-Ala)-OH
7	H- G t A G A t C A C t-(β-Ala)-OH
8	H- T T A T T A T T A T A T-(β-Ala)-OH
9	H- T T A T T A T T A T A t-(β-Ala)-OH
10	H- t t a t t a t t a t a t-(β-Ala)-OH

A/T = *aeg*PNA-A/T, a/t = *aep*PNA –A/T

DNA sequences

11	5'- G C A A A A A A A A C G -3'
12	5'- G C A A A T A A A A C G -3'
13	5'- A T A T A A T A A T A A T -3'

This chapter encompasses an extensive study of the binding properties of the modified PNA oligomers (**1**-**10**) with complementary DNA sequences. It involved the use of various biophysical techniques like UV, CD and fluorescence spectroscopy. The binding stoichiometry of PNA:DNA complexes was ascertained by UV-mixing/titration experiments. Complex formation was evaluated by UV-Tm measurements and substantiated by gel electrophoresis data. The *aep*PNA monomer when incorporated into oligomers, enhanced binding of the modified PNAs to complementary DNA as compared to the *aeg*PNA-DNA binding. The association was found to be at least partially electrostatic, as evidenced by salt-dependent studies. Moreover, the binding is highly sequence-specific and is overwhelmingly dictated by the specific hydrogen bonding between the nucleobases, as illustrated by mismatch studies. Figure **1** indicates the UV-melting profiles of some of the PNA:DNA complexes. Mixed sequence oligomers were also synthesized as shown in the table above. These were subjected to biophysical studies and examined for complementary DNA binding affinity. The dissociation-reassociation kinetics were also compared by utilizing UV-Tm hysteresis experiments. The observed results are discussed with respect to PNA structures, along with implications and potential for future work.

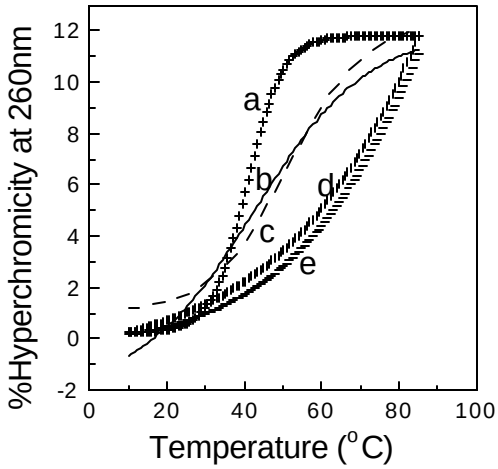


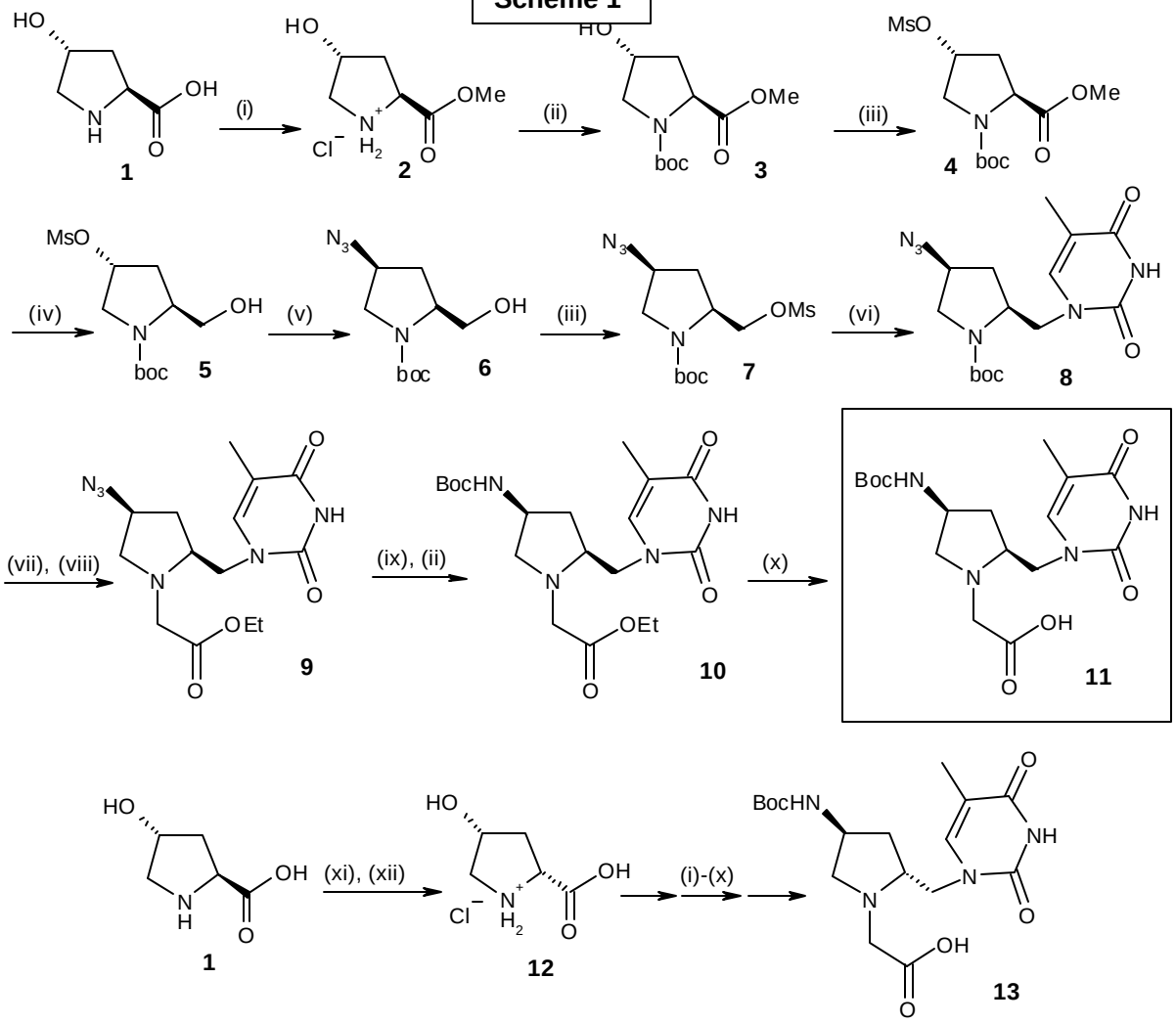
Figure 1: Melting profiles of DNA:PNA$_2$ complexes for the 2*S* modification a: **11:5**, b: **11:1**, c: **11:2**, d: **11:3**, e: **11:4**.

Chapter 4: Synthesis and hybridization application of 4-(*S*)-(N-Boc-amino)-2-(*S/R*)-(purinyl/pyrimidinyl-methyl)-pyrrolidine-N1-acetic acid based PNA oligomers

Our efforts towards generating positively charged PNA analogues were further extended by the use of 4-(*S*)-(N-Boc-amino)-2-(*S/R*)-(purinyl/pyrimidinyl-methyl)-pyrrolidine-N1-acetic acid monomers. This modification allows the introduction of rigidity into the PNA backbone in the form of a pyrrolidine ring, while simultaneously allowing a certain degree of flexibility in the linker to the nucleobase. Here, the β-carbon atom of the ethylenediamine moiety and the α-carbon of the linker to the

nucleobase are bridged by a methylene group. The nucleobase is attached to the C2 position of the pyrrolidine ring through a methylene linker. The pyrrolidine ring nitrogen is a tertiary amine and bears a positive charge at physiological pH. This greatly aids in improving the solubility of the monomer as well as the PNA oligomers containing these

Scheme 1



(i) $SOCl_2$, MeOH (ii) Boc-N_3, Et_3N, dioxane/H_2O (iii) Ms-Cl, pyridine (iv) $NaBH_4$: LiCl, 1:1, EtOH (v) NaN_3, DMF (vi) Thymine, K_2CO_3, 18-crown-6, DMF (vii) TFA:CH_2Cl_2, 1:1 (viii) Br-CH_2-COOEt, aq. Na_2CO_3 (ix) H_2, Pd-C, MeOH (x) NaOH, MeOH/H_2O (xi) AcOH, Ac_2O (xii) HCl

monomeric units in aqueous systems. This chapter comprises the synthesis of 4-(*S*)-(N-Boc-amino)- 2- (*S/R*)- (purinyl/pyrimidinyl- methyl)- pyrrolidine- N1- acetic acid monomers. The synthesis was carried out using 4-(*R*)-hydroxy-2-(*S*)-proline as the starting synthon (**Scheme 1**). The 4-hydroxyl group was mesylated after the carboxyl and amino functions were protected as methyl ester **2** and N-Boc derivatives

respectively. The reduction of the methyl ester was followed by conversion to the corresponding azide **6**, with inversion of the C4-stereocenter. The nucleobase was attached at the C2-methyl position *via* a mesyl intermediate **7**. The pyrrolidine ring nitrogen was alkylated after Boc-deprotection. Hydrogenation and Boc-protection furnished 4-(*S*)-(N-Boc-amino)-2-(*S*)-thymin-1-ylmethyl)-pyrrolidine-N1-ethyl acetate **10**. Subsequent hydrolysis of the ester function gave the desired monomer with free carboxylic acid, **11**, The 4-(*S*)-(N-Boc-amino)-2-(*R*)-(thymin-1-ylmethyl)-pyrrolidine-N1-acetic acid **13** was prepared after inversion of the C2-stereocenter in **1**, following the same series of reactions as above. The monomers bearing other nucleobases were also synthesized using the above synthetic route. All monomers were satisfactorily characterized by NMR and mass spectroscopy. These monomers are suitable for incorporation into PNA oligomers by solid phase peptide synthesis employing the well-known Boc-protection methodology. The biophysical studies of the complexes of such oligomers with complementary DNA sequences in order to evaluate their binding affinity and specificity are described.

ABBREVIATIONS

β-ala	β-alanine
A	Adenine
aeg	Aminoethylglycine
aep	Aminoethylprolyl
ala	Alanine
ap	Antiparallel
C	Cytosine
Cbz	Benzyloxycarbonyl
CD	Circular Dichroism
dA	2'-Deoxyadenosine
dC	2'-Deoxycytidine
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
DCU	Dicyclohexyl urea
dG	2'-Deoxyguanosine
DIAD	Diisopropylazodicarboxylate
DIPCDI	Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMF	N, N-dimethylformamide
DNA	2'-deoxynucleic acid
ds	Double stranded DNA
eda	Ethylenediamine
EDTA	Ethylenediaminetetraacetic acid
FAB	Fast Atom Bombardment
Fmoc	9-Fluorenylmethoxycarbonyl
FPLC	Fast Protein Liquid Chromatography
g	gram
G	Guanine
gly	Glycine
h	hours
his	Histidine
HOBt	1-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography

Hz	Hertz
IR	Infra Red
MALDI-TOF	Matrix Assisted Laser Desorption Ionization – Time of Flight
MF	Merrifield Resin
mg	milligram
MHz	Megahertz
µM	micromolar
ml	milliliter
mM	millimolar
mmol	millimoles
MS	Mass Spectrometry / Mass Spectrum
N	Normal
nm	nanometer
NMR	Nuclear Magnetic Resonance
p	Parallel
PCR	Polymerase Chain Reaction
PPh_3	Triphenylphosphine
PNA	Peptide Nucleic Acid
Pro	Proline
R	Rectus
RNA	Ribonucleic acid
r.t.	room temperature
S	Sinister
ss	Single strand / single stranded
T	Thymine
t-Boc	tert-Butoxycarbonyl
TEA	Triethylamine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethanesulfonic acid
THF	Tetrahydrofuran
UV-Vis	Ultraviolet-Visible

CHAPTER 1

INTRODUCTION

1.1. INTRODUCTION

The recognition of DNA or RNA sequences by complementary oligonucleotides is a central feature of biotechnology and is important for hybridization-based biological applications. Such recognition underlies widely used experimental techniques and diagnostic protocols and makes it possible to consider antisense- or antigene-based inhibition as a practical approach to therapeutics. Zamecnik and Stephensen were the first to propose the use of synthetic antisense oligonucleotides for therapeutic purposes.[1] The specific inhibition is based on the Watson-Crick base-pairing between the heterocyclic bases of the antisense oligonucleotide and of the target nucleic acid.

Various cellular processes can be inhibited depending on the site at which the oligonucleotide hybridizes to the target nucleic acid (Figure 1). For an 'antisense' oligonucleotide to be able to inhibit translation, it must reach the interior of the cell unaltered. The requirements for this are the stability of the oligonucleotide towards extra- and intra-cellular enzymes and equally importantly, its ability to traverse the cell membrane. Once within the cytoplasm, it must bind the target mRNA with sufficient affinity and high specificity. In addition, it should possess an adequate half-life in order to elicit its action. The toxicity of this oligonucleotide should also be negligible to the cell. In the conceptually similar 'antigene' approach, the therapeutic oligonucleotide is targeted to the complementary DNA sequence to tackle the problem at its origin.

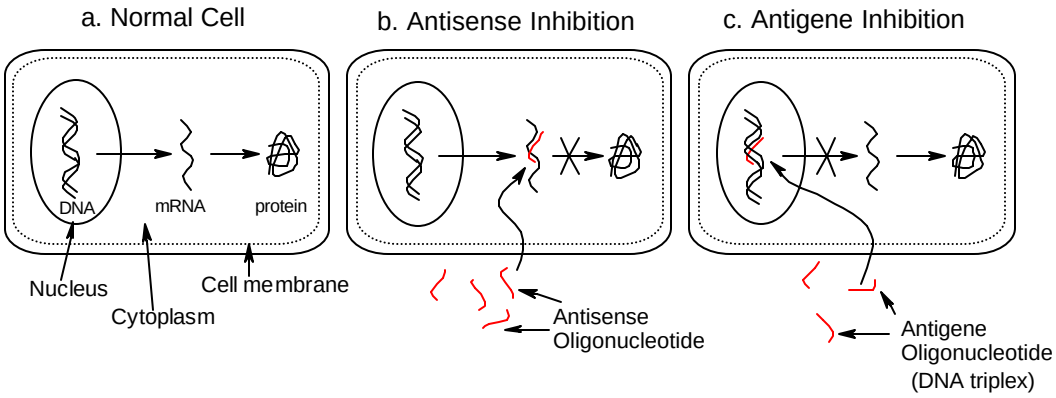


Figure 1: Principle of action of antisense and antigene oligonucleotides.

In order to meet all the requirements of successful antisense/antigene oligonucleotides, it is necessary for normal oligonucleotides to be chemically modified in a suitable manner. To address the combined task of improving the rate, affinity or specificity of oligonucleotide recognition, increasing the membrane permeability and resistance to nuclease digestion, several modifications of DNA have been attempted[2,3] (Figure 2). These include modification of the sugar-phosphate backbone and/or the

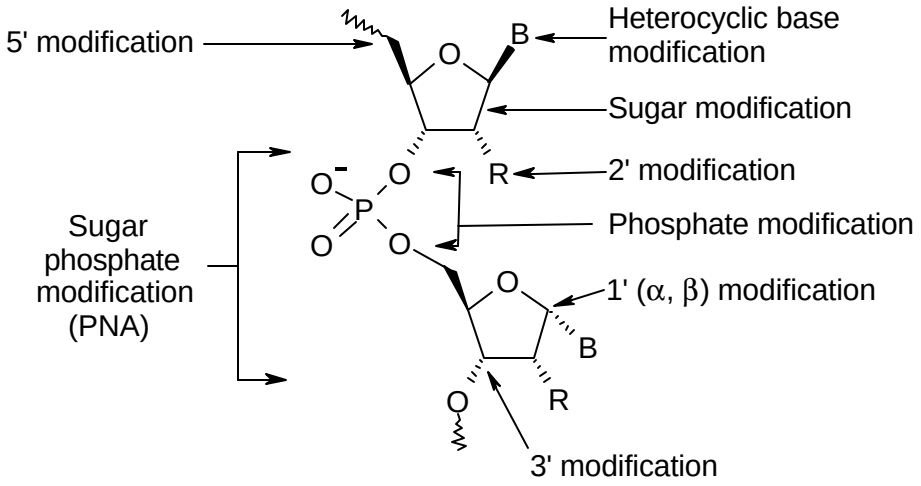


Figure 2. Structurally possible DNA modification sites.

nucleobases. The phosphate modifications such as the phosphorothioates, phosphorodithioates, methylphosphonates, phosphoramidates and phosphotriesters form the first generation 'antisense' oligonucleotides and have already shown promising results,[4] with one drug based on the phosphorothioates already approved.[5] Most modifications replaced the phosphodiester linkage by other four-atom W-X-Y-Z chains[3,6] (Figure 3). Of these, a few bind complementary DNA/RNA fairly well, but none have exhibited the potency to be an effective drug.

The replacement of the ribose sugar by hexose or carbocycles has not been very successful[7] in terms of specificity of binding/hybridization. However, morpholino oligomers, where the monomers are linked through neutral carbamate linkages[8] or

through phosphoramidate linkages[3] (Figure 4a), have shown promising antisense activity as they have superior permeability properties.

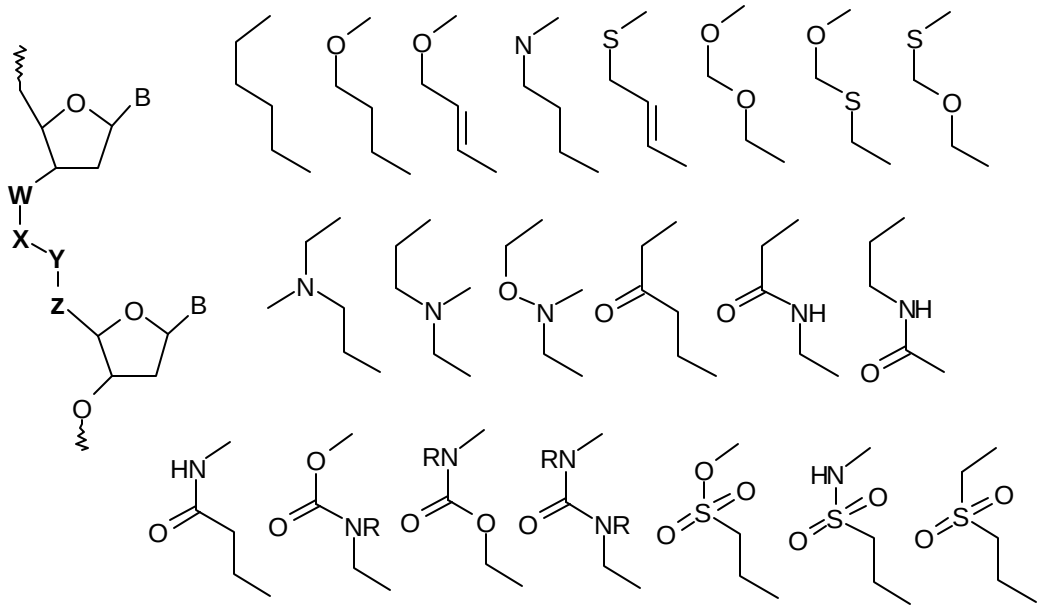


Figure 3: Phosphodiester Linkage Modifications

The locked nucleic acids (LNAs)[9] invented by Wengel *et al.* were found to exhibit unprecedented stability of their complexes with complementary DNA and RNA. They are also stable to 3'-exonucleolytic degradation and possess good water solubility. LNAs are oligonucleotides containing one or more 2'-O, 4'-C-methylene-β-D-ribofuranosyl nucleotides (Figure 4b). The conformational preorganization of LNA could be instrumental in imparting the enhanced binding affinity to DNA. However, more biochemical investigations are required to evaluate their success in the antisense/antigene context.

The most successful outcome of the chemical modification approach has probably been the more recently reported peptide nucleic acids (PNAs)[10] (Figure 5). PNAs are homomorphous DNA analogues in which 2-aminoethylglycine moieties replace the normal phosphodiester backbone. The nucleobases are connected to the amino nitrogens of this backbone by a methylene carbonyl linker, which renders PNA neutral

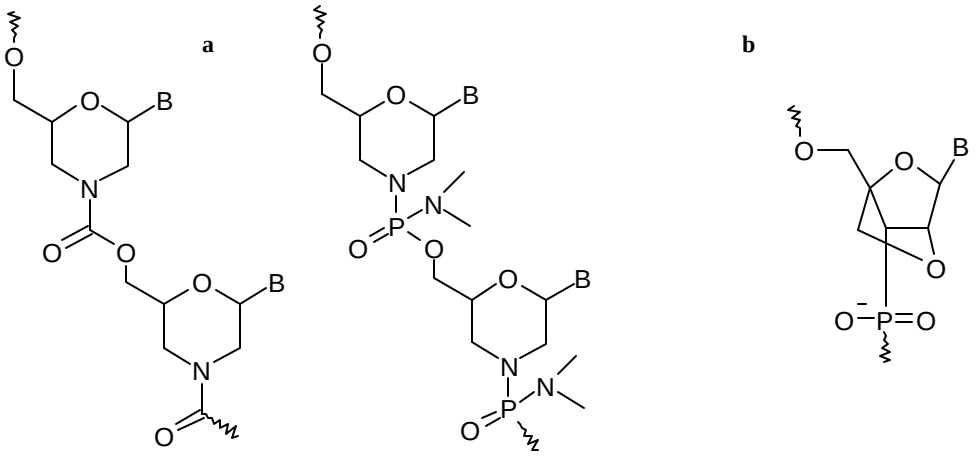


Figure 4. **a**. Morpholino oligomers and **b**. LNA

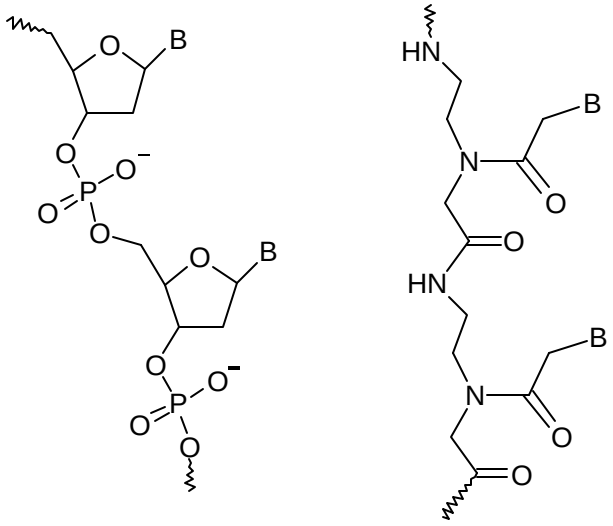
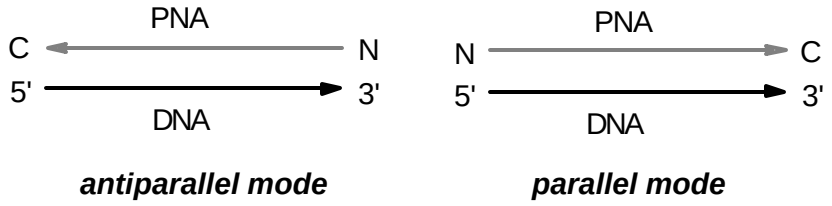


Figure 5. DNA and PNA structures

and achiral. In spite of this dramatic alteration, PNAs recognize complementary DNA/RNA by Watson-Crick hydrogen bonding and thus, are true DNA mimics in terms of base-pair recognition. PNAs lack 3' to 5' polarity and consequently, can bind in either parallel or antiparallel modes, the antiparallel mode being slightly preferred over the parallel one.[11] The antiparallel mode refers to the instance when the PNA 'N' terminus lies towards the 3'- end and the 'C' terminus, towards the 5'- end of the complementary DNA/RNA oligonucleotide. Likewise, the parallel mode of binding is

said to occur when the PNA 'N' terminus lies towards the 5'- end with the 'C' terminus towards the 3'- end of the complementary DNA/ RNA oligonucleotide, Scheme 1.

Scheme 1. Parallel and antiparallel modes of PNA-DNA binding

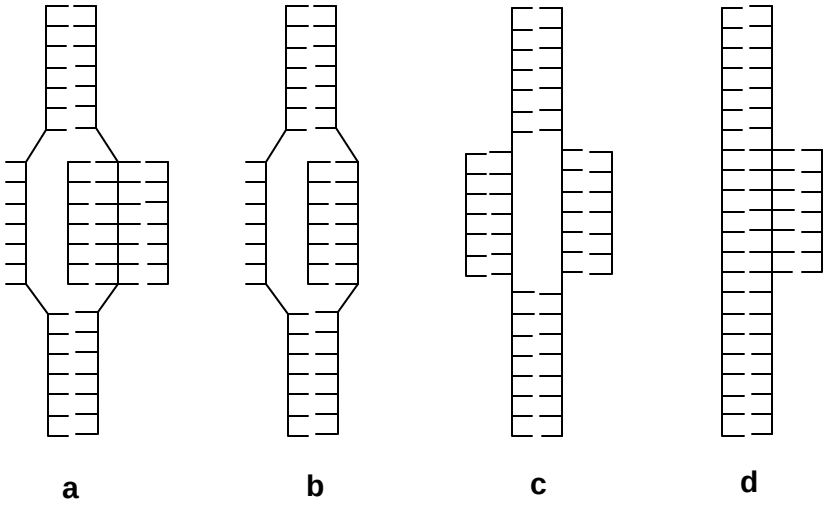


1.2. STRUCTURE OF PNA:DNA COMPLEXES

Homopyrimidine PNAs bind to double stranded DNA targets not by triplex formation as observed with the corresponding homomorphous DNA, but by a highly unusual and interesting mechanism of strand invasion (Scheme 2). Homopyrimidine polyT PNAs bind to the complementary DNA polyA strand forming a PNA_2:DNA triplex and displacing the DNA poly-T strand forming a P-loop or a D-loop structure (Scheme 2a).[12] Homopurine PNA can invade a target DNA duplex (although with a lower efficiency) and form a PNA:DNA duplex, with displacement of one strand of the original DNA duplex (Scheme 2b).[13] Recently, pseudo-complementary PNAs have been demonstrated to invade the target DNA duplex by double-duplex invasion, forming two PNA:DNA duplexes wherein, each PNA strand pairs with its complementary DNA strand (Scheme 2c).[14] In contrast to the strand displacement mode of binding, cytosine-rich homopyrimidine PNAs bind to target DNA duplexes as a third strand forming a PNA_2:DNA triplex (Scheme 2d).[15] The complexes formed by PNA with either DNA or RNA are in general, thermally more stable than the corresponding DNA:DNA or DNA:RNA complexes.[11,17] PNA:DNA duplexes are more stable when the purines are in the PNA rather than the DNA strand. Among the duplexes involving PNA, the generally observed thermal stability is of the order PNA:PNA > PNA:RNA > PNA:DNA.

The first report elucidating the structure of a PNA-nucleic acid hydrogen-bonded complex was put forth by Brown *et al.*[18] This NMR solution structure of a hexameric PNA, GAACTC, with complementary RNA revealed a 1:1 complex that is an antiparallel, right-handed double helix with Watson-Crick pairing similar to the 'A' form structure of RNA duplexes. The achiral PNA backbone was found to assume a distinct conformation upon binding to the chiral RNA.

Scheme 2. Strand Invasion Complexes



a. Triplex invasion **b.** Duplex invasion **c.** Double duplex invasion **d.** Third strand binding forming a PNA:DNA_2 complex

This was followed by a crystal structure of a PNA_2:DNA triplex.[19] The PNA hairpin used was discovered to give a 'P-type' helix that differed from previously reported nucleic acid structures. This helix was underwound, with a base tilt similar to B-form DNA. The bases were even more displaced from the helix than in A-form DNA. The deoxyribose sugars all have a C3'-endo conformation with an average interphosphate distance of 6.0 Å, similar to A-form DNA. This conformation is consistent with the observation that PNAs, including hairpins, bind more tightly to RNA than DNA.[20] The tilt of the base triplets is however, similar to that of B-form DNA.

Another structure of a PNA:DNA duplex derived from NMR data[21] and an X-ray crystal structure of a PNA:PNA duplex[22] were also subsequently reported. The PNA:DNA duplex was found to be very similar to the B-conformation of DNA, but preferred a unique different helix form, the P-form. This helix is very wide (28Å diameter) with a large pitch (18base-pairs) and the base-pairs are almost perpendicular to the helix axis.

In general, the PNA backbone seems to be flexible enough to adapt to its hybrid DNA/RNA to a large extent (Figure 6). The oligonucleotide in the PNA:RNA and the PNA:DNA duplexes adopts a conformation close to its natural 'A' and 'B' form respectively in terms of sugar puckering, while the helix parameters have both 'A' and 'B' form characteristics.

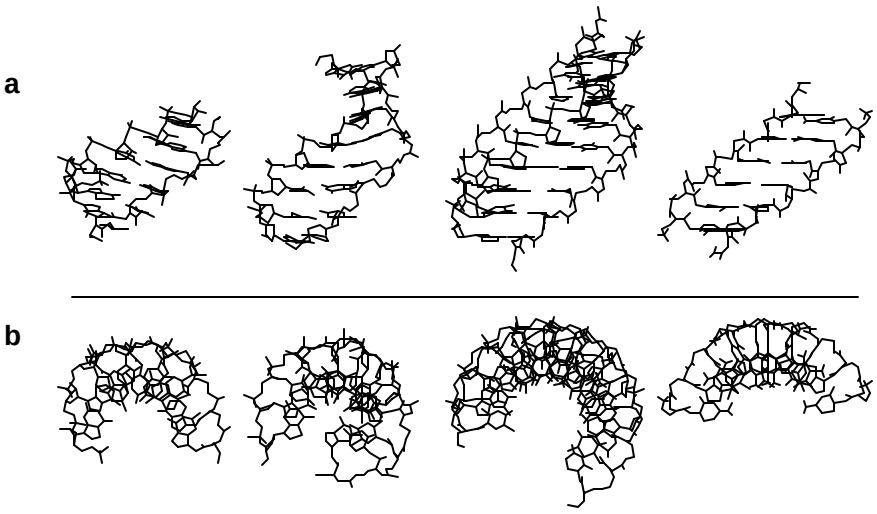


Figure 6 Structures of PNA complexes shown in side view (**a**) and top view (**b**).The complexes from left to right are PNA:RNA, PNA:DNA, PNA:DNA:PNA and PNA:PNA[21]

1.3. BIOLOGICAL APPLICATIONS OF PNA

An endearing feature of PNA applications is the ease of synthesis of the monomeric units from readily available starting materials and the facile adaptability to solid phase peptide synthesis procedures involving both, the Boc as well as the Fmoc protection

strategies. PNA hybridizes with complementary DNA/RNA sequences with high affinity and specificity. Moreover, it is stable to cellular enzymes like nucleases and proteases.[24] These features of PNA, coupled with its interesting properties like strand invasion[25-27] make it an important candidate for biological applications in therapeutics and diagnostics.

1.3.1. Transcription Regulation

The special displacement binding of PNA to complementary DNA has been utilized in either arresting or promoting transcription of genes, by a careful choice of binding sequence and site.

The PNA_2:DNA complex formed by PNA-T_{10} arrested transcription elongation,[28] with the PNA bound to the template strand. Only a marginal effect was observed with PNA bound to the non-template strand. The double-stranded (ds) DNA targets were positioned downstream from phage T3 or T7 promoters in pBluescriptKS$^+$ derived plasmids and transcription was carried out using the T3 or T7 RNA polymerases. With PNAs T_5CT_4 and $T_2CT_2CT_4$ (Figure 7), transcription elongation was arrested, but at 1-3 nucleotides inside the PNA binding site.

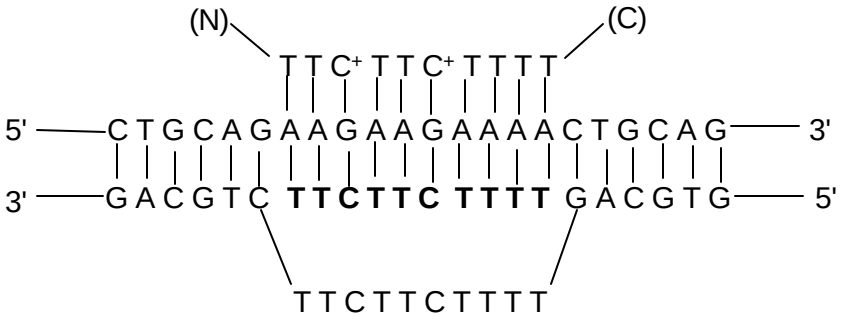


Figure 7. Transcription Elongation Arrest

PNA can be used as a transcription factor to positively control the corresponding ds target DNA that can be considered as an artificial transcription promoter[29] (Figures 8, 9). The loop structures formed by homopyrimidine PNAs when binding to

complementary ds DNA by strand displacement are recognized by RNA polymerase. The enzyme then initiates RNA transcription from PNA:ds DNA complexes at an efficiency comparable to that of the strong *Escherichia coli lac*UV5 promoter.

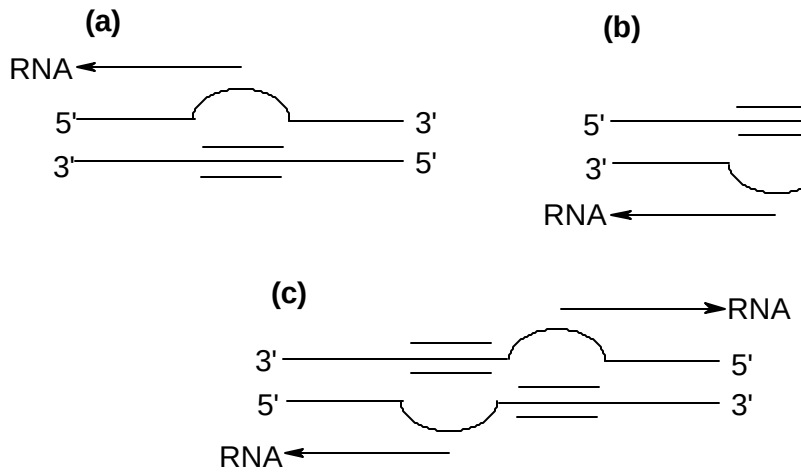


Figure 8. Schematic representation of the RNA transcription sites by reverse transcriptase primer extension showing the transcribed RNA from the PNA promoters

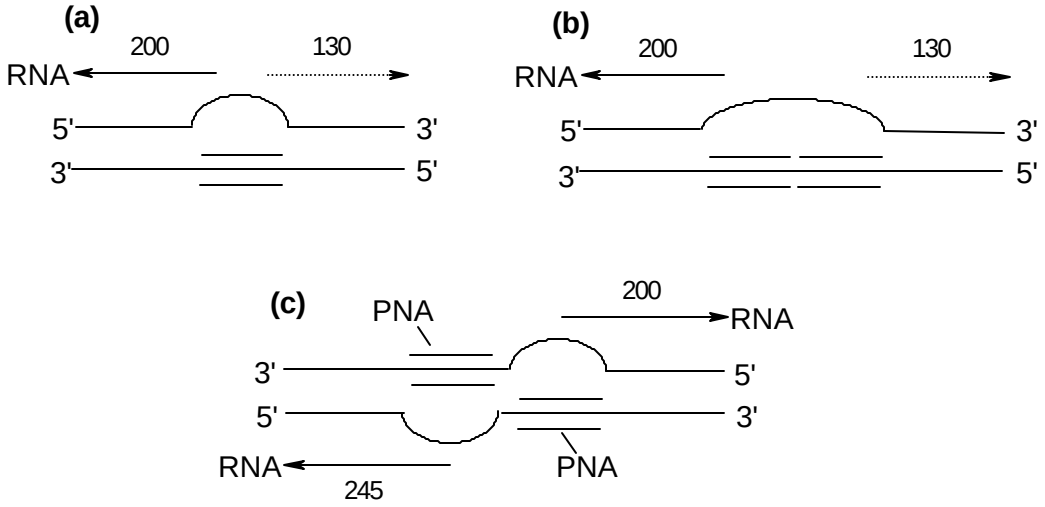


Figure 9 Transcription initiation from PNA:DNA strand displacement loops. A schematic diagram of *in vitro* transcription from purified DNA fragments containing a single **(a)** or a double [two sites in *cis* **(b)** or two sites in *trans* **(c)**] PNA target showing the displacement loops and the start and possible directions of RNA synthesis indicated by arrows (full line: observed).

Specific transcription initiation was also found to take place at the PNA loop in a eukaryotic system using nuclear extracts from rat spleen cells (Figure 10).

Recently, induction of gene expression of a reporter gene construct was demonstrated both *in vitro* and *in vivo* using PNAs designed to bind to the 5' flanking region of the γ-globin gene.[30] PNA-mediated induction of endogeneous γ-globin gene expression was also demonstrated in K562 human erythroleukemia cells.

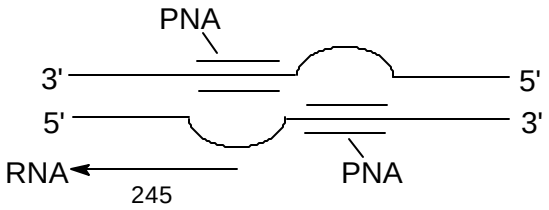


Figure 10. Transcription induced by PNA in a eukaryotic system. A schematic representation showing the strand displacement complex and the RNA transcription in only one direction that was detected.

1.3.2. Translation Inhibition

In vitro inhibition of translation and transcription of the promyelocytic leukemia/ retinoic acid receptor α (PML/RARα; P/R) hybrid gene present in acute promyelocytic leukemia cells was obtained with PNA, which proved to be approximately 40 times more active than antisense oligonucleotides representing the same sequence.[30] In another example, *in vitro* translation arrest was observed.[31] Incubation of PNA with the target mRNA showed concentration-dependent termination of translation elongation products while the random PNA control sequence with the same base composition did not cause any significant translation arrest. The antisense effect observed has been attributed to the steric blocking caused by the PNA with the mRNA target in the form of a PNA:RNA complex.

1.3.3. PCR- based Applications

1.3.3a PNA-directed PCR clamping[32]

PNAs recognize and bind to complementary nucleic acid sequences with higher affinity than the corresponding DNA oligomers and cannot function as primers for DNA polymerases. Thus, a PNA:DNA complex can effectively block the formation of a PCR product when the PNA is targeted to one of the PCR primer sites (Figure 11). This has been shown to allow selective amplification/suppression of target sequences that differ by only one base pair, allowing the direct analysis of a single base mutation by PCR. Interestingly, clamping can operate efficiently even with incomplete binding of PNA to its DNA targets. Clamping could also be accomplished when the PNA target site was located in between the two PCR primers.

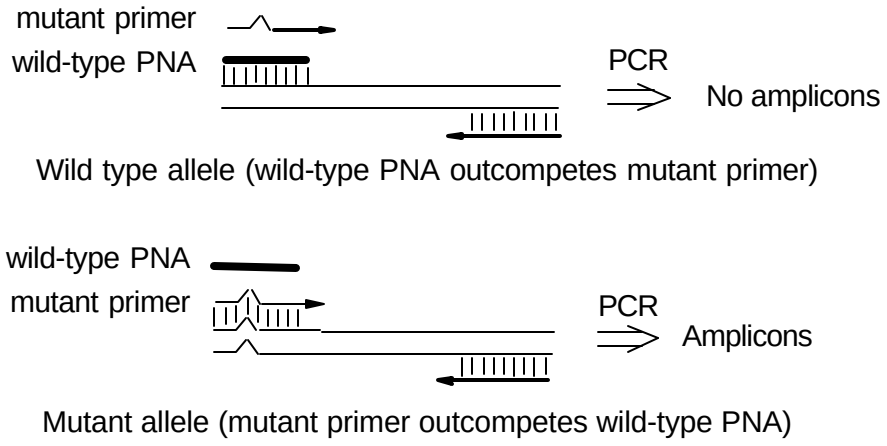


Figure 11. Schematic illustration of the PCR clamping method where wild-type PNA competes for binding to a common target with mutant PCR primer, or vice versa.

When the PNA and PCR primer target sites overlap, clamping operates by 'primer exclusion'. When the target site is located at a distance from the PCR primer sites, clamping is expected to operate by preventing read-through by the *Taq* polymerase ('elongation arrest'). When the PNA target is located adjacent to the PCR primer site, clamping may operate either by preventing polymerase access to the PCR primer and/or by preventing initiation of primer elongation.

1.3.3b PNA primer extension.[33] PNAs were demonstrated to serve as primers at least for certain DNA polymerases and reverse transcriptases, even though they lack the negatively charged phosphate residues of the natural nucleic acid substrates that were presumed to be necessary for binding to the enzymes *via* highly conserved amino acid residues. For this purpose, the PNA was derivatized to carry a 5'-amino-5'-deoxythymidine unit at its 'C' terminus (Figure 12a). This primer was not accepted by some DNA polymerases [e.g., phage T4, phage T7 exo (Sequenase 2.0)], while it was accepted by a few others [e.g., Klenow fragment of DNA pol I (*Escherichia coli*) and Vent DNA pol (*Thermococcus litoralis)*].

A similar monocharged primer was synthesized[34] (Figure 12b), but was found to have a lower efficiency than the previous primer due to stronger destabilization and structural perturbation of the duplex at the PNA-DNA junction.

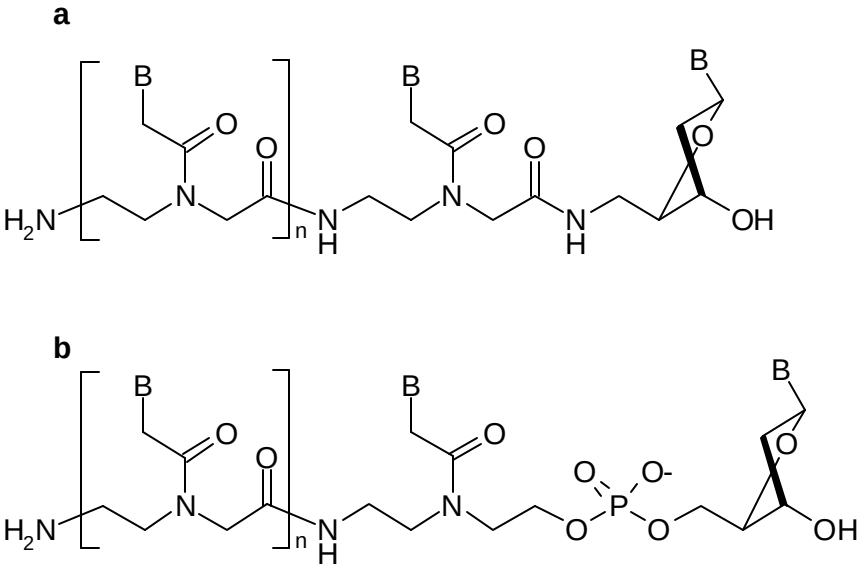


Figure 12. PNA primers

1.3.4. Inhibition of Human Telomerase

Telomerase is a rIbonucleoprotein[35] and possesses an RNA component that can be targeted to effect inhibition of enzyme activity. The designed PNAs were introduced

into cells by transfection using cationic lipids.[36] These PNAs were directed to non-template regions of the telomerase RNA.The problems due to the RNA secondary structure were overcome by intercepting the RNA component prior to holoenzyme assembly, leading to efficient inhibition of telomerase.

The RNA template of telomerase was targeted by peptide conjugated derivatives of a PNA pentamer[37] (Figure 13). It was shown that the presence of cationic peptides at the 'N' terminus of the PNA resulted in enhanced inhibition of telomerase activity. The inhibition was dependent on the specificity of PNA recognition.

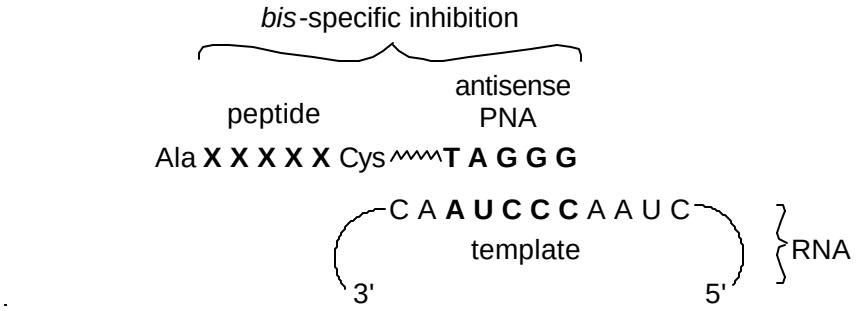


Figure 13. Design of PNA-peptide conjugates for inhibition of human telomerase.

PNAs complementary to the 11-base template of hTR were shown to be potent inhibitors of human telomerase *in vitro*[38] and PNAs were found to be 10-50 times more efficient inhibitors in comparison with phosphorothioate oligomers.

1.3.5. Isolation of mRNA

PNAs composed of *trans*-4-hydroxy-L-proline - based monomers and phosphono PNA monomers were utilized to achieve improved recovery of mRNA molecules with secondary structure at their 3' end as well as RNAs with short polyA tails.[39] By this method, mRNA free of genomic DNA contamination could be isolated.

PNA has also been utilized to capture ds DNA of a particular sequence of interest by affinity capture in the form of linear, non-supercoiled molecules. The classical biotin-streptavidin recognition is utilized for this process[40] (Figure 14).

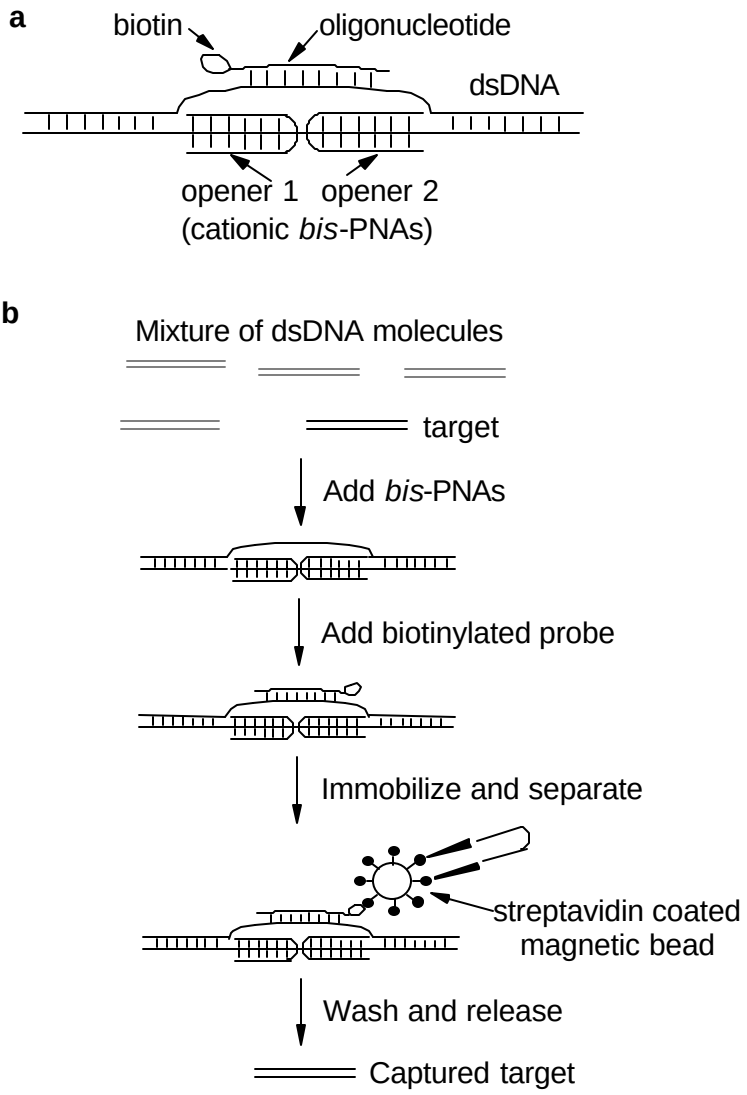


Figure 14. **a** The PD-loop consists of duplex DNA, an ODN and two PNA 'openers'. The ODN binds to the complementary DNA target *via* Watson-Crick pairing and carries biotin to provide capture on the affinity support. **b** The key steps of the procedure for dsDNA biomagnetic isolation.

1.3.6. As Artificial Restriction Enzymes

PNAs in combination with a non-specific nuclease, such as S1 nuclease, have been used as artificial restriction enzymes to cut target DNA at desired positions.[41] Double stranded DNA is cleaved at a site created by PNA strand displacement (Figure 15). This cleavage efficiency is enhanced more than 10 fold when a tandem PNA site is

targeted, and additionally if this site is in *trans* rather than *cis* orientation. Thus, the single strand specific nuclease S1 behaves like a pseudo restriction endonuclease.

The tethering of an artificial nuclease like Gly-Gly-His to PNA exhibits a similar effect. Binding to complementary DNA and in the presence of Ni, cleaves the DNA duplex in its proximity.[42]

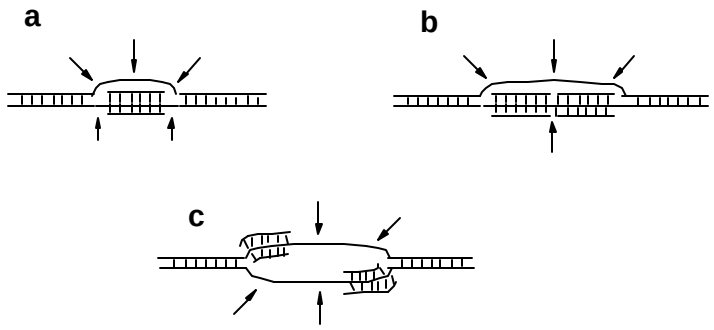


Figure 15. Schematic model for PNA-targeted S1 nuclease ds cleavage of DNA. **a** Single target . **b** Double target in *cis* orientation. **c** Double target in *trans* orientation. Arrows indicate S1 attack.

1.3.7. As a Primordial Genetic Material

The currently widely-accepted theory of the origin of biomolecular life implies that RNA evolutionarily preceded all the other biomolecules, *viz*., DNA, proteins and carbohydrates.[43] The theory gained credibility because RNA provides the link between DNA and proteins. Moreover, RNAs have been discovered to possess catalytic activity (Ribozymes)[44]. However, RNA is highly unstable and it is suspect how prebiotic life could have relied on such a fragile molecule as its genetic material. Miller[45] and Oro[46] demonstrated that under the conditions prevailing on the primitive Earth, nucleobases and amino acids can be easily obtained, whereas, ribose and nucleosides are extremely difficult to obtain under the same conditions. Recently, it was shown[47] that the PNA precursors are possible prebiotic products. In addition, information transfer between PNA and RNA is also possible, although with low efficiency.[48]

1.3.8. Plasmid Labeling

In this approach, a highly fluorescent plasmid DNA is made by hybridizing fluorescently labeled PNA to it. Importantly, the plasmid is neither functionally nor conformationally altered. For this, a PNA homopurine binding site was cloned into a reporter gene plasmid in a region that is not involved in transcription regulation so that PNA-based probes could bind to the plasmid without affecting reporter gene expression. The PNA clamp conjugated to reporter molecules like biotin, fluorescein or rhodamine did not affect the supercoiled conformation, nuclease sensitivity or transcription ability of the plasmid.[49] This method was employed to study the biodistribution of the plasmid upon transfection into cells. By using this system in a plasmid expressing green fluorescent protein (GFP), it was possible to simultaneously follow the delivery of the DNA and the expression of its transgene in real time in living cells.

1.4. CHEMICAL MODIFICATIONS OF PNA

The major limitations of the therapeutic application of PNAs are their poor solubility in aqueous media due to self-aggregation, insufficient cellular uptake and ambiguity in orientational selectivity of binding.[10b,c] Various attempts have been made to address these problems in an effective manner in the recent past.

Improvement of the aqueous solubility of PNAs has been achieved by the introduction of charges in the molecule (Figure 16), or by the introduction of ether linkages in the backbone. Positive charges are endowed on PNA by linking a terminal lysine residue[11,50,51] or by the introduction of a positive charge in the backbone effected by replacing the acetamide linker to the nucleobase by a flexible ethylene linker.[52] Cationic guanidinium linkages have also recently been introduced into PNA and were found to improve its binding affinity with complementary nucleic acid sequences.[53]

Making the PNA anionic also aided in increasing its water-solubility as in the case of the phosphonate analogues (pPNA, Figure 17), but was accompanied by a decrease in the binding affinity to complementary nucleic acid sequences.[54-58]

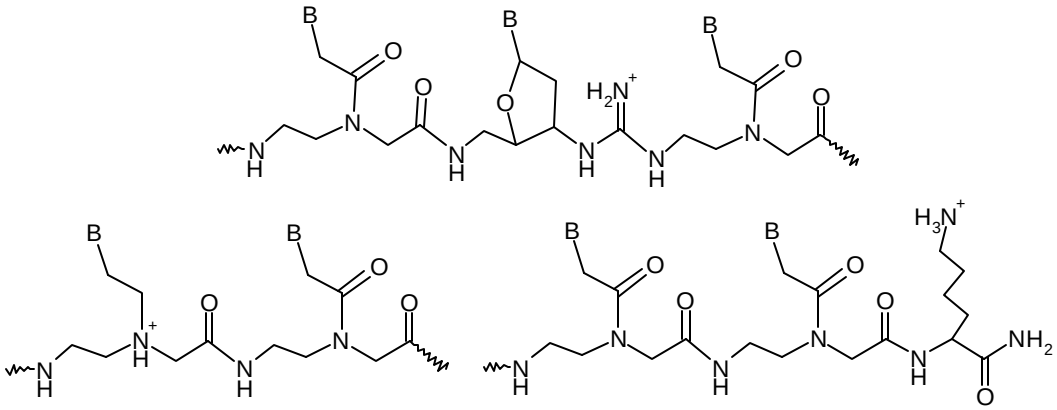


Figure 16. Positively charged PNA

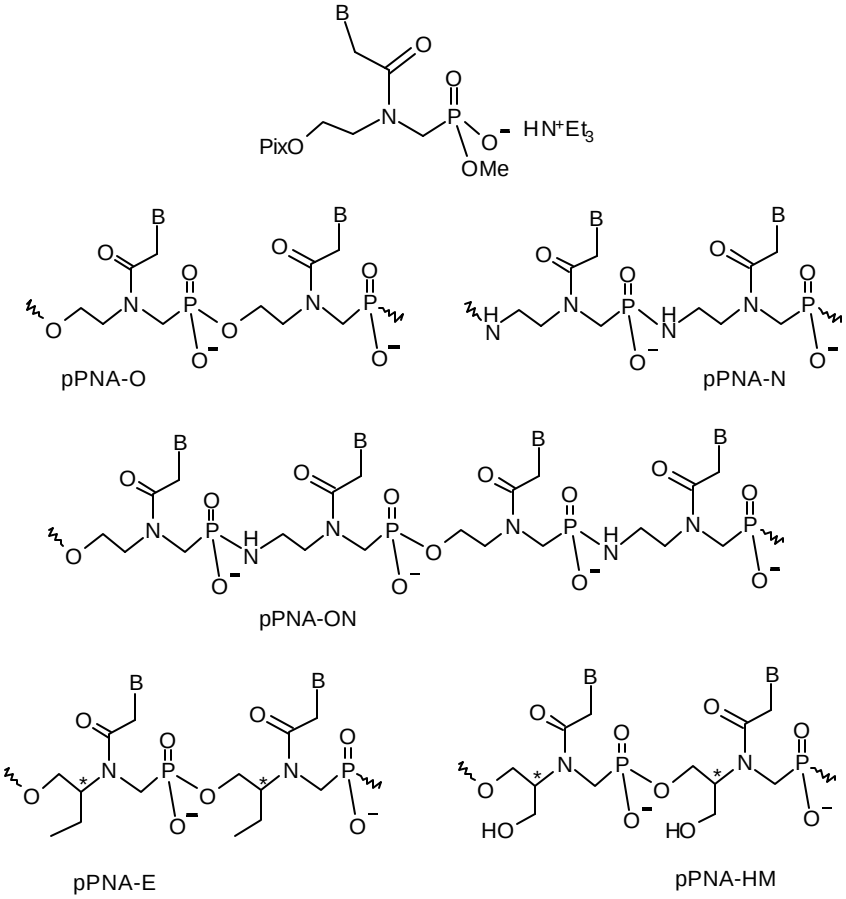


Figure 17. Phosphonate PNA

Efimov *et al.*[54] have described the synthesis of homopyrimidine pPNA oligomers containing *N*-(2-hydroxyethyl)-phosphonoglycine (pPNA-O), or *N*-(2-aminoethyl)-phosphono glycine (pPNA-N). These oligomers did form complexes with complementary DNA/RNA, but with a much lower stability than PNA. Increased water solubility was also achieved by putting an ether linkage in the PNA backbone[59,60] (OPNA, Figure 18).

The issue of insufficient cellular uptake has been addressed by conjugating the PNA with various 'transfer' adducts, which is discussed in a later section.

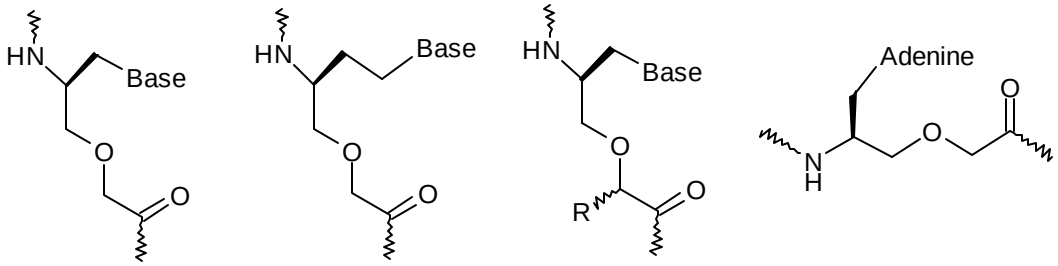


Figure 18. Ether-linked PNA (OPNA)

Ring structures have been introduced into the PNA backbone in several reports. One such effort involving the morpholino oligomers has been described in an earlier section. Another six-membered backbone ring structure is the glucose (hexose)-derived GNA, based on sugar nucleoamino acid monomers[61] (Figure 19). These oligomers were more water-soluble than PNA and bound DNA with a better sequence discrimination ability. Thermodynamic parameters have suggested an entropic gain in GNA due to a pre-organized scaffold.[62]

1, 2-cyclohexylamino and spirocyclohexyl rings (Figure 20 a and b respectively) have been employed in monomers by Lagriffoule *et al.*[63] and Maison *et al.*[64] respectively. PNAs that bear an (*S,S*) cyclohexyl ring in the aminoethyl part of the PNA backbone hybridize with complementary DNA similar to aminoethylglycyl PNA, while those with an (*R,R*) cyclohexyl ring significantly decrease the binding properties.

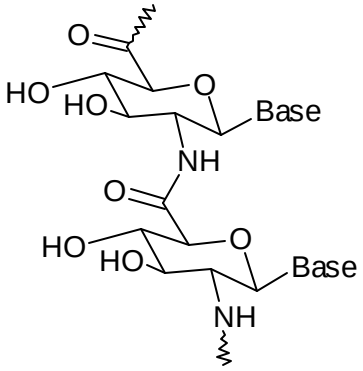


Figure 19. GNA

There have been reports of inclusion of five-membered heterocycles in the PNA backbone too. The foremost in this area was made by Gangamani *et al.*[65] who introduced constraint and chirality simultaneously in the form of a pyrrolidine ring (Figure 21 a). Unfortunately, homooligomers of this unit did not bind complementary DNA. However, interestingly, inclusion of even a single 4aminoproline unit in a PNA sequence either at the 'N' terminal or in the interior, not only led to stabilization of derived hybrids with DNA, but also led to significant discrimination in the orientation of binding.[66] Jordan *et al.*[67,68] then synthesized homopyrimidine oligomers with alternating pyrrolidyl (L-4-*trans*-aminoproline-based) and aminoethylglycyl units (Figure 21 b) and found them to bind complementary DNA better than homooligomers of aminoethylglycyl units. In contrast, backbones derived partially from L-4-*cis*-aminoproline or D-4-*trans*-aminoproline had a reverse effect and destabilized the corresponding complexes with complementary nucleic acids. Pyrrolidine based monomers were also synthesized by Lowe *et al.*[69]

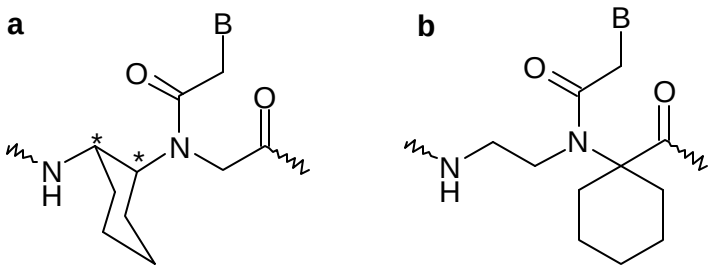


Figure 20. PNAs with hexyl rings in the backbone

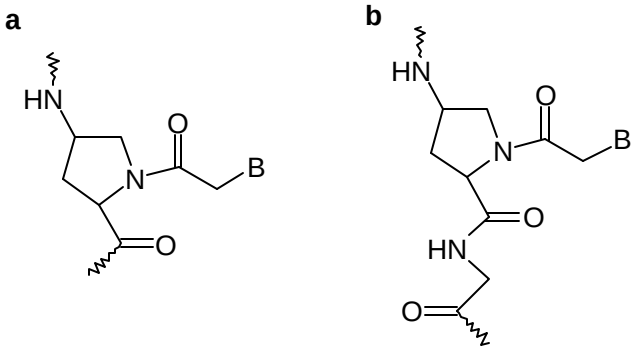


Figure 21. 4-aminoproline-based PNA.

Recently, Hickman *et al.*[70] reported the synthesis and binding properties of a T5-pentamer pyrrolidine oligonucleotide mimic (POM, Figure 22). These oligomers formed stable triplexes with DNA and RNA and exhibited binding selectivity for RNA over DNA. The pyrrolidine ring nitrogen is positively charged, leading to added attributes.

This was followed by work of Püschl *et al.*[71] who reported that a decamer comprising the (2*R*,4*S*) adenine unit of this modification formed a DNA_2:PNA triplex, with the PNA strand forming the central Watson-Crick strand.

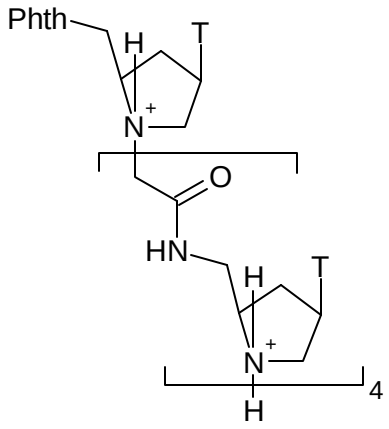


Figure 22**. Pyrrolidine-amide oligonucleotide mimic.**

In order to resolve the orientational ambiguity of binding, several attempts have been made to introduce chirality into PNA. These include covalent linking of a terminal chiral amino acid like lysine[72] or by incorporating chiral centers in the PNA backbone.[63,65,73-75] These were generally found to form slightly less stable complexes

with complementary DNA than the unmodified achiral PNA.[63] The destabilization was more pronounced for backbones containing amino acids with bulky apolar side chains[76] (e.g., those derived from phenylalanine or leucine).

Monomers derived from negatively charged amino acids like aspartic acid and glutamic acid, though helped to improve the aqueous solubility, further destabilized their complexes with DNA, probably due to electrostatic repulsion from the negatively charged phosphate groups.[49]

It was supposed that the D-monomers would induce a preferred right-handedness and the L-monomers, a preferred left-handedness in the PNA strand.[77] As a consequence, DNA being right-handed would bind preferentially with the right-handed PNAs. This supposition was indirectly confirmed by circular dichroism studies of the complexes between cyanine dyes and PNA-complementary nucleic acid complexes.

It was reported very recently that the incorporation of three adjacent D-Lysine residues as a 'chiral box' leads to complementary DNA recognition in exclusively antiparallel fashion and shows no affinity for the parallel target DNA.[78]

Besides the various modifications described above that introduced charge, conformational constraint or chirality in the neutral, flexible, achiral PNA backbone, several attempts to modify the parent aminoethylglycyl structure of PNA are documented. Some of these are discussed herewith.

Early modifications involved the extension of the PNA structure by a methylene group in the aminoethyl, glycyl[79] or nucleobase-linker[80] segments (Figure 23). The significantly lowered stabilities of the derived PNA:DNA hybrids pointed to the importance of constrained flexibility in the backbone.

Another mode of modification was the inversion of the intra- and inter-residue amide bonds leading to the retro-inverso,[81-83] peptoid[84] and heterodimeric[85] analogues (Figure 24). With the exception of the heterodimer analogue, these exhibited a lower potency

for duplex formation with complementary DNA/RNA suggesting that other subtle requirements like hydration and dipole-dipole interactions etc. may play an important role in the process.

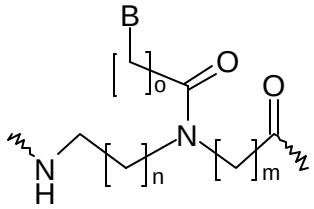


Figure 23. Extension of the PNA structure by methylene groups.

α-Amino acids carrying the nucleobases either on the side-chain[86-88] or on the α-carbon atom[89,90] were used as the monomer building blocks of modified PNAs. Studies on the duplexation of these modified PNAs with DNA are not encouraging.[91]

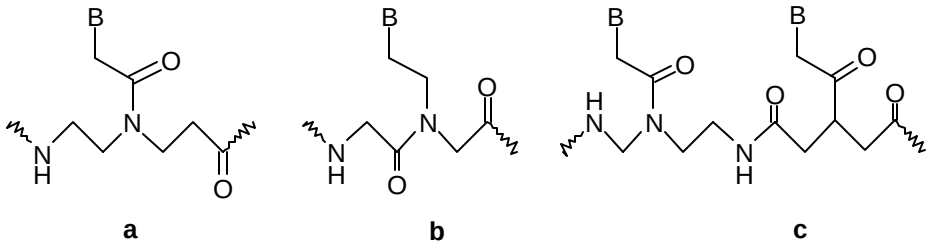


Figure 24. **a**. Retero-inverso, **b**. Peptoid and **c**. Heterodimeric PNA.

1.5. MODIFIED NUCLEOBASES

Non-natural nucleobases could aid the understanding of the recognition process between the natural nucleobase-pairs in terms of factors such as hydrogen bonding and internucleobase stacking. They could also generate new recognition motifs with potential applications in diagnostics. Only a few nucleobase modifications have been reported in the PNA context (Figure 25).

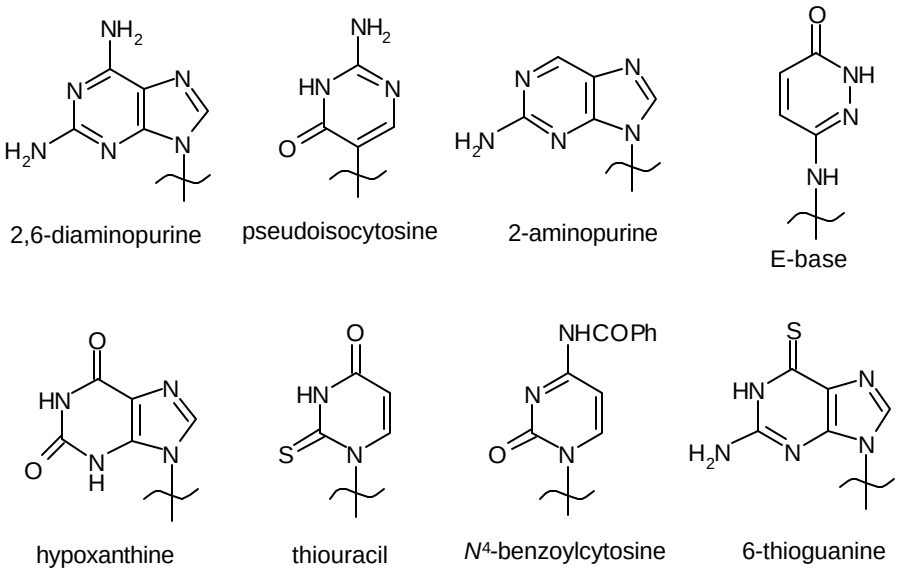


Figure 25. Modified nucleobases.

2,6-Diaminopurine offers increased affinity and selectivity for thymine[92] and pseudoisocytosine mimics the C^+ recognition pattern[93] for triplex formation. 2-Aminopurine[94,95] can hydrogen-bond with uracil and thymine in the reverse Watson-Crick mode and being inherently fluorescent, can be used to study the kinetics of the hybridization process with complementary nucleic acids.

The E-base,[96] hypoxanthine,[97] *N*[4]-benzoylcytosine[98] and 6-thioguanine[99] represent some more examples of modified nucleobases. Thiouracil along with 2,6-diaminopurine has been utilized as a non-natural base-pair in PNA-DNA recognition and was shown for the first time to lead to a phenomenon termed as 'double duplex invasion'.[14]

1.6. PNA CONJUGATES

Covalent hybrids of PNA with other molecules have been constructed to overcome the limitations of PNAs such as aggregation, solubility and cell uptake and to impart abilities to enable therapeutic applications like RNase H activation, cell uptake, etc.

1.6.1. PNA-DNA Chimerae

Several PNA-DNA chimerae[100] have been reported till date. Conjugation of PNA to the 5'-end of DNA led to PNA-(5')-DNA chimerae,[101-105] while conjugation to the 3'-end led to the DNA-(3')-PNA chimerae.[106,107] An advantage of attaching the PNA to the 3'-end of DNA is imparting stability towards the most common 3'-exonucleases.[108] Other advantages of such chimerae are their improved solubility in aqueous media, improved cellular uptake and a lower tendency to self-aggregate. The thermal stability of the complexes of these with complementary nucleic acids was however, lower than that of the complexes with PNA.

Chimerae were also synthesized utilizing 4-hydroxyproline-based linkers. Of these, chimerae comprising the L-*trans* unit were the most promising.[109]

1.6.2. PNA-Peptide Chimerae

There are varied reports in literature of the conjugation of PNA with peptides and proteins to gain an application advantage in biological systems. For example, the presence of cationic peptides at the N-terminus of the PNA resulted in an enhanced inhibition of human telomerase activity.[110] Another example of PNA-peptide chimerae can be found in the PNA-NLS peptide corresponding to the SV40 core nuclear localization signal. This conjugate increased the nuclear uptake of oligonucleotides and enhanced the transfection efficacy of plasmids.[111]

Shuttle proteins upon conjugation with PNA were used to 'smuggle' the target PNA into cells and across the cellular membranes.[112] This significantly increased the inhibition of target RNA expression compared to PNAs alone.

Cationic peptides linked to PNAs were found to enhance the strand invasion capability of PNAs into target DNA duplexes.[113] This complex formation was also found

to prime DNA strand elongation by oligonucleotide-peptide conjugates at sequences where elongation was hitherto undetected.

1.6.3. PNA-Liposome Chimerae

PNAs were conjugated to lipophilic groups and incorporated into liposomes.[114] As predicted, these enhanced the cellular uptake and distribution. These favourable properties increased with the introduction of an amino side-chain into the PNA backbone.

1.6.4. PNA-Polyamine Conjugates

PNAs were conjugated to polyamines like ethylene diamine and spermine in order to improve their solubility properties in aqueous media.[115] Spermine accelerated the association of PNA with anionic DNA by electrostatic interaction. It was envisaged that the conjugation of biologically relevant polyamines such as spermine would enhance the cellular uptake of PNAs through polyamine receptor mediated mechanisms. However, such studies are yet to be reported.

1.7. PRESENT WORK

The preceding sections give an overview of the peptide nucleic acids, PNAs, DNA analogues with a homomorphous but chemically different backbone consisting of *N*-(2-aminoethyl)-glycine units in contrast to the sugar-phosphate backbone of DNA. In spite of this, PNAs bind to complementary nucleic acid oligomers obeying the Watson-Crick hydrogen bonding rules for PNA:DNA duplexes and Hoogsteen hydrogen bonding mode for third strand binding in a triplex. The attractive binding properties of PNAs, both in terms of affinity and specificity, coupled with their strand invasion potential have

promoted PNA as a useful tool in molecular biology, diagnostics, and as a possible candidate for antisense/ antigene drug therapy.

The major factors restricting the applications of PNA have been its poor water solubility, insufficient cell uptake, self-aggregation and ambiguity in the binding orientation. Moreover, the strand invasion phenomenon is restricted to low salt concentrations.

In order to overcome these limitations, several modifications of PNA have been carried out. PNAs have also been linked to helper molecules in various chimerae in an endeavour to improve its favourable properties.

The work presented in this thesis involves the design, synthesis and biophysical evaluation of these backbone modified, chiral, charged PNA analogues- *aep*PNA, N7G-*aep*PNA and pyrrolidyl PNA.

Chapter 2

This chapter describes the synthesis of a novel modified PNA monomer (*aep*PNA), which was envisaged to confer constrained flexibility on the relatively more flexible PNA backbone. The modification introduces two chiral centers per unit in addition to a positive charge. The cationic aminoethylprolyl monomer was conceived by bridging the α-carbon atom of the glycine moiety in PNA and the β'-carbon of the linker to the nucleobase by a methylene group, with a simultaneous omission of the α'-carbonyl group (Figure 26). The chiral monomers were synthesized from the naturally occurring and easily available 4(*R*)-hydroxy-2(*S*)-proline. The synthesis of the chiral monomers bearing each of the four natural nucleobases is described. In addition, the non-natural N7-guanine monomer has also been synthesized. These aminoethylprolyl PNA monomers have been incorporated into PNA oligomeric sequences by solid phase peptide synthesis. Cleavage of the synthesized oligomers from the solid support, their

subsequent purification procedures, followed by suitable characterization is also detailed.

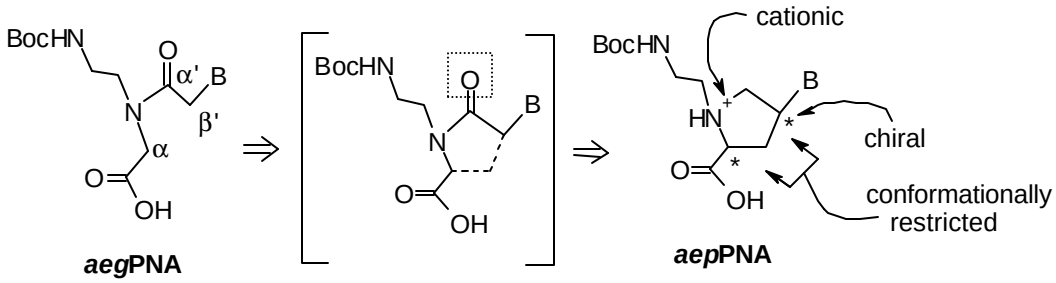


Figure 26. Aminoethylprolyl PNA

Chapter 3

This Chapter is divided into two sections.

Section A

In this chapter, the binding properties of the synthesized PNA oligomers containing the aminoethylprolyl monomers are extensively studied. The biophysical effects of the modification have been elucidated by sequentially increasing the number of modified units in the oligomer. The binding affinities of the PNA oligomers for complementary DNA sequences have been evaluated by temperature-dependent UV and CD spectroscopic studies. The phenomenon of strand invasion of DNA duplexes by PNA oligomers has been demonstrated by fluorescence spectroscopy. Gel retardation experiments provided additional proof of complexation of the PNA oligomers with complementary DNA sequences. The results of the above studies are discussed along with implications and potential for future work.

Section B

This section is devoted to the application of the aminoethylprolyl backbone in resolving the orientational ambiguity of PNA binding to complementary DNA

sequences. Bis-PNAs incorporating N7-guanine have been synthesized with both the aminoethylglycyl as well as the aminoethylprolyl backbone. The selectivity in their binding properties with complementary DNA sequences in the parallel and antiparallel modes have been studied and the interesting results are discussed.

Chapter 4

The PNA modification described in this Chapter is a result of our further efforts to generate positively charged chiral PNA monomers. The modification allows the introduction of rigidity into the flexible PNA backbone in the form of a pyrrolidine ring, while simultaneously allowing a certain degree of flexibility in the linker to the nucleobase. Conceptually, the monomer can be derived from the basic *aeg*PNA structure by bridging the β''-carbon atom of the ethylenediamine moiety and the α'-carbon atom of the linker to the nucleobase by a methylene group, with the removal of the carbonyl group (Figure 27). The nucleobase is attached to the pyrrolidine ring *via* a flexible methylene group. The synthesis of these chiral monomeric units bearing nucleobases is described. Solid phase peptide synthesis of PNA oligomers incorporating these monomeric units at different pre-determined positions has been carried out. Oligomers were cleaved with spermine to yield spermine-conjugated PNAs, and their properties with unconjugated oligomers are compared. The complexes of these oligomers with complementary DNA sequences have been studied for their binding affinity and specificity by UV spectroscopy. The results are discussed along with their implications and scope for further work.

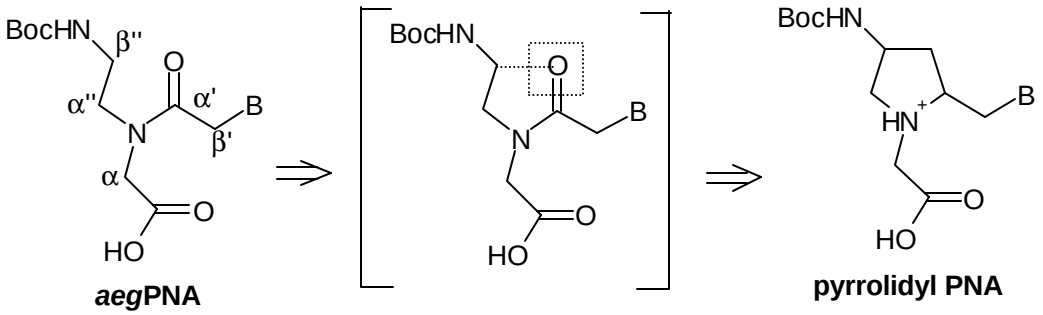


Figure 27. Pyrrolidyl PNA

1.8. REFERENCES

1. (a) Stephensen, M. L.; Zamecnik, P. C. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 285. (b) Zamecnik, P. C.; Stephensen, M. L. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 280.
2. Miller, P. S. in *Bioorganic Chemistry: Nucleic Acids* Hecht, S. M., ed., Oxford University Press) **1996**, pp 347-374.
3. Nielsen, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*,167.
4. Milligan, J.; Matteucci, M. D.; Martin, J. C. *J. Med. Chem.* **1993**, *36*, 1923.
5. (a) Matteucci, M. D.; Wagner, R. W. *Nature* **1996**, *384*, 2103. (b) Glazer, V. *Genet. Engg. News.* **1996**, *16*, 1, 16, 17 & 21.
6. Mesmaeker, A.D; Waldner, A.; Sanghvi, Y. S.; Lebreton, J. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 395.
7. Miculka, C.; Windhab, N.; Eschenmoser,A.; Sherer, S.; Quinkert, G. *PCT Int. Appl. WO 9915509 A2* **1999**, pp43.
8. (a) Adrianov, M.; Buryakova, A.; Choob, M.; Efimov, V. *Nucleosides Nucleotides* **1999**, *18*, 1427- 1428. (b) Stirchak, E. P.; Summerton, J. E.; Weller, D. D. *Nucleic Acids Res.* **1989**, *17*, 6129.
9. (a) Wengel, J. *Acc. Chem. Res.* **1999**, *32*, 301. (b) Håkansson, A. E.; Koshkin, A. A.; Sørensen, M. D.; Wengel, J. *J. Org. Chem.* **2000**, *65*, 5161. (c) Kværnø, L.; Kumar, R.; Dahl, B. M.; Olsen, C. E.; Wengel, J. *J. Org. Chem.* **2000**, *65*, 5167.
10. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500. (b) Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481-1485. (c) Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5. (d) Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem.* **2000**, *4*, 931.

11. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566.
12. Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Mol. Recog.* **1994**, *7*, 165.
13. Nielsen, P. E.; Christensen, L. *J. Amer. Chem. Soc.* **1996**, *118*, 2287.
14. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci.* USA **1999**, *96*, 11804.
15. Praseuth, D.; Grigoriev, M.; Guieysse, A -L.; Pritchard, L. L.; Harel-Bellan, A.; Nielsen, P. E.; Hélène, C. *Biochem. Biophys. Acta* **1996**, *1309*, 226.
16. Wittung, P.; Nielsen, P.; Nordén, B. *Biochemistry* **1997**, *36*, 7973.
17. Jensen, K. K.; Ørum, H.; Nielsen, P. E. Norden, B. *Biochemistry* **1997**, *36*, 5072.
18. Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777.
19. Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science* **1995**, *270*, 1838.
20. (a) Moore, M. S.; Blobel, G.; *Cell*, **1992**, *69*, 939. (b) Stewart, M. *Cell Biol.* **1992**, *3*, 267. (c) Powers, M. A.; Forbes, D. J. *Cell* **1994**, *79*, 931.
21. Eriksson, M.; Nielsen, P. E. *Nature Struct. Biol.* **1996**, *3*, 410.
22. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nature Struct. Biol.* **1997**, *4*, 98.
23. Erikkson, M.; Nielsen, P. E. *Quart. Rev. Biophysics* **1996**, *29*, 369.
24. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sönnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1310.
25. Peffer, N. J.; Hanvey, J. C.; Bisi, J. E.; Thomson, S. A.; Hassman, F. C.; Noble, S. A.; Babiss, L. E. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10648.
26. Møllegaard, N. E.; Christensen, L. *J. Am. Chem. Soc.* **1996**, *118*, 2287.
27. Wittung, P.; Nielsen, P.; Norden, B. *J. Am. Chem. Soc.* **1996**, *118*, 7049.
28. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Gene* **1994**, *149*, 139.
29. Møllegaard, N. E.; Buchardt, O.; Egholm, M.; Nielsen, P. E. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3892.

30. Wang, G.; Xu, X.; Pace, B.; Dean, D. A.; Glazer, P. M.; Chan, P.; Goodman, S. R.; Shokolenko, I. *Nucleic Acids Res.* **1999**, *27*, 2806.

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

32. Ørum, H.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanley, C. *Nucleic Acids Res.* **1993**, *21*, 5332.

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

34. Lutz, M. J.; Will, D. W.; Breipohl, G.; Benner, S. A.; Uhlmann, E. *Nucleosides Nucleotides* **1999**, *18*, 393.

35. Morin, G. B. *Cell* **1989**, *59I*, 521.

36. Hamilton, S. E.; Simmons, C. G; Kathiriya, I. S.; Corey, D. R. *Chem. & Biol.* **1999**, *6*, 343.

37. Harrison, J. G.; Frier, C.; Laurant, R.; Dennis, R.; Raney, K. D.; Balasubraminian, S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1273.

38. Norton, J. C.; Piatyszek, M. A.; Wright, W. E.; Shay, J. W.; Corey, D. R. *Nature Biotech.* **1996**, *14*, 615.

39. Phelan, D.; Hondorp, K.; Choob, M.; Efimov, V.; Fernandez, J. *XIV Intl. RT in Nucleosides & Nucleotides* **2000**.

40. Boffa, L. C.; Carpaneto, E. M.; Allfrey, V. G. *Proc. Natl. Acad. Sci.* **1995**, *92*, 1901.

41. Demidov, V. V.; Bukanov, N. O.; Frank-Kamenetskii, M. D. *Pept. Nucleic Acids* **1999**, 175.

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

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

44. Altman, S. *Compr. Nat. Prod. Chem.* **1999**, *6*, 109.

45. Miller, S. L. *Science* **1953**, *117*, 528.

46. Oro, J. *Biochem. Biophys. Res. Commun.* **1960**, *2*, 407.

47. Nelson, K. E.; Levy, m.; Miller, S. L.; *Proc. Natl. Acad. Sci.* USA **2000**, *97*, 3868.

48. Böhler, C.; Nielsen, P. E.; Orgel, L. E. *Nature* **1995**, *376*, 578.

49. Zelphati, O.; Liang, X.; Ho bart, P.; Felgner, P. L. *Human Gene Ther.* **1998**, *10*, 15.

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

51. Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999**, 197.

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

53. Barawkar, D. A.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 10418.

54. Efimov, V. A.; Choob, M. V.; Buryakova, A. A.; Chakhmakhcheva, O. G. *Nucleosides Nucleotides* **1998**, *17*, 1671.

55. Efimov, V. A.; Buryakova, A. A.; Choob, M. V.; Chakhmakhcheva, O. G. *Nucleosides Nucleotides* **1999**, *18*, 1393.

56. van der Laan, A. C.; Strömberg, R.; van Boom, J. H.; Kuyl-Yeheskiely, E. *Tet. Lett.* **1996**, *37*, 7857.

57. Peyman, A.; Uhlmann, E.; Wagner, K.; Augustin, S.; Breipohl, G.; Will, D. W.; Schäffer, A.; Wallmeier, H. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2636.

58. Kehler, J.; Henriksen, U.; Vekbjerg, H.; Dahl, O. *Bioorg. Med. Chem.* **1998**, *6*, 315.

59. Garcia-Echeverria, C.; Husken, D.; Chiesi, C. S.; Altmann, K -H. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1123.

60. Kuwahara, M.; Arimitsu, M.; Sisido, M. *J. Am. Chem. Soc.* **1999**, *121*, 256.

61. Goodnow Jr., R. A.; Richou, A -R.; Tam, S. *Tetrahedron Lett.* **1997**, *38*, 3195.

62. Goodnow Jr., R. A.; Tam, S.; Pruess, D. L.; McComas, W. W. *Tetrahedron Lett.* **1997**, *38*, 3199.

63. Lagriffoule, P.; Wittung, P.; Eriksson, M.; Jensen, K. K.; Norden, B.; Buchardt, O.; Nielsen, P. E. *Chem. Eur. J.* **1997**, *3*, 912.

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

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

66. Gangamani, B. P.; D'Costa, M.; Kumar, V. A.; Ganesh, K. N. *Nucleosides Nucleotides* **1999**, *18*, 1409.

67. Jordan, S.; Schwemler, C.; Kosch, W.; Kretschmer, A.; Schwenner, E.; Stropp, U.; Mielke, B. *Bioorg. Med. Chem. Lett.* **1997**, 681.

68. Jordan, S.; Schwemler, C.; Kosch, W.; Kretschmer, A.; Stropp, U.; Schwenner, E.; Mielke, B. *Bioorg. Med. Chem. Lett.* **1997**, 687.

69. (a) Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans I* **1997**, 539. (b) Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans I* **1997**, 547. (c) Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans I* **1997**, 555.

70. Hickman, D. T.; King, P. M.; Cooper, M. A.; Slater, J. M.; Micklefield, J. *Chem. Commun.* **2000**, 2251.

71. Püschl, A.; Tedeschi, T.; Nielsen, P. E. *Org. Lett.* **2000**, *2*, 4161.

72. Kim, S. H.; Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

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

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

75. Petersen, K. H.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 793.

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

77. Smith, J. O.; Olson, D. A.; Armitage, B. A. *J. Am. Chem. Soc.* **1999**, *121*, 2686.

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

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

80. Hyrup, B.; Egholm, M.; Nielsen, P. E.; Wittung, P.; Nordén, B.; Buchardt, O. *J. Am. Chem. Soc.* **1994**, *116*, 7964.

81. Krotz, A. H.; Buchardt, O.; Nielsen, P. E. *Tetrahedron Lett.* **1995**, *36*, 6937.

82. Krotz, A. H.; Buchardt, O.; Nielsen, P. E. *Tetrahedron Lett.* **1995**, *36*, 6941.

83. Krotz, A. H.; Larsen, S.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem.* **1998**, *6*, 1983.

84. Almarison, O.; Bruice, T. C.; *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9542.

85. Lagriffoul, P. H.; Egholm, M.; Nielsen, P. E.; Berg, R. H.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1081.

86. De Koning, H.; Pandit, U. K. *Recl. Trav. Chim. Pays-Bas* **1971**, *90*, 1069.

87. Howarth, N. M.; Wakelin, L. P. G. *J. Org. Chem.* **1997**, *62*, 5441.

88. Lenzi, A.; Reginato, G.; Taddei, M.; Trifilieff, E. *Tetrahedron Lett.* **1995**, *36*, 1717.

89. Shah, V. J.; Cerpa, R.; Kunta, I. D.; Kenyon, G. L. *Bioorg. Chem.* **1996**, *24*, 201.

90. Bergmeier, S. C.; Fundy, S. L. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 3135.

91. Savithri, D.; Leumann, C.; Scheffold, R. *Helv. Chim. Acta* **1996**, *79*, 288.

92. Haaima, G.; Hansen, H. F.; Christensen, L.; Dahl, O.; Nielsen, P. E. *Nucleic Acids Res.* **1997**, *25*, 4639.

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

94. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *J. Chem. Soc. Chem. Commun.* **1997**, 1913.

95. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Biochem. Biophys. Res. Commun.* **1997**, *240*, 778.

96. Eldrup, A. B.; Dahl, O.; Nielsen, P. E. *J. Am. Chem. Soc.* **1997**, *119*, 11116.

97. Timar, Z.; Bottka, S.; Kovacs, L.; Penke, B. *Nucleosides Nucleotides* **1999**, *18*, 1131.

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

99. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 2796.

100. Uhlmann, E. *Biol. Chem.* **1998**, *379*, 1045.

101. van der Laan, A. C.; Meeuwenoord, N. J.; Kuhl-Yeheskiely, E.; Oosting, R. S.; Brands, R.; van Boom, J. H. *Recl. Trav. Chim. Pays-Bas* **1995**, *114*, 295.

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

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

104. Finn, P. J.; Gibson, N. J.; Fallon, R.; Hamilton, A.; Brown, T. *Nucleic Acids Res.* **1996**, *24*, 3357.

105. van der Laan, A. C.; Havenaar, P.; Oosting, R. S.; Kuyl-Yeheskiely, E.; Uhlmann, E.; van Boom, J. H. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 663.

106. Petersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O.; *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1119.

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

108. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543.

109. Verheijen, J. C.; van Roon, A-M. M.; van der Laan, A. C.; van der Marel, G. A.; van Boom, J. H. *Nucleosides Nucleotides* **1999**, *18*, 493.

110. Harrison, J. G.; Freier, C.; Laurant, R.; Dennis, R.; Raney, K. D.; Balasubramanian, S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1273.

111. Brandén, L. J.; Mohamed, A. J.; Smith, C. I. E. *Nature Biotech.* **1999**, *17*, 784.

112. Pooga, M.; Soomets, U.; Hallbrink, M.; Valkna, A.; Saar, K.; Rezaki, K.; Kahl, U.; Hao, JX.; Xu, XJ.; Wisenfeld-Hallin, Z.; Hokfelt, T.; Bartfai, T.; Langel, U. *Biotech.* **1998**, *16*, 857.

113. Ishihara, T.; Corey, D. R. *J. Am. Chem. Soc.* **1999**, *121*, 2012.

114. Nielsen, P. E.; Knudsen, H. *PCT Int. Appl. WO 9853801 A1 3* **1998**, 60pp.

115. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, *55*, 177.

CHAPTER 2

CHEMICAL SYNTHESIS OF 1-(*N*-Boc-aminoethyl)-4(*S*)-(purinyl/pyrimidinyl)-2(*S*/*R*)-proline: CYCLIC MONOMERS FOR SYNTHESIS OF CONFORMATIONALLY CONSTRAINED POSITIVELY CHARGED PNAs

2.1. INTRODUCTION

As described in Chapter 1, the search for effective antigene/antisense agents has led to the development of novel DNA/RNA analogues in the past decade[1-8] and the most prominent outcome of this search is the Peptide Nucleic Acids (PNA)[9-14]. In this class of compounds, the entire negatively charged sugar-phosphate backbone of DNA is replaced by a neutral and achiral polyamide backbone consisting of N-(2-aminoethyl)glycine units. The nucleobases are attached to the backbone through a conformationally rigid tertiary acetamide linker (Figure 1). The internucleobase distances in PNA are conserved, allowing its binding to the target DNA/RNA sequences with high sequence specificity and affinity. Moreover, PNA is stable to cellular enzymes like nucleases and proteases.[15] However, major limitations of the therapeutic application of PNAs are their poor solubility in aqueous media due to self-aggregation, insufficient cellular uptake and ambiguity in orientational selectivity of binding.[16,17]

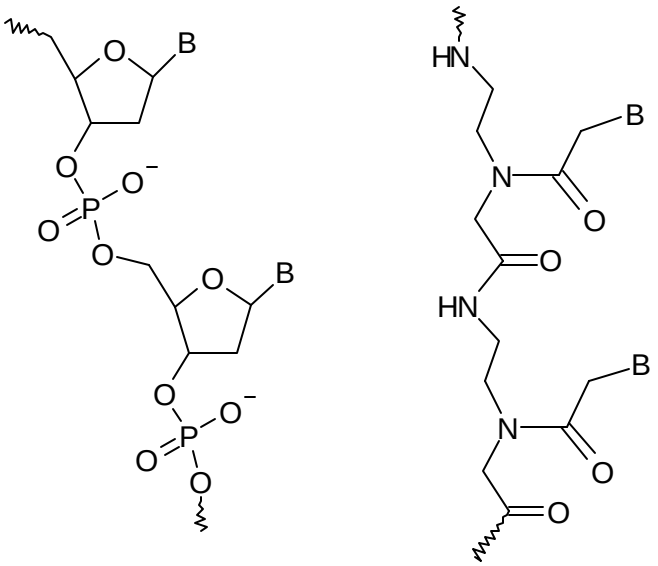


Figure 1. The basic structures of DNA and PNA

In the recent past, there have been many attempts to address these problems in a effective manner. Introduction of chirality into PNA by linking chiral amino acids,[18] peptides,[19] as well as oligonucleotides[20,21] in the PNA backbone was aimed at improving the orientational selectivity of binding. Chirality was also imparted to PNA by using chiral amino acids in the backbone itself.[22,23] Most of these efforts had only marginal desired effects on the hybridization properties of PNA. A five-membered pyrrolidine ring introduced in PNA to impart the necessary structural rigidity and chirality for orientational selectivity (Figure 2a) was found to be ineffective in improving the PNA:DNA binding in the desired manner.[24] However, alternating units with glycine enhanced the binding affinity for DNA[25-28]. Introduction of positive charges in the PNA backbone[29,30] was found to improve its aqueous solubility. Considerable interest is emerging in making positively charged PNAs as they are expected to possess superior ability to strand invade complementary DNA sequences.[31]

Hyrup *et al*.[29] have introduced a positive charge in the PNA backbone at the expense of the conformational rigidity of the tertiary amide linkage. The nucleobase is attached to the polyamide backbone *via* a flexible ethylene linker instead of an acetamide linker (Figure 2b). Although this improved the aqueous solubility of PNA, the detrimental effect on stability of the PNA:DNA complexes stressed the importance of rigid preorganization of the PNA structure for effective binding to ss/ds DNA.

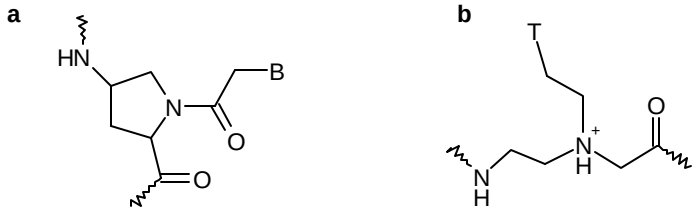


Figure 2. **a**. 4-Aminoprolyl PNA and **b**. Eth-T PNA

Very recently, Hickman *et al.*[32] reported the synthesis and binding properties of a T_5-pentamer pyrrolidine-amide oligonucleotide mimic (Figure 3). The nucleobase is attached to the C4 position of the pyrrolidine ring and the ring nitrogen is alkylated to yield a tertiary amine that is positively charged. This report was followed by that of Püschl *et al.*[33], who prepared the (2*R*,4*S*) adenine analogue. The homooligomer was found to form a DNA_2:PNA triplex in this case. These oligomers formed stable triplexes with DNA and RNA and exhibited binding selectivity for RNA over DNA. Though the pyrrolidine ring nitrogen is positively charged, increasing the salt concentration resulted in only slightly higher Tm values. This is in contrast to other cationic modified oligonucleotides that show a marked decrease in the stability of their complexes with RNA and DNA at higher salt concentration. This was attributed to a reduction in the electrostatic attraction between the oppositely charged backbones owing to the charge compensation by the salt.

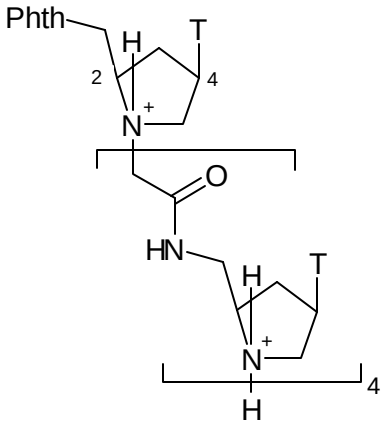


Figure 3. Pyrrolidine-amide Oligonucleotide Mimic

2.1.1. Present Work: Rationale

The present work is directed towards a modification of the PNA backbone, that has no configurational or conformational constraints.[34] It is intended to modulate its flexibility towards striking a balance between the flexibility and rigidity in the backbone

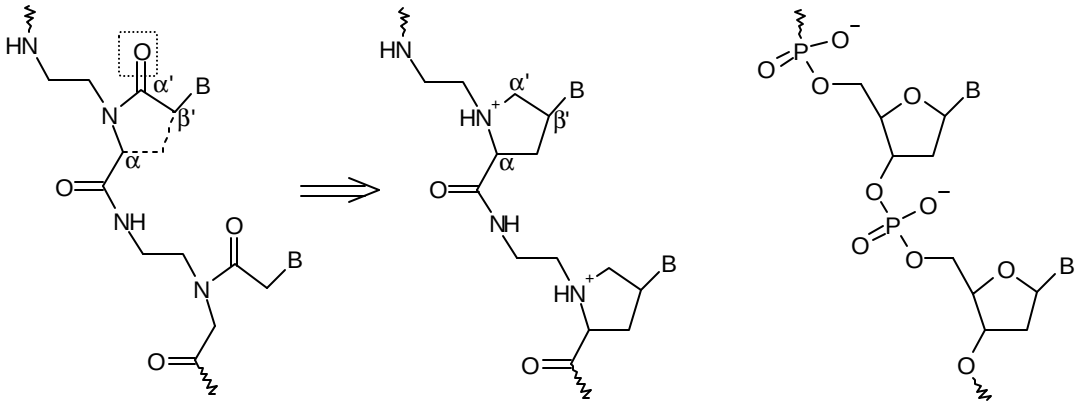


Figure 4. *Aep* PNA

and attain optimum binding with complementary nucleic acid sequences. The modification is conceptually derived from the aminoethylglycyl (*aeg*) PNA structure by bridging the α-carbon atom of the glycine moiety and the β-carbon atom of the linker to the nucleobase by a methylene group, with a simultaneous omission of the α-carbonyl group (Figure 4). The structure so derived is homomorphous with that of DNA as evident from the figure. This modification can thus be considered an analogue of both PNA and DNA. In addition to introducing two chiral centres per unit, it also confers a positive charge at physiological pH by virtue of the tertiary nitrogen atom of the pyrrolidine ring. An added advantage of this monomer would be its improved solubility in aqueous media compared to the unmodified aminoethylglycyl PNA. Through this modification, chirality, controlled flexibility and positive charge are simultaneously conferred on the PNA backbone to influence its physico-chemical and biological properties. The monomers are synthesized starting from the easily available, naturally occurring 4(*R*)-hydroxy-2(*S*)-proline.

The objectives of this chapter are:

(i) Synthesis of 1-(*N*-Boc-aminoethyl)-4(*S*)-(A/ T/ G/ C)-2(*S*)-proline (i.e., L-*cis*) and 1-(*N*-Boc-aminoethyl)-4(*S*)-(A/ T/ G/ C)-2(*R*)-proline (i.e., D-*trans*) diastereoisomers for use in PNA synthesis (Figure 5).

(ii) Synthesis of aminoethylglycyl PNA monomers (A/ T/ G/ C).

(iii) Solid phase synthesis of oligomers incorporating aminoethylglycyl and/or aminoethylprolyl PNA monomers.

(iv) Cleavage from the solid support and purification of the oligomers along with their characterization.

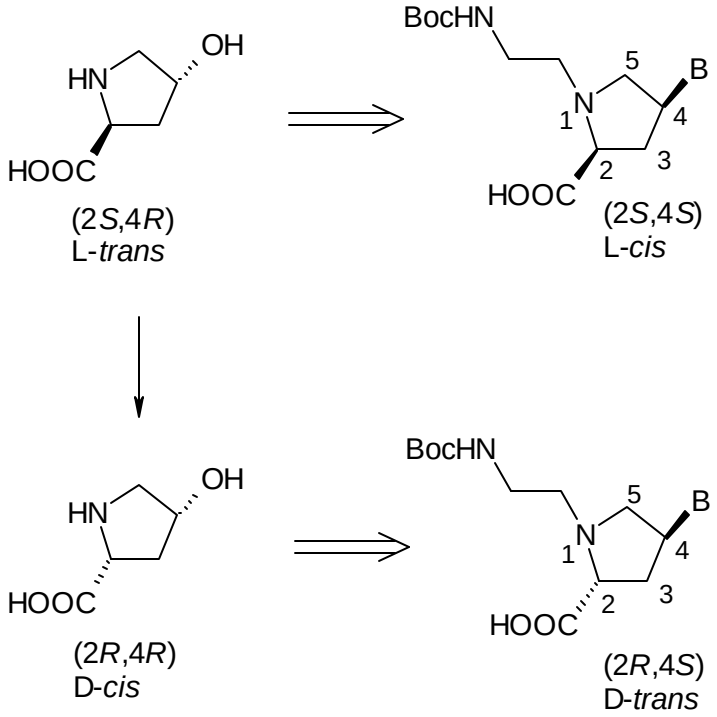


Figure 5

2.2. RESULTS AND DISCUSSION

The aminoethylprolyl thymine PNA monomers were synthesized from the 1-(*N*-Boc-aminoethyl)-4(*R*)-hydroxy-2(*S*/*R*)-proline methylester by Mitsunobu reaction[9] involving base N1 alkylation at C4. The attachment of the other nucleobases (A/G/C) was carried out *via* a 4(*R*)-*O*-mesyl intermediate as depicted in Figure 6. The nucleobases

(A/T/G/C) possess groups like the imide NH and exocyclic amines that are reactive in Mitsunobu conditions and hence had to be protected. Thymine was protected at N3 by a benzoyl group and the exocyclic amines of cytosine, adenine and guanine were protected as the corresponding benzyloxycarbonyl, benzoyl and isobutyryl derivatives respectively (Scheme 1).

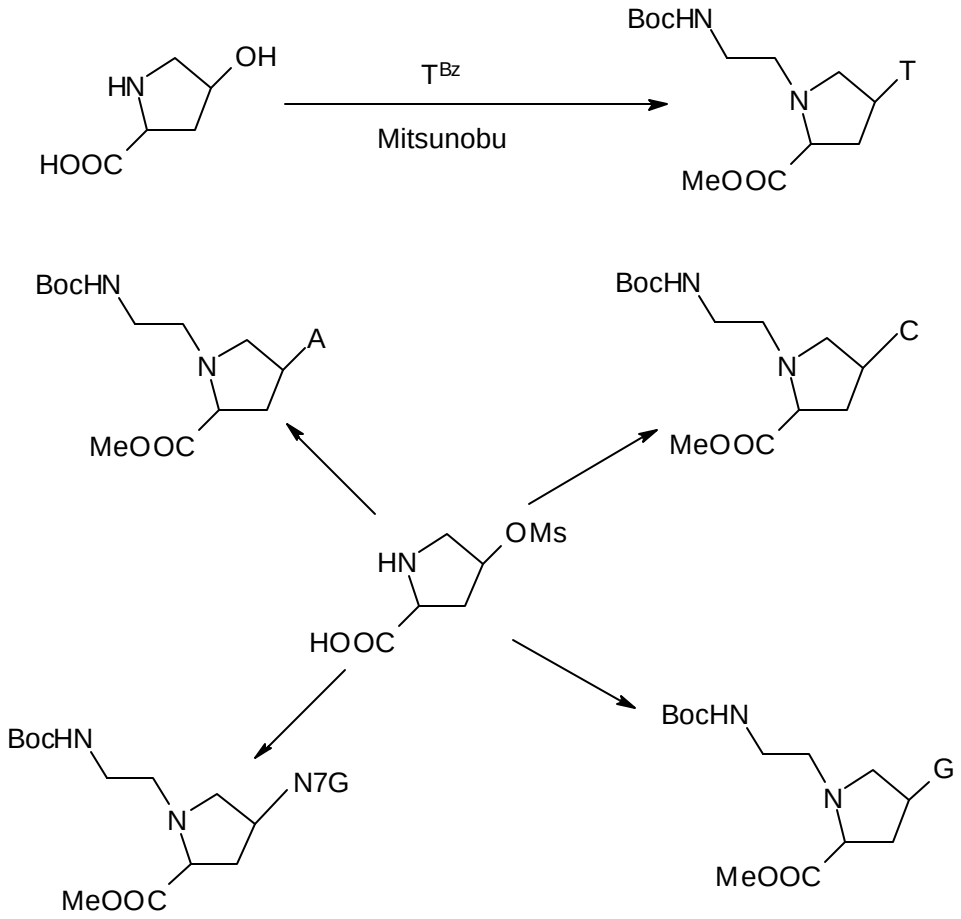


Figure 6. A schematic synthesis of the *aep* PNA-A/G/C monomers

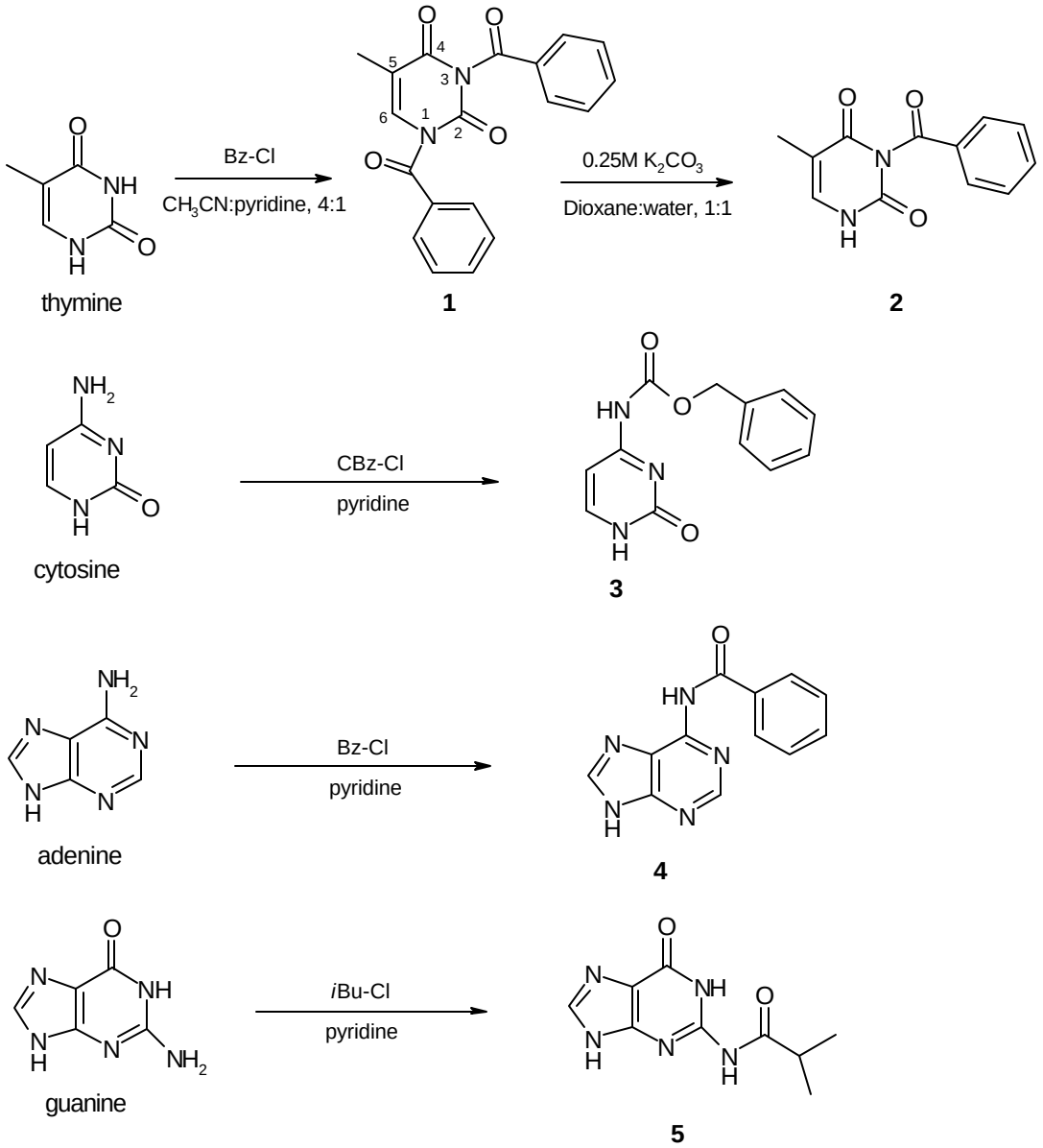
2.2.1. Synthesis of the Protected Nucleobases

2.2.1a N3-benzoylthymine[35] ***(2).*** Thymine was treated with benzoyl chloride in acetonitrile:pyridine (4:1) to obtain the N1, N3-dibenzoyl thymine derivative **1**. The N1-benzoyl group, being more labile than the imide N3-benzoyl group, was preferentially

hydrolyzed using 0.25M K_2CO_3 in dioxane: water (1:1), to yield the N3-benzoyl thymine **2** (Scheme 1).

2.2.1b ***N^4-benzyloxycarbonylcytosine***[36] ***(3)***. Cytosine was treated with benzyloxycarbonyl chloride in dry pyridine to get the desired product, N^4-benzyloxycarbonylcytosine (**3**) (Scheme 1).

Scheme 1. Protection of the exocyclic amino groups of the nucleobases



*2.2.1c N^6-benzoyladenine[37] (**4**).* Adenine, upon treatment with benzoyl chloride in dry pyridine gave the N^6-benzoyl adenine **4** in good yield, which was obtained as white crystals (Scheme1).

*2.2.1d N^2-isobutyrylguanine[38] (**5**).* Guanine was treated with isobutyryl chloride in dry pyridine to yield N^2-isobutyrylguanine (**5**) (Scheme 1).

2.2.2. Synthesis of Aminoethylprolyl PNA Monomers

2.2.2a 1-(N-Boc-aminoethyl)-4(S)-(N3-benzoylthymin-1-yl)-2(S/R)-proline. The synthesis of 1-(*N*-Boc-aminoethyl)-4(*S*)-(N3-benzoylthymin-1-yl)-2(*S*)-proline methyl ester **10** was achieved starting from the versatile, naturally occurring amino acid, 4(*R*)-hydroxy-2(*S*)-proline **6**, which bears three different functional groups that can be utilized as handles for combinatorial derivatization (Figure 7). This was achieved by a simple synthetic route comprising three steps (Scheme 2). 4(*R*)-hydroxy-2(*S*)-proline **6** was converted to its methyl ester hydrochloride by refluxing in methanol containing $SOCl_2$. The resulting 4(*R*)-hydroxy-2(*S*)-proline methyl ester **7** isolated as its hydrochloride was then alkylated at the ring nitrogen by (*N*-Boc)-2-aminoethyl bromide **13** in the presence of K_2CO_3 in dry acetonitrile-DMF to obtain 1-(*N*-Boc-aminoethyl)-4(*R*)-hydroxy-2(*S*)-proline methyl ester **8**. The alkylating agent **13** was prepared from 2-aminoethanol **11**, the amino group of which was first protected as a Boc-derivative

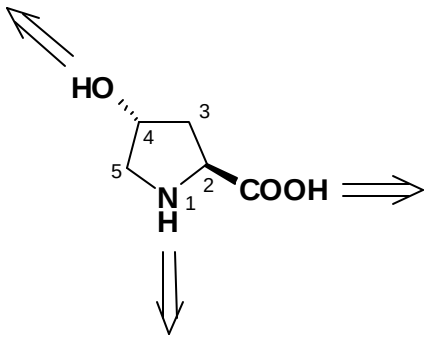
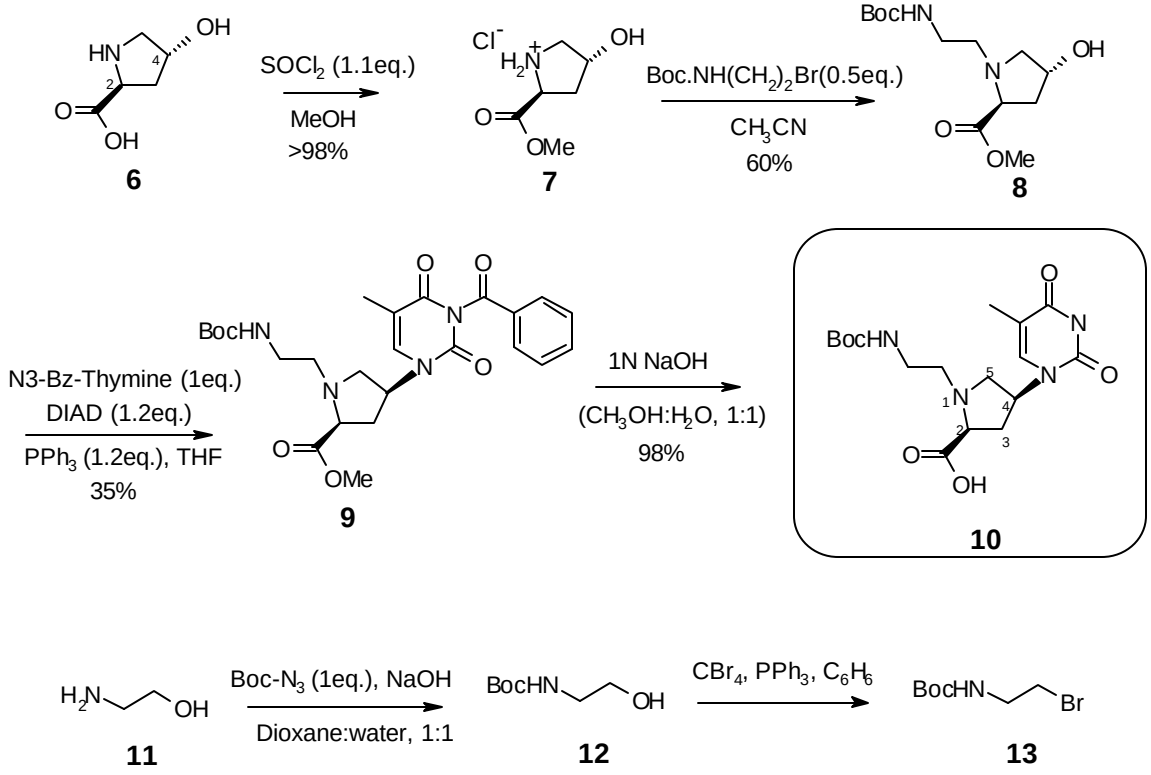


Figure 7. 4(R)-hydroxy-2(S)-proline (**6**)

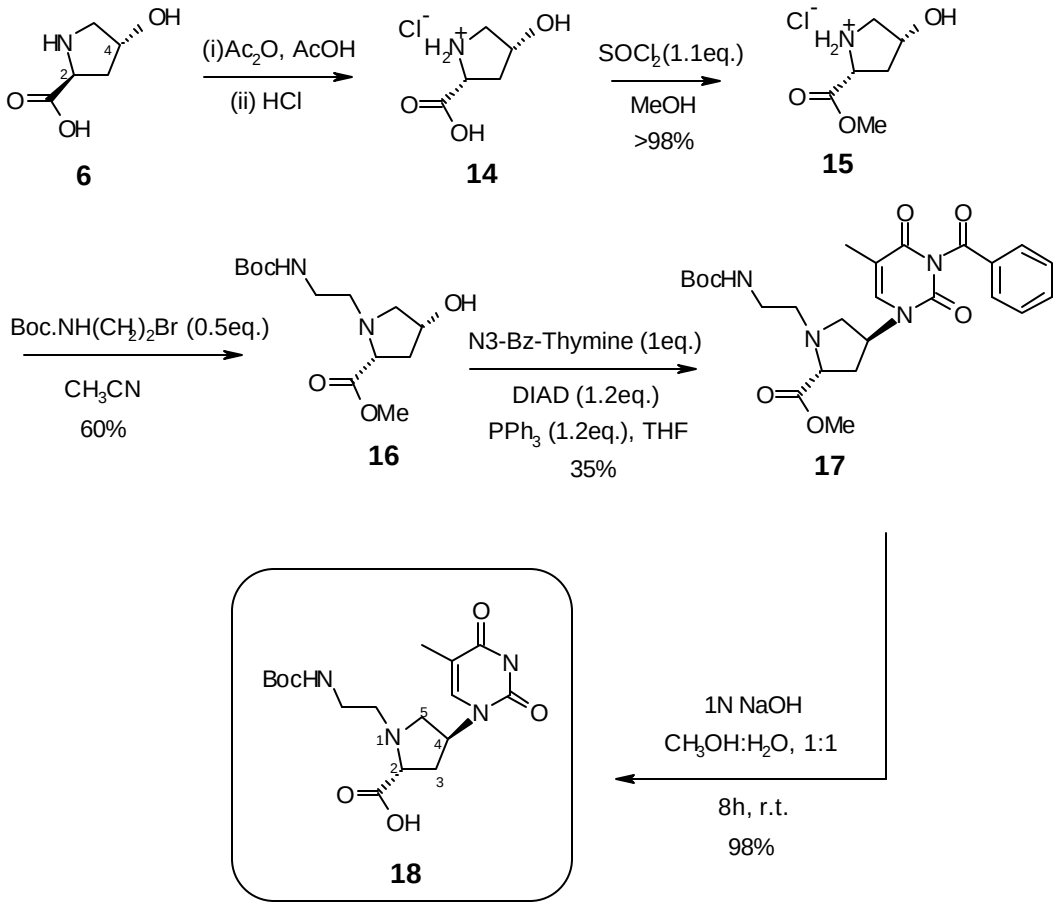
followed by the conversion of the hydroxyl group to the corresponding bromide utilizing carbon tetrabromide and triphenylphosphine in benzene. The 1-(*N*-Boc-aminoethyl)-4(*R*)-hydroxy-2(*S*)-proline methyl ester **8** was then subjected to Mitsunobu reaction conditions involving N3-benzoylthymine as the incoming nucleophile,[25] to yield 1-(*N*-Boc-aminoethyl)-4(*S*)-(N3-benzoylthymin-1-yl)-2(*S*)-proline methyl ester **9**. This reaction proceeds with inversion of stereochemistry at the C4 to yield the 4(*S*) product. A downfield shift in the H4-NMR signal from δ4.40 to δ5.25 and the appearance of a characteristic peak at δ8.10 of the H6 proton of thymine confirmed the identity of the Mitsunobu product. Thus, this scheme involving a single inversion step results in the overall transformation of 4(*R*)-hydroxy-2(*S*)-proline to 1-(*N*-Boc-aminoethyl)-4(*S*)-(N3-benzoylthymin-1-yl)-2(*S*)-proline methyl ester **9**. The ester function was hydrolyzed to get the (2*S*,4*S*) *aep*PNA-T monomer **10**.

Scheme 2. Synthesis of the (2*S*,4*S*) *aep*PNA-T monomer



The synthesis of the (2*R*,4*S*) diastereomer, 1-(*N*-Boc-aminoethyl)-4(*S*)-(N3-benzoylthymin-1-yl)-2(*R*)-proline methyl ester **18** was carried out from the same starting material, *viz*., 4(*R*)-hydroxy-2(*S*)-proline **6** and involved an initial epimerization step at C2. This was achieved by treating the 4(*R*)-hydroxy-2(*S*)-proline with acetic acid and acetic anhydride followed by the hydrolysis of the resulting lactone with 2N HCl to give 4(*R*)-hydroxy-2(*R*)-proline hydrochloride (**14**) in good yield.[39] Following a similar set of reactions as for the 2(*S*) isomer, *viz*., esterification of the carboxylic acid, alkylation of the proline ring nitrogen with 2-(*N*-Boc)-aminoethyl bromide and Mitsunobu replacement of the 4-hydroxyl group by N3-benzoylthymine involving inversion of the C4-stereocenter, yielded the 1-(*N*-Boc-aminoethyl)-4(*S*)-(N3-benzoylthymin-1-yl)-2(*R*)-

Scheme 3. Synthesis of the (2*R*,4*S*) *aep*PNA-T monomer

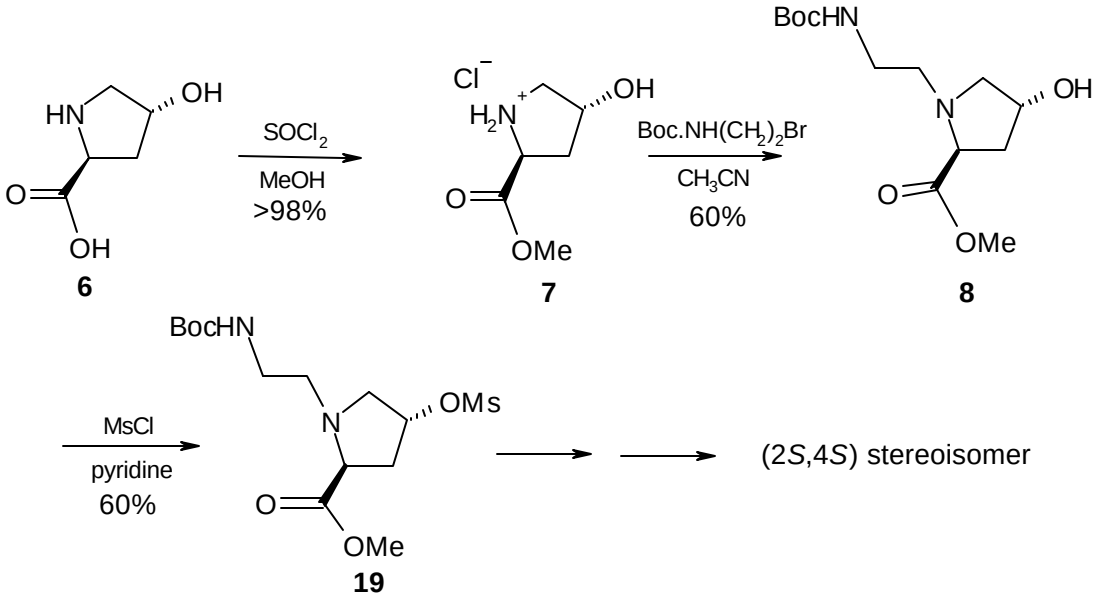


proline methyl ester **17** (Scheme 3). Upon hydrolysis of the ester, the (2*R*,4*S*) *aep*PNA-T monomer **18** is obtained. The synthesis of this monomer comprises inversion of stereochemistry at both the stereocenters, C2 and C4, to give the (2*R*,4*S*) isomer.

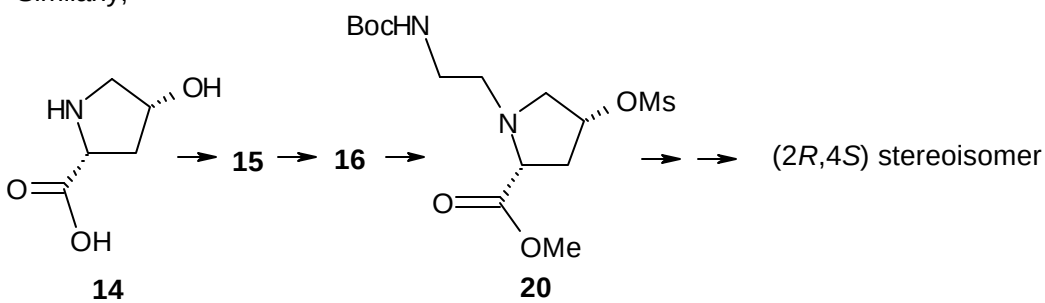
2.2.3. Synthesis of *Aep*PNA-(A/G/C) Monomers

The procedure utilized for the *aep*PNA-T monomer involving the attachment of the nucleobase at C4 of the pyrrolidine ring *via* Mitsunobu reaction did not work for the other nucleobases, *viz.* A, G and C. Lowe *et al*.[25] have reported the attachment of these nucleobases *via* the 4-*O*-tosyl derivative. The synthesis of the *aep* monomeric units

Scheme 4. Synthesis of the common intermediate for the *aep*PNA-A/G/C monomers



Similarly,



bearing these nucleobases (A/G/C) was carried out *via* a common intermediate, the 1-(*N*-Boc-aminoethyl)-4(*R*)-*O*-mesyl-2(*S*)-proline methyl ester for the 2(*S*) stereoisomers and the 1-(*N*-Boc-aminoethyl)-4(*R*)-O-mesyl-2(*R*)-proline methyl ester for the 2(*R*) stereoisomers (Scheme 4).The 1-(*N*-Boc-aminoethyl)-4(*R*)-*O*-mesyl-2(*S*)-proline methyl ester **19** was prepared by treating the 1-(*N*-Boc-aminoethyl)-4(*R*)-hydroxy-2(*S*)-proline methyl ester **8** with mesyl chloride (methane sulphonyl chloride) in pyridine (Scheme 4). The appearance of a characteristic signal at δ3.0 and a downfield of the H4 signal from δ4.4 to δ5.2 shift in the ^{1}H NMR indicated the formation of the mesyl derivative. The mesyl derivative **19** was the precursor to the (2*S*, 4*S*) *aep*PNA-(A/G/C) monomers.

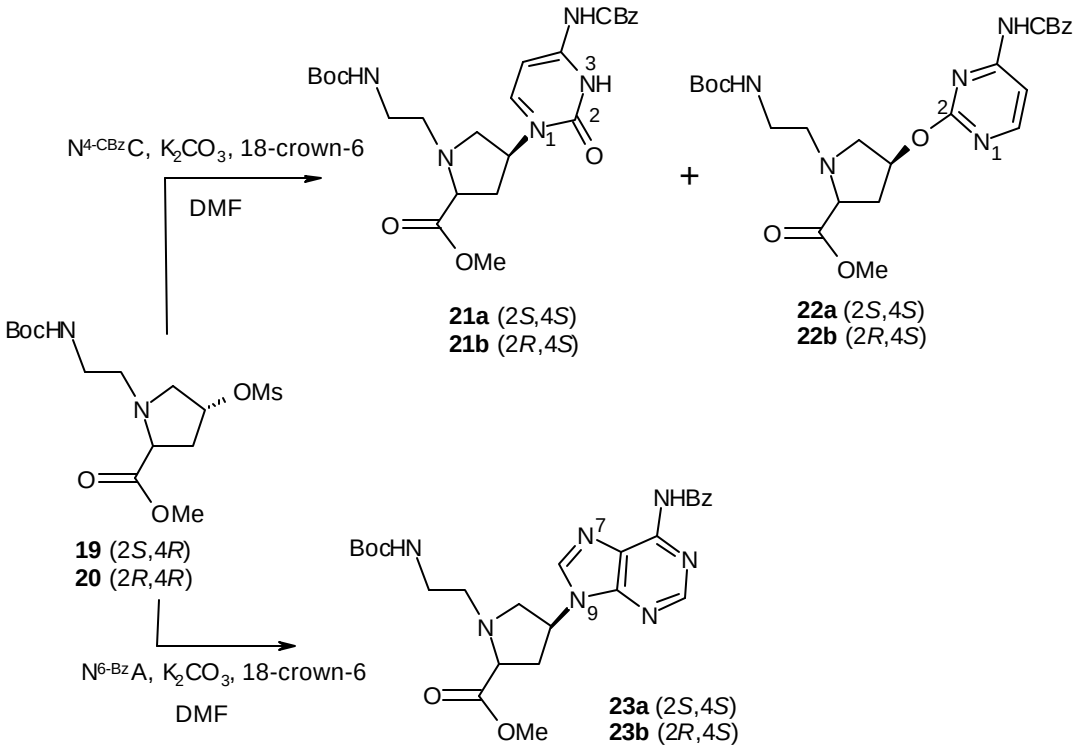
Similarly, the 1-(*N*-Boc-aminoethyl)-4(*R*)-O-mesyl-2(*R*)-proline methyl ester **20** was derived from the 1-(*N*-Boc-aminoethyl)-4(*R*)-hydroxy-2(*R*)-proline methyl ester **14** as described for the (2*S*,4*R*) diastereoisomer. This mesyl derivative served as the precursor to the (2*R*,4*S*) *aep*PNA-(A/G/C) monomers.

*2.2.3a 1-(N-Boc-aminoethyl)-4(S)-(N^{4}-benzyloxycarbonyl cytosin-1-yl)-2(S/R)-proline methyl ester (**21**).* The synthesis pathway for 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^{4}-benzyloxycarbonyl cytosin-1-yl)-2(*S*)-proline methyl ester (**21a**) consisted of four steps. These steps start from esterification of the carboxylic acid moiety followed by alkylation of the proline ring nitrogen with 2-(*N*-Boc)-aminoethyl bromide (**13**). The next step involved conversion of the C4-hydroxyl group to its corresponding mesyl derivative (**19**) by treatment of **8** with mesyl chloride in pyridine (Scheme 4). 1-(*N*-Boc-aminoethyl)-4(*R*)-O-mesyl-2(*S*)-proline methyl ester (**19**) was stirred with N^{4}-benzyloxycarbonyl cytosine, K_2CO_3 and a catalytic amount of 18-crown-6 in DMF to give the 4(*S*)-(N^{4}-benzyloxycarbonyl cytosin-1-yl) derivative (**21a**) (Scheme 5). This step is accompanied by an inversion of stereochemistry at the C4-stereocenter owing to the S_N2 displacement reaction. The 4(*S*)-(N^{4}-benzyloxycarbonyl-cytosin-2-yloxy) derivative (**22a**) was also obtained as a side-product of the above reaction in 20% yield. The N1-

and O2- substituted products are distinguishable in ^{1}H NMR by the chemical shifts of the H5 and H6 protons of cytosine, viz. δ7.25 & δ7.60 for H5 and δ8.45 & δ8.35 for H6 in the N1- and O2-substituted derivatives respectively. The O2-substituted product is also characterized by a downfield shift of the H4 signal from δ5.20 in the N1-substituted derivative to δ5.35. Thus, as in the synthesis of the (2*S*,4*S*) isomer bearing N3-benzoylthymine as the nucleobase, this scheme for the synthesis of the 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonyl cytosin-1-yl)-2(*S*)-proline methyl ester involves the inversion of stereochemistry at one stereocenter, i. e., C4.

1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonyl cytosin-1-yl)-2(*R*)-proline methyl ester (**21b**) was prepared from 4(*R*)-hydroxy-2(*S*)-proline (**6**), and this synthesis involved two inversion steps, one at each stereocenter (C2 & C4). Prior to the N-alkylation step using 2-(*N*-Boc)-aminoethyl bromide (**13**), the C2-stereocenter was

Scheme 5. Synthesis of the protected *aep*PNA-C & A monomers



epimerized by treatment with acetic anhydride and acetic acid, followed by hydrolysis of the lactone by 2N HCl. The C4-hydroxyl group was converted to its mesyl derivative (**20**). The C4-stereocenter was then inverted by S_N2 displacement of the 4-*O*-mesyl group by N^4-benzyloxycarbonyl cytosine. As in the case of the (2*S*,4*S*) diastereomer, this step yielded the cytosine O2-substituted derivative **22b** as a side-product in 20% yield. The downfield position of the H4 proton resonance in the 1H NMR spectrum of the O2-substituted product (δ5.30) compared to that in the N1-substituted product (δ4.90). The ^{1}H chemical shifts of cytosine H6 (δ8.35 and δ7.90 respectively) and H5 protons (δ7.60 and δ7.25 respectively) confirmed the identity of O2- and N1-substituted products. Thus, two inversions from (2*S*, 4*R*) hydroxy proline gave the protected (2*R*,4*S*) cytosine monomer (**21b**).

*2.2.3b 1-(N-Boc-aminoethyl)-4(S)-(N^6-benzoyladenin-9-yl)-2(S/R)-proline methyl ester (**23**).* 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^6-benzoyladenin-9-yl)-2(*S*)-proline methyl ester (**23a**) was synthesized from 4(*R*)-hydroxy-2(*S*)-proline (**6**) in four steps (Scheme 5): (I) esterification of the carboxylic acid function, (ii) alkylation of the pyrrolidine ring nitrogen with 2-(*N*-Boc)-aminoethyl bromide (**13**), (iii) conversion of the C4-hydroxyl group to the corresponding 4O-mesyl derivative (**19**) and (iv) reaction of 1-(*N*-Boc-aminoethyl)-4(*R*)-O-mesyl-2(*S*)-proline methyl ester with N^6-benzoyladenine in the presence of anhydrous K_2CO_3/18-crown-6. The last step was accompanied by an inversion of stereochemistry as a consequence of the S_N2 displacement at C4. In ^{1}H NMR, the H4 proton exhibited a downfield shift from δ5.20 to δ5.40 as a consequence of alkylation. No adenine N7-substituted product was obtained in this reaction.

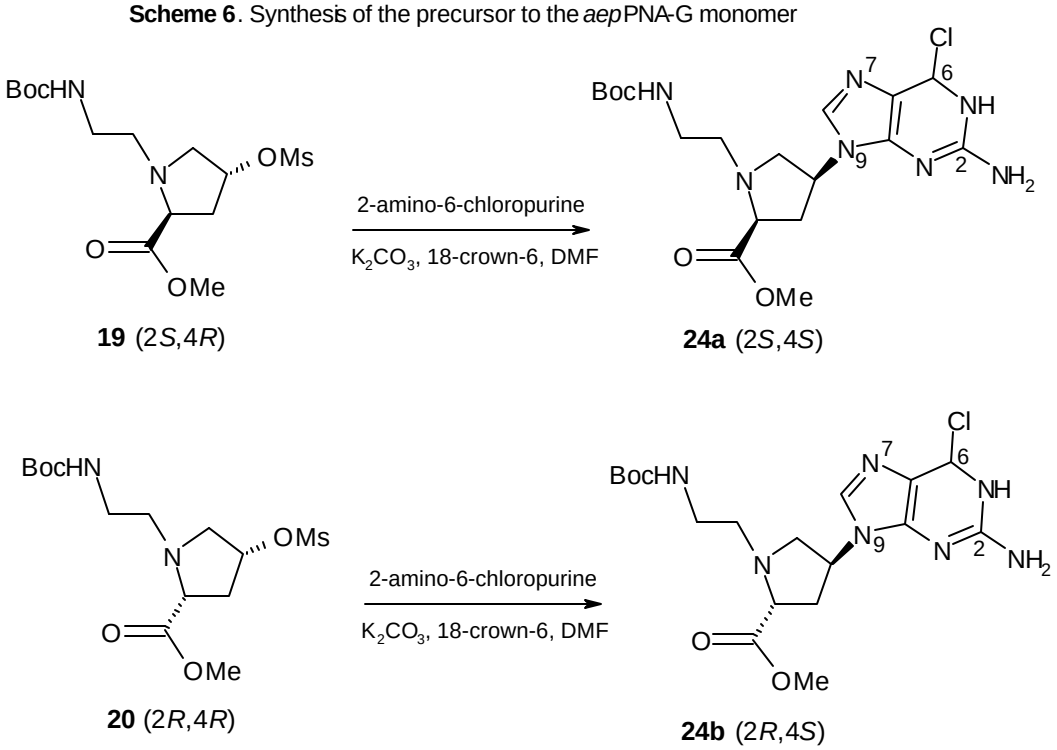
The synthesis of the diastereomer **23b** was achieved as follows: 4(*R*)-hydroxy-2(*S*)-proline (**6**) was epimerized at the C2 stereocenter by treating with acetic anhydride and acetic acid, followed by hydrolysis with 2N HCl. The carboxyl function was esterified, and subsequent N-alkylation yielded the 1-(*N*-Boc-aminoethyl)-4(*R*)-hydroxy-2(*R*)-

proline methyl ester (**16**). The 4(*R*)-O-mesyl derivative (**20**) obtained upon reaction with mesyl chloride in pyridine was stirred with N^6-benzoyladenine, anhydrous K_2CO_3 and 18-crown-6 in DMF, when the product 1(*N*-Boc-aminoethyl)-4(*S*)-(N^6-benzoyladenin-9-yl)-2(*R*)-proline methyl ester (**23b**) was obtained in good yield. The structure of the product was confirmed by the appearance of the characteristic benzoyl proton resonance peaks in the ^{1}H NMR spectrum, *viz*. a doublet at δ8.00 and two triplets at δ7.56 and δ7.45 corresponding to the *o*-, *p*- and *m*-CH respectively, and singlets at δ8.76 and δ8.10 of the H2 and H8 protons of adenine. S_N2 inversion at the C4 stereocenter that occurs during this step leads to the 2*R* derivative, 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^6-benzoyladenin-9-yl)-2(*R*)-proline methyl ester is obtained after two inversion steps.

*2.2.3c 1-(N-Boc-aminoethyl)-4(S)-(2-amino-6-chloropurin-9-yl)-2(S/R)-proline methyl ester (**24**).* 1-(*N*-Boc-aminoethyl)-4(*S*)-(2-amino-6-chloropurin-9-yl)-2(*S*)-proline methyl ester (**24a**) was synthesized from 4(*R*)-hydroxy-2(*S*)-proline (**6**) as a precursor to the guanine monomer. The 1-(*N*-Boc-aminoethyl)-4(*R*)-O-mesyl-2(*S*)-proline methyl ester (**19**) was obtained after a similar series of reactions as in (Scheme 5) involving (i) esterification of the carboxyl function, (ii) alkylation of the pyrrolidine ring nitrogen and (iii) mesylation of the C4 hydroxyl group using mesyl chloride in dry pyridine (Schemes 2 & 4). Upon stirring with 2-amino-6-chloropurine, anhydrous K_2CO_3 and 18-crown-6 in dry DMF, the mesyl derivative **19** was converted to 1-(*N*-Boc-aminoethyl-4(*S*)-(2-amino-6-chloropurin-9-yl-2(*S*)-proline methyl ester (**24a**), with S_N2 inversion of the C4 stereocenter (Scheme 6). This was evident from the appearance of a characteristic signal of the H8 of 2-amino-6-chloropurine in the ^{1}H NMR spectrum at δ8.35. Thus, a single inversion at C4 afforded the (2*S*, 4*S*) product from the starting (2*S*, 4*R*) compound.

The synthesis of the diastereomeric 1-(*N*-Boc-aminoethyl)-4(*S*)-(2-amino-6-chloropurin-9-yl)-2(*R*)-proline methyl ester (**24b**) was achieved starting from 4(*R*)-hydroxy-2(*S*)-proline (**6**) in five steps and included two inversions, one at C4 and the other, at C2. The 1-(*N*-Boc-aminoethyl)-4(*R*)-*O*-mesyl-2(*R*)-proline methyl ester (**20**) was prepared as described earlier. This, on stirring with 2-amino-6-chloropurine, anhydrous K_2CO_3 and 18-crown-6 in dry DMF, afforded 1-(*N*-Boc-aminoethyl)-4(*S*)-(2-amino-6-chloropurin-9-yl)-2(*R*)-proline methyl ester **24b** (Scheme 6). The H8 resonance of 2-amino-6-chloropurine appeared as a singlet at δ7.85 in the ^{1}H NMR spectrum as against δ8.35 in the (2*S*,4*S*) diastereomer. Thus, two inversions, one at C2 by epimerization using acetic acid and acetic anhydride followed by hydrolysis with HCl, and one at C4 by S_N2 displacement reaction, yielded the (2*R*,4*S*) product from the starting (2*S*,4*R*) compound. This product is the precursor to the guanine monomer, which can be obtained by alkaline hydrolysis.

Scheme 6. Synthesis of the precursor to the *aep* PNA-G monomer

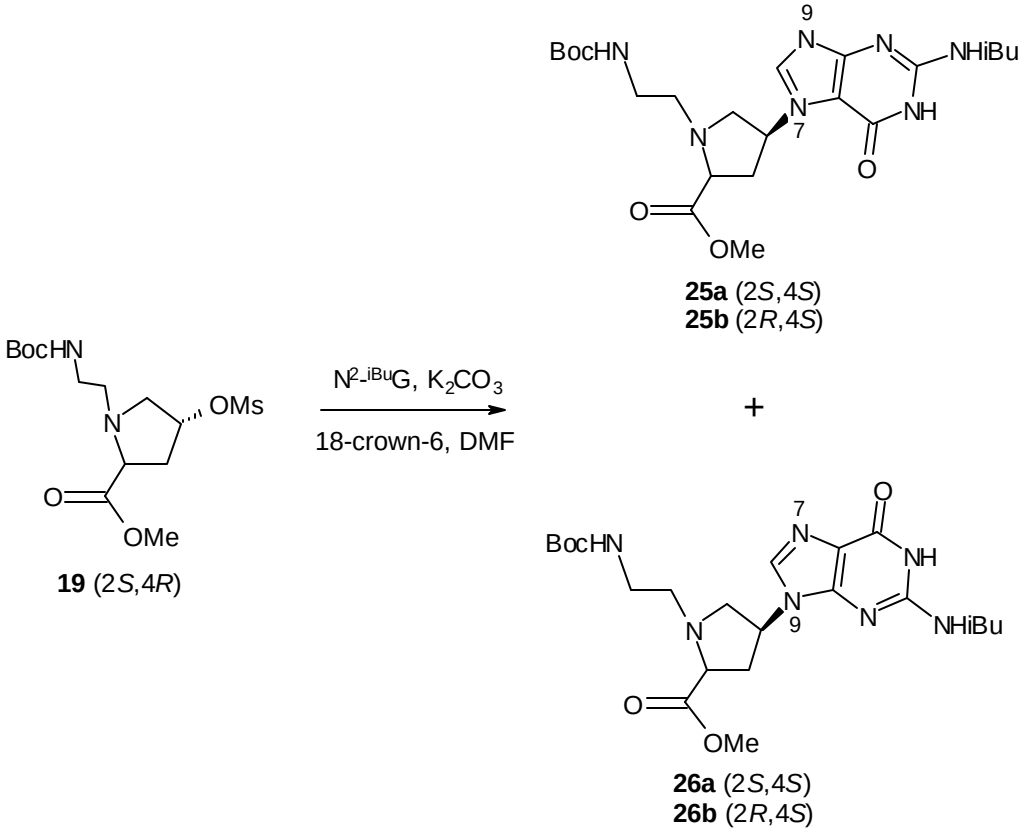


19 (2*S*,4*R*) **24a** (2*S*,4*S*)

20 (2*R*,4*R*) **24b** (2*R*,4*S*)

*2.2.3d 1-(N-Boc-aminoethyl)-4(S)-(N^2-isobutyrylguanin-7-yl)-2(S)-proline methyl ester (**25a**).* The guanine N7-substituted (2*S*,4*S*) derivative **25a** was prepared from 4(*R*)-hydroxy-2(*S*)-proline (**6**) by a similar set of reactions as in Scheme 7, involving a single inversion step. These comprised esterification of the carboxylic acid function, alkylation of the pyrrolidine ring nitrogen, followed by the treatment with mesyl chloride in pyridine to get the 1-(*N*-Boc-aminoethyl)-4(*R*)-*O*-mesyl-2(*S*)-proline methyl ester (Schemes 2 & 4). This mesylate (**19**) was stirred with N^2-isobutyrylguanine and anhydrous K_2CO_3 in dry DMF, when the 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyrylguanin-7-yl)-2(*S*)-proline methyl ester was obtained as the major product (Scheme 7). The reaction is an S_N2 displacement, which yields the product with inverted stereochemistry at the C4 stereocenter. The more polar 1-(*N*-Boc-aminoethyl)-

Scheme 7. Synthesis of the protected *aep* PNA-G monomers

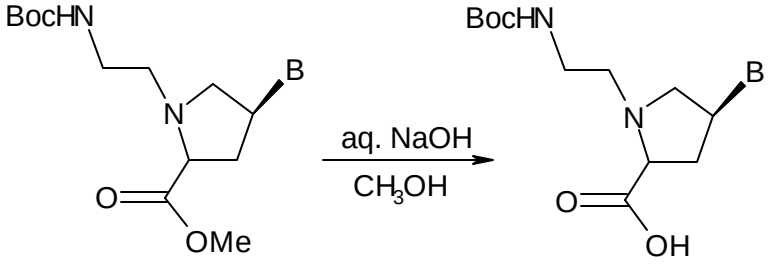


4(*S*)-(N^2-isobutyrylguanin-9-yl)-2(*S*)-proline methyl ester (**26a**) was also isolated as the minor product. The ^{1}H NMR spectrum of **25a** showed a downfield shift in the resonance of the H4 proton in from δ5.20 to δ5.65. These two products differed in the ^{1}H NMR spectra mainly in their guanine H8 resonances, *viz*. δ8.40 and δ8.10 in the N7- and N9-substituted products respectively. Thus, the (2*S*,4*S*) product (**25a**) was synthesized from the starting (2*S*,4*R*) hydroxy proline **6** in four steps.

2.2.4. Hydrolysis of esters

Solid phase synthesis of *aep*PNA requires N-protected free carboxylic acids. Towards this, the methyl ester groups of the *aep*PNA monomers were subjected to saponification by sodium hydroxide in a water-methanol mixture to yield the corresponding carboxylic acids (Scheme 8).

Scheme 8. Hydrolysis of the methyl esters of the *aep*PNA monomers



9,17 B = T^{Bz}
21 B = C^{CBz}
23 B = A^{Bz}
24 B = 2-amino-6-chloropurine
26 B = G^{iBu}
25 B = N7G^{iBu}

10, 18 B = T
27 B = C^{CBz}
28 B = A^{Bz}
29 B = G
30 B = G^{iBu}
31 B = N7G^{iBu}

*2.2.4a Synthesis of 1-(N-Boc-aminoethyl)-4(S)-(thymin-1-yl)-pyrrolidine-2(S/R)-carboxylic acid (**10**, **18**).* The hydrolysis of the methyl ester and the N3-benzoyl group of thymine in 1-(*N*-Boc-aminoethyl)-4(*S*)-(thymin-1-yl)-2(*S*/*R*)-proline methyl ester (**9** & **17**) were achieved in a single step, upon treatment with 1N NaOH in methanol: water,

(1:1). The ester function was found to be hydrolyzed within 10 min. However, the cleavage of the benzoyl group was complete only after 9h. Neutralization of the excess alkali with Dowex 50 H^+ and work-up gave 1-(*N*-Boc-aminoethyl)-4(*S*)-(thymin-1-yl)-pyrrolidine-2(*S*/*R*)-carboxylic acid (**10/18**) in quantitative yield.

*2.2.4b 1-(N-Boc-aminoethyl)-4(S)-(N^4-benzyloxycarbonylcytosin-1-yl)-pyrrolidine -2(S/R)-carboxylic acid (**27**).* The methyl ester in 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonylcytosin-1-yl)-2(*S*/*R*)-proline methyl ester (**21**) was hydrolyzed by treatment with aqueous methanolic NaOH. The benzyloxycarbonyl protecting group on the cytosine N^4 remains unaffected during this treatment. However, if the reaction is continued for a longer time, partial hydrolysis of the benzyloxycarbonyl group was found to occur. Hence, the excess alkali was immediately neutralized upon hydrolysis of the ester, to obtain 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonylcytosin-1-yl)-pyrrolidine-2(*S*/*R*)-carboxylic acid (**27**) as the appropriate cytosine monomer for use in solid phase peptide synthesis.

*2.2.4c 1-(N-Boc-aminoethyl)-4(S)-(N^6-benzoyladenin-9-yl)-pyrrolidine-2(S/R)-carboxylic acid (**28**).* 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^6-benzoyladenin-9-yl)-2(*S*/*R*)-proline methyl ester (**23**) upon treatment with 1N NaOH in aqueous methanol underwent ester hydrolysis within 10 min to afford the corresponding 1-(*N*-Boc-amionethyl)-4(*S*)-(N^6-benzoyladenin-9-yl)-pyrrolidine-2(*S*/*R*)-carboxylic acid **28** in quantitative yield. The benzoyl group was left unaffected, yielding the appropriately protected monomer for solid phase synthesis.

*2.2.4d 1-(N-Boc-aminoethyl)-4(S)-(guanin-9-yl)-pyrrolidine-2(S/R)-carboxylic acid (**29**, **30**).* The aminoethylprolyl guanine monomer **30** was obtained by two different routes. The first was from 1-(*N*-Boc-aminoethyl)-4(*S*)-(2-amino-6-chloropurin-9-yl)-2(*S*/*R*)-proline methyl ester (**24**). This derivative, upon treatment with NaOH in

aqueous methanol, initially underwent ester hydrolysis within 10 min., followed by conversion of the 6-chloro function to the 6-oxo function. The oxidation reaction occurred completely after 24h for the (2*R*,4*S*) isomer, while the (2*S*,4*S*) isomer required 72h for complete oxidation. The transformation of the nucleobase from 2-amino-6-chloropurine to guanine, i.e., 2-amino-6-oxopurine, was unambiguously confirmed by the UV spectra (Figure 8). The 2-amino-6-chloropurine derivative **24** exhibited a characteristic peak ~310nm, which was absent in the guanine derivative **29**. The guanine derivative **29**, on the other hand, sported a peak at ~280nm, that was absent in the precursor, 2-amino-6-chloropurine derivative **24**. The conversion was also evident from the appearance of a carbonyl resonance at δ167.8 in the ^{13}C NMR spectrum in the guanine monomer, which was absent in the 2-amino-6-chloropurine monomer. The second route was from 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyrylguanin-9-yl)-2(*S*)-proline methyl ester (**26**). This derivative, upon saponification with NaOH in water-methanol for 10min gave the 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyrylguanin-9-yl)-pyrrolidine-2(*S*)-carboxylic acid (**30**), which was used in the solid phase synthesis of oligomers. The N7-guanine derivative, 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyrylguanin-7-yl)-pyrrolidine-2(*S*)-carboxylic acid (**31**) was obtained in a similar manner from 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyryl-guanin-7-yl)-2(*S*)-proline methyl ester (**25a**) upon hydrolysis using aqueous methanolic NaOH.

2.2.5. Synthesis of Aminoethylglycyl PNA Monomers

The sequential addition of *aep*PNA units into *aeg*PNA was carried to enable a systematic study of their effect on the binding properties of PNA. For this purpose, the synthesis of unmodified *aeg*PNA monomers is necessary and was carried out according to the literature procedures.[40]

The synthesis was carried out starting from the easily available 1,2-diaminoethane (**32**) (Scheme 9). This was treated with Boc-azide to give the mono-protected derivative **33**, which was achieved by using a large excess of 1,2-diaminoethane over the Boc-azide in high dilution. The di-Boc derivative was also obtained in very small amounts (< 5%), but being insoluble in water, could be removed by filtration. The *N1*-Boc-1,2-diaminoethane was then subjected to *N*-alkylation using ethylbromoacetate and KF-Celite in dry acetonitrile. The use of KF-Celite[41] was found to be advantageous over K_2CO_3 both, in terms of the yield of the product, as well as the ease of work-up. Upon completion of the reaction, the KF-Celite was filtered off and the product isolated in the crude form after concentration of the filtrate. The aminoethylglycine **33** obtained was further treated with chloroacetyl chloride[42] to yield the corresponding chloro derivative

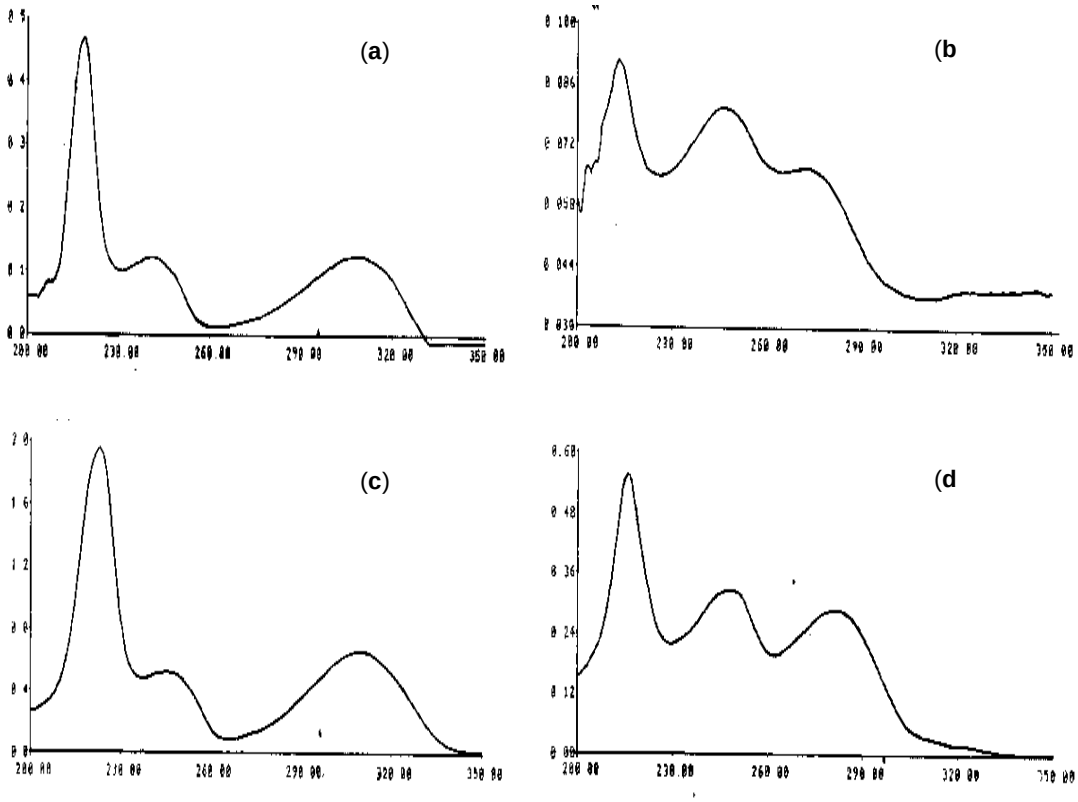
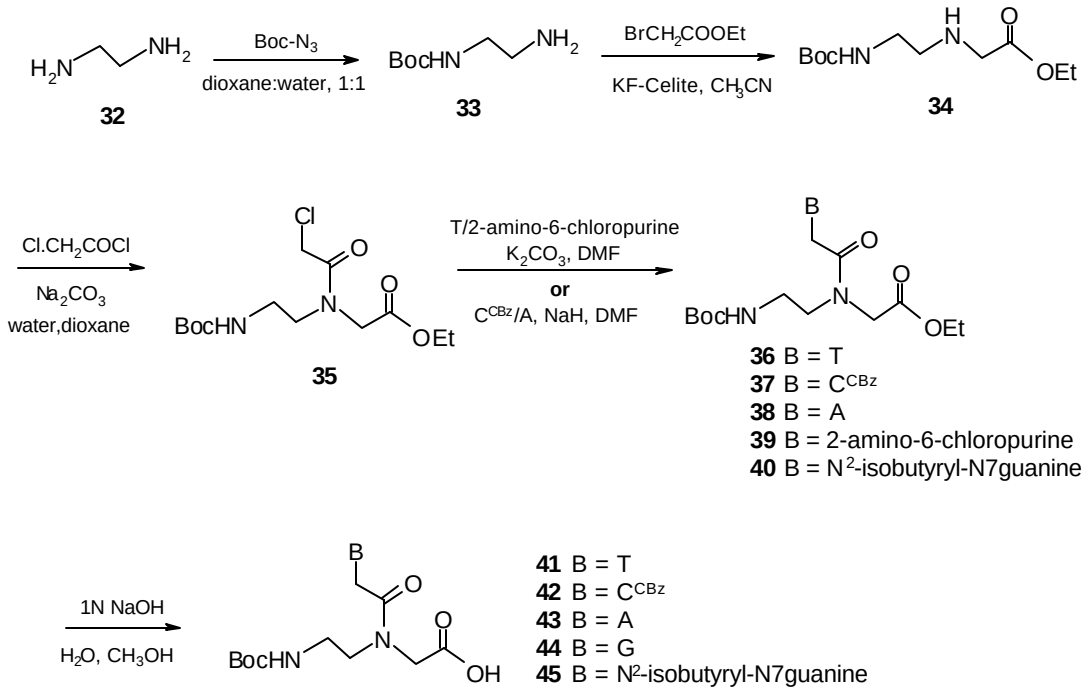


Figure 10. UV scans of the free nucleobases (**a**). 2-amino-6-chloropurine and (**b**). guanine and the *aep* PNA units (**c**). (2*S*,4*S*) compound **24a** and (**d**). (2*S*,4*S*) compound **29**

35 in good yield. The use of triethylamine as the base in this reaction gave poor yields. However, when Na_2CO_3 was used as the base in aqueous dioxane, the desired product was obtained in good yield. The ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate (**35**) was used as a common intermediate in the preparation of all the PNA monomers and can be stored under refrigeration. The alkylation of ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate with thymine and cytosine is regiospecific. Thymine was reacted with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate using K_2CO_3 as a base to obtain the *N*-(Boc-aminoethylglycyl)-thymine ethyl ester **36** in high yield. In the case of cytosine, the N^4-amino group was protected as its benzyloxycarbonyl derivative, and used for alkylation employing NaH as the base to provide the N1-substituted product (**37**). Although adenine is known to undergo both N7- and N9-substitution, N7-alkylation was not observed when NaH was used as the base. It reacted with adenine forming sodium adenylide, which was then reacted with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate to obtain *N*-(Boc-aminoethylglycyl)-adenine ethyl ester (**38**) in moderate yield. The alkylation of 2-amino-6-chloropurine with ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate was facile with K_2CO_3 as the base and yielded the corresponding *N*-(Boc-aminoethylglycyl)-(2-amino-6-chloropurine)-ethyl ester (**39**) in excellent yield. All the compounds exhibited ^1H and ^{13}C NMR spectra consistent with the reported data. The ethyl esters were hydrolyzed in the presence of NaOH to give the corresponding acids (**41** - **45**), which were used for solid phase synthesis. The need for the exocyclic amino groups of adenine and guanine to be protected was eliminated, as these were found to be unreactive under the conditions used for peptide coupling.

The *N*-(Boc-aminoethylglycyl)-(N^2-isobutyrylguanin-7-yl)ethyl ester (**40**) was prepared by reacting compound **36** with N^2-isobutyrylguanine in the presence of potassium carbonate. The ester function was subsequently hydrolyzed to yield the acid **45**.

Scheme 9. Synthesis of the *aeg*PNA monomers



2.2.6. pH Titration

Since the *aep*PNA monomers carry a tertiary nitrogen in the pyrrolidine ring that has the potential to get protonated, thus conferring a positive charge, it is important to know the pKa of this group. Therefore, a pH titration experiment was carried out to determine the exact pKa of the pyrrolidine ring nitrogen atom.

The titration of the *aep*PNA-T monomer **10** after Boc-deptrotection, with NaOH was performed in order to determine the pKa of the pyrrolidine ring nitrogen. The pH titration curve exhibited three pKas, the first one at 2.92 pH corresponding to the carboxylic acid, the second at 6.7 pH corresponding to the pyrrolidine ring nitrogen and the third at 10.93 corresponding to the primary amine (Figure 9).

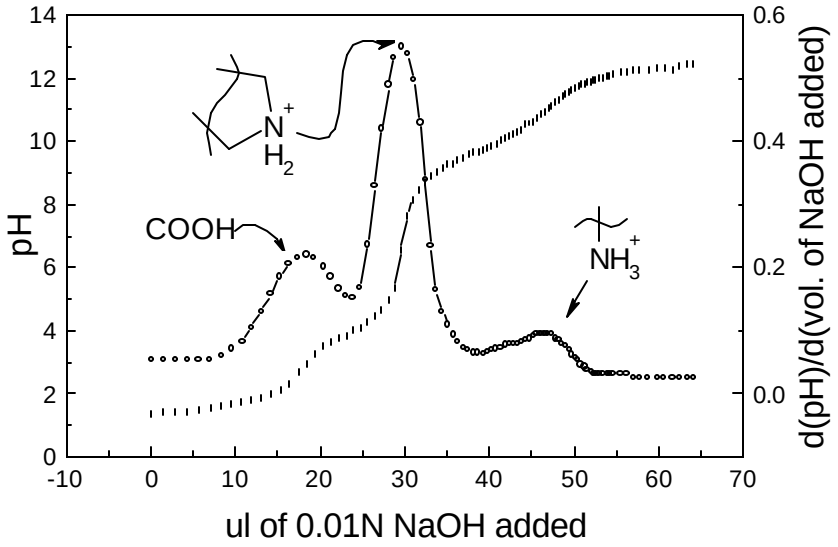


Figure 9. pH titration curve of (2*S*,4*S*) *aep*PNA-T monomer **10** with 0.01N NaOH

2.2.7. Solid Phase Peptide Synthesis

2.2.7a General Protocols for PNA Synthesis. Solid phase peptide synthesis protocols can be easily applied to the synthesis of PNAs. The ease of handling and scale-up procedures have made possible the synthesis of PNAs including those incorporating a large number of analogues in an endeavour to improve its favourable binding and biological properties.

As is the case with solid phase peptide synthesis, PNA synthesis is also done conveniently from the 'C' terminus to the 'N' terminus. For this, the monomeric units must have their amino functions suitably protected, and their carboxylic acid functions free. The most commonly used N-protecting groups for solid phase peptide synthesis are the *t*-butyloxycarbonyl (Boc) and the 9-flurorenylmethoxycarbonyl (Fmoc) groups. The Fmoc protection strategy has a drawback in PNA synthesis since a small amount of acyl migration has been observed under basic conditions from the tertiary amide to the free amine formed during piperidine deprotection.[43] Hence, the Boc-protection strategy was selected for the present work. The amino function of the monomers was protected as the corresponding Boc-derivative and the carboxylic acid function was

free to enable coupling with the resin-linked monomer. The diisopropylcarbodiimide (DIPCDI)/1-hydroxybenzotriazole (HOBt) activation strategy was employed for the coupling reaction.[44] Merrifield resin was selected as the solid polymeric matrix on which the oligomers were built. The first amino acid is linked to this matrix *via* a benzyl ester linkage. This can be cleaved either with a strong acid to yield the C-terminal free carboxylic acid, or with an amine to afford the C-terminal amide.

All the oligomers of the present work were synthesized manually on Merrifield resin. β-Alanine was selected as the linker amino acid. Being achiral, it would not interfere with the chirality-induced spectral properties of the pyrrolidyl monomeric units that bear two chiral centers each. Its contribution to the hydrophobicity of PNA is also negligible since it has only a short alkyl chain. *N*-Boc-β-alanine was linked to the resin by a benzyl ester linkage *via* the formation of its cesium salt.[45] The loading value of β-alanine on the resin was determined by the non-destructive picrate assay.[46,47] The resin loading was suitably lowered by partial capping of the free amino groups obtained after Boc-deprotection as N-acetates. The free uncapped amino groups on the resin were estimated once again by the picrate assay prior to commencing solid phase synthesis.

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

(i) Deprotection of the N-protecting Boc- group using 50% TFA in CH_2Cl_2.

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

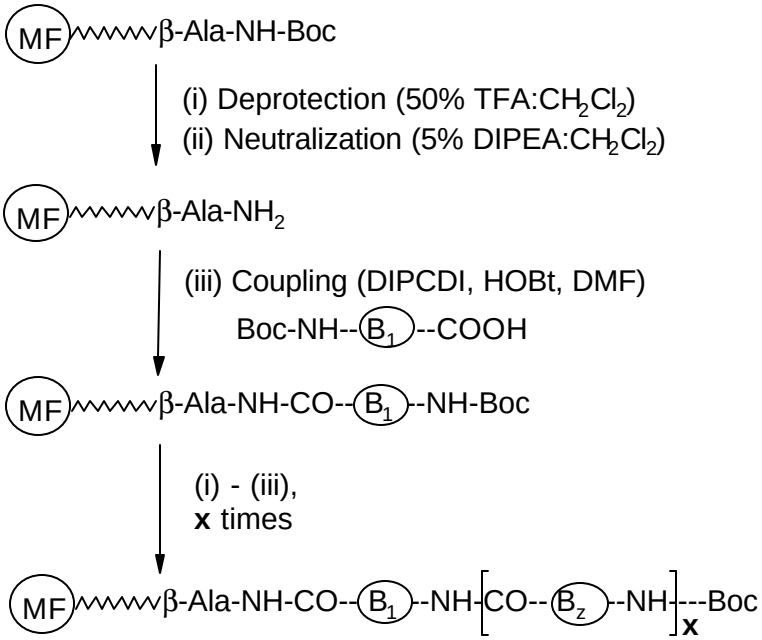
(iii) Coupling of the free amine with the free carboxylic acid group of the incoming monomer (3 to 4 equivalents). The coupling reaction was carried out in the presence of diisopropylcarbodiimide (DIPCDI) and 1-hydroxybenzotriazole (HOBt), used as a suppressing agent of racemization, in DMF or NMP as the solvent. The deprotection of the *N*-Boc protecting group and the coupling reaction were monitored by Kaiser's test.[48]

The Boc-deprotection step gives a positive Kaiser's test, wherein the resin beads as well as the solution are blue in colour. On the other hand, upon completion of the coupling reaction, the Kaiser's test is negative, the resin beads remaining colourless.

(iv) Capping of the unreacted amino groups using acetic anhydride in pyridine:CH_2Cl_2.

A typical synthesis cycle is depicted in Scheme 10.

Scheme 10. Schematic representation of solid phase peptide synthesis using the Boc-protection strategy



2.2.7b Synthesis of Cationic Aminoethylprolyl Peptide Nucleic Acids

(i) Homopyrimidine Sequences

The control aminoethylglycyl (*aeg*) PNA T_8 oligomer was first synthesized following the Boc- protection strategy outlined above. Chiral, positively charged aminoethylprolyl (*aep*) PNA-T monomer- containing sequences were subsequently made incorporating the *aep*PNA-T unit at one or more pre-determined positions within the octamer. The series of octamer sequences comprising aminoethylglycyl PNA-T and/or aminoethylprolyl PNA-T units is detailed in table 1. The capping step at the end of each

coupling cycle was not deemed necessary, as the coupling reaction was monitored manually to have gone to completion with a very high coupling efficiency (greater than 90%).

With a view to exploring the end-effects of the *aep*PNA monomeric unit, a single *aep* unit was introduced at either the 'N' or 'C' terminus (Table 1, entry 2 - 5). The study of the effect of introduction of increasing number of *aep* units on the stability and selec

Table 1. Resin-linked PNA Sequences Synthesized by Solid Phase Synthesis

Entry	Resin-linked PNA Oligomer
1	MF--β-ala- TTTTTTTT -Boc
2	MF--β-ala- TTTTTTt -Boc
3	MF--β-ala- TTTTTTt -Boc
4	MF--β-ala- tTTTTTTT -Boc
5	MF--β-ala- t TTTTTTT -Boc
6	MF--β-ala- t TTT t TTT -Boc
7	MF--β-ala- t TTT t TTT -Boc
8	MF--β-ala- t T t T t T t T -Boc
9	MF--β-ala- t T t T t T t T -Boc
10	MF--β-ala- t t t t t t t t -Boc
11	MF--β-ala- t t t t t t t t -Boc
12	MF--β-ala- TTTTTTTT-Lys -Boc
13	MF--β-ala-Lys - TTTTTTTT -Boc
14	MF--β-ala- TTTTTTTT -CF(OPiv)$_2$
15	MF--β-ala- t t t t t t t t -CF(OPiv)$_2$
16	MF--β-ala- AAAAAAAA -Boc
17	MF--β-ala- TTATTATTATAT -Boc
18	MF--β-ala- TTATTATTATA t -Boc
19	MF--β-ala- TTATTATTATA t -Boc
20	MF--β-ala- t t a t t a t t a t a t -Boc
21	MF--β-ala- TCACTAGATG -Boc
22	MF--β-ala- TCAC t AGATG -Boc
23	MF--β-ala- TCAC t AGATG -Boc
24	MF--β-ala- t CAC t AGA t G -Boc
25	MF--β-ala- t CAC t AGA t G -Boc
26	MF--β-ala- TCA c TAGATG -Boc
27	MF--β-ala- TCA c TAGATG -Boc
28	MF--β-ala- TCACTAG a TG -Boc
29	MF--β-ala- TCACTAG a TG -Boc
30	MF--β-ala- TCACTA g ATG -Boc
31	MF--β-ala- TCACTA g ATG -Boc
32	MF--β-ala- TTTCTCT -Lys -Aha-Lys -Aha-Lys - TCTCTTT -Boc
33	MF--β-ala- TTTCTCT -Lys -Aha-Lys -Aha-Lys - T c T c TTT -Boc
34	MF--β-ala- TTTCTCT -Lys -Aha-Lys -Aha-Lys - T^{7}GT^{7}GTTT -Boc
35	MF--β-ala- TTTCTCT -Lys -Aha-Lys -Aha-Lys - T ^{7}g T ^{7}g TTT -Boc
36	MF--β-ala- TTTCTCT -Lys -Aha-Lys -Aha-Lys - TGTGTTT -Boc

MF = Merrifield resin; Uppercase letters denote *aeg*PNA units. Lowercase letters

This was envisioned to supply data on the efficiency of the aminoethylprolyl backbone at accommodating purines/pyrimidines and exposing the differences in the *aep*PNA monomers bearing the different nucleobases. The synthesized sequences are listed in table 1.

2.2.8. Cleavage of the PNA Oligomers from the Solid Support

2.2.8a Cleavage from the Resin by Strong Acid. The cleavage of peptides from the Merrifield resin by strong acids like trifluoromethane sulphonic acid (TFMSA)-trifluoroacetic acid (TFA) yields peptides with free carboxylic acids at their 'C' termini.[49] The synthesized PNA oligomers were cleaved from the resin using this procedure to obtain sequences bearing β-alanine free carboxylic acids at their 'C' termini (**46** - **60**, **63** - **83**). After commencing the cleavage reaction, aliquots were removed after 30min, 1h, 2h and 24h, the peptides isolated and analyzed by HPLC. A cleavage time of ~2h at room temperature was found to be optimum. The peptides from aliquots removed prior to 2h indicated incomplete deprotection. This was evident from the greater number of peaks with a higher retention time as observed by analytical HPLC. These peaks were absent in the aliquots removed after 2h. The exocyclic amino groups of cytosine, adenine and guanine, if protected as benzyloxycarbonyl, are also cleaved during this process. If the N6-exocyclic amino group of adenine is, however, protected as a benzoyl group, then its deprotection must be carried out under alkaline conditions employing ammonia or ethylenediamine. The ammonia treatment is carried out at 55 ^{o}C for 16h. These conditions have proved to be harsh on the peptide, as they are conducive to degradation of the peptide by successive 'N' -terminal cleavage.[27]

The relatively milder conditions involve treatment with ethylenediamine in ethanol at room temperature overnight. This protocol was followed for oligomers containing N6-benzoyl adenine *aep* monomeric units, and was carried out prior to their cleavage with TFMSA-TFA.

2.2.8b Cleavage by Aminolysis. The resin-bound PNA oligomer was treated with spermine at 55 °C, 18h to yield oligomers with sperminamide at their 'C' terminii (**61**, **62**).

2.2.9. Purification of the PNA Oligomers

All the cleaved oligomers were subjected to initial gel filtration. These were subsequently purified by reverse phase FPLC on a semi-preparative C8 RP column by gradient elution using an ascending gradient of acetonitrile in water containing 0.1% TFA, or by isocratic elution in 10% acetonitrile- water containing 0.1% TFA on a semi-preparative HPLC RP C4 column. In some cases, FPLC did not produce a clean single peak profile. Hence, the sample was heated at ~80 °C for 4-5 min and then injected to destroy any secondary structure that might exist. Samples containing multiple *aep*PNA units, and therefore, multiple positive charges were suspended in buffer containing 0.1%TFA and allowed to stand for 2-3h prior to injection and chromatography. The purity of the oligomers was then checked by reverse phase HPLC on a C18 RP column and confirmed by MALDI-TOF mass spectroscopic analysis. Some representative HPLC profiles and mass spectra are shown in Figures 10 and 11. The PNA sequences obtained are listed in Table 2.

2.2.10. Synthesis of Complementary Oligonucleotides

The oligodeoxynucleotides (**84** - **98**, Table 3) were synthesized on a Pharmacia Gene Assembler Plus DNA synthesizer using the standard β-cyanoethyl phosphoramidite chemistry. The oligomers were synthesized in the 3' 5' direction on a CPG solid support, followed by ammonia treatment.[50] The oligonucleotides were de-salted by gel filtration, their purity ascertained by RP HPLC on a C18 column to be

more than 98% and were used without further purification in the biophysical studies of PNA.

Table 2. PNA Sequences

PNA	Sequence Composition
46	H- TTTTTTTT -NH-$(CH_2)_2$-COOH
47	H- **t** TTTTTTT -NH-$(CH_2)_2$-COOH
48	H- **t** TTTTTTT -NH-$(CH_2)_2$-COOH
49	H- TTTTTTT **t** -NH-$(CH_2)_2$-COOH
50	H- TTTTTTT **t** -NH-$(CH_2)_2$-COOH
51	H- TTT **t** TTT **t** -NH-$(CH_2)_2$-COOH
52	H- TTT **t** TTT **t** -NH-$(CH_2)_2$-COOH
53	H- T **t** T **t** T **t** T **t** -NH-$(CH_2)_2$-COOH
54	H- T **t** T **t** T **t** T **t** -NH-$(CH_2)_2$-COOH
55	H- **t t t t t t t t** -NH-$(CH_2)_2$-COOH
56	H- **t t t t t t t t** -NH-$(CH_2)_2$-COOH
57	H-Lys - TTTTTTTT -NH-$(CH_2)_2$-COOH
58	H- TTTTTTTT -Lys -NH-$(CH_2)_2$-COOH
59	Fluorescein-CO- TTTTTTTT -NH-$(CH_2)_2$-COOH
60	Fluorescein-CO- **t t t t t t t t** -NH-$(CH_2)_2$-COOH
61	Fluorescein-CO-TTTTTTTT-NH-$(CH_2)_2$-CO-NH-$(CH_2)_3$-NH-$(CH_2)_4$-NH-$(CH_2)_3$-NH_2
62	Fluorescein-CO-**tttttttt**-NH-$(CH_2)_2$-CO-NH-$(CH_2)_3$-NH-$(CH_2)_4$-NH-$(CH_2)_3$-NH_2
63	H- AAAAAAAA -NH-$(CH_2)_2$-COOH
64	H- TATATTATTATT -NH-$(CH_2)_2$-COOH
65	H- **t** ATATTATTATT -NH-$(CH_2)_2$-COOH
66	H- **t** ATATTATTATT -NH-$(CH_2)_2$-COOH
67	H- **t a t a t t a t t a t t** -NH-$(CH_2)_2$-COOH
68	H- GTAGATCACT -NH-$(CH_2)_2$-COOH
69	H- GTAGA **t** CACT -NH-$(CH_2)_2$-COOH
70	H- GTAGA **t** CACT -NH-$(CH_2)_2$-COOH
71	H- G **t** AGA **t** CAC **t** -NH-$(CH_2)_2$-COOH
72	H- G **t** AGA **t** CAC **t** -NH-$(CH_2)_2$-COOH
73	H- GT **a** GATCACT -NH-$(CH_2)_2$-COOH
74	H- GT **a** GATCACT -NH-$(CH_2)_2$-COOH
75	H- GTA **g** ATCACT -NH-$(CH_2)_2$-COOH
76	H- GTA **g** ATCACT -NH-$(CH_2)_2$-COOH
77	H- GTAGAT **c** ACT -NH-$(CH_2)_2$-COOH
78	H- GTAGAT **c** ACT -NH-$(CH_2)_2$-COOH
79	H- TTTCTCT -Lys -Aha-Lys -Aha-Lys - TCTCTTT -NH-$(CH_2)_2$-COOH
80	H- TTT **c** T **c** T -Lys -Aha-Lys -Aha-Lys - TCTCTTT -NH-$(CH_2)_2$-COOH
81	H- TTT^{7}*G*T^{7}*G*T -Lys -Aha-Lys -Aha-Lys - TCTCTTT -NH-$(CH_2)_2$-COOH
82	H- TTT^{7}***g***T^{7}***g***T -Lys -Aha-Lys -Aha-Lys - TCTCTTT -NH-$(CH_2)_2$-COOH
83	H- TTTGTGT -Lys -Aha-Lys -Aha-Lys - TCTCTTT -NH-$(CH_2)_2$-COOH

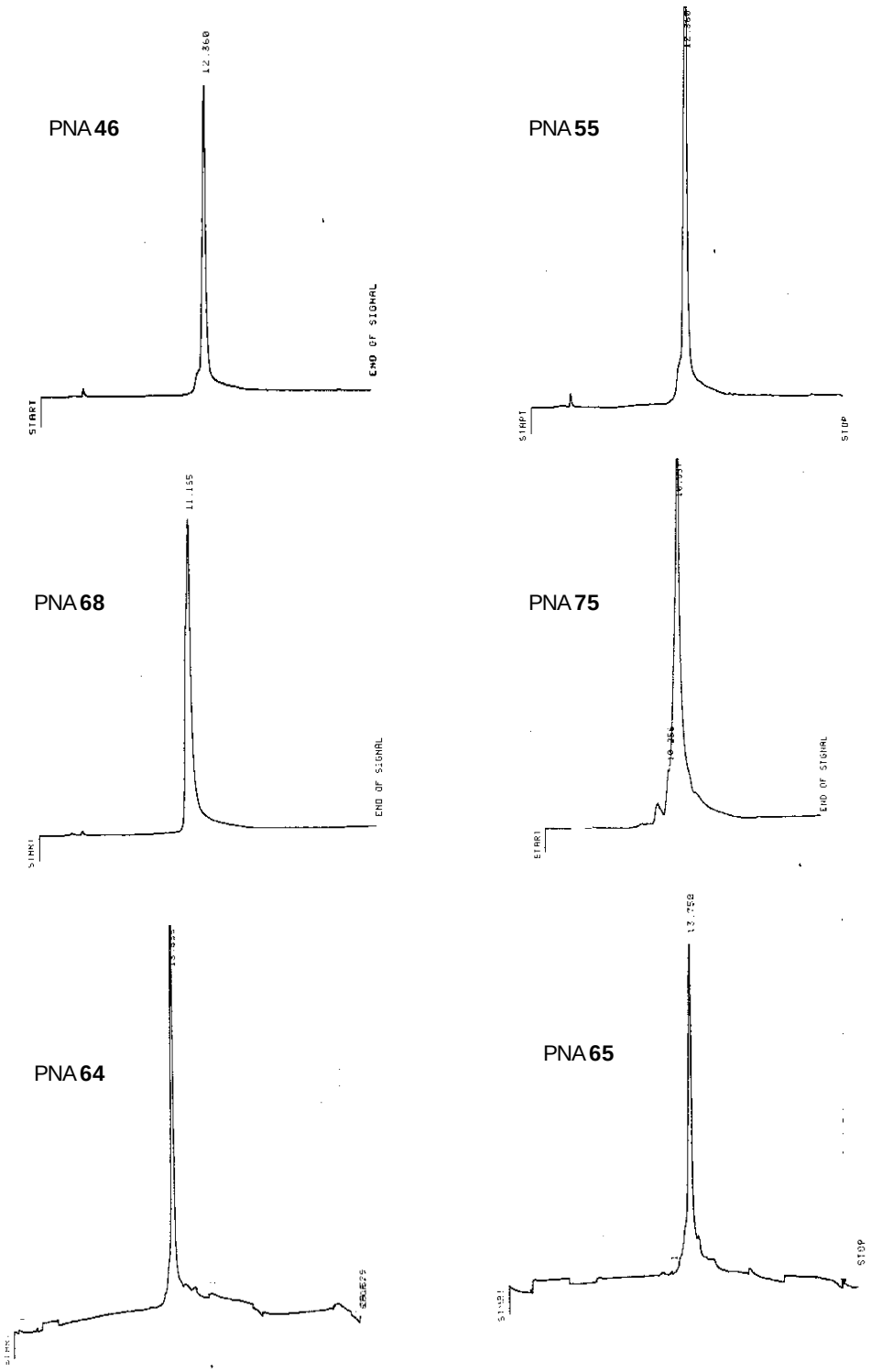
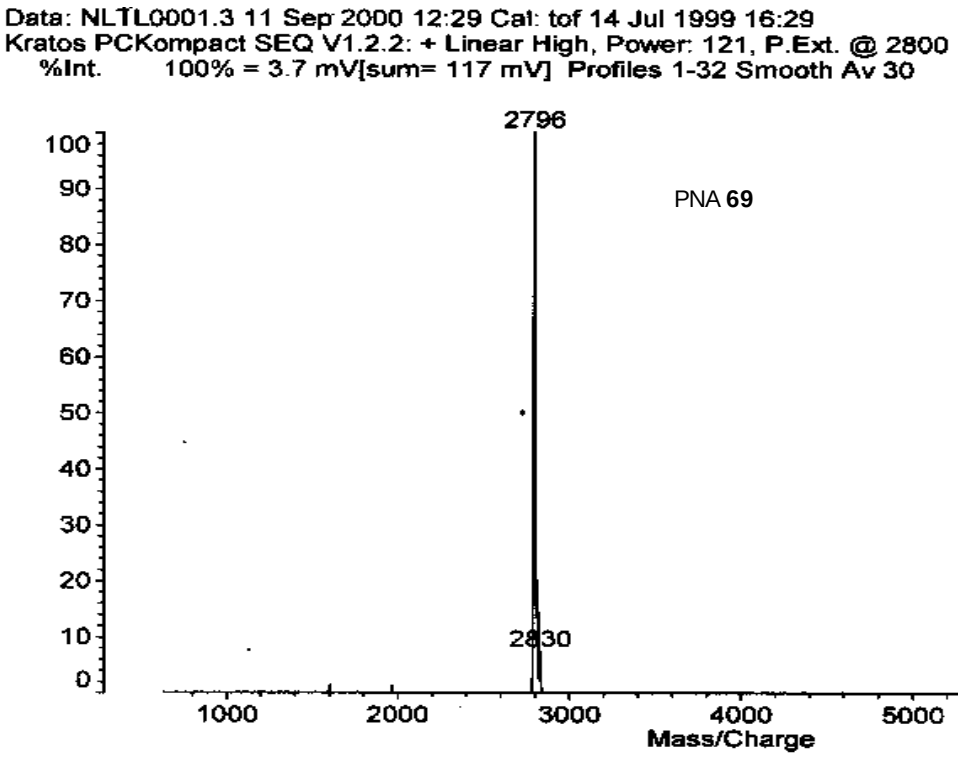


Figure 10. Representative HPLC profiles of *aeg* and *aep*PNA oligomers



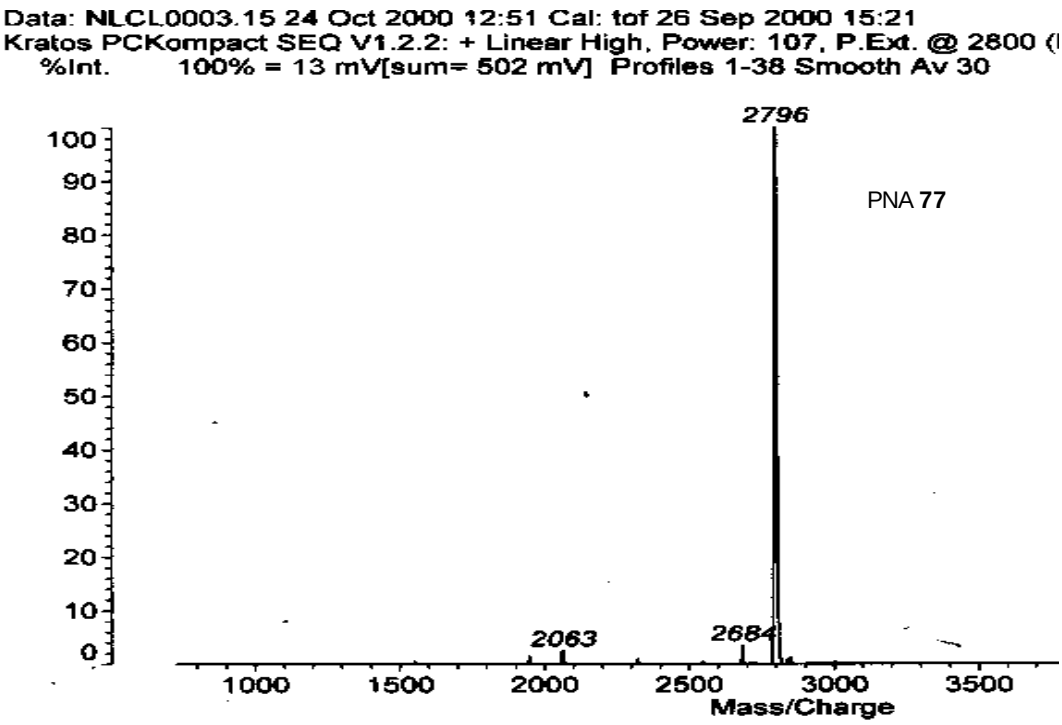


Figure 11. MALDI-TOF of representative *aep*PNAs **69** and **77**

Table 3. DNA Oligomer Sequences

DNA	Oligomer Sequences 5' → 3'
84	G C A A A A A A A A C G
85	T T T T T T T T
86	G C A A A T A A A A C G
87	A A T A A T A A T A T A
88	A T A T A A T A A T A A
89	G C A A A G A G A C T
90	T C A G A G A A A C G
91	G C A A A T A G A C T
92	A G T G A T C T A C
93	A G T G A T C C A C
94	A G T G A T A T A C
95	A G T G T T C T A C
96	A G T T A C C T A C
97	C A T C T A G T G A
98	G T A G A T C A C T

2.3. CONCLUSIONS

This Chapter describes the synthesis of novel *aep*PNA monomers. The *aep*PNA-T monomers were synthesized by direct Mitsunobu reaction of the 4-hydroxy compound with N3-benzoylthymine, whereas the *aep*PNA-A/G/C monomers were synthesized *via* a common mesyl intermediate. The *aep*PNA monomers were successfully incorporated into PNA oligomers using solid phase peptide synthesis methodology. Upon cleavage and subsequent purification, oligomers that were ~98% pure were obtained for use in the biophysical studies with DNA. The DNA sequences required were synthesized using the standard phosphoramidite chemistry.

The next Chapter describes the biophysical studies (UV, CD and gel retardation assays of *aep*PNA:DNA hybrids.

2.4. EXPERIMENTAL

The chemicals used were of laboratory or analytical grade. All the solvents used were purified according to the literature procedures.[51] All the reactions were monitored for completion by TLC. Usual work-up implies sequential washing of the organic extract with water and brine followed by drying over anhydrous sodium sulphate and evaporation under vacuum.

Column chromatography was performed for purification of compounds on LOBA chemie silica gel (100- 200 mesh). TLCs were carried out on pre-coated silica gel GF_{254} aluminium sheets (Merck 5554). TLCs were run in either dichloromethane with an appropriate quantity of methanol or in petroleum ether with an appropriate quantity of added ethyl acetate for most compounds. Free acids were chromatographed on TLC using a solvent system of *iso*-propanol: acetic acid: water in the proportion 9: 1: 1. The compounds were visualized with UV light and/ or by spraying with Ninhydrin reagent subsequent to Boc-deprotection (exposing to HCl vapors) and heating.

^{1}H(200 MHz/300MHz) and ^{13}C (50MHz) NMR spectra were recorded on a Bruker ACF 200 spectrometer fitted with an Aspect 3000 computer and all the chemical shifts are referred to internal TMS for ^{1}H and chloroform-d for ^{13}C. The chemical shifts are quoted in δ (ppm) scale. In compounds that bear a tertiary amide group, splitting of NMR signals was observed due to the presence of rotamers. In such cases, the major isomer is designated as 'maj' and the minor isomer, 'min'.

Optical rotations were measured on a JASCO DIP-181 polarimeter and CD spectra were recorded on a JASCO J715 spectropolarimeter.

Mass spectra were recorded on a Finnigan-Matt mass spectrometer, while MALDI-TOF spectra were obtained from a KRATOS PCKompact instrument.

N3-benzoylthmine[35] ***(2)***

To a stirred solution of thymine (5.0g, 40mmol) in dry acetonitrile (40ml) and dry pyridine (10ml) in an ice-bath, benzoyl chloride (11.2ml, 96mmol), was added dropwise. Stirring was continued at room temperature overnight, when dibenzoyl thymine (**1**) was found to be present. This was converted to the monobenzoyl derivative **2** by treating with 0.25M K_2CO_3 in dioxane: water, 1:1 (75ml) and monitored by TLC. The dioxane was removed under vacuum. The solid obtained was filtered and washed with water. The pure product **2** was obtained upon desiccation.

N^4-benzyloxycarbonylcytosine[36] ***(3)***

Cytosine (1.0g, 9.0mmol) was suspended in dry pyridine (100ml) at 0 ^oC. CBz-Cl (3.2ml, 22.5mmol) was added and the reaction was stirred under nitrogen overnight. The pyridine suspension was evaporated to dryness. Water (10ml) and dilute HCl was added to bring the pH to 4.0. The resulting white precipitate was filtered off, washed with water and partially dried under vacuum. The wet precipitate was boiled in absolute ethanol (10ml), cooled to 0 ^oC, filtered, washed thoroughly with ether and dried under vacuum (Yield = 1.1g, 50%).

N^6-benzoyladenine[37] ***(4)***

Benzoyl chloride (12.7ml, 110mmol) was added to a suspension of adenine (5.0g, 37mmol) in dry pyridine (50ml). The reaction mixture was refluxed for 2h. The solvent was removed under vacuum. The crude product was triturated with warm $NaHCO_3$ solution, which caused the separation of an oily phase. When the aqueous suspension was shaken with $CHCl_3$, the amide precipitated. The crystals were filtered off and washed with water (Yield = 7.5g, 85%)

*N^2-isobutyrylguanine[38] (**5**)*

To an ice-cold stirred solution of guanine (1.0g, 6.6mmol) in dry pyridine (10ml) was added dropwise isobutyryl chloride (0.83ml, 7.9mmol). Stirring was continued overnight. The solvent was evaporated under vacuum. The residue was re-crystallized from boiling ethanol: water, 1:1 to yield the product **5** (1.0g, 73.5%).

*N-Boc-2-aminoethanol (**12**)*

To a cooled, stirred solution of 2-aminoethanol **11** (5.9ml, 0.098mol) and sodium hydroxide (3.6g, 0.098mol) in water-dioxane (1:1), was added drop-wise Boc-azide (10ml, 0.081mol). The reaction was stirred at room temperature overnight. The dioxane was then removed under vacuum, and the water layer extracted several times with ethyl acetate. The organic layer was dried over sodium sulphate and then, evaporated to dryness under vacuum to get the product 2-*N*-Boc-aminoethanol **12** (5.37g, 41% yield), which was used in further steps without purification.

*2-N-Boc-aminoethylbromide (**13**)*

To an ice-cooled solution of 2-*N*-Boc-aminoethanol **12** (2.5g, 15.5mmol) and carbon tetrabromide (7.5g, 22.7mmol) in dry benzene (25ml), was added triphenyl phosphine (4.7g, 17.9mmol) in small portions over 10-15min. After 30 min. solvent was evaporated and the product immediately purified by silica gel column chromatography to get 2-*N*-Boc-aminoethylbromide **13** (5.53g, 79%).

^{1}H NMR ($CDCl_3$) δ: 4.95(br s, 1H, N*H*), 3.50(m, 4H, (CH_2)$_2$), 1.45(s, 9H, C(CH_3)$_3$)

*1-(N-Boc-aminoethyl)-4-(R)-hydroxy-2-(S)-proline methyl ester (**8**)*

A mixture of 4-(*R*)-hydroxy-2-(*S*)-proline methyl ester hydrochloride **7** (3.24g, 17.9mmol), 2-(*N*-Boc)-aminoethylbromide **13** (2.0g, 8.9mmol) and anhydrous potassium carbonate (3.69g, 26.7mmol) were stirred together in DMF: acetonitrile (1:1) at room temperature for 72 h under argon atmosphere. After completion of reaction as

indicated by TLC, solvents were removed *in vacuo*. The residue was taken in water and extracted with ethyl acetate (4x 30ml). The organic layer was dried over sodium sulphate and concentrated to get the crude product, which was purified by silica gel column chromatography. The pure product **13** was obtained in 56% yield.

^{1}H NMR ($CDCl_3$) δ: 5.30(br s, 1H, N*H*), 4.40(m, 1H, *H*4), 3.70(s, 3H, OCH_3), 3.50(t, 1H, *H*5), 3.38(dd, 1H, *H*5'), 3.10(dd, 2H, Boc-NH-CH_2), 2.70(br m, 4H, *H*2, Boc-NH-CH$_2$-CH_2, O*H*), 2.50(dd, 1H, *H*3), 2.10(m, 1H, *H*3'), 1.40(s, 9H, C(CH_3)$_3$)

*1-(N-Boc-aminoethyl)-4-(R)-hydroxy-2-(R)-proline methyl ester (**16**)*

A mixture of 4-(*R*)-hydroxy-2-(*R*)-proline methyl ester hydrochloride **15** (3.24g, 17.9mmol), 2-(*N*-Boc)-aminoethylbromide **13** (2.0g, 8.9mmol) and anhydrous potassium carbonate (3.69g, 26.7mmol) were stirred together in DMF: acetonitrile (1:1) at room temperature for 72 h under argon atmosphere. After completion of reaction as indicated by TLC, solvents were removed *in vacuo*. The residue was taken in water and extracted with ethyl acetate (4x 30ml). The organic layer was dried over sodium sulphate and concentrated to get the crude product, which was purified by silica gel column chromatography. The pure product **16** was obtained in 56% yield from 2-(*N*-Boc)-aminoethyl bromide.

^{1}H NMR ($CDCl_3$) δ: 5.30(br s, 1H, N*H*), 4.25(br s, 1H, *H*4), 3.72(s, 3H, OCH_3), 3.37(m, 1H, *H*5), 3.40-3.05(m, 4H, *H*5', Boc-NH-CH_2, O*H*), 2.65(m, 3H, *H*2, Boc-NH-CH$_2$-CH_2), 2.40(m, 1H, *H*3), 1.92(dd, 1H, *H*3'), 1.40(s, 9H, C(CH_3)$_3$)

^{13}C NMR ($CDCl_3$) δ: 174.9(*C*OOCH$_3$), 156.1(*C*OOC(CH$_3$)$_3$), 78.6(*C*(CH$_3$)$_3$), 70.0(O*C*H$_3$), 64.6(*C*4), 61.6(*C*5), 53.9(*C*3), 51.9(*C*2), 39.0(Boc-NH-(*C*H$_2$)$_2$), 28.3((*C*H$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(N3-benzoylthymin-1-yl)-2-(S)-proline methyl ester (**9**)*

To a stirred solution of 1-(*N*-Boc-aminoethyl)-4-(*R*)-hydroxy-2-(*S*)-proline methyl ester **8** (0.9g, 3.2mmol), N3-benzoylthymine **2** (0.74g, 3.2mmol) and triphenyl phosphine (1g, 3.8 mmol) in dry THF (10ml) at room temperature, was added dropwise diisopropylazodicarboxylate (DIAD, 0.76ml, 3.8mmol). After completion of the reaction as indicated by TLC (24h), the solvent was removed *in vacuo* and residue purified by silica gel column chromatography to get the pure product **9** (0.6g, 38%).

^{1}H NMR ($CDCl_3$) δ: 8.10(s, 1H, *T-H*6), 7.90(d, 2H, Bz, *o*-C*H*), 7.65(t, 1H, Bz-*p*-C*H*), 7.50(t, 2H, Bz, *m*-C*H*), 5.25(m, 2H, Boc-N*H*, *H*4), 3.80(s, 3H, OCH_3), 3.25(m, 4H, Boc-NH-CH_2, *H*5, *H*5'), 2.80(m, 3H, Boc-NH-CH_2-CH_2, *H*2), 2.60(m, 1H, *H*3), 2.00(s, 3H, *T*-CH_3), 1.95(m, 1H, *H*3'), 1.45(s, 9H, (CH_3)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(N3-benzoylthymin-1-yl)-2-(R)-proline methyl ester (**17**)*

To a stirred solution of 1-(*N*-Boc-aminoethyl)-4-(*R*)-hydroxy-2-(*R*)-proline methyl ester **16** (0.9g, 3.2mmol), N3-benzoylthymine **2** (0.74g, 3.2mmol) and triphenyl phosphine (1g, 3.8 mmol)in dry THF (10ml) at room temperature, was added dropwise diisopropylazodicarboxylate (DIAD, 0.76ml, 3.8mmol). After completion of the reaction as indicated by TLC (24h), the solvent was removed *in vacuo* and residue purified by silica gel column chromatography to get the pure product **17** (0.47g, 30%).

^{1}H NMR ($CDCl_3$) δ: 7.92(d, 2H, Bz-*o*-C*H*), 7.65(t, 1H, Bz-*p*-C*H*), 7.50(t, 2H, Bz-*m*-C*H*), 7.40(s, 1H, *T-H*6), 5.25(m. 1H, *H*4), 4.90(br t, 1H, Boc-N*H*), 3.90(dd, 1H, *H*5), 3.72(s, 3H, OCH_3), 3.30(m, 3H, *H*5', Boc-NH-CH_2), 2.92(dd, 1H, *H*2), 2.78(t, 2H, Boc-NH-CH_2-CH_2), 2.55(m, 1H, *H*3), 2.20(m, 1H, *H*3'), 2.00(s, 3H, *T*-CH_3), 1.45(s, 9H, C(CH_3)$_3$)

^{13}C NMR ($CDCl_3$) δ: 172.5(*C*$OOCH_3$), 169.0(*C*$OOC(CH_3)_3$), 162.5(*T-C*2), 155.8(Ph-*C*O), 149.5(*T-C*4), 137.5(*T-C*6), 134.8, 131.3, 130.0 & 128.9(Ph), 111.1(*T-C*5), 78.8($C(CH_3)_3$), 63.1(*C*4), 55.5(Boc-NH-CH_2), 53.7(*C*2), 51.4(OCH_3), 50.6(*C*5), 38.7(Boc-NH-CH_2-CH_2), 35.1(*C*3), 28.1($C(CH_3)_3$), 12.3(*T-C*H_3)

*1-(N-Boc-aminoethyl)-4-(R)-O-mesyl-2-(S)-proline methyl ester (**19**)*

To a stirred ice-cooled solution of 1-(*N*-Boc-aminoethyl)-4-(*R*)-hydroxy-2-(*S*)-proline methyl ester **8** (1.66g, 5.9mmol) in dry pyridine (15ml), was added dropwise methane sulphonyl chloride (0.55ml, 7.1mmol). After 2h, upon completion of reaction, the pyridine was removed under vacuum and the residue was taken in water and extracted with ethyl acetate (5x15ml). The organic layer was dried over sodium sulphate and concentrated to get the crude product **19**, which was purified by silica gel column chromatography (0.7g, 33%).

^{1}H NMR ($CDCl_3$) δ: 5.20(m, 2H, N*H*, *H*4), 3.70(s, 3H, OCH_3), 3.65(t, 1H, *H*5), 3.50(dd, 1H, *H*5'), 3.15(m, 2H, Boc-NH-CH_2), 3.00(s, 3H, SO_2-CH_3), 2.85(dd, 1H, *H*2), 2.75(m, 2H, Boc-NH-CH_2-CH_2), 2.35(m, 2H, *H*3, *H*3'), 1.45(s, 9H, $C(CH_3)_3$)

*1-(N-Boc-aminoethyl)-4-(R)-O-mesyl-2-(R)-proline methyl ester (**20**)*

To a stirred ice-cooled solution of 1-(*N*-Boc-aminoethyl)-4-(*R*)-hydroxy-2-(*R*)-proline methyl ester **16** (1.25g, 4.3mmol) in dry pyridine (10ml), was added dropwise methane sulphonyl chloride (0.4ml, 5.2mmol). After 2h, upon completion of reaction, the pyridine was removed under vacuum and the residue was taken in water and extracted with ethyl acetate (4x15ml). The organic layer was dried over sodium sulphate and concentrated to get the crude product **20**, which was purified by silica gel column chromatography (0.7g, 44%).

^{1}H NMR ($CDCl_3$) δ: 5.27(m, 1H, *H*4), 5.20(m, 1H, Boc-N*H*), 3.74(s, 3H, OCH_3), 3.40(m, 2H, *H*5, *H*5'), 3.20(m, 2H, Boc-NH-CH_2), 3.05(s, 3H, SO_2-CH_3), 2.85(m, 2H, Boc-NH-CH_2-CH_2), 2.60(m, 2H, *H*2, *H*3), 2.35(m, 1H, *H*3'), 1.45(s, 9H, (CH_3)$_3$)

^{13}C NMR ($CDCl_3$) δ: 172.6(*C*OOCH$_3$), 155.6(*C*OOC(CH_3)$_3$), 78.6(*C*(CH_3)$_3$), 78.2(O*C*H$_3$), 63.6(*C*4), 58.3(*C*5), 53.1(Boc-NH-*C*H$_2$), 51.7(SO_2-*C*H$_3$), 38.6(Boc-NH-CH_2-*C*H$_2$), 38.2(*C*2), 36.1(*C*3), 28.0((*C*H$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(N^4-benzyloxycarbonyl-cytosin-1-yl)-2-(S)-proline methyl ester (**21a**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4-(*R*)-O-mesyl-2-(*S*)-proline methyl ester **19** (0.6g, 1.6mmol), N^4-benzyloxycarbonyl-cytosine **3** (1.0g, 4.1mmol), anhydrous potassium carbonate (1.1g, 8.2mmol) and 18-crown-6(0.17g, 0.5mmol) in anhydrous DMF(6ml) was stirred under nitrogen atmosphere at 70°C for 5h. The solvent was completely removed under vacuum and the residue purified by silica gel column chromatography to get the pure product (0.25g, 30% yield) as the major product.

^{1}H NMR ($CDCl_3$) δ: 8.45(d, 1H, *C*-*H*6), 7.40(s, 5H, Ph), 7.25(d, 1H, *C*-*H*5), 5.35(t, 1H, Boc-N*H*), 5.20(s, 3H, Ph-CH_2, *H*4), 3.75(s, 3H, OCH_3), 3.20(br m, 4H, *H*5, *H*5', Boc-NH-CH_2), 2.80(m, 2H, Boc-NH-CH_2-CH_2), 2.60(m, 1H, *H*2), 1.90(m, 2H, *H*3, *H*3'), 1.45(s, 9H, C(CH_3)$_3$)

In addition to this, 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonyl-cytosin-2-O-)-2(*S*)-proline methyl ester **22a** was obtained as a side-product (0.19g, 22% yield).

1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonyl-cytosin-2-O-)-2(*S*)-proline methyl ester (**22a**)

^{1}H NMR ($CDCl_3$) δ: 8.35(d, 1H, *C*-*H*6), 7.60(d, 1H, *C*-*H*5), 7.40(s, 5H, Ph), 5.35(m, 2H, Boc-N*H*, *H*4), 5.20(s, 2H, Ph-CH_2), 3.70(s, 3H, OCH_3), 3.35(m, 2H, *H*5, *H*5'),

3.15(m, 2H, Boc-NH-C*H*$_2$), 2.80(m, 2H, Boc-NH-CH$_2$-C*H*$_2$), 2.50(m, 1H, *H*3), 2.20(m, 1H, *H*3'), 1.45(s, 9H, C(C*H*$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(N^4-benzyloxycarbonyl-cytosin-1-yl)-2-(R)-proline methyl ester (**21b**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4-(*R*)-O-mesyl-2-(*R*)-proline methyl ester **20** (0.13g, 0.35mmol), N^4-benzyloxycarbonyl-cytosine **3** (0.22g, 0.89mmol), anhydrous potassium carbonate (0.25g, 1.78mmol) and 18-crown-6(0.04g, 0.11mmol) in anhydrous DMF(2ml) was stirred under nitrogen atmosphere at 70^{o}C for 5h. The solvent was completely removed under vacuum and the residue purified by silica gel column chromatography to get the pure product **21b** (0.11g, 60% yield) as the major product.

^{1}H NMR ($CDCl_3$) δ: 7.90(d, 1H, *C-H*6), 7.35(s, 5H, Ph), 7.25(d, 1H, *C-H*5), 5.40(m, 1H, Boc-N*H*), 5.20(s, 2H, Ph-C*H*$_2$), 4.90(m, 1H, *H*4), 3.95(dd, 1H, *H*5), 3.74(s, 3H, OC*H*$_3$), 3.40(m, 1H, *H*5'), 3.25(m, 2H, Boc-NH-C*H*$_2$), 2.95(m, 1H, *H*2), 2.77(m, 2H, Boc-NH-CH$_2$-C*H*$_2$), 2.60(m, 1H, *H*3), 2.15(m, 1H, *H*3'), 1.45(s, 9H, C(C*H*$_3$)$_3$)

The less polar fractions were concentrated to obtain the O^2-substituted minor product **22b** (0.03g, 16%).

1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonyl-cytosin-2-*O*-)-2(*R*)-proline methyl ester (**22b**)

^{1}H NMR ($CDCl_3$) δ: 8.35(d, 1H, *C-H*6), 7.60(d, 1H, *C-H*5), 7.40(s, 5H, Ph), 5.40(m, 1H, Boc-N*H*), 5.30(m, 1H, *H*4), 5.20(s, 2H, Ph-C*H*$_2$), 3.75(s, 3H, OC*H*$_3$), 3.65(m, 2H, *H*5, *H*5'), 3.15(m, 2H, Boc-NH-C*H*$_2$), 2.80(m, 1H, *H*2), 2.65(m, 2H, Boc-NH-CH$_2$-C*H*$_2$), 2.30(m, 2H, *H*3, *H*3'), 1.45(s, 9H, C(C*H*$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(N^6-benzoyladenin-9-yl)-2-(S)-proline methyl ester (**23a**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4-(*R*)-O-mesyl-2-(*S*)-proline methyl ester **19** (0.2g, 0.55mmol), N^6-benzoyladenine **4** (0.16g, 0.66mmol), anhydrous potassium carbonate (0.09g, 0.66mmol) and a catalytic amount of 18-crown-6 (0.06g, 0.2mmol) was stirred in dry DMF (3ml) at 75°C for 48h. After completion of the reaction, the solvent was removed under vacuum. The residue was re-dissolved in water and extracted with chloroform (4x 5ml). The organic layer was then dried over anhydrous sodium sulphate and concentrated to get crude product, which was purified by silica gel column flash chromatography to obtain the pure product **23a** (0.18g, 65%).

^{1}H NMR ($CDCl_3$) δ: 9.27(br s, 1H, Benzamide-N*H*), 8.76(s, 1H, *A-H2*), 8.70(s, 1H, *A-H8*), 8.20(m, 2H, Bz, *o*-C*H*), 7.56(m, 3H, Bz, *p*-C*H*, *m*-C*H*), 5.40(m, 1H, *H4*), 5.15(m, 1H, Boc-N*H*), 3.74(s, 3H, OCH_3), 3.60-2.60(m, 9H, *H5*, *H5'*, *H3*, *H3'*, *H2*, Boc-NH-(CH_2)$_2$), 1.45(s, 9H, C(CH_3)$_3$)

^{13}C NMR ($CDCl_3$) δ: 173.4(*C*OOCH$_3$), 165.0(Benzamide *C*O), 155.9(*C*OOC(CH$_3$)$_3$), 151.9(*A-C2*), 151.6(*A-C6*), 149.4(*A-C4*), 142.4(*A-C8*), 133.6, 132.4, 128.5 and 127.9(Arom), 122.6(*A-C5*), 79.1(*C*(CH$_3$)$_3$), 64.3(OCH$_3$), 59.0(*C5*), 53.6(Boc-NH-*C*H$_2$), 52.2(*C4*), 51.8(*C2*), 39.0(*C3*), 36.9(Boc-NH-CH$_2$-*C*H$_2$), 28.2((*C*H$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(N^6-benzoyladenin-9-yl)-2-(R)-proline methyl ester (**23b**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4-(*R*)-O-mesyl-2-(*R*)-proline methyl ester **20** (0.35g, 0.9mmol), N^6-benzoyladenine **4** (0.28g, 1.2mmol), anhydrous potassium carbonate (0.15g, 1.2mmol) and a catalytic amount of 18-crown-6 (0.1g, 0.3mmol) was stirred in dry DMF (4ml) at 75°C overnight. After completion of the reaction, the solvent was removed under vacuum. The residue was re-dissolved in water and extracted with dichloromethane (5x 5ml). The organic layer was then dried over anhydrous sodium

sulphate and concentrated to get crude product, which was purified by silica gel column flash chromatography to obtain the pure product **23b** (0.29g, 60%).

^{1}H NMR ($CDCl_3$) δ: 9.27(br s, 1H, Benzamide-N*H*), 8.76(s, 1H, *A*-*H*2), 8.10(s, 1H, *A*-*H*8), 8.00(d, 2H, Bz, *o*-C*H*), 7.56(t, 1H, Bz, *p*-C*H*), 7.45(t, 2H, Bz, *m*-C*H*), 5.27(m, 2H, *H*4, Boc-N*H*), 3.93(m, 1H, *H*2), 3.74(s, 3H, OCH_3), 3.60(m, 2H, *H*5, *H*5'), 3.20(m, 2H, *H*3, *H*3'), 2.83(m, 2H, Boc-NH-CH_2), 2.60(m, 2H, Boc-NH-CH_2-CH_2), 1.42(s, 9H, C(CH_3)$_3$)

^{13}C NMR ($CDCl_3$) δ: 172.9(*C*OOCH$_3$), 164.9(Benzamide *C*O), 155.9(*C*OOC(CH$_3$)$_3$), 152.0(*A*-*C*2), 151.6(*A*-*C*6), 149.6(*A*-*C*4), 141.7(*A*-*C*8), 133.6, 132.3, 128.4 and 127.8(Arom), 126.0(*A*-*C*5), 79.0(*C*(CH$_3$)$_3$), 70.4(Boc-NH-*C*H$_2$), 63.4(OCH$_3$), 56.7(*C*5), 53.2(*C*4), 51.8(*C*2), 39.0(*C*3), 35.6(Boc-NH-CH$_2$-*C*H$_2$), 28.2((*C*H$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(2-amino-6-chloropurin-9-yl)-2-(S)-proline methyl ester (**24a**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4-(*R*)-O-mesyl-2-(*S*)-proline methyl ester **19** (0.22g, 0.6mmol), 2-amino-6-chloropurine (0.12g, 0.7mmol), anhydrous potassium carbonate (0.1g, 0.7mmol) and a catalytic amount of 18-crown-6 (0.06g, 0.2mmol) was stirred in dry DMF (3ml) at 75^{o}C overnight. After completion of the reaction, the solvent was removed under vacuum. The residue was re-dissolved in water and extracted with dichloromethane (3x10ml). The organic layer was then dried over anhydrous sodium sulphate and concentrated to get crude product, which was purified by silica gel column chromatography to obtain pure product **24a** (0.08g, 30%).

^{1}H NMR ($CDCl_3$) δ: 8.35(s, 1H, 2-amino-6-chloro-purine, *H*8), 5.20(br s, 2H, NH_2), 5.10(m, 2H, *H*4, Boc-N*H*), 3.75(s, 3H, OCH_3), 3.30(m, 4H, *H*5, *H*5', Boc-NH-CH_2), 2.90(m, 3H, *H*2, Boc-NH-CH_2-CH_2), 2.70(m, 1H, *H*3), 2.15(m, 1H, *H*3'), 1.45(s, 9H, C(CH_3)$_3$)

^{13}C NMR ($CDCl_3$) δ: 173.2(*C*OOCH$_3$), 158.9(2-amino-6-chloro-purine, *C*6), 155.9(*C*OOC(CH$_3$)$_3$), 153.3(2-amino-6-chloro-purine, *C*2), 150.8(2-amino-6-chloro-purine, *C*4), 141.5(2-amino-6-chloro-purine, *C*8), 124.7(2-amino-6-chloro-purine, *C*5), 79.2(*C*(CH$_3$)$_3$), 64.4(OCH3), 58.9(*C*5), 53.5(Boc-NH-*C*H$_2$), 52.2(*C*4), 51.6(*C*2), 38.9(*C*3), 36.7(Boc-NH-CH$_2$-*C*H$_2$), 28.3((*C*H$_3$)$_3$)

*1-(N-Boc-aminoethyl)-4-(S)-(2-amino-6-chloropurin-9-yl)-2-(R)-proline methyl ester (**24b**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4(*R*)-O-mesyl-2(*R*)-proline methyl ester **20** (0.25g, 0.7mmol), 2-amino-6-chloropurine (0.14g, 0.8mmol), anhydrous K_2CO_3 (0.11g, 0.8mmol) and a catalytic amount of 18-crown-6 (0.07g, 0.2mmol) in dry DMF (3ml) was stirred at 70°C under nitrogen atmosphere, 48h. The solvent was removed under vacuum and the crude product, purified by flash column chromatography to get the pure product **24b** (0.11g, 37% yield) as a white foam.

^1H NMR ($CDCl_3$) δ: 7.85(s, 1H, 2-amino-6-chloro-purine, *H*8), 5.20(m, 4H, N^2*H*$_2$, *H*4, N*H*), 3.92(t, 1H, *H*2), 3.76(s, 3H, OC*H*$_3$), 3.60-3.00(br m, 4H, *H*5, *H*5', *H*3, *H*3'), 2.84(t, 2H, Boc-NH-C*H*$_2$), 2.56(t, 2H, Boc-NH-CH$_2$-C*H*$_2$), 1.46(s, 9H, C(C*H*$_3$)$_3$).

*1-(N-Boc-aminoethyl)-4-(S)-(N^2-isobutyrylguanin-7-yl)-2-(S)-proline methyl ester (**25a**)*

A mixture of 1-(*N*-Boc-aminoethyl)-4-(*R*)-O-mesyl-2-(*S*)-proline methyl ester **19** (0.15g, 0.4mmol), N^2-isobutyrylguanine **5** (0.20g, 0.9mmol) and anhydrous potassium carbonate (0.13g, 0.9mmol) in dry DMF (2ml) was kept stirring at 60°C overnight. The solvent was completely removed under vacuum and the pure product obtained (0.06g, 31%) after column chromatography.

^1H NMR ($CDCl_3$) δ: 12.15(br s, 1H, *G*-N1*H*), 9.30(br s, 1H, iBu-N*H*), 8.40(s, 1H, *G*-*H*8), 5.65(m, 1H, *H*4), 5.20(m, 1H, Boc-N*H*), 3.75(s, 3H, OC*H*$_3$), 3.30(m, 4H, *H*5, *H*5',

Boc-NH-C*H*$_2$), 2.80(m, 5H, Boc-NH-CH$_2$-C*H*$_2$, *H*2, (CH$_3$)$_2$C*H*, *H*3), 2.20(m, 1H, *H*3'), 1.45(s, 9H, C(C*H*$_3$)$_3$), 1.27(d, 6H, (C*H*$_3$)$_2$CH).

^{13}C NMR ($CDCl_3$) δ: 179.6 (*C*OCH(CH$_3$)$_2$), 173.1 (*C*OOCH$_3$), 156.4 (G-*C*2), 156.0 (G-*C*4), 153.3 (*C*OOC(CH$_3$)$_3$), 147.5 (G-*C*6), 142.3 (G-*C*8), 111.5 (G-*C*5), 79.2 (*C*(CH$_3$)$_3$), 64.7 (*C*H(CH$_3$)$_2$), 59.5 (*C*4), 57.9 (*C*5), 53.7 (Boc-NH-*C*H$_2$), 52.0 (*C*2), 37.6 (Boc-NH-CH$_2$-*C*H$_2$), 35.9 (*C*3), 28.4 ((*C*H$_3$)$_3$), 19.0 ((*C*H$_3$)$_2$).

The more polar minor product, 1-(*N*-Boc-aminoethyl)-4-(*S*)-(N^2-isobutyrylguanin-9-yl)-2-(*S*)-proline methyl ester (**26a**), was isolated in 10% yield (0.02g).

^{1}H NMR ($CDCl_3$) δ: 12.20(br s, 1H, *G*-N1*H*), 9.55(br s, 1H, iBu-N*H*), 8.10(s, 1H, *G*-*H*8), 5.40(m, 1H, Boc-N*H*), 5.00(m, 1H, *H*4), 3.75(s, 3H, OC*H*$_3$), 3.30(m, 4H, *H*5, *H*5', Boc-NH-C*H*$_2$), 2.80(m, 5H, Boc-NH-CH$_2$-C*H*$_2$, *H*2, (CH$_3$)$_2$C*H*, *H*3), 2.25(m, 1H, *H*3'), 1.45(s, 9H, C(C*H*$_3$)$_3$), 1.27(d, 6H, (C*H*$_3$)$_2$CH)

*1-(N-Boc-aminoethyl)-4(S)-(thymin-1-yl)-pyrrolidine-2(S/R)-carboxylic acid (**10**, **18**)*

To a solution of 1-(*N*-Boc-aminoethyl)-4(*S*)-(N3-benzoylthymin-1-yl)-2(*S*/*R*)-proline methyl ester **9**/**17** (0.42g, 0.84mmol) in methanol (2ml), was added aqueous 2N NaOH (2ml). TLC analysis after 10min. indicated the absence of the starting material as a result of hydrolysis of the methyl ester function. The reaction mixture was further stirred overnight, when TLC indicated the presence of a lower moving spot. The excess NaOH was neutralized by Dowex-50 H$^+$ resin, which was then, filtered off. The methanol from the filtrate was removed under vacuum and the residue was taken up in water. This was washed with ethyl acetate before concentrating it to dryness to obtain the product **10**/**18** (0.32g, quantitative yield) as a white solid foam.

*1-(N-Boc-aminoethyl)-4(S)-(N^4-benzyloxycarbonylcytosin-1-yl)-pyrrolidine-2(S/R)-carboxylic acid (**27**)*

An aqueous solution of 2N NaOH (1ml) was added to a solution of 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonylcytosin-1-yl)-2(*S*/*R*)-proline methyl ester **21** and stirred at room temperature for 10min, when complete hydrolysis of the ester was found to have occurred, as evident from the TLC analysis. The reaction was immediately neutralized using Dowex H^+ resin. The resin was filtered off and the filtrate, evaporated to get the product **27** (0.06g, quantitative yield) as a hygroscopic white solid.

*1-(N-Boc-aminoethyl)-4(S)-(N^6-benzoyladenin-9-yl)-pyrrolidine-2(S/R)-carboxylic acid (**28**)*

1-(*N*-Boc-aminoethyl)-4(*S*)-(N^6-benzoyladenin-9-yl)-2(*S*/*R*)-proline methyl ester **23** (0.05g, 0.1mmol) was dissolved in methanol (0.5ml). To this was added an aqueous solution of 2N NaOH (0.5ml). The reaction was stirred at room temperature for 10min, when all the starting material was consumed. The pH of the solution was brought down to 7.0 using Dowex H^+ resin, which was subsequently filtered off. The filtrate was concentrated to give the pure product **28** in good yield (0.04g, 82%).

*1-(N-Boc-aminoethyl)-4(S)-(guanin-9-yl)-pyrrolidine-2(S/R)-carboxylic acid (**29**)*

To a solution of 1-(*N*-Boc-aminoethyl)-4(*S*)-(2-amino-6-chloropurin-9-yl)-2(*S*/*R*)-proline methyl ester **24** (0.05g, 0.11mmol) in methanol (1ml), was added an aqueous solution of 2N NaOH (1ml). The solution was stirred at room temperature. The ester was found be hydrolyzed within 10min, while the oxidation of the 6-chloro function to the 6-oxo one was complete in 24h for the 2(*R*) isomer and 72h for the 2(*S*) isomer respectively. The solution was brought to neutral pH by the addition of Dowex H^+ resin,

which was then filtered off, and the filtrate, concentrated to get the desired product **29** in excellent yield (0.43g, 93%).

(2*R*,4*S*)

^{1}H NMR (Pyridine-D_5) δ: 8.13 (s, 1H, *G*-N*H*), 7.30 (s, 1H, *G*-*H*8), 5.50 (br, 3H, Boc-N*H*, *G*-N^{2}*H*$_2$), 5.25 (m, 1H, *H*4), 4.05 (m, 1H, *H*5), 3.50 (m, 3H, *H*5', Boc-NH-C*H*$_2$), 3.05 (m, 3H, *H*2, Boc-NH-CH$_2$-C*H*$_2$), 2.70 (m, 2H, *H*3, *H*3'), 1.45 (s, 9H, (C*H*$_3$)$_3$).

^{13}C NMR (Pyridine-D_5) δ: 175.4 (*C*OOH), 167.8 (*G*-*C*6), 160.5 (*G*-*C*2), 156.4 (*C*OOC(CH$_3$)$_3$), 154.6 (*G*-*C*4), 138.1 (*G*-*C*8), 115.7 (*G*-*C*5), 78.1 (*C*(CH$_3$)$_3$), 64.6 (*C*4), 57.3 (*C*5), 52.8 (Boc-NH-*C*H$_2$), 52.5 (*C*2), 39.5 (Boc-NH-CH$_2$-*C*H$_2$), 35.8 (*C*3), 28.3 ((*C*H$_3$)$_3$).

(2*S*, 4*S*)

^{1}H NMR (Pyridine-D_5) δ: 8.80 (s, 1H, *G*-N*H*), 7.85 (s, 1H, *G*-*H*8), 5.00 (br, 4H, Boc-N*H*, *G*-N^{2}*H*$_2$, *H*4), 3.35 (br m, 5H, *H*5, *H*5', *H*2, Boc-NH-C*H*$_2$), 2.55 (br m, 4H, Boc-NH-CH$_2$-C*H*$_2$, *H*3, *H*3'), 1.45 (s, 9H, (C*H*$_3$)$_3$).

*1-(N-Boc-aminoethyl)-4(S)-(N^{2}-isobutyryl-guanin-7-yl)-pyrrolidine-2(S)-carboxylic acid (**31**)*

The 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^{2}-isobutyrylguanin-7-yl)-2(*S*)-proline methyl ester **25a** (0.05g, 0.1mmol) was taken in methanol (1ml) and 2N aqueous NaOH (1ml) was added. Stirring was continued for 10min, after which, the excess alkali was neutralized with Dowex H^{+} resin, and the filtrate, washed with ethyl acetate before concentrating under vacuum to get the product **31** in good yield (0.04g, 82%).

*N1-(Boc)-1,2-diaminoethane (**33**)*

1,2-diaminoethane **32** (20g, 0.33mol) was taken in dioxane: water (1:1, 500ml) and cooled in an ice-bath. Boc-azide (5g, 35mmol) in dioxane (50ml) was slowly added with stirring and the pH was maintained at 10.0 by continuous addition of 4N NaOH. The

mixture was stirred for 8h and the resulting solution was concentrated to 100ml. The N1, N2-di-Boc derivative not being soluble in water, precipitated, and it was removed by filtration. The corresponding N1-mono-Boc derivative was obtained by repeated extraction from the filtrate in ethyl acetate. Removal of solvents yielded the mono-Boc-diaminoethane **33** (3.45g, 63%).

^{1}H NMR ($CDCl_3$) δ: 5.21 (br s, 1H, N*H*), 3.32 (t, 2H, J=8 Hz), 2.54 (t, 2H, J=8 Hz), 1.42 (s, 9H).

*Ethyl N-(2-Boc-aminoethyl)-glycinate (**34**)*

The *N1*-(Boc)-1,2-diaminoethane **33** (3.2g, 20mmol) was treated with ethylbromoacetate (2.25ml, 20mmol) in acetonitrile (100ml) in the presence of K_2CO_3 (2.4g, 20mmol) and the mixture was stirred at ambient temperature for 5h. The solid that separated was removed by filtration and the filtrate was evaporated to obtain the ethyl *N*-(2-Boc-aminoethyl)-glycinate **34** (4.3g, 83%) as a colourless oil.

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

*Ethyl N-(Boc-aminoethyl)-N-(chloroacetyl)-glycinate (**35**)*

The ethyl *N*-(2-Boc-aminoethyl)-glycinate **34** (4.0g, 14mmol) was taken in 10% aqueous Na_2CO_3 (75ml) and dioxane (60ml). Chloroacetyl chloride (6.5ml, 0.75mmol) was added in two portions with vigorous stirring. The reaction was complete within 5 min. The reaction mixture was brought to pH 8.0 by addition of 10% aqueous Na_2CO_3 and concentrated to remove the dioxane. The product was extracted from the aqueous layer with dichloromethane and was purified by column chromatography to obtain the ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **35** as a colourless oil in good yield (4.2g, 80%).

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

*N-(Boc-aminoethylglycyl)-thymine ethyl ester (**36**)*

Ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **35** (4.0g, 11.6mmol) was stirred with anhydrous K_2CO_3 (1.56g, 11.8mmol) in DMF with thymine (1.4g, 11.2mmol) to obtain the desired compound **36** 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 s, 1H, *T*-N*H*), 7.05 (min) & 6.98 (maj) (s, 1H, *T-H*6), 5.65 (maj) & 5.05 (min) (br s, 1H, N*H*), 4.58 (maj) & 4.44 (min) (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), 1.28 (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-(Boc-aminoethylglycyl)-(N^4-benzyloxycarbonyl cytosine)ethyl ester (**37**)*

A mixture of NaH (0.25g, 6.2mmol) and N^4-benzyloxycarbonyl cytosine **3** (1.24g, 6.2mmol) was taken in DMF and stirred at 75°C till the effervescence ceased. The mixture was cooled and ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **35** (2.0g, 6.2mmol) was added. Stirring was then continued at 75°C to obtain the cytosine monomer, *N*-(Boc-aminoethylglycyl)-(N^4-benzyloxycarbonyl cytosine)ethyl ester **37**, in moderate yield (1.62g, 50%).

^{1}H NMR ($CDCl_3$) δ: 7.65 (d, 1H, *C-H6*, J=8Hz), 7.35 (s, 5H, Ar), 7.25 (d, 1H, *C-H*5, J=8Hz), 5.70 (br s, 1H, N*H*), 5.20 (s, 2H, Ar-*C*H_2), 4.71 (maj) & 4.22 (min) (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-(Boc-aminoethylglycyl)-adenine ethyl ester (**38**)*

NaH (0.25g, 6.1mmol) was taken in DMF (15ml) and adenine (0.8g, 6.1mmol) was added. The mixture was stirred at 75°C till the effervescence ceased and the mixture

was cooled before adding ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **35** (2.0g, 6.1mmol). The reaction mixture was heated once again to 75°C for 1h, when TLC analysis indicated the disappearance of the starting ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate. The DMF was removed under vacuum and the resulting thick oil was taken in water and the product, extracted in ethyl acetate. The organic layer was then concentrated to obtain the crude product, which was purified by column chromatography to obtain the pure *N*-(Boc-aminoethylglycyl)-adenine ethyl ester **38**.

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

*N-(Boc-aminoethylglycyl)-2-amino-6-chloropurine ethyl ester (**39**)*

A mixture of 2-amino-6-chloropurine (1.14g, 6.8mmol), K_2CO_3 (0.93g, 7.0mmol) and ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **35** (2.4g, 7.0mmol) were taken in dry DMF (20ml) and stirred at room temperature for 4h. K_2CO_3 was removed by filtration, and the DMF, by evaporation under reduced pressure. The resulting residue was purified by column chromatography to obtain the *N*-(Boc-aminoethylglycyl)-2-amino-6-chloropurine ethyl ester (**39**) in excellent yield (2.65g, 98%).

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

*N-(Boc-aminoethylglycyl)-(N^2-isobutyryl-N7-guanine)-ethyl ester (**40**)*

A mixture of N^2-siobutyrylguanine **5** (0.6g, 3.1mmol), K_2CO_3 (0.43g, 3.1mmol) and ethyl *N*-(Boc-aminoethyl)-*N*-(chloroacetyl)-glycinate **35** (0.5g, 3.1mmol) were stirred in

dry DMF at room temperature overnight. The reaction mixture was filtered and the filtrate evaporated under reduced pressure. The resulting residue was purified by column chromatography to get the *N*-(Boc-aminoethylglycyl)-(N^2-isobutyryl-N7-guanine) ethyl ester **40** in 50% yield (0.39g).

^{1}H NMR ($CDCl_3$) δ: 12.25 (s, 1H, iBu-N*H*), 10.13 (s, 1H, G-N^1-*H*), 7.90 (s, 1H, G-*H*8), 5.85 (m, 1H, Boc-N*H*), 5.33 (maj) & 5.13 (min) (s, 2H, C*H*$_2$-G), 4.40 - 4.05 (m, 4H, N-C*H*$_2$-COOC*H*$_2$CH$_3$), 3.60 (m, 2H, Boc-NH-C*H*$_2$), 3.40 (maj) & 3.30 (min) (m, 2H, Boc-NH-CH$_2$-C*H*$_2$), 2.55 (m, 1H, C*H*(CH$_3$)$_2$), 1.45 (min) & 1.40 (maj) (s, 9H, (C*H*$_3$)$_3$), 1.25 (m, 9H, CH$_2$C*H*$_3$, CH(C*H*$_3$)$_2$).

^{13}C NMR ($CDCl_3$) δ: 180.2 (*C*OOCH(CH$_3$)$_2$), 169.4 (*C*OOEt), 167.2 (G-CH$_2$-*C*O), 162.9 (G-*C*2), 156.4 (G-*C*4), 153.3 (*C*OOC(CH$_3$)$_3$), 148.0 (G-*C*6), 145.2 (G-*C*8), 112.2 (G-*C*5), 79.6 (*C*(CH$_3$)$_3$), 61.5 (*C*H(CH$_3$)$_2$), 48.7 (N-*C*H$_2$-COOEt), 47.1 (OCH$_2$CH$_3$), 38.7 (Boc-NH-*C*H$_2$), 35.9 (Boc-NH-CH$_2$-*C*H$_2$), 29.6 ( ), 28.4 ((*C*H$_3$)$_3$), 18.9 ((*C*H$_3$)$_2$), 14.1 (CH$_2$*C*H$_3$).

Hydrolysis of the ethyl ester functions of PNA monomers

General method

The ethyl esters were hydrolyzed using 2N aqueous NaOH (5ml) in methanol (5ml) and the resulting acid was neutralized with activated Dowex-H^+ 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 (**41** - **45**) in excellent yield (>85%).

Picric Acid Estimation of Resin Functionalization

The typical procedure for estimation of the loading value of the resin was carried out with 5mg of the resin and comprised the following steps:

The resin was swollen in dry CH_2Cl_2 for at least 30min. The CH_2Cl_2 was drained off and a 50% solution of TFA in CH_2Cl_2 was added (1ml x 2), 15 min each. After washing

thoroughly with CH_2Cl_2, The TFA salt was neutralized with a 5% solution of DIPEA in CH_2Cl_2 (1ml x 3, 2 min each). The free amine was treated with a 0.1M picric acid solution in CH_2Cl_2 (2ml x 2, 3min each). The excess picric acid was eliminated by extensively washing the resin with CH_2Cl_2. The adsorbed picric acid was displaced form the resin by adding a solution of 5% DIPEA in CH_2Cl_2. The eluant was collected and the volume was made up to 10 ml with CH_2Cl_2 in a volumetric flask. The absorbance was recorded at 358nm in ethanol and the concentration of the amine groups on the resin was calculated using the molar extinction coefficient of picric acid as 14,500 $cm^{-1}M^{-1}$ at 358 nm.

Kaiser's Test

Kaiser's test was used to monitor the Boc-deprotection and amide coupling steps in the solid phase peptide synthesis. Three solutions were used, viz. (1) Ninhydrin (5.0 g) dissolved in ethanol (100 ml), (2) Phenol (80 g dissolved in ethanol (20 ml) and (3) KCN: 2 ml of a 0.001M aqueous solution of KCN in 98 ml pyridine).

To a few beads of the resin to be tested taken in a test tube, were added 3-4 drops of each of the three solutions desribed above. The tube was heated at 100 °C for ~5 min, and the colour of the beads was noted. A blue colour on the beads and in the solution indicated successful deprotection, while colourless beads were observed upon completion of the amide coupling reaction. The blank solution should remain yellow.

Cleavage of the PNA oligomers from the solid support

A typical cleavage reaction was carried out with 5 or 10mg of resin-bound PNA oligomer. The resin-bound PNA oligomer (10mg) was stirred in an ice-bath with thioanisole (20µl) and 1,2-ethanedithiol (8µl) for 10min, TFA (120µl) was added and stirring was continued for another 10min. TFMSA (16µl) was added and stirring continued for 2h. The reaction mixture was filtered through a sintered funnel. The

residue was washed with TFA (3 x 2ml) and the combined filtrate and washings were evaporated under vacuum and co-evaporated with ether, avoiding heating during this process. The residue was precipitated using dry ether and centrifuged to obtain a white pellet. The pellet was re-dissolved in methanol (~0.1ml) and re-precipitated by adding ether. The pellet collected after centrifugation was subjected to this re-precipitation process at least thrice, when a white precipitate was obtained of the crude PNA oligomer.

Gel Filtration

The crude PNA oligomer obtained after ether precipitation was dissolved in water (~0.5ml) and loaded on a gel filtration column. This column consisted of G25 Sephadex and had a void volume of 1ml. The oligomer was eluted with water and ten fractions of 1ml volume each were collected. The presence of the PNA oligomer was detected by measuring the absorbance at 260nm. The fractions containing the oligomer were freeze-dried. The purity of the cleaved crude PNA oligomer was determined by RP HPLC on a C18 column. If found to be above 90%, the oligomers were sued as such for experiments without further purification. If the purity was not satisfactory, the oligomers were purified by HPLC/FPLC.

FPLC

The crude PNA oligomers were dissolved in water containing 0.1% TFA, the starting buffer for injection. The polypyrimidine T8 sequences were purified using a gradient of 0 to 50% buffer B in 30 min at a flow rate of 1.0 ml/min, where buffer A = water with 0.1% TFA and buffer B = 60% CH3CN in water containing 0.1% TFA. The mixed sequence PNAs eluted earlier and hence had to be purified using a gradient of 0 to 30% B in 30 min at a flow rate of 1.0 ml/min, when good resolution of the peaks was

obtained. The purity of the oligomer after FPLC was ascertained by HPLC on a C18 RP column.

HPLC

The purity of the PNA oligomers was ascertained on an analytical RP C18 column using a gradient of 5 to 80% CH_3CN in water containing 0.1% TFA at a flow rate of 1.5 ml/min.

HPLC purification of the mixed sequence decamers was carried out on a semi-preparative RP C4 column using isocratic elution at a flow rate of 6.0 ml/min. The eluent was varied between 8 to 12% CH3CN in water containing 0.1% TFA to obtain optimum separation of the constituent peaks. The oligomers so collected were re-checked for purity by analytical HPLC as described above.

MALDI-TOF Mass Spectrometry

Literature reports the analysis of PNA purity by MALDI-TOF mass spectrometry[52] in which several matrices have been explored, *viz*. sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), CHCA (α-cyano-4-hydroxycinnamic acid) and DHB (2,5-dihydroxybenzoic acid). Of these, sinapinic acid was found to give the best signal to noise ratio with all the other matrices typically producing higher molecular ion signals.

For all the MALDI-TOF spectra recorded for the *aep*PNAs reported in this Chapter, sinapinic acid was used as the matrix and was found to give satisfactory results.

2.5. REFERENCES

1. Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543-584.
2. *Antisense Research and Applications* (Eds.: Crooke, S. T.; Lebleu, B.), CRC Press, Boca Raton, FL, **1993**.
3. *Antisense Oligonucleotides- Chemical Modifications.* Uhlmann, E.; Peyman, A. in *Encyclopaedia of Molecualr Biology and Biotechnology* (Ed.: Meyer, E.), VCH, New York, **1995**.
4. De Mesmaeker, A.; Haner, R.; Mertin, P.; Moser, H. E. *Acc. Chem. Res.* **1995**, *28*, 366.
5. *Methods in Molecular Biology, Vol. 20: Protocols for Oligonucleotides and Analogs* (Ed.: Agrawal, S.), Humana Press, New Jersey, Totowa, Chapter 16, **1993**, pp. 355-389.
6. Nielsen, P. E.; *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167.
7. De Mesmaeker, A.; Altmann, K. -H.;Waldner, A.; Wendeborn, S. *Curr. Opin. Struct. Biol.* **1995**, *5*, 343.
8. Stirchak, E. P.; Summerton, J. E.; Weller, D. D.; *Nucleic Acids Res. Dev. Proc. Symp.* **1989**, *17*, 6129.
9. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Science* **1991**, *254*, 1497.
10. Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioconj. Chem.* **1994**, *5*, 3.
11. Hyrup, B.; Egholm, M.; Nielsen, P. E. *Bioorg. Med.Chem. Lett.* **1996**, *6*, 5.
12. Good, L.; Nielsen, P. E. *Antisense & Nucleic Acids Drug Dev.* **1997**, *7*, 431.
13. Nielsen, P. E.; Haaima, G. *Chem. Soc. Rev.* **1997**, 73.
14. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 2796.
15. Demidov, V. V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Sönnichsen, S. H.; Nielsen, P. E. *Biochem, Pharmacol.* **1994**, *48*, 1310-1313.

16. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5.

17. Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A; Ricca, D. J.; Hassman, 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.

18. Kim. S. H.; Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

19. Koch, T.; Naesby, M.; Wittung, P.; Jørgensen, M.; Larsson, C.; Buchardt, O.; Stanley, C. J.; Nordén, B.; Nielsen, P. E.; Ørum, H. *Tetrahedron* **1995**, *36*, 6933.

20. Peterson, K. H.; Jensen, D. K.; Nielsen, P. E.; Egholm, M.; Buchardt, O. *Bioorg. Med. Chem. lett.* **1995**, *5*, 1119.

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

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

23. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Biochem. Biophys. Res. Commun.* **1997**, *240*, 778.

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

25. Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. I* **1997**, 539.

26. Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. I* **1997**, 547.

27. Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. I* **1997**, 555.

28. Lowe, G.; Vilaivan, T.; Westwell, M. S. *Bioorg. Chem.* **1997**, *25*, 321.

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

30. Linkletter, B. A.; Szabo, I. E.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 3888.

31. Ishihara, T.; Corey, D. R. *J. Am. Chem. Soc.* **1999**, *121*, 2012.

32. Hickman, D. T.; King, P. M.; Cooper, M. A.; Slater, J. M.; Micklefield, J. *Chem. Commun.* **2000**, 2251.

33. Püschl, A.; Tedeschi, T.; Nielsen, P. E. *Org. Lett.* **2000**, *2*, 4161.

34. Egholm, M.; Christensen, L.; Dueholm, L. D.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, *23*, 217.

35. Cruickshank, K. A.; Jiricny, J.; Reese, C. B. *Tetrahedron Lett.* **1984**, *25*, 681.

36. Brown, D. M.; Todd, A.; Varadarajan, S. *J. Chem. Soc.* **1956**, 2384.

37. Bullock, M. W.; Hand, J. J.; Stokstad, E. L. R. *J. Org. Chem.* **1957**, *22*, 568.

38. Jenny, T. F.; Schneider, K. C.; Benner, S. A. *Nucleosides Nucleotides* **1992**, *11*, 1257.

39. Robinson, D. S.; Greenstein, J. P. *J. Biol. Chem.* **1952**, *195*, 383.

40. Buchardt, O.; Egholm, M.; Nielsen, P. E.; Berg, R. H. *Int. PCT Appl. wo 92/20702* **1992**.

41. Ando, T.; Yamawaki, J. *Chem. Lett.* **1979**, 45.

42. Chloroacetyl chloride was prepared from chloroacetic acid and thionyl chloride, by refluxing and then distillation.

43. 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.

44. 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.

45. Gisin, B. F. *Helv. Chim. Acta* **1970**, *56*, 1476.

46. Merrifield, R. B.; Stewart, J. M.; Jernberg, N. *Anal. Chem.* **1966**, *38*, 1905.

47. Erickson, B. W.; Merrifield, R. B. **1976** in *Solid Phase Peptide Synthesis. In the Proteins* Vol. II, 3rd ed.; Neurath, H. and Hill, R. L. eds., Academic Press, New York, pp 255.

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

49. Fields, G. B.; Fields, C. G. *J. Am. Chem. Soc.* **1991**, *113*, 4202.

50. Barawkar, D. A.; Rajeev, K. G.; Kumar, V. A.; Ganesh, K. N. *Nucleic Acids Res.* **1996**, *24*, 1229.

51. 27 Perrin, D. D.; Amarego, W. L. F. in *Purification of Laboratory Chemicals, 3rd edition,* **1989**, Pergamon Press..

52. Bulter, J. M.; Jiang-Baucom, P.; Huang, M.; Belgrader, P.; Girard, J. *Anal. Chem.* **1996**, *68*, 3283.

2.6. Appendix

- Compound **8**, ^{13}C and ^{1}H NMR
- Compound **9**, ^{1}H NMR
- Compound **10**, ^{1}H and ^{13}C NMR
- Compound **16**, ^{1}H and ^{13}C NMR
- Compound **17**, ^{13}C NMR
- Compound **18**, ^{1}H and ^{13}C NMR
- Compound **19**, ^{1}H NMR
- Compound **20**, ^{1}H NMR
- Compound **21a**, ^{1}H and ^{13}C NMR
- Compound **21b**, ^{1}H NMR
- Compound **23a**, ^{1}H and ^{13}C NMR
- Compound **23b**, ^{1}H NMR
- Compound **24a**, ^{1}H and ^{13}C NMR
- Compound **24b**, ^{1}H and ^{13}C NMR
- Compound **25a**, ^{1}H NMR
- Compound **30a**, ^{1}H NMR
- Compound **30b**, ^{1}H NMR
- Compound **40**, ^{1}H NMR
- Compounds **24a** and **23a**, FAB-MS
- Compounds **21a**, FAB-MS

CHAPTER 3

BIOPHYSICAL STUDIES OF 1-(*N*-Boc-aminoethyl)-4(*S*)-(purinyl/pyrimidinyl)-2(*S*/*R*)-proline-CONTAINING OLIGOMERS

SECTION A

3.1. INTRODUCTION

The purpose of designing antisense/antigene molecules is to achieve (i) high binding affinity and specificity to the target nucleic acid, (ii) stability to cellular enzymes, (iii) a long-enough half-life within the system to allow manifestation of its effect, and (iv) non-toxicity to the system in which its activity is desired. All newly designed and synthesized molecules are hence subjected to various biophysical and biochemical studies to evaluate their potential as antisense/antigene agents. In the recent past, PNA[1-4] has emerged as a potential antisense agent as it possesses many of the desirable qualities like high binding affinity and specificity to target sequences and stability to cellular enzymes. However, the major hurdle to its applicability is posed by its low aqueous solubility due to intermolecular self-aggregation and intramolecular self-organization. In addition, PNA is poor in diiferentiating between the parallel and antiparallel binding modes perhaps due to its achiral nature, leading to undesirable non-specific binding in cellular systems.

3.2. RATIONALE AND OBJECTIVES OF THE PRESENT WORK

The preceding Chapter reports the synthesis of chiral, conformationally constrained PNA analogues, the aminoethylprolyl (*aep*) PNA. These *aep* PNA units were introduced into achiral aminoethylglycyl (*aeg*) PNA at various positions in order to study the effect of structural constraint and chirality in influencing the binding properties to target nucleic acids in terms of strength, specificity and directionality of binding. The positive charge resulting on the tertiary pyrrolidine ring nitrogen atom at physiological pH presents an added advantage of increasing the solubility of the oligomers in aqueous media.

The objective of this Chapter is to carry out a systematic study of the chiral, cationic aminoethylprolyl PNA backbone in terms of its ability to form specific base-paired complexes (both, duplexes and triplexes) with target nucleic acids. In order to explore the triplex forming potential, polypyrimidine sequences were constructed and to enable the study of duplexes, mixed purine-pyrimidine sequences were employed with mixed *aeg/aep* backbones. This Chapter addresses the biophysical studies of *aep*PNAs and their hybrids with complementary nucleic acids using various techniques such as UV spectroscopy, CD spectrophotometry, fluorescence and gel mobility shift assays.

3.3. PRESENT WORK

The PNA oligomers employed for the biophysical studies presented in this Section are listed in Tables 1, 2 and 3.

Table 1. Polypyrimidine and Complementary Polypurine PNA Oligomer Sequences

PNA	Sequence Composition	
46	H- TTTTTTTT -NH-$(CH_2)_2$-COOH	*aeg*PNA control
47	H- **t** TTTTTTT -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*) *aep*PNA unit at N. T.
48	H- **t** TTTTTTT -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*) *aep*PNA unit at N. T.
49	H- TTTTTTT **t** -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*) *aep*PNA unit at C. T.
50	H- TTTTTTT **t** -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*) *aep*PNA unit at N. T.
51	H- TTT **t** TTT **t** -NH-$(CH_2)_2$-COOH	two (2*S*,4*S*) *aep*PNA units
52	H- TTT **t** TTT **t** -NH-$(CH_2)_2$-COOH	two (2*R*,4*S*) *aep*PNA units
53	H- T **t** T **t** T **t** T **t** -NH-$(CH_2)_2$-COOH	alternating *aeg*-(2*S*,4*S*) *aep*PNA units
54	H- T **t** T **t** T **t** T **t** -NH-$(CH_2)_2$-COOH	alternating *aeg*-(2*R*,4*S*) *aep*PNA units
55	H- **t t t t t t t t** -NH-$(CH_2)_2$-COOH	(2*S*,4*S*) *aep*PNA homooligomer
56	H- **t t t t t t t t** -NH-$(CH_2)_2$-COOH	(2*R*,4*S*) *aep*PNA homooligomer
63	H- AAAAAAAA -NH-$(CH_2)_2$-COOH	complementary *aeg*PNA

t = (2*S*,4*S*) *aep*PNA-T; **t** = (2*R*,4*S*) *aep*PNA-T; T = *aeg*PNA-T; A = *aeg*PNA-A

Table 2. Mixed Base PNA Comprising A and T Nucleobases

PNA	Sequence Composition	
64	H- TATATTATTATT -NH-$(CH_2)_2$-COOH	*aeg*PNA (control)
65	H- **t** ATATTATTATT -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*)*aep*PNA-T unit at N. T.
66	H- **t** ATATTATTATT -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*)*aep*PNA-T unit at N. T.
67	H-**t a t a t t a t t a t t** -NH-$(CH_2)_2$-COOH	(2*R*,4*S*) *aep*PNA homooligomer

t = (2*S*,4*S*) *aep*PNA-T; **t**/**a** = (2*R*,4*S*) *aep*PNA-T/A; T/A = *aeg*PNA-T/A

Table 3. Mixed Base PNA Sequences Comprising A, T, G & C Nucleobases

PNA	*Sequence Composition*	
68	H- GTAGATCACT -NH-$(CH_2)_2$-COOH	*aeg*PNA (control)
69	H- GTAGA **t** CACT -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*) *aep*PNA-T unit
70	H- GTAGA **t** CACT -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*) *aep*PNA-T unit
71	H- G **t** AGA **t** CAC **t** -NH-$(CH_2)_2$-COOH	three (2*S*,4*S*) *aep*PNA-T units
72	H- G **t** AGA **t** CAC **t** -NH-$(CH_2)_2$-COOH	three (2*R*,4*S*) *aep*PNA-T units
73	H- GT **a** GATCACT -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*) *aep*PNA-A unit
74	H- GT **a** GATCACT -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*) *aep*PNA-A unit
75	H- GTA **g** ATCACT -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*) *aep*PNA-G unit
76	H- GTA **g** ATCACT -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*) *aep*PNA-G unit
77	H- GTAGAT **c** ACT -NH-$(CH_2)_2$-COOH	one (2*S*,4*S*) *aep*PNA-C unit
78	H- GTAGAT **c** ACT -NH-$(CH_2)_2$-COOH	one (2*R*,4*S*) *aep*PNA-C unit

The DNA oligomers used in these studies are listed in Table 4.

Table 4. DNA Oligonucleotide Sequences

DNA	Oligomer Sequences 5' ® 3'	
For the Homopyrimidine PNA Sequences		
84	G C A A A A A A A A C G	complementary to PNA-T_8 **46-56** with CG clamps
85	T T T T T T T T	for DNA:DNA complex formation with DNA **84**
86	G C A A A T A A A A C G	mismatch DNA for PNAs **46-56**
For Mixed Base PNA Sequences Comprising A & T Bases		
87	A A T A A T A A T A T A	antiparallel DNA to PNAs **64-67**
88	A T A T A A T A A T A A	parallel DNA to PNAs **64-67**
For Mixed Base PNA Sequences Comprising A, T, G & C Bases		
92	A G T G A T C T A C	antiparallel DNA to PNAs **68-78**
93	A G T G A T C C A C	mismatch in DNA opposite A in PNAs **73** & **74**
94	A G T G A T A T A C	mismatch in DNA opposite G in PNAs **75** & **76**
95	A G T G T T C T A C	mismatch in DNA opposite T in PNAs **69-72**
96	A G T T A C C T A C	mismatch in DNA opposite C in PNAs **77** & **78**
97	C A T C T A G T G A	parallel DNA to PNAs **68-78**
98	G T A G A T C A C T	for DNA:DNA duplex formation with DNA **92**

3.4. BIOPHYSICAL SPECTROSCOPIC TECHNIQUES FOR STUDYING PNA-DNA INTERACTIONS

3.4.1. UV-Studies

Monitoring the UV absorption at 260nm as a function of temperature has been extensively used to study the thermal stability of nucleic acid systems and consequently, PNA:DNA hybrids as well. Increasing the temperature perturbs this system, inducing a structural transition by causing disruption of hydrogen bonds between the base-pairs, diminished stacking between adjacent nucleobases, and larger torsional motions in the backbone leading to a loss of secondary and tertiary structure. This is evidenced by an increase in the UV absorption at 260nm, termed as

'hyperchromicity'. The magnitude of hyperchromicity is a measure of the extent of the secondary structure present in nucleic acids. The process is co-operative and the plot of the absorbance at 260nm Vs the temperature is sigmoidal (Figure 1a). This also represents a two-state "all or none" model for nucleic acid melting, i.e., the nucleic acids exist in only two states, either as duplexes or as single strands and at varying temperatures, the relative proportions change. A non-sigmoidal (e.g., linear) transition with low hyperchromicity is a consequence of non-duplexation (non-complementation). In many cases, the transitions are broad and the exact Tms are obtained from the peak in the first derivative plots. This technique has provided valuable information regarding complementary interactions in nucleic acid hybrids involving DNA, RNA and PNA.[5]

The binding stoichiometry of nucleic acids can be determined from UV-mixing or UV-titration experiments. The UV-mixing experiments are carried out by mixing the appropriate oligomers in different mole ratios, keeping the total concentration constant. The UV-absorbance of these samples is plotted as a function of the mole fraction of one of the components, in what is termed as a Job's plot[6] (Figure 1b). The absorbance steadily decreases until all the strands present are involved in complex formation as a

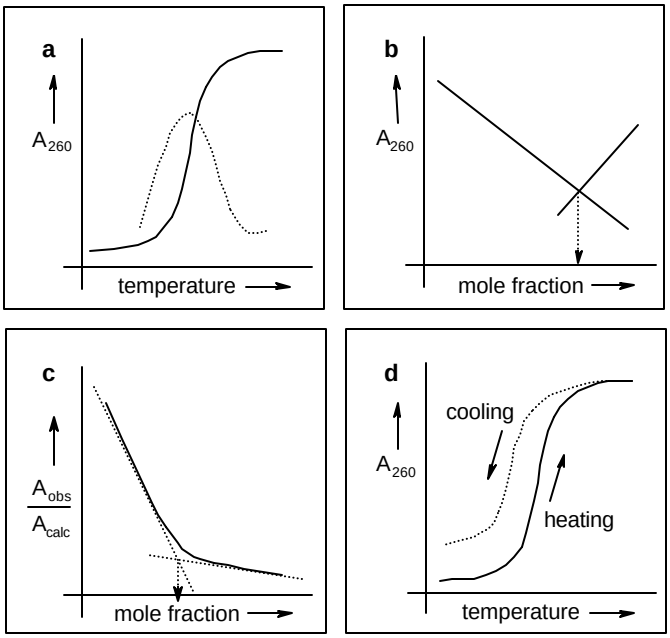


Figure 1. Schematic representation of **a**. UV-melting (thermal stability), **b**. UV-mixing, **c**. UV-titration (stoichiometry) and **d**. Hysteresis (rate of hybridization)

result of the hypochromic effect, and then rises afterwards when one strand is present in excess. The stoichiometry of the complexation is derived from the minimum in such a plot.

The stoichiometry of complexation can also be determined by UV-titration (Figure 1c). In this method, one of the strand involved in complexation Is sequentially added in aliquots to a fixed amount of the complementary component and the UV-absorbance is recorded at each addition. Upon successive addition of the complementary strand, complex formation results in hypochromicity which leads to a progressive decrease in the ratio of the observed to the calculated absorbance. After the first strand present in the buffer is exhausted, the absorbance reaches a plateau in the plot of the ratio of the observed to the calculated absorbance against the nucleic acid mole fraction and the point at which the plateau is reached indicates the stoichiometry of complexation.

PNA/DNA strands bearing charged groups can be tested for hysteresis by thermal dissociation Vs re-association plots. The experiment consists of recording the UV absorbance by first heating the duplexes/triplexes (UV-melting) followed by cooling the sample while recording the absorbance (re-association, cooling curve). In DNA:DNA complexes, the cooling curve does not follow the melting curve and exhibits a hysteresis (Figure 1d). This is due to the fact that the re-association of duplexes or triplexes is much slower than melting due to the interstrand repulsion on account of the negative phosphate groups. When one of the strands bears cationic charges, the net repulsion between the two strands is reduced, leading to a lower hysteresis in the heating-cooling plots.

The fidelity of base-pairing in the PNA:DNA complexes can be examined by challenging the PNA oligomer with a DNA strand bearing a mismatch at a desired site, preferably opposite the site of modification. The base mismatch leads to the absence of or incorrect hydrogen bonding between the bases and causes a drop in the measured melting temperature. A modification of the PNA structure is considered good if it gives

a significantly lower Tm with DNA sequences containing mismatches as compared to that with unmodified PNA. It is to be pointed out that in all biophysical experiments described herein, the modified PNAs are always evaluated against the unmodified control PNA.

Homopyrimidine thymine PNA sequences bind to the complementary homopurine DNA sequence forming PNA_2:DNA triplexes in which it is difficult to distinguish the PNA strand that binds to the central DNA strand by WC hydrogen bonding from that which binds by HG hydrogen bonding. Mixed base sequences form duplexes of antiparallel or parallel orientations that can be selected by proper design of the complementary DNA sequences. By convention, antiparallel PNA:DNA complexes are defined as those in which the 'N' terminal of the PNA faces the 3'-end of the DNA with the 'C' terminal facing the 5'-end and parallel PNA:DNA complexes are those in which the 'C' terminal of PNA faces the 3'-end of DNA with the 'N' terminal towards the 5'-end of the DNA[7] (Figure 2).

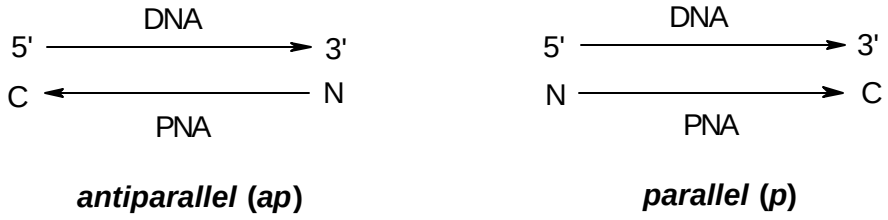


Figure 2. Schematic representation of the antiparallel and parallel modes of complexation of PNA with complementary DNA

3.4.2. Circular Dichroism

Circular Dichroism (CD) is a well-established tool used to study the conformational aspects of nucleic acids.[8,9] Upon comparison with reference samples, CD spectra can provide reliable and useful data concerning the conformational states of the system under study. However, CD does not give detailed structural data as obtained from X-ray crystallography or NMR, but it can be used as a complementary tool to UV

spectroscopy to evaluate the overall base-stacking patterns. The differences in secondary structure and handedness of helices can be differentiated conformationally as changes in CD profiles.

CD of nucleic acids arises predominantly as an effect of coupling between the transition moments of adjacent nucleobases due to continuous stacking. The PNA backbone is inherently achiral. However, PNA, a polyamide, can be expected to form helices *via* intramolecular hydrogen bonding leading to a racemic mixture of right- and left-handed helices and no net CD is observed.[10] Upon complexation with DNA/RNA, which is a chiral molecule, PNA:DNA/RNA duplexes/triplexes exhibit strong CD signals. Thus, the complex formed as a consequence of the binding of achiral PNA and chiral DNA leads to the formation of a chiral complex. CD thus, assumes importance in the characterization of such complexes.

3.5. RESULTS

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

3.5.1. Homopyrimidine PNA-T_8 Sequences: UV Studies

3.5.1a PNA:DNA Binding Stoichiometry. The UV-titration (Figure 3) experiments involving the stoichiometric addition of *aeg*PNA **46** or *aep*PNA **55** to DNA **84** led to a decrease in UV absorbance with both control and *aep* PNAs. The absorbance showed a similar saturation plateau around 2:1 stoichiometry suggesting the formation of a PNA_2:DNA triplex in both cases.

3.5.1b PNA_2:DNA Triplexes: UV-Tm Studies. The T_8 single strands (**46**- **56**) showed very little (<2%) change in absorbance at 260nm upon heating from 5 to 85 ^{o}C

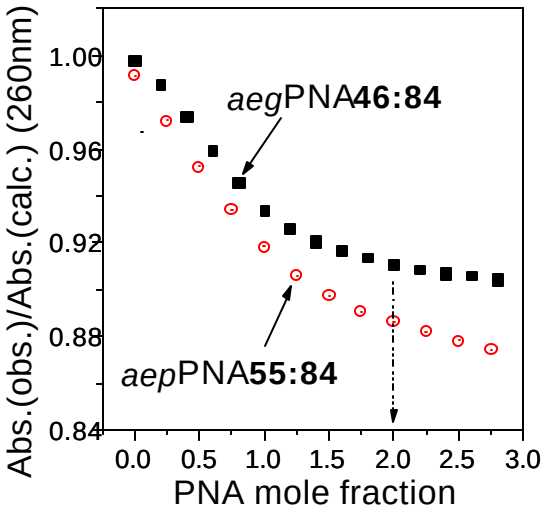


Figure 3. UV-titration of PNA:DNA complexes **46:84** and **55:84**

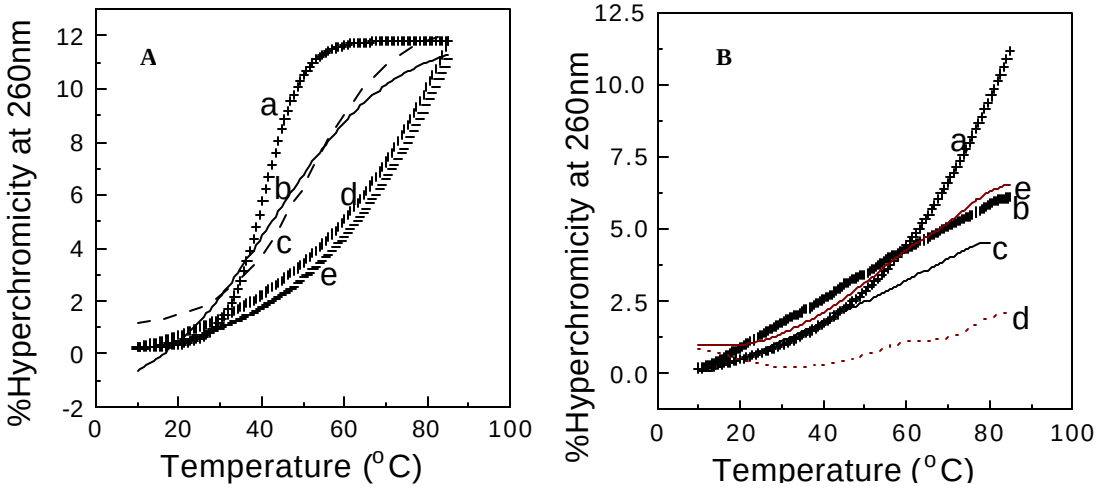


Figure 4. UV-Tm profiles of **A.** (2*S*,4*S*) *aep*PNA$_2$:DNA **84** complexes with PNA containing one (b), two (c), four (d), eight (e) or no (a, control) *aep*PNA units; **B.** *aep*PNA **55**:DNA **84** (a), mismatch complex **55**:**86** (b), single strand DNA **86** (c), single strand *aep*PNA **55** (d) and the addition spectrum of single strands c + d (e).

due to self-melting. The A_8 complementary DNA (**84**) showed a slightly higher increase in absorbance at 260nm (~3%). The shape of these plots was however, linear and not sigmoidal (Figure 4a).

Table 5 shows the Tm values for PNA$_2$:DNA complexes derived from various *aeg*PNA and *aep*PNA sequences of different stereochemistry and degree of modification. Generally, the UV-Tm values obtained for PNA:DNA complexes in 1:1 &

2:1 stoichiometry were almost identical and the plots of absorbance or % hyperchromicity at 260nm Vs temperature were sigmoidal, indicative of a two-state co-operative transition (Figure 4b). It is seen from Table 5 that changing the stereochemistry at C2 did not appreciably alter the stability of the various *aep*PNA_2:DNA complexes and the difference in Tm values with the (2S, 4S) and (2R, 4S) *aep*PNA units was only ~1-2 ^{o}C. A single *aep*PNA unit caused a destabilization of PNA_2:DNA complexes by ~5 ^{o}C when present at the 'N' terminus of the oligomer (Table 5, entry 1), while it had a negligible effect (ΔTm ~1-2 ^{o}C; Table 5, entry 2) when present at the 'C' terminus. Increasing the number of *aep*PNA units led to a progressive increase in the stability of the derived PNA_2:DNA complexes (Table 5, entries 3-5) with a stability increase of ~5 to 8 ^{o}C per *aep*PNA unit. The UV-melting plots of the PNA_2:DNA complexes with one and two modifications were sigmoidal (Figure 4A b,c) and those involving PNA with four or more *aep*PNA units failed to reach the upper plateau of the sigmoidal transition (Figure 4A d,e). Hence, the Tm of these complexes could not be determined (Table 5, entries 4,5). The Tm of the individual complexes of *aep*PNAs **53**- **56** with DNA **84** was more than 75 ^{o}C.

Table 5. UV-Tm (^{o}C) of PNA_2:DNA complexes*

	(2*S*,4*S*)	**(2*R*,4*S*)**	
Entry	**Tm**	**PNA_2:DNA**	**Tm**
1	38	**48:84**	37
2	44	**50:84**	42 (38)
3	51	**52:84**	54
4	>75	**54:84**	>75 (67)
5	>80	**54:86**	>80 (>74)
6	nd	**52:86**	nd
7	nd	**54:86**	nd
8	**46:84**	43 (38)	

*nd: not detected. Buffer: 10mM sodium phosphate, pH 7.4. Values in brackets: T_m(°C) with 100 mM NaCl. T_m values are accurate to (±)0.5 °C. Experiments were repeated at least thrice and the T_m values were obtained from the peaks in the first derivative plots.

3.5.1c Mismatch studies. The complexes of PNAs were constituted with DNA containing a mismatch base (Figure 5). The PNA:DNA complexes comprising the *aep*PNAs (**53**-**56**) and DNA **86** having a single mismatch gave linear, non-sigmoidal plots (Figure 3B b) and failed to show any peak in the first derivative plots. As a consequence, no melting temperature was detected for these complexes.

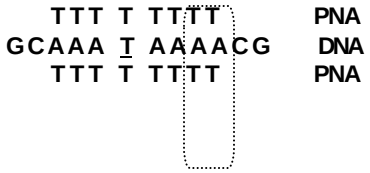


Figure 5. Mismatched PNA_2:DNA complex

3.5.1d Salt effects. The presence of salt (100mM NaCl) in the medium is known to destabilize the PNA_2:DNA complexes.[11] The DNA **84** complexes with control PNA **46** (entry 8) and PNA with one *aep*PNA unit, **50** (entry 2) exhibited ΔTm = ~ -5 ^{o}C upon salt addition. Higher substituted PNA_2:DNA complexes **54:84** with four *aep*PNA units; Table 5, entry 4 and **56:84** with eight *aep*PNA units; Table 5, entry 5 complexes were destabilized to a greater extent (ΔTm $\geq$ -8 ^{o}C) compared to the control (**46:84**) or monosubstituted *aep*PNADNA (**50:84**) complexes (Table 5, entries 8 & 2 respectively).

3.5.1e pH effects. The effect of pH on the stability of the *aep*PNA_2:DNA complexes was studied by constituting the hybrids of PNAs **46**, **48**, **51** & **52** with DNA **84** in sodium phosphate buffers at pH values ranging from 6.0 to 7.4. No significant change in Tm was observed in this pH range (ΔTm = $\pm$1 ^{o}C; Table 6).

Table 6. UV-Tm of PNA:DNA complexes in buffers of varying pH values

PNA_2:DNA complex	**Tm (°C)**		
	pH = 6.0	*pH = 6.8*	*pH = 7.4*
46:84	43.5	42.5	43.0
48:84	36.5	37.5	37.0
50:84	42.0	42.0	42.0
52:84	70.5	74.0	73.0

3.5.2. UV-Tm Studies in Duplexes

The oligothymine sequences described above were found to form triplexes In which the binding orientation (parallel-antiparallel) of the two PNA strands involved in complex formation remains inconsequential. In order to study the orientational preferences (parallel/antiparallel, Figure 2) of PNA:DNA binding, mixed purine-pyrimidine sequences were synthesized.

Table 7. UV-Tm (^{o}C) of PNA:DNA duplexes containing the nucleobases A & T

	(2*S*,4*S*)	**(2*R*,4*S*)**	
Entry	**Tm (^{o}C)**	**PNA:DNA**	**Tm (^{o}C)**
1	43.0	**66:87** (*ap*)	42.0
2	42.0	**66:88** (*p*)	44.0
Control *aeg*PNA_2:DNA complex			
3	**64:87**	43.5	
4	**64:88**	41.5	

Buffer: 10mM sodium phosphate, pH 7.4. Values in brackets: T_m(°C) with 100 mM NaCl. T_m values are accurate to (±)0.5 °C. Experiments were repeated at least thrice and the T_m values were obtained from the peaks in the first derivative plots.

The AT-rich mixed PNA sequences such as **64**-**67** have been demonstrated[12] to form duplexes with the complementary DNAs **87** (antiparallel, *ap*) and **88** (parallel, *p*). Hence, Tm studies of PNA:DNA duplexes of these oligomers containing *aep*PNA units were carried out. The duplexes were constituted by individually mixing equimolar

amounts of complementary achiral *aeg*PNA **64** and chiral *aep*PNAs (**65**, **66**) with DNA oligomers **87**/**88** in phosphate buffer.

The oligomers **65** & **66** bearing one (2*S*,4*S*) & (2*R*,*4S*) *aep*PNA-T unit respectively at their 'N' termini exhibited similar melting behaviour to the control *aeg*PNA **64** with DNA **87** (*ap*) (Table 7, entries 1 & 3). The parallel complex formed by the achiral PNA **64** with the complementary DNA **88** was 2 °C less stable than the antiparallel complex formed with DNA **87**. Similarly, the parallel complex formed by (2*S*, 4*S*) *aep*PNA **65:88** was 1 °C less stable than the corresponding antiparallel complex (**65**:**87**). However, PNA **66** bearing the (2*R*, 4*S*) *aep*PNA exhibited a slightly higher Tm for the parallel complex (with DNA **88**) than the antiparallel one (**66**:**87**) (ΔTm = 2 °C). Some select UV-Tm plots for parallel and antiparallel pairs are shown in Figure 6.

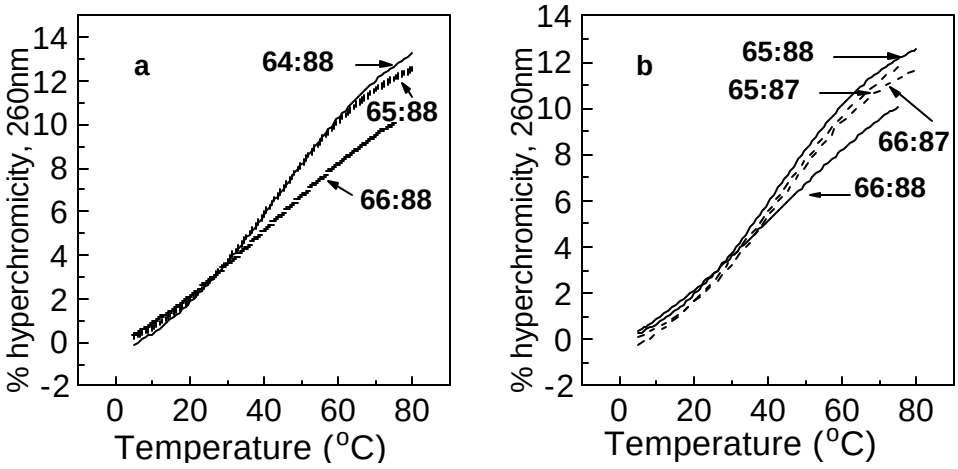


Figure 6. UV-Tm profiles of AT-rich PNA:DNA duplexes. **a**. Antiparallel complexes **b**. Parallel and antiparallel duplexes.

It is reported in the literature that the mixed base *aeg*PNA sequences containing all the four nucleobases, *viz*. A, T, G and C form 1:1 PNA:DNA duplexes.[13] Therefore, for all the melting experiments of the duplexes derived from *aeg*/*aep* PNA oligomers **68**-**78** with DNA **92** (*ap*) or **97** (*p*), the constituent strands were mixed in equimolar concentrations.

The results obtained in the UV-melting experiments are summarized in Tables 8 and 9. The following observations emerge from the data:

Table 8. UV-Tm data for the (2*S*,4*S*) *aep*PNA mixed base PNA:DNA duplexes

Entry	PNA	DNA 92 (*ap*)	DNA 97 (*p*)
		Tm (°C)	Tm (°C)
1		43.8	40.3
2		53.8	50.2
3		55.2	31.2
4		56.6	34.0
5		43.2	58.3

Table 9. UV-Tm data for the (2*R*,4*S*) *aep*PNA mixed base PNA:DNA duplexes

Entry	PNA	DNA 92 (*ap*)	DNA 97 (*p*)
		Tm (°C)	Tm (°C)
1		43.8	40.3
2		53.1	50.2
3		62.3	34.0
4		33.3	26.2
5		53.1	28.3

The mixed sequence control PNA **68** forms both parallel (**68:97**) and antiparallel (**68:92**) duplexes with DNA and the antiparallel duplex is only slightly stable by 3-5 °C over the parallel duplex (Table 8 & 9, entry 1). A similar behaviour of higher antiparallel duplex stability over the parallel duplex is noticed in the case of *aep*PNAs-A **73** (Table 8, entry 2) and **74** (Table 9, entry 2), with ΔTm = 3.5 °C. Significantly, both antiparallel (**73:92** & **74:92**) and parallel duplexes (**73:97** & **74:97**) containing the *aep*PNA-A units

were stabilized by ~10 °C compared to the duplexes with *aeg*PNA **68:92** & **68:97** respectively.

The antiparallel duplexes involving *aep*PNA-C units (**77:92** & **78:92**, entry 3 in Tables 8 & 9 respectively) exhibited considerable stabilization (ΔTm = 12 – 18 °C) over the control **68:92** while the parallel duplexes (**77:97** & **78:97**) were destabilized (ΔTm = -6 to -9 °C) compared to the control (entries 3 & 1 in Tables 8 & 9 respectively). The difference between the stabilities of antiparallel and parallel duplexes was therefore much higher ($\Delta Tm_{ap\text{-}p}$ = 24 – 28 °C) than that seen for the control ($\Delta Tm_{(ap\text{-}p)}$ = 3.5 °C). The effect of stereochemistry at C2 was not very prominent in this case and the *aep*PNA oligomeric duplexes bearing the 2*R* monomer were more stable than those with 2*S* (ΔTm = 3-6 °C)

In *aep*PNA-T and *aep*PNA-G-containing oligomeric duplexes, the differences between the two stereochemistries were evident. In case of the (2*S*,4*S*) *aep*PNA-T oligomer **69**, as in case of *aep*PNA-A and *aep*PNA-C, the antiparallel duplex **69:92** was stabilized by 12.8 °C over the control **68**:**92** (Table 8, entries 4 & 1). The corresponding (2*R*,4*S*) isomer, unlike any other *aep*PNA unit, destabilized the antiparallel duplex **70**:**92** by 10.5 °C (entry 4, Table 9) compared to the *aeg*PNA **68**:**92**. Both (2*S*,4*S* and 2*R*,4*S*) isomers destabilized the parallel duplexes **69:97** & **70:97** (entry 4 in Tables 8 & 9 respectively) compared to the control **68**:**97**, the effect being more pronounced in the (2*R*,4*S*) isomer **70** (ΔTm = -14 °C) than the (2*S*,4*S*) isomer **69** (ΔTm = -6 °C). The (2*S*,4*S*) *aep*PNA-T unit was thus able to differentiate between the antiparallel and parallel duplexes ($\Delta Tm_{ap\text{-}p}$ = 22 °C) more than the (2*R*,4*S*) isomer ($\Delta Tm_{ap\text{-}p}$ = 7 °C).

The (2*S*,4*S*) *aep*PNA-G unit was exceptional in stabilizing the parallel duplex **75:97** (ΔTm = 18 °C) compared to the antiparallel duplex **75:92** (Table 8, entry 5; Tm similar to control), unlike any other *aep*PNA unit. The (2*R*,4*S*) isomer stabilized the antiparallel duplex **76**:**92** (ΔTm = 9.3 °C) and destabilized the parallel duplex **76**:**97** (ΔTm = -12 °C;

Table 9, entry 5) compared to the controls **68:92** & **68:97** (Table 9, entry 1) respectively. Both isomers were thus, able to differentiate between antiparallel and parallel modes of binding. The (2*S*,4*S*) isomer favoured parallel duplexes **75:97** over antiparallel **75:92** ($\Delta Tm_{p\text{-}ap}$ = +15 °C), while the (2*R*,4*S*) isomer favoured the antiparallel **76:92** duplex over the parallel one **76:97** ($\Delta Tm_{ap\text{-}p}$ = +25 °C).

The data overall indicate base-dependent stability effects, in turn modulated by orientational and stereochemical factors. The most interesting results were obtained with the *aep*PNA-G units. The UV-melting profiles of the *aep*PNA:DNA duplexes are shown in Figure 7.

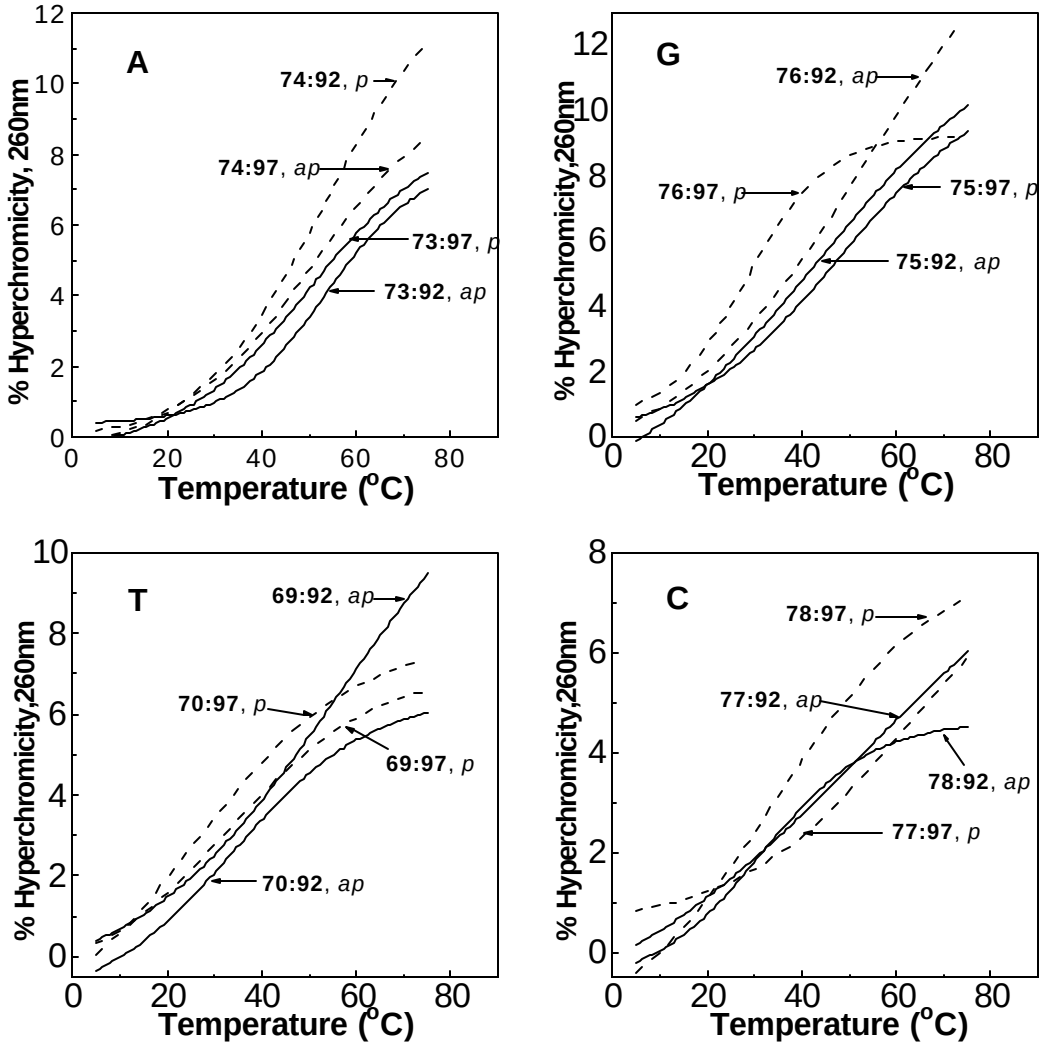


Figure 7. UV-melting profiles of parallel (---) and antiparallel (—) duplexes of *aep*PNA:DNA.

Hyperchromicity observations

A useful parameter of interest is the hyperchromicity changes accompanying the melting transitions which can be measured from UV-melting curves. This data can be derived from the UV-Tm plots and suggests two ranges of hyperchromicity for various *aep*PNAs, 6-9% and 9-12%. The duplexes of *aep*PNAs having purines generally exhibited higher hyperchromicity than those containing pyrimidines. Generally higher Tms were accompanied by larger % hyperchromicities though the hyperchromicity changes could not always be directly correlated with the thermal stabilities except in the case of *aep*PNA-C (Tables 8 & 9, entry 3),. However, in all cases, these were slightly lower than the control *aeg*PNAs. Since these reflect the extent of base-stacking, the observed pattern of changes suggest variable base stacking effects depending on the base type, stereochemistry and relative orientations of strands in duplexes. The percent hyperchromicity accompanying the melting of the mismatched PNA:DNA complexes were very much lower than that accompanying the melting of the fully complementary duplexes.

3.5.2a Mismatch Effects. The effect of mismatches on the stability of the PNA:DNA duplexes was studied using complementary DNA sequences that bore a single base mismatch at the site opposite the *aep*PNA unit. The design of the mismatch DNA sequences is depicted in Figure 8. The UV-Tm values obtained for the complexes of the PNA oligomers with DNA oligomers bearing a single mismatch opposite the site of the *aep*PNA unit are summarized in Table 10.

As seen from the ΔTm column in Table 10, the *aep*PNA:DNA (*ap*) complexes were destabilized to a greater extent In comparison with the corresponding *aeg*PNA:DNA antiparallel duplex, except that involving (2*R*,4*S*) *aep*PNA-T **70:95** (Table 10, entry 6). The *aep*PNA-A and *aep*PNA-T-containing complexes involving the (2*S*,4*S*) isomer were more sensitive to the mismatch (Table 10, entries 2 and 5; ΔTm = 9.4 & 29.7 ^{o}C

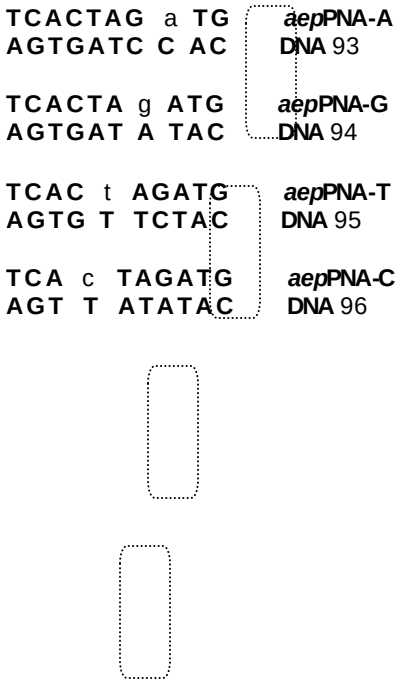


Figure 8. Mismatch DNA sequences for PNA:DNA complexes

for *aep*PNA-A **73:93** and *aep*PNA-T **69:95** respectively) compared to the (2*R*,4*S*) isomer (Table 10, entries 3 and 6; ΔTm = 5.7 & 3.3 °C for *aep*PNA-A **74:93** and *aep*PNA-T **70:95** respectively). The *aep*PNA-G (**75**, **76**) and *aep*PNA-C (**77**, **78**) – containing complexes showed a reverse trend to those containing *aep*PNA-A and *aep*PNA-T. Here, the complexes comprising oligomers involving the (2*R*,4*S*) isomer, **76:94** (entry 9) & **78:96** (entry 12) were more sensitive to the mismatch (ΔTm = 25.5 & 32.3 °C for *aep*PNA-G and *aep*PNA-C respectively) compared to the (2*S*,4*S*) isomer (ΔTm = 18.5 & 11.6 °C for *aep*PNA-G **75:94** (entry 8) and *aep*PNA-C **77:96** (entry 11) respectively).

The propensity of mismatch effects in the duplexes involving a single base mismatch at the site of the *aep*PNA unit followed the order G>T>C>A in comparison with the complexes formed by the oligomer with the respective DNA sequences.

Table 10. UV-Tm data of the mismatched *aep*PNA:DNA duplexes.

Entry			PNA:DNA	Tm (°C)	ΔTm (°C)
1	control		**68:93**	35.4	-8.4
2	*aep*PNA-A	**a**	**73:93**	44.4	-9.4
3	(A:C mismatch)	**a**	**74:93**	47.4	-5.7
4	control		**68:95**	36.8	-7.0
5	*aep*PNA-T	**t**	**69:95**	26.9	-29.7
6	(T:T mismatch)	**t**	**70:95**	30.0	-3.3
7	control		**68:94**	39.6	-4.2
8	*aep*PNA-G	**g**	**75:94**	24.7	-18.5
9	(G:A mismatch)	**g**	**76:94**	27.6	-25.5
10	control		**68:96**	36.8	-7.0
11	*aep*PNA-C	**c**	**77:96**	43.6	-11.6
12	(C:T mismatch)	**c**	**78:96**	30.0	-32.3

$\Delta Tm = Tm_{mismatch} - Tm_{ap}$

a, **t**, **g**, **c** = (2*S*,4*S*) *aep*PNA-A/T/G/C; **a**, **t**, **g**, **c** = (2*R*,4*S*) *aep*PNA-A/T/G/C

3.5.2b Dissociation Vs Re-association Kinetics

The UV-melting and re-association (by cooling) curves for hybrids of DNA with *aeg*PNA **68** and *aep*PNA **74** as a typical representative example are shown in Figure 9. It is seen that while the *aeg*PNA cooling curve is displaced significantly indicating a

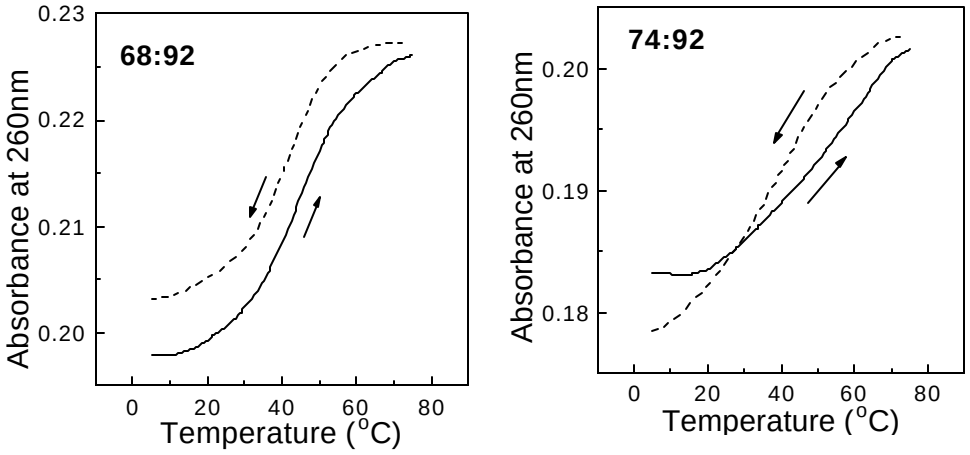


Figure 9. Hysteresis between the dissociation and re-association phenomena in *aeg*PNA (**68**) and *aep*PNA (**74**).

hysteresis, that of *aep*PNA exhibited overlapping curves. This suggests a faster re-association of *aep*PNA:DNA duplexes compared to the *aeg*PNA:DNA duplexes.

3.5.3. CD Studies

The CD spectra of the monomers (2*S*,4*S*) *aep*PNA-T **10**, (2*R*,4*S*) *aep*PNA-T **18**, (2*S*,4*S*) *aep*PNA-2-amino-6-chloropurine **24a** and (2*R*,4*S*) *aep*PNA-2-amino-6-chloropurine **24b** are shown in Figure 10, and these have clear differences. The (2*R*,4*S*) *aep*PNA-T monomer **10** displayed a minimum at 275nm and a maximum at 235nm with a cross-over point at 261nm. On the other hand, the (2*S*,4*S*) *aep*PNA-T unit **18** displayed a minimum at 265nm and a maximum at 228nm. As these are not enantiomers, the CD spectra are not expected to be mirror images of each other. The (2*S*,4*S*) and (2*R*,4*S*) *aep*PNA-2-amino-6-chloropurine monomers also exhibited CD spectra that were opposite in sign. The intensity of the spectra was negligible in the 260-280 region. The (2*S*,4*S*) *aep*PNA-2-amino-6-chloropurine monomer exhibited a minimum at 229nm, while the (2*R*,4*S*) diastereoisomer exhibited a maximum at 235nm.

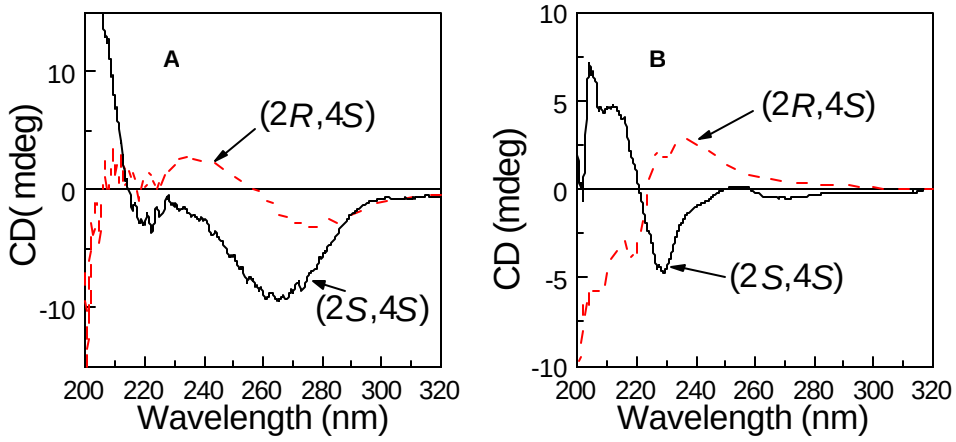


Figure 10. CD spectra of **A**. *aep*PNA-T monomeric units **10** (2*S*,4*S*) & **18** (2*R*,4*S*) and **B**. *aep*PNA-2-amino-6-chloropurine units **24a** (2*S*,4*S*) and **24b** (2*R*,4*S*)

Although the CD spectra of the *aep*PNA-T_8 single strands differed depending on the stereochemistry and the number of *aep*PNA units (Figure 11), upon complex formation with the complementary DNA oligomer, the CD exhibited by the complex was similar to

that of the control *aeg*PNA$_2$:DNA triplex (Figure 12). Thus, the CD spectra of the complexes formed by (2*S*,4*S*) *aep*PNA-T (e.g., **53:84**) and (2*R*,4*S*) *aep*PNA-T (e.g., **52:84**) were similar to that of the control *aeg*PNA **46:84**. The PNA$_2$:DNA triplexes exhibited maxima at 265nm and minima at 246nm (Figure 12).

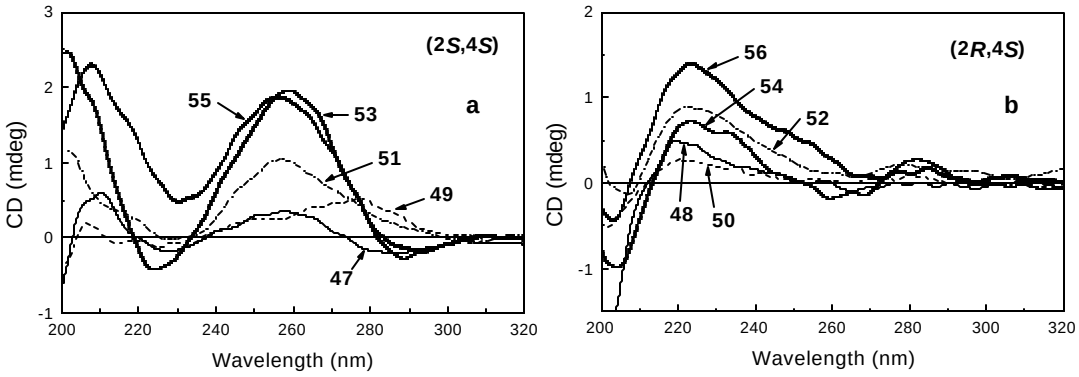


Figure 11. CD profiles of *aep*PNA-T$_8$ single strands. **a**. (2*S*,4*S*) *aep*PNA containing PNAs. **d**. (2*R*,4*S*) *aep*PNA containing PNAs.

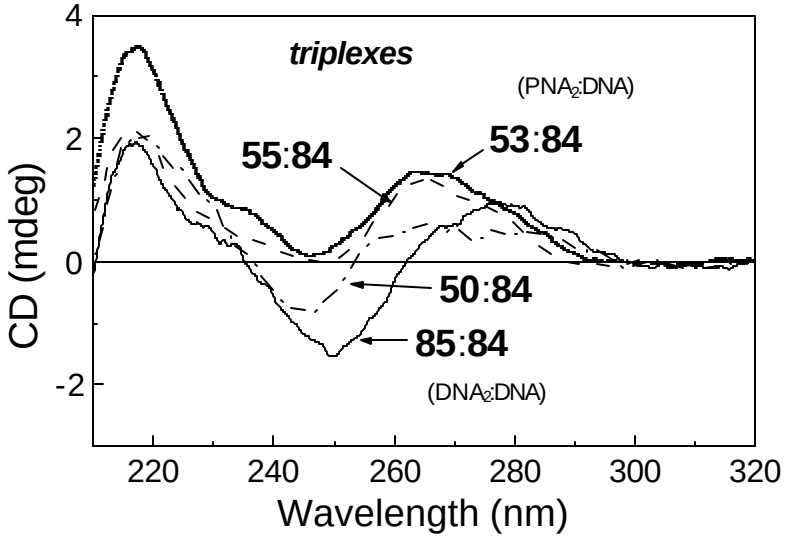


Figure12. CD profiles of PNA$_2$:DNA triplexes **53:84**, **50:84** & **55:84** in comparison with the DNA$_2$:DNA triplex **85:84**

The mixed base *aep*PNAs that were found to form duplexes with DNA by UV-melting were also analyzed by CD. The antiparallel complexes of PNAs **68**-**78** with DNA **92** exhibited similar patterns with maxima at ~270nm and ~220nm and a minimum at ~250nm (Figure 13, 14a). The CD spectra of the PNA:DNA complexes differed from the addition spectrum of the constituent single strands (e.g. Figure 14b) confirming their existence.

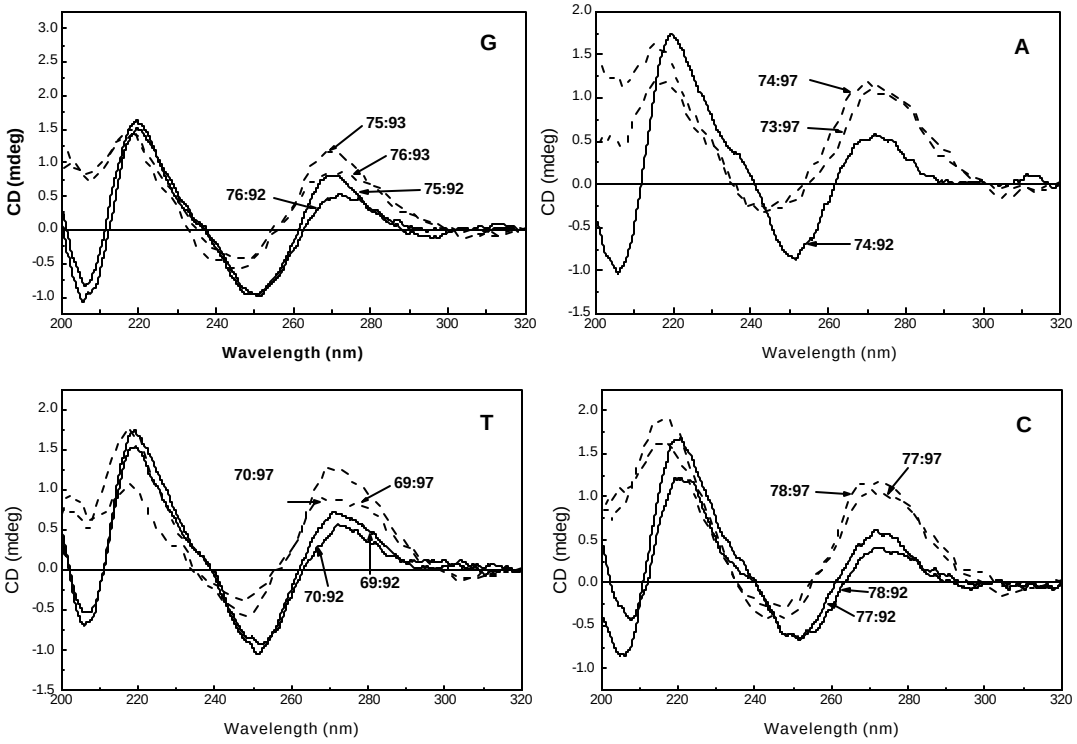


Figure 13. CD spectra of the parallel and antiparallel PNA:DNA duplexes.

The parallel PNA:DNA complexes of PNAs **69**-**78** with DNA **97** also exhibited CD spectra very similar to that of the control **46:97** and to each other with maxima ~270nm & ~220nm and a minimum at ~245nm (Figure 13, 14a). The band at ~270nm was more intense than that at ~245nm, in contrast to the antiparallel complexes, where the ~250nm band was more intense than the ~270nm band. As compared to the antiparallel complexes with DNA **92**, the parallel complexes with DNA **97** displayed a more intense CD signal for the (2*R*,4*S*) *aep*PNA-containing oligomers **70**, **72**, **74** & **78**

than the (2*S*,4*S*) *aep*PNA-containing oligomers **69**, **71**, **73** & **77**, with the exception of *aep*PNA-G in PNAs **76** & **75**.

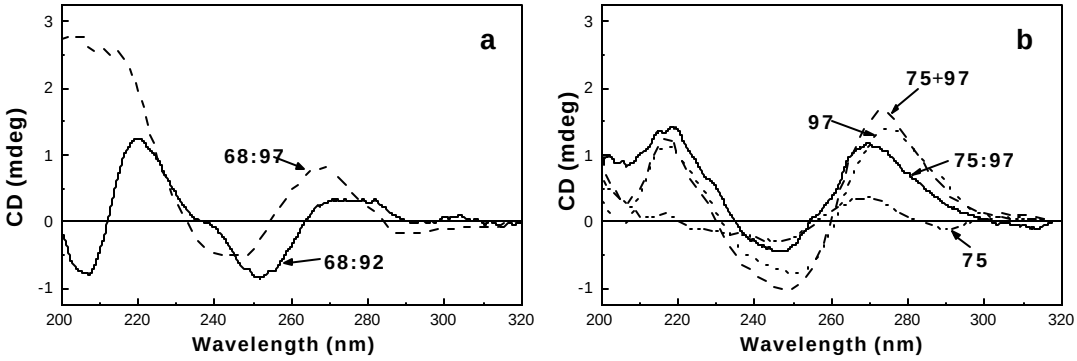


Figure 14. CD profiles of (**a**) parallel (---) and antiparallel (—) duplexes of control PNA ; (**b**) Observed (—) and addition (---) spectra of a representative PNA:DNA duplex along with the CD spectra of the constituent single strands.

3.5.4. Gel shift assays

The binding of *aep*PNAs to complementary DNA was also examined by gel retardation as shown in Figure 15. The PNA:DNA complexes derived from *aep*PNAs with even a single *aep*PNA unit e.g. **49:84** were significantly retarded in the gel (lane 3). The complexes involving PNAs with two *aep*PNA units e.g. **51:84** were retarded even more (lane 2) and remained close to the well in which they were loaded. The

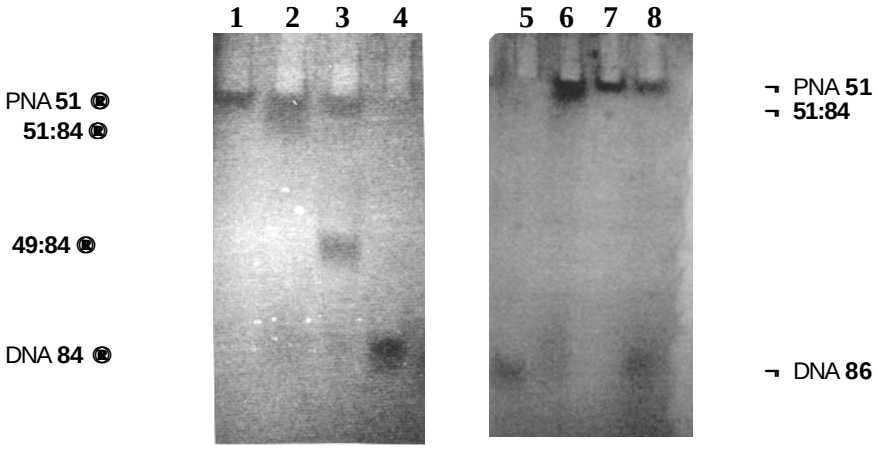


Figure 15. 15% Polyacrylamide Gel Electrophoresis of *aep*PNA:DNA complexes. Lanes **1** through **8**: PNA **51**, single strand; PNA **51**:DNA **84**; PNA **49**:DNA **84**; DNA **84**, single strand; DNA **86**, single strand; PNA **51**:DNA **84**; PNA **51**, single strand; PNA **51**:DNA **86**.

mismatch complex **51:86** was not detected (lane 8) and in this case both the constituent single strands involved **51** & **86** were observed individually in the gel lane.

3.6. DISCUSSION

The effect of PNA backbone modification in the form of *aep*PNA is expected to significantly affect the PNA:DNA thermal stability. The UV-Tm, fluorescence, CD and gel shift assay data presented in the last section suggests that the *aep*PNA:DNA interaction is significantly modulated.

3.6.1. UV-Spectroscopy

*Aeg*PNA homopyrimidine sequences comprising thymine units are known to form PNA_2:DNA triplexes.[14] Both UV-mixing and UV-titration experiments indicated a 2:1 binding stoichiometry (PNA_2:DNA) for PNA oligomers of alternating *aep-aeg* units **53** as well as holooligomers of *aep* units **55**. The percent hyperchromicity Vs temperature plots derived from the UV-melting data indicated a single transition, characteristic of PNA_2:DNA triplex melting, wherein both the PNA strands dissociate from the DNA strand simultaneously, in a single step. The stabilizing effect of the *aep*PNA units on the derived PNA_2:DNA complex progressively increased with an increase in the number of *aep*PNA units. More importantly, the specificity is retained in these cases. This was evident from the mismatch studies, where the presence of a single T-T mismatch in the center of the sequence made the PNA_2:DNA complexes very unstable. The PNA/DNA single strands, when subjected to the same temperature program as the PNA_2:DNA complexes, exhibited <3% change in absorbance, thus ruling out any significant contribution from single stranded ordering to the sigmoidal transition observed for the PNA_2:DNA triplexes. The tight binding of the T_8 *aep*PNAs with DNA was also seen in diagnostic gel mobility shift experiments, where even a single modification effected significant retardation.

To study the effect of pH on the stability of the PNA_2:DNA triplexes, the hybrids were constituted in 0.01M sodium phosphate buffers of different pH values in the range 6.0-7.4. No change in Tm ($\Delta Tm = \pm 1\ ^{o}C$) was detected in the above pH range. PNA_2:DNA complexes are known to be destabilized by the presence of salts and the stability of complexes of DNA with positively charged ligands is strongly salt-dependent.[11] The presence of salt in the medium caused a greater destabilization of the higher substituted *aep*PNA_2:DNA complexes, **54:84** & **56:84** ($\Delta Tm \geq 8\ ^{o}C$) compared to the complex of *aeg*PNA **46** with DNA **84** or monosubstituted *aep*PNA **48**:**84** (ΔTm = 4-5 ^{o}C). The higher salt dependence of the Tm of *aep*PNA_2:DNA complexes both in the alternating *aeg*-*aep* PNAs and in the *aep*PNA homooligomers is perhaps a consequence of the protonated proline ring nitrogen in the *aep*PNA units, leading to electrostatic destabilization under increased ionic strengths. The pH-titration plot of *aep*PNA-T monomeric unit exhibited three distinct transitions, wherein the pKas of the three functional groups that can be deprotonated *viz*., the carboxylic acid, the pyrrolidine ring nitrogen and the primary amine, were clearly resolved. The plot indicated a pKa ≈6.5 for the pyrrolidine nitrogen, substantiating its protonation status even at neutral pH. The large destabilization observed for the alternating *aep*-*aeg* oligomer (ΔTm per *aep*PNA unit ≈ +7-* ^{o}C) is at least partly due to electrostatic interaction between the positively charged backbone of *aep*PNA and the negatively charged sugar-phosphate DNA backbone. This is further amplified in the *aep*PNA homooligomers, which are polycationic. The main cause of *aep*PNA:DNA recognition still appears to be the specific hydrogen bonding between the complementary nucleobases of PNA(*aep*/*aeg*) and DNA, since even a single mismatch in the middle of the sequence was found to inhibit *aep*PNA:DNA complexation, as is the case with control *aeg*PNA. Both isomers studied here have the same spatial disposition of the attachment of the nucleobase to the proline ring at C4 equivalent to the nucleobase

attachment at C1' in DNA (Figure 16). Changing the stereochemistry at the other stereocenter, i.e., C2, located in the PNA backbone does not appreciably affect the stability of the *aep*PNA$_2$:DNA complexes. The structural changes caused by the difference in stereochemistry at C2 are probably accommodated due to the flexibility imparted to the backbone by the aminoethyl moiety flanking the proline ring. Such flexibility is lost when the proline nitrogen is part of an amide moiety.[15]

The trend observed in the Tm of the PNA:PNA complexes was the reverse to that observed with the PNA:DNA triplexes. In this case, a successive increase in the number of *aep*PNA units in the PNA oligomer led to a progressive decrease in the melting temperature as well as in the hyperchromicity upon melting. Thus, increasing the number of *aep*PNA units caused PNA:DNA complexation to be favoured, but PNA:PNA complexation to be disfavoured.

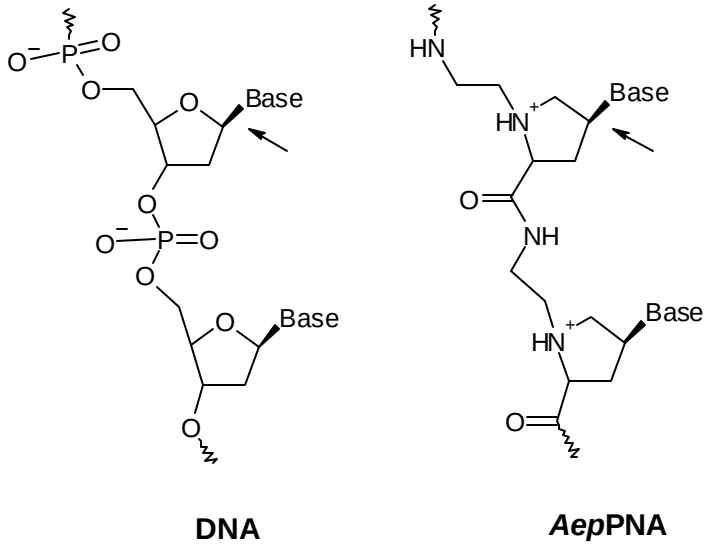


Figure 16. The similar spatial disposition of the nucleobase to the ring structure in DNA and *aep*PNA

The effect of backbone chirality of *aep*PNA while binding to complementary DNA sequences seems to be unimportant in such homopyrimidine (T_8) sequences since these bind to complementary DNA in both parallel (HG) and antiparallel (WC) orientations. The mixed purine-pyrimidine sequences (**69-78**) were constructed to

explore the effect of the *aep* backbone chirality on the directionality of binding in duplexes.

The presence of a single *aep*PNA-T unit at the terminus of a mixed sequence that is rich in adenine and thymine, e.g., **65** & **66** was not efficient in differentiating between the antiparallel and parallel binding orientations, as was the case with the *aeg*PNA **64** ($\Delta Tm \approx 2\ ^{o}C$).

The inclusion of single *aep*PNA units within a mixed sequence *aeg*PNA oligomer **68** bearing all the four nucleobases led to interesting nucleobase- and stereochemistry-dependent stabilization effects. The stabilization was at least partly due to electrostatic interactions. This was proved by the lower hysteresis in the UV-Tm dissociation-re-association plots compared to the control *aeg*PNA **68** (Figure 9).

The (2*S*,4*S*) *aep*PNA-T/A/C units in their respective positions in PNA **68** uniformly stabilized the antiparallel duplexes by 10-13 ^{o}C, while the duplex formed by *aep*PNA-G was as stable as the control **68**:**92**. The purine (2*S*,4*S*) *aep*PNA-A/G units also stabilized the parallel duplexes with DNA **97**.

The (2*R*,4*S*) *aep*PNA units exhibited base-dependent variations in both the parallel and antiparallel duplex stabilities. With the exception of the *aep*PNA-A unit, all the (2*R*,4*S*) *aep*PNA units destabilized the parallel duplexes.

The stereochemical preferences for the binding orientation were most prominent for the *aep*PNA-G units, while the *aep*PNA-A units were the least efficient in differentiating between the parallel and antiparallel modes.

The stringency of binding however, was maintained, as a single base mismatch at the site of the *aep*PNA unit caused a decrease in the Tm. This specific nucleobase recognition was even more stringent (ΔTm = 9-30 ^{o}C) than the control *aeg*PNA (ΔTm = 7-8 ^{o}C).

The duplexes formed by the purines, *aep*PNA-A/G exhibited hyperchromicity upon melting that was as good as the control duplexes (10-13%). The pyrimidines, *aep*PNA-T/C, on the other hand, showed a lower hyperchromicity (6-9%), which could be a consequence of less efficient nucleobase stacking. The pyrrolidine ring pucker and/or *syn*/*anti* orientation of the nucleobases that are directly attached to the ring may dictate the individual parallel/antiparallel preferences observed for the nucleobases on chiral *aep* units. It is known that the nature of the 4-substituent plays an important role in defining the pucker of the pyrrolidine ring in 4-substituted prolines.[16] The individual purines or pyrimidines present at the 4-position in *aep*PNA perhaps cause differential pyrrolidine ring puckers and consequent backbone conformational changes. The relatively better stabilizing effect of purines over the pyrimidines may arise from a better stacking effect in the resulting duplexes compared to the control.

3.6.2. CD Spectrophotometry

*Aeg*PNA is inherently achiral. However, upon complexation with complementary nucleic acids, it is rendered chiral and exhibits an induced CD signal. Chirality can be induced in the achiral PNA strand by linking chiral moieties like amino acids,[17] peptides,[18] or oligonucleotides[19-24] to the PNA termini. PNA has also been rendered chiral by the incorporation of chiral amino acids in its backbone.[25-34] The CD is predominantly an effect of coupling between the transition moments of the nucleobases as a result of their helical stacking.

The introduction of chiral monomers in the backbone allows the investigation of having specifically positioned stereogenic centers within the achiral PNA oligomer. Thus, it is of general interest to investigate how many stereogenic centers inserted at which positions of a PNA oligomer can induce a preferred handedness of the single or double stranded PNA, and if eventual stereochemical preorganization of PNAs can favour the DNA/RNA recognition process.[35] It has been suggested[36,37] that effective

mechanisms of inducing chirality would involve some immobilization of the rotation around the bonds of the α-carbon of the amino acid.[38]

A comparison of the structures of the complexes formed by PNA with complementary DNA/RNA and the corresponding DNA:DNA and DNA:RNA complexes suggested that PNA hybrids are right handed helices with a base-pair geometry not very much different from 'A' or 'B' form DNA.[9,38] The preferred handedness of the PNA:DNA duplexes seems to be dictated by the DNA. CD spectra of parallel and antiparallel DNA:PNA duplexes were distinctly different.

PNA-T_8 oligomers were demonstrated to form PNA_2:DNA triplexes with the DNA polypurine strand as the central strand. In such cases, CD supported the fact that a triplex is formed as the only PNA:DNA complex, and that it is a right-handed helix. The conformation of bases in the PNA_2:DNA triplex was found to be very similar to that of the conventional DNA_2:DNA T*A:T triplex.[14] The difference between the spectra for the PNA_2:DNA and the DNA_2:DNA triplexes suggests that the helical winding may somewhat differ between the two types of triplexes.

Of the two stereocenters present in each chiral *aep*PNA unit presented in this Chapter, the C2-stereocenter is present directly in the backbone (Figure 17) and hence, this is expected to exert a greater effect in inducing chirality in the oligomer backbone. The C4 stereocenter carrying the nucleobase may mostly affect the base stacking. However, since both are part of the pyrrolidine ring, the two roles may be correlated.

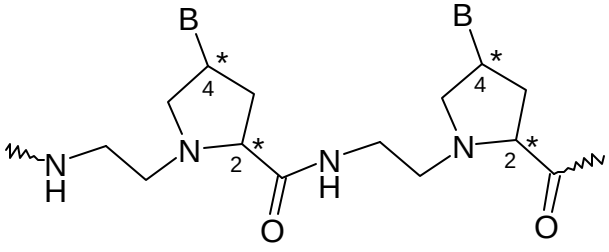


Figure 17. The *aep*PNA backbone showing the two chiral centers of each *aep*PNA unit.

With an increasing number of *aep*PNA units, the CD spectra of the corresponding oligomer showed a progressive increase in the intensity (Figure 11). Of the two diastereoisomers, the L-*cis* (2*S*,4*S*) isomer induced the stronger signal (Figure 11b) compared to the D-*trans* (2*R*,4*S*) isomer (Figure 11a). This indicated that the L-*cis* conformation was more efficient in inducing helicity in the oligomer. This is in contrast to earlier reports by Gangamani *et al*., where the D-*trans* isomer was found to possess the greater CD signal.[39]

The CD induced by the *aep*PNA units in the PNA single strands seemed to be inconsequential, since the *aep*PNA oligomers formed complexes (duplexes/triplexes) that gave very similar CD signals upon complexation with DNA. Significantly, these CD profiles were similar to those of the control achiral *aeg*PNA complexes. Thus, in a PNA:DNA complex, it is the CD of the DNA that dominates over any inherent CD of the PNA involved in the structure.

The CD signals of the duplexes of the *aep*PNA oligomers **69-78** with complementary DNA **92/97** were similar to those of the control *aep*PNA **68**. The backbone geometry (2*S*/*R*,4*S*) did not have any dramatic effect on the geometry of either parallel or antiparallel PNA:DNA duplexes.

3.6.3. Gel Retardation Assays

The gel mobility shift assays described in section 3.7.6 are confirmative proof for the formation of PNA_2:DNA triplexes. The formation of such a complex with DNA **84** resulted in significant retardation in the polyacrylamide gel electrophoresis. More importantly, the complexes constituted using DNA **86** with a single mismatch, failed to exhibit any complex and the two constituent strands were both observed separately in the gel. These results were in concordance with the data obtained from the UV-thermal melting studies with these oligomers. This underlined the specificity of the hydrogen-

bonding between the complementary base-pairs, which is of utmost importance for the application of oligonucleotide/PNA analogues in antisense/antigene therapy.

The *aep*PNAs exhibited enhanced solubility in aqueous media is probably a manifestation of the positive charge present in the monomeric units (pKa of the pyrrolidine ring nitrogen is ~6.7, which suggests that it should be at least partly protonated at pH 7.4). The PNAs containing *aep*PNA units remained in solution for a much longer time duration as compared to the uncharged *aeg*PNA oligomers, which precipitated out of solution as a result of self-aggregation.

3.7. CONCLUSIONS

The *aep*PNA units were shown to possess very interesting DNA-binding properties. The ability of PNAs containing these units in differentiating between parallel and antiparallel binding modes, which is incongruous in the PNA_2:DNA triplexes, becomes obvious in PNA:DNA duplexes.

The positive charge on the pyrrolidine nitrogen is at least partly responsible for the high binding affinity. The main contributing factor to the stability however, remains the specific hydrogen bonding between the A-T and G-C nucleobases. The fact that the *aep* units in mixed sequence duplexes exhibit stabilization effects dependent on the nucleobase and backbone chirality suggest that besides the positive charge and inter-nucleobase hydrogen bonding, other factors like the pyrrolidine ring pucker and/or *syn*/*anti* orientation of the nucleobases may play an important role. An added and important asset of this positively charged chiral backbone is its enhanced solubility in aqueous media.

In order to explain all the observed results, more work needs to be carried out, including conformational analysis of the pyrrolidine ring, etc. To investigate the

undiluted effect of the aminoethylprolyl backbone, homooligomeric *aep*PNA sequences need to be studied at length.

3.8. EXPERIMENTAL

3.8.1. UV studies

All the UV spectrophotometric studies were performed on a Perkin Elmer λ15 UV-VIS spectrophotometer equipped with a Julabo temperature programmer and a Julabo water circulator to maintain the temperature. The samples were degassed by purging nitrogen or argon gas through the solution for 2-3min prior to the start of the experiments. Nitrogen gas was purged through the cuvette chamber below 15 ^{o}C to prevent the condensation of moisture on the cuvette walls.

3.8.2. UV-titration

To a solution of DNA **84** in 0.01M sodium phosphate at pH 7.4, were added fixed portions of the complementary PNA oligomer **46**/**55**. The temperature of the circulating water was maintained at 10 ^{o}C (i. e., well below the melting temperature of the complexes) and the absorbance at each step was recorded at 267 and 277nm. This was plotted as a function of the PNA mole fraction.

3.8.3. UV-Tm

The PNA oligomers and the appropriate DNA oligomers were mixed together in stoichiometric amounts (2:1 PNA:DNA for oligothymine-T_8 PNAs or 1:1 for the duplex forming PNAs, *viz*., the mixed base sequences) in 0.01M sodium phosphate buffer, pH 7.4 to achieve a final strand concentration of either 0.5 or 1μM each strand. For the AT-rich PNAs, antiparallel complexes were constituted by employing DNA **88** while parallel complexes were constituted using DNA **87**. The antiparallel complexes involving the PNAs containing all the four nucleobases were constituted using DNA **92**, while DNA **97** was used to get the parallel complexes. The samples were heated at 85 ^{o}C for 5min

followed by slow cooling to room temperature. They were allowed to remain at room temperature for at least half an hour and refrigerated overnight prior to running the melting experiments. Each melting experiment was repeated at least thrice. The absorbance or the percent hyperchromicity at 260nm was plotted as a function of the temperature. The Tm was determined from the peaks in the first derivative plots and is accurate to $\pm 1\ ^{o}C$.

For the salt-dependence experiments, 100mM sodium chloride was added to the buffer before annealing the samples. In the pH-dependence experiments, the PNA:DNA complexes were constituted in 0.01M phosphate buffers of different pH values prior to annealing. In order to study the relative rates of melting and re-association, the complexes were held at 75 ^{o}C for 5min upon completion of the heating process and then cooled to 5 ^{o}C at a rate of 0.5 ^{o}C/min. As for the heating process, the absorbance values were recorded every minute and plotted as a function of the temperature along with the heating profile.

3.8.4. Mismatch studies

DNA **86** was used to probe the specificity of the PNA-oligothymine-T_8 interaction with DNA. The relevant PNA and DNA strands were mixed together in a 2:1 molar ratio and subjected to UV-melting.

The mixed sequence decamer specificity of binding to complementary DNA was probed by using a DNA sequence specific for each nucleobase. The DNA oligomers used were: DNA **93** for *aep*PNA-A; DNA **94** for *aep*PNA-G; DNA **95** for *aep*PNA-T and DNA **96** for *aep*PNA-C. The PNA and corresponding mismatched DNA sequence were mixed together in a 1:1 molar ratio and annealed prior to the melting experiments.

3.8.5. CD

CD spectra were recorded on a Jasco J-715 spectropolarimeter. The CD spectra of the PNA:DNA complexes and the relevant single strands were recorded in 0.01M

sodium phosphate buffer, pH 7.4. The temperature of the circulating water was kept below the melting temperature of the PNA:DNA complexes, i. e., at 10 ^{o}C.

The CD spectra of the oligothymine T_8 single strands were recorded as an accumulation of 8 scans from 320 to 195nm using a 1cm cell, a resolution of 0.1nm, band-width of 1.0nm, sensitivity of 2mdeg, response 2sec and a scan speed of 50nm/min. for the PNA_2:DNA complexes, spectra were recorded as an accumulation of 4 scans, response of 1sec and a scan speed of 200nm/min.

The CD spectra of the mixed base PNAs and the derived PNA:DNA duplexes were recorded as an accumulation of 3 scans and a scan speed of 200nm/min.

3.8.6. Gel shift experiments

The PNA (**49**, **51**) and DNA oligomers (**84**/**86**) utilized in the gel mobility shift assays were mixed together in the desired ratios in water. The sample was lyophilized to dryness and re-suspended in 10µl tris-EDTA buffer, pH 7.0. The sample was annealed as described earlier by heating at 85 ^{o}C for 5min followed by slow cooling to room temperature and refrigeration overnight. Prior to loading, 10µl glycerol in TBE buffer, pH 7.0, the gel-running buffer, was added and the sample, loaded on the gel. Bromophenol blue was used as the tracer dye, but was loaded in an adjacent well and not mixed with the sample.

Gel electrophoresis was performed on a 15% non-denaturing polyacrylamide gel (acrylamide:*bis*-acrylamide, 29:1) until the BPB migrated to three-fourths of the gel length. During electrophoresis, the temperature was maintained ~15 ^{o}C.

Bands of the single stranded PNA/DNA and the PNA:DNA complexes were visualized as dark bands by UV-shadowing, i.e., by illuminating the gel placed on a fluorescent thin-layer silica gel chromatographic plate, F_{254}, 20cm x 20cm using UV light.

SECTION B

3.9. INTRODUCTION

The neutral and achiral aminoethyl PNAs hybridize with high affinity and specificity to oligonucleotides obeying the Watson-Crick base-pairing rules and thus, are true mimics in terms of base-pair recognition.[17] As a consequence of their achiral nature, PNAs can hybridize to complementary DNA/RNA in both, the antiparallel (*ap*, with the PNA 'N' terminus towards the 3'-end and the 'C' terminus towards the 5'-end of the complementary DNA/RNA oligonucleotide) and parallel (*p*, with the PNA 'N' terminus towards the 5'-end and the 'C' terminus towards the 3'-end of the complementary DNA/RNA oligonucleotide) modes as shown in Figure 18, although the antiparallel complex is slightly more stable than the parallel one. While PNA:DNA duplexes are most stable in the antiparallel form, the parallel orientation seems to be preferred for PNA_2:DNA triplexes, if only one species of PNA is employed.[7]

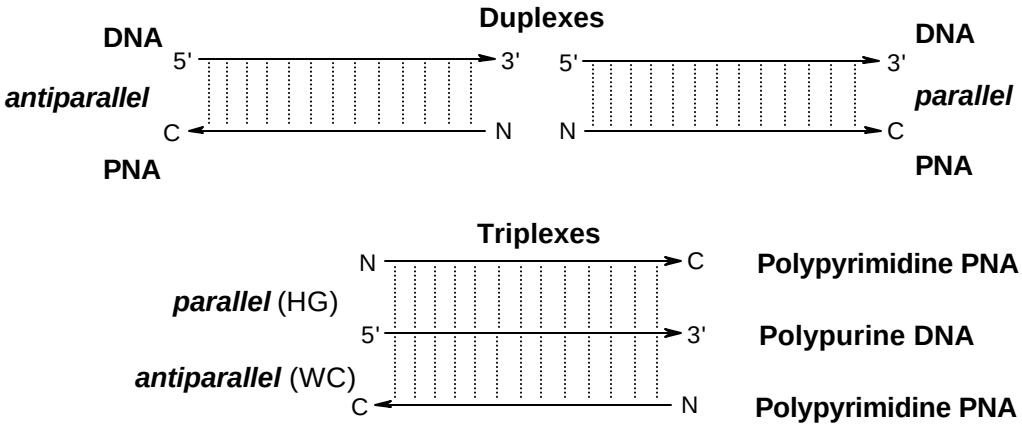


Figure 18. The parallel and antiparallel orientation of the PNA strands in a PNA:DNA duplex and a PNA_0:DNA triplex

PNA_2:DNA triplex formation[40] follows the rules of homopyrimidine DNA triplex formation. Thus, the most stable triplexes are formed when the Watson-Crick base-pairing PNA strand is in the antiparallel orientation relative to the DNA strand and the

Hoogsteen strand is in the parallel orientation relative to the DNA strand, i. e., the two strands are antiparallel to each other. [17]

Homopyrimidine PNAs that are rich in cytosines add to the double stranded target polynucleotides as Hoogsteen strands forming PNA:DNA_2 triplexes[1,7,40,41] These oligomers such as PNA-C_{10} or PNA-$(CT)_5$ do not invade double stranded polynucleotides by forming triplexes in a way analogous to that observed with thymine-rich PNAs.[42] This was concluded on the basis of LD and CD titration ratios that indicated 1: 1 stoichiometry of PNA to DNA. These reactions are significantly pH-dependent owing to the requirement of N3-protonation of cytosine in the third strand.

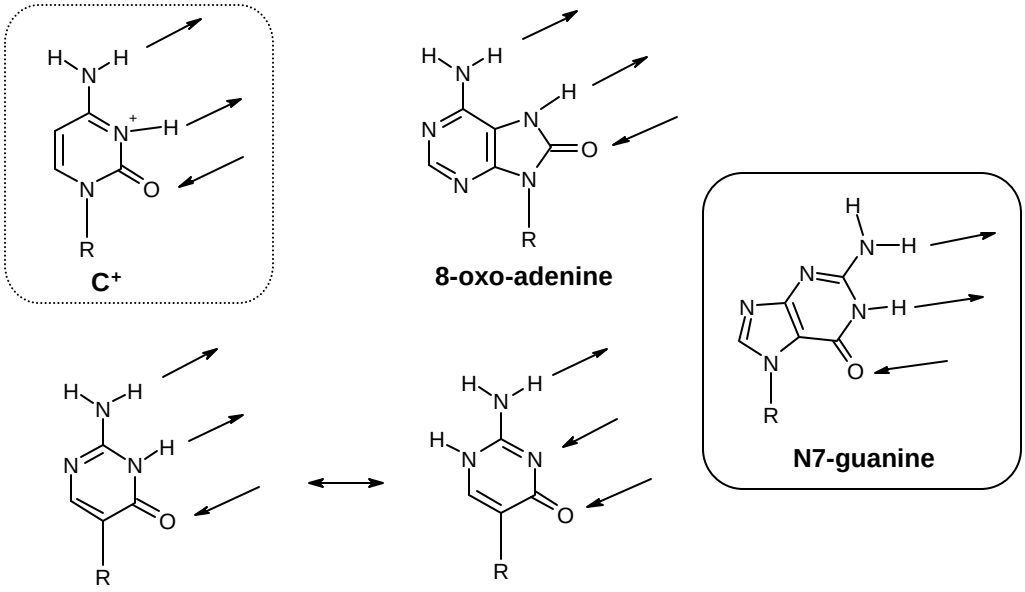


Figure 19. Neutral C^+ hydrogen bonding mimics.

In literature, reports exist of attempts to mimic the C^+ hydrogen bonding pattern by neutral bases to avoid the necessity of low pH for C-containnig PNA oligomers (Figure 19). One such mimic, 8-oxo-adenine has been employed to recognize guanine in the Hoogsteen mode.[43-45] Pseudoisocytosine (ψiC) is another C^+ mimic that has been utilized to achieve pH-independent Hoogsteen binding of the third strand to a G:C base

pair.[46,47] Pseudoisocytosine has also been introduced into PNA for pH-independent triplex formation.[48]

Employing bis-PNAs i.e., where the Watson-Crick PNA strand is connected by continuous synthesis *via* ethylene glycol type linkers to the Hoogsteen PNA strand[48,49] leads to an increase in the melting temperature of the triplexes. Optimal results in terms of pH-independence and thermal stability of the complexes formed were obtained when the cytosines in the Hoogsteen strand were replaced by pseudoisocytosine in comparison with the case when the pseudoisocytosine was placed in the Watson-Crick strand. In the complexes where the cytosine PNA strand is antiparallel to the DNA target (and thus the pseudoisocytosine strand is parallel), almost no pH dependence of the Tm was observed (Figure 20a). This indicated an orientation-directed selectivity in complex formation to align the Watson-Crick strand in the antiparallel configuration and the Hoogsteen strand in the parallel configuration as in Figure 20b.

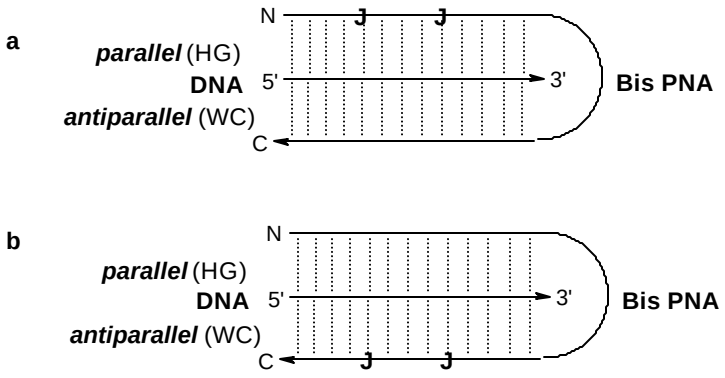


Figure 20. Triplex formation with bisPNAs carrying pseudoisocytosine as a neutral C^+ mimic.

N7-glycosylated guanine, a positional isomer of the naturally occurring N9-guanine is yet another nucleobase that provides bidentate hydrogen bonding to target G:C base pairs (Figure 21). It thus, mimics the C^+ hydrogen bonding pattern, but surpasses the requirement of low pH for complex formation. Nucleoside derivatives of N7-guanine

have been demonstrated to be effective in forming base triplets with dG:dC base pair targets.[50-52] In DNA, N7G was found to complex with a G:C base pair only when present in a parallel third strand, and not in an antiparallel one, i. e., by attaching the deoxyribose moiety at the N7 position of a guanine base, the third strand orientation in a G•G:C base triplet is reversed and becomes parallel to the purine Watson-Crick strand.[51]

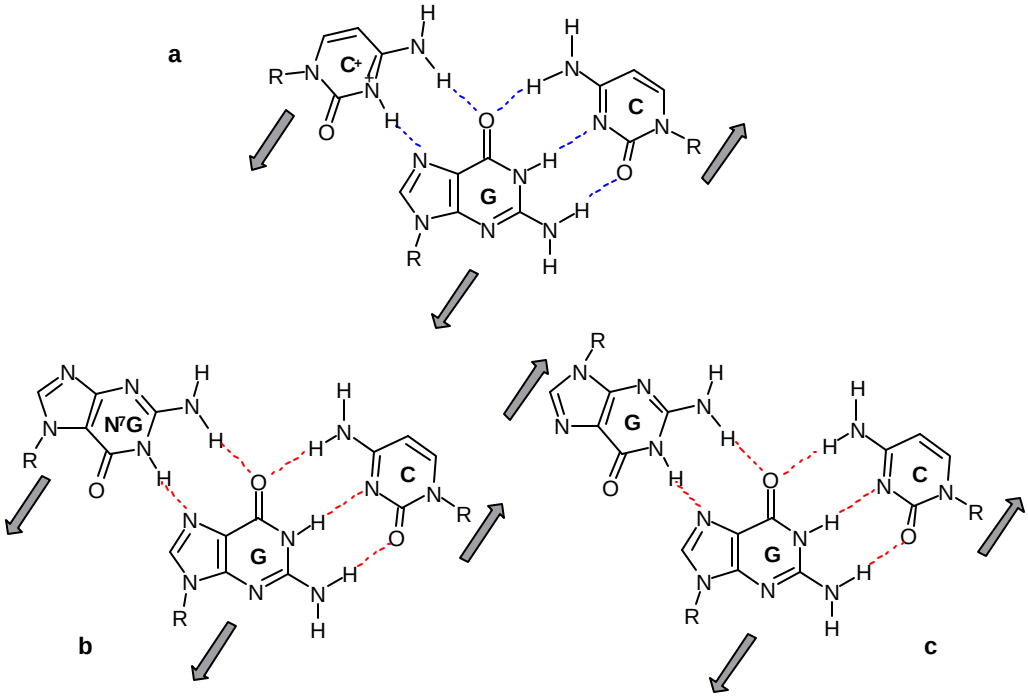


Figure 21.a. C^+, **b**. N7G and **c**. G hydrogen bonding patterns when present in the third strand of a DNA:DNA:DNA triple helix.

An acyclic nucleoside analogue bearing the N7-guanine nucleobase was also found to be able to form triplexes.[53] This analogue retains the heterocyclic portion of the molecule essential for base-base recognition, but lacks the rigid conformational features of the ribose/deoxyribose sugar (Figure 22). The triplex containing five adjacent acyclic N7G residues exhibited Tm values significantly higher than the triplex containing alternating G-C base pairs.

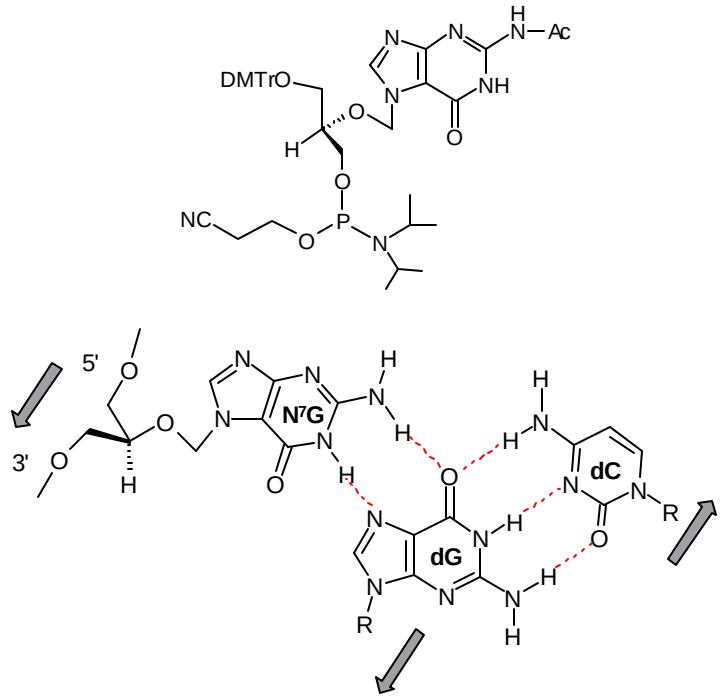


Figure 22. Acyclic N7G nucleoside analogue and its triplex formation.

3.10. RATIONALE FOR THE PRESENT WORK

In PNA_2:DNA triplexes, the DNA strand is engaged in simultaneous binding to the two PNA strands. Here, ambiguity persists in establishing the correlation of specific parallel, antiparallel PNA strands with Hoogsteen or Watson-Crick binding to DNA.[54] Introduction of a chiral identifier like *aep*PNA into one of the PNA strands may help determine their orientation and binding mode. In practice, this would lead to formation of PNA:DNA triplexes with multiple compositions. The uncertainty in stoichiometry of such mixed PNA_2:DNA complexes can be overcome by conjugating the two PNA strands through a linker to construct a hairpin that would only form triplexes with a 1:1 PNA:DNA composition. Linking the two PNA strands together in such a hairpin PNA reduces the entropy and converts binding into a bimolecular process. Positively charged PNAs have been demonstrated to possess superior strand invasion potential.[50,55] Hence, a positively charged linker including lysine residues was selected for the hairpin PNA sequences of interest. Section A of this Chapter reports the

excellent binding properties of *aep*PNAs with complementary DNA sequences. It was envisaged that this chiral PNA that is configurationally homomorphous with DNA, in conjunction with achiral *aeg*PNA, may induce orientational preferences in binding to DNA sequences.

Homopyrimidine PNA_2:DNA triplex formation is pH-dependent owing to the requirement of N3-protonation of cytosine in the Hoogsteen mode of recognition. N7G being a C^+ mimic, should circumvent the requirement of low pH for triplex formation and make effective triplexes at physiological pH.

3.11. OBJECTIVES

The objectives of this Section are:

(i) Synthesis of amionethylglycyl PNA-N7G monomer bearing N7-guanine as a C^+ mimic.

(ii) Synthesis of 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyrylguanin-7-yl)-2(*S*)-proline as a chiral N7G monomeric unit.

(iii) Design and solid phase peptide synthesis of bisPNA oligomers incorporating the *aeg*/*aep*PNA-N7G units.

(iv) A study of the pH-(in)dependence of triplex formation.

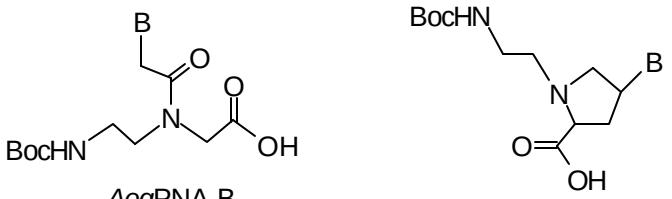
(v) A study of the influence of these oligomers on the binding orientation (*ap*/*p*).

3.12. WORK DONE

3.12.1. Synthesis of Protected Monomeric Amino Acids

The *N*-(Boc-aminoethyl)-(thymin-1-yl)-glycinate and *N*-(Boc-aminoethyl)-(N^4-benzyloxycarbonylcytosin-1-yl)glycinate were synthesized by reported procedures (See Chapter 2). The *N*-(Boc-aminoethyl)-(N^2-isobutyrylguanin-7-yl)glycinate was synthesized according to the procedure reported in Chapter 2, Section 2.2.5. The

aminoethylprolyl monomers i.e., 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^4-benzyloxycarbonyl cytosin-1-yl)-pyrrolidine-2(*S*)-carboxylic acid and 1-(*N*-Boc-aminoethyl)-4(*S*)-(N^2-isobutyrylguanin-7-yl)-pyrrolidine-2(*S*)-carboxylic acid were synthesized as described earlier (See Chapter 2, Section 2.2.3) starting from 4(*R*)-hydroxy-2(S)-proline.



*Aeg*PNA-B

41 B = T
42 B = C^{CBz}
45 B = $N7G^{iBu}$

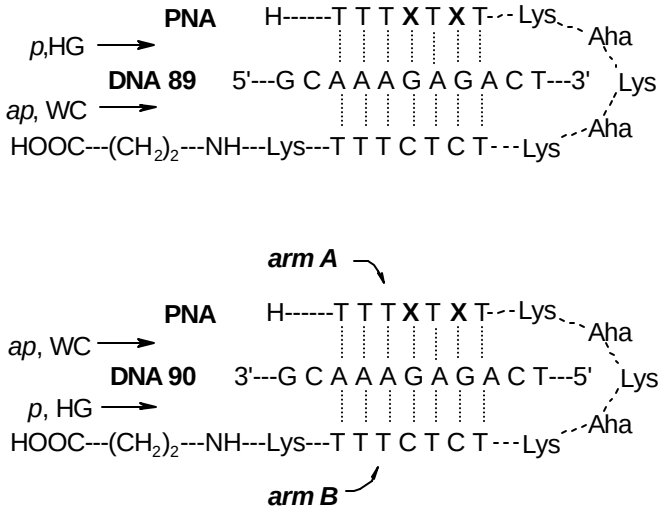
*Aep*PNA-B

27a B = C^{CBz}
31 B = $N7G^{iBu}$

Aeg and *aep* PNA monomeric units used in the synthesis of bisPNAs

(*N*-Boc)-ε-aminohexanoic acid (Aha) was prepared by treating ε-aminohexanoic acid with Boc-azide. N^{α}-Boc-Lysine with the side-chain ε-amino group protected as the 2-chlorobenzyloxy carbamate was purchased from Novasyn.

3.12.2. Design and Synthesis of Hairpin bisPNA Oligomers



X = *aeg*PNA-C, **PNA 79**

*aeg*PNA-N7G, **PNA 81**

*aep*PNA-C, **PNA 80**

*aep*PNA-N7G, **PNA 82**

Aha= ε-aminohexanoic acid

Figure 23. Design of hairpin bisPNAs.

The hairpin PNA sequences were designed as shown in the Figure 23. The *aeg*/*aep*PNA-N7G units are present in the arm A of the hairpin and bind to DNA **89** and DNA **90** in the parallel and antiparallel orientations respectively. Thus, arm B of the hairpin binds to DNA **89** and DNA **90** in the antiparallel and parallel orientations respectively. The PNA and DNA oligomers utilized in this study are listed in Tables 11 and 12 respectively.

Table 11. Hairpin bisPNA Oligomer Sequences

PNA	*Sequence Composition*
79	H- TTTCTCT -Lys-Aha-Lys-Aha-Lys- TCTCTTT -NH-$(CH_2)_2$-COOH
80	H- TTT **c** T **c** T -Lys-Aha-Lys-Aha-Lys- TCTCTTT -NH-$(CH_2)_2$-COOH
81	H- TTT^{7}*G*T^{7}*G*T -Lys-Aha-Lys-Aha-Lys- TCTCTTT -NH-$(CH_2)_2$-COOH
82	H- TTT^{7}**g**T^{7}**g**T -Lys-Aha-Lys-Aha-Lys- TCTCTTT -NH-$(CH_2)_2$-COOH
83	H- TTTGTGT -Lys-Aha-Lys-Aha-Lys- TCTCTTT -NH-$(CH_2)_2$-COOH

T/C/*7G*/G = *aeg*PNA-T/C/*7G*/G; **c**/^{7}**g** = *aep*PNA-C/N7G

Table 12. DNA Oligonucleotide Sequences

DNA	**Oligomer Sequences** 5' → 3'	
89	G C A A A G A G A C T	antiparallel DNA
90	T C A G A G A A A C G	parallel DNA
91	A G T C T C T T T G C	DNA complementary to **89** for duplexation

The PNA oligomers were synthesized utilizing the Boc-protection strategy outlined in Chapter 2. Each coupling cycle involving the deprotection, neutralization and coupling steps. was monitored for completion by the Kaiser's test. Lysine and ε-aminohexanoic acid of the linker loop were also coupled using the same procedure employing HOBt and diisopropylcarbodiimide as the coupling agents. Merrifield resin derivatized with β-alanine was used as the solid support.

The oligomer synthesis was carried out in a single assembly starting from the first arm (arm B) of the hairpin, the pentameric linker loop and the first thymine residue of

the second arm (arm A) of the hairpin. At this stage, the resin was partitioned into five portions. Of these, one portion was not extended further and was retained as such for control experiments. The four remaining portions were extended separately to include the *aeg*PNA-C, *aeg*PNA-N7G, *aep*PNA-C and *aep*PNA-N7G monomers respectively at the positions shown in Figure 23, followed by completion of the synthesis with the other *aeg*PNA monomers.

3.12.3. Cleavage from the Solid Support

The PNA oligomers were cleaved from the solid support using TFMSA[56] to yield oligomers with free 'C' terminal carboxylic acids. During this process, the cytosine exocyclic amino protecting groups, *viz*. the benzyloxycarbamate are cleaved liberating the amine. However, the N^2-isobutyryl groups of the N7-guanine units are left intact and had to be additionally subjected to treatment with aqueous ammonia at 55 ^{o}C for 16h to obtain the fully deprotected hairpin PNA oligomers.

3.12.4. Purification

The fully deprotected oligomers were initially desalted by size-exclusion chromatography over G25 Sephadex. They were subsequently purified by FPLC on a reverse phase C8 column using an ascending gradient of acetonitrile in water containing 0.1% TFA. The purity of the oligomers was re-checked by reverse phase HPLC on a C18 column and confirmed by MALDI-TOF mass spectrometry. Some representative HPLC profiles and mass spectra are shown in Figures 24 and 25 respectively.

3.12.5. UV-melting

The hairpin PNA sequences **79**-**83** were mixed with the appropriate DNA oligomer **89** or **90** in equimolar concentration and annealed prior to melting. The samples were

heated at a rate of 0.5 °C rise per minute and the absorbance at 260nm was recorded at every minute. The percent hyperchromicity at 260nm was plotted as a function of temperature and the melting temperature was deduced from the peak in the first derivative plots.

3.13. RESULTS

The Tm values of the complexes of the bis-PNAs with the complementary DNA strands are detailed in Table 11.

Table 11. UV-Tm data of the complexes formed by the bisPNAs with DNA.

Entry	PNA	Tm (°C)			
		DNA **89** (*p*)		DNA **90** (*ap*)	
		pH 7.4	*pH 5.8*	*pH 7.4*	*pH 5.8*
1	*aeg*PNA-C (**79**)	45	52.7	45	52
2	*aeg*PNA-N7G (**81**)	45	47	47	45
3	*aep*PNA-C (**80**)	37.5	47	48	52
4	*aep*PNA-N7G (**82**)	42	43	51	49

The control PNA **79** bearing *aeg*PNA-C in both arms bound to DNA **89** and DNA **90** with the same affinity (Table 11, entry 1) and the Tm gain at lower pH was 7 °C. The replacement of *aep*PNA-C in arm A by the C^+ mimic, N7G (bisPNA **81**), gave similar stabilities with DNA **89** and DNA **90** at both pH values (Table 11, entry 2, Figure 26c), with $\Delta Tm \approx 2$ °C at lower pH. No differences were thus seen in the binding of PNA **81** with DNA **89** or DNA **90**.

On the other hand, introduction of chiral *aep*PNA units led to a stronger binding with DNA **90** as compared to DNA **89** (Figure 26 a, b). With *aep*PNA-C **80** (ΔTm = 10 °C with DNA **89** & ΔTm = 4 °C with DNA **90** (Table 11, entry 3). With *aep*PNA-N7G, no Tm differences were seen for DNA **89** at pH 7.4 and 5.8 (Tm = 42-43 °C). Interestingly, PNA **82** exhibited maximum stability with DNA **90** at neutral pH (Tm = 51 °C), which is

as good as that shown by *aeg*PNA-C **79** at pH 5.8. In addition, as with *aeg*PNA **81**, the Tm values with *aep*PNA **82** at pH 7.4 & pH 5.8 differed only marginally.

3.14. DISCUSSION

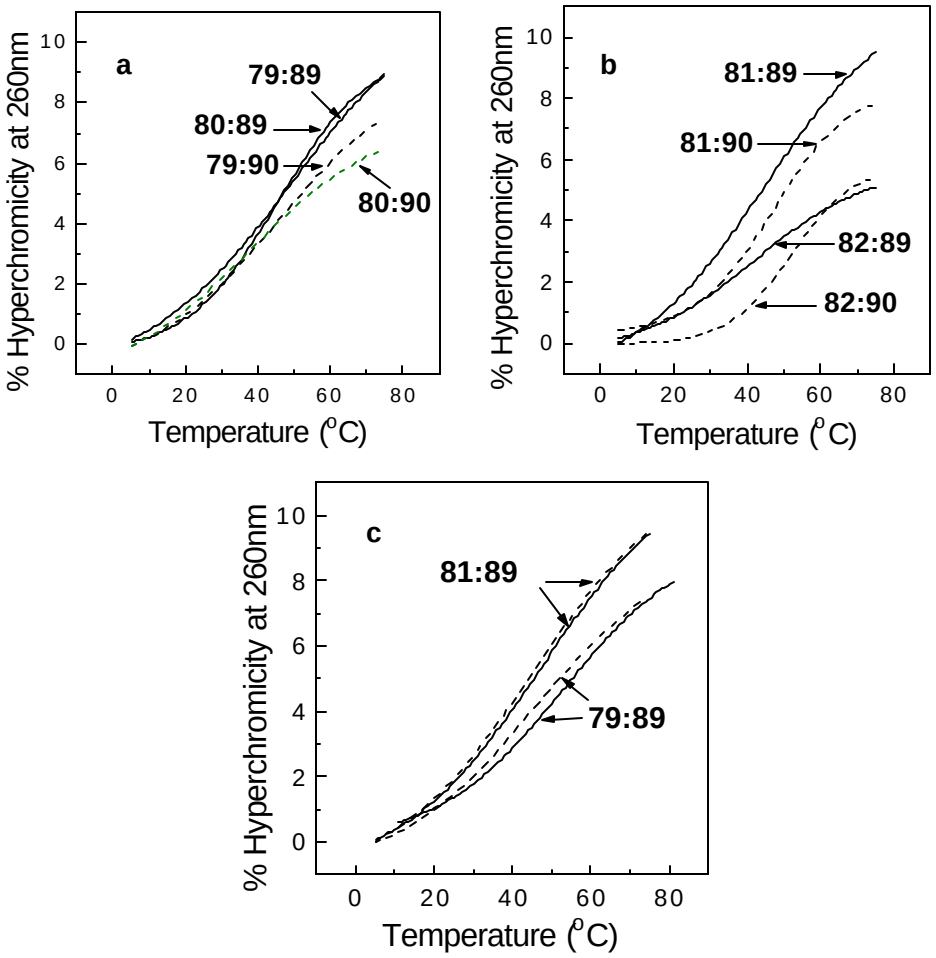


Figure 26. UV-Tm plots of the bisPNA:DNA complexes exhibiting the binding orientation bias with (**a**) *aep*PNA-C & (**b**) *aep*PNA-N7G and (**c**) the pH (in)dependence at pH 5.8 (—) and 7.4 (---).

The *aeg*PNA-C **79** and *aep*PNA-C-containing bis-PNA **80** both exhibited strong pH dependence, the complexes with the complementary DNA oligomer being more stable at acidic pH (Table 11, Entry 1). In the case of hairpin *aeg*PNA-C **79**, $\Delta Tm \approx 7\ ^{o}C$ was seen with both DNA **89** and DNA **90**. For hairpin PNAs containing *aep*PNA-C **80**, ΔTm

≈ 10 °C was observed with DNA **89** in comparison to 4 °C with DNA **90**. These results show that the bis-PNAs **79** & **80** conform to the rule of homopyrimidine PNAs forming triplexes, which requires the N3-protonated cytosine. This is further endorsed by the fact that replacement of 'C' in one arm of the hairpin by N7-guanine, a C^+ mimic, in both *aeg*PNA **81** and *aep*PNA **82** led to pH-independent complex formation with both DNA

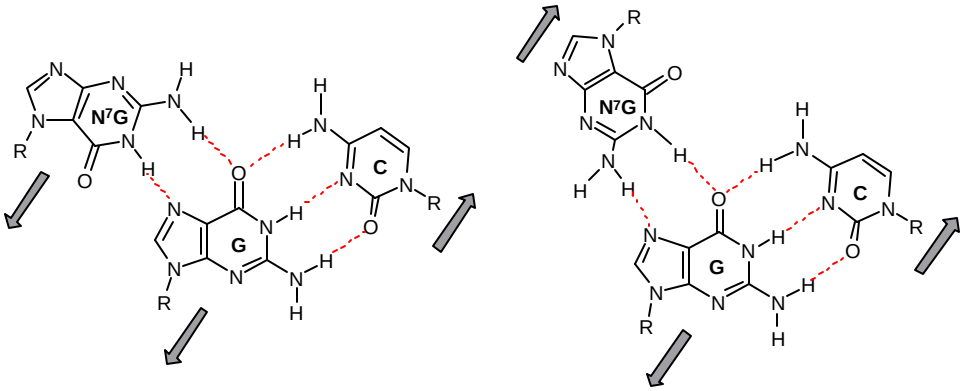


Figure 27. N7G third strand binding in the parallel and antiparallel HG modes.

89 and DNA **90** (ΔTm ≈ 1-2 °C). This is due to the fact that N7-guanine possesses the H-bonding motifs identical to those present in N3-protonated cytosine. Significantly, none of the *aeg*PNAs could distinguish between DNA **89** & DNA **90** at either pH. *Aeg*PNA-C and *aeg*PNA-N7G-containing bis-PNAs gave similar stabilities with the target DNA strands **89** and **90**. This perhaps arises from the capability of *aeg*PNA-N7G in recognizing a G:C base pair in the Hoogsteen mode without differentiation of the parallel or antiparallel orientation of the third strand (Figure 27).

The Tm results indicate that replacing the *aeg*PNA-C/*aeg*PNA-N7G units by the aminoethylprolyl backbone *aep*PNA-C/*aep*PNA-N7G introduced a significant bias in the DNA binding orientation. At physiological pH, *aep*PNA-C-containing bisPNAs bound to DNA **90** stronger than to DNA **89** (ΔTm ≈ 10 °C), indicating the orientation bias. At acidic pH, although this trend persisted, the difference in Tm was slightly lower (ΔTm ≈

5 °C). The significant differences seen in binding of *aep*PNAs **80** & **82** to DNA **89** & **90** could be attributed to the chirality of *aep*PNA-C and *aep*PNA-N7G units, that induces a better binding of DNA **90** with the *aep*PNA units. Although the *aep* units do not markedly increase the Tm of triplexes over that of the *aeg* units, the highest Tm in *aep* was comparable to that of *aeg* and most importantly, a distinction between the parallel and antiparallel orientations could be achieved.

It is possible that the cytosine N3 in *aeg*PNA-C and *aep*PNA-C units in the Hoogsteen strand may have different pKas. The presence of cationic lysine residues and the possible protonation of the pyrrolidine ring nitrogen in the *aep*PNA at physiological pH may lower the pKa of cytosine N3 in *aep*PNA-C causing less-pronounced stabilization effects at pH 5.8. Nevertheless, the observations in entry 3 (Table1) strongly indicate that the chiral *aep*PNA-C units in **80** stabilize the PNA:DNA complex **80**:**90** when they are present in the strand oriented in an antiparallel direction to the central strand. A similar preference for the antiparallel orientation of the *aep*PNA-N7G-containing units was observed for DNA **90** over DNA **89** at both acidic ($\Delta Tm \approx 9$ °C) and neutral ($\Delta Tm \approx 6$ °C) pH conditions (Entry 4, Table 11). Since N7G is known to recognize to G of a G:C basepair only through Hoogsteen bonding and being present in the arm A of PNA **82**, the arm B containing *aep*PNA-C should consequently interact in the Watson-Crick mode of binding (non-protonated N3 of C). This leads to non-protonated WC-mode C and no change observable in Tm at lower pH ($\Delta Tm \pm 1$ °C). For a better appreciation of *aep*PNA-C/N7G effects, replacement of lysine by non-ionic linkages is desirable.

3.15. CONCLUSIONS

The N7-guanine nucleobase has been successfully incorporated into PNA using simple chemistry and solid phase methodology. It has proved to be a good C^+ mimetic at neutral pH, with no difference in Tm being observed with DNA **89** and DNA **90** at

neutral and acidic pH. Thus, triplex formation is pH-dependent when *aeg*/*aep*PNA-C units are used, while *aeg*/*aep*PNA-N7G units permit triplex formation at physiological pH.

In addition, the introduction of *aep*PNA units in the backbone of the hairpin bisPNAs, which form triplexes, strongly influenced the recognition of DNA in an orientation selective manner. The more stable complexes were formed when the arm A was antiparallel to the central strand. Among all, the most stable triplexes were observed for *aep*PNA-N7G-containing bisPNAs at neutral pH with DNA **90**.

3.16. EXPERIMENTAL

The synthesis of the *aeg*PNA-T/C monomers and the *aep*PNA-N7G/C monomers was carried out according to the procedures described in Chapter 2.

3.16.1. UV-melting

The concentration of the PNA oligomers was calculated on the basis of the absorption at 260nm, assuming the molar extinction coefficients of the nucleobases to be as in DNA, i. e., T, 8.8 $cm^2/\mu mol$; C, 7.3 $cm^2/\mu mol$; G, 11.7 $cm^2/\mu mol$ and A, 15.4 $cm^2/\mu mol$. The hairpin bisPNA oligomer (**79**-**82**) and the relevant complementary DNA oligonucleotide (**89**/**90**) were mixed together in a 1:1 molar ratio in 0.01M sodium phosphate buffer, pH 5.8 or 7.4 to get a final strand concentration of 1μM. The samples were annealed by heating at 85 ^oC for 1-2 min, followed by slow cooling to room temperature, kept at room temperature for ~30 min and then, refrigerated overnight. UV experiments were performed on a Perkin Elmer λ15 UV-VIS spectrophotometer fitted with a Julabo temperature programmer and a Julabo water circulator. The samples were heated at a rate of 0.5 ^oC rise per minute and the absorbance at 260nm was recorded at every minute. The percent hyperchromicity at 260nm was plotted as a

function of temperature and the melting temperature was deduced from the peak in the first derivative plots.

3.17. REFERENCES

1. Nielsen, P. E.; egholm, M.; Buchardt, O. *Science* **1991**, *254*, 1497.
2. Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, 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. Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5.
4. Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem.* **2000**, *4*, 931.
5. Puglisi, J. D.; Tinoco, I. Jr. *Methods Enzymol.* **1989**, *180*, 304.
6. Barawkar, D. A.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 10418.
7. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Beg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677.
8. Gray, D. M.; Ratliff, R. L.; Vaughan, M. R. *Methods Enzymol.* **1992**, *211*, 389.
9. Gray, D. M.; Hung, S. – H.; Johnson, K. H. *Methods Enzymol.* **1995**, *246*, 19.
10. Wittung, P.; Nielsen, P. E.; Buchardt, O.; Egholm, M.; Nordén, B. *Nature* **1994**, *368*, 561.
11. Linkletter, B. A.; Szabo, I. E.; Bruice, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 3888.
12. *Chemical Synthesis and Structural Studies on Non Sugar-Phosphate Backbone Nucleic Acids: Oligoethylenoxy Linked Bis-Thymines and Prolyl Nucleic Acids*, Ph. D. dissertation of B. P. Gangamani **1997**, *University of Pune*, pp105
13. Nielsen, P. E.; Christensen, L. *J. Amer. Chem. Soc.* **1996**, *118*, 2287.
14. Kim, S. H.; Nielsen, P. E. ; Egholm, M.; Buchardt, O. *J. Am. Chem. Soc*. **1993**, *115*, 6477.
15. Lowe, G.; Vilaivan, T.; Westwell, M. S. *Bioorg. Chem.* **1997**, *25*, 321- 329

16. Panasik, N. Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262.

17. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordén, B.; Nielsen, P. E. *Nature* **1993**, 566.

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

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

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

21. Petersen, K. H.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 793.

22. Finn, P. J.; Gibson, N. J.; Fallon, A.; Hamilton, A.; Brown, T. *Nucleic Acids Res.* **1996**, *24*, 3357.

23. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 2796.

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

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

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

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

28. Lowe, G.; Vilaivan, T. *J. Chem. Soc. Perkin Trans. I* **1997**, 555.

29. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, *55*, 177.

30. Howarth, N. M.; Wakelin, L. P. G. *J. Org. Chem.* **1997**, *62*, 5441.

31. Altmann, K. –H.; Chiesi, C. S.; Garciá- Echeverría, C. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1119.

32. Zhang, L.; Min, J.; Zhang, L. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2903.

33. van der Laan, A. C.; van Amsterdam, I.; Tesser, G. I.; van Boom, J. H.; Kuyl-Yeheskiely, E. *Nucleosides & Nucleotides* **1998**, *17*, 219.

34. Fujii, M.; Yoshida, M.; Hidaka, J.; Ohtsu, T. *Chem. Commun.* **1998**, 717.

35. Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999**, 197.

36. Norden, B. *Chem. Scr.* **1975**, *7*, 14.

37. Norden, B. *Acta Chem. Scand.* **1972**, *26*, 111.

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

39. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, *55*, 177.

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

41. Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895.

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

43. Miller, P. S.; Bhan, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788.

44. Jetter, M. C.; Milligan, J. F.; Wadwani, S.; Moulds, C.; Froehler, B. C.; Matteucci, M. D. *Proc. Natl. Acad. Sci. USA* **1992b,** *89*, 3761.

45. Davison, E. C.; Johnsson, K. *Nucleosides Nucleotides* **1993**, *12*, 237.

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

47. Ono, A.; Ts'o, P. O. P.; Kan, L. J. *J. Org. Chem.* **1992**, *57*, 3225.

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

49. 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*, 831.

50. Rao, T. S.; Durland, R. H.; Revankar, G. R. *J. Heterocyclic Chem.* **1994**, *31*, 935.

51. Hunziker, J.; Priestley, E. S.; Brunar, H.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 2661.

52. Brunar, H.; Dervan, P. B. *Nucleic Acids Res.* **1996**, *24*, 1987.

53. St. Clair, A.; Xiang, G.; McLaughlin, L. W. *Nucleosides Nucleotides* **1998**, *17*, 925.

54. Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624.

55. Ishihara, T.; Corey, D. R. *J. Am. Chem. Soc.* **1999**, *121*, 2012.

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

3.18. APPENDIX

- CD spectra of (2*S*,4*S*) and (2*R*,4*S*) *aep*PNA single strands **69-78**

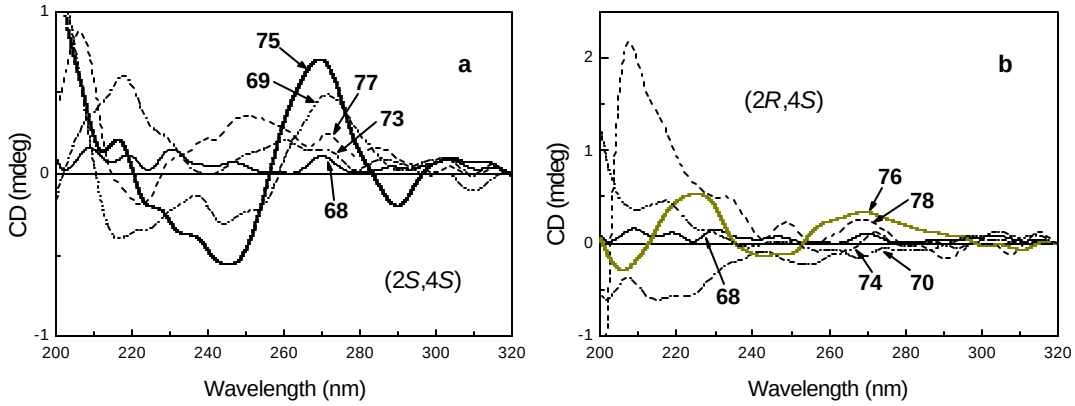


Figure 16. CD profiles of single strands **a**. (2*S*,4*S*) *aep*PNAs, **69**, **73**, **75**, **77** and **b**. (2*R*,4*S*) *aep*PNAs **70**, **74**, **76**, **78** mixed sequences and the achiral *aeg*PNA **68**.

CHAPTER 4

SYNTHEIS AND HYBRIDIZATION APPLICATION OF 4(S)-(N-Boc-amino)-2(S/R)-(purinyl/pyrimidinyl-methyl)-pyrrolidine-N1-acetic acid BASED PNA OLIGOMERS

4.1. INTRODUCTION

The very favourable binding affinity and specificity of PNA for complementary DNA/RNA[1] has sparked off the quest for compounds that would further improve its applicability by enhancing water-solubility and permeability across cell membranes. The basic aminoethylglycyl structure of PNA has been variously modified with this aim, with varying degrees of success. There have been only a few *de novo* modified oligonucleotides with positively charged backbones. In this context, the DNG-PNA chimera[2] (Figure 1a) comprising positively charged guanidinium linkages have been synthesized. This led to destabilization of the complexes with complementary nucleic acids, probably due to severe structural changes in the backbone of PNA. Another effort in this direction is embodied in the pyrrolidine-amide oligonucleotide mimic[3] (POM, Figure 1b) as mentioned in Chapter 2, where a kinetically selective binding was achieved for RNA over DNA. The (2*R*,4*S*) adenine monomer of the POM[4] was subsequently reported by Püschl *et al*. and the derived homodecamer was found to form a DNA_2:PNA triplex, with PNA as the central strand.

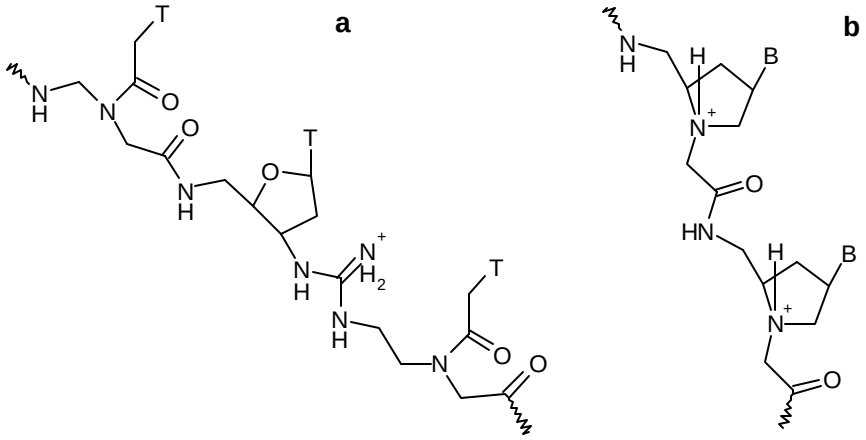


Figure 1. **a**. DNG-PNA chimera and **b**. Pyrrolidinamide Oligonucleotide Mimic

As reported in Chapter 2, positive charge was conferred on the PNA backbone upon replacing the rigid acetamide linker to the nucleobase by a relatively flexible ethylene

linker[5] (Eth-T PNA, Figure 2a). This, however, did not prove beneficial in terms of enhancing the thermal stability, although an improved water-solubility was achieved.

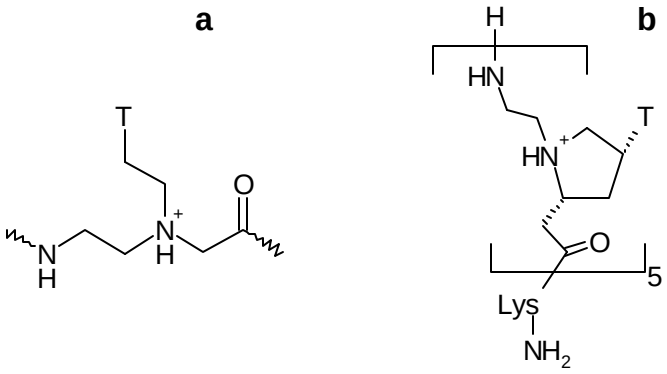


Figure 2. **a**. Eth-T PNA and **b**. (2*R*,4*R*) amionethylprolyl PNA-T_5

A highly successful modification of PNA in terms of its binding affinity and selectivity has been the aminoethylprolyl PNA reported in Chapters 2 & 3. Very recently, the synthesis of the (2*R*,4*R*) *aep*PNA-T monomer[6] and its incorporation into oligomers was carried out (Figure 2b). Interestingly, this pentamer did not show binding with polydeoxyadenylic acid (poly dA) and surprisingly bound to polyadenylic acid (poly A) forming a 2:1 hybrid with high affinity and specificity.

The above mentioned literature reports underlined the necessity for fine-tuning the PNA structure to achieve maximum advantage.

4.2. RATIONALE FOR THE PRESENT WORK

Efforts to achieve optimum fine-tuning of the PNA structure initiated in the preceding chapters are herewith extended in another direction. Retaining the concept of introducing a positive charge into the PNA backbone, the acetamide linker to the nucleobase is eliminated. Instead, structural/conformational constraint is introduced into the molecule in the form of a pyrrolidine ring. Here, the aminoethyl moiety of the PNA is restricted. This structure can be derived from aminoethylglycyl PNA by bridging the β''-carbon atom of the ethylenediamine segment and the α'-carbon atom of the

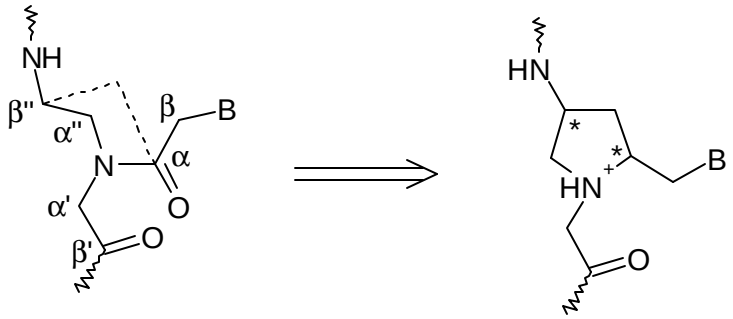


Figure 3. Derivation of the pyrrolidyl PNA structure from the parent *aeg* PNA structure

linker to the nucleobase by a methylene group, with a simultaneous omission of the α-carbonyl group (Figure 3). This leaves the nucleobase attached to the pyrrolidine ring *via* a flexible methylene group. The tertiary ring nitrogen atom is expected to bear a positive charge, thus enhancing the water-solubility. The institution of constraint in this way also creates two stereocenters per unit, which generates scope for introduction of chirality into oligomers. The modulation of the stereochemistry at these centres could lead to structures with inherent structural preferences. The monomeric units may be prepared from the versatile naturally-occurring 4(*R*)-hydroxy-2(*S*)-proline (Figure 4).

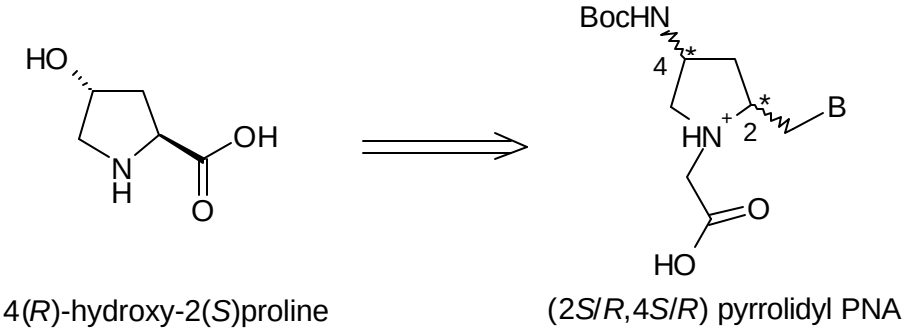
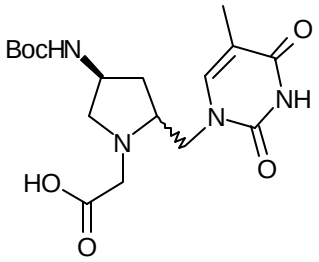


Figure 4. Schematic derivation of the pyrrolidyl PNA unit from the starting material, 4(*R*)-hydroxy-2(*S*)-proline

The conjugation of polycations such as spermine to oligonucleotides has been found to influence the properties of DNA duplexes and triplexes.[7,8] Spermine conjugation to PNA has also been reported to increase the water-solubility and binding efficiency of PNA due to the terminal positive charges.[9]

The objectives of this chapter are:

(i) Synthesis of 4(*S*)-(*N*-Boc-amino)-2(*S*/*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N*1-acetic acid diastereoisomers for PNA synthesis.



11 (2*S*,4*S*)
21 (2*R*,4*S*)

Pyrrolidyl PNA-T monomeric units

(ii) Solid phase peptide synthesis of oligomers incorporating 4(*S*)-(*N*-Boc-amino)-2(*S*/*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-acetic acid monomers in aminoethylglycyl PNA oligomers.

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

(iv) Biophysical studies of the PNA oligomers and their complexes with complementary nucleic acids.

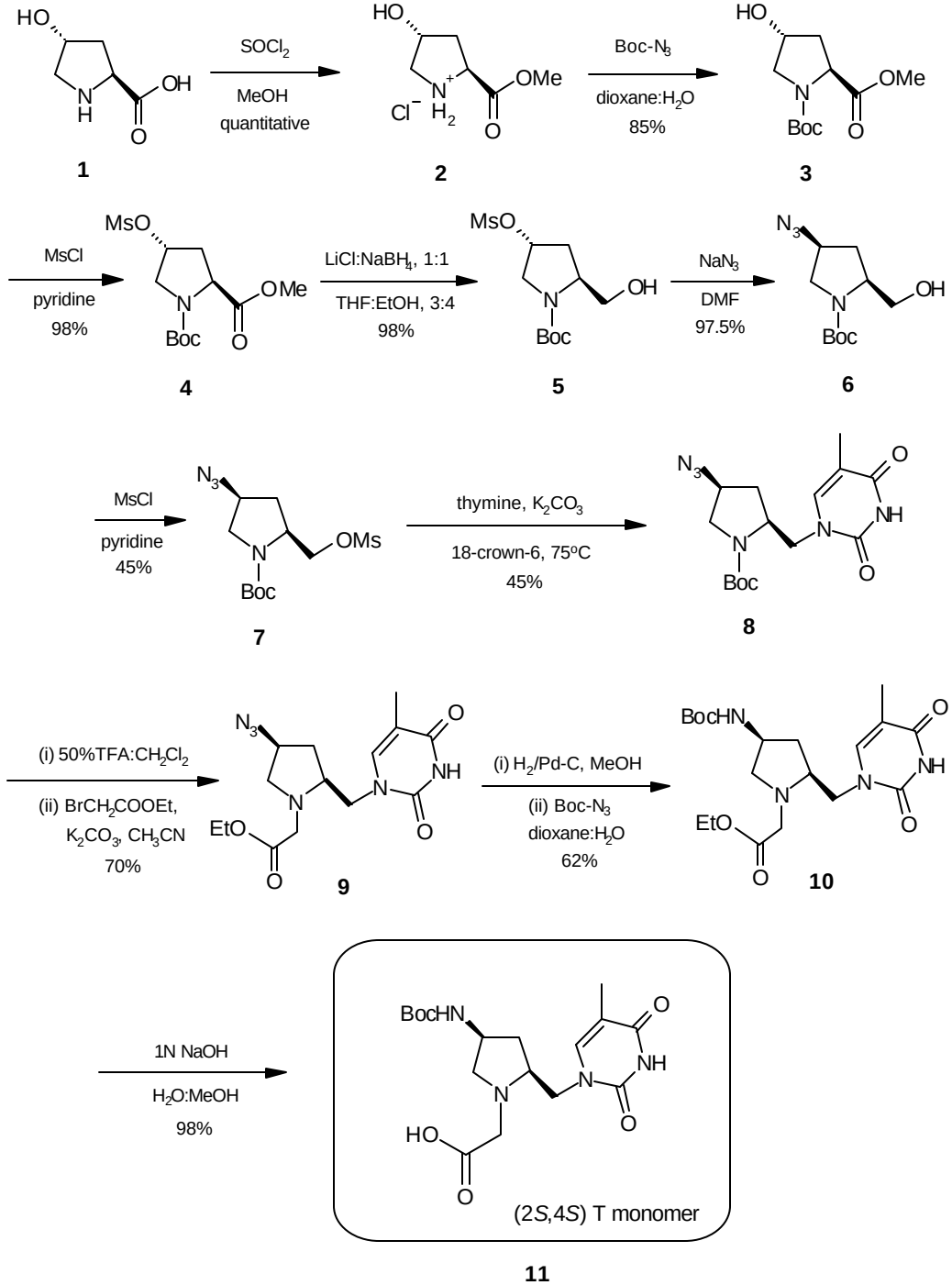
4.3. PRESENT WORK

4.3.1. Synthesis of the Pyrrolidyl PNA Monomers

4.3.1a Synthesis of 4(S)-(N-Boc-amino)-2(S)-(thymin-1-ylmethyl)-pyrrolidine-N1-acetic acid ***11***. The synthesis of 4(*S*)-(*N*-Boc-amino)-2(*S*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-acetic acid was carried out starting from 4(*R*)-hydroxy-2(*S*)-proline **1** (Scheme 1), further endorsing the versatility of this naturally occurring amino acid as a starting material in organic synthesis.[10]

4(*R*)-hydroxy-2(*S*)-proline was converted to its methyl ester by refluxing in methanol in the presence of thionyl chloride. The resulting 4(*R*)-hydroxy-2(*S*)-proline methyl ester

Scheme 1. Synthesis of 4(*S*)-(*N*-Boc-amino)-2(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-acetic acid



1 **2** **3** **4** **5** **6** **7** **8** **9** **10** **11**

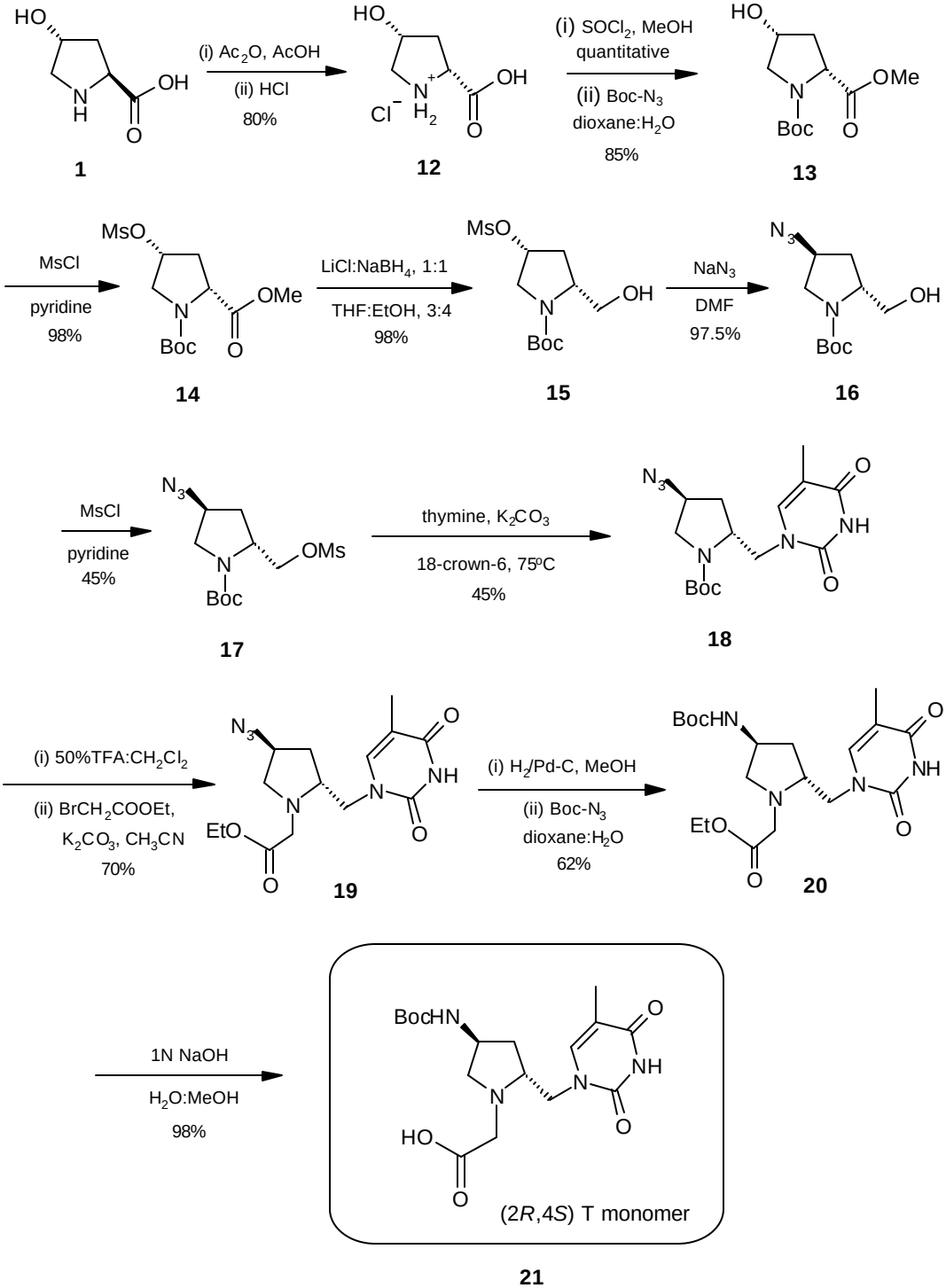
2 isolated as a hydrochloride, which upon subsequent treatment with Boc-azide in dioxane:water in the presence of triethylamine, yielded the carbamate-protected

pyrrolidine ring nitrogen to get 1-(*N*-Boc)-4(*R*)-hydroxy-2(*S*)-proline methyl ester **3**. The 4-hydroxyl group was converted to its corresponding mesyl derivative **4** upon treatment with methanesulphonylchloride in dry pyridine. The ester function in compound **4** was reduced to **5** by *in situ* generation of lithium borohydride[11] using an equimolar mixture of sodium borohydride and lithium chloride in THF:ethanol, 3:4. This was evident from the disappearance of the ester carbonyl resonance in the ^{13}C NMR and the OCH_3 singlet at δ 3.75 in the ^{1}H NMR spectrum. The 1-(*N*-Boc)-4(*R*)-*O*-mesyl-pyrrolidine-2(*S*)-methyl alcohol **5** was converted to the 4(*S*)-azido derivative by reaction with sodium azide in DMF. The presence of the azide group in **6** was confirmed by a sharp peak at 2100 cm^{-1} in the IR spectrum. This step is accompanied by a change in the stereochemistry at C4 as a result of S_N2 inversion. The primary hydroxyl group of this compound was then mesylated to **7** using mesyl chloride in dry pyridine and the success was evident from the appearance of a characteristic singlet at δ 3.03 in the ^{1}H NMR spectrum due to the sulphonyl methyl group. The mesyl derivative **7** upon reaction with thymine in the presence of K_2CO_3 and 18-crown-6 as a catalyst, furnished the 1-(*N*-Boc)-4(*S*)-azido-2(*S*)-(thymin-1-ylmethyl)-pyrrolidine **8**. This was accompanied by the appearance of the characteristic thymine peaks in the ^{1}H NMR spectrum, *viz*. at δ 7.00 (*T*-H6) and δ 1.90 (*T*-CH_3). The pyrrolidine ring nitrogen was deprotected using 50 % TFA:CH_2Cl_2 and the TFA salt was neutralized to the free amine. This being unstable, was immediately alkylated by ethylbromoacetate in the presence of potassium carbonate in dry acetonitrile to yield 4(*S*)-azido-2(*S*)-(thymin-1-ylmethyl)-pyrrolidine-*N*1-ethyl acetate **9** in good yield. The appearance of the signals corresponding to the ethyl ester (δ 4.15 and δ 1.25) in the ^{1}H NMR, and the carbonyl group (δ 170.2) in the ^{13}C NMR confirmed the transformation. The 4-azido function was reduced by catalytic hydrogenation using 10% Pd-C in methanol to obtain the free amine, which was isolated and immediately protected as its corresponding Boc-

derivative **10** using Boc-azide and triethylamine in dioxane:water. The reduction to amine was confirmed by the disappearance of the azide peak at 2100 cm^{-1} in the IR spectrum. The ethyl ester function in **10** was hydrolyzed using NaOH in aqueous methanol to yield the (2*S*,4*S*) pyrrolidine T-monomer **11** suitable for use in solid phase synthesis. This synthetic scheme involving twelve steps, including a single inversion at C4, results in the overall transformation of 4(*R*)-hydroxy-2(*S*)-proline to 4(*S*)-(*N*-Boc-amino)-2(*S*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-acetic acid **11**.

*4.3.1b Synthesis of 4(S)-(N-Boc-amino)-2(R)-(thymin-1-ylmethyl)-pyrrolidine-N1-acetic acid **21***. The synthesis of 4(*S*)-(*N*-Boc-amino)-2(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N*1-acetic acid **21** was also achieved starting from 4(*R*)-hydroxy-2(*S*)-proline **1** (Scheme 2). The C2-stereocenter was epimerized in the very first step[12] (Scheme 2) by treating the 4(*R*)-hydroxy-2(*S*)-proline with acetic acid and acetic anhydride to yield the *cis* lactone, which was hydrolyzed by 2N HCl without isolation to 4(*R*)-hydroxy-2(*R*)-proline hydrochloride **12** in good yield. This was followed by a similar series of reactions as with the 2(*S*) isomer. The carboxylic acid and amine groups were protected as the methyl ester and Boc-derivatives respectively by treatment with dry methanol in the presence of thionyl choride, followed by reaction with Boc-azide in dioxane:water in the presence of triethylamine. The hydroxyl group in 1-(*N*-Boc)-4(*R*)-hydroxy-2(*R*)-proline methylester **13** was transformed to its mesyl derivative **14** by reaction with mesyl chloride in pyridine. The methyl ester was reduced to the corresponding alcohol **15** using a 1:1 mixture of sodium borohydride and lithium chloride in THF:ethanol. The successful reduction was confirmed by the absence of the carbonyl resonance in ^{13}C NMR at δ 171.0 and the appearance of a broad multiplet corresponding to two additional protons at δ 3.50-3.90 in the ^{1}H NMR spectrum. S_N2 inversion at the C4-stereocenter was achieved by treating with sodium azide in DMF,

Scheme 2. Synthesis of 4(*S*)-(*N*-Boc-amino)-2(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-acetic acid



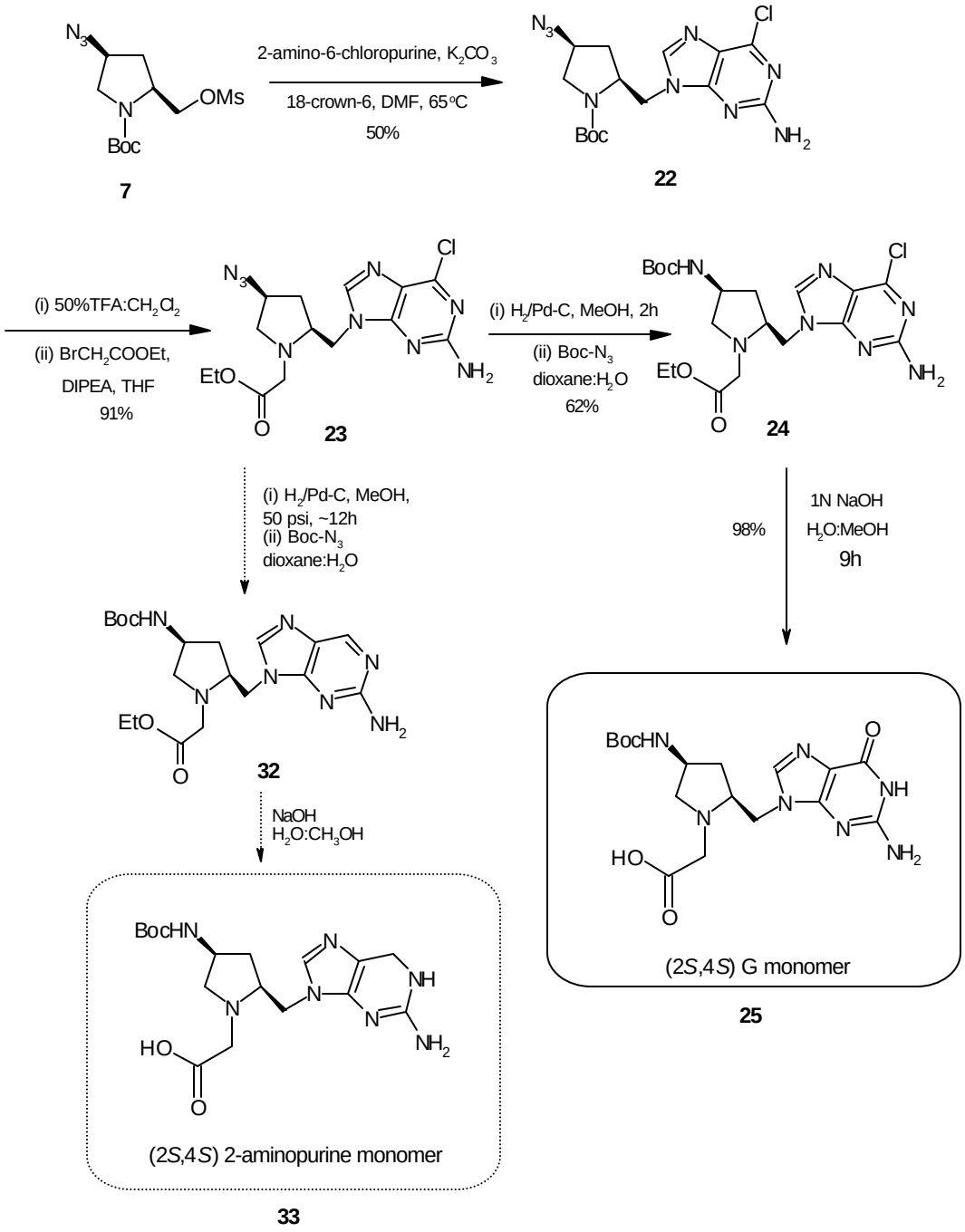
supported by the appearance of the characteristic azide signal at 2100 cm^{-1} in the IR spectrum, and an upfield shift of the H4 resonance from δ 5.18 in the 4(*R*)-*O*-mesyl derivative **15** to δ 3.70 in the 4(*S*)-azide **16**. The primary hydroxyl group at C2 was mesylated to furnish the intermediate **17**, which upon reaction with thymine in the presence of K_2CO_3 and the catalyst 18-crown-6, was transformed into 1-(*N*-Boc)-4(*S*)-azido-2(*R*)-(thymin-1-ylmethyl)-pyrrolidine **18**. The pyrrolidine ring nitrogen was deprotected with TFA:CH_2Cl_2 and subsequently alkylated using ethylbromoacetate to yield 4(*S*)-azido-2(*S*)-(thymin-1-ylmethyl)-pyrrolidne-*N1*-ethylacetate **19** in good yield. The 4(*S*)-azide group was converted to the 4-(*N*-Boc) derivative **20** after reduction with Pd-C and Boc-protection. The ethyl ester function was saponified to liberate the acid and yield the protected (2*R*,4*S*) pyrrolidine T-monomer **21** suitable for solid phase peptide synthesis.

This synthesis includes two inversion steps, one at the C2 and the other at C4 to yield 4(*S*)-(*N*-Boc-amino)-2(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-acetic acid.

*4.3.1.c. Synthesis of 4(S)-(N-Boc-amino)-2(S)-(guanin-9-ylmethyl)-pyrrolidine-N1-acetic acid **25**.* 4(*S*)-(*N*-Boc-amino)-2(*S*)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine-*N1*-ethyl acetate **24** was synthesized as a common precursor to the guanine and 2-aminopurine pyrrolidine PNA monomers (Scheme 3). The guanine monomer **25** can be reached by simple hydrolysis using NaOH, while the 2-aminopurine monomer **33** can be synthesized by catalytic hydrogenation. Due to the paucity of time, the 2-aminopurine monomer could not be synthesized and this section describes the synthesis of only the pyrrolidyl G-monomer.

The synthesis of the (2*S*,4*S*) isomer **25** of guanine was achieved starting from 4(*R*)-hydroxy-2(*S*)-proline **1** following a series of reactions that were similar to those employed for the (2*S*,4*S*) thymine monomer **11**. The intermediate **7** was converted to

Scheme 3. Synthesis of 4(*S*)-(*N*-Boc-amino)-2(*S*)-(guanin-9-ylmethyl)-pyrrolidine-*N1*-acetic acid **25** and 4(*S*)-(*N*-Boc-amino)-2(*S*)-(2-aminopurin-9-ylmethyl)-pyrrolidine-*N1*-acetic acid **33**



the 1-(*N*-Boc)-4(*S*)-azido-2(*S*)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine **22** upon reaction with 2-amino-6-chloropurine. The derivative **22** was converted into the 4(*S*)-azido-2(*S*)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine-*N*1-ethyl acetate **23** in excellent yield. The azide was reduced to the corresponding amine by controlled catalytic hydrogenation to keep the 2-amino-6-chloropurine moiety intact. The resulting amine at C4 being unstable, was immediately protected as its Boc derivative, yielding the fully protected compound **24**. The ester function in **24** was hydrolyzed using aqueous methanolic NaOH for an extended time period to achieve simultaneous oxidation to the 6-oxo-2-aminopurine derivative and yield the (2*S*,4*S*) pyrrolidyl G-monomer **25** suitable for use in solid phase synthesis.

4.3.2. Solid Phase Peptide Synthesis

In the synthetic strategy for the pyrrolidyl monomers, the Boc-protecting group was utilized for the amino group[13,14] and PNA oligomer synthesis was carried out from the 'C' to the 'N' terminus using monomers with free carboxylic acid functions and amino functions protected as Boc derivatives that are cleavable with TFA at the beginning of every cycle. The solid support used was Merrifield resin[15] that was derivatized with β-alanine using the cesium salt method[16] (See Chapter 2) to achieve a loading value of 0.5 meq/g resin. β-alanine is linked to the resin *via* a benzyl ester linkage enabling cleavage of the synthesized oligomer by either acid or amine to afford the 'C' terminal free acid or amide respectively.

The resin, upon attachment of the β-alanine spacer amino acid, was suitably down-loaded to attain a loading capacity of 0.1 to 0.2 meq/g resin (Please refer to Chapter 2), which has been reported to give optimum yields in PNA synthesis. The loading capacity of the resin was determined by the non-destructive picrate assay.[17] The synthesis was carried out by repetitive coupling cycles, each cycle comprising (I) deprotection of the Boc-amino group using TFA in dichloromethane to generate the

amine group as TFA salts. (ii) neutralization of the resulting TFA salt by DIPEA to liberate the amine. (iii) coupling of the resin-bound amino groups with the incoming amino acid using HOBt and DIPCDI as the condensing agents. The deprotection and coupling steps were monitored by the ninhydrin method using Kaiser's test.[18]

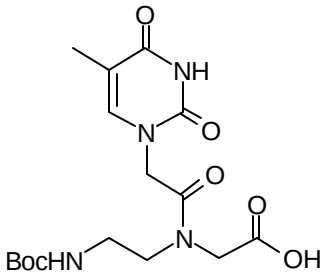
4.3.2a Synthesis of cationic pyrrolidyl peptide nucleic acids

Table 1. Resin-linked PNA Oligomers Synthesized.

Entry	Resin-linked PNA Oligomer	
1	MF--β-ala- T T T T T T T T -Boc	Control aegPNA
2	MF--β-ala- T T T T T T T **t** -Boc	One (2*S*,4*S*) pyrrolidyl unit at N.T.
3	MF--β-ala- T T T T T T T **t**-Boc	One (2*R*,4*S*) pyrrolidyl unit at N.T.
4	MF--β-ala- T T T T **t** T T T -Boc	One internal (2*S*,4*S*) pyrrolidyl unit
5	MF--β-ala- T T T T **t** T T T -Boc	One internal (2*R*,4*S*) pyrrolidyl unit
6	MF--β-ala- A A A A A A A A -Boc	Complementary to PNA-T_8

T/A = *aeg*PNA-T/A; **t** = (2*S*,4*S*) pyrrolidyl PNA-T; **t** = (2*R*,4*S*) pyrrolidyl PNA-T

For preliminary binding studies with complementary DNA sequences, PNA-T_8 octamers (Table 1) were synthesized. The achiral, neutral control aminoethylglycyl (*aeg*) PNA-T_8 oligomer was synthesized from the *aeg*PNA-T monomer[19] **34** as described in Chapter 2, Section 2.2.6.a The chiral, cationic pyrrolidyl units were incorporated into the *aeg*PNA oligomer **46** (Table 1, entry 1) at pre-determined positions.



*Aeg*PNA-T **34**

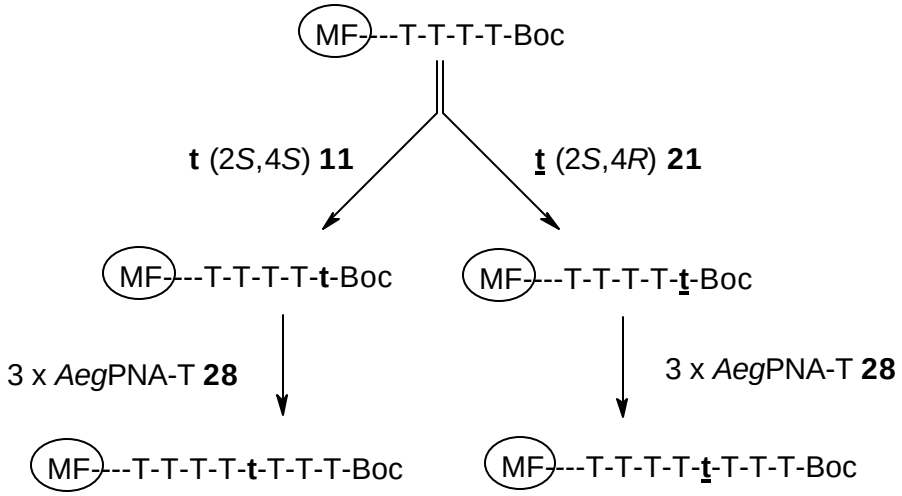
Synthesis of chiral, cationic pyrrolidyl PNAs bearing a terminal modified unit

The synthesis of the PNAs bearing one pyrrolidyl unit at the 'N' terminus (entries 2 & 3) was carried out from resins having an *aeg*PNA-T heptamer. The resin was divided into two equal portions and separately coupled with **11** (2*S*,4*S*) or **21** (2*R*,4*S*) to get the corresponding pyrrolidyl PNA oligomers (Table 1, entries 2 & 3 respectively). The coupling efficiency with the pyrrolidyl units was as good as with the *aeg* units (>90-95%), making the capping step after each coupling cycle unnecessary.

Synthesis of chiral, cationic pyrrolidyl PNAs bearing a pyrrolidyl unit in the interior of the sequence

The synthesis of the PNAs with the pyrrolidyl unit in the centre (entry 4 and 5, Table 1) was carried out using the resin bearing four *aeg*PNA-T units as in Scheme 4. The resin was split into two portions, which were individually coupled with either **11** (2*S*,4*S*) or **21** (2*R*,4*S*) pyrrolidyl PNA-T respectively and further extended by adding another three *aeg*PNA-T units to achieve the syntheses of the desired octamers. The coupling efficiency at each step was determined to be >90% and no capping by acetylation was necessary.

Scheme 4 Synthesis of cationic pyrrolidyl PNAs with a single pyrrolidyl unit in the center of the sequence

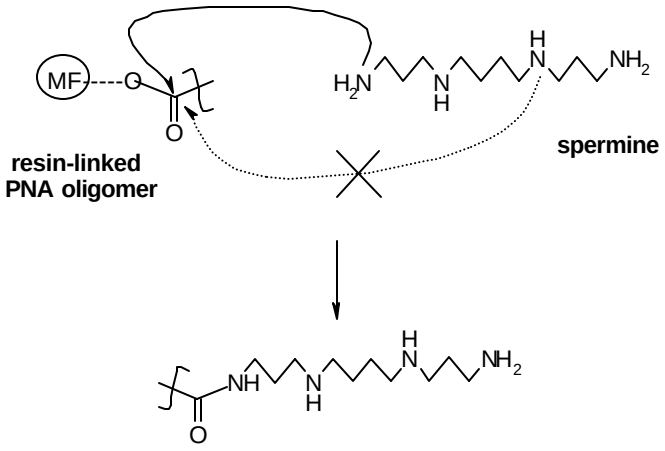


4.3.3. Cleavage of the Oligomers from the Solid Support

4.3.3a Cleavage by Acid. The synthesized PNA oligomers were cleaved from the solid support using the well-known TFMSA-TFA cleavage procedure.[15b] This yielded oligomers with free carboxylic acids at their 'C' termini (PNAs **26-29**). Since the oligomers cleaved were homothymine sequences having no exocyclic protecting groups employed for the other nucleobases, additional deprotection procedures like ammonia treatment were not necessary.

4.3.3b Cleavage by Aminolysis.[20] The resin-linked PNA oligomers (entries 2 and 3) were treated with neat spermine at 55 ^{o}C for 24h to obtain oligomers **30** and **31** as their carboxyl-sperminamides. The sperminamide is expected to arise from the attack of the primary amine rather than the secondary amino group due to steric reasons (Scheme 5), as has been established in earlier reports.[21]

Scheme 5. Aminolysis with spermine to afford the 'C' terminal sperminamide



4.3.4. Purification of the PNA Oligomers

Preliminary purification of the PNA oligomers from shorter sequences and low molecular weight impurities was achieved by gel filtration through Sephadex G25. These were subsequently purified by FPLC on a semi-preparative C8 reverse-phase column using an ascending gradient of acetonitrile in water containing 0.1% TFA. The purity of the oligomers was checked by HPLC on a C18 reverse-phase column and confirmed by MALDI-TOF mass spectroscopy. Some representative HPLC profiles and mass spectra are shown in Figures 5 and 6 respectively. The PNA sequences are listed in Table 2.

Table 2. PNA Oligomer Sequences.

PNA	Sequence Composition	
46	H- T T T T T T T T -NH-$(CH_2)_2$-COOH	AegPNA-T_8 (control)
63	H- A A A A A A A A -NH-$(CH_2)_2$-COOH	Complementray to PNA-T_8
tn 26	H- **t** T T T T T T T -NH-$(CH_2)_2$-COOH	(2*S*,4*S*) pyrrolidyl unit at N. T.
tn 27	H- **t** T T T T T T T -NH-$(CH_2)_2$-COOH	(2*R*,4*S*) pyrrolidyl unit at N. T.
tc 28	H- T T T **t** T T T T -NH-$(CH_2)_2$-COOH	Central (2*S*,4*S*) pyrrolidyl unit
tc 29	H- T T T **t** T T T T -NH-$(CH_2)_2$-COOH	Central (2*S*,4*S*) pyrrolidyl unit
tn-sp 30	H-**t**TTTTTTT-NH-$(CH_2)_2$-CONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$	(2*S*,4*S*) pyrrolidyl unit at the N.T. with 'C' terminal sperminamide
tn-sp 31	H-**t**TTTTTTT-NH-$(CH_2)_2$-CONH$(CH_2)_3$NH$(CH_2)_4$NH$(CH_2)_3NH_2$	(2*S*,4*S*) pyrrolidyl unit at the N.T. with 'C' terminal sperminamide

4.3.5. Synthesis of DNA Oligonucleotides

The DNA oligonucleotides **84-86** (Chapter 2) were synthesized on a Pharmacia Gene Assembler Plus automated synthesizer using the standard β-cyanoethyl phosphoramidite chemistry on a CPG solid support, followed by ammonia deprotection and cleavage.[22] The purity of the oligonucleotides was verified by reverse phase HPLC on a C18 column and was found to be >96%. Consequently, these oligonucleotides were used for the biophysical studies as such, without further purification.

Table 3. DNA Oligomer Sequences

DNA		Oligomer Sequences 5' ® 3'
84	G C A A A A A A A A C G	DNA complementary to PNA-T_8 **46** & **26**-**31** with CG clamps
85	T T T T T T T T	DNA complementary to **84**
86	G C A A A T A A A A C G	Mismatch DNA for PNAs **46** & **26**-**31**

4.4. BIOPHYSICAL STUDIES OF THE PYRROLIDYL PNAs AND THEIR HYBRIDS WITH COMPLEMENTARY NUCLEIC ACIDS

The efficacy of the PNA:DNA hybrids formed by pyrrolidyl PNA was evaluated by UV-spectroscopic and CD spectrophotometric experiments.

4.5. RESULTS

4.5.1. pH Titration

The titration curve of the pyrrolidyl-T monomeric unit **11** after the deprotection of the Boc-protecting group using TFA, with sodium hydroxide exhibited three distinct pKa values (Figure 4). These corresponded to the dissociation of the carboxylic acid at 3.54

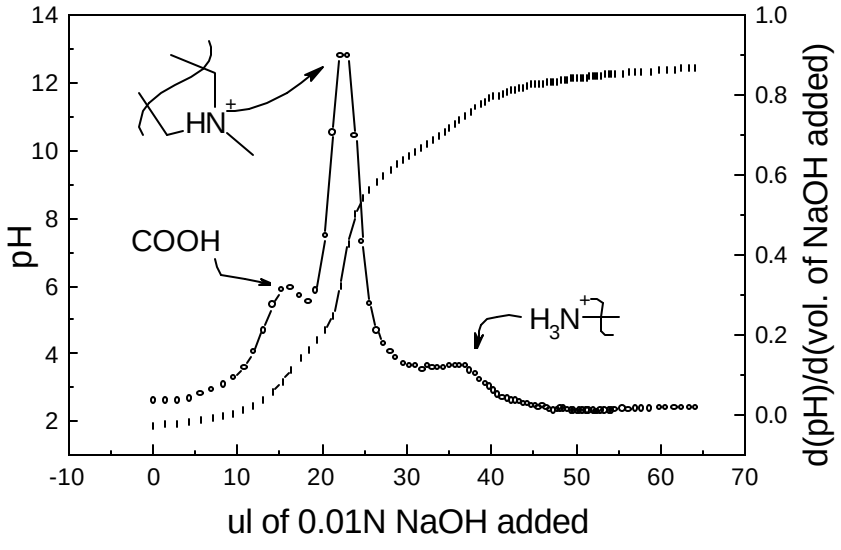


Figure 4. pH titration of (2*S*,4*S*) pyrrolidyl PNA-T unit **11** with NaOH

pH, the pyrrolidine ring deprotonation at 6.67 pH and the dissociation of the primary amine at 11.03 pH. The pKa obtained for the pyrrolidyl ring nitrogen was similar to that obtained for the *aep*PNA unit in Chapter 2. This suggested that the protonation status of the ring nitrogen in the pyrrolidyl and *aep*PNA units is similar. It would be at least partly protonated at physiological pH, i.e., the pH used in the melting experiments. The pH titration curve is shown in Figure 4.

4.5.2. UV-Tm

4.5.2a PNA:DNA Complexes

The PNA T_8 single strands **46** and **26-29** showed a very low hyperchromicity upon melting (~3-4%). The DNA single strands **84-86** also showed a hyperchromicity of ~5%. The UV-absorbance/% hyperchromicity plots of these oligomers with temperature however, were linear and failed to show any sigmoidal transition.

The stabilities of the PNA:DNA complexes differed depending on the stereochemistry in the diastereoisomer and also its position within the sequence (Table 4). Some representative UV-melting profiles are shown in Figure 5. The complexes of

Table 4. UV-Tm of PNA_2:DNA complexes

	(2*S*,4*S*)		**(2*R*,4*S*)**	
Entry	**PNA_2:DNA**	**Tm (°C)**	**PNA_2:DNA**	**Tm (°C)**
1	**tn 26:84**	59.1	**<u>t</u>n 27:84**	47.0
2	**tn-sp 30:84**	47.4	**<u>t</u>n-sp 31:84**	47.4
3	**tc 28:84**	28.4	**<u>t</u>c 29:84**	58.8
4	**tn 26:86**	44.3	**<u>t</u>n 27:86**	34.5
5	**tc 28:86**		**<u>t</u>c 29:86**	39.1
Control *aeg*PNA_2:DNA complex				
6	**46:84**		Tm = 43 °C	

tn, **<u>t</u>n** = 'N' terminal pyrrolidyl unit; **tn-sp**, **<u>t</u>n-sp** = 'N' terminal pyrrolidyl unit with 'C' terminal sperminamide; **tc**, **<u>t</u>c** = internal pyrrolidyl unit

the PNA oligomers with free 'C' terminal carboxylic acids **26-29** with DNA **84** differed from those involving the PNAs with 'C' terminal amides **30, 31**.

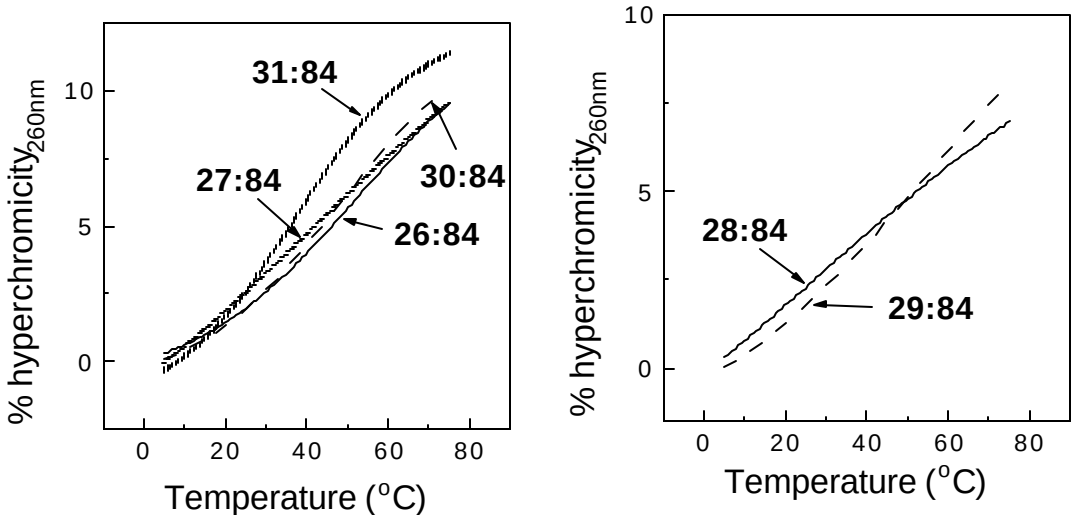


Figure 5.UV-Tm profiles of **a**. PNA:DNA complexes **26:84** & **27:84** wherein the PNAs exist as 'C' terminal carboxylic acids and in **30:84** & **31:84** as their 'C' terminal sperminamides **b**. PNA:DNA complexes **28:84** and **29:84** in which the PNA sequence carries a centrally located pyrrolidyl unit

The oligomer **27:84** with one (2*R*,4*S*) pyrrolidyl unit at its 'N' terminus gave a melting temperature stabilization of 4 °C over the control complex **46:84**. The (2*S*,4*S*) pyrrolidyl unit in **26** caused a much higher stabilization of 15 °C (Table 1, entry 1). The presence of the spermine moiety at the 'C' termini of the oligomers **30** & **31** caused an unexpected destabilization of the complexes with DNA **84** compared to the 'C' terminal free carboxylic acids **26** & **27** respectively (ΔTm = -11.7 & -5.6 °C for the (2*S*,4*S*) and the (2*R*,4*S*) isomers respectively; Table 4, entries 1 & 2).

The pyrrolidyl unit, when present in the center of the sequence exhibited UV-Tm effects that were reverse to those observed when this unit was present at the 'N' terminus of the oligomer. In this case, the (2*R*,4*S*) isomer in **29** stabilized the PNA:DNA complex **29:84** by 14.8 °C, while the (2*S*,4*S*) isomer in **28** destabilized the corresponding complex with DNA **84** by 15.6 °C (Table 4, entry 3).

The specificity of the PNA:DNA interactions was established by the mismatch studies with the DNA oligomer **86** bearing a T-T mismatch in the center of the sequence. The complexes of DNA **86** with the PNA oligomers **26-31** were found to be destabilized to a large extent (Table 4, entries 4 & 5). The oligomers **26** & **27** bearing the pyrrolidyl unit at their termini destabilized the complexes with the corresponding mismatch DNA **86** (ΔTm = -13.8 & -12.5 ^{o}C for the complexes **26:86** & **27:86** respectively). A representative example is shown in Figure 6.

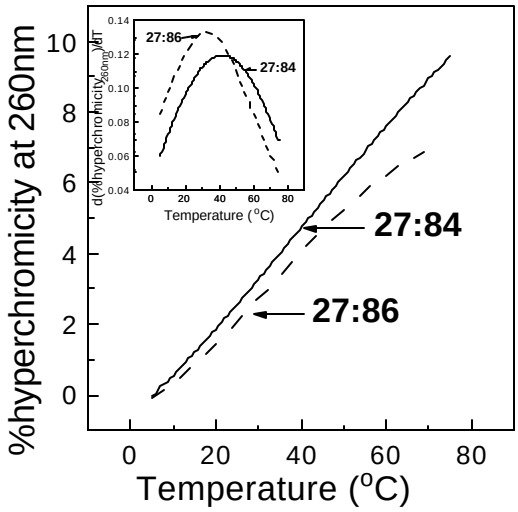


Figure 6. UV-Tm plots of the PNA:DNA complex **27:84** and the mismatched PNA:DNA complex **27:86**. The inset indicates the first derivatives of the UV-Tm plots

4.5.2b PNA:PNA Complexes

The complexes of the PNA oligomers **26-29** with the complementary PNA sequence **63** were found to possess a low Tm (Table 5). The complexes involving the PNAs **26** & **27** bearing one pyrrolidyl unit at their termini did not differ much in the melting temperatures (ΔTm = 1 ^{o}C) between the two (2*S*,4*S*) and (2*R*,4*S*) isomers **26:63** & **27:63** respectively. The Tm values were however, much lower than those obtained for the corresponding PNA:DNA complexes **26:84** & **27:84** respectively (ΔTm = -37.8 &

Table 5. UV-Tm of PNA:PNA complexes

	(2*S*,4*S*)		(2*R*,4*S*)	
Entry	**PNA:PNA**	**Tm (°C)**	**PNA:PNA**	**Tm (°C)**
1	**tn 26:63**	21.3	**tn 27:63**	
2	**tc 28:63**	21.7	**tc 29:63**	20.5

tn, **tn** = 'N' terminal pyrrolidyl unit; **tc**, **tc** = central pyrrolidyl unit

respectively). On the other hand, the hyperchromicity observed in the melting of the PNA:PNA complexes was almost double that seen for the analogous DNA:PNA complexes. A similar trend was observed in the case of the PNA oligomers with one pyrrolidyl unit in the center of the sequence, *viz*. **28:63** and **29:63** for the (2*S*,4*S*) and (2*R*,4*S*) isomers respectively (Figure 7). In this case also, the stability of the two complexes formed by PNAs **28** & **29** with PNA **63** differed by only 1 °C and were much lower than the stability of the complexes with the corresponding DNA **84** (ΔTm = -16.6 & -38.3 °C respectively for PNA **28** & **29**.

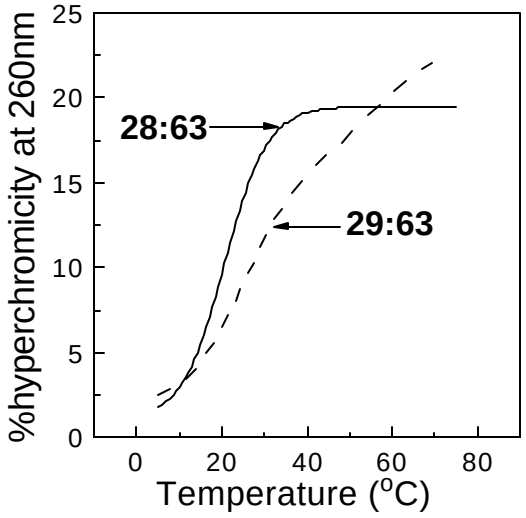


Figure 7. UV-Tm plots of PNA:PNA complexes **28:63** and **29:63**

4.5.3. CD

The PNA single strands **26** & **27** exhibited very low induced CD. The Induced CD spectra differed in sign, but were not exact mirror images of each other (Figure 8a).

The CD spectra of the PNA:DNA duplexes **26:84** & **27:84** were dominated by the CD spectrum of the constituent DNA single strand. The two complexes did not exhibit any differences in the maxima, minima or cross-over points (Figure 8b).

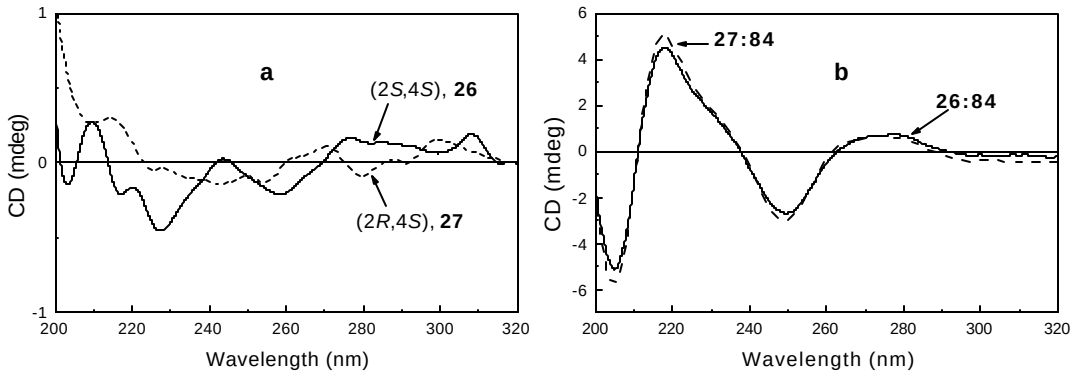


Figure 8. CD spectra of **a**. single strand PNAs **26** & **27** and **b**. PNA:DNA complexes **26:84** & **27:84**

4.6. DISCUSSION

The modification of the parent aminoethylglycyl PNA structure by the introduction of ring structures can be envisaged so as to make the resultant structure resemble either the basic DNA or the PNA structure (Figure 9). The modification of the *aeg*PNA structure discussed in Chapters 2 and 3 in the form of the *aep*PNA resembles the DNA structure with the nucleobases attached directly to a five-membered ring. The pyrrolidyl PNA, on the other hand, is more like the *aeg*PNA structure than DNA with the nucleobases present in a flexible side-chain and can be considered a true PNA mimic rather than a DNA mimic. The nucleobase is attached to the backbone *via* a two-atom spacer and the structural rigidity imposed by the tertiary amide in PNA is mimicked by the ring structure. This linker to the nucleobase is more flexible than in *aeg*PNA by

virtue of the absence of the rigid amide group, and the presence of the flexible ethylene spacer.

The pH titration of the pyrrolidyl PNA-T monomer with NaOH indicated the pKa of the pyrrolidyl ring nitrogen to be ~6.7. This suggested that this tertiary nitrogen atom would be at least partly protonated at physiological pH, i.e., the pH used in the melting experiments.

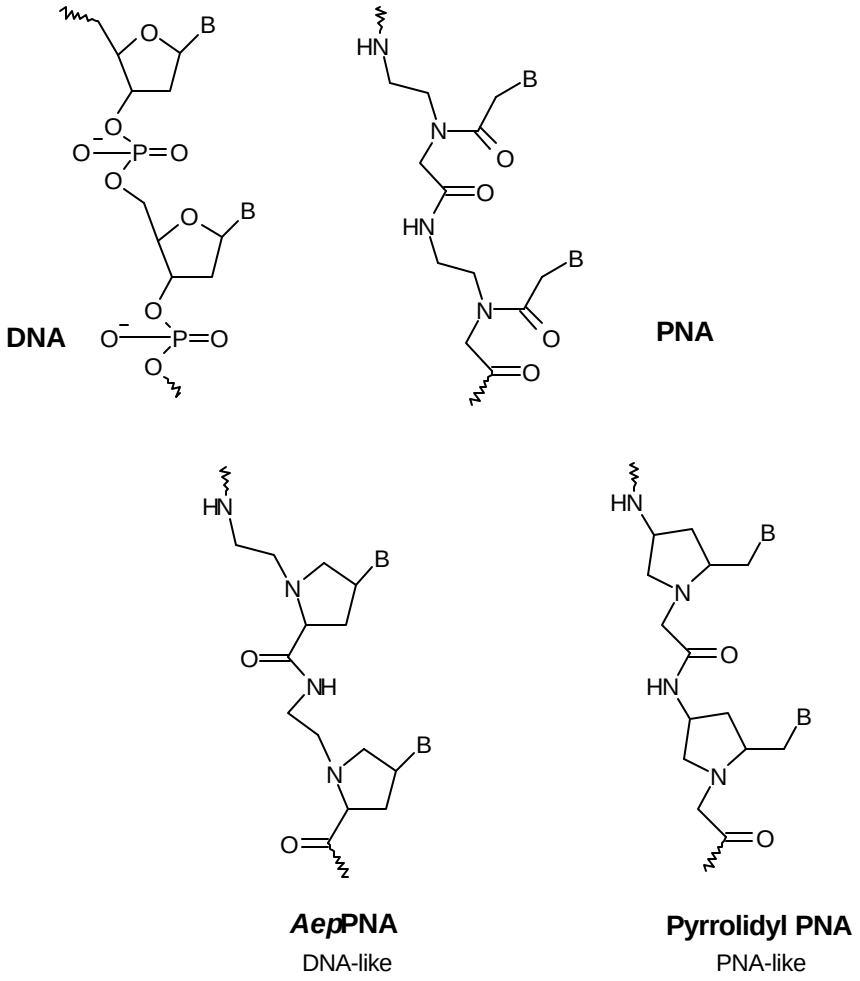


Figure 9 A comparison of the structures of *aep*PNA and pyrrolidyl PNA with those of the parent nucleic acids.

The Tm values obtained for the PNA:DNA complexes in 1:1 & 2:1 stoichiometry were almost identical and the plots of absorbance or percent hyperchromicity at 260nm Vs temperature were sigmoidal, indicative of a two-state co-operative transition, in

which the two PNA strands simultaneously dissociate from the DNA strand. These results were consistent with reports that *aeg*PNA homopyrimidine sequences are known to form PNA_2:DNA triplexes.[23]

The increased stability of the pyrrolidyl PNA:DNA complexes **26:84** & **27:84** can be explained by invoking the cationic nature of the pyrrolidyl ring nitrogen in the PNAs involved in forming these complexes. However, when these PNAs exist as their 'C' terminal sperminamides **30** & **31**, their complexes with DNA **84** are less stable (ΔTm = -6 to -12 °C). This seems to be a consequence of the terminal effects resulting from a charge repulsion between the positively charged sperminyl moiety and the protonated nitrogen of the pyrrolidine unit which face each other in a PNA_2:DNA triplex (Figure 10). The large difference between the PNAs **28** & **29** highlight the effects of stereochemical differences at C2 in the pyrrolidyl PNA units.

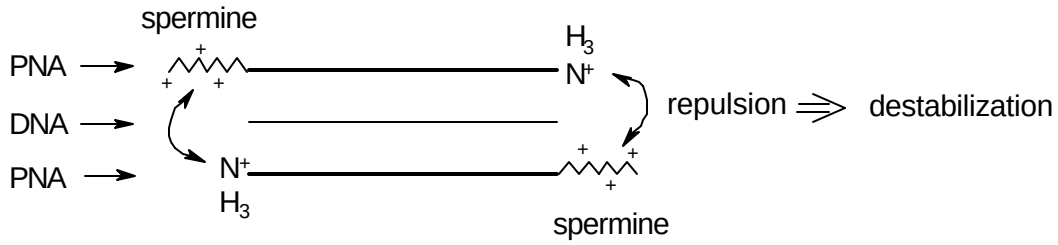


Figure 10. Terminal effects in 'C' terminal sperminamide and 'N' terminal amines of the PNAs in a PNA_2:DNA complex

The complex formation was stringent as was evident from the melting studies with a mismatch DNA sequence **86** introducing a T-T mismatch at a centrally located thymine unit in the PNA. The complexes so formed were highly unstable, leading to the conclusion that although the positive charge on the pyrrolidyl ring nitrogen does play a role in enhancing the binding to target DNA sequences, the main cause of the stabilization appears to be the specific hydrogen bonding between the complementary nucleobases in PNA and DNA.

The pyrrolidyl PNA:DNA complexes were much stabler than the corresponding PNA:PNA complexes. The hyperchromicity observed in the melting of these PNA:PNA complexes however, suggests that the base-stacking is more efficient in PNA:PNA in comparison with PNA:DNA. This could result from the pre-organization of the PNA structure before complexation itself as a manifestation of the chirality induced by the presence of two chiral centres in each pyrrolidyl PNA unit.

The pyrrolidyl PNAs exhibited enhanced solubility in aqueous media, probably a manifestation of the positive charge present in the monomeric units. The PNAs containing the pyrrolidyl PNA units remained in solution for a much longer time as compared to the uncharged *aeg*PNA oligomers, which precipitated out of solution as a result of self-aggregation.

4.7. CONCLUSIONS

The high binding affinity of the pyrrolidyl PNAs bearing one pyrrolidyl unit at the terminus is at least partly due to the protonated status of the pyrrolidyl ring nitrogen at physiological pH. The main reason for the induced stability however, remains the specific hydrogen bonding between the complementary nucleobases of PNA and DNA.

An important feature of the pyrrolidyl PNAs in terms of applicability for biological purposes is their improved water-solubility on account of the charged nature of the pyrrolidine ring nitrogen at physiological pH.

The work presented here is a preliminary study evaluating the efficiency of the pyrrolidyl PNAs at forming complexes with complementary nucleic acids. Further work is needed to completely exploit the encouraging results obtained in these studies such as capping of the termini, synthesis of monomers carrying other nucleobases, synthesis of oligomers comprising mixed *aeg* and pyrrolidyl backbones, etc..

4.8. EXPERIMENTAL

All chemicals used were of reagent or analytical grade. All the solvents used were purified according to the literature procedures[24] and the reactions were monitored for completion by TLC.

Column chromatography was performed for purification of compounds on LOBA chemie silica gel (100- 200 mesh). TLCs were carried out on pre-coated silica gel GF_{254} aluminium sheets (Merck 5554). TLCs were run in either dichloromethane with an appropriate quantity of methanol or in petroleum ether with an appropriate quantity of added ethyl acetate for most compounds. Free acids were chromatographed on TLC using a solvent system of iso-propanol: acetic acid: water in the proportion 9:1:1. The compounds were visualized with UV light and/ or by spraying with Ninhydrin reagent subsequent to Boc-deprotection (exposing to HCl vapors) and heating.

^{1}H (200 MHz/ 300MHz) and ^{13}C (50MHz) NMR spectra were recorded on a Bruker ACF 200 spectrometer fitted with an Aspect 3000 computer and all the chemical shifts are referred to internal TMS for ^{1}H and chloroform-d for ^{13}C. The chemical shifts are quoted in δ (ppm) scale. In compounds that bear a tertiary amide group, splitting of NMR signals was observed due to the presence of rotamers. In such cases, the major isomer is designated as 'maj' and the minor isomer, 'min'.

Optical rotations were measured on a JASCO DIP-181 polarimeter.

Mass spectra were recorded on a Finnigan-Matt mass spectrometer, while MALDI-TOF spectra were obtained from a KRATOS PCKompact instrument.

UV experiments were performed on a Lambda 15 UV-VIS spectrophotometer. Temperature dependent experiments were carried out by making use of a Julabo water circulator fitted with a temperature programmer.

1-(N-Boc)-4-(R)-O-mesyl-2(S)-proline benzyl ester

To a stirred solution of 1-(*N*-Boc)-4(*R*)-hydroxy-2(*S*)-proline benzyl ester (5.0g, 15.5mmol) in dry pyridine (40ml), was added dropwise, methane sulfonyl chloride (1.4ml, 18.7mmol). Upon completion of the reaction after 2h, the solvent was completely evaporated under vacuum. The residue was taken up in water and the product extracted in dichloromethane. The organic layer was dried over sodium sulphate and evaporated to give the desired product in almost quantitative yield (6.1g, 98%).

^{1}H NMR ($CDCl_3$) δ: 7.35 (s, 5H, Ph), 5.25 (m, 1H, *H*4), 5.15 (s, 2H, CH_2-Ph), 4.50 (m, 1H, *H*2), 3.30 (m, 2H, *H*5, *H*5'), 3.05 (s, 3H, SO_2CH_3), 2.60 (m, 1H, *H*3), 2.25 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 171.6 (maj) & 171.3 (min) (*C*OOBn), 153.5 (min) & 153.0 (maj) (*C*OOC$(CH_3)_3$), 78.2 (min) & 77.9 (maj) (*C*4), 66.6 (CH_2-PH), 57.1 (*C*2), 52.5 (min) & 51.9 (maj) (*C*5), 38.0 (SO_2CH_3), 36.9 (maj) & 35.8 (min) (*C*3), 27.9 (min) & 27.7 ($(CH_3)_3$).

1-(N-Boc)-4(R)-O-mesyl-pyrrolidine-2(S)-methyl alcohol ***5***

From the corresponding methyl ester:

To a stirred mixture of dry ethanol (8ml) and dry THF (6ml), was added lithium chloride (0.05g, 11.6mmol) and sodium borohydride (0.43g, 11.6mmol) and stirring continued for ~ 10 min under Argon atmosphere. A solution of 1-(*N*-Boc)-4(*R*)-*O*-mesyl-2(*S*)-proline methyl ester **4** (1.50g, 4.6mmol) in THF (1ml) was added dropwise and stirring was continued at room temperature under argon atmosphere for 6h, when TLC indicated the disappearance of the starting material. The reaction was then cooled on ice and acidified till pH 4.0 with a 10% aqueous solution of citric acid. The reaction mixture was stirred in an ice-bath for ~10min, after which, the solvents were removed

under vacuum. The residue was taken up in water and the product extracted in ethyl acetate. The organic layer was then dried over sodium sulphate and evaporated to get the desired product **5** as a white solid in quantitative yield (1.34g, 98%).

^{1}H NMR ($CDCl_3$) δ: 5.23 (br s, 1H, *H*4), 4.10 (m, 1H, *H*2), 3.30 (br m, 4H, *H*5, CH_2OH), 3.05 (s, 3H, SO_2CH_3), 2.60 (m, 1H, *H*3), 2.25 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

*1-(N-Boc)-4(S)-azido-pyrrolidine-2(S)-methyl alcohol **6***

A mixture of 1-(*N*-Boc)-4(*R*)-*O*-mesyl-pyrrolidine-2(*S*)-methyl alcohol **5** (1.40g, 4.7mmol) and sodium azide (1.1g, 16.6mmol) in dry DMF (15ml) was stirred at 60°C overnight, when the reaction was found to be complete by TLC inspection. The solvent was completely removed *in vacuo* and the residue, taken up in water. The product **6** was extracted in ethyl acetate and was obtained after the organic layer was dried over sodium sulfate and evaporated to dryness under vacuum (1.12g, 97.5%).

^{1}H NMR ($CDCl_3$) δ: 4.05 (m, 2H, *H*4, CH_2OH), 3.70 (m, 4H, *H*5, *H*5', CH_2OH), 3.30 (dd, 1H, *H*2), 2.30 (m, 1H, *H*3), 1.70 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 156.0 ($COOC(CH_3)_3$), 80.7 ($C(CH_3)_3$), 66.4 (CH_2OH), 58.9 (*C*4), 58.0 (*C*2), 52.1 (*C*5), 33.8 (*C*3), 28.4 (($CH_3)_3$).

*1-(N-Boc)-4(S)-azido-2(S)-(methyl-O-mesyl)-pyrrolidine **7***

To a stirred cooled mixture of 1-(*N*-Boc)-4(*S*)-azido-pyrrolidine-2(*S*)-methyl alcohol **6** (1.0g, 4.1mmol) and triethylamine (1.44ml, 10.3mmol) in dry dichloromethane (10ml), was added dropwise, methane sulfonyl chloride (0.8ml, 10.3mmol) dropwise. Stirring was continued at room temperature overnight. The solvent was completely removed under vacuum. The residue was taken up in water and the crude product obtained by extraction in dichloromethane, followed by drying of the organic layer sodium sulfate

and evaporation under vacuum. Silica gel column chromatography yielded the pure product **7** (0.6g, 45%).

^{1}H NMR ($CDCl_3$) δ: 4.45 (dd, 1H, *H*4), 4.15 (br m, 3H, *H*2, $CH_2OSO_2CH_3$), 3.65 (m, 1H, *H*5), 3.35 (m, 1H, *H*5'), 3.03 (s, 3H, SO_2CH_3), 2.30 (m, 1H, *H*3), 2.13 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 153.7 ($COOC(CH_3)_3$), 80.3 ($C(CH_3)_3$), 68.9 ($CH_2OSO_2CH_3$), 59.0 (*C*2), 54.8 (*C*4), 51.7 (*C*5), 36.8 (SO_2CH_3), 32.6 (*C*3), 28.0 ($(CH_3)_3$).

1-(N-Boc)-4(S)-azido-2(S)-(thymin-1-ylmethyl)-pyrrolidine ***8***

A mixture of 1-(*N*-Boc)-4(*S*)-azido-2(*S*)-(methyl-*O*-mesyl)-pyrrolidine **7** (0.25g, 0.8mmol), anhydrous potassium carbonate (0.11g, 0.8mmol), thymine (0.10g, 0.9mmol) and 18-crown-6 (0.08g, 0.2mmol) in dry DMF (3ml) was stirred at 75°C under argon atmosphere overnight. The solvent was completely removed under vacuum and the crude residue, purified by silica gel column chromatography to obtain the pure product **8** (0.32g, 46%).

^{1}H NMR ($CDCl_3$) δ: 9.50 (min) & 9.25 (maj) (s, 1H, *T*-N$_3$*H*), 7.00 (s, 1H, *T*-*H*6), 4.25 (m, 3H, CH_2-*T*, *H*4), 3.90- 3.50 (br m, 2H, *H*5, *H*5'), 3.40 (m, 1H, *H*2), 2.25 (m, 1H, *H*3), 2.05 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 164.5 (*T*-*C*2), 154.3 ($COOC(CH_3)_3$), 151.4 (*T*-*C*4), 140.9 (*T*-*C*6), 110.1 (*T*-*C*5), 80.3 ($C(CH_3)_3$), 59.9 (*C*2), 55.4 (*C*4), 52.1 (*C*5), 50.5 (CH_2-*T*), 33.8 (*C*3), 28.0 ($(CH_3)_3$), 12.0 (*T*-CH_3).

4(S)-azido-2(S)-(thymin-1-ylmethyl)-pyrrolidine-N1-ethyl acetate ***9***

1-(*N*-Boc)-4(*S*)-azido-2(*S*)-(thymin-1-ylmethyl)-pyrrolidine **8** (0.2g, 0.6mmol) was stirred in a solution of TFA:dichloromethane, 1:1 (1ml) for 30 min, when TLC indicated total deprotection of the ring nitrogen. The solvent was completely removed in vacuo and the residue, co-evaporated with dichloromethane (2 x 1ml). The 4(*S*)-azido-2(*S*)-

(thymin-1-ylmethyl)-pyrrolidine was immediately used in the subsequent reaction. It was taken up in dry acetonitrile (2ml), to which, anhydrous potassium carbonate (0.16g, 1.2mmol) and ethylbromoacetate (0.08ml, 0.7mmol) were added. The reaction was complete after 2h, as evident from TLC examination. The solvent was completely evaporated under vacuum and the crude product purified by silica gel column chromatography to get the pure product **9** (0.13g, 70%).

^{1}H NMR ($CDCl_3$) δ: 9.45 (s, 1H, *T*-N3*H*), 7.24 (s, 1H, *T*-*H*6), 4.15 (m, 3H, OC*H*$_2$CH$_3$, *H*4), 3.75 (m, 2H, C*H*$_2$COOEt), 3.55-3.10 (m, 4H, *H*5, *H*5', C*H*$_2$-*T*), 3.00 (dd, 1H, *H*2), 2.35 (m, 1H, *H*3), 1.90 (s, 3H, *T*-C*H*$_3$), 1.75 (m, 1H, *H*3'), 1.25 (t, 3H, OCH$_2$C*H*$_3$).

^{13}C NMR ($CDCl_3$) δ: 170.2 (*C*OOEt), 164.6 (*T*-*C*2), 151.3 (*T*-*C*4), 142.1 (*T*-*C*6), 109.1 (*T*-*C*5), 60.2 (OC*H*$_2$CH$_3$), 59.8 (*C*4), 57.6 (N-C*H*$_2$), 56.6 (*C*2), 53.6 (*C*5), 50.6 (*C*H$_2$-*T*), 34.2 (*C*3), 13.7 (OCH$_2$C*H*$_3$), 11.7 (*T*-C*H*$_3$).

4(S)-(N-Boc-amino)-2(S)-(thymin-1-ylmethyl)-pyrrolidine-N1-ethyl acetate ***10***

The azide function in 4(*S*)-azido-2(*S*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-ethyl acetate **9** (0.13g, 0.4mmol) was reduced by hydrogenation using 10% Pd-C (0.03g) as a catalyst and an H_2 gas pressure of 50 psi in a Parr hydrogenation apparatus. The reduction was complete within 4h, as monitored by TLC. The catalyst was filtered off and the filtrate evaporated to dryness and immediately set for the next reaction. The residue was taken up in dioxane:water, 1:1 (1ml). To this, was added triethylamine (0.14ml, 1.0mmol) and Boc-azide (0.07ml, 0.5mmol). The reaction was stirred at room temperature overnight. The solvents were removed under vacuum and the residue, taken up in water. The crude product was extracted in chloroform. The organic layer was dried over sodium sulfate and evaporated to get the crude product, which was purified by silica gel column chromatography to obtain the pure product **10** (0.10g, 62%).

^{1}H NMR ($CDCl_3$) δ: 9.30 (s, 1H, *T*-N3*H*), 7.32 (s, 1H, *T*-*H*6), 5.00 (d, 1H, Boc-N*H*), 4.20 (q, 2H, OCH_2CH$_3$), 4.10 (m, overlapping signal, 1H, *H*4 & 1H, N-CH_2), 3.55- 2.85 (m, 4H, CH_2-*T*, *H*5, *H*5'), 2.75 (m, 1H, *H*2), 2.30 (m, 1H, *H*3), 1.95 (s, 3H, *T*-CH_3), 1.60 (m, 1H, *H*3'), 1.45 (s, 9H, (CH_3)$_3$), 1.30 (t, 3H, OCH$_2$CH_3).

^{13}C NMR ($CDCl_3$) δ: 170.8 (*C*OOEt), 164.7 (*T*-*C*2), 155.5 (*C*OOC(CH$_3$)$_3$), 151.8 (*T*-*C*4), 141.9 (*T*-*C*6), 110.2 (*T*-*C*5), 79.5 (*C*(CH$_3$)$_3$), 61.2 (*C*4), 60.9 (OCH$_2$CH$_3$), 60.0 (N-*C*H$_2$), 54.8 (*C*5), 49.9 (*C*H$_2$-*T*), 49.6 (*C*2), 35.9 (*C*3), 28.4 ((*C*H$_3$)$_3$), 14.3 (OCH$_2$*C*H$_3$), 12.2 (*T*-*C*H$_3$).

1-(N-Boc)-4(R)-O-mesyl-2(R)-proline methyl ester ***14***

To an ice-cooled solution of 1-(*N*-Boc)-4(*R*)-hydroxy-2(*R*)-proline methyl ester **13** (2.0g, 8.2mmol) in dry pyridine (15ml) was added dropwise, methane sulfonyl chloride (1.3ml, 16.3mmol). Stirring was continued at room temperature. The reaction was complete in 4h. The solvent was completely removed *in vacuo* and the residue, taken up in water. The product was extracted in chloroform. The chloroform extracts were dried over sodium sulfate and evaporated to get the product **14**, which was obtained in the pure form after silica gel column chromatography (2.5g, 95%).

^{1}H NMR ($CDCl_3$) δ: 5.32 - 5.22 (m, 1H, *H*4), 4.52 - 4.35 (m, 1H, *H*2), 3.85 - 3.73 (m, 2H, *H*5, *H*5'), 3.75 (s, 3H, COOCH_3), 3.05 (s, 3H, SO_2CH_3), 2.72 - 2.52 (m, 1H, *H*3), 2.35 - 2.17 (m, 1H, *H*3'), 1.48 & 1.43 (s, 9H, (CH_3)$_3$).

1-(N-Boc)-4(R)-O-mesyl-pyrrolidine-2(R)-methyl alcohol ***15***

Sodium borohydride (0.76g, 20.4mmol) and lithium chloride (0.87g, 20.4mmol) were added to a cooled stirred mixture of dry THF (9ml) and dry ethanol (12ml). To this was added dropwise a solution of 1-(*N*-Boc)-4(*R*)-*O*-mesyl-2(*R*)-proline methyl ester **14** (2.64g, 8.2mmol) in THF:ethanol, 3:4 (2ml) and stirring was continued at room temperature overnight under argon atmosphere. The reaction was cooled on ice and

acidified till pH 4.0 with a 10% aqueous solution of citric acid. After stirring for 10min, the solvents were removed under vacuum and the residue was taken up in water. The product was extracted in ethyl acetate and the organic layer evaporated to dryness after drying over sodium sulfate, to get the product **15** (1.9g, 80%).

^{1}H NMR ($CDCl_3$) δ: 5.18 (m, 1H, *H*4), 4.10 (m, 1H, *H*2), 3.90 - 3.55 (m, 4H, *H*5, *H*5', CH_2OH), 3.25 (br s, CH_2OH), 3.05 (s, 3H, SO_2CH_3), 2.40 (m, 1H, *H*3), 2.07 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

*1-(N-Boc)-4(S)-azido-pyrrolidine-2(R)-methyl alcohol **16***

A mixture of 1-(*N*-Boc)-4(*R*)-*O*-mesyl-pyrrolidine-2(*R*)-methyl alcohol **15** (1.18g, 4mmol) and sodium azide (0.91g, 13.9mmol) in dry DMF (10ml) was stirred at 60°C overnight, when the reaction was found to be complete upon TLC inspection. The solvent was completely removed *in vacuo*, the residue was taken up in water, and the product, extracted in chloroform. The organic solvents were dried over sodium sulphate and evaporated to get the crude product, which was purified by silica gel column chromatography to get pure 1-(*N*-Boc)-4(*S*)-azido-pyrrolidine-2(*R*)-methyl alcohol **16** (0.94g, 90%).

^{1}H NMR ($CDCl_3$) δ: 4.60 (br s, 1H, *OH*), 4.05 (m, 2H, CH_2OH), 3.70 (m, 1H, *H*4), 3.50 (m, 3H, *H*5, *H*5', *H*2), 2.10 (m, 1H, *H*3), 1.85 (m, 1H, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 154.8 ($COOC(CH_3)_3$), 79.5 ($C(CH_3)_3$), 63.8 (CH_2OH), 58.4 (*C*4), 57.4 (*C*2), 51.3 (*C*5), 33.4 (*C*3), 27.6 (($CH_3)_3$).

*1-(N-Boc)-4(S)-azido-2(R)-(methyl-O-mesyl)-pyrrolidine **17***

To an ice-cooled solution of 1-(*N*-Boc)-4(*S*)-azido-pyrrolidine-2(*R*)-methyl alcohol **16** (0.8g, 3.3mmol) in dry pyridine (10ml) was added drop-wise methane sulfonyl chloride (0,77ml, 9.9mmol). Stirring was continued at room temperature for 2h when the absence of starting material on TLC indicated the completion of the reaction. The

solvent was completely removed under vacuum and the residue, taken up in water. The crude product was extracted in dichloromethane, the organic layer dried over sodium sulfate, evaporated to dryness and purified by silica gel column chromatography to get the pure product **17** (0.9g, 86%).

^{1}H NMR ($CDCl_3$) δ: 4.60 (maj) & 4.45 (min) (m, 1H, *H*4), 4.20 (m, 3H, *H*2, $CH_2OSO_2CH_3$), 3.65 & 3.45 (m, 2H, *H*5, *H*5'), 3.00 (s, 3H, SO_2CH_3), 2.40 (min) & 2.20 (maj) (m, 2H, *H*3, *H*3'), 1.45 (s, 9H, $(CH_3)_3$).

*1-(N-Boc)-4(S)-azido-2(R)-(thymin-1-ylmethyl)-pyrrolidine **18***

A mixture of 1-(*N*-Boc)-4(*S*)-azido-2(*R*)-(methyl-*O*-mesyl)-pyrrolidine **17** (0.90g, 2.8mmol), thymine (0.35g, 2.8mmol), anhydrous potassium carbonate (0.39g, 2.8mmol) and 18-crown-6 (0.29g, 0.8mmol) in dry DMF (10ml) was stirred at 70°C, 7h. The solvent was then evaporated under vacuum. The crude product was purified by silica gel column chromatography to get the pure product **18** (0.64g, 65%).

^{1}H NMR ($CDCl_3$) δ: 8.90 (br s, 1H, *T*-N3*H*), 6.85 (s, 1H, *T*-*H*6), 5.15 (s, 1H, *H*4), 3.90 - 3.10 (m, 5H, *H*5, *H*5, *H*2, CH_2-*T*), 2.15 (m, 2H, *H*3, *H*3'), 1.80 (s, 3H, *T*-CH_3), 1.45 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 163.9 (*T*-*C*2), 153.5 ($COOC(CH_3)_3$), 151.1 (*T*-*C*4), 139.5 (*T*-*C*6), 109.3 (*T*-*C*5), 79.3 ($C(CH_3)_3$), 58.0 (*C*2), 54.6 (*C*4), 51.0 (*C*5), 48.0 (CH_2-*T*), 34.0 (*C*3), 27.0 (($CH_3)_3$), 11.1 (*T*-CH_3).

*4-(S)-azido-2-(R)-(thymin-1-ylmethyl)-pyrrolidine-N1-ethylacetate **19***

1-(*N*-Boc)-4(*S*)-azido-2(*R*)-(thymin-1-ylmethyl)-pyrrolidine **18** (0.75g, 2.1mmol) was stirred in trifluoroacetic acid:dichloromethane, 1:1 (7ml) at room temperature for 45min to afford the pyrrolidine N-deprotected TFA salt. The solvents were completely removed, the residue was co-evaporated with dichloromethane and immediately set for N-alkylation. To an ice-cooled solution of the above TFA salt and anhydrous K_2CO_3

(0.60g, 4.3mmol) in dry acetonitrile (10ml) was added dropwise, ethylbromoacetate (0.3ml, 2.5mmol), under argon atmosphere. Stirring was continued at room temperature for 2h, when no starting material was observed by TLC. The reaction mixture was evaporated to dryness under vacuum and purified by silica gel column chromatography to get the pure product **19** (0.50g, 69%).

^{1}H NMR ($CDCl_3$) δ: 9.85 (s, 1H, *T*-N3*H*), 7.35 (s, 1H, *T*-*H*6), 4.15 (q, 2H, OC*H*$_2$CH$_3$), 4.00 (m, 1H, *H*4), 3.85 (dd, 1H, *H*5), 3.70 (m, 1H, *H*5'), 3.60- 3.20 (m, 5H, *H*2, C*H*$_2$-*T*, N-C*H*$_2$), 2.60 (dd, 1H, *H*3), 2.05 (m, 1H, *H*3'), 1.90 (s, 3H, *T*-C*H*$_3$), 1.25 (t, 3H, OCH$_2$*CH*$_3$).

4-(S)-(N-Boc-amino)-2-(R)-(thymin-1-ylmethyl)-pyrrolidine-N1-ethyl acetate ***20***

4-(*S*)-azido-2-(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-ethylacetate **19** (0.11g, 0.33mmol) was hydrogenated (50psi H_2) in methanol (5ml) in the presence of 10% Pd-C (20mg) as a catalyst. The azide was completely reduced within 4h, as evident upon TLC examination. The catalyst was filtered off and the filtrate evaporated to dryness under vacuum to afford 4-(*S*)-amino-2-(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-ethyl acetate (0.1g, quantitative yield), which was immediately used in the further reaction. The free amine (0.10gm, 0.33mmol) was stirred with triethylamine (0.1ml, 0.7mmol) in 1,4-dioxane:water (1:1, 2ml). Boc-azide (0.1ml, 0.7mmol) was added dropwise and the reaction was stirred at room temperature, 8h, when TLC indicated the completion of the reaction. The solvents were completely removed under reduced pressure and the residue taken up in water. The product was extracted in dichloromethane. The organic extracts were dried over sodium sulphate and evaporated to give the crude product **20**, which was purified by silica gel column chromatography (0.12g, 89%).

^{1}H NMR ($CDCl_3$) δ: 9.00 (s, 1H, *T*-N3*H*), 7.15 (s, 1H, *T*-*H*6), 5.18 (d, 1H, Boc-N*H*), 4.17 (q, 2H, OC*H*$_2$CH$_3$), 3.78 (m, 1H, *H*4), 3.65- 3.25 (m, 5H, *H*2, *H*5, *H*5', N-C*H*$_2$), 2.55

(m, 1H, *H*3), 1.90 (m, 1H, *H*3'), 1.85 (s, 3H, *T*-CH_3), 1.45 (s, 9H, (CH_3)$_3$), 1.28 (t, , 3H, OCH$_2$CH_3).

4-(S)-(N-Boc-amino)-2-(R)-(thymin-1-ylmethyl)-pyrrolidine-N1-acetic acid ***21***

To a solution of 4-(*S*)-(*N*-Boc-amino)-2-(*R*)-(thymin-1-ylmethyl)-pyrrolidine-*N1*-ethyl acetate **20** (0.11g, 0.27mmol) in methanol (1ml), was added 2N aqueous NaOH (1ml). The reaction was stirred at room temperature for 10min, when TLC analysis indicated the complete absence of the starting ester. The excess NaOH was neutralized by Dowex H^+ resin, which was then filtered off. The filtrate was evaporated to get the product as a white foam (0.10g, 98%).

^{1}H NMR ($CDCl_3$) δ: 7.50 (s, 1H, *T*-*H*6), 4.30 (m, 2H, N-CH_2COOH), 3.95 (m, 4H, CH_2-*T*, *H*5, *H*5'), 3.65 (m, 1H, *H*4), 3.40 (m, 1H, *H*2), 2.75 (m, 1H, *H*3), 2.00 (m, 1H, *H*3'), 1.85 (s, 3H, *T*-CH_3), 1.35 (s, 9H, (CH_3)$_3$).

^{13}C NMR ($CDCl_3$) δ: 171.1 (*C*OOH), 167.3 (*T*-*C*2), 158.3 (*C*OO(CH$_3$)$_3$), 154.5 (*T*-*C*4), 144.0 (*T*-*C*6), 112.0 (*T*-*C*5), 82.4 (*C*(CH$_3$)$_3$), 66.5 (*C*4), 61.3 (N-*C*H$_2$COOH), 57.0 (*C*5), 48.1 (*C*2), 47.4 (*C*H$_2$-*T*), 32.7 (*C*3), 28.3 ((*C*H$_3$)$_3$), 12.1 (*T*-*C*H$_3$).

1-(N-Boc)-4-(S)-azido-2-(S)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine ***22***

A mixture of 1-(*N*-Boc)-4(*S*)-azido-2(*S*)-(methyl-*O*-mesyl)-pyrrolidine **7** (0.10g, 0.31mmol), anhydrous potassium carbonate (0.22g, 1.56mmol), 2-amino-6-chloropurine (0.13g, 0.78mmol) and 18-crown-6 (0.03g, 0.09mmol) in dry DMF was stirred at 65 ^{o}C under argon atmosphere overnight. The solvent was completely removed under vacuum and the crude residue, purified by column chromatography subsequent to filtering and evaporation of the filtrate. The pure product **22** (0.06g) was obtained in 50% yield.

^{1}H NMR ($CDCl_3$) δ: 7.75 (s, 1H, *2-amino-6-chloropurine*, *H*8), 5.40 (br s, 2H, N^2H_2), 4.45 (m, 2H, *H*5, *H*5'), 4.20 (m, 2H, CH_2-*2-amino-6-chloropurine*), 3.60 (m, 1H, *H*4), 3.35 (m, 1H, *H*2), 2.10 (m, 2H, *H*3, *H*3'), 1.40 (s, 9H, $(CH_3)_3$).

^{13}C NMR ($CDCl_3$) δ: 159.2 (*2-amino-6-chloropurine*, *C*2), 154.1 (*2-amino-6-chloropurine*, *C*4 & $COOC(CH_3)_3$), 151.0 (2-*amino-6-chloropurine*, *C*6), 142.6 (*2-amino-6-chloropurine*, *C*8), 124.8 (*2-amino-6-chloropurine*, *C*5), 80.7 ($C(CH_3)_3$), 59.6 (*C*2), 56.5 (*C*4), 52.0 (CH_2-*2-amino-6-chloropurine*), 45.2 (*C*5), 33.7 (*C*3), 28.1 (($CH_3)_3$).

4(S)-azido-2(S)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine-N1-ethylacetate

23

To a solution of 1-(*N*-Boc)-4(*S*)-azido-2(*S*)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine **22** (0.08g, 0.20 mmol) in dry CH_2Cl_2 (1ml) was added TFA (1ml) and the reaction was stirred at room temperature for 10min, when TLC indicated a complete absence of starting material **22**. The TFA salt was neutralized with DIPEA and the resulting free amine was immediately subjected to alkylation using ethylbromoacetate (0.025 ml, 0.20 mmol) in dry THF (1 ml) in the presence of DIPEA (0.053 ml, 0.31 mmol). The amine was completely consumed within 2h, upon which, the solvents were removed under vacuum and the crude product purified by silica gel column chromatography to get the pure product **23** in excellent yield (91%).

^{1}H NMR ($CDCl_3$) δ: 7.90 (s, 1H, *2-amino-6-chloropurine*, *H*8), 5.20 (s, 2H, N^2H_2), 4.10 (m, 5H, *H*5, *H*5', *H*4, CH_2CH_3), 3.45 (m, 2H, C2-CH_2), 3.20 (m, 2H, CH_2COOEt), 3.05 (m, 1H, *H*2), 2.30 (m, 1H, *H*3), 1.80 (m, 1H, *H*3'), 1.25 (t, 3H, CH_3)

4(S)-(N-Boc-amino)-2(S)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine-N1-ethylacetate ***24***

4(*S*)-azido-2(*S*)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine-*N1*-ethylacetate **23** (0.07g, 0.18mmol) was reduced to the 4(S)-amine by hydrogenation at room

temperature for 2h. The resulting amine was treated with Boc-azide (0.03ml, 0.22mmol) in dioxane:water, 1:1 (1ml) in the presence of triethylamine (0.03ml, 0.22mmol) to get the crude product that was purified by silica gel column chromatography to get pure **24** (0.02g, 25% yield).

^{1}H NMR ($CDCl_3$) δ: 7.95 (s, 1H, 2-*amino-6-chloropurine*, *H*8), 5.28 (s, 2H, N^2H_2), 5.13 (m, 1H, Boc-N*H*), 4.14 (m, 5H, *H*5, *H*5', *H*4, CH_2CH_3), 3.48 (m, 1H, *H*2), 3.28 (m, 2H, C2-CH_2), 2.93 (m, 2H, CH_2-COOEt), 2.51 (m, 1H, *H*3), 2.28 (m, 1H, *H*3'), 1.37 (s, 9H, $(CH_3)_3$), 1.26 (t, 3H, CH_2CH_3).

4(S)-(N-Boc-amino)-2(S)-(guanin-9-ylmethyl)-pyrrolidine-N1-acetic acid ***25***

4(*S*)-(*N*-Boc-amino)-2(*S*)-(2-amino-6-chloropurin-9-ylmethyl)-pyrrolidine-*N1*-ethylacetate **24** (0.02g, 0.04mmol) was suspended in methanol:water, 1:1 (1ml) and stirred overnight at room temperature. The excess alkali was neutralized using Dowex H^+ resin and the product was obtained after filtration and concentration of the filtrate.

pH titration

5mg of the (2*S*,4*S*) pyrrolidyl PNA-T monomer **11** was treated with TFA in CH_2Cl_2 to get the deprotected amine. Co-evaporation was carried out with CH_2Cl_2 to eliminate the excess TFA. It was dissolved in water (2ml) and 0.01N NaOH was added in small aliquots and the pH was recorded after each addition. The pH was plotted as a function of the NaOH added and the pKas were determined from the peaks in the first derivative plot.

Solid Phase Peptide Synthesis

The PNA oligomers were synthesized by solid phase peptide synthesis by repetitive cycles each comprising (I) Boc-deprotection with 50% TFA:CH_2Cl_2 (ii) neutralization of the TFA salt formed using a 5%solution of DIPEA in CH_2Cl_2 to yield the free amine (iii)

coupling of the resin-bound amine with the incoming amino acid possessing a free carboxylic acid group. The coupling agents used were HOBt and DIPCDI and were added in 3-4 times excess over the active amine groups of the support. The resin was thoroughly washed in between each step using a mixture of solvents, *viz*. CH_2Cl_2 and DMF. The deprotection and coupling steps were monitored by Kaiser's test. This test was positive and resulted in a dark blue colour on the resin beads after Boc-deprotection. Conversely, upon completion of the coupling reaction, the test was negative, the resin beads remaining colourless, indicative of a very high coupling efficiency (>90%).

Cleavage by Acid[15b]

A typical cleavage reaction was carried out on 10mg resin-linked PNA. To the resin-linked PNA, were added thioanisole (20µl) and 1,2-ethane-dithiol (8µl). The reaction was stirred for 10min after which, TFA was added (120µl). After further stirring for 10min, TFMSA (16µl) was added and the reaction stirring was continued for 1.5-2h. The resin was filtered off using a funnel fitted with a sintered glass frit. The resin was washed 2-3 times with TFA and the combined washings were evaporated under vacuum. These were co-evaporated 2-3 times with dry ether to eliminate the excess TFA. The residue was taken up in a minimum amount of methanol and excess ether was added to precipitate the PNA. Centrifugation caused the separation of the PNA in the form of a white precipitate. The dissolution in methanol and precipitation by ether was repeated thrice to obtain a clean white precipitate of the crude PNA oligomer.

Cleavage by Aminolysis[20]

To the resin-linked PNA (10mg), was added an excess of spermine (60mg) and the mixture was heated at 55-60 ^{o}C for ~48h, when the PNA oligomer was obtained as its 'C' terminal sperminamide.

Purification of the PNA Oligomers

Preliminary purification of the PNA oligomers was carried out by gel filtration through G25 Sephadex. The crude PNA oligomers obtained by acid cleavage after ether precipitation or by aminolysis were dissolved in approximately 0.5ml water and passed through the G25 Sephadex column. Fractions of 1ml volume each were collected by elution with water. The presence of the PNA oligomer was detected by recording the absorbance at 260nm.

The PNA oligomers were further purified by FPLC on a reverse-phase C-8 column if required. This was achieved by using an ascending gradient of acetonitrile in water containing 0.1% TFA. The purity of the oligomers after FPLC was re-checked by HPLC and confirmed by MALDI-TOF mass spectroscopy.

UV- studies

The concentration of the PNA oligomers was calculated based on their absorption at 260nm and using the molar extinction co-efficients as for DNA, *viz*. 8.8 $cm^2/\mu mol$ for T, 15.4 $cm^2/\mu mol$ for A, 7.3 $cm^2/\mu mol$ for C and 11.7 $cm^2/\mu mol$ for G.[22]

PNA complexes with the corresponding complementary nucleic acid oligomers were constituted by mixing together stoichiometric quantities of the involved oligomers **46** & **26-31** with DNA **84** in 0.01M sodium phosphate buffer. The samples were annealed by first heating at 85 ^oC for 2 min followed by slow cooling to room temperature. They were kept at room temperature for at least half an hour and refrigerated overnight prior to running the melting experiments. Heating was carried out at a rate of 0.5 ^oC/min and the absorbance at 260nm was recorded every minute. The Tm of the PNA:DNA complexes was determined from the peaks in first derivative plots of the absorbance or percent hyperchromicity Vs temperature graphs.

Mismatch studies were performed with the complexes of PNAs **26-31** with DNA **86**. The sample preparation and execution of the experiments was the same as for the PNA:DNA complexes involving DNA **84**.

PNA:PNA complexes were constituted by mixing together equimolar quantities of PNAs **26-29** with the complementary PNA homopurine oligomer **63** in 0.01M sodium phosphate buffer, pH 7.4 and the samples were annealed as described above.

CD experiments

CD spectra were recorded on a JASCO J715 spectropolarimeter. Each spectrum was recorded with a scan-speed of 200nm/min from 320nm to 195nm keeping a resolution of 0.1nm, a band-width of 1.0nm, sensitivity of 2mdeg and response 2s as an accumulation of 8 scans.

The samples for CD were prepared in the manner used for UV-Tm studies by heating at 85 ^{o}C for 5min, followed by slow cooling to room temperature and refrigeration overnight prior to recording the CD spectra. The temperature was maintained below the Tm of the complexes as determined by UV measurements, i.e., at 10 ^{o}C.

4.9. REFERENCES

1. (a) Nielsen, P. E.; Egholm, M.; Buchardt, O. *Science* **1991**, *254*, 1497. (b) Ganesh, K. N.; Nielsen, P. E. *Curr. Org. Chem*. **2000**, *4*, 931.
2. Barawkar, D. A.; Bruice, T. C. *J. Am. Chem. Soc*. **1999**, *121*, 10418.
3. Hickman, D. T.; King, P. M.; Cooper, M. A.; Slater, J. M.; Micklefield, J. *Chem. Commun.* **2000**, 2251.
4. Püschl, A.; Tedeschi, T.; Nielsen, P. E. *Org. Lett.* **2000**, *2*, 4161.
5. Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 1083.

6. Vilaivan, T.; Khongdeesameor, C.; Harnyuttanakorn, P.; Westwell, M.; Lowe, G. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2541.
7. (a) Schmid, N.; Behr, P. *Tetrahedron Lett*. **1993**, *36*, 1447. (b) Sund, O.; Puri, N.; Chattopadhaya, J. *Tetrahedron* **1996**, *52*, 12275.
8. (a) Barawkar, D. A.; Rajeev, K. G.; Kumar, V. A.; Ganesh, K. N. *Nucleic Acids Res*. **1996**, *24*, 1229. (b) Prakash, T. P.; Barawkar, D. A.; Kumar, V. A.; Ganesh, K. N. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1733. (c) Barawkar, D. A.; Kumar, V. A.; Ganesh, K. N. *Biochem. Biophys. Res. Commun.* **1994,** *205*, 1665.
9. Gangamani, B. P.; Kumar, V. A.; Ganesh, K. N. *Tetrahedron* **1999**, *55*, 177.
10. Remuzon. P. *Tetrahedron* **1996**, *52*, 13803.
11. Hamada, Y.; Shioiri, T. *Tetrahedron Lett.* **1982**, *23*, 1193.
12. Robinson, D. S.; Greenstein, J. P. *J. Biol. Chem.* **1952**, *195*, 383.
13. 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.
14. 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. Pept. Sci.* **1995**, *3*, 175.
15. (a) Merrifield, R. B.; Stewart, J. M.; Jernberg, N. *Anal. Chem.* **1966**, *38*, 1905. (b) Erikson, B. W.; Merrifield, R. B. in *Solid Phase Synthesis in the Proteins* Vol. II, 3rd Ed., H. Neurath & R. L. Hill, eds., Academic Press, New York **1976**, pp 255.
16. Gisin, B. F. *Helv, Chim. Acta* **1970**, *56*, 1476.
17. (a) Hodges, R. S.; Merrifield, R. B. *Anal. Biochem.* **1975**, *65*, 241. (b) Arad, O.; Houghten, R. A. *Peptide Res.* **1990**, *3*, 42.
18. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34*, 595.
19. Buchardt, O.; Egholm, M.; Nielsen, P. E.; Berg, R. H. *PCT no. WO 92/20702* **1992**.

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

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

22. *Oligonucleotide Synthesis. A Practical Approach* **1984**, Ed. M. J. Gait, IRL Press, Oxford, Washington D. C.

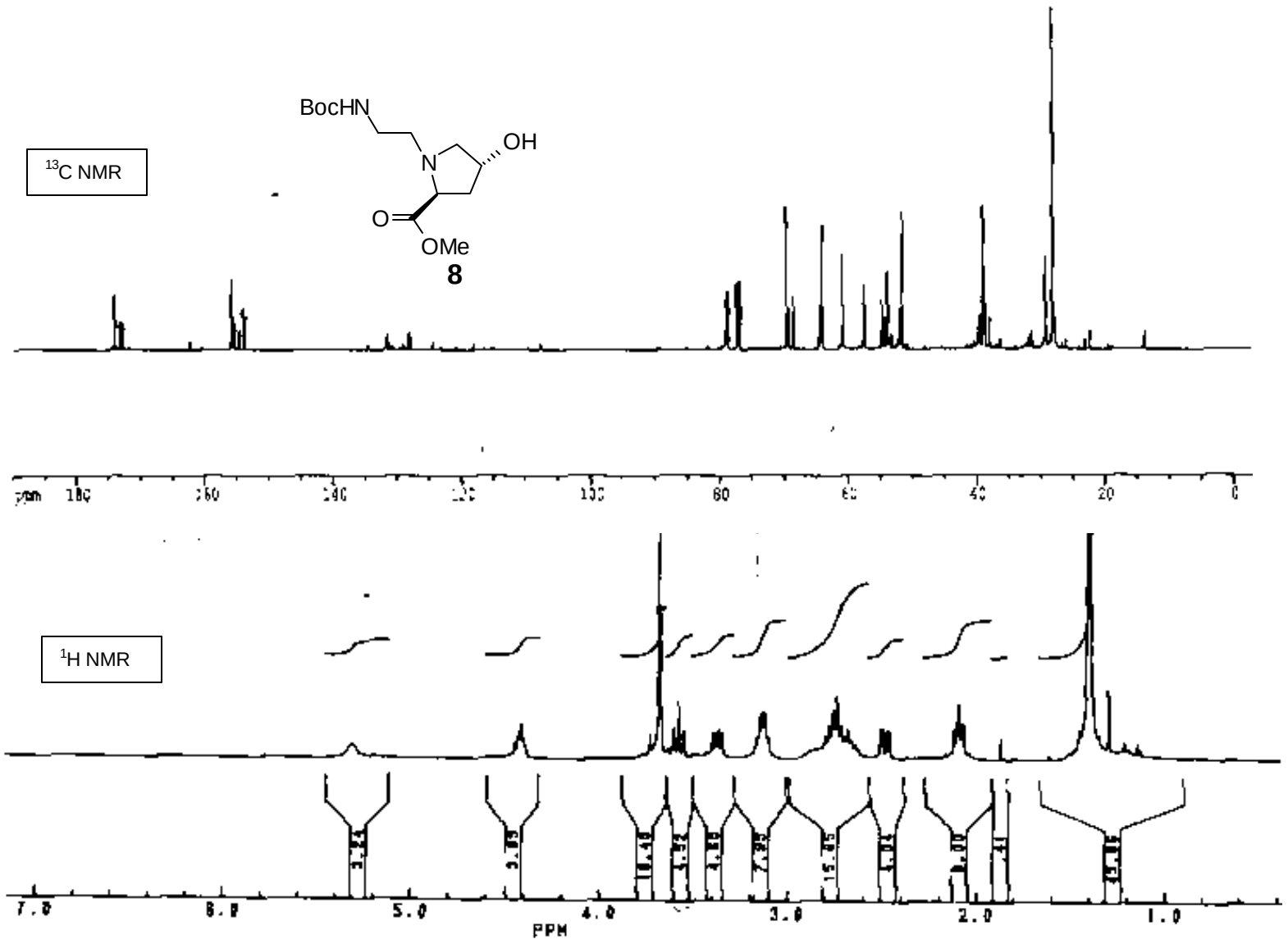
23. Kim, S H.; Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

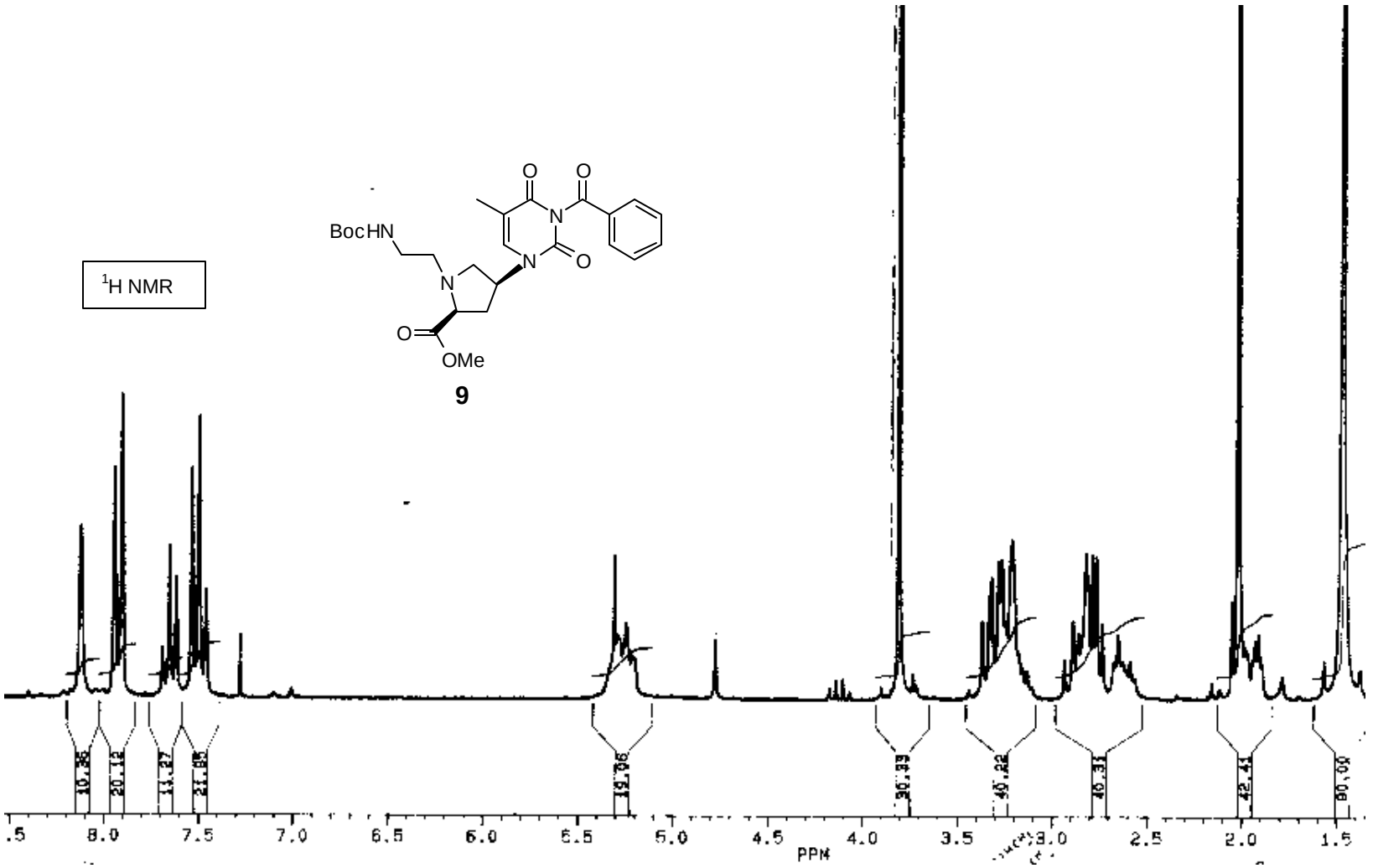
24. Armarego, W. L. F.; Perrin, D. D. in *Purification of Laboratory Chemicals*, Butterworth-Heineman **1996**, The Bath Press, Bath, Great Britain.

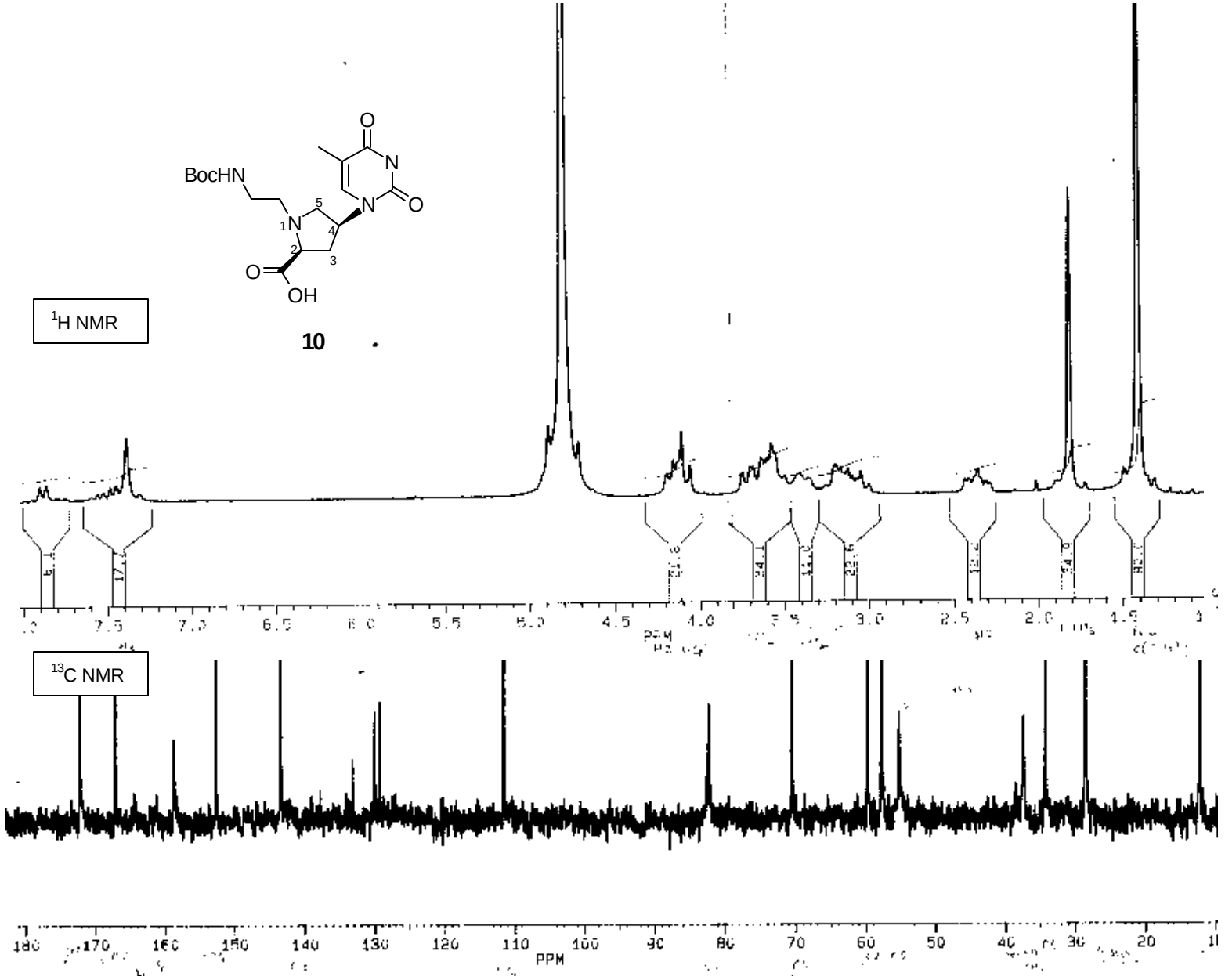
4.10. APPENDIX

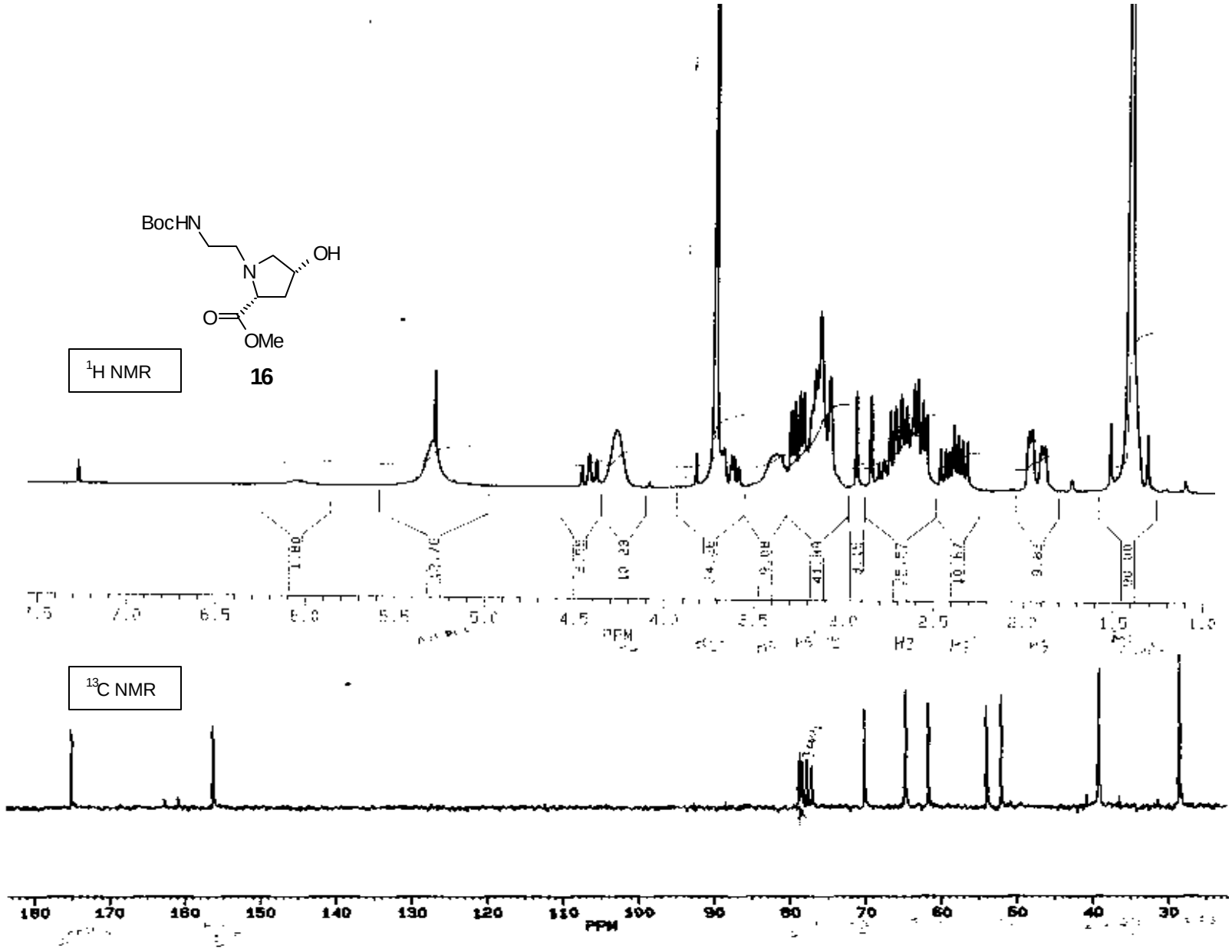
- Compound **5**, ^{1}H NMR
- Compound **6**, ^{1}H NMR
- Compound **7**, ^{1}H NMR
- Compound **8**, ^{1}H and ^{13}C NMR
- Compound **9**, ^{1}H and ^{13}C NMR
- Compound **10**, ^{1}H and ^{13}C NMR
- Compound **11**, ^{1}H and ^{13}C NMR
- Compound **16**, ^{1}H and ^{13}C NMR
- Compound **17**, ^{1}H NMR
- Compound **18**, DEPT and ^{13}C NMR
- Compound **19**, ^{1}H NMR
- Compound **22**, ^{1}H and ^{13}C NMR
- Compound **23**, ^{1}H NMR
- Compound **24**, ^{1}H NMR

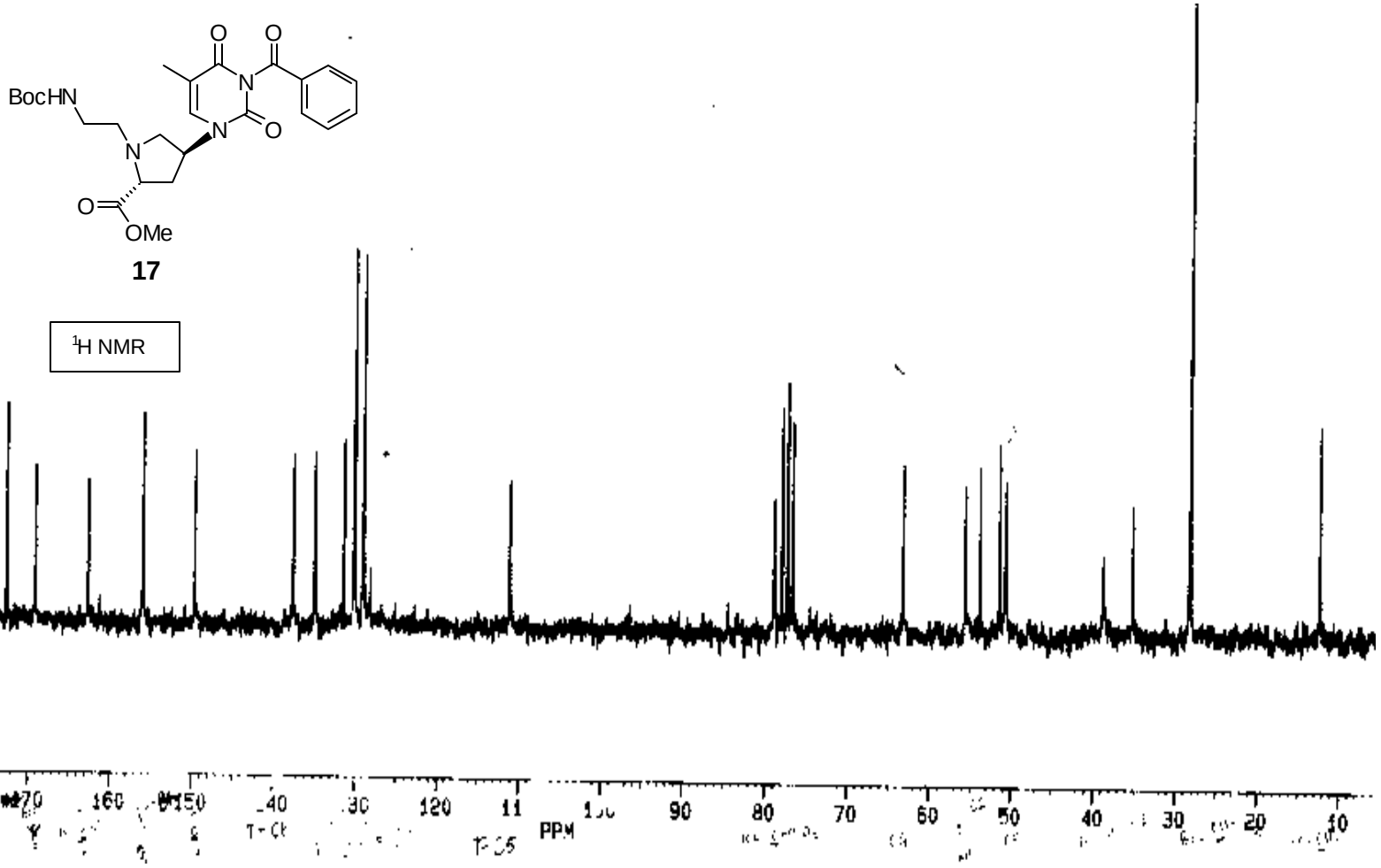
Appendix to Chapter 2

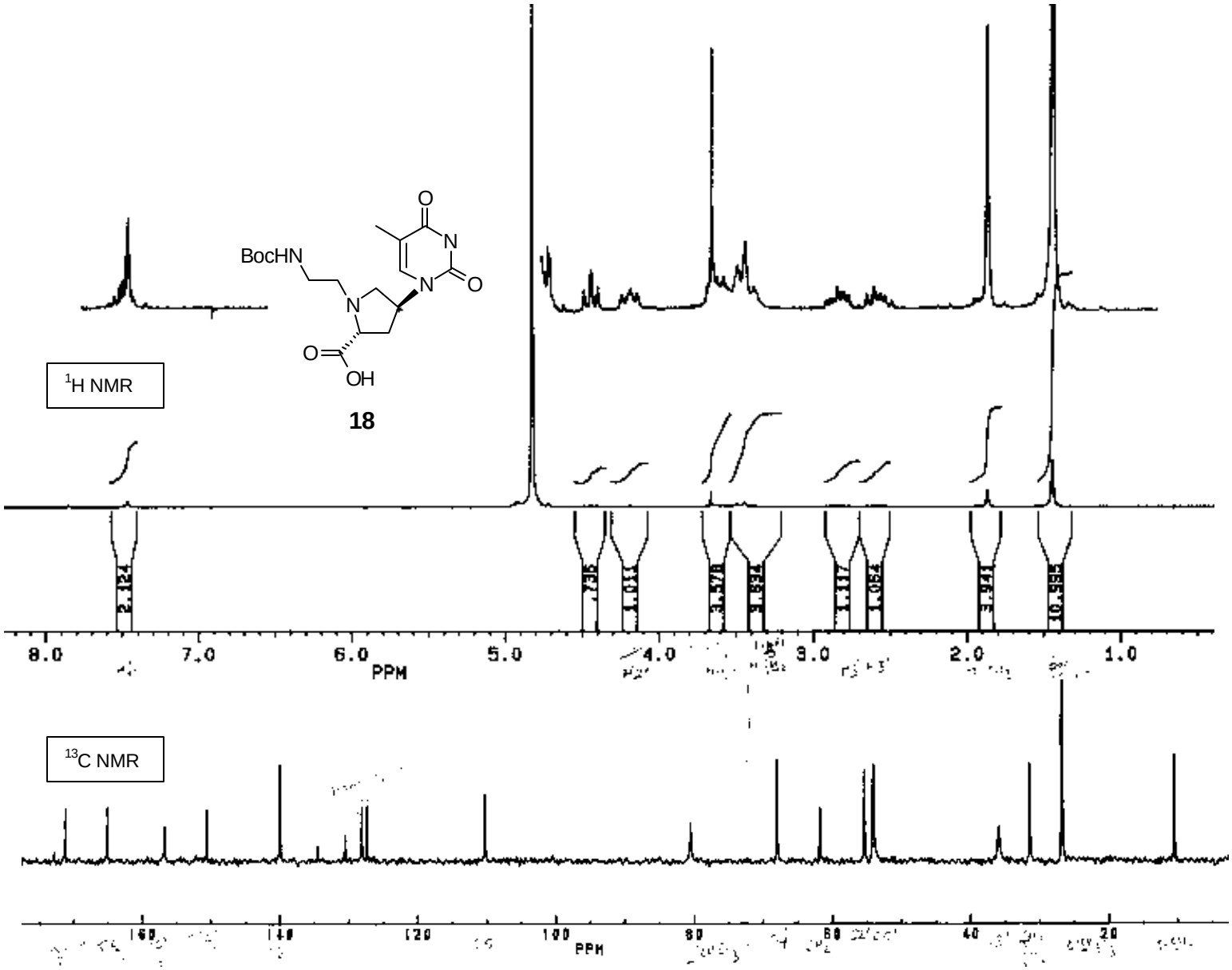

BocHN
N
OH
O
OMe
8
13C NMR
1H NMR
PPM

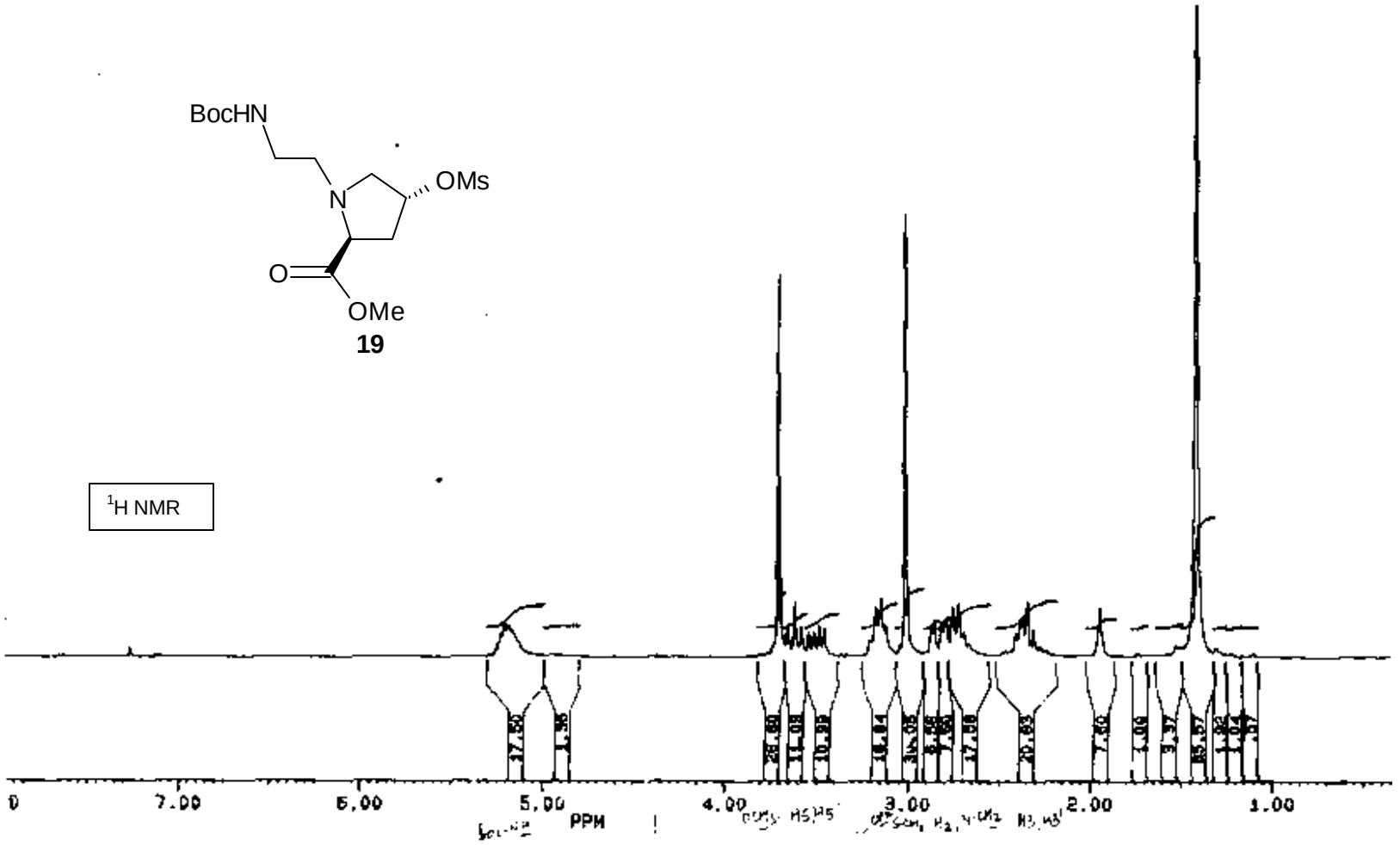

^{1}H NMR
BocHN
O
O
N
N
O
N
O
OMe
9
10.36
20.12
11.27
21.85
19.76
36.33
40.22
40.31
42.41
80.00
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
PPM

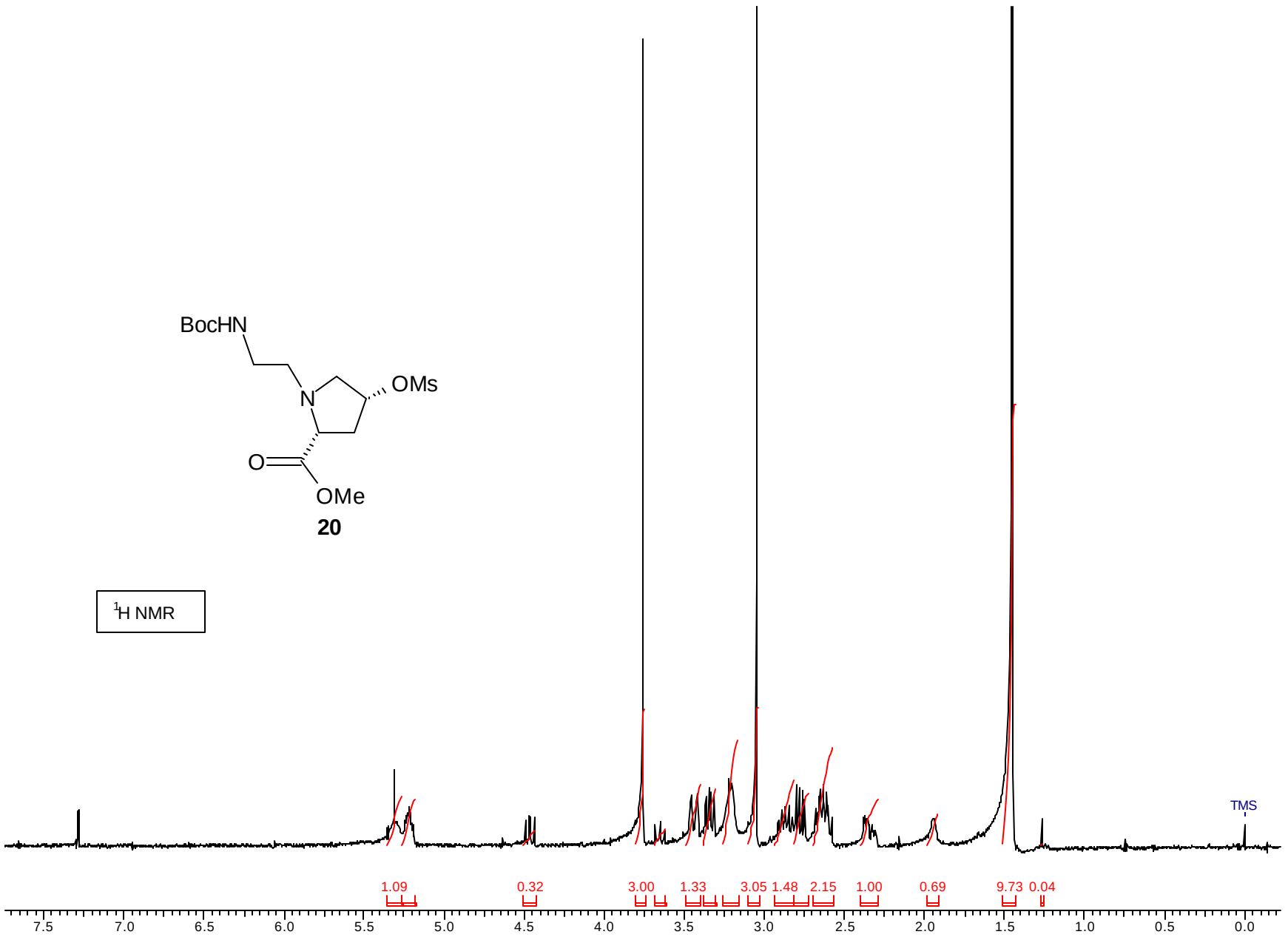

BocHN
O
N
N
O
1N
5
4
2
3
O
OH
10
1H NMR
13C NMR
PPM
PPM

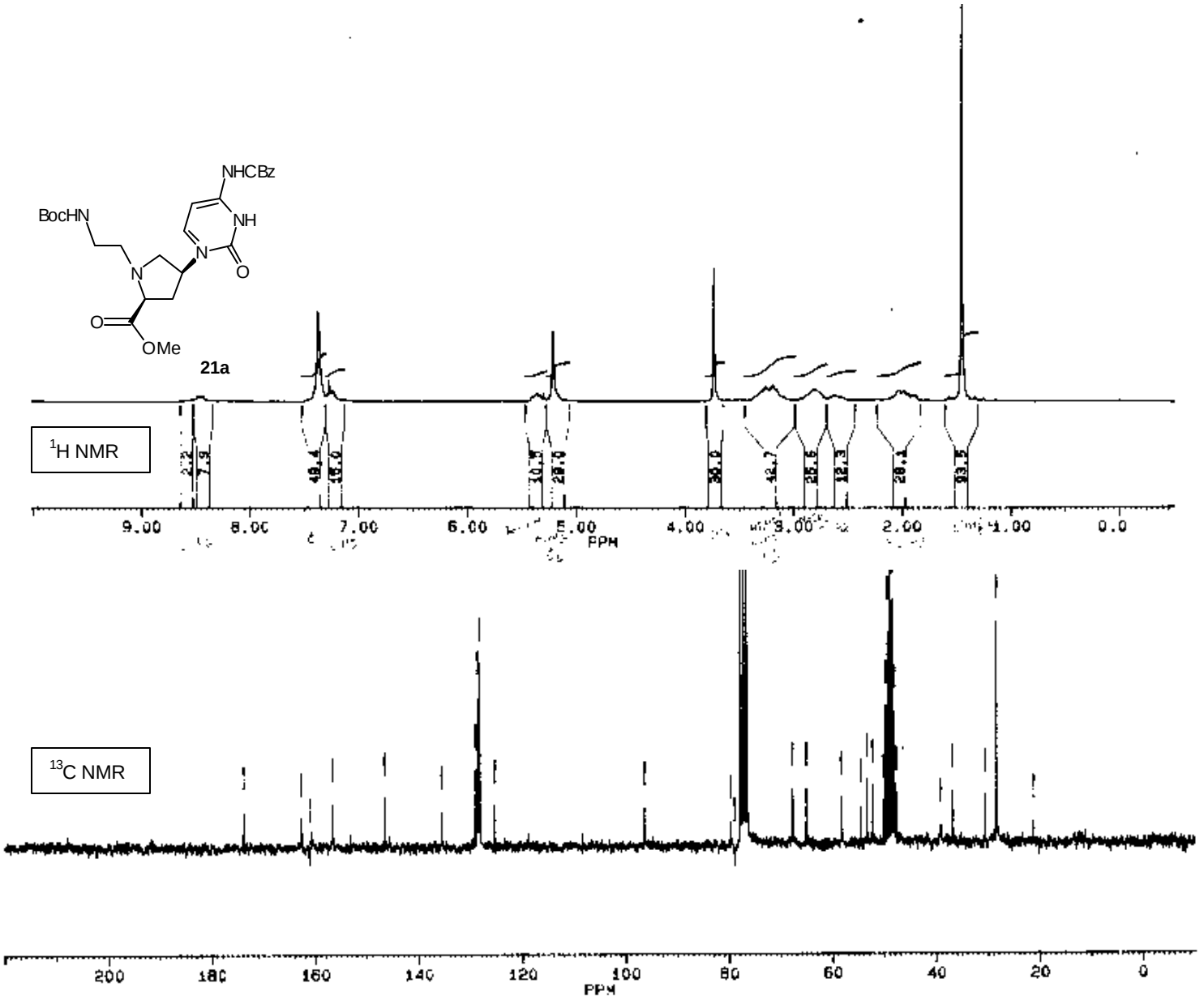

BocHN
N
OH
O
OMe
16
^{1}H NMR
^{13}C NMR
PPM

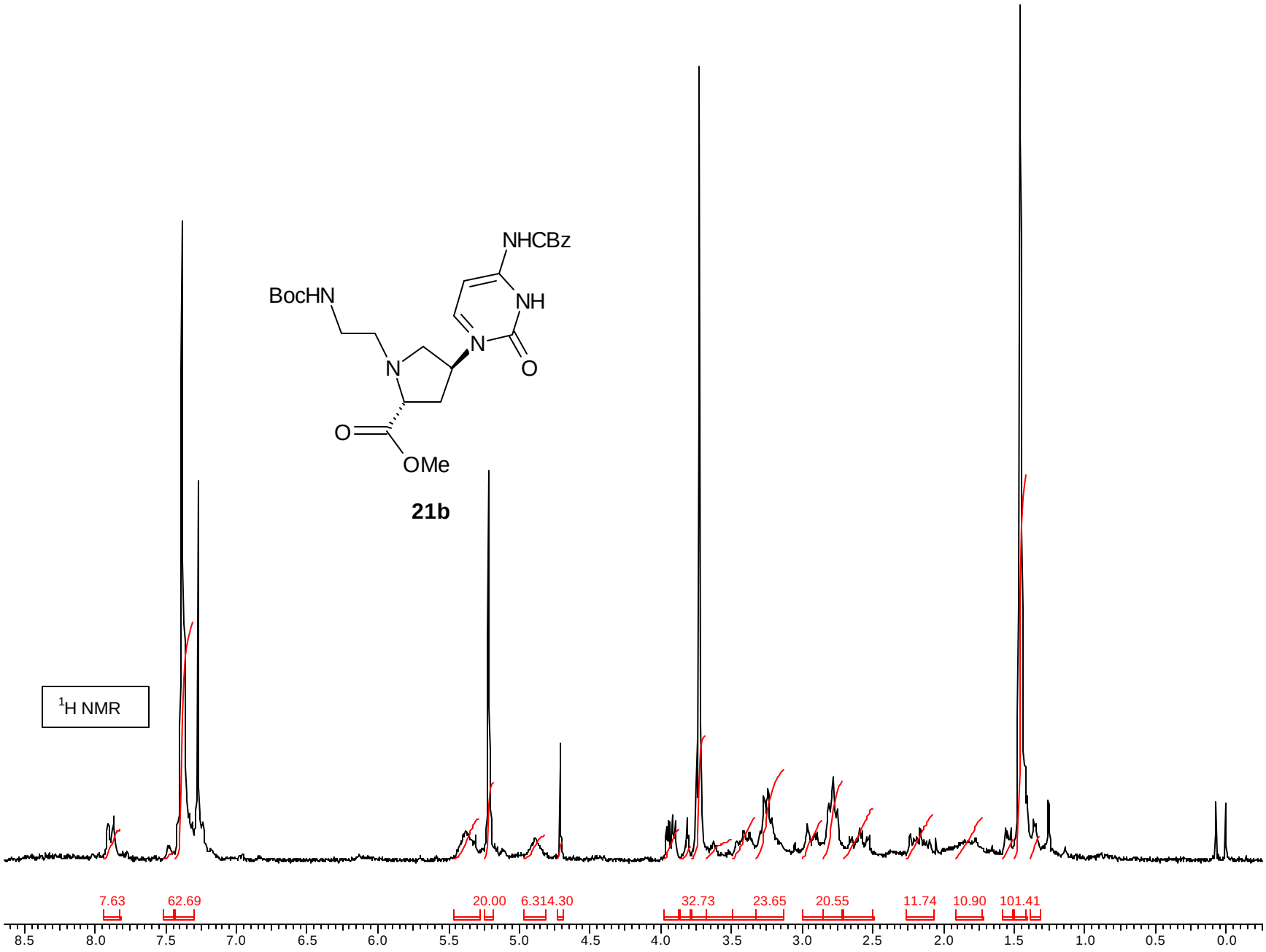

BocHN
N
O
O
N
N
O
O
OMe
17
^{1}H NMR
PPM

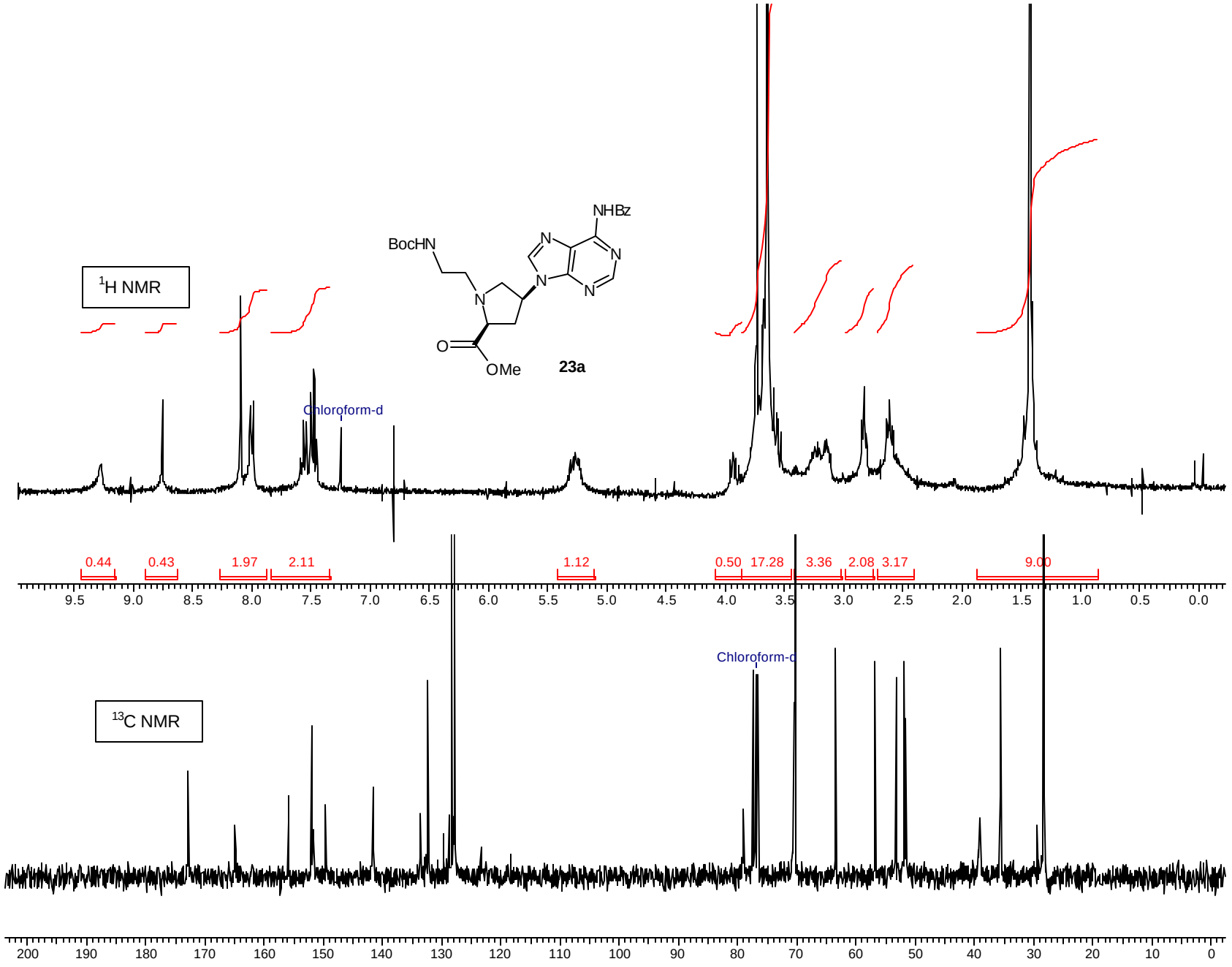

BocHN
O
N
N
O
N
O
OH
18
1H NMR
2.124
.735
1.011
3.578
3.894
1.117
1.064
3.941
10.385
8.0
7.0
6.0
PPM
5.0
4.0
3.0
2.0
1.0
13C NMR
160
140
120
100
PPM
80
60
40
20

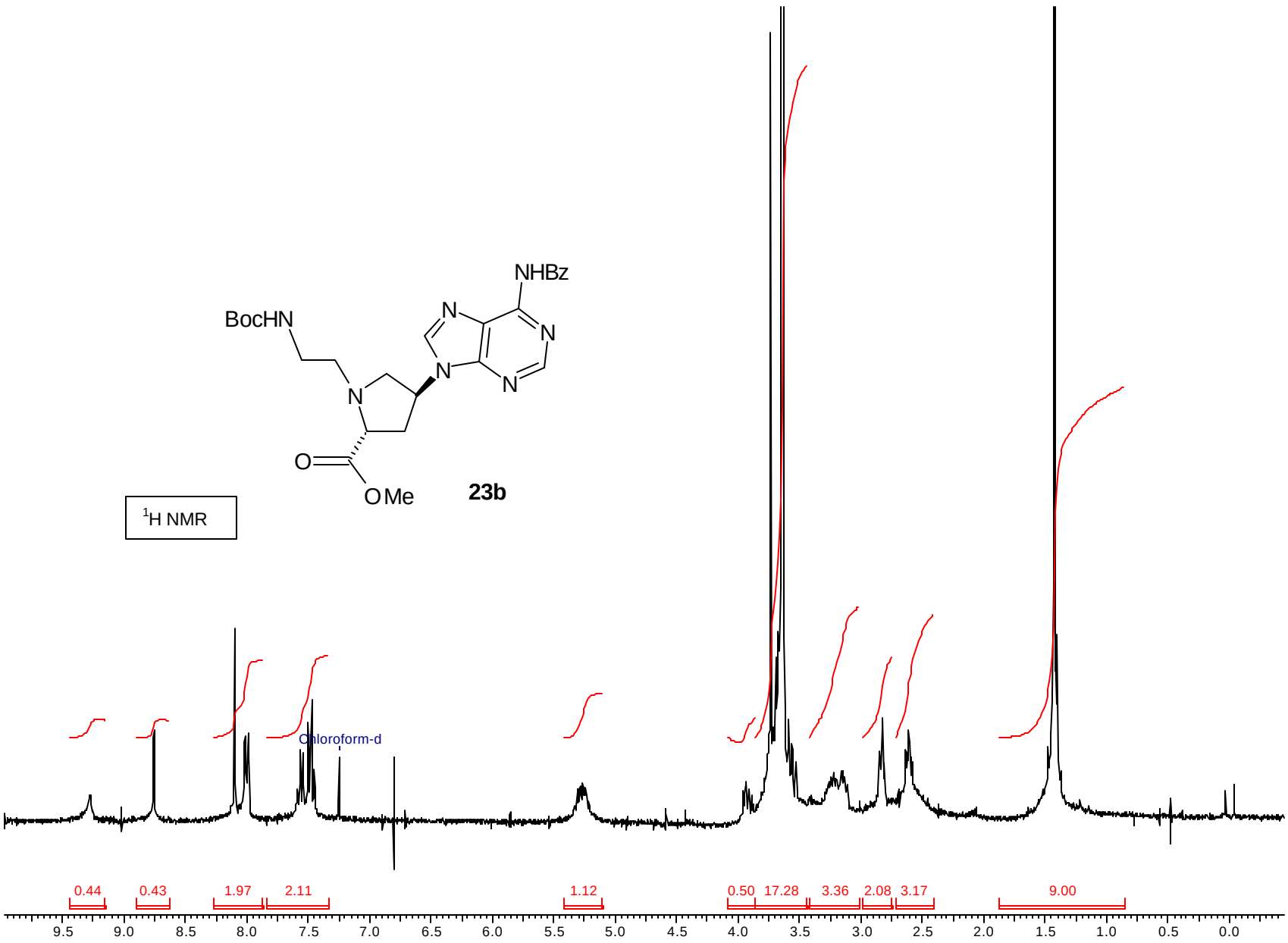

BocHN
N
OMs
O
OMe
19
^{1}H NMR
0
7.00
6.00
5.00
PPM
4.00
3.00
2.00
1.00

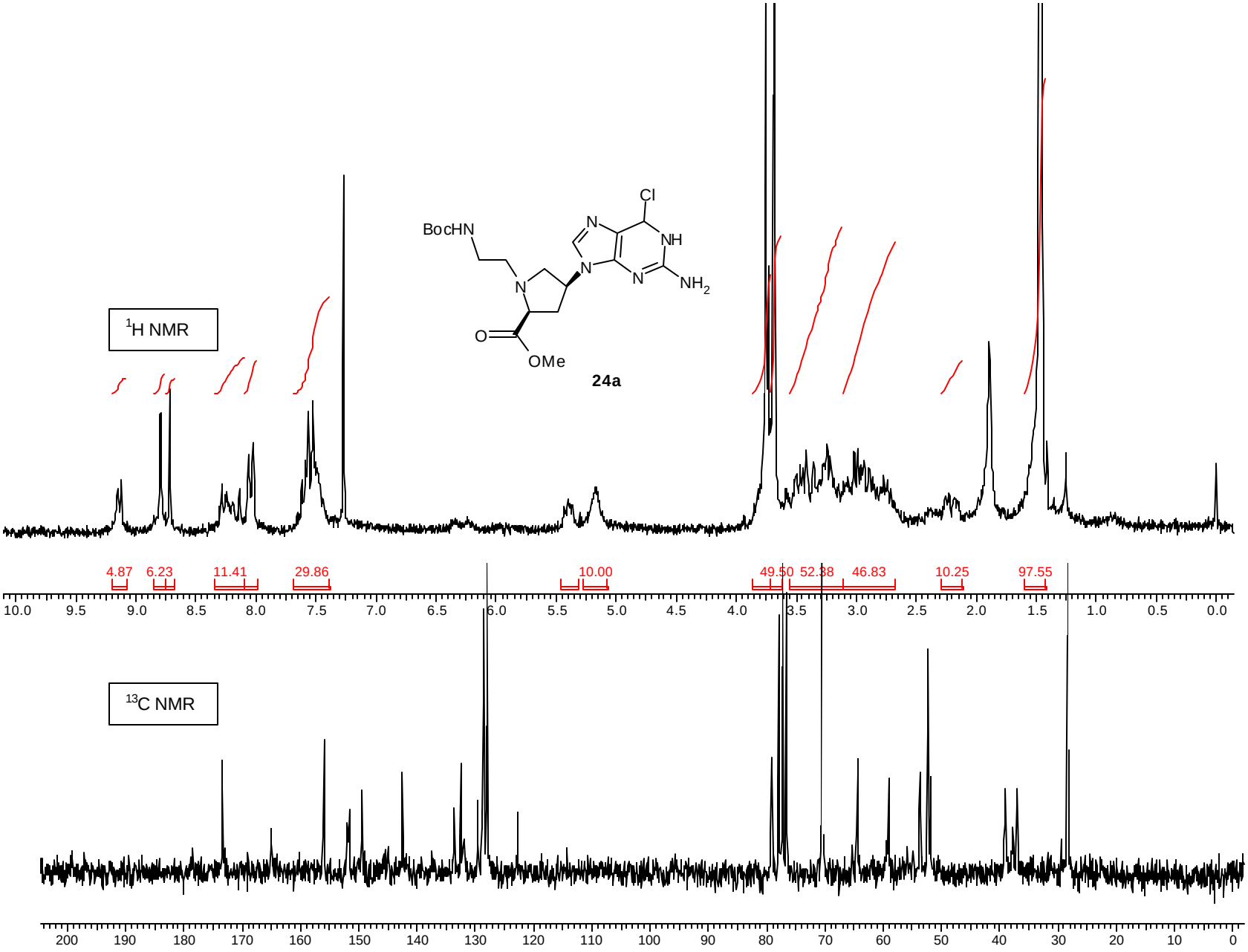

BocHN
N
OMs
O
OMe
20
1H NMR
TMS
1.09
0.32
3.00
1.33
3.05
1.48
2.15
1.00
0.69
9.73
0.04
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

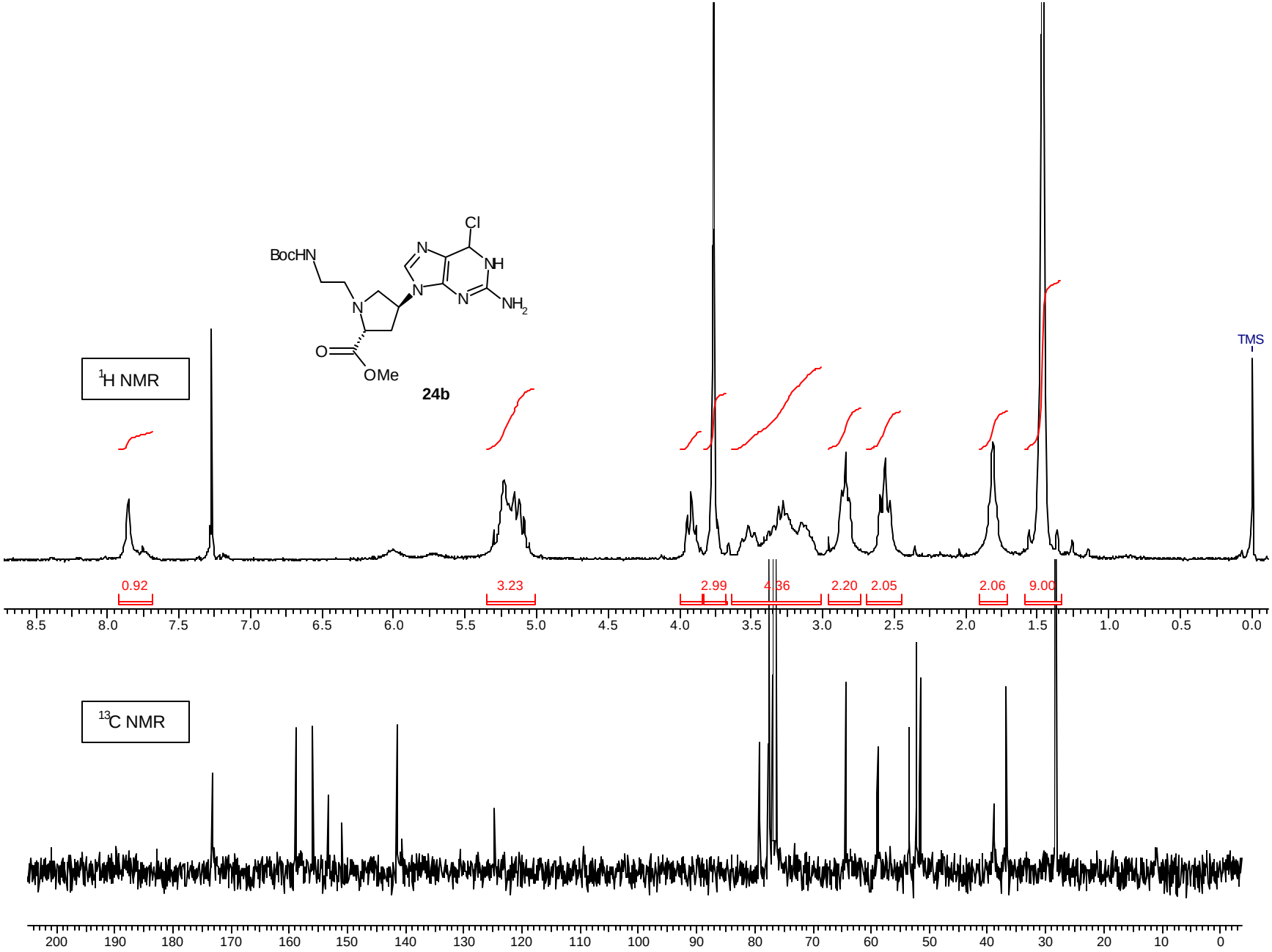

NHCBz
BocHN
NH
N
N
O
O
OMe
21a
^{1}H NMR
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.0
PPM
^{13}C NMR
200
180
160
140
120
100
80
60
40
20
0
PPM

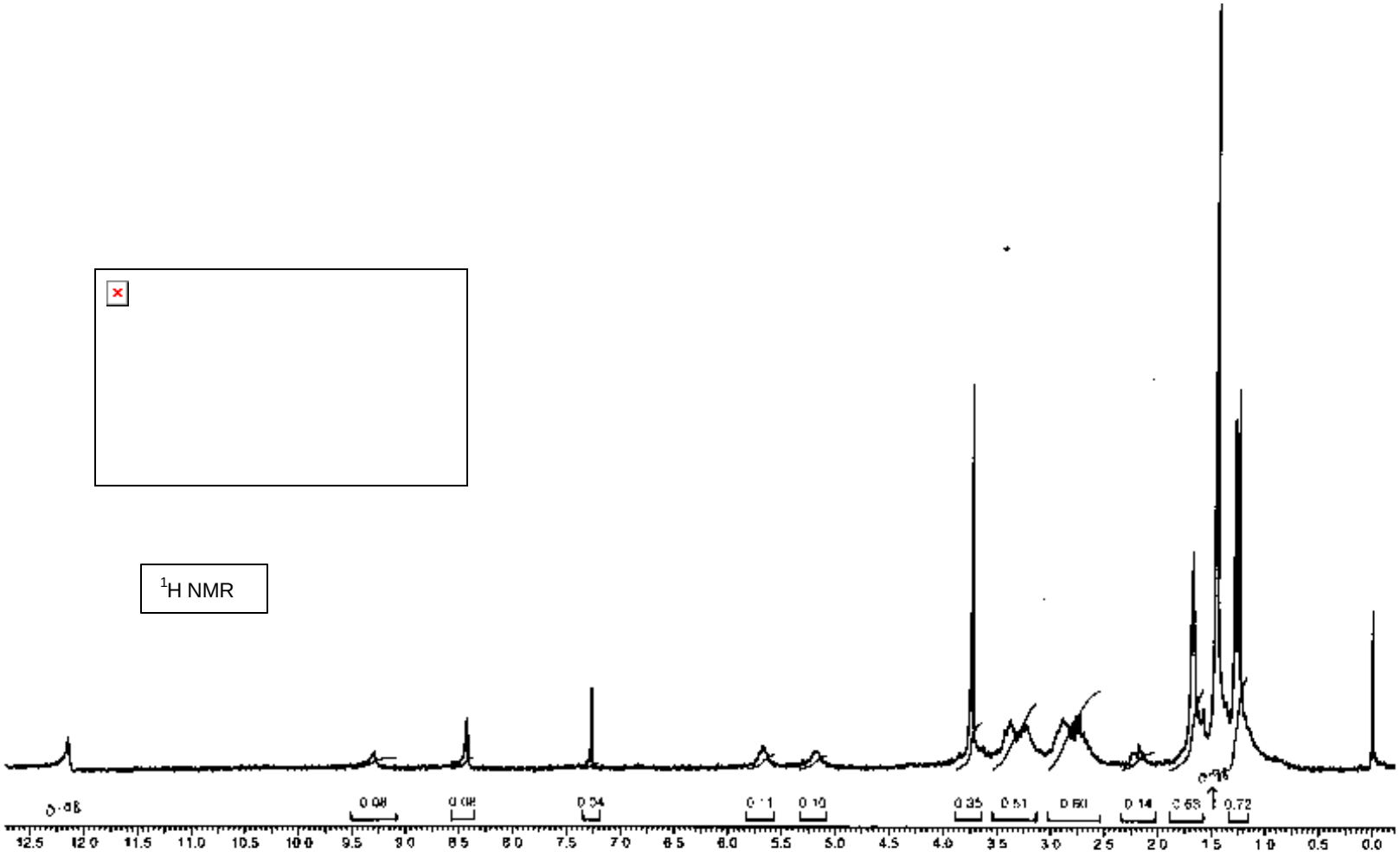

NHCBz
BocHN
NH
N
N
O
O
OMe
21b
^{1}H NMR
7.63
62.69
20.00
6.31
4.30
32.73
23.65
20.55
11.74
10.90
101.41
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

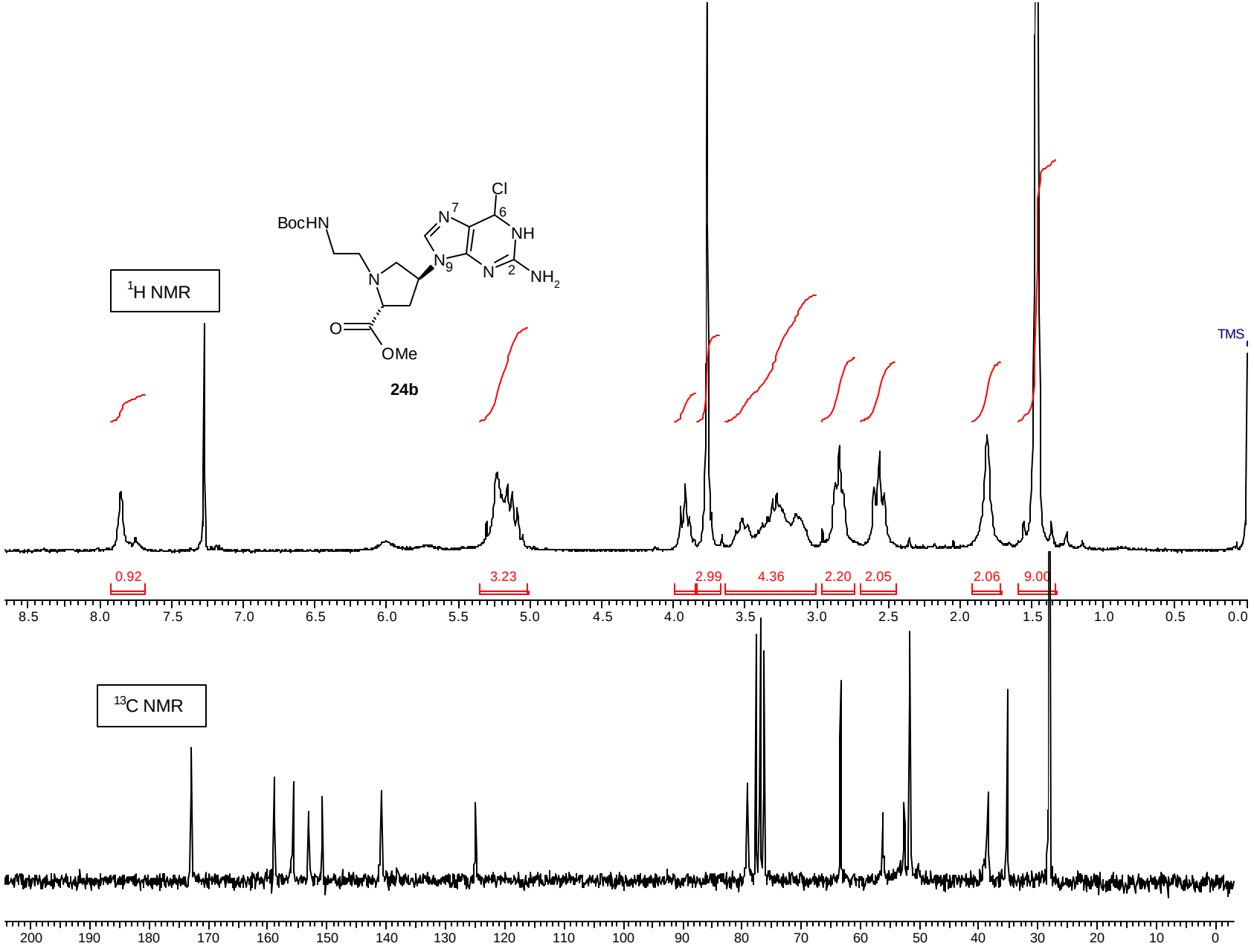

^{1}H NMR
BocHN
NHBz
N
N
N
N
N
O
OMe
23a
Chloroform-d
0.44
0.43
1.97
2.11
1.12
0.50
17.28
3.36
2.08
3.17
9.00
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
^{13}C NMR
Chloroform-d
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

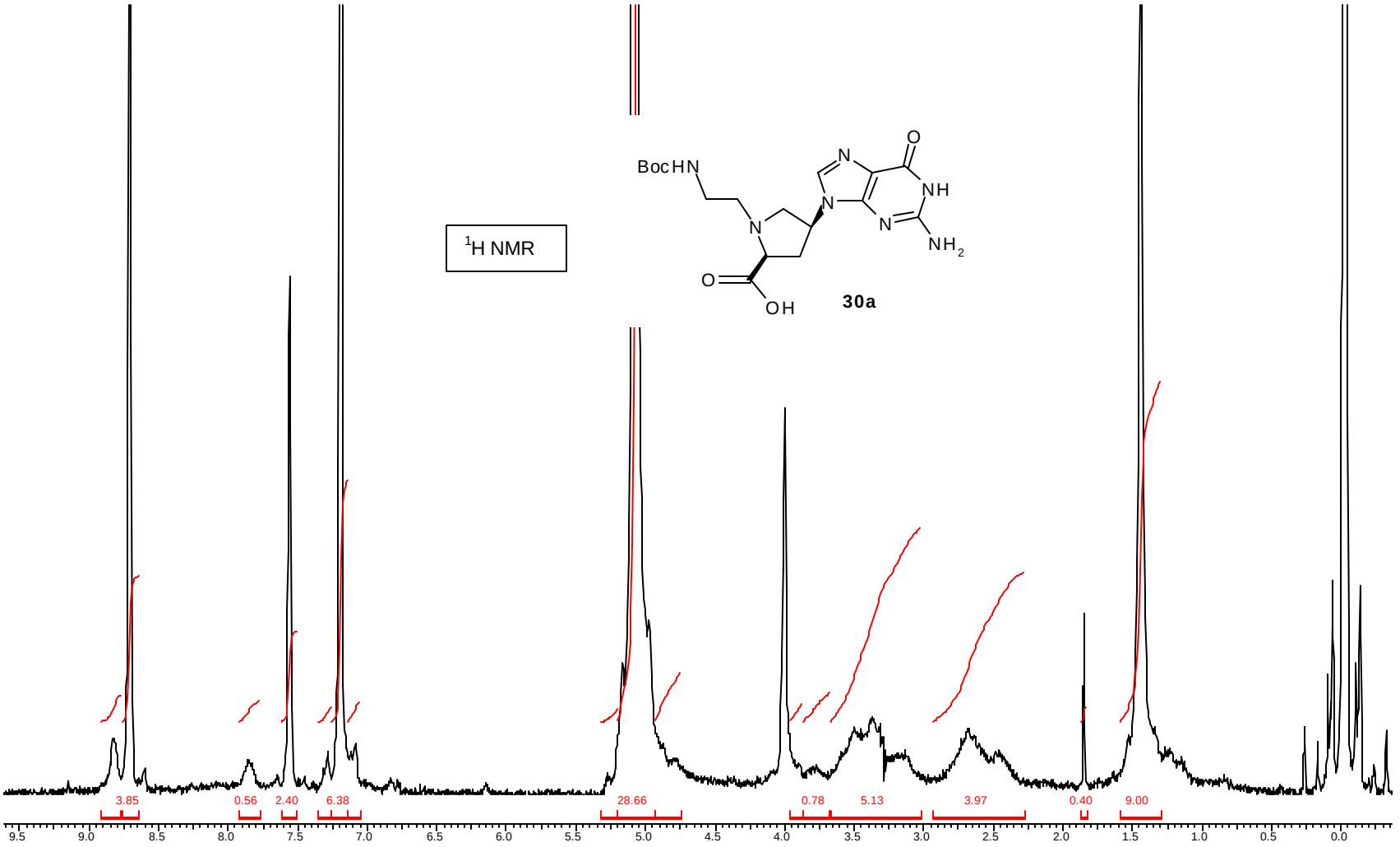

NHBz
BocHN
N
N
N
N
N
O
O Me
23b
^{1}H NMR
Chloroform-d
0.44
0.43
1.97
2.11
1.12
0.50
17.28
3.36
2.08
3.17
9.00
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

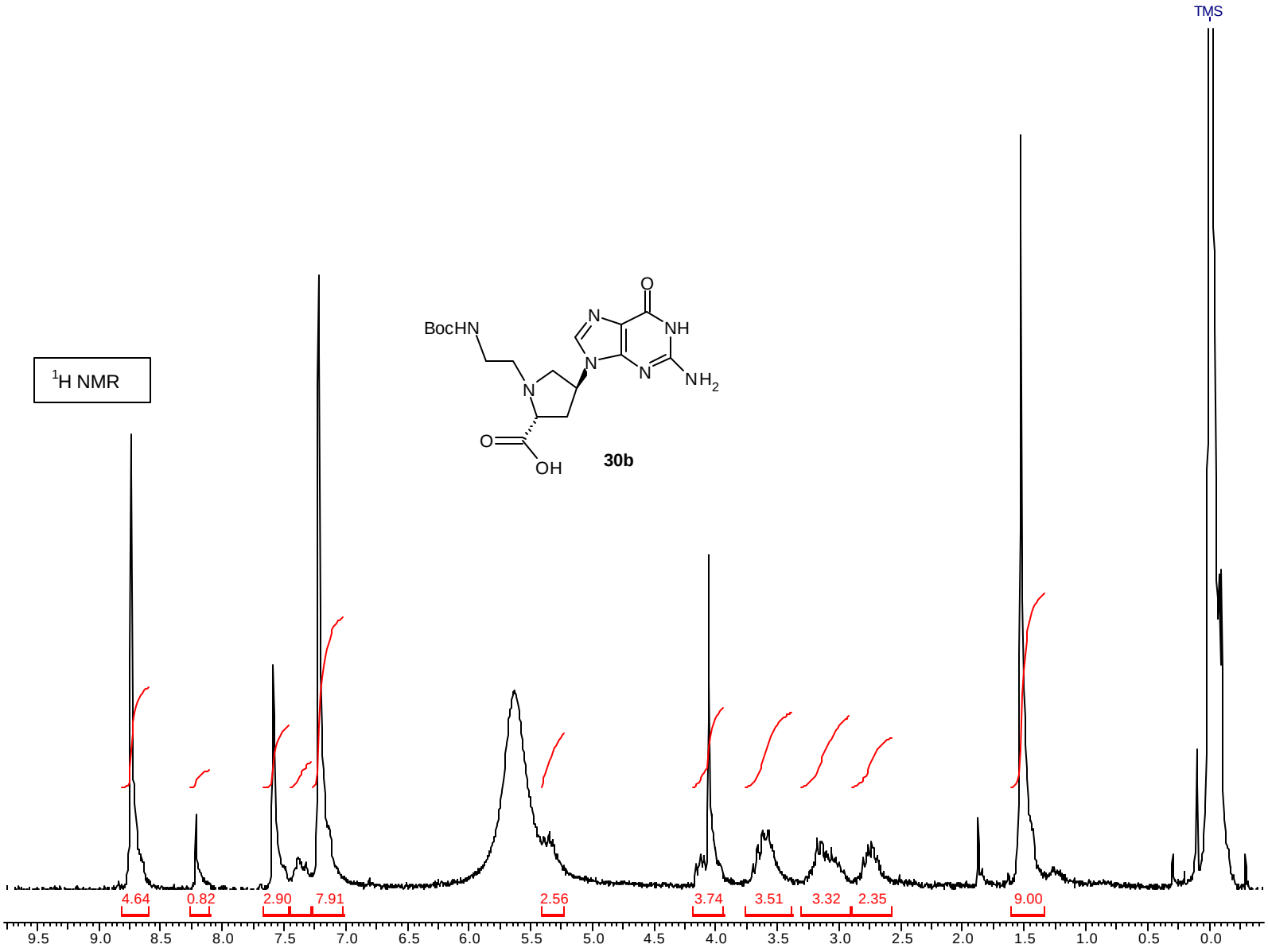

^{1}H NMR
BocHN
Cl
N
NH
N
N
N
NH$_2$
O
OMe
24a
4.87
6.23
11.41
29.86
10.00
49.50
52.38
46.83
10.25
97.55
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
^{13}C NMR
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

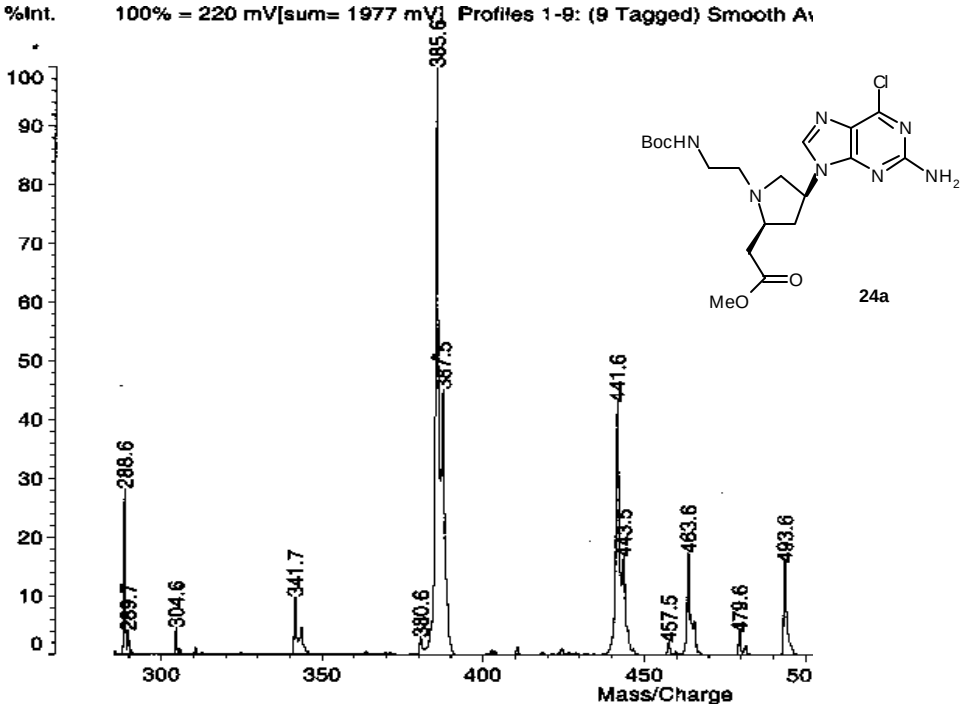

BocHN
Cl
N
NH
N
N
NH2
N
O
OMe
24b
1H NMR
TMS
0.92
3.23
2.99
4.36
2.20
2.05
2.06
9.00
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
13C NMR
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

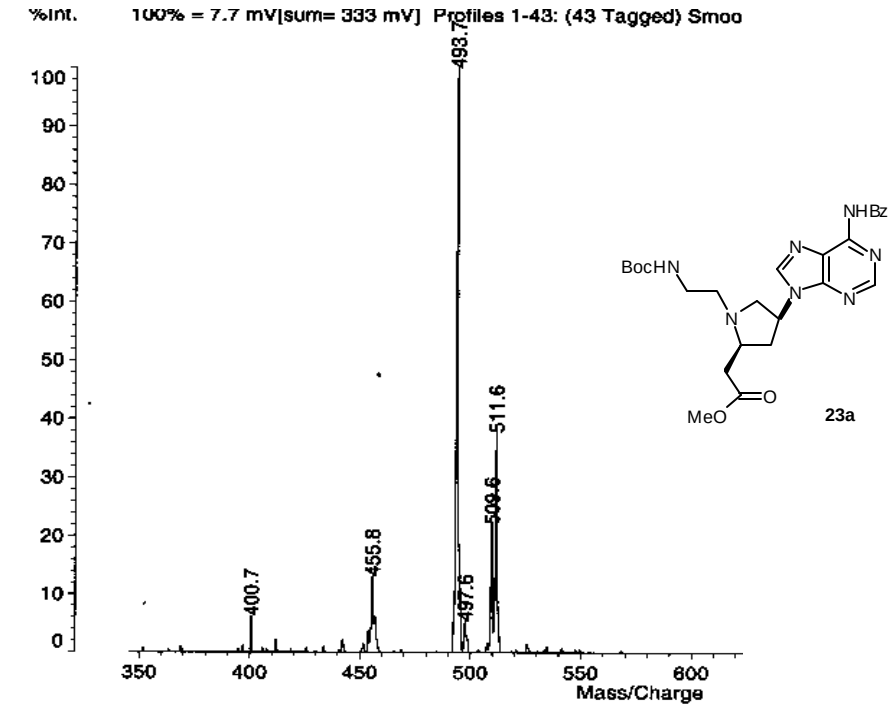

^{1}H NMR

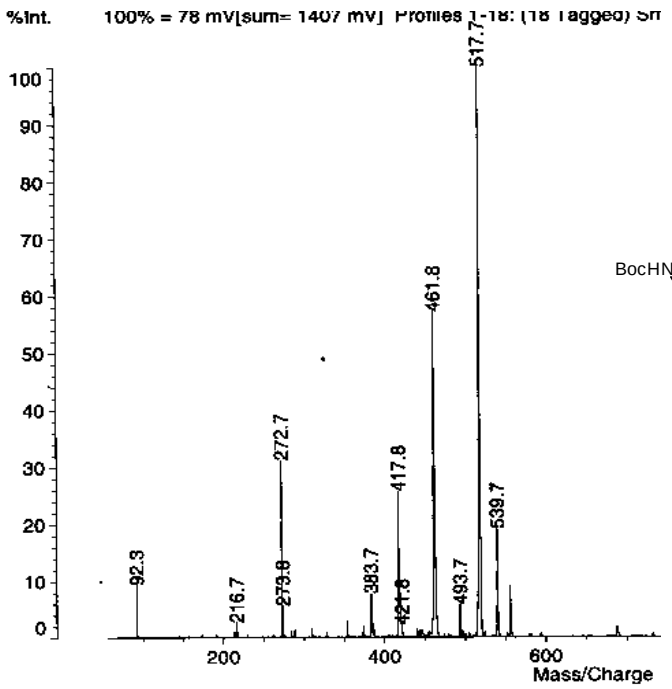

^{1}H NMR
BocHN
Cl
6
7
N
NH
N
9
N
2
NH$_2$
N
O
OMe
24b
TMS
0.92
3.23
2.99
4.36
2.20
2.05
2.06
9.00
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
^{13}C NMR
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

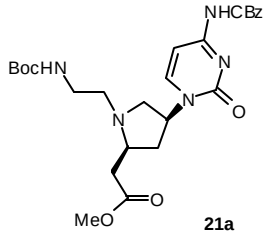

1H NMR
BocHN
O
N
NH
N
N
N
NH2
O
OH
30a
3.85
0.56
2.40
6.38
28.66
0.78
5.13
3.97
0.40
9.00
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0


TMS
^{1}H NMR
BocHN
O
N
NH
N
N
N
NH$_2$
O
OH
30b
4.64
0.82
2.90
7.91
2.56
3.74
3.51
3.32
2.35
9.00
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0


%Int.
100% = 220 mV[sum= 1977 mV] Profiles 1-9: (9 Tagged) Smooth A
385.6
387.5
441.6
288.6
341.7
443.5
463.6
493.6
304.6
380.6
457.5
478.6
Mass/Charge
BocHN
Cl
NH2
MeO
24a

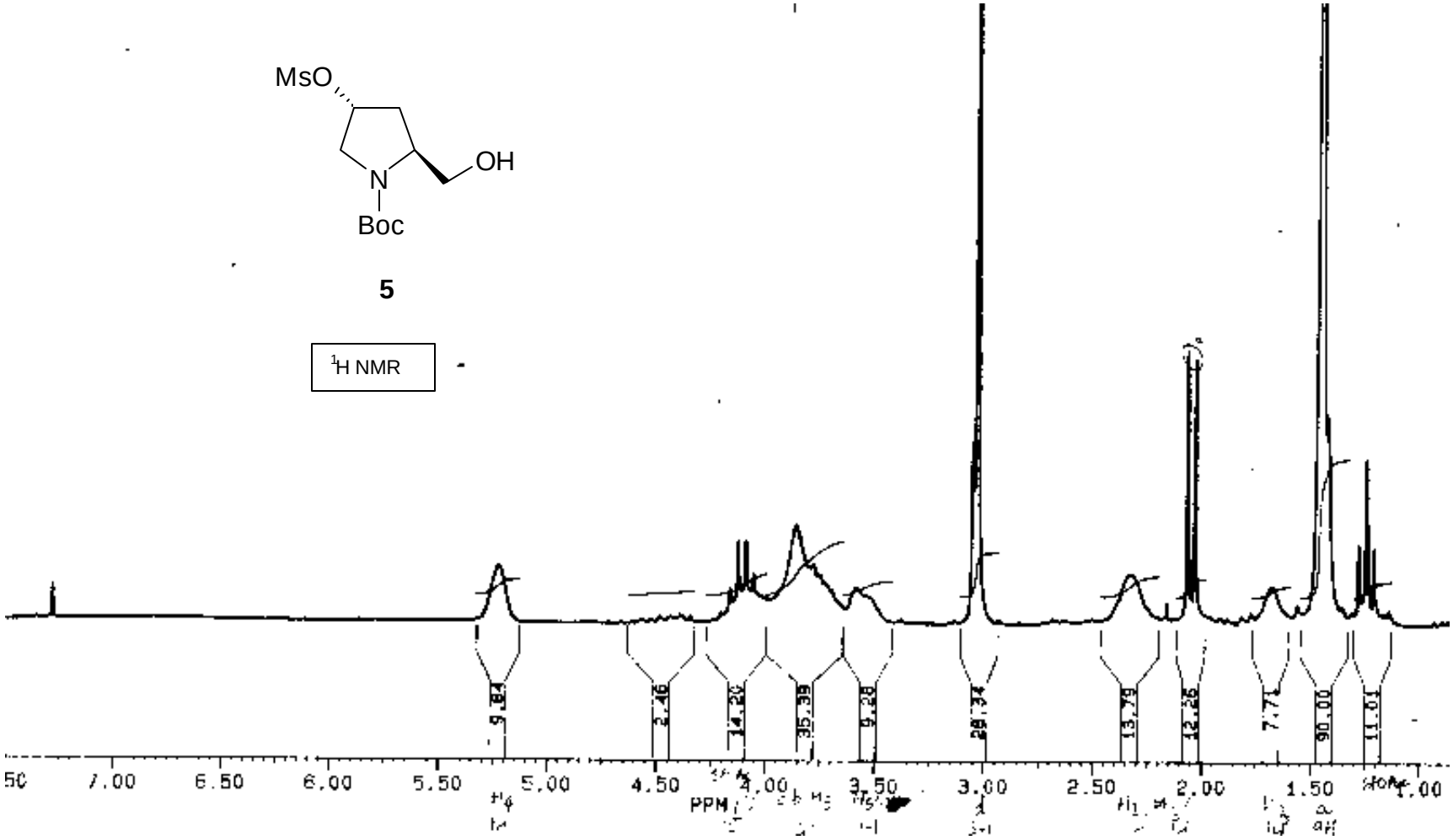

%Int.
100% = 7.7 mV[sum= 333 mV] Profiles 1-43: (43 Tagged) Smoo
493.7
511.6
509.6
455.8
400.7
497.6
Mass/Charge
BocHN
NHBz
MeO
23a

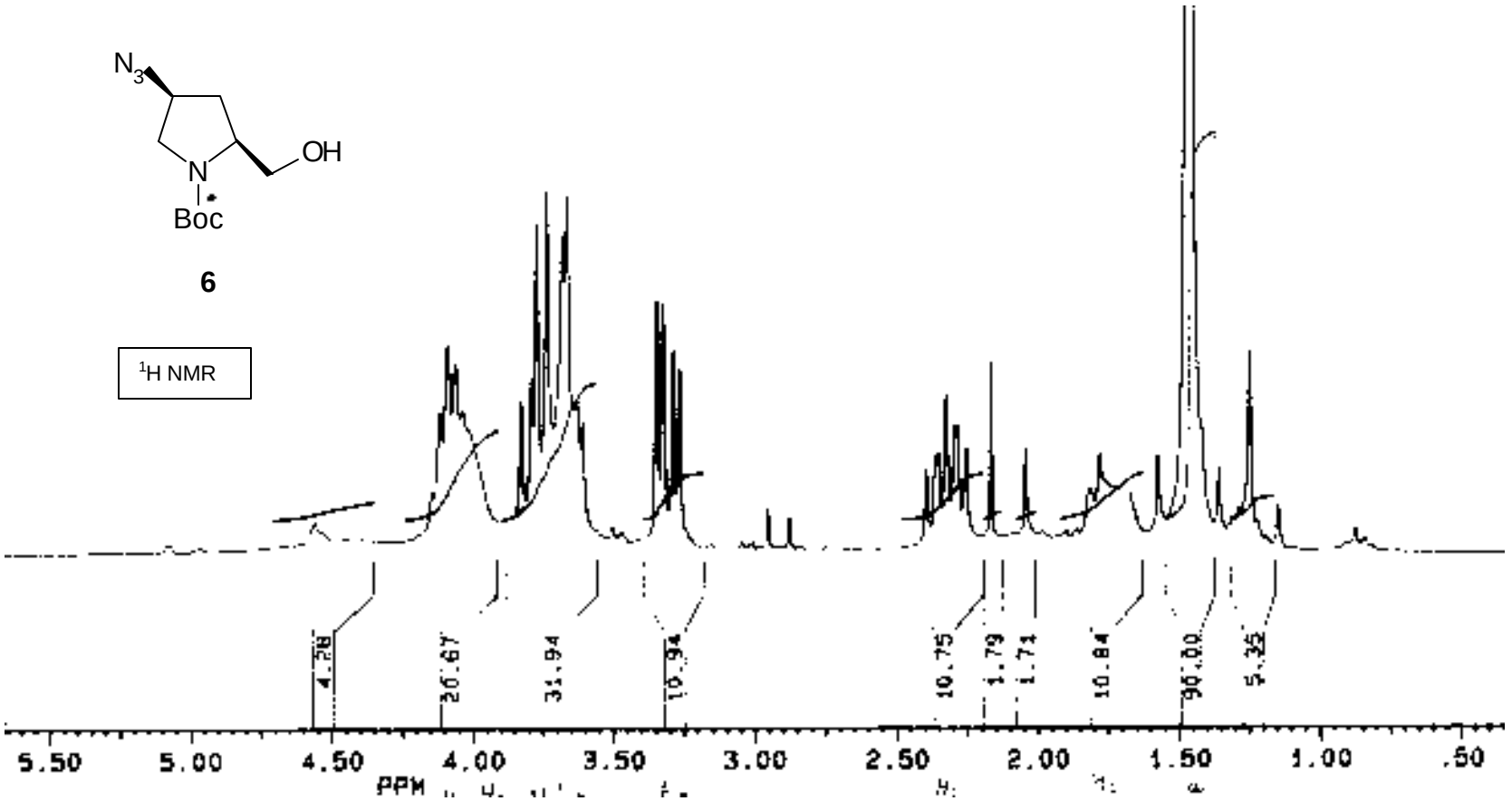

%Int.
100% = 78 mV[sum= 1407 mV]
100
90
80
70
60
50
40
30
20
10
0
92.3
216.7
272.7
279.8
383.7
417.8
421.8
461.8
493.7
517.7
539.7
200
400
600
Mass/Charge

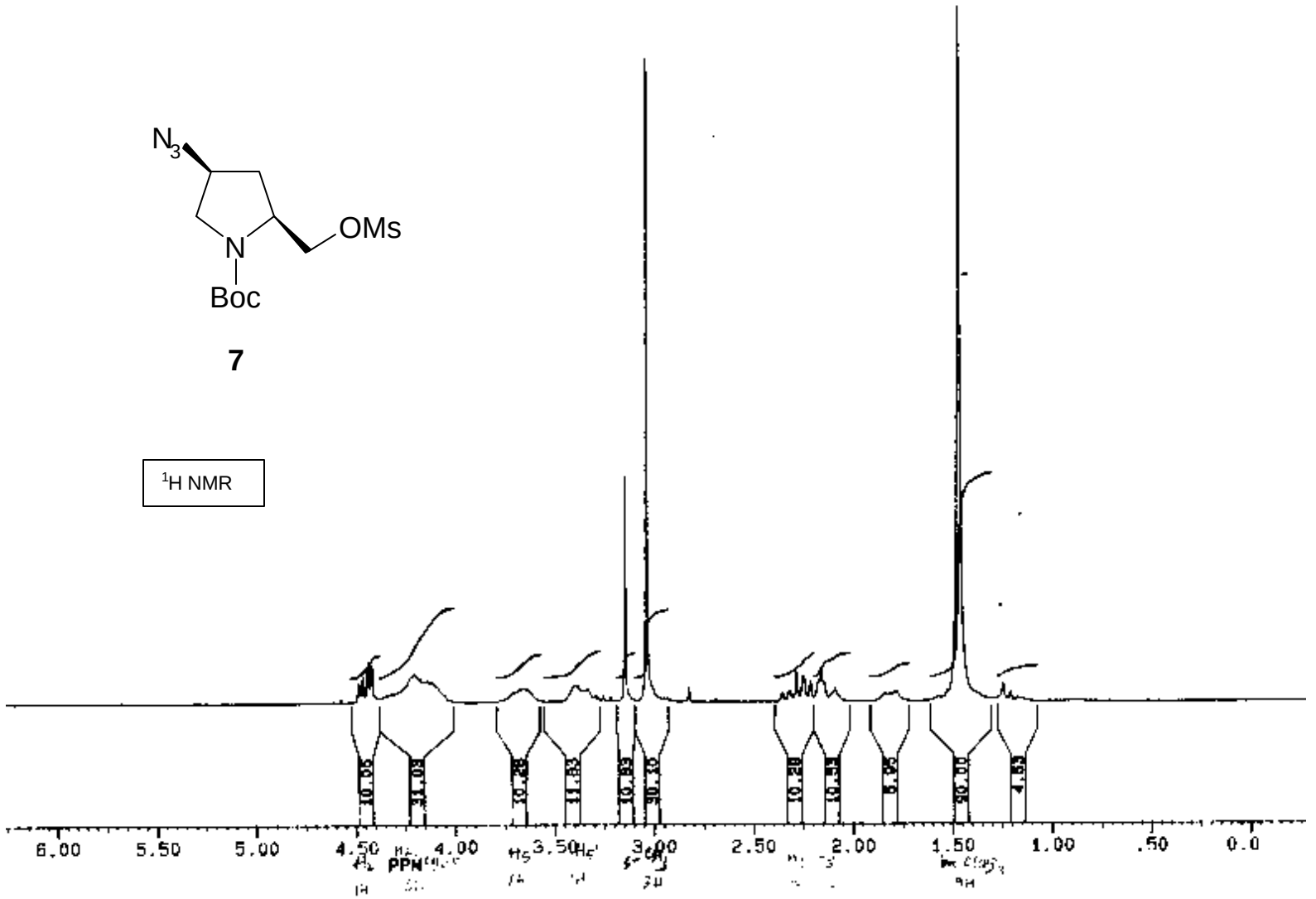

NHCBz
BocHN
N
N
N
O
O
MeO

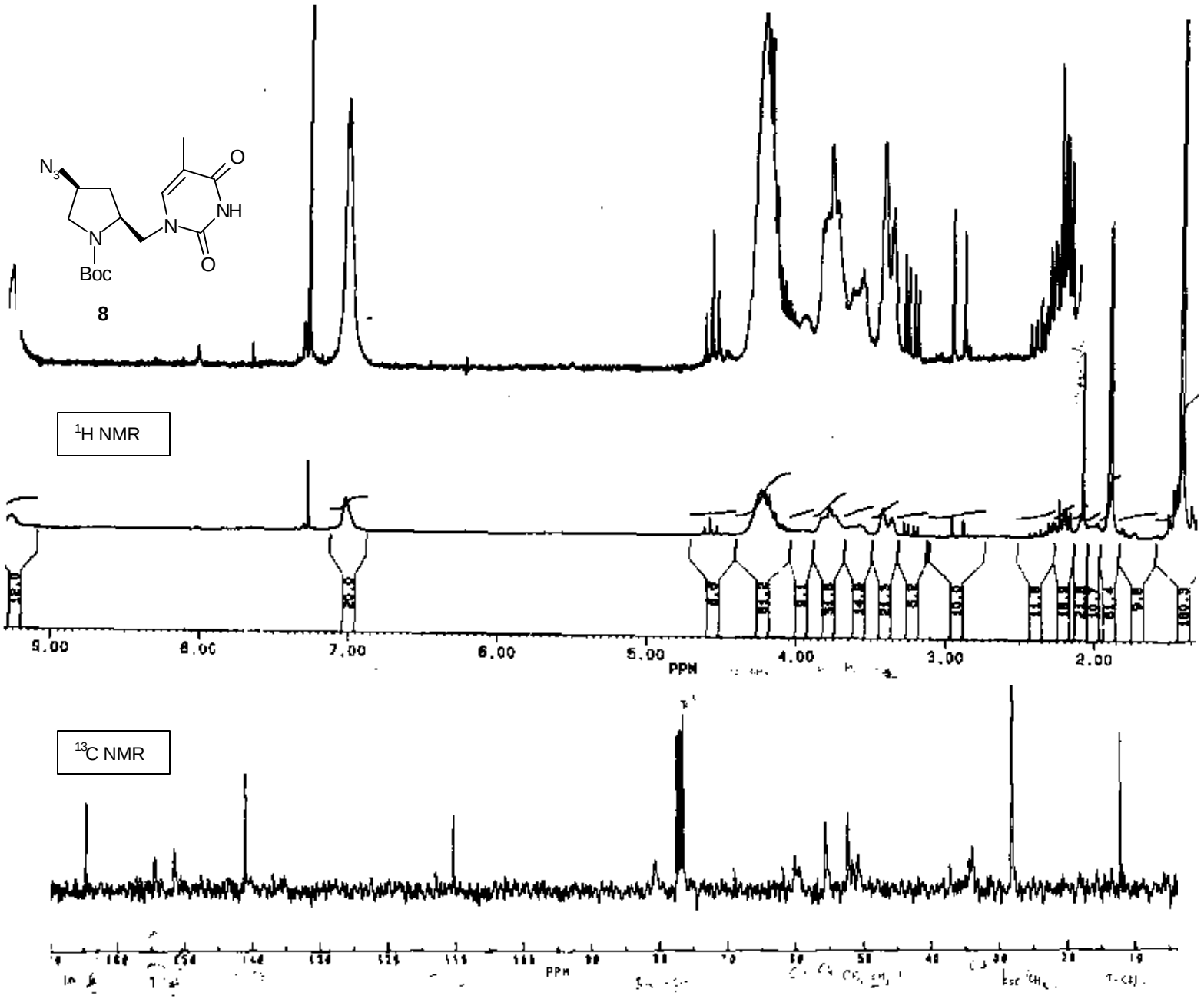
21a

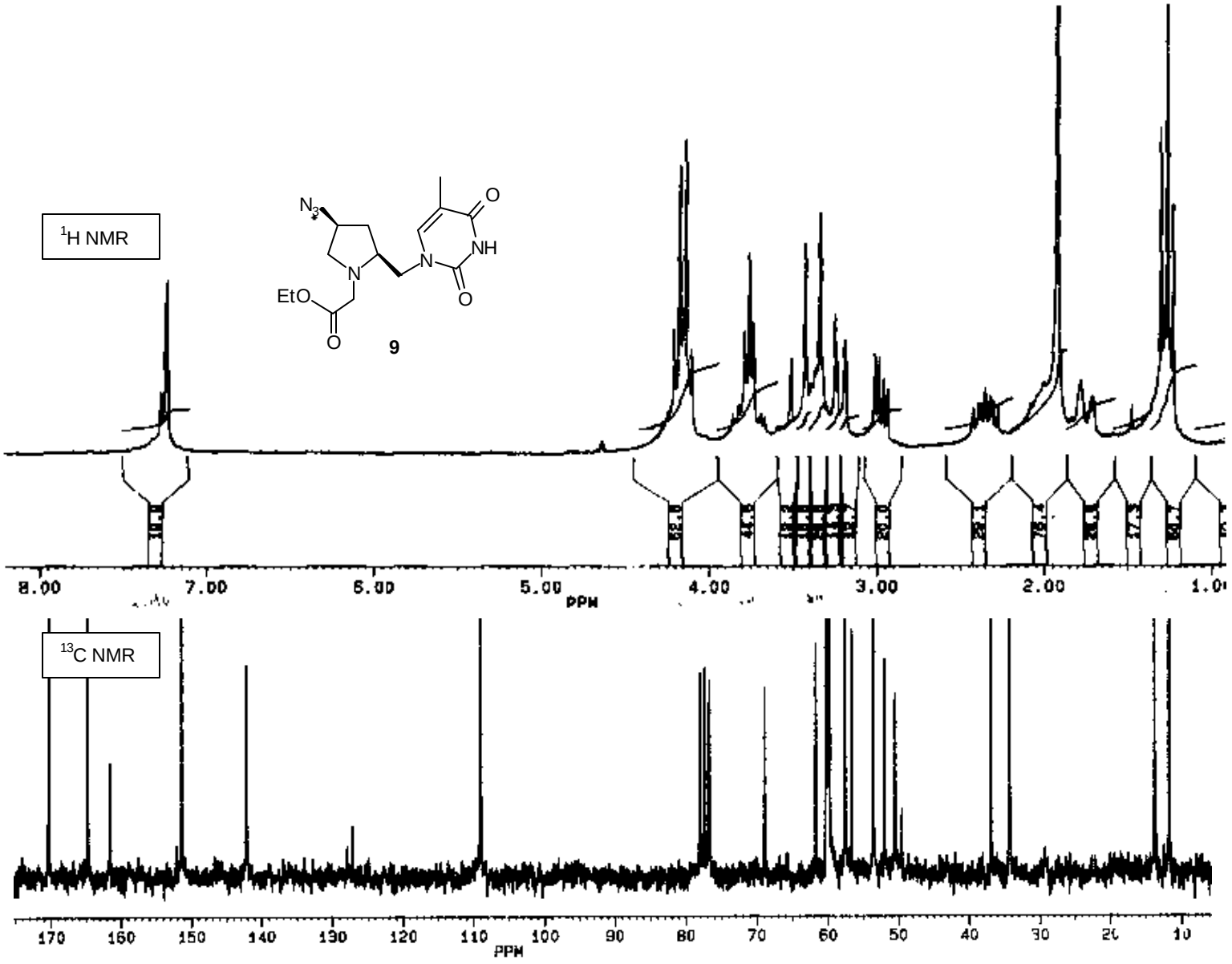
Appendix to Chapter 4

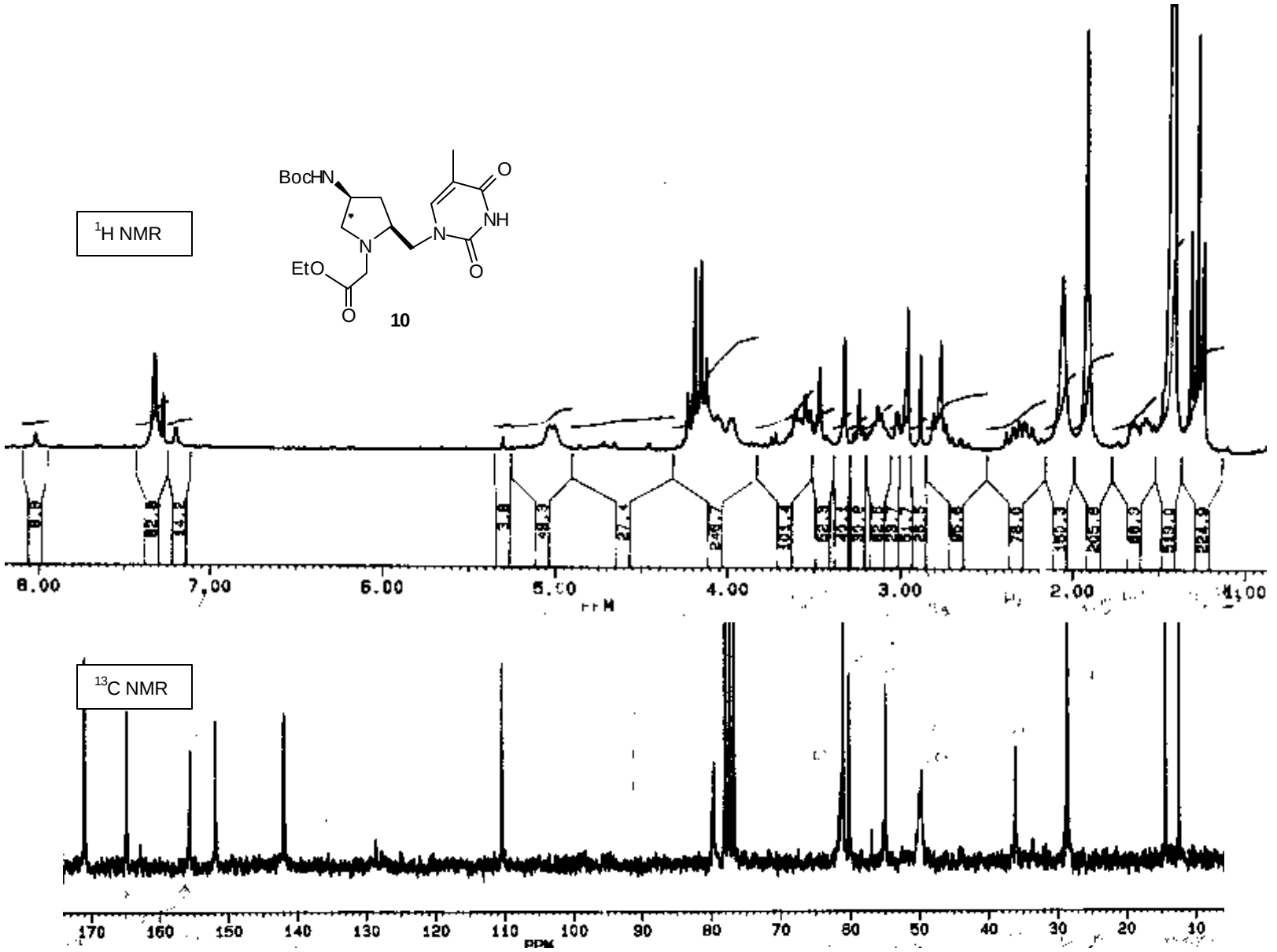

MsO
OH
N
Boc
5
^{1}H NMR
9.84
2.46
14.20
35.39
9.28
28.34
13.79
12.26
7.71
90.00
11.01
7.00
6.50
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
PPM

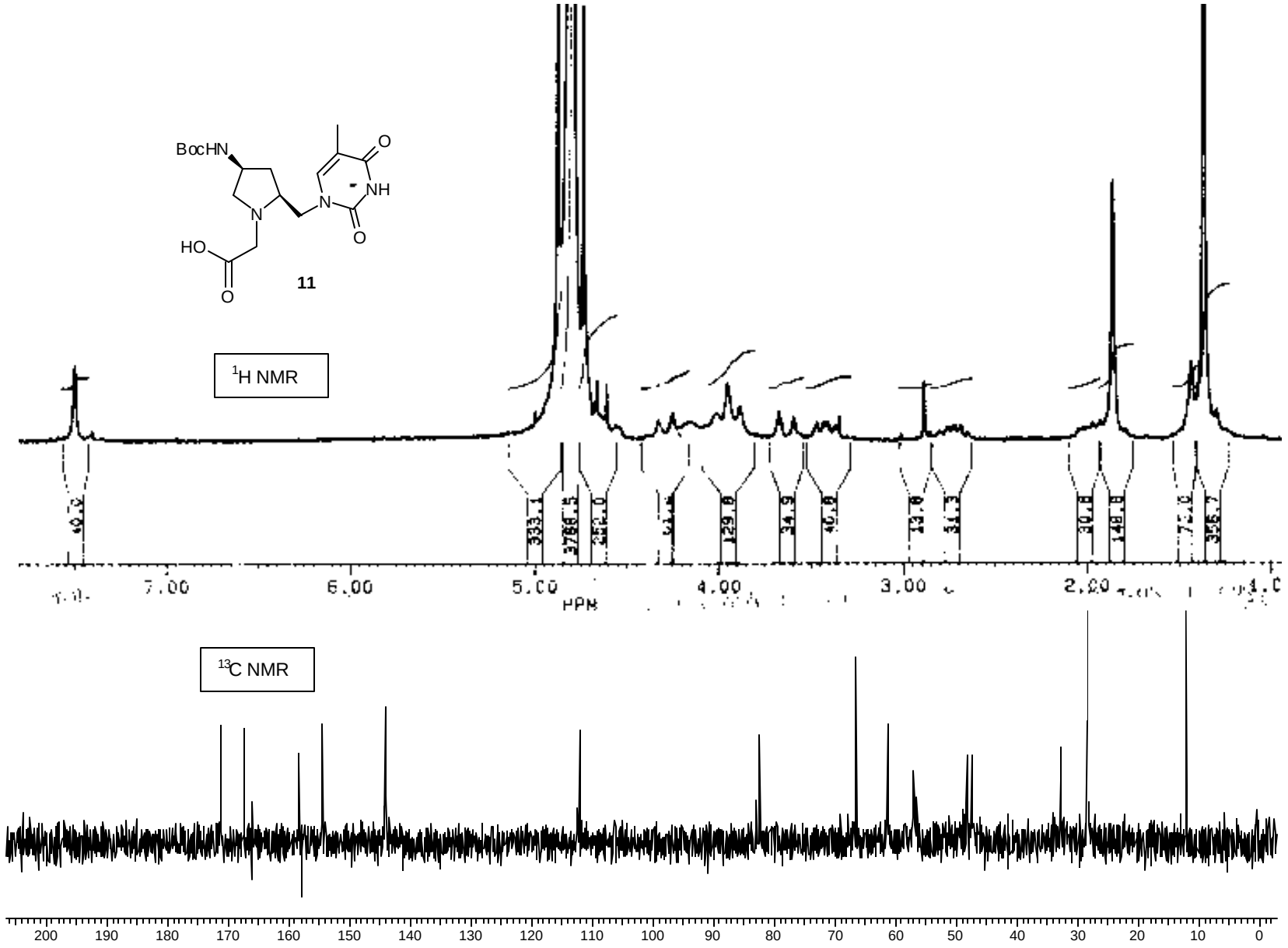

N3
OH
N
Boc
6
1H NMR
4.26
26.67
31.94
10.94
10.75
1.79
1.71
10.84
90.00
5.35
5.50
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
.50
PPM

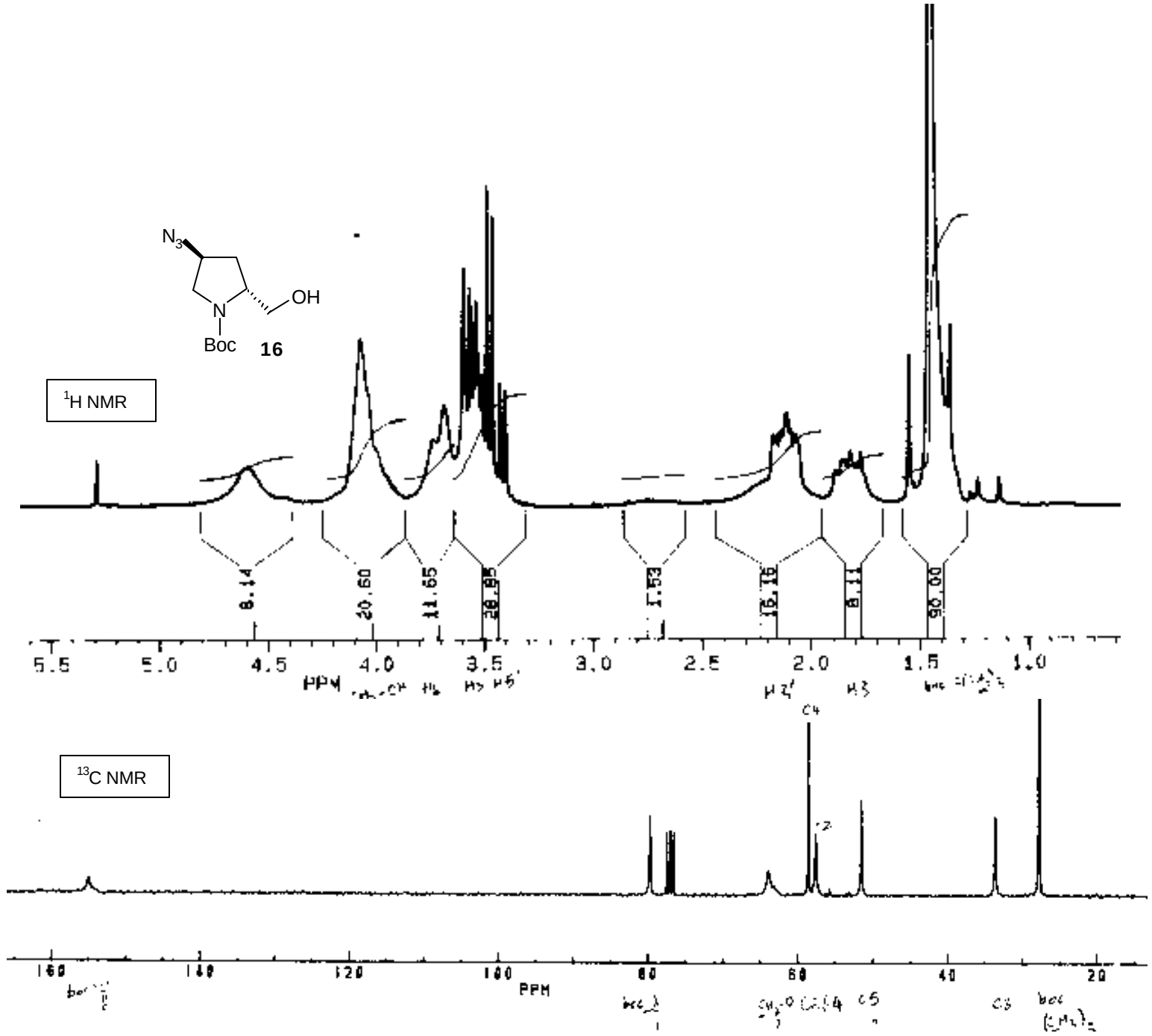

N3
OMs
N
Boc
7
1H NMR
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
.50
0.0
PPM

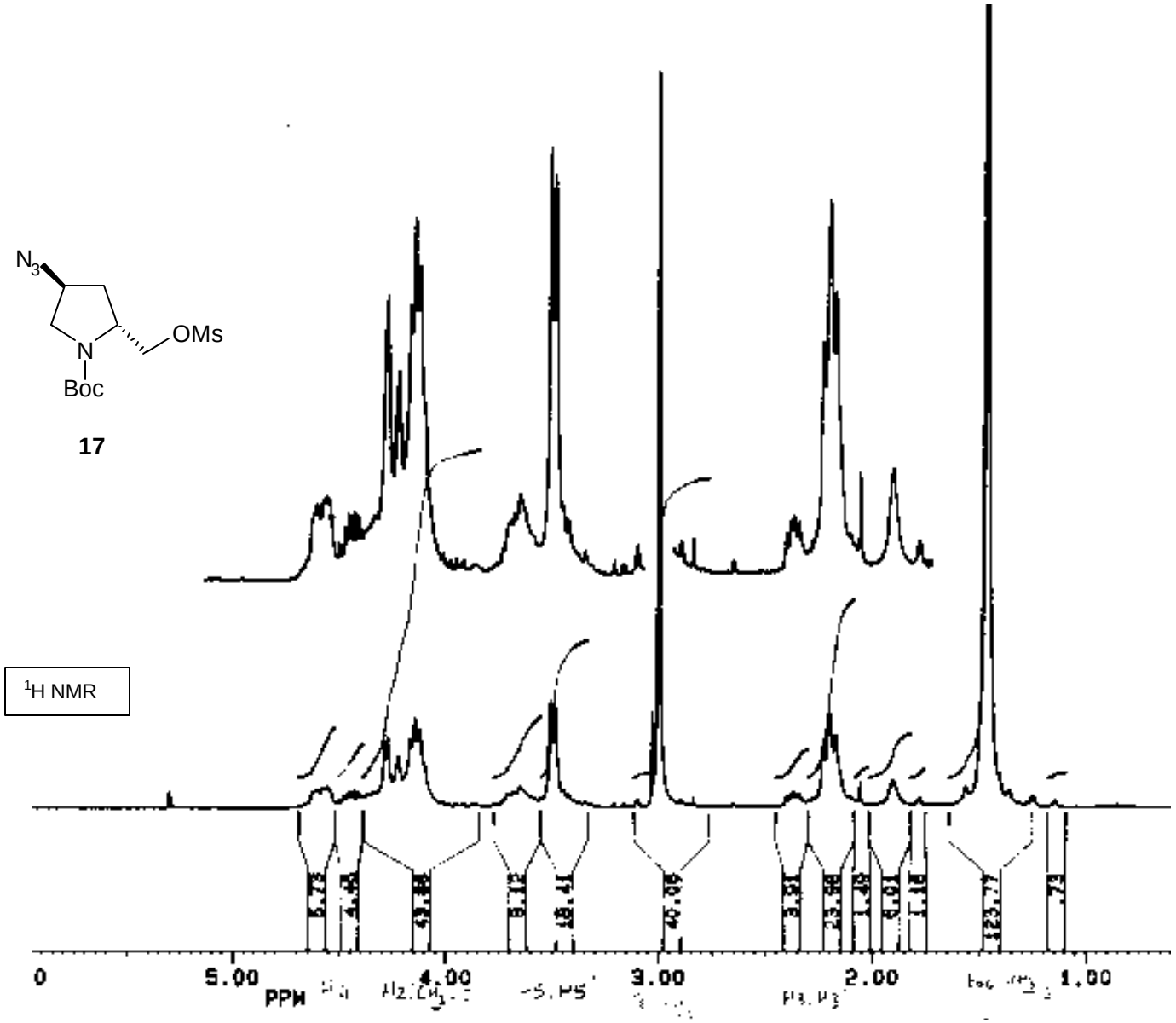

N3
O
NH
N
N
O
Boc
8
1H NMR
9.00
8.00
7.00
6.00
5.00
PPM
4.00
3.00
2.00
13C NMR
PPM

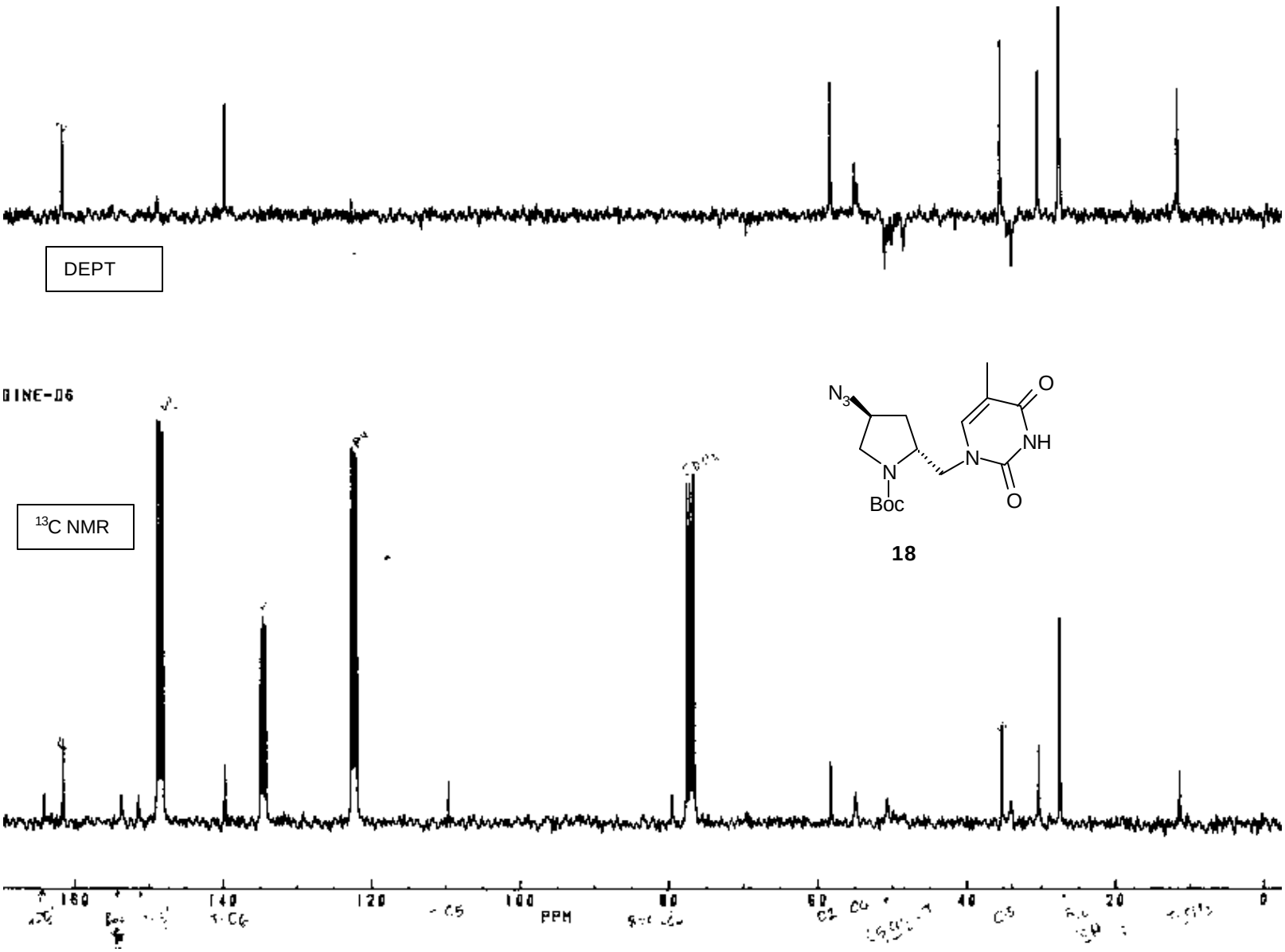

^{1}H NMR
N_3
O
NH
N
N
O
EtO
O
9
8.00
7.00
6.00
5.00
PPM
4.00
3.00
2.00
1.0
^{13}C NMR
170
160
150
140
130
120
110
PPM
100
90
80
70
60
50
40
30
20
10


^{1}H NMR
BocHN
EtO
O
NH
N
O
10
8.00
7.00
6.00
5.00
4.00
3.00
2.00
^{13}C NMR
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
ppm


BocHN
O
NH
N
N
O
HO
O
11
^{1}H NMR
7.00
6.00
5.00
4.00
3.00
2.00
PPM
^{13}C NMR
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

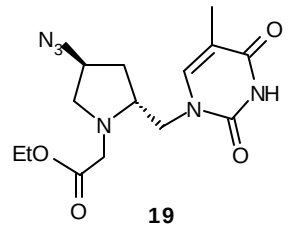

N3
OH
N
Boc
16
1H NMR
8.14
20.60
11.65
20.85
1.53
15.15
8.11
90.00
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
PPM
13C NMR
160
140
120
100
80
60
40
20
PPM


N3
OMs
N
Boc
17
1H NMR
0
5.00
PPM
4.00
3.00
2.00
1.00

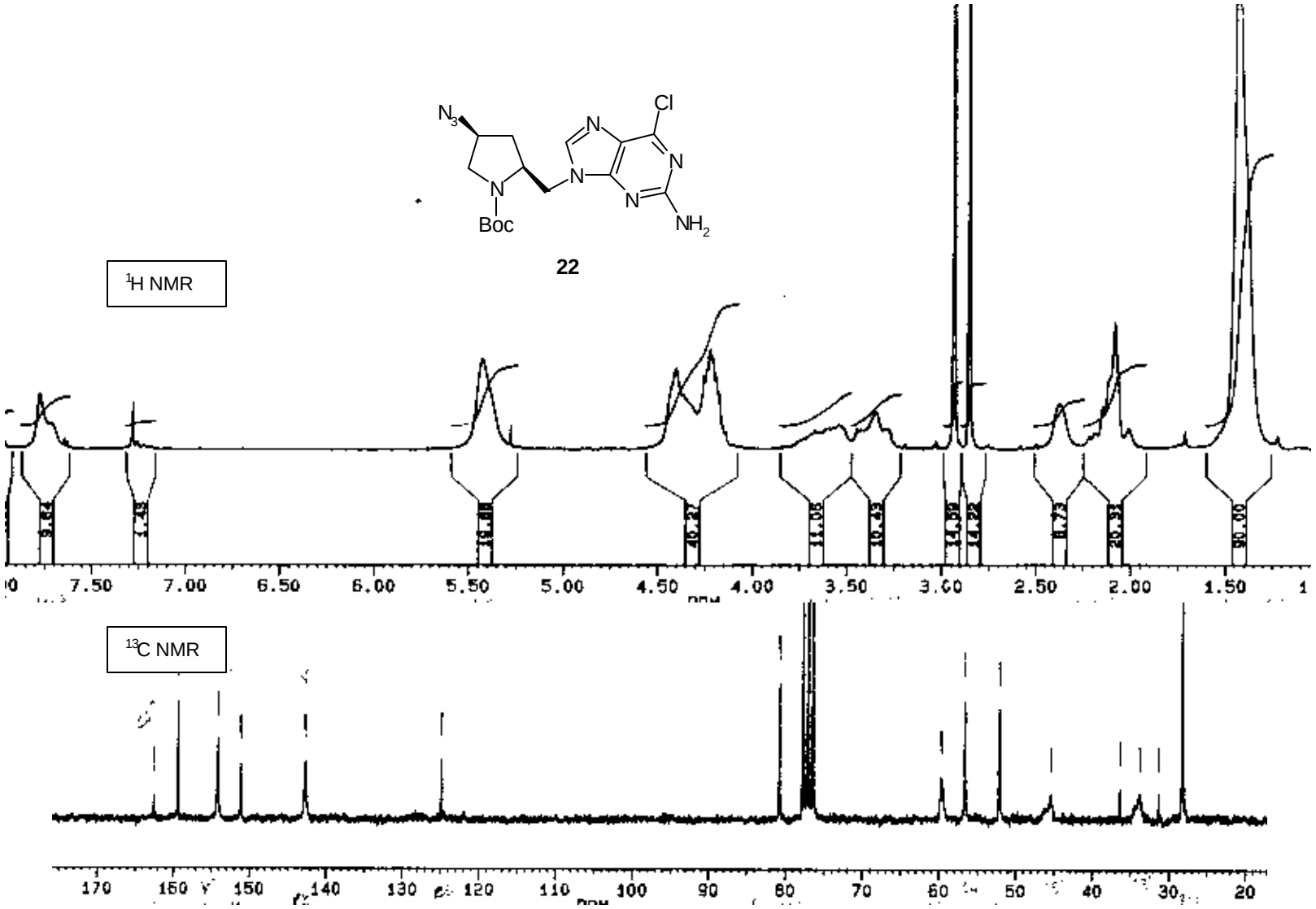


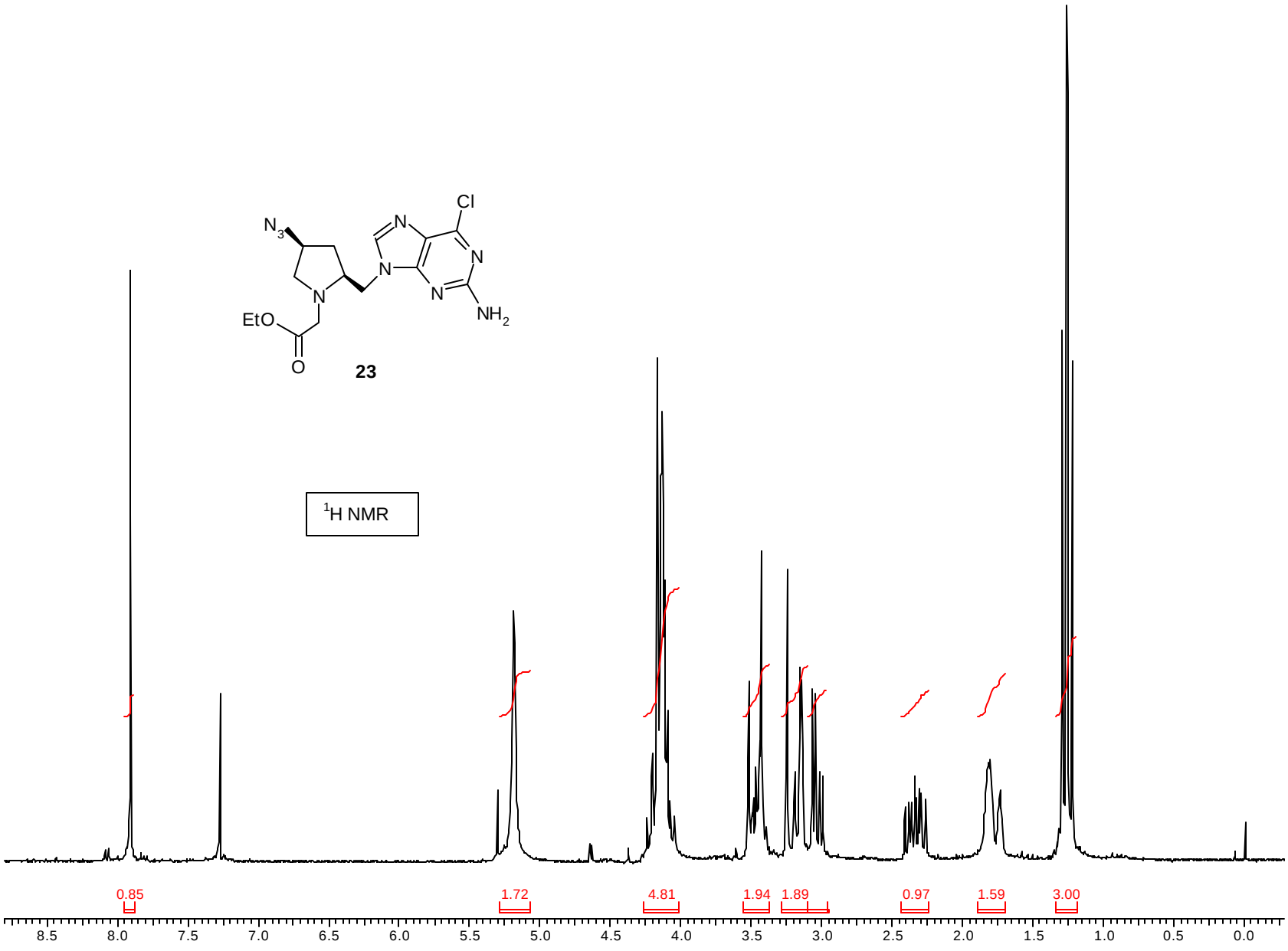
18

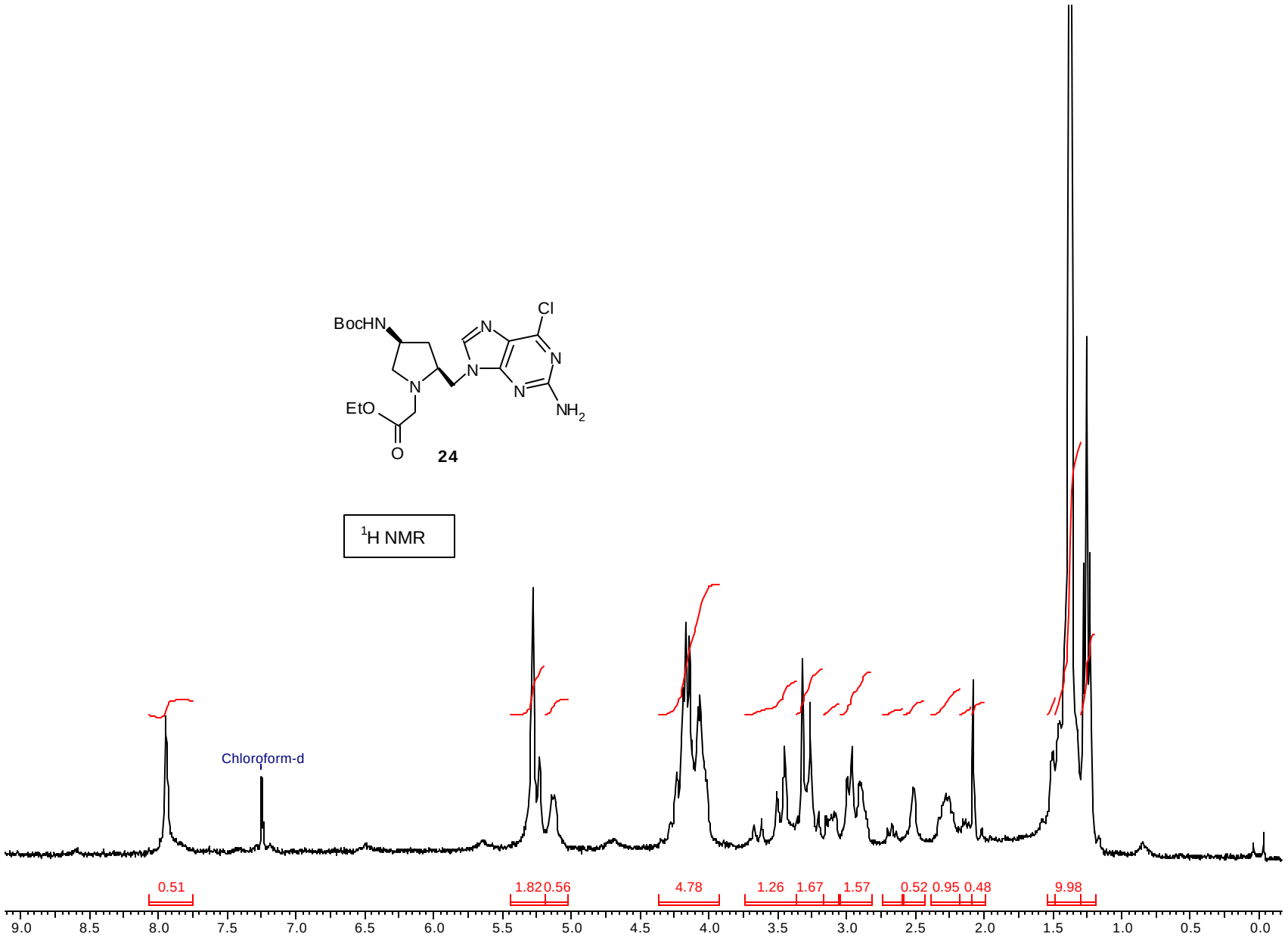
^{1}H NMR



19


N3
Cl
N
N
N
N
N
N
Boc
NH2
22
1H NMR
13C NMR


Cl
N3
N
N
N
N
N
NH2
EtO
O
23
1H NMR
0.85
1.72
4.81
1.94
1.89
0.97
1.59
3.00
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0


BocHN
Cl
N
N
N
N
N
EtO
O
NH2
24
1H NMR
Chloroform-d
0.51
1.82
0.56
4.78
1.26
1.67
1.57
0.52
0.95
0.48
9.98
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0