CHAPTER 1

Introduction
1.1 Introduction

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

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

The major limitation in using natural oligonucleotides as therapeutic agents is that they are rapidly degraded by cellular nucleases. Furthermore, the ability of the DNA strand to cross a cellular membrane is poor. As a result, a significant number of modifications of the oligonucleotide structure have been made to make them stable towards cellular enzymes.\(^8\) The various possible sites of modifications on DNA sequence are shown in Figure 1.
The phosphate modifications such as phosphorothioates (2a), phosphorodithioates (2b), methylphosphonates (2c), phosphoramidates (3) and phosphotriesters (4) form the first generation 'antisense' oligonucleotides (Figure 2) which have already shown promising results.\textsuperscript{9} Among these, the phosphorothioates are currently in phase III clinical trials.\textsuperscript{10,11} There are various modifications on the sugar as well as on the base moiety which give highly stable complexes as well as remarkable stability towards cellular degradations.

Figure 1: Position and types of oligonucleotide modifications.

2a: $R = S$, $R' = O$
2b: $R, R' = S$
2c: $R = \text{Me}$

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

![Chemical structures](image)

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

The replacement of ribose sugar by hexose had a major setback since the corresponding oligonucleotides 5 do not show any cooperative binding with natural DNA. They still show linear complexation with a complementary hexose nucleotide oligomer. The carbocyclic analogues also form linear duplexes instead of double helical structure which are structurally different. Only a few attempts to replace the ribose-
phosphate backbone have been so far successful. These include morpholino oligomers (Figure 4) wherein the monomers are linked through the carbamate linkages\textsuperscript{13} (structure 6) and show good stability towards complementary oligonucleotides. The stability of these oligomers is attributed to the neutral nature of the backbone. The replacement of the carbamate linkage with phosphoramidate 7 (Figure 4) exhibits higher stability on hybridization with complementary DNA.\textsuperscript{8}

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

Some exclusive characteristics of PNA such as strand displacement, have important biological applications in therapeutics and diagnostics (see Section 1.4).\textsuperscript{15} Since the monomeric units are achiral and the synthesis of the oligomers involves standard solid phase peptide synthesis, it is easy to obtain these oligomers in large scale. Homopyrimidine PNA sequences give highly stable complexes with DNA due to the formation of 1:2 (DNA:PNA) triplex complexes. The molecular modeling studies\textsuperscript{16,17} have
shown that PNA single strand has an ordered structure that allows the bases to be properly stacked in a helical form. The neutral nature of the backbone makes it less soluble in water and results in the formation of aggregates. This problem was overcome by introduction of a lysine amino acid at the C-terminus.\(^\text{14}\)

\[
\begin{align*}
\text{H}_2\text{NGly} & \quad \text{C} \quad \text{C} \quad \text{T} \quad \text{C} \quad \text{C} \quad \text{T} \quad \text{T} \quad \text{C} \quad \text{C} \quad \text{C} \quad \text{CONH}_2 \\
5' & \quad \text{G} \quad \text{G} \quad \text{A} \quad \text{G} \quad \text{G} \quad \text{A} \quad \text{A} \quad \text{G} \quad \text{G} \quad \text{G} \quad 3' \\
\text{H}_2\text{NGly} & \quad \text{C} \quad \text{C} \quad \text{T} \quad \text{C} \quad \text{T} \quad \text{T} \quad \text{T} \quad \text{C} \quad \text{C} \quad \text{C} \quad \text{CONH}_2 \\
\text{Triplex formation of (PNA)}_2:\text{DNA in parallel orientation}
\end{align*}
\]

\[
\begin{align*}
\text{H}_2\text{NGly} & \quad \text{C} \quad \text{C} \quad \text{T} \quad \text{C} \quad \text{C} \quad \text{T} \quad \text{T} \quad \text{C} \quad \text{C} \quad \text{C} \quad \text{CONH}_2 \\
3' & \quad \text{G} \quad \text{G} \quad \text{A} \quad \text{G} \quad \text{G} \quad \text{A} \quad \text{A} \quad \text{G} \quad \text{G} \quad \text{G} \quad 5' \\
\text{H}_2\text{NGly} & \quad \text{C} \quad \text{C} \quad \text{T} \quad \text{C} \quad \text{T} \quad \text{T} \quad \text{T} \quad \text{C} \quad \text{C} \quad \text{C} \quad \text{CONH}_2 \\
\text{Triplex formation of (PNA)}_2:\text{DNA in antiparallel orientation}
\end{align*}
\]

\[
\begin{align*}
\text{H}_2\text{NGly} & \quad \text{C} \quad \text{G} \quad \text{A} \quad \text{C} \quad \text{T} \quad \text{T} \quad \text{A} \quad \text{C} \quad \text{G} \quad \text{C} \quad \text{CONH}_2 \\
5' & \quad \text{G} \quad \text{C} \quad \text{T} \quad \text{G} \quad \text{A} \quad \text{A} \quad \text{C} \quad \text{G} \quad \text{C} \quad \text{G} \quad 3' \\
\text{PNA} : \text{DNA Duplex formation in parallel orientation}
\end{align*}
\]

\[
\begin{align*}
\text{H}_2\text{NGly} & \quad \text{C} \quad \text{G} \quad \text{A} \quad \text{C} \quad \text{T} \quad \text{T} \quad \text{A} \quad \text{C} \quad \text{G} \quad \text{C} \quad \text{CONH}_2 \\
3' & \quad \text{G} \quad \text{C} \quad \text{T} \quad \text{G} \quad \text{A} \quad \text{A} \quad \text{T} \quad \text{G} \quad \text{C} \quad \text{G} \quad 5' \\
\text{PNA} : \text{DNA Duplex formation in antiparallel orientation}
\end{align*}
\]

The PNA binding to DNA with NH\(_2\) terminus pointing towards 5' end of DNA is designated as parallel and NH\(_2\) end pointing towards 3' end is known as antiparallel. PNA is known to bind to DNA in both antiparallel and parallel orientation. In the triplex form parallel orientation is preferred over antiparallel, but in duplex, antiparallel is preferred to parallel mode. The following section discusses the structural features of PNA complexes with DNA and PNA.
1.2 Structure and hybridization properties

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

1.2.1b Bis-PNAs: Since the PNAs are made of neutral backbone their solubility and hybridization are heavily dependent on the salt concentration. The rate of hybridization is very low at physiological salt concentration although they form stable complexes. In order to improve this, two PNA strands which form components of triple helical structure with
DNA were linked together to reduce entropy and convert the binding into a bimolecular process. Bis-PNAs synthesized using a spacer linker like triethylene glycol or a lysine analogue\textsuperscript{25} do form highly stable complexes with the kinetics of complexation increasing under physiological salt concentrations.

![Bis-PNA diagram]

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

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

![Hydrogen-bonding patterns]

\textbf{Figure 6}
Bis-PNA linked through a short peptide chain was synthesized and the crystal structure of this bis-PNA:DNA triplex is solved. The studies on these systems revealed the occurrence of different structural motif which is different from the known stable helical forms that nucleic acids normally adopt. The structure of the triplex was a hybrid of B and A forms of DNA and designated as P-form. The P-form helix has a base tilt similar to B-form DNA, where the paired bases are displaced from the helix axis even more than in A form DNA. The structure of this triplex arises from 2:1 complexation of PNA with DNA. Interestingly, the phosphates of the DNA backbone formed hydrogen bonds to amide protons present in the PNA backbone of the Hoogsteen strand. In a recent report, Wittung et al. have demonstrated that a cytosine rich homopyrimidine PNA adds to double stranded polynucleotide targets through Hoogsteen strand forming PNA(py):DNA(pu):DNA(py) motif of triplex. This development can be exploited to form efficient antigenic agents.

1.2.2 PNA:DNA and PNA:RNA duplexes

PNA sequences made of all four base pairs are shown to form stable duplexes with DNA. Although the antiparallel orientation is the preferred mode of binding, they do form stable duplexes in parallel orientations also. The CD spectra of these complexes show base pair complementarity and are similar to the right handed helical B-DNA. The detailed structural information about the helical arrangement of such PNA:DNA hybrids have been revealed by NMR solution structural studies using 2-D NOESY and 2-D COSY experiments. A hexameric PNA, GAACCTC formed a 1:1 complex with complementary antiparallel RNA with Watson-Crick-CRICK base pairing similar to the A-form of RNA duplexes. The results show that no intramolecular hydrogen bonding is present on the PNA backbone as proposed by Bruice et al. The A-form conformation places the PNA tertiary amide carbonyl group in a position isoteric to the RNA C2', facilitating maximum exposure of the backbone carbonyl to water. Similarly the solution structure of PNA:DNA duplex with antiparallel orientation was solved by NMR studies. The PNA:DNA complex adopts a B-like structure with 2'-deoxysugar in the C2'-endo conformation. The observed base pairing was more like the Watson-Crick base pairing in
DNA. A mixed sequence PNA containing homopurines was also shown to bind dsDNA by initial formation of duplex with strand displacement, that was detected by photofootprinting experiments. Furthermore, the alternating TG-PNA sequences were shown to bind DNA targets and form a strand displacement complex containing PNA:DNA duplex. The strand invasion binding of PNA is a highly sequence dependent process and more studies are required to elucidate a general criteria for strand invasion duplex formation.

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

The basic molecular recognition criterion in DNA and its analogues is the base pairing pattern. PNA forms stable duplex and triplex with DNA and interestingly similar isomorphous structures can be formed with complementary PNA. Nielsen et al. have shown that such PNA:PNA duplexes are formed with high stability (more stable than PNA:DNA duplex) and these exhibit helical structures similar to the DNA double helix. The binding of these complementary PNA also follow the same rule as in DNA where the antiparallel mode of binding shows much higher stability compared to the parallel mode. The formation of these helical structures, as studied by the change in absorbance as a function of time, is instantaneous. The presence of a C-terminal lysine residue induces chirality to the helical structure and this was observed by CD wherein L and D lysine at the carboxy terminus give exactly opposite handedness to the helical PNA. The evolution of CD signal of the helix formation as a function of time is much slower compared to the kinetics observed by UV-absorption as a function of time. This was explained to be a result of the slow reorientation of the instantly formed duplex due to the induction of chirality by the terminal lysine. The PNA duplex with no chiral moiety failed to show CD band because of the presence of a racemic mixture of both right and left handed helices. The presence of a chiral lysine at the terminus does not affect the binding properties of the PNA with DNA. The (PNA)₃ triplexes [(PNA-T₁₀)₂:PNA-A₁₀] are also formed and found to be more stable than the PNA₂:DNA triplexes. The introduction of chiral cyclohexane within the backbone reduces the affinity but no preferential binding between the two enantiomers was observed (see later section 1.3.1).
1.2.4 Strand Displacement

The most interesting property of PNA is the formation of PNA₂:DNA triplexes containing homopurines through strand displacement.¹⁴ The high stability of the complexes does not favor dissociation. Since PNAs can bind in either orientation, the formation of the triplex takes place readily. These triplexes are so stable that the homopyrimidine/homopurines bind to DNA duplex by formation of a D-loop (Figure 7) which was shown unambiguously by photofootprinting experiments.¹⁴ The unpaired DNA strand present in the D-loop is highly susceptible to attack by S1 nuclease and oxidation with permanganate.³⁸ The strand displacement has also been visualized by electron microscopy and furthermore, it was shown that the binding of PNAs to closed-circular DNA results in unwinding of the double helix by approximately one turn per ten bases. The formation of D-loop by strand displacement is highly salt dependent, being efficient at low salt concentration (less than 50 mM).³⁹ However, once the strand displacement complex is formed, the salt concentration can be raised to more than 0.5 M NaCl without disrupting the complex. The low salt concentration might have a role in destabilizing the dsDNA and initiate PNA binding. A very slow dissociation of PNA from dsDNA was monitored by adding excess of an oligonucleotide complementary to PNA strand in a strand displacement complex.

Figure 7: Formation of D-loop by PNA strand invasion.
The design of bis-PNA where the two PNA strands are clamped by a positively charged lysine-aminohexyl linker lead to considerably faster and more efficient binding to the duplex.\textsuperscript{25} These strand displacement complexes were formed at relatively higher salt concentration (100 mM NaCl). The cationic charge on the linker increases the solubility and its interaction with the negatively charged DNA backbone leads to higher stability and faster formation of the duplexes.

1.3 PNA Modifications and PNA Conjugates

1.3.1 PNA backbone modifications

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

![Figure 8](image.png)
The oligomers prepared using these extended backbone units did not form any stable complexes with complementary duplexes. Incorporation of a single monomeric unit into PNA oligomer also showed a large destabilization though the selectivity was maintained. From this it is evident that the aminoethylglycine backbone is the most optimum structure. The methylene group insertion showed more destabilization when present in the acetyl linker part as in structure 12.

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

The above observation that certain structural modifications on the PNA backbone do not affect the binding properties, led to the introduction of chirality into the backbone to improve the orientational preferences. The introduction of alanine in the backbone was envisaged to control the direction of binding. Such a replacement with L-alanine 16 led to a slight destabilization whereas D-alanine 16 showed marginal stability. In both cases antiparallel orientation was preferred over parallel mode. Though introduction of
methyl group in the backbone does not alter the hybridization properties, bulky replacements like isoleucine 19 were shown to destabilize the duplex formation. A further study where glycine was replaced with L or D lysine as in 18 also showed stabilization of duplexes with D-lysine incorporation, indicating that D-amino acids are perhaps better tolerated than the L-counterparts. Though the introduction of lysine increased solubility, the electrostatic interaction did not show much contribution in duplex stability. The introduction of acidic amino acid as in 20 did show destabilization. The incorporation of a single backbone modified PNA monomer having the nucleobase thymine and a chiral cyclohexyl moiety in its backbone into PNA-T<sub>10</sub> showed preferential selectivity in forming a triplex with the purine strand of PNA-A<sub>10</sub> as compared to the corresponding DNA-A<sub>10</sub> strand.

1.3.2 PNA with nucleobase attached to α-C of glycine

The replacement of the sugar-phosphate backbone by a peptide carrying nucleobase was envisaged in the early 1970s. The developments in this direction were motivated by isolation of naturally occurring peptides containing a uracil base known as willuridine (22). There are quite a few reports on the synthesis of polythymine peptides in
the 70's but unfortunately, systematic binding studies with DNA were not done at that time. After many efforts finally Nielsen et. al.\textsuperscript{16} invented an optimum replacement for sugar-phosphate backbone with peptidic analogue. In this section, the efforts to synthesize PNA analogues using amino acid units carrying the nucleobase on the α-position are discussed.

![Chemical structures](image)

Figure 11

One of the interesting modifications is the preparation of Boc(serCH\textsubscript{2}-T)-OMe (21), which contains the O-C-N linkage to the nucleobase.\textsuperscript{48} This amino acid monomer coupled alternatively with glycine gave tetrapeptide with two thymine units. Taddie et. al.\textsuperscript{49} have used homoalanyl unit 24 and prepared both A and T monomers which were assembled alternatively with glycine to obtain a homochiral oligomer that forms a self-complementary duplex. They showed that a 10-mer duplex had much lower \( T_m \) compared to PNA:PNA duplexes. The work was extended by Shah et al.\textsuperscript{50} where they have prepared homopyrimidine oligomers similar to those reported by Taddie et al.\textsuperscript{49} but with alternating \( R \) and \( S \) configuration using 24. These heterochiral PNA oligomers do form stable duplexes with DNA.

Diederichsen et al.\textsuperscript{51} have used a thymine analogue of willuridine 23 which is also an alanyl thymine amino acid to prepare homooligomers. These have alternating \( R \) and \( S \) configuration but do not contain a glycine unit as in the earlier cases. They form linear self-complementary complexes with alternating A and T units. An important observation is
that the geometry is favorable for the formation of non-WC type A-A complexes which are more stable than the corresponding WC type A-T complexes. Further, to improve upon the A-T binding ability, an oligomer consisting of two alternating structures were prepared using 23 and 24, with a heterostereochemistry. These were shown to form stable self-complementary complexes with A-T type of base pairing.

![Chemical Structures](image)

**Figure 12**

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

1.3.3 PNA- Analogue

1.3.3a PNA: DNA Chimeras: It is a well established fact that PNAs bind to complementary DNA with high efficiency and specificity and are stable towards cellular enzymes (protease and nuclease). A combination of PNA and DNA may reduce the self-aggregation of the PNA and further increase its bioavailability. Such chimeras might also provide RNase H activation upon complexation with mRNA in antisense applications or even function as a primer for DNA polymerase for use in PCR. Nielsen et. al. have used linker 28 (Figure 13) to conjugate DNA with PNA. However, this type of head to tail orientation (5' DNA 3'-N-PNA-OH) did not show any significant thermal stability. The duplexes formed by such chimeras showed very low stability compared to DNA:DNA duplexes.

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

![28 and 29](image)

Figure 13
Uhlmann et al.\textsuperscript{64} have used a combination of both the linker units 28 and 29 mentioned earlier and synthesized PNA:DNA chimeras. They have demonstrated that these chimeras have preferential binding in antiparallel orientation compared to parallel orientation. The utility of these PNA as antisense agents is still not clear because of their poor cellular uptake, while PNA:DNA chimeras can cross the cell membrane with equal efficiency as an oligonucleotide and are resistant to 3'-exonucleases activity, due to the presence of PNA at the 3'end.

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

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

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

The primary goal of design and synthesis of PNA is for their potential biological applications, with the advantage being stability and efficient antiguene action by binding to dsDNA. But the actual oligomer showed more promising properties by strand displacement binding to the dsDNA.\textsuperscript{16} Since the PNA:DNA binding affinity is very high, a lower concentration of PNA is required to get the desired effect. They bind to supercoiled DNA with a higher affinity resulting in uncoiling.\textsuperscript{67} The exceptional high stability of PNA\textsubscript{2}-DNA triplexes enable strand displacement to take place upon targeting of double stranded DNA. The homopyrimidine PNA oligomers displace the pyrimidine strand of complementary dsDNA targets upon formation of PNA\textsubscript{2}-DNA triplex with homopurine strand which is shown to be salt dependent (<50mM) and has slight preference for parallel binding. Similar strand displacement complexes are reported to be formed by mixed sequence of PNA with dsDNA resulting in to the PNA-DNA duplex and displacement of other strand of target DNA.\textsuperscript{37}

1.4.1 Stability towards cellular enzymes

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

1.4.2 Transcription arrest by PNA

The binding of PNA to dsDNA is highly sequence specific and once bound, they form a strand displacement complex. It was found that PNA-T\textsubscript{10}-lys-NH\textsubscript{2} bound downstream from the promoter region caused transcription elongation arrest at the site
only where the PNA was bound to the template strand in presence of RNA polymerase T₃. There was no effect when the PNA was bound to the non-template strand. This selectivity indicated that independent targeting of mRNA with antisense sequence or dsDNA with sense sequence is possible with PNA in vivo.⁵⁹ It was also shown that the primer extension of Taq polymerase is arrested by PNA-T₁₀ binding on the template DNA. The transcription arrest was significantly low in case of PNA containing one mismatch and virtually no transcription arrest occurred with PNA containing two mismatches. The best effect was obtained by using PNA-T₁₀ and the arrest was less efficient upon the introduction of C monomers as in PNA-T₃CT₄ or PNA- T₃CT₂CT₄.⁷⁰ This was attributed to a weak binding of C monomer in the triplex form at physiological pH due to its non-protonation. PNA-T₈ and T₆ show lower efficiency of transcription arrest since the binding is weak. From the arrested transcripts it was observed that the polymerase transcribed two or three nucleotides into the PNA binding site. These results show that efficient transcription elongation arrest can be achieved by PNA targeting of the template DNA strand. In most of these experiments the PNA:DNA strand displacement complex were formed in low salt concentration and then transferred to the transcription buffer.

Similar transcription elongation arrest with eukaryote RNA polymerase enzyme was also observed. Again stable complexes of PNA with RNA-A₁₀ sequence (PNA₂:RNA) was shown to cause inhibition of reverse transcriptase (RNaseH, MMLV).⁷⁰ This type of inhibition is not exhibited by DNA forming triplex with RNA. While the binding of PNA to dsDNA is highly salt dependent, its hybridization with ssDNA and RNA is not significantly affected by salt concentration.

1.4.3 PNAs as transcription promoters

The strand displacement binding of PNA has unlimited scope. The D-loop formed by the binding of PNA to dsDNA are very much like transcription initiation complexes. Transcription initiation complexes are formed by helicases by unwinding of the helix with polymerase and transcription factors. The D-loop formed by PNA-T₁₀ binding is around 12 base pairs which is similar to the natural complexes (Figure 14). It is shown that RNA polymerase of E. coli recognizes the D-loop as the template and transcribes the DNA
strand. The efficiency of the transcription is increased by having more than one PNA binding to the DNA strand adjacent to each other. These results also imply that the PNA can be used as a general sequence specific gene activator (i.e., as synthetic transcription factor). It is not certain under what conditions the translation arrest or transcription is preferred when the PNA binds to dsDNA. This can be successfully used by manipulating the length of sequence, where longer PNA sequences are required for transcription promoters.

![Diagram of D-loop mechanism](image)

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

1.4.4 Translation arrest by PNA

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

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

1.4.5 PNA directed PCR clamping

The higher stability and sequence specificity of PNA-DNA complexes compared with DNA-DNA complexes has been exploited in expanding the array of PCR based applications. PNA can be directed at any site of template target DNA for different purpose. Competition between the primer and PNA for binding site can inhibit or even abolish PCR amplification. The careful design of PNA and primer sequences along with optimization of PCR conditions are shown to identify single base mutations. The targeting of PNA sequence adjacent or away to primer site blocks the extension by DNA polymerase leading to inhibition of PCR amplification. The most efficient and discriminative clamping was observed with homopyrimidine sequences due to the extraordinary stable PNA-DNA triplex formation causing inhibition of polymerase at any PNA binding site. The extension of PNA primers by polymerase was studied recently by Lutz et al. The DNA polymerases from phage T4 and T7, Thermus aquaticus, Thermus flavus, Thermus ubiquitos, Pyrococcus furiosus and DeepVent polymerase as well as HIV reverse transcriptase do not recognize the PNA primers and no primer extension products were observed. However, DNA polymerase I (E. coli) and Vent DNA polymerase
(Thermococcus litoralis) did produce primer extended products using PNA (Figure 15) as the primer. At the same time PNA:DNA chimeras containing 50% PNA were not recognized by the enzymes as they were blocked at such hybrid structures. The recognition of PNA by polymerase is 100 times lower compared to its recognition of DNA primers, nevertheless primer extensions do take place.

![Figure 15: Primers used in PCR reaction recognised by polymerase.](image)

1.4.6 Isolation of specific active genes by PNA strand invasion

The use of biotinylated PNA has been employed in the isolation of specific active genes by strand invasion. The chromatin fragments of interest containing CAG triplets\(^7\) were challenged with complementary PNA at low ionic strength and PNA-DNA hybrids were isolated by density gradient centrifugation followed by capture on streptavidin coated magnetic particles. The precise mechanism of this molecular binding is unclear as under the conditions employed PNA is not expected to engage in stable Watson-Crick-Hoogsteen type triplex formation, and no control experiments were performed.

1.4.7 PNA as an artificial restriction enzyme

The strand displacement complex of PNA provides an opportunity to use these complexes and cut the DNA at a desired position. PNA in combination with nuclease S1 have been used as artificial restriction enzymes.\(^7\) Once the strand displacement complex is formed, the displaced strand of DNA is susceptible to nuclease activity and is easily
digested by S1 nuclease into well defined products. Depending on the strand to be cleaved, the PNA can be selectively directed to the other strand. Tethering of artificial nuclease can also give similar effect as demonstrated by attachment of gly-gly-his to PNA. In the presence of Ni, this PNA with gly-gly-his, bin cleaves the DNA duplex part close to its position.

![DNA sequences and PNA structure](image)

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

### 1.4.8 PNA as inhibitor of human telomerase

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

1.4.9 PNA as primordial genetic material

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

1.5 DNA T-T photodimers

The deleterious effects of ultraviolet (UV) radiation on the genetic material of living cells are well known.\textsuperscript{83} The most susceptible sites in DNA to UV radiation, are the adjacentely stacked pyrimidine bases.\textsuperscript{84} Any mutation in DNA due to photodamage, if not repaired correctly, would lead to cell death or cancer. The skin cancer is known to be caused by damage to DNA by sunlight.\textsuperscript{85} Patients suffering from the genetic disorder, \textit{Xeroderma pigmentosum}, characterized by a defective repair mechanism of DNA photodamage are reported to be most sensitive to sunlight.\textsuperscript{83} Presently, the research in the
field of photochemistry of nucleic acid has acquired special status due to a fast depleting ozone layer and the consequent mutagenic and carcinogenic effects of solar UV radiation increasingly felt by the biosphere. The most significant photoproducts of DNA, both in terms of easy formation and adverse effects, are the cyclobutane dimers of thymine. The possible cyclobutane photodimers from DNA based on their stereochemistry can be classified into cis-syn and trans-syn isomers. The former is an irradiation product of either duplex or single stranded DNA (ssDNA) whereas the latter is formed only from ssDNA or at the junctions of DNA chains with different conformations. Organisms have developed various repair mechanisms to combat DNA damage. In a most ingenious method, the energy of the damaging radiation, in near UV region, is being utilized by DNA photolyase enzyme system to monomerize the pyrimidine photodimers. Photolyases were often found to have comparatively higher or even exclusive specificity towards cis-syn rather than trans-syn thymine dimers. As a result, studies have been oriented heavily on understanding of the mechanism of comparatively easily available cis-syn thymine dimers of natural dinucleotides or model compounds. Such studies are not many for trans-syn thymine photodimers. Recently, it has been demonstrated that E. coli DNA photolyase does repair trans-syn thymine dimers, but with $10^4$ fold less efficiency compared to repair of cis-syn thymine dimers. Such a subdued mechanism present in the cell for the repair of trans-syn dimers deem them extremely lethal to life.

1.6 Present Work

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

The poor cell permeability of PNA and bioavailability have been the major hurdle for the prospects of PNA as a drug. They are known to bind in both antiparallel and
parallel mode, with antiparallel mode being more stable. Above all PNAs have low solubility and strand displacements occur at low concentrations (50 mM). To overcome these drawbacks, a study of properties of PNA analogues with modified backbone is necessary and would greatly contribute to unravel the structure-activity relations of PNA with DNA.

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

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

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


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


            New York John Wiley and sons.