CHAPTER I

INTRODUCTION

1.1 Introduction to peptide nucleic acid (PNA)

The double helix of DNA is nature's the most simple and elegant solution to the problem of storing, retrieving, and transferring the genetic information of living organism. DNA has many important characteristics that allow it to perform these functions. Two of the most important properties are the specificity and reversible nature of the hydrogen bonding between complementary base pair, properties which allow the strands of the double helix to unwound and then rewound in exactly the same configuration. The field of life science realized early on the important implications of these traits. If DNA could be chemically synthesized, then the base sequences of genes could be studied and manipulated at will. Synthetic oligonucleotides are now indispensable tools for life scientists, with many applications in molecular biology, genetic diagnostics, and most likely also soon in medicine [1].

In the last few years, attempts to optimize the properties of oligonucleotides have resulted in the synthesis and analysis of a huge variety of new oligonucleotides derivatives [2] with modifications to the phosphate group, the ribose, or the nucleobase. The most radical change to the natural structure of DNA, however, was made by a Danish group of Buchardt, Nielsen, Egholm and Berg [3-6], who replaced the entire sugar-phosphate backbone by an *N*-(2-aminoethyl)-glycine-based polyamide structure (**Figure 1.1**). The astonishing discovery that these polyamide nucleic acids or "peptide nucleic acids" (PNAs) bind with higher affinity to complementary nucleic acid than their natural counterparts [7], and obey the Watson-Crick base-pairing rules resulted in the rapid establishment of a new branch of research focused on diagnostic and therapeutic applications of this highly interesting compound class [8-11]. The ability of PNAs to displace one strand of a DNA double-helix, very difficult process with natural oligonucleotides, was reported [12]. In addition the combination of PNA and DNA to form PNA/DNA chimeras results in

new structures, which, in addition to excellent binding, can also perform some typical biological functions, such as a primer function for DNA polymerases [13]. Furthermore, self-organizing structures of PNAs are of interest in material sciences, as well as in the study of evolution of potential alternative life-forms [13-15].

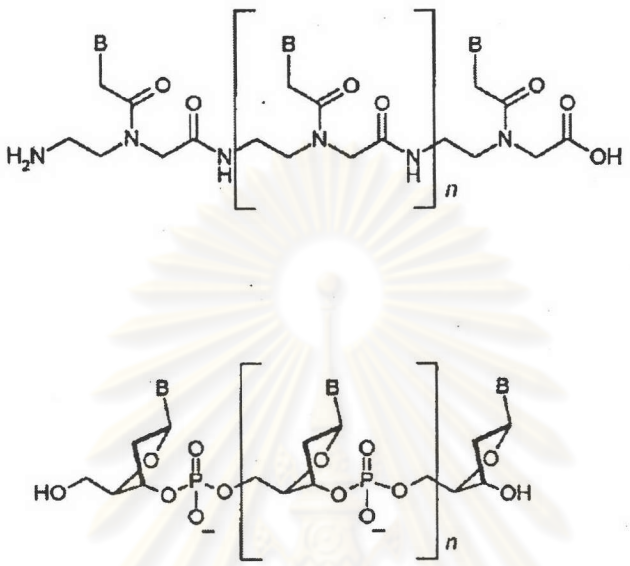

Figure 1.1 Structures of PNA and DNA. B=nucleobase

1.2 Structure and properties of peptide nucleic acid (PNA)

The structure of PNA is remarkably simple (**Figure 1.1**), consisting of repeating *N*-(2-aminoethyl)-glycine units linked by amide bonds. The purine (Adenine : A, Guanine : G) and pyrimidine (Cytosine : C, Thymine : T) bases are attached to the backbone by methylenecarbonyl linkages. Unlike DNA or DNA analogs, PNAs do not contain any (pentose) sugar moieties or phosphate groups. By conventions, PNAs sequences are written like peptides, with the N-terminus at the left-hand position and the C-terminus at the right **(Figure 1.2)** [16]. Besides the obvious structural difference, PNA is set apart from DNA in that the backbone of PNA is acyclic, achiral and neutral. PNAs can bind to complementary nucleic acids in both antiparallel (C-terminus of PNA facing 5′-OH of DNA) and parallel orientation (N-terminus of PNA facing 5′-OH of DNA) [17]. **(Figure 1.3)**

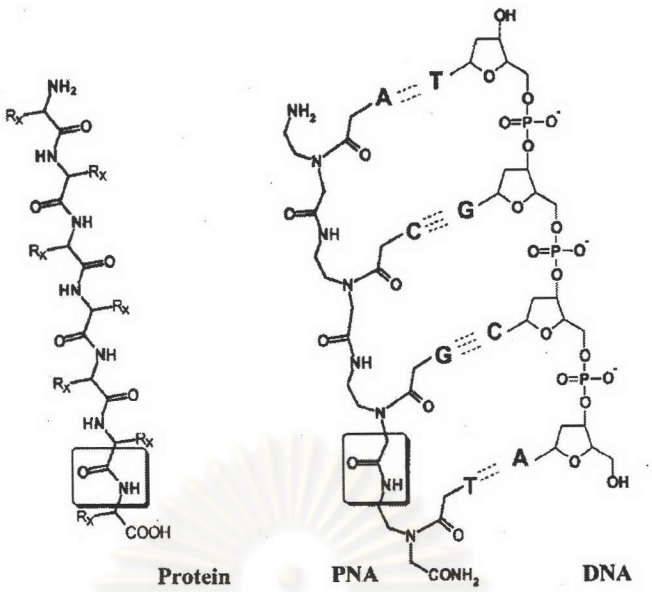

Figure 1.2 Chemical structures of a protein (peptide) (where Rx is an amino acid side chain), a PNA molecule, and a DNA molecule. The amide (peptide) bond characteristic for both protein and PNA is boxed in.

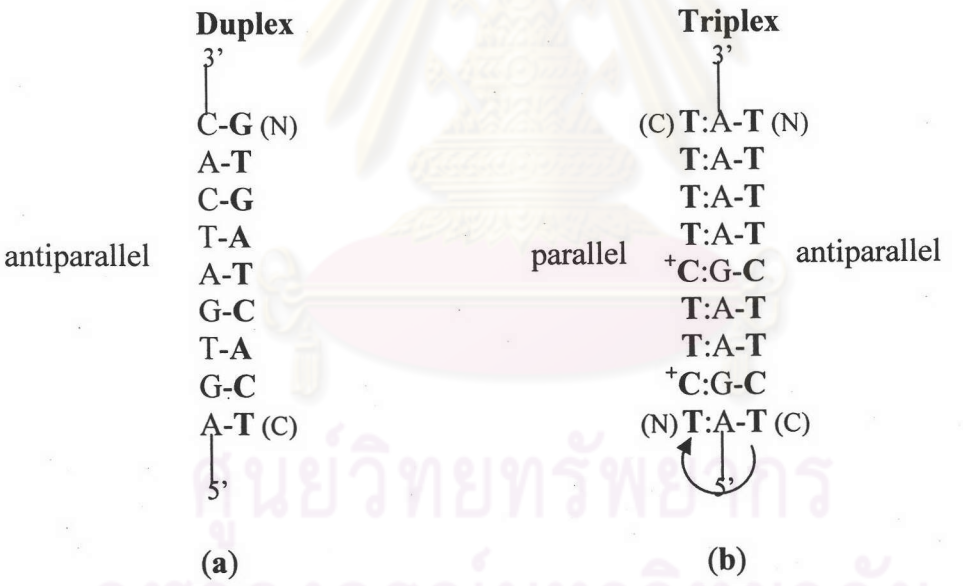

Figure 1.3 Schematic representations: **(a)** PNA-DNA duplex in the antiparallel mode (3′-end of the DNA facing the amino-terminal of the PNA); **(b)** $(PNA)_2$·DNA triplex in the preferred binding mode with antiparallel Watson-Crick strand and parallel Hoogsteen strand.

The neutral character of the PNA backbone resulted in the stronger binding between complementary PNA·DNA strands than between complementary DNA·DNA stands at low to medium ionic strength. This is attributed to the lack of charge repulsion between the PNA strand and the DNA strand. Interestingly, not only is the affinity higher for PNA·DNA duplexes, but the specificity was also found to be higher. The Watson-Crick base pairing rules are strictly observed in hybrids of PNA and nucleic acid. Furthermore, a mismatch in a PNA·DNA duplex is generally more destabilizing than a mismatch in a DNA·DNA duplex [18].

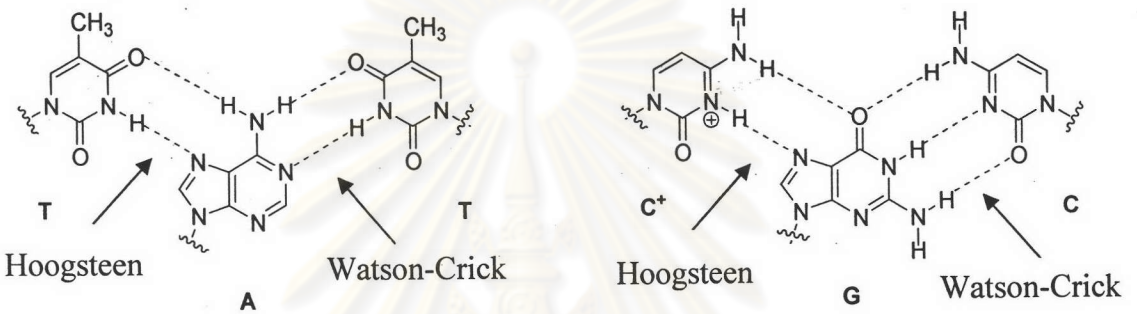

Figure 1.4 Hydrogen bonding via Watson-Crick and Hoogsteen base pairing

The possible mechanism of complex formation was proposed to involve: i) a transient partial opening of the PNA binding site on double strand DNA and incorporation of one PNA molecule with formation of an intermediate PNA·DNA duplex (DNA breathing) *via* formation of a transient PNA·DNA Watson-Crick duplex, which is subsequently trapped by a second PNA [19-21], forming Hoogsteen base pairing (**Figure 1.4**) and ii) formation of a vary stable $(PNA)_2$·DNA triplex [22]. Duplex forming between homopurine PNA to a complementary homopyrimidine target in double stand DNA was found to proceed through the duplex invasion to form PNA·$(DNA)_2$ (**Figure 1.5 (c)**). Preliminary studies revealed that this complex could be in either of parallel and antiparpllel mode, but antiparallel binding orientation of two purine strands gave a higher T_m value [23]. Cytosine rich homopyrimidine PNAs bond to guanine strand of double stand DNA and become triplex depending on N^3-protonation and Hoogsteen base pairing without strand displacement (**Figure 1.5 (a)**) [19]. Modes of Complexation of PNA to DNA could be summarized in **Figure 1.5**

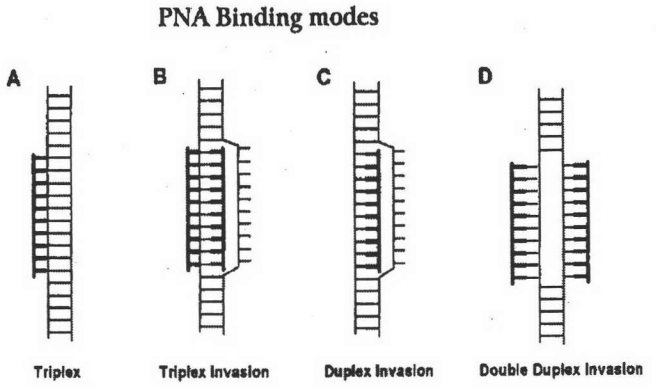

Figure 1.5 schematic drawing of complexes formed upon targeting double strand DNA with PNAs (bold): (**a**) a conventional triplex seems to be formed only with cytosine rich homopyrimidine PNA (**b**) triplex invasion complex is the most stable and is for P-loop; (**c**) duplex invasion complex can be formed by homopurine PNA; (**d**) double duplex invasion complex requires non-natural nucleobase in the PNAs strand;

The preference of $(PNA)_2$·DNA triplex helix formation has been supported by molecular mechanic calculation, comprising of intra- and inter- molecular hydrogen bonding as stabilizing forces [24,25]. Although formation of intramolecular hydrogen bonding (between NH-glycyl amide proton and acetyl linker carbonyl) would seem to be reasonable, only intermolecular hydrogen bonding between each amide NH and phosphate oxygen of backbone was observed in practice as confirmed by X-ray crystal structure [26], and NMR spectroscopy [27].

The three dimensional structure of at least four PNA complexes have been determined (**Figure 1.6**). A hexamer PNA·RNA duplex [28] and an octamer PNA·DNA duplex [29] were solved by NMR methods, while an undecamer $(PNA)_2$·DNA triple [30] and a hexamer PNA·PNA duplex [31] were solved by X-ray crystallography. It can be concluded from these structures that PNA is able to adapt well to its nucleic acid partner, as the RNA strand in the PNA·RNA duplex is essentially in an A-form conformation, while the DNA strand in the PNA·DNA duplex adopts a B-like conformation, It is, however, equally clear that PNA prefers to adopt a unique helical conformation (named the P-from [30,31]) that is distinctly different from other nucleic acid helices. This conformation is dominating in the $(PNA)_2$·DNA triplex [30] and is clearly seen in the PNA·PNA duplex which is a vary

wide helix (28 Å) with almost twice the pitch (18 base pairs turn) of an A- or B- form helix (10-11 base pairs per turn).

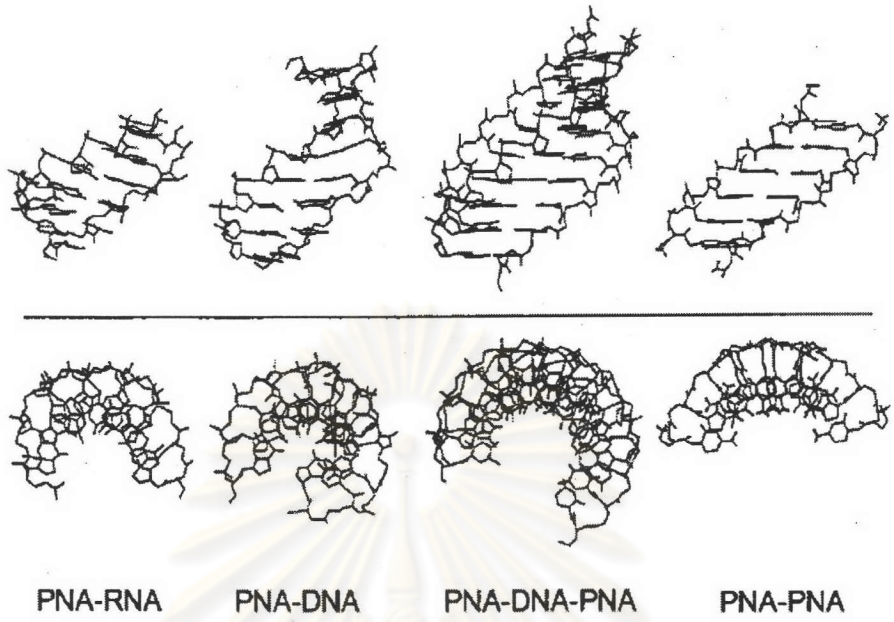

Figure 1.6 Three dimensional structures of various PNA complexes.

1.3 Applications of peptide nucleic acid (PNA)

Peptide Nucleic Acid (PNA) can mimic the most important feature of DNA, that is it can show the behavior of DNA and bind complementary nucleic acid strands. The neutral backbone of PNA results in stronger binding and greater specificity than would normally achieved with DNA. In addition, the unique chemical, physical and biological properties of PNA have been exploited to produce powerful biomolecular tools, antisense and antigene agent, molecular probes and biosensors.

1.3.1 Antigene and antisense application of peptide nucleic acid (PNA)

Peptide nucleic acids are promise candidates for development of gene therapeutic drugs. In principle, two general strategies can be adapted to design these drugs (**Figure 1.7**). Oligonucleotides or their analogs are designed to recognize and hybridize to the complementary sequences in a particular gene of interest whereby they should interfere with the transcription of the particular gene (antigene strategy). Alternatively, nucleic acid analogs can be designed to recognize and hybridize to complementary sequences in mRNA and thereby inhibit its translation (antisense

strategy). PNAs are chemically and biologically stable molecules and have significant effects on replication, transcription, and translation processes, as revealed from *in vitro* experiments. As an example, PNA-dsDNA strand displacement complexes can inhibit protein binding and block RNA polymerase elongation [32].

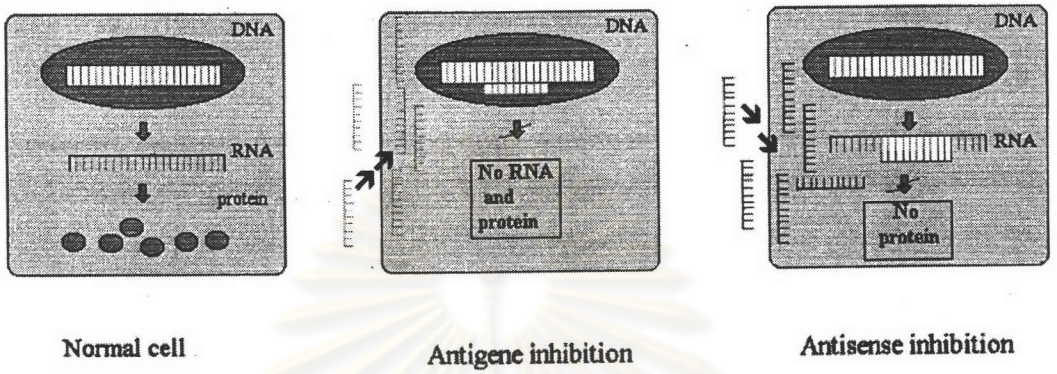

Figure 1.7 Antigene and antisense strategy is illustrated. An antigene oligomer (e.g. PNA) could bind to a complementary sequence in the DNA and inhibit transcription of the gene. On the other hand, cells can also be treated with an antisense oligonucleotide, and hybridization to a specific mRNA sequence can inhibit the expression of a protein at the level of translation.

1.3.2 Diagnostic application of peptide nucleic acid (PNA)

The high-affinity to nucleic acids of PNA oligomers has led to the development of the most important applications of PNA, namely as a diagnostic probe for detecting genetic mutation. These mismatch analyses employed PNA's unique hybridization properties. The screening for genetic mutations by capillary electrophoresis highlights some of the recent developments related to the use of PNA as a probe to detect genetic mutations and corresponding mismatch analysis [32].

1.3.3 PNAs as tools in biotechnology

The DNA biosensor techonology holds promise for rapid and cost-effective detection of specific DNA sequences. A single-stranded nucleic acid probe is immobilized onto optical, electrochemical, or mass sensitive transducers to detect the complementary (or mismatched) strand in a sample solution. The response from the

hybridization event is converted into a useful electrical signal by the transducer. PNA has been used as a probe for sequence-specific biosensors and highlight some of the promise it holds to work as the recognition layer in DNA biosensors [32].

1.4 Structures modified of peptide nucleic acid (PNA)

The success of PNA attracted much attention from many research groups to search for alternative structures which may behave similar to PNA or even have extra features currently not present or remained to be developed in Nielsen's PNA such as solubility, ability to penetrate cell membranes and specific direction of binding.

According to thermodynamic consideration, the hybridization of PNA to complementary oligonucleotide caused the enthalpy gain and significant entropy loss due to the conformational restriction of both PNA and DNA as a result of complexation. It can be envisioned that the free energy gain of hybrid formation may be decreased by reducing the entropy loss. One way to solve this problem is by locking the conformation of PNA to be the same as in the PNA·DNA hybrid [24]. For this reason, many research groups have paid attention focusing of structure preorganization of modified PNA such as introducing a rigid backbone to form conformational constrained PNA. A PNA analogue based on 4-aminoprolyl backbone with two chiral centers at C-2 and C-4 have been synthesized (**Figure 1.8b**) [33]. Homothymine hexamers of three diastereomers (*trans*-D/L, *cis*-L) failed to form any complexes with dA_6. Conversely, including such PNAs unit either at the N-terminus or within the strand of the original PNA resulted in oligomers which are able to form complexes with complementary oligonucleotides in both parallel and antiparallel orientation with higher T_m compared to the original PNA [34]. The other interesting proline system of *cis*-L and *trans*-L isomer of both 4-aminoproline and pyrrolidine-2-carboxylic acid derivative carrying nucleobase at C-4 position were studied which could not form the stable complexes (**Figure 1.8b** and **c**) [35]. However, homothymine of these 4-6 mer PNA alternating sequence with original PNA have higher the melting temperature than original PNA about 6-7 °C when the conformation on the pyrrolidine ring was *trans*-L [36]. Modified PNA containing glycylproline as a monomeric building block with *cis*-stereochemistry has been

synthesized (**Figure 1.8c**) [37-39]. These modified chiral PNAs were shown to interact strongly with complementary DNA and RNA [40].

These modified chiral PNAs and their hybrid with oligonucleotides, however, generally has poor solubility in aqueous media, making biological studies difficult. Attempts to improve the solubility of PNAs have so far met with variable success [41,42]. It have recently reported a synthesis and binding studies of the deoxy-analogue of prolylglycyl PNA, i.e. the glycine being replaced by an aminoethyl linker [44,46,47]. Replacing the glycine carbonyl group in the prolylglycyl PNA of aminoethylprolyl (*aep*) PNA (**Figure 1.8d**) with a methylene group should create a more confomationally flexible backbone while the conformation of the side chain is still restricted. Increasing conformational flexibility of the backbone might decrease the binding affinity due to the increased entropy loss upon hybridization, but it should allow the *aep*PNA to adopt a wider range of conformations than the prolylglycyl PNA. Combination of the two factors may decrease or increase the binding affinity of the resulting *aep*PNA to its complementary oligonucleotide, depending on how close the conformation of the *aep*PNA in the hybrid is to that in its native state. In addition, the basic proline nitrogen atom should be at least partially protonated under physiological conditions and should attract the negatively charged phosphate group of DNA so providing further stabilization of the hybrid formed with natural DNA. This idea has been exploited in Nielsen's type PNA by replacing the methylene carbonyl linker which joined the base and the backbone to an ethylene linker which joined the base and the backbone to an ethylene linker [43] so that the α-nitrogen atom became basic.

At least 3 research groups have reported synthesis and hybridization study of *aep*PNA (**Figure 1.8d**) with nucleic acids.

Kumar and co-workers reported the synthesis of two of the four possible stereoisomers of *aep*PNA, namely (2*S*,4*S*) and (2*R*,4*S*). They also reported site-specific incorporation into Nielsen's PNA oligomers by solid phase synthesis and hybridization properties of these PNAs with complementary DNA. They reported that the homothymine octamer *aep*PNA with (2*S*,4*S*) and (2*R*,4*S*) stereochemistry formed a very stable $(PNA)_2$·DNA triplex with complementary DNA (T_m > 80 °C) [44,45].

On the other hand, Liu and his co-workers reported the synthesis of similar (2*S*,4*S*)- *aep*PNA oligomers via a different method [46]. In contrast to Kumar's

report, the hybridization of decamer *aep*PNA with (2*S*,4*S*) stereochemistry according to Liu showed no detectable binding with complementary DNA.

Vilaivan and his colleagues has reported, another synthesis of a new stereochemistry (2*R*,4*R*)-*aep*PNA oligomer containing thymine together with its hybridization properties of (2*R*,4*R*)-*aep*PNA oligomer [47]. They reported that the thymine decamer *aep*PNA binds to complementary RNA to form a 2:1 hybirid with high affinity (T_m = 53°C), whereas no binding to the complementary DNA was observed.

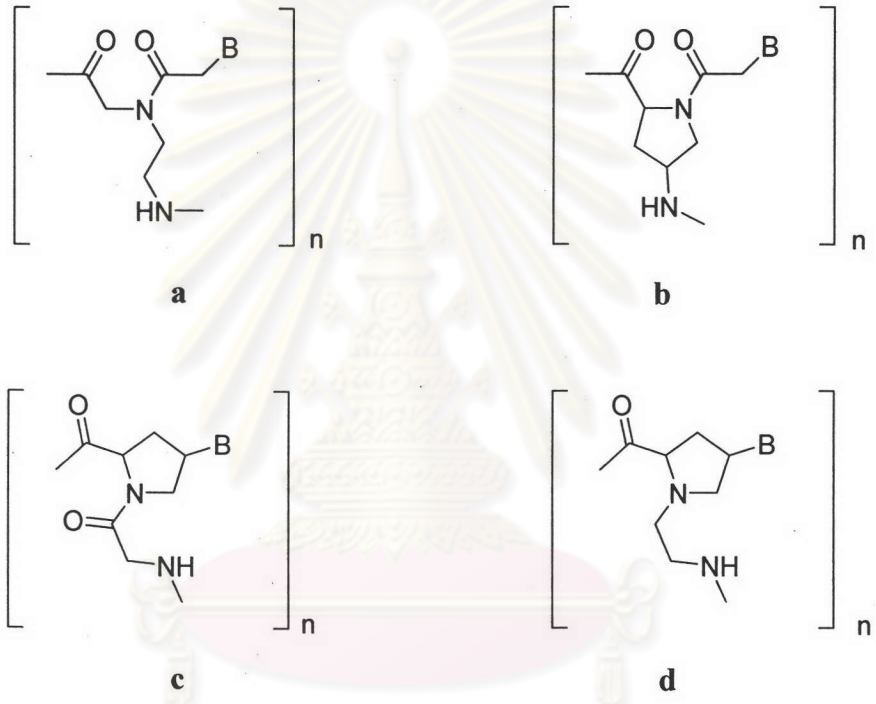

Figure 1.8 Structure of modified PNA (**a**) aminoethylglycine (*aeg*) PNA (**b**) 4-aminoprolyl PNA (**c**) prolylglycyl PNA (**d**) aminoethylprolyl (*aep*) PNA

According to the reports by the three research groups, the hybridization properties of *aep*PNA oligomers with complementary DNA showed some differences, and probably stereochemically dependent. None of these research groups has yet synthesized all four possible stereoisomers and perform a critical comparison. Moreover, the methods for syntheses of *aep*PNA previously reported, the methods were complicated and provided only moderate yield. It is therefore interesting to improve the efficiency in *aep*PNA syntheses and to study their hybridization properties.

1.5 Objectives of this research

The objective of this research is to improve the efficiency in aminoethylprolyl (*aep)*PNA monomer carrying "*cis*-D" or (2*R*,4*S*), "*cis*-L" or (2*S*,4*S*), "*trans*-D" or (2*R*,4*S*) and "*trans*-L" or (2*S*,4*R*) absolute stereochemistry. Oligomerizations of the PNA up to 10-mer will be carried out by solid phase peptide synthesis and preliminary study of binding stability these *aep*PNA oligomers with complementary oligonucleotides studied.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย