CHAPTER I

INTRODUCTION

1.1 Peptide nucleic acid (PNA)

Perhaps there is no molecule that is more important in the molecular processes in living system than the DNA double helix. All fundamental functions of life are encoded by this simple but very elegant molecule: the storage, the transfer (replication), as well as the translation (expression) of the genetic information. This is based on the precise order of only four constituents (nucleosides): deoxyadenosine (dA), deoxyguanosine (dG), thymidine (dT), and deoxycytosine (dC). Two complementary nucleotides from two strands of DNA(dA/dT, and dG/dC) form a Watson-Crick base pair by hydrogen bonding which leads to association of the two strands of DNA to from the famous double helix structure. The base pairing is further stabilized by π - π interaction and hydrophobic interaction, the so called π - π or base stacking. The negative charge phosphate groups lie on the outside of the double helix and are in direct contact with water, hiding the rather hydrophobic base pairs inside.

Figure 1.1 Nucleobase pair recognition by Watson-Crick hydrogen bonding.

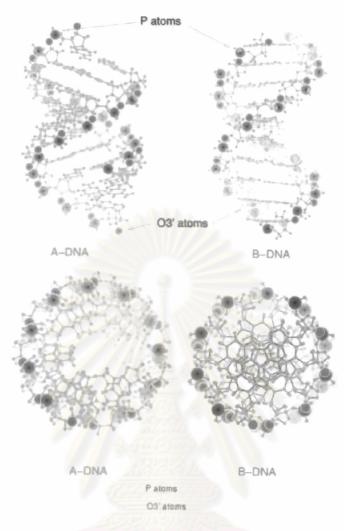


Figure 1.2 Secondary structure of DNA in A and B forms.[1]

Oligodeoxynucleotides are short strands of DNA, usually of a few-tens base. They are usually obtained by chemical synthesis. The important feature of oligonucleotides is that they can hybridize with another strand of DNA or oligonucleotide with a complementary sequence in the same way as DNA. As a result, oligonucleotides play a central role in a range of molecular biology techniques, including the use as a probe for hybridization in DNA/RNA sequence determination and as primers for Polymerase Chain Reaction (PCR). Since the binding of oligonucleotides to DNA and RNA is governed by the standard base pairing rules, the design of an agent that can directly interact with cellular DNA and RNA to interfere their normal functions, the so called antisense agent, should be straightforward. Natural oligonucleotides have been shown to exhibit both antisense and antigene properties in vitro.[2-4] However, they are rapidly degraded by nucleases in vivo.

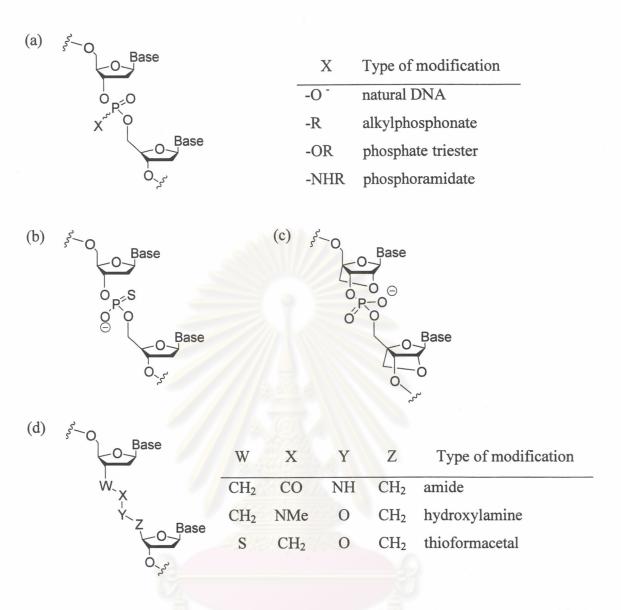


Figure 1.3 Examples of modified oligonucleotides (a) phosphate-modified oligonucleotide, (b) phosphorothioate (c) locked nucleic acid (LNA) (d) modified oligonucleotide without a phosphate

Many laboratories have appreciated this fact and developed novel oligonucleotide mimics to enhance hybridization, increase resistance to nuclease digestion and improve membrane permeability. Attempts to overcome this serious obstacle resulted in an impressive number of oligonucleotide analogues having been synthesized during the last two decades (**Figure 1.3**).[5-6] Most of the novel motifs contain relatively modest changes to either the nucleotide bases or to the phosphate backbone. However, one rather radical modification, peptide nucleic acid or poly amide nucleic acid (PNA), represents a dramatic departure from standard oligonucleotide chemistry.

Peptide nucleic acids were first described by Nielsen *et al.* in 1991.[7] A replacement of the natural deoxyribose phosphate backbone of DNA by a repeating *N*-(2-aminoethyl)-glycine (aeg) unit (**Figure 1.4**) resulted in a DNA-compatible skeleton. Attachment of nucleobases to that skeleton through methylenecarbonyl linkers afforded a DNA mimic with several remarkable properties.

Figure 1.4 Chemical structures of a DNA molecule and an aegPNA molecule

PNA binds with high affinity to its complementary nucleic acids forming a PNA·DNA hybrid that is more stable than the corresponding DNA·DNA hybrid. Unlike other DNA-binding peptides or proteins, the interaction is highly sequence specific and could be simply predicted based on the sequence of the PNA and The DNA following the Watson—Crick base pairing rules.[8]

The replacement of normal DNA skeleton has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis. Since the first report of PNA, much effort has been devoted to exploring the general DNA mimicking properties of PNAs such as their structure and their physical properties, including the use as a sequence specific ligand for binding double standed DNA.[9]

1.2 PNA-DNA Hybridization

A large amount of research effort focused on the structure and physical properties of PNA and its hybrids with nucleic acids. Physical techniques, including high-resolution NMR and X-ray crystallography have been used to analyze the major families of PNA complex. Two duplex structures, a PNA-RNA [10] and a PNA-DNA [11] duplex were solved by NMR, and the structures of a PNA-PNA duplex [12] and (PNA)₂·DNA triplex [13] were solved from X-ray crystallography. These results show that the PNA is flexible enough to able to adapt to its oligonucleotides partner since the conformation of the RNA strand in the PNA-RNA duplex is essential the A-from; while that of the DNA strand in the PNA-DNA duplex is closer to the B-from.[14]

Hybridization between DNA or RNA oligonucleotides is destabilized by electrostatic repulsion between the negatively charge phosphate backbones of the complementary strands. By contrast, PNA oligomers can bind strongly and with high sequence discrimination to complementary oligomers of DNA, RNA or another DNA and in general the thermal stability of hybrids (generally expressed as melting temperature: $T_{\rm m}$) for identical sequences follows the order: PNA-PNA > PNA-RNA > PNA-DNA.[15] Because of the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes due to the uncharged backbone of PNA, the stability and rate of PNA-DNA hybridization were improved. The *aeg*PNA can bind to its complementary nucleic acids in both antiparallel and parallel orientations (**Figure 1.5**).[16] However, the binding in antiparallel orientation has a marginally higher stability.

พูนยวทยทวพยากว จุฬาลงกรณ์มหาวิทยาลัย

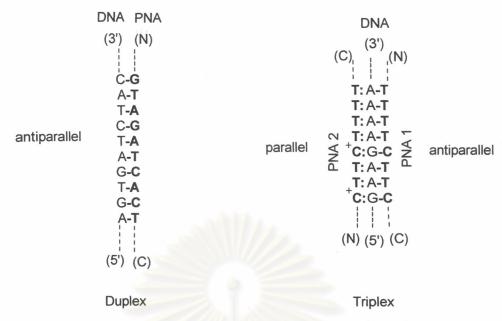


Figure 1.5 Schematic representation: (a) PNA·DNA duplex in antiparallel mode;

(b) (PNA)₂·DNA triplex in the preferred binding mode with antiparallel Watson-Crick strand and parallel Hoogsteen strand.

Figure 1.6 Hydrogen bonding via Watson-Crick and Hoogsteen base pairing.

One important feature of *aeg*PNA (1) is the tendency to form triple helices (PNA)₂·DNA especially when the PNA sequence is rich in pyrimidine bases (T, C). It comes as a great surprise that targeting of a homopyrimidine PNA to a double-stranded DNA resulted in a novel strand displacement complex which is triple helical in nature. Thymine-rich PNA binds to double stranded DNA to form a (PNA)₂·DNA complex, leaving another thymine-rich strand of DNA free as a loop structure. A preliminary study revealed that this complex could be in either of parallel and antiparallel mode, but antiparallel binding orientation of two purine stands (N-terminal PNA facing 3'-OH of DNA) gave a higher T_m value.[17] Cytosine-rich homopyrimidine PNAs bound to the guanine-rich strand of a double stranded DNA

and became triplex with N^3 -protonation and Hoogsteen base pairing without causing a strand displacement.[18]

From an X-ray crystal structure of a (PNA)₂·DNA triplex,[13] in addition to the normal Watson-Crick base-pairing, the second strand of PNA was bound by a Hoogsteen base-pairing (**Figure 1.6**). It also shows specific interactions (H-bonding) between each amide N-H of the backbone of the Hoogsteen PNA strand and the phosphate oxygen of the DNA backbone thereby also contributing to the stability. Consequently, (PNA)₂·DNA triplexes are of extremely high thermal stability.

The binding mechanism of the triple helix formation was proposed as summarized in Figure 1.7. The mechanism of the binding process involves: 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 and ii) formation of a very stable (PNA)₂·DNA triplex.[19]

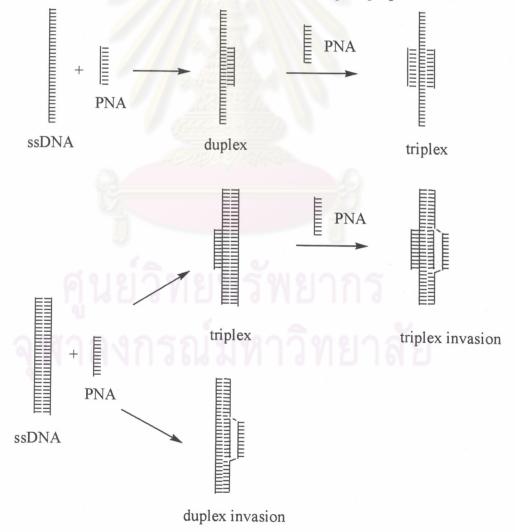


Figure 1.7 Schematic representation of modes of binding of PNA to ssDNA and ds DNA.

In contrast, mixed purine-pyrimidine PNA forms only duplex with DNA by the usual Watson-Crick base-pairing. The duplex invasion is not so readily as in the pyrimidine-rich PNA due to the high stability of DNA·DNA hybrids containing C·G base pairing, especially under a high salt condition.[20]

1.3 Application of peptide nucleic acid (PNA)

PNA is a DNA mimic that has outstanding potential for recognizing both double-stranded and single-stranded DNA. The neutral backbone of PNA results in stronger binding affinity and greater sequence specificity than would normally achieve with DNA. Moreover, the unique chemical and biophysical properties of PNA have been exploited making DNA a powerful platform to be developed as biomolecular tools, antisense and antigene agents, molecular probes and as biosensors.[21]

1.3.1 Therapeutics application of PNA

Due to its DNA and RNA binding capabilities, PNA has obvious potential in therapeutics, both for development of gene therapeutic drugs and for target validation of medicines, as well as for functional genomics studies in molecular biology. In principle, two general strategies can be adapted to design these drugs (**Figure 1.8**). In the first strategy, oligonucleoides 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 the nucleic acid analogs can be designed to recognize and hybridize to complementary sequences in mRNA and thereby inhibit its translation (antisense strategy). [20]

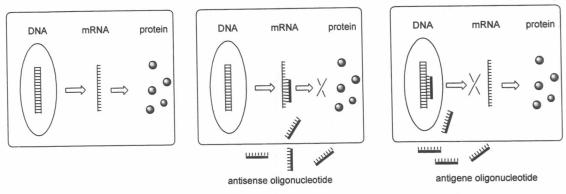


Figure 1.8 Schematic representations of the antisense inhibition and antigene inhibition. An antisene agent (e.g. PNA) binds to an mRNA containing the complementary sequence and can inhibit the expression of a protein at the level of translation. On the other hand, an antigene agent binds directly to a complementary sequence in the DNA and thereby inhibit transcription of the gene to the mRNA.

PNA·RNA duplexes are not substrates for RNAse-H, antisene inhibition of translation by PNA is therefore mechanistically different from that of phosphorothiotes. Consequently, sensitive targets identified for phosphothioate oligonuccleotides are not necessarily expected to be good targets for PNA. Indeed, sensitive RNA targets for PNA oligomers are presumably targets at which the PNA can physically interfere with mRNA function, such as ribosome recognition, scanning or assembly, whereas ribosomes involved in translation elongation appear much more robust.[22] Interestingly, it was recently demonstrated that intro-exon splice junctions are very sensitive targets for PNA antisense inhibition as correct mRNA splicing is prevented.[23] Thus in antisense experiments with PNA as with other DNA analogues and mimics, it is advisable to perform a mRNA scanning by testing a series of PNAs targeting different regions of the mRNA.

In 2005, Janowski et al. introduced PNA into the cell by complexing it with a partially complementary oligodeoxynucleotide and a cationic lipid.[24] The lipid promotes internalization of the DNA, whereas the PNA enters as cargo and is subsequently released. The antigene peptide nucleic acids (agPNAs) recognize transcription start sites within chromosomal DNA and block gene expression. These PNAs are the first synthetic molecules shown to be capable of recognizing transcription start sites inside cells.

1.3.2 PNA in diagnostics

The excellent hybridization properties of PNA oligomers combined with its unique chemistry has been exploited in a variety of genetic diagnostic techniques. For instance, fluorescene *in situ* hybridization (PNA-FISH), PNA probes for *in situ* hybridization yield superior signal to noise ratios and often allow milder washing procedures resulting in morphologically better samples. Thus PNA-FISH techniques have been developed for quantitative telomere analyses [25-26], chromosome painting [27] and viral and bacterial diagnostics both in medical as well as environmental samples.[28]

DNA biosensor technologies are thus currently under intense investigation owing to their great promise for rapid and low-effective detection of specific DNA sequences. These technologies commonly rely on the immobilization of a single-stranded (ss) DNA probe onto optical, electrochemical, or mass-sensitive transducers to recognize the complementary (or mismatch) DNA strand in a sample solution. The response from the hybridization event is converted into a detectable signal electrical signal, like fluorescent etc. PNA has been used as a probe for sequence-specific biosensor. Some of the promise it holds to work as the recognition layer in DNA biosensors has been highlighted.[29]

Human genomics and mitochondrial DNA contain large numbers of single-nucleotide polymorphism (SNPs), many of which are linked to known diseases. In a system developed for detection of SNPs, the fully matching PNA·DNA duplex protected from enzymatic digestion, whereas DNA in PNA·DNA duplex involving a mismatch are efficiently hydrolyzed by these enzyme. Therefore, one-base alteration at potential SNPs sites is visually detected by using a 3-3'-diethylthiadicarbocyanine dye which changes its color upon binding to PNA·DNA duplex. The color different is vivid in the visible-light region. The solution of fully matching PNA·DNA duplex exhibits purple color, while one mismatch between PNA and DNA shows its intrinsic blue color.[30]

A method for SNPs detection that couples the sensitivity and accuracy of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS) with the specificity and stability of PNA·DNA hybridization has also been developed.[31] Following PCR amplification, the method is applied directly to polymorphisms located within human mitochondrial DNA (mtDNA) and in genomic

DNA using human leukocyte antigen (HLA)DQα polymorphisms. The PNA probes were hybridized to single-stranded PCR-amplified DNA immobilized onto magnetic particles. Following washing, the immobilized PNA·DNA complex is then directly analyzed by MALDI-TOF mass spectrometry. The PNA·DNA hybrid denature during the MALDI process, providing the free PNA which can be detected by its mass.[32]

1.3.3 PNAs as tools in biotechnology

The polymerase chain reaction (PCR) has been widely used for various molecular genetic application including the amplification of variable number of tandem repeat (VNTR) loci for the purpose of genetic typing. Small PNA oligomers are used to block the template, and the latter becomes unavailable for intra- and interstrand interaction during reassociation. On the other hand, the primer extension is not blocked; during this extension, the polymerase displaces the PNA molecules from the template and the primer is extended toward completion of reaction.[20] PNA oligomers can be used to silent PCR amplifications in single mutation analyses.[33] This technique is so powerful that it is possible to obtain a specific signal from a single mutation oncogene in the presence of a 1,000-10,000 fold excess of the non-mutated wild type normal gene.[34-35]

1.4 Structural preorganization PNA: An approach to improve binding affinity

The success of PNA has attracted much attention from many research groups to search for alternative structures which may behave similar to PNA or may 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 this 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.[36] For

this reason, many research groups have focused on structural preorganization of PNA such as introducing a rigid backbone to form a conformationally constrained PNA.

One concept for preorganization of PNA is based on introduction of a methylene group to bridge between the aminoethylglycyl backbone and the methylene carbonyl side chain of aegPNA [(1) in Figure 1.9] to form various five- or sixmembered nitrogen heterocyclic analogues.[37] A methylene bridge between the C-5 atom of the aminoethyl part and the C-2 atom of the glycine part of the aegPNA resulted in 4-aminoprolyl PNA [(2) in Figure 1.9], with introduction of two chiral centers.[38] Homothymine hexamers of three diastereomers (trans-D/ trans-L and cis-L) failed to from any complexes with dA₆. The chiral PNAs containing monosubstitution with prolyl units either at the N-terminus or in the interior are capable of forming duplex with complementary DNA in parallel ($T_{\rm m} = 40$ °C) and antiparallel ($T_{\rm m}$ = 42 °C) modes.[39] A methylene bridge inserted between the C-2 atom of the glycine unit and the C-2' atom of the nucleobase linker of the aegPNA resulted in a novel chiral prolylgylcyl PNA [(3) in Figure 1.9]. [40-42] Another interesting PNA system derived from the cis-L and trans-L isomers of both 4-aminoproline and pyrrolidine-2-carboxylic acid derivative carring nucleobases at the C-4 position were studied.[43] However, these modified chiral PNAs and their hybrid with oligonucleotides, generally has poor solubility in aqueous media, making biological studies difficult.[44]

Attempts to improve the solubility of PNAs have so far met with variable success. [45-46] In a recent report, a synthesis and binding study of a pyrrolidinyl PNA with an aminoethyl linker has been described. [47] The glycine carbonyl group in the prolylglycyl PNA was replaced with a methylene group to form aminoethylprolyl (aep) PNA [(4) in Figure 1.9]. This PNA system should make a rather conformationally flexible backbone while the conformation of the side chain is still restricted. Though the increased entropy loss upon hybridization can result from increasing conformational flexibility of the backbone or the decreasing the binding affinity, this flexibility should allow the aepPNA to adopt a wider range of conformations than the prolylglycyl PNA. In addition, the positive charge and the tertiary nitrogen of proline should attract the negatively charged phosphate group of DNA, providing further stabilization of the hybrid formed with natural DNA.

Kumar et al. [48] reported the synthesis of two of the four possible stereoisomers of aepPNA, namely the (2S,4S)- and (2R,4S) isomers. The homothymine

octamer aepPNAs in both diastereomers were reported to form very stable (PNA)₂·DNA triplex with a complementary DNA ($T_{\rm m} > 80$ °C). On the other hand, Vilaivan and co-worker [49] reported another synthesis of a new stereoisomer of (2R,4R)-aepPNA oligomer. The thymine decamer (2R,4R)-aepPNA was found to bind with poly(rA) (the complementary RNA) to form a 2:1 hybrid with high affinity ($T_{\rm m} = 53$ °C), whereas no binding to poly(dA) (the complementary DNA) was observed

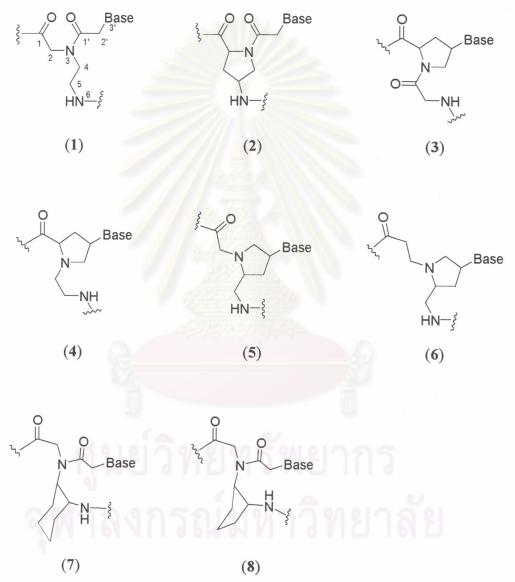


Figure 1.9 Structure of modified PNA (1) aminoethylglycine (aeg) PNA, (2) 4-aminoprolyl PNA, (3) prolylglycyl PNA, (4) aminoethylprolyl (aep) PNA, (5) pyrrolidinyl PNA, (6) backbone extened pyrrolidinyl (bep) PNA, (7) aminocyclohexylglycyl (ch) PNA and (8) aminocyclopentylglycyl (cp)PNA

Kumar and co-worker [50] has further constrained the flexibility in the aegPNA by insertion of a methylene bridge to link the C-4 atom of aminoethyl part and C-2' atom of the methylene carbonyl linker. The modified PNA [(5) in Figure 1.9] also bears a positive charge due to the presence of a tertiary sp³ nitrogen atom. The incorporation of the modified monomer with (2S,4S) stereochemistry into homothymine oligomers and mixed pyrimidine oligomers indicated a decreased binding efficiency of these oligomers with the targets DNA/RNA sequence. A backbone extened pyrrolidine PNA (bepPNA) [(6) in Figure 1.9] was later studied. The bepPNA monomer in chimeric and homooligomeric PNAs induced bining selectivity for RNA over DNA. Incorporation of the modified units at the terminals (C-/N-) exerts very weak effect on preorganized conformation and allowed binding with DNA as well as RNA.[51] A cis-(1S,2R/1R,2S) aminocyclohexanyl PNA (chPNA) [(7) in Figure 1.9], was designed so that the monomer dihedral angle matches the value (65°) found in PNA·RNA duplexes, rather than that in PNA·DNA duplexes (141°). This PNA shows an interesting preference for binding to RNA over DNA. The related cis-(1S,2R/1R,2S) aminocyclopentanyl PNA (cpPNA) [(8) in Figure 1.9] with a smaller dihedral angle value (25°) bind to both RNA and DNA with higher affinity compare to aegPNA and chPNA.[52-53]

Appella and coworkers [54] has reported a related trans-(1S,2S/1R,2R)-cyclopentane PNA (tcypPNA) [(8) in Figure 1.9]. One or more trans-(1S,2S)-cyclopentanediamine units incorporated into the aegPNA backbone significantly increase binding affinity and sequence specificity to complementary DNA. When the stereochemistry is changed to (1R,2R), there is no melting transition was observed in the UV analysis.

In most of the work reported by Kumar group and others, usually only some modified pyrrolidinyl monomers were inserted into the core structure of aegPNA at various positions. The hybridization results are not easily interpreted in these cases since the effect on T_m is not only dependent on the structure modification but also to the number and position of the modified residues in the original PNA strand. Consequently comparison and interpretation of these hetero-oligomeric PNA should be done with great care.

1.5 PNA containing β-amino acids: A novel concept of structurally preorganized PNA

In 1998, Gellman introduced the term foldamer [55-56] to describe any polymer with a strong tendency to adopt a specific secondary or tertiary structure especially β-amino acid. Conformationally rigidified β-peptide such as trans-2aminocyclohexane carboxylic acid (trans-ACHC) and trans-2-aminocyclopentane carboxylic acid (trans-ACPC) (Figure 1.10) revealed the perfect helical structure in solid state (Figure 1.11). Their study showed that trans-ACHC can form 14-helix and trans-ACPC can form 12-helix via intramolecular H-bonding between nitrogen and carbonyl group of amide bond. Thus, β-amino acid seem to have a potential to form a helix form even in short oligomer and it might help about pre-organization in thermodynamic aspect to PNA oligomer. In 2003, Vilaivan and co-worker developed a novel pyrrolidinyl PNA carrying an N-amino-N-methylglycine) [(10) in Figure 1.12] as a spacer in backbone. [57] Unfortunately, the part of β-amino acid (N-amino-N-methylglycine) in this PNA system did not increase the binding affinity between PNA and DNA because of the higher entropy loss on binding. When consider at structure, the designed PNA bearing a more flexible β -amino acid different to ACHC or ACPC which more rigid from the cyclic structure. So the cyclic β -amino acid have been studied and developed to use in PNA oligomer such as (S, S) trans-2-amino cyclopentane carboxylic by Vilaivan.[58]

Figure 1.10 Structure of monomeric and oligomeric trans-ACHC and trans-ACPC

Crystal structures of the hexamer and octamer of *trans*-ACPC confirm the predicted helical conformation (Figure 1.11).

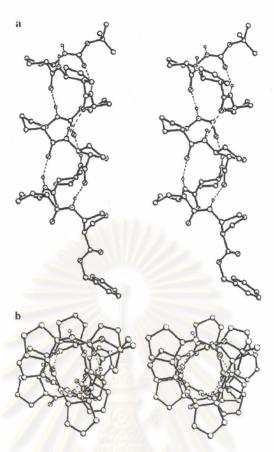


Figure 1.11 Solid-state structure of trans-ACPC octamer from two perspectives: (a) perpendicular to the helix axis and (b) along the helix axis.[56]

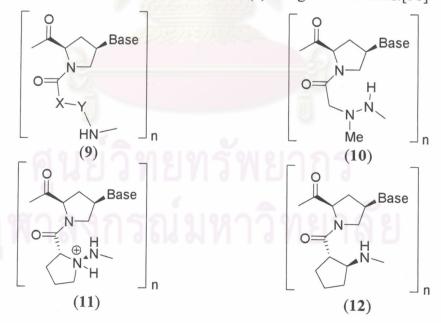


Figure 1.12 Structure of modified PNA bearing various β-amino acid spacer: (9) β-amino acid, (10) N-amino-N-methylglycine (11) D-aminopyrrolidine-2-carboxylic acid (D-Apc), (12) trans-(1S,2S)-2-aminopentanecarboxylic acid (ssACPC)

Since 2001 Vilaivan [57,59,60] has reported novel modified PNA various acyclic and cyclic β-amino acids. It was observed that a favorable geometry for hybridization to the complementary oligonucleotide is possible, at least with some specific configurations. Only the (R,R) isomer carring a D-aminopyrrolidine-2carboxylic acid spacer (D-Apc)) [(11) in Figure 1.12] could form a very stable hybrid with its complementary oligodeoxynucleotide. The stoichiochemistry studied by UV and CD spectroscopy revealed that only 1:1 hybrid was formed. Furthermore, D-Apc PNA system appeared to form more stable hybrid to DNA over RNA. This property had never been reported in others PNA. Indeed, most other PNA, including aegPNA, bind stronger with complementary RNA. Later Vilaivan and co-workers [58] has proposed a new the conformationally rigid pyrrolidine PNA based on prolyl-2aminocyclopentane carboxylic acid (ACPC) backbones [(12) in Figure 1.12]. The PNA containing (R,R) proline and (1S,2S) ACPC can form a very stable 1:1 complex with complementary oligonucleotide similar to D-Apc PNA. The T_m value was estimated to be > 85 °C indicating that the PNA can form a much more stable hybrid with DNA than DNA and most other PNAs including aegPNA and D-Apc PNA with the same sequence. It was proposed that the absence of positive contribution from electrostatic interaction in D-Apc PNA is compensated by the more structurally rigid ssACPC backbone. According to these good initial results, it was aimed to continue investigation of ssACPC PNA with different length and base sequence to evaluate its scope and limitation. The synthetic methodology of ssACPC PNA containing all four DNA bases (A, T, G and C) should also be developed further with the aim to obtain a highly pure PNA in large quantities for the studies and for future application.

1.6 Objective of this research

An objective of this research is to optimize the synthesis protocol for ssACPC PNA (12) bearing all four nucleobases (A, T, C and G) to improve the yield, decrease reaction time and facilitate purification. Another objective is study the binding properties of these mixed-base PNA with DNA.