CHAPTER I

Introduction

1.1. Noncovalent Interaction

Many basic processes of life depends on the noncovalent association and dissociation of biological molecules. The specificity of noncovalent interaction in biological molecules is the basic principle for molecular recognition in nature. The noncovalent interaction is one of the most fundamental regulatory mechanisms in the cell [1-3]. Biological signals are produced when proteins interact with other proteins, peptides, oligonucleotides, nucleic acids, lipids, metal ions, polysaccharides or small organic molecules. An example of protein interactions is protein oligomerization, which is believed to cause an improved stability against proteolysis and thermal degradation. Interactions between other biological macromolecules are also very common, such as between DNA-DNA, DNA-RNA, DNA-peptide, DNA-drugs, etc. Indeed, the basis of genetic information storage, transfer and expression relies on such interactions [3-7].

1.2. Introduction to Peptide Nucleic Acid (PNA)

Deoxyribonucleic acid (DNA) is a material that stores and transfers genetic information to the next generation of most organisms. DNA structure is composed of repeating units. Each unit incorporates three components, namely: 1) Nitrogenous base (purine or pyrimidine); 2) Phosphate group; 3) Deoxyribose (sugar). They are connected together through phosphodiester linkages. Peptide nucleic acids (PNA) are DNA mimics with a pseudopeptide backbone. Despite the lack of the sugar-phosphate backbone, PNA is an excellent structural mimic of DNA or RNA. PNA can form more stable duplex structures with Watson-Crick base pairs and triplex structures with Hoogsteen base pairs (Figure 1.1) [3, 4, 6, 8-10].

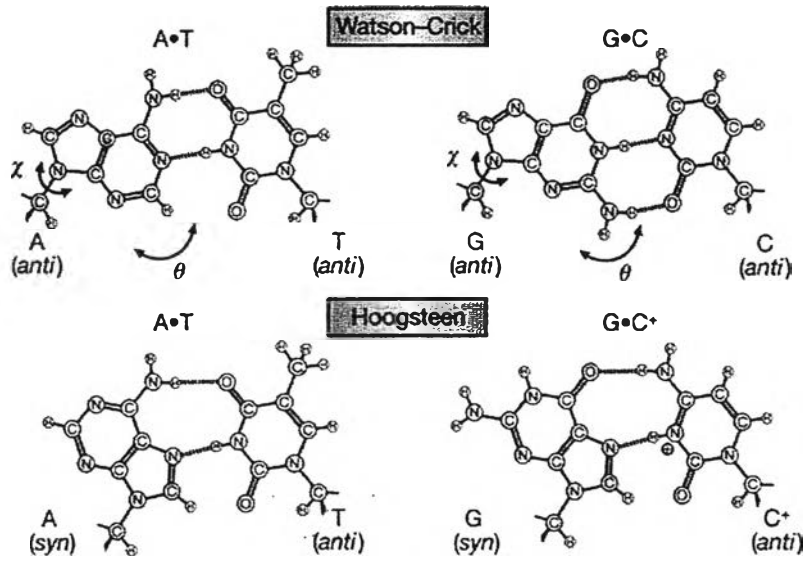

Figure 1.1 Chemical structure of a Watson-Crick and Hoogsteen base pair (http://en.wikipedia.org/wiki/Hoogsteen_base_pair).

1.2.1. PNA-DNA Hybridization

The first PNA was synthesized by Nielsen and co-workers in 1991. The PNA has a 2-aminoethyl-glycine unit in place of the normal phosphodiester backbone of DNA. This results in a chiral and uncharged DNA mimic (Figure 1.2) [3, 11, 12].

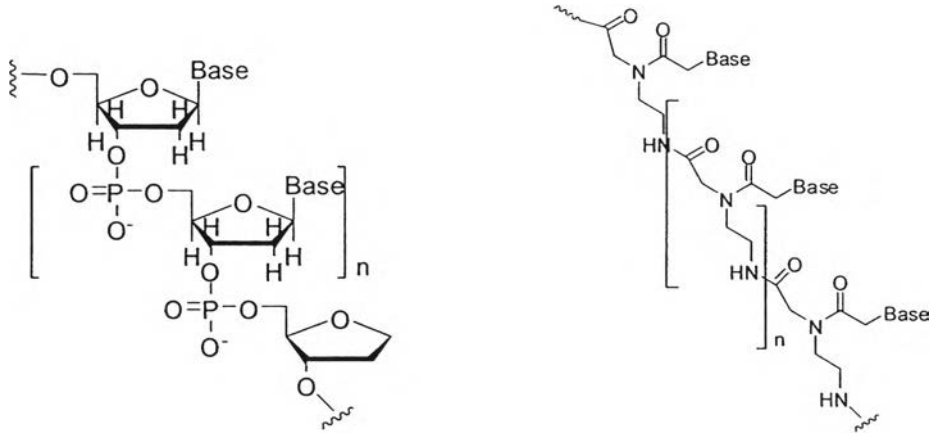

Figure 1.2 Chemical structures of a DNA and an aegPNA molecule.

The PNA and DNA form PNA-DNA hybrids with higher affinity than complementary DNA-DNA, DNA-RNA or RNA-RNA. The thermal stability follows as PNA-PNA > PNA-RNA > PNA-DNA > RNA-DNA > DNA-DNA, respectively [1, 5, 13, 14]. In addition, PNA can also form hybrids with double stranded DNA by a new strand inversion mechanism to form looped structures as shown in Figure 1.3 [11, 15-17].

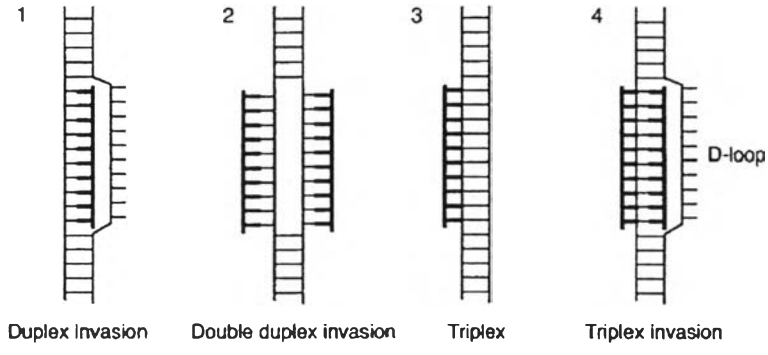

Figure 1.3 PNA binding modes for double stranded DNA [3].

1.2.2. Applications of Peptide Nucleic Acid

PNA can be used in a variety of fields, including biotechnological and medical applications. PNA-DNA hybridization can be used in genetic sequence determination, modulation of PCR analysis, biosensors and antisense/antigene technology. These applications mostly rely on Nielsen's aegPNA system since it is the only one that is presently commercially available. The aegPNA is known to bind to DNA in both parallel and antiparallel directions.

1.2.2.1. Antisense and Antigene Therapy

The base-pairing of two strands of DNA to form a double helical hybrid is among the most important molecular recognition event in nature. In the normal cell, the expression of genetic information consists of two processes, namely transcription and translation. In the transcription process, the DNA strand serves as the template for the synthesis of messenger ribonucleic acid (mRNA). In the translation process, the mRNA sequence determines the order of the amino acid that will be linked together by ribosomes to form a protein. This generates a protein with a specific structure and function. Inhibition of gene expression can be achieved by sequence-

specifically blocking the mRNA that act as the template for the protein production by a short oligonucleotide sequence. This will result in interference of the normal translation process - so called antisense approach. In a related concept, the antigene oligonucleotide can be targeted to the DNA sequence to inhibit the transcription into mRNA, and this its translation to proteins. These concepts are illustrated in Figure 1.4 [5, 9, 18].

Figure 1.4 Antigene and antisense strategy. An antigene oligomer 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 oligomer, and hybridization to a specific mRNA sequence can inhibit the expression of a protein at the level of translation [1].

The highly specific recognition through the pairing of the four nucleobases is very important for the development of antisense and antigene oligonucleotides. Natural oligonucleotides have been demonstrated to exhibit both antisense and antigene properties in vitro. However, they are rapidly degraded by nucleases in vivo. This stimulates the development of several novel oligonucleotide mimics to overcome this obstacle. The modification can improve some properties such as the binding affinity to complementary nucleic acids and membrane permeability. Several oligonucleotide analogues, including PNA [9, 18] have been successfully used as antisense and antigene agents that were superior to natural DNA.

1.2.2.2. PNA Probes in Nucleic Acid Biosensors

The sequence-specific recognition of a DNA target by another DNA strand ("probe") can be used for DNA sequence analysis (Figure 1.5). Recently, PNA has been used as a probe instead of DNA to improve the performance of the analysis. In these DNA biosensors, single-stranded PNA probes are often conjugated with optical, electrochemical or mass-sensitive transducers to allow detection of the

complementary (mismatch) DNA strand. The PNA can be labelled and used as a tag encoding the structure of the attached molecule of a defined base sequence. Several probes can be placed in the same piece of biosensor so that several regions of the DNA target can be simultaneously analyzed [19, 20].

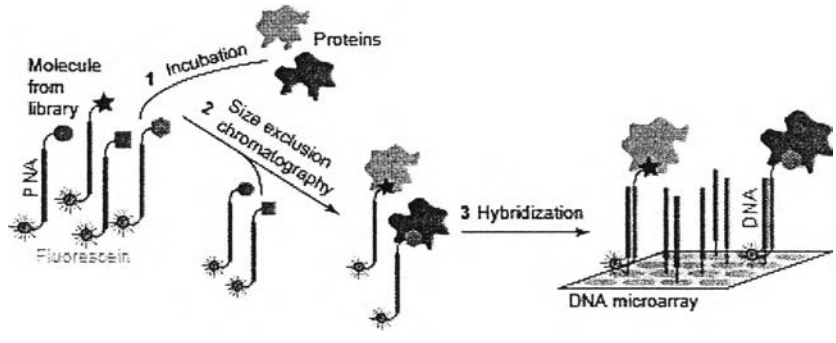

Figure 1.5 Molecule sorting on spatially addressable microarrays.

1.2.2.3. PCR Technique

The polymerase chain reaction (PCR) technique can be adapted to analyze a genetic variation called single nucleotide polymorphism (SNP), which is very useful in the medical and forensic fields. The DNA template carrying a target sequence and two oligonucleotide primers are added together with the deoxyribonucleotide triphosphates and a heat-stable DNA polymerase. The DNA template molecule is denatured by a brief heating to separate the two strands by breaking the hydrogen bonds that hold them together. The mixture is allowed to cool to a lower temperature, allowing the separated DNA template molecule to anneal with the primers that are complementary to either of the template's end. The primers are extended by the action of DNA polymerase that synthesizes a complementary sequence in the 5' to 3' direction starting from each of the primers. Polymerization continues until the end of the template is reached. The mixture is heated again to separate the DNA strands and the cycle is repeated. Each new strand then acts as a template for the next cycle of synthesis. In this way, the number of DNA increases exponentially (Figure 1.6) [21]. The success of the PCR technique was developed in two parts: a technology for detecting PCR products in fluorescence solution and developed a thermocycler.

We can measure PCR products has come from measurements of the fluorescence signal, which is the third method using SYBR Green I Dye, Fluorescence Resonance Energy Transfer (FRET) and Molecular Beacons [48-50].

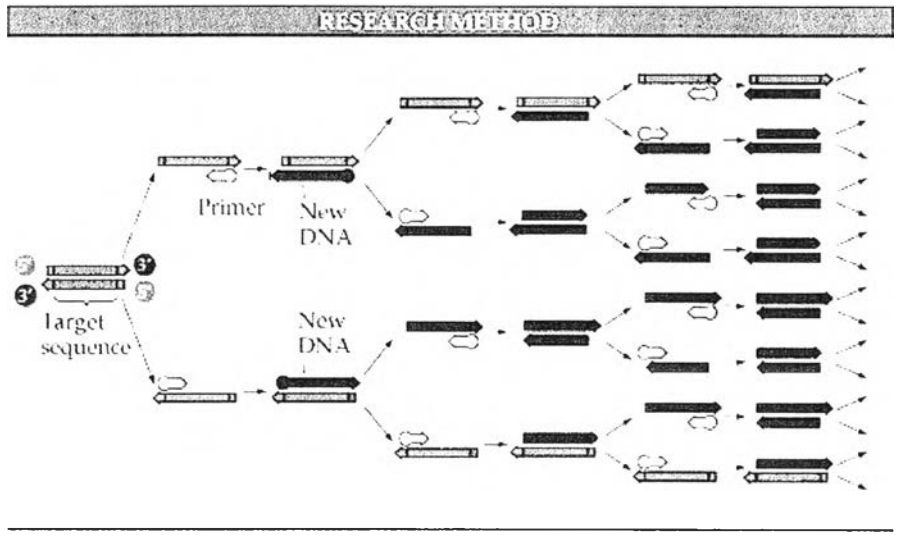

Figure 1.6 An illustration of how PCR works (Purves, Sadava, et al., 2001).

1.2.3. Theory of Mass Spectrometry

1.2.3.1. General Mass Spectroscopy

Mass spectrometry is one of the most powerful analytical techniques for analysis of biological molecules and their interactions. It is a high performance tool that has the ability to identify or detect structural information of unknown samples. Mass spectrometry has been used extensively in biochemistry (analysis of proteins, peptides, oligonucleotides, polysaccharides, lipids and virus particles), pharmaceutical applications (drug discovery, drug metabolism, combinatorial chemistry, pharmacokinetics), clinical applications (neonatal screening, diagnosis of diseases, hemoglobin analysis drug testing), environment (water and air quality, food contamination, heavy metals), geology (isotopic composition, oil composition), forensic science (identification of unknown samples) and industry (monitoring of process streams) [8, 22, 23].

Mass spectrometry can analyze samples both qualitatively and quantitatively. The analysis process starts from the charged sample compounds from which ions are then measured. The process includes ionization, ion separation, recording of the ions according to their mass-to-charge ratio (m/z) and abundance. Thus, a mass spectrometer can be divided into three fundamental parts: the ionization source, the mass analyzer, and the detector. In the ionization source, the solution sample was changed to sample ion form via the ion focusing system. The sample molecule was introduced into the ion source. The sample was changed to ions by protonation and deprotonation. There are many ways to create ions, which depend on the nature of the sample or what kind of mass spectrometry data is required. Samples to be analyzed by mass spectrometry can be solid, liquid and gas. The ions in the gas phase are generated by electrostatic forces. Ions in the gas phase go into a mass analyzer and are separated according to their mass-to-charge ratio and the latter is detected. The results of the data processing of mass spectrometry will reveal the molecular mass or structural information [22, 24-27].

1.2.3.2 The Electrospray Ionization Process

Electrospray ionization (ESI) is a technique of ionization at atmospheric pressure. It is suitable for analyzing samples in the form of preformed ions (the ion in the solvent). The sample in the form of the solution is passed through the small capillary tube at a rate of 10-20 uL/min. This solution is sprayed out of an electric field probe with high voltage (3-6 kV) transmission to the end of the capillary tube. The uniform droplets will have a positive charge or a negative charge depending on the end cap polarity on the counter electrode. When these droplets pass through the ionization process, ion evaporation occurs very rapidly. The ions of the sample are changed to ions in the gas phase. In the next step the ions are induced by the ion optic into the mass analyzer. The sample is analyzed with a chance of a multiply-charged ion in ESI induced to an easier ion optic solvent. The process of the sample preparation in the form of the ion can be achieved by simple reactions. When using acetic acid or formic acid, the solvent is about 1-5% (proton donor). Using an electron acceptor solvent is better for the analysis of the electron donor. In the ion evaporation conditions, the solution must be sprayed in the gas phase into the droplet profile. It's important to choose a solvent with low surface tension. Pure water has a high surface tension. It is not a suitable solvent for ESI, but the addition of an organic solvent with low surface tension makes it even more suitable for the

ESI technique, because solvents with high surface tension make it more difficult to spray very small droplets than those with low surface tension. Therefore, the voltage to the end caps capillary and counter electrode is often adjusted upwards in the case of solvents. The solvent with high surface tension to increase its density of ions in the sample and increase the ion yield in the ESI can be achieved by increasing the concentration of the sample. The case of increasing the flow rate did not increase the ion yield. The ESI technique is suitable for unstable compounds which cannot vaporize easily. Another reason why ESI is widely popular is due to the charge (Z) that has multiple values. It is possible to analyze compounds with high molecular weight by using a mass spectrometer to analyze the m/z which is not very high (500-3000). The ESI technique is a technique of very high accuracy that allows the molecular weight of proteins and nucleic acids to be determined. This is because the software used to manage the data obtained from ESI often uses the deconvolution algorithm to calculate the molecular weight by changing many mass peaks at different protonation states to one peak that corresponds to the molecular weight [8, 15, 16, 24-30].

1.2.3.3 Collision Induced Dissociation

Electrospray ionization mass spectrometry (ESI-MS) is a soft ionization process. The study of noncovalent interaction, including PNA-DNA complexes can be done with mass spectrometry using collision induced dissociation (CID) [13, 18, 26-28, 31-38]. The working principle of the CID can be explained as follows. The energy was changed from the kinetic energy of the collision into the internal energy of the ions. The ions are stimulated to reach a higher vibrational energy level, finally rupturing the bond issue. The processes induced by collisions are called collision activation or collision induced dissociation (CID). Mechanisms that occur during collisions leading to dissociation of the ions vary according to factors such as the size and energy of the ions. In the case of high kinetic energy (keV), the ionic dissociation mechanism may be due to the internal energy of the ions. The energy allows bond or mechanical dissociation resulting from the kinetic energy to be converted into vibrational direct energy. The two mechanisms show direct excitation of the electronic mode and direct excitation of the vibrational mode, respectively. In the case of ions with low kinetic energy (<100eV) the dissociation mechanism is usually caused by collisions. During the collision of complex ions with the collision gas complex, energy at a high levels is converted into vibrational energy by the ion

dissociation. In the case of an induced dissociation of the ion energy, low energy CID fragmentation efficiency is greater when the collision gas is large. But if the ion is not large, the use of helium gas as a collision gas which is too large will cause the breakdown of the mass spectrum into even more complexes. If the ion is large or if more fragmentation is desired then Ar or N_2 should be used [22, 36, 38-42].

1.2.3.4. Research Examples Related to the ESI-MS for the Determination of Noncovalent Interaction

Suparpprom and co-workers [43] proposed a new pyrrolidinyl peptide nucleic acid bearing prolyl-2-aminocyclopentanecarboxylic acid (ACPC) backbone called acpcPNA (Figure 1.7). They studied the hybridization of acpcPNA and DNA using UV-vis, CD and ESI-MS. The high T_m value indicated a strong DNA binding affinity.

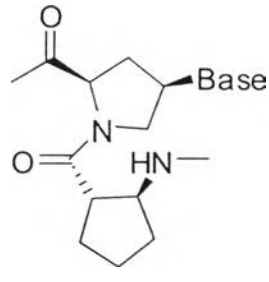

Figure 1.7 Structure of acpcPNA

Arno Wortmann and co-workers [17] determined the binding constant of high-affinity protein-ligand complexes by electrospray ionization mass spectrometry and ligand competition. This group measured the binding constant of PNA-ligand complexes by electrospray ionization mass spectrometry. The method is validated with ligands binding to avidin and applied to ligands binding to p38 mitrogen-activated protein kinase. The experiment demonstrated that ESI-MS can be used to directly study the binding constant of protein-ligand complexes and other noncovalent complexes. In addition, the ESI-MS has the advantage of high resolution, high sensitivity and MS-MS capabilities. The disadvantage of using this method of measurement was the long equilibration time to establish equilibrium in the case of very high affinity ligands.

Thomas J.D. Jørgensen and co-workers [31] studied a collision-induced dissociation of noncovalent complexes between vancomycin and antibiotics (ristocetin or pseudoaglycoristocetin) with peptide ligand stereoisomers to show an evidence for molecular recognition in the gas phase. This group studied the noncovalent interaction of vancomycin antibiotics with peptide ligands (Figure 1.9). In solution, the vancomycin group interacts stereospecifically to peptides with the C-terminal sequence –L-Lys–D-Ala–D-Ala. Substitution by an L-Ala at either of the two C-terminal residues causes a dramatic decrease in the binding affinity to the antibiotics. In negative ion mode of ESI-MS, the CID results show that the complex complexes formed between vancomycin with -L-Ala-L-Ala ligand (1:1) was more stable than -D-Ala-D-Ala ligand in the gas phase. The difference in stability of the complex hybridization was attributed to the effect of the specificity of the noncovalent interaction. The CID technique studies the noncovalent interaction compounds by calculating the energy of interaction with noncovalent interaction of the equation $E_{CM} = E_{LAB}[m_s/(m_s+m_p)]$.

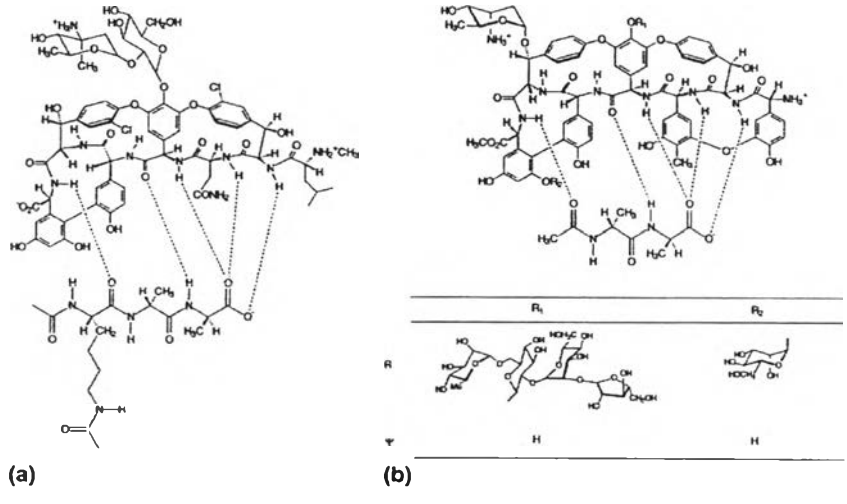

Figure 1.9 Noncovalent complex form (a) between vancomycin with -L-Lys-D-Ala-D-Ala, (b) between antibiotics with -L-Lys-D-Ala-D-Ala.

Jürg M. Daniel and co-workers [25] quantitatively determined noncovalent binding interactions using soft ionization mass spectrometry. In this research, the difference in stability of noncovalent interaction between a solution phase and a gas

phase was studied (Figure 1.10). The solution phase used methods such as melting curves, titration experiments or competition experiments. The gas phase applied methods such as cone voltage-driven dissociation, collision-induced dissociation, blackbody infrared radiative dissociation, or thermal dissociation. In the experiments it was shown that the MS methods can measure the dissociation energies in the gas phase and solution phase. The main reason is that electrostatic and dipolar noncovalent interactions are strengthened in the absence of solvent shielding, while other noncovalent interactions, in particular hydrophobic interactions, become less important in the absence of a solvent.

Type of noncovalent interaction	Formula	Name
Charge–charge	$\frac{Q_1 Q_2}{4\pi\varepsilon\varepsilon_0 r}$	Coulomb energy
Charge–dipole (fixed dipole)	$-\frac{Qu\cos\theta}{4\pi\varepsilon\varepsilon_0 r^2}$	
Charge–dipole (freely rotating dipole)	$-\frac{Q^2u^2}{6(4\pi\varepsilon\varepsilon_0)^2kTr^4}$	
Dipole–dipole (fixed dipole)	$-\frac{u_1u_2}{4\pi\varepsilon\varepsilon_0 r^3}(2\cos\theta_1\cos\theta_2 - \sin\theta_1\cos\phi\sin\theta_2)$	
Dipole–dipole (freely rotating dipole)	$-\frac{u_1^2u_2^2}{3(4\pi\varepsilon\varepsilon_0)^2kTr^6}$	Keesom energy (van der Waals energy $\propto 1/r^6$)
Charge–nonpolar	$-\frac{Q^2\alpha}{2(4\pi\varepsilon\varepsilon_0)^2r^4}$	
Dipole–nonpolar (fixed dipole)	$-\frac{u^2\alpha(1+3\cos^2\theta)}{2(4\pi\varepsilon\varepsilon_0)^2r^6}$	
Dipole–nonpolar (freely rotating dipole)	$-\frac{u^2\alpha}{(4\pi\varepsilon\varepsilon_0)^2r^6}$	Debye energy (van der Waals energy $\propto 1/r^6$)
Nonpolar–nonpolar	$-\frac{3}{2}\frac{\alpha_{01}\alpha_{02}}{(4\pi\varepsilon\varepsilon_0)^2r^6}\frac{I_1I_2}{I_1+I_2}$	London dispersion energy (van der Waals energy $\propto 1/r^6$)
Hydrogen bond	Special, directed interaction	
Hydrophilic interaction	Special interaction	
Hydrophobic interaction	Special interaction	

Adapted from Israelachvili [32]. Q = charge, u = dipole, r = distance, α = polarizability, ε = dielectric constant, I = first ionization potential, θ = angle between dipole and vector connecting the interacting particles, ϕ = polar angle of second dipole.

Figure 1.10 Types of noncovalent interaction [27].

1.3. The Objective in this Research

In this research we study the condition of an electrospray ionization mass spectrometer (ESI-MS) to observe non-covalent PNA-DNA duplexes. Next, we determine the stability of noncovalent interactions from the energy of the hybridization between PNA-DNA complexes using collision induced dissociation (CID) in ESI-MS. The energy data from ESI-MS will be compared with solution phase T_m data obtained from UV-vis spectroscopy.